Title
FERMI-LAT SENSITIVITY AND CONSTRAINTS ON DARK MATTER SIGNAL FROM THE MILKY WAY HOST HALO AND ITS SUBSTRUCTURE

Permalink
https://escholarship.org/uc/item/1m3394tm

Author
Anderson, Brandon

Publication Date
2012

Peer reviewed|Thesis/dissertation
FERMI-LAT SENSITIVITY AND CONSTRAINTS ON DARK MATTER SIGNAL FROM THE MILKY WAY HOST HALO AND ITS SUBSTRUCTURE

A dissertation submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

PHYSICS

by

Brandon Anderson

June 2012

The Dissertation of Brandon Anderson
is approved:

__________________________
Professor Robert Johnson, Chair

__________________________
Professor Stefano Profumo

__________________________
Professor Steve Ritz

Dean Tyrus Miller
Vice Provost and Dean of Graduate Studies
# Table of Contents

List of Figures vi
List of Tables x
Abstract xi
Dedication xiii
Acknowledgments xiv
1 Introduction 1

I Background

2 Instrumentation 6
2.1 Introduction .............................................. 6
2.2 History of Gamma-Ray Astronomy .............................. 7
  2.2.1 Early Missions .................................... 8
  2.2.2 CGRO: EGRET ................................... 8
  2.2.3 Movement to Solid State ........................... 13
2.3 Design .................................................. 15
  2.3.1 Silicon Detectors .................................. 15
  2.3.2 Converters ........................................ 15
  2.3.3 Calorimeter ....................................... 16
  2.3.4 Segmented ACD ................................... 17
  2.3.5 Modular ........................................... 17
2.4 Event Reconstruction .................................. 17
  2.4.1 Readout ............................................ 17
  2.4.2 Modeling ......................................... 18
  2.4.3 Reconstruction .................................. 19
  2.4.4 Calibration ....................................... 20
2.4.5 Background Rejection ........................................... 20
2.4.6 Pass History and Future ....................................... 21
2.4.7 Performance and Systematics ................................. 25

3 DM Particle 27
  3.1 Introduction ..................................................... 27
  3.2 The Standard Model ........................................... 28
  3.3 The WIMP Miracle .............................................. 29
    3.3.1 Thermal Equilibrium ..................................... 30
    3.3.2 Freeze-Out and Relic Density ............................ 32
  3.4 Signatures ..................................................... 36
    3.4.1 Annihilation ............................................... 37
    3.4.2 Decay ..................................................... 47
  3.5 Current Limits ................................................. 48
    3.5.1 Indirect Detection ........................................ 48
    3.5.2 Direct Detection .......................................... 50
    3.5.3 Collider ................................................... 51

4 Cosmology 57
  4.1 Introduction ................................................... 57
  4.2 Measurements of Dark Matter Distribution ..................... 58
    4.2.1 Direct Gravitational Evidence .......................... 58
    4.2.2 Cosmological Evidence .................................. 59
  4.3 N-Body Simulation: State of the Art .......................... 63
    4.3.1 Methodology .............................................. 65
    4.3.2 Results ................................................... 68
  4.4 N-Body Simulation: Data Confrontation ........................ 79
    4.4.1 Galaxy Clusters .......................................... 79
    4.4.2 Core Versus Cusp ........................................ 79
    4.4.3 Subhalo Abundance ....................................... 84

II Research 87

5 Substructure in Via Lactea II 88

6 Analysis of the Galactic Halo 115
  6.1 Introduction ................................................... 115
  6.2 Modeling ........................................................ 116
    6.2.1 DM Signal ................................................. 116
    6.2.2 Isotropic Background .................................... 117
    6.2.3 Galactic Background ...................................... 117
  6.3 Data Confrontation ............................................ 126
    6.3.1 Cosmic Rays .............................................. 126
6.3.2 Gamma Rays .................................................. 130
6.3.3 Pseudo-Experiments ....................................... 149
6.3.4 Model Population ........................................... 155
6.4 Results .......................................................... 162
6.5 Discussion ..................................................... 171

7 Conclusion ......................................................... 173

A Classification Trees ............................................ 188

B SuperSymmetry .................................................. 192
   B.1 Supersymmetry ............................................. 192
   B.2 Neutralino .................................................. 193
## List of Figures

1.1 Weak interaction processes and methods of WIMP detection. Image courtesy of The Max Planck Institute for Particle and Astroparticle Physics. 3

2.1 EGRET, Image courtesy of NASA. 9

2.2 COS-B and EGRET, All-Sky Surveys (Pinkau 2009). 11

2.3 EGRET “excess,” shown as colored bars. Top plot shows default GALPROP model in blue, with inverse Compton (IC), bremsstrahlung, $\pi^0$-decay, and extragalactic (EB) components. The second plot shows a GALPROP model adjusted to match the EGRET data (Strong et al. 2004). 12

2.4 Fermi Large Area Telescope. Image from www-glast.stanford.edu. 14

2.5 Fermi-LAT P7v6 instrument response functions: effective area and PSF as a function of energy (The Fermi-LAT Collaboration 2009). 22

3.1 Particles of the Standard Model, Image Credit: AAAS. 28

3.2 S-channel WIMP Annihilation. 35

3.3 Number density and relic density vs. time and temperature. (Feng 2010) 36

3.4 Natural values for thermal relic WIMP mass. 37

3.5 100 GeV WIMP annihilation spectra for pure channels (Dario Serpico and Hooper 2009). 41

3.6 Virtual Internal (t-channel) Bremsstrahlung. 44

3.7 VIB-dominant Regions of MSSM Space, (Bringmann et al. 2008). 45

3.8 VIB Spectra, (Bringmann et al. 2012). 46

3.9 Sommerfeld Enhancement Factor vs. $\epsilon \equiv m_\phi/(\alpha_\chi m_\chi)$ (Feng et al. 2010). 47

3.10 Upper limits on $\langle \sigma v \rangle$ from combined Fermi dwarf analysis (Abdo et al. 2010c). 50

3.11 Upper limits on $\langle \sigma v \rangle_{\chi \chi \rightarrow \gamma \gamma}$ from EGRET and the Fermi-LAT (Vertongen and Weniger 2011). 51
3.12 Upper limits on $\langle \sigma v \rangle_{\chi \chi \rightarrow b \gamma}$ using the Fermi telescope. Curves show limits from dwarf (diagonal lines), and all-sky (curved) analyses, compared with thermal production limits for three values of the mass splitting, $\mu$ (Bringmann et al. 2012). .............................................. 52

3.13 Upper limits on DM lifetime from Fermi cluster observations (Dugger et al. 2010). ................................................................. 53

3.14 Upper limits on WIMP nuclear scattering cross section from various experiments (Yao et al. 2006). .................................................. 54

3.15 LHC Constraints on WIMP spin-independent scattering cross section compared with limits from direct detection experiments (Akula et al. 2011). ........................................................................... 55

3.16 Combined limits on CSSM parameter space using direct, indirect, and collider measurements (Profumo 2011). .............................. 56

4.1 Gravitational Lensing. Image credit: NASA, ESA & L. Calada ...... 59

4.2 CMB, as seen by the WMAP (Larson et al. 2011). ......................... 60

4.3 CMB Angular Power Spectrum (Larson et al. 2011). ..................... 62

4.4 $\Lambda$CDM Fit to WMAP Data (Larson et al. 2011). ........................ 62

4.5 Joint WMAP, SDSS, and BBN Constraints (Tegmark et al. 2004). . 64

4.6 N-Body initial conditions for the size of initial density fluctuations, $\sigma_8$, relic matter density, $\Omega_M$, and experimentally favored regions. See Klypin et al. (2011). ................................................................. 69

4.7 Structure in the Millenium Simulation (Springel et al. 2008). .......... 72

4.8 Number of VL2 subhalos above $V_{\text{max}}$ within 400, 100, and 50 kpc (descending curves) of the galactic center. The dotted line shows a power law, $\propto M_{\text{sub}}^{-3}$ (Diemand et al. 2008). ......................... 73

4.9 As Figure 4.8, including the recent Bolshoi data (Klypin et al. 2011). 74

4.10 Aquarius Mass Function (Springel et al. 2008). ........................... 74

4.11 Subhalo distribution in the VL2 and Bolshoi simulations. Solid line and points depict normalized VL2 smooth number densities, respectively. Dot-dashed and dashed represent Bolshoi’s. The faster drop-off of Bolshoi number density is presumably numerical (Klypin et al. 2011). ................................................................. 77

4.12 Simulated concentration vs. mass at z=0. Thick dashed lines enclose the intrinsic scatter (Bullock et al. 2001). ................................. 78

4.13 Autocorrelation function: DM vs Galaxies (Jenkins et al. 1998). ...... 80

4.14 Inner slopes, $\alpha_m$, of low-mass galaxies as a function of the innermost measured point of the rotation curve. Data appears to favor pseudo-isothermal models (red) over NFW profiles (blue) (Spekkens et al. 2005). 82

4.15 Missing satellites in the Milky Way and M31 (Klypin et al. 1999). ... 85

6.1 Fermi-LAT CR background rate for two different instrument response functions (Abdo et al. 2010d). ........................................... 118
6.2 Comparison of an exemplary GALPROP model with local measurements of the Boron to Carbon ratio. The blue regions encompass the uncertainty due to solar modulation. ........................................ 127
6.3 As Figure 6.2 but for the ratio of Beryllium isotopes. .............................. 128
6.4 As Figure 6.2 but for the proton spectrum. ........................................... 129
6.5 Annular binning scheme, including mask of the galactic plane. The bottom plot shows the signal-to-noise ratio vs. annulus for the 100 GeV WIMP and LAT data. Note that the cumulative S/N ratio is roughly constant by annulus 6. ............................................. 131
6.6 Best-fitting uncorrected GALPROP model: spectrally integrated spatial fit. .......................................................... 134
6.7 Best-fitting uncorrected GALPROP model: spatially integrated spectral fit. .................................................................. 135
6.8 Best-fitting uncorrected GALPROP model: spatially integrated spectral residuals. .................................................. 136
6.9 Best-fitting uncorrected GALPROP model: spatial and spectral residuals. 137
6.10 Approximate systematic induced by applying outer galaxy spectral correction to the inner galaxy for a single GALPROP model. Also shown are polynomial fits to the systematic of up to fourth order. ......... 143
6.11 Best-fitting outer galaxy-corrected GALPROP model: spectrally integrated spatial fit. .......................................................... 144
6.12 Best-fitting outer galaxy-corrected GALPROP model: spatially integrated spectral fit. .......................................................... 145
6.13 Best-fitting outer galaxy-corrected GALPROP model: spatially integrated spectral residuals. The blue curves depict the outer galaxy correction and the fitted polynomial which represents the correction uncertainty. ...... 146
6.14 Best-fitting outer galaxy-corrected GALPROP model: spatial and spectral. 147
6.15 Pseudo-experiment results with no DM signal: recovered scale. .......... 151
6.16 Pseudo-experiment results including a 100 GeV WIMP signal: recovered scale. .............................................................. 152
6.17 Pseudo-experiment results including a 100 GeV WIMP signal: detection significance. ...................................................... 153
6.18 Pseudo-experiment 95% upper limits including a 100 GeV $bb$ WIMP DM signal. ................................................................. 154
6.19 Pseudo-experiment 95% upper limits and detection significance with no DM signal. Fits use the generator model for a range of WIMP masses. .......................................................... 155
6.20 The difference between randomly sampled non-linear parameters (blue), and those chosen by a CR-fit-trained classification tree (red). Note that it is not possible to see correlations this way, something that makes the CT invaluable. ...................................................... 157
6.21 Cosmic and gamma-ray fit space. Blue dots represent individual GALPROP models, and contours show equivalent combined fits (TS, see Eqs. 6.6 and 6.8). Models favor DM over the null hypothesis as they move up and to the right, with the best-fitting model defining “TS max.”

6.22 Cosmic and gamma-ray fit space, including a set of 0.5%-σ gaussian variant models (yellow) around the global maximum from the initial sampling.

6.23 Partially marginalized GALPROP models: change in TS vs. preferred DM normalization.

6.24 Spread of TS for models resulting from a 5%-σ variation of GALPROP parameters. Models to the right of the dashed line are improved fits.

6.25 Profile Likelihood for 100 GeV WIMP. Individual green curves represent partially maximized Likelihoods, \( \tilde{L}(\beta, \theta_{DM}) \), and the solid black line shows the Profile Likelihood. Vertical dashed lines mark the absolute minimum and the 95% confidence upper limit.

6.26 Results of 100 GeV DM fit. Top plot shows spatial and spectral residuals, middle the contribution of DM in fraction of total counts, and the bottom two plots show the integrated spatial and spectral component breakdowns.

6.27 As Figure 6.26, but for 200 GeV DM.

6.28 As Figure 6.26, but for 300 GeV DM.

6.29 As Figure 6.26, but for 500 GeV DM.

6.30 Significance and 95% confidence upper limits for \( \langle \sigma v \rangle \) of WIMPs with mass ranging from 100 to 600 GeV. The dashed line denotes the thermal relic cross section.

6.31 As Figure 6.30, but including limits derived from dwarf galaxy observations (Abdo et al. 2010c).

B.1 Standard Model along with Superpartners.
List of Tables

2.1 LAT Pass History ................................................................. 23

6.1 Nuisance parameters of the diffuse background model. Those marked linear are scalable post-propagation; others require a full recalculation to adjust. Ranges marked with an * are penalized in the fit with a gaussian set to 1-σ at the indicated limits. ........................................ 125

6.2 Parameters of best-fitting GALPROP model with no corrections or DM. 138

6.3 Parameters of best-fitting GALPROP model with outer galaxy correction and no DM. ......................................................... 148
Abstract

Fermi-LAT Sensitivity and Constraints on Dark Matter Signal from the Milky Way Host Halo and its Substructure

by

Brandon Anderson

Despite overwhelming evidence of its existence, the exact nature of Dark Matter (DM) remains unknown. Theoretically well-motivated, Weakly Interacting Massive Particles (WIMPs) make strong candidates for fulfilling the DM role; they provide gravitational force without participating in the electromagnetic interactions that would make them conventionally visible. Even so, their weak interactions en masse could allow them annihilate or decay into substantial quantities of visible particles, including gamma rays. Continuing observations from the Fermi Large Area Telescope (LAT) provide an opportunity to detect or rule out this possibility. Powerful modern N-body simulations like Via Lactea II make predictions for the distribution and consequent signal of Milky Way (MW) DM substructure for a given WIMP model. Without the fortuitous placement of nearby DM, however, detecting annihilation from a clump would be difficult. Indeed none have yet been conclusively observed. The much larger Host Halo, in whose gravity well the MW resides, constitutes a much more reliable source of possible DM gamma rays. This signal competes with the complex and uncertain background produced by cosmic-ray interactions in the MW medium. A search for DM Host Halo
photons including these background systematics gives constraints on the nature of the DM WIMP.
To my family,
Acknowledgments

The text of this dissertation includes a reprint of the following previously published material:

“Fermi-LAT Sensitivity to Dark Matter Annihilation in VIA Lactea II Substructure”

Anderson, B., Kuhlen, M., Diemand, J., Johnson, R. P., & Madau, P.


The co-authors listed in this publication have given permission to reproduce this material. This author contributed to the analysis in the following way:

- Calculated DM spectra and used them with projected density maps obtained by the other authors to simulate ten years of Fermi-LAT observation.
- Modeled and simulated backgrounds.
- Wrote and executed codes which searched the resulting simulated data for DM signal and assigned statistical significance and a Test Statistic for angular extent.
- Wrote most of the analysis and results sections.
- Made all plots except Figure 1.

I would like to thank Joel Primack for useful discussion and donation of computer time on the local Pleiades cluster. Jan Conrad, Igor Moskalenko, Troy Porter, and many others of the Fermi Collaboration for advice and discussion. Michael Dormody, Taylor Aune, Daniel Hansen, John Kehayias, and Robin Crough for always hearing my
ideas. Aaron Sander and Jonathan Cornell for coding contributions. Brian Winer for his persistence, support, and analysis contributions. Finally, Robert Johnson for his great example and always sticking up for me.
Chapter 1

Introduction

Eighty years have passed since evidence first began to reveal the subdominant nature of visible matter in the universe. Since then we have come to know with confidence that the distribution of all luminous material depends on the existence of an unidentified type of weakly (at best) interacting “dark matter,” or DM. Not only does DM exist, today it represents approximately four times more total mass than all known types of matter combined. Its discovery is as revolutionary as learning that the Earth revolves around the Sun – suddenly placing everything we know into the perspective of a much wider universe. It is no surprise then, that these eighty years have seen an extraordinary proliferation of science attempting to fill in our newfound lack of knowledge. From novel techniques of measuring or simulating the DM distribution, to the postulation and search for new particles or objects that make up its composition, we have made incredible progress.

Despite such progress, the fundamental nature of DM remains unknown. One
of the most popular and well-motivated classes of theory contends that DM is composed of a weakly-interacting, massive ($\mathcal{O}(\text{GeV})$) particle known commonly as a WIMP. The WIMP’s motivation is two-fold, and so neat it is often referred to as the “WIMP Miracle.” Firstly, if such a particle existed, our understanding of cosmology dictates that it would have been distributed in a way that accounts for DM gravitational evidence. And second, problems with the Standard Model (SM) of particle physics independently hint that new particles might exist at the GeV scale. This coincidence alone makes the WIMP worthy of special attention, but the possibility that it is in fact weakly, and not just gravitationally, interacting gives it another excellent feature – testability.

The weak interactions of the WIMP would come in many forms - see Figure 1.1. Searches pursue all possible avenues, with increasingly sensitive equipment emerging constantly to pick up where the last left off. Representing one of the largest leaps in sensitivity over its predecessors, the recently launched Fermi Large Area Telescope (LAT) provides some of the best knowledge we have on extraterrestrial gamma rays. If the WIMP exists and can annihilate or decay into SM particles, it is possible that areas of high DM concentration could induce radiation detectable by the Fermi-LAT.

Although there are theoretical realizations of the WIMP that result in a unique, or “smoking gun,” spectrum that would stand out against conventional sources, the most generic theories predict a hard, continuous spectrum with a sharp cutoff. Because such a spectrum could be confused with non-DM sources, astronomical searches for continuum DM emission benefit from knowing something about the morphology of the signal. Gravitational evidence such as stellar velocity dispersions can only tell us the
The best knowledge of the DM distribution on local scales is statistical, and comes from N-body simulations.

N-body simulations discretize the better-constrained early DM distribution and allow the constituent “particles” to evolve gravitationally and self-consistently to the present day. Studies of their resultant distribution give us statistical knowledge of the location of DM within our galaxy. Succinctly, they generically predict that galaxies like the Milky Way (MW) sit at the center of massive and enormous (200 kpc radius) Host Halos of DM. These Hosts are in turn populated with a wealth of subhalos: smaller, self-bound DM structures.

Figure 1.1: Weak interaction processes and methods of WIMP detection. Image courtesy of The Max Planck Institute for Particle and Astroparticle Physics.
This dissertation employs the information garnered from such N-body simulations to make predictions and search for the gamma radiation resulting from the galactic annihilation of generic WIMP candidates using the Fermi-LAT. Specifically, Chapter 5 presents a study of the Fermi-LAT detectability of MW DM substructure as predicted by the Via Lactea II simulation (Diemand et al. 2008). The results are somewhat pessimistic, and so Chapter 6 turns to the Host Halo itself, and details an analysis of whether or not any emission it provides stands out against the background of cosmic-ray-induced diffuse gamma radiation. Chapters 2 through 4 provide a review of the necessary subjects, and finally Chapter 7 summarizes the results.
Part I

Background
Chapter 2

Gamma-Ray History and Instrumentation

2.1 Introduction

Thought to have mass on the GeV scale, WIMPs by their nature possess enough energy to create gamma rays. WIMP annihilation, occurring mostly in areas of dense DM concentration could provide a signal that travels cleanly through the galaxy to Earth. To measure it, we look to the relatively new field of gamma-ray astronomy. This chapter outlines the history of the field, including highlights of some previous DM searches and then describes in some detail the workings of the most recent instrumental addition - the Fermi-LAT. This knowledge will be useful in understanding the sensitivities and systematics that play key roles in Chapters 5 and 6.
2.2 History of Gamma-Ray Astronomy

The year 2011 marked the fiftieth anniversary of the first gamma-ray detection by the Explorer XI satellite. Fifty years used to be a very short time in astronomy. It took fifty years to advance an order of magnitude in magnification from Galileo’s first refracting telescope to Newton’s reflector. After Edison’s original proposal to look for radio waves from the Sun, it took 43 years for Jansky to detect extraterrestrial radio photons. Now, fifty years from that first humble data set of twenty-two photons from Explorer XI, the Fermi-LAT telescope continuously accumulates them at a rate of several per second. Moore’s law applies! That is, we achieve a detection rate today equivalent to simply doubling the effective area of the Explorer XI every two years since 1961. Like any investment, it makes you wish we had started sooner. But of course, we had to wait for the space program.

The galaxy is remarkably transparent to gamma radiation, i.e. it rarely loses energy or changes direction. Although this fact allows the radiation to reach Earth relatively unadulterated, our atmosphere (thankfully!) is almost completely opaque. For the highest energy gamma rays, telescopes on the ground can reconstruct the initial photon using Cherenkov radiation from the resultant shower of particles in the atmosphere. Below 50 GeV or so, the shower becomes too faint and the technique is no longer viable without an impractically large telescope. Gamma rays below this threshold must be observed from space.
2.2.1 Early Missions

Having already observed cosmic rays (CR), scientists predicted that the high energy particles would interact with the interstellar medium and produce galactic diffuse gamma radiation. The Explorer XI mission confirmed this in 1962. The instrument functioned by converting incoming gamma rays into $e^+e^-$ pairs (whose charged interactions are much easier to measure) and then reconstructing the photons. Faced with an enormous background of protons, the satellite had to have a plastic scintillator shield that would detect incoming charged particles. This was known as an anti-coincidence detector, or ACD.

Subsequent missions in the same energy range followed the Explorer XI prototype. The next generation of space telescopes, which included COS-B and the ill-fated SAS 2, introduced spark chambers as the pair conversion and amplification mechanism. SAS 2 discovered gamma emission from the Geminga pulsar before its power supply failed half a year into the mission. COS-B, launched in 1975, remained in the sky for seven years and completed a map of the entire Milky Way, including 22 sources.

