SEEKING THE LIGHT IN THE DARK:
QUESTS FOR IDENTIFYING
DARK MATTER

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By
Chun Yu Ng 吳震宇, BSc
Graduate Program in Physics

The Ohio State University
2016

Dissertation Committee:
Professor John F. Beacom, Advisor
Professor Amy L. Connolly
Professor Robert J. Perry
Professor Todd A. Thompson
The night sky is a beautiful display of stars and galaxies. We have come a long way to realize that they are made with substances that can be produced and studied on Earth. However, it has been discovered that those substances make up only 5% of the observable Universe, with the remaining 95% being mysterious substances called dark matter and dark energy, both of which have never been observed directly. Their nature is among the most profound questions in modern science, and unquestionably holds the key to the fundamentals of the Universe and laws of physics. In this dissertation, I discuss a series of papers related to studies of dark matter. I revisit the problem of dark matter annihilation in the extragalactic background radiation, and show that they are sensitive to the properties of the smallest dark matter halos. I show that the newly discovered high-energy astrophysical neutrinos can be used to test secret neutrino interactions through their propagation in the Cosmic Neutrino Background. I discuss how we use the Fermi-GBM to search for sterile neutrino dark matter in a region of parameter space that is not probed otherwise. I discuss a novel method for testing dark matter annihilation/decay signals with a line spectrum. Lastly, I discuss new and interesting results from gamma-ray observations of the Sun, and how this is related to future dark matter searches from the Sun.
To those who loved me
Acknowledgments

May it be a light for you in dark places, when all other lights go out.

Galadriel

Thanks to all the people whom I had the pleasure to meet during graduate school, OSU will always be a special place to me.

Among fellow graduate students, I probably spent too much time debating with Ranjan Laha, but at least some of that was about physics. Shirley Li was always incredibly helpful and constructive, especially when I was most desperate. I also had the privilege of working closely with three younger students, Eric Speckhard, Bei Zhou, and Chris Cappiello, who always reminded me of the old Chinese saying, “fear the young.” I thank Andrea Albert for introducing me to CCAPP; joining CCAPP ended up being one of the best decision I have ever made. And I could not have imagined that the first student I interacted with at OSU would teach me, among other things, about superbowl and Thanksgiving, and then ended up being my academic neighbor. It is a pleasure meeting you, Joe McEwen.

I am also indebted to the kindness of many CCAPP postdocs (many are professors now!), especially when I just started. Thanks Shunsaku Horiuchi, Sheldon Campbell, Carsten Rott, Kohta Murase, Basudeb Dasgupta, and Jennifer Gaskins, all of whom I had the privilege to collaborate with. I am sure we will cross paths again. Thanks Mauricio Bustamante and Tim Linden for always being always helpful in Astro lunch and making the after-lunch lunch thing a reality! I look forward to collaborating with you soon.

I thank the many faculty who taught and guided me in the graduate school, especially
my thesis committee members, who each year encouraged me and reassured me that I am worthy of being a graduate student. Thanks Annika Peter and Chris Hirata for being such valuable additions to CCAPP, especially Annika for always being kind, optimistic, and such a delight to collaborate with. I am very thankful of Savvas Koushiappas and Julie McEnery for letters of recommendation. And a special thanks to Lisa, for holding CCAPP together!

Of course, what I am today is impossible without my advisor, John Beacom, a true mentor, not just in research but also in life. His enthusiasm in physics is always infectious, and his jokes are sometimes funny. I will always remember our late night conversations, which in many occasions materialized into interesting ideas! Thanks for showing me there is always a way forward.

And last but not least, how could I not thank all my dearest friends and families, who always reminded me the joy of life when there were seemingly none. I am especially thankful to the many women who raised me and loved me. This is for you.
VITA

2010 ................................................. BSc. Physics, The Chinese University of Hong Kong

2010 – 2015 ........................................ Graduate Research Associate, The Ohio State University

2015 – 2016 ........................................ Presidential Fellow, The Ohio State University

Publications

Galactic Center radio constraints on gamma-ray lines from dark matter annihilation,
Ranjan Laha, Kenny C. Y. Ng, Basudeb Dasgupta, and Shunsaku Horiuchi,

Resolving Small-Scale Dark Matter Structures Using Multi-Source Indirect Detection,
Kenny C. Y. Ng, Ranjan Laha, Sheldon Campbell, Shunsaku Horiuchi, Basudeb Das-
gupta, Kohta Murase and John F. Beacom,

Cosmic Neutrino Cascades from Secret Neutrino Interactions,
Kenny C. Y. Ng and John Beacom,

Improved Limits on Sterile Neutrino Dark Matter using Full-Sky Fermi Gamma-Ray Burst Monitor Data,
Kenny C. Y. Ng, Shunsaku Horiuchi, Jennifer M. Gaskins, Miles Smith, Robert Preece,

Dark Matter Velocity Spectroscopy,
Eric G. Speckhard, Kenny C. Y. Ng, John F. Beacom, Ranjan Laha,
First Observation of Time Variation in the Solar-Disk Gamma-Ray Flux with Fermi, 

Kenny C. Y. Ng, John F. Beacom, Annika H. G. Peter, Carsten Rott, 

Fields of Study

Major Field: Physics

Studies in Theoretical Astroparticle Physics: Dark matter, neutrino, cosmic rays, and gamma rays
# Table of Contents

Abstract ................................................................. iii
Dedication ........................................................................ iv
Acknowledgments ............................................................ vi
Vita ................................................................................. vi
List of Figures .................................................................... xi
List of Tables ..................................................................... xix

## Chapters

1 Introduction
   1.1 The dark matter puzzle ............................................. 1
   1.2 Dark matter candidates .............................................. 5
      1.2.1 WIMP dark matter .............................................. 5
      1.2.2 Sterile neutrino dark matter ................................. 9
   1.3 Dark matter indirect detection ................................. 10
      1.3.1 Solar dark matter search .................................... 13
   1.4 My work ................................................................. 13
      1.4.1 Dark matter annihilation in the observable Universe 13
      1.4.2 Probing CνB and νSI through high-energy neutrino propagation 14
      1.4.3 Searching for Sterile Neutrino Dark Matter with Fermi-GBM 15
      1.4.4 Dark matter velocity spectroscopy ...................... 16
      1.4.5 The solar gamma ray puzzle ............................... 17

2 Resolving Small-Scale Dark Matter Structures Using Multi-Source Indirect Detection 19
   2.1 Introduction ............................................................ 19
   2.2 IGRB from DM annihilation ....................................... 22
      2.2.1 Isotropic Galactic component ............................. 22
      2.2.2 Extragalactic component .................................... 32
      2.2.3 EBL attenuation and redshift distribution ............. 43
      2.2.4 Isotropic Galactic vs. Extragalactic .................... 44
   2.3 DM constraints ......................................................... 45
      2.3.1 Gamma-ray line – 130 GeV DM .......................... 47
      2.3.2 Continuum – 10-30 GeV light DM ...................... 56
   2.4 Summary and Outlook ............................................... 57
### 6 First Observation of Time Variation in the Solar-Disk Gamma-Ray Flux with Fermi

6.1 Introduction .............................................. 143
6.2 The Sun observed using Fermi-LAT ...................... 146
   6.2.1 Outline of the Analysis .......................... 146
   6.2.2 Data Selection and Cuts .......................... 147
   6.2.3 Background Estimation .......................... 151
   6.2.4 Solar-Disk Flux Spectrum ......................... 154
   6.2.5 Time Variation of the Solar-Disk Flux ............ 161
   6.2.6 Anticorrelation of the Solar-Disk Flux with Solar Activity .......... 162
6.3 Theoretical Overview and Observational Outlook ........... 164
   6.3.1 Physics of Solar-Disk Gamma Rays—The Cosmic-Ray Framework .... 164
   6.3.2 Prospects for Solar-Disk Gamma-Ray Theory ........... 167
   6.3.3 Prospects for the Inverse Compton Component .......... 168
   6.3.4 Prospects for Fermi and Future Space Missions .......... 169
   6.3.5 Prospects for Ground-Based Telescopes ............... 170
6.4 Conclusions .............................................. 172

### 7 Conclusions

174
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 [LEFT:] Velocity dispersion versus Galactocentric radius for stars in the Milky Way. The dashed and dot-dashed lines are the gravitational potential of the Galactic disk and bulge, while the dotted line is that of dark matter. Dark matter is required to explain the data points, and it dominates the potential at large radius. Plot taken from Ref. [7]. [RIGHT:] An image of the Bullet cluster, overlaid with the X-ray map (red) and the gravitational lensing map (blue). The lensing map traces the total mass, and shows that most of the matter in the cluster is dark. The fact that X-ray mass (mostly hot gas) lags behind the dark matter constrains the self-interaction of dark matter. Picture taken from APOD [8].</td>
<td>2</td>
</tr>
<tr>
<td>1.2 [TOP:] Primordial element predictions from Big Bang Nucleosynthesis (BBN) as a function of the baryon density ( \Omega_b ). Measurements of the elements are represented by the boxes. The deuterium observation puts the strongest constraint on baryon density from BBN. Plot taken from PDG [9]. [BOTTOM:] The angular power spectrum of the temperature fluctuations in the Cosmic Microwave Background (CMB) measured by the WMAP mission compared to the best-fit cosmological model that contains dark matter. The excellent agreement between the theory and data proves the success of the model. Plot taken from ref. [10]. An important conclusion from BBN and CMB about dark matter is that it cannot interact strongly with photons.</td>
<td>3</td>
</tr>
<tr>
<td>1.3 Schematic diagram that summarizes the WIMP paradigm, where dark matter (( \chi )) and some Standard Model particles (( f )) engage in 2-to-2 interactions. The horizontal direction sets the thermal equilibrium condition in the early Universe, and determines the final dark matter abundance. The forward, backward, and top-down directions represent indirect detection, collider searches, and direct detection.</td>
<td>6</td>
</tr>
</tbody>
</table>
1.4 The freeze-out mechanism, arguably the most popular way to produce the dark matter cosmological abundance. The y-axis parametrize the dark matter energy density, while the x-axis is the mass, $m$, over the temperature, $T$; time flows toward the right. At high temperature, dark matter follows the equilibrium density, $n_{eq}$. Once the temperature dropped below the dark matter mass, the production of dark matter would be suppressed, and its density would decrease exponentially due to rapid annihilation into Standard Model particles. This would continue until $T \sim m/20$, when the dark matter annihilation rate fall below the expansion rate of the Universe. At this point dark matter annihilation became inefficient and its density would “freeze out”. The final abundance mainly depends on the annihilation cross section. As shown, the observed dark matter abundance corresponds to the weak-scale cross section. Plot taken from Ref. [11].

1.5 Bounds on the mass ($M_1$) and mixing ($\sin^2 \theta$) of sterile neutrinos, assumed to be dark matter, to active neutrinos. The parameter space is bound from all sides. See text for a more detailed description. Plot taken from Ref. [12].

2.1 The normalized line of sight integral of the DM density squared (the J-factor) as a function of the viewing angle, $\psi$. The J-factor for EIN, NFW and ISO profiles for the smooth halo are shown with dotted lines. The contributions of of substructures to the J-factors for the LOW substructure case, assuming $M_{\text{min}} = 10^{-6} M_\odot$, are shown with dotted-dashed lines. The total J-factor (smooth + LOW substructure) for just the one case (NFW), is shown with a bold solid line.

2.2 Same as Fig. 2.1, but for the HIGH substructure case.

2.3 The formation redshift, $z_f$, versus $M_{\text{min}}$, for the first generation halos. The formation redshift is obtained by requiring the linear mass variance be equal to the characteristic overdensity, $\sigma_L (M, z) = 1.686$ [13, 14]. For $\sigma_L (M, z)$, we use the fitting functions of Eisenstein and Hu [15], which are evaluated and normalized with the Planck cosmological parameters [16]. For illustration, we also show $z_f$ for the $n_s = 1$ case as well as the extrapolated results from simulations by Ludlow et al. [17] (we take $z_f$ to be $z_{-2}$).

2.4 The characteristic density of the first generation halos, $\rho_{\text{max}}$, versus $M_{\text{min}}$, for the corresponding cases of $z_f$ in Fig. 2.3 using Eq. 2.8. The substructure boost is approximately proportional to $\rho_{\text{max}}$ (Eq. 2.6).

2.5 The normalized halo mass function $(1/\rho_{\chi}) M^2 dn/dM$ plotted versus $M$ for redshift $z = 0, 2, 4$. The halo mass function as a function of the linear mass variance is given by P12 [18]. The redshift evolutions of the fitting parameters are given by Tinker et al. [19].

2.6 The concentration parameter, $c(M, z)$, plotted against $M$ for redshift $z = 0, 2, 4$. The concentration mass relation as a function of the linear mass variance is again given by P12 [18]. For comparison, we also show the concentration if we simply extend the concentration-mass relation to small scales using the analytic function given in P12.
2.7 The clumping factor at $z = 0$ versus $M_{\text{min}}$. Using the P12 [18] Halo Model, we obtain the clumping factor without substructure enhancement. Adding the K10 [20] substructure model, we show the substructure-enhanced clumping factor for LOW and HIGH cases. For comparison, we also show the clumping factor if we simply extrapolate the concentration relation in P12, the clumping factor with G12 [21] substructure model, and the clumping factor using the extrapolated halofit [22] non-linear power spectrum, following S12 [23]. In this work, we consider the LOW and HIGH scenarios as the conservative and optimistic substructure cases.

2.8 The attenuation fraction of monochromatic gamma-ray signals from cosmological DM annihilation sources (Eq. 4.12) versus the emitted photon energy for different EBL models. We consider the EBL model from Gilmore et al. [24], “Best Fit 06” from Kneiske et al. [25], and the “Lower-Limit” from Kneiske and Dole [26]. All models considered are within 2-$\sigma$ of the Fermi measurement [27]. We adopt the Gilmore model throughout.

2.9 The extragalactic DM annihilation redshift distribution (Eq. 2.21). We show the distributions for the P12 Halo Model, P12 Halo Model with K10 LOW, and HIGH substructure models. The upper set of three lines uses $M_{\text{min}} = 10^{-6} M_\odot$, while the the lower set uses $M_{\text{min}} = 10^6 M_\odot$. The shape of the distribution varies mildly in different scenarios. Most of the DM annihilation signal comes from small redshifts.

2.10 The IGRB spectrum measured by Fermi. We show the published Fermi IGRB data and the extended single power-law fit from Abdo et al. [28], and also data points from Fermi preliminary results [29]. The EBL-attenuated power law is adapted from [30].

2.11 The combined (isotropic Galactic + extragalactic) DM signal for LOW and HIGH substructures for the 130 GeV DM with annihilation channel $\chi\chi \rightarrow \gamma\gamma$. Superposed are the IGRB spectra from Fig. 2.10. The individual isotropic Galactic and extragalactic components for the LOW substructure case are shown in black dotted lines. All DM components are evaluated with $\sigma v = 2 \times 10^{-27} \text{cm}^3 \text{s}^{-1}$ and $M_{\text{min}} = 10^{-6} M_\odot$. The DM signals are convolved with 10% Gaussian smearing to take into account the energy resolution of Fermi-LAT. For visualization, we also show the total spectra (isotropic Galactic + extragalactic + Abdo 2010 fit) for LOW and HIGH cases.

2.12 The “clumping factor-$\sigma v$” parameter space plane. It is bounded from below by the Minimum DM clustering and from the right by the relic abundance requirement. The allowed region for 130 GeV DM is fixed by the GC observation (vertical green band). The blue solid line is obtained using the total IGRB flux (100% IGRB). The bending feature notes the transition into where the isotropic Galactic component dominates (Eq. 2.23). We show the translation from clumping factor to $M_{\text{min}}$ using the LOW (HIGH) substructure case, with solid green (dashed red) marks on the allowed parameter space. The blue dot-dashed line represents the parameter space that IGRB could probe for a more detailed analysis (10% IGRB).

2.13 Same as Fig. 2.11, but with 30 GeV $\chi\chi \rightarrow b\bar{b}$ and $\sigma v = 2.2 \times 10^{-26} \text{cm}^3 \text{s}^{-1}$. We consider the prompt photon spectrum only.
2.14 Same as Fig. 2.12, but with 30 GeV $\chi\chi \rightarrow \bar{b}b$. The green parameter space are taken from the “best-fit spatial model” from [31].

2.15 Same as Fig. 2.11, but with 10 GeV $\chi\chi \rightarrow \tau^+\tau^-$ and $\sigma v = 2.2 \times 10^{-26} \text{ cm}^3 \text{s}^{-1}$. We consider the prompt spectrum only.

2.16 Same as Fig. 2.12, but with 10 GeV $\chi\chi \rightarrow \tau^+\tau^-$. The green parameter space are taken from the “best-fit spatial model” from [31].

3.1 Present constraints and future sensitivity to $\nu_{\text{SI}}$ in terms of neutrino coupling, $g$, and mediator mass, $M$, with diagonal dotted contours shown for values of the dimensionful coupling $G$. The blue shaded regions are excluded by astrophysical and cosmological considerations based on SN 1987A [32], BBN [33], and the CMB [34, 35]. The pink dashed lines indicate flavor-dependent limits based on laboratory measurements of meson and lepton decays [36]; we consider only the weakest limit, for $\nu_\tau$, to be robustly excluded for all flavors, and it is shaded. The red shaded region is excluded based on measurement of Z-boson decay [37]. The gray shaded region indicates the non-perturbative regime. The orange lines are contours of unit optical depth for different initial neutrino energies (Eq. 3.10), indicating the approximate boundary of the parameter space above which IceCube is sensitive to $\nu_{\text{SI}}$. The squares represent the example parameters (given in Table 4.1) used in our calculations.

3.2 The effects of $\nu_{\text{SI}}$ on an emitted line spectrum at 1 PeV with SFR evolution, for different assumed resonance energies, as labeled. The solid lines are for free-streaming neutrinos, the dotted lines are for $\nu_{\text{SI}}$ with regeneration of two neutrinos (the beam and target), and the dashed lines are for $\nu_{\text{SI}}$ with no regeneration.

3.3 The effects of $\nu_{\text{SI}}$ on an emitted continuum spectrum that is consistent with IceCube data. The solid line indicates the free-streaming case ($\gamma = 2$ and $E_{\text{cut}} = 10^7 \text{ GeV}$), while the other lines are for four models of $\nu_{\text{SI}}$, with the parameters defined in Table 4.1.

3.4 Same as Fig. 3.3, but with different values of $E_{\text{cut}}$ for the emitted spectrum. The solid lines are the free streaming case ($\gamma = 2$ and values of $E_{\text{cut}}$ as labeled), and the dashed lines are for Model A $\nu_{\text{SI}}$.

3.5 Same as Fig. 3.3, but with different values of the spectral index. The solid lines are for the free-streaming cases (values of $\gamma$ as labeled and $E_{\text{cut}} = 10^7 \text{ GeV}$), and the dashed lines are for Model B $\nu_{\text{SI}}$.

3.6 The predicted spectra of cascade events (histogram), taking into account realistic detector effects, compared to the 988 days of existing IceCube data (points). Selected curves from Fig. 3.3 are shown, along with an $E^{-2}$ power law with no intrinsic cutoff.

3.7 The same as Fig. 3.6, except that the calculations are for 988 days of an ideal efficiency 1-Gton detector, the energy bins are half as wide, and the range extends to lower energies.

4.1 The J-factor, $J(\psi)$ (Eq. (4.4)), as a function of half opening angle $\psi$ relative to the GC, for four Milky Way dark matter halo profiles.
4.2 The effective area for det-7 NaI detector versus the detector zenith angle relative to the detector normal, for three example energy bins. Points are data from GBM calibration files, and the anomalous dips in the effective area come from blockages from other satellite components. Solid lines are fits to the data neglecting the dips.

4.3 The effective area versus the incident photon energy for three incident angles. The points are obtained from GBM calibration files and the lines are linear interpolations of the points.

4.4 The convolved J-factor (Eq. (4.12)) versus the opening angle with respect to the GC, \( \psi(0, j) \), defined by where the detector normal is pointing. The difference between different profiles is drastically reduced. A small energy dependence is introduced from the effective area. The vertical dotted line denotes the boundary of our ROI. Shown in grey are the theoretical J-factors from Fig. 4.1.

4.5 (Left) Simulated dark matter counts rate maps in Galactic coordinates for several line energies, taking into account detector response. For each map, the assumed line energy is contained in the energy bin shown and \( \sin^2 \theta \) (as labeled) are chosen to approximately match the observed counts in the same bin. (Right) The final counts rate sample from 4 years of data from the NaI detector, which corresponds to 4.6 million seconds (\( \sim 53 \) days) of live time after data cuts. All of the sky maps are pixelated into 768 HEALPix pixels. The pixel position corresponds to the pointing direction of the detector normal. The grey pixels are where no observing time is registered after the selection cuts.

4.6 A sample of the raw photon count rates for the GBM-NaI detector from 20-12-2008, from 344 keV to 471 keV, where the counts rate is dominated by the cosmic ray induced background. The blue line indicates the location of the satellite in geomagnetic latitude (in absolute value). The epochs with no data are when the satellite is physically in the SAA. The red hashed regions represent the cuts on orbits that pass through the SAA. The grey shaded regions illustrate the cuts on geomagnetic latitude; the data removed by this cut is colored blue. In the end, only black data points in the white regions are used for analysis.

4.7 The counts spectrum for the final data sample for both GC ROI (\( \psi < 60^\circ \)) and anti-GC ROI (\( \psi > 120^\circ \)) chosen to have the same solid angle. The dominant component is a power-law plus various background lines. The overall difference in normalization is due to the higher exposure towards the GC than the anti-GC direction. The additional excess at low energies towards the GC region suggests the rise of the astrophysical component. The vertical dotted line indicates the energy bins used for the spectral analysis.

4.8 The blue points are the measured data with error bars indicating the 5% systematic error. The blue line is the best fit model to the data with only the power-law component. The red line shows the best fit model when including the line signal with 95% upper limit amplitude. The red arrow indicates the central energy of the line signal.
4.9 The black solid line is the 95% C.L. upper limit from the spectral analysis. The green (yellow) shaded region shows the 68% (95%) intervals from the Monte Carlo simulations.

4.10 The conservative upper limit from the flux analysis (Black) and the 95% C.L. upper limit from the spectral analysis (Blue) for the decay rate of sterile neutrino dark matter. The hashed regions are excluded by the corresponding analyses.

4.11 Constraints from X-ray missions on sterile neutrino dark matter decays, which depends on the mixing angle, $\sin^2(2\theta)$, and the mass, $m_s$. In the shown mass range, the best previous constraints are set by observations of cosmic X-ray background (CXB) from HEAO-1 [38] and the Milky Way (MW) halo from INTEGRAL [39]. The lower bound of the parameter space (black-hashed region) is valid for the model $\nu$MSM [12]. The flux (spectral) analysis limit derived from this work is shown in black-dashed (blue) line.

5.1 **Top:** How DM, astrophysical, and detector lines shift with Galactic longitude is starkly different. **Bottom:** For DM signals at positive longitude, our motion through the non-rotating DM halo yields a negative LOS velocity and thus a blue shift. In contrast, for astrophysical lines (e.g., from gas), co-rotation in the disk leads to a positive LOS velocity and thus a red shift. These signs reverse at negative longitude. Detector lines have zero shift.

5.2 Comparison of received spectra for DM and gas (G2). The emitted spectra are taken to have equal flux and to be centered at 3.5 keV before velocity effects. The line profiles include velocity dispersion and shift effects, as well as the energy resolution of Astro-H. Vertical bands indicate the 1-σ centroid uncertainties after 2-Ms observations. For contrast, the brown line in the figure and inset shows the same signal if Astro-H had the energy resolution of XMM.

5.3 LOS velocity for DM and various gas models (the realistic version of Fig. 5.1). Uncertainties are computed assuming 2-Ms Astro-H exposures on each point.

5.4 Radial velocity dispersion profiles for the MW and M31. Shaded vertical bands indicate the range of radii that contribute 90% of the signal along $\psi(l = 20^\circ, |b| = 5^\circ)$ in the MW and $\psi = 1^\circ$ in M31; the radius ranges for the other directions discussed in the text are similar. Note that the lower bounds of these ranges are the smallest $r$ probed by these directions.

5.5 Intrinsic ($\sigma_{DM}$) and effective ($\sigma_{eff}$) LOS velocity dispersion profiles for the MW and M31 as a function of $\psi/\psi_S$, the scaled angle relative to the center of each system. For the MW, $\psi_S = 50^\circ$, while for M31, $\psi_S = 2.5^\circ$; these scalings were chosen for display purposes. Intrinsic widths are determined by integrating the spectrum along the LOS using the radial velocity dispersion profiles given in the previous section. Effective widths include detector energy resolution. The increase in the LOS dispersion at small angles in M31 is due to the rising radial dispersions shown in Fig. 5.4; for equally small (scaled) angles in the MW, the LOS dispersion decreases because only radii < 1 kpc, where the radial dispersion is decreasing, contribute.
5.6 LOS velocity profiles for DM and HI gas [40] in M31. DM error bars are calculated assuming 2-Ms exposures with Astro-H and only CXB and detector backgrounds.

6.1 **Left:** Stacked photon counts map of the Sun ROI in 10–100 GeV. **Right:** Same, but for a fake-Sun ROI (in this example, trailing the Sun in its path by +180 days), which is used to measure the diffuse background. The exposures of the two ROIs differ by $\lesssim 2\%$. (Maps for $>0.1$ GeV are shown in Fermi2011.) Visually, the solar-disk component (comparable in extent to the size of the Sun, as marked) is obvious; that of the IC component (decreasing with angle) is more subtle. The numbers of photons within $1.5^\circ$ of the center are 175 versus 19; the numbers in $1.5^\circ–9^\circ$ are 844 versus 710.

6.2 Total gamma-ray flux between 1 GeV and 1.8 GeV within $1.5^\circ$ from the Sun versus time. Each bin corresponds to one week of observation, starting from 2008-08-07 (week 010). Periods that coincide with a bright solar flare are labeled with red squares; these are removed from the analyses. The horizontal grey band shows the resulting 6-year combined flux and its uncertainty.

6.3 **Left:** Angular distribution of the integrated intensity from 1–10 GeV in the Sun ROI. Black points show the observed data with statistical uncertainties only. Colored histograms show the fitted results for the signal and two backgrounds (the estimate of the diffuse background incorporates independent data from the fake-Sun ROIs). The inset shows the same data with smaller angular bins, but without the two solar components (note the different vertical scale). **Right:** Same, but for 10–100 GeV (note the lower flux).

6.4 Energy spectrum of the solar-disk flux. Blue squares and statistical uncertainties (systematic uncertainties, not shown, are $\approx 10\%$) are the results of our analysis with 6 years of data. Black dots and combined statistical and systematic uncertainties are the Fermi2011 results with 1.5 years of data. The green band shows the predicted flux range from the SSG1991 nominal model.

6.5 Energy spectrum of the solar-disk flux, separated into three periods, each of two years. The solar disk flux from first two years is consistent with Fermi2011, while the 1–10 GeV data shows a significant reduction in later periods.

6.6 **Left:** For several sources, the ratio of the 1–10 GeV flux in each year to its 6-year average (a time-independent source would fluctuate around unity). The solar-disk component (blue squares) demonstrates a clear decreasing trend and anticorrelation with the smoothed sunspot number, a tracer of solar activity. Other sources (points displaced for clarity) should be and are consistent with being time-independent; see the text for details. **Right:** Same, but for 10–100 GeV. No obvious trend is observed for the disk component, but the uncertainties are large.
6.7 Energy spectrum of gamma rays from the Sun. Blue squares are the total solar gamma-ray flux (solar disk + IC) within 1.5° of the Sun with only statistical error bars. Black dots are the solar-disk-only component from Fermi2011. The grey band shows the solar-disk-only component found in this work. Green circles are the estimated diffuse background within 1.5° of the Sun. The differential point-source sensitivities for HAWC [41] and LHAASO [42, 43] are shown.
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Parameters for the benchmark scenarios.</td>
<td>75</td>
</tr>
<tr>
<td>4.1 Dark matter profile parameters for widely adopted dark matter profiles in the literature. Our canonical profile is the NFW profile.</td>
<td>91</td>
</tr>
<tr>
<td>6.1 For each energy bin, as defined, the total number of photons within 1.5° of the center of the Sun, the rounded best-fit number of photons due to the solar-disk signal, and the significance (√TS) of the solar-disk flux detection.</td>
<td>159</td>
</tr>
</tbody>
</table>
One of the most fascinating aspects of humanity is our interest in the laws of nature. Throughout history, we have invented many ideas to explain the world around us. We have evolved from the concepts of the four classical elements (earth, water, air, and fire) to the construction of the Standard Model of particle physics that has withstood numerous rigorous scientific tests. This, together with Einstein’s general relativity, can comfortably explain all the physical phenomena within the solar system. However, when we gazed deep into outer space, we discovered that the majority of the matter in the Universe cannot be described by the Standard Model. This matter, owing to its lack of interactions with light, is called dark matter.

In this section, I give an (admittedly incomplete) overview of the lines of evidences for dark matter, discuss some currently popular ideas on its identity, and describe how for the past few years I worked towards advancing these areas.

1.1 The dark matter puzzle

The concept of dark matter can be traced back to at least the 18th century. From the motion of the planet Uranus, astronomers deduced that a new planet must exist, which eventually led to the discovery of the planet Neptune. In a similar way, the planet Vulcan was suggested to explain the orbital data of Mercury. However, we eventually realized that the apparently abnormal orbit of Mercury was due to the inaccuracy of Newtonian gravity, fixed by the theory of general relativity, rather than a new planet. (See Ref. [44] for a more complete account of new matter versus new forces.) Often, when astrophysical observations and laws of physics appear to be in conflict, new matter or physics are required. The modern story of dark matter started in a similar fashion.
Most people credit the Swiss astronomer Fritz Zwicky for pioneering the discovery of dark matter [45]. In the 1930s, using the virial theorem to estimate the velocities of galaxies in the Coma galaxy cluster, he found that the observed velocities of the galaxies were more than ten times higher than his estimates. A possible reason is that the cluster contains a significant amount of gravitating matter that is not galaxies.

The interest on missing matter reached a new height in the 1970s, when the study of galaxy rotation curves by pioneers such as Vera Rubin [46] showed that invisible matter could also be present within galaxies. Using multi-wavelength electromagnetic observations, one can construct a mass model of the galaxy with different populations of stars and gases. These mass models predict that the rotational speed should decrease with distance at the outskirt of galaxies (see Fig. 1.1, left). In other words, one expects falling rotation curves. However, observations show that galaxies have relatively flat rotation curves, and galaxies
Figure 1.2: [TOP:] Primordial element predictions from Big Bang Nucleosynthesis (BBN) as a function of the baryon density $\Omega_b$. Measurements of the elements are represented by the boxes. The deuterium observation puts the strongest constraint on baryon density from BBN. Plot taken from PDG [9]. [BOTTOM:] The angular power spectrum of the temperature fluctuations in the Cosmic Microwave Background (CMB) measured by the WMAP mission compared to the best-fit cosmological model that contains dark matter. The excellent agreement between the theory and data proves the success of the model. Plot taken from ref. [10]. An important conclusion from BBN and CMB about dark matter is that it cannot interact strongly with photons.
rotate faster than they should if they had only luminous matter. This not only is strong
evidence of having invisible gravitating matter in the galaxy, but it also shows that this
matter extends much further than the visible matter in the galaxy. Today we believe that
most of the luminous matter in galaxies resides in diffuse halos of dark matter.

The 1990s marked the beginning of precision cosmology. This revealed that the Universe
contains a significant amount of a new form of matter that is unlike anything that we have
seen. The discovery of the Cosmic Microwave Background (CMB) shows that the Universe
was once very dense and hot, so much that at some point the Universe was made of a soup
of protons, neutrons, electrons, photons, and neutrinos. When the Universe expanded and
cooled, protons and neutrons started merging and forming stable light elements, a process
called Big Bang Nucleosynthesis (BBN). From observations of these elements, in particular
deuterium, it was found that matter contained in the Standard Model only constitutes 5%
of the total energy budget of the Universe (see Fig. 1.2, top). This, together with the
anisotropies of CMB [10] (Fig. 1.2, bottom) and nearby supernova observations [47], form
the cornerstone of the ΛCDM (cosmological constant, Λ, plus Cold Dark Matter) model,
known as the the standard cosmological model. In ΛCDM, baryons only contribute about
5% of the total Universe, dark matter contributes about 25%, and the rest is an even more
mysterious substance called dark energy. For a more detailed discussion on dark energy, see
Ref. [48].

Just like the dark planets, it is possible, albeit quite difficult, to modify the theory of
gravity to reconcile all observations [49]. Here, we work under the assumption that dark
matter is made of a new type of particle not contained in the Standard Model. Though
we do not yet know its identity, the lines of evidence presented above actually tells us a
surprising amount about what dark matter cannot be.

From the abundance and distribution of galaxies, we know that dark matter must have
been non-relativistic during cosmological structure formation; hence the term cold dark
matter. From BBN and CMB data, we know that dark matter cannot carry electric
charge (like electrons), and hence does not interact strongly with light. This is why it
is truly dark and cannot be explained by anything in the Standard Model. From rotation
curves and colliding clusters (see Fig. 1.1), we know that dark matter cannot interact with itself as strongly as normal gases, hence it appears to be collisionless. Lastly, the fact that galaxies are still bound together suggests that dark matter is stable over the age of the Universe, in the sense that they have not all decayed into relativistic particles.

Though we strongly believe that dark matter exists, all of the evidence comes from gravitational interactions. To understand the particle physics of dark matter, we need to detect signatures that can tell us about its particle properties, which is also likely to be connected to other deep problems in physics (hierarchy problem, baryogenesis, neutrino mass, etc).

In the next section, I discuss some generic dark matter candidates, which are particularly relevant to the work in this thesis.

1.2 Dark matter candidates

1.2.1 WIMP dark matter

The Standard Model of particle physics, extended to include neutrino masses, provides an incredibly successful description of particle physics up to at least the electroweak scale at around TeV. However, new physics may be lurking around just beyond the weak interaction scale, and physics at this scale arguably can provide a natural explanation to the observed dark matter abundance. In this case, dark matter could be a Weakly Interacting Massive Particle (WIMP). This WIMP paradigm provides a unified description of dark matter production and observational signatures.

WIMP phenomenology

Without specifying how the dark matter, $\chi$, is coupled to the Standard Model particles, $f$, we can parameterize their interaction through a 2-to-2 diagram, as shown in Fig. 1.3. This summarizes some of the interesting dark matter phenomenology:

$\chi\chi \leftrightarrow ff$: In the horizontal direction are the equilibrium reactions between dark matter annihilation to and creation by Standard Model particles in the early Universe. This
Figure 1.3: Schematic diagram that summarizes the WIMP paradigm, where dark matter ($\chi$) and some Standard Model particles ($f$) engage in 2-to-2 interactions. The horizontal direction sets the thermal equilibrium condition in the early Universe, and determines the final dark matter abundance. The forward, backward, and top-down directions represent indirect detection, collider searches, and direct detection.
determines the production of WIMPs through the freeze-out mechanism, described below.

χχ → ff: The forward reaction describes the late-time annihilation of dark matter particles in the Universe when they occasionally meet each other. This is the basis of dark matter indirect detection, described in more detail in Sec. 1.3.

χχ ← ff: The backward reaction describes the production of dark matter particle from Standard Model particle interactions, such as in particle colliders. Of course, the produced final state dark matter particles are not observable, but it is possible to infer their production through initial-radiation and missing-energy signatures.

χf → χf: Last but not least, the top-down reaction describes elastic scattering between dark matter and Standard Model particles. It is possible to detect these interactions in direct-detection underground experiments.

**Freeze-out mechanism**

Assuming dark matter is coupled to the Standard Model in the early Universe, at a temperature much higher than the dark matter mass, dark matter and the Standard Model plasma were in thermal equilibrium, via χχ ⇌ ff. As the Universe expanded and the temperature dropped below the dark matter mass, eventually the primordial plasma would not have had enough energy to produce any dark matter. The equilibrium reaction then became one-sided as dark matter production was forbidden, i.e., χχ → ff. Therefore, the dark matter abundance would decrease exponentially with time, until at some point it became difficult for the dark matter to find its annihilation partner. This happens when the annihilation rate falls below the expansion rate of the Universe. Thus, the dark matter energy density would approach an asymptotic value, or “freeze out”.

