Astrophysical Probes of Particle Dark Matter

Title

Permalink
https://escholarship.org/uc/item/1xs489xb

Author
Kwa, Anna

Publication Date
2017

License
CC BY 4.0

Peer reviewed|Thesis/dissertation
UNIVERSITY OF CALIFORNIA, IRVINE

Astrophysical Probes of Particle Dark Matter

DISSERTATION

submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in Physics

by

Anna Kwa

Dissertation Committee:
Professor Manoj Kaplinghat, Chair
Professor Kevork Abazajian
Professor James Bullock

2017
DEDICATION

To my family
# TABLE OF CONTENTS

## LIST OF FIGURES

- v

## LIST OF TABLES

- x

## ACKNOWLEDGMENTS

- xii

## CURRICULUM VITAE

- xiii

## ABSTRACT OF THE DISSERTATION

- xv

## 1 Introduction

1.1 Indirect detection of weakly interacting massive particles (WIMPs) ........ 4
1.2 Self-interacting dark matter ........................................ 8

## 2 Spatial Non-Uniformity of the Galactic Center Excess

2.1 Background .............................................. 12
2.2 Methods .................................................. 15
   2.2.1 Fit components ....................................... 16
   2.2.2 Fit procedure ....................................... 22
2.3 Results .................................................. 24
   2.3.1 Systematics associated with background model components ............ 24
   2.3.2 The GCE spectrum in the galactic center versus the inner galaxy .... 27
   2.3.3 The GCE spatial profile and radial distribution ...................... 29
   2.3.4 Spatial uniformity of the galactic center excess spectrum ............ 31
2.4 Discussion ................................................ 34
   2.4.1 The GCE high-energy spectrum above 10 GeV .......................... 34
   2.4.2 Is a point source population favored over a smooth annihilation profile as the origin of the GCE? ........................................ 41
2.5 Conclusions ................................................ 43

## 3 A New Gamma-ray Source Consistent with Upscattered Starlight

3.1 Background .............................................. 51
3.2 Methods .................................................. 52
3.3 Results .................................................. 53
3.4 Interpretation and Discussion ..................................... 57
# Lepton-Flavor Violating Mediators

## Background

4.1 Background

4.2 Models of a Lepton-Flavor Violating Mediator

4.2.1 Dark Sector Interactions

4.2.2 Self-Interacting Dark Matter

4.3 The Fermi-LAT $\gamma$-Ray Excess

4.3.1 Photons from Leptons

4.3.2 Fit to Fermi $\gamma$-ray Excess

4.4 The AMS-02 $e^\pm$ Spectrum

4.5 Conclusions & Outlook

## Hidden Sector Hydrogen as Dark Matter

5.1 Background

5.2 SIDM cross sections at low velocities

5.3 Atomic Dark Matter Model

5.3.1 Quantum formalism

5.3.2 Hyperfine interaction

5.3.3 Scattering

5.4 Applications of Atomic Dark Matter

5.4.1 Cosmological considerations

5.4.2 Cross sections, lifetimes, and structure formation

5.4.3 SIDM halo profiles

5.4.4 Minimum halo masses

5.5 Consequences of Inelastic Scattering

5.5.1 Comparison of viscosity and upscattering cross sections

5.5.2 Halo cooling

5.6 Additional Considerations at the Cluster Scale

5.6.1 Upscatterings to the $n = 2$ excited state

5.6.2 Ionization in the late universe

5.7 Conclusions

## Conclusions

# Bibliography
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>The diversity of rotation curves across a range of spiral galaxy masses. Each panel contains 14 colored galaxy rotation curves that are selected to have similar flat rotation velocities ( v_{\text{flat}} ) at their furthest radial data points. From left to right, the corresponding ( v_{\text{flat}} ) bins are 79 – 91 km/s, 139 – 172 km/s and 239 – 315 km/s, spanning most of the mass range of the galaxies considered in this work. The points with error bars are rotation curve data from the SPARC dataset [159]. Solid lines are the halo model fits to the data using a SIDM model with cross section ( \sigma/m = 3 \text{ cm}^2/\text{g} ).</td>
</tr>
<tr>
<td>2.1</td>
<td>Total observed counts map with labelled regions of interest. Each ROI will be referred to hereafter using its label from this figure. We will collectively refer to the entirety of the ROIs excluding the innermost 7° × 7° as the ‘inner galaxy’ (orange). The innermost 7° × 7° ROI is referred to as the ‘galactic center’ (blue). The overlap of the Fermi bubble template used in this analysis (c.f. Sec. 2.2.1) with the ROIs is shown as the gray overlay. Not shown here are the farther latitude N2/S2/bubble N/bubble S ROIs which were used to constrain the Fermi bubble spectrum and estimate the extent of the GCE signal.</td>
</tr>
<tr>
<td>2.2</td>
<td>Spectra for the GCE and spatially extended background model components in the galactic center (top row) and combined inner galaxy ROIs (bottom row). Fits were performed with the GALPROP IC and ( \pi^0 )+bremsstrahlung templates free to vary independently of each other in each energy bin. Red ★ symbols denote the best fitting background model in the respective regions.</td>
</tr>
<tr>
<td>2.3</td>
<td>Same as the top row of Fig. 2.2, but with the fits performed with the IC and ( \pi^0 )+bremsstrahlung templates constrained to have the same relative normalizations to each other as predicted by GALPROP. The dashed magenta line plots the GCE spectrum from Fig. 2.2 (where the IC and ( \pi^0 )+bremsstrahlung were fit separately) for comparison. Note that the dashed comparison NFW annihilation spectrum is indeed plotted in the model F panel but is difficult to see because of its close overlap with the solid magenta NFW spectrum (where the diffuse templates are fixed relative to each other).</td>
</tr>
</tbody>
</table>
2.4 Best-fit GCE spectra in the galactic center (black squares) and inner galaxy (orange triangles) regions, shown for varied GALPROP diffuse models (rows) and NFW density profile slopes $\gamma$ (columns). The spectrum of the GCE in the inner galaxy is shown for the sum of all inner galaxy ROIs. Normalizations are scaled to show the expected flux for the entire GCE template ($35^\circ \times 35^\circ$), such that the normalization for the two ROIs will match if they are consistent with originating from a single NFW-distributed source. Also shown are the exponential cutoff parameterized fits (gray dashed line) to the galactic center spectrum. The panel with the light gray (orange) background denotes the NFW slope and diffuse background combination with the highest likelihood fit for the galactic center (inner galaxy) ROI as recorded in Tab. 2.3.

2.5 Change in test statistic value for inner galaxy fits as a function of NFW profile slope $\gamma$. Individual panels show results obtained using GALPROP diffuse models A/E/F. Each line shows $-\Delta TS$ calculated using the fits in energy bins $< 1.9$ GeV (blue dashed line), 1.9-10.0 GeV (green dot-dashed line), $> 10.0$ GeV (red dotted line) as well as the full energy range (solid indigo line). For a given GALPROP background, this allows providing a visualization of which energy bins are driving the fit towards the preferred NFW slope.

2.6 Radial distribution of the GCE residual flux in energy bins 0.7–1.9 GeV, 1.9–10.0 GeV, and 10.0 GeV–200 GeV. The top row plots the GCE residuals for the best fitting profile slopes obtained when fitting with diffuse model backgrounds A/E/F. For all models and ROIs this corresponds to $\gamma = 1.1$, with the exception of the inner galaxy model F fit which is plotted for $\gamma = 0.9$. The bottom row plots the GCE residual for the fits with $\gamma = 1.1$ (0.9) in the galactic center (inner galaxy) and model F diffuse background (solid black steps) against the expected radial distribution of emission from NFW-like sources of varying profile slopes (colored lines) comparison.

2.7 Best-fit GCE spectra in the galactic center ROI and the 6 inner galaxy ROIs, shown for fits with diffuse backgrounds model A (dark blue), E (purple), and F (light blue). Additional regions N2/S2 are also plotted. Panels are arranged to reflect location on the sky. We show our results for the case of a GCE template with profile slope $\gamma = 1.1$, though our findings are consistent for all template slopes. Normalizations are scaled such that each subplot shows the expected flux for the entire GCE template ($35^\circ \times 35^\circ$). Energy coordinates of the data points are slightly offset for visibility purposes.

2.8 Left: Comparison of the GCE spectrum in the galactic center (black circles) and inner galaxy (orange squares) to the spread of globular cluster MSP spectra (dashed gray lines) from Ref. [70]. Also shown are the inferred MSP spectra from Ref. [70] derived from stacked observations of globular clusters (solid green line) and individual MSPs (dashed red line). The spectra are normalized to match each other at 2 GeV for ease of comparison. Right: Comparison of the inner galaxy GCE spectrum (orange squares) to our estimations of the prompt (dotted line) and IC (dashed line) components of the potential unresolved MSP population.
2.9 Simulated photon-count distribution in the combined inner galaxy ROIs between 1.9–10.0 GeV of an NFW source with $\gamma = 1.1$ (dashed blue line) vs. a population of unresolved point sources (solid pink line) with the best-fit source count function from Ref. [158]. Pixel size is 0.2° per pixel. The spectrum and total flux of each simulated source is chosen to match our total observed flux in this energy range. For comparison, we also show the photon-count distribution of the full residual (solid gray line) for the case of GALPROP model F and $\gamma = 1.1$. Note that the zero-count bin is not shown.

3.1 Shown in the left column are the residual photons (top) and best fit model (bottom) associated with the projected interstellar radiation field template in its peak intensity bin, $0.303 \text{ GeV} < E_\gamma < 0.372 \text{ GeV}$, where the residual map has been smoothed with a Gaussian of $\sigma = 0.9^\circ$ (to roughly account for the point-spread function). The middle column shows the residual photons (top) and best fit model (bottom) associated with the projected dark matter density squared template in its peak intensity bin, $1.59 \text{ GeV} < E_\gamma < 1.95 \text{ GeV}$, where the residual map has been smoothed with a Gaussian of $\sigma = 0.4^\circ$. The right column shows the residual photons (top) and best fit model (bottom) associated with the 20 cm radio map in the same energy bin and with the same smoothing as the middle row. Residual and model maps have the same color scale for each row. This analysis used ULTRACLEAN-class photons.

3.2 The residual spectra (points with errors) and best-fit model spectra for the projected interstellar radiation field (golden triangle), gas (pink circle), and dark matter density squared (blue square) templates. ULTRACLEAN class photons are used for this analysis.

3.3 Shown here is an example 8 GeV dark matter annihilation model with equal branching to all charged leptons, $e^\pm, \mu^\pm, \tau^\pm$, with the residual spectra of the prompt GCE (blue square), IC (golden triangle), and bremsstrahlung (pink circle) sources. The blue (dashed) GCE spectrum is is determined by the particle mass and annihilation rate fit to the observations. The solid predicted resultant spectra for this annihilation channel’s IC (golden) and bremsstrahlung (pink) cases are in solid lines. ULTRACLEAN class photons are used for this analysis.

4.1 Diagrams showing annihilation to mediators and possible off-diagonal mediator decay modes with off-shell $\ell, \ell' = \mu$ or $\tau$. The dot represents the Fermi vertex from an off-shell $W$. 

vii
4.2 The predicted combined prompt and ICS gamma-ray spectra for each scenario: \(\chi \bar{\chi} \rightarrow 2\phi\) (left) and \(\chi \bar{\chi} \rightarrow 3\phi\) followed by the lepton-flavor violating decays indicated in each panel. The grey shaded region represents the Fermi collaboration’s \(\gamma\)-ray excess spectrum bounded by its estimated systematic error when fit as a parameterized form to the entire energy range of the data. Each color-coded band corresponds to a set of \(\{m_\chi, m_\phi\}\) with \(m_\phi\) varying in the range \([m_{\text{heavy}}, m_\chi]\) (left) and \([m_{\text{heavy}}, \frac{2}{3}m_\chi]\) (right). The spectra are calculated assuming a halo profile slope of \(\gamma_{\text{NFW}} = 1.0\) and the annihilation cross sections indicated in each figure. For a steeper halo profile of \(\gamma_{\text{NFW}} = 1.2\) cross sections are a factor of \(\sim 3\) smaller.

4.3 Same as Fig. 4.2, but now the grey shaded region represents the \(\gamma\)-ray excess spectrum bounded by its estimated systematic error when fit in independent energy bins, as reported by the Fermi collaboration. Our predicted combined prompt and ICS gamma-ray spectra for each scenario: \(2 \rightarrow 2\) (left) and \(2 \rightarrow 3\) (right) \(\chi\chi\) annihilations to \(\Phi\), followed by one of the following decays: \(\phi \rightarrow \tau e\) (up), \(\tau\mu\) (middle), and \(\mu\epsilon\) (bottom). Each color-coded band corresponds to a set of \(\{m_\chi, m_\phi\}\) with \(m_\phi\) varying in the range \([m_{\text{heavy}}, m_\chi]\) (left) and \([m_{\text{heavy}}, \frac{2}{3}m_\chi]\) (right).

4.4 Predicted \(e^\pm\) spectra at Earth’s position for each scenario: \(\chi \bar{\chi} \rightarrow 2\phi\) (left) and \(\chi \bar{\chi} \rightarrow 3\phi\) (right) followed by one the lepton-flavor violating decays indicated in each panel. The grey shaded region represents the bounds from AMS-02 electron and positron spectra while each color-coded band corresponds to a set of \(\{m_\chi, m_\phi\}\) with \(m_\phi\) varying in the range \([m_{\text{heavy}}, m_\chi]\) (left) and \([m_{\text{heavy}}, \frac{2}{3}m_\chi]\) (right). The dashed lines shown in the bottom row correspond to the prediction for AMS-02 spectra if an NFW profile slope of \(\gamma_{\text{NFW}} = 1.2\) is assumed. This leads to a lower annihilation cross section of \(1.1 (1.4)\) times the relic density for \(\chi \bar{\chi} \rightarrow 2(3)\phi\).

5.1 Left: A galaxy (D631-7) where the dark matter halo density profile has a shallow rise, and is able to be fit well with a SIDM halo model with \(\sigma/m = 3\text{cm}^2/\text{g}\) (upper left) but not with a collisionless NFW halo (upper right). Right: A galaxy (UGC 08490) with a steeply rising dark matter density profile, in which the rotation curve may be well-fit using either a SIDM halo model with \(\sigma/m = 3\text{cm}^2/\text{g}\) (lower left) or a collisionless NFW halo (lower right).

5.2 Energy-level diagram for the \(n = 1\) ground state of dark hydrogen. Hyperfine interactions break the degeneracy between the \(F_{A,B} = 0\) and \(F_{A,B} = 1\) states for the dark atoms \(A\) and \(B\). The labels \(b, c,\) and \(d\) correspond to \(m_F = -1, 0,\) and \(1,\) respectively.

5.3 Parameter space scan comparing the viscosity cross section (5.19), which involves only hyperfine ground-state atoms, to the total viscosity cross section in the elastic approximation, used in Ref. [76]. The left plot shows \(\sigma_V\) with \(E_{hf} = 10^{-5}E_0\), while the right plot shows \(\sigma_V\) with \(E_{hf} = 10^{-4}E_0\). The spin averaging factors for \(\sigma_V\) and \(\sigma_{\text{tot},V}^{(\text{elastic})}\) are different, so we do not necessarily expect \(\sigma_V\) to be less than \(\sigma_{\text{tot},V}^{(\text{elastic})}\).
5.4 Viscosity scattering cross sections are calculated at the velocities of interest at each point in the $m_H-\alpha$ plane and then used to determine which areas either satisfy target cross sections or are in tension with observations. The hyperfine splitting is fixed to $E_{hf} = 10^{-5} E_0$ in the left panel and $E_{hf} = 10^{-4} E_0$ in the right panel. The vertical hatched grey area is disfavored by measurements of cluster halo ellipticities and corresponds to the region where $\sigma_V/m_H > 1 \text{ cm}^2/\text{g}$ for velocities $v_{\text{rms}} = 1000 \text{ km/s}$. We show contours of the ionization fraction $\chi_e = 10^{-1}, 10^{-2}, \text{ and } 10^{-3}$; we consider $\chi_e \lesssim 10^{-2}$ sufficiently low to ignore dark ions, which excludes a large portion of the displayed parameter space for $E_{hf} = 10^{-5} E_0$. Points within the cross-hatched green region satisfy $0.5 \text{ cm}^2/\text{g} < \sigma_V/m_H < 5 \text{ cm}^2/\text{g}$ for velocities $v_{\text{rms}} = 30-100 \text{ km/s}$, which approximates the condition for cores to form in lower-mass halos. Points within the solid orange region provide the best-fit viscosity cross sections for cores to form in relaxed cluster halos. The dashed lines show contours of constant minimum halo mass for values of $E_{hf}, m_H$, and $\alpha$ in our model, assuming $\xi = 0.6$. Lower, allowed values of $\xi$ lead to smaller minimum halo masses.

5.5 Viscosity (red curves) versus upscattering (blue curves) cross sections per unit mass as a function of halo velocity. Because of the short decay time scale for the hyperfine excited state, we assume that both initial particles are in the ground state. The value of the hyperfine splitting is fixed to $E_{hf} = 10^{-5} E_0$ and $E_{hf} = 10^{-4} E_0$ in the left and right panels, respectively. Each line is drawn randomly from values of $m_H$ and $\alpha$ lying within the overlapping regions of Fig. 5.4 corresponding to parameters which lead to cores consistent with observations in both cluster-scale halos as well as dwarf- to LSB-scale halos.

5.6 The left subpanels in each figure show the constrained (vertical grey hatched) and target (solid orange and cross-hatched green) areas of parameter space. The right subpanels show the ratio of the inelastic viscosity cross section to the viscosity cross section as a function of $\alpha$. The left and right figures are shown for fixed hyperfine splittings of $E_{hf} = 10^{-5}$ and $10^{-4}$, respectively.

5.7 We compare the effects of SIDM heating and cooling mechanisms in atomic dark matter halos by plotting the ratios of outward energy flow lost through cooling over the inward heat flow from scatterings at a radius of $r = 0.5 r_s$ in the halo, which is approximately the radius at which the inward heat flow due to scatterings and the outward energy loss due to upscatterings are greatest, as well as $r = r_s$. The left and right figures in both rows are shown for fixed hyperfine splittings of $E_{hf} = 10^{-5}$ and $10^{-4}$, respectively. The lower mass halo plotted in each panel (black triangles) corresponds to the smallest halos in which upscatterings to the hyperfine excited state are not suppressed by low particle velocities. We also show the cooling to heating ratios at cluster scales for cooling through hyperfine excitations (magenta circles) and $n = 2$ excitations (blue squares). See Sec. 5.6.1 for details and discussion regarding our estimation of the $n = 2$ cooling rate.
2.1 Input parameters for our set of three GALPROP diffuse background models. We use the same scale radius $r_D=20 \text{ kpc}$ and Alfvén speed $v_A=32.7 \text{ km s}^{-1}$ for all models. The scale height $z_D$ is given in units of kpc. The diffusion coefficient $D_0$ is given in units $\times 10^{28} \text{ cm}^2 \text{ s}^{-1}$. The convection velocity gradient $dv/dz$ is given in km s$^{-1}$ kpc$^{-1}$. The cosmic-ray source distribution is taken from either the measured supernova remnant (SNR) distribution [68] or the Lorimer pulsar distribution [162]. (Both of these cosmic-ray distributions approach zero at the galactic center and are in all likelihood severely underestimating the cosmic-ray source density in the innermost kpc. We discuss the implications of this deficiency in Sec. 2.3.1.) The power law index of the electron (proton) injection spectrum above rigidity $2.18 \text{ (11.3)} \text{ GV}$ is given by $\alpha_e(\alpha_p)$. The electron (proton) cosmic-ray injection spectrum is normalized to $N_e(N_p)$ in units of $\times 10^{-9}$ cm$^{-2}$ sr$^{-1}$ s$^{-1}$ MeV$^{-1}$ at $34.5 \text{ (100)} \text{ GeV}$. The first set of three digits in the magnetic field model are $B_0 \times 10\mu\text{G}$, the second set of three digits are $r_c \times 10 \text{ kpc}$, and the last set of three digits are $z_c \times 10 \text{ kpc}$. ISRF normalization factors are given for the optical, IR, and CMB components respectively. The gas spin temperature $T_S$ is in units of K. A full description of the parameters may be found in Refs. [10, 65].

2.2 Our regions of interest, as defined by range in galactic longitude $l$ and latitude $b$. Angular areas in steradians are also given, although all our results for best-fit GCE flux in each ROI are normalized to display the total expected flux (GeV s$^{-1}$ cm$^{-2}$) from the $35^\circ \times 35^\circ$ GCE template, based on the observed flux for each individual ROI. Note that the farthest latitude regions ‘bubble N/S’ and ‘N2/S2’ are not included in the GCE analysis and thus not shown in Fig. 2.1; they are included solely for the purpose of constraining the Fermi bubbles’ spectrum and testing the extent of the GCE signal.

2.3 Relative differences in the test statistic (TS) value of the NFW template, given for all combinations of NFW template/diffuse background in both the galactic center and combined inner galaxy ROIs. The differences in test statistic values are given relative to the combination of NFW density slope and diffuse background with the highest TS values: $\gamma=1.1$/model A in the galactic center ROI and $\gamma=0.8$/model F in the inner galaxy ROIs.
4.1 Summary of Figs. 4.2 and 4.3. The annihilation cross-section, \( \langle \sigma v \rangle \) is given for \( \gamma_{NFW} = 1.0 \) (1.2). The \( \mu e \) modes are in tension with the AMS-02 positron bound unless one takes the contracted \( \gamma_{NFW} = 1.2 \) profile. The range of mediator masses are given in (4.3).

5.1 List of interaction channels with associated quantum numbers. The level for dark atoms A and B corresponds to the labeling in Fig. 5.2. The horizontal lines exhibit the block diagonal nature of the potential \( \mathcal{V}(x) \).

5.2 Definition of various spin-averaged cross sections.
ACKNOWLEDGMENTS

I would like to give a special thank you to Manoj Kaplinghat, who has been a tremendous advisor and source of guidance and support. I am also grateful to the postdocs and faculty whose mentorship and collaboration has been very important to me: Jennifer Johnson, Kev Abazajian, Shunsaku Horiuchi, Kim Boddy, Iftah Galon, Flip Tanedo, Annika Peter, Sheldon Campbell, and Haibo Yu.

Thanks to my fellow graduate students, officemates, friends, and colleagues for making graduate school fun. Thanks also to everyone else—faculty, staff, and postdocs—in the department for your support and camaraderie.

To my parents and sister—Mom, Dad, and Mei-en, thanks for being there for me since day zero and for your encouragement in all of my pursuits. Last, but definitely not least, thank you to my husband Preston for being my number one supporter throughout everything.

This research has been funded by NSF GRFP Grant No. DGE-1321846 and a UC President’s Dissertation Year fellowship.

Portions of this work were published in the Journal of Cosmology and Astroparticle Physics, Physics Review D, and the Journal of High Energy Physics.
CURRICULUM VITAE

Anna Kwa

EDUCATION

Doctor of Philosophy in Physics 2017
University of California, Irvine

Bachelor of Science in Physics & Astronomy 2011
The Ohio State University

RESEARCH EXPERIENCE

Graduate Research Assistant 2012–2017
University of California, Irvine

TEACHING EXPERIENCE

California State Summer School in Math & Science Summer 2014–2017
University of California, Irvine

Lecturer, Physics 3A Summer Session 2017
University of California, Irvine

Teaching Assistant 2012–2014
University of California, Irvine

AWARDS

President’s Dissertation Year Fellowship 2017
National Science Foundation Graduate Fellowship 2014-2017
President’s Dissertation Year Fellowship 2017

REFEREED JOURNAL PUBLICATIONS

Lepton-Flavor Violating Mediators 2016
with Iftah Galon and Phillip Tanedo, JHEP 1703 (2017) 064

Hidden Sector Hydrogen as Dark Matter: Small-scale Structure Formation Predictions and the Importance of Hyperfine Interactions 2016
with Kim Boddy, Manoj Kaplinghat, and Annika Peter, Phys.Rev. D94 (2016) no.12, 123017
Investigating the Uniformity of the Excess Gamma rays towards the Galactic Center Region
with Shunsaku Horiuchi and Manoj Kaplinghat, JCAP 1611 (2016) no.11, 053

Discovery of a New Galactic Center Excess Consistent with Upscattered Starlight
with Kevork Abazajian, Nicholas Canac, and Manoj Kaplinghat, JCAP 1507 (2015) no.07, 013

SELECTED PRESENTATIONS

TeV Particle Astrophysics
SIDM as an explanation for the diversity of rotation curves
2017

APS April Meeting
Tails and Clumps in the Galactic Center Excess
2017

Los Alamos Particle Physics Seminar
Astrophysical Searches for Dark Matter: a light sampler
2017

AAS Winter Meeting
Atomic Dark Matter
2017

Harvard Particle Theory Seminar
Atomic Dark Matter
2016

UC Davis Cosmology Seminar
Atomic Dark Matter
2016

Fermi Symposium
Discovery of a New Galactic Center Excess Consistent with Upscattered Starlight
2014
ABSTRACT OF THE DISSERTATION

Astrophysical Probes of Particle Dark Matter

By

Anna Kwa

Doctor of Philosophy in Physics

University of California, Irvine, 2017

Professor Manoj Kaplinghat, Chair

Although many independent astrophysical and cosmological observations imply the existence of a new dark matter particle, experimental searches have so far been unable to provide any conclusive evidence of such a particle. The study of dark matter and its behavior in astrophysical systems is a promising directions for current and future searches, as the sheer scale of these systems allows physicists to probe regions of parameter space that are unreachable in laboratory or collider experiments. This thesis will describe several ways in which we may test dark matter models using astrophysical observations. Chapters 2, 3, and 4 will describe the use of gamma-ray and cosmic-ray data in testing and constraining WIMP dark matter models. Chapter 5 will discuss dark hydrogen as a possible self-interacting dark matter candidate, and provide testable predictions for future observations.
Chapter 1

Introduction

Multiple independent measurements have found that visible matter accounts for less than 20% of the total mass in the Universe [112]. This discrepancy may be explained if the majority of the matter in the Universe consists of ‘dark matter’, a new, exotic particle that interacts gravitationally with visible matter but does not emit or absorb light. Much effort has been expended in laboratory and collider experiments to detect evidence of non-gravitational dark matter interactions with Standard Model particles, as observation of any such interactions would help to identify the dark matter particle and further our knowledge of its particle properties.

So far, no experiments have yielded any conclusive discoveries of the dark matter particle [72]. The current status of null results does not rule out the possibility of detection in future experiments: these are extremely difficult endeavors which require exquisite sensitivities and/or long experimental lifetimes, as the cross sections for dark matter-Standard Model interactions are miniscule at most. To overcome the difficulties associated with searching for evidence of such rare interactions, the ideal experiment (1) involve a large amount of dark matter and (2) have a long run time. In these two regards, astrophysical systems may
be thought of as large-scale dark matter laboratories, as they contain huge masses of dark matter and have been ‘running’ for cosmological timescales.

This thesis addresses various methods by which dark matter particle theories may either (1) be tested using current astronomical data, or (2) make testable predictions for future observations. The dissertation is comprised of two parts, each addressing a particular approach to searching for dark matter particle interactions via astronomical observations. Chapters 2 and 3 will describe high-energy searches for gamma-rays resulting from dark matter annihilations in our Milky Way. Chapter 5 will describe the puzzling mismatches between observations and theoretical predictions of dark matter structure on galactic scales, and how the dark hydrogen model with scattering ’self-interactions’ may provide an explanation for these discrepancies as well as testable predictions for the future.

To date, our only evidence for the existence of dark matter has come from astrophysical/cosmological observations. The first hints of the ‘missing mass’ problem came from dynamical motions of luminous matter in gravitationally bound systems. Most famously, Zwicky’s measurements of the Coma cluster’s virial mass and Rubin’s observations of rotation curve flattening in spiral galaxies suggested that the majority of the mass in astrophysical systems was nonluminous [237, 203, 204].

These dynamical mass measurements did not necessarily imply a need for a new particle beyond the Standard Model, as the missing mass might be accounted for by non-luminous baryonic matter in the form of massive compact halo objects (MACHOs), such as rocky planets. However, the MACHO hypothesis was mostly ruled out in the 1990s by gravitational microlensing surveys [24], which did not observe the number of MACHO lens events that would be expected if this discrepancy was due to MACHOs—no more than a few percent of the total Milky Way mass may be comprised of MACHOs.

Observations of the cosmic microwave background (CMB), the primordial deuterium abun-
dance, and merging clusters also provide strong evidence in favor of a new, massive subatomic dark matter particle. We may infer the baryon fraction of the Universe—i.e. the fraction of matter that interacts through the electromagnetic force and is coupled to the photons before recombination—from the relative amplitudes of the peaks in the angular temperature power spectrum of the CMB. The baryon density is approximately 16% of the total matter density of the Universe, which implies that the remainder of the matter content is non-baryonic dark matter [18]. Measurements of the baryon fraction (through the baryon-photon ratio) may also be derived from the primordial deuterium abundance from Big Bang Nucleosynthesis; these are also in agreement with the CMB results and also imply the need for particle dark matter [82]. Weak gravitational lensing has shown that the bulk of the mass in merging galaxy clusters does not scatter or interact with the other cluster during the merger, which tells us that the dark matter in these systems behaves much differently than ordinary baryonic matter [79].

The following sections are not a comprehensive overview of the vast landscape of proposed high-energy theories of particle dark matter; instead, we focus on two broad categories of dark matter theories.¹

- Weakly interacting massive particle (WIMPs): For the purposes of this thesis, the defining feature of WIMP dark matter is a weak scale ($O(10^{-26})$ cm$^3$/s) cross section for interaction between WIMP particles. This range of cross sections allows for annihilations of dark matter particles in astrophysical systems to produce observable fluxes of high-energy gamma-ray photons. Ch. 2-4 describe an anomalous excess in gamma-ray observations of our Milky Way’s center, which might be a signature of these WIMP annihilations.

- Self-interacting dark matter (SIDM): Dark matter particles can exchange kinetic energy by scattering off of each other. The cross section for these self-interactions can be

¹It should be noted these two categorizations are not mutually exclusive. For example, Ref. [111]
quite large \((\sigma/m \sim \mathcal{O}(1) \text{ cm}^2/\text{g} \sim \mathcal{O}(1) \text{ barn/GeV})\). This class of models makes unique predictions for the distribution of dark matter in astrophysical systems. Ch. 5 describes the predictions that atomic dark matter—a specific model of SIDM—makes for the density profiles of galaxies and how these may be tested with observations.

1.1 Indirect detection of weakly interacting massive particles (WIMPs)

Supersymmetric extensions of the Standard Model (SM) have received a great deal of attention in the past few decades from the high-energy particle theory community. These models are appealing because they offer solutions to unresolved issues in the SM (e.g. the hierarchy problem and gauge coupling unification) while also predicting the existence of a stable, neutral, massive particle that is an ideal candidate to play the role of cold and collisionless dark matter [140]. The lightest supersymmetric particle is stable on cosmological timescales and massive \((\gtrsim \text{TeV})\); such supersymmetric dark matter candidates are referred to as WIMPs (weakly-interacting massive particles). The term ‘weakly-interacting’ in reference to the assumed cross section for WIMPs’ hypothesized interaction with themselves and the SM. This presumed scale of interaction strength is motivated by the ‘WIMP miracle’: if the WIMP velocity-averaged cross section \(\langle \sigma v \rangle\) for self-annihilation in the early Universe is of order the weak scale \(\mathcal{O}(10^{-26}) \text{ cm}^3/\text{s}\), then the resulting relic density of dark matter after thermal freeze-out matches the observed present-day abundance of dark matter [216].