2.2.2 CGRO: EGRET

2.2.2.1 Design

Mounted on the Compton Gamma-Ray Observatory (CGRO), the Energetic Gamma Ray Experiment Telescope (EGRET, Fig. 2.1) advanced the field yet another order of magnitude in size and sensitivity, operating from 1991-2000. Incident gamma
rays interacted in the upper tantalum foils of the instrument, sending $e^+e^-$ pairs down through high voltage gas-filled spark chambers. Freed electrons in the chambers would then cascade onto anode wires, forming a detectable level of secondary ionization. Directional information came from the upper twenty-eight layers, where the spark chambers were densely packed. Then, after crossing the first of two time-of-flight (TOF) scintillators, the gamma-induced $e^+e^-$ pair passed through a larger set of spark chambers spaced further apart to better separate strongly beamed high-energy events. A second TOF recorder completed the traversal time measurement and high-Z NaI crystals at the bottom of the telescope converted the remaining electromagnetic shower into scintillated light which served to measure the total event energy. This energy information, combined with a multiple scattering analysis of the $e^+e^-$ tracks (accounting for leakage and material loss), determined momenta to reconstruct the original gamma-ray direction.

Because the preceding chain of events could also be precipitated by an incident CR, the whole instrument was shielded by a monolithic ACD, just like Explorer XI.
The plastic shield would scintillate when penetrated by a CR, issuing a veto flag to the system and preventing the trigger and resultant shower. Being monolithic, however perfectly hermetic, the EGRET ACD gave no positional information about the veto-ing CR. This proved to be a detriment, as the occasional stray particle or photon produced by a gamma-ray-induced shower hitting the bottom of the instrument could travel back up through the telescope and trigger the ACD. An estimated 50% of photons at 10 GeV were lost to this effect (Thompson et al. 1993). Despite this, the rejection system performed admirably, reducing the would-be-overwhelming (10^5 times the number of incoming gamma rays) background to a manageable level.

2.2.2.2 Indirect DM Searches

EGRET sported a 0.5 sr. field of view (FOV) and a 1500 cm^2 effective area. It performed an all-sky survey during its first sixteen month phase, refining and expanding the SAS-2/COS-B picture (Fig. 2.2). Measurements of the inner galaxy flux above 1 GeV exceeded the expectations of the most natural CR-driven models by over a factor of two (Hunter et al. 1997). This caused quite a stir and was seen by some as possible evidence for WIMP annihilation. Although a DM signal could account for the missing flux (de Boer et al. 2005), so could fine-tuning the CR model. The latter was largely achieved by adjusting the average electron source flux. Default models had naturally tuned this to reproduce the spectrum measured at Earth. Of course, this is not a bad place to start, but because the low-mass particles lose energy rapidly, such a measurement will always be very sensitive to the proximity of the nearest emitters.
Measurements at Earth may not represent the average, and some range of variation is acceptable. Moskalenko et al. exploited this range (Strong et al. 2004) and explained the gamma-ray excess through much less exotic means (Fig 2.3).

Beyond the diffuse emission, EGRET amassed 271 sources in its third catalogue, an order of magnitude more than found with COS-B (Hartman et al. 1999). With over 120 of these unidentifiable\(^1\), another exiting possibility arose: the detection of DM signal originating in dark (baryon-free) subhalos. Depending on their proximity and boost factors (Ch. 5) such objects could indeed emit flux comparable to that from the host halo. Randomly distributed through the galaxy, some could appear at angles far from the galactic plane, contending with a much lower background than the host

\(^1\)Often due to large error boxes with multiple associations. See Siegal-Gaskins et al. (2009).
Figure 2.3: EGRET “excess,” shown as colored bars. Top plot shows default GALPROP model in blue, with inverse Compton (IC), bremsstrahlung, $\pi^0$-decay, and extragalactic (EB) components. The second plot shows a GALPROP model adjusted to match the EGRET data (Strong et al. 2004).
halo. There were no “smoking gun” signals, but assuming all plausible unidentified sources were in fact DM implied an upper limit on the annihilation rate (Flix et al. 2005). For a 100 GeV WIMP, this bounded the velocity-averaged cross section, $\langle \sigma v \rangle$, to $5 \times 10^{-24} \text{cm}^3\text{s}^{-1}$, a significant push toward the natural freeze-out value of $3 \times 10^{-26}$ (See Ch. 3).

### 2.2.3 Movement to Solid State

Up to this point, the spark chamber was the centerpiece of pair-conversion gamma-ray science. In use since the 1930’s, it was already a very mature technology. In fact, by the time of EGRET’s launch in 1991, high-energy particle physics (BaBar, NA11/32, CDF, DELPHI, etc.) had moved on, using solid-state detectors or drift chambers. Given the need to house and replenish the gas alone, spark chambers make fragile and short-lived space experiments. The advantages of moving to solid state for a next-generation gamma-ray telescope were thus championed and ultimately brought to fruition by a collaboration of physicists from high energy particle (HEP) and astronomy backgrounds\(^2\). The resulting instrument, the Fermi Large Area Telescope (LAT, Fig. 2.4), carried on the pair-conversion tradition with several new key features:

- Modular Silicon Tracker (TKR)
- Segmented ACD
- CsI Calorimeter (CAL)

\(^2\)Fermi Collaboration (FC)
• Full Instrument Modeling

As a proof-of-concept and pointing instrument, the Italian Space Agency launched an instrument similar to a single tower from the LAT design called AGILE in 2007. Close on its heels, the Fermi Gamma-Ray Space Telescope (FGST) launched a year later, with sixteen towers, a calorimeter, and an additional array of NaI and BGO detectors known as the Gamma-ray Burst Monitor (GBM). After an initial period of on-orbit calibration the telescope entered an all-sky survey mode (occasionally interrupted by pointed observations) that it continues today.
2.3 Design

2.3.1 Silicon Detectors

Aside from being suited to a space environment, another great advantage afforded by silicon detectors in a pair conversion experiment is that they are self-triggering. In order to trigger the high voltage needed to amplify the tiny amount of ionization created by an event in EGRET’s gas spark chambers, a separate time-of-flight system (a pair of vertically separated scintillators) had to register particles as moving downward (Thompson et al. 1993). This was vital in reducing the huge Earth albedo background, but confined the experiment geometrically and left EGRET with a greatly reduced FOV. In addition, any inactive material in a pair conversion instrument acts as a source of background; protons interacting in it can produce gamma rays outside the ACD and are thus irreducible.

Silicon strips, on the other hand, provide constant sensitivity. In the case of the LAT, three successive layers (in the x-y plane) reporting voltages over threshold are sufficient to trigger a readout. This meant the satellite could trim the excess triggering material, lowering irreducible background and opening up the acceptance angle so that the final product could boast a FOV covering roughly 20% of the sky.

2.3.2 Converters

Interspersed with the silicon strips, 16 planes of tungsten convert the gamma rays into $e^+e^-$ pairs. Optimizing the thickness of these planes is non-trivial. A thicker
foil increases the chances of early conversion, making sure each event deposits energy in the maximum number of silicon layers. A thinner one lowers the multiple scattering, and thereby enhances the point spread function. The LAT team compromised with a hybrid of the two: 0.36 radiation lengths of thin converters in the top (front) of the instrument and 0.72 lengths of thick ones in the bottom (back). This way, all photons are given the chance of first converting in the thin foil (enhanced PSF) region, but if they pass through, they are caught by the thicker backup foils further down. This is especially important for higher-energy gamma rays, as they have low flux and it is important to convert as many as possible. Since multiple scattering is suppressed (E^{-1}) for these high energy showers, they produce straighter tracks anyway, compensating in the PSF for more material near the vertex.

2.3.3 Calorimeter

After converting and tracking through the LAT, each burgeoning shower encounters the CAL - a block of 1536 cubic rectangle Thallium-doped Cesium Iodide crystals. The crystals are clear, and transmit scintillated light from the quickly growing shower to PIN photodiodes on opposite ends of the long dimension. The asymmetry of the readout implies the location of the shower center in that dimension, which when combined with the physical location of each crystal, gives a reasonable 3-D coordinate to attach each energy deposit to.

Including both the TKR and CAL, the LAT is 10.1 radiation lengths thick, making it often incapable of containing the entire shower. Fortunately, because the
shower profile decays slowly and predictably beyond the maximum, as long as the CAL captures the early profile shape, a reliable fit can be made for the tail.

2.3.4 Segmented ACD

Like all of its predecessors, the LAT contends with a tremendous flux of CR background. To reduce this, the LAT also uses a plastic ACD, with one notable modification: it is segmented rather than monolithic. This keeps backsplash particles (Sec. 2.2.2.1), which do not align with the shower that caused them, from issuing a veto.

2.3.5 Modular

The TKR itself is also segmented into 16 identical towers. Although this adds about 10% of the effective area as dead space, the redundancy and cost reduction benefits are crucial to making the LAT a practical space mission. Furthermore, longer silicon strips would reduce efficiency and increase noise.

2.4 Event Reconstruction

2.4.1 Readout

Limited by downlink rate, the LAT must perform a good deal of onboard filtering before reporting an event. At the most basic level, triggers are based on primitive logic conditions, such as hits in three x-y planes in a row, high/low energy thresholds
in the CAL, or at a constant rate for calibration purposes. Each tower measures these values and buffers the signal in a Tower Electronics Module (TEM), sending a request to a small processor, known as the Global-Trigger Electronics Module (GEM) when an event condition is satisfied. Including data from the ACD, the GEM then makes a decision on whether or not to read out the instrument. If it is satisfied, the GEM sends out a signal and reads all the instrument electronics, sending their data to an Event Builder Module (EBM), which constructs the full event. The EBM forwards the event to one of two 115.5 MHz Event Processor Units, which apply filter algorithms to reduce the background rate from up to 4 kHz to a level compatible with the downlink rate, 400 Hz.

2.4.2 Modeling

Following the practice of thorough instrument simulation in the manner of its HEP heritage, the FC assembled a mass model in order to simulate the full satellite using the Geant4 Monte-Carlo package (Agostinelli et al. 2003). The model is highly detailed, with about 54,000 distinct volume elements, and tuned with beam test calibration runs to accurately reproduce the digital readout ensuing from any given event. With it, the reconstruction and background rejection software could be optimized with the instrument on the shelf, avoiding the cost and limitations of continued beam tests.
2.4.3 Reconstruction

Reconstructing an event is very simple in principle. Simply launch test particles at the virtual instrument from various angles and energies until the readout matches that from the real satellite. Whichever combination of particle, energy, and angle reproduces the data should correspond to the actual event. Of course, the entire phase space is too enormous for modern computers, and therefore we need clever algorithms to generate only the likely test particles given some subset of the full instrument readout.

The most common of the algorithms used for the LAT basically draws a line between the centroid of the CAL readout and the first hits. If other deposits were made on that line, a Kalman fit (Frühwirth et al. 2000) is performed, tracing the ray from top to bottom, searching for hits in each successive layer within projected covariant error ellipses. This method is less accurate when simultaneous events or backsplash draw the CAL centroid away from the true track center. A future solution is discussed in Section 2.4.6.3. Even when the CAL records no energy, a combinatoric pattern recognition algorithm tries replacing the CAL centroid with each event in the next lowest level and does a decent job of salvaging the reconstruction.

The LAT determines energy by fitting a profile to the EM shower development in the CAL. The already-computed event trajectory is used to parametrically estimate the energy loss from leakage and deposition in the inactive portions of the instrument.
2.4.4 Calibration

Two spare towers were taken to SLAC, CERN, and GSI to put the LAT simulations to the test. The model required very little adjustment, with the furthest exception being the CAL calibration, found to be roughly 7% low (The Fermi-LAT Collaboration 2009). The LAT as a whole was calibrated using sea-level muons and again for the first 60 days of orbit, generally by pointed observations of the Earth’s limb emission. See Appendix A and Chapter 6 for additional calibration discussion.

2.4.5 Background Rejection

After a set of intuitive analysis cuts, a statistical tool known as a classification tree (CT) aids in distinguishing signal from background events and assigning quality to the remaining tracks. See Appendix A for a more complete description of their workings. Trained with a Monte-Carlo model of the on-orbit all-sky flux that includes the expected CR background, the CTs refine the 1:300 downlinked signal-to-noise ratio to an impressive 1:10^6. This is all while maintaining 75% of the gamma-ray events.

The rejection level is somewhat arbitrary and can be tuned, depending on the scientific S/N requirements. To accommodate this without requiring full user reconstructions of the data, the FC releases several ready-to-go sets of events with differing rejection levels. These are known as Event Classes. Depending on the data release, there are several options, but the most common/general are “transient”, “source”, and “diffuse.” These three correspond to progressively (and hierarchically) tighter rejection schemes named in a way suggestive of the science where they are most natural. Diffuse
(most restrictive), for example, suits an all-sky study of DM coming from the Galactic Halo, as in Chapter 6, where the integrated angle provides so many photons that some extra rejection loss is acceptable.

To perform any kind of scientific analysis, raw photon data from any event class must be transformed into a flux. A conversion requires at least three elements: instrument pointing history and estimates of the time and orientation-dependent instrument acceptance and effective area. The FC derives energy dependent averages of the acceptance and effective area from the Monte-Carlo simulations and releases them as official instrument response functions (IRFs) corresponding to each Event Class. Two of these important functions can be seen in Figure 2.5.

2.4.6 Pass History and Future

Background rejection cuts are complex and based on measurement and calibration of the instrument as well as the formal details of the reconstruction. Replacing or altering variables used to train the CTs changes the reconstruction. Scaling the CAL energy reconstruction to match beam testing requires a reworking. Blasting into space changes alignments and connections enough to warrant a new calibration and subsequent reconstruction version. These examples, among many others, mark the “Pass” history of the instrument: a trail of continuously improving versions of the reconstruction software. The analyses detailed in this document make use of Passes 6 and 7 (P6 and P7) only. The rest of the section outlines how each Pass came about, its shortcomings and its improvements over predecessors.
Figure 2.5: Fermi-LAT P7v6 instrument response functions: effective area and PSF as a function of energy (The Fermi-LAT Collaboration 2009).
2.4.6.1 Pass 6

The label P6 refers to the state of reconstruction at the time of launch (P6v1) and after an initial sixty-day period of validations, the primary scheme (P6v2 through v11) for the next three years. Version three was the first to incorporate flight measurements on top of beam and muon tests conducted on the ground. See Table 2.1 for more detail on each revision.

Table 2.1: LAT Pass History

<table>
<thead>
<tr>
<th>Version</th>
<th>Major Updates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pass 6</strong></td>
<td></td>
</tr>
<tr>
<td>v1</td>
<td>State of the Art</td>
</tr>
<tr>
<td>v3</td>
<td>Overlay solution to ghost events</td>
</tr>
<tr>
<td>v7*</td>
<td>Data-derived PSF</td>
</tr>
<tr>
<td>v9*</td>
<td>CTBCORE issue resolved</td>
</tr>
<tr>
<td>v11</td>
<td>Phi and livetime-dependent effective area</td>
</tr>
<tr>
<td><strong>Pass 7</strong></td>
<td></td>
</tr>
<tr>
<td>v4*</td>
<td>Double King function MC PSF</td>
</tr>
<tr>
<td>v7</td>
<td>Ghost tracks recovered</td>
</tr>
</tbody>
</table>

*unofficial

The appearance of *ghost tracks*, or residual energy overlap between events, stands out as the most significant revision prompted by the initial data. Ghost events occur because it takes a finite amount of time for energy to be collected from the LAT. Therefore, each background event leaves a small window of time where residual energy will contaminate any signal readout. This generally confuses the tracking algorithms, which were based on simulations with no overlap, and frequently results in the rejection of a perfectly good gamma ray. Actual recovery of these events had to wait until Pass 7,
since it required CT retraining and an eventual re-write of the reconstruction software (P8).

With ghost tracks eliminating some appreciable fraction of the signal, the P6v1 IRFs incorrectly overestimated the effective area. The FC introduced a quick fix by adding random background (periodic trigger) data to the already simulated signal which would interfere with reconstruction of good events and therefore provide an estimate of the lost photons. This new reconstruction version was known as P6v3, and the accounted-for losses reduced the estimated effective area by roughly ten percent (at 1 GeV). Further corrections to P6 included more accurate functional forms for the point spread function (PSF) and a scale factor correction to the effective area to account for an important cut variable (CTBCORE) that the MC reproduced poorly.

2.4.6.2 Pass 7

Pass 7 fixed the ghost problem at a deeper level, retraining the CTs and re-calculating the MC efficiencies with background overlays. This recovered much of the lost signal, significantly increasing the effective area at energies below 200 MeV. P7 also introduced an improved MC PSF function.

2.4.6.3 Pass 8

Still in progress, Pass 8 represents a complete reworking of the reconstruction. The update promises to flag and remove ghost hits during the track-finding, which will recover yet more events and improve the PSF. The mass model in the G4 simulation
will be overhauled. New methods for track finding will further rein in the tails of the PSF. The calorimeter localization will be tighter, allowing for multiple CAL clusters rather than the overall centroid.

These and other improvements will significantly alter the character of the LAT data. The expected reduction of the PSF tails alone will precipitate effects in nearly every science analysis. The DM projects described in this thesis are no exception. A sharper image will help to decide whether unidentified potential DM substructures are in fact associated with multi-wavelength sources, and if they are spatially extended. The improved localization will make for a better source catalogue, enabling better masking and modeling when investigating isotropic backgrounds, lowering systematic error in the Halo analysis. And that is just from the PSF improvement! Improvements on the energy systematics (Section 2.4.7) could be of even greater importance.

2.4.7 Performance and Systematics

2.4.7.1 Effective Area

The LAT effective area can be thought of as the product of the physical (geometrical) area and the efficiency of the pattern recognition and background rejection cuts. The mass model in the G4 simulation, known from construction and calibrated with beam tests, gives a good understanding of the physical area. Validating the efficiency of the cuts is more difficult. An ideal testing ground requires not only a strong, well-known calibration signal, but the full and complex on-orbit CR background. Bright pulsars fulfill these requirements, with the added bonus that their periodic signal allows
for off-pulse background separation.

To do this, the FC studied the Vela pulsar, a well-characterized and very bright gamma-ray source. As discussed in Appendix A, CTs determine the important variables and where to place event cuts. Validating the efficiency means simulating the pulsar and checking each CT cut individually to make sure it produces the same effect on the real signal that it does on the MC. The study assigned a systematic uncertainty to the effective area, of “10% below 100 MeV, decreasing to 5% at 560 MeV, and increasing to 20% at 10 GeV and above.” (Parent 2010)

2.4.7.2 Residual Background

A pair conversion telescope with inactive material will always suffer from irreducible CR background. The telescope cannot distinguish, for example, between a gamma ray generated in a distant AGN and one that came from a proton interaction in the outer instrument. Fortunately for galactic studies like the ones presented here, this background should average out to be completely isotropic. This makes eliminating it from most DM searches feasible, as the DM usually has a distinct morphology. See Chapter 6 for more discussion.
Chapter 3

The Dark Matter Particle

3.1 Introduction

The best theoretical description of the fundamental subatomic nature of our universe is known as the Standard Model (SM). Despite its success in accommodating all experimental observations, it has some shortcomings. Reformulating our understanding has the potential to go beyond simply remedying these. New theories offer solutions to outstanding questions in fields once thought to be unrelated. This chapter describes the problems with the SM and how extended theories relate to Dark Matter (DM) (Sec. 3.2), reviews the motivation for Weakly Interacting Massive Particles (WIMPs) (Sec. 3.3), and finally examines possible observable signatures (Sec. 3.4).
3.2 The Standard Model

The SM developed in the 1960’s from an effort to unify a plethora of newly discovered particles into an irreducible set of fundamental constituents. There are at least 61 such constituents, including all flavor varieties, cleanly categorized into fermionic “matter” particles, force-carriers, and the Higgs (Figure 3.1). Aside from the Higgs, each one has been confirmed experimentally.

While greatly successful, the SM is not perfect. For one, it does not include gravity. It also exhibits many seemingly arbitrary parameters (masses, mixing angles, etc.), and some that require “unnatural” fine-tuning to reproduce observations. Particularly, strong limits on the neutron electric dipole moment and the low physical mass of the Higgs boson necessitate a delicate balancing of the theory (Feng 2010).
These unsettling problems cast enough doubt on their own to motivate an expansion of our fundamental views. But perhaps even more urgent, our present understanding of cosmology and the role of DM cannot function without a particle whose properties match nothing conceivable within the SM (See Ch. 4). This fact drastically changed the character of the problem. Rather than setting straight some somewhat minor model deficiencies, beyond-the-SM (BSM) theories were catapulted into the spotlight, potentially rescuing the field of cosmology from a would-be dead end.

There are now many theories that fix some of the problems of the SM while simultaneously offering a candidate DM particle. Their mechanisms range from adding extra fields/particles to entire new dimensions. Most make predictions, but ones that are just on or beyond the threshold of testability. With so many options, it is appealing to focus on the most generic properties of DM candidates and use them to motivate discovery of the new particle(s). From there, more specific tests can use the particle to refine the theory space.

BSM theories generically predict new physics at the weak scale. A new particle with mass of that order works well with cosmology. And so this prediction, the WIMP, marks the starting point for most (including this work) efforts to discover BSM physics.

3.3 The WIMP Miracle

The current $\Lambda$CDM\(^1\) cosmological view requires that the total energy budget of the universe contains roughly 23% cold (non-relativistic) DM ($\Omega_\chi = 0.23$, Ch. 4). It

\(^1\)Dark Energy ($\Lambda$) plus Cold Dark Matter (CDM)
is called dark because all astronomical (lack of) evidence argues that it be weakly interacting at best, and certainly incapable of participating in the strong or electromagnetic interactions that would make it visible to us. Aside from possessing these properties and maintaining consistency with other measurements that would reveal effects from its presence (e.g. Big Bang nucleosynthesis, stellar evolution, etc. See Taoso et al. (2008).), there remains considerable freedom in the type of particle(s) that makes up the DM. The WIMP is just one of these, but because it simultaneously helps with SM problems, they receive the most attention.

The WIMP works as DM by not only being hard to detect, but by distributing itself exactly where it should be, and in the proper quantity. That is, if WIMPs were with us from the Big Bang, they would have dropped out of thermal equilibrium with the right abundance and velocity to explain all subsequent DM-driven structure formation. So natural, this solution which solves both physics problems at once was dubbed the “WIMP Miracle.” The remainder of this section is devoted to a quick overview of the cosmological WIMP abundance calculation.

