

SEARCHES FOR BEYOND THE STANDARD MODEL PHENOMENA WITH THE  
HAWC DETECTOR

By

Joseph Andrew Lundeen II

A DISSERTATION

Submitted to  
Michigan State University  
in partial fulfillment of the requirements  
for the degree of

Physics – Doctor of Philosophy

2021

## ABSTRACT

### SEARCHES FOR BEYOND THE STANDARD MODEL PHENOMENA WITH THE HAWC DETECTOR

By

Joseph Andrew Lundeen II

The universe is known to be dominated by a class of matter known as dark matter. No known particles possess the necessary characteristics to make up this dark sector. Possible extensions to the Standard Model produce promising candidates, in particular Weakly Interacting Massive Particles (WIMPs). Under this hypothesis, it may be possible to probe dark matter with high-energy gamma-rays. The High Altitude Water Cherenkov (HAWC) detector is a contemporary experiment sensitive to multi-TeV cosmic gamma rays and able to observe approximately two thirds of the sky in any given day. With these properties, HAWC is able to search for high-mass classes of WIMPs from the dark matter halo surrounding the Milky Way galaxy and set strong constraints on dark matter annihilation and decay into gamma rays.

## ACKNOWLEDGMENTS

Thank you to my advisor, Dr. Kirsten Tollefson, for being my mentor throughout this process. And another thank you to Dr. Pat Harding for being my honorary other advisor. Both of your guidance has been invaluable in navigating the physics world.

I would also like to thank the entire HAWC collaboration. Thanks especially to Dr. Tolga Yapici and Dr. Sam Marinelli for their patience with me as a new and confused grad student. Also thanks to Dr. Andrea Albert, Dr. Chad Brisbois, Dr. Kelly Malone, and Dr. Henrike Fleischhack, for tolerating me as a not-so-new but still somewhat confused grad student. Speaking of graduate students, thank you to Alison Peisker for being the best cubicle-mate of all time, and for all the shared laughs even after she'd moved on to a bigger and better office!

Thank you to my parents, Ed and Christine Lundeen. And a very special thank you to Candyce Hill. Your ongoing support has meant more to me than I can even begin to describe (well, I suppose I could but this dissertation doesn't need to be any longer). I am so glad to have you here at this moment, and for all time.

# TABLE OF CONTENTS

LIST OF TABLES . . . . .	vii
LIST OF FIGURES . . . . .	ix
CHAPTER 1 INTRODUCTION . . . . .	1
CHAPTER 2 WIMP DARK MATTER AND GAMMA RAYS . . . . .	3
2.1 Introduction . . . . .	3
2.2 Evidence for Dark Matter . . . . .	3
2.2.1 Rotation Curves . . . . .	3
2.2.2 Gravitational Lensing . . . . .	4
2.2.3 Cosmic Microwave Background . . . . .	5
2.3 The Standard Model . . . . .	6
2.4 Lambda Cold Dark Matter and WIMPs . . . . .	8
2.5 Detection Approaches . . . . .	9
2.6 WIMPs and Gamma Rays . . . . .	12
2.7 Potential Search Targets . . . . .	15
2.8 Motivation For High-Mass WIMPs and Large Targets . . . . .	16
CHAPTER 3 AIR SHOWERS AND THE HAWC DETECTOR . . . . .	19
3.1 Cosmic Gamma Rays and Atmospheric Air Showers . . . . .	19
3.2 HAWC Description . . . . .	22
3.3 Hardware and Construction . . . . .	23
3.4 Data Acquisition System . . . . .	24
3.5 Calibration . . . . .	27
3.6 The Outriggers . . . . .	27
CHAPTER 4 EVENT RECONSTRUCTION AND ANALYSIS . . . . .	30
4.1 Primary Reconstruction . . . . .	30
4.1.1 Core Fitting . . . . .	30
4.1.2 Angle Fitting . . . . .	31
4.2 Event Characterization . . . . .	33
4.2.1 $f_{\text{hit}}$ Binning . . . . .	33
4.2.2 Spatial Binning and Angular Resolution . . . . .	34
4.2.3 Energy Binning . . . . .	35
4.3 Background Rejection . . . . .	36
4.3.1 Gamma-Hadron Cuts . . . . .	36
4.3.2 Direct Integration . . . . .	39
4.3.3 Background with the Energy Estimators . . . . .	39
4.4 The Neural Network Energy Estimator . . . . .	40
4.5 Neural Network Re-optimization . . . . .	42
4.6 Forward Folding and Likelihood Analysis . . . . .	44

CHAPTER 5	LIV: AN APPLICATION OF THE ENERGY ESTIMATORS . . . . .	47
5.1	Introduction . . . . .	47
5.2	Lorentz Invariance Violation . . . . .	47
5.3	Derivation of Upper Limits . . . . .	48
5.4	Searching for LIV . . . . .	50
5.4.1	High Energy Sources . . . . .	50
5.4.2	Limit Technique . . . . .	51
5.4.3	Fitting Details . . . . .	54
5.5	Results . . . . .	56
5.5.1	Individual Results . . . . .	58
5.5.2	Systematics . . . . .	59
5.5.3	LIV Limits . . . . .	62
CHAPTER 6	SEARCHES FOR DARK MATTER SUB-STRUCTURE . . . . .	64
6.1	Introduction . . . . .	64
6.2	Missing Satellites . . . . .	64
6.3	Searching for Dark Dwarfs in HAWC Data . . . . .	65
6.3.1	Search Method . . . . .	65
6.3.2	Fit results . . . . .	67
6.4	Characteristic Upper Limits Across The Sky . . . . .	70
6.4.1	Characteristic Limits . . . . .	71
6.4.2	Statistical Variations and Expected Limits . . . . .	74
6.5	HAWC Sensitivity to Modeled Dark Dwarfs . . . . .	76
6.5.1	Modeling Galactic Substructure . . . . .	76
6.5.2	Detection Thresholds . . . . .	80
6.5.3	J-factor Sensitivity . . . . .	83
6.6	Summary . . . . .	83
CHAPTER 7	DARK MATTER IN THE GALACTIC MAIN HALO . . . . .	84
7.1	Introduction . . . . .	84
7.2	Spatial Profiles . . . . .	84
7.3	Large-Scale Background Estimation . . . . .	87
7.3.1	Background Anisotropy and Limits of Direct Integration . . . . .	87
7.3.2	The $\alpha$ -factor . . . . .	88
7.4	Region of Interest (ROI) Selection and Source Modeling . . . . .	91
7.4.1	Selecting an Optimal ROI . . . . .	91
7.4.2	Extended Source Convolution . . . . .	93
7.4.3	Electroweak Corrections . . . . .	95
7.5	Signal Contamination of Background . . . . .	96
7.5.1	Effective Flux . . . . .	96
7.5.2	Estimation of Contamination Factors . . . . .	98
7.6	Systematics . . . . .	99
7.6.1	Detector Systematics . . . . .	99
7.6.2	Pointing Systematic . . . . .	101
7.6.3	Spectral Systematics . . . . .	101

7.7	Results . . . . .	104
7.7.1	Fit Results . . . . .	104
7.7.2	Limits . . . . .	104
7.7.3	Statistical Effects . . . . .	109
7.8	Conclusion . . . . .	111
CHAPTER 8 SUMMARY . . . . .		116
APPENDIX . . . . .		118
BIBLIOGRAPHY . . . . .		125

## LIST OF TABLES

Table 4.1:	Bin definitions for $f_{\text{hit}}$ binning scheme. . . . .	33
Table 4.2:	Neural net energy binning labels and edges. Note that this is part of a 2D binning scheme combined with the $f_{\text{hit}}$ bins shown in Table. 4.1. . . . .	36
Table 4.3:	Variables and definitions used in the NN energy estimator. . . . .	41
Table 5.1:	Fitted locations of HAWC sources emitting above 56 TeV and used to constrain LIV. . . . .	51
Table 5.2:	HAWC sources with p-values for LIV detection and lower limits on maximum photon energy in TeV. . . . .	58
Table 5.3:	Color-coded table of the 68% and 95% statistical bands on the 95% CL limits on $E_c$ . All the lower limits fall well within the 95% band, with most falling within the 68% band. . . . .	59
Table 5.4:	Effect of systematic uncertainties on the 95% CL lower limits on $E_c$ . All values are reported in TeV. “BP” refers to the inclusion of a broad pulse in the detector simulation, “Qunc” refers to the size of the charge uncertainty, “Thresh” refers to the PMT triggering threshold and the various run indices are different estimates of PMT efficiency (see Ref. [33] for details). The “Swapped” and “Shifted” labels refer respectively to changing to the less-favored spectral form and moving the source center to that fitted using all bins rather than just the highest energy bins. Finally, “Energy Scale” refers to a 10% possible effect of the HAWC absolute energy scale compared to the HESS detector. The data in this table was published in [46]. . . . .	61
Table 5.5:	Overall effects of systematic errors on $E_c$ for the combined limits. . . . .	62
Table 5.6:	Limits on LIV parameters derived from lower limits on the maximum observed photon energy, $E_c$ . Note that limits on parameters derived from photon-splitting ( $3\gamma$ ) can only be reported for individual sources, since they depend on the propagation distance. . . . .	63

Table 6.1:	Location and spatial data for potential dark dwarfs considered in this analysis. Radius refers to the angular size of the disk source hypothesis (zero corresponding to a point source). All four of these sources were seen in the 2HWC catalog with maximum of $TS > 25$ [34] and no additional resolved sources were found in the dataset used in this analysis. Only one of these sources (2HWCJ 1040+308) still remains at $TS > 25$ in the 760 day dataset. . . . .	66
Table 6.2:	Summary of test statistics (TS) and $\Delta$ TS for the three flux models considered: power law (PL), cut-off power law (CU), and the best-fit dark matter spectrum (DM). See Fig. 6.1 for details on best-fit spectra for each source. . . . .	70
Table 6.3:	Parameters used in the assumed dark matter density profiles (Eq. 6.3 and Eq. 6.4). $R_{\odot}$ and $\rho_{\odot}$ are respectively the distance from the sun to the Galactic center and the local dark matter density of the solar system. The scale radius $r_s$ is the radius such that $\rho(r_s) = \rho_s$ and $\alpha$ determines the slope of the Einasto profile (the Burkert profile does not set the logarithmic slope $\alpha$ as a free parameter). . . . .	77
Table 6.4:	Parameters used in the CLUMPY simulation of the Galactic dark matter substructure mass function and spatial distribution. $M_{min}$ and $M_{max}$ are respectively the minimum and maximum subhalo masses considered (in units of the solar mass $M_{\odot}$ ) and $n$ is the slope of the power law assumed for the mass probability function. The spatial distribution is modeled by Eq. 6.3 with the $r_s$ and $\alpha$ values reported here and $\rho_s$ chosen to normalize to unity. The values reported here are taken from Ref. [70] and Ref. [71]. . . . .	80
Table 7.1:	Summary of systematic uncertainty sizes on the 95% CL limits for annihilation and decay. The dominating systematic in both cases is the choice of density profile, followed by systematics arising from detector simulation. Note that these are only estimates, and the actual contribution of each varies from spectrum to spectrum. No overall uncertainty is reported since neither choice of density profile is the used as the nominal case. . . . .	111



## LIST OF FIGURES

Figure 2.1:	Rotation curve showing star velocities as a function of distance from the center, $R$ , for the galaxy M33 compared with the expectation from the observed luminous component. The expected curve falls off as $1/\sqrt{R}$ , while the actual curve flattens out. This implies the presence of a “dark” component that interacts only gravitationally. Figure from Ref. [5]. See also Ref. [4] for details on how these measurements were obtained. . . . .	4
Figure 2.2:	A false color image showing the baryonic gaseous component of the Bullet Cluster overlaid with contours of the gravitational lensing probe. The two clusters pictured are undergoing a merger event where the gas in the smaller bullet-shaped object collided with and passed through the larger gas cloud at a reduced speed compared to the collisionless cluster components. This shows a clear displacement between the X-ray emitting gaseous component of the two clusters and the centers of mass, indicating most of the mass is contained within collisionless halos that dominate the cluster masses. Figure from Ref. [6]. . . . .	5
Figure 2.3:	Table of Standard Model (SM) elementary particles showing their organization into fermions and bosons, as well the subdivisions between them. Figure taken from <a href="https://en.wikipedia.org/wiki/Standard_Model">https://en.wikipedia.org/wiki/Standard_Model</a> . . . . .	7
Figure 2.4:	Simulated plot of thermal freeze-out in the $\Lambda$ CDM framework for various assumed velocity-weighted cross sections showing the dark matter number density approaching a constant in the present day. Estimates of the current dark matter density (see Eq. 2.1) allow for a prediction of the velocity-weighted cross section. Figure from Ref. [15]. . . . .	10
Figure 2.5:	Feynman diagram showing the interaction behind the various dark matter detection approaches. “DM” refers to dark matter particles, while “SM” refers to Standard Model particles. The large blue circle is a stand-in for the theorized physical interactions between dark matter and the Standard Model and the arrows along each axis show the direction one should follow to trace the interaction probed by each detection class. Following a given arrow from start to finish shows the initial particles of an interaction scattering and resulting in the final product particles of that interaction. . . . .	11

Figure 2.6:	Sample map showing one simulation of the smooth halo and substructure of the Milky Way halo. The color scale shows the logarithm of the J-factor (Eq. 2.4) integrated over a single pixel width of solid angle (about 0.25 square degrees in this map). The high J-factor at the Galactic Center is clearly visible with substructure distributed throughout the sky. The details for generating this simulation are discussed further in Chapters 6 and 7. . . . .	13
Figure 2.7:	Example dark matter energy spectra for two annihilation channels ( $b\bar{b}$ and $\tau^+\tau^-$ ), assuming a dark matter mass of 100 TeV and showing the characteristic hard cutoff feature. These two channels are typically taken as representative of dark matter gamma-ray spectra, with $b\bar{b}$ falling off most quickly with energy and $\tau^+\tau^-$ falling off slowest with energy. Most other dark matter spectra fall between these two cases. . . .	14
Figure 2.8:	95% confidence upper limits on dark matter annihilation in $b\bar{b}$ (left) and $\tau^+\tau^-$ (right) channels from the FERMI-LAT observation of the dwarf spheroidal galaxies. Below assumed dark matter masses of 100 GeV, these constraints are lower than the predicted thermal cross section of the $\Lambda$ CDM model (Eq. 2.1), therefore completely excluding WIMP dark matter with these masses. Figure taken from Ref. [26]. . . . .	16
Figure 2.9:	95% confidence upper limits on dark matter annihilation in the $W^+W^-$ (left) and $\tau^+\tau^-$ (right) channels from the HESS observation of the Galactic Center. The model $W^+W^-$ channel in the HESS analysis behaves analogously to the $b\bar{b}$ channel and plotted here as a representative soft spectrum. These results are able to exclude WIMPs of certain multi-TeV masses assuming annihilation into harder channels like $\tau^+\tau^-$ , but depend on assumptions about the exact behavior of the density profile. Figure taken from Ref. [28]. . . . .	18
Figure 3.1:	Illustration of a gamma-ray induced extensive air shower. The shower propagates itself via successive pair production and bremsstrahlung. The shower footprint widens over time due to the finite opening angles of successive interactions. . . . .	20
Figure 3.2:	Illustration of a hadronic air shower. Unlike the gamma-ray shower shown in Fig. 3.1, this shower contains pions, neutrinos, and other non-electromagnetic components. . . . .	21

Figure 3.3:	Illustration of a Cherenkov light cone with opening angle $\theta$ produced in a material. Here, $c$ is the speed of light in vacuum, $n$ in the index of refraction of the material, $\beta c$ is the speed of the radiating particle, and $t$ is the amount of time the particle has been propagating in the medium. Created by Arpad Horvath and used under a Creative Commons BY-SA 2.5 license. . . . .	22
Figure 3.4:	A photograph of the HAWC detector showing the main array and the mountain Pico de Orizaba in the background. . . . .	23
Figure 3.5:	Illustration of a shower secondary producing Cherenkov light in a HAWC WCD (left), and a map of the HAWC layout (right). Image taken from Ref. [33]. . . . .	25
Figure 3.6:	Illustration of the conversion of PMT voltage over time to an FEB output signal, showing both possible voltage thresholds. The time over threshold (ToT) is used later in the reconstruction process. Figure from [35]. . . . .	26
Figure 3.7:	Map of the recently completed array of smaller outrigger tanks (represented by black dots) superimposed on a map of the main array. Figure from Ref. [38]. . . . .	28
Figure 3.8:	Schematic of a single outrigger tank showing the dimensions. The outriggers are smaller than the main tanks and contain only a single PMT, allowing them to cover more area at a lower cost than extending the main array. Photo credit to Michael Schneider. . . . .	29
Figure 4.1:	Illustration of an air shower arriving at HAWC with the arrival angle denoted by $\theta$ and core location marked with a red cross. This figure illustrates how the edge of the shower is curved rather than flat due to secondary interactions, and how particle density decreases further from the central shower axis as predicted by the NKG function (Eq.3.1). Figure from Ref. [36]. . . . .	31
Figure 4.2:	Plot of a HAWC event showing the lateral charge distribution across the array. The core reconstructed using the functional form in Eq.4.1 is marked with a red star. The details in the title refer to the reconstructed arrival coordinates and the HAWC DAQ run from which the event is displayed, as well as variables used to separate gamma-like and hadron-like events (see Sec. 4.3.1). Figure from Ref. [33]. . . . .	32

Figure 4.3:	Scaled histogram of the fitted true energy distribution of each $f_{\text{hit}}$ bin as shown in Table 4.1. Figure is taken from Fig. 3 of Ref. [33] and was generated from Monte Carlo simulations assuming a power law spectrum $\frac{d\Phi}{dE} \sim E^{-2.66}$ for a single transit of a source at a declination of 20 degrees. All fitted distributions have been normalized to have unity amplitude for visualization purposes. . . . .	34
Figure 4.4:	Comparison of a likely cosmic-ray (left) and likely gamma-ray (right) lateral charge distributions with the SFCF (Eq. 4.1) functional form fitted. The gamma-ray shower has one distinct core and falls off fairly symmetrically about its central axis. In contrast, the cosmic-ray shower contains multiple smaller peaks of charge far from the main core. The presence of charge peaks far from the shower core is quantified by compactness in Eq. 4.2, while the axial symmetry can be seen in the PINCNess (Eq. 4.3) moving average, with large deviations in the cosmic-ray shower indicating an asymmetric charge distribution. Figure from Ref. [33].	38
Figure 4.5:	Bias and resolution of the neural net (NN) and ground parameter (GP) energy estimators evaluated on current Monte Carlo data. Figure taken from Ref. [41]. . . . .	43
Figure 4.6:	Bias and resolution of the newly re-optimized neural network (NNV3) trained on the most recent HAWC Monte Carlo, compared with the performance of the old neural network (NNV2) trained on older Monte Carlo. The data is plotted for events on (left) and off (right) the array, and with (top) and without (bottom) a zenith cut of 45 degrees. In all cases, the re-optimized (NNV3) neural network trained on a flat energy spectrum performs the best at all energy scales above 1 TeV. . . . .	45
Figure 5.1:	Significance maps in the vicinity of the four high-energy sources for bins above 56 TeV. These maps were made using the likelihood-derived significance following a fitted power law of index 2.0. . . . .	52
Figure 5.2:	Effect of softening the hard cutoff by both the HAWC energy resolution and the additional smoothing required to avoid a discontinuous likelihood function. For illustration, the cutoffs are superimposed on a power law spectrum with an index of -2.7, and arbitrary flux and energy units. Figure from [36]. . . . .	55
Figure 5.3:	Plots of best-fit spectra for each source compared with spectra expected were a hard cutoff detected at 100 TeV. The shaded regions show the statistical uncertainties on the fits. As expected, the presence of a hard cutoff in the HAWC energy range would have had statistically significant effect on the spectra were it detected, even with the additional softening added. . . . .	56

Figure 5.4:	95% CL Lower limits on $E_c$ as a function of chosen radius for the four high energy sources considered in the analysis. The optimal radius (marked with a green dot) is the smallest one such that all the highest energy photons are included and the lower limit is maximized. Each optimal radius is on the order of the angular extent of the source, where it is expected that all high-energy photons are incorporated into the fit with minimal background. . . . .	57
Figure 5.5:	Distribution of p-values derived from 1000 pseudo-experiments on the Crab Nebula. As expected, the one-sided nature of the hard cutoff creates a delta function at 1. . . . .	60
Figure 6.1:	Best-fit results in $f_{\text{hit}}$ bins for the four targeted sources showing the three fitted spectra and TS values. The counts do not reflect the full spatial-likelihood analysis of HAWC and are for visualization purposes. The best-fit TS numbers are from the full HAWC likelihood calculation. Each spectrum is scaled by the estimated uncertainty in data (as is the case in the lower-left plot of Fig. 6.2), showing that each best-fit hypothesis is comparable in goodness of fit. The same spatial hypothesis was assumed for each as the 2HWC catalog (see Table 6.1) [34] . . . . .	68
Figure 6.2:	Sample best-fit spectra plotted both in energy (top) and $f_{\text{hit}}$ bins (bottom) to one of the unassociated sources (2HWC J1309-054) considered in this analysis for its best-fit simple power law, cut-off power law, and dark matter spectrum. While the spectra appear distinct in energy-space, they are of comparable goodness of fit to the data in $f_{\text{hit}}$ bins. The top right plot also shows the statistical uncertainties of the best fits as correspondingly-colored shaded regions, indicating that the range of allowed fits overlaps substantially. We show the $f_{\text{hit}}$ bin fits both in terms of raw counts (left) and scaled by the estimated error in the data (right). The $f_{\text{hit}}$ bin plots (bottom) are calculated summing events over a fixed angular disk in the sky rather than the full spatial-likelihood calculation. These are for visualization purposes only. The shown best-fit spectra and TS (Eq. 4.10) values are those from the full spatial-likelihood calculation. . . . .	69
Figure 6.3:	Significance (as defined in the 2HWC catalog [34]) distribution of the grid points from all strips of declination used in the characteristic limit calculations (Observed) superimposed over the distribution expected from background-only (Expected). The likelihoods used for this plot were calculated assuming a 10 TeV, $b\bar{b}$ channel dark matter spectrum. The distribution has a mean of zero and standard deviation of one, consistent with the expectation from background-only, indicating no excess remains after resolved sources from the catalog are removed and that the sample is an unbiased representation of the remaining sky. . . . .	72

Figure 6.4:	Characteristic upper limits on $J\langle\sigma v\rangle$ for the $b\bar{b}$ and $\tau^+\tau^-$ channels at each declination assuming various values of the dark matter mass $M$ . The points plotted show the expected 95% confidence level upper limit for points at each declination. The most constraining limits are found at declinations directly overhead for HAWC. Individual dwarf upper limits will vary depending on statistical fluctuations at their locations. . . . .	73
Figure 6.5:	Characteristic upper limits on $\langle\sigma v\rangle$ for the $b\bar{b}$ and $\tau^+\tau^-$ channels assuming a dwarf galaxy with $J = 10^{19} \text{ GeV}^2 \text{ cm}^{-5} \text{ sr}$ were discovered at each declination. These are obtained from scaling the limits in Fig. 6.4 by this J-factor. For reference, the canonical thermal $\langle\sigma v\rangle$ value is $2.2 \times 10^{-26} \text{ cm}^{-3} \text{ s}^{-1}$ [60]. Individual upper limits will depend on statistical fluctuations at the location of a discovered dwarf and its measured J-factor. . . . .	74
Figure 6.6:	Statistical variation (68% and 95%) of the characteristic upper limits on $J\langle\sigma v\rangle$ as well as the median (expected) value, plotted as a function of declination. These values were calculated assuming a $b\bar{b}$ annihilation channel spectrum with a dark matter mass of 2 TeV. The error bar size is roughly independent of declinations and spectral assumptions, so only this sample is shown. . . . .	75
Figure 6.7:	A comparison of the density profile behavior as a function of distance from the Galactic Halo center. The Einasto (cuspy) profile differs by nearly three orders of magnitude from the Burkert (cored) profile towards the center. . . . .	78
Figure 6.8:	Sample map showing one simulation of the smooth halo and substructure of the Milky Way halo, flattened into a Mollweide projection. This map is in celestial coordinates and is truncated so as to correspond to the HAWC field-of-view. The color scale shows the logarithm of the J-factor (Eq. 2.4) integrated over a single pixel width of solid angle ( $0.082 \text{ deg}^2$ ). The high J-factor at the Galactic Center is clearly visible with substructure distributed throughout the sky. . . . .	81
Figure 6.9:	Estimated HAWC detection threshold $\langle\sigma v\rangle$ values for a sample of dark matter simulated substructure. Substructure models for a range of dark matter particle masses, annihilation channels and halo mass models are shown. The plotted points show the $\langle\sigma v\rangle$ such that 50% of possible realizations yielded a TS of 25 or greater for at least one dark dwarf. . . . .	82
Figure 7.1:	A comparison of the density profile behavior as a function of distance from the Galactic Halo center. The Einasto (cuspy) profile differs by nearly three orders of magnitude from the Burkert (cored) profile towards the center. . . . .	86

Figure 7.2:	Relative intensity map of the small-scale cosmic-ray anisotropy. These deviations from the isotropic background estimation can show up as excesses even in the gamma-ray maps, due to the fact that the data is dominated by hadrons. This effect must be accounted for in this analysis. Figure take from Ref. [74]. . . . .	88
Figure 7.3:	Significance map of the HAWC field of view assuming an extended disk source with a radius of 5 degrees and a power law with index -2.7. The Galactic Plane and known HAWC sources are visible as well as the cosmic-ray anisotropy regions A, B, and C defined in Fig. 7.2. The area around the Galactic Plane shows a negative bias of unknown origin, indicating a failure of the direct integration to properly estimate the background. . . . .	89
Figure 7.4:	Significance map of the HAWC field of view without gamma/hadron cuts assuming an extended disk source with a radius of 5 degrees and a power law with index -2.7. This map is dominated by hadrons due to the lack of cuts. It lacks the negative bias around the Galactic Plane seen in Fig. 7.3, although it contains larger positive and negative regions due to the large-scale cosmic-ray anisotropy discussed in Ref. [74]. . . . .	90
Figure 7.5:	Significance map of HAWC field of view assuming an extended disk source with a radius of 5 degrees and a power law with index -2.7 and the $\alpha$ -background technique for $f_{\text{hit}}$ bins 3-9. The Galactic Plane and known HAWC sources are visible as well as a new unidentified source of excess. This map lacks the strong negative bias of that shown in Fig. 7.3. . . . .	92
Figure 7.6:	Relative sensitivity assuming a 10 TeV $b\bar{b}$ annihilation spectrum. The top figure assumes an Einasto spatial profile while the bottom figure assumes a Burkert spatial profile. These figures demonstrate HAWC's sensitivity to dark matter emission from the Galactic halo. Points with more constraining upper limits will be included in the regions of interest for the full analysis and excluded from the background estimation. . . . .	94
Figure 7.7:	The sub-ROIs used to efficiently convolve the Galactic Halo spatial profile with the HAWC detector response using flat sky projections. Each ROI is a different color in an arbitrary scale. To avoid contamination from normal matter gamma-ray sources, the galactic plane and known gamma-ray emission mechanisms are masked out (illustrated by rendering in dark blue). . . . .	96

Figure 7.8:	Comparison of the expected dark matter energy spectra with and without electroweak corrections for both the $b\bar{b}$ and $W^+W^-$ channels, assuming a mass of 100 TeV and arbitrary flux units. The $W^+W^-$ channel gains a secondary peak from the electroweak correction due to the $W$ boson having a full unit charge. In contrast the $b\bar{b}$ channel, with only 1/3 unit charge, is only slightly effected. . . . .	97
Figure 7.9:	Expected contamination factors in each $f_{\text{hit}}$ bin averaged in declination bands assuming $\langle\sigma v\rangle$ values of $10^{-21}$ (top left), $10^{-22}$ (top right), $10^{-23}$ (bottom left), $10^{-2}$ (bottom right), all in units of $\text{cm}^2\text{s}^{-1}$ . The expected contamination behaves identically for the different $\langle\sigma v\rangle$ values except for one that is nearly twenty orders of magnitude outside the physically realistic parameter space (bottom right). Therefore, the contamination can be treated as independent of overall flux level. . . . .	100
Figure 7.10:	Expected contamination factors in each $f_{\text{hit}}$ bin assuming an Einasto spatial profile, a 100 TeV dark matter mass, and a $b\bar{b}$ spectrum as a function of $\langle\sigma v\rangle$ for the highest (left) and lowest (right) declinations considered. These plots show no dependence on overall flux level, measured by choice of $\langle\sigma v\rangle$ . . . . .	101
Figure 7.11:	Expected contamination factors in each $f_{\text{hit}}$ bin averaged in declination bands for spectra assuming 10 TeV (top left), 20 TeV (top right), 50 TeV (bottom left), and 100 TeV (bottom right) for the dark matter mass $M$ . These plots hold the dark matter annihilation channel ( $b\bar{b}$ ) and spatial profile (Burkert) constant across all cases in order to isolate the effect of different assumed dark matter masses. These plots show the contamination does depend at least somewhat on the dark matter mass. Note that the discontinuous behavior in Bin 9 for the 10 TeV spectrum arises due to the relatively low energy cutoff, leaving very few gamma-rays that will reconstruct in the higher $f_{\text{hit}}$ bins. . . . .	102
Figure 7.12:	Expected contamination factors in each $f_{\text{hit}}$ averaged in declinations bands for the $b\bar{b}$ (left) and $\tau^+\tau^-$ (right) channels. The dark matter mass (100 TeV) and spatial profile (Burkert) are held constant between these plots. These two channels are chosen because they are, respectively, the softest and hardest spectra thereby probing the extremes of the channel parameter space. These plots show the contamination depends fairly strongly on the channel. . . . .	103



- Figure 7.13: Contamination factors in each  $f_{\text{hit}}$  averaged in declination bands assuming a Burkert profile (left) and an Einasto profile (right), and holding the assumed mass at 100 TeV and annihilation channel to  $b\bar{b}$ . As expected, the assumed density profile has a strong effect on the contamination, with the strongly-cusped Einasto profile showing lower contamination than the cored Burkert profile. . . . . 103
- Figure 7.14: Significance ( $\sqrt{\text{TS}}$ ) as a function of dark matter mass and channel for annihilation assuming a Burkert profile (top left), annihilation assuming an Einasto profile (top right), decay assuming a Burkert profile (bottom left), and decay assuming an Einasto profile (bottom right). No spectrum shows significant evidence of gamma-ray emission, although a small positive significance is shown at the lower masses due to the residual unidentified excess in the lower bins. . . . . 105
- Figure 7.15: 95% CL upper limits on the velocity-weighted annihilation cross-section  $\langle\sigma v\rangle$  for the  $b\bar{b}$  (top left),  $\tau^+\tau^-$  (top right),  $t\bar{t}$  (bottom left), and  $W^+W^-$  (bottom right) dark matter spectra assuming the Einasto (red) and Burkert (blue) spatial profiles. The correspondingly colored shaded regions are the systematic uncertainty bands. The correspondingly colored shaded regions are the systematic uncertainty bands. The corresponding limits from the previous HAWC Galactic Halo analysis [24] are also plotted for comparison. . . . . 106
- Figure 7.16: Additional 95% CL upper limits on the velocity-weighted annihilation cross-section  $\langle\sigma v\rangle$  for the  $e^+e^-$  (top left),  $\mu^+\mu^-$  (top right), and  $Z^0Z^0$  (bottom) dark matter spectra assuming the Einasto (red) and Burkert (blue) spatial profiles. The correspondingly colored shaded regions are the systematic uncertainty bands. The corresponding limits from the previous HAWC Galactic Halo analysis [24] are also plotted for comparison. 107
- Figure 7.17: 95% CL lower limits on the decay lifetime  $\tau$  for the  $b\bar{b}$  (top left),  $\tau^+\tau^-$  (top right),  $t\bar{t}$  (bottom left), and  $W^+W^-$  (bottom right) dark matter spectra assuming the Einasto (red) and Burkert (blue) spatial profiles. The correspondingly colored shaded regions are the systematic uncertainty bands. The corresponding limits from the previous HAWC Galactic Halo analysis [24] are also plotted for comparison. . . . . 108
- Figure 7.18: Additional 95% CL lower limits on the decay lifetime  $\tau$  for the  $e^+e^-$  (top left),  $\mu^+\mu^-$  (top right), and  $Z^0Z^0$  (bottom) dark matter spectra assuming the Einasto (red) and Burkert (blue) spatial profiles. The correspondingly colored shaded regions are the systematic uncertainty bands. The corresponding limits from the previous HAWC Galactic Halo analysis [24] are also plotted for comparison. . . . . 109

Figure 7.19: Comparison of 95% upper limits when the spatial template is shifted in declination by the pointing systematic. The limits are indistinguishable at the higher masses and minimally effected at the lower dark matter masses. Therefore the pointing effect is treated as negligible. . . . .	110
Figure 7.20: Comparison of dark matter annihilation upper limits with and without the electroweak correction for the $b\bar{b}$ (left) and $W^+W^-$ (right) channels. As expected, the $W^+W^-$ channel shows a strong improvement with the EW correction, while the $b\bar{b}$ channel is only weakly effected by it. . . . .	110
Figure 7.21: Statistical bands for 95% CL upper limits on dark matter annihilation into the $b\bar{b}$ (left) and $\tau^+\tau^-$ (right) channels assuming a Burkert spatial profile. In all cases, the actual limits fall well within the expected range from statistical fluctuations. . . . .	112
Figure 7.22: Statistical bands for 95% CL upper limits on dark matter annihilation into the $b\bar{b}$ (left) and $\tau^+\tau^-$ (right) channels assuming an Einasto spatial profile. In all cases, the actual limits fall well within the expected range from statistical fluctuations. The expected limits from the previous HAWC (Fermi Bubble) analysis [24] are also plotted to more clearly show the improvement in sensitivity from this analysis independent of statistical fluctuation effects. . . . .	112
Figure 7.23: Statistical bands for 95% CL lower limits on dark matter decay into the $b\bar{b}$ (left) and $\tau^+\tau^-$ (right) channels assuming a Burkert spatial profile. In all cases, the actual limits fall well within the expected range from statistical fluctuations. . . . .	113
Figure 7.24: Statistical bands for 95% CL lower limits on dark matter decay into the $b\bar{b}$ (left) and $\tau^+\tau^-$ (right) channels assuming an Einasto spatial profile. In all cases, the actual limits fall well within the expected range from statistical fluctuations. The expected limits from the previous HAWC (Fermi Bubble) analysis [24] are also plotted to more clearly show the improvement in sensitivity from this analysis independent of statistical fluctuation effects. . . . .	113
Figure 7.25: Statistical bands on the 95% CL upper limits on the velocity-weighted annihilation cross-section $\langle\sigma v\rangle$ for the $b\bar{b}$ (top left), $\tau^+\tau^-$ (top right), $t\bar{t}$ (bottom left), and $W^+W^-$ (bottom right) dark matter spectra assuming an Einasto profile. This is a supplementary figure meant to demonstrate the all channels are within statistical expectations for a given case. . . . .	114

- Figure 7.26: Additional statistical bands on the 95% CL upper limits on the velocity-weighted annihilation cross-section  $\langle\sigma v\rangle$  for the  $e^+e^-$  (top left),  $\mu^+\mu^-$  (top right), and  $Z^0Z^0$  (bottom) dark matter spectra assuming an Einasto profile. This is a supplementary figure meant to demonstrate that all channels are within statistical expectations for a given case. . . . . 115
- Figure A.1: A screenshot of the main HOMER home page. It includes all essential summary information such as total scaler rates and temperatures. The tabs on the left hand side are links to pages with further monitoring details. 121
- Figure A.2: Sample HOMER sub-page plot showing the CPU load on the site computers. Included also are fields for plotting a custom time range and the option to download a printable version. Many plots such as this exist for reading historic monitoring data. . . . . 122
- Figure A.3: High voltage control pages for both the main array(top) and the outriggers (bottom). The buttons on these pages allow for channels to be easily toggled on and off. This page exists only in the on-site version of HOMER to prevent access by non-experts. . . . . 123
- Figure A.4: Sample Slack alert sent by one of the many automatic monitoring scripts at the HAWC site. This alert indicates a FEB is running hot and includes a timestamp of when the high temperature was detected. To avoid saturating the channel, alerts are sent with a decreasing frequency if the same problematic state is detected multiple polls in a row. . . . . 124

# CHAPTER 1

## INTRODUCTION

Contemporary theoretical physics has been remarkably successful in describing the observable universe. The theory of relativity is able to describe the gravitational motion of both large celestial bodies and smaller-scale objects, while the Standard Model of particle physics characterizes the many known classes of particles into a coherent picture that is strongly consistent with observations. Unexplained phenomena nevertheless persist. The large-scale behavior of galaxies suggests the presence of a dark type of matter that cannot be observed optically and possesses inferred properties that are incompatible with any particle in the Standard Model picture. Meanwhile, the Standard Model is known to be incomplete and possible extensions suggest natural candidates for a dark matter particle. Modern colliders have nonetheless been unable to detect any of these hypothetical particles in laboratory conditions.

This dissertation will explore these open problems from a particle astrophysics perspective. In particular, it will detail a set of analyses performed by the High Altitude Water Cherenkov (HAWC) experiment that aim to constrain explanatory models of new physics. The second chapter will explain the observational basis for dark matter and the theoretical models proposed to explain it. It will focus in particular on the potential of dark matter detection from astrophysical gamma-ray signals. In the third chapter, this work will turn to methods of detecting these gamma rays on Earth. This chapter will cover atmospheric air showers originating from cosmic rays and how these showers can be leveraged to observe gamma-ray signals. It will then transition into the implementation of these techniques via the construction of the HAWC experiment and conclude with a description of the HAWC hardware configuration. Further details about the remote control and monitoring framework can be found in an appendix at the end of this dissertation.

In the fourth chapter, the data reconstruction algorithms used by HAWC will be ex-

plained as well as the statistical basis for interpreting HAWC results in terms of physical models. A particular focus will be given to gamma-ray energy reconstruction via a neural network technique, the current advantages of this approach, and promising improvements. Chapter five will be an aside describing how this energy estimation technique allows for probing Einstein's theory of relativity at extremely high photon energies. We will derive robust constraints on possible violations of Lorentz invariance by the statistically significant detection of multi-hundred TeV gamma rays.

The focus will then return to dark matter in the sixth chapter, where we will use HAWC data to search for gamma-ray signals from undetected dark matter substructure formation in the Milky Way galaxy. Numerical models of dark matter structure formation will be introduced and used to constrain the existence of undetected dark matter sub-halos. This analysis will also yield useful information on the HAWC sensitivity to dark matter signals as a function of sky location. With these models and sensitivity information, Chapter Seven will outline an optimized search for dark matter in the main Milky Way Galactic Halo by selecting an ideal region of interest to make use of HAWC's capabilities. This analysis results in constraints on dark matter that are robust to the large systematic uncertainties arising from modeling of dark matter spatial distributions. The final chapter will then conclude the dissertation with a summary of the presented results and a brief discussion of future analyses in light of recent improvements to the HAWC detector hardware and reconstruction algorithms.

## CHAPTER 2

### WIMP DARK MATTER AND GAMMA RAYS

#### 2.1 Introduction

The mass of the known universe is dominated by a dark component that is not optically observable. First observed by Fritz Zwicky's 1930s study of the Coma Cluster, evidence of this dark matter has gradually accumulated over the years. Its effects can be seen in phenomena such as galactic velocity dispersion and gravitational lensing through galaxy clusters [1]. These effects cannot be explained by normal, luminous matter within the known laws of gravity. This leads to the hypothesis that halos of dark matter surround the luminous components, providing the mass necessary to explain these effects [2]. Current estimates place the dark matter contribution to the mass of the universe at 86% [14]. However, the composition of the dark matter remains unknown.

#### 2.2 Evidence for Dark Matter

##### 2.2.1 Rotation Curves

One of the most historically prominent pieces of evidence for dark matter halos is rotation curve data from spiral galaxies. Following the universal law of gravitation and assuming a dominant visible matter component, the expected orbital velocity of a given star in a galaxy should trend as  $1/\sqrt{r}$ , where  $r$  is the distance from the galaxy center. However, observations by Rubin et al [3] show a leveling off of the stellar velocities as a function of  $r$ , in contradiction to the expectation from a simple gravitational analysis. Further observations of galaxies such as M33 have consistently confirmed this behavior as illustrated in Fig. 2.1 and discussed in Refs. [4, 1]. This behavior can be explained by the presence of a non-luminous halo surrounding the stars, which contributes an additional gravitational force that roughly

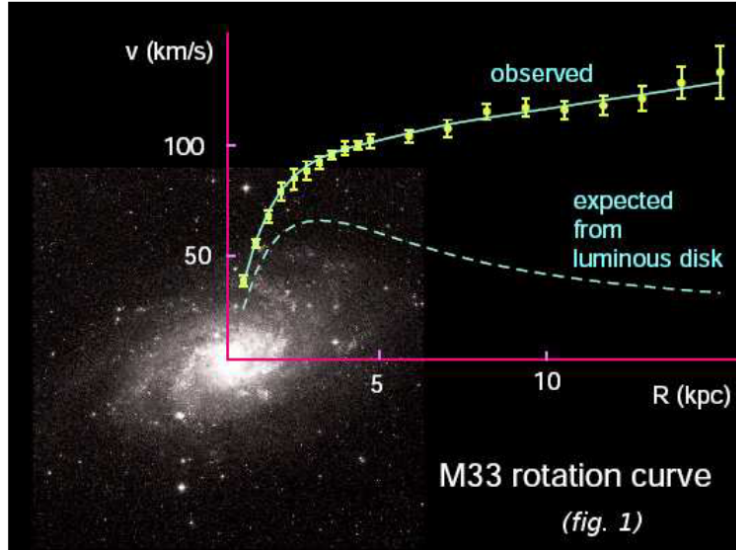


Figure 2.1: Rotation curve showing star velocities as a function of distance from the center,  $R$ , for the galaxy M33 compared with the expectation from the observed luminous component. The expected curve falls off as  $1/\sqrt{R}$ , while the actual curve flattens out. This implies the presence of a “dark” component that interacts only gravitationally. Figure from Ref. [5]. See also Ref. [4] for details on how these measurements were obtained.

cancels the fall off in the gravitational force due to the stars.

## 2.2.2 Gravitational Lensing

Further evidence for dark matter exists in the form of gravitational lensing observations of galaxy clusters. When compared with optical data, gravitational lensing shows divergences between the distribution of total mass and the observed distribution of observable matter. This is best explained by the majority of mass being located in a dark, collisionless halo [1].

One of the most famous and striking cases is that of the cluster merger event 1E0657-558, often called the Bullet Cluster [6]. Observations by the Chandra experiment of the two main clusters of 1E0657-558 show two distinct objects with fairly well-defined centers. However, the use of weak gravitational lensing to map the gravitational potential of the same objects show a substantial offset between the centers of mass and observable cluster centers as mapped by X-ray emission. The famous figure is shown in Fig. 2.2. While the X-ray emitting gaseous matter of the galaxy clusters collided and formed the eponymous

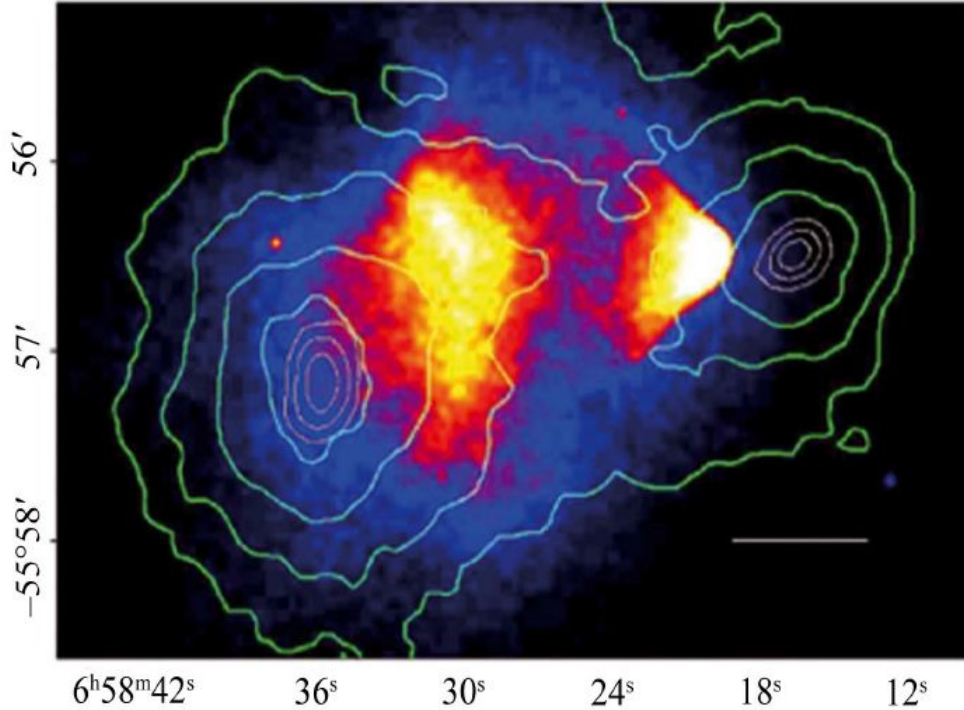


Figure 2.2: A false color image showing the baryonic gaseous component of the Bullet Cluster overlaid with contours of the gravitational lensing probe. The two clusters pictured are undergoing a merger event where the gas in the smaller bullet-shaped object collided with and passed through the larger gas cloud at a reduced speed compared to the collisionless cluster components. This shows a clear displacement between the X-ray emitting gaseous component of the two clusters and the centers of mass, indicating most of the mass is contained within collisionless halos that dominate the cluster masses. Figure from Ref. [6].

bullet shape as the smaller object slowly passed through the larger object, the majority of the mass from each original cluster continued to pass through unobstructed and emerged ahead of the corresponding gas cloud. This suggests the presence of a large collisionless and non-luminous component of the galactic mass that became offset of the luminous matter during the merge process.

### 2.2.3 Cosmic Microwave Background

A third piece of evidence for dark matter comes in the form of features in the cosmic microwave background, or CMB. First observed by McKeller in 1940 [7], and then again in 1965 by



Penzias and Wilson [8], the CMB is an all-permeating background of photons originating from the early hot state of the universe. As the universe cooled and matter recombined into nuclei and atoms, space became transparent to these early photons, allowing them to propagate without scattering off charged particles. The ensuing expansion of the universe then attenuated these photon wavelengths down to the microwave level, resulting in the CMB as observed today and leaving a powerful probe of the early state of the universe.

The so-called “last scattering” of the CMB photons before recombination left a persistent mark on the background radiation that contains important information on the universe’s composition. This takes the form of temperature anisotropies at various angular scales, as measured by the Planck experiment [9]. Analyses of these anisotropy peaks and troughs reveal the matter density universe to be approximately 86% dark matter, consistent with the dark matter-dominated rotation curves of galaxies.

## 2.3 The Standard Model

The Standard Model of particle physics (SM) is the current unifying framework by which all known elementary particles are classified. The SM also describes three of the four fundamental forces: the weak force, the electromagnetic force, and the strong force. It is based on the development of quantum field theory (QFT) throughout the 20th century and has been immensely successful in describing particle behavior [10].

In the SM, all elementary particles are classified based on the fundamental forces through which they interact. All particles can be divided into fermions, spin  $1/2$  particles that make up most ordinary matter, and integer-spin bosons. These elementary particles are the building blocks of all known composite particles which in turn make up macroscopic matter, driven by the interaction mediated through the force-carrying particles [10]. We will briefly lay out the further subdivisions and classes of these elementary particles below. A chart showing the organization of the SM particles for reference can be found in Fig. 2.3.

