

UNIVERSITY of CALIFORNIA  
Santa Barbara

**A Low-threshold Analysis of Data from the  
Cryogenic Dark Matter Search Experiment**

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

Doctor of Philosophy

in

Physics

by

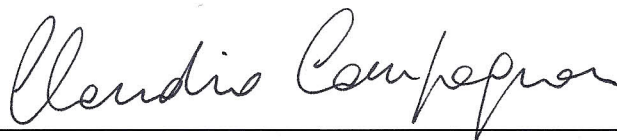
Raymond A. Bunker III

Committee in charge:

Professor Harry Nelson, Chair  
Professor Claudio Campagnari  
Professor Mark Srednicki

December 2011

The dissertation of Raymond A. Bunker III is approved:

A handwritten signature in cursive script, reading "Claudio Campagnari".

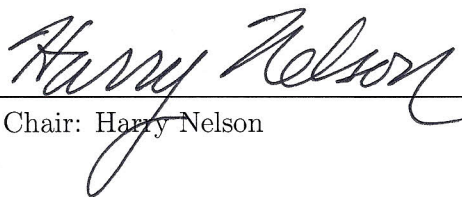
---

Claudio Campagnari

A handwritten signature in cursive script, reading "Mark Srednicki".

---

Mark Srednicki

A handwritten signature in cursive script, reading "Harry Nelson".

---

Chair: Harry Nelson

December 2011



**A Low-threshold Analysis of Data from the  
Cryogenic Dark Matter Search Experiment**

© Copyright 2011

by

Raymond A. Bunker III

It was a dark and stormy night...

## Acknowledgements

Foremost, I owe my advisor Harry Nelson a debt of gratitude for his unwavering support and patience. I judge my graduate career a success if I have learned to be half the scientist he is. This thesis is the culmination of over a decade of scientific endeavor. Without the benefit of Harry's experience and insight it would likely have taken two decades to reach its current maturity. His uncanny ability to simplify and prioritize critical concepts redirected my efforts toward more productive ends more times than I can recall.

I would like to thank the High Energy Physics (HEP) group, whose open and friendly support through the years I have come to take for granted. It has been an absolute pleasure to work with the world-class faculty: David Caldwell, Claudio Campagnari, Joe Incandela, Ben Monreal, Jeff Richman, Dave Stuart, and Mike Witherell. Special thanks to the top-notch engineering and support staff: Sam Burke, Dan Callahan, Debbie Cedar, Dave Hale, Susanne Kyre, Lap Leung, and Dean White. They deserve credit for all the tedious, labor-intensive, and skilled (real) work that went into the successful construction and operation of our science experiments. It would have been infinitely more difficult and a lot less fun without them!

I would also like to thank the CDMS collaboration. My experience with collaborative science before joining CDMS was frustrating at best. I have benefited from witnessing a collaborative style of friendly and productive camaraderie, while taking advantage of a nearly bottomless pool of experience and knowledge to help further my own research interests. I would especially like to thank Dan Bauer for his guidance over the years. I have always been able to count on his sound and reasonable advice. Particular thanks to Richard Schnee and Steve Yellin for their tireless help with the data analysis presented in this thesis, and to Rupak Mahapatra and Joel Sander for their support and friendship. There are many other present and past CDMS collaborators that helped bring this thesis to fruition: Dan Akerib, Scott Armel, Laura Baudis, Paul Brink, Blas Cabrera, Clarence Chang, Jodi Cooley, Mike Crisler, Fritz DeJongh, Mike Dragowski, Don Driscoll, Jeff Filippini, Rick Gaitskell, Sunil Golwala, Jeter Hall, Don Holmgren, Sharmila Kamat, Vuk Mandic, David Moore, Sae Woo

Nam, Robert Nelson, Walter Ogburn, Eric Ramberg, Angela Reisetter, Tarek Saab, Bernard Sadoulet, Chris Savage, Dennis Seitz, Andrew Sonnenschein, Kyle Sundqvist, John-Paul Thompson, Gensheng Wang, and Betty Young. The list is so long that I am sure to have missed a few. Nevertheless, the entire CDMS collaboration, past and present, has had a tremendous influence on my education and training as an experimental physicist.

Finally, I extend heart-felt thanks and love to my wonderfully supportive family who have waited patiently and without complaint for me to finish my education. My parents and grandparents have encouraged my every step along the way, providing the emotional support every son and grandson so needs and deserves. And a great big, huge thank you to my sister, her husband and my nephew. This thesis would not exist without your love and support.

# Curriculum Vitæ

## Raymond A. Bunker III

### Education

- University of California, Santa Barbara (UCSB)
  - Ph.D., Physics *September 2011*
  - M.A., Physics *March 2004*
- Massachusetts Institute of Technology (MIT)
  - B.S., Mathematics with Physics Minor *June 1997*

### Research Experience

- Graduate Student Researcher  
UCSB High Energy Physics Group
  - Cryogenic Dark Matter Search (CDMS) *June 1998 to September 2011*
  - Neutron Multiplicity Meter *September 2006–2011*
- Undergraduate Research Opportunities Program  
MIT Laboratory for Nuclear Science
  - Out-of-Plane Spectrometer System (OOPS) *June 1995 to May 1997*
  - Axion Dark Matter Experiment (ADMX) *Summer of 1994*

### Teaching Experience (UCSB)

- Teaching Assistant
  - Physics 125: Elementary Particle Physics *Spring 2000*
  - Physics 127: Analog Electronics *Fall 2000*
- Lab Instructor
  - Physics 6: Introductory Experimental Physics *Spring 1998*
  - Physics 128: Advanced Experimental Physics *Fall 1997 & Winter 1998*

## Selected Publications

- CDMS and EDELWEISS Collaborations (Z. Ahmed *et al.*)  
*Combined limits on WIMPs from the CDMS and EDELWEISS experiments*  
Physical Review **D84**, 011102 (2011), doi:10.1103/PhysRevD.84.011102
- CDMS Collaboration (Z. Ahmed *et al.*)  
*Results from a low-energy analysis of the CDMS II germanium data*  
Physical Review Letters **106**, 131302 (2011), doi:10.1103/PhysRevLett.106.131302
- CDMS Collaboration (Z. Ahmed *et al.*)  
*Search for inelastic dark matter with the CDMS II experiment*  
Physical Review **D83**, 112002 (2011), doi:10.1103/PhysRevD.83.112002
- CDMS Collaboration (D. S. Akerib *et al.*), *corresponding author*  
*Low-threshold analysis of CDMS shallow-site data*  
Physical Review **D82**, 122004 (2010), doi:10.1103/PhysRevD.82.122004
- CDMS Collaboration (Z. Ahmed *et al.*)  
*Dark matter search results from the CDMS II experiment*  
Science **327**, 1619 (2010), doi:10.1126/science.1186112
- CDMS Collaboration (Z. Ahmed *et al.*)  
*Analysis of the low-energy electron-recoil spectrum of the CDMS experiment*  
Physical Review **D81**, 042002 (2010), doi:10.1103/PhysRevD.81.042002
- CDMS Collaboration (Z. Ahmed *et al.*)  
*Search for axions with the CDMS experiment*  
Physical Review Letters **103**, 141802 (2009), doi:10.1103/PhysRevLett.103.141802
- CDMS Collaboration (D. S. Akerib *et al.*)  
*Limits on spin-dependent WIMP-nucleon interactions from CDMS*  
Physical Review **D73**, 011102 (2006), doi:10.1103/PhysRevD.73.011102
- CDMS Collaboration (D. S. Akerib *et al.*)  
*First results from CDMS in the Soudan Underground Laboratory*  
Physical Review Letters **93**, 211301 (2004), doi:10.1103/PhysRevLett.93.211301
- CDMS Collaboration (D. S. Akerib *et al.*)  
*New results from the CDMS experiment*  
Physical Review **D68**, 082002 (2003), doi:10.1103/PhysRevD.68.082002

## Abstract

### A Low-threshold Analysis of Data from the Cryogenic Dark Matter Search Experiment

by

Raymond A. Bunker III

Although dark matter appears to constitute over 80% of the matter in the Universe, its composition is a mystery. Astrophysical observations suggest that the luminous portions of the Galaxy are embedded in a halo of dark-matter particles. Weakly Interacting Massive Particles (WIMPs) are the most studied class of dark-matter candidates and arise naturally within the context of many weak-scale supersymmetric theories. Direct-detection experiments like the Cryogenic Dark Matter Search (CDMS) strive to discern the kinetic energy of recoiling nuclei resulting from WIMP interactions with terrestrial matter. This is a considerable challenge in which the low (expected) rate of WIMP interactions must be distinguished from an overwhelming rate due to known types of radiation.

An incontrovertible positive detection has remained elusive. However, a few experiments have recorded data that appear consistent with a low-mass WIMP. This thesis describes an attempt to probe the favored parameter space. To increase sensitivity to low-mass WIMPs, a low-threshold technique with improved sensitivity to small energy depositions is applied to CDMS shallow-site data. Four germanium and two silicon detectors were operated between December 2001 and June 2002, yielding 118 days of exposure. By sacrificing some of the CDMS detectors' ability to discriminate signal from background, energy thresholds of  $\sim 1$  and  $\sim 2$  keV were achieved for three of the germanium and both silicon detectors, respectively. A large number of WIMP candidate events are observed, most of which can be accounted for by misidentification of background sources. No conclusive evidence for a low-mass WIMP signal is

found. The observed event rates are used to set upper limits on the WIMP-nucleon scattering cross section as a function of WIMP mass. Interesting parameter space is excluded for WIMPs with masses below  $\sim 9 \text{ GeV}/c^2$ . Under standard assumptions, the parameter space favored by interpretations of other experiments' data as low-mass WIMP signals is partially excluded, and new parameter space is excluded for WIMP masses between 3 and  $4 \text{ GeV}/c^2$ .



# Contents

<b>List of Figures</b>	<b>xv</b>
<b>List of Tables</b>	<b>xxviii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 The Standard Model . . . . .	2
1.2 The Standard Cosmology . . . . .	5
1.2.1 Theoretical Framework . . . . .	6
1.2.2 Thermal History . . . . .	11
1.3 Observational Evidence . . . . .	15
1.3.1 Big Bang Nucleosynthesis . . . . .	16
1.3.2 The Cosmic Microwave Background Radiation . . . . .	23
1.3.3 Type Ia Supernovae . . . . .	30
1.3.4 Baryon Acoustic Oscillations . . . . .	37
1.3.5 Cosmic Concordance . . . . .	44
<b>2 Dark Matter</b>	<b>49</b>
2.1 Mass-to-Light Ratios . . . . .	50
2.2 Galactic Dark Matter . . . . .	54
2.2.1 Spiral Galaxy Rotation Curves . . . . .	55
2.2.2 The Tully-Fisher Relation . . . . .	70
2.2.3 Elliptical Galaxies . . . . .	73
2.3 Dark Matter in Galaxy Clusters . . . . .	78
2.3.1 The Virial Theorem . . . . .	81
2.3.2 X-Ray Emission . . . . .	89
2.3.3 Gravitational Lensing . . . . .	95
2.3.4 The Sunyaev-Zel'dovich Effect . . . . .	102

2.4	Dark Matter Candidates . . . . .	105
2.4.1	Baryons . . . . .	109
2.4.2	Neutrinos . . . . .	121
2.4.3	Weakly Interacting Massive Particles . . . . .	126
2.5	Supersymmetry . . . . .	131
2.5.1	General Overview . . . . .	132
2.5.2	Motivation . . . . .	135
2.5.3	The MSSM, LSP & Neutralino . . . . .	140
2.5.4	Neutralino Mass . . . . .	147
<b>3</b>	<b>Dark Matter Detection</b>	<b>155</b>
3.1	Production at Particle Colliders . . . . .	156
3.2	Indirect Detection . . . . .	163
3.2.1	Electrons & Positrons . . . . .	165
3.2.2	Gamma Rays . . . . .	178
3.3	Direct Detection . . . . .	186
3.3.1	Approximate Deposited Energy . . . . .	186
3.3.2	Scattering Rates & Energy Spectra . . . . .	188
3.3.3	Backgrounds & Shielding . . . . .	206
3.3.4	Discrimination . . . . .	223
3.3.5	Low-mass WIMP Signals? . . . . .	229
<b>4</b>	<b>The CDMS Experiment</b>	<b>241</b>
4.1	Introduction . . . . .	241
4.2	Detector Technology . . . . .	244
4.2.1	Ionization Measurement . . . . .	246
4.2.2	Phonon Measurement . . . . .	255
4.2.3	Recoil Energy & Ionization Yield . . . . .	266
4.2.4	Alternate Recoil-Energy Estimators . . . . .	272
4.2.5	Position Reconstruction . . . . .	273
4.3	The Shallow-site Installation . . . . .	278
4.3.1	The Stanford Underground Facility . . . . .	279
4.3.2	Passive Shielding . . . . .	282
4.3.3	Active Shielding—The Muon Veto . . . . .	284
4.3.4	The Dilution Unit, Cryostat, & Tower Assembly . . . . .	304
4.3.5	Room-Temperature Electronics & The DAQ . . . . .	309

<b>5</b>	<b>A Low-Threshold Analysis</b>	<b>315</b>
5.1	Run 21 . . . . .	316
5.1.1	Data Samples . . . . .	317
5.2	Thresholds . . . . .	324
5.2.1	Hardware Trigger . . . . .	324
5.2.2	Software Threshold . . . . .	328
5.2.3	Viable Low-Threshold Detectors . . . . .	333
5.3	Analysis Cuts . . . . .	334
5.3.1	Data Quality . . . . .	335
5.3.2	Event Bursts . . . . .	339
5.3.3	Fiducial Volume . . . . .	343
5.3.4	Single Scatters . . . . .	349
5.3.5	Veto Anticoincidence . . . . .	350
5.3.6	Nuclear-Recoil Band . . . . .	352
5.4	Energy Scale & Resolution . . . . .	357
5.4.1	Electron-Recoil Energy Response . . . . .	358
5.4.2	Nuclear-Recoil Energy Scale . . . . .	373
5.4.3	Zero-Energy Resolution . . . . .	377
5.5	Combined Detection Efficiencies . . . . .	380
<b>6</b>	<b>Results</b>	<b>383</b>
6.1	WIMP Candidates . . . . .	383
6.2	Backgrounds . . . . .	384
6.3	Exclusion Limits . . . . .	391
6.4	Systematic Studies . . . . .	401
6.4.1	Galactic Escape Velocity . . . . .	401
6.4.2	Nuclear-Recoil Energy Scale Revisited . . . . .	403
<b>A</b>	<b>The Veto Front-End Board</b>	<b>413</b>
<b>B</b>	<b>Threshold Efficiencies</b>	<b>435</b>
B.1	Hardware Thresholds . . . . .	435
B.2	Software Thresholds . . . . .	449
<b>C</b>	<b>Analysis Cuts</b>	<b>461</b>
C.1	Fiducial Volume . . . . .	461
C.2	Nuclear-Recoil Band . . . . .	475

<b>D</b>	<b>Energy Resolution</b>	<b>485</b>
D.1	Electron-Recoil Energy Response . . . . .	485
<b>E</b>	<b><math>^{252}\text{Cf}</math> Spectral Shapes</b>	<b>491</b>
E.1	Differential Scattering Rate . . . . .	491
E.2	Differential Number Density . . . . .	493
E.3	Elastic Scattering Cross Section . . . . .	496
E.4	Elastic Scattering Angular Probability . . . . .	497
E.5	Recoil Spectra . . . . .	499
<b>F</b>	<b>Detection Efficiencies</b>	<b>503</b>
<b>G</b>	<b>WIMP Candidates</b>	<b>511</b>
	<b>Bibliography</b>	<b>515</b>

# List of Figures

1.1	Combined measurement of the cosmological constant's equation of state and the Hubble constant . . . . .	10
1.2	Radiation, matter and dark-energy densities as a function of redshift . . . . .	11
1.3	Evolution of the abundances of light elements as a function of time and temperature . . . . .	18
1.4	Big Bang nucleosynthesis predictions for the abundances of light elements as a function of the baryon-to-photon ratio . . . . .	20
1.5	The Wilkinson Microwave Background Anisotropy Probe's all-sky map of the cosmic microwave background anisotropy . . . .	25
1.6	Cosmic microwave background power spectrum . . . . .	27
1.7	Hubble diagram for Type Ia supernovae . . . . .	35
1.8	Type Ia supernovae constraint on the nature of dark energy . .	36
1.9	Spatial-redshift correlation function for a sample of luminous, red galaxies from the Sloan Digital Sky Survey . . . . .	43
1.10	Residual power spectrum of large-scale density fluctuations in the spatial-redshift distribution of nearly one million nearby galaxies . . . . .	44
1.11	Combined astrophysical constraints for the dark-energy density and equation of state as a function of the total matter density .	46
1.12	Combined astrophysical constraints on the nature of dark energy	47
2.1	Early rotation curve for the Andromeda Galaxy . . . . .	57
2.2	Universal rotation curves for spiral galaxies in bins of luminosity	61
2.3	Compilation of rotational velocities as a function galactocentric radius for the Milky Way . . . . .	64

2.4	Tully-Fisher relation between galactic luminosity and rotational velocity . . . . .	72
2.5	Mass profile for the dark-matter-dominated elliptical galaxy NGC 7785 . . . . .	76
2.6	X-ray temperature as a function of stellar velocity dispersion for elliptical galaxies . . . . .	77
2.7	Radial dependence of elliptical-galaxy mass-to-light ratios . .	78
2.8	Mass-to-light ratio as a function of velocity dispersion for galaxy groups . . . . .	80
2.9	Mass-to-light ratios for differently sized galaxy systems as a function of their virial masses . . . . .	86
2.10	Galaxy-cluster and -group mass-to-light ratios (versus virial mass) compared to cosmological models . . . . .	88
2.11	Evolution of the cluster-mass function . . . . .	94
2.12	Optical mass-to-light ratios derived from weak gravitational-lensing measurements of cluster mass . . . . .	100
2.13	False-color photograph of the Bullet cluster, an example of a galaxy-cluster collision . . . . .	101
2.14	Baryon-mass fraction from stars and intracluster gas as a function of total cluster mass . . . . .	110
2.15	MACHO mass fraction limits as a function of microlensing mass	118
2.16	Feynman diagrams of 1-loop radiative corrections to the Higgs mass . . . . .	136
2.17	Running (as a function of energy) of the Standard Model coupling constants with and without supersymmetry . . . . .	139
2.18	Neutralino-nucleon elastic scattering cross section predictions as a function of neutralino mass . . . . .	146
2.19	Comparison of mSUGRA predictions for the neutralino cross section and mass to an MSSM model with a light neutralino . .	150
2.20	Combined experimental constraints on the mSUGRA neutralino mass versus the ratio of the vacuum expectation values of the Higgs bosons . . . . .	151
3.1	Schematic of the Compact Muon Solenoid (CMS) detector at the Large Hadron Collider (LHC) . . . . .	157
3.2	SUSY exclusion limits on universal scalar- and gaugino-mass parameters in the context of mSUGRA . . . . .	161

3.3	Measurements of the cosmic-ray positron fraction, indicating an excess of high-energy positrons possibly due to dark-matter annihilations . . . . .	167
3.4	Measurements of the high-energy (absolute) flux of cosmic-ray positrons and electrons . . . . .	171
3.5	Measurements of the total $e^\pm$ cosmic-ray flux compared with a pulsar explanation of the high-energy tail . . . . .	175
3.6	Measurements of the cosmic-ray positron fraction compared to models in which pulsars are the source of the high-energy excess	176
3.7	Comparison of the EGRET and Fermi-LAT diffuse Galactic gamma-ray spectra . . . . .	180
3.8	Indirect evidence for a light WIMP from gamma-ray emission from the Galactic center . . . . .	184
3.9	Allowed regions of local halo density and halo-core radii for different halo profiles . . . . .	193
3.10	The Helm/Lewin-Smith nuclear form factor for a variety of target nuclei . . . . .	199
3.11	The recoil-energy spectrum expected from WIMPs scattering with Ge and Si targets . . . . .	202
3.12	Average deposited energy expected for various targets as a function of WIMP mass . . . . .	203
3.13	Underground muon intensity as a function of depth . . . . .	209
3.14	Underground flux of gamma rays as a function of gamma-ray energy . . . . .	212
3.15	Neutron flux for different neutron sources as a function of depth	218
3.16	The flux of high-energy neutrons versus neutron energy . . . . .	222
3.17	An example electron-recoil versus nuclear-recoil discrimination plot for a direct-detection experiment . . . . .	226
3.18	Annual-modulation fit to the DAMA/LIBRA residual event rate . . . . .	231
3.19	Spin-independent WIMP-nucleon cross sections allowed by the DAMA annual-modulation signal as a function of WIMP mass .	232
3.20	Differential event rate observed by the CoGeNT Ge diode at the Soudan Underground Laboratory in a 56 day exposure . . . . .	235
3.21	Joint allowed region for the DAMA annual modulation and CoGeNT excess . . . . .	238

4.1	Schematic of the CDMS ZIP-detector ionization electrodes and readout circuit . . . . .	249
4.2	Demonstration of ZIP-detector phonon-based pulse-shape discrimination . . . . .	258
4.3	Schematic of phonon detection with quasiparticle-trap-assisted transition-edge sensor . . . . .	260
4.4	Images and schematics of the CDMS II ZIP-detector phonon sensors . . . . .	261
4.5	Example of superconducting transition for ZIP-detector tungsten . . . . .	263
4.6	Schematic of the readout circuit used in the ZIP-detector phonon measurement . . . . .	264
4.7	Demonstration of the CDMS ionization-yield discrimination parameter as a function of recoil energy . . . . .	267
4.8	Compilation of ionization-yield measurements for nuclear recoils in Ge compared to Lindhard theory . . . . .	270
4.9	Demonstration of ZIP-detector $xy$ -position reconstruction . . . .	274
4.10	Demonstration of how the degeneracy in the ZIP-detector $xy$ -position reconstruction can be broken . . . . .	276
4.11	Comparison of electron-recoil ionization yield before and after position correction of the phonon energy . . . . .	277
4.12	Schematic of the Stanford Underground Facility's layout . . . .	281
4.13	Diagram of the passive and active shielding configuration employed at the CDMS shallow-site . . . . .	283
4.14	Detailed geometry of the shallow-site muon-veto counters . . . .	286
4.15	Schematic representation of the measurement used to map the muon-veto counters' light-collection efficiencies . . . . .	287
4.16	Light-collection maps for two of the shallow-site muon-veto counters . . . . .	289
4.17	Energy spectra for two shallow-site muon-veto channels . . . . .	290
4.18	Schematic representation of analog readout of muon-veto signals	293
4.19	Energy spectra for all 8 analog-readout muon-veto channels . . .	295
4.20	Photograph of the prototype circuit board used for analog readout of the muon veto during Run 21 . . . . .	297
4.21	Muon-like ionization-energy spectra for the Tower 1 detectors . .	299
4.22	Muon-veto bottom-side coincident nuclear recoils and muon-tagging efficiency . . . . .	302



4.23	Schematic of the CDMS shallow-site dilution refrigerator and cryostat . . . . .	306
4.24	Schematic of a ZIP-detector tower assembly . . . . .	308
5.1	Example phonon noise traces from an aberrant class of randomly triggered events . . . . .	322
5.2	Estimate of the hardware trigger efficiency for a representative Ge ZIP . . . . .	325
5.3	Nuclear-recoil band centroid and associated error envelope for a representative Ge ZIP . . . . .	326
5.4	Statistical uncertainties associated with the trigger-efficiency estimate for the Z2 3V data . . . . .	327
5.5	Time dependence of the phonon-energy software threshold for the Z4 3V data . . . . .	330
5.6	Estimate of the average efficiency of the phonon-energy software threshold for the Z4 3V data . . . . .	331
5.7	Demonstration of the $\chi^2$ data-quality cut for a typical detector .	339
5.8	Periods of unusually high trigger rates (event bursts) from the Z2 6V WIMP-search data . . . . .	340
5.9	Demonstration of the cut used to remove event bursts from the Z2 6V WIMP-search data . . . . .	341
5.10	Optimization of the event-burst cut levels for the Z2 6V data . .	342
5.11	Fits to $Q$ -outer noise distributions for a low- and high-energy bin of $Q$ -inner energy for the Z6 6V data . . . . .	344
5.12	The $Q$ -outer $2\sigma$ noise band as a function of $Q$ -inner energy for the Z6 6V data . . . . .	346
5.13	Estimate of the Z6 6V fiducial-volume cut's selection efficiency as a function of recoil energy . . . . .	348
5.14	Distribution of arrival times for muon-veto hits as a function of ZIP-detector recoil energy . . . . .	351
5.15	Bin-wise estimate of the ionization-yield mean and width for Z5 6V nuclear recoils . . . . .	353
5.16	Nuclear-recoil band fits for the viable low-threshold detectors . .	355
5.17	Estimate of the $2\sigma$ nuclear-recoil band's selection efficiency for Z5 6V data . . . . .	356
5.18	Characterization of the phonon- and ionization-energy noise distributions for the 3V data . . . . .	359

5.19	Two-dimensional preselection of the low-energy events used to reconstruct the 1.29 keV line from the Ge 6V data . . . . .	361
5.20	Fits to the 1.29 keV peak in the 6V WIMP-search data for the Ge detectors . . . . .	363
5.21	Fits to the 8.98 and 10.36 keV peaks in the 3V WIMP-search data for the Ge detectors . . . . .	365
5.22	Fits to the 66.7 keV peak in the 6V WIMP-search data for the Ge detectors . . . . .	366
5.23	Effect of recalibrating the electron-recoil energy scale on the detection efficiencies for the Z2 6V data . . . . .	368
5.24	Phonon resolution as a function of total phonon energy for the viable low-threshold Ge detectors . . . . .	370
5.25	Energy dependence of the viable low-threshold Ge detectors' recoil-energy resolutions . . . . .	372
5.26	Comparison of observed and simulated recoil-energy spectra for $^{252}\text{Cf}$ nuclear recoils . . . . .	375
5.27	Comparison of the Ge and Si subthreshold noise distributions to the observed WIMP-candidate event rates . . . . .	379
5.28	Combined detection efficiencies for the Z6 3V data . . . . .	381
6.1	Combined WIMP-candidate event rates for the Ge and Si detector ensembles . . . . .	385
6.2	Ionization yield versus recoil energy for the Z5 6V WIMP candidates . . . . .	386
6.3	Ionization yield versus recoil energy for the Z4 3V WIMP candidates . . . . .	388
6.4	Ionization yield versus recoil energy for the Z6 6V WIMP candidates . . . . .	390
6.5	Comparison of the serialization and averaging methods for combining WIMP-search data from different detectors . . . . .	397
6.6	Low-threshold WIMP-nucleon exclusion limits . . . . .	399
6.7	Low-threshold exclusion limits divided according to limit-setting detector . . . . .	400
6.8	Low-threshold exclusion limits versus Galactic escape velocity . . . . .	402
6.9	Spectral comparison of simulated and observed $^{252}\text{Cf}$ nuclear recoils for the Ge and Si detector ensembles . . . . .	404

6.10	$\chi^2$ statistic versus scaling factor for a one-parameter energy-scale correction . . . . .	405
6.11	Recoil-energy dependence of the best-fit scale factor for the one-parameter energy-scale correction . . . . .	407
6.12	$\chi^2$ statistic for a two-parameter energy-scale correction . . . . .	408
6.13	Spectral comparison of simulated and observed $^{252}\text{Cf}$ nuclear recoils following a two-parameter energy-scale correction . . . . .	409
6.14	Comparison of low-threshold limits with and without nuclear-recoil energy-scale correction . . . . .	410
6.15	Effect of energy-scale correction uncertainty on the low-threshold exclusion limits . . . . .	411
A.1	Diagram of the input-buffer stage of the shallow-site muon-veto analog-readout circuit . . . . .	418
A.2	Diagram of the integrating-amplifier stage of the shallow-site muon-veto analog-readout circuit . . . . .	419
A.3	Diagram of the output-driver stage of the shallow-site muon-veto analog-readout circuit . . . . .	420
A.4	Diagram of the voltage regulation for the shallow-site muon-veto analog-readout circuit . . . . .	421
A.5	Diagram of the input-buffer stage of the deep-site muon-veto front-end board . . . . .	422
A.6	Diagram of the integrating-amplifier stage of the deep-site muon-veto front-end board . . . . .	423
A.7	Diagram of the output-driver stage of the deep-site muon-veto front-end board . . . . .	424
A.8	Diagram of the voltage regulation for the deep-site muon-veto front-end board . . . . .	425
A.9	All layers of the deep-site muon-veto front-end board's PCB layout . . . . .	426
A.10	PCB layout for the top side of the deep-site muon-veto front-end board . . . . .	427
A.11	PCB layout for the bottom side of the deep-site muon-veto front-end board . . . . .	428
A.12	PCB layout for the power layer of the deep-site muon-veto front-end board . . . . .	429

A.13	PCB layout for the ground layer of the deep-site muon-veto front-end board . . . . .	430
A.14	PCB layout for the surface-mount footprints on the top side of the deep-site muon-veto front-end board . . . . .	431
A.15	Silk-screen pattern for the top side of the deep-site muon-veto front-end board . . . . .	432
A.16	Digital photograph of a fully assembled deep-site muon-veto front-end board . . . . .	433
B.1	Trigger-efficiency estimate for the Z2 6V data . . . . .	436
B.2	Nuclear-recoil band centroid and associated error envelope for the Z2 6V data . . . . .	436
B.3	Trigger-efficiency statistical uncertainties for the Z2 6V data . .	437
B.4	Trigger-efficiency estimate for the Z3 3V data . . . . .	437
B.5	Nuclear-recoil band centroid and associated error envelope for the Z3 3V data . . . . .	438
B.6	Trigger-efficiency statistical uncertainties for the Z3 3V data . .	438
B.7	Trigger-efficiency estimate for the Z3 6V data . . . . .	439
B.8	Nuclear-recoil band centroid and associated error envelope for the Z3 6V data . . . . .	439
B.9	Trigger-efficiency statistical uncertainties for the Z3 6V data . .	440
B.10	Trigger-efficiency estimate for the Z4 3V data . . . . .	440
B.11	Nuclear-recoil band centroid and associated error envelope for the Z4 3V data . . . . .	441
B.12	Trigger-efficiency statistical uncertainties for the Z4 3V data . .	441
B.13	Trigger-efficiency estimate for the Z4 6V data . . . . .	442
B.14	Nuclear-recoil band centroid and associated error envelope for the Z4 6V data . . . . .	442
B.15	Trigger-efficiency statistical uncertainties for the Z4 6V data . .	443
B.16	Trigger-efficiency estimate for the Z5 3V data . . . . .	443
B.17	Nuclear-recoil band centroid and associated error envelope for the Z5 3V data . . . . .	444
B.18	Trigger-efficiency statistical uncertainties for the Z5 3V data . .	444
B.19	Trigger-efficiency estimate for the Z5 6V data . . . . .	445
B.20	Nuclear-recoil band centroid and associated error envelope for the Z5 6V data . . . . .	445
B.21	Trigger-efficiency statistical uncertainties for the Z5 6V data . .	446

B.22	Trigger-efficiency estimate for the Z6 3V data . . . . .	446
B.23	Nuclear-recoil band centroid and associated error envelope for the Z6 3V data . . . . .	447
B.24	Trigger-efficiency statistical uncertainties for the Z6 3V data . .	447
B.25	Trigger-efficiency estimate for the Z6 6V data . . . . .	448
B.26	Nuclear-recoil band centroid and associated error envelope for the Z6 6V data . . . . .	448
B.27	Trigger-efficiency statistical uncertainties for the Z6 6V data . .	449
B.28	Time dependence of the phonon-energy software threshold for the Z1 3V data . . . . .	450
B.29	Time dependence of the phonon-energy software threshold for the Z1 6V data . . . . .	450
B.30	Time dependence of the phonon-energy software threshold for the Z2 3V data . . . . .	451
B.31	Estimate of the phonon-energy software threshold efficiency for the Z2 3V data . . . . .	451
B.32	Time dependence of the phonon-energy software threshold for the Z2 6V data . . . . .	452
B.33	Estimate of the phonon-energy software threshold efficiency for the Z2 6V data . . . . .	452
B.34	Time dependence of the phonon-energy software threshold for the Z3 3V data . . . . .	453
B.35	Estimate of the phonon-energy software threshold efficiency for the Z3 3V data . . . . .	453
B.36	Time dependence of the phonon-energy software threshold for the Z3 6V data . . . . .	454
B.37	Estimate of the phonon-energy software threshold efficiency for the Z3 6V data . . . . .	454
B.38	Time dependence of the phonon-energy software threshold for the Z4 6V data . . . . .	455
B.39	Estimate of the phonon-energy software threshold efficiency for the Z4 6V data . . . . .	455
B.40	Time dependence of the phonon-energy software threshold for the Z5 3V data . . . . .	456
B.41	Estimate of the phonon-energy software threshold efficiency for the Z5 3V data . . . . .	456

B.42	Time dependence of the phonon-energy software threshold for the Z5 6V data . . . . .	457
B.43	Estimate of the phonon-energy software threshold efficiency for the Z5 6V data . . . . .	457
B.44	Time dependence of the phonon-energy software threshold for the Z6 3V data . . . . .	458
B.45	Estimate of the phonon-energy software threshold efficiency for the Z6 3V data . . . . .	458
B.46	Time dependence of the phonon-energy software threshold for the Z6 6V data . . . . .	459
B.47	Estimate of the phonon-energy software threshold efficiency for the Z6 6V data . . . . .	459
C.1	Example fits to $Q$ -outer noise distributions for Z2 3V data . . .	462
C.2	The $Q$ -outer $2\sigma$ noise band for the Z2 3V data . . . . .	462
C.3	Estimate of the fiducial-volume cut's selection efficiency for the Z2 3V data . . . . .	463
C.4	Example fits to $Q$ -outer noise distributions for Z2 6V data . . .	463
C.5	The $Q$ -outer $2\sigma$ noise band for the Z2 6V data . . . . .	464
C.6	Estimate of the fiducial-volume cut's selection efficiency for the Z2 6V data . . . . .	464
C.7	Example fits to $Q$ -outer noise distributions for Z3 3V data . . .	465
C.8	The $Q$ -outer $2\sigma$ noise band for the Z3 3V data . . . . .	465
C.9	Estimate of the fiducial-volume cut's selection efficiency for the Z3 3V data . . . . .	466
C.10	Example fits to $Q$ -outer noise distributions for Z3 6V data . . .	466
C.11	The $Q$ -outer $2\sigma$ noise band for the Z3 6V data . . . . .	467
C.12	Estimate of the fiducial-volume cut's selection efficiency for the Z3 6V data . . . . .	467
C.13	Example fits to $Q$ -outer noise distributions for Z4 3V data . . .	468
C.14	The $Q$ -outer $2\sigma$ noise band for the Z4 3V data . . . . .	468
C.15	Estimate of the fiducial-volume cut's selection efficiency for the Z4 3V data . . . . .	469
C.16	Example fits to $Q$ -outer noise distributions for Z4 6V data . . .	469
C.17	The $Q$ -outer $2\sigma$ noise band for the Z4 6V data . . . . .	470
C.18	Estimate of the fiducial-volume cut's selection efficiency for the Z4 6V data . . . . .	470

C.19	Example fits to $Q$ -outer noise distributions for Z5 3V data . . .	471
C.20	The $Q$ -outer $2\sigma$ noise band for the Z5 3V data . . . . .	471
C.21	Estimate of the fiducial-volume cut's selection efficiency for the Z5 3V data . . . . .	472
C.22	Example fits to $Q$ -outer noise distributions for Z5 6V data . . .	472
C.23	The $Q$ -outer $2\sigma$ noise band for the Z5 6V data . . . . .	473
C.24	Estimate of the fiducial-volume cut's selection efficiency for the Z5 6V data . . . . .	473
C.25	Example fits to $Q$ -outer noise distributions for Z6 3V data . . .	474
C.26	The $Q$ -outer $2\sigma$ noise band for the Z6 3V data . . . . .	474
C.27	Estimate of the fiducial-volume cut's selection efficiency for the Z6 3V data . . . . .	475
C.28	Binwise estimate of the ionization-yield mean and width for Z2 3V nuclear recoils . . . . .	476
C.29	Estimate of the $2\sigma$ nuclear-recoil band's selection efficiency for Z2 3V data . . . . .	476
C.30	Binwise estimate of the ionization-yield mean and width for Z2 6V nuclear recoils . . . . .	477
C.31	Estimate of the $2\sigma$ nuclear-recoil band's selection efficiency for Z2 6V data . . . . .	477
C.32	Binwise estimate of the ionization-yield mean and width for Z3 3V nuclear recoils . . . . .	478
C.33	Estimate of the $2\sigma$ nuclear-recoil band's selection efficiency for Z3 3V data . . . . .	478
C.34	Binwise estimate of the ionization-yield mean and width for Z3 6V nuclear recoils . . . . .	479
C.35	Estimate of the $2\sigma$ nuclear-recoil band's selection efficiency for Z3 6V data . . . . .	479
C.36	Binwise estimate of the ionization-yield mean and width for Z4 3V nuclear recoils . . . . .	480
C.37	Estimate of the $2\sigma$ nuclear-recoil band's selection efficiency for Z4 3V data . . . . .	480
C.38	Binwise estimate of the ionization-yield mean and width for Z4 6V nuclear recoils . . . . .	481
C.39	Estimate of the $2\sigma$ nuclear-recoil band's selection efficiency for Z4 6V data . . . . .	481

C.40	Binwise estimate of the ionization-yield mean and width for Z5 3V nuclear recoils . . . . .	482
C.41	Estimate of the $2\sigma$ nuclear-recoil band's selection efficiency for Z5 3V data . . . . .	482
C.42	Binwise estimate of the ionization-yield mean and width for Z6 3V nuclear recoils . . . . .	483
C.43	Estimate of the $2\sigma$ nuclear-recoil band's selection efficiency for Z6 3V data . . . . .	483
C.44	Binwise estimate of the ionization-yield mean and width for Z6 6V nuclear recoils . . . . .	484
C.45	Estimate of the $2\sigma$ nuclear-recoil band's selection efficiency for Z6 6V data . . . . .	484
D.1	Characterization of the phonon- and ionization-energy noise dis- tributions for the 6V data . . . . .	486
D.2	Preselection of 1.29 keV events for the 3V data . . . . .	486
D.3	Fits to the 1.29 keV peak in the 3V WIMP-search data for the Ge detectors . . . . .	487
D.4	Fits to the 8.98 and 10.36 keV peaks in the 6V WIMP-search data for the Ge detectors . . . . .	488
D.5	Fits to the 66.7 keV peak in the 3V WIMP-search data for the Ge detectors . . . . .	489
E.1	Spectrum of neutron energies emitted by $^{252}\text{Cf}$ . . . . .	494
E.2	Spectrum of $^{252}\text{Cf}$ neutron energies incident upon the ZIP de- tectors following simulated transport through the shielding . . .	495
E.3	Elastic scattering cross sections as a function of incident neutron energy for Ge and Si targets . . . . .	497
E.4	Angular probability for neutrons to scatter from $^{74}\text{Ge}$ . . . . .	498
E.5	Angular probability for neutrons to scatter from $^{28}\text{Si}$ . . . . .	499
E.6	Numerically calculated $^{252}\text{Cf}$ differential event rates . . . . .	500
E.7	Comparison of observed and numerically calculated $^{252}\text{Cf}$ recoil spectra for the Si detector ensemble . . . . .	501
E.8	Comparison of observed and numerically calculated $^{252}\text{Cf}$ recoil spectra for the Ge detector ensemble . . . . .	502
F.1	Combined detection efficiencies for the Z2 3V data . . . . .	504



F.2	Combined detection efficiencies for the Z2 6V data . . . . .	504
F.3	Combined detection efficiencies for the Z3 3V data . . . . .	505
F.4	Combined detection efficiencies for the Z3 6V data . . . . .	505
F.5	Combined detection efficiencies for the Z4 3V data . . . . .	506
F.6	Combined detection efficiencies for the Z4 6V data . . . . .	506
F.7	Combined detection efficiencies for the Z5 3V data . . . . .	507
F.8	Combined detection efficiencies for the Z5 6V data . . . . .	507
F.9	Combined detection efficiencies for the Z6 6V data . . . . .	508
F.10	Average total detection efficiency for the Ge detector ensemble .	508
F.11	Average total detection efficiency for the Si detector ensemble .	509
G.1	Ionization yield versus recoil energy for the Z2 3V WIMP can- didates . . . . .	511
G.2	Ionization yield versus recoil energy for the Z2 6V WIMP can- didates . . . . .	512
G.3	Ionization yield versus recoil energy for the Z3 3V WIMP can- didates . . . . .	512
G.4	Ionization yield versus recoil energy for the Z3 6V WIMP can- didates . . . . .	513
G.5	Ionization yield versus recoil energy for the Z4 6V WIMP can- didates . . . . .	513
G.6	Ionization yield versus recoil energy for the Z5 3V WIMP can- didates . . . . .	514
G.7	Ionization yield versus recoil energy for the Z6 3V WIMP can- didates . . . . .	514

# List of Tables

1.1	Fundamental fermions of the Standard Model of particle physics	3
1.2	Cosmological parameters of the best-fit $\Lambda$ CDM model to the WMAP 7-year CMB power spectrum . . . . .	31
2.1	Mass-to-light ratios ( $\Upsilon$ ) and inferred matter densities for several classes of differently sized objects . . . . .	53
3.1	Values of the free parameters assumed for calculating direct-detection scattering rates . . . . .	194
3.2	The discrimination techniques and optimal sensitivities for several of the world's leading direct-detection experiments . . . . .	228
4.1	CDMS II Tower 1 ZIP-detector properties . . . . .	245
5.1	Summary of the Run 21 data samples . . . . .	321
5.2	Summary of the phonon-energy hardware and software thresholds . . . . .	333
5.3	Summary of the veto-cut parameters and detection efficiency . .	352
5.4	Results of an energy-scale correction for the viable low-threshold Ge detectors . . . . .	367
5.5	Summary of fits to the ionization resolution's energy dependence for the viable low-threshold Ge detectors . . . . .	371
5.6	Summary of the recoil-energy noise resolutions for the viable low-threshold detectors . . . . .	378
6.1	Summary of the WIMP-search exposure and total number of observed candidate events for each viable low-threshold detector and bias-voltage run . . . . .	384

6.2	Tally of the percentages of WIMP candidates due to various sources of background . . . . .	392
A.1	List of capacitors for the deep-site muon-veto front-end board .	416
A.2	Parts list (other than capacitors) for the deep-site muon-veto front-end board . . . . .	417

# Chapter 1

## Introduction

Throughout history the scientific community has endeavored to unravel the mysteries of the Universe, often with no better motivation than simple human curiosity. The focus of this thesis is one such pursuit of the unknown, an attempt to experimentally observe a new form of matter. Although motivated by a healthy dose of curiosity, the true impetus behind this investigation is nearly a century’s worth of compelling scientific evidence for the existence of an exotic form of matter in the Universe.

In this chapter I briefly review the fundamental constituents of the known forms of matter, as well as the Universe in which they exist. The former is well described by the Standard Model (SM) of particle physics, while a consistent picture of the latter has come to be termed the Standard Cosmology [1]. The theoretical framework and general features of the Standard Cosmology are presented, followed by a detailed discussion of several types of supporting astrophysical data. When interpreted in the context of the Standard Cosmology, the observational data make a compelling case for the existence of a non-SM form of matter known as “dark matter” [2].

## 1.1 The Standard Model

Most ordinary (stable) matter is in the form of interstellar and intergalactic hydrogen and helium (often referred to simply as “gas”), while common terrestrial matter and star-stuff are far less abundant. At the heart of all ordinary matter are bound states of a surprisingly small number of subatomic particles: electrons and up- and down-type quarks. Furthermore, the mass of conventional matter is principally in the form of quark bound states called nucleons, with an insignificant ( $<0.03\%$ ) contribution from electrons. This is remarkable; nearly all the (known) matter in the Universe can be attributed to two quark types and the gluon-mediated interactions that bind them into nucleons. Lets review the components of the SM theory that so successfully describes these particles and (by extension)  $>99.97\%$  of the known (stable) matter in the Universe.

The SM is a quantum field theory that describes the properties of a relatively small number of fundamental fermions and their gauge-boson-mediated interactions. The composite particles that are the building blocks of ordinary matter, as well as higher order resonant states (produced during particulate collisions either in the laboratory or in nature) are understood as bound states of these fermions. The bound state configurations, quantum numbers, lifetimes, and probabilities to decay into other states can be predicted according to their SM boson-mediated interactions.

The SM’s fundamental fermions are summarized in Table 1.1 and occur in three families, or generations, of leptons and quarks. The first generation fermions include the stable particles from which all known matter is comprised, while the second and third generations consist of more massive (and generally transient) particles.

The quarks interact with one another via the gluon-mediated strong force. The strong-force part of the SM, known as quantum chromodynamics (QCD), is an  $SU(3)$  gauge theory. Quarks therefore carry one of three possible strong-force charges, called color. The symmetry of the gauge group implies that color is carried from one quark (or antiquark) to another via eight types of massless

### 1.1. THE STANDARD MODEL

Table 1.1: The fundamental fermions of the Standard Model (SM) of particle physics. Ordinary (stable) matter occurs primarily in the form of baryonic bound states of the first-generation quarks (nucleons). For more details regarding SM fermions and the bosons that mediate their interactions see [3].

Generation	I	II	III
Quarks	Up ( $u$ )	Charm ( $c$ )	Top ( $t$ )
	Down ( $d$ )	Strange ( $s$ )	Bottom ( $b$ )
Leptons	$e$ -Neutrino ( $\nu_e$ )	$\mu$ -Neutrino ( $\nu_\mu$ )	$\tau$ -Neutrino ( $\nu_\tau$ )
	Electron ( $e$ )	Muon ( $\mu$ )	Tau ( $\tau$ )

gluon (one less than nine due to unitarity). A phenomenon known as color confinement prevents objects with a net color from existing in nature, with the (probable) exception of the few moments following the Big Bang during which the Universe was dense and hot enough to disfavor color-neutral bound states. At the present epoch, all colored objects are bound into color-neutral composite particles known as hadrons, which consist of either a quark and an antiquark (mesons), three quarks (baryons), or three antiquarks (antibaryons). Hadronic states of the heavier quark generations tend to decay into leptons, photons, and lighter first-generation quark states. First-generation quarks are found in nature in the form of the lightest baryonic states: the proton and the neutron (nucleons). Since the rest mass of ordinary matter is almost entirely accounted for by the mass of these baryons, conventional matter is often referred to as baryonic matter.

The total symmetry of the SM can be understood as a product of gauge theories:  $SU(3) \otimes SU(2) \otimes U(1)$ , where  $SU(3)$  is the QCD part of the theory, and  $SU(2) \otimes U(1)$  is the electroweak part known as the Glashow-Salam-Weinberg theory (GSW). GSW theory describes additional fermion-fermion interactions with four massless bosons. In nature the symmetry of the electroweak gauge

## CHAPTER 1. INTRODUCTION

group is spontaneously broken (electroweak symmetry breaking), causing the electroweak force to split into two lower-energy effective forces: the weak and the electromagnetic forces. The resulting weak-force carriers are the massive  $W^\pm$  and  $Z$  bosons, while the electromagnetic force is propagated via the massless photon.

It is (strongly) speculated that electroweak symmetry breaking is driven by the Higgs mechanism and gives rise to an additional boson called the Higgs [4, 5, 6]. The Higgs is a sort of Holy Grail of particle physics as it is the only (fundamental) particle of the theory that has yet to be observed. It is unique in that it is the SM's only fundamental scalar and is the mechanism by which the theory's fermions acquire mass. The existence of a Higgs boson also explains why the photon is massless while the weak bosons are massive. It is possible that the Higgs has been generated in colliders but overlooked due to the low production rate relative to competing background processes. If true, its discovery at the Large Hadron Collider (LHC) should be imminent [7, 8].

It is worth emphasizing that the SM contains stable fermions that only interact via the weak force. When first postulated, the three neutrino flavors were believed to be massless particles. Large-volume neutrino observatories have since successfully confirmed the existence of neutrino mass through measurements of neutrino mixing [9]. In other words, the neutrino flavor eigenstates listed in Table 1.1 do not diagonalize the neutrinos' time-evolution Hamiltonian. Rather, each flavor oscillates into the other flavors according to a time-dependent probability and is related to the time-stable (mass) eigenstates through a mixing matrix whose off-diagonal elements have been measured to be nonzero. Weakly interacting and massive, SM neutrinos are therefore a type of dark matter.<sup>1</sup> However, as will be discussed in more detail in Section 2.4.2, they account for only a trivial fraction of the nonbaryonic dark matter in the Universe. Nonetheless, neutrino mass establishes an interesting precedent for the existence of stable weakly-interacting particles with nonzero mass.

---

<sup>1</sup>Although SM neutrinos are technically massless, neutrino mass can be included in the SM with a relatively simple extension [10, 11]. It is therefore (often) not thought of as physics beyond the SM.

## 1.2. THE STANDARD COSMOLOGY

### 1.2 The Standard Cosmology

In relatively small quantities ordinary matter arranges itself according to the subatomic physics of the SM, with little regard to gravitational effects. As greater and greater quantities of matter are brought together, however, the gravitational force becomes increasingly important and eventually causes matter to clump, a process that occurs at almost all astrophysical length scales. Physical cosmology is the branch of physics that attempts to understand the large-scale structure of matter (and radiation) in the Universe and how it has evolved from its origins to present day, as well as what its eventual fate will be. Theoretical cosmological models are tested against astrophysical observations in the hope that a progressively more accurate picture of the Universe will emerge. As I will discuss shortly, although the Universe's baryonic matter and SM radiations play critical roles in our current understanding, they represent minority components of the Universe, acting like tracers for far more abundant but unknown forms of exotic matter and energy.

Einstein's formulation of general relativity marked the beginning of modern cosmology by establishing a robust framework with which to understand the dynamical evolution of spacetime, matter, and radiation [12]. Since then, the fundamental ingredients of any serious cosmological model have included: 1) Einstein's equation which relates the geometry of the Universe with its matter and energy content; 2) a metric which describes the symmetries of the model; and 3) an equation of state which specifies the physical properties of the model's matter and energy content. Dozens of cosmological models have been constructed according to this recipe. Though instructive, nearly all have been systematically eliminated as inconsistent with the observed Universe and therefore unphysical.

In this section I review the construction of and evidence supporting the most successful cosmological model to date. The Lambda-Cold Dark Matter model ( $\Lambda$ CDM) [13, 14] of Big Bang cosmology has become known as the Standard Cosmology because it is the simplest model that is in general agreement with astrophysical observations [15, 16]. By incorporating dark matter and a



cosmological constant into the Big Bang scenario, it (generally) succeeds in describing the Universe’s thermal history, light-element abundances, accelerating expansion, electromagnetic background radiation, and large-scale structure.

### 1.2.1 Theoretical Framework

The first ingredient of the Standard Cosmology is the Einstein gravitational field equation of general relativity, which can be succinctly expressed as a second-order differential equation of tensors [12]:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}\mathcal{S} = -\frac{8\pi G_N}{c^4}\mathcal{T}_{\mu\nu} + \Lambda g_{\mu\nu}. \quad (1.2.1)$$

$R_{\mu\nu}$  and  $\mathcal{S}$  are the Ricci tensor and scalar (obtained through contraction of the Riemann curvature tensor, see [17] for example), while  $g_{\mu\nu}$  is the metric tensor in which the properties of spacetime are encoded. The left-hand side of the Einstein equation describes the evolution of  $g_{\mu\nu}$  and therefore spacetime itself.  $G_N$  is Newton’s constant, and  $\mathcal{T}_{\mu\nu}$  is the energy-momentum tensor.  $\Lambda$  is an allowed constant term known as the cosmological constant. As will be discussed in more detail below, despite its somewhat sorted history, the cosmological constant is an accepted (and unknown) ingredient of the Standard Cosmology. The right-hand side of Equation 1.2.1 can be thought of as a collection of source terms; spacetime evolves according to the energy content on the right-hand side. For a more complete discussion of Einstein’s equation see [17] or [18].

The second ingredient is effectively a well motivated ansatz for the space-time metric  $g_{\mu\nu}$ . The simplifying principles of spatial homogeneity and isotropy are combined to yield the most general form for  $g_{\mu\nu}$  that respects the implied spacetime symmetries, regardless of general relativity theory. Spatial homogeneity is a form of the Copernican principle in which we admit that our particular position within the Universe is not specially distinguished in any way, and nor are any other spatial locations. Spatial isotropy (essentially) imposes spherical symmetry about any given observational position. Spatial homogeneity and isotropy are closely connected concepts for which there is

## 1.2. THE STANDARD COSMOLOGY

(so far) no evidence to suggest they are not valid principles of symmetry in the Universe when it is viewed on a suitably large scale. While it is difficult to measure homogeneity, spatial isotropy is strongly supported by measurements of the temperature of the cosmic background radiation, which is the same to one part in  $10^4$  everywhere in the sky (see Section 1.3.2 for further details). The resulting Robertson-Walker metric (named for Robertson [19] and Walker [20] for being the first to derive it) is traditionally expressed as the line element

$$ds^2 = -c^2 dt^2 + a(t)^2 \left( \frac{dr^2}{1 - k_c r^2} + r^2 (d\theta^2 + \sin^2 \theta d\phi^2) \right). \quad (1.2.2)$$

Equation 1.2.2 represents the squared distance between two infinitesimally separated points in a four-dimensional spacetime that includes a single time dimension  $t$  and three spatial dimensions  $r$ ,  $\theta$  and  $\phi$ . The principles of homogeneity and isotropy have determined  $g_{\mu\nu}$  up to three possibilities ( $k_c = -1$ ,  $0$  or  $1$ ) and an arbitrary (positive) function  $a(t)$ . For  $k_c = 0$  the spatial part is easily recognized as the line element for a flat three-dimensional space (expressed in spherical coordinates), where  $r$  is a positive number that runs from  $0$  to  $\infty$ , and  $\theta$  and  $\phi$  are the usual azimuthal and polar angles, respectively. The spatial parts for  $k_c = 1$  and  $k_c = -1$  represent the analogous spherical coordinates for closed and open spaces, respectively. In the closed case, the coordinate  $r$  is cyclic and runs from  $0$  to  $2\pi$  as well. The role of the function  $a(t)$  can easily be understood if one considers two co-moving spatial points at some initial time  $t_1$  separated by a distance  $d_1$ . At an arbitrary time later,  $t = t_1 + \delta t$ , the distance between the two points is given by  $d(t) = a(\delta t) d_1$ .  $a(t)$  is therefore a time-dependent scale factor.

We can learn more about the evolution of the scale factor by substituting the Robertson-Walker metric into Equation 1.2.1 and applying the same principles of homogeneity and isotropy to the energy-momentum tensor. One of the resulting equations is (see [18] for a full derivation)

$$H^2(t) = \left( \frac{\dot{a}}{a} \right)^2 = \frac{8\pi G_N}{3} \rho_{\text{total}} - \frac{k_c c^2}{a^2}, \quad (1.2.3)$$

## CHAPTER 1. INTRODUCTION

where  $\rho_{\text{total}}$  represents the average energy density of the Universe (including contributions from the cosmological constant as well as from all forms of matter and radiation).  $H(t)$  is called the Hubble parameter (defined as the ratio of  $\dot{a}(t)$  to  $a(t)$ ) and is a measure of the rate at which spacetime is expanding or contracting. Equation 1.2.3 is typically referred to as the Friedmann equation despite differing from the forms Friedmann originally considered in [21] and [22].<sup>2</sup>

It is interesting to note that there is a particular value of  $\rho_{\text{total}}$  that results in a flat ( $k_c = 0$ ) Universe for all values of the Hubble parameter at all times; the “critical density”

$$\rho_c \equiv \frac{3H^2}{8\pi G_N}. \quad (1.2.4)$$

It is customary to rewrite the Friedmann equation in terms of the critical density by defining the parameter  $\Omega_0 \equiv \rho_{\text{total}}/\rho_c$ , yielding

$$\Omega_0 - 1 = \frac{k_c c^2}{H^2 a^2}. \quad (1.2.5)$$

This form is instructive as it allows us to draw the following conclusions regarding the relationship between the Universe’s energy content and its spatial curvature: 1)if the energy density exceeds the critical density ( $\Omega_0 > 1$ ), the spatial curvature is closed because  $k_c$  is strictly positive and the  $k_c = 1$  metric applies; 2)if the energy density equals the critical density ( $\Omega_0 = 1$ ), the Universe is flat because  $k_c = 0$ ; and 3)if the energy density is less than the critical density ( $\Omega_0 < 1$ ), the spatial curvature is open because  $k_c$  is strictly negative and thus the  $k_c = -1$  metric is the correct solution.

To see how the energy content of the Universe evolves according to the Friedmann equation, it is common practice to split  $\Omega_0$  into its constituents (matter, radiation and cosmological constant) and express Equation 1.2.5 in terms of the redshift parameter  $z$ . For radiation emitted far from our observation point here on Earth, the redshift parameter relates the observed wavelength  $\lambda_{\text{obs}}$  to

---

<sup>2</sup>Both Equation 1.2.2 and 1.2.3 are commonly attributed to Alexander Friedmann, Georges Lemaître, Howard Percy Robertson and Arthur Geoffrey Walker (FLRW, FRW, RW or FL metric and equation) in deference to the significant contributions each made to the development of the Standard Cosmological model [1].

## 1.2. THE STANDARD COSMOLOGY

the emitted wavelength  $\lambda_{\text{emit}}$  and provides a measure of the distance to the radiation source. The difference in the scale factor  $a(t)$  between the source and observation points can also be expressed in terms of  $z$ :

$$z \equiv \frac{\lambda_{\text{obs}}}{\lambda_{\text{emit}}} - 1 = \frac{a(t_{\text{obs}})}{a(t_{\text{emit}})} - 1. \quad (1.2.6)$$

Following the derivation in [23], the Friedmann equation can be rewritten as

$$\frac{H^2(z)}{H_0^2} = [\Omega_\Lambda + \Omega_K(1+z)^2 + \Omega_{\mathcal{M}}(1+z)^3 + \Omega_R(1+z)^4], \quad (1.2.7)$$

where  $\Omega_{\mathcal{M}}$ ,  $\Omega_R$  and  $\Omega_\Lambda$  refer respectively to the present-day matter, radiation, and cosmological-constant (fractional) densities and sum to  $\Omega_0$ .  $\Omega_K = -k_c c^2 / a_0^2 H_0^2$  is simply a convenient form for the curvature-sensitive part of the equation.  $H_0$  is the present day value of the Hubble parameter (see Figure 1.1) and is typically measured in km per second per megaparsec ( $\text{km s}^{-1} \text{Mpc}^{-1}$ ). To distinguish it from  $H(t)$ ,  $H_0$  is usually referred to as the Hubble *constant*.

Almost by sleight of hand, the third and final ingredient necessary for a complete cosmological model has been introduced; to obtain Equation 1.2.7 the equations of state for the model's energy content have been applied. An equation of state is typically expressed as the dimensionless ratio of pressure,  $p$ , to energy density,  $\rho$ ;  $w = p/\rho$ . The energy density evolves as a function of the scale factor and  $w$ :

$$\rho \propto a^{-3(1+w)}. \quad (1.2.8)$$

For nonrelativistic matter (*e.g.*, baryons in stars and cosmic gas)  $w = 0$  and  $\rho_{\mathcal{M}} \propto a^{-3}$ , leading to the  $z^3$  dependence of the matter term in Equation 1.2.7; the density of matter dilutes (increases) in direct proportion to the expanding (contracting) volume it occupies. This makes intuitive sense; if the amount of matter is held constant, its density should scale according to the volume. The equation of state for ultra-relativistic matter (*e.g.*, photons and neutrinos) is  $w = 1/3$ . Since  $\rho_R \propto a^{-4}$ ,  $\Omega_R$  evolves with an extra factor of  $z$  in Equation 1.2.7; radiation not only dilutes (increases) proportional to the volume expansion (contraction), but its wavelength redshifts (blueshifts) as well.

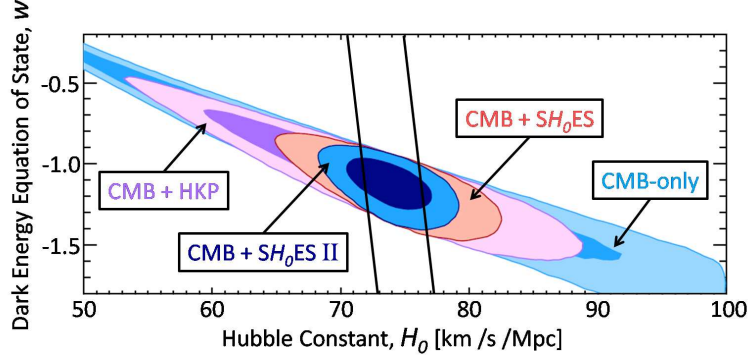


Figure 1.1: Combined measurement of the cosmological constant’s equation of state ( $w$ ) and the present day value of the Hubble parameter ( $H_0$ ). An analysis of the WMAP 7-year cosmic microwave background data [24] (see Section 1.3.2) provides a degenerate constraint (outermost contours) that is improved by the measurement of  $H_0$  (nearly vertical solid lines),  $73.8 \pm 2.4 \text{ km s}^{-1} \text{ Mpc}^{-1}$ , by the  $SH_0ES$  II program [25], resulting in the two innermost contours and corresponding to  $w = -1.08 \pm 0.10$ . Also shown are similar contours derived from the previous  $SH_0ES$  result [26] and the Hubble Key Project (HKP) [27]. The 68.3% and 95.4% confidence-level regions are given for each combination as similarly colored inner and outer contours, respectively. Figure adapted from [25].

Equation 1.2.7 assumes that the cosmological constant is true to its name and has a  $w = -1$  equation of state. As demonstrated in Figure 1.1, current astrophysical evidence is consistent with this assumption. The cosmological constant’s defining characteristic is that its density either does not vary with time or varies so slowly that it appears to be constant. Often referred to as zero-point or vacuum energy, it has come to be known colloquially as “dark energy” [28], and measuring its equation of state as precisely as possible is an active field of research [29]. Evidence that  $w$  is not exactly  $-1$  or that it varies slowly with time could provide a clue to understanding the (as yet) unknown mechanism behind the present-day nonzero value of  $\Omega_\Lambda$ . One possibility known as “quintessence” posits that the dark energy is a slowly evolving, self-interacting scalar field for which  $w \in (-1, 0)$  [30]. Current constraints on the time varying component of  $w$  are discussed further in Sections 1.3.3 and 1.3.5.

## 1.2. THE STANDARD COSMOLOGY

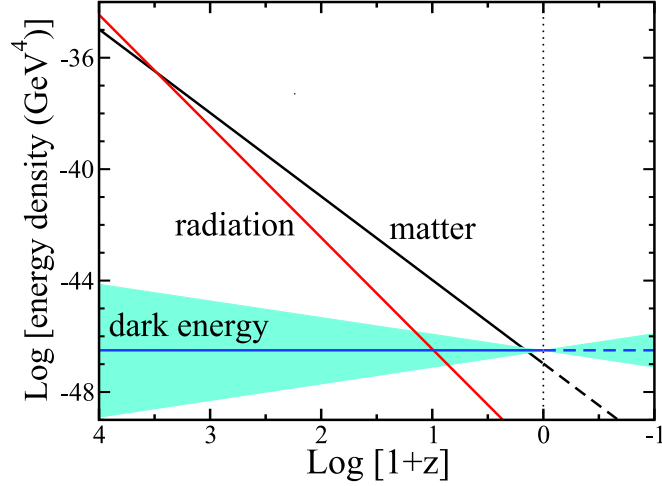


Figure 1.2: Radiation, matter and dark energy densities as a function of redshift, illustrating how the Universe is first radiation dominated (large redshifts to the left), then matter dominated, and finally dark-energy dominated (slightly positive and negative redshifts to the right). The dark-energy density is assumed to be constant ( $w = -1$ ) with a possible 20% uncertainty indicated by the shaded region. Figure taken from [31].

### 1.2.2 Thermal History

The differing  $z$  dependencies in Equation 1.2.7 for matter, radiation and the cosmological constant provide a method for disentangling their respective contributions to  $\Omega_0$  through astrophysical observations at different redshifts. As is illustrated in Figure 1.2, the history of the Universe therefore divides into three distinct epochs during which a different component of  $\Omega_0$  dominates the evolution of the scale factor.

## Radiation Dominated

Radiation (*e.g.*, photons and neutrinos) dominated the evolution of the scale factor in the earliest moments following the Big Bang<sup>3</sup> ( $z \gtrsim 3000$ ), and its strictly positive equation of state results in a Universe that expanded more and more slowly as a function of time ( $a \propto t^{1/2}$ , therefore  $\ddot{a} < 0$ ). The exact nature of the energy density directly following the Big Bang is not well understood, but is believed to have been in the form of an extremely dense, relativistic plasma in which the gravitational, strong, and electroweak forces had equal (or at least similar) strengths. A proper description of this single, ultra-high-energy force is the goal of grand unified theories (*e.g.*,  $SO(10)$  [34] or trinification [35]). As the Universe cooled the forces separated and following electroweak symmetry breaking the energy density became well described by the SM. Furthermore, baryogenesis [36, 37] created a matter-antimatter asymmetry that provided nucleons (instead of antinucleons) for the formation of hydrogen and other light elements. During this first period several significant events occurred (all three of which are discussed in more detail in Section 1.3):

- *Neutrino Decoupling:* Although large-volume neutrino observatories have successfully measured and confirmed the existence of neutrino mass [9], since neutrino energies were much greater than their rest masses in the early Universe, they were effectively a form of massless radiation. Neutrinos were in thermal equilibrium up to approximately one tenth of a second following the Big Bang, at which time the rate of their interactions with the other weakly-interacting matter in the thermal bath dropped below the rate at which the scale factor was expanding. The temperature of the Universe at this time was  $\sim 3.5 \times 10^{10}$  K, and  $kT$  was therefore  $\sim 3$  MeV [38]. A significant consequence of neutrino decoupling is that

---

<sup>3</sup>Part of the Standard Cosmology is the popular conjecture that the Universe underwent an inflationary period [32] of exponential expansion during the first fraction of a second following the Big Bang, successfully explaining the simultaneous size, isotropy, homogeneity and flatness of the Universe without violating causality. If true, the Universe was temporarily matter dominated during the inflationary expansion, becoming once again radiation dominated through the reheating caused by the decay of the inflaton [33].

## 1.2. THE STANDARD COSMOLOGY

the weak processes that maintain thermal equilibrium between protons and neutrons quickly turned off as the decoupled neutrinos continued to cool ( $T_\nu \propto 1/a$ ). Approximately one second later, the neutron-to-proton ratio became fixed (save for a gradual decrease due to the  $\sim 15$  minute neutron half-life) and played a critical role in the next stage.

- *Nucleosynthesis*: Light nuclei began to form once the temperature of the Universe cooled to less than 2.23 MeV, or  $\sim 2.6 \times 10^{10}$  K, below which the average nucleon energy is less than the deuteron binding energy and the crucial reaction



can occur. However, due to the large number density of the photon background, it was not until the universe cooled to  $\sim 1$  MeV (or  $\sim 1 \times 10^9$  K) that the number of deuteron-dissociating photons fell below the number of nucleons and Equation 1.2.9 could produce a stably increasing deuteron density [38]. Most nucleosynthesis occurred approximately 100 s after the Big Bang, and the abundances of light nuclei (*e.g.*, D, He and Li) eventually stabilized, or “froze-out” after roughly half an hour (right side of Figure 1.3). These abundances place strong constraints on the baryonic contribution to  $\Omega_{\mathcal{M}}$  (see Section 1.3.1 for further details).

- *Recombination*: As the Universe cooled to below the electron binding energy of hydrogen (13.6 eV), neutral hydrogen began to form out of the plasma of electrons and protons. The rate of photodissociation was finally outpaced by the expansion at a characteristic temperature of about 0.3 eV, resulting in the recombination of protons and electrons into neutral hydrogen (and traces of other light elements) several hundred thousand years after the Big Bang [39]. Prior to recombination the Universe was effectively opaque to electromagnetic radiation due to Thomson scattering of the photon background by free electrons. Following recombination the free-electron density dropped dramatically, causing an equally dramatic increase in the photon mean free path and resulting in a Universe that is (now) transparent to light. Recombination is often referred



to as the surface of last scattering since the photons that emerge from this event traverse the Universe largely unmolested. Due to the relentless expansion of spacetime, the surface of last scattering is visible today as a uniform glow of microwave photons with a characteristic temperature of  $\sim 23.5$  meV, or roughly 2.73 K, and is referred to as the Cosmic Microwave Background (CMB) radiation. Furthermore, CMB fluctuations ( $\Delta T/T$ ) on the order of one part in 100,000 have been measured and provide the earliest glimpse of matter-density fluctuations in the Universe (see Section 1.3.2 for further details).

## Matter Dominated

Matter (*e.g.*, baryons and dark matter) dominated the evolution of the scale factor for redshifts between roughly 3000 and 0.5. The Universe expanded more rapidly than during the radiation-dominated epoch, with  $a \propto t^{2/3}$ . The density perturbations from the surface of last scattering continued to grow, eventually condensing into stars and galaxies. Between recombination and the formation of the first stars—a prolonged period of a few hundred million years commonly referred to as the “dark ages”—the only significant source of light was the (now diffuse) CMB radiation. The first stars, galaxies and quasars to form radiated energetically enough to initiate a period of neutral hydrogen reionization that lasted for hundreds of millions of years ( $z_{\text{reion}} \simeq 10$ ). By this time, the expansion had diluted the distribution of matter sufficiently for the Universe to remain largely transparent to light despite the reionized hydrogen. However, reionization caused an  $\sim 10\%$  opacity that can be seen in the pattern of CMB fluctuations we observe today. Supernovae that occurred late in the matter-dominated epoch provide some of the most convincing evidence that the expansion of the Universe deviates from Hubble’s law [40], in which a galaxy’s redshift is linearly proportional to its distance (see Section 1.3.3 for further details). In fact, the present-day expansion is accelerating, indicating a need for a nonzero cosmological constant and leading to the following epoch.

### 1.3. OBSERVATIONAL EVIDENCE

#### Dark Energy Dominated

The evolution of the scale factor appears to have transitioned into a dark-energy-dominated era at  $z \simeq 0.5$ . During this phase the expansion accelerates, eventually yielding to an  $a \propto e^{Ht}$  behavior. If the dark energy's equation of state is truly constant ( $\dot{w} = 0$ ), the Universe will continue expanding in this manner for time eternal. The evolution of gravitationally bound structures (*e.g.*, galaxies) will become increasingly more complicated and nonlinear, while unbound structures (*e.g.*, galaxy superclusters) will gradually disperse until all the matter of the Universe is (effectively) isolated into highly evolved, causally disconnected island galaxies with vast gulfs of empty space in between. Observations of the distribution of matter at dark-energy-dominated redshifts (see Section 1.3.4 for further details) compliment observations of light streaming from the previous two epochs, helping to unravel the composition of  $\Omega_0$ , and suggesting that the Universe is a surprisingly dark place.

## 1.3 Observational Evidence

In Section 1.2.1 the three basic ingredients of the Standard Cosmology were combined to yield a prescription for the evolution of the energy content of the Universe (Equation 1.2.7) and how it relates to and affects the geometry (and evolution) of spacetime. An assumption regarding the model's energy content was made, and it was segregated into three pieces: radiation ( $\Omega_R$ ), matter ( $\Omega_M$ ) and a cosmological constant ( $\Omega_\Lambda$ ). This division is well motivated by observational evidence and delineates even further upon closer inspection. Perhaps the most remarkable aspect of the Standard Cosmology is that  $\sim 95\%$  of its energy content is of an unknown and dark composition. In this section I review some of the astrophysical measurements supporting the Standard Cosmology and its somewhat peculiar energy budget, with emphasis on its so-called dark sector.

### 1.3.1 Big Bang Nucleosynthesis

Big Bang nucleosynthesis (BBN) is a standardized theoretical framework with which the abundances of light nuclei formed in the early Universe can be predicted with high reliability. BBN converts the roughly equal number of protons and neutrons that existed just moments after the Big Bang ( $t \ll 0.1$  s) and transforms them into (effectively) stable densities of  $^4\text{He}$ , deuterium (D),  $^3\text{He}$ , and  $^7\text{Li}$  several minutes later ( $t \simeq 30$  minutes). Aside from the well known nuclear cross sections involved, BBN abundance calculations depend solely on the baryon number density  $n_b$ , and can therefore be compared to experimental observations of the light-element abundances to constrain the baryonic component of  $\Omega_{\mathcal{M}}$ . Since more detailed and pedagogical discussions of BBN can be found elsewhere (see, *e.g.*, [41], [42] or [43]), I will simply review a few of the basic concepts here.

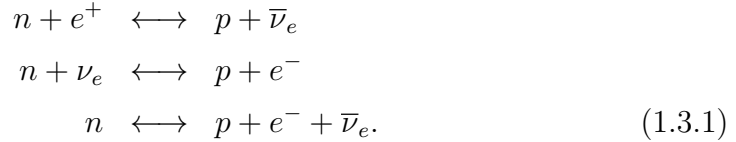
Standard BBN calculations rely on the following simplifying, but reasonable assumptions [44]:

- The expanding Universe following the Big Bang is spatially isotropic and homogeneous. It is in a radiation-dominated epoch during which contributions to  $\Omega_0$  from dark matter and dark energy are negligible.
- The Universe starts out hot enough for the protons and neutrons to be in thermal equilibrium, such that there are an approximately equal number of each nucleon (explained below).
- The fundamental particles and their interactions are governed by the SM (discussed in Section 1.1), and the baryon asymmetry (absence of antibaryons) already exists. Furthermore, photons, neutrinos, and (for a limited time) electrons and positrons dominate  $\Omega_0$ .
- Nuclear and fundamental particle properties (*e.g.*, masses, couplings, cross sections and lifetimes) are the same today as they were in the early Universe when BBN took place.

### 1.3. OBSERVATIONAL EVIDENCE

Since the early Universe was radiation dominated, the equation of state (and therefore the expansion rate and temperature) depended on the number of relativistic degrees of freedom  $g^*$  [45], which in turn determined when certain processes could and could not occur. BBN is thus sensitive to the number of relativistic particle species present in the early Universe. Photons and the three SM neutrino species contribute to  $g^*$  throughout BBN, whereas electrons and positrons contribute for only the first second. This dependence on  $g^*$  led to one of the early successes of BBN theory; BBN calculations accurately predicted the number of light-neutrino species [46, 47] well before it was precisely measured from the invisible width of the  $Z$  boson at electron-positron colliders many years later [48, 49, 50, 51].

The first stage of BBN involves the determination of the neutron-to-proton ratio ( $n/p$ ). At sufficiently high temperatures ( $kT \gg 1$  MeV), neutrinos are in thermal equilibrium and the following processes readily proceed in either direction:



These processes fix the nucleon number densities and are highly sensitive to the  $\sim 15$  minute neutron half-life. While  $kT \gg 1$  MeV, neutrinos interact frequently enough to maintain  $(n/p) \simeq 1$ . As indicated by the left-most vertical band in Figure 1.3, when the rate of weak interactions becomes outpaced by the expansion  $\sim 0.1$  s after the Big Bang, neutrinos chemically decouple. At a characteristic temperature of approximately 0.8 MeV, the processes in Equation 1.3.1 turn off and after  $\sim 1$  s ( $n/p$ ) freezes out and can be estimated by the Boltzmann factor,

$$(n/p) = e^{-\Delta mc^2/kT} \simeq \frac{1}{5}, \tag{1.3.2}$$

where  $\Delta m$  is the difference in mass between the two nucleons ( $\sim 1.3$  MeV). This stage is often referred to as ( $n/p$ ) decoupling (second vertical band from the left in Figure 1.3). At roughly the same time electrons and positrons

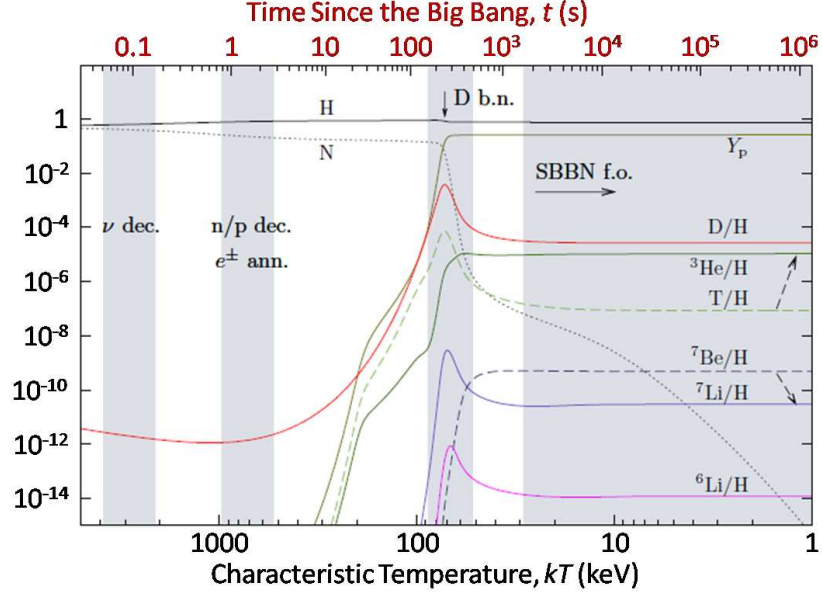


Figure 1.3: Evolution of the abundances of light elements as a function of time (top axis) and characteristic temperature (bottom axis). From left to right, the blue bands indicate significant stages in Big Bang nucleosynthesis: neutrino decoupling, freeze-out of the neutron-to-proton ratio, the deuteron bottleneck, and freeze-out of the light-element abundances. The proton (H) and neutron (N) abundances are given relative to the total number of baryons, and  $Y_p$  denotes the  $^4\text{He}$  mass fraction. Figure adapted from [44].

annihilate, and the  $e^\pm$  degrees of freedom no longer contribute to  $g^*$ , slowing the expansion and reheating the photon background. Although the average temperature is now below the deuteron binding energy, the temperature profile of background photons has a tail of photons energetic and numerous enough to photodissociate deuterons nearly as fast as they can be produced. The overall effect is a prolonged period of relative inactivity following  $(n/p)$  decoupling that lasts for  $\sim 100$  s. As the photons continue to lose energy to the expansion, the decay of neutrons causes  $(n/p)$  to fall to about  $1/7$  before deuteron production truly takes off.

### 1.3. OBSERVATIONAL EVIDENCE

At a characteristic temperature of  $\sim 100$  keV BBN enters a stage referred to as the deuteron bottleneck, indicated in Figure 1.3 by the third vertical band from the left. All nuclei heavier than a deuteron depend directly on the deuteron number density. Once it is sufficiently large, the reactions that produce helium and lithium nuclei ignite, and free neutrons are very quickly bound (primarily) into  ${}^4\text{He}$  nuclei. To surprisingly good accuracy, the  ${}^4\text{He}$  mass fraction ( $Y_p$ ) can be estimated from  $(n/p)$ :

$$Y_p \simeq \frac{2(n/p)}{1 + (n/p)} \bigg|_{kT \simeq 100 \text{ keV}} = \frac{1}{4}. \quad (1.3.3)$$

To a lesser extent,  ${}^3\text{He}$  and  ${}^7\text{Li}$  are produced, and a few light radioactive nuclei are formed as well (*e.g.*, tritium,  ${}^7\text{Be}$  and  ${}^6\text{Li}$ ). Production of heavier elements ( $A > 7$ ) is highly suppressed because the Coulomb barrier to such reactions is simply too large. Half an hour after the Big Bang the nuclear reactions turn off and the light-element abundances freeze-out. The BBN process from beginning to end is summarized in Figure 1.3.

BBN predictions of the abundances of light nuclei are typically calculated as a function of the baryon-to-photon ratio, a quantity that effectively tracks the baryon density in the early Universe. It is precisely known from measurements of the CMB (discussed in more detail in the next section):

$$\eta = \frac{n_b - n_{\bar{b}}}{n_\gamma} \simeq \frac{n_b}{n_\gamma} = (6.23 \pm 0.17) \times 10^{-10}, \quad (1.3.4)$$

where  $n_b$ ,  $n_{\bar{b}}$ , and  $n_\gamma$  are the baryon, antibaryon and photon number densities, respectively [52]. Figure 1.4 gives a compilation of BBN predictions as a function of  $\eta$  for the abundance of  ${}^4\text{He}$  in terms of mass fraction ( $Y_p$ ) and for the abundances of D,  ${}^3\text{He}$ , and  ${}^7\text{Li}$  in terms of number densities relative to hydrogen. The central predictions are indicated by black curves surrounded by semi-horizontal  $1\sigma$  uncertainty bands. The uncertainties are dominated by eleven (key) strong-interaction rates and the uncertainty in the neutron lifetime [56]. The  $1\sigma$  uncertainty on the value of  $\eta$  quoted in Equation 1.3.4 is indicated by the vertical band. The CMB measurement of  $\eta$  coupled with such precise BBN predictions results in a fully (and tightly) constrained theory for

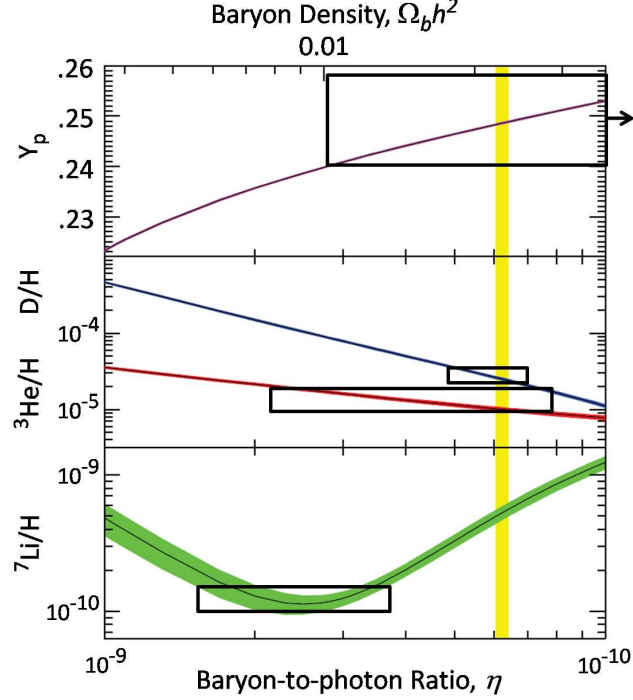


Figure 1.4: BBN predictions for the abundances of light elements as a function of the baryon-to-photon ratio (bottom axis) and baryon density (top axis). The  $^4\text{He}$  abundance ( $Y_p$ ) is given in terms of its baryonic mass fraction, while the others are in terms of number densities relative to hydrogen. The semi-horizontal bands represent the BBN predictions' (black lines)  $1\sigma$  uncertainties, while the vertical band represents the uncertainty in the baryon density inferred from the WMAP 5-year CMB data [52]. The black boxes indicate the  $1\sigma$  confidence intervals associated with recent measurements. Figure taken from [45] and augmented with recent measurements from (top to bottom) [53, 45, 54, 55].

the light-element abundances. Also indicated in Figure 1.4, measurements of these four abundances provide a powerful test of the theory.

Direct measurements of the deuterium abundance can be made through observations of its isotope-shifted Lyman- $\alpha$  absorption. The most reliable abun-

### 1.3. OBSERVATIONAL EVIDENCE

dance measurement comes from observations of seven low-metallicity, high-redshift-quasar absorption systems, and yields a value for D/H of  $(2.82 \pm 0.21) \times 10^{-5}$  [57]. Following [45], a somewhat larger uncertainty derived from the sample variance seems appropriate to account for systematic errors, yielding the  $1\sigma$  confidence interval  $D/H = (2.82 \pm 0.53) \times 10^{-5}$ . This measurement is indicated by the smallest black box in Figure 1.4 and compares favorably with the prediction inferred from the CMB measurement of  $\eta$ ;  $D/H = (2.49 \pm 0.17) \times 10^{-5}$  [58].

The data used to measure the  ${}^4\text{He}$  abundance are far more rich and easily obtained since the value of  $Y_p$  is so large, yet yield a less precise determination of  $\eta$  (and therefore  $\Omega_b$ ). The measurements are complicated by production and ejection of  ${}^4\text{He}$  by stellar processes. Traditionally, metal-poor regions absent of star formation such as clouds of ionized hydrogen (known as H II regions) in dwarf galaxies are the most reliable. Observations of the emissions from ionized  ${}^4\text{He}$  recombining in these gas clouds yield a value of  $Y_p = 0.249 \pm 0.009$  [53], which is remarkably close to the rough estimate given in Equation 1.3.3. As can be seen in Figure 1.4, it also agrees well with the BBN prediction based on the CMB value of  $\eta$ ;  $Y_p = 0.2486 \pm 0.0002$  [58].

The primordial  ${}^7\text{Li}$  abundance presents somewhat of a mystery; there is a significant discrepancy between the observed value and the CMB-assisted BBN prediction that might be an indication of new physics or nontrivial systematics. Aside from its production during BBN, low levels of lithium are produced via cosmic-ray-assisted  $\alpha$ -particle fusion and in Type II supernovae [59]. Furthermore, the amount of stellar nonprimordial lithium appears to correlate with a star’s metallicity. Consequently, the Milky Way’s metal-poor Pop II stars<sup>4</sup> exhibit the most consistently low levels of lithium, a phenomenon usually referred to as the “Spite plateau” [60]. Fitting a model for the total lithium as

---

<sup>4</sup>Generally the stars within the Milky Way are classified according to their heavy-element abundance, or metallicity. Roughly 2% are Population I (Pop I) stars that tend to occupy positions within the galactic disk and have relatively high metallicities. They tend to be young, hot and luminous. Population II (Pop II) stars are at the other end of the spectrum and tend to populate the stellar halo surrounding the center of the galaxy (often referred to as the bulk or bulge).



## CHAPTER 1. INTRODUCTION

a function of metallicity to observations from both low- and high-metallicity stars, the primordial  ${}^7\text{Li}$  abundance can be estimated by extrapolating the fit to zero metallicity, resulting in a value for  ${}^7\text{Li}/\text{H} = (1.23^{+0.34}_{-0.16}) \times 10^{-10}$  [55]. As is clear from Figure 1.4, the discrepancy with the CMB-assisted BBN prediction is severe,  ${}^7\text{Li}/\text{H} = (5.24^{+0.71}_{-0.62}) \times 10^{-10}$  [58]. The resolution to this apparent contradiction is an active area of research. Possibilities include systematic investigations into stellar-temperature estimates (upon which the lithium abundance is highly sensitive—see, *e.g.*, [61, 62]) and  ${}^7\text{Li}$  depletion in the surface of stars via convective dynamics [63]. More intriguing is the possibility that this lithium problem is an indication of the type of new physics beyond the SM (such as supersymmetry, see Section 2.5) that also produces a particulate dark-matter candidate [64].

Although  ${}^3\text{He}$  is more abundant than  ${}^7\text{Li}$ , when coupled to BBN predictions it provides a far less reliable measurement of the baryon density. Like  ${}^4\text{He}$ , the abundance can be estimated from observations of H II regions. The authors in [54] find a range for  ${}^3\text{He}/\text{H} = (0.9\text{--}1.9) \times 10^{-5}$ , which is consistent with the CMB-assisted BBN prediction (see Figure 1.4). However, unlike  ${}^7\text{Li}$ , the chemical evolution of  ${}^3\text{He}$  since it was originally produced during BBN is unknown. Without a working model for its post-BBN production or depletion,  ${}^3\text{He}$  abundance measurements from nearby H II regions are not useful for predicting the primordial abundance. The  ${}^3\text{He}$  abundance is instead used to better understand stellar evolution by nailing down its primordial value using the CMB value of  $\eta$  and comparing observations of H II regions to planetary nebulae.

In summary, BBN refers to a set of highly constrained calculations for predicting the abundances of light nuclei ( $A < 8$ ) as a function of a single parameter, the baryon-to-photon ratio. Based on a minimal set of well supported assumptions, during a period of approximately 30 minutes beginning a split second after the Big Bang, BBN consolidates a fraction of the Universe’s baryons (primarily in the form of free nucleons) into stable abundances of helium, lithium, and deuterium nuclei. Observations of these abundances today can be used to place constraints on the contribution to  $\Omega_{\mathcal{M}}$  from bary-

### 1.3. OBSERVATIONAL EVIDENCE

onic matter. Measurement of the deuterium abundance provides the tightest and most reliable constraint, implying a value of  $\Omega_b \simeq 0.035\text{--}0.052$ . As will be discussed in the following section, analysis of CMB data provide tighter constraints, but are more model dependent. When combined with measurements of the total matter density  $\Omega_{\mathcal{M}}$ , the BBN prediction for  $\Omega_b$  provides one of the most compelling arguments for the existence of nonbaryonic dark matter.

#### 1.3.2 The Cosmic Microwave Background Radiation

Perhaps the richest source of cosmological information and consequently the most convincing body of support for the Standard Cosmology is derived from observations of the cosmic microwave background radiation. Penzias and Wilson's original (somewhat accidental) discovery in 1965 of a uniform glow of microwaves with a characteristic temperature of  $T = 3.5 \pm 1.0$  K everywhere in the sky revolutionized the field of cosmology, lending credence to concepts like spatial isotropy, the Big Bang and inflation [65]. A rich history of theoretical speculation and experimental discovery followed their groundbreaking achievement, including a refinement of the CMB temperature ( $T \simeq 2.73$  K) and the definitive discovery of a CMB anisotropy ( $\Delta T/T \simeq 6 \times 10^{-6}$ ) by the COBE satellite nearly 30 years later [66].

In the last two decades, refined measurements of the CMB anisotropy and its polarization modes by ground-based telescopes (*e.g.*, VIPER [67] and TOCO [68]), interferometers (*e.g.*, CBI [69] and DASI [70]), as well as balloon-born instruments (*e.g.*, MAXIMA [71] and BOOMERANG [72]) have (arguably) turned experimental cosmology into a precision science. However, while these technologies continue to make significant contributions to the field, only the satellite-born Wilkinson Microwave Anisotropy Probe (WMAP) [73] has provided a detailed all-sky map of the anisotropy with which the Standard Cosmology can be comprehensively tested. In this section I review how observations of the CMB can be interpreted to measure the basic cosmological parameters, with emphasis on recent measurements based on the WMAP 7-

## CHAPTER 1. INTRODUCTION

year data. My treatment of this subject is necessarily brief; an entire Ph.D. thesis could be devoted to understanding and interpreting WMAP data and it would still be only a partial survey of the subject. I refer the interested reader to a series of papers by the WMAP collaboration that explore the collection, processing and implications of their 7-year CMB data in far greater detail [24, 74, 75, 76, 77, 78].

Before jumping into cosmological-parameter measurements, we must first understand a little about what CMB data is and how it is obtained. In the case of the WMAP satellite, microwave radiation is collected with a pair of back-to-back  $\gtrsim 2\text{ m}^2$  Gregorian-style telescopes that feed a series of horns attached to radiometers designed to make multifrequency differential measurements of the CMB radiation [79]. Rather than measure the absolute temperature of the CMB *à la* Penzias and Wilson, CMB temperature differences between two points in the sky separated by  $180^\circ$  are measured to approximately  $\mu\text{K}$  accuracy. The instrument aboard WMAP is commonly referred to as a differential microwave radiometer (DMR). The DMR design has the obvious advantage that it is largely insensitive to the temperature or radiative properties of its host satellite (similar to how the noise performance of a transistor-based amplifier is improved when the transistors share a common substrate and therefore temperature), making it largely insensitive to a variety of nontrivial systematic temperature variations. WMAP orbits the Sun in the Sun-Earth  $L_2$  Lagrange point where it is partially shielded from solar radiation, and where it rotates and precesses quickly enough for its DMR to scan microwave-temperature differences over  $\sim 30\%$  of the sky every hour. Due to its orbit, however, the instrument requires six months in order to access the entire sky, resulting in highly redundant scans of the entire sky in five different frequencies. The WMAP collaboration have released data four times, once after the first fully operational solar orbit (WMAP first-year data [73]) and once every two years since then (WMAP 3-year [80], 5-year [81], and 7-year data [74]).

### 1.3. OBSERVATIONAL EVIDENCE

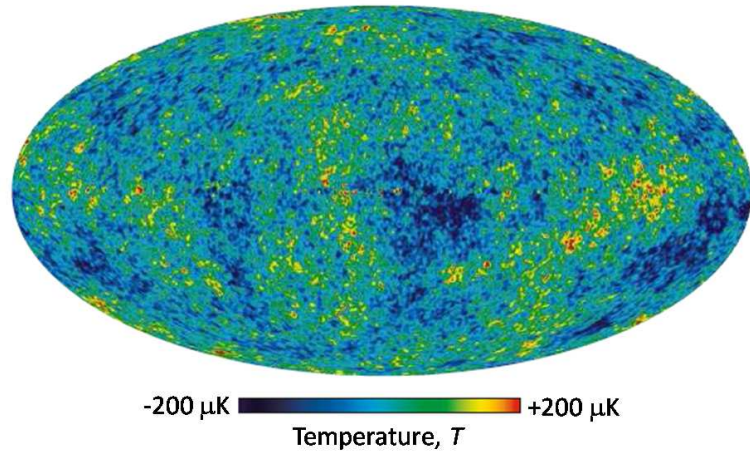


Figure 1.5: The Wilkinson Microwave Anisotropy Probe 5-year data all-sky map of the cosmic microwave background primary anisotropy. A Mollweide equal-area projection is used to display the entire sky in galactic coordinates, with temperature differences given in units of thermodynamic temperature. The most recent primary anisotropy maps from the 7-year data release tend to be published in the form of differences relative to this map (see, *e.g.*, Figure 2 in [77]). The differences are consistent with pixel noise, slight calibration errors and an expected change in the Earth’s dipole signature; by eye they are nearly indistinguishable. Figure taken from [81].

The data for each frequency band are converted to an all-sky map that is typically displayed in a Mollweide equal-area projection<sup>5</sup> in Galactic coordinates, with the temperature differences given in units of CMB thermodynamic temperature [82]. At this stage, the maps not only include the primordial-temperature differences from the surface of last scattering (“primary anisotropy”), but also include features due to diffuse galactic emission, point sources such as planets and nearby galaxies, and a large dipole signature caused by the motion of the Earth relative to the cosmic rest frame. The reason

---

<sup>5</sup>The Mollweide projection is a type of coordinate transformation commonly used for maps of the globe or the sky. It accurately represents area while tending to distort angles and shapes. For CMB anisotropy maps, the galactic plane runs horizontally through the middle of the map.

## CHAPTER 1. INTRODUCTION

the maps are recorded at five different frequencies is to remove these foreground features. The technique for extracting the primary anisotropy is described in [83] and results in a single map like the one shown in Figure 1.5.

In order to fit a cosmological model to CMB data, the primary-anisotropy map is (typically) decomposed into spherical harmonics,  $Y_{lm}$ ;

$$T(\hat{\mathbf{n}}) = \sum_{l,m} a_{lm} Y_{lm}(\hat{\mathbf{n}}), \quad (1.3.5)$$

where  $\hat{\mathbf{n}}$  is a unit direction vector and represents the angular position of a map pixel. The angular power spectrum at multipole moment  $l$  is given by an average over the moments  $m$ ;

$$C_l = \frac{1}{2l+1} \sum_{m=-l}^l |a_{lm}|^2. \quad (1.3.6)$$

The power spectrum is usually plotted as a function of  $l$  in terms of the squared temperature anisotropy

$$(\Delta T_l)^2 = l(l+1)C_l/2\pi. \quad (1.3.7)$$

The error bars in Figure 1.6 show the power spectrum derived from the WMAP 7-year CMB data for multipoles up to  $\gtrsim 1000$  [24], augmented with higher multipole moments derived from data obtained by the ACBAR [84] and QUaD [85] experiments. The spectrum can be thought of as the amount of power stored in small- and large-scale fluctuations in the CMB temperature, where low multipole moments represent large angular scales (*e.g.*,  $l < 100$  corresponds to  $\theta \gtrsim 2^\circ$ ) and high multipole moments represent relatively small angular scales (*e.g.*,  $l > 1000$  corresponds to  $\theta \lesssim 0.2^\circ$ ). Since these temperature differences represent fluctuations in the temperature of the surface of last scattering, they correspond to the matter-density fluctuations in the early Universe that eventually evolved into the structures we see today.

Prior to discussing the details of the best-fit cosmological model shown in Figure 1.6, it is instructive to qualitatively explore the structure of the power spectrum. For the first two thousand multipole moments the shape is primarily

### 1.3. OBSERVATIONAL EVIDENCE

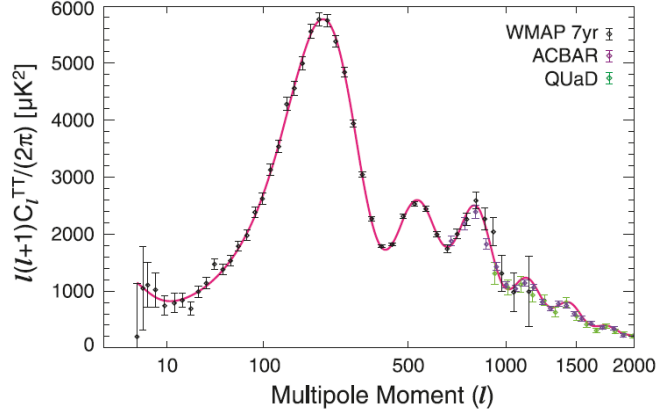


Figure 1.6: The angular power spectrum of the CMB primary temperature anisotropy as a function of multipole moment. The black error bars up to  $l \simeq 1200$  are derived from the WMAP 7-year data [76], while the lighter colored error bars for  $l \geq 690$  are derived from data obtained by the ACBAR [84] and QUaD [85] experiments. The WMAP errors for  $l$  up to about 600 are consistent with cosmic variance, while those for larger multipoles are dominated by instrument noise. The solid curve represents the best-fit  $\Lambda$ CDM model to only the WMAP data. Figure taken from [24].

due to three effects: the Sachs-Wolfe effect [86], acoustic oscillations, and Silk damping [87]. The following discussion follows the review by Scott and Smoot in [88].

The power spectrum is approximately flat for (nearly) the first hundred multipole moments and corresponds to gravitational fluctuations on scales larger than the cosmic horizon. Two points in space separated by a distance greater than the horizon scale are not causally connected within the age of the Universe. Such large-scale fluctuations have therefore had no time to evolve; they reflect the nature of the earliest gravitational perturbations. These perturbations cause slight gravitational redshifts and blueshifts to the CMB photons at the surface of last scattering due to corresponding matter overdensities and underdensities, resulting in a weak CMB anisotropy for  $l < 100$  commonly referred to as the Sachs-Wolfe effect [86]. The shape of the power spectrum

## CHAPTER 1. INTRODUCTION

for  $l < 100$  has sensitivity to a cosmological parameter known as the spectral index of density perturbations,  $n_s$ ; if  $n_s \simeq 1$  then the initial perturbations are scale-invariant and result in a CMB power spectrum that is nearly constant for large angular scales. This feature between  $l$  of 10 and 100 is commonly referred to as the Sachs-Wolfe plateau.

For multipoles less than  $l = 10$ , there is a related contribution to the power spectrum caused by the integrated Sachs-Wolfe effect. At the largest scales there is a bit of extra power in the CMB anisotropy due to gravitational redshifting and blueshifting of the CMB photons as they travel between the surface of last scattering and the Earth, causing a gentle rise in the power spectrum as  $l$  decreases from 10 down to 1. In a Universe without a cosmological constant the equation of state following recombination is effectively zero and unchanging. The gravitational potentials that the CMB photons encounter on their way to our instruments do not evolve significantly as the photons pass through them. Consequently, CMB photons will gain (blueshift) and lose (redshift) equal but opposite gravitational kicks as they pass through these wells, remaining unchanged overall. In a dark-energy-dominated Universe, however, the cosmological constant causes the gravitational wells (voids) to evolve, and the CMB photons receive a net gain (loss) as the wells (voids) become shallower (deeper) during their passage. The overall effect is that if the CMB photons traverse a dark-energy-dominated period before they are detected, the CMB anisotropy will have some sensitivity to  $\Omega_\Lambda$  in the low  $l$  part of the power spectrum.

The peaked structure of the power spectrum for  $l \gtrsim 100$  is due to acoustic oscillations in the baryon-photon fluid prior to recombination. Acoustic oscillations can be understood as a competition between the tendency for baryonic matter to clump as it falls into gravity wells caused by the initial gravitational perturbations, and the tendency for the radiation pressure of the tightly coupled photon background to oppose the clumping. These prerecombination harmonic oscillations in the baryon-photon fluid cause time variations in the fluid temperature with a frequency characteristic of the speed of sound (or acoustics) in the fluid. Following recombination, the photon background decouples and

### 1.3. OBSERVATIONAL EVIDENCE

the phases of the acoustic oscillations freeze out, resulting in the pattern of harmonic peaks we see in the CMB power spectrum.

The power spectrum's acoustic peaks are sensitive to several cosmological parameters. The curvature of spacetime depends on the peaks' angular positions and is sensitive to  $\Omega_K$  (and under some scenarios to  $\Omega_0$ ). An open ( $k_c = -1$ ) Universe has peaks shifted toward higher multipoles relative to a closed ( $k_c = 1$ ) geometry. The peaks' amplitudes relative to the Sachs-Wolfe plateau provide a measure of the reionization optical depth  $\tau$ . The scattering of  $100\tau\%$  of the CMB photons following reionization ( $z < z_{\text{reion}}$ ) partially erases the CMB anisotropy for angular scales less than a few degrees, causing a reduction of the acoustic peaks' amplitudes by a factor of  $e^{-2\tau}$  relative to the Sachs-Wolfe plateau. Furthermore, the absolute and relative amplitudes of the first three peaks is highly sensitive to the baryonic ( $\Omega_b$ ) and nonbaryonic ( $\Omega_c$ ) matter densities in the early Universe. For example, a larger value of  $\Omega_b$  causes the baryon-photon fluid to fall deeper into the initial gravitational perturbations, thereby enhancing the compressional acoustic oscillations (odd numbered peaks) relative to the anticompressional acoustic oscillations (even numbered peaks).

Acoustic peaks at higher and higher  $l$  exhibit a damped behavior, eventually dying out for multipoles greater than  $\sim 2000$ . This is a direct consequence of the nonzero time scale over which recombination occurs and is called Silk damping [87]. Effectively, the surface of last scattering has a thickness. CMB observations measure the average temperature of photons from throughout this thickness, which tends to smear out the small-scale anisotropies and therefore dampen the power spectrum for angular scales less than about a tenth of a degree.

Fitting a specific cosmological model to the CMB power spectrum is complicated by the subtle interplay of the many effects that give rise to its shape. Cosmologists have developed a number of sophisticated computer programs to calculate theoretical models and perform likelihood fits to the power spectrum (*e.g.*, CMBFAST [89], CAMB [90], and RECFAST [91]). The best-fit curve in Figure 1.6 makes use of several such codes in concert to find the simplest known



model that reasonably fits the WMAP 7-year power spectrum (dark error bars up to  $l \simeq 1200$  only). The so-called minimal  $\Lambda$ CDM is a six-parameter model for which the geometry of the Universe is assumed to be flat ( $\Omega_K = 0$ ),  $\Omega_\Lambda$  is nonzero, the dark-energy equation of state is a constant ( $w = -1$ ), and  $\Omega_M$  has a nonbaryonic and cold dark-matter (CDM) component ( $\Omega_c$ ) in addition to a baryonic component ( $\Omega_b$ ).<sup>6</sup> The model derives its name from the latter two assumptions and has become the reference cosmology against which ideas for new physics are tested. Table 1.2 summarizes the best-fit parameters as well as a few derived quantities.

For the purposes of this thesis, the most important results from Table 1.2 are the values of  $\Omega_b$  and  $\Omega_c$ . If we ignore the  ${}^7\text{Li}$  anomaly, the value of  $\Omega_b$  coupled with BBN predictions and observations of the light-element abundances firmly establish the amount of baryonic matter in the Universe at roughly 5% of the critical density. However, to fully explain the pattern of temperature fluctuations that was imprinted upon the sky when the primordial photon background emerged from the surface of last scattering requires a significant—23% of the critical density—nonbaryonic-matter density  $\Omega_c$ . Furthermore, this nonbaryonic matter cannot have coupled strongly to the baryon-photon fluid prior to recombination or have been moving at relativistic speeds, implying that  $\Omega_c$  is comprised of a non-SM type of nonrelativistic (cold) and nonluminous (dark) matter, or cold dark matter.

### 1.3.3 Type Ia Supernovae

While fits to the CMB power spectrum provide support for the Standard Cosmology through observations of the early Universe ( $z > 1000$ ), there is also a large body of corroborating evidence from a type of (relatively) nearby astrophysical source ( $z \lesssim 2$ ) known as a Type Ia supernova (SNe Ia) [92]. As the name implies, SNe Ia are a subcategory of exploding stars believed to be the result of binary systems in which a white dwarf accretes mass from a companion

---

<sup>6</sup>Fits to the WMAP 7-year power spectrum in which additional  $\Lambda$ CDM parameters are allowed to float (such as  $\Omega_K$  and  $w$ ) are consistent with the minimal, 6-parameter  $\Lambda$ CDM model assumptions, but yield less accurate values of the basic parameters [76].

### 1.3. OBSERVATIONAL EVIDENCE

Table 1.2: The six best-fit parameters of the minimal  $\Lambda$ CDM model for the fit to the WMAP 7-year power spectrum shown in Figure 1.6, as well as a few derived parameters.  $h = H_0/(100 \text{ km/s/Mpc})$  is a unitless form of the Hubble constant and is implicitly determined by the flat spacetime constraint ( $\Omega_b + \Omega_c + \Omega_\Lambda = 1$ ) imposed by the fit. Values and their  $1\sigma$  errors taken from [76].

Symbol	Description	Best-fit Value
<i>Model parameters</i>		
$10^2 \Omega_b h^2$	Baryon density	$2.249^{+0.056}_{-0.057}$
$\Omega_c h^2$	Dark-matter density	$0.1120 \pm 0.0056$
$\Omega_\Lambda$	Dark-energy density	$0.727^{+0.030}_{-0.029}$
$10^9 \Delta_{\mathcal{R}}^2$	Curvature perturbation amplitude	$2.43 \pm 0.11$
$n_s$	Spectral index	$0.967 \pm 0.014$
$\tau$	Reionization optical depth	$0.088 \pm 0.015$
<i>Derived parameters</i>		
$t_0$	Age of the Universe (Gyr)	$13.77 \pm 0.13$
$H_0$	Hubble parameter ( $\text{km s}^{-1} \text{Mpc}^{-1}$ )	$70.4 \pm 2.5$
$\Omega_b$	Baryon density	$0.0455 \pm 0.0028$
$\Omega_c$	Dark-matter density	$0.228 \pm 0.027$
$z_{\text{eq}}$	Matter-radiation equality redshift	$3196^{+134}_{-133}$
$z_{\text{reion}}$	Reionization redshift	$10.6 \pm 1.2$

## CHAPTER 1. INTRODUCTION

star until it nears the Chandrasekhar limit of  $1.38 \mathcal{M}_\odot$  [93]. As the mass approaches  $1.38 \mathcal{M}_\odot$ , the electron degeneracy pressure that prevents a white dwarf from collapsing further can no longer balance the star's inward gravitational self-attraction. The white dwarf collapses, initiating a brief but intense period (typically a few seconds) during which its core material (typically carbon and oxygen) undergoes a runaway fusion reaction that generates enough energy to completely unbind the star. Since the accretion mechanism causes the build up of a consistent amount of stellar material before the  $1.38 \mathcal{M}_\odot$  limit is reached, the peak light output (or luminosity) of the subsequent explosion is highly uniform from one SNe Ia to the next. This makes SNe Ia excellent “standard candles” with which to test the Universe’s distance versus redshift relation [94].

The idea is to compare the light observed (or flux,  $F$ ) for a series of objects known to have similar intrinsic luminosities ( $L$ ) at a variety of cosmic distances (or redshifts). In a geometrically flat Universe, the observed flux is directly proportional to the luminosity and inversely proportional to the square of the distance to the object. More generally, the observed flux depends on the curvature of spacetime; lower (higher) flux is observed for an open (closed) geometry. One can define the luminosity distance

$$d_L(z) \equiv \sqrt{\frac{L}{4\pi F}} = (1+z)r(z), \quad (1.3.8)$$

where  $r(z)$  is the comoving distance to an object at redshift  $z$  and depends on the curvature parameter  $k_c$ :

$$\begin{aligned} r(z) &= c \int_0^z \frac{dz'}{H(z')} & (k_c = 0), \text{ or} \\ r(z) &= X\left(\sqrt{|k_c|}c \int_0^z \frac{dz'}{H(z')}\right) / \sqrt{|k_c|} & (k_c \neq 0), \end{aligned} \quad (1.3.9)$$

where  $X(x) = \sin(x)$  for  $k_c = 1$ , and  $X(x) = \sinh(x)$  for  $k_c = -1$  [31]. The relationship between flux and redshift for a collection of standard candles with known intrinsic luminosities (like SNe Ia) is therefore sensitive to a number

### 1.3. OBSERVATIONAL EVIDENCE

of cosmological parameters. This is most easily demonstrated for a flat Universe for which the dark-energy equation of state is a constant (*e.g.*,  $w = -1$ ). Combining Equations 1.2.7, 1.3.8, and 1.3.9 for  $k_c = 0$  yields,

$$d_L(z) = \frac{c(1+z)}{H_0} \int_0^z \frac{dz'}{\sqrt{\Omega_\Lambda + \Omega_{\mathcal{M}}(1+z')^3 + \Omega_R(1+z')^4}}. \quad (1.3.10)$$

Since the value of  $\Omega_R$  is well known from measurements of the CMB, distances to standard candles (for this particular example) depend on  $\Omega_{\mathcal{M}}$ ,  $\Omega_\Lambda$  and  $H_0$ . Furthermore, for  $z \ll 1$  and  $\Omega_R \ll 1$ , to leading order in  $z$  Equation 1.3.10 reduces to

$$\begin{aligned} d_L(z) &\simeq \frac{c(1+z)}{H_0} \int_0^z \frac{dz'}{\sqrt{1 + 3z'\Omega_{\mathcal{M}}}} \\ &= \frac{2c(1+z)}{3\Omega_{\mathcal{M}}H_0} \left[ \sqrt{1 + 3z\Omega_{\mathcal{M}}} - 1 \right] \\ &\simeq \frac{cz}{H_0}(1+z). \end{aligned} \quad (1.3.11)$$

This is easily recognized as Hubble’s law in which the comoving distance is directly proportional to redshift:  $r(z) = cz/H_0$  [40]. For nearby SNe Ia ( $z \lesssim 0.1$ ) the relationship between distance and redshift is therefore sensitive to only  $H_0$ . Measuring  $H_0$  with low- $z$  standard candles was the Hubble Space Telescope’s so-called “key project” [27]. More recently, the  $SH_0ES$  program used several hundred Cepheid variable stars<sup>7</sup> to calibrate nearby SNe Ia luminosities and estimate  $H_0 = 73.8 \pm 2.4 \text{ km s}^{-1} \text{ Mpc}^{-1}$  [25] (see Figure 1.1).

SNe Ia are usually discovered with optical telescopes (*e.g.*, the Hubble Space Telescope [27]) by scanning known galaxies and galaxy clusters for brightening

---

<sup>7</sup>Cepheid variables are very bright, pulsating stars whose luminosity varies periodically with time. Cepheids are commonly categorized into several subclasses according to their masses, metallicities, and evolutionary histories. Classical (Type II) Cepheids are massive (low mass) Pop I (Pop II) stars that pulsate with a periodicity ranging between days and months. Their intrinsic luminosities can be reliably measured based on their pulsation periods, making them excellent standard candles for measuring cosmic distances out to a few tens of megaparsecs. Hubble’s discovery in 1929 that the Universe is expanding was based on Cepheids [40].

## CHAPTER 1. INTRODUCTION

point sources that mark the beginning of supernovae explosions. Once a candidate event has been identified, several multicolor (or photometric) followup observations are made—ideally several times over the course of a few months—to measure its apparent magnitude as a function of time, or “light curve.” Due to the highly uniform mechanism by which SNe Ia light is produced, their light curves share a common (characteristic) shape, which is what makes them suitable for use as standard candles. Through comparison to other SNe Ia, a candidate’s light curve is used to validate its intrinsic luminosity and measure its apparent magnitude. Spectrographic observations of the candidate’s host galaxy are typically used to estimate its redshift.

In addition to the experimental challenge of picking these somewhat rare events out of the sky, using SNe Ia as standard candles is complicated by a number of systematic uncertainties that can affect the observational data. Systematics include but are not limited to: light contamination from the host galaxy, light absorption by interstellar material in the host and Milky Way galaxies, light extinction by the Earth’s atmosphere, distortion due to gravitational lensing, and instrumental aberrations. Due to the  $z$  dependencies in Equation 1.3.10, relatively high- $z$  ( $z \gtrsim 0.5$ ) SNe Ia are required to gain sensitivity to cosmological parameters other than  $H_0$ . Many of the systematic effects are increasingly difficult to handle for high- $z$  observations. Consequently, convincing evidence (based on SNe Ia) for a nonzero  $\Omega_\Lambda$  did not emerge until 1998 when techniques for observing and analyzing SNe Ia matured sufficiently for surveys to include redshifts up to  $z \lesssim 1$  (*e.g.*, the High- $z$  Supernova Search Team [95] and the Supernova Cosmology Project [96]).

Further systematics arise when combining data sets from different teams’ observations to form a high statistics sample over a broad range of redshifts. There have been several efforts to compile SNe Ia data into self-consistent and systematic-free catalogs suitable for measuring cosmological parameters (see, *e.g.*, [106, 105, 114, 99, 98, 115, 116]). The most recent and comprehensive such collection (the “Union2.1” compilation [110]) is shown in Figure 1.7 in terms of a quantity known as the distance modulus,  $\mu_0$ .  $\mu_0$  is defined as the difference between an object’s apparent ( $m_{\text{mag}}$ ) and absolute magnitudes ( $M$ ),

### 1.3. OBSERVATIONAL EVIDENCE

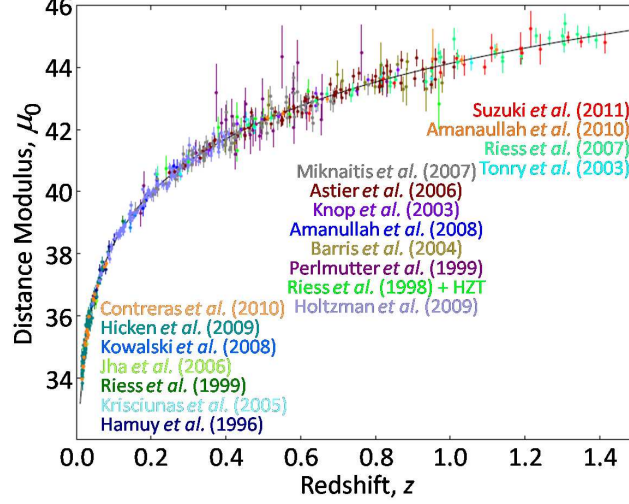


Figure 1.7: The Hubble diagram (distance modulus versus redshift) for the Union2.1 Type Ia supernovae compilation. The curve represents the best-fit minimal  $\Lambda$ CDM cosmological model and is consistent with the  $\Lambda$ CDM fit to the CMB power spectrum shown in Figure 1.6. The references for the different colored error bars are (listed from left to right and top to bottom as authors and years) [97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 96, 95, 109, 110, 111, 112, 113]. Figure adapted from [110].

and is related to the luminosity distance;

$$\mu_0(z) \equiv m_{\text{mag}} - M = 5 \log_{10}(d_L/10 \text{ pc}), \quad (1.3.12)$$

where  $m_{\text{mag}}$  and  $M$  are proportional to the logarithms of the object's observed flux ( $F$ ) and intrinsic luminosity ( $L$ ), respectively. In terms of redshift and the comoving cosmic distance  $r(z)$ , Equation 1.3.12 can be recast as

$$\mu_0(z) = 5 \log_{10} [(1+z)r(z)/pc] - 5, \quad (1.3.13)$$

a relationship commonly referred to as the Hubble diagram. Figure 1.7 is the most complete SNe Ia-based Hubble diagram to date.

The curve in Figure 1.7 represents the best-fit minimal  $\Lambda$ CDM cosmological model (flat geometry and constant dark-energy equation of state  $w = -1$ ) to

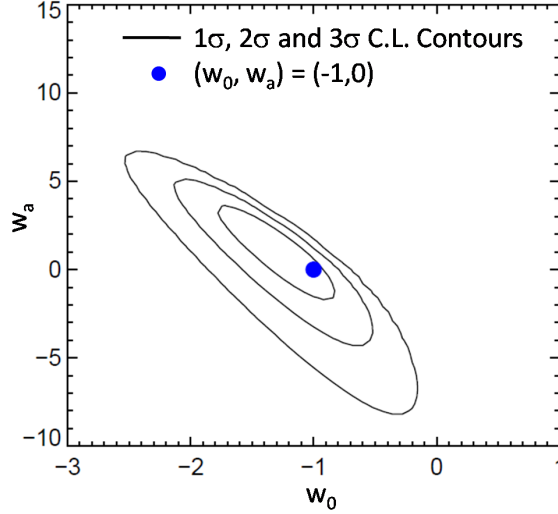


Figure 1.8: Two-dimensional constraint on the dark-energy equation of state from a cosmological-model fit to Type Ia supernovae data, indicating consistency with the minimal  $\Lambda$ CDM model in which  $w_0 = -1$  and  $w_a = 0$ . A nonzero  $w_a$  would indicate that the dark-energy equation of state evolves as a function of redshift. The contours (from small to large) are the  $1\sigma$ ,  $2\sigma$  and  $3\sigma$  confidence level regions. Figure adapted from [114].

the SNe Ia data. The best-fit values for the total amount of matter and dark energy in the Universe are

$$\begin{aligned}\Omega_{\mathcal{M}} &= 0.295^{+0.043}_{-0.040}, \quad \text{and} \\ \Omega_{\Lambda} &= 0.705^{+0.040}_{-0.043},\end{aligned}\tag{1.3.14}$$

where the errors are 68.27% confidence intervals and include both statistical and systematic uncertainties [110]. Comparing these values to those in Table 1.2, the CMB and SNe Ia best-fit  $\Lambda$ CDM models are consistent to within the measurement errors, despite the use of such completely different techniques.

SNe Ia data can also be used to constrain the dark energy's equation of state  $w$ . A fit to SNe Ia data in which  $w$  is allowed to float (a constant possibly different from  $-1$ ) yields somewhat less accurate but consistent values of  $\Omega_{\mathcal{M}}$  and  $\Omega_{\Lambda}$ , and  $w = -1.001^{+0.348}_{-0.398}$ . Though not terribly accurate, such tests of  $w$

### 1.3. OBSERVATIONAL EVIDENCE

are anticipated to be a useful tool for learning more about the nature of dark energy [29]. The problem generalizes further if one allows  $w$  to vary with time. A common parameterization is to expand  $w$  in terms of the scale factor  $a$ :

$$w(a) = w_0 + w_a(1 - a), \quad (1.3.15)$$

where  $w_0$  and  $w_a$  are constants [117]. Rewriting  $w$  in terms of  $z$ , Equation 1.3.10 becomes

$$d_L(z) = \frac{c(1+z)}{H_0} \int_0^z \frac{dz'}{\sqrt{\Omega_\Lambda e^{3W(z')} + \Omega_{\mathcal{M}}(1+z')^3 + \Omega_R(1+z')^4}}, \quad (1.3.16)$$

where  $W(z) = \int_0^z [1 + w_0 + w_a(z'/(1+z'))] [1+z']^{-1} dz'$  [88]. Consequently, if the value of  $w$  depends on redshift, it might be possible to discover a nonzero  $w_a$  with a larger sample of high- $z$  SNe Ia. Current constraints based solely on SNe Ia are rather loose and still consistent with the minimal  $\Lambda$ CDM model for which  $(w_0, w_a) = (-1, 0)$ . An example from [114] in which a modest SNe Ia compilation is used in combination with an assumed value of the matter density ( $\Omega_{\mathcal{M}} = 0.27 \pm 0.03$ ) is given in Figure 1.8. These contours can be tightened by including constraints from other astrophysical evidence (*e.g.*, fits to the CMB power spectrum). In general, combining constraints from multiple types of astrophysical evidence tends to yield more precise cosmological-parameter estimates (see, *e.g.*, Figure 1.1). I explore this further in Section 1.3.5.

#### 1.3.4 Baryon Acoustic Oscillations

Baryon acoustic oscillations (BAO) typically refers to the clustering of nearby galaxies at certain length scales caused by gravitational perturbations in the early Universe. That is, the acoustic peaks in the CMB power spectrum (see, *e.g.*, Figure 1.6) are also visible as baryon overdensities at relatively nearby redshifts in the form of distance and redshift correlations in the spatial distribution of the number of galaxies.

Consider for example a single spherical gravitational fluctuation prior to recombination; the early Universe is well approximated by a linear superposition of many such perturbations. The perturbation begins as a collection of



## CHAPTER 1. INTRODUCTION

similarly sized overdensities of baryons, photons, and dark matter. Prior to recombination ( $z_{\text{rec}} < 1100$ ) the baryons and photons are tightly coupled into a relativistic fluid with a characteristic speed of sound  $c_s$  determined by the ratio of their densities [52]:

$$c_s = \frac{c}{\sqrt{3(1 + R_{b/\gamma})}}, \quad (1.3.17)$$

where  $c$  is the speed of light and

$$R_{b/\gamma} \equiv 3\rho_b/4\rho_\gamma = \frac{3}{4} \frac{\Omega_b}{\Omega_\gamma(1+z)}. \quad (1.3.18)$$

The overdensity of photons causes a build up of radiation pressure that forces a baryon-photon acoustic wave to travel outward in the form of a spherical shell of overdensity, leaving behind a dark-matter core. Following recombination the photon background decouples and the radiation pressure drops; the expanding baryon acoustic wave stalls while the CMB photons continue on. The acoustic wave therefore reaches a maximum radius that is imprinted on the distribution of baryonic matter as a spherical shell of excess density. Over the course of the next few hundred million years, the shell of baryonic matter and the dark-matter core interact gravitationally. Some dark matter is pulled out to the shell boundary, while a large fraction of the baryonic matter is “dragged” back toward the center of the perturbation. This period is commonly referred to as the baryon-drag epoch [118]. By the end of the dark ages ( $z \simeq 10$ ) there is an increased probability that stars and galaxies will form in the center and along the outer layer of the expanded acoustic oscillation, increasing the odds of finding galaxies separated by a distance corresponding to its radius. Consequently, there is a characteristic BAO scale  $s_{\text{bao}}$  that can be used as a “standard ruler” for measuring cosmological parameters since it persists in the distribution of galaxies across a wide range of redshifts.

The BAO scale is set by the size of the sound horizon at recombination;

$$s_{\text{bao}} = \int_{z_{\text{rec}}}^{\infty} \frac{c_s}{H(z)} dz \quad (1.3.19)$$

### 1.3. OBSERVATIONAL EVIDENCE

and can be estimated from CMB-derived cosmological parameters. Assuming that the contribution from dark energy is negligible, the Hubble parameter can be written (approximately) as

$$\begin{aligned} H(z) &\simeq H_0 \sqrt{\Omega_{\mathcal{M}}(1+z)^3 + \Omega_R(1+z)^4} \\ &= \sqrt{\Omega_{\mathcal{M}} H_0^2 (1+z)^3} \sqrt{1 + \frac{1+z}{1+z_{\text{eq}}}}, \end{aligned} \quad (1.3.20)$$

where  $z_{\text{eq}} = \Omega_{\mathcal{M}}/\Omega_R$  is the redshift of matter-radiation equality. Integrating Equation 1.3.19 gives

$$s_{\text{bao}} \simeq \frac{1}{H_0 \sqrt{\Omega_{\mathcal{M}}}} \frac{2c}{\sqrt{3z_{\text{eq}} R_{\text{eq}}}} \ln \left[ \frac{\sqrt{1+R_{\text{rec}}} + \sqrt{R_{\text{rec}} + R_{\text{eq}}}}{1 + \sqrt{R_{\text{eq}}}} \right], \quad (1.3.21)$$

where  $R_{\text{rec}}$  and  $R_{\text{eq}}$  are the baryon-to-photon density ratios at recombination and matter-radiation equality, respectively [119]. Using the values for  $z_{\text{eq}}$  and  $\Omega_b$  from Table 1.2,  $R_{\text{eq}}$  can be estimated from the present-day photon density with Equation 1.3.18.  $\Omega_\gamma$  is known from the blackbody temperature of the CMB; for  $T_{\text{cmb}} = 2.725$  K and  $h = 0.7$ ,  $\Omega_\gamma = 4.988 \times 10^{-5}$  [24], and  $R_{\text{eq}} \simeq 0.22$ . Similarly, based on  $z_{\text{rec}} = 1090$  [24],  $R_{\text{rec}} \simeq 0.63$ . The remaining parameters required to evaluate Equation 1.3.21 are available in Table 1.2 and give a value for  $s_{\text{bao}} \simeq 146$  Mpc.

The true scale is somewhat larger because subsequent to recombination the momentum of the baryons causes the acoustic waves to continue to expand for a short time. The redshift  $z_d$  at which the acoustic waves finally reach their maximum size (corresponding to the beginning of the baryon-drag epoch) is also well determined by the CMB power spectrum;  $z_d \simeq 1020$  [52]. Substituting  $R_d$  for  $R_{\text{rec}}$  in Equation 1.3.21 yields  $s_{\text{bao}} \simeq 153$  Mpc.

Measurements of the clustering of nearby galaxies provide a complementary probe of the BAO scale. Clustering can be measured along ( $\parallel$ ) and across ( $\perp$ ) the line of site, translating into redshift ( $\Delta z$ ) and angular ( $\Delta\theta$ ) scales, respectively. As a function of redshift, the BAO scale is given by  $s_{\text{bao}} = c\Delta z/H(z)$  [120]. The redshift clustering scale  $\Delta z$  therefore constrains the product of  $s_{\text{bao}}$  and  $H$ . To understand the angular BAO measurement the

## CHAPTER 1. INTRODUCTION

concept of angular diameter distance  $d_A$  is useful. If an object of size  $D$  is aligned perpendicularly to our line of sight and subtends an angle  $\theta$  at our point of observation,  $d_A$  is simply the ratio of  $D$  to  $\theta$ . In general,  $d_A$  depends on the cosmology of the Universe and is related to the comoving and luminosity distances (Equation 1.3.8);

$$d_A(z) = \frac{d_L(z)}{(1+z)^2} = \frac{r(z)}{(1+z)}. \quad (1.3.22)$$

In terms of  $d_A$ , the BAO scale is  $s_{\text{bao}} = d_A(1+z)\Delta\theta$  [120]. The angular clustering scale  $\Delta\theta$  therefore constrains the ratio of  $s_{\text{bao}}$  to  $d_A$ . The combined measurement of the redshift and angular scales (in the form of  $\Delta z/\Delta\theta$ ) constrains the product of  $H$  and  $d_A$ , and is known as an Alcock-Paczynski (AP) test [121]. At nearby redshifts ( $z < 1$ ) this product has very little sensitivity to the radiation density and is given approximately by (for a flat geometry with constant dark-energy equation of state  $w = -1$ )

$$H(z)d_A(z) \simeq \frac{c\sqrt{1+3\Omega_{\mathcal{M}}z+\mathcal{O}(z^2)}}{1+z} \int_0^z \frac{dz'}{\sqrt{1+3\Omega_{\mathcal{M}}z'+\mathcal{O}((z')^2)}}, \quad (1.3.23)$$

which depends entirely on  $\Omega_{\mathcal{M}}$ . More generally, AP tests are also sensitive to the geometry of spacetime and the properties of the dark energy [119]. However, dark-energy sensitivity requires relatively high- $z$  BAO observations for which it is considerably more difficult to acquire large data sets [122, 123]. At present, it is the AP test's  $\Omega_{\mathcal{M}}$  sensitivity through low- $z$  measurements of galaxy clustering that are the most cosmologically significant.

At nearby redshifts the BAO scale manifests itself as an increased probability (relative to a random distribution) for finding two galaxies in the sky separated by  $\sim 153$  Mpc. A common method for searching for this  $\sim 10\%$  effect is to measure the two-point correlation function  $\xi(r)$ , a statistical measure of the excess clustering of baryonic matter on a given scale relative to a uniform distribution of matter with the same average density. To better understand the meaning of  $\xi$ , consider a fluctuation in the baryon density  $\delta(\vec{x}) = (\rho(\vec{x}) - \rho_b)/\rho_b$ , where  $\rho_b$  is the average density and  $\rho(\vec{x})$  is the density at spatial location  $\vec{x}$ .

### 1.3. OBSERVATIONAL EVIDENCE

$\delta(\vec{x})$  represents the fractional deviation of the baryon density as a function of spatial position. The two-point correlation function is then simply an average over all space of this density times itself displaced by some distance  $\Delta\vec{x}$ ;

$$\xi = \langle \delta(\vec{x})\delta(\vec{x} + \Delta\vec{x}) \rangle. \quad (1.3.24)$$

Spatial homogeneity ensures that the two-point correlation function depends on only the spatial separation  $\Delta\vec{x}$  and not the particular position  $\vec{x}$ , while the spherical symmetry implied by spatial isotropy removes the angular dependence in  $\Delta\vec{x}$ .  $\xi$  is therefore a function of the scalar separation  $|\Delta\vec{x}|$ , which I will simply call  $r$ . Since  $r$  can extend either along or across the line of site,  $\xi$  is commonly referred to as the spatial-redshift correlation function.

There have been a number of attempts (with mixed results) to measure the spatial-redshift correlation function using spectrographic (multicolor) galaxy surveys (see, *e.g.*, [124, 125, 126, 127, 128]). The earliest surveys revealed that  $\xi$  is well approximated by a power law at short distance scales:

$$\xi(r) = \left(\frac{r_1}{r}\right)^\zeta, \quad (1.3.25)$$

where  $r_1 \simeq 7.5 \text{ Mpc}$  and  $\zeta \simeq 1.8$  [18]. Evidence for acoustic oscillations at scales larger than  $\sim 15 \text{ Mpc}$  was at first subtle, with  $\xi$  deviating only slightly from Equation 1.3.25 in a manner sensitive to the value of  $\Omega_{\mathcal{M}}$  [129].

The expected  $\sim 153 \text{ Mpc}$  feature in the correlation function was not found until the advent of two very large, modern photometric galaxy surveys: the Sloan Digital Sky Survey (SDSS) [130] and the Two-degree Field Galaxy Redshift Survey (2dFGRS) [131]. The goal of these kinds of surveys is to identify and measure the redshifts for as many galaxies as possible, over as much of the sky as possible. The SDSS accomplishes this with an automated 2.5 m telescope at the Apache Point Observatory in New Mexico. With a  $3^\circ$  field of view capable of simultaneous observations of  $\gtrsim 600$  objects at a time, it operates in a search-followup mode similar to the technique used to find SNe Ia; the sky is scanned to locate galaxies and then followup photometric observations are

## CHAPTER 1. INTRODUCTION

made with a spectrograph to measure their redshifts. The SDSS project intends to eventually characterize roughly one million galaxies over  $\sim 10,000 \text{ deg}^2$  (out of  $41,253 \text{ deg}^2$ ), or  $1/4$  of the sky. The 2dFGRS used the  $2^\circ$  degree field of view on the Anglo-Australian Telescope (capable of  $\sim 400$  simultaneous observations) and a similar technique to obtain redshifts for  $\sim 250,000$  galaxies over approximately  $1500 \text{ deg}^2$ .

Using a subset of nearly 50,000 luminous red galaxies (LRG) from the SDSS main galaxy sample [132], the authors in [129] find clear evidence for a peak in the spatial-redshift correlation function between roughly 140 and 155 Mpc. Their LRG sample includes redshifts from 0.16 to 0.47 and covers nearly  $4000 \text{ deg}^2$ .  $\xi$  as a function of comoving distance (in terms of the dimensionless Hubble constant  $h$ ) for their LRG selection is compared to a few different cosmological models in Figure 1.9. The models with both baryons and dark matter follow the data at all scales, while a model with no baryonic matter (and therefore no acoustic oscillations) fails to reproduce the peak.

More recently the authors in [118] have derived improved BAO-based constraints by combining a truly staggering number of galaxies from the 2dFGRS catalog and an updated SDSS catalog. They select  $\sim 700,000$  nearby galaxies ( $\langle z \rangle = 0.12$ ) from the seventh SDSS data release [133], as well as  $\sim 80,000$  LRGs covering a range of redshifts from 0.2 to 0.5. To increase galaxy statistics for  $z < 0.3$ , nearly 150,000 galaxies from the 2dFGRS catalog are also included. In total, just under 900,000 galaxies covering  $9100 \text{ deg}^2$  of the sky were analyzed. The data are presented in terms of the residual power spectrum of galaxy density fluctuations in overlapping redshift slices in Figure 1.10. The power spectrum is simply the Fourier transform of the correlation function:

$$P(k_\lambda) = \int d^3r \xi(r) \exp\left(-i\vec{k}_\lambda \cdot \vec{r}\right), \quad (1.3.26)$$

where  $k_\lambda = |\vec{k}_\lambda|$  is the wavenumber and is typically measured in units of  $h \text{ Mpc}^{-1}$ . When expressed in terms of  $P(k_\lambda)$ , the peak in Figure 1.9 manifests as a modulation with a characteristic wavelength of  $\sim 2\pi/(153 \text{ Mpc}) \simeq 0.06 h \text{ Mpc}^{-1}$ . This wavenumber modulation is visible in Figure 1.10 and com-

### 1.3. OBSERVATIONAL EVIDENCE

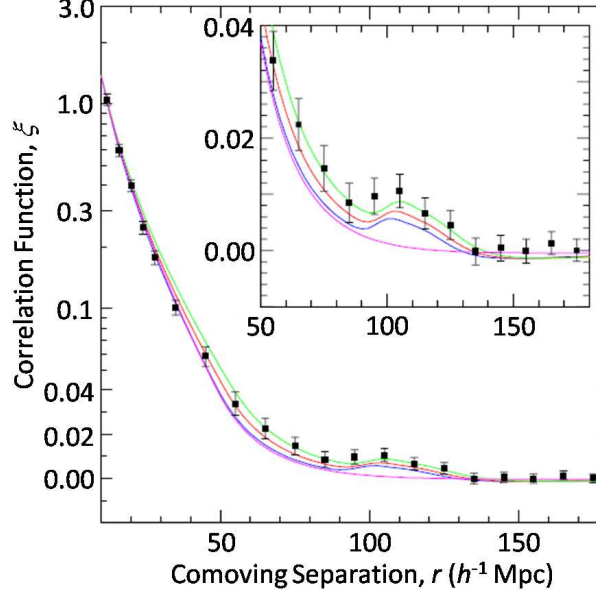


Figure 1.9: The spatial-redshift correlation function as a function of comoving-distance scale (in terms of the dimensionless Hubble constant  $h$ ) for a subset of the Sloan Digital Sky Survey main galaxy sample [132]. The characteristic baryon acoustic oscillations scale is clearly visible as a peak at  $\gtrsim 100 h^{-1}$  Mpc. The curves correspond to  $\Lambda$ CDM models with differing baryon and dark-matter densities. The top three have  $\sim 5\%$  baryons and  $\Omega_{\mathcal{M}}h^2 = 0.12$  (top, green),  $\Omega_{\mathcal{M}}h^2 = 0.13$  (red), and  $\Omega_{\mathcal{M}}h^2 = 0.14$  (bottom with peak, blue), while the bottom curve without the peak has no baryons and  $\Omega_{\mathcal{M}}h^2 = 0.105$ . Clearly, both baryons and dark matter are required to reproduce the data. Note that the vertical scale is both logarithmic (top two thirds with 2-digit axis labels) and linear (bottom third with 3-digit axis labels). The inset gives a zoomed in view of the linear region. Figure adapted from [129].

compares favorably to the  $\Lambda$ CDM model shown (for which  $h = 0.72$ ,  $\Omega_b = 0.043$  and  $\Omega_{\mathcal{M}} = 0.25$ ). The data in both Figures 1.9 and 1.10 show clear signs of the  $\sim 153$  Mpc BAO scale in the clustering of nearby galaxies and are consistent with a significant nonbaryonic contribution to  $\Omega_{\mathcal{M}}$ .

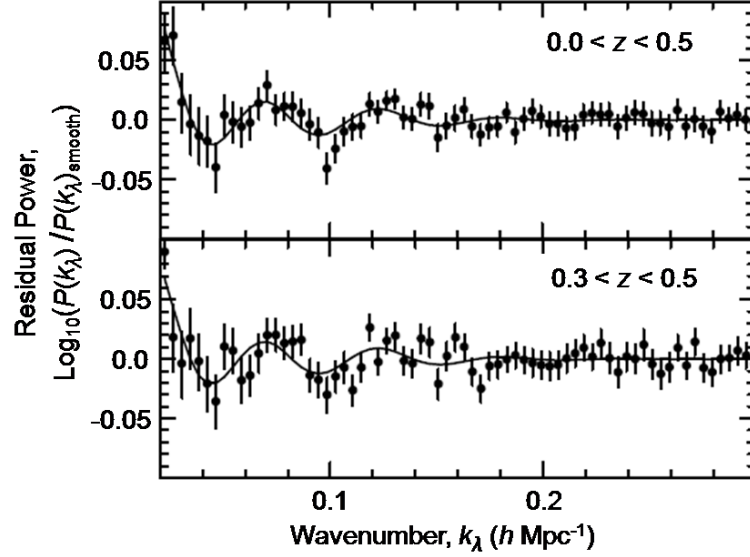


Figure 1.10: The residual power spectrum of density fluctuations in the spatial-redshift distribution of nearby galaxies compared to the power spectrum for a  $\Lambda$ CDM cosmological model with  $\Omega_b = 0.043$ ,  $h = 0.72$ , and  $\Omega_{\mathcal{M}} = 0.25$ . The smooth component of the  $\Lambda$ CDM model has been divided out to make the modulation due to the characteristic scale of baryon acoustic oscillations more apparent. The upper panel shows the residuals for  $\sim 900,000$  galaxies with redshifts between 0 and 0.5 selected from the Sloan Digital Sky Survey [133] and Two-degree Field Galaxy Redshift Survey [131] catalogs, while the lower panel shows the residuals for an  $\sim 70,000$  galaxy subset with redshifts between 0.3 and 0.5. Figure adapted from [118].

### 1.3.5 Cosmic Concordance

Neither SNe Ia data nor observations of the BAO scale are individually sensitive enough to fully constrain the minimal  $\Lambda$ CDM cosmological model. As discussed in the previous two sections, BAO data are primarily sensitive to  $\Omega_{\mathcal{M}}$ , whereas SNe Ia data are primarily sensitive to the properties of the dark energy. Fits to the CMB power spectrum are capable of fully constraining all six  $\Lambda$ CDM parameters, but only with precision when zero spatial curvature ( $\Omega_K = 0$ ) and a constant dark-energy equation of state ( $w = -1$ ) are

### 1.3. OBSERVATIONAL EVIDENCE

assumed. To probe these assumptions and constrain additional cosmological parameters, it is common practice to combine constraints derived from several data sources. The resulting parameter estimates are not only more precise, but also serve as powerful consistency checks in those cases where the parameters are overconstrained.

At the time of this writing the most up-to-date and  $\Lambda$ CDM-sensitive astrophysical measurements are:

- **CMB:** The CMB anisotropy power spectrum derived from the WMAP 7-year data and shown in Figure 1.6 (Larson *et al.* 2011 [76]).
- **SNe:** The Hubble diagram shown in Figure 1.7 for the Union2.1 compilation of Type Ia supernovae (Suzuki *et al.* 2011 [110]).
- **BAO:** The residual power spectrum of nearby density fluctuations derived from a 2dFGRS and SDSS combined galaxy sample and shown in Figure 1.10 (Percival *et al.* 2010 [118]).
- **$H_0$ :**  $SH_0ES$  II program estimate of the present-day value of the Hubble parameter derived from Cepheid-calibrated nearby SNe Ia and shown in Figure 1.1 (Riess *et al.* 2011 [25]).

In the context of the minimal  $\Lambda$ CDM model, combining all four yields estimates of the total matter and dark-energy densities that are twice as precise as the CMB-only estimates in Table 1.2, and nearly a factor of three better than the SNe Ia-only estimates in Equation 1.3.14:

$$\begin{aligned}\Omega_{\mathcal{M}} &= 0.271 \pm 0.014, \text{ and} \\ \Omega_{\Lambda} &= 0.729 \pm 0.014,\end{aligned}\tag{1.3.27}$$

where the errors represent  $1\sigma$  confidence intervals.<sup>8</sup> The individual and combined  $\Omega_{\Lambda}$  versus  $\Omega_{\mathcal{M}}$  constraints are shown in Figure 1.11 and demonstrate an astonishing consistency that is sometimes referred to as cosmic concordance [134]; a mutually consistent overlap region was not guaranteed *a priori*.

---

<sup>8</sup>Throughout this thesis all errors represent  $1\sigma$  confidence intervals unless otherwise stated.



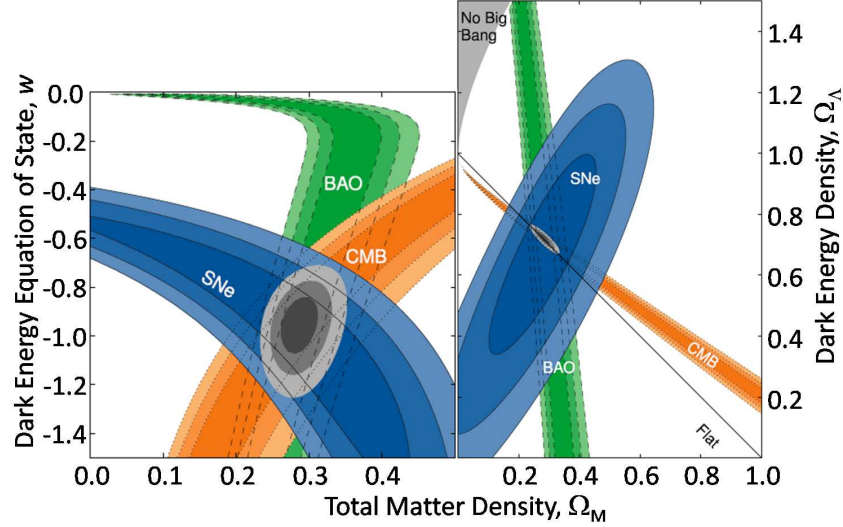


Figure 1.11: Individual (colored contours) and combined (gray contours) cosmological constraints based on cosmic microwave background (CMB, orange) [76], Type Ia supernovae (SNe, blue) [110], and baryon acoustic oscillations (BAO, green) [118] data. The right panel shows minimal  $\Lambda$ CDM constraints in the  $\Omega_\Lambda$  versus  $\Omega_M$  plane, while the left panel shows  $w$ CDM constraints in the  $w$  versus  $\Omega_M$  plane. The contours represent (big to small) 99.7%, 95.4% and 68.3% confidence regions. Figure adapted from [110].

Combining the CMB, BAO and  $H_0$  results yields improved estimates of the baryon and cold dark-matter densities [24]:

$$\begin{aligned}\Omega_b &= 0.0458 \pm 0.0016, \quad \text{and} \\ \Omega_c &= 0.229 \pm 0.015.\end{aligned}\tag{1.3.28}$$

Ignoring the  ${}^7\text{Li}$  anomaly, Equation 1.3.28 is also consistent with the value of  $\Omega_b$  derived from comparing BBN predictions with measurements of the primordial abundances of light nuclei (Figure 1.4). As evidence for a significant cold dark-matter density, Equations 1.3.27 and 1.3.28 and Figure 1.4 are extremely compelling.

As discussed toward the ends of Sections 1.2.1 and 1.3.3, it is possible that the dark-energy equation of state is neither constant nor exactly equal to  $-1$ .

### 1.3. OBSERVATIONAL EVIDENCE

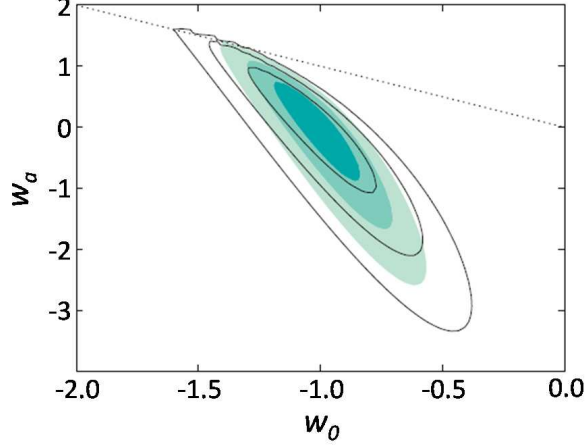


Figure 1.12:  $w_z$ CDM  $w_a$  versus  $w_0$  combined constraints based on cosmic microwave background [76], Type Ia supernovae [110], baryon acoustic oscillations [118], and  $H_0$  [25] data, with (solid black contours) and without (shaded contours) SNe Ia systematic errors. The region above the dotted line ( $w_0 + w_a > 0$ ) violates early matter domination. The contours represent (big to small) 99.7%, 95.4% and 68.3% confidence regions. Figure adapted from [110].

Furthermore, it might also be true that the Universe has a nonzero spatial curvature. The CMB, SNe, BAO and  $H_0$  data can be combined to constrain  $w_0$ ,  $w_a$  and  $\Omega_K$  using the following, slightly extended  $\Lambda$ CDM models:

- $o\Lambda$ CDM:  $\Omega_K$  is allowed to be nonzero ( $w_0 = -1$  and  $w_a = 0$ ).
- $w$ CDM:  $w$  is a constant not necessarily equal to  $-1$  ( $\Omega_K = 0$  and  $w_a = 0$ ).
- $w_z$ CDM:  $w$  is allowed to vary with time ( $\Omega_K = 0$ ).
- $ow_z$ CDM:  $\Omega_K$  is allowed to be nonzero, and  $w$  can vary with time.

The  $o\Lambda$ CDM and  $w$ CDM constraints are entirely consistent with the minimal  $\Lambda$ CDM model and yield  $\Omega_K = 0.002 \pm 0.005$  and  $w_0 = -1.013^{+0.068}_{-0.073}$ , respectively [110]. The combined and individual  $w$  versus  $\Omega_{\mathcal{M}}$   $w$ CDM constraints are shown in Figure 1.11. The  $w_z$ CDM constraints are also consistent with the minimal  $\Lambda$ CDM model and show no evidence for a time varying dark-energy

## CHAPTER 1. INTRODUCTION

equation of state. Figure 1.12 gives the  $w_a$  versus  $w_0$  contours in the context of the  $w_z$ CDM model. The combined  $ow_z$ CDM constraints provide the first hint that the minimal  $\Lambda$ CDM model's curvature and dark energy assumptions might be invalid [110]:

$$\begin{aligned}\Omega_K &= 0.027^{+0.012}_{-0.011}, \\ w_o &= -1.198^{+0.100}_{-0.112}, \text{ and} \\ w_a &= 1.19^{+0.13}_{-0.13}.\end{aligned}\tag{1.3.29}$$

While intriguing, these estimates are slightly misleading. The authors in [110] note that in some cases the confidence intervals have significant non-Gaussian tails. For example, while the  $1\sigma$  lower limit on  $w_a$  is well above zero at 1.06, the  $2\sigma$  lower limit is well below zero at -1.21. The  $ow_z$ CDM constraints are all within  $2\sigma$  of the minimal  $\Lambda$ CDM parameters and are therefore consistent with statistical fluctuations.