### 3.3.1 Thermal Equilibrium

In the first few nanoseconds of the universe, deep in the radiation era at temperatures $T \gg M_\chi$, WIMPs would have existed in thermal and chemical equilibrium. That is, their interaction rates were greater than the expansion rate of the universe and thus maintained contact with the thermal bath. Pair processes are generally the only ones considered here, for two reasons. First, any process that produces a single DM
particle (e.g. \(ZZ \rightarrow \chi\)) allows decay as the time-reversed equivalent. While this could form an acceptable balance during equilibrium, if it happened with any appreciable cross section, the DM would have long since decayed away after the universe cooled below the production temperature. Second, popular Supersymmetry (SUSY, Appendix B) theories often include a new discrete symmetry known as R-parity which forbids production except in pairs and decay except to lighter DM.

Allowing only pair production and annihilation, we can begin by using an equation of equilibrium to solve for the WIMP abundance,\(^2\( n\chi \), as a function of time and coupled SM pair abundance, \( n_{SM} \).

\[
a^{-3} \frac{d(n\chi a^3)}{dt} = n^{(0)}_{\chi} n^{(0)}_{\chi} \langle \sigma v \rangle \left( \frac{n_{SM} n_{SM}}{n^{(0)}_{SM} n^{(0)}_{SM}} - \frac{n^{(0)}_{\chi} n^{(0)}_{\chi}}{n_{SM} n_{SM}} \right)
\]

The scale factor, \(a\), accounts for the expansion of the universe, \(\langle \sigma v \rangle\) represents the velocity-averaged cross section, and any \((0)\) denotes an equilibrium value. Here the change in co-moving number density per time, \(\frac{d(n\chi a^3)}{dt}\), basically depends on the relative abundance of particles in each side of the annihilation/production reaction, \(SM \leftrightarrow \chi \chi\). An overabundance of SM will produce a positive \(\frac{d(n\chi a^3)}{dt}\) and vice-versa, causing the system to self-adjust – hence equilibrium.

From here on we will assume that the WIMP is Majorana, or its own anti-particle. Although this is not necessarily the case, it is common for SUSY WIMPs (Appendix B), and will significantly reduce our notation. If we further assume that

\(^2\)Number per unit volume. See (Dodelson 2003) for more details.
$M_{SM} \ll M_\chi$\textsuperscript{3} then the SM particles will still be deep in equilibrium at the temperature when DM drops out. So setting $n_{SM} = n_{SM}^{(0)}$ and $\chi = \bar{\chi}$ leaves

$$a^{-3} \frac{d(n_\chi a^3)}{dt} = \langle \sigma v \rangle \left( (n_\chi^{(0)})^2 - n_\chi^2 \right)$$ \hspace{1cm} (3.2)

### 3.3.2 Freeze-Out and Relic Density

To move on from Equation 3.2, we must add temperature dependence to account for the expanding and cooling universe. At equilibrium, where the chemical potential is zero, the Boltzmann number density is defined to be

$$n_\chi^{(0)} = g_\chi \int \frac{d^3p}{(2\pi)^3} e^{-E_\chi/T},$$ \hspace{1cm} (3.3)

where $g_\chi$, $p$, and $E_\chi$ are the species degeneracy, momentum, and energy of the WIMP. In the limiting case where WIMPs are relativistic ($M_\chi \ll T$), approximating the exponential with the first series term reduces the equation to

$$n_\chi^{(0)} = g_\chi \frac{T^3}{\pi^2}.$$ \hspace{1cm} (3.4)

This drops intuitively in relation to the cooling and expanding universe. At the other limit, where the WIMP cools to non-relativistic speeds, we cannot avoid the exponential factor, and falling temperature suppresses the number density much faster than in Eq. 3.4:

\textsuperscript{3}Or at least that the SM particles decay quickly.
\[ n_\chi^{(0)} = g_\chi \left( \frac{m_\chi T}{2\pi} \right)^{3/2} e^{-m_\chi / T} \tag{3.5} \]

Physically, the rapid drop in \( n_\chi^{(0)} \) comes from the fact that the SM particles that collide to produce new DM no longer possess the kinetic energy to overcome the mass difference, except in the far tails of the velocity distribution.

With both temperature regimes covered, let us revisit Eq. 3.2. Any time the temperature drops, so does \( n_\chi^{(0)} \), making the left side of the equation negative. Negative derivative means the actual number distribution, \( n_\chi \), will also drop, attempting to stay in balance. The question is this: given \( T, M_\chi, \langle \sigma v \rangle \), and the expansion rate of the universe, \( H(T) \), can the actual number distribution track the equilibrium value, or will it fall behind?

If WIMPs could stay in equilibrium, with no way of producing more, they would disappear. What makes WIMPs interesting is that a particle with that particular mass and \( \langle \sigma v \rangle \) stays in equilibrium at high temperature, but once non-relativistic, cannot annihilate fast enough to keep up with the equilibrium density. Thus, they freeze out of the mix, locked at some relic density. Further details of the calculation are available in Dodelson (2003) and Kolb and Turner (1990), but the bottom line is that freeze-out occurs at a temperature \( T_f \simeq M_\chi / 20 \), and accounting for the subsequent cosmological evolution leaves us with a current relic density of

\[ \Omega_\chi = \frac{M_\chi T_0^3}{T_f \rho_c M_{pl}} \langle \sigma v \rangle^{-1}, \tag{3.6} \]

\(^4\)Let us assume briefly that \( \langle \sigma v \rangle \) has no temperature dependence.
where $T_0$ and $\rho_c$ are the temperature and critical density\(^5\) today, and $M_{pl}$ is the plank mass. The close relationship between $T_f$ and $M_\chi$ leaves $\Omega_{DM}$ largely independent of the WIMP mass. To reproduce the measured $\Omega_{DM}$ (Ch. 4), $\langle \sigma v \rangle$ must have a value on the order of $3 \times 10^{-26}$ cm\(^3\) s\(^{-1}\). \(^6\)

Although the mass cancels out of the relic density equation, $\langle \sigma v \rangle$ can still carry a mass dependence. The calculation of $\langle \sigma v \rangle$ is formally comprised of an integral over all particle momenta (subject to conservation laws) times the Boltzmann probability of their energy and the amplitude for the process, $|M|^2$. Specifying the amplitude ties us to a particular WIMP theory, and so it is more illuminating and less tedious to proceed on dimensional grounds.

In natural units ($\hbar = c = k_B = 1$), $\langle \sigma v \rangle$ is simply eV\(^{-2}\). If we assume $M_\chi$ is the only dimensionful (eV) quantity determining the cross section (Feng 2010), then it only works if

$$\langle \sigma v \rangle \propto \frac{1}{M_\chi^2} \quad (3.7)$$

Now to fill in the relevant dimensionless dependencies. For velocity, since the WIMPs freeze out at low temperature (with respect to their mass), they are non-relativistic and we can use a partial wave expansion to get

$$\langle \sigma v \rangle \propto \Sigma_p a_p v^p \quad (3.8)$$

\(^5\)For a flat universe. This also ignores a logarithmic correction – see Kolb and Turner (1990).
\(^6\)With exclusively s-wave annihilation.
Estimating the velocity at freeze-out to be $v_f = (3T_f/2M_\chi)^{1/2} \simeq 0.27$ (Bertone 2010), and assuming the average collision velocity to be of the same order, we see that by $p = 4$ the contribution is suppressed on the order of $10^{-3}$. And so the first two ($s$ and p-wave) are generally sufficient, and we can ignore higher powers.⁷

A basic tree level annihilation diagram has two weak vertices (Fig. 3.2). Their corresponding factors are multiplied and then squared bringing the interaction gauge coupling, $g_{weak}$, in at the fourth power. With the last addition of a geometrical factor we have

$$\langle \sigma v \rangle = \frac{g_{weak}^4}{16\pi^2 M_\chi^2} (1 + a_2v^2) \quad (3.9)$$

Adopting this relation gives some constraints on $M_\chi$. If we increase it too much, $\langle \sigma v \rangle$ will drop, and the WIMPs fall quickly out of thermal equilibrium, with a relic density that is too high. The opposite goes for decreasing the mass. See Figure 3.3 for an idea of how this happens as a function of time. Including some uncertainty in $g_{weak}$ and the ratio of $s$ to p-wave annihilation, we are left with a range of $M_\chi$ able to

---
⁷This is not valid near resonances or thresholds. We are also ignoring co-annihilations, a circumstance where scattering can convert the WIMP into a slightly heavier particle with a higher annihilation cross section, effectively keeping the DM in equilibrium longer.
supply all the needed DM: approximately 0.1 - 1 TeV. If WIMPs supply only a fraction of the DM (i.e., there are multiple types) the range expands even further (Fig. 3.4).

3.4 Signatures

If the DM particle really is weakly interacting, the same processes that held it in thermal equilibrium could make it detectable today. Annihilation, which occurs in proportion to the square of the particle density (see Section 3.4.1), could be significant in areas of high DM concentration. Production requires high energy and competes with all other channels, making it incredibly rare naturally, but feasible in a high energy collider. If the WIMP dominant at the time of freeze-out is unstable on a timescale relevant to the age of the universe, decay products are also a possibility. And finally, WIMPs
could scatter directly on normal matter, imparting recoil energy which is the focus of the direct detection field. The work in Chapters 5 and 6 focuses on the signatures of annihilation only.

3.4.1 Annihilation

3.4.1.1 Rate

The annihilation rate, $R$, is a function of the WIMP mass, density, and velocity-averaged cross section ($M_X$, $\rho$, $\langle \sigma v \rangle$). To start, we know that it has linear cross-sectional dependence, which simply scales the target area.

$$R \propto \langle \sigma v \rangle$$  \hspace{1cm} (3.10)
Thinking of the number density, \( n_\chi \), as the probability of finding a particle at any given location means that the likelihood of finding two particles at that location is proportional to \( n_\chi^2 \) (assuming their locations are not correlated). Now,

\[
R \propto \langle \sigma v \rangle n_\chi^2. \tag{3.11}
\]

We cannot forget the combinatorics and so must add a factor of \( 1/2 \) to adjust for double counting. And because we usually infer DM presence by its gravity, or the potential well it creates, it is more useful to refer to the mass density, \( \rho_\chi = n_\chi M_\chi \). With those last changes, the equation is finished.

\[
R = \frac{1}{2} \langle \sigma v \rangle \frac{\rho_\chi^2}{M_\chi^2}. \tag{3.12}
\]

Note that although we considered a mass-dependent cross section in Section 3.3.2, we are now assuming it to be constant. This is common, as any strict mass dependence would be a guess.

### 3.4.1.2 Channels and Hadronization

We can now discuss what happens to a WIMP after it annihilates. Obviously this is a key factor for prospective indirect detection methods. According to perturbation theory, two-body final states will have the highest branching ratios (they are tree level processes). These include \( W^+W^- \), \( ZZ \), and a whole slew of \( jj \): neutrino, quark, and lepton pairs. Gluon and photon pairs are also possible, the latter receiving (despite not being tree-level) much attention (Sec. 3.4.1.4).
Even with little specific knowledge of the WIMP, we can still make some further general statements on which pair channels are favored. Cold DM will be at low velocity and the only viable channels have daughter mass less than $M_\chi$. For the expected $M_\chi$ range, this is only an issue for the heavy gauge bosons and Higgs particles with $M > 80$ GeV. But as long as the WIMP is heavy enough, any bosonic channel could be important.

Fermions require a bit more care. Given that WIMPs today have low relative velocities and that they feel only short range weak interactions, they will annihilate in the lowest possible angular momentum state, s-wave. If we consider spin-1/2 Majorana WIMPs (e.g. neutralinos), then their initial spin state must be antisymmetric, giving a total angular momentum of $J = 0$. Conserving $J$ means the product fermions must also have opposed spins. In other words, the $f\bar{f}$ pair resulting from the annihilation must both have either right or left handed helicity – they cannot be mixed.

Let us consider the case where $f$ and $\bar{f}$ have right-handed helicity ($u_\uparrow$), i.e. their spins are aligned with their momenta. For massive particles, helicity breaks down into right and left-handed chiral components, $u_R$ and $u_L$. The right-handed helicity state expands into

$$u_\uparrow = \frac{1}{2} \left( 1 + \frac{\vec{p}}{E + m} \right) u_R + \frac{1}{2} \left( 1 - \frac{\vec{p}}{E + m} \right) u_L, \quad (3.13)$$

where $\vec{p}$, $E$, and $m$ are the fermion momentum, energy, and mass. Now, the weak force only interacts with the left-handed chiral components of normal particles, and the right-handed of antiparticles. Therefore, the weak $f$ interaction strength depends on
\[ 1 - \frac{|\vec{p}|}{E + m}, \quad (3.14) \]

and the \( \mathcal{F} \) on

\[ 1 + \frac{|\vec{p}|}{E + m}. \quad (3.15) \]

Clearly, the fermion mass plays some role here. The \( f \) component in particular goes to zero when \( |\vec{p}| = E + m \) (\( m \ll E \)). The same thing happens when the products start with left-handed helicity, except it is the \( \mathcal{F} \) component that goes to zero. This all amounts to a suppression of the annihilation process of the order \( (M_f/M_\chi)^2 \), known as “helicity suppression.” Third generation (top and bottom) quarks are therefore favored channels and considered more generic than other species, with the top receiving less attention because it could actually have a mass higher than the WIMP.

Whatever channel it takes, even if the initial annihilation pair is unstable, it will hadronize or otherwise decay through progressively lighter SM particles, and eventually result in long-lived and potentially detectable products. Depending on the environment (local fields, targets, etc.) in which the shower takes place, interactions of these resultant particles can also contribute significantly to the signal.

3.4.1.3 Gamma Radiation

WIMP decay chains can result in gamma radiation, particularly if a \( \pi^0 \) is involved (99% likely to decay into \( 2\gamma \)). From rest, a \( \pi^0 \) will generate two 67 MeV photons. Annihilation product pions have some velocity distribution, however, that smears the
Figure 3.5: 100 GeV WIMP annihilation spectra for pure channels (Dario Serpico and Hooper 2009).

photon energy up through the gamma band all the way up to $M_\chi$. Readily available (Sjöstrand et al. 2008) Monte-Carlo simulation gives us the average total gamma-ray output per annihilation. Figure 3.5 shows the average spectrum for a 100 GeV WIMP annihilating exclusively into $b\bar{b}$, along with a few other exemplary channels (Dario Serpico and Hooper 2009).

There are many stable final state annihilation products besides photons. Leptons can be very plentiful, and with enough energy, contribute significant gamma ra-
diation as they propagate and interact. To get a feel, consider inverse Compton (IC) emission from electrons and muons scattering on the infrared starlight (ISRF) and radio Cosmic Microwave Background (CMB) \([O(10^{12})]\), and \(1.6 \times 10^{11} \text{Hz}\), respectively.

The average frequency, \(\langle \nu \rangle\), of a photon up-scattered by a lepton of mass \(M_\ell\) and energy \(E_\ell\) is given by (Yao et al. 2006)

\[
\langle \nu \rangle = \frac{4\nu_0}{3} \left( \frac{E_\ell}{m_\ell c^2} \right)^2,
\]

where \(\nu_0\) is the photon’s original frequency. Now, an electron or muon from a stationary \(W^-\) decay will have roughly 40 GeV, and therefore scatter a CMB or ISRF photon up to around 1 or 100 MeV, respectively. Being so heavy, the \(W^-\) will rarely have an appreciable velocity (at least for \(\sim 100\) GeV WIMPs), and so this process would seem just shy of producing gamma radiation. Despite this, very heavy WIMPs, especially those constructed to be “leptophillic”\(^8\), are theoretically possible (Fox and Poppitz 2009) and often investigated (e.g. (Dugger et al. 2010). A TeV scale WIMP, annihilating directly to electrons, for example, could easily scatter CMB photons up to a GeV or more.

In addition to IC scattering, bremsstrahlung, spallation, and synchrotron can all add to the gamma output. Calculating such radiation requires a bit of extra work, as the products must be propagated in a realistic environment. Fortunately, code exists for doing just that within the Milky Way (Strong et al. 2004). One needs only to provide a spatial and spectral lepton source map.

\(^8\)Interacts preferentially with leptons, i.e. \(\chi \chi \rightarrow l^+l^-\).
3.4.1.4 Special Cases

The generic picture of WIMP annihilation results in a pretty nondescript ($\pi^0$-dominated) gamma-ray signature. That is, compared with most of the gamma-ray signals in the sky, a gently varying power-law bump hardly stands out. Detecting it requires either very low background, or very good models. Beyond the most basic assumptions, however, some very simple and plausible theory additions can dramatically alter the character or strength of the signal, making it both easier to detect/rule out. While there are many, three of the most popular are direct-to-photon annihilation, internal bremsstrahlung, and Sommerfeld enhancement.

Annihilation directly into photons is a tantalizing prospect, not only because it results in a completely unmistakeable line feature, but the photon energy would very accurately determine $M_\chi$, something a broad signal would leave in ambiguity. The only drawback is that all the $\chi \chi \rightarrow \gamma \gamma$ diagrams involve a loop and are appropriately suppressed by a factor of $10^{-4}$–$10^{-1}$. This makes them secondary channels, leading to very small signals, unless something prevents annihilation by the more preferred channels discussed earlier. There have been many searches, nonetheless, e.g. Abdo et al. (2010b).

Slightly less striking than a line, but for some models the dominant source of gamma rays, are hard photons from virtual internal bremsstrahlung (VIB) (Bringmann et al. 2008; Cannoni et al. 2010). The term refers to a process where WIMPs anni-

---

9 Internal Bremsstrahlung is a related interesting process, particularly for Kaluza Klein DM. (Bergström et al. 2005)
hilate via t-channel (Figure 3.6), exchanging a particle which radiates a photon. Not intuitively a dominant process, it becomes important in models where normal channels are suppressed. For example, consider the helicity suppression of s-wave (2-body) annihilation in Section 3.4.1.2. A three-body decay ($\chi \chi \rightarrow SM\gamma$) breaks the symmetry and avoids suppression, particularly at lower masses.

There are several regions of theory space where VIB photons dominate. See Fig. 3.7 for a view of the $M_\chi$ vs gaugino fraction Minimal Supersymmetric Standard Model (MSSM) parameter space, where the dark areas indicate VIB-dominance (Bringmann et al. 2008). Favoring the same process, the character of the spectrum remains roughly constant across the model space. See Figure 3.8 for some toy-model examples (Bringmann et al. 2012). Note the line-like feature just shy of $M_\chi$. This comes from resonance of the virtual particle mass with $M_\chi$. Searches for such a signal are ongoing, and probably also indirectly subject to constraints set by simple line searches.

Finally, affecting only the rate of annihilation (not the products), we have Sommerfeld enhancement. With the s-wave channel suppressed by helicity constraints
Figure 3.7: VIB-dominant Regions of MSSM Space, (Bringmann et al. 2008).
and the p-wave by low $[\mathcal{O}(10^{-3})]$ velocity and a short range force, it is difficult to garner a strong DM signal without evoking density enhancements, or clumps, in the DM distribution (Ch. 4). Adding an attractive force\(^\text{10}\) between the WIMPs increases their annihilation rate without changing the current density: this is Sommerfeld enhancement.

Specifically, the attractive force commonly considered comes from interaction with a light carrier, forming a Yukawa potential between the WIMPs. This leads to an enhancement of $\sigma v$ which depends on the velocity in the following manner, (Feng 2010):

$$S^0 = \frac{\pi \alpha_x / v}{1 - e^{-\pi \alpha_x / v}}.$$  \hspace{1cm} (3.17)

When velocity is much smaller than the fine structure constant, $\alpha_x$, the relation becomes

\(^{10}\text{Much stronger than gravity.}\)
Figure 3.9: Sommerfeld Enhancement Factor vs. $\epsilon \equiv m_\phi / (\alpha_\chi m_\chi)$ (Feng et al. 2010).

proportional to $v^{-1}$. Using the expected velocity today, the enhancement would be $S_0 \simeq 20$.

Recall from Equation 3.6, that $\Omega_\chi$ depends on $\langle \sigma v \rangle$. Although freeze-out happens at essentially the same time ($S_0$ is small for high velocities), the attraction has strong effects on the annihilation rate prior to kinetic decoupling, the velocity distribution, and therefore $\Omega_\chi$. Adding further complication, there are resonances in the ratio of $M_\chi$ to the force carrier that drastically boost $S_0$ (Fig. 3.9). All of this must be balanced to get back the correct relic density, and is often done maximally for searches trying to explain excesses.

3.4.2 Decay

A DM particle can have a finite lifetime, provided it is long enough to allow for structure formation. Decay could be from an excited state, to a lighter type of DM,
or directly into SM particles. The last would be impossible with an unbroken R-parity (Appendix B), but there are many models (including SUSY ones) that can avoid such a constraint (Chen and Kamionkowski 2004). Whatever the mode, the products and timescale at which they are released must maintain consistency with the entire history of the universe.\footnote{In some cases, the introduction of a late-decaying heavy particle can actually offer an explanation to an otherwise unsolved problem. For example, the under (over)abundance of $^7\text{Li}$ ($^6\text{Li}$) (Fields 2011). Late decay can also heat up DM halos, fixing cusps and over-predictions of satellite galaxies (Bell et al. 2010). Decaying DM with a finely-tuned lifetime can even explain the CR excesses in Sec. 3.4.1.4 (Papucci and Strumia 2010).}

Provided it can maintain consistency, currently decaying DM deserves attention from the same indirect detection methods brought to bear on annihilation. If the decay mode produces gamma signatures roughly equivalent to those in Section 3.4.1.2, we are left with only one major difference in the predicted signals. This is that the rate of decay does not depend on the DM density squared (Eq.3.12), but simply on the density. Decay is therefore much less sensitive to “hot-spots” of DM concentration, and in general provides a more diffuse signal.

### 3.5 Current Limits

#### 3.5.1 Indirect Detection

Despite representing the bulk of the mass in our galaxy, with the exception of a few dense regions (e.g. Sun, subhalos), the DM concentration at Earth’s galactic radius is very low ($\rho_E = 0.386 \text{ GeV/cm}^3$ (Catena and Ullio 2010)). This is approximately $1/100$ the local density of Hydrogen (Puyoo et al. 1997), and plugging that in to Equation 3.12,
along with some natural other values, gives a local annihilation rate of about 1/hour – over the entire volume of the Earth! This pushes our gaze further away, to include areas of higher density, such as the galactic center. Such regions are also natural sources of gamma rays, however, and so any such search must model or otherwise subtract the background. This dictates the m.o. of the indirect detection field: seek regions where the DM somehow stands out. So far, despite the investigation of many such promising regions, we have no convincing evidence for an astronomical indirect detection of WIMP annihilation or decay. What tantalizing evidence there has been (de Boer et al. 2005; Finkbeiner 2004; Meade et al. 2010) can be equally explained with more mundane processes (Kaplinghat et al. 2009; Linden et al. 2010).