Fermions are further subdivided in quarks and leptons. Quarks come in six distinct

## Standard Model of Elementary Particles

		three generations of matter (fermions)			interactions / force carriers (bosons)	
		I	II	III		
mass		$\approx 2.2 \text{ MeV}/c^2$	$\approx 1.28 \text{ GeV}/c^2$	$\approx 173.1 \text{ GeV}/c^2$	0	$\approx 124.97 \text{ GeV}/c^2$
charge		$\frac{2}{3}$	$\frac{2}{3}$	$\frac{2}{3}$	0	0
spin		$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	0
		<b>u</b> up	<b>c</b> charm	<b>t</b> top	<b>g</b> gluon	<b>H</b> higgs
	<b>QUARKS</b>	<b>d</b> down	<b>s</b> strange	<b>b</b> bottom	<b><math>\gamma</math></b> photon	
		$\approx 4.7 \text{ MeV}/c^2$	$\approx 96 \text{ MeV}/c^2$	$\approx 4.18 \text{ GeV}/c^2$	0	
		$-\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$	0	
		$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	
		<b>e</b> electron	<b><math>\mu</math></b> muon	<b><math>\tau</math></b> tau	<b>Z</b> Z boson	
	<b>LEPTONS</b>	$\approx 0.511 \text{ MeV}/c^2$	$\approx 105.66 \text{ MeV}/c^2$	$\approx 1.7768 \text{ GeV}/c^2$	$\approx 91.19 \text{ GeV}/c^2$	
		-1	-1	-1	0	
		$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	
		<b><math>\nu_e</math></b> electron neutrino	<b><math>\nu_\mu</math></b> muon neutrino	<b><math>\nu_\tau</math></b> tau neutrino	<b>W</b> W boson	
		$< 1.0 \text{ eV}/c^2$	$< 0.17 \text{ MeV}/c^2$	$< 18.2 \text{ MeV}/c^2$	$\approx 80.39 \text{ GeV}/c^2$	
		0	0	0	$\pm 1$	
		$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	
						<b>SCALAR BOSONS</b>
						<b>GAUGE BOSONS</b> <b>VECTOR BOSONS</b>

Figure 2.3: Table of Standard Model (SM) elementary particles showing their organization into fermions and bosons, as well the subdivisions between them. Figure taken from [https://en.wikipedia.org/wiki/Standard\\_Model](https://en.wikipedia.org/wiki/Standard_Model).

flavors (along with corresponding anti-quarks of opposite charge), interact via all three SM forces, and are primarily found arranged into composite particles known as hadrons. Odd-numbered quark hadrons are called baryons, examples of which include the protons and neutrons of atomic nuclei, while even-numbered hadrons are known as mesons. Leptons do not interact via the strong force and consist of charged leptons such as electrons, and very light neutral leptons called neutrinos, which are known to interact only through the weak force. All fermions furthermore are categorized into three generations, with each generation containing a pair of quarks, a charged lepton, and a neutrino, along with the corresponding anti-particles. The corresponding particles of each generation are identical in electric charge but differ in mass, with each successive generation containing heavier particles.

Bosons consist of the spin-1 gauge bosons that carry the three SM forces, as well as a single spin-0 scalar boson known as the Higgs. The gauge bosons consist of the gluons,

carrier of the strong force, the  $W^+$ ,  $W^-$ , and  $Z^0$  which carry the weak force, and the familiar photon that carries the electromagnetic force. The Higgs boson eluded detection until 2012, when the ATLAS [11] and CMS [12] experiments jointly announced its discovery, adding a further strong confirmation of the SM.

The closest SM particle to a viable dark matter candidate is the neutrino. Neutrinos are neutral, colorless particles that are indeed “dark” due to not interacting electromagnetically. However, neutrinos are also extremely light and are expected to have been accelerated to relativistic velocities shortly after the Big Bang. Therefore, a sufficient abundance of neutrinos to explain the observed dark matter density would lead to a hot dark matter or HDM scenario. Numerical n-body simulations have shown that HDM is incompatible with galaxy formation as currently understood, meaning neutrinos cannot be the primary component of dark matter [2].

## 2.4 Lambda Cold Dark Matter and WIMPs

In the absence of a viable SM particle candidate for dark matter, alternative explanations are now sought beyond the Standard Model. To be consistent with the observed large-scale structure of the universe, most particle candidates are within the Lambda Cold Dark Matter, or  $\Lambda$ CDM, framework. In this picture, the majority of the universe’s energy density consists of dark energy that causes the accelerating expansion rate of the universe, with  $\Lambda$  being used to denote dark energy corresponding the cosmological constant in general relativity. Dark matter is the second most dominant component, making up about 85% of the mass of the universe with a roughly constant density and low temperature [13].

In the  $\Lambda$ CDM model, as the universe cools, the density of dark matter is expected to approach a constant, illustrated in Fig. 2.4. The present-day constant or “relic” density  $\Omega_\chi$  is determined by the following equation:

$$\Omega_\chi = \frac{\rho_\chi}{\rho_{crit}} = \frac{2h^{-2} \times 10^{-9} \text{GeV}^{-2}}{\langle \sigma v \rangle}, \quad (2.1)$$

where  $\langle \sigma v \rangle$  is the velocity-weighted cross section of two-particle dark matter annihilation,  $\rho_\chi$

is the energy density of dark matter, and  $\rho_{crit}$  is the critical density at which the universe has zero curvature [1]. Based on the most recent data available from the Planck experiment [14], this formula yields a prediction for the relic cross section of  $\langle\sigma v\rangle_{\text{therm}} = 3 \times 10^{-26} \text{cm}^2 \text{s}^{-1}$ .

Although no SM particle is able to reproduce the observed relic density in the  $\Lambda$ CDM framework, well-motivated extensions of the SM do produce particles compatible with  $\Lambda$ CDM. One of the most popular candidates are Weakly Interacting Massive Particles, or WIMPs. As indicated by the name, these are hypothesized particles with a non-zero mass that interact via a weak-scale force and in many theoretical frameworks are their own anti-particle [2] [15]. This weak-scale self-annihilation naturally reproduces the observed relic density while maintaining the cold dark matter state of the known universe. WIMPs are therefore one of the most promising dark matter candidates and many contemporary experiments aim to find direct evidence of them.

## 2.5 Detection Approaches

Three approaches currently exist for detecting dark matter: direct detection, collider detection, and indirect detection. Each method relies on a different class of dark matter and Standard Model particle interaction. The particle physics aspect of these approaches is best summarized by seeing them as different directions on a Feynman diagram, as shown in Fig. 2.5.

Direct detection aims to observe interactions between dark matter and luminous matter. Experiments designed for direct detection include LZ, XENONnT, and SuperCDMS [16]. It is based on the principle that Earth would be in motion with respect to the WIMP dark matter halo surrounding the Milky Way galaxy. With enough time and targets, it may be possible to observe scattering of nucleons and WIMPs due to the relative motion between them even if the cross-section is extremely low. Experiments of this type rely on removing as many SM-on-SM interaction opportunities as possible by constructing low-radiation environments filled with inert gasses. If enough SM sources of radiation have been

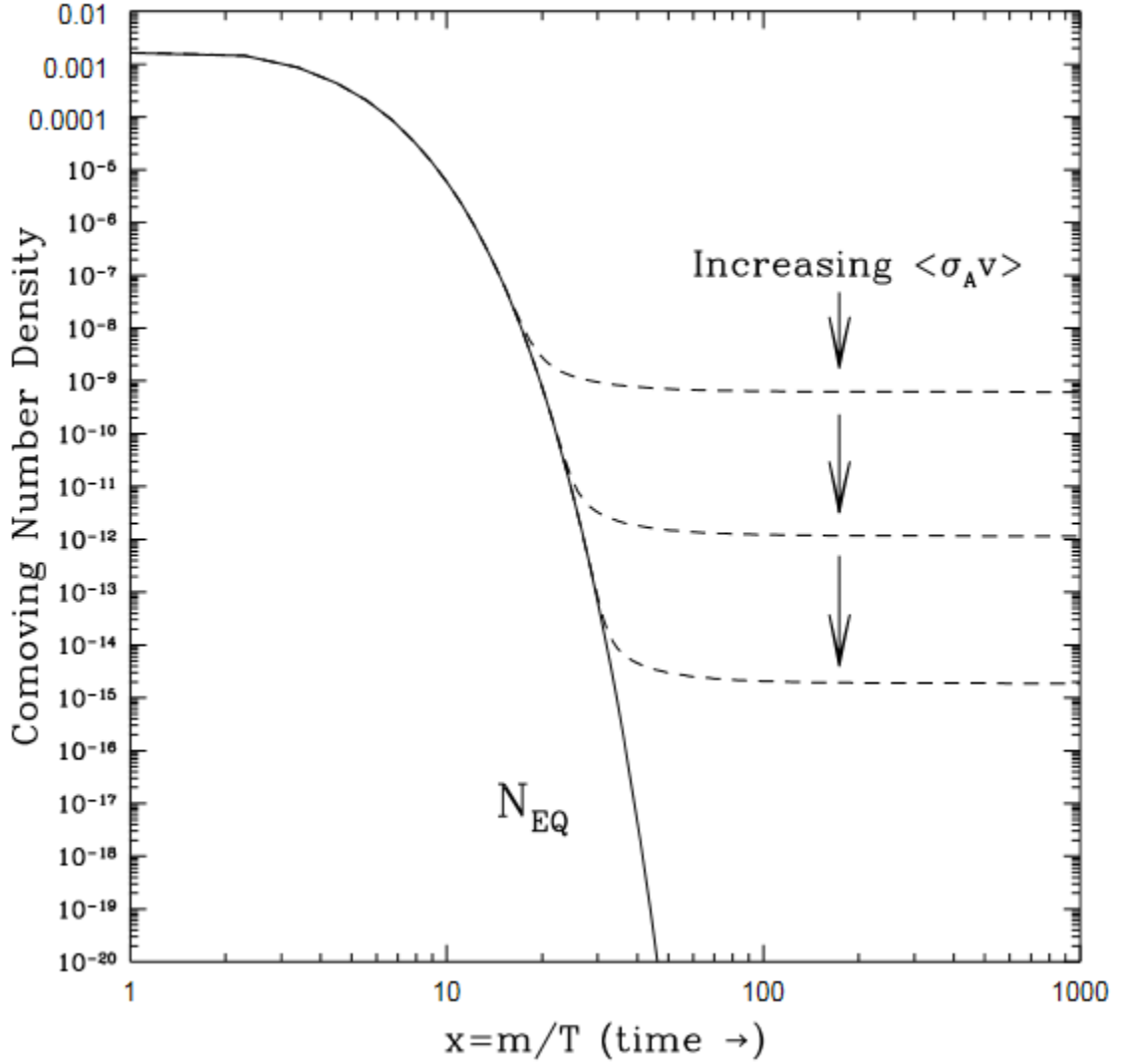


Figure 2.4: Simulated plot of thermal freeze-out in the  $\Lambda$ CDM framework for various assumed velocity-weighted cross sections showing the dark matter number density approaching a constant in the present day. Estimates of the current dark matter density (see Eq. 2.1) allow for a prediction of the velocity-weighted cross section. Figure from Ref. [15].

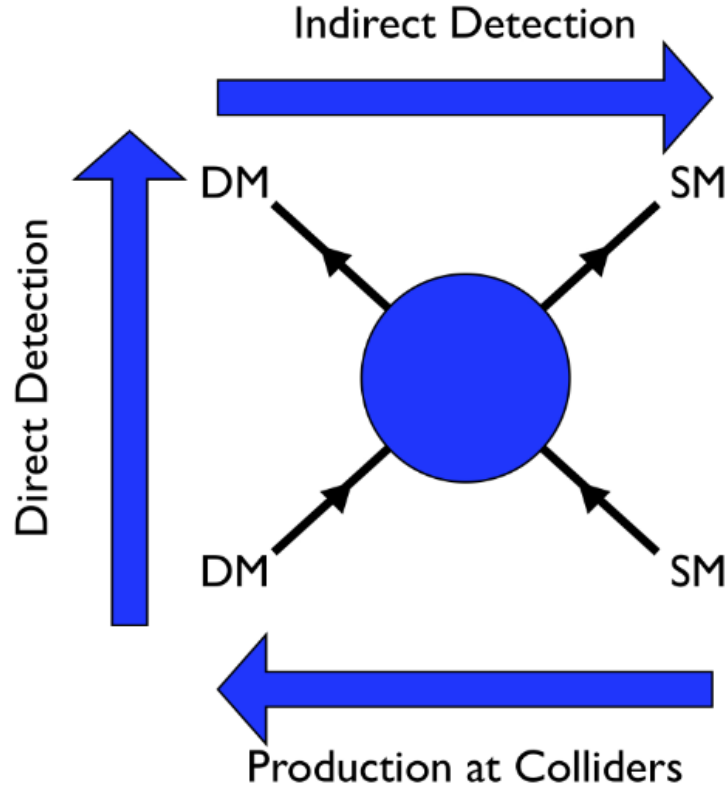


Figure 2.5: Feynman diagram showing the interaction behind the various dark matter detection approaches. “DM” refers to dark matter particles, while “SM” refers to Standard Model particles. The large blue circle is a stand-in for the theorized physical interactions between dark matter and the Standard Model and the arrows along each axis show the direction one should follow to trace the interaction probed by each detection class. Following a given arrow from start to finish shows the initial particles of an interaction scattering and resulting in the final product particles of that interaction.

removed, the target nuclei should be sufficiently isolated that any scattering interactions could only come from WIMPs that passed sufficiently deep into the chamber. As of this writing, no experiment has found evidence of dark matter scattering off nucleons [16].

Collider detection refers to attempts to produce dark matter particles through interactions between SM particles in accelerators. Currently, the ATLAS [17] and CMS [18] experiments at the Large Hadron Collider of CERN are actively searching for created dark matter during collider runs. Although neither experiment is expected to be sensitive enough to directly detect produced dark matter, the creation of a non-SM particle can be inferred

by conservation of momentum. A significant amount of missing transverse momentum (the component of particle momentum perpendicular to the collider beam) in the sum of the products following a collision could be evidence of a dark matter-like particle being produced and passing through the detector unseen. Such evidence has yet to be detected though efforts to produce dark matter particles in colliders continue [19].

Indirect detection searches look for byproducts of WIMP interactions that result in observable, SM particles. Because dark matter over-densities are detectable via gravitational effects, it is possible to perform targeted searches for expected products of dark matter interaction. Moreover, the expected dark matter annihilation cross-section calculated from Eq. 2.1 means that indirect searches have a well-defined parameter space to probe, and are able to rigorously exclude dark matter at different mass scales if they probe the thermal scale with no detection. Indirect detection will be the main focus of this thesis.

## 2.6 WIMPs and Gamma Rays

Astrophysical indirect detection relies on the fact that under the WIMP hypothesis, dark matter annihilation can produce Standard Model particles through weak interactions. For sufficiently high (GeV scale or greater) assumed WIMP masses, these particles will in turn produce gamma-ray photons mainly via pion decay, but also through inverse Compton scattering of photons off produced electron-positron pairs [20]. The expected differential gamma-ray flux,  $d\Phi$ , per unit energy  $dE$  from a dark matter halo is described by the following equations, respectively modeling annihilation and decay:

$$\frac{d\Phi}{dE_{\text{ann}}} = \frac{J\langle\sigma v\rangle}{8\pi M^2} \frac{dN(M, \text{channel})}{dE} \quad (2.2)$$

$$\frac{d\Phi}{dE_{\text{decay}}} = \frac{D}{4\pi M\tau} \frac{dN(M, \text{channel})}{dE}, \quad (2.3)$$

where  $\langle\sigma v\rangle$  is the velocity-weighted annihilation cross section,  $\tau$  is the decay lifetime,  $M$  is the dark matter particle mass, and the  $J$  and  $D$  factors contain the astrophysical information of the assumed dark matter density profile. See Ref. [20] for a detailed derivation of these functional forms.

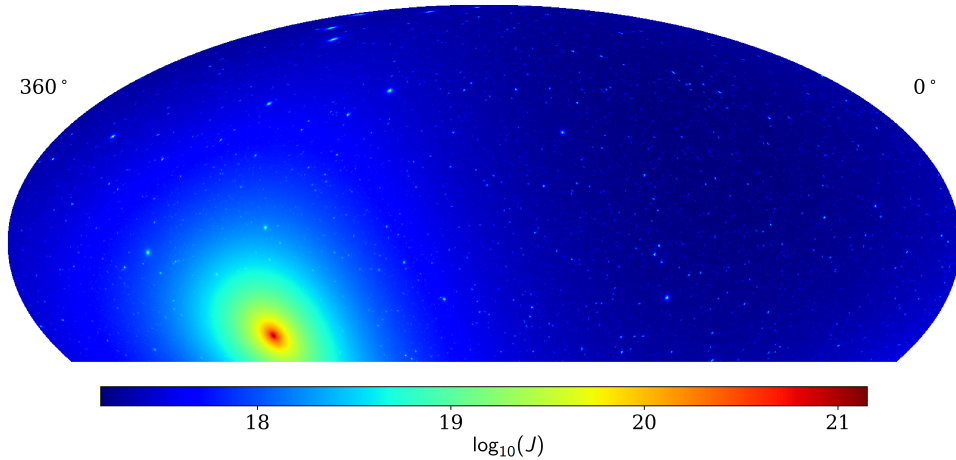


Figure 2.6: Sample map showing one simulation of the smooth halo and substructure of the Milky Way halo. The color scale shows the logarithm of the J-factor (Eq. 2.4) integrated over a single pixel width of solid angle (about 0.25 square degrees in this map). The high J-factor at the Galactic Center is clearly visible with substructure distributed throughout the sky. The details for generating this simulation are discussed further in Chapters 6 and 7.

The  $J$ -factor and  $D$ -factor are defined respectively as:

$$J = \int \int \rho_{dm}^2(l, \Omega) dl d\Omega \quad (2.4)$$

$$D = \int \int \rho_{dm}(l, \Omega) dl d\Omega, \quad (2.5)$$

where  $\rho_{dm}$  is the dark matter density profile (typically containing both a smooth component and a contribution from sub-halos),  $dl$  is the line of sight to the halo, and  $\Omega$  is the solid angle of the observation. The J-factor contains a density-squared factor due to annihilation requiring two particles to interact, while the D-factor includes only one power of the density profile since decay is a single-particle process. A sample map of J-factors from the Galactic Halo with simulated substructure over-densities is shown in Fig. 2.6.

The quantity  $\frac{dN}{dE}$  in Eq. 2.2 and Eq. 2.3 is the gamma-ray spectrum from a single dark matter annihilation [13]. It is typically calculated using Monte Carlo simulations assuming each dark matter interaction produces only one class of SM particle, referred to as a “chan-



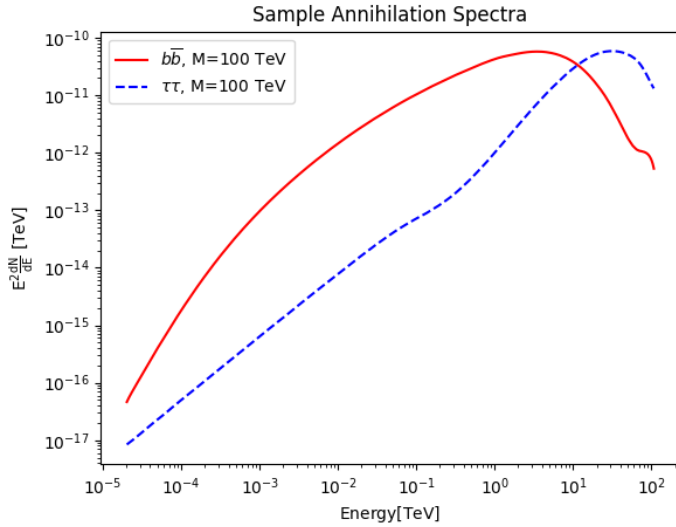


Figure 2.7: Example dark matter energy spectra for two annihilation channels ( $b\bar{b}$  and  $\tau^+\tau^-$ ), assuming a dark matter mass of 100 TeV and showing the characteristic hard cutoff feature. These two channels are typically taken as representative of dark matter gamma-ray spectra, with  $b\bar{b}$  falling off most quickly with energy and  $\tau^+\tau^-$  falling off slowest with energy. Most other dark matter spectra fall between these two cases.

nel”. Since the branching fractions into various channels are unknown, these calculations typically assumed a 100% branching ratio into a given channel to avoid introducing a high number of additional free parameters. More complicated spectra can be computed by simply summing contributions from these single-channel spectra assuming a given set of branching ratios. One tool for calculating these quantities is PYTHIA, a set of code libraries meant to simulate elementary particle interactions and record the final products [21]. An example of dark matter spectral shapes generated using PYTHIA 8.2 is shown in Fig. 2.7. These energy spectra all show a characteristic cutoff at the assumed dark matter mass, a unique spectral feature not expected to occur in spectra arising from SM gamma-ray production mechanisms. This hard cutoff is therefore a tell-tale sign of dark matter production of gamma rays, and is the feature indirect detection aims to observe to verify the presence of WIMP dark matter.

## 2.7 Potential Search Targets

Regions with known dark matter over-densities are expected to produce a gamma-ray excess. Any dense star-forming region with an associated dark matter halo can be considered a target, so long as sufficient stellar velocity data or N-body simulation results are available to estimate the density profile.

A popular class of targets is the set of dwarf spheroidal galaxies. These structures, as the name suggests, are small galaxies with a regular spherical shape. Their sphere-like shape makes calculating the J and D-factors from stellar velocity data fairly simple, and their sparse stellar content means there is very little gamma-ray background from visible-matter production mechanisms. However, many of the dwarfs expected to yield the highest fluxes from dark matter production suffer from substantial systematic uncertainties due to these sparse stellar populations and the difficulty of obtaining accurate stellar velocity measurements [22].

Larger galaxies and galaxy clusters are another common target. These include the Virgo Cluster, the M31 galaxy [23], and the local Milky Way Galactic Halo [24]. The greater distance of the extra-galactic targets compared to the dwarfs is balanced by a larger corresponding halo, and they often yield very high expected fluxes. Because of the higher stellar population, it is often necessary to consider contributions from normal-matter emission when searching for signs of WIMPs. Moreover, these objects are typically highly extended on the angular scale of current instruments and suffer from substantial systematic uncertainties in the flux arising from the dark matter density profile. The Galactic Halo is particularly sensitive to the behavior of the density profile towards the core. It is possible to mitigate these uncertainties by considering regions of the halo further away from the core, where the different candidate density profiles behave similarly. This requires considering a very large region of interest (ROI) in order to obtain sufficient sensitivity, and such an analysis will be performed in Chapter 7.

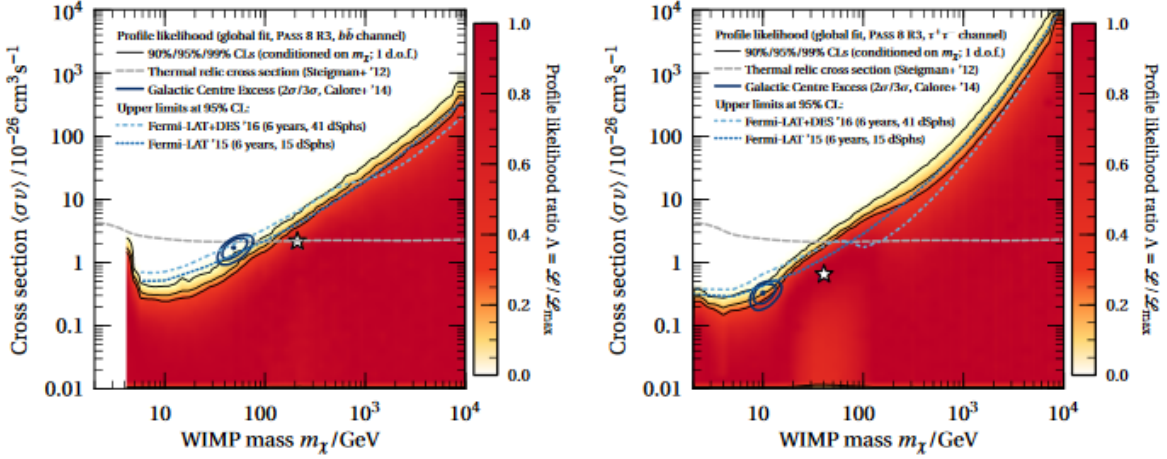


Figure 2.8: 95% confidence upper limits on dark matter annihilation in  $b\bar{b}$  (left) and  $\tau^+\tau^-$  (right) channels from the FERMI-LAT observation of the dwarf spheroidal galaxies. Below assumed dark matter masses of 100 GeV, these constraints are lower than the predicted thermal cross section of the  $\Lambda$ CDM model (Eq. 2.1), therefore completely excluding WIMP dark matter with these masses. Figure taken from Ref. [26].

## 2.8 Motivation For High-Mass WIMPs and Large Targets

Several active classes of experiments exist to perform indirect dark matter searches. The Fermi Large Area Telescope (Fermi-LAT) experiment is a satellite gamma-ray detector that performed a series of robust searches for dark matter with GeV masses. It has a very wide field of view and can detect gamma rays without interference from atmospheric effects due to being located on a satellite [25]. A combined analysis of dwarf spheroidal galaxies by Fermi-LAT has already probed the thermal cross section predicted by Eq. 2.1 at the GeV energy scale [26]. As shown in Fig. 2.8, these results exclude dark matter annihilation with the predicted thermal cross section at a 95% confidence for assumed particle masses below 100 GeV, therefore motivating the need to search for WIMPs at the TeV mass scale.

Multi-TeV mass WIMPS will produce fairly “hard” energy spectra with a large number of high-energy photons (in contrast, “soft” spectra tend to fall off faster with energy and produce relatively few high-energy photons). There are considerably fewer astrophysical sources of TeV gamma rays than lower-energy gamma rays and X-rays, almost all of which

are associated with the Galactic Plane. Therefore, the background of gamma rays from SM production mechanisms is far less of a concern at the highest energies, making it possible to probe even highly star-rich sources such as the Milky Way Galactic Halo. One class of experiment sensitive to emission from these heavier particle masses are Imaging Air Cherenkov Telescopes or IACTs. These instruments detect the effects of cosmic gamma rays in the atmosphere. They benefit from excellent angular and energy resolution, allowing them to resolve the energy spectra of sources with sufficient precision to detect the unique spectral features associated with dark matter. However, they suffer from limited time windows for observation, only being able to collect data during dark nights, and their angular resolution comes at the cost of narrow fields of view [27].

With these features, the HESS IACT has probed the thermal scale for certain channels in the TeV dark matter mass range based on an analysis of the Milky Way Galactic Center [28]. Shown in Fig. 2.9, this analysis excludes thermal dark matter at the low TeV mass scale. However, the HESS analysis relies on the assumption that the dark matter density profile peaks sharply towards the Galactic Center (also known as being “cuspy”), due to its narrow field of view only allowing it to observe the very center of the galaxy. It is therefore quite sensitive to systematic uncertainties on the exact behavior of the Galactic dark matter density profile. This systematic will be explored further in Chapters 6 and 7.

A third class of instrument sensitive to indirect dark matter signals are ground-based extensive air shower (EAS) detectors. Like IACTs, these experiments are highly sensitive to multi-TeV gamma rays and are able to probe high-mass WIMP dark matter. While they suffer from lower angular resolution, they have the advantage of a near-100% livetime, being able to run at all times of day regardless of ambient light level [29]. They can easily perform stacked analyses like those done by Fermi-LAT due to having a comparably large field of view, but at higher assumed WIMP masses that have yet to be probed as thoroughly by current observations. They also benefit from wide fields of view that allow for observations of highly extended sources that are much larger than those able to be probed by an IACT.

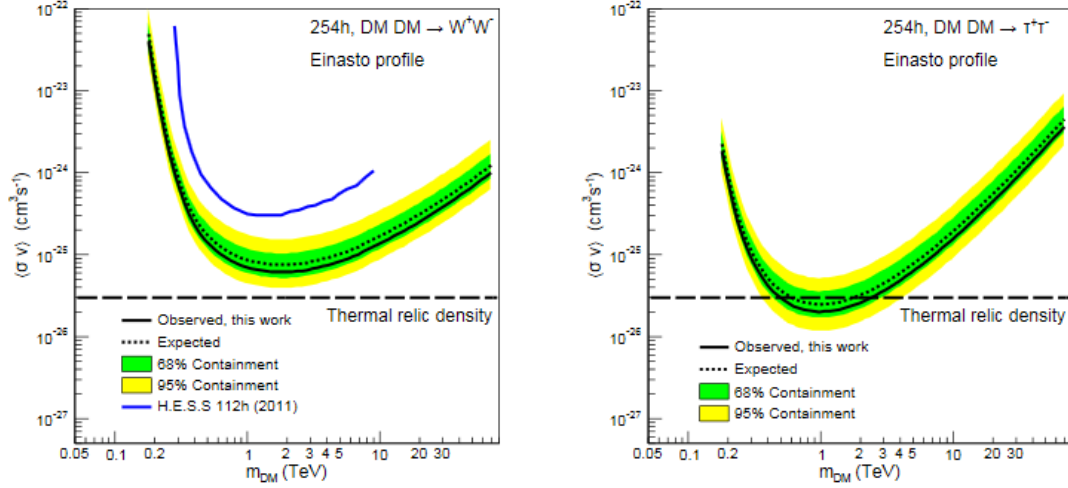


Figure 2.9: 95% confidence upper limits on dark matter annihilation in the  $W^+W^-$  (left) and  $\tau^+\tau^-$  (right) channels from the HESS observation of the Galactic Center. The model  $W^+W^-$  channel in the HESS analysis behaves analogously to the  $b\bar{b}$  channel and plotted here as a representative soft spectrum. These results are able to exclude WIMPs of certain multi-TeV masses assuming annihilation into harder channels like  $\tau^+\tau^-$ , but depend on assumptions about the exact behavior of the density profile. Figure taken from Ref. [28].

In particular, an EAS detector can perform searches in the Galactic Halo that do not require knowing the exact density profile behavior towards the center, due to being able to observe wider regions of the Halo further from the center. One such EAS detector, the High Altitude Water Cherenkov (HAWC) observatory, will be the main focus of this thesis. The next chapter will introduce the mechanisms by which HAWC detects cosmic gamma rays and explain the detector construction and hardware.

## CHAPTER 3

### AIR SHOWERS AND THE HAWC DETECTOR

#### 3.1 Cosmic Gamma Rays and Atmospheric Air Showers

Historically, the term “cosmic ray” refers to charged particles interacting with Earth’s atmosphere. These particles consist almost entirely of hadrons, primarily protons, but also include a small fraction of heavier nuclei. Originally discovered by Victor Hess, they made their presence known to him by an unexpected increase in the radiation measured by a Geiger counter as it moves higher above ground level [30]. Because of their charged nature, cosmic rays are deflected by the interstellar magnetic fields permeating space and do not arrive in the same direction as the line of sight to their acceleration source. Particles arriving at Earth also include neutral particles such as gamma rays. Due to the lack of deflection from magnetic fields, these gamma rays do arrive along the path to their sources and are of primary importance in particle astrophysics.

When a cosmic ray or gamma ray interacts with the atmosphere, typically at a height of 15 to 20 km above sea level, it scatters into a set of secondary particles. These new particles continue to scatter off the atmosphere into a greater number of lower-energy particles, creating a phenomenon known as an extensive air shower. As described by Heitler, gamma-ray showers are characterized by successive creation of electron-positron pairs, each of which produces more gamma rays via bremsstrahlung, the radiation produced when charged particles are decelerated by the presence of atoms [31]. Due to the opening angles of various interactions, the shower will widen over time (see Fig. 3.1 for an example). The shower will continue to propagate until the secondary particle energies fall below the ionization threshold, at which point they will gradually be reabsorbed by the atmosphere.

The particle density as a function of distance from the shower axis is determined by a complex system of cascade relations that describes the particle production mechanisms.

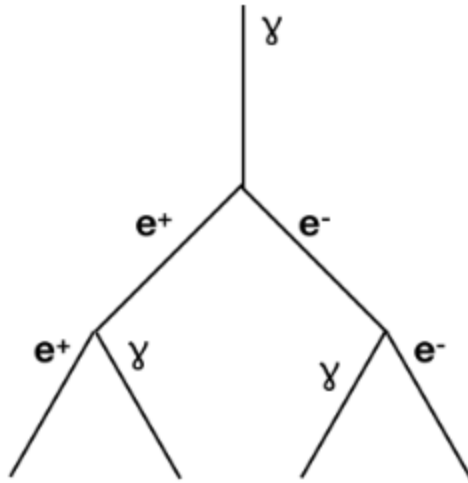


Figure 3.1: Illustration of a gamma-ray induced extensive air shower. The shower propagates itself via successive pair production and bremsstrahlung. The shower footprint widens over time due to the finite opening angles of successive interactions.

An approximate solution is obtained in Ref. [32] by Nishimura, Kamata, and Greisen. The resulting solution is called the NKG function after the authors' initials and has the form:

$$f(r) \sim (r/r_m)^{s-2}(1 - r/r_m)^{s-4.5}, \quad (3.1)$$

where  $r$  is the distance from the shower axis,  $r_m$  is the Moliere radius (the distance from the shower axis expected to contain 90% of an air shower's energy deposition, which is about 124 m at the elevations relevant to this dissertation [30]), and  $s$  is the shower age. The shower age  $s$  characterizes the longitudinal development of the shower as a function of primary gamma-ray energy  $E_\gamma$  and is given by:

$$s = \frac{3}{1 + 2y/t}, \quad (3.2)$$

where  $y = \log(\frac{E_\gamma}{E_c})$ ,  $E_c$  is the critical energy where the shower begins to die off due to ionization (about 80 MeV in the atmosphere) and  $t$  is the ratio of the path length to the radiation length [30].

Charged cosmic rays also produce extensive air showers [30]. Unlike gamma-ray showers, hadronic showers contain non-electromagnetic contributions due to the presence of strong

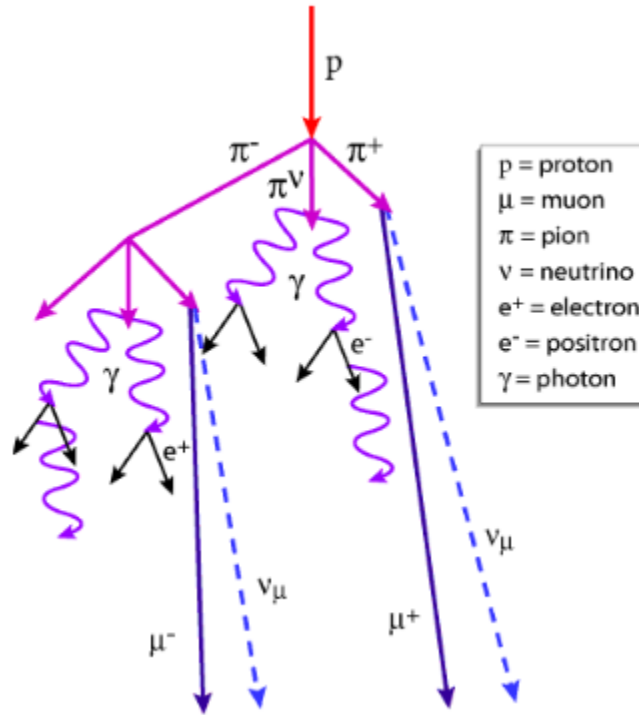


Figure 3.2: Illustration of a hadronic air shower. Unlike the gamma-ray shower shown in Fig. 3.1, this shower contains pions, neutrinos, and other non-electromagnetic components.

interactions, as illustrated in Fig. 3.2. In particular, cosmic-ray showers will produce charged pions and muons as they propagate rather than simple electron-positron pairs and photons. These additional features allow for the discrimination between cosmic-ray and gamma-ray showers by detectors.

If the shower secondaries are traveling faster than the group velocity of the radiation they produce through bremsstrahlung, the radiation builds up into a cone of light. This phenomenon, called the Cherenkov effect, is illustrated in Fig. 3.3. This Cherenkov light is a tell-tale signal of air shower secondaries passing through a medium and is the primary mechanism by which ground-based observatories detect gamma-ray showers. IACTs, as discussed in the previous chapter, rely on detecting Cherenkov light in the atmosphere itself, which is what leads to their limited livetime since a high amount of ambient light makes it impossible to detect the Cherenkov radiation. An alternative approach that avoids this



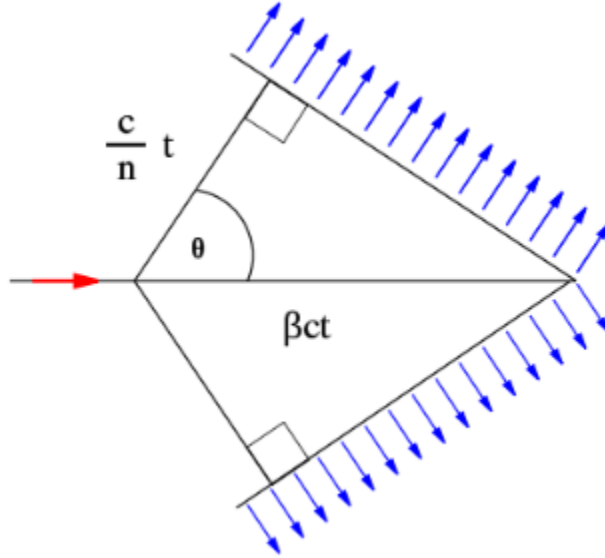


Figure 3.3: Illustration of a Cherenkov light cone with opening angle  $\theta$  produced in a material. Here,  $c$  is the speed of light in vacuum,  $n$  in the index of refraction of the material,  $\beta c$  is the speed of the radiating particle, and  $t$  is the amount of time the particle has been propagating in the medium. Created by Arpad Horvath and used under a Creative Commons BY-SA 2.5 license.

issue is to construct an environment insulated from ambient light at a high enough altitude that shower secondaries will be able to propagate through it before the shower completely dies off through ionization. This allows for the detection of air showers at all times of day, since secondaries entering this environment will still produce Cherenkov light, particularly if they are made to pass through water, which has an even higher index of refraction than the atmosphere. This is known, appropriately, as the Water Cherenkov Technique and is the mechanism of cosmic gamma-ray detection this dissertation will focus on.

### 3.2 HAWC Description

The High Altitude Water Cherenkov (HAWC) detector is a gamma-ray observatory located at Sierra Negra, Mexico. Consisting of an array of 300 Water Cherenkov Detectors (WDCs) and covering an area of  $22,000 \text{ m}^2$ , it is one of the most sensitive instruments to multi-TeV gamma rays currently operating. Shower secondaries moving faster than the group velocity of photons in water (about  $2.26 \times 10^8 \text{ m/s}$ ) will produce Cherenkov light cones in



Figure 3.4: A photograph of the HAWC detector showing the main array and the mountain Pico de Orizaba in the background.

the WCDs, which are then detected with four photomultiplier tubes (PMTs) mounted at the bottom of each tank. Timing information is used to reconstruct the angle of arrival, and the distribution of charge across the array is used to reconstruct various properties of the shower primary [33]. HAWC is sensitive to gamma rays with energies of at least 300 GeV and over 100 TeV. It is well-suited for detecting possible signals from multi-TeV mass WIMPs. In addition, HAWC operates on a near-continuous duty cycle with a roughly 2 sr instantaneous field of view that makes it ideal for performing survey-style observations [34]. This wide field of view allows HAWC to collect data from 2/3 of the sky each day as the Earth rotates. A photograph of the HAWC array is shown in Fig. 3.4.

### 3.3 Hardware and Construction

Each WCD in the HAWC detector is a cylindrical tank of diameter 7.3 m and height 5 m. Inside each tank is a black plastic bladder constructed to prevent reflection of Cherenkov light and be opaque to outside light pollution. These bladders are filled with 4.5 m of water,

purified such that Cherenkov photons will have an attenuation length longer than the tank size (see Fig. 3.5 for a single-tank schematic). Three of the four PMTs at the bottom of each WCD are 8-inch diameter Hamamatsu R5912 series PMTs repurposed from the previous Milagro experiment once located near Los Alamos, New Mexico [29]. The fourth PMT in the center is a Hamamatsu R7081, 10 inches in diameter and with nearly twice the quantum efficiency of the smaller PMTs. This difference in response between the two classes of PMT is taken into account in the reconstruction algorithms described in Chapter 4. All PMTs are connected to an individual high voltage meant to match the respective PMT gains, with each being on the order of 1700 V.

Roughly in the center of the array is the counting house, where all PMTs signals are processed. The counting house contains all the electronics needed for signal processing as well as computers used for remote control and monitoring of the detector functions (more details in Appendix A). Each PMT is connected to the counting house via a 610 ft RG-59 cable (the cable length is kept constant for each PMT to ensure uniformity in response delays), which feeds the signal from a PMT into one of the 16 slots on each of the front-end boards (FEBs), which also serve to administer the high voltage to each PMT. The FEBs integrate and amplify the signal and serve as the first step in the data acquisition system, or DAQ.

### 3.4 Data Acquisition System

Due to the remote location of HAWC, the DAQ is set up to be as automated and robust as possible, able to function for days or even weeks at a time without human intervention. The DAQ process starts once a PMT signal enters an FEB, where it is sent to two comparators with separate thresholds. This results in a signal with either 2 or 4 transitions depending on if one or both thresholds were reached. The triggering process is illustrated in Fig. 3.6 and further details can be found in Ref. [35].

The time over threshold (or ToT) of a PMT signal is measured by one of ten CAEN

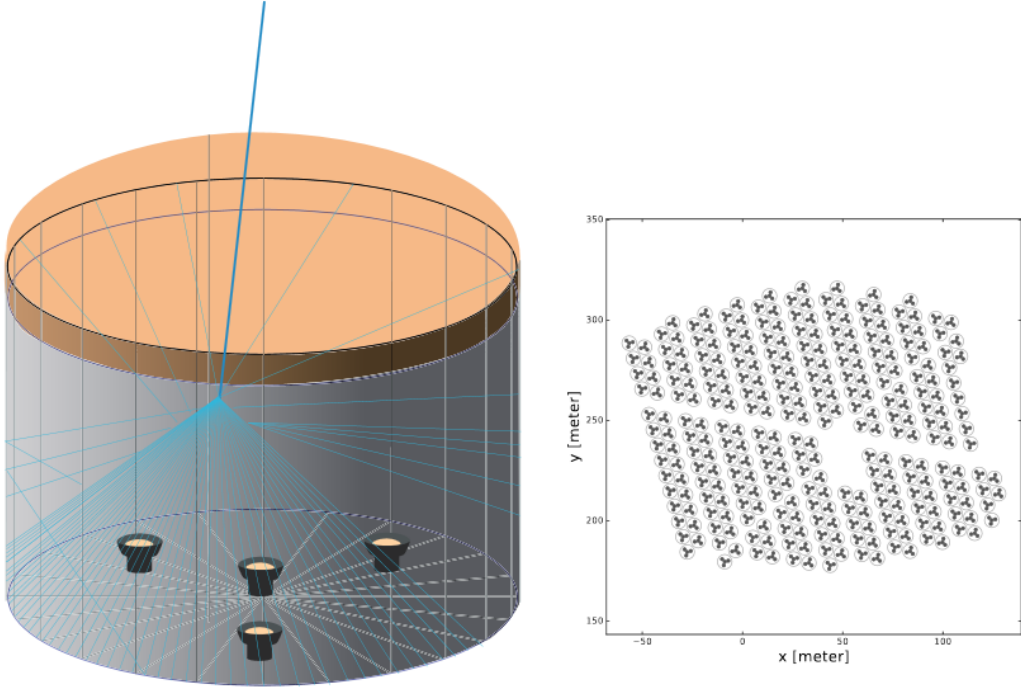


Figure 3.5: Illustration of a shower secondary producing Cherenkov light in a HAWC WCD (left), and a map of the HAWC layout (right). Image taken from Ref. [33].

V1190A-2eSST Time to Digital Converters (TDCs), each of which handles up to 128 PMT channels. Each TDC contains two thresholds that correspond to the approximate number of photoelectrons (PEs) produced by a PMT, and the signal will contain four or two transitions depending on whether the voltage reached both thresholds or only the lower one (see Fig. 3.6). The lower of the two thresholds corresponds to about 0.25 PE, while the higher threshold corresponds to 4 PEs. Each TDC has a time resolution of 98 ps, and outputs timing information in batches every  $25 \mu\text{s}$ , with a  $1 \mu\text{s}$  overlap to avoid missing triggers along the edges. The TDCs are themselves controlled by an array of General Electric XVB602-13240010 single-board computers (SBCs), which run a CentOS operating system equipped with a custom program called Readout Process. Each TDC is set to trigger the Readout Process protocol every 25 data outputs, which then transfers the TDC timing data from the SBC to the main Reconstruction Client via an internal Gigabit Ethernet connection.

The Reconstruction Client is a program that runs on several dedicated computers, col-

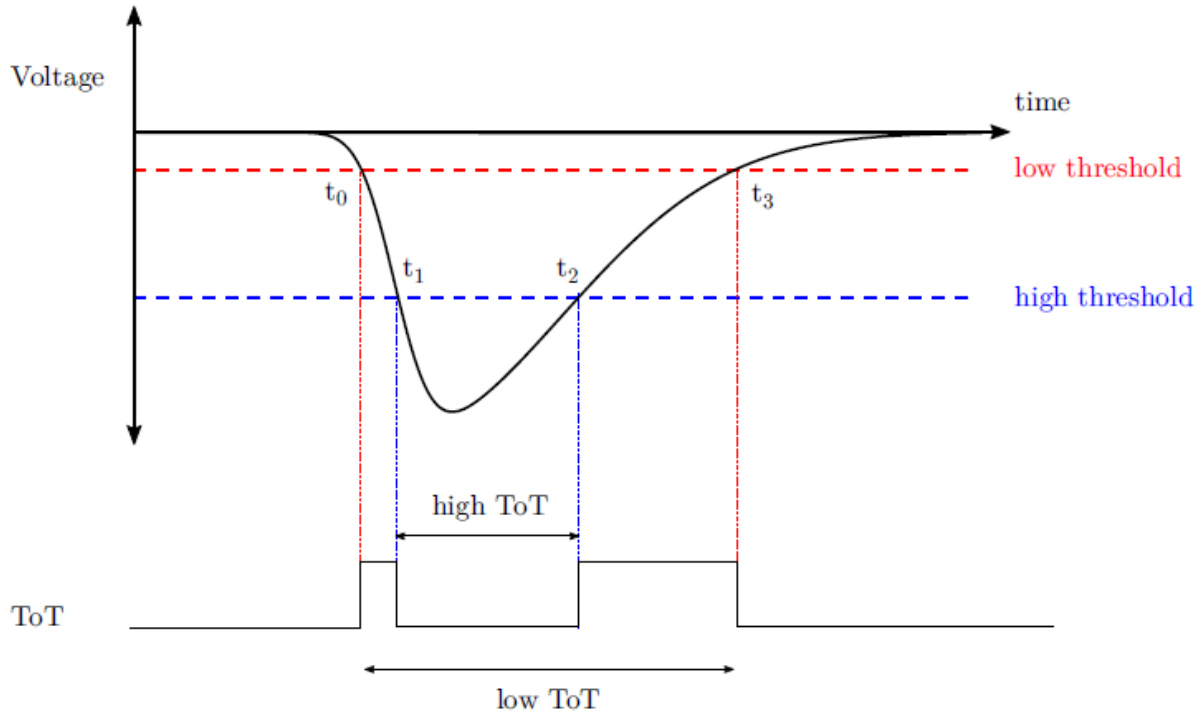


Figure 3.6: Illustration of the conversion of PMT voltage over time to an FEB output signal, showing both possible voltage thresholds. The time over threshold (ToT) is used later in the reconstruction process. Figure from [35].

lectively known as the online nodes, and begins the process for reconstructing air shower information from the PMT data. See Ref. [35] for the details of the reconstruction process and its outputs.