The freeze-out mechanism is shown quantitatively in Fig. 1.4. Importantly, the final dark matter abundance mainly depends on the annihilation cross section, ⟨σv⟩, and only mildly on the mass. The abundance of dark matter is usually represented by its energy
Figure 1.4: The freeze-out mechanism, arguably the most popular way to produce the dark matter cosmological abundance. The y-axis parametrize the dark matter energy density, while the x-axis is the mass, $m$, over the temperature, $T$; time flows toward the right. At high temperature, dark matter follows the equilibrium density, $n_{eq}$. Once the temperature dropped below the dark matter mass, the production of dark matter would be suppressed, and its density would decrease exponentially due to rapid annihilation into Standard Model particles. This would continue until $T \sim m/20$, when the dark matter annihilation rate fall below the expansion rate of the Universe. At this point dark matter annihilation became inefficient and its density would “freeze out”. The final abundance mainly depends on the annihilation cross section. As shown, the observed dark matter abundance corresponds to the weak-scale cross section. Plot taken from Ref. [11].
density relative to the total energy density of the Universe, $\Omega_\chi$. It can be approximately described by [11]

$$\left( \frac{\langle \sigma v \rangle}{\text{cm}^3\text{s}^{-1}} \right) \Omega_\chi h^2 \sim 2.4 \times 10^{-27}, \quad (1.1)$$

where $h$ is about 0.7. For $\Omega_\chi h^2 = 0.11$ [10], it corresponds to $\langle \sigma v \rangle = 2.2 \times 10^{-26}$ cm$^3$s$^{-1}$, which is roughly $(100 \text{ GeV})^{-2}$. Hence, it is considered natural for new physics just beyond the weak scale to be responsible for dark matter. This annihilation cross section also sets a benchmark value for dark matter indirect detection, which is also sensitive to the annihilation cross section.

1.2.2 Sterile neutrino dark matter

The discovery of neutrino masses suggests that rich new physics could be contained in the neutrino sector [50]. Right-handed neutrinos, which carry no Standard Model quantum numbers and are hence sterile, could exist and play a role in giving neutrinos mass. In this case, a new mass state containing mostly sterile states could be the dominant dark matter.

Production through mixing

As opposed to WIMPs, sterile neutrinos have even weaker couplings to the Standard Model particles, and were never in thermal equilibrium in the early Universe. Hence, the freeze-out scenario does not apply.

Instead, the simplest sterile neutrino dark matter production scenario is the non-resonant production mechanism, first suggested by Dodelson and Widrow [51]. The Universe would first start with zero dark matter abundance. At some point in the early Universe, while the active neutrinos were in thermal equilibrium and were constantly being produced, there would be a small probability that the active neutrinos would convert to sterile ones. The lost active neutrino in this case would be quickly replenished by the thermal plasma to ensure thermal equilibrium. As a result, a population of sterile neutrinos would slowly build up and could conceivably explain the bulk of the observed dark matter. To produce the observed dark matter abundance, the mixing angle between active and
sterile neutrinos is required to be [52]

\[ \sin^2 2\theta \simeq 10^{-7.1} \left( \frac{m_\chi}{\text{keV}} \right)^{-1.8}. \] (1.2)

In addition, just like active neutrino mixing in dense material is modified compared to the vacuum case, sterile neutrino production in the early Universe could be affected by the presence of a large primordial lepton asymmetry. This scenario, first suggested by Shi and Fuller [53], is also called resonant production mechanism, as the sterile–active neutrino mixing was enhanced through a resonance, similar to the MSW effect for solar neutrinos in the Sun. Importantly, because the effective mixing is enhanced, one can produce sufficient dark matter density with a smaller vacuum mixing angle compared to the Dodelson–Widrow scenario.

Fig. 1.5 summarizes the situation for sterile neutrino dark matter. The available parameter space is bounded from above and from the right by the Dodelson–Widrow curve (NRP, black solid line) and the X-ray constraints (described in detail in Chap. 4 and Chap. 5). It is also bounded from the left for being too “warm” and from below for not being able to explain all of the observed dark matter abundances through the Shi–Fuller mechanism. Therefore, sterile neutrino dark matter produced by mixing has a finite parameter space and can be tested definitively.

1.3 Dark matter indirect detection

In this section, I focus specifically on dark matter indirect detection [54], a powerful technique for searching for particle signatures of dark matter with astrophysical data. Indirect detection can be used to search for both WIMP and sterile neutrino dark matter, and is the main physics driver for some of the work described in this thesis.

In the current cosmological epoch, most dark matter resides in gravitationally bound objects—dark matter halos. If dark matter is WIMP-like, then it is possible for one dark matter particle to find a partner and annihilate into observable standard model particles just like they did in the early Universe. If dark matter is a sterile neutrino or similar,
Figure 1.5: Bounds on the mass ($M_1$) and mixing ($\sin^2 2\theta_1$) of sterile neutrinos, assumed to be dark matter, to active neutrinos. The parameter space is bound from all sides. See text for a more detailed description. Plot taken from Ref. [12].
it can decay into observable messengers. Late-time dark matter annihilation and decay therefore provide a means to test the particle properties of dark matter, such as its mass and couplings to the Standard Model. This approach is called indirect detection.

The flux of messengers produced by dark matter decay or annihilation is given by

\[
\frac{dF}{dE} = \frac{1}{4\pi} \frac{\Gamma}{m_\chi} \frac{dN}{dE} \int d\Omega \int d\ell \rho_\chi [r(\ell)] \quad \text{(Decay)}
\]

or

\[
\frac{dF}{dE} = \frac{1}{4\pi} \frac{\sigma v}{2m_\chi^2} \frac{dN}{dE} \int d\Omega \int d\ell \rho_\chi^2 [r(\ell)] \quad \text{(Annihilation)}.
\]

Here, \(dF/dE\) is the number of particles per unit area, time, and energy. We can roughly separate these equations to two parts. The first part contains the particle properties of the dark matter, including mass \(m_\chi\), decay rate \(\Gamma\) or annihilation cross section \(\sigma v\), and the decay spectrum \(dN/dE\). The second part describes the astrophysics, including the observational size of the object \(d\Omega\) and the dark matter density profile of the observed object \(\rho_\chi\), which is integrated through a line-of-sight integral.

An important distinction between dark matter decay and annihilation is their different dependence on the dark matter density. For decay, the flux is simply proportional to the dark matter density. For annihilation, the flux is proportional to the square of the dark matter density, since particles must meet each other. As a result, the annihilation rate from dense regions is higher, which enhances its detectability. On the other hand, uncertainties in the dark matter density distribution are also amplified in this case.

These two descriptions are applicable to chargeless annihilation products, such as gamma rays and neutrinos. For charged final states, such as electrons and positrons, the observed spectrum would be modified by particle propagation in the Galactic environment. One can also look for secondary emissions induced by the final states. For example, if the final states were electron and positrons, then, in a dense stellar environment, they could produce gamma rays through inverse-Compton scattering or bremsstrahlung. In some cases, secondary emissions can yield comparable or even better limits than direct messengers.
1.3.1 Solar dark matter search

For WIMP dark matter, the combination of WIMP-nucleon scattering ($\chi f \rightarrow \chi f$) and WIMP self annihilation ($\chi \chi \rightarrow ff$) make the Sun an important dark matter detector [55, 56, 57, 58].

First, similar to normal direct detection experiments, dark matter particles can elastically scatter with the nucleons in the Sun. The key insight is that, if the final-state dark matter particle falls into the potential well of the Sun, then it would form a gravitationally bound state with the Sun. Subsequently, this dark matter particle would continue to scatter with nucleons in the Sun, lose energy, and eventually settle in the core of the Sun. Eventually, a population of dark matter particles will be trapped in the core of the Sun and start to annihilate. Of course, the only annihilation product that can escape the Sun are neutrinos. The Sun, therefore, is also an attractive target for neutrinos telescopes to detect dark matter.

1.4 My work

In this section, I give a high-level introduction to the motivations and key ideas behind the work contained in this dissertation.

1.4.1 Dark matter annihilation in the observable Universe

In Chap. 2, we study the gamma-ray signal from dark matter annihilation in the entire observable Universe.

As described in Eq. 1.3, one can search for WIMP dark matter by looking for gamma rays from WIMP annihilation in known dark matter halos, such as the Milky Way, dwarf galaxies, galaxy clusters, etc. Alternatively, annihilation in all dark matter halos in the Universe, most of which are too small to host stars, would contribute to the extragalactic Isotropic Gamma-Ray Background (IGRB).

The IGRB has long been considered an attractive target for dark matter indirect detection, mainly because the annihilation signal is expected to be greatly boosted by the
small halos and subhalos in the Universe. These small halos are ubiquitous in the cold dark matter Universe, and more importantly, they are denser than the large halos due to their early formation time (when the background density of the Universe is higher). The high density is favorable to dark matter annihilation.

To accurately determine the extragalactic dark matter annihilation flux, we need to know the properties of these small dark matter halos. Unfortunately, they are completely invisible in standard electromagnetic surveys, since they are not massive enough to host stars. Even worse, it is difficult to use numerical N-body simulations to simulate the large-scale Universe while resolving the small halos, which could be as low mass as Earth. Past attempts to estimate the extragalactic dark matter annihilation rate relied on extrapolation, sometimes on multiple quantities. Different groups, with different methodologies, often gave annihilation rates that differ by many order magnitudes [59, 60, 61]. It is therefore difficult to perform robust dark matter studies with the IGRB.

In our work, we take a careful look at this problem, and make two significant contributions. We first point out that the often used power-law extrapolation of the concentration parameter is too naive and optimistic. Instead, by correlating the density of the small dark matter halos and their formation time, we are able to accurately compute the densities of small dark matter halos by calibrating to N-body simulations. Given that the extragalactic annihilation can now be more accurately evaluated, we point out that it is possible to probe the minimum dark matter halo mass by correlating the IGRB to any potential dark matter signals from other indirect detection sources. The minimum dark matter halo mass is sensitive how and when dark matter kinetically decoupled from the Standard Model plasma in the early Universe. Even in the case of non-detection from the IGRB, our approach can be used to constrain these dark matter properties.

1.4.2 Probing $C_{νB}$ and $ν_{SI}$ through high-energy neutrino propagation

In Chap. 3, we digress a little from the study of the dominant and unknown dark matter abundance, and look at a subdominant, yet guaranteed part of the dark matter sector: the Cosmic Neutrino Background ($C_{νB}$). We show that the recently discovered high-energy
astrophysical neutrinos in IceCube can be used to probe the $\nu$B and secret neutrino interactions ($\nu\nu \rightarrow \nu\nu$, denoted as $\nuSI$).

In the Standard Model, the $\nu$B, similarly to the CMB, is the thermal neutrino relic of the early Universe, produced essentially through the freeze-out mechanism. $\nu$B is notoriously difficult to detect directly due to the small cross section of low energy neutrinos. The most promising method to infer their presence is to detect their suppression on the cosmic structure formation. Cosmologists have thus set the most stringent constraints on the sum of neutrino masses, and have high hopes of a detection in the near future.

However, with the discovery of the high-energy astrophysical neutrinos by IceCube, we point out that the propagation of high-energy neutrinos through the $\nu$B can be used to test $\nu$B itself and $\nuSI$. Nominaly, the neutrino-neutrino interaction cross section is too small to see any effect on the flux of high-energy astrophysical neutrinos. However, some neutrino-mass models, required to explain the extraordinarily small, yet non-zero, neutrino masses, predict large $\nu$SIs.

We demonstrate that the observation of high-energy astrophysical neutrinos itself already sets new constraints on $\nuSI$. We model the propagation of high-energy neutrinos through the $\nu$B using a propagation equation, which, for the first time, takes into account both attenuation and regeneration of neutrinos simultaneously. If the $\nuSI$ is mediated by new particles at the MeV scale, absorption features could also be present in the final spectrum. Interestingly, this may be able to explain the gap and/or cutoff seen in the IceCube event spectrum.

1.4.3 Searching for Sterile Neutrino Dark Matter with Fermi-GBM

In Chap. 4, we perform a novel analysis with Fermi-GBM to search for signals of the decay of sterile neutrino dark matter.

Although sterile neutrino dark matter does not couple directly to Standard Model particles, heavier sterile neutrinos can decay into lighter active neutrinos through neutrino mixing. In particular, the radiative channel $\chi \rightarrow \nu_a + \gamma$, where a sterile neutrino decays into an active neutrino plus a photon, is the most promising channel for detection. This
is because two-body decays involving photons produce line spectra. This sharp feature allows experiments to more effectively separate the signal from most backgrounds, which are typically smooth in their energy spectra.

As a result, sensitive X-ray telescopes, such as INTEGRAL, Chandra, XMM-Newton, and Suzaku have all been used to set stringent constraints on sterile neutrino dark matter. However, one particular X-ray energy range, 10–25 keV (or, equivalently, 20–50 keV dark matter mass), is outside the sensitivity of these instruments. In fact, the best dark matter constraint in this energy range came from HEAO-1, which flew in the late 1970s. This leaves a troublesome gap in the sterile neutrino dark matter parameter space.

We point out that the Gamma-ray Burst Monitor (GBM), the secondary instrument on board of the Fermi Gamma Ray Space Telescope (Fermi), is a promising tool for sterile neutrino dark matter searches. It not only covers the aforementioned energy window, but with its large field of view, large effective area, and decent energy resolution, it can potentially improve the constraint significantly.

The main difficulty is that, as the name of the instrument suggests, GBM was designed to search and localize transients sources in the sky, whereas the dark matter emission is diffuse and steady. After significant amount of efforts in cleaning the GBM data, we are able to produce a diffuse X-ray sky map with GBM. Not only is this the only sky map at this energy range, it also clearly detects the galactic astrophysical component. With this data set, we are able to set a new constraint on sterile neutrino dark matter, improving the limit by a factor of about ten compared to previous constraints.

1.4.4 Dark matter velocity spectroscopy

In Chap. 5, we propose a novel technique for diagnosing any tentative dark matter line signals with sensitive next-generation instruments.

As mentioned before, dark matter annihilation or decay producing a line spectrum in the final state has been considered the most sensitive channel for dark matter detection due to relatively easy background rejection. This so-called “smoking-gun signature,” however, is not always clear, due to systematic effects. In the X-ray band, there are also astrophysical
lines from atomic transitions, which, if modeled incorrectly, could be confused with dark matter signal.

This situation is particularly interesting given that recently a tentatively dark matter line signal at 3.5 keV was discovered from various astrophysical sources by several groups [62, 63]. It is difficult to definitively determine the origin of this line with current instruments and data. Naturally, hope lies on next-generation instruments, such as *Astro-H*, *Micro-X*, and *Athena*. These instruments offer significant improvement on energy resolution, at the $10^{-3}$ level, which could be essential for settling the question of the origin of the 3.5 keV line.

In this work, we propose a novel set of observations and signatures that could cleanly separate lines of dark matter, astrophysical, and detector origin. The key insight here is realizing that with $10^{-3}$ energy resolution, one is sensitive to the speed of dark matter particles and gases (which produce astrophysical X-ray lines) in our galaxy. Importantly, due to the intrinsic difference between how dark matter and gas move around in the galaxy, they would produce completely opposite Doppler shifts for a line signal. This means that, along some directions, dark matter particles are collectively moving towards us, while gas particles are collectively moving away from us. Instruments like *Astro-H* can pick up these shifts just like a speed camera, and, hence, determine the nature of the line. This technique, first realized in this work, not only could settle the case for the 3.5 keV line, but is also applicable to other dark matter line searches in the future, provided that instruments with the required precision are available.

### 1.4.5 The solar gamma ray puzzle

In Chap. 6, we perform an observational study of the puzzling gamma rays from the Sun and discover two interesting features. These gamma rays will play a key role in precisely understanding the backgrounds for dark matter searches from the Sun.

As discussed in 1.3.1, WIMP dark matter can be captured in the core of the Sun and annihilate into neutrinos, which has also been considered as smoking-gun signatures of dark matter. However, high-energy cosmic rays interacting with the solar atmosphere would also produce neutrinos, similar to how Earth’s atmospheric neutrinos are produced. These solar
atmospheric neutrinos are a problematic background for near-future solar WIMP searches with neutrino telescopes. To better understand this background, one must then understand how cosmic rays interact with the Sun.

In the seminal work of Seckel, Stanev, and Gaisser (denoted as SSG1991 [64]), they provided the first and only calculation that modeled the effect of solar magnetic fields on cosmic-ray interactions in the solar atmosphere. They also predicted that these interactions can produce detectable gamma rays from the Sun. More than twenty years later, these gamma rays were detected precisely by the Fermi. The gamma-ray observation was found to be in disagreement with the SSG1991 prediction: the gamma-ray Sun is about ten times brighter than expected!

It is now clear that the problem of how cosmic rays interact with the Sun needs to be revisited. This is required to understand the gamma-ray data, and more importantly, to understand the neutrino background for solar dark matter searches.

As a first step toward this ultimate goal, we use Fermi to improve the solar gamma-ray observation. First, we are able to detect gamma rays that are 10 times more energetic than the official Fermi analysis. In addition, we also discover that the solar gamma-ray flux varies with time and anticorrelates with solar activity. Though this feature is not surprising, it was not theoretically predicted and its amplitude is not easily understood, such as through cosmic-ray modulation observed on Earth. These features will surely provide important clues for understanding the underlying physics of cosmic rays interacting with the Sun. Lastly, we point out that ground-based water Cherenkov telescope, such as HAWC and LHAASO, can provide unique insights about solar gamma rays in the TeV range.
Chapter 2

Resolving Small-Scale Dark Matter Structures Using Multi-Source Indirect Detection

The extragalactic dark matter (DM) annihilation signal depends on the product of the clumping factor, \( \langle \delta^2 \rangle \), and the velocity-weighted annihilation cross section, \( \sigma v \). This “clumping factor–\( \sigma v \)” degeneracy can be broken by comparing DM annihilation signals from multiple sources. In particular, one can constrain the minimum DM halo mass, \( M_{\text{min}} \), which depends on the mass of the DM particles and the kinetic decoupling temperature, by comparing observations of individual DM sources to the diffuse DM annihilation signal. We demonstrate this with careful semi-analytic treatments of the DM contribution to the diffuse Isotropic Gamma-Ray Background (IGRB), and compare it with two recent hints of DM from the Galactic Center, namely, \( \sim 130 \) GeV DM annihilating dominantly in the \( \chi \chi \rightarrow \gamma \gamma \) channel, and \( (10-30) \) GeV DM annihilating in the \( \chi \chi \rightarrow b \bar{b} \) or \( \chi \chi \rightarrow \tau^+ \tau^- \) channels. We show that, even in the most conservative analysis, the Fermi IGRB measurement already provides interesting sensitivity. A more detailed analysis of the IGRB, with new Fermi IGRB measurements and modeling of astrophysical backgrounds, may be able to probe values of \( M_{\text{min}} \) up to \( \sim 1 \, M_\odot \) for the 130 GeV candidate and \( \sim 10^{-6} \, M_\odot \) for the light DM candidates. Increasing the substructure content of halos by a reasonable amount would further improve these constraints.

The contents of this chapter were published in [2].

2.1 Introduction

The observed universe is well explained by the \( \Lambda \)CDM cosmological model (\( \Lambda \) Cold Dark Matter). A large fraction, \( \Omega_\Lambda \), of its energy density is in the form of enigmatic dark energy, and the rest, \( \Omega_M \), is mostly non-relativistic matter and a tiny fraction of relativistic particles.
A major fraction of $\Omega_M$ has no detectable electromagnetic interactions, thus is termed Dark Matter (DM). From its gravitational effects on different length scales, DM is determined to have an energy density fraction $\Omega_\chi$. The particle nature of DM is largely unknown.

Identifying the fundamental particle nature of DM is one of the most important problems in contemporary science. A well-motivated DM candidate is the generic Weakly Interacting Massive Particle (WIMP), produced as a thermal relic in the early universe [65, 66, 67]. DM that self-annihilates at the electroweak scale naturally produces the observed DM abundance. The precise value of the thermally averaged total annihilation cross section that determines the DM abundance depends on several parameters [11]. A larger value will delay chemical decoupling, which would underproduce DM relative to the observed abundance, and vice versa.

After freeze-out, DM will continue to self-annihilate but at a cosmologically negligible rate. At the present epoch, DM is non-relativistic and is no longer thermally distributed. As a result, the velocity-weighted cross section (or simply annihilation cross section), $\sigma v$, which controls the annihilation rate now, could be a function of relative velocity and thus depends on the phase space of the DM structures. In this work, we consider the simplest case where $\sigma v$ is velocity independent over the relevant range of velocities of cosmic DM. In this case (s-wave), if the total value of $\sigma v$ now differs from the value of thermally-averaged $\langle \sigma v \rangle$ that determines the relic abundance, it could imply a dominantly p-wave annihilation cross section [68, 69], non-trivial velocity dependence of $\sigma v$ [70, 71], or some special thermal scenarios [72].

DM self-annihilation opens up the possibility of remotely detecting its annihilation products from concentrated DM sources, i.e., indirect detection. Together with directly detecting nuclear recoils in underground experiments, DM production in collider experiments, and DM influence of astrophysical systems, these four types of DM detection provide crucial and complementary information on the particle nature of DM [7].

Indirect detection is a powerful way to detect DM. However, it suffers from problems of low signal-to-noise ratios due to large and complicated astrophysical backgrounds. One strategy is to search for smoking-gun signatures that would allow for effective separation
between background and signal. Examples of such signatures are spectral lines [73, 74, 75, 76], spectral cut-offs [77, 78, 79, 80, 81], or distinct anisotropy signals [82, 83, 84]. Since annihilation signals are proportional to the DM density squared [54, 85], it is advantageous to search for these signatures from regions where DM is clustered, e.g., the Galactic Center (GC), dwarf galaxies, galaxy halos, galaxy clusters, or the diffuse signal from annihilation in all the DM structures in the Universe.

The diffuse extragalactic DM annihilation signal is particularly difficult to predict robustly [59, 60, 61]. It depends not only on the self-annihilation cross section and DM density distribution within halos, but also on the statistics of cosmological DM halos such as the halo abundances and their concentrations at small scales. Dense DM structures are expected to be present, whether they are isolated or residing within halos as substructures. They span down to the smallest possible bound DM objects with mass $M_{\text{min}}$.

The value of $M_{\text{min}}$ corresponds to the cutoff of the matter power spectrum, $k_{\text{max}}$, which is usually set by either the free-streaming scale after kinetic decoupling [13] or the scale of acoustic oscillation with the radiation fields [86]. These scales depend on the DM mass and its elastic coupling to the cosmic background particles. Parameter scans of some supersymmetric models show a large range of possibilities, $10^{-12} M_\odot < M_{\text{min}} < 10^{-3} M_\odot$ [87, 88].

Direct observation of microhalos is very difficult because they are not massive enough to host stars. Current gravitational lensing probes are only sensitive to relatively massive halos ($> 10^6 M_\odot$) [89, 90, 91]. Nanolensing [92, 93] or proper motion detection [94, 95] might be able to probe smaller scales. The presence of microhalos, however, changes the clustering property of DM structures, which is encoded in the clumping factor, $\langle \delta^2(z) \rangle$ [59, 60, 61, 23], defined as the mean of the matter overdensity squared. The clumping factor boosts the annihilation rate relative to the mean background density, and is completely degenerate with the effect of $\sigma v$ for extragalactic diffuse DM annihilation signals. Both the annihilation cross section and the clumping factor are important DM parameters to be determined.

In this work, we demonstrate how to break this degeneracy and constrain both $M_{\text{min}}$ and $\sigma v$ by comparing the diffuse Isotropic Gamma-ray Background (IGRB) [28, 96] with
tentative DM annihilation signals from the GC [97, 98, 99, 31]. These excesses of events from GC might be DM signals, astrophysical phenomena, or experimental artifacts. It is important to scrutinize them as much as possible. We therefore consider them as a proof of principle as well as a test. Multiple-source analyses for DM indirect detection have proven to be invaluable for constraining DM candidate signals [100, 101].

Throughout this work, we use \( M_x = M/10^2 M_\odot \) and cosmological parameters from the Planck mission (\( \Omega_\Lambda = 0.6825, \Omega_M = 0.3175, \Omega_\chi = 0.1203 h^{-2}, h = 0.6711, \) and the Hubble constant, \( H_0 = 100 h \text{ km s}^{-1} \text{ Mpc}^{-1}, n_s = 0.96, \sigma_8 = 0.8344 \)) [16].

In Sec. 2.2, we calculate the contribution of DM annihilation signals in the IGRB, showing the dependence of annihilation signals on \( M_{\text{min}} \). In Sec. 2.3, we discuss the constraints obtained by combining the GC and IGRB observations for DM candidate events. Lastly, we summarize in Sec. 2.4.

### 2.2 IGRB from DM annihilation

The diffuse IGRB is the isotropic component of the gamma-ray sky, in principle mostly contributed by unresolved extragalactic astrophysical sources. The Fermi Gamma-Ray Space Telescope measures the IGRB by careful reductions of the Galactic astrophysical components, astrophysical sources, and detector backgrounds [28]. In the presence of DM annihilation, it contains an irreducible isotropic Galactic component and the diffuse extragalactic component [102, 103]. In this section, we discuss each component and their dependence on DM substructures.

#### 2.2.1 Isotropic Galactic component

Since Fermi is embedded in the Milky Way (MW), an irreducible isotropic contribution of DM self-annihilation to the IGRB comes from the MW halo. We first review the case of DM annihilation in the MW.

The smooth DM density distribution in the MW, \( \rho_x^{\alpha\beta\gamma}(r) \), can be parametrized by the following form [104],

\[
\rho_x^{\alpha\beta\gamma}(r) = \frac{c}{\sqrt{2\pi} \sigma} e^{-\frac{c^2 r^2}{2 \sigma^2}} \sum_{i=1}^{n} a_i \phi_{i}(r) \psi_{i}(\theta, \phi) \]

where \( c \) is the scale parameter, \( \sigma \) is the width parameter, \( \phi_i \) and \( \psi_i \) are the radial and angular functions, and \( a_i \) are the coefficients.
\[ \rho^\alpha\beta\gamma(r) = \rho_\odot \left[ \frac{r}{r_\odot} \right]^{-\gamma} \left[ \frac{1 + (r_\odot/r_s)^\alpha}{1 + (r/r_s)^\alpha} \right]^{\frac{2-\gamma}{\alpha}}, \]  

(2.1)

where \( r \) is the galactocentric distance, \( \rho_\odot = 0.4 \pm 0.1 \text{ GeV cm}^{-3} \) is the DM density in the solar neighborhood \([105, 106]\), \( r_\odot = 8.5 \text{ kpc} \) is the solar distance to the GC, and \( r_s \) is the scale radius. The shape of the profile is determined by the parameters, \( \alpha, \beta, \gamma, \) and the scale radius, \( r_s \). The commonly used NFW profile in the MW takes the values \( \{\alpha, \beta, \gamma, r_s\} = \{1, 3, 1, 20 \text{ kpc}\} \); the cored isothermal (ISO) profile takes \( \{2, 2, 0, 3.5 \text{ kpc}\} \).

Another profile favored by recent simulations is the Einasto (EIN) profile \([107]\),

\[ \rho^\text{Ein}_\gamma(r) = \rho_\odot \exp\left( -\frac{2}{\alpha_E} \frac{r}{r_\odot} \frac{r^{\alpha_E}}{r_s^{\alpha_E}} \right), \]  

(2.2)

with \( \alpha_E = 0.17 \) and \( r_s = 20 \text{ kpc} \).

The gamma-ray (number flux) intensity due to the Galactic Halo DM self annihilation, \( I_\gamma^G(E_0) \), is

\[
I_\gamma^G(E_0) = \frac{dN_\gamma}{dA d\ell_0 d\Omega dE_0} = \frac{\sigma v r_\odot \rho_\odot^2}{8\pi m_\chi^2} J(\psi) \frac{dN_\gamma}{dE_0} = 3.7 \times 10^{-10} \text{ cm}^{-2} \text{s}^{-1} \text{ sr}^{-1} \text{ GeV}^{-1} \times J(\psi) \]  

(2.3)

where \( m_\chi \) is the DM mass, \( E_0 \) is the observed photon energy, and \( dN_\gamma/dE_0 \) is the photon energy spectrum per annihilation. The so-called J-factor, \( J(\psi) \), is the dimensionless line of sight integral of the density squared, and depends on the DM distribution in the Galactic halo, including halo substructures.

**Galactic smooth halo**

The J-factor for the smooth DM density distribution for an observer within the halo, as a function of the angle between the line of sight and the GC, \( \psi \), is

\[
J_S(\psi) = \frac{1}{r_\odot \rho_\odot^2} \int_0^{\ell_{\text{max}}} \rho_\chi^2(r(\psi, \ell)) d\ell.
\]  

(2.4)
Figure 2.1: The normalized line of sight integral of the DM density squared (the J-factor) as a function of the viewing angle, $\psi$. The J-factor for EIN, NFW and ISO profiles for the smooth halo are shown with dotted lines. The contributions of substructures to the J-factors for the LOW substructure case, assuming $M_{\text{min}} = 10^{-6} M_\odot$, are shown with dotted-dashed lines. The total J-factor (smooth + LOW substructure) for just the one case (NFW), is shown with a bold solid line.
The integration limit is determined by the size of the MW DM halo: \( \ell_{\text{max}} = \sqrt{R^2 - r_\odot^2 \sin^2 \psi + r_\odot \cos \psi} \), where \( R = 200 \text{kpc} \) is the halo’s virial radius. The J-factor is largely insensitive to the exact value of \( R \). The galactocentric distance is \( r(\psi, \ell) = \sqrt{r_\odot^2 - 2\ell r_\odot \cos \psi + \ell^2} \). To compare the theoretical expectations with detector observables, one simply average the J-factor over the detector angular resolution or the field of view. The J-factors for smooth halos are shown in Fig. 2.1 and Fig. 2.2.

![Figure 2.2: Same as Fig. 2.1, but for the HIGH substructure case.](image)

In principle, the isotropic Galactic component of the IGRB from DM annihilation is
given by the zeroth component of the spherical harmonic decomposition, or equivalently the average of the field of view of the observation. This might be complicated by the masking of the sky (e.g., the bright Galactic plane) and all the background reductions performed by the Fermi Collaboration. We therefore take the most conservative estimate by taking the constant J-factor from the Anti-GC \((\psi = \pi)\),

\[
J_{iso} \equiv \frac{1}{d\Omega} \int d\Omega J_S(\pi) = J_S(\pi). \tag{2.5}
\]

A more detailed analysis for determining the isotropic Galactic component, possibly by including a DM template to the Fermi IGRB analysis, would further improve our DM constraint.

**Galactic substructure enhancement**

In \(\Lambda\)CDM, structures form hierarchically. The smallest DM halos are expected to form first. Some of these small halos subsequently merge and eventually may live in large host halos of galaxies or clusters. During structure formation, the small halos that are captured by larger halos are tidally disrupted and their low-density outer layers are stripped. The dense cores, however, could very well survive and become subhalos of the main halo \([108, 109, 110]\) (however, also see \([111]\)). We collectively define all of these surviving DM clumps to be substructures. High resolution simulations are beginning to resolve substructures down to the resolution limit \([112, 107]\). These substructures can lead to many interesting DM phenomenologies \([14]\).

Smaller DM structures tend to have higher concentrations \([113]\), which can be understood by their earlier formation time at which the background density is higher. Therefore, although substructures may not occupy much of the total volume of a halo, they could significantly enhance the DM annihilation rate of a halo.

To describe the substructure boost to the isotropic Galactic component of the IGRB, we use the theoretical model proposed by Kamionkowski and Koushiappas \([114]\). This model was later calibrated to high-resolution simulations by Kamionkowski, Koushiappas, and Kuhlen \([20]\,\text{hereafter K10}\), and therefore can be used to calculate the boost of the
Figure 2.3: The formation redshift, $z_f$, versus $M_{\text{min}}$, for the first generation halos. The formation redshift is obtained by requiring the linear mass variance be equal to the characteristic overdensity, $\sigma_L(M, z) = 1.686$ \cite{13, 14}. For $\sigma_L(M, z)$, we use the fitting functions of Eisenstein and Hu \cite{15}, which are evaluated and normalized with the Planck cosmological parameters \cite{16}. For illustration, we also show $z_f$ for the $n_s = 1$ case as well as the extrapolated results from simulations by Ludlow et al. \cite{17} (we take $z_f$ to be $z_{-2}$).
Figure 2.4: The characteristic density of the first generation halos, $\rho_{\text{max}}$, versus $M_{\text{min}}$, for the corresponding cases of $z_f$ in Fig. 2.3 using Eq. 2.8. The substructure boost is approximately proportional to $\rho_{\text{max}}$ (Eq. 2.6).
Galactic annihilation rate relative to a smooth halo density profile for the MW. The Galactic local boost factor as a function of $r$, for a velocity independent $\sigma_v$, is

$$B(r) = f_s(r) e^{\delta f} + [1 - f_s(r)] \frac{\alpha}{1 - \frac{\rho_{\max}}{\rho_{\chi}(r)}}^{1-\alpha} - 1,$$

(2.6)

where $f_s(r)$ is the fraction of the volume that would be occupied by the smooth halo component, and $\rho_{\max}$ is the highest DM substructure density. The DM substructure fraction, $1 - f_s$ is

$$1 - f_s(r) = \kappa \left( \frac{\rho_{\chi}(r)}{\rho_{\chi}(100 \text{ kpc})} \right)^{-0.26}.$$  

(2.7)

The first term in $B(r)$ is the boost from the smooth halo component, by taking into account random fluctuations in the density. The second term describes the boost due to substructures. The substructure fraction is normalized by the parameter $\kappa$. Following K10, these parameters are determined to be $\delta_f = 0.2$, $\alpha = 0$, and $\kappa = 0.007$.

It has been pointed out in Fornasa et al. [115] that the original K10 model tends to give a conservative substructure enhancement compared to other studies [116, 21, 117], mainly due to different methodologies. This can be reconciled by increasing the substructure survival fraction parameter, $\kappa$, from 0.007 to 0.15 – 0.2. In subsequent discussions, we refer to $\kappa = 0.007$ as the LOW substructure case as a conservative estimate, and $\kappa = 0.18$ as the HIGH substructure case as an optimistic scenario.

The boost factor is approximately proportional to the characteristic density of the minimum halo mass, $\rho_{\max}$. It depends on the natal concentration, $c_0$, the formation redshift, $z_f$, and the mass, $M_{\min}$, of the first generation halos [20],

$$\rho_{\max}(M_{\min}) = \frac{1}{12} \frac{c_0^3}{\ln(1 + c_0) - \frac{c_0}{1+c_0}} \Delta \rho_c(z_f(M_{\min})),$$

(2.8)

where $\Delta = 200$ is the halo over-density. $\rho_c(z) = \rho_c(0) \mathcal{H}^2(z)$, where $\rho_c(0) = 1.05 \times 10^{-5} h^2 \text{GeV cm}^{-3}$ is the critical density, and $\mathcal{H}^2(z) = \Omega_{\Lambda} + \Omega_M (1 + z)^3$ is the Hubble function squared. The dependence of $M_{\min}$ in $\rho_{\max}$ is mainly on $z_f(M_{\min})$, as the natal
concentration is fairly constant for \( z_f > 5 \) [118]. We follow K10 and take \( c_0 \) to be 3.5.

Parameter scans of some supersymmetric models show that \( 10^{-12} M_\odot < M_{\text{min}} < 10^{-3} M_\odot \) [87, 88]. Different models can drastically change the prediction for the value of \( M_{\text{min}} \). We consider \( 10^{-12} \) and \( 10^0 M_\odot \) as the lower and upper extreme cases for CDM, and adopt \( 10^{-6} M_\odot \) as the reference value. We consider \( 10^6 M_\odot \) unlikely for simple Cold DM models. Such a high value would require special DM physics (e.g. see Ref. [119], and also [120, 33]) and is within the sensitivity of gravitational lensing probes [89, 90, 91].

To estimate the value of \( \rho_{\text{max}}(M_{\text{min}}) \), we need to know the corresponding \( z_f \). For the first generation halos, this can be estimated using cosmological perturbation theory [13, 14], since they are the first nonlinear structures of the Universe. Then \( z_f \) is implicitly defined by \( \sigma_L(M, z_f) = 1.686 \), where 1.686 is the characteristic over-density of the 1-\( \sigma \) linearized density fluctuation when halo collapse has occurred (see Ref. [121] and reference therein). \( \sigma_L(M, z) \) is the linear mass variance defined by

\[
\sigma_L^2(M, z) = \int_0^{\infty} W^2(kR) \Delta_L^2(k, z) \frac{dk}{k},
\]

where \( W(kR) \) is the Fourier transform of the top hat window function, \( \Delta_L^2 \) is the dimensionless linear power spectrum, and \( R \) is the comoving length scale. The mass of the collapsed halos can be estimated by \( M \simeq (4/3) \pi R^3 \rho_c(z_f) \). We evaluate and normalize the mass variance using the fitting formula by Eisenstein and Hu [15], according to the Planck cosmological parameters [16]. We also take into account the non-unity of the spectral index \( (n_s = 0.96, \) without running), which is measured by the Planck collaboration with high significance. The effect of the slight tilt is amplified at small scales that we are interested in. Varying the index by approximately the 1-\( \sigma \) Planck limit \( (n_s = 0.96 \pm 0.01) \) yields a 5\% change in \( z_f \) for \( M_{\text{min}} = 10^{-6} M_\odot \), which translates into a 15\% change for \( \rho_{\text{max}} \). We have considered only the 1-\( \sigma \) density fluctuations which collapse into halos. Higher-\( \sigma \) density fluctuations will collapse even earlier, and are thus denser, but they are correspondingly rarer.