WIMP dark matter is also motivated from an astrophysical standpoint as it fits quite naturally into the ΛCDM (‘collisionless, cold dark matter’) paradigm. Currently, ΛCDM is the most widely-accepted model of the cosmos, in which Λ refers to cosmological constant dark energy. ΛCDM has been remarkably successful in predicting the statistical properties of fluc-
tuations in the cosmic microwave background [18] as well as the formation and distribution of structure in astrohphysical systems [214]. The descriptor ‘cold’ is taken to mean that the dark matter particle is non-relativistic during and after the start of structure formation. Dark matter particles in the ΛCDM model are effectively collisionless, i.e., they cannot exchange kinetic energy amongst themselves through self-scattering. Note that the ΛCDM cosmology does not assume any particular particle theory of dark matter; it is consistent with any dark matter particle with the generic properties of being cold and collisionless. However, since WIMPs are a subset of dark matter candidates with these properties, and they are well-motivated and well-studied in high energy particle physics, the term ‘WIMP’ is often used interchangeably with CDM when discussing dark matter in an astrophysical context.

ΛCDM: Predictions for dark matter halo densities

The ΛCDM model can be used to make many different predictions for the distribution and structure of astrohphysical dark matter; this overview on the following predictions as they are most relevant to the work in this dissertation.

N-body simulations of CDM structure formation predict that galaxies form at the centers of extended, spheroidal ‘halos’ of dark matter. The radial density profile $\rho(r)$ of dark matter halos in these simulations is found to follow the functional form

$$\rho(r) = \frac{\rho_s}{(\frac{r}{r_s})^\gamma (1 + \frac{r}{r_s})^{3-\gamma}}$$

which depends on the inner slope $\gamma$, scale radius $r_s$, and scale density $\rho_s$ [172]. The density profile described by Eqn. 1.1 is commonly referred to as the NFW profile. Given the mass of the NFW halo, the parameters $r_s$ and $\rho_s$ may also be estimated (to within some scatter) from the results of cosmological simulations [98]. In the limit of $r \ll r_s$, the density of an NFW
halo goes as $\rho(r) \propto r^{-\gamma}$ and rises steeply with decreasing radius. The inner slope parameter is taken to be $\gamma \approx 1$ from the results of N-body simulations [196, 174]. This predicts that CDM halos have centrally dense ‘cusps’ at their centers (as opposed to lower-density ‘cores’, which will be discussed later in Sec. 1.2).

**WIMP dark matter annihilation in astrophysical systems**

The average dark matter density of the Universe today is too low for WIMPs to annihilate at any observable rate. However, the central regions of NFW dark matter halos are dense enough for WIMP annihilation to continue into the present day. Two WIMP particles in the center of a dark matter halo may annihilate with each other, producing Standard Model final states. Depending on the resulting Standard Model particles, they might then hadronize and/or decay to produce high-energy gamma-ray emission. Thus, one method for detecting WIMP dark matter is to look at the central regions of dark matter halos in order to search for excess gamma-ray emission which may originate from WIMP annihilations. These searches are referred to as ‘indirect detection’ experiments\(^2\).

Our Milky Way’s galactic center is predicted to be the brightest potential source of gamma-rays from dark matter annihilation. For comparison, the the expected WIMP annihilation flux from the galactic center is approximately $10^3$ times brighter than the expected flux from the next brightest targets in the night sky, M31 and the dwarf satellite Segue I. Much attention has therefore been focused on studying the galactic center at gamma-ray energies in the hopes of detecting an excess originating from WIMP annihilation.

\(^2\)Other WIMP dark matter experiments are classified as direct detection (laboratory searches for scatterings between WIMPs and a SM nucleus or electron) or collider searches (searches for signatures of WIMP production from SM particles in an accelerator).
The Galactic Center Excess

Multiple studies of the galactic center have found strong evidence for an excess of gamma-ray emission beyond what is predicted by models of astrophysical gamma-ray emission [116, 6, 61, 126, 127, 6, 117, 164, 130, 4, 5, 236, 87, 61, 22, 16]. This finding has sparked significant interest as a potential dark matter signature as the galactic center excess has been unable to be modeled using any known source of astrophysical gamma-rays, such as cosmic ray interactions in dust and gas or inverse Compton upscattering. The observed properties of the gamma-ray excess were in line with expectations for a WIMP annihilation signal:

- The spatial morphology of the excess emission is consistent with the radial profile of prompt gamma-ray emission from an NFW dark matter halo.

- The spectrum of the excess emission is well-fit with the calculated spectra of WIMP dark matter with a mass $m_\chi \sim 10 - 40$ GeV promptly annihilating into $b\bar{b}$ quarks or $\tau^\pm$ leptons.

- The inferred annihilation cross section that would result in the observed excess flux is consistent (within a factor of few) to the velocity-averaged cross section $\langle \sigma v \rangle \simeq 2 - 3 \times 10^{-26}$cm$^3$/s that would also result in the observed relic density of dark matter in the Universe.

However, it is also possible that this excess might be produced by an unresolved population of $\mathcal{O}(1000)$ millisecond pulsars in the galactic center [3, 158].

Chapters 2-3 will describe indirect searches for dark matter using data from the Fermi gamma-ray space telescope. In particular, they focus on the ‘Galactic Center excess’ (GCE), an anomalous source of gamma-rays that might be attributable to WIMP dark matter annihilations in the Milky Way. Ch. 2 will describe a detailed characterization of the GCE.
gamma-ray excess and the systematics associated with modeling the astrophysical backgrounds. Ch. 3 will describe an infrared-correlated extended source of gamma-rays in the galactic center and its potential association with a high-energy electron population originating from the same source as the GCE.

1.2 Self-interacting dark matter

Does CDM fail at small scales?

As previously mentioned, the ΛCDM cosmological model has proven to be extremely successful in predicting the large scale structure of the universe. However, at smaller length scales (on the order of kiloparsecs), the results of CDM simulations seem to be in tension with our observations [57]. There are multiple areas of tension between CDM predictions and observations at these scales which motivate consideration of dark matter models with more complex interactions than standard WIMPS. This thesis focuses on two related issues—the ‘core-cusp problem’ and the ‘diversity problem’—that are related to the inner density profiles of dark matter halos.

The ‘core-cusp problem’ arises due to discrepancies between predictions of CDM halos’ radial density profiles and their observed density profiles. CDM N-body simulations predict ‘cuspy’ radial density profiles which are centrally dense and steeply rising with decreasing radius ($\rho \sim r^{-1}$) [173, 58, 95, 215]. On the other hand, observations of rotation curves in dwarfs and low surface brightness (LSB) galaxies imply that their dark matter halos are ‘cored’ within radii of $\sim$1 kpc, having roughly constant density ($\rho \sim$ constant) [171, 189, 89, 88, 218, 208, 212, 154, 90, 96, 180, 17].

Additionally, galactic rotation curves show a variety of behavior in the inner parts even across systems with similar halo and stellar masses. The ‘diversity problem’ describes the
inability of CDM simulations to produce a similarly wide range of inner dark matter density profiles. Fig. 1.1 illustrates the diversity of galactic rotation in similar halo mass bins.

Figure 1.1: The diversity of rotation curves across a range of spiral galaxy masses. Each panel contains 14 colored galaxy rotation curves that are selected to have similar flat rotation velocities $v_{\text{flat}}$ at their furthest radial data points. From left to right, the corresponding $v_{\text{flat}}$ bins are $79 - 91$ km/s, $139 - 172$ km/s and $239 - 315$ km/s, spanning most of the mass range of the galaxies considered in this work. The points with error bars are rotation curve data from the SPARC dataset [159]. Solid lines are the halo model fits to the data using a SIDM model with cross section $\sigma/m = 3 \text{ cm}^2/\text{g}$.

**Potential solutions to small-scale problems in CDM**

One proposed solution to CDM small-scale issues is to include the baryonic feedback effects from supernovae and/or black holes in simulations [118, 55, 163, 215, 179, 185, 148, 198, 228]. If the energy injected into the baryonic matter by these processes significantly alters the baryons’ gravitational potential, the density profile of the dark matter in the halo can be lowered in response to this fluctuating baryonic potential. In this way, baryonic feedback may be able to affect the dark matter distribution in smaller galaxies, even if there is no coupling between the Standard Model baryons and the dark matter particle. However, realistic feedback models are difficult and resource-intensive to implement in cosmological simulations; it is yet unclear whether baryonic effects are able to alleviate the aforementioned small-scale CDM issues across the range of galaxies simulated [106, 227, 188, 92, 91, 184, 186].

Another possible solution is to consider self-interacting dark matter (SIDM) particle models that allow for self-interactions between particles as a means for transferring kinetic en-
ergy [213]. Scatterings between SIDM particles in a halo can transfer heat from the outer halo into the colder, denser central regions; after the SIDM halo has reached equilibrium, the resulting central core region is hotter and less dense than in an equivalent CDM halo.

Chap. 5 will discuss atomic dark matter, a specific SIDM model which may resolve small-scale structure puzzles while also making unique predictions for the halo mass function and the velocity-dependence of SIDM scattering cross sections.
Chapter 2

Spatial Non-Uniformity of the Galactic Center Excess

Based on Horiuchi et al., JCAP 1611 (2016) no.11, 053 [132].

We perform a composite likelihood analysis of subdivided regions within the central $26^\circ \times 20^\circ$ of the Milky Way, with the aim of characterizing the spectrum of the gamma-ray galactic center excess in regions of varying galactocentric distance. Outside of the innermost few degrees, we find that the radial profile of the excess is background-model dependent and poorly constrained. The spectrum of the excess emission is observed to extend upwards of 10 GeV outside $\sim 5^\circ$ in radius, but cuts off steeply between 10–20 GeV only in the innermost few degrees. If interpreted as a real feature of the excess, this radial variation in the spectrum has important implications for both astrophysical and dark matter interpretations of the galactic center excess. Single-component dark matter annihilation models face challenges in reproducing this variation; on the other hand, a population of unresolved millisecond pulsars contributing both prompt and secondary inverse Compton emission may be able to explain the spectrum as well as its spatial dependency. We show that the expected differences in the
photon-count distributions of a smooth dark matter annihilation signal and an unresolved point source population are an order of magnitude smaller than the fluctuations in residuals after fitting the data, which implies that mismodeling is an important systematic effect in point source analyses aimed at resolving the gamma-ray excess.

2.1 Background

Fermi Large Area Telescope (LAT) observations towards the Milky Way center have revealed a spatially extended source of gamma rays in excess of the modeled astrophysical backgrounds [116, 222, 126, 127, 6, 117, 164, 130, 4, 5, 236, 87, 61, 22]. This ‘galactic center excess’ (GCE) has so far been found to be robust against variations in background modeling [81, 164, 236, 4, 61, 22]. Possible explanations for this excess include weakly-interacting massive particle (WIMP) dark matter annihilations, unresolved millisecond pulsars (MSPs), and cosmic-ray outbursts from the galactic center.

The interpretation of the GCE as emission from dark matter annihilations has raised considerable interest due to the findings of Refs. [116, 6, 117, 164, 4, 87, 5, 61, 22] that point out (1) the spatial morphology of the GCE is consistent with that WIMP annihilations in a Navarro-Frenk-White (NFW) dark matter halo, (2) the spectrum of the GCE is consistent with predictions for WIMP annihilations into Standard Model particles, and (3) the annihilation cross sections required to fit the modeled spectra to the data are of the same order as the weak-scale annihilation cross section that results in the observed relic abundance of dark matter. However, the excess emission may also be attributed to astrophysical sources. A large unresolved population of millisecond pulsars remains a viable astrophysical explanation for the excess emission [6, 169, 117, 192, 232, 231, 52, 182, 183]: the typical MSP spectrum, as observed in globular clusters, is consistent with the observed GCE spectrum [3], and the spatial distribution of low-mass X-ray binaries (which are thought to be an earlier evolu-
tionary phase of MSPs) is consistent with an NFW-like power law, at least in M31 \cite{232}.

Additionally, the central regions of the Milky Way have experienced violent eruptions in the past, as evidenced by the lobed structures of the Fermi bubbles emanating from the galactic center \cite{217}; this history of burst activity has motivated authors to consider cosmic-ray injection events as another possible astrophysical explanation for the GCE \cite{66, 191, 110, 69}.

Any spectral or spatial variation (or lack thereof) in the signal would be of critical importance in discerning amongst the possible origins of the GCE. For example, a prompt dark matter annihilation signal—where the subsequent decay and hadronization of the Standard Model products occurs quickly and the ensuing gamma rays are emitted at the site of annihilation—should have an intensity directly proportional to the square of the dark matter density profile, and the spectrum should be independent of sky position. On the other hand, if the GCE was at least partially produced through inverse Compton (IC) scattering from a population of high-energy leptons—which is possible in the cases of MSPs \cite{192, 231}, dark matter annihilations to leptons \cite{155, 60, 145, 156}, or a leptonic cosmic-ray outburst—its spectrum and intensity would be dependent on cosmic-ray diffusion processes in the central Milky Way as well as the interstellar radiation field (ISRF). We might thus expect to observe some variation in the spectral shape and normalization as a function of sky position if the GCE source was (at least partially) leptonic.\footnote{Hadronic cosmic-ray outbursts may also impinge upon gas and produce gamma rays through subsequent $\pi^0$ decays and bremsstrahlung processes; however, the gamma rays produced in this scenario will trace the gas distribution and thus have a disk-like, not spherical, morphology.}

The GCE has been observed within the innermost few degrees of the Milky Way \cite{116, 6, 117, 164, 4, 87, 5} (hereafter referred to as the ‘galactic center’) as well as the region immediately exterior to the galactic center (hereafter referred to as the ‘inner galaxy’) \cite{87, 61}. In this paper, we compare the best-fit spectra and morphologies across multiple regions, including the galactic center, using consistent diffuse background models and fitting procedures between

\footnote{The population in M31 is used instead of the Milky Way as the current INTEGRAL catalog of low-mass X-ray binaries in the Milky Way bulge has substantial completeness concerns \cite{44}.}
analyses of each region. The key idea is to use the morphologies of the diffuse backgrounds to constrain the spectrum and thereby investigate the spatial uniformity and photon-count distribution of the excess. We discuss our results in terms of their implications for both dark matter and astrophysical interpretations of the GCE. In particular, we investigate (1) the presence of a power law-like feature in the GCE spectrum, with emission extending upwards of $\sim 20$ GeV \cite{61, 22} and (2) the consistency of the GCE with a population of unresolved MSPs \cite{32, 158}.

Analyses of the excess in the inner galaxy report a power law-like high-energy tail in the GCE spectrum beyond 10 GeV, possibly extending upwards of $\sim 20 - 100$ GeV \cite{60, 61, 22}. If the GCE is a bona fide signal from dark matter annihilations, multiple particle properties (mass, annihilation primaries, branching ratios) are encoded within the shape of its gamma-ray spectrum. Inclusion or exclusion of the high-energy tail in the GCE spectrum can greatly affect the best fit dark matter mass and annihilation channel(s). If the GCE has an astrophysical origin, the presence of high-energy emission could inform us about the processes that gave rise to it and perhaps rule out certain scenarios. In Sections 2.3.4 and 2.4.1 we investigate whether the high-energy tail of the GCE spectrum originates from the same source that produces its spectrum below $\sim 10$ GeV.

Recent results support an unresolved point source origin for the GCE and indicate that such sources may be able to account for the entirety of the excess in the inner galaxy \cite{158, 32}. If so, this would strongly imply that most, if not all, of the GCE signal is produced by millisecond pulsars, not dark matter annihilation. This would be evidence of an as-yet-undiscovered pulsar population at the galactic center with exciting implications for astronomy across the electromagnetic spectrum. The MSP interpretation of the excess would also set strong upper limits on the WIMP annihilation cross section. In Section 2.4.2 we attempt to determine whether our findings are suggestive of either an unresolved point source distribution or annihilation in a smooth NFW halo.
2.2 Methods

We use approximately 73 months of Pass 7 data from the Fermi-LAT taken between August 2008 and September 2014.\textsuperscript{3} We use CLEAN-class photon events and the Pass 7 reprocessed instrument response functions. The photon events range from 700 MeV–200 GeV and are binned in 8 logarithmically spaced bins from 700 MeV–10 GeV and 3 high-energy bins from 10–200 GeV. We use larger bin sizes above 10 GeV to compensate for the lower photon counts at high-energies. We apply a maximum zenith angle cut of 90° to avoid contamination from Earth limb emission. We define our ‘inner galaxy’ regions of interest (ROIs) and ‘galactic center’ ROIs in Tab. 2.2 as well as Fig. 2.1.

Many studies of the galactic center emission, including this work, rely on spatial template-based analyses in which the GCE spectrum is fit alongside the spectra of the background diffuse emission and point sources. The strengths of such analyses lie in their ability to effectively subtract out the bulk of the astrophysical backgrounds from the data. However, this method inherently introduces systematic effects by assuming a given spatial profile for each diffuse background source. Potentially large errors in the best-fit GCE spectrum may arise if the spatial templates assumed for the backgrounds differ considerably from the true background emission. For this reason we test three different diffuse background models (described below in Sec. 2.2.1) to estimate the systematic error in the GCE spectrum due to the uncertainty in our assumptions about the astrophysical background. Our results are valid and robust under the three models tested below, but note that the range of backgrounds tested here is more limited than used in previous works [236, 61].

\textsuperscript{3}This study uses the Pass 7 data as it was commenced before the public release of the Pass 8 dataset. Ref. [161] show that using Pass 7 versus Pass 8 data has a negligible effect on the GCE spectrum, including the higher energies.
2.2.1 Fit components

The gamma-ray observations are modeled as a combination of the following source templates:

**Diffuse gamma-ray background:** The primary diffuse astrophysical gamma-ray background is produced by the following processes:

- Neutral pion ($\pi^0$) decay: Neutral pions are produced when hadronic cosmic rays impinge upon clouds of gas in the interstellar medium (ISM). The $\pi^0$'s subsequently decay into pairs of high-energy photons.

- Bremsstrahlung radiation: High-energy electron cosmic rays interact with gas in the ISM.

- Inverse Compton radiation: High-energy electron cosmic rays upscatter lower energy background starlight photons in the interstellar radiation field.

We model the gamma-ray emission from the above processes using the WebRun interface of the GALPROP (version 54) cosmic-ray propagation code [2, 194, 223], which computes the diffusion and energy losses for a chosen set of propagation parameters and outputs the resultant gamma-ray skymap templates and spectra. For a given diffusion model, we generate the emission templates from $\pi^0$ decay and bremsstrahlung separately and then combine them into a single $\pi^0$+bremsstrahlung diffuse template. The spatial distribution of background emission from $\pi^0$ decay and bremsstrahlung radiation is very similar because both processes require the same gas cloud target. If individual templates are included for the $\pi^0$ and bremsstrahlung emission, the large degeneracies between the two spatial morphologies would make it difficult for the likelihood maximization to correctly fit the spectrum of each component. We therefore fit a single, combined $\pi^0$+bremsstrahlung diffuse template in each
energy bin to avoid the inclusion of two templates with largely degenerate morphologies. The $\pi^0$-bremsstrahlung and IC components are fit independently of each other.

To test the robustness of our results, we repeat the analysis with three different diffuse gamma-ray background models generated using GALPROP. We use models selected from the suite of diffuse backgrounds tested by Ref. [61] in their systematic analysis of the GCE signal. For consistency and ease of comparison between works, we refer to the background models using the same labelling (A/E/F) as in Ref. [61]. The variations in the input parameters for our diffuse backgrounds are listed in Tab. 2.1. GALPROP model A is chosen for testing as it is ‘tuned’ such that the recovered best-fit template normalizations after fitting to the data agree well with the GALPROP prediction. Model F is chosen as it was found to provide the highest likelihood fit in the inner galaxy between $2^\circ < |b| < 20^\circ$. We chose to test GALPROP model E as an extreme case: the low diffusion coefficient $D_0$ in this model leads to a large bump in the IC spectrum below 10 GeV as well as different spatial morphologies compared to models A and F. The effects of fitting with this extreme background are further discussed in Sec. 2.3.1. For a detailed description of the effects of varying diffuse model parameters on the characterization of the GCE, see Refs. [61, 22].

GCE template: The GCE is well-fit by annihilation signals based on NFW profiles [173], which approximate cold dark matter halo densities in N-body simulations. We therefore base our set of GCE templates upon the signal morphology that is predicted for annihilations in an NFW halo, which is proportional to the density squared, integrated along the line of sight. It should be noted that the spatial profile we assume in our template model for the GCE is not unique to dark matter annihilations, and may also be consistent with a central MSP population [6]. The dark matter density profiles in N-body simulations have been found to follow the functional form

$$\rho(r) = \frac{\rho_s}{(r/r_s)\gamma[1 + r/r_s]^{(3-\gamma)}}.$$  \hspace{1cm} (2.1)
Table 2.1: Input parameters for our set of three GALPROP diffuse background models. We use the same scale radius $r_D=20$ kpc and Alfvén speed $v_A=32.7$ km s$^{-1}$ for all models. The scale height $z_D$ is given in units of kpc. The diffusion coefficient $D_0$ is given in units $\times 10^{28}$ cm$^3$ s$^{-1}$. The convection velocity gradient $dv/dz$ is given in km s$^{-1}$ kpc$^{-1}$. The cosmic-ray source distribution is taken from either the measured supernova remnant (SNR) distribution [68] or the Lorimer pulsar distribution [162]. (Both of these cosmic-ray distributions approach zero at the galactic center and are in all likelihood severely underestimating the cosmic-ray source density in the innermost kpc. We discuss the implications of this deficiency in Sec. 2.3.1.) The power law index of the electron (proton) injection spectrum above rigidity 2.18 (11.3) GV is given by $\alpha_e$($\alpha_p$). The electron (proton) cosmic-ray injection spectrum is normalized to $N_e$($N_p$) in units $\times 10^{-9}$ cm$^{-2}$ sr$^{-1}$ s$^{-1}$ MeV$^{-1}$ at 34.5 (100) GeV. The first set of three digits in the magnetic field model are $B_0 \times 10$ µG, the second set of three digits are $r_c \times 10$ kpc, and the last set of three digits are $z_c \times 10$ kpc. ISRF normalization factors are given for the optical, IR, and CMB components respectively. The gas spin temperature $T_S$ is in units of K. A full description of the parameters may be found in Refs. [10, 65].

<table>
<thead>
<tr>
<th>Model</th>
<th>$z_D$</th>
<th>$D_0$</th>
<th>$dv/dz$</th>
<th>CR Source</th>
<th>$\alpha_e/\alpha_p$</th>
<th>$N_e/N_p$</th>
<th>B-field</th>
<th>ISRF</th>
<th>$T_S$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>4</td>
<td>5.0</td>
<td>50</td>
<td>SNR</td>
<td>2.43/2.47</td>
<td>2.00/5.8</td>
<td>0900500020</td>
<td>1.36/1.36/1.0</td>
<td>150</td>
</tr>
<tr>
<td>E</td>
<td>4</td>
<td>2.0</td>
<td>0</td>
<td>SNR</td>
<td>2.43/2.39</td>
<td>0.40/4.9</td>
<td>0501000020</td>
<td>1.0/1.0/1.0</td>
<td>150</td>
</tr>
<tr>
<td>F</td>
<td>6</td>
<td>8.3</td>
<td>0</td>
<td>PLS</td>
<td>2.42/2.39</td>
<td>0.49/4.8</td>
<td>0501000020</td>
<td>1.0/1.0/1.0</td>
<td>$10^5$</td>
</tr>
</tbody>
</table>

The density profile is normalized to the local dark matter density at the solar position, $\rho_\odot = 0.3$ GeV cm$^{-3}$ [235], with scale radius $r_s = 23$ kpc. The log slope of the NFW density profile asymptotes to the inner slope $\gamma$ as $r$ approaches the halo center. In regions lying outside the central few degrees, such as our inner galaxy ROIs, the density slope begins to deviate from the asymptotic inner value $\gamma$. Thus, there is some degeneracy between the NFW inner profile slope $\gamma$ and the scale radius $r_s$ as the region of interest moves away from the galactic center. In outlying regions (but not the galactic center), a GCE template with a shallower inner slope $\gamma$ and smaller scale radius $r_s$ may be similar in morphology to a template with a steeper inner slope and larger scale radius. Since our aim is to describe the morphology preferred by the inner galaxy excess (rather than infer the parameters of the NFW-like profile), we fix the scale radius to $r_s = 23$ kpc.

For a source originating from Majorana dark matter annihilations with velocity-averaged cross section $\langle \sigma v \rangle$, the differential flux received along a line of sight towards galactic coordi-
\[
\frac{d\Phi(l, b)}{dE} = \frac{1}{4\pi m_X^2} \frac{\langle \sigma v \rangle}{2} J(l, b) \frac{dN_\gamma}{dE},
\]  
(2.2)

where \( m_X \) is the dark matter particle mass, and \( dN_\gamma/dE \) is the gamma-spectrum per annihilation. The quantity \( J(l, b) \), commonly referred to as the ‘J-factor’, depends on the astronomical dark matter distribution and is equal to the mass density squared, integrated over the line of sight \( x \) through \((l, b)\):

\[
J = \int_{l.o.s} \rho^2(r_G)(x, l, b) \, dx
\]  
(2.3)

where \( r_G = \sqrt{R_\odot^2 - 2xR_\odot \cos(l) \cos(b) + x^2} \) is the distance from the galactic center. For our ROIs close to the Milky Way center, \( r \ll r_s \) so that \( \rho(r) \propto r^{-\gamma} \). We use the value \( R_\odot = 8.25 \) kpc for the solar distance to the galactic center.

As previously discussed, one of our goals is to test whether the GCE in the galactic center and inner galaxy regions can be described with a single NFW annihilation profile. We use NFW annihilation templates with inner slopes \( \gamma = \{0.9, 1.0, 1.1, 1.2, 1.3\} \) when fitting the GCE in the galactic center and inner galaxy to test whether any of these profiles yields consistent fluxes between various ROIs. We also test a template with \( \gamma = 0.8 \) for the GALPROP model F background in the inner galaxy, which is the one case we find where the likelihood favors shallower NFW profiles.

**Fermi bubbles:** The Fermi bubbles are a diffuse, lobed source extending up to \( \sim 50^\circ \) North and South in latitude from the galactic center [217]. The bubbles are found by Ref. [12] to have a hard spectrum of \( dN/dE \propto E^{-1.9\pm0.2} \) with a high-energy cutoff around 100 GeV. We employ a flat emission template with edges defined as in Ref. [217] because our GALPROP-generated diffuse backgrounds do not model this extended source. The use of a uniform spatial template is motivated by Refs. [217, 12]’s finding that the bubbles’
intensity is approximately flat in projection. As the bulk of the Fermi bubbles’ emission lies at farther latitudes outside our ROIs, we use additional regions defined by \(330^\circ \leq l \leq 20^\circ\), \(20^\circ \leq |b| \leq 35^\circ\) in the northern and southern galactic hemispheres to externally constrain their spectrum within the ROI. These regions were chosen to lie outside of \(b = 20^\circ\) to avoid overlap with other subregions. Although the northern lobe of the bubble template extends into our galactic center ROI, we do not include this template as part of the galactic center fit as (1) its spatial profile becomes uncertain at low latitudes and (2) its flux per steradian is subdominant to other extended components in the galactic center.

**20 cm gas template:** We include a gas template for the galactic ridge structure as previously described in Refs. [21, 233, 164, 4, 5]. This emission, which is correlated with 20 cm radio emission as well as \(\sim\)TeV gamma-ray emission in the central region, has been interpreted by Refs. [233, 164, 4, 5] as bremsstrahlung emission from a population of high-energy electrons in the galactic center.

**WISE 3.4 \(\mu m\) template:** We include a template tracing the infrared starlight emission in the galactic center. This component was interpreted in Ref. [5] to be IC emission (in excess of the GALPROP predicted IC flux) from a population of high-energy leptons. We do not find any flux associated with this template in our analysis, which is consistent with our previous finding in Ref. [5] which found that the spectrum of the component had a steep cutoff around \(\sim\)400–500 MeV, which is below the minimum energy of our analysis. This is also found in the work of Ref. [156] using self-consistent GALPROP modeling of the additional IC emission. Thus, we do not show any spectrum for this component in Figs. 2.2 and 2.3.

**Point sources:** We include point sources from the Fermi 3FGL point source catalog [9] that lie within or near the regions of interest (ROIs). Sources listed in the catalog with significance \(\sigma > 5.0\) are free to have their spectra varied.
Isotropic gamma-ray background: We do not assume a fixed spectrum for the isotropic extragalactic background, but fit the normalization of a uniform isotropic background template independently in each energy bin. We note that Refs. [4, 87, 5] find evidence for an isotropic or close to isotropic component in the innermost $7^\circ \times 7^\circ$ that is somewhat brighter than the Fermi collaboration’s standard extragalactic isotropic background template [14]. Refs. [4, 5] fit this component with either an additional isotropic template (Ref. [4]) or a nearly isotropic ‘new diffuse’ (ND) template (Ref. [5]). Hence, we test the case where the isotropic component in the galactic center region is allowed to vary separately from the isotropic component in the inner galaxy.

Figure 2.1: Total observed counts map with labelled regions of interest. Each ROI will be referred to hereafter using its label from this figure. We will collectively refer to the entirety of the ROIs excluding the innermost $7^\circ \times 7^\circ$ as the ‘inner galaxy’ (orange). The innermost $7^\circ \times 7^\circ$ ROI is referred to as the ‘galactic center’ (blue). The overlap of the Fermi bubble template used in this analysis (c.f. Sec. 2.2.1) with the ROIs is shown as the gray overlay. Not shown here are the farther latitude N2/S2/bubble N/bubble S ROIs which were used to constrain the Fermi bubble spectrum and estimate the extent of the GCE signal.
Table 2.2: Our regions of interest, as defined by range in galactic longitude \( l \) and latitude \( b \). Angular areas in steradians are also given, although all our results for best-fit GCE flux in each ROI are normalized to display the total expected flux \((\text{GeV s}^{-1} \text{ cm}^{-2})\) from the \( 35^\circ \times 35^\circ \) GCE template, based on the observed flux for each individual ROI. Note that the farthest latitude regions ‘bubble N/S’ and ‘N2/S2’ are not included in the GCE analysis and thus not shown in Fig. 2.1; they are included solely for the purpose of constraining the Fermi bubbles’ spectrum and testing the extent of the GCE signal.

### 2.2.2 Fit procedure

We use the Composite2 tool within the Fermi Science Tools Python interface [1] to perform a composite likelihood analysis of multiple ROIs simultaneously for each energy bin. This allows for any number of chosen model parameters—e.g. flux normalization of the diffuse background components in the chosen energy bin—to be tied across multiple ROIs, while still allowing for the possibility that other extended sources—e.g. the GCE template—might be fit with different normalizations between ROIs. We constrain the normalization of the extended astrophysical sources (GALPROP \( \pi^0 \) + bremsstrahlung diffuse, GALPROP IC diffuse, Fermi bubbles, and isotropic background templates) to be the same throughout all ROIs in Tab. 2.2. The origin of the GCE is yet unknown and we do not presume that
it must be fit with a single spectrum and template normalization across all regions. Thus, we allow the GCE template to be fit with different normalizations in the individual inner galaxy and galactic center ROIs shown in Fig. 2.1.