# Chapter 2

## Dark Matter

The Standard Cosmology and supporting observational evidence discussed in the previous chapter argue strongly in favor of the existence of dark matter, but suggest very little regarding its exact nature; the dark matter has a significant nonbaryonic component ( $\Omega_c$ ) that is largely nonrelativistic at recombination and does not couple very strongly with baryons or photons in the early Universe. Aside from (maybe) neutrinos (a possibility addressed in Section 2.4.2), there are clearly no stable Standard Model particles that meet these requirements, begging the question: what exactly is the dark matter? This unresolved mystery is generally referred to as the dark-matter problem.

Historically, the dark-matter problem predates the emergence of the Standard Cosmology, going back to the 1930s and Fritz Zwicky's early observations of the Coma galaxy cluster [135, 136]. In this chapter I will discuss this as well as a few other types of observational evidence that have traditionally been associated with the dark-matter problem. While these dark-matter-specific forms of evidence can be used for only very rough estimates of the amount of dark matter, they tend to be more direct and less model dependent than those of the previous chapter. They also provide fairly convincing evidence that dark matter populates our galaxy and solar neighborhood, where it might be detectable by terrestrial detectors.

In the second half of the chapter I survey potential solutions to the dark-matter problem, including a few of the most popular dark-matter candidates.

Despite the nearly irrefutable cosmological evidence that the dark matter is primarily nonbaryonic, both baryonic and nonbaryonic explanations are explored. The chapter ends by introducing the Weakly Interacting Massive Particle (WIMP) [137], followed by a review of supersymmetry as a theory beyond the SM that naturally contains a WIMP candidate, the neutralino [2, 138, 139].

## 2.1 Mass-to-Light Ratios

The oldest and perhaps most fundamental argument for the existence of dark matter is based on observations of mass-to-light ratios. The mass-to-light ratio  $\Upsilon$  is the ratio of a system's mass  $\mathcal{M}$  to its luminosity  $L$  and is typically measured in solar units:

$$\Upsilon_{\odot} \equiv \frac{\mathcal{M}_{\odot}}{L_{\odot}}, \quad (2.1.1)$$

where  $\mathcal{M}_{\odot} = 1.99 \times 10^{30}$  kg and  $L_{\odot} = 3.85 \times 10^{26}$  W are the solar mass and luminosity, respectively [3].  $\Upsilon$  measurements depend strongly on the distance scales over which they are measured [140], with compact systems (*e.g.*, the solar neighborhood and galactic cores) tending to yield lower values than the largest gravitationally bound objects (*e.g.*, galaxy clusters). The usual interpretation is that the smaller objects are dominated by their visible-matter densities (*e.g.*, baryons in stars and interstellar gas), while the larger systems have an additional dark component (*e.g.*, nonbaryonic dark matter) contributing to their matter densities but not to their luminosities; there is more mass than can be seen.

At first glance, any discrepancy between large- and small-scale  $\Upsilon$  measurements might seem contradictory. All light-emitting structures within a  $\Lambda$ CDM Universe form from dark-matter density perturbations and should therefore include both visible and dark-matter components. The resolution to this apparent paradox has to do with how dark matter is spatially distributed relative to visible matter. As will be discussed in Section 2.2.1, dark matter appears to occupy very large volumes (*e.g.*, spherical halos hundreds of kiloparsecs across), whereas visible matter tends to concentrate into more compact objects (*e.g.*,

## 2.1. MASS-TO-LIGHT RATIOS

galactic discs tens of kiloparsecs across). Consequently, on short distance scales the dark matter is effectively a homogeneous energy density that is so sparsely distributed relative to visible matter that it does not factor significantly into mass (or density) estimates. Mass-to-light ratios for smaller systems are indicative of the Universe’s visible-matter density and can provide only lower bounds on the total matter density  $\Omega_{\mathcal{M}}$ . As larger and larger volumes are considered, the enclosed dark-matter density becomes increasingly influential, and the corresponding mass-to-light ratios are more representative of  $\Omega_{\mathcal{M}}$ . Differences between large- and small-scale  $\Upsilon$  measurements can therefore be used to infer the presence and rough amount of dark matter in the Universe’s largest structures.

To this end, it is useful to define the critical mass-to-light ratio  $\Upsilon_c$ , the ratio of the critical density  $\rho_c$  to the average luminosity density of the Universe. Introduced in Section 1.2.1, the critical density is

$$\rho_c \equiv \frac{3H_0^2}{8\pi G_N} = 2.77 \times 10^{11} h^2 \mathcal{M}_\odot \text{ Mpc}^{-3} \simeq 6 \text{ keV}/c^2/\text{cm}^3, \quad (2.1.2)$$

where  $h = 0.738$  is the dimensionless Hubble constant. The average luminosity density can be measured from galaxy surveys like those mentioned in Section 1.3.4 (see [18] for a detailed discussion). Based on a subset of nearby ( $z = 0.1$ ) galaxies from the SDSS catalog, the average luminosity density is [141]

$$\mathcal{L} = (1.84 \pm 0.04) \times 10^8 h L_\odot \text{ Mpc}^{-3}. \quad (2.1.3)$$

Hence, the critical mass-to-light ratio in units of  $\Upsilon_\odot$  is

$$\Upsilon_c = \frac{\rho_c}{\mathcal{L}} \simeq 1500 h \Upsilon_\odot. \quad (2.1.4)$$

Comparing a system’s mass-to-light ratio to this reference value yields an estimate of the Universe’s matter density under the assumption that the system is representative of the Universe as a whole;

$$\Omega_{\mathcal{M},\text{system}} = \frac{\rho_{\text{system}}}{\rho_c} = \Upsilon_{\text{system}}/\Upsilon_c. \quad (2.1.5)$$

## CHAPTER 2. DARK MATTER

This technique for estimating the mass-density parameter is known as the Oort method [142, 143]. Again, due to differences in the spatial distributions of visible and dark matter, Oort-method estimates are typically lower limits. Furthermore, the method is highly sensitive to the value of  $\mathcal{L}$ , and the galaxies from which  $\mathcal{L}$  is estimated are (somewhat naïvely) assumed to fairly sample the luminosity of the Universe without regard to their surrounding environments. Despite these shortcomings, the Oort method still provides a useful tool for inferring the presence of dark matter. The precise values obtained for  $\Omega_{\mathcal{M}}$ , however, should be taken with a grain of salt.

The mass-to-light ratio for the local stellar neighborhood, for which matter and luminosity densities can be measured specifically for the stellar component, provides a rough estimate of the Universe’s luminous-matter density. The stars and gas within  $\sim 1$  kpc of the Sun form a column of stellar and gaseous baryonic matter that spans the Milky Way’s galactic disc and is commonly referred to as the “solar cylinder.” Integrating the volumetric mass and luminosity densities along the height of the solar cylinder yields surface densities for the local galactic disc. The methodology for extracting these surface densities from luminosity and astrometry catalogs of nearby stars is a well developed science (see, *e.g.*, [144, 145, 146]). Thanks primarily to improved stellar distance measurements made by the Hipparcos satellite [147] and Hubble Space Telescope [27], the surface mass density of the solar cylinder is known to within  $\sim 10\%$  of  $50 \mathcal{M}_{\odot} \text{pc}^{-2}$  [148]. The gas fraction ( $\sim 14 \mathcal{M}_{\odot} \text{pc}^{-2}$  [149]) is thought to be particular to our local neighborhood and therefore not necessarily representative of baryons at large. The stellar component ( $\sim 36 \mathcal{M}_{\odot} \text{pc}^{-2}$ ), however, can be considered more universal and has a surface luminosity density of  $\sim 24 L_{\odot} \text{pc}^{-2}$  [150]. The luminous mass-to-light ratio for the solar neighborhood is therefore  $\Upsilon_{*} \simeq 1.5 \Upsilon_{\odot}$ , where the  $*$  denotes stellar material. Assuming  $\Upsilon_{*}$  is representative of luminous matter everywhere, we can extrapolate to the Universe as a whole to obtain an estimate of its luminous-matter density:

$$\Omega_{*} = \frac{\Upsilon_{*}}{\Upsilon_c} \simeq 0.001 h^{-1}, \quad (2.1.6)$$

or  $\gtrsim 0.1\%$  of the critical density. Although baryons comprise about 5% of the

## 2.1. MASS-TO-LIGHT RATIOS

Table 2.1: Mass-to-light ratio ( $\Upsilon$ ) ranges are listed for several classes of differently sized objects in units of the solar mass-to-light ratio  $\Upsilon_\odot$  and (in most cases) in terms of the dimensionless Hubble constant  $h$ . Each system’s corresponding matter density relative to the critical density is also listed. “Luminous Cores” refers to the bright central parts of galaxies, while “Small Groups” refers to bound objects consisting of a small number of galaxies like binary systems.  $\Upsilon$  values taken from [151].

System	Mass-to-Light Ratio, $\Upsilon/\Upsilon_\odot$	$\Omega_{\mathcal{M},\text{system}} = \Upsilon/\Upsilon_c$
Solar Neighborhood	1–3	0.0009–0.0027
Luminous Cores	(10–20) $h$	0.0067–0.0133
Elliptical Galaxies	(30–200) $h$	0.02–0.13
Small Groups	(60–180) $h$	0.04–0.12
Galaxy Clusters	(200–500) $h$	0.13–0.33

Universe’s energy density, only a small fraction are bound into stars and are actively luminous. Even assuming a 100% error on  $\Upsilon_*$ , Equations 1.3.28 and 2.1.6 indicate that  $\gtrsim 95\%$  of the Universe’s baryons are dark. As we will see later in this chapter, many are found in clouds of interstellar and intergalactic gas. These baryons are not truly dark thanks to a variety of electromagnetic interactions that result in observable photons, often at nonoptical wavelengths and emitted in far lower levels than stellar photons. However, when interpreting astrophysical evidence that directly suggests the gravitational influence of dark matter, it is prudent to consider dark baryons as a possible (even if only partial) explanation.

The mass-to-light ratios for larger systems vary depending on their sizes as well as the methods used to estimate their masses and luminosities. The details for some of these methods are explored for galaxies and galaxy clusters in Sections 2.2 and 2.3, respectively. Table 2.1 summarizes approximate



$\Upsilon$  ranges and corresponding matter densities (relative to the critical density) for a variety of object classes. If these systems were composed primarily of baryonic matter, one would expect the mass-to-light ratios to be more similar. The small- to large-scale progression in the inferred matter densities clearly indicates an increasingly significant dark component. Considering the combined cosmological constraint on  $\Omega_b$  from the previous chapter (Equation 1.3.28), the case for nonbaryonic dark matter in galaxies and clusters is almost undeniable.

## 2.2 Galactic Dark Matter

The formation of structure in the Standard Cosmology suggests (somewhat indirectly) that gravitationally bound systems grow from dark-matter density fluctuations in the early Universe. The most ubiquitous visible structures (larger than stars) at the present epoch are galaxies, numbering well over 100 billion and visible in all directions in the sky. If the Standard Cosmology is an accurate description of the Universe, in addition to their stellar and gaseous content, galaxies should harbor significant (even dominant) dark-matter components. Evidence for galactic dark matter (specifically Milky Way dark matter) is critical if the dark matter is to be detected by terrestrial detectors. In this section I review the observational evidence that the luminous portions of galaxies are embedded in dark-matter halos. Where appropriate, specific evidence for the Milky Way will be discussed.

Although dark matter cannot (yet) be observed directly, the principle behind inferring its presence in galaxies is simple. Imagine a gravitational potential resulting from a system of bound matter. The depth and extent of the potential depend on both the amount and spatial distribution of the system's matter content. If a test particle is placed within the potential and allowed to gravitationally stabilize into an orbit at some distance  $r$  from the center, it will rotate with a velocity that depends on the total amount of matter  $\mathcal{M}(r)$

## 2.2. GALACTIC DARK MATTER

contained within the sphere defined by the orbit;

$$v(r) = \sqrt{\frac{G_N \mathcal{M}(r)}{r}}. \quad (2.2.1)$$

The more matter (whether luminous or dark) that is distributed interior to the test particle's orbit, the greater its velocity. The distribution of matter in a gravitational potential can therefore be inferred by observing the velocities of test particles as a function of their distances from the center of the system. In the case of galaxies, stars and gas can be used as test particles; if dark matter makes a significant contribution to a galactic gravitational potential, it should be detectable in the galaxy's stellar and gaseous velocities.

Inferring the presence of dark matter in galaxies is (in practice) a bit more complicated than the procedure outlined above. I will restrict the following discussion to the two most common galaxy morphologies: spiral galaxies (like the Milky Way) and elliptical galaxies. In each case I review techniques for probing galactic gravitational potentials, summarizing with mass-to-light ratio estimates for comparison with the values in Table 2.1.

### 2.2.1 Spiral Galaxy Rotation Curves

#### Pioneering Work

The earliest evidence for the presence of dark matter in spiral galaxies is based on spectrographic observations of gas in the nearby Andromeda Galaxy (also called M31) by Vera Rubin and Kent Ford in the early 1970s [152, 153]. M31 was easily the best candidate for these early studies. While a detailed understanding of the Milky Way would be ideal, observations of the Milky Way are more difficult due to our position within it. M31 is close enough for the ground-based (optical) telescopes of the 1970s (and earlier) to spatially resolve individual regions, making it possible to map its velocity field as a function of position. M31 is also a spiral galaxy whose mass is similar to the Milky Way. Any conclusions based on one should therefore (roughly) extend to the other.

## CHAPTER 2. DARK MATTER

Rubin and Ford’s measurements rely on spectroscopy of specific optical lines (*e.g.*,  $H\alpha$  and  $H\beta$  emission due to electronic transitions in hydrogen) from several dozen distinct regions (*e.g.*, H II gaseous regions) that are spatially distributed along M31’s visible disc and beyond. If a region is moving toward (away from) our point of observation, the optical lines observed from it are blueshifted (redshifted) in proportion to its velocity. After correcting for the motion of M31 as a whole—moving toward the Milky Way on average—and any nonradial component, each region’s circular velocity is plotted as a function of its distance from the center of the galaxy. The resulting relation is commonly referred to as a galactic rotation curve. Rubin and Ford’s M31 rotation curve is shown in Figure 2.1.

The interesting feature of Figure 2.1 is that the circular velocities appear to approach a constant very far from the center of the M31 ( $\gtrsim 15$  kpc); the rotation curve is said to be flat. This is intriguing considering the regions furthest from the center of M31 are well separated from its visible disc and bulge. The discs and bulges of spirals are their dominant visible structures and tend to fall off exponentially as a function of distance from the galactic center [154]. Judging from Figure 1 in [152], M31’s visible matter becomes subdominant at a distance of  $\sim 12$  kpc from its center. If M31’s luminous matter were the only contributor to its gravitational potential, the rotation curve would exhibit a  $1/\sqrt{r}$  “Keplerian” fall-off in accordance with Equation 2.2.1. Instead, the rotational velocities of M31’s large-radius gaseous regions appear to be elevated, implying (somewhat marginally) the presence of a dark-matter component.

To be fair, Rubin and Ford’s optical measurements were preceded by low-resolution radio-telescope observations of the 21 cm hydrogen line [155, 156]. The lowest electron-orbital state of neutral hydrogen exhibits hyperfine splitting due to the spins of the proton and electron. The state in which their spins are aligned has slightly more energy—equivalent to a photon with a 21 cm wavelength—than the ground state in which their spins are antiparallel. Although the spin-flip transition from one hyperfine level to the other is highly suppressed, it is easily observable by radio telescopes due to the large quantities

## 2.2. GALACTIC DARK MATTER

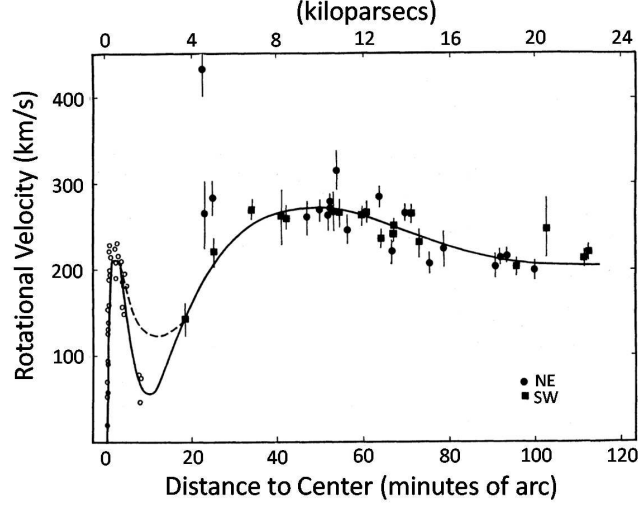


Figure 2.1: Early rotation curve for the Andromeda Galaxy (M31) in which the rotational velocities of several dozen distinct gaseous regions are plotted as a function of distance to the center of the galaxy in units of minutes of arc (kiloparsecs) along the bottom (top). The open circles are derived from a narrow N II emission line (658.3 nm), while the filled in squares and circles are primarily derived from H $\alpha$  emission. The NE versus SW distinction between the filled circles and squares indicates the direction away from the center of M31 along which the corresponding gaseous regions were observed. The solid and dashed curves are polynomial fits to the data. Note how the fits turn up at large distances (as required by the data), indicating a flat rotation curve and hinting at the presence of dark matter. Figure taken from [152].

of neutral hydrogen in the Universe and, in particular, in galaxies. The low probability of the transition gives the 21 cm line a very narrow natural width, making it ideal for extracting velocities due to doppler shifts. The angular resolution of radio telescopes prior to Rubin and Ford's measurements were still too crude to compete with their optical counterparts. Early M31 rotation curves derived from 21 cm observations generally agreed with the data in Figure 2.1, but were too coarse to rule out a Keplerian fall-off. However, radioastronomy has come a long way in the past several decades; modern rotation curves rely heav-

ily on the 21 cm technique when optical wavelengths are obscured or too weak, as is the case far from galactic centers as well as for much of the Milky Way.

## The Universal Rotation Curve

Although the early evidence for galactic dark matter was tentative and largely qualitative, it spawned a rich and active field of research into the nature and distribution of matter (both luminous and dark) in galaxies (see, *e.g.*, [157, 158, 159, 160]). While there is some disagreement regarding modeling techniques, irregular cases and specific conclusions (see, *e.g.*, [161, 158, 162], respectively), galactic rotation curves (typically) exhibit a few common features: 1) their profiles depend primarily on total galactic luminosity; 2) they are (approximately) flat well beyond the edge of the galactic disc, implying the presence of a dark halo (DH); and 3) interior to the edge of the visible disc, they are dominated by the DH for the lowest-luminosity spirals and by the disc for the highest-luminosity spirals. As a case study, the following discussion focuses on Persic and Salucci’s 1996 seminal work on universal rotation curves (URC) for spiral galaxies [163].

The dominant visible feature of (most) spiral galaxies is a thin, luminous disc (LD) of stars and gas that rotates about a central axis perpendicular to the disc.<sup>1</sup> If  $r$  is the galactocentric radius, the disc of a typical spiral exhibits a sur-

---

<sup>1</sup>A common spiral harbors a variety of visible structures that are usually subdominant (in terms of mass) to its luminous disc. Most contain a tightly packed—typically no more than a few kiloparsecs in diameter—central halo of older stars referred to as a galactic bulge and thought to be the result of galaxy-galaxy interactions. On average, galactic bulges contribute negligibly to a spiral’s rotation curve at large radii; for simplicity bulges are not included in this section’s URC discussion. However, they can be important when interpreting an individual spiral’s rotational velocities. Some spirals also possess a central, gravitationally unstable bar-like structure that is a roughly linear concentration of stars and gas spanning a few kiloparsecs in the plane of the disc. A bar can be treated as a perturbation to a spiral’s disc [164], a perturbation that is neglected here because it should smooth out when averaging over an ensemble of galaxies. The Milky Way appears to have a bar [165] that may be responsible for the dip in its rotation curve at  $\sim 3$  kpc (see, *e.g.*, Figure 2.3 and [164]). Extended regions of neutral atomic hydrogen (HI regions) are common to spirals as well and are often found at large radii. Their low densities contribute negligibly to rotation curves, but are useful tracers due to spectral emissions such as the hydrogen 21 cm line.

## 2.2. GALACTIC DARK MATTER

face brightness that falls off exponentially as  $r$  increases (see, *e.g.*, Figure 3.12 in [18]) and is reasonably well described by

$$I_{\text{ld}}(r) \propto e^{-r/r_D}, \quad (2.2.2)$$

where  $r_D$  is the disc scale.  $r_D$  is often expressed in terms of a galaxy's optical radius  $r_{\text{opt}} \equiv 3.2r_D$ , the radius that contains 83% of its light. Although the optical radii of spiral galaxies vary from a few kiloparsecs to several tens of kiloparsecs, a characteristic spiral has  $r_{\text{opt}} \simeq 10 \text{ kpc}$  and a total luminosity  $L_* \simeq 2.5 \times 10^{10} L_{\odot}$ . Assuming the matter density of a spiral's disc follows its surface brightness and falls off exponentially with the same characteristic scale, an expression for the LD mass as a function of radius can be obtained by integrating over the disc;

$$\begin{aligned} \mathcal{M}_{\text{ld}}(r) &= \int_{r' \leq r} \Sigma_{\text{ld}}(r') dA = \int_0^{2\pi} d\phi \int_0^r r' \Sigma_0 e^{-r'/r_D} dr' \\ &= 2\pi \Sigma_0 r_D^2 \left( 1 - e^{-r/r_D} - \frac{r}{r_D} e^{-r/r_D} \right), \end{aligned} \quad (2.2.3)$$

where  $\Sigma_0$  is the central areal matter density. The contribution to a spiral's galactic rotation curve from its LD is now easily obtained by substituting this expression for  $\mathcal{M}_{\text{ld}}(r)$  into Equation 2.2.1;

$$v_{\text{ld}}(r) = r_D \left[ \frac{2\pi \Sigma_0 G_N}{r} \left( 1 - e^{-r/r_D} - \frac{r}{r_D} e^{-r/r_D} \right) \right]^{-1/2}, \quad (2.2.4)$$

which falls off as  $1/\sqrt{r}$  for radii beyond the edge of the disc.

A flat rotation curve beyond  $r_{\text{opt}}$  implies a mass component that increases linearly with radius beyond the extent of the disc and is therefore inconsistent with the LD distribution;  $\mathcal{M}(r) \propto r$  at large  $r$ . The most popular solution is to assume that spiral galaxies are embedded in large halos of nonluminous matter, composed of either dark baryons (a possibility considered further in Section 2.4.1) or a collection of unknown exotic particles. The simplest DH density profile,  $\rho_{\text{dh}}(r)$ , that satisfies this condition (in the large- $r$  limit) is spherically

## CHAPTER 2. DARK MATTER

symmetric and falls off as  $1/r^2$ , such that

$$\begin{aligned}
 \mathcal{M}_{\text{dh}}(r) &= \int_{r' \leq r} \rho_{\text{dh}}(r') dV \\
 &= \int_0^{2\pi} d\phi \int_{-1}^1 d(\cos(\theta)) \int_0^r \frac{\rho_0}{(r')^2} (r')^2 dr' \\
 &= 4\pi\rho_0 r,
 \end{aligned} \tag{2.2.5}$$

where the central matter density  $\rho_0$  is proportional to the value of  $v^2$  where the rotation curve is flat. Obviously  $\rho_{\text{dh}} = \rho_0/r^2$  (known as an isothermal sphere DH profile) is an incomplete description since it becomes infinite at zero radius where the density is expected to be dominated by luminous matter. A common correction adds a halo-core radius  $a_r$  to the denominator such that

$$\rho_{\text{dh}}(r) \propto \frac{1}{r^2 + a_r^2}, \tag{2.2.6}$$

retaining the desired large- $r$  behavior while remaining finite at zero. The DH contribution to the rotation curve is then

$$v_{\text{dh}}^2(r) \propto \left(1 - \frac{a_r}{r} \tan^{-1} \frac{r}{a_r}\right). \tag{2.2.7}$$

The DH represented by Equations 2.2.6 and 2.2.7 is often referred to as a softened isothermal sphere. Most spirals are well approximated by a combination of LD and DH densities, with rotation-curve contributions given roughly by Equations 2.2.4 and 2.2.7, respectively.

Persic and Salucci combine rotation curves from optical and radio spectroscopy of  $\sim 1100$  spiral galaxies in [163]. Each rotation curve is scaled by its value at the optical radius,  $v(r_{\text{opt}})$ , expressed in terms of the dimensionless radius  $x_r \equiv r/r_{\text{opt}}$ , and sorted into one of 11 luminosity bins (according to their  $I$ -band absolute magnitudes  $M_I$ ). The resulting average rotation curves for 6 of their luminosity bins are shown in Figure 2.2 and are generally (but not exactly) flat beyond the optical radius ( $x_r \gtrsim 1$ ), suggesting the presence of DHs. They share a common profile with radial-dependent slopes that vary monotonically as a function of luminosity. Rubin and her collaborators recognized this trend

## 2.2. GALACTIC DARK MATTER

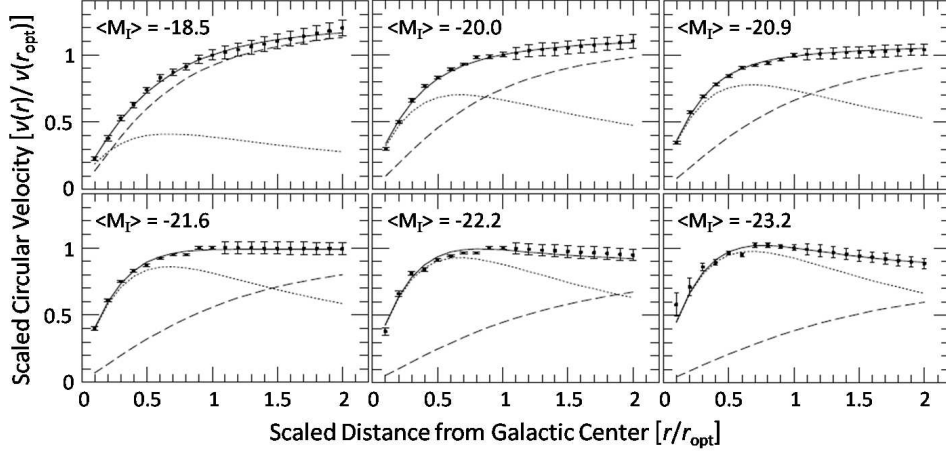


Figure 2.2: Average rotation curves (error bars) for 50–100 spiral galaxies of similar luminosity are shown in each panel, with the average  $I$ -band magnitude ( $\langle M_I \rangle$ ) indicated for each ensemble. Fits to the luminous-disc (dotted lines) and dark-halo (dashed lines) contributions add in quadrature to yield universal rotation curves (solid lines) that follow the data remarkably well. The data (spanning a range that covers a factor of  $\sim 75$  in luminosity) and fits are scaled to the rotational velocity at the optical radius,  $v(r_{\text{opt}})$ , and are plotted as a function of the dimensionless radius  $x_r \equiv r/r_{\text{opt}}$ . Discrepancies between the fits and data at low radius for the higher-luminosity bins (lower panels) are likely due to the influence of central, luminous bulges not accounted for by the fits. Regardless, each panel’s rotation curve clearly demonstrates the need for a dark-halo contribution beyond the optical radius. Figure adapted from [163].

in the mid-1980s and proposed a luminosity-dependent universal rotation curve for spiral galaxies [166, 167] that was later validated by Persic and Salucci with a sample of  $\sim 60$  spirals [168]. Based on the much larger sample in [163] and represented in Figure 2.2, the respective LD and DH URC components are

$$v_{\text{ld}}(r) = v(r_{\text{opt}}) \left[ \beta_d \frac{1.97 x_r^{1.22}}{(x_r^2 + 0.78^2)^{1.43}} \right]^{1/2}, \quad (2.2.8)$$

and

$$v_{\text{dh}}(r) = v(r_{\text{opt}}) \left[ (1 - \beta_d)(1 + \alpha_r^2) \frac{x_r^2}{(x_r^2 + \alpha_r^2)} \right]^{1/2}, \quad (2.2.9)$$



## CHAPTER 2. DARK MATTER

where  $\beta_d$  is the disc mass fraction at the optical radius and  $\alpha_r$  is the DH core radius in units of  $r_{\text{opt}}$ . This form for  $v_{\text{ld}}$  approximates an exponential disc and is only valid for  $0.04r_{\text{opt}} < r \leq 2r_{\text{opt}}$ . The full URC is given by the quadrature sum of  $v_{\text{ld}}$  and  $v_{\text{dh}}$ , and fits the average rotation-curve data in Figure 2.2 for

$$\begin{aligned}\beta_d &= 0.72 + 0.44 \log_{10}\left(\frac{L}{L_*}\right), \quad \text{and} \\ \alpha_r &= 1.5 \left(\frac{L}{L_*}\right)^{1/5}.\end{aligned}\tag{2.2.10}$$

This URC fit follows the data with remarkable accuracy considering the lowest- and highest-magnitude bins span a factor of  $\sim 75$  in luminosity. The LD contribution is an insufficient explanation of the data between  $r_{\text{opt}}$  and  $2r_{\text{opt}}$  for all ensemble averages, with the DH making an increasingly significant contribution as luminosity decreases.

The URC fit can be used to estimate the contributions to a spiral's total mass from its DH and LD components in terms of luminosity [154]. To this end, it is necessary to define an effective DH edge. It is common practice to use  $r_{200}$ , the radius at which the DH density is 200 times the critical density. Based on the URC represented by Equations 2.2.8, 2.2.9 and 2.2.10, Persic and Salucci find

$$r_{200} = 250 \left(\frac{L}{L_*}\right)^{0.2} \text{ kpc},\tag{2.2.11}$$

interior to which the total DH and LD masses are

$$\begin{aligned}\mathcal{M}_{\text{dh}} &\simeq 2 \times 10^{12} \left(\frac{L}{L_*}\right)^{1/2} \mathcal{M}_{\odot}, \quad \text{and} \\ \mathcal{M}_{\text{ld}} &\simeq 1 \times 10^{11} \left(\frac{L}{L_*}\right)^{1.3} \mathcal{M}_{\odot},\end{aligned}\tag{2.2.12}$$

respectively [154]. The total mass-to-light ratio for a spiral galaxy is then simply

$$\Upsilon_{\text{spiral}} = \Upsilon_{\text{dh}} + \Upsilon_{\text{ld}},\tag{2.2.13}$$

## 2.2. GALACTIC DARK MATTER

where

$$\begin{aligned}\Upsilon_{\text{dh}} &= \frac{\mathcal{M}_{\text{dh}}}{L} \simeq 80 \left( \frac{L_*}{L} \right)^{1/2} \Upsilon_{\odot}, \quad \text{and} \\ \Upsilon_{\text{ld}} &= \frac{\mathcal{M}_{\text{ld}}}{L} \simeq 4 \left( \frac{L}{L_*} \right)^{0.3} \Upsilon_{\odot}.\end{aligned}\tag{2.2.14}$$

For an  $L_*$ -luminosity spiral, the corresponding matter density relative to the critical density (using Equation 2.1.5) is  $\Omega_{\mathcal{M}} \simeq 0.074$ . This exceeds the allowed budget for baryons in the Standard Cosmology, an indication that spiral-galaxy DHs are composed of something other than dark baryons. It is also evident that only a fraction of the total dark matter preferred by the cosmological constraint on  $\Omega_c$  (Equation 1.3.28) is bound in spiral galaxies;  $\Omega_{\text{spiral}} \lesssim 0.1$ .

### The Milky Way

Although Persic and Salucci’s URC provides convincing evidence of DHs in spiral galaxies, it does not necessarily extend to our own galaxy. The possibility of directly detecting dark matter with terrestrial detectors relies on the existence of a DH in the Milky Way. In this section I review rotation-curve and mass estimates for the Milky Way and explore implications regarding the existence of a Galactic DH. The evidence presented here constitutes the best reason for believing dark matter populates our local neighborhood and might therefore be detectable here on Earth.

A compilation of rotational velocities for the Milky Way is shown in Figure 2.3. For radii interior to the Sun, the same kind of spectroscopy-based techniques described above are used to map the velocity profile. For radii beyond the Sun, however, doppler-shift-derived tangential velocities are less reliable and tend to be inaccurate (*e.g.*, notice the diamonds’ large error bars in Figure 2.3). Our position within the Galactic disc—assumed to be at a galactocentric radius of 8 kpc in Figure 2.3 and often referred to as the “solar circle”—is roughly equal to the Milky Way’s optical radius and is therefore near the outer edge of the LD. This is unfortunate as the DH’s contribution to

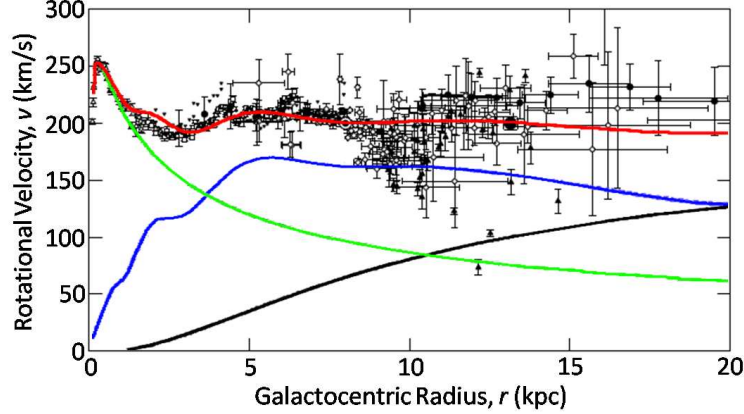


Figure 2.3: A compilation of rotational-velocity measurements as a function of galactocentric radius for the Milky Way Galaxy. From left to right (roughly), the different symbols are: HI 21 cm line observations ( $\triangle$ , [169]); CO 2.2 mm line ( $\square$ , [170]); HI 21 cm line ( $\blacktriangledown$ , [171]); H II and CO spectroscopy ( $\diamond$ , [171, 172]); carbon-star spectroscopy (filled triangles, [173]); HI disc thickness ( $\bullet$ , [174, 175]); and the big circle at  $\sim 13$  kpc is from precise parallax and velocity observations of variable stars in the S269 star forming region [176]. The Sun is indicated by the encircled bullet, with all data scaled to reflect a solar position and velocity of  $(r_0, v_\odot) = (8.0 \text{ kpc}, 200 \text{ km/s})$ . The data are available at <http://www.ioa.s.u-tokyo.ac.jp/~sofue/mw/rc2009/>. A three component rotation curve fit (red line) is (somewhat crudely) overlaid, with the individual contributions from the luminous bulge (green line), disc and spiral arms (blue line) and dark halo (black line) indicated as well. The fit includes only the HI data from [174] at large radii. Figure adapted from [164].

the rotation curve is expected to be most significant at radii beyond the disc where circular-velocity measurements are difficult.

One of the most reliable indications of a flat rotation curve beyond the solar circle results from observations of the S269 star forming region. A type of triangulation technique called parallax is used to estimate the distance to S269 from observations of its variable stars with the Very long-baseline interferometry Exploration of Radio Astrometry (VERA) telescope array [177, 176]. Stellar parallax refers to the change in a nearby object's angular position in

## 2.2. GALACTIC DARK MATTER

the sky relative to a set of distant light sources due to the motion of the Earth. The ratio of the observational baseline (the diameter of the Earth’s orbit) to the parallax gives accurate distances for nearby stars up to several kiloparsecs away. The VERA S269 parallax is one of the smallest ever measured, yielding a 5% estimate of the distance to S269. Spectroscopy of the 22 GHz maser emission from S269’s variables allows determination of the region’s rotational velocity. The final result is shown as a single (large) filled-in circle at  $\sim 13$  kpc in Figure 2.3 and shares a common rotational velocity of  $\sim 200 \text{ km s}^{-1}$  with the Sun despite being 5 kpc further away from the center of the Galaxy. The data strongly support a flat rotation curve from approximately 3–18 kpc.

To better understand the size of the Milky Way’s DH, it is instructive to fit for the individual contributions to its rotation curve. The rise in rotational velocities toward the center of the Galaxy is indicative of a nontrivial contribution from a luminous bulge (LB). To fit the data at all radii, it is therefore necessary to include an LB component in addition to the LD and DH. A popular LB model has a surface brightness that follows the de Vaucouleurs law [178], for which the surface mass density decreases quasi-exponentially with radius;

$$\Sigma_{\text{lb}}(r) \propto \exp \left[ -7.67 \left( \left( \frac{r}{r_e} \right)^{1/4} - 1 \right) \right], \quad (2.2.15)$$

where  $r_e$  is the half-light radius and is  $\sim 0.5$  kpc for the Milky Way [164]. To follow the wavy nature of the data between  $\sim 1$ –10 kpc, it is useful to model the density fluctuations caused by the Milky Way’s two spiral arms by adding a perturbation to the LD’s exponential surface-density profile such that

$$\Sigma_{\text{ld}}(r) \propto e^{-r/r_D} (1 + \Delta(r, \theta)), \quad (2.2.16)$$

where  $\theta$  is the polar angle in the plane of the disc, and  $\Delta$  is a sinusoidal function whose argument varies linearly with  $\theta$  and logarithmically with radius (see [164] for more details). The DH density is assumed to be of the form given in Equation 2.2.6. A simultaneous fit to the LB, LD and DH densities yields the individual and combined rotation curves shown in Figure 2.3. Aside from the dip at  $\sim 3$  kpc—thought to be caused by a bar (not accounted for

in the fit) that joins the two spiral arms at the center of the Galaxy—the fit generally follows the data and can be used for (rough) global estimates of the Milky Way’s mass parameters. The total mass within a 20 kpc sphere is  $2 \times 10^{11} \mathcal{M}_{\odot}$  ( $\pm 10\%$ ), with  $\sim 60\%$  due to the DH density (see Table 3 in [164] for further details). Although the DH contribution is subdominant and essentially unnecessary interior to the solar circle, to reproduce the flatness of the Milky Way’s rotation curve at larger radii requires a significant dark-matter density.

The shape of the DH fit shown in Figure 2.3 implies that this total-mass estimate for the Milky Way is a lower limit. Persic and Salucci’s URC suggests that an  $L_*$ -luminosity spiral has a DH that extends to  $r_{200} \simeq 250$  kpc. Even if the Milky Way’s luminosity is an order of magnitude lower, Equation 2.2.11 predicts  $r_{200} \simeq 160$  kpc, well in excess of the 20 kpc radius used above to derive the (clearly crude) 60% DH total-mass fraction. The principle shortcoming is a lack of high-radius observations due to a paucity of gas and stars in the LB and LD beyond 20 kpc. Fortunately, the extended halo of the Milky Way is not entirely dark. It hosts a number of globular clusters<sup>2</sup> and satellite galaxies whose velocities can be used (instead of blindly extending the URC fit beyond 20 kpc) for mass estimates of the Galaxy. Though highly model dependent and accurate to only  $\sim 50\%$ , estimates derived from astrometry of the Milky Way’s globular clusters and satellite galaxies yield a total mass of  $\sim 2 \times 10^{12} \mathcal{M}_{\odot}$  [179, 180].

An alternative method for judging the extent of the DH is to gauge the depth of the Milky Way’s gravitational potential using high-velocity stars to estimate its escape velocity. Unlike rotational velocities which have limited sensitivity to the mass exterior to the solar circle, the escape velocity is sensitive to the total mass of the Galaxy. Returning to the idea of a test particle in a gravitational potential, the escape velocity  $v_{\text{esc}}$  is the speed required for an

---

<sup>2</sup>Globular clusters are somewhat mysterious overdensities of stars usually found within galactic halos. They often consist of several hundred thousand low-metallicity Pop II stars tightly bound into a spherical ball tens of parsecs across and reaching a density of 100–1000 stars per cubic parsec near their cores. The Milky Way is believed to harbor  $\sim 200$  globular clusters, while the largest galaxies at the centers of galaxy clusters can possess in excess of 10,000.

## 2.2. GALACTIC DARK MATTER

object to have sufficient kinetic energy to (just) overcome a system's (full) gravitational potential energy:

$$\frac{1}{2}mv_{\text{esc}}^2 = \frac{G_N m \mathcal{M}_{\text{tot}}}{r_{\text{vir}}}, \quad (2.2.17)$$

where  $m$  is the mass of the test particle,  $\mathcal{M}_{\text{tot}}$  is the total mass of the system and  $r_{\text{vir}}$  its virial radius. The virial theorem is discussed in Section 2.3.1 in the context of galaxy clusters; for now it suffices to note that a common convention is to take  $r_{\text{vir}} = r_{200}$  and to consider the virial radius the (effective) dynamical extent of the DH. The escape velocity therefore depends on only the total mass and size of the Galaxy;

$$v_{\text{esc}} = \sqrt{\frac{2G_N \mathcal{M}_{\text{tot}}}{r_{200}}}. \quad (2.2.18)$$

$v_{\text{esc}}$  can be statistically inferred from the high-value tail of the stellar-velocity distribution and, when coupled to a specific DH model, can be used to estimate  $\mathcal{M}_{\text{tot}}$  for the Milky Way.

The principle challenge in estimating the escape velocity is to find a statistically significant sample of high-velocity stars. Most stellar catalogs consist primarily of stars in the LD that are locked into (essentially) circular orbits. They possess velocities that reflect the mass interior to their orbits and rarely exceed  $\sim 200 \text{ km s}^{-1}$ . For this reason, high-velocity stars within the halo of the Galaxy are better suited for estimating its escape velocity. Furthermore, halo stars are more likely to possess large radial velocities, which can be measured more accurately than tangential velocities.

Out of  $\gtrsim 52,000$  stars in the RAdial Velocity Experiment (RAVE) 2006 [181] and Beers *et al.* 2000 [182] catalogs, Smith and collaborators identify 33 halo stars with radial velocities in excess of  $300 \text{ km s}^{-1}$  [183]. Using a likelihood method that models the tail of the Milky Way's stellar-velocity distribution as a power law with an unknown exponent, they estimate (with a prior on the exponent justified by galactic simulations)

$$v_{\text{esc}} = 544 \text{ km s}^{-1}, \quad (2.2.19)$$

## CHAPTER 2. DARK MATTER

which is the median value of their best-fit 90% confidence interval:

$$498 < v_{\text{esc}} < 608 \text{ km s}^{-1}. \quad (2.2.20)$$

Unless otherwise stated, the escape velocity throughout this thesis is taken to be  $544 \text{ km s}^{-1}$ .

The  $498 \text{ km s}^{-1}$  lower limit is itself convincing proof of the existence of a DH. From Equations 2.2.1 and 2.2.18 it is clear that the escape and rotational velocities are related by a factor of  $\sqrt{2}$ . An expression for the escape velocity in terms of the Sun's rotational velocity ( $v_{\odot}$ ) is then

$$\begin{aligned} v_{\text{esc}}^2 &= 2v_{\odot}^2 + 2G_N \int_{r_0}^{r > r_0} \frac{\rho(r)}{r} dV \\ &= 2v_{\odot}^2 + 8\pi G_N \int_{r_0}^{r_{200}} \rho(r) r dr, \end{aligned} \quad (2.2.21)$$

where  $r_0$  is the radius of the solar circle. The currently favored value of  $v_{\odot}$  is  $220 \pm 20 \text{ km s}^{-1}$  [184]. The  $240 \text{ km s}^{-1}$  upper limit gives a value for  $\sqrt{2}v_{\odot} \simeq 340 \text{ km s}^{-1}$ . This is significantly less than  $498 \text{ km s}^{-1}$ , implying that the second term on the right-hand side of Equation 2.2.21 makes a nontrivial contribution to the matter density outside the solar circle. Beyond the solar circle, the LB and LD are too sparse to cause such a large discrepancy between  $v_{\text{esc}}$  and  $\sqrt{2}v_{\odot}$ . Invoking a DH with properties similar to those described previously resolves the apparent contradiction.

To estimate the total mass of the Milky Way from its escape velocity requires specification of a mass-distribution model. Assuming LB and LD models with total masses similar to those obtained from the rotation-curve fit in Figure 2.3, Smith *et al.* employ a variety of DH models to estimate the total mass and virial radius of the Milky Way. A popular model is the Navarro, Frenk and White (NFW) halo profile in which

$$\rho_{\text{dh}} \propto \frac{1}{r(a_r + r)^2}. \quad (2.2.22)$$

Notice that the NFW density is infinite at zero radius; this is what is known as a cuspy DH. Allowing for adiabatic contraction of the NFW dark halo due to

## 2.2. GALACTIC DARK MATTER

gravitational interactions with baryons in the LB and LD [185], Smith *et al.* find

$$\begin{aligned}\mathcal{M}_{\text{tot}} &= 1.42_{-0.54}^{+1.14} \times 10^{12} \mathcal{M}_{\odot}, \text{ and} \\ r_{200} &\simeq 305_{-45}^{+66} \text{ kpc},\end{aligned}\tag{2.2.23}$$

consistent with estimates derived from globular-cluster and satellite-galaxy velocities;  $\mathcal{M}_{\text{tot}} \simeq (2 \pm 1) \times 10^{12} \mathcal{M}_{\odot}$  [179, 180]. Extending Persic and Salucci's URC results to the Milky Way, an estimate for its luminosity can be derived from Equation 2.2.11:

$$\begin{aligned}L_{\text{tot}} &= \left( \frac{r_{200}}{250 \text{ kpc}} \right)^5 L_{*} \\ &\simeq 2.7 L_{*} = 6.8 \times 10^{10} L_{\odot}.\end{aligned}\tag{2.2.24}$$

Combining Equations 2.2.23 and 2.2.24 gives a total mass-to-light ratio for the Milky Way of  $\sim 10\text{--}40 \Upsilon_{\odot}$ .



### 2.2.2 The Tully-Fisher Relation

A superior method for estimating galactic luminosities from rotational velocities is to use the Tully-Fisher Relation (TFR) [186]. Over the years the term TFR has been loosely used to refer to the tight correlation between the value of a disc galaxy’s rotational velocity (where its rotation curve is flat, call it  $v_{\text{flat}}$ ) and global parameters such as its total baryonic mass, stellar mass or luminosity. Tully and Fisher originally observed a relationship between luminosity (in terms of absolute magnitude) and  $v_{\text{flat}}$ . Regardless of its particular form, a TFR contains information similar to Persic and Salucci’s URC. However, instead of using a specific galactic-mass model to relate global parameters to luminosity, a global parameter such as luminosity is simply plotted as a function of  $v_{\text{flat}}$  for an ensemble of galaxies. When fit as a function of  $v_{\text{flat}}$ , the resulting relation is referred to as a forward TFR. When fit as a function of the global parameter (*e.g.*,  $v_{\text{flat}}$  vs. luminosity), the relation is commonly referred to as an inverse TFR. The forward TFR between luminosity and  $v_{\text{flat}}$  is less model dependent than the URC fit and should provide a more reliable estimate for the Milky Way’s luminosity than the estimate in Equation 2.2.24.

To gain a better understanding of how  $v_{\text{flat}}$  (and therefore the TFR) relates to a galaxy’s DH, it is useful to consider the simplest possible DH model, a spherically symmetric isothermal sphere of ideal gas (see [18] for a detailed discussion). For the purposes of this discussion, whether the gas particles are dark baryons or nonbaryonic dark matter is irrelevant. An isothermal velocity distribution is isotropic. It has a velocity dispersion  $\sigma$  that is independent of radius and relates the gas pressure to its density;<sup>3</sup>

$$p = \rho(r)\sigma^2. \quad (2.2.25)$$

Hydrostatic equilibrium requires the pressure force to balance the gravitational

---

<sup>3</sup>Note that the symbol  $\sigma$  is used to represent several unrelated quantities throughout this thesis. Rather than awkwardly invent a new symbol each time a  $\sigma$  is called for by the literature for a particular subject, I chose to conform to popular conventions. Particular care should be taken to note the context in which  $\sigma$  is used in each instance. In many cases a subscript helps to avoid confusion.

## 2.2. GALACTIC DARK MATTER

force (per unit volume), such that

$$-\frac{dp}{dr} = -\sigma^2 \frac{d\rho}{dr} = \frac{G_N \mathcal{M}(r)}{r^2} \rho(r), \quad (2.2.26)$$

where  $\mathcal{M}(r)$  is obtained by integrating  $\rho(r)$  over a sphere with radius  $r$ . The solution to Equation 2.2.26 is nontrivial and generally requires a numerical rather than analytic approach. However, at large radii the density approaches  $\rho(r) \propto 1/r^2$ . Substituting this form into Equation 2.2.26, an expression for the total mass as a function of radius can be found without knowing the proportionality constant;

$$\mathcal{M}(r) = \frac{2\sigma^2 r}{G_N}. \quad (2.2.27)$$

Comparing this to Equation 2.2.1, the value of a galactic rotation curve where it is flat can be related to an isothermal DH's velocity dispersion;

$$v_{\text{flat}} = \sqrt{2}\sigma. \quad (2.2.28)$$

The TFR between luminosity and  $v_{\text{flat}}$  directly relates a galaxy's luminosity to a fundamental property of its DH. That the DH's velocity dispersion is larger for more luminous galaxies reflects the gravitational influence of a larger DH. Disc galaxies with larger DH's have deeper gravitational wells that are (apparently) more effective at causing visible matter to collapse into luminosity-producing star forming regions.

The  $I$ -band luminosity TFR for a sample of  $\lesssim 100$  disc-dominated galaxies selected from the SDSS 2002 main galaxy catalog [132] is shown in Figure 2.4. The best-fit forward TFR relation is [187]

$$L_I = (1.8 \pm 0.1) \left( \frac{v_{2.2}}{149.6 \text{ km s}^{-1}} \right)^{2.60 \pm 0.13} \times 10^{10} L_{\odot}, \quad (2.2.29)$$

where  $v_{2.2}$  is the rotational velocity at 2.2 times the disc-scale length and can be considered equivalent to  $v_{\text{flat}}$ . Substituting  $v_{2.2} = v_{\odot} = 220 \pm 20 \text{ km s}^{-1}$  into Equation 2.2.29 yields a prediction for the Milky Way's total luminosity;

$$L_{\text{tot}} \simeq (5.0 \pm 0.5) \times 10^{10} L_{\odot}, \quad (2.2.30)$$

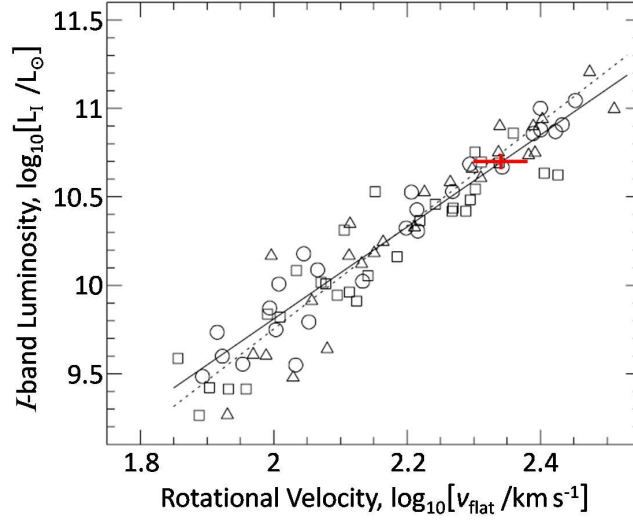


Figure 2.4: The Tully-Fisher Relation (TFR) between the logarithm of the  $I$ -band luminosity (in solar units) and the logarithm of the rotational velocity at 2.2 times the disc-scale length for  $\lesssim 100$  disc-dominated galaxies from the SDSS 2002 main galaxy catalog [132], overlaid with the best-fit forward TFR given by Equation 2.2.29. Circles (squares) represent galaxies that are more blue (red) relative to the average color of the full sample, while triangles are of intermediate color. The luminosity range corresponding to the Milky Way’s  $220 \pm 20 \text{ km s}^{-1}$  rotational velocity is indicated by a red cross. The best-fit inverse TFR is shown as a dotted line. Figure adapted from [187].

which is indicated by the red cross in Figure 2.4 and is similar to Equation 2.2.24. Dividing  $\mathcal{M}_{\text{tot}}$  from Equation 2.2.23 by Equation 2.2.30, yields a mass-to-light ratio for the Milky Way of  $\sim 15\text{--}55 \Upsilon_{\odot}$ . Extending to the Universe as a whole, this corresponds to a matter density of 1.5–5% of the critical density. This is roughly consistent with the URC results quoted in Equation 2.2.14. If the relative contributions to total mass implied by Equation 2.2.12 hold true for the Milky Way,  $>90\%$  of the Milky Way’s total mass is from dark particles (whether baryonic or exotic) in its halo.

## 2.2. GALACTIC DARK MATTER

### 2.2.3 Elliptical Galaxies

Though less relevant than spirals regarding the local dark-matter density, elliptical galaxies have traditionally played an important role in our understanding of galactic dark matter. Ellipticals are ellipsoidal in shape, ranging from perfect spheres to highly flattened ellipsoids, and exhibit a wider range of sizes and total masses than spirals. Despite their Hubble-sequence classification as early-type galaxies, ellipticals are thought to be near the endpoint of the galactic-evolution scale, consisting primarily of older, yellow-red Pop II stars. They exhibit a nearly featureless surface brightness, not dissimilar to the LBs of spirals, with luminosity profiles that are (generally) well described by Sérsic's  $1/n$  law [188] in which the surface brightness is given by

$$I(r) \propto \exp \left[ -b_n \left( \left( \frac{r}{r_e} \right)^{1/n} - 1 \right) \right]. \quad (2.2.31)$$

$n$  is a variable index and the constant  $b_n$  is chosen such that  $r_e$  is the radius containing one-half of the projected light:  $b_n \simeq 2n - 0.324$  [189]. Substituting  $n = 4$  returns the de Vaucouleurs law used in the previous section to model the LB of the Milky Way. The most luminous ellipticals tend to have larger Sérsic indices, while dwarf ellipticals have exponential luminosity profiles similar to the discs of spirals with  $n \simeq 1$  [190].

While techniques similar to those used for spirals are available for luminosity estimates, measuring an elliptical's total dynamical mass is less straightforward. Ellipticals are pressure-supported, triaxial stellar systems in which most of the support against gravitational collapse is due to essentially random motions rather than ordered rotation, as in the discs of spirals. Inferring an elliptical's mass distribution from stellar kinematics is complicated by geometric and rotational effects as well as by possible velocity-dispersion anisotropies [191]. Emissions from gaseous regions have been used to trace mass distributions for elliptical galaxies. Spectroscopic observations of H $\alpha$  lines from warm gas (in their cores) and the 21 cm line from neutral hydrogen (at larger radii) indicate mass-to-light ratios that increase radially, due (presumably) to the presence of dark halos [192].

## CHAPTER 2. DARK MATTER

Most ellipticals harbor halos of x-ray-emitting ionized gas that extend to radii well beyond their central stellar regions and can be used to trace their mass distributions. The free electrons and ions undergo thermal bremsstrahlung in which inter-electron and electron-ion collisions release photons with energies that scale according to their gravitationally-induced velocities. The host galaxy's gravitational potential exerts enough force on the gas to induce velocities characterized by x-ray temperatures ( $kT \simeq 1$  keV). The precise interplay between an elliptical's total mass and x-ray temperature can be understood by considering its gas content to be in hydrostatic equilibrium (as in Equation 2.2.26). The gas pressure is related to its density through the ideal gas law;

$$p_g(r) = \frac{kT(r)\rho_g(r)}{\mu_g m_p}, \quad (2.2.32)$$

where  $T(r)$  is the x-ray temperature of the gas,  $\mu_g \simeq 0.6$  is the mean molecular weight [193] and  $m_p$  is the mass of the proton (such that the product  $\mu_g m_p$  gives the mean mass per ionized-gas particle). Substituting this into Equation 2.2.26 gives

$$-\frac{d}{dr} \left( \frac{kT(r)\rho_g}{\mu_g m_p} \right) = \frac{G_N \mathcal{M}(r)}{r^2} \rho_g, \quad (2.2.33)$$

which can be rearranged to express the total mass in terms of the radial-density and -temperature gradients of the gas [191];

$$\mathcal{M}(r) = -\frac{kT(r)r}{G_N \mu_g m_p} \left( \frac{d \ln \rho_g}{d \ln r} + \frac{d \ln T(r)}{d \ln r} \right). \quad (2.2.34)$$

The distribution of gas in ellipticals is often close to isothermal such that the second term in Equation 2.2.34 is negligible compared to the density gradient, and the system is well described by a single (global) temperature  $T$ . When modeling specific elliptical galaxies, it is common practice to parameterize the gas according to a  $\beta$ -model [194] with density

$$\rho_g \propto \frac{1}{(1 + (r/r_c)^2)^{3\beta/2}}, \quad (2.2.35)$$

where  $\beta$  and  $r_c$  are free parameters. Fitting the  $\beta$ -model to an elliptical's x-ray surface-brightness profile and temperature allows determination of its mass

## 2.2. GALACTIC DARK MATTER

profile via Equation 2.2.34. Since an elliptical's halo gas contributes negligibly to its gravitational potential,  $\mathcal{M}(r)$  can be modeled independently of  $\rho_g$  as a sum of stellar and DH components. For the isolated elliptical NGC 7785, the authors in [191] model the stellar component according to Sérsic's law in Equation 2.2.31, with  $n = 3.8$  and  $r_e = 5.6$  kpc. They employ the Burkert DH model [195] in which the density is given by

$$\rho_{\text{dh}} = \frac{\rho_0 a_r^3}{(r + a_r)(r^2 + a_r^2)}, \quad (2.2.36)$$

and for which the mass profile has the form

$$\mathcal{M}_{\text{dh}}(r) = 6.4\rho_0 a_r^3 \left[ \ln(1 + r/a_r) - \tan^{-1}(r/a_r) + \frac{1}{2} \ln(1 + (r/a_r)^2) \right]. \quad (2.2.37)$$

Their best-fit model is compared to NGC 7785's x-ray-determined mass profile in Figure 2.5. The evidence for a significant dark halo is dramatic.

A less model dependent argument for the existence of dark halos in ellipticals can be derived from the relationship between their x-ray temperatures and stellar velocity dispersions ( $\sigma_*$ ). If an elliptical's mass is due entirely to stellar material, its stars and gas should be characterized by similar velocity dispersions. Relating the temperature of an ideal gas to its velocity dispersion through Equations 2.2.25 and 2.2.32, the square of the stellar velocity dispersion should therefore be proportional to the temperature of the gas;

$$\sigma_*^2 = \frac{kT}{\mu_g m_p}. \quad (2.2.38)$$

However, if a dark halo contributes nontrivially to an elliptical's gravitational potential, the x-ray temperature of ionized gas in its halo will be elevated relative to the temperature expected from the influence of the stellar component alone. Since an elliptical's stellar component is so centrally concentrated, where the dark-halo density tends to be less influential, the stellar velocity dispersion will not necessarily reflect the full gravitational potential of the extended halo. Consequently, the relationship between  $T$  and  $\sigma_*$  is expected to be softer in the

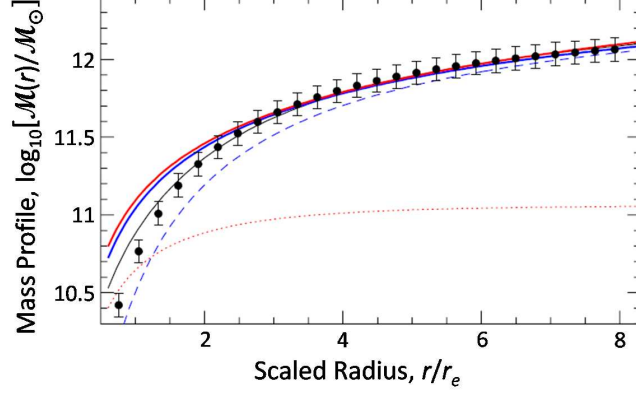


Figure 2.5: Mass profile for the isolated elliptical galaxy NGC 7785 as measured from its x-ray temperature and surface brightness (error bars) as a function of galactocentric radius scaled to the half-light radius of the best-fit  $1/n$  Sérsic stellar component (red dotted line), with  $r_e = 5.6$  kpc and  $n = 3.8$ . The best-fit (combined) stellar and dark-halo mass profile (black solid line) includes a significant contribution from a Burkert halo (blue dashed line). The total mass profiles for an NFW halo (blue solid line) and for the Modified Newtonian Dynamics (MOND) [196, 197] framework (red solid line) are shown as well. Figure taken from [191].

dark-halo scenario, with  $T \propto \sigma_*^\kappa$  and  $\kappa < 2$ . This is precisely what Davis and White observe for a sample of 27 elliptical galaxies with x-ray temperatures between  $\sim 0.5$  and 2 keV [198]. The data are shown in Figure 2.6 compared to Equation 2.2.38, and clearly exhibit temperatures in excess of those expected from their stellar velocity dispersions. The best-fit  $T$  versus  $\sigma_*$  relationship prefers  $\kappa = 1.45 \pm 0.20$ , indicating the presence of dark halos in these elliptical galaxies. This relationship between gas temperature and stellar velocity dispersion is often expressed in terms of the  $\beta$ -model's exponent [199];

$$\beta = \frac{\mu_g m_p \sigma_*^2}{kT}, \quad (2.2.39)$$

the ratio between the stellar kinetic energy and the thermal energy of the gas. For a stellar-dominated elliptical  $\beta \simeq 1$ , whereas for an elliptical with an extended dark halo  $\beta < 1$ .  $\beta \simeq 0.5$  for Davis and White's sample.

## 2.2. GALACTIC DARK MATTER

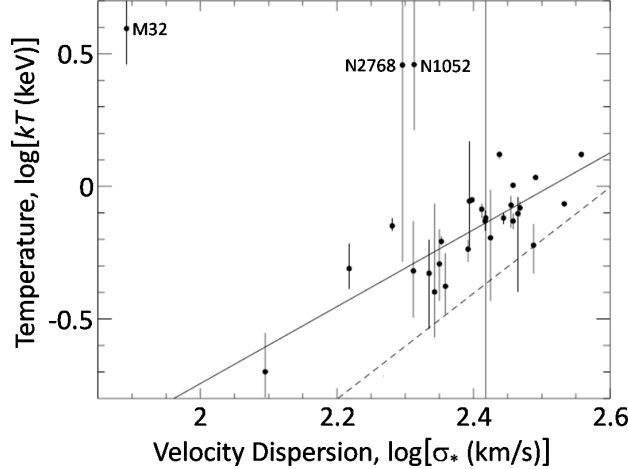


Figure 2.6: The x-ray temperatures for  $\sim 30$  elliptical galaxies versus their stellar velocity dispersions. The data prefer  $T \propto \sigma_*^{1.45 \pm 0.20}$  (solid line) over the  $T \propto \sigma_*^2$  relationship (dashed line) in which the ellipticals' gravitational potentials are due entirely to their stellar content, indicating the presence of dark halos. The three  $kT > 2$  keV galaxies (with labels) were excluded from the fit due to suspected temperature contributions from x-ray binaries or active galactic nuclei (rather than by diffuse, ionized halo gas). Figure taken from [198].

Unlike spirals, generalization of the properties of elliptical-galaxy dark halos has traditionally been difficult due to their morphological variability. It has become common practice to consider ellipticals in statistically-limited subclasses, or on a case-by-case basis. The advent of modern x-ray observatories (*e.g.*, the Chandra [201] and XMM-Newton [202] satellites) has (somewhat) improved the situation by allowing for detailed studies of elliptical-galaxy temperature profiles for larger samples. Analyses based on Chandra and XMM-Newton x-ray data tend to find  $0.5 < \beta < 1$  (see, *e.g.*, Figure 9 in [203] or Figure 11 in [200]). The improved resolution of these modern x-ray observations also allows deduction of the radial dependence of elliptical-galaxy mass-to-light ratios. As illustrated for several ellipticals in Figure 2.7, they (almost universally) tend to increase with radius, suggesting the presence of dark halos [204].



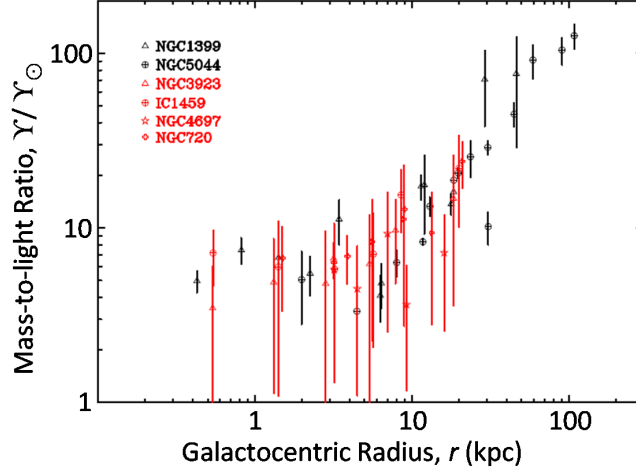


Figure 2.7: Mass-to-light ratios (in solar units) for several elliptical galaxies as a function of galactocentric radius. Data are derived from Chandra and XMM-Newton x-ray temperature profiles and optical photometry. The black error bars correspond to extended x-ray galaxies (EXGs) in which the x-ray temperatures and intensities increase with radius, while the red error bars correspond to compact x-ray galaxies (CXGs) in which the temperature profiles are constant or decline with radius. The presence of dark halos in the two EXG ellipticals is unmistakable. Figure taken from [200].

## 2.3 Dark Matter in Galaxy Clusters

Although evidence for the existence of galactic dark matter is convincing, the matter density inferred from spiral- and elliptical-galaxy mass-to-light ratios using the Oort method ( $\Omega_{\mathcal{M}} \lesssim 0.1$ ) is roughly a factor of 3 lower than is preferred by the Standard Cosmology ( $\Omega_{\mathcal{M}} \simeq 0.3$ ). A significant fraction of the Standard Cosmology's nonbaryonic dark-matter density is unaccounted for in direct observations of the gravitational effects of dark matter on galactic scales. Galactic-matter distributions are evidently not representative of the Universe at large. To make the connection between the cosmologically preferred and locally observed matter densities requires sampling spatial volumes large enough for the Standard Cosmology's homogeneity assumption to be valid.

### 2.3. DARK MATTER IN GALAXY CLUSTERS

Galaxy “groups” are the next largest gravitationally bound systems in the Universe. Over half of the nearest galaxies appear to be members of groups [205]. A typical group consists of several galaxies (2–50) occupying a spatial volume a few megaparsecs across, with a total mass of  $\sim 10^{12}–10^{13} \mathcal{M}_{\odot}$ . The Milky Way and Andromeda galaxies are the largest members of the Local Group, which hosts  $\gtrsim 30$  galaxies and spans  $\gtrsim 2$  Mpc [206]. In addition to their galactic content, groups contain baryons in the form of intergalactic gas and appear to possess significant dark halos. Mass-to-light ratios for a sample of  $\sim 200$  galaxy groups are plotted as a function of their galactic velocity dispersions in Figure 2.8. The average mass-to-light ratio for this sample is consistent with  $\Omega_{\mathcal{M}} \gtrsim 0.1$  [207]. As is clear from the spread of values in Figure 2.8, applying the Oort method to groups is problematic. This is partly due to their low member statistics and the correspondingly high probability of skewed results due to contamination by nonmember galaxies (sometimes called “interlopers”) [208]. Furthermore, the virial method (see below) used to estimate the group masses in Figure 2.8 relies on the assumption that they are gravitationally relaxed, or virialized. Groups, however, often possess highly irregular morphologies that suggest they are far from virialization [209]. It is also likely that groups are generally too small to be considered representative of the Universe as a whole.

Galaxy clusters are the largest gravitationally bound objects in the Universe. Like groups, they consist of galaxies, intracluster gas and dark matter, all orbiting a common center of mass in a roughly spherical configuration. They can contain as few as 50 or as many as several thousand galaxies. Clusters (or groups) are said to be “rich” or “poor” if they have more or fewer member galaxies than average, respectively. Dense clusters tend to have a higher fraction of elliptical and lenticular<sup>4</sup> galaxies, while sparse clusters favor spirals [210]. Clus-

---

<sup>4</sup>According to the Hubble-sequence classification scheme, lenticular galaxies are an intermediate galaxy type that falls between the elliptical and spiral types. Lenticulars have prominent bulges *and* discs, but lack well-defined spiral arms. Regarding stellar content, lenticulars have more in common with ellipticals than spirals; they tend to contain older stars, few or no active star forming regions, and are often difficult to distinguish (observationally) from highly flattened ellipticals. Due to their similarity, lenticulars and ellipticals are commonly referred to (somewhat ironically) as early-type galaxies, while spirals are sometimes called late-type galaxies.

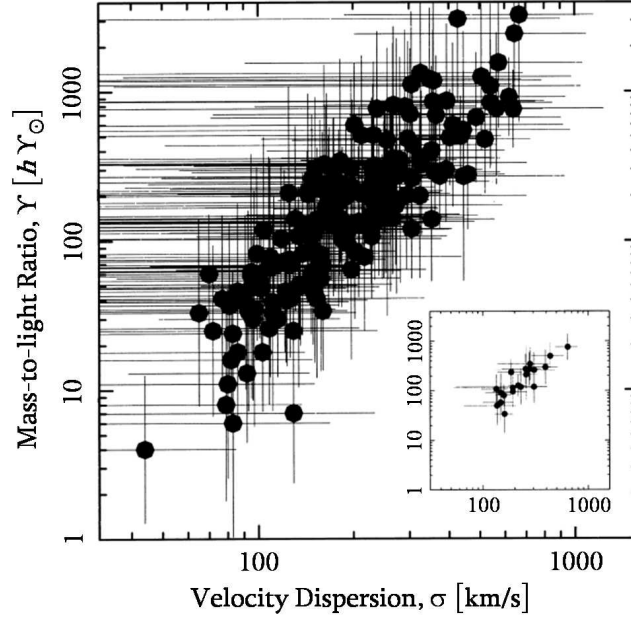


Figure 2.8: Mass-to-light ratios (in solar units scaled to the dimensionless Hubble constant) for a sample of  $\sim 200$  galaxy groups plotted versus their galactic velocity dispersions. The large spread of values is indicative of the systematic difficulties inherent to galaxy-group virial-mass estimates. Rich groups with at least six member galaxies are displayed in the inset. Figure taken from [207].

ters have total masses that range from  $\sim 10^{13} \mathcal{M}_{\odot}$  to several times  $10^{15} \mathcal{M}_{\odot}$ , and diameters that vary from a few to tens of megaparsecs. For further details regarding the general properties of galaxy clusters, I refer the interested reader to Biviano’s excellent review (given from a historical perspective with over 500 references) of cluster-based science in [211].

As was the case for galactic dark matter, the study of cluster dark matter relies on accurately measuring total dynamical mass. Since mass is not a direct observable for clusters, a mass proxy is used to infer the extent of their gravitational potentials. There are four widely used methods for estimating cluster mass via a proxy: 1) the virial theorem makes use of galactic velocity dispersion; 2) x-ray emission utilizes the x-ray temperature or luminosity of in-

### 2.3. DARK MATTER IN GALAXY CLUSTERS

tracluster ionized gas; 3) gravitational lensing uses luminosity distortions; and 4) the Sunyaev-Zel'dovich (SZ) effect [212] relies on observations of CMB photons. A truly staggering body of literature exists for each technique. Rather than attempt a comprehensive review, for each method I will outline the technique and discuss a few representative results. Where appropriate, implications for the Standard Cosmology are discussed. Similar to the BAO results in Section 1.3.4, the global properties of clusters are primarily sensitive to  $\Omega_{\mathcal{M}}$  due to the limited number of cluster observations at high redshifts. However, when combined with the CMB and BBN preferred value of  $\Omega_b$ , cluster observations confirm the existence of a significant nonbaryonic-matter density pervading the large-scale structure in the Universe.

#### 2.3.1 The Virial Theorem

The virial theorem relates a cluster's kinetic and potential energies, allowing deduction of the cluster's dynamical mass from its (relatively) easily observed galactic velocities. The following discussion (loosely) follows the treatment in [18]. Consider a cluster of galaxies with masses  $m_i$  and individual, nonrelativistic velocities  $\vec{v}_i$ . The cluster density is position dependent,  $\rho \equiv \rho(\vec{x})$ , and has an average density  $\rho_b$ . The Hamiltonian for the cluster can be written as a sum of its kinetic and potential energies;

$$\mathcal{H} = \frac{1}{2} \sum_{\forall i} m_i a^2 v_i^2 - \frac{1}{2} G_N a^5 \int dV_1 dV_2 \frac{(\rho(\vec{x}_1) - \rho_b)(\rho(\vec{x}_2) - \rho_b)}{|\vec{x}_1 - \vec{x}_2|}, \quad (2.3.1)$$

where the factor of 1/2 in front of the potential-energy term corrects for double counting in the integral, and  $a$  is the scale factor introduced in Section 1.2.1. Equation 2.3.1 can be rewritten in terms of the kinetic and potential energies per unit mass;

$$\mathcal{H} = \mathcal{M}_{\text{tot}} (\mathcal{K} + \mathcal{W}), \quad (2.3.2)$$

where

$$\begin{aligned}
 \mathcal{M}_{\text{tot}} &= \sum_{\forall i} m_i = a^3 \int dV \rho_b, \\
 \mathcal{K} &= \frac{1}{2} \frac{\sum m_i a^2 v_i^2}{\sum m_i} = \frac{1}{2} \sigma^2, \quad \text{and} \\
 \mathcal{W} &= -\frac{1}{2} G_N a^2 \rho_b \int dV \frac{\xi(x)}{x}.
 \end{aligned} \tag{2.3.3}$$

$\xi$  is the dimensionless two-point correlation function as in Section 1.3.4, and  $\sigma$  is the velocity dispersion of the cluster's mass components. Since the velocity dispersion is a mass-weighted average, it can be derived from a representative subcomponent of the cluster. Provided the cluster's galaxies fairly sample the total-mass distribution,  $\sigma$  can be estimated from galactic velocities alone. The energy equation equates the total and partial time derivatives of the Hamiltonian;

$$\frac{d\mathcal{H}}{dt} = \frac{\partial \mathcal{H}}{\partial t}, \tag{2.3.4}$$

where the partial derivative is performed at fixed positions,  $\vec{x}_i$ , and momenta,  $\vec{p}_i = m_i a^2 \vec{v}_i$ . The right-hand side of Equation 2.3.4 can be evaluated by noting that

$$\begin{aligned}
 \mathcal{W} &\propto a^2 \rho_b = a^2 \frac{\mathcal{M}_{\text{tot}}}{a^3 \int dV} \propto a^{-1}, \quad \text{and} \\
 \mathcal{K} &= \frac{1}{2\mathcal{M}_{\text{tot}}} \sum_{\forall i} \frac{|\vec{p}_i|^2}{2m_i a^2} \propto a^{-2},
 \end{aligned} \tag{2.3.5}$$

such that  $\partial \mathcal{W} / \partial t = -\dot{a} \mathcal{W} / a$  and  $\partial \mathcal{K} / \partial t = -2\dot{a} \mathcal{K} / a$ . Equation 2.3.4 can therefore be written as

$$\frac{d}{dt}(\mathcal{K} + \mathcal{W}) + \frac{\dot{a}}{a}(2\mathcal{K} + \mathcal{W}) = 0, \tag{2.3.6}$$

a form commonly referred to as the Layzer-Irvine equation [213]. If the cluster has relaxed to dynamical equilibrium it is said to be virialized, and the first term in Equation 2.3.6 is effectively zero. Under the assumption of virialization, the Layzer-Irvine equation reduces to

$$2\mathcal{K} = \sigma^2 = \mathcal{W} \propto \frac{\mathcal{M}_{\text{vir}}}{r_{\text{vir}}}, \tag{2.3.7}$$

### 2.3. DARK MATTER IN GALAXY CLUSTERS

where  $r_{\text{vir}}$  is the cluster-centric radius inside of which virialization holds true, and  $\mathcal{M}_{\text{vir}}$  is the corresponding enclosed (total) dynamical mass of the cluster. Equation 2.3.7 is known as the virial theorem.

Among the earliest virial measurements were Zwicky and Smith’s mass estimates for the Coma and Virgo clusters, respectively [135, 136, 214]. Scaled to the currently accepted value of the Hubble constant, they found  $\mathcal{M}_{\text{Coma}} \gtrsim 3.3 \times 10^{14} \mathcal{M}_{\odot}$  and  $\mathcal{M}_{\text{Virgo}} \simeq 7.7 \times 10^{14} \mathcal{M}_{\odot}$  [208]. Zwicky deduced a mass-to-light ratio for the Coma cluster more than two orders of magnitude in excess of the known value (at the time) of the stellar mass-to-light ratio. He briefly commented that the discrepancy merited further investigation [136]. Smith compared his estimate of the Virgo cluster’s mass to Hubble’s estimate for the average mass of a galaxy, or what they referred to at the time as a “nebula.” He found a similarly sized discrepancy and concluded [214],

The difference between this result and Hubble’s value for the average mass of a nebula apparently must remain unexplained until further information becomes available. . . It is also possible that both values are essentially correct, the difference representing internebular material, either uniformly distributed or in the form of great clouds of low luminosity surrounding the nebulae. . .

The nature of Smith’s internebular material is a mystery to this day. These early applications of the virial theorem are now recognized as the first hints that dark matter pervades the large-scale structure of the Universe.

In practice, application of the virial theorem is complicated by the fact that a cluster’s full 3-dimensional structure is inaccessible to the observer. The principle observables are the projected number-density profile,  $N(R)$ , and the line-of-sight velocity dispersion of the cluster’s member galaxies,  $\sigma_p(R)$ . A popular technique for disentangling the virial mass from  $N(R)$  and  $\sigma_p(R)$  is the Jeans method (see [215] for a detailed derivation). Except where otherwise noted, the following discussion follows Biviano’s summary of the Jean’s method in [216], and spherical symmetry is assumed. The first step of the (direct) Jeans method involves deprojecting the observables. The number-density profile can

be deprojected via a straightforward Abel inversion [217];

$$n(r) = -\frac{1}{\pi} \int_r^\infty \frac{dN(R)}{dR} \frac{dR}{\sqrt{R^2 - r^2}}, \quad (2.3.8)$$

where  $r$  and  $R$  are the 3-dimensional and projected 2-dimensional radii, respectively. Deprojecting the velocity dispersion is complicated by the anisotropy profile

$$\beta_a(r) \equiv 1 - \frac{\sigma_t^2(r)}{\sigma_r^2(r)}, \quad (2.3.9)$$

where  $\sigma_t$  and  $\sigma_r$  are the tangential and radial velocity dispersions, respectively. For an isotropic velocity distribution  $\beta_a = 0$ , and the radial velocity dispersion is given by

$$\sigma_r^2(r) = -\frac{1}{\pi n(r)} \int_r^\infty \frac{d[N(R)\sigma_p^2(R)]}{dR} \frac{dR}{\sqrt{R^2 - r^2}}. \quad (2.3.10)$$

Deprojecting  $\sigma_p(R)$  into either  $\sigma_r$  or  $\sigma_t$  for a nonzero anisotropy profile is possible but significantly more complicated (see, *e.g.*, Equations 10 and 29–33 in [218]). With the observables deprojected, the mass profile can be determined from a specialized form of the collisionless Boltzmann equation [219] known as the stationary Jeans equation [218];

$$\nabla \cdot (n \hat{\sigma}^2) = -n \nabla \Phi, \quad (2.3.11)$$

where  $\Phi$  is the gravitational potential and  $\hat{\sigma}^2$  is the velocity-dispersion tensor. Equation 2.3.11 is valid provided the cluster is gravitationally relaxed. It is essentially a differentiated form of the virial theorem in which the mass profile of a system ( $\Phi \propto \mathcal{M}(r)$ ) is related to its velocity-dispersion profile. After imposing spherical symmetry, the stationary Jeans equation can be rewritten to express the mass profile in terms of the (deprojected) number-density, radial velocity-dispersion, and anisotropy profiles;

$$\mathcal{M}(r) = -\frac{r\sigma_r^2(r)}{G_N} \left( \frac{d \ln n(r)}{d \ln r} + \frac{d \ln \sigma_r^2(r)}{d \ln r} + 2\beta_a(r) \right). \quad (2.3.12)$$

Generally,  $\mathcal{M}(r)$  cannot be deduced directly from observations because  $\beta_a$  is unknown. Either  $\mathcal{M}(r)$  can be determined given  $\beta_a(r)$ , or  $\beta_a(r)$  can be

### 2.3. DARK MATTER IN GALAXY CLUSTERS

determined given  $\mathcal{M}(r)$ . Sometimes referred to as the “mass-anisotropy” degeneracy, this is usually an intractable problem with the direct Jeans method outlined above [216]. A popular solution is to take an inverse approach and assume functional forms for  $n(r)$ ,  $\sigma_r(r)$ , and  $\beta_a(r)$  [220]. The observed number-density and radial velocity-dispersion profiles are matched to these forms by finding the best-fit parameters that satisfy

$$N(R)\sigma_p^2(R) = 2 \int_R^\infty \left(1 - \beta_a(r)\frac{R^2}{r^2}\right) \frac{n(r)\sigma_r^2(r)rdr}{\sqrt{r^2 - R^2}}. \quad (2.3.13)$$

The mass profile can then be derived from Equation 2.3.12. A number of inverse Jeans methods of this sort can be found in the literature. A more general technique developed by van der Marel *et al.* assumes a 5-parameter form for  $N(R)$ , a constant value for  $\beta_a$ , and a mass-density profile

$$\rho(r) = \frac{\rho_0}{(r/a_r)^\gamma [1 + (r/a_r)]^{3-\gamma}}, \quad (2.3.14)$$

where  $\rho_0$ ,  $a_r$  and  $\gamma$  are free parameters that describe the properties of a cluster’s dark halo [221]. When  $\gamma = 1$  this expression for  $\rho(r)$  is identical to the NFW profile in Equation 2.2.22. A third method models  $N(R)$  and  $\beta_a(r)$ , and assumes  $\rho(r) \propto n(r)$  [222].

Each inverse Jeans method has certain advantages as well as shortcomings, but all tend to overestimate the virial mass of a cluster. The exclusion of a cluster’s outermost galaxies—observational data tend to not extend to the very edge of the cluster—is an  $\sim 20\%$  effect [223]. Most modern virial-mass estimates include the so-called “surface-term” correction (see, *e.g.*, Equation 8 in [222]) to account for the excluded galaxies. A more serious systematic has to do with an apparent disparity between the velocity distributions of early- and late-type member galaxies (*e.g.*, ellipticals and spirals, respectively). While early-type galaxies are largely virialized with isotropic orbits, late-type galaxies are often found in radial orbits that are indicative of recent infall into the cluster and are therefore in violation of the virialization assumption [224]. The inclusion of late-type galaxies in a sample tends to bias  $\sigma_p$  high and can cause as much as a 25% overestimate of  $\mathcal{M}_{\text{vir}}$  [225].



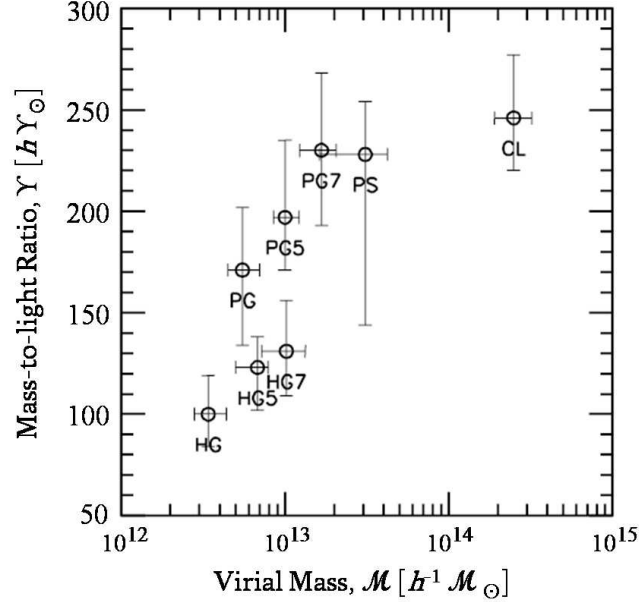


Figure 2.9: Mass-to-light ratios for a variety of galaxy-system samples as a function of their virial masses. The different labels represent: CL, sample of 119 clusters; PS, sample of 43 rich groups and poor clusters; HG, sample of 475 groups identified with a hierarchical technique; and PG, sample of 513 groups identified with a percolation technique (see [226] for further details regarding the hierarchical and percolation group-finding methods). The PG and HG labels followed by the numbers 5 or 7 are subsets with at least 5 or 7 galaxy members, respectively. The error bars are 90% confidence intervals, while the circles indicate mean values. Figure taken from [227].

Average virial-mass estimates for several galaxy-system samples are shown in Figure 2.9. Each error bar represents a particular cluster or group sample, with labels as explained in the following list:

- **CL**, 119 clusters selected from [222], each with at least 30 member galaxies. Galactic luminosities for 89 clusters were derived from the COSMOS catalog [228], while for 52 clusters the Revised APS catalog [229] was used. The 22 clusters with magnitudes in both catalogs were used to cross-calibrate luminosities for the full sample.

### 2.3. DARK MATTER IN GALAXY CLUSTERS

- **PS**, a total of 43 poor systems consisting of 36 poor clusters (with 4 or more member galaxies) selected from [230] and 7 rich groups selected from [231]. Luminosities were obtained from the COSMOS and Revised APS catalogs.
- **HG (PG)**, 475 (513) loose groups chosen from Giuricin *et al.*'s Nearby Optical Galaxy catalog [226] based on hierarchical (percolation) identification of groups from the Lyon-Meudon Extragalactic Database of galaxies [232]. Rich subsets are indicated in Figure 2.9 by the addition of a 5 or 7 in the label, indicating that all groups within the subset have at least 5 or 7 member galaxies. The HG and PG samples are not mutually exclusive.

Virial-mass estimates for the CL and PS samples were performed using the inverse Jeans methods outlined in [222] and [233], and include the surface-term correction mentioned above. Estimating mass for the HG and PG samples is complicated by their low member statistics and uncertainty concerning the degree to which the virialization assumption is valid. To account for the HG and PG groups' dynamical states, their virial-mass estimates were further corrected according to simulations of the collapse of galaxy systems within an expanding spacetime [234]. The average accuracy of the final mass estimates for all samples is approximately 40%.

Figure 2.9 also shows each sample's mass-to-light ratio. The requisite luminosity measurements are no less complicated than the virial-mass estimates; I refer the interested reader to [227] for further details. The CL sample consists of the largest systems considered in the study and is therefore the most likely to be representative of the Universe as a whole. Applying the Oort method to its mass-to-light ratio yields  $\Omega_{\mathcal{M}} \simeq 0.15\text{--}0.18$ , a range that is still somewhat short of the value preferred by the Standard Cosmology ( $\Omega_{\mathcal{M}} \simeq 0.27$ ). The value of the luminosity density used to estimate  $\Upsilon_c$  (see Equations 2.1.3 and 2.1.4) could be the source of the discrepancy. Values as low as  $\mathcal{L} = 1.65 \times 10^8 h L_{\odot} \text{ Mpc}^{-3}$  [235] and as high as  $\mathcal{L} = 2.9 \times 10^8 h L_{\odot} \text{ Mpc}^{-3}$  [236] are consistent with the  $1\sigma$  confidence intervals reported in the literature. Cor-

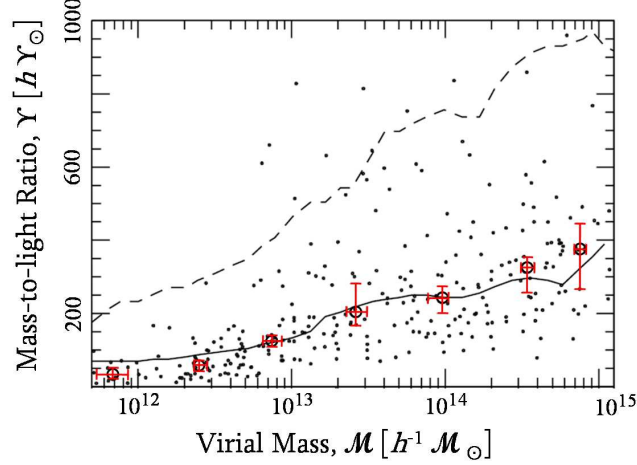


Figure 2.10: Individual (black dots) and average (90% C.L. error bars) mass-to-light ratios as a function of virial mass for a combined sample of 294 galaxy clusters and groups, compared to Kauffmann *et al.*'s  $\Lambda$ CDM (solid line) and  $\tau$ CDM (dashed line) theoretical predictions [239]. The former is roughly consistent with the Standard Cosmological model, while the latter is a high-density model with  $\Omega_{\mathcal{M}} = 1$ . Figure taken from [227].

recting the CL Oort estimate to the latter yields an upper limit of  $\Omega_{\mathcal{M}} \simeq 0.27$ , which agrees with the cosmologically preferred value to within measurement errors. A more recent measurement based on high-quality photometric data for  $\sim 10,000$  galaxies from the Millennium Galaxy Catalogue [237] finds  $\mathcal{L} = (1.99 \pm 0.17) \times 10^8 h L_{\odot} \text{Mpc}^{-3}$  [238]. The corresponding upper limit on  $\Omega_{\mathcal{M}}$  based on Girardi *et al.*'s CL mass-to-light ratio is  $\Omega_{\mathcal{M}} \simeq 0.21$ .

A more likely resolution to this apparent discrepancy between the cosmologically preferred and cluster-derived values of  $\Omega_{\mathcal{M}}$  is that the Oort method used to derive the latter is too naïve in its assumptions [227]. The data in Figure 2.9 appear to support a correlation between  $\Upsilon$  and  $\mathcal{M}_{\text{vir}}$ . To test this notion, Girardi *et al.* define a combined sample of 294 clusters and groups that includes the entire CL and PS samples, and all groups common to the both the PG and HG samples that have at least 5 members. Fitting in the mass versus

### 2.3. DARK MATTER IN GALAXY CLUSTERS

luminosity plane, they find [227]

$$\frac{\mathcal{M}_{\text{vir}}}{\mathcal{M}_{\odot}} = 10^{-1.6 \pm 0.4} \left( \frac{L}{L_{\odot}} \right)^{1.34 \pm 0.03}. \quad (2.3.15)$$

The statistical significance for a luminosity exponent  $>1$  is undeniable; as shown in Figure 2.10, the data support a mild correlation between  $\Upsilon$  and  $\mathcal{M}_{\text{vir}}$ . Clearly the Oort method is insufficient for reconciling this trend with the cosmologically preferred value of  $\Omega_{\mathcal{M}}$ . Simulations of CDM structure formation, however, successfully reproduce the observed behavior [239, 240]. In particular, Kauffmann *et al.*'s  $\Lambda$ CDM theoretical prediction compares favorably with the virial-mass-binned mass-to-light ratios of Girardi *et al.*'s combined sample. The comparison is shown in Figure 2.10 and is based on a value for  $\mathcal{L} \simeq 2 \times 10^8 h L_{\odot} \text{Mpc}^{-3}$  [241], for which the Oort method yields  $\Omega_{\mathcal{M}} \simeq 0.18$  for the CL sample. Using the lower value for  $\mathcal{L}$  given in Equation 2.1.3 causes the  $\Lambda$ CDM prediction to overestimate  $\Upsilon$  by  $\sim 50\text{--}100\%$  while retaining the overall trend as a function of virial mass [240] (see, *e.g.*, Figure 12 in [227]).

#### 2.3.2 X-Ray Emission

Like elliptical galaxies, clusters are permeated by x-ray emitting ionized gas that can be used to trace their gravitational potentials. The methodology for estimating cluster mass from x-ray temperature is essentially the same as was described in Section 2.2.3 for elliptical galaxies. Assuming the gas is in hydrostatic equilibrium, the mass profile can be derived from Equation 2.2.34 by modeling the gas density according to the  $\beta$ -model given in Equation 2.2.35. However, due to the relative size of clusters compared to galaxies, the angular resolution available with current x-ray observatories makes it possible to characterize the temperature's radial dependence in far greater detail than is possible for elliptical galaxies;  $dT/dr$  can be precisely measured. This opens up the possibility for more sophisticated analyses.

One use of cluster-mass estimates derived from x-ray temperatures is as an independent confirmation of optically-based virial estimates, thereby substantiating any conclusions based on the latter regarding the Standard Cosmology

and dark matter. Obtaining enough statistics for such a comparison can be difficult as there is no guarantee that clusters identified from optical surveys (*e.g.*, the SDSS [130]) will have corresponding signatures in x-ray surveys (*e.g.*, the ROSAT (*Röntgensatellit* [242, 243]) All Sky Survey, or RASS [244]) and *vice versa*. Girardi *et al.* compare x-ray- and virial-mass estimates for 66 clusters and find reasonable agreement between the two techniques (see, *e.g.*, Figure 5 in [222]). Popesso *et al.* make a similar comparison based on a smaller but more accurately measured sample of 16 clusters. They find a tendency for virial masses to be overestimated relative to their x-ray counterparts, but only by a factor of a few (see, *e.g.*, Figure 1 in [225]).

It is worth noting that a cluster’s x-ray luminosity  $L_x$  is also a useful mass proxy. Analyses of large samples of galaxy groups and clusters support the interchangeability of x-ray luminosity and temperature through the simple scaling relation  $L_x \propto T^3$  [245, 246, 247]. This is a manifestation of clusters’ so-called “self-similar” nature. Since gravity has no preferred scales, if gravity alone determines the thermodynamic properties of the intracluster gas, clusters of different sizes should be scaled versions of one another, or self-similar [248]. With representative measurements from the literature, the expected self-similar scaling relations for nearby clusters are [249, 250]:

- $\mathcal{M}_{\text{tot}} \propto T^{3/2}$ , the  $\mathcal{M}$ – $T$  relation. The logarithmic slope of this relation has been measured to be roughly consistent with the expected 3/2 exponent (see, *e.g.*, Figures 2–3 in [251] or Figure 9 in [252]).
- $L_x \propto T^2$ , the  $L_x$ – $T$  relation. As mentioned above, the logarithmic slope of this relation has been measured to be  $\sim 3$  (see, *e.g.*, Figure 13 in [253], Figure 8 in [246], or Figure 1 in [247]).
- $L_x \propto \mathcal{M}_{\text{tot}}^{4/3}$ , the  $L_x$ – $\mathcal{M}$  relation. The logarithmic slope of this relation has been measured to be  $\sim 1.3$ – $1.8$  (see, *e.g.*, Figure 6 in [245], Figure 11 in [254], or Figure 7 in [255]).

The observed variation in the logarithmic slopes of these scaling relations appears to be due (in part) to the inclusion or exclusion of x-ray observations

### 2.3. DARK MATTER IN GALAXY CLUSTERS

from the clusters' central-most regions. The intracluster gas harbored within the cores of clusters is evidently dense enough for nongravitational effects to be important, violating the naïve gravity-only assumption from which the above self-similar relations are derived. For example, if the central regions of clusters are assumed to have been preheated prior to cluster collapse, subsequent radiative cooling can lead to lower temperatures for a given (x-ray) luminosity [256]. When this is taken into account in structure-formation models, the expected  $L_x$ – $\mathcal{M}$  and  $L_x$ – $T$  logarithmic slopes respectively increase to 11/6 and 11/4, in better agreement with observations [257].

More generally, cluster self-similar relations depend on the underlying cosmology of the Universe [258]. For example, the  $\mathcal{M}$ – $T$  relation at arbitrary redshift is

$$\mathcal{M}_{\text{tot}} \propto \frac{T^{3/2}}{E(z)}, \quad (2.3.16)$$

where  $E(z)$  is the ratio of the Hubble parameter to the Hubble constant (originally defined in Equation 1.2.7). For a flat spacetime with constant dark-energy equation of state,

$$E(z) = \sqrt{\Omega_{\mathcal{M}}(1+z)^3 + \Omega_{\Lambda}(1+z)^{3(1+w)}}. \quad (2.3.17)$$

The radiation term in Equation 1.2.7 has been dropped from this definition of  $E(z)$  because  $\Omega_R$  and the redshifts at which clusters are observed are small enough for it to be negligible relative to the other terms. Observations of x-ray temperatures and luminosities for clusters over a range of redshifts can thus be used to trace cosmic evolution and constrain cosmological parameters.

A widely used method for constraining cosmological parameters with x-ray-derived cluster masses is through construction of the cluster-mass function  $\mathcal{N}( > \mathcal{M}, z )$ —the number of clusters with mass greater than  $\mathcal{M}$  as a function of  $\mathcal{M}$  (the total cluster mass) and redshift. In a series of papers by Vikhlinin *et al.* [259, 260, 261], the cluster-mass function is constructed from temperature-profile and luminosity observations made with the Chandra x-ray observatory [201] for a high- $z$  sample ( $\bar{z} \simeq 0.55$ ) of 37 clusters and a low- $z$

sample ( $\bar{z} \simeq 0.05$ ) of 49 clusters. Cluster masses are estimated from scaling relations between total mass and three mass proxies:

- The  $\mathcal{M}$ – $T$  relation is calibrated using average x-ray temperatures measured over an annulus that excludes a central region  $\sim 3/10$  of the virial radius across to avoid the effects of radiative cooling. In this case, cluster masses are estimated according to the method introduced for elliptical galaxies in Section 2.2.3. The logarithmic slope of the scaling relation is measured to within  $\sim 7\%$  of  $3/2$ , the expected self-similar value (see, *e.g.*, Figure 7 in [260]). Also, the  $E(z)$  scaling in Equation 2.3.16 is confirmed.
- The  $\mathcal{M}$ – $\mathcal{M}_g$  relation between total cluster mass and the total mass of the intracluster gas is estimated.  $\mathcal{M}_g$  is obtained by integrating the gas density as derived from multi-parameter  $\beta$ -model fits (see, *e.g.*, Equation 2 in [260]) to the observed x-ray-flux profiles. In this case, the total mass is estimated according to a method commonly referred to as the baryon-mass fraction technique. To leading order, provided the baryonic content of the intracluster gas is representative of the Universe as a whole, the *total* cluster mass can be approximated as

$$\mathcal{M}_{\text{tot}} \simeq \mathcal{M}_g f_b^{-1}, \quad (2.3.18)$$

where  $f_b = \Omega_b/\Omega_{\mathcal{M}}$  is the cosmic baryon fraction and is specified according to independent cosmological evidence. The actual procedure used is a bit more complicated as it has to account for an apparent deficit of baryons in the intracluster gas (relative to the Universe as a whole). From Equations 1.3.27 and 1.3.28, the cosmologically preferred value of the baryon fraction is  $f_b = 0.17 \pm 0.01$ , while Vikhlinin *et al.* find that  $f_b$  exhibits a mild correlation with  $\mathcal{M}_{\text{tot}}$  for their cluster samples, ranging from  $\sim 0.08$  for clusters with  $\mathcal{M}_{\text{tot}} \simeq 10^{14} h^{-1} \mathcal{M}_{\odot}$  to  $\sim 0.12$  for clusters with  $\mathcal{M}_{\text{tot}} \simeq 10^{15} h^{-1} \mathcal{M}_{\odot}$  (see, *e.g.*, Figure 9 in [260]). Considering intracluster-gas x-ray observables are insensitive to the stellar baryons locked within a cluster’s member galaxies, the discrepancy is not too surprising. Regardless, the measured high-mass asymptotic value of  $f_b$  is

### 2.3. DARK MATTER IN GALAXY CLUSTERS

an important result. As was alluded to in Sections 2.1 and 2.2, galactic (luminous) baryon densities fall short of the cosmologically preferred value of  $\Omega_b$ .  $f_b \simeq 0.12$  indicates that most of the Universe's baryons are located in the hot x-ray emitting gas clouds of clusters, and not in stars and galactic gas.

- The  $\mathcal{M}$ – $Y_x$  relation is estimated from the product of the previous two.  $Y_x = T\mathcal{M}_g$  corresponds to the total thermal energy of the intracluster gas. It also corresponds to the integrated low-frequency Sunyaev-Zel'dovich flux (discussed further in Section 2.3.4). To distinguish the two very different techniques for measuring  $Y$ , I will use  $Y_x$  to refer to the thermal energy as measured with x-ray observables and  $Y_{sz}$  to refer to the SZ flux as estimated from CMB observations. The logarithmic slope of the  $\mathcal{M}$ – $Y_x$  relation is found to agree to within  $\sim 5\%$  of the self-similar prediction [262]:

$$\mathcal{M}_{\text{tot}} \propto Y_x^{3/5} E(z)^{-2/5}. \quad (2.3.19)$$

The  $L_x$ – $\mathcal{M}$  scaling relation is measured but not used to (directly) estimate mass (see, *e.g.*, Figure 12 in [260]). Rather, it plays a critical role in converting the sampled cluster volumes from functions of luminosity to functions of mass. For each redshift slice, the mass and volume estimates are combined using a likelihood technique to obtain the cluster-mass functions shown in Figure 2.11.

Also shown in Figure 2.11 are  $\Lambda$ CDM fits to the cluster mass functions. A detailed discussion of the  $\Lambda$ CDM functional form used to fit the data is beyond the scope of this thesis. I refer the interested reader to Voit's review on the subject in [258]. Vikhlinin *et al.* use a universal form derived by Tinker *et al.* from simulations of dark-halo formation in a flat  $\Lambda$ CDM cosmology [263]:

$$\frac{d\mathcal{N}}{d\mathcal{M}} = f(\sigma) \frac{\Omega_{\mathcal{M}} \rho_c (1+z)^3}{\mathcal{M}} \frac{d \ln \sigma^{-1}}{d\mathcal{M}}, \quad (2.3.20)$$

where  $\sigma$  is the root-mean-square (rms) variance of density fluctuations at an arbitrary distance scale and redshift, and

$$f(\sigma) = C_0 [(\sigma/c_2)^{-c_1} + 1] \exp(-c_3/\sigma^2) \quad (2.3.21)$$



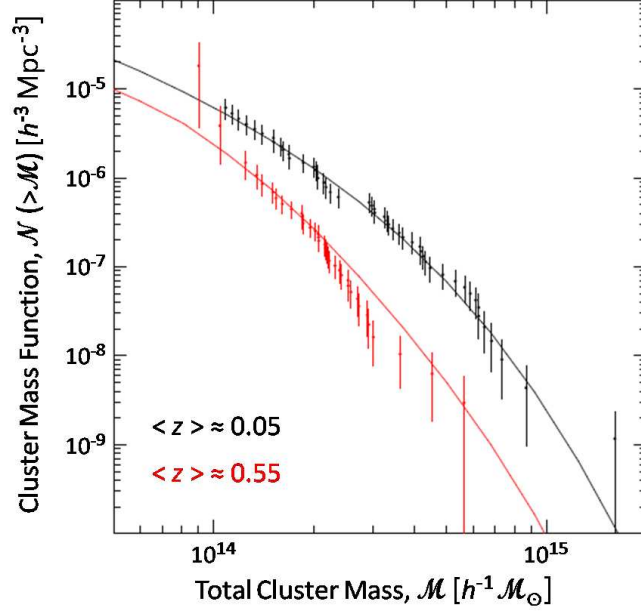


Figure 2.11: The cluster-mass function for high- $z$  (red error bars) and low- $z$  (black error bars) clusters. Cosmological-model fits (solid lines) are shown as well. The overall agreement supports the minimal  $\Lambda$ CDM parameters reported in Chapter 1, despite the  $\sim 2\sigma$  deficit of high- $z$ , high-mass clusters in the data. Figure taken from [261].

is a universal form for which  $C_0$ ,  $c_1$ ,  $c_2$ , and  $c_3$  are constants determined by simulation. The normalization of this cluster-mass function is sensitive to  $\sigma$ , while the shape is sensitive to  $\Omega_M$ . Implicit to the data and fits in Figure 2.11 are assumed values for  $h$ ,  $\Omega_M$ , and  $\Omega_\Lambda$ , while the overall normalization is determined by fitting for the cosmological parameter  $\sigma_8$ , the value of  $\sigma$  smoothed over an  $8 h^{-1}$  Mpc scale and evaluated at  $z = 0$ . The normalization of the low- $z$  mass-function suggests  $\sigma_8 \simeq 0.8$ , in agreement with the value preferred by the minimal  $\Lambda$ CDM parameters in Table 1.2 ( $\sigma_8 \simeq 0.81$  [24]). The shape of the low- $z$  mass function prefers  $\Omega_M \simeq 0.26$ , while the relative change in the shape of the high- $z$  mass function prefers  $\Omega_M \simeq 0.34$  [261]. In either case, the data suggest that clusters are dominated by nonbaryonic dark-matter halos.

### 2.3. DARK MATTER IN GALAXY CLUSTERS

The cluster-mass function can also be constructed as a function of x-ray luminosity or temperature, in which case it is referred to as the cluster-luminosity or -temperature function, respectively. Mullis and collaborators examine the evolution of the cluster-luminosity function for a sample  $\sim 200$  galaxy systems, and find consistency with the evolution of structure expected within the context of the Standard Cosmology [264]. Patrick constructs the cluster-temperature function from a few dozen low- $z$  and a few dozen high- $z$  clusters (see, *e.g.*, Figure 2 in [265]). The evolution of his temperature function is consistent with a  $\Lambda$ CDM model in which  $\Omega_{\mathcal{M}} \simeq 0.28$ ,  $\Omega_{\Lambda} \simeq 0.98$ , and  $\sigma_8 \simeq 0.7$ .

#### 2.3.3 Gravitational Lensing

The idea that photons are deflected by gravitational fields in a manner that might be useful for measuring the masses of celestial objects dates back to a series of papers by Chwolson (1920), Einstein (1936) and Zwicky (1937) [266, 267, 136]. It was not until the 1990s, however, that the technique became practical for measuring cluster masses [268, 269, 270]. The premise is simple; light from a distant galaxy is deflected by the gravitational field of a cluster, causing a distortion of the galaxy's image that depends on the cluster's mass. The technique has several advantages over x-ray and virial methods as it is insensitive to a cluster's dynamical state or particular composition. Baryon-mass fraction, intracluster-gas hydrostatic equilibrium, and degree of member-galaxy virialization do not affect lensing measurements. Furthermore, the large galaxy catalogs available from wide-field optical surveys such as the SDSS [130] are well suited to lensing analyses. Consequently, gravitational lensing provides an independent method for confirming cluster masses that avoids several of the systematic uncertainties and statistical limitations of other techniques.

Gravitational lensing occurs when a background source's image is distorted as its photons pass through the gravitational field of a massive object (referred to as a lens). The most famous such distortions are beautiful deep-field images from the Hubble Space Telescope [27] of galaxies bent into arcs or Einstein rings (see, *e.g.*, Figure 3 in [271]). Such dramatic distortions are the result of

what is known as strong gravitational lensing; the light from a source (*e.g.*, a galaxy or quasar) is deflected nonlinearly by a particularly strong gravitational field (*e.g.*, the dense cores of massive clusters or galaxies), often resulting in multiple foreground images of the source. The details of how a source image is distorted depend not only on the size and mass of the lens but also on the angular diameter distances to the lens and the source. Strong-lensing images are sensitive to the underlying cosmology of the Universe through the latter. When at least two sources at different redshifts, say  $z_{s1}$  and  $z_{s2}$ , are strongly lensed by a cluster at redshift  $z_l$ , the cosmological parameters to which the images are sensitive can be disentangled from the properties of the lens provided there are at least two images for each source [272]. In this case, strong-lensing data can be used to constrain the cosmological parameters  $\Omega_{\mathcal{M}}$ ,  $\Omega_{\Lambda}$ , and  $w$  through construction of the so-called “cosmological family ratio” [273]

$$\Xi(z_l, z_{s1}, z_{s2}; \Omega_{\mathcal{M}}, \Omega_{\Lambda}, w) = \frac{d_A(z_l, z_{s1})}{d_A(0, z_{s1})} \frac{d_A(0, z_{s2})}{d_A(z_l, z_{s2})}, \quad (2.3.22)$$

where

$$d_A(z_1, z_2) = \frac{c}{H_0(1+z_2)} \int_{z_1}^{z_2} \frac{dz}{E(z)}, \quad (2.3.23)$$

and  $E(z)$  is given by Equation 2.3.17. In an analysis of 12 sources strongly lensed by the cluster Abell 1689 into 28 distinct images and spanning  $z \simeq 1$ –5, the authors in [274] derive combined  $w$ – $\Omega_{\mathcal{M}}$  constraints that are roughly orthogonal to and consistent with constraints derived from the WMAP CMB power spectrum and the x-ray-derived cluster-mass function (see, *e.g.*, Figure 2 in [274]).

While strong gravitational lensing is a useful tool for studying cosmology, it is not an entirely appropriate mass proxy for large cluster samples. Extracting the mass profile of a cluster from strong-lensing data typically involves a complicated nonlinear analysis of deep-field optical images. Useful information can be acquired on a cluster-by-cluster basis, but the technique is not readily applicable to the large statistical samples available from wide-field surveys. Furthermore, a cluster’s total mass does not necessarily contribute to a strong lens. It is only at the dense cores of clusters (or galaxies) that gravitational

### 2.3. DARK MATTER IN GALAXY CLUSTERS

fields are strong enough to cause the characteristic arcs and rings associated with strong lensing. In some cases, a strong-lensing analysis can be performed as an independent confirmation of the central-most region of a cluster’s mass profile [275, 276].

The mass profile for an ensemble of clusters can be probed statistically through a technique known as weak gravitational lensing. While there is no concrete definition of weak lensing, it generally refers to the very slight changes in ellipticity and orientation of a background galaxy’s image due to the gravitational field of a massive foreground object (such as a cluster). On an individual basis, a weak-lensing distortion provides very little information regarding a cluster’s mass profile due to the large variety of galaxy shapes and possible orientations. Typical weak-lensing signals for individual galaxy images amount to changes on the order of  $\sim 1\%$  (or less) to the image’s original shape, while the intrinsic variability among galaxy shapes is  $\sim 30\%$  [277]. The weak-lensing signal-to-noise for a typical galaxy image is therefore too small to be useful on an individual basis. However, when a large enough ensemble of field galaxies is considered, the galaxies’ intrinsic shapes and random orientations average away. Any correlations among the ellipticities and alignments of such an ensemble’s galaxy images indicate the presence of a lens and provide a sensitive statistical measure of its gravitational field. On a cluster-by-cluster basis, even these ensemble correlations are prone to large systematic uncertainties. For example, lensing due to voids or other clusters along or closely adjacent to the observational line-of-sight can affect measurements of individual cluster masses in an unpredictable manner [278, 279]. A common workaround is to combine weak-lensing signals *en masse*—an ensemble of lenses, each with an ensemble of sources—such that these effects are also averaged away. For a large enough sample of clusters, the combined volume of space between the observer and source galaxies is so large that it is representative of the Universe as a whole. Intervening voids and clusters that might adversely affect an individual lensing measurement will have no net effect on the average lensing measurement of the ensemble because the density contained in the relevant spatial volume is homogenous and isotropic. This sort of ensemble averaging is often called “stacking.”

## CHAPTER 2. DARK MATTER

The extraction of cluster masses from stacked weak-lensing images relies on a complicated formalism that is well beyond the scope of this thesis. I refer the interested reader to Bartelmann and Schneider’s extensive review in [280] for a detailed treatment of the subject. The following discussion briefly highlights a technique developed by Bernstein and Jarvis that has been widely adopted by the community and is described in [277]. The principle weak-lensing observable is the shape of a source galaxy’s image. The observed shape can be described by the ellipticity components  $e_+$  and  $e_\times$ , the E- and B-mode ellipticities relative to the center of the lensing cluster, respectively.  $e_+$  is measured tangentially with respect to concentric circles centered on the lensing cluster, and is naturally a function of the projected cluster-centric radius  $R$ .  $e_\times$  is measured radially at  $45^\circ$  relative to the  $e_+$  tangents. On average, the  $e_\times$  component is expected to be zero if the galaxy-image distortions are due entirely to the lensing cluster’s gravitational field [281]. Measurement of  $e_\times$  provides a useful crosscheck on nonlensing sources of distortion that might contribute to both the E and B modes (*e.g.*, atmospheric effects or instrument aberrations). The observed E-mode ellipticity is related to the intrinsic shape of the galaxy image,  $e_+^{\text{int}}$ , through the tangential shear;

$$\gamma_T = \frac{1}{2\mathcal{R}} (e_+ - e_+^{\text{int}}), \quad (2.3.24)$$

where  $\mathcal{R}$  is a galaxy’s responsivity to shear and can be measured on average from an ensemble of galaxy images (see [277] for further details). An estimate for  $\gamma_T(R)$  is obtained by averaging Equation 2.3.24 azimuthally in bins of  $R$ . The contribution from the intrinsic ellipticity falls out as  $\langle e_+^{\text{int}} \rangle$  is expected to be zero provided the number of galaxy images in each bin of  $R$  is large enough.  $\gamma_T$  is related to the cluster’s projected matter density according to

$$\Delta\Sigma(R) \equiv \gamma_T(R) \times \Sigma_{\text{crit}} = \bar{\Sigma}(< R) - \bar{\Sigma}(R), \quad (2.3.25)$$

where  $\bar{\Sigma}(< R)$  is the average projected matter density within the projected disc interior to  $R$ , and  $\bar{\Sigma}(R)$  is the average within the projected annulus defined by a given bin of  $R$  [282]. The geometry of the lens-observer system is encoded in

### 2.3. DARK MATTER IN GALAXY CLUSTERS

the quantity [281]

$$\Sigma_{\text{crit}}^{-1} = \frac{4\pi G_N d_A(z_l, z_s) d_A(0, z_l)}{c^2 d_A(0, z_s)}, \quad (2.3.26)$$

where  $z_l$  and  $z_s$  are the redshifts of the lens and source, respectively. In addition to the weak-lensing images from which  $\gamma_T$  is estimated, photometric redshift measurements are required for the source galaxies and lensing cluster in order to estimate  $\Sigma_{\text{crit}}$  and thereby convert the tangential shear to a projected matter overdensity. The full 3-dimensional matter overdensity can be unfolded from  $\Delta\Sigma$  according to [282]

$$\Delta\rho(r) \equiv \rho(r) - \bar{\rho} = -\frac{1}{\pi} \int_r^\infty \frac{dR}{\sqrt{R^2 - r^2}} \frac{d\Sigma}{dR}, \quad (2.3.27)$$

where

$$-\frac{d\Sigma}{dR} = \frac{d\Delta\Sigma}{dR} + 2\frac{\Delta\Sigma}{R}, \quad (2.3.28)$$

and  $\bar{\rho}$  is the mean density of the Universe. Equations 2.3.27 and 2.3.28 are only valid provided the geometry of the lens is spherically symmetric, an accurate approximation for a large enough stack of clusters. Finally, the average 3-dimensional mass profile for the cluster stack is given by

$$\mathcal{M}(r) = \pi r^2 \Delta\Sigma(r) + 2\pi \int_r^\infty R \Delta\rho(R) \left[ \frac{r^2}{\sqrt{R^2 - r^2}} - 2 \left( R - \sqrt{R^2 - r^2} \right) \right] R dR, \quad (2.3.29)$$

the total excess mass within a spherical volume of radius  $r$  [283]. To a very good approximation,  $\mathcal{M}(r)$  represents the average total mass of the stacked clusters.

This technique is applied to a selection of  $\sim 28$  million SDSS galaxy images in a series of papers by Sheldon *et al.* [282, 285, 283]. Over 130,000 clusters are identified and sorted according to their number of member galaxies, or richness (not to be confused with the number of lensed source galaxies). Each richness bin represents an independent cluster stack for which an average mass estimate is obtained using the method described above. The corresponding average luminosity is estimated using standard methods (see [283] for details).

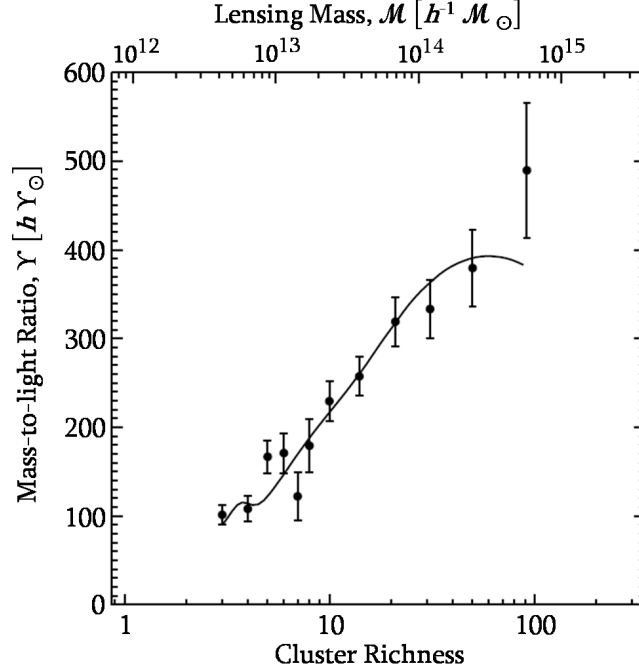


Figure 2.12: The average mass-to-light ratios for over 100,000 galaxy systems identified from the SDSS [130] and binned according to the number of member galaxies, or richness (bottom axis), and plotted as a function of total mass (top axis) as inferred from weak gravitational-lensing measurements. A model based on simulations by Tinker *et al.* (described in [284]) is roughly overlaid for comparison (solid line). The asymptotic, high-mass value of the mass-to-light ratio approximates the matter density of the Universe;  $\Omega_{\mathcal{M}} \simeq 0.21\text{--}0.28$ . Figure taken from [283].

The resulting mass-to-light ratios as a function of cluster richness are shown in Figure 2.12. For comparison, a theoretical prediction based on large-scale structure simulations (including dark matter) by Tinker *et al.* is roughly overlaid (but not fit). The spatial volume associated with the richest stacks is large enough that the associated mass-to-light ratio is expected to be representative of the Universe as a whole. Applying the Oort method to the asymptotic, high-mass value of  $Y$  in Figure 2.12 yields  $\Omega_{\mathcal{M}} \simeq 0.21\text{--}0.28$ , consistent with the

### 2.3. DARK MATTER IN GALAXY CLUSTERS

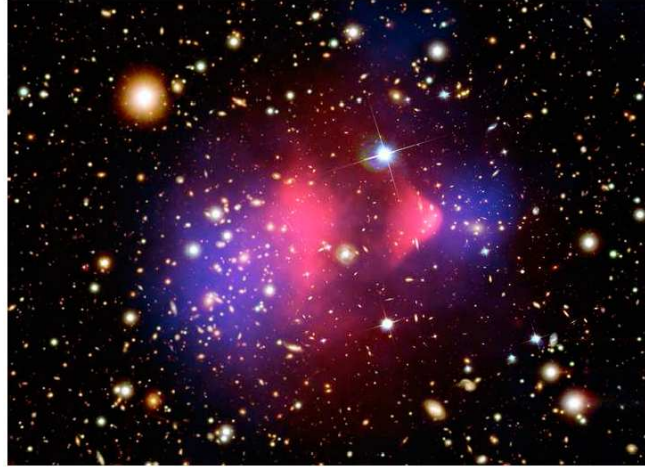


Figure 2.13: False-color image of the interacting cluster 1E 0657–558, also known as the Bullet cluster. Two intracluster-gas clouds (pink) appear to lag behind their respective clusters’ main matter distributions (blue). The former are inferred from a Chandra x-ray image, while the latter are reconstructed surface-mass density maps from a weak-lensing analysis. The mass-to-light ratios associated with the weak-lensing analysis are consistent with significant nonbaryonic dark-matter halos. The prevailing interpretation is that the clusters’ member galaxies and dark halos passed through one another relatively intact, while their intracluster-gas clouds were stripped away by drag forces such that they appear to lag behind. Photograph taken from [286].

cosmologically preferred value and indicative of significant nonbaryonic dark halos on cluster scales.

#### The Bullet Cluster

The interacting galaxy cluster 1E 0657–558, also known as the Bullet cluster, is one of the most famous applications of the weak-lensing technique. An analysis of its Chandra [201] x-ray image by Markevitch and collaborators in 2002 first deduced the presence of a bow shock propagating in front of a cloud of intracluster gas in the shape of a bullet [287]. Their x-ray temperature



measurements are depicted in false-color pink in Figure 2.13. The inferred intracluster-gas densities and velocities can be interpreted as the aftermath of the collision of two galaxy clusters. The clusters essentially collisionless member galaxies appear well separated from the fluid-like x-ray-emitting intracluster gas (see, *e.g.*, Figure 1 in [288]). Furthermore, weak-lensing estimates of the two clusters' mass profiles indicate that the majority of their matter densities (depicted in false-color blue in Figure 2.13) are also well separated from their respective intracluster-gas clouds [289]. Mass-to-light ratios based on these weak-lensing estimates are consistent with those shown in Figure 2.12 and inconsistent with the cluster's similarly (spatially) distributed member galaxies. The basic conclusion is that the two clusters' member galaxies and dark halos passed through one another relatively intact, while their intracluster-gas clouds were stripped away by drag forces such that they appear to lag behind. Most agree that the Bullet cluster provides compelling evidence for the existence of particulate dark matter as an explanation of the dark-matter problem (rather than a modification of gravitational theory).

### 2.3.4 The Sunyaev-Zel'dovich Effect

Yet another method for estimating cluster masses is via the thermal Sunyaev-Zel'dovich (tSZ) effect [212]. Due to space considerations, I will only briefly review this technique. For detailed discussions, I refer the interested reader to the reviews by Rephaeli and Birkinshaw [290, 291]. As CMB photons pass through a cluster they can inverse Compton scatter from free electrons in the intracluster gas. The free electrons are relatively hot and tend to increase the frequency of the CMB radiation, resulting in an increased (decreased) number of high-frequency (low-frequency) CMB photons while conserving the total flux. A decrement in CMB intensity is expected for wavelengths  $\gtrsim 1.3$  mm, while an increment is anticipated for shorter wavelengths in excess of  $\sim 250$   $\mu$ m, below which the tSZ effect is expected to be small [292]. For a fixed CMB frequency  $\nu$ , it can be shown that the resulting fractional change in CMB intensity is

### 2.3. DARK MATTER IN GALAXY CLUSTERS

approximately given by<sup>5</sup>

$$\frac{\Delta I(x)}{I_0} = Y_{sz} \left( \frac{x^4 e^x}{(e^x - 1)^2} \left( x \frac{e^x + 1}{e^x - 1} - 4 \right) + \delta(x, T_e) \right), \quad (2.3.30)$$

where  $T_e$  is the electron temperature (equal to the x-ray temperature of the intracluster gas),  $x \equiv 2\pi\hbar\nu/kT_{\text{cmb}}$  is the dimensionless frequency of the CMB radiation, and  $I_0 = 2(kT_{\text{cmb}})^3/(2\pi\hbar c)^2$  [294].  $\delta(x, T_e)$  is a correction term that is zero for nonrelativistic electrons but can be important for hot clusters (see, *e.g.*, [295] for further details). The normalization is equivalent to the total thermal energy of the intracluster gas and is given by

$$Y_{sz} = \frac{\sigma_T}{m_e c^2} \int kT_e n_e dl, \quad (2.3.31)$$

where  $\sigma_T \simeq 0.67 \times 10^{-24} \text{ cm}^2$  is the Thomson scattering cross section and  $n_e$  is the number density of the free electrons in the intracluster gas [296]. The integral in Equation 2.3.30 is performed along the line-of-sight, and  $dl$  is the proper distance along the path through the cluster.  $Y_{sz}$  is commonly referred to as the Comptonization parameter and, as alluded to in Section 2.3.2, is an excellent low-scatter proxy for total cluster mass [297]. It is common practice to model the electron number density according to a  $\beta$ -model such that

$$n_e(r) = n_{e0} \left( 1 + \frac{r^2}{r_c^2} \right)^{-3\beta/2}, \quad (2.3.32)$$

where  $r_c$  is the core radius and  $n_{e0}$  is the central number density. For an isothermal gas of electrons, substituting this expression for  $n_e$  into Equation 2.3.31 and evaluating the line integral gives

$$Y_{sz}(\theta) = Y_0 \left( 1 + \frac{\theta^2}{\theta_c^2} \right)^{(1-3\beta)/2}, \quad (2.3.33)$$

---

<sup>5</sup>An additional term not included in Equation 2.3.30 arises from a cluster's peculiar velocity (relative to the CMB radiation). The so-called kinetic Sunyaev-Zel'dovich (kSZ) effect is proportional to the average line-of-sight streaming velocity of the intracluster gas  $v_p$ , the Thomson optical depth  $\tau = \sigma_T \int n_e dl$ , and a spectral shape that depends on the dimensionless CMB frequency  $x$  (see, *e.g.*, Equations 10–12 in [290]). For particularly hot clusters, the kSZ effect's spectral shape is also (slightly) sensitive to the electron temperature  $T_e$  [293]. The tSZ effect vanishes for  $\nu = 217 \text{ GHz}$ , while the kSZ effect's dependence on frequency is generally weak; the two effects can be disentangled with multi-frequency CMB observations.

where

$$Y_0 = n_{e0} \sigma_T r_c \sqrt{\pi} \frac{kT_e}{m_e c^2} \frac{\Gamma(\frac{3}{2}\beta - \frac{1}{2})}{\Gamma(\frac{3}{2}\beta)}, \quad (2.3.34)$$

$\theta$  is the angle between the direction of interest and the center of the cluster, and  $\theta_c = r_c/d_A$  is the angular core radius of the cluster [291]. Cluster tSZ measurements are sensitive to the underlying cosmology of the Universe through  $\theta_c$ 's dependence on the angular diameter distance  $d_A$ . If CMB observations of a cluster's SZ intensity are made at multiple frequencies, the free parameters  $T_e$ ,  $n_{e0}$  and  $r_c$  can be simultaneously constrained, the tSZ and kSZ effects can be disentangled, and the intracluster-gas mass  $\mathcal{M}_g$  can be estimated from Equation 2.3.32. The total cluster mass follows from Equation 2.3.18. tSZ measurements thus provide an independent method for studying cluster mass and  $\Omega_{\mathcal{M}}$ . Multi-frequency SZ observations for cluster ensembles spanning a wide range of redshifts have the potential to vastly improve constraints on the dark-energy equation of state [298].

While a number of SZ studies have been published detailing projected cosmological sensitivity and complementarity to other methods [299, 300, 301], instruments capable of the type of CMB observations required to study clusters are yet in their infancy. The DMR aboard the WMAP satellite is well suited to wide-field mapping of the CMB sky, but its relatively large angular resolution (10's of arcminutes) is (generally) too coarse for studying clusters [302, 303]. However, data from WMAP and first-generation SZ instruments have been used to study the suitability of  $Y_{\text{sz}}$  as a proxy for total cluster mass [304, 305].

The current generation of SZ telescopes have the arcminute resolution and deep-field capabilities necessary to identify and characterize clusters efficiently. Two of the most promising are the 10 m South Pole Telescope (SPT) [306] located in Antarctica and the 6 m Atacama Cosmology Telescope (ACT) [307] located in the Chilean Atacama Desert. Both detect CMB intensity bolometrically at three frequencies with multi-element (1000 or more) arrays of transition-edge sensors (TES-based detectors are discussed further in Chapter 4). Previously unknown clusters are now being discovered for the first time

## 2.4. DARK MATTER CANDIDATES

from SZ observations made with these highly sensitive telescopes [308, 309, 310]. Though limited in scope, the first SPT and ACT SZ-selected cluster catalogs [311, 312]—roughly two dozen clusters each—have been used to derive improved cosmological constraints on the dark-energy equation of state and  $\sigma_8$  (see, *e.g.*, Figure 5 in [311] and Figure 6 in [313]). Furthermore, the expected self-similar  $Y_{\text{SZ}}-Y_x$  and  $Y_{\text{SZ}}-\mathcal{M}$  scaling relations have been (roughly) confirmed using SPT and ACT clusters for which high-quality x-ray data are available [314, 315, 310], validating the potential of the tSZ technique for studying the large-scale clustering of dark matter in the Universe.

## 2.4 Dark Matter Candidates

From the cosmological evidence presented in Chapter 1 it is clear that  $\sim 27\%$  of the Universe’s energy density is in the form of some kind of gravitating matter. The same cosmological evidence suggests that  $\sim 85\%$  of this matter is nonbaryonic and dark. The arguments presented above demonstrate unequivocally the influence of dark matter in galaxies and clusters, but distinguish less clearly the division between its baryonic and nonbaryonic components. The value of  $\Omega_{\mathcal{M}}$  inferred from clusters is well in excess of the preferred value of  $\Omega_b$ , roughly coinciding with the cosmological-matter budget. On galactic scales, however, the need for nonbaryonic dark matter is less obvious. In most cases, a combination of dark baryons and systematic uncertainties could explain galactic dark-matter phenomena and still be consistent with the value of  $\Omega_b$  derived from BBN and CMB observations. On the other hand, the baryon-mass fraction for clusters indicates that most of the Universe’s baryons are in clouds of intracluster gas and not in galaxies, encouraging the possibility that a significant fraction of the galactic dark matter is nonbaryonic. In any case, it seems prudent to consider both baryonic and nonbaryonic explanations of the dark-matter problem.

There are nearly as many hypothetical dark-matter candidates as there are theorists who propose such things, far too many to describe in detail here. I will concentrate on the most popular (and some might argue well motivated)

SM and exotic possibilities. A few of the dark-matter candidates not covered, but for which there are extensive discussions in the literature, are enumerated in the following list (in no particular order) with references for those readers interested in these (somewhat esoteric) solutions. This list is a mere sampling of the theorists imaginations and is by no means comprehensive; a number of excellent reviews are available in the literature that discuss these and other candidates in greater detail [316, 23, 2, 317].

- **Axions:** As a solution to the strong-CP problem [318, 319, 320], axions are one of the most well motivated dark-matter candidates. If they constitute the dark matter, they would have been produced athermally in the early Universe, allowing them to be simultaneously light and cold. Although largely invisible, axions might be detectable through conversion to photons in a strong magnetic field [321]. Searches for axionic dark matter employ high- $Q$  resonant cavities to enhance the expected conversion rate and target axion masses of  $\sim 1\text{--}100\,\mu\text{eV}\,c^{-2}$  [322, 323]. For further details see the Particle Data Group’s excellent review in [88].
- **MOND:** Modified Newtonian dynamics is a phenomenological approach in which the evidence for dark matter is explained by a modification of the gravitational force on large distance scales, rather than by the gravitational influence of particulate matter. The idea was originally proposed by Milgrom in the 1980s to explain dark matter in galaxies and clusters [196, 197, 324]. A recent review by Milgrom compares MOND and  $\Lambda$ CDM cosmological predictions [325].
- **MOG:** Modified theory of gravity in which fundamental constants such as  $G_N$  are allowed to vary with space and time [326]. In some ways, MOG is a generalized version of MOND. Like MOND it is a modification of gravitational theory that attempts to explain the dark-matter problem without exotic particles. A recent review by Moffat and Toth compares the growth of structure in the early Universe for the MOG and  $\Lambda$ CDM paradigms [327].

## 2.4. DARK MATTER CANDIDATES

- **Mirror Matter:** The violation of parity in SM weak interactions has led some theorists to postulate that parity could be conserved as a global symmetry provided there exist mirror particles of opposite handedness for every SM particle [328, 329, 330]. The only fundamental SM boson identical to its mirror partner would be the graviton. SM and mirror matter would therefore interact gravitationally, but otherwise only very weakly through hypothetical fifth-force bosons or kinetic mixing of ordinary and mirror bosons [331]. The dark matter would then be explained by stable (or long lived) mirror-matter particles or bound states, such as mirror hydrogen atoms [332, 333]. Mirror matter is a specialized version of a hidden-sector (or hidden-valley) physics theory in which part of the theory is “hidden” from the SM due to a broken symmetry (see, *e.g.*, [334, 335, 336, 336]).
- **Asymmetric Dark Matter:** A number of proposals suggest that the dark-matter density is tied to the baryon asymmetry in the early Universe. A common feature of these so-called asymmetric dark-matter models is the prediction of a relatively light weakly-interacting dark-matter candidate with a mass of a few  $\text{GeV}/c^2$  (see, *e.g.*, [337, 338, 339]).
- **WIMPzillas:** Theories in which dark-matter production is tied to inflation have been proposed that predict the athermal generation of very heavy, weakly interacting massive particles ( $>10^{13} \text{GeV}/c^2$ ) known colloquially as WIMPzillas [340, 341, 342]. Some have proposed decaying WIMPzillas as a source of ultra-high-energy cosmic rays [343].
- **Kaluza-Klein Particles:** Originally proposed as a five-dimensional unified theory of gravitation and electromagnetism in the 1920s [344, 345], Kaluza-Klein (KK) theories postulate the existence of extra dimensions as a mechanism for unifying general relativity and the SM. Although the original version of the theory was largely ignored, within the last decade a number of interesting variants have come into focus (*e.g.*, the

Arkani-Hamad, Dimopoulos and Dvali (ADD) [346, 347] and Randall-Sundrum models [348]). Of particular interest for dark matter is the Universal Extra Dimensions (UED) model [349] in which SM particles are allowed to propagate through one or more compactified extra dimensions, resulting in a hierarchical KK tower of states for each SM particle. The lightest such KK state is a natural dark-matter candidate [350, 351, 352].

- **Q-balls:** In supersymmetric extensions of the SM (see Section 2.5) a type of non-topological soliton is generically predicted that would have been produced in the early Universe and might be stable [353]. A spherical and coherent combination of a potentially large number of squarks, leptons and Higgs fields, Q-balls could comprise part or all of the dark matter if nature is indeed supersymmetric [354, 355].
- **superWIMPs:** Of the particle candidates mentioned thus far, the superweakly interacting massive particles (superWIMPs) in supersymmetric theories (*e.g.*, the gravitino) are perhaps the most disturbing possibility from the direct-detection perspective. If such a particle is the lightest supersymmetric particle it would be a viable dark-matter candidate that interacts so weakly that no foreseeable detector technology will have the requisite sensitivity for a positive detection [356].
- **LIPs:** Although a dark-matter candidate with a standard electrical charge is essentially ruled out [357, 317], there remains the possibility for a dark-matter particle with a fractional charge. There are a number of names in the literature that basically correspond to the same concept: lightly ionizing particles (LIPS), fractionally charged particles, and milli-charged massive particles (milliCHAMPs). Some mirror-matter models contain a viable LIP, as do extensions to the SM in which the SM neutrino possesses a fractional charge. [358, 359].

## 2.4. DARK MATTER CANDIDATES

- **SIMPs:** Although the idea has largely fallen out of favor, it might be possible that a fraction of the dark matter is in the form of strongly interacting massive particles (SIMPs) [360]. The cosmologically favored parameter space is ruled out [361], but SIMPs are interesting even if they are not the dark matter [362].

### 2.4.1 Baryons

As pointed out in Section 2.1, the luminous baryon density in galactic discs accounts for only a small fraction ( $\lesssim 5\%$ ) of the total baryon density in the Universe. At the same time, a typical spiral galaxy’s mass-to-light ratio corresponds to a matter density that is similar in magnitude to the cosmologically preferred value of  $\Omega_b$  (see Sections 2.2.1 and 2.2.2). This begs the question of whether dark baryons are sufficient to explain the galactic dark-matter problem. In this section I explore the two most obvious possibilities for baryonic dark matter in galaxies: hydrogen in clouds of gas or frozen snowballs, and baryons bound into stellar remnants or Jupiter-like objects.

It is worth emphasizing that despite the apparent match between  $\Omega_{\text{galaxy}} \simeq 0.05$  and the similarly sized (global) value of  $\Omega_b$ , measured values of the baryon-mass fraction for clusters indicate that most of the Universe’s (low- $z$ ) baryons are in the form of (hot) ionized intracluster gas, and not as dark baryons in the halos of galaxies. As illustrated in Figure 2.14, the gas-mass fraction in clusters approaches the value of  $f_b$  ( $\sim 0.17$ ) preferred by  $\Lambda$ CDM fits to the CMB power spectrum, an indication that the cosmic baryon budget is primarily spent on intracluster gas. While this does not preclude the possibility that dark baryons might constitute part of the Milky Way’s dark halo, it strongly suggests that a baryon-dominated dark halo is unlikely; the Milky Way is not located in the center of a massive cluster where its halo would be more likely to share mass with a dense region of intracluster gas. However, an obvious question remains: what fraction of the Milky Way’s dark halo can be attributed to dark baryons?



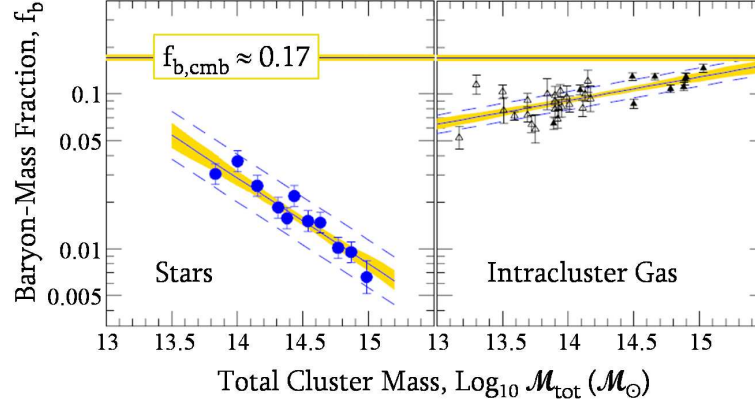


Figure 2.14: The contribution to the baryon-mass fraction from stars (left panel) and intracluster gas (right panel) in galaxy systems, compared to the global value preferred by  $\Lambda$ CDM fits to the CMB power spectrum (horizontal line at  $f_b \simeq 0.17$ ). The stellar fractions are based on virial estimates in [363] (blue, closed circles), while the gas fractions are estimated from Chandra x-ray temperatures in [364] (black, open triangles) and [365] (black, closed triangles). Model fits (slanted solid lines) are shown with  $1\sigma$  uncertainties (shaded regions). For large cluster masses the value of the gas fraction is an indication that most of the Universe’s baryons are located in intracluster-gas clouds, and not in the dark halos of galaxies like the Milky Way. Figure adapted from [363].

## Hydrogen

The least exotic and perhaps most obvious galactic dark-matter candidate is hydrogen. Let’s first consider a dark halo consisting of clumps of frozen hydrogen, sometimes called “snowballs.” Snowballs are electrostatically bound but not gravitationally collapsed, distinguishing them from dwarf stars or Jupiter-like objects (discussed below). This definition places an upper limit on the mass of an individual snowball;  $m_{sb} \lesssim 0.005 \mathcal{M}_\odot \simeq 10^{31} \text{ g}$  [366]. Also known as the Jeans’ mass [367], this limit corresponds to the minimum mass required for an interstellar cloud of hydrogen and helium to undergo gravitational collapse. If the snowball density is that of solid hydrogen,  $\rho_{sb} \simeq 0.07 \text{ g cm}^{-3}$  [88], it would occupy a spherical volume with radius  $r_{sb} \lesssim 3 \times 10^{10} \text{ cm}$ , or roughly half

## 2.4. DARK MATTER CANDIDATES

the radius of the Sun. A cold, neutral object of this size (or smaller) would be incredibly difficult to detect observationally in the halo of the Milky Way. Current exoplanet missions [368, 369] can detect objects this size, but only in orbit around nearby stars where snowballs would quickly evaporate. Snowball-like objects in the Milky Way’s halo with masses between  $\sim 10^{-3}$  and  $10^{21}$  g, however, are essentially ruled out by the frequency with which the Earth encounters interstellar meteors, by the number of craters on the moon, and by the lack of any orbit-crossing (interstellar) comets in the last 400 years [370]. Unfortunately, these limits leave open the possibility of snowballs with  $m_{sb} \lesssim 10^{-3}$  g or  $m_{sb} \gtrsim 10^{21}$  g.

Regardless of observational limits, Hegyi and Olive argue that snowballs could not have survived evaporation by the CMB [371, 372]. Their line of reasoning is quite general and also leads to a minimum snowball size. If snowballs are electrostatically bound, the binding energy per molecule is  $\sim 5$  eV [373]. Discussed in more detail in the next chapter, the currently accepted (three-dimensional) rms velocity of an object in the Milky Way’s dark halo is  $v_h = 270 \pm 25$  km s $^{-1}$  [374], implying an average kinetic energy per snowball molecule of  $\sim 750$  eV. Snowballs would clearly be too loosely bound to survive collisions with one another. A limit on their minimum size and mass can be derived by limiting their rate of collision during the age of the halo,  $t_h$ , otherwise they would have broken apart into hydrogen gas and would no longer be snowballs;

$$\Gamma_c = n_{sb} \sigma_{sb} v_h < \frac{1}{t_h}, \quad (2.4.1)$$

where  $n_{sb} = \rho_h / m_{sb}$  is the snowball number density,  $\rho_h$  is the average density of the Milky Way’s halo, and

$$\sigma_{sb} = \pi r_{sb}^2 = \pi \left( \frac{3m_{sb}}{4\pi\rho_{sb}} \right)^{2/3} \quad (2.4.2)$$

is the cross section of a sphere of solid hydrogen. Rearranging Equation 2.4.1 yields a lower limit on  $m_{sb}$  in terms of halo parameters and  $\rho_{sb}$ ;

$$m_{sb} > \frac{9\pi}{16} (\rho_h v_h t_h)^3 \rho_{sb}^{-2}. \quad (2.4.3)$$

## CHAPTER 2. DARK MATTER

A crude estimate for the halo density can be obtained by uniformly distributing its total mass over a spherical volume. Assuming a total halo mass of  $\sim 10^{12} \mathcal{M}_\odot$  within a 100 kpc radius [375],  $\rho_h \simeq 1.6 \times 10^{-26} \text{ g cm}^{-3}$  (or  $\sim 10 \text{ MeV } c^{-2} \text{ cm}^{-3}$ ).<sup>6</sup> Based on the ratio of  $^{238}\text{U}$  (14 Gyr half-life) to  $^{232}\text{Th}$  (4.5 Gyr half-life) inferred from the spectra of the oldest metal-poor halo stars,  $t_h > 10^{10} \text{ yr} \simeq 3 \times 10^{17} \text{ s}$  [377]. The minimum snowball mass is therefore  $\sim 1 \text{ g}$ , and the minimum radius is  $\sim 1.5 \text{ cm}$ . Consequently, snowballs with masses less than  $10^{-3} \text{ g}$  are unlikely because their number density would be too large for them to have avoided collisions and survived in tact to the present day. An independent argument based on a careful treatment of a snowball's thermal properties limits their mass to be greater than  $\sim 10^{-6} \mathcal{M}_\odot \simeq 2 \times 10^{27} \text{ g}$  [378]. Less substantial snowballs would be evaporated by cosmic rays as they passed through the disc of the Milky Way.

Ultimately, the snowball concept is problematic because they must form in a collisionless manner by condensing out at a time when the density of the Universe was approximately equal to the current halo density. It can be shown that this corresponds to a redshift no smaller than  $z \simeq 3.5$ , when the temperature of the CMB was  $\sim 9.5 \text{ K}$  [372]. Solid hydrogen's high vapor pressure prevents it from existing in equilibrium with the gaseous state at these densities and temperatures; snowballs would immediately begin sublimating after formation. The evaporation time scale is given approximately by

$$\tau_{\text{ev}} \simeq \frac{e^{b_0/kT}}{f_0}, \quad (2.4.4)$$

where  $b_0/k = 91 \text{ K}$  corresponds to the energy required to free a hydrogen molecule from a snowball's surface and  $f_0 \simeq 10^{12} \text{ s}^{-1}$  is the frequency at which a hydrogen lattice vibrates [372]. Snowball formation would therefore compete

---

<sup>6</sup>This value for the halo density is a bit misleading due to the naïve (and incorrect) assumption that the total halo mass is uniformly distributed. As pointed out in Section 2.2.1, the flatness of spiral-galaxy rotation curves suggests a halo density that is inversely proportional to the square of the distance from the center of the galaxy (at large radii). Realistic halo models yield a local dark-matter density of  $0.2\text{--}0.8 \text{ GeV } c^{-2} \text{ cm}^{-3}$  [376]. However, the average density used here,  $\rho_{sb} \simeq 10 \text{ MeV } c^{-2} \text{ cm}^{-3}$ , should be sufficient for setting a conservative lower limit on the mass of a galactic-halo snowball.

## 2.4. DARK MATTER CANDIDATES

with an extremely rapid evaporation rate, characterized by  $\tau_{\text{ev}} \simeq 10^{-8}$  s, due to the 9.5 K thermal bath of CMB photons. It can further be shown that in order for snowballs to populate the Milky Way's present-day halo despite such rapid evaporation they must start out with radii no smaller than  $\sim 10^{16}$  cm [372]. At the density of solid hydrogen, however, this corresponds to a minimum mass of  $\sim 3 \times 10^{47}$  g, well in excess of the Jeans mass and therefore past the limit of being considered electrostatically bound.

Another possibility is that the halo consists of cold hydrogen gas (either molecular or atomic). The collapse time of a spherical halo with average density  $\rho_h$  is

$$\tau_c = \sqrt{\frac{3\pi}{32G_N\rho_h}} \simeq 5 \times 10^8 \text{ yr}, \quad (2.4.5)$$

which is far shorter than the age of the Milky Way. If hydrogen gas populates the halo it must therefore be in hydrostatic equilibrium. Under the simplifying assumption that the gas is distributed in an isothermal sphere ( $\rho_h \propto 1/r^2$ ), its temperature can be estimated from Equation 2.2.33:

$$T = \frac{\mu_g m_p G_N \mathcal{M}_h}{2kr_h} \simeq 1.5 \times 10^6 \text{ K}. \quad (2.4.6)$$

A halo consisting entirely of cold hydrogen gas is therefore ruled out; to remain in equilibrium and not collapse into star forming regions such a large gas cloud would have to exist in an ionized form that is inconsistent with observations (see, *e.g.*, Chapter 23 in [18] for further details).

The above argument does not preclude a population of smaller molecular ( $\text{H}_2$ ) or atomic ( $\text{HI}$ ) hydrogen clouds distributed throughout the halo. Recall that much of the rotation-curve data in Figure 2.3 is derived from 21 cm line observations from  $\text{HI}$  regions. Molecular hydrogen is difficult to detect, but can be inferred from CO emission lines as the ratio of CO to  $\text{H}_2$  in molecular clouds is believed to be a constant. In both cases, although hydrogen gas extends to larger radii than the Milky Way's main stellar population, it appears to be confined (for the most part) to within several kiloparsecs of the galactic plane [379, 380, 381, 382]. While the exact distribution of  $\text{HI}$  and  $\text{H}_2$  in the

Milky Way is still only poorly understood, all indications are that neutral gas contributes  $\lesssim 2\%$  to the total mass of the halo [383] (see, *e.g.*, Figure 2 in [384]).

### Dwarf Stars and Jupiter-like Objects

A popular baryonic dark-matter solution is that the halo is populated with ultra-low-luminosity (or entirely dark), gravitationally collapsed massive objects, commonly referred to as MAssive Compact Halo Objects (MACHOs) [385]. MACHO candidates include black holes, stellar remnants that are too small and dim to be directly observed (except locally), and Jupiter-like objects with little to no intrinsic source of luminosity. Our own solar system harbors several excellent examples of the latter in the form of the gas giants Jupiter, Saturn, Neptune and Uranus. Planets bound to stars, however, are not viable dark-matter candidates as their masses are but small perturbations to their parent stars. Furthermore, their spatial distribution is coincident with the Milky Way's stellar population and therefore inconsistent with an extended dark halo. In order to reproduce the Galaxy's flat rotation curve, a large number of MACHOs beyond the extent of the LD and LB is necessary.