In Fig. 2.3, we show \( z_f \) as a function of \( M_{\text{min}} \) for \( n_s = 0.96 \). For comparison, we also show the case for \( n_s = 1 \) and an extrapolation from the simulation of Ludlow et al. [17].
The hierarchical nature of structure formation is apparent in this plot, with the smaller halos forming earlier. In Fig. 2.4, we show the corresponding $\rho_{\text{max}}$ evaluated using Eq. 2.8.

To incorporate the effect of substructure, we insert the boost factor into the line of sight integral to obtain the J-factor with substructure enhancement, $J_B(\psi)$,

$$J_B(\psi) = \frac{1}{r_\odot \rho_\odot} \int_0^{\ell_{\text{max}}} \rho_\chi^2 [r(\psi, \ell)] \cdot B [r(\psi, \ell)] d\ell.$$  \hfill (2.10)

In Fig. 2.1 and Fig. 2.2, we show the effect of substructure on the J-factor for the LOW and HIGH substructure boost cases, respectively.

It is well known that the J-factor near the GC is very profile dependent [102]. However, substructures have relatively small enhancements to the J-factor at the GC, since substructures are more susceptible to tidal effects in high density regions. Therefore, DM signals from the GC can be considered to be substructure independent. The K10 substructure model qualitatively reflects this feature. However, the calibration to simulation inevitably breaks down near the GC, due to finite resolution effects [20]. Since details at the GC have no effect to our result, we assume the K10 model is valid at all regions.

As a result, any $\sigma v$ extracted from GC analysis is subjected to profile dependence, but independent of the underlying substructure assumptions. On the other hand, the J-factor is practically profile independent at large angles. We therefore find the isotropic Galactic component depends mostly on the substructure content of the halo, but not the density profile. The substructure enhancement for the isotropic Galactic component depends sensitively on the survival fraction $\kappa$. For the LOW (HIGH) substructure case, the boost is at most a factor of 1.5 (10).

It is also interesting to see that at $\sim 30^\circ$, the DM signal is the least uncertain relative to both density profile [102, 116] and substructure scenarios. Therefore, one would ideally prefer to detect Galactic DM annihilation from such angles to minimize the astrophysical uncertainty on DM density distribution.
2.2.2 Extragalactic component

The gamma-ray (number flux) intensity from extragalactic DM self-annihilation, \( I_{\gamma}^{EG}(E_0) \), is given by the cosmological line of sight integral,

\[
I_{\gamma}^{EG}(E_0) = \frac{\sigma v}{8\pi} \int \frac{v_c dz}{H_0 H(z) (1 + z)^3} \left( \frac{\bar{\rho}_\chi(z)}{m_\chi} \right)^2 \frac{dN_\gamma(E)}{dE} e^{-\tau(z,E_0)}
\]

Where \( E \) is the center-of-momentum frame energy given by \( E = E_0 (1 + z) \), \( v_c \) is the speed of light, and \( \bar{\rho}_\chi(z) \) is the cosmological mean DM density. The clumping factor, \( \langle \delta^2(z) \rangle \), which measures the cosmologically averaged DM density squared, relative to the mean DM density squared, \( \langle \delta^2(z) \rangle = \langle \rho_\chi^2(z) \rangle / \bar{\rho}_\chi^2(z) \). High-energy gamma rays propagating through intergalactic space will suffer attenuation due to the Extragalactic Background Light (EBL). This effect is included in the attenuation factor, \( e^{-\tau(E_0,z)} \) (see Sec. 2.2.3).

The clumping factor is the main theoretical astrophysical uncertainty in evaluating the expected DM annihilation intensity. We review how to evaluate the clumping factor using the Halo Model approach, with or without substructures in massive halos. We also review how to evaluate the equivalent quantity using the Power Spectrum approach, which is complementary to the halo model approach in terms of theoretical uncertainties.

Halo Model approach with smooth halos only

The clumping factor for smooth halos, \( \langle \delta_S^2(z) \rangle \), can be calculated using the Halo Model framework [59, 60],

\[
\langle \delta_S^2(z) \rangle = \frac{\langle \rho_\chi^2(z) \rangle}{\bar{\rho}_\chi^2(z)} = \frac{1}{\bar{\rho}_\chi^2(z)} \int dM \frac{dn}{dM}(z) \int_{r<R} dV \rho_\chi^2(\rho_s, r_s)
\]

where \( \bar{\rho}_\chi(z) \) is the cosmological mean DM density.
Figure 2.5: The normalized halo mass function \((1/\rho_\chi)M^2 dn/dM\) plotted versus \(M\) for redshift \(z = 0, 2, 4\). The halo mass function as a function of the linear mass variance is given by P12 [18]. The redshift evolutions of the fitting parameters are given by Tinker et al. [19].
Figure 2.6: The concentration parameter, \( c(M, z) \), plotted against \( M \) for redshift \( z = 0, 2, 4 \). The concentration mass relation as a function of the linear mass variance is again given by P12 [18]. For comparison, we also show the concentration if we simply extend the concentration-mass relation to small scales using the analytic function given in P12.
where \( \frac{dn}{dM} \) is the distinct halo mass function in physical units, which describes the number density of bound objects with mass \( M \) at a particular redshift. These objects are distinct halos — their centers are not inside the virial radius of larger halos. The density profile of a particular halo, \( \rho_\chi(\rho_s, r_s) \), is characterized by its scale density, \( \rho_s \), and scale radius, \( r_s \), which in turn depends on the halo mass, \( M \), and redshift, \( z \). In Eq. 2.12, \( \langle \rho^2 \rangle \) denotes the volume average of halos of all masses of the density squared and \( \bar{\rho} \) denotes the cosmic mean density.

We use the distinct halo mass function from Prada et al. ([18], hereafter P12). The P12 halo mass function is obtained by fitting to four cosmological simulations. The fitting functional form follows from the Press-Schechter theory and its extensions [122, 123]. The halo mass function describes the full hierarchy of distinct cosmological DM halos down to \( M_{\text{min}} \), and the cosmology dependence enters through the linear mass variance, \( \sigma_L \) (Eq. 2.9). For the redshift dependence of the fitting parameters, we follow those from Tinker et al. [19]. In Fig. 2.5 we show the halo mass function for several redshifts. The low mass dependence is slightly harder than the critical \( M^{-2} \) behavior.

The volume integral of the density squared can be simplified using two halo mass relations, which convert the \( \{\rho_s, r_s\} \) dependence to only the concentration parameter, \( c(M, z) = R/r_s \). The first one is

\[
M = \frac{4}{3} \pi R^3 \Delta \rho_c(z),
\]

where \( R \) is the virial radius of the halo. The second halo mass relation is

\[
M = \int_{r < R} dV \rho_\chi(\rho_s, r_s)
\]

which is integrated up to the virial radius.

The clumping factor can now be written as

\[
\langle \delta_3^2(z, M_{\text{min}}) \rangle = \frac{1}{\Omega_X} \int_{M_{\text{min}}}^{\infty} dM \frac{\mathcal{H}(z)^2}{(1+z)^3} \frac{1}{\bar{\rho}_\chi(z)} \frac{dn}{dM} \frac{M \Delta}{3} \times \int d\hat{c} P(c, \hat{c}) \hat{c}^3 \frac{I_2(\hat{c})}{I_1(\hat{c})^2},
\]

where \( \mathcal{H}(z) \) is the Hubble parameter at redshift \( z \), and \( \bar{\rho}_\chi(z) \) is the cosmic mean density at redshift \( z \).
where we have introduced the dimensionless integral

\[ I_n(c) = \int_0^R (dr/r_s) (r/r_s)^2 (\rho (r)/\rho_s)^n \]

and the log-normal distribution, \( P(c, \tilde{c}) \), with constant 1-\( \sigma \) deviation \( \sigma_{\log_{10}} = 0.13 \) (or \( \sigma_{\ln} = 0.3 \)) [124, 125] around the mean concentration parameter, \( c(M, z) \). We simplify the formalism by defining the effective cut-off in the Halo Mass function to be the minimum halo mass, \( M_{\text{min}} \), thus ignoring objects with masses below \( M_{\text{min}} \) [126].

We argue that this definition of \( M_{\text{min}} \) is effectively equivalent to the \( M_{\text{min}} \) in the Galactic substructure calculation. The smallest substructure mass in halos may be less than the smallest cosmological halo mass because of tidal disruption in merging. But the relevant part of DM annihilation, which is the maximum density of the substructures, can be assumed to be unaffected by tidal disruptions [108, 109, 110].

The last ingredient we need is the mean concentration parameter \( c(M, z) \), which is a quantitative measure of halo concentrations. We use the analytic function from P12, which is derived from cosmological simulations and agrees well with cluster observations. The P12 result for \( c(M, z) \) shows a remarkably tight relation with the linear matter mass variance, \( \sigma_L(M) \), for which we again use the linear mass variance given by Eisenstein and Hu with the Planck cosmology. It is intuitive that halo concentrations would tightly correlate with the linear mass variance, since the latter is intimately related to halo formation [113].

We show the concentration-mass relation for NFW profiles in Fig. 2.6 for \( z = 0, 2, 4 \). For comparison, we also show the concentration if we simply extend the fitting function for \( \sigma_L(M) \) from P12 to small scales. Recent microhalo simulations have shown that a \( \sim 10^{-7} M_\odot \) first generation halo has concentration \( 57 < c < 84 \) at redshift zero [127], with a mean value of 72. We find that a naive substitution of the linear mass variance from [15] slightly underestimates the concentration at small mass scales. Therefore, we change one of the fitting parameter (\( c \) in Eq. (16) in [18]) from 1.022 to 1.05. The resulting concentration increases from 67 to 73 at \( 10^{-7} M_\odot \), with negligible changes at large scales. From Fig. 2.6, we can see the concentrations grow slower than a naive power-law extrapolation to small halo masses. This “bending” of concentration is also recently noted by Ludlow et al. [128] and by Sánchez-Conde and Prada [129].

The rising concentration at large mass scales is a novel feature from the simulation [130].
This feature, though interesting, has no effect on our calculation due to the rapidly falling halo mass function at the corresponding mass and redshift.

In Fig. 2.7, we show the clumping factor as a function of $M_{\text{min}}$, using the P12 model with $\sigma_L(M)$ from [15]. At the extreme case of $M_{\text{min}} = 10^6 M_\odot$, the P12 model yields $\sim 4 \times 10^4$, consistent with similar evaluations [100, 77, 131]. This mass scale is within the simulation limits [131], and is also within the reach of gravitational lensing probes [89, 90, 91]. Therefore, we consider this to be the minimum DM clustering value, a lower bound to the clumping factor.

Microhalo simulations show that first generation halos have a steeper inner slope than the normal NFW profiles [132, 127]. This would enhance the annihilation signal from microhalos. One may also be interested in the profile dependence of the clumping factor. It is not straightforward, however, to change the density profile in calculating the clumping factor, since the value of the concentrations extracted from simulations depend on the assumed profile [18]. Nonetheless, the clumping factor is expected to be relatively insensitive to the density profile. For example, the total annihilation luminosity from the MW halo only experience a change of -20% or +30%, if isothermal or Einasto profiles are used. For simplicity, we use the NFW profile for all the evaluations.

**Halo Model approach with substructures**

In the above calculation, we assume each halo in the halo mass function has a universal smooth DM density profile. However, in addition to the cosmological isolated small halos, substructures within halos also contribute to the clumping factor.

Unlike the Galactic case, the observer is outside of all the halos observed and each halo has different mass and size. To incorporate substructure effects, we extend the K10 substructures model to different halo sizes, following the approach taken by Sánchez-Conde et al. [133] (also see Fornasa et al. [115]). To recalibrate the K10 model for different halo sizes, the substructure fraction needs to be modified,

$$1 - f_s(r) = \kappa \left( \frac{\rho_N(r)}{\rho_N(3.56 \times r_s)} \right).$$

(2.16)
Figure 2.7: The clumping factor at $z = 0$ versus $M_{\text{min}}$. Using the P12 [18] Halo Model, we obtain the clumping factor without substructure enhancement. Adding the K10 [20] substructure model, we show the substructure-enhanced clumping factor for LOW and HIGH cases. For comparison, we also show the clumping factor if we simply extrapolate the concentration relation in P12, the clumping factor with G12 [21] substructure model, and the clumping factor using the extrapolated halofit [22] non-linear power spectrum, following S12 [23]. In this work, we consider the LOW and HIGH scenarios as the conservative and optimistic substructure cases.
In doing so we have assumed the same radial dependence of $f_s(r)$ for all halo masses. The factor of 3.56 is the conversion factor from the Milky Way halo to the size of the simulation from K10. The local boost factor would enter inside the dimensionless integral $I_2$ in Eq. 2.15.

The substructure enhanced clumping factor, $\langle \delta^2_B(z, M_{\text{min}}) \rangle$, is

\begin{equation}
\langle \delta^2_B(z, M_{\text{min}}) \rangle = \frac{1}{\Omega_\chi} \int_{M_{\text{min}}} dM \frac{\mathcal{H}(z)^2}{(1+z)^3} \frac{1}{\bar{\rho}_\chi(z)} \frac{dn}{dM} \frac{M \Delta}{3} \int d\hat{c} P(c, \hat{c}) \hat{c}^3 \tilde{I}_2 \frac{I_2}{I_1},
\end{equation}

where

\begin{equation}
\tilde{I}_2(c, M_{\text{min}}) = \int_0^c d\hat{c} \frac{r}{r_s} \left( \frac{r}{r_s} \right)^2 \left( \frac{\rho_\chi(r)}{\rho_s} \right)^2 \cdot B(r, M_{\text{min}}).
\end{equation}

Recall that the $M_{\text{min}}$ dependence enters $B(r, M_{\text{min}})$ through $\rho_{\text{max}}$.

In Fig. 2.7, we show the clumping factor with substructures, $\langle \delta^2_B(0, M_{\text{min}}) \rangle$, for the LOW and HIGH substructure cases. For the LOW case, the substructure boost is small. For the HIGH case, the substructure boost is more important, where the enhancement ranges from about a factor of 2 in large $M_{\text{min}}$ to a factor of 6 in the smallest $M_{\text{min}}$.

For extragalactic halos one can also use the substructure boost by Gao et al. ([21], hereafter G12), which corresponds to the HIGH case in cluster scale. The substructure fraction for HIGH is tuned to match the G12 boost factor for an individual cluster scale halo [134, 135, 136], with $M_{\text{min}}$ assumed to be $10^{-6} M_\odot$. Taking into account the scaling factor of the G12 boost factor to $M_{\text{min}}$ ($\propto M_{\text{min}}^{-0.226}$ [116, 134]), the clumping factor for G12 is shown in Fig. 2.7. The clumping factor is even higher than the HIGH case for $M_{\text{min}}$ less than about $10^{-6} M_\odot$, but opposite otherwise. We see that even for the HIGH substructure case, the clumping factor is more conservative than the the cases where power-law extrapolation is used, due to a slower increase in small-scales.

The shape of the clumping factor is inherited from both the shape the halo mass function and the concentration-mass relation. The characteristic shape of the P12 clumping factors are due to the slower increase in concentration in smaller masses (Fig. 2.6), which ultimately traces back to the flattening of $\Delta^2_L$ in $\sigma_L(M)$ (Eq. 2.9). Physically, this reflects the property that small halos over a large range of mass formed in a relatively small period of time. As
a result, when predicting the clumping factor (i.e., the extragalactic DM annihilation flux),
decreasing $M_{\text{min}}$ leads to only a small increase in the clumping factor. In contrast, when
constraining $M_{\text{min}}$, a small improvement on the flux limit would lead to a large improvement
on the limit for $M_{\text{min}}$.

Additional clustering of DM can also be achieved by density spikes near Black Holes
[137, 138] or adiabatic contraction of DM halos [139]. On the other hand supernova feedback
might introduce a core to the density profiles for larger size halos [140, ?, 141]. Warm or
mixed DM, DM interactions with themselves [142, 143, 144], or with other particles [119] can
also significantly change dark matter density distributions. These effects deserve detailed
studies and are outside the scope of this work.

**Power Spectrum approach**

We have shown how to obtain the clumping factor using the Halo Model formalism. It
can alternatively be obtained from the $r \to 0$ limit of the two-point correlation function,
$\langle \delta(x + r)\delta(x) \rangle$, as shown by Serpico et al. ([23], hereafter S12). It can be expressed as an
integral of the non-linear power spectrum, $\Delta_{\text{NL}}(k, z)$,

$$
\langle \delta^2(z) \rangle = \lim_{r \to 0} \int_{0}^{k_{\text{max}}} \frac{dk}{k} \sin kr \frac{kr}{k} \Delta_{\text{NL}}(k, z),
$$

(2.19)

where $k_{\text{max}}$ is the cut-off of the non-linear power spectrum and corresponds to $M_{\text{min}}$ by
$M_{\text{min}} \approx (4/3)\pi(\pi/k_{\text{max}})^3 \bar{\rho}_X$.

Using the Power Spectrum approach, one has the obvious advantage that many un-
certainties of the Halo Model are collectively reflected in the non-linear part of the power
spectrum. The constraints from DM observation can be related to constraints on the shape
and cut-off of the non-linear power spectrum $\Delta_{\text{NL}}(k, z)$. It is, however, difficult to probe
the small-scale non-linear regime in theoretical treatments and simulations, and the physics
in Fourier space is more difficult to translate to physics in real space.

Nonetheless, the power spectrum approach is appealing for its simplicity and different
systematics. We selected the *halofit* model [22] following S12 and extrapolate it to the scales
relevant for our discussion of $M_{\text{min}}$. The resulting clumping factor is shown in Fig. 2.7. The
Power Spectrum result roughly agrees with the Halo Model approach.

Figure 2.8: The attenuation fraction of monochromatic gamma-ray signals from cosmological DM annihilation sources (Eq. 4.12) versus the emitted photon energy for different EBL models. We consider the EBL model from Gilmore et al. [24], “Best Fit 06” from Kneiske et al. [25], and the “Lower-Limit” from Kneiske and Dole [26]. All models considered are within 2-σ of the Fermi measurement [27]. We adopt the Gilmore model throughout.
Figure 2.9: The extragalactic DM annihilation redshift distribution (Eq. 2.21). We show the distributions for the P12 Halo Model, P12 Halo Model with K10 LOW, and HIGH substructure models. The upper set of three lines uses $M_{\text{min}} = 10^{-6} M_{\odot}$, while the lower set uses $M_{\text{min}} = 10^6 M_{\odot}$. The shape of the distribution varies mildly in different scenarios. Most of the DM annihilation signal comes from small redshifts.
2.2.3 EBL attenuation and redshift distribution

In this section we discuss the effect of the EBL attenuation and the general redshift behavior of the extragalactic DM annihilation signal.

EBL attenuation

It is evident from Eq. 2.11 that all the astrophysical and cosmological uncertainties are contained in the combination, \((1 + z)^3 (⟨\delta^2(z)⟩/H(z)) e^{-\tau(z,E_0)}\). In particular, only \(\tau(z,E_0)\) depends on the nature and energy of the messenger, therefore also on the mass and the annihilation channel of DM particles. It is instructive to explore the effect of the EBL attenuation by looking at the ratio of the attenuated total flux to the unattenuated photon flux, \(\eta(E_γ)\), for monoenergetic photon emission,

\[
\eta(E_γ) = \frac{\int dE_0 \int dz \frac{(1+z)^3}{H(z)} δ(E_0(1+z) - E_γ) e^{-\tau(z,E_0)}}{\int dE_0 \int dz \frac{(1+z)^3}{H(z)} δ(E_0(1+z) - E_γ)},
\]

where \(δ(E_0(1+z) - E_γ)\) is the Dirac-delta function connecting the observed energy and the emitted energy. This factor represents the relative flux suppression due to EBL absorption, according to DM clumping evolution.

In Fig. 2.8 we show \(\eta(E_γ)\) for a few different EBL models that are compatible with the latest Fermi results [27]. We use the P12 Halo Mass model with \(M_{\text{min}} = 10^{-6} M_{⊙}\) to evaluate \(\eta(E_γ)\). Different models share the same generic feature that attenuation affects annihilation signals with gamma-ray above \(\sim 50\) GeV, and the amount of attenuation is fairly consistent for different models. Throughout this work we adopt the Gilmore fixed model [24], which has a slightly lower EBL compared to other models. The EBL mildly attenuates the 130 GeV DM, but has virtually no effect on light DM.

Redshift distribution

Another interesting quantity to see is the redshift distribution of the DM annihilation signal. We quantify this by defining the dimensionless quantity \(ξ(z)\),

43
\[ \xi(z) = \frac{\frac{d}{dz} \int dE_0 I_{\gamma}^{EG}(E_0)}{\frac{d}{dz} \int dE_0 I_{\gamma}^{EG}(E_0) \big|_{z=0}} \]
\[ = \frac{(1 + z)^3 \langle \delta^2(z) \rangle}{\mathcal{H}(z) \langle \delta^2(0) \rangle} \frac{1}{1 + z}. \quad (2.21) \]

Physically, \( \xi(z) \) is the relative DM annihilation signal per redshift interval. To make the discussion independent of particle physics, we integrate out the energy spectrum, which results in the additional factor of \( 1/(1 + z) \) at the end. This factors accounts for the energy binning effect or equivalently the cosmological time dilation, as the energy-time element is redshift invariant (\( dtdE = dt_0dE_0 \)). We also neglect the attenuation factor, \( \tau \), to make the discussion independent of the EBL model and the annihilation products being observed.

We show \( \xi(z) \) in Fig. 2.9 for three cases: the P12 smooth Halo Model and, the P12 model with K10 substructure for LOW and HIGH cases. For a fixed \( M_{\text{min}} \), substructure has minor effect on the distribution. Varying the value of \( M_{\text{min}} \) also changes the shape slightly. In all cases, \( \xi(z) \) is peaked at redshift zero. In terms of implications for detection prospects, not only does low-\( z \) region have a larger flux, the less redshifted energy also means the signal is more detectable. This argument is even stronger if there is a considerable cosmic attenuation effect.

The shape of \( \xi(z) \) determines the gamma-ray profile of DM annihilation signals before detector smearing and cosmic attenuation. It encodes the redshift evolution of DM density distribution. Therefore, the signal profile with energy could be a probe for the cosmic structure evolution. The effect is however secondary to the signal strength, and we encourage future works to explore this possibility.

### 2.2.4 Isotropic Galactic vs. Extragalactic

One can compare the relative importance of the isotropic Galactic component to the extragalactic component (see also [102]),

\[ \frac{I_{\gamma}^{EG}}{I_{\gamma}^{G}} \sim \left[ \frac{\langle \delta^2 \rangle}{8 \times 10^4} \right] \left[ \frac{0.4}{J} \right]. \quad (2.22) \]
The J-factor for the isotropic Galactic component has a robust lower limit, $J \sim 0.4$, for the case of no substructure at anti-GC. In this case, the extragalactic component will be comparable to the isotropic Galactic component, with $\langle \delta^2 \rangle \sim 8 \times 10^4$. This corresponds to $M_{\text{min}} \sim 10^3 M_\odot$ for LOW substructure case, or $\sim 10^5 M_\odot$ for HIGH substructure case.

The isotropic Galactic component naturally breaks the “clumping factor–$\sigma v$” degeneracy. This can be seen by the following schematic equation,

$$
\Phi^{\text{IGRB}} \propto \sigma v \left( \langle \delta^2 \rangle \Phi^\text{EG}_0 + J \Phi^G_0 \right),
$$

where $\Phi^{\text{IGRB}}$ is the total DM annihilation contribution to IGRB and $\Phi^\text{EG}_0$ ($\Phi^G_0$) is the extragalactic (isotropic Galactic) component properly normalized to factor out the dependence of $\sigma v$, $\langle \delta^2 \rangle$, and $J$. For large $\langle \delta^2 \rangle$, or small $M_{\text{min}}$, the isotropic Galactic component is negligible and the $\langle \delta^2 \rangle \sigma v$ degeneracy is apparent. For small $\langle \delta^2 \rangle$, or large $M_{\text{min}}$, the isotropic Galactic component dominates. In particular for large $M_{\text{min}}$, the substructure enhancement to the J-factor is small, thus the degeneracy is naturally broken. In this case, however, the information about $M_{\text{min}}$ is lost unless the isotropic component is subtracted.

### 2.3 DM constraints

In the previous section, we present the DM annihilation contribution to the IGRB from both the isotropic Galactic and extragalactic components. We consistently take into account the substructure enhancement using the K10 model.

In this section, we show that comparing the signals from the GC and the IGRB can break the degeneracy of the small-scale cutoff ($M_{\text{min}}$) with the annihilation cross section, thus testing both cosmology and particle physics scenarios. We illustrate this using two DM candidate scenarios, representing the narrow line and the broad continuum classes. For simplicity, we focus the discussion of the clumping factor using the Halo Model approach with K10 substructure only.

The energy spectrum of the IGRB measured by Fermi is shown in Fig. 2.10. We show the data points and the single power law fit, naively extrapolated, from the published Fermi
Figure 2.10: The IGRB spectrum measured by Fermi. We show the published Fermi IGRB data and the extended single power-law fit from Abdo et al. [28], and also data points from Fermi preliminary results [29]. The EBL-attenuated power law is adapted from [30].
result [28], as well as preliminary result from Fermi [29]. The attenuated power law is adapted from Murase et al. [30], who considered the case that the IGRB is composed by unresolved astrophysical sources with star formation evolution. One can see the preliminary data set shows hints of spectral softening in high energies and is closer to the attenuated power law than the extrapolated power law. This could potentially lessen the Very High Energy Excess problem [145, 30], and it adds support to the hypothesis that the observed spectrum is extragalactic in origin, validating the background reduction procedure by Fermi. The attenuated power law represents one of the simplest astrophysical-only IGRB spectra, normalized to lower energy points. It shows the theoretical limit of using the IGRB to constrain DM signals, if only flux information is used. In this work, we conservatively derive constraints using the extended power law fit, which is well above the curve with attenuation.

2.3.1 Gamma-ray line – 130 GeV DM

Recently, gamma-ray line with energy $\sim 130$ GeV were reported towards the GC in Fermi-LAT data at high statistical significance, and clusters at less significance [97, 98, 146, 147, 148, 149]. So far, no such signals have been seen detected from dwarf galaxies [150]. There are astrophysical explanations for these events ([151], but also see [152]), as well as interesting instrumental effects [153, 154, 155, 156]. Radio measurements of this candidate seems promising [157, 1] as a fairly model-independent check. The line signal, if interpreted as DM, requires the annihilation cross section to be $\sigma v_{\chi\chi \rightarrow \gamma \gamma} = (1 - 2) \times 10^{-27}$ cm$^3$ s$^{-1}$. This value of $\sigma v$ is higher than normally expected [158], but could be a manifestation of DM physics [159]. The morphology of the signal is best fit with the Einasto profile, but is also consistent with the NFW profile [104]. The Fermi collaboration has confirmed the feature, but with lower significance and a small shift in energy to 133 GeV [160], mostly due to a better modeling of the detector response to monochromatic photons. The nature of this feature is currently inconclusive, and actions are advocated to quickly resolve the situation [161]. In this section we assume the feature is due to DM annihilation, and refer to it as the 130 GeV DM.
Figure 2.11: The combined (isotropic Galactic + extragalactic) DM signal for LOW and HIGH substructures for the 130 GeV DM with annihilation channel $\chi\chi \rightarrow \gamma\gamma$. Superposed are the IGRB spectra from Fig. 2.10. The individual isotropic Galactic and extragalactic components for the LOW substructure case are shown in black dotted lines. All DM components are evaluated with $\sigma_v = 2 \times 10^{-27} \text{cm}^3\text{s}^{-1}$ and $M_{\text{min}} = 10^{-6} M_\odot$. The DM signals are convolved with 10% Gaussian smearing to take into account the energy resolution of Fermi-LAT. For visualization, we also show the total spectra (isotropic Galactic + extragalactic + Abdo 2010 fit) for LOW and HIGH cases.
Figure 2.12: The “clumping factor-$\sigma v$” parameter space plane. It is bounded from below by the Minimum DM clustering and from the right by the relic abundance requirement. The allowed region for 130 GeV DM is fixed by the GC observation (vertical green band). The blue solid line is obtained using the total IGRB flux (100% IGRB). The bending feature notes the transition into where the isotropic Galactic component dominates (Eq. 2.23). We show the translation from clumping factor to $M_{\text{min}}$ using the LOW (HIGH) substructure case, with solid green (dashed red) marks on the allowed parameter space. The blue dot-dashed line represents the parameter space that IGRB could probe for a more detailed analysis (10% IGRB).
One can predict the contribution to the IGRB from such a DM particle given the information from GC. We show a representative case in Fig. 2.11. Assuming $M_{\text{min}} = 10^{-6} M_\odot$ and $\sigma v = 2 \times 10^{-27} \text{cm}^3 \text{s}^{-1}$, we show the combined isotropic Galactic and extragalactic DM components for both LOW and HIGH substructure cases. For the extragalactic component, we integrate up to redshift 4 to cover the interesting energy range. Both features are obtained by convolving the intrinsic spectrum using 10% energy resolution with Gaussian smearing.

![Image of graph showing photon spectrum]

Figure 2.13: Same as Fig. 2.11, but with 30 GeV $\chi \chi \to b \bar{b}$ and $\sigma v = 2.2 \times 10^{-26} \text{cm}^3 \text{s}^{-1}$. We consider the prompt photon spectrum only.
We first consider the most conservative constraint of DM annihilation from the IGRB. This can be obtained by requiring the total DM signal to not overshoot the total flux of the IGRB. Recall that for the extragalactic component, the clumping factor is degenerate with the annihilation cross section. In addition, the clumping factor correlates with the Galactic substructure boost through their dependence on $M_{\text{min}}$. Therefore, the general constraint is a surface on the $M_{\text{min}}$-$\sigma v$-$m_\chi$ space. For a specific DM case, like the 130 GeV DM, we can condense the $m_\chi$ dimension. Lastly, the resultant $M_{\text{min}}$-$\sigma v$ plot would depend on the underlying model of the DM density distribution. A more convenient treatment is to construct the clumping factor versus $\sigma v$ plot. In that case, most of the model dependence moves to the interpretations of the parameter space. For a pure extragalactic component, a constant flux would be represented by a straight diagonal line in the clumping factor-$\sigma v$ plane, representing complete degeneracy.

In Fig. 2.12, we show one of the main results of this work. The observed IGRB flux defines a line in the clumping factor-$\sigma v$ plane, as labeled by the 100% IGRB line, above which the DM signal exceeds the IGRB total flux, and thus is robustly excluded. Superposed are two independent constraints. The plane is bounded from below by minimal DM clustering, conservatively defined by $M_{\text{min}} = 10^6 M_\odot$, and bounded from the right by the relic abundance criterion (the precise value of $\langle \sigma v \rangle$ is mass dependent and is $\sim 2.2 \times 10^{-26}$ cm$^3$ s$^{-1}$ [11], for the mass range that we are interested in).

The degeneracy between the clumping factor and $\sigma v$ is apparent in the parameter space where the extragalactic component dominates the isotropic Galactic component. As one increases the value of $M_{\text{min}}$, the decrease of the clumping factor is much faster than the decrease of the boost factor for the isotropic Galactic component. When $\langle \delta^2 \rangle$ falls below $\sim 8 \times 10^4$, the Galactic component begins to dominate the extragalactic component (Eq. 2.23), resulting in near-independence of the flux on the value of $\langle \delta^2 \rangle$, and hence the bending feature. The required value of $\sigma v$ for the 130 GeV DM is labelled by the vertical green band. On the green band, we map the clumping factor to $M_{\text{min}}$ using the LOW substructure model (solid green labels). In this conservative scenario, the constraint is below $10^{-12} M_\odot$, and thus can be considered as unconstrained. But for the HIGH case (dotted red labels), it is
probing near $10^{-12} M_\odot$. For simplicity, we conservatively ignored the larger enhancement for the isotropic Galactic component for the HIGH substructure case, which matters only when the isotropic Galactic component dominates.

![Graph showing clumping factors and DM clustering](image)

Figure 2.14: Same as Fig. 2.12, but with 30 GeV $\chi\chi \rightarrow b\bar{b}$. The green parameter space are taken from the “best-fit spatial model” from [31].

Any additional non-DM component will significantly improve the constraint. This extra component could come from different unresolved astrophysical sources, including star-forming galaxies, blazers, etc (e.g., see [162, 163]).
One can also distinguish DM signals from the data itself. DM annihilation signals usually show a sharp spectral cut-off near the DM mass. Such features, if present in the data, should be easily isolated from any underlying background that behaves like power laws. The distinct anisotropy feature of DM annihilation is also a powerful tool to distinguish the DM signal from non-DM components, even down to $\sim 10\%$ level [164].

Therefore, IGRB DM sensitivity can potentially reach $10\%$ of its total flux using either better background estimation, spectral analysis, or anisotropy analysis. We label this by the $10\%$ IGRB line in Fig. 2.12, above which is the parameter space we think realistic IGRB analysis can probe. One can see that with $10\%$ DM sensitivity, IGRB can probe up to $\sim 1 M_\odot$, the upper extreme for most of the cold DM scenarios.

In addition, we have conservatively taken the isotropic Galactic component to be from the Anti-GC. The IGRB analysis, however only uses photons from $\sim 80\%$ of the sky ($|b| > 10^\circ$) [28]. A realistic estimation of how much the Galactic Halo DM component is contaminating the IGRB probably requires a detailed study by adding a DM template to the IGRB analysis. To estimate that analysis, we consider using the Galactic Pole ($J(\frac{\pi}{2})$), where the sky is least contaminated by the Galactic foreground. In this case, the isotropic Galactic component constraint, represented by the bending features in Fig. 2.12, is improved by a factor of $\sim 2$ (the constraints lines shift to the left by a factor of $2$). All of the $130$ GeV DM parameter space will be probed by the $10\%$ IGRB line in this case.

Last but not least, we used the spectral fit to the Fermi data given by Ref. [28] to derive all the constraints. With more and better data [165], as already partly demonstrated by Ref. [29], one can expect the overall constraint can be improved significantly soon, especially if the high energy spectral softening is confirmed.

We therefore conclude that a DM IGRB analysis in the near future can realistically probe all of the parameter space of the $130$ GeV DM, even in the conservative substructure case. Such an analysis contains slightly different systematics than the GC DM search, since the GC region is excluded from the IGRB measurement.
Figure 2.15: Same as Fig. 2.11, but with 10 GeV $\chi\chi \rightarrow \tau^+\tau^-$ and $\sigma v = 2.2 \times 10^{-26} \text{cm}^3 \text{s}^{-1}$. We consider the prompt spectrum only.
Figure 2.16: Same as Fig. 2.12, but with 10 GeV $\chi\chi \rightarrow \tau^+\tau^-$. The green parameter space are taken from the “best-fit spatial model” from [31].
2.3.2 Continuum – 10-30 GeV light DM

For many DM models, DM annihilating into quarks or leptons is more favored than monochromatic photons, since the latter may be loop-suppressed. Therefore, annihilations typically produce a broad gamma-ray continuum. Much attention has been paid to the low energy spectrum observed by the Fermi-LAT at the GC where unexplained excess photons are observed [166, 167, 168, 99, 31, 169, 170], which may be incompatible with being unresolved astrophysical sources [171] (but also see [172]). To obtain the GC excess, the complicated GC astrophysical emission needs to be subtracted. The resulting excess is therefore subject to large systematic uncertainties.

If interpreted as signals from DM annihilation, these excesses are generally favored by \( \chi\chi \rightarrow b\bar{b} \) or \( \chi\chi \rightarrow \tau^+\tau^- \) at \( \sigma v \sim 10^{-26} \text{ cm}^3 \text{ s}^{-1} \) and mass 10 – 30 GeV. The profiles favored by the excesses are usually more cuspy than NFW (typically \( \gamma \sim 1.3 \)). The cuspy profile has no impact on the calculation of the isotropic Galactic component. It can constitute an extra boost to the clumping factor. We conservatively continue considering the NFW case for the extragalactic component.

Similar to the 130 GeV DM, we test the compatibility of these DM annihilation channels using the IGRB. For definiteness, we use the best fit parameters from Ref. [31], which are \( m_\chi = 30 \text{ GeV} \) for the \( b\bar{b} \) and \( m_\chi = 10 \text{ GeV} \) for \( \tau^+\tau^- \) channels, respectively. We consider a range of \( \sigma v \) with \((0.8 - 2.2) \times 10^{-26} \text{ cm}^3 \text{ s}^{-1} \) for \( b\bar{b} \) and \((0.3 - 2) \times 10^{-26} \text{ cm}^3 \text{ s}^{-1} \) for \( \tau^+\tau^- \) given by the best-fit spatial model (\( \gamma = 1.2 \)) in the above reference.

In Fig. 2.13 and 2.15, we show the spectra of \( \chi\chi \rightarrow b\bar{b} \) and \( \chi\chi \rightarrow \tau^+\tau^- \) together with the IGRB data, in the same format as Fig. 2.11. We adopt \( \sigma v = 2.2 \times 10^{-26} \text{ cm}^3 \text{ s}^{-1} \) and \( M_{\text{min}} = 10^{-6} M_\odot \). The gamma-ray spectra are obtained using Pythia [173].