We perform purely spatial fits to the data within each independent energy bin, i.e., we do not require the modeled sources to follow any fixed spectral shape or parameterized functional form across multiple energy bins. This is also true for the galactic diffuse templates and isotropic background, which are typically constrained to have a fixed spectral shape. We note that the GALPROP code does give a prediction of the spectrum for each diffuse background component; however, we do not constrain the normalization of the diffuse templates to follow the GALPROP-predicted spectral shapes when fitting.

Our choice of methodology does entail some caveats. As previously mentioned, the $\pi^0$+bremsstrahlung and IC diffuse backgrounds are not fixed to the broadband spectral shapes predicted by GALPROP for each of the models. Therefore, the best-fit spectrum for either of these diffuse components may be unphysical in the sense that it does not necessarily correspond to the GALPROP parameters that produce the spatial profile it is associated with. In principle, fitting the background in independent energy bins allows more freedom for the diffuse backgrounds to absorb the GCE component. However, we find that for extreme background model parameters—such as in model E, where the low diffusion coefficient leads to a large modification in the IC component—the modeled background is a poor fit to the actual gamma-ray diffuse background, which causes the fitting procedure to lower the normalization of the background model in favor of increasing the normalization of the GCE or other extended templates (see Sec. 2.3.1). We also note that this analysis does not include a template for the large scale feature Loop I in the northern galactic sky. If this omission affected the derived GCE spectrum in the inner galaxy, we would expect to observe lower intensity in the best-fit GCE spectrum in our northern ROIs relative to the south; however, we show in Fig. 2.7 that this is not the case.
2.3 Results

In Sec. 2.3.1 we describe the systematic variations in the best-fit GCE spectra associated with the diffuse background model components, and in particular, the GALPROP-generated IC templates. Sec. 2.3.2 compares and contrasts the best-fit GCE spectra in the galactic center and combined inner galaxy regions. In Sec. 2.3.3 we discuss how the choice of GALPROP background model affects the best-fit NFW slope and the GCE residual radial profile. Finally, we examine the spatial uniformity of the GCE across the separate ROIs in Sec. 2.3.4.

2.3.1 Systematics associated with background model components

Fig. 2.2 shows the spectra of the NFW template and diffuse background model components for the three different GALPROP model backgrounds. GCE spectra are shown for the cases where the excess was fit using an NFW template with slope $\gamma = 1.1$, which was found to be the favored value of $\gamma$ in all but one of the fits (see Sec. 2.3.3). From the upper row of Fig. 2.2 we see that when the spectral shapes of the isotropic and IC diffuse backgrounds are not fixed (but are instead allowed to vary in normalization in independently-fit energy bins), these components are severely under-fit in the galactic center ROI. This indicates that the generic diffuse background models calculated by the GALPROP code are not able to adequately model the spatial distribution of diffuse emission in the innermost $\sim\text{kpc}$ of the galaxy.

The default spatial distributions for cosmic-ray injection used to model the IC emission in GALPROP are peaked between $r \sim 2 - 5\text{ kpc}$ (depending on the model used). As pointed out by Refs. [110, 64, 65], insertion of a strong source of cosmic rays at the galactic center affects the diffuse background modeling and could thus also affect the derived characteristics of the gamma-ray excess. Using a specialized, local model for the IC emission close to the galactic
center, the Fermi collaboration [22] finds that this component is strongly enhanced relative to previous diffuse background models; this suggests that the spatial models of cosmic-ray lepton injection in GALPROP are deficient within the innermost kpc and do not produce an accurate representation of the IC emission there.

![Diagram](image)

Figure 2.2: Spectra for the GCE and spatially extended background model components in the galactic center (top row) and combined inner galaxy ROIs (bottom row). Fits were performed with the GALPROP IC and $\pi^0$+bremsstrahlung templates free to vary independently of each other in each energy bin. Red ★ symbols denote the best fitting background model in the respective regions.

The dropout of the diffuse galactic center IC background below 10 GeV raises the concern that the under-modeling of this component might be causing photons from this source to be falsely attributed to the GCE, and that a significant portion—if not all—of the GCE in the galactic center is simply misattributed IC background emission. To test whether this is the case here, we combine the IC and $\pi^0$+bremsstrahlung template into a single template which matches the GALPROP prediction for each component. We then repeat the bin-by-bin template fitting in the galactic center ROI using this single IC+$\pi^0$+bremsstrahlung
Figure 2.3: Same as the top row of Fig. 2.2, but with the fits performed with the IC and $\pi^0$+bremsstrahlung templates constrained to have the same relative normalizations to each other as predicted by GALPROP. The dashed magenta line plots the GCE spectrum from Fig. 2.2 (where the IC and $\pi^0$+bremsstrahlung were fit separately) for comparison. Note that the dashed comparison NFW annihilation spectrum is indeed plotted in the model F panel but is difficult to see because of its close overlap with the solid magenta NFW spectrum (where the diffuse templates are fixed relative to each other).

diffuse template. The results of these fits are shown in Fig. 2.3. By constraining the IC background component to be fixed to its predicted intensity relative to the bright, more easily-fit $\pi^0$+bremsstrahlung component, we are able to recover a physically realistic spectrum for the IC background in the galactic center.

We see in Fig. 2.3 that fitting with a combined IC+$\pi^0$+bremsstrahlung diffuse template causes the GCE spectrum to change at most by a factor of two downwards compared to the case where the IC template normalization was allowed to vary freely. We therefore caution that there may be some degeneracy between the GCE and GALPROP IC components, depending on the chosen background model. The comparison of the GCE spectrum with and without the IC+$\pi^0$+bremsstrahlung templates tied is also shown in Fig. 2.3, where the dashed magenta line is the NFW annihilation template spectrum from Fig. 2.2.

In our inner galaxy fits, we find that the GCE spectrum has a more pronounced bump as well as a slightly higher peak normalization at $\sim$2 GeV in fits where the model E diffuse background was used. In the bottom row of Fig. 2.2 we see that this bump feature in the GCE spectra is accompanied by a corresponding dip in the IC diffuse background at the same energies (relative to the best-fit IC spectrum in the other model backgrounds). This suggests
that the GALPROP-generated spatial templates for the IC diffuse background at energies $\lesssim 2$–3 GeV are a very poor description of the true IC emission in the inner galaxy—so much so that the likelihood fitting procedure finds that a large fraction of the GALPROP-predicted IC emission is better fit by the NFW template than the model E IC template.

### 2.3.2 The GCE spectrum in the galactic center versus the inner galaxy

Fig. 2.4 shows the GCE spectrum in the galactic center versus the combined inner galaxy ROIs for all combinations of background diffuse models and NFW templates. We highlight the panels in Fig. 2.4 which correspond to the highest likelihood background model and NFW template combinations as recorded in Tab. 2.3 for the galactic center (gray) and inner galaxy (light orange). The flux in each region is scaled by the J-factor of the entire NFW template divided by the J-factor of the plotted ROI. Thus, all subplots show the expected flux for the entire GCE template ($35^\circ \times 35^\circ$), which allows for easier comparison between different regions: if both the galactic center and inner galaxy are consistent with a single NFW-like source, then their data points should have the same normalization in Fig. 2.4. With this scaling applied, it is apparent that for the best-fitting GALPROP backgrounds and NFW profile slopes, the peak intensity of the observed GCE spectrum in the galactic center and inner galaxy regions is consistent with originating from a single NFW source.

It is also evident from Fig. 2.4 that the shape of the GCE spectrum in both the galactic center and inner galaxy ROIs is remains consistent throughout the various combinations of GALPROP backgrounds and NFW profiles used in this analysis. The spectrum in the galactic center agrees with the results of previous studies confined to the innermost few degrees of the Milky Way [4, 5, 117], where the GCE had a steep cutoff before $\sim 10$–20 GeV. In contrast to our galactic center results, the GCE in the inner galaxy does not exhibit any
Figure 2.4: Best-fit GCE spectra in the galactic center (black squares) and inner galaxy (orange triangles) regions, shown for varied GALPROP diffuse models (rows) and NFW density profile slopes $\gamma$ (columns). The spectrum of the GCE in the inner galaxy is shown for the sum of all inner galaxy ROIs. **Normalizations are scaled to show the expected flux for the entire GCE template** ($35^\circ \times 35^\circ$), such that the normalization for the two ROIs will match if they are consistent with originating from a single NFW-distributed source. Also shown are the exponential cutoff parameterized fits (gray dashed line) to the galactic center spectrum. The panel with the light gray (orange) background denotes the NFW slope and diffuse background combination with the highest likelihood fit for the galactic center (inner galaxy) ROI as recorded in Tab. 2.3. 28
spectral cutoff and still shows significant flux at energies $\gtrsim 10$ GeV. This high-energy tail, as referenced in Sec. 2.1, is also robust to model variations and is present in all combinations of diffuse backgrounds and NFW templates tested here. We further discuss the significance of this finding in Sec. 2.4.1.

### 2.3.3 The GCE spatial profile and radial distribution

Tab. 2.3 gives the change in NFW test statistic value ($TS = -2\Delta \ln L$) for each combination of GALPROP diffuse model and NFW template slope. The differences in the TS values are given relative to the model with the highest TS value for the NFW template. In the galactic center ROI, we find that the data is best fit with GALPROP diffuse background A and NFW template slope $\gamma=1.1$. However, if using the typical cut of $TS>25$ to determine significance, $\gamma=1.1$ is not significantly favored over $\gamma=1.2$ in the galactic center. Within the inner galaxy, the highest GCE TS values correspond to fits using GALPROP model F. The NFW template in the inner galaxy favors shallower profiles with $\gamma \leq 0.9$ when fitting with the model F diffuse background. However, we find that $\gamma=1.1$ is the best fit NFW template in the inner galaxy when using the less-favored background models A and E. We thus conclude that the slope of the NFW density profile is poorly constrained in the inner galaxy, and variations in diffuse background modeling can have large effects on the best-fit NFW profile slope in that region.

In Fig. 2.5 we plot the change in TS value as a function of NFW slope $\gamma$ for each GALPROP diffuse model fit in the inner galaxy. The total $-\Delta TS$ is broken down into its contributions from energy bins below 1.9 GeV, 1.9–10 GeV, and above 10 GeV. For inner galaxy fits using GALPROP diffuse model F, we see that the preference for shallow profile slopes is most strongly driven by the low energy end of the GCE spectrum below 2 GeV.

Our weak constraint on the NFW profile slope $\gamma$ in the inner galaxy is seemingly in contrast
with the findings of previous works which strongly favor spatial profiles for the GCE with \( \gamma \sim 1.1–1.3 \) in the inner galaxy, with little to no dependence on background modeling [87, 61]. We attribute this discrepancy to differences in the minimum galactocentric distance used by various authors to define their region of interest. Ref. [61] derive their constraints on \( \gamma \) by analyzing the region defined by \( 2^\circ \leq b \leq 20^\circ \) and \( l \leq 20^\circ \), while Ref. [87] use \( 1^\circ \leq b \leq 20^\circ \) and \( l \leq 20^\circ \). Our ‘galactic center’ ROI—within which we find that \( \gamma \) is consistently 1.1–1.2 for all GALPROP diffuse models—overlaps with the regions in these works between latitudes of \( 1^\circ – 2^\circ \leq b \leq 3.5^\circ \). As we show in the top row of Fig. 2.6, the radial profile of the excess below 10 GeV is largely insensitive to changes in the GALPROP background model out to \( r \sim 5 – 6^\circ \).

Fig. 2.6 shows the radial distribution of flux in the GCE-associated residual for energies below 1.9 GeV, 1.9–10.0 GeV, and above 10 GeV. The top row of Fig. 2.6 plots this quantity for all three GALPROP backgrounds with the best-fit NFW profile residuals in the galactic center and inner galaxy. Below 10 GeV, the radial profile of the GCE within a 6° radius shows little variation with changes in background modeling. The systematic effects associated with the diffuse modeling become more apparent outside of this radius, especially below 1.9 GeV,
where the differences between the radial profiles for fits with backgrounds A/E/F are larger than the error bars of each radial bin.

In the bottom row of Fig. 2.6 we plot the expected radial profiles for NFW haloes with density slopes of $\gamma = 0.8 - 1.3$ along with the observed radial profile of the best-fit GCE residuals for GALPROP model F. The normalizations of the NFW profiles are adjusted to best fit the entire radial range of GCE residual data points for energy bins below 1.9 GeV and from 1.9–10.0 GeV. For the energy bin above 10.0 GeV, the curves for varying NFW profiles are fit to the inner galaxy data points ($r \geq 5^\circ$). No single NFW template is able to fit all the radial bins; for example, between 1.9–10 GeV, all the NFW profiles tend to over-predict flux between $4^\circ - 5^\circ$, with shallower (steeper) profiles under(over)-predicting flux at lower radii and over(under)-predicting flux at higher radii. At energies above 10 GeV, the radial profile of the GCE residual is decidedly non-NFW-like due to a drop in flux within $4^\circ$. This is consistent with our earlier observation that the GCE spectrum in the galactic center drops off steeply around 10 GeV, while its spectrum has a high-energy tail in the inner galaxy ROIs.

2.3.4 Spatial uniformity of the galactic center excess spectrum

In Fig. 2.4 we identify a high-energy tail in the GCE spectrum which is present in combined inner galaxy ROIs (as seen in Refs. [61, 22]), but not in the galactic center ROI. Here, we compare the spectra between individual inner galaxy ROIs (as defined in Tab. 2.2) to explore the spatial uniformity of the GCE spectrum and its high-energy tail.

Fig. 2.7 shows the spectrum of the GCE-associated residual (observed counts—full model + best fit GCE model) for the subregions defined in Fig. 2.1. Also included are two additional ROIs N2 and S2 (defined in Tab. 2.2) which lie at farther latitudes ($11^\circ < |b| < 19^\circ$) than the N/S ROIs used in the inner galaxy analysis. We plot these additional ROIs in
Fig. 2.7 to check (1) the approximate spatial extent of the GCE and (2) whether the GCE spectrum in the farther-latitude regions, if detected, is consistent with the spectrum at closer galactocentric radii. Note that the N2 and S2 regions are not included in the analysis of the GCE in the inner galaxy. The error bars are shown for the statistical uncertainties in the binned fluxes, while the systematic uncertainties associated with the background modeling may be roughly estimated by the spread in the GCE spectrum fit with the different background models.

The spectrum in each ROI subplot is normalized in the same way as in Fig. 2.4, where the GCE flux is scaled by the J-factor of the entire NFW template divided by the J-factor of the smaller ROI. The average normalizations between ∼1–5 GeV in the GC and the separate inner galaxy ROIs (with the exception of the SW ROI) are in agreement within a factor of two. If taken at face value, this difference in GCE normalizations across the separate ROIs may be interpreted as a rough estimate of the uncertainty in the axis ratio of the GCE’s projected spatial distribution. Interestingly, the normalizations in the NW/NE/SE ROIs are higher on average than those in the N/S ROIs (although still overlapping within statistical and systematic uncertainties), which may hint at some degree of compression along the longitudinal axis in the GCE spatial profile.

We are unable to identify any specific source or extended feature in the SW ROI as the cause of this discrepancy. It is a strong possibility that the differences in the SW GCE spectrum are due to mismodeling of the diffuse background in that region, as the IC templates calculated by GALPROP are symmetric in $l$ and $b$ and are thus unable to capture the non-axisymmetric variations in the true background diffuse emission.

At farther distances from the galactic center, the GCE spectrum in the S2 ROI appears roughly similar in shape to the GCE spectrum in the inner galaxy, but has a slightly lower overall normalization. The GCE spectrum in the N2 ROI is consistent with zero flux for two out of the three diffuse backgrounds tested. These results may indicate that the GCE’s
intensity profile is well-described by an NFW profile out to angular distances of $\sim 10 - 12^\circ$, outside of which it falls off more steeply. However, the expected flux from an NFW profile at radii $>10^\circ$ is quite low and thus any determination of the GCE spectrum at these larger radii is subject to great uncertainty.

Within the framework of this analysis, we find that the high-energy tail of the GCE is a large-scale spatial feature that is present in all of the inner galaxy ROIs, with the exception of the SW ROI. We previously noted that the high-energy spectrum in the inner galaxy is in sharp contrast to the galactic center, where no GCE emission is observed at energies $\gapprox 10 - 20$ GeV. In Fig. 2.7 we show that (with the exception of the SW ROI) the GCE spectrum in all inner galaxy ROIs falls by a factor of roughly 2–3 between its peak at $\sim 2$ GeV and the highest energy bin (44.7–200.0 GeV). The high-energy GCE emission is prevalent throughout the inner galaxy, and is not a result of one region heavily biasing the combined inner galaxy spectrum.

It is possible that the high-energy tail of the GCE spectrum is simply misattributed flux from the Fermi bubbles. This hypothesis is supported by the absence of the tail in the low-latitude galactic center ROI and the similarity of the hard bubble spectrum to the GCE spectrum above 10 GeV (both are roughly $\sim E^{-2}$). The bubble morphology becomes uncertain at low latitudes, and may perhaps cover a larger fraction of the inner galaxy ROIs than assumed in our template [12, 165]. We consider it unlikely that the Fermi bubbles are responsible for the majority of the high-energy GCE flux if we assume that the rough bubble template used in this analysis is a reasonable approximation for the extent of the bubbles in the inner galaxy ROIs. As can be seen in Fig. 2.1, the fraction of each individual ROI in the inner galaxy covered by the bubble template ranges from $\sim 0.1$ (SE) up to 1.0 (N). Even if the bubble template used was not fully accurate in tracing the bounds of the bubbles, we would still expect to observe varying normalizations in the high-energy tail between different ROIs if these photons were in fact originating from the Fermi bubbles (i.e. the spectrum above 10
GeV would have the highest intensity in the N ROI and the lowest in the SE ROI). We do not observe any correlation between the intensity of the spectrum above 10 GeV in Fig. 2.7 and the fraction of each ROI overlapping with the bubble template.

2.4 Discussion

2.4.1 The GCE high-energy spectrum above 10 GeV

In Figs. 2.4 and 2.7 we explore the spatial dependence of the GCE spectrum and confirm previous findings by Refs. [61, 22] of GCE-associated emission upwards of \( \sim 20 \) GeV in the inner galaxy region. We find that this high-energy tail is not present in the GCE spectrum within the galactic center ROI; within approximately 5\(^\circ\) it has a spectral cutoff between \( \sim 5\)–10 GeV, while outside of this region the high-energy tail becomes a prominent feature for \( r \gtrsim 7^\circ \) – 8\(^\circ\). This spectral feature (or lack thereof) is robust to variations in the density profile of the NFW template and persists through variations in the diffuse \( \pi^0 \) decay + bremsstrahlung and IC background templates.

This difference between the high-energy GCE spectra in the galactic center versus the inner galaxy may be construed as either (1) a systematic effect associated with uncertainties in the background modeling, (2) an intrinsic spatial variation in the source contributing to the GCE spectrum, or (3) the signature of multiple sources with different spatial profiles. We do not find evidence of the former in this work, as the high-energy inner galaxy GCE spectrum above 10 GeV is recovered in all the GALPROP models and GCE spatial templates we used (Fig. 2.4). Ref. [22] do find that the GCE spectrum in the innermost 15\(^\circ\) × 15\(^\circ\) shows a dependence on background modeling: their ‘index-scaled’ diffuse emission models result in a softer GCE spectrum above 10 GeV than their ‘intensity-scaled’ models. When fit to an exponential cut-off functional form, the GCE spectrum cuts off before 10 GeV in the index-
scaled background fits; however, when the GCE spectrum is fit as a power law in individual energy bins, the power law-like high-energy tail is present for all background models (albeit with a softer index in for the index-scaled cases).

Here we will assume that (1) is not the case and discuss what implications might follow for interpretations of the galactic center excess as either dark matter annihilation or an astrophysical source. Under this assumption, the $\gtrsim 20$ GeV emission implies that a MSP population or dark matter source would need to produce both a prompt gamma-ray component (peaking in the $\sim$GeV range) as well as a hard leptonic component (which produces the high-energy tail above 10 GeV through IC scattering).

**Dark matter annihilation and the GCE spectrum**

We will first consider the case of dark matter annihilation producing the GCE. The simplest dark matter annihilation models fit the excess with prompt gamma-ray emission\(^4\), without the need for the primary annihilation products to produce secondary gamma-ray emission through the environment-dependent processes of IC or bremsstrahlung scattering. Prior to claims of the excess emission extending beyond $\sim 10$ GeV in energy, the GCE spectrum was most commonly fit with WIMPs annihilating into $\sim 10$ GeV $\tau$ leptons or $\sim 40$ GeV $b$ quarks [116, 126, 127, 6, 117, 164, 4, 61]. However, the gamma-ray spectra of these oft-mentioned $\tau^+\tau^-$ and $b\bar{b}$ annihilation modes cut off sharply by about 10–20 GeV, which is difficult to reconcile with the inner galaxy GCE spectrum we observe beyond those energies (although consistent with the GCE in the galactic center). Producing a high-energy tail through prompt dark matter annihilation alone is still possible, but requires models such as prompt annihilation of WIMPs into nonrelativistic Higgs ($m_\chi \simeq 126$ GeV) [60, 19].

\(^4\)Here, ‘prompt’ emission refers to the gamma rays produced through the hadronization and/or decays of the primary annihilation products as well as higher order corrections to the dark matter annihilation diagrams.
The prompt gamma-ray spectrum is only dependent on the particle physics involved in the dark matter annihilations and subsequent Standard Model hadronizations and/or decays. Assuming these processes are independent of environment, we would expect the GCE spectrum to be spatially uniform if it was due to prompt dark matter annihilation. In contrast, we observe that the GCE spectrum has a power law-like tail at high-energies in the inner galaxy but not the galactic center. If this discrepancy is a true feature of the GCE and not a systematic error, it disfavors the interpretation of the GCE as emission from prompt dark matter annihilation.

Alternatively, the GCE could be produced through secondary emission from Standard Model annihilation products, as discussed in Sec. 2.1. Secondary emission from IC or bremsstrahlung processes is dependent on the environment and may result in a spatially varying GCE spectrum. Previous authors have fit the GCE spectrum with IC scattering off of dark matter annihilation products [145, 60]. However, neither of these proposed IC scenarios are capable of producing a high-energy tail at larger galactocentric radii while suppressing it at lower radii because the IC target density decreases with distance from the galactic center.

In order for dark matter annihilations to reproduce the spatial variation we observe in the high-energy GCE spectrum, the spectrum of the secondary $e^+e^-$'s would need to have a cutoff energy higher than that of the primary annihilation products producing the prompt emission. The simplest dark matter annihilation models that fit the GCE with a combination of prompt and IC emission from a single annihilation channel (e.g. prompt annihilation into muons accompanied by IC scattering of the secondary $e^+e^-$ for $m_\chi \simeq 60$–70 GeV [60]) have secondary $e^+e^-$ spectra with cutoff energies below that of the primaries. Dark matter annihilation through multiple channels including direct annihilation into electrons may produce a harder electron spectrum, but the branching ratios and cross sections for that channel are tightly constrained by AMS-02 limits on electron-positron spectral lines features [39, 54].
Our arguments outlined above consider only single-component dark matter models. It may be possible to explain both the excess and its spatial variation through a model with two dark matter particles, with the higher-mass particle annihilating preferentially into leptons. We also note another possibility that the high-energy tail arises from an astrophysical source while the bulk of the excess below 10 GeV is due to dark matter annihilation. In this case the dark matter interpretation is not expected to be significantly different from that considered in the early papers [116, 126, 127, 6, 117, 164, 4] and so we do not discuss it further here.

**Leptonic cosmic-ray outbursts and the GCE spectrum**

Given that the Fermi bubbles are evidence of an extremely energetic outburst in the Milky Way’s past, it is possible that the GCE may originate from one or more cosmic-ray outbursts [66, 191, 110, 69]. Interpretations of the GCE as the product of burst events tend to focus on models dominated by leptonic, rather than hadronic, cosmic rays, as gamma-ray emission following a hadronic outburst would be strongly correlated with the gas distribution in the plane of the disk and thus would not be consistent with the approximately spherical morphology of the GCE [66].

The cosmic-ray electron spectrum changes with distance from the galactic center due to diffusion and energy losses. The combination of a spatially varying electron spectrum and interstellar radiation field should lead to a similarly non-uniform GCE spectrum. In this regard, our finding of a radially varying GCE spectrum would seem to support the leptonic outburst scenario. However, it is still difficult to explain why the GCE high-energy tail would be present in the inner galaxy but not the galactic center, as cosmic-ray propagation outwards from a central source would result in the opposite effect. If the GCE is due to multiple outbursts, the harder spectrum at farther radii might be produced through a burst with a considerably harder injection spectrum than the more recent bursts which contribute to the excess at lower radii.
An additional complication in modeling the GCE with cosmic-ray outbursts is replicating the steeply rising GCE spatial profiles within small radii. The angular profiles of cosmic-ray GCE models are flat at low radii, and multiple recent fine-tuned bursts (within a few hundreds of years) are needed to produce a GCE profile that continues to rise at low radii [69].

**Millissecond pulsars and the GCE spectrum**

The similarity of the $\sim2$ GeV bump in the GCE spectra reported in Refs. [127, 6, 117, 164, 4, 5, 87] to Fermi observations of resolved MSPs and globular clusters (which host populations of millisecond pulsars) motivates the interpretation of the signal as emission from an unresolved population of MSPs concentrated at the galactic center [3, 6, 169, 192, 232, 231, 156, 183]. Ref. [183] also discuss the possible contribution to the GCE from young pulsars. We will focus this discussion on MSPs as the potential unresolved young pulsar population is concentrated in the plane of the disk and thus would not produce the roughly spherical profile of the GCE.

In the left panel of Fig. 2.8 we compare the GCE spectrum in both the galactic center and inner galaxy ROIs to the prompt spectra of MSPs as measured in Ref. [70]. Although the spectral shape of the excess in the galactic center is consistent with the spectra of the stacked globular clusters and stacked individual MSPs, the high-energy tail of the GCE spectrum in the inner galaxy is a distinct departure from the sharp cutoffs at $\lesssim10$ GeV in the stacked spectra. The inner galaxy GCE spectrum is barely consistent with prompt MSP emission—while the typical globular cluster or MSP spectrum cuts off before 10 GeV, there are a handful of globular clusters (M5, M62, NGC 6624, NGC 6752) whose parameterized spectra predict gamma-ray emission at energies above 15 GeV. Thus, one could claim that the high-energy tail of the GCE in the inner galaxy is not entirely inconsistent with prompt MSP emission, if using the outliers in the globular cluster sample as a comparison. However, these clusters also have very large uncertainties in their fitted spectra: the 68% lower confidence interval
on the spectral cutoff energy lies below 15 GeV for all of the aforementioned outliers.

Furthermore, it may be possible for a MSP population to produce the variation in the high-
energy tail through a combination of prompt emission from the MSPs themselves as well as 
secondary IC emission from the $e^+e^-$ injected by the MSPs into the interstellar medium. 
Ref. [192] point out that for certain cosmic-ray propagation parameters in their models, 
the secondary IC emission from MSPs is subdominant to the prompt signal within $\sim 2^\circ$ in 
latitude but becomes comparable to—or even greater than—the prompt emission at latitudes 
outside this range. If the electron injection spectrum has a high cutoff energy above $\sim 100$ 
GeV, the secondary IC emission would extend beyond 10 GeV and could give rise to the 
prominent high-energy feature in the inner galaxy. A comparison of the GCE spectra in the 
galactic center and inner galaxy ROIs may in fact be suggestive of this: the inner galaxy GCE 
spectrum resembles a composite of the galactic center spectrum with a harder IC $\sim E^{-2}$ 
power law spectrum extending past 100 GeV.

We explore this possibility in the right panel of Fig. 2.8 by plotting the inner galaxy GCE 
spectrum against the combined prompt and IC components that might arise from an unre-
solved MSP population. The ‘prompt’ component of this model is taken to be the best-fit 
exponential cutoff parameterization of the GCE in the galactic center (for GALPROP model 
A and $\gamma = 1.1$) with a freely floating normalization. The ‘IC’ component is fit as an expo-
nential cutoff spectrum with the normalization, index, and cutoff energy free to vary. The 
combination of the best-fit ‘prompt+IC’ spectrum is shown as the heavy solid line in the 
right-hand side of Fig. 2.8.

The best-fit IC parameterization has a power-law index $\Gamma = 1.55$ and a cutoff energy of 
68 GeV, indicating that a MSP origin would require a very hard spectrum spectrum of 
outgoing electrons with energies up to $\mathcal{O}(10)$ times greater than the maximum energy of the 
prompt gamma-ray emission. Such an injection spectrum may be achievable through one of 
several mechanisms proposed to accelerate MSP electrons to energies $>100$ GeV. Ref. [120]
find that offsets between the polar cap and magnetic dipole axis in MSPs can produce $e^+e^-$ cascade pairs with energies up to $\mathcal{O}(100)$ times greater than in young pulsars; however, the outgoing pair spectra in these models have softer power-law indices than the gamma-ray IC fit in Fig. 2.8 might require. A harder $\sim$TeV electron spectrum may be produced through reacceleration of $e^+e^-$ at intrabinary shock fronts within MSP systems [221] or in pulsar wind shocks [35, 36, 231].

The large uncertainties in MSP electron injection spectra leave room for a hard $e^+e^-$ population to produce both the GCE spectrum above 10 GeV and its spatial variation. An unresolved population of MSPs is therefore a compelling explanation for the GCE with (1) the prompt gamma-ray emission dominating the spectrum between $\sim 1 - 5$ GeV, accounting for the similarity between the galactic center and inner galaxy GCE spectra in this energy range and (2) the spatially dependent IC emission dominating the inner galaxy spectrum above $\sim 20$ GeV. Future work (in progress) will model the combined prompt+IC spectrum from a $\sim r^{-(1.6 - 2.2)}$ MSP distribution as a function of galactocentric distance and investigate how this population contributes to the WMAP haze.$^5$

Radio and microwave observations provide some constraints on leptonic emission from MSPs, but are highly dependent upon the parameters assumed for the cosmic-ray injection and propagation [71, 54, 99]. Radio observations utilizing the upcoming Square Kilometer Array (SKA) will provide greater power to detect more MSPs in addition to any synchotron emission associated with their secondary IC emission [211, 63]. If the GCE high-energy tail is indeed IC radiation from an electron population with $\gtrsim$TeV energies, the possible extension of this component into photon energies above the Fermi-LAT sensitivity range may be detectable by next-generation of TeV-scale gamma-ray observatories such as the Cherenkov Telescope Array (CTA) [231].

$^5$Ref. [146] study young pulsars as a possible source for the WMAP haze; however, it is hard to reproduce the latitudinal extent of the haze with the disk-like young pulsar distribution.
Of course, the GCE might originate from a combination of multiple astrophysical sources. The inner $\lesssim 5^\circ$ of the GCE may be due to prompt emission from an unresolved population of MSPs (which produce the spectrum with a $\sim 10$ GeV cutoff that we observe in the galactic center) while the dominant contribution to the harder gamma-ray GCE spectrum at larger radii comes from one or more leptonic cosmic-ray bursts. In this combined scenario, the GCE spatial profile in the galactic center arises because of the $\sim r^{-2.2}$ distribution of the putative MSP population, and there is no need to invoke a series of recent outbursts to explain the steep rise in the centralmost regions.