The entire process of running on the DAQ and initial reconstruction is performed by the Experiment Control framework. Experiment Control is an autonomous system to monitor and adjust the DAQ setting as well as stop and start data collection runs. It also reports and logs significant errors in the operation of the detector and will automatically attempt to restart the experiment when a failure causes it to crash. Run crashes are caused primarily by lightning storms and recovery is almost always quick, on the order of minutes. This system also contains the software libraries for remote monitoring of the experiment and gathering of transient alerts, detailed in Appendix A.

### 3.5 Calibration

The ToT information read by the DAQ is used to estimate the number of photoelectrons (PEs) produced in a triggered PMT, which is in turn an estimator of the charge deposited in a PMT during an event. This requires calibration of both the time it takes a PMT voltage to reach the lower threshold (called slewing time) and the relationship between a given PMT ToT and the number of PEs produced [36]. Both these classes of calibration are performed using laser measurements.

For the slewing time calibration, a laser is fired into each PMT through an array of filters of varying transmittance. Slewing time as a function of the ToT is then fit using a sum of two exponential terms and one linear term, where the linear term dominates for the majority of times and exponential terms account for edge effects at high and low ToT.

Charge calibration is performed by instead firing several laser pulses through each filter and measuring how many times the PMT triggered, where a trigger is associated with a given pulse if they coincide by less than  $2 \mu\text{s}$ . The resulting fraction of PMT triggers to laser pulses serves as an estimator of the probability of either a laser photon or a cosmic ray triggering the PMT. Comparing this fraction with the trigger rate from cosmic rays alone (measured by simply observing the PMT when the laser is not firing) allows an estimate of the number of PEs produced by the laser as a function of intensity. See Chapter 3 of Ref. [36] for further details on both classes of calibration.

### 3.6 The Outriggers

A secondary array of tanks was recently completed with the goal of augmenting the HAWC sensitivity. This addition consists of smaller tanks known as “outriggers,” with the goal being to extend the effective area of the detector with a sparser array that covers more area [37]. An image of the outriggers surrounding the main array is shown in Fig. 3.7. The individual tanks contain only a single PMT and are therefore not as sensitive to individual shower secondaries as the main array (see Fig. 3.8 for a schematic of a single outrigger tank).

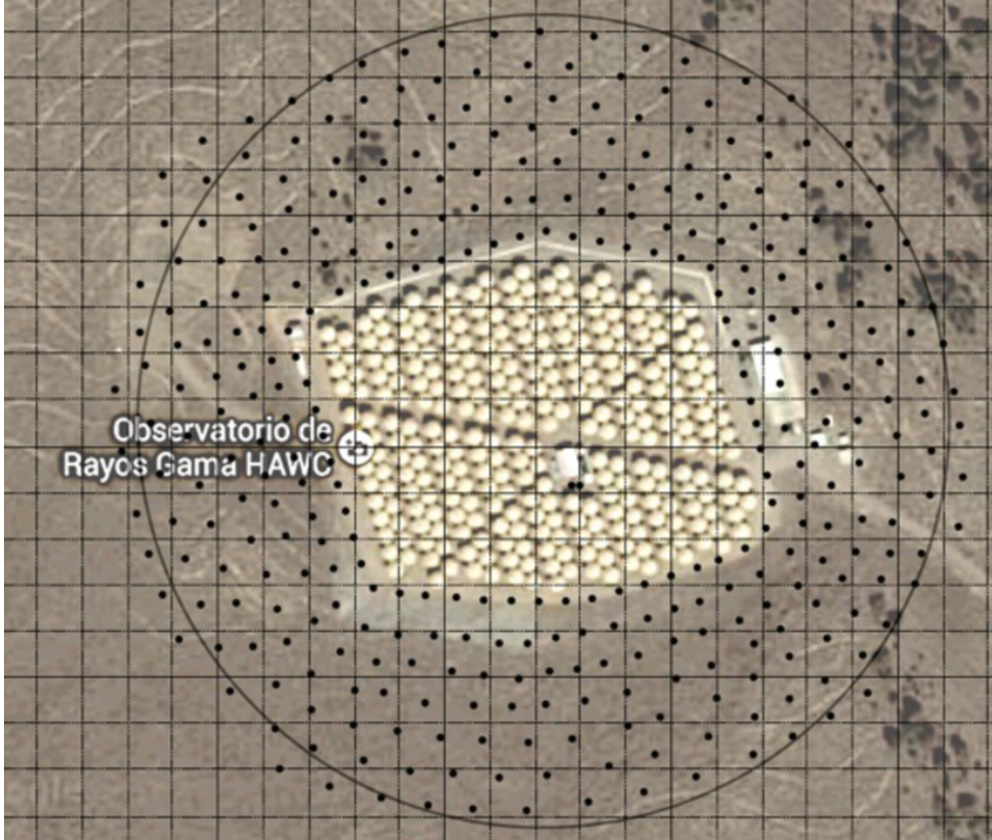


Figure 3.7: Map of the recently completed array of smaller outrigger tanks (represented by black dots) superimposed on a map of the main array. Figure from Ref. [38].

The increased effective area from the outriggers will allow for more effective reconstruction algorithms, especially for high energy showers with footprints much larger than the main array. They are currently collecting data and the HAWC collaboration is in the process of incorporating them into the reconstruction algorithms. Once fully online, the outrigger data will allow for improved energy resolution, as well as more effective algorithms for locating the shower core and reconstructing the arrival angle. The analyses in this dissertation are among those that stand to benefit from incorporation of the outriggers and future work will focus on extending these techniques to benefit from the outriggers. However, at the time of writing, none of the algorithms or analyses discussed use outrigger data.

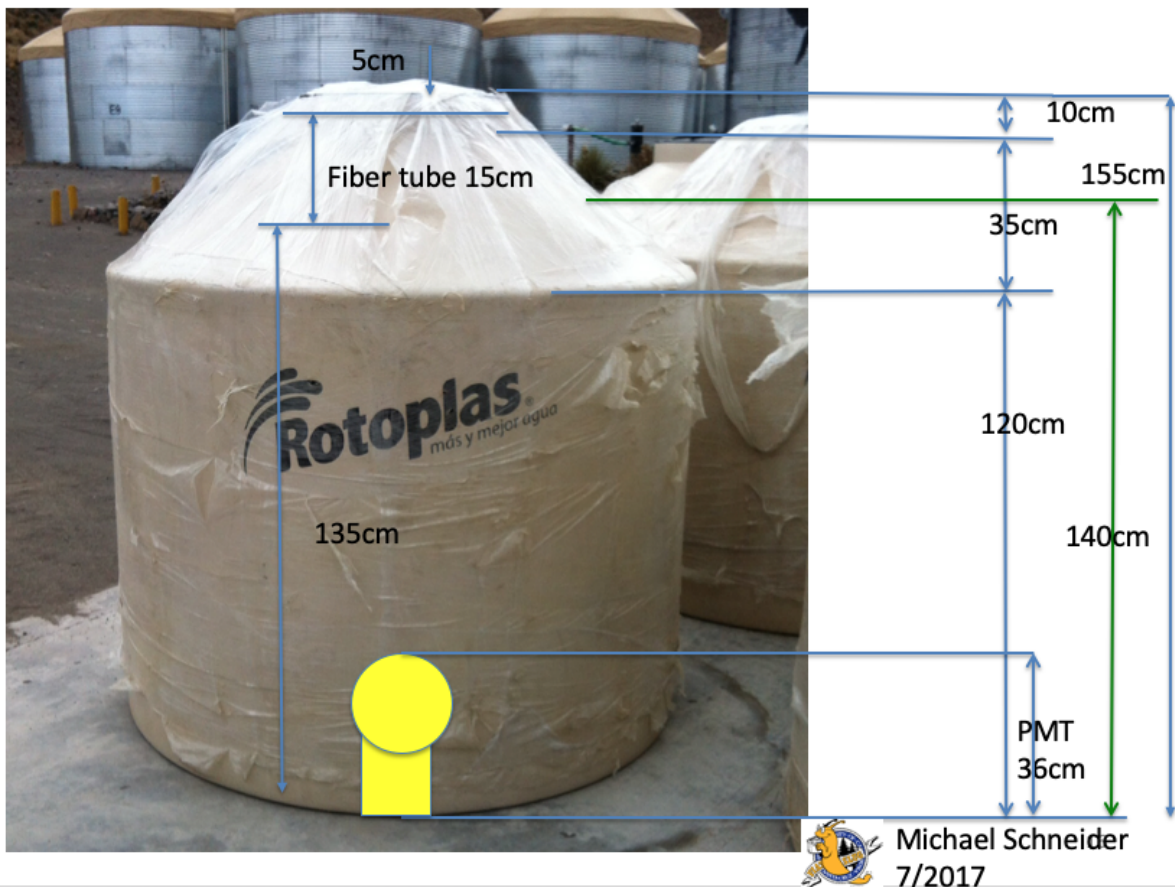


Figure 3.8: Schematic of a single outrigger tank showing the dimensions. The outriggers are smaller than the main tanks and contain only a single PMT, allowing them to cover more area at a lower cost than extending the main array. Photo credit to Michael Schneider.



## CHAPTER 4

### EVENT RECONSTRUCTION AND ANALYSIS

#### 4.1 Primary Reconstruction

##### 4.1.1 Core Fitting

The first step in reconstructing an air shower from HAWC data is to determine the core of the lateral charge distribution across the array. This core location is then used in many successive stages of the reconstruction. An arriving air shower is illustrated in Fig. 4.1, showing the shower core as well as features such as the arrival angle and curvature. A sample HAWC event showing charge of such an event distributed across the array is shown in Fig. 4.2, as well as the type of data that must be fit to get resulting reconstructed core.

An algorithm called Super Fast Core Fitter (SFCF) is currently used to find the shower core. The SFCF performs a fit for the core coordinates assuming a functional form for the effective charge distribution. This form is chosen such that it will have the same center as the NKG function (Eq. 3.1) expected to describe the distribution while avoiding numerical issues arising from the pole at the core position in the NKG function. The form chosen is:

$$\rho(r) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right) + \frac{N}{(0.5 + r/r_m)^3}, \quad (4.1)$$

where  $\sigma = 10\text{m}$ ,  $N = 5 \times 10^{-5}$ , and  $r_m = 124\text{m}$  is the Moliere radius at the HAWC elevation. This function parameterizes the charge distribution as a Gaussian close to the center represented by the first term, but with a longer tail component represented by the second term.  $\sigma$  therefore corresponds to the usual width of a Gaussian, while  $N$  is chosen to be a small number that ensures the tail component will be subdominant [33]. The sample HAWC event displayed in Fig. 4.2 shows the reconstructed core from the SFCF marked with a red star, as well as the lateral charge data used in the reconstruction.

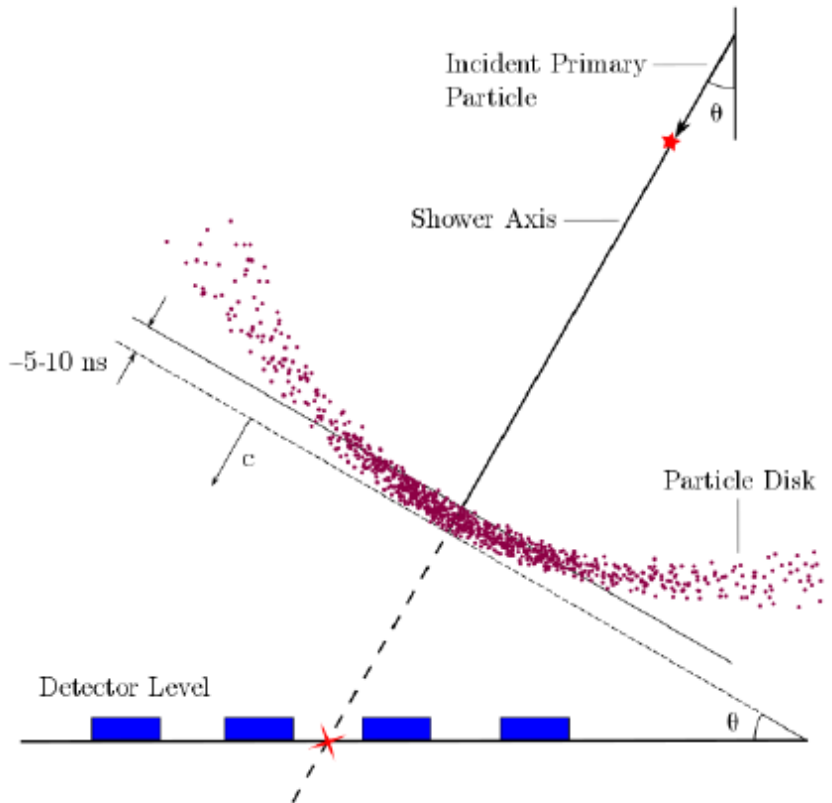


Figure 4.1: Illustration of an air shower arriving at HAWC with the arrival angle denoted by  $\theta$  and core location marked with a red cross. This figure illustrates how the edge of the shower is curved rather than flat due to secondary interactions, and how particle density decreases further from the central shower axis as predicted by the NKG function (Eq.3.1). Figure from Ref. [36].

#### 4.1.2 Angle Fitting

Once the core is found, the shower angle of arrival is fit. This angle of arrival corresponds to the arrival direction of the primary gamma ray, and therefore to the path back to the gamma-ray source. The angle is fit based on the relative timing of PMT hits. Before fitting the shower, it is necessary to correct for the non-flat nature of the shower plane. This curvature arises from both multiple-scattering of secondary particles far from the core, and timing delays due to lower available sampling at the edges and can be seen in the illustration in Fig. 4.1. Simulation is used to derive a correction for this curvature [33] as a function of

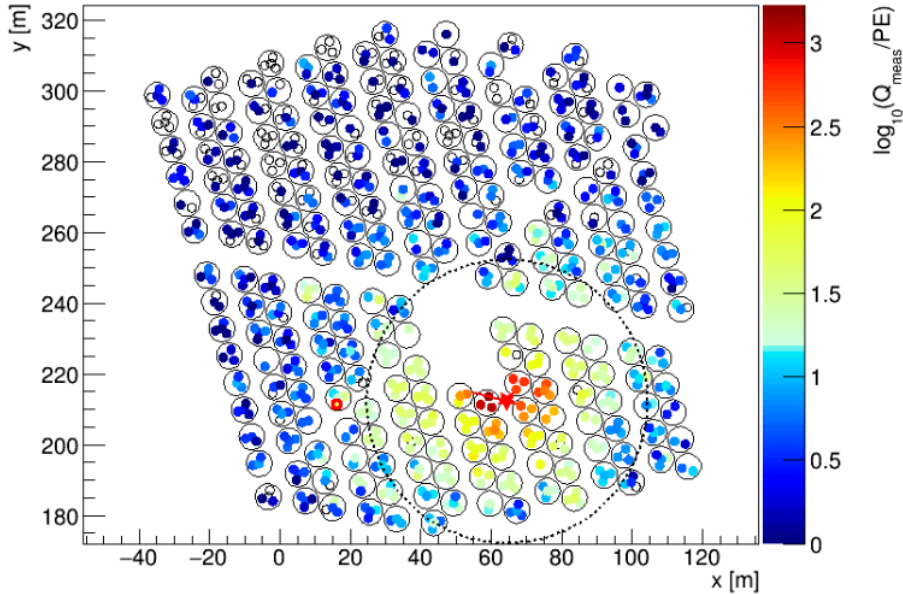


Figure 4.2: Plot of a HAWC event showing the lateral charge distribution across the array. The core reconstructed using the functional form in Eq.4.1 is marked with a red star. The details in the title refer to the reconstructed arrival coordinates and the HAWC DAQ run from which the event is displayed, as well as variables used to separate gamma-like and hadron-like events (see Sec. 4.3.1). Figure from Ref. [33].

shower parameters. Once the curvature is estimated, the arrival direction is fit using a  $\chi^2$  minimization as demonstrated in [39].

The angle fit procedure results in a shower arrival direction parameterized by its zenith and azimuth angles relative to the HAWC array. These local coordinates must then be converted into a celestial coordinate system in order to correlate the arrival direction with the location of gamma-ray sources on the celestial sphere. HAWC analyses primarily use the equatorial coordinate system, which is based on projecting the terrestrial longitude and latitude coordinates out into space and specifies points in terms of right ascension (RA) and declination (Dec). RA is defined as the angular distance eastward along the equator with respect to the position of the Sun during the Vernal Equinox, and is analogous to longitude. In this same vein, Dec is the angular distance from a point to the plane of the equator, analogous to latitude.

Bin	Lower Edge (% of PMTs hit)	Upper Edge (% of PMTs hit)
1	6.7	10.5
2	10.5	16.2
3	16.2	24.7
4	24.7	35.6
5	35.6	48.5
6	48.5	61.8
7	61.8	74.0
8	74.0	84.0
9	84.0	100.0

Table 4.1: Bin definitions for  $f_{\text{hit}}$  binning scheme.

## 4.2 Event Characterization

### 4.2.1 $f_{\text{hit}}$ Binning

HAWC events are primarily binned by  $f_{\text{hit}}$ , which is the ratio of PMTs participating in a shower event ( $n_{\text{hit}}$ ) to available PMTs in array ( $N_{\text{ChAvail}}$ ). Typically,  $N_{\text{ChAvail}}$  is around 1000 PMTs, with PMTs being occasionally removed for maintenance and re-calibration [33]. This binning aims to both estimate the primary shower energy (due to higher energy events tending to produce larger shower footprints which will then trigger more PMTs) and the angular resolution of the detector (due to the higher precision available to the angle fitter with more participating PMTs). The binning logic for  $f_{\text{hit}}$  is summarized in Table 4.1.

It should be noted that  $f_{\text{hit}}$  is imperfectly correlated with primary energy. Atmospheric attenuation effects will cause a shower to lose energy before reaching HAWC and have a smaller shower footprint and therefore lower  $f_{\text{hit}}$  value. This often results in high energy events being placed in lower  $f_{\text{hit}}$  bins and causes  $f_{\text{hit}}$  to have a broad energy resolution. This can be quantified by binning Monte Carlo simulations of gamma-ray sources where the true energy of events is known into  $f_{\text{hit}}$  space, with Fig. 4.3 showing such an example. It should be noted that the resulting true energy distributions in each bin are dependent on both the declination the modeled source transits at and the assumed shape of the energy spectrum.

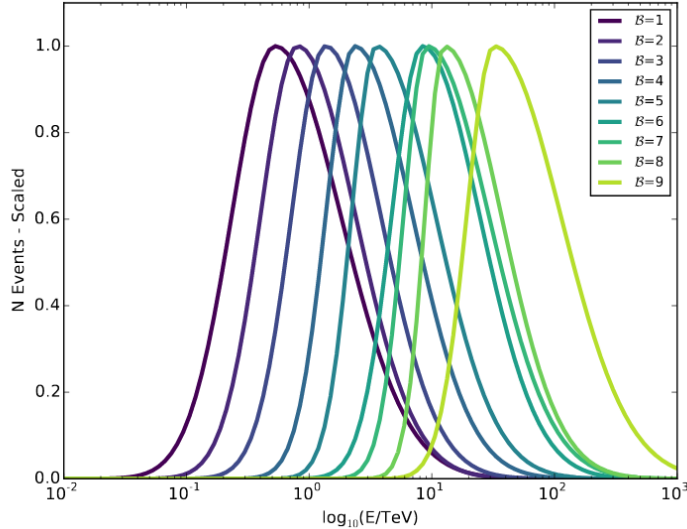


Figure 4.3: Scaled histogram of the fitted true energy distribution of each  $f_{\text{hit}}$  bin as shown in Table 4.1. Figure is taken from Fig. 3 of Ref. [33] and was generated from Monte Carlo simulations assuming a power law spectrum  $\frac{d\Phi}{dE} \sim E^{-2.66}$  for a single transit of a source at a declination of 20 degrees. All fitted distributions have been normalized to have unity amplitude for visualization purposes.

This is due to the fact that farther from overhead sources produce events at larger zenith angles (the angle between the shower arrival direction and the point directly overhead from the detector). Showers that arrive further from overhead of HAWC will be more attenuated by the atmosphere before reaching HAWC than overhead events, meaning it is possible for high energy showers to occur in low  $f_{\text{hit}}$  bins. Although  $f_{\text{hit}}$  does possess sufficient energy resolution for many HAWC analyses, more rigorous energy estimation techniques will be discussed in Sec. 4.2.3.

#### 4.2.2 Spatial Binning and Angular Resolution

Events are further categorized into spatial bins, which can be thought of as dividing the HAWC field of view into pixels of a fixed area that cover some range in RA and Dec coordinates. These bins are defined using the HEALPix protocol, with the pixel resolution characterized by the NSIDE parameter, where the total number of pixels dividing the sky

is equal to  $12 \times \text{NSIDE}^2$  and NSIDE is always equal to a power of two [40]. Most HAWC analyses use spatial pixel schemes with NSIDE=1024 for an approximate pixel width of 0.05 degrees, though analyses less concerned with precise source location will use lower NSIDE values for efficiency when fitting.

The spatial bin size is meant to be smaller than the HAWC angular resolution to avoid unnecessarily diluting the ability of HAWC to find a source’s position. HAWC’s angular resolution is characterized by the width of the point spread function, or PSF, which is the expected spatial distribution of events originating from a truly point-like source. The PSF is estimated by Monte Carlo simulations of point-like sources, where simulated events are reconstructed using the angle fitter and binned both spatially and in  $f_{\text{hit}}$ , analogously to the data. These distributions are then fitted assuming a sum of two Gaussians with separate widths and normalizations. Simulations have shown the PSF to be a function of  $f_{\text{hit}}$ , as expected since it is correlated to angular resolution, with the resulting 68% width being about one degree in the lowest bin and less than 0.2 degrees in the highest bin [33]. Moreover, the best-fit PSF is also a function of declination, with slightly better angular resolution for events overhead. This effect is accounted for by fitting the PSF in each  $f_{\text{hit}}$  bin at 5-degree declination intervals and interpolating the fit parameters for spatial bins that fall between them. This results in a continuous estimate of angular resolution that is valid for all spatial and analysis bins.

### 4.2.3 Energy Binning

Because of the shortcomings of using  $f_{\text{hit}}$  as an energy estimator, two techniques have recently been developed with the aim of more accurately estimating the shower energy. Both take into account the charge distribution across the array and the event zenith angle. The first, called the ground parameter (GP), is based on evaluating a modified NKG function to directly calculate the shower energy. Instead of the NKG function in Eq. 3.1, a modified form is used with an extra factor of  $1/r$  introduced so that the function corresponds to energy density

Bin	Lower Edge (TeV)	Upper Edge (TeV)
a	0.316	0.562
b	0.562	1.00
c	1.00	1.78
d	1.78	3.16
e	3.16	5.62
f	5.62	10.0
g	10.0	17.8
h	17.8	31.6
i	31.6	56.2
j	56.2	100
k	100	177
l	177	316

Table 4.2: Neural net energy binning labels and edges. Note that this is part of a 2D binning scheme combined with the  $f_{\text{hit}}$  bins shown in Table. 4.1.

rather than particle density [41]. This function is then fit using the PMT charge data and evaluated at 40 m from the shower center to yield an estimate of the primary energy. See Ref. [42] for further details. The second is based on a neural network machine learning algorithm (called NN from here on). This work will focus on the NN estimator with more details later in Section 4.5. Because  $f_{\text{hit}}$  estimates the angular resolution of the detector, events are binned both in energy and in  $f_{\text{hit}}$  when an analysis uses these energy estimators. The binning scheme for both estimators is shown in Table. 4.2. Note that analyses that do not require higher energy resolution do not use these energy estimators, and the work in Chapters 6 and 7 of this dissertation will use data binned only in  $f_{\text{hit}}$ .

## 4.3 Background Rejection

### 4.3.1 Gamma-Hadron Cuts

The majority of air showers detected by HAWC come from charged cosmic rays. Gamma-ray and cosmic-ray events can be separated by the different characteristics of the lateral charge distributions they create across the array. The primary difference is that the hadronic showers will produce muons through strong interactions, leading to a different shower profile

morphology. The lateral charge distribution of a cosmic-ray event is plotted side by side with a gamma-ray event for comparison in Fig. 4.4.

Two variables are used to characterize the relative likelihood that a shower primary was a hadron. The first is compactness ( $C$ ), defined as follows:

$$C = n_{\text{hit}}/\text{CxPE}_{40}, \quad (4.2)$$

where  $n_{\text{hit}}$  is the number of PMTs that participated in the event, and  $\text{CxPE}_{40}$  is the maximum effective charge (PMT-recorded charge with a correction factor to account for the different responses of the 8-inch and 10-inch PMTs) observed in a PMT at least 40 meters away from the shower core. This quantity aims to detect the presence of muons in the shower, which will cause a large response in a PMT even if it is far from the shower core and most of the secondary particles. The smaller the compactness, the more likely an event is to be hadronic in nature [33] [29].

The second variable used aims to characterize the axial symmetry of the shower footprint and is referred to as PINCness (short for Parameter for Identifying Nuclear Cosmic-rays, and referred to from here on as  $P$ ).  $P$  is defined as follows:

$$P = \frac{1}{N} \sum_{i=1}^N \frac{(\log Q_i - \langle \log Q_i \rangle)}{\sigma_i}, \quad (4.3)$$

where  $Q_i$  is the effective charge measured by the  $i$ th PMT,  $N$  is the number of participating PMTs, and  $\sigma_i$  is the error in the logarithm of  $Q_i$  (determined from a study of the gamma ray-rich Crab Nebula region). The average associated with each  $Q_i$  is determined from the mean charge in PMTs within a 5 meter wide annulus containing the  $i$ th PMT, centered on the shower core [33].

Cuts on  $C$  and  $P$  are used to filter out hadron-like events from the data. The values of these cuts are chosen to maximize the expected significance of gamma-ray spectra as modeled in Monte Carlo simulations. This increase in significance is estimated by the Q-factor:

$$Q = \frac{S}{\sqrt{B}}, \quad (4.4)$$



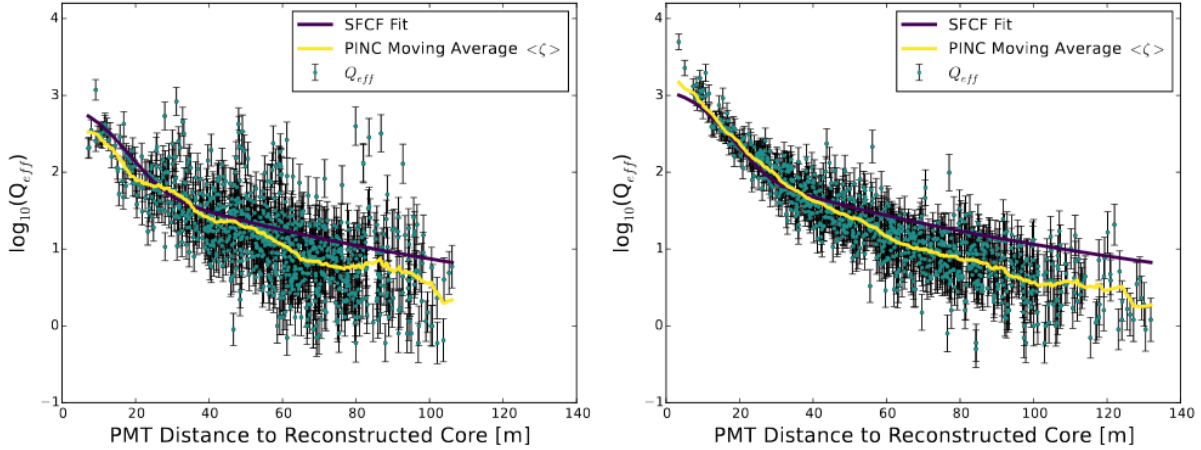


Figure 4.4: Comparison of a likely cosmic-ray (left) and likely gamma-ray (right) lateral charge distributions with the SFCF (Eq. 4.1) functional form fitted. The gamma-ray shower has one distinct core and falls off fairly symmetrically about its central axis. In contrast, the cosmic-ray shower contains multiple smaller peaks of charge far from the main core. The presence of charge peaks far from the shower core is quantified by compactness in Eq. 4.2, while the axial symmetry can be seen in the PINCness (Eq. 4.3) moving average, with large deviations in the cosmic-ray shower indicating an asymmetric charge distribution. Figure from Ref. [33].

where  $S$  is the number of gamma rays expected to pass the cuts, and  $B$  is the expected number of hadronic background events, and the cuts resulting in the maximum Q-factor are considered the optimum [33]. The number of signal events  $S$  is found by applying the cuts to simulated gamma-ray events. However, hadronic showers are known to have poor agreement between observed data and Monte Carlo samples in the distributions of the variables  $C$  and  $P$ . This makes it difficult to determine the number of hadronic events expected to pass from Monte Carlo. Instead,  $B$  is estimated by applying the cuts to actual HAWC data. Due to the very high relative abundance of charged cosmic rays to gamma rays (up to a factor of  $10^4$  depending on energy), the remaining data after cuts is still dominated by hadronic events and approximately equals the expected background.

### 4.3.2 Direct Integration

With optimized gamma-hadron cuts, 99% of hadronic events are rejected at the highest energies [33]. However, as stated above, due to the high relative abundance of charged cosmic rays over gamma rays, the data after cuts is still dominated by background, especially at the lower energies. To estimate the remaining background after cuts, a technique called direct integration is used. This approach relies on the fact that cosmic ray arrival directions are approximately isotropic due to the effects of interstellar magnetic fields (Chapter 3). Because of this, and the much higher relative abundance of cosmic rays to gamma rays, an average of the HAWC event rate over time in bins of comparable sensitivity is approximately equal to the expected cosmic-ray background. Specifically, the background in a given spatial bin is estimated by averaging the rate of events within a chosen time interval, assuming the detector is stable (see Ref. [43] for further details). For most HAWC analyses, events are integrated over two-hour intervals to estimate the associated background [33].

This integration process is performed separately in each  $f_{\text{hit}}$  bin, as well as for each energy bin if the energy estimators are considered. This results in an individual estimate of the expected background in all analysis bins. Due to the relative sparseness of both the higher  $f_{\text{hit}}$  bins and spatial bins at the edges of the field of view, the background from direct integration is not always sufficiently smooth. To address this, an additional smoothing is performed by averaging the background over a 0.5 degree radius [33].

### 4.3.3 Background with the Energy Estimators

The 2D binning scheme used for the energy analysis (Table 4.2) reduces the statistical content of each bin. In particular, the highest energy bins are often very sparsely populated. Because of this, statistics are often insufficient for the background to be estimated via direct integration. Instead, a technique called background randomization is used for any bins with a low number of events. This technique allows the background to be estimated by averaging over the entire data set in a given bin and, increasing the statistics available and yielding a

smoother background than direct integration for the sparsely-populated bins.

For each bin, a two-dimensional histogram of hour angle (the local time coordinate analog to right ascension) and declination is created using the entire data set. This distribution is then sampled 10000 times, with each resulting pair being combined with the time from one randomly chosen event from the data set, and the pixel in which this event would have fallen is incremented. The background map is then scaled to have the appropriate normalization to match the real data. See Refs. [42] and [41] for additional details. As with direct integration, this background estimate is again smoothed by averaging over 0.5 degrees.

#### 4.4 The Neural Network Energy Estimator

This section gives a more detailed discussion of the neural network (NN) energy estimator. The NN approach uses a multi-layer perceptron network defined using the Toolkit for Multivariate Analysis (TMVA) package available in the ROOT software library [44]. The goal is to create a map between the observable shower variables and gamma-ray primary energy. The NN is trained using reconstructed Monte Carlo simulations in which the true energy is known. The optimal weights  $w_i$  are found by minimizing the error function defined as:

$$D = \frac{1}{2} \sum_{i=1}^{N_{\text{events}}} u_i [\log_{10} \hat{E}(x_i, w_i) - \log_{10} E_i]^2, \quad (4.5)$$

where  $\hat{E}$  is the reconstructed energy as a function of the NN input variable  $x_i$  and weights  $w_i$ ,  $E_i$  is the MC-true energy, and  $u_i$  is the relative importance of each training event. The observable input variables used for the NN can be broken up into three distinct categories: multiplicity, containment, and atmospheric attenuation. All input variables are listed in Table. 4.4.

The multiplicity category refers to the magnitude of the signal measured by the detector. The  $f_{\text{hit}}$  variable used in the main binning scheme is the most prominent variable here, and is defined the same as in the 1-D binning scheme (Table 4.1). In addition to  $f_{\text{hit}}$ , the fraction of tanks with at least one triggered PMT in the event,  $f_{\text{Tank}}$ , is used. Including this second

Variable	Definition	Category
$f_{\text{hit}}$	$n_{\text{hit}}/N_{\text{ChTot}}$	Multiplicity
$f_{\text{Tank}}$	Fraction of tanks hit	Multiplicity
$\log(\text{CoreAmp})$	Amplitude of function used in core fit	Multiplicity
$r_{\text{core}}$	Distance between shower core and detector center	Containment
$\theta$	Zenith Angle	Attenuation
Annulus 1	Fraction charge > 10 m from shower axis	Attenuation
Annulus 2	Fraction charge > 10 m and < 20 m from shower axis	Attenuation
Annulus 3	Fraction charge > 20 m and < 30 m from shower axis	Attenuation
Annulus 4	Fraction charge > 30 m and < 40 m from shower axis	Attenuation
Annulus 5	Fraction charge > 40 m and < 50 m from shower axis	Attenuation
Annulus 6	Fraction charge > 50 m and < 60 m from shower axis	Attenuation
Annulus 7	Fraction charge > 60 m and < 70 m from shower axis	Attenuation
Annulus 8	Fraction charge > 70 m and < 80 m from shower axis	Attenuation
Annulus 9	Fraction charge > 80 m from shower axis	Attenuation

Table 4.3: Variables and definitions used in the NN energy estimator.

multiplicity estimator was found to improve performance during initial development of the NN [36]. Finally, the amplitude of the SFCF core fit (Eq. 4.1) is in principle a more rigorous estimate of the multiplicity than  $f_{\text{hit}}$ , since it is less sensitive to the shower containment within the array boundaries. However, this is actually one of the least important variables, only having a true impact on the highest-energy events, which are more poorly contained by the array [36].

Containment refers to how much of the event was captured by the array. This effect is measured by the distance  $r_{\text{core}}$  between the array center and the shower core. Showers with cores closer to the center of the array will naturally be better contained and trigger more of the PMTs, whereas a shower of the same size with a core close to an array edge will not trigger nearly the same number of PMTs despite having the same footprint size. Adding  $r_{\text{core}}$  as an input variable accounts for this effect [36].

The last class of variables account for the atmospheric attenuation of the shower. The shower zenith angle  $\theta$  is used since showers that trigger at larger zenith angles must traverse through a larger atmospheric distance and will suffer greater attenuation than showers orig-

inating from overhead. The remaining nine variables consist of the fraction of total charge distributed across the array in subsequent annuli. These variables aim to estimate the shower age (Eq. 3.2) by parameterizing the lateral distribution of charge along the shower axis. The annuli are defined in the reconstructed shower plane rather than on the detector plane to more directly measure the true lateral charge distribution, rather than the one projected onto the ground. Each of the first eight annuli extends for ten meters, starting at the shower axis, while the ninth annulus contains all the charge deposited ninety meters or greater from the axis.

Two quantities are used to evaluate the performance of the energy estimators. The bias, defined as:

$$\text{bias} = \langle \log_{10} \hat{E} - \log_{10} E \rangle, \quad (4.6)$$

and the resolution defined as:

$$\text{res} = \sqrt{\langle (\log_{10} \hat{E} - \langle \log_{10} \hat{E} \rangle)^2 \rangle}, \quad (4.7)$$

where the average in both cases runs over all events in a given  $\hat{E}$  bin. If added in quadrature, the bias and resolution are equal to the error function minimized by NN optimization (up to a factor of 1/2). The bias and resolution are plotted for both the NN and GP estimators in Fig. 4.5. Note that both energy estimators suffer from large bias at the lowest energies. This is due to a trigger effect arising from imposing a minimum threshold of triggered PMTs in order for an event to be binned in  $\hat{E}$  space. At the lowest energies, only events that trigger a larger than usual fraction of PMTs are included, which are disproportionately likely to be binned at higher than true energy [36].

## 4.5 Neural Network Re-optimization

HAWC recently completed a reprocessing of archival data with an improved class of algorithms. Among these is a new core fitter and an improved protocol for removal of noise hits from shower events. This effort also included a re-optimization of the neural network energy estimator. Although the resulting data from this reconstruction effort was unavailable at the

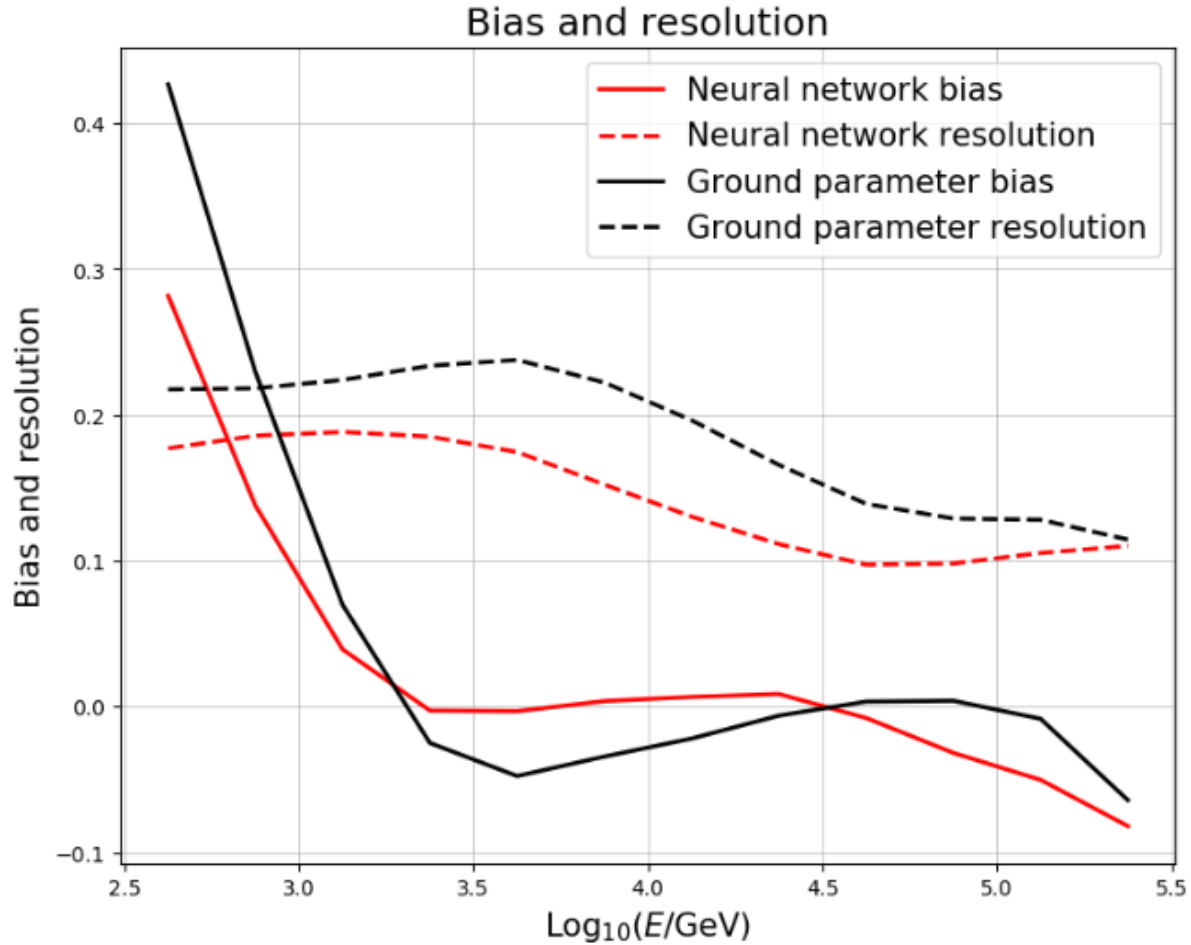


Figure 4.5: Bias and resolution of the neural net (NN) and ground parameter (GP) energy estimators evaluated on current Monte Carlo data. Figure taken from Ref. [41].

time the analyses in this dissertation were completed, these improvements will contribute to future work with HAWC.

Because the new core fitter is not attached to a functional form of the lateral charge distribution, it is not able to return an associated amplitude and the  $\log(\text{CoreAmp})$  variable is no longer available. Although previously found to be only weakly correlated with energy, the amplitude is necessary to achieve optimal performance at the highest energies. Instead, the amplitude of the lateral charge distribution fit from the GP energy estimator is used. This was found to yield comparable performance to the old core fit amplitude.

The other major change was the training sample used. The spectrum of the Monte

Carlo sample falls off with the square of the gamma-ray energy and therefore consists of far more events at lower energies. As discussed above, the lowest energies suffer from large positive bias due to the trigger threshold in  $f_{hit}$  space. This assigns disproportionately higher importance to these events, forcing a negative bias at the higher energies, where many of the analyses that use the NN are focused. One theoretical way to counteract this would be to train two NNs; one for low energies and one for high energies. However, this is impossible to implement in practice because it would require knowing the true energy before selecting which NN to use for evaluating the event. Instead, the spectrum is re-weighted to remove the importance placed on the low-energy events. Specifically, when minimizing the error function,  $u_i = E_i^2$  is used (see Eq. 4.5). This effectively makes the training sample uniform in energy space and removes the negative bias at the intermediate and high energies at the cost of slightly more positive bias at the low energies. See Fig. 4.6 for a plot comparing the neural net performance with and without this re-optimization.

## 4.6 Forward Folding and Likelihood Analysis

HAWC uses a forward folding approach to reconstruct the true energy spectrum of a source from the binned data. Rather than attempting to deconvolve the various reconstruction effects from the observed data, simulation is used to map a given energy spectrum to an expected detector response. These detector response histograms are then used to transform assumed true energy spectra into reconstructed spectra in either 1D ( $f_{hit}$ ) only or 2D ( $f_{hit}$  and  $\hat{E}$ ) space, depending on the needs of a given analysis. These transformed spectra are also spatially binned into the same scheme as used by the data, with the PSF used to compute the expected number of events in each spatial bin considering HAWC's angular resolution.

Fits are performed using a maximum likelihood approach. The likelihood in each bin is assumed to be Poisson-distributed. Therefore, the likelihood for a given assumed signal is:

$$\log(L) = \sum_{n_{bins}} \log(L_i), \quad (4.8)$$

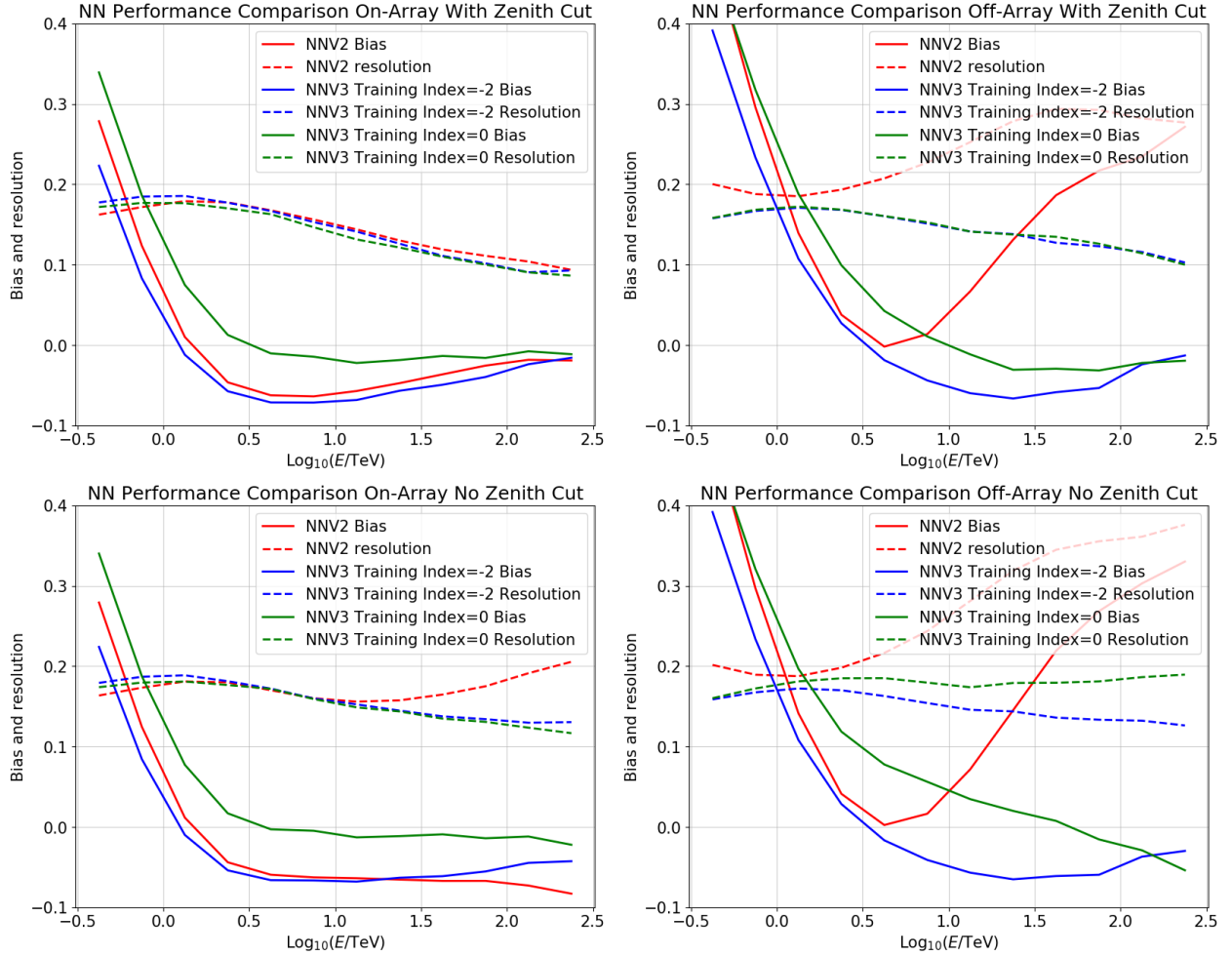


Figure 4.6: Bias and resolution of the newly re-optimized neural network (NNV3) trained on the most recent HAWC Monte Carlo, compared with the performance of the old neural network (NNV2) trained on older Monte Carlo. The data is plotted for events on (left) and off (right) the array, and with (top) and without (bottom) a zenith cut of 45 degrees. In all cases, the re-optimized (NNV3) neural network trained on a flat energy spectrum performs the best at all energy scales above 1 TeV.



$$L_i = \frac{(S_i + B_i)_i^N e^{-(S_i+B_i)}}{N_i!}, \quad (4.9)$$

where  $N_i$ ,  $B_i$ , and  $S_i$  are respectively the observed counts, estimated background, and expected signal in the  $i$ th bin for a given set of fit parameters. Note that the index here runs over spatial bins,  $f_{\text{hit}}$  bins, and  $\hat{E}$  bins (if the energy estimators are being used).

The best fit spectrum for a given source is found by varying the free parameters in the model and recomputing the resulting  $S_i$  counts until  $L$  is maximized. The statistical significance is then computed using the test statistic (TS) defined as:

$$TS = 2 \log (L_{max}/L_0), \quad (4.10)$$

where  $L_{max}$  is the maximum likelihood resulting from the fit, and  $L_0$  is the likelihood for a background-only model (where  $S_i = 0$  in all bins).

Since the background-only hypothesis is contained entirely within the signal plus background hypothesis, Wilks' theorem can be used to interpret the TS. Wilks' theorem shows that if the background-only hypothesis is true, in the high-statistics limit the TS will follow a  $\chi^2$  distribution with degrees of freedom equal to the difference in the degrees of freedom between the signal-hypothesis and background-only hypothesis [45]. Therefore, the appropriate  $\chi^2$  distribution for a given signal hypothesis can be used to transform the TS into a measure of significance in which the background-only model is rejected.