The spectrum of possible MACHO masses is surprisingly large. Although the Jeans' mass is a suitable upper limit for electrostatically bound objects, it does not necessarily translate into a MACHO-mass lower limit. Uranus is a mere  $5 \times 10^{-5} \mathcal{M}_{\odot}$ , well below the  $\sim 10^{-3} \mathcal{M}_{\odot}$  Jeans' limit. Clearly there is a mechanism by which smaller-mass gravitationally bound objects can form. An obvious possibility is the fragmentation of planet-like objects from the circumstellar discs of newly formed stars. However, in order for the requisite number of MACHOs to have been produced in this manner, thousands of Jupiter-like objects would have to be ejected during the formation of a typical solar system. Nevertheless, a suitable mechanism might exist that evades the Jeans' limit. A much weaker limit can be derived by considering the rate at which hydrogen escapes the surface of a gravitationally collapsed object. In order to survive total hydrogen evaporation during the age of the Galaxy, a MACHO must have a mass in excess of  $10^{-7} \mathcal{M}_{\odot}$  [386].

## 2.4. DARK MATTER CANDIDATES

The upper limit on mass depends on the type of MACHO. Heavy Jupiter-like objects (often referred to as brown dwarfs [387, 388]) are limited to no more than  $0.08 \mathcal{M}_\odot$  [18], otherwise hydrogen fusion would take place at their cores and they would be visible as conventional stars. If a brown dwarf’s mass is close enough to this limit, it might have a low intrinsic luminosity due to deuterium fusion, but will generally be too dim to observe directly. Stellar remnants such as white dwarfs, black holes and neutron stars can have masses in excess of  $\mathcal{M}_\odot$ . However, most stellar remnants are excluded from contributing significantly to the Milky Way’s halo. For example, the formation of too many white dwarfs would contaminate the local environment—through ejection of their outer layers—with overabundances of heavy elements. In particular, the abundances of carbon and nitrogen inferred from observations of Lyman- $\alpha$  absorption lines imply  $\Omega_{\text{wd}} \lesssim 0.0003$  [389]. While it is possible that heavier elements like carbon are not efficiently ejected during the violent (and some might argue poorly understood) formation of stellar remnants [390], a looser constraint of  $\Omega_{\text{wd}} \leq 0.003$  can be derived based on the well understood abundances of D and  $^4\text{He}$  [391]. Neutron stars and stellar-mass black holes are similarly restricted. Their larger masses make the problem of heavy-element contamination even worse. Hydrodynamic calculations of the collapse of very massive stars, however, indicate that black holes with masses in excess of  $\sim 100 \mathcal{M}_\odot$  might be viable MACHO candidates [392, 393, 394].

There is also a window of masses between  $\sim 0.08 \mathcal{M}_\odot$  and  $0.2 \mathcal{M}_\odot$  belonging to a class of very dim hydrogen burning MACHO candidates known as red dwarfs. Early attempts to directly resolve red dwarfs in the Milky Way’s halo detected only a handful of candidates [395, 396]. A careful analysis of these candidates by Graff and Freese concluded that red dwarfs contribute no more than 1% to the mass of the halo [397]. More recent studies have essentially confirmed their findings [398, 399].

Taking the above arguments into consideration, if the Milky Way’s halo consists entirely of MACHOs, the most likely allowed range of MACHO masses is  $10^{-7} - 10^{-1} \mathcal{M}_\odot$  and  $\gtrsim 100 \mathcal{M}_\odot$ . In summary, the possible candidates and their approximate mass ranges are:

- **Brown Dwarfs** and smaller planet-sized objects with  $\mathcal{M}_{\text{bd}} \in 10^{-7} - 10^{-1} \mathcal{M}_{\odot}$  could comprise a significant fraction of the halo.
- **Red Dwarfs** with  $\mathcal{M}_{\text{rd}} \in 0.08 - 0.2 \mathcal{M}_{\odot}$  constitute no more than 1% of the halo's mass.
- **Stellar Remnants** with masses  $\lesssim 100 \mathcal{M}_{\odot}$  constitute no more than  $\sim 10\%$  of the halo's mass. These include white dwarfs ( $\sim 0.6 \mathcal{M}_{\odot}$ ), neutron stars ( $\sim 1.5 \mathcal{M}_{\odot}$ ) and stellar-mass black holes ( $\gtrsim 1.5 \mathcal{M}_{\odot}$ ).
- **Very Massive Black Holes** with masses  $\gtrsim 100 \mathcal{M}_{\odot}$  could constitute a significant fraction of the halo's mass.

A technique for observing MACHOs in the Milky Way's halo was proposed in a paper by Paczyński in 1986 [400]. Despite their lack of luminosity, MACHOs in the halo should be detectable when their gravitational fields deflect light from background stars. Unlike clusters and galaxies, an individual MACHO's mass is far too small for strong or even weak gravitational-lensing distortions to be visible. Instead, the detection technique relies on an effect known as microlensing. As a MACHO passes between us and a source star, its gravitational potential causes a temporary increase in the apparent magnitude of the star by focusing a greater solid angle of the star's emitted light onto our telescopes. The amplification factor is given by

$$\mathcal{A} = \frac{1 + 2(r_{\text{E}}/r_d)^2}{\sqrt{1 + 4(r_{\text{E}}/r_d)^2}}, \quad (2.4.7)$$

where  $r_d$  is the angular distance (as viewed from Earth) between the lens and the source star (or the impact parameter),

$$r_{\text{E}} \equiv \sqrt{\frac{4G_N \mathcal{M}_{\text{macho}} d_s x(1-x)}{c^2}} \quad (2.4.8)$$

is a characteristic scale referred to as the Einstein radius,  $d_s$  is the distance to the source star, and  $x$  is a fraction between 0 and 1 such that  $x d_s$  is the

## 2.4. DARK MATTER CANDIDATES

distance to the MACHO [400]. The convention is to define the width of a microlensing event as the time during which  $r_d < r_E$ , equivalent to observing an amplification factor in excess of 1.34 (set  $r_d$  equal to  $r_E$  in Equation 2.4.7). Microlensing results are usually expressed in terms of the microlensing optical depth

$$\tau_{\text{opt}} = \frac{4\pi G_N d_s^2}{c^2} \int_0^1 \rho(x) x(1-x) dx, \quad (2.4.9)$$

where  $\rho(x)$  is the density of MACHOs in the halo [401].  $\tau_{\text{opt}}$  is the probability that a given source star will be amplified by more than a factor of 1.34 at a given time—the instantaneous microlensing probability. Since the distribution of dark matter in the Milky Way’s halo is generally unknown, it is common practice to assume a specific halo model for the purpose of inter-comparing microlensing results. The MACHO community typically models  $\rho(x)$  according to a softened isothermal sphere (see Equation 2.2.6) with core radius  $a_r = 5 \text{ kpc}$  [385, 402]. If MACHOs contribute a fraction  $f$  (by mass) of the dark-matter in the Milky Way’s halo, according to this model the expected microlensing optical depth for source stars in the Large (Small) Magellanic Cloud is  $\tau_{\text{opt}} \simeq 4.7f \times 10^{-7}$  ( $6.6f \times 10^{-7}$ ) [401].<sup>7</sup> For a softened isothermal sphere of MACHOs with a  $270 \text{ km s}^{-1}$  rms velocity, the average duration of an ensemble of microlensing events is related to the lensing mass:

$$\langle t_E \rangle \simeq 70 \left( \frac{\mathcal{M}_{\text{macho}}}{\mathcal{M}_{\odot}} \right)^{1/2} \text{ days}, \quad (2.4.10)$$

where  $t_E$  is one half the time during which  $\mathcal{A} > 1.34$ —the time it takes the MACHO to travel an angular distance across the sky equal to the Einstein radius [401]. Searching for MACHOs boils down to monitoring individual stars in a nearby galaxy for temporary brightenings in their apparent magnitudes. Short duration events of less than a few weeks are indicative of brown dwarfs

---

<sup>7</sup>The Large and Small Magellanic Clouds (commonly abbreviated as LMC and SMC, respectively) are two of Milky Way’s closest satellite galaxies, a mere 50–60 kpc from the Sun. Microlensing source stars are commonly chosen from these nearby locations because individual stars are easily resolved and bright, but can still be treated as point-like background sources.



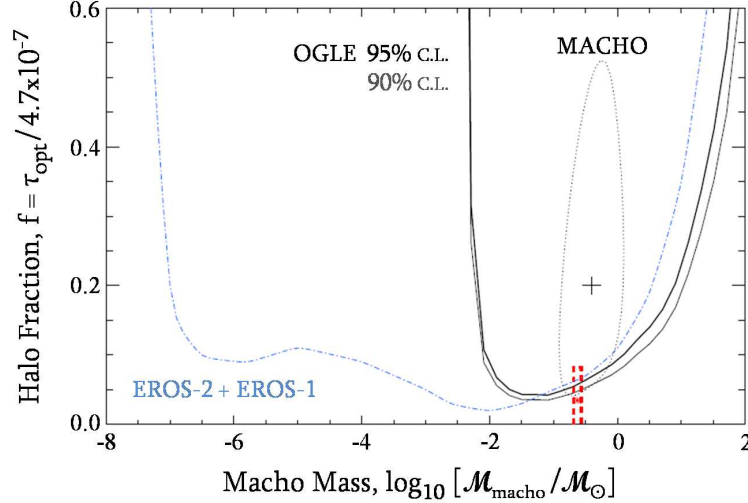


Figure 2.15: Gravitational-microlensing limits from the EROS [401], MACHO [402], and OGLE [403] collaborations on the fraction ( $f$ ) of the Milky Way’s halo (by mass) due to MACHOs as a function of MACHO mass. The lack of microlensing events toward the LMC in the EROS data is consistent with a halo consisting of no more than  $\sim 10\%$  MACHOs (blue dash-dotted line, 95% C.L.) over a wide range of masses, while the 13–17 LMC events observed by the MACHO collaboration prefer  $f \simeq 20\%$  (+ surrounded by 95% C.L. dotted contour). The 2 LMC events in the OGLE data have been interpreted both in terms of MACHOs in the halo of the Milky Way (red dashed lines) and as background due to LMC self-lensing (light and dark solid lines, 95% and 90% C.L., respectively). Figure taken from [403].

with  $\mathcal{M}_{\text{macho}} \lesssim 0.1 \mathcal{M}_{\odot}$ , while longer-duration events are more likely to be due to red dwarfs or stellar remnants with  $\mathcal{M}_{\text{macho}} \gtrsim 0.1 \mathcal{M}_{\odot}$ .

Several collaborations have monitored stars in the Magellanic Clouds for years at a time with varying degrees of sensitivity in order to search for the characteristic microlensing signatures of MACHOs in the halo of the Milky Way. Their results are varied and in some cases at odds with one another. However, it is almost universally true that microlensing results are inconsistent with a dark halo consisting entirely of MACHOs. The following list summarizes the main results.

## 2.4. DARK MATTER CANDIDATES

- The **MACHO** collaboration monitored  $\sim 12$  million stars in the LMC for nearly 6 years and discovered 13–17 microlensing candidates with an expected background of 2–4, depending on selection criteria. No long-duration candidates were discovered, ruling out a Milky Way halo composed entirely of MACHOs with masses of  $\sim 1\text{--}30 \mathcal{M}_\odot$ , but setting only loose constraints on the possible fraction [404]. The shorter duration candidates are consistent with an LMC optical depth of  $\tau_{\text{opt}} \simeq 1.2 \times 10^{-7}$  and characteristic time scales of  $2 < t_E < 400$  days, implying a halo consisting of  $\sim 20\%$  MACHOs (by mass) with  $\mathcal{M}_{\text{macho}} \simeq 0.4 \mathcal{M}_\odot$ . Their best-fit halo fraction,  $f$ , is shown as a function of  $\mathcal{M}_{\text{macho}}$  in Figure 2.15. Despite being model dependent, their results exclude a 100% MACHO halo at the 95% confidence level [402].
- In the first phase of the *Expérience de Recherche d'Objets Sombres* project (**EROS-1**) a search for both long- and short-duration microlensing events was conducted over the course of 4 years. Two events were observed with lensing masses consistent with  $\mathcal{M}_{\text{macho}} \simeq 0.1 \mathcal{M}_\odot$ , and a 95% confidence-level upper limit of 20% (of total halo mass) was placed on MACHOs with  $\mathcal{M}_{\text{macho}} \in 10^{-7}\text{--}0.02 \mathcal{M}_\odot$  [405]. The second phase (EROS-2) monitored  $\sim 33$  million stars in both Magellanic clouds for nearly 7 years. Based on data from a subsample of their 7 million brightest source stars they identified 9 new candidates, for a total of 11 EROS-1 and EROS-2 events. However, followup observations indicate that both EROS-1 and 8 of the EROS-2 candidates are consistent with being due to variable stars or supernovae, reducing the final selection to a single event. From this single microlensing candidate the optical depth toward the LMC is limited to  $\tau_{\text{opt}} < 0.36 \times 10^{-7}$  (95% confidence level) and the halo contribution of MACHOs with  $\mathcal{M}_{\text{macho}} \in 10^{-7}\text{--}1 \mathcal{M}_\odot$  is limited to no more than  $\sim 10\%$  [401]. The combined EROS limits are compared to the best-fit region from the MACHO collaboration's 13–17 events in Figure 2.15. The EROS data limit MACHOs with  $\mathcal{M}_{\text{macho}} \simeq 0.4 \mathcal{M}_\odot$  to contributing no more than 8% to the Milky Way's halo.

- The Optical Gravitational Lensing Experiment (**OGLE**) monitored both Magellanic clouds for an eight year period. Based on two candidate events, they measure an LMC optical depth of  $\tau_{\text{opt}} = (0.16 \pm 0.12) \times 10^{-7}$ , which can be interpreted in terms of a halo composed of 1–5% MACHOs with  $\mathcal{M}_{\text{macho}} \simeq 0.2 \mathcal{M}_{\odot}$  [403]. However, the OGLE collaboration argues that these events are consistent with an optical depth background of  $\tau_{\text{opt}} = 0.1 \times 10^{-7}$ , expected from self-lensing by objects within the LMC [406]. Instead, they set an upper limit on the halo mass fraction for MACHOs with  $\mathcal{M}_{\text{macho}} \simeq 0.01\text{--}10 \mathcal{M}_{\odot}$  that agrees with the EROS results and is shown in Figure 2.15. MACHOs with  $\mathcal{M}_{\text{macho}} \simeq 0.4 \mathcal{M}_{\odot}$  are limited to contributing no more than 7% of the halo’s mass at 95% confidence. In a separate paper, the OGLE collaboration describes their measurement of 3 SMC microlensing events, which results in an SMC optical depth of  $\tau_{\text{opt}} = (1.30 \pm 1.01) \times 10^{-7}$  that they claim is consistent with the expected background due to SMC self-lensing [407]. The limits from their SMC data do not significantly improve upon the LMC limits shown in Figure 2.15.

The origin of the MACHO collaboration’s 13–17 candidates has been the focus of much debate within the community. Although one of their candidates was identified as a variable star in the OGLE data [403], reanalysis of the their data and followup observations have essentially confirmed their original findings [408, 409]. The analysis by Bennett in [408] reduces the number of candidates to 10 for an LMC optical depth of  $\tau_{\text{opt}} = (1.0 \pm 0.3) \times 10^{-7}$ , adjusting the halo fraction preferred by the MACHO data down to  $\sim 16\%$  for MACHOs with  $\mathcal{M}_{\text{macho}} \simeq 0.4 \mathcal{M}_{\odot}$ . The LMC statistics observed by all three microlensing collaborations are low enough that agreement between the experiments is possible but unlikely. It can be said with confidence, however, that brown dwarfs and planet-like objects with  $\mathcal{M}_{\text{macho}} \in 10^{-7}\text{--}10^{-1} \mathcal{M}_{\odot}$  contribute no more than  $\sim 10\%$  to the mass of the Milky Way’s halo. Furthermore, considered together, the microlensing data appear roughly consistent with the above mentioned limits concerning red dwarfs ( $< 1\%$ ) and stellar remnants with masses less than  $\sim 100 \mathcal{M}_{\odot}$  ( $< 10\%$ ). Limits on very massive MACHOs with  $\mathcal{M}_{\text{macho}} \gtrsim 30 \mathcal{M}_{\odot}$

## 2.4. DARK MATTER CANDIDATES

are lacking, leaving open the possibility that the halo consists of very massive black holes. Nevertheless, given the case for nonbaryonic dark matter presented in Chapter 1 as well as the lack of evidence for a significant baryon density in the Milky Way’s halo, more exotic solutions to the dark-matter problem must be explored.

### 2.4.2 Neutrinos

The only known nonbaryonic dark-matter candidates are SM neutrinos. As mentioned in Section 1.1, the flavor eigenstates of SM neutrinos have been experimentally proven to oscillate from one to another, implying nonzero off-diagonal mass-matrix terms that correspond to nonzero neutrino masses [9]. It has long been argued that SM neutrinos with masses  $m_\nu \simeq 30 \text{ eV}/c^2$  could account for a significant fraction of the Universe’s dark matter [410]. This is easy to understand in light of an expected (present-day) relic density of  $\gtrsim 100$  neutrinos per cubic centimeter for each neutrino species with  $m_\nu \lesssim 1 \text{ MeV}/c^2$  [411]. In this section I briefly review some arguments against SM neutrinos as the dominant form of dark matter.

Since light and heavy neutrinos have very different properties in the early Universe, it is important to establish an approximate neutrino-mass scale. The least model dependent and accurate estimates of neutrino mass come from measuring the endpoint and shape of the tritium  $\beta$ -decay spectrum. Tritium is an unstable isotope of hydrogen with a nucleus consisting of two neutrons in addition to the usual proton. It decays with an  $\sim 12 \text{ yr}$  half-life to a  $^3\text{He}$  nucleus by converting one of its neutrons to a proton via  $\beta$  decay:

$$^3_1\text{T} \longrightarrow ^3_2\text{He}^+ + e^- + \bar{\nu}_e. \quad (2.4.11)$$

Due to the relative rest masses of the decay products, the  $^3\text{He}$  daughter is produced at rest (relative to the tritium parent), while the  $18.6 \text{ keV}$  decay energy is split between the electron and antineutrino [412]. The  $\bar{\nu}_e$  is so weakly interacting that it will escape with its share of the decay energy undetected (from most detectors). With a suitable spectrometer and large enough tritium source,

however, the distribution of electron decay energies can be measured. Both the endpoint and the shape of this tritium  $\beta$ -decay spectrum are sensitive to the square of the mass of the antineutrino,  $\Delta m_{\bar{\nu}_e}^2$ . To date, a (convincing) finite neutrino mass has yet to be measured in this manner, with current upper limits suggesting  $m_{\bar{\nu}_e} < 2.3 \text{ eV } c^{-2}$  at 95% confidence [413, 414]. The KARlsruhe TRItium Neutrino (KATRIN) experiment will take data for five years beginning in 2012 and promises an order of magnitude improvement over current limits, with sensitivity to  $m_{\bar{\nu}_e} \simeq 0.3 \text{ eV } c^{-2}$  [415].

Direct measurements of the masses of the muon and tau neutrinos yield much weaker limits, with  $m_{\nu_\mu} \lesssim 0.2 \text{ MeV}/c^2$  and  $m_{\nu_\tau} \lesssim 20 \text{ MeV}/c^2$  [88]. The large-volume neutrino observatories with which neutrino oscillations are characterized, however, are sensitive to neutrino-mass differences. If we assume for the sake of argument that  $2.3 \text{ eV}/c^2$  represents the mass of the lightest neutrino, upper limits on the masses of the two heavier species can be derived from the squared neutrino-mass differences measured by these experiments.

The main purpose of neutrino-oscillation experiments is to characterize the  $3 \times 3$  Pontecorvo-Maki-Nakagawa-Sakata (PMNS) mixing matrix [416, 417] that relates the neutrino flavor eigenstates to the neutrino mass eigenstates ( $\nu_1$ ,  $\nu_2$  and  $\nu_3$ ). The PMNS matrix is commonly parameterized by a CP-violating phase,  $\delta_{\text{CP}}$ , and the three mixing angles  $\theta_{12}$ ,  $\theta_{23}$  and  $\theta_{13}$ , each tied to the probability for neutrinos to oscillate from one flavor to another. Traditionally,  $\theta_{12}$  has been associated with an apparent deficit in electron-neutrino flux from the Sun. This deficit was first measured by Davis and collaborators with a chlorine-based detector at the Homestake Mine in South Dakota, where they detected  $\sim 1/3$  of the electron neutrinos expected to be produced by reactions in the Sun, giving rise to the so-called solar-neutrino problem [418]. Many years and many  $\nu_e$ -disappearance experiments later, the solar neutrino problem is now understood to be the result of electron neutrinos oscillating into another flavor while in transit from the Sun (or an accelerator target). The oscillation probability is governed by the large mixing angle  $\theta_{12} \simeq 34^\circ$ , and corresponds to a squared mass difference between the first two neutrino mass

## 2.4. DARK MATTER CANDIDATES

eigenstates of  $\Delta m_{12}^2 = (7.6 \pm 0.2) \times 10^{-5} \text{ eV}^2/c^4$  [419]. The mixing angle  $\theta_{23}$  is associated with a deficit of muon neutrinos (relative to electron neutrinos) expected from the byproducts of cosmic-ray interactions in the atmosphere. The Super-Kamiokande  $\nu_\mu$ -disappearance experiment in Japan confirmed the existence of neutrino-flavor oscillations by providing the first precision measurement of  $\theta_{23}$  [9]. The size of the corresponding squared mass difference is much larger than  $\Delta m_{12}^2$ , implying that the atmospheric-neutrino problem is due to muon neutrinos oscillating into a different flavor than solar electron neutrinos. As is the case for solar neutrinos, the mixing angle is large;  $\sin^2(2\theta_{23}) > 0.92$  [420]. The most accurate measure of the squared mass difference,  $\Delta m_{23}^2 = (2.43 \pm 0.13) \times 10^{-3} \text{ eV}^2/c^4$ , comes from the MINOS  $\nu_\mu$ -disappearance experiment at the Soudan Mine in northern Minnesota [421]. The value of  $\delta_{\text{CP}}$  is currently unknown, while a nonzero value of  $\theta_{13}$  ( $\sim 1\text{--}10^\circ$  at 90% confidence) appears to have been measured for the first time by the T2K  $\nu_e$ -appearance experiment [422].

From the neutrino-oscillation results it is clear that SM neutrinos are light. If we stretch the experimental errors to the  $2\sigma$  level and assume the lightest neutrino has a mass of  $2.3 \text{ eV}/c^2$ , the mass of the heaviest neutrino cannot exceed  $\sim 2.36 \text{ eV}/c^2$ . A crude limit on the sum of the neutrino masses is therefore

$$\sum m_\nu \lesssim 7 \text{ eV}/c^2, \quad (2.4.12)$$

where the sum is over neutrino flavors. To translate this into a limit on the present-day neutrino energy density,  $\Omega_\nu$ , an estimate of the neutrino number density is needed. During the early stages of BBN, neutrinos remain in equilibrium with other forms of radiation (*e.g.*, photons, electrons and positrons) provided the rate of their weak-scale interactions is greater than the expansion rate of the Universe:  $\Gamma_{\text{weak}} > H$ . It can be shown that this is approximately true for

$$(kT)^3 > \sqrt{\frac{8\pi^3 g^*}{90}} \frac{(m_W c^2)^4}{\mathcal{M}_P c^2}, \quad (2.4.13)$$

where  $\mathcal{M}_P = \sqrt{\hbar c/G_N} = 1.22 \times 10^{19} \text{ GeV}/c^2$  is the Planck mass,  $m_W \simeq$

$80 \text{ GeV}/c^2$  is the mass of the  $W$  boson, and

$$g^* = \sum g_{\text{boson}}^* + \frac{7}{8} \sum g_{\text{fermion}}^*, \quad (2.4.14)$$

is the number of relativistic degrees of freedom [151, 88]. Photons contribute 2 degrees of freedom to the bosonic sum, while electrons, positrons and the 3 neutrino species contribute a total of 10 degrees of freedom to the fermionic sum, resulting in  $g^* = 43/4$ . Neutrinos will therefore remain in equilibrium until the temperature of the Universe cools to below  $kT \simeq 1 \text{ MeV}$ . This is an important dividing line regarding the relic density of light versus heavy neutrinos. The number density of heavy neutrinos with masses  $\gtrsim 1 \text{ MeV}/c^2$  will be exponentially suppressed due to self-annihilations [423]. Light neutrinos, however, will decouple as radiation and will have a relic density nearly equal to the density of CMB photons;

$$n_\nu = \frac{\rho_\nu}{\sum m_\nu} = \frac{3}{11} n_\gamma, \quad (2.4.15)$$

where the suppression factor of  $3/11$  accounts for the reheating of the photon background during the BBN  $e^\pm$  annihilation stage that occurs after neutrino decoupling [151]. The photon number density can be estimated by integrating a black-body spectrum characterized by temperature  $T_{\text{cmb}}$  (see Chapter 6 in [18] for further details);

$$n_\gamma = \frac{2}{\pi^2} \left( \frac{kT_{\text{cmb}}}{\hbar c} \right)^3 \int_0^\infty \frac{x^2 dx}{e^x - 1} \simeq 410.5 \text{ cm}^{-3}. \quad (2.4.16)$$

An expression for the present-day energy density of light neutrinos follows from a combination of Equations 2.4.15 and 2.4.16 and the critical density in Equation 2.1.2:

$$\Omega_\nu = \frac{\rho_\nu}{\rho_c} \simeq \frac{\sum m_\nu}{94 h^2 \text{ eV } c^{-2}}. \quad (2.4.17)$$

The neutrino-mass limit from Equation 2.4.12 therefore implies  $\Omega_\nu \lesssim 0.14$  for  $h = 0.738$ .

The neutrino contribution to the energy density can be further limited by considering the effect light neutrinos have on density fluctuations in the early

## 2.4. DARK MATTER CANDIDATES

Universe, and on the subsequent clustering of galaxies at later times. Following inflation, curvature perturbations begin to grow into the regions of over- and underdensity that will later give rise to the CMB anisotropy and structures we observe today. During these radiation-dominated times, light neutrinos have large velocities due to their small masses (relative to the high temperatures), and are therefore capable of free-streaming over large distances. If  $\Omega_{\mathcal{M}}$  is dominated by light neutrinos, their free-streaming capability will tend to inhibit (or “wash out”) density contrasts (both in neutrinos and baryons) on small scales [424, 425]. This can have a pronounced effect on the shape of the CMB power spectrum and the level of small-scale galaxy clustering at late times [426, 427]. Furthermore, if neutrinos with masses less than  $\sim 30 \text{ eV}/c^2$  dominate the matter density, N-body simulations of the nonlinear growth of structure indicate the formation of too much large-scale structure [428]. Such neutrino-driven structure formation occurs in a top-down fashion, where large structures form first and smaller structures like galaxies fragment off at later times ( $z \lesssim 1$ ) [424, 429]. The problem with this top-down scenario is that galaxies have been observed out to redshifts as large as  $z \simeq 8.6$  [430]. In general, to correctly simulate the observed spatial-redshift distribution of galaxies requires  $\Omega_{\mathcal{M}}$  to be dominated by a nonrelativistic particle species in the early Universe [431], which is perhaps the main reason for the popularity of the cold dark-matter scenario. The astrophysical evidence discussed in the previous chapter (CMB, SNe, BAO and  $H_0$ —see Section 1.3.5) suggests that the sum of the SM neutrino masses is less than  $\sim 1 \text{ eV}/c^2$  [432, 24]. The authors in [433] further restrict  $\sum m_\nu$  by studying the clustering of  $> 700,000$  LRGs in the seventh SDSS data release [133] as a function of angular scale and redshift. At small angular scales their power spectrum exhibits too much galaxy clustering to be consistent with  $\sum m_\nu > 1 \text{ eV}/c^2$ , and when combined with other cosmological constraints their analysis indicates

$$\sum m_\nu < 0.281 \text{ eV}/c^2. \quad (2.4.18)$$

The corresponding neutrino-density limit from Equation 2.4.17 is  $\Omega_\nu < 0.005$ , a nearly trivial fraction ( $\sim 2\%$ ) of the total dark-matter density. The three SM neutrino species are clearly an inadequate solution to the dark-matter problem.



### 2.4.3 Weakly Interacting Massive Particles

It should be clear by this point that dark baryons are insufficient to explain the dark-matter problem, whether on a cosmic or galactic scale. Furthermore, a dark-matter density dominated by particles that were relativistic during the early stages of structure formation—hot dark matter—is disfavored because the corresponding top-down structure-formation scenario is at odds with the spatial-redshift distribution of small-scale structure in the Universe. All the evidence appears to lead to a dark-matter candidate that lies outside the SM, is weakly (or very weakly) interacting, and is nonrelativistic during the early stages of structure formation. These conditions are satisfied by a large (and popular) class of hypothetical candidates known as Weakly Interacting Massive Particles (WIMPs) [137]. WIMPs are stable (or very long lived) particles that interact weakly (and gravitationally) and are individually massive enough to have frozen-out at a relatively high temperature in the early Universe, allowing them to cool to nonrelativistic velocities and seed structure formation in a manner consistent with observational evidence. There are a number of theories beyond the SM that contain a natural WIMP candidate. Before discussing any specific examples, however, it is instructive to review the provocative (and general) weak-scale relationship between the annihilation cross section and relic density of cosmologically interesting WIMPs ( $\Omega_{\text{wimp}} \simeq 1$ ), what many consider to be their defining property and the principle reason they are a well motivated class of candidates. The following discussion loosely follows the line of reasoning presented in Chapter 18 of [18], but also mimics the standard “WIMP coincidence” argument found elsewhere in the literature [2, 138, 434].

Suppose the dark matter is composed of a WIMP, call it  $\chi$  with mass  $m_\chi$ , that is in thermal equilibrium in the early Universe; that is, provided  $kT \gg m_\chi c^2$ , the number of WIMPs is maintained at a level roughly equal to the photon background by rapid conversion to lighter species through annihilations ( $\chi + \bar{\chi} \rightarrow l + \bar{l}$ , where  $l$  could be an electron, quark or other particle in thermal equilibrium) and *vice versa* ( $l + \bar{l} \rightarrow \chi + \bar{\chi}$ ). As the Universe cools, the number of WIMPs drops exponentially because WIMP production

## 2.4. DARK MATTER CANDIDATES

becomes kinematically disfavored. Eventually, the rate of annihilations slows to an effective stop as the WIMPs cool to nonrelativistic speeds and the space-time expansion dilutes their numbers. This occurs when the expansion and WIMP-annihilation rates are roughly equal, and is commonly referred to as freeze-out because the absolute number of WIMPs ceases to change appreciably:

$$\Gamma_A = \langle \sigma_A v \rangle_f n_{\chi,f} \simeq H(t_f), \quad (2.4.19)$$

where  $n_\chi$  is the WIMP number density,  $\sigma_A$  is the annihilation cross section,  $v$  is their relative velocity, the brackets represent thermal averaging, and the subscript  $f$  denotes values at the freeze-out time  $t_f$ .

The balance between WIMP annihilation, production, and dilution is governed by the Boltzmann equation:

$$\frac{dn_\chi}{dt} + 3Hn_\chi = -\langle \sigma_A v \rangle [n_\chi^2 - n_{\chi,\text{eq}}^2], \quad (2.4.20)$$

When the rate of annihilations is sufficiently low (or effectively zero), the evolution of the number density is controlled by the second term on the left-hand side, and  $n_\chi$  dilutes according to the expansion of the Universe;  $n_\chi \propto a^{-3}$ . The terms on the right-hand side represent particle annihilations (first term) and the reverse process of WIMP production (second term), and are approximately equal when WIMPs are in thermal equilibrium. Assuming the WIMPs follow a Maxwell-Boltzmann distribution and are nonrelativistic (such that an individual particle's energy is well approximated by  $\mathcal{E} = m_\chi c^2 + p^2/2m_\chi$ ), an expression for the number density at thermal equilibrium is

$$\begin{aligned} n_{\chi,\text{eq}} &= g_\chi^* \int \frac{d^3p}{(2\pi\hbar)^3} e^{-\mathcal{E}/kT} \\ &= \frac{g_\chi^*}{(2\pi\hbar)^3} \int d^3p e^{-p^2/2m_\chi kT} e^{-m_\chi c^2/kT} \\ &= 4 \left( \frac{m_\chi kT}{2\pi\hbar^2} \right)^{3/2} e^{-m_\chi c^2/kT}, \end{aligned} \quad (2.4.21)$$

where the WIMPs have been assumed to be Dirac particles such that the particle degrees of freedom  $g_\chi^* = 4$ . The arguments presented here work equally

## CHAPTER 2. DARK MATTER

well, however, for Majorana particles in which  $\chi = \bar{\chi}$  and  $g_\chi^* = 2$  (see [138] for further details). The number density at freeze-out is given by Equation 2.4.21 evaluated at the freeze-out temperature  $T_f$ . In fact, Equation 2.4.21 is quite general; it gives the number density (at  $T_f$ ) of any heavy particle species that was once in thermal equilibrium. Following freeze-out, WIMPs are a form of nonrelativistic matter and their density will scale as  $\rho_\chi \propto a^{-3}$  (see Equation 1.2.8). Their present-day density is therefore related to  $n_{\chi,f}$  according to

$$\Omega_\chi = \frac{\rho_{\chi,0}}{\rho_c} = \frac{n_{\chi,f} m_\chi}{(1+z_f)^3 \rho_c}, \quad (2.4.22)$$

where  $z_f$  is the freeze-out redshift.  $\Omega_\chi$  is commonly referred to as the relic density because it is simply a scaled version of the density at freeze-out and therefore represents the number of WIMPs left over following thermal production in the early Universe.

An expression for  $z_f$  can be derived by invoking conservation of entropy, thereby relating quantities at freeze-out to present-day quantities (at  $z = 0$ ). Now and in the early Universe, the entropy density  $S$  is dominated by the relativistic degrees of freedom. Consider the entropy associated with a volume of black-body radiation at temperature  $T$ . Adding an increment of heat  $\Delta U$ , while holding the volume and temperature fixed, increases its entropy by  $\Delta S = \Delta U/T$ . It is straightforward to integrate the black-body distribution of photon frequencies to obtain an expression for  $U$  in terms of  $T$ ;

$$\begin{aligned} U &= \int_0^\infty U(\omega) d\omega \int_0^\infty \frac{\hbar}{\pi^2 c^3} \frac{\omega^3 d\omega}{e^{\hbar\omega/kT} - 1} \\ &= \frac{\pi^2 (kT)^4}{15 (\hbar c)^3} = \frac{4\sigma_B}{c} T^4, \end{aligned} \quad (2.4.23)$$

where  $\sigma_B \equiv (\pi k^2)^2 / (60 \hbar^3 c^2)$  is the Stefan-Boltzmann constant and  $U = 4\sigma_B T^4 / c$  is often referred to as the Stefan-Boltzmann law. Utilizing the Stefan-Boltzmann law, the entropy density can be written as

$$S = \int \frac{dU}{T} = \frac{16\sigma_B}{c} \int T^2 dT = \frac{16}{3c} \sigma_B T^3, \quad (2.4.24)$$

## 2.4. DARK MATTER CANDIDATES

which generalizes to

$$S_f = \frac{16}{3c} \sigma_B T_f^3 \frac{g_f^*}{2} \quad (2.4.25)$$

for an arbitrary number of relativistic degrees of freedom  $g_f^*$  at temperature  $T_f$ . The present-day entropy density is dominated by the three neutrino species (two spin degrees of freedom each) and the two components of the electromagnetic field. Recalling from Equation 2.4.14 that each fermionic degree of freedom contributes  $7/8$  as much as each bosonic degree of freedom, the present-day entropy density can therefore be split into a term due to the CMB and a term due to the neutrino background;

$$S_0 = \frac{16}{3c} \sigma_B \left( 2T_{\text{cmb}}^3 + 6T_\nu^3 \frac{7}{8} \right) / 2, \quad (2.4.26)$$

It can be shown that the CMB and neutrino temperatures are related according to (see Chapter 6 in [18] for derivation)

$$T_\nu = \left( \frac{4}{11} \right)^{1/3} T_{\text{cmb}}. \quad (2.4.27)$$

Finally,  $S_f$  can be related to  $S_0$  by noting that the entropy density per comoving volume is a conserved quantity;  $a^3 S$  is a constant. When expressed in terms of redshift (see Equation 1.2.6), this conservation condition requires

$$\frac{S_f}{(1+z_f)^3} = \frac{S_0}{(1+z_0)^3} = S_0. \quad (2.4.28)$$

Putting the previous four numbered equations together yields

$$(1+z_f) = \left( \frac{11g_f^*}{43} \right)^{1/3} \frac{T_f}{T_{\text{cmb}}}. \quad (2.4.29)$$

The final ingredient needed to relate the WIMP-annihilation cross section and relic density is an expression for the Hubble parameter at freeze-out. Since the Universe is radiation dominated when WIMPs freeze-out, the Hubble parameter is well approximated by

$$H^2(t_f) \simeq \frac{8\pi G_N}{3} \rho_{R,f}, \quad (2.4.30)$$

where  $\rho_{R,f}$  is the radiation density at  $t_f$ . Generalizing the Stefan-Boltzmann law to an arbitrary number of relativistic degrees of freedom and noting that  $\rho_R = U/c^2$ , the Hubble parameter at  $t_f$  is given by

$$H(t_f) = \left( \frac{16\pi G_N \sigma_B}{3c^3} \right)^{1/2} T_f^2 \sqrt{g_f^*}. \quad (2.4.31)$$

An expression for  $\langle \sigma_{Av} \rangle_f$  follows from combining this with Equations 2.4.19, 2.4.22, and 2.4.29:

$$\begin{aligned} \langle \sigma_{Av} \rangle_f &= \left[ \frac{16\pi G_N \sigma_B}{3c^2} \right]^{1/2} \frac{43T_{\text{cmb}}^3}{11\sqrt{g_f^*}} \frac{1}{\Omega_\chi \rho_c} \frac{m_\chi}{T_f} \\ &\simeq \frac{10^{-27}}{\sqrt{g_f^*}} \frac{1}{\Omega_\chi h^2} \frac{m_\chi}{T_f} \text{cm}^3 \text{s}^{-1}, \end{aligned} \quad (2.4.32)$$

where  $m_\chi$  and  $T_f$  are in GeV in the second line. A useful relationship between WIMP mass and freeze-out temperature can be obtained by combining Equation 2.4.21 (evaluated at  $T_f$ ) with Equations 2.4.22 and 2.4.29:

$$\frac{m_\chi}{T_f} = 18.9 + \frac{3}{2} \ln \left( \frac{m_\chi}{T_f} \right) + \ln \left( \frac{m_\chi}{\Omega_\chi h^2 g_f^*} \right), \quad (2.4.33)$$

where  $m_\chi$  and  $T_f$  are again in GeV. If WIMPs are cosmologically interesting and dominate the present-day dark-matter density,  $\Omega_\chi h^2 \simeq 0.1$ . Furthermore, for freeze-out temperatures between  $\sim 100$  MeV and 100 GeV,  $g_f^*$  varies from  $\sim 10$  to 100 (see, *e.g.*, Figure 3 in [138] or Figures 1–3 in [434]). The argument of the logarithm in the right-most term is therefore of order unity provided the WIMP has a GeV-scale mass. In this scenario, Equation 2.4.29 requires  $m_\chi/T_f \simeq 20$ –30, and Equation 2.4.32 gives the WIMP density in terms of the annihilation cross section:

$$\Omega_\chi h^2 \simeq \frac{10^{-26} \text{cm}^3 \text{s}^{-1}}{\langle \sigma_{Av} \rangle}. \quad (2.4.34)$$

It is interesting to note that the annihilation cross section of a weakly interacting particle is given approximately by

$$\langle \sigma_{Av} \rangle_{\text{weak}} \simeq \frac{(\alpha \hbar)^2}{m_W^2} c \simeq 10^{-27} \text{cm}^3/\text{s}, \quad (2.4.35)$$

## 2.5. SUPERSYMMETRY

where  $\alpha \simeq 1/137$  is the fine-structure constant [138]. Consequently, Equation 2.4.34 implies that cosmologically interesting WIMPs have weak-scale interaction rates.

This result is often referred to as the WIMP coincidence; GeV-scale WIMPs will have a present-day density equal to the cosmologically preferred value of  $\Omega_c$  provided they have weak-scale interaction rates. There is no *a priori* reason for the dark-matter density to be tied to the weak scale. The order-of-magnitude estimates outlined above have naturally led to a plausible solution to the dark-matter problem in the form of a weakly interacting and massive particle. Furthermore, the favored mass range of cosmologically interesting WIMPs (1–1000 GeV/ $c^2$ ) is also the energy scale at which new physics is expected to explain the SM’s mass-hierarchy problem [435, 436]. WIMPs are well motivated from the point of view of cosmology and particle physics!

## 2.5 Supersymmetry

A WIMP-dominated dark-matter density implies that the SM is incomplete; there are no massive weakly interacting SM particles that are also stable (or very long lived). The SM’s three neutrino species possess most of the requisite properties, but as discussed in Section 2.4.2 neutrino-mass limits suggest that SM neutrinos are a form of hot dark matter. They cannot constitute a significant fraction of  $\Omega_c$  due to the prevalence of small-scale structure in the Universe. However, the precedent for a weakly interacting stable particle with nonzero mass should not be taken lightly. The existence of a fundamental particle that is in many respects very similar to a WIMP has been demonstrated. It is not much of a stretch to imagine that a heavier version of something like an SM neutrino could exist, particularly considering the unexplained hierarchy of fermion masses in the SM.

This line of reasoning naturally leads to the possibility of a sterile [437, 438] or fourth generation neutrino [439, 440]. While these are both interesting (and valid) WIMP candidates, their *ad hoc* addition to the SM does little to improve the larger theoretical framework. Given the success of the SM, if it is

to be extended, the additional content should not only provide a viable dark-matter candidate but also address as many of the SM's other deficiencies as possible. Perhaps the two most notable such issues are the SM mass hierarchy and the unification scale of the strong and electroweak forces. Supersymmetric extensions address both of these concerns (and more) while providing an ideal WIMP candidate. In this section I briefly review supersymmetry (SUSY) from a general point of view as well as describe a specific SUSY model that contains a relatively light WIMP. From here on, the focus of this thesis will be on the detection of WIMP dark matter, with particular emphasis on what are considered light WIMPs ( $m_\chi \simeq 1\text{--}10\text{ GeV}/c^2$ ).

### 2.5.1 General Overview

As described in Section 1.1, the SM is a quantum field theory based on principles of symmetry. It describes particles as manifestations of fields that are distinguished from one another by conserved (or partially conserved) quantum properties. In most cases, the quantum properties can be related directly to a symmetry of nature. Perhaps the most obvious is the symmetry of electric charge—usually represented by the charge-conjugation operator  $\hat{C}$ —that governs the relationship between matter and antimatter. For each particle with charge  $q$ , charge symmetry requires a nearly identical particle with charge  $-q$ . For example, charge conjugation relates an electron to a positron:

$$\hat{C} |e^-\rangle \longrightarrow |e^+\rangle. \quad (2.5.1)$$

Other SM symmetries include parity ( $\hat{P}$ ), time reversal ( $\hat{T}$ ) and gauge invariance. Of course, not all of these are perfectly obeyed symmetries. The highest-energy form of the SM Lagrangian, however, can be constructed by considering only terms that are consistent with symmetries observed in nature. Symmetry clearly plays a central role in our understanding of physical reality!

At the most basic level, SUSY can be similarly understood; it is an additional symmetry of nature that relates one type of particle to another. In the case of SUSY, the symmetry is between bosons and fermions, a spin-statistics

## 2.5. SUPERSYMMETRY

symmetry. If truly a symmetry of nature, SUSY will manifest itself as a spectrum of particles. For each boson in the SM, SUSY predicts a fermionic superpartner and *vice versa*. The spectrum of so-called superpartners is related to the SM through supersymmetric operators. The number and specific properties of these operators is model dependent. A theory with a single SUSY generator is typically described as  $N = 1$  SUSY (see, *e.g.*, [441] for further details). The most basic SUSY operator might be a direct transformation from an SM particle to a superpartner, call it  $\widehat{Q}$ . Under this transformation, for example, a spin-1/2 electron becomes a spin-0 selectron:

$$\widehat{Q} |e^-, j = 1/2\rangle \longrightarrow |\widetilde{e}^-, j = 0\rangle. \quad (2.5.2)$$

The nomenclature for SUSY particles is straightforward. Superpartners of SM bosons have similar names but with the suffix *ino* added to the end (*e.g.*, Higgsino and photino), while superpartners of SM fermions are named according to their SM counterparts but with the letter *s* added to the beginning (*e.g.*, squark and sneutrino). As in Equation 2.5.2, a superpartner is typically represented by placing a tilde over the corresponding SM symbol.

While the SM makes a very clear distinction between matter (fermions) and force carriers (bosons), SUSY predicts bosonic matter and fermionic force carriers. Just as quantum mechanics blurred the distinction between waves and particles in the early 1900s, SUSY attempts to blur the distinction between particles and interactions. This has a profound consequence regarding the possibility of a quantum theory of gravity. Gravitation is a purely attractive force. Within a quantum field-theory framework such a force must be transmitted by an even-integer field, in this case the spin-2 graviton. The electromagnetic force can be either attractive or repulsive and is carried by an odd-integer field, the spin-1 photon. Direct symmetry transformations between even- and odd-integer fields are forbidden by so-called “no go theorems” (see, *e.g.*, [442] for further details). This is one of the primary difficulties in unifying gravitation with the other forces in the SM. SUSY bridges the gap, so to speak, by connecting the graviton to the photon through a spin-3/2 gravitino. SUSY’s natural faculty for unifying matter and the force carriers through superpartner



intermediaries is one of its principle attractions, often making it a key (and some might argue necessary) ingredient in attempts to combine general relativity and the SM. There is also a certain ascetic appeal to integrating the spin degree of freedom into the theoretical framework, something that is lacking in the SM.

So why has SUSY not been discovered? There is an important difference between a symmetry like charge conjugation and SUSY. In the case of charge conjugation the operator  $\hat{C}$  does not change the mass of the particle, only the electric charge (*e.g.*, the electron and positron have the same mass). The symmetry is said to be unbroken. If SUSY were an unbroken symmetry of nature, all the superpartners would have masses equal to their SM counterparts. A spin-0 selectron with a  $511 \text{ keV}/c^2$  mass, however, has not been observed in nature or in the laboratory. With the center-of-mass energies available at past and present particle accelerators and colliders, such a particle would have been easily created and detected. This is generically true for the superpartners of any SUSY model. If SUSY exists it must be a very badly broken symmetry, and the symmetry-breaking scale and spectrum of superpartner masses is apparently beyond the energies currently accessible in the laboratory. Spontaneously broken symmetries are nothing new to particle physics. The symmetry that unifies the electromagnetic and weak forces is spontaneously broken at the weak scale, giving rise to weak bosons with GeV-scale masses. Perhaps supersymmetry is also a spontaneously broken symmetry of nature, and the SUSY-breaking scale is such that the masses of the superpartners are elevated beyond the few hundred  $\text{GeV}/c^2$  currently accessible in the laboratory.

Since its discovery in the early 1970s [443, 444, 445] and despite the lack of any experimental evidence for its existence, SUSY has blossomed into a rich (and favored) field of theoretical particle physics. I cannot hope to do justice to the SUSY theorists' ingenuity and diligence over the past four decades, not to mention the multitude of specific SUSY models that have been constructed and explored. In what follows I will attempt to outline a few of the main motivations for searching for SUSY. I apologize ahead of time if the discussion

## 2.5. SUPERSYMMETRY

comes across as oversimplified; I am but a humble experimentalist and some of these concepts are beyond my experience. For more general and detailed SUSY discussions I refer the interested reader to three reviews from the mid-1980s that I found particularly instructive. The highly technical review by Nilles in [441] pays particular attention to the relationship between SUSY and supergravity; gravity appears as a natural component of field theories that invoke SUSY as a local (as opposed to global) symmetry. The review by Haber and Kane in [446] is more general as well as more accessible, with particular emphasis on experimental constraints resulting from laboratory searches for SUSY. Finally, Sohnius' review in [442] is perhaps the most mathematically complete and contains an interesting historical perspective in the introduction. For a more recent and pedagogical reference, Baer and Tata's *Weak Scale Supersymmetry* might be more appropriate [447].

### 2.5.2 Motivation

#### The Higgs Mass

One of the most unsettling features of the SM is the Higgs mass. As mentioned in Section 1.1, the Higgs is responsible for electroweak symmetry breaking and is the SM's only fundamental scalar. As such, it has some unique properties. All particles in the SM receive corrections to their masses due to so-called loop diagrams. This is demonstrated for the Higgs at the 1-loop level in Figure 2.16. These radiative corrections diverge logarithmically for SM fermions, and are manageable within the theory through renormalization. For the Higgs, however, the 1-loop correction diverges quadratically. Stated in another way, if the SM is a low-energy approximation of a more fundamental theory that comes into play at some scale  $\Lambda_{\text{uv}}$ , the 1-loop correction to the Higgs mass is given by

$$\delta m_{\text{Higgs}}^2 \propto g^2 \Lambda_{\text{uv}}^2, \quad (2.5.3)$$

where  $g$  is a coupling constant. Although there is no guarantee that a finite ultraviolet cut-off exists, a couple of natural new physics scales are the

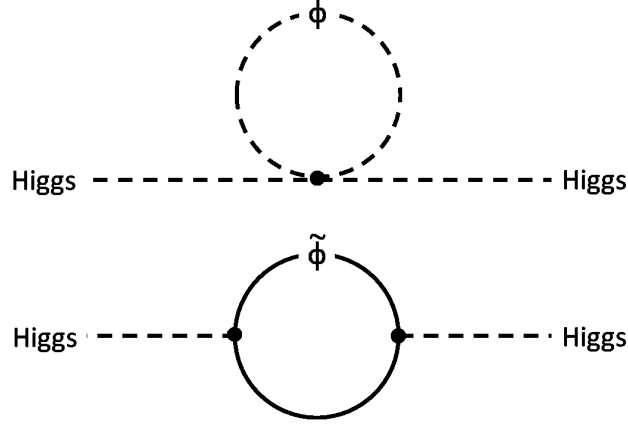


Figure 2.16: Feynman diagrams of 1-loop radiative contributions to the Higgs propagator. The top diagram represents the quadratically divergent contribution to the Higgs mass from its scalar self-correction, while the bottom diagram represents the additional, opposite sign correction (also quadratically divergent) from the fermionic superpartner to the SM scalar. In SUSY extensions to the SM, both diagrams contribute to the Higgs mass such that the quadratic divergence cancels, preserving the weak-scale at the few hundred GeV level without fine tuning.

gauge-force unification scale at  $\mathcal{M}_{\text{gut}} \simeq 10^{16} \text{ GeV}/c^2$  and the Planck scale at  $\mathcal{M}_p \simeq 10^{19} \text{ GeV}/c^2$ . In any case, in the absence of an intermediate physics scale, the Higgs mass should be effectively infinite due radiative self-corrections. However, the Higgs vacuum expectation value drives electroweak symmetry breaking and sets the weak scale at  $E_{\text{weak}} \simeq 200 \text{ GeV}$ . The measured values of the weak-boson masses give empirical proof that radiative corrections have not driven the weak scale to the GUT scale or beyond. The only way to prevent the Higgs mass from diverging to the next natural scale is to fine tune the coupling constant  $g$  to one part in  $\Lambda_{\text{uv}}^2/E_{\text{weak}}^2$ . If we assume  $\Lambda_{\text{uv}} \simeq \mathcal{M}_{\text{gut}} c^2$ ,  $g$  has to be fine tuned to 1 part in  $\sim 10^{28}$ ! This is often referred to as the naturalness or hierarchy problem; how can  $g$  be so finely controlled so as to maintain the hierarchy between the weak scale and the next natural scale?

## 2.5. SUPERSYMMETRY

One solution is to add SUSY to the SM. In SUSY extensions, the Higgs mass receives corrections from both the SM field (top of Figure 2.16) and its superpartner (bottom of Figure 2.16). Since the SM field and its superpartner have opposite spin statistics, their 1-loop diagrams enter into the Higgs mass correction with opposite signs:

$$\delta m_{\text{Higgs}}^2 \propto g^2 (\Lambda_{\text{uv}}^2 + m_\phi^2) - g^2 (\Lambda_{\text{uv}}^2 + m_{\tilde{\phi}}^2) = g^2 (m_\phi^2 - m_{\tilde{\phi}}^2), \quad (2.5.4)$$

where  $\phi$  is the SM scalar field and  $\tilde{\phi}$  is its fermionic superpartner. The cancellation is not exact since SUSY is a broken symmetry, *i.e.*,  $m_\phi \neq m_{\tilde{\phi}}$ . In order for the weak scale to be preserved at the few hundred GeV level without fine tuning, SUSY should come into play at scale of  $\Lambda_{\text{uv}} \simeq E_{\text{weak}}/g$ , or a few TeV. The new SUSY particles would naturally have masses just below this scale, making any stable superpartner a cosmologically interesting WIMP candidate.

### Gauge Force Unification

A generic feature of quantum field theories is the so-called “running” of the coupling constants. The idea of a coupling constant was introduced above, where a factor of  $g^2$  was included in Equations 2.5.3 and 2.5.4 to represent the interaction points (or vertices) in the 1-loop diagrams. Coupling constants are associated with a theory’s force carriers and measure the strength of the associated interactions. They show up explicitly as dimensionless coefficients of terms in a theory’s Lagrangian, and they are not truly constant. In general, a theory’s coupling constants are a function of the energy transferred from one particle to another across a given vertex; that is, coupling constants run either up or down in strength as a function of energy.

## CHAPTER 2. DARK MATTER

It is customary to write the coupling constants associated with the SM gauge interactions as

$$\begin{aligned}\alpha_1 &= \frac{5}{3} \frac{g'^2}{4\pi} = \frac{5\alpha}{3\cos^2\theta_W}, \\ \alpha_2 &= \frac{g^2}{4\pi} = \frac{\alpha}{\sin^2\theta_W}, \quad \text{and} \\ \alpha_3 &= \frac{g_s^2}{4\pi},\end{aligned}\tag{2.5.5}$$

where  $\alpha$  (without a subscript) is the fine-structure constant and  $\theta_W$  is the weak-mixing angle that arises from electroweak symmetry breaking. Recall that the SM's gauge symmetry can be expressed as the group product  $SU(3) \otimes SU(2) \otimes U(1)$ .  $g_s$ ,  $g$  and  $g'$  are the coupling constants associated with the  $SU(3)$  (strong),  $SU(2)$  (weak) and  $U(1)$  (electromagnetic) gauge groups, respectively, and  $g$  and  $g'$  are related through the weak-mixing angle:  $g'/g = \tan\theta_W$ . It is well established experimentally that  $\alpha_1$  is correlated with the interaction energy  $Q$ , while  $\alpha_2$  and  $\alpha_3$  are both anticorrelated with  $Q$ . That the strong force gets weaker as the distance between colored objects decreases (or the energy increases) is known as asymptotic freedom. At sufficiently high energies, strongly interacting, colored fields can escape color confinement because  $\alpha_3 \rightarrow 0$  as  $Q \rightarrow \infty$  (see, *e.g.*, Figure 10 in [448]). This is a direct result of the running of the strong coupling constant.

The idea of unification is that at a sufficiently high energy the SM's coupling constants will meet at a common value; the forces of nature (excluding gravitation) will unify into a single strong-electroweak force. Of course, *a priori* there is no guarantee that the couplings will converge to a single value. To test the idea within the framework of the SM, a renormalization scheme has to be specified. A commonly used scheme is the so-called modified minimal subtraction ( $\overline{MS}$ ) scheme [449, 450]. A detailed discussion of renormalization is beyond the scope of this thesis. In short, the couplings are specified at some starting point and then run up to higher energies by solving the renormalization-group (RG) equations (see, *e.g.*, [451] for further details). In the  $\overline{MS}$  scheme, the starting point is the  $Z^0$  mass, where  $Q \simeq 91$  GeV. From a simultaneous fit to

## 2.5. SUPERSYMMETRY

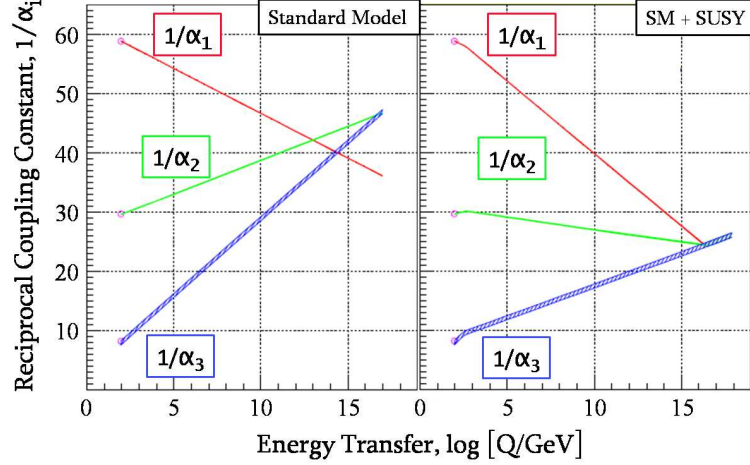


Figure 2.17: The running of the Standard Model’s three gauge couplings as a function of interaction energy with (right panel) and without (left panel) supersymmetry. Within the context of the Standard Model the couplings do not meet at a single point and gauge-force unification is ruled out at high significance. A supersymmetric extension to the Standard Model in which the superpartners have TeV-scale masses allows the couplings to meet at the Grand Unified Scale (GUT) scale of  $\mathcal{M}_{\text{gut}} \simeq 10^{16} \text{ GeV}/c^2$ . Figure adapted from [451].

data measured at the LEP and Tevatron colliders, the starting-point values of the couplings are  $\alpha_1 \simeq 0.017$ ,  $\alpha_2 \simeq 0.034$ , and  $\alpha_3 \simeq 0.12$  [88]. The solution to the RG equations is

$$\frac{1}{\alpha_i(Q^2)} = \frac{1}{\alpha_i(m_Z^2 c^4)} - b_i \log \left( \frac{Q^2}{m_Z^2 c^4} \right), \quad (2.5.6)$$

where  $m_Z \simeq 91 \text{ GeV}/c^2$  and the coefficients  $b_i = (41/10, -19/6, -7)$  [451]. This result is shown in the left panel of Figure 2.17. Clearly the SM couplings do not run to a single point; unification within the SM is ruled out at the  $8\sigma$  confidence level!

One of the major motivations for extending the SM with SUSY is the possibility of gauge-force unification. Although this is largely an ascetic pursuit, it is not without precedent. After all, the weak and electromagnetic forces

are understood as low-energy approximations of a higher-energy unified interaction, the electroweak force. The right panel in Figure 2.17 demonstrates that a unified strong-electroweak force is indeed possible if SUSY is included in the SM at the TeV scale. In the example shown, the superpartners have TeV-scale masses that modify the slopes of the coupling constants' energy dependence (at  $Q \simeq 1$  TeV) such that the strength of the unified gauge interaction is  $\alpha_{\text{gut}} \simeq 0.04$  at  $Q \simeq 10^{16}$  GeV. This is the motivation for the GUT scale given in the previous section. Some might consider the failure of unification within the SM a near miss, a failure that new physics at the TeV scale is likely to correct.

### 2.5.3 The MSSM, LSP & Neutralino

To understand why SUSY is relevant to the dark-matter problem it is instructive to review a specific model. A full description of any SUSY model is beyond the scope of this thesis, but a partial sketch of the so-called Minimal Supersymmetric Standard Model (MSSM) is still useful. For a formal introduction to the MSSM, I refer the interested reader to Chapter 12 in Mohapatra's *Unification and Supersymmetry* textbook [452]. The MSSM's guiding principle is that it contains the smallest field content required to give rise to the SM fields we know and love, and is therefore the simplest to write down. It is the model's field content that will be the focus of the discussion here.

In terms of ordinary, nonsuperpartner fields, the MSSM contains all the SM fields and an additional Higgs doublet. The additional degrees of freedom in the Higgs sector are necessary to give mass to both up- and down-type quarks and to preserve supersymmetry following electroweak symmetry breaking without introducing anomalies into the theory. The Higgs bosons are still responsible for electroweak symmetry breaking, which results in five physical degrees of freedom in the Higgs sector. Two are electrically neutral and CP even, often referred to as the light ( $h^0$ ) and heavy ( $H^0$ ) Higgs scalars, two are charged Higgs bosons ( $H^\pm$ ), and the fifth is a neutral pseudoscalar (CP odd) commonly denoted  $A^0$ .

## 2.5. SUPERSYMMETRY

The particle spectrum of superpartners is constructed to match a SUSY degree of freedom to each SM degree of freedom. Since the SM fermions each have two spin degrees of freedom (and a scalar has only one), the MSSM contains two squarks and two sleptons for each quark and lepton, respectively. For example, to compliment the SM's up quark, the MSSM contains left- ( $\tilde{q}_L^u$ ) and right-handed ( $\tilde{q}_R^u$ ) up-squark interaction eigenstates. Antiquarks, leptons and antileptons are similarly duplicated (*e.g.*, the  $\bar{\tau}$  is accompanied by the  $\tilde{\tau}_R^*$  and  $\tilde{\tau}_L^*$ , where the  $*$  indicates a superpartner antiparticle). Most models allow for mixing between the left- and right-handed sfermion degrees of freedom. The resulting squark(slepton)-mass eigenstates are sometimes denoted  $\tilde{q}_1$  ( $\tilde{l}_1$ ) and  $\tilde{q}_2$  ( $\tilde{l}_2$ ). As right-handed neutrinos and left-handed antineutrinos are absent from the SM, only 3 sneutrinos ( $\tilde{\nu}$ ) and 3 antisneutrinos ( $\tilde{\nu}^*$ ) are required to complete the list of sfermions.

The superpartners to the gauge bosons are generically referred to as gauginos. Since the gluon is a spin-1 particle and the gluino ( $\tilde{g}$ ) is a spin-1/2 particle, there are an equal number of each in the MSSM. In the SM, prior to electroweak symmetry breaking, the  $SU(2) \otimes U(1)$  gauge degrees of freedom are represented by four massless bosons, commonly denoted  $W^i$  ( $i=1, 2$  or  $3$ ) and  $B$ . The Higgs mechanism spontaneously breaks the  $SU(2) \otimes U(1)$  symmetry, resulting in the weak and electromagnetic mass eigenstates known as the  $W^\pm$ , the  $Z^0$ , and the  $\gamma$ . The  $SU(2) \otimes U(1)$  gauginos are also mixed into linear combinations that result in mass eigenstates. The main difference is that the gauginos also mix with any Higgsinos that share the same quantum numbers.<sup>8</sup> The two charged Higgsinos ( $\tilde{H}^\pm$ ) mix with the charged  $W$ -inos ( $\tilde{W}^\pm$ ), resulting in four charginos ( $\tilde{\chi}_1^\pm$  and  $\tilde{\chi}_2^\pm$ ). The neutral Higgsinos mix with the  $B$ -ino and the  $W^3$ -ino, resulting in four neutralinos denoted by  $\tilde{\chi}_i^0$  ( $i=1, 2, 3$  or  $4$ ). The neutralinos are labeled according to mass, with the  $\tilde{\chi}_1^0$  the lightest and the  $\tilde{\chi}_4^0$  the heaviest.

---

<sup>8</sup>Gluinos can only mix among themselves as none of the Higgsinos or other gauginos are colored. It is common in the literature to reserve the name gaugino for the noncolored gauge-boson superpartners, the  $W$ -inos and the  $B$ -ino.



## CHAPTER 2. DARK MATTER

A well known feature of the SM is its conservation of fermion number. For example, any SM interaction in which there are no fermions in the initial state must have an equal number of fermions and antifermions in the final state. In supersymmetric field theories, the concept of fermion number conservation generalizes to something known as  $R$ -symmetry (see, *e.g.*, [446] for further details). I have intentionally avoided a discussion of how supersymmetry is broken in the MSSM as it is the most complicated part of the model. However, since SUSY has yet to be observed, it is clear that the MSSM must include a mechanism for breaking SUSY and giving the superpartners masses that exceed current experimental limits. One of the results of SUSY breaking is that the continuous  $R$ -symmetry is partially broken, leaving behind a discrete symmetry known as  $R$ -parity.  $R$ -parity is characterized by the quantum number

$$R = (-1)^{3b+l+2j}, \quad (2.5.7)$$

where  $j$ ,  $b$  and  $l$  are the spin, baryon and lepton quantum numbers.  $R = +1$  for ordinary particles and  $R = -1$  for superpartners. Although it is possible to construct a theory that also violates this discrete symmetry, experimental constraints involving baryon- and lepton-number violating processes (*e.g.*, the lifetime of the proton) place severe restrictions on  $R$ -parity violation. If SUSY is a part of nature, it would appear that  $R$ -parity is a perfect (or very nearly perfect) symmetry. This has an important consequence regarding the stability of at least one of the MSSM mass eigenstates described above. In the absence of  $R$ -parity conservation, there is nothing to stop superpartners created in the early Universe from decaying to lighter stable states of ordinary matter and radiation as the Universe cools. Conservation of  $R$ -parity means that sparticles decay into an odd number of superpartners in addition to ordinary matter. This prevents all of the superpartners from being unstable. There has to remain a Lightest Supersymmetric Particle (LSP) that the other superpartners decay down to.  $R$ -parity prevents the LSP from decaying into ordinary matter and radiation except through annihilation with its antiparticle (or with itself in the case of a Majorana LSP). A stable LSP is precisely what makes SUSY interesting from a dark-matter perspective. It would have been produced in

## 2.5. SUPERSYMMETRY

the early Universe as described in Section 2.4.3, with the expansion of the Universe eventually halting LSP annihilations and leaving behind a cosmologically interesting LSP density if the mass of the LSP is in the GeV to TeV range.

An obvious question is which of the MSSM superpartners is likely to be the LSP. Charged, uncolored sparticles like sleptons and charginos are unlikely because they would bind to ordinary matter to form stable, heavy states. These heavy states would be intermixed with normal matter, and searches for anomalously heavy protons have effectively ruled out an electrically charged LSP over most of the cosmologically interesting parameter space [453, 454]. Squarks and gluinos might also bind into heavy states. If these states are electrically charged, they are similarly constrained by the heavy proton searches. However, there is no guarantee that colored superpartners would not form neutral hadronic states that need not mix with ordinary matter, and might thus avoid the experimental limits. Many argue that a squark or gluino LSP is disfavored on theoretical grounds; the gluino is expected to be heavier than the lightest neutralino and the squarks are expected to be heavier than the lightest slepton [138]. A colored LSP is therefore not considered likely, but is possible. Further arguments against an electrically charged or colored LSP can be found in [455]. A sneutrino LSP would naïvely appear to be an ideal candidate for WIMP dark matter. However, their range of cosmologically interesting masses ( $550 \text{ GeV}/c^2 \lesssim m_{\tilde{\nu}} \lesssim 2.3 \text{ TeV}/c^2$ ) was essentially ruled out by early dark-matter direct-detection experiments [456]. That leaves the neutralino as the candidate favored by most to be the LSP. In particular, if the lightest neutralino  $\tilde{\chi}_1^0$  is the LSP, it is an ideal WIMP candidate since it is weakly interacting (by definition), massive (due to SUSY breaking), and stable (due to  $R$ -parity conservation). Furthermore, the neutralino parameter space predicted by (most) SUSY models is still largely unexplored (see, *e.g.*, Figure 2.18).

As mentioned above, the MSSM's neutralinos are linear combinations of the  $B$ -ino,  $W^3$ -ino, and the two neutral Higgsinos ( $\tilde{H}_1^0$  and  $\tilde{H}_2^0$ ). These four sparticles form the basis from which the neutralino-mass eigenstates are constructed.

The MSSM neutralino mass matrix can be written as<sup>9</sup>

$$\mathcal{M}_{\tilde{\chi}^0} = \begin{pmatrix} m_1 & 0 & -m_Z c_\beta s_W & m_Z s_\beta s_W \\ 0 & m_2 & m_Z c_\beta c_W & -m_Z s_\beta c_W \\ -m_Z c_\beta s_W & m_Z c_\beta c_W & 0 & -\mu \\ m_Z s_\beta s_W & -m_Z s_\beta c_W & -\mu & 0 \end{pmatrix}, \quad (2.5.8)$$

where  $m_Z$  is the mass of the  $Z^0$ ,  $m_1$  and  $m_2$  are the gaugino-mass parameters,  $\theta_W$  is the weak-mixing angle,  $\mu$  is the Higgsino-mass parameter, and  $\tan \beta_h$  is the ratio of the vacuum expectation values (VEVs) of the Higgs bosons. The lightest neutralino can be expressed as

$$\tilde{\chi}_1^0 = N_{11}\tilde{B} + N_{12}\tilde{W}^3 + N_{13}\tilde{H}_1^0 + N_{14}\tilde{H}_2^0, \quad (2.5.9)$$

where the coefficients  $N_{ij}$  are entries in the  $4 \times 4$  matrix that diagonalizes the neutralino-mass matrix (see Appendix A in [2] for detailed expressions). Since our interest is primarily in the LSP, for simplicity (and unless otherwise noted) I will denote the  $\tilde{\chi}_1^0$  as  $\chi$  from here on. Furthermore, the term neutralino will be used to refer to the lightest neutralino.

There are three distinct possibilities regarding the content and general properties of the neutralino. The  $\chi$  can be described by its gaugino and Higgsino content with the parameters

$$\begin{aligned} f_g &= |N_{11}|^2 + |N_{12}|^2, \text{ and} \\ f_h &= |N_{13}|^2 + |N_{14}|^2, \end{aligned} \quad (2.5.10)$$

the gaugino and Higgsino fractions, respectively. When  $\mu \gg m_1$  and  $m_2$ ,  $f_g \gg f_h$  and the neutralino is gaugino dominated—specifically, when  $f_g > 0.99$ . The relationship between  $m_1$  and  $m_2$  is determined by the RG equations. If the coupling constants and gaugino masses are unified at the GUT scale, at the electroweak scale

$$m_1 = \frac{5}{3}m_2 \tan^2 \theta_W \simeq \frac{1}{2}m_2. \quad (2.5.11)$$

---

<sup>9</sup>I have use a shorthand notation to represent the cosine and sine trigonometric functions in order fit Equation 2.5.8 onto one line. The letters  $c$  and  $s$  with a subscript  $W$  ( $\beta$ ) refer to the  $\cos \theta_W$  ( $\cos \beta_h$ ) and the  $\sin \theta_W$  ( $\sin \beta_h$ ), respectively.

## 2.5. SUPERSYMMETRY

In this scenario the neutralino's mass is given by the lighter of the two gaugino-mass parameters,  $m_1$ , and the neutralino is primarily  $B$ -ino in nature. Since the SM's photon is a linear combination of the  $B$  and  $W^3$  fields, a gaugino dominated neutralino is sometimes referred to as a photino. Another possibility is that  $\mu \ll m_1$  and  $m_2$ , in which case  $f_h \gg f_g$  and the neutralino is Higgsino dominated—specifically, when  $f_g < 0.01$ . Finally, if  $\mu$  and  $m_1$  are similarly sized such that  $0.01 \leq f_g \leq 0.99$ , the neutralino is both Higgsino- and gaugino-like, and is considered mixed.

The composition of the neutralino is important when calculating its scattering (with ordinary matter) and annihilation cross sections within the context of a specific model. Certain mixtures, for example, might favor large spin-dependent scattering cross sections with terrestrial matter (as opposed to spin-independent), or perhaps annihilations into down-type rather than up-type quark pairs. Such considerations can affect the materials chosen to construct detectors, or the spectrum of photons targeted by observatories. That is to say, if the dark matter is composed of neutralinos, the specific SUSY annihilation and scattering channels associated with different neutralino mixtures can have important consequences regarding dark-matter detectability. This is demonstrated in Figure 2.18, where predictions for the neutralino-nucleon elastic scattering cross section are plotted as a function of neutralino mass for a particular MSSM model with different neutralino mixtures. On average, a gaugino-like neutralino has a smaller mass and larger cross section than a Higgsino-like neutralino [457], making it easier to detect.

Another reason for showing Figure 2.18 is to introduce the parameter space that is the primary target of direct-detection dark-matter experiments, cross section versus mass. At the most basic level, a direct-detection experiment attempts to measure the rate at which dark-matter particles in the Milky Way's halo scatter from nuclei in a terrestrial detector. If the detected rate of interactions is consistent with zero (or with expected background rates), the duration of the measurement, the size of the detector, and the detector's energy threshold can be interpreted in terms of an upper limit on the neutralino-nucleon scattering cross section as a function of neutralino mass. This is precisely what

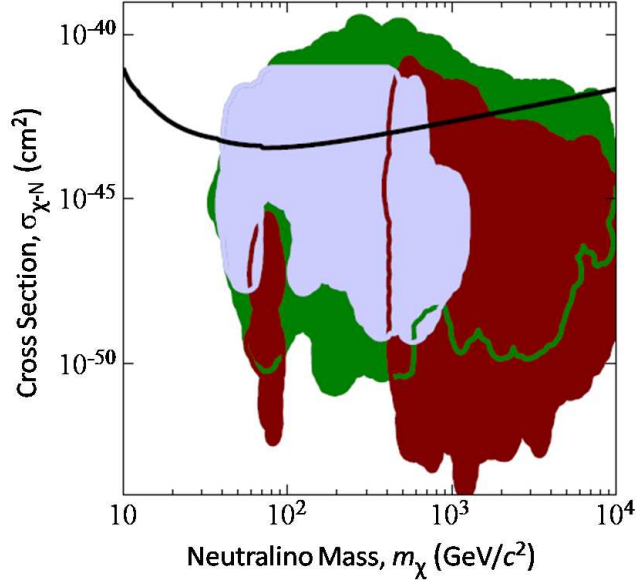


Figure 2.18: Neutralino-nucleon elastic scattering cross-section predictions (filled regions, 90% C.L.) as a function of neutralino mass. The three regions are the results of scanning allowed parameter space for a constrained MSSM model in which the neutralino is Higgsino dominated (disjoint dark red regions), gaugino dominated (smaller light blue region), and mixed (larger green region). The model calculations are courtesy of Gondolo *et al.*'s DarkSUSY code [458]. The solid (black) line stretching from left to right is an upper limit from the Cryogenic Dark Matter Search II experiment [459]; neutralino parameter space above this curve is excluded at 90% confidence. This curve is close to the sensitivity limit of the current generation of direct-detection experiments, and is therefore a useful benchmark. Figure adapted from [460].

is shown in Figure 2.18 (black, solid line that stretches from left to right) for data from the Cryogenic Dark Matter Search II (CDMS II) experiment [459]. In the absence of a positive detection, the goal is to exclude as much of the neutralino parameter space as possible. The CDMS exclusion curve in Figure 2.18 is near the sensitivity limit of the current generation of direct-detection experiments. Direct detection, the CDMS experiment, and the derivation of a

## 2.5. SUPERSYMMETRY

neutralino cross-section limit from direct-detection data are discussed in much greater detail in the following chapters.

Due to space and time constraints I will not describe specific neutralino scattering channels in detail. For further information I refer the interested reader to Appendices B and C in [2], and the related discussion in [138]. Rather, the neutralino interaction cross section with normal matter will be treated as an unknown parameter to be explored experimentally. As is clear from Figure 2.18, not much is gained by detailed knowledge of a particular model's scattering channels anyway. Predictions of the neutralino-nucleon cross section span roughly a dozen orders of magnitude! Models that predict cross sections well below current experimental sensitivities are common in the literature. From an experimental point of view, it is sufficient to assume that the neutralino's cross section is very small, and thus strive to construct as sensitive a detector as possible.

The focus of this thesis is the neutralino mass. The standard type of constrained MSSM calculation shown in Figure 2.18 clearly disfavors a neutralino mass less than  $\sim 30 \text{ GeV}/c^2$ . In fact, nearly all recent scans of the neutralino parameter space disfavor a neutralino mass less than  $\sim 40 \text{ GeV}/c^2$  [461, 462, 463]. I digress a bit in the following section to discuss neutralino mass and an MSSM model that predicts a relatively light neutralino. It is worth spending some time to understand the difference between models with and without light neutralinos. Nevertheless, all model calculations are intended mainly as a guide. They indicate that the neutralino cross sections and masses investigated in this work are feasible from a theoretical point of view.

### 2.5.4 Neutralino Mass

Despite being minimal, the general MSSM has over 100 free parameters. Testing a theory with so many unknown quantities is incredibly difficult. It is therefore common practice to reduce the number of free parameters through ascetically motivated assumptions. The idea is to start by assuming that the MSSM takes the simplest form possible. While such a model might be unre-

alistic, it vastly simplifies comparison of experimental results. If experiments manage to completely rule out the allowed parameter space for the most constrained version of the MSSM, some of the assumptions are relaxed and the number of free parameters is increased until a higher-dimensional parameter space is found that places SUSY just beyond current experimental sensitivities. This “procedure” is likely to continue until SUSY is discovered. Unfortunately, it is not as straightforward a prescription as it sounds. Because SUSY is popular and different theorists favor different assumptions, there is a substantial proliferation of methods for constraining the MSSM in the literature. In this section I discuss how different methods for constraining the MSSM affect the allowed mass of the neutralino.

### mSUGRA and the LEP $m_\chi$ Lower Limit

One of the most widely studied constrained MSSM models is minimal SuperGRAvity (mSUGRA) (see, *e.g.*, [464] for formal details). In mSUGRA, several seemingly reasonable assumptions reduce the number of free parameters added to the SM by the addition of SUSY to 5. The assumptions are summarized in the following list.

- **Gauge coupling unification** requires that the SM’s three coupling constants meet at a common value at the GUT scale as indicated in Figure 2.17;  $\alpha_1(\mathcal{M}_{\text{gut}}) = \alpha_2(\mathcal{M}_{\text{gut}}) = \alpha_3(\mathcal{M}_{\text{gut}})$ .
- **Unification of the gaugino masses** requires that the  $W$ -ino,  $B$ -ino and gluino share a common mass,  $m_{1/2}$ , at the GUT scale. In many SUSY theories this condition is related to gauge coupling unification.
- **Universal scalar masses** sets the masses of the Higgs scalars and the sfermions to a common value,  $m_0$ , at the GUT scale. This is perhaps the most far reaching and least justifiable assumption.
- **Universal trilinear couplings** sets the values of certain coupling coefficients in the part of the Lagrangian that breaks SUSY to be equal at the GUT scale.

## 2.5. SUPERSYMMETRY

When these assumptions are carried through electroweak symmetry breaking, the only additional free parameters in mSUGRA that were not already a part of the SM are  $\tan \beta_h$  (the ratio of the Higgs VEVs),  $m_{1/2}$ ,  $m_0$ , the trilinear coupling  $A_0$ , and the sign of the Higgsino mass parameter  $\mu$ . mSUGRA model calculations (often referred to as scans) in which these free parameters are allowed to vary (but are restricted to be consistent with experimental data) can be converted into predictions for the neutralino mass and cross section. A recent scan of the allowed neutralino parameter space for mSUGRA is shown in Figure 2.19, where the lightest neutralino predicted has a mass  $\sim 40\text{--}50 \text{ GeV}/c^2$  [465].

The principle reason for the mSUGRA neutralino-mass lower limit in Figure 2.19 is the inclusion of constraints derived from searches for SUSY at particle colliders. To date, some of the most sensitive searches for SUSY have been performed with data acquired at the Large Electron-Positron (LEP) collider [469] in Geneva, Switzerland. These searches generally fall into one of three categories: 1) attempts to produce and observe superpartners directly; 2) measurements of missing energy indicative of the production (and escape) of neutralinos; and 3) improved measurements of SM interactions to look for anomalous behavior. The following list gives a sampling of a few of the studies conducted with the  $\sim 200 \text{ GeV}$  (center-of-mass energy)  $e^\pm$  collisions at the LEP facility.

- A search for sleptons with the ALEPH detector [470].
- Searches for superpartners with the DELPHI detector [471].
- Searches for SUSY with the L3 detector [472].
- Searches for new particles with the OPAL detector [473].
- Searches for the SM Higgs boson by the LEP Higgs Working Group [474].
- Electroweak measurements and constraints by the LEP Electroweak Working Group [475].

Unfortunately, none of these searches found any evidence for physics beyond the SM. If superpartners were produced at the LEP collider, the rates were too low



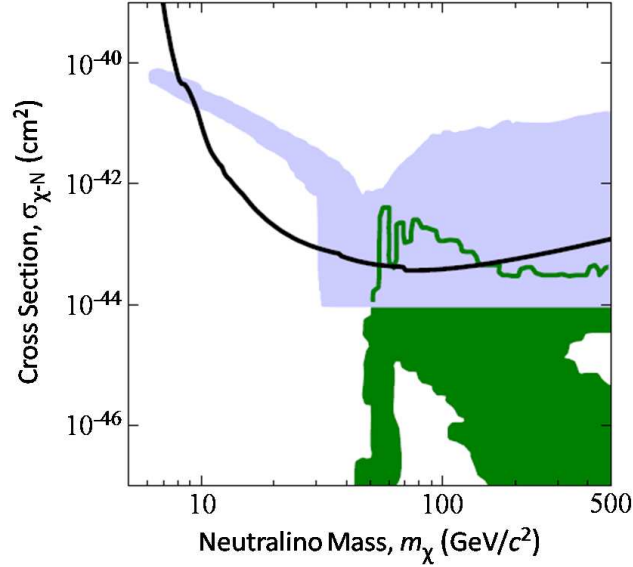


Figure 2.19: Comparison of constrained MSSM model predictions (filled regions, 90% C.L.) for the neutralino-nucleon cross section as a function of neutralino mass. Results from scans of the mSUGRA parameter space (green, lower-right region), that include experimental constraints from searches for SUSY at particle colliders, set a lower limit on the neutralino mass of  $\sim 40\text{--}50 \text{ GeV}/c^2$  [465]. Those constraints do not apply if some of mSUGRA’s assumptions are relaxed, resulting in a funnel of light-neutralino models with relatively large cross sections (light blue, funneled region extending to low masses) [466, 467, 468]. Note that the light-neutralino model calculations are cut off at  $\sigma_{\chi-N} = 10^{-44} \text{ cm}^2$  because the data are unavailable; it is likely that smaller cross sections are also possible. Neutralinos with masses  $\lesssim 7 \text{ GeV}/c^2$  are predicted that just evade the CDMS II upper limit [459] (black, solid line), motivating a targeted search for light neutralinos. Figure adapted from [460].

to observe above SM backgrounds. In the absence of a positive detection, the LEP data can be used to constrain mSUGRA free parameters. Furthermore, when all of the relevant data from the four main LEP detectors (ALEPH, DELPHI, L3 and OPAL) are interpreted in the context of mSUGRA, a lower limit on the neutralino mass can be derived [476]. The limit depends (in part) on assumptions made regarding some of the mSUGRA free parameters, and

## 2.5. SUPERSYMMETRY

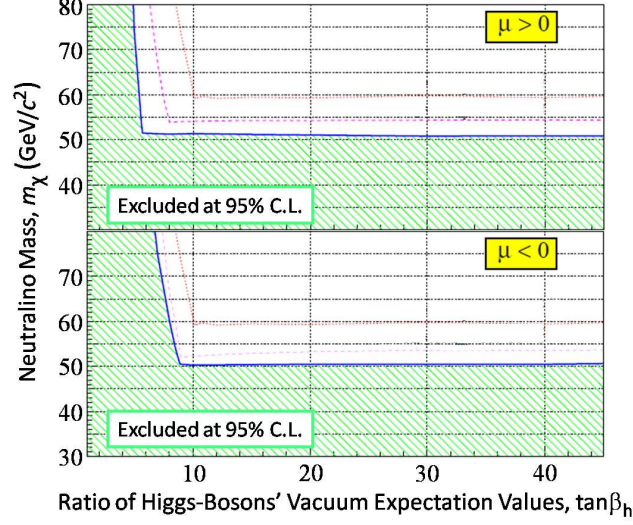


Figure 2.20: Constraints on the mSUGRA neutralino mass versus the ratio of the vacuum expectation values of the Higgs bosons ( $\tan \beta_h$ ) from a combined analysis of data taken at the LEP facility. mSUGRA’s universal trilinear coupling  $A_0$  is allowed to take any value for the shaded exclusion regions (outlined by a blue, solid line and for which  $m_{\text{top}} = 175 \text{ GeV}/c^2$ ), while  $A_0 = 0$  for the barely visible dashed (pink,  $m_{\text{top}} = 180 \text{ GeV}/c^2$ ) and dotted (red,  $m_{\text{top}} = 175 \text{ GeV}/c^2$ ) lower limits. The universal scalar-mass parameter  $m_0$  is limited to  $< 1 \text{ TeV}/c^2$ , while  $\text{sign}(\mu) = +1$  for the upper panel and  $\text{sign}(\mu) = -1$  for the lower panel. All things considered, the LEP data exclude mSUGRA neutralinos with masses  $\lesssim 50 \text{ GeV}/c^2$  at 95% confidence. Figure taken from [476].

is shown in Figure 2.20 in the  $m_\chi$ – $\tan \beta_h$  plane for both possible values of  $\text{sign}(\mu)$ . The lightest mSUGRA neutralino allowed by the LEP data has a mass  $\gtrsim 50 \text{ GeV}/c^2$  (at 95% confidence), which is why the mSUGRA predictions shown in Figure 2.19 do not include lighter neutralinos.

Similar neutralino-mass lower limits have been derived for slightly less constrained MSSM models. The same authors responsible for the limits in Figure 2.20 have also analyzed the combined LEP data set in the context of the so-called constrained MSSM (CMSSM) model (characterized by 6 free param-

eters), and find  $m_\chi \geq 47 \text{ GeV}/c^2$  [477]. In an analysis by the ALEPH collaboration, the combined LEP data are used to explore a more general (but still constrained) MSSM model (with nonuniversal trilinear couplings at the GUT scale), obtaining  $m_\chi > 42.4 \text{ GeV}/c^2$  at 95% confidence [478].

### Light Neutralinos

The LEP neutralino-mass lower limits are well known and are so often quoted that some believe them to be absolute. It is worth remembering that the  $m_\chi \gtrsim 40 \text{ GeV}/c^2$  limit is only valid provided nature's manifestation of SUSY obeys a set of far reaching, unproven assumptions at a grand-unification scale that may or may not exist. Regardless of how reasonable and ascetically pleasing unified masses and universal couplings might seem, it is important to explore other possibilities until experimental evidence suggests otherwise.

In particular, there are a number of studies in the literature in which grand-unification schemes disfavor a common value for the gaugino masses at the GUT scale. Most of these theories attempt to envelop the SM into an  $SU(5)$  symmetry group that includes local SUSY. The overall theoretical framework is similar to mSUGRA, but the gaugino masses do not unify at the GUT scale. In some cases, unification of the coupling constants at the GUT scale implies fixed-ratio relationships between the gaugino masses that are not one-to-one [479, 480], while in others very little can be said regarding the relationship between the masses of the  $SU(3)$ ,  $SU(2)$  and  $U(1)$  gauginos [481, 482]. From an experimental point of view, the existence of such models is sufficient to justify exploring the consequences of relaxing the gaugino-mass unification assumption.

This is precisely the motivation behind the Light Neutralino Model (LNM) developed by Bottino and collaborators [483]. First introduced in 2003 [466, 467, 468], the LNM model is similar to the mSUGRA framework in that it is a constrained version of the MSSM. However, the simplifying assumptions employed to reduce the number of MSSM free parameters are less restrictive:

- Unification of the gauge couplings is the same as in mSUGRA.
- The gaugino masses need not share a common value at the GUT scale.

## 2.5. SUPERSYMMETRY

- The squark masses share a common value at the GUT scale:  $m_{\tilde{q}i} \equiv m_{\tilde{q}}$ .
- The slepton masses are unified at the GUT scale:  $m_{\tilde{l}i} \equiv m_{\tilde{l}}$ .
- All trilinear couplings are set to zero except for the third-generation sfermions:  $A_{\tilde{b}} = A_{\tilde{t}} \equiv A m_{\tilde{q}}$  and  $A_{\tilde{\tau}} \equiv A m_{\tilde{l}}$ .

As a result, the LNM has 8 independent parameters (in addition to the SM's 19):  $\mu$ ,  $\tan\beta_h$ , the mass of the pseudoscalar Higgs  $m_A$ ,  $m_{\tilde{q}}$ ,  $m_{\tilde{l}}$ ,  $A$ ,  $m_1$  and  $m_2$ . The latter two are the  $U(1)$  and  $SU(2)$  gaugino-mass parameters. As mentioned in the previous section, when  $\mu \gg m_1$  and  $m_2$ , the neutralino is gaugino dominated, and the neutralino mass is set by  $m_1$  if the gauge couplings are unified at the GUT scale. Additionally, the lightest chargino mass is set by  $m_2$ . In mSUGRA, the gaugino-mass unification assumption requires that  $m_1 \simeq m_2/2$  at the electroweak scale (see Equation 2.5.11). The LEP neutralino-mass limits depend critically on this relationship. The LEP data directly constrain the lightest chargino mass to be  $\gtrsim 100 \text{ GeV}/c^2$ , which translates into a neutralino-mass lower limit of  $\sim 50 \text{ GeV}/c^2$ . In the LNM, the gaugino masses need not be related in a two-to-one ratio at the electroweak scale. The LEP data still constrain the mass of the lightest chargino, but no longer constrain the neutralino mass. When astrophysical constraints on the dark-matter density are taken into account, neutralinos as light as  $\sim 6 \text{ GeV}/c^2$  can be accommodated within the LNM framework. The results of a scan of the allowed LNM  $m_{\chi}-\sigma_{\chi-N}$  parameter space are shown in Figure 2.19. A remarkable “funnel” of low-mass neutralino models with relatively large cross sections is predicted, the tip of which just eludes the CDMS II limit, providing motivation for a targeted search for low-mass-neutralino dark matter.

Independent scans of the neutralino parameter space for constrained MSSM models with nonunified gaugino masses roughly confirm Bottino *et al.*'s LNM results. Two groups, in particular, scan the neutralino parameter space by employing nearly identical models and experimental constraints. Belanger *et al.* find neutralino masses as low as  $6 \text{ GeV}/c^2$  are possible provided the Higgs pseudoscalar is light ( $m_A < 200 \text{ GeV}/c^2$ ) and  $\tan\beta_h$  is large [484] (see, *e.g.*,

Figure 14 in [484] for the equivalent of Figure 2.19). Calibbi *et al.* find neutralino masses as low as  $8 \text{ GeV}/c^2$  are possible provided  $\tan \beta_h$  is large [485] (see, *e.g.*, Figure 9 in [485]). However, it should be noted that recent results from the Compact Muon Solenoid (CMS) collaboration appear to rule out values of  $\tan \beta_h$  larger than  $\sim 30$  in searches for light neutral Higgs bosons ( $m_A \simeq 100 \text{ GeV}/c^2$ ) [486]. Both the Belanger and Calibbi neutralino-mass lower limits become more strict for low values of  $\tan \beta_h$ , increasing to  $\sim 15$ – $30 \text{ GeV}/c^2$ . In an analysis with few theoretical assumptions, Hooper and Plehn find a similar MSSM neutralino-mass lower limit:  $m_\chi \geq 18 \text{ GeV}/c^2$  [487]. On the other hand, the authors in [488, 489] argue that massless neutralinos are possible if astrophysical constraints on dark matter are relaxed.

Even if light MSSM neutralinos are ruled out, it has been shown that within the framework of the Next to Minimal Supersymmetric Standard Model (NMSSM), neutralino masses as low as  $1 \text{ GeV}/c^2$  are consistent with the LEP bounds (see, *e.g.*, Figure 12 in [490]). In addition to the field content of the MSSM, the NMSSM includes a new gauge singlet. The lightest NMSSM neutralino is a linear combination of the four MSSM components (see Equation 2.5.9) and the superpartner of the new singlet, the singlino. As described in [491], this extra degree of freedom allows GeV-mass neutralinos to easily evade current experimental constraints.

## Chapter 3

# Dark Matter Detection

Based on the evidence provided in the previous two chapters, the case for a significant (and pervasive) nonbaryonic dark-matter density in the Universe is compelling. Perhaps the most well motivated dark-matter candidate is a neutral relic that lies outside the realm of the SM and might interact weakly with normal matter, a WIMP. Theories of physics beyond the SM that contain a stable WIMP abound in the literature. In particular, the lightest neutralino of most supersymmetric models is an ideal candidate. Unfortunately, even the most highly constrained SUSY models cannot accurately predict the properties of the neutralino (see, *e.g.*, Figures 2.18 and 2.19). A broad experimental program tasked specifically with searching for and characterizing the dark matter is therefore of the utmost importance. In this chapter I review some of the techniques employed in the effort to detect WIMP dark matter. They generally fall into one of three categories: 1) direct production at a particle collider; 2) indirect detection of WIMP-annihilation products; and 3) direct detection of WIMPs interacting with terrestrial matter. Where appropriate, how a detection technique depends on the mass of the WIMP is emphasized. Potential evidence for a light WIMP is discussed as well.

### 3.1 Production at Particle Colliders

While attempts to produce and detect new particles that lie outside the SM have grown into a colossal, multi-national effort, the technique is (in spirit) quite simple. Particles such as electrons and protons (and/or their antiparticles) are accelerated to near the speed of light and smashed together. If the center-of-mass energy of the collision is sufficiently large, the particles will scatter inelastically, converting a fraction of the initial-state momentum into the rest masses of heavier particles. In the case of particle colliders, the collision points are surrounded by highly segmented detectors capable of tracking the electrically charged collision products. Large magnetic fields applied near the collision point cause charged particles to leave curved tracks, allowing identification of their masses and other properties. Calorimeters are usually interspersed with the trackers to detect photons and neutral hadrons (*e.g.*, light mesons such as the  $\pi^0$ ,  $\eta$  and  $\rho$  are common). The heaviest particles tend to be the most transient, typically decaying (almost immediately) to lighter particles whose tracks must be reconstructed in order to deduce the properties of the parent. Modern particle detectors are designed specifically to track heavy particles that might travel only a few millimeters from the collision point before decaying, while also tracking longer lived decay products that might travel many meters. Minimum ionizing muons, in particular, can penetrate many kilometers through dense materials. Muons are the primary reason for the tremendous size of modern detectors. A schematic of the CMS detector at the LHC is shown in Figure 3.1, illustrating the design of a modern particle detector.

Many of the SM's fundamental particles were first created and detected in the laboratory using the method described above (*e.g.*, the weak bosons and third-generation quarks). With the advancements in particle-acceleration technology over the last several decades, larger center-of-mass energies have become available for the production of heavier and heavier states, a direct consequence of Einstein's famous  $E = mc^2$  relationship. The drive to push collision energies to ever higher levels in the pursuit of new, heavy states is

### 3.1. PRODUCTION AT PARTICLE COLLIDERS

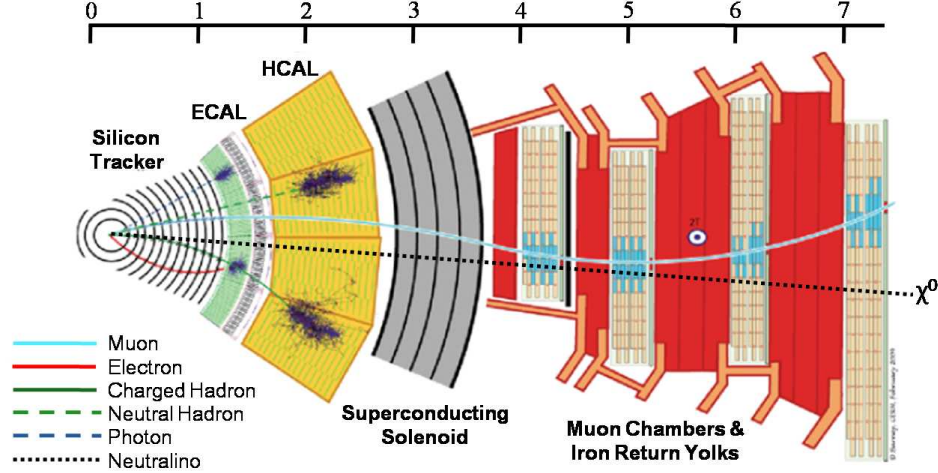


Figure 3.1: A cross-sectional schematic view of the Compact Muon Solenoid (CMS) detector at the Large Hadron Collider (LHC). The physical scale is given in meters along the top. The collision point for the LHC’s twin proton beams is surrounded (from left to right) by a silicon tracker, an electromagnetic calorimeter (ECAL), a hadronic calorimeter (HCAL), a superconducting solenoid that generates a 4 T magnetic field within the silicon tracker (perpendicular to the plane of the page), and a series of muon chambers interspersed with iron slabs that act as the return yoke for the magnetic field. A variety of particle interactions are depicted, demonstrating the intended utility for each detector segment. As indicated, neutralinos (or stable WIMPs in general) would escape the detector without interacting with any of the detector elements. Figure adapted from [492] and courtesy of CERN.

often referred to as the “energy frontier.” Although center-of-mass energies of several TeV are now possible, no particles with masses greater than the top quark have yet been discovered ( $m_t = 172 \text{ GeV}/c^2$  [88]).

Part of the difficulty has to do with production rates; heavy particles near the energy frontier are not created as often as collections of (less interesting) lighter particles. This means that despite TeV-scale collision energies, the majority of the tracks recorded at colliders are due to well known SM particles, so-called SM “backgrounds.” Consequently, to increase the likelihood of creating



### CHAPTER 3. DARK MATTER DETECTION

a heavy state, colliders operate at high luminosity—as many particle collisions per second as can be managed—which in turn requires advanced hardware and software capable of handling large data rates.

Searches for new particles can involve the reduction of many millions of recorded “events” into statistical distributions that describe the kinematics and particle content of the collisions (*e.g.*, transverse momentum). Monte Carlo packages like GEANT4 [493] are used to simulate the same kinematical distributions under the hypothesis that the SM is a sufficient description of reality. The simulated distributions represent the SM backgrounds, and are often combined using sophisticated statistical techniques to find a region in the kinematical parameter space that is expected to have relatively few background events while remaining sensitive to new physics. A statistically significant excess in the data above the expected rate of SM backgrounds might be an indication that a new particle was created in a handful of events.

Detection of WIMPs at a collider is complicated by their long lifetimes and weak-scale interaction rates. As indicated in Figure 3.1, for example, if a neutralino were produced in a collision it would escape the detector without depositing energy in any of the detector elements. Fortunately, since neutrinos behave in exactly this manner, a method for searching for “missing energy” (denoted  $\cancel{E}$ ) has been developed specifically to search for long-lived (or stable) particles capable of escaping direct detection. In the case of neutralino production, if  $R$ -parity is conserved, at least two neutralinos will escape and at least  $2m_\chi c^2$  of the initial-state energy will be missing in the tally of detected particle energies. Care must be taken to avoid confusion with missing energy due to neutrinos. Events with lepton tracks are commonly excluded from neutralino searches because they are often the result of weak decays with final-state neutrinos. A further complication arises when attempting to detect light neutralinos. Each detector element has a finite energy resolution. When the corresponding energy variances are summed over all of an event’s detectable decay products, they can conspire to produce an undermeasurement that mimics missing energy. With the energies available at modern facilities like the

### 3.1. PRODUCTION AT PARTICLE COLLIDERS

LHC, mismeasuring a few hundred GeV is difficult. However, finite-resolution effects can easily lead to a false missing energy of several GeV, or even several tens of GeV. Consequently, even if the  $\gtrsim 6 \text{ GeV}/c^2$  neutralinos predicted by the LNM model (see Section 2.5.4) are produced at the LHC, they are likely to be incredibly difficult to detect.

As an example, let's consider a search for SUSY recently conducted by the CMS collaboration [494], utilizing data acquired in proton collisions at the LHC's (current) 7 TeV center-of-mass energy. The most probable neutralino production channels at the LHC are indirect; neutralinos are created through the production and subsequent decay of heavier, colored sparticles (*e.g.*, squark-squark, squark-gluino and gluino-gluino pair production). Recall that if  $R$ -parity is conserved, each sparticle created will eventually decay into an LSP. Assuming the neutralino is the LSP, pair produced squarks and gluinos will decay into a pair of neutralinos as well as some number of quarks and gluons. The colored SM particles immediately pull other colored particles out of the vacuum so as to create color-neutral hadrons, a violent process that results in hadronic-particle "jets." The event signature is therefore two or more jets and missing energy. Furthermore, since the jets are decay products of (presumably) heavy sparticles, the total jet energy is expected to be (on average) larger than for SM backgrounds (see, *e.g.*, Figure 1 in [494]).

A first pass selection of events (or preselection) starts at the hardware level, and is further narrowed down by a series of software "cuts" on kinematic quantities deduced "offline" from reconstructed tracks and calorimetrically measured energies. The hardware trigger requires the total transverse jet energy,  $H_T$ , to exceed 150 GeV (in the plane perpendicular to the proton beams), resulting in many millions of events in  $35 \text{ pb}^{-1}$  of integrated luminosity. The data sample is reduced to  $\sim 4.7$  million events by further restricting  $H_T > 250 \text{ GeV}$  with an offline cut, increasing the likelihood of events with multiple jets. Only events with two or more jets and no lepton tracks are considered. The two leading jets for each event are required to have transverse energies  $E_T > 100 \text{ GeV}$ , while any additional jets must have  $E_T > 50 \text{ GeV}$ . Restrictions are placed on the jet topologies, ensuring they are well defined and have energies fully contained

### CHAPTER 3. DARK MATTER DETECTION

in the hadronic calorimeter. To reduce the contribution of the SM multi-jet background, an additional cut rejects any events with  $H_T < 350$  GeV, yielding a final preselection of  $\sim 900,000$  events.

The final event selection employs a variable called  $\alpha_T$  that has been shown to be an effective discriminator against the SM multi-jet background in searches for SUSY [495, 496].  $\alpha_T$  is most easily understood in the context of “di-jet” events, in which there are exactly two jets:

$$\alpha_T \equiv E_T^{j2}/\mathcal{M}_T, \quad (3.1.1)$$

where  $E_T^{j2}$  is the transverse energy of the least energetic of the two jets, and  $\mathcal{M}_T$  is the transverse mass of the di-jet event (see [494] for further details). For perfectly back-to-back di-jet events with perfectly measured jet energies,  $\alpha_T = 0.5$ . Mismeasured jet energies will tend to yield  $\alpha_T < 0.5$ , while events with true missing energy can have  $\alpha_T > 0.5$ . For events with more than two jets, two pseudo-jets are formed by clustering the jets into two groups such that the difference between the pseudo-jets’ transverse energies is minimized. Simulations of the multi-jet background that include the full CMS detector geometry and performance indicate that the SM multi-jet background is effectively removed by requiring  $\alpha_T > 0.55$ . In light of the  $H_T > 350$  GeV preselection cut, this  $\alpha_T$  cut is equivalent to requiring a total transverse missing energy  $> 140$  GeV. Although this does not translate directly into neutralino-mass sensitivity, it is easy to see that the neutralino detection efficiency for this analysis falls off as neutralino mass decreases.

The final event selection for  $\alpha_T > 0.55$  yields 37 events with  $31 \pm 5$  expected from SM backgrounds.  $\sim 20$  events are expected from the SM multi-jet background, 4 from decays of the  $Z^0$  to two neutrinos, 4 from decays involving  $W$ s and jets, and 3 from production of top-quark pairs (with subsequent decay to jets). Despite the slight excess, the number of observed events is statistically consistent with the expected backgrounds. The data therefore agree with the SM and are inconsistent with the production of neutralinos (and other sparticles). In the absence of a positive detection, it is common practice to use the data to constrain a particular SUSY model. Figure 3.2 shows the resulting ex-

### 3.1. PRODUCTION AT PARTICLE COLLIDERS

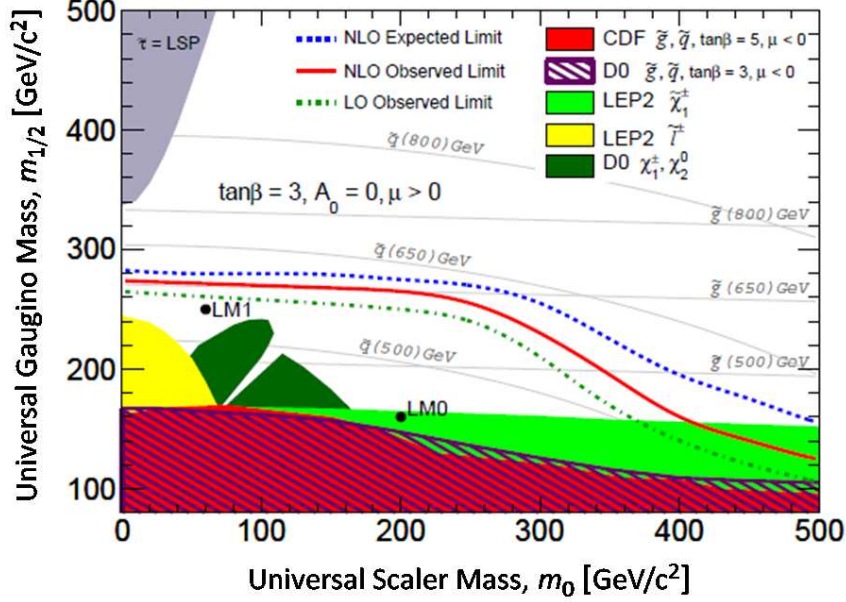


Figure 3.2: mSUGRA exclusion limits on the universal scalar- and gaugino-mass parameters for  $\tan \beta_h = 3$ ,  $A_0 = 0$  and  $\text{sign}(\mu) > 0$ . A limit derived from a multi-jet plus missing energy search for SUSY with the CMS detector (red, solid line) is compared to sensitivity calculations that include leading-order (LO) Feynman diagrams (green, dash-dotted line) and next-to-leading-order diagrams (blue, dashed line). Models in which the combination of  $m_{1/2}$  and  $m_0$  fall below the limits are excluded at 95% confidence, including two SUSY benchmark models known as LM0 and LM1 [495] (dots). Exclusion regions based on older data recorded with detectors at the Tevatron (D0 [497, 498] and CDF [499]) and LEP [476] facilities are also shown. The gray region in the upper-left corner is excluded because in those models the LSP is the stau. Curves corresponding to particular gluino and squark masses (light gray, solid lines) are given as well. mSUGRA squarks with masses less than  $500 \text{ GeV}/c^2$ , for example, are disfavored by the CMS limit. Figure taken from [494].

clusion limits for mSUGRA in the  $m_{1/2}$ – $m_0$  plane, with  $\tan \beta_h = 3$ ,  $A_0 = 0$ , and  $\text{sign}(\mu) > 0$ . For further details regarding ongoing searches for SUSY and other so-called “exotica” at the LHC, I refer the interested reader to [500]. Prospects for discovering a light neutralino in particular are discussed in [501, 502].

### CHAPTER 3. DARK MATTER DETECTION

The exclusion limits shown in Figure 3.2 represent searches for SUSY, rather than the type of targeted search for neutralino dark matter represented by the CDMS II limit in Figures 2.18 and 2.19. In general, searches for the production of sparticles (or other exotica) at colliders are not specifically dark-matter searches. While it is hoped that a discovery of physics beyond the SM at the LHC will provide a hint to help unravel the solution to the dark-matter problem, there is no guarantee that a long-lived WIMP produced at a collider is the dark matter. The discovery of SUSY and neutralinos at the LHC would be compelling evidence, but a detection of relic neutralinos in the halo of the Milky Way would be needed to convince the community that dark matter had indeed been produced in the laboratory. Ideally, the detection of a long-lived WIMP at a collider would be accompanied by a WIMP signal from either (or both) types of dark-matter experiments discussed in the remainder of this chapter. In this sense, high-energy physics experiments and indirect- and direct-detection dark-matter experiments are complementary.

## 3.2 Indirect Detection

“Indirect detection” of dark matter typically refers to the observation of ordinary forms of matter and radiation that result from dark-matter pair-annihilation. If, as suspected, the dark matter is composed of thermally produced WIMP relics, the annihilation cross section is exceedingly small (as discussed in Section 2.4.3). Based on evidence for dark matter in galaxies and clusters (see, *e.g.*, Sections 2.2 and 2.3), WIMPs are generally too sparse for the low (expected) rate of WIMP annihilations to produce a flux of annihilation products large enough to distinguish from astrophysical backgrounds. Consequently, indirect dark-matter searches tend to target regions of space in which the WIMP density is expected to be the greatest. Assuming there exists a WIMP density sufficient to cause a detectable annihilation signal, the situation is further complicated by the possibility of unknown astrophysical backgrounds not associated with dark matter; it is difficult to know with certainty that an indirect signal is truly due to WIMP annihilations. As a result, although an incontrovertible indirect dark-matter signal has yet to be uncovered, data from indirect measurements have caused great excitement and controversy within the community.

The detectable products of WIMP annihilation depend on several factors. If the WIMP source is too far away, the solid angle subtended by the flux of annihilation products will be too small to be detectable above backgrounds. Nearby sources for which the WIMP density is expected to be large offer the best chance. The cores of the Milky Way and its satellite galaxies are prime targets, which means the WIMPs will be nonrelativistic when they annihilate, otherwise they would have escaped their host galaxy’s gravitational potential. The energy available to the annihilation products is therefore roughly equal to twice the mass of the WIMP. Under most scenarios, only particles lighter than WIMPs will be among the debris of an annihilation, which means SM particles. In the case of neutralino dark matter, for example, annihilations can result in the direct production of: 1) fermion pairs through  $Z^0$  or pseudoscalar-Higgs exchange; 2) weak-boson pairs through neutralino or chargino exchange; 3) pairs of

### CHAPTER 3. DARK MATTER DETECTION

Higgs bosons (or a Higgs and weak boson) through Higgs, neutralino or chargino exchange; or 4) photon pairs (or a photon and weak boson) through more complicated exchanges involving loops with charginos, fermions, sfermions, weak bosons, and/or Higgses. For annihilations in which pairs of particles with identical masses are produced, the only final states permitted are those with fermions, weak bosons, and Higgses whose masses are less than  $m_\chi$  (as well as the two-photon final state). For mixed final states, if the mass of the neutralino is in between the masses of the two final-state particles, annihilations are possible in which a particle is produced that is more massive than the neutralino (*e.g.*, if the neutralino is lighter than the  $Z^0$ , annihilation into a photon and  $Z^0$  could still occur). I refer the interested reader to Appendix B in [2] for further details regarding specific diagrams and annihilation-rate calculations involving neutralinos. Since all but the lightest fermions are unstable, as are the weak and Higgs bosons, most of these annihilation products will decay immediately to lighter secondaries. Among the detectable primary annihilation products, photon and neutrino pairs are particularly interesting because they are neutral and either exactly or effectively massless; both will travel basically unmolested from the source of annihilations and will have energies nearly equal to the mass of the dark-matter particle. The narrow spectral line(s) expected from WIMPs annihilating directly into photons provides an almost unmistakable dark-matter signature [503]. Unfortunately, the expected rate for (most) dark-matter candidates to annihilate directly into photons is very small; there are no “tree-level” Feynman diagrams contributing to direct photon production, only diagrams involving one or more loops [2]. Detecting annihilations via neutrinos is also expected to be difficult due to their weak-interaction rates. For the same reason, however, they provide the only known channel through which WIMP annihilations in the center of the Sun [504] (or the Earth [505]) might be detectable. Secondary photons and neutrinos would be more difficult to discriminate from astrophysical backgrounds because their expected spectrum of energies is less well defined. However, sources of high-energy photons and neutrinos are rare enough that an indirect dark-matter signal might still be discernible if WIMPs are massive enough. Antimatter primary and secondary

### 3.2. INDIRECT DETECTION

annihilation products (*e.g.* positrons, antiprotons and antideuterons) are also interesting because astrophysical sources of antimatter are rare. In the remainder of this section I briefly review a few of the experimental efforts to measure these indirect WIMP signals, including a signal (detected in gamma rays) from the center of the Milky Way that might be consistent with annihilations of a light WIMP.

#### 3.2.1 Electrons & Positrons

The flux of cosmic-ray electrons and positrons has been the focus of considerable debate in the literature recently. This is primarily due to a series of measurements from satellite- and balloon-born instruments with sensitivity to the positron and/or total  $e^\pm$  flux. The positron flux is particularly interesting as there are not many known astrophysical sources of positrons with energies greater than a few GeV. Postulated sources include positrons created in the magnetospheres of nearby pulsars or as the result of dark-matter annihilations, while a known “conventional” source is expected in which positrons are created when cosmic-ray nuclei interact inelastically with interstellar gas. These are also potential sources of high-energy electrons. The study of electrically charged cosmic rays is complicated by any radiation or magnetic fields they might encounter, including the CMB, the Galactic magnetic field, and the Sun’s magnetosphere. Simulations of the production and transport of conventional  $e^\pm$  cosmic rays through the Milky Way predict a featureless energy spectrum that falls off rapidly with energy [506]. The PAMELA [507] and ATIC [508] experiments have recently measured high-energy features that appear to be inconsistent with this fall off, and might be consistent with dark-matter annihilations. A more accurate (and even more recent) measurement of the total  $e^\pm$  flux by the Fermi-LAT [509] experiment is featureless, but harder than expected. The following is a review of these seemingly contradictory measurements and their relevance as potential indirect dark-matter signals.



### The PAMELA Positron Excess

The Payload for Antimatter-Matter Exploration and Light-nuclei Astrophysics (PAMELA) is a satellite-born apparatus designed specifically for the identification of electrically charged cosmic rays of antimatter [507, 510]. In many respects, it is similar to a high-energy particle-physics detector. The main instrument is a spectrometer that is comprised of a silicon tracker embedded in an  $\sim 0.5$  T magnetic field, and is capable of measuring a particle’s momentum and the sign (and absolute value up to  $Z = 6$ ) of its electric charge. A series of plastic scintillators above and below the spectrometer act as a hardware trigger, picking out only particles traveling through the front of the detector and measuring a particle’s transit time through the apparatus. More plastic scintillators surround the spectrometer so that events contaminated by charged particles entering the sides of the apparatus can be rejected (“vetoed”). Following the spectrometer, a calorimeter distinguishes positrons (antiprotons) from protons (electrons) with high efficiency. This is a critical part of the measurement, as the flux of protons (electrons) exceeds the expected flux of positrons (antiprotons) by approximately three (two) orders of magnitude in the energy range of interest. Furthermore, the calorimeter can measure the energy of electrons that cause electromagnetic showers, providing a check on the energies measured by the spectrometer. PAMELA’s ability to identify antimatter from like-charged matter is aided by an additional plastic scintillator that is sensitive to charged shower fragments escaping out the back of the calorimeter. The final active element is a  $^3\text{He}$  neutron detector at the end of the apparatus.  $^3\text{He}$ ’s relatively large neutron-capture cross section gives sensitivity to any neutrons produced by showers in the calorimeter. As hadronic showers are expected to yield 10–20 times as many neutrons as electromagnetic showers, the  $^3\text{He}$  neutron detector further improves PAMELA’s ability to separate antimatter signals from their like-charged backgrounds.

The PAMELA apparatus was launched in June of 2006 aboard a Russian satellite, and has successfully measured the fluxes of both galactic and solar cosmic rays over a wide range of energies [521]. From a data set including

### 3.2. INDIRECT DETECTION

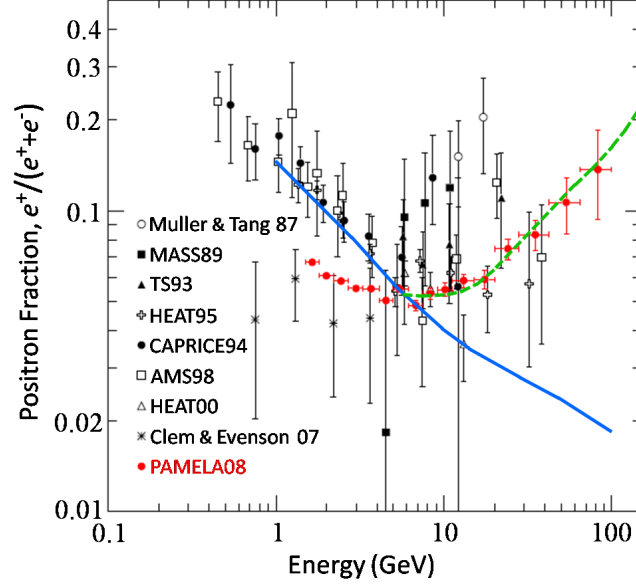


Figure 3.3: Comparison of measurements of the cosmic-ray positron fraction—flux of positrons relative to the total flux of positrons and electrons—as a function of energy. References for the measurements in the legend are (from top to bottom, respectively) [511, 512, 513, 514, 515, 516, 517, 518, 519]. The discrepancy below  $\sim 5$  GeV is due to the time varying solar modulation of the cosmic-ray flux. Above 10 GeV the PAMELA data exhibit a compelling rise that is inconsistent with a theoretical model [506] (blue, solid line) of conventional positron production due to inelastic scattering of cosmic rays from interstellar gas. A model (green, dashed line), in which high-energy positrons result from neutralinos annihilating in nearby clumps [520], roughly reproduces the PAMELA data. Figure adapted from [519] and [520].

$\gtrsim 10^9$  event triggers taken over the course of  $\sim 500$  days, just over 151,000 electrons and nearly 9,500 positrons with energies from 1.5 to 100 GeV were identified [519]. The data are plotted in Figure 3.3 as a function of energy in terms of the cosmic-ray positron fraction—the flux of positrons divided by the total flux of positrons and electrons. The positron fraction is more robust against systematic uncertainties than the absolute positron flux; uncertainties that are common to both positrons and electrons will cancel in the flux ratio. Further-

### CHAPTER 3. DARK MATTER DETECTION

more, provided an instrument has equal efficiency for detecting electrons and positrons, detailed knowledge of its geometric acceptance is unnecessary since this also factors out of the positron fraction. The authors in [519] were careful to select only events for which the response to electrons and positrons is expected to be the same.

For comparison, the data from a number of older experiments are also shown in Figure 3.3. Below  $\sim 5$  GeV there is a clear discrepancy between recent measurements (PAMELA [519] and Clem & Evenson [518]) and measurements made prior to 2000. This is apparently explained by a reversal of the Sun's magnetic field. Measurements by the Balloon-born Experiment with a Superconducting Spectrometer [522] (BESS), before and after the turn of the millennium, clearly demonstrate an increase in antiproton cosmic rays (relative to protons) that has been interpreted in terms of a modulation of the Sun's magnetic field [523]. Apparently, the polarity of the field reversal is such that (after the reversal) negatively (positively) charged cosmic rays are being preferentially swept toward (away from) the Earth's path, causing an increase in the  $\bar{p}/p$  ratio. This would have the opposite effect on the positron fraction, particularly at low energy where the positrons and electrons are more susceptible to deflection by the Sun's magnetic field.

Above about 5 GeV, the PAMELA data roughly agree with previous measurements. For the first time, however, the statistical accuracy of the data is sufficient to discern a clear rise in the positron fraction for energies greater than  $\sim 10$  GeV. The main source of positrons in the Milky Way is thought to be from inelastic collisions between cosmic-ray nuclei and interstellar gas. As indicated in Figure 3.3, a theoretical model of conventional positron production faithfully reproduces the positron fraction between  $\sim 5$  and 8 GeV, but is far too soft to explain the PAMELA data out to 100 GeV.

Although not as statistically significant, previous measurements of the positron fraction caused considerable speculation in the literature regarding potential sources of positrons with energies  $\gtrsim 10$  GeV. Supersymmetric models in which neutralino dark matter annihilates into  $W$  bosons or  $\tau$  leptons (with subsequent decays leading to positrons) appear capable of producing an

### 3.2. INDIRECT DETECTION

excess of high-energy cosmic-ray positrons if the dark-matter density is somehow “boosted” above that expected from relic-density calculations [524, 525]. A boost factor might be reasonable if the dark matter is clumpy rather than diffuse. Nearby clumps of neutralinos could provide the requisite positron excess without overproducing other annihilation products (*e.g.*, antiprotons and gamma rays) that are not observed [526, 520]. Transport of positrons through the radiation and magnetic fields of the Milky Way generally softens their energy spectrum. Consequently, a relatively massive neutralino is needed to produce positrons with energies as large as 100 GeV. A clumpy neutralino model from [520] is shown in Figure 3.3 that reproduces the high-energy PAMELA data provided  $m_\chi \simeq 800 \text{ GeV}/c^2$  and the neutralino clumps are no more than  $\sim 1 \text{ kpc}$  away.

A pulsar-based explanation of the PAMELA positron excess that does not involve dark matter has also gained favor in the literature. A pulsar is a rapidly rotating neutron star that emits beams of electromagnetic radiation from its magnetic poles. A pulsar’s magnetic field can be quite large (surface fields of  $> 10^8 \text{ T}$  are not uncommon), and when combined with its rapid rotation can cause large electric fields capable of accelerating electrons along magnetic field lines. The acceleration experienced by these electrons can be sufficient to cause synchrotron photons energetic enough to create electron-positron pairs that lead to electromagnetic showers and further positron production [527]. Alternatively, a single electron ejected from a pulsar’s surface might experience a potential drop sufficient to cause synchrotron radiation and subsequent positron production. Pulsar models that produce cosmic-ray positrons with energies in excess of 100 GeV are feasible [528, 529]. A nearby population of pulsars might partially or fully explain the PAMELA positron excess, an idea I explore in more detail following the description of the Fermi-LAT  $e^\pm$  flux measurement (see below).

### The ATIC Electron-Positron Bump

The Advanced Thin Ionization Calorimeter [508, 530] (ATIC) is a balloon-born experiment intended to measure the absolute flux of cosmic rays arriving at the Earth's upper atmosphere. The instrument design is similar to the PAMELA apparatus; several layers of plastic scintillator select particles entering the front of the instrument, and are followed by a calorimeter that differentiates hadronic from electromagnetic showers and measures event energies. One striking difference is the lack of a spectrometer and magnetic field. The ATIC apparatus targets a higher-energy part of the cosmic-ray spectrum. The payload-size limitations of a balloon-born experiment make the measurement of curved tracks for particles with energies up to a few TeV impractical. Instead, no attempt is made to differentiate matter and antimatter; the magnitude of a cosmic ray's charge is measured, but not its sign. A pixelated-silicon-matrix detector, at the very front of the apparatus, efficiently separates cosmic rays with charge  $|q| \geq 2e$  from protons, electrons and their antiparticles (for which  $|q| = e$ ), and from gamma rays (for which  $|q| = 0$ ). Several layers of graphite, interspersed between the plastic scintillators at the front of the instrument, induce a particle cascade (or shower), whereupon the cosmic ray's energy is subdivided among many secondary particles. The shower propagates more or less directly to the calorimeter where it develops further in the calorimeter's dense bismuth germanate (BGO) scintillating crystals. Electromagnetic cascades initiated by gamma rays, electrons and positrons do not spread out as much in the calorimeter as hadronic showers caused by protons (or antiprotons). Furthermore, electromagnetic showers tend to deposit most of their energy in the calorimeter, whereas hadronic showers are still developing as they exit the back of the ATIC apparatus. These differences in event topology allow for efficient selection of electrons and positrons from the combined  $|q/e| = 1$  event sample. As the rate of high-energy cosmic-ray positrons represents (at most) a few percent of the total electron and positron flux, the absolute flux measurement to which the ATIC apparatus is sensitive is (effectively) the flux of high-energy cosmic-ray electrons.

### 3.2. INDIRECT DETECTION

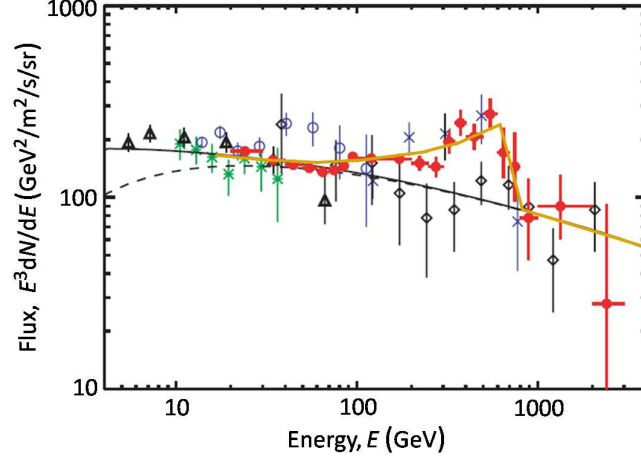


Figure 3.4: Measurements of the combined differential flux of cosmic-ray electrons and positrons as a function of energy, scaled by a factor of  $E^3$  such that the flux is in units of  $\text{GeV}^2$  per square-meter per second per steradian. The ATIC spectrum [531] (red, filled circles) is compared with data from emulsion chambers [532] (black diamonds) and the AMS [515] (green stars), HEAT [533] (black triangles), BETS [534] (blue, open circles), and PPB-BETS [535] (blue crosses) experiments. The expected spectrum of cosmic-ray electrons and positrons from conventional sources (black, solid line) fails to reproduce the ATIC and PPB-BETS data between  $\sim 300$  and  $800$  GeV. The effect of the Sun’s modulating magnetic field (black, dashed line) is also indicated. The high-energy excess could be the result of dark-matter annihilations in the halo of the Milky Way. Annihilations of a  $\sim 600$  GeV Kaluza-Klein particle [536] or  $\sim 800$  GeV neutralino [520] might yield a feature in the  $e^\pm$  spectrum (roughly indicated by the orange, solid line) similar to the ATIC excess under certain restricted scenarios. Figure adapted from [531].

The ATIC experiment recorded valuable data during two balloon flights launched from Antarctica [531]. The resulting combined flux of electrons and positrons with energies between  $\sim 20$  GeV and 3 TeV is shown in Figure 3.4 as a function of electron energy,  $E$ . Because the spectrum of cosmic-ray electrons falls off sharply as  $E$  increases, the differential flux—number per

### CHAPTER 3. DARK MATTER DETECTION

square meter per second per steradian per GeV—is multiplied by a factor of  $E^3$  to make any features in the spectrum more obvious. A computer package called GALPROP [537] is used to simulate the conventional  $e^\pm$  production via cosmic-ray interactions in the interstellar medium, as well as their transport through the radiation and magnetic fields of the Milky Way. The predicted spectrum is in good agreement with the ATIC data (as well as data from previous measurements) over much of the observed energy range. The ATIC and PPB-BETS [535] (Balloon-borne Electron Telescope with Scintillating fibers, BETS [538], aboard the Polar Patrol Balloon, PPB [539]) spectra, however, exhibit an excess between  $\sim 300$  and  $800$  GeV that cannot be explained by conventional sources. The significance of the “bump” in the ATIC data is  $4\text{--}6\sigma$  (above the GALPROP spectrum), depending on the statistical method employed.

Annihilating Kaluza-Klein dark matter [536] in the Milky Way’s dark halo has been proposed as the source of the ATIC bump [531]. In this scenario, KK particles with  $\sim 600$  GeV/ $c^2$  masses annihilate directly into electron-positron pairs, yielding a delta function (centered at the mass of the KK particle) of initial electron and positron energies. Subsequent transport through the Milky Way smears the delta function to lower energies, while preserving a sharp drop in flux at  $\gtrsim 600$  GeV [536]. The total  $e^\pm$  flux expected for this model is illustrated in Figure 3.4. Despite the boost factor of about 200 required to fit the ATIC data, the KK scenario is slightly easier to manage than a neutralino explanation. The sharp drop at approximately the mass of the KK particle is a difficult feature to reproduce with neutralino annihilations. Annihilation directly into electron-positron pairs for such massive neutralinos is heavily suppressed, with production of top-quark, weak-boson, and Higgs-boson pairs dominant. Subsequent decay to positrons and electrons results in a much broader energy spectrum, because some of the annihilation energy is carried away by other decay products. However, the same clumpy neutralino model put forth as a possible explanation of the PAMELA excess (see Figure 3.3) also appears capable of explaining the ATIC data (see, *e.g.*, Figure 2 in [520]).

### 3.2. INDIRECT DETECTION

#### The Fermi-LAT Electron-Positron Flux

The Large Area Telescope (LAT) is the primary instrument aboard the Fermi Gamma-Ray Space Telescope (Fermi) satellite [509]. Although the LAT is in many respects a multi-purpose instrument, it is tasked specifically with measuring the flux of photons (arriving at the Earth from the Universe at large) with energies between 20 MeV and 300 GeV. Prior to the first Fermi-LAT data release [540], the spectrum of photons with energies in excess of 10 GeV was largely unexplored. The extraordinary accuracy of the Fermi-LAT gamma-ray data should lead to an improved understanding of a number of astrophysical phenomena, some of which are: 1) unidentified gamma-ray sources and the diffuse gamma-ray background; 2) particle acceleration near active galactic nuclei, pulsars, supernovae remnants and the Sun; 3) gamma-ray bursts and transients; and 4) gamma rays from dark-matter annihilations. The latter will be discussed further in the next section. The design of the LAT apparatus also gives it unprecedented sensitivity to the high-energy  $e^\pm$  cosmic-ray flux. The LAT detection technique and its first measurement of the  $e^\pm$  flux are the focus of the remainder of this section. Comparison with previous observations and models of electron-positron production will be discussed as well. Despite qualitative spectral differences, a common pulsar-driven explanation of the PAMELA positron excess, the ATIC  $e^\pm$  bump, and the Fermi-LAT  $e^\pm$  spectrum seems likely.

The LAT is what is known as a pair-conversion telescope. The high-energy portion of the electromagnetic spectrum cannot be reflected or refracted as in a conventional telescope. Instead, high-energy photons are converted to electron-positron pairs by a high- $Z$  target material. The charged particles resulting from pair conversion can then be tracked in order to reconstruct the incoming direction of the original gamma ray. The LAT employs 16 layers of tungsten foil to induce pair conversion. Their summed thickness corresponds to just over a single radiation length for gamma rays in the energy range of interest, ensuring reasonable pair-conversion efficiency. Spreading the target thickness over several thin layers prevents  $e^\pm$  multiple scattering from obscuring the be-



### CHAPTER 3. DARK MATTER DETECTION

gining of the charged tracks. Charged-particle tracking is accomplished with  $xy$ -position sensitive layers of silicon strip detectors following each tungsten layer. The layers of tungsten and silicon are followed by a segmented calorimeter constructed from scintillating CsI crystals. The charged pair-conversion products induce particle cascades in the calorimeter, allowing measurement of the original gamma ray's energy. Similar to the ATIC and PAMELA experiments, electromagnetic and hadronic events are differentiated by their topological signatures in the calorimeter. The entire apparatus is surrounded by 89 plastic-scintillator tiles that act as a charged particle veto. Advanced electronics allow the LAT to run in a scanning mode, where event triggers are formed from real-time information from the silicon layers. Further details regarding the LAT instrumentation and science objectives can be found in [509].

It is easy to see how the LAT apparatus is also sensitive to the cosmic-ray electron-positron flux. An incident electron or positron passing through the silicon layers will trigger a readout of the calorimeter, while a corresponding signal in the plastic-scintillators allows incident gamma rays to be distinguished from charged cosmic rays. Hadronic events due to protons, antiprotons and heavier nuclei are distinguished by the broader profiles of the showers they induce in the calorimeter. The LAT's energy response to gamma rays translates to a sensitivity to electrons and positrons with energies between  $\sim 20$  GeV and 1 TeV. During the second half of 2008, the Fermi-LAT recorded more than four million  $e^\pm$  events in this energy range. The resulting total  $e^\pm$  flux is compared with previous measurements in Figure 3.5. Unlike the ATIC and PPB-BETS spectra, the high-energy portion of the Fermi-LAT spectrum is featureless. However, because the Fermi-LAT spectrum is considerably harder than expected from conventional diffuse  $e^\pm$  production [541], it is statistically consistent with previous measurements.

The hardness of the Fermi-LAT  $e^\pm$  spectrum is supported by even higher-energy data recored by the High Energy Stereoscopic System (HESS) experiment [549]. The HESS experiment is a ground-based array capable of imaging the Cherenkov radiation produced when high-energy gamma rays interact in the Earth's atmosphere. A multi-national effort is underway in which this kind

### 3.2. INDIRECT DETECTION

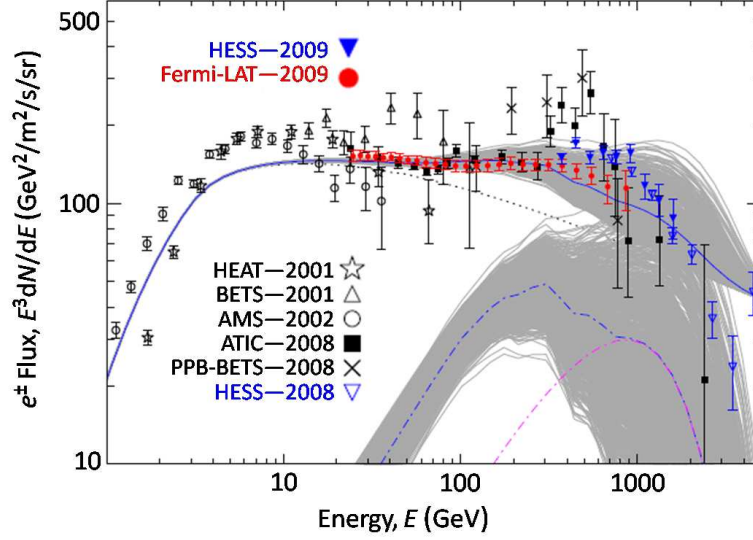


Figure 3.5: Compilation of the most recent measurements of the total electron-positron cosmic-ray flux (scaled by  $E^3$ ) as a function of energy, compared with the expected contribution from conventional  $e^\pm$  production and transport in the Milky Way (black, dotted line) as simulated using GALPROP [541]. References for the labeled measurements are (from top to bottom, respectively) [542, 543, 533, 534, 544, 531, 545, 546]. The AMS and PPB-BETS spectra are updated versions of those shown in Figure 3.4. The expected  $e^\pm$  contribution from pulsars in the ATNF catalog [547] for a representative choice of model parameters (blue/dark dash-dotted line) compares favorably with the measured spectra when added to the expected conventional contribution (blue, solid line), with the contribution from the Monogem pulsar (magenta/light dash-dotted line) indicated for that particular case. Varying the pulsar-model parameters (gray/light solid lines) results in an uncertainty, indicated for both the pulsar-only contribution and the total flux, that might explain the unexpected hardness of the Fermi-LAT and HESS spectra. Figure taken from [548].

of imaging atmospheric Cherenkov telescope is the primary tool for studying gamma-ray astronomy at very high energies ( $>100$  GeV). For further information I refer the interested reader to the reviews in [550] and [551]. Like the LAT apparatus, the HESS array also has sensitivity to high-energy cosmic-ray elec-

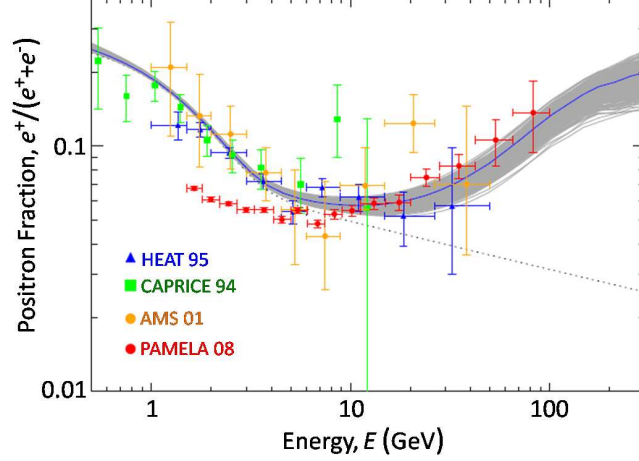


Figure 3.6: Select measurements of the cosmic-ray positron fraction as a function of energy, compared with the expected contribution from conventional positron production and transport through the Milky Way (black, dotted line) as simulated using GALPROP [541]. When the positron contribution expected from all pulsars in the ATNF catalog [547] (for the same representative choice of model parameters as in Figure 3.5) is added to the conventional contribution, the result (blue/dark solid line) compares favorably with the positron excess observed above  $\sim 10$  GeV. Below  $\sim 5$  GeV, the discrepancy between the model and the PAMELA [519] data is due to the inclusion of the pre-2000 solar-modulation effect. The spectra from the HEAT [513], CAPRICE [514] and AMS [544] experiments are derived from data recorded prior to 2000. While the PAMELA, HEAT and CAPRICE spectra are the same as in Figure 3.3, the AMS data have been rebinned. As in Figure 3.5, variation of the pulsar model's parameters results in an uncertainty region (gray/light solid lines). Figure taken from [548].

trons and positrons, but from  $\sim 300$  GeV to 5 TeV. Contrary to the spectrum expected for conventionally produced electrons and positrons, the  $E^3$ -scaled HESS  $e^\pm$  spectrum agrees with (but is even harder than) the Fermi-LAT spectrum, remaining flat out to roughly 1 TeV (see Figure 3.5).

As discussed in the context of the PAMELA positron excess, the magnetospheres of nearby pulsars are probable sources of high-energy positrons and

### 3.2. INDIRECT DETECTION

electrons. The authors in [548] identify two pulsars in particular from the Australia Telescope National Facility (ATNF) radio-pulsar catalog [547] that appear to contribute significantly to the high-energy  $e^\pm$  flux reaching the Earth: the Monogem [547] and Geminga [552] pulsars. The Monogem pulsar’s modeled contribution is indicated in Figure 3.5 (magenta/light dash-dotted line), as is the (approximate) summed contribution of both pulsars (blue/dark dash-dotted line). When the latter is added to the featureless spectrum of conventionally produced electrons and positrons, the total reproduces the hardness of Fermi-LAT spectrum. High-energy positrons from these same two pulsars have been proposed independently as an explanation of the PAMELA positron excess [553]. The pulsar plus conventional production model shown in Figure 3.5 is also shown in terms of the positron fraction in Figure 3.6, and compares favorably with the PAMELA data. A common explanation in terms of conventional plus pulsar-driven electron and positron production works well for the Fermi-LAT and PAMELA spectra.

The agreement between all experimental data is improved with a more complete model that includes contributions from all pulsars in the ATNF catalog that meet certain criteria. Only nearby, “mature” pulsars are expected to contribute to the high-energy  $e^\pm$  flux. When electrons and positrons are created in a pulsar’s magnetosphere, it takes many thousands of years for them to escape and become part of the interstellar cosmic-ray flux.  $\sim 150$  pulsars from the ATNF catalog are old enough ( $> 5 \times 10^4$  yr) and close enough ( $< 3$  kpc distant) to contribute, with dominant contributions from Monogem and Geminga. The two-pulsar model discussed above included a choice of model parameters that provides the best fit to the Fermi-LAT and PAMELA spectra. That particular choice was intended only as a representative model. The uncertainty bands (gray/light solid lines) shown in Figures 3.5 and 3.6 are the result of randomly varying the free parameters within reasonable (motivated) limits (see [548] for further details) for the many-pulsar model. All of the high-energy  $e^\pm$  flux and positron-fraction measurements are statistically consistent with these bands. Furthermore, the pulsar-model uncertainties are almost certainly greater than depicted in Figures 3.5 and 3.6. Because pulsars are incredibly difficult to de-

tect unless their beams happen to cross the Earth’s path, it is likely that not all nearby pulsars are accounted for in the ATNF catalog. A significant population of radio-quiet gamma-ray pulsars was recently discovered [554] that is also absent from the ATNF catalog. In any case, pulsar (plus conventional) production of electrons and positrons appears to simultaneously explain the PAMELA, ATIC, PPB-BETS, Fermi-LAT and HESS high-energy spectra. An annihilating dark-matter interpretation is still possible, but is highly constrained (as discussed in [548]) to models with (artificially) boosted densities of relatively heavy ( $\sim 0.5 \text{ TeV}/c^2$  to a few  $\text{TeV}/c^2$ ) dark-matter particles.

### 3.2.2 Gamma Rays

Similar to the  $e^\pm$  searches, the hunt for an indirect dark-matter signal via a gamma-ray signature has left a colorful trail through the literature over the past decade. Prior to the first Fermi-LAT data release, the most sensitive probe of the high-energy gamma-ray flux was the Energetic Gamma-Ray Experimental Telescope’s [555] (EGRET) measurement of diffuse Galactic gamma rays with energies between  $\sim 100 \text{ MeV}$  and  $10 \text{ GeV}$  [556]. As illustrated in the left panel of Figure 3.7, when compared to a GALPROP simulation of conventional gamma-ray production (labeled background), the EGRET spectrum exhibits an excess for energies  $\gtrsim 1 \text{ GeV}$ . The background spectrum of high-energy gamma rays arriving at the Earth is the result of cosmic-ray interactions (primarily protons and electrons) with the Milky Way’s interstellar gas and radiation field [557, 558]. As indicated in Figure 3.7, inelastic  $\pi^0$  production with subsequent decay to photons dominates the energy range of interest. When combined with photons from bremsstrahlung and inverse Compton scattering, an isotropic flux of gamma rays is expected that falls smoothly and rapidly for gamma-ray energies  $\gtrsim 1 \text{ GeV}$ . The EGRET data are in clear disagreement with the conventional model. Augmenting the conventional GALPROP spectrum with gamma-rays from annihilations of  $\sim 60 \text{ GeV}/c^2$  WIMPs reproduces the EGRET high-energy data provided the WIMP density is boosted by a factor of  $\sim 100$  [559]. While the EGRET excess appears to possess the req-

### 3.2. INDIRECT DETECTION

quisite features of an indirect dark-matter signal, some have argued that the implied rate of annihilations should be accompanied by a flux of cosmic-ray antiprotons that is not observed [560]. Counterarguments also exist in which the conventional (isotropic) model for the propagation of cosmic-ray antiprotons is modified such that the expected flux of annihilation antiprotons escapes detection via anisotropic deflection away from the Milky Way’s luminous disc [561, 562].

The first Fermi-LAT measurement of the diffuse Galactic gamma-ray flux has essentially ruled out the possibility that dark-matter annihilations are responsible for the EGRET excess [540]. The Fermi-LAT spectrum is shown in the right panel of Figure 3.7. Despite a slight excess of gamma rays with  $\sim$ GeV energies, the Fermi-LAT spectrum is consistent with an *a priori* conventional model of diffuse gamma-ray production in the Milky Way. The absolute scale for the Fermi-LAT spectrum is substantially reduced compared to the EGRET flux because only a limited range of Galactic latitudes is included in the former. In light of the more accurately measured Fermi-LAT spectrum, an explanation of the EGRET excess in terms of instrumental effects [563, 564] or unresolved point sources [565] seems likely. Furthermore, the consistency demonstrated in Figure 3.7, between the Fermi-LAT diffuse Galactic gamma-ray flux and the conventional model, severely constrains neutralino-annihilation models capable of producing the high-energy positron and/or electron fluxes observed by other experiments [566] (and displayed in Figures 3.5 and 3.6).

In the absence of an indirect dark-matter signal, the Fermi-LAT diffuse Galactic spectrum becomes an important measurement of the Milky Way’s “foreground” gamma-ray emission. With the Fermi-LAT’s ability to reconstruct a gamma ray’s arrival direction, extragalactic objects suspected of harboring large dark-matter densities can be specifically targeted. However, diffuse production in the Milky Way is the dominant source of gamma rays arriving at the Earth. The expected flux of gamma rays from dark-matter annihilations at the center of a cluster or satellite galaxy, for example, might be an order of magnitude (or more) lower than the diffuse foreground flux. Fitting for and subtracting a well understood foreground spectrum generally improves sensitiv-

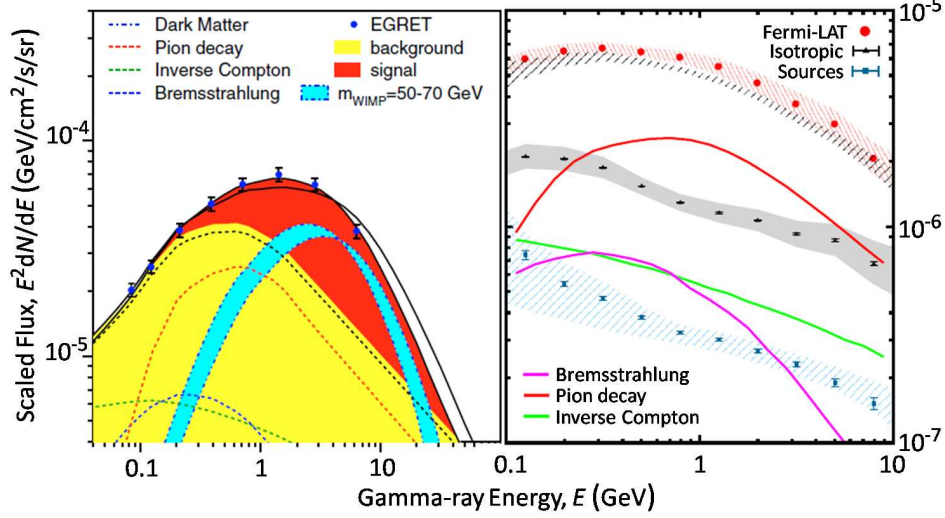


Figure 3.7: *Left*: The EGRET  $E^2$ -scaled diffuse Galactic gamma-ray flux as a function of energy (blue dots), compared to a GALPROP simulation of conventional gamma-ray production (yellow shaded). Dashed lines indicate the dominant conventional photon production channels: inelastic pion production and decay (red), bremsstrahlung (blue), inverse Compton scattering (green), and the sum of the three (black). When an additional component from annihilations of a 50–70  $\text{GeV}/c^2$  WIMP (blue shaded region bounded by dash-dotted lines) is added to the background model, the total gamma-ray flux (black solid lines) agrees well with the EGRET data. The 50  $\text{GeV}/c^2$  WIMP model predicts a larger flux at 1 GeV (and is a better fit) than the 70  $\text{GeV}/c^2$  WIMP model. *Right*: The Fermi-LAT diffuse Galactic gamma-ray spectrum (red dots and hashed uncertainty band) for a limited range of Galactic latitudes is well explained by a model of conventional photon production (black hashed region). The model’s dominant production channels are indicated (colored solid lines). The Fermi-LAT spectrum from gamma-ray point sources (blue squares and hashed uncertainty band) and isotropic extragalactic flux (black triangles and gray/solid uncertainty band) are also shown. The latter includes unresolved Galactic and extragalactic point sources. Figure adapted from [559] and [540].

ity to potential extragalactic (or simply nondiffuse) annihilation signals. Since the publication of the first Fermi-LAT foreground spectrum, indirect searches

### 3.2. INDIRECT DETECTION

for dark-matter via an extragalactic (or nondiffuse) gamma-ray signature have proliferated considerably in the literature. No extragalactic dark-matter annihilation signals have been discovered thus far. Several recent results are summarized below:

- Dwarf spheroidal galaxies are among the most dark-matter dominated structures observed in the Universe, with mass-to-light ratios that exceed 1000 in some cases (see, *e.g.*, Table 1 in [567]). Nearby dwarf spheroidals therefore make ideal targets to search for an indirect dark-matter signal. The LAT has measured the flux of gamma rays in the direction of several of the Milky Way’s dwarf spheroidal satellite galaxies, and no significant gamma-ray emission (not associated with foreground sources) has been observed. The null results are used to constrain the pair-annihilation cross section of WIMPs under a variety of popular scenarios [567, 568].
- Clusters and groups are potentially bright sources of pair-annihilation gamma rays. After nearly a year of Fermi-LAT observations, no excess gamma-ray emission has been observed from the direction of any extragalactic galaxy systems. These null results have been used to derive (some of the most) stringent constraints on pair-annihilation models proposed to (simultaneously) explain the PAMELA positron excess and the hardness of the Fermi-LAT  $e^\pm$  spectrum [569, 570].
- The extragalactic diffuse spectrum of gamma rays arriving at the Earth is expected to be sensitive to dark-matter annihilations in the Universe at large. The first attempt to unravel the (subdominant) extragalactic diffuse flux from the Fermi-LAT data is shown in Figure 3.7 (black triangles with solid uncertainty band). A more detailed analysis is presented in [571] (see, *e.g.*, Figure 3 in [571]), where the extragalactic diffuse spectrum is found to be consistent with a power law and considerably softer than similar attempts to extract an equivalent spectrum from the EGRET data [572, 573]. The featureless profile of the Fermi-LAT extragalactic diffuse gamma-ray spectrum has also been used to constrain a variety of dark-matter annihilation scenarios [574, 575].