2.4.2 Is a point source population favored over a smooth annihilation profile as the origin of the GCE?

Recent works have attempted to determine whether the GCE originates from a smoothly distributed NFW annihilation source or a population of $\mathcal{O}(1000)$ faint point sources with fluxes below the Fermi-LAT detection sensitivity. Ref. [32] employ a wavelet decomposition analysis of Fermi-LAT data and find that the photon clustering structure is compatible with the estimated radial distribution and spectrum of a faint MSP population. Ref. [158] use non-Poissonian photon-count statistics to differentiate between the signal produced by a smooth NFW source versus a unresolved point source distribution. Their analysis favors a point source origin for the GCE, where all of the excess might be explained by a source-count distribution with a sharp decline just below the Fermi detector sensitivity ($\sim 1 - 2 \times 10^{-10}$ ph cm$^{-2}$ s$^{-1}$).

Here, we explore whether the photon-count distribution of the inner galaxy GCE residual allows us to distinguish between the unresolved point sources and dark matter annihilation scenario for the GCE source. We simulate a population of unresolved point sources in our inner galaxy ROIs with a radial distribution $\sim r^{-2.2}$. Point source fluxes are drawn from
a source-count function \( \frac{dN}{d\Phi} \) \([\text{sources/(ph cm}^{-2} \text{s}^{-1})]\) modeled by a broken power law of the form \( \Phi^{-\alpha_1(\alpha_2)} \) below (above) the break flux \( \Phi_b \), where \( \Phi \) is the flux \([\text{ph cm}^{-2} \text{s}^{-1}]\) per point source between 1.9–11.9 GeV. We use the parameters \( \{\Phi_b = 2.16 \times 10^{-10} \text{ ph cm}^{-2} \text{s}^{-1}, \alpha_1 = -0.57, \alpha_2 = 29.5\} \) from the source-count function found by Ref. [158] to be the best fit point source population. The spectrum and total flux of the simulated point source distribution is chosen to match the inner galaxy GCE spectrum for \( \gamma=1.1 \) and the model F diffuse background. We find that \( \sim1900 \) point sources are required within an 18° radius to match our inner galaxy GCE spectrum and flux using this source-count function. About 300 of these sources lie within a 10° radius with |\( b \)| > 2°, which corresponds to the region within which Ref. [158] require \( 203^{+109}_{-68} \) point sources to explain the excess.

We use the \textit{gtobssim} tool [1] to simulate photon events originating from (1) the point source population described above and (2) annihilation signal from an NFW source with \( \gamma = 1.1 \) and the same spectrum as the point sources. The observations are simulated using the same timeframe and cuts as described in Sec. 2.2 and include the effect of the instrument’s energy-dependent PSF. In Fig. 2.9 we plot the distribution of photon-counts per pixel for pixels with non-zero counts. We show both of the simulated observations, as well as the photon-counts per pixel for the full residuals of the fit using \( \gamma = 1.1 \) and the model F background.

As expected, the simulated observations (solid lines) in Fig. 2.9 show that the smooth annihilation source has more pixels with low photon-counts (1 ph/pixel) while the point source distribution has more pixels with higher photon-counts (\( \geq 2 \) ph/pixel). The statistical method described in Ref. [157] takes advantage of this difference to determine whether a point source population or smooth NFW halo is the true source of the GCE.

By comparison, we find that the amplitudes of the positive and negative pixel-count distributions for the full residuals (gray lines) are an order of magnitude larger than the differences between the point source population and NFW halo annihilation. Over- and under-subtractions in the gamma-ray residuals may be due to multiple issues in the modeling, and are not nec-
necessarily due to real gamma-ray emitting features. We do not presume to understand the underlying causes of the large positive and negative fluctuations in our residuals.\textsuperscript{6} Lacking an understanding of the effects that give rise to the photon-count distribution of the residuals, it is possible that small-scale spatial structure in the mismodeling may be erroneously interpreted as sub-threshold point sources. Our analysis of the photon-count distribution is not sufficiently sophisticated to estimate the extent to which this mismodeling may affect current sub-threshold point source analyses, but it provides a visual demonstration of this systematic uncertainty.

\section*{2.5 Conclusions}

We find that the inclusion of an extended, spherically symmetric gamma-ray source with an NFW-like radial profile of $\sim r^{-(1.6-2.2)}$ strongly increases the fit likelihood obtained through the template fitting procedure within $\sim 10^\circ - 15^\circ$ of the galactic center. The galactic center excess spectrum obtained through the likelihood template fitting procedure is reasonably robust to variations in the NFW and background diffuse model templates used in this work, even in extreme cases where the background modeling is likely an unphysical description of the true extended gamma-ray sources. These findings are in agreement with many previous studies of the excess [116, 222, 126, 127, 6, 117, 164, 130, 4, 5, 236, 87, 61, 22, 81]. If we compare the galactic center and inner galaxy spectra for $\gamma = 1.1$, the peak normalizations are consistent in both ROIs.

When varying the GALPROP-generated IC and $\pi^0$+bremsstrahlung background models, we find that the galactic center ROI within $\lesssim 4^\circ - 5^\circ$ is consistently best fit with an NFW profile slope around $\gamma = 1.1$. Outside of this radius, however, the best fit NFW profile in the

\textsuperscript{6}For a thorough description of the intricacies of modeling the diffuse emission towards the galactic center, see Ref. [22]. We note that the residual counts per pixel area shown in Ref. [22] are of similar order to this work, despite their use of specialized background modeling tuned to flatten the residuals.
inner galaxy is poorly constrained \((\gamma \lesssim 0.8 - 1.1)\), with a heavy dependence on the choice of diffuse background model (Tab. 2.3). Our results suggest that previous works \([87, 61]\), in which the profile slope was found to be well-constrained to \(\gamma \sim 1.1 - 1.2\) when using a \(\sim 20^\circ\) ROI about the galactic center, may have been driven strongly by the inclusion of the innermost few degrees of the excess during the template fitting procedure.

The most noticeable difference between the GCE in the galactic center versus the inner galaxy regions is the hardening of its spectrum at galactocentric radii above \(\sim 5^\circ - 6^\circ\) (Figs. 2.4, 2.6). In the inner galaxy, we observe a power law-like tail in the GCE spectrum extending upwards of 100 GeV, while the spectrum in the galactic center has a steep falloff at \(\sim 10\) GeV.

The inner galaxy high-energy tail above 10 GeV is found in all but one of the inner galaxy ROIs and is robust to variations in the diffuse background models and the GCE spatial templates used in this work. This presence and intensity of this high-energy component is roughly consistent across most of the inner galaxy ROI and shows no obvious azimuthal asymmetry (Fig. 2.7). Upon examination of the radial distribution of GCE photons above \(\sim 10\) GeV, we see that this high-energy spectral feature is roughly consistent with an NFW annihilation profile outside of a \(\sim 5^\circ - 6^\circ\) radius, but does not exhibit the steep rise in brightness towards lower radii that we observe for the GCE photons below 10 GeV.

If the full energy range of the GCE emission in the centralmost few degrees as well as the outlying regions is assumed to arise from a single source, then a single component dark matter annihilation model cannot account for the spatial variation of the high-energy GCE emission above 10 GeV. Of course, it is possible that the high-energy tail and bulk of the excess below 10 GeV are due to two different sources, in which case there is no difficulty in explaining the excess below 10 GeV as arising from dark matter annihilation.

We attempt to use the photon-count distribution of the GCE residual to distinguish between
the scenarios of dark matter annihilation in a smooth NFW halo and an unresolved population of MSPs. However, this effort is inconclusive as the amplitude of the full residuals is greater than the GCE amplitude by a factor of $\sim$few in almost all photon count bins (Fig. 2.9). We thus caution that the residual photon count distribution resulting from mis-modeling of the data may be a confounding factor when using photon count statistics to search for point source populations.

Although we are unable to confirm the existence of an unresolved MSP point source population, it remains a compelling explanation because of its close match with the GCE spectrum below 10 GeV and the potential for the MSP population to produce a secondary leptonic component at energies significantly higher than that of the prompt gamma-ray emission. If the MSP electron injection spectrum is sufficiently hard and extends upwards of $\sim 100$ GeV, the GCE emission above 10 GeV may be attributable to IC scattering of these high-energy electrons. The spatial variation of the high-energy tail of the GCE spectrum described in this work would then follow as a natural result of electron propagation. Looking towards the future, our understanding of the true source(s) of the GCE will be greatly advanced by combining multiwavelength observations with the ongoing efforts involving realistic modeling of cosmic-ray propagation and novel statistical analyses.

Acknowledgements

We thank Kevork Abazajian, Sheldon Campbell, Oscar Macias and Simona Murgia for useful discussions and comments, as well as Farhad Yusef-Zadeh for sharing the 20 cm radio map used as a template in this work. M.K. is supported by NSF Grant No. PHY-1316792. A.K. is supported by NSF GRFP Grant No. DGE-1321846.

\footnote{As this manuscript was being prepared we became aware of Ref. [65], which studies the dependence of the GCE spectra upon a set of physically-motivated background diffuse models. There are some similarities in results between Ref. [65], particularly with regards to the potential degeneracies between the GCE and GALPROP IC components as well as the background-dependence of the best-fit GCE morphology.}
Table 2.3: Relative differences in the test statistic (TS) value of the NFW template, given for all combinations of NFW template/diffuse background in both the galactic center and combined inner galaxy ROIs. The differences in test statistic values are given relative to the combination of NFW density slope and diffuse background with the highest TS values: $\gamma=1.1$/model A in the galactic center ROI and $\gamma=0.8$/model F in the inner galaxy ROIs.

<table>
<thead>
<tr>
<th>$\gamma_{\text{NFW}}$</th>
<th>model A</th>
<th>model E</th>
<th>model F</th>
<th>model A</th>
<th>model E</th>
<th>model F</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>0.0</td>
</tr>
<tr>
<td>0.9</td>
<td>-44.0</td>
<td>-81.1</td>
<td>-112.4</td>
<td>-349.6</td>
<td>-976.1</td>
<td>-21.2</td>
</tr>
<tr>
<td>1.0</td>
<td>-14.1</td>
<td>-49.4</td>
<td>-105.2</td>
<td>-342.1</td>
<td>-952.7</td>
<td>-32.6</td>
</tr>
<tr>
<td>1.1</td>
<td>0.0</td>
<td>-33.6</td>
<td>-101.7</td>
<td>-314.9</td>
<td>-932.1</td>
<td>-31.0</td>
</tr>
<tr>
<td>1.2</td>
<td>-4.5</td>
<td>-36.2</td>
<td>-103.1</td>
<td>-353.6</td>
<td>-946.8</td>
<td>-50.9</td>
</tr>
<tr>
<td>1.3</td>
<td>-34.2</td>
<td>-63.5</td>
<td>-119.7</td>
<td>-356.7</td>
<td>-942.6</td>
<td>-81.8</td>
</tr>
</tbody>
</table>

Figure 2.6: Radial distribution of the GCE residual flux in energy bins 0.7–1.9 GeV, 1.9–10.0 GeV, and 10.0 GeV–200 GeV. The top row plots the GCE residuals for the best fitting profile slopes obtained when fitting with diffuse model backgrounds A/E/F. For all models and ROIs this corresponds to $\gamma = 1.1$, with the exception of the inner galaxy model F fit which is plotted for $\gamma = 0.9$. The bottom row plots the GCE residual for the fits with $\gamma = 1.1$ (0.9) in the galactic center (inner galaxy) and model F diffuse background (solid black steps) against the expected radial distribution of emission from NFW-like sources of varying profile slopes (colored lines) comparison.
Figure 2.7: Best-fit GCE spectra in the galactic center ROI and the 6 inner galaxy ROIs, shown for fits with diffuse backgrounds model A (dark blue), E (purple), and F (light blue). Additional regions N2/S2 are also plotted. Panels are arranged to reflect location on the sky. We show our results for the case of a GCE template with profile slope $\gamma = 1.1$, though our findings are consistent for all template slopes. Normalizations are scaled such that each subplot shows the expected flux for the entire GCE template ($35^\circ \times 35^\circ$). Energy coordinates of the data points are slightly offset for visibility purposes.
Figure 2.8: Left: Comparison of the GCE spectrum in the galactic center (black circles) and inner galaxy (orange squares) to the spread of globular cluster MSP spectra (dashed gray lines) from Ref. [70]. Also shown are the inferred MSP spectra from Ref. [70] derived from stacked observations of globular clusters (solid green line) and individual MSPs (dashed red line). The spectra are normalized to match each other at 2 GeV for ease of comparison. Right: Comparison of the inner galaxy GCE spectrum (orange squares) to our estimations of the prompt (dotted line) and IC (dashed line) components of the potential unresolved MSP population.
Figure 2.9: Simulated photon-count distribution in the combined inner galaxy ROIs between 1.9–10.0 GeV of an NFW source with $\gamma = 1.1$ (dashed blue line) vs. a population of unresolved point sources (solid pink line) with the best-fit source count function from Ref. [158]. Pixel size is 0.2° per pixel. The spectrum and total flux of each simulated source is chosen to match our total observed flux in this energy range. For comparison, we also show the photon-count distribution of the full residual (solid gray line) for the case of GALPROP model F and $\gamma = 1.1$. Note that the zero-count bin is not shown.
Chapter 3

A New Gamma-ray Source Consistent with Upscattered Starlight

Based on Abazajian et al., JCAP 1507 (2015) no.07, 013 [5].

We present a new extended gamma ray excess detected with the Fermi Satellite Large Area Telescope toward the Galactic Center that traces the morphology of infrared starlight emission. Combined with its measured spectrum, this new extended source is approximately consistent with inverse Compton emission from a high-energy electron-positron population with energies up to about 10 GeV. Previously detected emissions tracing the 20 cm radio, interpreted as bremsstrahlung radiation, and the Galactic Center Extended emission tracing a spherical distribution and peaking at 2 GeV, are also detected. We show that the inverse Compton and bremsstrahlung emissions are likely due to the same source of electrons and positrons. All three extended emissions may be explained within the framework of a model where the dark matter annihilates to leptons or a model with unresolved millisecond pulsars in the Galactic Center.
3.1 Background

The Fermi Gamma Ray Space Telescope Large Area Telescope (Fermi LAT) has observed with unprecedented detail the “heart of darkness” of our Galaxy: its gravitational center. The past few years have revealed that there are a large number of new point sources [178] as well as new diffuse emission [233]. In addition, a large extended source, the Galactic Center Extended (GCE), has been detected by a number of groups [116, 126, 127, 6, 87] and is robust to uncertainties in the diffuse emission foregrounds in the region [117, 4, 236, 61]. The high-energy radiative processes that produce gamma rays are often commensurate with production or acceleration of related relativistic charged particle cosmic rays. Astrophysical processes include diffusive shock acceleration, magnetic reconnection, “one-shot” acceleration across high-voltage electric fields, and many other possibilities. Another source that can produce both cosmic rays and gamma rays are the products from candidate dark matter particle annihilation or decay. The significance of the Galactic center as a bright source for dark matter annihilation photons and cosmic rays has been known for some time [40].

High-energy charged particles, deposited either from astrophysical sources or dark matter annihilation, experience various propagation and energy-loss processes in the Galactic Center region. There has been recent work discussing how bremsstrahlung and inverse Compton (IC) effects could alter the prompt spectra coming from dark matter annihilation [74, 155]. What we show here, for the first time, is that these separate components—prompt, bremsstrahlung and IC—can be separated with morphological as well as spectral information. In particular, we report the discovery of a new extended component of the gamma-ray emission toward the Milky Way Galactic Center that is spectrally and morphologically consistent with a population of electron-positron ($e^\pm$) cosmic rays producing gamma rays by upscattering starlight through the IC process. Secondly, we confirm the presence of an emission consistent with bremsstrahlung radiation, and find that this emission can be produced by the same population of $e^\pm$ impinging on the high-density gas in the Galactic Center. Lastly, we confirm the
presence of a GCE source (peaking around 2 GeV) that has a centrally-peaked morphology consistent with dark matter annihilation. We show that the IC, bremsstrahlung and GCE components could all have originated from the products of dark matter annihilation. This explanation is not unique in that an unresolved population of millisecond pulsars or two independent astrophysical sources could produce these signals.

3.2 Methods

We use Fermi Tools version v9r33 to study Fermi LAT data from August 2008 to June 2014 (approximately 70 months of data). We use Pass 7 rather than Pass 7 Reprocessed instrument response functions since the diffuse map associated with the latter have strong caveats for use with new extended sources. Our procedure is similar to those described in Ref. [4]. We simultaneously fit the amplitude and spectrum of point sources from the 2FGL catalog [178], plus four other point sources in the ROI, as described below, in our region of interest (ROI) $7^\circ \times 7^\circ$ around the GC centered at $b = 0, \ell = 0$. We use $0.2 - 100$ GeV photons in 30 logarithmically-spaced energy bins. To enhance spatial resolution, we use ULTRACLEAN-class photons binned in an Aitoff projection into pixels of $0.1^\circ \times 0.1^\circ$.

We include the 20 cm radio template as a tracer of gas to account for the bremsstrahlung emission as has been done previously [233, 164, 4]. To test the possibility of IC emission from starlight due to this same population of $e^\pm$, we use the 3.4 $\mu$m template for stellar light from the WISE mission [230]. Among the templates tested, this had the least obscuration of stellar light in the ROI; the results from other templates studied are discussed later. Our goal in using the 3.4 $\mu$m template is to test whether the IC component’s morphology might be approximated by it; we do not presume that this template is an exact morphological description of the putative IC emission. As an example, if the diffusion length of $e^\pm$ is significantly less than the ROI dimensions, then the IC emission will track the morphology.
of the $e^\pm$ source more.

We use a $14^\circ \times 14^\circ$ template because of the broad PSF of Fermi-LAT producing contributions outside of the ROI, particularly at low energies consistent with the IC photons. As in Ref. [4], we also include the New Diffuse (ND) map whose intensity increases with angle away from the GC, which was interpreted as accounting for additional gas not captured in the 20 cm map. We have optimized the morphology of the GCE excess and ND templates to their best-fit profiles, as in Ref. [4]. To optimize the GCE excess, we used templates of $\rho(r)^2$ projected along the line-of-sight with $\rho(r) \propto r^{-\gamma}(r+r_s)^{-(3-\gamma)}$ and found that $\gamma = 1$ provided the best-fit. The best-fit new diffuse template increases with projected distance from the Galactic Center, $\theta$, as $\theta^{0.3}$. It is worth noting that our GCE template is somewhat less steep than found previously, $1.1 - 1.4$ for the inner slope of the density profile [4, 117, 87].

### 3.3 Results

In our analysis, we include two previously discovered point sources, 1FGL J1744.0-2931c and “bkgA” [233], and furthermore discover two new point sources, PS1 and PS2, at $\ell, b$ of $356.616^\circ, 1.009^\circ$ and $356.829^\circ, -0.174^\circ$ with large test statistic (TS) values of 168 and 140, respectively. PS1 has a spectrum consistent with a Log-Parabola, is near numerous X-ray and radio sources including, e.g., 1RXS J173331.6-311522, and could be in the plane of the Galaxy or extragalactic. PS2 has a spectrum consistent with a power law, and is near the supernova remnant G356.8-00 and compact radio source G356.9+0.1 [202]. They are at the edge of our ROI, and our conclusions regarding the IC, bremsstrahlung and GCE sources are not qualitatively affected by their inclusion. We choose to keep them in our models.

All the 4 extended sources (GCE, ND, IC, Bremsstrahlung) were given generic log-parabola

\[ TS \equiv 2\Delta \ln L, \] where $\Delta L$ is the difference of the best fit likelihood with and without the source. For point sources, a value of $TS = 25$ is detected at a significance of just over $4\sigma$ [178].
spectral forms with four free parameters each. We detect the WISE 3.4 μm template at very high significance of TS = 197.0. The previously studied sources were also detected at high significance. The GCE was detected with TS = 207.5, bremsstrahlung was detected with TS = 97.2. These sources and best fit models are shown in Fig. 3.1, and the resulting residual spectra and best-fit log-parabola models are shown in Fig. 3.2.

Figure 3.1: Shown in the left column are the residual photons (top) and best fit model (bottom) associated with the projected interstellar radiation field template in its peak intensity bin, $0.303 \text{ GeV} < E_\gamma < 0.372 \text{ GeV}$, where the residual map has been smoothed with a Gaussian of $\sigma = 0.9^\circ$ (to roughly account for the point-spread function). The middle column shows the residual photons (top) and best fit model (bottom) associated with the projected dark matter density squared template in its peak intensity bin, $1.59 \text{ GeV} < E_\gamma < 1.95 \text{ GeV}$, where the residual map has been smoothed with a Gaussian of $\sigma = 0.4^\circ$. The right column shows the residual photons (top) and best fit model (bottom) associated with the 20 cm radio map in the same energy bin and with the same smoothing as the middle row. Residual and model maps have the same color scale for each row. This analysis used ULTRACLEAN-class photons.

In addition to the IC and Bremsstrahlung signatures of a population of high energy $e^\pm$, the 20
cm radio emission is also consistent with the synchrotron emission from the same population of electrons with correlated implications for the ionization and temperature of the molecular gas [233]. The fact that the bremsstrahlung emission traces the 20 cm (synchrotron) map indicates that the magnetic field is frozen into the gas. When we replaced the 20 cm map with a CO map, which contains dense molecular structures along the plane, the bremsstrahlung excess was not detected.

We again emphasize that we do not expect the WISE template to be an exact morphological description for the proposed IC emission from GCE-associated electrons. However, the WISE template’s high TS value does indicate that it is indeed a reasonable approximation for the IC component. We tested two other templates for the IC component. With a 100 µm dust template map [206], we were able to detect essentially the same IC spectrum with almost the same TS value. This indicates that the IC emission traces a disky template (thicker than the bremsstrahlung emission) but that there are considerable uncertainties in determining the correct morphology due to the poor angular resolution at energies below 500 MeV. With
a 2MASS J-band (1.2 \(\mu m\)) template [210], the significance of the IC detection was much lower (TS = 98.4); this is likely due to the large variable dust attenuation evident in the J-band map. For both, we observe the notable feature of the IC spectral cutoff at \(\sim 1\) GeV. This is a distinctive spectral feature of the IC excess that distinguishes it from the GALPROP-calculated IC emission contained within the diffuse background model.

To test the robustness of our results, we repeat our analysis using different diffuse backgrounds generated using the GALPROP code [2]. We tested two models in the extreme case where parameters were chosen with the intent to increase the IC emission predicted in the diffuse background, thus increasing the possibility that some or all of the excess might be absorbed into the background. For the first of these extreme cases, we chose a model with a very low diffusion parameter, which results in increased IC emission at lower energies. For the second extreme case we tested a model with a factor of 1.5 increase in optical and IR ISRF normalizations. We also tested two models with more standard parameters that were found by [61] to be good fits to the data in the inner galaxy. In all tests, the IC and GCE components were recovered with very similar spectra and uncertainties as when the Pass 7 model background was employed. The bremsstrahlung component was similar in most test cases; the only notable difference was that when employing a background model with extremely low diffusion coefficient, the bremsstrahlung spectrum did not show a cut-off as in Fig. 3.2, but it still had the same flux at GeV energies. These findings support our claim that the excesses are robust, and not an artifact of using the Pass 7 diffuse model. Interested readers may find the details of these diffuse background model checks in the following appendix.

We also test the dependence of our main results on the extended source templates included when modelling the data. We confirm that the spectra of the IC and GCE components remain more or less unchanged when the data is fit without the MG and ND components. We also substitute the HESS collaboration’s \(\sim\)TeV residual map of the Galactic Ridge [21]
in place of the 20 cm map used as the bremsstrahlung template, and find that the GCE and IC spectra again remain very consistent with the results shown in Fig. 3.2, while the HESS map spectrum has a slightly lower energy cutoff (around 1 GeV) than the 20 cm map spectrum (around 2–4 GeV) but with similar peak normalization. Details of these checks may also be found in the appendix.

3.4 Interpretation and Discussion

We discuss here how the detected IC emission is consistent with arising from the same population of $e^\pm$ as that producing the bremsstrahlung emission. In addition, we show that the GCE, IC, and bremsstrahlung emission can all arise from dark matter annihilation to leptons.

Apart from the dark matter interpretation [127], the GCE has been proposed to be emission from millisecond pulsars (MSPs) [3, 6], episodic hadronic [66], or episodic leptonic cosmic ray injection [191]. Pulsars have the right conditions to produce energetic $e^\pm$ cosmic rays [115] and hence, in principle, MSPs could explain all three excesses: the GCE excess due to the gamma-ray emission from their outer magnetosphere, and the IC and bremsstrahlung resulting from the $e^\pm$ that are produced along with gamma-rays in cascades [121]. Hadronic emission is less promising because it has trouble with the observed symmetry of the GCE. The IC emission must arise from a leptonic channel, perhaps secondary $e^\pm$ produced due to hadronic interactions or a separate channel such as leptonic cosmic ray emission from star formation activity.

The fact that bremsstrahlung and the GCE spectrum could originate from a broken spectrum of $e^\pm$ resulting from dark matter annihilation has been discussed previously [160, 233]. Below we argue that the bremsstrahlung and IC spectra may naturally be related to the same $e^\pm$
Figure 3.3: Shown here is an example 8 GeV dark matter annihilation model with equal branching to all charged leptons, $e^\pm, \mu^\pm, \tau^\pm$, with the residual spectra of the prompt GCE (blue square), IC (golden triangle), and bremsstrahlung (pink circle) sources. The blue (dashed) GCE spectrum is determined by the particle mass and annihilation rate fit to the observations. The solid *predicted* resultant spectra for this annihilation channel’s IC (golden) and bremsstrahlung (pink) cases are in solid lines. **ULTRACLEAN** class photons are used for this analysis.

population, which in turn could be connected to the origin of the GCE excess.

Let us consider a population of $e^\pm$ with energy $E_c$. The resulting IC photons have typical energies of $(E_c/m_e)^2hc/(1 \text{ } \mu m)$, where we have used the fact that the spectral energy distribution of the interstellar radiation field peaks around a micron. Assuming $E_c = 10$ GeV results in an IC spectrum that cuts off rapidly by 1-2 GeV. The bremsstrahlung spectrum for the same population is broader in energy and extends up to $E_c$. Both these predictions are qualitatively consistent with the spectra shown in Fig. 3.2. To test the consistency of the spectra with this simple picture further, we build a simplified model of diffusion and energy loss in the Galactic Center.

The $e^\pm$ in the products created by dark matter annihilation lose energy through three distinct process [80]: (1) IC, which leads to upscattering of the interstellar radiation field (ISRF) photons, (2) bremsstrahlung (Br) radiation off the gas, and (3) synchrotron radiation in the
Galactic magnetic field. We focus on the first two components in this letter. The differential flux of photons for these two components may be written as,

$$E \frac{dN_{IC, Br}}{dE} = \int_{\text{FOV}} \frac{d\Omega}{4\pi} \int_{\text{LOS}} d\ell \int_{E_{\text{min}}}^{m} \frac{dE_e}{dE} \frac{dn_e}{dE} \frac{dP_{IC, Br}}{dE}$$

(3.1)

where FOV and LOS indicate integration over the field-of-view and line-of-sight respectively, $dP_{IC}/dE$ and $dP_{Br}/dE$ are the differential power emitted per electron due to IC and bremsstrahlung processes. For bremsstrahlung, we include energy losses from atomic H and He. To get the source energy distribution of electrons, positrons and gamma rays, we use the software PPPC4DMID [73]. The number density of electrons and positrons per unit energy, $dn_e/dE_e$, is computed after including diffusion and energy losses according to the prescriptions in Refs. [94, 74].

To propagate the $e^\pm$, we assume a spatially constant diffusion coefficient $K(E) = K_0 E^\delta$, with $K_0$ and $\delta$ set to the ”MED” model [94] (often used for diffusion in the local neighborhood). The diffusion process is largely unconstrained in the Galactic Center and variations away from the assumed parameters have significant effects on the magnitude and spatial profile of the IC and Br signals. Also, the analytic description for diffusion does not allow for spatially-varying energy loss terms and we have assumed average values for the ISRF energy density and the gas density to create the model spectra for comparison. These choices, however, serve to illustrate our two main points that (a) the IC and bremsstrahlung spectrum could be due to the same lepton population, and (b) a single mechanism could explain the morphologies, strengths and spectra of the three distinct Galactic Center extended excesses.

For the average gas density and magnetic field strength, we assume $3 \text{ cm}^{-3}$ and $3 \mu \text{G}$, which are reasonable given the large uncertainties at the Galactic Center [155]. We use the radiation density of ISRF photons included with GALPROP v50 [2, 194]. Since our FOV is $\pm 0.5$ kpc of the Galactic Center, we use the value of the ISRF energy spectrum tabulated for
$R = 0, Z = 0.25$ kpc in GALPROP v50 as the average over the region contributing to the IC flux.

Our final estimates for the IC and bremsstrahlung excesses are,

$$E \frac{dN_{IC,Br}}{dE} = \frac{\text{FOV}}{4\pi} \ell_{IC,Br} \int_{E_{\text{min}}}^{m} dE_e \left\langle \frac{d\nu_e}{dE_e} \right\rangle \left\langle \frac{dP_{IC,Br}}{dE} \right\rangle$$

where $\left\langle \frac{dP_{IC,Br}}{dE} \right\rangle$ are computed using the average ISRF and gas densities and $\left\langle \frac{d\nu_e}{dE_e} \right\rangle$ is averaged over the inner 0.5 kpc (in keeping with the small FOV). The factors $\ell_{IC}$ and $\ell_{Br}$ depend on the details of the deprojected ISRF and gas densities. For a consistent solution we expect them to be $O$(kpc).

The GCE, IC and bremsstrahlung spectra in the case of a minimal “democratic” $e^\pm : \mu^\pm : \tau^\pm = 1 : 1 : 1$ annihilation channel is shown in Fig. 3.3 for particle mass $m_\chi = 8$ GeV and annihilation cross section $\langle \sigma v \rangle = 3.6 \times 10^{-26}$ cm$^3$s$^{-1}$. The best-fit dark matter mass when fitting to the GCE excess is closer to 7 GeV. In this model, the gamma rays from the $\tau^\pm$ dominate the prompt flux and explain the GCE excess. In Fig. 3.3, we have shown IC and bremsstrahlung model spectra using $\ell_{Br} = \ell_{IC} = 1.3$ kpc. We caution the reader that no attempt has been made to fit to all three components simultaneously.

The value of the cross section used to create the spectra in Fig. 3.3 is ruled out by AMS-02 constraints on WIMP annihilation to leptonic channels [39]. However, the required cross section depends sensitively on the assumed density profile. For example, if we assume a scale radius ($r_s$) of 10 kpc and $\gamma = 1.2$ (keeping the local density unchanged at 0.3 GeV/cm$^3$), the required cross section is a factor of 10 smaller. We have checked that such a profile is consistent with the expectation that the dark matter halo undergoes adiabatic contraction [43] due to the formation of the disk and bulge of stars. On the particle physics side, some of this tension may be relieved by considering annihilation through a vector mediator, which
softens the final $e^\pm$ spectrum.

### 3.5 Conclusions

We have detected a new excess in gamma rays toward the Galactic Center that spatially traces starlight intensity. The spectrum of this new source is consistent with that produced by high energy electrons and positrons with energies up to about 10 GeV, upscattering starlight. The population of electrons and positrons required to produce such an inverse Compton emission would also produce bremsstrahlung radiation due to interactions with the dense gas at the Galactic Center. Further studies are required to examine the physical implications of this high energy electron population and perform more detailed modeling of the predicted IC excess morphology and spectrum. We detect a gamma-ray excess tracing 20 cm radio map and show that its flux spectrum is consistent in both shape and amplitude with bremsstrahlung radiation from the same population of electrons and positrons. We show that the Galactic Center extended excess that peaks around 2 GeV is also detected at high significance and that a dark matter model with annihilation to leptons may provide a consistent explanation for all three excesses.