A likelihood approach is also used to derive lower and upper limits of a given confidence on fit parameters. This is done by finding:

$$\Delta LL = \log(L_{limit}) - \log(L_{max}) \quad (4.11)$$

such that  $\Delta LL$  corresponds to a given confidence interval, based on the appropriate  $\chi^2$  distribution. This likelihood-based approach is the one used for all subsequent analyses in this dissertation.

## CHAPTER 5

### LIV: AN APPLICATION OF THE ENERGY ESTIMATORS

#### 5.1 Introduction

This chapter uses the improved energy resolution available from the NN energy estimator as discussed in Chapter 4 to search for Lorentz Invariance Violation (LIV). Lorentz invariance is the symmetry underlying Einstein's theory of relativity. It imposes that all reference frames should obey the same physical laws regardless of their relative velocities. All observations up to the present day have confirmed Lorentz invariance is an exact symmetry and that the theory of relativity holds. This analysis will probe for LIV at very high energy scales and attempt to confirm if relativity holds even in these extreme conditions. The results of this analysis were published in Physical Review Letters and can be read in full at [46].

#### 5.2 Lorentz Invariance Violation

The Lorentz-invariant dispersion relation for a particle of energy  $E$ , momentum  $p$  and mass  $m$  is given by:

$$E^2 - (pc)^2 = (mc)^2. \quad (5.1)$$

Moving to units where  $c = \hbar = 1$  (used for the duration of this dissertation), and considering the case where  $m = 0$ , the dispersion relation for a photon becomes:

$$E = p. \quad (5.2)$$

If Lorentz invariance is violated at some high energy scale, then Eq. 5.1 will include additional terms. These terms are best expressed as a power series of the momentum  $p$  with some set of coefficients. For the case of a photon, the modified dispersion relation can be expressed as:

$$E_\gamma^2 - p_\gamma^2 = \pm |\alpha_n| p_\gamma^{n+2}, \quad (5.3)$$

where  $\alpha_n$  is the lowest order non-zero coefficient of the power series in  $p$ .

In the case where the sign of the LIV term is negative, it becomes possible for photons to decay via pair-production [47] [48]. Over large astrophysical distances, the decay probability approaches one, and above some energy no astrophysical photons will be observed. Therefore, simply observing photons above a given energy scale at high-confidence is sufficient to set upper limits on the LIV coefficients  $\alpha_n$ .

### 5.3 Derivation of Upper Limits

The upper limits on  $\alpha_n$  corresponding to a given maximum photon energy are derived as follows. First, consider the modified dispersion relation rewritten with the photon momentum as  $k$  instead of  $p_\gamma$  to more easily distinguish it from other momenta that will enter into the calculation:

$$E_\gamma^2 = k^2(1 + \alpha_n k^n). \quad (5.4)$$

Using this modified relation, consider the decay of a photon into an electron-positron pair. By conservation of energy, and letting  $p$  represent the momentum of the positron produced in the decay,  $\theta$  be the half-angle between the electron-positron pair, and  $m_e$  be the electron mass, Eq. 5.4 becomes:

$$k\sqrt{1 + \alpha_n k^n} = \sqrt{m_e^2 + p^2} + \sqrt{m_e^2 + |\mathbf{k} - \mathbf{p}|^2} \quad (5.5)$$

$$= \sqrt{m_e^2 + p^2} + \sqrt{m_e^2 + k^2 + p^2 - 2kp\cos(\theta)}. \quad (5.6)$$

This equation can be solved for  $p$  by squaring both sides and rearranging into a form to which we can apply the quadratic formula. The two admitted solutions are:

$$p = \frac{\alpha_n k^{n+1} \cos(\theta) \pm \sqrt{\alpha_n^2 k^{2n+1} \cos^2(\theta) - 4(\sin^2(\theta) + \alpha_n k^n)(1 + \alpha_n k^n)m_e^2}}{2\alpha_n k^n + 2\sin^2(\theta)}. \quad (5.7)$$

In order for the decay to be allowed,  $p$  must be a real number. The minimum value of  $\alpha_n$  needed to guarantee a positive discriminant and ensure a real-valued  $p$  is therefore:

$$\alpha_n \geq \frac{4m_e^2}{k^n(k^n - 4m_e^2)}. \quad (5.8)$$

Since the LIV coefficients  $\alpha_n$  are expected to be small,  $k \approx E_\gamma$ , Eq. 5.8 implies that observing a photon with energy greater than or equal to some maximum observed energy  $E_c$  yields a corresponding upper limit on  $\alpha_n$  of at least:

$$\alpha_n \geq \frac{4m_e^2}{E_c^n(E_c^n - 4m_e^2)}. \quad (5.9)$$

Limits on additional physical parameters also follow from Eq. 5.9. For example, many theories of LIV relate the  $\alpha_n$  coefficients to the quantum gravity energy  $E_{LIV}$  by:

$$E_{LIV} = \alpha^{-1/n}, \quad (5.10)$$

where  $E_{LIV}$  is the energy at which quantum gravity effects are expected to dominate, and should be on the order of the Planck scale ( $10^{-28}$  eV) [47]. Note that the corresponding constraint on  $E_{LIV}$  is dependent on the expansion order of the  $\alpha$  coefficient measured, which is denoted below using superscript notation. Putting together the above equations leads to the following relations between  $E_{LIV}$  and the maximum observable astrophysical gamma-ray energy  $E_c$ :

$$E_{LIV}^{(1)} \gtrsim 9.57 \times 10^{23} \text{eV} \left( \frac{E_c}{\text{TeV}} \right)^3, \quad (5.11)$$

$$E_{LIV}^{(2)} \gtrsim 9.78 \times 10^{17} \text{eV} \left( \frac{E_c}{\text{TeV}} \right)^2. \quad (5.12)$$

These equations can therefore be used to relate an observed universal hard cutoff energy directly to the LIV parameters.

A second LIV decay process was also considered:  $\gamma \rightarrow N\gamma$ , known as photon-splitting. The dominant splitting process is the photon decay into three photons ( $3\gamma$ ) [48], and the derivation of limits from this phenomenon is shown here. If photon-splitting is permitted, the decay rate for the  $3\gamma$  process is:

$$\Gamma_{\gamma \rightarrow 3\gamma} = 5 \times 10^{-14} \frac{E_\gamma^{19}}{m_e^8 E_{LIV}^{(2)10}}, \quad (5.13)$$

as shown in Ref [48]. While the decay rate is nominally smaller than that of photon decay, this process is kinematically allowed whenever  $E_\gamma^2 > p_\gamma^2$ . It becomes significant when

photons propagate through cosmological distances and also predicts a cutoff at the highest energy part of the photon spectra of astrophysical sources. Because these photons originate from distant sources, the mean free path of a photon is equal to the distance,  $L$ , between the source and observer, such that  $L \Gamma = 1$ , with  $\Gamma$  translated to units of  $kpc^{-1}$ . The corresponding LIV limit, as a function of the highest photon energy, is given by,

$$E_{LIV}^{(2)} > 3.33 \times 10^{19} \text{eV} \left( \frac{L}{\text{kpc}} \right)^{0.1} \left( \frac{E_c}{\text{TeV}} \right)^{1.9}, \quad (5.14)$$

where  $E_c$  is again the maximum photon energy observed. Therefore limits on  $E_{LIV}$  are obtained by observing a true maximum gamma-ray energy associated with an astrophysical source.

## 5.4 Searching for LIV

### 5.4.1 High Energy Sources

To search for LIV, the four highest-energy sources in the HAWC catalog published in Ref. [49] are considered. These sources all have been shown to produce photons above 100 TeV. Due to its high significance at all energies, the Crab Nebula is also included. As established above, certain forms of LIV are expected to cause a hard cutoff in the photon energy spectrum. We check the source spectra for evidence of such a cutoff within the HAWC energy range. If no cutoff is found, detecting photons above a given energy inherently constrains the energy scale at which LIV occurs.

The analysis that initially established these sources as emitting high energy photons was performed using the ground parameter energy estimator [49]. Due to its better expected resolution at high energies, the neural network energy is used instead to search for an LIV cutoff. Because of known systematic differences between the estimators, the source location may be different in the neural network data used here.

To check this, a maximum-likelihood fit of the position analogous to Ref. [49] is performed.

Source	RA (deg)	Dec (deg)
eHWC J1825-134	276.41	-13.45
eHWC J1907+063	286.96	6.25
eHWC J0534+220 (Crab)	83.60	22.05
eHWC J2019+368	304.94	36.74

Table 5.1: Fitted locations of HAWC sources emitting above 56 TeV and used to constrain LIV.

To do so, the differential energy spectrum  $\frac{d\Phi}{dE}$  is fixed to a power law form:

$$\frac{d\Phi}{dE} = \Phi_0(E/E_0)^{-\gamma}, \quad (5.15)$$

where the index  $\gamma$  is fixed to 2.0, the pivot energy  $E_0$  is fixed to 10 TeV, and the flux normalization  $\Phi_0$ , RA, and Dec of the source are free. The center of the source location is considered to be the coordinates from the best fit. In addition, due to the potential for multiple sources to be clustered together, the high-energy component is isolated by considering only those bins with reconstructed energy  $\hat{E} > 56$  TeV. The resulting source locations are summarized in Table 5.1. Fig. 5.1 shows significance maps made with the 2.0 index in the vicinity of source locations for all bins above 56 TeV.

#### 5.4.2 Limit Technique

Because of the finite HAWC energy resolution, the hard cutoff predicted by LIV will not appear as a true hard cutoff in the observed data. Some events will reconstruct to higher or lower bins than correspond to their true energies. Because astrophysical spectra tend to fall off sharply as a function of energy, the highest energy bins will have a substantial content of photons from lower energies. This phenomenon, known as bin-migration, can be accounted for as part of the forward folding approach as described in Section 4.6. Each source’s known spectral functional form is modified to include a hard cutoff at some energy scale. This adds the maximum gamma-ray energy,  $E_c$ , as a fit parameter. The fits are then redone with this additional degree of freedom and the maximum-likelihood best-fit is found.

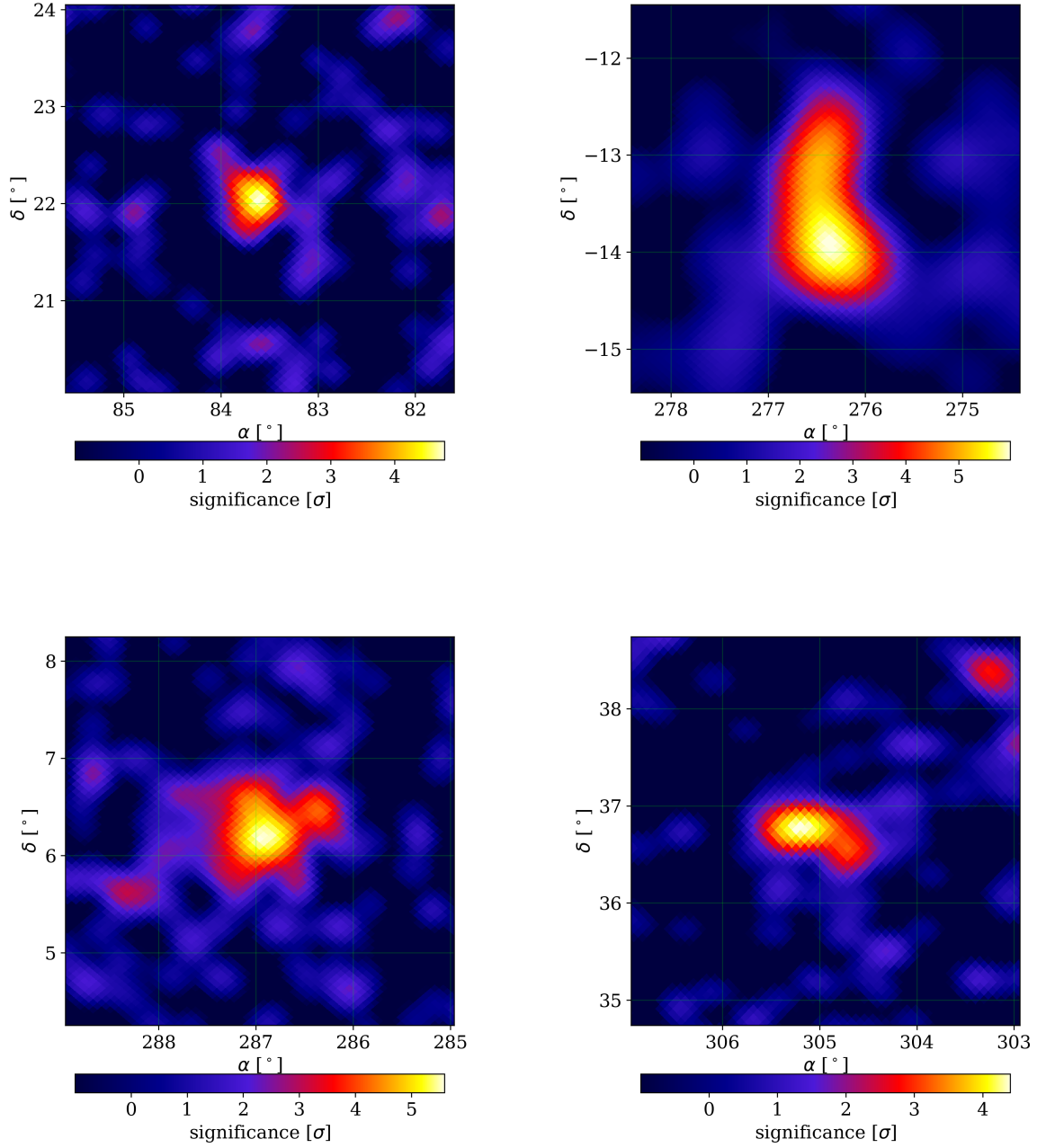


Figure 5.1: Significance maps in the vicinity of the four high-energy sources for bins above 56 TeV. These maps were made using the likelihood-derived significance following a fitted power law of index 2.0.

The spectra are modeled either using an exponentially cut-off power law as in Eq. 5.16 (for eHWC J1825-134) or as the log parabola in Eq. 5.17 (for all other sources). The spectral form for the cut-off power law is:

$$\frac{d\Phi}{dE} = \Phi_0(E/E_0)^{-\gamma} e^{-E/X}, \quad (5.16)$$

where  $\Phi_0$  and  $\gamma$  are again normalization and index with additional parameter  $X$  as the exponential cutoff energy. The spectral form for a log parabola is:

$$\frac{d\Phi}{dE} = \Phi_0(E/E_0)^{-\alpha-\beta \ln(E/E_0)}, \quad (5.17)$$

where the  $\alpha$  and  $\beta$  control the inflection of the log parabola and  $\Phi_0$  is the normalization. In these fits, the pivot energy  $E_0$  is set to 7 TeV for the Crab, and 10 TeV for all other sources. The remaining spectral parameters are free and optimized using a maximum likelihood method. To check for LIV, the fits are re-done with the assumed source spectra including an additional floating hard cutoff energy,  $E_c$ .

These spectral forms were found to maximize the likelihood for their corresponding sources in Ref. [49], and were also found to be the best fitting in the neural-net data. Each source is first fit with all parameters free to check if a finite hard cutoff is preferred. If no hard cutoff is found, lower limits will be set on the cutoff energy  $E_c$  from each source.

This analysis uses a profile-likelihood method to determine preference for a finite hard cutoff as well as to set lower limits. In this approach, all spectral parameters ( $\alpha$ ,  $\beta$ ,  $X$ , etc.) besides  $E_c$  are treated as nuisance parameters. Let  $\mathbf{P}$  denote the vector consisting of all nuisance parameters for a given spectrum. The profile likelihood  $L_p$  is defined as:

$$L_p(E'_c) = \max_{\mathbf{P}} L(\mathbf{P}', E'_c), \quad (5.18)$$

meaning the maximum likelihood with all nuisance parameters set free and  $E_c$  fixed to a given value.

Using profile likelihood the quantity  $D$  is defined as :

$$D(E_c) = 2 \frac{\max_{E_c, \mathbf{P}} L}{L_p(E_c)}, \quad (5.19)$$



where the numerator is the likelihood maximized for all parameters (including  $E_c$ ) set free.

Based on Wilks' theorem [45],  $D(E_c)$  will follow a  $\chi^2$  distribution with one degree of freedom. Therefore,  $D(\infty)$  is interpreted as a measure of statistical preference for a finite hard cutoff over no (i.e. an infinite) hard cutoff. Lower limits on  $E_c$  are set by finding  $D(E_c)$  such that a given confidence interval is reached based on this expected  $\chi$ -squared distribution. As explained by Chernoff in Ref. [50], because the null case is at the edge of a parameter space ( $E_c \rightarrow \infty$ ), the likelihood is inherently one-sided. Therefore, the thresholds for 95% and  $3\sigma$  confidence bounds are respectively 2.71 and 7.71.

Since LIV is expected to occur at the same energy scale regardless of the source, and the photon decay energy  $E_c$  depends only on the LIV parameters, a combined likelihood analysis can be performed using all four sources. In this approach, the quantity  $D(E_c)$  is summed together over all sources and the significance and limits computed using this summed profile likelihood.

### 5.4.3 Fitting Details

Naively adding a hard cutoff as a free parameter has the effect of making the profile likelihood function piece-wise continuous. This is due to the nature of the forward folding algorithm. Each data bin has a true energy histogram associated with it, and these histograms themselves have a finite bin width. Because the hard cutoff is an inherently discrete parameter, this produces a discontinuous likelihood when convolved with the discrete true energy histogram. One possible fix to this issue is to increase the bin size of the true energy histograms, but this was found to result in poor fit convergence. Instead, the hard cutoff was modified into a slightly softened cutoff by convolving it with a normal distribution with of standard deviation of 0.1 in  $\log_{10}(E)$  space. This makes the hard cutoff sufficiently continuous to properly convolve with the energy resolution without becoming degenerate with the other fit parameters. See Fig. 5.2 for an illustration of the energy-resolution convolution's effect on the hard cutoff. As illustrated in Fig. 5.3, a spectrum softened cutoff is still distinguishable

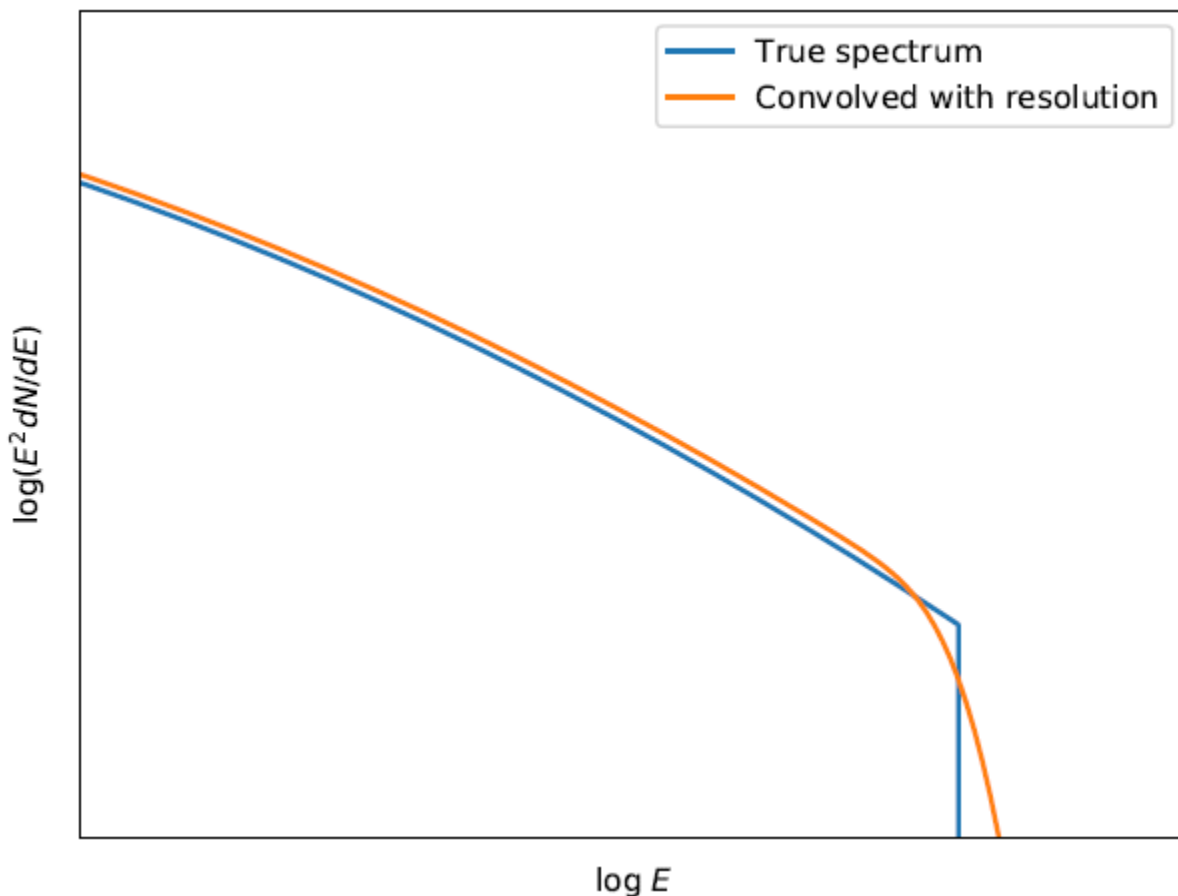


Figure 5.2: Effect of softening the hard cutoff by both the HAWC energy resolution and the additional smoothing required to avoid a discontinuous likelihood function. For illustration, the cutoffs are superimposed on a power law spectrum with an index of -2.7, and arbitrary flux and energy units. Figure from [36].

from the unmodified spectra even given statistical uncertainties on the fits.

In addition, to desensitize the fits to known systematic errors in the modeling of the HAWC point spread function (PSF), the standard HAWC spatial binning approach is not used. Instead of convolving the source morphologies with the HAWC PSF and summing the likelihood in each spatial bin, the total data, background, and expected signal are summed within a given angular radius centered on the source location. The optimal radius for sensitivity to LIV is expected to be reached when all of the highest energy photons from a source are included in the fit. The limit will then grow less constraining for successively larger radii as more background is added to the fit without any new high energy photons. Therefore,

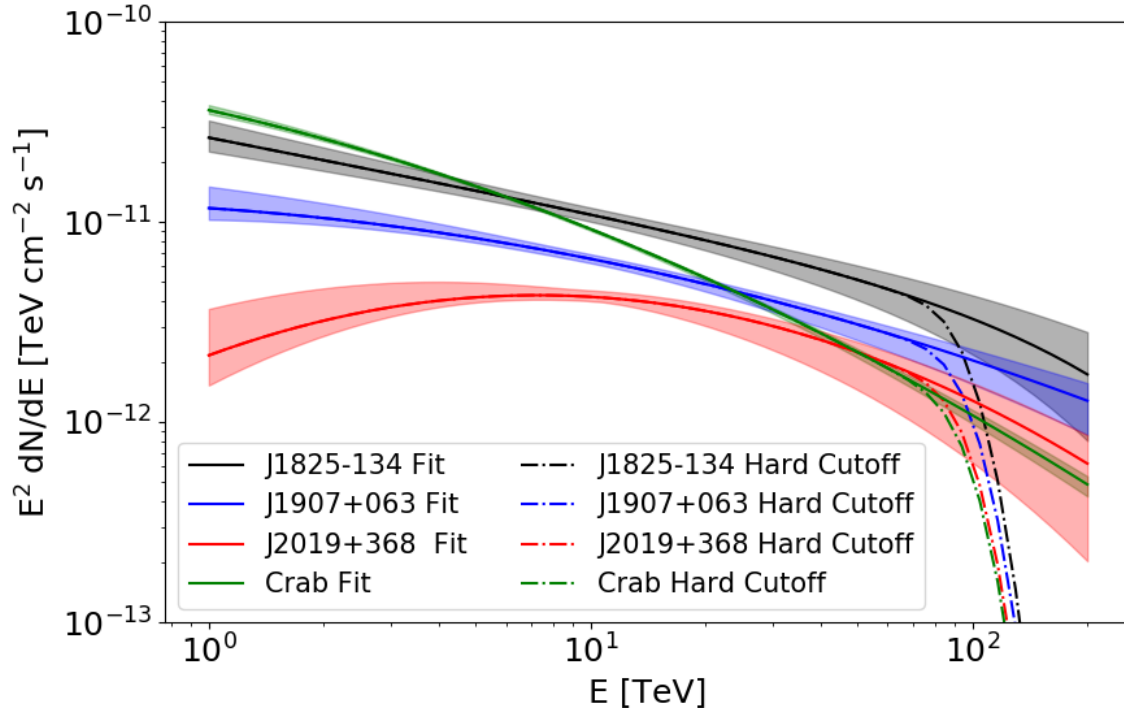


Figure 5.3: Plots of best-fit spectra for each source compared with spectra expected were a hard cutoff detected at 100 TeV. The shaded regions show the statistical uncertainties on the fits. As expected, the presence of a hard cutoff in the HAWC energy range would have had statistically significant effect on the spectra were it detected, even with the additional softening added.

to find the optimal radius the lower limits on  $E_c$  are computed for successively larger radii until the limit reaches a maximum. A plot showing the limits as a function of radius for each source, with the optimum marked, is shown in Fig. 5.4

## 5.5 Results

Fits with the hard cutoff added were run on each of the four candidate sources considered. No source showed statistically significant evidence of a hard cutoff and therefore no LIV phenomena were detected. The resulting p-values and lower limits at both 95% and  $3\sigma$  confidence are shown in Table 5.2.

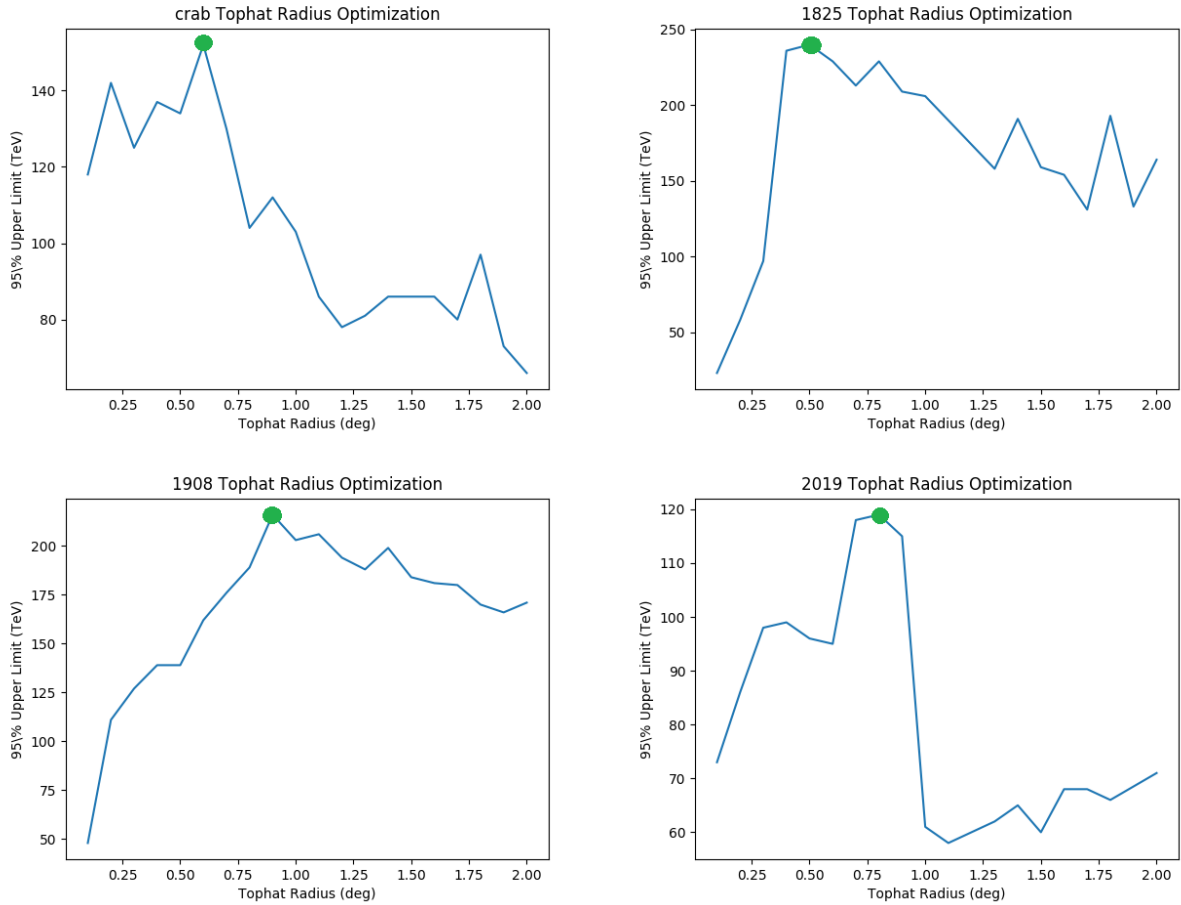


Figure 5.4: 95% CL Lower limits on  $E_c$  as a function of chosen radius for the four high energy sources considered in the analysis. The optimal radius (marked with a green dot) is the smallest one such that all the highest energy photons are included and the lower limit is maximized. Each optimal radius is on the order of the angular extent of the source, where it is expected that all high-energy photons are incorporated into the fit with minimal background.

Source	p-value	$E_c(95\%)$	$E_c(3\sigma)$
eHWC J1825-134	1.000	244	158
eHWC J1907+063	0.990	218	162
eHWC J0534+220 (Crab)	1.000	152	104
eHWC J2019+368	0.828	120	88

Table 5.2: HAWC sources with p-values for LIV detection and lower limits on maximum photon energy in TeV.

### 5.5.1 Individual Results

It should be noted that the inherently one-sided nature of  $E_c$  requires a specific interpretation of the p-values in Table 5.2. Typically, a null detection corresponds to a p-value distribution centered about 0.5, with statistical fluctuations being able to both increase and decrease the p-value. However, the value for  $E_c$  corresponding to the null case is infinite, and inherently at an edge of the parameter space. Therefore, only downward fluctuations in the significance can be detected, with upward fluctuations producing a result indistinguishable from the null case. The p-value distribution for the null case instead consists of a delta function at 1 and a flat distribution extending to 0. This modified distribution then yields an expected p-value close to 1, which is what is observed for each source. Therefore, these p-values are interpreted as indicating a lack of evidence for a finite hard cutoff arising from LIV.

To check the limits are within statistical expectations, a series of pseudo-experiments are run for each source. The best-fit spectrum (with no LIV cutoff) for each source was injected into simulated HAWC data, and then 100 pseudo-maps were generated following Poisson fluctuations of the flux and background. The limits were then re-computed from each pseudo-experiment to check the actual limits were consistent with the distributions found. The 95% and 68% bounds from the simulated distributions are shown in Table 5.3.

These pseudo-experiments allow the interpretation of the p-values to be checked by extracting the true p-value distribution. The distribution for the Crab Nebula is shown in Fig. 5.5. As expected, the distribution contains an added delta function at 1. The clustering

Brazil Bands						
100 Trials						
Source	95% High	68% High	Central	68% Low	95% Low	Actual
Crab	214	162	145	133	83	152
J1825-134	300	232	220	145	92	244
J1908+063	277	235	210	176	139	218
J2019+367	241	158	136	121	78	120
Combined	301	279	240	213	140	285

Table 5.3: Color-coded table of the 68% and 95% statistical bands on the 95% CL limits on  $E_c$ . All the lower limits fall well within the 95% band, with most falling within the 68% band.

of the p-values at 1 is evidence LIV is therefore disfavored.

### 5.5.2 Systematics

Systematic uncertainties arising from the HAWC detector simulation as described in Refs. [33] and [41] are first considered. These include modeling of the PMT response, charge uncertainty, and timing calibration. These systematic effects are derived by redoing the forward-folding analysis with Monte Carlo sets that vary the simulation parameters and obtaining new limits in each case.

Another class of uncertainties considered is the source model. The functional form chosen to model each source’s emission was purely from the maximum likelihood and was not derived from an a priori justification. Therefore effect of changing each source’s assumed spectral shape (either log parabola or cut-off power law) to an alternative shape is another systematic. An additional modeling systematic exists in the chosen source location. As described above, the two energy estimators used by HAWC yield different best-fit coordinates for the high-energy flux. Therefore, the effect of centering the source on the location derived from the ground parameter maps is considered as another possible systematic.

Finally, the effect of uncertainties in the HAWC absolute energy scale is considered. This is estimated by comparing the HAWC Crab spectrum to that measured by the HESS

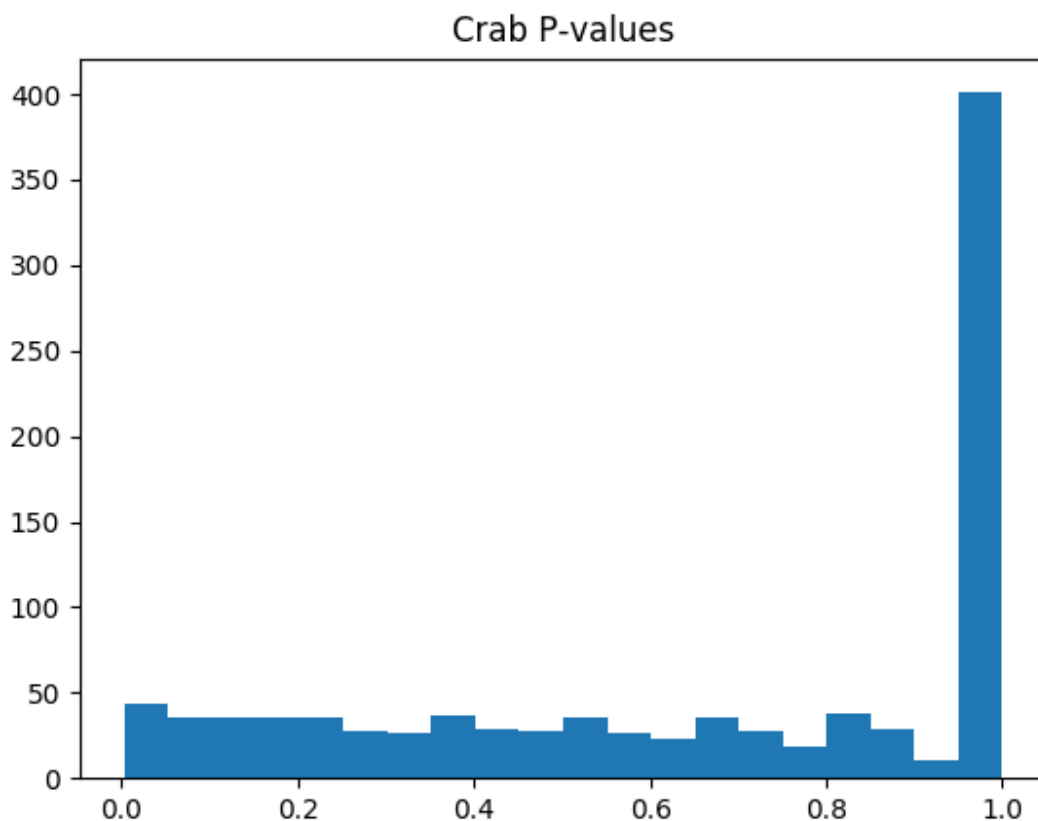


Figure 5.5: Distribution of p-values derived from 1000 pseudo-experiments on the Crab Nebula. As expected, the one-sided nature of the hard cutoff creates a delta function at 1.

experiment. Because both experiments will have separate systematics, it is not possible to derive an exact estimate of the HAWC energy scale uncertainty from this alone. However, it is possible to obtain an upper bound on this systematic by considering a “worst-case” scenario where all sources of error fall solely on HAWC. This highly-conservative estimate gives an upper bound of approximately 6% on the error in the HAWC energy scale and on the  $E_c$  lower limits.

The effects of each of these systematics are shown for both the individual and combined lower limits on  $E_c$  in Table 5.4. The cumulative percent errors (obtained by quadrature addition) on the limits are then summarized in Table 5.5. Note that only the low and high ends of each detector simulation uncertainty are considered in the total error calculation.

Systematic	Crab	J1825-134	J1908+063	J2019+367	Combined
Nominal	152	244	218	120	285
0 BP	159	258	224	123	296
0.5 BP	151	244	218	120	288
0.75 BP	149	237	214	119	282
1.5 BP	141	222	214	119	272
0 Qunc	151	239	216	121	285
0.6 Qunc	145	231	212	119	280
0.15Thresh	151	239	216	121	285
0.25Thresh	148	239	216	121	282
run004255	152	233	214	121	280
run005214	143	239	214	118	277
run005689	158	235	210	116	277
run006283	144	239	216	119	280
run006801	148	248	218	120	282
run007578	145	237	216	119	280
Swapped	131	246	210	124	288
Shifted	152	253	218	115	281
Energy Scale	137	228	197	104	268

Table 5.4: Effect of systematic uncertainties on the 95% CL lower limits on  $E_c$ . All values are reported in TeV. “BP” refers to the inclusion of a broad pulse in the detector simulation, “Qunc” refers to the size of the charge uncertainty, “Thresh” refers to the PMT triggering threshold and the various run indices are different estimates of PMT efficiency (see Ref. [33] for details). The “Swapped” and “Shifted” labels refer respectively to changing to the less-favored spectral form and moving the source center to that fitted using all bins rather than just the highest energy bins. Finally, “Energy Scale” refers to a 10% possible effect of the HAWC absolute energy scale compared to the HESS detector. The data in this table was published in [46].



Simulation	-4% to +7%
Spectrum choice	-1%
Source location	-1%
Energy scale	-6%
Overall	-7% to + 7%

Table 5.5: Overall effects of systematic errors on  $E_c$  for the combined limits.

### 5.5.3 LIV Limits

Limits on the LIV parameters are computed using the more constraining combined results when possible. All limits are reported with associated systematic uncertainties in Table 5.6. Limits derived from photon splitting are reported only for individual sources, since these results depend on both  $E_c$  and the distance  $L$  to the source. We also propagate the systematic uncertainties on  $E_c$  into the LIV limits. Note that the effect of errors in the source distance on the photon-splitting limits are neglected. Because the limit here scales as  $L^{1/10}$ , the effect is negligible compared to that of systematics on  $E_c$ . These limits are several orders of magnitude more constraining than any previously published and verify that Lorentz invariance holds up to the highest currently measured gamma-ray energies [46].

Source	$E_c$	L	$\alpha_0$	$\alpha_1$	$\alpha_2$	$\alpha_2(3\gamma)$	$E_{\text{LIV}}^{(1)}$	$E_{\text{LIV}}^{(2)}$	$E_{\text{LIV}}^{(2)}(3\gamma)$
	TeV	kpc	$10^{-17}$	$10^{-32}\text{eV}^{-1}$	$10^{-48}\text{eV}^{-2}$	$10^{-48}\text{eV}^{-2}$	$10^{31}\text{eV}$	$10^{23}\text{eV}$	$10^{23}\text{eV}$
<b>eHWC J1825-134</b>	244	1.55	$1.75_{1.50}^{2.00}$	$7.19_{5.68}^{8.70}$	$295_{212}^{378}$	$0.70_{0.5}^{0.9}$	$1.39_{1.10}^{1.68}$	$0.58_{0.50}^{0.66}$	$12_{10.3}^{14.0}$
<b>eHWC J1907+063</b>	218	2.37	$2.2_{1.89}^{2.51}$	$10.1_{7.98}^{12.22}$	$462_{333}^{591}$	$0.99_{0.71}^{1.27}$	$0.99_{0.78}^{1.20}$	$0.47_{0.40}^{0.54}$	$10.1_{8.7}^{12.0}$
<b>eHWC J0534+220 (Crab)</b>	152	2	$4.5_{3.89}^{5.15}$	$29.7_{23.46}^{35.94}$	$1960_{1411}^{2509}$	$4.01_{2.89}^{5.13}$	$0.34_{0.27}^{0.41}$	$0.23_{0.20}^{0.26}$	$4.99_{4.3}^{6.0}$
<b>eHWC J2019+368</b>	120	1.8	$7.25_{6.24}^{8.27}$	$60.4_{47.72}^{73.08}$	$5040_{3629}^{6451}$	$10.1_{7.27}^{12.93}$	$0.17_{0.13}^{0.21}$	$0.14_{0.12}^{0.16}$	$3.15_{2.7}^{4.0}$
<b>Combined</b>	285	-	$1.29_{1.11}^{1.47}$	$4.51_{3.56}^{5.46}$	$158_{114}^{202}$	-	$2.22_{1.75}^{2.69}$	$0.8_{0.69}^{0.91}$	-

Table 5.6: Limits on LIV parameters derived from lower limits on the maximum observed photon energy,  $E_c$ . Note that limits on parameters derived from photon-splitting ( $3\gamma$ ) can only be reported for individual sources, since they depend on the propagation distance.

## CHAPTER 6

### SEARCHES FOR DARK MATTER SUB-STRUCTURE

#### 6.1 Introduction

This chapter returns to the search for dark matter in the gamma-ray regime using HAWC data. This analysis attempts to detect gamma-ray emission from as of yet undetected dark matter substructure within the Galactic Halo. These results will then be used to estimate the HAWC sensitivity to both Galactic substructure specifically and dark matter-like spectra in general, as a function of source coordinates. The analysis in this chapter was published in the *Journal of Cosmology and Astrophysics* (Ref. [51]) and will be used in Chapter 7 to determine an optimal region of interest (ROI) for a dark matter search in the main Galactic Halo.

#### 6.2 Missing Satellites

Searches for dark matter signals typically focus on regions known to be heavily dark matter dominated, such as the dwarf spheroidal galaxies surrounding the Milky Way [52]. These dwarf galaxies form as a result of dark matter substructure (sub-halos) within the Milky Way halo. These sub-halos form a gravitational nucleus around which stars can form, leading to the formation of a dwarf galaxy. However, models of the evolution of the early universe predict there to be far more of these sub-halos than are currently observed. Most simulations predict hundreds or even thousands of dwarfs around each galaxy; yet only about one hundred have been observed around the Milky Way. This is known as the missing satellites problem [53]. It should be noted that the inclusion of baryonic matter in these simulations reduces the amount of expected substructure, but the discrepancy still remains [54].

It is theorized that lower-mass dark matter sub-halos are not able to attract visible matter to form dwarf galaxies. Much of the substructure is then hidden in the form of “dark dwarfs”

- nearby dark matter halos with no luminous counterpart that would escape current optical observation [22] [55]. Under the WIMP hypothesis, dark matter annihilation in these unseen dwarfs would produce a gamma-ray excess due to the overdensity of dark matter compared to the main Galactic Halo. Therefore, it may be possible to detect dark matter signals from these previously unknown sub-halos by performing an all-sky survey for gamma-ray signals.

This analysis will investigate resolved HAWC gamma-ray sources that might be associated with dark matter sub-halos. Finding no sources that significantly preferred the dark matter hypothesis, upper limits on dark matter annihilation are calculated for an unbiased sample of the HAWC sky, assuming a dark dwarf were to be discovered. Using this information in combination with simulations of dark matter substructure formation, HAWC’s sensitivity to detections of dark dwarf signals is then estimated.

## 6.3 Searching for Dark Dwarfs in HAWC Data

### 6.3.1 Search Method

This analysis begins by searching 760 days of HAWC data for gamma-ray excesses that could potentially originate from dark dwarfs. Any detected sources in the 2HWC catalog will be considered if they were found at least 5 degrees away from the Galactic Plane with no known astrophysical counterpart reported [34]. A source was considered “detected” in the 2HWC catalog if the likelihood-based test statistic (TS) from Eq. 4.10 was 25 or greater for any of the considered source models. Because these sources had no known normal-matter association (defined as being within the HAWC angular resolution of a known gamma-ray production mechanism [34]) and are so far off the Galactic plane, each one is a potential dark dwarf. Note that although all of these sources have a TS greater than 25 in the 507 day dataset used for the 2HWC catalog, three of these sources have TS values less than 25 in the new data set of 760 days. No additional sources with TS greater than 25 were found in the 760 day data set.

Four sources were found to satisfy these criteria and are listed in Table 6.1, where RA is

Name	RA (°)	Dec (°)	radius (°)
2HWC J0819+157	124.98	15.79	0.5
2HWC J1040+308	160.22	30.87	0.5
2HWC J1309-054	197.31	-5.49	0
2HWC J1829+070	277.34	7.03	0

Table 6.1: Location and spatial data for potential dark dwarfs considered in this analysis. Radius refers to the angular size of the disk source hypothesis (zero corresponding to a point source). All four of these sources were seen in the 2HWC catalog with maximum of  $TS > 25$  [34] and no additional resolved sources were found in the dataset used in this analysis. Only one of these sources (2HWCJ 1040+308) still remains at  $TS > 25$  in the 760 day dataset.

the right ascension, Dec is the declination, and the radius is the angular size of the disk source hypothesis used in the 2HWC catalog. A radius of zero indicates a point source (smaller than the HAWC angular resolution). Sources with larger angular extent would indicate dark dwarfs that are either close to Earth or have large spatial extent. Because this analysis uses a different data set than the 2HWC catalog (which consisted of 507 days of data rather than 760), the maximum TS values for the power law hypothesis will differ from those reported in the catalog consistent with fluctuations from additional data [33]. If the charged cosmic ray background fluctuates to a higher value, but not the signal, this can lower the TS relative to that reported in the catalog, which is the case for three of the four sources (see Table 6.2 in the following section.)

To test these sources for possible dark matter signals, the fits of spectra from dark matter annihilation (Eq. 2.2) are compared to those predicted by astrophysical spectra (Eq. 6.1 and Eq. 6.2).

$$\frac{d\Phi}{dE} = \Phi_0(E/E_0)^{-\gamma} \quad (6.1)$$

$$\frac{d\Phi}{dE} = \Phi_0(E/E_0)^{-\gamma} e^{-E/X} \quad (6.2)$$

Eq. 6.1 is a simple power law where the flux is proportional to energy raised to a power  $\gamma$ , while Eq. 6.2 is a cut-off power law with the addition of an exponentially decaying factor and the cut-off energy  $X$ . In both cases,  $\Phi_0$  is the normalization and  $E_0$  is the pivot energy, that

is fixed to 7 TeV. This choice of parameterization has no effect on the actual spectral shape but does affect the correlation between the analysis bins and is chosen in order to minimize the correlation (see Ref. [33]). In the case of the dark matter hypothesis, no assumptions are made about the density profiles and therefore the  $J$ -factor (Eq. 2.4) cannot be independently computed. Instead,  $\langle\sigma v\rangle J$  is treated as a single free parameter, with the dark matter mass being the other.

For each of the models, the null hypothesis is that the spectrum is background-only. Each alternative hypothesis is tested separately and the TS measured. The dark matter hypothesis is checked to see if it produces a TS at least as high as the most favored (highest TS) power law. If this is the case, further analysis with a closer examination of the likelihood distributions is used to determine if the dark matter hypothesis is statistically favored and the observed emission a possible detection of a dark dwarf.

### 6.3.2 Fit results

For each source, possible fitted forms are a power law, cut-off power law and dark matter spectrum. The fits are performed using the Multi-Mission Maximum Likelihood framework (3ML) [56] [57]. The dark matter fits have two free parameters: the  $J\langle\sigma v\rangle$  pre-factor and  $M$ , the dark matter mass. Various spectra are generated assuming different dark matter mass and annihilation channels. Then the  $J\langle\sigma v\rangle$  that maximizes the likelihood is found in each case. The best fits for the spectra are shown in Fig. 6.1 and summarized in Table 6.2; only the dark matter spectrum with the highest TS is plotted in each case. For illustrative purposes, these plots are shown in  $f_{\text{hit}}$  space with statistical uncertainties rather than in forward-folded energy space. Fig. 6.2 shows several more detailed plots for one source in particular (2HWC J1309-054), both in  $f_{\text{hit}}$  and as a function of energy.