## CHAPTER 3. DARK MATTER DETECTION

- Simulations of CDM galaxy formation appear to naturally produce a number of dark-matter subhalos. Some of the Milky Way's subhalos might still orbit the Milky Way as dark-matter satellites (with no visible components). If sufficiently massive and nearby, these subhalos might appear as gamma-ray point sources in the Fermi-LAT data (due to dark-matter annihilations), with no associated emission at other wavelengths. The authors in [576] have identified several dozen Fermi-LAT point sources that are roughly consistent with the expected properties of dark-matter subhalos. However, the information available is currently insufficient for any to be clearly identified, leading to constraints on the dark-matter annihilation cross section under certain scenarios.

### The Galactic Center & Light Neutralinos

The center of the Milky Way is perhaps the most obvious target for an indirect dark-matter search. Although the spatial distribution of Galactic dark matter is not known precisely, a universal property of dark-halo models is a maximal density at the Galactic center (see, *e.g.*, Equations 2.2.6, 2.2.22, and 2.2.36). A highly general parameterization models the halo density as

$$\rho(r) = \frac{\rho_0}{(r/a_r)^\gamma [1 + (r/a_r)^\delta]^{(\epsilon-\gamma)/\delta}}, \quad (3.2.1)$$

where  $a_r$  is the halo-core radius, and  $\delta$ ,  $\epsilon$  and  $\gamma$  are (nonnegative) free parameters. Note that this is simply a generalization of the halos described by Equations 2.2.6 ( $\delta = \epsilon = 2$  and  $\gamma = 0$ ) and 2.2.22 ( $\delta = \gamma = 1$  and  $\epsilon = 3$ ). The  $\gamma$  parameter is the most relevant to a potential annihilation signal from the Galactic center, as it controls the inner slope of the halo profile. A nonzero value of  $\gamma$  means the halo is cuspy, leading to a core density that is technically infinite. For  $\gamma = 1$  and an annihilation cross section of  $\sim 10^{-27} \text{ cm}^3/\text{s}$  (see Section 2.4.3), the density at the core of the Milky Way is sufficient to yield an expected flux of hundreds of gamma rays per year (arriving at the Fermi satellite) due to dark-matter annihilations [577].

### 3.2. INDIRECT DETECTION

The main challenge is to disentangle this annihilation signal from gamma rays originating from other astrophysical sources. In addition to the diffuse foreground emission, any gamma-ray point sources that are coincident on the sky with the Galactic center can cause confusion. The gamma-ray spectrum for a particularly bright point source in the direction of the Galactic center, for example, has been measured by a number of ground-based atmospheric Cherenkov telescopes [578, 579]. Fortunately, the spatial distribution associated with the dark-matter density described by Equation 3.2.1 translates to an angular distribution—degrees away from the Galactic center along the Milky Way’s luminous disc—of gamma rays that can be used to differentiate an annihilation signal from the astrophysical backgrounds (or foregrounds). In particular, the angular distribution associated with diffuse foreground emission should be less concentrated at the Galactic center than the annihilation signal expected from a cuspy dark halo, while the latter will tend to be more extended than for gamma-ray point sources.

This is precisely the approach taken by Hooper and Goodenough in [577] (and first introduced in [580]). They focus on gamma rays from the inner  $10^\circ$  of the Galactic center that were recorded by the Fermi-LAT during its first two years of operation. Based on the flux from the outer  $2\text{--}10^\circ$ , they deduce three distinct classes of background and foreground gamma-ray emission: 1) diffuse emission along the luminous disc; 2) spherically symmetric, diffuse emission associated with the Milky Way’s luminous bulge; and 3) emission from point sources. Similar to the spectra shown in Figure 3.7, the diffuse components—dominated by gamma rays from neutral pion decay, inverse Compton scattering, and bremsstrahlung—exhibit spectral shapes well described by GALPROP simulations of conventional gamma-ray production and transport. Gamma-ray emission from 69 point sources are modeled according to information in the Fermi First Source Catalog [581]. As illustrated in the upper-left panel of Figure 3.8, the gamma-ray flux from the outer region of the Galactic center is well described by the conventional model. In contrast, as is clear in the lower-left panel of Figure 3.8, the innermost  $\sim 1^\circ$  of the Galactic center exhibits a significant excess (relative to the background model) of  $\sim 1\text{--}5$  GeV gamma rays. This

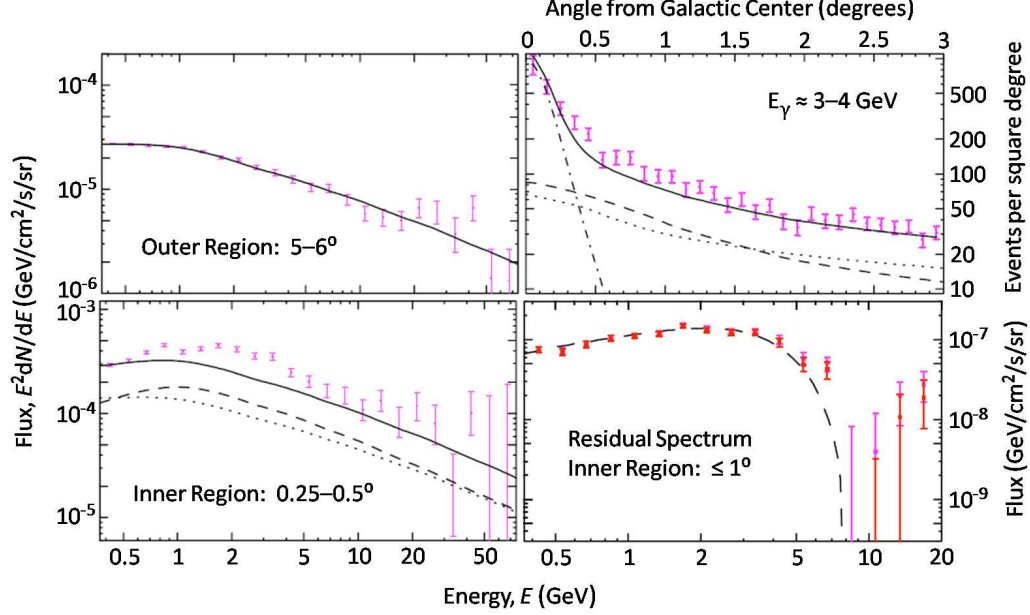


Figure 3.8: *Upper left*:  $E^2$ -scaled spectrum of gamma-ray emission (magenta error bars) from the Galactic center for a 5–6° angular bin, compared to a conventional model of gamma-ray production (solid). *Lower left*: Gamma-ray spectrum (magenta error bars) for a 0.25–0.5° angular bin, compared to the total background model (solid) and its disc (dashed) and bulge (dotted) components. There is a clear excess for gamma rays with energies between  $\sim 1$ –5 GeV. *Upper right*: Number of gamma rays with energies between  $\sim 3$  and 4 GeV (magenta error bars) as a function of angular distance from the Galactic center, compared to a model (solid) that includes background components due to diffuse gamma-ray emission in the disc (dashed) and bulge (dotted), as well as emission due to annihilating WIMPs (dash dotted). *Lower right*: The residual gamma-ray flux from the innermost region of the Galactic center (error bars), compared to the spectrum of gamma rays expected for  $\sim 9$  GeV/ $c^2$  WIMP annihilations (dashed). The red (magenta) error bars include (do not include) a correction based on HESS observations. Figure adapted from [577].

excess has the same spherical symmetry that characterizes the second background component listed above. However, as shown in the upper-right panel of

### 3.2. INDIRECT DETECTION

Figure 3.8, the angular distribution of the excess is far more concentrated at the Galactic center than either of the diffuse background components. The excess gamma-ray flux from the innermost region appears consistent with emission from a spherically symmetric component distributed as in Equation 3.2.1 with  $\gamma \simeq 1.2\text{--}1.3$ . Subtracting the combined background model from the spectrum of gamma-ray emission from the innermost degree of the Galactic center results in the residual spectrum shown in the lower-right panel of Figure 3.8. As indicated, a model in which  $7\text{--}10\text{ GeV}/c^2$  WIMPs annihilate (primarily) into tau and  $b$ -quark pairs (with subsequent decay to photons) reproduces the residual spectrum. While the agreement between the residual spectrum and the dark-matter annihilation model is quite good, it is worth noting that the Galactic center is a chaotic environment. In particular, emission from the Milky Way’s supermassive black hole is not fully understood, and could partially explain the excess from the innermost region observed in the Fermi-LAT data [577]. Nevertheless, in light of the direct evidence for a light WIMP that will be discussed in the next section, the indirect evidence discussed here helps to further motivate the targeted search for low-mass WIMPs that is the focus of this thesis.

### 3.3 Direct Detection

The evidence presented in the previous chapter suggests that the luminous portions of the Milky Way are embedded in a halo of dark-matter particles. As the Earth rotates with the Sun about the center of the Galaxy, it passes through an apparent “wind” of dark matter. If WIMPs constitute this dark-matter wind, they will scatter very rarely off the nuclei in terrestrial matter. First proposed by Goodman and Witten in [582], dark-matter “direct detection” refers to the prospect of observing (and measuring) the kinetic energy, in the keV range, of recoiling nuclei following a WIMP-nucleus scattering event. The primary challenge is to distinguish the low (expected) WIMP-interaction rate from an overwhelming background rate due to interactions of known forms of radiation. Dark-matter direct detection has grown into a global effort in which a variety of technologies are employed to discriminate against (and moderate) these backgrounds. In this section I review the standard methods for predicting WIMP-interaction rates and deposited-energy spectra, as well as some of the methods used to distinguish the expected WIMP signal from backgrounds. Evidence for a light WIMP based on interpretations of the DAMA/LIBRA [583] and CoGeNT [584] experiments’ data is also presented.

#### 3.3.1 Approximate Deposited Energy

Before delving into detailed interaction-rate calculations, it is instructive to make a rough estimate of what to expect in terms of deposited energy. The principle factors determining the amount of energy deposited by a WIMP are its velocity and mass, and the mass of the target nucleus. If dark-matter particles are to remain gravitationally bound to the Milky Way, they must have halo velocities that do not exceed the Milky Way’s escape velocity. The best measurement of the escape velocity was presented in Section 2.2.1 (see Equations 2.2.19 and 2.2.20), and limits the velocity of a halo object to  $\lesssim 600$  km/s. On average, however, WIMPs in the solar vicinity are expected to have velocities (relative to the Earth) of  $\sim 200$  km/s, or  $v/c \simeq 10^{-3}$  [374]. WIMP

### 3.3. DIRECT DETECTION

interactions with terrestrial matter are therefore nonrelativistic, and can be characterized by a de Broglie wavelength of

$$\lambda = \frac{2\pi\hbar}{m_\chi v} \sqrt{1 - (v/c)^2} \simeq \frac{2\pi\hbar}{m_\chi v}, \quad (3.3.1)$$

where  $m_\chi$  is the WIMP mass. A  $10 \text{ GeV}/c^2$  WIMP in the solar neighborhood, for example, has a de Broglie wavelength of  $\sim 2 \times 10^{-13} \text{ m}$ . The nuclear radius for an atom of terrestrial matter can be approximated by

$$r_{\text{nucleus}} \simeq A^{1/3} (1.2 \times 10^{-15} \text{ m}), \quad (3.3.2)$$

where  $A$  is the nucleon number [585]. The largest nuclear diameters are therefore  $\lesssim 1.5 \times 10^{-14} \text{ m}$ , significantly smaller than the WIMP de Broglie wavelength; a WIMP in the Milky Way's halo (generally) has too little momentum to probe nuclear structure. Incapable of resolving individual nucleons, WIMPs will scatter coherently from the entire nucleus. Following a scattering event, the energy deposited by a WIMP is therefore equal to the kinetic energy of the recoiling nucleus. WIMP-nucleus scattering is similar to the elastic collisions between billiard balls. If the mass of the WIMP and target nucleus are nearly equal and they collide head on, the target nucleus can acquire the WIMP's full velocity, and

$$E_{\text{recoil}} \simeq \frac{1}{2} m_{\text{nucleus}} v^2 \simeq \frac{A}{2} \text{ keV}. \quad (3.3.3)$$

Detector materials currently used in direct-detection experiments include nuclei as light as helium [586] ( $A \simeq 4$ ) and as heavy as tungsten [587] ( $A \simeq 184$ ). The approximate range of expected energy depositions is therefore (at most) tens of keV, a range easily accessible to modern particle detectors. Note that to arrive at this result I have intentionally glossed over a number of the finer details (discussed below). In the spirit of an order-of-magnitude estimate, however, Equation 3.3.3 is not far wrong for a wide range of WIMP masses and target materials.

### 3.3.2 Scattering Rates & Energy Spectra

The current goal of direct-detection experiments is to measure a nonzero rate of WIMP interactions in a terrestrial detector. An incontrovertible WIMP signal has yet to be observed. In the absence of a positive detection, the null results of different experiments are usually compared in terms of upper limits on the WIMP-nucleon cross section (see, *e.g.*, Figures 2.18 and 2.19). However, the true observable is the interaction rate with a particular detector material, which depends on the material’s nuclear structure and several loosely constrained properties of the Milky Way’s dark halo. Intercomparison of different experiments necessitates a standard framework built around a standard set of assumptions. The most commonly used standard framework is described below, and (generally) follows Lewin and Smith’s concise summary in [588].

#### The Standard Halo

Nontrivial uncertainties arise in the expected WIMP-nucleus scattering rate due to uncertainties in the distribution of dark matter in the Milky Way. However, these uncertainties are shared by all direct-detection experiments, and (for most WIMP masses) have a common-mode effect on derived cross-section limits. Consequently, the convention is to assume a standard dark-halo model when intercomparing direct-detection results, the so-called “spherical-cow,” isothermal halo. Unless otherwise stated, the standard halo model described here will be used throughout the remainder of this thesis.<sup>1</sup> For information on how alternative halo models affect detection rates, I refer the interested reader to [589].

A variety of halo models (in the form of density profiles) have been discussed in this thesis. The concept of a spherically symmetric isothermal sphere was first introduced in Section 2.2.1 as a potential explanation for the flatness of spiral-galaxy rotation curves, and again in Section 2.2.2 to demonstrate the  $\sqrt{2}$

---

<sup>1</sup>The Galactic escape velocity is one of the Milky Way’s least well known parameters that is relevant to the low-mass WIMP sensitivity of a direct-detection experiment. The direct-detection uncertainty on the WIMP-nucleon cross-section that results from varying the escape velocity will be explored in Chapter 6.

### 3.3. DIRECT DETECTION

relationship between circular velocity and velocity dispersion. Equation 2.2.6 gives the density for a softened isothermal sphere with core radius  $a_r$ , Equation 2.2.22 for the cuspy NFW halo, Equation 2.2.36 for a Burkert halo, and Equations 2.3.14 and 3.2.1 for generalized, spherically symmetric halos. This landscape of models highlights the uncertainty associated with the spatial distribution of dark matter in gravitationally bound systems, and can be a bit confusing. Nearly all of them, however, reduce to a  $\rho \propto 1/r^2$  or  $1/r^3$  behavior in the large- $r$  limit. The proliferation of halo models in the literature primarily reflects uncertainty in core densities (addressing, *e.g.*, whether or not halos are cuspy), and in the density profiles inferred from different astrophysical data (addressing, *e.g.*, whether or not a universal profile exists that describes galaxies and clusters equally well). Direct-detection experiments are sensitive to the local value of the halo density,  $\rho_0 \equiv \rho(r_0 \simeq 8 \text{ kpc})$ , but not to its radial dependence.<sup>2</sup> Since the Earth's position is well separated from the core of the Milky Way, an isothermal sphere is actually a reasonable representation (from a direct-detection point of view) for any of the halo models mentioned thus far.

In general, a halo of WIMPs can be described by its differential number density:

$$dn_\chi \propto f(\vec{v}, \vec{x}) d^3v d^3x, \quad (3.3.4)$$

where  $f(\vec{v}, \vec{x})$  is the full phase-space WIMP distribution, specified by three velocity ( $\vec{v}$ ) and three spatial ( $\vec{x}$ ) coordinates. For an isothermal sphere (characterized by temperature  $T$ ) of WIMPs with mass  $m_\chi$ ,  $f(\vec{v}, \vec{x})$  is given by the Boltzmann distribution:

$$\begin{aligned} f(v, r) \propto e^{-\mathcal{E}/kT} &= e^{-(m_\chi v^2/2 + m_\chi \Phi(r))/kT} \\ &= e^{-(v^2 + 2\phi(r))/2\sigma^2}, \end{aligned} \quad (3.3.5)$$

where  $\Phi(r)$  is the gravitational potential per unit mass,  $v \equiv |\vec{v}|$ ,  $r \equiv |\vec{x}|$ , and  $\sigma = \sqrt{kT/m_\chi}$  is the one-dimensional velocity dispersion (as in Section 2.2.2). Since the Earth's radial position relative to this distribution is (effectively)

---

<sup>2</sup>Note that this definition of  $\rho_0$  differs from the central halo-density normalization constant first introduced in Equation 2.2.5. The convention for the direct-detection standard halo is to define  $\rho_0$  at the solar radius rather than at the Galactic center.



### CHAPTER 3. DARK MATTER DETECTION

fixed, the gravitational-potential factor can be absorbed into a normalization constant. The differential number density relevant to direct detection is therefore a Maxwellian velocity distribution, commonly written as:

$$dn_\chi = \frac{n_0}{k_f} f(\vec{v}, \vec{v}_\oplus) d^3v = \frac{n_0}{k_f} e^{-(\vec{v} + \vec{v}_\oplus)^2 / 2\sigma^2} d^3v, \quad (3.3.6)$$

where  $\vec{v}$  is the WIMP velocity relative to the Earth,  $\vec{v}_\oplus$  is the Earth's velocity relative to the WIMP distribution,

$$k_f = \int_{|\vec{v} + \vec{v}_\oplus| \leq v_{\text{esc}}} f(\vec{v}, \vec{v}_\oplus) d^3v \quad (3.3.7)$$

is a normalization constant, and

$$n_0 = \int_0^{v_{\text{esc}}} dn_\chi \quad (3.3.8)$$

is the mean WIMP number density. Note that all velocity integrals are truncated by the Galactic escape velocity. It can be shown that  $k_f$  is independent of  $\vec{v}_\oplus$ , and is given by (see, *e.g.*, Appendix A in [588])

$$k_f = (2\pi\sigma^2)^{3/2} \left( \text{erf}\left(\frac{v_{\text{esc}}}{\sqrt{2}\sigma}\right) - \sqrt{\frac{2}{\pi}} \frac{v_{\text{esc}}}{\sigma} e^{-v_{\text{esc}}^2/2\sigma^2} \right), \quad (3.3.9)$$

where

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt. \quad (3.3.10)$$

In the  $v_{\text{esc}} \rightarrow \infty$  limit,  $\text{erf}(v_{\text{esc}}/\sqrt{2}\sigma) \rightarrow 1$  and  $k_f \rightarrow k_0 = (2\pi\sigma^2)^{3/2}$ .

Under the standard isothermal-halo assumption, the local distribution of WIMPs is fully specified by Equations 3.3.6 and 3.3.9 up to four parameters: the Galactic escape velocity  $v_{\text{esc}}$ , the WIMP velocity dispersion  $\sigma$ , the velocity of the Earth  $\vec{v}_\oplus$ , and the local WIMP number density  $n_0$ . The escape velocity preferred by the direct-detection community has varied somewhat over the past two decades. While Lewin and Smith use  $v_{\text{esc}} = 600$  km/s [588], a value of 650 km/s is also common in the literature [374]. The best-fit value from the study of high-velocity halo stars in [183] (see Section 2.2.1) has been used to

### 3.3. DIRECT DETECTION

quote several recent direct-detection results, and is likely to become the new standard. Consequently, unless otherwise stated, I will use  $v_{\text{esc}} = 544 \text{ km/s}$ . In any case, for a wide range of experimental conditions and WIMP masses, detection rates are not very sensitive to the exact value of the escape velocity because the most probable WIMP velocity is substantially smaller. From Equation 3.3.6, the most probable WIMP velocity is  $\sqrt{2}\sigma$ , which for an isothermal sphere is equal to the asymptotic value of the circular velocity (see Equation 2.2.28). The standard value for the circular velocity is  $v_{\odot} = 220 \text{ km/s}$  [184], such that  $\sigma = v_{\odot}/\sqrt{2} \simeq 156 \text{ km/s}$ . To avoid confusing the velocity dispersion  $\sigma$  with the cross-section symbols I will use later in this section, all formulae will be quoted in terms of  $v_{\odot}$  from here on. Note that the term *velocity dispersion* is used somewhat loosely in the literature, typically referring to one of three things: the one-dimensional (line-of-sight) rms velocity  $\sigma \simeq 156 \text{ km/s}$ , the three-dimensional rms velocity  $v_{\text{rms}} = \sqrt{3/2}v_{\odot} \simeq 270 \text{ km/s}$ , and the most probable velocity  $v_{\odot} \simeq 220 \text{ km/s}$ .

The velocity of the Earth relative to the nonrotating reference frame of the WIMP distribution can be split into three parts:

$$\vec{v}_{\oplus} = \vec{v}_{\odot} + \vec{u}_{\odot} + \vec{u}_{\oplus}, \quad (3.3.11)$$

where  $\vec{v}_{\odot}$  is the velocity of the Galactic disc,  $\vec{u}_{\odot}$  is the velocity of the Sun relative to the disc, and  $\vec{u}_{\oplus}$  is the velocity of the Earth relative to the Sun. In Galactic coordinates,<sup>3</sup>

$$\vec{v}_{\odot} = (0, 220, 0) \text{ km/s} \quad (3.3.12)$$

is the circular velocity (same as above) common to all stars in the solar neighborhood, and

$$\vec{u}_{\odot} = (9, 12, 7) \text{ km/s} \quad (3.3.13)$$

represents the so-called “proper motion” of the Sun relative to disc [590]. The full expression for  $\vec{u}_{\oplus}$  is a bit complicated (see, *e.g.*, Appendix B in [588]).

---

<sup>3</sup>The  $x$  direction in Galactic coordinates points toward the Galactic center, while the  $y$  direction is the direction of Galactic rotation, and the  $z$  direction is perpendicular to the previous two and points away from the plane of the disc toward the north Galactic pole.

### CHAPTER 3. DARK MATTER DETECTION

However, the  $y$  component of  $\vec{v}_\odot$  dominates the vector sum in Equation 3.3.11 such that the  $x$  and  $z$  components of  $\vec{u}_\odot$  and  $\vec{u}_\oplus$  need not be considered. A convenient form for the  $y$  component of  $\vec{u}_\oplus$  is given by Equation 3.6 in [588]:

$$u_{\oplus,y} \simeq 15 \sin(2\pi t) \text{ km/s}, \quad (3.3.14)$$

where  $t$  is elapsed time (in years) relative to March 2<sup>nd</sup>. Ignoring the  $x$  and  $z$  components, the relative WIMP-Earth velocity is therefore well approximated by

$$v_\oplus \simeq 232 + 15 \cos\left(2\pi \frac{t - 152.5}{365.25}\right) \text{ km/s}, \quad (3.3.15)$$

where  $t$  is now elapsed time (in days) since January 1<sup>st</sup>. The oscillatory part of  $v_\oplus$  causes a few percent modulation in the (expected) WIMP-nucleus scattering rate, referred to within the direct-direction community as “annual modulation.” If a direct-detection experiment manages to measure a nonzero rate of WIMP interactions, observation of this annual modulation is a potentially powerful method for confirming that the signal is indeed due to WIMPs in the Milky Way’s halo. Annual modulation will be discussed further in Section 3.3.5 with regard to the DAMA/LIBRA experiment. In the absence of a signal, it is customary to interpret direct-detection results relative to the average expected WIMP detection rate, calculated using  $\langle v_\oplus \rangle = 232 \text{ km/s}$ .

The final parameter required to fully specify the standard halo is the mean WIMP number density  $n_0$ , usually expressed in terms of the WIMP mass (which is allowed to vary) and the local WIMP matter density:  $n_0 = \rho_0/m_\chi$ .  $\rho_0$  depends on the total halo mass, and on how it is spatially distributed. As discussed in the previous chapter (see Section 2.2.1), there is significant uncertainty in the mass of the halo. It is common to assume that the Galactic mass (within a 100 kpc radius) is  $\lesssim 10^{12} \mathcal{M}_\odot$ , and that the luminous portions constitute  $\sim 10\%$  of the total. This yields  $\mathcal{M}_{\text{halo}}(r < 100 \text{ kpc}) \lesssim 9 \times 10^{11} \mathcal{M}_\odot$ , which is roughly consistent with estimates in the literature [376]. To determine the local halo density from  $\mathcal{M}_{\text{halo}}$ , a functional form for the density profile must be chosen. The authors in [376] have constrained the halo density using the general form given in Equation 3.2.1 for four choices of the exponents  $\delta$ ,  $\epsilon$  and

### 3.3. DIRECT DETECTION

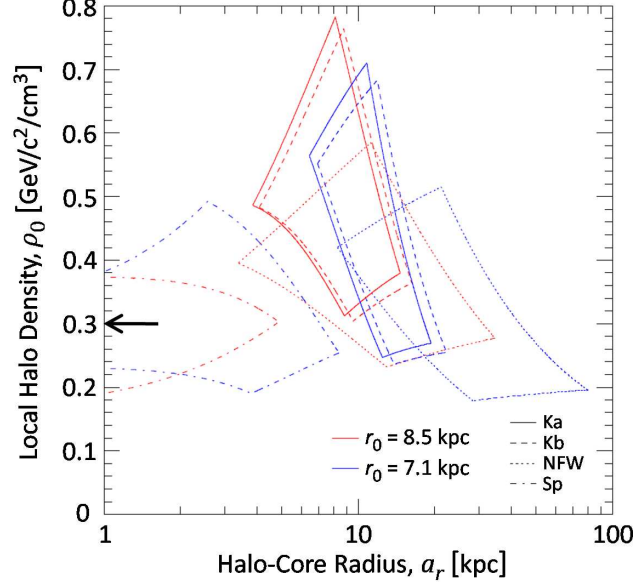


Figure 3.9: Allowed values of the Milky Way’s local halo density and halo-core radius for four versions of the density profile given in Equation 3.2.1, and for two values of the Sun’s galactocentric radius. The four models are:  $[\delta, \epsilon, \gamma] = [2, 3, 0.2]$  (Ka);  $[2, 3, 0.4]$  (Kb);  $[1, 3, 1]$  (NFW); and  $[2, 2, 0]$  (Sp). The arrow indicates  $\rho_0 = 0.3 \text{ GeV}/c^2/\text{cm}^3$ , the local halo density used in the standard halo model. Figure adapted from [376].

$\gamma$ . The results are shown in Figure 3.9 as a function of the halo-core radius for two different values of the Sun’s galactocentric radius, and indicate  $\rho_0 \simeq 0.2\text{--}0.8 \text{ GeV}/c^2/\text{cm}^3$ . The model most similar to the standard halo discussed here is the softened isothermal sphere (labelled Sp in the figure) with  $a_r = 1$ , for which  $\rho_0 \simeq 0.2\text{--}0.4 \text{ GeV}/c^2/\text{cm}^3$ . The central value,  $\rho_0 = 0.3 \text{ GeV}/c^2/\text{cm}^3$ , is the halo density used by the majority of the direct-detection community, and is the value I will use throughout the remainder of this thesis. Although it is likely that the Milky Way’s true halo-density profile resembles more closely one of the cuspy models (for which larger values of  $\rho_0$  are probable), the standard halo’s relatively small  $\rho_0$  is a conservative choice: assuming a smaller value of  $\rho_0$  leads to a less constraining WIMP-nucleon upper limit. Adopting a conser-

## CHAPTER 3. DARK MATTER DETECTION

Table 3.1: Values of the standard-halo parameters and target-material properties assumed for calculating WIMP-nucleon differential scattering rates in this thesis.

Symbol	Description	Value [reference]
<i>Halo parameters</i>		
$v_{\text{esc}}$	Galactic Escape velocity	544 km/s [183]
$v_{\odot} = \sqrt{2}\sigma$	Characteristic WIMP velocity	220 km/s [184]
$\langle v_{\oplus} \rangle$	Average Earth velocity	232 km/s [374]
$\rho_0$	Local halo density	$0.3 \text{ GeV}/c^2/\text{cm}^3$ [376]
<i>Target properties</i>		
$m_N$	Average nucleon mass	$0.9315 \text{ GeV}/c^2$ [88]
$A_{\text{Ge}}$	Average # of Ge nucleons	72.64 [88]
$A_{\text{Si}}$	Average # of Si Nucleons	28.0855 [88]

vative  $\rho_0$  helps ensure that the sensitivity of a direct-detection experiment is not overstated. While not so critical for the intercomparison of direct-detection experiments, it can affect comparison with indirect-detection and collider results. The standard-halo parameters used to interpret the results of this thesis (described in Chapter 6) are summarized in Table 3.1.

### WIMP-Nucleus Differential Scattering Rate

The rate at which WIMPs are expected to interact with a terrestrial detector is proportional to the product of the WIMP-nucleus cross section,  $\sigma_{\chi-T}$ , and the WIMP flux. The WIMP flux (relative to the target) is simply the product of the WIMP number density and WIMP-Earth relative velocity. The scattering

### 3.3. DIRECT DETECTION

rate per unit target material is therefore

$$dR_\chi = \frac{N_A}{A} \sigma_{\chi-T} dn_\chi v, \quad (3.3.16)$$

where  $N_A$  is Avogadro's number (number of atoms per mole) and  $A$  is the average number of nucleons per atom of target material (number of grams per mole). To obtain the absolute scattering rate due to Galactic WIMPs, one must consider the dependence of the cross section on the energy transferred during a scattering event (which depends on the WIMP velocity), and integrate over all possible WIMP velocities.

The absolute scattering rate, however, is less interesting than the differential scattering rate  $dR_\chi/dE_R$  (also referred to as a deposited-energy spectrum), where the subscript  $R$  is an abbreviation for recoil (as in Equation 3.3.3) such that  $E_R$  denotes the energy transferred from a WIMP to a target nucleus—the nuclear-recoil energy. Although  $dR_\chi/dE_R$  is essentially a featureless exponential (as we will see shortly—see, *e.g.*, Figure 3.11), it can differ substantially from the energy spectrum deposited by a source of background radiation. Consequently, a statistical comparison of spectral shapes can be used to distinguish WIMPs from backgrounds.

The differential scattering rate is the principle observable of (most) direct-detection experiments. Combining Equations 3.3.6 and 3.3.16, it can be written as

$$\frac{dR_\chi}{dE_R} = \frac{N_A \rho_0}{A k_f m_\chi} \int d^3v f(\vec{v}, \vec{v}_\oplus) v \frac{d\sigma_{\chi-T}}{dE_R}, \quad (3.3.17)$$

where the differential cross section,  $d\sigma_{\chi-T}/dE_R$ , is a function of both  $E_R$  and  $v$ . It is customary to separate  $d\sigma_{\chi-T}/dE_R$  into a (constant) point-like WIMP-nucleus cross section,  $\sigma_0$ , and an energy-dependent nuclear form factor  $F$  (discussed further below) [374];

$$\frac{d\sigma_{\chi-T}}{dE_R} = \frac{\sigma_0}{E_R^{\max}} F^2(q), \quad (3.3.18)$$

where  $E_R^{\max}$  is the maximum recoil energy for a given WIMP velocity,  $q \equiv \sqrt{|\vec{q}|^2} = \sqrt{2m_T E_R}$  is the momentum transferred during the scattering event,

### CHAPTER 3. DARK MATTER DETECTION

and  $m_T$  is the nuclear mass. In the WIMP-nucleus center-of-mass frame, the nuclear-recoil energy is given by

$$E_R = \frac{m_r^2}{m_T} v^2 (1 - \cos \theta^*), \quad (3.3.19)$$

where  $\theta^*$  is the center-of-mass scattering angle, and

$$m_r \equiv \frac{m_\chi m_T}{m_\chi + m_T} \quad (3.3.20)$$

is the WIMP-nucleus reduced mass. Clearly,  $E_R^{\max} = 2m_r^2 v^2 / m_T$ , such that

$$\frac{dR_\chi}{dE_R} = \frac{N_A \rho_0 \sigma_0 m_T}{2 A k_f m_\chi m_r^2} F^2 \left( q = \sqrt{2m_T E_R} \right) \int_{v_{\min}}^{v_{\max}} d^3 v \frac{f(\vec{v}, \vec{v}_\oplus)}{v}, \quad (3.3.21)$$

where  $v_{\max}$  corresponds to the  $|\vec{v} + \vec{v}_\oplus| \leq v_{\text{esc}}$  requirement, and

$$v_{\min} = \frac{q}{2m_r} = \sqrt{\frac{m_T E_R}{2m_r^2}} \quad (3.3.22)$$

is the minimum WIMP velocity that yields the given recoil energy  $E_R$ .

Before quoting the expression for the fully general differential spectrum, it is instructive to consider the limit in which  $\vec{v}_\oplus = 0$  and  $v_{\text{esc}} \rightarrow \infty$ . In this case, the velocity integral in Equation 3.3.21 is isotropic in  $v$  and easily evaluated:

$$\left. \frac{dR_\chi}{dE_R} \right|_{\substack{\vec{v}_\oplus=0 \\ v_{\text{esc}} \rightarrow \infty}} = \frac{2N_A \rho_0}{\sqrt{\pi} A m_\chi} \sigma_0 v_\odot \frac{e^{-E_R/E_0 r_m}}{E_0 r_m} F^2 \left( q = \sqrt{2m_T E_R} \right), \quad (3.3.23)$$

where the limiting form  $k_0 = (\pi v_\odot^2)^{3/2}$  has been substituted for  $k_f$ ,  $E_0 = m_\chi v_\odot^2 / 2$  is the most probable kinetic energy of a Galactic WIMP, and

$$r_m \equiv \frac{4m_r^2}{m_\chi m_T} = \frac{4m_\chi m_T}{(m_\chi + m_T)^2} \quad (3.3.24)$$

is the dimensionless WIMP-nucleus reduced-mass parameter. If we consider the case in which  $F^2 = 1$ , Equation 3.3.23 can be integrated over recoil energy to yield the absolute WIMP-nucleus scattering rate

$$R_0 \equiv \int_0^\infty dE_R \left. \frac{dR_\chi}{dE_R} \right|_{\substack{\vec{v}_\oplus=0 \\ v_{\text{esc}} \rightarrow \infty \\ F^2(q)=1}} = \frac{2N_A \rho_0}{\sqrt{\pi} A m_\chi} \sigma_0 v_\odot. \quad (3.3.25)$$

### 3.3. DIRECT DETECTION

Rewriting Equation 3.3.23 in terms of  $R_0$  yields a simplified form for the WIMP-nucleus differential scattering rate:

$$\left. \frac{dR_\chi}{dE_R} \right|_{\substack{\vec{v}_\oplus=0 \\ v_{\text{esc}} \rightarrow \infty}} = \frac{R_0}{E_0 r_m} e^{-E_R/E_0 r_m} F^2 \left( q = \sqrt{2m_T E_R} \right). \quad (3.3.26)$$

This limiting case provides some useful insight into the expected WIMP signal without having to resort to numerical calculations. As promised, the differential spectrum as a function of nuclear-recoil energy is essentially an exponential. The ability to detect small energy depositions is therefore critical to the sensitivity of a direct-detection experiment; scattering events with the lowest recoil energies are expected to be the most numerous. While a realistic form factor, WIMP-Earth velocity and Galactic escape velocity are important considerations, they (generally) cause only minor modifications to this exponential behavior.

For nonzero  $\vec{v}_\oplus$  and finite  $v_{\text{esc}}$ , the general expression for the differential scattering rate is (Equations 3.12 and 3.13 in [588])<sup>4</sup>

$$\frac{dR_\chi}{dE_R} = \frac{k_0}{k_f} \left[ \left. \frac{dR_\chi}{dE_R} \right|_{v_{\text{esc}} \rightarrow \infty} - \frac{R_0}{E_0 r_m} F^2(q) e^{-v_{\text{esc}}^2/v_\odot^2} \right], \quad (3.3.27)$$

where

$$\left. \frac{dR_\chi}{dE_R} \right|_{v_{\text{esc}} \rightarrow \infty} = \frac{R_0}{E_0 r_m} \frac{\sqrt{\pi} v_\odot}{4v_\oplus} F^2(q) \left[ \text{erf} \left( \frac{v_{\text{min}} + v_\oplus}{v_\odot} \right) - \text{erf} \left( \frac{v_{\text{min}} - v_\oplus}{v_\odot} \right) \right], \quad (3.3.28)$$

and  $q = \sqrt{2m_T E_R}$ . Although not immediately apparent, the exponential behavior (as a function of  $E_R$ ) is hidden in the combination of error functions. Use of these equations requires the following additional information:

---

<sup>4</sup>Note that while taking the  $v_{\text{esc}} \rightarrow \infty$  limit of Equation 3.3.27 yields Equation 3.3.28, simply setting  $v_\oplus = 0$  in Equation 3.3.28 does not return the limiting form given in Equation 3.3.26. A quick application of L'Hôpital's rule and the fundamental theorem of calculus in the  $v_\oplus \rightarrow 0$  limit, however, confirms their consistency.



### CHAPTER 3. DARK MATTER DETECTION

- A target material, specified by its atomic mass number  $A$ , must be chosen. Picking a value for  $A$  also specifies the nuclear mass  $m_T = Am_N$ . The average nucleon mass,  $m_N$ , is listed in Table 3.1 together with the values of  $A$  for the two detector materials studied in this thesis. Some detectors have a multi-element composition. It is common practice to compute the differential rate separately for each element by evaluating Equation 3.3.27 for each  $A$  and reducing it by a (multiplicative) molecular-mass fraction  $f_A$ . The Cryogenic Rare Event Search with Superconducting Thermometers (CRESST II) [591], for example, employs calcium tungstate crystals ( $\text{CaWO}_4$ ) that are characterized by three target materials: calcium with  $A = 40.1$  and  $f_{\text{Ca}} \simeq 0.14$ ; tungsten with  $A = 183.8$  and  $f_{\text{W}} \simeq 0.64$ ; and oxygen with  $A = 16$  and  $f_{\text{O}} \simeq 0.22$ .
- The zero-momentum transfer WIMP-nucleus cross section,  $\sigma_0$ , and the WIMP mass,  $m_\chi$ , are the unknown quantities that direct-detection experiments attempt to measure or constrain. Due to the variety of nuclei used in different experiments, the convention is to scale  $\sigma_0$  down to the WIMP-nucleon cross section according to [592]

$$\sigma_0 = A^2 \frac{m_r^2}{m_{r'}^2} \sigma_{\chi-N}, \quad (3.3.29)$$

where  $m_{r'}$  is the WIMP-nucleon reduced mass (obtained by replacing  $m_T$  with  $m_N$  in Equation 3.3.20), and  $\sigma_{\chi-N}$  is the spin-independent (or scalar) WIMP-nucleon cross section. In the case of neutralino dark matter, the dominant scalar scattering channels (within the MSSM) involve WIMP-quark scattering via exchange of a squark or neutral Higgs scalar. Scalar neutralino-gluon elastic scattering is also possible, but only through lower-rate Feynman diagrams involving quark and/or squark loops (see, *e.g.*, Appendix C in [2]). As is common practice in the field, I have assumed that WIMPs have equal scalar couplings to protons and neutrons. Because WIMPs are expected to scatter coherently from an entire nucleus, the contribution from each nucleon is summed before squar-

### 3.3. DIRECT DETECTION

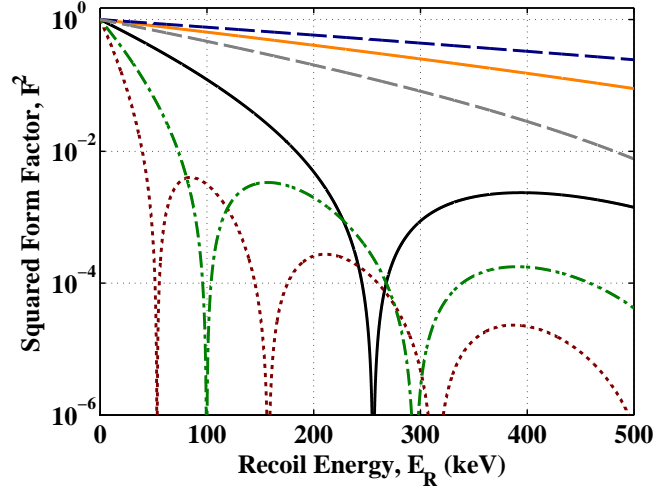


Figure 3.10: The Helm/Lewin-Smith nuclear form factor (squared) as a function of nuclear-recoil energy for several detector materials: Ge (dark/black solid), Si (light/orange solid), W (dark red dotted), I (green dash-dotted), Ne (dark/blue dashed), and Ar (light/gray dashed). Due to the quasi-exponential nature of the WIMP-nucleus differential scattering rate, the energy range of most interest for dark-matter direct detection is  $E_R < 100$ .

ing the scattering amplitude, giving rise to the  $A^2$  coherence factor.<sup>5</sup> By factoring out the dependence on  $A$ , Equation 3.3.29 facilitates the comparison between experiments with different detector materials. In the absence of a WIMP signal, the differential scattering rate observed by a direct-detection experiment is used to constrain  $\sigma_{\chi-N}$  as a function of  $m_\chi$  (see, *e.g.*, Figures 2.18 and 2.19).

<sup>5</sup>WIMPs might also scatter from nuclei via a spin-dependent (or axial-vector) interaction, in which case the  $A^2$  coherence factor in Equation 3.3.29 is not entirely appropriate. For the spin-dependent case, I refer the interested reader to Section 2.2.2 in [592]. In this thesis only the spin-independent WIMP-nucleon cross section is explored. For simplicity and unless otherwise stated, any hereafter mention of the WIMP-nucleon cross section (or use of the symbol  $\sigma_{\chi-N}$ ) is a reference to the WIMP-nucleon spin-independent elastic scattering cross section.

### CHAPTER 3. DARK MATTER DETECTION

- A nuclear form factor  $F(q)$  must be specified. The form factor accounts for the nonzero momentum transfer that occurs during a WIMP-nucleus scattering event. If an incoming WIMP has a large enough momentum relative to a target nucleus of sufficient size, the WIMP can start to probe the structure of the nucleus, and the WIMP-nucleus scattering coherence partially breaks down.  $F$  is a dimensionless measure of the loss of coherence, varying from 0 to 1 depending on  $q$  and  $A$ , and normalized such that  $F(q = 0) = 1$ . If the nuclear ground state is assumed to be (roughly) spherical, the form factor can be approximated by the Fourier transform of a spherically symmetric mass profile  $\rho_T$ :

$$F(q) = \frac{1}{m_T} \int d^3r \rho_T(r) \exp(-i\vec{q} \cdot \vec{r}). \quad (3.3.30)$$

Because it is difficult to probe nuclear-mass distributions directly,  $\rho_T$  is typically modeled according to the distribution of electric charge in the nucleus:

$$\rho_T(r) = \frac{m_T}{Ze} \rho_+(r), \quad (3.3.31)$$

where  $Z$  is the number of protons,  $e$  is the electric charge, and  $\rho_+$  is the charge density. Accurate form factors derived from measurements of  $\rho_+$  using electron-scattering data are available in the literature for several nuclei [593]. The standard in the direct-detection community, however, is to use the Helm form factor [594], an analytic form that works reasonably well for any  $A$ :

$$F(q) = 3 \left[ \frac{\sin(x_q) - x_q \cos(x_q)}{x_q^3} \right] \exp\left(-\frac{q^2 s_T^2}{2\hbar^2}\right), \quad (3.3.32)$$

where  $x_q \equiv qr_T/\hbar$  is a dimensionless quantity,  $r_T$  is an effective nuclear radius (*a la* Equation 3.3.2), and  $s_T$  is the nuclear-skin thickness. Lewin and Smith use muon scattering data from [595] to derive a general parameterization for  $r_T$ :

$$r_T = \sqrt{c_T^2 + \frac{7}{3}\pi^2 a_T^2 - 5s_T^2} \text{ fm}, \quad (3.3.33)$$

### 3.3. DIRECT DETECTION

where  $a_T = 0.52 \text{ fm}$ ,  $s_T = 0.9 \text{ fm}$ , and  $c_T = (1.23A^{1/3} - 0.6) \text{ fm}$  [588]. Equation 3.3.32 combined with Equation 3.3.33 is often referred to as the Helm/Lewin-Smith form factor, and is plotted for several  $A$  in Figure 3.10 as a function of  $E_R$ .

With the expression for  $dR_\chi/dE_R$  given in Equation 3.3.27 (and accompanying equations), the values in Table 3.1, and the Helm/Lewin-Smith form factor, the standard framework used to calculate WIMP deposited-energy spectra (and detection rates), for given values of  $\sigma_{\chi-N}$  and  $m_\chi$ , is fully specified. As an example, the spectra resulting from 5 and 100  $\text{GeV}/c^2$  WIMPs scattering from Ge and Si targets are shown in Figure 3.11. The abrupt high-energy cut off that each distribution exhibits is due to the finite value of the Galactic escape velocity. Otherwise, the expected signal is essentially an exponential as a function of  $E_R$ . It is clear from the upper panel in Figure 3.11 that sensitivity to light WIMPs necessitates a low threshold for detecting recoil energy.

It is useful to review a few numerical results to get a feel for how the detectable event rate and average deposited energy depend on target material and WIMP mass. In the limiting case of  $v_\oplus = 0$  and  $v_{\text{esc}} \rightarrow \infty$ , the absolute detectable event rate is given by  $R_0$ , which can be written in the following convenient form:

$$R_0 = \frac{387.5}{Am_\chi} \left[ \frac{\rho_0}{0.3 \text{ GeV}/c^2/\text{cm}^3} \right] \left[ \frac{\sigma_0}{10^{-36} \text{ cm}^2} \right] \left[ \frac{v_\odot}{220 \text{ km/s}} \right] \text{ kg}^{-1} \text{ d}^{-1}, \quad (3.3.34)$$

where  $A$  is in grams per mole and  $m_\chi$  is in  $\text{GeV}/c^2$ . Lets consider two of the neutralino models from Figure 2.19 that just evade the CDMS II upper limit: a light neutralino with  $m_\chi = 7 \text{ GeV}/c^2$  and  $\sigma_{\chi-N} \simeq 7 \times 10^{-41} \text{ cm}^2$ , and a heavy neutralino with  $m_\chi = 500 \text{ GeV}/c^2$  and  $\sigma_{\chi-N} \simeq 10^{-43} \text{ cm}^2$ . Using Equation 3.3.29 for a germanium target,  $\sigma_0 \simeq 2 \times 10^{-35} \text{ cm}^2$  for the former and  $\sim 2 \times 10^{-36} \text{ cm}^2$  for the latter. With the values for  $\rho_0$  and  $v_\odot$  from Table 3.1, the approximate detectable event rate expected for the light (heavy) neutralino is  $\sim 15 \text{ kg}^{-1} \text{ d}^{-1}$  ( $\sim 0.02 \text{ kg}^{-1} \text{ d}^{-1}$ ). It would seem at first glance that the light neutralino ought to be much easier to detect. To understand why low-mass WIMPs with relatively large cross sections continue to evade detection, con-

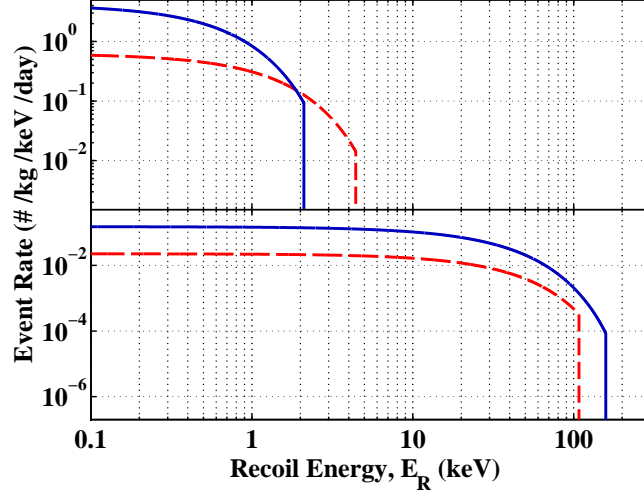


Figure 3.11: Expected differential event rates for  $5 \text{ GeV}/c^2$  (upper panel) and  $100 \text{ GeV}/c^2$  (lower panel) WIMPs scattering from Ge (blue/solid lines) and Si (red/dashed lines) targets within the standard direct-detection framework described in the text, and for an arbitrarily chosen WIMP-nucleon cross section of  $1 \times 10^{-41} \text{ cm}^2$ . Event rates are given in events per kilogram per keV per day, a convenient unit for direct-detection experiments that is sometimes referred to as the differential rate unit (dru) [588]. As indicated in the upper panel, sensitivity to light WIMPs depends critically on an experiment's energy detection threshold. Figure taken from [596].

sider the (approximate) average energy deposition

$$E_0 r_m = \left[ \frac{m_\chi}{1 \text{ GeV}/c^2} \right] \left[ \frac{v_\odot}{220 \text{ km/s}} \right]^2 \frac{4m_\chi m_T}{(m_\chi + m_T)^2} \times 0.27 \text{ keV}. \quad (3.3.35)$$

For a Ge target, the average energy deposited by a  $7 \text{ GeV}/c^2$  WIMP is  $\sim 0.5 \text{ keV}$ , compared to  $\sim 57 \text{ keV}$  for a  $500 \text{ GeV}/c^2$  WIMP. A typical energy threshold for a Ge-based experiment is 10–20 keV, which is the reason the higher-rate light-neutralino model evades the CDMS II limit in Figure 2.19. A more accurate average deposited-energy calculation is shown as a function of WIMP mass in Figure 3.12 for a variety of target materials. Light nuclear targets provide a better match to the mass of light WIMPs, making larger energy depositions

### 3.3. DIRECT DETECTION

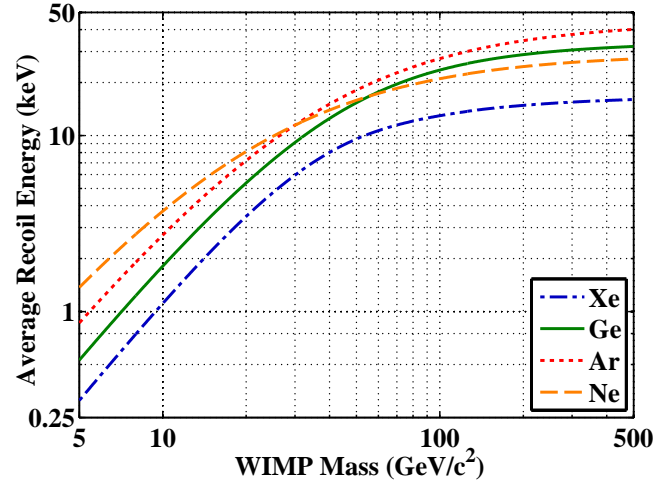


Figure 3.12: Average deposited recoil energy, calculated with the standard framework described in the text, for various nuclear targets as a function of WIMP mass. Light nuclei have a distinct advantage (in terms of energy) for detecting light WIMPs. While very heavy nuclei (such as Xe) benefit (in terms of rate) from larger coherence factors, their form factors (see, *e.g.*, Figure 3.10) cause a reduction in the average recoil energy that is particularly evident for heavy WIMPs.

kinematically favorable. This advantage relative to heavier nuclei is countered by the reduced event rate associated with a smaller coherence factor. Also note that for very heavy targets (such as Xe), the form factor causes a nontrivial reduction in the average recoil energy.

#### Experimental Considerations

There are a number of experimental considerations involving detector size and performance that affect the sensitivity of a direct-detection experiment. Without specifying a particular detector technology, I can describe these in only a general sense here. I will elaborate further in Chapters 5 and 6 for the specific experiment and analysis studied in this thesis.

### CHAPTER 3. DARK MATTER DETECTION

I have already alluded to the most important experimental consideration with respect to low-mass WIMP sensitivity, the recoil-energy detection threshold  $E_{\text{th}}$ . A threshold is (typically) established to prevent an overwhelming event rate due to low-energy backgrounds. Although a large number of valid WIMP scatters are expected to deposit small energies, they are of no use to the experimentalist if they cannot be distinguished as such. In particular, despite the common use of low-noise amplifiers operated at low temperature, all direct-detection experiments are subject to electronic noise of one form or another—an inescapable reality of experimental physics. In most experiments, noise manifests as a Gaussian distribution (with width  $\sigma$ ) centered at  $E_R \simeq 0$ . Setting  $E_{\text{th}}$  many  $\sigma$  above zero helps ensure a low probability for electronic-noise events to be misidentified as potential WIMP events. As mentioned above, energy thresholds of  $\sim 10$ – $20$  keV are common. In some cases, however, a particular detector technology might enable a sub-keV threshold, or might limit WIMP sensitivity by imposing a relatively large threshold. For example, the CRESST collaboration achieved  $E_{\text{th}} = 0.6$  keV with their sapphire ( $\text{Al}_2\text{O}_3$ ) crystals [597], while the Wimp ARgon Programme (WARP) strives for a 30 keV threshold using a liquid-argon target [598].

Additionally, differential scattering rates depend on energy resolution. An experiment’s energy resolution is a measure of how accurately it can measure a given recoil energy. Spectra calculated with the framework described above are idealized, perfect-resolution versions of what an experiment can actually measure. In reality, the observed spectrum will be “smeared” by an experiment’s finite energy resolution  $\sigma_E$ :

$$\frac{d\tilde{R}_\chi}{dE_R}(E_R) = \frac{1}{\sqrt{2\pi\sigma_E^2}} \int \frac{dR_\chi}{dE_R}(E'_R) \exp\left[\frac{(E'_R - E_R)^2}{-2\sigma_E^2}\right] dE'_R, \quad (3.3.36)$$

where  $d\tilde{R}_\chi/dE_R$  and  $dR_\chi/dE_R$  denote the observed and ideal spectra, respectively. Equation 3.3.36 is simply the prescription for a Gaussian convolution. Because the ideal spectrum is essentially an exponential, smearing it with a Gaussian causes only a minor change in the observed behavior as a function of recoil energy. The most important effect has to do with the interplay be-

### 3.3. DIRECT DETECTION

tween resolution and threshold. Finite energy resolution near threshold causes sub-threshold energy depositions to occasionally exceed  $E_{\text{th}}$ , and *vice versa*. An important consequence is that an experiment with a given threshold will have (some) sensitivity to a low-mass WIMP whose ideal spectrum cuts off just below  $E_{\text{th}}$ . With most modern detectors,  $\sigma_E$  is small enough that the effect of finite resolution is inconsequential when  $E_{\text{th}} \gtrsim 5 \text{ keV}$ , and need not be considered. For thresholds of a few keV or less, however, resolution effects can become important, particularly regarding sensitivity to sub-10 GeV/ $c^2$  WIMPs.

Finally, when comparing the absolute number of events in an experimentally observed spectrum to a particular WIMP model, the latter must be scaled by the detection efficiency and exposure of the former. The exposure is simply the product of the detector size  $\mathcal{M}_D$  (in kilograms) and the duration of the experiment  $\tau_{LT}$  (in days).  $\tau_{LT}$  is often referred to as the “live time,” hence the  $LT$  subscript. Multiplying  $d\tilde{R}_\chi/dE_R$  by the exposure converts the differential spectrum from units of events  $\text{kg}^{-1} \text{keV}^{-1} \text{d}^{-1}$  to events per keV. The detection efficiency,  $\varepsilon \equiv \varepsilon(E_R)$ , is a dimensionless function of  $E_R$  that varies from 0 to 1. In general,  $\varepsilon$  accounts for any experimental effects that reduce WIMP sensitivity but do not depend on the precise nature of the WIMP-nucleus interaction. For example, a reduced detection efficiency ( $<100\%$ ) might arise because of periods of poor data quality (effectively reducing  $\tau_{LT}$ ), or because part of the detector does not work properly (effectively reducing  $\mathcal{M}_D$ ). Average detection efficiencies as low as 40% and as high as 80% are not uncommon. After all is said and done, the differential spectrum observed by a direct-detection experiment is used to constrain a particular WIMP model by comparing the former to

$$\frac{d\tilde{R}_\chi}{dE_R}(E_R) \varepsilon(E_R) \mathcal{M}_D \tau_{LT} \Theta(E_R - E_{\text{th}}), \quad (3.3.37)$$

where  $\Theta$  is the step function ( $\Theta = 1$  for  $E_R \geq E_{\text{th}}$ , and  $\Theta = 0$  for  $E_R < E_{\text{th}}$ ).



### 3.3.3 Backgrounds & Shielding

The standard framework outlined above provides the means to calculate the expected rate of WIMP interactions with a terrestrial detector. Consider again the low-mass WIMP model, for which  $m_\chi = 7 \text{ GeV}/c^2$  and  $\sigma_{\chi-N} = 7 \times 10^{-41} \text{ cm}^2$ , incident on a Ge target. For the sake of argument, let's assume  $\langle \varepsilon \rangle = 50\%$  and  $E_{\text{th}} = 1 \text{ keV}$ . In this oversimplified and somewhat optimistic scenario, the total rate obtained by integrating Equation 3.3.37 is  $\sim 4$  events per kilogram per day. For the  $500 \text{ GeV}/c^2$  neutralino considered above, the equivalent rate is  $\sim 0.01 \text{ events kg}^{-1} \text{ d}^{-1}$ . In any case, it is fair to assume that the WIMP rate will be no more than a few events per kilogram-day of exposure, and possibly several orders of magnitude less.

A critical question is whether or not this low event rate can be distinguished (or isolated) from the much larger rate due to known forms of background radiation. Although most detectors have some means of differentiating backgrounds from signal, consider for the moment that events due to WIMPs and backgrounds are indistinguishable. Furthermore, consider an experiment in which the number of interactions in the target can be measured, but the magnitude of the energy depositions cannot—sensitivity to  $R_\chi$  but not  $dR_\chi/dE_R$ . This is known as a counting experiment. In the absence of background events (“background free”) WIMP sensitivity scales linearly with exposure ( $\propto \mathcal{M}_D \tau_{LT}$ ). For example, if 0 events are observed (and ignoring questions of efficiency and energy dependence), the 90% C.L. upper limit on the WIMP-nucleus scattering rate is  $2.3/\mathcal{M}_D \tau_{LT}$ ;<sup>6</sup> increasing the exposure probes ever smaller WIMP-nucleon cross sections:  $\sigma_{\chi-N} \propto 1/\mathcal{M}_D \tau_{LT}$ . Once a nonzero event rate is observed, however, additional exposure no longer yields improvement. With no ability to distinguish signal from background, the observed rate cannot be attributed to WIMPs. The best one can do in this situation

---

<sup>6</sup>2.3 is the one-sided 90% confidence interval for a Poisson random variable given that 0 has been observed, *i.e.*, 2.3 is the 90% C.L. error on 0. The principle observable of a counting experiment—the number of observed counts (equal to the product of the rate and exposure)—is a Poisson random variable. For a large number of events, the  $1\sigma$  Poisson error is given by the square root of the counts. For further information see [88].

### 3.3. DIRECT DETECTION

is to count events until the statistical error on the observed rate is small compared to the rate itself, at which point the WIMP-nucleus event rate is limited to the measured background rate. The experiment is said to be “background limited.”

The obvious problem with the above counting experiment is that it could never discover WIMPs. Consequently, a major part of the direct-detection experimental program is to gain a detailed knowledge of *all* possible backgrounds. Utilizing independent information, the expected rate of backgrounds can be predicted (*e.g.*, via simulations), and a nonzero background rate can be subtracted to improve WIMP sensitivity. In this scenario, the residual number of observed events (following background subtraction) will increase by the square root of the exposure (the Poisson error on the number of background events), and the upper limit on the WIMP cross section will improve in proportion to  $1/\sqrt{\mathcal{M}_D\tau_{LT}}$ .

Additionally, most direct-detection experiments have the ability to measure recoil-energy spectra. Spectral-shape differences between the deposited-energy distributions expected for backgrounds and WIMPs can be used to identify recoil-energy ranges in which the number of background events is minimal compared to the expected WIMP signal, leading to improved WIMP-nucleon cross-section limits (with increased exposure) despite the occurrence of background events. One such technique is Yellin’s “optimum interval” method for calculating upper limits [599]. Application of the optimum interval method will be discussed further in Chapter 6.

Of course, the best thing is to remain background free such that WIMP sensitivity scales in direct proportion to exposure. There are two general methods employed to achieve a background-free environment for direct-detection experiments: reduction and rejection. Some of the discrimination techniques used to achieve the latter are described below in Section 3.3.4. Reduction involves the identification of background sources, and the application of appropriate types of shielding. In this section I describe several of the main classes of background radiation, and discuss the techniques most commonly used to shield against them.

### Cosmic Rays

The most pervasive background is due to secondaries from cosmic rays interacting in the Earth’s upper atmosphere.  $\sim 90\%$  of all cosmic rays are protons, which collide inelastically with particles in the atmosphere to create charged and neutral mesons. The mesons quickly decay to a mix of particles—muons, electrons, positrons, photons, pions and neutrinos—that can reach the surface of the Earth. Some cosmic-ray primaries also survive transport through the atmosphere (albeit degraded in energy). I will focus on the muon flux here. Excluding neutrinos (discussed below), the others are easily shielded with a relatively modest thickness of material. The range of a 1 GeV proton in rock, for example, is  $\lesssim 1$  m [88]. The flux of muons at the surface is equivalent to roughly one per second passing through your hand if you were to hold it out palm up ( $\sim 1 \text{ min}^{-1} \text{ cm}^{-2}$  through a horizontal surface for  $E_\mu \gtrsim 1 \text{ GeV}$  [88]). Such a large rate would drown out any potential WIMP signal in a direct-detection experiment conducted at sea level.

The primary defense against muons is to conduct experiments in underground laboratories such that an overburden of rock moderates the muon flux to a manageable level. The majority of the muons that reach sea level are minimum ionizing,  $\langle E_\mu \rangle \simeq 4 \text{ GeV}$  [88], and can be characterized by an energy loss of  $\gtrsim 2.6 \text{ MeV cm}^2 \text{ g}^{-1}$  (increasing with increasing  $E_\mu$  due to radiative losses) [600]. Furthermore, there is a nontrivial flux of muons at sea level with TeV-scale energies. Underground laboratories must be many hundreds of meters deep for the overburden to provide an effective shield against such energetic and penetrating particles. Figure 3.13 shows the muon flux as a function of depth for a compilation of flux measurements taken at laboratories across the globe. The depth is measured in meters water equivalent (m.w.e.) to normalize different rock compositions to a single standard. For reference, a common density for rock is  $\sim 2.5\text{--}3 \text{ g/cm}^3$ , while water has a density of almost exactly  $1 \text{ g/cm}^3$ . For more information regarding the flux of muons at sea level, underground and underwater, I refer the interested reader to [601].

### 3.3. DIRECT DETECTION

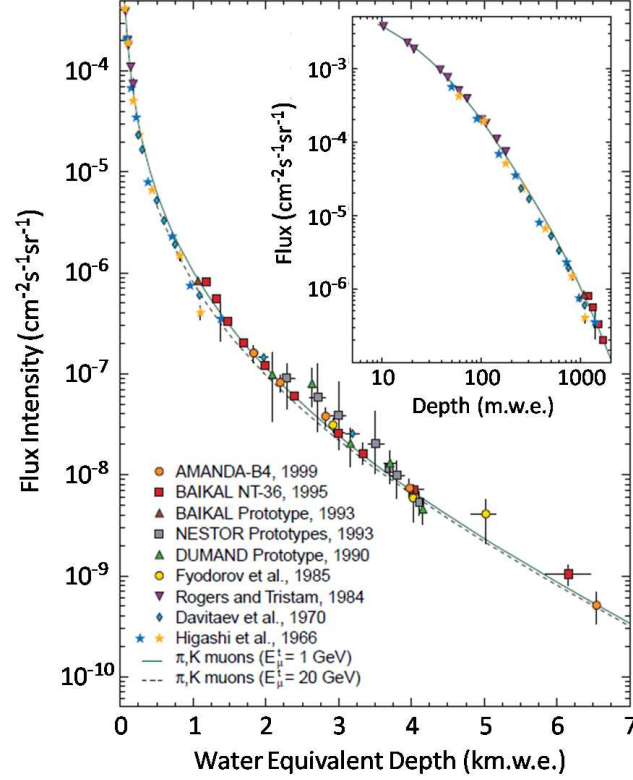


Figure 3.13: Compilation of muon-flux measurements from underground and underwater laboratories across the globe, as a function of depth in kilometers water equivalent (km.w.e.). The inset provides a zoomed in view of the muon intensity for shallow depths (in meters water equivalent). For further information regarding individual measurements and the flux models (solid and dashed lines), see [601]. Figure taken from [601].

As is clear from Figure 3.13, it is impossible to completely escape the muon flux using passive shielding techniques. Consequently, it is common practice to surround a direct-detection experiment with an active shield that serves as a “muon veto,” tagging any muons that survive passage through a lab’s overburden. Muon vetoes usually rely on conventional technologies for converting ionization losses to scintillation (*e.g.*, plastic or liquid scintillators) or Cherenkov photons (*e.g.*, water tanks) that are easily detected using photo-

## CHAPTER 3. DARK MATTER DETECTION

multiplier tubes. When a muon passes through the veto, its presence is used to reject any activity in the main dark-matter target as potentially due to cosmogenics (and not WIMPs). Ideally, a muon veto provides full  $360^\circ$  coverage, *i.e.*, it is hermetic. However, at sufficient depth, the muon flux becomes primarily vertical, and an umbrella of ionizing detectors can serve as an effective veto.

The atmospheric-, solar- and cosmic-neutrino fluxes cannot be moderated by any amount of overburden. There is also a geo-neutrino flux resulting from decays of radioisotopes distributed throughout the volume of the Earth, as well as several sources of man-made neutrinos (*e.g.*, nuclear reactors). Neutrinos can scatter coherently from the nuclei in a direct-detection target, depositing keV-scale nuclear-recoil energies if sufficiently energetic. Such a neutrino-nucleus scattering event would be indistinguishable from a WIMP-nucleus interaction. The flux of neutrinos arriving underground therefore represents an irreducible background for direct-detection experiments. Fortunately, the neutrino background is not expected to become problematic until dark-matter experiments achieve sensitivities equivalent to  $\sigma_{\chi-N} \simeq 10^{-46} \text{ cm}^2$ , at which point a ton-year exposure for a detector with  $E_{\text{th}} \simeq 1 \text{ keV}$  might result in tens of neutrino events [602]. For further information on the prospect of using direct-detection technology to detect neutrino-nucleus coherent scattering, I refer the interested reader to [602, 603].

### Radioactivity

Every direct-detection experiment is subject to a substantial flux of radiation resulting from radioisotope decays. Sources include the rock from which an underground laboratory is excavated, the materials used to construct an experiment (*e.g.*, mechanical supports and shielding materials), the target material itself, and (worst of all) the very air we breath. In short, all materials are radioactive at some level. The dominant forms of radiation associated with terrestrial radioactivity can be divided into two classes: electrically charged particles (and photons) that interact primarily with the electrons in a ma-

### 3.3. DIRECT DETECTION

terial, and neutral particles that interact with a material’s nuclei. The only product of radioactivity that fits the latter description is the neutron. Neutrons are an important enough background to be discussed separately (see below). Electrons, positrons, alpha particles, and photons comprise the other type of radioactivity. These can be further classified as either internal or external. The former refers to radiation emitted by isotopes embedded in or near the direct-detection target, while the latter refers to radiation incident from the outside of the apparatus. Since electrons, positrons and alpha particles are easily shielded, the only external background that needs to be considered is the incident flux of photons. Furthermore, only those isotopes that emit the most energetic gamma rays (“gammas”) need be considered to understand how much shielding is necessary to moderate *all* external photons to a manageable level. The internal situation is more complicated, and is largely specific to a particular experiment’s choice of materials. There are a few commonalities that I briefly review below.

For the external gamma-ray background, the most important isotopes are  $^{238}\text{U}$ ,  $^{232}\text{Th}$  and  $^{40}\text{K}$ . All three have half-lives of billions of years and are commonly found at the level of a few parts per million (or less) in typical rock (see, *e.g.*, Table 4.5 in [604]). Bulk materials used to construct mechanical-support structures (*e.g.*, steel, aluminum and concrete) can also have significant concentrations of these isotopes.  $^{40}\text{K}$  usually decays to  $^{40}\text{Ca}$  via emission of an electron. However, it can also decay to  $^{40}\text{Ar}$  via electron capture, resulting in the emission of a 1.46 MeV gamma. Both  $^{238}\text{U}$  and  $^{232}\text{Th}$  are the parents of long “decay chains” that involve the subsequent decay of roughly a dozen isotopes before finally arriving at stable isotopes of lead. Most of the gamma-ray emission from the uranium chain is due to the beta decay of  $^{214}\text{Bi}$  near the end of the decay sequence, with notable high-energy lines at 2.2 and 2.45 MeV. The thorium chain’s gamma-ray emission is predominately from beta decays of  $^{228}\text{Ac}$  (near the beginning of the decay sequence) and  $^{208}\text{Tl}$  (at the end of the decay sequence), with notable lines at 911 keV from the former and 2.6 MeV from the latter. The  $^{208}\text{Tl}$  2.6 MeV line is the most energetic gamma emitted by the decay of a naturally oc-

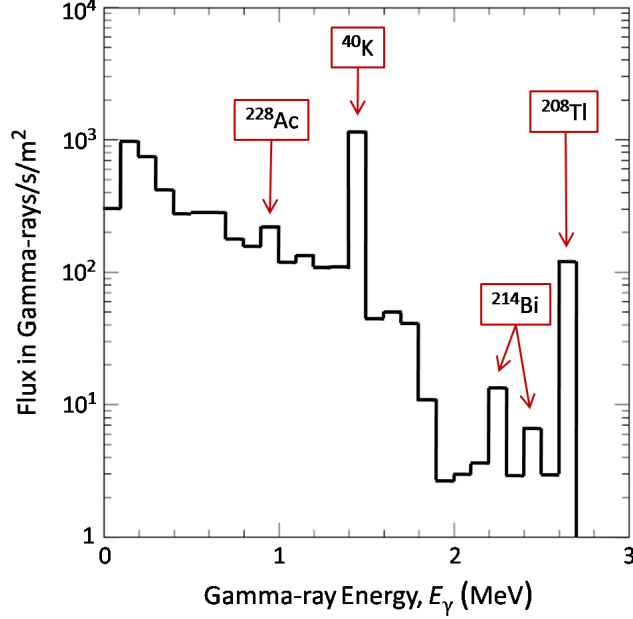


Figure 3.14: The flux of gamma rays (as a function of gamma-ray energy) emitted by the walls and floor of an experimental hall in the Soudan Underground Laboratory in northern Minnesota due to decays of  $^{40}\text{K}$  and various isotopes in the  $^{238}\text{U}$  and  $^{232}\text{Th}$  decay chains. The isotopes responsible for the most prominent lines are indicated. Figure adapted from [600].

curing radioisotope. As an example, the combined spectrum of gamma-ray energies emitted by the rock and concrete in one of the experimental halls in the Soudan Underground Laboratory is shown in Figure 3.14. For further information regarding particular isotopes and their decay channels, I refer the interested reader to [412] (from which much of the above information was obtained).

Shielding a dark-matter target against the external gamma-ray background involves surrounding it with layers of one or more materials of sufficient thickness to attenuate the most dangerous  $\sim 2\text{ MeV}$  gammas. MeV-scale gammas interact primarily via Compton scattering in low- $Z$  materials, and also via the

### 3.3. DIRECT DETECTION

photoelectric effect and Rayleigh scattering in high- $Z$  materials.<sup>7</sup> The attenuation process (typically) occurs via successive Compton scatters followed by photoelectric absorption. Unfortunately, the mean free path of gammas (in most materials) is close to maximal for  $E_\gamma \gtrsim 1$  MeV, and is characterized by an absorption length of 15–25 g/cm<sup>2</sup> (see, *e.g.*, Figure 27.16 in [88]). Except for the lowest- $Z$  materials, this absorption length is largely independent of  $Z$ . Consequently, for a given thickness, high-density materials provide the best shielding. Osmium, platinum, and gold all have densities of  $\sim 20$  g/cm<sup>3</sup>, but are rare and therefore prohibitively expensive. With a density of  $\sim 11$  g/cm<sup>3</sup>, lead is easily the most dense material that is readily available and relatively cheap. It is the most common material from which gamma shields are constructed.  $\sim 10$  cm of lead is sufficient to reduce the flux of dangerous gammas to  $\sim 1\%$  of the incident intensity.

Another motivation for choosing lead is that it can be found with relatively low levels of its own radioisotopes. Whenever materials are introduced in close proximity to a direct-detection target (usually for the purpose of shielding external radiation), one must consider their levels of radioactivity (a form of internal radioactivity). Most freshly mined (“virgin”) lead has negligible concentrations of dangerous isotopes. The exception is  $^{210}\text{Pb}$ . Nearly all radioactivity associated with virgin lead can be traced back to this isotope. The decays of  $^{210}\text{Pb}$  and its daughter isotopes result in the emission of x rays, electrons and alphas (see [605] and references therein for further details). Each of these forms of radiation are problematic due to their proximity to the dark-matter target. Despite their extremely small mean free path, the alphas are energetic enough to eject neutrons (discussed further below). The x rays and electrons are commonly shielded by an extra layer (internal to the main gamma shield) of even lower activity material. One op-

---

<sup>7</sup>Compton and Rayleigh scattering are similar processes that occur at different energy scales. In Compton scattering ( $1\text{ keV} \lesssim E_\gamma \lesssim 100\text{ MeV}$ ), gammas scatter incoherently from individual electrons in a material, typically imparting enough energy to cause ionization or even electron ejection. Ejected electrons are often referred to as “knock-on electrons” or delta rays. In Rayleigh scattering ( $E_\gamma \lesssim 10\text{ keV}$ ), gammas scatter coherently from entire atoms without causing ionization or excitation.



### CHAPTER 3. DARK MATTER DETECTION

tion is ancient lead that has been underground or underwater long enough for the  $^{210}\text{Pb}$  half-life of  $\sim 22$  years to have reduced the isotope to an immeasurable level. Unfortunately, ancient lead is rare, making it an expensive option if available. Sources include shipwrecked European galleons that have been underwater for several hundred years, and lead from ancient Roman aqueducts. Another low-activity material is electroformed copper. The density of copper is similar to lead, and it can be made extremely pure. For a comparison of copper and lead activities see [606]. The balance between shielding needs and cost usually results in a direct-detection experiment's gamma shield being constructed primarily from virgin lead, with an innermost layer of high-purity copper and possibly a thin layer of ancient lead as well. If a gamma shield's materials have been carefully screened to ensure low levels of radioactivity, an underground differential event rate on the order of 0.1–1 events/kg/keV/d is possible for  $E_R < 100\text{ keV}$  (see, *e.g.*, Figure 2 in [607]).

Perhaps the most difficult background to protect against is radon gas. One of the products of the  $^{238}\text{U}$  decay chain is  $^{222}\text{Rn}$ . Because of its  $\sim 4$  day half-life,  $^{222}\text{Rn}$  readily diffuses through the Earth's crust, and can accumulate in dangerous concentrations in underground laboratories. The concentration of  $^{210}\text{Pb}$  in virgin lead is due to the diffusion of  $^{222}\text{Rn}$  into the lead ore before it is mined. When the radon decays,  $^{218}\text{Po}$  gets embedded in the lead, and quickly decays (successively) through  $^{214}\text{Pb}$ ,  $^{214}\text{Bi}$ , and  $^{214}\text{Po}$  to  $^{210}\text{Pb}$ . Only lead that has been isolated (in a concentrated form) underground or underwater for hundreds of years is protected from its bulk being contaminated by this invasive  $^{222}\text{Rn}$  to  $^{210}\text{Pb}$  pathway. More troublesome is that the longest-lived product of radon decay,  $^{210}\text{Pb}$ , easily attaches itself to aerosols, and thus pervades the air in any laboratory environment. Unless particular care is taken to store dark-matter targets in radioisotope-free environments (*e.g.*, under a nitrogen purge), contact with laboratory air (both underground and at the surface) can lead to  $^{210}\text{Pb}$  surface contamination. Due to its invasive nature, radon gas is a form of internal radioactivity that all direct-detection experiments struggle to control.

### 3.3. DIRECT DETECTION

An experiment's innermost shielding layers and dark-matter target can also be contaminated by cosmogenically produced radioisotopes. This usually occurs when a material is unavoidably exposed to cosmic rays at sea level during fabrication, testing and shipping. For example, when copper is exposed to cosmic rays, long-lived cobalt isotopes (among others) can be produced via spallation. Spallation is a process in which a cosmic ray causes several nucleons to be ejected from a nucleus, often transmuting it into a lower- $Z$  or  $-A$  radioactive isotope (see, *e.g.*, Table III in [606] for a list of  $^{63}\text{Cu}$  spallation products). The isotopes produced in a dark-matter target obviously depend on its composition. All naturally occurring Ge isotopes (particularly  $^{70}\text{Ge}$ ), for example, are subject to transmutation to  $^{68}\text{Ge}$  and  $^{60}\text{Co}$  [608].  $^{68}\text{Ge}$  is particularly problematic because it is long lived ( $\sim 271$  day half-life), and emits either 10.4 or 1.3 keV in x rays or Auger electrons when it decays. Because it is embedded in the dark-matter target, these keV-scale radiations are unshielded and result in energy depositions similar to those expected for WIMPs. To minimize cosmogenic activation, dark-matter targets and shielding materials are usually stored underground when not in use. If necessary, materials can be produced (or grown) underground so as to never expose them to the hadronic component of the sea-level cosmic-ray flux.

#### Neutrons

Neutrons with MeV-scale kinetic energies will scatter from the nuclei in a dark-matter target to produce keV-scale energy depositions that are indistinguishable (on an event-by-event basis) from WIMP-nucleus scattering events. For this reason, fission neutrons from radioactive sources are an invaluable tool for calibrating a dark-matter target's response to nuclear-recoil energy depositions (discussed further in Chapter 6 and Appendix E). Lower- and higher-energy neutrons can also lead (albeit less directly) to keV-scale energy depositions. Like neutrinos, neutrons are an irreducible background for which there are several sources contributing to the flux in an underground laboratory.

### CHAPTER 3. DARK MATTER DETECTION

Unlike neutrinos, the neutron-nucleus interaction rate is too large to be ignored. Fortunately, there are a variety of techniques available for characterizing and shielding against this important background.

Neutrons are often classified according to their kinetic energy (or temperature). The nuclear-power industry has nearly a dozen such categories based on precise temperature delineations. For direct detection, however, neutrons naturally fall into one of three classes according to how they interact with the nuclei in a material: 1)thermal (or low-energy) neutrons will either scatter elastically or be captured, but have insufficient velocities to cause nuclear-recoil energies (in a dark-matter target) over threshold; 2)medium-energy neutrons tend to scatter elastically, and can cause energy depositions similar to those expected from WIMPs; and 3)fast (or high-energy) neutrons are most likely to interact inelastically, often causing the ejection of several lower-energy neutrons. Each category has the potential to contribute sensitivity-limiting background levels to a direct-detection experiment's event rate.

Technically, thermal neutrons have energies that correspond to room temperature— $kT$  of tens of meV or  $T \simeq 300$  K. The direct-detection community tends to use the term thermal (somewhat loosely) to refer to neutrons with energies less than  $\sim 10$  keV. Low-energy neutron is perhaps a more appropriate moniker. To understand the reason for the dividing line between low- and medium-energy neutrons, consider the recoil energy imparted by a neutron with energy  $E_n$  when it scatters elastically from a target with  $A$  nucleons:

$$E_R = \frac{2A}{(1+A)^2} E_n (1 - \cos \theta^*), \quad (3.3.38)$$

where  $\theta^*$  is the center-of-mass scattering angle.<sup>8</sup> Such an energy deposition can be confused with a WIMP-nucleus scattering event only if  $E_R > E_{\text{th}}$ . The maximum recoil energy that can be imparted to the lightest possible nuclear target ( $A = 1$ ) is 10 keV for the above definition of thermal. Consequently, thermal neutrons pose no (direct) danger to an experiment with

---

<sup>8</sup>Equation 3.3.38 is an approximation in which the neutron mass has been set equal to the average nucleon mass. The mass of an unbound neutron is  $\sim 1\%$  larger than  $m_N$ .

### 3.3. DIRECT DETECTION

a nominal 10 keV threshold. Furthermore, to take advantage of the  $A^2$  coherence factor, most direct-detection experiments employ large- $A$  targets. A 0.5 keV threshold would be needed, for example, before energy depositions from 10 keV neutrons could potentially be confused with WIMP events in a Ge target. Note that 10 keV need not be considered a rigid dividing line between low- and medium-energy neutrons. An alternate definition of low energy might be any neutrons incapable of depositing an over-threshold  $E_R$ , in which case the definition depends on an experiment's particular values of  $A$  and  $E_{\text{th}}$ .

Although thermal neutrons cannot directly cause WIMP-like energy depositions, they can be captured by nuclei in or near a dark-matter target to create radioactive isotopes. Neutron-capture cross sections for many nuclei compete with elastic scattering cross sections for  $E_n \lesssim 10$  keV, and often exhibit prominent resonances. There are a series of resonances, for example, in the  $^{70}\text{Ge}$  capture cross section for neutrons with 1–10 keV energies [609]. When a neutron is captured on  $^{70}\text{Ge}$ ,  $^{71}\text{Ge}$  is created.  $^{71}\text{Ge}$  is an unstable isotope that decays in the same manner as the cosmogenically produced  $^{68}\text{Ge}$  (described above), but with a shorter half-life of  $\sim 11$  days. The process of rendering a material radioactive through exposure to thermal neutrons is called neutron activation. Care must be taken to prevent a dark-matter target from becoming too activated (through overexposure) when using a neutron source to calibrate its nuclear-recoil response. The consequences of overexposure are explored further in the final two chapters of this thesis.

The sources of low- and medium-energy neutrons are basically the same: cosmic rays, muon capture, and radioactivity. As shown in Figure 3.15, cosmic rays are the dominant source of neutrons at sea level. Note that the cosmic-ray neutron flux has been divided into two categories here: thermal and fast. Since a modest overburden of  $\sim 10$  m.w.e. is sufficient to reduce the flux of cosmic-ray neutrons to trivial levels, they are a concern primarily with respect to exposure of a dark-matter target during production, testing, and handling, *i.e.*, before the target can be moved underground. In this case, two categories

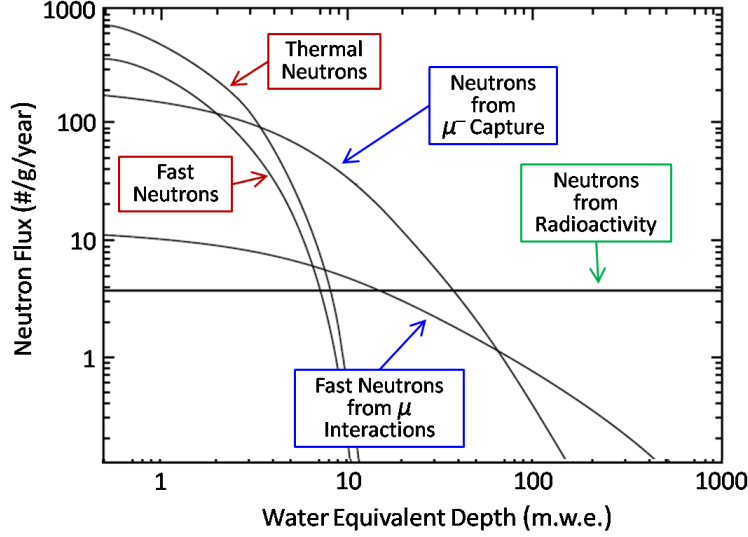


Figure 3.15: Neutron flux subdivided according to source as a function of depth. At sea level, the thermal and fast cosmic-ray neutron fluxes (dark red labels) are dominant. At depths  $\gtrsim 5$  m.w.e., neutrons resulting from muon interactions in the overburden (blue labels) are more prevalent, while the flux of MeV-scale neutrons from natural radioactivity (green label) is roughly constant as a function of depth, making it the dominant source of neutrons for depths  $\gtrsim 100$  m.w.e. Figure adapted from [604].

are sufficient to describe the danger to a material. Thermal neutrons can lead to activation via neutron capture, while fast neutrons can cause transmutation via spallation. Both processes often produce long-lived radioisotopes that will limit an experiment’s sensitivity if efforts are not taken to minimize sea-level exposure.

For laboratories at intermediate depths (tens of m.w.e.), the dominant source of neutrons is from the capture of slow, negatively charged muons on nuclei in the overburden. When the soft component of the sea-level muon flux ranges out in the overburden (sometimes referred to as “stopping muons”), the positively charged fraction will decay as free muons ( $\mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu$ ), while the negative muons will very quickly ( $\sim 10^{-13}$  s) become captured into bound

### 3.3. DIRECT DETECTION

states with nuclei in the rock [610]. Based on data derived from the study of  $\mu^-$  capture in noble-gas mixtures, the probability for capture on a particular nucleus appears to scale in proportion to the product of its number density and  $Z^{2/3}$  [611]. When a  $\mu^-$  is captured, it orbits (as an electron would) for a short time before it either decays (as a free muon would:  $\mu^- \longrightarrow e^- + \bar{\nu}_e + \nu_\mu$ ) or cascades in toward the nucleus to convert a proton to a neutron [612]:

$$\mu^- + p \longrightarrow n + \nu_\mu. \quad (3.3.39)$$

Because the rate of the latter process competes with the free-decay rate, the mean life of a captured muon is shorter than that of a free muon. Furthermore, the more protons there are in a nucleus, the more likely it is that a captured muon will interact with a proton before it can decay. For example, the mean life of a muon captured by a hydrogen nucleus ( $\sim 2195$  ns) is only a few nanoseconds shorter than for a free muon ( $\sim 2197$  ns), while for an iron nucleus it is an order of magnitude smaller ( $\sim 200$  ns) [613]. The most common materials found in rock are oxygen, silicon and aluminum, for which the captured-muon life times are  $\sim 1800$  ns,  $\sim 800$  ns, and  $\sim 900$  ns, respectively. Consequently, a fair fraction of negatively charged stopping muons result in the conversion of a proton to a neutron.

The neutrons resulting from muon capture are often created with significant kinetic energy. The capturing nucleus dispels this energy in one of two ways, depending on how deep into the nucleus the proton conversion occurred. In less than a quarter of muon-capture proton conversions, the neutron is created at the nuclear surface and is directly ejected. These neutrons are characterized by an exponential differential number distribution with a 7–15 MeV decay constant [604]. Consequently, a small fraction of muon captures result in the ejection of high-energy neutrons ( $E_n \gtrsim 10$  MeV). Most neutrons created through muon capture, however, are generated inside the capturing nucleus where the neutron's energy is shared with the other nucleons. In this case, the nucleus de-excites via neutron evaporation, a process in which several MeV-scale neutrons effectively boil off the nucleus. Because this is the dominant production mechanism, neutrons resulting from  $\mu^-$  capture are primarily medium energy.

### CHAPTER 3. DARK MATTER DETECTION

At depths  $\gtrsim 100$  m.w.e., natural radioactivity is the principle neutron source. Neutrons are created when heavy elements undergo spontaneous fission (SF), and in  $(\alpha, n)$  reactions. The  $^{238}\text{U}$  and  $^{232}\text{Th}$  decay chains are the main sources of both. The only naturally occurring elements that undergo SF at a sufficient rate to be observed are  $^{238}\text{U}$ ,  $^{235}\text{U}$  and  $^{232}\text{Th}$ . Their SF branching fractions are incredibly small, and amount to the following half-lives [614]:

$$\begin{aligned}\tau_{1/2\text{-SF}}(^{238}\text{U}) &= 8.2 \times 10^{15} \text{ years}, \\ \tau_{1/2\text{-SF}}(^{235}\text{U}) &= 1.03 \times 10^{19} \text{ years, and} \\ \tau_{1/2\text{-SF}}(^{232}\text{Th}) &= 1.2 \times 10^{21} \text{ years.}\end{aligned}\tag{3.3.40}$$

$^{238}\text{U}$  is therefore the dominant (natural) source of fission neutrons. Each  $^{238}\text{U}$  SF creates  $\sim 2$  free neutrons [615], such that 1 kg of natural uranium emits  $\sim 16$  neutrons/s [604]. For rock (or shielding materials) with a 1 ppm uranium impurity, this translates to a few neutrons per year per gram.

The other source of neutrons from radioactivity is less direct, but of greater concern. In an  $(\alpha, n)$  process, an energetic alpha particle (from the decay of some isotope) interacts with a nearby nucleus to eject one or more neutrons (usually no more than 2). For most nuclei found in rock, the threshold energy for an  $(\alpha, n)$  reaction to proceed is larger than the energy of the most energetic alpha available from natural radioactivity—the decay of  $^{212}\text{Po}$  (near the end of the thorium chain) produces an 8.8 MeV alpha. For example, the  $(\alpha, n)$  thresholds for  $^{28}\text{Si}$  and  $^{16}\text{O}$  are  $\sim 11$  and 15.5 MeV, respectively [616]. However, some common elements have lower-energy thresholds (the  $^{27}\text{Al}$  threshold, *e.g.*, is  $\sim 3.8$  MeV), and are thus susceptible to  $(\alpha, n)$  reactions. The neutrons from  $(\alpha, n)$  reactions outnumber those from  $^{238}\text{U}$  SF by nearly an order of magnitude for some rock compositions (see, *e.g.*, Table 4.5 in [604]).  $(\alpha, n)$  neutrons from an underground laboratory’s cavern walls are not the only concern. As mentioned above, the presence of  $^{210}\text{Pb}$  on the surfaces of a dark-matter target and its innermost shielding layers is difficult to prevent due to the invasive nature of radon gas. The  $^{210}\text{Pb}$  decay sequence includes the  $\sim 5$  MeV alpha decay of  $^{210}\text{Po}$ . 5 MeV is sufficient to instigate  $(\alpha, n)$  reactions on three of the five naturally occurring isotopes in a Ge target.

### 3.3. DIRECT DETECTION

Neutrons from radioactivity and  $\mu^-$  capture primarily fall into the medium-energy category, and are therefore capable of causing over-threshold recoil energies in a direct-detection target. If left unshielded, these medium-energy neutrons would quickly limit an experiment's WIMP sensitivity. Equation 3.3.38 holds the key to shielding this background. On average, a neutron loses  $2A/(A+1)^2$  of its initial energy following an elastic scattering event. Clearly, energy loss is optimized by minimizing  $A$ . When neutrons elastically scatter with hydrogen nuclei, for example, they lose half their energy (on average). When they scatter from lead nuclei, on the other hand, the average energy loss is  $\lesssim 1\%$ . Consequently, neutron shields are typically constructed from materials rich in hydrogen (*e.g.*, water and plastics). Consider, for example, a polyethylene shield. Polyethylene is a high-density plastic created through polymerization of ethylene ( $\text{C}_2\text{H}_4$ ). It is relatively cheap, radiopure, hydrogen rich, and easy to work with. The mean free path of a neutron in polyethylene is

$$\lambda = \left[ \rho_{\text{poly}} \frac{N_A}{\mathcal{M}_{\text{ethene}}} (n_C \sigma_C + n_H \sigma_H) \right]^{-1}, \quad (3.3.41)$$

where  $n_H = 4$  ( $n_C = 2$ ) is the number of H (C) atoms in an ethylene molecule,  $\mathcal{M}_{\text{ethene}} = 14 \text{ g/mole}$  is the molecular weight of ethylene,  $\rho_{\text{poly}} \simeq 1 \text{ g/cm}^3$  is the density of polyethylene, and  $\sigma_H$  ( $\sigma_C$ ) is the cross section for neutrons to scatter elastically from H (C). Since carbon is a much less effective moderator than hydrogen, its contribution can be safely dropped from Equation 3.3.41 for our purposes here. The elastic scattering cross section for a 1 MeV neutron on hydrogen is  $\sim 4 \times 10^{-24} \text{ cm}^2$  [609]. The average distance a 1 MeV neutron is likely to travel in polyethylene before encountering a hydrogen atom and losing half its energy is therefore  $\sim 3 \text{ cm}$ .  $\sigma_H$  is  $\sim 50\%$  larger for a 500 keV neutron, and twice as large for a 250 keV neutron. The neutron elastic scattering cross section continues to increase as  $E_n$  decreases until it reaches a value of  $20 \times 10^{-24} \text{ cm}^2$  for  $E_n \simeq 10 \text{ keV}$ . Taking into account the corresponding decrease in the mean free path, a 1 MeV neutron will scatter (on average) 7 times before its energy falls below 10 keV, during which time it will have penetrated an average of  $\sim 10 \text{ cm}$  of polyethylene. Since the spectrum of neutron



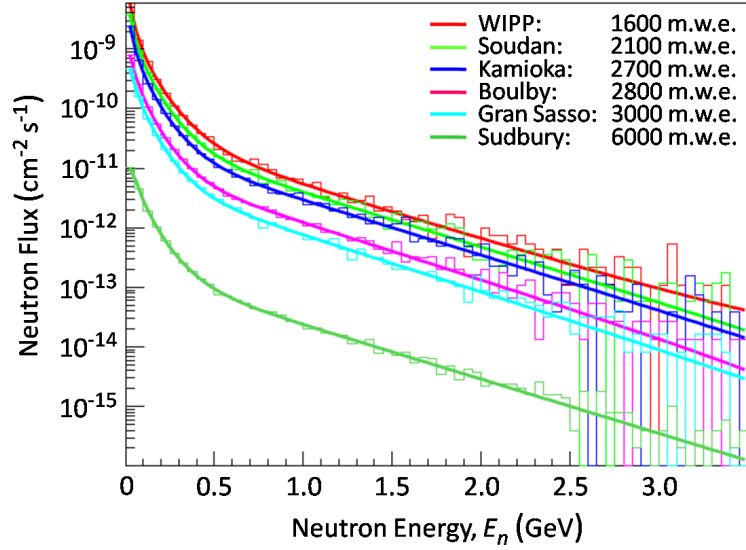


Figure 3.16: The flux of high-energy neutrons expected (from simulations) for several underground laboratories of varying depth, as a function of neutron energy. Because the production of high-energy neutrons is tied to the muon flux, which decreases with increasing overburden (see, *e.g.*, Figure 3.13), the absolute flux of high-energy neutrons in underground laboratories also decreases with increasing depth. The depth of each underground facility is indicated in m.w.e. in the legend. Figure adapted from [617].

energies from radioactivity extends to several MeV, a 20–30 cm-thick polyethylene neutron shield (or slightly thicker water shield) is a common feature of a direct-detection experiment.

The flux of high-energy neutrons is subdominant at all depths. However, no practical amount of hydrogen-rich neutron shielding can protect an experiment from this component of the neutron background. High-energy neutrons are a product of hadronic and electromagnetic showers initiated by cosmic-ray muons in a laboratory’s overburden, or in an experiment’s shielding materials. Those created in shielding materials are easily tagged by a muon veto due to the passage of the parent muon. Many created in the rock just above an experiment are also tagged by an active veto because they are often accompanied by electri-

### 3.3. DIRECT DETECTION

cally charged shower fragments. There remains an incident flux, however, that can evade detection by active shielding elements. High-energy neutrons punch through materials much like a rock thrown into a swimming pool. When they interact, they tend to scatter inelastically and can cause the ejection of several lower-energy neutrons from high- $Z$  materials like lead. Consequently, high-energy neutrons can penetrate an active muon veto undetected, punch through an experiment's passive neutron shielding, and generate dangerous MeV-scale secondary neutrons (near the dark-matter target) when they finally interact with an experiment's gamma shield. In this sense, the underground differential flux of high-energy neutrons shown in Figure 3.16 represents an irreducible background. The figure demonstrates the only practical reduction method: increase the overburden by moving to a deeper laboratory. Fortunately, neutrons tend to scatter multiple times in a dark-matter target, whereas WIMPs will scatter only once. A segmented (or multi-module) detector therefore has a fair chance of rejecting events induced by high-energy neutrons. Furthermore, by simulating the ratio of single- to multiple-scatter neutron events expected for a particular geometry, the singles rate can be measured from the multiples rate. If a nonzero singles rate is observed, the neutron component can be subtracted using the multiple-scatter event rate. The current sensitivity of direct-detection experiments is such that a few thousand m.w.e. of overburden is sufficient to reduce the high-energy neutron background to a negligible level. If the next generation of experiments is to be background-free, however, they might have to instrument their neutron shields to actively detect (and veto) neutron secondaries, thereby taking advantage of the tendency for high-energy neutrons to create multiple secondaries that tend to multiple scatter.

#### 3.3.4 Discrimination

Despite the best efforts to shield direct-detection targets from the backgrounds described above, some level of incident radiation is unavoidable. One of the best shielded experiments ever conducted, for example, was the Heidelberg Dark Matter Search [618] (HDMS), in which a differential event rate

### CHAPTER 3. DARK MATTER DETECTION

$\lesssim 0.5 \text{ keV}^{-1} \text{ kg}^{-1} \text{ d}^{-1}$  was observed (for  $E_R \gtrsim 10 \text{ keV}$ ) with a  $^{73}\text{Ge}$  enriched ionization detector (see, *e.g.*, Figure 10 in [619]). Most experiments are subject to similar (or worse) event rates due to unshielded backgrounds. In the case of HDMS, the observed spectrum—attributed to backgrounds—was modeled and subtracted to improve the experiment’s WIMP sensitivity. Unfortunately, background models are accompanied by systematic errors that eventually limit the improvement gained through subtraction (despite increased exposure). To continue to probe ever smaller WIMP-nucleon cross sections, an alternate approach to dealing with backgrounds is necessary; the passive and active shielding techniques described above are inadequate if the WIMP-nucleus scattering rate turns out to be less than about one event per kilogram-day of exposure.

Unlike WIMPs, most backgrounds—photons, alphas, electrons and positrons—will scatter from a target’s electrons. Therein lies the key to rejecting them. To further reduce backgrounds, many experiments actively discriminate nuclear recoils from electron recoils. The most common discrimination techniques involve simultaneously measuring energy via two methods, at least one of which behaves differently for electron and nuclear recoils. Two of the following three are often measured in tandem:

- **Phonons:** Solid dark-matter targets usually exhibit a highly regular crystalline structure as a result of their purity. When a particle interacts with a target it causes vibrations, and in a crystal these vibrations occur in quantized units known as phonons. Phonons can either be detected directly when they interact with specialized electrodes at a detector’s surfaces (discussed further in the next chapter), or indirectly with thermometers when they thermalize and cause the temperature of the crystal to increase. In either case, for a given amount of deposited energy, a phonon (or heat) measurement provides an equal measure of the recoil energy for electron and nuclear recoils.
- **Ionization:** A fraction of the energy deposited when a particle interacts with a material goes into the creation of electron-hole pairs, or ionization. If an electric field is applied across the detector, the electrons and holes

### 3.3. DIRECT DETECTION

can be drifted to opposite sides and detected with standard ionization electrodes. Electron recoils cause more ionization than nuclear recoils for a given amount of deposited energy. If a WIMP deposits 50 keV in a Ge target, for example, it is expected to cause  $\sim 1/3$  the ionization signal a 50 keV x ray would cause.

- **Scintillation:** In certain materials, a small amount of the energy deposited during a scattering event goes into the generation of scintillation light, which can be detected with standard photon detectors (*e.g.*, photomultiplier tubes). Generally, for a given amount of deposited energy, electron recoils result in more scintillation light than nuclear recoils.

Instrumenting a detector to measure any two of the above provides an experiment with an event-by-event discrimination parameter. The CRESST II experiment mentioned in Section 3.3.2, for example, simultaneously measures a phonon and scintillation signal for each scattering event. The ratio of the scintillation to phonon energies (“light yield”) is larger for electrons recoils than for nuclear recoils. In the light-yield versus recoil-energy plane, x-ray and neutron calibration data can be used to construct electron- and nuclear-recoil bands, respectively. Events due to WIMPs will enter into the nuclear-recoil band, while events due to (most) backgrounds will enter into the electron-recoil band. An example “discrimination plot” is shown in Figure 3.17 for CRESST II calibration data. The separation between electron and nuclear recoils is close to perfect. Through a combination of shielding and discrimination, the expected number of background events in the WIMP signal region is nearly zero. Experiments that use discrimination can be effectively background-free despite a nonzero flux of unshielded radiation incident upon their targets.

Another promising method for discriminating backgrounds is to construct a detector that is insensitive to electron recoils. This is sometimes referred to as a threshold detector. The discrimination principle relies on the tendency for nuclear recoils to deposit energy over shorter distances than electron recoils, *i.e.*,  $dE/dx$  for nuclear recoils is larger than for electron recoils. Some mate-

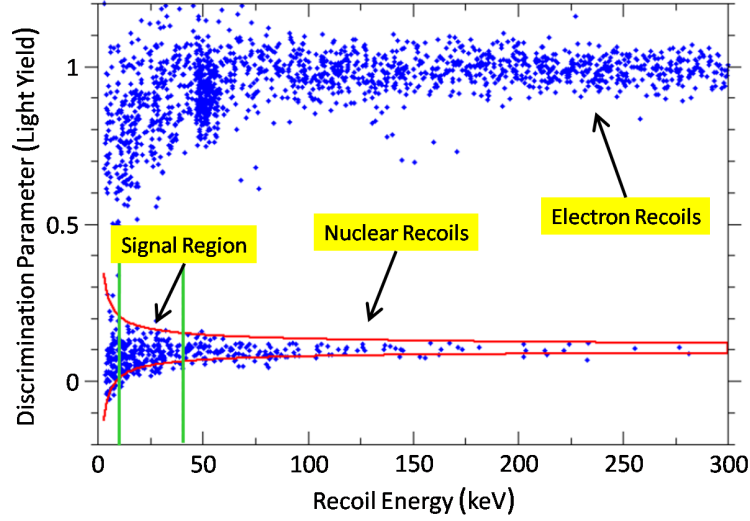


Figure 3.17: An example electron- versus nuclear-recoil discrimination plot, in which light yield (scintillation to phonon ratio) is plotted as a function of recoil energy for calibration data (blue dots) from the CRESST II experiment. A nuclear-recoil band (red/horizontal lines) indicates where 80% of nuclear-recoils due to neutrons scattering from oxygen nuclei are expected. As indicated, electron-recoil backgrounds are well separated from nuclear-recoil events by the light-yield discrimination parameter. The vertical lines (green), together with the nuclear-recoil band, indicate the low-energy region where WIMP events are expected to occur. In the absence of a neutron source, this WIMP-signal region is almost background-free. Figure adapted from [587].

rials can be tuned (via careful pressure, temperature, and/or volume control) to be insensitive to particle interactions in which  $dE/dx$  is too small. The superheated  $\text{CF}_3\text{I}$  used by the Chicagoland Observatory for Underground Particle Physics [620] (COUPP) is one such compound. When a scattering event with sufficient  $dE/dx$  occurs in the superheated fluid, the recoiling nucleus acts as a nucleation site for the growth of a bubble. Bubbles from potential WIMP-nucleus scatters are recorded by monitoring the fluid with video cameras through a transparent container. In such a detector, photons and electrically charged backgrounds need not be shielded because they cannot cause sufficient

### 3.3. DIRECT DETECTION

$dE/dx$  to trigger bubble growth. A significant overburden and neutron shielding is still necessary to reduce the neutron background. Also, care must be taken to avoid contamination of the target material with alpha-emitting radioisotopes. Energetic alpha decays can cause the daughter nucleus to recoil with sufficient energy to cross the  $dE/dx$  threshold. There are currently three experiments that use this discrimination technique to directly search for dark matter. In addition to COUPP, there is a superheated  $C_4F_{10}$  droplet detector called the Project In Canada to Search for Super-Symmetric Objects [621] (PICASSO), and there is a  $C_2ClF_5$  detector called the Superheated Instrument for Massive Particle Experiments [622] (SIMPLE).

There are far too many direct-detection experiments and discrimination techniques to present a comprehensive review here. However, they tend to fall into one of three basic categories: liquid-noble targets, solid-state (crystalline) detectors, and superheated droplet detectors. Most, though not all, use some form of discrimination. A summary of several of the most sensitive discrimination technologies is given in Table 3.2. Experiments that do not use discrimination accept relatively large background rates (usually) in favor of low recoil-energy thresholds, and the corresponding enhanced possibility of observing the annual modulation expected in the WIMP-nucleus scattering rate due to the Earth's motion around the Sun (see, *e.g.*, Section 3.3.5 below). Although such an experiment is technically background-limited, the small modulated event rate expected from WIMPs can still be discerned provided the backgrounds do not also modulate.

Before moving on to the next section, it is worth noting a potentially powerful method for unambiguously demonstrating the Galactic-WIMP origin of a nuclear-recoil signal in the presence of a nonzero event rate due to backgrounds. The technique involves measuring the direction of the recoiling nucleus following a scattering event, thereby measuring the incoming direction of the incident particle. Due to the Earth's daily rotation with respect to the WIMP wind, the rate of nuclei recoiling in certain directions (with respect to the Earth's equator) is expected to modulate by nearly an order of magnitude every 12 hours [641]. This is a much larger signal than the annual modulation, and

### CHAPTER 3. DARK MATTER DETECTION

Table 3.2: The discrimination techniques used by several of the world’s most sensitive direct-detection experiments. “Charge” and “Light” are abbreviations for ionization and scintillation, respectively. Sensitivities are quoted in terms of the indicated 90% C.L. cross-section upper limit (in pb) at the indicated WIMP mass (in GeV), and are grouped according to interaction type.

Experiment	Target–Discrimination	Sensitivity (pb, GeV)
<i>Spin Independent, <math>\sigma_{\chi-N}</math> (nucleon)</i>		
CDMS II [623]	Ge–Charge/Phonon	$3.8 \times 10^{-8}$ , 70 [459]
COUPP [620]	CF <sub>3</sub> I– $dE/dx$	$\gtrsim 1 \times 10^{-7}$ , 60 [624]
CRESST II [591]	CaWO <sub>4</sub> –Light/Phonon	$\sim 6 \times 10^{-7}$ , 40 [625]
EDELWEISS II [626]	Ge–Charge/Heat	$4.4 \times 10^{-8}$ , 85 [627]
WARP [598]	Ar–Charge/Light	$\sim 7 \times 10^{-7}$ , 90 [628]
XENON100 [629]	Xe–Charge/Light	$7.0 \times 10^{-9}$ , 50 [630]
ZEPLIN III [631]	Xe–Charge/Light	$8.1 \times 10^{-8}$ , 60 [632]
<i>Spin Dependent, <math>\sigma_{\chi-p}</math> (proton)</i>		
COUPP [620]	CF <sub>3</sub> I– $dE/dx$	$\sim 5 \times 10^{-3}$ , 50 [624]
KIMS [633]	CsI(Tl)–Pulse Shape	$\lesssim 2 \times 10^{-1}$ , 80 [634]
PICASSO [621]	C <sub>4</sub> F <sub>10</sub> – $dE/dx$	$1.6 \times 10^{-1}$ , 24 [635]
SIMPLE [622]	C <sub>2</sub> ClF <sub>5</sub> – $dE/dx$	$4.2 \times 10^{-3}$ , 35 [636]
<i>Spin Dependent, <math>\sigma_{\chi-n}</math> (neutron)</i>		
CDMS II [623]	Ge–Charge/Phonons	$\lesssim 2 \times 10^{-2}$ , 50 [637]
XENON10 [638]	Xe–Charge/Light	$\sim 6 \times 10^{-3}$ , 30 [639]
ZEPLIN III [631]	Xe–Charge/Light	$1.9 \times 10^{-2}$ , 55 [640]

### 3.3. DIRECT DETECTION

there are no known backgrounds that can mimic it. Unfortunately, the ability to measure tracks caused by recoiling nuclei has yet to be proven feasible on a scale that is competitive with other technologies. For a recent review of directional dark-matter detection, I refer the interested reader to [642].

#### 3.3.5 Low-mass WIMP Signals?

The direct-detection literature has been the focus of considerable debate over the past few years. The excitement (or controversy) stems from data recorded by two experiments that appear to be mutually consistent with a Galactic-WIMP interpretation. Due to its low mass, the preferred WIMP model (just) evades nearly all the experimental constraints listed in Table 3.2. The favored  $\sim 7 \text{ GeV}/c^2$  mass is also consistent with the WIMP model proposed by Hooper and Goodenough to explain an excess in the flux of gamma rays from the Galactic center (see Section 3.2.2 above). Furthermore, the favored cross section is roughly consistent with the lowest-mass neutralinos allowed by the LNM framework described in Section 2.5.4 (see, *e.g.*, Figure 2.19). Although this confluence of experimental data and theory is provocative, it appears that most of the community is withholding judgment until a more sensitive experiment can test the favored parameter space. This thesis is an attempt at such a test. In the remainder of this section I describe the hunting ground by reviewing evidence for a low-mass WIMP from the DAMA/LIBRA and CoGeNT direct-detection experiments.

#### The DAMA/LIBRA Annual Modulation

The first hint of a Galactic-WIMP direct detection actually goes back many years to early data from one of the Dark Matter (DAMA) group's experiments at the Laboratori Nazionali del Gran Sasso (LNGS). The DAMA group has used a number of different scintillation targets over the past two decades to investigate dark matter [643, 644],  $\beta\beta$  decay [645, 646], and other rare-event processes [647, 648]. The data from their 100 kg NaI(Tl) array [649] (commonly referred to as DAMA/NaI) was first analyzed in terms of an annual



### CHAPTER 3. DARK MATTER DETECTION

modulation in a 1998 publication [650]. The setup consisted of a dozen rectangular, high-purity NaI crystals doped with Tl to improve their scintillation-photon yield for the low recoil energies expected from WIMP interactions [651]. Located  $\sim 3000$  m.w.e. underground, the crystals—monitored by two photomultiplier tubes each—were surrounded by an outer neutron shield (10 cm of polyethylene around 1.5 mm of Cd) and an inner gamma shield (15 cm Pb around 10 cm of Cu) to reduce the event rate due to backgrounds to  $\sim 1$ – $2$  events/kg/keV/d. Although their experimental strategy did not include electron-recoil discrimination, it benefited from a large detector size and a relatively low 2 keVee energy threshold. Note that 2 keVee is an electron-equivalent recoil energy, and is not the same as the nuclear-recoil energy that a WIMP would deposit. For a given energy deposition, electron recoils generate more scintillation light than nuclear recoils. To convert to nuclear-recoil energy, the number of keVee must be divided by the quenching factor for either Na or I. Although there is some uncertainty in these numbers (see, *e.g.*, Section II in [652]), the most commonly used quenching factors are  $Q_{\text{Na}} = 0.3$  and  $Q_{\text{I}} = 0.09$  [649], corresponding to nuclear-recoil energy thresholds of 6.7 keV and 22.2 keV for Na and I, respectively.<sup>9</sup> Their lowest-energy data (2–12 keVee) from the first year of operations exhibits a marginal hint of a modulation riding on top of their total event rate. This modulation was confirmed with an additional year of exposure [653], and turns out to have the phase expected for the Galactic-WIMP annual-modulation signature. The two-year modulated event rate of  $\sim 0.015 \pm 0.01$  events/kg/keV/d (between 2 and 4 keVee), when interpreted in terms of WIMPs scattering from iodine nuclei, favors  $m_\chi \simeq 60 \text{ GeV}/c^2$  and  $\sigma_{\chi-N} \simeq 7 \times 10^{-6} \text{ pb}$  (see, *e.g.*, Figure 6 in [653]).

---

<sup>9</sup>Throughout Section 3.3 I have somewhat loosely interchanged the terms *recoil energy* and *nuclear-recoil energy*, and quoted any related numerical values in units of keV. Up to this point, there has been no need for further clarification because the intended quantity was always nuclear-recoil energy. To clearly differentiate between nuclear-recoil and electron-equivalent recoil energy, units of keVr (or keVnr) are often used for the former, while units of keVee are used for the latter. In this thesis, unless otherwise stated or obvious from context, units of keV are used for nuclear-recoil energy, while units of keVee are used for electron-equivalent recoil energy.

### 3.3. DIRECT DETECTION

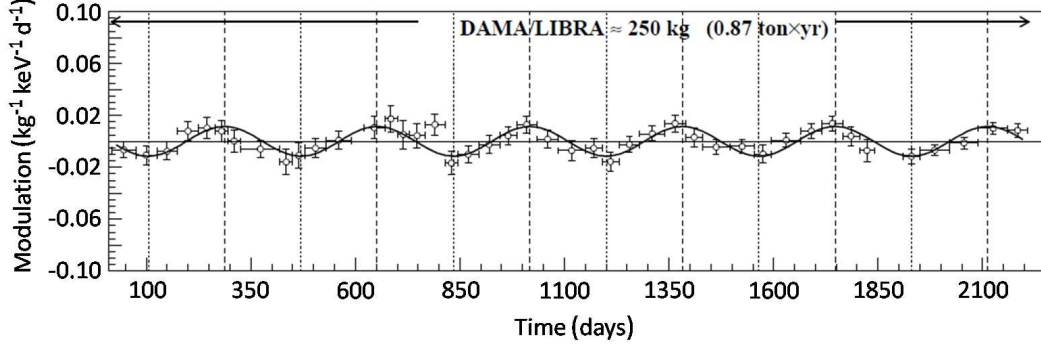


Figure 3.18: Annual-modulation fit to the DAMA/LIBRA residual event rate for 2–6 keVee energy depositions occurring over a six-year period. The fit has the phase expected for the Galactic-WIMP annual modulation signature, leading many to interpret the  $\sim 0.01$  events/kg/keV/d modulation amplitude as evidence of the first dark-matter direct detection. Figure taken from [657].

Such a large cross section for a  $60 \text{ GeV}/c^2$  WIMP is in direct conflict with nearly all the spin-independent constraints listed in Table 3.2. To reconcile their modulation signal with the null results of other experiments, the DAMA collaboration have explored alternate halo models, scattering from Na nuclei, spin-dependent scattering, and ion channeling (see, *e.g.*, [654] and [655]). In their presentation of the final 7-year DAMA/NaI  $\sim 108,000 \text{ kg d}$  exposure, they interpreted their  $\sim 0.02 \text{ event/kg/keV/d}$  modulation amplitude (in the 2-6 keVee bin) under a wide variety of scenarios [656]. For spin-independent couplings, they derived an allowed region in the  $m_\chi - \sigma_{\chi-N}$  plane that extends from  $3 \times 10^{-5}$  to  $10^{-3} \text{ pb}$  at  $m_\chi = 4 \text{ GeV}/c^2$ , from  $2 \times 10^{-7}$  to  $6 \times 10^{-5} \text{ pb}$  at  $m_\chi = 30 \text{ GeV}/c^2$ , and from  $5 \times 10^{-6}$  to  $1.5 \times 10^{-5} \text{ pb}$  at  $m_\chi = 500 \text{ GeV}/c^2$  (Figure 9 in [656] shows the full region). The low-mass portion of this allowed region is the most difficult to test, and continues to evade experimental constraints.

To confirm and better measure their annual-modulation signal, the DAMA collaboration constructed a larger detector called the Large sodium Iodide Bulk for RAre processes [583] (usually referred to as DAMA/LIBRA). The new setup makes use of the same shielding configuration at the same deep-site installa-

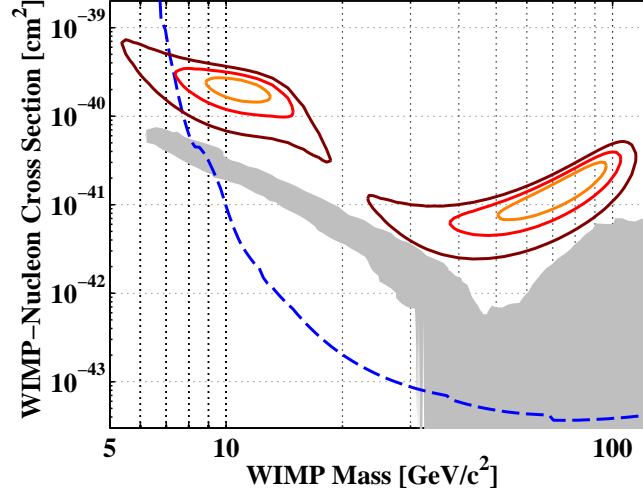


Figure 3.19: Allowed regions (solid contours) corresponding to the modulated event rate observed in the combined DAMA/NaI and DAMA/LIBRA  $\sim 1.2$  ton-year exposure (as interpreted by Savage *et al.* in [659]), compared to the CDMS II limit [459] (90% C.L., blue dashed line) and LNM calculations [468] (90% C.L., gray shaded region). The different contours correspond to the 90% (orange),  $3\sigma$  (red), and  $5\sigma$  (dark red) confidence levels.

tion, but employs 25 higher-purity NaI(Tl) crystals for a total target mass of  $\sim 250$  kg. They have published data from this new apparatus twice, once after the first four years of operation in 2008 [658], and again in 2010 to report on a total of six annual cycles [657]. A fit to their 6-year residual event rate is shown in Figure 3.18. The combined DAMA/NaI and DAMA/LIBRA data set represents an exposure of  $\sim 1.2$  ton-years, spans 13 annual cycles, and exhibits a self-consistent annual modulation (both in phase and amplitude) for 2–6 keVee energy depositions. The main difference between the DAMA/NaI and DAMA/LIBRA modulation signals is that the amplitude of the latter is roughly half as large as the former. The difference is consistent with an  $\sim 2\sigma$  statistical fluctuation.

While each of the DAMA/NaI publications includes at least one figure in which the modulated event rate is converted to an allowed region in the  $m_\chi$ –

### 3.3. DIRECT DETECTION

$\sigma_{\chi-N}$  plane, the two DAMA/LIBRA papers leave interpretation in terms a particular framework up to the community. For the standard framework described in Section 3.3.2, Savage and collaborators have published perhaps the most conservative (and widely used) DAMA allowed regions [660, 661, 659]. The results from their spin-independent analysis of the combined DAMA/NaI and DAMA/LIBRA 1.2 ton-year exposure are shown in Figure 3.19. The higher-mass region associated with WIMPs scattering from iodine nuclei is ruled out by the CDMS II limit. It is the lower-mass region, corresponding to scattering from Na nuclei, that warrants further investigation. The tension between the CDMS II limit and this low-mass region is clearly severe. As pointed out in [662], however, an energy-dependent Na quenching factor (as in [663]) can move the favored region to even lower masses.

The DAMA annual modulation has also been interpreted under alternative hypotheses. The most common involves the same standard framework applied to spin-dependent WIMP-nucleon interactions. As in the spin-independent case, there are low- and high-mass allowed regions corresponding to scattering from Na and I nuclei, respectively (see, *e.g.*, Figures 6 and 7 in [660]). For the case of pure neutron spin-dependent couplings, the XENON10 [639] and CDMS II [637] experiments basically rule out both regions, while for pure proton couplings, the favored parameter space is excluded by data from the SIMPLE experiment [636]. As with spin-independent couplings, uncertainty in the Na and I quenching factors could shift the DAMA regions to lower WIMP masses, thereby relaxing the apparent tension with the null results of other experiments. More exotic possibilities such as leptophilic [664, 665] and inelastic [666, 667] dark matter have also been explored.

#### The CoGeNT Excess

As the name implies, the original intent of the Coherent Germanium Neutrino Technology [584] (CoGeNT) was to establish a compact alternative (to large-volume observatories) for the detection of neutrinos. The design employs a p-type point contact (PPC) electrode attached to a high-purity Ge

### CHAPTER 3. DARK MATTER DETECTION

diode to achieve an unprecedented combination of low-energy resolution and detection threshold for a kg-mass ionization detector. Such a detector has the potential to detect the flux of antineutrinos from nuclear reactors at a rate of hundreds of events per day provided the event rate due to backgrounds can be reduced to  $\lesssim 1$  event/kg/keV/d for a sufficiently low threshold (see, *e.g.*, Figure 5 in [584]).

The same features that make the CoGeNT a good neutrino detector, also make it a viable tool for the direct detection of Galactic WIMPs. An  $\sim 1/2$  kg prototype was constructed and installed at a shallow site (330 m.w.e.) in Chicago, and achieved a threshold of 0.4 keVee during an 8.4 kg d exposure [668]. The Ge quenching factor—sometimes referred to as ionization yield and discussed further in the next chapter—is  $\sim 0.2$  at 0.4 keVee (see, *e.g.*, Figure 4 in [669]). The threshold for nuclear-recoil energy depositions due to WIMPs was therefore 2 keV. Thanks to this low threshold, the relatively modest shallow-site exposure of the CoGeNT prototype was sufficient to exclude most of the low-mass WIMP parameter space favored by the DAMA annual modulation, even despite a near-threshold differential event rate of  $\sim 50$  events/kg/keV/d (see, *e.g.*, Figure 2 in [668]).

The background levels observed with the prototype prompted the CoGeNT collaboration to construct a similar PPC Ge diode, but with particular care taken to reduce internal sources of radioactivity. The new 440 g detector was installed at the Soudan Underground Laboratory ( $\sim 2100$  m.w.e. overburden) inside a passive shielding configuration similar to the shallow-site setup: an inner gamma shield consisting of 15 cm of normal lead exterior to 5 cm of low-activity lead, and an outer neutron shield consisting of 30 cm of polyethylene exterior to 0.5 cm of borated neutron absorber [670]. While the setup also includes a plastic-scintillator muon veto, it was not used to reject candidate events because events (in the Ge target) correlated with veto activity (apparently) do not accumulate in excess at the low energies of interest for a WIMP search. The differential spectrum resulting from a 56 day exposure is shown in Figure 3.20. The observed events are due primarily to decays of cosmogenically activated radioisotopes in the Ge target. Relative to the

### 3.3. DIRECT DETECTION

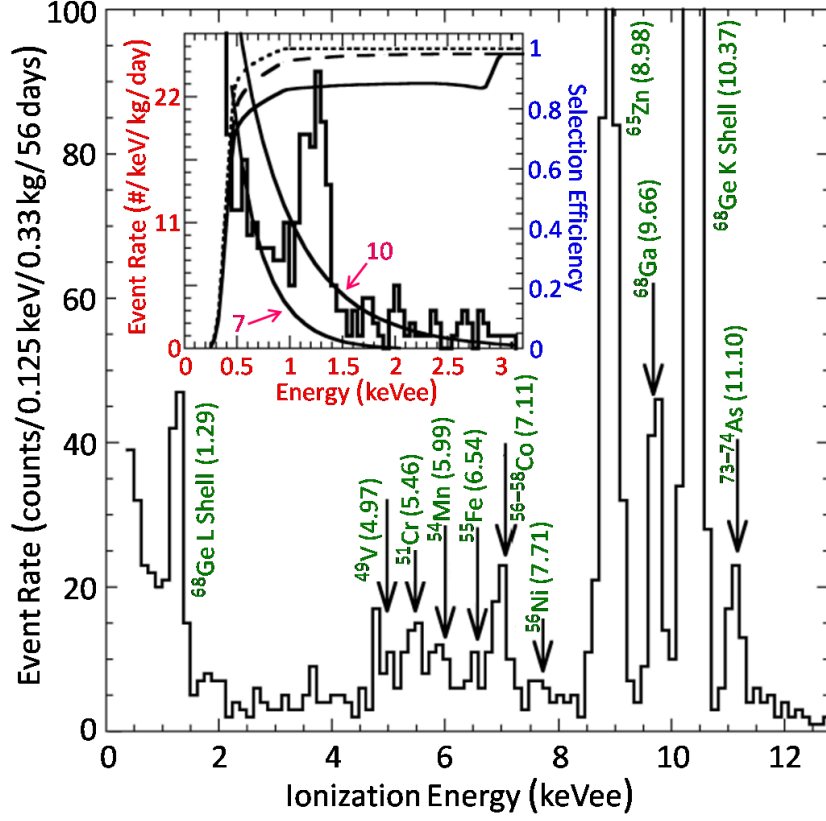


Figure 3.20: The differential spectrum recorded by the second CoGeNT PPC Ge diode during a 56 day exposure. The observed event rate is due primarily to decays of cosmogenically activated radioisotopes in the detector itself. The most prominent lines are indicated (green labels). The inset highlights the low-energy spectrum (red labels) used to search for Galactic WIMPs. Note that the event-rate units for the main spectrum and the inset differ by a factor of 2.5. Spectra expected for 7 and 10  $\text{GeV}/c^2$  WIMPs with  $\sigma_{\chi-N} = 10^{-4}$  pb are overlaid for comparison (magenta labels). A WIMP-like event excess that cannot be attributed to backgrounds is observed just above the 0.4 keVee threshold. The efficiencies of the hardware trigger and the cuts used for the final event selection are shown as well: trigger (dotted), trigger & microphonics cut (dashed), trigger & microphonics & rise-time cut (solid). Figure adapted from [670].

### CHAPTER 3. DARK MATTER DETECTION

spectrum obtained during the shallow-site exposure of the prototype, the near-threshold event rate represents nearly an order-of-magnitude reduction in backgrounds.

A type of discrimination has been used to obtain the final event selection shown in Figure 3.20. The CoGeNT detector exhibits a surface dead layer. Scattering events occurring within this surface layer can result in reduced ionization signals, possibly causing energy depositions due to WIMPs and backgrounds to be undermeasured. Not only does this reduce sensitivity to low-mass WIMPs, but it can also shift high-rate backgrounds to lower energies (where WIMPs are expected). Fortunately, ionization pulses recorded for interactions in the dead layer exhibit longer rise times than events in the crystal’s bulk. This long versus short rise-time behavior allows surface events to be discriminated from bulk events via pulse-shape discrimination (PSD). Application of a PSD cut, designed to retain 90% of all bulk events, effectively removes the outer 2 mm-thick surface of the detector, reducing the target mass to a “fiducial volume” of 330 g. Additionally, a modest PSD cut (described in [668]) is used to remove events caused by microphonics—laboratory vibrations that can mimic keV-scale energy depositions. The WIMP detection efficiencies associated with the rise-time and microphonics PSD cuts (as well as the hardware trigger) are shown in the inset of Figure 3.20.

Nearly all of the deep-site CoGeNT spectrum can be attributed to backgrounds. However, the CoGeNT collaboration lacks an explanation for the exponential rise in the event rate for energy depositions below  $\sim 2$  keV. As indicated in the inset of Figure 3.20, such an exponential spectrum just above threshold is precisely what is expected from low-mass WIMPs scattering from the target. Furthermore, if interpreted in terms of spin-independent WIMP-nucleon interactions, the favored parameter space is tantalizingly close to the low-mass DAMA allowed region in Figure 3.19. In [670], the authors correctly point out that their evidence is marginal. The probability that it is due to WIMPs is similar to the likelihood of the null hypothesis. Nevertheless, they give an allowed region that corresponds roughly to  $m_\chi = 8\text{--}10 \text{ GeV}/c^2$  and  $\sigma_{\chi-N} = 5\text{--}6 \times 10^{-5} \text{ pb}$  (see Figure 4 in [670] for full region).

### 3.3. DIRECT DETECTION

Although they do not overlap under standard assumptions, the similarity between the DAMA and CoGeNT regions has understandably generated excitement within the field, causing many to draw slim connections where they otherwise might not have. For example, the two events observed in the final run of the CDMS II experiment [459] (from which the CDMS II limit in Figure 3.19 is partially derived) have also been interpreted as evidence for a low-mass WIMP [670, 671, 672, 673]. It is arguable whether it makes sense to do this with only 2 events. In light of independent information from other experiments, perhaps exploration of the possibility is justified. In any case, due to the low recoil energies of these 2 events, in some scenarios the corresponding allowed region overlaps with the parameter space favored by the CoGeNT excess (again, see Figure 4 in [670]).

There are a number of studies in the literature that attempt to find a common explanation for the DAMA annual modulation and the CoGeNT excess. Among the possibilities explored are standard spin-independent elastic scattering [670, 673], asymmetric dark matter [671], generalized WIMP-proton and -neutron couplings with momentum dependence [674], pure-proton spin-dependent inelastic scattering [672], mirror dark matter [675], and more complicated multi-particle scenarios [676]. In most of these papers the authors find mutually consistent parameter space for the DAMA and CoGeNT data only if a significant ion-channeling effect is allowed for the former. It was first pointed out by Sekiya *et al.* in 2003 that a direct-detection experiment’s energy scale might be affected by ion channeling [677]. Early models for NaI suggest that a nontrivial fraction of recoiling nuclei following a WIMP-nucleus scattering event might be “channeled” between the rows and planes of a detector’s crystalline lattice, allowing them to move more freely through the crystal and deposit more energy via ionization (much like an electron recoil) [655, 678]. For small energy depositions, these early models predict that close to the full recoil energy of a WIMP interaction should be deposited as ionization:  $Q_{\text{Na}}$  and  $Q_{\text{I}} \rightarrow 1$  as  $E_R \rightarrow 0$  keV (see, *e.g.*, Figure 4 in [655]). On average, this beneficially lowers DAMA’s effective energy threshold for nuclear recoils, and shifts the DAMA allowed regions to lower masses and cross sections (see, *e.g.*,



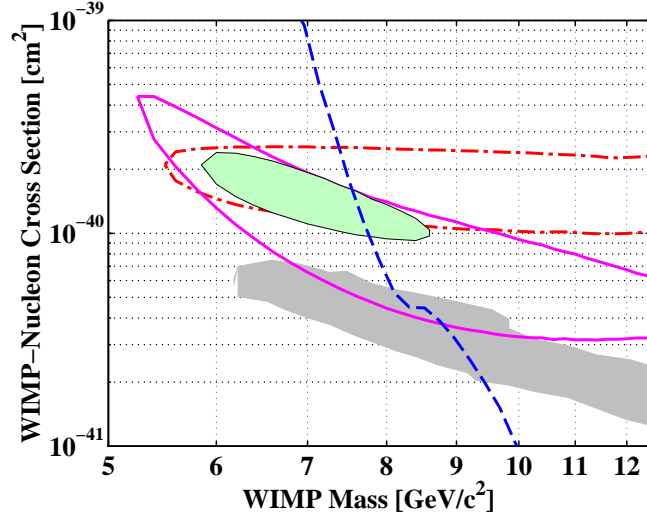


Figure 3.21: 99% confidence level allowed regions (as derived in [652]) corresponding to the DAMA annual modulation (red dash-dotted contour) and the CoGeNT excess (magenta solid contour), compared to the CDMS II limit [459] (90% C.L., blue dashed line) and LNM calculations [468] (90% C.L., gray shaded region). The low-mass WIMP models that are mutually consistent with the DAMA and CoGeNT regions at 99% confidence (oval-shaped, light-green shaded region) are the target of this thesis.

Figure 5 in [660] and Figure 7 in [655]). It is this ion-channeling shift (or some fraction thereof) that has been widely used to find a joint allowed region for the DAMA annual modulation and CoGeNT excess. Unfortunately, the more sophisticated models developed by Bozorgnia *et al.* indicate that ion channeling is much less pronounced than originally imagined [679, 680]. The DAMA contours shown in Figure 3.19 include the effect of ion channeling (according to Bozorgnia *et al.*'s model), and differ only very slightly from the corresponding regions that do not include channeling (see, *e.g.*, Figure 4 in [659]). Under most scenarios, the effect of ion channeling does not lead to a mutually consistent DAMA and CoGeNT allowed region.

Without invoking channeling, the authors in [652] offer what is perhaps the simplest combined analysis of the DAMA and CoGeNT data that yields

### 3.3. DIRECT DETECTION

a joint allowed region in terms of standard spin-independent WIMP-nucleon interactions. By considering uncertainty in the Ge and Na quenching factors, they find that the DAMA and CoGeNT contours partially overlap. Their individual and joint allowed regions are compared to the final CDMS II limit in Figure 3.21. The joint 99% confidence contour, centered at  $m_\chi \simeq 7 \text{ GeV}/c^2$  and  $\sigma_{\chi-N} \simeq 1.5 \times 10^{-4} \text{ pb}$ , is the low-mass WIMP parameter space targeted by the analysis described in this thesis.



# Chapter 4

## The CDMS Experiment

### 4.1 Introduction

The Cryogenic Dark Matter Search (CDMS) is a direct-detection experiment that has searched for WIMP dark matter for nearly two decades. During this time the experiment has progressed through two phases, distinguished by their respective detector technologies. The Berkeley Large Ionization- and Phonon-mediated (BLIP) germanium detectors employed during the first phase (CDMS I) are described in detail elsewhere (see, *e.g.*, the Ph.D. theses of Ning Wang [681], Tom Shutt [682], Walter Stockwell [683], Andrew Sonnenschein [684], and Sunil Golwala [685]), and are not the focus of this thesis. The salient feature of the BLIP technology was an ability to simultaneously measure an ionization and phonon signal for each scattering event, allowing event-by-event discrimination of electron recoils from nuclear recoils. The phonon measurement was achieved with two neutron-transmutation-doped (NTD) germanium thermistors [686] attached to each detector. Because the phonons had to first thermalize before a corresponding rise in crystal temperature could be measured by the NTDs, the BLIP phonon (or heat) measurement was somewhat slow—tens of milliseconds. In many respects, the upper limits on the WIMP-nucleon cross section presented in [687] represent the culmination of the CDMS I

## CHAPTER 4. THE CDMS EXPERIMENT

project. This 2002 paper was also the final CDMS publication featuring data from the BLIP detectors.

The second phase, CDMS II, employed Z-sensitive Ionization- and Phonon-mediated (ZIP) detectors [688, 689, 690]. As the name implies, ZIPs are also capable of the simultaneous measurement of ionization and phonons that enables electron-recoil discrimination. Unlike BLIPs, ZIPs are larger and more numerous, use both Si and Ge as targets, and (most importantly) measure phonons while they are still athermal—sub-millisecond time scales. Measuring the phonons before they thermalize has a number of advantages, some of which are discussed below. For details regarding the development of the ZIP-detector technology, I refer the interested reader to Sae Woo Nam’s Ph.D. thesis [691].

The two CDMS phases are also differentiated by the location and depth of their experimental installations. The CDMS I experiment was conducted exclusively at the Stanford Underground Facility (SUF) shallow site, characterized by a modest overburden of  $\sim 17$  m.w.e. Detailed treatments of the backgrounds at this site can be found in Angela Da Silva [692] and Thushara Perera’s [693] Ph.D. theses. The CDMS II experiment began operations at the same shallow site with its first “tower” of six ZIP detectors (Tower 1). The final exposure of Tower 1 at the shallow site, collectively termed “Run 21,” resulted in the first ZIP-exclusive WIMP-nucleon upper limit [694], and is the subject of Tarek Saab and Don Driscoll’s Ph.D. theses [695, 696].

The sensitivities of the final CDMS I and first CDMS II exposures were both limited by the shallow-site neutron background (see, *e.g.*, [697, 687, 694]). Consequently, the CDMS II effort was relocated to a new deep-site installation at the Soudan Underground Laboratory. The full compliment of CDMS II ZIP detectors was deployed in stages. The first deep-site run observed zero events in an  $\sim 50$  day exposure of Tower 1, resulting (at the time) in the world’s strongest limit on the spin-independent WIMP-nucleon cross section [698, 623] (further details can be found in Vuk Mandic, Clarence Chang, and Gensheng Wang’s Ph.D. theses [699, 700, 701]). During the second stage, two towers of ZIP detectors were exposed for  $\sim 75$  days, and (again) no sig-

#### 4.1. INTRODUCTION

nal due to Galactic WIMPs was observed (see, *e.g.*, Angela Reisetter, Joel Sander, and Walter Ogburn’s Ph.D. theses [702, 703, 704]). Data from the one- and two-tower runs were interpreted in terms of both spin-independent and (neutron-only) spin-dependent interactions, resulting in 90% C.L. upper limits of  $\sigma_{\chi-N} < 1.6 \times 10^{-7}$  pb [705] and  $\sigma_{\chi-n} \lesssim 10^{-1}$  pb [706] for  $m_\chi = 60 \text{ GeV}/c^2$ , respectively. During the final stage, the full compliment of CDMS II detectors—19 Ge and 11 Si ZIPs in a five-tower configuration—was operated for a two year period. An analysis of the first  $\sim 9$  months of the five-tower exposure observed zero WIMP candidates [637] (see also Raul Hennings-Yeomans, Jeffrey Filippini, and Catherine Bailey’s Ph.D. theses [707, 708, 709]), while the two candidates mentioned in the previous chapter were identified in the analysis of the final exposure [459] (see also Tobias Bruch, Matthew Fritts, and Zeesh Ahmed’s Ph.D. theses [710, 711, 712]). The final combined result from the CDMS II experiment is the WIMP-nucleon cross-section upper limit I have used as a benchmark in the previous two chapters (see, *e.g.*, Figures 2.18, 2.19, 3.19, and 3.21).

As stressed in the previous chapter, a direct-detection experiment’s sensitivity to low-mass WIMPs is critically dependent on its energy threshold. Among the publications and theses referenced above, the threshold (generally) varies between 7 and 20 keV. The lowest threshold achieved (prior to the analysis in this thesis) was for the shallow-site result in [694]. Four of the six detectors in the first half of the Run 21 exposure were analyzed with a 5 keV threshold, yielding unprecedented sensitivity (at the time) to WIMPs with  $m_\chi \lesssim 20 \text{ GeV}/c^2$ . As with many experiments, the choice of a particular threshold is (typically) driven by the desire for a background-free analysis. Due to finite-resolution effects, the ability of the CDMS discrimination parameter to reject electron recoils becomes degraded at low energies. Lowering a detector’s threshold below  $\sim 10$  keV invites the possibility that electron recoils will be misidentified as nuclear recoils. In order to maximize WIMP sensitivity, previous CDMS results have used the thresholds necessary to remain as close to background-free as possible.

## CHAPTER 4. THE CDMS EXPERIMENT

In most cases, published CDMS thresholds are a factor of 5 to 10 larger than required by the hardware trigger. Consequently, there is a substantial low-energy exposure that has yet to be analyzed. This thesis represents a first attempt to lower the CDMS thresholds as much as possible, and thus utilize for the first time the small energy depositions previously ignored. In contrast to previous results, the signal region is not expected to be background-free. By maximizing detection efficiency at low energy, the aim is to increase sensitivity to the low-mass WIMP parameter space favored by the joint analysis of the DAMA and CoGeNT data shown in Figure 3.21. To this end, the full Run 21 shallow-site exposure of the first tower of CDMS II detectors is considered. This thesis is thus an awkward mix of CDMS II detectors deployed at the CDMS I installation. Because the Tower 1 ZIP detectors and the SUF shallow-site facility have been covered in detail elsewhere, this chapter provides only the brief description necessary to set the scene for the analysis presented in the next chapter.

### 4.2 Detector Technology

CDMS II ZIP detectors are 1 cm-thick, 7.62 cm-diameter cylindrical crystals of Ge or Si. Tower 1 consists of four Ge and two Si ZIPs arranged in a vertical stack, with adjacent detectors separated by  $\sim 2.2$  mm with no intervening material. Operating detectors in such close proximity has the advantage that events due to backgrounds will often deposit energy in more than one detector. WIMP interactions are expected to be so localized and infrequent that multiple-scatter events can be safely rejected without loss of signal efficiency. Based on the above dimensions, one would expect the mass of a Ge (Si) ZIP to be 243 g (106 g). The thicknesses of the Tower 1 Ge substrates, however, varied from  $\sim 9.2$  mm to  $\sim 9.7$  mm. Also, the detectors do not have perfectly circular profiles. As indicated in Figure 4.1, small segments are missing at five positions along each detector's circumference. The flattened edges facilitate detector handling as a substrate is processed into a viable ZIP; a perfectly circular object is more difficult to hold (and align) than one with a slightly

## 4.2. DETECTOR TECHNOLOGY

Table 4.1: The first six CDMS II ZIPs are listed in order of their relative positions (top to bottom) within Tower 1. Each detector’s name, material, substrate thickness, and mass is indicated.

Name	Material	Thickness (mm)	Mass (g)
Z1	Ge	$9.65 \pm 0.05$	$230.5 \pm 1.2$
Z2	Ge	$9.53 \pm 0.23$	$227.6 \pm 5.5$
Z3	Ge	$9.18 \pm 0.05$	$219.3 \pm 1.2$
Z4	Si	$10.0 \pm 0.05$	$104.5 \pm 0.5$
Z5	Ge	$9.18 \pm 0.05$	$219.3 \pm 1.2$
Z6	Si	$10.0 \pm 0.05$	$104.5 \pm 0.5$

flattened profile. Relative to a perfect, 7.62 cm-diameter circle, the area of a semi-circular ZIP surface is smaller by 1.6%. ZIP-detector masses are therefore somewhat less than implied by the nominal thickness and diameter so often quoted in CDMS publications. The name, material, thickness, mass, and relative position of each Tower 1 detector is listed in Table 4.1.

As mentioned above, the most important feature of a ZIP detector is its ability to discriminate electron recoils from nuclear recoils. To understand how this is possible, consider the processes that occur when a particle interacts with a Ge or Si crystal. Following a scattering event, most of the energy imparted to the crystal is deposited as a spectrum of high-frequency athermal phonons (“primary phonons”). These are localized lattice vibrations that quickly proliferate as they downgrade in frequency and travel toward the surfaces of the detector. Along the path of the recoiling electron or nucleus, electron-hole pairs are created as well. The number depends on the total recoil energy as well as the type of interaction; nuclear recoils are less efficient at creating electron-hole pairs than electron recoils. An electric field applied across the crystal will cause the electrons and holes to drift to opposite sides where surface electrodes can be used to measure an ionization signal. The drifted electrons (holes) will eventu-



ally recombine with holes (electrons) at the electrodes (or in the bulk), at which point the deposited energy that went into the creation of the charge carriers is released as phonons (“recombination phonons”). Consequently, each scattering event results in a population of phonons—primary plus recombination—that is representative of the interaction’s full recoil energy. This is equally true for electron and nuclear recoils. For a given recoil energy, the ratio of the crystal’s ionization and phonon responses is therefore smaller for a nuclear recoil than for an electron recoil. This is the basis of the ZIP discrimination technology. Exactly how phonon and ionization signals are measured with ZIP detectors is described in greater detail in the following sections, as well as how they are used to reconstruct an event’s recoil energy, interaction type, and topology.

### 4.2.1 Ionization Measurement

In most respects, the ZIP-detector ionization measurement is achieved with conventional techniques for counting the number of electron-hole pairs created when a particle deposits energy in a semiconductor. The most important departure from conventional ionization detectors (like the HDMS Ge diode mentioned in Section 3.3.4, *e.g.*) is that ZIPs are operated at a much lower temperature—tens of mK rather than 77–300 K. In this section I briefly review the appearance and collection of ionization in a semiconductor, details specific to the ZIP-detector ionization electrodes and readout, and the main challenge to collecting ionization at sub-Kelvin temperatures. A more in depth treatment of ionization collection at very low temperatures can be found in Tom Shutt’s Ph.D. thesis [682]. Although tailored to the CDMS I BLIP-detector technology, most of his discussion is equally applicable to ZIP detectors.

#### **Ionization in Semiconductors**

The physics of ionization in a pure semiconductor at absolute zero is relatively straightforward. Valence electrons are attached to lattice sites and will not (of their own accord) jump the energy gap (“band gap”) to enter the conduction band. Consequently, application of an electric field does not cause a

## 4.2. DETECTOR TECHNOLOGY

current to flow. When a particle interaction supplies the energy needed to elevate electrons to the conduction band, a current proportional to the number of electron-hole pairs flows across the crystal. For a given recoil energy, the size of the current depends on the amount of energy it takes to produce an electron-hole pair, and whether it was an electron or nuclear recoil. A semiconductor's band gap is mildly temperature dependent, and is well approximated by the Varshni relation [713]:

$$E_g(T) = E_0 - \frac{\alpha T^2}{T + \beta}, \quad (4.2.1)$$

where  $E_0$ ,  $\alpha$  and  $\beta$  are constants that depend on the type of semiconductor.  $\alpha$  is typically on the order of  $10^{-4}$  eV/K, while  $\beta$  is roughly 100–1000 K [714]. Since CDMS ZIPs are operated at  $\sim 20$  mK, the band gap is effectively equal to  $E_g(T = 0 \text{ K}) \simeq E_0$ . The respective band-gap energies of Ge and Si at 0 K are 0.74 eV and 1.17 eV [715]. One might therefore suspect that a 10 keV electron recoil would produce  $\sim 10,000$  electron-hole pairs. This turns out to be a serious overestimate. The average deposited energy per electron-hole pair created, commonly denoted  $\epsilon$ , is actually a factor of a few larger than the band gap (in most semiconductors). For Ge (Si) substrates,  $\epsilon = 3.0$  (3.8) eV at mK temperatures [716, 717], and is a few percent less at 100 K (see, *e.g.*, Figure 5 in [718]). In other words, the number of electron-hole pairs created (by an electron recoil) is simply the ratio of the recoil energy to  $\epsilon$ :  $N_Q = E_R/\epsilon$ . On average, nuclear recoils result in about 1/3 as many charge carriers [719, 720] (nuclear-recoil ionization yield is discussed further in Section 4.2.3).

To understand why  $\epsilon$  is so large, we must delve further into the subtlety of how energy is deposited during a particle interaction. Put forth by Klein in [721, 718], the current best model posits two phases. For clarity, consider an electron-recoil interaction. In the first phase the incoming particle elevates electrons into the conduction band with sufficient momentum to cause a cascade of secondary ionization. The resulting cloud of charge carriers are initially so energetic that interactions with the crystal lattice result in the emission of optical phonons. Once the energy of the charge car-

riers falls below a threshold energy for the production of new charge carriers, a second phase occurs in which their residual kinetic energy is dissipated in the form of acoustic phonons. The first phase is nearly instantaneous, while the second occurs on a time scale of  $\sim 10$  ps. The violence of the energy deposition and corresponding phonon production is the reason  $\epsilon$  is larger than the band-gap energy. For a more detailed discussion, I refer the interested reader to Appendix C in Walter Ogburn’s Ph.D. thesis [704].

### Electrodes & Readout

A schematic of the ZIP-detector ionization electrodes and readout is shown in Figure 4.1. Two ionization electrodes are deposited onto the bottom surface of each detector. A circular inner electrode (“ $Q$  inner”) covers most of the physical area, and is encircled by a thin annular electrode (“ $Q$  outer”). Thanks to the close spacing of detectors within the tower, the inner sections of the top and bottom surfaces are subject (via direct line of sight) to a very small solid angle of external background radiation. The outer sections (including the side walls), however, are more exposed. Background particles resulting from decays of radiocontaminants on the copper surfaces of the tower’s mechanical support structure are more likely to be incident upon these outer sections. The  $Q$ -outer electrode grants a level of protection against such events, while the  $Q$ -inner electrode is used to identify interactions occurring in the central part of the crystal—the fiducial volume.