While this consistency with three excesses in terms of signal strengths, spectra and morphologies is remarkable, other astrophysical explanations exist that deserve detailed investigations. Infrared, radio and gamma ray data as incorporated in this work has allowed complicated high-energy emissions seen toward the Milky Way’s Galactic Center to be disentangled, and this multiwavelength approach may help us to further elucidate the true origin of these excesses.
Chapter 4

Lepton-Flavor Violating Mediators

Based on Galon, Kwa, & Tanedo, JHEP 1703 (2017) 064 [111]

We study a scenario in which dark matter annihilates into mediators $\varphi$, which then decay into pairs of differently flavored leptons. This work is motivated by the desire to produce the gamma-ray excess observed at the galactic center while simultaneously avoiding tensions with other experimental results. We model the galactic center excess as an inverse Compton scattering signal arising from mediator decays into $e^\pm/\mu^\pm/\tau^\pm$ pairs of different flavor, and check whether the final $e^\pm$ spectra produced via these dark matter annihilations are in tension with the $e^\pm$ fluxes observed by the Alpha Magnetic Spectrometer (AMS-02) at Earth's position.

4.1 Background

Scenarios where dark matter is a thermal relic that interacts directly with the Standard Model are typically constrained by a range of complementary experimental searches [29].
On the other hand, if dark matter is secluded from the Standard Model and only inter-
acts through a light mediator, then one may obtain the observed relic density from thermal 
freeze out while parametrically suppressing signatures in direct detection and collider ex-
periments [195]. Direct annihilation into on-shell mediators sets the dark matter–mediator 
couplings, while the mediator–Standard Model couplings may be chosen small enough to ex-
plain the null results from direct detection and collider experiments. Typically one assumes 
that the mediator decays into Standard Model states: if the mediator is stable, then it is a 
dark matter component that couples directly to the Standard Model. If, on the other hand, 
it decays into lighter hidden sector states, then those states are constrained by cosmological 
limits on the number of light degrees of freedom. Dark matter continues to annihilate in 
the present day and the Standard Model byproducts of the ensuing mediator decays may be 
observable.

One possible signature consistent with this framework is the recent excess of $\gamma$-rays from 
the Galactic Center observed by independent analyses of the Fermi Large Area Telescope 
(LAT) data [116, 126, 3, 49, 127, 6, 117, 130, 133, 181, 164, 4, 87, 236, 61, 60, 62] and 
directly by the Fermi-LAT collaboration [22]. Alternative explanations include unresolved 
pulsars [3, 6, 125, 4, 232, 192, 32, 158, 168, 129] or cosmic ray outbursts [191, 66, 67, 69]. 
The excess can be fit to effective theories that describe the annihilation of dark matter into 
pairs of Standard Model particles [47, 25, 135, 136, 41, 19]. Intriguingly, the total flux of 
excess $\gamma$-rays is consistent with the $s$-wave dark matter annihilation cross-section required 
to produce the observed relic density after thermal freeze out. Early fits to the energy 
spectrum preferred 40 GeV dark matter annihilating to $b$-quarks or 10 GeV dark matter 
annihilating into $\tau$-leptons; however, later studies found that masses up to $\mathcal{O}(100$ GeV) and 
a range of final states are allowed after accounting for the systematic uncertainties in the 
astrophysical background models [236, 61, 60, 19, 62, 22]. When dark matter annihilates 
into on-shell mediators in the secluded dark matter framework, the target space is shifted 
towards heavier dark matter and larger annihilation cross-sections [46, 8, 167, 197, 101].
Most proposals to explain the excess from dark matter annihilations focus on γ-ray emission from bremsstrahlung and π⁰ decays of annihilation products. These processes produce prompt γ-rays at the site of annihilation with intensities directly proportional to the square of the dark matter density. This predicts a similar signal in dwarf spheroidal galaxies which are rich in dark matter and have simpler astrophysical backgrounds than the Galactic Center [7]. The non-observation of a clear excess in dwarf spheroidal galaxies [97, 13] is typically a tension in dark matter interpretations of the γ-ray excess, indications of a possible excess in Reticulum II [113, 128] notwithstanding.

Lacroix, Boehm, and Silk [155] pointed out that another mechanism by which the Galactic Center excess might be generated is through the inverse Compton scattering (ICS) of final-state electrons and positrons with infrared starlight. The energetic e⁺e⁻ pairs up-scatter the low-energy photons into the GeV range. Recently, Calore et al. and Kaplinghat et al. proposed the possibility that these electron pairs may result from the decay of on-shell mediators [60, 145]. Up-scattering of starlight in dwarf satellites produces significantly less ICS emission at GeV energies due to their much weaker interstellar radiation field. This removes the tension between the Fermi Galactic Center and dwarf γ-ray observations. In its simplest form, however, this scenario is in tension with a different astrophysical observation. Direct dark matter annihilation into e⁺e⁻ pairs produces a line in the local e⁺e⁻ spectrum that is observable by the Alpha Magnetic Spectrometer (AMS-02) telescope. The absence of such a line requires a mechanism to soften the primary e⁺e⁻ spectrum. Kaplinghat, Linden, and Yu realize this in the secluded dark matter scenario in which the annihilations into light mediators smears the spectrum of daughter electrons and positrons. When there is a hierarchy in the dark matter χ and mediator φ masses, m_χ ≫ m_φ ≫ m_e, the production spectrum of electrons is smeared from a line at m_χ to a box from 0 to m_χ. In this way, one may attempt to hide the electron–positron spectrum by smearing it out within the AMS-02 error bars. The benchmark model in [145] invokes 50 GeV dark matter annihilating into 100 MeV spin-1 mediators. As an additional feature, these masses automatically furnish the
ingredients for a self-interacting dark matter solution to outstanding small-scale structure anomalies as reviewed in [220, 147]. The resulting dark matter annihilation cross-section is smaller than that of a thermal relic in the visible sector and thus requires the dark sector temperature to differ from the Standard Model at freeze out.

In this work, we introduce an alternative class of models that produce a $\gamma$-ray signal in the galactic center while avoiding bounds from the non-observation of such a signal in dwarf galaxies. Like Calore et al. and Kaplinghat et al., dark matter annihilates into on-shell mediators and the $\gamma$-ray signal is produced from the up-scattering of starlight. In our case, however, the mediator is a spin-0 particle which decays into different-flavor lepton pairs, $\tau\mu$, $\tau e$ or $\mu e$. The heavier leptons subsequently decay into electrons so that the final $e^+e^-$ spectrum is softer which allows it to better fit within the error bars of the AMS-02 observations. In the models with $\tau$ couplings, the hadronic $\tau$ decays yield prompt photons which contribute to the $\gamma$-ray excess in the absence of an interstellar radiation field; this may re-introduce some tension with the dwarfs, though this tension is lessened here compared to direct annihilation models. As observed in [8, 197], the leading $s$-wave contribution to dark matter annihilation into pseudoscalars is the $2 \rightarrow 3$ process that further softens the spectrum of ensuing the Standard Model decay products. We show that the resulting $\gamma$-ray spectra are consistent with the Fermi-LAT excess for an annihilation cross-section that is compatible with the correct dark matter relic density even when the dark sector and Standard Model have the same temperature at freeze out. In the case where the mediator also contains a spin-0$^+$ component, this framework may still furnish a self-interacting dark matter solution to small-scale structure anomalies.

### 4.2 Models of a Lepton-Flavor Violating Mediator

We present our model in this section; the interactions are summarized in Fig. 4.1.
4.2.1 Dark Sector Interactions

Fermionic dark matter $\chi$ is proposed to interact with a spin-0 mediator, $\varphi$, through the interactions,

$$\mathcal{L}_{\varphi\chi} = \frac{1}{2} y_S \eta \varphi \bar{\chi} \chi + i \frac{y_P}{2} \eta \varphi \bar{\chi} \gamma^5 \chi$$

$\eta = 1$ (2) for Majorana (Dirac) $\chi$. (4.1)

If $\varphi$ is complex, one must also include $\mathcal{L}_{\varphi\chi}^\dagger$. Unlike the case of a vector mediator for which there is no coupling to Majorana dark matter, the spin-0 mediator can couple to either Dirac or Majorana fermions. In writing these interactions, we assume that the fermion mass is manifestly real so that the pseudoscalar interaction is physical. See, for example, Ref. [102] for a demonstration of how a complex fermion mass term—say, if $\langle \varphi \rangle$ and $y_P \neq 0$—would shift the pseudoscalar interaction upon performing a chiral rotation to make the mass term real. For the remainder of this manuscript we assume that $\langle \varphi \rangle = 0$.

In the secluded dark matter scenario where annihilation goes into on-shell mediators, the couplings of (4.1) control the annihilation rate independently of the mediator–Standard Model couplings. A useful benchmark is the annihilation cross-section required for $\chi$ to reproduce the observed dark matter density assuming that it was initially in thermal equilibrium with
the Standard Model [216],

\[
\langle \sigma v \rangle_{\text{rel.}} = 2.2 \eta \times 10^{-26} \text{cm}^3/\text{s}.
\] (4.2)

We present targets for these couplings implicitly as a ratio of the required present day annihilation cross-section, \( \langle \sigma v \rangle \) set by the Fermi \( \gamma \)-ray excess, to the thermal relic cross-section, \( \langle \sigma v \rangle_{\text{rel.}} \). In doing so, we cancel all dependence on \( \eta \) in our results. The extent to which \( \langle \sigma v \rangle \) is compatible with \( \langle \sigma v \rangle_{\text{rel.}} \) is a useful metric of attractiveness for our models. Compatibility is simplest when annihilation is dominantly s-wave since higher partial waves are highly velocity suppressed in the present day. In this case, the dark matter and mediator masses are in the range

\[
m_{\chi} \sim \mathcal{O}(10 - 100 \text{ GeV}) \quad m_{\ell_i} + m_{\ell_j} < m_{\varphi} < m_{\chi} \left( \frac{2}{3} m_{\chi} \right)
\] (4.3)

where the value in parentheses corresponds to the \( 2 \rightarrow 3 \) annihilation.

We assume that the mediator mass \( m_{\varphi} \) is sufficiently light such that these annihilation modes are accessible. From the point of view of softening the positron spectrum, one may wish to consider mediator masses lighter than one of the leptons to which it couples. This forces the \( \varphi \) decay to be multi-body and further softens the spectrum. We do not consider this possibility because the constraints from two-body decays of heavy leptons [23, 34], typically imply that the mediator is too long-lived and would be ruled out by cosmological bounds.

### 4.2.2 Self-Interacting Dark Matter

The scalar coupling \( y_S \) in (4.1) generates a long-range Yukawa potential between dark matter particles. This is the key ingredient for how dark matter self-interactions address small scale structure anomalies [220]. The lightness of the mediator introduces a velocity-dependence
in the self-scattering cross-section; this affects the dark matter halo profile on the scales of
dwarf galaxies, while remaining consistent with constraints from galaxy cluster mergers [122].
In the model by Kaplinghat, Linden, and Yu, a vector mediator, $V$, produces the Fermi $\gamma$-
ray excess by the inverse Compton scattering of $\chi \bar{\chi} \rightarrow VV \rightarrow 4\nu$ [145]. The vector mass,
$m_V \sim O(10 - 100 \text{ MeV})$, and coupling to dark matter, with a transfer cross-section of
$\sigma_T \sim O(0.5 - 50 \text{ cm}^2/\text{g})$, were found to be of the correct size to realize this self-interacting
dark matter target region. This came at the cost of some tension with the thermal relic
cross-section, $\langle \sigma v \rangle_{\text{rel}}$, so that they invoke a different dark sector temperature [105].

In the scalar models here, only the parity-even dark sector interaction in (4.1) mediates
a Yukawa potential. The parity-odd interaction mediates a spin dependent potential that
scales as $e^{-m_\phi r}/r^3$ [38]; this is not expected to have a significant effect on astrophysical
dynamics. We thus observe that the scenario with a pure pseudoscalar mediator ($y_S = 0$)
does not realize the self-interacting dark matter target region. On the other hand, in the
scenario where $\phi$ has mixed parity, s-wave $\chi \bar{\chi} \rightarrow \phi \phi^{(*)}$ annihilation depends on both $y_S$ and
$y_P$. This introduces some freedom to choose $y_S$ to realize a large self-interaction cross-section
and then separately choose $y_P$ to select the annihilation cross-section, $\langle \sigma v \rangle$.

Our models also differ from Kaplinghat et al. because the minimum mediator mass scale is
set by the heavier lepton to which the mediator couples. Thus the lightest mediator mass we
consider is $m_\phi \sim m_\mu = 106 \text{ MeV}$ which is accessible for mediators with $\mu \nu$ couplings. Observe
that this mass is near the heavy limit of mediator masses that are compatible with solving
small scale structure anomalies [220]. We then expect that the case where the mediator
couples to a $\tau$ are typically incompatible with the self-interacting dark matter target region.

While a detailed study of the dark matter self-interactions in this model is beyond the scope
of this paper, the benchmark results in Ref. [220] already demonstrate the key properties.
We specifically note that the dark matter and mediator masses considered here populate the
numerically difficult resonant regime where consistency with the self-interacting dark matter
target region is plausible but very sensitive to the precise values of $m_{X,\varphi}$. For this reason, in this manuscript we focus on the compatibility of the Fermi $\gamma$-ray excess in our scenario with the thermal relic cross-section without invoking a different dark sector temperature. We leave the details of the dark sector self-interactions—which we emphasize are automatic in our constructions—for separate work.

4.3 The Fermi-LAT $\gamma$-Ray Excess

4.3.1 Photons from Leptons

The spectrum of photons to be identified with the Fermi $\gamma$-ray excess originate from two sources:

1. Prompt photons from the final state leptons or

2. Up-scattered starlight from the inverse Compton scattering (ICS) of $e^\pm$ produced through the $\varphi$ decay.

This is in contrast to models where dark matter annihilates predominantly into quarks or gluons. In that case the photons are a result of $\pi^0 \rightarrow \gamma\gamma$ decays from the showering and hadronization of the final state partons. Since the $\tau$ has a large hadronic decay width ($\sim 65\%$), its spectrum of prompt photons is similar to that of quarks and gluons. In contrast, electrons and muons (and leptonically decaying taus) typically produce a smaller flux of prompt photons, but can yield a large number of up-scattered photons from inverse Compton scattering in a stellar radiation field. The scattered photon energy, $E'_\gamma$ can be approximated

69
in terms of the incoming photon energy, $E_\gamma$ and the scattering electron energy, $E_e$

$$E'_\gamma \approx \left( \frac{E_e}{m_e} \right)^2 E_\gamma .$$

(4.4)

In the $\chi\bar{\chi} \rightarrow VV \rightarrow 4e$ model of Kaplinghat et al., $m_\chi \sim 10$ GeV so that $(E_e/m_e) \sim 10^4$. We see that $O(10 \text{ eV})$ starlight is then up-scattered to $O(\text{GeV})$, corresponding to the characteristic scale of the Fermi-LAT $\gamma$-ray excess. In contrast, the $e^\pm$ energy spectrum resulting from the scalar mediator decays $\varphi \rightarrow \bar{\ell}_i \ell_j (i \neq j)$ and subsequent decays to electrons is softer as part of the energy is deposited in neutrinos. Working in the limit $m_\ell \ll m_\varphi$, the average $e^\pm$ energies in the $e\mu, e\tau, \mu\tau$ scenarios, are approximately suppressed by $\sim \frac{2}{3}, \frac{11}{18}, \frac{5}{18}$ respectively\footnote{for $\tau$s we have only taken into account the leptonic decay modes} with respect to a $\varphi \rightarrow e^+e^-$ scenario. We therefore expect that the average ICS photon energy to be approximately an order of magnitude softer in the flavor violating case.

In order to properly estimate the photon spectrum, we use the Mathematica package PPC [73, 56]. As an input, PPC requires the flavor dependent energy spectrum distributions of the leptonic annihilation products,

$$\frac{dN_{\ell_j}}{dE_{\ell_j}} = \int dE_\varphi \frac{dN_\varphi}{dE_\varphi} \frac{dN_{\ell_j}(E_\varphi)}{dE_{\ell_j}} ,$$

(4.5)

where $dN_\varphi/dE_\varphi$ is the spectrum of mediators and $dN_{\ell_j}(E_\varphi)/dE_{\ell_j}$ is the spectrum of $j$-type leptons produced in the decay of a mediator with energy $E_\varphi$. We consider the two annihilation modes mentioned above, $\chi\bar{\chi} \rightarrow 2\varphi$ and $\chi\bar{\chi} \rightarrow 3\varphi$. In the first case, the $\varphi$ energy spectrum is monochromatic, $dN_\varphi/dE_\varphi = \delta(E_\varphi - m_\chi)$, so that the boost of lepton energies from the $\varphi$ rest-frame is straightforward, and given by

$$E_\ell = \gamma E^0_\ell + \sqrt{\gamma^2 - 1}|p^0_\ell| \cos \theta ,$$

(4.6)
where $E_\ell^0$ and $p_\ell^0$ are the lepton energy and three-momentum in the $\varphi$ frame, and $\gamma = E_\varphi/m_\varphi$ is the boost to the lab frame. The lepton energy distribution is then box-shaped because the cosine of the azimuthal angle $\cos \theta$ is uniformly distributed over its range. For annihilation into three mediators, these distributions are implemented following the discussion in Ref. [197] (see [166, 101] for related discussions on cascade decays from mediators).

Given $\frac{dN}{dE_\ell_j}$, Pythia [209] results are interpolated in PPPC to calculate the secondary prompt photon energy spectrum,

$$
\frac{dN_\gamma}{dE_\gamma} = \sum_j \int dE_\ell_j \frac{dN_{\ell_j}}{dE_\ell_j} \frac{dN_\gamma^\ell(E_\ell_j)}{dE_\gamma},
$$

(4.7)

by accounting for photon final state radiation, and for decays of hadrons to photons in the case of $\tau$s. Here $dN_\gamma^\ell(E_\ell)/dE_\gamma$ is the spectrum of photons produced from a lepton $\ell$ with energy $E_\ell$. In addition, PPPC decays muons and taus and calculates the total $e^\pm$ energy spectrum in the galactic center region. In turn, this spectrum is used as the initial $e^\pm$ injection spectrum from which the resulting propagation, energy losses, and ICS flux may be calculated [155]. In calculating the ICS spectrum of photons, we use the MED set of diffusion parameters as described in [94] and a Navarro-Frenk-White (NFW) halo profile for the dark matter distribution with inner profile slope of $\gamma_{\text{NFW}} = 1.0$ and a local dark matter density of $\rho_\odot = 0.4 \text{ GeV cm}^{-3}$ [205]. If one uses a contracted NFW profile with a higher inner slope of $\gamma_{\text{NFW}} \approx 1.2$, as suggested in Refs. [4, 87, 61] and used by Kaplinghat et al., the required cross-section is reduced by a factor of $\sim 3$ to match the $\gamma$-ray excess intensity. Varying the diffusion parameters across the range of uncertainties as described in [94] does not significantly affect the resultant ICS spectrum. However, their theoretical uncertainties are set by modeling the propagation from across much longer ranges than our region of interest. The actual range of uncertainties on diffusion parameters for galactic center are unknown and may well be much larger.
4.3.2 Fit to Fermi $\gamma$-ray Excess

Figs. 4.2 and 4.3 show the photon spectrum prediction for the $15^\circ \times 15^\circ$ region of the sky centered at galactic coordinates $(l, b) = (0, 0)$. We considered each lepton flavor model and each dark matter annihilation mode separately. We tested several $m_\chi$ benchmarks, each of which is plotted in a different color. In each benchmark, the range of $m_\varphi$ masses considered is accounted for by the thickness of each plotted color. The color edges interpolate the range $m_\varphi \in [m_\ell^{\text{heavy}}, m_\chi (\frac{2}{3} m_\chi)]$ for the two-(three-) $\varphi$ annihilation mode.

Each plot is shown with a fixed benchmark annihilation cross-section. The spectrum scales linearly with this cross-section $\langle \sigma v \rangle$ and quadratically with the local dark matter density $\rho_\odot$. To aid in rescaling estimates, we also provide the cross-sections for a contracted NFW profile $\gamma_{\text{NFW}} = 1.2$ which would correspond to the same spectrum normalizations as in the $\gamma_{\text{NFW}} = 1.0$ halo. Note that the contracted profile only contains a rescaling by the $J$-factor. Because the interstellar radiation field is not uniform in the region of interest, it is possible that the contracted profile may lead to a change in the ICS spectrum. The range of cross-sections can be interpreted as an estimate of uncertainty when comparing to the thermal relic cross-section.

For comparison to the observed Fermi-LAT spectrum, we plot in grey the systematic error band of the $\gamma$-ray excess spectrum as defined in the Fermi collaboration study of the $\gamma$-ray emission from the galactic center [22]. The Fermi collaboration provides two different estimates of the excess $\gamma$-ray spectrum and its systematic uncertainty:

1. The first fits the excess as a parameterized exponential cutoff spectrum across the entire energy range of the data. This is shown in Fig. 4.2.

2. The second fits the $\gamma$-ray spectrum in independent energy bins. This is shown in Fig. 4.3.
Figure 4.2: The predicted combined prompt and ICS gamma-ray spectra for each scenario: 
$\chi \chi \rightarrow 2\varphi$ (left) and $\chi \chi \rightarrow 3\varphi$ followed by the lepton-flavor violating decays indicated in each panel. The grey shaded region represents the Fermi collaboration’s $\gamma$-ray excess spectrum bounded by its estimated systematic error when fit as a parameterized form to the entire energy range of the data. Each color-coded band corresponds to a set of $\{m_\chi, m_\varphi\}$ with $m_\varphi$ varying in the range $[m_{\text{heavy}}, m_\chi]$ (left) and $[m_{\text{heavy}}, \frac{2}{3}m_\chi]$ (right). The spectra are calculated assuming a halo profile slope of $\gamma_{\text{NFW}} = 1.0$ and the annihilation cross sections indicated in each figure. For a steeper halo profile of $\gamma_{\text{NFW}} = 1.2$, cross sections are a factor of $\sim 3$ smaller.
Figure 4.3: Same as Fig. 4.2, but now the grey shaded region represents the \( \gamma \)-ray excess spectrum bounded by its estimated systematic error when fit in independent energy bins, as reported by the Fermi collaboration. Our predicted combined prompt and ICS gamma-ray spectra for each scenario: \( 2 \rightarrow 2 \) (left) and \( 2 \rightarrow 3 \) (right) \( \chi \chi \) annihilations to \( \varphi \)s, followed by one of the following decays: \( \varphi \rightarrow \tau e \) (up), \( \tau \mu \) (middle), and \( \mu e \) (bottom). Each color-coded band corresponds to a set of \( \{ m_\chi, \ m_\varphi \} \) with \( m_\varphi \) varying in the range \([ m_{\text{heavy}}, \ m_\chi ]\) (left) and \([ m_{\text{heavy}}, \ 2^3 m_\chi ]\) (right).
Flavor-dependence of $\gamma$-ray spectra. The $\mu e$ final states result in much harder $\gamma$-ray spectra than $\tau e$ or $\tau \mu$ final states. In the case of $\mu e$ final states, the ICS contributes the majority of the $\gamma$-ray flux at lower, $O(1 \text{ GeV})$, energies while the prompt contribution dominates at higher, $O(10 \text{ GeV})$, energies. In contrast, the ICS $\gamma$-ray flux in the cases of $\tau e$ and $\tau \mu$ final states only constitutes a small fraction of the low energy spectrum, while the total signal is dominated by the prompt photon ‘bump’, which peaks between $2 - 10 \text{ GeV}$ before a spectral cutoff. This is because the hadronic $\tau$-decays allow for these annihilation channels to produce a much higher flux of prompt photons.

Comparison to Fermi spectra. Our models are able to reasonably reproduce the parameterized Fermi $\gamma$-ray excess spectrum in Fig. 4.2. The $\tau e$ and $\tau \mu$ model achieve this with dark matter masses of $m_\chi \sim 20 - 40 \text{ GeV}$, while for the $\mu e$ case, slightly higher masses of $m_\chi \sim 40 - 100 \text{ GeV}$ are required. This result corresponds to the previous point: muons and electrons produce less prompt emission than $\tau$s. The dark matter mass for the $\mu e$ final state must be higher than in the $\tau e$ or $\tau \mu$ cases in order for the resulting electron spectrum to be hard enough to produce the Fermi $\gamma$-ray excess through ICS. We note that although the gamma-ray spectra produced in the $\mu e$ models may lie within the systematic error band defined by the Fermi collaboration’s parametric fits, they are generally harder at high energies and do not have the characteristic peak at $\sim 2 - 4 \text{ GeV}$ that is typically found in template analyses of the $\gamma$-ray excess.

The ‘cinched’ shapes of the enveloped range of gamma-ray emission in the $\tau e$ channels of Figs. 4.2 and 4.3 arise due to a sampling effect: the spectra of the outgoing $\tau$'s and $e$'s do not change significantly until the mediator mass approaches its allowed minimum (the heavier lepton mass). The spectra for the three heavier mediator masses in the $\tau e$ case are very similar to each other, leading to an enveloped range that is effectively defined by only two unique spectra.
On the other hand, Fig. 4.3 shows that our computed spectra are worse at fitting the Fermi collaboration’s γ-ray spectra obtained through bin-by-bin fits. The spectrum derived through fitting the data in individual energy bins displays an extended, power-law-like tail at energies ≥ 10 GeV; this was observed in [60, 62] and has recently been explored further in [132]. Our theoretical gamma-ray spectra all cut off sharply around $\mathcal{O}(m_\chi)$ and thus cannot reproduce this spectral feature. One can interpret the difference between the parameterized and bin-by-bin fits as a qualitative assessment of the uncertainty in the target region for the spectral fit.

**Compatibility with relic abundance.** For $\tau e$ and $\tau \mu$ final states, we are able to produce the observed Galactic Center excess flux with an annihilation cross-section $\langle \sigma v \rangle$ roughly equal to the relic density cross-section, (4.2). For $\mu e$ final states, the annihilation cross-section must be 3–4 times higher than the canonical relic cross-section in order to match the intensity of the Fermi Galactic Center excess. This is again related to the fact that there are substantially less prompt photons in the $\mu e$ scenario, hence the larger annihilation rate needed to account for the excess. We point out, however, the $\mu e$ states are brought back into consistency with (4.2) if one instead invokes a contracted NFW profile.

**Dwarf spheroidal bounds.** Dwarf spheroidals are satellite galaxies that are rich in dark matter but with relatively little stellar matter. As a result, they typically set the strongest bounds on models of the galactic center excess that rely on prompt photon emission [13]. This is avoided when the $\sim$GeV photons are produced through the inverse scattering of starlight because the dwarfs have a weak interstellar radiation field. Thus the $\mu e$ models are able to completely evade the dwarf bounds.

For $\tau e$ and $\tau \mu$ channels, the decay into two differently flavored leptons means that the cross-section for annihilation to $\tau$'s is half of the total annihilation cross-section. We also note...
that the range of dark matter particle masses considered here for annihilations to mediators are generally higher than the best-fit masses in the case of direct annihilations, since either four or six SM leptons are produced per annihilation in these models (instead of two in the direct scenario). The dwarf constraints on the annihilation cross-section into $\tau$’s are roughly $\sim 1.5 - 2$ times weaker at $m_\chi \sim 20 - 40$ GeV compared to $m_\chi \sim 10$ GeV, which is often quoted as the best-fit mass for direct annihilations into $\tau$’s. These two effects combine to partially alleviate the existing tensions with dwarf constraints on prompt gamma-ray flux from annihilation into $\tau$’s; the cross-sections required in our model for $\tau e$ and $\tau \mu$ channels are within 1$\sigma$ of the limits from Ref. [13].

**Comparison to Kaplinghat et al.** We briefly compare our results to the $\chi \bar{\chi} \rightarrow VV \rightarrow 4e$ scenario; in doing so, we may highlight the differences in the lepton-flavor violating case and the role of uncertainties in astrophysical parameters. Kaplinghat et al. [145] found that the annihilation mode to $4e$ fits the Galactic Center excess for an annihilation cross-section of $\langle \sigma v \rangle \approx \langle \sigma v \rangle_{\text{rel.}}/7$. By comparison, our $2(\mu e)$ annihilation mode is found to fit with $\langle \sigma v \rangle \approx 3\langle \sigma v \rangle_{\text{rel.}}$. The difference between these factors are predominantly from three sources:

1. Our result uses a standard $\gamma_{\text{NFW}} = 1.0$ dark matter halo profile while Kaplinghat et al. use a contracted $\gamma_{\text{NFW}} = 1.2$ profile. As shown in the plot, the difference in $J$ factors is approximately a factor of 3.

2. The discussion in Sec. 4.3.1 shows that electrons and muons produce very different ICS spectra. This is in contrast to prompt photon production where the mass difference can be ignored at sufficiently high energies. This introduces an $\mathcal{O}(\text{few})$ difference between the flavor violating and flavor conserving modes.

3. Finally, there are some $\mathcal{O}(\text{few})$ differences in the target region, Kaplinghat et al. use the fit by [61] while we use the envelopes from the Fermi collaboration [22].
4.4 The AMS-02 $e^\pm$ Spectrum

The AMS-02 experiment may be capable of detecting electrons and positrons produced in dark matter annihilations [39]. AMS observations thus far have found that the $e^\pm$ spectra are smoothly varying, with no line-like or bump features [20]. We constrain the range of particle properties allowed in our model by requiring that the flux of $e^\pm$ produced through annihilations and propagated to Earth’s position in the Milky Way must be low enough to avoid producing any such features in the observed $e^\pm$ spectra. To do so, we use the DRAGON 3D cosmic ray propagation code along with the propagation setup described in Ref. [109].

The diffusion coefficient is assumed to depend on particle rigidity as

$$D(\rho) = \beta^{-0.4} D_0 \left( \frac{\rho}{\rho_0} \right)^{\delta} \quad \text{with} \quad D_0 = 3 \times 10^{28} \text{ cm}^2 \text{s}^{-1} , \quad \rho_0 = 3\text{ GV} , \quad \delta = 0.6 . \quad (4.8)$$

We assume the same halo parameters and annihilation cross-section for each case of lepton final states and number of mediators as in the previous section.

Fig. 4.4 shows the expected $e^\pm$ energy spectrum from $\chi\bar{\chi}$ annihilation to $\varphi$ followed by the decay $\varphi \rightarrow \bar{\ell}_i \ell_j$. Following the notation of the previous section, results are presented for each annihilation mode and each of the three $\varphi$ decay models separately. We fix the dark matter annihilation rate of each model to roughly match the observed Galactic Center $\gamma$-ray excess flux. We regard a model as consistent with the AMS observations if for all energies, the $e^\pm$ flux predicted by the model is lower than the total size of the error bars given in Ref. [20] for the $e^+$ and $e^-$ binned fluxes at that energy.