No significant improvement in the TS for the dark matter spectra is found compared to the power laws. In each case, the fits are either comparable or slightly worse (lower TS) for the dark matter hypothesis. It should also be noted that the TS values are nearly

indistinguishable between the two power law hypotheses, indicating that the sources lack sufficient statistical power to differentiate even nested spectral shapes. Therefore, these fits do not favor a dark matter hypothesis. However, it is important to note that these results do not exclude the possibility of dark matter signals from the surveyed sources; they simply do not favor the dark matter hypothesis over other astrophysical spectra.

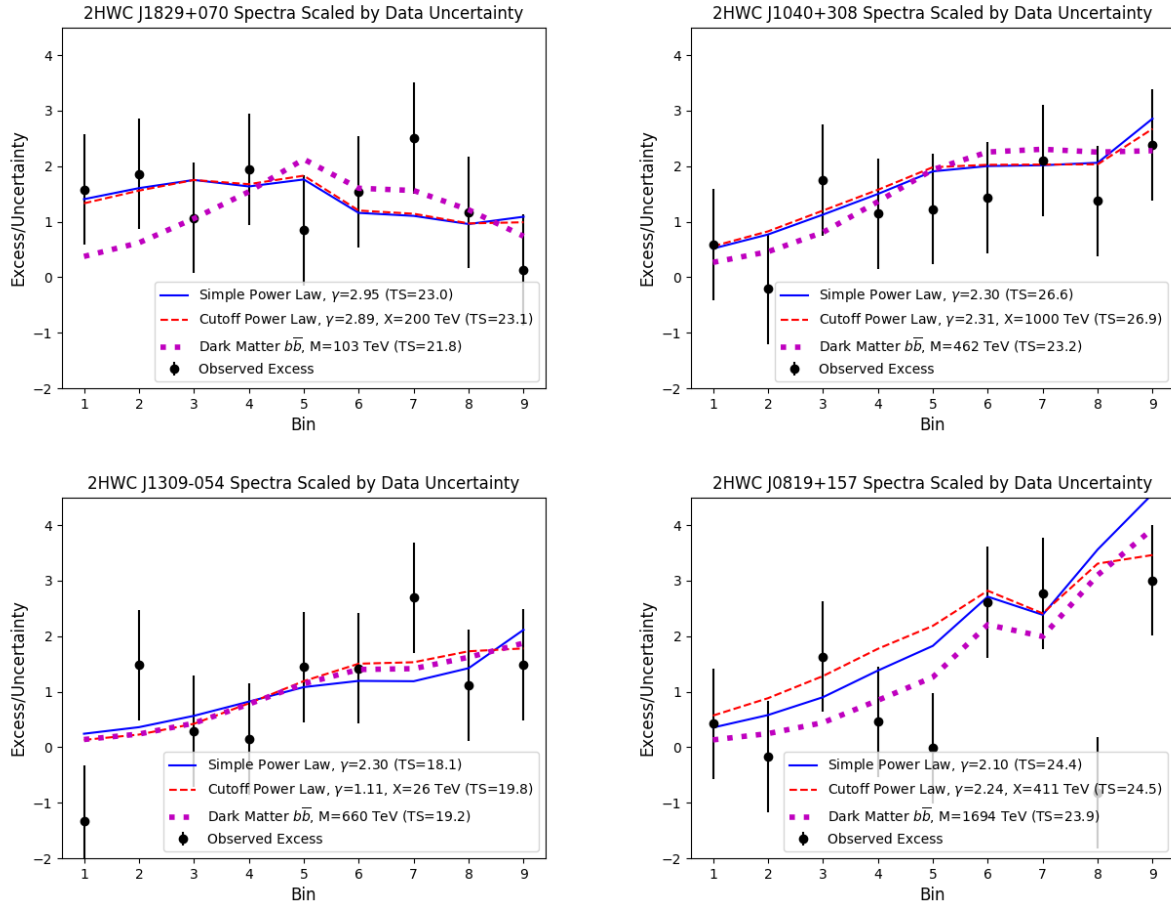


Figure 6.1: Best-fit results in  $f_{\text{hit}}$  bins for the four targeted sources showing the three fitted spectra and TS values. The counts do not reflect the full spatial-likelihood analysis of HAWC and are for visualization purposes. The best-fit TS numbers are from the full HAWC likelihood calculation. Each spectrum is scaled by the estimated uncertainty in data (as is the case in the lower-left plot of Fig. 6.2), showing that each best-fit hypothesis is comparable in goodness of fit. The same spatial hypothesis was assumed for each as the 2HWC catalog (see Table 6.1) [34]

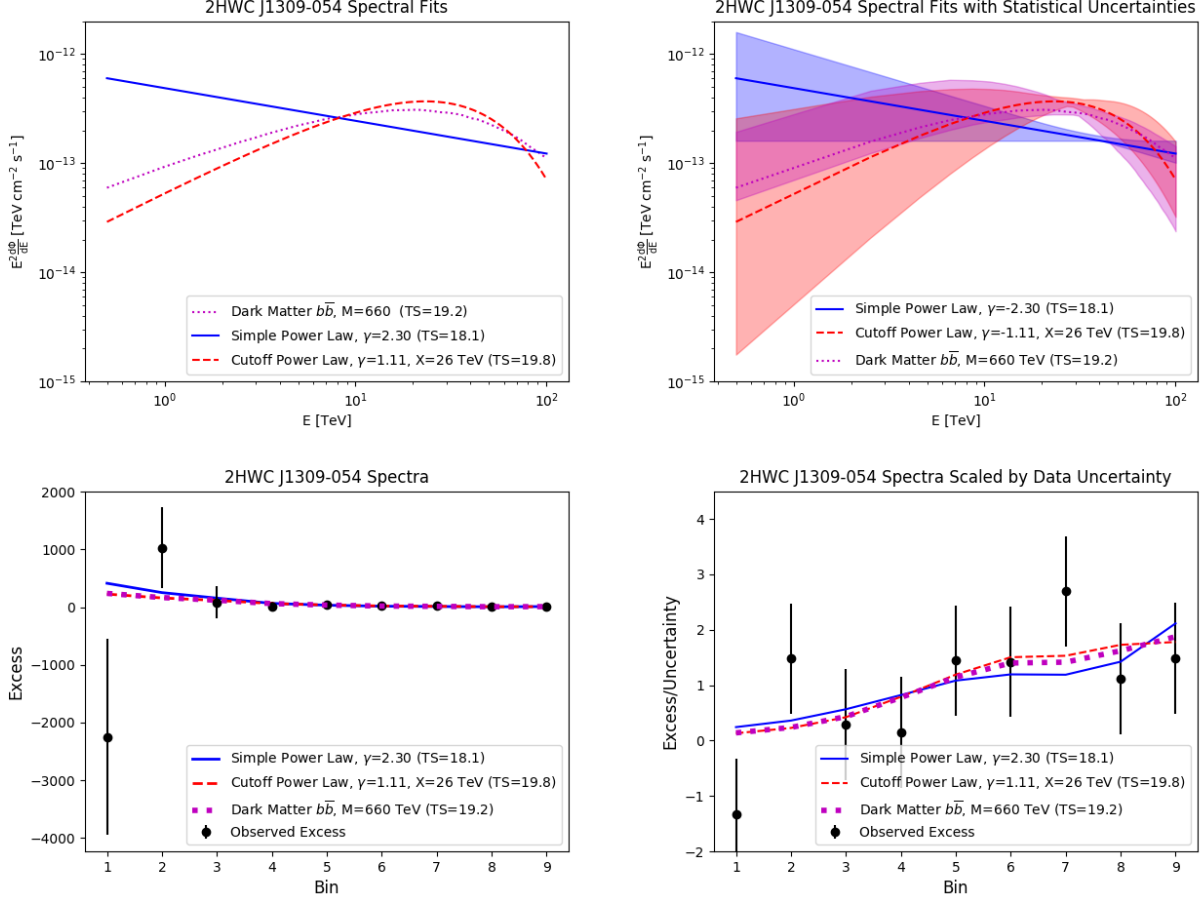


Figure 6.2: Sample best-fit spectra plotted both in energy (top) and  $f_{\text{hit}}$  bins (bottom) to one of the unassociated sources (2HWC J1309-054) considered in this analysis for its best-fit simple power law, cut-off power law, and dark matter spectrum. While the spectra appear distinct in energy-space, they are of comparable goodness of fit to the data in  $f_{\text{hit}}$  bins. The top right plot also shows the statistical uncertainties of the best fits as correspondingly-colored shaded regions, indicating that the range of allowed fits overlaps substantially. We show the  $f_{\text{hit}}$  bin fits both in terms of raw counts (left) and scaled by the estimated error in the data (right). The  $f_{\text{hit}}$  bin plots (bottom) are calculated summing events over a fixed angular disk in the sky rather than the full spatial-likelihood calculation. These are for visualization purposes only. The shown best-fit spectra and TS (Eq. 4.10) values are those from the full spatial-likelihood calculation.



Name	TS (PL)	TS (CU)	TS (DM)	$\Delta$ TS (DM-PL)	$\Delta$ TS (DM-CU)
2HWC J0819+157	24.4	24.5	23.9	-0.5	0.6
2HWC J1040+308	26.6	26.9	23.2	-3.3	-3.7
2HWC J1309-054	18.1	19.8	19.2	1.1	-0.6
2HWC J1829+070	23.0	23.1	21.8	-1.2	-1.3

Table 6.2: Summary of test statistics (TS) and  $\Delta$  TS for the three flux models considered: power law (PL), cut-off power law (CU), and the best-fit dark matter spectrum (DM). See Fig. 6.1 for details on best-fit spectra for each source.

## 6.4 Characteristic Upper Limits Across The Sky

With no clear detections of dark matter sub-halos in the HAWC data, this analysis will proceed to estimate characteristic upper limits on dark matter annihilation within HAWC’s sky-coverage. Characteristic upper limits refer to the average dark matter velocity-weighted cross section,  $\langle\sigma v\rangle$ , that HAWC can exclude at 95% confidence as a function of declination, assuming a dwarf with known J-factor were to be discovered at that declination and no gamma-ray excess was observed. The characteristic limits are estimated by finding the average flux which could be excluded by HAWC at 95% confidence for each declination. If a dark dwarf were discovered by another instrument, these characteristic limits could be used to give the expected value of the corresponding dark matter annihilation upper limit. Note that for any one given discovery of a dwarf, fluctuations in the background at that location would vary the corresponding limit from the characteristic limit.

To calculate the characteristic upper limits, a grid is constructed for  $0^\circ < \text{RA} < 360^\circ$  in steps of one degree and  $-10^\circ < \text{Dec} < 55^\circ$  in steps of five degrees. The declination spacing of five degrees is chosen to account for the HAWC detector response to different declinations, which does not substantially change on the scale of five degrees. The right-ascension spacing is chosen due to the HAWC PSF (see Sec. 4.2.2), which has a width of approximately one degree in the lowest  $f_{\text{hit}}$  bin [33]. A spacing of one degree ensures each sampled point is roughly independent (little overlap in the PSF) while still obtaining

a representative sample of points at each declination. These calculations are performed assuming point source hypotheses since the expected characteristic size of a dwarf is on the scale of the HAWC PSF width, although some dark dwarfs may have slightly larger extent [58] [59]. Note that the points analyzed in Sec. 6.3.1 were drawn from a search of all points in the entire HAWC sky [34], and all possible sources of dark matter have already been identified so there is no chance an additional  $TS = 25$  excess will be missed in this analysis.

To avoid contamination from luminous matter gamma-ray emission, this search excludes any points within five degrees of known luminous matter gamma-ray sources reported in the 2HWC catalog [34], as well as the Galactic plane. As the HAWC PSF has a width of approximately one degree at the lowest energies (and narrows at higher energies), five degrees is chosen in order to exclude both the main emission and the tails due to the PSF as well as emission from highly extended sources [33]. An example significance distribution is shown in Fig. 6.3 and is consistent with background, indicating that the exclusion procedure has removed significant source contamination and is using a sufficiently unbiased sample of the sky. The fit is then performed at each sampled grid point with spectra from a range of assumed dark matter masses and annihilation channels. As the dark dwarfs are expected to be relatively low-mass and/or far away, each fit is performed using a simple point source model. Since no assumptions are made about the density profiles or distances to the satellites, the characteristic limits are on  $J\langle\sigma v\rangle$  rather than just  $\langle\sigma v\rangle$ .

#### 6.4.1 Characteristic Limits

Since the HAWC sensitivity is highly declination-dependent, the characteristic upper limits are reported as both a function of dark matter mass and declination, rather than showing the individual limits for each sampled point. These characteristic 95% CL upper limits are computed using the distribution of fitted fluxes. Each best-fit  $J\langle\sigma v\rangle$  value in a given declination is placed in a histogram and the characteristic upper limit on  $J\langle\sigma v\rangle$  is found to be that which is greater than 95% of the best-fit values. In practice, this is equivalent

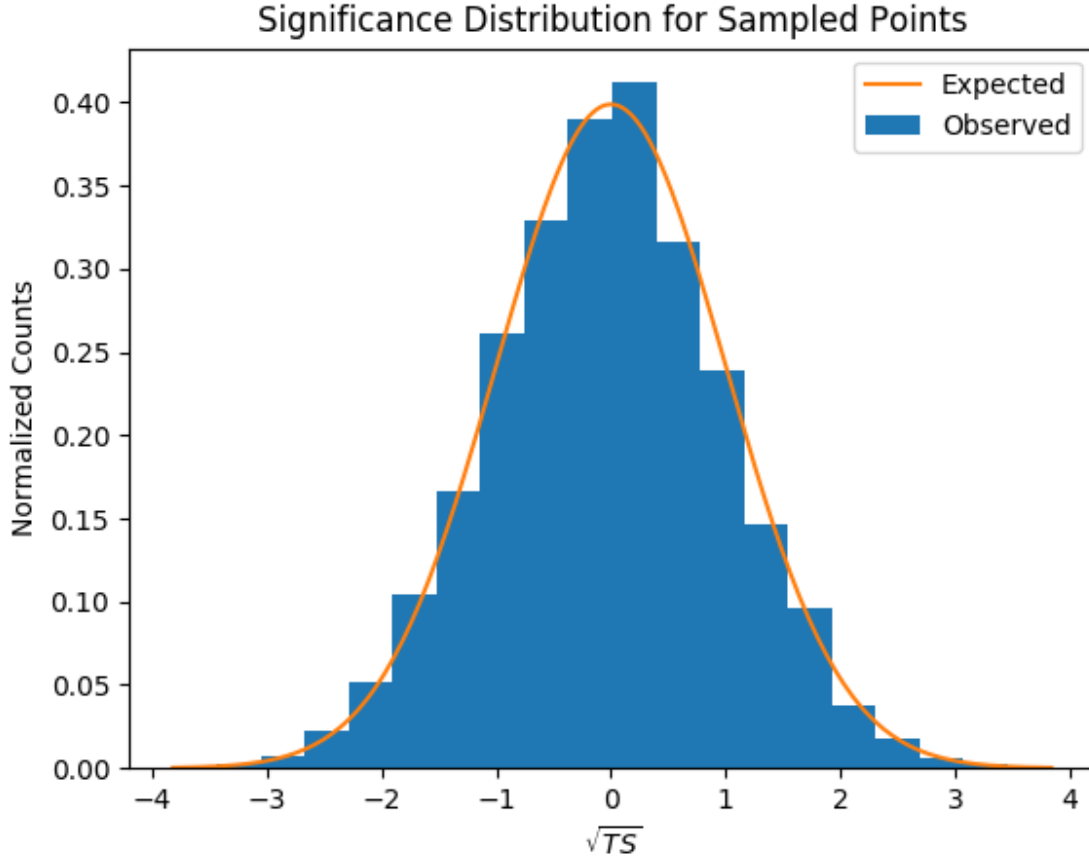


Figure 6.3: Significance (as defined in the 2HWC catalog [34]) distribution of the grid points from all strips of declination used in the characteristic limit calculations (Observed) superimposed over the distribution expected from background-only (Expected). The likelihoods used for this plot were calculated assuming a 10 TeV,  $b\bar{b}$  channel dark matter spectrum. The distribution has a mean of zero and standard deviation of one, consistent with the expectation from background-only, indicating no excess remains after resolved sources from the catalog are removed and that the sample is an unbiased representation of the remaining sky.

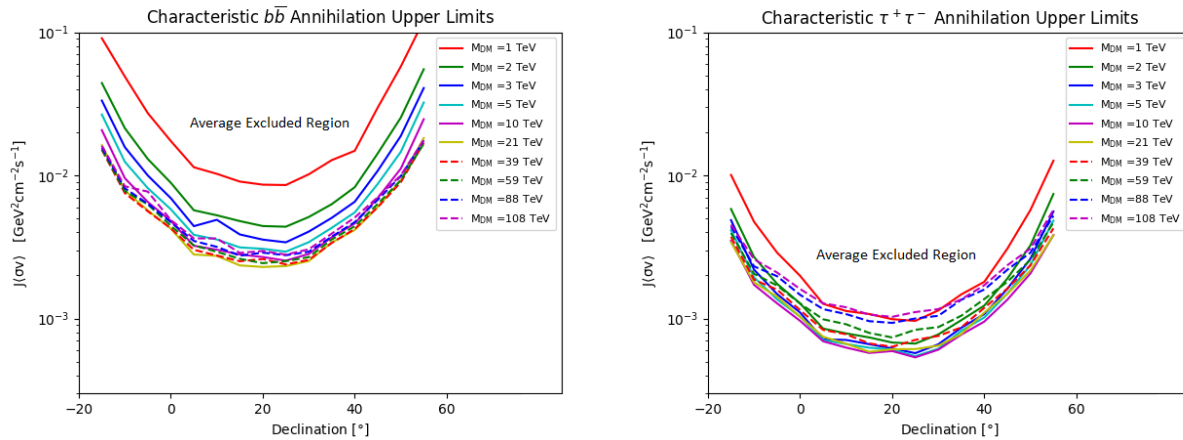


Figure 6.4: Characteristic upper limits on  $J\langle\sigma v\rangle$  for the  $b\bar{b}$  and  $\tau^+\tau^-$  channels at each declination assuming various values of the dark matter mass  $M$ . The points plotted show the expected 95% confidence level upper limit for points at each declination. The most constraining limits are found at declinations directly overhead for HAWC. Individual dwarf upper limits will vary depending on statistical fluctuations at their locations.

to averaging the individual limits obtained above. These characteristic limits are plotted in Fig. 6.4. Because this analysis is ultimately concerned with estimating sensitivity rather than constraining individual dark matter spectra results are reported for only the  $b\bar{b}$  and  $\tau^+\tau^-$  channels, because they are the softest and hardest spectra and therefore are representative of the parameter space. With these characteristic limits, if a new dwarf galaxy were to be discovered and its J-factor measured, the expected  $\langle\sigma v\rangle$  upper limit can be found by scaling the limits in Fig. 6.4 by the corresponding J-factor. An example is shown in Fig. 6.5 assuming a fairly massive dwarf galaxy (with  $J = 10^{19} \text{ GeV}^2 \text{ cm}^{-5}$ ) was discovered at each sampled declination. Assuming an even more massive or nearby dwarf with  $J = 8.7 \times 10^{19} \text{ GeV}^2 \text{ cm}^{-5}$  was discovered at HAWC’s best declination ( $20^\circ$ ) the expected upper limits would become competitive with the HAWC combined dwarf-spheroidal upper limits [52]. For example, assuming the  $b\bar{b}$  channel and a dark matter mass of 10 TeV, the upper limit on  $\langle\sigma v\rangle$  is  $3 \times 10^{-23} \text{ cm}^3 \text{ s}^{-1}$ . In order to become competitive with the corresponding MAGIC upper limits from observations of Segue 1 ( $4.33 \times 10^{-24} \text{ cm}^3 \text{ s}^{-1}$ ) [61], the J-factor would need to be at least  $6 \times 10^{19} \text{ GeV}^2 \text{ cm}^{-5}$ . Note that the actual upper limit derived

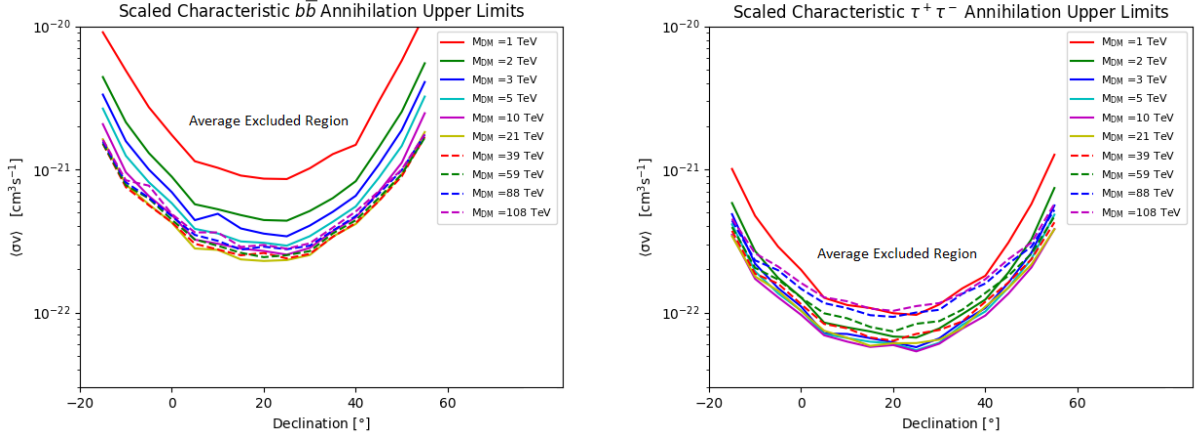


Figure 6.5: Characteristic upper limits on  $\langle\sigma v\rangle$  for the  $b\bar{b}$  and  $\tau^+\tau^-$  channels assuming a dwarf galaxy with  $J = 10^{19} \text{ GeV}^2 \text{ cm}^{-5} \text{ sr}$  were discovered at each declination. These are obtained from scaling the limits in Fig. 6.4 by this J-factor. For reference, the canonical thermal  $\langle\sigma v\rangle$  value is  $2.2 \times 10^{-26} \text{ cm}^{-3} \text{ s}^{-1}$  [60]. Individual upper limits will depend on statistical fluctuations at the location of a discovered dwarf and its measured J-factor.

from a discovered dwarf will depend on fluctuations in the HAWC data at the dwarf location (see the analysis in Ref. [52]). The limits shown here characterize the sensitivity of HAWC to dark matter emission at different declinations and show how constraining these future analyses are expected to be.

#### 6.4.2 Statistical Variations and Expected Limits

To verify the characteristic limits, the expected statistical variation in the upper limits is also calculated. To do so, the upper limits are re-calculated at each declination using a series of pseudo-maps generated using our measured background. Each pseudo-map is created by injecting a simulated signal drawn from a Poisson distribution about the measured background in each bin. The pseudo-map limits are placed in a histogram and the boundaries that contain 68% and 95% of the distribution are found, that demonstrates the possible statistical variation of true limits compared to the characteristic limits. Simulated pseudo-maps are used rather than the actual limit distribution because there are not enough independent points in each declination bin to rigorously find the statistical variation. The characteristic

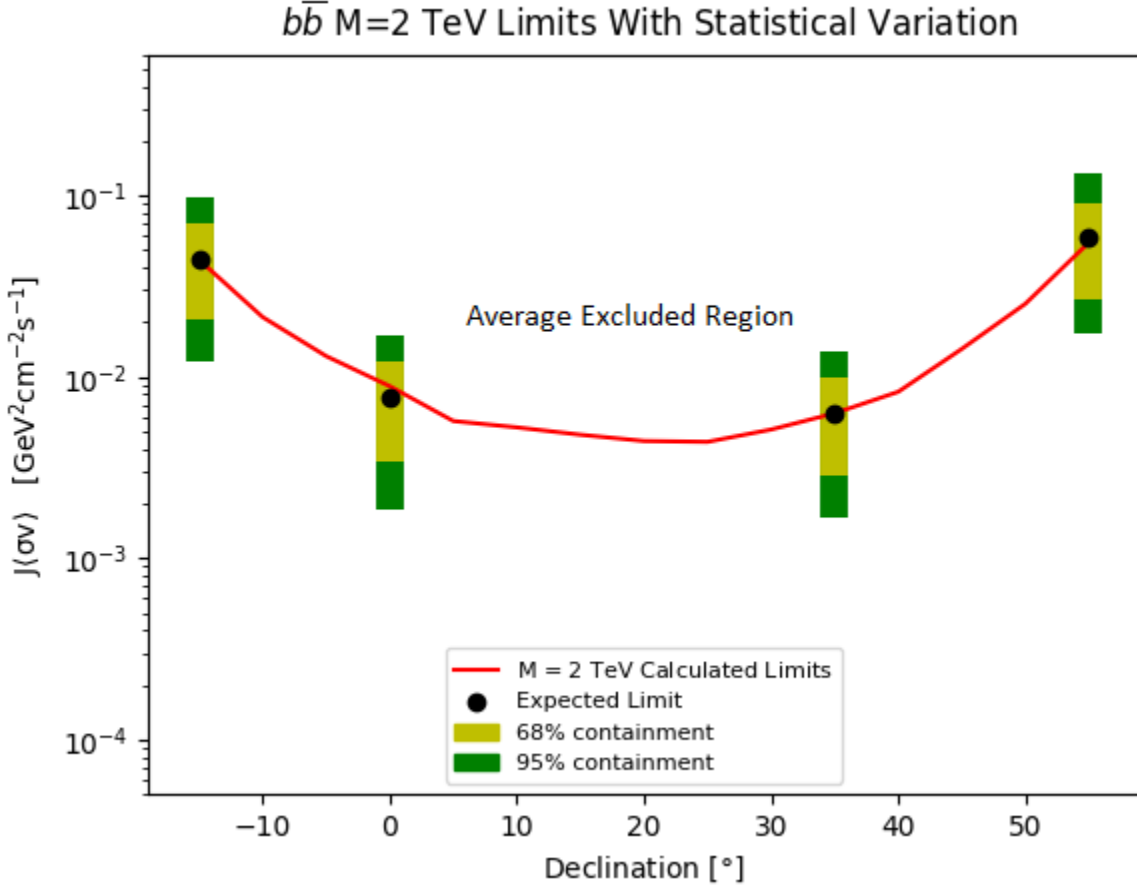


Figure 6.6: Statistical variation (68% and 95%) of the characteristic upper limits on  $J\langle\sigma v\rangle$  as well as the median (expected) value, plotted as a function of declination. These values were calculated assuming a  $b\bar{b}$  annihilation channel spectrum with a dark matter mass of 2 TeV. The error bar size is roughly independent of declinations and spectral assumptions, so only this sample is shown.

limits calculated in Sec. 6.4.1 are consistent with the expected (median) pseudo-map limits, as would be expected from a background-dominated distribution.

The resulting statistical variations and median (expected) limits are shown in Fig. 6.6, superimposed over the corresponding calculated limits from Fig. 6.4. The size of the statistical error bars was found to be roughly independent of the spectral assumptions and declinations. Therefore, the plot shows only a sample of error bars for select declinations and assuming a 2 TeV dark matter mass and a  $b\bar{b}$  annihilation spectrum.

The characteristic limits as a function of declination reflect the HAWC sensitivity curve [33], with the most constraining limits found at points that transit directly overhead. This analysis gives an estimate of the HAWC sensitivity specifically to dark matter spectra. In the following section, this estimate is used in conjunction with simulations of substructure to compute the HAWC sensitivity to detection of gamma-ray flux from an ensemble of dark dwarfs. It should be emphasized that the results in this section show characteristic limits consistent with expected limits within a declination band. The individual upper limits on flux from newly discovered dwarf galaxies will have statistical fluctuations as shown in Fig. 6.6.

## 6.5 HAWC Sensitivity to Modeled Dark Dwarfs

The HAWC sensitivity estimates of Sec. 6.4 can be applied to ensembles of dark dwarfs, not just individual dark dwarfs. The second part of this analysis calculates the HAWC ability to detect dark dwarfs based on models of dark matter sub-structure across the sky. This will give an estimate of how likely HAWC is to observe a dark dwarf in its field-of-view.

### 6.5.1 Modeling Galactic Substructure

The CLUMPY (version 2015.06) software package is used to obtain dark matter substructure models. CLUMPY models both the main (smooth) dark matter halo as well as dark matter sub-halos (clumps) and computes the corresponding J-factor at each modeled point [62]. As the substructure distribution is not perfectly constrained by current observations, a set of characteristic halo model samples are considered with 100 Monte Carlo trials in each case. For each trial CLUMPY creates a HEALPix [40] map of the J-factor associated with each point given a user defined integration angle, that is used to determine the expected halo locations.

The exact behavior of the dark matter density profile towards the center of a halo is not well-constrained, so two different parameterizations that are consistent with numerical simulations and observational data are considered. Many fits to numerically simulated halos favor the Einasto density profile which is characterized by a cuspy shape towards the halo

$R_{\odot}$ (kpc)	$\rho_{\odot}$ (GeV/cm <sup>3</sup> )	$r_s$ (kpc)	$\alpha$
8	0.4	15.7	0.17

Table 6.3: Parameters used in the assumed dark matter density profiles (Eq. 6.3 and Eq. 6.4).  $R_{\odot}$  and  $\rho_{\odot}$  are respectively the distance from the sun to the Galactic center and the local dark matter density of the solar system. The scale radius  $r_s$  is the radius such that  $\rho(r_s) = \rho_s$  and  $\alpha$  determines the slope of the Einasto profile (the Burkert profile does not set the logarithmic slope  $\alpha$  as a free parameter).

center [63] [64] and is given by:

$$\rho(r) = \rho_s e^{\frac{-2}{\alpha}[(r/r_s)^{\alpha}-1]}, \quad (6.3)$$

where  $\rho_s$  is a normalization constant on the dark matter mass density determined by the total halo mass,  $r_s$  is the characteristic scale radius of the halo and  $\alpha$  determines the profile's curvature. Note that in the case of the Galactic halo CLUMPY determines this constant internally using the distance from the sun to the galactic center ( $R_{\odot}$ ) and the local dark matter density ( $\rho_{\odot}$ ) [62]. All parameter values used are reported in Table 6.3.

Empirical measurements of dark matter density profiles through observation of the baryonic component of the halos tend to favor a Burkert profile [65], shown in Eq. 6.4 and characterized by a cored center that lacks a sharp peak.

$$\rho(r) = \frac{\rho_s}{(1 + r/r_s)(1 + (r/r_s)^2)} \quad (6.4)$$

Here,  $\rho_s$  and  $r_s$  are the density normalization and scale radius and are given the same values as in the Einasto profile for the main halo component (see Table 6.3). A comparison of the two profile behaviors in the main Galactic Halo is plotted in Fig. 6.7 for illustration purposes and shows the strong divergence in behavior towards the halo center. This sharp divergence leads to extremely large systematic uncertainties on the expected dark matter gamma-ray flux when only the center of a halo is considered, and is also a substantial systematic when integrating the contribution from an entire halo. Therefore, both profile possibilities will be considered in this analysis. For the simulated sub-halos,  $\rho_s$  and  $r_s$  are determined as



## Dark Matter Radial Profiles

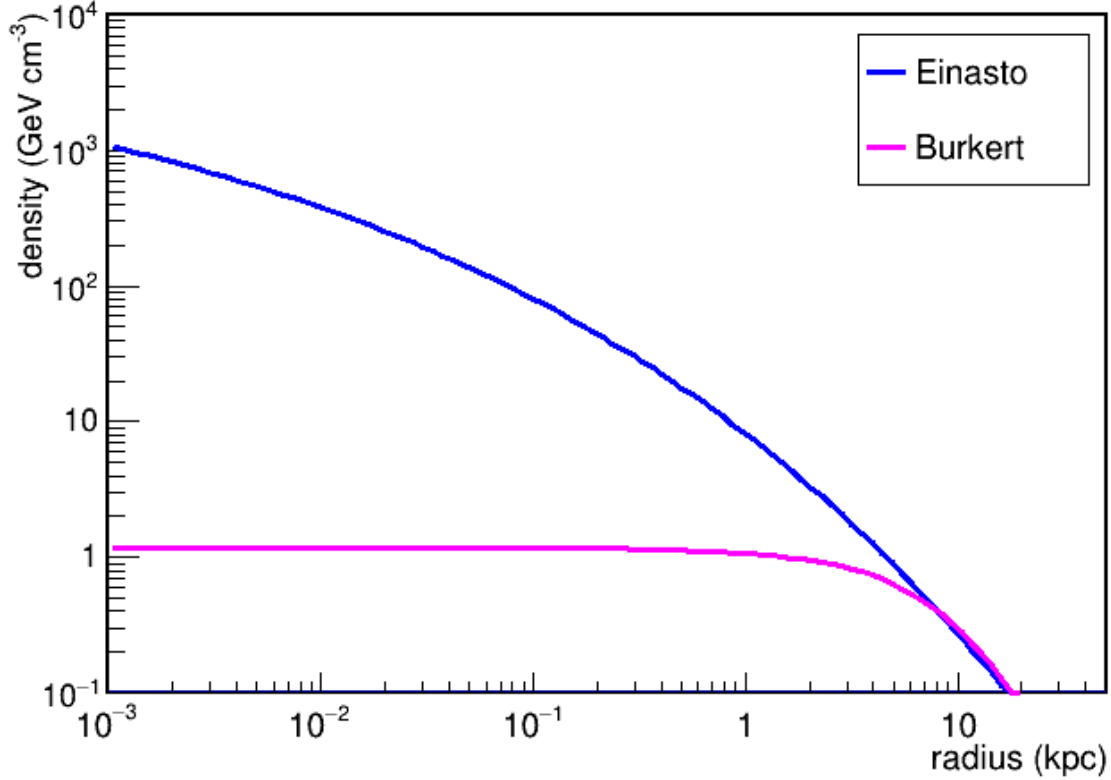


Figure 6.7: A comparison of the density profile behavior as a function of distance from the Galactic Halo center. The Einasto (cuspy) profile differs by nearly three orders of magnitude from the Burkert (cored) profile towards the center.

a function of the total sub-halo mass  $M_{sub}$  using the concentration parameter formalism defined in Ref. [66]. In this framework, the halo shape parameters are related to the total halo mass by the concentration parameter  $c_{vir}$  defined as:

$$c_{vir}(M_{sub}) = R_{vir}(M_{sub})/r_{-2} \quad (6.5)$$

where  $r_{-2}$  is the radius at which the logarithmic slope of the halo is equal to -2 and  $R_{vir}(M_{sub})$  is the virial radius of the halo. The virial radius is given as the following function of  $M_{sub}$ :

$$R_{vir}(M_{sub}) = \left( \frac{M_{sub}}{(4\pi/3)\Delta_{vir}\Omega_m\rho_c} \right)^{1/3} \quad (6.6)$$

where the values of the constants are given in Ref. [66]. Various models exist for the functional

form of  $c_{vir}(M_{sub})$  and this analysis uses the one provided by Ref. [66]:

$$\ln(c_{vir}) = \sum C_i \ln\left(\frac{M_{sub}}{M_\odot}\right) + \ln\left(\frac{1}{1+z}\right) \quad (6.7)$$

where  $M_\odot$  is the solar mass,  $z$  is the halo red-shift (zero for the Galactic Halo), and the coefficients  $C_i$  are taken from Eq. 3 of Ref. [67]. It should be noted that the fitted model to Eq. 6.7 was made assuming the dark matter density follows the NFW profile [67] of the form:

$$\rho(r) = \frac{\rho_s}{\frac{r}{r_s}\left(1 + \frac{r}{r_s}\right)^2}, \quad (6.8)$$

where  $\rho_s$  and  $r_s$  are again the density normalization and scale radius. The NFW profile has a steeper central cusp than the Einasto profile (Eq. 6.3) and has historically been fit to N-body simulations of dark matter that exclude baryons [68]. It should therefore be kept in mind that the coefficients could differ had the simulation assumed an Einasto or Burkert profile. Recent work indicates sub-halos may be more concentrated than the main component [69]. The model reported here is therefore the most conservative possible, as more concentrated sub-halos would produce larger gamma-ray excesses.

For a given halo of mass  $M_{sub}$ ,  $R_{vir}$  and  $c_{vir}$  are calculated and used to determine  $r_{-2}$  using Eq. 6.5. The corresponding  $r_s$  is then obtained from  $r_{-2}$  depending on the functional form of the density profile. For the Einasto profile,  $r_s = r_{-2}$ ; for the Burkert profile,  $r_s \approx r_{-2}/1.52$ . Once  $r_s$  is determined for a given halo mass,  $\rho_s$  is determined by normalizing the integrated density profile to the total mass. Put in equation form:

$$\rho_s = \frac{M_{sub}}{4\pi \int_0^{R_{vir}} f(r)r^2 dr} \quad (6.9)$$

where  $f(r)$  is the functional form of the density profile (Eq. 6.3 or Eq. 6.4).

Following fits to N-body simulations, the distribution of sub-halo masses is assumed to follow a power law form ( $\frac{dN}{dM_{sub}} \sim M_{sub}^{-n}$ ), where  $n$  is the power law index, and the minimum and maximum sub-halo masses are set based on the results of these simulations [70]. The mass probability distribution is then normalized to unity over the chosen range.

$M_{min} (M_{\odot})$	$M_{max} (M_{\odot})$	$n$	$r_s$ (kpc)	$\alpha$
$10^{-6}$	$10^8$	1.9	199	0.69

Table 6.4: Parameters used in the CLUMPY simulation of the Galactic dark matter substructure mass function and spatial distribution.  $M_{min}$  and  $M_{max}$  are respectively the minimum and maximum subhalo masses considered (in units of the solar mass  $M_{\odot}$ ) and  $n$  is the slope of the power law assumed for the mass probability function. The spatial distribution is modeled by Eq. 6.3 with the  $r_s$  and  $\alpha$  values reported here and  $\rho_s$  chosen to normalize to unity. The values reported here are taken from Ref. [70] and Ref. [71].

The parameters used here are summarized in Table 6.4 and the resulting total fraction of dark matter mass contained in the substructure is then 18%. It should be noted that the values determined here were obtained using a set of cosmological parameters that differ from those most recently obtained by the Planck experiment [14], and the fraction contained within sub-halos may be smaller. The spatial probability distribution of the dark dwarfs is modeled by the same functional form as the Einasto profile (Eq. 6.3), where  $\rho_s$  is chosen such that the profile is normalized to unity [71]. Only dark dwarfs with simulated J-factors  $> 10^{16} \text{ GeV}^2 \text{ cm}^{-5} \text{ sr}$  are considered. In both cases, the sub-halo density profiles have the same functional form as the main halo profile [62]. A sample simulation is shown in Fig 6.8. This map was generated assuming an Einasto profile for the main halo and sub-halos and is one of the 100 trials used. Under this profile assumption, one would expect to observe roughly 50% more high J-factor ( $> 10^{19} \text{ GeV}^2 \text{ cm}^{-5} \text{ sr}$ ) dwarfs than have been observed.

### 6.5.2 Detection Thresholds

For each CLUMPY trial, HAWC’s sensitivity to gamma-ray signals from the generated J-factor map of the sky is computed. As the HAWC sensitivity is dependent on declination, the sky is binned based on this sensitivity into declination bands of five degrees. For each dark dwarf the corresponding likelihood profile for a given mass and channel is generated as a function of assumed  $\langle\sigma v\rangle$ , using the observed background from a randomly chosen pixel within the corresponding declination band.

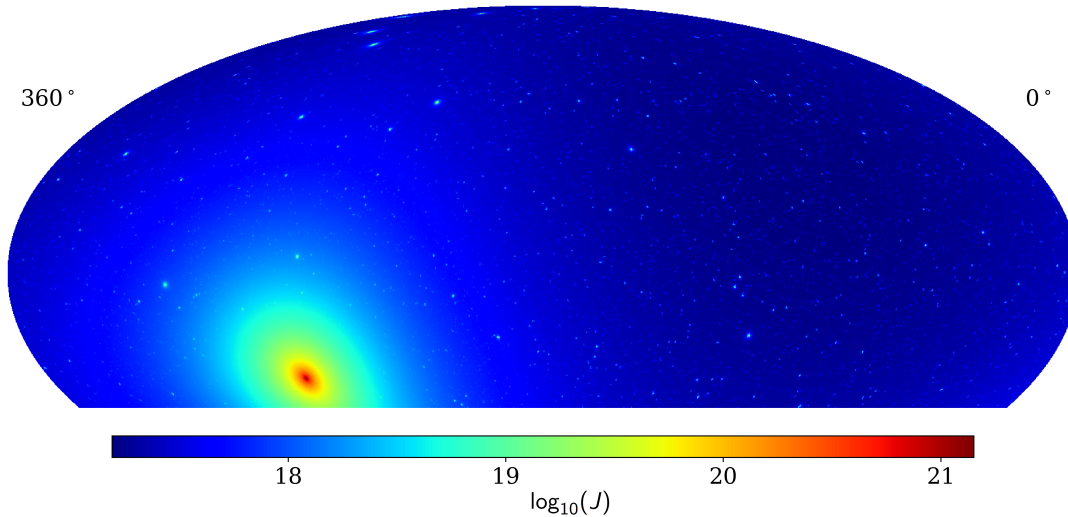


Figure 6.8: Sample map showing one simulation of the smooth halo and substructure of the Milky Way halo, flattened into a Mollweide projection. This map is in celestial coordinates and is truncated so as to correspond to the HAWC field-of-view. The color scale shows the logarithm of the J-factor (Eq. 2.4) integrated over a single pixel width of solid angle ( $0.082 \text{ deg}^2$ ). The high J-factor at the Galactic Center is clearly visible with substructure distributed throughout the sky.

The velocity-weighted cross section  $\langle\sigma v\rangle$  is then calculated such that at least one of these simulated sub-halos' TS values change by 25 from the background-only case. This definition matches that used in the 2HWC catalog (i.e. a source is detected if  $\text{TS} \geq 25$  [34].) See Appendix A of Ref. [52] for additional details on evaluating TS values with HAWC.

This process is repeated for each of the one-hundred CLUMPY trials, resulting in an ensemble of  $\langle\sigma v\rangle$  sensitivity values. Fig. 6.9 reports the median of these values as a function of dark matter mass and annihilation channel. The data in Fig. 6.9 should then be interpreted as the  $\langle\sigma v\rangle$  value such that HAWC would have a 50% chance of detecting dark dwarf emission at  $\text{TS} = 25$  within the ensemble of possible realizations of dark matter substructure. Only the  $b\bar{b}$  and  $\tau^+\tau^-$  channels are considered since they are representative of the overall population of annihilation channels. It should be noted that these results are not intended to be as constraining as the 95% upper limits shown for a hypothetical new dwarf in Fig. 6.5.

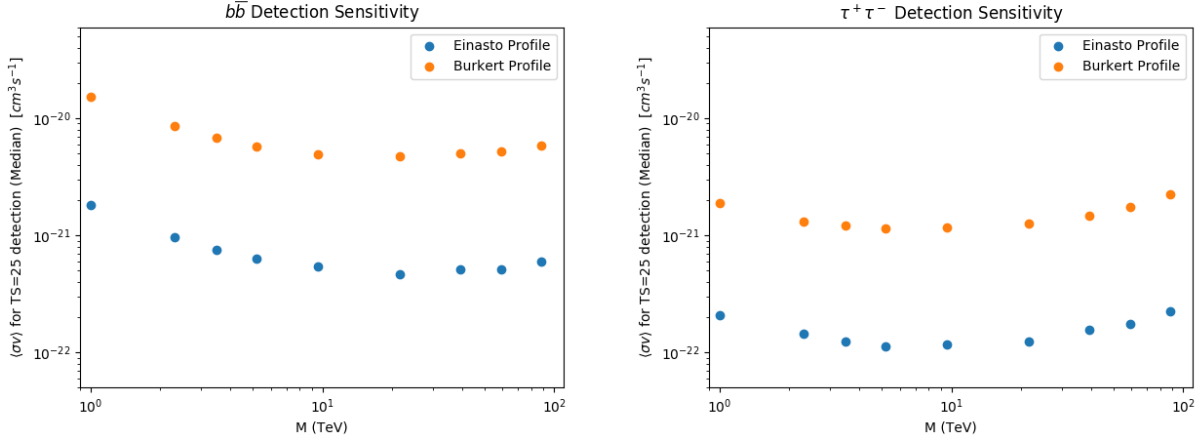


Figure 6.9: Estimated HAWC detection threshold  $\langle\sigma v\rangle$  values for a sample of dark matter simulated substructure. Substructure models for a range of dark matter particle masses, annihilation channels and halo mass models are shown. The plotted points show the  $\langle\sigma v\rangle$  such that 50% of possible realizations yielded a TS of 25 or greater for at least one dark dwarf.

Resolving a source at  $\text{TS} = 25$  requires a much larger  $\langle\sigma v\rangle$  than can be excluded at  $\text{TS} = 2.7$  (the corresponding number for 95% one-sided limits).

In the background-only case, ignoring effects from fluctuations, the TS is approximately proportional to  $\langle\sigma v\rangle^2$ . This can be seen by performing a Taylor expansion on the TS as a function of scale factor (Eq. A4 of Ref. [52]). Therefore, the sensitivity for this analysis is expected to be a factor of approximately  $\sqrt{25}/\sqrt{2.7} \approx 3.08$  less constraining than the 95% CL threshold if considered with the same declination and J-factor. To verify this expectation the results of one of the one-hundred CLUMPY trials is examined, assuming a dark matter mass of 10 TeV. In this trial the  $\langle\sigma v\rangle$  threshold was

$6.06 \times 10^{-22} \text{ cm}^{-3}\text{s}^{-1}$  and resulted in a  $\text{TS} = 25$  detection of a simulated dwarf with  $J = 10^{19.23} \text{ GeV}^2 \text{ cm}^{-5} \text{ sr}$  at a declination of five degrees. The corresponding characteristic upper limit from Fig 6.4, scaled by this J-factor, gives a 95% CL limit on  $\langle\sigma v\rangle$  of  $1.94 \times 10^{-22} \text{ cm}^{-3} \text{ s}^{-1}$ . The  $\text{TS} = 25$  case is a factor of 3.1 times less constraining than the 95% CL case, matching the estimated change.

### 6.5.3 J-factor Sensitivity

Another possible interpretation of the results in Sec 6.5.2 is to determine the minimum J-factor required for a dark dwarf to be detected (assuming some  $\langle\sigma v\rangle$  value and dark matter model). Rather than associating each likelihood profile with a simulated J-factor, one can instead fix the value of  $\langle\sigma v\rangle$  and compute J such that the TS reaches 25. Here, a sample calculation is shown for the  $\tau^+\tau^-$  channel assuming a dark matter mass of 10 TeV (the best-constrained channel according to Fig. 6.4), and a velocity-weighted cross-section of  $10^{-24} \text{ cm}^3 \text{ s}^{-1}$  (roughly the highest value consistent with the upper limits set in Ref. [52]). Using TS profiles generated at a declination of 20 deg (the most sensitive declination bin), the J-factor required for a TS of 25 for this spectrum is found to be  $5.79 \times 10^{20} \text{ GeV}^2 \text{ cm}^{-5} \text{ sr}$ . For comparison the Segue 1 dwarf spheroidal galaxy has an observed J-factor of  $1.1 \times 10^{19} \text{ GeV}^2 \text{ cm}^{-5} \text{ sr}$  [61]. Given that Segue 1 is at a distance of 23 kpc, and that the J-factor is approximately proportional to the square of distance, a dwarf of comparable mass and scale would need to be at most  $\sim 3$  kpc away to guarantee a detection by HAWC.

## 6.6 Summary

This analysis results in an estimate of HAWC’s ability to set dark matter limits from discovered dwarf galaxies. Since no assumptions are made about the dark matter density profiles for the hypothetical dwarfs, these characteristic limits were done for  $J\langle\sigma v\rangle$  rather than  $\langle\sigma v\rangle$ . The studies done in this chapter will allow for the determination of optimal regions for subsequent dark matter searches and characterize the expected sensitivity of future HAWC searches.