Without an obvious signal, we place upper limits on the DM signal flux, $R$ (eq 3.12). By itself, this is not very useful feedback to particle theorists, since it depends not only on the density, but its integral over the line of sight. Removing this dependence requires knowledge of the three-dimensional DM distribution. Fortunately, great strides in the field of cosmology now give us a fairly consistent picture of where and how DM concentrates (Ch. 4). Using this knowledge distills the astronomical information ($R$), allowing us to place limits directly on combinations of $M_\chi$, $\langle \sigma v \rangle$, and channel.

Because of the “WIMP Miracle,” great efforts go into continually improving detectors, models, and techniques push indirect sensitivity to the level of the thermal relic cross section. Only recently have studies begun to reach into this territory, placing upper limits that put pressure on the natural theoretical value (Abdo et al. 2010c; Bringmann et al. 2012; Chen and Kamionkowski 2004; Dugger et al. 2010; Vertongen 49)
3.5.2 Direct Detection

Even though low DM density at Earth precludes a strong local annihilation or decay signal, given a sufficiently dense target with low background, it may be possible to induce detectable levels of direct nuclear scattering. To this end, detectors shield large volumes of high-Z (for cross section resonance) materials and search for energy
Figure 3.11: Upper limits on $\langle \sigma v \rangle_{\chi \chi \rightarrow \gamma \gamma}$ from EGRET and the Fermi-LAT (Vertongen and Weniger 2011).

Like indirect detection, the field has generated some promising hints, but mainly only limits on the scattering cross section (Fig. 3.14 (Yao et al. 2006)). Although indirectly related to the annihilation cross section, constraints on $\langle \sigma_p \rangle$ will soon push into model space, potentially detecting or ruling out values of $M_\chi$.

### 3.5.3 Collider

Starting in March 2010, the Large Hadron Collider (LHC) began operations with record (3.5 TeV) energy proton beams. Built for high-mass particle searches, the same cross section that is the focus of direct detection experiments, enables WIMP
Figure 3.12: Upper limits on $\langle \sigma v \rangle_{\chi \chi \rightarrow b\bar{b}\gamma}$ using the Fermi telescope. Curves show limits from dwarf (diagonal lines), and all-sky (curved) analyses, compared with thermal production limits for three values of the mass splitting, $\mu$ (Bringmann et al. 2012).
Figure 3.13: Upper limits on DM lifetime from Fermi cluster observations (Dugger et al. 2010).

production. A WIMP would generally make itself known by leaving the detector with a large unaccounted-for mass. 12

Because they measure related cross sections, collider and direct detection experiments are very complimentary. Figure 3.15 (Akula et al. 2011) shows their joint constraints on mSUGRA (Appendix B) parameter space. Indirect constraints are so far much less comparable, at least in the same regions of theory space (Profumo 2011). See, for example, Figure 3.16 where, in CMSSM (Appendix B), interesting regions of the universal scalar ($m_0$) and gaugino ($m_{1/2}$) mass space are barely touched by indirect searches. Should there be a detection, however, the overlap of collider and direct

\footnote{Note that this leaves no way of knowing whether the WIMP is stable beyond the brief period it traverses the detector.}
Figure 3.14: Upper limits on WIMP nuclear scattering cross section from various experiments (Yao et al. 2006).
Figure 3.15: LHC Constraints on WIMP spin-independent scattering cross section compared with limits from direct detection experiments (Akula et al. 2011).

searches will break model degeneracies and help to identify the theory behind the new particle.
Figure 3.16: Combined limits on CSSM parameter space using direct, indirect, and collider measurements (Profumo 2011).
Chapter 4

Dark Matter Cosmology

4.1 Introduction

Recent decades have witnessed the maturation of cosmology as a precision field, now able to accurately measure critical features of our universe as a whole. These include the overall distribution of energy among various forms and the evolution of the early universe. The arrival of powerful computers facilitated the growth of one of the newest branches of cosmology - DM N-body simulation. Such simulations utilize the fact that the universe is mostly composed of DM, which can be approximated to only interact gravitationally, to accurately simulate the spatial evolution of matter while ignoring all complex baryonic processes. These simulations are convergent in many ways and provide universal statistical information on the distribution of DM, even on the sub-galactic scale. Such information is key to guiding indirect gamma-ray searches: giving probable distributions used to make quantifiable predictions on annihilation rates. This
4.2 Measurements of Dark Matter Distribution

4.2.1 Direct Gravitational Evidence

Early DM evidence came from measuring the velocity dispersion of groups of astronomical objects (van de Hulst 1994; Zwicky 1937) and finding that they were higher than expected, given a gravitational potential well based on luminous matter. Take a cluster of galaxies for example. Estimating the gas distribution, and the mass of each individual galaxy by its luminosity or rotation curve implies a cluster density profile, or potential well. The virial theorem tells us that the average kinetic energy for each galaxy should equal half the potential its cluster radius. On the contrary, observations consistently show galaxies moving faster than they should, implying some hidden mass.

Such results are typical, for both galaxies within clusters, and stars within galaxies. Using the virial theorem in the other direction, measured velocities can reconstruct the mass of what must be there, rather than just what we see. Doing so gives a rather startling picture of the way the universe is laid out – it is totally dominated (4:1) by DM! Some dwarf galaxies boast DM-to-baryon mass ratios as high as 1000:1.

General relativity tells us that so much mass will affect not only the velocity of
baryonic matter, but the paths of photons as well. And indeed, the field of gravitational lensing not only bears out the DM hypothesis, but provides an effective way of mapping it where luminous tracers are scarce. Gravitational lensing works just like a conventional lens, except that space-time bends the photon paths, rather than media boundaries (Fig. 4.1). Beyond dense, compact, “strong lens” objects, weak lensing\(^1\) enables mass reconstruction without the fortuitous placement of foreground matter. Some famous cases (such as the Bullet cluster) have a reconstructed mass distribution that is clearly spatially distinct from the luminous matter.

### 4.2.2 Cosmological Evidence

Clearly dominant on the galactic scale, DM must have had a huge influence on the evolution of the universe as a whole. Indeed, all current successful cosmological models rely on the contribution of non-interacting mass to match measurements on

\(^1\)Uses statistics from many small lensing effects.
the largest of scales. Temperature fluctuations in the Cosmic Microwave Background (CMB) stand out as the most revolutionary and precision measurements used to constrain cosmological parameters. The CMB is made up of photons from the “surface of last scattering,” a term that refers to the region in time and space where the primordial plasma cooled to form atoms and the universe became effectively transparent to electromagnetic radiation. This happened approximately 379,000 years after the Big Bang ($z \sim 1400$), and is the furthest back we have “seen” (directly observed). Accurately measured by the Wilkinson Microwave Anisotropy Probe (WMAP), the CMB is an incredibly isotropic\(^2\) signal with a near-perfect black body spectrum with $T=2.726$ K. See Figure 4.2 for an all-sky image (Larson et al. 2011).

Although the scale in Fig. 4.2 is tiny [$\mathcal{O}(10^{-5})$ fluctuations], it clearly shows some anisotropy. Expanded into spherical harmonics, it yields the power spectrum shown in Figure 4.3 (Larson et al. 2011). The fact that there is so much power in the lower moments is actually remarkable, when considering how very far apart regions only a few degrees apart on the sky were at that point in time. Assuming that initial

\(^2\)After correcting for motion of the Milky Way (MW) and contamination by the galactic plane.
perturbations were caused by quantum fluctuations alone would have resulted in a CMB with almost zero spatial correlation; each region of space would have developed according to an initial condition that had nothing to do with its neighbors. Current cosmological theory relies on a process known as inflation to rectify this.

Inflation is an empirically motivated process that kicks in around $10^{-36}$ seconds after the big bang. At this point, the regions that show correlation in the CMB are still in contact. Rather than allowing further quantum fluctuations to wash out these correlations as the universe expands, inflation drastically expands the universe (faster than light), pushing the regions so far apart that their correlations become frozen in. They are no longer close enough to anything to form new domains and after inflation ends they evolve in parallel to what we see as the CMB.

With the addition of both inflation and DM, it is possible to fit the CMB using an analytical model with several simple parameters, e.g. the scale of initial fluctuations ($\sigma_8$), redshift of the transition from radiation to matter dominance ($z_{eq}$), and the density of dark (baryonic) matter ($\Omega_{c(b)}$). A maximum Likelihood fit determines most of these parameters to a high degree of accuracy. WMAP constraints alone give us the parameters in Figure 4.4, which lead to the excellent fit in Figure 4.3 (Larson et al. 2011). The fit suggests we live in a universe composed of 22% cold DM, 4.5% baryons, and 73% dark energy.

\[^3\text{Some are degenerate.}\]
Figure 4.3: CMB Angular Power Spectrum (Larson et al. 2011).

![CMB Angular Power Spectrum](image)

**Table 3**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Seven-year Fit</th>
<th>Five-year Fit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^2 \Omega_m h^2$</td>
<td>$2.258^{+0.027}_{-0.036}$</td>
<td>$2.273 \pm 0.062$</td>
</tr>
<tr>
<td>$\Omega_k h^2$</td>
<td>$0.1109 \pm 0.0056$</td>
<td>$0.1099 \pm 0.0052$</td>
</tr>
<tr>
<td>$\Omega_m h^2$</td>
<td>$0.734 \pm 0.029$</td>
<td>$0.742 \pm 0.030$</td>
</tr>
<tr>
<td>$\Delta_L$</td>
<td>$(2.43 \pm 0.11) \times 10^{-9}$</td>
<td>$(2.41 \pm 0.11) \times 10^{-9}$</td>
</tr>
<tr>
<td>$n_s$</td>
<td>$0.965 \pm 0.014$</td>
<td>$0.963^{+0.014}_{-0.014}$</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>$0.088 \pm 0.015$</td>
<td>$0.087 \pm 0.017$</td>
</tr>
</tbody>
</table>

**Derived parameters**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Seven-year Fit</th>
<th>Five-year Fit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Omega_b$</td>
<td>$13.75 \pm 0.13$ Gyr</td>
<td>$13.69 \pm 0.13$ Gyr</td>
</tr>
<tr>
<td>$H_0$</td>
<td>$71.0 \pm 2.5$ km s$^{-1}$ Mpc$^{-1}$</td>
<td>$71.9^{+2.5}_{-2.5}$ km s$^{-1}$ Mpc$^{-1}$</td>
</tr>
<tr>
<td>$\sigma_8$</td>
<td>$0.801 \pm 0.020$</td>
<td>$0.796 \pm 0.026$</td>
</tr>
<tr>
<td>$\Omega_c$</td>
<td>$0.0449 \pm 0.0028$</td>
<td>$0.0441 \pm 0.0030$</td>
</tr>
<tr>
<td>$\Omega_k$</td>
<td>$0.022 \pm 0.026$</td>
<td>$0.024 \pm 0.027$</td>
</tr>
<tr>
<td>$z_{eq}$</td>
<td>$3196^{+125}_{-135}$</td>
<td>$3176^{+129}_{-180}$</td>
</tr>
<tr>
<td>$\theta_{moon}$</td>
<td>$10.5 \pm 1.2$</td>
<td>$11.3 \pm 1.4$</td>
</tr>
</tbody>
</table>

**Note.** *Models fit to WMAP data only. See Komatsu et al. (2011) for additional constraints.*

Figure 4.4: ΛCDM Fit to WMAP Data (Larson et al. 2011).
4.3 N-Body Simulation: State of the Art

Section 4.2 reviewed the gravitational evidence for DM and why we need it to understand the CMB. Of course, the CMB is not the final word on whether a cosmological model is successful. The same theory must be consistent with observations leading up to and including the present time. Predictions here are difficult; even ignoring all forces other than gravity, the density perturbations that give rise to the CMB leave the regime of linear evolution and proceed in an analytically intractable way.

Enter the science of N-Body simulation. Modern computational power enables us to calculate non-linear evolution by approximating the matter distribution as a set of massive\(^4\) discrete particles experiencing only gravitational forces. The benefits of such calculations have been many-fold. Requiring the least resolution, large scale structure simulation came first and provided important confirmation of ΛCDM, even removing some of the parameter degeneracies left by the CMB.\(^5\) As computing power increased, simulations grew more precise, able to resolve DM structure on the sub-galactic scale. We can now predict local DM distributions, at least in a statistical sense, without the aid of luminous matter. This is something lacking, by definition, in current observational techniques.

\(^4\)\(O(10^8 \rightarrow 10^{12})\) M\(_\odot\), depending on the scale of structure and desired resolution.

\(^5\)See Figure 4.5, where Sloan Digital Sky Survey (SDSS, a measurement of the spatial power spectrum of over 200,000 galaxies) results help to constrain dark and baryonic matter densities (Tegmark et al. 2004).
Figure 4.5: Joint WMAP, SDSS, and BBN Constraints (Tegmark et al. 2004).
4.3.1 Methodology

4.3.1.1 Initial Conditions

Section 4.2.2 stated that density perturbations began as quantum fluctuations in an inflationary universe. These perturbations, $\delta$, in the density, $\rho$, compared with the average, $\bar{\rho}$,

$$\delta = \frac{\rho - \bar{\rho}}{\bar{\rho}}$$ (4.1)

are thought to take the form of a Gaussian random field with a scale-invariant spectrum (Dodelson 2003). That is, performing a spatial Fourier transform on the gravitational potential gives a power spectrum of length scale $k$, where $k^3 P(k)$ is constant. This is a technical way of saying that there will be similar fluctuations on all scales. Perturbations grow as the universe evolves, but while they remain small it is possible to (semi) analytically calculate the evolution of the individual modes. Generating initial conditions for an N-Body simulation consists of carrying out this evolution as far as possible$^6$ and then placing particles to approximate the resulting density/velocity field.

Calculating density mode evolution works reliably out to something like $z \sim 100$, using two distinct approximation regimes. The first is the linear approximation, when $\delta \ll 1$. Here, all modes are assumed to be uncorrelated and evolve obeying$^7$

---

$^6$Subject to the constraint that the field reproduce the CMB.

The modes do not evolve identically, as Eq. 4.2 seems to imply. Those with too large of a scale extend beyond the causal horizon, and remain flat and un-evolving. As the universe expands, modes come inside the horizon sequentially, and thus evolution begins at a different time for each one (Dodelson 2003).

Equation 4.2 remains valid while $\delta \ll 1$, until correlations between modes begin to creep in. Here we switch to a second regime, typically known as the Zel’dovich approximation (Zel’Dovich 1970). This scheme treats the universe as a fluid, incorporating velocity and spatial expansion into a matrix of evolution along principle axes. It is most famous for predicting the sequential contraction of these axes, so that the universe forms pancake-like structures first, which then contract into filaments and globules. The approximation works reliably to roughly $z = 100$, where $\delta \sim 1$. From there, the calculation again becomes unreliable and the endpoint serves as initial conditions for the N-Body.

4.3.1.2 Ignoring Baryons

Major N-Body undertakings ignore the effects of baryons. Certainly they are a second order effect from a gravitational standpoint, and so make little difference on large scales. They require far more intense computation, and their inclusion would come at the cost of tremendous reduction in the simulation scope and/or resolution. At small scales, however, baryonic effects could be substantial (see Sec. 4.4.2). As
N-Body simulations converge, the inclusion of baryons is becoming more commonplace (e.g. Weinberg et al. (2008)).

4.3.1.3 Distributing Resolution

Omitting all baryonic processes, DM behaves like a collision-less fluid with a continuous distribution – these are subatomic particles, after all. Approximating this as an N-Body means calculating the gravitational potential by discretizing this fluid into larger bodies. Some flexibility with the resolution keeps the simulation efficient. A smooth, low-gradient region requires less sampling (in both time and space) than a dense, rapidly changing one. And from far away, the coarsely sampled region produces the same potential, maintaining accuracy of the global tidal field. Beyond the computational priorities, resolution must be distributed intelligently, i.e. on the most scientifically interesting structures. So what do we care most about?

From the perspective of indirect DM detection, the ultimate goal of N-Body simulation should be to generate a realistic model of the (sub)structure in the MW halo. The exact distribution will differ from our own, so we need an ensemble of high-resolution halos that match the observable characteristics of the MW. Their convergent properties then make predictions for our own galaxy. From a DM standpoint, the observable properties of the MW are presently quite vague. We know the total mass to be roughly $10^{12} \, M_\odot$ (McMillan 2011), and that we are fairly isolated, with no recent major mergers and the nearest halo located at M31 – 0.79 Mpc away.

Even with such loose constraints, it is impossible to predict beforehand which
simulated halos will arrive at $z=0$ resembling the MW. So how do we apply extra resolution to \textit{just} the MW-like halos? The answer is to run the simulation twice. An initial, low-resolution run establishes which halos will have the right mass, position, and merger history. Then a re-simulation from initial conditions can apply increased resolution to the regions that will form MW-like halos. Coarse sampling elsewhere saves on compute time, but produces an accurate tidal field for the high-resolution regions to evolve in.

4.3.2 Results

4.3.2.1 Via Lactea and Aquarius

On the scale of a galaxy and smaller, the most powerful simulations available resolve halos\(^8\) using hundreds of millions of particles. Those with this resolution so far are known as the Aquarius (Springel \textit{et al.} 2008), Via Lactea (VL1 and VL2, Diemand \textit{et al.} (2008)), and GHALO (Stadel \textit{et al.} 2009). Within these halos, they resolve hundreds of thousands of substructures, with masses as low as $10^4 \, M_\odot$. Results are consistent between simulations after accounting for differences in their initial conditions (Diemand and Moore 2011). See Figure 4.6 for a comparison of these conditions with the most current favored region.

\(^8\)The inner $r < r_{200}$, where $r_{200}$ encloses the part of the halo with a density greater than 200 times the mean background density.
Figure 4.6: N-Body initial conditions for the size of initial density fluctuations, $\sigma_8$, relic matter density, $\Omega_M$, and experimentally favored regions. See Klypin et al. (2011).
4.3.2.2 Interpretation: Defining Halos & Subhalos

Deriving useful results from raw N-Body particle data begins with halo identification. A halo is essentially any DM structure whose particles are gravitationally bound. If bound to a more massive structure, they are referred to as subhalos. Today, there are dozens of halo-finder algorithms utilizing various fairly convergent techniques (Knebe et al. 2011). Once found, we still face some ambiguity in assigning halos a radius and mass. These can be avoided by using $V_{\text{max}}$, the maximum circular velocity of a bound particle, as a proxy for mass.

$$V_{\text{max}} = \sqrt{\frac{GM_{<r_{\text{max}}}}{r_{\text{max}}}} \quad (4.3)$$

Additionally, the concentration, or average density inside $V_{\text{max}}$, succinctly describes whether the mass in Eq. 4.3 is diffuse or all packed in the center. Normalizing by the critical density (for spherical collapse), makes it $z$-independent.

$$c_V = \frac{\rho(r_{\text{max}})}{\rho_{\text{crit}}} = \frac{2}{H_0^2} \left( \frac{V_{\text{max}}}{r_{\text{max}}} \right)^2 \quad (4.4)$$

4.3.2.3 Density Profiles

One of the earliest and yet most ubiquitous results of N-Body simulation is the so-called “universal density profile,” discovered by Navarro et al. (1996). It seems that all halos, big or small, share a self-similar distribution of mass with the form,
\[ \rho(r) = \frac{\rho_0}{\frac{r}{R_S} \left(1 + \frac{r}{R_S}\right)^2}, \]  

(4.5)

where \( R_S \) and \( \rho_0 \) are the scale radius and density. At low radii, the function diverges \( \propto r^{-1} \) (to a point, obviously). Further out, it falls more slowly, like \( r^{-3} \). Similar functional forms introduced since the NFW boast better fits, but all have the same basic properties.

What happens at the innermost halo radii remains a point of debate. Never quite resolved in a discrete simulation, the increasing application of computer power has failed to converge on a slope. Recent studies argue that the profile will continually flatten with decreasing radii (Bertone 2010). See Section 4.4.2 for further discussion, including the possible effects of baryons.

### 4.3.2.4 Mass/Velocity Function

In \( \Lambda \)CDM, the smallest subhalos are the first to form. From there, the agents of gravity, tidal forces and dynamical friction, strip and re-assemble small structures into bigger ones, resulting in huge filaments, dotted with massive halos like our own. See Figure 4.7 for an eyeful. Bound structures can retain a strongly peaked core, highly resistant to disruption. Therefore small structure does not entirely give way to large - they coexist. Because indirect signals depend critically on such “hot spots”, much effort has gone into quantifying the number and character of surviving subhalos. Studies show that in general, even at \( z=0 \), light subhalos greatly outnumber heavy ones. The surviving distribution falls approximately as a power law \( \propto V_{\text{max}}^{-3} \), or \( \propto M^{-2} \) (Diemand
While data from the recent Bolshoi (Klypin et al. 2011) simulation confirms the power law fit to much more massive subhalos than VL2 studied (Fig. 4.9), as of yet there are no studies which comprehensively probe substructures below $10^4 \, M_\odot$. In principle, DM should clump all the way down to the free-streaming scale\(^9\) – around $10^{-6} \, M_\odot$! That leaves about ten orders of magnitude of extrapolation for any power law fit. See Figure 4.10 to get a feel for how far this is, and how the overall halo mass fraction depends on the power law fit.

Fitting the mass function to VL2 data with a cutoff at $10^{-6} \, M_\odot$ puts about 60% of a MW-like host halo’s mass in substructures below the simulation resolution. Substructures themselves must be similarly “boosted” to account for the increased an-

---

\(^9\)Perturbations so small that the constituent mass will “stream” out before it can collapse.
Figure 4.8: Number of VL2 subhalos above $V_{\text{max}}$ within 400, 100, and 50 kpc (descending curves) of the galactic center. The dotted line shows a power law, $\propto M_{\text{sub}}^{-3}$ (Diemand et al. 2008).
Figure 4.9: As Figure 4.8, including the recent Bolshoi data (Klypin et al. 2011).

Figure 4.10: Aquarius Mass Function (Springel et al. 2008).
nihilation rate in the cores of their own sub-subhalo population. As the extrapolation involves some uncertainty, however, most studies set WIMP annihilation cross-section (or decay rate) limits extremely conservatively with no substructure boost at all. While this assuredly makes for a hard upper bound, it is difficult to believe that there exists zero substructure below the resolution of $10^3$ $M_\odot$ particles. Including such a boost would significantly drop the limits shown in (Ch. 3), casting doubt on standard WIMPs with a thermal relic cross section.