We find that for $\tau e$ and $\tau \mu$ final states, dark matter masses above $\sim 20$ GeV are not excluded by AMS observations, and can also produce a $\gamma$-ray signal consistent with the Fermi excess. For $\mu e$ final states, all potential $e^\pm$ spectra studied here are in tension with the AMS positron bounds when $\gamma_{\text{NFW}} = 1.0$. This tension may be reduced by either
Figure 4.4: Predicted $e^\pm$ spectra at Earth’s position for each scenario: $\chi \bar{\chi} \rightarrow 2\varphi$ (left) and $\chi \bar{\chi} \rightarrow 3\varphi$ (right) followed by one the lepton-flavor violating decays indicated in each panel. The grey shaded region represents the bounds from AMS-02 electron and positron spectra while each color-coded band corresponds to a set of $\{m_\chi, m_\varphi\}$ with $m_\varphi$ varying in the range $[m_{\text{heavy}}, m_\chi]$ (left) and $[m_{\text{heavy}}, 2m_\chi]$ (right). The dashed lines shown in the bottom row correspond to the prediction for AMS-02 spectra if an NFW profile slope of $\gamma_{\text{NFW}} = 1.2$ is assumed. This leads to a lower annihilation cross section of 1.1 (1.4) times the relic density for $\chi \bar{\chi} \rightarrow 2(3)\varphi$. 

79
1. increasing the dark matter mass beyond 90–100 GeV, or

2. considering a steeper dark matter halo density profile.

Increasing the mass comes at the expense of hardening the $\gamma$-ray signal and introduces tension with the Fermi result. On the other hand, a slightly steeper NFW inner profile slope or $\gamma_{\text{NFW}} = 1.1 - 1.3$ was suggested in [87, 61] for the Fermi $\gamma$-ray excess. When assuming a steeper profile, the annihilation cross-section needed to reproduce the observed excess brightness decreases by up to a factor of $\sim 5$. Since the local dark matter density is held fixed as the density profile slope changes, this lower annihilation rate results in a decrease in the dark matter contribution to the AMS $e^\pm$ spectrum. The potential dark matter contribution to the local $e^\pm$ spectrum is dominated by the flux from annihilations near the solar neighborhood; changes to the profile near the galactic center have little effect on this measurement. The dashed lines in the $\mu e$ plots of Fig. 4.4 show the upper contour of the $e^\pm$ spectra for the estimated reduced annihilation cross-sections with a contracted $\gamma_{\text{NFW}}$ profile. One can see that the factor of $\sim 3$ in the cross-section allows these modes to avoid the AMS $e^\pm$ bound.

4.5 Conclusions & Outlook

In this manuscript we examined a class of models where dark matter interacts with the Standard Model through spin-0 mediators with chiral, flavor violating interactions to leptons. For a range of mediator masses, this setup realizes the secluded dark matter scenario where the relic abundance and indirect detection annihilation rates are controlled by one set of couplings, while direct detection, collider bounds, and low-energy searches are controlled by a separate set of couplings.

We have shown that in the dark sector, one is able to simultaneously achieve a thermal
Table 4.1: Summary of Figs. 4.2 and 4.3. The annihilation cross-section, $\langle \sigma v \rangle$ is given for $\gamma_{\text{NFW}} = 1.0 \ (1.2)$. The $\mu e$ modes are in tension with the AMS-02 positron bound unless one takes the contracted $\gamma_{\text{NFW}} = 1.2$ profile. The range of mediator masses are given in (4.3).

<table>
<thead>
<tr>
<th>Dark parity of $\varphi$</th>
<th>Annihilation</th>
<th>$\ell_i \ell_j$</th>
<th>$m_\chi$/GeV</th>
<th>$\langle \sigma v \rangle / \langle \sigma v \rangle_{\text{rel.}}$</th>
<th>AMS-02</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not parity eigenstate</td>
<td>$\tau e$</td>
<td>$20 - 40$</td>
<td>$1 \ (0.4)$</td>
<td>$\checkmark$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\tau \mu$</td>
<td>$20 - 40$</td>
<td>$1 \ (0.4)$</td>
<td>$\checkmark$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\mu e$</td>
<td>$30 - 90$</td>
<td>$3 \ (1.1)$</td>
<td>$\times \ (\checkmark)$</td>
<td></td>
</tr>
<tr>
<td>Pseudoscalar (parity-odd)</td>
<td>$\tau e$</td>
<td>$20 - 40$</td>
<td>$1 \ (0.4)$</td>
<td>$\checkmark$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\tau \mu$</td>
<td>$20 - 40$</td>
<td>$1 \ (0.4)$</td>
<td>$\checkmark$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\mu e$</td>
<td>$40 - 100$</td>
<td>$4 \ (1.4)$</td>
<td>$\times \ (\checkmark)$</td>
<td></td>
</tr>
</tbody>
</table>

relic and the observed Fermi $\gamma$-ray excess without causing tension with the measured AMS-02 positron spectrum. For final states containing $\tau$s, the $\gamma$-ray excess is mostly produced through prompt gamma-ray emission from hadronic $\tau$ decays, with some contribution at lower energies from inverse Compton scattering of the interstellar radiation field off of electrons from $\mu$ or $e$ final states. We find that even in these cases, the annihilation cross section to $\tau$s needed to reproduce the observed excess is consistent with bounds from dwarf spheroidals, within the uncertainties. For $\mu e$ final states, the excess is predominantly produced by inverse Compton scattering. Because dwarf spheroidals have much weaker interstellar radiation fields, the inverse Compton scattering contribution to the excess helps to alleviate (or eliminate, for $\mu e$ interactions) tensions of the $\gamma$-ray excess with non-observations in dwarf spheroidals. For the case of $\mu e$ interactions, however, this typically requires assuming a contracted dark matter halo profile.

The spectra of the Standard Model byproducts of dark matter annihilation are softened because the decay goes through on-shell mediators. This smearing helps the $e^+e^-$ spectrum to fit within the error bars of the AMS experiment. We have commented that the parameters for the $\mu e$ final state appear to be consistent with the target region for a self-interacting dark matter solution to small scale structure anomalies. The dark sector interactions are
summarized in Table 4.1.

One unique feature of the chiral lepton-flavor violating interactions is that the bounds on the Standard Model couplings are weaker than direct flavor-conserving interactions. In other words, the lepton-flavor violating decay mode allows one to mix the $\gamma$-ray spectra of different final states while avoiding—rather combining—the bounds from flavor conserving decays into those final states. We have shown that the upper bounds on the Standard Model interactions in this scenario come from the forward–backward asymmetry in $e^+e^- \rightarrow f \bar{f}$ and from the anomalous magnetic moments of the muon and electron. For the case of a $\tau\mu$ interaction, one can simultaneously explain the $(g - 2)_\mu$ anomaly. We have shown that the chiral flavor structure of the mediator–Standard Model interaction suppresses bounds from conventional charged lepton flavor violation experiments and direct detection experiments. We explain that these suppressions are straightforward to understand from the point of view of a spurious $L_i - L_j$ symmetry that is respected by the mediator when it is complex. Finally, we point out possible opportunities in dark photon experiments and collider searches that are motivated by this class of mediator models.

We emphasize that while we have benchmarked our models for the Fermi $\gamma$-ray excess, the class of models are independently meaningful as an example of light, weakly-coupled new physics that can play an important role in both Standard Model and dark matter phenomenology and that are able to avoid current constraints.
Chapter 5

Hidden Sector Hydrogen as Dark Matter


We study the atomic physics and the astrophysical implications of a model in which the dark matter is the analog of hydrogen in a secluded sector. The self-interactions between dark matter particles include both elastic scatterings as well as inelastic processes due to a hyperfine transition. The self-interaction cross sections are computed by numerically solving the coupled Schrödinger equations for this system. We show that these self-interactions exhibit the right velocity dependence to explain the low dark matter density cores seen in small galaxies while being consistent with all constraints from observations of clusters of galaxies. For a viable solution, the dark hydrogen mass has to be in 10–100 GeV range and the dark fine-structure constant has to be larger than 0.01. This range of model parameters requires the existence of a dark matter–anti-matter asymmetry in the early universe to set the relic abundance of dark matter. For this range of parameters, we show that significant cooling losses may occur due to inelastic excitations to the hyperfine state and subsequent decays,
with implications for the evolution of low-mass halos and the early growth of supermassive black holes. Cooling from excitations to higher $n$ levels of dark hydrogen and subsequent decays is possible at the cluster scale, with a strong dependence on halo mass. Finally, we show that the minimum halo mass is in the range of $10^{3.5}$ to $10^{7} M_{\odot}$ for the viable regions of parameter space, significantly larger than the typical predictions for weakly interacting dark matter models. This pattern of observables in cosmological structure formation is unique to this model, making it possible to rule in or rule out hidden sector hydrogen as a viable dark matter model.

5.1 Background

While the standard ΛCDM cosmological model with collisionless, cold dark matter (CDM) is successful in explaining the observed large-scale structure in the Universe, there are many puzzles on galactic scales yet to be explained convincingly by a CDM-based scenario. $N$-body simulations of CDM structure growth predict cuspy radial density profiles ($\rho \sim r^{-1}$) with high central densities [173, 58, 95, 215]; however, observations of rotation curves reveal cored ($\rho \sim$ constant) or otherwise low-density inner regions in dark matter dominated galaxies, from dwarfs to low surface brightness (LSB) galaxies [171, 189, 88, 218, 208, 212, 154, 90, 96, 180, 17]. Galaxy clusters also show evidence for a deficit of dark matter within the effective stellar radius of the central galaxy, with the mass profile outside being consistent with CDM predictions [177, 175, 176]. The dark matter density profiles in dwarf spheroidal galaxies (dSphs) are a subject of current debate, with various studies finding that the stellar data for various dSphs is most consistent with a core [226, 27, 138, 26, 28], or a cusp [139, 200], or is unconstrained [53]. However, it seems clear that these galaxies are less dense than expected in pure CDM models [50, 51, 187]. The inclusion of supernovae and/or black hole feedback processes in cosmological simulations may ameliorate these anomalies in dwarf
galaxies by significantly altering the central gravitational potential \([118, 55, 163, 215, 179, 185, 148, 198, 228]\). However, it is unclear whether such effects are simultaneously able to affect the halo structure to the extent observed in low-mass \((M_* \sim 10^6-10^7 M_\odot)\) isolated dwarf galaxies \([106, 227, 188]\) and low surface brightness galaxies \([92]\). It is also not clear if the diversity of cores observed inferred from rotation curves can be explained in the context of these models \([91, 184, 186]\).

Self-interacting dark matter (SIDM) is an attractive solution \([213]\) to these anomalies that works across the range of mass scales under consideration, from dwarf galaxies to galaxy clusters \([201]\). In such a scenario, scatterings between dark matter particles allow for energy to be transferred from the hotter outer regions of the halo into the colder innermost regions. SIDM halos thus have hotter cores with higher velocity dispersions than the cold, cuspy interiors of collisionless dark matter halos, which lack a mechanism to heat the inner cusp into a core. \(N\)-body simulations \([225, 201]\) and analytic models based on these simulations find that the aforementioned issues in small-scale structure (on the dSph and LSB scales) may be alleviated if the dark matter is strongly self-interacting with a hard-sphere scattering cross section of \(0.5 \text{ cm}^2/\text{g} \lesssim \sigma/m \lesssim 5.0 \text{ cm}^2/\text{g} \) \([234, 100, 147, 224]\), where \(m\) is the mass of the dark matter. In order to produce cores of radius 10–50 kpc in cluster-sized halos in which the relative dark matter particle velocity is \(v \sim 1000 \text{ km/s}\), the required cross section is \(\sigma \sim 0.1 \text{ cm}^2/\text{g} \) \([147]\). We are thus motivated to consider SIDM models with velocity-dependent cross sections that are suppressed at cluster-scale velocities. Upper limits on the SIDM cross section may also be derived through the observed ellipticities of cluster-scale halos \(\sigma/m \lesssim 1 \text{ cm}^2/\text{g} \) \([190]\) and the measured center-of-mass offsets in merging cluster systems \(\sigma/m < 0.47 \text{ cm}^2/\text{g} \) \([122]\), but we find both these constraints to be weaker than those obtained from measurements of the inner density profiles of galaxy clusters \([147]\).

Atomic dark matter \([114, 170, 103, 142, 37, 84, 76, 108]\), in which the features of Standard Model (SM) hydrogen are copied to a dark sector, has all the features required of an SIDM
model to solve the small-scale puzzles. We consider a dark proton and dark electron, which are charged under an unbroken U(1) gauge group and may combine to form dark hydrogen. If the formation of dark hydrogen bound states is efficient, these dark atoms constitute approximately all of the dark matter. Since the dark hydrogen is a composite particle with an extended, finite size, its self-interaction cross section can be naturally large, as required for SIDM.

In this work, we consider atomic dark matter that exists today exclusively in bound states—dark recombination was fully complete. This model has uniquely testable phenomenology due to the ability of atomic dark matter to dissipate energy. We calculate and explore the cosmological consequences of both collisional scattering (which transfers energy between between dark atoms) and inelastic hyperfine upscattering (which results in energy loss through excitations and subsequent decays). These cross sections are velocity dependent, allowing the self-interactions to modify the halo profile to varying degrees in different astrophysical systems. The general trend is that the cross sections of particles in dwarf halos with characteristic velocities of \( \sim 40 \text{ km/s} \) will be larger than those in cluster halos with characteristic velocities of \( \sim 1000 \text{ km/s} \), which allows for regions of parameter space in which this model may resolve the aforementioned issues in small-scale structure. The heating rate from scatterings as well as the cooling losses from inelastic collisions vary widely depending on both the model parameters and the radial position in a halo of a given mass. The combined effects of both types of scatterings may thus lead to nontrivial effects on dark matter halo structure and evolution. At higher particle energies, additional atomic interactions, such as collisional excitations to the \( n = 2 \) state and ionization, may begin to affect the structure of cluster-sized halos.

For this interesting range of parameter space, where we see competing effects from collisional heating and cooling processes on the evolution of halos, we find additional features in the small-scale halo mass function that allow us to distinguish atomic dark matter from CDM.
cosmologically. Coupling between the dark matter and dark radiation produces dark acoustic oscillations, which are weakly constrained by measurements of the matter power spectrum and the cosmic microwave background (CMB) [83]. The most interesting effect of acoustic oscillations in the dark plasma would be a cutoff in the matter power spectrum set by the size of the fluctuations entering the horizon before the time of matter-radiation decoupling, resulting in a minimum dark matter halo mass that is significantly larger than in the typical weak-scale models without hidden sectors.

Altering small-scale structure with SIDM neither assumes nor requires any interactions between dark matter and SM particles beyond gravitational interactions; thus, we take a minimal approach and seclude the dark sector from the visible sector. Atomic dark matter has been presented in other contexts, such as a mirror universe [170, 107] and supersymmetry [37]. The effect of hidden sector dissipative dark matter on small-scale structure has been previously studied in Refs. [108, 107]. We note that our approach differs from prior works [108, 107]: the cored profiles in this work result from the collisional scatterings of the neutral dark atoms, whereas the density profiles in Ref. [108] are shaped by a combination of bremsstrahlung cooling processes in the dark sector as well as energy injection from visible supernovae [which is made possible through the inclusion of a kinetic mixing interaction between the dark U(1) and the SM hypercharge]. Kinetic mixing has also been used to explain DAMA and CoGeNT [142, 143, 78, 77, 108] and the 3.5 keV line [75].

The chapter is structured as follows. In Sec. 5.2 we describe how the estimated target SIDM cross section at low velocities is obtained from fitting to rotation curves. In Sec. 5.3 we describe the atomic dark matter model and the scattering properties of dark hydrogen. In Sec. 5.4 we consider dark hydrogen as an SIDM candidate and determine the parameter space allowed to accommodate SIDM and satisfy cosmological constraints. In Sec. 5.5 we investigate how inelastic scattering processes can affect halo formation. In Sec. 5.6 we discuss the possibility of upscatterings to the $n = 2$ excited state as well as collisional ionization in
cluster-scale halos. We conclude in Sec. 5.7.

## 5.2 SIDM cross sections at low velocities

As described in Chap. 1, one source of tension between CDM predictions and observations is the diversity of rotation curve shapes when considering galaxies with similar halo masses. Previous works have shown that SIDM halo models are able to resolve this tension through a combination of the following two effects[144, 147, 141].

1. Scatter in the cosmological concentration-mass relation (which is directly interchangable with the $V_{\text{max}} - r_{\text{max}}$ relation): At fixed halo mass, this scatter will lead to an associated scatter in the characteristic SIDM core radius $r_1$. The variety in the core sizes is reflected in the galaxies’ inner rotation curves.

2. Differences in baryonic mass distribution: As described in Ref. [144], the baryonic potential influences the thermalized cores of SIDM halos and is therefore also important in the determination of the core size $r_1$.

We use two different methods to fit SIDM halo models to rotation curves while also taking into account the effect of the baryons’ gravitational potential on the SIDM halo. Our data is taken from the SPARC dataset [159]. We interpolate between a grid of SIDM halo model templates and also fit using a Markov Chain Monte Carlo (MCMC) code. The rotation curves in our sample are best fit with self-interaction cross sections $\sigma/m$ of order $\mathcal{O}(1) \text{cm}^2/\text{g}$. Unlike rotation curve fits using collisionless WIMP/CDM halo models, SIDM halo models are able to fit rotation curve data across a diverse range of ‘cored’ and ‘cusped’ central density profiles. For an example of the range of rotation curves that can be fit using an SIDM halo model with $\sigma/m = 3 \text{cm}^2/\text{g}$, see Fig. 1.1. Fig. 5.1 below demonstrates two galaxies of similar mass which may both be well-fit by an SIDM halo model but not by a collisionless halo model.
Figure 5.1: Left: A galaxy (D631-7) where the dark matter halo density profile has a shallow rise, and is able to be fit well with a SIDM halo model with $\sigma/m = 3\text{cm}^2/\text{g}$ (upper left) but not with a collisionless NFW halo (upper right). Right: A galaxy (UGC 08490) with a steeply rising dark matter density profile, in which the rotation curve may be well-fit using either a SIDM halo model with $\sigma/m = 3\text{cm}^2/\text{g}$ (lower left) or a collisionless NFW halo (lower right).

These results, along with previous studies’ findings from N-body simulations [201, 100], motivate the ‘target’ viscosity cross section of $\sigma/m \simeq 0.5 - 5\text{cm}^2/\text{g}$ at velocities $v_{\text{rms}} = 30 - 100$ km/s that is used in this work.

### 5.3 Atomic Dark Matter Model

We begin this section by describing the properties and parameters of a secluded dark atomic model. We then present the formalism for dark hydrogen-hydrogen scattering and show results for the scattering cross section, obtained by numerically solving the Schrödinger equation. Our calculations agree with a previous detailed study for elastic scattering [76]. We further improve upon the basic model by incorporating inelastic processes that arise from hyperfine interactions. In subsequent sections, we show that the hyperfine splitting of ground-state dark hydrogen provides a rich phenomenological framework to study structure
We consider dark matter in a secluded sector that mimics the properties of SM hydrogen. The dark sector has two elementary particles: a dark electron with mass $m_e$ and a dark proton with mass $m_p$. These particles have opposite charge under an unbroken U(1) gauge, and they interact with a strength given by a dark fine-structure constant $\alpha$. The dark electron and dark proton may combine to form a neutral dark hydrogen atom with mass $m_H$ and a binding energy $B_H = \alpha^2 \mu_H / 2$, where $\mu_H$ is the reduced mass of the dark electron-proton system. It is convenient to parametrize the theory in terms of the following:

$$\mu_H = \frac{m_e m_p}{m_e + m_p}, \quad R \equiv \frac{m_p}{m_e}, \quad f(R, \alpha) \equiv \frac{m_H}{\mu_H} = R + 2 + \frac{1}{R} - \frac{\alpha^2}{2}. \quad (5.1)$$

Without loss of generality, we set $R > 1$. If $R \gg 1$, we may ignore the contribution from the binding energy so that $f(R)$ is a function of $R$ only,

$$f(R) \approx R + 2 + R^{-1}. \quad (5.2)$$

All mass and coupling variables refer to dark-sector quantities and are not used in this paper to refer to their visible-sector counterparts. However, since the analysis in this section closely follows that of SM hydrogen, dark hydrogen can have the same atomic properties of SM hydrogen by setting the model parameters appropriately. In fact, we can further diminish the distinction between the generic dark atomic model and SM hydrogen by expressing all dimensionful quantities in terms of the atomic energy and length scales

$$E_0 = \alpha^2 \mu_H, \quad a_0 = (\alpha \mu_H)^{-1}. \quad (5.3)$$

In doing so, we may adapt numerical results from the visible sector to the dark sector by adjusting these atomic scales. For the remainder of this section, we express quantities in
terms of atomic units for full generality.

5.3.1 Quantum formalism

We are interested in the interaction between two \( n = 1 \) ground-state dark hydrogen atoms, where \( n \) is the principal quantum number. The formalism is adapted from SM hydrogen [151, 239], and we outline the procedure here. An essential tool for simplifying the computation of a molecular wave function is the Born-Oppenheimer (BO) approximation, in which we write the total wave function as

\[
\Psi = \sum_{\gamma} \psi_{\gamma}(x) \phi_{\gamma}(x, y),
\]

where \( x \) is the relative separation (in atomic units) of the dark protons and \( y \) is a collective coordinate (in atomic units) for the dark electrons—we discard the center-of-mass motion. The subscript \( \gamma \) is a shorthand notation for the quantum numbers that define a set of basis states. Although, in principle, \( \gamma \) runs over all possible states, we truncate it to include only \( n = 1 \) atomic ground states. In writing Eq. (5.4), we treat \( \psi_{\gamma}(x) \) as the nuclear wave function for the dark proton, and it has no explicit dependence on the motion of the dark electrons. On the other hand, the electronic wave function \( \phi_{\gamma}(x, y) \) is dependent on the relative position of the dark protons. Accordingly, the Hamiltonian of the system separates into a part that describes only the relative motion of the protons and a part that encompasses the motion of the electrons and all Coulomb interactions. To solve the Schrödinger equation for \( \Psi \), the dark protons are initially held fixed, leaving just the electronic part of the Schrödinger equation. By repeatedly solving for \( \phi_{\gamma} \) under various nuclear configurations, the electronic BO eigenvalues \( \epsilon_{\gamma}(x) \) form a potential energy surface, which depends on the distance between the dark protons. These eigenvalues receive higher-order corrections from vibrational and rotational nuclear motion and from relativistic electronic motion. The validity of the BO
approximation relies on the dark proton being sufficiently heavier than the dark electron; the error of the approximation is $\sim R^{-3/2}$ [219] and so should be near and below the percent level for $R \gtrsim 20$.

We use existing calculations [152, 207, 229] of the BO eigenvalues for SM, ground-state molecular hydrogen. The states are labeled by the total electronic spin $S$. The $S = 0$ spin-singlet state $X^1\Sigma_g^+$ has a BO eigenvalue $\epsilon_0(x)$ and positive electronic parity; the $S = 1$ spin-triplet state $b^3\Sigma_u^+$ has a BO eigenvalue $\epsilon_1(x)$ and negative electronic parity. The explicit forms of $\epsilon_0(x)$ and $\epsilon_1(x)$ that we use are found in Ref. [76], with the exception of $\epsilon_0(x)$ in the range $0.3 < x < 12$, for which we interpolate tabulated results in Ref. [229]. We identify these states with our electronic states $\phi_\gamma(x,y)$, modulo rotated configurations of the dark protons [240].

With the BO eigenvalues at hand, we reincorporate the kinetic energy of the dark protons. Solving the full Schrödinger equation reduces to solving a one-dimensional (1D) Schrödinger equation for the relative nuclear motion in the potential $\epsilon_\gamma(x)$. In ket notation, we write the basis for the total wave function as $|SM_SIM_I\rangle$, where $S$ and $I$ are the total electronic and nuclear spins, respectively, and $M_S$ and $M_I$ are their associated $z$-axis projections. The label $\gamma$ runs over 16 states (or scattering channels) for two ground-state dark hydrogen atoms. Since the potential depends only on the nuclear separation $x$, we may expand the nuclear wave function in terms of partial waves,

$$\psi_\gamma(x) = \sum_{l,m} x^{-1} [F_\gamma(x)]_l Y_{lm}(\theta, \phi) , \quad (5.5)$$

where $[F_\gamma(x)]_l$ is the partial wave radial amplitude. The Schrödinger equation we must solve is

$$\left\{ \frac{d^2}{dx^2} - \frac{l(l+1)}{x^2} + f(R) [E - \epsilon_\gamma(x)] \right\} [F_\gamma(x)]_l = 0 \quad (5.6)$$
for each channel $\gamma$ with angular momentum $l$ and energy $E$ (in atomic units). Employing a more succinct notation, we discard the label $\gamma$ and express the Schrödinger equation in vector/matrix notation. We group the amplitudes into a single vector $F_l$ of length 16, whose row entries correspond to the various channels $\gamma \leftrightarrow |SM_S IM_I\rangle$. The potential then becomes a diagonal $16 \times 16$ matrix $V(x)$, whose entries are $\epsilon_{0,1}(x)$ for corresponding channels with $S = 0, 1$. The Schrödinger equation becomes

$$\left\{ \frac{d^2}{dx^2} - \frac{l(l+1)}{x^2} + f(R) \left[ E - V(x) - W \right] \right\} F_l = 0 ,$$

where we have included an additional constant $16 \times 16$ potential matrix $W$ in anticipation of the next section, but here we set $W = 0$.

### 5.3.2 Hyperfine interaction

We now incorporate a hyperfine interaction of the form

$$\hat{H}_{hf} = E_{hf} \left( \mathbf{I}_A \cdot \mathbf{S}_A + \mathbf{I}_B \cdot \mathbf{S}_B \right) ,$$

between the nuclear and electronic spins of atoms $A$ and $B$ involved in the scattering process. The interaction creates an energy splitting

$$\frac{E_{hf}}{E_0} = \frac{2}{3} g_e g_p \alpha^2 \frac{1}{f(R)} ,$$

of the $n = 1$ ground state into a hyperfine ground state and a hyperfine excited state. We set the Landé-$g$ factors of the dark electron and dark proton to be $g_e = 2$ and $g_p = 2$, respectively. The hyperfine excited state is unstable with a decay width

$$\Gamma = \frac{1}{3} \frac{\alpha E_{hf}^3}{m_e^2} .$$
The basis $|SM_SIM_I⟩$ is not ideal for this interaction, so we perform a change of basis [239] to $|F AMAF MB⟩$, where $\hat{F}_A = \hat{I}_A + \hat{S}_A$ is the total angular momentum of atom $A$ with spin projection $M_A$ (and similarly with atom $B$). Table 5.1 lists the quantum numbers for each channel, and Fig. 5.2 shows a schematic energy-level diagram for the ground state dark hydrogen atom with a hyperfine splitting. The hyperfine potential

$$[W]_{F_AM_AM_BF_AM_B}^{F'_AM'_AF'_BM'_B} = \delta_{F'_A F_A} \delta_{F'_B F_B} \delta_{M'_A M_A} \delta_{M'_B M_B} \frac{E_{hf}}{2E_0}[F_A(F_A+1)+F_B(F_B+1)-3]$$ (5.11)

is a diagonal matrix in this basis, while the change of basis induces off-diagonal elements in $V(x)$. As a result, the Schrödinger equation (5.7) becomes a system of 16 coupled differential equations. However, the selection rule $\Delta M = 0$, where $M = M_A + M_B$, allows $V(x)$ to be written as a block-diagonal matrix with four\(^1\) submatrices, whose form is given explicitly in Ref. [239]. The horizontal lines in Table 5.1 indicate which sets of channels correspond to different submatrices. Thus, we may solve Eq. (5.7) by solving multiple systems of fewer coupled equations, which is more computationally efficient.

### 5.3.3 Scattering

The cross section (in atomic units) from the state $|j⟩ ≡ |F AMAF MB⟩$ to $|i⟩ ≡ |F'_AM'_AF'_BM'_B⟩$ is [238]:

$$\sigma(j \rightarrow i) = \frac{\pi}{2k_j^2} \sum_l (2l + 1) \left| (T_l)_{ij} + (-1)^l (\tilde{T}_l)_{ij} \right|^2,$$ (5.12)

where $k_j$ is the wave number (in atomic units), $S_l = 1 + T_l$ is the $S$ matrix, and $\tilde{i}$ denotes swapping the labels on (or the quantum numbers of) dark atoms $A$ and $B$. The wave number depends on which hyperfine energy level the dark atoms occupy; at a given energy $\mathcal{E}$, a dark

\(^1\)Channels 15 and 16 have different values of $M$, but they are grouped in a single submatrix to maintain consistency with Ref. [239].
atom in the excited state will have a smaller velocity than one in the ground state. The wave number $k_j$ relates the relative velocity $v_j$ to the asymptotic energy of the channel $j$ via

$$k_j = \frac{1}{2} f(R) \frac{v_j}{\alpha} = \sqrt{f(R) \left[ E - (\mathcal{W})_{jj} \right]}.$$  

(5.13)

If there is sufficient energy such that $k_j^2 > 0$, then channel $j$ is open and accessible; otherwise, it is closed.

Instead of using the cross section for specific channels, a more useful quantity is the spin-averaged cross section for a particular type of process. Based on the change in the total angular momentum $\hat{F} = \hat{F}_A + \hat{F}_B$, we group spin-averaged cross sections into seven categories: $\Delta F = \pm 2$, $\Delta F = \pm 1$, and three separate groups of $\Delta F = 0$. Table 5.2 lists these cross
sections and their corresponding scattering channels. Note that the channels in each group have the same initial and final velocities. The spin-averaged cross sections \( \sigma_{+} \) and \( \sigma_{++} \) represent one and two dark atoms, respectively, in the hyperfine ground state upscattering to the excited state. Similarly, \( \sigma_{-} \) and \( \sigma_{--} \) represent one and two dark atoms downscattering from the excited state to the ground state. The remaining spin-averaged cross sections involve dark atoms that either do not make an energy transition or simply swap energy states. Atoms remain in the ground state for \( \sigma_{gg} \) and in the excited state for \( \sigma_{ee} \). For \( \sigma_{ge} \), one dark atom is in the ground state while the other is in the excited state. There are some processes not listed in Table 5.2; they are zero due to selection rules or because total angular momentum is not conserved.

The total spin-averaged cross section \( \sigma_{\text{tot}} \) is obtained by summing over all individual cross sections in Eq. (5.12) and dividing by 16, if all channels are open. There are also variations of the total cross section that incorporate nonuniform weights for the angular integration. The

<table>
<thead>
<tr>
<th>Cross section</th>
<th>Transition ((\Delta F))</th>
<th>Processes</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_{++} )</td>
<td>2</td>
<td>( aa \rightarrow cc, bd, db )</td>
</tr>
<tr>
<td>( \sigma_{--} )</td>
<td>-2</td>
<td>( cc, bd, db \rightarrow aa )</td>
</tr>
<tr>
<td>( \sigma_{+} )</td>
<td>1</td>
<td>( ac, ca \rightarrow bd, db ) ( ad, da \rightarrow cd, dc ) ( ab, ba \rightarrow bc, cb )</td>
</tr>
<tr>
<td>( \sigma_{-} )</td>
<td>-1</td>
<td>( bd, db \rightarrow ac, ca ) ( cd, dc \rightarrow ad, da ) ( bc, cb \rightarrow ab, ba )</td>
</tr>
<tr>
<td>( \sigma_{gg} )</td>
<td>0</td>
<td>( aa \rightarrow aa )</td>
</tr>
<tr>
<td>( \sigma_{ee} )</td>
<td>0</td>
<td>( cc, bd, db \rightarrow cc, bd, db ) ( cd, dc \rightarrow cd, dc ) ( bc, cb \rightarrow bc, cb ) ( dd \rightarrow dd ) ( bb \rightarrow bb )</td>
</tr>
<tr>
<td>( \sigma_{ge} )</td>
<td>0</td>
<td>( ac, ca \rightarrow ac, ca ) ( ad, da \rightarrow ad, da ) ( ab, ba \rightarrow ab, ba )</td>
</tr>
</tbody>
</table>

Table 5.2: Definition of various spin-averaged cross sections.
momentum-transfer cross section $\sigma_{\text{tot},T}$ weights the cross section by the fractional longitudinal momentum transferred in the scattering process, thereby suppressing forward scattering; whereas the viscosity cross section $\sigma_{\text{tot},V}$ weights the cross section by the fractional transverse energy transfer, thereby suppressing both forward and backward scattering.

$$\sigma_{\text{tot},T} = \int \frac{d\sigma_{\text{tot}}}{d\Omega} (1 - \cos \theta) d\Omega$$

$$\sigma_{\text{tot},V} = \int \frac{d\sigma_{\text{tot}}}{d\Omega} \sin^2 \theta d\Omega .$$

We may apply these variations to the individual cross sections in Eq. (5.12), under the assumption that the scattering particles are identical. For ease of notation, we define $(T_{\text{eff}})_l^{ij} \equiv (T_l)_{ij} + (-1)^l(T_l)_{ij}$. To obtain $\sigma_V(j \rightarrow i)$, we replace $(2l+1) \rightarrow (l+1)(l+2)/(2l+3)$ and $T_{\text{eff}}^l \rightarrow T_{l+2}^{\text{eff}} - T_l^{\text{eff}}$. The normalization ensures that $\sigma_{\text{tot},V}$ has the proper limits for pure $s$-wave scattering, for which the differential cross section is isotropic. For $\sigma_{\text{tot},T}$, however, the part of the integrand with the additional $\cos \theta$ yields zero, so $\sigma_{\text{tot},T} = \sigma_{\text{tot}}$ [153]. The suppression of forward scattering is compensated by the identical backward scattering, resulting in no change from $\sigma_{\text{tot}}$.