## CHAPTER 7

### DARK MATTER IN THE GALACTIC MAIN HALO

#### 7.1 Introduction

The Galactic Halo is the closest large dark matter halo known. This proximity means the region towards the Galactic Center is expected to yield an extremely high dark matter gamma-ray flux and is a promising region for probing WIMP signals. As alluded to in Chapter 2, the HESS experiment has produced strong limits on dark matter for TeV-scale WIMP masses [28]. However, those results rely on assuming a steep central cusp at the center of the Galactic Halo density profile and are unable to probe a cored profile. Another analysis by HESS attempted to probe a cored shape by artificially flattening the center of an otherwise cuspy spatial model [72]. This model still relied on behavior consistent with otherwise cuspy models and is not consistent with observations that indicate a truly cored model of dark matter halos [72].

HAWC has published a search for dark matter signals from a region of the Galactic Halo [24]. Unlike the HESS analysis, the wide field of view of HAWC allows for this region to be probed regardless of density profile model (although it still produces a systematic uncertainty as explained below) [24]. The analysis in this chapter improves and expands on the published HAWC results, using a more robust approach to fitting as well as a larger region of interest (ROI). This should allow for a more constraining set of limits to be produced, while still benefiting from HAWC's wide field of view.

#### 7.2 Spatial Profiles

As discussed in Chapters 2 and 6, the exact behavior of the Galactic halo density profile is not well constrained towards the center, so different parameterizations consistent with numerical simulations and observational data will be assumed. Here, the density profile

models are restated now in the context of the main halo rather than theoretical sub-halos as in Chapter 6.

The two possible density profiles considered are again the Einasto and Burkert profiles. Many fits to numerically simulated halos consisting only of dark matter favor the Einasto density profile which is characterized by a sharp cusp towards the halo center [63] [64] and given by:

$$\rho(r) = \rho_s e^{\frac{-2}{\alpha}[(r/r_s)^\alpha - 1]}, \quad (7.1)$$

where  $\rho_s$  is a normalization constant on the dark matter mass density determined by the total halo mass,  $r_s$  is the characteristic scale radius of the halo and  $\alpha$  determines the profile's curvature, which is fixed to a value of 0.17 for the Galactic Halo [62]. Observations of dark matter halos favor a Burkert density profile [65] that lacks the center cusp and instead flattens towards the center (referred to as being “cored” in contrast to “cuspy”). The Burkert profile is given by:

$$\rho(r) = \frac{\rho_s}{(1 + r/r_s)(1 + (r/r_s)^2)}, \quad (7.2)$$

where  $\rho_s$  and  $r_s$  are the density normalization and scale radius. Additionally, more recent N-body simulations that include baryonic matter favor a cored shape parameterized by Eq. 7.2. The addition of the baryonic matter disk results in an expected flattening of the central peak, causing the cored profiles favored by observation [73].

The behavior of the different density profiles is illustrated in Fig. 7.1 for the Galactic Halo. Towards the center, the profile behavior diverges considerably. Therefore, any J-factor computed using only the region close to the center will have large systematic uncertainties (up to six orders of magnitude) due to the density profile. In the case of the cored Burkert profile, it becomes nearly impossible to resolve an excess from the center if the region of interest is too small. Because this profile lacks a sharp peak, the flux in the “off” or background region will contain a comparable amount of dark matter emission as the “on” region. Therefore, any excess due to dark matter would cancel itself if the background estimation relies on an



## Dark Matter Radial Profiles

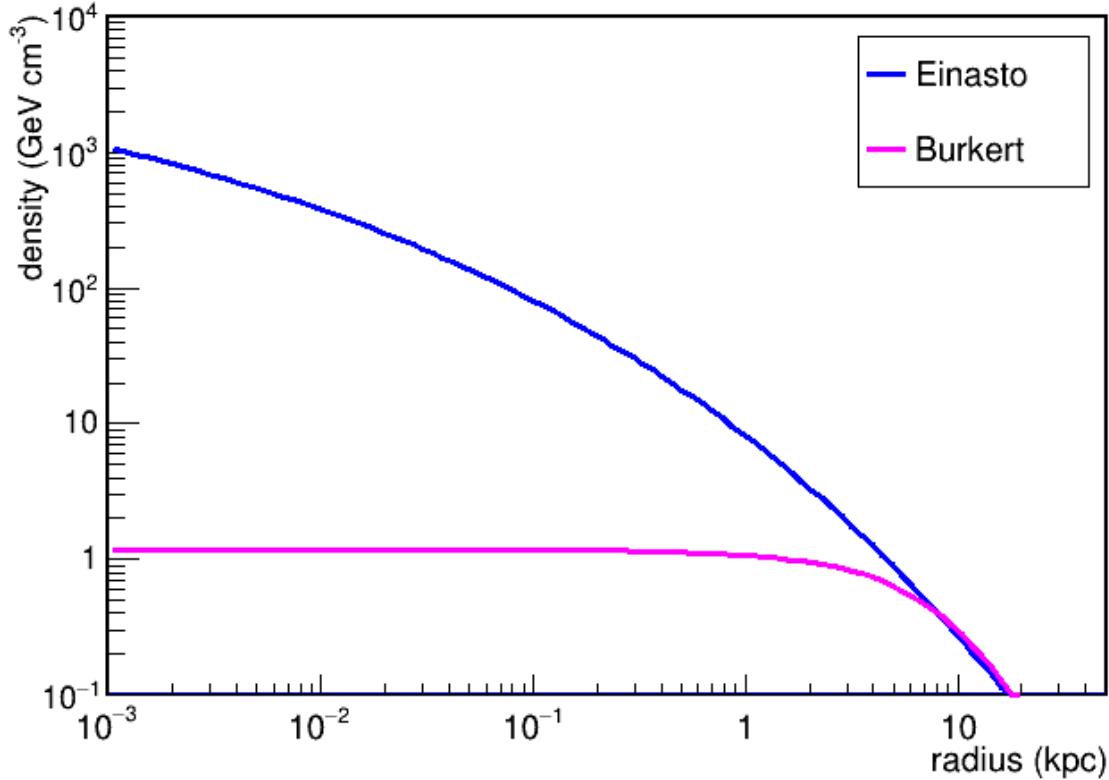


Figure 7.1: A comparison of the density profile behavior as a function of distance from the Galactic Halo center. The Einasto (cuspy) profile differs by nearly three orders of magnitude from the Burkert (cored) profile towards the center.

“off” region in proximity to the center. Indeed, this is why the HESS analysis in Ref. [28] depends on the assumption of a cuspy profile such as Eq. 7.1.

Using HAWC to consider a larger region of interest (ROI) farther away from the Galactic Center substantially mitigates the density profile systematic, since then the cored and cuspy profiles do not differ as strongly. In the cored case, the sensitivity lost from the lack of an assumed central cusp is also mitigated by the larger ROI, allowing more theoretical fluxes to be observed over a larger area. In addition, the HAWC all-sky coverage allows for background-estimation techniques that do not depend on simultaneous observation of regions close to the ROI, allowing an excess to be resolved even in the cored case favored by current observational evidence.

## 7.3 Large-Scale Background Estimation

### 7.3.1 Background Anisotropy and Limits of Direct Integration

The direct integration technique used for background estimation in most HAWC analyses and described in Sec. 4.3.2 typically averages events within a 2-hour window. Resolving a highly-extended source such as the Galactic Halo requires that events be averaged over a full 24 hours in sidereal time, as explained in Refs. [74] and [24]. In addition, direct integration assumes that the background is isotropic in a given declination band. However, at the highly-extended scales of this analysis, the background estimation becomes sensitive to anisotropies in the cosmic-ray arrival direction. This includes both the large-scale cosmic-ray anisotropy due to the direction of the Earth’s revolution about the Sun, and the small-scale cosmic-ray anisotropy due to unknown effects that create local excesses and deficits in the cosmic-ray background [74].

Direct integration over a full 24-hour sidereal day results in background artifacts, As shown in Fig. 7.3. The small scale anisotropy regions (see Fig. 7.2) introduce deviations from the isotropic component on the order of one part in  $10^4$ . When summed over a large number of spatial bins, this introduces spurious excess and deficit regions as illustrated in Fig. 7.3. In addition, around the Galactic Plane the significance is biased strongly negative. The cause of this effect is unknown, but may be related to the limits of direct integration encountered in Ref. [74]. The effect is not correlated with any strong cosmic-ray anisotropy regions shown in Fig. 7.2. Re-making the significance map without gamma/hadron cuts (which results in the data being almost entirely cosmic rays) does not show this same excess, as can be seen in Fig. 7.4. The presence of these features therefore indicates that 24-hour direct integration overestimates the hadronic content of the gamma-ray data in the vicinity of the Galactic Plane and cannot be used to accurately estimate the background in this analysis.

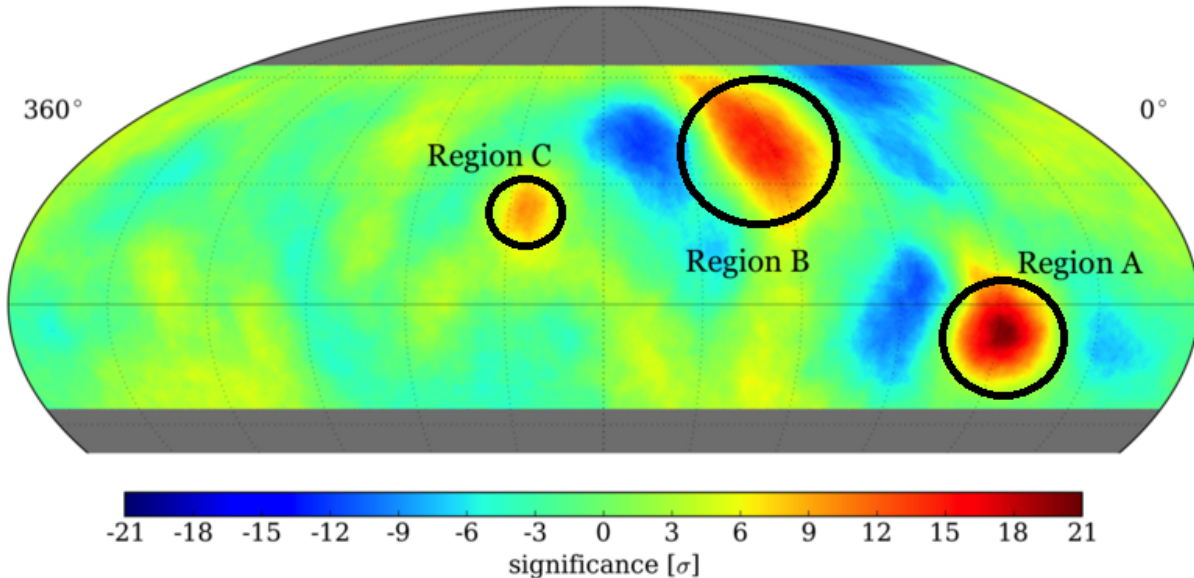


Figure 7.2: Relative intensity map of the small-scale cosmic-ray anisotropy. These deviations from the isotropic background estimation can show up as excesses even in the gamma-ray maps, due to the fact that the data is dominated by hadrons. This effect must be accounted for in this analysis. Figure taken from Ref. [74].

### 7.3.2 The $\alpha$ -factor

An alternative approach to large-scale background estimation has been developed by Pooja Surajbali [75]. This approach is independent of direct integration and is specifically optimized for large-scale sources where background anisotropies are expected to be important. The software and documentation for this technique is available to HAWC collaborators at: [https://gitlab.com/hawc-observatory/sandboxes/psurajbali/alpha\\_factor](https://gitlab.com/hawc-observatory/sandboxes/psurajbali/alpha_factor).

The technique uses two datasets; one with standard gamma-hadron cuts applied and another with reversed cuts meant to pass hadrons, thereby creating a dataset consisting purely of cosmic rays. Both maps are created using the HEALPix scheme with  $\text{NSIDE}=256$  (see Sec. 4.2.2). This approach then attempts to map the behavior of the pure cosmic-ray maps back to the distribution of hadronic background in the standard maps. A series of calculations are then performed to derive what is called the  $\alpha$ -factor (analogous to the

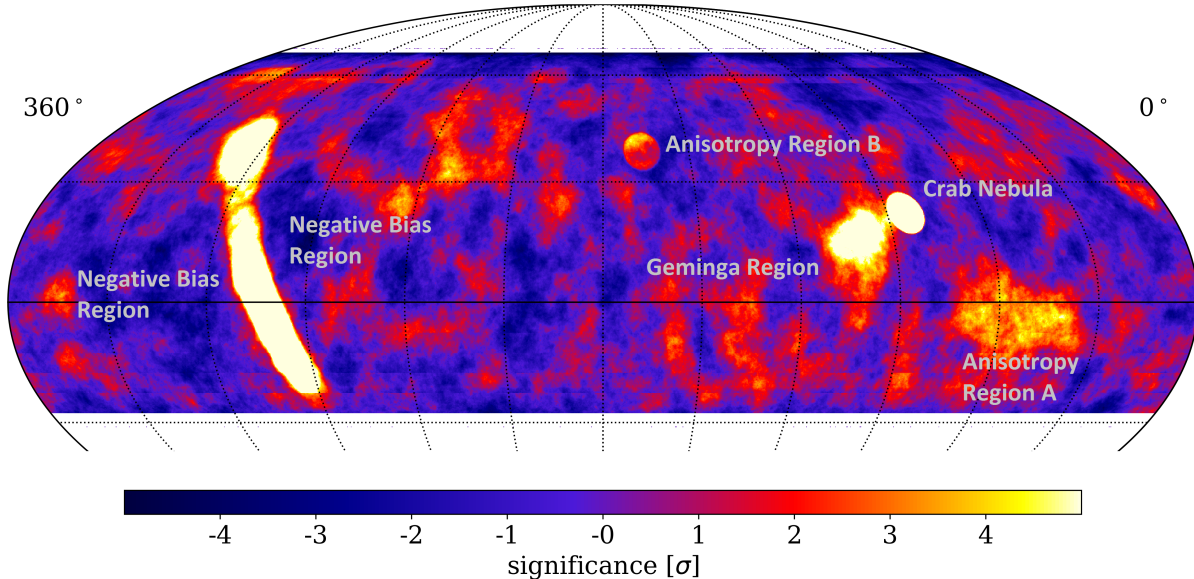


Figure 7.3: Significance map of the HAWC field of view assuming an extended disk source with a radius of 5 degrees and a power law with index -2.7. The Galactic Plane and known HAWC sources are visible as well as the cosmic-ray anisotropy regions A, B, and C defined in Fig. 7.2. The area around the Galactic Plane shows a negative bias of unknown origin, indicating a failure of the direct integration to properly estimate the background.

exposure factor used by Li and Ma [76]). This  $\alpha$ -factor is used to compute the estimated hadronic background in the  $i$ th bin of the standard maps, where  $i$  runs over both spatial bins and analysis bins. This technique has only been developed for the  $f_{\text{hit}}$  binning scheme so the newer energy estimators will not be used in this analysis. In addition, only  $f_{\text{hit}}$  bins 3 through 9 will be considered due to known biases with the  $\alpha$ -factor in the first two bins. The  $\alpha$ -factor is defined by the following equation:

$$N_i^{BKG} = \alpha_i \times H_i, \quad (7.3)$$

where  $H_i$  is the content of the hadron maps in a given pixel and analysis bin, and  $\alpha_i$  is the pixel-by-pixel and bin-by-bin computed  $\alpha$ -factor.

To compute  $\alpha_i$ , it is decomposed into the RA-dependent and Dec-dependent terms as

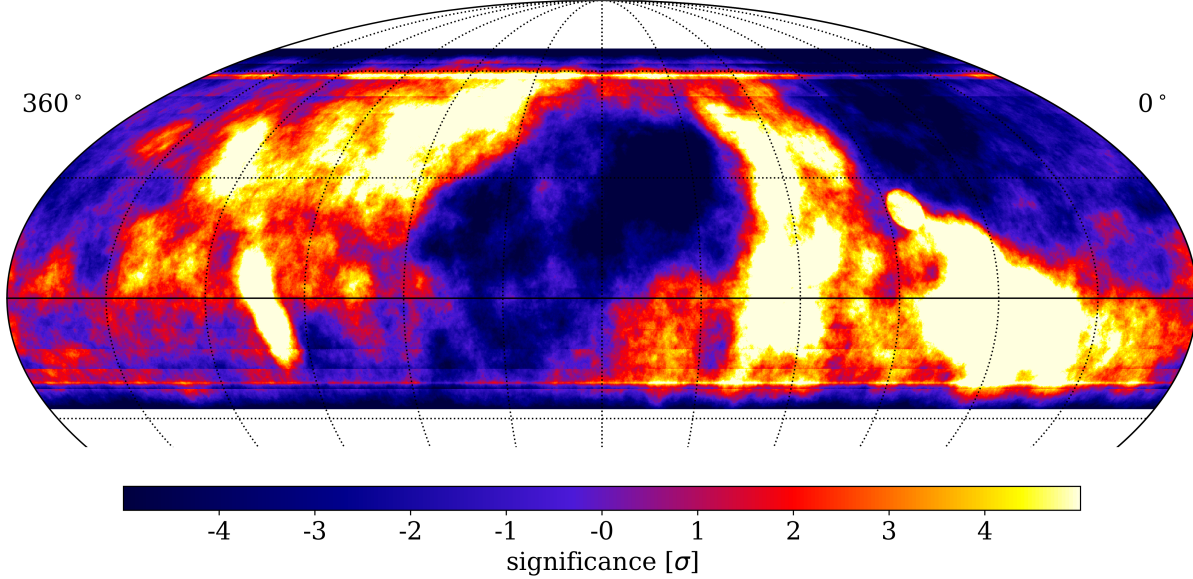


Figure 7.4: Significance map of the HAWC field of view without gamma/hadron cuts assuming an extended disk source with a radius of 5 degrees and a power law with index -2.7. This map is dominated by hadrons due to the lack of cuts. It lacks the negative bias around the Galactic Plane seen in Fig. 7.3, although it contains larger positive and negative regions due to the large-scale cosmic-ray anisotropy discussed in Ref. [74].

follows:

$$\alpha_i(\text{RA}, \text{Dec}) = a_i(\text{Dec}) \times b_i(\text{RA}), \quad (7.4)$$

where the  $i$  index runs over both spatial and analysis bins. Assuming that each component is independent of the other, the  $a_i(\text{Dec})$  component is calculated by:

$$a_i(\text{Dec}) = \frac{G_{\text{Dec}}}{H_{\text{Dec}}}, \quad (7.5)$$

where  $G_{\text{Dec}}$  and  $H_{\text{Dec}}$  are the counts averaged within a 0.5 degree declination window of a given pixel in the gamma and hadron maps, respectively. Since the HAWC event rate is highly dependent on declination, a 0.5 degree binning was chosen that is on the order of the size of a single spatial bin to minimize mixing of counts between different declinations. The RA-dependence,  $b_i(\text{RA})$ , is expected to depend on the pixel-by-pixel count ratio  $\frac{G_i}{H_i}$  but

with the declination-dependence factored out. The equation is then:

$$b_i(\text{RA}) = S_G \left( \frac{G_i}{H_i} \div a_i(\text{dec}) \right), \quad (7.6)$$

where  $S_G$  is a Gaussian smoothing function with width  $\sigma$  (in degrees) determined by the  $f_{\text{Hit}}$  index  $k$  as follows:

$$\sigma = 2k + 15. \quad (7.7)$$

Since the average counts in each map falls off in the higher  $f_{\text{hit}}$  bins, the smoothing width is made wider in these bins to gather additional statistics. See Ref. [75] for further details.

A significance map made with the  $\alpha$ -background shown in Fig. 7.5 does not contain the strong negative regions seen in Fig. 7.3, indicating it does not suffer from the overestimation of background seen in 24-hour direct integration and correctly accounts for the cosmic-ray anisotropy in the background. However, it does contain a new source of excess just off the north side of the Galactic Plane, as shown in Fig. 7.5. This source is morphologically inconsistent with dark matter emission originating from the main Galactic Halo and will need to be removed from the ROI when the search is performed.

## 7.4 Region of Interest (ROI) Selection and Source Modeling

### 7.4.1 Selecting an Optimal ROI

As is the case with direct integration, the  $\alpha$ -background approach requires removing pixels expected to contain excess from the background estimation calculations. This will be the region of interest (ROI) expected to yield the most sensitivity for HAWC. The HAWC sensitivity to dark matter as a function of declination was estimated by using the procedure outlined in Chapter 6 and published in Ref. [51]. The characteristic upper limits from this analysis are on  $J\langle\sigma v\rangle$  rather than  $\langle\sigma v\rangle$  and are independent of the assumed dark matter density profile. In order to estimate the best ROI, taking into account the dark matter morphology, these characteristic limits are combined with the simulations of the CLUMPY density profile discussed in Sec. 6.5.2 and Ref. [51]. To do this, a HEALpix sky-map of characteristic

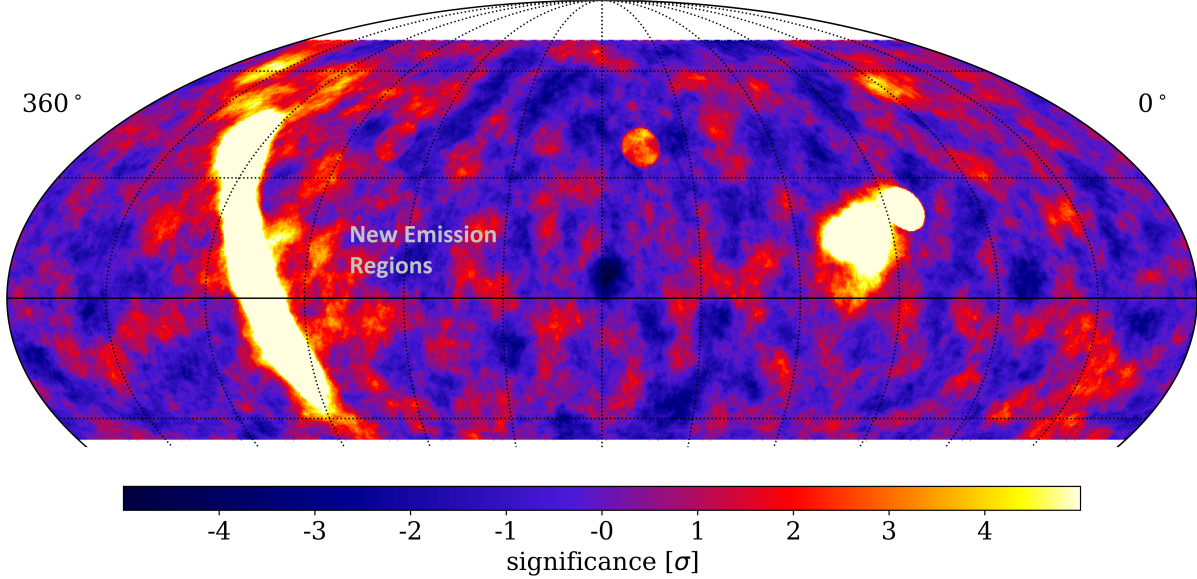


Figure 7.5: Significance map of HAWC field of view assuming an extended disk source with a radius of 5 degrees and a power law with index  $-2.7$  and the  $\alpha$ -background technique for  $f_{\text{hit}}$  bins 3-9. The Galactic Plane and known HAWC sources are visible as well as a new unidentified source of excess. This map lacks the strong negative bias of that shown in Fig. 7.3.

upper limits is constructed and is defined as:

$$\langle\sigma v\rangle_{95\% \text{Char}} = \frac{(J\langle\sigma v\rangle)_{95\% \text{Char}}}{J_{\text{sim}}} \quad (7.8)$$

where  $(J\langle\sigma v\rangle)_{95\% \text{Char}}$  is the characteristic limit associated with a given pixel's declination and  $J_{\text{sim}}$  is the simulated J-factor derived from CLUMPY. As the characteristic limits from Chapter 6 were calculated in 5-degree declination intervals, they are linearly interpolated to cover all pixels for this analysis. Since the characteristic upper limit  $(J\langle\sigma v\rangle)_{95\% \text{Char}}$  depends on the choice of dark matter spectrum used, it is necessary to choose a channel and assumed mass to calculate the sensitivity. For the purposes here, a  $b\bar{b}$  channel annihilation and a dark matter mass of 10 TeV are chosen as a fairly representative spectrum. The choice of spectrum has a negligible effect on the ROI.

The optimal ROI may in general depend on the choice of density profile. Since cuspy profiles peak sharply towards the Galactic center, additional sensitivity could be gained by considering

pixels at the extreme southern edge of the HAWC field of view. In contrast, cored profiles are expected to derive almost all of their sensitivity from regions overhead, since the increase in expected flux towards the center is not enough to dominate loss of sensitivity at high zenith angles. The resulting plots are shown from both a cuspy (Einasto) and cored (Burkert) profile in Fig. 7.6. Since the quantity used to estimate relative sensitivity takes the form of an expected upper limit, smaller values indicate better expected sensitivity.

The optimal ROI chosen contains the Galactic Center despite it lying at the extreme southern edge of HAWC’s field of view. This is to be expected since under a cuspy-halo assumption the expected flux peaks more quickly towards the Galactic Center than the HAWC sensitivity falls off. However, the center lies in the Galactic plane and is therefore subject to contamination from known visible-matter sources. Given the high number of detected HAWC sources associated with Standard Model emission in the Galactic plane, it is best to remove this region from the final ROI. Additionally, all TeV emitting sources with known Standard Model associations in the TeVCat catalog [77] are removed to minimize possible contamination from sub-threshold sources. An additional constraint is made so the ROI excludes all pixels above the equator (Dec=0) in order to avoid the new excess found in the  $\alpha$ -factor analysis shown in Fig. 7.5. Finally, the ROI is selected by choosing the most permissive sensitivity cut such that the ROI at its widest spans no more than 180 degrees. This final constraint is added to prevent the ROI from becoming arbitrarily wide and creating declinations where no pixels remain outside the ROI for use in the background calculation. Under these constraints both density profiles yield approximately the same optimal ROI due to their shapes converging at distances far from the halo center (which is again excluded a priori since it is part of the Galactic Plane). Therefore, the ROI is chosen using only the Einasto profile sensitivity. The necessary cut to satisfy these conditions is  $\log_{10} \left( \frac{\langle \sigma v \rangle_{95\% \text{Char}}}{\text{cm}^3 \text{s}^{-1}} \right) < -17.8$ . The chosen ROI is shown Fig. 7.7 with the excluded regions in dark blue. This is the region that will be probed for Galactic dark matter emission and also excluded from the background calculation.

#### 7.4.2 Extended Source Convolution

To compute the expected flux from such a large extended source, it is necessary to convolve the HAWC PSF (see Sec. 4.2.2) and detector response with the source spatial profile. This can be done



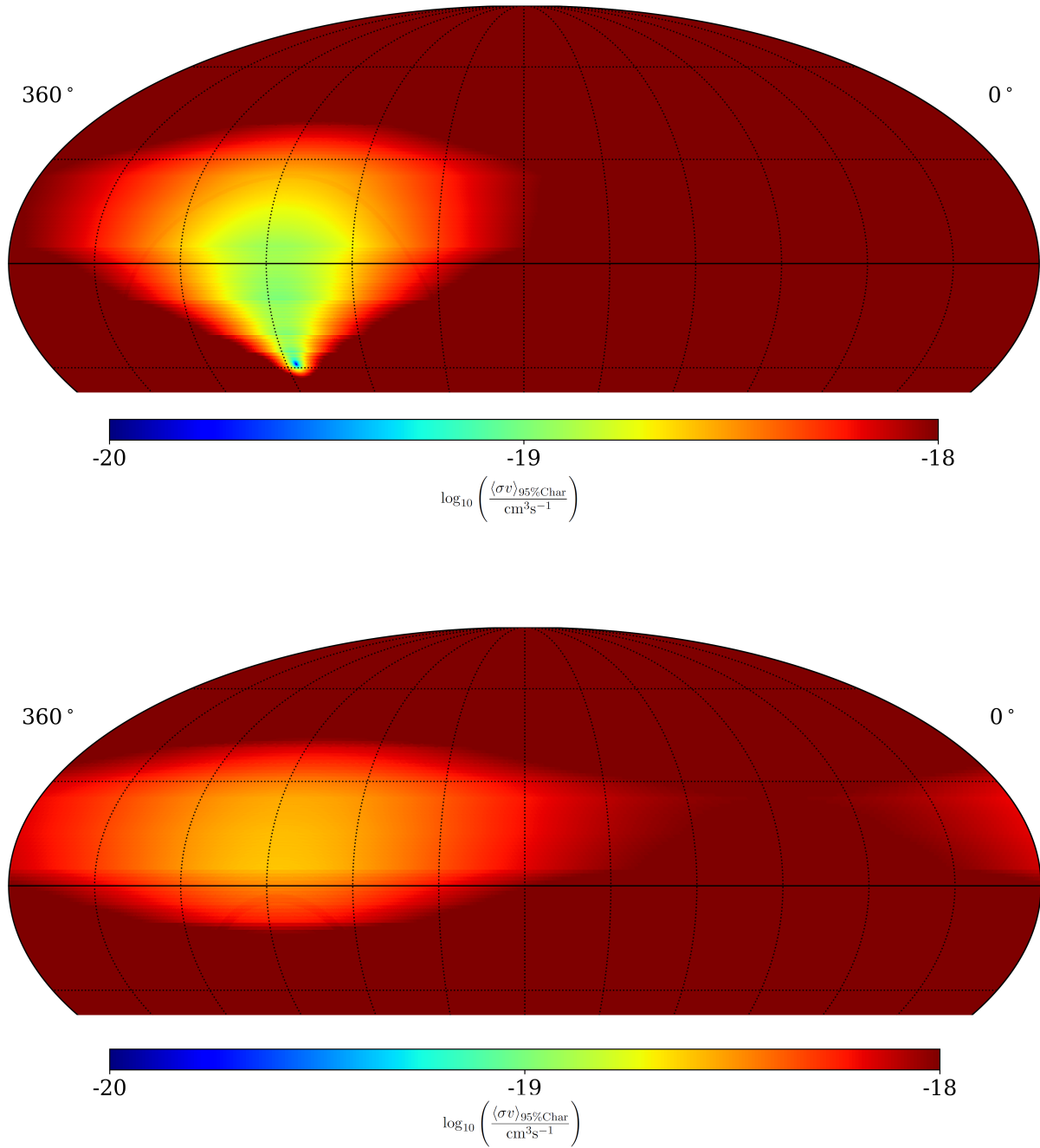


Figure 7.6: Relative sensitivity assuming a 10 TeV  $b\bar{b}$  annihilation spectrum. The top figure assumes an Einasto spatial profile while the bottom figure assumes a Burkert spatial profile. These figures demonstrate HAWC’s sensitivity to dark matter emission from the Galactic halo. Points with more constraining upper limits will be included in the regions of interest for the full analysis and excluded from the background estimation.

using a spherical harmonic decomposition of the PSF, which is much more efficient than brute-force but still requires a large amount of computation time for a highly extended source. Convolutions on a sphere can be sped up substantially by projecting the source region onto a flat sky using the Aitoff projection technique [78]. For relatively small portions of the sphere, any distortions from the projection to Cartesian space become negligible. In flat space, the convolutions can be done with the Fast Fourier Transform (FFT) algorithm rather than a decomposition into spherical harmonics. This is roughly a factor of 100 times quicker than the spherical harmonic approach and allows extended sources' expected fluxes to be computed efficiently. The HAWC Accelerated Likelihood (HAL) implements this approach and is used for all fits in this section [78].

Due to the large extent of the Galactic Halo ROI, its convolution with the HAWC detector response cannot be computed using a single flat sky projection. Instead, the halo is broken up into a set of smaller ROIs each with its own separate flat sky projection. The convolution is then performed separately over each sub-ROI and the resulting models linked to recover the full Galactic Halo extent. To perform a fit the likelihood approach detailed in Sec. 4.6 and applied in Chapters 5 and 6 is generalized by simply summing the likelihood contribution from each sub-ROI to obtain the full likelihood profile. The regions used are illustrated in Fig. 7.7.

### 7.4.3 Electroweak Corrections

While the previous studies cited in this dissertation used spectra simulated with the Pythia [21] software, this analysis derives its dark matter spectra from a more recent framework. This analysis uses spectra computed with the Poor Particle Physicist's Cookbook (PPPC) model as derived by Cirelli et. al. [79]. These spectra are becoming standard within the indirect dark matter detection community due to their ease of use and incorporation of additional corrections. In particular, the PPC spectra include corrections arising from the electroweak coupling. These corrections are relatively small for most channels considered, but have a substantial effect on the charged gauge boson ( $W^+W^-$ ) channel. Here the electroweak correction introduces a secondary peak just before the hard cutoff for sufficiently large assumed dark matter masses, substantially increasing the expected flux at high energies. To illustrate this, a plot comparing sample spectra for the  $b\bar{b}$  and  $W^+W^-$  with and without electroweak corrections is shown in Fig. 7.8. This additional

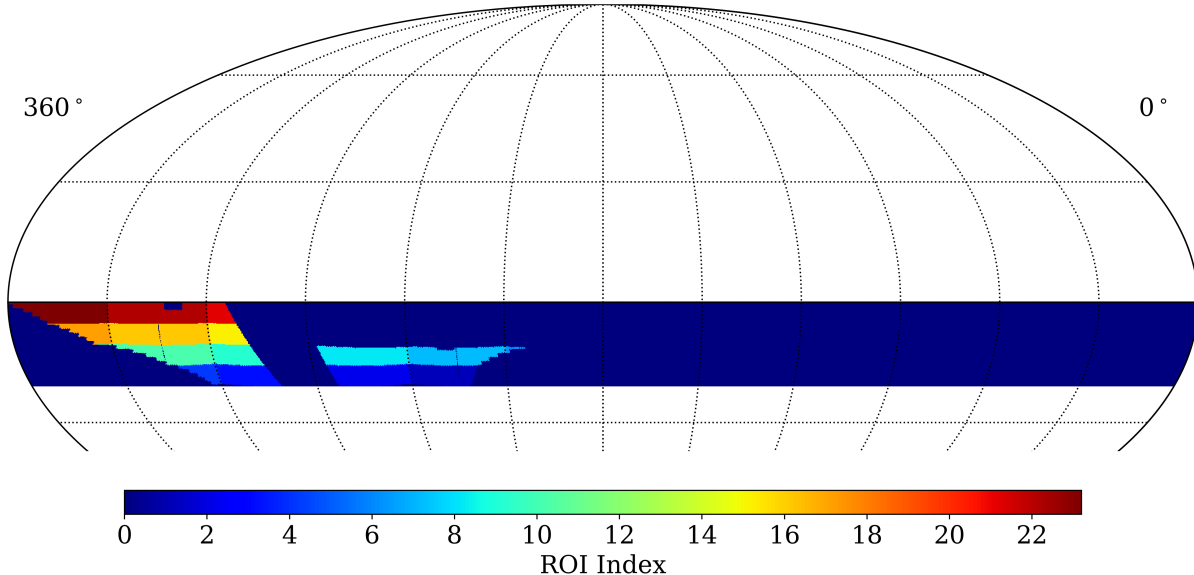


Figure 7.7: The sub-ROIs used to efficiently convolve the Galactic Halo spatial profile with the HAWC detector response using flat sky projections. Each ROI is a different color in an arbitrary scale. To avoid contamination from normal matter gamma-ray sources, the galactic plane and known gamma-ray emission mechanisms are masked out (illustrated by rendering in dark blue).

flux from photon radiation created by the electroweak effects is expected to yield substantially more constraining limits for the  $W^+W^-$  channel than the previous HAWC Galactic Halo analysis, especially when combined with the additional data and optimized ROI.

## 7.5 Signal Contamination of Background

### 7.5.1 Effective Flux

In the event of a source as highly-extended as the Galactic Halo, the estimated background of the data maps will contain a substantial component of gamma-rays. It is therefore only possible to resolve gamma-ray signals in excess of the background gamma-ray component, which will be referred to as effective flux (in contrast to the true total gamma-ray flux originating from hypothetical Galactic dark matter). The previous HAWC analysis of the Galactic Halo had to account for this effect when using direct integration to estimate the background as seen in Ref. [24].

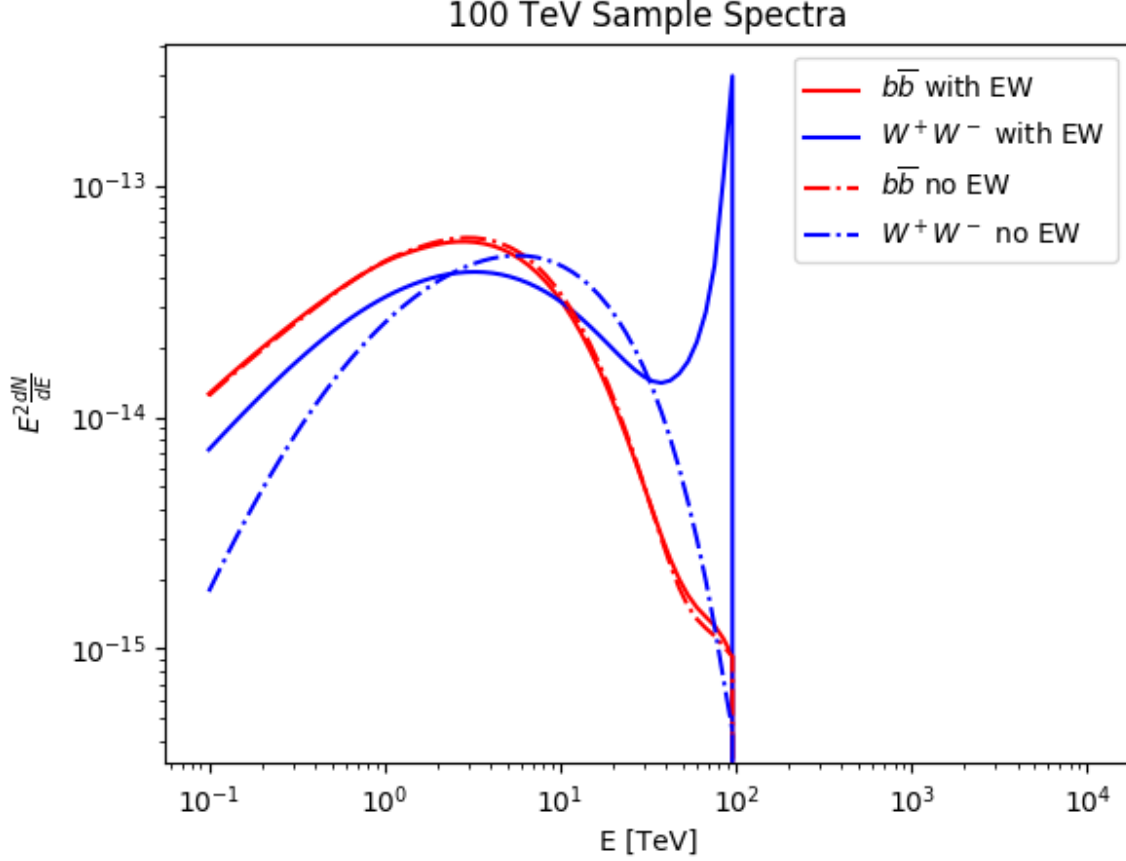


Figure 7.8: Comparison of the expected dark matter energy spectra with and without electroweak corrections for both the  $b\bar{b}$  and  $W^+W^-$  channels, assuming a mass of 100 TeV and arbitrary flux units. The  $W^+W^-$  channel gains a secondary peak from the electroweak correction due to the  $W$  boson having a full unit charge. In contrast the  $b\bar{b}$  channel, with only  $1/3$  unit charge, is only slightly effected.

The contamination,  $C$ , is defined as the fraction of gamma-rays that are classified as part of the estimated background  $N^{\text{bkg}}$ . To derive the expected effective flux in a given bin  $\mu^{\text{sig}}$  in terms of the contamination factor  $C$ , consider decomposing the contents of the alpha-background into gamma and hadron content:

$$N^{\text{bkg}} = N^{\text{CR}} + CN^\gamma, \quad (7.9)$$

where  $N^{\text{CR}}$  and  $N^\gamma$  are respectively the number of hadrons and the number of gamma-rays present in a given pixel. The background is therefore composed both of hadrons and some fraction of gamma-rays, quantified by the contamination  $C$ .

Analogously, the true background  $\mu^{\text{bkg}}$  is decomposed into hadron and gamma components:

$$\mu^{\text{bkg}} = \mu^{\text{CR}} + C\mu^\gamma. \quad (7.10)$$

Note that neither  $\mu^{\text{CR}}$  or  $\mu^\gamma$  are observable quantities. Instead, it is only possible to observe the total background and the fraction of gamma-rays that appear as an excess above the background level. In order to be consistent with the new definition of  $\mu^{\text{bkg}}$ , the definition of the effective flux normalization,  $\mu^{\text{sig}}$ , must instead be:

$$\mu^{\text{sig}} = (1 - C)\mu^\gamma. \quad (7.11)$$

This definition is chosen so that the sum of  $\mu^{\text{sig}}$  and  $\mu^{\text{bkg}}$  is still an estimator for the total number of gamma and hadron events in the data maps. Before performing the fits it is therefore necessary to correct the expected number of gamma-rays  $\mu^\gamma$  by the factor  $(1 - C)$  to transform it into the observable effective flux  $\mu^{\text{sig}}$ , as this is the quantity which can be estimated from the data.

## 7.5.2 Estimation of Contamination Factors

The  $C$  factors needed to correct the interpreted excess in each pixel cannot be determined a-priori, as was done for the J-factors in Ref. [24]. Instead, they must be determined for a given spectral shape and spatial profile. This correction is in general expected to be spectrally-dependent.

To estimate the  $C$  factors, the  $\alpha$  background is recomputed using gamma-ray maps injected with simulated dark matter spectra. For a given simulated spectrum, the HAWC forward-folding technique described in Sec. 4.6 is used to compute the expected number of gamma-rays in each  $f_{\text{hit}}$  and spatial bin  $N_i^\gamma$ . These simulated counts are then added to the corresponding bins in the actual observed gamma-ray maps. The  $\alpha$ -factor is then recomputed using the new gamma-ray map that includes the simulated dark matter spectrum, leaving the hadron maps unmodified. The cuts used to make the hadron maps filter out gamma-rays efficiently enough that the additional simulated contribution is negligible. The  $C$  factor in each bin is then estimated as:

$$C_i = \frac{N_{\text{inj}}^{\text{bkg}} - N^{\text{bkg}}}{N_i^\gamma}, \quad (7.12)$$

where  $N_{\text{inj}}^{\text{bkg}}$  is the background is estimated from using the  $\alpha$ -factor technique on the injected maps,  $N^{\text{bkg}}$  is the background from the non-injected maps, and  $N_i^\gamma$  is the number of gamma-rays from the

injected dark matter spectrum. Because  $N_i^\gamma$  is explicitly dependent on the choice of dark matter spectrum, this process must be done separately for each channel, assumed mass, and spatial profile that will be explored.

It must be established that the  $C$  factor is not strongly dependent on overall flux level (determined by the assumed value of  $\langle\sigma v\rangle$ ). If this is the case, it becomes extremely tedious to do a true fit since every trial value of the cross-section will have a different  $C$  map. A series of test cases were used to determine which spectral parameters effect the  $C$  factor. The results are plotted in Figures 7.9, 7.10, 7.11, 7.12, and 7.13, separated into different  $f_{\text{hit}}$  bins. Each  $C$  factor plotted is averaged in a given declination bin for visualization purposes, since one would expect declination to be the spatial coordinate with the strongest dependence. Fortunately, Fig. 7.9 and Fig. 7.10 do not show any dependence on the  $\langle\sigma v\rangle$  value chosen for the spectrum. The  $C$ -factors can therefore be treated as independent of flux for the purpose of this analysis. All other tested variables show at least some effect on  $C$  and the analysis will therefore assign each tested spectrum and spatial profile its own unique  $C$  map to account for signal contamination.

## 7.6 Systematics

### 7.6.1 Detector Systematics

All detector Monte Carlo systematics considered in HAWC's 2019 observation of the Crab Nebula [41] and in Chapter 5 are considered here as well. These include PMT efficiency, threshold, charge uncertainty, and the broad pulse effect (listed individually in Table 5.4). These are taken into account by simply re-running the analysis (including the computation of the  $C$  factors) with the associated Monte Carlo detector response for each systematic. In addition, to account for the 10% flux systematic also cited in Ref. [41], an addition 10% uncertainty is added in quadrature onto the bands. As the  $C$  factors do not depend on the overall flux level, they do not need to be recalculated for the flux systematic. The aggregate effects of these systematics are summarized in Table 7.1 at the end of this chapter.

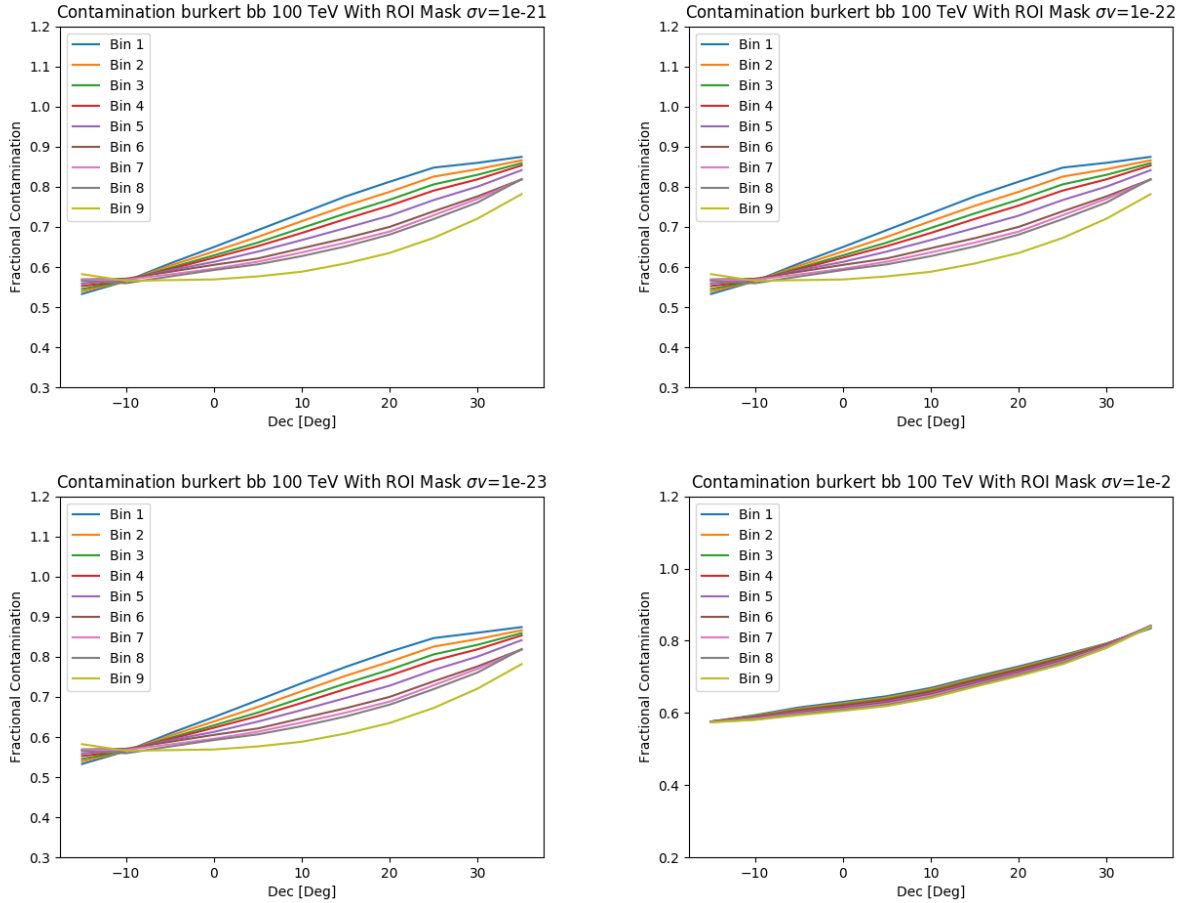


Figure 7.9: Expected contamination factors in each  $f_{\text{hit}}$  bin averaged in declination bands assuming  $\langle\sigma v\rangle$  values of  $10^{-21}$  (top left),  $10^{-22}$  (top right),  $10^{-23}$  (bottom left),  $10^{-2}$  (bottom right), all in units of  $\text{cm}^2\text{s}^{-1}$ . The expected contamination behaves identically for the different  $\langle\sigma v\rangle$  values except for one that is nearly twenty orders of magnitude outside the physically realistic parameter space (bottom right). Therefore, the contamination can be treated as independent of overall flux level.