4.3.2.5 Subhalo Distribution

Having established how much substructure exists, we must also understand where it resides. Here, intuition serves well. As a halo builds, in-falling subhalos must slow down due to dynamical friction and lose mass to tidal stripping. They should become more concentrated with each pass (only the core remains intact), and leave a stream of unbound DM behind them. Multiple passes bring the subhalo further inward, increasing both effects until it disrupts entirely, joining with the host.

Following that line of reasoning, one would then expect a few concentrated subhalos near the core, and more diffuse ones further away. We cannot intuit their number distribution as a function of radius, however, without knowing something of the accretion history. For example, if the subhalo infall rate was higher long ago, we would expect proportionally more subhalos in the inner halo, and vice versa. So from here we must abandon our thought experiment and trust the simulations.

---

10 Gravitational drag.
After some debate on the subject (Klypin et al. 2011; Kuhlen et al. 2008; Pieri et al. 2008), interpretation of the subhalo distribution appears convergent. Because of tidal stripping, subhalos selected by mass are strongly anti-biased. That is, the massive ones mostly exist at large halo radii, in contrast to the smooth halo mass which is peaked at the center. Quantitatively (Kuhlen et al. 2008), the number of subhalos, \( n_{\text{sub}} \), at a given mass, \( M \), and halo radius, \( r \), obey the relation

\[
\frac{dn_{\text{sub}}}{dM}(r) \propto \rho(r)r.  \tag{4.6}
\]

Or, assuming an NFW profile (with scale density and radius, \( \rho_0 \) and \( R_s \)),

\[
\frac{dn_{\text{sub}}}{dM}(r) = \frac{\rho_0}{\left(1 + \frac{r}{R_s}\right)^2}.  \tag{4.7}
\]

Selecting subhalos by \( V_{\text{max}} \) rather than mass leads to an entirely different result. Since subhalo cores stay mostly intact, they retain enough of a gravity well to drive high-velocity orbits despite losing much of their outer layer. See Figure 4.11 (Klypin et al. 2011), where a re-analysis of the VL2 halo and Bolshoi simulations shows a subhalo population which closely follows the smooth component. Only below 0.3 \( R_{\text{vir}} \) does the population finally dwindle, falling off as subhalos dissociate.

### 4.3.2.6 Uncertainties: Scatter and Convergence

As discussed in Section 4.3.1.3, high-resolution N-Body depends on selecting a MW-like coarse halo for re-simulation. Knowing little about our own galaxy’s DM distribution, we must select from a large set of equally appropriate candidate halos.
Figure 4.11: Subhalo distribution in the VL2 and Bolshoi simulations. Solid line and points depict normalized VL2 smooth number densities, respectively. Dot-dashed and dashed represent Bolshoi’s. The faster drop-off of Bolshoi number density is presumably numerical (Klypin et al. 2011).
Figure 4.12: Simulated concentration vs. mass at z=0. Thick dashed lines enclose the intrinsic scatter (Bullock et al. 2001).

The scatter in their properties corresponds to a systematic error inherent to any study directly using high-resolution results. Bullock et al. (2001) make a thorough study of this, finding substantial scatter, as seen in the mass-concentration plot of Figure 4.12.

Earlier discussion also included the expectation that DM forms (and retains) structure all the way down to masses at the free-streaming limit. While it is unlikely that any single simulation will ever explicitly answer this question, the trend has continually gone in the right direction. As resolution increases, N-body reveals smaller substructure. One solution might come from using the saved data of previous projects as a coarse tidal field where individual subhalos can be re-simulated, much like the halos in VL2.
and Aquarius.

4.4 N-Body Simulation: Data Confrontation

4.4.1 Galaxy Clusters

N-Body simulations match CMB data by design. Galaxy clusters are the next-largest-scale observable structure, and a logical starting point for the comparison of simulation with reality. This works on a purely statistical basis, namely with autocorrelation functions, $\xi(r)$, which describe the probability of finding galaxies (halos) separated by a distance, $r$. A high $\xi$ for a particular distance means that the universe clumps strongly at that scale. Early investigations (Klypin et al. 1996) found strong deviations at the Mpc scale, where DM appeared to clump much more strongly than galaxies. Higher resolution simulations have mitigated the discrepancy since then, but still show unexplained bias structure relative to observations (See Jenkins et al. (1998), Fig. 4.13). Most recently, further simulation and investigation suggest that observational systematics might account for the rest (Angulo et al. 2012).

4.4.2 Core Versus Cusp

Given that explaining galactic rotation curves is a primary DM motivation (Sec 4.2.1), it stands to reason that they should be consistent with the results of N-body simulations. Recall that the simulations predict a “universal density profile” (Sec. 4.3.2.3), usually $\propto r^{-1}$ in the inner regions where baryons reside. This “cuspy”
Figure 4.13: Autocorrelation function: DM vs Galaxies (Jenkins et al. 1998).
profile results in an virialized inner velocity slope $v \propto r^{1/2}$, which is spectroscopically measurable. That is, red/blue shifts of emission lines like Hα (656.281 nm) give doppler velocities of far away galaxies as a function of radius. When corrected for the orientation of the galaxy, we have a velocity profile which in turn implies a density profile. Again, the question is simply: do the measured density profiles match N-body predictions?

Despite many attempts to reconcile them, the answer is still, no. Observations of dwarf and low surface brightness (LSB) galaxies imply inner density profiles more similar to a pseudo-isothermal (PI) sphere (de Blok 2010) or constant density core ($\rho \sim r^0$). See Figure 4.14 for example, where Hα curves of 165 low-mass galaxies (Spekkens et al. 2005) apparently prefer these “cored” over “cusped” profiles. Cusps at the cluster level are also strongly disfavored by gravitational lensing measurements (Sand et al. 2002), provided the potential well is spherically symmetric.

Reluctant to question a theory as successful as ΛCDM, the low-radius convergence of N-body profiles came under scrutiny. Do we actually have the resolution to make claims about the inner, say, 0.1 kpc of a DM-only simulated halo? Studies of the simulations themselves (Navarro et al. 2010) revealed that the innermost slopes were actually non-convergent, dropping below the classic $r^{-1}$ to slopes as shallow as $r^{-0.85}$ before disappearing out of resolution. This certainly mitigates the discrepancy, but does not eliminate it.

Perhaps the most intuitive answer is that N-body does not include baryons. Indeed, the discrepancy only seems to exist in the inner halo, exactly where the baryons are. Of course, simulations omit them because they are incredibly complex; star for-
Figure 4.14: Inner slopes, $\alpha_{m}$, of low-mass galaxies as a function of the innermost measured point of the rotation curve. Data appears to favor pseudo-isothermal models (red) over NFW profiles (blue) (Spekkens et al. 2005).
mation, supernova feedback, magnetic fields, radiation pressure, and many other effects come into play, all consuming resources and introducing their own systematics. This also makes it nearly impossible to guess beforehand what effect the baryons will have - even to whether they will make the core-cusp problem better or worse! Ejecting many baryons at once (high star formation rate) could puff up the central density spike (Gelato and Sommer-Larsen 1999) into a core, as could dynamical friction from gas clouds in the early universe (El-Zant et al. 2001). On the other hand, hydrogen gas can radiatively cool, which lets it condense into the core. If this happens slowly, the cusp will be exaggerated by conserving $rM(r)$ (where $M(r)$ is the total internal mass) and thereby pulling the DM inward in a process known as adiabatic contraction (Blumenthal et al. 1986).

In short, the problem is far from solved. Perhaps future simulations will be powerful enough to include baryonic effects. This will clarify whether the core-cusp problem comes from the simplicity of N-body simulations to date, or if in fact the problem lies with ΛCDM. Either way, the shape of the density profile matters quite a bit for indirect searches, especially where signal depends on $\rho^2$ (CHAPTER REF). In lieu of a solution, it seems prudent to use both profiles when searching for astronomical signals, though no instrument yet has the angular resolution at γ-wavelengths to resolve this inner region.
4.4.3 Subhalo Abundance

Section 4.3.2.4 detailed N-Body findings regarding the mass (or \( V_{\text{max}} \)) function of subhalos at \( z=0 \). Evidently, we ought to expect the MW to house many thousands of subhalos, from very tiny to \( V_{\text{max}} \gtrsim 30 \text{ km s}^{-1} \). All DM signal aside, such structures can contain stars, whose orbital velocities give away their invisible content. Dwarf spheroidal galaxies fit the bill exactly, but their baryonic content is very limited, making them difficult to detect – especially at large halo radii.\(^{11}\) Even so, early comparisons using the subhalo density \( \text{per volume} \) revealed an actual dwarf galaxy abundance at low \( V_{\text{max}} \) ten times lower than that predicted by N-Body (Klypin \textit{et al.} (1999), Fig. 4.15).

Recently the SDSS dramatically expanded the catalogue of known MW satellites (Adelman-McCarthy \textit{et al.} 2007). Correcting for exposure, Koposov \textit{et al.} (2008) predicts that there ought to be \( \sim 85 \) dwarf galaxies with magnitude \( M_V > -2 \). While significant, this cannot fully account for the so-called “missing satellites” predicted by simulation. From there, the explanation can follow one of two paths. In one, we do not observe the N-Body predicted satellites because they do not contain enough baryons. As the subhalos still represent significant gravity wells, something must have prevented their accretion. For example, if subhalos of different mass accreted their baryons at different times, events like re-ionization could have stunted the process for some and not others (Bullock \textit{et al.} 2000). On the other path, we do not observe the dwarf galaxies because the substructure does not exist. Something in ΛCDM must therefore give way,

\(^{11}\)Recall that the baryonic MW is roughly 20 kpc across, while its host halo has a virial radius more than ten times that size.
Figure 4.15: Missing satellites in the Milky Way and M31 (Klypin et al. 1999).
reducing the number of small DM subhalos. Such a scenario goes beyond the scope of this study.

The problem is of course degenerate; it may well be that there are both fewer DM satellites than we currently predict, and their baryon capture is somehow retarded. Upcoming surveys (LSST in particular) should help by further expanding the satellite catalogue.
Part II

Research
Chapter 5

Substructure in Via Lactea II

As discussed in Chapter 4, N-body DM simulations predict a wealth of substructure within larger galaxy-hosting halos. Their distribution closely follows the host mass, making the inner radii where the Earth resides a region of high subhalo abundance.\(^1\) The likelihood that we live in close proximity to an individual subhalo is proportionately high. And so, although any individual subhalo would emit far less signal than its host, it is possible that its nearness would provide us with substantial flux. Also, because they are spherically distributed, they could appear away from the galactic plane, where the background is much lower.

This reasoning makes a blind\(^2\) substructure search an important early study for any new experiment. That is, a little luck could make a DM substructure signal more overt than that from a more involved, background-dependent search like the one detailed in Chapter 6. In the event that no substructures are observed, it will be

\(^1\)Between \(10^3\) and \(10^4\) times the average subhalo number density.
\(^2\)Unassociated with conventional sources.
important to have an accurate prediction of their signal so that limits can be placed. For these reasons, a study of simulated VL2 substructure signal, as detected by the Fermi-LAT, was completed by this author and collaborators in 2010 and is presented here in its entirety.

It is interesting to note that the Fermi Collaboration has since performed and published a search for such “dark” satellites, where it found only two possible candidates in one year of data that were likely associated with conventional sources (Ackermann et al. 2012). Such a result is consistent with the predictions made here, and so does not rule out the possibility that the accumulation of further data will reveal substructure signal.
Fermi-LAT Sensitivity to Dark Matter Annihilation in Via Lactea II

Substructure

Brandon Anderson\textsuperscript{1}, Michael Kuhlen\textsuperscript{2}, Jürg Diemand\textsuperscript{3,4}, Robert Johnson\textsuperscript{1}, & Piero Madau\textsuperscript{1,3}

anderson@physics.ucsc.edu, mqk@astro.berkeley.edu

ABSTRACT

We present a study of the ability of the Fermi Gamma-ray Space Telescope to detect dark-matter annihilation signals from the Galactic subhalos predicted by the Via Lactea II N-body simulation. We implement an improved formalism for estimating the boost factor needed to account for the effect of dark-matter clumping on scales below the resolution of the simulation, and we incorporate a detailed Monte Carlo simulation of the response of the Fermi-LAT telescope, including a simulation of its all-sky observing mode integrated over a ten year mission. We find that for WIMP
masses up to about 150 GeV/$c^2$ in standard supersymmetric models with
\(\langle \sigma v \rangle = 3 \times 10^{-26} \text{cm}^3\text{s}^{-1}\), a few subhalos could be detectable with $> 5$ standard deviations significance and would likely deviate significantly from the appearance of a point source.

Subject headings: dark matter – Galaxy: structure – gamma rays: observations

1. Introduction

Even 75 years after the first observational evidence for a non-luminous form of matter (Zwicky 1933) we know remarkably little about the nature of the hypothesized dark matter particle. A promising way forward is indirect detection, whereby ground and space-based observatories are searching for the products of pair annihilations of dark matter particles, such as neutrinos, relativistic positrons, or gamma-rays.

The gamma-ray signal in particular has received a lot of attention in recent years, owing in part to the launch of the Fermi Gamma-ray Space Telescope. Numerous papers have discussed whether the Large Area Telescope (LAT) aboard Fermi will be sensitive enough to detect this signal and how to differentiate it from conventional astrophysical sources, e.g. Baltz et al. (2007). It appears that a detection is challenging but feasible for a wide range of plausible physics models (Baltz et al. 2008). The Galactic Center (GC) has the greatest and closest concentration of dark matter in the Local Group and as such is a promising target (Berezinsky et al. 1994; Bergström et al. 1998; Cesarini et
al. 2004; Jeltema & Profumo 2008). It is, unfortunately, also a very active region, with a bright diffuse flux of gamma-rays from cosmic-ray interactions as well as being filled with numerous hard X-ray and gamma-ray sources such as supernova remnants, pulsar wind nebulae, X-ray binaries, etc. (Aharonian et al. 2006; Kuulkers et al. 2007), which complicate attempts to search for a DM annihilation signal. Alternatively, the diffuse gamma-ray signal from the Milky Way host halo out beyond several degrees from the GC has been suggested as the most promising (Stoehr et al. 2003; Springel et al. 2008), but it will be difficult to disentangle from the poorly constrained diffuse Galactic gamma-ray background arising from cosmic-ray interactions with interstellar hydrogen and the interstellar radiation field (Strong et al. 2004). Instead, the centers of Galactic subhalos may prove to be the most detectable and least ambiguous sources of gamma-rays from DM annihilations. These could be either subhalos hosting dwarf satellites (Strigari et al. 2008) or one of the many dark subhalos predicted by numerical simulations (e.g. Diemand et al. 2008; Kuhlen 2009).

Subhalos as sources have recently been considered by Kuhlen et al. (2008), who used the Via Lactea II simulation (VL2) in conjunction with a realistic treatment of the expected backgrounds to show that a handful, even up to a few dozen, subhalos should be detectable with the Fermi-LAT at more than 5σ significance for a standard weakly interacting DM particle with a mass between 50 and 500 GeV and a cross section in the range $\langle \sigma v \rangle \sim 10^{-26} - 10^{-25} \text{cm}^3 \text{s}^{-1}$. That analysis assumed homogenous sky-coverage, an angular resolution of 9′, and used a pre-launch estimate of the energy-dependent effective area of the LAT. The present paper updates and improves upon that analysis.
by running the same VL2 all-sky emission maps and the diffuse background predictions through a Monte Carlo simulation of the Fermi-LAT instrument, taking into account the time-dependent sky coverage and an energy-dependent instrument response and angular resolution. Additionally we introduce an improved treatment of the boost factor due to the DM substructure below the resolution limit of the N-body simulation.

2. Methods

The VL2 simulation is one of the highest resolution cosmological simulations of the formation of a Milky-Way-scale dark matter halo to date (Diemand et al. 2008). It employs just over one billion $4,100 \, M_{\odot}$ particles to model the formation of a $M_{200} = 1.93 \times 10^{12} \, M_{\odot}$ Milky-Way size halo and its substructure. It resolves over 50,000 subhalos today within the host’s $r_{200} = 402$ kpc (the radius enclosing an average density 200 times the cosmological mean matter density). As described in more detail in Kuhlen et al. (2008), the simulated dark matter distribution was used to construct all-sky maps of the annihilation flux for a set of observers located 8 kpc from the host halo’s center. These maps consist of $2400 \times 1200$ pixels equally spaced in longitude and the cosine of colatitude, corresponding to a solid angle per pixel of $4.4 \times 10^{-6}$ sr, and form the basis for our studies assessing the capability of the Fermi-LAT to detect signals from individual subhalos. To correct for the artificially low central densities of poorly resolved subhalos, the surface brightness from the central region of each subhalo was increased on a pixel-by-pixel basis to match the expected surface brightness of an NFW (Navarro, Frenk, & White 1996) halo with the subhalo’s measured $V_{\text{max}}$ and $r_{V_{\text{max}}}$. Note that
assuming an NFW profile is somewhat conservative: using the density profiles measured
directly in large N-body simulations (e.g. VL2, Springel et al 2008) or one of the newer
fitting functions instead of NFW leads to an increase of about 30% in the halo luminosity
(Diemand & Moore 2010).

In an attempt to account for the increase in luminosity due to clumping of the
dark matter distribution below the scales resolved by the simulation, we “boost” the
total flux from a subhalo of mass $M$ by a factor $B(M)$, determined from an analytic
model (described in Kuhlen et al. 2008) that depends on the slope $\alpha$ and low mass
cutoff $m_0$ of the subhalo mass function. In Kuhlen et al. (2008) this resulted in a
factor of $\sim 2$ difference in the number of detectable subhalos depending on the values
of these uncertain parameters. The prescription applied there, however, does a poor
job of accounting for the expected radial distribution of this boost. By summing $(1 +
B(M))\rho_i m_i$ (where $m_i$ and $\rho_i$ are the mass and density of the $i^{th}$ simulation particle)
over all the subhalo’s particles within a given pixel, the radial dependence of the boost
effectively followed that of the subhalo’s smooth luminosity profile, whereas it really
should have followed the radial profile of the subhalo population. As such it overly
boosted the bright central region, and the results for the most strongly boosted scenarios
($\alpha = 2.0$) were overly optimistic.

In this work we apply the boost factor in a way that more appropriately accounts
for the radial distribution of subhalos.\footnote{The numerical simulations do not have sufficient resolution to address directly the radial distribution of sub-subhalos within subhalos, so we are guided by the radial distribution of subhalos within the host} The boosted luminosity of a subhalo of mass $M$
is now given by

$$L = \sum m_i (\rho_i + B(M)\langle \rho \rangle) .$$ (1)

The sum is over all of the subhalo’s particles, and \(\langle \rho \rangle = (\sum m_i \rho_i)/M\) is the particle-mass-weighted mean density of the subhalo. Note that just as in the original prescription, the total subhalo luminosity equals \((1 + B(M))\) times the smooth luminosity, but that the radial dependence of the boosted component (i.e. \(m_i B(M)\langle \rho \rangle\)) now follows the mass, not the luminosity. See Figure 1 for a comparison of the different boost intensity profiles. This new prescription may still lead to an overly centrally concentrated boost, since the radial number density profile of subhalos is known to be anti-biased with respect to the host’s mass density profile (Diemand et al. 2007). Note however that the size of this anti-bias depends strongly on the way a subhalo sample is defined. Springel et al. 2008 used samples selected by the present subhalo mass, which show the largest anti-bias, but are irrelevant for the luminosity distribution of substructure. Tidal mass loss affects subhalo masses much more than subhalo luminosities. The mass of subhalos near the halo center is reduced the most, introducing a strong anti-bias into any mass selected sample. But the samples relevant here consist of subhalos that, despite the tidal effects, remain above a certain luminosity, not above a certain mass. Such luminosity selected subhalo samples show practically no anti-bias. Only within about 5% of the virial radius does the subhalo number density profile become shallower than the mass density profile (Figure 6 in Diemand & Moore (2010)). There, \(\rho(r) > 200\bar{\rho} > 20B(M)\bar{\rho}\). In other words, Eqn. (1) follows the real luminosity distribution very well and the correction in...
the very inner parts over-estimates the total luminosity by a few percent at most. We use a boost with \((\alpha, m_0) = (2.0, 10^{-6} M_\odot)\), normalized to have 10% of the mass in clumps containing between \(10^{-5}\) and \(10^{-2}\) times the total mass. This provides an upper bound for the boost factor, and we include the overly pessimistic unboosted case as a lower bound.

### 2.1. Observation Simulation

The Fermi-LAT observation simulation program, *gtobssim*, that is part of the LAT Science Tools package supported by the Fermi Science Support Center (FSSC), uses parameterized instrument response functions (based on detailed Monte-Carlo simulations backed up by beam testing) to approximate the response of the LAT instrument in orbit.\(^2\) For each source, the user provides the spacecraft pointing history, a FITS sky map of the sources, and the total photon flux. Simulating operation in “sky-survey” mode (nearly uniform exposure) over a ten year period, we generate realistic diffuse background predictions as well as predictions for the WIMP (Weakly Interacting Massive Particle) gamma-ray signal.

The simulation includes the dependence of the LAT effective area on the viewing angle and photon energy, after accounting for all selection effects, including the trigger, the on-board filter, and the extensive offline analysis used to reduce cosmic-ray background to a low level. The simulation of the point-spread function (PSF) accounts for

\(^2\)http://Fermi.gsfc.nasa.gov/ssc
the dependencies on inclination angle and energy and also includes a parametrization of
the significant non-Gaussian tails. Conversions in the thin versus thick tungsten foils are
separately parameterized, an important detail considering that almost half of the LAT
effective area is from thick-foil conversions, for which the angular resolution is roughly a
factor of two worse. The observation simulation includes the rocking of the instrument
toward the orbital poles to improve uniformity of the all-sky exposure, and it also takes
into account dead time, in particular the passages through the South Atlantic Anomaly,
during which triggering of the LAT is disabled (Atwood et al. 2009).

We consider WIMPs that are capable of pair annihilation, for example a hypotheti-
cal stable supersymmetric partner of the gauge and Higgs bosons, namely the neutralino.
The WIMP remains a promising candidate for the role of DM, in particular for indi-
rect detection by astronomical observations (Bergström 2000). After choosing a mass,
$M_\chi$, and a thermally-averaged, velocity-weighted annihilation cross section, $\langle \sigma v \rangle$, we use
DarkSUSY (Gondolo et al. 2004) to simulate by Monte-Carlo the gamma ray spectrum
that results from WIMP annihilation and the subsequent fragmentation. In keeping
with the model used in Kuhlen et al. (2008), we compute this spectrum for WIMP
masses ranging from 50 to 500 GeV, assuming a 100% branching ratio into $b\bar{b}$ quarks
and $\langle \sigma v \rangle = 3 \times 10^{-26}\text{cm}^3\text{s}^{-1}$. Fig. 3 shows how a 100 GeV spectrum compares with the
backgrounds in different regions of the sky.