In Fig. 2.14 and 2.16, we show the corresponding constraints in the clumping factor-\( \sigma v \) plane. The \( \tau^+\tau^- \) channel constraint is slightly better than the \( b\bar{b} \) due to the slightly smaller DM mass and harder spectrum. The conservative constraint, by requiring the DM signals do not overshoot the total IGRB flux, is given by the 100% IGRB line. Using the conservative LOW substructure case, we see that both \( b\bar{b} \) and \( \tau^+\tau^- \) are unconstrained. For
the HIGH substructure case, the 100% IGRB line would carve into parameter space near $10^{-12} M_\odot$, for the $\tau^+\tau^-$ channel.

Similar to the 130 GeV DM case, any extra component in the IGRB or any method in isolating the potential DM signal from background can significantly shrink the allowed parameter space. The spectra of $\bar{b}b$ and $\tau^+\tau^-$ are not as sharp as the monochromatic photon channel, but they do have a cutoff in the spectrum near the DM mass. The light DM annihilation channels also enjoy higher statistics compared to the 130 GeV DM, which would benefit both the spectral and anisotropy analyses. We therefore also consider 10% DM sensitivity as realistic for the light DM, as shown by the 10% IGRB line in Fig. 2.14 and 2.16. In that case, even for the conservative LOW substructure case, the IGRB can probe near $10^{-6} M_\odot$, and even up to $10^0 M_\odot$ for higher $\sigma v$ regions. As a result, the IGRB is also promising in constraining optimistic substructure and small $M_{\text{min}}$ DM scenarios.

For both channels, we only considered the prompt photon emission and ignored secondary processes such as synchrotron, bremsstrahlung and inverse Compton emissions. These components depend on the astrophysical environments such as photon density and magnetic field. Adding these components would improve the constraints on the specific channels. For more thorough treatments of these processes, see, e.g., Ref. [174, 175, 176]. The DM candidates we considered, were obtained from fits without taking into account secondary emissions. As a result, for easier comparison, we neglect these extra components.

2.4 Summary and Outlook

2.4.1 Summary

We study the effect of Dark Matter (DM) microhalos on DM annihilation signals in the Isotropic Gamma-Ray Background (IGRB). We demonstrate that using substructure-dominated systems and multi-source observations together can constrain the minimum halo mass and annihilation cross section separately. We show that using the IGRB leads to interesting sensitivity for testing tentative signals from the Galactic Center (GC).

We consider the case of DM annihilation contributing to the IGRB. Motivated by Prada
et al. ([18], P12), we extend their results using a physically-motivated cosmological variable, $\sigma_L(M, z)$, with the latest Planck cosmology. As a result, we obtain a halo concentration description that fits well to both large-scale observations and small-scale microhalo simulations. Adding the substructure model of Kamionkowski, Koushiappas, and Kuhlen [20], we consistently take into account the effect of DM substructures on the isotropic Galactic and extragalactic signals of DM annihilation. For a given substructure scenario, the IGRB DM contribution then only depends on the minimum halo mass, $M_{\text{min}}$, set during the epoch of kinetic decoupling, and the annihilation cross section, $\sigma v$.

We show that using the IGRB alone, the DM constraint suffers from the “clumping factor-$\sigma v$” degeneracy. We propose a new perspective by constructing the “clumping factor-$\sigma v$” plots, where this problem is explicit for any particular DM case (Figs. 2.12, 2.14, 2.16). The degeneracy can be broken by adding information from an independent measurement, thus yielding information for both $M_{\text{min}}$ and $\sigma v$. This is potentially the only method to observationally constraint $M_{\text{min}}$ for cold DM cosmologies.

We demonstrate this idea by comparing the Fermi-measured IGRB to two tantalizing DM gamma-ray indirect detection candidates from the GC. One is the $\sim 130$ GeV DM in the $\chi\chi \rightarrow \gamma\gamma$ channel. The other is $(10 - 30)$ GeV light DM in the $\chi\chi \rightarrow b\bar{b}$ or $\chi\chi \rightarrow \tau^+\tau^-$ channels. We show that, in the most conservative case, where the substructure fraction is low and DM annihilation is allowed to saturate the IGRB flux, DM analyses using the IGRB are reaching interesting sensitivity for $M_{\text{min}}$.

We further argue that it is unlikely that DM annihilation signals would dominate the IGRB. Taking into account unresolved astrophysical sources can reduce the allowed DM contribution to the IGRB. Utilizing the spectral and anisotropy feature of DM annihilation signals, one could further limit the IGRB DM component. We show that if 10% DM sensitivity can be achieved by a more detailed analysis using the IGRB, one should be able to recover the 130 GeV DM signal, while the more clumpy cases can be probed for light DM. The rapid improvement of the limit on $M_{\text{min}}$ reflects the physical expectation that concentrations increase progressively slower with decreasing scales, as shown by the P12 concentration-mass relation. Last but not least, we use the single power law fit from the
Fermi published result, which only uses 2 years of data. Imminent Fermi updates on the IGRB with better data in terms of background reduction and higher statistics would further improve our result.

2.4.2 Outlook

We only focus on the velocity independent DM annihilation case. One can generalize and probe the velocity dependent DM candidates (e.g., [70, 71]). In that case, in addition to the DM spatial distribution, one can probe the DM velocity distribution as well. The relevant quantity for the extragalactic component would be $\langle \rho_\chi^2 \sigma v \rangle$ [68, 69]. This is analogous to the clumping factor, but also takes into account the velocity distribution.

We demonstrate the benefits of comparing GC and IGRB for DM annihilation signals, but one need not stop there. In principle, one can do global analyses including multiple DM sources, e.g., observations from Dwarf Galaxies or Galaxy Clusters etc. It can further disentangle different dependencies like halo profiles, substructure content, substructure evolution history, etc.

We have reached the era where many astrophysical probes are reaching the relevant parameter space for simple WIMP DM indirect detections. We anticipate that in the future cross correlation of multiple astrophysical observations will become more and more important in DM indirect detection.
Chapter 3

Cosmic Neutrino Cascades from Secret Neutrino Interactions

The first detection of high-energy astrophysical neutrinos by IceCube provides new opportunities for tests of neutrino properties. The long baseline through the Cosmic Neutrino Background (CνB) is particularly useful for directly testing secret neutrino interactions ($\nu$SI) that would cause neutrino-neutrino elastic scattering at a larger rate than the usual weak interactions. We show that IceCube can provide competitive sensitivity to $\nu$SI compared to other astrophysical and cosmological probes, which are complementary to laboratory tests. We study the spectral distortions caused by $\nu$SI with a large s-channel contribution, which can lead to a dip, bump, or cutoff on an initially smooth spectrum. Consequently, $\nu$SI may be an exotic solution for features seen in the IceCube energy spectrum. More conservatively, IceCube neutrino data could be used to set model-independent limits on $\nu$SI. Our phenomenological estimates provide guidance for more detailed calculations, comparisons to data, and model building.

The contents of this chapter were published in [3].

3.1 Introduction

Neutrinos are mysterious. The discovery of neutrino mass and mixing established physics beyond the standard model. With rapid improvements in experimental sensitivity, neutrinos might soon reveal more dramatic new physics. This could include signatures that depend on neutrino mass, e.g., neutrino decay, neutrino magnetic moments, or neutrinoless double beta decay. The weak interactions of neutrinos make them unique messengers for studying astrophysical systems. The extreme scales of astrophysics allow tests of neutrino properties
far beyond what is possible in the laboratory, and may reveal new interactions that shed
light on the origin of neutrino mass and other important questions.

The term “secret neutrino interactions” (νSI) indicates new physics that couples neu-
trinos to neutrinos. A wide variety of models have already been considered, and some have
implications for neutrino masses. A way to characterize these models is by their mediator
mass. For massless mediators, such as in Majoron models [177, 178, 179, 180, 181], there is
at least one stable new particle. For very heavy mediators, one can use an effective theory
to study the phenomenology of a class of models [32, 182, 183, 37]. In between, the mediator
mass is more moderate, and could induce resonances [184, 185, 186, 187, 188, 189, 190]. In
some models, the neutrinos also interact with dark matter [191, 192, 193, 194, 195, 196,
119, 197, 198, 199, 200].

It is challenging to directly test νSI through neutrino-neutrino scattering. Sufficiently
high flux or volume densities of neutrinos for any interactions to occur only exist in astro-
physical systems. Even there, the difficulty is revealing (using neutrinos!) the signatures of
such interactions. So far, only νSI interactions much stronger (in a sense explained below)
than weak interactions have been constrained. Given the difficulty of probing νSI in the
laboratory, it is therefore interesting to consider more model-independent probes, such as
those from astrophysics and cosmology.

One direct probe of νSI uses astrophysical neutrinos as a beam and the Cosmic Neutrino
Background (CνB) as a target. Kolb and Turner (hereafter KT87) [32] utilized the detection
of astrophysical neutrinos from SN 1987A. The agreement of the detected signal with the
standard expectation of no neutrino scattering en route yields robust constraints on νSI.
KT87 established a phenomenological approach by considering general interactions with
mediator masses either much smaller or much larger than the interaction energy. Their
constraints could be applied to many possible νSI models.

The first detection of astrophysical neutrinos since SN 1987A is the 37 events detected
by IceCube [201, 202, 203]. One expects a steady stream of more events in the near future,
so the precision will improve quickly. The angular distribution of the events suggests that
most, if not all, of them are extragalactic in origin. Compared to the SN 1987A neutrinos,
the IceCube neutrinos have a much longer baseline through the CνB, making them more sensitive to νSI; they have much higher energies, making them more powerful probes for massive mediators; and they have a diffuse (many-source) origin, thus averaging out the uncertainties for individual sources.

We take advantage of this new opportunity and explore the sensitivity of IceCube to νSI. With minimal assumptions about the interaction to reduce model dependence, we show that there are regions of parameter space where νSI could cause significant distortions to the neutrino spectrum. Because the flux is diffuse and the shape and normalization without interactions are not known, we must look for distortions to the spectrum that have characteristic shapes. This favors interactions with strong energy dependence, especially due to a resonance.

Our method generalizes earlier work, going beyond the pure attenuation considered in KT87 and Refs. [204, 205] as well as the simplified treatment of regeneration considered in Refs. [187, 188]. We improve upon these by using the propagation equation to describe the interaction of a neutrino beam with the CνB in the presence of strong νSI. Besides attenuation, this properly takes into account regeneration as well as multiple scattering of the parent and daughter neutrinos, i.e., a cascade.

In Sec. 3.2, we consider existing νSI constraints. In Sec. 3.3, we discuss the effects of νSI on astrophysical neutrino spectra. We conclude in Sec. 6.4. Throughout, we use cosmological parameters for which the matter density fraction is Ω_M = 0.3, the cosmological constant density fraction is Ω_Λ = 0.7, and the Hubble function is \( H(z) = H_0 \sqrt{\Omega_\Lambda + \Omega_M (1 + z)^3} \), where \( H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1} \).

### 3.2 Secret Neutrino Interactions

We first review existing constraints on νSI for a phenomenological scalar interaction term, \( L_{\text{int}} \sim g\phi\bar{\nu} \nu \). Fig. 4.11 shows the parameter space of neutrino coupling \( g \) to a new mediator \( \phi \) with mass \( M \). While these parameters vary depending on the specific model, type of coupling, number of new mediators, etc., this figure gives a broad comparison of different
constraints. The accuracy is adequate, especially considering the many orders of magnitude
on the axes. We do not consider the vector mediator case, since the laboratory constraints
are much stronger than for the scalar case [120, 206].

The range in mediator mass is chosen to span from the KT87 constraint to those near
the weak scale, focusing on the mass range that has been of particular interest in recent
work, e.g., Refs. [119, 198, 199]. Coincidentally, this is where the IceCube sensitivity is
greatest, as we show below. The range in coupling is chosen from the boundary of the non-
perturbative regime to where IceCube loses sensitivity. It would be possible to extend the
figure to smaller masses and couplings, showing interesting features in some of the limits,
but that would detract from our focus, so we just describe those features in the text.

There are three kinematic regimes in which constraints can be derived, depending on
how the mediator mass, \(M\), compares to the interaction energy in the center of momentum
frame, \(\sqrt{s}\). These are where the mediator mass is small (i.e., like a Majoron, where the
mediator mass is zero or negligible), comparable to the interaction energy (i.e., where the
energy dependence of the cross section depends on the mediator mass, possibly through a
resonance), or large (i.e., an effective theory where the mediator mass has been integrated
out). Constraints derived assuming extremely small or large mediator masses cannot be
applied beyond their domains of applicability.

An effective theory description is appropriate when \(M \gtrsim \sqrt{s}\) and \(g \lesssim 1\). Then these
parameters can be characterized in a combination analogous to the Fermi constant for
low-energy weak interactions, i.e., a dimensionful coupling

\[
G \equiv \frac{1}{\sqrt{4\pi}} \frac{g^2}{M^2}.
\]  

(3.1)

Constraints on \(\nu\)SI are sometimes quoted using just \(G\). This does not provide as much
information as a region in the \(g-M\) plane, because of the degeneracy in \(g/M\) and the
unspecified limits of applicability. In Fig. 4.11, we plot some diagonal contours of constant
\(G\). The fact that the line given by \(G \sim G_F \sim 10^{-5} \text{ GeV}^{-2}\) is not the same as the single
point \((M \sim 100 \text{ GeV}, g \sim 1)\) that defines the weak interactions illustrates our caution about
characterizing \(\nu\)SI with only \(G\).
One general framework for directly testing $\nu$SI is to use neutrinos from sources that travel a long distance through the $C\nu B$. The only possibly relevant standard model process is the Z-burst scenario [207, 208], where a high energy neutrino interacts with the $C\nu B$ through a Z-boson resonance. However, the required neutrino energy is extremely high, $\sim 10^{14}$ GeV, and neutrinos of such energy may not exist; the cross section at lower energies is much smaller. Any significant neutrino self-interaction observed at lower energies must be due to $\nu$SI.

In terms of specific limits, neutrinos detected from SN 1987A were the first and, until recently, only direct detection of neutrinos from astrophysical sources beyond the Sun. Requiring that these neutrinos travel through the $C\nu B$ without scattering leads to a robust upper limit on the cross section. Had the interaction strength been larger, the neutrinos would have scattered to lower energy and fallen below the detector sensitivity [32]. The limit from KT87 corresponds to $G \lesssim 10^8$ GeV$^{-2}$. The average supernova neutrino energy is $\sim 10$ MeV and that of the $C\nu B$ is $\sim 10^{-1}$ eV (assuming small but degenerate neutrino masses), making $\sqrt{s} \sim 10^{-3}$ MeV, so the applicability of this limit would end below a vertical boundary at $M \sim 10^{-3}$ MeV (not shown). For a massless neutrino, this boundary would be at $M \sim 10^{-4}$ MeV, because the average neutrino energy is $\sim 10^{-3}$ eV.

Another general framework for directly testing $\nu$SI is through their effects on a gas with a high neutrino density. In the early universe [209, 210, 211, 212, 213] and in core-collapse supernovae [214, 215, 216, 217, 218, 219, 220, 221, 222, 223], the conditions are so extreme that even standard model scale neutrino self-scattering and their self-induced matter mixing potential are important. In the early universe, $\nu$SI could cause neutrinos to annihilate or decay into light particles, modifying the expansion history [224, 225, 226, 227, 33]; change the free-streaming property of neutrinos during photon decoupling [184, 228, 229, 34, 35]; or induce new mixing effects [230, 198]. In supernovae, the effect of elastic scattering on neutrino escape time [231] is irrelevant [232], but $\nu$SI could cause a phase transition [233, 234], change the cooling process [235, 236, 237, 192, 238], or induce non-standard flavor mixing [239].

There are several specific limits. In the early universe, if the $\nu$SI mediator is not too
Figure 3.1: Present constraints and future sensitivity to $\nu$SI in terms of neutrino coupling, $g$, and mediator mass, $M$, with diagonal dotted contours shown for values of the dimensionful coupling $G$. The blue shaded regions are excluded by astrophysical and cosmological considerations based on SN 1987A [32], BBN [33], and the CMB [34, 35]. The pink dashed lines indicate flavor-dependent limits based on laboratory measurements of meson and lepton decays [36]; we consider only the weakest limit, for $\nu_\tau$, to be robustly excluded for all flavors, and it is shaded. The red shaded region is excluded based on measurement of Z-boson decay [37]. The gray shaded region indicates the non-perturbative regime. The orange lines are contours of unit optical depth for different initial neutrino energies (Eq. 3.10), indicating the approximate boundary of the parameter space above which IceCube is sensitive to $\nu$SI. The squares represent the example parameters (given in Table 4.1) used in our calculations.
massive, it could be in thermal equilibrium, changing the number of relativistic degrees of freedom \([224, 33]\), which can be tested through Big Bang Nucleosynthesis (BBN). We show the “maximally conservative” case from Ref. \([33]\), which assumed vector boson mediators. The BBN limits extend down to \(g \sim 10^{-8}\). The presence of \(\nu\)SI can also change the free-streaming property of the \(\nu\)B, which can leave an imprint on the observed Cosmic Microwave Background (CMB). Strong constraints on \(\nu\)SI have been set in Refs. \([34, 35]\).

In the mediator mass range we focus on, the constraint was derived assuming a heavy mediator, and is \(G \lesssim 100 \text{ GeV}^{-2}\), which is much more stringent than the SN 1987A limit. In Ref. \([34, 35]\), the \(\nu\)B is constrained to be free-streaming until redshift \(\sim 2 \times 10^5\), where \(\sqrt{s} \sim 10^{-4} \text{ MeV}\). Therefore, the domain of applicability of the CMB limit would end at a vertical boundary (not shown) at \(M \sim 10^{-4} \text{ MeV}\).

Limits on \(\nu\)SI can also be set by observations of laboratory processes. Even if the neutrinos are not detected directly, their presence can be inferred by precise measurements of other particles. For example, in the presence of \(\nu\)SI, a mediator can be produced by bremsstrahlung from an external neutrino \([240, 36, 120]\); a massive mediator will then decay into other particles \([182, 183, 37]\). In Majoron-type models, the best laboratory constraints come from meson and lepton decays, e.g., Refs. \([240, 36]\), but they depend on the particular flavor coupling, \(g_{\alpha\beta}\), where \(\alpha, \beta = e, \mu, \tau\), and are valid only up to the mass of the meson or lepton, e.g., kaons or tau leptons. The couplings involving \(\nu_\tau\) are the least constrained, so we regard them as the most robust. Accordingly, they are shaded in Fig. 4.11 to indicate exclusion for all flavors.

An flavor-independent constraint can be obtained from Z-boson decay. If a heavy mediator is assumed, the limit is quite strong, \(G \sim G_F\), as shown in Refs. \([182, 183, 37]\), though the domain of applicability of that effective theory calculation ends below the Z-boson mass. We emphasize that though this is nominally a very strong limit, it does not rule out all of the parameter space above the diagonal line \(G \sim G_F \sim 10^{-5} \text{ GeV}^{-2}\). If the calculation is extended to allow a light mediator, following Ref. \([120]\), the result for the scalar mediator case (not shown) is comparable to the \(g_{\alpha\tau}\) constraint in Fig. 4.11 for mediator masses below the mass of Z-boson.
This combination of constraints shows a window of parameter space in the MeV range where model-independent astrophysical or cosmological constraints could be improved. The IceCube sensitivity, shown approximately by the three orange lines in Fig. 4.11 and calculated in the next section, lies in this region. Because both the astrophysical neutrino beam and the CνB targets are expected to contain all flavors of neutrinos, the IceCube sensitivity is complementary to the flavor-dependent laboratory limits.

### 3.3 Astrophysical Neutrino Interactions with the CνB

#### 3.3.1 Sensitivity estimate for νSI

We first make an order-of-magnitude estimate of the sensitivity of IceCube to νSI with cross section \(\sigma_{\nu\nu}\); we present a more detailed calculation in later sections. If the neutrino mass scale is \(\simeq 0.1\) eV, target CνB neutrinos have degenerate masses and are nonrelativistic today. Because of neutrino mixing, all flavors of neutrinos and antineutrinos should be present with comparable fractions in both the beam and target. To be conservative, we assume that only one species (flavor or mass eigenstate) of neutrinos and antineutrinos in the beam interacts, each with only half of the targets of a given species, so the target density is \(n_t \simeq 56\) cm\(^{-3}\) [241, 242]. A typical distance for astrophysical sources is the Hubble length, \(c/H_0 \simeq 4\) Gpc. The optical depth for νSI interactions is \(\tau \simeq n_t \sigma_{\nu\nu} c/H_0\). For νSI to affect neutrino propagation appreciably, one would require \(\tau \simeq 1\), and therefore

\[
\sigma_{\nu\nu} \simeq 1 \times 10^{-30}\text{ cm}^2.
\]  

(3.2)

This is an necessary, but not sufficient, condition for the effects of νSI on astrophysical neutrinos to be observed. The actual νSI sensitivity for an neutrino telescope depends on model details such as the resonance energy and detector details such as the analysis energy range. A larger \(\sigma_{\nu\nu}\) would severely affect the incoming neutrino beam, because the unattenuated fraction scales as \(e^{-\tau}\).

If the interaction is through a heavy mediator, then \(\sigma_{\nu\nu} \simeq G^2 s \simeq G^2 2 E m_{\nu}\). For the PeV neutrinos detected by IceCube, this leads to a scale
This nominal sensitivity on $G$ is a factor of $\sim 25$ times below than the CMB limit \cite{35}. Thus, the detection of astrophysical neutrinos gives an exciting opportunity to test $\nu$SI.

### 3.3.2 Neutrino propagation: free streaming case

In the usual case, neutrinos from cosmic sources travel to the detector without interaction. The standard model neutrino self-interaction cross section grows linearly with neutrino energy when the center of mass energy is below the mass of the $Z$ boson. It is $\sim 10^{-41}\text{cm}^2$ \cite{243, 244, 245, 246} for a $1\text{PeV}$ neutrino scattering on C$\nu$B ($0.1\text{eV}$ masses), much smaller than the sensitivity estimated above. Neutrino-CMB scattering is even more suppressed \cite{247}. Near $1\text{PeV}$, the neutrino-nucleon cross section is $\sim 10^{-33}\text{cm}^2$ \cite{248, 249}, but the nucleon density is only $\sim 10^{-7}\text{cm}^{-3}$, due to the small baryon asymmetry, so the interaction probability is also negligible. (For electrons as targets, both the cross section and number density are small.) Therefore, any interactions of PeV neutrinos during propagation must be due to interactions beyond the standard model.

To describe the neutrino beam, we define the co-moving number density at a time $t$ by $n(t)$ and the differential (in energy $E$) number density by $\tilde{n}(t,E) \equiv dn(t,E)/dE$. The observable neutrino number flux is

$$J(E) \equiv \frac{dN_\nu}{dAdtd\Omega dE} = \frac{c}{4\pi} \tilde{n}(0,E).$$

The evolution of the number density is described by the propagation equation \cite{250, 244, 251, 252, 253, 145}. In the free-streaming limit, that is

$$\frac{\partial \tilde{n}(t,E)}{\partial t} = \frac{\partial}{\partial E}(b\tilde{n}(t,E)) + \mathcal{L}(t,E).$$

The first term on the right takes into account the continuous energy loss rate $b = H(t)E$ due to redshift, and the second is the differential number luminosity density of the sources. Throughout this work, we solve the propagation equation numerically in redshift variables,
conservatively taking the initial condition $\tilde{n}(z_{\text{max}} = 4, E) = 0$.

In the free-streaming case, $J(E)$ has a convenient closed form in the redshift variable [252, 251], given by

$$J(E) = \int_0^{z_{\text{max}}} \frac{cdz}{4\pi H(z)} \mathcal{L}(z, E(1 + z)),$$

(3.6)

following from the simple relationship between emitted and observed energy as a function of redshift. For the source term, we assume a universal emission spectrum with a factorized form,

$$\mathcal{L}(z, E) = W(z) L_0(E),$$

(3.7)

where $L_0(E)$ is the differential number luminosity for each source and $W(z)$ the redshift evolution of the source density, assumed to follow the star formation rate (SFR) [254, 255]. The term $H(z)$ increasingly suppresses the importance of flux contributions from higher redshifts. The SFR evolution, which rises by an order of magnitude between $z = 0$ and $z = 1$ and then begins to fall, overcomes this effect until $z = 1$, so $z = 1$ is the typical redshift of the most relevant cosmic sources.

### 3.3.3 Neutrino propagation: $\nu$SI case

When neutrinos interact with the CnB, the effects can be calculated by adding terms to the propagation equation, so that

$$\frac{\partial \tilde{n}(t, E)}{\partial t} = \frac{\partial}{\partial E} (b\tilde{n}(t, E)) + \mathcal{L}(t, E)$$

$$- c n_t \sigma \tilde{n}(t, E) + c n_t \int_{E}^{\infty} dE' \tilde{n}(t, E') \sum_i \frac{d\sigma_i}{dE}. $$

(3.8)

We assume that the neutrino (or other particle) targets are non-relativistic; if they are not, their energy distribution needs to be taken into account in the propagation equation. The third term on the right accounts for attenuation at a given energy due to scattering with cross section $\sigma(E)$. The fourth term accounts for particle regeneration from one energy to another, including when an incoming particle of energy $E'$ is down-scattered to a lower energy $E$ but not lost and when a target particle with their rest mass energy is up-scattered
to energy $E$ to join the beam. The distributions of secondary particles are described by the differential cross sections $d\sigma_i(E, E')/dE$, where $i$ denotes each process. Here, we only include down-scattering and up-scattering with neutrino targets; this could be generalized. The net effect is therefore a distortion of the beam spectrum in a way that conserves energy but not particle number.

This propagation equation automatically takes into account the re-scattering of secondary particles, analogous to electromagnetic cascades for high energy cosmic gamma rays [250, 145], $\nu_\tau$ regeneration in matter [256, 257, 258], and ultrahigh energy cosmic ray propagation [244, 251, 252, 253]. As far as we know, cascade calculations have not been done for neutrino-neutrino interactions. (A similar formalism for supernova neutrinos interacting with dark matter appeared in preprint while we were finalizing this work [200].)

We assume that there are only active neutrinos in the beam and target, that all species are comparably populated by mixing, and that this happens long before any effects due to propagation. We assume the neutrino masses are all $\simeq 0.1$ eV and that only one species of $\nu + \bar{\nu}$ interacts, each with half of the targets of a given species, so $n_t(z) = 56(1 + z)^3 \text{ cm}^{-3}$

We ignore the possibility of flavor changes in scattering. We take a generic form for the total and differential cross sections to minimize model dependence. The propagation equation and our calculations could be generalized to account for changes in the assumptions, and we discuss below what happens when some of them are relaxed.

We focus on elastic scattering of neutrinos in the $s$ channel. For generality, we assume that the cross section takes a Breit-Wigner form,

$$
\sigma(E) = \frac{g^4}{4\pi} \frac{s}{(s - M^2)^2 + M^2 \Gamma^2},
$$

where $s = 2E m_\nu$ and the decay width of the mediator is $\Gamma = g^2 M/4\pi$. With this form, we generalize the phenomenological approach of KT87 to include the possibility that a resonance dominates the cross section. In that case, the $t$ channel contribution can be neglected. In the off-resonance case, neglecting the $t$ channel does not change the results much for the cases considered here. For the differential cross section, we take a flat distribution in $E$ between zero and $E'$, which corresponds to the case of a scalar mediator.
Vector mediators have a different distribution, but we do not consider this case due to strong constraints [120, 206].

This form of the cross section parametrizes all three kinematic regimes of how the mediator mass compares to the interaction energy in the center of momentum frame. For a very light mediator, the cross section is decreasing with neutrino energy, \( \sigma \simeq g^4/(4\pi s) \), while for a very heavy mediator, the cross section is increasing with neutrino energy, \( \sigma \simeq g^4 s/(4\pi M^4) \). These two limits correspond to the massless and massive mediator cases considered in KT87. For the former, the cross section is independent of \( M \); for the latter, it is degenerate in \( g/M \). Between these two limits, the cross section is peaked at the resonance energy defined by \( s = 2E_{\text{res}}m_\nu = M^2 \), where the cross section is regulated by its decay width and is \( \sigma = 4\pi/M^2 \).

Figure 4.11 shows all three of these behaviors in the optical depth for neutrino scattering. We define this as

\[
\tau(E|g, M) = c \int_0^1 dz \frac{n_t(0)}{H(z)} (1 + z)^2 \sigma(E|g, M),
\]

where \( z = 1 \) is a typical redshift for cosmic sources. For simplicity, we ignore the redshift dependence of \( \sigma(E) \) inside the integral, which would slightly broaden the range of \( M \) for which a resonance could occur. The factor \((1 + z)^2\) comes from the target density evolution factor \((1 + z)^3\) and a factor of \((1 + z)^{-1}\) from \(|dt/dz|\). Taking redshift into account only increases \( \tau \) by about a factor of 3; it would matter more if high-redshift sources were dominant. We show contours of \( \tau = 1 \) for \( E = 0.01, 0.1 \) and 1 PeV in Fig. 4.11. Above the contours (\( \tau > 1 \)), the effect of scattering increases exponentially with \( \tau \), which increases as \( g^4 \) for fixed \( M \). Near the sharp dips, the realistic sensitivity would be reduced by the effects of detector energy resolution. We emphasize that we use \( \tau \) just for illustration; for our main result, we calculate spectra using Eq. (3.8).

### 3.3.4 Line emission with \( \nuSI \)

Before considering astrophysical scenarios with broad energy spectra, it is instructive to show the effects of \( \nuSI \) on a mono-energetic neutrino spectrum. For rough consistency with
the IceCube data, we choose 1 PeV for the line energy and define the flux in the same units as the deduced IceCube spectrum.

In Fig. 3.2, we compare cases with free streaming, $\nu$SI with attenuation only, and $\nu$SI with attenuation and regeneration. The energy dependence of the $\nu$SI effects depends on the neutrino-neutrino cross section. Using the general form of Eq. (3.9), we choose example resonance energies $E_{\text{res}}$ well above, near, and well below the emission energy of 1 PeV. For each $E_{\text{res}}$, the couplings are tuned so that $\sim e^{-1}$ of the energy spectrum is unattenuated after propagation.

For the free-streaming case, as in Eq. (3.6), the spectrum of neutrino energies simply reflects the distribution of source redshifts through the relation $E = 1 \text{ PeV}/(1 + z)$. The edge at 1 PeV is from emission nearby, the peak near 0.5 PeV from emission near $z = 1$, and lower energies from emission at still-higher redshifts. As noted, the rapid rise of the SFR between $z = 0$ and $z = 1$ [254, 255] overcomes the suppression due to the Hubble expansion, i.e., less volume per redshift interval, until near $z = 1$, where the SFR begins to flatten and then fall.

With $\nu$SI, interpreting the spectrum shape is more complicated. Attenuation is energy-dependent, and regeneration moves particles to lower energies while increasing their numbers. We obtain the resultant spectra by numerically solving the propagation equation, Eq. (3.8). We show the effects of attenuation alone as a comparison for the full calculation that includes regeneration. For regeneration, the total energy carried in neutrinos is conserved. We checked our numerical results by comparing the total energy in the spectrum to that of the free-streaming case, finding agreement at the percent level. Energy conservation corresponds to area conservation in a plot of $E^2dN/dE$ with log energy bins. In Fig. 3.2, the area conservation is apparent, with the exception of the bottom panel, where we cut off the figure before showing the whole regenerated spectrum.

For the top panel, where $E_{\text{res}} = 5$ PeV, the cross section increases with energy, which produces a flattish spectrum of regenerated neutrinos. For the middle panel, where $E_{\text{res}} = 0.5$ PeV, the resonance energy is at the peak of the unattenuated spectrum, which causes significant absorption there and a pileup of regenerated events at slightly lower energy.
Figure 3.2: The effects of $\nu$SI on an emitted line spectrum at 1 PeV with SFR evolution, for different assumed resonance energies, as labeled. The solid lines are for free-streaming neutrinos, the dotted lines are for $\nu$SI with regeneration of two neutrinos (the beam and target), and the dashed lines are for $\nu$SI with no regeneration.
Figure 3.3: The effects of $\nu$SI on an emitted continuum spectrum that is consistent with IceCube data. The solid line indicates the free-streaming case ($\gamma = 2$ and $E_{\text{cut}} = 10^7 \text{ GeV}$), while the other lines are for four models of $\nu$SI, with the parameters defined in Table 4.1.
Importantly, the absorption dip is broadened by the redshift effects. Neutrinos emitted with the same energy at different redshifts reach the resonance energy at different redshifts, smearing out the resonance in received energy (this is analogous to the redshift smearing of a monoenergetic emission line). This broadening is helpful for detection, as a narrow feature would be difficult to observe with realistic detector energy resolution. For the lower panel, where \( E_{\text{res}} = 0.05 \) PeV, the cross section is decreasing at energy higher than the resonance, so regenerated neutrinos will continue to interact until they fall below the resonance energy, where the energy dependence changes, forming a true cascade.

### 3.3.5 Astrophysical scenarios with \( \nu \text{SI} \)

We now consider more realistic astrophysical scenarios that are compatible with IceCube measurements [201, 202, 203, 259, 260, 261]. We assume that a generic astrophysical flux can be described by \( \mathcal{L}_0(E) \propto E^{-\gamma}e^{-E/E_{\text{cut}}} \). An \( E^{-2} \) power-law is a typical astrophysical neutrino spectrum. IceCube detected no events above about 2 PeV. A cutoff in the spectrum is required to explain this, because the effective area is rising, especially due to the Glashow resonance. We first choose \( E_{\text{cut}} = 10^7 \) GeV. We normalize the spectrum to about the level seen by IceCube, \( E^2J(E) \sim 10^{-8} \) GeV cm\(^{-2}\) s\(^{-1}\) sr\(^{-1}\) for neutrinos plus antineutrinos for each flavor. Not only is this consistent with IceCube data, it is also predicted from many astrophysical scenarios [262].

In Table 4.1, we provide details for the four example points identified in Fig. 4.11. These points are chosen to represent the regions of parameter space where \( \nu \text{SI} \) can have appreciable effects on the IceCube data while being consistent with the most robust limits.

In Fig. 3.3, we show the results for a continuum injection spectrum, for the free-streaming

<table>
<thead>
<tr>
<th>( g )</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>( M ) [MeV]</td>
<td>100</td>
<td>10</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>( \tau ) (1 PeV)</td>
<td>( \sim 0.7 )</td>
<td>( \sim 0.6 )</td>
<td>( \sim 0.2 )</td>
<td>( \sim 0.002 )</td>
</tr>
<tr>
<td>( E_{\text{res}} ) [PeV]</td>
<td>50</td>
<td>0.5</td>
<td>0.045</td>
<td>0.005</td>
</tr>
</tbody>
</table>
Figure 3.4: Same as Fig. 3.3, but with different values of $E_{\text{cut}}$ for the emitted spectrum. The solid lines are the free streaming case ($\gamma = 2$ and values of $E_{\text{cut}}$ as labeled), and the dashed lines are for Model A $\nu$SI.
case and the four benchmark $\nu$SI cases including regeneration. For Models B, C, and D, the presence of a resonance causes a dip and a pileup of the cascaded neutrinos right below the dip. For Model A, the spectrum cutoff is steepened and a small pileup is produced. These features are potentially large enough to be observed, and may even explain the gap seen in the IceCube spectrum at moderate energies.

The lowest energy events currently observed by IceCube are \( \simeq 0.03 \) PeV, which means small mediator masses like those for Model D are difficult to observe through an obvious dip. Of course, even if the resonance energy is below the detector energy threshold, its effects can be observed if the coupling is large enough. At the other extreme, the largest mediator masses that can be observed through an obvious dip depend on the highest observed neutrino energies. Similarly, even larger mediator masses can be probed if the coupling is large enough.

In Fig. 3.4, we show the effects of extending the spectrum to higher energies with and without $\nu$SI. Without $\nu$SI, these spectra extend to energies well above what IceCube has observed, and therefore are unrealistic. However, it is possible that the emitted spectrum does extend to high energies, but $\nu$SI lead to the observed cutoff near 1 PeV. We use Model A as an example, which has a high mediator mass and a resonance energy of 50 PeV. An accompanying cascade bump occurs at \( \sim 1 \) PeV, with the height of the bump reflecting the energy carried by higher-energy neutrinos that down-scattered.

It might be possible that IceCube is not seeing a dip below PeV energies, but rather a bump near 1 PeV. This would require a higher $\nu$SI energy cutoff and therefore a larger mediator mass than that of Model A. This scenario, however, is in tension with other constraints because a larger coupling is needed for a larger mediator mass to maintain the same interaction strength. Nonetheless, this could be an interesting scenario [263], and we discuss the impact on the measured event spectrum more in the next subsection.

In Fig. 3.5, we show the effects of changing the spectral index. We use Model B as an example, which has its resonance energy slightly below the spectrum cutoff. We normalize the spectra to be the same at 2 PeV. The cascade bump makes the spectrum harder below the absorption dip. The $\gamma = 1.4$ case roughly mimics an astrophysical spectrum
with a $p\gamma$ origin (e.g., [264, 265]), as opposed to a flatter power-law spectrum with a $pp$
origin (e.g., [266]). For this case, the spectrum with $\nu$SI can have twin bumps, separated
by an absorption dip.