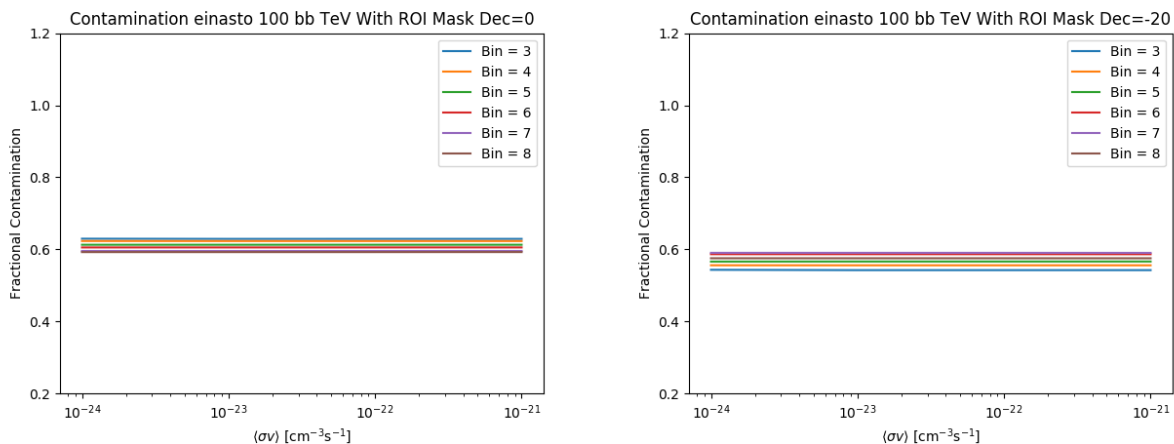


Figure 7.10: Expected contamination factors in each  $f_{\text{hit}}$  bin assuming an Einasto spatial profile, a 100 TeV dark matter mass, and a  $b\bar{b}$  spectrum as a function of  $\langle\sigma v\rangle$  for the highest (left) and lowest (right) declinations considered. These plots show no dependence on overall flux level, measured by choice of  $\langle\sigma v\rangle$ .

### 7.6.2 Pointing Systematic

In addition, the effect of a possible pointing systematic between HAWC and other experiments in explored. To address this, upper limits are recalculated in the most extreme case where the pointing error is entirely in the declination coordinate and takes the maximal realistic value of 0.5 degrees. This is implemented in the model by simply rotating the CLUMPY template by 0.5 degrees along the declination axis. The fit and limit calculations (including the estimation of the  $C$  factors) are then repeated with this modified template. Due to the extremely large size of the ROI in comparison to the expected pointing error size, this systematic is not expected to appreciably effect the results.

### 7.6.3 Spectral Systematics

The most substantial systematic is the assumed dark matter density profile. As detailed in Sec. 6.4 and 7.2, the flux has a very large systematic depending on the choice of density profile. As neither the Burkert (Eq. 7.2) or Einasto (Eq. 7.1) profiles are strongly favored by current observations of the Galactic Center, both will be reported separately with neither being considered the nominal case. HAWC is uniquely suited to probe the Galactic Halo in spite of this systematic. Because



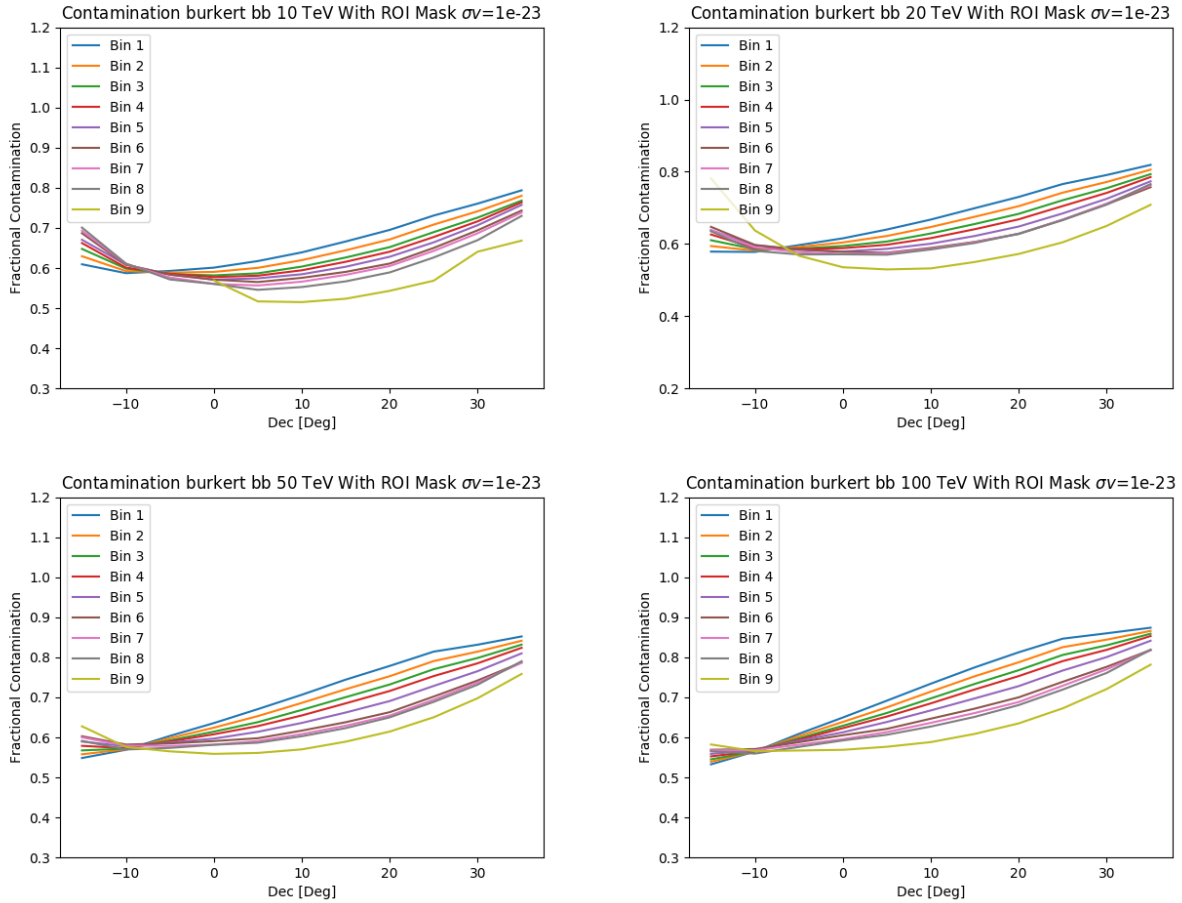


Figure 7.11: Expected contamination factors in each  $f_{\text{hit}}$  bin averaged in declination bands for spectra assuming 10 TeV (top left), 20 TeV (top right), 50 TeV (bottom left), and 100 TeV (bottom right) for the dark matter mass  $M$ . These plots hold the dark matter annihilation channel ( $b\bar{b}$ ) and spatial profile (Burkert) constant across all cases in order to isolate the effect of different assumed dark matter masses. These plots show the contamination does depend at least somewhat on the dark matter mass. Note that the discontinuous behavior in Bin 9 for the 10 TeV spectrum arises due to the relatively low energy cutoff, leaving very few gamma-rays that will reconstruct in the higher  $f_{\text{hit}}$  bins.

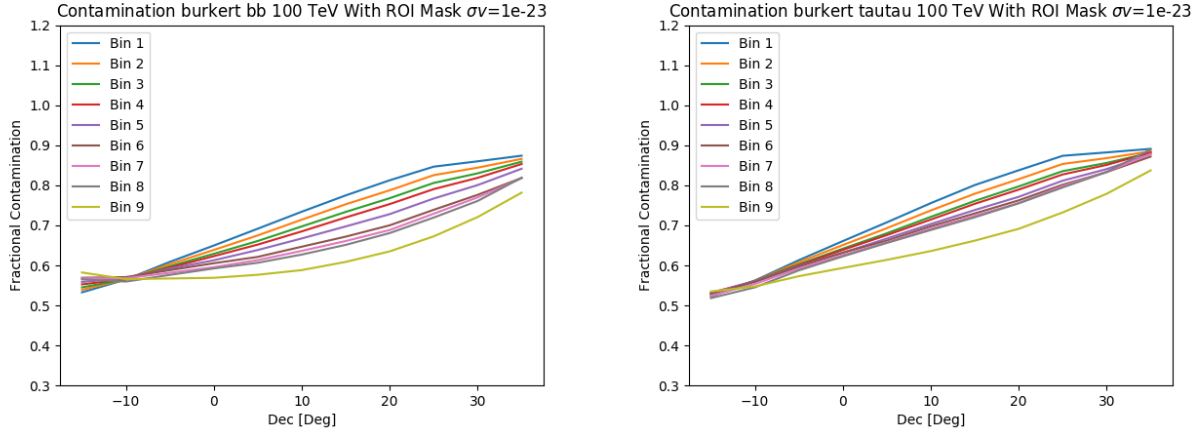


Figure 7.12: Expected contamination factors in each  $f_{\text{hit}}$  averaged in declination bands for the  $b\bar{b}$  (left) and  $\tau^+\tau^-$  (right) channels. The dark matter mass (100 TeV) and spatial profile (Burkert) are held constant between these plots. These two channels are chosen because they are, respectively, the softest and hardest spectra thereby probing the extremes of the channel parameter space. These plots show the contamination depends fairly strongly on the channel.

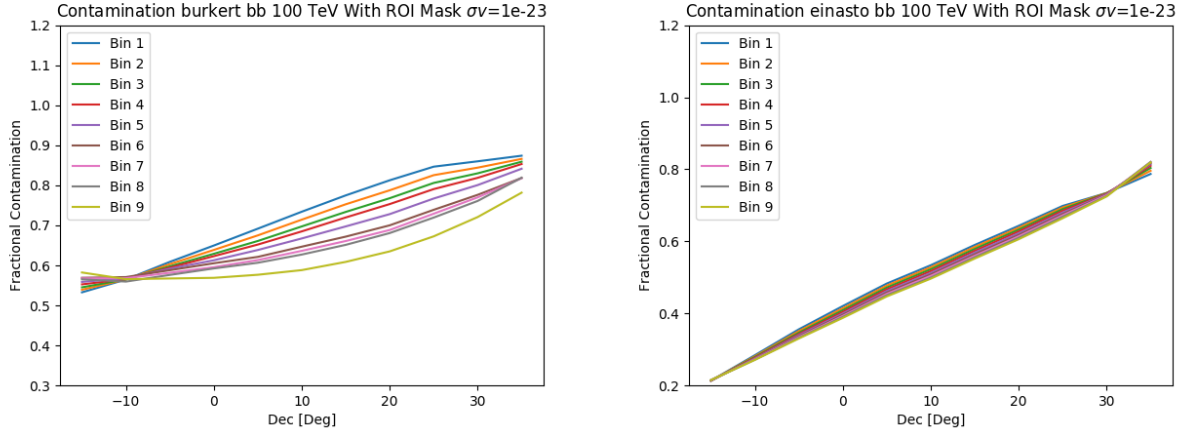


Figure 7.13: Contamination factors in each  $f_{\text{hit}}$  averaged in declination bands assuming a Burkert profile (left) and an Einasto profile (right), and holding the assumed mass at 100 TeV and annihilation channel to  $b\bar{b}$ . As expected, the assumed density profile has a strong effect on the contamination, with the strongly-cusped Einasto profile showing lower contamination than the cored Burkert profile.

HAWC is able to observe a larger region of the Halo away from the center, our limits are far less sensitive to the density profile behavior than observations focused directly on the Galactic Center. Finally, although this will not be reported as an additional systematic, a test case is considered with the electroweak correction removed in order to check that the effect of including it behaves as expected.

## 7.7 Results

### 7.7.1 Fit Results

With an ROI selected and known contamination removed, the results of the dark matter search are now shown. The significance ( $\sqrt{\text{TS}}$ ) values for various spectra are plotted in Fig. 7.14. No significant emission is found for any spectrum. Therefore, it is necessary to set 95% CL upper limits on the cross-section  $\langle\sigma v\rangle$  and 95% CL lower limits on the decay lifetime  $\tau$ .

### 7.7.2 Limits

The limits are set using the same likelihood-based limit-setting procedure as in Chapter 6. Again, the TS quantity will asymptotically follow a  $\chi$ -squared distribution with one degree of freedom. The limits are then set by increasing the free parameter corresponding to flux normalization (the cross-section  $\langle\sigma v\rangle$  for annihilation and reciprocal lifetime  $1/\tau$  for decay) from the maximum-likelihood value until the likelihood decreases by an appropriate amount for a given confidence level. To obtain 95% CL limits, 2.7 is again chosen as the appropriate  $\Delta\text{TS}$  value. For annihilation this again results in upper limits on  $\langle\sigma v\rangle$ . However, for decay this corresponds to lower limits on  $\tau$ , since it is inversely proportional to the flux normalization. The annihilation upper limits are shown in Fig. 7.15 and Fig. 7.16, while the decay lower limits are shown in Fig. 7.17 and Fig. 7.18

As expected, the density profile choice has the largest effect on the limits and is the dominant systematic. However, the difference is only on the order of a factor of 2-3, rather than the many orders of magnitude where the Galactic Center is the only part of the halo considered. Neither choice of profile is treated as the nominal case here, so all other systematics are computed independently for the limits arising in both cases. The uncertainty bands on Figs. 7.15, 7.16, 7.17, and 7.18 are

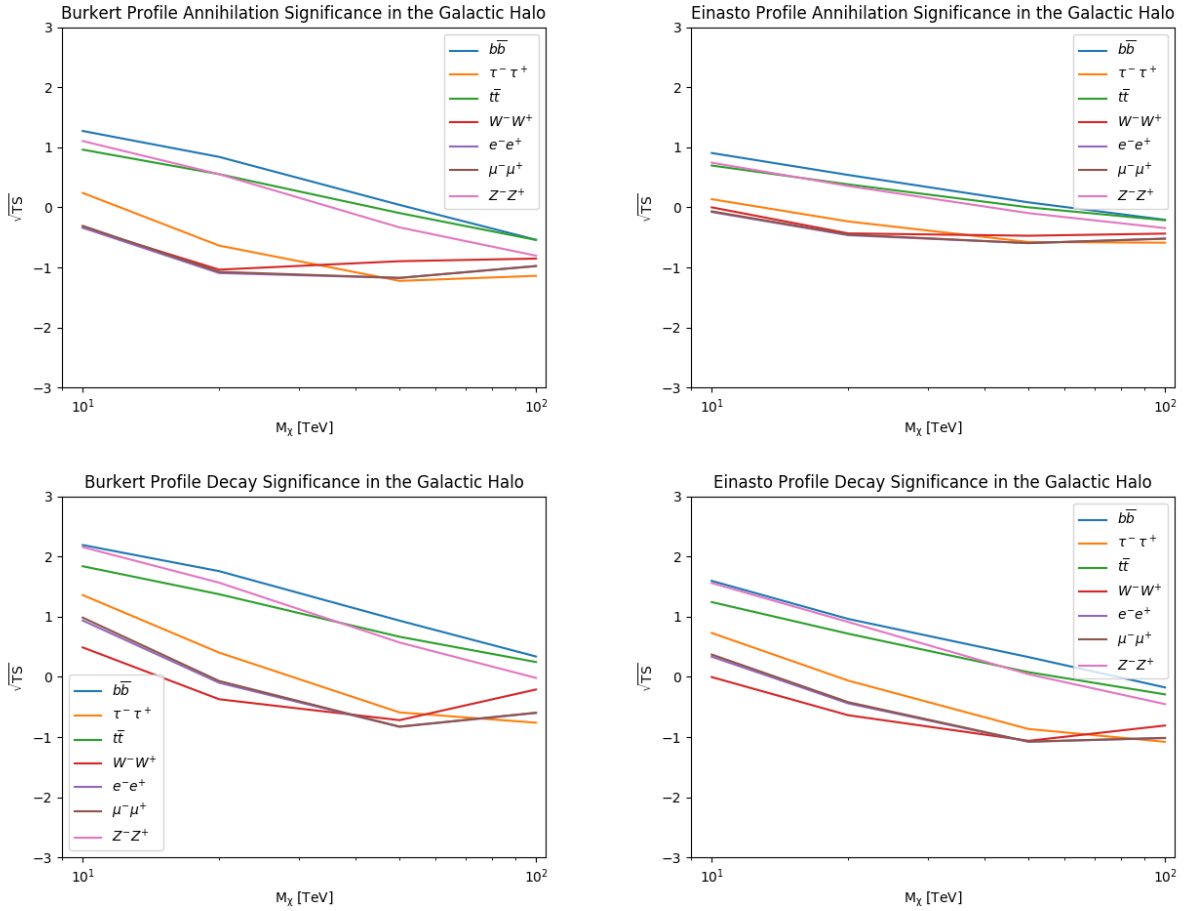


Figure 7.14: Significance ( $\sqrt{\text{TS}}$ ) as a function of dark matter mass and channel for annihilation assuming a Burkert profile (top left), annihilation assuming an Einasto profile (top right), decay assuming a Burkert profile (bottom left), and decay assuming an Einasto profile (bottom right). No spectrum shows significant evidence of gamma-ray emission, although a small positive significance is shown at the lower masses due to the residual unidentified excess in the lower bins.

computed by simply adding the contribution from each systematic in quadrature. This quadrature addition is done separately for the positive and negative halves of the bands. As the Monte Carlo contains many variations for certain simulation parameters, in particular the PMT thresholds (see Table 5.4), only the largest positive and negative changes from a given detector systematic are considered as contributions. The overall contribution of the Monte Carlo to the systematics is summarized in Table 7.1.

As the steeper Einasto profile is expected to be the most sensitive to small changes in the template, this will be the test case for the pointing systematic. The  $\tau^+\tau^-$  annihilation channel is

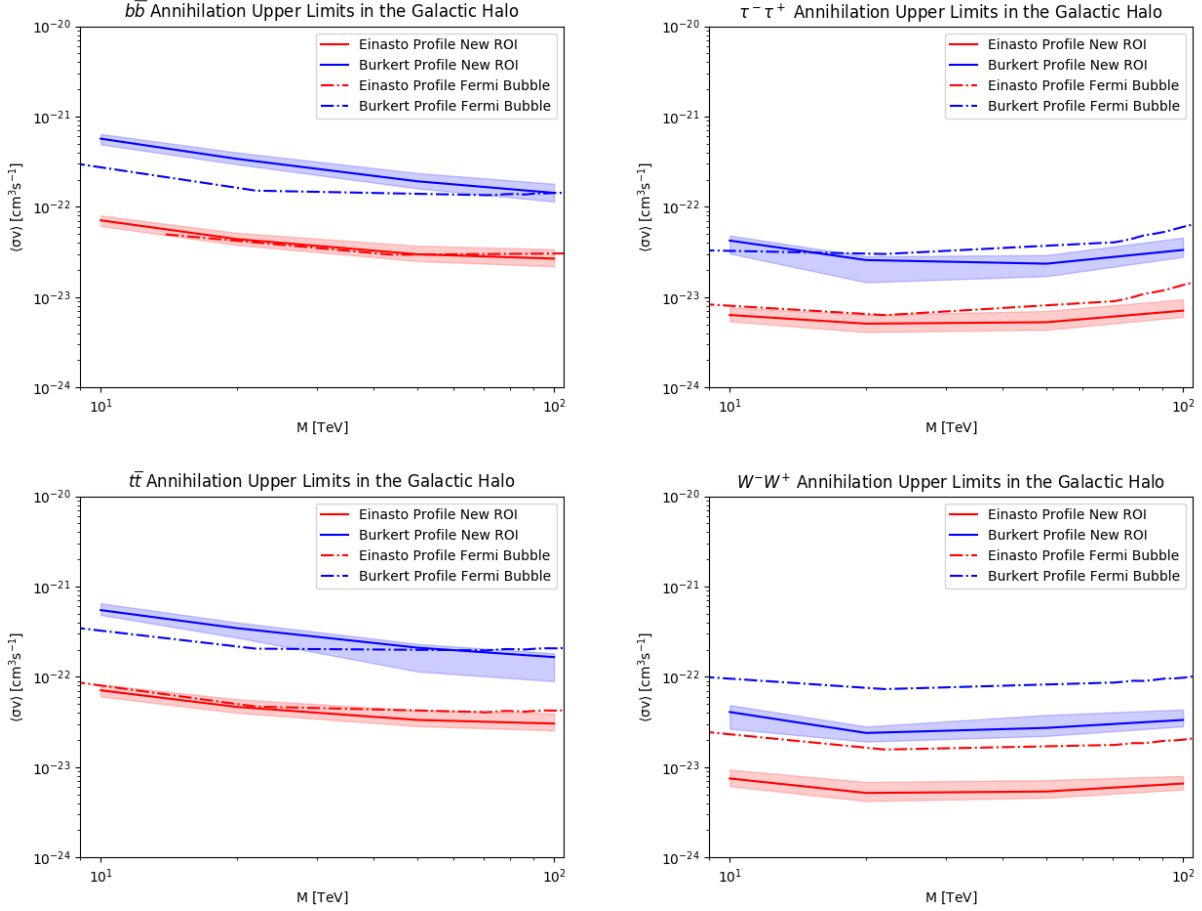


Figure 7.15: 95% CL upper limits on the velocity-weighted annihilation cross-section  $\langle\sigma v\rangle$  for the  $b\bar{b}$  (top left),  $\tau^+\tau^-$  (top right),  $t\bar{t}$  (bottom left), and  $W^+W^-$  (bottom right) dark matter spectra assuming the Einasto (red) and Burkert (blue) spatial profiles. The correspondingly colored shaded regions are the systematic uncertainty bands. The corresponding limits from the previous HAWC Galactic Halo analysis [24] are also plotted for comparison.

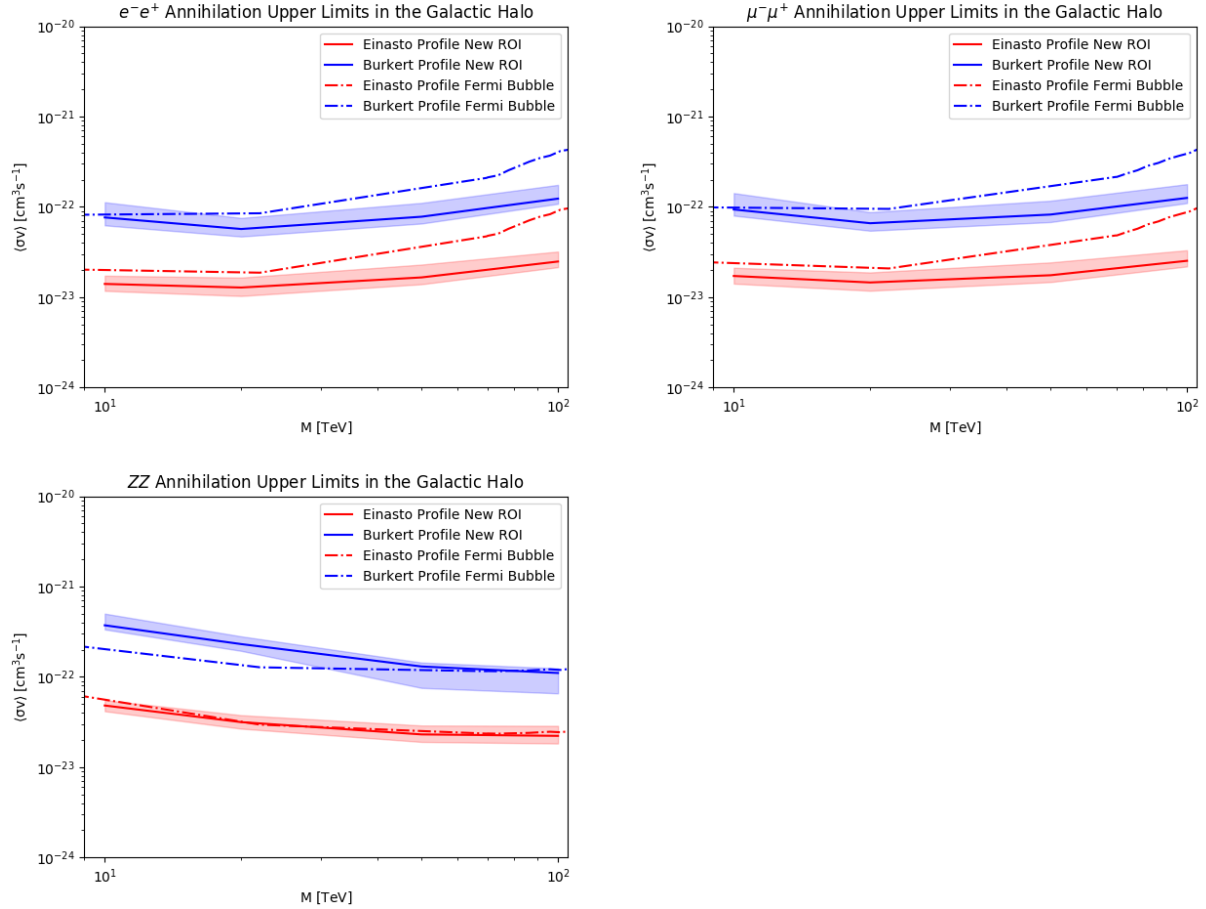


Figure 7.16: Additional 95% CL upper limits on the velocity-weighted annihilation cross-section  $\langle\sigma v\rangle$  for the  $e^+e^-$  (top left),  $\mu^+\mu^-$  (top right), and  $Z^0Z^0$  (bottom) dark matter spectra assuming the Einasto (red) and Burkert (blue) spatial profiles. The correspondingly colored shaded regions are the systematic uncertainty bands. The corresponding limits from the previous HAWC Galactic Halo analysis [24] are also plotted for comparison.

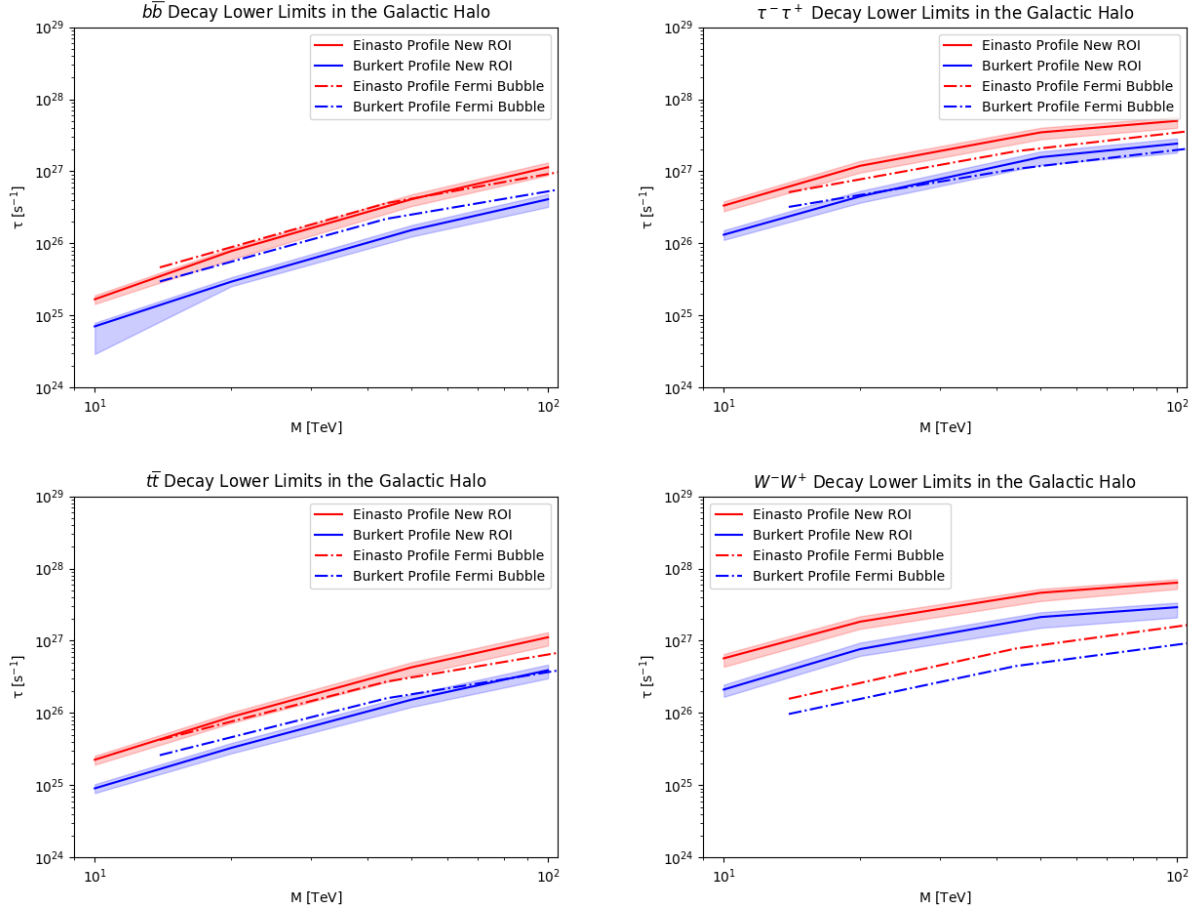


Figure 7.17: 95% CL lower limits on the decay lifetime  $\tau$  for the  $b\bar{b}$  (top left),  $\tau^+\tau^-$  (top right),  $t\bar{t}$  (bottom left), and  $W^+W^-$  (bottom right) dark matter spectra assuming the Einasto (red) and Burkert (blue) spatial profiles. The correspondingly colored shaded regions are the systematic uncertainty bands. The corresponding limits from the previous HAWC Galactic Halo analysis [24] are also plotted for comparison.

chosen for the energy spectrum, due to it having the highest expected flux in the higher  $f_{\text{hit}}$  bins where the HAWC PSF is narrowest compared to the pointing error. The resulting upper limits computed with a 0.5 degree shift downwards in declination of the spatial profile are shown compared with the unshifted limits in Fig. 7.19. As expected, the limits are nearly identical to those without the shif, having an effect of less than 1%. It is therefore concluded that the pointing systematic has a negligible effect on these results compared to the other systematics (see Table 7.1).

Finally, the effect of removing the electroweak correction on the  $W^+W^-$  and  $b\bar{b}$  spectra is tested. The results are plotted in Fig. 7.20. As expected, only the  $W^+W^-$  spectrum is strongly

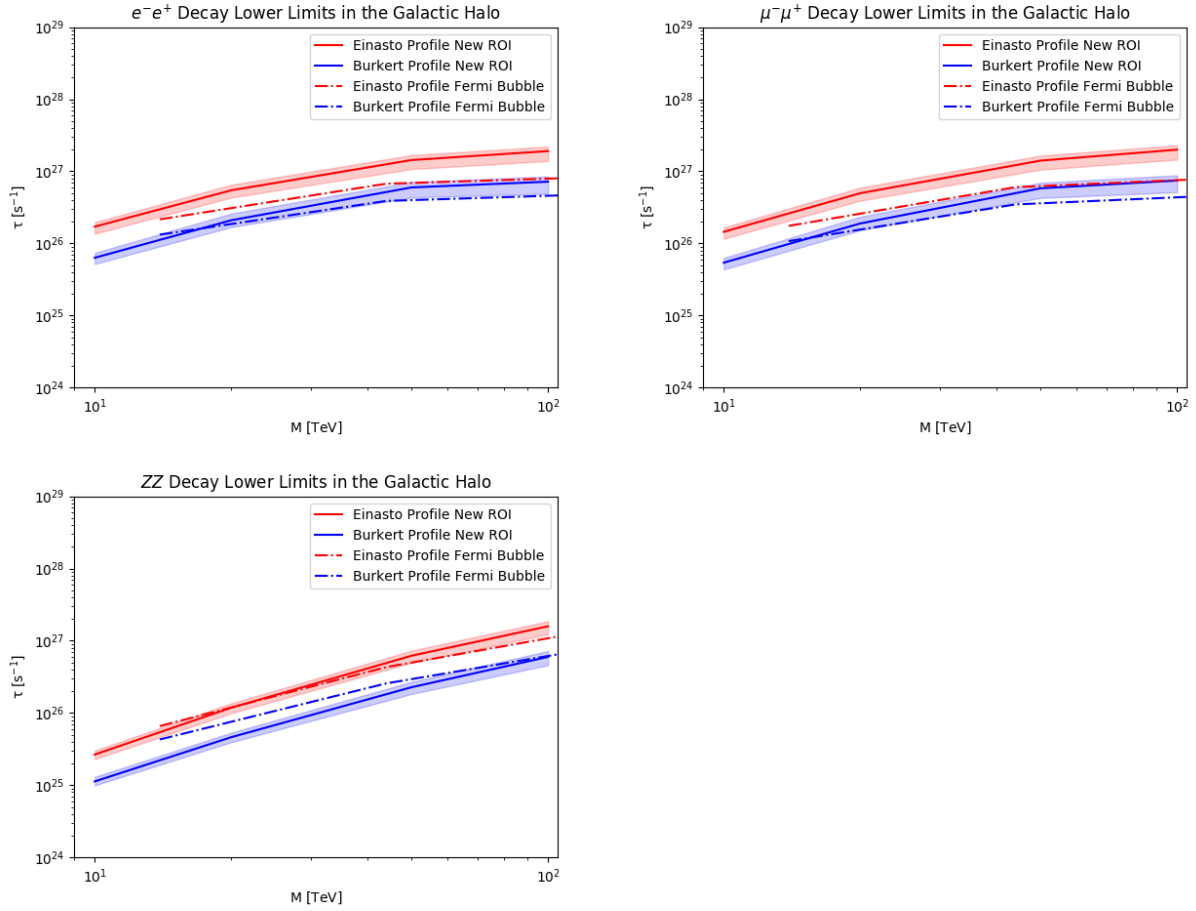


Figure 7.18: Additional 95% CL lower limits on the decay lifetime  $\tau$  for the  $e^+e^-$  (top left),  $\mu^+\mu^-$  (top right), and  $Z^0Z^0$  (bottom) dark matter spectra assuming the Einasto (red) and Burkert (blue) spatial profiles. The correspondingly colored shaded regions are the systematic uncertainty bands. The corresponding limits from the previous HAWC Galactic Halo analysis [24] are also plotted for comparison.

driven by the electroweak correction due the presence of a strong secondary peak. The  $b\bar{b}$  spectrum shows only a small change driven by the slightly increased expected photon count.

### 7.7.3 Statistical Effects

Finally, the effects of statistical fluctuations on the limits are computed. For this process 500 pseudomaps are generated, where the data in each bin is replaced with a randomly drawn value from a Poisson distribution with a mean equal to the estimated background in that bin. The 95% upper and lower limits are recomputed using these pseudomaps in place of the actual data.



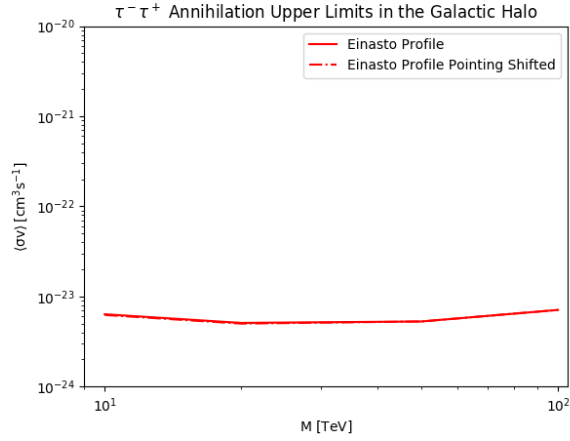


Figure 7.19: Comparison of 95% upper limits when the spatial template is shifted in declination by the pointing systematic. The limits are indistinguishable at the higher masses and minimally effected at the lower dark matter masses. Therefore the pointing effect is treated as negligible.

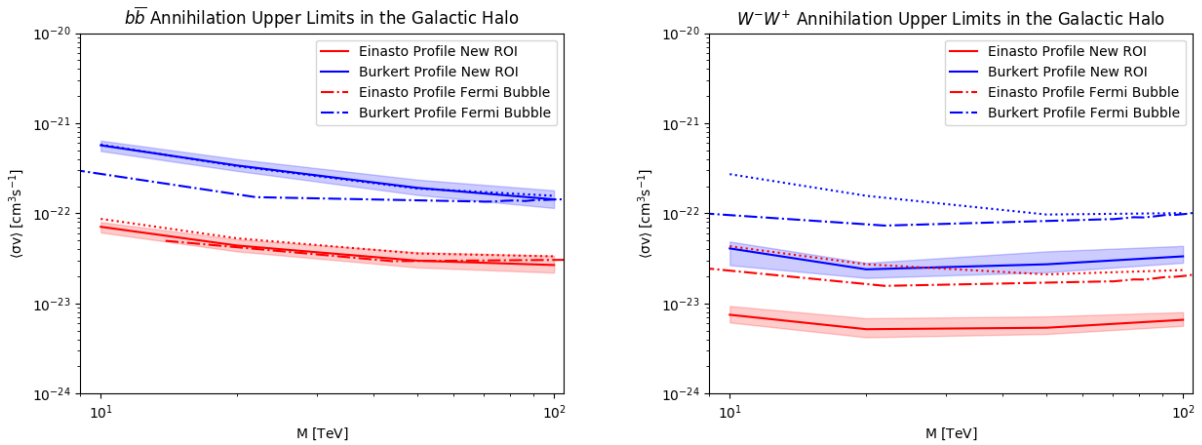


Figure 7.20: Comparison of dark matter annihilation upper limits with and without the electroweak correction for the  $b\bar{b}$  (left) and  $W^+W^-$  (right) channels. As expected, the  $W^+W^-$  channel shows a strong improvement with the EW correction, while the  $b\bar{b}$  channel is only weakly effected by it.

<b>Systmatic</b>	<b>Annihilation</b>	<b>Decay</b>
<b>Density Profile</b>	<b>factor of <math>\approx 2.5</math></b>	<b>Factor of <math>\approx 1.5</math></b>
<b>Simulation</b>	<b>-11% to +23%</b>	<b>-23 % to +11%</b>
<b>Flux Scale</b>	<b>-10% to +10%</b>	<b>-10% to +10%</b>
<b>Pointing</b>	<b>less than 1%</b>	<b>less than 1%</b>

Table 7.1: Summary of systematic uncertainty sizes on the 95% CL limits for annihilation and decay. The dominating systematic in both cases is the choice of density profile, followed by systematics arising from detector simulation. Note that these are only estimates, and the actual contribution of each varies from spectrum to spectrum. No overall uncertainty is reported since neither choice of density profile is the used as the nominal case.

The median of the resulting distribution results in “expected” limits; i.e. the limits one would expect assuming the null hypothesis of a background-only model is true. As in Chapter 6, the 68% and 95% containment bands are computed about this expectation. Assuming the background was well-modeled in this analysis, the limits should be no further from the expected case than the 95% band, being influenced only by Poisson-fluctuations rather than an undetected source or mismodeled background.

As this process is computationally intensive, the statistical bands for only two sample spectra,  $b\bar{b}$  and  $\tau^+\tau^-$ , are plotted for all masses and spatial profiles. These are, respectively, (most low-energy dominated) and hardest (most high-energy dominated) channels and therefore span the extreme edges of the channel parameter space. The results are plotted for annihilation and decay, as well as the two spatial profiles considered, in Figs. 7.21, 7.22, 7.23, and 7.24. All of the plotted limits fall well within the 95% band, and is therefore concluded that the limits are consistent with expectation. The expected limits from the previous HAWC analysis of the Fermi Bubble region are also plotted for comparison when available (Figs. 7.22 and 7.24). To further verify the limits are within expectation across all channels, a complete set of plots was made for the Einasto profile annihilation case and is shown in Figures 7.25 and 7.26 at the end of this chapter.

## 7.8 Conclusion

Although no significant evidence of dark matter gamma-ray emission was found, the new constraints on dark matter annihilation and decay derived from this analysis show a marked improvement over

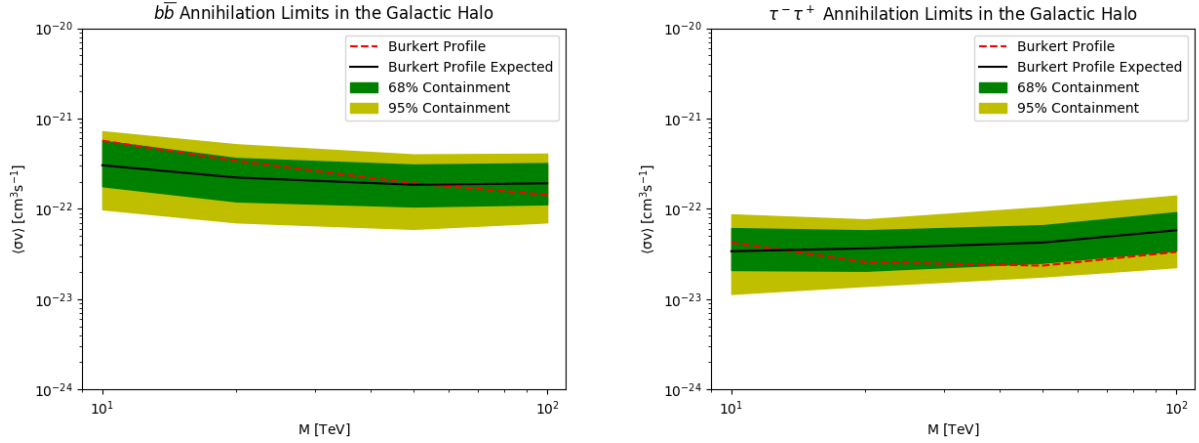


Figure 7.21: Statistical bands for 95% CL upper limits on dark matter annihilation into the  $b\bar{b}$  (left) and  $\tau^+\tau^-$  (right) channels assuming a Burkert spatial profile. In all cases, the actual limits fall well within the expected range from statistical fluctuations.

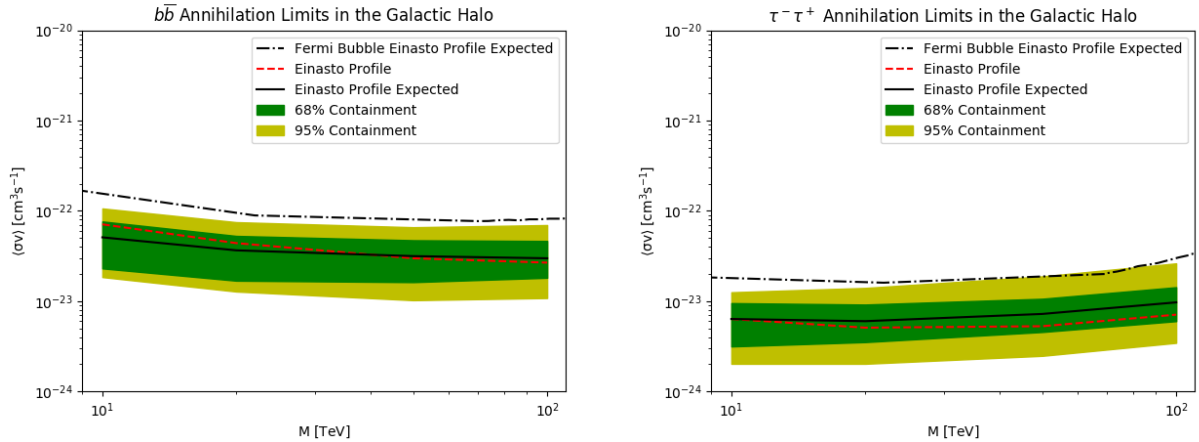


Figure 7.22: Statistical bands for 95% CL upper limits on dark matter annihilation into the  $b\bar{b}$  (left) and  $\tau^+\tau^-$  (right) channels assuming an Einasto spatial profile. In all cases, the actual limits fall well within the expected range from statistical fluctuations. The expected limits from the previous HAWC (Fermi Bubble) analysis [24] are also plotted to more clearly show the improvement in sensitivity from this analysis independent of statistical fluctuation effects.

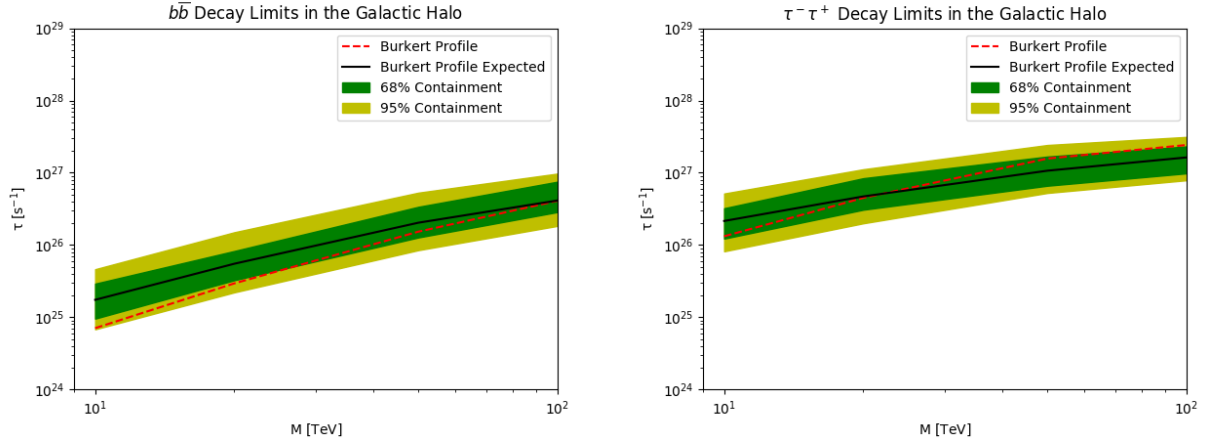


Figure 7.23: Statistical bands for 95% CL lower limits on dark matter decay into the  $b\bar{b}$  (left) and  $\tau^+\tau^-$  (right) channels assuming a Burkert spatial profile. In all cases, the actual limits fall well within the expected range from statistical fluctuations.

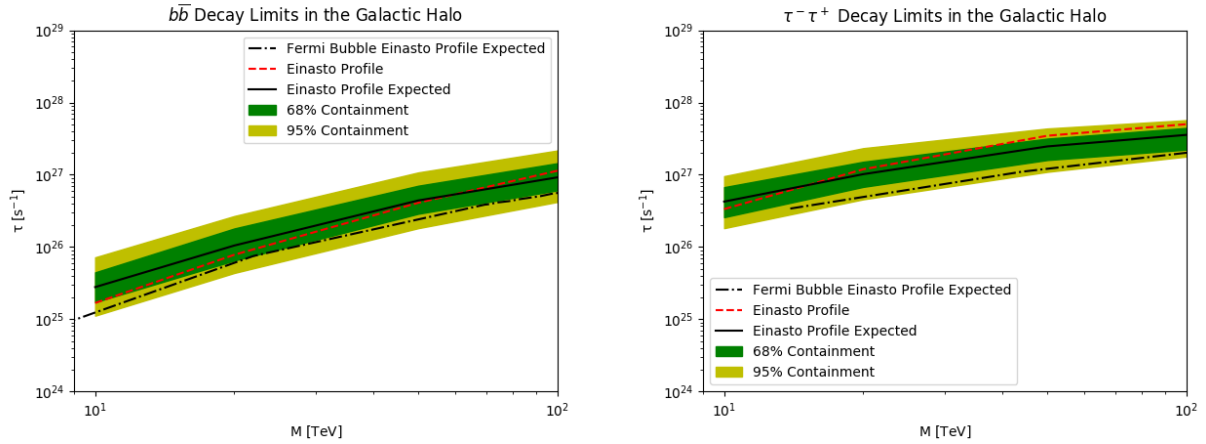


Figure 7.24: Statistical bands for 95% CL lower limits on dark matter decay into the  $b\bar{b}$  (left) and  $\tau^+\tau^-$  (right) channels assuming an Einasto spatial profile. In all cases, the actual limits fall well within the expected range from statistical fluctuations. The expected limits from the previous HAWC (Fermi Bubble) analysis [24] are also plotted to more clearly show the improvement in sensitivity from this analysis independent of statistical fluctuation effects.