An estimate of the fiducial volume can be derived from the dimensions of the crystal and ionization electrodes. The 6.88 cm-diameter  $Q$ -inner electrode is separated (radially) from the  $\sim 2$  mm-wide  $Q$ -outer electrode by a 1 mm-wide trench. A naïve calculation of the  $Q$ -inner fractional coverage yields an expected fiducial volume of 81.5%. Events occurring beneath the inner half of the trench, however, are more likely to cause ionization that is drifted to the  $Q$ -inner electrode. The effective diameter of the  $Q$ -inner electrode is thus closer to 6.98 cm, increasing the fiducial volume to 83.9%. This is a slight under-

## 4.2. DETECTOR TECHNOLOGY

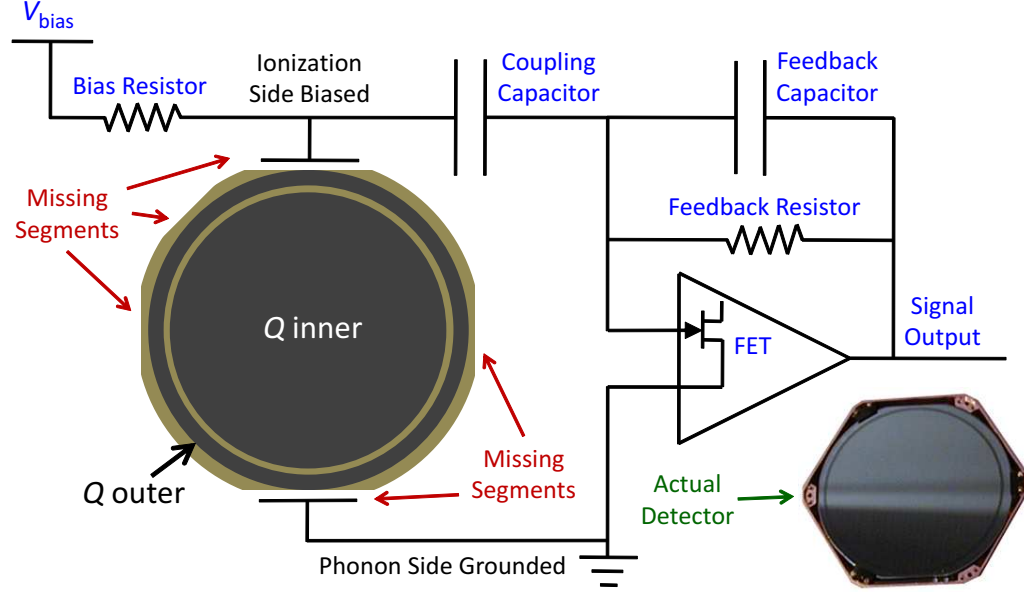


Figure 4.1: Schematic of the ZIP inner and outer ionization electrodes, and the low-noise charge-amplifier circuit (with Field-Effect Transistor (FET) front ends) used to readout ionization pulses. A separate charge-amplifier channel is used for each electrode such that each scattering event produces independent  $Q$ -inner and -outer signals. The dimensions of the electrodes and missing segments have been exaggerated for illustrative purposes. A photograph of the ionization side of an actual ZIP detector is shown at the lower right.

estimate because (as discussed above) the detector profiles are not perfectly circular. Taking into account the reduced surface area due to the five missing segments, the expected fiducial volume is  $\sim 85.3\%$ . In this thesis I adopt the (very slightly) conservative CDMS convention of rounding this number down to 85%.

The ionization measurement is accomplished by holding the phonon side at ground while a bias voltage between  $-3$  and  $-6$  V is applied to the ionization electrodes, causing liberated holes and electrons to drift to opposite sides. A charge equal to  $eN_Q$  accumulates on the feedback capacitor shown in Figure 4.1 through the action of a charge-sensitive amplifier. The build up (over a few  $\mu\text{s}$ )

## CHAPTER 4. THE CDMS EXPERIMENT

and subsequent decay (over  $\sim 100 \mu\text{s}$ ) of charge on the feedback capacitor results in an ionization pulse. The relatively quick response of the ZIP ionization measurement makes the ionization-pulse start time the most accurate indicator of when a scattering event occurred. Independent  $Q$ -inner and -outer signals are digitized for each event. A software optimal-filter template fit measures the pulse height (and start time) of each digitized signal offline (see Chapter 6 and Appendix B in [685] for further details regarding optimal filtering and pulse fitting in general). In this way an “ionization energy” is reconstructed for each channel.

One of the advantages of working at cryogenic temperatures is that the electronic noise of the readout circuit shown in Figure 4.1 is quite low. At finite temperatures there can be sufficient thermal energy to populate the conduction band and cause a “leakage current” in the presence of an electric field. Fluctuations in the leakage current due to shot noise—Poisson fluctuations in the number of electrons contributing to the current—can be large enough to mask the ionization caused by particle interactions. From the perspective of the readout circuit, a ZIP detector is basically a capacitor. The rms fluctuation in voltage across a capacitor ( $C$ ) at temperature  $T$  is given by  $V_{\text{rms}} = \sqrt{kT/C}$  [722], which corresponds to a charge fluctuation of

$$Q_{\text{rms}} = \sqrt{kTC}. \quad (4.2.2)$$

For a nominal detector capacitance of 100 pF [695], thermal fluctuations contribute  $\sim 4000$  charge carriers at 300 K. This corresponds (for an electron recoil) to a  $1\sigma$  noise width of  $\sim 4000e/\sqrt{2} \simeq 10 \text{ keV}$  for Ge and Si semiconductors. Even at the temperature of liquid nitrogen—77 K is a common operational temperature for conventional ionization detectors—thermal fluctuations would contribute  $\sim 5 \text{ keV}$  in noise at the  $1\sigma$  level. To prevent confusion of noise pulses with true particle interactions, an ionization threshold of perhaps  $4\text{--}6\sigma$  is necessary. Such a large threshold would severely limit a dark-matter detector’s sensitivity to the small energy depositions expected from WIMP-nucleus interactions, particularly considering that even smaller ionization signals are expected for nuclear recoils. Consequently, conventional ionization detectors

## 4.2. DETECTOR TECHNOLOGY

are typically constructed as p-n diodes, and operated with a large reverse bias voltage. This configuration creates a charge-free depletion region that serves as the active detector volume.

Due to the low 20 mK operating temperature, the CDMS approach is fundamentally different. At 20 mK, rms thermal fluctuations contribute only about 33 charge carriers, which translates to a  $1\sigma$  thermal-noise width of  $\sim 70$  eV (90 eV) for Ge (Si) ZIPs. With such low intrinsic noise levels, the Ge and Si substrates do not have to be constructed as diodes. Furthermore, a relatively modest electric field of  $\sim 200$  mV/cm is sufficient to achieve nearly full charge collection [682]. The use of a small field has an important consequence with respect to the phonon measurement. As charge carriers are drifted across a crystal they cause lattice vibrations known as “Neganov-Luke” phonons (or “drift heat”) [723, 724]. These are in addition to the primary and recombination phonons mentioned above. The energy required to produce drift heat comes from the electric field permeating the crystal, and not directly from the energy deposited during a scattering event. Consequently, the addition of drift heat to the phonon population represents an amplification of an event’s recoil energy. The amount of drift heat produced is directly proportional to the number of charge carriers and the bias voltage:

$$E_{\text{Luke}} = eVN_Q = \frac{eVQ}{\epsilon} \text{ keV}, \quad (4.2.3)$$

where  $Q$  represents the measured ionization energy in keV,  $V$  is the absolute value of the bias voltage (in volts),  $e$  is the charge of the proton, and  $\epsilon$  is as given above for Ge and Si (in eV). For a Ge electron recoil, for example, drift heat doubles (triples) the phonon signal for  $V = 3$  (6) V. This drift-heat amplification makes the phonon signal a more sensitive indicator of a particle interaction (in terms of recoil energy). The benefit of amplifying small energy depositions in this manner, however, must be weighed against the performance of the discrimination parameter. As the bias voltage is increased, the drift heat associated with the thermal ionization noise is amplified as well, and the resolution of the ratio of the ionization and phonon signals becomes degraded. If the bias voltage is too large, the ability to discriminate electron from nuclear

## CHAPTER 4. THE CDMS EXPERIMENT

recoils is compromised at low energy. CDMS has determined empirically that an electric field of 3 (4) V/cm is about optimal for Ge (Si) ZIPs. Although I will not discuss this in detail, one of the purposes of Run 21 was to determine if there might be any advantage to operating ZIPs with a larger bias voltage. Consequently, during the first half of Run 21, a  $-3$  ( $-4$ ) V bias was used for Ge (Si) ZIPs, while during the second half a  $-6$  V bias was used (for both Ge and Si).

Despite their low-noise Field-Effect Transistor (FET) front ends, the dominant source of noise in the ionization measurement is from the charge-sensitive operational amplifiers (“op amps”). The gate capacitance of a CDMS FET is comparable to the detector capacitance. However, the FETs are rendered nonoperational at temperatures below  $\sim 100$  K. Consequently, they must be isolated thermally from the rest of the readout circuit and the detector. Rather than sitting at 20 mK, the FETs are coupled to the 4 K stage of the cryostat (discussed below), where they are allowed to self heat to their optimal operating temperature of  $\sim 130$  K. All things considered, the 200–400 eV baseline ( $1\sigma$ ) ionization noise achieved for the Tower 1 detectors during Run 21 is quite remarkable. Energy resolution is discussed further in the next chapter.

### Impurities & Neutralization

Crystal impurities play an important role in the collection of ionization in a semiconductor. Impurities come in two varieties: donors and acceptors. The former contribute electrons to the conduction band more easily than the main semiconductor material, while the latter accept conduction electrons more easily. In other words, it requires far less energy—typically on the order of 10 meV relative to the  $\sim 1$  eV band gap of Ge and Si—for a donor (acceptor) to contribute an electron (hole) to the conduction (valence) band. Semiconductors with more donors than acceptors are called n-type, while those with more acceptors are p-type. The level of impurity is usually quantified in terms of the difference in number density between acceptors and donors:  $n_a - n_d$ . In a

## 4.2. DETECTOR TECHNOLOGY

conventional ionization detector, the presence of too many acceptors or donors can significantly alter the semiconductor’s conductivity and lead to unacceptable leakage currents. On the other hand, in some applications highly doped semiconductors are desirable. Si with an impurity level of parts per thousand (“degenerately doped”), for example, is often used in integrated circuits as a conductor (instead of metal).

Even the ultra-pure substrates from which ZIPs are produced contain non-trivial impurity densities. The p-type substrates used to fabricate CDMS Ge ZIPs, for example, are characterized by  $n_a - n_d \simeq 6 \times 10^{10} \text{ cm}^{-3}$  [725]. At 20 mK, the slightly modified band structure caused by the impurities in ultra-pure Ge or Si is not a direct concern with respect to the number of electron-hole pairs created during a scattering event, or with respect to the semiconductor’s conductivity. There simply is not enough thermal energy to significantly populate the conduction band ( $kT \simeq 2 \mu\text{eV} \ll 10 \text{ meV}$ ). The impurity sites are a problem for an altogether different reason. Ionized impurities can trap charge carriers as they are drifted across a detector, causing reduced ionization signals. An electron recoil with a reduced ionization signal might be misidentified as a nuclear recoil. Ionized impurities therefore have a direct effect on a ZIP’s discrimination performance.

The problem originates at room temperature, where the available thermal energy causes a significant population of the conduction (valence) band by electrons (holes). Since impurities more easily donate electrons or holes, they are almost completely ionized at room temperature. As a crystal is cooled the number of free charge carriers is Boltzmann suppressed. Normal Ge and Si atoms return to a neutral state in which their valence band is full (of electrons). An equal number of donors and acceptors, however, find it energetically favorable to swap an electron and hole rather than return to a neutral state. In the CDMS p-type Ge, since acceptors outnumber donors, this results in an equal density of  $n_d$  ionized donors *and* acceptors. While the majority of the impurity sites ( $\mathcal{O}(10^{10}) \text{ cm}^{-3}$  acceptors) freeze into a neutral state, on the order of  $10^9 \text{ cm}^{-3}$  donors and acceptors freeze into an ionized state. With so many islands of ionization (“charge traps”) permeating the crystal, a significant frac-



## CHAPTER 4. THE CDMS EXPERIMENT

tion of the charge carriers created during a scattering event become trapped as they are drifted to the surface electrodes. For an electron recoil, there are only  $\sim 330$  (260) charge carriers created per keV of deposited energy in Ge (Si). For nuclear recoils, only  $\sim 100$  charge carriers are created (on average) per keV of recoil energy. In order to preserve the integrity of the keV-scale ionization signals expected from low-mass WIMPs, full charge collection is essential. Charge traps can be overcome by large electric fields, but at the expense of degraded discrimination performance due to an overwhelming amount of drift heat.

An alternate solution is to somehow neutralize the ionized impurities after the crystal has been cooled to the 20 mK base temperature. Neutralization is accomplished by creating electron-hole pairs (throughout the crystal) in the absence of a bias field. The free charges will either neutralize nearby ionized impurities, or immediately recombine. Since there is very little thermal energy at 20 mK, the fully neutral state is basically stable. In practice, electron-hole pairs are created by exposing the detectors to a strong source of ionizing radiation. CDMS uses photons from light emitting diodes (LEDs). In particular, infrared LEDs emit photons with the  $\sim 1$  eV necessary to overcome the Ge and Si band gaps. The detectors are “baked” with these infrared photons for several hours until they are completely neutralized. During the course of normal operations, they are further neutralized every few hours for minute long periods. This maintains neutralization despite the continual exposure to ionizing (background) radiation (while biased). The LEDs are “flashed” during these periodic neutralizations so that the detectors are not heated too far above base temperature, allowing the search for WIMPs to resume within minutes. As I will discuss in the next chapter (see Section 5.3.2), these LED flashing periods were (indirectly) linked to the single largest source of low-energy background events in the Run 21 exposure.

## 4.2. DETECTOR TECHNOLOGY

### 4.2.2 Phonon Measurement

The CDMS II phonon sensors are one of the truly unique features of a ZIP detector. As mentioned above, the phonons are measured before they thermalize, providing a much faster response than possible with a more conventional heat measurement. The ZIP phonon technology goes by the name Quasiparticle-trap-assisted Electrothermal-Feedback Transition-edge sensors (QETs), and is described in detail in a number of papers and CDMS theses (see, in particular, [688, 689, 690, 691, 695, 726, 727]). Many years of ingenious, hard work went into the design, fabrication, and testing of the Tower 1 phonon sensors. I cannot take any credit here; most of this work was done by Stanford University researchers, and was slightly before my time. As such, I also cannot hope to do justice to all the intricacies involved. The following is thus similar to the above discussion of the ionization measurement; I briefly review the appearance of phonons in semiconductor crystals, and their collection with ZIP QETs.

#### Phonons in Semiconductor Crystals

When a particle deposits energy in a semiconductor crystal, it gives rise to multiple phonon populations. The majority of an event's recoil energy goes into the production of primary phonons at the interaction point. These are localized lattice vibrations whose spectrum of frequencies is peaked at several THz. Because the semiconductors are held at such low temperatures, phonons due to thermal vibrations are highly suppressed. The primary phonons therefore rarely scatter from thermal phonons, and instead undergo anharmonic decay as they travel away from the interaction point [728]. Anharmonic decay is a process whereby a phonon downgrades in energy by spontaneously splitting into two lower-frequency phonons. Although their number quickly increases, energy is conserved, and the fraction of the original interaction's recoil energy carried by the primary-phonon population does not change. Multi-THz phonons also undergo isotopic scattering in Ge and Si semiconductors—naturally occurring Ge (Si) consists of 5 (3) stable isotopes. This is a process in which a high-

## CHAPTER 4. THE CDMS EXPERIMENT

frequency phonon encounters a different isotope in the crystal's lattice, causing the phonon to scatter elastically. Unlike anharmonic decay, isotopic scattering only changes a phonon's direction of propagation. As a result of these two processes, the primary phonons propagate in a quasidiffuse manner from the interaction point until their frequencies fall below  $\sim 0.6$  THz (1.2 THz) in Ge (Si) [695, 728]. At this point the cross sections for anharmonic decay and isotopic scattering have fallen off so much that the phonon mean free path becomes larger than the size of the crystal. These lower-frequency phonons are described as “ballistic” because they travel in straight lines to the crystal's surfaces. The time scale for primary phonons generated in the bulk of the crystal to arrive at the detector surfaces is several  $\mu$ s. If they encounter a bare surface (no electrodes), they will simply reflect back into the crystal without loss of energy. As will be described in more detail below, the QETs cover only a fraction of a ZIP's phonon side. Furthermore, the probability that a phonon is absorbed by a QET is only  $\sim 30\%$  (averaged over frequency). Consequently, the collection of primary phonons from the bulk typically occurs over tens of  $\mu$ s.

As mentioned above, there are also phonons associated with the ionization produced during a scattering event: recombination and Neganov-Luke phonons. In the absence of an electric field, the recoil energy spent on creating electron-hole pairs would be released into recombination phonons right away. The bias field of a few volts/cm prevents this from happening near the interaction point, and instead causes the electrons and holes to drift to opposite sides of the crystal. As the electric field attempts to accelerate the charge carriers to velocities in excess of the speed of sound in the semiconductor, a phenomenon similar to Cherenkov radiation occurs, and the charges shed energy by emitting Neganov-Luke phonons. The energy required to create this drift heat is pulled from the electric field, and is therefore not part of the recoil energy deposited at the original interaction point. When the electrons and holes eventually recombine at the surface electrodes, their creation energy is released as recombination phonons. Due to interactions with the metallic electrodes, these recombination phonons very quickly become ballistic. Since it takes several hundred nanosec-

## 4.2. DETECTOR TECHNOLOGY

onds for the charges to drift across the crystal, the time scale for the arrival of recombination and Neganov-Luke phonons to reach the crystal surfaces is only a few  $\mu\text{s}$ .

An important consequence of the ZIP athermal phonon measurement is an unexpected ability to discriminate events occurring in a detector’s bulk from those occurring near its surfaces (discussed in detail, *e.g.*, in Roland Clarke’s Ph.D. thesis [729]). The primary phonons from interactions that occur near a surface interact more quickly with the metallic electrodes, and thus downgrade to ballistic frequencies more quickly. The collection time for phonons from surface events is therefore shorter than for bulk events. The corresponding shorter rise time of a surface-event phonon signal (relative to a bulk event) allows for a type of pulse-shape discrimination (PSD). Surface events are problematic because charge carriers produced in such close proximity to an electrode will often recombine before they can be drifted across the crystal. Electron-recoil surface events therefore exhibit smaller ionization signals and can be mistaken for nuclear recoils. This is demonstrated in Figure 4.2 for calibration data from the final 5-tower CDMS II exposure. Bulk electron (nuclear) recoils occur near the top (bottom) of the graph due to their large (small) ionization-to recoil-energy ratios, demonstrating the performance of the standard CDMS discrimination parameter (described below). Electron recoils occurring at the surface of the detector, however, span the region in between (along the vertical axis) as a result of their reduced ionization signals, and are thus more difficult to distinguish from nuclear recoils. Fortunately, the phonon-based PSD (along the horizontal axis) allows bulk nuclear recoils (to the right) to be separated from these otherwise degenerate surface events.

The limiting background for the final 5-tower CDMS II exposure turned out to be surface events (rather than high-energy neutrons as originally anticipated). The main source of these events is from  $^{210}\text{Pb}$  contamination on the towers’ inner Cu surfaces and on the detectors themselves—a result of radon exposure. When  $^{210}\text{Pb}$  decays it typically emits an electron in the energy range of interest (17 keV or 63 keV). Such low-energy electrons cannot penetrate into a detector’s bulk. The resulting reduced ionization and normal phonon signals

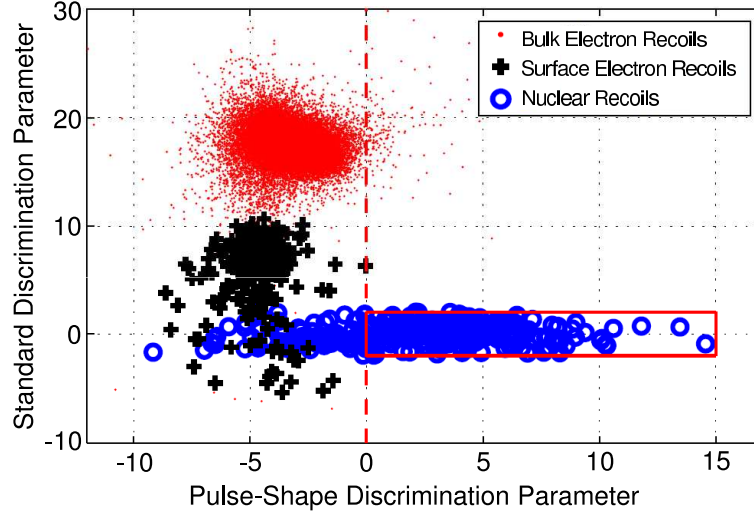


Figure 4.2: Demonstration of ZIP-detector electron-recoil (vertical axis) and surface-event (horizontal axis) discrimination for calibration data from the final 5-tower CDMS II exposure. Bulk electron recoils (red dots) caused by gammas from a  $^{133}\text{Ba}$  source are well separated (vertically) from nuclear recoils (blue circles) caused by fission neutrons from a  $^{252}\text{Cf}$  source, demonstrating the standard CDMS discrimination technology. Gammas from the  $^{133}\text{Ba}$  source that interact at the detector surfaces (black pluses), however, span the region in between, and can be confused with nuclear recoils. Phonon-based pulse-shape discrimination (horizontal axis) allows events in the bulk (to the right) to be distinguished from these otherwise degenerate surface events. The box (solid lines) indicates the signal region where bulk nuclear recoils are expected. Both discrimination parameters are plotted in normalized units—number of standard deviations relative to the signal region. Figure adapted from [459].

can easily mimic the type of event expected from a WIMP-nucleus interaction. The phonon-based PSD allows a fiducial volume to be defined across a detector's thickness (in addition to radially via the  $Q$ -inner electrode). In this way, the main CDMS II analyses remained nearly background-free at the cost of exposure.

In the low-energy analysis presented in the next chapter, no attempt is made to use phonon-based PSD. The PSD demonstrated in Figure 4.2 only works

## 4.2. DETECTOR TECHNOLOGY

effectively for recoil-energies in excess of 10-20 keV (depending on detector). In order to include the lowest-energy recoils without shrinking the fiducial volume to zero thickness, the surface-event background must be accepted into the signal region. Again, a philosophy of maximizing detection efficiency at low energy is adopted for the chance to probe previously untested low-mass WIMP parameter space. Consequently, I will not discuss surface-event rejection any further.

### The ZIP-Detector QETs

The ZIP-detector phonon measurement begins when a ballistic phonon interacts with one of the many Al collector fins photolithographed onto the top side of each detector. The 300 nm-thick Al films are held well below their superconducting transition ( $T_c \simeq 1.2$  K [730]), such that many of their electrons are coupled into Cooper pairs [731]. The superconducting gap—the energy required to break a Cooper pair into 2 quasiparticles—of  $340 \mu\text{eV}$  (in Al) corresponds to a phonon frequency of 84 GHz [695]. A majority of the phonons reaching a detector’s surface (following a scattering event) have frequencies of several hundred thousand GHz. When such high-frequency phonons interact with Cooper pairs, they create quasiparticles with sufficient energy (a few meV each) to initiate cascades of quasiparticle production. The quasiparticles then diffuse down the length of the collector fin to a  $4 \mu\text{m}$ -wide region in which a 35 nm-thick W film overlaps the Al. When they scatter inelastically in this Al-W transition region, the quasiparticles become trapped by its smaller energy gap. Unable to return to the Al collector fin, the quasiparticles continue to a W-only region that is held at its superconducting transition. Finally, the quasiparticles transfer their energy to the W film’s electron system, causing an increase in temperature that increases the W film’s resistance. The change in resistance can be measured by biasing the W and monitoring the flow of current with a low-noise readout circuit (described below). A diagram illustrating this phonon-detection process is shown in Figure 4.3.

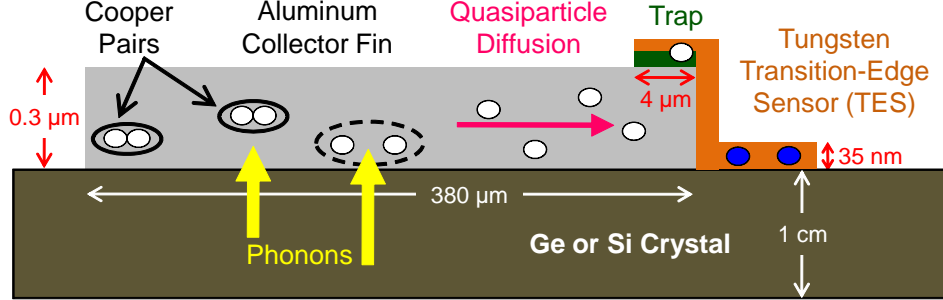


Figure 4.3: Cartoon of how a quasiparticle-trap-assisted transition-edge sensor works. Ballistic phonons resulting from a particle interaction in the Ge or Si substrate travel to the surface of the detector where they encounter an aluminum collector fin. The Al is held below its superconducting transition such that many of its electrons are coupled into Cooper pairs. Phonons can dissipate their energy by breaking these Cooper pairs into quasiparticles, which then diffuse toward an Al-W transition region where they become trapped. The W film is held at its superconducting transition such that the change in temperature due to the quasiparticles causes a corresponding increase in the tungsten’s resistance.

The first generation of ZIPs employed a design in which nearly all of a detector’s phonon side was covered by Al collector fins. Such a configuration turns out to be nonoptimal because the average distance a quasiparticle has to diffuse (before encountering an Al-W transition region) is a factor of  $\sim 3.5$  larger than the quasiparticle diffusion length ( $\sim 180 \mu\text{m}$  [727]). Quasiparticles are unlikely to travel more than 1–2 diffusion lengths before recombining back into Cooper pairs. Consequently, the original design suffered from a quasiparticle-collection inefficiency of  $\sim 95\%$ . The signal-to-noise in the phonon channel is critically dependent on the number of collected quasiparticles per keV of deposited energy. These early ZIP detectors therefore had limited sensitivity to the small energy depositions expected from WIMP-nucleus interactions.

The current design is the result of a quasiparticle-collection optimization, and is described in greater detail in Tarek Saab’s Ph.D. thesis [695]. Each

## 4.2. DETECTOR TECHNOLOGY

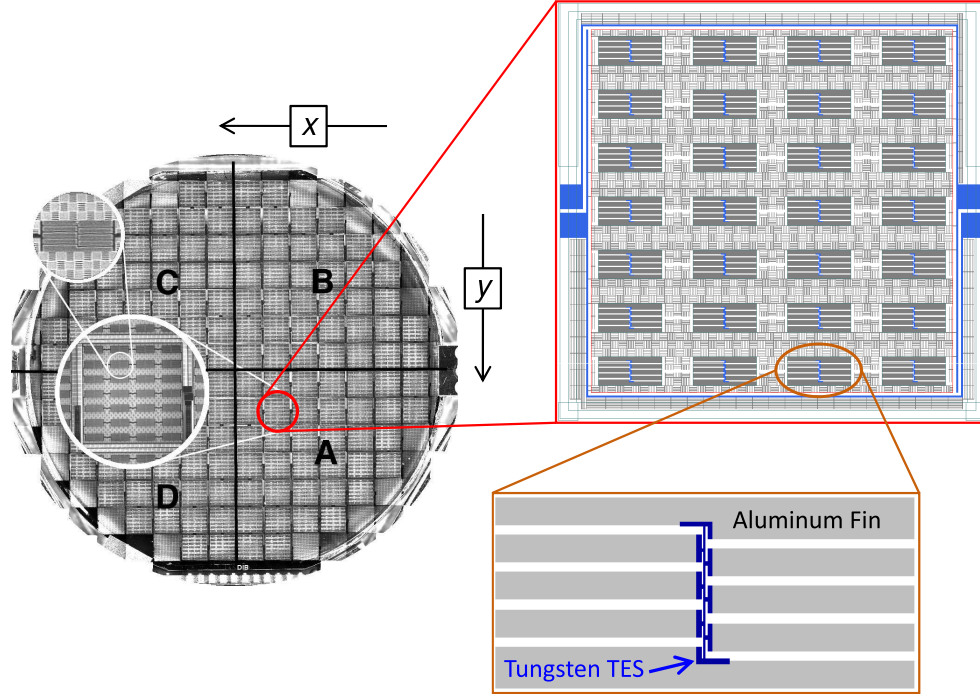


Figure 4.4: Images and schematics of the ZIP-detector phonon sensors. *Left*: Digital photograph of the arrangement of QET cells into quadrants ( $4 \times 37$   $5 \times 5 \text{ mm}^2$  cells), with zoomed in views of an individual QET cell (larger inset) and an individual QET structure (smaller inset). *Upper right*: Detailed schematic revealing the  $7 \times 4$  array of QETs within a (Ge-ZIP) QET cell. *Lower right*: Schematic of a QET structure showing how the W transition-edge sensor attaches to its Al collector fins. The QETs within a cell and the 37 cells within a quadrant are electrically connected, yielding a total of four phonon channels per detector. Figure adapted from [695].

QET structure consists of ten  $\sim 380 \times 60 \mu\text{m}^2$  Al collector fins attached to an  $\sim 250 \times 1 \mu\text{m}^2$  W transition-edge sensor (TES) (as shown at the lower right in Figure 4.4). For the Tower 1 Ge (Si) ZIPs, 28 (24) of these QET structures are arranged in a  $7 \times 4$  ( $6 \times 4$ ) array to form a  $5 \times 5 \text{ mm}^2$  QET cell (upper right of Figure 4.4 shows the cell for a Ge ZIP, *e.g.*). The top surface of each ZIP is patterned with 148 cells that are divided into electrically connected quadrants,



## CHAPTER 4. THE CDMS EXPERIMENT

resulting in four independent phonon channels per detector (left side of Figure 4.4). This arrangement of QETs yields a quasiparticle-collection efficiency of nearly 25%.

A critical feature of the ZIP-detector phonon measurement is the use of ElectroThermal Feedback (ETF) (hence the  $E$  in QET). As alluded to above, the small change in temperature caused by quasiparticles transferring their energy to the electrons in a W TES is measured as a change in resistance. The largest change in resistance (for a given rise in temperature) occurs at the superconducting transition. This is demonstrated in Figure 4.5 for a W sample with characteristics typical of the W used in ZIP-detector QETs. Note that the superconducting transition is well above the 20 mK temperature at which the detectors are operated. This is necessary for ETF to function properly. The W TESs are held near their superconducting transition through the application of a stiff bias voltage.<sup>1</sup> In this way, the tendency for the W to cool to the temperature of the (Ge or Si) substrate is balanced by the bias current via joule heating. If tuned properly, this equilibrium between joule heating and the cooling power of the thermal bath is highly stable. A small rise in temperature causes a reduction in joule heating—due to an increase in resistance and corresponding reduction in current—that temporarily allows heat to leak from the W electron system into the thermal bath faster than it is being replaced. Once the W resistance has decreased sufficiently, the bias current increases to the level necessary for joule heating to restore equilibrium. For further details regarding ETF, I refer the interested reader to [732, 691, 695].

There are a few important considerations that contribute to the target  $T_c$  for ZIP-detector TESs. The first is that the  $T_c$  must be considerably larger than the temperature of the substrate in order for ETF to work properly and efficiently. If the W  $T_c$  is too close to the temperature of the substrate, the ETF time scale becomes comparable to the intrinsic thermal decay time of the

---

<sup>1</sup>ZIP-detector transition-edge sensors are actually biased toward the low- $T$  side of their superconducting transition to increase the dynamic range of their response. The nominal resistance at the biasing point is  $\sim 200$  m $\Omega$  for a W  $T_c \simeq 80$  mK and a transition width of a few mK.

## 4.2. DETECTOR TECHNOLOGY

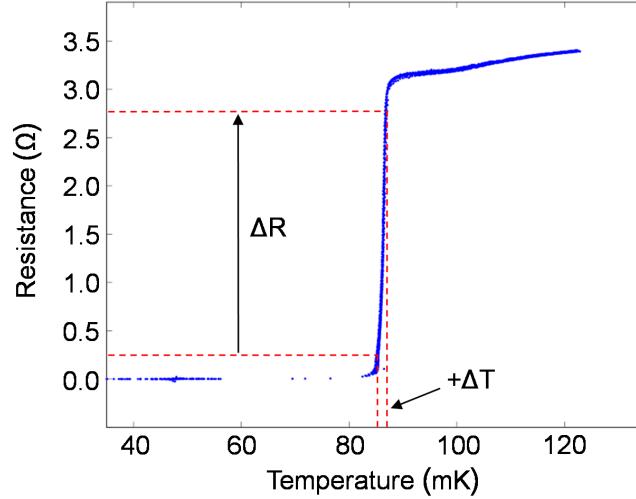


Figure 4.5: Example superconducting transition for a tungsten sample with characteristics typical of the W used in ZIP-detector QETs. As indicated, a small rise in temperature causes a large change in resistance. By holding the W at the low- $T$  side of its superconducting transition via electrothermal feedback, the W is converted to a highly sensitive thermometer with a short recovery time. Figure adapted from [695].

W, allowing heat to be lost into the substrate before it can be detected. Similar to quasiparticle-collection inefficiency, this leads to a reduction in the phonon sensor’s signal-to-noise, and sets a requirement that  $T_c \gg 20$  mK. On the other hand, the noise resolution of a TES scales in proportion to  $T_c^2$  to  $T_c^3$  (see, *e.g.*, Kent Irwin’s Ph.D. thesis [733]). The CDMS collaboration has determined empirically that a  $T_c \simeq 70$ – $80$  mK provides a nearly optimal balance between these two competing factors.

W deposited on a detector surface tends to form in a mixture of two crystalline phases: the body-centered cubic  $\alpha$ -W phase with  $T_c \simeq 10$  mK, and the A15  $\beta$ -W phase with  $T_c \simeq 1$ – $4$  K [734]. The amount of the  $\beta$  phase appears to depend on the level of oxygen present during the deposition. As a result of  $\alpha$ - and  $\beta$ -phase mixing, directly following deposition, the  $T_c$  of a ZIP-detector TES tends to fall in the 100–150 mK range. This satisfies the ETF requirement,

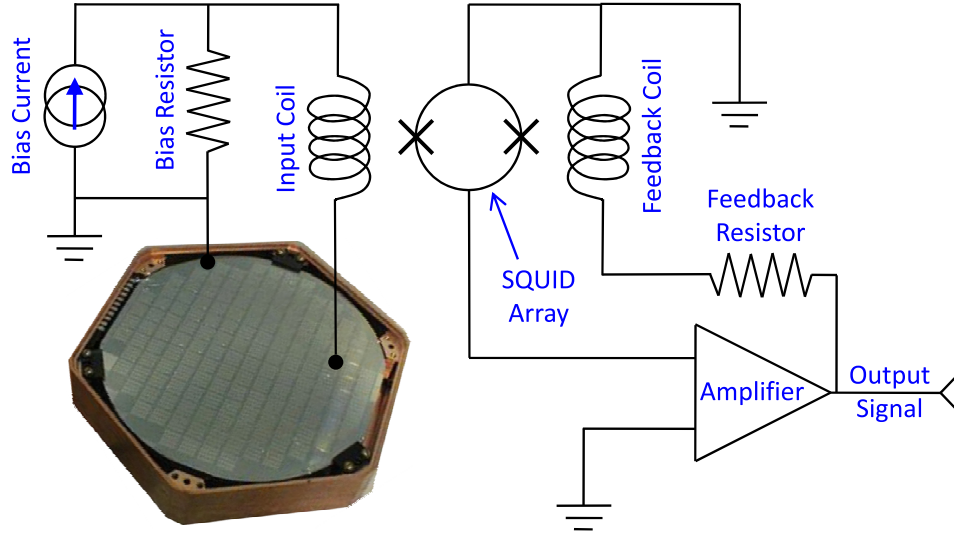


Figure 4.6: Schematic of the biasing and SQUID-based amplification circuit used to readout the ZIP-detector QETs. See main text in this section for details of operation.

but is too large from a noise-performance perspective. Furthermore, the  $T_c$  on any given ZIP exhibits a 20–40 mK gradient across the face of the detector. Such a gradient causes a variation in phonon pulse shape across the face of the detector, which leads to a position-dependent systematic uncertainty in the phonon measurement. As discussed in the context of the ionization measurement, a degradation in signal resolution can compromise a detector’s ability to discriminate electron recoils from nuclear recoils. Fortunately, a contingent within the CDMS collaboration has developed a method for correcting the gradient while lowering the  $T_c$ . They recognized that the  $W T_c$  can be suppressed through the addition of metal impurities with a procedure known as ion implantation [735, 736]. ZIP-detector TESs are thus tuned to a nearly uniform  $T_c \simeq 80$  mK.

The change in current caused by quasiparticles interacting with a TES held (via ETF) near its superconducting transition is readout with the circuit shown in Figure 4.6. The TES biasing circuit is depicted as well. The main

## 4.2. DETECTOR TECHNOLOGY

requirement of this readout circuit is to amplify the phonon signal without contributing significantly to its zero-energy noise resolution. The primary sources of noise are intrinsic to the QET and its biasing circuit. Any nonzero resistances at finite temperature will exhibit Johnson-Nyquist noise due to thermal agitation of charge carriers [737, 738]. This kind of thermal noise (like shot noise) is nearly white and exhibits an amplitude distribution that is approximately Gaussian [739]. It is commonly measured in terms of the corresponding rms current fluctuation in units of amps per square root of bandwidth (*i.e.*, in units of nA or pA per  $\sqrt{\text{Hz}}$ ). There are three nearly equal sources of Johnson-Nyquist noise in the QET and its biasing circuit: 1) the  $\sim 200\text{ m}\Omega$  resistance and 80 mK temperature of the biased TES; 2) the  $20\text{ m}\Omega$  biasing (or “shunt”) resistor coupled to the 600 mK stage of the cryostat; and 3) the few  $\text{m}\Omega$  of parasitic resistance coupled to the 4 K stage of the cryostat. Each of these contributes  $\sim 4\text{ pA}/\sqrt{\text{Hz}}$  of electronic noise to the phonon measurement, which adds in quadrature for a total of  $\sim 7\text{ pA}/\sqrt{\text{Hz}}$  [695].

As pictured in Figure 4.6, an array of Superconducting QUantum Interference Devices (SQUIDs)—inductively coupled to the TES biasing circuit—provides the current-sensing element required for low-noise, high-bandwidth amplification of QET signals. A change in current through the input coil induces magnetic flux in the SQUID, causing a change in voltage across its terminals. In response, a low-noise amplifier (with FET front ends) drives a current through the feedback coil that cancels the change in magnetic flux through the SQUID. The result is an output voltage equal to  $NR_{\text{fb}}I_{\text{TES}}$ , where  $N = 10$  is the ratio of the number of turns in the input to output coils,  $R_{\text{fb}} = 1\text{ k}\Omega$  is the feedback resistance, and  $I_{\text{TES}}$  is the change in current through the TES. The actual SQUID circuitry is considerably more complex as it has to be able to modify the amount of magnetic flux trapped in the SQUID array (see, *e.g.*, Sae Woo Nam’s Ph.D. thesis [691]). The particular dc-SQUIDs used by CDMS are described in detail in [740, 741]. The important thing to note is that they contribute (in quadrature) only  $\sim 2\text{ pA}/\sqrt{\text{Hz}}$  of electronic noise to the phonon channel. During normal operations, the combined noise performance of the phonon measurement (including the QET and biasing circuit) tends to vary

from  $\sim 10\text{--}20\text{ pA}/\sqrt{\text{Hz}}$ , depending on detector, frequency and day. The  $1\sigma$  phonon-signal noise resolution (averaged over Run 21) of  $\sim 100\text{--}250\text{ eV}$  (for all but one detector) is comparable to the ionization resolution.

### 4.2.3 Recoil Energy & Ionization Yield

For each scattering event, the ionization measurement allows inference of an ionization energy  $Q$  from the total number of liberated charge carriers, and the phonon measurement enables deduction of the energy  $P_{\text{total}}$  from the total phonons. Neither of these is an accurate representation of the energy deposited during a WIMP-nucleus interaction— $Q$  because of nuclear-recoil quenching, and  $P_{\text{total}}$  because of drift-heat amplification. In order to constrain WIMP models, observed and expected differential event rates must be cast in terms of the same observable. Typically this is the recoil energy  $E_R$  (as described in Section 3.3). As discussed in Section 3.3.5, for experiments like DAMA/LIBRA and CoGeNT that perform a single quenched energy measurement,  $E_R$  is obtained by dividing by the nuclear-recoil quenching factor. An unfortunate consequence is that any systematic uncertainty in the quenching factor is incorporated into the recoil-energy estimate, and will ultimately translate to uncertainty in the  $m_\chi\text{--}\sigma_{\chi\text{--}N}$  plane. This is perhaps the main reason it is unclear if the DAMA/LIBRA annual modulation and CoGeNT excess are consistent. The CDMS method for constructing recoil energy is fundamentally different. Rather than rely on a possibly uncertain quenching model,  $E_R$  is deduced directly from a combination of energy measurements. Traditionally this is done by subtracting the drift heat (see Equation 4.2.3) from the total phonon signal on an event-by-event basis:

$$E_R = P_{\text{total}} - \frac{eV}{\epsilon}Q, \quad (4.2.4)$$

where (again)  $V$  is the absolute value of the detector bias voltage,  $e$  is the charge of the proton, and  $\epsilon = 3.0$  ( $3.8$ ) eV for Ge (Si). This method yields an accurate recoil energy for both electron and nuclear recoils without *a priori* knowledge of the type of interaction. The main drawback is that errors on both  $Q$  and  $P_{\text{total}}$

#### 4.2. DETECTOR TECHNOLOGY

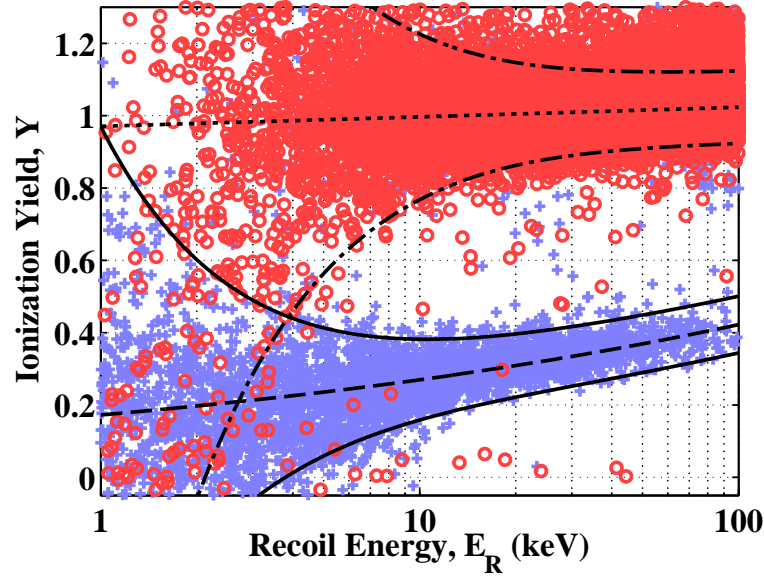


Figure 4.7: Demonstration of the CDMS ionization-yield discrimination parameter as a function of recoil energy for calibration data from Z3 (operated with a  $-6\text{ V}$  bias). Electron recoils induced by gammas from a  $^{60}\text{Co}$  source (circles) cluster near  $Y = 1$ , while nuclear recoils induced by neutrons from a  $^{252}\text{Cf}$  source (pluses) exhibit  $Y \simeq 0.2\text{--}0.4$ . The various lines indicate the average ionization yield—dotted line for electron recoils and dashed line for nuclear recoils—and the  $2\sigma$  “bands” within which 95% of all electron (dash-dotted lines) and nuclear recoils (solid lines) occur. The discrimination is nearly perfect above 10 keV, is somewhat worse from 2–10 keV, and breaks down completely below 2 keV.

are incorporated into the estimate. These errors, however, are statistical rather than systematic, and there is a well-defined recipe for their inclusion in the calculation of WIMP-nucleon cross sections (discussed further in Chapter 6).

The ZIP detectors’ two sensor technologies also provide a method for discriminating nuclear recoils from electron recoils. The discrimination parameter is a dimensionless quantity called “ionization yield,” and is defined as the ratio

## CHAPTER 4. THE CDMS EXPERIMENT

of ionization to recoil energy:

$$Y \equiv \frac{Q}{E_R} = \frac{Q}{P_{\text{total}} - eVQ/\epsilon}, \quad (4.2.5)$$

As demonstrated in Figure 4.7,  $Y$  is larger for electron recoils than for nuclear recoils, and provides near-perfect event-by-event discrimination for recoil energies in excess of 10 keV. Figure 4.7 differs from a traditional CDMS “yield plot” (see, *e.g.*, Figure 1 in [698]) in that a logarithmic scale has been used to display the recoil-energy axis. This highlights the low-energy region, where the ionization yield broadens due to irreducible electronic noise in the ionization and phonon measurements. A ZIP detector’s ability to discriminate gradually degrades as the recoil energy is decreased below  $\sim 10$  keV, and discrimination power is lost altogether for recoil energies below about 2 keV.

The energy scale for  $Q$  and  $E_R$  is calibrated using electron recoils caused by photon sources (described in detail in the next chapter). A number of spectral lines resulting from decays of internal radioisotopes (*e.g.*, the 1.3 and 10.4 keV lines from decays of  $^{68}\text{Ge}$  and  $^{71}\text{Ge}$ ) and from calibration sources (*e.g.*, the 1.17 and 1.33 MeV lines from a  $^{60}\text{Co}$  source) make this a relatively straightforward procedure. By construction, the resulting ionization yield for electron recoils is (on average) equal to unity (dotted line in Figure 4.7, *e.g.*).

The nuclear-recoil response is considerably more difficult to gauge. The primary reason is that exposing a detector to a source of neutrons typically results in a quasi-exponential energy spectrum; there are no spectral features with well-defined energies. In practice, CDMS has traditionally applied the energy scale derived from electron recoils directly to nuclear recoils, and used Monte Carlo simulations to check the validity of the resulting nuclear-recoil energy scale. The comparison of simulated and observed recoil-energy spectra for nuclear recoils will be discussed in Chapter 6 (see also Appendix E).

Another method for checking the nuclear-recoil energy scale is to compare the expected and measured ionization yields. The energy dependence of the ionization yield for nuclear recoils can be measured directly with ZIP detectors using fission neutrons from a  $^{252}\text{Cf}$  source. It is observed to smoothly

## 4.2. DETECTOR TECHNOLOGY

increase from  $\sim 0.2$  at  $E_R = 2 \text{ keV}$  to  $\sim 0.4$  at  $100 \text{ keV}$  (dashed line in Figure 4.7, *e.g.*). Theoretically, nuclear-recoil ionization yield in semiconductors is well described by Lindhard theory [719, 720], in which a material's stopping power is understood in terms of velocity-dependent electronic and nuclear components. Low-velocity nuclear recoils deposit most of their energy through nuclear interactions, while high-velocity electron recoils prefer electronic interactions. Since ionization results from electronic excitations, the ionization yield for nuclear recoils (relative to electron recoils) is smaller. This, of course, is the basis of CDMS electron-recoil discrimination. Lewin and Smith have simplified Lindhard theory to the following set of equations [588]:

$$\begin{aligned} Y(E_R) &= \frac{\kappa g(\varepsilon_z)}{1 + \kappa g(\varepsilon_z)}, \\ g(\varepsilon_z) &= 3\varepsilon_z^{0.15} + 0.7\varepsilon_z^{0.6} + \varepsilon, \\ \kappa &= 0.133 Z^{2/3} A^{-1/2}, \quad \text{and} \\ \varepsilon_z &= 11.5 E_R Z^{-7/3}, \end{aligned} \tag{4.2.6}$$

where  $E_R$  is in keV. A number of neutron-scattering experiments have attempted to test the recoil-energy dependence predicted by these equations. Figure 4.8 shows a compilation of measurements obtained with various Ge detectors compared to the Lindhard-theory prediction. The average ionization yield for a representative Ge ZIP (Z5 operated with a  $-3 \text{ V}$  bias) is shown as well. The comparison roughly confirms that the ZIP-detector energy scale, as derived from electron recoils, is appropriate for nuclear recoils. Again, a more thorough inspection of the nuclear-recoil energy scale will be discussed in the final chapter.

The general agreement shown in Figure 4.8 is important for another reason. Due to resolution effects, the ionization yield for nuclear recoils is difficult to measure with ZIP detectors for recoil energies  $\lesssim 2 \text{ keV}$ . The average ionization yield for the detector shown in Figure 4.8 was measured directly only between 2 and  $100 \text{ keV}$ . For the analysis described in the next chapter, a sub- $2 \text{ keV}$  relationship between  $Y$  and  $E_R$  is needed to estimate the energy dependence of a few detection efficiencies for the lowest-energy nu-



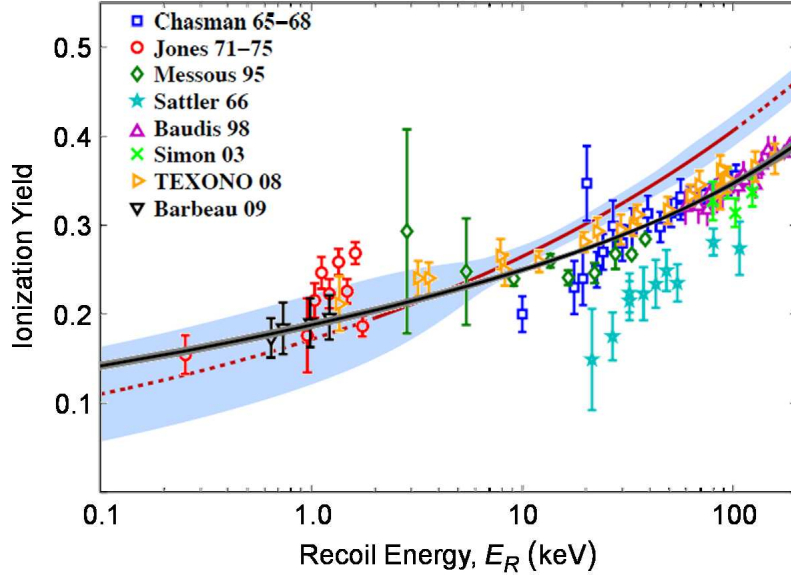


Figure 4.8: Compilation of ionization-yield measurements for neutrons scattering in Ge (data points with error bars) compared to Lindhard theory (black/solid line), and the average ionization yield as measured (between 2 and 100 keV) with a representative Ge ZIP (dark red/solid line). The latter is extrapolated below 2 keV and above 100 keV (dark red/dotted lines) via a power law such that  $Y = 0$  at 0 keV. Note that the associated uncertainty band (light blue/shaded region) differs from the  $2\sigma$  nuclear-recoil band shown in Figure 4.7; it represents the 95% C.L. statistical uncertainty in the determination of the average, whereas the band in Figure 4.7 is due to the phonon and ionization channels' nonzero resolutions. The experimental data is from (top to bottom in the legend) [742, 743, 744, 745, 669, 746, 747, 748, 749, 750, 751]. Similar neutron-scattering data are available for Si in [752, 753, 754, 755]. Figure adapted from [756].

clear recoils. Given the low-energy agreement between Lindhard theory and data from other experiments, an extrapolation via a power law from 2 keV to lower energies such that  $Y = 0$  at 0 keV (dotted line in Figure 4.8, *e.g.*) provides a reasonable approximation. Note that at such low energies,  $E_R$  is less than but nearly equal to  $P_{\text{total}}$ ; there is very little drift heat associated

## 4.2. DETECTOR TECHNOLOGY

with a low-energy nuclear recoil. Consequently, the low-energy ionization-yield extrapolation used to estimate  $E_R$  represents a small correction to  $P_{\text{total}}$ . Although there is clearly some uncertainty in this correction, it is necessarily small since the correction itself is small. Furthermore, since the extrapolation underestimates the ionization yield relative to Lindhard theory, it is likely that it results in a slight overestimate of  $E_R$ . Overestimating  $E_R$ —in estimates of near-threshold detection efficiencies—effectively increases the recoil-energy threshold  $E_{\text{th}}$ , which reduces sensitivity to low-mass WIMPs and is thus a conservative choice. I make a concerted effort throughout this thesis to use conservative (but fair) estimates (with regards to low-mass WIMP sensitivity) whenever uncertainty regarding the ZIP detectors’ low-energy response arises.

The small amount of drift heat associated with low-energy nuclear recoils brings to light another issue. Nuclear recoils with  $E_R$  of only a few keV will appear to deposit very small ionization energies—a few hundred eV (or less). The irreducible electronic noise in the ionization channel is similar in size. Consequently, for the lowest-energy nuclear recoils, the ionization signals are often indistinguishable from electronic noise. The corresponding phonon signals are (on average) a factor of 5 (or more) larger, which makes it possible to include Ge (Si) recoil energies as low as  $\sim 1$  (2) keV in the analysis described in the next chapter. It also means that an ionization-energy threshold cannot be used. CDMS has traditionally implemented a threshold on  $Q$  in order to avoid misidentification of electron recoils as nuclear recoils due to ionization noise (for small  $E_R$ ). This severely limits detection efficiency for  $E_R \lesssim 7$  keV. Despite the possibility that electron recoils might invade the WIMP-search signal region, the low-threshold analysis presented in the next chapter avoids any minimum requirement on  $Q$  in order to preserve near-threshold detection efficiency and the chance to probe low-mass WIMP parameter space.

#### 4.2.4 Alternate Recoil-Energy Estimators

The recoil-energy estimate given in Equation 4.2.4 is the primary method used to define (and measure efficiencies for) the data-selection cuts described in the next chapter. I refer to this estimate as “ $Q$ -corrected” recoil energy because it uses the ionization energy  $Q$  on an event-by-event basis. Again, note that this method incorporates errors on both  $Q$  and  $P_{\text{total}}$ , and therefore results in an estimate that has worse energy resolution than either  $Q$  or  $P_{\text{total}}$  alone. Additionally, the inclusion of  $Q$  in both the numerator and denominator of Equation 4.2.5 introduces significant correlations between  $E_R$  and  $Y$ . Throughout the remainder of this thesis, the recoil energy and ionization yield are always  $Q$  corrected unless otherwise stated.

$E_R$  can also be estimated from  $P_{\text{total}}$  alone by scaling it according to the average ionization yield measured from calibration samples (*e.g.*, the dotted and dashed lines in Figure 4.7). This method works provided the type of interaction (electron or nuclear recoil) is specified *a priori*. Consider the electron-recoil calibration data in Figure 4.7, for example. Since the average ionization yield is equal to 1,  $Q$  for electron recoils is equal to  $E_R$  on average. An expression for recoil energy in terms of  $P_{\text{total}}$  can therefore be obtained by replacing  $Q$  with  $E_R$  in Equation 4.2.4 and solving for  $E_R$ :

$$E_R = P_{\text{total}} - \frac{eV}{\epsilon} E_R \longrightarrow E_R = \frac{P_{\text{total}}}{1 + eV/\epsilon}. \quad (4.2.7)$$

For Ge (Si) ZIPs operated with a  $-3$  ( $-4$ ) V bias,  $E_R = P_{\text{total}}/2$ , whereas for a  $-6$  V bias  $E_R = P_{\text{total}}/3$  ( $2P_{\text{total}}/5$ ). I refer to this estimate as the recoil energy corrected by electron-recoil ionization yield, or “ $Y_{\text{ER}}$ -corrected” recoil energy. Since it does not include event-by-event uncertainty in  $Q$ , it has superior resolution relative to  $Q$ -corrected recoil energy, making it particularly useful when studying a ZIP-detector’s response to x-ray and gamma-ray sources. For example,  $Y_{\text{ER}}$ -corrected recoil energy facilitates comparison of the energy dependence of a detector’s phonon resolution for different bias voltages.

## 4.2. DETECTOR TECHNOLOGY

A similar recoil-energy estimate can be derived from  $P_{\text{total}}$  for nuclear recoils. The main difference is that the drift heat associated with nuclear recoils varies nonlinearly as function of recoil energy. Let  $\langle Y_{\text{NR}} \rangle (E_R)$  represent the average ionization yield (as a function of  $E_R$ ) for nuclear recoils as measured from  $^{252}\text{Cf}$  neutron calibrations (dashed line in Figure 4.7, *e.g.*). I define the recoil energy corrected by nuclear-recoil ionization yield, or “ $Y_{\text{NR}}$ -corrected” recoil energy, according the following implicit relationship:<sup>2</sup>

$$E_R = \frac{P_{\text{total}}}{1 + \langle Y_{\text{NR}} \rangle (E_R) \text{ eV}/\epsilon}. \quad (4.2.8)$$

The resulting estimate includes electronic noise from only the phonon channel, and is thus more appropriate than  $Q$ -corrected recoil energy for estimating (nuclear-recoil) quantities that depend on only the phonon signal. In particular, the hardware and software energy thresholds (described in the next chapter) depend on the phonon signal alone. In order to interpret the associated threshold efficiencies in terms of a WIMP model, it is convenient to convert them to  $Y_{\text{NR}}$ -corrected recoil energy.

### 4.2.5 Position Reconstruction

The segmentation of a ZIP-detector’s phonon sensors into four quadrants provides a method for reconstructing an event’s position in the plane parallel to the detector’s top and bottom surfaces (“ $xy$  position”). The  $x$  and  $y$  directions are as indicated in Figure 4.4 (*i.e.*, events occurring underneath phonon channels C and D (A and B) are in the  $+x$  ( $-x$ ) plane, and channels A and D (B and C) correspond to the  $+y$  ( $-y$ ) plane). In practice, an event’s  $xy$  position can be measured from differences in either phonon-signal start times or pulse heights as follows:

---

<sup>2</sup>Note that this expression can be made explicit if  $\langle Y_{\text{NR}} \rangle$  is instead measured as a function of  $P_{\text{total}}$ . Despite being slightly awkward to implement, I use the implicit relationship to convert  $P_{\text{total}}$  to  $Y_{\text{NR}}$ -corrected recoil energy in this thesis. The explicit form was adopted for a refined low-energy analysis of CDMS II deep-site Ge data in [757].

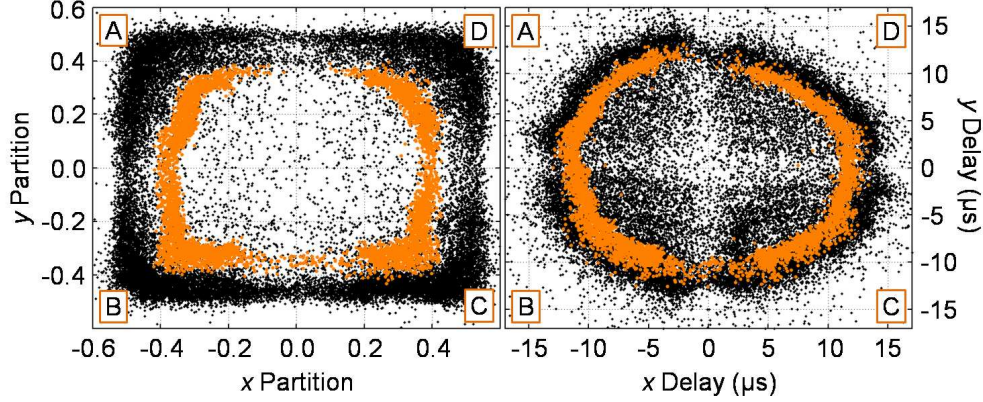


Figure 4.9: Demonstration of a representative Si ZIP’s ability to reconstruct  $xy$  position from the  $x$ - and  $y$ -partition (left panel) and -delay parameters (right panel). Each data point corresponds to an electron recoil induced by a uniformly illuminating  $^{60}\text{Co}$  gamma-ray source (dark/black dots). Events with more significant  $Q$ -outer than -inner signals (light/orange dots)—corresponding to events occurring near the edge of the detector—appear at smaller than expected partition and delay values due to a position-dependent phonon response (caused by a residual  $T_c$  gradient in the W TESs).

- **$x$  &  $y$  delays:** A start-time delay is measured for each phonon signal relative to the start time of the (more prompt)  $Q$ -inner signal. The “ $x$  delay” (“ $y$  delay”) is defined as the difference between the shortest such time (among the 4 phonon channels) and the start-time delay for the horizontally (vertically) adjacent channel. For example, if phonon-channel B has the shortest delay relative to the  $Q$ -inner signal, the  $x$  delay is the difference between the channel-B and -C delays. The  $x$  ( $y$ ) delay (in  $\mu\text{s}$ ) is negative if the shortest start-time delay occurs in channels A or B (B or C), and positive otherwise.

## 4.2. DETECTOR TECHNOLOGY

- **x & y partitions:** The total phonon energy is calculated from the sum of the four phonon channels’ pulse heights. The “ $x$  partition” (“ $y$  partition”) is defined as the fraction contributed from channels C and D (A and D) minus the fraction contributed from channels A and B (B and C). For example, if  $P_{\text{total}} = P_a + P_b + P_c + P_d = 2 + 1 + 3 + 4 = 10$  keV, the  $x$  ( $y$ ) partition is 0.4 (0.2).

Example  $xy$ -delay and -partition plots are shown in Figure 4.9 for a representative Si ZIP (Z4 operated with a  $-4$  V bias). The former roughly reproduces the circular shape of a ZIP detector’s profile (as well as the phonon-channel borders), while the latter maps  $xy$  positions to a distinctive square shape (sometimes referred to as a “box plot”).  $xy$ -delay plots for Ge ZIPs exhibit delays that are approximately twice as long, reflecting the slower speed of sound in Ge. The ability to reconstruct event position provides a useful diagnostic for monitoring the performance of a ZIP detector’s phonon channels. Additionally, any WIMP-candidate events will be spread uniformly in the  $xy$  plane, whereas events resulting from radiocontamination are often localized.

### Position Correction

The position reconstruction demonstrated in Figure 4.9 is not a one-to-one mapping of (physical) event location to either the  $xy$ -delay or -partition plane—there is degeneracy. This is evident from the (apparent) location of events exhibiting larger  $Q$ -outer than -inner signals (light/orange dots in Figure 4.9, *e.g.*). Such events most likely occurred directly beneath the  $Q$ -outer electrode, and were thus physically located at the largest possible radii in the true  $xy$  plane. Their reconstructed locations, however, appear at unexpectedly small values of the delay and partition parameters.

This degeneracy is caused by  $T_c$  gradients in the ZIP detectors’ W TESs. Despite efforts to control the variation in  $T_c$  via ion implantation, most of the Tower1 detectors exhibit residual  $\sim 10$  mK  $T_c$  gradients. Furthermore, one detector (Z1) was commissioned prior to the development of the ion-implantation technique, and exhibits a large 30–40 mK gradient.  $T_c$  gradi-

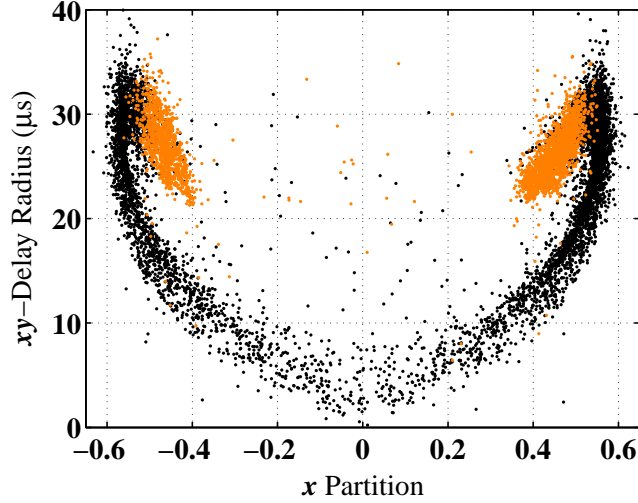


Figure 4.10:  $xy$ -delay radius as a function of the  $x$ -partition parameter for a restricted range of  $y$ -partition values ( $-0.25$  to  $0.25$ ), demonstrating how the degeneracy in the  $xy$ -position reconstruction can be broken for a representative Ge ZIP. Each data point corresponds to an electron recoil from a  $^{60}\text{Co}$  gamma-ray calibration. Events occurring under the  $Q$ -inner electrode (dark/black dots) are distinguished from those occurring under the  $Q$ -outer electrode (light/orange dots), highlighting the degeneracy of the delay and partition parameters when not used in conjunction.

ents can result in different phonon pulse heights for same-energy events occurring at different (physical)  $xy$  locations (*i.e.*,  $P_{\text{total}} = P_{\text{total}}(x, y)$ ). The  $xy$  variation in  $P_{\text{total}}$  contributes systematic uncertainty to recoil-energy and ionization-yield estimates, which results in degraded resolution and discrimination power.

The degeneracy in the delay and partition plots can be broken if the two types of position information are combined. One possible combination is shown in Figure 4.10, where the quadrature sum of the  $x$  and  $y$  delays (“delay radius”) is plotted as a function of the  $x$ -partition parameter for a representative Ge ZIP (Z5 operated with a  $-3$  V bias). Events near the center of the detector exhibit a well-behaved, quasi-linear correlation between the delay and partition pa-

## 4.2. DETECTOR TECHNOLOGY

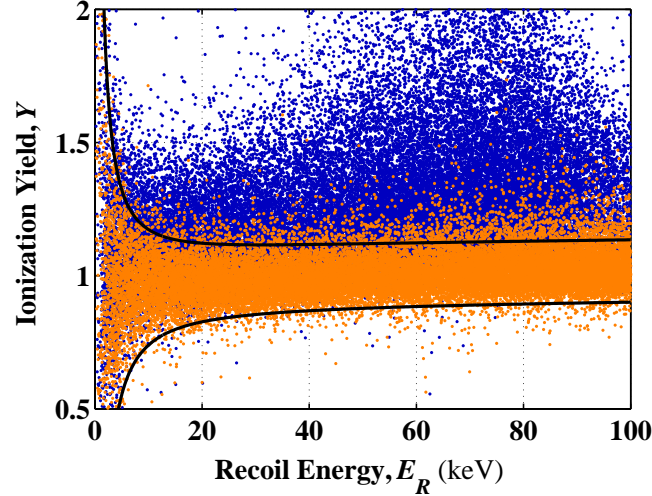


Figure 4.11: Comparison of electron-recoil ionization yield before (dark/blue dots) and after (light/orange dots) position correcting the phonon signals. The data are from  $^{60}\text{Co}$  gamma-ray calibrations for a Ge ZIP with a small  $T_c$  gradient (Z5 operated with a  $-6\text{ V}$  bias). The superior resolution of the position-corrected ionization-yield and recoil-energy estimates is obvious. The  $2\sigma$  electron-recoil band (solid lines) derived from the position-corrected quantities is shown as well.

rameters. Events located near the detector’s side wall—as indicated by events occurring underneath the  $Q$ -outer electrode—share common values of the delay radius and  $x$  partition with events located partway between the edge and the center; individually, the delay and partition parameters are double valued. The combination, however, clearly breaks the degeneracy by mapping each (physical) event location to a unique position in the delay-radius versus  $x$ -partition plane. More generally, a “position manifold” can be constructed. The manifold is embedded in a three-dimensional space in which the  $x$ - $y$  plane is defined by the  $x$ - and  $y$ -partition parameters, and a normalized delay radius—scaled to have values similar to the partition parameters—serves as the  $z$  coordinate. In this way, each physical  $xy$  position across the face of a ZIP detector is mapped to a unique set of delay-radius and partition coordinates.



## CHAPTER 4. THE CDMS EXPERIMENT

Blas Cabrera and Clarence Chang developed a method that uses this position manifold to correct the phonon signals as a function of event location (*i.e.*, to make the phonon response more uniform). For each well-behaved event from a gamma-ray calibration, manifold coordinates are calculated and assigned a set of correction values. The correction values are computed by averaging the phonon response over the event’s nearest neighbors on the manifold. The manifold positions (with their associated correction values) then serve as a “lookup table” for the event-by-event correction of phonon signals. This type of ZIP-detector “position correction” was first implemented for the Tower 1 detectors during Run 21. The main purpose was to correct Z1’s rise-time information to allow for reasonable performance of the phonon-based PSD. While the analysis presented in the next chapter does not use rise-time information, it benefits from the improved energy resolution of the corrected phonon signals. As demonstrated in Figure 4.11, the position correction results in superior resolution in the  $E_R$ - $Y$  plane, even for detectors with relatively modest  $T_c$  gradients. Consequently, the recoil-energy and ionization-yield estimates used in this thesis are based on the position-corrected  $P_{\text{total}}$ . The position correction was improved and extended for later exposures at the CDMS deep site. Some of the details can be found in Clarence Chang’s Ph.D. thesis [700].

### 4.3 The Shallow-site Installation

In this section I describe the infrastructure of the CDMS shallow-site installation. In order to reduce the flux of cosmic rays at sea level, the setup was located in an underground tunnel at the SUF. Both passive- and active-shielding elements helped to further moderate and reject the remaining flux of background radiation, providing the type of low-background environment required by a direct-detection experiment. The roughly cubical configuration occupied  $\gtrsim 5\text{ m}^3$ , and was the result of simulation and design work that is described in detail in Angela Da Silva’s Ph.D. thesis [692]. Maintenance of the active shielding—a plastic-scintillator muon veto—was my primary responsi-

### 4.3. THE SHALLOW-SITE INSTALLATION

bility during on-site shifts. Consequently, I will discuss the muon veto in somewhat more detail than the other components of the installation. A copper cryostat attached to a dilution refrigerator provided the low-temperature environment necessary for ZIP-detector operation. The design and construction of the custom-built cryostat (the “Icebox”) are described in detail in Peter Barnes’s Ph.D. thesis [758]. A variety of room-temperature electronics combined with a LabVIEW-based data-acquisition system allowed for experimental control and monitoring, signal amplification and shaping, and event triggering and readout.

The success of the shielding configuration at creating a low-background environment despite the modest depth of the installation was remarkable, at times allowing CDMS to compete (in terms of WIMP sensitivity) with experiments located much further underground. For example, the CDMS shallow-site result in [687] demonstrated comparable sensitivity to the 4800 m.w.e.-deep EDELWEISS (contemporary) result in [759]. In many respects, the superb performance of the shallow-site shielding set a new standard in the field of dark-matter direct detection.

#### 4.3.1 The Stanford Underground Facility

The shallow-site installation was located  $\sim 11$  m underground in Tunnel A at the SUF. The earth overburden is equivalent to  $\sim 17$  m of water, effectively stopping the hadronic component of the cosmic-ray flux while reducing the muon flux by a factor of  $\sim 5$ . The SUF’s three tunnels were originally excavated to serve as beam dumps for the Hansen Experimental Physics Laboratory’s (HEPL) linear accelerator, and are located at the south end of the HEPL End Station III building on the Stanford University campus. As shown in Figure 4.12, Tunnel A was extended to accommodate the CDMS I apparatus. The very end was partitioned (via a vinyl curtain) from the rest of the experimental hall and maintained as a clean area. The passive and active shields were located here, with the cryostat containing the ZIP detectors situated inside. Access to this area was restricted to individuals in full cleanroom garments.

## CHAPTER 4. THE CDMS EXPERIMENT

The room-temperature electronics, some general-purpose workspace, and the dilution refrigerator (as well as several of its associated components) occupied the remainder of the tunnel. Tunnel B housed vacuum pumps and gas canisters needed for the operation of the dilution refrigerator, and contained a modestly shielded area for storage of calibration sources (*e.g.*, a wax enclosure for the  $^{252}\text{Cf}$  source). Unused low-background materials (*e.g.*, extra lead and copper) were stored in Tunnel C to prevent cosmogenic activation by hadronic cosmic rays.

The primary disadvantage of the shallow-site installation was an irreducible neutron background that ultimately limited sensitivity to Galactic WIMPs, and led to the construction of a deep-site installation at the Soudan Mine in northern Minnesota. It is worth noting that the shallow site had a number of advantages not replicated at the deep site. Foremost was its accessibility. Easy 24-hour access at a location central to many of the institutions in the CDMS collaboration was critical during the development and testing of the ZIP-detector technology. Office space at the north end of the HEPL End Station III provided a convenient location for collaborative data analysis, and was close enough to the experimental apparatus to immediately make modifications to the hardware upon discovery of a malfunction. The overall result was a more efficient use of manpower than can be achieved at a remotely located, limited-access deep site. A limited-access installation also requires more sophisticated (and expensive) hardware and software to enable remote control and monitoring, complicating the experimental setup. Furthermore, sovereignty over the laboratory space at the shallow site allowed us to control all aspects of the experiment. One of the benefits of this access and control was a lower-noise electrical environment than has ever been achieved at the deep site, allowing for sub-keV hardware triggers for some of the Tower 1 Ge detectors during Run 21.

#### 4.3. THE SHALLOW-SITE INSTALLATION

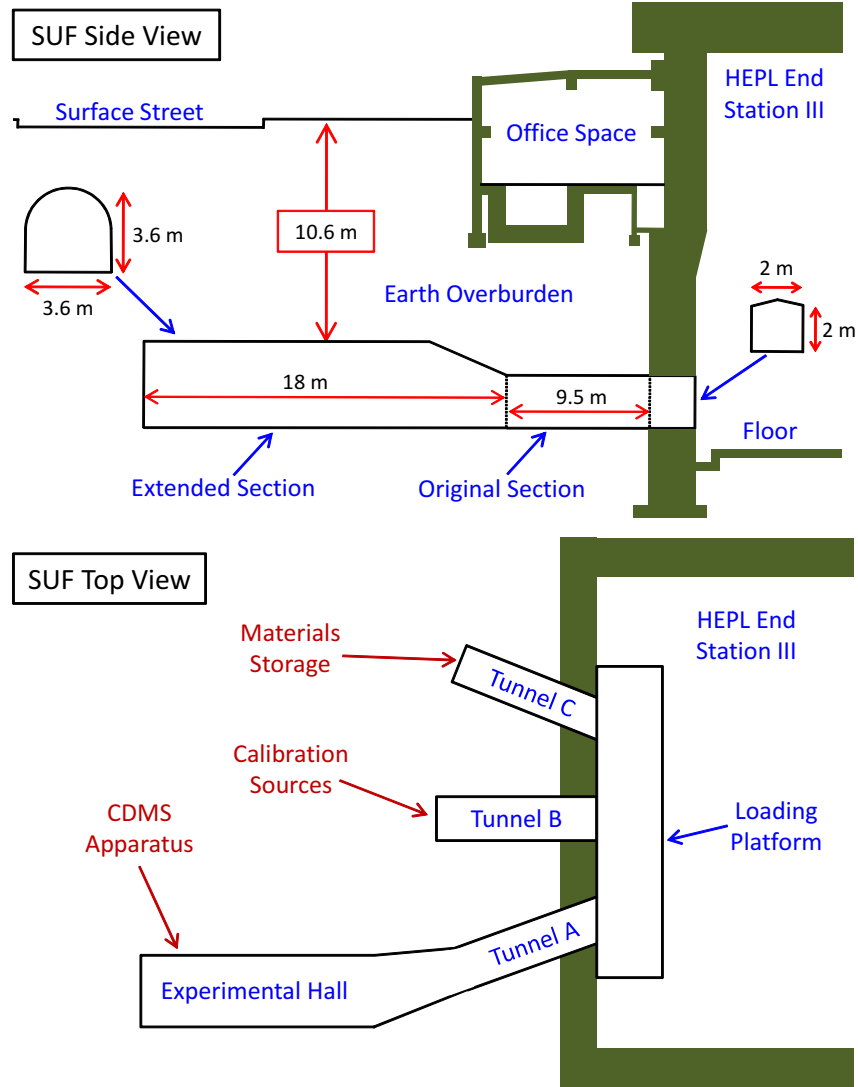


Figure 4.12: Layout of the Stanford Underground Facility (SUF). *Top*: Side view of the main experimental hall located at one end of the Hansen Experimental Physics Lab (HEPL) End Station III building on the Stanford University campus. The earth overburden is equivalent to  $\sim 17$  m of water. *Bottom*: Top-down view of the SUF's three tunnels. Figure adapted from [692].

### 4.3.2 Passive Shielding

As described in Section 3.3.3, the purpose of passive shielding is to moderate the gamma-ray and neutron backgrounds. This was achieved at the shallow site with tightly packed layers of lead and polyethylene. A cross-sectional view of the setup is shown in Figure 4.13. Immediately surrounding the detector cold volume (a multi-canned copper cryostat—see below) was a 25 cm-thick outer polyethylene shield to moderate the low- and medium-energy components of the neutron background (*e.g.*, neutrons from fissionable radionuclides and from  $\mu^-$  capture in the surrounding rock). Exterior to the outer polyethylene, an outer 15 cm-thick low-activity lead shield attenuated gammas from radioactivity in the surrounding rock (*e.g.*, the  $^{238}\text{U}$  and  $^{232}\text{Th}$  decay chains—see Figure 3.14), and in the materials used to construct the experiment (*e.g.*,  $^{60}\text{Co}$  is a radiocontaminant commonly found in steel). Lead and polyethylene located within the innermost 20 mK cold volume provided further shielding. These inner shields consisted of an additional 11 kg of polyethylene inside a 1 cm-thick ancient-lead pail. The inner polyethylene was not part of the original shielding configuration, but was added to reduce the neutron background measured during the final exposure of the CDMS I BLIP detectors [687]. Its addition allowed confirmation that the nuclear recoils observed by the BLIP detectors were neutrons rather than WIMPs; the Tower 1 ZIP detectors—operated during Run 21 with the inner polyethylene in place—observed the expected reduction (by a factor of  $\sim 2$ ) of the shallow-site neutron background [694].

Several penetrations in the outer lead and polyethylene shields allowed access to the interior of the apparatus. Built into the cryostat were two stems that provided access to the detector cold volume. As shown in Figure 4.13, the cold stem (“C-STEM”) penetrated near the bottom of the northern side, and provided a means to attach the cryostat to the dilution refrigerator. The electronics stem (“E-STEM”) penetrated near the top of the northern side, and allowed signals to be sent to and retrieved from the ZIP detectors. Additionally, there was a small hole with a removable plug in the south side of the outer lead shield, allowing (mostly) unattenuated exposure of the detectors

#### 4.3. THE SHALLOW-SITE INSTALLATION

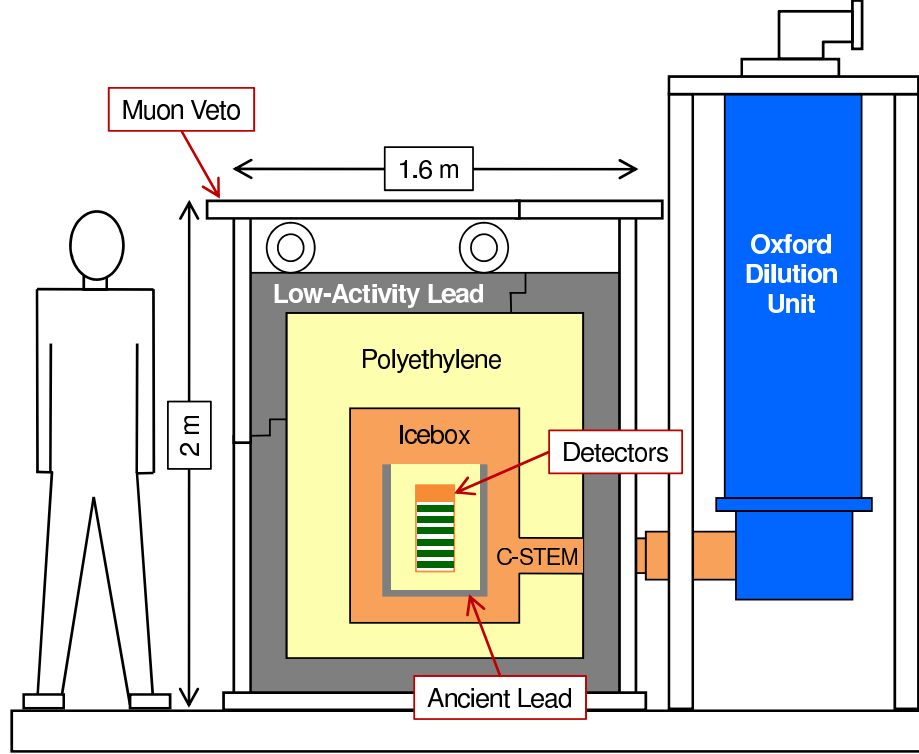


Figure 4.13: Cross-sectional view of the CDMS shallow-site shielding configuration. From the inside to the outside, the detectors were surrounded by an inner polyethylene shield, an ancient-lead pail, a copper cryostat (labeled Icebox), an outer polyethylene shield, an outer lead shield, and a plastic-scintillator muon veto. The Icebox connected to the dilution refrigerator via a cold stem (labeled C-STEM) that penetrated through the northern side of the outer shielding layers. Also shown is the region of the outer lead shield (along the top and south sides) that rolled away via steel wheels in order to access the interior. Figure adapted from [685].

to gamma-ray calibration sources. In order to conduct effective nuclear-recoil calibrations with neutrons from a  $^{252}\text{Cf}$  source, the top and south sides of the outer polyethylene shield had to be partially removed. This was a somewhat involved operation. First, a hand-cranked crane was used to move and suspend

some of the muon-veto counters. A section of the outer lead shield (depicted in Figure 4.13) was then rolled back onto a steel support structure via a set of steel tracks and wheels. With access to the interior of the lead shield, it was then a simple matter to unstack several of the 1 inch-thick polyethylene sheets. Due to the danger of activating the detectors, neutron calibrations were necessarily (and thankfully) infrequent.

### 4.3.3 Active Shielding—The Muon Veto

A significant flux of cosmic-ray muons penetrates the SUF’s modest overburden. From the inset in Figure 3.13, the muon flux is  $2\text{--}3 \times 10^{-3} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}$  at a depth of 17 m.w.e. This equates to  $\sim 500$  muons per second incident upon the CDMS I apparatus. While muons passing through a ZIP detector are easily identified by the large number of charge carriers they create, those that interact in the passive shielding (but not in the detectors) are dangerous because of their ability to generate secondary neutrons that can mimic WIMP-nucleus interactions.  $\lesssim 400$  such neutrons per kilogram-day of exposure are expected for the shallow site’s passive-shielding configuration [685]. A near-perfect efficiency for detecting muons with the active-shielding layer described below was therefore critical to the success of the WIMP searches conducted at the SUF.

#### Physical Description

Offline rejection of muon-coincident ZIP-detector interactions was achieved by surrounding the outer lead shielding with a (nearly) hermetic 5 cm-thick plastic-scintillator muon veto. The east and west (south and north) sides were each covered by one (two) scintillator counter(s). Scintillation light from particle interactions in each of these counters was detected with two photomultiplier tubes (PMTs) optically coupled to bars of wavelength shifter (WLS) slightly spaced from the counter’s edges. There were two circular holes in the shared border of the north-side counters to accommodate the C-STEM and E-STEM. Since underground muon fluxes are largely vertical, the top side of any muon veto is the most important. Consequently, the top side of the shallow-site ap-

### 4.3. THE SHALLOW-SITE INSTALLATION

paratus was covered by three counters instrumented with a larger number of PMTs: two main counters each readout with four PMTs coupled to WLS bars, and a smaller “crack” counter placed above the border between the two main counters and readout directly by two PMTs coupled to its east and west ends. Additionally, there were four side-by-side counters that spanned the bottom side (from north to south), and were each monitored directly by a single PMT. These bottom counters were separated by small gaps ( $\sim 20\%$  of the bottom-side area) to accommodate the steel I-beams upon which the passive shielding was built. Detailed geometries for the 13 muon-veto counters are shown in Figure 4.14.

The muon veto’s principle of operation is remarkably simple. Minimum ionizing muons deposit (on average) at least 10 MeV in the 5 cm-thick plastic scintillator, while gammas from the surrounding rock and other materials have an endpoint of  $\sim 2.6$  MeV (from the decay of  $^{208}\text{Tl}$  in the  $^{232}\text{Th}$  chain). Signals due to muons can therefore be differentiated from those due to gammas via simple pulse-height discrimination. Since gammas will not penetrate the lead shields, being able to distinguish between the two types of activity is important; there is no benefit in rejecting ZIP-detector events that are accidentally coincident with the higher-rate gamma-ray activity. The 26 analog PMT signals were reduced to 13 channels (one for each muon-veto counter) via summing amplifiers. 30 mV discriminators converted the 13 analog signals to logical pulses that were continuously monitored by a history-buffer unit. The resulting record of muon-veto hits was recorded to an event record whenever a ZIP-detector caused an experimental trigger. The veto signals were also connected to a variety of other electronics that did not contribute directly to ZIP-detector event records. These were used to monitor each counter’s event rate as a function of time, control each PMT’s bias voltage, and diagnose and “tune” the PMTs’ gains.

Muon-veto hits typically occurred tens of  $\mu\text{s}$  prior to correlated ZIP-detector activity (*i.e.*, the veto signals were even more prompt than the ZIP-detector ionization signals—this will be discussed further in the next chapter). Consequently, a ZIP-detector interaction is classified as “veto coincident” if there



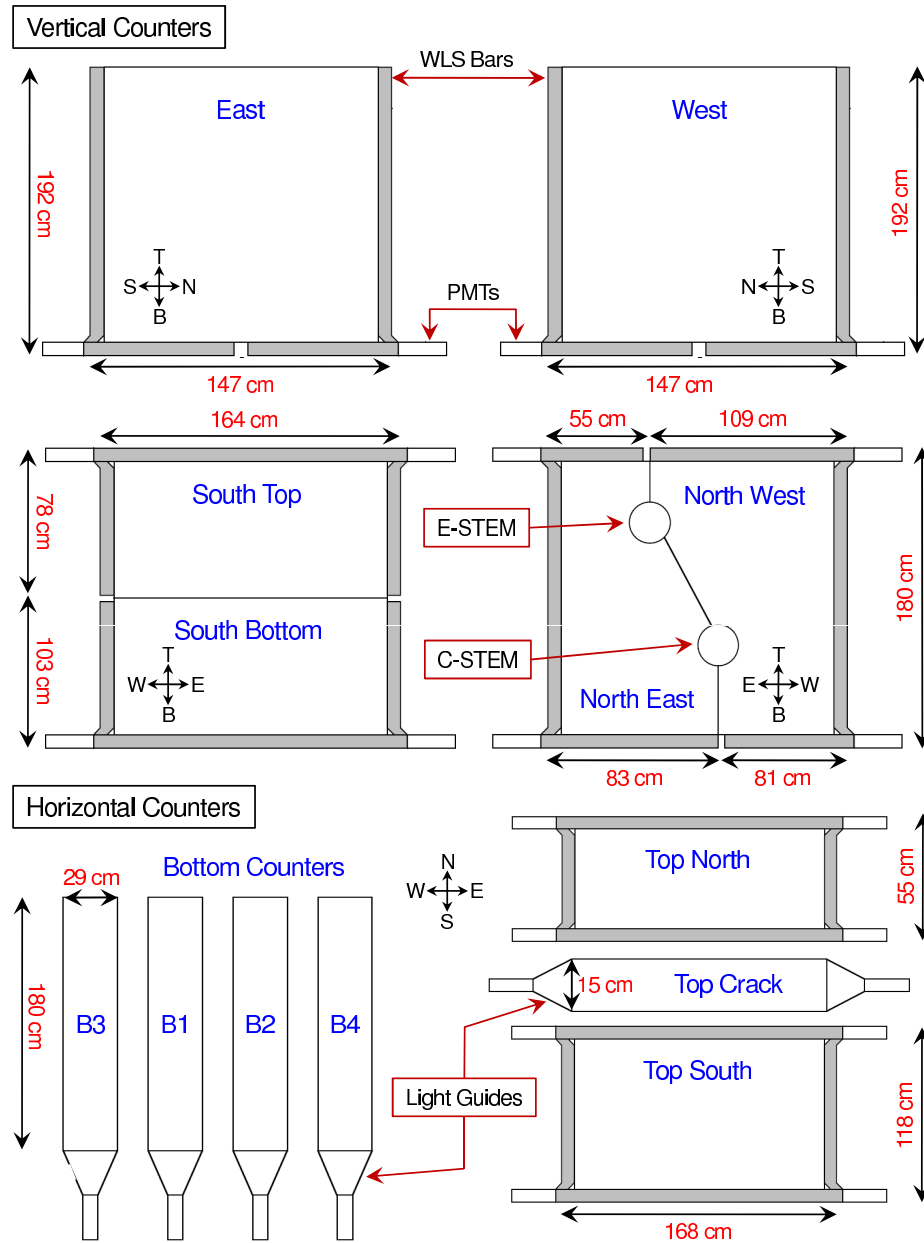


Figure 4.14: Detailed geometry of the CDMS shallow-site plastic-scintillator muon-veto counters. Figure adapted from [692].

#### 4.3. THE SHALLOW-SITE INSTALLATION

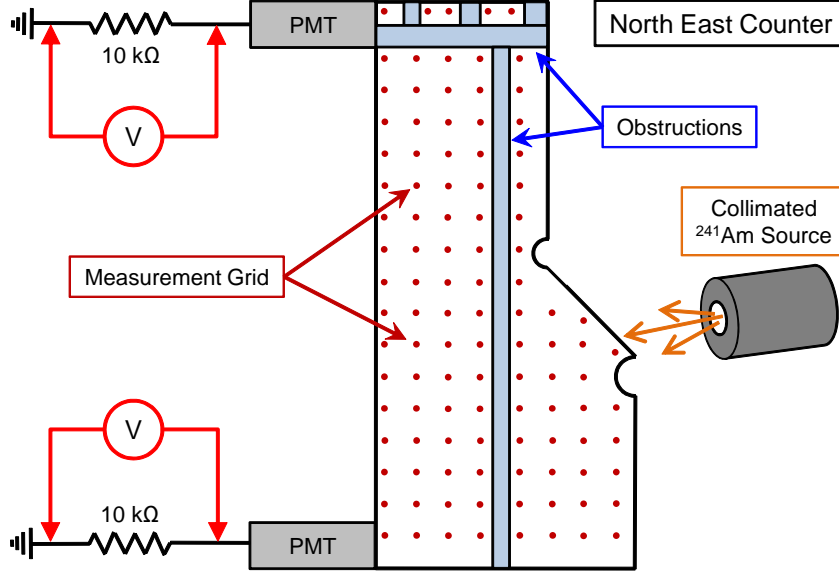


Figure 4.15: Schematic representation of how light-collection maps were measured for each muon-veto counter (north east counter pictured here). A collimated  $^{241}\text{Am}$  source was moved from point to point along a measurement grid. Activity induced by the source’s gamma emission was readout for each PMT (at each grid point) as an average current across a  $10\text{ k}\Omega$  resistor. The resulting maps (see, *e.g.*, Figure 4.16) allowed identification of regions with poor light collection, and provided a method to balance the relative gains among a counter’s PMTs.

were any hits in the muon-veto counters within a  $50\text{--}80\text{ }\mu\text{s}$  window—depending on detector and phonon pulse height—prior to the event trigger. If not, it is classified as “veto anticoincident,” and might be a viable WIMP candidate. The size of the rejection window, coupled with the rate of activity in the veto, determines the amount of dead time associated with rejecting veto-coincident events. For a  $50\text{--}80\text{ }\mu\text{s}$  rejection window and a nominal muon rate of  $500\text{ Hz}$ , the expected dead time is only  $\sim 2.5\text{--}4\%$ .

### Light-Collection Efficiency & Maintenance

Unfortunately, the geometry shown in Figure 4.14 turned out to be less than optimal with regards to light collection. This was particularly true for the somewhat awkward geometries of the two counters on the north side, and the large counters on the east and west sides. Additionally, the small air gap between the WLS bars and the edges of the scintillators—necessary for efficient collection of scintillation light—was compromised in places due to how tightly the counters fit together. As discussed below, this turned out to have important consequences regarding veto-associated dead time, and helped motivate the development of a more sophisticated readout system.

I originally diagnosed each counter’s light-collection efficiency with the measurement represented schematically in Figure 4.15. A collimated  $^{241}\text{Am}$  source was held up to a series of test points placed periodically across each counter’s outer surface. Activity induced by the source’s gamma emission ( $\langle E_\gamma \rangle \simeq 30\text{ keV}$ ) was readout from each PMT as an average current across a  $10\text{ k}\Omega$  resistor. The resulting light-collection maps for two of the counters are shown in Figure 4.16. These maps not only revealed regions of poor light collection, but were particularly useful for balancing the relative gains among a counter’s PMTs.

In some cases the maps helped to identify PMTs that needed to be replaced. A common malady involved He gas—boil-off from the daily transfer of liquid He to the dilution refrigerator—diffusing through the glass of the PMTs. The gas compromises the vacuum inside a PMT, causing a phenomenon known as afterpulsing [760] that slowly degrades the PMT’s gain as the concentration of He increases. The high voltage on affected PMTs has to be increased to offset the loss of gain and maintain sensitivity to muons, which also increases the rate of pulses not associated with muons (*e.g.*, pulses due to electronic noise and afterpulsing). A PMT can eventually become so saturated with He that it must be operated near its maximum rated high voltage. At this point it will contribute disproportionately to the overall rate of muon-veto activity, and it makes sense to simply replace it with a new PMT.

### 4.3. THE SHALLOW-SITE INSTALLATION

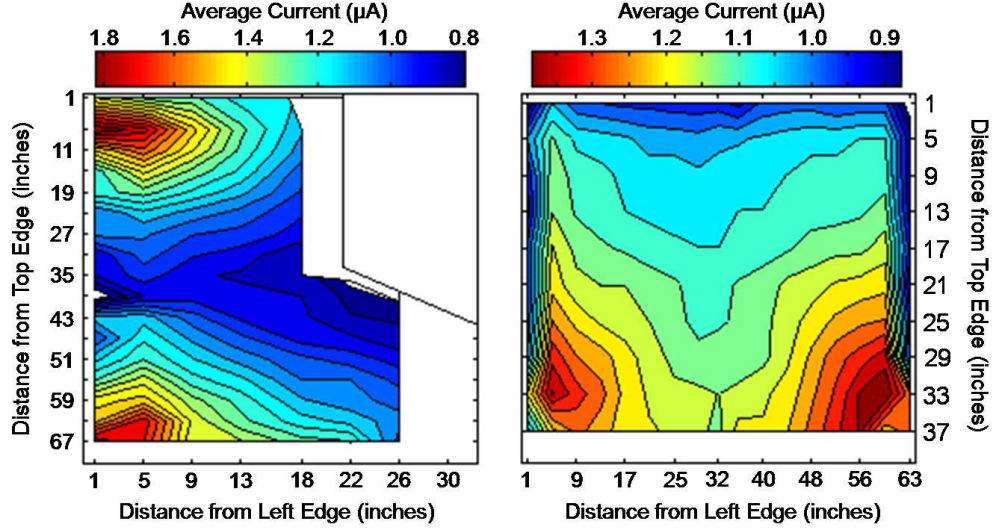


Figure 4.16: Light-collection maps for the north east (left panel) and south bottom (right panel) muon-veto counters. The two axes in each plot represent distance from the top and left edges of the muon-veto counter. For the north east counter, the rough outline of the counter’s shape is shown as well. Color represents the summed current readout of the two PMTs with the method shown in Figure 4.15. The regions close to the counters’ PMTs exhibit superior light collection (high currents). For the north east counter in particular, the regions of poor light collection required the PMTs to be operated at very high gains to ensure efficient tagging of muons.

Even after replacing several PMTs and tuning high voltages to spread light collection as evenly as possible across each counter, a large disparity between low- and high-efficiency regions was unavoidable for many of the vertically oriented counters. A consequence of this disparity is a degraded ability to distinguish muons from gammas. This is demonstrated in Figure 4.17, where the separation between muons and gammas is clearly much worse for the north-side counters than for the well-behaved crack counter. To ensure efficient tagging of muons passing through regions of poor light collection (*i.e.*, to prevent the tail of muons at low pulse height from slipping below the 30 mV discriminator threshold), the PMTs were operated at very high gains. This caused regions

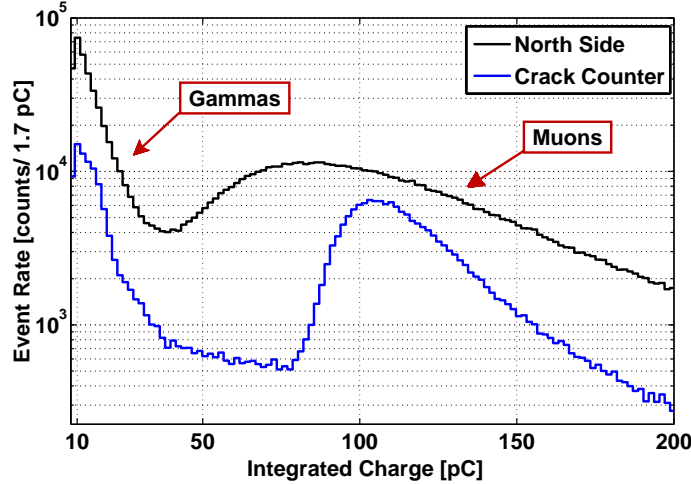


Figure 4.17: Energy spectra for two muon-veto channels, demonstrating the separation between interactions caused by muons and gamma rays. The degraded ability of the north-side counters to distinguish muons from gammas is obvious in comparison to the crack counter, whose simple geometry (and lack of WLS bars) exhibited superior light-collection efficiency. The spectra were acquired with the upgrade to the muon-veto readout electronics described below. Because the energies are derived from integrated veto pulses (hence the units of charge along the  $x$  axis), the separation between muons and gammas shown here is likely better than was achieved with the simple pulse-height discrimination employed by the original readout electronics.

with superior light collection (such as those in close proximity to the PMTs) to be partially sensitive to environmental gammas, increasing the veto-related dead time due to rejection of ZIP-detector events that were accidentally coincident with gamma-ray activity in the muon veto. The average rate of muon-veto interactions with pulse heights in excess of 30 mV was maintained near 5 kHz during Run 21, corresponding to a ZIP-detector dead time of  $\sim 20$ –30% (discussed in more detail in the next chapter). Preserving muon-tagging efficiency while limiting the muon veto's overall rate was a constant struggle.

### 4.3. THE SHALLOW-SITE INSTALLATION

#### Electronics Upgrade

The original muon-veto readout electronics were designed to generate a record of whether or not there was a veto pulse in excess of 30 mV at a given time. In this sense, the readout electronics provided a digital (yes or no) history of veto hits. The advantage of such a system is that the veto-related part of each ZIP-detector event record is remarkably simple, making offline deduction of correlations between ZIP-detector and muon-veto activity relatively straightforward. However, there is also an obvious disadvantage. As discussed above, there was a significant rate of gamma-induced pulses exceeding the 30 mV discriminator thresholds. Without (analog) knowledge of a veto hit's pulse height, there is no way to distinguish muons from gammas offline. The corresponding higher rate of veto-coincident ZIP-detector events was necessary for the counters exhibiting regions of poor light collection. For some of the horizontally oriented counters, however, the light collection was spread evenly enough that gammas were well separated from muons (see, *e.g.*, the crack-counter spectrum in Figure 4.17). Tuning such counters to be fully sensitive to muons while minimizing the number of gamma-induced pulses  $>30$  mV is not as straightforward as it might sound. In fact, to ensure a high tagging efficiency for muons, in most instances it is best to be cautious and use higher gains than are actually necessary. This prevents changes in the performance of the PMTs and readout electronics from compromising tagging efficiency, but increases the amount ZIP-detector dead time. For example, the dc-output levels of the analog summing amplifiers had a tendency to drift by a few mV relative to the 30 mV discriminator levels. Tuning the PMTs to higher gains ensured sensitivity to muons during times when this drift effectively increased the discriminator thresholds. An analog history of veto activity would allow some of the resulting ZIP-detector dead time to be recovered offline.

Another disadvantage involves muon-veto pulse shapes. The large amount of scintillation light caused by a muon often reflects many times before finding its way to a PMT. For those counters with large and/or awkward geometries, the resulting PMT pulses can be multi-peaked; the muon pulse is effectively

## CHAPTER 4. THE CDMS EXPERIMENT

spread out in time. The maximum pulse height (usually defined by the leading peak) can be small, while the total area (or charge) is large. In these instances, pulse height is not a reliable indicator of the amount of deposited energy. A pulse induced by a gamma-ray interaction might have a similar pulse height but only a single peak, and thus a smaller area. Simple pulse-height discrimination is therefore not an effective method for differentiating gamma-induced pulses from this class of small-pulse-height, large-area muon pulses.

These failings of the original readout electronics motivated my investigation into an alternative system. The basic idea is to incorporate an analog history of muon-veto activity into each ZIP-detector event record, an analog history based on PMT-pulse integrals (total charge) rather than pulse heights (maximum voltage). In concept, the system I developed closely resembles the readout of the ZIP-detector ionization electrodes shown in Figure 4.1. The main difference is that the PMT pulses are faster than the ionization signals, and require higher-bandwidth amplifiers. Additionally, the veto signals had to be conditioned (stretched in time) such that they could be recorded with a spare 5 MHz waveform digitizer we had on hand. A raw PMT pulse might rise to its maximum amplitude in 1–2 ns, and then decay exponentially in as few as 10 ns. Directly digitizing such pulses (and integrating offline) would require a sampling speed on the order of 1 GHz, which is not only prohibitively expensive for more than a few channels, but also results in excessively large data records. The alternative is to integrate the pulses in real time (with hardware rather than software), and allow the integrated charge to decay over the course of a few  $\mu$ s. The resulting pulses can then be accurately sampled with a much slower digitizer.

To test the new system, the circuit represented schematically in Figure 4.18 was incorporated into the muon-veto readout electronics during Run 21. The 26 PMT signals were combined via the original system’s summing amplifiers to reduce the number of channels to 8 (one for each of the muon veto’s six sides, one for the crack counter, and an OR of all channels). Multiplexing was necessary because the spare digitizer had only 8 input channels. Although they had some undesirable properties, the summing amplifiers’ multiple outputs per channel

### 4.3. THE SHALLOW-SITE INSTALLATION

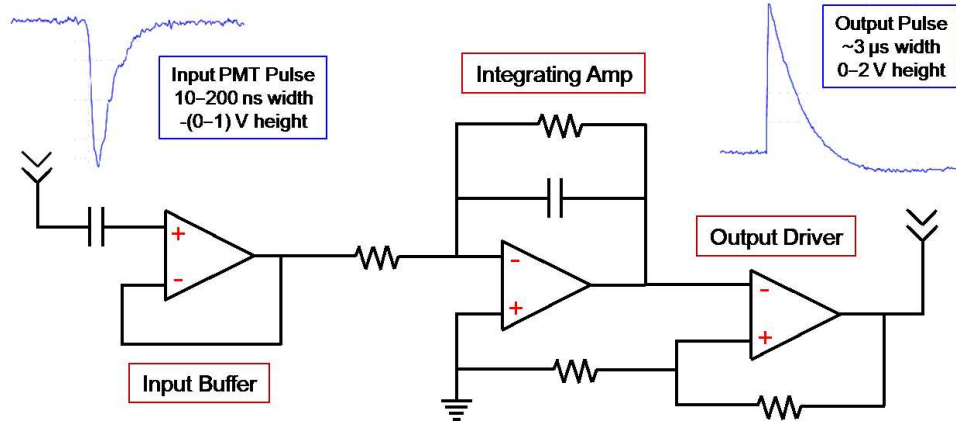


Figure 4.18: Schematic representation of the circuit used to integrate, stretch and amplify raw PMT pulses to be digitized into an analog history of muon-veto activity. Example input (left side) and output (right side) pulses are shown as well. See main text and Appendix A for further details.

provided the extra signal pathways required to incorporate the new electronics without modifying the original readout system. The new system was still in the prototyping and testing phase during Run 21, and not yet ready to supplant the original readout electronics. The signals from each channel were buffered by a capacitively coupled op amp (in a unity-gain follower configuration) to avoid any trouble with the summing amplifiers' drifting output levels. They were then integrated onto the feedback capacitor of a low-noise ( $6 \text{ nV}/\sqrt{\text{Hz}}$ ), high-bandwidth (400 MHz) op amp with FET front ends. The high bandwidth allowed pulses at the amplifier's input to be integrated on a time scale of only a few ns (*i.e.*, the charge on the feedback capacitor was continuously updated). Consequently, the multi-peaked muon pulses described above were fully integrated before any significant decay could occur. This is an important feature of the circuit; its output pulse heights are reliable indicators of the (full) energies deposited in the veto counters. The integrated charge slowly decayed with a time constant of a few  $\mu\text{s}$  (determined by the product of the feedback resistance and capacitance), and the resulting pulse was driven by an additional op amp to the waveform digitizer. The configuration of the final-stage amplifier was



## CHAPTER 4. THE CDMS EXPERIMENT

tailored to the 2 V input range of the digitizer. Raw PMT pulses with 0 to  $-1$  V amplitudes and 10–200 ns widths were thus integrated, stretched and amplified into 0–2 V pulses with  $\sim 3 \mu\text{s}$  widths. The latter were continuously digitized with a sampling period of 200 ns into a circular memory buffer. Following a ZIP-detector trigger, the digitizers continued to sample for a short time before the full-length traces were recorded to an event record. Offline, integrated PMT-pulse charges (in pC) were deduced by measuring digitized amplitudes (in volts), dividing by the gain of the final-stage amplifier, and multiplying by the feedback capacitance of the integrating amplifier.<sup>3</sup> In this way, an analog pre- and post-trigger history of muon-veto activity was reconstructed for each ZIP-triggered event.

The energy spectra shown in Figure 4.17 were derived from the new analog histories. Figure 4.19 shows spectra for all 8 of the instrumented channels for the first few months of successful operation during Run 21. Despite the difficulties described in the previous section, the separation between gamma- and muon-like energy depositions is quite good for most channels. Nevertheless, the qualitative difference between vertically and horizontally oriented counters speaks for itself. For the reasons described above, it is likely that the gamma-muon separation (versus integrated charge) is better than was achieved with the original readout system’s simple pulse-height discrimination. Imperfectly balanced gains for a larger number of PMTs accounts for the top-side spectrum’s relatively broad muon peak. Also shown are events reconstructed with the analog information that did not register with the older digital system (red histograms). One of the advantages of the new system is that it is inherently thresholdless, allowing reconstruction of veto activity for which the raw PMT pulse heights were smaller than 30 mV. In practice, however, an 8 pC software threshold was applied to prevent identification of veto sig-

---

<sup>3</sup>The value of the feedback capacitor shown in Figure 4.18 was 50 pF. A few sources of nearby and stray capacitance increased this to an effective value of 68–69 pF. Also, the gain of the final-stage amplifier was set to +2. However, the amplified pulses were driven through  $200 \Omega$  resistors into  $50 \Omega$  coaxial lines, and then terminated across  $50 \Omega$  resistors (at the digitizer’s inputs). This combination of resistors acted as a voltage divider, such that the overall gain of the circuit was  $2/5$ .

### 4.3. THE SHALLOW-SITE INSTALLATION

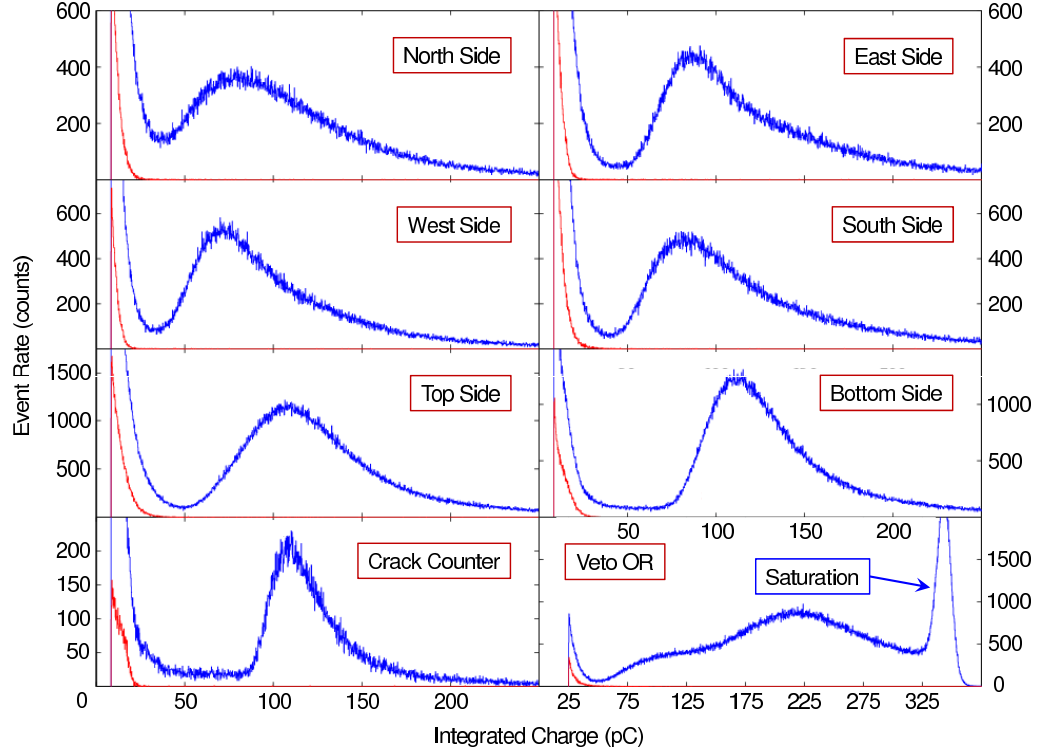


Figure 4.19: Energy spectra for all 8 analog muon-veto channels (blue histograms), reconstructed from the first few months of data recorded during Run 21, and compared to integrated charges for which there was no activity in the digital muon-veto history (red histograms). The differences between the blue and red histograms indicate the 30 mV discriminator thresholds (in terms of integrated charge) employed by the original readout electronics. Note that the  $x$  scales are identical for all but the lower-right panel. Veto-OR events are characterized by larger integrated charges because most muons deposit energy in more than one counter; the muon peak at  $\gtrsim 100$  pC ( $\gtrsim 200$  pC) corresponds to muons depositing energy in one (two) counter(s). The OR's expanded scale also shows a saturation peak, corresponding to the maximum output voltage of the integration circuit ( $340 \text{ pC} / (5/2) / 68 \text{ pF} = 2 \text{ V}$ ).

## CHAPTER 4. THE CDMS EXPERIMENT

nals due to electronic noise (25 pC for the OR). This corresponds roughly to a (single-peaked) 30 mV PMT pulse (in a  $50\,\Omega$  line) that has a rise time of a few ns and decays exponentially with an  $e^{-1}$  time of  $\sim 10$  ns. Note that the 30 mV discriminator threshold—the difference between the red and blue histograms in Figure 4.19—does not appear as a sharp cut off in terms of integrated charge. This is because the relationship between pulse height and area is not linear.

The analog histories acquired with the new readout system were not used (directly) in any of the Run 21 WIMP-search analyses. This was not because of a lack of performance. The muon-tagging efficiencies achieved with the analog and digital records were comparable. Probably the main reason the new data were not used is that it took time for the CDMS collaboration to be convinced that the system was robust. After all, I assembled the prototype circuit board by hand (shown in Figure 4.20), and Run 21 was its trial period. Additionally, analog histories were not available for the first part of the run because I was a bit late in bringing the new system on line. Another factor was that the gamma-muon separation for most of the charge spectra was not sufficient to confidently recover ZIP-detector dead time (relative to the digital histories) without compromising the muon-tagging efficiency. One could imagine defining veto-coincident events according to charge thresholds determined individually for each analog channel. For example, a 50 pC charge threshold would be sufficient to tag  $\sim 100\%$  of the muons in the bottom side and crack counter, while minimizing the number of the accidentally tagged gammas. For the other 6 channels, however, the low-charge muon tails extend well into the gamma distributions. In the end, despite the richer content of the analog histories, regions of poor light collection prevented effective discrimination of gammas from muons for events with  $\lesssim 50$  pC of integrated charge in 6 of the 8 analog channels.

The analog histories were not used for the analysis presented in the next chapter for a technical reason. The software I developed to analyze digitized veto traces had a predetermined look-back time.  $40\,\mu\text{s}$  was believed to be a sufficient window of time to find all possible ZIP-detector and muon-veto

### 4.3. THE SHALLOW-SITE INSTALLATION

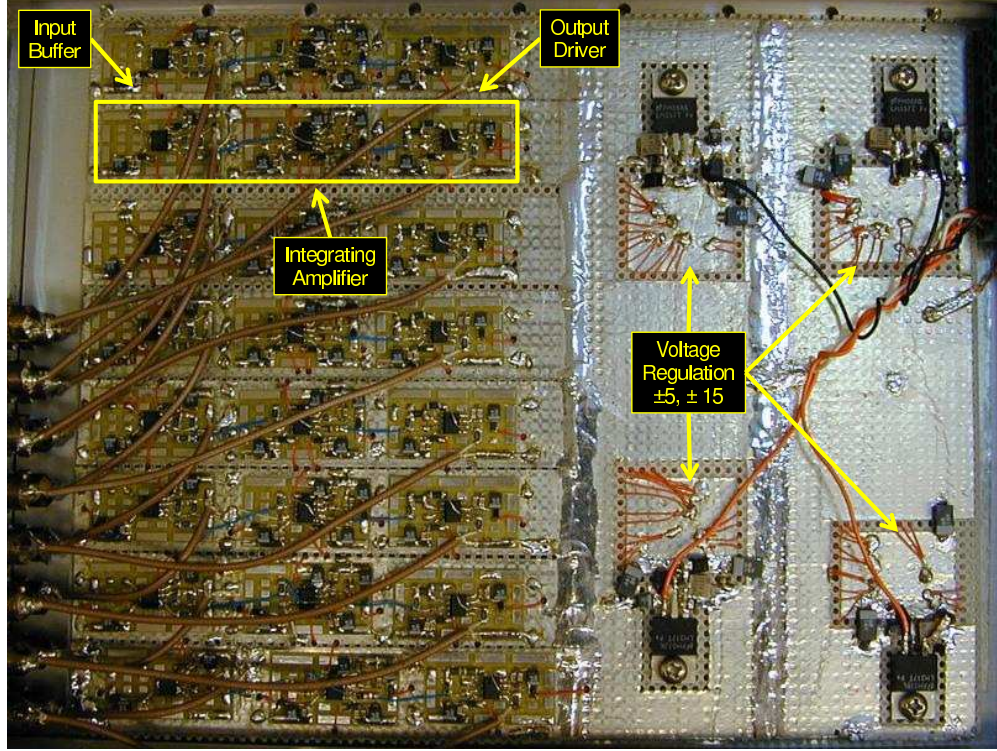


Figure 4.20: Photograph of the 8-channel prototype used for analog readout of the muon veto during Run 21. A single channel is highlighted, with the three amplification stages indicated. The regulation circuitry for the  $\pm 5$  V and  $\pm 15$  V power lines (required by the op amps) is pointed out as well. The entire board was packaged in a single-width NIM module, with power provided by the backplane of a NIM crate, and input and output signals made available via LEMO connectors on the module’s face plate. See Appendix A for further details.

correlations. I later discovered that (for some detectors) a  $70\text{--}80\,\mu\text{s}$  rejection window was necessary for events in which the ZIP-detector recoil energy was less than  $\sim 3.5\,\text{keV}$  (discussed further in the next chapter). Given the minimal dead-time recovery potential of the analog histories, it made more sense to use the digital histories rather than reanalyze the Run 21 data set with a modified veto-trace algorithm.

## CHAPTER 4. THE CDMS EXPERIMENT

One might be tempted to judge the electronics upgrade a failure based on the above comments. However, it was a complete success in the sense that it demonstrated the functionality of the circuit, and the potential for near-perfect separation of gamma- and muon-like energy depositions in veto counters with sufficiently uniform light-collection efficiency. Consequently, the CDMS II deep-site muon veto—constructed from 40 smaller plastic-scintillator counters each monitored directly by a single PMT (*i.e.*, no WLS bars)—was instrumented with quasi-redundant readout electronics based on the analog and digital schemes used at the shallow site. The main difference between the shallow-site and deep-site muon-veto electronics is that each PMT is readout individually, eliminating the need for the summing amplifiers. An extra follower was incorporated at the input-buffer stage of the circuit in Figure 4.18 to provide signals to the digital readout chain. For further information regarding the use of the deep-site analog histories, I refer the interested reader to Joel Sander’s Ph.D. thesis [703]. For posterity’s sake, the custom circuits used for analog readout of the shallow- and deep-site muon vetoes are documented in detail in Appendix A.

### Muon-Tagging Efficiency

The veto’s efficiency for tagging muons can be measured in a variety of ways. In this section I describe the use of muon-, neutron-, and gamma-like ZIP-detector events for measuring the tagging efficiency. In each case, the basic method involves the selection of a sample of ZIP-detector events that should have associated activity in the muon veto. Efficiencies are derived by comparing the number of events with correlations in the muon-veto digital histories to the total sample size. In some cases, an alternate method is used to estimate inefficiencies based on the number of events with correlations in only the bottom counters (relative to the total number with bottom-side correlations). This is viewed as an effective method for determining inefficiency in the other 5 sides because the rate of upward-going (relative to downward-going) muons is effectively zero (*i.e.*, any muon that hits a bottom counter should have reg-

### 4.3. THE SHALLOW-SITE INSTALLATION

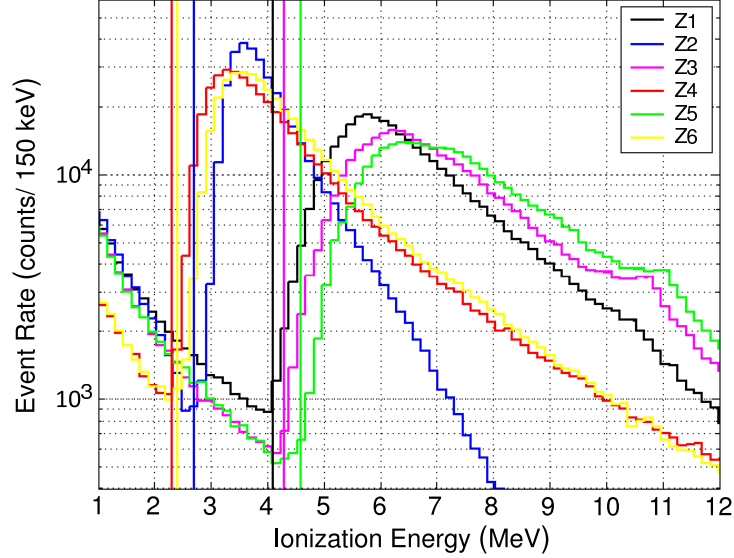


Figure 4.21: Muon-like ionization-energy spectra (histograms) for the Tower 1 detectors, compared to the minimum ionization energies (vertical lines) used to define the ZIP-detector through-going muon sample. The data are from the second half of Run 21 during the period in which the ZIPs were operated with a  $-6$  V bias. The muon peaks appear at lower energies than expected because the ZIP detector energy response is nonlinear, and was not calibrated to yield accurate MeV-scale ionization energies.

istered in one of the side or top counters). Since the events used to measure this sort of inefficiency were actually detected by the bottom counters (which cover  $\sim 80\%$  of the bottom side), the resulting “bottom-only efficiencies” are likely to be lower limits on the true tagging efficiency.

The most obvious measure of tagging efficiency uses events that are consistent with the large energies expected from ZIP-detector through-going muons. If the veto counters are in good condition, nearly all such events should have corresponding veto hits. The distributions of muon-like energy depositions for the Tower 1 detectors (as measured by their ionization channels) are shown in

## CHAPTER 4. THE CDMS EXPERIMENT

Figure 4.21 for a subset of Run 21 data.<sup>4</sup> Note that the average energy deposition expected for a muon passing through a Ge (Si) ZIP is  $\gtrsim 13$  ( $\sim 6$ ) MeV. The muon peaks in Figure 4.21 appear at lower energies because the ZIP detectors' energy scales are nonlinear outside the energy range of interest for WIMP-nucleus interactions, and were not calibrated to yield accurate ionization energies in excess of  $\sim 1$  MeV. The exact energy estimates are unimportant in this instance, as a sample of muons is easily defined relative to the muon-peak locations. A ZIP-detector muon sample was defined according the minimum ionization energies indicated in Figure 4.21. The fraction for which there was correlated activity (in at least one veto channel) in a  $40\ \mu\text{s}$  window (prior to the ZIP-detector trigger time) corresponds to a muon-tagging efficiency of  $99.92 \pm 0.01\%$ , while the bottom-only efficiency is  $99.88 \pm 0.01\%$ .<sup>5</sup>

It was realized during Run 21 that there is a low rate of interactions that deposit muon-like energies in the ZIP detectors, but are not necessarily veto coincident. These events arise when thermal neutrons capture on materials near the detectors (primarily Cu, Pb, Ge, and Si), sometimes creating a highly excited isotope that quickly de-excites via emission of a gamma-ray cascade. Although the associated gamma-ray energies are typically smaller than  $\sim 1$  MeV, they occasionally extend up  $\sim 10$  MeV (see, *e.g.*, the prompt gamma-ray neutron-activation data available from [761]). Since thermal neutrons are not necessarily correlated with activity in the muon veto, gamma-cascade events might be a source of bias in the above efficiency estimates. Even if the neutrons are generated by muons interacting in the passive shielding, it takes on the order of  $100\ \mu\text{s}$  for them to thermalize and be captured. Gammas resulting from such captures might not appear as veto coincident because the associated muon is likely to have interacted in the veto more than  $40\ \mu\text{s}$  before the ZIP-

---

<sup>4</sup>Ionization energy is used to define ZIP-detector through-going muons because the phonon measurement saturates at  $Y_{\text{ER}}$ -corrected recoil energies of only a few MeV for most detectors. Muons tend to create so much phonon energy that the TESs run out of dynamic range and saturate. For the Tower 1 Ge (Si) detectors, the phonon response to muons is fully (partially) saturated.

<sup>5</sup>Muon-tagging efficiencies are averaged over the entire Run 21 data set, and the errors represent binomial 90% confidence intervals (unless otherwise stated).

### 4.3. THE SHALLOW-SITE INSTALLATION

detector trigger time. Prompt gammas in the 7–8 MeV range from activated Cu and Pb isotopes (in particular) have large enough intensities to (potentially) affect the muon-tagging efficiencies at the level a few hundredths of a percent.

There are a couple of possibilities for removing the gamma-cascade events from the muon sample. One is to simply increase the ZIP-detector minimum ionization energies (used to select the muons in Figure 4.21) beyond the maximum energy a gamma cascade is likely to have. The difficulty with this approach is that the energy scale is unreliable for such high energies, making it difficult to judge where the minimum ionization-energy cuts should be placed without severely reducing the size of the muon sample. Fortunately, there is another handle; the cascade events tend to interact in a single ZIP, while muons usually interact in two or more. Further requiring ZIP-detector through-going muons to have deposited muon-like ionization energies in at least 2 ZIPs causes an insignificant reduction in statistics. The muon-tagging efficiency for this refined sample is  $99.97 \pm 0.01\%$ , while the bottom-only efficiency is  $99.90 \pm 0.01\%$ . Unfortunately, considering the reduced acceptance (in terms of solid angle), these estimates are clearly biased toward testing the efficiency of the top-side counters, for which it seems reasonable to conclude that the muon-tagging efficiency during Run 21 was at least 99.9%.

It is also possible to use veto-coincident nuclear recoils to estimate the tagging efficiency. In this case, the bottom-only method must be used to ensure exclusion of the shallow site's irreducible neutron background; high-energy neutrons can punch through the muon veto without registering, and then create secondary neutrons in the passive shielding that can result in ZIP-detector nuclear recoils. Although some punch-through neutrons are tagged by the muon veto due to the backsplash of radiation that can occur when they hit the outer-lead shielding, the muon veto was not (specifically) designed to reject them, and it is therefore not entirely fair to include them in estimates of the veto's muon-tagging efficiency. Instead, this estimate uses hits in the bottom-side counters to ensure that a muon passed through the shielding, increasing the likelihood that any correlated ZIP-detector nuclear recoils were caused by neu-



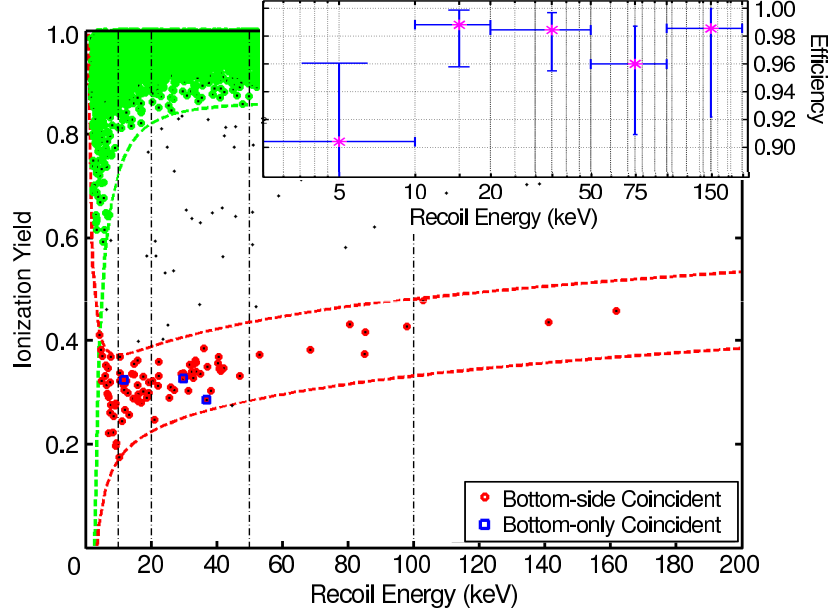


Figure 4.22: Selection of muon-veto bottom-coincident nuclear recoils (red circles) for a representative ZIP detector (Z5 operated with a  $-3$  V bias). The events that hit only a bottom counter (blue squares) can be used to estimate the veto’s muon-tagging inefficiency. The corresponding efficiency is shown in the inset (averaged over all ZIPs and the entire Run 21 data set). Unlike the efficiencies quoted in the main text, the vertical errors are 95% C.L. The horizontal error bars indicate the limits of the recoil-energy bins. Also shown in the main figure are the electron- (green dashed lines) and nuclear-recoil bands (red dashed lines). While most electron recoils (green circles) occur near an ionization yield of 1, note the increased tendency for some to leak toward the nuclear-recoil band at low energies (between-band black dots), possibly contaminating the lowest-energy bin of the tagging-efficiency estimate.

trons resulting from muons interacting in the passive shielding layers. This is precisely the event topology the muon veto was designed to reject. The expected rate of these events can be reduced from a few hundred to less than one (per kilogram-day of exposure) provided the muon-tagging efficiency is at least 99.9% [685].

### 4.3. THE SHALLOW-SITE INSTALLATION

Figure 4.22 demonstrates the selection of the bottom-coincident nuclear-recoil sample for a representative ZIP. Events must be consistent with being a nuclear recoil (in one of the ZIP detectors), but inconsistent with being an electron recoil. The merging of the nuclear- and electron-recoil bands at low energy (and the use of an ionization-energy threshold) limits the sample to recoil energies in excess of 5–8 keV (depending on detector). The lack of high-energy nuclear recoils further limits the sample to recoil energies  $\lesssim 200$  keV. For each of the recoil-energy bins indicated in Figure 4.22 (vertical dash-dotted lines in the main figure), a muon-tagging inefficiency is constructed by comparing the number of nuclear recoils with hits in only a bottom counter to the total number in the bin. The corresponding tagging efficiency (averaged over all ZIPs) is shown in the inset of Figure 4.22 as a function of recoil energy. When averaged over recoil energy, the efficiency is  $97.0^{+1.1}_{-1.5}\%$ . If the lowest-energy bin—possibly contaminated by surface events leaking into the nuclear-recoil band—is excluded from the average, the estimate increases to  $97.9^{+0.9}_{-1.4}\%$ .

It is also possible for muons to produce secondary particles that result in ZIP-detector electron recoils. Consequently, a tagging-efficiency estimate can be derived by considering bottom-coincident electron recoils. The event selection is virtually identical to the nuclear-recoil sample, except events must be (in)consistent with being a (nuclear) electron recoil. This sample differs from the through-going muon sample because the ZIP-detector ionization energies considered are much smaller ( $< 200$  keV). The bottom-only muon-tagging efficiency for the electron-recoil sample is  $99.50^{+0.05}_{-0.06}\%$ .

A possible explanation for the differences in tagging efficiency derived from the muon, electron-recoil, and nuclear-recoil samples is that the top counters are more efficient than the side counters. If true, the reasoning is as follows. Nearly all through-going muons hit one of the top counters, and thus provide an overestimate of the overall tagging efficiency, biasing it toward the higher efficiency of the top counters. Veto-coincident electron recoils (primarily gammas) originate from a greater internal volume than is probed by through-going muons. Consequently, since the parent muon did not necessarily pass through a top counter, they provide a more sensitive measure of the side counters' in-

efficiencies. Veto-coincident nuclear recoils will originate from an even greater internal volume; muon-induced neutrons can diffuse (and multiply) in a fashion that covers much larger distances than gammas. The nuclear-recoil sample is therefore the least biased toward the top counters. Furthermore, it may provide the only measure of inefficiency associated with the unvetoes E- and C-STEM penetrations (*i.e.*, a muon passes through a stem hole and one of the gaps between the bottom counters). This might explain why the efficiency estimate based on the nuclear-recoil sample is so much lower than the other estimates. Again, keep in mind that the bottom-only estimates are lower limits since the events used were actually detected by the one of the bottom counters. All things considered, the shallow-site muon veto's tagging efficiency was consistent with  $\sim 99\%$  during Run 21.

#### 4.3.4 The Dilution Unit, Cryostat, & Tower Assembly

As mentioned above, in order for the phonon sensors to function properly, the detector substrates must be cooled to well below the  $\sim 80$  mK superconducting transition of the tungsten TESs. A 20 mK base temperature was achieved at the shallow-site installation with an Oxford KelvinOx 400-S dilution refrigerator. Dilution refrigerators take advantage of a peculiar phase separation that occurs when a mixture of  $^3\text{He}$  and  $^4\text{He}$  is cooled to below 860 mK. The two isotopes separate into a lighter phase that is rich in  $^3\text{He}$ , and a heavier phase that is dilute in  $^3\text{He}$  (and rich in  $^4\text{He}$ ). Below about 2 K,  $^4\text{He}$  begins to turn into a zero-viscosity superfluid via Bose-Einstein condensation, whereas  $^3\text{He}$  does not because it has an overall half-integer spin. The  $^3\text{He}$  atoms in the dilute phase therefore behave effectively like a gas because the much higher concentration of  $^4\text{He}$  ( $\gtrsim 90\%$ ) is essentially noninteracting. The  $^3\text{He}$  atoms in the rich phase, however, behave as a liquid. As  $^3\text{He}$  moves across the phase boundary, a type of evaporative cooling occurs because its enthalpy in the two phases is different. The function of a dilution refrigerator is to cool a mixture of the two isotopes via a series of chambers, tubes, heat exchangers, and vacuum pumps such that this process of moving  $^3\text{He}$  across the phase boundary

### 4.3. THE SHALLOW-SITE INSTALLATION

can occur continuously. The reduction in temperature is typically characterized by six stages: the outermost surface at room temperature ( $\sim 300$  K), a liquid-nitrogen bath at 77 K, a liquid- $^4\text{He}$  bath at 4 K, the “still” at 600 mK, a cold plate at 50 mK, and the mixing chamber (containing the phase boundary) at 5 mK. The still is connected to the mixing chamber, and is used to remove  $^3\text{He}$  from the dilute phase to prevent saturation. Both the still and cold plate are coupled to heat shields that surround the mixing chamber. Similarly, the liquid-helium and -nitrogen baths encompass the colder stages and act as heat shields. A dilution refrigerator’s five internal stages can each be used to cool experimental components. For a more detailed discussion of the practical aspects of operating a dilution refrigerator, I refer the interested reader to [762].

For most low-temperature setups, the experimental apparatus is located directly beneath the dilution unit to facilitate direct thermal contact to its mixing chamber (see, *e.g.*, Figure 9 in [762]). Unfortunately, the components from which dilution units are constructed are not radiopure. Furthermore, the Oxford 400-S (shown schematically in Figure 4.23) is nearly 2 m tall. Considering the SUF’s space constraints, it would have been impractical (not to mention awkward and expensive) to include the refrigerator as part of the shielded inner volume of the CDMS I apparatus.

A radiopure custom cryostat, called the Icebox, and the “dog-leg” design shown in Figure 4.23 were used to create a detector cold volume well separated from the refrigerator (by the shielding layers described above), while simultaneously providing a thermal link to its six temperature stages. The Icebox consists of six concentric cans, each constructed from oxygen-free electronic (OFE) copper (99.99% pure). OFE copper is a relatively inexpensive, low-activity, and high-thermal-conductivity material, making it ideal for low-temperature and -background applications. The total Cu thickness of the Icebox also provided additional gamma-ray shielding. The refrigerator’s temperature stages were extended to the Icebox’s five outer cans via the C-STEM’s nested copper tubes, attaching to side flanges near their bases. A solid copper rod, coupled to the dilution unit’s mixing chamber via a cold finger, ran down the middle of the

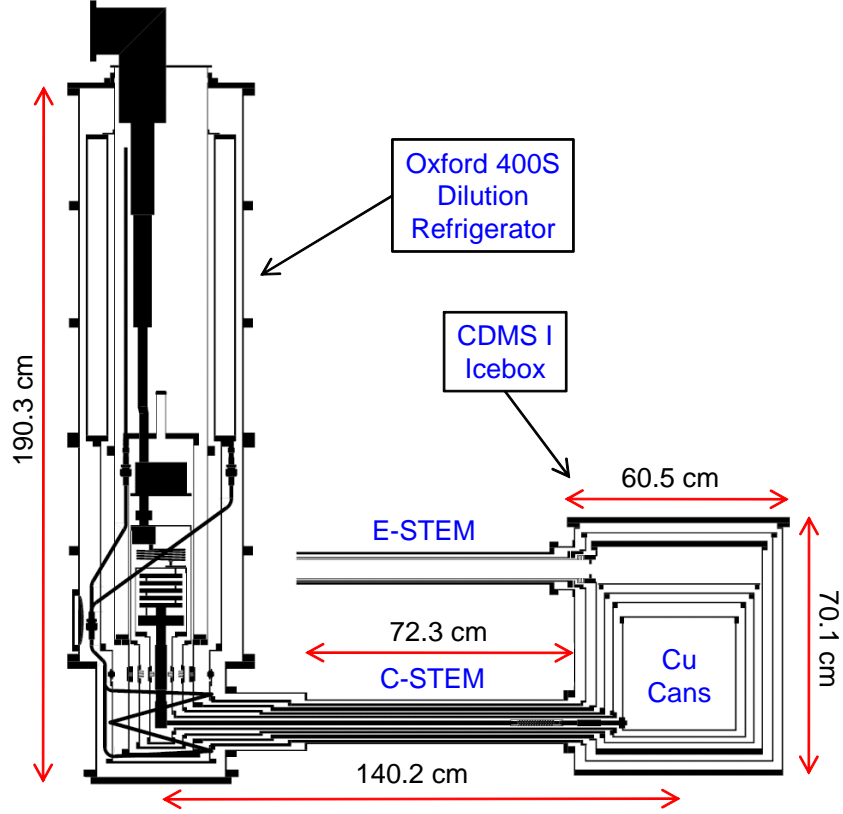


Figure 4.23: Cross-sectional schematic (side view) of the copper cryostat (Icebox) and Oxford KelvinOx 400-S dilution refrigerator employed at the CDMS shallow-site installation. The Icebox is cooled via a series of copper tubes known as the cold stem (C-STEM), while the electronics stem (E-STEM) allows passage of signals from within the Icebox to room temperature. Figure adapted from [685].

innermost C-STEM tube and was heat sunk (via flexible copper braid) to the innermost Icebox can. This inner can provided the ZIP-detector base temperature, and was typically maintained at  $\sim 20$  mK. The C-STEM was located near the bottom of the Icebox such that there was no direct line of sight between the detectors—suspended near the middle of the inner can—and the

### 4.3. THE SHALLOW-SITE INSTALLATION

dilution unit's radioactive components. Again, for further information on the design, fabrication and testing of the Icebox, see Peter Barnes's Ph.D. thesis [758].

The E-STEM's nested concentric tubes attached to side flanges on the Icebox's three outermost cans. The E-STEM contained copper striplines for carrying signals to and from the ZIP detectors.<sup>6</sup> This was one of the most challenging aspects of the thermal design; signals had to penetrate from room to cryogenic temperatures without overloading the refrigerator. Fortunately, it was possible to dissipate most of the associated heat load by heat sinking the striplines to the liquid-nitrogen bath at the 77 K stage. Cooling power is proportional to  $T^4$ . Consequently, the system was designed such that the greatest heat loads were coupled to the liquid-nitrogen and -helium baths. The dominant loads were due to blackbody radiation (peaked in the infrared) transferred from warmer to cooler stages. Due to the large temperature differences among the outer four stages, the primary loads on the 77 K, 4 K, and 600 mK stages were from infrared leakage. Additionally, the FETs used to amplify detector signals were coupled to the 4 K stage, while the slight amount of heat associated with the SQUIDS was handled by the 600 mK stage. The resulting consumption of liquid He required cryogens to be added to the dilution refrigerator on a daily basis, naturally breaking WIMP-search data runs into  $\sim 24$  hour periods.

The detector tower assembly, represented schematically in Figure 4.24, is a critical element of the experimental setup, providing the mechanical support and thermal and electrical connections required to house and operate the ZIP detectors and their readout electronics in a multi-staged low-temperature environment. The cryogenic electronics were attached to the top of the tower, which occupied the volume near the top of the 4 K can (shown in Figure 4.23 opening directly into the E-STEM). Through a series of heat sinks and mechanical supports, the tower extended through (and thermally coupled to) the

---

<sup>6</sup>The striplines are actually  $\sim 3$  m-long, 2.5 cm-wide strips of Kapton, a low-activity material [763] that is thermally stable across a wide range of temperatures, and is often used in the fabrication of flexible printed circuits. A thin,  $18\text{ }\mu\text{m}$ -thick layer of copper was etched onto the surface in a pattern that yielded 50 conductors for the transmission of signals to and from the ZIP detectors. A more complete description is available in [685].

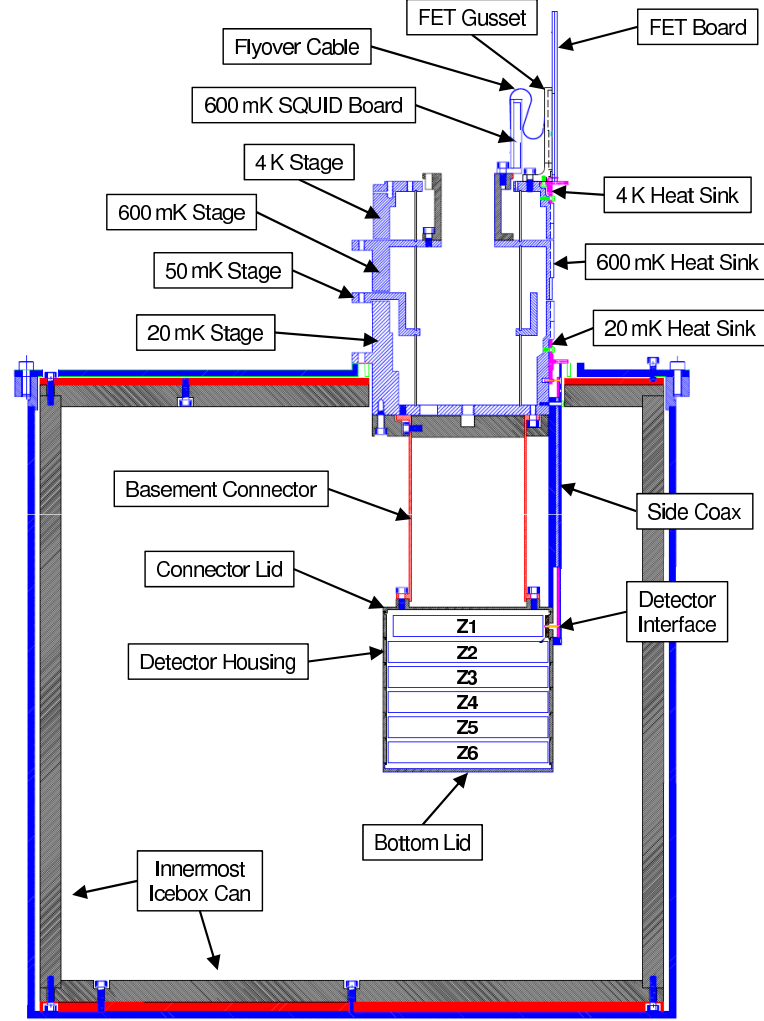


Figure 4.24: Schematic representation of a ZIP-detector tower assembly in relation to the innermost can of the Icebox. One of six SQUET (SQUID and FET) cards is depicted at the top of the tower. The flyover cable allowed transmission of signals from the SQUID board to the FET board without significant thermal coupling between the 600 mK and 4 K stages. Copper striplines then carried signals from near the top of the FET boards through the E-STEM to room-temperature electronics. Figure adapted from [695].

### 4.3. THE SHALLOW-SITE INSTALLATION

lids of the innermost Icebox cans. Just inside the lid of the innermost can, the ZIP detectors—arranged in a vertical stack inside a copper housing—were suspended via an  $\sim 10$  cm-long cylindrical tube (“basement connector”) down into (roughly) the middle of the can. Phonon and ionization signals were carried via a “side coax” from a detector interface near the top of the detector housing to the lid of the of innermost can, where they were continued to the cryogenic electronics at the top of the tower. A comprehensive discussion of each of the detector-tower components can be found in Chapter 5 of Sunil Golwala’s Ph.D. thesis [685].

#### 4.3.5 Room-Temperature Electronics & The DAQ

Rounding out the CDMS I experimental setup were several racks of room-temperature electronics and a LabVIEW-based data-acquisition system (DAQ). These were a mix of custom and commercial components that provided the electronic and software controls required to configure and monitor the active detector elements (ZIPs and veto counters), as well as condition their signals to form ZIP-triggered event records. I have already described in detail the muon-veto analog and digital histories, and will thus focus on the ZIP-related components in this section. A comprehensive discussion of the individual electronic components and DAQ software routines can be found in Sunil Golwala’s Ph.D. thesis [685].

At their room-temperature ends, the striplines connected to a rack of electronics containing custom Front-End Boards (FEBs). Each FEB contained the circuitry required to configure, monitor, and readout a single ZIP detector. This included detector biasing, SQUID biasing and control, controls for flashing the infrared LEDs (used for detector neutralization), individual analog readout (and amplification) of the phonon and ionization signals, and digital logic for remote communication with the DAQ software.

Phonon and ionization signals were transmitted from the FEBs to custom Receiver/Trigger/Filter (RTF) boards. The main purpose of these boards was to generate logical pulses (or triggers) from the ZIP-detector signals, while



## CHAPTER 4. THE CDMS EXPERIMENT

preserving (and further conditioning) them for eventual digitization. To accomplish this, the RTF boards—each connected to a single FEB and thus a single ZIP—created two copies of each phonon and ionization channel. The individual analog signals from one set were tailored to the dynamic range of the same type of waveform digitizer (Joerger VTR812) used to record veto traces. For ZIP-detector signals, the digitizers were configured to a 1.25 MHz sampling speed. Upon receipt of an event trigger, they were programmed to record a total of 2048 samples, yielding  $\sim 1.6$  ms-long traces. As mentioned previously, these traces were analyzed offline with a software optimal filter, resulting in a total of six reconstructed energies per detector per event. The other set of copies were combined to form a single ionization and a single phonon signal per detector. Each of these analog sums was subjected to a bandpass filter and compared to low- and high-level discriminator thresholds. Like the FEBs, the RTF boards included remote-communication logic, which allowed the discriminator levels to be controlled with the DAQ software. The resulting RTF-board logical pulses were continuously monitored with a time-stamp module. In this way, pre- and post-trigger histories of ZIP-detector phonon and ionization hits were included in each event record. Additionally, the RTF-board logical pulses were sent to a Trigger-Logic Board (TLB). The TLB was responsible for generating the experimental trigger from a configurable pattern of low- and high-level phonon and ionization hits.

As discussed previously, the phonon signal is an overmeasure of recoil energy due to drift-heat amplification, while the ionization signal is either an equal measure (for electron recoils) or an undermeasure (for nuclear recoils). The phonon measurement therefore provides the most sensitive indication of an energy deposition in one of the ZIP detectors. Consequently, when searching for WIMPs, the TLB was configured to OR the RTF boards' phonon low-level logical pulses across all six detectors to form the experimental trigger. The corresponding RTF discriminator thresholds were set as low as possible without allowing too many triggers from electronic-noise fluctuations. Careful tuning of these trigger levels ensured that the  $\lesssim 1$  Hz trigger rate (while searching for WIMPs) was dominated by true particle inter-

### 4.3. THE SHALLOW-SITE INSTALLATION

actions, while simultaneously maximizing the likelihood of an experimental trigger from the low-energy signals expected from WIMP-nucleus scattering events.

Note that the triggering phonon energy is slightly different than the reconstructed phonon energy. The 400 Hz and 5 kHz poles of the RTF boards' bandpass filters were chosen to resemble the software optimal filtering as closely as possible. However, the triggering phonon energy is the result of an analog sum that happens in real time, and is thus incapable of compensating for relative start-time differences among the four phonon channels. Offline, the phonon traces are fit independently, and four individual reconstructed energies are deduced. The total reconstructed phonon energy is formed from their sum, and is independent of the associated traces' start times. Consequently, the reconstructed phonon energy is equal to or (slightly) greater than the triggering phonon energy. The efficiency of the hardware trigger (a function of the triggering energy) exhibits a peculiar dependence on the reconstructed energy as a result. For previous CDMS analyses, this somewhat subtle aspect of the triggering was unimportant because recoil energies were limited to levels well above the energy at which the hardware trigger attains 100% efficiency. For the low-threshold analysis presented in the next chapter, the inclusion of trigger-level recoil energies requires a careful measurement of the hardware trigger efficiency. Since the triggering phonon energy is not available offline, reconstructing the trigger efficiency is a nontrivial aspect of the analysis.

The DAQ software ran on three Power Macintosh clones, and consisted of several custom-written LabVIEW software modules with embedded C and Java subroutines (as well as commercially-available hardware drivers) capable of interfacing with the room-temperature electronics. The DAQ provided what was referred to as slow and fast control of the experiment. The former was used to configure the experimental components (*e.g.*, PMT high voltages, FEB bias voltages, and RTF-board discriminator levels), and continuously monitor various rates and voltages (about once per minute) to determine if the experiment was operating within nominal parameters. The fast-control

## CHAPTER 4. THE CDMS EXPERIMENT

software was used to record calibration and WIMP-search data, as well as coordinate periodic LED flashing and the acquisition of randomly triggered events (“random triggers”). This primarily involved an event-based loop in which the DAQ armed the digitizers and synchronized (in time) the history buffers, read them out to an event record whenever a trigger was issued from the TLB, and then started over again. A data run began with 200 initial random triggers, and was followed by triggered data acquisition via the event loop in a 4 hour cycle. At the end of each 4 hour period the event loop was suspended, and the detectors were flashed with the infrared LEDs to neutralize any charge traps. Additionally, the DAQ activated the dilution refrigerator’s temperature and pressure sensors—otherwise disabled to prevent noise pickup on the ZIP-detector signals—and monitored the cryogenic system for the  $\sim 15$  minutes it took to flash the LEDs and subsequently allow the detectors to cool sufficiently for the phonon sensors to again function properly. The DAQ was also configured to intersperse randomly generated software triggers during each 4 hour period such that  $\sim 5\%$  of all triggers were random, providing a significant sample of “empty” events for characterizing the ZIP detectors’ zero-energy response. While searching for WIMPs, this cycle of 4 hours of triggered data acquisition, followed by 15 minute servicing periods, was automatically repeated until it was necessary to halt the DAQ in order to add cryogenics to the refrigerator. Each  $\sim 24$  hour data-taking period (or “series”) was named according to its unique starting date and time. The “series number” was constructed from a numerical representation of the year, month, day and time (in that order). For example, *020304-1346* refers to a series that was started on the 4<sup>th</sup> of March in 2002, and began at 1:46 pm (PST). When exposing the ZIPs to calibration sources, it was often desirable to break the data series into shorter periods (*e.g.*, to prevent overexposure of the ZIPs to neutrons). Through the heroic efforts of Dan Bauer, this (now outdated) LabVIEW-based system was capable of sustained data acquisition rates  $\gtrsim 3$  Hz, which was more than adequate for the low WIMP-search trigger rate, but limited the rate at which calibration data could be acquired.