![Graph](image-url)

Figure 3.5: Same as Fig. 3.3, but with different values of the spectral index. The solid lines
are for the free-streaming cases (values of $\gamma$ as labeled and $E_{\text{cut}} = 10^7\text{ GeV}$), and the dashed
lines are for Model B $\nu$SI.
3.3.6 Detection prospects

Here we assess the prospects for detecting distorted neutrino spectra in IceCube. We focus on cascade events, because they have a large signal to background ratio and because they reflect the underlying neutrino energy spectrum better than track events [267]. Cascade events with energy $\lesssim 1$ PeV are caused by $\nu_e$ and $\nu_\tau$ charged-current reactions, plus a small contribution from all-flavor neutral-current reactions, while $\nu_\mu$ charged-current reactions cause only track events. We follow [259] and compute the cascade energies deposited in the detector, taking into account the different mean cascade energy for different interaction modes.

We comment on one possible flavor scenario using the cases depicted in Fig. 3.3 as an example. We assume the initial flavor ratio to be $\nu_e:\nu_\mu:\nu_\tau = 1:2:0$ for both neutrino and anti-neutrinos, and we take the only non-zero $\nu$SI coupling to be $g_{\tau\tau}$, in light of strong constraints on other flavors. The $\nu$SI mean free path is much longer than the neutrino oscillation lengths, so it is safe to assume that neutrinos propagate as an incoherent mixture of mass states with ratio $\nu_1:\nu_2:\nu_3 \simeq 1:1:1$ before they interact with the $\nu_\tau$ content of the C$\nu$B. For an imagined case where every mass state was $1/3 \nu_\tau$, then each state would be depleted equally, though with $1/3$ of the interaction strength compared to cases shown above, where we considered just the one flavor scenario.

Realistically, because the $\nu_\tau$ fraction is non-negligible but different in each of the three mass states, each interacts with the C$\nu$B with the modified cross section $|U_{i\tau}|^2\sigma_{\nu\nu}$, where $U_{i\alpha}$ is the standard neutrino mixing matrix. The two up- and down-scattered $\nu_\tau$ states rapidly mix to the mass state ratio $\sim 0.3:1.2:1.5$, which will appear in the detector with flavor ratios $\sim 0.6:1.1:1.3$. The reduction of $\nu_e$ cascade events is mostly compensated by the increase in $\nu_\tau$ cascade events. We neglect the accumulated flavor effects for multiple regenerations, and assume the final flavor ratio remains $\sim 1:1:1$. Finally, in order for each mass state to have the spectrum shown in Fig. 3.3, we would need to compensate the factor $|U_{i\tau}|^2$ by increasing $\sigma_{\nu\nu}$. The smallest element is $|U_{1\tau}|^2 \sim 0.1$, therefore at most it suffices to increase $g$ by $\sim 1.8$ for all flavor spectra to have at least the same degree of spectral
Figure 3.6: The predicted spectra of cascade events (histogram), taking into account realistic detector effects, compared to the 988 days of existing IceCube data (points). Selected curves from Fig. 3.3 are shown, along with an $E^{-2}$ power law with no intrinsic cutoff.
distortion as in Fig. 3.3. Considering the constraints on our benchmark models, this would be viable for all but Model A.

In Fig. 3.6, we show the binned cascade event spectrum detected by IceCube. We consider only events with deposited energy above $10^5$ GeV, because below that energy the background events shown in the IceCube paper are comparable to the signal. For com-

Figure 3.7: The same as Fig. 3.6, except that the calculations are for 988 days of an ideal efficiency 1-Gton detector, the energy bins are half as wide, and the range extends to lower energies.
parison, we show the expected number of events from an unbroken $E^{-2}$ power law, the same with $E_{\text{cut}} = 10^7$ GeV, and the $\nu$SI spectra for Models B and C from Fig. 3.3. We assume $E^2 J(E) = 10^{-8}$ GeV cm$^{-2}$ s$^{-1}$ sr$^{-1}$ below the cutoff, which is consistent with the IceCube data [203]. We take the live time to be 988 days, and use the effective detector mass from [202], which takes into account how the sensitivity drops at low energy, and flattens at $\sim 0.4$ Gton at high energy. We take the neutrino cross section from [249], and assume 10% detector energy resolution. We do not show results for Models A and D because they do not show appreciable differences in this energy range.

Given the low statistics, all spectra in Fig. 3.6 describe the data points reasonably well. However, with more exposure, IceCube should be able to distinguish these $\nu$SI cases from a power-law spectrum. Similarly, the flux spectra shown in Fig. 3.4 and Fig. 3.5 might also be distinguishable in the future. Other than associating the gap with a resonance dip, it is also interesting to ask if the observed events near $\sim 1$ PeV energies can be caused by a bump-like feature as shown in Fig. 3.4. This hypothesis can be tested by fitting a precise cascade spectrum with $\nu$SI parameters such as cross section and mediator mass. It is interesting to note that the required parameter are not too far away from laboratory constraints [263].

We anticipate future IceCube analyses can extend to lower energies with better background rejection. In Fig. 3.7, we show the cascade event spectrum for a 1-Gton detector with the 988 days and the same detector energy resolution as above, but with 10 bins per decade to match the 10% energy resolution. We consider perfect detector efficiency. We see that Models B and C cause distinct spectral features that are detectable. However, it is clearly challenging to distinguish Models A and D.

The flavor phenomenology described above varies mildly with the uncertainty in mixing parameters. In principle, it can be self-consistently incorporated into three-flavor Boltzmann equations by generalizing Eq. (3.8). Another class of scenarios we would like to mention is for the coupling to be in the mass basis instead of the flavor basis. The spectral evolution in that case would also depend on the mass difference as well as mass hierarchy. In this work we only wish to highlight the phenomenology of spectral distortion by $\nu$SI. We defer the comprehensive analysis for studies of concrete models, which may also find interesting
flavor effects.

3.3.7 Further discussion

The results above demonstrate that $\nu$SI could appreciably affect the spectrum detected by IceCube, perhaps in ways that could explain some of its features. However, it is too soon to make definite statements. The most obvious point is that the IceCube data is presently sparse but will soon improve in both statistics and the types of searches (cascades and tracks, diffuse and point sources, etc.). Combined with multi-messenger studies, this will help identify the origin of the events and thus information about their emitted spectrum. Once there is more data, more detailed calculations will be warranted. Those could explore a wider range of theoretical possibilities for the $\nu$SI scenarios.

As described above, we assume that just one species (flavor or mass eigenstate) of neutrinos and antineutrinos experiences $\nu$SI. This is because the laboratory limits are strong for $\nu_e$ and $\nu_\mu$, but weak for $\nu_\tau$. Thus our calculations are nominally for $\nu_\tau + \bar{\nu}_\tau$. However, the situation is more complex. Because of the vast distances, astrophysical neutrinos propagate as incoherent mass eigenstates, and all of the mass eigenstates have an appreciable $\nu_\tau$ fraction. Whether the primarily $\nu$SI couplings are to flavor or mass eigenstates is model-dependent. The effects of $\nu$SI that we illustrate for one species will be diluted by the lack of $\nu$SI for the other species, but the details are model dependent. A closer look at how the laboratory and astrophysical studies together constrain the different flavor or mass eigenstates is needed. Flavor ratios for astrophysical neutrinos may be an especially important test.

If the neutrino masses are not degenerate, then resonances could occur at different energies, which would lead to more complicated spectra or possibly even cancelations between dips and bumps. In the case that the lightest neutrino is relativistic today, there would be non-negligible thermal broadening. There could also be model-dependent details that complicate the discussion, including by having more than one mediator, by coupling to dark matter, or by having more general final states. We took both final states to be active neutrinos. If only one is, then the absorption dip and spectral cutoff would be unaffected but
the cascade bump would be reduced. If neither is, then only the absorption dip or spectral
cutoff is observable.

3.4 Conclusions and outlook

Neutrinos may still hold surprises, and $\nu$SI are among the possibilities. Their effects can
be probed directly through neutrino-neutrino scattering — provided that we have detected
neutrinos from astrophysical sources traveling through the C$\nu$B. Until recently, this was only
possible with the SN 1987A data [32]. The detection of high-energy neutrinos by IceCube
has opened a new frontier in neutrino astronomy, which provides new opportunities for
probing $\nu$SI. Because the IceCube sources appear to be extragalactic, the column density of
neutrino targets is much greater than for SN 1987A; because the energies are much larger,
a wider range of $\nu$SI parameters can be probed; and, because the observed flux is diffuse,
that averages out the peculiarities of individual sources.

The observed IceCube spectrum contains interesting features, which include a gap at
moderate energies, a possible excess near 1 PeV, and a cutoff at slightly higher ener-
gies [202, 201, 203]. Given the current statistics, these features are consistent with standard
model expectations with simple astrophysical assumptions [259, 260, 261]. It is, however,
interesting to consider exotic explanations such as $\nu$SI, pseudo-Dirac neutrinos [268], or
Lorentz-invariance violation [269, 270].

We perform the first study of $\nu$SI in the context of the detected IceCube spectrum and
its features. Using a phenomenological approach for the interactions, we show that IceCube
is sensitive to an interesting range of $\nu$SI parameters that evades the most robust of the
laboratory limits and is more sensitive than other astrophysical or cosmological techniques.
We provide an improved calculation using the propagation equation, the first for high-
energy neutrinos to take into account $\nu$SI through attenuation, regeneration, and multiple
scattering. Solving the propagation equation numerically, we show $\nu$SI could generate
spectral distortions such as a dip, bump, or cutoff large enough to mimic the features seen
in the IceCube spectrum. Although $\nu$SI might be able to explain some features of the
observed data, it is too soon to draw such conclusions. We expect the IceCube spectrum
will become more precise in the near future by improved statistics and analysis. With that,
more detailed phenomenological studies and associated model-building will be possible.

An expected — but not yet observed — source of high-energy astrophysical neutrinos is
produced through the energy losses of ultra-high-energy cosmic rays propagating through
the CMB [271, 272]. Once these cosmogenic neutrinos [273] are observed, it will be possible
to test \( \nu \)SI using calculations similar to those presented here. Although the cosmogenic
neutrino spectrum is not a simple power law, its shape is reasonably well predicted. For
an energy of \( \sim 10^{10} \) GeV, a resonance with the C\( \nu \)B would probe mediator masses near
\( M \sim 10^3 \) MeV, which are not well constrained (see Fig. 4.11). We do not show the curves
for optical depth \( \tau \) for this case; their shape is similar to that for PeV neutrinos, but
displaced to larger mediator masses and couplings (for a heavy mediator, the sensitivity is
\( g/M \sim 0.4/(10^3 \text{ MeV}) \)). Could \( \nu \)SI explain the non-observation of cosmogenic neutrinos?
While pushing their spectrum to lower energies could be consistent with IceCube data, the
required coupling is relatively large, \( g \sim 1 \).

Our calculations are for a diffuse flux, which is consistent with IceCube data. If point
sources are observed, the effects of deflection and delay should be noted (these are irrelevant
for the diffuse flux). The Lorentz factor \( \gamma \) of the center of momentum frame is \( \sim 10^8 \) for
1 PeV neutrinos scattering on neutrinos of mass 0.1 eV. For one scattering, the deflection
is \( \Delta \theta \sim 10^{-8} (10^8/\gamma) \), which is tiny, and the time delay is \( \Delta t \sim 10 \text{ s} (10^8/\gamma)^2 \), which might
not be negligible in some cases. These effects would be increased by multiple scattering,
ultimately washing out transient and point sources into a steady diffuse flux. For reasonable
couplings, these effects are not relevant for the PeV neutrinos.

For low-energy neutrinos from a nearby supernova, these effects could be much more
important. The delay is smaller by \( \sim 10^6 \) due to the closer distance but larger by \( \sim 10^8 \) due
to the change in \( \gamma^2 \), making \( \Delta t \sim 10^3 \) s. KT87 [32] defined their constraint by changes in the
energy due to energy loss, which requires assumptions about the total energy in neutrinos
and the energy spectrum. The same constraint can be obtained by the simpler time delay
argument, which only requires an assumption about the total energy in neutrinos.
The IceCube neutrino telescope has opened a new age in neutrino astronomy, as well as providing a way to directly test $\nu$SI. Complementary constraints should also be developed for neutrinos in the early universe and core-collapse supernovae. In those settings, even weak-scale neutrino-neutrino collisions and mixing from the self-induced potential are important. The rapid advance of precision cosmology and perhaps a lucky detection of a Milky Way supernova might reveal more secrets about neutrinos.
Chapter 4  
Improved Limits on Sterile Neutrino Dark Matter using Full-Sky Fermi Gamma-Ray Burst Monitor Data

A sterile neutrino of $\sim$keV mass is a well motivated dark matter candidate. Its decay generates an X-ray line that offers a unique target for X-ray telescopes. For the first time, we use the Gamma-ray Burst Monitor (GBM) onboard the Fermi Gamma-Ray Space Telescope to search for sterile neutrino decay lines; our analysis covers the energy range 10–25 keV (sterile neutrino mass 20–50 keV), which is inaccessible to X-ray and gamma-ray satellites such as Chandra, Suzaku, XMM-Newton, and INTEGRAL. The extremely wide field of view of the GBM enables a large fraction of the Milky Way dark matter halo to be probed. After implementing careful data cuts, we obtain $\sim$53 days of full sky observational data. We observe an excess of photons towards the Galactic Center, as expected from astrophysical emission. We search for sterile neutrino decay lines in the energy spectrum, and find no significant signal. From this, we obtain upper limits on the sterile neutrino mixing angle as a function of mass. In the sterile neutrino mass range 25–40 keV, we improve upon previous upper limits by approximately an order of magnitude. Better understanding of detector and astrophysical backgrounds, as well as detector response, will further improve the sensitivity of a search with the GBM.

The contents of this chapter were published in [4].

4.1 Introduction

Right-handed neutral fermions (henceforth sterile neutrinos) arise in many extensions of the Standard Model in explaining the observed flavor oscillations of active neutrinos, and
yield an extremely rich phenomenology (for recent reviews, see, e.g., [274]). Sterile neutrinos
may be produced in core-collapse supernovae [275], providing a new mechanism for explosion
[276], and may explain the origin of strong neutron star kicks [277, 278, 279]. The sterile
neutrino can modify big bang nucleosynthesis [280, 281, 282], assist reionization [283, 284,
285, 286, 287, 288], and affect neutrino oscillations [289].

Moreover, it has been noted that sterile neutrinos can contribute the entirety of the
observed dark matter density. They could be produced in the early universe via oscillation
mechanisms, including non-resonantly [51] or resonantly with active neutrinos [53],
or alternatively via non-oscillation mechanisms, such as decays of heavy particles (see
[290, 291, 292, 293, 294, 295] for some of the scenarios). Sterile neutrinos could be a warm
or cold dark matter candidate [296, 52, 297, 298, 299]. In addition, some sterile neutrino
dark matter models can explain the baryon asymmetry in the Universe [300, 301, 302, 12].

For sterile neutrinos produced via oscillations to be a viable dark matter candidate, they
typically have mass in the 1 – 100 keV range [51, 53, 296]. They can radiatively decay into
an active neutrino and a photon [303, 304]. The photon carries half of the total energy,
and therefore lies in the X-ray energy range. The photon line is strongly distinct from most
astrophysical and detector backgrounds, which have smooth energy spectra. An exception
is line emissions from hot gases and activated detector materials. While the decay lifetime
must be comparable to the age of the Universe in order to ensure that sterile neutrinos
remain a viable dark matter candidate, the decay in high concentrations of dark matter,
e.g., centers of galaxies, clusters of galaxies, and dwarf spheroidal galaxies, can lead to
an appreciable X-ray flux. The large expected flux in many targets, coupled with the
spectral and morphological characteristics of the signal, make searches with X-rays a very
powerful approach for testing sterile neutrino dark matter scenarios (see [305, 12, 306] for
a comprehensive discussion of various searches).

The X-ray constraint on sterile neutrino dark matter was first obtained using the Cosmic
X-ray Background (CXB) [307]. Subsequently a more detailed analysis was carried out in
[308], where the authors considered a set of galaxy clusters, two spiral galaxies, and the
Cosmic X-ray Background (CXB), and used observations by Chandra and XMM-Newton.
Since then, a host of other sources have been explored (for a full discussion, see, e.g., review articles [305, 12]), including other galaxy clusters such as Coma [309, 310], the distant A520 [311], and the Bullet [312]; nearby galaxies such as Andromeda [313, 314, 315, 316] and M33 [317]; additional analysis of the CXB [318, 38, 319, 320]; more recently the Milky Way satellites including the Large Magellanic Cloud [38], Ursa Minor [321, 322], Draco [323], Willman I [324], andSegue I [325]; and, finally, the nearest dark matter concentration, the Milky Way galaxy [38, 321, 326, 319, 327, 39, 328].

Several possible line detections consistent with sterile neutrino dark matter decay have been reported. A line feature was observed in Willman I that could be interpreted as the decay of a sterile neutrino of mass $m_s \approx 5$ keV and $\sin^2 \theta \approx 10^{-9}$, where $\theta$ is the mixing angle between the sterile and active neutrinos [324] (but, see [329, 330]). In another study, X-ray line ratios showed an excess that could be interpreted as arising from the decays of 17 keV sterile neutrinos with $\sin^2 \theta \approx 10^{-12}$ [328]. Most recently, an anomalous X-ray line was detected from galaxy clusters and Andromeda [62, 63] (also see [331, 332, 333, 334, 335, 336, 337, 338]), which can be interpreted as the decay of 7 keV sterile neutrinos [339].

In this work, we use the Gamma-ray Burst Monitor (GBM) onboard the Fermi Gamma-Ray Space Telescope to search for X-ray lines. Notable advantages of the GBM include its all-sky coverage, which allows the entire Milky Way dark matter halo to be explored, and large effective area, yielding a very high statistics data set. The energy range of the GBM extends from 8 keV up to 40 MeV, conveniently filling a gap in energy above the range of previously considered X-ray satellites and below the range of INTEGRAL. Therefore, in this work we focus on this unexplored photon energy range $E_\gamma = 10$–25 keV ($m_s = 20$–50 keV). We consider the Milky Way because of its proximity and well-studied dark matter distribution, and because the GBM detectors are more sensitive to large scale diffuse emission such as from the Galactic halo, due to GBM’s large field of view (FOV) and poor angular resolution.

We describe the expected X-ray signal from sterile neutrino dark matter decays in Sec. 4.2. The GBM instrument and the dark matter signal modeling in the context of the GBM detectors are presented in Sec. 4.3. The data reduction procedures are described
in Sec. 4.4. In Sec. 4.5, we describe the line search analysis and the procedure used to obtain limits on sterile neutrino decays. We summarize in Sec. 4.6. Throughout this work, we adopt cosmological parameters from Planck [16], where $H_0 = 100 h \, \text{km/s/Mpc}$, $h = 0.673$, $\Omega_{\Lambda} = 0.685$, $\Omega_M = 0.315$, $h(z) = \sqrt{\Omega_{\Lambda} + \Omega_M (1 + z)^3}$, the dark matter fraction $\Omega_{DM} = 0.265$, and $\rho_c = 1.05 \times 10^{-5} h^2 \, \text{GeV cm}^{-3}$.

4.2 Expected signal flux

![Figure 4.1: The J-factor, $J(\psi)$ (Eq. (4.4)), as a function of half opening angle $\psi$ relative to the GC, for four Milky Way dark matter halo profiles.](image-url)
Table 4.1: Dark matter profile parameters for widely adopted dark matter profiles in the literature. Our canonical profile is the NFW profile.

<table>
<thead>
<tr>
<th>Profile</th>
<th>α</th>
<th>β</th>
<th>γ</th>
<th>Rs [kpc]</th>
</tr>
</thead>
<tbody>
<tr>
<td>NFW</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>cNFW</td>
<td>1</td>
<td>3</td>
<td>1.15</td>
<td>23.7</td>
</tr>
<tr>
<td>Cored isothermal (ISO)</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>3.5</td>
</tr>
<tr>
<td>Einasto (EIN)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>20</td>
</tr>
</tbody>
</table>

The primary decay channel of sterile neutrinos is into three light active neutrinos. The radiative decay into an active neutrino and a photon that we are interested in is suppressed by a factor of $27\alpha/8\pi \approx 1/128$ relative to the primary decay channel \cite{304}, and has a decay rate \cite{303, 308}

\[
\Gamma_s \simeq 1.36 \times 10^{-32} \text{s}^{-1} \left( \frac{\sin^2 2\theta}{10^{-10}} \right) \left( \frac{m_s}{1\text{keV}} \right)^5,
\]

where we have assumed a Majorana sterile neutrino (for a Dirac sterile neutrino the decay rate is halved). The energy luminosity of decay photons arising from a sterile neutrino dark matter clump of mass $M_{DM}$ is given by $L_\gamma = E_\gamma (M_{DM}/m_s) \Gamma_s$, where $E_\gamma = m_s/2$ is the photon energy, and equals

\[
L_\gamma \simeq 1.2 \times 10^{38} \text{erg s}^{-1} \left( \frac{M_{DM}}{10^{11} M_\odot} \right) \left( \frac{\sin^2 2\theta}{10^{-10}} \right) \left( \frac{m_s}{10^{10}\text{keV}} \right)^5,
\]

for a typical galaxy-size dark matter halo mass. It can be immediately appreciated that this is comparable to the total luminosity of astrophysical X-rays in the Milky Way in the $2 - 10$ keV range, $\sim 10^{39} \text{erg s}^{-1}$ \cite{340}, or the total Milky Way diffuse emission in the same energy range, $\sim 10^{38} \text{erg s}^{-1}$ \cite{341}.

The photon intensity (number flux per solid angle) of sterile neutrino dark matter decay coming from an angle $\psi$ away from the Galactic Center (GC) consists of both the Galactic and the extragalactic components,

\[
I(\psi, E) \equiv \frac{dN}{dAdTd\Omega dE} = \frac{\rho_\odot R_\odot}{4\pi m_s \tau_s} \mathcal{J}(\psi) \frac{dN}{dE} + \frac{\Omega_{DM} \rho_c}{4\pi m_s \tau_s H_0} \int \frac{dz}{h(z)} \frac{dN}{dE'},
\]

\[
= \frac{\rho_\odot R_\odot}{4\pi m_s \tau_s} \left[ \mathcal{J}(\psi) \frac{dN}{dE} + R_{EG} \int \frac{dz}{h(z)} \frac{dN}{dE'} \right],
\]

91
where $\tau_s = 1/\Gamma_s$ is the lifetime, $\rho_\odot = 0.3 \text{ GeV cm}^{-3}$ is the local dark matter mass density, $R_\odot = 8.5 \text{ kpc}$ is the Sun’s distance to the GC, and $dN/dE = \delta(E - m_s/2)$ is the dark matter decay spectrum. The first term in the bracket is the Galactic component. The so-called J-factor, $\mathcal{J}(\psi)$, is the integral of the dark matter mass density $\rho$ in the Milky Way halo along the line-of-sight,

$$\mathcal{J}(\psi) = \frac{1}{\rho_\odot R_\odot} \int_0^{\ell_{\text{max}}} d\ell \, \rho(\psi, \ell),$$

where $\ell_{\text{max}}$ is the outer limit of the dark matter halo. We assume the dark matter distribution is spherically symmetric about the GC, hence

$$\rho(\psi, \ell) = \rho(r_{\text{GC}}(\psi, \ell)) = \rho \left( \sqrt{R_\odot^2 - 2 \ell R_\odot \cos \psi + \ell^2} \right).$$

The value of $\ell_{\text{max}}$ differs depending on the adopted halo model, but the contribution to $\mathcal{J}(\psi)$ from beyond $\sim 30 \text{ kpc}$ is negligible. We adopt $\ell_{\text{max}} = 250 \text{ kpc}$ in this work.

The second term in the bracket of Eq. (4.3) describes the isotropic extragalactic component, where $E' = E(1 + z)$. The factor $R_{\text{EG}}$ roughly compares the contribution of the extragalactic component versus the Galactic component, up to the shape of the energy spectrum.

$$R_{\text{EG}} \equiv \frac{c}{H_0} \frac{\Omega_{\text{DM}} \rho_c}{\rho_\odot R_\odot} \approx 2.$$  

Normally, the extragalactic component can be ignored as typically the analysis region is chosen to be a small patch of the sky where the Galactic component is much larger (e.g., the GC, where $\mathcal{J} \gg 1$). However, in our case, the large FOV of the GBM makes the extragalactic component non-negligible.

The dark matter density profile $\rho(r)$ of the Milky Way is not precisely known, in particular at small Galactic radius. We consider several fitting functions that capture the results of numerical simulations of dark matter halo profiles, which can be parameterized by the following form,

$$\rho^{\alpha\beta\gamma}(r) = \rho_\odot \left( \frac{r}{R_\odot} \right)^{-\gamma} \left[ \frac{1 + (R_\odot/R_s)^\alpha}{1 + (r/R_s)^\alpha} \right]^{(\beta-\gamma)/\alpha},$$

where parameters for commonly used profiles are summarized in Table 4.1. Another profile favored by recent simulations is the Einasto profile,
\[
\rho_{\text{Ein}}(r) = \rho_\odot \exp\left( -\frac{2}{\alpha_E} \frac{r^{\alpha_E} - R_{s}^{\alpha_E}}{R_{s}^{\alpha_E}} \right),
\]

(4.8)

with \( \alpha_E = 0.17 \) and scale radius \( R_s = 20 \) kpc. These profiles differ mainly at small Galactic radius. The first three profiles have constant logarithmic slopes at small radii, which are described by the \( \gamma \) factor. The Einasto profile has the same slope as the NFW profile at the scale radius, but the slope decreases as the radius decreases.

In Fig. 4.1, we show the J-factor \( J(\psi) \) for each dark matter profile as a function of the angle \( \psi \) viewed away from the GC. The differences between profiles are relatively small, because the density \( \rho \) appears linearly in the decay flux (as opposed to in the annihilation flux where the density appears quadratically). We use the NFW profile as our canonical profile in this work. As will be shown in Sec. 4.3.2, the impact of varying the profile is minimal after taking into account the detector response and the FOV. Thus the sterile neutrino constraint obtained using GBM is robust against dark matter profile uncertainties.

A crude estimate of the expected number of photons \( \nu_\gamma \) per unit time \( T \) from Galactic dark matter decay is

\[
\frac{d\nu_\gamma}{dT} \sim 20 \, \text{s}^{-1} \left( \frac{A_{\text{eff}} \Omega}{20 \pi \text{ cm}^2 \text{ sr}} \right) \left( J_{60} \right) \times \\
\left( \frac{\sin^2 2\theta}{10^{-11}} \right) \left( \frac{m_\gamma}{20 \text{ keV}} \right)^4,
\]

(4.9)

where we use representative values for the effective area and solid angle, the J-factor at \( \psi = 60^\circ \), \( J_{60} \), and a nominal sterile neutrino mixing angle. It is immediately clear that even a small fraction of the total Fermi-GBM live time can yield significant number of signal photons.

### 4.3 Instrument and signal modeling

#### 4.3.1 GBM Instrumentation

The GBM consists of 14 detectors: 12 NaI detectors, each operating over energies from 8 keV to 1 MeV, and 2 BGO detectors, each operating over energies from 200 keV to 40 MeV. The NaI detectors are located on the corners and sides of the spacecraft, with different
orientations, and they together provide a nearly complete coverage of the occulted sky. At any given time, typically 3–4 NaI detectors view the Earth within 60 degrees of the detector zenith, i.e., their FOV is occulted by the Earth.

Figure 4.2: The effective area for det-7 NaI detector versus the detector zenith angle relative to the detector normal, for three example energy bins. Points are data from GBM calibration files, and the anomalous dips in the effective area come from blockages from other satellite components. Solid lines are fits to the data neglecting the dips.

Not all of the NaI detectors are best suited for dark matter searches. At first consideration, det-0 and det-6 would seem to be the best detectors to use since they are aligned
close to the LAT zenith ($\simeq 20^\circ$ offset). However, we find that significant parts of the FOV of these two detectors are actually blocked by the LAT itself. Also, half of the detectors are pointed towards the Sun all the time, and X-ray emissions from the Sun contaminate their low energy spectrum. Lastly, some detectors are pointed sideways, i.e. $\simeq 90^\circ$ relative the LAT-zenith, which suffer large FOV blockage from the Earth. Ruling out these detectors, only det-7 and det-9 seem to be suitable, which are $\simeq 45^\circ$ relative the LAT-zenith. Upon inspection, we observe an anomalous spectral feature in the low energy spectrum of det-9 compared to other detectors. As a result, we use det-7 as our fiducial detector for analysis. As will be shown below, this analysis is systematically limited, rather than statistically limited. Using only one detector for this analysis also avoids introducing systematic uncertainties from combining multiple sets of data from different detectors.

The NaI detectors have a wide FOV, as seen in Fig. 4.2, which shows the effective area versus the detector zenith angle, $\theta$, for the det-7 detector. We obtain the GBM effective area data from detector response matrix files (GS-008). Each file contains the effective area as a function of energy, for a specific detector zenith and azimuthal angle. In Fig. 4.2, each point denotes the effective area extracted versus the corresponding zenith angle from the detector response file. Beyond about $40^\circ$, we observe anomalous dips in certain azimuthal direction at all energies, which is presumably caused by blockages from satellite components in the FOV of the detector. We remove these anomalous dips by requiring adjacent bins deviate no more than $\sim 10\%$. After this procedure, the angular dependence of the effective area can be well-described by cosine functions, as shown by the solid lines.

In Fig. 4.3, we show the energy dependence of the effective area for three representative zenith angles. The points are obtained from the detector response matrix files, and are chosen from a specific azimuthal angle at which the detector FOV is not blocked. Using the cosine fits described above, the energy dependence is obtained by linear interpolation of the model in energy. Thus, we obtain an azimuthally symmetric model of the effective area (i.e., it depends only on zenith angle and energy). Shown in solid lines in Fig. 4.3, are the model for the given zenith angles. The FOV blockages slightly reduce the detector sensitivity towards a particular azimuthal angle. This effect, however, is not expected to introduce
Figure 4.3: The effective area versus the incident photon energy for three incident angles. The points are obtained from GBM calibration files and the lines are linear interpolations of the points.
spurious spectral features in energy spectrum, since the blockages affect all energies. We thus neglect these blockages and use the smooth angular fits to model the expected signal in the next section.

One important feature of the NaI detectors is that they are limited in their photon-tracking capabilities, i.e., one cannot simply obtain the photon flux as a function of the incidence direction for a specific position on the sky. In other words, the tradeoff for the large FOV is poor angular resolution. Earth occultation techniques can be employed to obtain photon direction for point source studies [342, 343], but this technique has not yet been demonstrated for diffuse emissions. Fortunately, the lack of photon tracking is not very problematic for sterile neutrino dark matter decay searches due to the large angular extent of the expected emission. However, this does mean that one cannot accurately construct an intensity sky map of the GBM data (Eq. (4.3)). As a result, one needs to properly model the signal taking into account the detector response to match the observable. In this case, the instrumental observable is counts rate (number of photons per second), as a function of the NaI detector pointing direction.

4.3.2 Expected Signal Modeling

Given the sterile neutrino decay photon intensity $I(\psi, E)$, we compute the expected number of photons, $\nu_{i,j}$, for energy bin $i$ and detector sky-pointing direction $j$. The expected number of photons per observing time, $T_j$, from a particular detector pointing direction is then

$$\frac{d\nu_{i,j}}{dT_j} = \int_{E_{i}^{\text{min}}}^{E_{i}^{\text{max}}} dE \int_{2\pi} d\Omega(\theta) \int d\tilde{E} \left\{ I(\psi, \tilde{E}) G(E, \tilde{E}) A_{\text{eff}}(\tilde{E}, \theta) \right\} ,$$

(4.10)

where $E_{i}^{\text{max}}$ and $E_{i}^{\text{min}}$ are the boundaries of the energy bin $i$. We integrate over the hemisphere the NaI detector points at, i.e., over the detector zenith angle $\theta$, and attribute all the photons to pixel $j$. A position on the sky with an angle relative to the GC, $\psi$, is related to the detector zenith angle and the pixel that the detector points at through $\psi \rightarrow \psi(\theta, j)$.

The pointing direction of the detector is therefore defined by $\psi(0, j)$. The factor $G(E, \tilde{E})$
takes into account the energy resolution of the NaI detector, which we model as a Gaussian
with width given by the pre-launch calibrations [344, 345]. The energy resolution is about
10% for our analysis range. And lastly, $A_{\text{eff}}(E, \theta)$ is our NaI detector effective area model,
which is a function of energy and the detector zenith angle, as in Figs. 4.2, 4.3.

Using the Dirac-delta function for the energy spectrum, the expected signal is

$$\frac{d\nu_{i,j}}{dT_j}(m_s) = \frac{\rho_\odot R_\odot}{4\pi m_s \tau} \int_{E_{i \min}}^{E_{i \max}} dE \left\{ \mathcal{N}\left(\frac{m_s}{2}\right) \tilde{J}\left(\frac{m_s}{2}, j\right) G\left(E, \frac{m_s}{2}\right) + R_{\text{EG}} \int d\tilde{E} \int d\Omega A_{\text{eff}} G(E, \tilde{E}) \Theta\left(\frac{m_s}{2} - \tilde{E}\right) \right\},$$

where $\tilde{J}(E, j)$ is the “convolved J-factor”, $\mathcal{N}(E)$ is a normalizing factor, $h(E) = \sqrt{\Omega_\Lambda + \Omega_M (m_s/(2E))^3}$, $\Theta$ is the Heaviside step function. We have suppressed the argument of $A_{\text{eff}}$ for simplicity.

The convolved J-factor is defined as

$$\tilde{J}(E, j) = \frac{\int J(\psi) A_{\text{eff}}(E, \theta) d\Omega(\theta)}{\mathcal{N}(E)},$$

which takes into account the effect of detector response; it represents the J-factor defined by
detector pointing directions. It depends on the detector pointing direction through $\psi(\theta, j)$.
The normalization factor $\mathcal{N}(E) \equiv 2\pi A_{\text{eff}}(E, 0)$ captures the energy behavior of the effective
area. The normalization of this factor is unimportant as it cancels itself when obtaining
dark matter decay fluxes/limits. In Fig. 4.4, we compare the convolved J-factor with the
normal J-factor defined in Eq. (4.4). Once the detector response is taken into account, the
difference between different profiles decreases drastically even for pointing directions very
close to the GC. E.g., for 10–11 keV bin, the difference in the convolved J-factor between
NFW versus EIN and ISO is $\lesssim 1\%$. Therefore, systematic uncertainties due to the choice
of dark matter profile are minimal.

In the left column of Fig. 4.5, we show the modeled dark matter maps for the NaI
detector from the Milky Way halo for several energies. We pixelate the sky into 768 pixels
of equal solid angle using the HEALPix scheme\textsuperscript{1}, i.e., each pixel corresponds to a solid angle of $\Delta \Omega \simeq 1.6 \times 10^{-2}$ sr. We use the Milky Way contribution from Eq. (4.11), which takes into account detector energy and angular response. The extragalactic component only adds a constant value to the signal map. We choose the line energy to be at the center of the chosen energy bin. The decay rates for the sterile neutrino scenarios are chosen to approximately match the count rates of the corresponding data maps (right column, described below). By construction, the Milky Way dark matter contribution is spherically symmetric, and the large angular extent of the signal is due to the large FOV and poor angular resolution of the NaI detectors.

### 4.4 Data selection and reduction

In this section, we describe the data reduction procedures to improve the data quality and the cuts designed to reduce various backgrounds. At the end we obtain a data set that can be compared to Eq. (4.11) to obtain limits for sterile neutrino dark matter decay.

We use GBM daily data from 12-AUG-2008 to 31-DEC-2012, a total of 1601 days. We use the \textit{CSPEC} data (GS-002) with nominal 4.096 s time resolution and 128 channels in energy from 5 to 1402 keV (the first and last few energy bins are not usable). We then devise several cuts to improve the data quality. The goal is to obtain a data set that is representative of the diffuse sky emission as observed by the GBM NaI detectors, while minimizing various types of backgrounds. The most dominant source of background is due to cosmic rays interacting with the satellite, directly activating the detector or triggering the detector through delayed radioactive decays of the satellite material.

To this end, we employ the following cuts:

- \textit{LAT} cut.

We first select data sets that are suitable for analysis using data flags from Fermi-LAT weekly photon files: \texttt{LAT_CONFIG=1, LAT_MODE=5, DATAQUAL=1, ROCK_ANGLE<50, SAA=F}. The first three conditions ensure the detector configuration and data quality

\textsuperscript{1}http://healpix.jpl.nasa.gov [346]
Figure 4.4: The convolved J-factor (Eq. (4.12)) versus the opening angle with respect to the GC, $\psi(0,j)$, defined by where the detector normal is pointing. The difference between different profiles is drastically reduced. A small energy dependence is introduced from the effective area. The vertical dotted line denotes the boundary of our ROI. Shown in grey are the theoretical J-factors from Fig. 4.1.
are suitable for scientific analysis. The fourth condition ensures that the Earth is not in front of the LAT’s FOV, which is approximately, but not exactly, the FOV of the NaI detector (we address this in the Earth cut below). The last condition excludes the times when the satellite is inside the South Atlantic Anomaly (SAA), where the high cosmic ray activity significantly increases the radioactivity of the satellite. The GBM detectors are turned off during SAA passage, hence the observed counts are zero in these time periods.