5.4 Applications of Atomic Dark Matter

We now examine the atomic dark matter model of Sec. 5.3 in a cosmological and astrophysical context. First, we determine the necessary conditions for dark matter to be in the form of $n = 1$ ground state dark hydrogen. Then, with calculations of scattering cross sections from the previous section at hand, we examine the atomic dark matter model as an SIDM candidate. Instead of parametrizing the model in terms of those listed in Eq. (5.1), we opt for the quantities $R$, $\alpha$, and $m_H$; furthermore, since Eq. (5.9) relates $E_{\text{hf}}/E_0$ to $\alpha$ and $R$, we may use $E_{\text{hf}}/E_0$ in place of $R$ as a free parameter. We consider two specific cases for the hyperfine
splitting, $E_{\text{hf}} = 10^{-4}E_0$ and $E_{\text{hf}} = 10^{-5}E_0$, and determine what regions of the remaining parameter space can address the small-scale structure puzzles. These values are chosen such that the cooling from inelastic scatterings may lead to interesting observable effects in dwarf-scale to cluster-scale halos. Larger values of $E_{\text{hf}} \gtrsim 10^{-3}E_0$ require correspondingly larger relative particle velocities in order to upscatter to the hyperfine excited state; for $E_{\text{hf}} = 10^{-3}$ this means that cooling effects would only occur in halos with $v_{\text{rms}} > 260$ km/s.

If the hyperfine splitting $E_{\text{hf}}$ is decreased below $10^{-6}E_0$, the energy losses from upscatterings become negligible and our collisional cross sections approach those obtained in the elastic approximation. Splittings of $E_{\text{hf}} \sim 10^{-4}$ are of particular interest to us as they lead to large swaths of $m_\text{H} - \alpha$ parameter space in which collisional heating effects can solve the aforementioned small-scale structure puzzles.

In the following analysis, we work under the simplifying assumptions that the dark matter halos are completely neutral and are not affected by excitations to the $n = 2$ state. For the reasons that we discuss later on, the effects of excitations to the $n = 2$ (see Sec. 5.5.2) state as well as collisional ionizations (see Sec. 5.6.2) may become non-negligible at galaxy cluster scales for $E_{\text{hf}} = 10^{-5}E_0$, the smaller of the two hyperfine splittings considered here. Without a more detailed treatment of the dark atomic physics, it is possible that for hyperfine splittings $\sim 10^{-5}E_0$ these effects may change the cluster-scale halo structure in such a way as to become inconsistent with current observations. To be conservative, one should not interpret the results below for $E_{\text{hf}} = 10^{-5}E_0$ as predictions, but should instead use them as a comparison to the more straightforward $E_{\text{hf}} = 10^{-4}E_0$ case to see how the results are affected for different hyperfine splittings.
5.4.1 Cosmological considerations

Our goal is to uncover the cosmological phenomenology of neutral atomic dark matter in which the dominant inelastic scattering mode is through dark hyperfine transitions. We must first map out the region of parameter space that is allowed for this model.

Under what circumstances does dark matter today consist of dark hydrogen bound states, with no dark ions remaining? In analogy with SM hydrogen, there is a period of dark recombination in the early universe, and for large regions of parameter space, the majority of dark ions do indeed form into a neutral bound atomic state. We assume the Universe has no overall dark charge, so perfect recombination would result in no remaining dark ions. Most of the recombination occurs in the range \( 0.007 \lesssim T_D/B_H \lesssim 0.01 \), where \( T_D \) is the temperature of the dark radiation bath [84]. At the end of dark recombination, the residual ionization fraction \( \chi_e \) is given by [84]

\[
\chi_e \sim 2 \times 10^{-16} \frac{\xi}{\alpha^6} \left( \frac{0.11}{\Omega_{DM} h^2} \right) \left( \frac{m_H}{\text{GeV}} \right) \left( \frac{B_H}{\text{keV}} \right),
\]

where \( \xi \equiv (T_{D,L}/T_{V,L}) \) is the ratio of the dark radiation temperature to the visible-sector CMB temperature in the present-day late universe. The number density of dark matter particles changes with \( m_H \), where larger masses correspond to lower number densities and hence a later recombination redshift and a larger ionization fraction. A higher binding energy, \( E_0/2 \), results in a larger ionization fraction because recombination is less efficient. However, the dependence on \( \alpha \) is much stronger: if \( \alpha \) is too small, the interaction between ions is simply not strong enough for them to attract one another and form bound states.

Constraints in the \( m_H-\alpha \) plane are plotted in Fig. 5.4 below for different values of \( \chi_e \). Postrecombination, we will require \( \chi_e \lesssim 0.01 \). With this simplification, we assume all the dark matter is in the form of dark hydrogen and do not consider processes such as dark ion-ion or ion-hydrogen scattering. Furthermore, we avoid dark particles with long-range forces, which
affect structure formation [142, 143]. Note that dark hydrogen typically recombines to the $n = 2$ state and not the $n = 1$ state analyzed in Sec. 5.3. However, the lifetime of the $n = 2$ state is very short ($\ll 1$ year [84]), so we expect all dark hydrogen to have settled into its $n = 1$ ground state by the time of structure formation.

A complication for our desired dark hydrogen model is that dark hydrogen can potentially form molecules and affect halo cooling. The formation of dark molecular hydrogen $H_2$ may occur through (1) neutral-neutral dark atom processes $H + H \rightarrow H_2 + \gamma$ or (2) processes requiring a dark electron or proton catalyst ($e^- + H \rightarrow H^- + \gamma$, $H^- + H \rightarrow H_2 + e^-$), $(H^+ + H \rightarrow H_2^+ + \gamma, H_2^+ + H \rightarrow H_2 + H^+)$, where $\gamma$ is the dark photon. The first type of process is very suppressed due to the fact that it must occur through a quadrupole transition [124]. The second type of process requires a free ionized population and thus may occur before or during recombination; however, at these times there are enough dark Lyman-Werner photons to photodissociate the dark $H_2$. Hence, although a very small amount of dark molecular formation is possible, we do not consider it further. However, in the visible sector, we know that even small traces of SM molecular hydrogen dramatically affect the cooling of gas in the first dark matter minihalos: for high-redshift ($z = 23$) minihalos of masses $5 \times 10^5$–$10^6 M_\odot$, a SM molecular hydrogen fraction of $\lesssim 10^{-3} \rightarrow$ can cool the innermost regions and precipitate gravitational collapse [119]. Analogously, even a small amount of dark $H_2$ present may allow for much more efficient cooling in halos from excitations of the rotational and vibrational modes. A comprehensive treatment of dark $H_2$ formation would be necessary to investigate this effect.

Finally, we consider the abundance and temperature of the dark sector, as both can dramatically affect the expansion history of the Universe and clustering of dark matter. We assume there is a dark matter–antimatter asymmetry [143], and the full annihilation of dark antiparticles in the early universe yields the correct relic abundance of dark particles. Since we assume the dark sector is overall charge neutral (i.e., an equal number of dark protons
and electrons) by the time of recombination, the abundance is controlled by the heavier dark protons. Thus, $\alpha$ is bounded from below by requiring that the dark antiprotons annihilate efficiently. For $p + \bar{p} \rightarrow 2\gamma$ annihilation in the nonrelativistic regime, the velocity-averaged annihilation cross section is given by

$$\langle \sigma v \rangle = \frac{\pi \alpha^2}{m_p^2} = 3.66 \times 10^{-25} \text{cm}^3/\text{s} \left( \frac{\alpha}{0.01} \right)^2 \left( \frac{m_p}{100 \text{ GeV}} \right)^{-2}.$$  

(5.17)

For Dirac dark matter, the thermal relic annihilation cross section is $\langle \sigma v \rangle \approx 4.4 \times 10^{-26} \text{cm}^3/\text{s}$ [216]. Thus, for efficient annihilation of antiprotons, we have the following lower bound on $\alpha$:

$$\alpha > 0.0035 \left( \frac{m_p}{100 \text{ GeV}} \right).$$  

(5.18)

We also need to include the annihilation to the hidden electrons and positrons, which would weaken the lower limit. However, this is not required because these limits are much less constraining than our lower limits on $\alpha$ derived by imposing $\chi_e < 0.01$ on the late-time ionization fraction. We will also find that in most of the parameter space where the SIDM phenomenology could be relevant for the small-scale puzzles, the $\alpha$ value will be larger than the lower limit in Eq. (5.18). Given these constraints, the dark matter abundance must be set by a dark matter–antimatter asymmetry.

Dark photons contribute to the effective number of relativistic degrees of freedom $N_{\text{eff}}$, measured at the time of last scattering and during big bang nucleosynthesis (BBN). Since the dark sector is secluded from the visible sector, the temperatures in each are naturally allowed to differ. If $\xi \lesssim 0.65$ (at $\sim 1\sigma$), then we avoid BBN bounds on $N_{\text{eff}}$ for the range of $m_H, \alpha$, and $E_{hf}$ considered here [84]. One may attempt to motivate a natural value for $\xi$ by allowing the visible and dark sectors to interact, for instance, via a kinetic mixing term $\frac{1}{2} \epsilon_k F^\mu_\nu F^{\mu\nu}$, which would give the SM electrons a charge of $\epsilon_k e$ under the hidden $U(1)$. Then, the two sectors could come into thermal equilibrium through the process $e_{\text{SM}} + \gamma_{\text{SM}} \leftrightarrow e_{\text{SM}} + \gamma$ at
some temperature $T$ if the condition $T^2/M_{Pl} = \alpha_{SM} \alpha^2 T$ is met in the thermal bath [131]. However, direct detection constraints from the LUX experiment place strong constraints on the mixing parameter $\epsilon_k \lesssim 2 \times 10^{-10}$ for the preferred regions of parameter space determined below [93]. For the range of $\epsilon_k$ small enough to satisfy direct detection constraints, the condition for achieving thermal equilibrium is not reached prior to the freeze-out of SM $e^+e^-$ annihilation. The resulting low SM electron density causes the equilibration process between the dark and visible sectors to be inefficient, so the sectors do not achieve thermal equilibrium; thus, we do not have a well-motivated value to assume for the ratio $\xi$ of their present-day temperatures. If initially set by inflationary reheating, the visible and dark sectors could have different temperatures depending on the inflaton couplings to the respective sectors. We use the value $\xi = 0.6$ in the following work. By using a value close to the upper limit on $\xi$, the contours we show later on in Sec. 5.4.4 may be interpreted as approximate upper limits on the minimum halo mass.

In summary, our requirement that dark matter consists exclusively of dark hydrogen bound states means that we only consider $\chi_e < 0.01$. In order not to exceed the tight BBN constraints on the light degrees of freedom in the early universe, we require that the temperature of the dark sector be no greater than approximately 0.65 times the temperature of the visible sector. Both of these constraints are easily met with a secluded dark U(1) sector.

### 5.4.2 Cross sections, lifetimes, and structure formation

One of the difficulties in determining the effects of SIDM models on structure formation is that it is unclear how the microphysical scattering can be represented by macroscopic simulation particles in $N$-body experiments or in more general macroscopic descriptions of halos. In this section, we advocate the use of the viscosity cross section of $n = 1$ ground-state dark hydrogen atoms to model the microphysics of atomic dark matter in the evolution
of dark matter halos. We first show that nearly all dark atoms should be in their hyperfine ground state, and then motivate our choice of the viscosity cross section to model the scattering-induced energy flow in a dark matter halo.

Dark hydrogen may be in either its hyperfine ground or excited state. For the hyperfine splittings we consider here, the time scale in Eq. (5.10) for decays from the hyperfine excited state to the ground state is always much less than the time scale for excitation via upscattering. Specifically, for parameter space of interest in Sec. 5.4.3, the excited state lifetime is always $\Gamma^{-1} \ll 1 \text{ Gyr}$; a particle that upscatters into the excited state almost always emits a dark photon and returns to the ground state before scattering with another particle. Hence, we focus on cross sections in which both particles are initially in the ground state.

A dark matter halo can be altered if interactions occur that transfer momentum: elastic scatterings between atoms allow for heat to flow into the cold halo interior and increase the velocity dispersion of the inner halo relative to the CDM case. The momentum-transfer cross section $\sigma_{\text{tot},T}$ (introduced in Sec. 5.3.3) is commonly used in the literature to describe astrophysical constraints on the SIDM cross section, since it suppresses the far-forward scattering case of $\theta = 0$, which is equivalent to no interaction occurring. However, $\sigma_{\text{tot},T}$ preferentially weights backward scattering, which also does not change the velocity distribution away from the CDM case, despite the fact that a large amount of momentum is transferred. An alternative is instead to use the viscosity cross section $\sigma_{\text{tot},V}$ to favor scattering in the transverse direction. While this choice may not be the fully correct quantity to use for comparison with SIDM constraints from simulations, we argue that it better captures the relevant SIDM physics at lowest order compared to the momentum-transfer cross section. Additionally, the viscosity cross section is more physically well motivated to use with identical particles, for which forward and backward scattering are identical, and we agree with Ref. [76] that the use of $\sigma_{\text{tot},T} = \sigma_{\text{tot}}$ for identical particles [114, 153, 137, 84] is unwarranted. Although the overall structure of the halo should be insensitive to the quantum mechanical nature of
individual dark matter particles, it is reassuring that the viscosity cross section provides a consistent description for SIDM limits on both the macroscopic and the microscopic scales. Thus, we consider $\sigma_{gg}$, $\sigma_{++}$, and their viscosity counterparts. Recall from Sec. 5.3.3 that a single excitation from the scattering of two ground-state atoms does not occur because of selection rules (for $aa \rightarrow ab, ad$) or because total angular momentum is not conserved at a fixed orbital angular momentum (for $aa \rightarrow ac$). Instead of the total $\sigma_{\text{tot},V}$, it is

$$\sigma_V \equiv \sigma_{gg,V} + \sigma_{++,V}$$ (5.19)

that we use to compare with target SIDM cross sections in Sec. 5.4.3.

The use of the viscosity cross section is complicated by introducing a hyperfine splitting, and $\sigma_{\text{tot},V}$ may not be an appropriate quantity to use with SIDM either. Keep in mind that the constraints from SIDM simulations assume elastic scattering—kinetic energy is conserved. There may be regions of parameter space where $\sigma_{++,V}$ is a substantial contribution to $\sigma_V$ at energies near the hyperfine threshold, resulting in a significant loss in kinetic energy. In this case, applying SIDM constraints using $\sigma_V$ is not necessarily a valid comparison, and we discuss this issue further in Sec. 5.5.

Figure 5.3 shows the ratio of $\sigma_V$ (with $E_{hf} = 10^{-4}E_0$ and $E_{hf} = 10^{-5}E_0$) to $\sigma_{\text{tot},V}^{(\text{elastic})}$ (with $E_{hf} = 0$) and demonstrates how the inclusion of the hyperfine interaction affects the cross section, as compared to Ref. [76]. Since we assume all dark atoms are in their hyperfine ground state, the spin-averaging factor for $\sigma_V$ is unity; meanwhile, $\sigma_{\text{tot},V}^{(\text{elastic})}$ contains more contributions from the other scattering channels, but it has a spin-averaging factor of $1/16$. Thus, in our comparison of $\sigma_V$ and $\sigma_{\text{tot},V}^{(\text{elastic})}$, it is not generically true that $\sigma_V$ is the strictly smaller quantity. At higher energies, the two cross sections are comparable, so using either results in very similar regions of acceptable SIDM parameter space. At lower energies, there is a resonant effect from the scattering of low-velocity particles; these particles can
exchange multiple dark photons and form quasibound states, resulting in an enhancement of the scattering cross section. In this regime, $\sigma_V$ can be larger or smaller than the total elastic viscosity cross section by a factor of a few, and the resulting shapes of the acceptable SIDM regions will differ. Although the SIDM constraints might look similar between our model and its counterpart in the elastic approximation, the crucial difference comes from the potentially significant energy losses that result from hyperfine transitions. We discuss this issue further in Sec. 5.5.

Figure 5.3: Parameter space scan comparing the viscosity cross section (5.19), which involves only hyperfine ground-state atoms, to the total viscosity cross section in the elastic approximation, used in Ref. [76]. The left plot shows $\sigma_V$ with $E_{hf} = 10^{-5} E_0$, while the right plot shows $\sigma_V$ with $E_{hf} = 10^{-4} E_0$. The spin averaging factors for $\sigma_V$ and $\sigma_{tot,V}^{(elastic)}$ are different, so we do not necessarily expect $\sigma_V$ to be less than $\sigma_{tot,V}^{(elastic)}$.

5.4.3 SIDM halo profiles

We now show how we use the viscosity cross section to identify interesting regions of dark atom parameter space. Self-interacting dark matter models are motivated by their ability to produce cored density profiles in dark matter halos below $\sim 10^{11} M_\odot$ (corresponding to $v_{rms} \lesssim 100$ km/s), thereby relieving tensions between the predictions for small-scale
structure from collisionless $N$-body simulations and the inferred halo profiles of observed galaxies [201, 100]. Simulations of galaxies in this mass range find that hard-sphere scattering cross sections of $\sigma/m \sim 0.5$–5 cm$^2$/g are capable of reproducing the cored density profiles of low surface-brightness galaxies (and perhaps dwarf spheroidals) [201, 100]. We thus require that the atomic dark matter models in our allowed region of parameter space result in cross sections $\sigma_Y/m_H$ within this range for halos characterized by velocities of $v_{\text{rms}} = 30$–$100$ km/s, which roughly corresponds to halo masses of $5 \times 10^9$–$10^{11}$ $M_\odot$.

One may also calculate a target range for the velocity-dependent viscosity cross section at higher velocities of $v_{\text{rms}} \sim 1000$ km/s using the core sizes of cluster-mass halos. The scattering cross section required to produce a core of radius $r_c$ may be approximated by assuming that the size of the cluster core is equal to the maximum radius at which the average dark matter particle will scatter at least once during the lifetime of the halo,

$$t_{\text{scatter}}(r_c) = t_{\text{age}} = \left( \frac{16 \sigma_Y(v)}{3 \pi m_H \rho(r_c) v_{\text{rms}}(r_c)} \right)^{-1},$$

(5.20)

where $\rho(r)$ is the dark matter density at radius $r$. The 1D velocity dispersion in a halo $v_{\text{rms}}$ and the average relative collisional velocity $v$ between particles are related by $v \approx \sqrt{2}v_{\text{rms}}$. Reference [147] sets the cluster age $t_{\text{age}} = 5$ Gyr and uses the halo profiles reported for the set of relaxed galaxy clusters in Ref. [176] to derive the cluster-scale cross sections; they find that the observed core sizes may be reproduced if the cross section is $\sim 0.1$ cm$^2$/g at cluster velocities $v_{\text{rms}} \sim 1000$ km/s. We use the inferred cluster-scale cross sections, velocities, and uncertainties from Ref. [147] and require that our atomic dark matter models must have a viscosity cross section within this range at velocities $v_{\text{rms}} \sim 1000$ km/s.

While target cross sections for velocity-dependent SIDM may be obtained using the core sizes in dark matter halos across a wide range of characteristic velocities, upper limits on the scattering cross section at high velocities may be derived from observations of cluster-
scale systems. Constraints on SIDM cross sections at larger scales may be derived from measurements of merging galaxy clusters \cite{86, 122}, displacements of galaxies from cluster centers long after merging \cite{150}, and ellipticities of dark matter halos \cite{59, 199}. A constraint on \( \sigma V / m_H \lesssim 1 \text{ cm}^2/\text{g} \) at the cluster scale may be derived from the observed ellipticities of cluster halo profiles inferred through gravitational lensing \cite{199, 190}: if the SIDM cross section is too high, then repeated scatterings of particles in cluster halos will transform the halo shape from a triaxial ellipsoid into a sphere. We use this value as an upper limit on \( \sigma V / m_H \) at cluster scales in the analysis below.

Another constraint on self-interaction cross sections may be derived from the lack of observed deceleration of the dark matter components due to a drag force in systems of merging clusters, although for a completely different type of cross section than can be produced in atomic dark matter models. Reference \cite{122} uses observations of multiple merging cluster systems to place an upper limit of \( \sigma / m < 0.47 \text{ cm}^2/\text{g} \) at relative velocities \( v \sim 900 \text{ km/s} \). We note that this limit is not a true constraint on our parameter space, as it assumes an SIDM model where anisotropic and frequent scattering events with low momentum exchange give rise to a drag force with a \( v^2 \) dependence within merging clusters.

At the high end of the galactic mass scale, past works \cite{84, 104} have also used the halo ellipticity of the elliptical galaxy NGC 720 as inferred through x ray observations \cite{59} to set an upper bound on the scattering cross section. However, we do not use NGC 720 to set an upper limit at its velocity scale \( v_{\text{rms}} \sim 250 \text{ km/s} \) for the following reasons \cite{190}: (1) particles in such halos may on average undergo multiple scatterings over the halo lifetime while still preserving ellipticities, and (2) NGC 720 is the only object at this mass scale which has so far been measured to have an elliptical halo, and the scatter in simulated halo shapes may still allow for this halo to be accommodated as an outlier even if \( \sigma V / m_H \gtrsim 0.1 \text{ cm}^2/\text{g} \).

In Fig. 5.4 we show the regions of \( m_H - \alpha \) parameter space which satisfy the target cross section ranges for dwarf and low surface brightness galaxies, as well as galaxy clusters. The
Figure 5.4: Viscosity scattering cross sections are calculated at the velocities of interest at each point in the $m_H$–$\alpha$ plane and then used to determine which areas either satisfy target cross sections or are in tension with observations. The hyperfine splitting is fixed to $E_{hf} = 10^{-5} E_0$ in the left panel and $E_{hf} = 10^{-4} E_0$ in the right panel. The vertical hatched grey area is disfavored by measurements of cluster halo ellipticities and corresponds to the region where $\sigma_V/m_H > 1 \text{ cm}^2/\text{g}$ for velocities $v_{\text{rms}} = 1000 \text{ km/s}$. We show contours of the ionization fraction $\chi_e = 10^{-1}$, $10^{-2}$, and $10^{-3}$; we consider $\chi_e \lesssim 10^{-2}$ sufficiently low to ignore dark ions, which excludes a large portion of the displayed parameter space for $E_{hf} = 10^{-5} E_0$. Points within the cross-hatched green region satisfy $0.5 \text{ cm}^2/\text{g} < \sigma_V/m_H < 5 \text{ cm}^2/\text{g}$ for velocities $v_{\text{rms}} = 30$–$100 \text{ km/s}$, which approximates the condition for cores to form in lower-mass halos. Points within the solid orange region provide the best-fit viscosity cross sections for cores to form in relaxed cluster halos. The dashed lines show contours of constant minimum halo mass for values of $E_{hf}, m_H$, and $\alpha$ in our model, assuming $\xi = 0.6$. Lower, allowed values of $\xi$ lead to smaller minimum halo masses.

Cross-hatched green areas denote the region where $0.5 \text{ cm}^2/\text{g} < \sigma_V/m_H < 5 \text{ cm}^2/\text{g}$ in halos with $v_{\text{rms}} = 30$–$100 \text{ km/s}$, while the solid orange shaded areas denote the 68% confidence limit for the best-fit region where the viscosity cross sections can reproduce cluster core sizes. The resonant structure of the scattering cross section is evident in the green curves, since the associated velocities are quite low; the orange curves, on the other hand, do not probe the resonant regime. The vertical hatched grey region is excluded by the cluster halo shape constraint of $\sigma_V/m_H < 1 \text{ cm}^2/\text{g}$ in halos with $v_{\text{rms}} > 1000 \text{ km/s}$. After imposing the constraint on $\sigma_V/m_H$ from cluster halo shapes, the regions of allowed parameter space in
Fig. 5.4 extend to lower masses ($\lesssim 100$ GeV) than those given in Refs. [84] and [76]. The overlap between the cross-hatched green region and the solid orange region encloses the values of $m_H$ and $\alpha$ which give the desired viscosity cross sections at both of the velocity scales of interest. For $E_{\text{hf}} = 10^{-5} E_0$ there is very little overlap region, and much of the parameter space shown is excluded by our requirement that the Universe is neutral: $\chi_e \lesssim 0.01$. For both cases, the calculation of the preferred SIDM regions in green and orange may not be reliable for $\chi_e > 0.01$.

### 5.4.4 Minimum halo masses

Prior to kinetic decoupling, the coupling between the dark radiation and the dark matter affects the growth of density perturbations and suppresses the small-scale matter power spectrum relative to predictions from CDM. The dark radiation-matter coupling leads to dark sector analogs of the phenomena of diffusion (or Silk) damping and baryon acoustic oscillations. The growth of density perturbations below the damping scale $r_D$ is damped, leading to a cutoff in the matter power spectrum at small scales. For a detailed review and explanation of these effects, see Refs. [42, 103, 84, 83, 85].

One may associate a minimum dark halo mass $M_{\text{min}}$ with the smallest perturbation mode below which the growth of structure is suppressed, given by [42]

$$M_{\text{min}} \simeq 0.1 \left( \frac{T_{\text{dec}}}{\text{MeV}} \right)^{-3} M_\odot .$$

(5.21)

The minimum halo mass is set by the decoupling temperature $T_{\text{dec}}$, which in turn is set by the physics of the atomic dark matter. Unlike the equivalent scenario involving SM hydrogen, the dark atoms do not necessarily become transparent to dark photons soon after recombination—for high enough values of $\alpha$, the contribution to the opacity from Rayleigh
scattering between photons and neutral atoms may keep the dark plasma opaque even if it is not ionized [84]. There are thus two expressions for $T_{\text{dec}}$, depending on whether Compton or Rayleigh scattering provides the dominant contribution to the dark matter opacity prior to decoupling. If Rayleigh scattering dominates, we use the following equation for $T_{\text{dec}}$ [84]:

$$T_{\text{dec}}^{\text{Rayleigh}} \simeq 7 \times 10^{-4} B_H \left[ \frac{1}{\alpha^{\frac{6}{3}} \xi^3} \left( \frac{B_H}{\text{keV}} \right) \left( \frac{m_H}{\text{GeV}} \right) \right]^\frac{1}{3},$$

(5.22)

where $\xi$ is the ratio of the dark radiation temperature to the visible sector CMB temperature in the present-day late universe. As mentioned in Sec. 5.4.1, we take this ratio to be $\xi = 0.6$ so that the minimum halo masses calculated are approximately upper limits.

If Compton scattering is the dominant source of opacity, we follow the method of Ref. [83] to solve for the scale factor at decoupling $a_{\text{dec}}$ and then convert this to the temperature $T_{\text{dec}}$. We approximate that the dark electrons and photons decouple when the expansion rate begins to exceed the Thomson scattering rate, i.e. when $H \simeq n_{\text{H}} \chi_e \sigma_{\text{Thomson}}$. This leads to the following equation for the scale factor at decoupling $a_{\text{dec}}$:

$$a_{\text{dec}}^3 + \Omega_R \Omega_m^2 a_{\text{dec}}^2 = \frac{1}{\Omega_m h^2} \left[ \epsilon_D \alpha \xi^3 \left( \frac{B_H}{\text{eV}} \right)^{-1} \left( \frac{m_H}{\text{GeV}} \right)^{-\frac{1}{3}} \right]^2.$$

(5.23)

The constant $\epsilon_D$ is obtained by fitting to the numerically calculated ionization fraction and thermal evolution of the dark sector and is approximately $\epsilon_D \sim 8 \times 10^{-3}$ for the range of $\alpha$ considered here [83]. We find that the Rayleigh scattering case dominates for the parameter space considered here, and thus the minimum halo masses shown in Fig. 5.4 are calculated using the decoupling temperatures given by Eq. (5.22).

We calculate the minimum halo masses in our region of $m_H$-$\alpha$ parameter space and show these as dashed contours in Fig. 5.4. In general, the minimum halo masses in the allowed region of parameter space lie below the current observational limits. The highest minimum halo mass that is not ruled out by the constraints from cluster shapes is $M_{\text{min}} \sim 10^{7.5} M_\odot$ for
hyperfine splittings of $E_{hf} = 10^{-5}E_0$, or $M_{\text{min}} \sim 10^5 M_\odot$ for $E_{hf} = 10^{-4}E_0$. If the parameter space is constrained by demanding that velocity-dependent elastic cross sections produce cores at both high and low halo masses, then this would lead to a prediction of a minimum halo mass of around $M_{\text{min}} \sim 10^7 M_\odot$ for fixed $E_{hf} = 10^{-5}$, or $M_{\text{min}} \sim 10^{3.5-5} M_\odot$ for fixed $E_{hf} = 10^{-4}$. Halos of these sizes may be observed in next-generation substructure lensing or galactic tidal stream surveys [123, 48].

The temperature ratio $\xi$ may take on different values in various inflationary reheating scenarios (as long as the BBN constraint of $\xi \lesssim 0.65$ is still satisfied). From Eqs. (5.21) and (5.22), $M_{\text{min}} \propto \xi^{9/5}$ if $T_{\text{dec}}$ is set by the dominance of Rayleigh scatterings: lower values of $\xi$ result in lower minimum halo masses. For a sufficiently large difference in potential values of $\xi$ (e.g. 0.1 versus the upper limit of 0.65), the minimum halo mass can vary by an order of magnitude. If one were able to observe the matter power spectrum cutoff associated with $M_{\text{min}}$ in an atomic dark matter scenario, this measurement could be translated into a lower bound on the dark to visible sector temperature ratio in various regions of the $(E_{hf}, m_H, \alpha)$ parameter space.

In summary, we find regions of dark atom parameter space that satisfy observed constraints on the elastic SIDM cross section on scales from dwarfs to clusters. Because dark acoustic oscillations lead to a dark-sector temperature-dependent truncation of the halo mass function on potentially observable scales, it may be possible to fully constrain the atomic dark matter model with halo core sizes and the halo power spectrum.