### 4.3. THE SHALLOW-SITE INSTALLATION

The DAQ packaged raw data into 500- or 1000-event chunks, which it moved to a cluster of Linux-based PCs for reduction with a collection of algorithms (in custom-written MATLAB code) called *DarkPipe*. The purpose of this code was to reduce the raw trace data and history-buffer patterns to a set of “Reduced Quantities” (RQs) for further analysis. RQs summarize the physical characteristics of an event, while reducing the associated disk space by an order of magnitude. Among the RQs derived from phonon traces, for example, are pulse heights, start times, rise times, and a few prepulse baseline quantities (to determine the presence of multiple pulses or excessive electronic noise). History-buffer records were converted to a variety of relational pre- and post-trigger times, with associated bit patterns (“masks”) that can be dereferenced to determine which channels of which ZIP detectors or veto counters had a hit at that time. A second-pass analysis package called *PipeCleaner* (also custom-written MATLAB code) was used to convert ZIP-detector pulse-height RQs to useful energy units via the application of calibration factors, and to calculate derived quantities such as ionization yield and the  $xy$ -position parameters. These second-pass quantities were referred to as “Relational Reduced Quantities” (RRQs). During Run 21, the position correction discussed in Section 4.2.5 was added as a final step in the *PipeCleaner* analysis, yielding additional RRQs. The analysis presented in the next chapter is based on a combination of *DarkPipe* and *PipeCleaner* RQs and RRQs derived from the WIMP-search and calibration data recorded during Run 21.



## Chapter 5

### A Low-Threshold Analysis

As mentioned in the previous chapter, traditional CDMS WIMP-search analyses have employed recoil-energy thresholds well above the hardware trigger level, ignoring the lowest-energy recoils in favor of optimal sensitivity to the favored neutralino models for which the WIMP mass is  $>40 \text{ GeV}/c^2$  (see, *e.g.*, Figure 2.18). With the evidence presented in Chapter 3, there is ample motivation for lowering the CDMS analysis threshold in order to target lower-mass WIMPs. Unfortunately, for  $E_R \lesssim 10 \text{ keV}$ , finite-resolution effects are expected to degrade the discrimination performance of the ZIP detectors (see, *e.g.*, Figure 4.7), resulting in possibly significant backgrounds in the low-energy signal region. Nevertheless, given the relatively large WIMP-nucleon cross sections favored by the DAMA/LIBRA and CoGeNT data (see, *e.g.*, Figure 3.21), even a background-limited CDMS WIMP-search analysis has the potential to be sensitive to the favored parameter space. The primary challenge is to preserve sufficient low-energy detection efficiency, while preventing electronic-noise fluctuations from entering the signal region. Consequently, although the Run 21 exposure of the first tower of ZIP detectors is known to be limited by the shallow-site neutron background, the superb zero-energy noise resolution achieved during this run make it a viable data set for exploring low-mass WIMP parameter space.

## CHAPTER 5. A LOW-THRESHOLD ANALYSIS

In this chapter I present a first attempt to include recoil energies down to  $\sim 1$  (2) keV in a WIMP-search analysis using Ge (Si) ZIP detectors.<sup>1</sup> After a brief review of the Run 21 data samples, I discuss the phonon-energy thresholds associated with the hardware trigger and a software cut used to reject events due to electronic-noise fluctuations. Then I describe the analysis cuts (*i.e.*, WIMP-candidate selection cuts) and their associated efficiencies for detecting WIMP-like events (*i.e.*, nuclear recoils). Since it is difficult to distinguish nuclear recoils from electron recoils at the lowest energies considered, a simplified set of analysis cuts whose detection efficiencies have little to no energy dependence is adopted. Finally, the ZIP-detector energy response to nuclear and electron recoils is discussed, both in terms of absolute scale and resolution. Particular care is taken to understand the zero-energy resolution, as this plays an important role in the calculation of WIMP-nucleon upper limits. The results of the low-threshold analysis described below are presented in Chapter 6.

### 5.1 Run 21

The Tower 1 detectors were originally commissioned at the CDMS shallow-site installation during the second half of 2001. After the tower was installed and the Icebox and shielding layers reassembled, there was a long period of cooling the detectors (and inner shielding layers) to the 20 mK base temperature. This was followed by SQUID tuning to obtain optimal noise performance, several days of LED baking to neutralize charge traps, and tuning of the RTF-board discriminator thresholds. Additionally, a variety of gamma-ray sources (*e.g.*,  $^{137}\text{Cs}$ ,  $^{57}\text{Co}$ , and  $^{60}\text{Co}$ ) were used to establish the ionization and phonon energy scales (for electron recoils), as well as compile sufficient events to construct pulse templates for the optimal-filter fitting routine. The experiment appeared to achieve a tuned and stable state after a period of  $\sim 2$  months (September and October of 2001). Unfortunately, in early November it was discovered that superfluid  $^4\text{He}$  had leaked into the detector cold volume, form-

---

<sup>1</sup>Actually, much of the material presented in this chapter and the next first appeared in [596], a publication featuring the results of the low-threshold analysis discussed here.

## 5.1. RUN 21

ing films that gradually coated the ZIPs and caused a damping of the phonon response. This abruptly ended what is referred to as Run 21a. No data from this commissioning period were used in the analysis presented below.

After warming the apparatus to extract He from the cold volume and fix the leak, the experiment was cooled back to base temperature and data taking resumed in the beginning of December. Having already gone through the LED baking and detector tuning program once, the experiment was successfully configured in a much shorter period of time. Gammas from a  $^{137}\text{Cs}$  source were used to recalibrate the ZIP-detector energy response, and neutrons from a  $^{252}\text{Cf}$  source induced nuclear-recoil events from which new pulse templates were constructed. The acquisition of viable WIMP-search data began during the third week of December. The experiment operated stably for nearly 7 months, acquiring a combination of WIMP-search and calibration data ( $^{60}\text{Co}$  and  $^{252}\text{Cf}$ ). In the third week of June 2002 the dilution refrigerator crashed, bringing Run 21b to a close. After a few weeks of repairs, additional neutron and gamma-ray calibration data were acquired during the month of July, but no further WIMP-search data were recorded. This final month of operation is referred to Run 21c.

### 5.1.1 Data Samples

#### WIMP Search

The WIMP-search data recorded during Run 21b were split into two operational periods. The Ge (Si) detectors were initially operated with a  $-3$  ( $-4$ ) V bias voltage (applied to the ionization electrodes) for 94 days between December 2001 and April 2002. This data set yielded just over 66 live days, and will henceforth be referred to as the “3V data.” As mentioned in the beginning of the last chapter, the 3V data were originally analyzed with a 5 keV recoil-energy threshold, and the resulting exclusion limit on the spin-independent WIMP-nucleon cross section—which included statistical subtraction of the neutron background—is (to this day) one of the most constraining for WIMP masses between 10 and 20 GeV/ $c^2$  [694] (see also Tarek Saab’s Ph.D. thesis [695]). During



## CHAPTER 5. A LOW-THRESHOLD ANALYSIS

the second operational phase, all ZIPs were operated with a  $-6$  V bias voltage over a period of 74 days between April and June 2002, yielding an additional 52 live days. These data, henceforth referred to as the “6V data,” were analyzed with a 10 keV recoil-energy threshold in Don Driscoll’s Ph.D. thesis [696]. The reason for modifying the bias voltage was to explore the dependence of the ionization-yield and surface-event discrimination parameters on drift heat. Despite superior charge-collection efficiency for events occurring near the detector surfaces, the 6V data exhibited a degraded ability to reject surface events via the phonon-based PSD. Overall, the performance of the CDMS discrimination parameters was slightly better for the 3V data. Consequently, except in a few special cases (see, *e.g.*, [764]), Ge (Si) ZIPs were operated with a  $-3$  ( $-4$ ) V bias voltage during the WIMP searches conducted at the CDMS II deep-site installation.

For the low-threshold analysis presented below, the entire Run 21b WIMP-search data set is considered. Since phonon-based PSD is not used, and the  $Y$ -based electron-recoil discrimination breaks down at low energy regardless of bias voltage, the slight difference in discrimination performance between the 3V and 6V data sets is unimportant; significant backgrounds are expected in the signal regions for both. However, the relative difference in drift-heat amplification between the two bias voltages results in different phonon energy scales. Discussed further in Section 5.4, this has an important effect on the recoil-energy resolution. Overall, the detector response is different for the two bias voltages, making it necessary to analyze the 3V and 6V data sets separately. In the following sections, everything is effectively calculated twice for each detector, once for the 3V data and once for the 6V data.

### Calibration Data

A mixture of various types of calibration data were recorded during all three phases of Run 21. Most of the data from Run 21a were taken in relatively short series to test adjustments made to the experimental apparatus, and are therefore either statistically insignificant or simply unreliable. Again, none of

## 5.1. RUN 21

these data were used in the following analysis. At the beginning of Run 21b, roughly half a million  $^{137}\text{Cs}$  events were used to test detector neutralization, and to establish the calibration factors used by *PipeCleaner* to convert pulse-height RQs (in volts) to energy RRQs (in keV). Calibration of the ZIP-detector phonon and ionization energy scales involved comparison of observed and simulated  $^{137}\text{Cs}$  Compton-recoil differential event rates, and is discussed at length in Tarek Saab’s Ph.D. thesis [695]. Other than relying on the  $^{137}\text{Cs}$ -based calibration factors to calculate RRQs, I did not directly use the  $^{137}\text{Cs}$  data from Run 21b in the analysis that follows. Other methods for gauging the ZIP-detector energy scales that make use of WIMP-search and  $^{252}\text{Cf}$  data are described below and in the next chapter.

The primary electron-recoil calibration data were acquired by exposing the detectors to relatively high-energy gamma rays from a  $^{60}\text{Co}$  source. The  $\gtrsim 1$  MeV gammas from this source were energetic enough to fully illuminate each ZIP detector despite the inner shielding layers. A gated trigger was used to select events with energies between the low- and high-level RTF-board (phonon) discriminator thresholds, promoting acquisition of electron recoils in the energy range of interest for WIMP-nucleus interactions. Three  $^{60}\text{Co}$  calibrations were performed for each bias voltage. A relatively short exposure was conducted at the beginning of Run 21b, and was followed by more extensive exposures mid-way through and after the WIMP search. The  $^{60}\text{Co}$  data series provided the large statistics required to construct the lookup table used in the event-by-event position correction described in Section 4.2.5. Additionally, each detector’s electron-recoil band in the  $E_R$ - $Y$  plane (see, *e.g.*, Figure 4.7) was calculated from these data.<sup>2</sup>

Fission neutrons from a  $^{252}\text{Cf}$  source were used to calibrate the detectors’ response to nuclear recoils. Neutron calibrations were conducted just before

---

<sup>2</sup>The electron-recoil bands are not directly applicable to the WIMP-search signal region described in this chapter. However, some analysis cuts are naturally a function of  $Q$  (*e.g.*, the fiducial-volume cut). When estimating the energy dependence of such a cut’s efficiency, it is sometimes useful to define a pure sample of electron recoils. In these instances I used data cuts based on the 3V and 6V electron-recoil bands calculated in Don Driscoll’s Ph.D. thesis [696] to select electron recoils.

## CHAPTER 5. A LOW-THRESHOLD ANALYSIS

and after the WIMP searches for each bias voltage. The ionization-yield distributions obtained from these data were used to calculate a nuclear-recoil band for each detector (see Section 5.3.6 below), helping to define the WIMP-search signal region. Additionally, the  $^{252}\text{Cf}$  data series were invaluable for estimating nuclear-recoil detection efficiencies. Unfortunately, to avoid activating the Ge detectors, the neutron data sets are necessarily small. Low nuclear-recoil statistics result in significant uncertainty in the determination of a number of quantities, the foremost of which is the nuclear-recoil energy scale. Nevertheless,  $^{252}\text{Cf}$  calibrations—particularly those conducted between the 3V and 6V WIMP searches—were (indirectly) responsible for a significant fraction of the events observed in the low-energy signal region. These issues are discussed further in the next chapter.

### Empty Events

Empty (or zero-energy) events can be considered a third type of calibration data (in addition to gamma-ray and neutron calibrations) and, as alluded to above, are particularly important for a low-threshold analysis. Empty events are needed primarily for two reasons: to characterize the distribution of subthreshold noise pulses, and for estimates of near-threshold detection efficiencies. The former is required to confidently exclude noise events from the WIMP-search signal region (via a software threshold), and to properly include near-threshold finite-resolution effects—which are dominated by the zero-energy resolution—in the calculation of WIMP-nucleon exclusion limits (as discussed toward the end of Section 3.3.2). The importance of the latter is self-evident; the efficiency for detecting WIMPs must be understood at all energies considered, and is particularly critical near the detector thresholds with regards to low-mass WIMP sensitivity. That empty (or nearly empty) events are needed for this is a consequence of the low thresholds considered here.

Although the RTF-board thresholds were tuned to occasionally trigger on electronic-noise fluctuations, the rate was far too low to characterize near-

### 5.1. RUN 21

Table 5.1: Summary of the various data samples acquired during Run 21b and Run 21c. The numbers of random triggers and viable (non-random) events are listed. The majority of the  $^{137}\text{Cs}$  data was used to test detector neutralization for which the choice of bias voltage is unimportant, hence the large disparity between the numbers of 3V and 6V  $^{137}\text{Cs}$  events. The reason the random-trigger percentages are not exactly 5% is due to the initial 200 random triggers recorded at the beginning of each data series.

Description	Total Viable	Randoms	Comments
3V WIMP search	4,476,533	253,911 (5.4%)	66.12 live days
6V WIMP search	3,705,012	209,120 (5.3%)	51.66 live days
$^{60}\text{Co}$ gammas	3,203,708	172,034 (5.1%)	51% 3V and 49% 6V
$^{252}\text{Cf}$ neutrons	464,094	25,694 (5.3%)	31% 3V and 69% 6V
$^{137}\text{Cs}$ gammas	520,056	32,497 (5.9%)	9% 3V and 91% 6V

threshold phenomena. Furthermore, such events are naturally biased away from 0 keV toward positive energies, making them unsuitable for the study of zero-energy resolution. This is the reason random triggers were interspersed throughout the WIMP search—to provide a periodic sampling of the electronic-noise environment without the trigger bias (“nontriggered data”). The numbers of random triggers and viable (non-random) events obtained for each category of data are summarized in Table 5.1. The 3V and 6V WIMP searches provided the primary samples of randomly triggered data. The rate of true particle interactions during WIMP searches is so low that randomly triggered events usually consisted of only noise fluctuations, with a low probability of measuring a nonzero energy in any of the detectors’ sensors. This makes randomly triggered events particularly useful for determining daily noise levels—their primary use in the analysis described below—but inadequate for efficiency estimates since so few have energies near the detector thresholds.

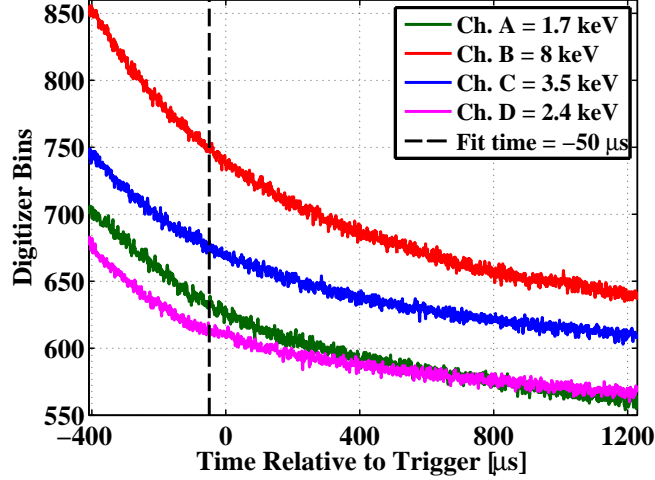


Figure 5.1: Example phonon noise traces (Z5 3V data) from an aberrant class of randomly triggered events. The full-length digitized traces for all four phonon channels exhibit large negative slopes that confuse the optimal-filter fitting routine, resulting in a gross overestimate of each trace’s energy. The vertical, dashed line indicates the  $-50\ \mu\text{s}$  limit of the time window within which the primary fitting routine searched for pulses. Since the maximum pulse height within this window is at the negative limit, all the fits for this type of event have  $-50\ \mu\text{s}$  start times.

The randomly triggered data were also plagued by a low rate of unrepresentative noise traces, apparently caused by a timing artifact associated with the software implementation of the random trigger. As demonstrated for a Ge ZIP in Figure 5.1, the individual phonon traces for these aberrant events exhibit large negative slopes. The traces appear to be the falling edges of phonon pulses resulting from a true particle interaction that occurred earlier in time. Because they partially resemble the pulse shape described by the optimal-filter pulse template, they were mistakenly assigned relatively large energies. The primary fitting routine did not look for pulses occurring more than 50 (25)  $\mu\text{s}$  prior to the trigger time for Ge (Si) detectors. As a result, it identified false peaks at  $-50$  ( $-25$ )  $\mu\text{s}$  for this class of events in the Ge (Si) detectors. A total phonon energy of 15.6 keV was returned by the optimal filter for the traces

### 5.1. RUN 21

in Figure 5.1. True phonon pulses with such a large total energy would be easily visible by eye in the time domain, but there are clearly no rising edges in this case. This population of noise traces contributes non-Gaussian tails to each detector’s underlying electronic-noise distribution (when constructed from randomly triggered events). If these distributions are scaled to an exposure equivalent to the WIMP-search data and subjected to the WIMP-search cuts (described below), the resulting rate of noise pulses in the signal region is overestimated by approximately four orders of magnitude. Consequently, the randomly triggered data were not used to measure the detectors’ zero-energy resolutions.

Zero-energy resolution and threshold efficiency were studied with an alternate sample of data that naturally avoids the trigger bias, but is acquired during normal, triggered readout. Although most WIMP-search events were characterized by an energy deposition in a single ZIP detector, traces for all six detectors were recorded and analyzed. I refer to the data from the detectors that did not cause the event trigger as “other-detector triggers.” It turns out that these data are a rich source of information regarding each detector’s subthreshold noise distribution. Most importantly, they constitute a high-statistics sample that is both fair and representative. There are two types of other-detector triggers: those for which the main event trigger was caused by a detector either directly above or below, and those for which the triggering detector was not an adjacent detector. Non-Gaussian tails are avoided when the latter are used to construct a detector’s subthreshold noise distribution, making nonadjacent other-detector triggers ideally suited for measuring zero-energy resolution. Adjacent other-detector triggers sometimes contained small energy depositions due to true multiple-detector interactions. Because this class of events includes a sampling of phonon energies up to and exceeding the detector (hardware and software) thresholds, it is useful for estimating near-threshold efficiencies. Some of these events exhibit energies in excess of their detector’s RTF-board (low-level phonon) discriminator threshold, and are not truly nontriggered data. These are events in which a delayed trigger occurred during an event readout instigated by another detector. Records of these de-

layed triggers are available in the ZIP detectors’ logical-pulse-based post-trigger histories, and are the basis of the hardware trigger efficiency estimates described in the next section.

## 5.2 Thresholds

WIMP candidates were required to pass hardware and software phonon-energy thresholds. As mentioned in the previous chapter, in order to preserve low-energy detection efficiency, no limitation was placed on a candidate event’s ionization energy. The hardware thresholds were defined by the RTF-board discriminator levels, while the software thresholds were derived (roughly) on a daily basis from the randomly triggered data. In this section I present estimates of the detection efficiencies associated with these thresholds, as well as a description of the derivation and stability of the software thresholds. Note that threshold efficiencies are functions of phonon energy alone. Consequently, they are first calculated as a function of total phonon energy, and then converted to  $Y_{\text{NR}}$ -corrected recoil energy via the average ionization yields (for nuclear recoils) measured in Section 5.3.6. This is an important step as it casts the efficiencies in terms of an equivalent recoil energy—for use in the calculation of WIMP-nucleon exclusion limits—without artificially introducing noise from the ionization channel. Figures demonstrating each type of estimate for a representative ZIP (and bias voltage) are shown here, with the corresponding figures for the other detectors (and bias voltage) given in Appendix B.

### 5.2.1 Hardware Trigger

The hardware trigger is naturally a function of the triggering phonon energy. Other-detector triggers from WIMP-search data are used to measure the associated detection efficiencies as a function of reconstructed phonon energy. As explained in Section 4.3.5, the triggering and reconstructed phonon energies are slightly different. Consequently, the hardware threshold efficiencies as a function of the latter are not simple step functions. The first step in the

## 5.2. THRESHOLDS

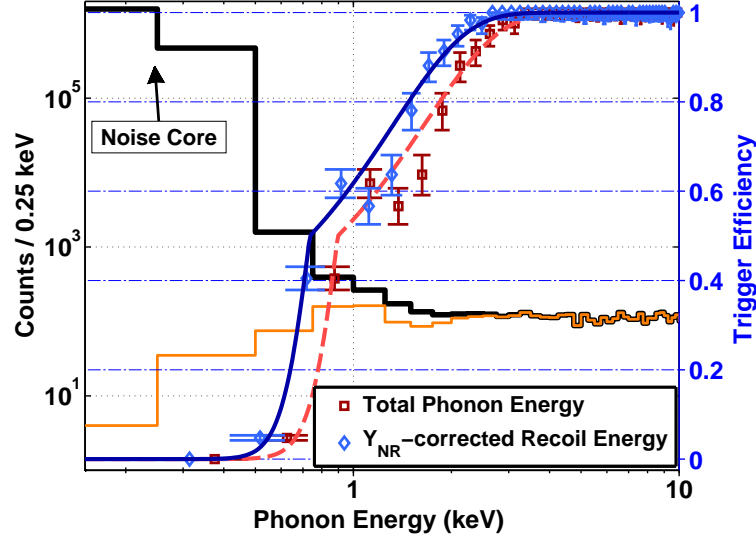


Figure 5.2: Estimate of the hardware trigger efficiency for a representative Ge ZIP (Z2 3V). The distribution of reconstructed phonon energies for successful triggers (light/orange, thin histogram) is divided by the distribution of all reconstructed energies (dark/black, thick histogram), yielding the trigger efficiency as a function of total phonon energy (red squares with  $1\sigma$  error bars) and, following a drift-heat subtraction, as a function of  $Y_{\text{NR}}$ -corrected recoil energy (blue diamonds with error bars). The functional form given in Equation 5.2.1 is fit to the data—as a function of total phonon energy (red dashed line) and  $Y_{\text{NR}}$ -corrected  $E_R$  (blue solid line)—providing a parameterization that accurately reproduces the trigger efficiency at arbitrary energy. For this detector, the hardware trigger is 100% efficient above  $\sim 3$  keV, and the hardware threshold (at 50% efficiency) is 0.74 keV (in terms of  $Y_{\text{NR}}$ -corrected  $E_R$ ).

efficiency estimate for a given detector involves evaluating the reconstructed phonon energy (via optimal-filter fitting) in the  $50\mu\text{s}$  following the time of each other-detector trigger. The majority of these energies follow a Gaussian distribution centered at zero that is consistent with electronic noise (“noise core”). However, nonzero energies are occasionally reconstructed and form an approximately uniform distribution that extends from the noise core to higher energies. Reconstructed energies for which there was an associated hit in the



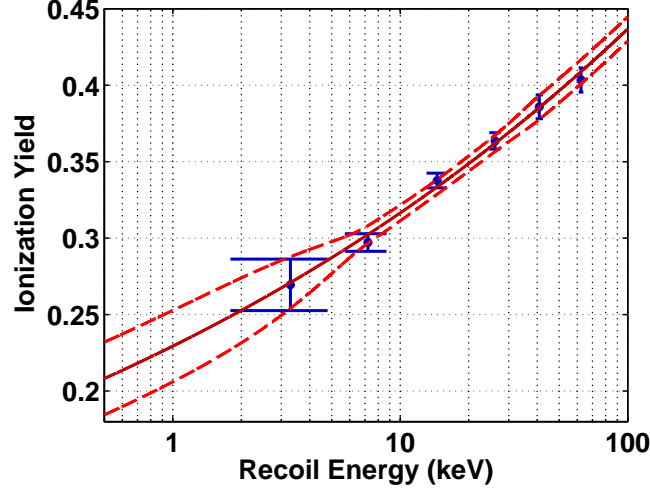


Figure 5.3: The nuclear-recoil band centroid (solid line) and associated  $1\sigma$  error envelope (dashed lines) for a representative Ge ZIP (Z2 3V), derived from a power-law fit (discussed further in Section 5.3.6) to ionization-yield averages (error bars) for neutron-calibration data.

ZIP-detector post-trigger histories form a subset whose distribution is indicative of successful triggers, while the full set of reconstructed energies includes all triggers (successful and unsuccessful). The hardware trigger efficiency is obtained by dividing the former by the latter. This is demonstrated for a representative Ge ZIP in Figure 5.2, where the efficiency is plotted in terms of total reconstructed phonon energy (red error bars) and  $Y_{\text{NR}}$ -corrected recoil energy (blue error bars). The latter is obtained by subtracting the nuclear-recoil drift heat corresponding to the average ionization yield in Figure 5.3 (the nuclear-recoil band “centroid”).

The trigger efficiencies for most detectors (and bias voltages) are well fit by a split-width error function:

$$\begin{aligned} \text{Efficiency}(E \geq A_3) &= \frac{1}{2} [1 - \text{erf}(A_1 (A_3 - E))] \\ \text{Efficiency}(E < A_3) &= \frac{1}{2} [1 - \text{erf}(A_2 (A_3 - E))], \end{aligned} \quad (5.2.1)$$

where  $A_{1-3}$  are free parameters, and  $E$  represents either total phonon energy

## 5.2. THRESHOLDS

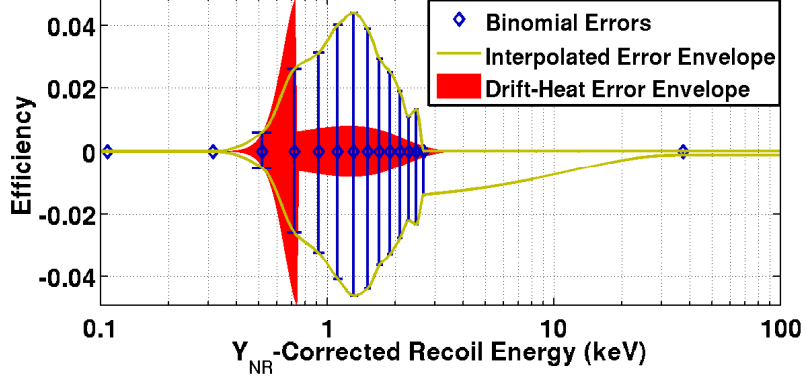


Figure 5.4: Statistical uncertainties associated with the trigger-efficiency estimate (as a function of  $Y_{\text{NR}}$ -corrected  $E_R$ ) shown in Figure 5.2 (for Z2 3V data). A  $1\sigma$  error envelope (solid lines) valid at arbitrary energy is interpolated from the efficiency estimate’s binomial errors (error bars). The uncertainty associated with the drift-heat subtraction (shaded region) is shown as well. The vertical scale is in terms of absolute efficiency relative to the best-fit trigger efficiency (blue solid line in Figure 5.2).

or  $Y_{\text{NR}}$ -corrected recoil energy. To accurately represent the Z4 6V and Z5 (3V and 6V) trigger efficiencies, the following multiplicative factor was added to both lines in the above functional form:

$$\frac{1}{2} [1 - \text{erf}(A_4 (A_5 - E))], \quad (5.2.2)$$

where  $A_4$  and  $A_5$  are additional free parameters. The resulting fits are shown in Figure 5.2 and Appendix B. The reason for the split-width behavior is likely because the triggering phonon energy is less than or equal to the reconstructed phonon energy, while the less common, more complicated behavior is probably related to RTF-board instability or time-varying electronic noise. In any case, the exact reasons the trigger efficiencies are well described by these particular parameterizations is largely unimportant; what matters most is that they accurately follow the data.

The statistical errors associated with the trigger-efficiency estimates (as a function of  $Y_{\text{NR}}$ -corrected recoil energy) are composed of two parts. The first

is due to the binomial confidence intervals derived from comparing the numbers of successful and total triggers in each phonon-energy bin. These are the error bars shown in Figure 5.2. As demonstrated in Figure 5.4, I interpolated along the upper and lower limits of these binomial errors to derive a  $1\sigma$  trigger-efficiency error envelope. The second part of the error estimate is related to uncertainty in the determination of the nuclear-recoil band centroid (dashed lines in Figure 5.3), which causes uncertainty in the drift-heat subtraction used to convert total phonon energy to  $Y_{\text{NR}}$ -corrected recoil energy. The corresponding trigger-efficiency uncertainty is also shown in Figure 5.4. The total statistical error associated with the trigger-efficiency estimate is obtained from the quadrature sum of the binomial and drift-heat error envelopes.

### 5.2.2 Software Threshold

A reconstructed phonon-energy threshold (“software threshold”) was used to reject events caused by electronic-noise fluctuations. Since electronic noise in the phonon measurement is expected to exhibit a Gaussian behavior about 0 keV (see, *e.g.*, the discussion at the end of Section 4.2.2), software thresholds were derived from Gaussian fits to the phonon-energy noise cores. By requiring WIMP candidates to exceed a threshold of  $6\sigma$  above the noise-core mean, the probability of a noise-fluctuation event in a detector’s signal region is effectively zero.

Traditionally, CDMS analyses have calculated noise thresholds from the average behavior of the randomly triggered data over the course of an entire run. This is fine provided the noise levels do not deviate significantly from one data series to the next, or if a relatively large recoil-energy threshold is applied (as is common). For example, with a  $6\sigma$  noise threshold of  $\sim 1$  keV, a noise fluctuation would have to exceed  $\sim 60\sigma$  to enter the signal region of an analysis with a 10 keV recoil-energy threshold. In the analysis presented here, an arbitrary recoil-energy threshold is not used. Instead, the hardware trigger and software noise thresholds are allowed to define the lowest energies considered.

## 5.2. THRESHOLDS

Consequently, particular care must be taken to prevent temporary periods of elevated noise from contaminating the low-energy signal region. Under nominal operating conditions,  $>6\sigma$  noise fluctuations are highly improbable ( $\sim 1$  in  $10^9$ ). However, it is possible for one or more of a detector's phonon channels to temporarily become noisier due to SQUID instability. There is an art to tuning the CDMS SQUIDs for optimal noise performance. Sometimes the SQUIDs appeared stable, but later required additional tuning. Each phonon channel's noise performance was checked on a daily basis (in terms of  $\text{pA}/\sqrt{\text{Hz}}$  versus frequency) to monitor SQUID stability. Nevertheless, there were a few Run 21 data series during which a few detectors (Z1 and Z4) exhibited truly elevated noise levels. Because these periods were short, the *average* noise-core widths were not greatly affected. The corresponding average software thresholds are thus unable to prevent misidentification of noise events as WIMP candidates for the noisiest data series.

There are two approaches one can take to prevent such noisy periods from contributing unwanted events to the WIMP-search signal region. One is to simply exclude the offending data series. Although this results in a reduction of live time, it is a reasonable strategy. Alternatively, daily-acquired noise traces can be used to derive time-dependent software thresholds, effectively increasing a detector's  $6\sigma$  level during any noisy periods. This approach makes better use of the available (noise-free) low-energy detection efficiency, and is the method I adopt here. The randomly triggered data for each data series were fit independently, and a series-dependent  $6\sigma$  software threshold was calculated for each detector from the resulting collection of noise-core means and widths.<sup>3</sup> The resulting thresholds are plotted as a function of time for the Z4 3V data in Figure 5.5. During the 3V WIMP search, this detector experienced one week of elevated noise due to SQUID-related problems. Only Z1's software threshold was less stable, varying by as much as 50%. Most of the other detectors'  $6\sigma$  levels were extremely stable, with the software thresholds for the Z3 3V and

---

<sup>3</sup>The non-Gaussian tails discussed in Section 5.1.1 were mostly avoided by constraining the noise-core fits to the distributions' central regions. In any case, inclusion of non-Gaussian behavior in the fits would result in slightly larger (and therefore conservative) noise thresholds.

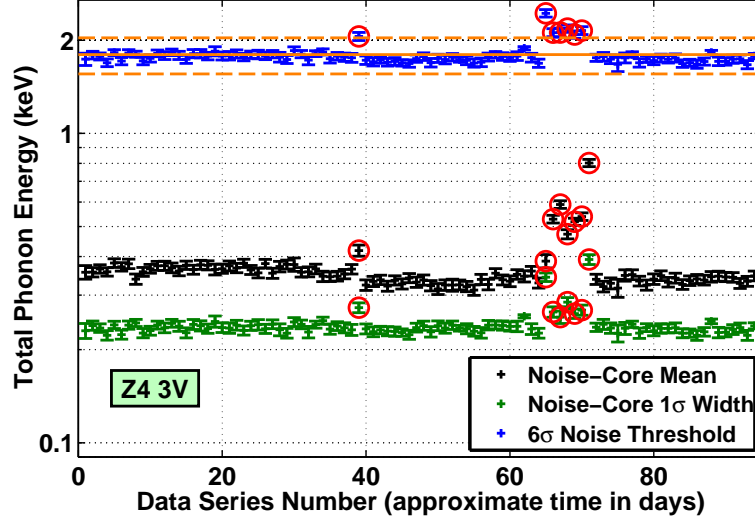


Figure 5.5: Time dependence of the phonon-energy  $6\sigma$  software threshold used to reject events due to noise fluctuations (for the Z4 3V data), compared to the average threshold (solid line). For each data series, the noise-core mean and  $1\sigma$  width from which the software threshold is calculated are shown as well. Data series characterized by  $6\sigma$  thresholds greater (less) than the average  $7\sigma$  ( $5\sigma$ ) threshold (dashed lines) are highlighted by red circles, indicating periods of elevated (reduced) phonon noise.

Z2 (3V and 6V) data varying by less than 5% from day to day (consistent with expected statistical fluctuations). Figures for the other detectors are provided in Appendix B.

Note that the noise-core mean in Figure 5.5 is significantly greater than 0 keV. The main fitting routine was programmed to identify phonon-trace pulses by looking for the maximum positive deviation in a  $-50$  to  $200$  ( $-25$  to  $100$ )  $\mu\text{s}$  window (relative to the trigger time) for Ge (Si) detectors. This caused a slight bias toward positive energies, yielding noise-core means of  $\sim 0.2$ – $0.5$  keV, depending on detector. An alternate algorithm forced fits at the trigger time, resulting in a “zero-delay” phonon-energy estimator that does not suffer from this bias. The reconstructed phonon energy returned by the main algorithm was used to define the software thresholds because this fitting method

## 5.2. THRESHOLDS

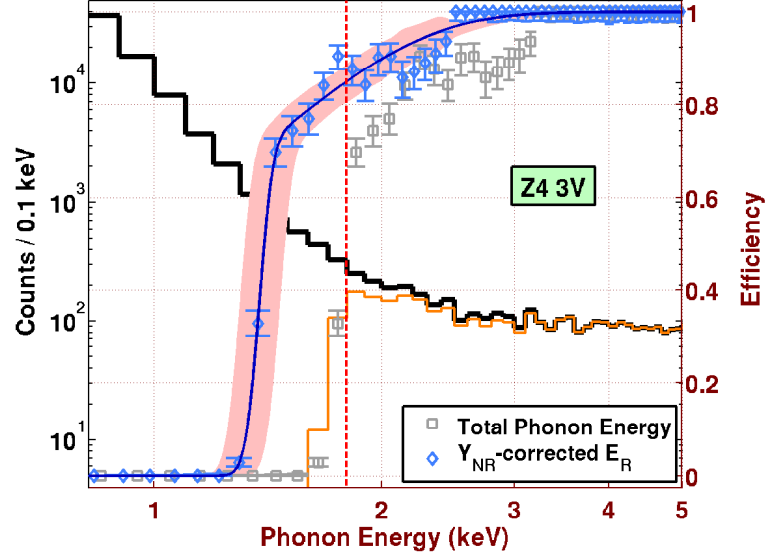


Figure 5.6: An estimate of the average detection efficiency associated with the  $6\sigma$  software threshold for the Z4 3V data is obtained as a function of total phonon energy (gray squares with  $1\sigma$  error bars) by dividing the distribution of energies for events passing the software threshold (light/orange, thin histogram) by the distribution of all reconstructed energies (dark/black, thick histogram). The average software threshold in terms of total phonon energy (red dashed line) is indicated as well. Following a drift-heat subtraction, the efficiency as a function of  $Y_{\text{NR}}$ -corrected  $E_R$  (light blue diamonds with error bars) is fit by a 5-parameter functional form (blue solid line). The total  $1\sigma$  statistical uncertainty associated with the efficiency estimate (shaded region) includes uncertainty related to the drift-heat subtraction as well as binomial statistics.

more closely resembles how the hardware trigger functions. The zero-delay phonon energies are ideally suited to measuring zero-energy resolution, and are used in Section 5.4.

The efficiencies associated with the  $6\sigma$  software thresholds were estimated using the same other-detector triggered data used to measure the hardware trigger efficiencies. The method is virtually identical, with the distribution of energies corresponding to successful hardware triggers replaced by the dis-

## CHAPTER 5. A LOW-THRESHOLD ANALYSIS

tribution of energies exceeding the series-dependent software threshold. As with the hardware-efficiency estimates, the software threshold efficiencies were first calculated as a function of total phonon energy and then converted to  $Y_{\text{NR}}$ -corrected recoil energy. Unlike the hardware trigger, the software thresholds are inherently a function of reconstructed phonon energy. Their efficiencies are therefore expected to closely resemble step functions. Time variations in the software thresholds' Gaussian fit parameters caused the transition from 0 to 100% efficiency to have a small nonzero width for all but one of the viable low-threshold detectors (see Section 5.2.3). Consequently, the data for most detectors are well fit by a simple two-parameter error function:

$$\text{Efficiency}(E) = \frac{1}{2} [1 - \text{erf}(A_1 (A_2 - E))], \quad (5.2.3)$$

where  $A_1$  and  $A_2$  are free parameters and  $E$  represents  $Y_{\text{NR}}$ -corrected  $E_R$ . Among the viable low-threshold detectors, only Z4's 3V data exhibited significant time variations in its software threshold, requiring the more complicated 5-parameter functional form described by Equations 5.2.1 and 5.2.2 to accurately reproduce its efficiency. Binomial and drift-heat error envelopes were estimated in the same manner as for the hardware efficiencies. The total statistical uncertainty associated with the software threshold efficiency—quadrature sum of the binomial and drift-heat error envelopes—is included in each software-efficiency figure (instead of as individual error envelopes in a separate figure). This is demonstrated for the more complicated behavior exhibited by the Z4 3V data in Figure 5.6, and for the other detectors in Appendix B.

The hardware and software thresholds (at 50% efficiency) and their  $1\sigma$  errors are compared in Table 5.2. For all but the Z4 3V data, the software thresholds rise to 100% efficiency more quickly and at lower energies than the hardware efficiencies. Consequently, the combined (hardware and software) threshold efficiency for each detector differs from its hardware efficiency only for very low energies. Except for the Z4 3V data, the combined thresholds (at 50% efficiency) are equal to the hardware thresholds listed in Table 5.2. The combined threshold (at 50% efficiency) for the Z4 3V data is  $1.68^{+0.09}_{-0.08}$  keV,

## 5.2. THRESHOLDS

Table 5.2: Summary of the phonon-energy thresholds (at 50% efficiency) with  $1\sigma$  statistical uncertainties in terms of  $Y_{\text{NR}}$ -corrected recoil energy. For each viable low-threshold detector the hardware and software thresholds are listed for both bias voltages. Except for the Z4 3V data, the combined thresholds are equal to the hardware thresholds listed here. The combined threshold for the Z4 3V data is  $1.68^{+0.09}_{-0.08}$  keV.

Detector	Hardware (keV)		Software (keV)	
	3V	6V	3V	6V
Z2	$0.74^{+0.07}_{-0.02}$	$0.67^{+0.09}_{-0.02}$	$0.63 \pm 0.01$	$0.58 \pm 0.02$
Z3	$1.13^{+0.07}_{-0.04}$	$1.12^{+0.07}_{-0.05}$	$0.82 \pm 0.02$	$0.72 \pm 0.03$
Z4	$1.52^{+0.11}_{-0.09}$	$1.48^{+0.11}_{-0.08}$	$1.40^{+0.09}_{-0.08}$	$1.31^{+0.08}_{-0.07}$
Z5	$1.00^{+0.05}_{-0.04}$	$0.91^{+0.06}_{-0.05}$	$0.73 \pm 0.02$	$0.63^{+0.04}_{-0.03}$
Z6	$1.32^{+0.07}_{-0.05}$	$1.34^{+0.08}_{-0.05}$	$1.20 \pm 0.05$	$1.16 \pm 0.04$

and is the largest among the accepted low-threshold detectors. The energy dependence of the combined threshold efficiencies and their statistical errors are discussed further in Section 5.5.

### 5.2.3 Viable Low-Threshold Detectors

Z1 was rejected as a low-threshold detector due to significantly larger hardware and software thresholds relative to the other Ge detectors. Recall from Section 4.2.5 that Z1 suffered from a particularly strong dependence of phonon pulse height on  $xy$  position because its tungsten  $T_c$  gradient was untreated by ion implantation. The gamma-calibration position correction was unable to fully compensate for the resulting position- and energy-dependent response function. Z1's position-corrected energy resolution and associated  $6\sigma$  noise thresholds are 3 to 4 times larger than those of the other Ge detectors. Addi-



tionally, to prevent excessive noise triggers, Z1's RTF-board (low-level phonon) discriminator threshold was set 2 to 3 times higher than the other Ge detectors' discriminator thresholds. Consequently, this detector's low-energy detection efficiency was seriously diminished, and its ionization yield-based discrimination breaks down at a higher energy. Considering that low-mass WIMP sensitivity is critically dependent on a detector's detection threshold, data from Z1 would contribute very little to the reach of the analysis described here, while allowing a disproportionate number of background events to leak into the signal region (due to degraded  $Y$ -based discrimination at low energy). Z1 was therefore only used to veto multiply scattering events. Z2, Z3 and Z5 (Z4 and Z6) were selected as viable low-threshold Ge (Si) detectors based on their superior low-energy phonon response, both in terms of resolution and threshold. Data from these detectors are the focus of the remainder of this thesis.

### 5.3 Analysis Cuts

WIMP candidates were required to pass a set of analysis cuts intended to select events from the highest-quality, veto-anticoincident data that are consistent with being single-scatter nuclear recoils occurring within the fiducial volume of one of the ZIP detectors. Together with the above thresholds and a recoil-energy upper limit of 100 keV, these cuts define the WIMP-search signal region of this analysis. In this section I describe how the cuts were formed, and present estimates of their detection efficiencies. Unlike the above thresholds, several of the analysis cuts depend either partially or exclusively on the event-by-event ionization measurement, while others exhibit no direct dependence on either  $Q$  or  $P_{\text{total}}$ . The detection efficiencies of the former are most naturally expressed in terms of  $Q$ -corrected recoil energy, while for the latter the issue of energy dependence is trivial. Where appropriate, figures demonstrating detector-dependent cut definitions and efficiency estimates are shown here for a representative detector and bias voltage, with the corresponding figures for the other detectors and bias-voltage runs provided in Appendix C.

### 5.3. ANALYSIS CUTS

#### 5.3.1 Data Quality

Data-quality cuts were used to reject events exhibiting excessive electronic noise in one or more of a ZIP detector’s ionization or phonon channels, events with multiple pulses occurring within the same event (“pileup”), and events for which any part of the data record was incomplete or inconsistent. The former two were suppressed by requiring the variances of the phonon and ionization traces’ prepulse baselines to be within  $5\sigma$  of the average behavior for randomly triggered data, and based on the performance of the optimal-filter fits to the ionization signals. Faulty data records were removed by identifying instances in which some part of the DAQ had lost synchronization. Most of these conditions are inherently independent of the reconstructed  $Q$  and  $P_{\text{total}}$  values, while the optimal-filter criterion might depend on an event’s ionization energy. In this section I review these cuts and their associated efficiencies. The combined detection efficiency for the data-quality cuts is  $99\pm 1\%$  for the entire Run 21 WIMP-search exposure, and exhibits no significant dependence on energy.

#### Prepulse Baselines

Each ZIP-detector trace consisted of a  $400\ \mu\text{s}$ -long pre-trigger and  $\gtrsim 1.2\ \text{ms}$ -long post-trigger record. Pulses resulting from the triggering interaction tended to start within a few tens of  $\mu\text{s}$  of the event trigger. Consequently, the primary pulse-finding routine only considered the last 50 (25)  $\mu\text{s}$  of the pre-trigger part of a Ge (Si) trace when searching for pulses (in addition to 200 (100)  $\mu\text{s}$  of the post-trigger record), leaving over 300  $\mu\text{s}$  of prepulse baseline as an indication of the electronic-noise environment or any ZIP-detector activity occurring immediately before the event trigger. Although highly unlikely due to the low WIMP-search trigger rate, true pileup occurred a few times during Run 21. In these instances, three (or more) interactions transpired close enough in time that the trigger of one (or more) happened during DAQ dead time ( $\gtrsim 250\ \text{ms}$  per event); under the right conditions, pulses sometimes appeared in a detector’s prepulse baselines. The main pulse-finding routine was effectively blind

## CHAPTER 5. A LOW-THRESHOLD ANALYSIS

to this part of the trace.<sup>4</sup> The prepulse baselines provided a means to reject such multiple-pulse event records, which was particularly important for ensuring the quality of high-rate calibration data. Furthermore, events occurring during transient periods of excessive electronic noise were rejected based on their prepulse-baseline behavior.

For each event, *DarkPipe* calculated the average and standard deviation of the first  $320\ \mu\text{s}$  of each ZIP-detector trace, and stored the resulting prepulse-baseline quantities as independent RQs. Using a subset of randomly triggered data from which any (accidentally) non-empty events had been removed,  $5\sigma$  prepulse-baseline thresholds (mean plus 5 standard deviations) were constructed for all 36 ZIP-detector channels. WIMP-search events with prepulse baselines exceeding one (or more) of the  $5\sigma$  thresholds were classified as poor-quality data and were rejected. Except for Z2's phonon channel C, the same  $5\sigma$  levels were used throughout the entirety of Run 21. On the 25<sup>th</sup> of December 2001, the electronic noise on Z2's phonon channel C changed for the worse, requiring its  $5\sigma$  level to be increased by  $\sim 30\%$ .

Two methods were used to estimate the combined efficiency of the prepulse-baseline cuts for each detector and bias-voltage run. The first measured efficiency as the fraction of WIMP-search live time retained by the cuts, while the second used the percentage of events retained. For the viable low-threshold detectors, the results are consistent with a range of 99.8–99.9%. Since prepulse quantities are uncorrelated with the triggering phonon- and ionization-pulse amplitudes, the efficiency of these cuts is independent of energy.

---

<sup>4</sup>Large pre-trigger pulses can take several hundred  $\mu\text{s}$  to return to baseline levels (see, *e.g.*, Figure 5.1), in which case the triggering pulses might be superimposed on top of falling edges. In these instances the main pulse-finding routine is partially sensitive to pileup because the falling edges of the preceding pulses can cause poor optimal-filter fits, and correspondingly large  $\chi^2$  values.

### 5.3. ANALYSIS CUTS

#### The Error Mask

Part of each event record included a bit pattern (or mask) constructed by the DAQ to indicate various error conditions. Among the possible errors, two regularly occurred during Run21 at very low rates. Both were related to a loss of synchronization between the DAQ and one of the digital-history modules. The most common was a result of the muon veto's  $\sim 5$  kHz rate overflowing the veto history buffer while waiting for an event trigger to occur. The time-stamp module in question had a maximum capacity (per event) corresponding to 64,000 veto hits. Consequently, if the time between ZIP triggers was more than  $\sim 13$  s, there was a fair chance that the veto buffer would fill up. When this occurred, the DAQ issued an error code indicating that the veto history for the current event was no longer properly synchronized with ZIP-detector activity. For the  $\lesssim 1$  Hz ZIP-detector trigger rate during the Run21 WIMP searches, this was expected to (and did) occur roughly once in every 10,000 events. A similar but less frequent error indicated a loss of synchronization between the DAQ and the ZIP-detector history buffer.

There were also intermittent periods during Run21 in which one of the muon veto's PMTs caused the overall veto rate to exceed 100 kHz, resulting in a higher than expected rate of veto-synchronization errors. This appeared to be due to abnormally long (in time) and malformed PMT pulses caused by breakdown in one of the older, helium-saturated PMTs. These high-rate episodes had little affect on the 3V WIMP search, but increased the rate of veto-synchronization errors during the 6V WIMP search to 3–4 per 1,000 events.

Any events with a nonzero error-mask value were rejected as WIMP candidates. The efficiency of this cut was estimated using the method applied to the prepulse-baseline cuts, and is consistent with  $>99.9\%$  for the 3V data and  $\sim 99.5\%$  for the 6V data.

### Ionization $\chi^2$

The final data-quality cut is based on the  $\chi^2$  values of the optimal-filter fits to the ionization signals.<sup>5</sup> These provide a statistical measure of the quality (or goodness) of the fits, with abnormally large  $\chi^2$  values indicating poor fits. Poor fits can occur when there is pileup in the fitting routine’s search window, or when traces are uncharacteristically noisy. In both cases, the optimal-filter pulse (and noise) templates inadequately describe the trace shapes, and therefore cannot accurately reproduce the ionization energy. Clearly, such events should be excluded from consideration as WIMP candidates.

Events with abnormally large  $\chi^2$  values were rejected as WIMP candidates based on an energy-dependent cut derived from  $^{60}\text{Co}$  calibration data:

$$\chi^2 \geq C_z Q_{\text{sum}}^2 + 5000, \quad (5.3.1)$$

where  $C_z = 0.3$  (0.125) for Z2 and Z5 (Z4 and Z6) and 0.25 for Z3, and  $Q_{\text{sum}}$  is the total ionization energy ( $Q_{\text{inner}}$  plus  $Q_{\text{outer}}$ ). The cut was designed to accept virtually all events, excluding only exceptionally large  $\chi^2$  values. Figure 5.7 shows the cut in relation to electron recoils from the 3V WIMP-search data for a representative Ge ZIP, as well as an estimate of the cut’s efficiency as a function of ionization energy.<sup>6</sup> The latter is estimated in bins of  $Q_{\text{sum}}$  by comparing the number of events passing the cut to the total number in a given bin. The efficiency was also tested with nuclear recoils from  $^{252}\text{Cf}$  calibrations. For each detector, bias voltage, and data type the efficiency of the  $\chi^2$  cut is independent of energy and consistent with  $>99.8\%$  for events in the fiducial volume (*i.e.*,  $Q_{\text{outer}}$  signals consistent with noise—see Section 5.3.3 below), and only slightly less efficient otherwise.

---

<sup>5</sup>Ideally, a similar cut based on the optimal-filter fits to the phonon signals would have been part of the data-quality cuts. However, there were bugs in the *DarkPipe* code used to calculate the  $\chi^2$  values for the phonon fits, and it was not deemed important enough to justify a lengthy reprocessing of the Run 21 raw data with a corrected algorithm.

<sup>6</sup>Note that because the equivalent plots for the other detectors, bias voltages, and data types look virtually identical to Figure 5.7 (aside from differences in statistics), they are not included in Appendix C.

### 5.3. ANALYSIS CUTS

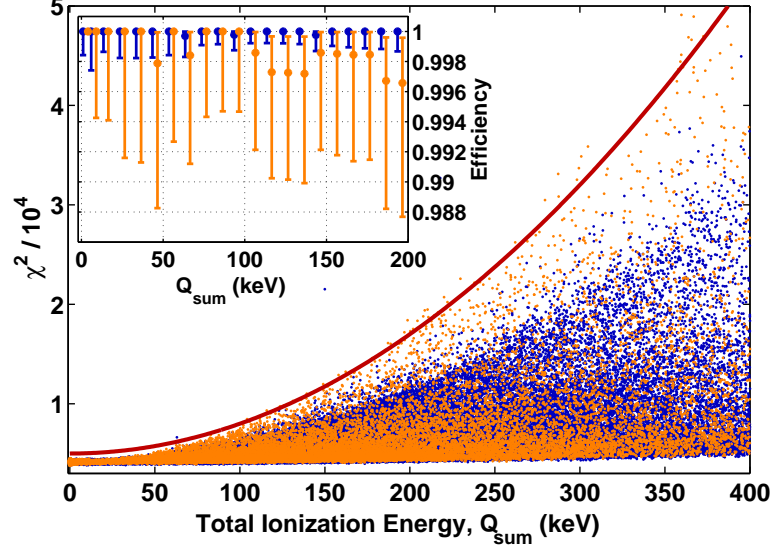


Figure 5.7: The data-quality cut (solid line) used to remove abnormally large  $\chi^2$  values as a function of total ionization energy, compared to electron recoils (dots) from 3V WIMP-search data for a representative Ge ZIP (Z5). Events occurring in the detector’s fiducial volume (dark/blue dots) are distinguished from those occurring under the  $Q$ -outer electrode (light/orange dots). *Inset*: Estimate of the cut’s efficiency in bins of ionization energy. Colors correspond to those in the main figure.

#### 5.3.2 Event Bursts

For the majority of the Run 21 WIMP search, the cuts described above were sufficient to ensure high data quality. Unfortunately, the 6V data were additionally plagued by intermittent periods of high trigger rates (“event bursts”) due to sub-2 keV pulses in Z2. As demonstrated in Figure 5.8, many of these followed cryogenic and detector servicing periods, during which detector temperatures were elevated due to exposure to photons from infrared LEDs (to maintain detector neutralization—see Sections 4.2.1 and 4.3.5). Apparently, the tungsten  $T_c$  was somewhat lower for Z2’s TESs than for the other detectors. In particular, it appears that the residual  $T_c$  gradient for Z2’s tungsten

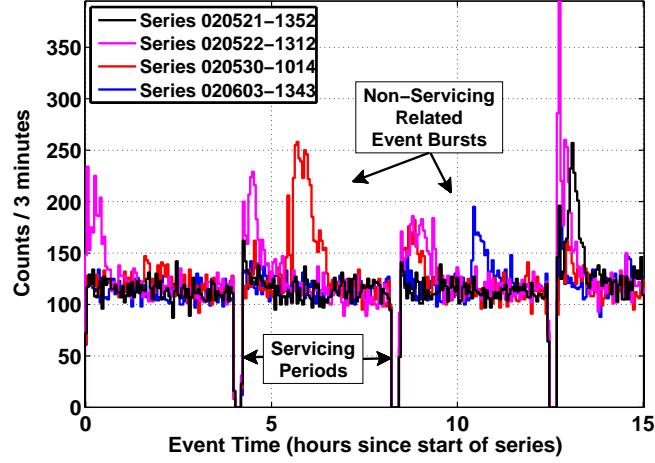


Figure 5.8: Event-rate histograms for a few Z2 6V WIMP-search data series, demonstrating a correlation between high trigger rates and detector servicing periods. Note that a few of the event bursts appear uncorrelated with detector servicing.

was such that the  $T_c$  was lowest toward  $-y$  positions (phonon channels B and C), and highest toward  $+y$  positions (channels A and D). Having a lower  $T_c$  means that the TESs in question would have to be cooled to lower temperatures to again function properly following LED flashing. Evidently, the 15 minute-long servicing periods did not allow enough time for the Z2 phonon sensors to fully recover. Histograms similar to Figure 5.8 were checked for each detector and bias-voltage run; event bursts were observed for only the Z2 6V data.

This correlation between detector temperature and event rate implies that the Z2 event bursts were not due to physical recoils in the detector. Consequently, they were considered periods of poor data quality. Figure 5.8 also shows a few event bursts that did not directly follow an LED flashing period. Isolating the event bursts was therefore more complicated than simply identifying an amount of time following each detector servicing period. One option is to reject the entire Z2 6V WIMP search. However, among the viable low-threshold detectors, Z2 had the lowest thresholds, which partially explains its

### 5.3. ANALYSIS CUTS

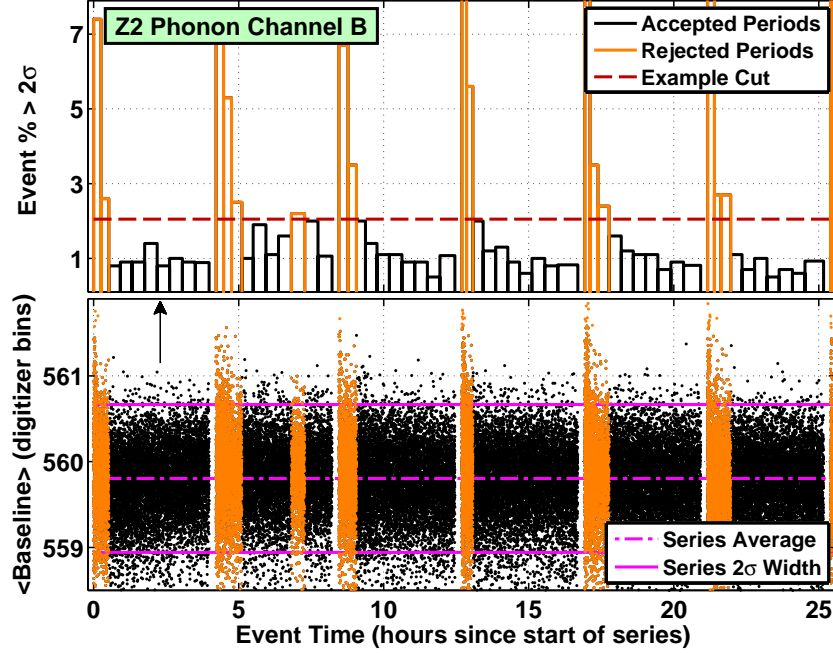


Figure 5.9: Demonstration of the cut used to remove event bursts from the Z2 6V WIMP-search data. *Bottom:* The average prepulse baseline (APB) for phonon channel B is plotted as a function of event time for each event in a data series (dots), with the average and  $2\sigma$  width of the collection of APBs indicated as well. *Top:* The percentage of events for which the APB is greater than the series mean plus  $2\sigma$  width is histogrammed in 1000-event bins. Periods during which the histogram exceeds an optimized cut level are rejected as poor-quality data.

sensitivity to the event-burst phenomenon. Relative to the other detectors, Z2’s WIMP-search exposure is expected to be among the most sensitive to low-mass WIMPs, making it worth while to preserve as much of it as possible.

Upon careful inspection, Z2’s phonon traces were discovered to have slightly elevated prepulse baselines during event-burst episodes. This is demonstrated in Figure 5.9 for channel B, where the average prepulse baseline (APB) is plotted as a function of event time. During event-burst episodes, there was a



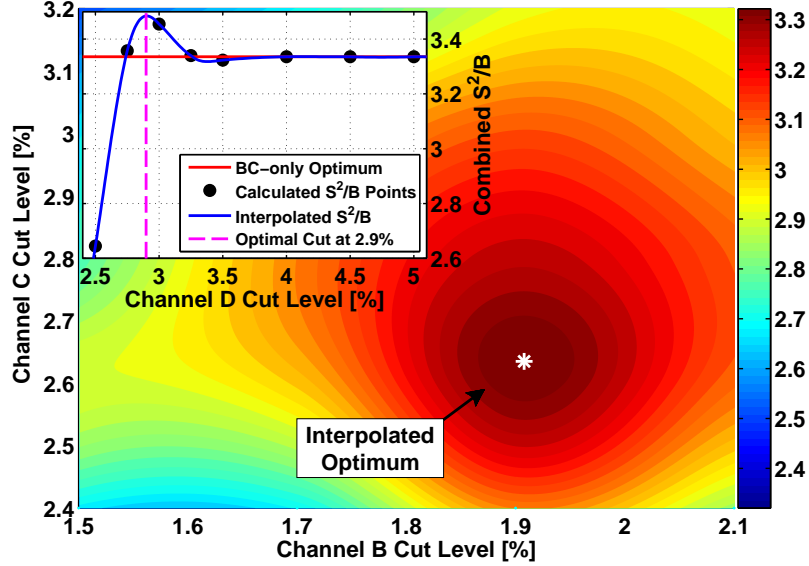


Figure 5.10: The contours represent a simultaneous optimization of the event-burst cut levels for phonon channels B and C, with color indicating the value of the  $S^2/B$  statistic. An optimal combination is found at 1.9% (2.6%) for channel B (C). *Inset:* Several cut levels for channel D are scanned in combination with the channel B and C optimal levels, and a maximum is found that improves the overall performance of the event-burst cut. A similar scan was performed for channel A, but no further improvement was discovered.

larger percentage of events with an elevated APB than during other times. The phonon-baseline standard-deviation RQ for events occurring during these times was largely unaffected, allowing them to slip past the  $5\sigma$  prepulse-baseline cut described above. An event-burst discrimination parameter was derived from the percentage of events for which the APB exceeded a  $2\sigma$  threshold in 1000-event bins (also demonstrated in Figure 5.9). Event-burst periods were identified as times during which this discrimination parameter exceeded some cut level, and were excluded from Z2’s WIMP-search exposure. Z2 was only used to veto multiply scattering events during the rejected times.

The event bursts were not spread equally among Z2’s phonon channels. Consequently, each channel required a different cut level. An  $S^2/B$ -style op-

### 5.3. ANALYSIS CUTS

timization was performed to fairly decide the best cut levels. Although the statistically most proper procedure would be to optimize all four channels simultaneously, such a multi-dimensional optimization is logistically complicated. Instead, since most of the event bursts were characterized by  $xy$  positions in phonon quadrants B and C, a simplified procedure was used in which the cut levels for channels B and C were optimized simultaneously, followed by a search for an improved overall optimum through inclusion of channel D and then A. Signal ( $S$ ) was estimated according to the fraction of live time retained with a given set of cut levels, while background ( $B$ ) was estimated as the fraction of events remaining in a background sample for the same cut levels. The hour of data following each LED flashing period was used as the background sample. Since the event bursts during these times were clearly correlated with elevated detector temperatures, the use of these events should not result in any significant bias against a possible WIMP signal.<sup>7</sup> The resulting  $S^2/B$  statistic is normalized (and dimensionless) such that a value greater than one indicates an advantageous combination of cut levels. The results of the optimization are summarized in Figure 5.10. Optimal cut levels were identified for channels B, C and D, but not channel A. The resulting event-burst cut reduced Z2's 6V WIMP-search exposure from 52 to 20 days, while reducing the corresponding number of events in its signal region by a factor of  $\sim 40$ . Without this cut, the Z2 event bursts would have been the dominant source of background, roughly tripling the Ge WIMP-candidate event rate.

#### 5.3.3 Fiducial Volume

WIMP candidates were required to pass a fiducial-volume cut formed from the ionization signals. This cut rejected events occurring near the detector edges, where charge collection was less efficient, and where there was a higher likelihood of interactions due to the decay of surface contaminants on the

---

<sup>7</sup>In hindsight, events occurring within Z2's 6V WIMP-search signal region should have been excluded from the background sample. However, the full background sample of  $\sim 900,000$  events is much larger than the  $\sim 3,000$  that occur in the signal region. Any bias introduced by this oversight should therefore be very small.

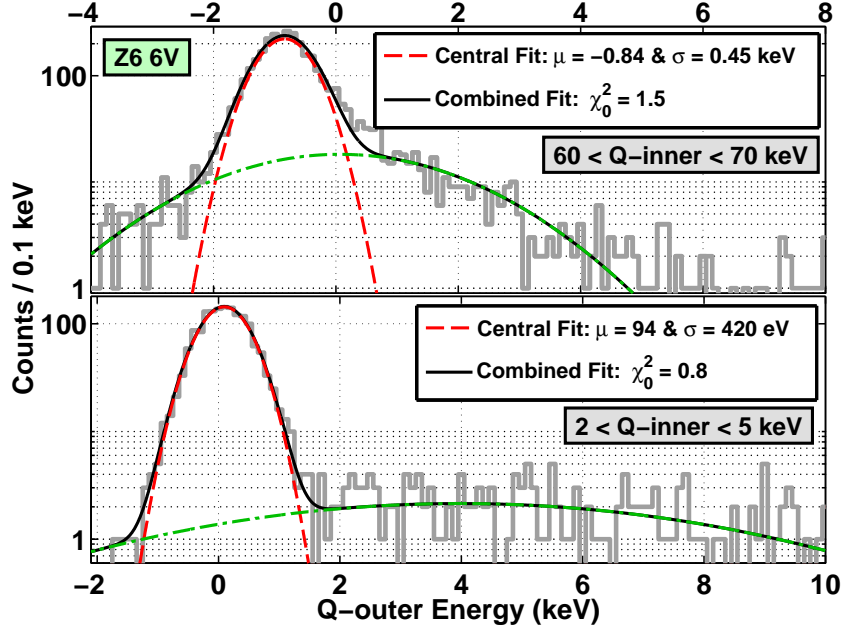


Figure 5.11: Example  $Q$ -outer ( $Q_o$ ) noise-distribution fits for low-energy (bottom) and high-energy (top) bins of  $Q$ -inner ( $Q_i$ ) energy for the Z6 6V data. For each  $Q_i$  bin, two Gaussians are fit simultaneously to the  $Q_o$  noise distribution (histogram): a central fit (dashed line) provides the  $Q_o$  mean and width for constructing the  $Q_o$  noise band (as in Figure 5.12), and a broader fit (dash-dotted line) prevents the low- and high-value tails from biasing the central fit.

detectors' copper housings. The geometry of the ionization electrodes was designed with this purpose in mind. The principle behind the cut is quite simple: the  $Q$ -inner signal estimates an event's ionization energy, while the  $Q$ -outer signal is required to be consistent with electronic noise; events with significant  $Q$ -outer signals most likely occurred directly beneath the  $Q$ -outer electrode and were thus outside the fiducial volume. Unfortunately, deriving such a cut in practice was complicated by cross talk between the two ionization electrodes, which resulted in a small, energy-dependent anticorrelation between the  $Q$ -inner and -outer signals for each detector and bias-voltage run.

### 5.3. ANALYSIS CUTS

The dependence of the  $Q$ -outer noise level on the  $Q$ -inner signal was modeled by constructing a  $Q$ -outer “noise band.”  $Q$ -outer noise events for WIMP-search single scatters (see Section 5.3.4 below) were sorted into 14 bins of  $Q$ -inner energy. In each bin, the mean  $Q$ -outer energy and  $1\sigma$  width were extracted with a Gaussian fit to a histogram of the selected  $Q$ -outer noise events. This is demonstrated for a low- and high-energy bin for a representative detector in Figure 5.11.<sup>8</sup> In most cases, the  $Q$ -outer noise distributions exhibit tails that are not well described by a single Gaussian. The positive-energy tails are likely a mixture of true  $Q$ -outer events (not in the fiducial volume) and events for which the ionization energy was shared between the two electrodes. The negative-energy tails are discussed below. A second Gaussian was fit simultaneously in order to prevent the tails from affecting the fit parameters for the central part of the noise distribution.

The collection of means and widths derived from the (central)  $Q$ -outer noise fits were parameterized as a function of  $Q$ -inner energy. This is demonstrated in Figure 5.12, where the Gaussian means and  $2\sigma$  widths are overlaid onto a scatter plot of  $Q$ -outer versus  $Q$ -inner energy for the Z6 6V WIMP-search data. The means are well fit by a 5-parameter functional form:

$$Q_o = \sqrt{A_1^2 + A_2^2 (Q_i - A_3)^2} + A_4 (Q_i - A_3) + A_5 (Q_i - A_3)^2, \quad (5.3.2)$$

where  $A_{1-5}$  are free parameters, and  $Q_i$  ( $Q_o$ ) is the  $Q$ -inner ( $Q$ -outer) energy. As demonstrated in the inset of Figure 5.12, the widths were fit separately with a 6-parameter functional form:

$$\sigma_{Q_o} = B_5 + B_6 \ln(1 + \sum_{n=1}^4 B_n Q_i^n), \quad (5.3.3)$$

where  $B_{1-6}$  are free parameters and  $\sigma_{Q_o}$  is the  $1\sigma$   $Q$ -outer noise width. Although these functional forms are a bit contrived, they fit quite well down to  $Q_i \simeq 0$  keV. As with the threshold efficiencies, a parameterization that

---

<sup>8</sup>There are a total of 140 Gaussian fits to binned  $Q$ -outer noise distributions for the viable low-threshold detectors, far too many to include here. Instead, example low- and high-energy fits in the style of Figure 5.11 are provided for the other detectors and bias-voltage runs in Appendix C.

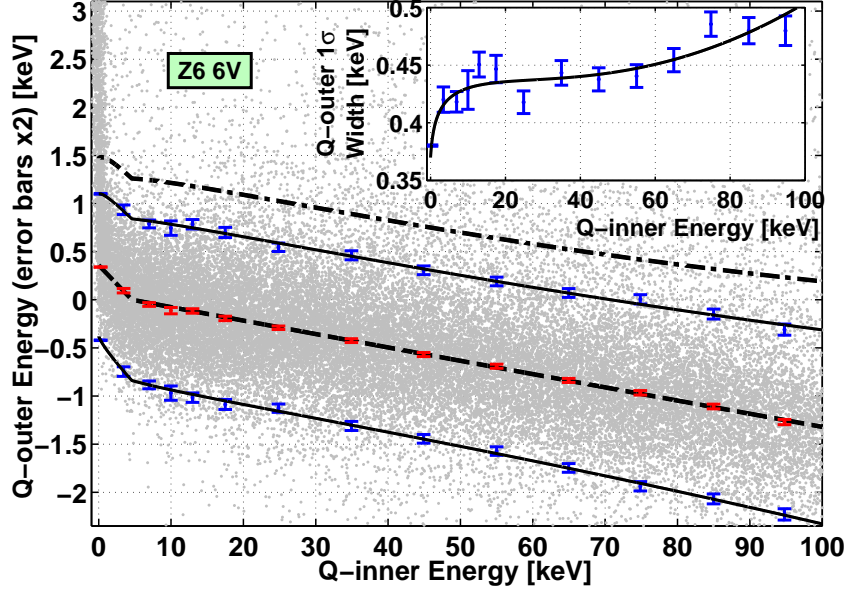


Figure 5.12: The  $Q$ -outer  $2\sigma$  noise band (solid lines) as a function of  $Q$ -inner energy for the Z6 6V data. The means (red error bars) and  $2\sigma$  widths (blue error bars) from the  $Q_o$  noise-distribution fits (as in Figure 5.11) are compared to the individual  $Q_o$  versus  $Q_i$  data points (gray dots) from which they were derived. The average  $Q_o$  noise level was fit using Equation 5.3.2 (dashed line), while the  $Q_o$   $1\sigma$  widths (error bars in inset) were fit using Equation 5.3.3 (solid lines in main figure and inset). The fiducial-volume cut rejected events exceeding the  $+2\sigma$  limit of the  $Q_o$  noise band. The  $3\sigma$  upper limit (dash-dotted line) used to estimate the cut's selection efficiency is shown as well.

is accurate at arbitrary energy is of primary importance. Judging by Figure 5.12 and the others like it in Appendix C, Equations 5.3.2 and 5.3.3 allow for a remarkably accurate description of the  $Q$ -outer noise band, especially considering the kink between 0 and 10 keV that most of the detectors exhibit.

The negative-energy tails in Figure 5.11 are an artifact of a cross-talk correction performed to account for signals induced by the phonon sensors of the

### 5.3. ANALYSIS CUTS

adjacent detector.<sup>9</sup> This correction worked properly for most events. However, close inspection of Figure 4.4 reveals that the phonon-sensor coverage near the detector edges is not complete. Consequently, for events occurring where there was no corresponding phonon-sensor coverage on the bottom surface of the adjacent detector, the cross talk was smaller. The cross-talk correction oversubtracted from the  $Q$ -outer signals for this class of events, pushing them toward negative values inconsistent with expected noise fluctuations. These are otherwise viable events that should be included as part of the fiducial volume. However, many are rejected if the fiducial volume is restricted to include only events within the  $\pm 2\sigma Q_o$  noise band. To avoid the corresponding reduction in efficiency, the  $-2\sigma$  limit of the  $Q_o$  noise band was not used. Events were considered to have occurred within a detector’s fiducial volume provided their ionization signals did not exceed the  $+2\sigma Q_o$  noise-band limit. The expected efficiency for selecting fiducial-volume events with this cut is therefore  $\sim 97.7\%$ . Note that this is the expected “selection efficiency” of the cut, and not the overall detection efficiency associated with the fiducial volume.

The efficiency of the  $+2\sigma Q_o$  noise band for selecting fiducial-volume events was tested using electron recoils from  $^{60}\text{Co}$  calibrations. The data were sorted according to the same bins of  $Q_i$  used above (minus the zero-energy bin), and the efficiency was estimated by comparing the numbers of events below the  $+2\sigma$  and  $+3\sigma Q_o$  noise-band limits in each bin.<sup>10</sup> The resulting efficiency is a function of  $Q_i$ . Nuclear recoils from  $^{252}\text{Cf}$  calibrations were used to convert the efficiencies to  $Q$ -corrected recoil energy. The method depends on the assumption that nuclear and electron recoils have equal efficiencies for same-energy ionization signals. If true, then for each electron-recoil  $Q_i$  bin, an equivalent bin can be constructed from the average (bin center) and standard deviation (bin width) of the  $Q$ -corrected recoil energies for nuclear recoils

---

<sup>9</sup>Recall that (except for Z1) each detector’s ionization electrodes are separated from the phonon sensors on the bottom side of the detector directly above by  $\sim 2$  mm.

<sup>10</sup>The number of events below the  $+3\sigma$  limit was taken as the denominator of the efficiency estimate, and does not include 100% of all events in the fiducial volume for a given  $Q_i$  bin. The efficiencies were multiplied by 0.9987 to account for the missing events.

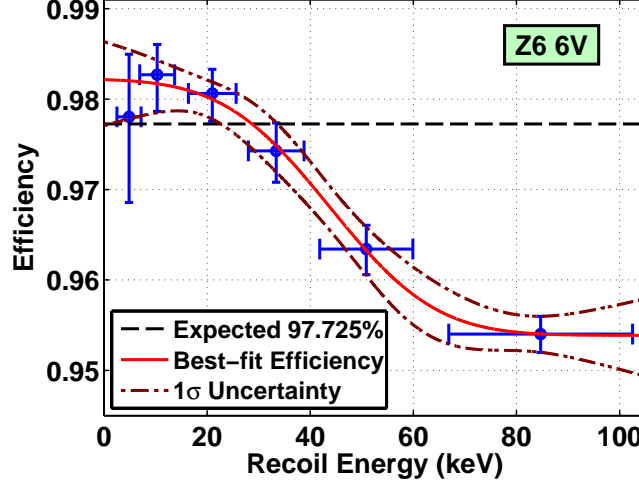


Figure 5.13: Estimate of the Z6 6V fiducial-volume cut’s selection efficiency (error bars) as a function of recoil energy, compared to the expected efficiency. The functional form given in Equation 5.3.4 reproduces the slight dependence on energy, and was also used to derive the error envelope.

whose ionization energies fall within the  $Q_i$  bin. Each recoil-energy bin is then assigned the efficiency of its corresponding electron-recoil  $Q_i$  bin. Since nuclear-recoil ionization energies are quenched relative to electron recoils, only the first six  $Q_i$  bins are needed to span the recoil-energy range of interest for WIMP-nucleus ineteractions ( $E_R < 100$  keV). The resulting selection efficiency’s slight dependence on recoil energy is well described by a 4-parameter error function:

$$\text{Efficiency} = A_1 - A_2 \text{erf}[A_3 (E_R - A_4)], \quad (5.3.4)$$

where  $A_{1-4}$  are free parameters and  $E_R$  is  $Q$ -corrected recoil energy.<sup>11</sup> The fit for the Z6 6V fiducial-volume cut’s selection efficiency is shown in Figure 5.13 with its associated  $1\sigma$  statistical uncertainty. In this case, the error envelope was derived from the fit’s covariance matrix.

<sup>11</sup>The functional form in Equation 5.3.4 was used for all but one detector and bias voltage; a constant-efficiency hypothesis provides an adequate description of the Z5 3V fiducial-volume cut’s selection efficiency (see Figure C.21).

### 5.3. ANALYSIS CUTS

The overall detection efficiency associated with the fiducial volume was estimated by combining the measured selection efficiencies with the physical coverage of the  $Q$ -inner electrode. The average fiducial-volume detection efficiencies agree well with the expected  $\sim 83\%$ , varying between  $\sim 81\%$  (Z4 3V data) and  $\gtrsim 83\%$  (Z2 3V data).<sup>12</sup> Note that for recoil energies  $\lesssim 4$  keV, a nuclear-recoil event's ionization signals (both  $Q_i$  and  $Q_o$ ) are difficult to distinguish from electronic noise. Consequently, the fiducial-volume cut's ability to reject events near the detector edges is expected to gradually degrade for recoils energies less than 4 keV, and the overall fiducial-volume detection efficiency is expected to increase. Unfortunately, this behavior cannot be measured with the ZIP detectors. I made the conservative choice to linearly extrapolate the efficiency for recoil energies less than 4 keV to match the efficiency at 4 keV, resulting in a loss of low-energy efficiency relative to the true (and unmeasurable) efficiency of this analysis.

#### 5.3.4 Single Scatters

WIMP-nucleus interactions are expected to be so localized and infrequent that only events caused by backgrounds will deposit significant energy in more than one detector during a scattering event. To reduce multiply scattering backgrounds, WIMP candidates were therefore required to pass a single-scatter criterion. Because the total phonon signal provides the most sensitive indication of a particle interaction in a ZIP detector, the single-scatter cut is based on the  $6\sigma$  software thresholds described in Table 5.2, Section 5.2.2, and Appendix B. A single scatter is defined as an event for which only one detector had a reconstructed phonon energy exceeding its software threshold. Events with a significant energy deposition in more than one detector were thus rejected. As mentioned previously, the Z1 data—as well as the Z2 data rejected by the event-burst cut—were included when determining an event's single-scatter status.

---

<sup>12</sup>This can be stated in another way that is perhaps more clear. The selection efficiency of the fiducial-volume cut is  $\sim 95$ – $98\%$  (depending on detector), which reduces the fiducial volume from 85% to  $\sim 81$ – $83\%$ .



The efficiency of the single-scatter cut is  $\sim 100\%$  for sufficiently large recoil energies (greater than a few keV). The trigger rate during WIMP searches was so low that the probability of two or more (physically uncorrelated) pulses occurring within one of the 1.6 ms-long event records is negligible ( $\sim 1$  in 1000).<sup>13</sup> For recoil energies near threshold, the cut’s efficiency is given by the software threshold efficiency (as in Figure 5.6).

### 5.3.5 Veto Anticoincidence

Events with activity in the muon veto were removed from consideration with a veto-anticoincidence cut. The nature of this cut—based on the muon-veto digital histories—is described in detail in Section 4.3.3. Basically, any events for which a veto hit occurred too close in time to the event trigger were rejected. Previous CDMS shallow-site WIMP-search analyses used a cut that rejected event triggers occurring less than  $25\text{--}40\,\mu\text{s}$  after the last veto hit (see, *e.g.*, [687] and [694]).  $40\,\mu\text{s}$  is sufficient for recoil energies  $\gtrsim 3.5\,\text{keV}$ , for which the phonon pulses rise past the RTF-board discriminator levels relatively quickly. The smaller pulses associated with lesser energies, however, must rise to nearly their full amplitudes before passing the hardware threshold to cause an event trigger. As demonstrated in Figure 5.14, this results in greater differences between veto- and phonon-pulse arrival times. Veto-coincident events with near-threshold phonon energies exhibit correlations as much as  $80\,\mu\text{s}$  in advance of their event triggers. Consequently, to ensure high muon-tagging efficiency, an extended pre-trigger time window was rejected for low-energy events. For most detectors, a  $70\text{--}80\,\mu\text{s}$ -long period was used for low-energy events, and a more modest  $50\text{--}55\,\mu\text{s}$ -long period otherwise. Z4 was the only detector for which  $50\,\mu\text{s}$  was sufficient to reject even the lowest-energy veto-coincident events. The cut parameters are summarized in Table 5.3.

---

<sup>13</sup>Note that two or more pulses occurring within a single event record is a more general event type than the pileup described in Section 5.3.1. While the data-quality cuts are likely to reject any events with two or more pulses in the traces for a single detector, they are insensitive to the occurrence of two or more pulses occurring in two or more detectors.

### 5.3. ANALYSIS CUTS

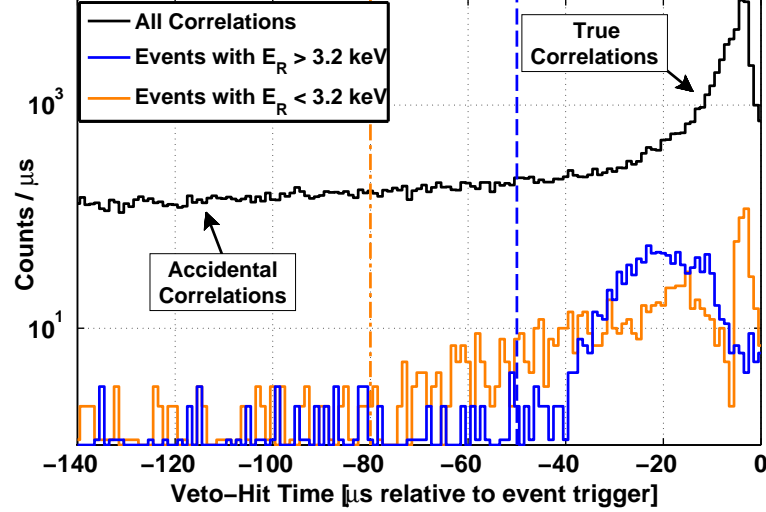


Figure 5.14: Distribution of arrival times for muon-veto hits for Z3 3V WIMP-search data. True correlations between ZIP-detector and muon-veto activity form a peak close to the event-trigger time at  $0 \mu\text{s}$ , while accidental correlations are characterized by an exponential distribution with a slope corresponding to the veto’s overall  $\sim 5 \text{ kHz}$  rate. Low-energy ZIP-detector events exhibit correlations further in advance of the event trigger, requiring a stricter anticoincidence cut (dash-dotted line) than required for events with sufficiently large phonon pulses (dashed line).

As discussed in Section 4.3.3, the dead time associated with the veto cut was significant. This is because the veto’s overall rate of activity during the 3V (6V) WIMP search was  $5047 \pm 7$  ( $5004 \pm 7$ ) Hz. When combined with the parameters listed in Table 5.3, these rates yield a veto-cut detection efficiency that varies from  $\sim 67$  to  $78\%$ , depending on detector, energy, and bias voltage. Application of this cut results in the single largest loss of detection efficiency for the WIMP-search analysis described here.

## CHAPTER 5. A LOW-THRESHOLD ANALYSIS

Table 5.3: Summary of the veto-cut parameters and corresponding detection efficiencies for the viable low-threshold detectors. The  $1\sigma$  statistical uncertainties for the efficiency estimates are 0.04% and 0.03% for the 70–80 and 50–55  $\mu$ s-long cuts, respectively.

Detector	Cut Parameters	3V Efficiency	6V Efficiency
Z2	80 $\mu$ s for $E_R \leq 3.25$ keV	66.78%	67.01%
	55 $\mu$ s for $E_R > 3.25$ keV	75.76%	75.94%
Z3	80 $\mu$ s for $E_R \leq 3.15$ keV	66.78%	67.01%
	50 $\mu$ s for $E_R > 3.15$ keV	77.70%	77.86%
Z4	50 $\mu$ s $\forall E_R$	77.70%	77.86%
Z5	80 $\mu$ s for $E_R \leq 3.0$ keV	66.78%	67.01%
	50 $\mu$ s for $E_R > 3.0$ keV	77.70%	77.86%
Z6	70 $\mu$ s for $E_R \leq 3.75$ keV	70.24%	70.45%
	50 $\mu$ s for $E_R > 3.75$ keV	77.70%	77.86%

### 5.3.6 Nuclear-Recoil Band

Finally, to discriminate against the otherwise overwhelming rate of electron recoils, WIMP candidates were required to have an ionization yield consistent with being a nuclear recoil. As discussed in Section 4.2.3, the ionization yield for nuclear recoils varies with recoil energy, monotonically increasing from zero at 0 keV to  $\sim 0.4$  at 100 keV. Nuclear recoils induced by fission neutrons during the  $^{252}\text{Cf}$  calibrations were used to parameterize each detector’s ionization-yield response as a function of recoil energy. Similar to the fiducial-volume cut, a  $2\sigma$  nuclear-recoil band was constructed, and any events with an ionization yield falling outside their detector’s band were rejected.

To parameterize the nuclear-recoil band’s energy dependence, single scatters from the  $^{252}\text{Cf}$  data were sorted into six bins of recoil energy spanning  $E_R = 2$ –100 keV. In each bin, a maximum-likelihood fit to the distribution of low-value

### 5.3. ANALYSIS CUTS

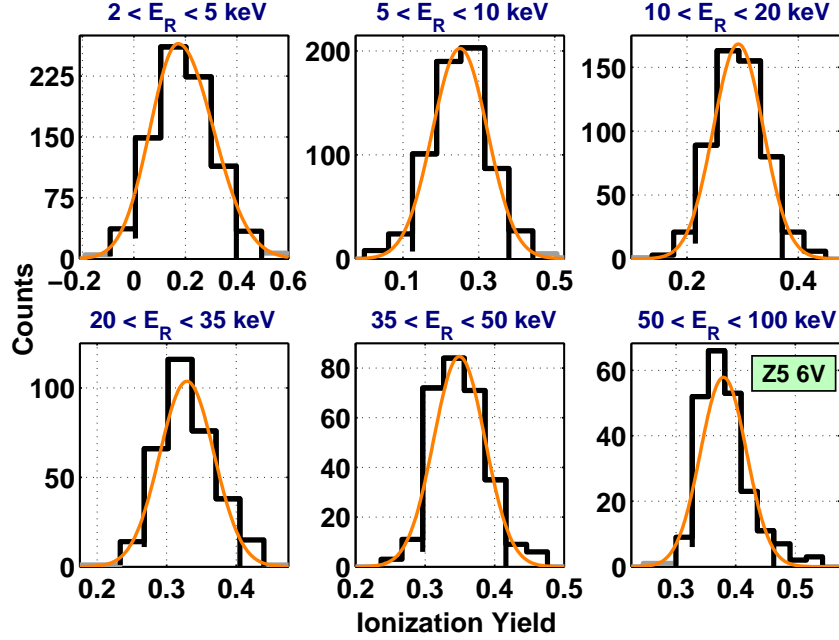


Figure 5.15: Recoil-energy binned, low-value ionization-yield distributions for Z5 6V  $^{252}\text{Cf}$  calibration data, consistent with nuclear recoils induced by neutrons. In each bin, the Gaussian parameters resulting from the fits described in the main text (represented by the thin, light/orange lines) are compared to a histogram of the data from which they were derived (dark/black lines). Data not included in the fits are indicated as well (thick, light/gray lines).

ionization yields—those corresponding to the nuclear recoils—was used to extract a  $1\sigma$  Gaussian width. Similarly, for all but the lowest-energy bins, average ionization yields were determined from Gaussian maximum-likelihood estimates (MLE). For the (two) lowest-energy Ge (Si) bin(s), the ionization-yield distribution(s) was (were) slightly asymmetric. To obtain an accurate estimate of the average ionization yield for these bins, split-width Gaussians were fit to

## CHAPTER 5. A LOW-THRESHOLD ANALYSIS

ionization-yield histograms using a least-squares fitting method.<sup>14</sup> The resulting fits are depicted in Figure 5.15 for a representative detector, with figures for the other detectors provided in Appendix C. Note that in those cases where a maximum-likelihood fit was used, the histograms and Gaussian curves are merely convenient representations of the data and MLEs.

For each detector and bias-voltage run, the above bin-wise estimate of the average ionization yield was parameterized as a function of recoil energy according to a simplified form of Lindhard *et al.*'s theory for nuclear-recoil ionization yield in semiconductor crystals:

$$Y = A_1 E_R^{A_2}, \quad (5.3.5)$$

where  $A_1$  and  $A_2$  are free parameters, and  $E_R$  is  $Q$ -corrected recoil energy. Similarly, an inverse-squared form was fit to the ionization-yield widths:

$$\sigma_Y = \sqrt{B_1^2 + \frac{B_2^2}{E_R^2}}, \quad (5.3.6)$$

where  $B_1$  and  $B_2$  are free parameters. The results are shown in Figure 5.16 for the viable low-threshold detectors. The fitted means are the same as the nuclear-recoil band centroids (*a la* Figure 5.3) used in Section 5.2 to convert total phonon energy to  $Y_{\text{NR}}$ -corrected recoil energy. Note, however, that the fitted widths differ from the centroids' statistical uncertainties. The  $1\sigma$  widths reflect the spread in ionization yield caused by finite resolution, whereas the centroids' statistical uncertainties reflect the uncertainty in the determination of the average ionization yield.

---

<sup>14</sup>Note that except for the maximum-likelihood fits used to fit the ionization-yield distributions in this section, all curve fitting in this thesis was performed using the MATLAB least-squares fitting routine *lsqnonlin*. I used MLEs here because of low nuclear-recoil statistics. Least-squares fits require the ionization yields to be binned (within a given recoil-energy bin), and can yield results that depend strongly on the choice of bins when statistics are low. MLEs are derived from unbinned data, and tend to yield more robust results in the low-statistics limit.

### 5.3. ANALYSIS CUTS

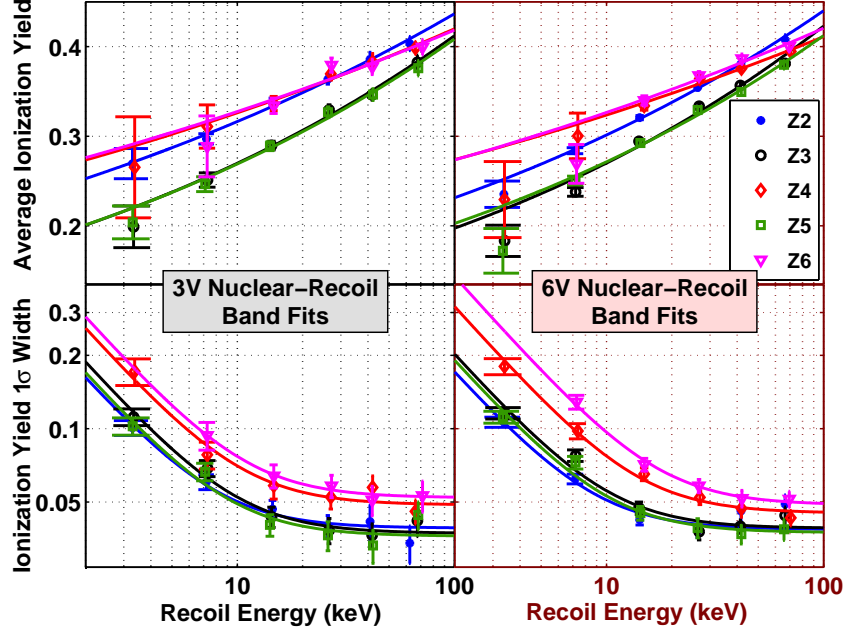


Figure 5.16: Nuclear-recoil band fits for the viable low-threshold detectors. *Upper panels:* The two-parameter power law given by Equation 5.3.5 is fit to a bin-wise estimate of the average ionization yield (error bars). *Lower panels:* The functional form given in Equation 5.3.6 is fit to the corresponding ionization-yield  $1\sigma$  widths. Note that the lowest-energy (2–5 keV) bin was not included in the fits to the Z6 means and widths. The resulting parameterizations (solid lines) for the 3V (left panels) and 6V (right panels) data are used to construct  $2\sigma$  bands for selecting nuclear recoils from the WIMP-search data.

The fits were used to construct  $2\sigma$  nuclear-recoil bands (as in Figure 4.7). An event is considered a nuclear recoil if its ionization yield falls within its detector’s band. This is demonstrated for  $^{252}\text{Cf}$  data in Figure 5.17, in which the nuclear-recoil band for the Z5 6V data is compared to the data from which it was derived. By construction, the efficiency for selecting nuclear recoils (and therefore WIMPs) is expected to be  $\sim 95.5\%$ . As indicated in Figure 5.17 (and Appendix C), the same  $^{252}\text{Cf}$  data were used to directly measure the selection

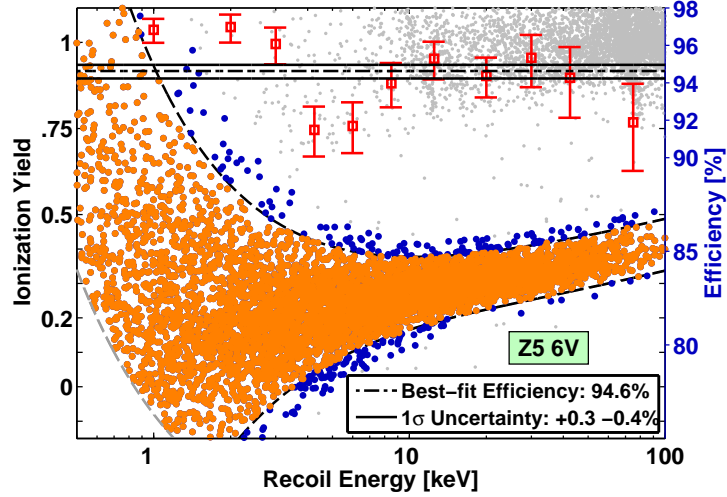


Figure 5.17: The  $2\sigma$  nuclear-recoil band (dark dashed lines) for the Z5 6V data, compared to the  $^{252}\text{Cf}$  single scatters from which it was derived (+ and dots). The band’s selection efficiency is estimated in bins of recoil energy by comparing the events in the  $2\sigma$  band (light/orange dots) to the those in the  $3\sigma$  band (dark/blue dots). As indicated, a constant efficiency is fit to the resulting binomial confidence intervals (error bars). Because events in the nuclear-recoil band are partially excluded by the  $6\sigma$  phonon-energy software threshold (light dashed line) at low energy, the efficiency estimate for the lowest-energy bin (leftmost error bar) was not included in the constant-efficiency fit.

efficiency by comparing the numbers of events in the  $2\sigma$  and  $3\sigma$  nuclear-recoil bands. As expected for an accurate parameterization, no significant dependence on energy was observed. A constant-efficiency hypothesis was fit to each efficiency estimate as a function of recoil energy, yielding a nuclear-recoil cut efficiency of  $92.9\text{--}95.6\pm 0.3\text{--}0.8\%$ , depending on detector and bias voltage.

## 5.4 Energy Scale & Resolution

Before exploring the results of the WIMP-search selection outlined above, it is worth verifying and further characterizing the ZIP-detector energy response. Because the sensitivity of this analysis is so critically dependent on the detection thresholds, even a slight miscalibration of the energy scale for nuclear recoils could potentially cause the low-mass WIMP sensitivity to be overstated. Furthermore, as discussed toward the end of Section 3.3.2, the lowest-mass WIMP sensitivity is partially dependent on near-threshold energy resolution; electronic-noise fluctuations occasionally allow below-threshold energy depositions to cause experimental triggers (and *vice versa*). Similar to the absolute energy scale, over- or underestimating the energy resolution could introduce significant (and possibly nonconservative) errors in the WIMP-nucleon exclusion limits presented in the next chapter, particularly at low WIMP masses.

In this section I first review the electron-recoil energy scale and energy-dependent ionization and phonon resolutions. However, the procedures used do not work equally well for the Si and Ge detectors and, ultimately, the energy response to nuclear recoils is most important. Consequently, I also review a quick check of the nuclear-recoil energy scale by comparing  $^{252}\text{Cf}$  calibration data to Monte Carlo simulations, an analysis that is expanded upon in the next chapter. Finally, the most accurate possible zero-energy resolutions for the  $Q$ - and  $Y_{\text{NR}}$ -corrected recoil-energy estimators are derived from the nonadjacent other-detector triggered data. These are the resolutions used to calculate the WIMP-nucleon exclusion limits presented in the next chapter. As with the sections describing the thresholds and analysis cuts, there are too many figures to include here. Where appropriate, figures for the other bias-voltage run are provided in Appendix D.



### 5.4.1 Electron-Recoil Energy Response

As mentioned in Section 5.1.1, the electron-recoil energy scales were calibrated at the beginning of Run 21b using the 662 keV gammas emitted by a  $^{137}\text{Cs}$  source. This sort of calibration is highly dependent on the spectral endpoint, which (in this case) is at a much higher energy than of interest for WIMP-nucleus interactions. Consequently, there is no guarantee that the calibration is accurate at lower energies. Even a slight nonlinearity in the energy response could lead to large calibration errors at low energy.

Fortunately, for the Ge detectors there are several lines from the decay of internal radioisotopes that span the energy of interest. Using these lines, I checked the phonon and ionization energy scales, and characterized the energy dependence of the energy resolution. Unfortunately, these methods do not work for the Si detectors for which there are no well defined spectral features. However, since the initial calibration using  $^{137}\text{Cs}$  gammas was performed in the same manner for all detectors, it is reasonable to assume that the results presented here for the Ge detectors provide an indication of the accuracy of the initial energy scale for the Si detectors as well.

#### Fitting the Ge Spectral Features

For each of the features in the following list, Gaussian fits to electron-recoil distributions in the WIMP-search data were used to extract means and  $1\sigma$  widths:<sup>15</sup>

- Randomly triggered “noise blobs” at zero energy for the Ge *and* Si detectors;
- 1.29 keV peak from the decay of  $^{68}\text{Ge}$  and  $^{71}\text{Ge}$  via electron capture from the L shell, with the former produced cosmogenically ( $\sim 271$  d half-life) and the latter from thermal neutron capture on  $^{70}\text{Ge}$  ( $\sim 11.4$  d half-life);

---

<sup>15</sup>The expected line energies and isotope half-lives were taken from [412].

#### 5.4. ENERGY SCALE & RESOLUTION

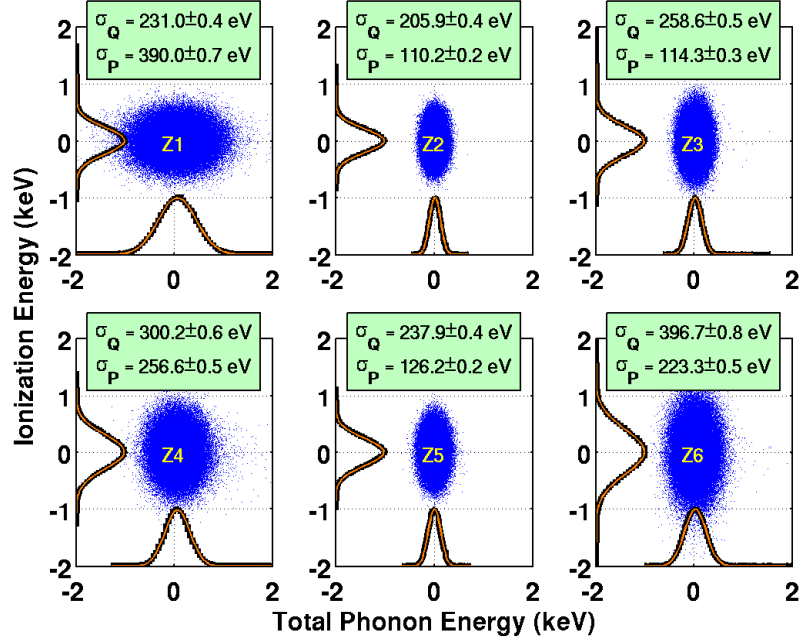


Figure 5.18: Scatter plots of (zero-delay) ionization versus total phonon energy for randomly triggered 3V data. For each detector, Gaussian fits (light/orange lines) were used to extract the means and widths of the projected noise distributions (black lines). The resulting  $1\sigma$  noise resolutions are indicated.

- 8.98 keV peak from the decay of  $^{65}\text{Zn}$ —cosmogenically produced from Cu and Ge isotopes and long lived ( $\sim 244$  d half-life)—via electron capture from the K shell;
- 10.36 keV peak from the decay of  $^{68}\text{Ge}$  and  $^{71}\text{Ge}$  via electron capture from the K shell; and
- 66.7 keV peak from the decay of the cosmogenically induced—typically via medium-energy neutron capture on  $^{72}\text{Ge}$ —metastable  $^{73m}\text{Ge}$  state (0.5 s half-life).

In the following discussion, I review the fits to each of these spectral features.

## CHAPTER 5. A LOW-THRESHOLD ANALYSIS

Fits to the so-called noise blobs—scatter plots of ionization versus total phonon energy—were primarily used to provide a zero-energy anchor for the determination of detector resolutions as a function of energy. In order to do this accurately, zero-delay ionization- and phonon-energy RRQs were used to construct noise distributions from the randomly triggered data.<sup>16</sup> As discussed in Section 5.2.2, the primary fitting routine yields an energy estimate that is biased toward positive energies. Additionally, this bias causes the width of the electronic-noise distribution to be underestimated by as much as 15%. When such means and widths are combined to form thresholds, the biases generally offset one another because they go in opposite directions. To obtain an accurate estimate of the noise resolution, however, the biased energy estimators are clearly inappropriate. The zero-delay RRQs avoid the bias by forcing the fit start time to occur at the trigger time, making identification of negative and positive fluctuations equally likely (as opposed to searching for the largest positive fluctuation). The zero-delay noise blobs for the 3V data are shown in Figure 5.18 (see Appendix D for the 6V data), with the ionization- and phonon-energy noise distributions projected onto the  $y$  and  $x$  axes, respectively. These projected distributions were fit independently, providing a zero-energy mean and  $1\sigma$  width for each detector and bias-voltage run.

Both  $^{71}\text{Ge}$  and  $^{68}\text{Ge}$  typically decay via electron capture. The remaining electrons find themselves bound to a Ga nucleus and quickly adjust to its slightly less positive electric field, resulting in a Ga isotope with a hole in one of its electron shells. The hole radiates outward, releasing the binding energy of the Ga shell via a cascade of x rays and/or Auger electrons. The time scale for the orbital electrons to adjust to the electric field of the Ga nucleus is effectively instantaneous relative to the time it takes the hole to move outward. Consequently, such decays are characterized by the Ga (rather

---

<sup>16</sup>The non-Gaussian tails discussed in Section 5.1.1 were mostly avoided by requiring the phonon-pulse start time found by the main fitting routine to be well separated from the  $-50(-25)\mu\text{s}$  minimum for the Ge (Si) detectors. Because this reduces the statistics and does not completely remove the non-Gaussian tails, the resolutions acquired in this manner were not used to calculate WIMP-nucleon exclusion limits. The results of the dedicated zero-energy resolution analysis presented in Section 5.4.3 were used instead.

#### 5.4. ENERGY SCALE & RESOLUTION

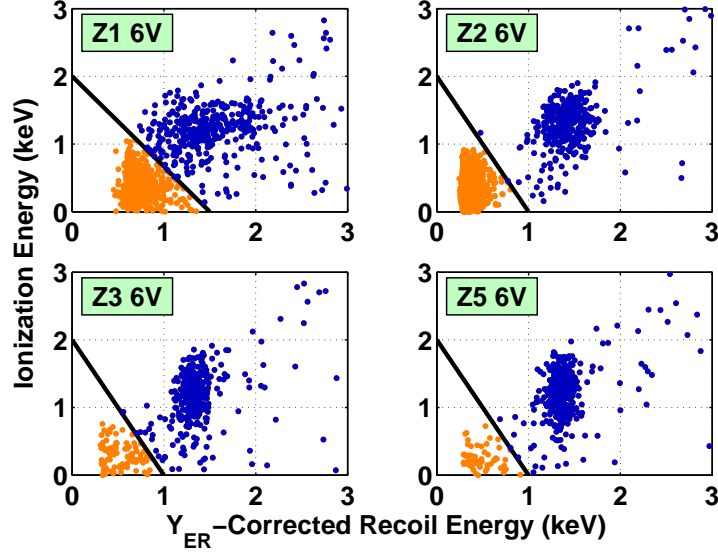


Figure 5.19: Scatter plot of ionization versus phonon energy, demonstrating how 6V WIMP-search events were preselected for the fits to the 1.29 keV line in the Ge detectors. Events associated with electronic noise (light/orange dots) are well separated from the 1.29 keV peak (dark/blue dots), allowing the preselection cuts to be chosen by eye (black lines).

than Ge) electron-shell binding energies. The most probable shells from which electrons are captured are the K and L, thus providing convenient lines at 10.36 and 1.29 keV, respectively. Similar decays occur for any cosmogenically produced  $^{65}\text{Zn}$ . However, because K-shell captures occur nearly an order of magnitude more often than L-shell captures, and there were relatively few  $^{65}\text{Zn}$  decays during Run 21, the lower-energy line from the L-shell decay of  $^{65}\text{Zn}$  was not observed. Furthermore, the 8.98 keV line resulting from K-shell decays was barely resolved due to its proximity to the higher-rate 10.36 keV line.

Obtaining an accurate fit to the 1.29 keV line is complicated by its proximity to the hardware threshold. In the ionization channel, the quasi-exponential rise of noise events toward zero energy was suppressed by preselecting events

## CHAPTER 5. A LOW-THRESHOLD ANALYSIS

in the ionization- versus phonon-energy plane.<sup>17</sup> This is demonstrated for the Ge detectors in Figure 5.19 for the 6V data. As shown in Figure 5.20, a Gaussian (plus linear background) was fit to a histogram of the selected ionization energies for each detector. For the corresponding phonon-energy distributions (in terms of  $Y_{\text{ER}}$ -corrected recoil energy), the hardware trigger efficiencies were folded into the fits to better follow the low-value tails. The results for the 6V data are shown in Figure 5.20 as well, with a similar figure for the 3V data provided in Appendix D. Note that the resolution of the phonon measurement is superior to the ionization measurement at these low energies (except for Z1), an indication of the superb low-noise performance of the phonon channel's SQUID-based readout.

The means and widths of the 8.98 and 10.36 keV lines were measured simultaneously. For the ionization measurement, the fits were relatively straightforward; a functional form including two Gaussians and a linear background was fit to the spectrum of 5–15 keV ionization energies for each Ge detector. The results for the 3V WIMP-search data are shown in Figure 5.21. Above a few keV, the phonon resolution degrades more rapidly than the ionization resolution. Consequently, the lower-rate 8.98 keV line is not well resolved in the phonon channel, appearing as a bump on the low-value tail of the 10.36 keV peak. In order to accurately characterize either peak, constraints had to be added to the fits: 1) the difference in mean peak energies was restricted to  $\sim 1.4 \pm 0.1$  keV; and 2) the amplitude of the 8.98 keV Gaussian was constrained by the ratio of the integrals of the Gaussians fitted to the corresponding 8.98 and 10.36 keV peaks in the ionization channel. The results of these fits for the 3V data are also shown in Figure 5.21, with a similar figure for the 6V data provided in Appendix D. Note the elevated rate of 10.36 keV events for the 6V WIMP-search data in Figure D.4. This is due to the extensive  $^{252}\text{Cf}$  calibrations performed between the 3V and 6V WIMP searches, which caused

---

<sup>17</sup>Only single scatters from WIMP-search data that passed the fiducial-volume and data-quality cuts were considered for the peak fits discussed in this section. Note that the pre-selection depicted in Figure 5.19 effectively rejects the Z2 event bursts clustered near zero energy, allowing the full 6V exposure to be utilized for the Z2 fits.

#### 5.4. ENERGY SCALE & RESOLUTION

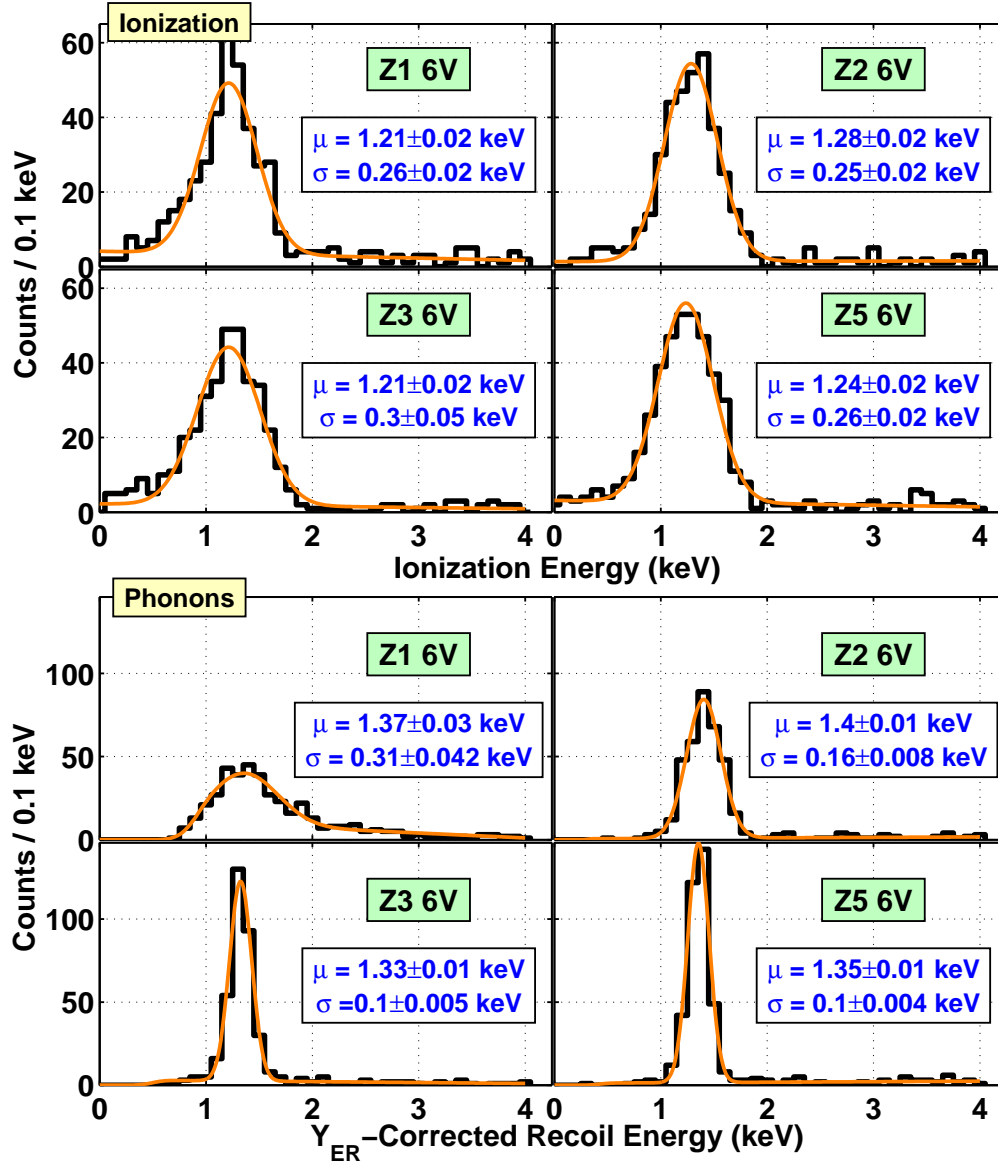


Figure 5.20: Gaussian fits (light/orange lines) to histograms of the 1.29 keV peak (dark/black lines) as observed by the Ge detectors' ionization (upper four panels) and phonon (lower four panels) channels for 6V WIMP-search data. The means ( $\mu$ ) and  $1\sigma$  widths obtained from the fits are indicated.

## CHAPTER 5. A LOW-THRESHOLD ANALYSIS

significant neutron activation of  $^{70}\text{Ge}$  atoms. As a result, the 8.98 keV peak was even more difficult to resolve in the 6V data. Constraints were needed to obtain accurate fits to the 6V phonon *and* ionization peaks. Both of the constraints used for the 3V phonon fits were applied to the former, while only the first was used for the latter.

At higher energy, but still within the energy range of interest, is the 66.7 keV line due to decays of the metastable  $^{73m}\text{Ge}$  state. The most probable decay mode actually involves successive emission of 53.4 and 13.3 keV photons as  $^{73m}\text{Ge}$  quickly de-excites via an intermediary energy level (with a few  $\mu\text{s}$  half-life) to the stable  $^{73}\text{Ge}$  state. This occurs effectively instantaneously relative to the time scales of the ionization and phonon measurements, and because both photons are emitted (and eventually absorbed) within the Ge substrates, these decays appear at 66.7 keV. For each Ge detector, the peak was easily characterized in the ionization and phonon channels via a Gaussian fit (plus linear background) over an energy range of  $\sim 30\text{--}90$  keV. This is demonstrated for the 6V WIMP-search data in Figure 5.22, with a similar figure provided in Appendix D for the 3V data.

### Energy-Scale Check

The average peak energies measured above were used to check the electron-recoil energy scales for the viable low-threshold Ge detectors. Toward this end, the observed locations of the 1.29, 10.36, and 66.7 keV peaks were used to gauge the accuracy of the energy scale at low, medium, and high energies of interest for WIMP-nucleus interactions. Due to the superior quality of the 10.36 keV fits, the 8.98 keV peak was not used here. For the three lines considered, there are significant deviations from the expected peak locations, with fewer than half of the observations consistent with  $2\sigma$  statistical fluctuations. However, as I hope to convince you with the following discussion, these discrepancies are such that the original  $^{137}\text{Cs}$ -calibrated energy scale is conservative with regards to low-mass WIMP sensitivity.

#### 5.4. ENERGY SCALE & RESOLUTION

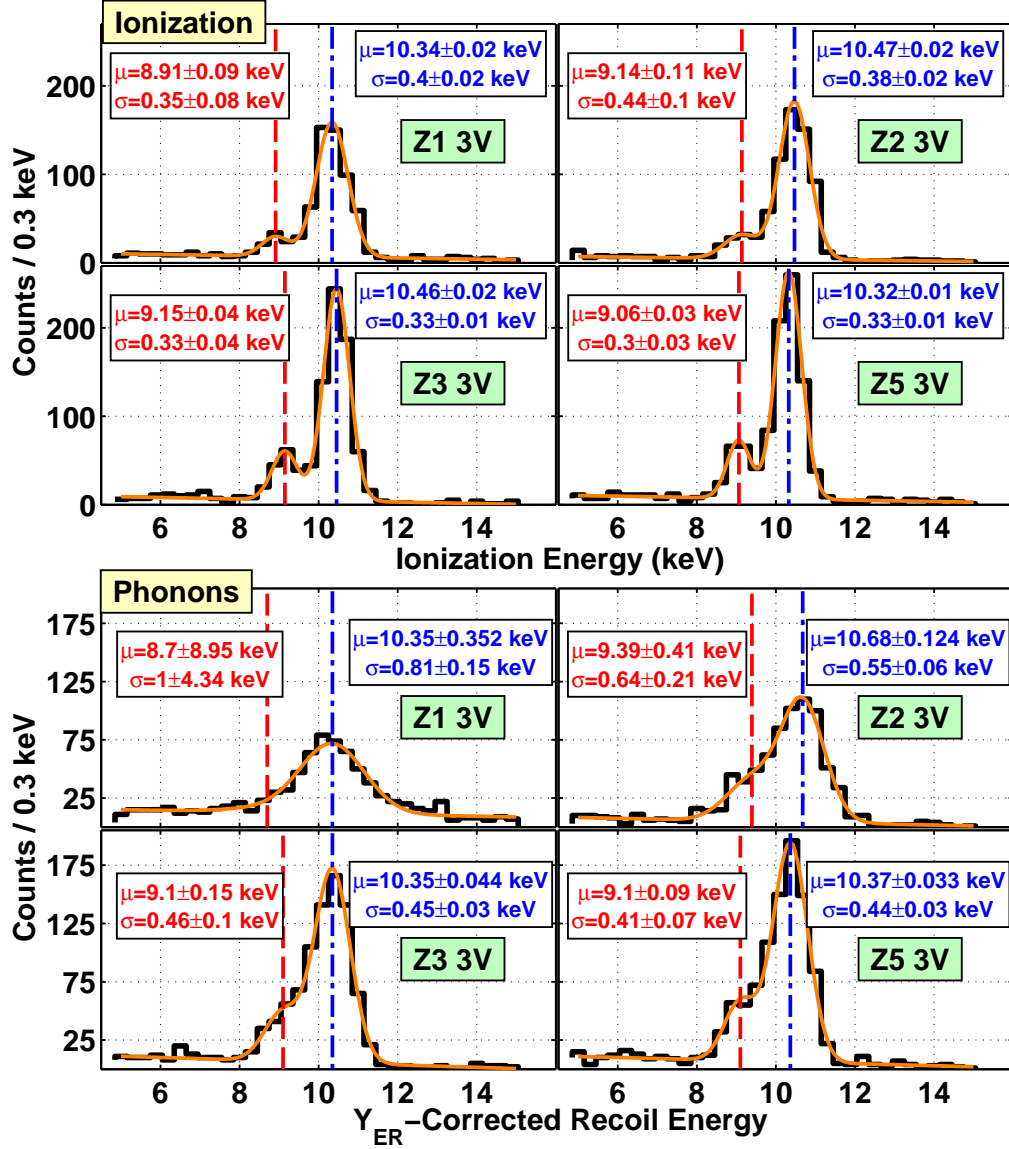


Figure 5.21: Gaussian fits (light/orange lines) to histograms of the 8.98 and 10.36 keV peaks (dark/black lines) as observed by the Ge detectors' ionization (upper four panels) and phonon (lower four panels) channels for 3V WIMP-search data. The means ( $\mu$ ) and  $1\sigma$  widths obtained from the fits are indicated.



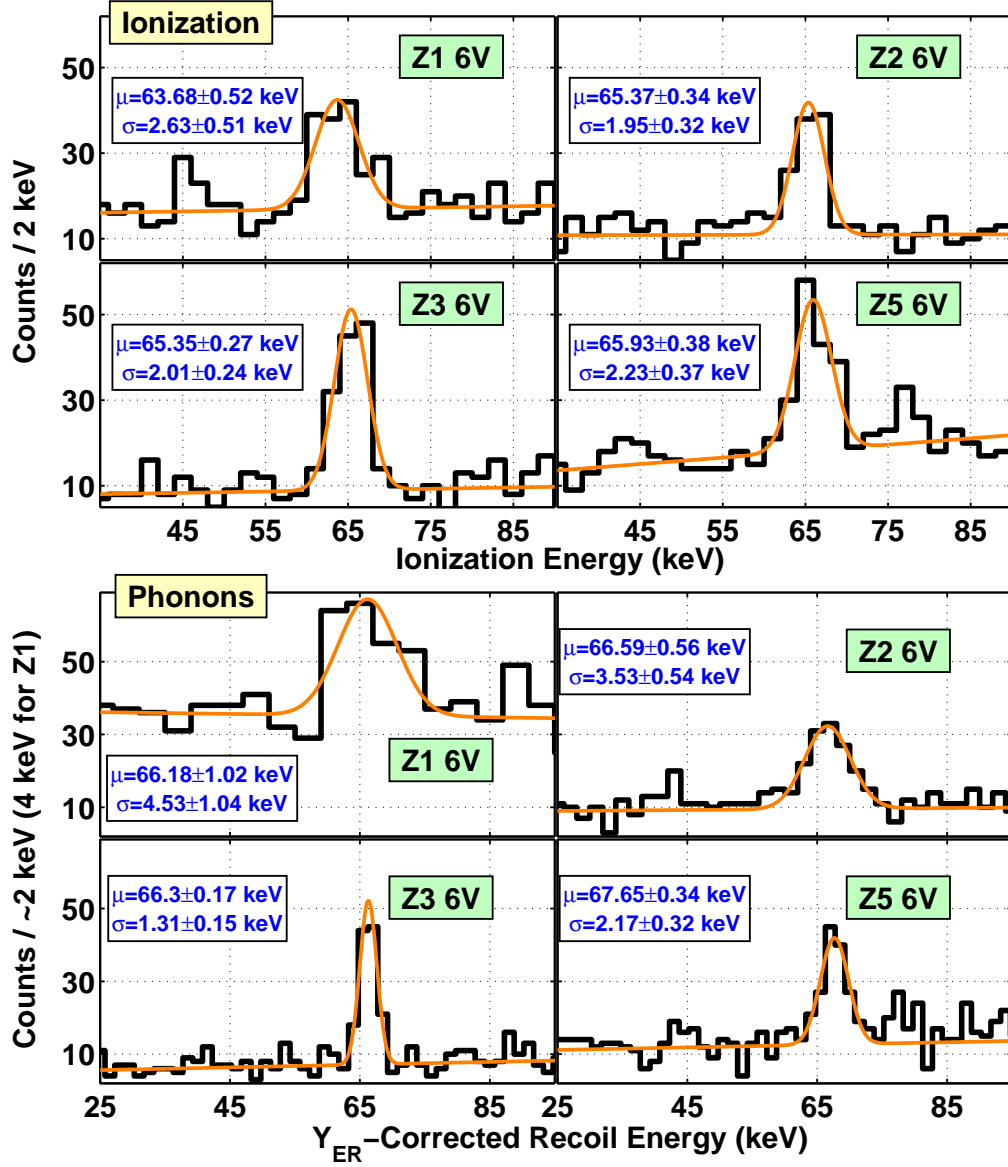


Figure 5.22: Gaussian fits (light/orange lines) to histograms of the 66.7 keV peak (dark/black lines) as observed by the Ge detectors' ionization (upper four panels) and phonon (lower four panels) channels for 6V WIMP-search data. The means ( $\mu$ ) and  $1\sigma$  widths obtained from the fits are indicated.

#### 5.4. ENERGY SCALE & RESOLUTION

Table 5.4: Results of an energy-scale correction based on the observed and expected locations of the 1.29, 10.36, and 66.7 keV peaks for the viable low-threshold Ge detectors. For each detector and bias voltage, the percent changes (relative to the original energy scale) required to correct 1, 10, and 50 keV  $Q$ - and  $Y_{\text{NR}}$ -corrected recoil energies are indicated. A negative (positive) sign means the corrected energy is less (greater) than the original estimate.

Detector	$Q$ -Corrected $E_R$		$Y_{\text{NR}}$ -Corrected $E_R$	
	3V Data	6V Data	3V Data	6V data
Percent Change at 1 keV				
Z2	−10.4%	−20.8%	−4.4%	−6.9%
Z3	−7.9%	−17.3%	−0.9%	−2.3%
Z5	−11.0%	−17.3%	+1.5%	−3.7%
Percent Change at 10 keV				
Z2	−4.9%	−12.5%	−3.0%	−5.3%
Z3	+1.1%	−1.4%	+0.02%	−1.0%
Z5	−0.6%	−8.1%	+0.05%	−3.3%
Percent Change at 50 keV				
Z2	+1.2%	−4.0%	+0.8%	−1.1%
Z3	−1.1%	−2.1%	−0.3%	+0.2%
Z5	−1.9%	−6.3%	+0.2%	−1.8%

Unfortunately, the peak energies could not be recalibrated to their expected values via a simple linear fit without introducing errors larger than those already observed; there is a nontrivial nonlinearity in the energy response. Instead, corrected energy scales were calculated in which the three peaks were shifted to their expected values and all energies in between were linearly in-

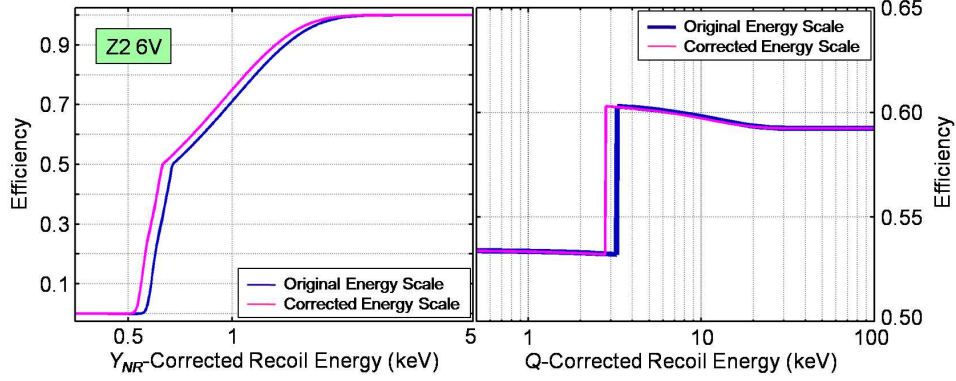


Figure 5.23: Comparison of the combined nuclear-recoil detection efficiencies—thresholds (left panel) and analysis cuts (right panel)—for the original and recalibrated Z2 6V energy scales.

terpolated to new values accordingly. In order to make this work for energies below 1.29 keV and above 66.7 keV, the energy scales were kept fixed at 0 and 100 keV. This is a valid assumption at 0 keV (zero is zero). Any systematic uncertainty introduced at 100 keV is restricted to energies  $>66.7$  keV due to the nature of the correction, and should therefore have little to no effect on low-mass WIMP sensitivity. The resulting corrected ionization and phonon energy scales were used to derive recalibrated energy scales in terms of  $Y_{\text{NR}}$ - and  $Q$ -corrected recoil energy, the two energy estimators relevant to the calculation of WIMP-nucleon exclusion limits. Table 5.4 summarizes the corresponding shifts in energy for a few example energies.

There are two important trends with regard to low-mass WIMP sensitivity: the original energy scales generally overestimate (underestimate) phonon (ionization) energies relative to the corrected scales, particularly at low energies. Near threshold, this corresponds to a modest overestimate of  $Y_{\text{NR}}$ -corrected recoil energy for most detectors (by a few %), and a more significant overestimate of  $Q$ -corrected recoil energy (by 10–20%). Correcting the former shifts the hardware and software thresholds to lower energies, while correcting the latter shifts the combined efficiency of the analysis cuts such that there is slightly

## 5.4. ENERGY SCALE & RESOLUTION

more efficiency for the detection of low-energy nuclear recoils. This is demonstrated for the Z2 6V data in Figure 5.23, the detector and bias-voltage run with the largest correction. On average, application of the corrected energy scales results in greater detection efficiency, and therefore stronger WIMP-nucleon exclusion limits. The improvement, however, is relatively minor. Rather than risk introducing systematic uncertainty related to the *ad hoc* nature of this recalibration, I made the slightly conservative choice to use the Ge detectors' original energy scales for the exclusion-limit calculations presented in the next chapter.

### Energy-Dependent Energy Resolution

The detectors' phonon and ionization resolutions are both energy dependent. This dependence was characterized for each viable low-threshold Ge detector by fitting the following functional form to the widths derived from the Gaussian fits described above:

$$\sigma_E = \sqrt{A_1^2 + A_2E + A_3^2E^2}, \quad (5.4.1)$$

where  $A_{1-3}$  are free parameters, and  $E$  is either ionization or phonon energy. The constant term corresponds to the zero-energy noise resolution, while the term proportional to  $\sqrt{E}$  is associated with (not-quite-Poisson) fluctuations in the number of phonons and charge carriers produced during an interaction of a given energy (the Fano factor [765]). The third term ( $\propto E$ ) is the result of the template-based pulse-fitting method used to estimate ZIP-detector event energies. The templates represent averaged pulse shapes. Variations in pulse shape (relative to the templates) cause the optimal-filter estimates to fluctuate about an event's true energy. The amplitude of these fluctuations grows linearly with energy, giving rise to the third term in Equation 5.4.1. This is supported by the observation that the phonon resolution degrades more quickly (with increasing energy) than the ionization resolution, a consequence of the position-dependent phonon pulse shapes caused by the detectors' residual  $T_c$  gradients.

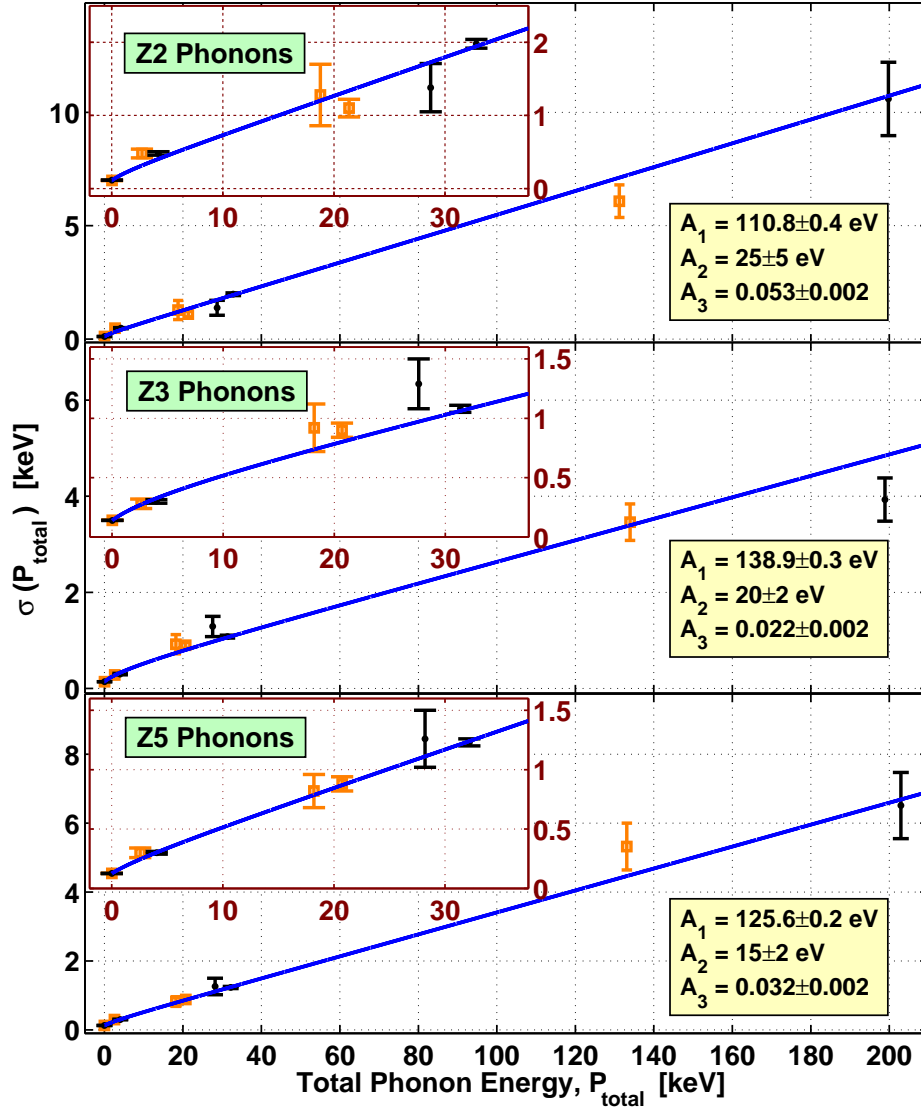


Figure 5.24: Fits to the phonon resolution as a function of total phonon energy ( $P_{\text{total}}$ ) for the viable low-threshold Ge detectors. In each case, the  $1\sigma$  widths resulting from the fits to the 3V (light/orange squares) and 6V (dark/black dots) spectral features were fit simultaneously using the functional form given in Equation 5.4.1, with the inset highlighting the low-energy region. The resulting  $A_{1-3}$  fit parameters are indicated.

#### 5.4. ENERGY SCALE & RESOLUTION

Table 5.5: Summary of the fit parameters resulting from fits of the form described by Equation 5.4.1 to the 3V and 6V ionization resolutions for the viable low-threshold Ge detectors. Note that while  $A_1$  and  $A_2$  are in units of energy,  $A_3$  is dimensionless.

Detector	Bias	$A_1$ (eV)	$A_2$ (eV)	$A_3$
Z2	3V	205.9 $\pm$ 0.4	4 $\pm$ 3	0.025 $\pm$ 0.004
	6V	201.8 $\pm$ 0.4	9 $\pm$ 3	0.026 $\pm$ 0.006
Z3	3V	258.6 $\pm$ 0.5	0.2 $\pm$ 1.3	0.019 $\pm$ 0.003
	6V	246.6 $\pm$ 0.5	0.0 $\pm$ 1.4	0.021 $\pm$ 0.004
Z5	3V	237.9 $\pm$ 0.4	0.9 $\pm$ 1.4	0.020 $\pm$ 0.003
	6V	238.0 $\pm$ 0.5	0.0 $\pm$ 3.6	0.028 $\pm$ 0.006

For each detector, a single fit was performed to the collection of 3V and 6V  $1\sigma$  widths as a function of total phonon energy. As a function of  $Y_{\text{ER}}$ -corrected recoil energy, the spectral features described above appear to occur at only five unique energies (0, 1.29, 8.98, 10.36, and 66.7 keV). However, due to the difference in drift-heat amplification between the two bias-voltage runs ( $\times 2$  ( $\times 3$ ) for 3V (6V) electron recoils), when the average peak locations are instead expressed in terms of  $P_{\text{total}}$ , the  $1\sigma$  widths occur at nine unique energies that span  $\sim 0$ –200 keV (the zero-energy peaks occur at 0 keV regardless of bias voltage). Consequently, the phonon resolution’s energy dependence is better constrained as a function of  $P_{\text{total}}$ . This is demonstrated in Figure 5.24, where fits of the form given in Equation 5.4.1 are compared to the collection of 3V and 6V peak widths for each viable Ge detector.

The ionization resolutions were fit separately for the 3V and 6V data. The use of different bias voltages likely resulted in slightly different charge-collection efficiencies for the two runs, resulting in differences in resolution as a function of energy. Consequently, independent fits are more appropriate than a combined

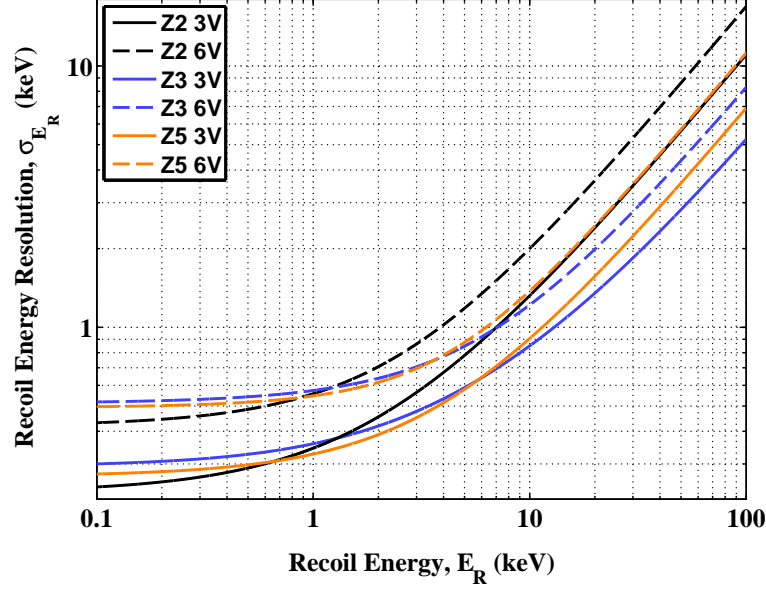


Figure 5.25: Energy dependence of the viable low-threshold Ge detectors' ( $Q$ -corrected) recoil-energy resolutions.

fit. Table 5.5 summarizes the fit parameters for each viable Ge detector, demonstrating that there are indeed significant (albeit slight) ionization-resolution differences between the two bias-voltage runs.

The ionization and phonon resolutions can be combined to estimate the energy dependence of the ( $Q$ -corrected) recoil-energy resolution according to the following prescription:

$$\begin{aligned} \sigma_E(E) &= \sqrt{\sigma_{P_{\text{total}}}^2(P_{\text{total}}) + \left(\frac{eV}{\epsilon}\right)^2 \sigma_Q^2(Q)} \\ &= \sqrt{\left(1 + \frac{eV}{\epsilon}\right)^2 \sigma_{E'}^2(E') + \left(\frac{eV}{\epsilon}\right)^2 \sigma_Q^2(Q)}, \end{aligned} \quad (5.4.2)$$

where  $E$  ( $E'$ ) represents  $Q$ -corrected ( $Y_{\text{ER}}$ -corrected) recoil energy, and the definitions for the other variables are the same as in Equation 4.2.4. The second line in Equation 5.4.2 facilitates the evaluation of  $\sigma_E(E)$  by placing the ioniza-

## 5.4. ENERGY SCALE & RESOLUTION

tion and phonon resolutions on the same scale as one another. The resulting  $Q$ -corrected recoil-energy resolutions are compared for the viable low-threshold Ge detectors for both bias-voltage runs in Figure 5.25. Note the sizable difference in recoil-energy resolution resulting from the difference in drift-heat amplification between the 3V and 6V runs.

Another important feature of Figure 5.25 is that the recoil-energy resolutions are dominated by the (constant) zero-energy resolutions near the detector thresholds. Consequently, not much is gained by including the full energy dependence when calculating WIMP-nucleon exclusion limits. For simplicity, only the zero-energy term will be considered from here on. This also allows the Si detectors—for which the energy dependence could not be measured—to be treated in an identical fashion to the Ge detectors. Out of curiosity, I calculated Ge exclusion limits with and without the resolution’s energy dependence. The difference is not visible by eye. Furthermore, for low WIMP masses, the exclusion limits calculated with the constant zero-energy resolution are very slightly conservative compared to those calculated with the resolution’s full energy dependence.

### 5.4.2 Nuclear-Recoil Energy Scale

As explained in Section 4.2.3, there are two methods available for checking the ZIP detectors’ nuclear-recoil energy scales. The comparison of nuclear-recoil band centroids to Lindhard theory (demonstrated in Figure 4.8) gives rough confirmation of the recoil-energy response. Alternatively, the differential event rates observed for nuclear recoils from  $^{252}\text{Cf}$  calibrations can be compared to Monte Carlo simulated spectra. Because the spectra are basically featureless (decaying) exponentials, this approach is limited by the available nuclear-recoil statistics. However, if the observed events are summed to produce mean Ge and Si recoil spectra, there is sufficient statistical power to make a meaningful test of the nuclear-recoil energy scale. In this section I outline a quick check that compares the decay constants of exponential fits to the observed and simulated spectra, thereby deriving scale factors with which to gauge the accuracy of



## CHAPTER 5. A LOW-THRESHOLD ANALYSIS

the energy scale for nuclear recoils. A more detailed and statistically rigorous approach is explored in the next chapter.<sup>18</sup>

All available  $^{252}\text{Cf}$  neutron-calibration data for the viable low-threshold detectors were used for this study, corresponding to a total Ge (Si) raw exposure of 1.22 (0.38) kilogram-days. Events were required to pass all of the criteria outlined in Section 5.3 (except for the event-burst cut, for which there is no equivalent for the  $^{252}\text{Cf}$  data)—single-scatter, veto-anticoincident nuclear recoils passing the data-quality and fiducial-volume cuts. The selected events were sorted into 5 keV recoil-energy bins spanning  $E_R = 5\text{--}100\text{ keV}$ . The resulting histogrammed event rates were corrected for the selection-cut efficiencies, and expressed in terms of absolute differential event rates (events per keV kg d). The six (four) individual Ge (Si) spectra—one for each detector and bias-voltage run—were combined to form a single representative recoil-energy distribution.

Corresponding spectra were generated with a GEANT3 [766] simulation of the  $^{252}\text{Cf}$  source and a detailed geometry of the shallow-site setup.<sup>19</sup> In short, the differential spectrum of neutron energies emitted by the  $^{252}\text{Cf}$  source (see, *e.g.*, Figure E.1) was transported through the shielding layers, resulting in a moderated and roughly exponential spectrum (with a decay constant of  $\sim 1\text{ MeV}$ ) of neutron energies incident upon the ZIP detectors (see, *e.g.*, Figure E.2). When these neutrons scattered from nuclei within the detectors, the kinetic energies of the recoiling Ge and Si nuclei were recorded. Note that because it is a simulation, recoil energies can be inferred directly from nuclear velocities without having to simulate the phonon or ionization response, thus bypassing a number of practical issues related to the operation of ZIP

---

<sup>18</sup>Two very similar tests of the nuclear-recoil energy scale—the quick one presented here and the more detailed one discussed in the next chapter—were conducted for practical reasons. The results of the quick test were available on a shorter time scale, which allowed them to be utilized in [596], the publication featuring the analysis presented in this thesis. The more detailed energy-scale check was performed after the fact. Fortunately, the former turns out to be conservative relative to the latter. The differences are discussed in the context of WIMP-nucleon exclusion limits toward the end of the next chapter.

<sup>19</sup>The GEANT3 Monte Carlo simulations were performed by Sharmila Kamat. Further details are available in her Ph.D. thesis [767].

#### 5.4. ENERGY SCALE & RESOLUTION

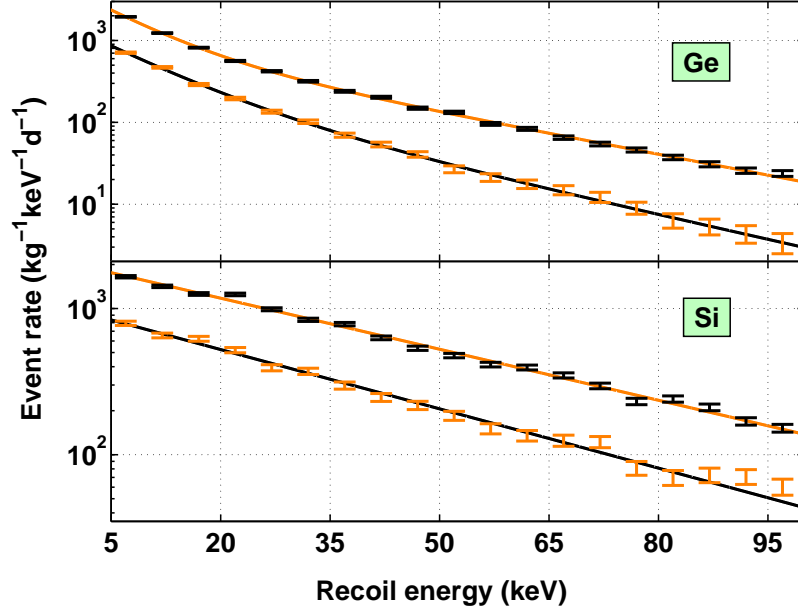


Figure 5.26: Comparison of experimentally measured (light/orange error bars) to Monte Carlo simulated (dark/black error bars) recoil-energy spectra for nuclear recoils from  $^{252}\text{Cf}$  calibrations. Exponential fits of the form given in Equation 5.4.3 are overlaid for both observed (dark/black lines) and simulated (light/orange lines) data. Despite the large discrepancy in absolute event rate due to a poor understanding of the strength of the shallow-site neutron source, the decay constants of the fitted curves can be compared to gauge the accuracy of the nuclear-recoil energy scale. Figure adapted from [596].

detectors. In principle, if the shielding layers, detectors, source and relevant physics processes are modeled with sufficient accuracy, the results of such a simulation should be indicative of the detectors' true recoil-energy response, providing a control sample against which the observed response can be compared. The recoil energies from a simulated Ge (Si) 1.16 (0.36) kg d exposure were binned and combined to form a single Ge (Si) recoil-energy distribution. Note that because the efficiency for selecting simulated nuclear recoils is perfect, the simulated sample has superior statistics despite the slightly smaller exposure.

## CHAPTER 5. A LOW-THRESHOLD ANALYSIS

The observed and simulated spectra were characterized by exponential fits:

$$\begin{aligned} \text{Ge functional form: } \frac{dR}{dE_R} &= A_1 e^{-E_R/A_2} + A_3 e^{-E_R/A_4}, \text{ and} \\ \text{Si functional form: } \frac{dR}{dE_R} &= A_1 e^{-E_R/A_2}, \end{aligned} \tag{5.4.3}$$

where  $A_{1-4}$  are free parameters. A numerical calculation of the expected shapes is presented in Appendix E, and does not agree exactly with the above hypothesis. However, detector resolution and lack of statistics smear the spectra sufficiently for Equation 5.4.3 to provide a reasonable description. Fits to the simulated and observed Ge and Si spectra are compared in Figure 5.26. There is a substantial discrepancy between the fitted decay constants for the Si spectra:  $A_2 = 32.1 \pm 0.6 \text{ keV}$  for the observed spectrum versus  $37.2 \pm 0.5 \text{ keV}$  for the simulated spectrum. This implies that the Run21 Si energy scale underestimates nuclear recoil energies by  $16 \pm 3\%$ . Since this is in the non-conservative direction, I adjusted the Si energy scale upwards by 16%, which has the effect of reducing the Si detectors' low-energy detection efficiencies. The combined detection efficiencies presented below in Section 5.5 reflect this correction. The fitted decay constants for the Ge spectra exhibit smaller discrepancies that go in opposite directions. The low-energy decay constants— $A_2 = 8.9 \pm 0.8 \text{ keV}$  for the observed spectrum versus  $7.3 \pm 0.5 \text{ keV}$  for the simulated spectrum—imply that the Run21 Ge energy scale overestimates the recoil energy for low-energy nuclear recoils (by  $18 \pm 9\%$ ), whereas the high-energy decay constants— $A_4 = 21.8 \pm 1.9 \text{ keV}$  for the observed spectrum versus  $25.1 \pm 0.7 \text{ keV}$  for the simulated spectrum—imply the opposite for higher-value recoil energies (by  $15 \pm 11\%$ ). Since the observed and simulated Ge energy scales are consistent at the  $1-2\sigma$  level, and the mismeasurement at low energies is in the conservative direction, I did not correct the Ge energy scale. This choice results in a slightly weaker WIMP-nucleon exclusion limit for low-mass WIMPs.

## 5.4. ENERGY SCALE & RESOLUTION

### 5.4.3 Zero-Energy Resolution

To obtain the most accurate possible zero-energy noise resolutions for use in the exclusion-limit calculations discussed in the next chapter, a dedicated analysis was performed to characterize the viable low-threshold detectors' sub-threshold noise distributions using the largest, highest-quality sample of empty events available. Two forms of energy resolution are relevant: the  $Y_{\text{NR}}$ -corrected and  $Q$ -corrected recoil-energy resolutions. The former is needed to properly include the effect of subthreshold energy depositions occasionally exceeding the phonon-based software and hardware thresholds due to electronic-noise fluctuations, and therefore does not include any noise contributions from the ionization channel. The effect of the latter on low-mass WIMP sensitivity is virtually negligible. However, since the combined efficiency of the analysis cuts (discussed below) exhibits a slight dependence on  $Q$ -corrected recoil energy, its intrinsic resolution includes ionization as well as phonon noise. Consequently, both types of recoil-energy resolution are needed to fully account for finite-resolution effects.

As discussed previously, the zero-delay ionization and phonon energy estimators for the nonadjacent other-detector triggered data provide the least biased sampling of the electronic noise environment. A specialized set of analysis cuts based on the zero-delay RRQs was developed to select the noise events directly responsible for occasionally promoting subthreshold WIMP-like energy depositions into the detectors' signal regions. The primary challenge is to select signal-like events without allowing the hardware or software thresholds to bias the selection. The other-detector triggered data naturally avoid the hardware-trigger bias, while the single-scatter criterion is modified slightly from the cut described in Section 5.3.4 to avoid the phonon-energy software threshold. The singles cut used here places no restriction on the phonon energy for the detector of interest, while rejecting any events with energies exceeding the  $6\sigma$  thresholds for the other detectors. The data-quality, fiducial-volume, veto-anticoincidence, and nuclear-recoil cuts are analogous to those described in Section 5.3. The resulting event selection represents a thresholdless WIMP-search signal region.

## CHAPTER 5. A LOW-THRESHOLD ANALYSIS

Table 5.6: The zero-energy  $1\sigma$  recoil-energy resolutions for the viable low-threshold detectors are listed for each bias-voltage run. All measurements are rounded to the nearest eV and have an accuracy of  $\sim 1\%$ . Table taken from [596].