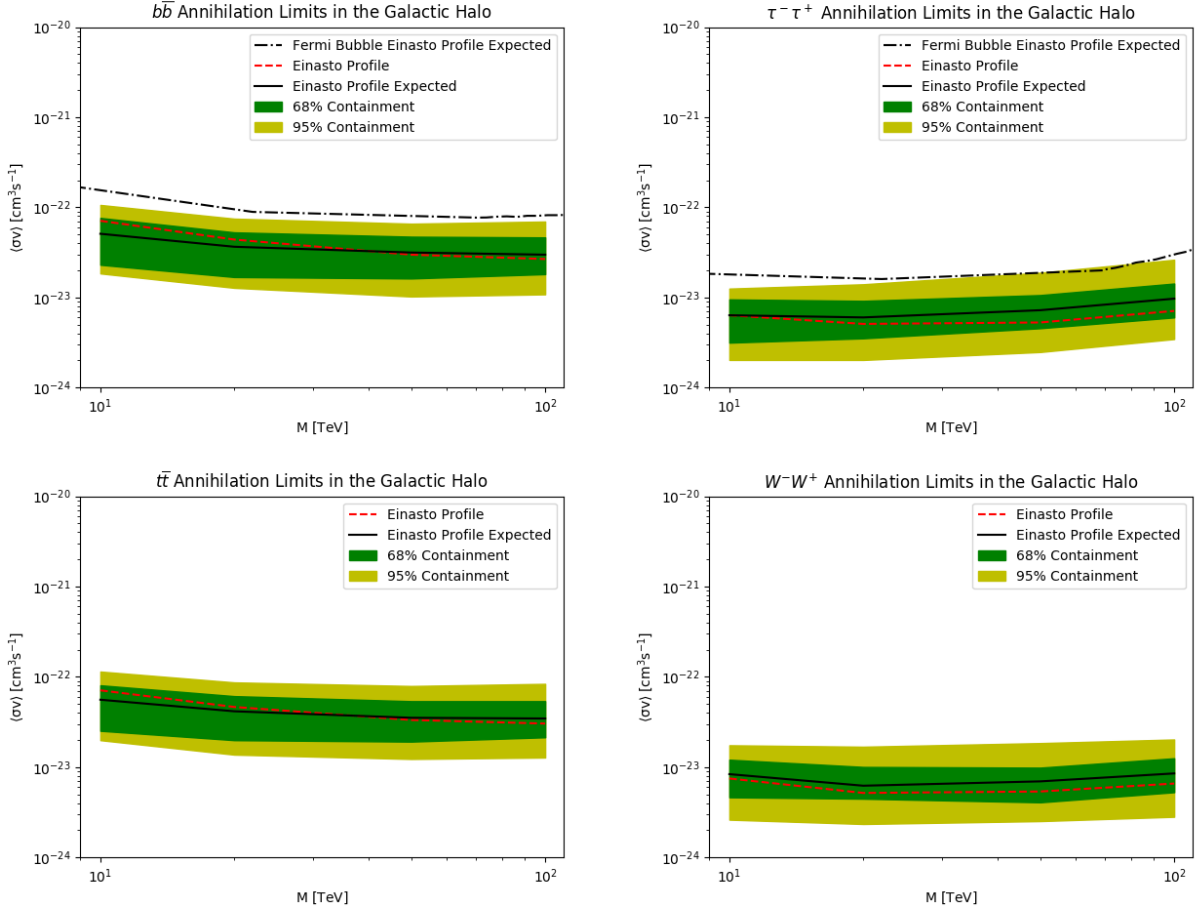


Figure 7.25: Statistical bands on the 95% CL upper limits on the velocity-weighted annihilation cross-section  $\langle\sigma v\rangle$  for the  $b\bar{b}$  (top left),  $\tau^+\tau^-$  (top right),  $t\bar{t}$  (bottom left), and  $W^+W^-$  (bottom right) dark matter spectra assuming an Einasto profile. This is a supplementary figure meant to demonstrate the all channels are within statistical expectations for a given case.

the previous published results in Ref. [24]. In addition, the inclusion of the electroweak corrections dramatically improves the results for the  $W^+W^-$  channel due to the higher relative flux at the highest energies. Furthermore, the background is now better estimated and does not contain the strong negative flux artifacts seen in [24]. As expected, the limits are relatively robust against the large uncertainties arising from the density profile choice. Rather than the multi-order-of-magnitude systematic that one would expect if only the Galactic Center were considered (see Fig 7.1), these results differ by less than a single order of magnitude between the profiles.

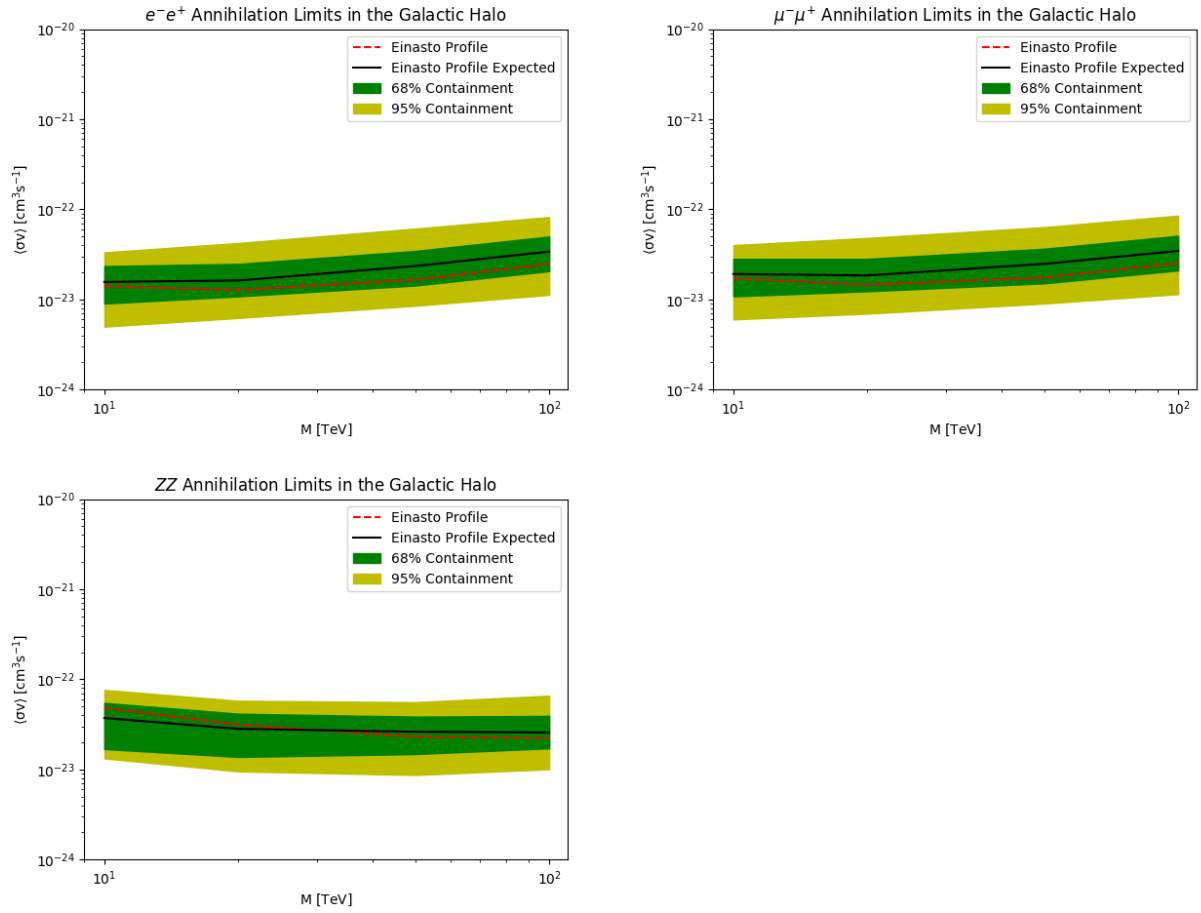


Figure 7.26: Additional statistical bands on the 95% CL upper limits on the velocity-weighted annihilation cross-section  $\langle\sigma v\rangle$  for the  $e^+e^-$  (top left),  $\mu^+\mu^-$  (top right), and  $Z^0Z^0$  (bottom) dark matter spectra assuming an Einasto profile. This is a supplementary figure meant to demonstrate that all channels are within statistical expectations for a given case.

## CHAPTER 8

### SUMMARY

This dissertation used data from the HAWC detector to probe cutting-edge physics beyond the Standard Model. The techniques by which HAWC is able to detect cosmic gamma rays were demonstrated and the many advantages of HAWC in probing ultra-high energy gamma-ray physics were detailed. It was shown how HAWC data can be used to explore unanswered questions such as the nature of dark matter and the limits of Lorentz invariance. In particular, a search for evidence of WIMP dark matter in the Milky Way Galactic Halo was performed. To accomplish this, simulations of the dark matter density profile were combined with estimates of the HAWC sensitivity to dark matter-like energy spectra. This allowed strong constraints on dark matter annihilation and decay from the Galactic Halo to be derived that are insensitive to the large uncertainties arising from systematics in the dark matter spatial distribution. Multi-hundred TeV photon spectra were also significantly detected from HAWC sources within the Galactic Plane. These results lead to the strongest constraints on Lorentz invariance violation to be published at the time of writing.

The work of this dissertation was made possible by the ongoing development of new algorithms and reconstruction techniques within the HAWC collaboration. Probing the Galactic Halo required the creation of a novel background estimation technique that relied on HAWC's wide field of view and strong ability to discriminate between gamma rays and cosmic rays. Meanwhile, the constraints on Lorentz invariance violation were enabled by the improved energy resolution from a machine learning technique. HAWC has recently completed a reprocessing of all archival data using an updated set of algorithms that can lead to compelling follow-up work on these results. Combining the new background technique with the re-optimized energy estimators will allow for Galactic dark matter to be probed at even higher masses, as well as for analyses that require precise energy resolution such as gamma-ray line searches.

Even more improvements can be made to these analyses by the inclusion of the outrigger tanks constructed around HAWC. A new reprocessing of HAWC data incorporating the outriggers into the reconstruction is currently in development. This will extend the dynamic range of HAWC to

even higher energies and allow for even better energy resolution, allowing HAWC to probe Lorentz invariance further. This improved energy resolution will also lead to stronger constraints on dark matter from the Galactic Halo at high masses, especially when combined with the improved energy estimators. In addition to these changes, HAWC is constantly collecting events and adding to its data archive, meaning its sensitivity to all classes of astrophysics will only improve with time.



## APPENDIX

## REMOTE CONTROL AND MONITORING

Maintaining HAWC's near 100% duty cycle requires a robust system for remote monitoring. Monitoring data must be transferred off site in an efficient manner that is robust to potential network outages. On-duty shifters must be able to view diagnostic information in a clear, user-friendly way that updates in real time so that experts can be notified of potential issues. These experts also require readily available tools to obtain more in-depth information, as well as archival monitoring data, to check for abnormal behavior and transients.

Two systems exist to meet these goals. Data collection is handled by the Advanced Tracking of HAWC Experiment Notifications and Alerts (ATHENA) system. This is a Python library designed for polling the experiment hardware and storing the information in MySQL [80] database tables. While the polling routine must be designed independently for each monitored component, ATHENA automatically structures the user-provided fields in a standardized way and comes with a built-in version control system should the definition of a data stream change. This allows new components to be added to the monitoring logic by simply adding a new polling script that registers the new data stream with the main ATHENA server and creates a data table for it. Previous monitoring systems were based on ASCII files with no standardized format, which often caused issues with parsing data. ATHENA deals with all formatting on the back end and guarantees any new data streams will be consistent and accessible. Most components are polled on the order of every minute, giving an effectively real-time view of the detector's health. Insertion into the database is then passed by a ZeroMQ [81] wrapper within the ATHENA server library. Approximately 8 Gb of data are written to the database per year.

ATHENA also handles transportation of monitoring data to the off-site computers where it can be easily accessed. This is done using a database synchronization routine custom-written in Python for ATHENA. This routine checks both the on-site and off-site versions of a given table and compares the time column of each row. It will then write a SQL query to copy only entries more recent than the latest entry in the off-site database, preventing unnecessary load on the network. Tables are synchronized sequentially one at a time in two loops; one for smaller tables and one for larger tables. The synchronization process is launched and completed roughly every minute, keeping

the off-site information updated at nearly the same cadence as the monitoring polling is performed. Any attempt to synchronize will time out and re-launch with approximately the same frequency (with a more permissive timeout for the larger tables), preventing a transient bad connection or slow table from interrupting the entire monitoring process. In the event of a longer network outage, a more brute-force script can be run by hand to move data off-site in smaller packets and catch up the monitoring more efficiently than a single synchronization pass. These optimizations mean only around 20 kilobytes of data must be sent during each synchronization instance to keep everything up to date, which is well within the tolerance of the network. When originally designed, this process was further complicated by an unreliable microwave link internet connection that was prone to outages. Unexpected outages could cause the sync routines to hang and would require they be restarted. A second master script run through the Crontab was implemented to automatically kill and restart a hanging sync script, as well as clean up duplicate sync processes. This process has become mostly redundant with the advent of a more reliable cable connection, but still exists for the sake of safety.

A user-facing display of the monitoring is handled by the HAWC Observatory Monitoring for Experiment and Reconstruction (HOMER) system. This a collection of webpages and apps generated using the PHP language for preprocessing HTML [82]. It uses the built-in SQL wrapper libraries for PHP to access the ATHENA database and sort data into arrays. The HOMER homepage contains a summary of the most critical site components as well as links to subpages with more details on individual aspects. The homepage is meant to be intuitive even to novice users and display the most critical information up front, with color codes to indicate whether a component is in a normal state or potentially problematic one. Critical information on the homepage includes the run status (for both the main DAQ and the outrigger nodes), electronics' temperatures, scaler rates (abnormalities in which indicate a possible detector malfunction), and overall status of the high voltage, as shown in Fig. A.1. HOMER also displays the age of the data in each stream, letting users know how up-to-date the current state is with more color-coded text if the update cadence is slower than expected.

The more detailed sub-pages for various monitoring streams rely on a Google Charts API for plotting [83]. While this API introduces an external dependency, it provides a very convenient

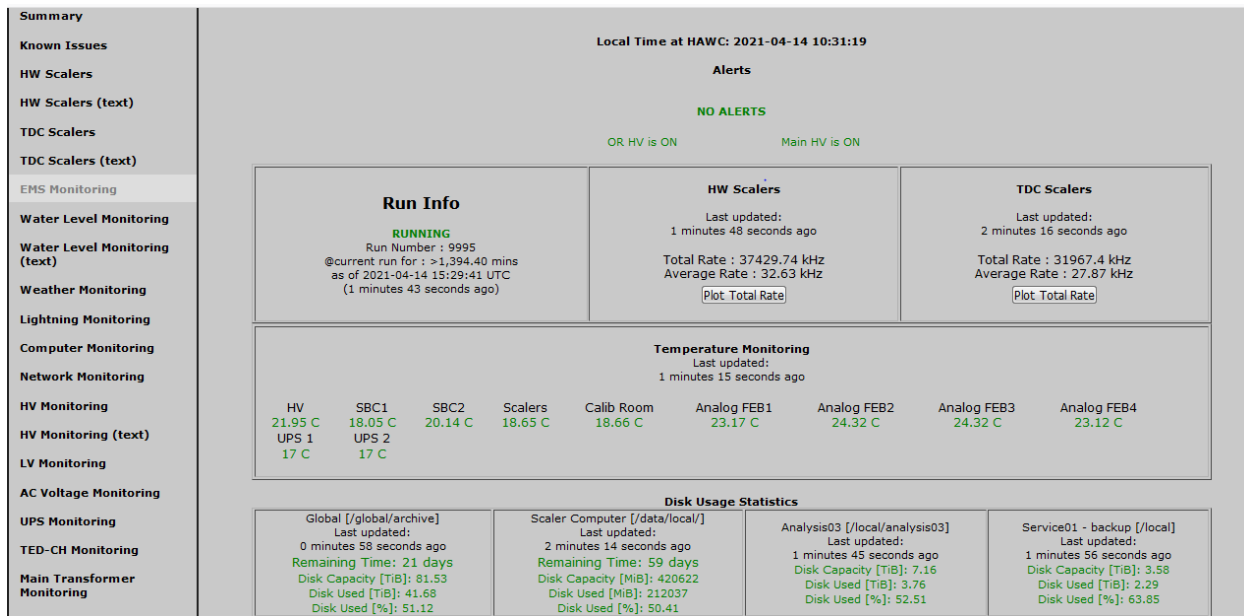


Figure A.1: A screenshot of the main HOMER home page. It includes all essential summary information such as total scaler rates and temperatures. The tabs on the left hand side are links to pages with further monitoring details.

way of making well-formatted plots that are optimized to be readable. Google Charts also includes routines for making plots dynamic, allowing users to easily zoom in or out and get the exact value of a data point by dragging the mouse over it. Each dedicated page also contains a menu for drawing plots with custom time-ranges as inputted by a user in order to easily access older data to take a long-term view of hardware performance when needed. See Fig. A.2 for a sample sub-page showing the various features available to users.

A second version of HOMER exists at the HAWC site with reduced features. It is meant to be both lighter-weight and independent of the Google APIs, allowing it to run even if the network connection is down albeit without the user-friendly features of Google Charts. Since it is only meant to be used by the crew and other experts on site, it is optimized for reliability and conciseness rather than accessibility by inexperienced users. The homepage contains only the critical information component and the side tabs link only to pages that only need to be accessed for on-site troubleshooting, such as scaler rates and electronics status. Since it runs on the same machine as the on-site database, it does not need to wait for a database sync to complete before updating and it is in real-time up to the polling frequency of the various ATHENA scripts. This

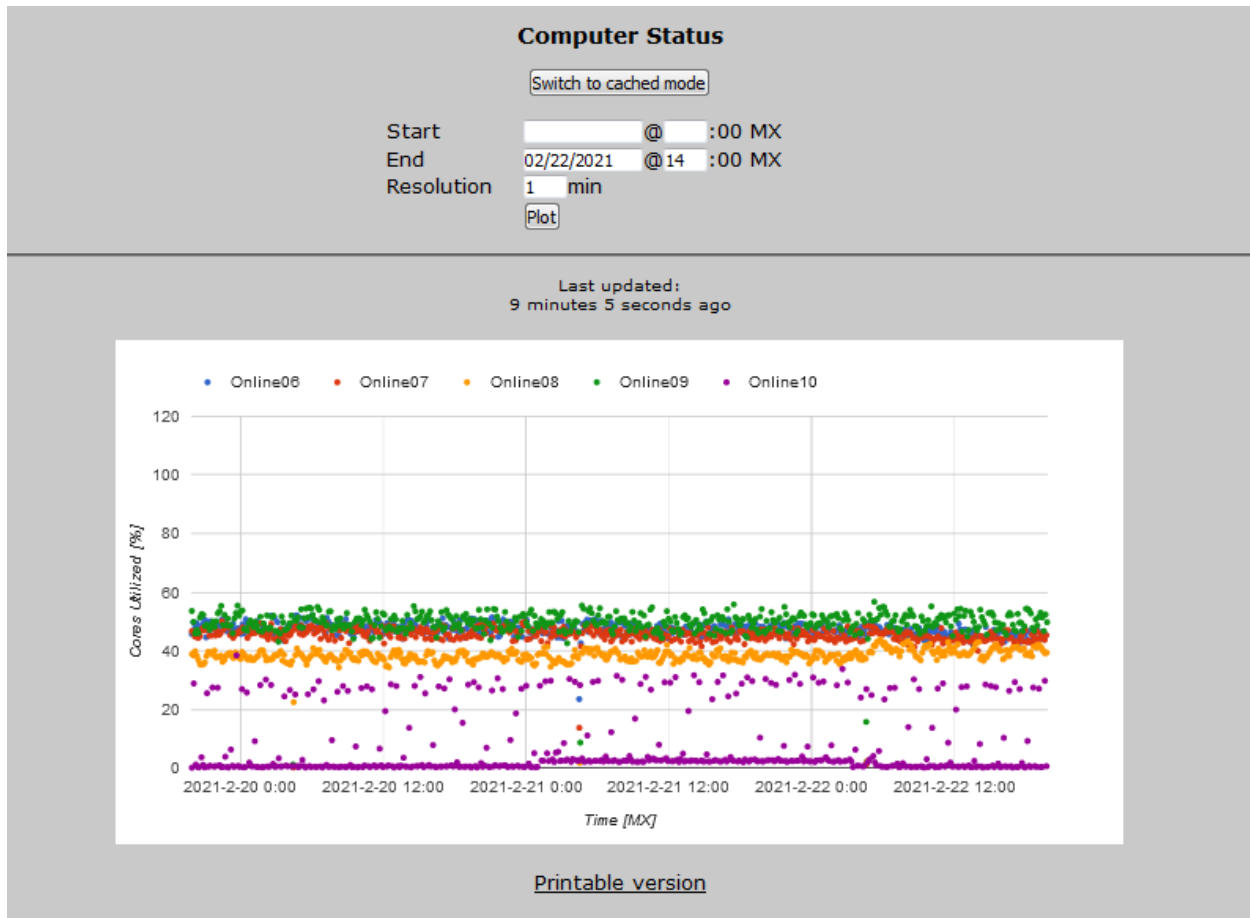


Figure A.2: Sample HOMER sub-page plot showing the CPU load on the site computers. Included also are fields for plotting a custom time range and the option to download a printable version. Many plots such as this exist for reading historic monitoring data.

system has proven vital during extended network outages so anyone physically at site is able to quickly diagnose the detector state. The plots here are generated by launching Python scripts once the SQL data is parsed, generating static figures. It also contains pages with JavaScript functions to turn high-voltage channels on and off. These pages directly access the high voltage modules to display status when loaded and have no intermediate dependency on the database polling. This allows any crew at site to quickly toggle high-voltage channels and view their status without relying on knowledge of the underlying scripts (see Fig. A.3 for a view of the control pages). These features are only available in the on-site version of HOMER to minimize the chance of an accidental channel trip by an off-site user or tampering in the event of a security breach of the off-site computers.

In addition to the HOMER system, a second web page known as the Online Plots is used for

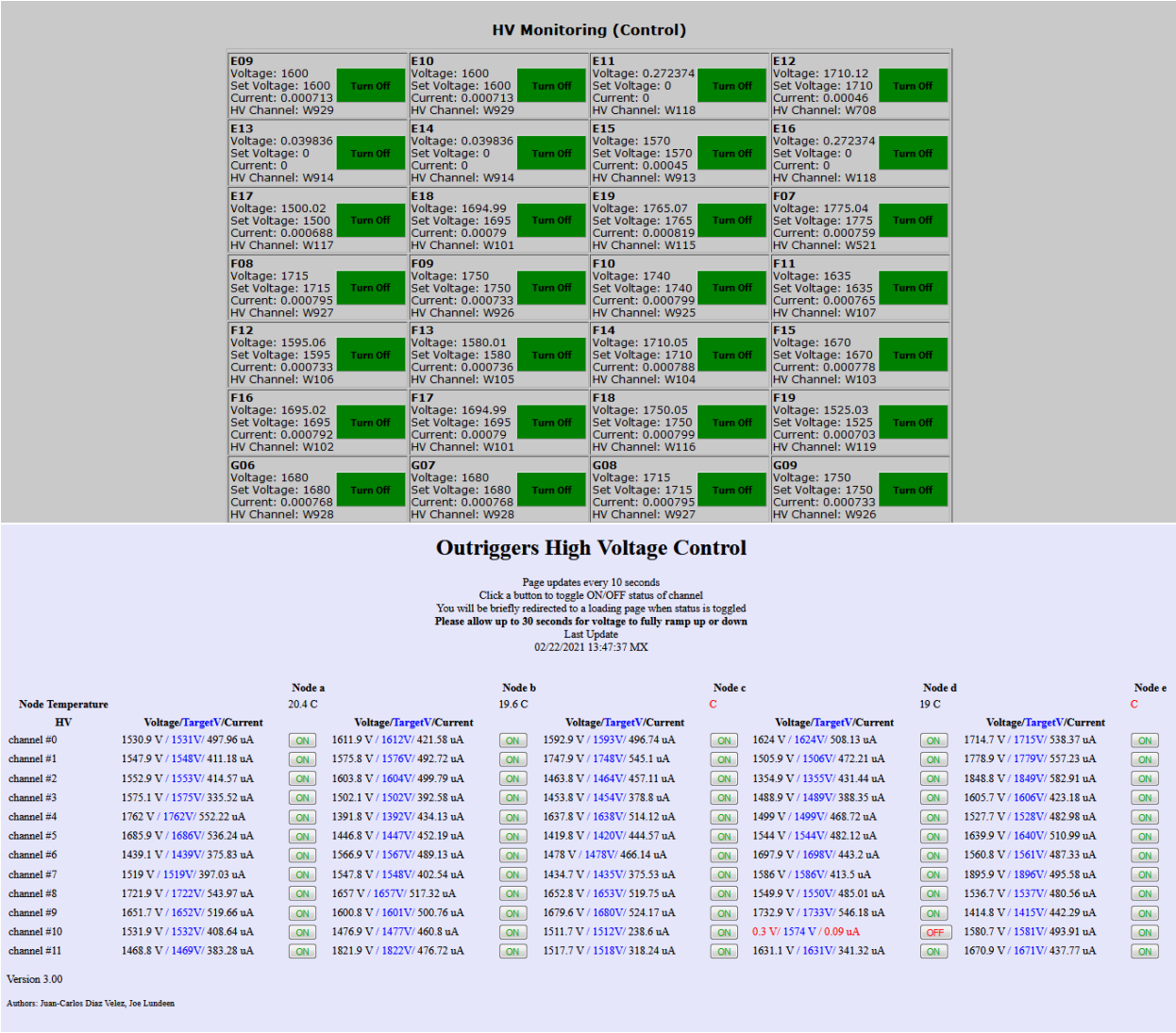


Figure A.3: High voltage control pages for both the main array(top) and the outriggers (bottom). The buttons on these pages allow for channels to be easily toggled on and off. This page exists only in the on-site version of HOMER to prevent access by non-experts.

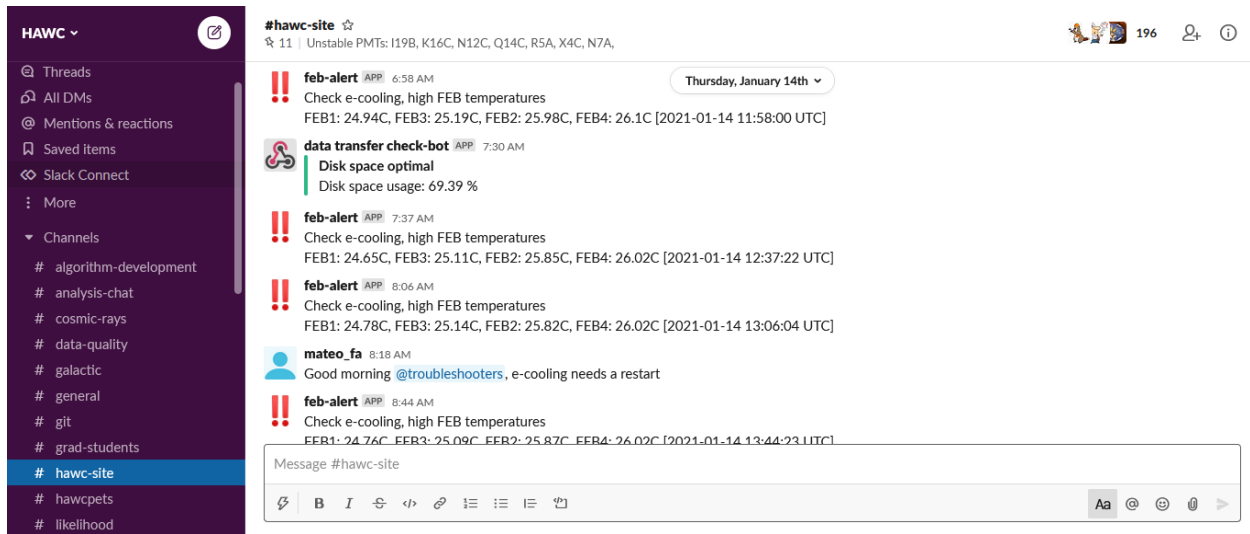


Figure A.4: Sample Slack alert sent by one of the many automatic monitoring scripts at the HAWC site. This alert indicates a FEB is running hot and includes a timestamp of when the high temperature was detected. To avoid saturating the channel, alerts are sent with a decreasing frequency if the same problematic state is detected multiple polls in a row.

monitoring. These pages consist of static figures and HTML text generated at site and transferred to the off-site monitoring using rsync. The plots on this page are meant to provide a lighter-weight snapshot of the detector status as well as give details from the online analysis such as the current significance of bright sources on a given day. It also includes displays of the main array and outrigger statuses and more technical diagnostic plots of various electronic components.

Another component of HAWC remote monitoring is the Slack alert system. This consists of a series of Python processes that poll various DAQ components in much the same way as the ATHENA scripts. However, instead of inserting results into a database, these scripts check for abnormal responses and send alerts to the HAWC Slack workspace if one is detected. Alerts of this type exist for the high voltage, low voltage, run status, AC voltage, FEB temperatures, output file sizes, and data copying. Each alert is constructed to send appropriate details (such as which high voltage channels have tripped) if a problem is detected, as well as to temporarily silence themselves after a certain number of alerts if no change in state is detected to prevent overloading the Slack channels. These alert experts to critical issues in real-time and allow a quick response to correct the problem. An example of one such alert displayed in the Slack workspace is shown in Fig. A.4.

## **BIBLIOGRAPHY**



## BIBLIOGRAPHY

- [1] Gianfranco Bertone, Dan Hooper, and Joseph Silk. Particle dark matter: Evidence, candidates and constraints. *Phys. Rept.*, 405:279–390, 2005.
- [2] Gerard Jungman, Marc Kamionkowski, and Kim Griest. Supersymmetric dark matter. *Phys. Rept.*, 267:195–373, 1996.
- [3] V. C. Rubin, Jr. Ford, W. K., and N. Thonnard. Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 (R=4kpc) to UGC 2885 (R=122kpc). 238:471–487, June 1980.
- [4] Edvige Corbelli and Paolo Salucci. The Extended Rotation Curve and the Dark Matter Halo of M33. *Mon. Not. Roy. Astron. Soc.*, 311:441–447, 2000.
- [5] Link to figure source: [https://www.researchgate.net/figure/M33-rotation-curve-5\\_fig1\\_333232727](https://www.researchgate.net/figure/M33-rotation-curve-5_fig1_333232727).
- [6] Douglas Clowe, Marusa Bradac, Anthony H. Gonzalez, Maxim Markevitch, Scott W. Randall, Christine Jones, and Dennis Zaritsky. A direct empirical proof of the existence of dark matter. *Astrophys. J. Lett.*, 648:L109–L113, 2006.
- [7] Andrew McKellar. Molecular Lines from the Lowest States of Diatomic Molecules Composed of Atoms Probably Present in Interstellar Space. *Publications of the Dominion Astrophysical Observatory Victoria*, 7:251, January 1941.
- [8] A. A. Penzias and R. W. Wilson. A Measurement of Excess Antenna Temperature at 4080 Mc/s. 142:419–421, July 1965.
- [9] P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, and et al. Planck2015 results. *Astronomy & Astrophysics*, 594:A13, Sep 2016.
- [10] David J Griffiths. *Introduction to elementary particles; 2nd rev. version*. Physics textbook. Wiley, New York, NY, 2008.
- [11] G. Aad, T. Abajyan, B. Abbott, J. Abdallah, S. Abdel Khalek, A.A. Abdelalim, O. Abdinov, R. Aben, B. Abi, M. Abolins, and et al. Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Physics Letters B*, 716(1):1–29, Sep 2012.
- [12] S. Chatrchyan, V. Khachatryan, A.M. Sirunyan, A. Tumasyan, W. Adam, E. Aguilo, T. Bergauer, M. Dragicevic, J. Erö, C. Fabjan, and et al. Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc. *Physics Letters B*, 716(1):30–61, Sep 2012.
- [13] Kevork N. Abazajian and J. Patrick Harding. Constraints on WIMP and Sommerfeld-Enhanced Dark Matter Annihilation from Observations of the Galactic Center. *JCAP*, 1201:041, 2012.

- [14] Aaron A Dutton and Andrea V Macciò. Cold dark matter haloes in the planck era: evolution of structural parameters for einasto and nfw profiles. *Monthly Notices of the Royal Astronomical Society*, 441(4):3359–3374, 2014.
- [15] M. Kamionkowski. WIMP and axion dark matter. In *ICTP Summer School in High-Energy Physics and Cosmology*, pages 394–411, 6 1997.
- [16] Jianglai Liu, Xun Chen, and Xiangdong Ji. Current status of direct dark matter detection experiments. *Nature Phys.*, 13(3):212–216, 2017.
- [17] The ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, 3(08):S08003–S08003, aug 2008.
- [18] The CMS Collaboration. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08004–S08004, aug 2008.
- [19] Oliver Buchmueller, Caterina Doglioni, and Lian Tao Wang. Search for dark matter at colliders. *Nature Phys.*, 13(3):217–223, 2017.
- [20] J. Patrick Harding. Dark matter indirect detection with gamma rays. *Journal of Physics: Conference Series*, 866:012007, jun 2017.
- [21] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An Introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159–177, 2015.
- [22] Louis E. Strigari, James S. Bullock, Manoj Kaplinghat, Juerg Diemand, Michael Kuhlen, and Piero Madau. Redefining the Missing Satellites Problem. *Astrophys. J.*, 669:676–683, 2007.
- [23] A. Albert et al. Search for Dark Matter Gamma-ray Emission from the Andromeda Galaxy with the High-Altitude Water Cherenkov Observatory. *JCAP*, 06:043, 2018. [Erratum: *JCAP* 04, E01 (2019)].
- [24] A.U. Abeysekara et al. A Search for Dark Matter in the Galactic Halo with HAWC. *JCAP*, 02:049, 2018.
- [25] M. Ackermann et al. The First Fermi-LAT Catalog of Sources Above 10 GeV. *Astrophys. J. Suppl.*, 209:34, 2013.
- [26] Sebastian Hoof, Alex Geringer-Sameth, and Roberto Trotta. A Global Analysis of Dark Matter Signals from 27 Dwarf Spheroidal Galaxies using 11 Years of Fermi-LAT Observations. *JCAP*, 02:012, 2020.
- [27] B. L. Dingus, D. B. Kieda, and M. H. Salamon, editors. *Proceedings, 26th International Cosmic Ray Conference (ICRC), August 17-25, 1999, Salt Lake City: Invited, Rapporteur, and Highlight Papers*, volume 516, New York, 2000. AIP.
- [28] H. Abdallah et al. Search for dark matter annihilations towards the inner Galactic halo from 10 years of observations with H.E.S.S. *Phys. Rev. Lett.*, 117(11):111301, 2016.
- [29] R. Atkins et al. Observation of TeV Gamma Rays from the Crab Nebula with Milagro Using a New Background Rejection Technique. 595, October 2003.

- [30] Maurizio Spurio. *Particles and astrophysics : a multi-messenger approach / Maurizio Spurio*. Astronomy and astrophysics library. Springer, Cham, Switzerland, 2015.
- [31] J. Matthews. A Heitler Model of Extensive Air Showers. *Astroparticle Physics*, 22(5):387 – 397, 2005.
- [32] T Antoni, W.D Apel, F Badea, K Bekk, K Bernlöhner, H Blümer, E Bollmann, H Bozdog, I.M Brancus, A Chilingarian, K Daumiller, P Doll, J Engler, F Feßler, H.J Gils, R Glasstetter, R Haeusler, W Hafemann, A Haungs, D Heck, T Holst, J.R Hörandel, K.-H Kampert, J Kempa, H.O Klages, J Knapp, D Martello, H.J Mathes, H.J Mayer, J Milke, D Mühlberg, J Oehlschläger, M Petcu, H Rebel, M Risse, M Roth, G Schatz, F.K Schmidt, T Thouw, H Ulrich, A Vardanyan, B Vulpescu, J.H Weber, J Wentz, T Wiegert, J Wochele, and J Zabierowski. Electron, muon, and hadron lateral distributions measured in air showers by the cascade experiment. *Astroparticle Physics*, 14(4):245 – 260, 2001.
- [33] A. U. Abeysekara et al. Observation of the Crab Nebula with the HAWC Gamma-Ray Observatory. *Astrophys. J.*, 843(1):39, 2017.
- [34] A. U. Abeysekara et al. The 2HWC HAWC Observatory Gamma Ray Catalog. *Astrophys. J.*, 843(1):40, 2017.
- [35] A. U. Abeysekara et al. Data acquisition architecture and online processing system for the HAWC gamma-ray observatory. *Nuclear Instruments and Methods in Physics Research A*, 888:138–146, April 2018.
- [36] Samuel Marinelli. *Constraints on Lorentz Invariance Violation with the HAWC Observatory*. PhD thesis, Michigan State University, 220 Trowbridge Rd, East Lansing, MI 48824, 3 2019.
- [37] Vincent Marandon, Armelle Jardin-Blicq, and Harm Schoorlemmer. Latest news from the HAWC outrigger array. *PoS, ICRC2019:736*, 2020.
- [38] Adiv González Muñoz. Latest news from the high altitude water cherenkov observatory. *Journal of Physics: Conference Series*, 730:012013, 07 2016.
- [39] A.U. Abeysekara et al. Sensitivity of the high altitude water cherenkov detector to sources of multi-teV gamma rays. *Astroparticle Physics*, 50-52:26 – 32, 2013.
- [40] K. M. Gorski, Eric Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelman. HEALPix - A Framework for high resolution discretization, and fast analysis of data distributed on the sphere. *Astrophys. J.*, 622:759–771, 2005.
- [41] A.U. Abeysekara et al. Measurement of the Crab Nebula at the Highest Energies with HAWC. *Astrophys. J.*, 881:134, 2019.
- [42] Kelly Malone. *A Survey of the Highest Energy Astrophysical Sources with the HAWC Observatory*. PhD thesis, Pennsylvania State University, State College, PA 16801, 10 2018.
- [43] Daniel Fiorino. *Observation of TeV-Energy Cosmic-Ray Anisotropy with the HAWC Observatory*. PhD thesis, University of Wisconsin, Madison, Madison, WI 53706, 2 2015.

- [44] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, H. Voss, M. Backes, T. Carli, O. Cohen, A. Christov, D. Dannheim, K. Danielowski, S. Henrot-Versille, M. Jachowski, K. Kraszewski, Jr. Krasznahorkay, A., M. Kruk, Y. Mahalalel, R. Ospanov, X. Prudent, A. Robert, D. Schouten, F. Tegenfeldt, A. Voigt, K. Voss, M. Wolter, and A. Zemla. TMVA - Toolkit for Multivariate Data Analysis. *arXiv e-prints*, page physics/0703039, Mar 2007.
- [45] S. S. Wilks. The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Ann. Math. Statist.*, 9(1):60–62, 03 1938.
- [46] A. Albert et al. Constraints on Lorentz Invariance Violation from HAWC Observations of Gamma Rays above 100 TeV. *Phys. Rev. Lett.*, 124:131101, Mar 2020.
- [47] H. Martínez-Huerta and A. Pérez-Lorenzana. Restrictions from lorentz invariance violation on cosmic ray propagation. *Physical Review D*, 95(6), Mar 2017.
- [48] Petr Satunin. New constraints on Lorentz Invariance violation from Crab Nebula spectrum beyond 100 TeV. *Eur. Phys. J.*, C79(12):1011, 2019.
- [49] A.U. Abeysekara et al. Multiple Galactic Sources with Emission Above 56 TeV Detected by HAWC. *Phys. Rev. Lett.*, 124(2):021102, 2020.
- [50] Herman Chernoff. On the distribution of the likelihood ratio. *Ann. Math. Statist.*, 25(3):573–578, 09 1954.
- [51] A. U. Abeysekara et al. Searching for Dark Matter Sub-structure with HAWC. *JCAP*, 1907(07):022, 2019.
- [52] A. Albert et al. Dark Matter Limits From Dwarf Spheroidal Galaxies with The HAWC Gamma-Ray Observatory. *Astrophys. J.*, 853(2):154, 2018.
- [53] Anatoly A. Klypin, Andrey V. Kravtsov, Octavio Valenzuela, and Francisco Prada. Where are the missing Galactic satellites? *Astrophys. J.*, 522:82–92, 1999.
- [54] S. Garrison-Kimmel, A. Wetzel, J. S. Bullock, P. F. Hopkins, M. Boylan-Kolchin, C.-A. Faucher-Giguère, D. Kereš, E. Quataert, R. E. Sanderson, A. S. Graus, and T. Kelley. Not so lumpy after all: modelling the depletion of dark matter subhaloes by Milky Way-like galaxies. *MNRAS*, 471:1709–1727, October 2017.
- [55] Stacy Y. Kim, Annika H. G. Peter, and Jonathan R. Hargis. Missing satellites problem: Completeness corrections to the number of satellite galaxies in the milky way are consistent with cold dark matter predictions. *Phys. Rev. Lett.*, 121:211302, Nov 2018.
- [56] Giacomo Vianello, Robert J. Lauer, Patrick Younk, Luigi Tibaldo, James M. Burgess, Hugo Ayala, Patrick Harding, Michelle Hui, Nicola Omodei, and Hao Zhou. The Multi-Mission Maximum Likelihood framework (3ML). 2015.
- [57] Giacomo Vianello. 3ml. <https://github.com/threeML/threeML>, 2016.
- [58] Dan Hooper and Samuel J. Witte. Gamma Rays From Dark Matter Subhalos Revisited: Refining the Predictions and Constraints. *JCAP*, 1704(04):018, 2017.

- [59] Ti-Lin Chou, Dimitrios Tanoglidis, and Dan Hooper. Resolving Dark Matter Subhalos With Future Sub-GeV Gamma-Ray Telescopes. *Phys. Dark Univ.*, 21:1–7, 2018.
- [60] Gary Steigman, Basudeb Dasgupta, and John F. Beacom. Precise Relic WIMP Abundance and its Impact on Searches for Dark Matter Annihilation. *Phys. Rev.*, D86:023506, 2012.
- [61] J. Aleksić et al. Optimized dark matter searches in deep observations of Segue 1 with MAGIC. *JCAP*, 1402:008, 2014.
- [62] Vincent Bonnivard, Moritz Hütten, Emmanuel Nezri, Aldée Charbonnier, Céline Combet, and David Maurin. CLUMPY : Jeans analysis, gamma-ray and neutrino fluxes from dark matter (sub-)structures. *Comput. Phys. Commun.*, 200:336–349, 2016.
- [63] Joachim Stadel, Doug Potter, Ben Moore, Jurg Diemand, Piero Madau, Marcel Zemp, Michael Kuhlen, and Vicent Quilis. Quantifying the heart of darkness with GALO - a multi-billion particle simulation of our galactic halo. *Mon. Not. Roy. Astron. Soc.*, 398:L21–L25, 2009.
- [64] Julio F. Navarro, Aaron Ludlow, Volker Springel, Jie Wang, Mark Vogelsberger, Simon D. M. White, Adrian Jenkins, Carlos S. Frenk, and Amina Helmi. The Diversity and Similarity of Cold Dark Matter Halos. *Mon. Not. Roy. Astron. Soc.*, 402:21, 2010.
- [65] A. Burkert. The Structure of dark matter halos in dwarf galaxies. *IAU Symp.*, 171:175, 1996. [Astrophys. J.447,L25(1995)].
- [66] J. Lavalle, Q. Yuan, D. Maurin, and X. J. Bi. Full Calculation of Clumpiness Boost factors for Antimatter Cosmic Rays in the light of Lambda-CDM N-body simulation results. Abandoning hope in clumpiness enhancement? *Astron. Astrophys.*, 479:427–452, 2008.
- [67] Miguel A. Sánchez-Conde and Francisco Prada. The flattening of the concentration–mass relation towards low halo masses and its implications for the annihilation signal boost. *Mon. Not. Roy. Astron. Soc.*, 442(3):2271–2277, 2014.
- [68] Julio F. Navarro, Carlos S. Frenk, and Simon D. M. White. A Universal density profile from hierarchical clustering. *Astrophys. J.*, 490:493–508, 1997.
- [69] Ángeles Moliné, Miguel A. Sánchez-Conde, Sergio Palomares-Ruiz, and Francisco Prada. Characterization of subhalo structural properties and implications for dark matter annihilation signals. *Mon. Not. Roy. Astron. Soc.*, 466(4):4974–4990, 2017.
- [70] Volker Springel, Jie Wang, Mark Vogelsberger, Aaron Ludlow, Adrian Jenkins, Amina Helmi, Julio F. Navarro, Carlos S. Frenk, and Simon D. M. White. The Aquarius Project: the subhalos of galactic halos. *Mon. Not. Roy. Astron. Soc.*, 391:1685–1711, 2008.
- [71] Lidia Pieri, Julien Lavalle, Gianfranco Bertone, and Enzo Branchini. Implications of High-Resolution Simulations on Indirect Dark Matter Searches. *Phys. Rev.*, D83:023518, 2011.
- [72] A. Abramowski et al. Constraints on an Annihilation Signal from a Core of Constant Dark Matter Density around the Milky Way Center with H.E.S.S. *Phys. Rev. Lett.*, 114(8):081301, 2015.
- [73] Carlo Nipoti and James Binney. Early flattening of dark matter cusps in dwarf spheroidal galaxies. *Monthly Notices of the Royal Astronomical Society*, 446(2):1820–1828, Nov 2014.

- [74] A. U. Abeysekara et al. Observation of Anisotropy of TeV Cosmic Rays with Two Years of HAWC. *Astrophys. J.*, 865(1):57, 2018.
- [75] Pooja Surajbali. *Observing large-scale structures in the gamma-ray sky*. PhD thesis, Ruprecht Karl University of Heidelberg, Heidelberg, Germany, 7 2020.
- [76] T-P Li and Y-Q Ma. Analysis methods for results in gamma-ray astronomy. *The Astrophysical Journal*, 272:317–324, 1983.
- [77] TeVCat homepage: <http://tevcat.uchicago.edu/>.
- [78] Giacomo Vianello. Hawc accelerated likelihood. [https://github.com/threeML/hawc\\_hal/](https://github.com/threeML/hawc_hal/), 2018.
- [79] Marco Cirelli, Gennaro Corcella, Andi Hektor, Gert Hutsi, Mario Kadastik, Paolo Panci, Martti Raidal, Filippo Sala, and Alessandro Strumia. PPPC 4 DM ID: A Poor Particle Physicist Cookbook for Dark Matter Indirect Detection. *JCAP*, 03:051, 2011. [Erratum: *JCAP* 10, E01 (2012)].
- [80] MySQL homepage: <https://www.mysql.com/>.
- [81] ZeroMQ homepage: <http://zeromq.org/>.
- [82] PHP homepage: <https://www.php.net/>.
- [83] Google Charts homepage: <https://developers.google.com/chart>.