5.5 Consequences of Inelastic Scattering

The constraints on atomic dark matter so far come from assuming elastic scattering. One of the most interesting aspects of dark atoms is that they have excited states, which admit
inelastic processes. In this section, we consider the magnitude of inelastic hyperfine scattering in our model. The net effect of two particles upscattering and then decaying back to the ground state is an overall loss in kinetic energy; this provides a mechanism for cooling in dark matter halos which may potentially counterbalance or dominate over the heating mechanism provided by elastic scatterings. Depending on the values of $E_{\text{hf}}$, $m_H$, and $\alpha$, either one or both of these effects may have a large influence on the evolution of a halo. In the following discussion, we investigate whether inelastic cooling effects may significantly impact the halo structure in any regions of the parameter space considered in this work.

5.5.1 Comparison of viscosity and upscattering cross sections

In Fig. 5.5 we plot examples of upscattering and viscosity cross sections per unit mass for values of $m_H$ and $\alpha$ lying within the allowed regions of parameter space in Fig. 5.4 which may produce cores in cluster, LSB, and dwarf halos consistent with observations. We plot these cross sections for multiple values of $E_{\text{hf}}$, $m_H$, and $\alpha$ in order to show the range of velocity-dependent behavior allowed in our model which may resolve small-scale structure issues. In the following discussion, we explore the phenomenology that may arise in different regions of parameter space due to the different velocity-dependent behaviors of the viscosity and upscattering cross sections.

We now compare the velocity dependence of the viscosity and upscattering cross sections at cluster-scale velocities $v_{\text{rms}} \sim 1000$ km/s. At velocities of order $O(100)$ km/s and lower, the viscosity cross section is generally comparable to or larger than the upscattering cross section; at higher velocity scales above 1000 km/s, the viscosity cross section decreases at a steeper rate such that it falls well below the upscattering cross section. We choose to plot cross sections and velocities in units of astrophysical observables in Fig. 5.5, though the same quantities can be plotted in terms of the geometric cross section $\pi a_0^2$ and atomic energy $E_0$. 

112
Figure 5.5: Viscosity (red curves) versus upscattering (blue curves) cross sections per unit mass as a function of halo velocity. Because of the short decay time scale for the hyperfine excited state, we assume that both initial particles are in the ground state. The value of the hyperfine splitting is fixed to $E_{\text{hf}} = 10^{-5} E_0$ and $E_{\text{hf}} = 10^{-4} E_0$ in the left and right panels, respectively. Each line is drawn randomly from values of $m_H$ and $\alpha$ lying within the overlapping regions of Fig. 5.4 corresponding to parameters which lead to cores consistent with observations in both cluster-scale halos as well as dwarf- to LSB-scale halos.

The shape of the viscosity cross section changes noticeably at two velocities or energies; these changes are more easily seen for the case where $E_{\text{hf}} = 10^{-4} E_0$. At very low energies, the viscosity cross section is $s$ wave and constant. There is a break in the viscosity cross section near $E = E_0 R^{-3/2}$, where the de Broglie wavelength of the atoms can probe the structure of the lowest-order Van der Waals interaction potential ($\sim 1/x^6$), and higher partial waves begin to contribute. For $R \sim 100$, this break occurs near $v \sim 100 \text{ km/s}$. In atomic units, $\sigma_V/\pi a_0^2$ scales roughly as $(E/E_0)^{-0.4}$. There is another power-law break in the viscosity cross section near $E \sim 0.1 E_0$ or $v \sim 1000 \text{ km/s}$ for $R \sim 100$. At these higher energies, higher order multipoles of the Van der Waals potential are able to be probed, and thus higher partial waves contribute to the viscosity cross section. The increased importance of higher partial waves cause the viscosity cross section to fall quickly as $\sigma_V/\pi a_0^2 \sim (E/E_0)^{-1.3}$. This behavior allows for the model to produce high viscosity cross sections at dwarf scales that decrease quickly enough with energy to become consistent with cluster observations at
higher velocities.

Depending on the values of $E_{\text{hf}}$ and $\alpha$, the inelastic upscattering cross section surpasses the viscosity cross section in halos with characteristic velocities as low as $v_{\text{rms}} \sim 300$ km/s, and can be over an order of magnitude larger at cluster scales. The ratio $\sigma_{++}/\sigma_V$ increases with decreasing $E_{\text{hf}}$; for $E_{\text{hf}} = 10^{-5}$ the upscattering cross section is $\mathcal{O}(10)$ times higher than the viscosity cross section at cluster scales. However, this does not necessarily mean that cooling is more efficient in atomic dark matter models with lower hyperfine splittings—although upscatterings may occur more frequently in these models, the smaller values of $E_{\text{hf}}$ mean that less energy is lost when the particles upscatter into the excited state and decay. We quantify the relative effectiveness of the heat transfer and energy loss mechanisms in the following Sec. 5.5.2.

The right panel of Fig. 5.5 demonstrates an interesting feature of the atomic dark matter model: if the average kinetic energy of two incoming particles in a halo is lower than the hyperfine splitting $E_{\text{hf}}$, then the dark atoms cannot upscatter to the hyperfine excited state and the inelastic cross section drops precipitously. The value of $E_{\text{hf}}$ thus sets a halo scale below which our mechanism for cooling is “turned off.” Dark atoms in halos with velocities below this scale may still be upscattered if they lie in the high-velocity tail of the velocity distribution, but the overall upscattering rate will be severely lowered by this effect. For hyperfine splittings of $E_{\text{hf}} = 10^{-4} E_0$, upscatterings are suppressed in halos with $v_{\text{rms}} \lesssim 40$ km/s. For hyperfine splittings of $E_{\text{hf}} = 10^{-5} E_0$, upscatterings are only suppressed for $v_{\text{rms}} \sim 1–2$ km/s, which corresponds to halos of mass $M_{\text{halo}} \sim 10^6 M_\odot$. If the cooling effects of collisional upscattering lead to observable effects in the structural evolution of the dark halo, then measurements of halo profiles below and above this turn-off velocity may allow us to infer a hyperfine splitting value in an atomic dark matter scenario.

In the right-hand subpanels of Fig. 5.6, we plot the fraction of the viscosity cross section $\sigma_V$ that arises from the inelastic viscosity cross section $\sigma_{++,V}$ at velocities of $v_{\text{rms}} = 40$ km/s.
and \( v_{\text{rms}} = 1000 \text{ km/s} \). We are particularly interested in the fraction \( \sigma_{++,V}/\sigma_V \) for the larger hyperfine splitting of \( E_{\text{hf}} = 10^{-4} E_0 \) at low velocities \( v_{\text{rms}} = 40 \text{ km/s} \): if this fraction is large, inelastic collisions are no longer approximately equivalent to elastic collisions in terms of momentum transfer, and comparison to existing SIDM constraints becomes difficult. From the right panel of Fig. 5.6, we see that for the majority of the target parameter space in this case (green triangles), the inelastic viscosity cross section contributes only a small (\( \lesssim 0.2 \)) fraction of the viscosity cross section at low velocities. We therefore assume that our comparison to existing SIDM cross section constraints are valid for these values of \( \alpha \) at which the viscosity cross section is close to the elastic viscosity cross section. We do note that there are narrow ranges of \( \alpha \) where the inelastic contribution to the viscosity cross section is significant (\( \gtrsim 0.8 \)). Nonetheless, we retain this definition for \( \sigma_V \) for two reasons: we want to include \( \sigma_{++,V} \) far from the hyperfine threshold where all scattering processes are approximately elastic, and it is not clear to what extent inelastic processes near threshold affect the SIDM constraints.

![Figures showing parameter space constraints and inelastic viscosity cross section](image)

Figure 5.6: The left subpanels in each figure show the constrained (vertical grey hatched) and target (solid orange and cross-hatched green) areas of parameter space. The right subpanels show the ratio of the inelastic viscosity cross section to the viscosity cross section as a function of \( \alpha \). The left and right figures are shown for fixed hyperfine splittings of \( E_{\text{hf}} = 10^{-5} \) and \( 10^{-4} \), respectively.
5.5.2 Halo cooling

We now quantify the energy loss rate due to hyperfine upscatterings. In what follows, we assume that halos are optically thin to the dark photons emitted following decays from the excited state. In the limit $E_\gamma \ll E_{\text{Ly} \alpha}$, which is valid for $E_\gamma = E_{\text{hf}}$, the Rayleigh scattering cross section for dark photons of energy $E_\gamma$ is approximately given by [84]

$$\sigma_{\text{Rayleigh}} \approx \frac{81\pi}{24} \left( \frac{\alpha}{m_e} \right)^2 \left( \frac{E_\gamma}{E_{\text{Ly} \alpha}} \right)^4.$$  \hspace{1cm} (5.24)

Since this quantity is negligible for dark photons with energies equal to the hyperfine splittings considered here, we assume that the emitted photons free stream out of the halo after emission.

Although the upscattering cross section can be large relative to the total viscosity cross section, the more important quantity to compare is the energy flow from each process. We now calculate the energy losses expected from inelastic upscatterings and compare this energy loss rate to the rate of inward heat flow due to dark atom-atom collisions. We outline below the net kinetic energy loss per particle expected per hyperfine upscattering and decay. We define the parameter

$$\varepsilon \equiv \frac{m_{H^*}^2 - m_H^2}{4m_H^2} = \frac{E_{\text{hf}}}{2m_H},$$  \hspace{1cm} (5.25)

where $m_{H^*} \equiv m_H + E_{\text{hf}}$ and $m_H$ refer to the excited and ground state masses, respectively. For the halos and parameter space studied here, $v_0 \ll 1$ and $\varepsilon \ll 1$, where $v_0$ is the incoming relative velocity of the colliding particles. In this limit, the net change in kinetic energy per particle per upscattering in the center-of-momentum frame is simply $\Delta KE_{\text{upscatter}} \approx -2m_H\varepsilon = -E_{\text{hf}}$. 

116
After the upscattered particle decays, its velocity in the lab frame is given by

\[ v_f^2 = 1 - \frac{1 - v_0^2}{1 + 4\varepsilon} \left(1 - \frac{2\varepsilon}{1 + 4\varepsilon (1 - v_0 \cos \theta)}\right)^{-2} \]  

(5.26)

where \( \theta \) is the angle between \( v_0 \) and the outgoing dark photon. The change in kinetic energy after undergoing this decay is

\[ \Delta KE_{\text{decay}} = 2\varepsilon m_H \left(8\varepsilon + 2v_0^2 - v_0 \cos \theta\right). \]  

(5.27)

After averaging over possible angles \( \theta \), this is a net increase—emitting a dark photon imparts a net positive kick velocity to the final ground state atom. However, the increase in kinetic energy from decay processes is \( \mathcal{O}(\varepsilon^2) \) or \( \mathcal{O}(\varepsilon v_0^2) \) (the dominant term depends on the value of \( E_{\text{hf}} \) and the halo in question), while the decrease from upscattering processes is \( \mathcal{O}(\varepsilon) \).

[The net \( \Delta KE_{\text{upscatter}} \) of both particles is still \( \mathcal{O}(\varepsilon) \) after shifting back to the lab frame.]

Henceforth, we will approximate the change in kinetic energy per particle per upscattering as \( \Delta KE \approx E_{\text{hf}} \) when investigating the regimes in which cooling effects become important.

The rate of energy loss in a thin shell of width \( dr \) at a radius \( r \) is given by

\[ 4\pi r^2 dr \Gamma_{\text{upscatter}} n_H \Delta E \approx 4\pi r^2 dr \sqrt{\frac{16 \sigma_{++}(v) v_{\text{rms}}(r) \rho^2(r) E_{\text{hf}}}{m_H^{2s}}}, \]  

(5.28)

where we have used the above reasoning to assume that the average energy lost per particle upscattering is approximately equal to \( E_{\text{hf}} \). The total amount of energy lost due to atomic upscatterings in a halo over its lifetime may be estimated by integrating Eq. (5.28) over the radius \( r \) and multiplying by the lifetime. We verify that the energy lost due to collisional cooling is never more than 0.001 times the total initial kinetic energy of the halo—for the range of parameters studied here, hyperfine upscatterings cannot disrupt the entire halo. However, as we demonstrate below, the energy losses from inelastic upscatterings can be up to 0.1–0.5 times the rate of inward heat flow from scatterings within the inner halo for
particular ranges in $\alpha$. Hyperfine upscatterings may therefore play an important role in the structural evolution of the inner halo if the atomic SIDM model parameters lead to significant cooling rates.

To calculate the rate of heat flow resulting from particle collisions, we treat the dark halo as a fluid with the luminosity $L$ at radius $r$ given by [30, 193]

$$\frac{L}{4\pi r^2} = -\kappa \frac{\partial T}{\partial r} = -\frac{3}{2} \frac{abv \sigma_v}{m_H} \left[ a \left( \frac{\sigma_v}{m_H} \right)^2 + \frac{b}{C} \frac{4\pi G \rho v^2}{\rho v^2} \right]^{-1} \frac{\partial v^2}{\partial r}. \tag{5.29}$$

The dimensionless coefficients $a$ (which describes hard sphere scattering)$^2$, $b$ (which describes the short mean-free-path regime), and $C$ (which describes the scale at which the transition between long- and short-mean-free path regimes occurs) are taken to be $a = \sqrt{16/\pi}, b = 25\sqrt{\pi}/32$, and $C \approx 0.75$ as in Ref. [193]. In Sec. 5.4.2, we explain why the viscosity cross section $\sigma_v$—as opposed to the transfer cross section $\sigma_T$—is the quantity that best describes the rate of events which result in a net transfer of energy. In line with this reasoning, we use our calculated values for $\sigma_v$ in Eq. (5.29) when calculating the rate of heat flow.

In Fig. 5.7 we show the ratio of heat lost through upscattering and decays to heat inflowing through collisional processes in a thin shell at radius $r = 0.5 \, r_s$ in low-mass and high-mass halos. The low-mass halo is chosen to be approximately the lowest-mass halo in which the inelastic upscattering rate is not suppressed by the average particle velocity being lower than the hyperfine splitting. For a hyperfine splitting of $10^{-4}E_0$ ($10^{-5}E_0$), this corresponds to a halo mass of $10^{10} \, M_\odot$ ($6 \times 10^6 \, M_\odot$). The high-mass halo corresponds to a cluster-scale halo with $M_{\text{halo}} = 10^{14} \, M_\odot$. We choose to plot this ratio at the radius $r = 0.5 \, r_s$ as this is roughly

\footnote{This value given for the coefficient $a$ in Eq. (5.29) assumes elastic scatterings. As noted previously in Sec. 5.5.1, $\sigma_v$ has contributions from both elastic and inelastic scattering cross sections. However, $\sigma_v$ can be considered as an approximately elastic cross section if 1) $E_{\text{hf}} \ll m_H v^2$, i.e. the hyperfine splitting is small compared to the initial energies of the interacting particles, or 2) the viscosity upscattering cross section $\sigma_{+,V}$ does not contribute significantly to $\sigma_v$. Either one or both of these conditions are met for a large majority of our favored regions in parameter space (see Fig. 5.6). We therefore consider the use of this value for $a$ to be reasonable.}
where the cooling and heating rates are both maximized in the halo.

We find that cooling is preferentially important for small halos relative to big halos. This is because the energy loss $\Delta KE \sim E_{\text{hf}}$ is fixed, while the typical kinetic energy per particle increases with increasing halo mass. Furthermore, the cooling and heating processes have different overall effects on the halo: heating the inner part of the halo is caused by a transference of energy within the halo, whereas the energy emitted as dark photons is presumably not reabsorbed and is instead permanently lost from the halo. In Fig. 5.7, the cases with high cooling rates have high heating rates as well, so the moderate cooling-to-heating ratios could be underestimating the overall importance of cooling.

The structural evolution of the halo in instances of non-negligible cooling effects is nontrivial and may be modeled using numerical integration methods. Evolving an atomic dark matter halo over cosmic time with the inclusion of dark cooling as well as the baryonic potential in the innermost region $r \lesssim 0.1 r_s$ is beyond the scope of this paper, but will be addressed in future work.

### 5.6 Additional Considerations at the Cluster Scale

In our above treatment of the interactions between neutral dark atoms, we do not consider the possibility of collisional ionizations or excitations to $n \geq 2$ states. This simplification is adequate for lower-mass halos in which dark matter particles have enough energy to excite the hyperfine state, but not enough energy to ionize or excite the $n = 2$ state through collisions. As halo masses increase and the typical particle velocities surpass the ionization and $n = 2$ excitation thresholds, these processes may potentially affect the halo structure. Below, we discuss the potential for these processes to affect cluster-scale halos.
Figure 5.7: We compare the effects of SIDM heating and cooling mechanisms in atomic dark matter halos by plotting the ratios of outward energy flow lost through cooling over the inward heat flow from scatterings at a radius of $r = 0.5 r_s$ in the halo, which is approximately the radius at which the inward heat flow due to scatterings and the outward energy loss due to upscatterings are greatest, as well as $r = r_s$. The left and right figures in both rows are shown for fixed hyperfine splittings of $E_{hf} = 10^{-5}$ and $10^{-4}$, respectively. The lower mass halo plotted in each panel (black triangles) corresponds to the smallest halos in which upscatterings to the hyperfine excited state are not suppressed by low particle velocities. We also show the cooling to heating ratios at cluster scales for cooling through hyperfine excitations (magenta circles) and $n = 2$ excitations (blue squares). See Sec. 5.6.1 for details and discussion regarding our estimation of the $n = 2$ cooling rate.

### 5.6.1 Upscatterings to the $n = 2$ excited state

Above a particular halo mass scale, particle velocities may be high enough to collisionally excite atoms into the $n = 2$ excited state, which would quickly decay back to the ground $n = 1$ state. This additional cooling mechanism may affect the halo structure if the relative particle velocities are above the threshold for upscattering one of the incident particles into
the $n = 2$ state,

$$v^2 > \frac{9}{8g_\text{e}g_\text{p}} \frac{E_{\text{hf}}}{E_0}.$$  \hspace{1cm} (5.30)

For hyperfine splittings of $E_{\text{hf}} = 10^{-5}E_0$, relative particle velocities $v \gtrsim 500 \text{ km/s}$ (corresponding roughly to a halo mass of $\sim 10^{13}M_\odot$) may result in upscattering to the $n = 2$ state. Relative velocities above $\sim 1600 \text{ km/s}$ are needed for hyperfine splittings of $E_{\text{hf}} = 10^{-4}E_0$, which may be reached in massive clusters ($M_{\text{halo}} \gtrsim 4 \times 10^{14}M_\odot$) or systems of merging clusters.

We use the following method to obtain approximate values for this cross section in order to estimate the potential cooling losses from $n = 2$ upscatterings. From the analytic derivation of cross sections for collisions between neutral ground-state SM hydrogen atoms presented in Ref. [33], we see that the $n = 2$ upscattering cross sections $\sigma_{n=2}(1s + 1s \rightarrow 1s + 2s/2p)$ may be written using $v/\alpha$ as the independent variable, with $\sigma_{n=2}$ in units of the geometric cross section $\pi a_0^2$. We then scale the experimental measurements of this cross section [134, 31] to estimate the collisional $n = 2$ upscattering cross sections for the dark hydrogen analogs.

Using this scaling, the $n = 2$ upscattering cross sections are typically much smaller than the hyperfine upscattering cross sections ($\sigma_{n=2} \lesssim 0.01 \sigma_{++}$).\footnote{Cross sections quoted in Secs. 5.6.1 and 5.6.2 make the simplifying assumption that all particle pairs have the same typical relative velocity for their position in the halo. If the cross section is appropriately averaged over the Maxwell-Boltzmann distribution and relative velocities below the minimum threshold for the interaction are excluded, results are consistent within $\sim 20\%$.} However, the energy lost per upscattering is much greater than [of order $(E_0/E_{\text{hf}})^{-1}$ times] the energy lost per hyperfine upscattering. Since the $n = 2$ cooling rate may thus be non-negligible, we estimate it using the expression on the left-hand side of Eq. (5.28) with $\Gamma_{\text{upscatter}} = n_H\sigma_{n=2}v$ and $\Delta E = \Delta E_{\text{Ly} \alpha} = 3/4 B_H$. We show this estimate of the $n = 2$ cooling rate over the heating rate in Fig. 5.7.

For hyperfine splittings of $E_{\text{hf}} = 10^{-5}E_0$, we find that the $n = 2$ cooling rate over the
collisional heating rate can be up to $\sim 0.1$ in cluster-scale halos of mass $10^{14} M_\odot$. Although the relative particle velocities in smaller halos are above the threshold for $n = 2$ upscattering, the cross section $\sigma_{n=2}$ for these interactions decreases with velocity in this regime such that this ratio is an order of magnitude lower for a $10^{13} M_\odot$ halo than for a $10^{14} M_\odot$ halo. For hyperfine splittings of $E_{\text{hf}} = 10^{-4} E_0$, halos must be at least $\sim 4 \times 10^{14} M_\odot$ in mass for enough particles to surpass the threshold velocity for $n = 2$ upscattering. For a $4 \times 10^{14} M_\odot$ halo, we find that the cooling-to-heating ratio from $n = 2$ upscattering is $\simeq 0.1$ for $\alpha \lesssim 0.04$ and decreases to $\simeq 0.02$ for $\alpha \simeq 0.1$. We therefore expect that while the cooling effects from $n = 2$ upscattering processes may be large enough to affect halo structure, they do not affect the evolution and growth of lower mass halos and only become significant at the cluster scale.

### 5.6.2 Ionization in the late universe

Once halos form, dark hydrogen remains intact if there is insufficient energy in the system for ionization: $(v/\alpha)^2 < 1/f(R)$. The particle velocities in a halo may be high enough to ionize the majority of dark atoms if the following condition is met:

$$ v^2 > \frac{3}{2g_e g_p} \frac{E_{\text{hf}}}{E_0}. $$  \hfill (5.31)

For the hyperfine splitting $E_{\text{hf}} = 10^{-5}(10^{-4}) E_0$, the above relation is satisfied for relative velocities of $v \gtrsim 580(1800)$ km/s: atoms have enough energy for ionization in isolated cluster halos if $E_{\text{hf}} \sim 10^{-5}$, or in merging clusters if $E_{\text{hf}} \sim 10^{-4}$. This raises the concern that dark matter halos above these velocities may contain a significant ionized component. However, the above condition is necessary but not sufficient to ionize the majority of the dark atoms in a halo—the cross section for collisional ionization $\sigma_i$ must also be high enough to allow for particles to experience such an interaction over the cluster lifetime or merger time.
In a similar manner in which we use the analytic expression from Ref. [33] to estimate the 
$n = 2$ excitation cross sections, we scale the experimental measurements of the collisional 
ionization cross section [134, 31] to estimate the collisional ionization cross sections for atomic 
dark hydrogen. At its maximum, this cross section is approximately the geometric cross 
section: $\sigma_{i,\text{max}} \simeq \pi a_0^2$. Thus, for hyperfine splittings of $E_{\text{hf}} = 10^{-4}$, the collisional ionization 
cross section is always $\sigma_i \lesssim 0.01 \text{ cm}^2/\text{g}$, and we do not expect cluster halos to be significantly 
ionized.

However, the geometric cross sections in our preferred region of parameter space for splittings 
of $E_{\text{hf}} = 10^{-5} E_0$ are large enough ($\pi a_0^2 \sim 0.2$) such that the collisional ionization cross 
section may be as large as $\sigma_i \sim 0.1 \text{ cm}^2/\text{g}$ for relative velocities above $v \sim 2000 \text{ km/s}$. 
Massive clusters or systems of merging clusters above these velocities may become ionized if 
the hyperfine splitting is of order $E_{\text{hf}} \sim 10^{-5} E_0$. Ionization may result in increased mass loss 
during mergers, cooling effects (due to recombination followed by emission of a photon), and a 
variety of possible scattering cross sections between ions, electrons, and atoms. The complex 
effects of ionization on the structural evolution of a halo are not included in the comparison of 
our results with existing cluster-scale observations, and we caution that hyperfine splittings 
of $E_{\text{hf}} = 10^{-5} E_0$ in this model may alter the dark matter structure at high mass scales to be 
inconsistent with observations. Again, for the aforementioned reason of low ionization cross 
section, this issue of late-time ionization does not significantly affect our results for splittings 
of $E_{\text{hf}} = 10^{-4}$.

5.7 Conclusions

In this paper, we have investigated a model of self-interacting dark matter that mimics the 
properties of atomic hydrogen. Dark matter in the late universe takes the form of dark 
hydrogen, which is neutral under a new U(1) gauge force. We do not assume a specific
interaction between this new U(1) and the SM for the predictions in this paper. The key features of our work are the inclusion of a hyperfine interaction, which induces an energy splitting in the ground state of dark hydrogen, and the calculation of the basic heat transport properties in halos, which allows us to identify the viable regions of parameter space where the small-scale puzzles can be solved.

Collisions of dark atoms in halos may induce hyperfine excitations, which then decay by emitting dark photons. Halo cooling from this upscattering and subsequent energy loss works against halo heating that occurs from the scattering processes. To study these effects on halo structure, we calculated the cross sections for dark hydrogen scattering over a wide range of parameter space, using techniques from standard hydrogen to aid in numerically solving the Schrödinger equation. The velocity dependence of the cross sections allows the heating and cooling mechanisms to operate differently on scales of dwarf spheroidal galaxies ($v_{\text{rms}} = 40 \text{ km/s}$) compared to scales of galaxy clusters ($v_{\text{rms}} = 1000 \text{ km/s}$).

We argue that the viscosity cross section where both the forward and backward scattering are suppressed is the better quantity, compared to the momentum-transfer cross section, to use when comparing to SIDM simulation results and observational constraints. The velocity dependence of the viscosity cross section shows a sharp drop for kinetic energies larger than about $0.1 \ E_0 \approx 0.1 \ a^2 m_H / R$ as contributions from higher partial waves become important. This allows the model to be consistent with cluster constraints. The typical cross section at $E = 0.1 \ E_0$ is roughly $10 \ a^2_0$ and scales approximately as $E^{-1.3}$ above these energies. For kinetic energies below $0.1 \ E_0$, we see a steady increase in the viscosity cross section with decreasing relative velocity, which implies that the scattering processes are very important in small halos. The viscosity cross section in this regime scales roughly as $E^{-0.4}$.

We have found regions of parameter space for the atomic dark matter model in which dark matter self-interactions can explain the measured core sizes in both dwarfs and clusters, while being consistent with all other observations including cluster halo shapes. The solutions
are not fine-tuned; for a hyperfine splitting that is about $10^{-4}E_0$, we find that much of the parameter space with $\chi_e < 0.01$ and dark hydrogen mass in the 10–100 GeV range is viable. In this part of parameter space, the dark matter is in atomic form and we find that cooling mechanisms are generically important for the structure of low-mass halos (masses below $10^{10} M_\odot$) but not important enough to completely disrupt these halos. An immediate consequence of this observation is that the collapse of small halos at early times will be affected by the cooling and, therefore, it is likely that the growth of the seeds of supermassive black holes will also be altered. We leave this discussion for another paper.

The kinetic energy of dark matter particles in galaxy clusters is large enough to allow for additional atomic physics. We find that collisional excitations to $n = 2$ and ionizations could be significant processes in galaxy clusters for $E_{\text{hf}} = 10^{-5}$. For $E_{\text{hf}} = 10^{-4}$, we show that the cooling rate due to these processes is subdominant to the heating rate and our predictions, which assume negligible scattering to $n = 2$ and fully atomic dark hydrogen, are robust. Thus, galaxy clusters are important astrophysical laboratories for testing atomic dark matter models.

The interactions between the dark matter and the light mediator in the early Universe modifies the kinetic decoupling of the dark matter. The kinetic decoupling temperature may be used to estimate the minimum halo mass in the universe. Assuming that the ratio of the hidden sector temperature to the visible photon temperature at late times is 0.6 (close to the maximum allowed by BBN constraints), we find that the range of halo minimum masses in the viable regions of parameter space are between $10^{3.5}$ and $10^7 M_\odot$. These minimum masses are smaller than the host masses of the currently observed dwarf galaxies, but much larger than the minimum masses predicted for dark matter in weak-scale theories. If the ratio of the temperatures is smaller (due to the fact that the two sectors were reheated to different temperatures and remained decoupled), then the minimum halo masses will be lower by a factor of $(\xi/0.6)^{9/5}$. 

125
In summary, we have shown that an analog of hydrogen in the hidden sector is a viable self-interacting dark matter candidate that can alleviate the small-scale structure formation puzzles, and the dissipative nature of atomic dark matter provides a phenomenologically rich foundation to make observational predictions.
Chapter 6

Conclusions

The search for dark matter is one of the major outstanding problems in modern physics. This thesis has described two broad classes of dark matter particle candidates—collisionless WIMP dark matter and self-interacting dark matter—and how each might be tested using astrophysical observations.

Chapters 2, 3, and 4 describe a potential signal of WIMP dark matter annihilation gamma-rays in the Milky Way Center (a. k. a. the ‘galactic center excess’). Chapter 5 describes the atomic dark matter model, in which scatterings between neutral dark hydrogen atoms are lead to the observed range of dark matter halo density profiles from dwarf to cluster scales.

In chapter 2, we characterize the spectrum of the galactic center excess and determine that the high-energy emission from this source ($\gtrsim 10$ GeV) appears to be spatially non-uniform. This finding has important implications for interpretations of its origin, as prompt dark matter annihilation would produce a spatially uniform spectrum, in contrast to the observed signal. Alternate sources such as inverse Compton upscattering of electrons injected by cosmic ray outbursts, millisecond pulsars, or WIMP annihilations to leptons may be able to produce spatially non-uniform spectra.
Chapter 3 describes the detection of an extended gamma-ray excess source that is correlated with the 3.4 $\mu m$ infrared emission in the central Milky Way. We describe how this new source could conceivably share the same source as the higher-energy galactic center excess: if dark matter is annihilating in the Milky Way center with equal branching ratios to each lepton flavor, then the prompt emission from $\tau$ lepton final states could produce the galactic center excess, while inverse Compton upscattering of the $e$ final states off of background starlight could produce the new source described in this chapter.

Chapter 4 presents a scenario where the galactic center excess is produced through WIMP annihilations to mediators, which subsequently decay into pairs of differently flavored leptons. Gamma-ray emission may be produced through prompt emission from the final $\tau^{\pm}$ states and/or inverse Compton emission from primary and secondary $e^{\pm}$. This model may relieve tension between the two-body annihilation cross section required to produce the observed flux of the galactic center excess and null observations of similar excesses in dwarf galaxies. Our model also avoids constraints on dark matter annihilation into $e^{\pm}$ from AMS-02 cosmic-ray measurements.

Finally, in chapter 5 we discuss a specific model, atomic dark matter, which is motivated as a solution to unresolved discrepancies between predictions from CDM simulations and the inferred dark matter distribution in galaxies within $\sim$kiloparsec scales of the central halo. We show how rotation curve fitting implies a scattering cross section of $\mathcal{O}(1)$ cm$^2$/g at velocities $< 100$ km/s, and outline the regions of atomic dark matter parameter space where this model is able to match the estimated core sizes in dwarf and low-surface brightness galaxies ($v_{\text{rms}} \simeq 30 - 100$ km/s while obeying constraints from galaxy clusters ($v_{\text{rms}} \gtrsim 1000$ km/s). We also make predictions for minimum halo masses, some of which may be testable in the future using gravitational lensing and tidal stream observations.

The vast sizes and scales involved in astrophysical systems make them ideal targets for dark matter searches. Indeed, if the coupling between the dark and Standard Model particle
sectors is non-existent or sufficiently small, astrophysical studies may be our only avenue through which to probe the particle nature of dark matter.