Detector	$Y_{\text{NR}}$ -corrected (eV)		$Q$ -corrected (eV)	
	3V	6V	3V	6V
Z2	95	88	223	387
Z3	126	114	282	477
Z4	208	196	398	538
Z5	113	102	258	450
Z6	185	179	434	678

When applied to the nonadjacent other-detector triggers, the selected events provide a fair sampling of the electronic noise I have so carefully worked to exclude from the signal region of the analysis described in this chapter. Consequently, these events can be used to predict the background rate due to electronic noise in addition to measuring the relevant noise resolutions.

For each detector and bias-voltage run, the selected zero-delay phonon energies were converted to  $Y_{\text{NR}}$ -corrected recoil energy and histogrammed. Similarly, the selected phonon and ionization energies were combined to form an equivalent distribution as a function of  $Q$ -corrected recoil energy. The width of each distribution was extracted with a Gaussian fit. The resulting  $1\sigma$  zero-energy noise resolutions are summarized in Table 5.6. Rather than present individual figures for each detector and bias-voltage run, the method is demonstrated in Figure 5.27 for the combined noise distributions as a function of  $Y_{\text{NR}}$ -corrected recoil energy. These combined noise cores were further subjected to the phonon-energy software thresholds and the hardware trigger efficiencies to obtain a prediction of the rate of noise events in the signal region of the main analysis. Nominal recoil-energy thresholds of 0.5 and 1 keV were employed for

#### 5.4. ENERGY SCALE & RESOLUTION

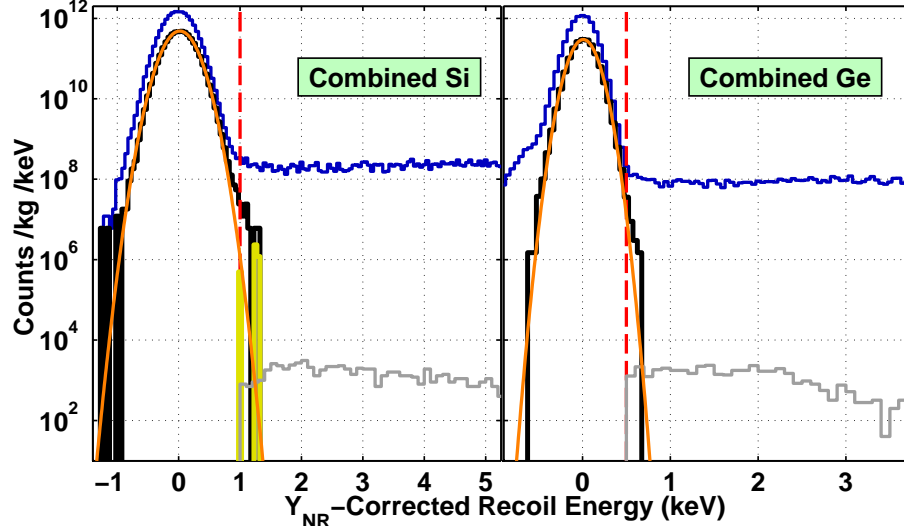


Figure 5.27: The combined noise-core distributions for the low-threshold Ge and Si detector ensembles as a function  $Y_{\text{NR}}$ -corrected recoil energy. In each case, the distribution of all noise events (thin, dark/blue histogram) is reduced by a specialized set of thresholdless WIMP-search cuts (thick, dark/black histogram), and a Gaussian fit (orange solid line) is used to extract the  $1\sigma$  recoil-energy noise resolution. The noise cores were further subjected to the phonon-energy thresholds, resulting in a prediction for the rate of noise events in the signal region (thick, light/yellow histogram) above nominal recoil-energy thresholds (red dashed lines). The actual WIMP-candidate event rates (described in more detail in the next chapter) are included for comparison (thin, light/gray histograms). The predicted rate of signal-region noise events for the Ge ensemble is several orders of magnitude less than the observed WIMP-candidate event rate, and is therefore not visible on this scale. For the Si ensemble, the noise core exhibits a slightly non-Gaussian high-value tail that is likely due to  $^{14}\text{C}$  contamination on Z6's bottom surface.

the Ge and Si detectors, respectively. The predicted noise-event background is so small for the Ge detectors that it is not visible on the scale plotted in Figure 5.27. For the Si detector ensemble, a rate of noise events is predicted that far exceeds the observed WIMP-candidate event rate. These events are likely a

result of  $^{14}\text{C}$  contamination on the bottom surface of Z6. This is the reason Z6 was placed at the bottom of the tower with its phonon side facing away from the adjacent detector. Beta decays of  $^{14}\text{C}$  emit a spectrum of electron energies that extends to 0 keV. Such events could easily leak into the noise-event selection described above, contributing preferentially to the noise core’s high-value tail. This is supported by the observation that the low-value tail of the combined Si noise core is smaller, and in better agreement with the Gaussian fit.

## 5.5 Combined Detection Efficiencies

The detection efficiencies for each detector and bias-voltage run naturally combine into two distinct curves: the threshold efficiency as a function of  $Y_{\text{NR}}$ -corrected recoil energy (including the hardware and software phonon-energy thresholds described in Section 5.2), and the combined efficiency of the analysis cuts as a function  $Q$ -corrected recoil energy. As will be discussed in the next chapter, to properly include finite-resolution effects when calculating WIMP-nucleon exclusion limits, it is necessary to apply these efficiencies separately (just as it was necessary to calculate two recoil-energy resolutions in the previous section). Furthermore, unlike previous CDMS upper limits, the calculation technique I use employs the individual detector efficiencies rather than efficiencies that have been averaged over the Ge and Si detector ensembles.<sup>20</sup> The combined efficiencies and their associated statistical uncertainties are shown for the Z6 3V data in Figure 5.28, with similar figures for the other detectors provided in Appendix F. Note that the 16% energy-scale correction described in Section 5.4.2 has been applied to the efficiencies for the Si detectors, effectively shifting them to the right. This increases their combined thresholds (at 50% efficiency) from  $\sim 1.5\text{--}1.7$  (1.3) keV for Z4 (Z6) to  $\sim 1.7\text{--}1.9$  (1.5) keV. The step in the combined efficiency of the analysis cuts at  $\sim 3\text{--}4$  keV is due to the

---

<sup>20</sup>Although the averaged efficiencies are not used to calculate exclusion limits, they are useful for other applications (such as calculating the WIMP-candidate absolute event rates for the Ge and Si detector ensembles). For this reason, they are included at the end of Appendix F.

### 5.5. COMBINED DETECTION EFFICIENCIES

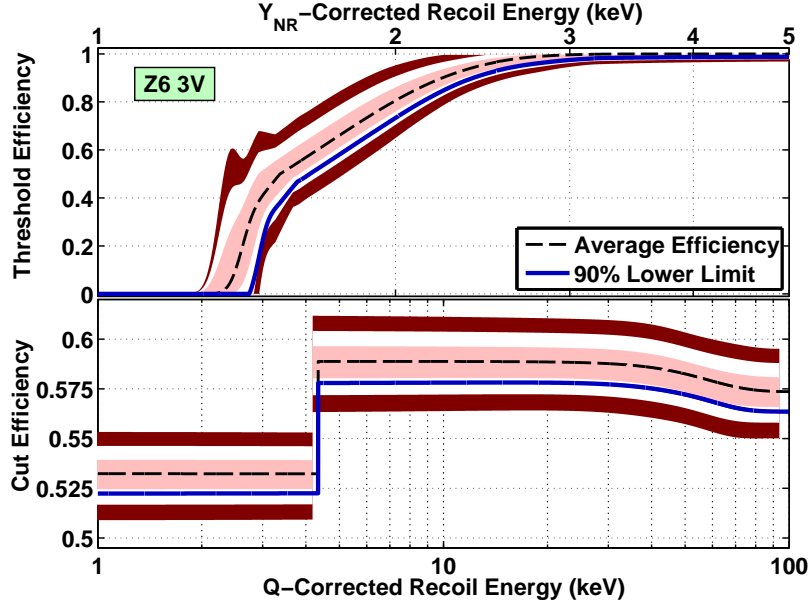


Figure 5.28: Combined detection efficiencies (dashed lines) for the Z6 3V phonon-energy thresholds (top panel) and analysis cuts (bottom panel), and their associated  $1\sigma$  (medium/light-red shaded regions),  $2\sigma$  (light/white shaded regions), and  $3\sigma$  (dark/dark-red shaded regions) statistical uncertainties. The 90% lower-limit efficiencies (thick solid lines) are used to calculate WIMP-nucleon exclusion limits.

larger muon-veto rejection window chosen for the lowest-energy events. In the spirit of being conservative, the 90% confidence-level (statistical) lower-limit efficiencies ( $1.28\sigma$  below the mean) are used to calculate the exclusion limits for this analysis. These reduced-efficiency curves are indicated in Figure 5.28 and Appendix F as well.





# Chapter 6

## Results

In this chapter I present the results of the low-threshold WIMP-search analysis described in the previous chapter. Since the selected events appear to be explained by a variety of background sources, I can claim no evidence for a WIMP signal. Instead, the observed event rates are used to derive upper limits on the spin-independent WIMP-nucleon cross section (under the standard assumptions described in Section 3.3). Systematic effects involving the Galactic escape velocity and the nuclear-recoil energy scale are explored as well.

### 6.1 WIMP Candidates

Following application of the analysis cuts (described in Section 5.3) and phonon-energy software thresholds (described in Section 5.2) to the WIMP-search data, a substantial residual event rate is observed in the low-threshold signal region. Since there is very little Ge (Si) detection efficiency for recoil energies less than 0.5 (1) keV, I restrict the Ge (Si) candidate-event selection to  $Q$ -corrected recoil energies between 0.5 (1) and 100 keV for convenience. Table 6.1 summarizes the WIMP-search (raw) exposure and total number of candidate events for each viable low-threshold detector and bias-voltage run. A total of 1080 Ge and 970 Si candidates are observed. The corresponding differential event rates are shown in Figure 6.1, where the spectra have been

Table 6.1: Summary of the WIMP-search (raw) exposure and total number of observed candidate events for each viable low-threshold detector and bias-voltage run. The Z2 6V exposure is reduced relative to the other Ge detectors as a result of the event-burst cut described in Section 5.3.2. Table adapted from [596].

Detector	3V WIMP Search		6V WIMP Search	
	Exposure (kg d)	Candidates	Exposure (kg d)	Candidates
Z2	15.05	159	4.59	67
Z3	14.50	129	11.33	349
Z4	6.91	130	5.40	125
Z5	14.50	174	11.33	202
Z6	6.91	401	5.40	314

successively corrected by the average efficiencies for the analysis cuts, and then by the average threshold efficiencies. These Ge and Si ensemble-averaged detection efficiencies are described in Appendix F.

## 6.2 Backgrounds

Although the quasi-exponential recoil spectra in Figure 6.1 resemble in shape the differential event rates expected for WIMP-nucleus interactions, it is far more likely that the candidate events are the result of several unrelated background processes. These include electron recoils leaking into the nuclear-recoil band, zero-ionization events,  $^{14}\text{C}$  contamination on the bottom surface of Z6, and nuclear recoils from cosmogenic neutrons. In this section I describe these backgrounds in more detail, and derive rough estimates for the fraction each source contributes to the totals listed in Table 6.1. Although most are easily identified by their ionization-yield versus recoil-energy distributions, the

## 6.2. BACKGROUNDS

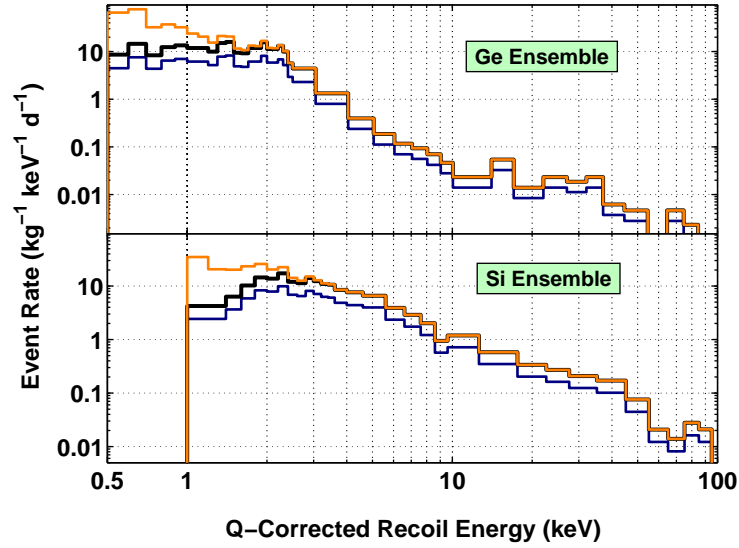


Figure 6.1: Combined WIMP-candidate event rates for the Ge and Si detector ensembles. The uncorrected rates (dark/blue thin histograms) selected directly by the analysis are compared to the efficiency corrected spectra. The latter are first corrected by the exposure-weighted, detector-averaged efficiencies of the analysis cuts (dark/black thick histograms), and then by the similarly averaged threshold efficiencies to obtain the absolute event rates (light/orange thin histograms). Figure adapted from [596].

type of detailed modeling necessary to subtract these backgrounds was not performed for this analysis. Instead, all backgrounds are accepted as viable WIMP candidates for the purpose of calculating upper limits on a WIMP signal. The resulting limits are both conservative and free of the systematic uncertainty that would arise from extrapolating background models into the poorly resolved low-energy region.

The most obvious source of background events is easily identified in a scatter plot of ionization yield versus recoil energy. This is demonstrated for the Z5 6V WIMP candidates in Figure 6.2, with similar figures provided for the other detectors in Appendix G. The well defined 1.29 keV line between 1 and 3 keV in recoil energy clearly accounts for a substantial number of the candidate events.

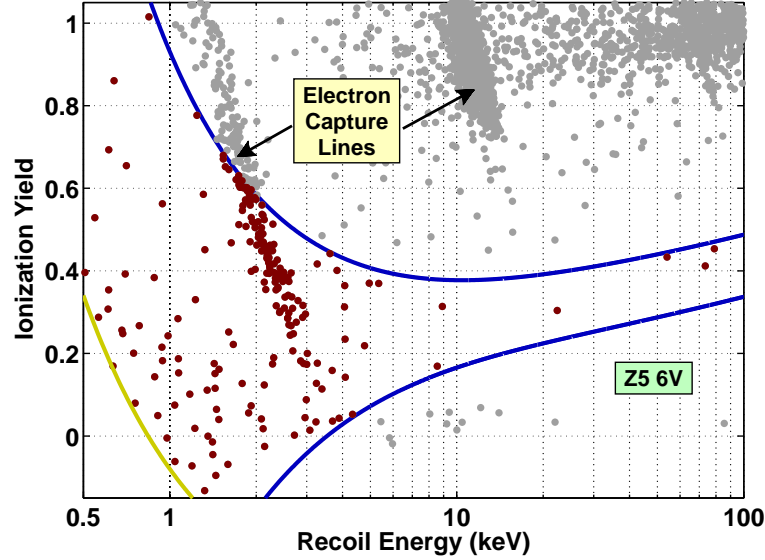


Figure 6.2: Ionization yield versus recoil energy for singles scatters passing the data-quality, fiducial-volume and veto-anticoincidence cuts (all dots) from the Z5 6V WIMP-search data. The signal region is outlined by the  $2\sigma$  nuclear-recoil band (dark/blue lines) and the  $6\sigma$  phonon-energy software threshold (light/yellow line). The 202 events inside this region (dark/red dots) are accepted as WIMP candidates. Note the obvious leakage of the 1.29 keV electron-capture line into the signal region.

On average, the x rays or Auger electrons from decays of  $^{68}\text{Ge}$  and  $^{71}\text{Ge}$  via L-shell electron capture are consistent with being electron recoils and have unit ionization yield. However, the signal-to-noise in both phonons and ionization is relatively poor at these energies, resulting in a low- $Y$  tail of electron recoils that extends well into the nuclear-recoil band.<sup>1</sup> The 1.29 keV line is internal to the Ge detectors. Consequently, the associated low- $Y$  leakage into the signal region is particular to the Ge detectors. The fits to the 1.29 keV peak described in Section 5.4.1 can be used to select these events independently, allowing their

<sup>1</sup>The 1.29 and 10.36 keV electron-recoil lines in Figure 6.2 are tilted with respect to the recoil-energy axis because the numerator and denominator of the ionization-yield expression are anticorrelated (see, *e.g.*, Equation 4.2.5).

## 6.2. BACKGROUNDS

contribution to the candidate event rate to be estimated. Low- $Y$  electron recoils from the 1.29 keV line account for  $\sim 20\%$  of the selected WIMP candidates for the Ge 3V data, and for  $\sim 1/3$  (Z3) to  $1/2$  (Z2 and Z5) for the 6V data. Again, the rate of 1.29 keV events is greater for the 6V data due to the extensive  $^{252}\text{Cf}$  calibrations conducted between the 3V and 6V WIMP searches, resulting in significant neutron activation of  $^{70}\text{Ge}$ .

The most pervasive background for all detectors is a population of events with nearly zero ionization yield. These “zero-charge” events have ionization signals indistinguishable from electronic noise, but are otherwise normal events with phonon energies as large as 100 keV. As is clear in the scatter plot of ionization yield versus recoil energy for the Z4 3V WIMP-search data in Figure 6.3, they are clustered just below the  $2\sigma$  nuclear-recoil band and leak into the low-energy signal region. For Z4 in particular, zero-charge events constitute a majority of the observed WIMP candidates, and are therefore this detector’s limiting background. Scatter plots of  $xy$  position (*a la* Figure 4.9) for the zero-charge events occurring outside the signal region indicate a preference for large radii, corresponding to the edges of the detectors. A possible explanation for their lack of ionization energy is that these events represent surface interactions along the detectors’ cylindrical side walls, where the electric field lines used to drift charge carriers do not span the detectors’ 1 cm thicknesses. Charge carriers produced by an interaction occurring near fields lines that terminate on a detectors’ edge (rather than on an electrode) would not drift across the crystal, resulting in no ionization signals. Any  $^{210}\text{Pb}$  contamination (from radon exposure) on the inside surfaces of the detectors’ copper housings would emit betas that could interact in this manner. This background is new to the CDMS signal region because past analyses have employed an ionization-energy threshold that naturally excludes the phase space zero-charge events populate. Unfortunately, placing a minimum requirement on ionization energy severely limits the efficiency for detecting WIMPs with recoil energies  $\lesssim 5$  keV. Consequently, this background must be accepted as signal for a chance to probe low-mass WIMP parameter space. The recoil-energy distribution for zero-charge events outside the signal region ( $E_R \gtrsim 10$  keV) is approximately exponential. Extrapolating

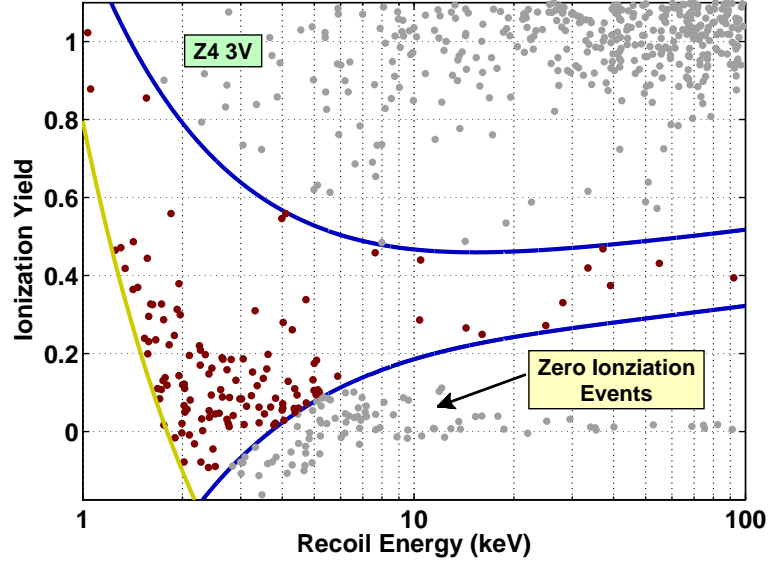


Figure 6.3: Ionization yield versus recoil energy for the Z4 3V WIMP-search data. The different features have the same meaning as in Figure 6.2. Of the 130 WIMP candidates, most appear to be associated with a population of zero-ionization events ( $\langle Y \rangle \simeq 0$ ) that extends from  $\sim 100$  keV down to the software threshold.

this exponential to lower energies indicates a zero-charge event contribution to the total WIMP-candidate event rate of 30–40% (when averaged over all detectors).

A significantly higher event rate was observed for Z6. This is most likely due to exposure to a  $^{14}\text{C}$  source during testing at one of the CDMS test facilities. Because the betas emitted by sources like  $^{14}\text{C}$  are easily collimated, they are useful for testing the detectors' ability to reconstruct  $xy$  position and for gauging the performance of the phonon-based PSD. The source was later discovered to have faulty encapsulation, resulting in an accidental contamination of the bottom side of Z6 with a low level of  $^{14}\text{C}$ . Again, this is the reason Z6 was placed at the bottom of the tower with its phonon side facing away from Z5. The betas emitted during  $^{14}\text{C}$  decays have an average (maxi-

## 6.2. BACKGROUNDS

mum) energy  $\sim 50$  (156) keV, and therefore span the energy range of interest for WIMP-nucleus interactions. Unfortunately, betas of these energies are not very penetrating, and will tend to interact entirely within a detector's  $\sim 10\ \mu\text{m}$  dead layer. As discussed in Section 4.2.2, charge carriers produced in a detector's dead layer have a tendency to recombine before they can be drifted across the crystal. Electron-recoil surface events are therefore characterized by reduced ionization signals that can lead to misidentification as nuclear recoils. As is clear from the scatter plot of ionization yield versus recoil energy for the Z6 6V WIMP-search data in Figure 6.4, these  $^{14}\text{C}$  betas populate the gap between the electron- and nuclear-recoil bands, with a substantial number leaking into the signal region. Similar to the zero-charge events, these surface events exhibit an exponential recoil-energy distribution that can be extrapolated into the signal region to estimate their contribution to the candidate event rate. When averaged over the two Si detectors and both bias-voltage runs,  $^{14}\text{C}$  betas account for  $\sim 40\%$  of the Si detectors' WIMP candidates.

General leakage of electron recoils into the nuclear-recoil band is a component of each detector's WIMP-candidate event rate. However, unlike the Z6  $^{14}\text{C}$  background and the Ge events associated with the 1.29 keV peak, the source is usually Compton scatters of photons. Finite resolution causes the  $Y$ -based electron-recoil discrimination to gradually break down with decreasing recoil energy, until the  $2\sigma$  electron- and nuclear-recoil bands eventually merge at a "crossover energy" that varies from  $\sim 3$  keV for the Z5 3V bands to  $\sim 7$  keV for the Z6 6V bands. This crossover energy is also correlated with the magnitude of the bias voltage; the 6V crossover energies are larger than their 3V counterparts, a consequence of degraded recoil-energy resolution due to the relative difference in Neganov-Luke phonon production. As a result, although the phonon-energy thresholds are generally lower for the 6V data (thanks to drift-heat amplification), the detectors' 3V signal regions exhibit lower backgrounds due to superior electron-recoil discrimination. Overall, Ge (Si) ZIPs perform better as low-threshold detectors when operated with a  $-3$  ( $-4$ ) V bias voltage. Above the crossover energies, data from  $^{60}\text{Co}$  calibrations can be used to estimate the leakage of electron recoils into the signal region. Scaling



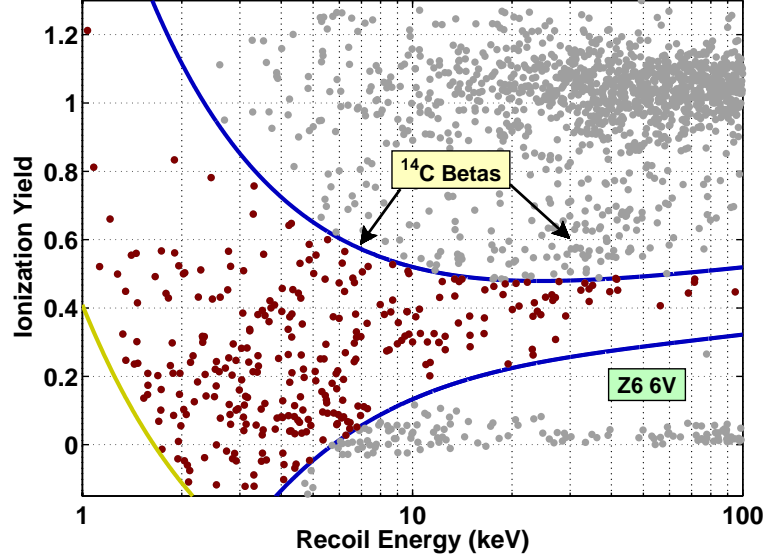


Figure 6.4: Ionization yield versus recoil energy for the Z6 6V WIMP-search data. The different features have the same meaning as in Figure 6.2. Of the 314 WIMP candidates, over half appear to be associated with the  $^{14}\text{C}$  contamination on the bottom surface of the detector.

these leakage rates to an exposure equivalent to the WIMP-search data results in an estimate of only a few events per detector. Below the crossover energies it is much more difficult to gauge the Compton electron-recoil background's contribution to the number of WIMP candidates. A rough scaling yields a guesstimate of 10–20%.

The small number of relatively high-energy signal events can be attributed to the cosmogenic-neutron background measured previously in [694]. A relatively high flux of cosmic-ray muons penetrated the SUF's modest overburden. As explained in Section 3.3.3, muons can occasionally break apart nuclei in materials exterior to the shielding layers (primarily in the surrounding rock), ejecting neutrons of sufficient energy to punch through the muon veto undetected and generate secondary neutrons in the passive shielding layers capable of mimicking WIMP-nucleus interactions. The expected rate of

### 6.3. EXCLUSION LIMITS

veto-anticoincident, single-scatter neutron interactions was simulated for the Ge detectors and the Run 21 shielding configuration for a previous analysis (see, *e.g.*, the dashed curve in Figure 2 in [694]). The absolute rate peaks at  $\sim 0.1$  events per kg keV d at 0.5 keV, and decays in an approximately exponential fashion to  $\sim 0.003$  events per kg keV d at 100 keV. If the average detection efficiency for the Ge detector ensemble (shown in Figure F.10) is applied to this simulated neutron event rate, and then scaled by the total Ge exposure ( $\sim 71$  kg d),  $\sim 66$  (or 6%) of the 1080 Ge WIMP candidates can be attributed to cosmogenic neutrons. Scaling this prediction by the relative neutron cross sections for Si and Ge targets (as well as by the relative Si and Ge efficiencies and exposures) yields an estimate of  $\sim 19$  neutrons in the Si signal region.

Table 6.2 summarizes the various contributions to the WIMP-candidate event rates for the Ge and Si detector ensembles. An accurate accounting of each source of background is difficult due to the highly confused low-energy region, particularly for the zero-charge and electron-recoil leakage backgrounds. However, a refined low-threshold analysis of 8 Ge ZIPs operated at the CDMS deep-site installation roughly confirms the observations reported here [757], supporting the conclusion that the observed event rates are not due to WIMP-nucleus interactions. Nevertheless, I include an “other” category in Table 6.2, allowing for the possibility that a fraction of the observed event rate could be due to WIMPs or unidentified sources of background. It is much more likely, however, that these unaccounted for events represent systematic uncertainty in the background estimates described above.

## 6.3 Exclusion Limits

Although the above discussion of background sources does not preclude the possibility that a fraction of the observed event rate is due to WIMP-nucleus interactions, I claim no evidence for a WIMP signal with this analysis. Without reliable models for each source of background, attempting to extract a WIMP signal from the observed event rate would be misleading at best. Instead, the data are used to calculate upper limits on the spin-independent

Table 6.2: Approximate tally of the percentages contributed by various sources of background to the 1080 Ge and 970 Si WIMP-candidate events observed in the low-threshold signal region. The “other” category accounts for the fraction of events not attributed to the background sources as described in the main text. Table adapted from [596].

	Ge Ensemble (%)	Si Ensemble (%)
1.29 keV Electron-Capture Line	32	0
Zero-Charge Events	30–40	30–40
Z6’s $^{14}\text{C}$ Contamination	0	40
General Electron-Recoil Leakage	10–20	10–20
Cosmogenic Neutrons	6	2
Other	2–22	0–18

WIMP-nucleon elastic scattering cross section (as a function of WIMP mass) under the conservative assumption that the candidate events may constitute a WIMP signal. The CDMS convention is to calculate separate exclusion limits for our two target materials. However, the relatively high-background event rates observed here lend themselves naturally to a novel exclusion-limit technique that allows data from different target materials to be combined into a single limit. The primary result of this thesis is therefore a combined Ge and Si exclusion limit based upon all 2050 signal events. Since the Ge data dominate the sensitivity of the main result, a secondary Si-only limit is also calculated. Although substantially weaker than the combined limit, the Si-only limit provides a useful indication of the Si detectors’ low-mass WIMP sensitivity.

There are a variety of statistical techniques available for calculating WIMP-nucleon exclusion limits, but all share at least one thing in common; observed event rates are compared to the event rates expected from a hypothetical WIMP model. To facilitate comparison with results from other direct-detection exper-

### 6.3. EXCLUSION LIMITS

iments, I use the standard framework and Galactic halo outlined in Section 3.3. The parameters for this model are summarized in Table 3.1. As discussed in Section 3.3, there is a fair amount of uncertainty regarding a few of these parameters, particularly the local WIMP density and the Galactic escape velocity. Varying the former causes exclusion limits to shift either up or down in a manner that is common to all direct-detection results (including potential signal regions). Consequently, while uncertainty in  $\rho_0$  might affect comparison of the limits calculated here to LSP searches at particle colliders or indirect-detection results, it has no bearing upon their sensitivity relative to other direct-detection experiments. Since my goal is to test the parameter space favored by Hooper *et al.*'s combined analysis of the DAMA/LIBRA and CoGeNT data, uncertainty in  $\rho_0$  is of no consequence. The effect of varying the Galactic escape velocity is explored in Section 6.4.1.

Steve Yellin's optimum interval method [599] has become one of the most widely adopted statistical techniques for calculating exclusion limits in the direct-detection community. The limits calculated here use a version that has been extended (by Yellin) to accommodate high statistics [768]. In short, the upper limit is determined from the energy interval in which the number of observed events is particularly low relative to the number expected from the hypothetical WIMP model. The intervals are constructed from the candidate event energies (as well as the signal region's minimum and maximum energies); all possible combinations (of any two) are considered. In general, the optimum interval varies as a function of WIMP mass, and may include only a fraction of the signal region's full range of energies. In the limit of zero signal events, the calculated limit reduces to a classical one-sided Poisson confidence interval. When setting a limit from a nonzero event rate, to compensate for selecting what is essentially the lowest-background portion of the signal region (relative to the expected signal—*i.e.*, the interval that sets the strongest limit), an appropriate statistical penalty is applied. The end result is an unbiased, frequentist 90% confidence-level upper limit. The widespread appeal of Yellin's optimum interval method is clear; it is especially effective at discriminating against backgrounds that are distributed differently from the expected signal.

## CHAPTER 6. RESULTS

In order to calculate a single limit from data obtained with multiple detectors and during two different operational periods, the data from the individual Ge and Si detectors have to be appropriately combined. The most common method involves representing the detector ensemble by a single averaged detector, with an exposure-weighted average efficiency. The efficiencies shown in Figures F.10 and F.11, for example, are the detection efficiencies one would use to calculate “averaged” limits for this analysis. Averaged limits make use of the entire WIMP-search exposure, and derive optimum intervals from a co-mingled set of candidate event energies. Figure 6.1 shows the Ge (Si) WIMP-candidate event rate that results from co-mingling the Ge (Si) detectors’ event energies. The averaging method is appropriate when the individual detectors within an ensemble have approximately equal sensitivity to WIMP-nucleus interactions. One instance in which this is true is a background-free analysis. In this case the sensitivity grows linearly with exposure, and the strongest limits are obtained by simply merging the individual detectors into a single averaged detector with a greater exposure.

If there are significant variations among the individual event rates within a detector ensemble (as is the case here), application of the averaging technique can result in overly conservative limits. Co-mingling data from higher-background detectors with those from lower-background detectors effectively pollutes the most sensitive energy intervals with a disproportionate number of events. Recently, Yellin proposed six alternative methods for combining data from different detectors (and even different experiments) [769]. Among these, two are particularly well suited in the case of high backgrounds: the “serialization” and “minimum limit” methods. Both methods sacrifice WIMP-search exposure in order to isolate the most sensitive individual-detector energy intervals. This is a trade-off that makes sense when the detectors’ signal regions are background-limited. The minimum limit method involves directly choosing the detector that yields the minimum (and therefore strongest) limit, and requires an additional statistical penalty to compensate for the associated bias. In the serialization method, energy intervals are separately pre-

### 6.3. EXCLUSION LIMITS

pared for each detector (see below), and then concatenated in an arbitrary order. Optimum intervals are selected from this serialized (rather than comingled) set of energy intervals. Because the limits are chosen from the full set of energy intervals, no additional statistical penalty is required as with the minimum limit method. This is the method I adopt here. In order to avoid possible bias, the concatenation order was selected before the effect of the order was known. Since the 3V and 6V signal regions, associated detection efficiencies and exposures, and WIMP candidates were derived separately, the two operational periods for each detector can be thought of as two independent detectors. Consequently, each detector appears twice in the serialization, once as a 3V detector and once as a 6V detector. I chose to place the 3V intervals before the 6V intervals, and then to order them according to the detectors' relative positions within the tower (from top to bottom). Provided there are no optimum intervals that span multiple detectors, the serialization order will not affect the result. This turned out to be the case for the limits calculated here. Additionally, since the energy intervals are prepared separately, different target materials can be included in the serialization, providing a natural method for combining the Ge and Si data.

The preparation of each detector's energy intervals involves several steps that must occur in a particular order to properly include the effect of nonzero energy resolution. Upper limits were calculated for 75 WIMP masses between 1 and  $100 \text{ GeV}/c^2$ . At each WIMP mass, and for each detector and bias voltage, the following recipe was used. The guiding principle is to ensure that the expected WIMP spectrum—to which the observed event rates are compared—is expressed in terms of the proper recoil-energy estimator before application of the two types of detection efficiency discussed in Section 5.5.

1. The halo model was used to predict the differential WIMP-nucleon scattering rate in terms of an ideal, perfect-resolution recoil energy (as in Figure 3.11, *e.g.*).

2. This ideal spectrum was then smeared (via a Gaussian convolution according to Equation 3.3.36) with the  $Y_{\text{NR}}$ -corrected recoil-energy resolution ( $\sigma_E$  in Equation 3.3.36) listed in Table 5.6. Recall that the hardware and software thresholds depend solely on the phonon signal. Consequently, the expected spectrum should include noise from only the phonon channel when the combined threshold efficiency is applied.
3. The expected spectrum, now as a function of  $Y_{\text{NR}}$ -corrected recoil energy, was reduced (multiplied) by the threshold efficiency (top panel of Figure 5.28, *e.g.*).
4. Ionization noise was added to the threshold-reduced result of the previous step by smearing it with a Gaussian resolution equal to the quadrature difference between the  $Q$ - and  $Y_{\text{NR}}$ -corrected recoil-energy resolutions listed in Table 5.6. This casts the expected WIMP spectrum in terms of the same recoil-energy estimator as the efficiency of the analysis cuts.
5. The threshold-reduced spectrum in terms of  $Q$ -corrected recoil energy is further reduced by the combined efficiency of the analysis cuts (lower panel in Figure 5.28, *e.g.*).
6. The doubly smeared and efficiency-reduced spectrum is scaled to the appropriate exposure listed in Table 6.1 to obtain an absolute differential rate in units of events per keV.
7. The detector's candidate event energies (in terms of  $Q$ -corrected recoil energy) are sorted from lowest to highest, and the spectrum from the previous step is integrated from threshold—0.5 and 1 keV for the Ge and Si detectors, respectively—to each event energy to form a cumulative distribution. This cumulative distribution reflects the likelihood of a WIMP interaction as a function of deposited energy, but has not yet been normalized to form a proper probability.

In the case of the combined Ge and Si (Si-only) limit, the resulting 10 (4) individual cumulative distributions were serialized according to the order de-

### 6.3. EXCLUSION LIMITS

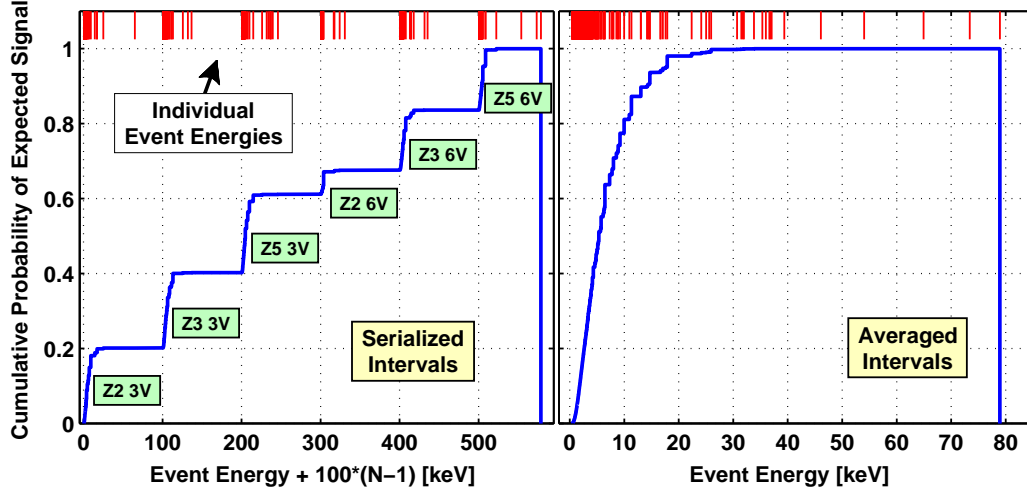


Figure 6.5: Comparison of the serialization (left panel) and averaging (right panel) methods for combining the Ge detectors' WIMP search data. For the serialization method, the cumulative probability distributions used by the optimum interval method to calculate exclusion limits are derived separately for each detector and bias voltage, and then concatenated as shown. The  $x$  axis reflects recoil energy that is offset by the detector's position within the serialization order ( $N = 1-6$ ). For the averaging method, a single cumulative distribution is constructed from the co-mingled set of event energies. In both cases, the individual event energies are indicated by the short vertical lines near the top of each plot.

scribed above. Each distribution receives a vertical offset that depends on its order in the serialization. The first distribution starts at zero, while the second starts at the maximum of the first (both in energy and counts). This continues down the order until the last detector, whose cumulative distribution is offset vertically by the total number of counts expected in all previous detectors. The resulting serialized distribution is then normalized to the total number of counts expected in all detectors, yielding a cumulative probability distribution. This is demonstrated for the 6 sets of Ge intervals in the left panel of Figure 6.5 for a  $20 \text{ GeV}/c^2$  WIMP. The Si intervals were not included in order to allow a fair comparison to the equivalent distribution one



would construct to calculate an averaged limit (right panel), for which the two target materials cannot be combined in such a direct manner. In any case, Figure 6.5 clearly illustrates the difference between the serialization method for combining detector data and the more conventional averaging method. Note the relative difference in the density of the lowest-energy intervals between the two methods. The serialization method’s more sparsely distributed low-energy intervals yield stronger exclusion limits for WIMP masses less than  $\sim 8 \text{ GeV}/c^2$ .

The combined Ge and Si WIMP-search data, as well as the Si-only data, were prepared in the manner described above, and the optimum interval method was used to derive corresponding 90% confidence-level upper limits on the spin-independent WIMP-nucleon elastic scattering cross section as a function of WIMP mass. These limits, which represent the final results of the low-threshold WIMP search featured in this thesis, are shown in Figure 6.6. The combined Ge and Si limit partially excludes the parameter space favored by Hooper *et al.*’s joint analysis of the DAMA/LIBRA and CoGeNT data. Exclusion limits from direct-detection experiments with competitive low-mass WIMP sensitivity are included in Figure 6.6 for comparison. The combined Ge and Si limit excludes new parameter space for WIMP masses between 3 and  $4 \text{ GeV}/c^2$ . The extension of these limits down to  $1 \text{ GeV}/c^2$  is explored in Section 6.4.1 below.

After all the effort dedicated to preserving as much low-threshold detection efficiency for as many detectors as possible, it is interesting to review which detectors are responsible for the limits shown in Figure 6.6. For each WIMP mass, the upper limits were (effectively) determined from an energy interval from a single detector and bias-voltage run. The detectors from which the limit-setting intervals were chosen are indicated in Figure 6.7 for WIMP masses between 2 and  $100 \text{ GeV}/c^2$ . The combined Ge and Si limit was derived entirely from Ge energy intervals. Consequently, a similar limit based on the Ge data only is virtually indistinguishable from the main result. The only Ge intervals that do not contribute are those for the Z3 6V data. This is likely a consequence of this detector’s relatively large 6V  $Q$ -corrected recoil-

### 6.3. EXCLUSION LIMITS

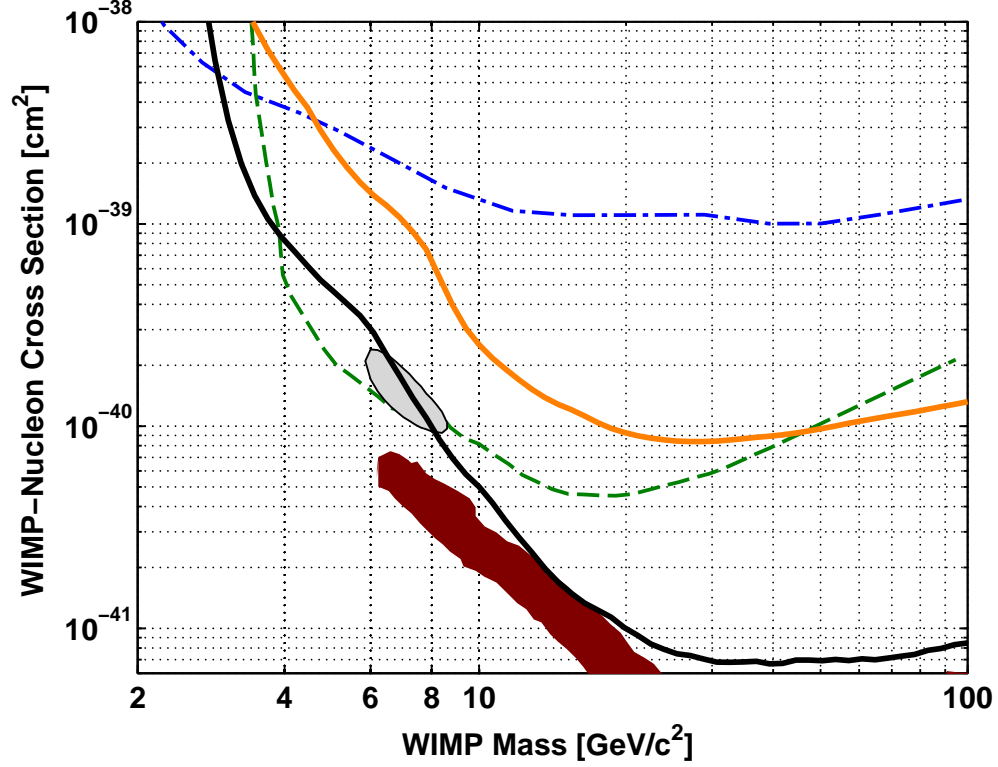


Figure 6.6: Comparison of 90% confidence-level upper limits from the combined Ge and Si (dark/black solid line, main result of this thesis) and Si-only data (light/orange solid line), with the potential signal region found by Hooper *et al.*'s joint analysis of the DAMA/LIBRA and CoGeNT data [652] (99% C.L., light/gray shaded region). Also shown are 90% C.L. exclusion limits from the CoGeNT [670] (dashed line) and CRESST [597] (dash-dotted line) experiments, as well as LNM theoretical predictions [468] (90% C.L., dark/red shaded region). Note that the CRESST and CoGeNT limits, as well as the Hooper *et al.* region, use larger Galactic escape velocities (600–650 km/s) than the 544 km/s value used in this thesis.

energy resolution, which resulted in worse electron-recoil discrimination at low energy, and a correspondingly larger observed event rate. Except for the very lowest-mass WIMPs considered, the main result for WIMP masses less than

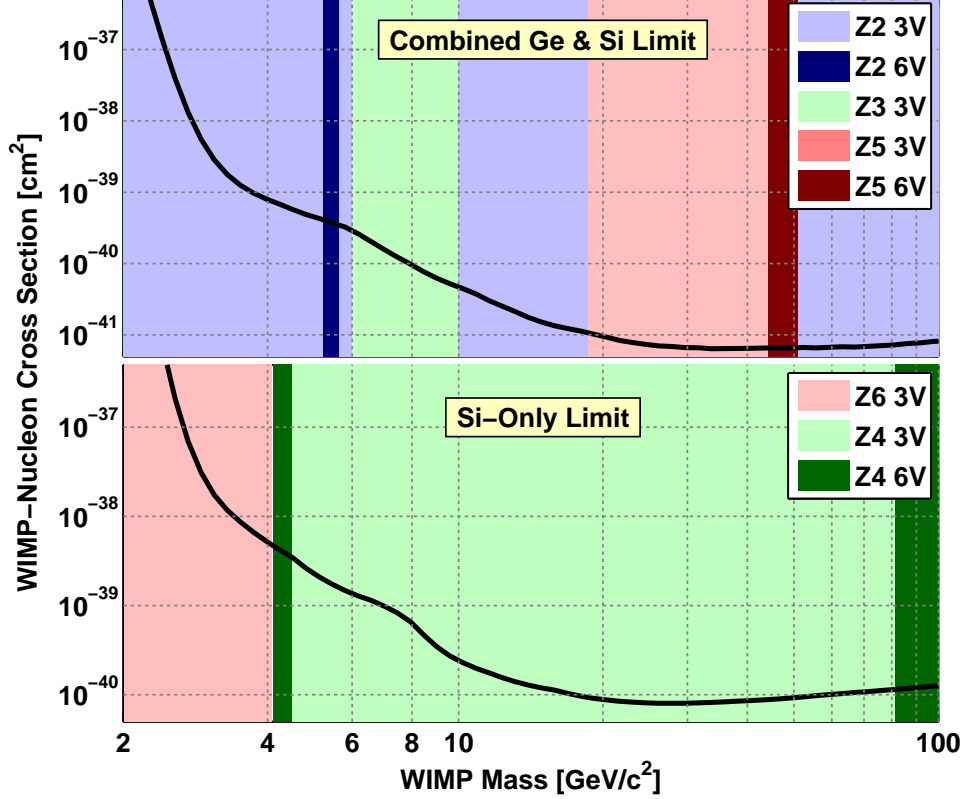


Figure 6.7: The combined Ge and Si (top panel) and Si-only (bottom panel) exclusion limits (solid lines) divided according to the limit-setting detector and bias-voltage run. For each shaded region, data from the indicated detector were used to determine the optimum intervals from which the limit over the corresponding range of WIMP masses was derived.

$2 \text{ GeV}/c^2$  was determined from Z2 3V intervals. For  $1\text{--}1.1 \text{ GeV}/c^2$  WIMPs, the limits were derived from Z5 3V intervals. Intervals from each Si detector and bias-voltage run contributed to the Si-only limit. Below  $2 \text{ GeV}/c^2$ , the Z6 3V (6V) intervals determined the sensitivity for WIMP masses between 1 and 1.3 (1.3 and 2)  $\text{GeV}/c^2$ . Despite Z4's substantially lower backgrounds, Z6's lower thresholds gave it superior sensitivity to the lowest-mass WIMPs considered.

## 6.4 Systematic Studies

Considering the variety of conservative choices made throughout this low-threshold analysis, it is difficult to imagine how the limits shown in Figure 6.6 might be overstated. If anything, the main result is overly conservative due to (in particular) the use of 90% lower-limit detection efficiencies (see Section 5.5), and the choice to include all observed signal events in the exclusion-limit calculations despite plausible background estimates that account for nearly the full event rate. Nevertheless, a number of systematic studies were performed to test the virility of the two limits. One of these I have already reviewed in Section 5.4.1, where the electron-recoil energy scale for the Ge detectors was found to be conservative. The exclusion limits were also tested for sensitivity to the lowest-energy trigger efficiency, and to the inclusion of the effect of near-threshold energy resolution. These studies, described in more detail in [596], revealed systematic uncertainties at the level of a few percent for  $1\text{ GeV}/c^2$  WIMPs, and no discernable effect for WIMP masses  $\gtrsim 2\text{ GeV}/c^2$ .

In this section I describe two additional systematic studies: variation of the WIMP model’s Galactic escape velocity, and a more detailed derivation of the ZIP-detector nuclear-recoil energy scale. The former has a pronounced effect on the exclusion limits for low WIMP masses, while the latter reveals yet another conservative tendency of the results presented above.

### 6.4.1 Galactic Escape Velocity

The low-threshold exclusion limits are shown for WIMP masses down to  $1\text{ GeV}/c^2$  in Figure 6.8. As is clear from the figure, the lowest-mass WIMP sensitivity is heavily dependent on the WIMP model’s assumed value for the Galactic escape velocity. Smith *et al.*’s derivation of the escape velocity using halo stars from the RAVE survey (discussed in Section 2.2.1) suggests values as low as  $\sim 500\text{ km/s}$  and as high as  $\sim 600\text{ km/s}$  at the 90% confidence level. Prior to this study, most direct-detection experiments either employed the  $600\text{ km/s}$

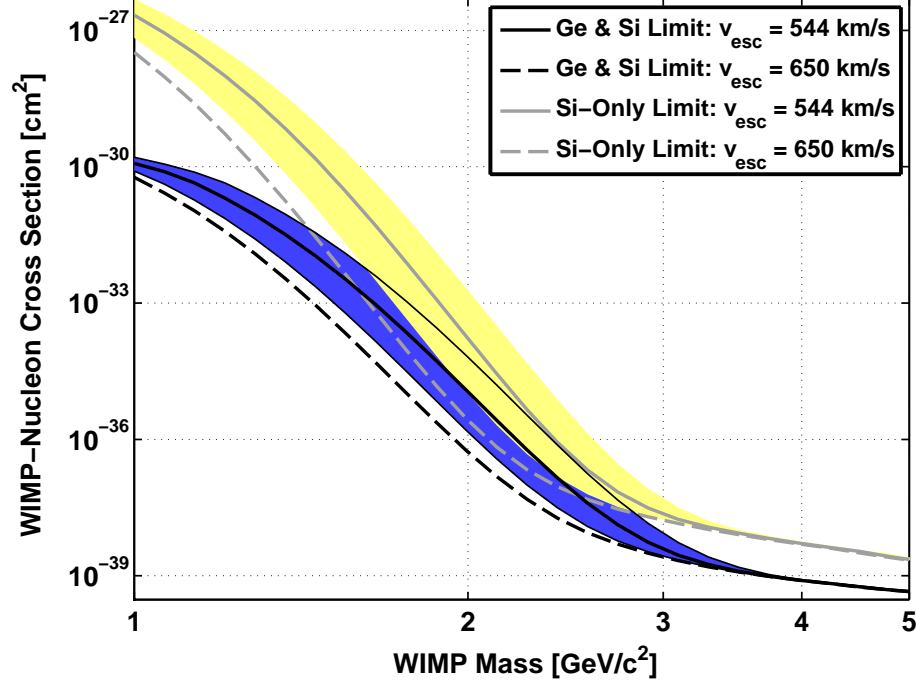


Figure 6.8: The low-threshold exclusion limits (solid lines) for WIMP masses down to  $1 \text{ GeV}/c^2$ , compared to equivalent limits for different values of the Galactic escape velocity. The shaded regions (dark/blue for Ge & Si data, and light/yellow for Si-only data) correspond to the 90% confidence interval derived by Smith *et al.* in [183]:  $498 < v_{\text{esc}} < 608 \text{ km/s}$ . Figure adapted from [596].

value indicated in [588] (*e.g.*, the CoGeNT limit and Hooper *et al.* signal region in Figure 6.6), or the larger,  $650 \text{ km/s}$  value adopted in [374] (*e.g.*, the CRESST limit in Figure 6.6). It is therefore instructive to explore the dependence of the low-threshold limits on escape velocity. Additional exclusion limits were calculated that correspond to Smith *et al.*'s 90% confidence interval, and to  $v_{\text{esc}} = 650 \text{ km/s}$ . Generally, WIMP sensitivity and the escape velocity are correlated, with the effect particularly pronounced for WIMP masses  $\lesssim 4 \text{ GeV}/c^2$ . The effect is nearly maximal for  $2 \text{ GeV}/c^2$  WIMPs, with uncertainty in the main (Si-only) result spanning over 1 (2) order(s) of magnitude. For WIMP

## 6.4. SYSTEMATIC STUDIES

masses larger than  $4 \text{ GeV}/c^2$  there is virtually no visible difference in the main result, while the Si-only limit is as much as 20% weaker between 6 and  $10 \text{ GeV}/c^2$  for the lowest escape velocity considered (relative to  $v_{\text{esc}} = 650 \text{ km/s}$ ). Consequently, the use of  $v_{\text{esc}} = 544 \text{ km/s}$  for the low-threshold limits in Figure 6.6 is slightly conservative relative to the larger values adopted by the other results in that figure.

### 6.4.2 Nuclear-Recoil Energy Scale Revisited

While the test of the nuclear-recoil energy scale in Section 5.4.2 provides a rough measure of the ZIP-detector recoil-energy response, it is based upon a slightly faulty assumption. As demonstrated by the numerical calculations in Appendix E, the differential event rates expected for  $^{252}\text{Cf}$  neutrons scattering from Ge and Si targets are not exactly exponential. In this section I explore a more statistically rigorous method for testing the nuclear-recoil energy scale using a chi-squared test. The Ge and Si energy scales are systematically varied via a two-parameter energy-scale correction, and chi-squared statistics are constructed from the simulated and observed spectra as a function of the correction parameters. The resulting statistics exhibit minima that are indicative of the best-fit energy scales for the Ge and Si detector ensembles. Corresponding low-threshold exclusion limits are calculated, and the effect of uncertainty in the best-fit correction parameters is explored.

The simulated and observed data sets used for this study are identical to those described in Section 5.4.2. However, to ensure a well behaved  $\chi^2$  statistic, the recoil-energy binning was modified to include more (fewer) low-energy (high-energy) bins. This spreads the statistics a bit more evenly among the bins, and ensures a minimum number of counts per bin. Consequently, the corresponding Poisson counting errors are well defined (even for the lowest-rate bins) and reasonably well balanced as a function of recoil energy. The resulting spectra for the Ge and Si detector ensembles are shown in Figure 6.9. Note that the integrated event rates for the observed spectra have been scaled to the integrated event rates of the simulated spectra to facilitate comparison.

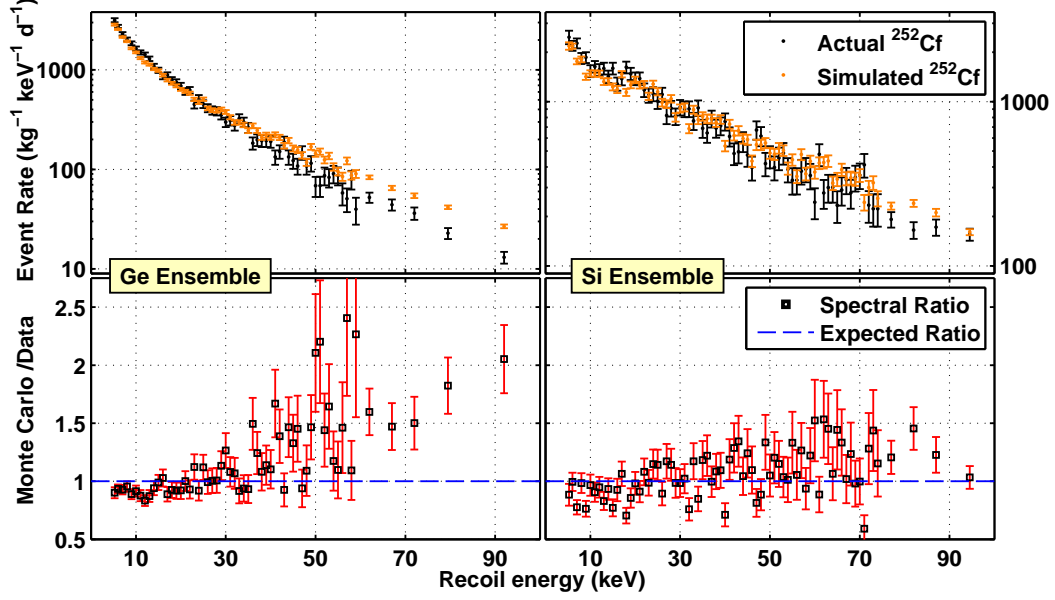


Figure 6.9: *Upper panels:* Spectral comparison of simulated (light/orange error bars) and observed (dark/black error bars) <sup>252</sup>Cf nuclear recoils, where the latter reflects the original calibration of the Run 21 energy scale. *Lower panels:* Ratios of the simulated to observed spectra (error bars), compared to the ratio expected (dashed lines) if the Ge and Si energy scales were in agreement with the Monte Carlo simulations.

It is clear from the spectral ratios shown in the lower panels that there are significant discrepancies for both detector ensembles.

To better quantify these discrepancies, a  $\chi^2$  statistic was constructed from the event rates, effectively reducing the two spectra for each detector ensemble to a single number. The statistic follows the following standard formula for the  $\chi^2$  distribution:

$$\chi^2 = \sum_{i=1}^k \left( \frac{X_i - \mu_i}{\sigma_i} \right)^2, \quad (6.4.1)$$

where the sum runs over the recoil-energy bins (*e.g.*,  $i = 1$  corresponds to the first bin from 5 to 5.5 keV),  $k = 60$  (74) is the total number of Ge (Si) bins,  $\mu_i$  is the simulated event rate in the  $i^{\text{th}}$  bin,  $X_i$  is the observed event rate in the

#### 6.4. SYSTEMATIC STUDIES

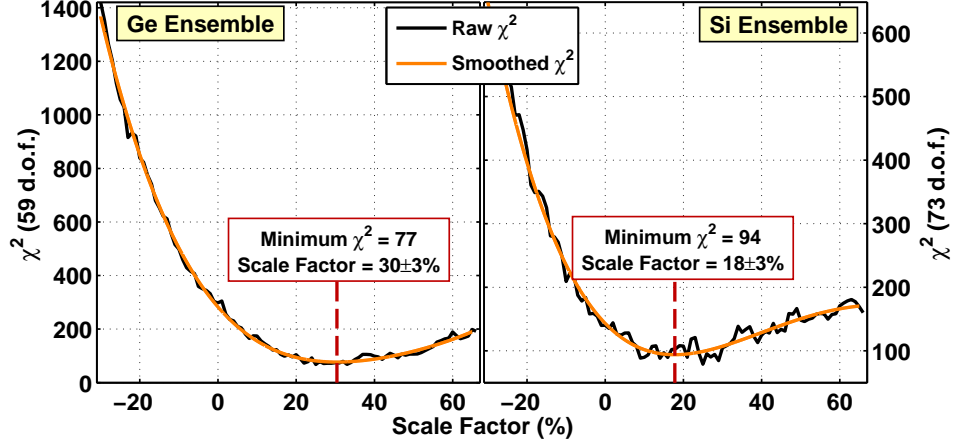


Figure 6.10:  $\chi^2$  statistic versus scaling factor for the one-parameter energy-scale correction described in the main text. The minima correspond to scale factors that improve the agreement between the observed and simulated  $^{252}\text{Cf}$  spectra.

$i^{\text{th}}$  bin, and  $\sigma_i$  is the quadrature sum of the observed and simulated Poisson errors in the  $i^{\text{th}}$  bin. Similar to Figure 6.9, the observed spectra are scaled by the ratio of the simulated to observed integrated event rates prior to formation of the  $\chi^2$  statistic. The better the agreement between the two spectra, the smaller its value will be. The value of the  $\chi^2$  statistic for the Ge (Si) spectra in Figure 6.9 is  $\sim 300$  ( $150$ ). This number is commonly expressed relative to the number of degrees of freedom (d.o.f.)—the total number of bins in this case—and is often referred to as the reduced chi-square:  $\chi_0^2 = \chi^2/\text{d.o.f.}$  The  $\chi_0^2$  value is a useful measure of goodness-of-fit; when it deviates significantly from unity (as is the case here), it is usually an indication that something is amiss.

To reconcile the observed and simulated event rates, the nuclear-recoil energy scale for the former was systematically adjusted. This was first done for a one-parameter correction, where the individual recoil energies from which the observed spectra are constructed were scaled by a simple multiplicative factor ranging from 0.7 to 1.65 in steps of 0.01. Correction factors less than 1



correspond to a downward scaling (by as much as 30%), while factors greater than 1 correspond to an upward scaling (by as much as 65%). For each correction, the rescaled recoil energies were histogrammed for each detector and bias-voltage run. Similar to the procedure outlined in Section 5.4.2, these individual spectra were corrected for their detector’s selection-cut efficiencies, and then combined to yield a single distribution for each detector ensemble. Note that the selection-cut efficiencies were also rescaled to reflect the adjusted energy scale. A  $\chi^2$  statistic was constructed as a function of the scaling parameter by comparing the corrected spectra to the simulated spectra using the method described above. The results for the Ge and Si detector ensembles are shown in Figure 6.10, and exhibit minima that correspond to respective upward scalings of  $30\pm3\%$  and  $18\pm3\%$ . The latter is consistent with the 16% energy-scale correction derived for the Si detectors in Section 5.4.2. The best-fit corrections, characterized by  $\chi_0^2 \simeq 1.3$ , imply that the original Run 21 energy scale is nonconservative with respect to low-mass WIMP sensitivity.

Although the results from Section 5.4.2 and Figure 6.10 are consistent for the Si ensemble, they are contradictory for the Ge ensemble. Recall from the comparison of best-fit exponential decay constants in Section 5.4.2 that the Ge energy scale appears to overestimate low-value recoil energies, and underestimate high-value recoil energies. If this is truly the case, the one-parameter correction described above is not sufficiently flexible to simultaneously correct low- and high-value recoil energies. In order to gauge the necessity for a more complicated two-parameter energy-scale correction, the above analysis was repeated for a series of successively smaller energy ranges. For the Ge (Si) ensemble, the range of recoil energies considered was systematically reduced from 5–100 keV to 5–10 (5–20) keV. The resulting dependence of the best-fit scale factor on recoil energy is shown in Figure 6.11. Although the evidence is a bit marginal due to poor statistics, the trends for *both* detector ensembles indicate that an energy-dependent energy-scale correction might improve the agreement between the simulated and observed  $^{252}\text{Cf}$  spectra.

#### 6.4. SYSTEMATIC STUDIES

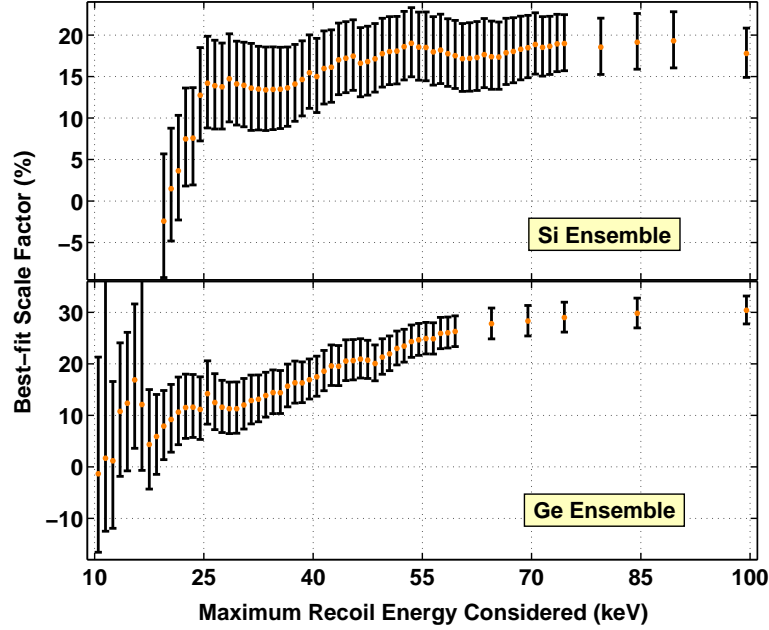


Figure 6.11: For the one-parameter energy-scale correction described in the main text, the best-fit scale factor is plotted as a function of the maximum recoil energy considered. As the recoil-energy range shrinks to include only the lowest-energy events, the scale factor required to match the observed and simulated event rates decreases, supporting the case for a two-parameter (energy-dependent) energy-scale correction.

The trends in Figure 6.11 motivated the following two-parameter energy-scale correction:

$$E_R \longrightarrow cM \left( \frac{E_R}{M} \right)^b, \quad (6.4.2)$$

where  $c$  is a scale factor similar to the one-parameter correction,  $b$  is an exponent that stretches (squeezes) the energy scale for values greater (less) than 1, and  $M = \sqrt{5 \times 100} \simeq 22$  keV is the geometric mean of the recoil-energy range considered. Including the geometric mean in this manner helps prevent correlations between  $c$  and  $b$ . Note that if  $c$  and  $b$  both equal 1, the original energy scale is returned. Additionally, Equation 6.4.2 reduces to the one-parameter correction if  $b$  is held constant and equal to one. The same procedure outlined

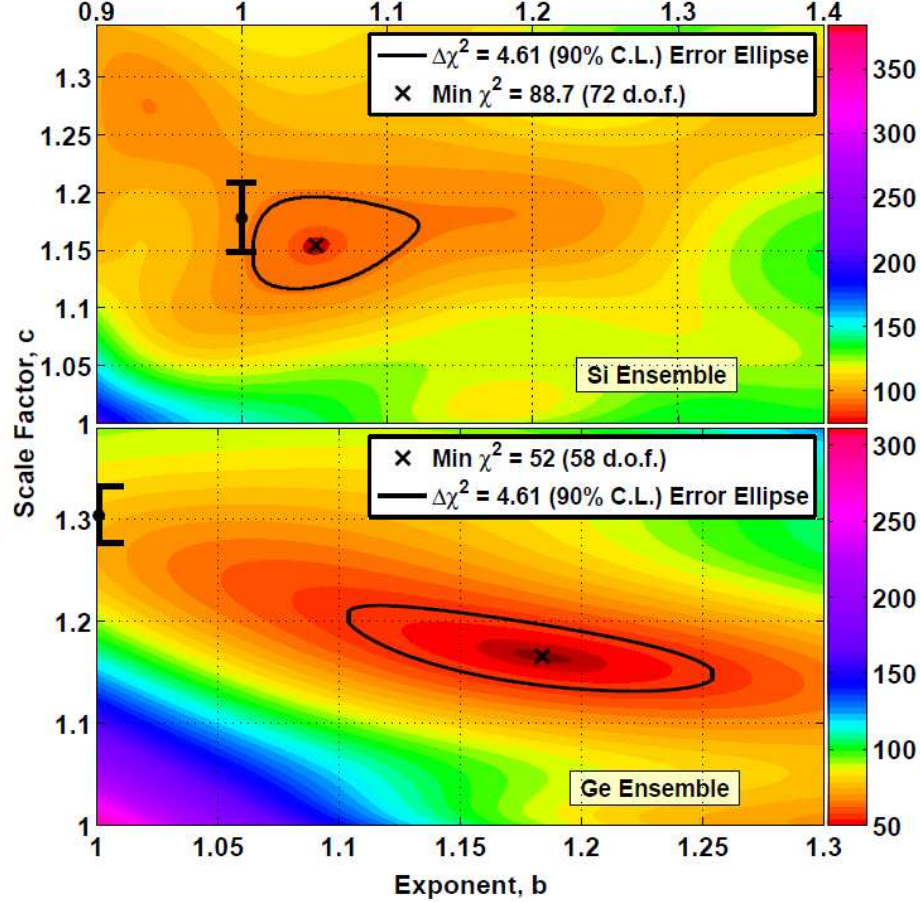


Figure 6.12:  $\chi^2$  statistic (values indicated by color) versus scaling factor and exponent for the two-parameter energy-scale correction described in the main text. Minima are identified that improve the agreement between the observed and simulated  $^{252}\text{Cf}$  spectra relative to the best-fit scale factors (dots with  $1\sigma$  error bars) from the one-parameter correction. The 90% C.L. error ellipses associated with the best-fit scale factors and exponents are indicated as well.

above for the one-parameter correction was used to calculate a  $\chi^2$  statistic over a two-dimensional grid of scale factors and exponents for each detector ensemble. Smoothed versions of the resulting two-dimensional  $\chi^2$  statistics are shown in Figure 6.12, with the best-fit scale factors from Figure 6.10 included

#### 6.4. SYSTEMATIC STUDIES

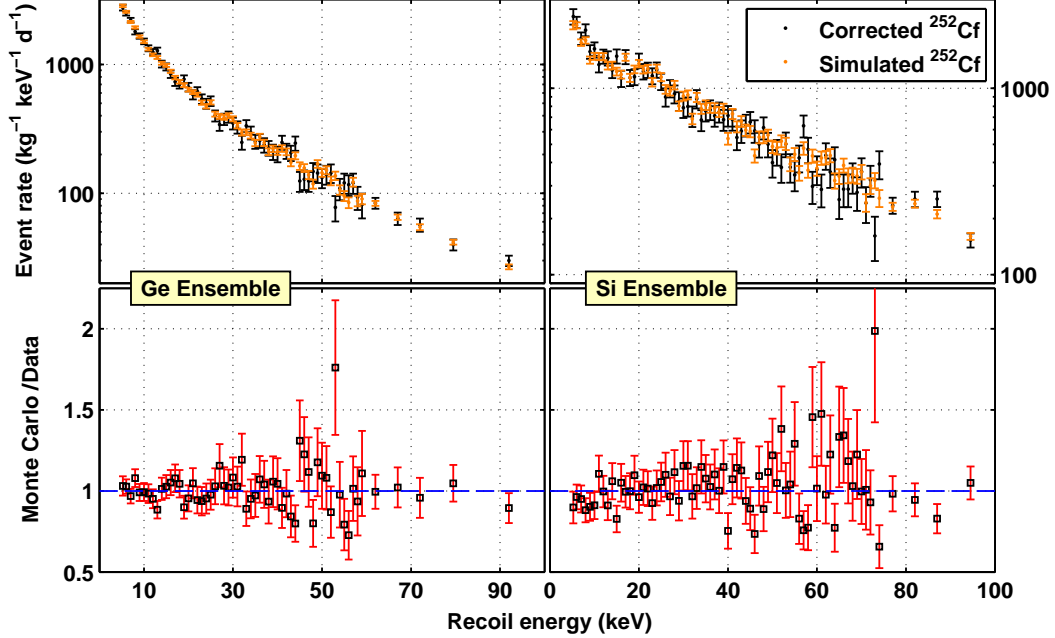


Figure 6.13: *Upper panels:* Spectral comparison of simulated (light/orange error bars) and observed (dark/black error bars)  $^{252}\text{Cf}$  nuclear recoils, where the latter reflects the corrected energy scale indicated in Equations 6.4.2 and 6.4.3. *Lower panels:* Ratios of the simulated to observed spectra (error bars) agree with expectation (dashed lines).

for comparison. Absolute minima are identified that yield lower  $\chi_0^2$  values than the corresponding one-parameter corrections;  $\chi_0^2 = 0.9$  (1.2) for the Ge (Si) detector ensemble. 90% confidence-level error ellipses for the best-fit parameters are derived as well. If the statistical uncertainties are projected onto the scale-factor and exponent axes in Figure 6.12, the best-fit parameters with  $1\sigma$  errors are:

$$\begin{aligned} \text{Ge Ensemble:} \quad & c = 1.166^{+0.019}_{-0.016} \quad \& \quad b = 1.184^{+0.032}_{-0.034}, \\ \text{Si Ensemble:} \quad & c = 1.154^{+0.018}_{-0.017} \quad \& \quad b = 1.051^{+0.023}_{-0.020}. \end{aligned} \quad (6.4.3)$$

The effect of applying this two-parameter correction is demonstrated in Figure 6.13, where the simulated and observed spectra are compared in the

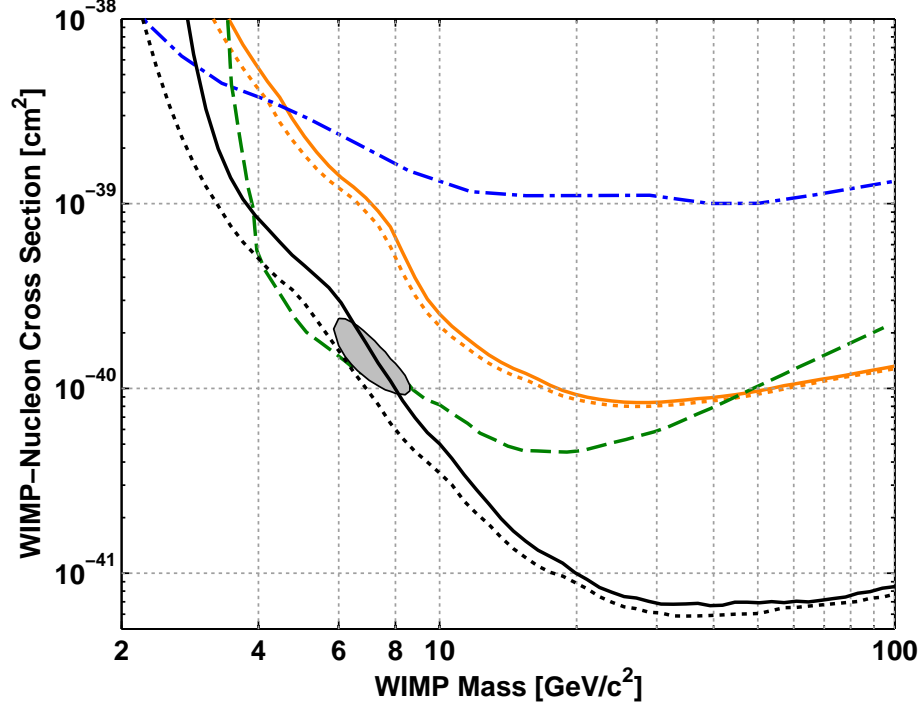


Figure 6.14: Comparison of the combined Ge and Si (dark/black solid line) and Si-only (light/orange solid line) exclusion limits, with those obtained following application of the best-fit two-parameter energy-scale correction described by Equations 6.4.2 and 6.4.3 (correspondingly colored dotted lines). The corrected Ge and Si limit just excludes Hooper *et al.*'s joint DAMA/LIBRA and CoGeNT signal region [652] (99% C.L., shaded region) at 90% confidence. Limits from the CoGeNT [670] (dashed line) and CRESST [597] (dash-dotted line) experiments are also indicated.

same manner as Figure 6.9. The spectral ratios are clearly much closer to unity, confirming the goodness-of-fit indicated by the relatively low  $\chi_0^2$  values quoted above. It is not immediately obvious how the correction will affect the low-threshold WIMP sensitivity. Application of Equations 6.4.2 and 6.4.3 to the original energy scale basically stretch it while holding recoil energies in the proximity of  $\sim 10$  keV fixed. Consequently, energies  $\lesssim 10$  keV are pushed lower, while energies  $\gtrsim 10$  keV are pushed higher. Since low-value recoil energies are

#### 6.4. SYSTEMATIC STUDIES

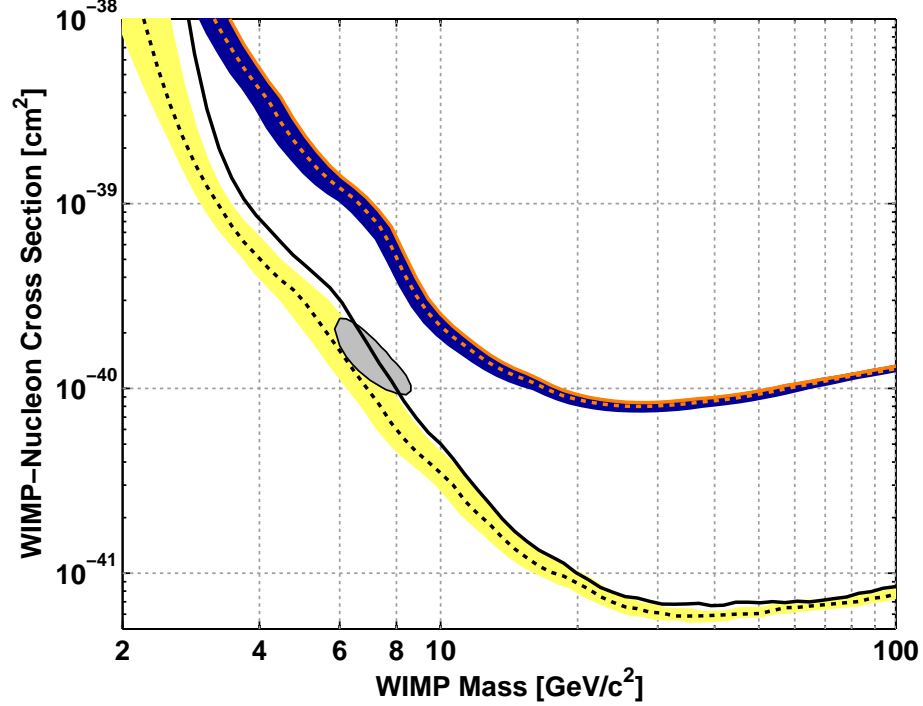


Figure 6.15: Comparison of the combined Ge and Si (dark/black solid line) and Si-only (light/orange solid line) exclusion limits, with those obtained following application of the best-fit two-parameter energy-scale correction described by Equations 6.4.2 and 6.4.3 (correspondingly colored dotted lines). Also shown are the 90% C.L. uncertainty bands associated with uncertainty in the energy-scale correction’s best-fit parameters for the Ge (light/yellow shaded region) and Si (dark/blue shaded region) detector ensembles, as well as Hooper *et al.*’s joint DAMA/LIBRA and CoGeNT signal region [652] (99% C.L., oval shaded region).

corrected to yet lower values, the threshold efficiencies are shifted to lower energies. The Ge detectors’  $\sim 1$  keV thresholds (at 50% efficiency), for example, are pushed down to  $\sim 0.7$  keV. The two-parameter correction should therefore yield improved sensitivity to low-mass WIMPs. When the correction is propagated through the relevant aspects of the low-threshold analysis, the exclusion limits shown in Figure 6.14 are obtained. Not only are both corrected limits signifi-

cantly stronger, but the combined Ge and Si limit just excludes Hooper *et al.*'s signal region at 90% confidence. Furthermore, the region of new parameter space excluded by the main result is extended to lower WIMP masses.

Unfortunately, our new found sensitivity comes with a bit of a catch; there is uncertainty in the energy-scale correction's best-fit parameters. To gauge the affect of this uncertainty on the low-threshold limits, the correction parameters were varied within the 90% C.L. error ellipses shown in Figure 6.12. For each detector ensemble, 16 sets of correction parameters were chosen, half from inside the ellipse boundary and half from the ellipse itself. These corrections were propagated through the low-threshold analysis, and corresponding exclusion limits were derived. The resulting 90% C.L. uncertainty bands are indicated in Figure 6.15, and clearly weaken the above claims against Hooper *et al.*'s joint signal region.

# Appendix A

## The Veto Front-End Board

The custom circuits used for analog readout of the muon-veto counters at the CDMS shallow- and deep-site installations are documented in this appendix. This includes detailed circuit diagrams, printed circuit board (PCB) layouts, and digital photographs for the veto Front-End Board (vFEB). For the deep-site vFEB, the component labels in the circuit diagrams are used to reference the parts lists provided in Tables A.1 and A.2. All circuit diagrams and layouts were created with the Cadence OrCAD software package. The deep-site PCB layouts are thanks to by Bruce Lyons.

This appendix is basically a collection of figures. Rather than annotate each figure individually, a summary is provided in the following list, with the figures appearing subsequently. This allows several of the larger figures to span an entire page, ensuring clarity of their finer details. Note that only the metric prefix is given for units of capacitance (F) and resistance ( $\Omega$ ) in the circuit diagrams (*e.g.*, 47 n refers to a capacitance of 47 nF).

- **Figure A.1:** Diagram of the input-buffer stage of the vFEB prototyped during Run 21. A high slew rate ( $3000 \text{ V}/\mu\text{s}$ ) op amp is configured as a unity-gain follower. Note that the input is capacitively coupled and terminated across  $50 \Omega$ . This circuit connects via the right side of the coupling resistor R3 to the integrating-amplifier stage at the location indicated in Figure A.2.



## APPENDIX A. THE VETO FRONT-END BOARD

- **Figure A.2:** Continuance of Figure A.1, detailing the integrating-amplifier stage of the vFEB prototype. Pulses arriving (from the previous stage) at the negative input of the high-bandwidth (400 MHz) op amp are integrated onto the feedback capacitor C21. The integrated charge then decays with a time constant determined by the product of C21 and R4. This circuit connects from between C8 and R5 to the output-driver stage at the location indicated in Figure A.3.
- **Figure A.3:** Diagram of the output-driver, final stage of the circuit continued from Figure A.2. A high-speed, high-output-current op amp—configured for a gain of +2—drives signals (from the previous stage) through a  $200\ \Omega$  resistor into a  $50\ \Omega$  line. When the output signals are terminated across  $50\ \Omega$ , this configuration yields an overall gain of  $+2/5$ .
- **Figure A.4:** Diagram of the circuitry used to regulate the  $\pm 24\ \text{V}$  ( $\pm 12\ \text{V}$ ) voltages on the NIM backplane down to the  $\pm 15\ \text{V}$  ( $\pm 5\ \text{V}$ ) power rails required by the vFEB's op amps. All diodes are type 1N4002.
- **Figure A.5:** Diagram of the input-buffer stage of the deep-site vFEB. This is largely the same as Figure A.1, except the second channel of the op amp is configured to fan-out a copy of the input for use with the digital-history readout chain. By configuring the fan-out to have +2 gain and driving it through the  $50\ \Omega$  resistor R35, the overall gain for a  $50\ \Omega$  termination is +1. Note that the PMT input is capacitively coupled and  $50\ \Omega$  terminated with a combination of discrete components that acts as a high-pass filter.
- **Figure A.6:** Continuance of Figure A.5, detailing the integrating-amplifier stage of the deep-site vFEB. The functionality is identical to Figure A.2.

- **Figure A.7:** Diagram of the output-driver, final stage of the deep-site vFEB circuit continued from Figure A.6. The functionality is identical to Figure A.3. Note that the coupling capacitor C113 and resistor R41 act as a high-pass filter.
- **Figure A.8:** Diagram of the deep-site vFEB's voltage-regulation circuitry. Other than a change in the values of C236, C240, C244 and C248 from  $68\mu\text{F}$  to  $100\mu\text{F}$ , and the addition of test points, these circuits are identical to those in Figure A.4. The test points provide convenient locations to monitor the regulators output voltages as they are tuned via the potentiometers R89, R91, R93 and R95.
- **Figure A.9:** All layers of the deep-site vFEB PCB layout. The PCB includes four physical layers: top, bottom, power, and ground. There is a layout layer for each, as well as a separate layer for the surface-mount footprints on the top side. The individual layout layers are provided in the next 6 figures, including a silk-screen pattern for the top side. Dimensions are in inches. Note the large heat-sink footprints surrounding the voltage regulators (near the top of this figure). The drill chart indicates the hole diameters for the penetrations (in red) required by the design.
- **Figure A.10:** PCB layout for the top side of the deep-site vFEB.
- **Figure A.11:** PCB layout for the bottom side of the deep-site vFEB.
- **Figure A.12:** PCB layout for the power layer of the deep-site vFEB.
- **Figure A.13:** PCB layout for the ground layer of the deep-site vFEB.
- **Figure A.14:** PCB layout for the surface-mount footprints on the top side of the deep-site vFEB.
- **Figure A.15:** Silk-screen pattern for the top side of the deep-site vFEB.
- **Figure A.16:** Digital photograph of a fully assembled deep-site vFEB packaged in a single-width NIM module.

## APPENDIX A. THE VETO FRONT-END BOARD

Table A.1: List of capacitors for the deep-site vFEB. In the first column, the numbers correspond to the component labels that begin with a  $C$  in Figures A.5, A.6, A.7, and A.8. In the third column, percentages refer to tolerances in the values given in the second column, while voltages indicate maximum ratings.

Reference	Description	Specifications	Package
88, 91, 94, 98, 101, 104, 107, 111, 115	0.1 $\mu\text{F}$ ceramic	50 V, 10%	SM 805
90, 93, 99, 100, 105, 106, 110, 114, 235, 239, 243, 247	10 nF ceramic	50 V, 10%	SM 805
92, 95, 97, 102, 103, 108, 112, 116	6.8 $\mu\text{F}$ tantalum	20 V, 20%	SM B size
89, 109	2.2 pF ceramic	50 V, 5%	SM 1206
96	51 pF ceramic	50 V, 5%	SM 1206
113	47 pF ceramic	50 V, 10%	SM 805
233, 237, 241, 245	1 $\mu\text{F}$ tantalum	35 V, 20%	SM B size
234, 238, 242, 246	10 $\mu\text{F}$ tantalum	16 V, 20%	SM B size
236, 240, 244, 248	100 $\mu\text{F}$ tantalum	16 V, 20%	SM D size

Table A.2: List of parts (that are not capacitors) for the deep-site vFEB.

Reference	Description	Specifications	Package
<i>Resistors</i> (prefix R)			
34, 38	100 $\Omega$ chip	1/8 W, 1%	SM 1206
35, 39	49.9 $\Omega$ chip	1/8 W, 1%	SM 1206
44	200 $\Omega$ chip	1/8 W, 1%	SM 1206
36, 37, 42, 43	510 $\Omega$ chip	1/4 W, 1%	SM 1206
40	100 k $\Omega$ chip	1/8 W, 1%	SM 1206
41	1.5 k $\Omega$ chip	1/8 W, 5%	SM 1206
89, 91, 93, 95	2 k $\Omega$ trim pot	1/4 W	4 mm
90, 92, 94, 96	120 $\Omega$ chip	1/4 W, 1%	SM 1206
<i>Others</i>			
D1–D8	rectifier diode	1 A, 100 V	DO-214
J1–J2	elbow socket	LEMO	PCB mount
J3	fixed socket	LEMO	bulkhead
L25–L32	0.22 $\mu$ H inductor	5% fixed	SM 805
TP1–TP4	PC test point		mini SM
U10	dual op amp	high speed	SO-8
U11	op amp	FET input	SO-8
U12	op amp	high speed	SO-8
U25–U26	adj regulator	+1.5 A	TO-220-3
U27–U28	adj regulator	–1.5 A	TO-220-3

## APPENDIX A. THE VETO FRONT-END BOARD

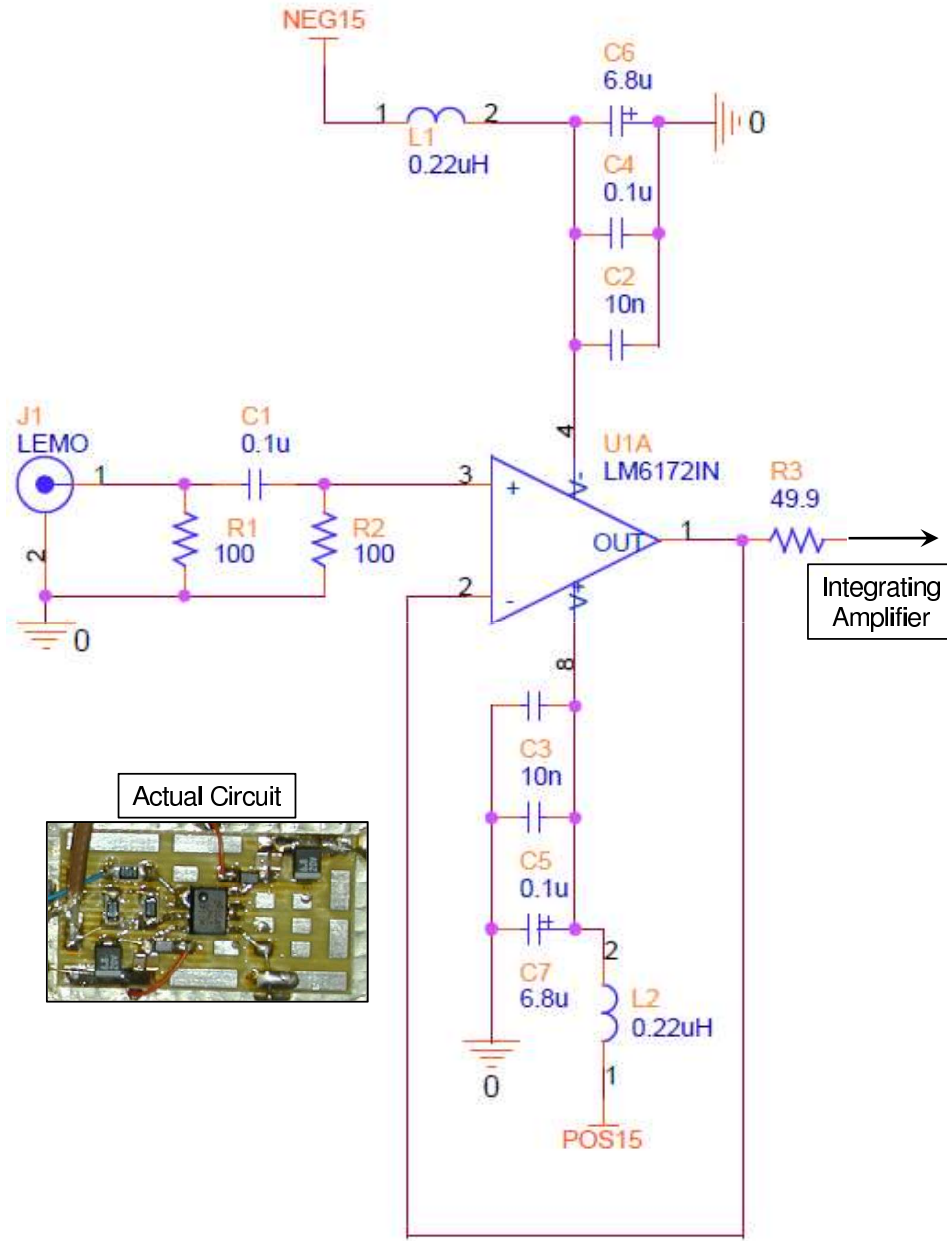


Figure A.1: Input-buffer stage of the prototype circuit.

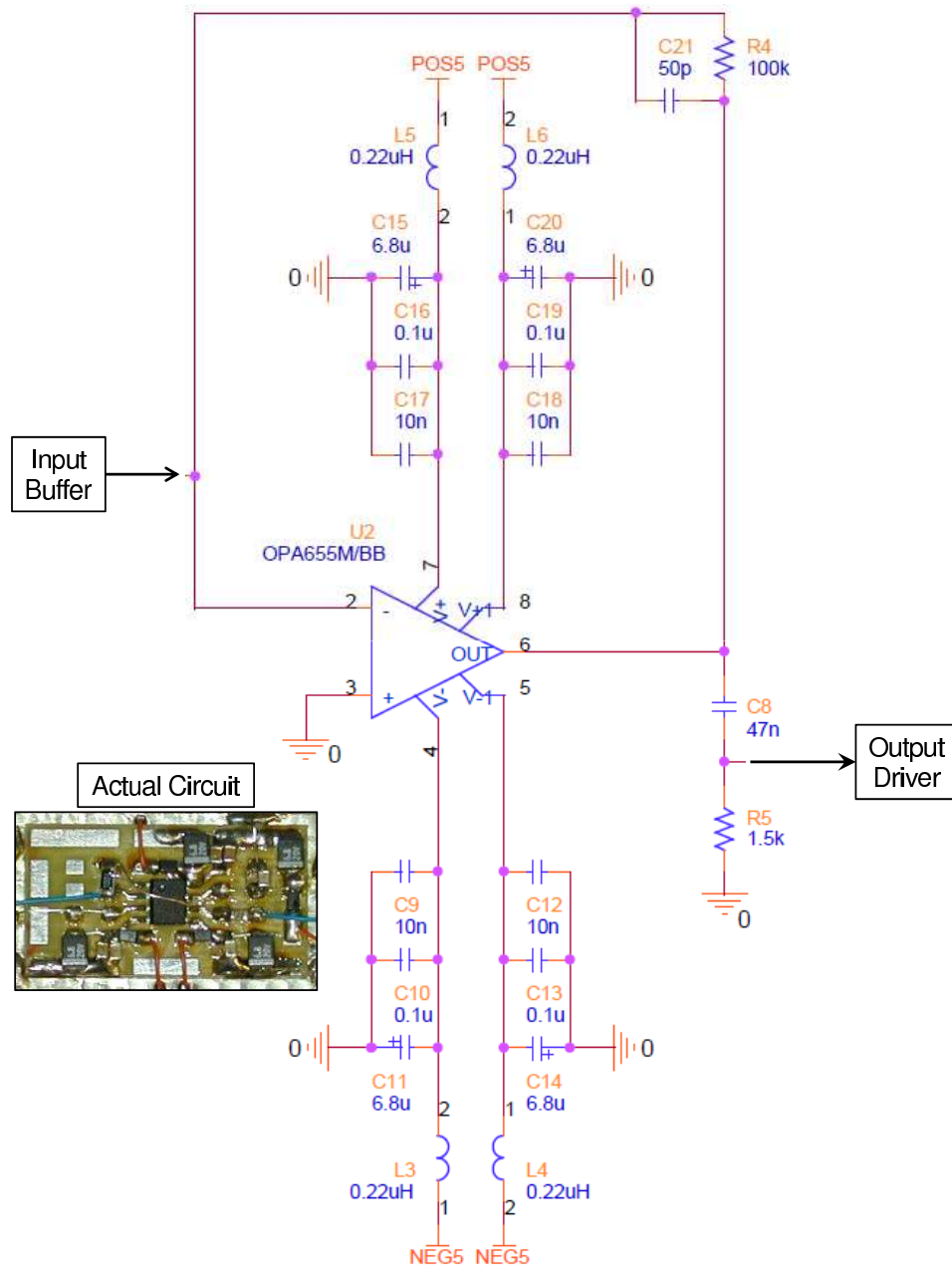


Figure A.2: Integrating-amplifier stage of the prototype circuit.

## APPENDIX A. THE VETO FRONT-END BOARD

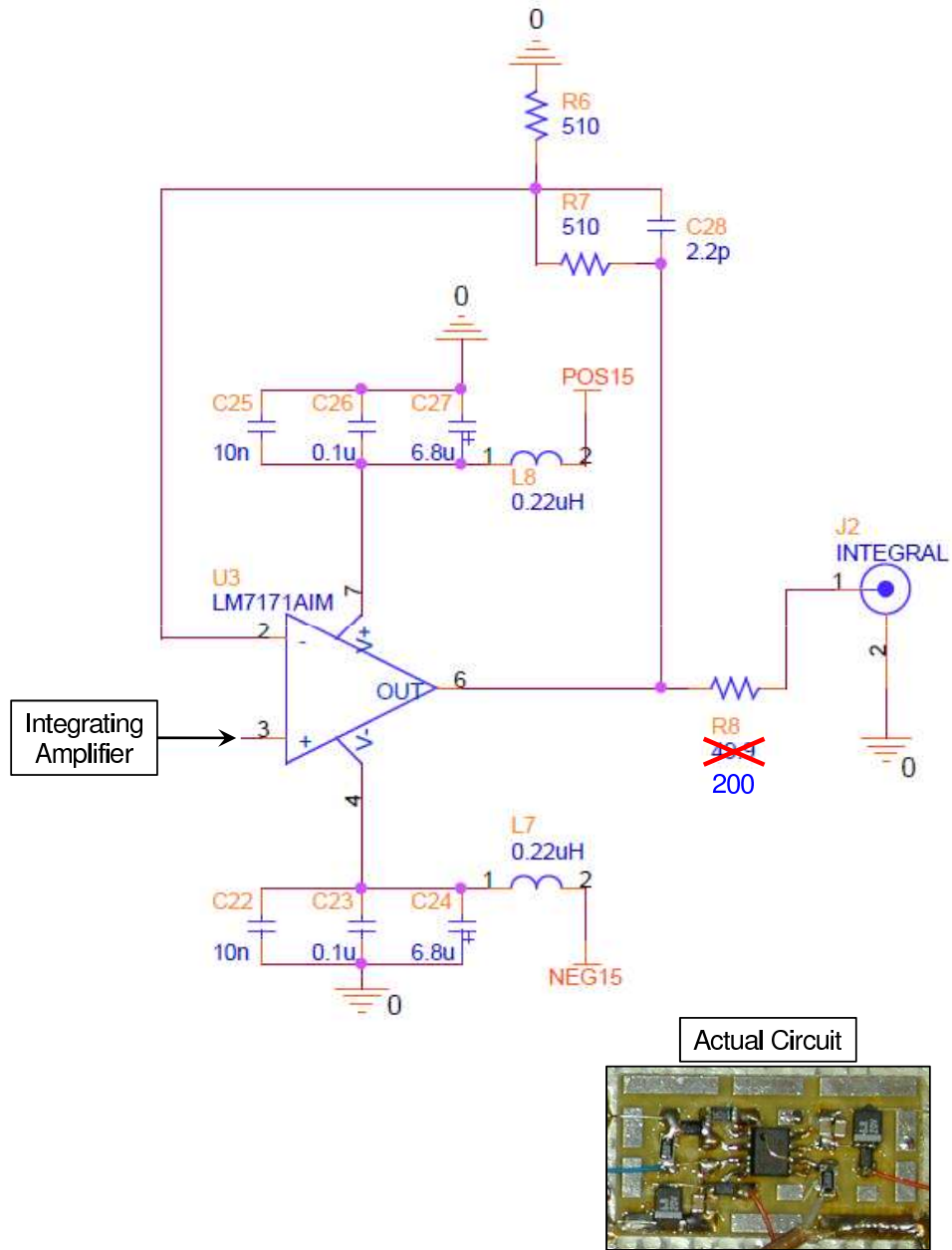


Figure A.3: Output-driver stage of the prototype circuit.

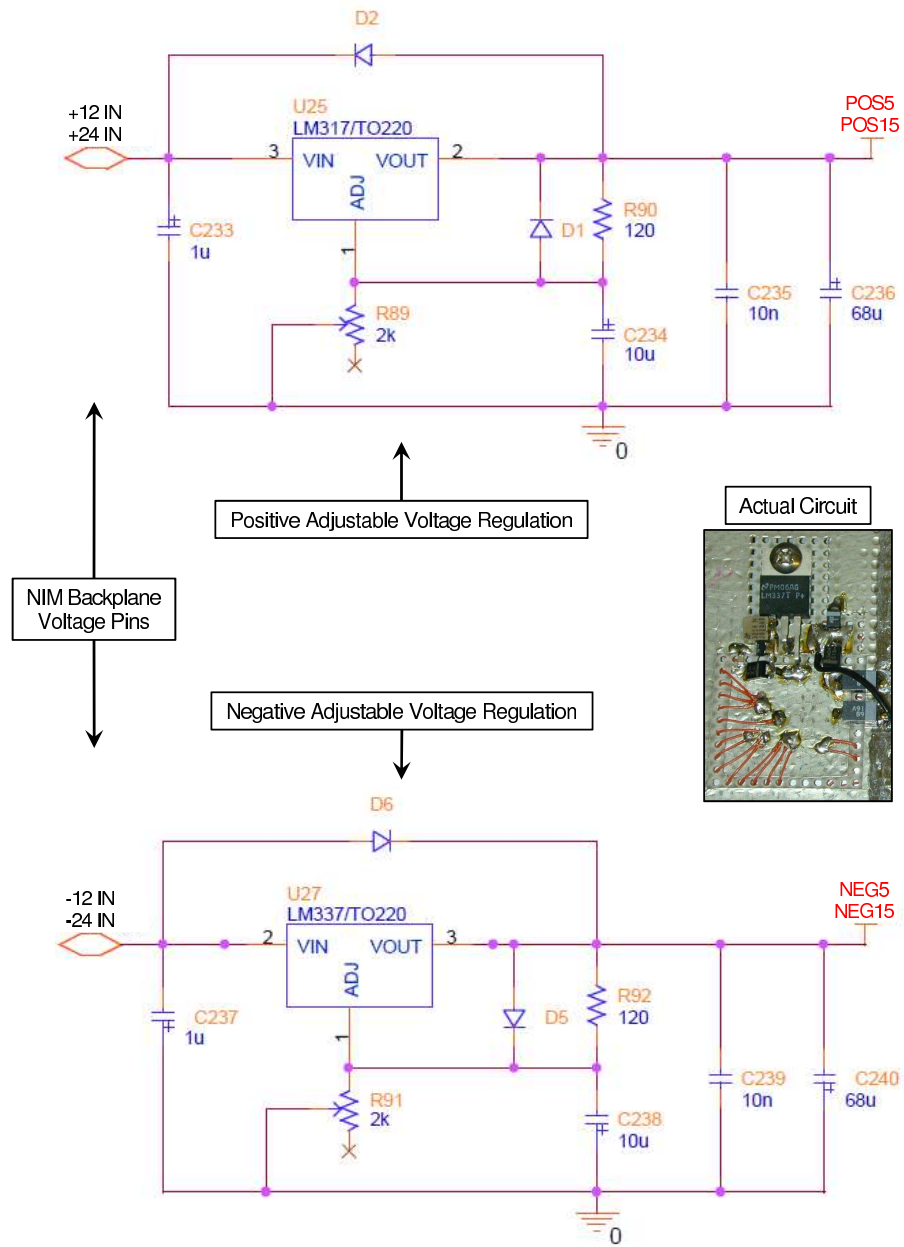


Figure A.4: Voltage regulation for the prototype circuit.



# APPENDIX A. THE VETO FRONT-END BOARD

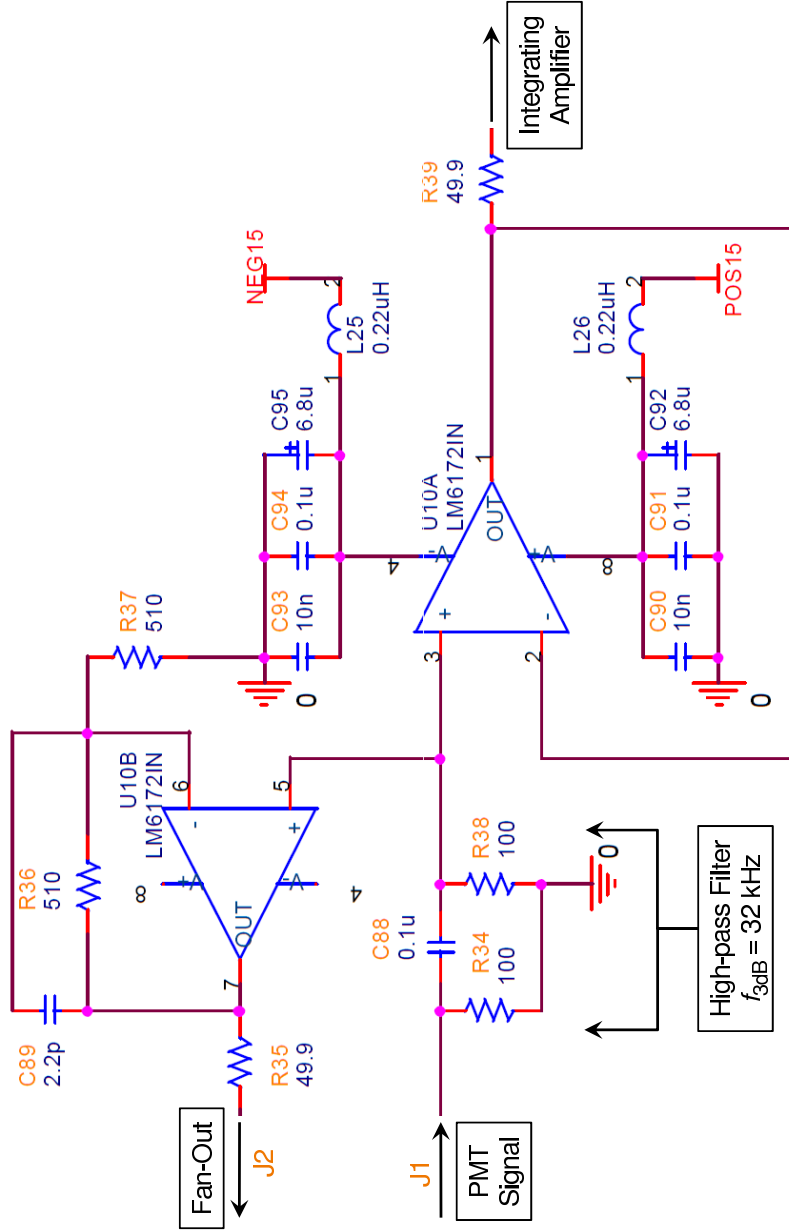


Figure A.5: Input-buffer stage of the deep-site vFEB.

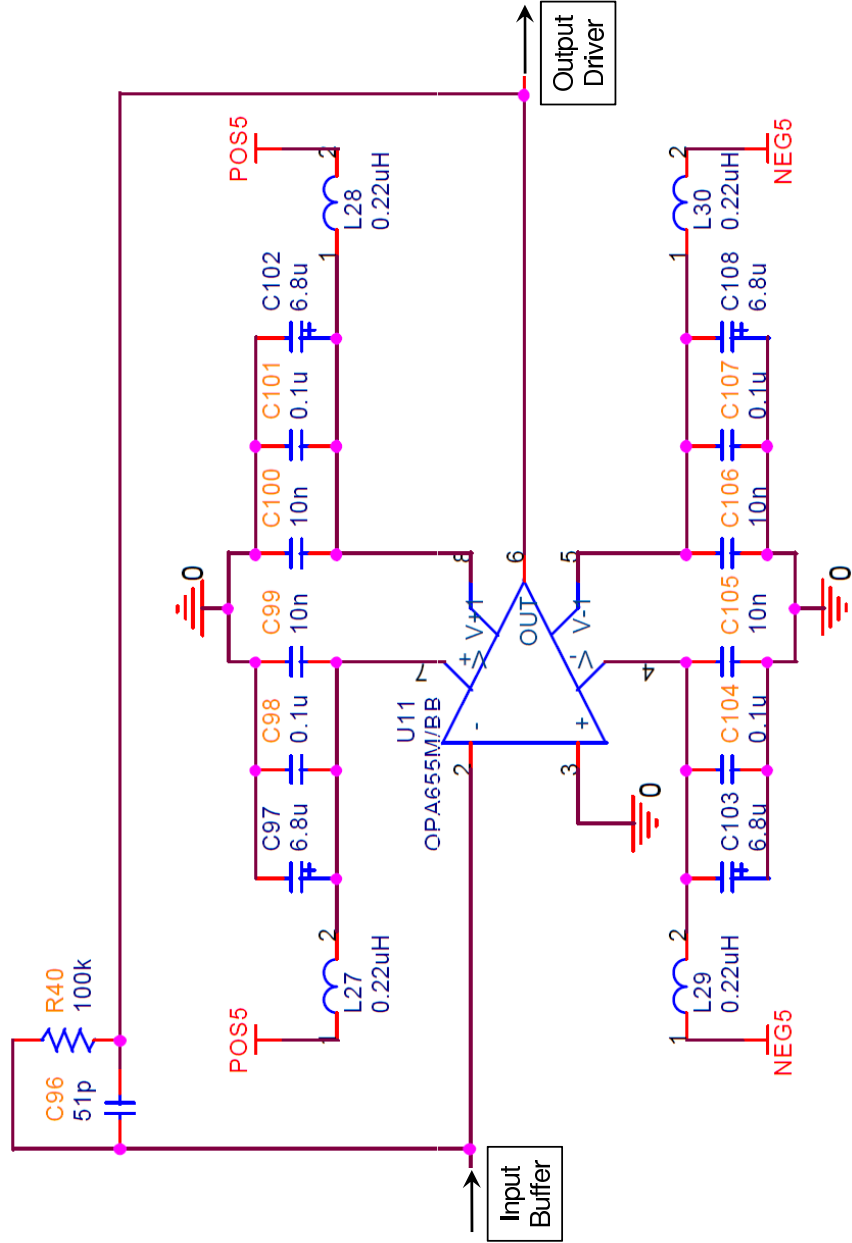


Figure A.6: Integrating-amplifier stage of the deep-site vFEB.

# APPENDIX A. THE VETO FRONT-END BOARD

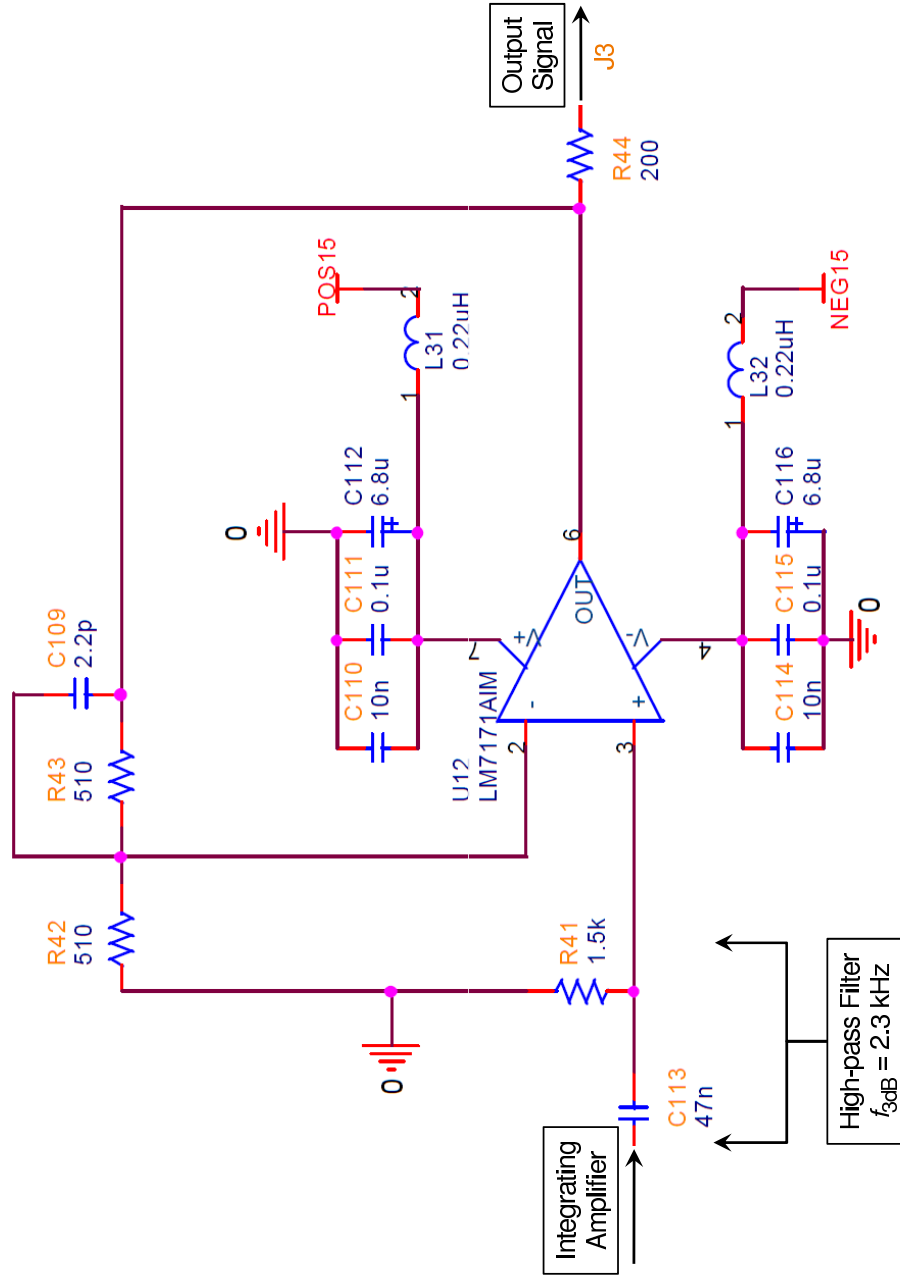
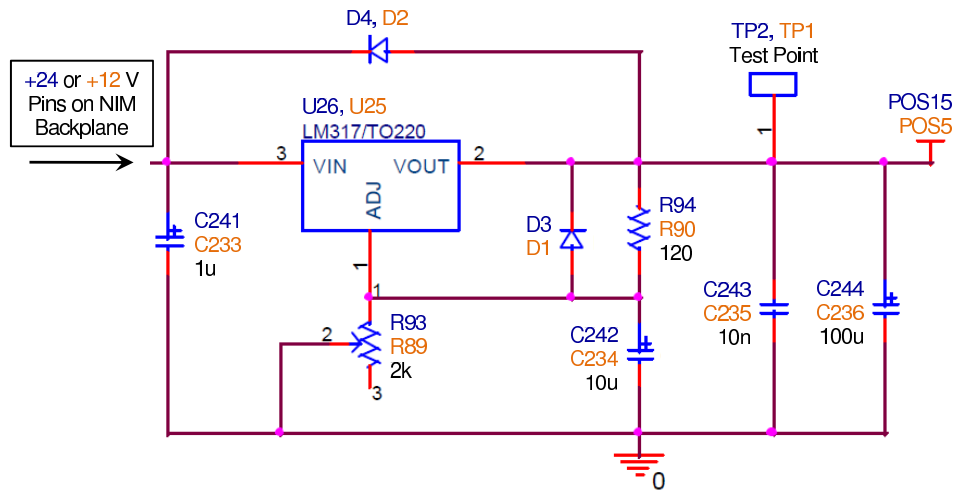


Figure A.7: Output-driver stage of the deep-site vFEB.

Positive Voltage Regulation: +24 (+12) V  $\rightarrow$  +15 (+5) V



Negative Voltage Regulation: -24 (-12) V  $\rightarrow$  -15 (-5) V

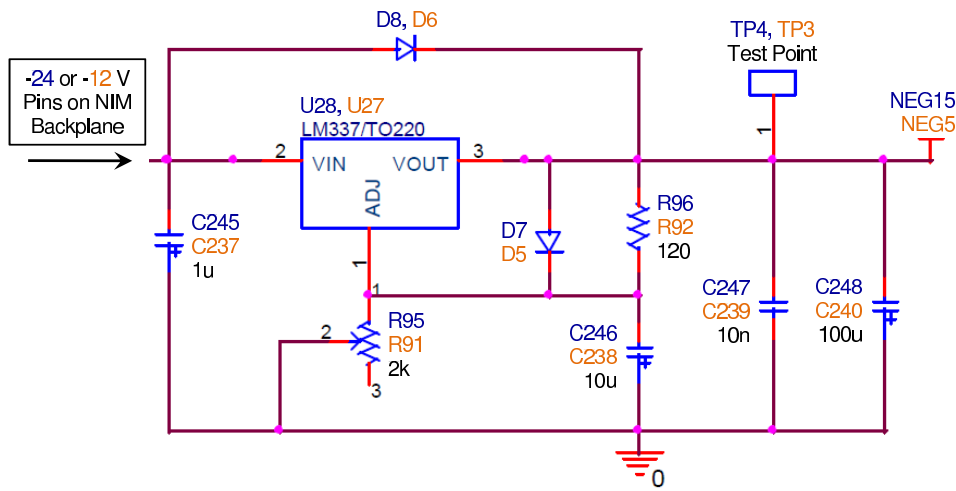


Figure A.8: Voltage regulation for the deep-site vFEB.

# APPENDIX A. THE VETO FRONT-END BOARD

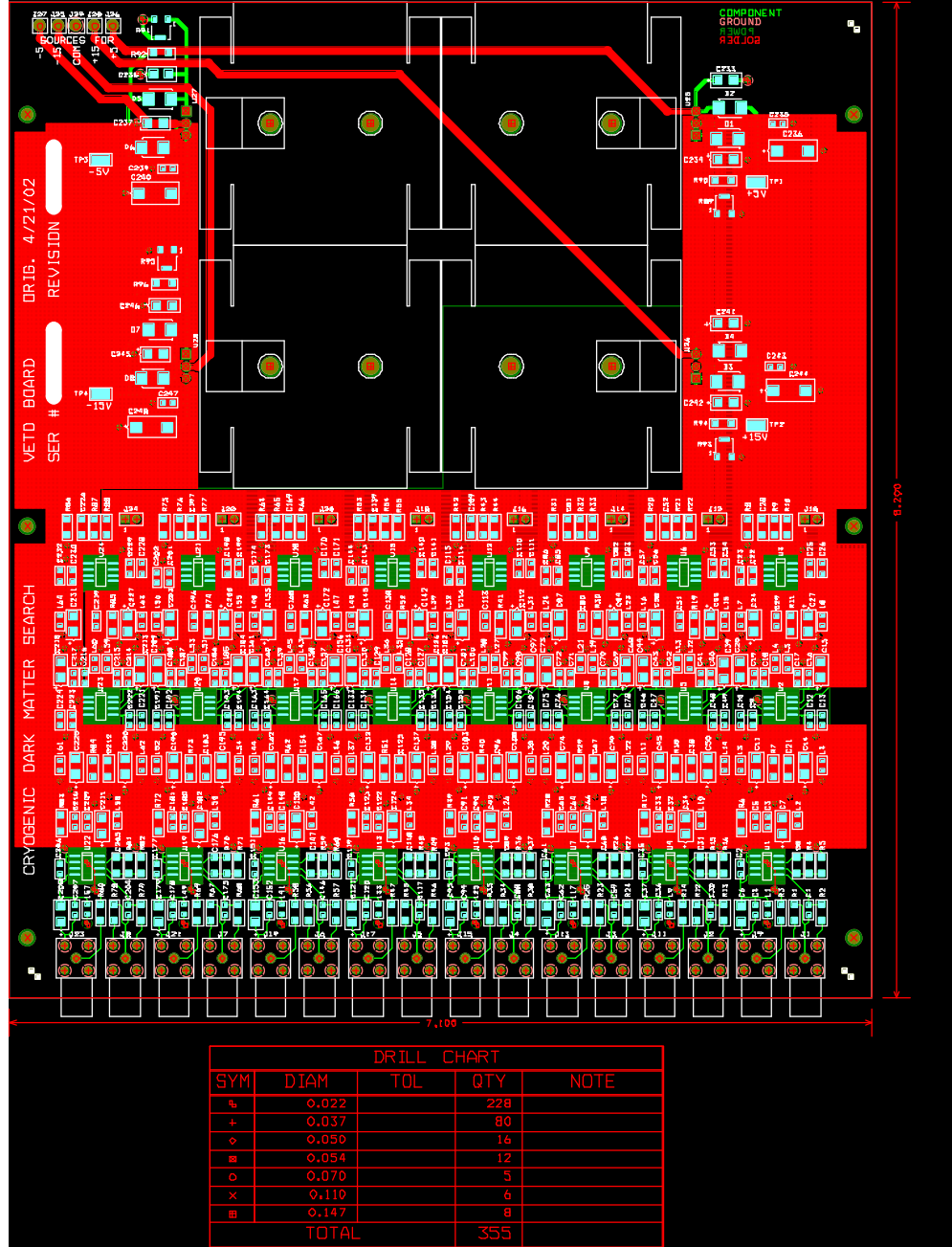


Figure A.9: All-layer PCB layout for the deep-site vFEB.

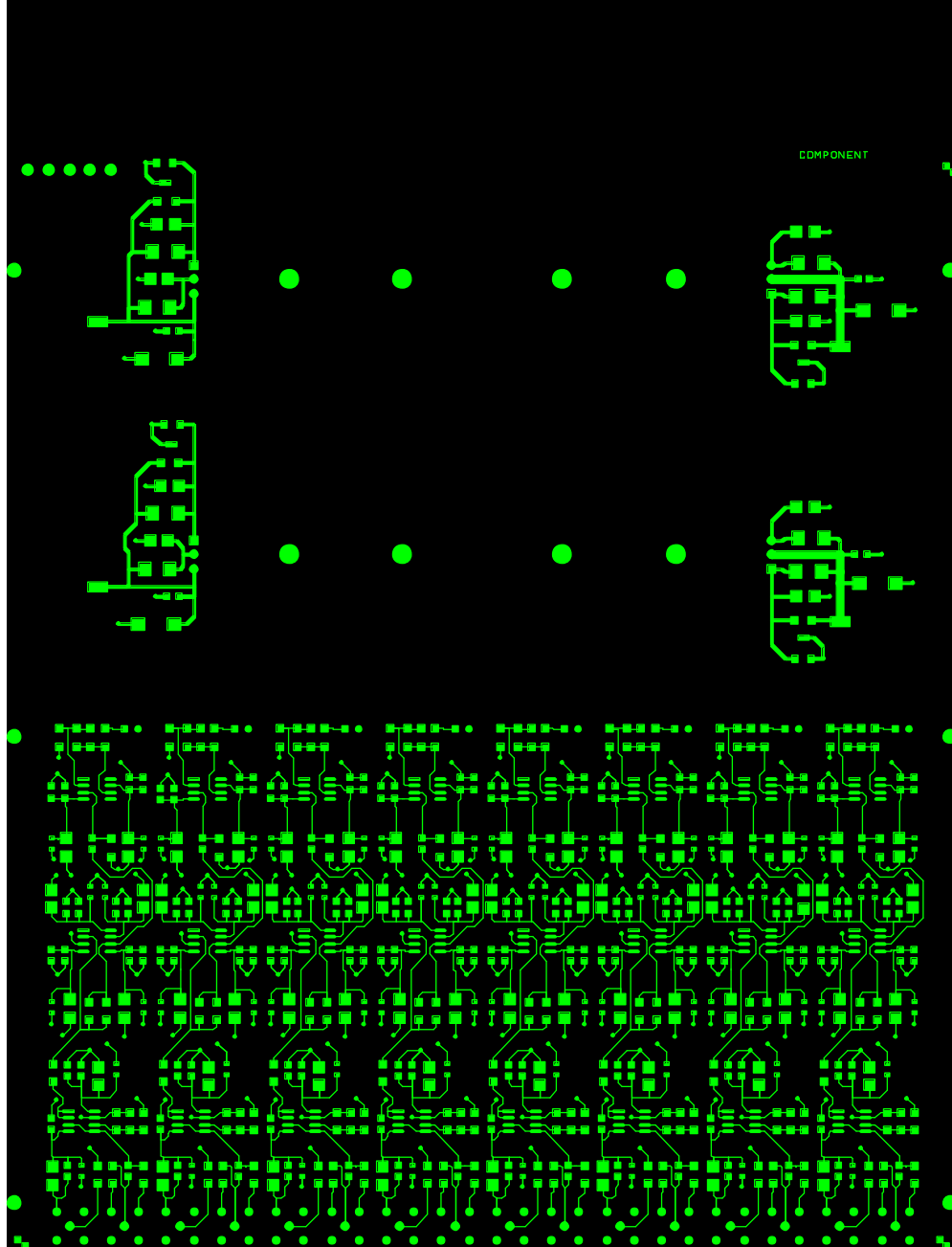


Figure A.10: Top-side layout for the deep-site vFEB.

APPENDIX A. THE VETO FRONT-END BOARD

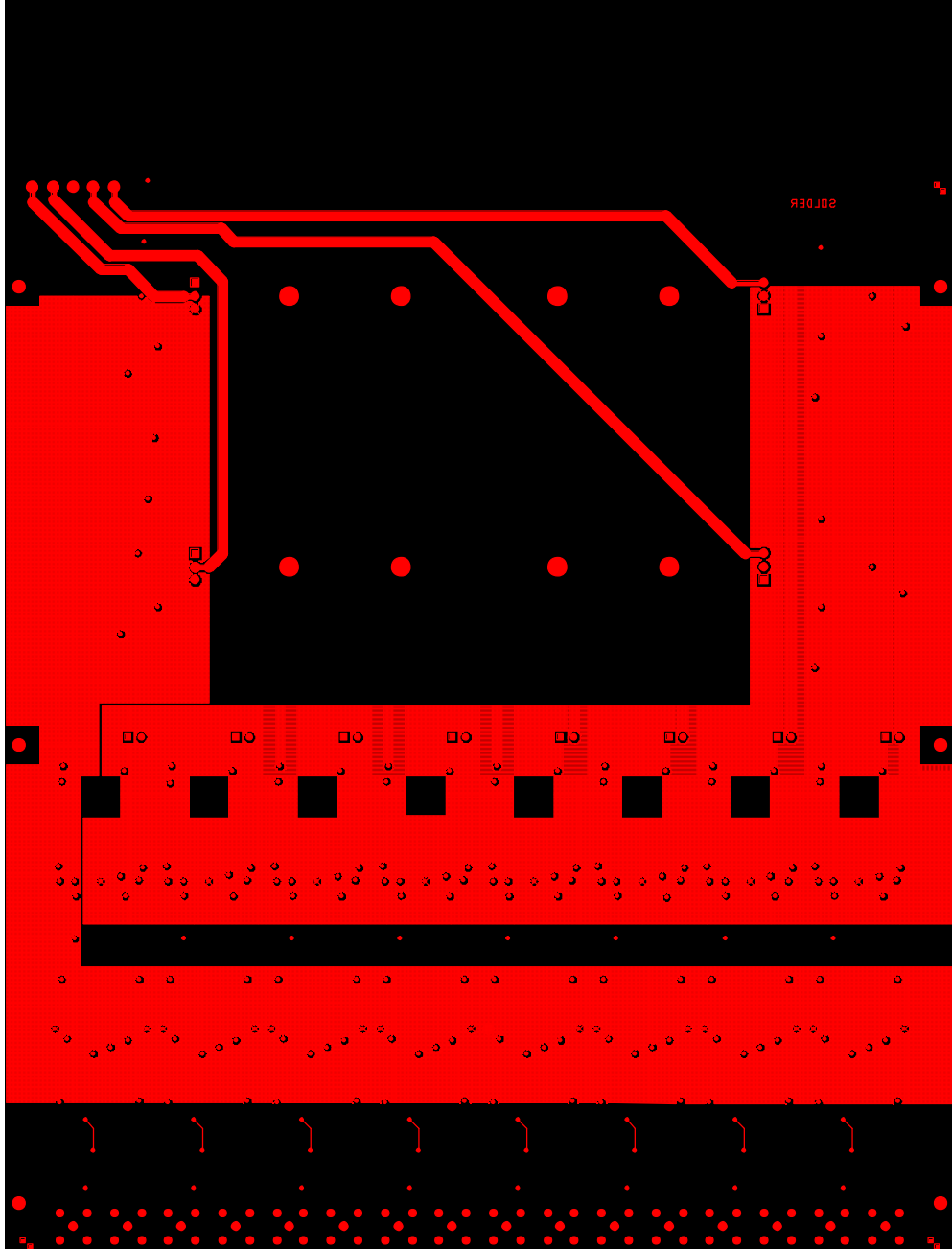


Figure A.11: Bottom-side layout for the deep-site vFEB.

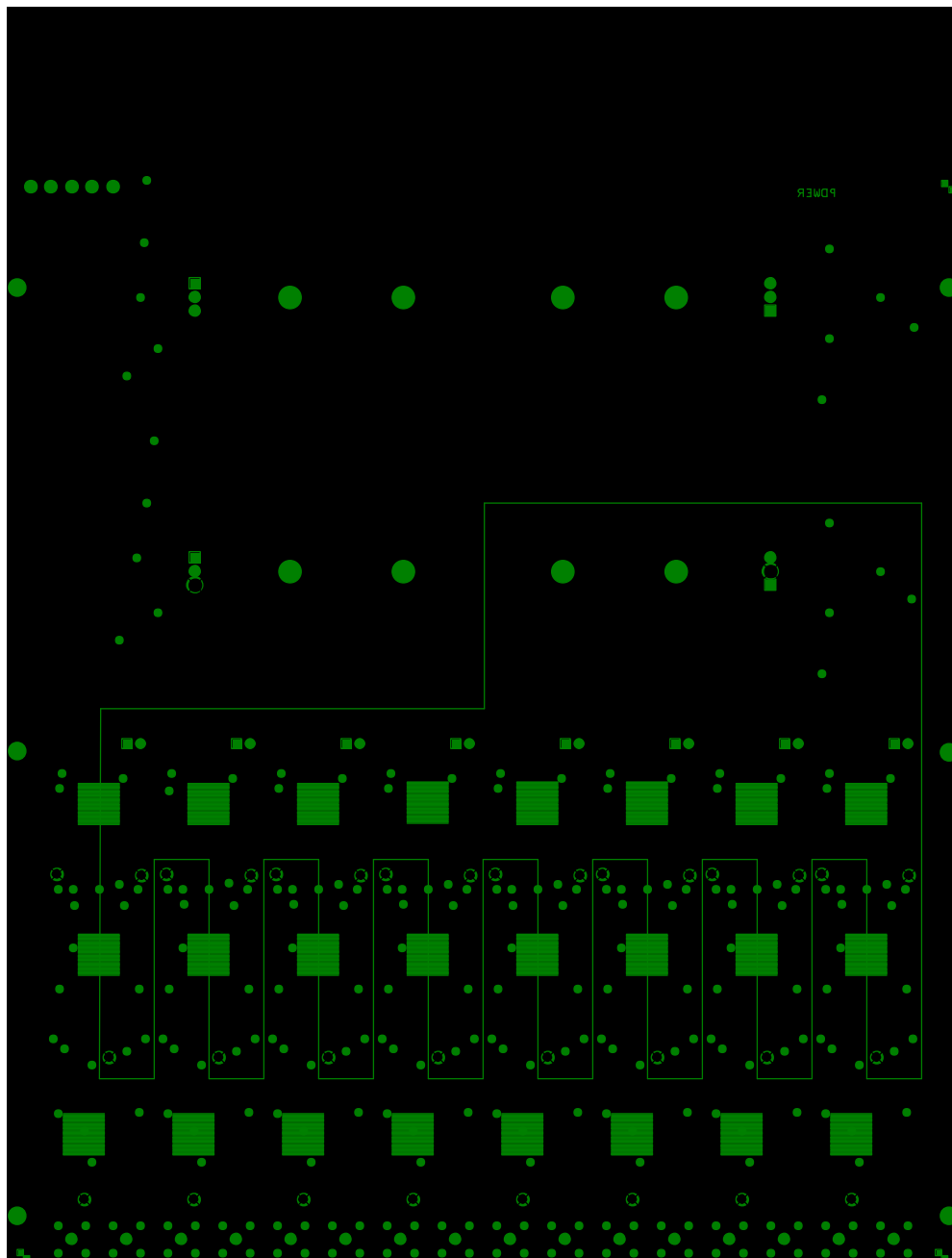


Figure A.12: Power-layer layout for the deep-site vFEB.



APPENDIX A. THE VETO FRONT-END BOARD

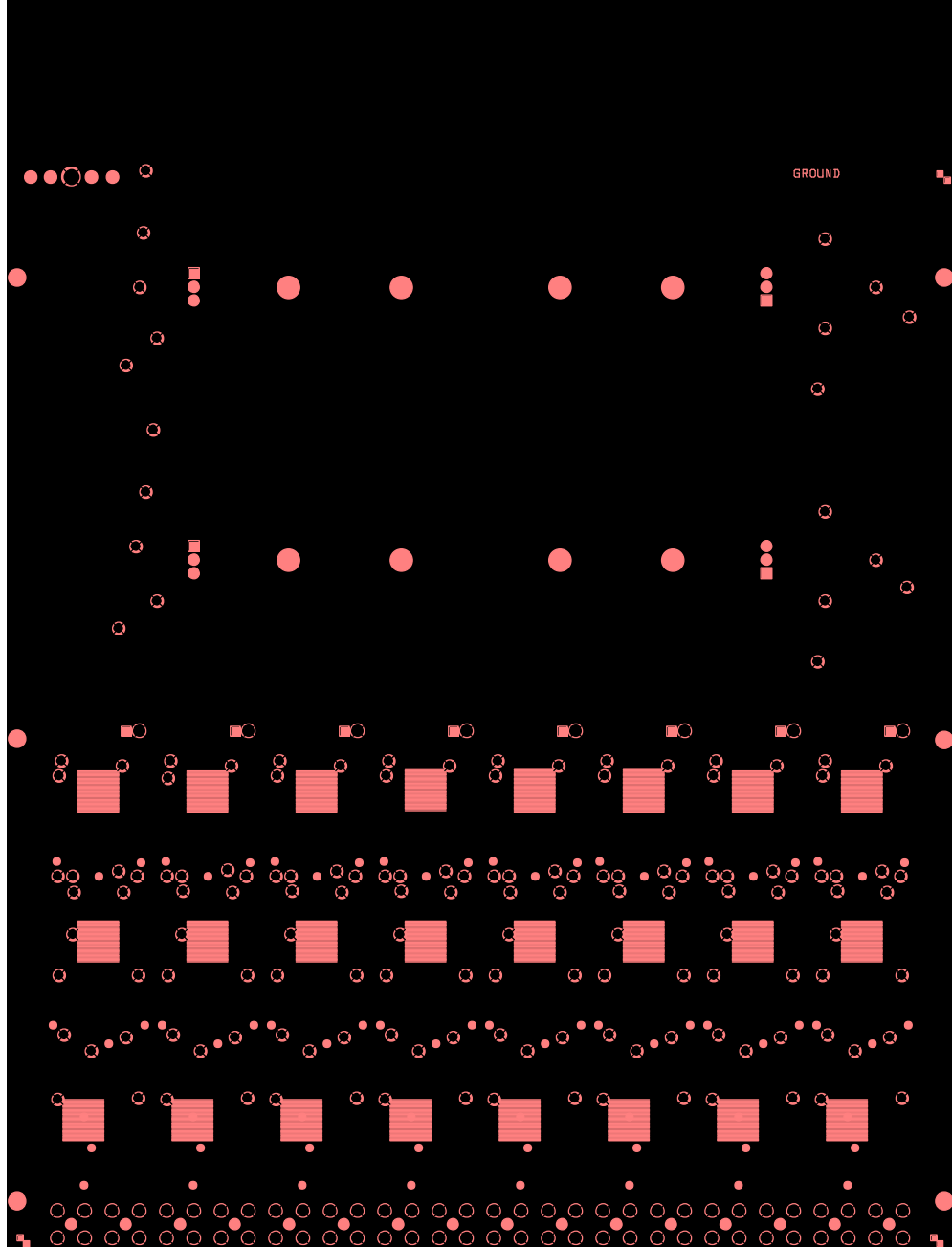


Figure A.13: Ground-layer layout for the deep-site vFEB.

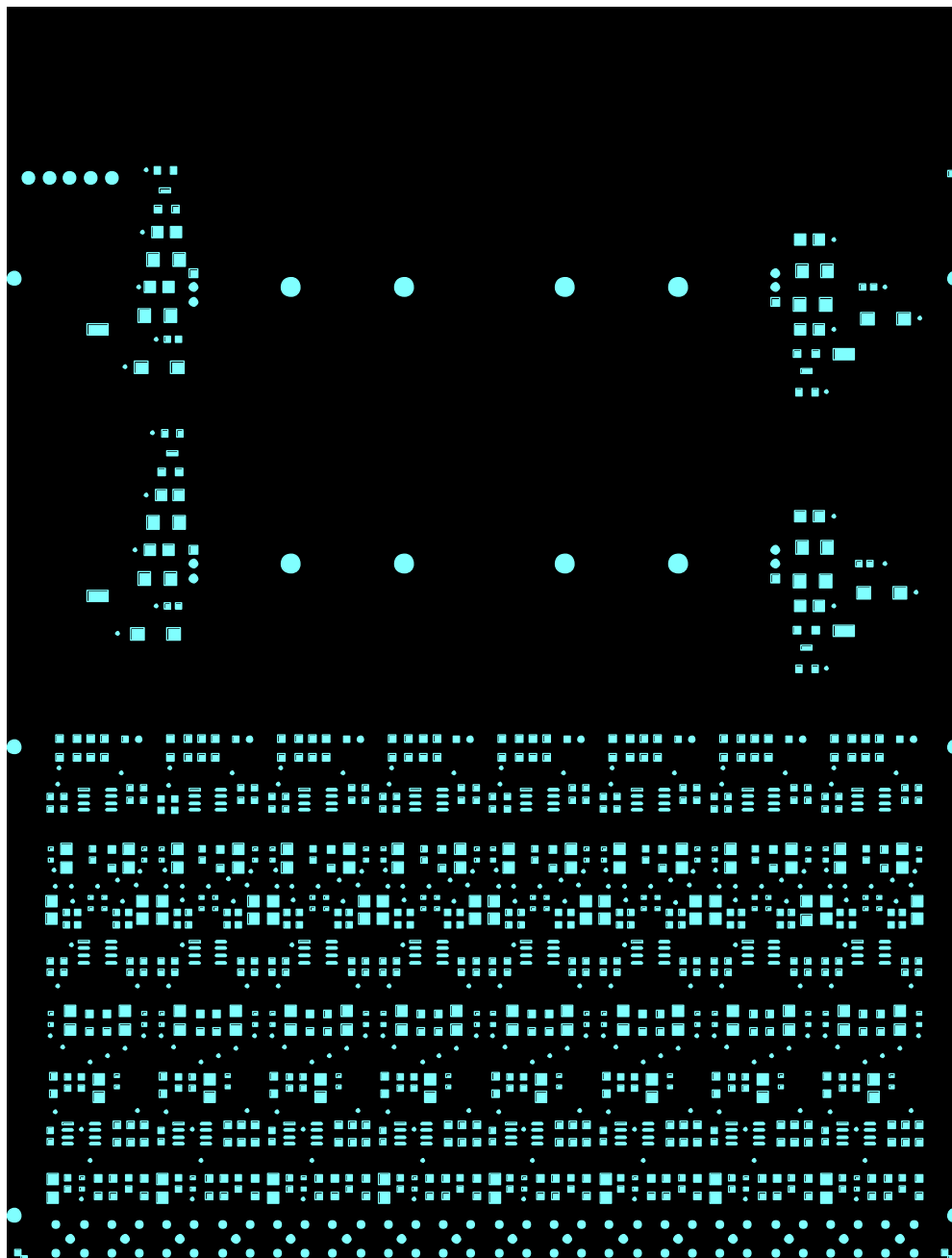


Figure A.14: Top-side surface-mount footprints for the deep-site vFEB.

## APPENDIX A. THE VETO FRONT-END BOARD

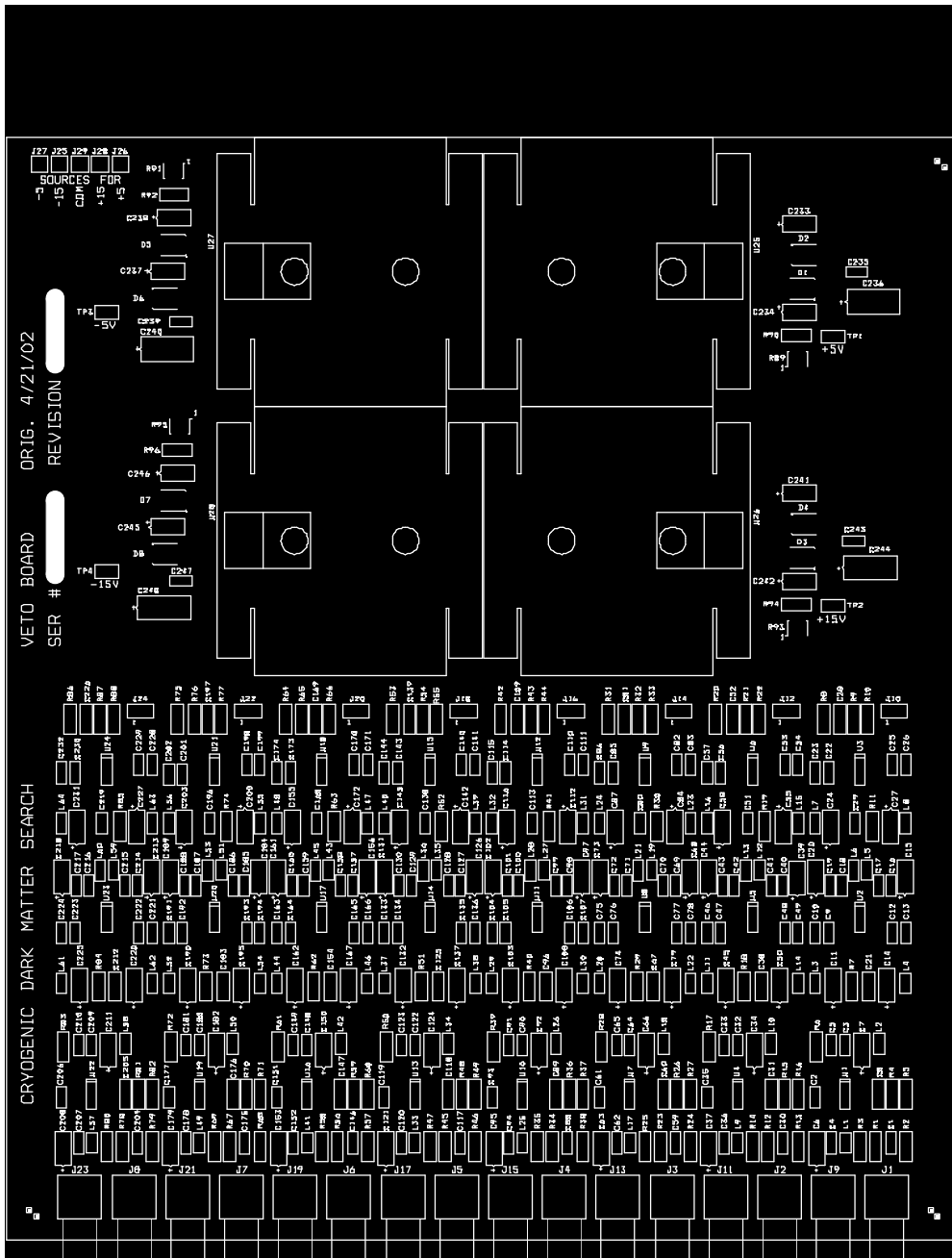


Figure A.15: Silk-screen pattern for the deep-site vFEB.



Figure A.16: Photograph of a completed deep-site vFEB.



# Appendix B

## Threshold Efficiencies

This appendix includes figures associated with the hardware and software threshold-efficiency estimates for the detectors not shown in Section 5.2. See the main text and figure captions in that section for further details.

### B.1 Hardware Thresholds

In this section there are three figures for each viable low-threshold detector and bias-voltage run (excluding the figures for the Z2 3V data, which appear in Section 5.2.1):

- Estimate of the hardware trigger efficiency (equivalent to Figure 5.2);
- The nuclear-recoil band centroid and  $1\sigma$  error envelope used to convert total phonon energy to  $Y_{\text{NR}}$ -corrected recoil energy (equivalent to Figure 5.3); and
- Trigger-efficiency statistical uncertainties (equivalent to Figure 5.4).

## APPENDIX B. THRESHOLD EFFICIENCIES

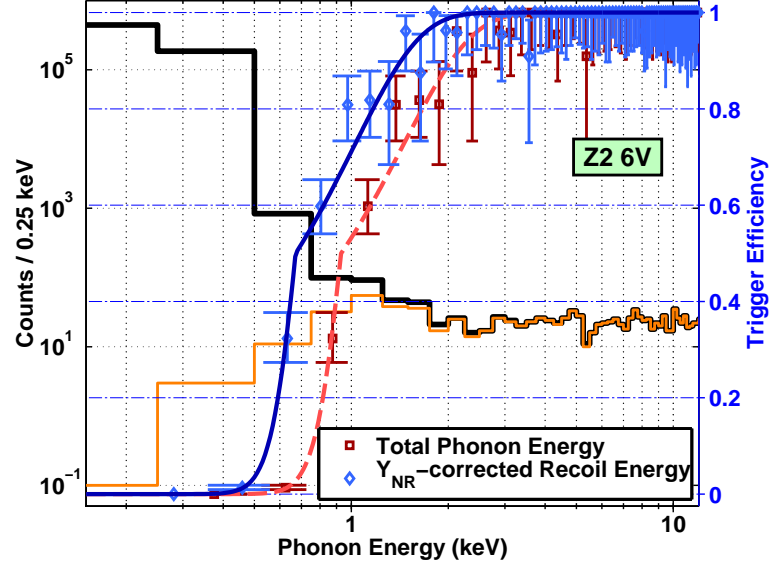


Figure B.1: Trigger-efficiency estimate for the Z2 6V data. The statistics are low due to the event-burst cut described in Section 5.3.2.

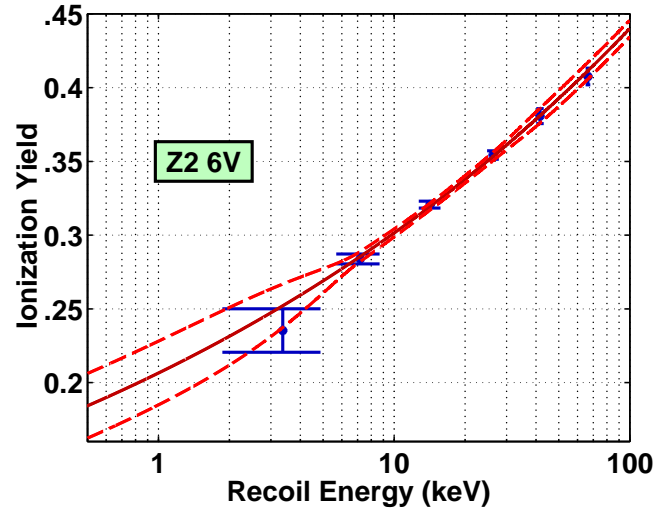


Figure B.2: The nuclear-recoil band centroid and associated error envelope for the Z2 6V data.

### B.1. HARDWARE THRESHOLDS

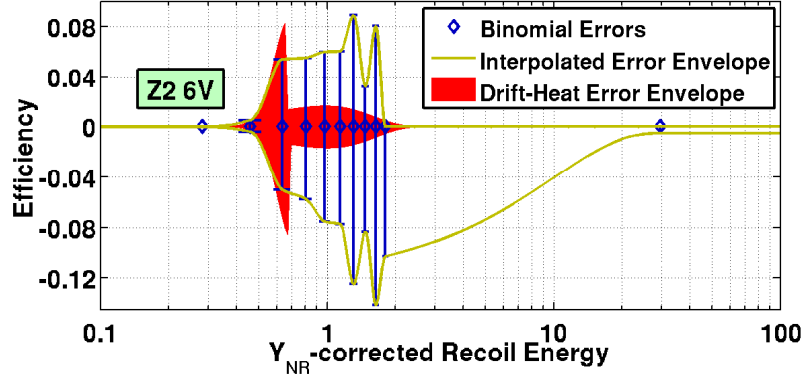


Figure B.3: Trigger-efficiency statistical uncertainties for the Z2 6V data.

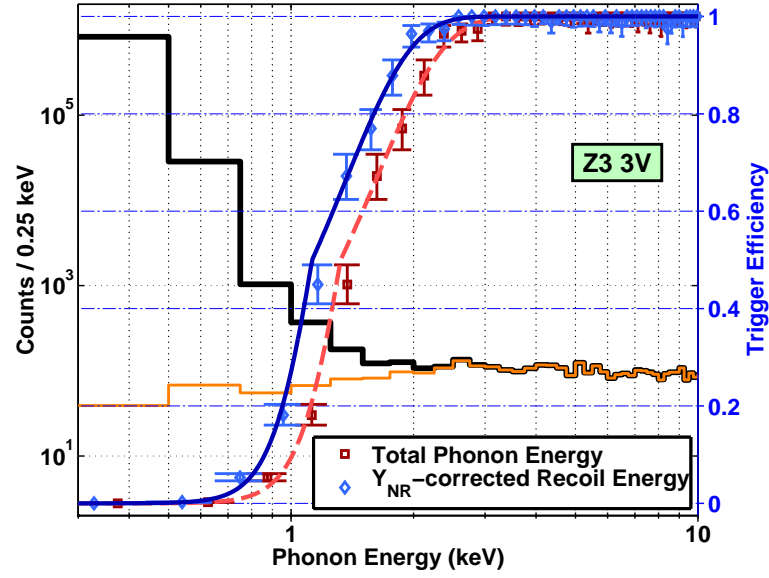


Figure B.4: Trigger-efficiency estimate for the Z3 3V data.