The LAT cuts alone, however, are insufficient for reducing background events, because of the different physical locations of the detectors on the satellite, different backgrounds, and the different technologies of the LAT and the GBM. We therefore develop new cuts specifically for the GBM.

- **Transient sources cut.**
  
  This cut removes the epochs when the GBM detectors detect transient sources, such as gamma-ray bursts, direct cosmic-ray hits, solar flares, Galactic X-ray transients, and magnetospheric events, etc. Though these transients only occupy a small fraction of the observation time, some of them can be bright enough to cause the data acquisition system to overflow.

- **Extended SAA cut.**

  The LAT cut does not completely remove events due to passages of SAA. This is because the satellite is intensively bombarded by cosmic rays during each passage through the SAA, leaving the satellite in a highly radioactive state even after leaving SAA. This effect is even more pronounced for consecutive passes through the SAA. In this case, there is insufficient time for the satellite to return to its normal radioactive state. As a result, orbits passing through the SAA consecutively induce anomalously high photon count rates even when the satellite is outside the SAA. We therefore apply cuts to remove the data collected between consecutive passages of the SAA, in addition to the times that the satellite is physically in the SAA, which are eliminated.
in the \textit{LAT} cut. Removing these orbits is important to reduce events originated from cosmic rays.

- \textit{Earth cut.}

Lastly we apply two cuts on the orientation and the position of the NaI detector relative to the Earth. We first require that the angle between the NaI detector normal and the vector directed from the Earth center to the satellite to be less than 50°. This is to reduce contamination from the Earth limb and occultation from the Earth itself. The next cut is on the geomagnetic coordinate. The high-altitude cosmic ray activity is directly correlated to the Earth’s magnetic field structure. The number of observed background events increases with geomagnetic latitude. To minimize this contamination, we select data only when the geomagnetic latitude is less than |20|°.

In Fig. 4.6, we show an example of the data and the cuts we adopt to improve the data quality. The data points are the counts rates on 20th December 2008 observed by det-7. Each dot corresponds to count rates measured over ~ 4 s. We select the energy range from 344 keV to 471 keV, where the data is dominated by the cosmic-ray-induced background.

The first feature that can be seen in Fig. 4.6 is the series of epochs with no count rate. This is because the detector was shut down when the satellite is in the SAA. These epochs are removed in the \textit{LAT cut}. It is also clear that the count rates are anomalously high even after the satellite leaves the SAA (i.e., right after the gap), due to the increased radioactivity of the satellite. These epochs are removed in the \textit{Extended SAA cut}. These two SAA related cuts are represented by the red hashed regions.

The second feature is the oscillatory shape during the middle of the day. Overlaying the data points we also plot the location of the satellite in geomagnetic latitude (blue line). One can see the count rates are correlated with the geomagnetic latitude. We therefore remove all the data recorded when the geomagnetic latitude is larger than |20|°. This cut is represented by the grey shaded region, and the removed data points are labelled in blue. The choice of a uniform |20|° geomagnetic latitude cut is a balancing act between maximizing sky coverage and reducing background. More sophisticated cuts may be possible.
Figure 4.5: **(Left)** Simulated dark matter counts rate maps in Galactic coordinates for several line energies, taking into account detector response. For each map, the assumed line energy is contained in the energy bin shown and $\sin^2 2 \theta$ (as labeled) are chosen to approximately match the observed counts in the same bin. **(Right)** The final counts rate sample from 4 years of data from the NaI detector, which corresponds to 4.6 million seconds ($\sim 53$ days) of live time after data cuts. All of the sky maps are pixelated into 768 HEALPix pixels. The pixel position corresponds to the pointing direction of the detector normal. The grey pixels are where no observing time is registered after the selection cuts.
Figure 4.6: A sample of the raw photon count rates for the GBM-NaI detector from 20-12-2008, from 344 keV to 471 keV, where the counts rate is dominated by the cosmic ray induced background. The blue line indicates the location of the satellite in geomagnetic latitude (in absolute value). The epochs with no data are when the satellite is physically in the SAA. The red hashed regions represent the cuts on orbits that pass through the SAA. The grey shaded regions illustrate the cuts on geomagnetic latitude; the data removed by this cut is colored blue. In the end, only black data points in the white regions are used for analysis.

The importance of our cuts for improving the data quality can be estimated in Fig. 4.6. The increased count rate right after SAA can be a factor of a few higher. Even the variation due to geomagnetic latitude can be up to a factor of two. Our Extended SAA cut and Earth cut is therefore necessary to reveal the astrophysical component, which is comparable to the detector background at low energies (shown below). Transient sources does not contribute significantly to the total counts, but they can dominate a particular sky pixel. Since they only contribute a small fraction of the live time, the Transient sources cut is very efficient.

The final data products obtained are observed counts and exposure time over 128 energy bins and 768 sky pixels. The total live time of the data product is \( \sim 4.6 \times 10^6 \) seconds (~53 days). Despite having a huge reduction from the raw data, we are still far from statistically limited, as will be shown below.

In the right panel of Fig. 4.5, we show the counts rate sky map for the labeled energy bins. At low energies, we observe a clear excess towards the GC. We interpret the excess as astrophysical (i.e., non-instrumental-related) emissions from the Milky Way. The astrophysical flux is about \( \sim 10^{-1} \text{cm}^{-2}\text{s}^{-1} \) if one extrapolate from the high energy observations [347, 348, 349], which matches the observed counts rate of about \( \sim 10\text{s}^{-1} \). The
observed excess towards the GC also shows a small north-south asymmetry, which probably reflects the underlying distribution of diffuse and discrete X-ray sources.

For the maps at high energies, the morphology is significantly more isotropic than at low energies, with small variations that trace orbital structure, as expected from cosmic-ray-induced backgrounds. For example, the two dark spots near the orbital pole in high energies is also seen in the low energy map.

As a result, we conclude the low energy data set consists of a mixture of astrophysical diffuse and point source emissions, plus residual cosmic-ray-induced background. The grey pixels in the maps represent positions on the sky that were not visited by the detector, and are excluded from the analysis.

Using the data sky map and the convolved J-factor $\tilde{J}(E, j)$, we can determine the region of interest (ROI) for our analysis. As $\tilde{J}(E, j)$ flattens out at small angles due to the poor detector angular resolution, as shown in Fig. 4.4, there is little benefit in choosing a small ROI. We carry out a signal-to-noise study to look for an optimal ROI angle. The morphology of the GC excess seen in low energies turns out to be comparable to the smoothed dark matter distribution, and the signal-to-noise is fairly insensitive to the choice of angle. This is a direct consequence of the poor angular resolution of the NaI detector. We conservatively choose a large ROI, which consists of pixels within $60^\circ$ from the GC, i.e., $\psi < 60^\circ$. With this selection, we have enough pixels to average out potentially spurious behavior in some individual pixels, and have more than enough statistics. Lastly, this ROI only minimally overlaps with the dark spot positions near the orbital poles.

In Fig. 4.7, we show the binned counts spectrum for the data sample in the GC ROI. As a comparison, we also show the spectrum for the anti-GC ROI ($\psi > 120^\circ$). The total observed time for the two samples are 975066 s and 911451 s, respectively, and this difference is the main reason the normalization differ in high energies. In general, the counts spectrum has a power-law behavior at high energies, as expected from cosmic-ray induced background. There are several prominent line features from excited energy levels of $^{127}$I at 57.6 keV and 202.9 keV, as well as the 511 keV line from positron annihilation from the atmosphere and nearby materials [345]. At low energies, the GC and anti-GC spectral shape starts
to deviate, and the difference in normalization increases compared to high energies. This indicates that the astrophysical component starts to appear in the GC sample.

4.5 Limits on sterile neutrinos

We present two limits on sterile neutrino decay lines. The first is a conservative limit based purely on flux comparison. The second uses the fact that the signal is a photon line, while the background flux is approximately a power-law within the search energy window.

4.5.1 Flux Analysis

The most robust constraint one can place on the amplitude of a sterile neutrino dark matter decay signal is to require that the expected signal counts do not exceed the total measured counts. For a set of dark matter masses, we compare, bin by bin, the predicted counts from sterile neutrino decay to the total counts measured. This approach therefore assumes the hypothesized signal dominates the observed spectrum without any assumptions about the detector and astrophysical background.

The expected signal counts are given by summing the count rates in all the individual sky pixels within the ROI, using Eq. (4.11), weighted by the actual observing time $T_j$ in each pixel,

$$\nu_i = \sum_j^{\text{ROI}} T_j \frac{dv_{i,j}}{dT_j}. \tag{4.13}$$

The measured photon counts data from all pixels in the ROI is

$$d_i = \sum_j^{\text{ROI}} N_{i,j}, \tag{4.14}$$

where $N_{i,j}$ is the number of counts in energy bin $i$ and pixel $j$ measured by the GBM detector.

We obtain the flux analysis limit on the decay rate, $\Gamma_s$, by requiring $\nu_i < d_i$ for all energy bins for each $m_s$. The limit obtained this way is very conservative. It is unlikely that sterile neutrino decay, which has a sharp spectral shape, would dominate a narrow energy range.
in the count spectrum while other components conspire to vanish in that particular energy range.

Figure 4.7: The counts spectrum for the final data sample for both GC ROI ($\psi < 60^\circ$) and anti-GC ROI ($\psi > 120^\circ$) chosen to have the same solid angle. The dominant component is a power-law plus various background lines. The overall difference in normalization is due to the higher exposure towards the GC than the anti-GC direction. The additional excess at low energies towards the GC region suggests the rise of the astrophysical component. The vertical dotted line indicates the energy bins used for the spectral analysis.
4.5.2 Spectral Analysis

The sensitivity to sterile neutrino dark matter decay can be improved dramatically using the observation that the sterile neutrino decay signal and the dominant background have different spectral shapes.

A simple background model

As shown above, our GC data sample contains an astrophysical component as well as internal detector backgrounds. The astrophysical contribution from the inner galaxy is dominated by points sources (all unresolved by GBM), and the energy spectrum was shown to be well described by a power law above 20 keV [347, 348, 349]. The internal detector background is a consequence of cosmic rays interacting with the satellite components, which retains the power-law behavior of the incoming cosmic rays. We therefore expect the energy spectrum to have a power law distribution.

A power-law spectrum is an even better approximation when we analyze the data in small energy windows. We consider 15 of such search windows, one for each line energy of interest. The line energies are the corresponding energies of the energy bin number 6 to 20 in GBM numbering scheme (labeled by $i_0$). For each search bin $i_0$, the search window contains a number of energy bins (labeled by $i$), where

$$\text{Max} \left( i_{\text{min}}, i_0 - \Delta w \right) < i < i_0 + \Delta w. \quad (4.15)$$

The window size is $\Delta w = 5$, which makes the window width on each side about 3–4$\sigma$ of the energy resolution at the line energy. For line energies near the low energy cutoff, we truncate the search window at the lowest usable energy bin, $i_{\text{min}} = 6$, which corresponds to a central bin energy of 9.3 keV. The signal line energy in such case is not located in the center bin of the search window.

With the power-law assumption for the non-dark matter components in each search window, the model photon counts spectrum therefore contains the dark matter signal component ($d\nu/dE$) and a power-law background component ($db/dE$),
\[ \frac{d\nu}{dE} = f_s \left\{ \delta(E - E_0)N(E) + \right. \]
\[ R_{\mathrm{EG}} \frac{\sum T_j \tilde{J}(E_0, j)}{\sum T_j \tilde{J}(E_0, j)} \int A_{\text{eff}}(E) d\Omega \frac{\theta(E_0 - E)}{E h(E)} \right\} ; \]

\[ \frac{d\nu}{dE} = \beta \left( \frac{E}{E_0} \right)^{-\gamma} N(E) , \] (4.17)

respectively, where \( E_0 \) is the energy of bin \( i_0 \). The factor \( N(E) \) takes into account the energy response of the effective area. The model has only three free parameters, \( f_s, \beta, \) and \( \gamma \). The factor \( f_s \) is the amplitude of the dark matter signal, which is the only parameter that we are interested in. The normalization and the spectral index of the background power law are thus treated as nuisance parameters, \( \Xi = (\beta, \gamma) \).

The total expected counts in energy bin \( i \) in the search window is then obtained by convolving with the detector energy resolution and integrating the model over the energy bin,
\[ \nu_i + b_i = \int_{E_{\text{min}}^i}^{E_{\text{max}}^i} dE \int G(\tilde{E}, E) \left( \frac{d\nu}{dE} + \frac{d\nu}{dE} \right) G(\tilde{E}, E) . \] (4.18)

Comparing the data model (Eq. (4.13)) with the expected signal (Eq. 4.18), the line amplitude \( f_s \) is related to sterile neutrino parameters by
\[ f_s = \frac{\rho_{\odot} R_{\odot}}{4\pi m_{s} \tau} \sum T_j \tilde{J}(E_0, j) \] (4.19)

\[ = \frac{8.6 \times 10^{-2} \text{cm}^{-2} \text{s}^{-1} \text{sr}^{-1} \left( \sin^2 \frac{2\theta}{10} \right)}{10^{-10} \left( \frac{m_{s}}{10 \text{ keV}} \right)^4} \times \sum T_j \tilde{J}(E_0, j) . \]

To search for a line signal from the data, it is important to understand the uncertainties associated with the measurement. We first consider the systematic uncertainty in the effective area of the detector, which is \( \sim 5\% \) according to the GBM collaboration [344, 345]. Note the quoted uncertainty is the total uncertainty for the effective area, which in principle can be two different kinds of uncertainty. The first kind is the overall uncertainty on the effective area across all energy bins, which affects the value of the flux obtained from data. The second kind is the uncorrelated errors between energy bins, which may introduce spurious spectral features even if the true flux spectrum and the true effective area are both
Figure 4.8: The blue points are the measured data with error bars indicating the 5% systematic error. The blue line is the best fit model to the data with only the power-law component. The red line shows the best fit model when including the line signal with 95% upper limit amplitude. The red arrow indicates the central energy of the line signal.
smooth in energy. For a spectral analysis, the uncorrelated error among energy bins is much more important than simply a normalization shift. In this work, we conservatively attribute all the 5% uncertainty to the uncorrelated errors. As a result, the model uncertainty for each energy bin is

$$\sigma_{A\text{eff}} = 0.05(\nu_i + b_i).$$

(4.20)

We then consider the statistical uncertainty. As shown in Fig. 4.7, the number of photons is enormous in the energy range that we are interested in. The statistical uncertainty in each bin is small, $\sqrt{N}/N < 10^{-3}$. Therefore, we safely ignore the statistical uncertainties in this work.

We adopt the method of maximum likelihood for fitting the counts spectrum for each search window. For each search bin $i_0$, we assume a Gaussian probability distribution function for each energy bin in the search window, giving the likelihood function:

$$L(f_s, \Xi | i_0) = \prod_i \frac{1}{\sqrt{2\pi}\sigma_{A\text{eff}}} e^{-\frac{(\nu_i + b_i - d_i)^2}{2\sigma_{A\text{eff}}^2}},$$

(4.21)

where the product is taken over the energy bins $i$ in the search window. Best fit parameters are obtained by maximizing the likelihood function, or equivalently minimizing its negative logarithm.

We first find the best-fit background only parameters, $\Xi_0$, where $f_s$ is set to zero. The implicitly defined $\Xi_0$ is given by

$$\lambda(f_s = 0, \Xi_0 | i_0) = \text{Min}\{-2\text{Log}L(f_s = 0, \Xi | i_0); \Xi\}.$$

(4.22)

We check whether the power-law only background model is a reasonable hypothesis by computing the reduced $\chi^2$ ($\chi^2$ per degree of freedom) for each search window. We find that the reduced $\chi^2$ ranges from 0.2 to 1.4 in our analysis range. We therefore conclude that the power-law only model plus the prescribed 5% systematic error can reasonably describe the data for each search window.

In Fig. 4.8, we show explicitly the 15 search windows for this analysis. The blue data points are the GBM data, and the assigned error bars are the 5% systematic uncertainty.
Figure 4.9: The black solid line is the 95% C.L. upper limit from the spectral analysis. The green (yellow) shaded region shows the 68% (95%) intervals from the Monte Carlo simulations.
The statistical errors are too small to be shown. The blue lines are the best fit count spectrum from the power-law only model described above. The apparent peculiar spectral features, such as those around 18 and 26 keV, are successfully captured by the power-law model when effective area and non-uniform energy bins are taken into account.

**Limits on dark matter decay rate**

To search or constrain the line signal, we use the so-called profile likelihood method [350]. We search for the best fit line amplitude by minimizing the negative log-likelihood with respect to all the model parameters,

$$\lambda(f_s, \Xi|i_0) = \text{Min}\{-2\text{Log} \mathcal{L}(f_s, \Xi|i_0); f_s, \Xi\}, \quad (4.23)$$

where $f_s$ is constraint to be non-negative. We observe no significant preference for the presence of the line signal. We then proceed to find the 95% C.L. one-sided upper limits on the dark matter signal amplitude, $f^9_{s5}$, by increasing the amplitude while continuously minimizing the log-likelihood function over the nuisance parameters, until it is 2.71 larger than the best fit log-likelihood,

$$\text{Min}\{-2\text{Log} \mathcal{L}(f^9_{s5}, \Xi|i_0); \Xi\} \equiv \lambda(f_{s0}, \Xi_0|i_0) + 2.71. \quad (4.24)$$

The best fit model when the line signal is at 95% upper limit is shown in Fig. 4.8 in red lines. The red arrow indicates the energy of the inserted X-ray line. Using $f^9_{s5}$ and Eq. (4.19), we then obtain the 95% C.L. upper limit of the dark matter decay rate.

We perform a Monte Carlo study to check the robustness of the limit. For each search window, we generate 100 mock data sets, according to the best fit power-law only parameters and the 5% systematic error with Gaussian probability distribution function. We perform the profile likelihood analysis to obtain the 95% upper limits for the mock data sets. In Fig. 4.9, we show the obtained upper limit from data and the 68% and 95% coverage of the limits from our Monte Carlo simulations.

Overall, we find that the observed limit is consistent with our Monte Carlo realizations at the 95% level. At about 28 keV sterile neutrino mass, we find the actual limit touches the
Figure 4.10: The conservative upper limit from the flux analysis (Black) and the 95% C.L. upper limit from the spectral analysis (Blue) for the decay rate of sterile neutrino dark matter. The hashed regions are excluded by the corresponding analyses.
95% lower bound of the expect limit. This is likely due to the data point at about 14 keV falls below what one would expect from a smooth power-law flux spectrum, as shown in the fits in Fig. 4.8. There are no known detector defects at this energy [344, 345], we thus consider this as a $\sim 2\sigma$ systematic downward fluctuation in the effective area model. This downward shift effectively means the data prefers a negative line, which results in the improved limit at this energy. It is also important to note that the limits are correlated due to the largely overlapping data points in adjacent search windows. As a result, limits from line energies close to 14 keV are all slightly improved. The simulated limits from our Monte Carlo realizations are not correlated with adjacent line energies, since the mock data sets are generated independently for each search window.

Finally, Fig. 4.10 shows the limits obtained on the decay rate from both the flux analysis and the spectral analysis, with the hashed region corresponding to the excluded parameter space. As expected, the spectral analysis produces a much stronger limit than the flux analysis. Since the presence of a line signal mostly only affects one energy bin, one would expect the spectral analysis limit is approximately given by the size of the error bars of the data points, and thus the spectral analysis limit is expected to be about 5% of the flux analysis limit. This is indeed the case in most of the mass range, except where the data prefers a negative line, as discussed above. At low energies, the spectral analysis limit deteriorates rapidly. This is due to the imposed lower cutoff of the search energy window. As the line energy approaches the boundary, the number of bins used for the fit is reduced accordingly, and the spectral shape becomes increasingly degenerate with the power-law shape. Both factors cause the limit from spectral analysis to deteriorate.

### 4.5.3 Limits on Sterile Neutrino Dark Matter

Using the upper limits on the decay rate, we derive the corresponding upper limit on the mixing angles for sterile neutrino dark matter. In Fig. 4.11, we show the constraint on the mixing angle–mass plane. For comparison, we also show the only limit in this energy range, obtained with CXB observations using *HEAO-1* [38], and with Milky Way observations using *INTEGRAL* [39]. Unsurprisingly, the flux analysis does not yield competitive limits.
However, in the mass range $m_s \sim 25 - 50$ keV, the spectral analysis improves the limit on the mixing angles by about an order of magnitude compared to the previously strongest limit.

Compared to the previous analysis [38], which uses HEAO-1 A4 Low Energy Detector data [351], our analysis improves mainly in: The GBM data has smaller error bars compared to the HEAO-1 data ($\sim 5\%$ vs $\sim 10\%$); We have employed several cuts to reduce cosmic rays induced backgrounds; The Milky Way halo yields a larger signal flux than the CXB alone; The GBM NaI detector has a slightly better energy resolution. These factors, each expected to give a factor of few improvement, all contribute to our improved limit.

4.6 Discussion and conclusion

4.6.1 Future Developments

In this work, we obtain competitive limits on sterile neutrino dark matter decay by analyzing GBM data. This is the first time the GBM data is used for a dark matter search, and we have obtained the strongest constraint available in the mass range $25 - 50$ keV (Fig. 4.11). Although we focus on the implications for the sterile neutrino, our limits can be applied to all dark matter candidates that produce a mono-energetic photon in the keV range, such as moduli dark matter, gravitino dark matter, and other candidates [352, 353, 354]. It is straightforward to constrain the parameter space for the corresponding dark matter candidates using the limit on the decay rate from Fig. 4.10, taking into account any normalization or energy scaling.

Our analysis uses simple data reduction, minimal background assumptions, and straightforward analysis procedures. Thus there are many ways our results can be improved with further study.

Firstly, the dominant uncertainty on the data currently comes from the energy behavior of the effective area. In this work, we use a constant 5\% uncertainty as quoted by the GBM collaboration, which we then validate against a power-law assumption. The limit can be improved if this uncertainty can be better quantified or calibrated, e.g., using known
Figure 4.11: Constraints from X-ray missions on sterile neutrino dark matter decays, which depends on the mixing angle, $\sin^2(2\theta)$, and the mass, $m_s$. In the shown mass range, the best previous constraints are set by observations of cosmic X-ray background (CXB) from HEAO-1 [38] and the Milky Way (MW) halo from INTEGRAL [39]. The lower bound of the parameter space (black-hashed region) is valid for the model $\nu$MSM [12]. The flux (spectral) analysis limit derived from this work is shown in black-dashed (blue) line.
Secondly, most of the observed counts come from cosmic-ray related events. In this work we do not attempt to model such background in detail. In principle, this background can be better understood with simulations of cosmic-ray interactions with the satellite that take into account the satellite geometry and composition. Additionally, one can characterize the cosmic-ray induced background by satellite positioning, either by using high energy observations where the data is background dominated, or using the Earth’s magnetic field information. One can possibly construct a template for cosmic-ray induced events, which allows background reduction using spatial information. A significant reduction of the cosmic-ray background can further improve the dark matter limit for all energies.

The next source of backgrounds for dark matter searches is those arising from astrophysical origins, which dominate at low energies. This background can be modeled using high resolution X-ray sky maps from other missions. One can generate an astrophysical template for the GBM detectors which can then be used to subtract the astrophysical contribution from the data.

Importantly, if eventually the data are reduced to a regime where statistical uncertainties become important, it is important to treat systematic and statistical uncertainty simultaneously [354].

In principle, the analysis can be extended to higher energies. The GBM-NaI detectors are sensitive up to 1 MeV and the GBM-BGO detectors extend to 40 MeV, which is even higher than INTEGRAL and complementary to COMPTEL. However, at higher energies, the NaI detectors start to observe photons from backward directions due to either re-scattering or penetrating photons. The BGO detectors are designed to be sensitive to both front and back directions. A dedicated analysis taking into account this detector response for signal modeling is therefore necessary.

4.6.2 Astrophysical Implications

It is important to highlight that our novel use of GBM successfully detects the Galactic astrophysical component at 10–20 keV energies, as shown in Fig. 4.5 and 4.7. Due to the
broad point spread function of the NaI detector, this seemingly diffuse astrophysical component contains all points source emission near the GC, and possibly some diffuse emission. GBM observations may be used to impose interesting limits on the total Galactic astrophysical emission, which would extend the results from INTEGRAL [347, 348, 349] down to $\sim 10$ keV.

To constrain the astrophysical component, it is necessary to further reduce the detector background (see suggestions in the previous section). The analysis procedures would also need to be modified for a continuum spectrum. Such an analysis and a detailed interpretation of the astrophysical component are beyond the scope of this work.

4.6.3 Conclusion

We use data obtained by a GBM NaI detector (det-7), on board Fermi to set limits on dark matter decaying into mono-energetic photons. We first perform a conservative flux analysis, based on comparing the total flux normalization of the data and the model. We then perform a spectral analysis that assumes the total of non-dark matter contributions exhibits a power-law flux spectrum within the search window. Our spectral analysis is able to improve the limit by about an order of magnitude compared to previous searches using CXB with HEAO-1 data for line energies in the $10$–$25$ keV energy range, or $20$–$50$ keV sterile neutrino mass.

Conventionally used to detect and locate transients such as gamma-ray bursts, this is the first time that the GBM has been used to construct an all-sky map and search for dark matter emissions. After performing careful background reduction procedures, we are able to detect the astrophysical component centered on the GC.

Dark matter searches with GBM benefit from the large sky coverage and good energy resolution. Although the GBM does not have excellent angular resolution, this is not a severe problem for decaying dark matter for which the expected Milky Way signal is appreciably extended. A unique advantage of the GBM-NaI detectors is that they are sensitive to an energy range that is too high for X-ray telescopes such as Chandra, Suzaku, and XMM-Newton, but too low for INTEGRAL, therefore filling an energy gap that was last probed
by \emph{HEAO-1} in the late 1970s.

The current search sensitivity is dominated by systematic uncertainties in the effective area. A better understanding of GBM detector response through simulation or calibration, as well as a better understanding of detector and astrophysical backgrounds, can further improve the data quality and resulting limits substantially.
Chapter 5

Dark Matter Velocity Spectroscopy

Dark matter decays or annihilations that produce line-like spectra may be smoking-gun signals. However, even such distinctive signatures can be mimicked by astrophysical or instrumental causes. We show that velocity spectroscopy—the measurement of energy shifts induced by relative motion of source and observer—can separate these three causes with minimal theoretical uncertainties. The principal obstacle has been energy resolution, but upcoming experiments will have the precision needed. As an example, we show that the imminent Astro-H mission can use Milky Way observations to separate possible causes of the 3.5-keV line. We discuss other applications.

The contents of this chapter were published in [5].

5.1 Introduction

What is the dark matter? Identification depends upon more than just observation of its bulk gravitational effects; distinct particle signatures are needed. Backgrounds make it difficult to pick out these signals, which are constrained to be faint. Among possible decay or annihilation signals, those with sharp spectral features, such as a line, are especially valuable.

Given that the stakes and difficulties are so profound, even such a “smoking-gun” signal may not be conclusive. A line could have other causes: astrophysical (baryonic) emission or detector backgrounds (or response effects). For example, the cause of the recently discovered 3.5-keV line is disputed [62, 63, 331, 333, 335, 334, 332, 355, 356, 336, 357, 358, 359]. This
problem is more general [324, 328, 98, 154, 360, 151, 146, 361, 161, 362] and will surely arise again. We need better evidence than just a smoking gun—we need to see it in motion.

5.2 Premise and Motivation

We propose a general method for distinguishing the possible causes of a sharp spectral feature. Consider a line of unknown cause—dark matter (DM), astrophysical, or detector—observed in the Milky Way (MW). Relative motion between source and observer leads to distinctive energy shifts as a function of line of sight (LOS) direction. Figure 5.1 illustrates this schematically. Because typical Galactic virial velocities are $\sim 10^{-3}c$, the Doppler shifts are only $\sim 0.1\%$. Though exploiting such shifts is a standard astronomical technique (e.g., Refs. [363, 364, 365, 366, 367, 368]), this is the first demonstration of their power for testing DM signals.

A potential target for DM velocity spectroscopy is the 3.5-keV line recently observed in MW, M31, and galaxy cluster spectra [62, 63, 333]. The line energy and flux can naturally be explained by sterile neutrino DM [51, 53, 339, 296, 290, 291, 369, 292, 370, 371, 372] (or other candidates [373, 374, 375, 376, 377, 378, 379, 380, 381, 382]). However, the significance of the line is disputed [331, 335, 334], and it has been argued that it can be explained by astrophysical emission [332, 356].

With present detectors, velocity spectroscopy of this line is impossible. Excitingly, the Soft X-Ray Spectrometer (SXS) on Astro-H (launch date early 2016) has a goal energy resolution of $\sigma_{\text{AH}} = 1.7\text{eV (4 eV FWHM)}$ [383, 384], which is at the 0.1% scale. We show that if this goal resolution is achieved, together with a reasonable exposure, Astro-H can identify the cause of the 3.5-keV line through precise measurements of the centroid energies in different directions. (More generally, this could be done in detectors with worse energy resolution but better statistics.)

We emphasize that the applicability of DM velocity spectroscopy is much more general. The purpose of this paper is to introduce a new concept to increase the power of DM searches and to spur innovation in detector design. We conclude by discussing several
Figure 5.1: Top: How DM, astrophysical, and detector lines shift with Galactic longitude is starkly different. Bottom: For DM signals at positive longitude, our motion through the non-rotating DM halo yields a negative LOS velocity and thus a blue shift. In contrast, for astrophysical lines (e.g., from gas), co-rotation in the disk leads to a positive LOS velocity and thus a red shift. These signs reverse at negative longitude. Detector lines have zero shift.
5.3 Usual DM Decay Signal

The differential intensity (flux per solid angle) from DM with mass $m_\chi$ and lifetime $\tau = 1/\Gamma$, decaying within the MW, is

$$\frac{dI(\psi, E)}{dE} = \frac{\Gamma}{4\pi m_\chi R_\odot \rho_\odot} J(\psi) \frac{dN(E)}{dE},$$  \hspace{1cm} (5.1)

where $R_\odot \simeq 8$ kpc and $\rho_\odot \simeq 0.4$ GeV cm$^{-3}$ \cite{385, 386, 387} are the distance to the Galactic center (GC) and local DM density. (We neglect the cosmologically broadened extra-galactic signal, which contributes negligibly in Astro-H’s narrow energy bins.) $J(\psi)$ is the dimensionless, astrophysical J-factor defined by the LOS integral

$$J(\psi) \equiv \frac{1}{R_\odot \rho_\odot} \int ds \rho_\chi(r|s, \psi|),$$  \hspace{1cm} (5.2)

where $\psi$ is the angle relative to the GC and is related to Galactic longitude and latitude via $\cos \psi = \cos l \cos b$. $dN(E)/dE$ is the photon spectrum.

The above treatment assumes that the astrophysical term, $J(\psi)$, and the photon spectrum, $dN(E)/dE$, are separable. However, for detectors with energy resolution $\lesssim 0.1\%$, this approximation is not valid because relative velocities between source and observer, and therefore the spectral shape, vary along the LOS.

5.4 Modified DM Spectrum

We first account for how the signal is broadened by DM velocity dispersion and second for how it is shifted due to bulk relative motion.

We take the DM halo of the MW to be spherically symmetric, in steady state, and to have no appreciable rotation. The last is expected from angular momentum conservation, as the baryons from the proto-halo have collapsed significantly, while the DM has not; this is confirmed by simulations \cite{388, 389}. Thus, $\langle \vec{v}_\chi \rangle = 0$.

DM particles do have non-zero velocity dispersion, determined by the total gravitational
potential of the halo [390, 391]. Assuming an isotropic velocity distribution \((\sigma_{v,r} = \sigma_{v,\phi} = \sigma_{v,\theta})\), so the total dispersion is \(3\sigma_{v,r}\), the radial velocity dispersion of DM is [390]

\[
\sigma_{v,r}^2(r) = \frac{G}{\rho_\chi(r)} \int_r^{R_{vir}} dr' \rho_\chi(r') \frac{M_{\text{tot}}(r')}{r'^2},
\]

(5.3)

where \(M_{\text{tot}}(r)\) is the total mass within a radius \(r\). Typical values at \(r \sim \text{few kpc}\) are \(\sigma_{v,r} \simeq 125\ \text{km s}^{-1}\).

To calculate \(\sigma_{v,r}(r)\), we adopt the mass model of Ref. [392], which fits a contracted DM and three-component baryon mass profile to MW rotation curve data; for more details see Supplemental Materials. The choice of mass model is not critical; kinematic results from other models agree within \(O(10\%)\) [385, 393].

The spectrum from a point along the LOS is the convolution of the intrinsic spectrum with the DM velocity distribution at that point. We assume a Maxwellian velocity distribution throughout the halo, which, at each point, yields a Gaussian distribution of the LOS velocity component. The modified spectrum from each point is

\[
\frac{d\tilde{N}(E, r[s, \psi])}{dE} = \int dE' \frac{dN(E')}{dE'} G(E - E'; \sigma_{E'}),
\]

(5.4)

where \(G(E; \sigma_E)\) is a Gaussian of width \(\sigma_E = (E/c)\sigma_{v,\text{LOS}}\). Based upon observations of the LOS velocity distribution of MW halo stars reported in [394], we take \(\sigma_{v,\text{LOS}}(r) \simeq \sigma_{v,r}(r)\) which implies \(\sigma_E = (E/c)\sigma_{v,r}(r[s, \psi])\).

The line shift follows from the LOS velocity, \(v_{\text{LOS}} \equiv ((\bar{v}_\chi - \bar{v}_\odot) \cdot \hat{r}_{\text{LOS}})\), where positive \(v_{\text{LOS}}\) indicates receding motion. For \(v_{\text{LOS}} \ll c\), the resultant energy shift is \(\delta E_{\text{MW}}/E = -v_{\text{LOS}}/c\).

The Sun follows a roughly circular orbit about the GC in the direction toward positive Galactic longitude at a speed \(v_\odot \simeq 220\ \text{km s}^{-1}\) [395]. (Recent work suggests \(v_\odot \simeq 240\ \text{km s}^{-1}\) [396, 397], which would strengthen our results.) The spectrum is therefore shifted by \(\delta E_{\text{MW}}(l, b)/E = +(v_\odot/c)\sin l \cos b\), which changes sign with \(l\). We neglect the solar peculiar velocity as well as Earth and satellite motions, as the in-plane components of these are each \(\sim 10\ \text{km s}^{-1}\) [398, 399, 400]; even in combination, these are at most a \(\sim 10\%\) effect, and they enter with distinctive timescales.

The final expression for the modified spectrum, including broadening and shifts, is
therefore
\[
\frac{dJ}{dE} = \frac{1}{R_\odot \rho_\odot} \int ds \rho_\chi(r[s, \psi]) \frac{d\tilde{N}(E - \delta E_{\text{MW}}, r[s, \psi])}{dE},
\]
(5.5)
so that Eq. (5.1) is altered by \( J(\psi) \frac{dN}{dE} \rightarrow d\frac{J(\psi, E)}{dE} \). The observed signal, which is the convolution of \( dJ/dE \) with the detector response, is nearly Gaussian and has an effective width \( \sigma_{\text{eff}} \).

### 5.5 Modified Astrophysical Spectrum

The details are slightly different for astrophysical lines.

The widths of astrophysical lines are primarily determined by the mass of the emitting atom and by the gas temperature; turbulent broadening is negligible [401]. For potassium at \( T = 2 \text{keV} \), the intrinsic line width is \( \sigma_{\text{gas}} \simeq 0.8 \text{eV} \), comparable to Astro-H’s goal resolution, \( \sigma_{\text{AH}} \simeq 1.7 \text{eV} \). The intrinsic width is weakly sensitive to the gas temperature and mass \( (\propto \sqrt{T/m}) \); any reasonable values of \( T \) and \( m \) give similar results.

For the shift of an astrophysical signal, we must account for co-rotation within the MW disc. (While there is a non-rotating, gaseous halo at the outskirts of the MW, its temperature, \( \sim 0.1 \text{eV} \) [402, 403, 404, 335], is far below the multi-keV temperatures needed to produce significant line emission near 3.5 keV [62]). For simplicity, we assume all baryons follow circular orbits about the GC with speed \( v_{\text{circ}}(r) = \sqrt{GM_{\text{tot}}(r)/r} \). With this circular speed and the hot gas distribution of Ref. [405], we compute the spectral shift by integrating the signal along the LOS with the contribution from each point weighted by the gas density. We call this fiducial model G2.