Since the energy spectrum of the DM signal has no correlation with the spatial
distribution, we calculate the overall flux for input to gtobssim as
\[ \Phi = \frac{1}{4\pi} \frac{\langle \sigma v \rangle}{2M^2} \int_{E_{TH}}^{M_\chi} \frac{dN}{dE} dE \int_{\Omega_{los}} \rho(l)^2 \, dl, \]

where \( \rho(l) \) is the density in the VL2 simulation, \( dN/dE \) is the annihilation spectrum, and \( E_{TH} = 500 \text{ MeV} \) is the minimum photon energy accepted in our analysis.

We repeated the simulations for each WIMP mass using VL2 subhalo sky maps corresponding to 10 random viewpoints at 8 kpc radius around the VL2 galaxy. A total of ten years of Fermi-LAT operation was simulated for each viewpoint. Figure 2 shows the full sky for an example simulation of the signal over a ten year observation.

The DM subhalo signal competes with a diffuse background from four categories: extragalactic, Galactic, host-halo DM, and unresolved DM. We simulate extragalactic diffuse emission using the power law fit determined in the recent Fermi-LAT analysis (Abdo et al. 2010). For Galactic emission, we use the GALPROP (Strong et al. 2000) (v50.1p) cosmic-ray propagation code to produce a map of diffuse gamma-rays originating from cosmic ray interactions with the Galactic interstellar medium and radiation field. The “optimized” GALPROP model (Strong et al. 2004) used in Kuhlen et al. (2008) relaxes constraints imposed by measurements of the local cosmic-ray flux, especially for cosmic-ray electrons, in order to account for the “GeV excess” seen by EGRET (Hunter et al. 1997). Since initial Fermi-LAT observations show no such excess (Abdo et al. 2009), we return to the “conventional” model, which keeps the constraints imposed by local cosmic-ray fluxes. We did not include the LAT residual cosmic-ray background in the model. It is roughly one third the extragalactic diffuse contribution and if included
would decrease our sensitivity to DM subhalo objects by an estimated 20%.

Galactic sources were not included in the background model. Since their gamma-ray signals are likely to overlap with subhalos, especially at low energy, they will reduce the sensitivity with respect to what is presented here. Furthermore, two or more point sources that are nearly coincident on the sky can mimic an extended object in the gamma-ray signal, producing a background to be dealt with when analyzing Fermi-LAT data. In general, when a subhalo candidate is detected it will be necessary to try to reject point source hypotheses through spectral analysis, analysis of the angular shape or source extension (see §2.2), searches for temporal variations, and multi-wavelength studies (Baltz et al. 2007).

In addition to the Galactic and extragalactic diffuse, the background includes photons from annihilation of DM residing in the smooth host halo. For the case where we boost the substructure, this smooth component also includes the extrapolated contribution due to unresolved subhalos. The total mass of the extrapolated substructure has the same normalization as the boost described in §2, but an anti-biased distribution due to tidal stripping and destruction of subhalos at inner radii. (Kuhlen et al. 2008)

We generated ten-year Fermi-LAT Monte Carlo simulations of all four diffuse background sources using \textit{gtobssim}, by the same procedure as used for the signal subhalos. All Monte-Carlo simulation counts were binned, by the Fermi-LAT Science Tool \textit{gtbin}, into four logarithmically spaced energy bands from 500 MeV to 300 GeV for use in §2.2.

In order to assign a detection significance to the region occupied by a particular
subhalo, we first find the total simulated background counts $\lambda$ in the region and then calculate the Poisson probability

$$P = \sum_{i=k}^{\infty} \frac{\lambda^i e^{-\lambda}}{i!}$$

for the background to fluctuate to a level equal to or greater than the observed counts $k$. From $P$ we derive a significance expressed in terms of standard deviations for a normal distribution:

$$S = \sqrt{2} \text{erf}^{-1}(1 - 2P).$$

(4)

For each subhalo we take all photons above 500 MeV in a region of interest (ROI) set at one degree, which corresponds roughly to the 68% containment angle in the LAT at 500 MeV. (Atwood et al. 2009) Fine-tuning this radius does not dramatically change the results given here. Finally, for a given map we quote $N_5$ ($N_3$), the number of subhalos above five (three) standard deviations significance.

### 2.2. Resolution of Angular Structure

One of the crosschecks needed to bolster the case for a DM subhalo candidate is to look for extended emission that is inconsistent with a point source. Since the LAT PSF has a strong energy dependence, we employ a likelihood-ratio analysis to compare for each significant subhalo the point-source hypothesis to a subhalo hypothesis. Instead of using a spatially unbinned event list, as in the previous section, we bin the sky into a grid of $2400 \times 1200$ pixels, equally spaced in Galactic longitude and latitude. For candidates near the galactic poles the counts are binned with a coordinate-system rotated 90 degrees.
in l and b to avoid distortion. Then for each VL2 subhalo, we convolve the VL2 “true” sky map with a parametrization of the LAT PSF and calculate the Poisson likelihood $L_1$ for the binned Monte-Carlo results. We repeat the calculation using a map of a point-source convolved with the PSF and calculate a second likelihood $L_2$ to arrive at a test statistic that compares the subhalo and point-source hypotheses:

$$TS = 2 \ln \left( \frac{L_1}{L_2} \right).$$

(5)

The PSF model used for the convolution is a simplification of what is in the \texttt{gtobssim} Monte Carlo. We assume that there is no azimuthal dependence around the source direction, while the radial distribution is given by the functional form used in the Science Tools to model the non-Gaussian tails of the LAT PSF:

$$p(r) = \frac{r}{\delta^2} \left( 1 - \frac{1}{\gamma} \right) \left[ 1 + \frac{1}{2\gamma} \left( \frac{r}{\delta} \right) \right]^{-\gamma}$$

(6)

where $\gamma = 2$. The 68% containment radius of this distribution is about $2.9\delta$. We match that radius in each energy bin to the documented LAT 68% containment angles (Atwood et al. 2009), taking into account the different PSF for conversions in thick tungsten versus thin and the relative LAT effective areas for the two conversion types. The convolved subhalos should fit well to their simulated counterparts, which we verified to be the case for the subhalos of interest.

The TS is only an indicator of how well the data could in principle statistically distinguish an extended subhalo from a point source. To rule out a point-source hypothesis at some confidence level would require something different, such as a goodness-of-fit test for that hypothesis. Success would rely on a thorough Monte-Carlo modeling of...
the PSF, including a full detector simulation (gtobssim uses only a parametrization of the response), plus understanding of any associated systematic problems. Furthermore, this method of calculating the $TS$ is not applicable to analysis of data, for which the true subhalo shape is unknown. Data analysis will require fitting to a subhalo model with a-priori unknown parameters and will also require dealing with backgrounds such as partially overlapping point sources, which are not considered here.

3. Results

Figure 4 shows our results for the number of detectable subhalos with significance greater than five (and three) versus WIMP mass. The two plots represent the different assumptions on the boost of the subhalo luminosity, as discussed in § 2. For each the shaded region indicates the range of variation among the ten different observer positions around the Galaxy with a dark line indicating the average.

We find that $N_5$ ranges from eight to zero over the range of $M_\chi$ and subhalo boosts. In all cases, there are no detections above 5$\sigma$ with a WIMP mass greater than 200 GeV. For comparison, the fiducial model considered in Kuhlen et al. (2008), where

$$(M_\chi/\text{GeV}, \langle\sigma v\rangle/(10^{-26}\text{cm}^3\text{s}^{-1}), \alpha, m_0) = (100, 3, 2.0, 10^{-6}) ,$$

predicted an $N_5$ ranging from 13 to 19. Scaling this down to account for ten years of orbit, including trigger and SAA dead time, leaves approximately 11 to 16 $N_5$. For that same DM setup, but with the new, more detailed LAT-specific analysis and the improved substructure boost implementation, we find here results for $N_5$ ranging from
zero to three with an average over all positions of 1.6. This difference comes largely from a lower effective area and angular resolution than that assumed in Kuhlen et al. (2008). Other factors include the usage of Poisson statistics rather than a simplified $S = N_s/\sqrt{N_b}$, and a factor of two (upward) error in the Kuhlen et al. (2008) flux calculation.

These detections are distributed roughly isotropically over the sky, as can be seen in Fig. 5. Requiring that a subhalo have resolvable angular extent (i.e. that $TS \geq 25$ ($\simeq 5\sigma$)), removes five of the 93 $N_5$ subhalos from the set of all 50 GeV Galactic positions, and three of 39 in the 100 GeV set. Overall, 95% of $S > 3$ detections have $TS > 25$. This drops to 68% for the case with no sub-substructure boost. A plot of the relation between TS and significance is given in Fig. 6.

While we plot $N_5$ here only as a function of $M_\chi$, the signal strength is also dependent on the annihilation cross section, $\langle \sigma v \rangle$, a quantity that has a plausible range spanning orders of magnitude. To a lesser extent, uncertainty in the the boost parameters $\alpha$ and $m_0$ and the nature of the central extrapolation of the density profile, i.e. whether an NFW or Einasto profile is assumed, can all make roughly factors $\sim 2$ differences in the $N_5$ (Kuhlen et al. 2008).

4. Discussion

We have estimated the sensitivity of the Fermi-LAT to WIMP annihilation gamma-rays from Galactic DM substructure predicted by the VL2 simulation. Using a thorough
instrument simulation, we add to predictions such as Kuhlen et al. 2008 a more accurate treatment of the LAT’s capabilities. Many of the details of the instrument’s behavior, e.g. reduced live-time due to the SAA, a different PSF in the thick and thin foil converters, etc. lead to a less optimistic predicted sensitivity. Combined with a less concentrated sub-substructure boost, these factors leave room for very few expected detectable subhalos over the LAT’s lifetime, given a conventional DM candidate.

Beyond the pure signal-to-noise ratio, resolution of the angular extension of a subhalo would go a long way to make it a convincing direct DM detection candidate. By computing a test-statistic for each possible predicted detection, we estimate the fraction which could be confidently deemed extended. Roughly 83% of our significant detections should have $\geq 5\sigma$ resolution of angular extent. This is a somewhat optimistic prediction as we use the true subhalo density profile in our likelihood ratio, instead of a generic (NFW, Einasto, etc..) profile as would probably be used in a real blind substructure search. Even so, based on this study of subhalos resolved in the VL2 simulation, it seems likely that a candidate bright enough to be significant will also appear extended.

If the Fermi-LAT does not resolve any significant DM substructure, flux limits will be of value. In this case it is important to understand the astrophysical uncertainties in predictions like the one given here. The strongest of these come from the extrapolations needed to compensate for finite numerical resolution in N-body simulations such as VL2, and the stochastic nature of our position in the Galaxy relative to individual subhalos. Simulating signals from multiple random positions at our Galactic radius gives an estimate of the latter uncertainty. We account for unresolved sub-subhalos

104
within subhalos resolved in VL2 by the boost correction Eqn. (1). Comparison with un-
boosted results illustrate that this effect is of some importance, while its exact amplitude
remains uncertain. As in Kuhlen et al (2008 and 2009) we have neglected any possibly
significant signals from subhalos that are too small to be resolved in the VL2 simulation
and we refer to these works for detailed discussions of their potential detectability: Pieri
et al. (2008); Siegal-Gaskins (2008); Ando (2009).

Note also that in models in which the pair annihilation of non-relativistic DM
particles is enhanced by the Sommerfeld effect (e.g. Arkani-Hamed et al. 2009), the
subhalo signal can be strongly enhanced owing to the lower velocity dispersion of the
DM in subhalos (Robertson & Zentner 2009; Lattanzi & Silk 2009). Even the first year
of Fermi-LAT data will strongly constrain many such scenarios (Kuhlen et al. 2009;

We look forward to the Fermi team’s upcoming analyses of the first year’s LAT
data, searching for a DM annihilation signal from dark subhalos. An analysis of the DM
signal from known dwarf galaxies is already published (Abdo et al. 2010).

Support for this work was provided by NASA through grants HST-AR-11268.01-
A1, NNX08AV68G (P.M.), and NASA Guest Investigator grant NNX08AV59G (B.A.).
This research used resources of the National Center for Computational Sciences at Oak
Ridge National Laboratory, which is supported by the Office of Science of the Depart-
ment of Energy under Contract DE-AC05-00OR22725. Fermi-LAT simulation tools were
provided by the Fermi-LAT collaboration and the Fermi Science Support Center.
REFERENCES


Ando, S. 2009, Phys. Rev. D, 80, 023520


Physics, 21, 267.


Kuhlen, M., 2010, Advances in Astronomy, Article ID 162083


Siegal-Gaskins, J. M. 2008, Journal of Cosmology and Astro-Particle Physics, 10, 40


Fig. 1.—: A comparison of the old and new boost prescriptions, performed before the observation simulation on the prominent sub halo near the right edge of the projection in Fig. 2d. The dotted line represents the new boost; the dashed and solid represent the old and un-boosted profiles, respectively.
Fig. 2.— All-sky $\log(\text{counts})$ maps for 100 GeV WIMP subhalo annihilation signal and the three background sources for ten years of Fermi-LAT orbit as seen by an observer along the intermediate halo axis. The second panel (2b) includes an extrapolation for subhalos which are unresolved in the VL2 simulation, which we call DM Diffuse.
Fig. 3.—: The relative contributions to the entire sky, including all subhalos (3a), and the inner five degrees of a single detectable subhalo (3b) for ten years of simulated Fermi data from a 100 GeV WIMP and the four background sources. The single subhalo is located at (l, b) = (195.45, -11.25). Bins are equally spaced in log(E).
Fig. 4.—: The number of subhalos above five (and three) standard deviations significance, \(N_5\) (\(N_3\)), as a function of WIMP mass, \(M_\chi\), for \(\langle \sigma v \rangle = 3 \times 10^{-26}\) cm\(^3\) s\(^{-1}\). The shaded regions show the range of variation among ten different observer positions, while the dashed and solid lines represent the average over all positions. The subhalos on the left have been boosted for unresolved substructure while those on the right have not. The simulations represent a Fermi-LAT observation time of ten years.
Fig. 5.—: The cumulative distribution of $N_3$ over galactic latitude for two choices of WIMP mass. The subhalos here are boosted and the shaded regions span the variations due to different observer positions around the Galaxy. The dashed lines represent the expected behavior for an isotropic distribution.
Fig. 6.—: The significance versus test statistic (TS) of $S \geq 3$ boosted subhalos, including all Galactic positions. The line marks the cut made at TS = 25. Subhalos above this line are significantly ($\simeq 5\sigma$) more extended than point sources.
Chapter 6

Analysis of the Galactic Halo

6.1 Introduction

Although DM substructure represents a very high possible signal-to-noise ratio, its detectability depends critically on the actual physical locations of the clumps. We rely on having a very close or very large object in an uncomplicated region of sky, i.e. not overlapping another source or the galactic plane (GP). So far, none have been conclusively observed (Ackermann et al. 2012), indicating that either DM does not shine brightly, or that nature was not kind to us with regard to the distribution of substructure. It seems prudent at this point to focus attention on searches where the nature of DM is the only unknown (as far as signal goes).

We know that all this substructure resides in a much larger Host Halo, where annihilation or decay could also take place. Depending on the shape of the inner profile, the Halo could produce more gamma rays than all resolvable substructure combined.
Unfortunately, the brightest region lies exactly atop the galactic center (GC), a notoriously complex and poorly understood region. The Halo is huge, however, and although less dense than the core, regions extending beyond the GP can also contribute significant signal. These regions are the focus of the search described by this chapter.

Because we expect DM Halo photons to be spread over the entire sky, we can statistically afford to mask the most complicated background regions: the GP and individual point sources. This allows us to choose simple and robust models that still capture the diffuse character of the multiple backgrounds. A Halo analysis is therefore somewhat of a compromise between strong signal and complex background: more signal than a dwarf galaxy search, but less background than an analysis of the GC.

6.2 Modeling

6.2.1 DM Signal

Chapter 4 describes the believed distribution of DM in the Milky Way, and Chapter 3 examines the possible gamma-ray signals that could originate from the WIMPs comprising the DM. Of all the possibilities afforded by the uncertainty in these two fields, one combination stands out as very conservative\(^1\) and yet still generic and interesting. This is a “naked,” or substructure-less Halo where WIMP annihilation takes place exclusively into the $b\bar{b}$ quark channel.

\(^1\)For avoiding false detections.
distribution, 
\[ \rho(r) = \rho_{\odot} \exp \left( -\frac{2}{\alpha} \left[ \left( \frac{r}{r_{\odot}} \right)^\alpha - 1 \right] \right) \]  
(6.1)

where \( \rho_{\odot} \) and \( r_{\odot} \) are the local DM density and galactic radius. Here we will take these to be 0.386 GeV cm\(^{-3}\) (Catena and Ullio 2010) and 8.5 kpc, respectively. The power law, \( \alpha \), is set to 0.17. The DarkSUSY (Gondolo et al. 2004) package again supplies the spectrum for the \( b \bar{b} \) annihilation for a range of WIMP masses. Further variations of DM, including more masses, channels, decay models, and amounts of substructure will be explored in a future Fermi Collaboration (FC) paper.

### 6.2.2 Isotropic Background

The simplest gamma-ray backgrounds to model have isotropic morphology, i.e. every direction on the sky delivers the same flux. There are two origins for such background, unresolved astrophysical sources outside the galaxy (EGB), and CR interactions in the outer, inactive portions of the Fermi-LAT. Both measured by the FC, these take the form, respectively, of a featureless power law with index \( \gamma = 2.41 \pm 0.05 \) and the on-orbit calibrated background rate shown in Figure 6.1 (Abdo et al. 2010d).

### 6.2.3 Galactic Background

Uncertainty in both its spectrum and morphology makes the galactic diffuse model nearly as unknown as the DM signal. The emission presumably originates with CR accelerated by galactic sources such as supernova remnants (SNR) or pulsars. As
Figure 6.1: Fermi-LAT CR background rate for two different instrument response functions (Abdo et al. 2010d).
these particles traverse through the gas, dust, and light of the galaxy, they lose energy in processes that produce gamma rays. GALPROP (Strong et al. 2004), a particle diffusion and interaction package, follows these processes from start to finish in a two-dimensional\textsuperscript{2} simplification of the Milky Way. Although we understand the particle interactions quite well, it is not possible to know the local conditions (magnetic fields, light and gas density, etc.) in every part of the galaxy. Therefore, codes like GALPROP will always suffer from systematic uncertainties in the parameters that define the environment in which its particles propagate. The remainder of this section outlines the nature and extent of these uncertainties, and Section 6.3.2.2 deals with their marginalization in the analysis.

\subsection*{6.2.3.1 Interaction Processes}

The dominant gamma-ray producing CR interaction processes in the Milky Way are pion production, bremsstrahlung, and inverse Compton scattering (ICS). Pion production in this sense refers generally to a CR-Hydrogen collision that results in $\pi^0$ that rapidly decay into gamma rays. This is especially sensitive to the assumed primary proton flux (Sec. 6.2.3.3) and the target gas density (Sec. 6.2.3.2). Bremsstrahlung means “braking radiation,” and occurs when a charged particle is accelerated by a strong electromagnetic field (another charged particle) and emits a photon. This process has dependencies similar to pion production, except that it includes electron and positron fluxes. Finally, ICS gamma rays come from charged particles up-scattering starlight or

\textsuperscript{2}Radius and height above the GP.
CMB photons. In the Milky Way, this flux is dominated by $e^+/e^-$ interactions and is thus highly sensitive to their source properties in addition to the collected photon target population, or interstellar radiation field (ISRF).

Pion production and bremsstrahlung are calculated within GALPROP to a degree of accuracy that reflects our best understanding of the processes. It saves quite a bit of compute time, however, to run GALPROP in a mode which approximates all ICS as an isotropic process. That is, averaging over all scattering angles, assuming no preferred angle of interaction. Of course this is not the case, since photons generally stream outward from the galactic plane and center, where they are up-scattered by randomly oriented electrons. Fortunately, this approximation can be somewhat corrected post-simulation. Running the full anisotropic calculation once for several combinations of galactic halo height ($Z_H$) and diffusion coefficient ($D_N$) and then finding the ratio of these maps to their isotropically approximated counterparts, provides a set of corrective templates. Each new isotropic model is then corrected by multiplying it with the template most similar in $Z_H$ and $D_N$.

6.2.3.2 Target Densities

There are the five main ingredients that comprise the Milky Way interstellar medium (ISM): stellar and CMB photons, and molecular (H$_2$), atomic (HI), and ionized (HII) hydrogen. Measurements of each of these carry some uncertainty in both morphology and overall normalization. Changes to the target density in any region lin-
early scale the interaction rates within it.¹³ GALPROP keeps track of the interactions in each medium and at a variety of galactic rings separately, allowing modifications to any model’s target densities to be made post-CR propagation by simple scaling. Fine resolution of the gas in the inner galactic rings only matters in the galactic plane, which we mask, and the outer rings contribute very little to the total emission. Therefore, we reduce the native GALPROP resolution of the galactic ring contributions to only three: inner (< 8 kpc), local (8-9 kpc), and outer (> 9 kpc).

The ISRF originates mainly with starlight and the CMB, see Porter and Strong (2005). Although measurements of it have uncertainties of their own, changes to the ISRF normalization are degenerate with the primary electron flux, a parameter described in Section 6.2.3.3. The CMB on the other hand has an isotropic distribution and a spectrum measured with no relevant uncertainty. For these two reasons, the ISRF is not given explicit uncertainty in the background model.

Hydrogen in the MW takes form and morphology according to its temperature. Cool and dense, H₂ clings to the galactic plane. Largely inert, its distribution is measured using the rotational carbon monoxide line as a proxy, a molecule thought to have a similar distribution. In fact, GALPROP uses CO maps directly as the H₂ distribution, but with a factor, $X_{CO}$, that scales the density appropriately. The actual uncertainty in $X_{CO}$ makes for a complex discussion (The Fermi-LAT Collaboration 2012), but here we simply allow it to take values of uniform prior within a factor of 0-2 of the nominal.

³Within the diffuse regime, that is. Target densities exceeding a critical value would begin to provide enough radiation lengths to stop the propagating CR.
model outside of the galactic plane is relatively minor.