Because the spatial and speed distributions of MW X-ray gas are uncertain, we compare to models in Ref. [367] with smaller and larger line shifts. G1 is based on the distribution of free e\(^-\) [406] and the MW rotation curve [407]. G3 is based on the observed distribution of \(^{26}\text{Al}\) gamma rays [367]. G1 and G2 are in good agreement with MW HI and CO data [363, 364]. Peak LOS velocities for G1, G2, and G3 are \( \simeq 50, 75, \) and \( 250 \text{km s}^{-1} \).
Figure 5.2: Comparison of received spectra for DM and gas (G2). The emitted spectra are taken to have equal flux and to be centered at 3.5 keV before velocity effects. The line profiles include velocity dispersion and shift effects, as well as the energy resolution of Astro-H. Vertical bands indicate the 1-σ centroid uncertainties after 2-Ms observations. For contrast, the brown line in the figure and inset shows the same signal if Astro-H had the energy resolution of XMM.
Figure 5.3: LOS velocity for DM and various gas models (the realistic version of Fig. 5.1). Uncertainties are computed assuming 2-Ms Astro-H exposures on each point.
5.6 Line Flux Detection

One prerequisite to detecting a spectral shift is that the number of signal events be non-zero. Another is that the background fluctuations be small in comparison. Though Astro-H has a small field of view (FOV), its excellent energy resolution strongly suppresses backgrounds for a line signal, so that even a small number of signal events can be significant.

Viewing directions $l \simeq 10^\circ - 40^\circ$ have advantages. First, the balance between decreasing signal flux and increasing energy shift at large $l$ is optimized. Second, theoretical uncertainties are minimized, as the DM density profile at $r \gtrsim$ few kpc is fixed by rotation curve data. Third, continuum astrophysical backgrounds are reduced; we reduce these further by going slightly off the Galactic plane, which minimally affects the DM signal.

The expected signal intensity is calculated from Eq. (5.1). For our DM example, this is

$$I(\psi) = 1.2 \times 10^{-8} \text{ cm}^{-2} \text{s}^{-1} \text{arcmin}^{-2} \times \left( \frac{\sin^2 2\theta}{7 \times 10^{-11}} \right) \left( \frac{m_\chi}{7 \text{ keV}} \right)^4 \left( \frac{\mathcal{J}(\psi)}{\mathcal{J}(l = 20^\circ, |b| = 5^\circ)} \right),$$

where we have integrated over energy in the line profile, calculated $\mathcal{J}(l = 20^\circ, |b| = 5^\circ) = 7.5$ using Ref. [392], and taken the DM parameters from Ref. [62]. For Astro-H, $\Omega_{\text{FOV}} = 9 \text{ arcmin}^2$ and (conservatively) $A_{\text{eff}} = 200 \text{ cm}^2$ [383, 384], so the expected number of events is

$$N_s(\psi) \simeq 43 \left( \frac{\mathcal{J}(\psi)}{\mathcal{J}(l = 20^\circ, |b| = 5^\circ)} \right) \left( \frac{t}{2 \text{ Ms}} \right).$$

This assumed exposure is large, but appropriate to the stakes (a potential discovery of DM) and the difficulties (the total exposure of XMM, Chandra, and Suzaku used in the 3.5-keV analyses is $\gtrsim 40 \text{ Ms}$ [62, 63, 333, 408, 320]). Furthermore, due to Astro-H’s excellent energy resolution, all pointings in a substantial fraction of the sky will help test the 3.5-keV line.

For continuum backgrounds, we consider only the contribution over the narrow energy range $\pm 2\sigma_{\text{eff}}$ centered at 3.5 keV. (We do not need to include the tails of nearby astrophysical lines, as they will be well-resolved, unlike in XMM.) One component of the background is due to the isotropic cosmic X-ray background (CXB) [409, 410, 411]. We conservatively adopt the total CXB flux (unresolved + resolved sources) $E d\Phi_{\text{CXB}}/dE = \ldots$
9.2 \times 10^{-7} (E/\text{keV})^{-0.4} \text{cm}^{-2} \text{s}^{-1} \text{arcmin}^{-2} \ [411]. Another background, due to hot gas in the MW, varies strongly with direction \ [412]. Finally, there are detector backgrounds due to intrinsic and induced radioactivities as well as cosmic-ray interactions; their intensity is expected to be comparable to that of the CXB \ [413]. For \(\psi(l = 20^\circ, |b| = 5^\circ)\), backgrounds contribute \(N_b \simeq 5.2 + 5.4 + 5.4 = 16\) events per 2 Ms within the \(\pm 2\sigma_{\text{eff}} \simeq \pm 4.8\) eV band centered at 3.5 keV, compared to \(N_s \simeq 41\).

We estimate the detection significance by the Poisson probability \(P(n \geq 57 | \mu = 16)\), which corresponds to a one-sided Gaussian probability \(> 7\sigma\).

### 5.7 Line Shift Detection

Detecting a line shift depends on how well the centroid of the line profile is determined. Backgrounds decrease the precision, but, as above, the energy resolution of Astro-H plays a critical role.

When backgrounds are absent, the uncertainty on the centroid is \(\sigma_{\text{eff}}/\sqrt{N_s}\). When they are present, the uncertainty becomes \(\delta E = C(R) \sigma_{\text{eff}}/\sqrt{N_s}\), where \(C(R)\) is a correction factor and \(R\) is defined by the background to signal ratio. We calculate the optimal \(C(R)\) using the Cramer-Rao theorem \[414, 415, 416\]. For \(\psi(l = 20^\circ, |b| = 5^\circ)\), \(C(R) \simeq 1.6\), so that the uncertainty in the LOS velocity is \(\delta v_{\text{LOS}} \simeq 50 \text{ km s}^{-1}\).

Figure 5.2 shows the line profiles at \(\psi(l = 20^\circ, |b| = 5^\circ)\) for a 3.5-keV emission line, due either to DM or gas. (A detector line would have zero shift). These profiles show how the energy spectra are shifted due to relative motion as well as broadened due to intrinsic dispersion and detector resolution. We show the uncertainties on the centroids, which are separated from each other and from zero in a 2-Ms exposure. With the energy resolution of XMM \[417\] (\(\sigma_{\text{XMM}} \simeq 47\) eV vs. \(\sigma_{\text{AH}} \simeq 1.7\) eV), the profiles are indistinguishable.

Figure 5.3 shows how the expected shifts vary with Galactic longitude, along with their uncertainties, assuming 2-Ms observations for each point. We show the DM signal uncertainties; for an astrophysical line of the same flux, the uncertainties are comparable because the effective widths are comparable (\(\sigma_{\text{eff}}^{\text{gas}} \simeq 160 \text{ km s}^{-1}, \sigma_{\text{eff}}^{\text{DM}} \simeq 200 \text{ km s}^{-1}\)); see Fig. 5.2.
For a detector line with zero intrinsic width, the effective width is $\sigma_{\text{eff}}^{\text{det}} \simeq 150 \text{ km s}^{-1}$, approximately a factor of $\sqrt{2}$ less than $\sigma_{\text{eff}}^{\text{DM}}$.

For each point in Fig. 5.3, it is easy to assess the probability that the expected DM signal could fluctuate to match that expected for an astrophysical or detector line, i.e., that a true DM signal could remain hidden. With two observations, at $l = \pm 20^\circ$, this scenario can be ruled out, relative to G2, at $\simeq 3.6\sigma$. This establishes that this technique has interesting sensitivity. Once there is data, one can assess the probability that an astrophysical or detector line could mimic a DM signal (for the same flux, $\delta_{\text{vLOS}}^{\text{gas}} \simeq \delta_{\text{vLOS}}^{\text{det}} \simeq \delta_{\text{vLOS}}^{\text{DM}}/\sqrt{2}$).

If the energy resolution is worse than the design goal, e.g., $\sigma_{\text{AH}} \simeq 2.1, 2.5, \text{ or } 3 \text{ eV}$, then the line shift significance is $\simeq 3.0, 2.4,$ or $1.9\sigma$ (the line flux significance is always $> 5\sigma$). This could be improved as $\sqrt{t}$ with more exposure (including non-dedicated pointings). We have not included the systematic uncertainty due to detector gain calibration, for which the goal is 0.4 eV [413]. This can be mitigated by comparing the energies of nearby astrophysical lines, especially at opposite longitudes.

### 5.8 Related Searches

Astro-H may be able to resolve the intrinsic width of a MW DM line. This would provide the first information on the large-scale DM velocity distribution, which is sensitive to DM particle properties [143] and to the presence of substructure [418, 419] (see Suppl. Mat.).

The 3.5-keV line has been detected in M31. Due to the relative motion between the Sun and M31, DM or astrophysical lines from M31 will have LOS shifts of $\simeq -300 \text{ km s}^{-1}$ [40]. Provided the line energy has been measured precisely in other data, we estimate that this blue shift could be detected with $> 5\sigma$ significance, making this an attractive way to test detector causes. Due to M31’s rotation, astrophysical lines are separated from DM lines by $\pm 200 \text{ km s}^{-1}$ around $\pm 1^\circ$, but, because the statistical uncertainties are large, they cannot be cleanly distinguished in 2 Ms; see Suppl. Mat. and Refs. [420, 421, 422, 423, 424, 425, 10]. The LMC [426] may also be an attractive target.

Cluster observations will also be important. For a single galaxy cluster, the cosmological
recession velocity can be measured with Astro-H using atomic line emission from the intracluster gas. The large internal velocities of the cluster will broaden astrophysical and DM lines (differently, which is a potential discriminator [62, 413]), making it somewhat more challenging to test line shifts (see above for how a larger line width can be compensated by larger statistics). For bright clusters, it may be possible to measure their velocities precisely, such that stacked observations increase the significance of the flux, as in Ref. [62], and also the velocity dependence of the centroid, allowing stronger tests of detector versus astrophysical or DM causes.

More speculatively, it may be possible to see the line in the extragalactic DM signal, if more astrophysical sources in the CXB are resolved, e.g., with eRosita [427, 428]. Furthermore, because we move at $\sim 400\,\text{km s}^{-1}$ with respect to the CMB, it may be possible to detect a dipole signature in DM line signal. Far-future observations may even detect a forest of sources in each LOS spectrum.

5.9 Conclusions

Even for a supposedly smoking-gun signal, such as a line, it may be difficult to distinguish between DM, astrophysical, or detector causes. We have shown that detectors with energy resolution $\lesssim 0.1\%$ can break this degeneracy using velocity spectroscopy, which has minimal theoretical uncertainties. We emphasize that our main goal is to point out this new and robust method for testing DM signals, which can be applied to any sharp feature, such as an edge or box [429, 430].

To demonstrate the potential of this technique, we have shown that Astro-H will be able to test the origin of the 3.5-keV line. In the future, other lines may be discovered. For lines at higher energy, the relative energy resolution of Astro-H improves. This unprecedented resolution will allow Astro-H to dramatically improve on existing sterile neutrino limits [308, 38, 313, 319, 310, 431, 327, 314, 39, 322, 12, 432, 433, 434, 316, 435, 4, 436, 437]. We encourage a dedicated study by the Astro-H Collaboration, once post-launch parameters are known, to give definitive answers on DM sensitivity over their full energy range.
We are encouraged by the expected 0.1% resolution of Astro-H in the range 0.3–12 keV, and the demonstrated 0.1% resolution of INTEGRAL-SPI in the range 20 keV to 8 MeV (including velocity spectroscopy of the 1.809-MeV line from $^{26}$Al [365, 366, 367]). Excitingly, the proposed X-ray mission ATHENA [438] and GeV gamma-ray mission HERD [439] have made achieving similar energy resolution a priority, which will improve existing limits [440, 441, 442, 74, 443, 75, 79, 444, 353, 2, 354, 445]. We encourage other missions to pursue this aggressively.
5.10 Supplemental Materials

5.10.1 Outline

We first briefly discuss the mass models and dispersion profiles used to derive the results presented in the main text. We then provide an expanded discussion of two additional applications of DM velocity spectroscopy, namely: probing the intrinsic DM dispersion profile using LOS observations and using velocity spectroscopy of M31 to test detector causes.

5.10.2 Radial Velocity Dispersion

To calculate the intrinsic broadening of a DM line, a galactic mass model must be adopted to determine the velocity dispersion profile. Below, we describe the mass models used in our analysis of the MW and M31.

Milky Way Mass Profile

We use model A1 of Ref. [392], which utilizes a DM halo determined by adiabatically contracting an initial NFW profile in the presence of baryons. We summarize key aspects of the model.

Before contraction, an NFW profile with scale radius $r_s = 21.5 \text{kpc}$ is assumed to coexist with three axisymmetric baryonic profiles roughly associated with the nucleus, bulge/bar, and disc of the galaxy. The total baryonic mass within a given radius is determined by the integration of the density profiles, with the addition of a central black hole of mass $m_{\text{BH}} = 2.6 \times 10^6 M_\odot$. The enclosed baryonic mass is

$$M_b(r) = m_{\text{BH}} + \int_0^r \int_{4\pi} dr' d\Omega \rho_b(r') r'^2.$$  \hspace{1cm} (5.8)

The final DM profile is determined by contracting the initial NFW profile in the presence of this baryonic mass distribution. The baryonic profiles are adiabatically contracted under the assumption that spherical shells of matter do not cross and that the DM particles follow circular orbits. This deepens the potential well and causes the DM to contract. Angular
momentum conservation then dictates the following equations:

\[
G [M_b(r_f) + M_{dm}(r_f)] r_f = GM_{halo}(r_i) r_i
\]

\[
M_{halo}(r_i) = M_{dm}(r_f) \frac{\Omega_b + \Omega_{dm}}{\Omega_{dm}},
\]

where \(M_{halo}(r_i)\) is the halo mass before contraction and \(\Omega_{dm}\) and \(\Omega_b\) are the dark and baryonic matter densities, taken to be in the ratio \(\Omega_{dm}/(\Omega_b + \Omega_{dm}) = 0.9\); more recent observations give \(\Omega_{dm}/(\Omega_b + \Omega_{dm}) = 0.84\) \[10\], which gives identical results.

These equations are solved numerically to give a final radius, \(r_f\), corresponding to a given initial radius, \(r_i\). The contracted profile has a normalization \(\rho_\chi(r = 8\, \text{kpc}) \simeq 0.4\, \text{GeV cm}^{-3}\).

The combined baryonic and contracted DM profiles are integrated to give the total mass enclosed within a given radius, \(M_{tot}(r)\):

\[
M_{tot}(r) = M_b(r) + M_{dm}(r),
\]

where \(M_{dm}(r)\) is the dark matter mass within a radius \(r\).

The velocity dispersion is determined by the potential well of the galaxy, which is, in general, non-spherical. We approximate the true mass distribution by the spherically averaged mass profile given above. This approximation has little impact outside of \(r \sim \text{few kpc}\) (where the DM becomes the dominant mass component), but greatly simplifies the calculation of the dispersion profile. Spherical symmetry allows for a simpler treatment of the Jeans equations \[390\] and, together with equilibrium and an isotropic velocity distribution, yields the expression for the radial velocity dispersion given in the main text.

This mass profile (Eq. 5.10) generates a rotation curve, \(v_{circ}(r) = \sqrt{GM_{tot}(r)/r}\), which is in good agreement with observations and a dispersion profile which agrees with results of previous papers \[392, 391\].

**M31 Mass Profile**

We use the mass model of Ref. \[424\]. Generalized Einasto profiles (given below) are used to describe the baryonic components
\[
\rho_b(a) = \rho_c \exp \left( -d_N \left( \left( \frac{a}{a_c} \right)^{1/N} - 1 \right) \right),
\]

(5.11)

with \( \rho_c, \ d_N, \ a_c, \) and \( N \) adjusted to match data. The baryonic mass model includes five components (nucleus, bulge, disc, young disc, and stellar halo). Together with the adopted NFW DM profile, the measured M31 rotation curve is reproduced well [424].

We also include a black hole of mass \( m_{BH} = 3.5 \times 10^7 M_\odot \) [392]; more recent observations suggest a slightly larger mass or \( 1.4 \times 10^8 M_\odot \) [422]. The inclusion of a central black hole yields larger velocity dispersions at small radii (\( \lesssim 10 \) pc), which increases the intrinsic width of DM lines arising from small angle LOS directions. However, because we focus on large angles (\( l \simeq 10^\circ - 40^\circ \) in the MW and \( \psi \simeq 0.5^\circ - 1.5^\circ \) in M31), we do not probe the region affected by the black hole, so its effect is negligible; we verified that our results were unmodified by this addition. Dispersions in M31 are comparable to those in the MW, but are systematically higher because of its larger mass and concentration.

Figure 5.4 shows the radial DM velocity dispersion profiles for the MW and M31. Vertical bands represent the range of radii that contribute 90% to the signal along \( \psi(l = 20^\circ, |b| = 5^\circ) \) in the MW and \( \psi = 1^\circ \) in M31.

### 5.10.3 LOS Velocity Dispersion

The velocity distribution of DM is of great interest both for the information it contains about the particle nature of DM and for its implications for direct and indirect detection experiments [400]. For example, models of self-interacting DM (SIDM) predict higher velocity dispersions near the centers of DM halos. By measuring the LOS velocity dispersion, it may be possible to constrain SIDM interaction cross-sections, particularly in clusters where deviations between SIDM and CDM dispersions are large [143, 425]. Additionally, because sub-halos generate smaller velocity dispersions, variations in line width along different LOS could help to constrain the size and distribution of DM substructure.

Because an observed DM signal will contain contributions from the entire LOS, and therefore a range of galactic radii, the full radial velocity dispersion cannot be probed directly. However, the observed LOS dispersion may still contain useful information. It is
Figure 5.4: Radial velocity dispersion profiles for the MW and M31. Shaded vertical bands indicate the range of radii that contribute 90% of the signal along $\psi(l = 20^\circ, |b| = 5^\circ)$ in the MW and $\psi = 1^\circ$ in M31; the radius ranges for the other directions discussed in the text are similar. Note that the lower bounds of these ranges are the smallest $r$ probed by these directions.
Figure 5.5: Intrinsic ($\sigma_{\text{DM}}$) and effective ($\sigma_{\text{eff}}$) LOS velocity dispersion profiles for the MW and M31 as a function of $\psi/\psi_S$, the scaled angle relative to the center of each system. For the MW, $\psi_S = 50^\circ$, while for M31, $\psi_S = 2.5^\circ$; these scalings were chosen for display purposes. Intrinsic widths are determined by integrating the spectrum along the LOS using the radial velocity dispersion profiles given in the previous section. Effective widths include detector energy resolution. The increase in the LOS dispersion at small angles in M31 is due to the rising radial dispersions shown in Fig. 5.4; for equally small (scaled) angles in the MW, the LOS dispersion decreases because only radii < 1 kpc, where the radial dispersion is decreasing, contribute.
Figure 5.6: LOS velocity profiles for DM and HI gas [40] in M31. DM error bars are calculated assuming 2-Ms exposures with Astro-H and only CXB and detector backgrounds.
natural to ask how well Astro-H may be able to reconstruct the intrinsic DM LOS dispersion, given the observed signal.

Figure 5.5 shows both the intrinsic and observed (assuming $\sigma_{\text{AH}} \simeq 1.7$ eV) LOS velocity width for a DM line in the MW and M31. Because the detector resolution is comparable to the intrinsic width, the detector response broadens the signal by a factor of $\simeq \sqrt{2}$.

In principle, if the energy resolution of Astro-H were known exactly, the intrinsic width of the DM line could be reconstructed precisely; assuming the signal and detector response are both Gaussian, the effective width is $\sigma_{\text{eff}}^2 = \sigma_{\text{AH}}^2 + \sigma_{\text{DM}}^2$, so that $\sigma_{\text{DM}}$ can be determined simply.

Of course, in practice, the resolution can never be known exactly. Assuming the goal uncertainty of 1 eV (the uncertainty is expected to be $\lesssim 2$ eV [413]), we estimate that the intrinsic width of a 3.5 keV DM line can be reconstructed with an uncertainty of $\simeq 40$ km s$^{-1}$; for higher energies the uncertainty in the width is smaller and scales as $E^{-1}$. See the Appendix of Ref. [413] for more details regarding uncertainty in detector energy resolution and intrinsic line width reconstruction.

More speculatively, using information about the strength of the signal along the LOS, it may be possible to construct a course-grained radial velocity dispersion profile from the LOS dispersion. For example, we see from the vertical bands in Fig. 5.4 that the range of radii that contributes to the $\psi(l = 20^\circ, |b| = 5^\circ)$ signal is narrow and that the dispersion of these points is directly reflected in the intrinsic LOS dispersion shown in Fig. 5.5. With additional pointings that probe different radii, it may be possible to constrain the radial dispersion profile using the measured line widths. This method would be most effective for small angles where the range of contributing radii is narrowest, although increased backgrounds would have to be overcome.

5.10.4 Velocity Spectroscopy of M31

DM velocity spectroscopy can also be applied to a signal observed from M31. Relative motion between the Sun and M31 produces a DM LOS velocity shift of $\simeq -300$ km s$^{-1}$ that is essentially independent of viewing angle. For astrophysical lines, one must also consider
the rotation of the M31 disc. This produces an additional LOS velocity shift that varies
strongly with viewing angle, separating the DM and astrophysical lines by \(\pm \simeq 200 \text{ km s}^{-1}\)
around \(\pm 1^\circ\) [40]. Detector lines are unshifted.

The large differences in LOS velocities between DM, astrophysical, and detector lines
make M31 a potentially powerful tool to probe the origin of spectral lines. However, large
LOS velocities are not by themselves sufficient to distinguish between these three causes; it
is also necessary that the uncertainty in the profile centroid be small in comparison to the
expected centroid separations.

As discussed in the main text, the uncertainty in the centroid is given by \(\delta E = C(R) \sigma_{\text{eff}}/\sqrt{N_s}\), where \(\sigma_{\text{eff}}\) is the observed line width, \(N_s\) is the number of signal events, and
\(C(R)\) is a correction factor that accounts for the presence of backgrounds. As can be seen in
Fig. 5.5, the observed widths of DM signals arising from M31 and the MW are expected to
be quite similar. However, the number of DM signal events in M31 is considerably smaller,
increasing the centroid uncertainty substantially.

Figure 5.6 shows the LOS velocities for DM, astrophysical and detector lines as a function
of the angular offset \(\psi\) from the center of M31. We show the error bars on a DM signal
assuming 2-Ms observations and only CXB and detector backgrounds. Astrophysical X-ray
emission in M31 is not well studied outside of \(\sim 0.5^\circ\), but is expected to be small [420,
421, 423]. However, even without including this background, it is clear that the significance
(\(\propto \sqrt{t}\)) with which DM and astrophysical signals can be differentiated is considerably smaller
than for the MW. However, the large differences between DM, astrophysical and detector
lines shifts could allow for cleaner separation of these causes, if uncertainties were reduced.
If MW observations of a line suggest a DM origin, several Ms would be well spent on M31
observations.

Perhaps the greatest utility of observing M31 is in its power to test detector causes of
a signal. This can be done most easily by looking directly at the center of M31. If the
line is DM or astrophysical in nature, the signal strength should be strong and the centroid
uncertainty correspondingly small, so that detector causes can be easily tested (Fig. 5.6).
Though we have shown error bars assuming 2-Ms observations, for this purpose, shorter
exposures will clearly suffice.
Chapter 6

First Observation of Time Variation in the Solar-Disk Gamma-Ray Flux with Fermi

The solar disk is a bright gamma-ray source. Surprisingly, its flux is about one order of magnitude higher than predicted. As a first step toward understanding the physical origin of this discrepancy, we perform a new analysis in 1–100 GeV using 6 years of public Fermi-LAT data. Compared to the previous analysis by the Fermi Collaboration, who analyzed 1.5 years of data and detected the solar disk in 0.1–10 GeV, we find two new and significant results: 1. In the 1–10 GeV flux (detected at $>5\sigma$), we discover a significant time variation that anticorrelates with solar activity. 2. We detect gamma rays in 10–30 GeV at $>5\sigma$, and in 30–100 GeV at $>2\sigma$. The time variation strongly indicates that solar-disk gamma rays are induced by cosmic rays and that solar atmospheric magnetic fields play an important role. Our results provide essential clues for understanding the underlying gamma-ray production processes, which may allow new probes of solar atmospheric magnetic fields, cosmic rays in the solar system, and possible new physics. Finally, we show that the Sun is a promising new target for ground-based TeV gamma-ray telescopes such as HAWC and LHAASO.

The contents of this chapter were published in [6].

6.1 Introduction

The Sun is well studied and understood with a broad set of messengers at different energies. For example, the optical photon and MeV neutrino spectra confirm a detailed picture of the Sun as a middle-aged G-type main-sequence star powered by nuclear fusion [446, 447]. However, the gamma-ray emission from the Sun is poorly understood. Precision studies of
the Sun at GeV energies are only now possible after the 2008 launch of the Fermi Gamma-Ray Space Telescope (Fermi).

Naively, one does not expect the quiet Sun (also known as the steady or the quiescent Sun) to produce an appreciable GeV gamma-ray flux. Even though the solar atmospheric temperature rises to millions of Kelvin in the corona, it corresponds to $\lesssim$ keV in energy. And, although solar flares can accelerate particles non-thermally, bright flares are rare and the highest-energy gamma ray observed from a flare is only $\simeq 4$ GeV [448, 449, 450, 451].

There are, however, two distinct processes involving cosmic rays that guarantee the continuous production of gamma rays from the vicinity of the Sun. The first contribution comes from the Inverse-Compton (IC) scattering of cosmic-ray electrons and positrons with solar photons [452, 453, 454]. The IC component appears as an extended halo ($\sim \mathcal{O}(10^\circ)$) around the Sun. The second contribution comes from the hadronic interaction of cosmic rays with the solar atmosphere (photosphere and chromosphere) [64]. The extent of this component has the angular size of the Sun ($\simeq 0.5^\circ$); we denote it (plus any potential non-cosmic-ray contribution) as the solar-disk component.

Theoretical estimation of both components requires taking into account the effects of solar magnetic activity. Magnetic fields carried by the solar wind modulate the fluxes of cosmic-ray particles in the solar system [455, 456, 457]. This effect is expected to be stronger for the solar-disk component than the IC component because of the much closer approach to the Sun for the parent cosmic rays. In addition, magnetic fields in the solar atmosphere [458, 459, 460] affect the solar-disk component. Seckel et al. [64] (denoted as SSG1991 in the following) showed that solar atmospheric magnetic fields could boost gamma-ray production through the magnetic reflection of the primary cosmic rays or their showers out of the Sun. Consequently, they estimated that the Sun could be detected by space-based gamma-ray telescopes.

The first experiment to have the sensitivity to detect quiet Sun gamma rays was the Energetic Gamma Ray Experiment Telescope (EGRET) [461]. A reanalysis of the EGRET data later reported the first detection of solar-disk gamma rays, but the flux uncertainties were large [462]. More recently, with the improved sensitivity of the Large Area Tele-
scope (LAT) on board Fermi, the IC and solar-disk components were each well measured at 0.1–10 GeV in Abdo et al. [463] (denoted as Fermi2011 in the following). The IC component was detected out to 20° from the Sun, and was found to be consistent with theoretical expectations [452, 453, 454]. Although the observed solar-disk component satisfies the theoretical upper bound derived in SSG1991 (the naive case), it is in complete disagreement with the nominal model of SSG1991, the one and only theoretical prediction: The observed flux is about one order of magnitude higher at all energies and the spectrum shape is flatter than predicted. This mismatch motivates new theoretical modeling and new observational studies of the solar-disk gamma rays. The latter is the focus of this study.

After Fermi2011, two key questions naturally surfaced concerning solar-disk gamma rays. First, does the solar-disk gamma-ray flux have a long-duration time variation? In Fermi2011, after comparing to the results from Ref. [462], it was pointed out that a significant variation of the solar-disk emission may be present. If such a variation is confirmed, and if it is related to the solar activity cycle, it could test the cosmic-ray origin of the gamma rays and help reveal their production mechanism. Second, does the Sun shine in gamma rays beyond 10 GeV? The last two data points from the Fermi2011 solar-disk energy spectrum suggest the spectrum might become softer at higher energy. Interactions of cosmic rays with solar magnetic fields are energy dependent; a spectral cutoff at high energy could reveal the end of magnetic field effects on the cosmic-ray interactions. It is only possible to answer these questions now because of the improved statistics and long time baseline (> 6 years) of the Fermi-LAT data set.

We aim to address these questions in this work, which is structured as follows: In Sec. 6.2, we present our analysis and findings. In Sec. III, we first provide a short overview of the hadronic solar gamma-ray production by cosmic rays. Then we discuss future prospects for both theory and observation. Seasoned readers on cosmic-ray theory can skip the overview (Sec. III A) and move on to the rest of the section. We conclude in Sec. 6.4.
6.2 The Sun observed using Fermi-LAT

6.2.1 Outline of the Analysis

Launched in 2008 on board Fermi, the LAT instrument is a pair-conversion gamma-ray detector sensitive to energies from about $10^{-2}$ GeV to $10^3$ GeV [464, 465]. Its large field of view allows it to survey the whole sky. With 1.5 years of data, Fermi2011 detected the solar-disk and IC components separately in 0.1–10 GeV. Since then, Fermi not only collected more data, but its quality has also improved. Fermi data are publicly available, which allows us to perform this study.

Due to the apparent motion of the Sun on the sky, one needs to trace its position continuously with time to produce a Sun-centered image. Because we focus on the solar-disk component, all other sources of emission are treated as backgrounds. There are two main backgrounds that need to be accounted for; both are small compared to the signal. The first is the diffuse background that consists of astrophysical emission (smeared due to the motion of the Sun) and the detector background. The second background (technically, a foreground) is the IC component in the line of sight. Both backgrounds can be estimated from the data.

We follow Fermi2011 by removing data near the Galactic plane and model the diffuse background using the fake-Sun method. In addition, we remove all the bright solar flares. To increase photon statistics, we relax the point-source cut and moon cut used Fermi2011. We study and take into account the possible systematics associated with this step.

To extract the solar-disk signal, we perform a likelihood analysis with the data binned in both energy and angle. This allows us to perform a simple and conservative analysis to characterize the main features of the signal. The accuracy goal of this analysis is limited by the systematic uncertainty of Fermi-LAT’s effective area, which is estimated to be about 10% [465], so we ignore uncertainties that are much less than that. We discuss possible ways to improve the analysis in Sec. 6.3.
6.2.2 Data Selection and Cuts

We choose our analysis energy range to be 1–100 GeV. Below 1 GeV, the point spread function (PSF) of Fermi-LAT deteriorates rapidly, making it difficult to isolate the solar-disk component (in addition, the Fermi Collaboration is performing a dedicated analysis at low energies [466]). Above 100 GeV, although we find 3 photons (up to \(\sim 300\) GeV) within 1° of the center of the Sun in the final photon map, it is difficult to estimate the background contribution due to the small number of photons.

We analyze the data using the Fermi science tools version v9r33p0\(^2\). We use the weekly P7REP data set from week 010 to week 321, which covers from 2008–08–07 to 2014–07–31. (Pass 8 data became available during the final stages of this work; we discuss this in Sec. 6.3.) To trace the Sun’s position, we divide each week into 40 identical time segments. Because the Sun moves \(\simeq 7°\) per week, its positional drift per time segment is \(\simeq 0.2°\). This is smaller than the diameter of the Sun (\(\simeq 0.5°\)) and the LAT PSF at 1 GeV (\(\simeq 1°\)). Above 10 GeV, the drift becomes comparable to the PSF (\(\simeq 0.1°\)), which we mitigate by using large angular bins in the likelihood analysis.

For each time segment, we adopt the standard data selection procedure recommended by the collaboration. We use \texttt{gtselect} to select photons from the \texttt{SOURCE} event class and to divide the events into eight energy bins of equal logarithmic width. We set the maximum zenith angle to be 100° to avoid photons coming from the luminous Earth limb [467, 468]. We select all photons within 10° of the Sun; to avoid potential edge effects, we define our region of interest (ROI) as a 9°-radius circle. The photon events are filtered using \texttt{gtmktime} with the keywords \texttt{DATA\_QUAL==1}, \texttt{LAT\_CONFIG==1}, and \texttt{ABS(ROCK\_ANGLE<52)}. The first two keywords ensure that the data quality is good enough for a point-source analysis; the last one requires that the spacecraft be within the range of rocking angles used during nominal sky-survey observations. The filtered photon events are binned into photon counts maps in equatorial coordinates using \texttt{gtbin} with a pixel size 0.1° × 0.1°. The photon maps are stacked to construct a single map for each energy bin.

\(^2\)http://fermi.gsfc.nasa.gov/ssc/data/analysis/software/
Figure 6.1: **Left:** Stacked photon counts map of the Sun ROI in 10–100 GeV. **Right:** Same, but for a fake-Sun ROI (in this example, trailing the Sun in its path by +180 days), which is used to measure the diffuse background. The exposures of the two ROIs differ by $\lesssim 2\%$. (Maps for $> 0.1$ GeV are shown in Fermi2011.) Visually, the solar-disk component (comparable in extent to the size of the Sun, as marked) is obvious; that of the IC component (decreasing with angle) is more subtle. The numbers of photons within 1.5° of the center are 175 versus 19; the numbers in 1.5°–9° are 844 versus 710.
To calculate the expected number of photons from an underlying intensity (flux per solid angle) distribution, we obtain the exposure map using `gt1tcube` and `gtexpcube2` with identical settings as for the photon maps, and using the `P7REP_SOURCE_V15` instrumental response function. The flux map is obtained by dividing the stacked photon map by the stacked exposure map. The total exposure in the ROI is about \( \simeq 10^{11} \text{ cm}^2 \text{s} \), and is spatially uniform at the \(~ 1\%\) level in 1–100 GeV.

To check our data selection procedures, we measure the gamma-ray flux from one of Fermi’s calibration sources, the Vela pulsar, which is the brightest steady astrophysical gamma-ray source above 0.1 GeV. We repeat the same data selection procedures, except for the time segments used to trace the Sun, to obtain the photon map and exposure map. The gamma-ray flux is estimated from the total flux within 1.5\(^\circ\) of Vela, after subtracting the background estimated from the 6\(^\circ\)–9\(^\circ\) region of the same ROI. The flux obtained is consistent with that in Ref. [469].

![Figure 6.2: Total gamma-ray flux between 1 GeV and 1.8 GeV within 1.5\(^\circ\) from the Sun versus time. Each bin corresponds to one week of observation, starting from 2008-08-07 (week 010). Periods that coincide with a bright solar flare are labeled with red squares; these are removed from the analyses. The horizontal grey band shows the resulting 6-year combined flux and its uncertainty.](image)

Following Fermi2011, we remove data when \(|b| < 30^\circ\), where \(b\) is the Galactic latitude. This avoids the bright diffuse and point-source emission from the Galactic plane. After this cut, the exposure time is reduced by \(\simeq 40\%\) and the total photons by \(\simeq 76\%\), consistent
with the values in Fermi2011. This cut is efficient for reducing background contamination, but is conservative because the Galactic plane emission decreases rapidly with Galactic latitude. We discuss in detail the remaining background components in Sec. 6.2.3.

In Fermi2011, data are excluded whenever a known point source or the Moon is within 20° of the Sun. In order to maximize the photon counts in high energy, we relax these cuts. Point sources are expected to increase the diffuse background by about 10%, which has minimal effect to our solar-disk-centric analysis. The Moon should not affect our analysis because its energy spectrum falls rapidly above 1 GeV [470]. We describe in the next section in detail how we handle the inclusion of background sources in the likelihood analysis. Imposing the point-source cut would reduce the exposure time by at least a factor of 3 (shown in Fermi2011 with 1FGL), making the high-energy analysis significantly more difficult. (The IC component has a smaller signal-to-noise ratio. As a result, the point-source cut is more important for an IC-centric analysis, as in Fermi2011.)

With the goal of searching for time variations in the solar-disk flux, we pay special attention to possible time-varying sources. The most important ones are solar flares [448, 449, 450]. During the period of bright solar flares, the flaring regions can emit a significant flux of gamma rays for a short period of time, thus contaminating the solar-disk signal and potentially changing the time profile of solar-disk flux. Only a few flares are expected to matter, as solar flares are typically dim beyond a few GeV. Another special source is the blazar 3C 279, which overlaps the coordinates of the Sun every October [471]. This blazar has a flux comparable to that of the Sun and the Sun stays about a day near its location, hence it would nominally contaminate the solar-disk component at the ∼1% level. However, when it is in a flaring state, it can temporarily be 100 times brighter [472, 473]. We check and find that the Sun was never nearby during the reported 3C 279 flares.

Figure 6.1 (left) shows the stacked photon map in 10–100 GeV. It is clear from the density and the brightness of the pixels the solar disk is observed. This is the first time that the Sun has been detected with > 10 GeV photons. Compared to the map shown in Fermi2011, which was for all photons above 0.1 GeV and is thus dominated by low-energy photons, this image is sharper due to the improved PSF at higher energies. The right panel
of Fig. 6.1 is a “fake-Sun” photon map, used as a background estimate, described in the next subsection.

Figure 6.2 shows the gamma-ray flux (1–1.8 GeV) within 1.5° of the Sun as a function of time. We label the time periods that contain solar flares detected by Fermi at greater than 10σ. Some anomalously bright periods are correlated with solar flares, most notably the ones in 7 March 2012 and 25 February 2014 (week 196 and 299). Beyond that, we do not observe any obvious excesses. For consistency, all labeled periods are removed from the Sun and fake-Sun analyses.

### 6.2.3 Background Estimation

#### Diffuse Background

Due to the motion of the Sun on the sky, all astrophysical emission is smeared to a diffuse and isotropic background. This includes truly diffuse as well as resolved and unresolved point-source emission. We denote this emission together with the detector background (misidentified cosmic rays) as the diffuse background.

We estimate the expected contribution of the diffuse background in the Sun ROI using the fake-Sun method described in Fermi2011. We repeat identical analyses (including all cuts) at positions where the Sun would have been +60, +90, +180, and −90 days away from the actual time. The fake Suns traverse the same paths through the sky as the Sun, which allows us to measure the diffuse background independently.

Figure 6.1 (right) shows the stacked photon map in 10–100 GeV for one of the fake Suns (+180 days). The Sun and fake-Sun ROIs have comparable exposures (≤ 2% difference). As a result, the small excess of photons away from the center of the Sun ROI already shows hints of the extended IC component, which becomes apparent when the angular distribution of the intensity is shown.

The combination of four fake Suns allows us to estimate the diffuse background with better than 10% statistical uncertainty. However, when comparing the individual fake-

---

3http://hesperia.gsfc.nasa.gov/fermi/lat/qlook/lat_events.txt
Sun background estimates, we observe, at the low end of our energy range, \(\simeq 10\%\) variations among the fake Suns, which is larger than their individual statistical uncertainties. Upon closer inspection, we found that this is driven by one particularly brighter fake-Sun ROI (+180), while the other three agree with each other at subpercent level. We check and do not find any significantly bright periods in this fake-Sun ROI. Therefore, this flux enhancement is likely due to one or several mild time-varying background sources, an arguably expected consequence of including point sources in the data set. We combine the four fake-Sun ROIs to estimate the diffuse background, and mitigate the potential background variation by adding a 10% systematic uncertainty to the diffuse background in the likelihood analysis. We also check our result using the background estimates without the +180 fake-Sun ROI. The difference is miniscule.

We compare our combined fake-Sun background estimate with that from Fermi2011, and find that our background estimate is higher by \(~10\%\) at the low energy end. Though this is consistent with systematic variation described above, it could also be explained by background sources. The average point-source contribution to the diffuse background can be estimated using the total high-latitude (\(|b| > 20^\circ\)) point-source intensity reported in the Fermi Isotropic Gamma-Ray Background analysis (see Fig. 8 in Ref. [474]). Comparing this to the diffuse background in the fake-Sun ROI (Fig. 3 in Fermi2011), point sources contribute about 10% of the total diffuse background, which matches the difference seen in our fake-Sun analysis versus that in Fermi2011. Because this extra small contribution affects both the Sun and fake-Sun ROIs, it is self-consistently modeled in the likelihood analysis. Nonetheless, we add an additional 10% systematic uncertainty to the diffuse background in the likelihood analysis. These systematic uncertainties (20% of our fake-Sun estimate) take into account all the potential systematics introduced to the diffuse background by including the point sources.

Lastly, the gamma-ray intensity of the fake-Sun ROIs are found to the uniform in radial direction. This is consistent with the finding from Fermi2011, which showed that the only source of anisotropy is the Galactic plane, which we have removed. This angular dependence allows us to separate the diffuse background from the signal and the IC component.
Figure 6.3: **Left:** Angular distribution of the integrated intensity from 1–10 GeV in the Sun ROI. Black points show the observed data with statistical uncertainties only. Colored histograms show the fitted results for the signal and two backgrounds (the estimate of the diffuse background incorporates independent data from the fake-Sun ROIs). The inset shows the same data with smaller angular bins, but without the two solar components (note the different vertical scale). **Right:** Same, but for 10–100 GeV (note the lower flux).
Inverse Compton Emission

In addition to the diffuse background, the extended IC component also contributes to the total emission in the Sun ROI. We model the IC component background using its distinctive angular distribution. Assuming the cosmic-ray electron density is homogeneous throughout the solar system, the IC component intensity is simply proportional to the column density of solar optical photons [452, 453, 454]. This description was found to be reasonable in Fermi2011, especially for gamma-ray energies above 1 GeV. With this assumption, we can approximate the IC intensity as \( \propto \alpha^{-1} \), where \( \alpha \) is the angular distance from the Sun. This distribution deviates from the true one [452, 453, 454] slightly at large angles, and is accurate at the \( \sim 5\% \) level at the edge of our ROI. In the angular region of the solar disk, the IC component is suppressed, which we take into account in the analysis (described below). Overall, small uncertainties of the shape of the IC component do not affect our results, as it is subdominant compared to the solar-disk emission in the inner 1.5°.

6.2.4 Solar-Disk Flux Spectrum

We use a multicomponent fit to extract the solar-disk component. This exploits the facts that the Sun is spatially concentrated (see Sec. 6.3.4 for discussion on resolving the Sun), the IC component is extended with a characteristic profile, and the diffuse background is spatially uniform. The angular information allows us to fit the components individually for each energy bin, without requiring any assumptions about the energy spectra.

We divide the Sun ROI into angular bins that are concentric rings of 1.5° width. This choice is guided by the PSF of Fermi-LAT, which is 0.8° at 1 GeV (68%). Because the PSF improves above 1 GeV and flattens out by \( \sim 10 \) GeV, the 1.5° bin ensures that the solar-disk component is always fully contained in the first angular bin. This criterion significantly simplifies the analysis. Moreover, our choice of the uniform 1.5° angular bin across all energies is conservative. The PSF of Fermi-LAT improves at high energies, so in principle one can afford a smaller angular bin at higher energy bins. However, we expect the improvement from such an analysis will be small, given that the diffuse background is
Figure 6.4: Energy spectrum of the solar-disk flux. Blue squares and statistical uncertainties (systematic uncertainties, not shown, are \( \approx 10\% \)) are the results of our analysis with 6 years of data. Black dots and combined statistical and systematic uncertainties are the Fermi2011 results with 1.5 years of data. The green band shows the predicted flux range from the SSG1991 nominal model.
Figure 6.5: Energy spectrum of the solar-disk flux, separated into three periods, each of two years. The solar disk flux from first two years is consistent with Fermi2011, while the 1–10 GeV data shows a significant reduction in later periods.
small. For simplicity, we use constant angular bins across all energy range.

With this angular binning, the distribution of the gamma-ray flux in the Sun ROI is modeled independently for each energy bin, as follows:

\[
\begin{align*}
  s_i &= s_1 \delta_{i1} \\
  b_{i\text{IC}} &= f_{IC} \sum_j E_{i,j} \alpha_{i,j}^{-1} \\
  b_{i\text{BKG}} &= f_{BKG} \sum_j E_{i,j}
\end{align*}
\]

where \( s_i, b_{i\text{IC}}, \) and \( b_{i\text{BKG}} \) are the modeled photon counts for the solar-disk signal, as well as the IC and diffuse backgrounds in angular bin \( i \). \( E_{i,j} \) is the exposure for a pixel \( j \) in bin \( i \) (with unit \([\text{cm}^2\text{s}\text{sr}]\)), and \( \alpha_{i,j} \) is the angular distance from the center to a pixel \( j \) in bin \( i \). The solar-disk component is described by a Kronecker delta function, \( \delta_{i1} \), which indicates that the first angular bin fully contains the solar-disk flux. The IC component is described by a normalization factor, \( f_{IC} \), times the total exposure weighted by \( \alpha^{-1} \). At the region of the solar disk (\( \alpha < 0.27^\circ \)), the IC component is strongly suppressed due to the anisotropy of the solar radiation and the occultation of the Sun [452, 453, 454]; we set the IC component to be zero in this region accordingly. The diffuse background component is radially isotropic, so it is only a normalization factor, \( f_{BKG} \), weighted by the total exposure.

For each energy bin, we perform a profile likelihood analysis [350, 475]. The likelihood function is a function of the signal parameter, \( s_1 \), and the nuisance parameters, \( f_{IC} \) and \( f_{BKG} \):  

\[
\mathcal{L}(s_1; f_{IC}, f_{BKG}) = G(f_{BKG}) \prod_i P(s_i + b_{i\text{IC}} + b_{i\text{BKG}} | d_i),
\]

where \( P \) is the Poisson probability for the model to yield the observed number of photons, \( d_i \). The product is taken over all angular bins. The Gaussian term, \( G(f_{BKG}) \), constrains the diffuse background from deviating too much from the value determined from the fake-Sun method. We take the variance of the Gaussian to be 20% of the combined fake-Sun flux estimate, and assume that it is uncorrelated between energy bins. The 20% systematic uncertainty conservatively combines the 10% variations among the individual fake Suns and the 10% difference we observe from our fake-Sun method compared to that from Fermi2011.
The best-fit diffuse background normalization in the Sun ROI is found to be within 10% of our fake-Sun estimate for all energy bins, which shows that the fake-Sun estimate is accurate and the choice of 20% variance for $G(f^{BKG})$ is conservative. The normalization of the IC component is conservatively set as a nuisance parameter. The final uncertainty of the extracted solar-disk component therefore includes the maximum normalization uncertainty of the IC component.

Figure 6.3 shows the angular distribution of the intensity in coarse energy bands, given by the number of photons in each angular bin divided by the total exposure. The data points represent the total observed intensity with statistical error bars only, and the colored histograms represent the fit for the three individual components. This simple model describes all features of the data well, and it is evident that the solar-disk component has a high signal-to-noise ratio.

For each energy bin, we obtain the best-fit model parameters by maximizing the likelihood function with respect to all model parameters. The uncertainty of the extracted solar-disk signal is found using the profile likelihood function, which is the likelihood function maximized over only the nuisance parameters. Assuming the signal parameter is Gaussian-distributed, the 1-$\sigma$ error bar of the signal is determined by where the log-profile likelihood function differs from the best-fit value by 1/2. This uncertainty determination procedure is exact when the sample size is large, but is found to be reasonable for fairly small sample sizes [475]. We check explicitly that the log-profile likelihood function behaves close to the expected parabolic shape, which verifies the Gaussian-distribution assumption. In addition to the uncertainties estimated above, the gamma-ray flux has an overall 10% systematic uncertainty from the effective area of the Fermi-LAT.

We check our result using the same 1.5-year time period as in Fermi2011. We find that our solar-disk component is consistent with that of Fermi2011, despite using different data sets (Pass 6 vs Pass 7), different energy and angular binning, different cuts, and a different analysis method. This supports our analysis choices.

For the full 6-year data set, we obtain a non-zero solar-disk signal in all eight energy bins from the likelihood analysis. The detection significance can be estimated from the
Table 6.1: For each energy bin, as defined, the total number of photons within 1.5° of the center of the Sun, the rounded best-fit number of photons due to the solar-disk signal, and the significance (\(\sqrt{\text{TS}}\)) of the solar-disk flux detection.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0–1.8</td>
<td>1468</td>
<td>961</td>
<td>20.5</td>
</tr>
<tr>
<td>1.8–3.2</td>
<td>914</td>
<td>628</td>
<td>17.7</td>
</tr>
<tr>
<td>3.2–5.6</td>
<td>448</td>
<td>329</td>
<td>13.6</td>
</tr>
<tr>
<td>5.6–10</td>
<td>188</td>
<td>133</td>
<td>8.5</td>
</tr>
<tr>
<td>10–18</td>
<td>92</td>
<td>67</td>
<td>6.7</td>
</tr>
<tr>
<td>18–32</td>
<td>55</td>
<td>42</td>
<td>5.9</td>
</tr>
<tr>
<td>32–56</td>
<td>16</td>
<td>10</td>
<td>2.6</td>
</tr>
<tr>
<td>56–100</td>
<td>12</td>
<td>7</td>
<td>2.3</td>
</tr>
</tbody>
</table>

test statistic (TS ≡ 2\(\Delta\log L\)), given by two times the difference between the best fit log-profile likelihood function and the one with the null hypothesis (\(s_1 = 0\)). The Gaussian significance, to good approximation, is given by \(\sqrt{\text{TS}}\) \([475]\). As a cross check, we obtain comparable best-fit parameters and uncertainties using a simple \(\chi^2\) and \(\Delta\chi^2\) analysis.

Table 6.1 summarizes our results, listing the energy bins, the total photon counts, and the best-fit numbers of photons in the solar-disk component, and \(\sqrt{\text{TS}}\). We find that the solar-disk component is significantly detected (\(> 5\sigma\)) up to \(\simeq 30\) GeV, and is detected (\(> 2\sigma\)) in each of the last two energy bins that go up to 100 GeV. The lower detection significance at \(> 30\) GeV is mainly due to not having enough statistics to distinguish the IC and solar-disk components. We discuss the total solar gamma-ray flux more in Sec. 6.3.5.

Figure 6.4 shows the energy spectrum of the solar-disk component obtained in our 6-year analysis with 1-\(\sigma\) error bars. The spectrum extends to 100 GeV without an obvious spectrum cutoff, though for energies \(\gtrsim 30\) GeV, the error bars are large. The spectrum can be roughly described as a single power law, \(\propto E^{-2.3}\), though the power-law fit is not particularly good. For comparison, we also show the solar-disk component found in Fermi2011 and the SSG1991 nominal model prediction on the solar-disk flux, where in the former the error bars include systematics. Comparing our result to that of Fermi2011, our analysis yields a similar spectrum with a lower normalization in the overlapping energy range. We find that this is because the underlying flux has a significant time variation, as
Figure 6.6: **Left:** For several sources, the ratio of the 1–10 GeV flux in each year to its 6-year average (a time-independent source would fluctuate around unity). The solar-disk component (blue squares) demonstrates a clear decreasing trend and anticorrelation with the smoothed sunspot number, a tracer of solar activity. Other sources (points displaced for clarity) should be and are consistent with being time-independent; see the text for details. **Right:** Same, but for 10–100 GeV. No obvious trend is observed for the disk component, but the uncertainties are large.
detailed in the next subsection.

Compared to the central value of the SSG1991 prediction, our 1–10 GeV result is still higher by a factor of about 5. The flux normalization of the solar-disk gamma-ray flux remains an unsolved puzzle. To provide more context on the physical implications of this disagreement, we discuss and provide more details about the SSG1991 model in Sec. 6.3.1.

6.2.5 Time Variation of the Solar-Disk Flux

Figure 6.5 shows the solar-disk gamma-ray flux energy spectrum obtained from our analysis when we divide the whole data set into two-year segments (52 weeks per “year”). In 1–10 GeV, a decreasing trend in flux is clearly observed. Above 10 GeV, the situation is unclear, due to the large error bars. The time modulation of flux above GeV is already hinted at in Fig. 6.2, where the 1–1.8 GeV data showed a slow decline over the course of 6 years.

To better quantify the time variation observed in 1–10 GeV, we first combine the data into two broad energy bands (1–10 and 10–100 GeV), and then find the flux ratio for each energy band, which is the integrated flux in each year relative to that averaged over 6 years. A time-independent source would fluctuate around unity.

Figure 6.6 shows the flux ratios in these two energy bands. In the 1–10 GeV band, the solar-disk component demonstrates a significant time variation, an overall decreasing trend, in which the extremes differ by about a factor of 2 to 3. We estimate the statistical significance of the time variation by testing the data against the null hypothesis (the underlying distribution is time independent) using a simple $\chi^2$-square test. The $\chi^2$-squares are 104 and 1.6 for 5 degrees of freedom for the 1–10 GeV 10–100 GeV band, respectively. This shows that the time-variation in the 1–10 GeV band is highly significant, while the 10–100 GeV data is consistent with being time independent.

We note that in Fig. 6.5, it can be seen that our 1–2 year result is compatible with the Fermi2011 spectrum in the overlapping energy range. Given that flux only slightly decreases from the first year to the second year, this shows that our analysis with 18-months of data is compatible with that of Fermi2011.
To make sure that the observed time variation is physical, we check the flux ratios of several gamma-ray sources as control samples. First, we consider one of the fake Suns (+180 days). We find the total gamma-ray flux within 1.5° of its center, as in our solar-disk analysis. This allows us to investigate possible fluctuations of the diffuse background. For both energy bands, we find that they are consistent with being time independent. Similar results are obtained when other fake Suns are used.

Second, we consider the gamma-ray flux from the Vela pulsar (a constant gamma-ray source), which we use to validate our data selection procedure in Sec. 6.2.2. This allows us to check for unknown systematics in data selection. The flux ratios of Vela demonstrate very small deviations from unity in both energy bands.

Third, we consider the total flux in the 3°–9° region from the Sun ROI, which allows us to check for peculiarities in the Sun ROI. The flux ratios are again consistent with being time independent for both energy bands.

None of the control samples demonstrates any systematic effects. This means that the observed signal time variation is robust, and is a feature of the underlying gamma-ray production processes. This variation and its amplitude was never quantitatively predicted and this is the first time it is clearly observed.

6.2.6 Anticorrelation of the Solar-Disk Flux with Solar Activity

We check whether the observed time variation is related to solar activity. Our analysis period coincides with solar cycle 24, which started with the solar minimum in 2009 and reached the solar maximum in 2014. In Fig. 6.6 we overlay the yearly smoothed sunspot number 4, which is a tracer of solar activity. Though the sunspot number and the solar-disk gamma-ray flux vary with different amplitudes, the trends are clearly opposite. In other words, the solar-disk gamma-ray flux anticorrelates with solar activity at least during the first half of the solar cycle 24.

This trend is also qualitatively consistent with the EGRET observation. The flux measured by Ref. [462] used data collected during 1991–1995, which is approximately the second

half of solar cycle 22, when solar activity was declining from the solar maximum. The antici-
correlation explains the smaller flux observed by Ref. [462] compared to Fermi2011, who
used data mainly from the solar minimum.

Before this work, there was no direct evidence showing that the solar-disk gamma-rays
are of cosmic-ray origin. Though only rare solar flares are found to accelerate particles
beyond 1 GeV, it may be possible that some yet-unknown solar processes continuously ac-
celerate particles up to the multi-GeV energy range. However, one expects these solar
processes would be correlated with solar activity, the opposite of the cosmic-ray frame-
work (detailed in the next section). The anticorrelation with solar activity found in the
solar-disk gamma-ray flux therefore strongly indicates that the bulk of the gamma-ray flux
is induced by cosmic rays. (Exploration of theoretical possibilities for the Sun itself to
generate gamma rays that mimic the observed time variation is beyond the scope of this
work.)

It is interesting to put the amplitude of this time variation into perspective, assuming
the cosmic-ray production mechanism. The progenitors of 1–10 GeV solar-disk gamma rays
are \(\sim 10–100\) GeV cosmic-ray protons. The time variation (or modulation) of the cosmic-ray
flux at Earth is known to anticorrelate with solar activity, in the same sense as the solar-disk
gamma-ray flux found in this work. The cosmic-ray flux modulation at Earth is frequently
described by the force field model with a single empirical parameter, the force field potential
\(\Phi\) [476, 477, 478]. The value of \(\Phi\) can be extracted from precision ground-based neutron
observations [479, 480, 481]. We obtain the corresponding values for our observation period
by averaging over the monthly values. In our analysis period, the maximum value of \(\Phi\)
was 630 MV in 2014 and the minimum was 300 MV in 2009. Taking these values, the
maximum cosmic-ray flux is larger than the minimum by about 15% at 10 GeV and 2% at
100 GeV. (For comparison, the extreme yearly values from 1964 to 2014 are about 1200 MV
and 270 MV, which corresponds to about 50% and 5% difference in the cosmic-ray flux
amplitude). This is too small to explain the amplitude seen in Fig. 6.6. This suggests that
one needs additional modulation of the cosmic-ray flux in the inner solar system, variations
in solar atmospheric magnetic fields that can affect cosmic rays of such high energies, or

163
perhaps both to explain the observed variation amplitude.

In fact, the Tibet air shower array found time variation in observations of $\sim 10$ TeV cosmic-ray shadows of the Sun. During the solar maximum, the cosmic-ray shadows are shallower than during the solar minimum [482]. This can be explained by coronal magnetic fields: cosmic rays are more severely deflected by the solar atmospheric magnetic fields during the solar maximum [482]. This implies that it is more difficult for cosmic rays to go deep into the solar atmosphere during the solar maximum, which is consistent with our solar-disk gamma ray observations.

The observation of time variation in the solar-disk gamma-ray flux therefore provides strong support for the cosmic-ray framework, which we discuss in detail in the next section.

6.3 Theoretical Overview and Observational Outlook

In this section, we first review the cosmic-ray framework, i.e., how solar-disk gamma rays can be produced from cosmic-ray interactions with the solar atmosphere. Experienced readers can skip the first part, and move on to the bulk of this section, where we discuss some future prospects on solar gamma-ray theory and observations.

6.3.1 Physics of Solar-Disk Gamma Rays—The Cosmic-Ray Framework

General Considerations

The physics involved in the production of solar-disk gamma rays is complicated by the effects of magnetic fields. To gain some physical insights, we describe some general cases and approximations, following SSG1991.

Cosmic-ray propagation from the interstellar medium to the surface of the Sun is known to be affected by solar magnetic fields carried by the solar wind. As a result, this propagation is also affected by solar activity [476, 477]. Generally, cosmic rays with energy $\lesssim 10$ GeV observed at the Earth are more suppressed when the Sun is more active. Additional modulation of cosmic rays may occur when they propagate from the Earth to the Sun.
Once cosmic rays reach the Sun, their motion is dominated by the magnetic fields in the corona and photosphere. The Larmor radius of cosmic rays near the surface of the Sun sets a reference energy scale, $E_c$. For cosmic-ray protons, taking the typical solar magnetic field strength, $B \sim 1 \text{ G}$, and setting the Larmor radius, $L$, to be the solar radius, $R_\odot \simeq 7 \times 10^5 \text{ km}$, yields

$$E_c \simeq 2 \times 10^4 \text{ GeV} \left( \frac{L}{R_\odot} \right) \left( \frac{B}{1 \text{ G}} \right).$$

A similar scale is obtained for sunspots, where the length scale is about $10^3$ times smaller, but the field strength is roughly $10^3$ times stronger. The range of $E_c$ was estimated in SSG1991 to be between $\simeq 3 \times 10^2 \text{ GeV}$ and $\simeq 2 \times 10^4 \text{ GeV}$. This scale separates the physics into three regimes: $E_p \gg E_c$, $E_p \ll E_c$, and $E_p \sim E_c$, where $E_p$ is the primary cosmic-ray energy.

When $E_p \gg E_c$, one can ignore the magnetic fields. Cosmic rays and their interaction products travel in straight trajectories following the initial cosmic-ray momentum. In this case, only gamma rays from the Sun limb are observable. The Sun limb is the thin layer of the outer solar atmosphere that has high enough column density for cosmic rays to interact, but not so much that gamma rays are unable to escape. This corresponds to a column density of $\mathcal{O}(1)$ hadronic interaction length, which is similar to the photon absorption length. The Sun limb component is non-zero but is argued in SSG1991 to be small; it should also inherit the primary cosmic rays’ spectral index ($\sim 2.7$). The Sun limb component is expected to be time-independent.

When $E_p \ll E_c$, cosmic rays propagate along solar atmospheric magnetic field lines. Inward-pointing (towards the Sun) cosmic rays are funneled into magnetic flux concentrations (or flux tubes) in the photosphere, where the field strength is stronger and the matter density is higher. Assuming adiabatic invariance, the inward-moving cosmic rays would be reflected by the magnetic field strength gradient (magnetic reflection). It is then possible for the cosmic rays to interact with the solar atmosphere on their way out and to produce gamma rays that point toward Earth. This mechanism, suggested in SSG1991, allows the whole solar disk to be involved in gamma-ray production, and thus enhances the flux. Be-
cause the effects of magnetic fields on cosmic-ray propagation are energy dependent, the spectral index of the resultant gamma-ray flux could deviate significantly from that of the primary cosmic-ray spectrum. During solar maxima, the strength of solar atmospheric magnetic fields increases [483], so the magnetic reflection of cosmic rays are expected to occur at higher altitudes, where the density is lower. This decreases the gamma-ray production efficiency during solar maxima compared to that during solar minima, which is qualitatively consistent with the time variation observed in this work.

When $E_p \sim E_c$, no simple approximation can describe the physics. The corresponding gamma-ray energy at $\sim 0.1E_c$ marks the transition from the low-energy regime to the high-energy regime. In other words, the gamma-ray flux, spectral index, and time-dependence should be intermediate between those of the two regimes above. It is interesting to note that the robust detection of the solar-disk component at 30 GeV and the non-observation of a spectral break in this work already requires that $E_c \gtrsim 300$ GeV, which is close to the lower bound estimated by SSG1991. Interestingly, the result from the Tibet air shower array shows that cosmic rays at $\sim 10$ TeV are still affected by solar atmospheric magnetic fields [482].

The SSG1991 Model

We now briefly describe the SSG1991 “naive” and “nominal” cases for the solar-disk gamma-ray flux (see Ref. [64] for details). The SSG naive calculation ignores all the propagation and magnetic-field effects, assumes 100% efficiency for cosmic-ray absorption in the solar surface, and counts all the gamma rays produced. The naive case, therefore, is an robust theoretical upper limit on how much solar-disk gamma rays can be produced by cosmic rays. It is not a physical model, and hence, it is not surprising that our flux and that from Fermi2011 is lower than this bound.

The appropriate comparison with data is using the SSG nominal model, shown in Fig. 6.4. In this case the cosmic-ray propagation was treated as a diffusion problem from the Earth to the Sun. Primarily concerning the $E_p \ll E_c$ case, all cosmic rays were assumed to land on magnetic flux tubes, and then reflected with some efficiency. With a chosen set
of diffusion parameters, the cosmic-ray absorption rate was determined, which is roughly 0.5%. Finally, the magnetically enhanced gamma-ray flux was obtained by integrating the gamma-ray yield with the absorption rate and the path length distribution. The upper edge of the green band in Fig. 6.4 corresponds to the extreme case where all the cascade products are charged and contribute to the gamma-ray production. The lower edge corresponds to the conservative case where all the cascade products are neutral, hence only primaries that interact after being reflected can contribute to the gamma-ray flux. These two cases bracket the theoretical uncertainty concerning the cascade development inside the flux tubes, but not other model ingredients.

6.3.2 Prospects for Solar-Disk Gamma-Ray Theory

As already discussed in Fermi2011, the SSG nominal model is unable to explain the observed gamma ray data. Our result, even if taken at solar maximum, is still inconsistent with the SSG nominal model. Therefore, it is necessary to revisit the modeling of the cosmic-ray framework. Most likely, new implementations of cosmic-ray physics and solar physics are needed. We will provide new theoretical investigations in our forthcoming papers.

There are several key observations that the new model needs to address. First, it needs to reexamine the effectiveness of solar magnetic fields in enhancing the gamma-ray flux at $E_p \ll E_c$. In particular, SSG1991 estimated $\sim 0.5\%$ of the total available cosmic-ray energy at the Sun is converted to gamma rays, but observations suggest $\sim 5\%$, modulo the time variation. Second, the high-energy gamma rays found in this work demand a proper treatment of the $E_p \sim E_c$ and $E_p \gg E_c$ regimes. Third, the time variation found in this work, as well as that from the Tibet air shower array, show that the model should track the variations of solar magnetic activity. Lastly, the model needs to quantitatively explain the observed amplitude of the time variation.

With an accurate model of gamma-ray production, solar gamma-ray observations can be used to constrain model ingredients and parameters, thus providing a new probe of solar atmospheric magnetic fields and of cosmic-ray propagation in the solar system. This is particularly promising given that many current and future instruments will have excellent...
sensitivity for continuously monitoring solar gamma rays.

With a sufficient understanding of the solar-disk gamma rays, it will be possible to use the Sun as a laboratory to test new physics. For example, a popular dark matter candidate is the Weakly Interacting Massive Particle (WIMP), which can accumulate and annihilate in the core of the Sun after being gravitationally captured ([55, 56, 57]; see Ref. [58] for a recent review). Typical WIMPs captured in the Sun generate negligible electromagnetic signals [484]. However, non-minimal physics, such as inelastic dark matter [485, 486, 487] and metastable mediators in the dark sector [488, 489, 71, 490], can significantly enhance the electromagnetic signatures [491, 492, 493, 494]. Understanding the standard model predictions is necessary to uncover or interpret any potential signatures from dark matter [495, 64, 496, 497, 498, 499]. For example, both the spectral information and time variation can be useful model differentiators. We will further discuss the implications of high-energy solar observations for new physics in our forthcoming papers.

6.3.3 Prospects for the Inverse Compton Component

In our analysis, the IC component is treated as a background. However, with new data releases, which improve both statistics and data quality, a more precise study of the IC halo component is also warranted. A minor tension between the data and the prediction for the IC component was found in Fermi2011, where the data seemed to be higher at 10 GeV than expected. A more precise measurement is needed to clarify the situation.

A new study of the IC component will allow one to use gamma rays to probe the cosmic-ray electron density in the solar system [453]. This is because the IC intensity is the product of the electron density and the photon density along the line of sight, with the latter being a known quantity. The IC component is therefore sensitive to electron densities from fairly close to the Sun to beyond the Earth’s orbit. In addition, if there is time variation in the IC component, its broad angular distribution may allow one to test the variation amplitude as a function of the distance to the Sun. These observations can help with understanding cosmic-ray modulation in the solar system, which despite many years of effort, is still under active investigation [455, 500, 478]. This approach is complementary
to solar-disk gamma-ray observations, which are strongly affected by the conditions of the solar atmosphere.

Similar to our analyses, it is also interesting to characterize the IC component beyond 10 GeV as well as search for long-term time variations. Because point sources are not removed, our analysis is not optimized for the IC component. With this caveat, we check the best-fit IC amplitude from our analysis, and we find no obvious time variation (only $\sim 20\%$ scatter around the mean). A more careful analysis is needed to provide a definitive statement. Analyzing the IC component is difficult at high energies, where statistics are low, and equally challenging at low energies for Fermi-LAT, where the PSF is $\sim 10^\circ$ at 100 MeV.

### 6.3.4 Prospects for Fermi and Future Space Missions

In this work, we use a straightforward analysis to characterize and robustly detect important features of the solar-disk gamma rays. Future analyses and observations, with more optimized analysis procedures and improved data sets, can yield more precise measurements or even discover new features. Below, we discuss some possible analysis improvements with Fermi.

At high energies, where statistics are low, one can use an unbinned analysis to fully utilize the information carried by each photon. In particular, better angular resolution at high energies may allow one to resolve the solar disk and locate hot spots (as for solar flares [449, 450]). On the other hand, the improved angular resolution also means that the solar disk can no longer be treated as a point source. One needs to take into account the fact that the astrophysical diffuse background and the IC component are reduced toward the solar disk [454]. This also means that the one should avoid using the stacking procedures performed in this work, which slightly smears the position of the Sun according to the length of each time segment. Instead, one should select the events and calculate the exposure in a solar-centric coordinate system.

For improving statistics, one can potentially develop more optimized cuts. For example, it is likely that the Galactic plane cut employed in this work can be improved, given that
the Galactic plane gamma-ray intensity drops rapidly with latitude and can in principle be modeled. This may improve the statistics by about a factor of 2. In addition, the new Fermi data release, Pass 8 [165], has a larger effective area and better angular resolution. Improving the statistics is particularly important for high-energy observations.

Next-generation space gamma-ray telescopes can further improve the solar-disk observations in both time and energy range. The apparent anticorrelation between the solar-disk gamma-ray flux and solar activity suggests that the flux should start to increase as we start to leave the solar maximum. This can be checked with near-future data from Fermi. Next-generation instruments, such as DAMPE [501], GAMMA-400 [502], and HERD [439], will allow the Sun to be monitored at the GeV range even beyond Fermi’s lifetime. Though in principle Fermi is sensitive to gamma rays down to 10 MeV, extracting the solar-disk signal is difficult due to the broad PSF. Future missions such as PANGU [503] and ComPair [504] can provide improved sensitivity in the MeV range. Low-energy observations could provide additional information on the time variation and probe potential leptonic components or even new solar-disk gamma-ray emission mechanisms.

### 6.3.5 Prospects for Ground-Based Telescopes

To expand solar gamma-ray observations into the TeV range and beyond, large ground-based experiments are required. It is impossible for air-Cherenkov telescopes to observe the Sun due the bright optical emission from the Sun itself. The Sun, therefore, is a unique target for water-Cherenkov telescopes such as HAWC and LHAASO.

To assess whether water-Cherenkov telescopes can detect the Sun, we consider the total solar gamma-ray flux, including both the solar-disk and IC components. We estimate this flux by finding the total flux within 1.5° of the Sun and subtracting the diffuse background. In this case, the Sun is detected at > 5σ in all eight energy bins. Assuming a single power-law spectrum, the total solar gamma-ray flux can roughly be described by $3.5 \times 10^{-8} (E/\text{GeV})^{-2.3} \text{GeV}^{-1} \text{cm}^{-2} \text{s}^{-1}$ in 1–100 GeV.

Figure 6.7 shows the total solar gamma-ray flux, the solar-disk-only component from Fermi2011, the solar-disk-only component found in this work, and the diffuse background
Figure 6.7: Energy spectrum of gamma rays from the Sun. Blue squares are the total solar gamma-ray flux (solar disk + IC) within 1.5° of the Sun with only statistical error bars. Black dots are the solar-disk-only component from Fermi2011. The grey band shows the solar-disk-only component found in this work. Green circles are the estimated diffuse background within 1.5° of the Sun. The differential point-source sensitivities for HAWC [41] and LHAASO [42, 43] are shown.
within 1.5° of the Sun. The total solar gamma-ray flux is clearly much larger than the diffuse background. For comparison, we show also the sensitivity of HAWC [41] and LHAASO [42, 43]. If the total solar gamma-ray flux follows the same spectral index to the TeV range, both HAWC and LHAASO should be able to detect the Sun.

The water-Cherenkov telescopes are in a unique position to probe solar gamma rays. In particular, they are sensitive to the $E_p \sim E_c$ and $E_p \gg E_c$ regimes. Either a detection or an upper limit from the water-Cherenkov telescopes can provide valuable information on gamma-ray production from the Sun.

6.4 Conclusions

Despite being the nearest star to us, much about the Sun’s gamma-ray emission is still poorly understood. Previous study by the Fermi collaboration, who used 1.5 years of data, precisely detected the solar-disk gamma rays in 0.1–10 GeV. However, the flux is about ten times brighter than predicted. Motivated by this puzzle, we focus on the solar-disk component, and use 6 years of public Fermi data to gain a better understanding of these gamma rays. We employ a straightforward and conservative analysis to search for new features in the gamma-ray flux.

Utilizing the improved photon statistics, we extend the observations to 100 GeV. As in Fermi2011, we find that the gamma-ray flux is higher than the central value of the SSG1991 prediction by about one order of magnitude in 1–10 GeV, modulo time variation. In addition, we detect the solar-disk component in 10–30 GeV at $>5\sigma$, and in 30–100 GeV at $>2\sigma$. This is the first time the Sun is detected above 10 GeV in gamma rays. There are no theoretical predictions for solar-disk gamma rays in this energy range. As a result, our observations demand further theoretical investigation.

Importantly, we find a significant time variation in the solar-disk gamma-ray flux over the analysis period, which apparently anticorrelates with solar activity. This is the first clear observation of such a time variation, though it was hinted at in earlier studies [462, 463]. This variation was not theoretically predicted, and its large amplitude deserve further inves-
tigation. Nonetheless, the anticorrelation with solar activity indicates that the bulk of the solar-disk gamma rays can be explained by cosmic-ray interactions in the solar atmosphere and the gamma-ray production process is strongly affected by the solar magnetic fields.

Future observations with Fermi and other instruments may provide even more information about gamma rays from the Sun. For example, the anticorrelation of the solar-disk gamma-ray flux with solar activity can be further confirmed with near-future Fermi data. In addition, our robust detection ($>5\sigma$) of the total solar gamma-ray flux shows that the Sun is a new and promising source for large water-Cherenkov gamma-ray telescopes, such as HAWC and LHAASO. Observations from water-Cherenkov telescopes can provide important insights on the gamma-ray production processes in the TeV range.

This work lays the observational foundation for our future theoretical work, where we will investigate in detail how cosmic rays interact with the Sun under the influence of solar magnetic fields. We will study the multi-messenger signatures from these high energy processes, their implications for solar physics, cosmic-ray physics, and new physics. Gamma-ray studies of the Sun are still in their infancy, but have already yielded interesting results. Future observations and the accompanying theoretical investigations may uncover even greater surprises.
In this thesis, we discuss work that is primarily focused on particle dark matter detection. It has become apparent that the solution of the dark matter problem requires more ingenuity than originally expected. The work contained in this thesis is aimed to produce new and sometime novel tests of dark matter.

In Chap. 2, we study the contribution of dark matter annihilation to the extragalactic gamma-ray background. In Chap. 3, we study the prospects of using high-energy astrophysical neutrinos to probe secret neutrino interactions. In Chap. 4, we present the novel usage of Fermi-GBM data in searching for sterile neutrino dark matter. In Chap. 5, we discuss the novel idea of dark matter velocity spectroscopy, which could help diagnose any dark matter signals that have a line spectrum. In Chap. 6, we study the puzzling gamma-ray emission of the Sun, providing important clues that might help understand the gamma-ray production mechanism. This will be an important step in characterizing the neutrino background for solar dark matter searches.

We believe these projects have all advanced the study of dark matter in some capacity. We have discussed obvious and important follow-up studies that will yield more interesting results in the near future.
Bibliography


190


196


199


200