Warmer than H2, HI diffuses much further out of the GP and the bremsstrahlung and $\pi^0$ production that take place within it contribute the bulk of the gamma-ray emission in our model. Detected by the 21-cm line, its uncertainty comes from the assumed spin temperature of the HI gas. The proton-electron spin alignment of a hydrogen atom causes hyperfine splitting at 21 cm, effectively giving the line some opacity\(^4\). The level of this opacity controls the inferred column density. Spin temperature can range from 85–10,000 K (The Fermi-LAT Collaboration 2012), corresponding to an overall HI density normalization factor of 0.9–1.07.

Finally we have HII, the proton half of a hot hydrogen plasma. Extending far out of the galactic plane, this component of the ISM may be low density, but it facilitates gamma-ray production from a huge portion of the sky. No longer electrically neutral, its density is measured by the dispersion of radio pulsar photons. We adhere to Gaensler et al. (2008), a study built on many such measurements, for an HII target distribution. This assumes a distribution out of the GP with the following form:

\[
n(z) = n_0 \exp\left( -\frac{z}{H_n} \right),
\]

where \( H_n = 1010^{+40}_{-170} \) pc and \( n_0 = 0.031^{+0.004}_{-0.002} \). Uncertainty in the scale height, \( H_n \), can be ignored, as it is much smaller than the annular binning scheme resolution described in Sec. 6.3.2.1. Accounting for the range of \( n_0 \) corresponds to an HII model normalization

\(^4\)For example, if half of the HI had its spins aligned and the other half anti-aligned, photons readily absorbed by one set would be less likely for the other. This effectively reduces the radiation lengths for a given column of gas.
uncertainty of 0.93–1.13.

6.2.3.3 CR Source Properties

SNR and pulsars are thought to be the progenitors of CR in the Milky Way, imparting energy to particles through Fermi acceleration or intense fields. Not many details are actually known, however, since the acceleration regions are small and difficult to observe, and the random nature of their propagation prevents all but the highest energy CR from pointing back to a source. A model of the galactic diffuse emission must therefore incorporate several uncertain parameters to cover the range of realistic possibilities. This analysis uses five: the electron injection index, two proton injection indices (assuming a fixed break), one describing the galactic CR source distribution, and finally the electron injection flux normalization.

Measurements at Earth\textsuperscript{5} constrain the electron and proton injection indices to the values listed in Table 6.1. Population studies of SNR and pulsars lead to different source distributions. The distribution parameter maps CR sources to either pure SNR/pulsars, or a hybrid of the two.

Lastly, we have uncertainty in the electron injection flux. Electrons traversing the galaxy lose energy extremely rapidly, with an approximate rate of (Pohl and Eichler 2010)

\[
E'(E) \approx -2.5 \times 10^{-14} \text{ GeV s}^{-1} \left( \frac{E}{10 \text{ GeV}} \right)^2
\]  

\textsuperscript{5}Including data from ACE, HEAO-3, AMS, the Fermi-LAT, and H.E.S.S. (The Fermi-LAT Collaboration 2012)
Proximity the nearest emitter can therefore strongly skew any local measurement. We have only measured this spectrum at Earth, and can hardly assume it represents the average at our galactic radius. In the past, deviations as strong as a factor of four from the local measurement were invoked to explain the local gamma-ray spectrum (Strong et al. 2004). Here we consider the more modest range of 0–2.

6.2.3.4 Propagation Conditions

Sections 6.2.3.2 and 6.2.3.3 detailed the range of uncertainty in CR sources and their targets. Simulating their propagation from one to the other requires four more parameters. First, the diffusion coefficient characterizes the rate at which CR scatter on magnetohydrodynamic inhomogeneities, and the thus the characteristics of their random walk; increasing it allows them to move faster and further through the galaxy. Second, because CR diffusion depends on the energy of the particle, we have a diffusion index, or $D_g$. Third, the Halo height, $Z_H$. CR that move outside this value escape the galaxy. Fourth and finally, the Alfvén velocity, $v_A$, or wave speed in the ionized MW medium is important for re-acceleration. Tightly constrained by local CR measurements (see Sec.6.3.1), all of these parameters start with the broad uncertainty ranges shown in Table 6.1 and are constrained by the fit.

---

6In a two dimensional, diffusive, non-convective environment. See Strong et al. (2007) for a thorough discussion.

124
Table 6.1: Nuisance parameters of the diffuse background model. Those marked linear are scalable post-propagation; others require a full recalculation to adjust. Ranges marked with an * are penalized in the fit with a gaussian set to 1-σ at the indicated limits.
6.3 Data Confrontation

6.3.1 Cosmic Rays

Tracing the propagation of each individual species, every GALPROP model is strongly constrained by local CR measurements. The ratio of Boron to Carbon, for example, provides a measure of the amount of material traversed between the source and measurement\(^7\). Comparing the simulated B/C ratio with data taken at Earth rules out models with a diffusion halo that is too large/dense or small/sparse. In a similar way, the ratio of the Beryllium isotopes Be\(^{10}\)/Be\(^{9}\) constrains the amount of time a CR can spend diffusing before it reaches Earth.\(^8\) These two ratios, along with the proton spectrum, comprise a discriminating \(\chi^2\) for each GALPROP model:

\[
\chi^2 = \Sigma_j \Sigma_i N_i \frac{(D_{ij} - T_{ij})^2}{\sigma_{ij}^2 + \Delta\phi_{ij}^2}.
\] (6.4)

Here D and T are model and data respectively, and the sum is over all points and experiments. The \(\Delta\phi^2\) term accounts for uncertainty in the solar modulation, or repellant force of the Sun on charged particles. Modulating according to the measured solar value during each experiment\(^9\) ± 100 MV defines the uncertainty values which are then added in quadrature to the experimental error bars.

Data from the following experiments represents our best knowledge of the local proton spectrum and the two nuclei ratios: HEAO-3, IMP, ATIC-2, CREAM, ACE,\(^7\) Boron is produced solely by spallation, so the more material traversed, the more Carbon is converted.\(^8\) Be\(^{10}\) is unstable and decays while Be\(^{9}\) does not.\(^9\) The modulation force changes with the solar cycle.
Figure 6.2: Comparison of an exemplary GALPROP model with local measurements of the Boron to Carbon ratio. The blue regions encompass the uncertainty due to solar modulation.

ISOMAX, AMS01, CAPRICE, and BESS. See Figures 6.2-6.4 for example fits from a single GALPROP model. The ATIC-2 data, used for its proton measurement, actually measures spectral details too fine for GALPROP to reproduce. Scaling its error bars up by a somewhat arbitrary factor of four allows the data to constrain the essentially power-law proton simulated spectrum without overly penalizing it.
Figure 6.3: As Figure 6.2 but for the ratio of Beryllium isotopes.
Figure 6.4: As Figure 6.2 but for the proton spectrum.
6.3.2 Gamma Rays

6.3.2.1 Fit Preparation

Before comparing with the LAT data, we must convolve the DM signal and each model component with the instrument response functions and exposure. The GARDIAN (The Fermi-LAT Collaboration 2012) program does just that – converts the model fluxes into counts anticipated for 32 months of P7v6 Fermi-LAT observation. Both the model and data are then binned into 12 annular spatial bins symmetric about the galactic center (see Fig. 6.5) and 46 logarithmic energy bins from 1 to 100 GeV. Noisy regions are masked: the GP ($-10 < b < 10$) and point sources with a mask that decreases in size with increasing energy according to the LAT PSF. The tails of this PSF extend (weakly) outside of any mask, however. Adding a simulation of these residuals using Fermi-LAT software\footnote{http://fermi.gsfc.nasa.gov/ssc/site_map.html} tuned to the Fermi First Source Catalogue (Abdo et al. 2010a) measured source parameters, to the model compensates for this.

6.3.2.2 Fitting Linear and Non-Linear Parameters

An ideal search for diffuse DM would fit both the signal and background models, including all free parameters, to the data simultaneously. Unfortunately, many of these parameters have non-linear effects on the model and require a full re-run of GALPROP to change. While it is theoretically possible to build GALPROP into the fitting procedure, it is computationally prohibitive. Therefore, fitting the gamma-ray sky becomes a hybrid between a continuous and a discrete process.
Figure 6.5: Annular binning scheme, including mask of the galactic plane. The bottom plot shows the signal-to-noise ratio vs. annulus for the 100 GeV WIMP and LAT data. Note that the cumulative S/N ratio is roughly constant by annulus 6.
For example, say we choose a set of non-linear parameters, $\vec{\beta}$, and use GAL-PROP to generate a model. Keeping the default linear parameters, $\vec{\alpha}$, this model matches the gamma-ray data with a Poisson Likelihood (Mattox et al. 1996), $L(\vec{\alpha}, \vec{\beta})$, that is the product of Poisson probabilities in each bin. A package such as MINUIT (James and Roos 1975) can then partially maximize the Likelihood by varying $\vec{\alpha}$. We can denote this as $L(\vec{\alpha}^*, \vec{\beta}, \theta_{DM})$, where $\theta_{DM}$ represents the DM normalization in the fit. Keeping $\theta_{DM}$ separate is important; the goal here is to measure detection significance as a function of it.

We cannot forget to account for the CR-fit, but doing so is as simple as multiplying by another probability. This combination is the total, partially maximized, Likelihood:

$$\tilde{L}(\vec{\alpha}^*, \vec{\beta}, \theta_{DM}) = P_{CR}(\vec{\beta}) \times L(\vec{\alpha}^*, \vec{\beta}, \theta_{DM}).$$

which can also be written as

$$\ln(\tilde{L}) = \ln(P_{CR}) + \ln(L) = \frac{\chi^2}{2} + \ln(L).$$

Now, if we form a null hypothesis, $\tilde{L}(\vec{\alpha}^*, \vec{\beta}, \theta_{DM} = 0)$, we have the essence of a Test Statistic, or measure of the DM fit significance.

$$\tilde{T}_S(\beta, \theta_{DM}) = 2\ln \left( \frac{\tilde{L}(\beta, \theta_{DM})}{\tilde{L}(\beta, \theta_{DM} = 0)} \right)$$

This is all well and good, but these Likelihoods are only partially maximized. The fitter
took care of $\vec{\alpha}$, but we still have to maximize $\vec{\beta}$ by generating discrete GALPROP realizations before the TS can be solely a function of $\theta_{DM}$. The resultant, fully marginalized TS is known as a Profile Likelihood,

$$
TS(\theta_{DM}) = 2 \ln \left( \frac{\hat{L}(\theta_{DM})}{L(\theta_{DM} = 0)} \right) = 2 \ln \left( \frac{\tilde{L}(\beta^*, \theta_{DM})}{\tilde{L}(\beta^*, \theta_{DM} = 0)} \right).$$  \hspace{1cm} (6.8)

Section 6.3.4 details the methods used to explore and maximize the $\vec{\beta}$-space. The TS($\theta_{DM}$) curve should behave like a $\chi^2$ with a single degree of freedom. Its maximum value corresponds to the detection significance and we can set 95% confidence upper limits where it deviates by 3.84 from there.

### 6.3.2.3 GALPROP Fit Quality

Before applying any sort of detection or limit-setting procedure, we must consider how well GALPROP reproduces the Fermi-LAT data. Calculating $\hat{L}(\vec{\alpha}^*, \vec{\beta}, \theta_{DM} = 0)$ for over 2500 different sets of $\vec{\beta}$ results in a best-fitting model with the parameters listed in Table 6.2. The match is not spectacular. Figures 6.6-6.9 show the spatial and spectral fits and residuals. Large residuals persist, particularly near the galactic center. There are also several 5-10%-sized spectral features, amounting to a $\chi^2$/d.o.f. of over 10.

It is not entirely surprising that the model is poor. GALPROP was conceived with the intent to match data from all separate constraints equally, rather than any of them very accurately. The version used here is also missing the Fermi “lobe” feature.
Figure 6.6: Best-fitting uncorrected GALPROP model: spectrally integrated spatial fit.
Figure 6.7: Best-fitting uncorrected GALPROP model: spatially integrated spectral fit.
Figure 6.8: Best-fitting uncorrected GALPROP model: spatially integrated spectral residuals.
Figure 6.9: Best-fitting uncorrected GALPROP model: spatial and spectral residuals.
<table>
<thead>
<tr>
<th>Target Densities</th>
<th>Scale/Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISRF</td>
<td>1</td>
</tr>
<tr>
<td>Ionized Hydrogen (HII)</td>
<td>1.13</td>
</tr>
<tr>
<td>Atomic Hydrogen (HI)</td>
<td>1.07</td>
</tr>
<tr>
<td>Molecular Hydrogen (H$_2$)</td>
<td>1.0</td>
</tr>
<tr>
<td>Isotropic</td>
<td>1.22</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CR Source Properties</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$e^+/e^-$ Flux</td>
<td>2.0</td>
</tr>
<tr>
<td>$e^+/e^-$ Index</td>
<td>2.42</td>
</tr>
<tr>
<td>Nucleon Index (Low)</td>
<td>2.01</td>
</tr>
<tr>
<td>Nucleon Index (High) ($n_g$)</td>
<td>2.43</td>
</tr>
<tr>
<td>Source Distribution</td>
<td>SNR</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Propagation Parameters</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Diffusion Coefficient</td>
<td>8.84e+28</td>
</tr>
<tr>
<td>Halo Height</td>
<td>8.44 kpc</td>
</tr>
<tr>
<td>Diffusion Index ($D_0$)</td>
<td>0.33</td>
</tr>
<tr>
<td>Alfvén Velocity</td>
<td>21.95 km/s</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fit Diagnostics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Gamma Rays: $\chi^2$/d.o.f.</td>
<td>10.54</td>
</tr>
<tr>
<td>Cosmic Rays: $\chi^2$/d.o.f.</td>
<td>1.75</td>
</tr>
</tbody>
</table>

Table 6.2: Parameters of best-fitting GALPROP model with no corrections or DM.
(Su et al. 2010), as its origins are still unclear. It is also likely that the LAT data carries an energy-dependent systematic error due to misestimation of the effective area. The error’s origin lies in the particle background rejection cuts made by the onboard filter and the trained CTs used in all LAT data preparation (Appendix A). Error in the efficiency estimate of any one of these cuts skews the effective area. The FC made some progress quantifying this uncertainty, using well-characterized steady sources such as pulsars to compare the efficiency of the cuts on real and MC data. The Fermi-LAT Collaboration (2009) limit the error to roughly 10% at 5 GeV and above.

Whatever the reason for the model quality, searching for DM using such a background would be problematic. Clearly some improvement must be made before we can proceed.

### 6.3.2.4 Outer Galaxy Correction

To account for effective area systematic and improve the raw GALPROP models, we will use something akin to common off-source background subtraction. As the off-source, we consider the entire outer galaxy, $90^\circ < l < 270^\circ$. Very few\footnote{A thousand photons, about 1/6 those from the inner galaxy.} DM photons are expected here, given the distributions discussed in Chapter 4, especially after the source masking scheme (Sec. 6.3.2.1) blocks out potential substructure signal. The region should still tell us something about the effective area systematic. Dividing the measured outer galaxy energy spectrum, $D_{\text{out}}(E)$, by that of the fitted model, $M_{\text{out}}(E)$, gives an ad-hoc energy-dependent correction to our model:
The outer galaxy data is actually a product of the true flux, \( d_{\text{out}}(E) \), and the effective area systematic, \( A_{\text{sys}}(E) \), leaving us with

\[
C_{\text{out}}(E) = \frac{D_{\text{out}}}{M_{\text{out}}}. \tag{6.9}
\]

The outer galaxy has model deficiencies of its own. Aside from the possible addition of a DM spectrum, \( \nu_{\text{dm}}(E) \), the preceding argument applies and we can describe the data/model ratio as,

\[
C_{\text{in}}(E) = \frac{(d_{\text{in}} + \nu_{\text{dm}})A_{\text{sys}}}{M_{\text{in}}}. \tag{6.11}
\]

If the inner and outer deficiencies were identical, we could use \( C_{\text{out}}(E) \) to perfectly correct the inner galaxy model. We must assume they can differ, though, and can represent the difference as a new systematic, \( C_{\text{sys}} \). Symbolically,

\[
C_{\text{out}}(E) = C_{\text{sys}} \times C_{\text{in}}. \tag{6.12}
\]

Including the previous equations the systematic can be represented as

\[
C_{\text{sys}}(E) = \frac{C_{\text{out}}}{C_{\text{in}}} = \frac{d_{\text{out}}A_{\text{sys}}/M_{\text{out}}}{(d_{\text{in}} + \nu_{\text{dm}})A_{\text{sys}}/M_{\text{in}}} = \frac{d_{\text{out}}M_{\text{in}}}{(d_{\text{in}} + \nu_{\text{dm}})M_{\text{out}}}. \tag{6.13}
\]

Assuming \( \nu_{\text{dm}} \ll d_{\text{in}} \) allows us to expand the denominator to first order.
\[ C_{\text{sys}}(E) \simeq \frac{d_{\text{out}} M_{\text{in}}}{d_{\text{in}} M_{\text{out}}} \left( 1 - \frac{\nu_{\text{dm}}}{d_{\text{in}}} \right) \] (6.14)

We now have a way of improving the fit in the on-source signal region, at the expense
of a new systematic. Explicitly applying it gives us

\[ M_{\text{in}} C_{\text{out}} = M_{\text{in}} C_{\text{sys}} C_{\text{in}} = C_{\text{sys}} D_{\text{in}} \] (6.15)

and then

\[ C_{\text{sys}} D_{\text{in}} \simeq \frac{d_{\text{out}} M_{\text{in}}}{d_{\text{in}} M_{\text{out}}} \left( 1 - \frac{\nu_{\text{dm}}}{d_{\text{in}}} \right) D_{\text{in}} = \frac{d_{\text{out}} M_{\text{in}}}{d_{\text{in}} M_{\text{out}}} (D_{\text{in}} - \Upsilon_{\text{dm}}) \] (6.16)

where \( \Upsilon_{\text{dm}} \) is the observed (post-\( A_{\text{sys}} \)) DM signal. So finally writing

\[ M_{\text{in}} C_{\text{out}} \simeq \frac{d_{\text{out}}}{d_{\text{in}}} \frac{M_{\text{in}}}{M_{\text{out}}} (D_{\text{in}} - \Upsilon_{\text{dm}}), \] (6.17)

makes it clear what we get out of applying the outer galaxy correction. We reproduce
exactly what the model should, \( (D_{\text{in}} - \Upsilon_{\text{dm}}) \), which is the observed data minus DM.

Any effective area systematic, being isotropic, is eliminated. In exchange, we add a
systematic which is the product of something we know, \( M_{\text{in}}/M_{\text{out}} \),\(^{12} \) and something we
cannot know, \( d_{\text{out}}/d_{\text{in}} \) – the ratio of the true outer/inner diffuse emission. Therefore,
to use \( C_{\text{out}} \) we must add uncertainty to the fit. The question is, what is an appropriate
form for the uncertainty?\(^{12}\)

\(^{12}\) \( M_{\text{in}} \) and \( M_{\text{out}} \), the spatially averaged inner and outer galaxy model spectra, usually differ by no
more than 10%.
We can get an idea of what $C_{sys}(E)$ looks like when assuming there is no DM signal which reduces the systematic to

$$C_{sys}(E) \sim \frac{D_{out}M_{in}}{D_{in}M_{out}} \quad (6.18)$$

Figure 6.10 shows this approximation to $C_{sys}(E)$ for the model set out in Table 6.2. It appears to be an effect on the order of 10-15%, with a smooth underlying shape. Other models show similar features. This means that adding a smooth function with free parameters to the fit can account for the new uncertainty. Adding a function with as few parameters as possible should be less likely to hide a DM signal. Fitting the systematic with polynomials from $O(1-4)$ reveals $O(3)$ as the clear favorite in terms of diminishing fit quality returns.

Let us recap the outer galaxy correction. In an effort to improve the energy residuals of natural GALPROP models and account for any possible effective area systematic, we split the galaxy in half and use the fit residuals in the outer galaxy to correct the inner galaxy model. Doing so cleans up effective area issues and improves the fit, but at the expense of adding a new systematic that comes from assuming the model problems in the outer galaxy are the same as those in the inner. Investigating an approximation of this systematic shows that it can be described with a third order polynomial. Therefore the fit procedure from here on will be to apply the outer galaxy correction and then fit including an $O(3)$ polynomial, free to float within a 10% envelope of unity. Figures 6.11-6.14 illustrate the fit improvement for the new best-fitting model, with parameters listed in Table 6.3.
Figure 6.10: Approximate systematic induced by applying outer galaxy spectral correction to the inner galaxy for a single GALPROP model. Also shown are polynomial fits to the systematic of up to fourth order.
Figure 6.11: Best-fitting outer galaxy-corrected GALPROP model: spectrally integrated spatial fit.
Figure 6.12: Best-fitting outer galaxy-corrected GALPROP model: spatially integrated spectral fit.
Figure 6.13: Best-fitting outer galaxy-corrected GALPROP model: spatially integrated spectral residuals. The blue curves depict the outer galaxy correction and the fitted polynomial which represents the correction uncertainty.
Figure 6.14: Best-fitting outer galaxy-corrected GALPROP model: spatial and spectral.
<table>
<thead>
<tr>
<th>Target Densities</th>
<th>Scale/Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISRF</td>
<td>1</td>
</tr>
<tr>
<td>Ionized Hydrogen (HII)</td>
<td>1.13</td>
</tr>
<tr>
<td>Atomic Hydrogen (HI)</td>
<td>[0.92, 1.07, 1.00]</td>
</tr>
<tr>
<td>Molecular Hydrogen (H$_2$)</td>
<td>1.0</td>
</tr>
<tr>
<td>Isotropic</td>
<td>1.43</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CR Source Properties</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$e^+/e^-$ Flux</td>
<td>0.86</td>
</tr>
<tr>
<td>$e^+/e^-$ Index</td>
<td>1.98</td>
</tr>
<tr>
<td>Nucleon Index (Low)</td>
<td>2.28</td>
</tr>
<tr>
<td>Nucleon Index (High) ($n_{g2}$)</td>
<td>2.26</td>
</tr>
<tr>
<td>Source Distribution</td>
<td>Pulsars</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Propagation Parameters</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Diffusion Coefficient</td>
<td>8.89+28</td>
</tr>
<tr>
<td>Halo Height</td>
<td>9.69 kpc</td>
</tr>
<tr>
<td>Diffusion Index ($D_g$)</td>
<td>0.5</td>
</tr>
<tr>
<td>Alfvén Velocity</td>
<td>21.80 km/s</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fit Diagnostics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Gamma Rays: $\chi^2$/d.o.f.</td>
<td>3.42</td>
</tr>
<tr>
<td>Cosmic Rays: $\chi^2$/d.o.f.</td>
<td>0.84</td>
</tr>
</tbody>
</table>

Table 6.3: Parameters of best-fitting GALPROP model with outer galaxy correction and no DM.
While the fit is certainly improved, from $\chi^2$/d.o.f. = 10 to 3.4, it is not ideal. The addition of a DM signal (with cross section at the thermal relic level) to the model still represents a small perturbation to the fit quality. If the problems with the model are of the same or greater scale, detections or upper limits on DM cannot be trusted. Therefore, as a final conservative step, we will scale up all the error bars until the best-fitting model (without DM) gives $\chi^2$/d.o.f. = 1. While this fails to take advantage of the full Fermi-LAT statistics, it makes no sense to do so while the background model is incapable of fitting well. It may be worthwhile to revisit this problem in the future, with a more developed model and better calibration of the effective area.

6.3.3 Pseudo-Experiments

In light of the corrections and systematics introduced in Sec. 6.3.2.4, it is important that we check the fit procedure for bias and artifacts. Assuming there is a DM signal, does the fitter find it? Does it get the mass and normalization right? Does it find DM when there is none? These are questions that require pseudo-experiments (PE), Monte-Carlo simulated data sets where we control the amount of signal. Fitting to many PE gives an estimation of our statistical biases.

The LAT data may well include a DM signal, and so cannot be used to generate PE. The only option then is to make Monte-Carlo realizations of GALPROP models. We must be careful here, for as outlined in Sec. 6.3.2.3, uncorrected GALPROP models differ substantially from LAT observations. Fitting a particular model to a PE based on itself is very different from fitting a poor model to data; it would be difficult to miss
a DM signal with a perfectly accurate background model!

There are several things we can do to make the PE more realistic. First, they are all given an energy residual corresponding to the average found in the outer galaxy of the actual data. Next, we can exclude any model that fits the pseudo-data better than the best fit to real data. Finally, because models which fit very poorly are unlikely to contribute to the Profile Likelihood, we require that the TS be within 50 of the maximum.\footnote{This is generous – most models within this will not contribute.} This defines a set of models whose fit to pseudo-data best represents the true situation and are referred to as “realistic” models from here on.

Fits to this set of models are made in exactly the same manner as for the data, including the outer galaxy correction and corresponding systematic polynomial. Figure 6.15 shows the recovered scale when the pseudo-data contains no DM. False signals are rare, with the 1-\(\sigma\) widths barely reaching past a signal scale factor of one (corresponds to thermal relic cross section). PE using the same model that generated the pseudo-experiment are also shown, predictably recovering the input with high accuracy.

If the PE are to be believed, it seems the procedure allows for few false detections. Inserting a 100 GeV \(b\bar{b}\) WIMP signal at the thermal relic cross section into the same PE results in Figure 6.16, where significance is also plotted in Figure 6.17. Realistic models do almost as well as the generator model, although with a long tail in the recovered signal scale. The high-scale results rarely have significance greater than 3\(\sigma\), however.

Finally, in the event of DM non-detection, we will try to set an upper limit,
Figure 6.15: Pseudo-experiment results with no DM signal: recovered scale.
Figure 6.16: Pseudo-experiment results including a 100 GeV WIMP signal: recovered scale.
Figure 6.17: Pseudo-experiment results including a 100 GeV WIMP signal: detection significance.
Figure 6.18: Pseudo-experiment 95% upper limits including a 100 GeV $b\bar{b}$ WIMP DM signal.

constraining the nature of the WIMP using the method described in Sec. 6.3.2.2. To do this, PE must demonstrate that our process does not set limits below the cross section of DM in the pseudo-data. Running the full analysis with the generator model fit to pseudo-data realizations that include a $\langle \sigma v \rangle = 3 \times 10^{-26}$ cm$^3$ s$^{-1}$ 100 GeV $b\bar{b}$ WIMP results in an average 95% confidence upper limit of $\langle \sigma v \rangle = 7.2 \times 10^{-26}$ cm$^3$ s$^{-1}$. See Figure 6.18. Fits done with realistic models are even more conservative. In other words, the fit procedure does not appear to set upper limits that violate the input signal.

Upper limits can also be set for PE that include no DM signal. Even with a perfect model, statistical fluctuations will always allow for some DM in the fit. Figure
6.3.4 Model Population

To be confident with a limit or detection, it is necessary to have sufficiently populated the model space such that the maximum TS($\theta_{DM}$) is the true maximum, and that the nearby sampling is dense enough to smoothly detect the profile out to the upper limit. Populating by sampling random $\vec{\beta}$ is highly inefficient, and doubly so,
since most combinations will fit either the CR or gamma rays very poorly. Here the predictive power of classification trees comes in handy (See Appendix A). An initial grid scan of $\vec{\beta}$ with a threshold $\chi^2$ as the categorical variable provides enough training that the CT can accurately predict whether a given GALPROP model will adhere to CR data. Figure 6.20 shows how the CT refines randomly generated $\vec{\beta}$, accepting only a subset likely to fit CR data reasonably well.

The introduction of the CT effectively removes a dimension from the relevant model space. To apply the Profile Likelihood requires models that are densely spaced in $\chi^2/2 + \Delta ln(L)$. Now, rather than generating models with both highly random $\chi^2$ and L, we sample only a tight band of $\chi^2$. This greatly speeds the rate at which model generation converges on a global maximum. See Figure 6.21 for the CR-gamma fit scatter plot resulting from a large number of CT-picked GALPROP models. The global maximum lies on the green line labeled “TS max.” Contours mark the levels at which models can contribute to progressively stronger upper limits.

After determining the $\vec{\beta}$ that results in the best combined Likelihood, random sampling to generate model density nearby makes no sense. It would take far too long, as most models will not contribute to the Profile Likelihood. Instead, varying $\vec{\beta}$ around the maximum with a simple 0.5%-σ gaussian generates sufficient density with a few hundred models (see Figure 6.22). Note that this is specific to the 100 GeV spectrum. Using a different mass shuffles the model space, and the $\vec{\beta}$ of the global maximum, although the important regions share many models.

The pseudo-experiments in Section 6.3.3 show that the actual Profile Likeli-
Figure 6.20: The difference between randomly sampled non-linear parameters (blue), and those chosen by a CR-fit-trained classification tree (red). Note that it is not possible to see correlations this way, something that makes the CT invaluable.
Figure 6.21: Cosmic and gamma-ray fit space. Blue dots represent individual GAL-PROP models, and contours show equivalent combined fits (TS, see Eqs. 6.6 and 6.8). Models favor DM over the null hypothesis as they move up and to the right, with the best-fitting model defining “TS max.”
Figure 6.22: Cosmic and gamma-ray fit space, including a set of 0.5%-σ gaussian variant models (yellow) around the global maximum from the initial sampling.
Figure 6.23: Partially marginalized GALPROP models: change in TS vs. preferred DM normalization.

The likelihood procedure is fairly robust. Since the Profile itself appears well-populated, (Sec. 6.4) the largest concern to the accuracy of the results is that the Likelihoods are in a false maximum of $\beta$. In other words, are there unexplored regions of parameter space that could lead to a very different result? While this possibility is difficult to rule out entirely, we can gain confidence that further gaussian variations on any of our existing GALPROP realizations would not significantly change the outcome.

To see this quickly, first see Figure 6.23, which depicts the TS of the GALPROP models as a function of their favored DM normalization. No 0.5%-$\sigma$ variation models are included in the plot, as they tend to deepen the -TS minimum wherever they are centered. Models that prefer higher $\theta_{DM}$ appear to have have worse gamma-ray fits.
Figure 6.24: Spread of TS for models resulting from a 5%-σ variation of GALPROP parameters. Models to the right of the dashed line are improved fits.

in general. Even by \( \theta_{DM} = 2 \), the log Likelihood has already risen above the current minimum by 60, a value that contributes directly to the TS. Obviously they do not contribute to the Profile Likelihood at the 95% confidence level, but the question at hand is whether a gaussian variation of any of these could fit well enough to provide a new minimum. Figure 6.24 provides an example of the TS range generated by a 5% gaussian variation on a single model. Most variant models actually give a worse fit to both gamma and cosmic rays. The TS improves by about 11 at most, and only about 10% of the models show any improvement at all. Although small variations on other models show slight differences, there are no drastic changes. It would seem the TS gap of 60 at \( \theta_{DM} = 2 \) is insurmountable to any 5% variation on a current model.
It is of course possible to vary at a higher percent, or to daisy-chain small variations, i.e. to vary the best variation and so on in an attempt reach a new minimum. Such methods are computationally expensive, however. Even if a daisy chain worked perfectly, requiring 10 models on average to find each improvement, it would take hundreds of GALPROP runs to cross 60 TS. And it is entirely likely that the chain would end before that, especially if the CR fit begins to degrade. Exploring every corner of the parameter space would be ideal, but at some point we must trust that the CT trained to predict CR fits avoided that region when searching for a global maximum for good reason.

### 6.4 Results

Applying the techniques of Section 6.3.2.2 to the region of dense sampling in Fig. 6.22 results in the Likelihood Profiles shown in Fig. 6.25. The TS is maximized for a 100 GeV WIMP with $\langle \sigma v \rangle = 9.3 \times 10^{-27}$ cm$^3$ s$^{-1}$, at 0.48$\sigma$ significance. Upper limits are set at the 95% confidence level and bound $\langle \sigma v \rangle$ to be less than $5.5 \times 10^{-26}$ cm$^3$ s$^{-1}$. The model density is high and the Likelihood Profile curvature appears convergent.

Residuals from the best-fitting model including 100 GeV DM are shown in Figure 6.26. There is very little visible difference from the best fit model without DM – the fit only adds about 500 DM photons, after all. The reduced $\chi^2$ actually suffers due to the marginal fit improvement coupled with an added free parameter. Other masses allow for progressively more signal. See Figures 6.27-6.29. Fits with the
Figure 6.25: Profile Likelihood for 100 GeV WIMP. Individual green curves represent partially maximized Likelihoods, $\tilde{L}(\beta, \theta_{DM})$, and the solid black line shows the Profile Likelihood. Vertical dashed lines mark the absolute minimum and the 95% confidence upper limit.
highest significance allow for several thousand DM photons, though in the bins where they contribute significantly to the model (>10%) and Likelihood, DM accounts for approximately 20 photons per bin.

High mass DM (>200 GeV) is more strongly favored, though no detection exceeds 3σ. Residuals are visibly reduced in the regions where these signals contribute, centered on roughly 30 GeV and annulus zero. For the best-fitting models not including DM, this area has a 10-15% residual (under-prediction of the model) that is mitigated by the addition of a hard photon spectrum near the galactic center. No mass is strongly favored; in fact everything >200 GeV has approximately the same (2.5σ) significance. Considering this lack of discriminatory power coupled with already weak detections, we can only set conservative limits on the WIMP cross section. See Figure 6.30 for a plot of significance and 95% confidence level upper limits as a function of WIMP mass. Figure 6.31 shows a comparison of said limits with those derived from the combined dwarf galaxy analysis mentioned in Chapter 3.
Figure 6.26: Results of 100 GeV DM fit. Top plot shows spatial and spectral residuals, middle the contribution of DM in fraction of total counts, and the bottom two plots show the integrated spatial and spectral component breakdowns.
Figure 6.27: As Figure 6.26, but for 200 GeV DM.
Figure 6.28: As Figure 6.26, but for 300 GeV DM.
Figure 6.29: As Figure 6.26, but for 500 GeV DM.
Figure 6.30: Significance and 95\% confidence upper limits for $\langle \sigma v \rangle$ of WIMPs with mass ranging from 100 to 600 GeV. The dashed line denotes the thermal relic cross section.
Figure 6.31: As Figure 6.30, but including limits derived from dwarf galaxy observations (Abdo et al. 2010c).
6.5 Discussion

The study described in this chapter sought to measure a possible generic WIMP signature originating from the diffuse Host Halo that houses the MW. Because such a signal competes with a very uncertain background – the galactic CR diffuse emission – we took care to marginalize as many potentially important parameters of the background model as possible. Although they fit well to a cross-check of local CR data, our naïve models resulted in poor fits to LAT-measured gamma rays. It is unclear to what extent this is due to inadequacy of the model itself or to a 10%-level energy dependent effective area systematic that may exist in the LAT data. Either way, the model required some correction before any search could be performed.

To improve the background models and account for any energy systematic, we split the sky into halves and used the residuals from the low-signal outer galaxy to correct the inner. Such a correction introduced its own systematic error, which we accounted for by adding a floating polynomial function to the fit in the inner galaxy. The outer galaxy correction cleaned up the energy residuals and was shown through pseudo-experiments to function with acceptably low bias.

Marginalizing all the free parameters proved to be a computationally challenging task, owing mainly to the complexity of GALPROP. A classification tree aided in this by learning and predicting which combinations of parameters would result in poor fits to CR data. Small gaussian variations of parameters near the best-fitting GALPROP model gave adequate population of the model space to construct a Profile Likelihood,
or marginalized Test Statistic. The TS then gave information on the significance and upper limits to DM signal seen by the LAT.

The results do not favor a DM detection. The upper limits derived are generally an order of magnitude above the thermal relic cross section (see Ch. 4). Although the LAT data has the statistical power to be much more constraining, we are limited at this point by the quality of the background model. Correcting it using the outer galaxy appears to help, with little introduced bias, but is no substitute for an improved model and for further calibration of the LAT effective area. Such a calibration could be performed using the bright and almost featureless gamma-ray emission that comes from CR interactions in the Earth’s limb. Better knowledge of the instrument effective area would remove the degeneracy in this study between the deficiencies of the instrument and background models, not to mention increase sensitivity to a signal expected to be a perturbation of the same or lesser order.

Regardless of the availability of any further knowledge of the LAT effective area, the work done here will be continued in an upcoming Fermi Collaboration publication. That study will expand the analysis to search for a broader range of WIMP masses, and channels, along with variations of the DM distribution.
Chapter 7

Conclusion

The identification of DM stands out as one of the most intriguing and difficult physics problems of our time. Able to shape the structure of the entire universe, DM is by definition difficult to detect directly. The launch of the Fermi-LAT satellite presented a great opportunity to make progress understanding the nature of DM. The detection of gamma rays resulting from the annihilation of the broad class of WIMP DM candidates would be an enormous stride towards this goal. Assuming the signal is not overwhelming (or it would have been obvious in previous experiments), great care must be taken in making the predictions used to test any DM hypothesis. Such tests rely on predictions of:

- The DM distribution.
- The Background.
- The Instrument Response.
All three predictions carry uncertainties that must be taken into account before claiming any discovery. The analysis in Chapter 5 focuses on the uncertainty in the distribution of DM to make predictions for the LAT detectability of MW substructure. These turn out to be greatly dependent on the relative location of such substructure to Earth. The LAT could see none over its lifetime, and it would not strongly constrain the nature of DM because the Earth might simply reside at a poor vantage point. One optimistic result is that, should we be in a position to detect a DM subhalo, it will likely appear as an extended source, which will aid in its identification.

There is less uncertainty in the DM distribution when considering the Host Halo – especially when omitting the inner region (< kpc). This advantage motivates the second analysis of this dissertation, presented in Chapter 6. Here the greatest challenges come from imprecise knowledge of the CR-induced background and the modeling of the LAT effective area. A combination of marginalizing the parameters controlling the background and an on-off source correction to account for the effective area made a reliable test possible. The test resulted in no DM detection for WIMPs with a range of masses annihilating into $b\bar{b}$ quarks, and set upper limits on their velocity-averaged cross section.

The power of the Fermi-LAT for both of these analyses is far from exhausted. Not only does the accumulation of further data make for better statistics, but the understanding of the instrument itself is continually improving. The upcoming “Pass 8,” in particular, a new set of reconstructed data and instrument response functions made by the Fermi Collaboration will make enough improvements to warrant a new
look at both studies. In addition, a new effective area calibration using an on-orbit source such as the Earth’s limb could greatly improve and clarify the analysis of the Host Halo. Even without improvements to the instrumental uncertainties, a future Fermi Collaboration study will expand on the Host analysis, adding searches for various WIMP masses, annihilation channels, and distributions.
Bibliography


N. F. Bell, A. J. Galea, and K. Petraki. Lifetime constraints for late dark matter decay. 


G. Jungman, M. Kamionkowski, and K. Griest. Supersymmetric dark matter.  

M. Kaplinghat, D. J. Phalen, and K. M. Zurek. Pulsars as a source of the WMAP haze.  


A. Klypin, *et al.* Where Are the Missing Galactic Satellites?  


M. Kuhlen, J. Diemand, and P. Madau. The Dark Matter Annihilation Signal from Galactic Substructure: Predictions for GLAST.  


M. Papucci and A. Strumia. Robust implications on dark matter from the first FERMI sky $\gamma$ map. *JCAP*, **3**:14, 2010. doi:10.1088/1475-7516/2010/03/014.


D. J. Thompson, *et al.* Calibration of the Energetic Gamma-Ray Experiment Telescope


Appendix A

Classification Trees

Classification Trees (CT) are a modern computational answer to problems with high dimensionality and non-homogeneity (Breiman et al. 1984). That is, a situation where a prediction depends on many variables with complex relationships that change depending on the position in phase space. The role is similarly filled by neural networks.

Ideally a CT works very simply. It consists of a set of variable-specific conditions that, when applied sequentially, sort data into subsets which are relatively pure in some characteristic. The tree can then predict that characteristic based on other properties of the data. The hard part of course is figuring out the conditions and their sequence.

Rejecting CR contamination from the LAT data serves as a concrete example. Given a set of events, we seek to answer a simple question: which of them are caused by CR? Some individual variables, such as those characterizing the CAL shower profile, are quite powerful and can produce fairly pure samples with simple universal cuts.
To achieve a $10^6 : 1$ rejection ratio, however, we must employ a host of less intuitive variables whose ensemble proves greater than the sum of its parts.

CT must be trained before they can function. This requires a data set where the answer (proton IDs in this case) is known, such as a Monte-Carlo or beam test. Handing the event sample to a computer, it of course proceeds blindly at first. It might, for example, start with the variable “TkrBlankHits,” a record of the number of hit clusters in silicon layers not adjacent to a converter. Cutting at an arbitrary value divides the sample into two branches. The goal is to get all the protons by themselves, and so the quality of the cut we just made can be evaluated on how much separation it generated. There are several ways to quantify this, but we will focus on the entropy:

$$ S = \sum_j -p_j \log_2(p_j) $$  \hspace{1cm} (A.1)

$p_j$ stands for the fraction of total ensemble in state $j$. For a question like ours with a binary answer, the sum is easy to write out.

$$ S = -p_p \log_2(p_p) - p_{np} \log_2(p_{np}) $$ \hspace{1cm} (A.2)

Here $p_p$ is the proton fraction of the branch and $p_{np}$ is the fraction from everything else. A perfect cut would set $p_p$ to one and the entropy, $S$, to zero. A useless cut would leave have $p_p = p_{np}$ and set $S = 1$. Each node in the tree is therefore evaluated on the basis of minimum entropy, or the purity of its branches. From there, “growing” a tree is simple. First, split the sample with the cut (out of all variables) that produces...
the least entropy. Then, for each successive level of the tree, repeat the procedure, scanning over all the remaining variables. The splitting stops when it becomes impossible to significantly increase the sample purity. Lastly, assign each terminal node, or “leaf,” a category (proton or not in this case) based on the ratio of the remaining events it holds.

A fully grown tree will funnel any new data down its branches and predict the category based on the final leaf. To be of scientific value, it needs an error, or misclassification estimate. An impure leaf provides some estimate of the error by itself. To go beyond this, it is helpful to split up the initial training data set into an ensemble. A tree is grown for each subset and then tested with the rest of the data. Since all the answers are known, this provides information on the misclassification rates.

Before we are done, another fact becomes important. Trees for a given problem have an optimal number of branches; beyond that, their misclassification rate increases. It turns out that overgrowing a tree and then pruning it is the best way to reach the optimal size (Breiman et al. 1984). One can use the ensemble to test the average accuracy of different sizes of trees. Once the best size is known, each overgrown tree should be pruned (collapsing branches) to the proper level. Rather than pruning evenly, assigning a each cut a cost proportional to the entropy allows the tree to retain its most effective branches. This makes a pruned tree better than one that is simply grown directly to the optimal size.

There are several ways to utilize the ensemble for prediction. It can be combined into an average tree or used as seeds for another round of growth. Perhaps most
simply, running each unknown event through all the trees in the ensemble and then statistically combining the results works well enough for the purposes within this study.
Appendix B

SuperSymmetry

B.1 Supersymmetry

Largely constructed to solve the hierarchy problem of the SM (CHAPTER REF), Supersymmetry (SUSY) has been at the forefront of particle physics since the introduction of the Minimal Supersymmetric Standard Model (MSSM) in 1981. At its heart, it postulates that for every elementary SM particle, there exists a “superpartner” that differs by half a spin unit. Half a spin means the difference between boson and fermion, and so every superpartner is of the opposite type. These contribute to the loop corrections of the hierarchy problem in opposition to their partners, canceling them out. If the symmetry were perfect, each superpartner would mirror its SM counterpart in both mass and quantum numbers. We know this is not the case, however, since no one has ever observed a superpartner. This implies that they exist at a mass scale generally unreachable in our present era.
There is considerable freedom in constructing a SUSY Lagrangian that satisfies its basic motivations (Chung et al. 2005), manifesting in over a hundred free parameters. These are often condensed or constrained using further arguments into more manageable versions, such as minimal supergravity (mSUGRA, also known as the constrained MSSM), which distills the theory freedom into five free parameters. Despite its additional freedom, the MSSM adds as few fields as possible, resulting in merely a doubling (plus one extra Higgs doublet) of the known elementary particles. See Figure B.1 for an illustration.

B.2 Neutralino

The famed neutralino arises from a mixing of the photon, Z, and neutral Higgs superpartners. The mixing results in four mass eigenstate neutral particles. These should be heavy and neutral, which makes them a WIMP. In order to be a solid DM candidate, however, they also need to be stable. Fortunately, the whole theory works
better\textsuperscript{1} when subjected to a new symmetry, denoted

\[ R = (-1)^{3(B-L)+2S}. \] (B.1)

Here \( B \) is the baryon number, \( L \) is the lepton, and \( S \) represents spin. For all SM particles, \( R=1 \), and so the 2S addition means every superpartner has \( R=-1 \). If the symmetry is strictly conserved, superpartners can only decay into superpartners and there must be a Lightest Supersymmetric Particle, or LSP. If this were one of the neutralino eigenstates, the theory delivers a perfect DM candidate.

\textsuperscript{1}Without R-parity, protons in the MSSM are unstable, decaying with a lifetime strongly ruled out by current data (Jungman \textit{et al.} 1996).