## APPENDIX B. THRESHOLD EFFICIENCIES

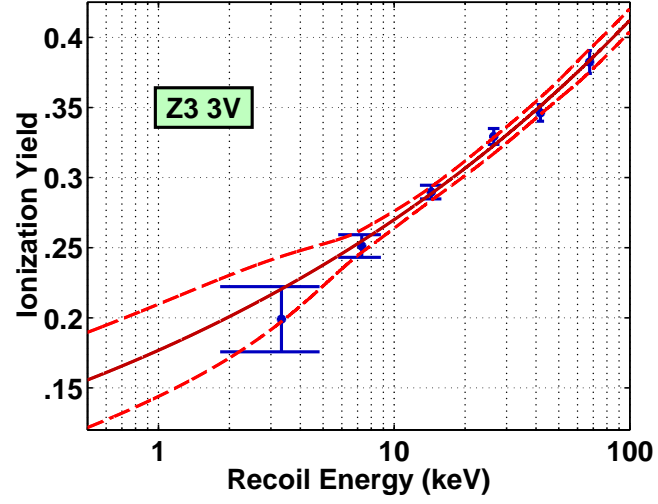


Figure B.5: The nuclear-recoil band centroid and associated error envelope for the Z3 3V data.

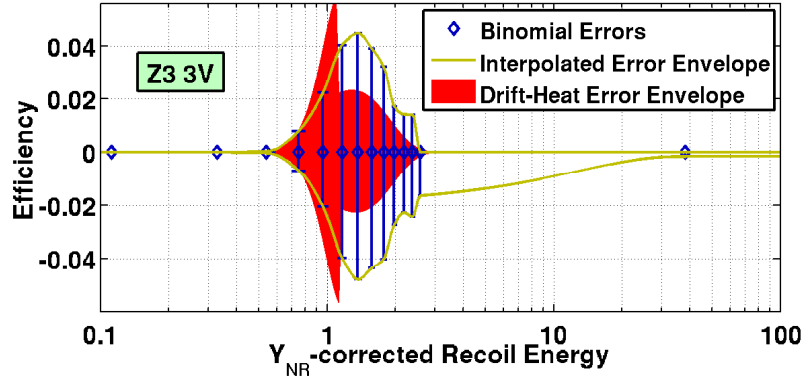


Figure B.6: Trigger-efficiency statistical uncertainties for the Z3 3V data.

### B.1. HARDWARE THRESHOLDS

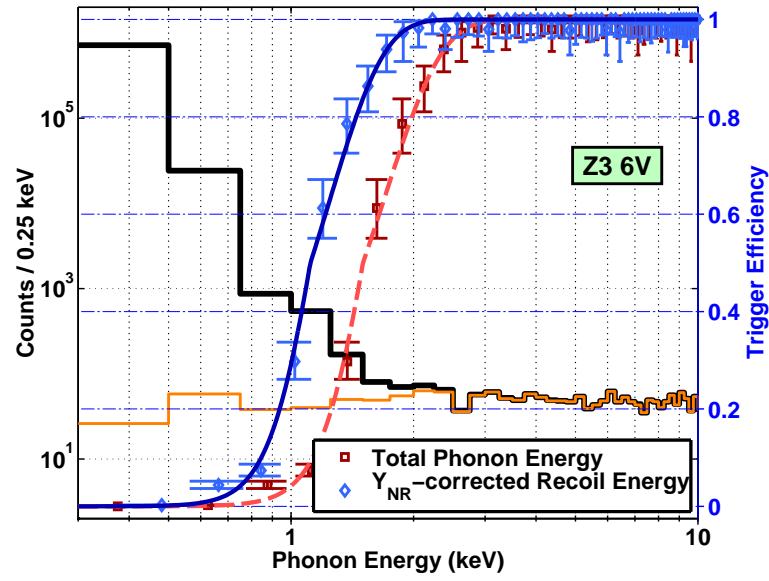


Figure B.7: Trigger-efficiency estimate for the Z3 6V data.

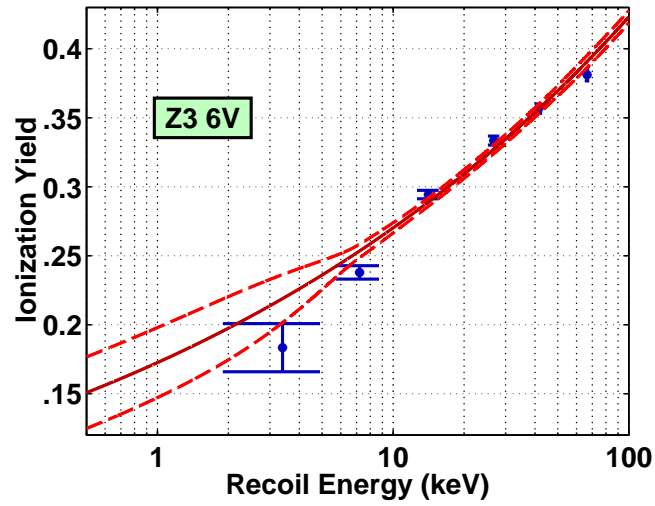


Figure B.8: The nuclear-recoil band centroid and associated error envelope for the Z3 6V data.

## APPENDIX B. THRESHOLD EFFICIENCIES

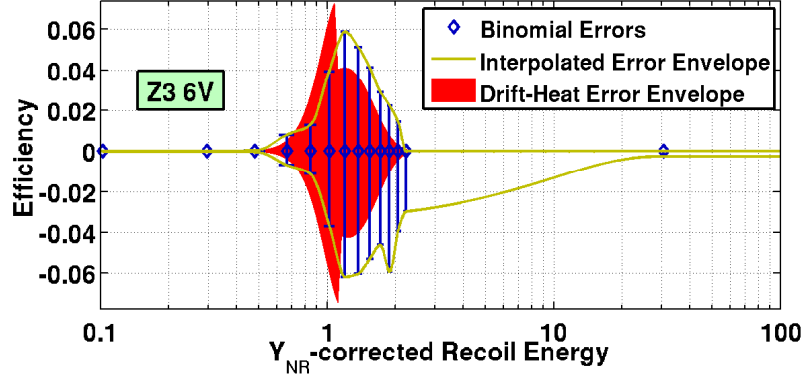


Figure B.9: Trigger-efficiency statistical uncertainties for the Z3 6V data.

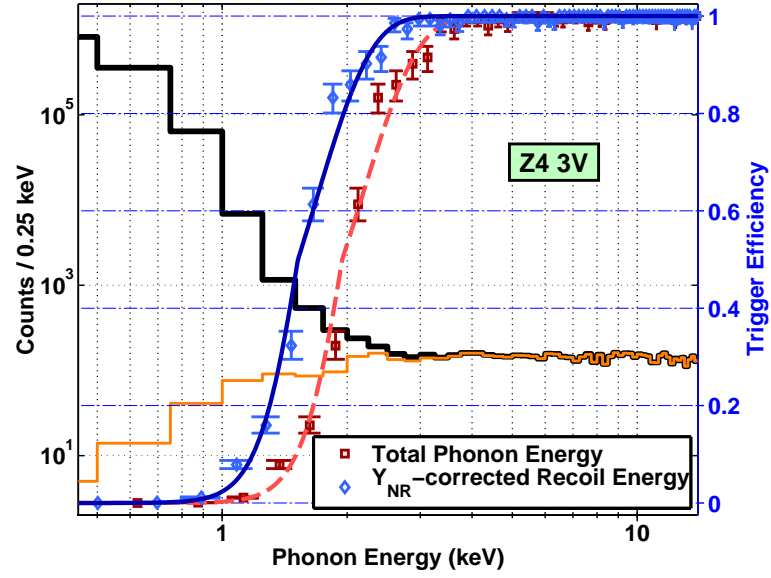


Figure B.10: Trigger-efficiency estimate for the Z4 3V data.

### B.1. HARDWARE THRESHOLDS

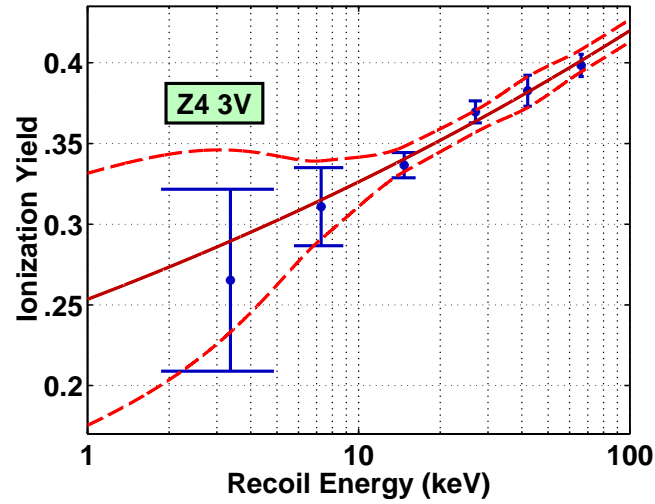


Figure B.11: The nuclear-recoil band centroid and associated error envelope for the Z4 3V data.

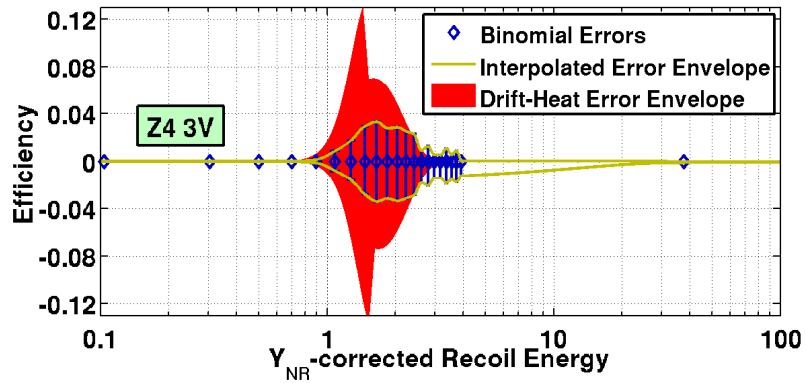


Figure B.12: Trigger-efficiency statistical uncertainties for the Z4 3V data.

## APPENDIX B. THRESHOLD EFFICIENCIES

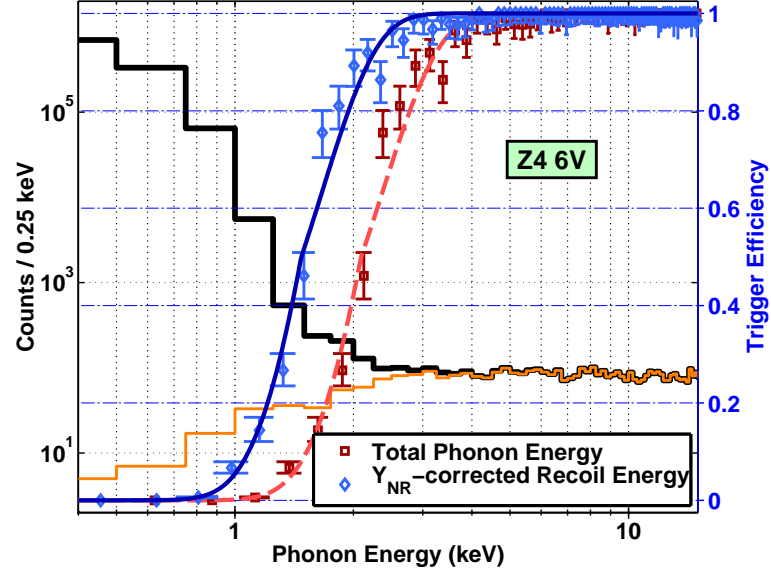


Figure B.13: Trigger-efficiency estimate for the Z4 6V data.

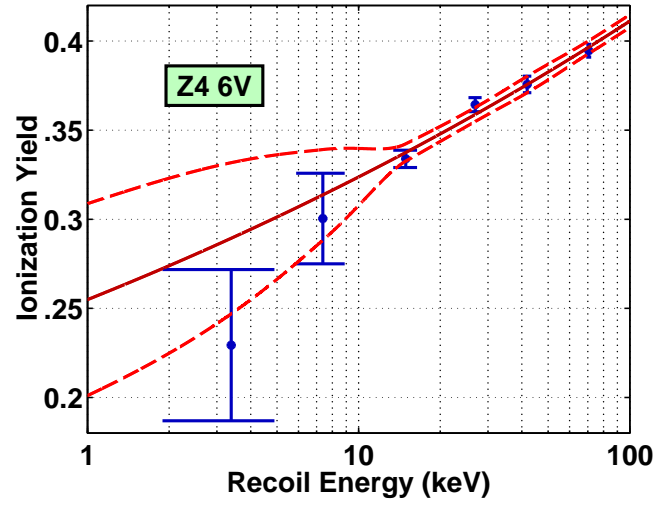


Figure B.14: The nuclear-recoil band centroid and associated error envelope for the Z4 6V data.

## B.1. HARDWARE THRESHOLDS

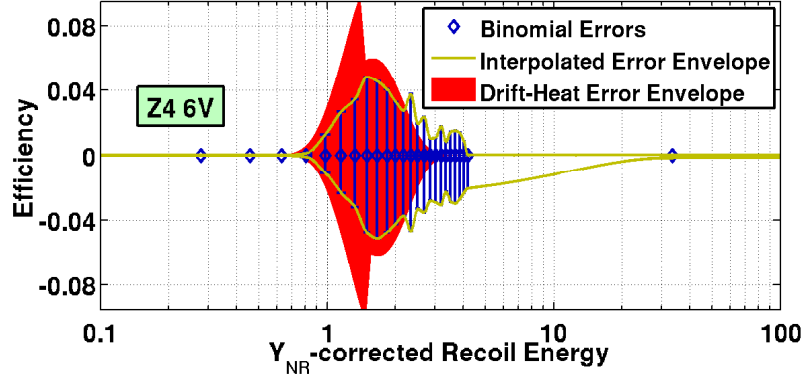


Figure B.15: Trigger-efficiency statistical uncertainties for the Z4 6V data.

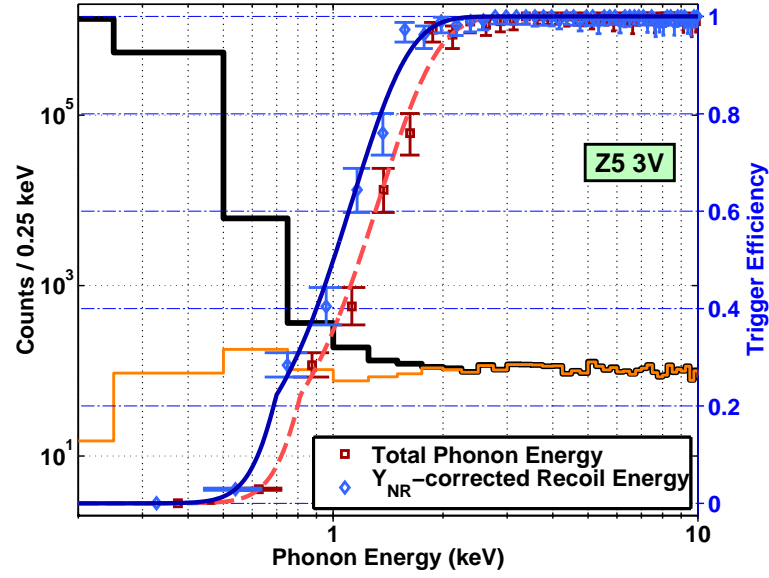


Figure B.16: Trigger-efficiency estimate for the Z5 3V data.

## APPENDIX B. THRESHOLD EFFICIENCIES

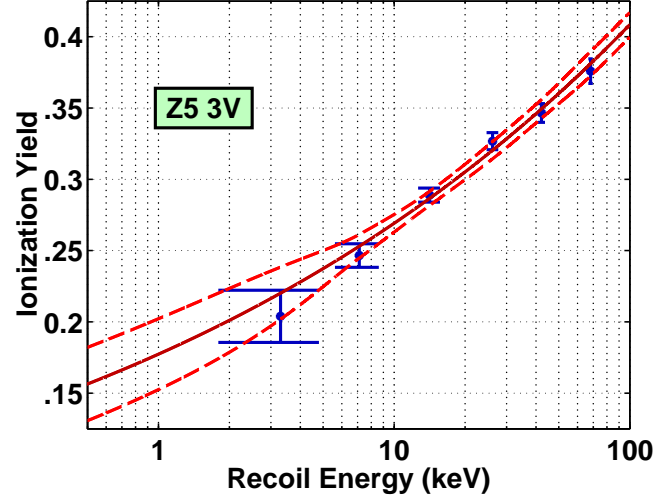


Figure B.17: The nuclear-recoil band centroid and associated error envelope for the Z5 3V data.

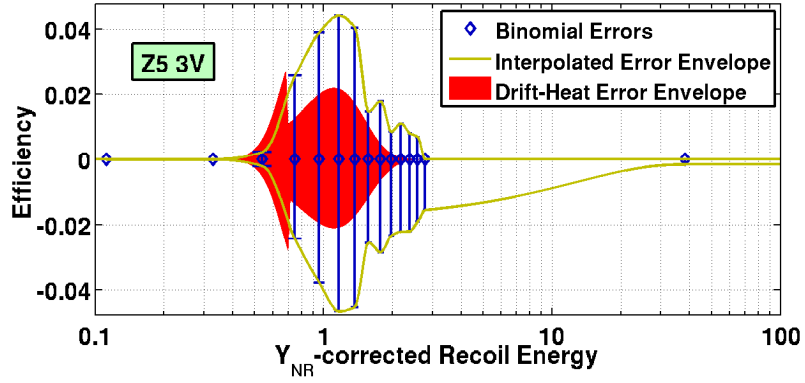


Figure B.18: Trigger-efficiency statistical uncertainties for the Z5 3V data.

### B.1. HARDWARE THRESHOLDS

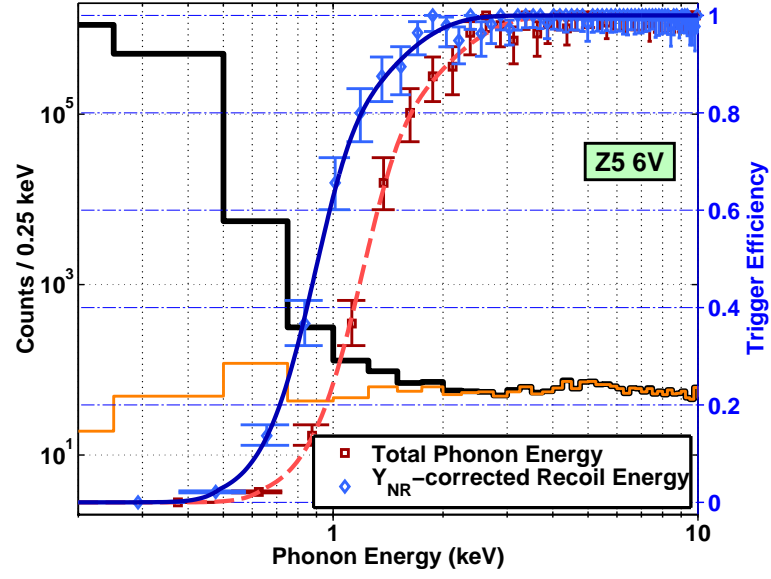


Figure B.19: Trigger-efficiency estimate for the Z5 6V data.

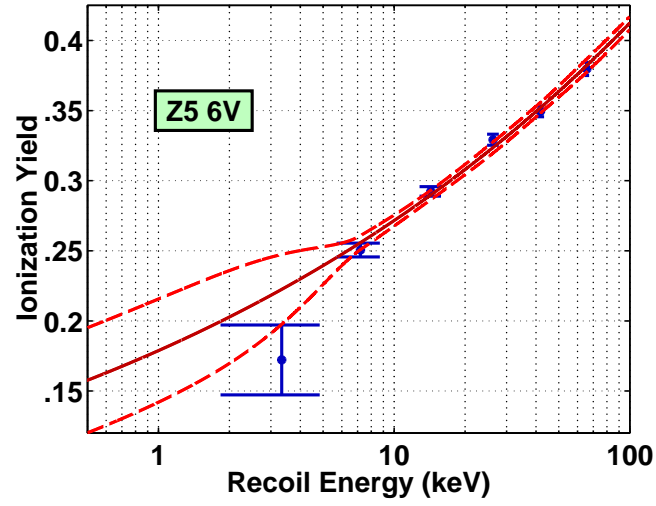


Figure B.20: The nuclear-recoil band centroid and associated error envelope for the Z5 6V data.



## APPENDIX B. THRESHOLD EFFICIENCIES

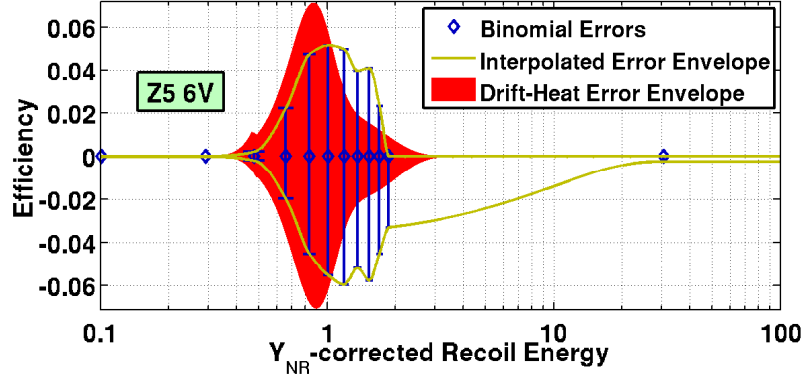


Figure B.21: Trigger-efficiency statistical uncertainties for the Z5 6V data.

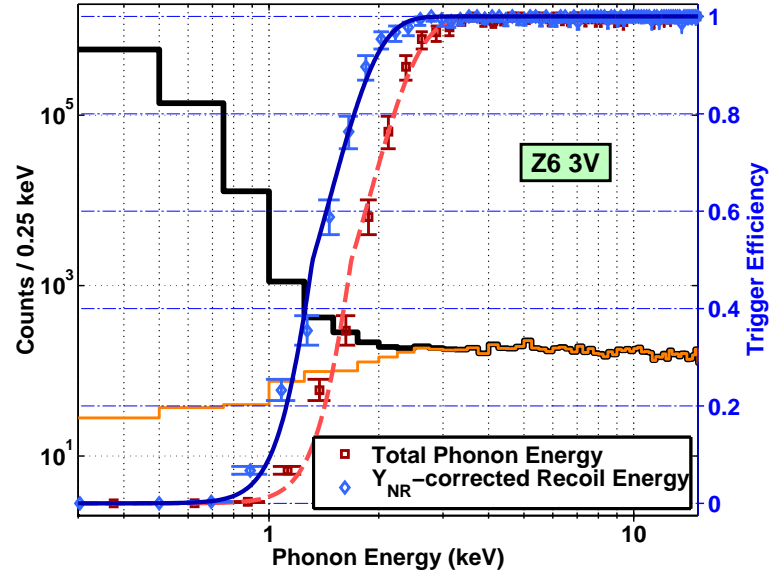


Figure B.22: Trigger-efficiency estimate for the Z6 3V data.

### B.1. HARDWARE THRESHOLDS

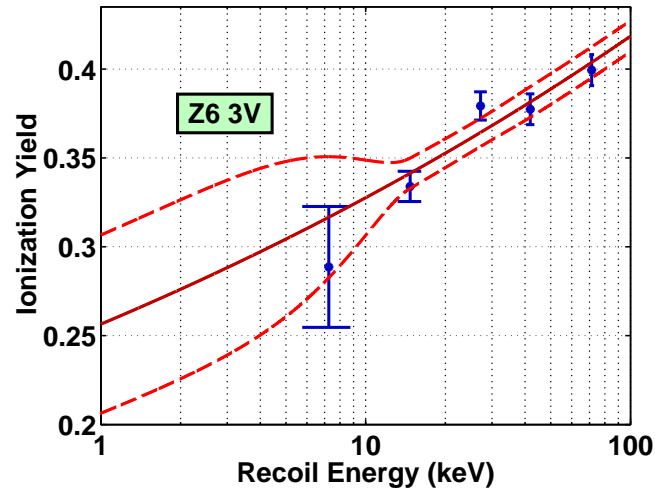


Figure B.23: The nuclear-recoil band centroid and associated error envelope for the Z6 3V data.

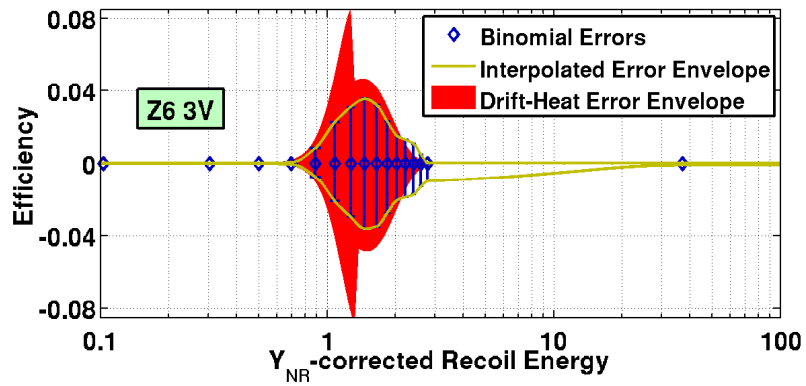


Figure B.24: Trigger-efficiency statistical uncertainties for the Z6 3V data.

## APPENDIX B. THRESHOLD EFFICIENCIES

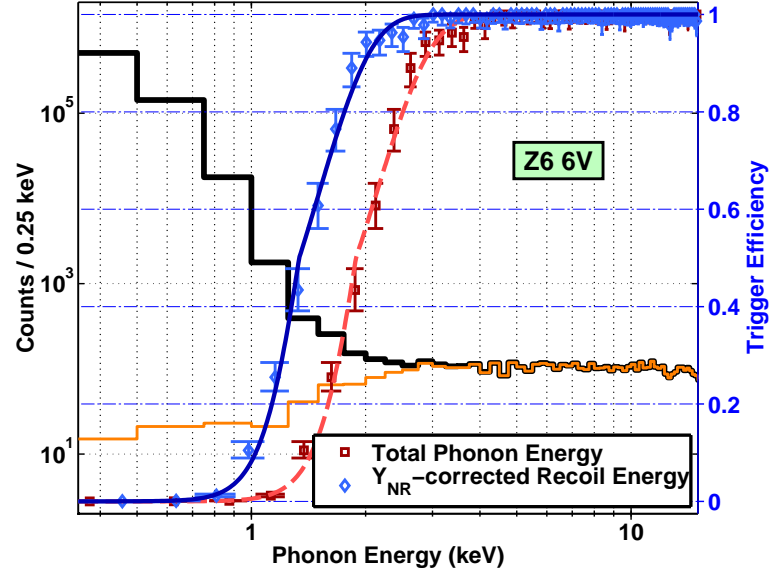


Figure B.25: Trigger-efficiency estimate for the Z6 6V data.

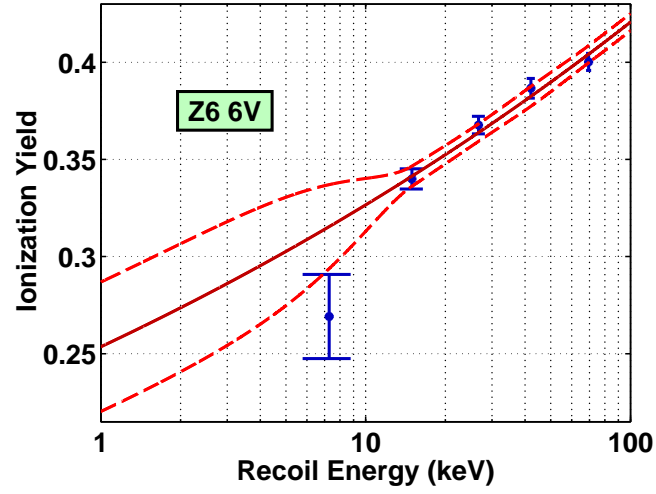


Figure B.26: The nuclear-recoil band centroid and associated error envelope for the Z6 6V data.

## B.2. SOFTWARE THRESHOLDS

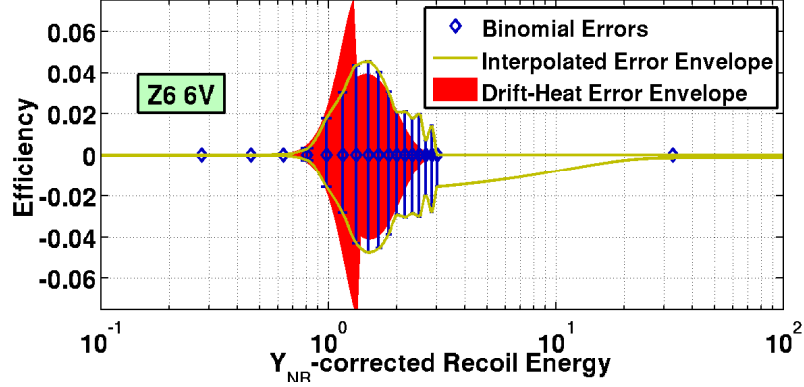


Figure B.27: Trigger-efficiency statistical uncertainties for the Z6 6V data.

## B.2 Software Thresholds

In this section there are two figures for each detector and bias-voltage run (excluding the figures for the Z4 3V data, which appear in Section 5.2.2):

- Time dependence of the  $6\sigma$  phonon-energy software threshold, including Z1 (equivalent to Figure 5.5); and
- Estimate of the software threshold efficiency and statistical uncertainty for the viable low-threshold detectors (*i.e.*, not including Z1—equivalent to Figure 5.6).

## APPENDIX B. THRESHOLD EFFICIENCIES

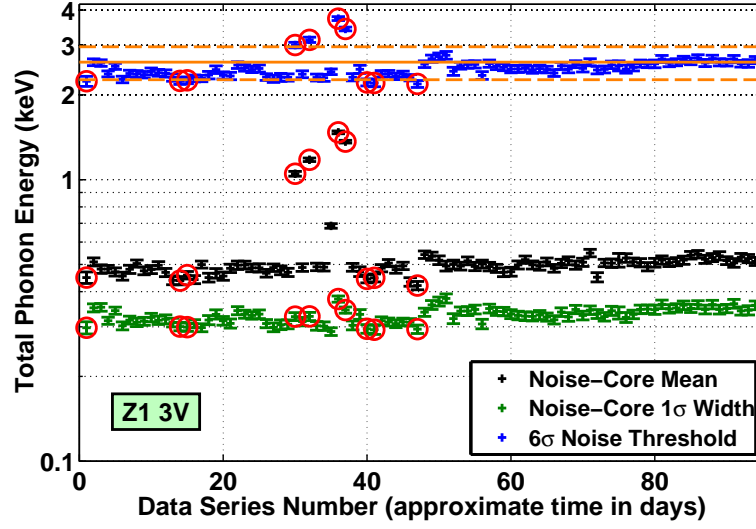


Figure B.28: Time dependence of the phonon-energy software threshold for the Z1 3V data.

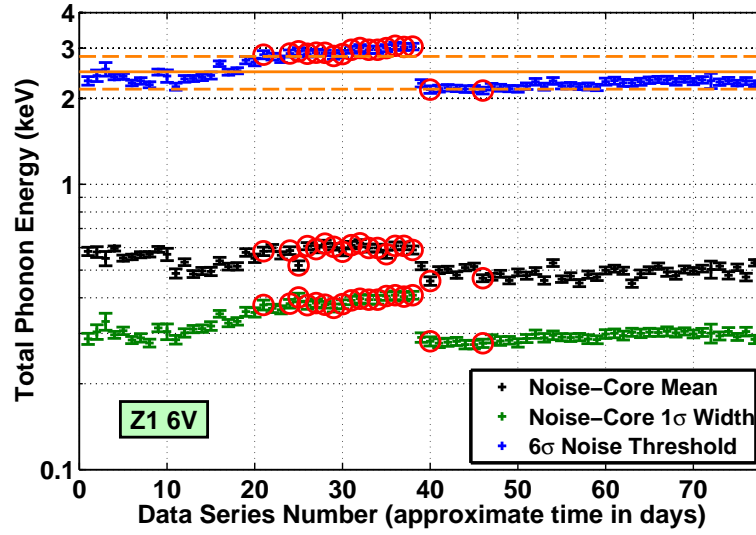


Figure B.29: Time dependence of the phonon-energy software threshold for the Z1 6V data.

## B.2. SOFTWARE THRESHOLDS

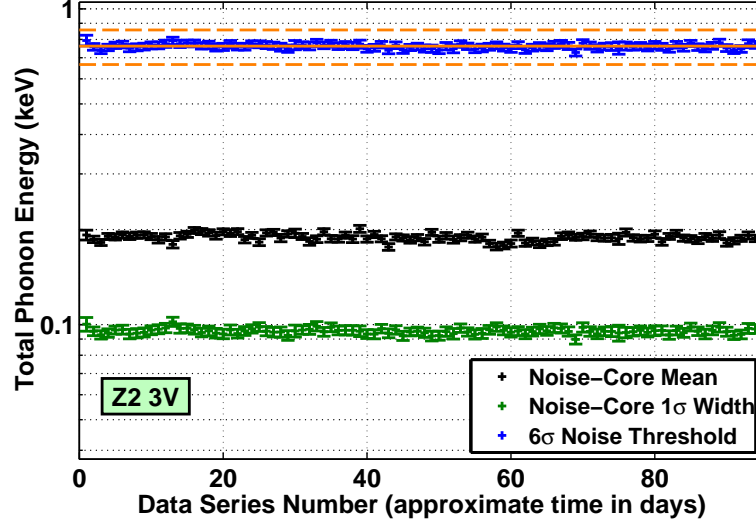


Figure B.30: Time dependence of the phonon-energy software threshold for the Z2 3V data.

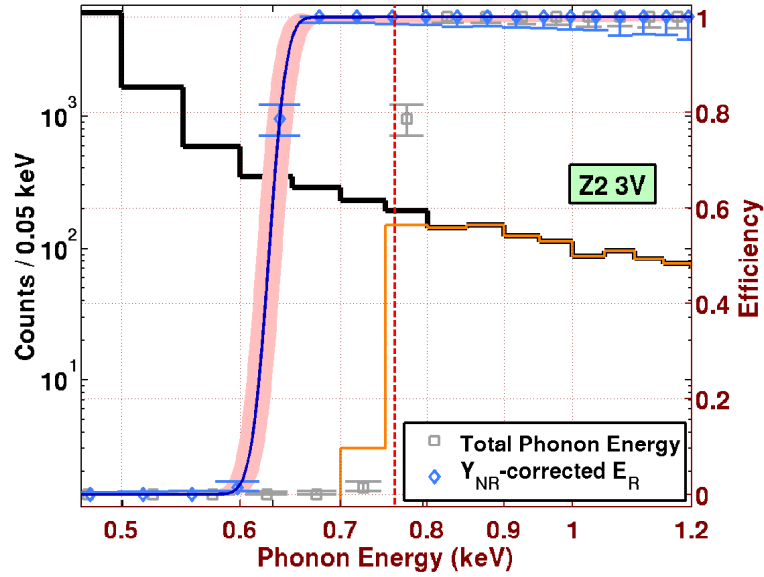


Figure B.31: Estimate of the phonon-energy software threshold efficiency for the Z2 3V data.

## APPENDIX B. THRESHOLD EFFICIENCIES

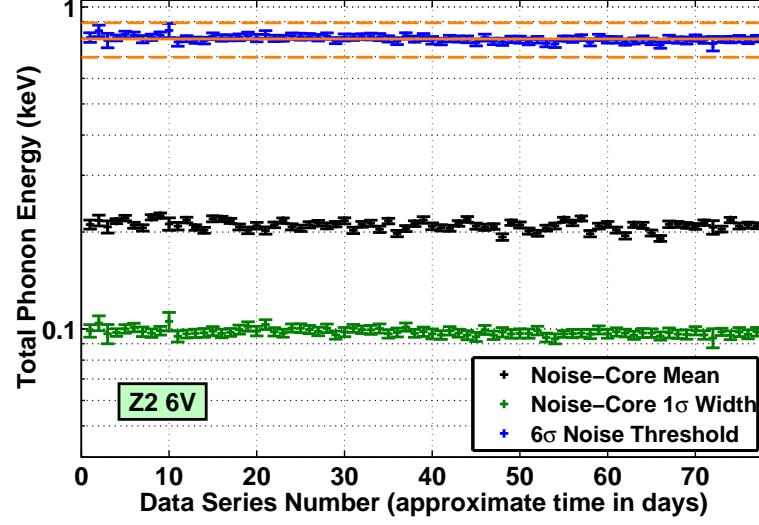


Figure B.32: Time dependence of the phonon-energy software threshold for the Z2 6V data.

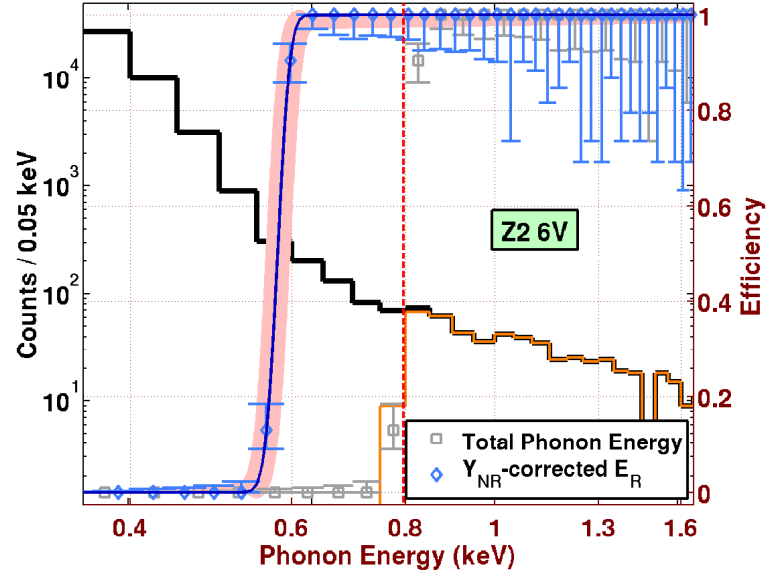


Figure B.33: Estimate of the phonon-energy software threshold efficiency for the Z2 6V data. The statistics are low due to the event-burst cut.

## B.2. SOFTWARE THRESHOLDS

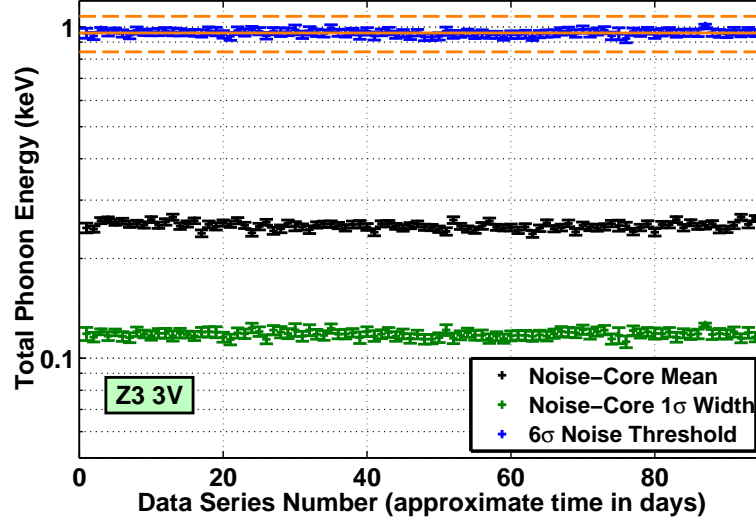


Figure B.34: Time dependence of the phonon-energy software threshold for the Z3 3V data.

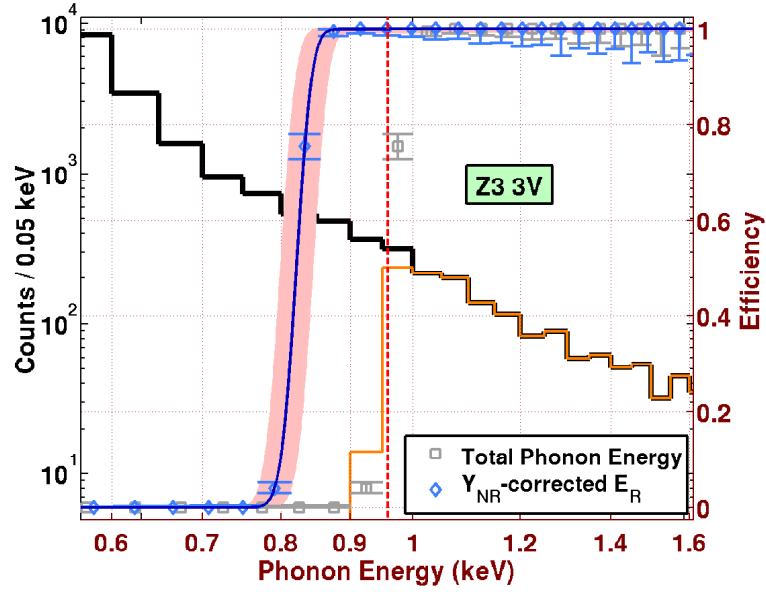


Figure B.35: Estimate of the phonon-energy software threshold efficiency for the Z3 3V data.



## APPENDIX B. THRESHOLD EFFICIENCIES

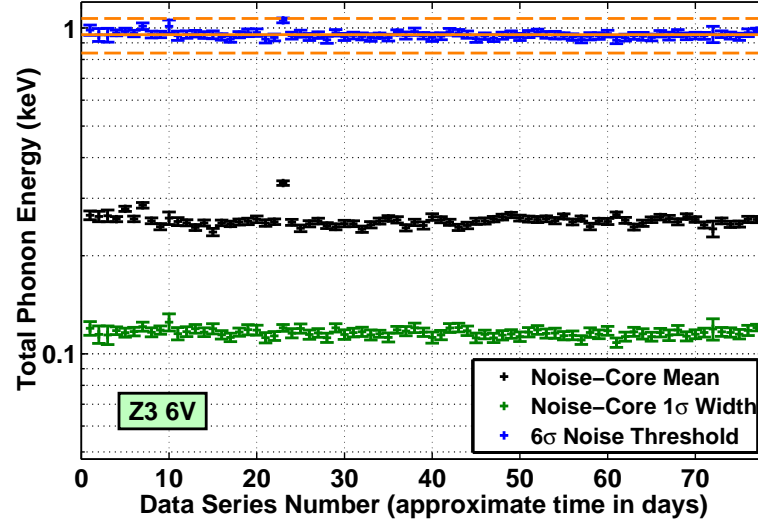


Figure B.36: Time dependence of the phonon-energy software threshold for the Z3 6V data.

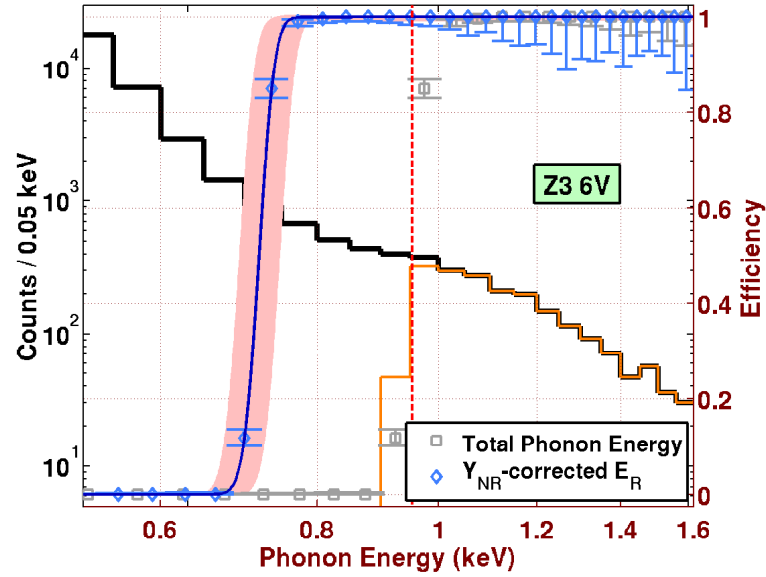


Figure B.37: Estimate of the phonon-energy software threshold efficiency for the Z3 6V data.

## B.2. SOFTWARE THRESHOLDS

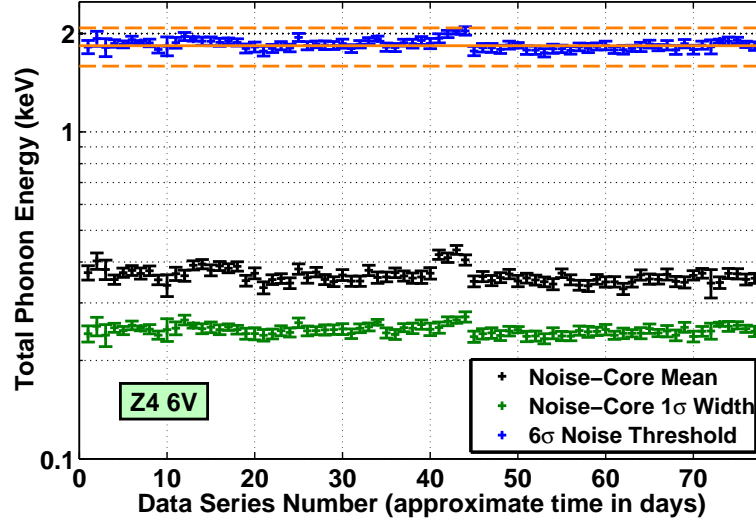


Figure B.38: Time dependence of the phonon-energy software threshold for the Z4 6V data.

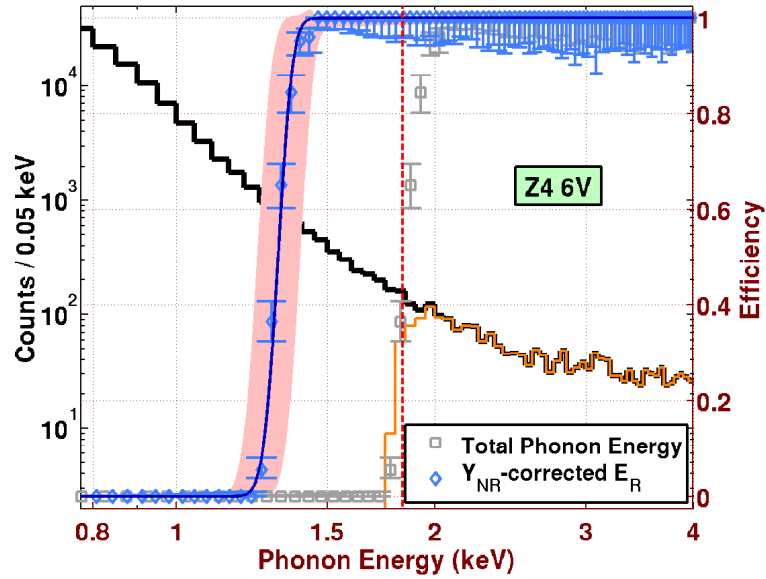


Figure B.39: Estimate of the phonon-energy software threshold efficiency for the Z4 6V data.

## APPENDIX B. THRESHOLD EFFICIENCIES

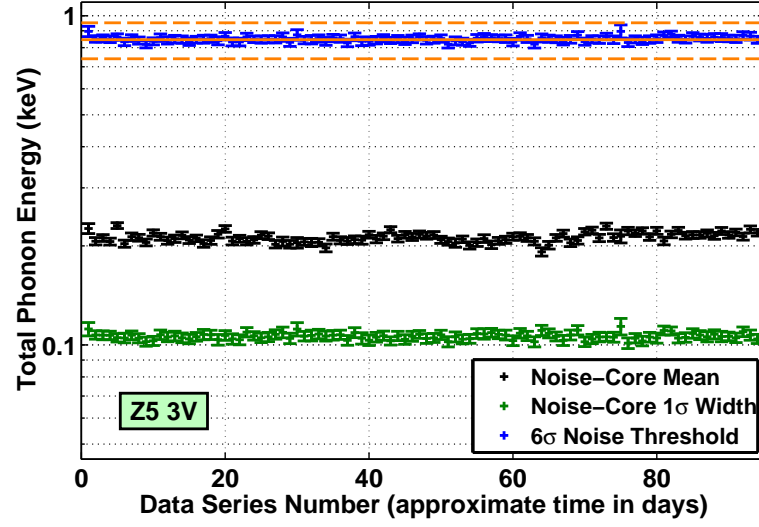


Figure B.40: Time dependence of the phonon-energy software threshold for the Z5 3V data.

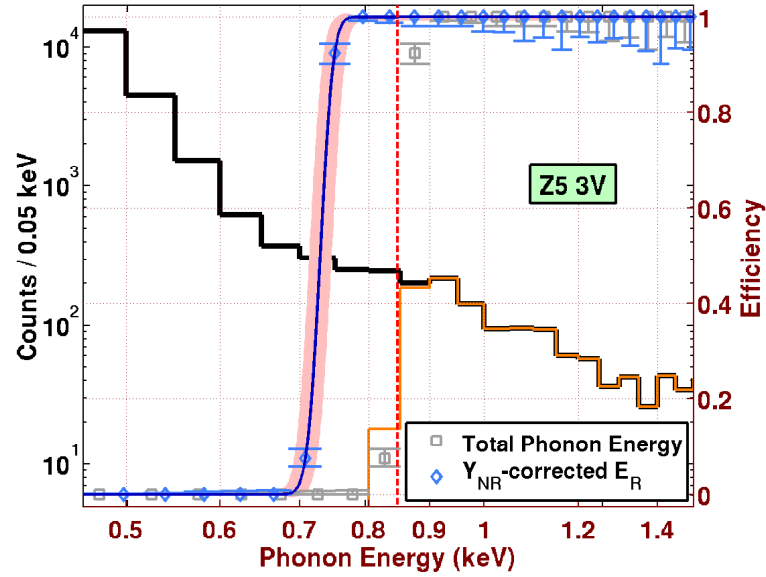


Figure B.41: Estimate of the phonon-energy software threshold efficiency for the Z5 3V data.

## B.2. SOFTWARE THRESHOLDS

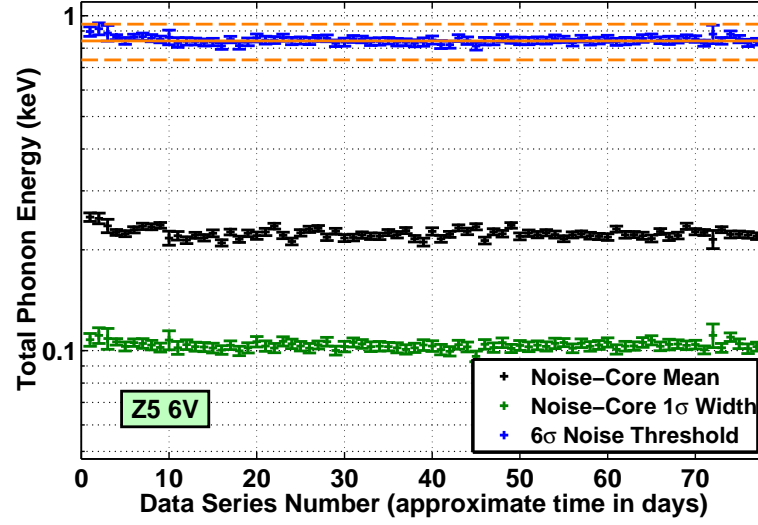


Figure B.42: Time dependence of the phonon-energy software threshold for the Z5 6V data.

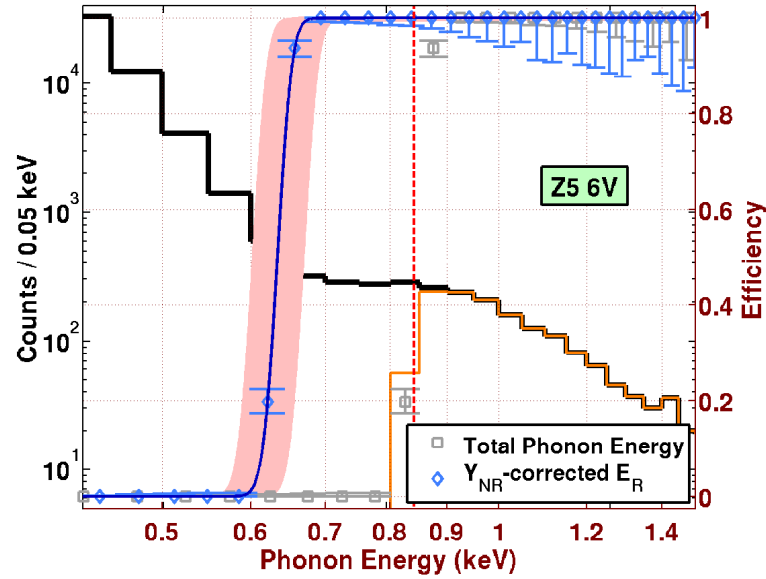


Figure B.43: Estimate of the phonon-energy software threshold efficiency for the Z5 6V data.

## APPENDIX B. THRESHOLD EFFICIENCIES

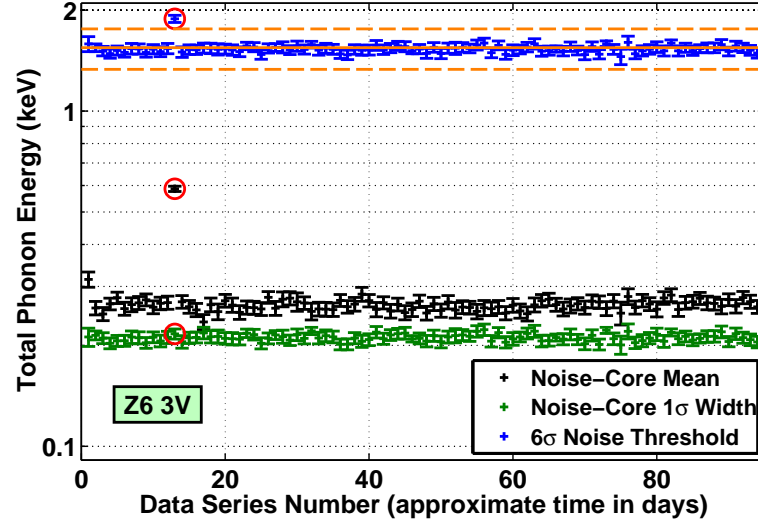


Figure B.44: Time dependence of the phonon-energy software threshold for the Z6 3V data.

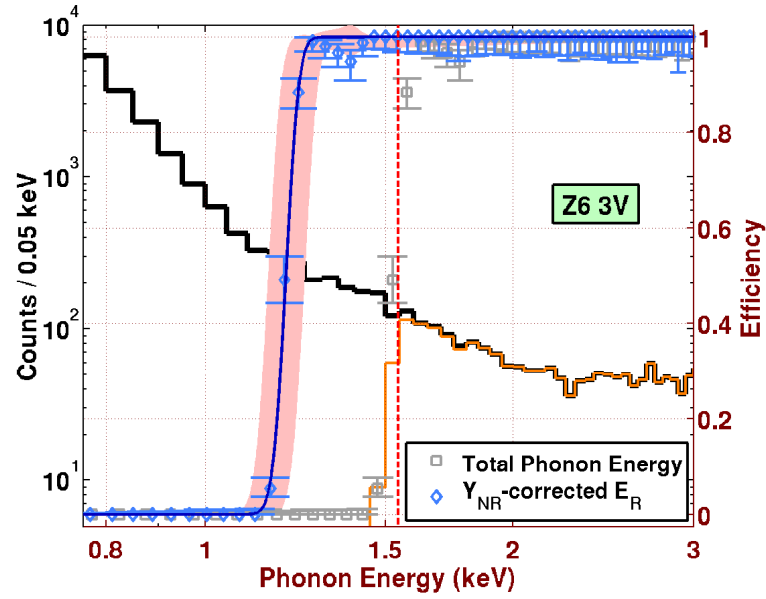


Figure B.45: Estimate of the phonon-energy software threshold efficiency for the Z6 3V data.

## B.2. SOFTWARE THRESHOLDS

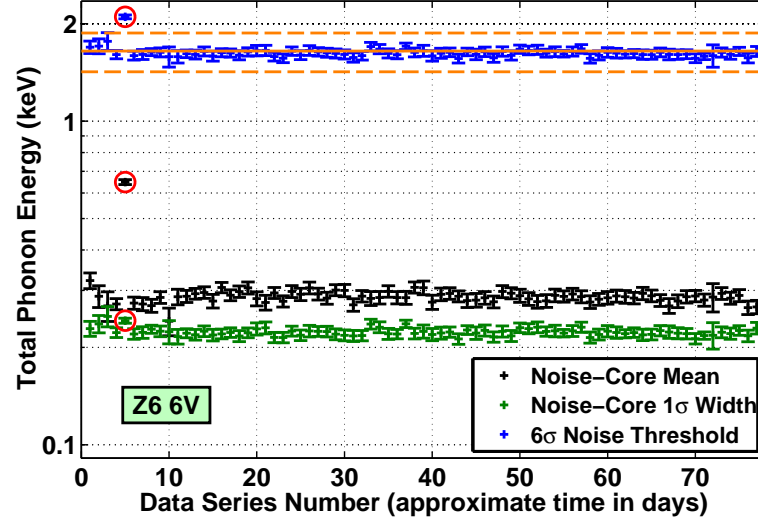


Figure B.46: Time dependence of the phonon-energy software threshold for the Z6 6V data.

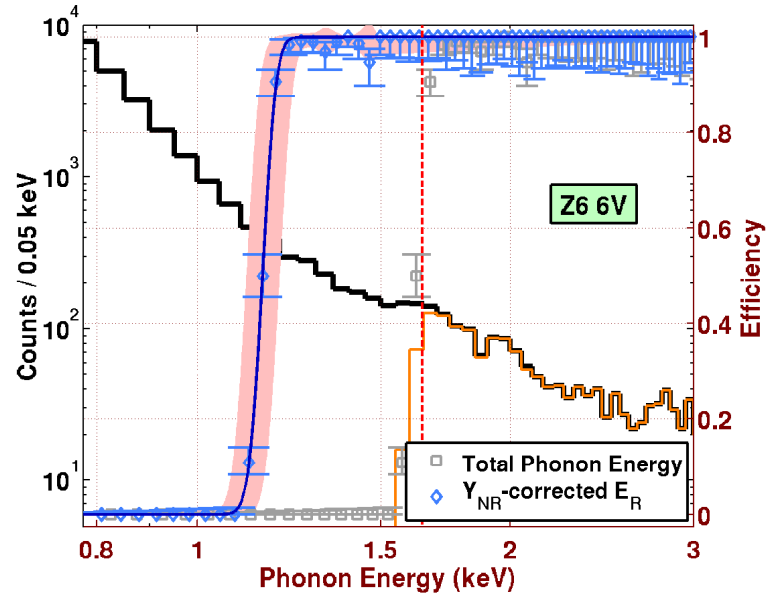


Figure B.47: Estimate of the phonon-energy software threshold efficiency for the Z6 6V data.



# Appendix C

## Analysis Cuts

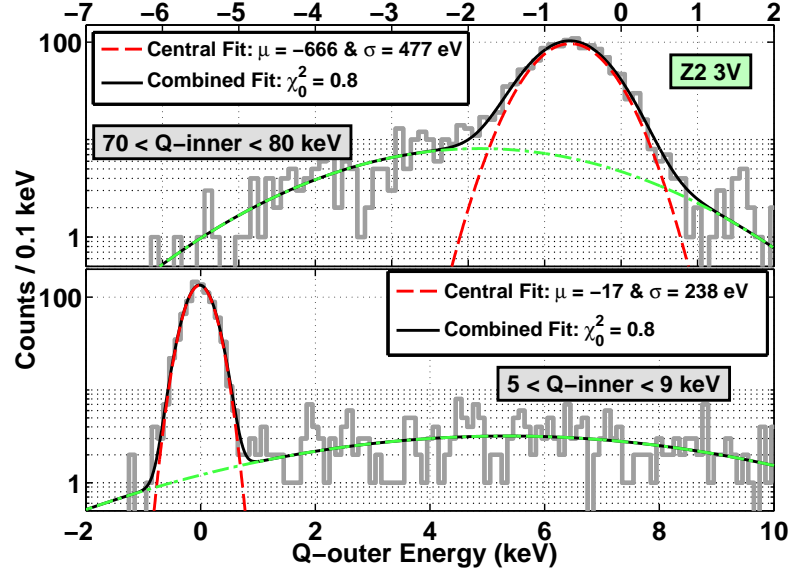
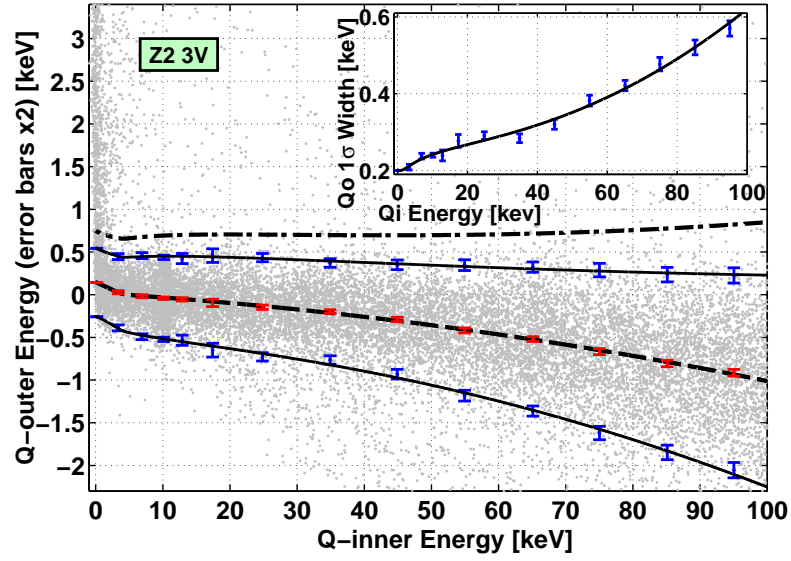
This appendix includes figures associated with the fiducial-volume and nuclear-recoil cuts, and estimates of their associated detection efficiencies, for the detectors not shown in Section 5.3. See the main text and figure captions in that section for further details.

### C.1 Fiducial Volume

There are three figures for each viable low-threshold detector and bias-voltage run (excluding the figures for the Z6 6V data, which appear in Section 5.3.3):

- Fits to the distribution of  $Q$ -outer ( $Q_o$ ) noise events for example low- and high-energy bins of  $Q$ -inner ( $Q_i$ ) energy (equivalent to Figure 5.11);
- The  $2\sigma$   $Q_o$  noise band (equivalent to Figure 5.12); and
- Estimate of the cut's selection efficiency (equivalent to Figure 5.13).



Figure C.1: Example fits to  $Q_o$  noise distributions for Z2 3V data.Figure C.2: The  $Q_o$   $2\sigma$  noise band for the Z2 3V data.

### C.1. FIDUCIAL VOLUME

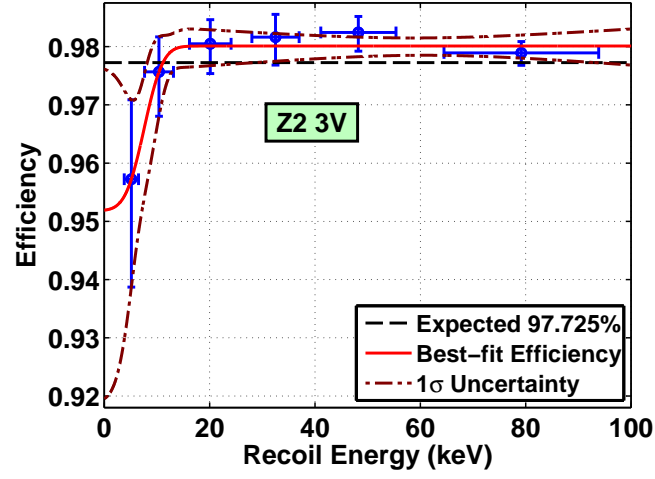


Figure C.3: Estimate of the fiducial-volume cut's selection efficiency for the Z2 3V data.

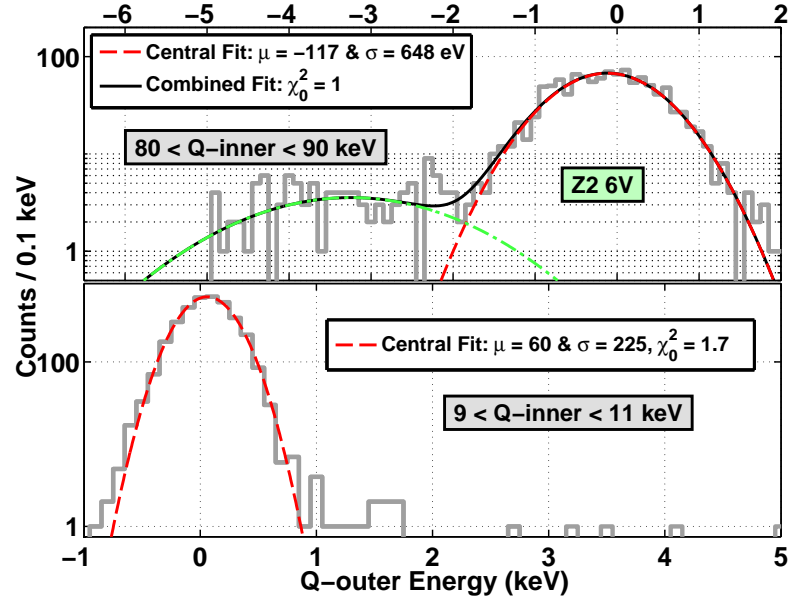


Figure C.4: Example fits to  $Q_o$  noise distributions for Z2 6V data.

# APPENDIX C. ANALYSIS CUTS

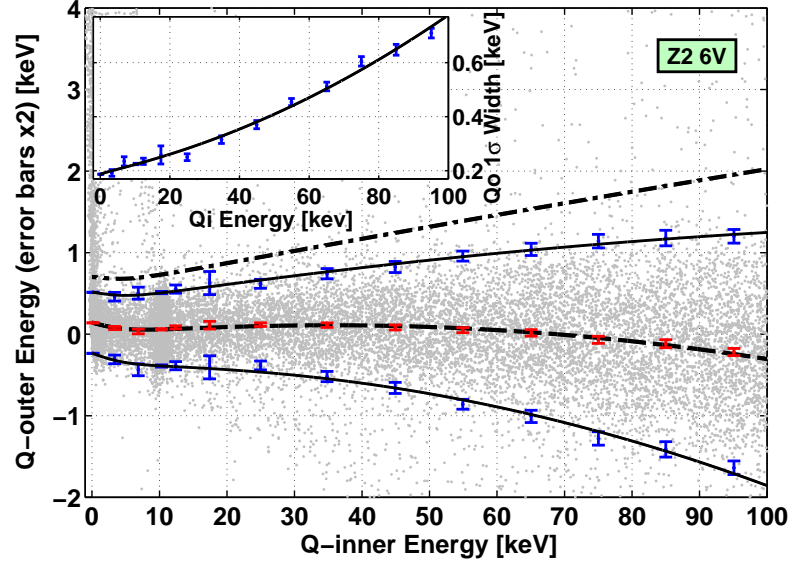


Figure C.5: The  $Q_o$   $2\sigma$  noise band for the Z2 6V data.

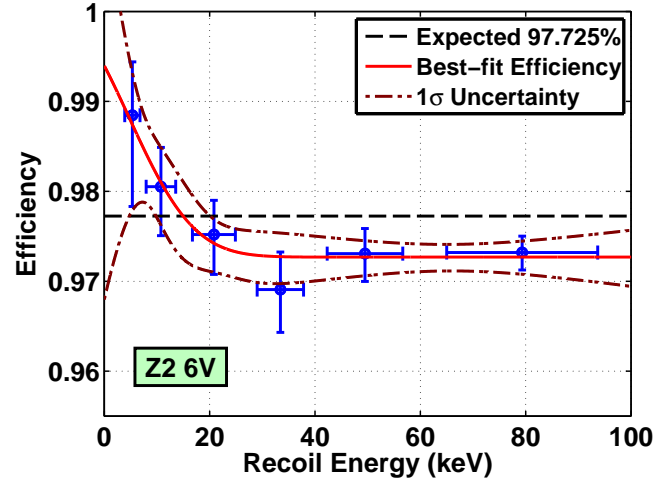


Figure C.6: Estimate of the fiducial-volume cut's selection efficiency for the Z2 6V data.

### C.1. FIDUCIAL VOLUME

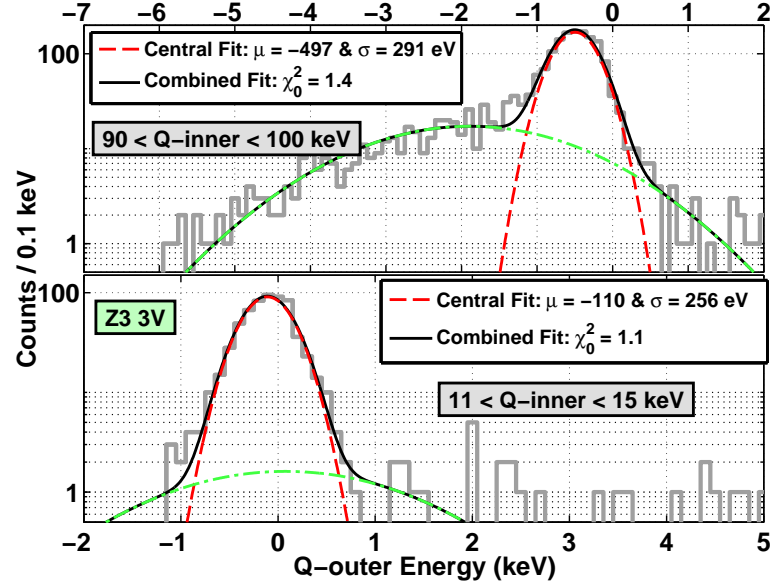


Figure C.7: Example fits to  $Q_o$  noise distributions for Z3 3V data.

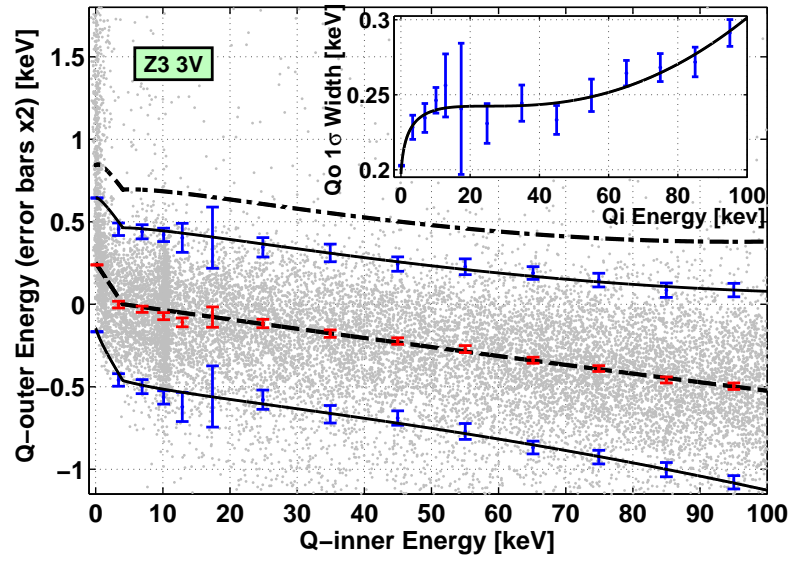


Figure C.8: The  $Q_o$   $2\sigma$  noise band for the Z3 3V data.

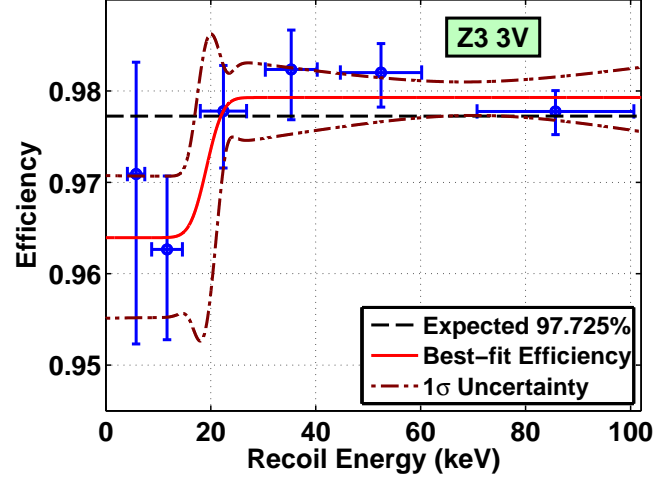


Figure C.9: Estimate of the fiducial-volume cut's selection efficiency for the Z3 3V data.

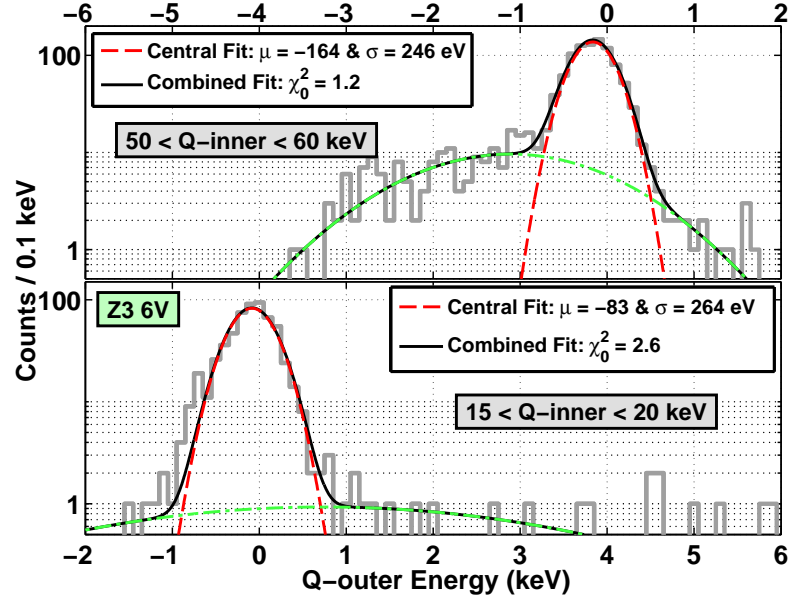


Figure C.10: Example fits to  $Q_o$  noise distributions for Z3 6V data.

### C.1. FIDUCIAL VOLUME

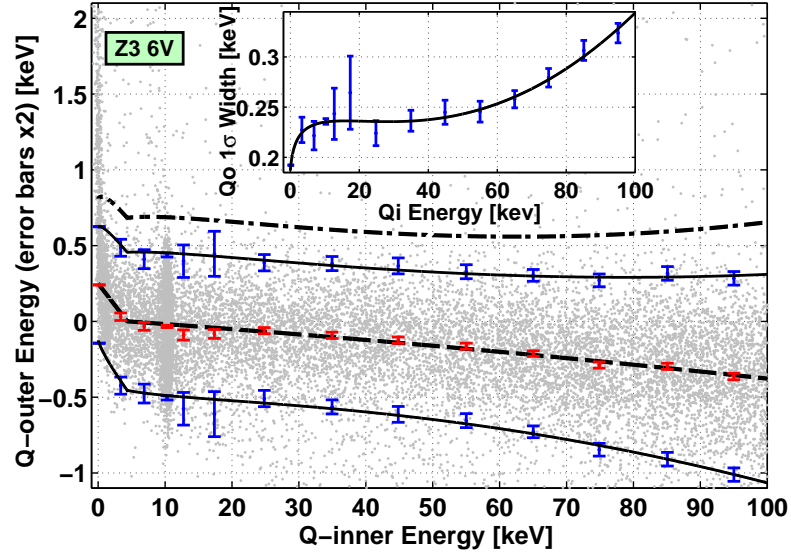


Figure C.11: The  $Q_o$   $2\sigma$  noise band for the Z3 6V data.

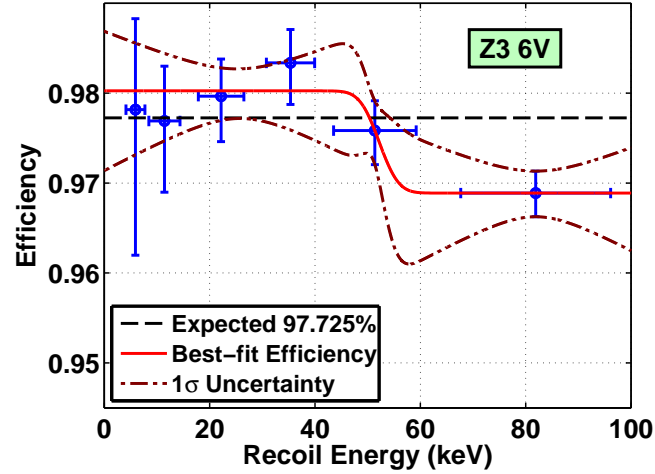
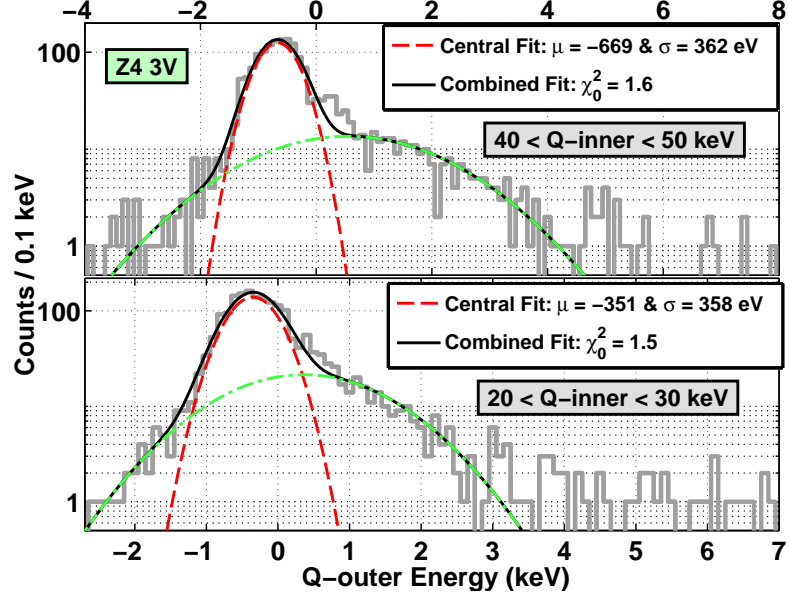
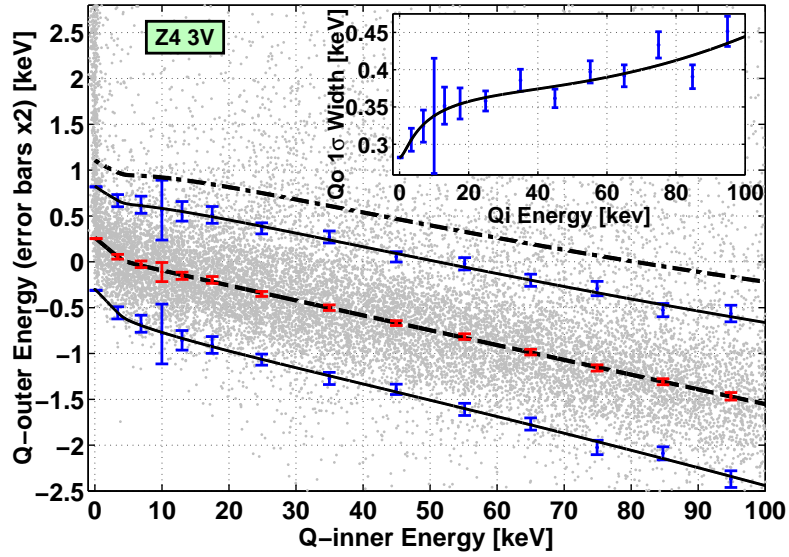


Figure C.12: Estimate of the fiducial-volume cut's selection efficiency for the Z3 6V data.


 Figure C.13: Example fits to  $Q_o$  noise distributions for Z4 3V data.

 Figure C.14: The  $Q_o$   $2\sigma$  noise band for the Z4 3V data.

### C.1. FIDUCIAL VOLUME

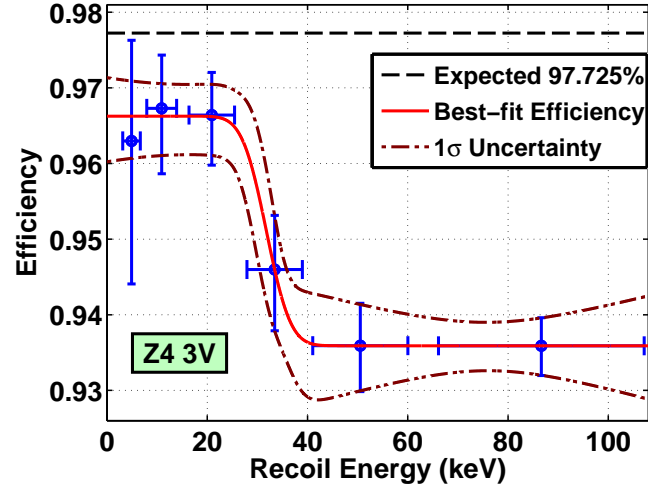


Figure C.15: Estimate of the fiducial-volume cut's selection efficiency for the Z4 3V data.

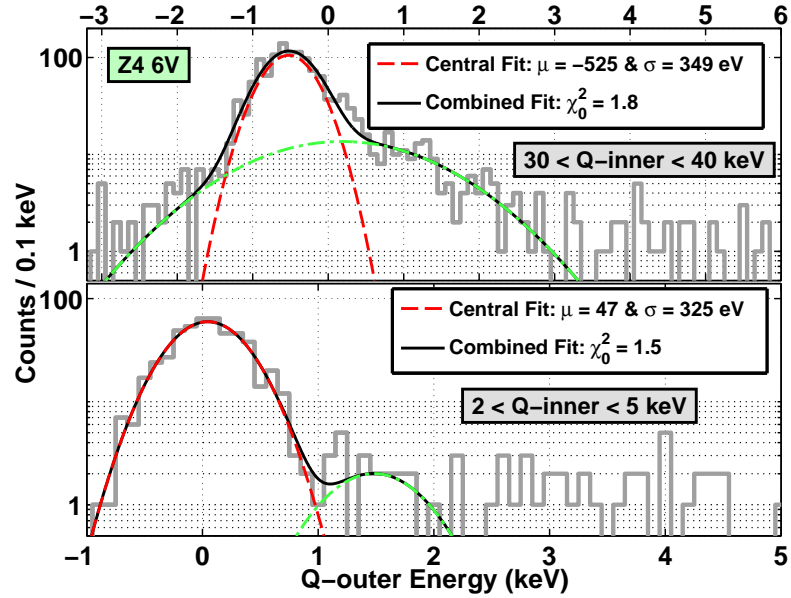


Figure C.16: Example fits to  $Q_o$  noise distributions for Z4 6V data.



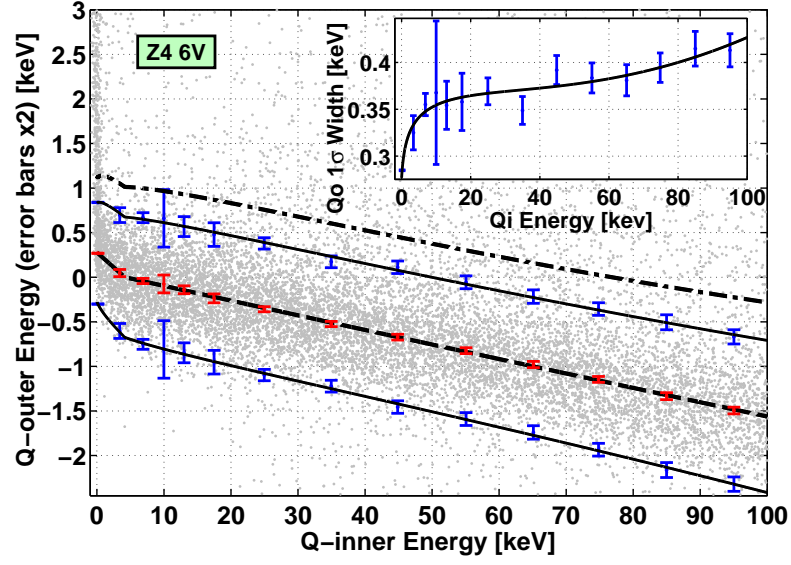
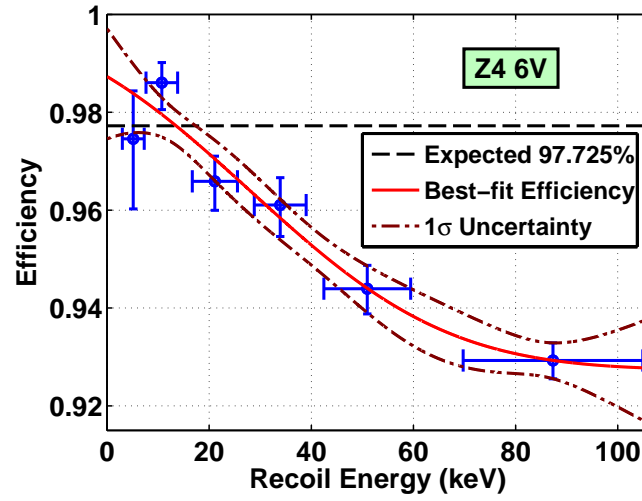

 Figure C.17: The  $Q_o$   $2\sigma$  noise band for the Z4 6V data.


Figure C.18: Estimate of the fiducial-volume cut's selection efficiency for the Z4 6V data.

### C.1. FIDUCIAL VOLUME

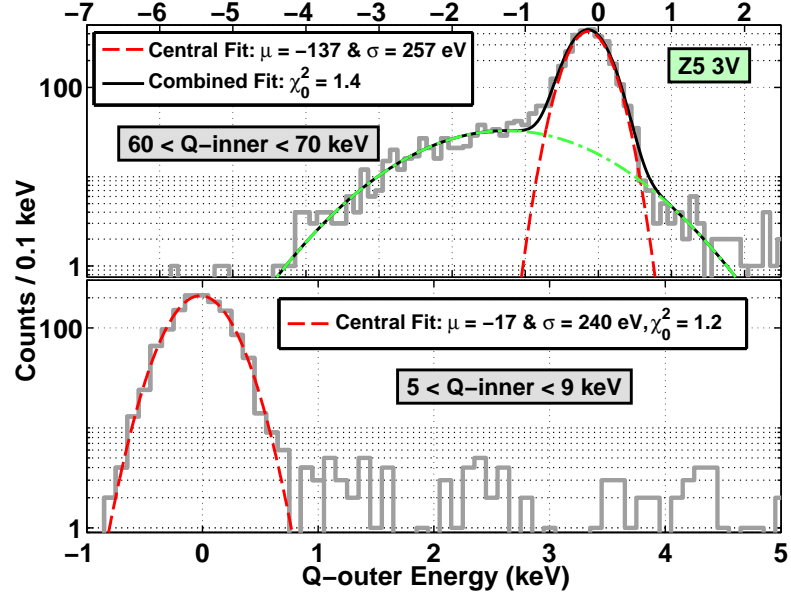


Figure C.19: Example fits to  $Q_o$  noise distributions for Z5 3V data.

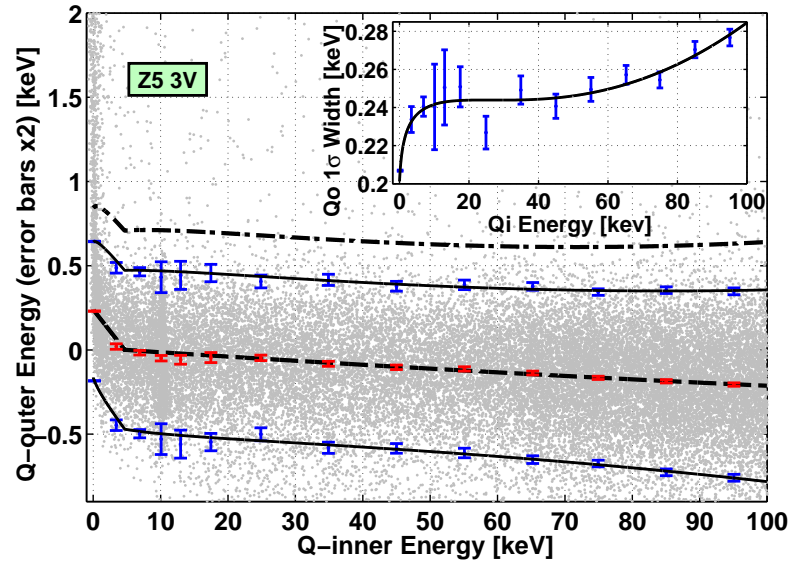


Figure C.20: The  $Q_o$   $2\sigma$  noise band for the Z5 3V data.

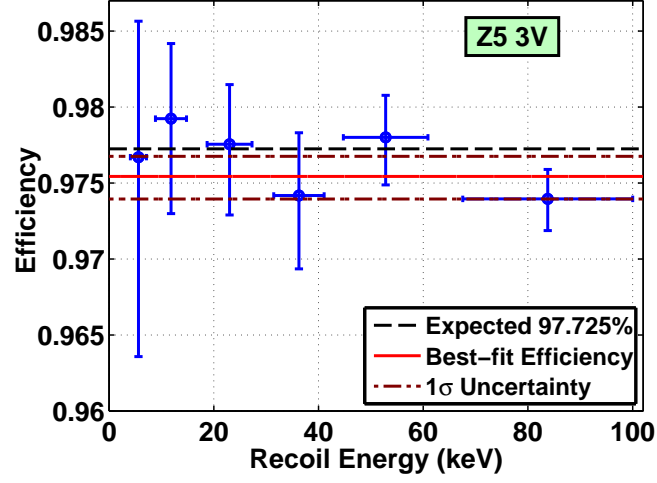


Figure C.21: Estimate of the fiducial-volume cut's selection efficiency for the Z5 3V data.

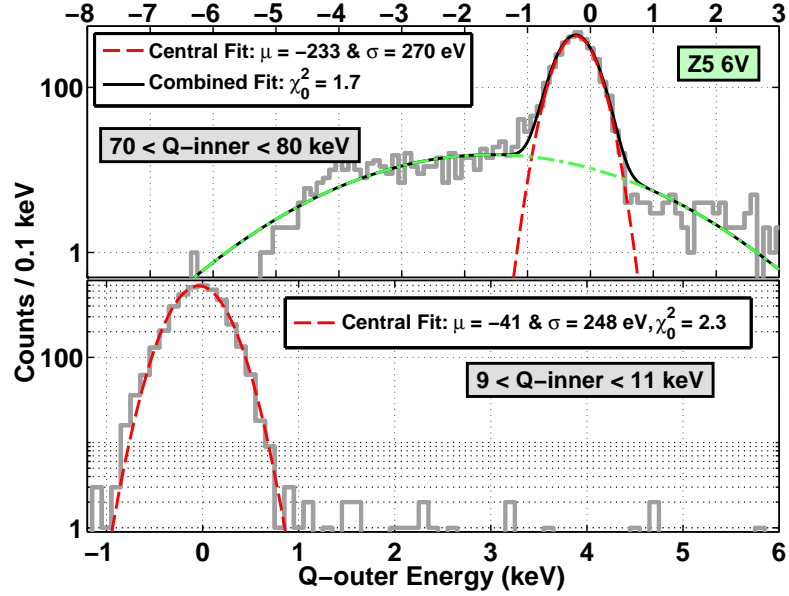


Figure C.22: Example fits to  $Q_o$  noise distributions for Z5 6V data.

### C.1. FIDUCIAL VOLUME

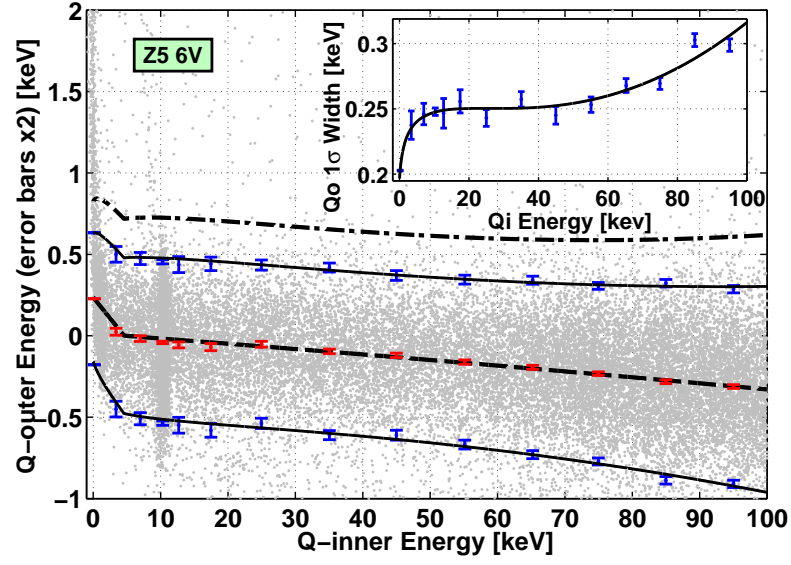


Figure C.23: The  $Q_o$   $2\sigma$  noise band for the Z5 6V data.

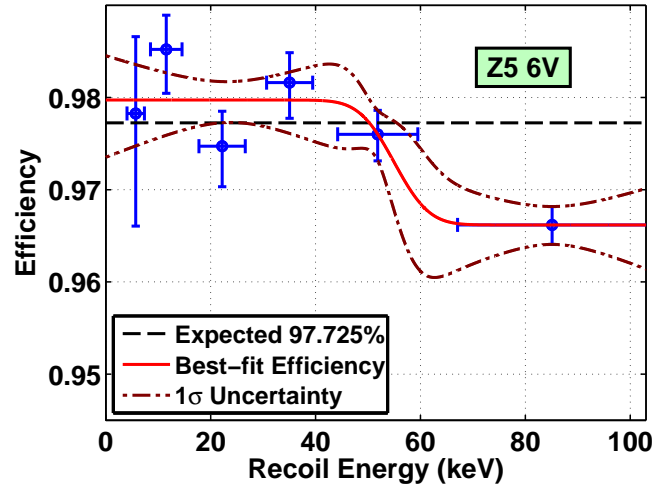


Figure C.24: Estimate of the fiducial-volume cut's selection efficiency for the Z5 6V data.

# APPENDIX C. ANALYSIS CUTS

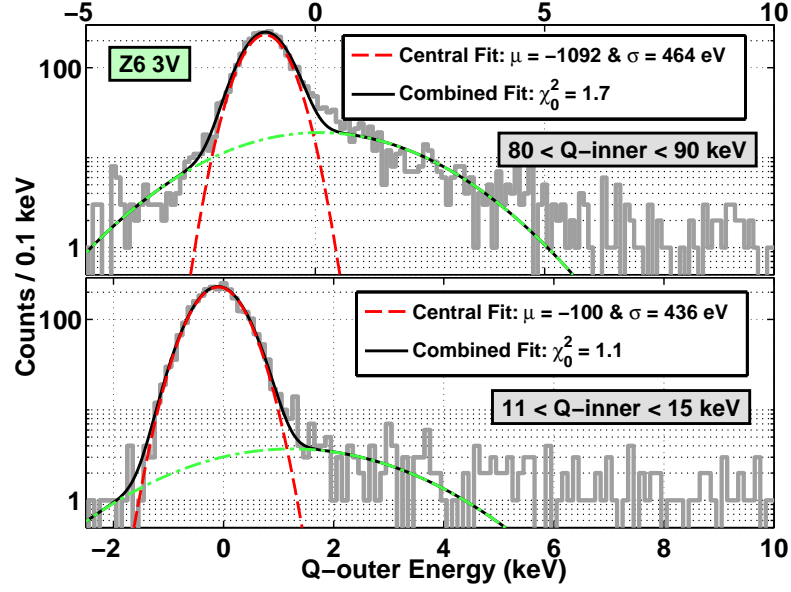


Figure C.25: Example fits to  $Q_o$  noise distributions for Z6 3V data.

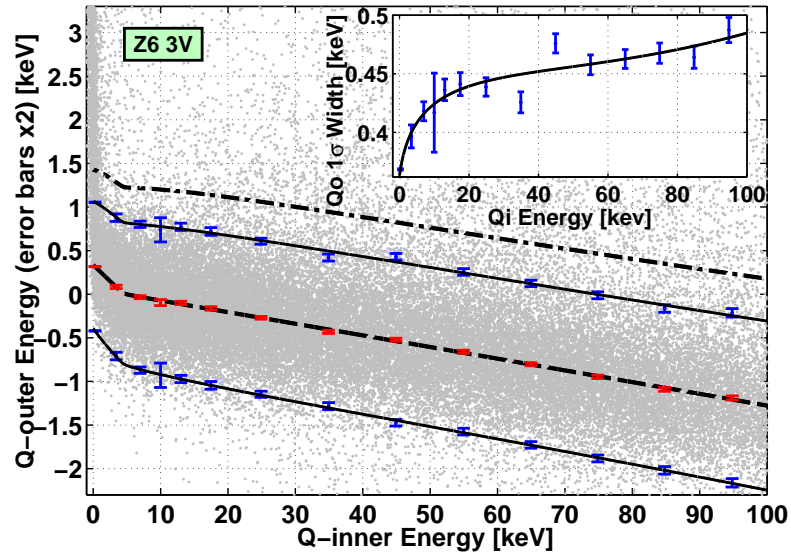


Figure C.26: The  $Q_o$   $2\sigma$  noise band for the Z6 3V data.

## C.2. NUCLEAR-RECOIL BAND

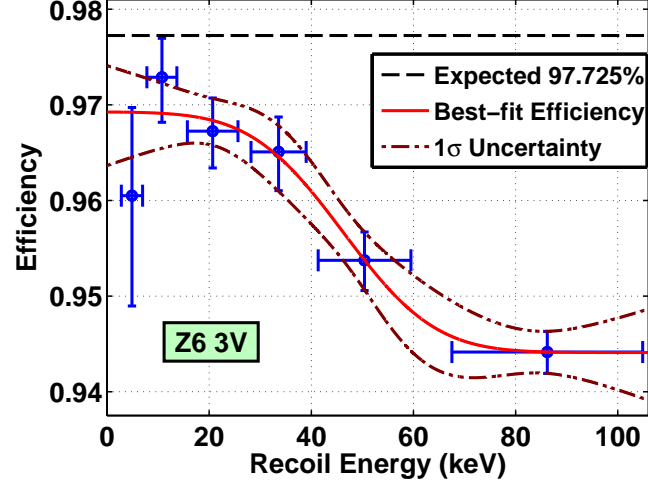


Figure C.27: Estimate of the fiducial-volume cut's selection efficiency for the Z6 3V data.

## C.2 Nuclear-Recoil Band

There are two figures for each viable low-threshold detector and bias-voltage run (excluding the figures for the Z5 6V data, which appear in Section 5.3.6):

- Fits to distributions of ionization yield in bins of recoil energy for  $^{252}\text{Cf}$  nuclear recoils (equivalent to Figure 5.15); and
- Efficiency estimate for the  $2\sigma$  nuclear-recoil band (equivalent to Figure 5.17).

# APPENDIX C. ANALYSIS CUTS

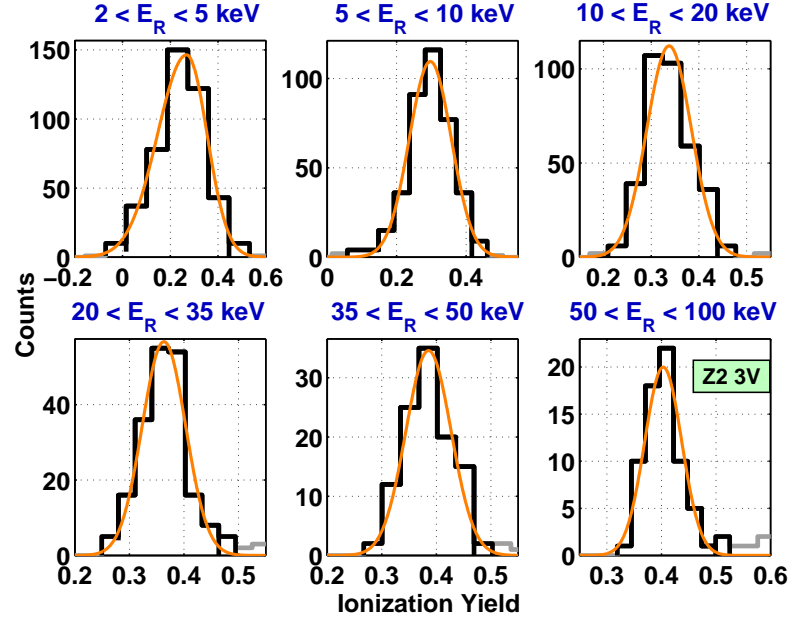


Figure C.28: Binwise estimate of  $\langle Y \rangle$  and  $\sigma_Y$  for Z2 3V nuclear recoils.

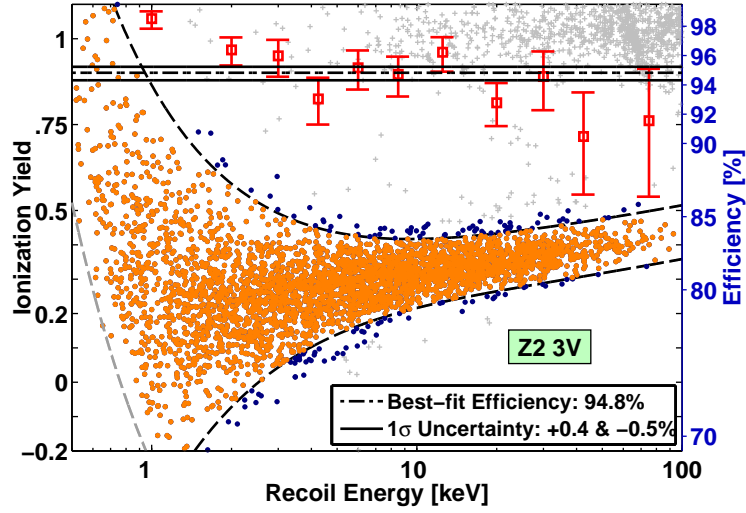


Figure C.29:  $2\sigma$  nuclear-recoil band selection efficiency for Z2 3V data.

## C.2. NUCLEAR-RECOIL BAND

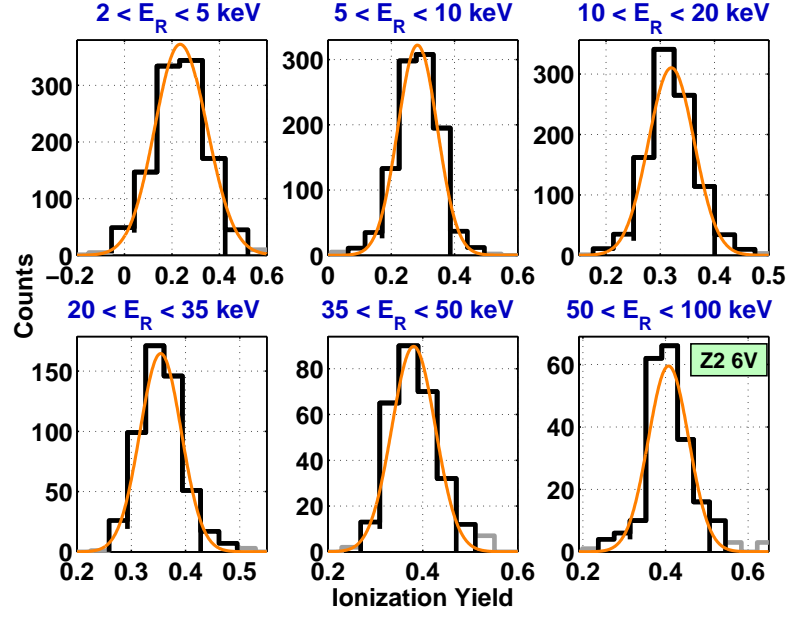


Figure C.30: Binwise estimate of  $\langle Y \rangle$  and  $\sigma_Y$  for Z2 6V nuclear recoils.

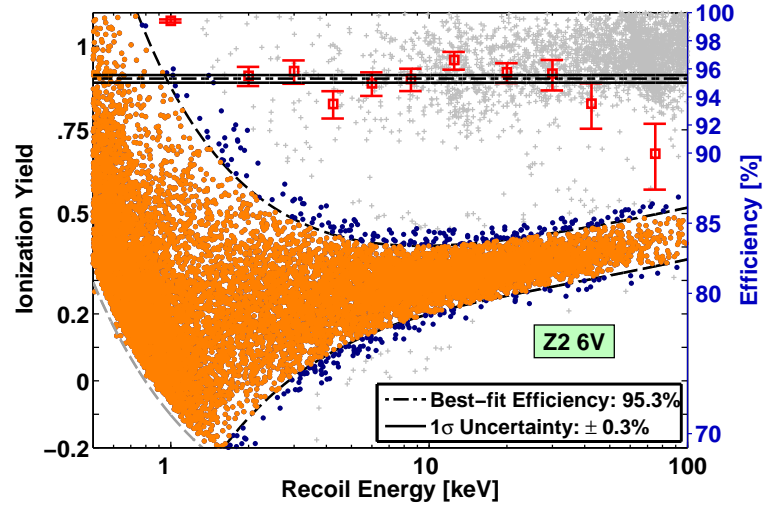


Figure C.31:  $2\sigma$  nuclear-recoil band selection efficiency for Z2 6V data.



# APPENDIX C. ANALYSIS CUTS

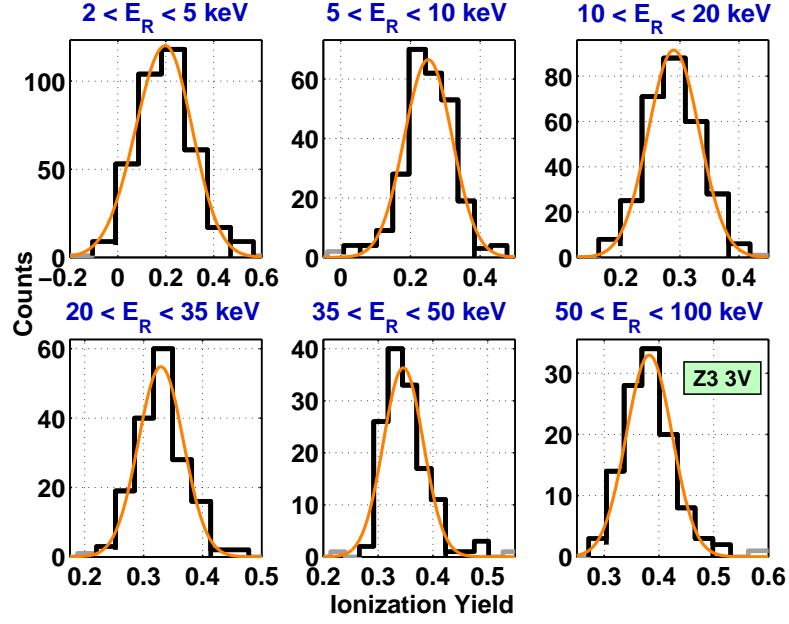


Figure C.32: Binwise estimate of  $\langle Y \rangle$  and  $\sigma_Y$  for Z3 3V nuclear recoils.

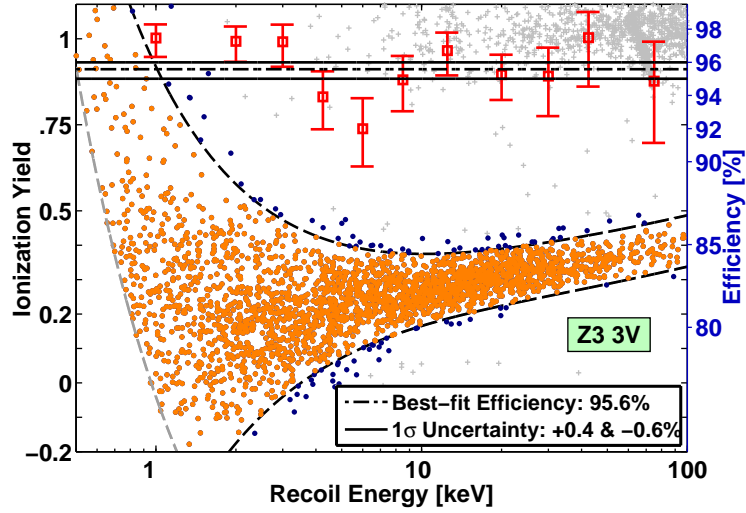


Figure C.33:  $2\sigma$  nuclear-recoil band selection efficiency for Z3 3V data.

## C.2. NUCLEAR-RECOIL BAND

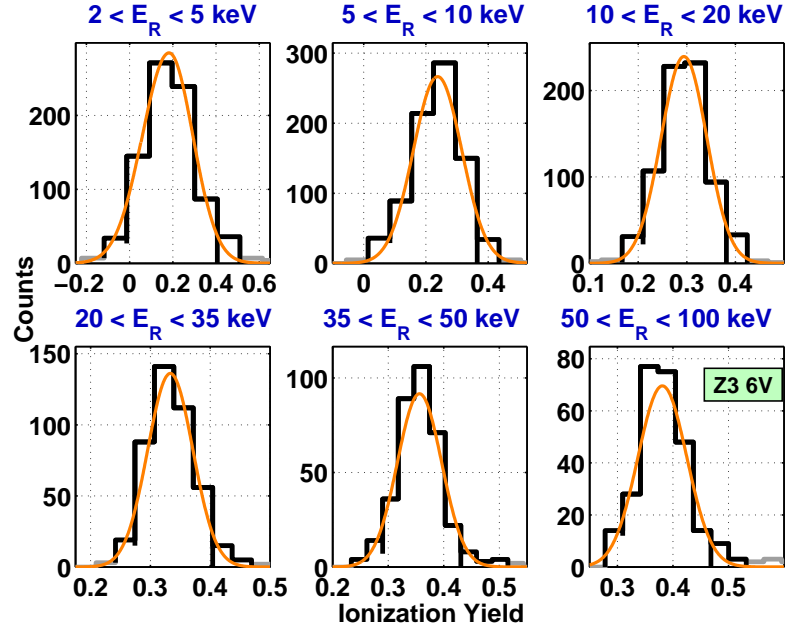


Figure C.34: Binwise estimate of  $\langle Y \rangle$  and  $\sigma_Y$  for Z3 6V nuclear recoils.

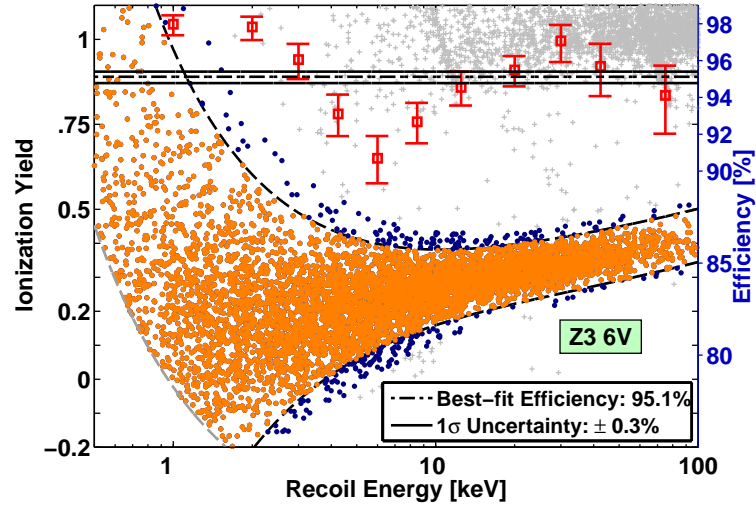


Figure C.35:  $2\sigma$  nuclear-recoil band selection efficiency for Z3 6V data.

# APPENDIX C. ANALYSIS CUTS

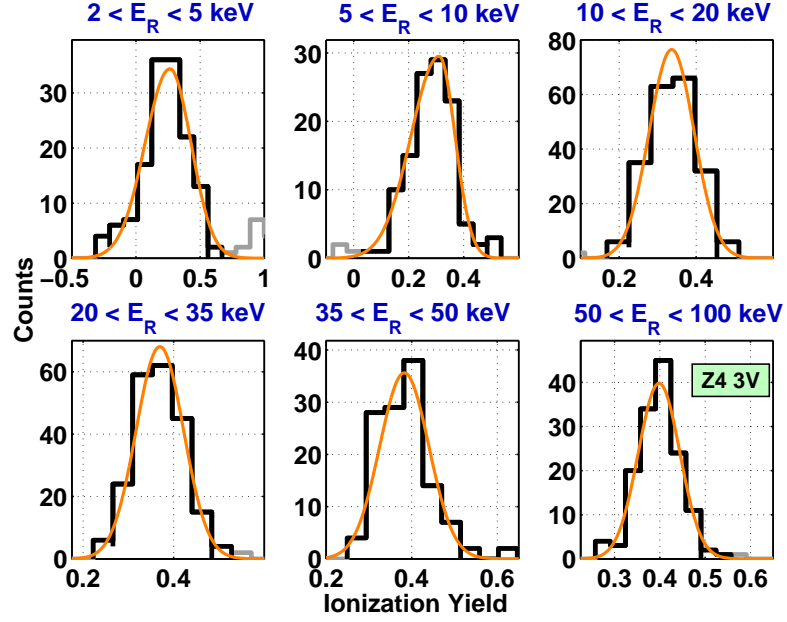


Figure C.36: Binwise estimate of  $\langle Y \rangle$  and  $\sigma_Y$  for Z4 3V nuclear recoils.

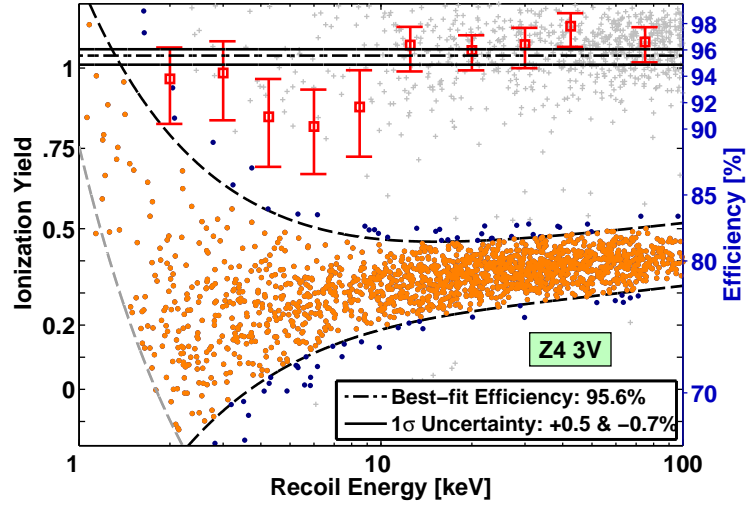


Figure C.37:  $2\sigma$  nuclear-recoil band selection efficiency for Z4 3V data.

## C.2. NUCLEAR-RECOIL BAND

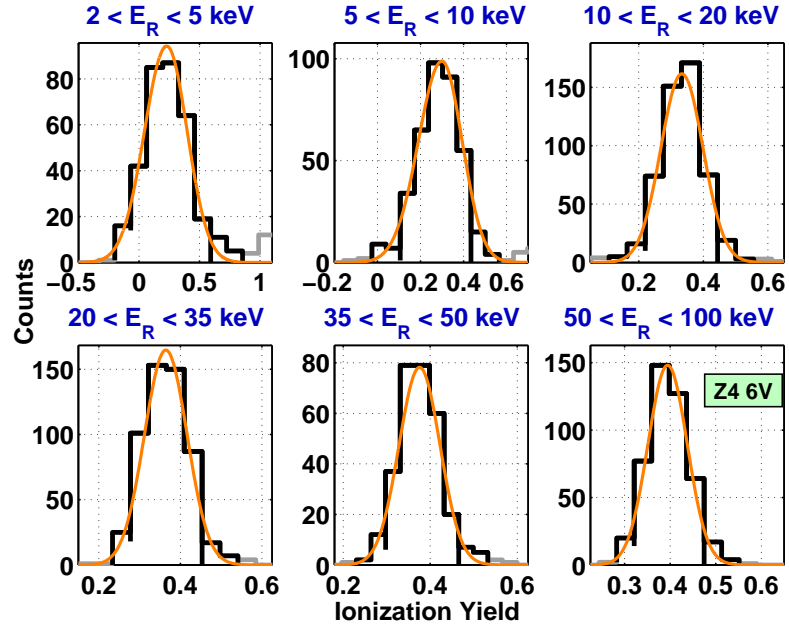


Figure C.38: Binwise estimate of  $\langle Y \rangle$  and  $\sigma_Y$  for Z4 6V nuclear recoils.

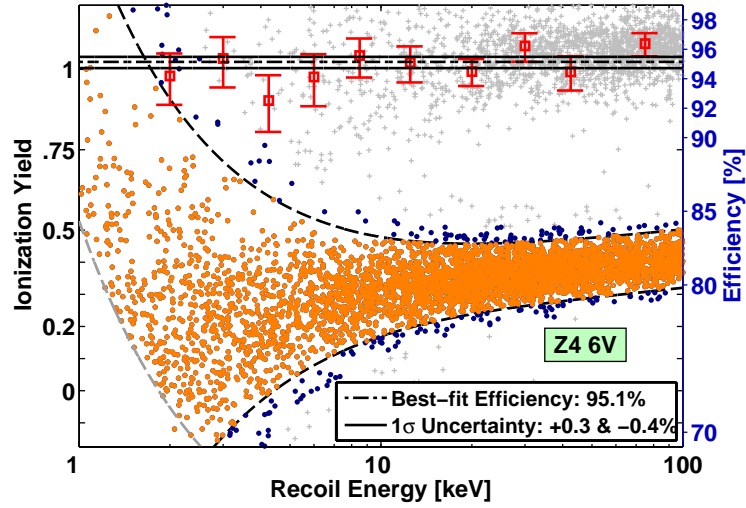


Figure C.39:  $2\sigma$  nuclear-recoil band selection efficiency for Z4 6V data.

# APPENDIX C. ANALYSIS CUTS

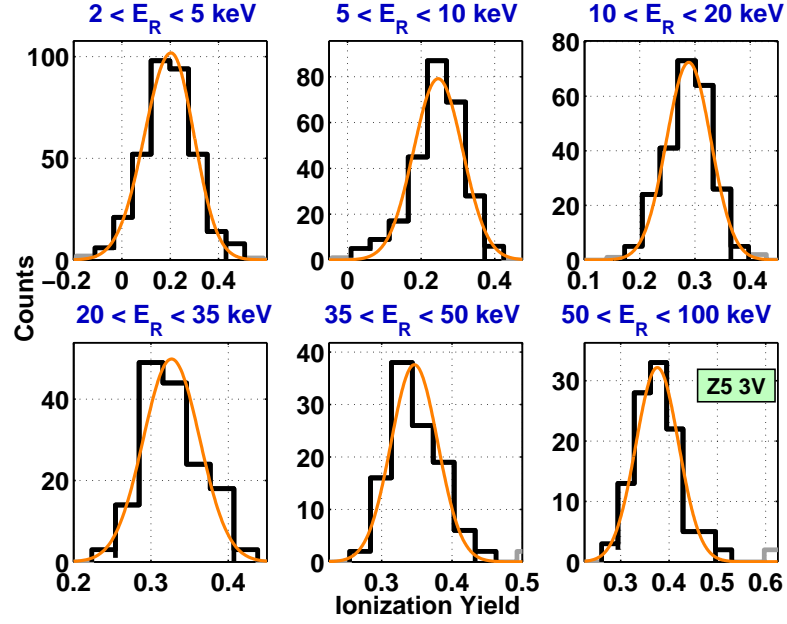


Figure C.40: Binwise estimate of  $\langle Y \rangle$  and  $\sigma_Y$  for Z5 3V nuclear recoils.

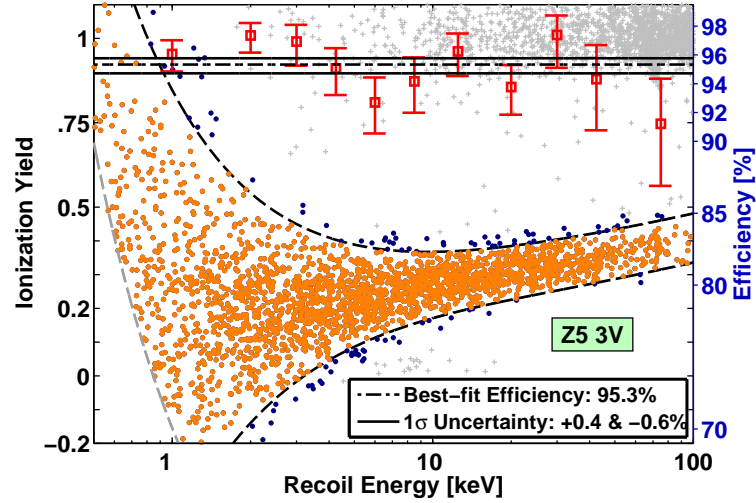


Figure C.41:  $2\sigma$  nuclear-recoil band selection efficiency for Z5 3V data.

## C.2. NUCLEAR-RECOIL BAND

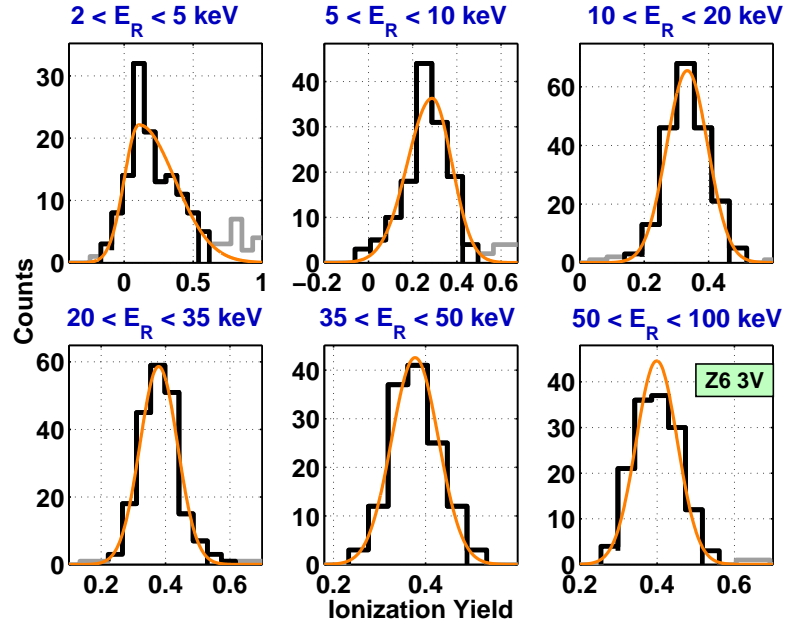


Figure C.42: Binwise estimate of  $\langle Y \rangle$  and  $\sigma_Y$  for Z6 3V nuclear recoils.

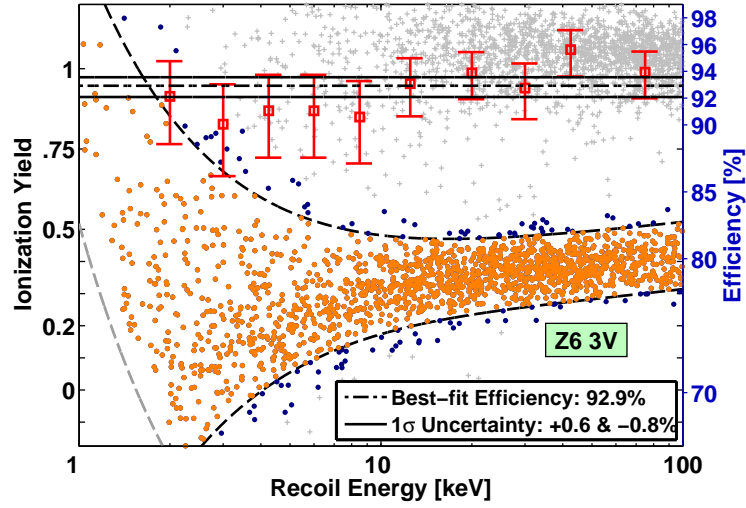


Figure C.43:  $2\sigma$  nuclear-recoil band selection efficiency for Z6 3V data.

# APPENDIX C. ANALYSIS CUTS

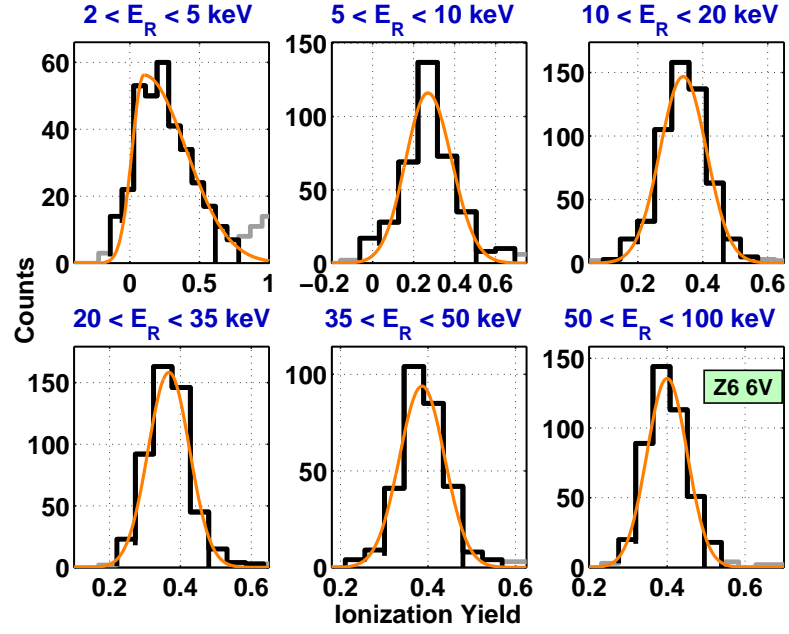


Figure C.44: Binwise estimate of  $\langle Y \rangle$  and  $\sigma_Y$  for Z6 6V nuclear recoils.

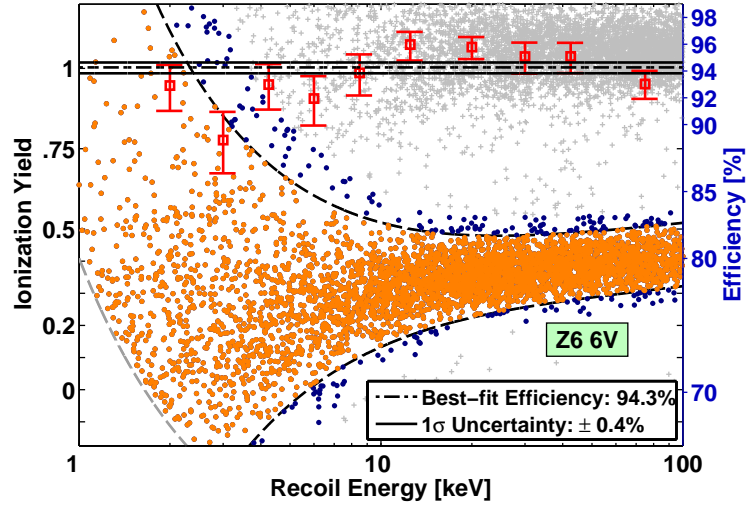


Figure C.45:  $2\sigma$  nuclear-recoil band selection efficiency for Z6 6V data.

# Appendix D

## Energy Resolution

This appendix includes additional figures related to the discussion of ZIP-detector energy resolution in Section 5.4. See the main text and figure captions in that section for further details.

### D.1 Electron-Recoil Energy Response

In most cases, the following figures are for the bias-voltage run not included in Section 5.4.1.



## APPENDIX D. ENERGY RESOLUTION

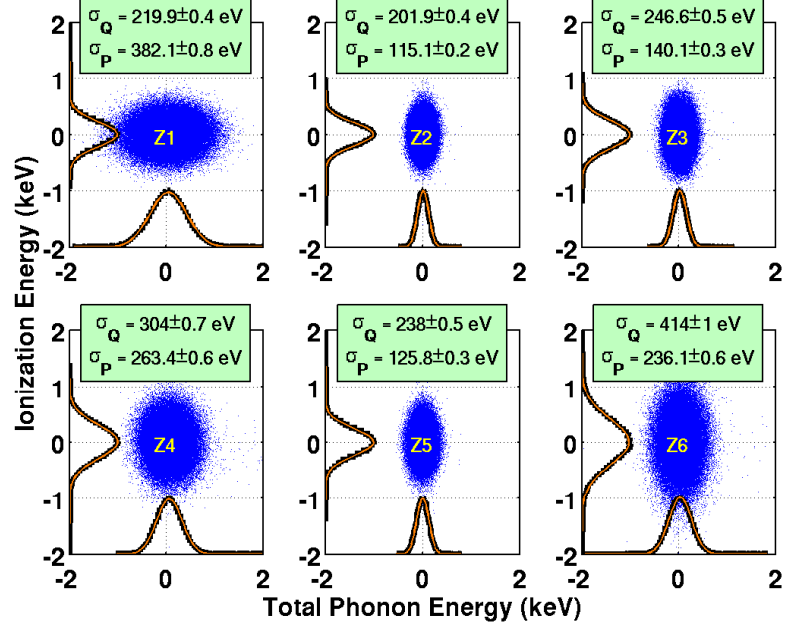


Figure D.1: Noise blobs for the 6V data (equivalent to Figure 5.18).

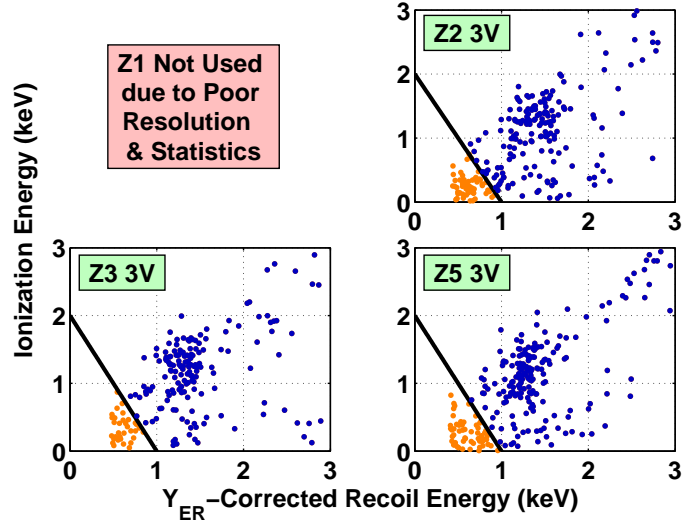


Figure D.2: 3V preselection of 1.29 keV events (equivalent to Figure 5.19).

### D.1. ELECTRON-RECOIL ENERGY RESPONSE

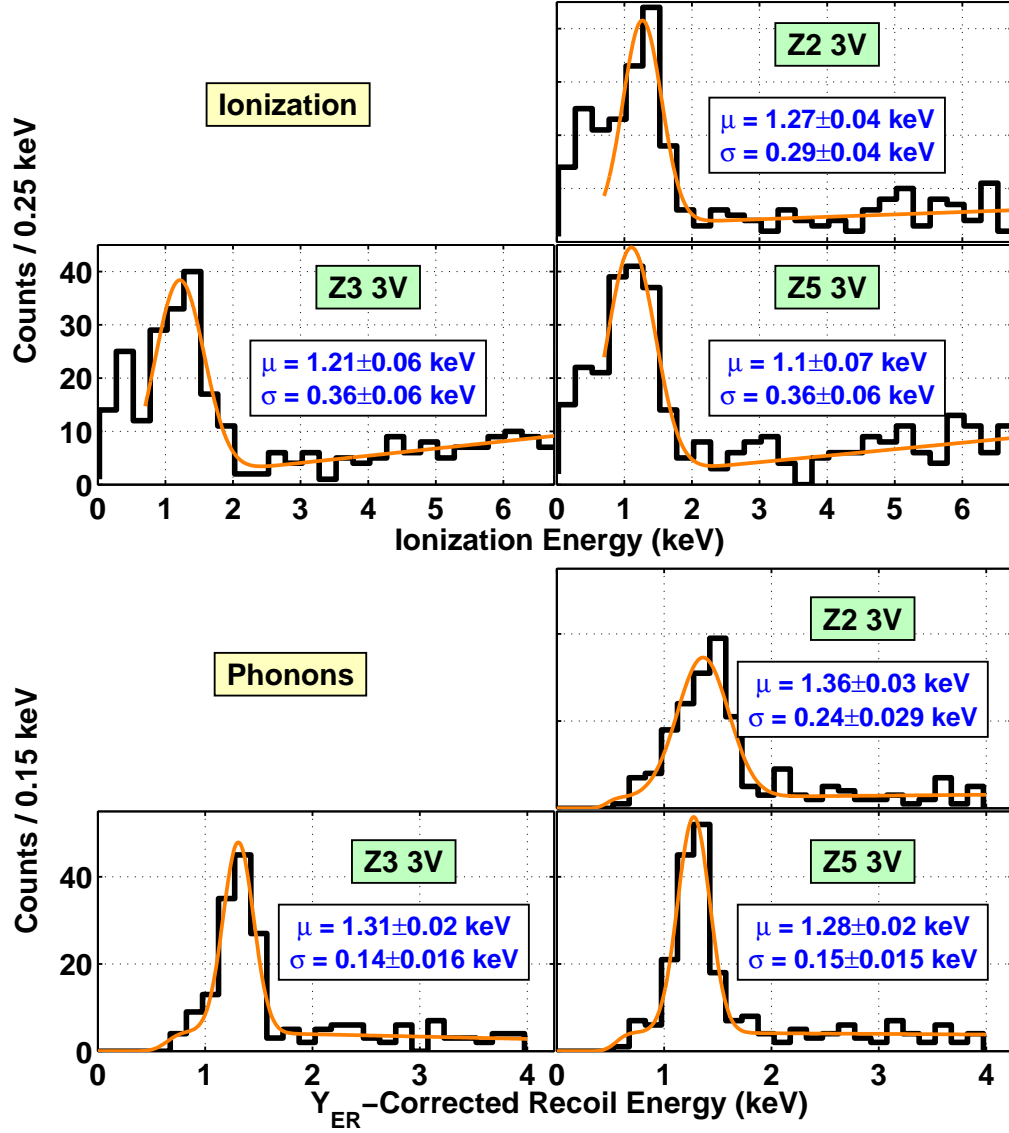


Figure D.3: Fits to the 1.29 keV peak in the 3V WIMP-search data (equivalent to Figure 5.20). To suppress the influence of the noise distribution at 0 keV, the ionization fits were restricted to energies  $>0.7 \text{ keV}$ .

# APPENDIX D. ENERGY RESOLUTION

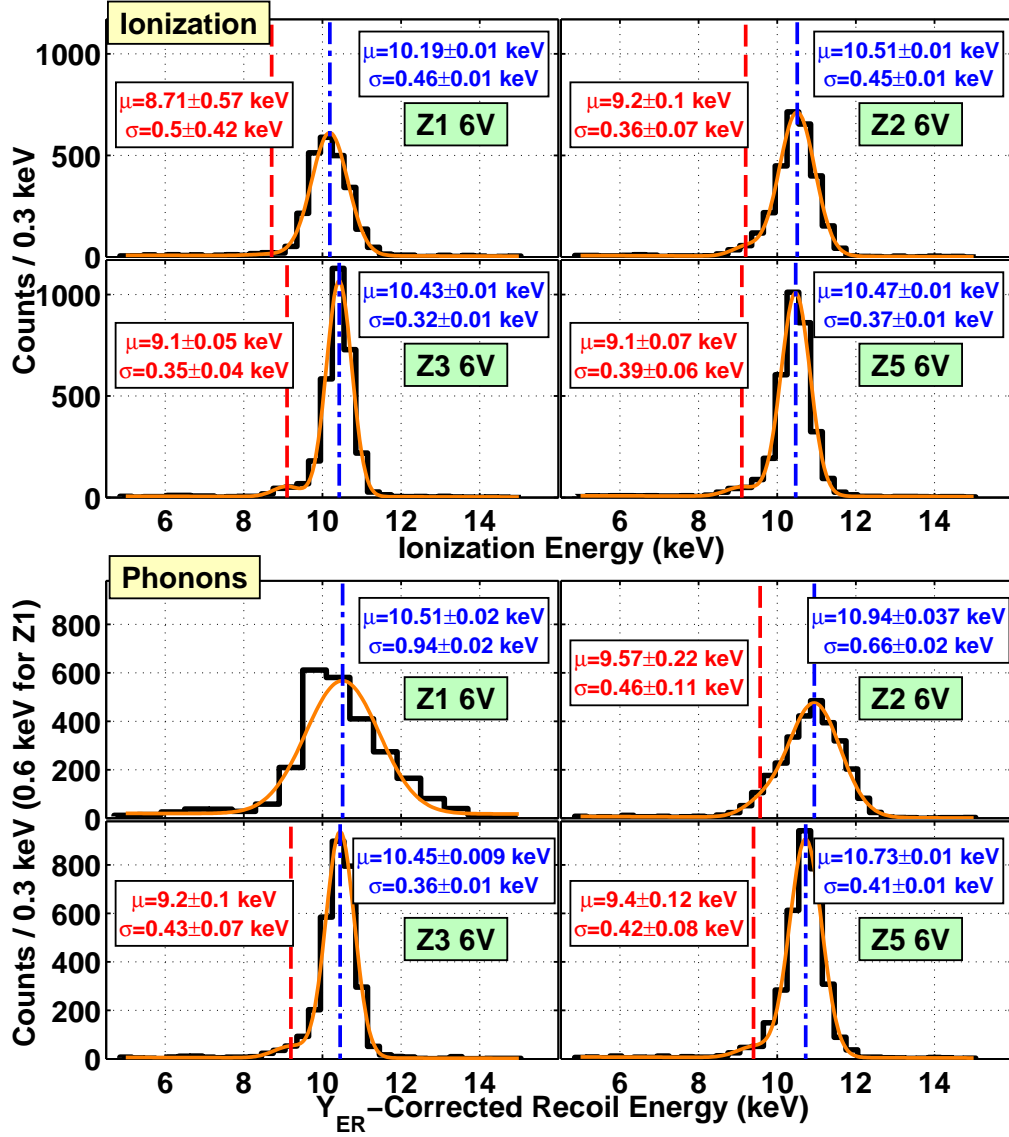


Figure D.4: Fits to the 8.98 and 10.36 keV peaks in the 6V WIMP-search data (equivalent to Figure 5.21). Due to poor resolution, the 8.98 keV peak could not be resolved in the Z1 phonon data.

### D.1. ELECTRON-RECOIL ENERGY RESPONSE

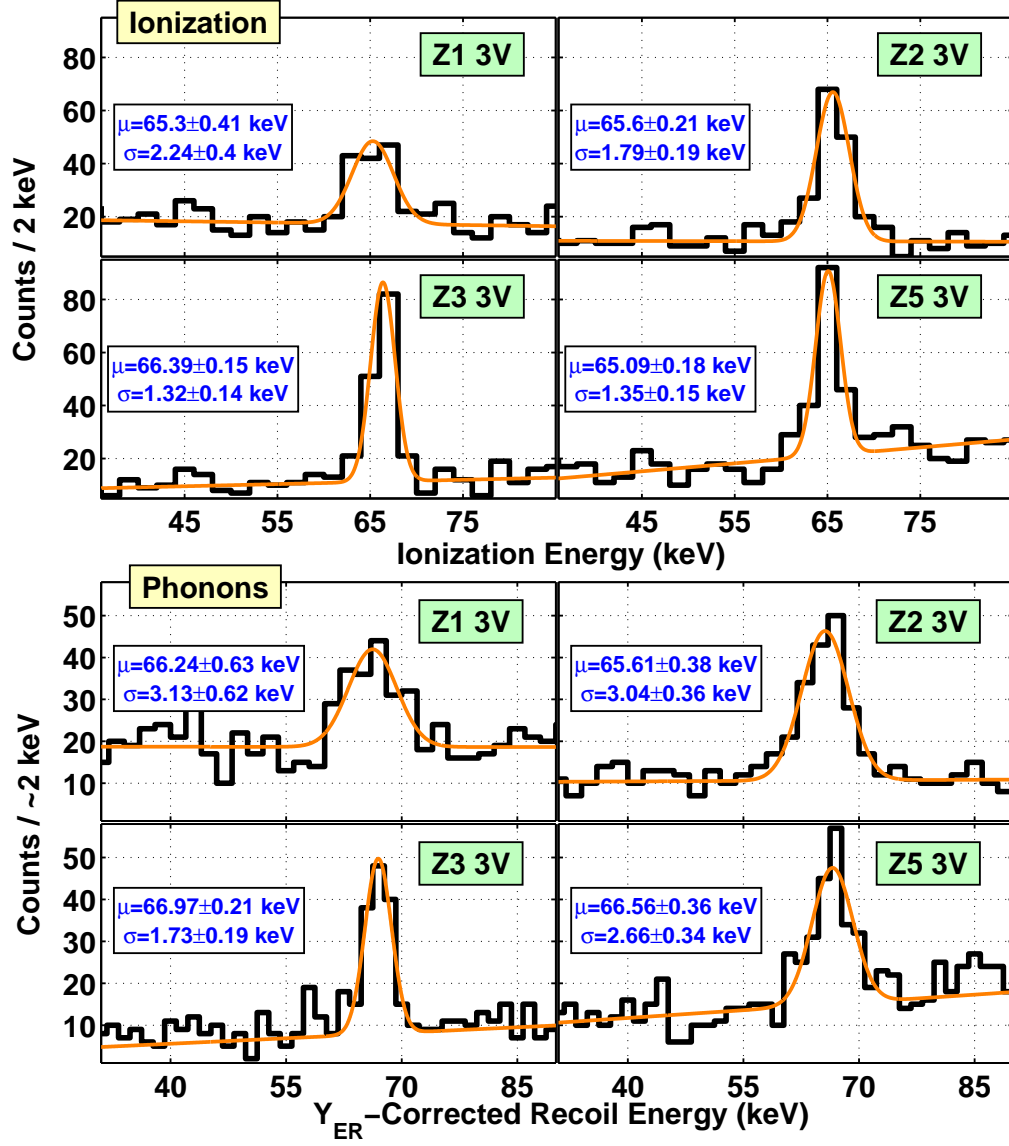


Figure D.5: Fits to the 66.7 keV peak in the 3V WIMP-search data (equivalent to Figure 5.22).



# Appendix E

## $^{252}\text{Cf}$ Spectral Shapes

In this appendix I review a numerical calculation of the spectral shapes expected from  $^{252}\text{Cf}$  neutrons scattering from Ge and Si ZIP detectors. As discussed in Sections 5.4.2 and 6.4.2, these spectra are not exactly exponential as a function of recoil energy. In particular, high-statistics GEANT4 [493] simulations performed for the CDMS deep-site shielding configuration have revealed a bump near 20 keV in the Si detectors'  $^{252}\text{Cf}$  spectra, which prompted the following calculations. Any distinguishing features in the  $^{252}\text{Cf}$  spectra could potentially be very useful for gauging the nuclear-recoil energy scale. In the following, I attempt to derive the precise recoil-energy shapes by using the same ENDF [770] neutron cross sections and angular probabilities used as inputs to the GEANT4 simulations.

### E.1 Differential Scattering Rate

The derivation of the differential scattering rate for neutrons scattering from nuclear targets is similar to the derivation for WIMPs presented in Section 3.3.2. However, the energy dependence is slightly different because the spectrum of incident energies is not defined according to a Maxwellian velocity distribution, but instead according to the energy distribution obtained by transporting the spectrum of energies emitted by the  $^{252}\text{Cf}$  source through the CDMS shielding

layers. The treatment of the elastic scattering cross section differs as well. In this section I outline a rough derivation aimed toward understanding the energy dependence. No attempt is made to derive the absolute normalization, and several (constant) multiplicative factors are neglected or dropped along the way.

The differential scattering rate for neutrons to scatter from a nuclear target (described by its atomic mass  $A$ ) is given by

$$\frac{dR}{dq^2} \propto \frac{d\sigma}{dq^2}(q^2, v) v n(v), \quad (\text{E.1.1})$$

where  $q^2$  is proportional to the transferred energy,  $v$  is the relative neutron-nucleus velocity,  $n$  is the velocity-dependent neutron number density, and  $\sigma$  is the energy- and velocity-dependent neutron-nucleus cross section. Note that Equation E.1.1 is true for a particular value of  $v$ . To get the correct recoil-energy shape, the right-hand side of Equation E.1.1 must be integrated over all possible velocities:

$$\frac{dR}{dq^2} \propto \int \frac{d\sigma}{dq^2} \frac{dn}{dv} dv = \int \frac{d\sigma}{dq^2} \frac{dn}{dE_i} dE_i \sqrt{E_i}, \quad (\text{E.1.2})$$

where the right-hand side is obtained via a change of variables from  $v$  to the incident neutron energy  $E_i \propto v^2$ . At this point it is useful to recall the elastic scattering relationship between the kinetic energy of the recoiling nucleus,  $E_R$ , the energy of the incident neutron, and the center-of-mass scattering angle  $\theta^*$  (see also Equation 3.3.38):

$$E_R = \frac{2A}{(1+A)^2} E_i (1 - \cos \theta^*), \quad (\text{E.1.3})$$

where, again,  $A$  is the atomic mass of the target.  $A = 28$  for a Si target, and is a mixture of 70, 72, 73, 74 and 76 for a Ge target. For nonrelativistic scattering,  $q^2 \propto E_R$ , and the differential cross section can be rewritten as

$$\frac{d\sigma}{dq^2} \propto \frac{d\sigma}{dE_R} = \frac{d\sigma}{d\Omega} \frac{\delta\Omega}{\delta E_R} \propto \frac{1}{E_i} \frac{d\sigma}{d\Omega}, \quad (\text{E.1.4})$$

## E.2. DIFFERENTIAL NUMBER DENSITY

since  $\Omega \propto \cos \theta^*$ , and  $\cos \theta^* \propto E_R/E_i$ . Noting that  $dR/dq^2 \propto dR/dE_R$ , and substituting the right-hand side of Equation E.1.4 into the right-hand side of Equation E.1.2, the differential scattering rate can be written as

$$\frac{dR}{dE_R} \propto \int \frac{d\sigma}{d\Omega} \frac{dn}{dE_i} \frac{dE_i}{\sqrt{E_i}}, \quad (\text{E.1.5})$$

where the integrand is now entirely in terms of  $E_i$  and the center-of-mass scattering angle. For a given value of  $E_R$ , the integral is restricted to combinations of  $E_i$  and  $\cos \theta^*$  that satisfy Equation E.1.3. Specifically, since  $\cos \theta^*$  varies from -1 to 1, the integral runs from  $(1 + A)^2 E_R / 4A$  to  $\infty$ .

Equation E.1.5 and the limits of integration noted above provide the framework necessary to calculate the shape (or energy dependence) of the differential event rate for a spectrum of neutrons to scatter from a nuclear target. All that remains is to specify the differential number density  $dn/dE_i$ , and the differential cross section  $d\sigma/d\Omega$ . The former is simply the spectrum of incident neutron energies, while the latter decomposes into two parts:

$$\frac{d\sigma}{d\Omega} \propto \sigma(E_i) P(\cos \theta^* | E_i), \quad (\text{E.1.6})$$

where  $\sigma(E_i)$  is the elastic cross section as a function of incident neutron energy (analogous to the WIMP-nucleus form factor described in Section 3.3.2) and  $P(\cos \theta^* | E_i)$  is the angular probability for a particular value of  $\cos \theta^*$  as a function of  $E_i$ . Three inputs are thus required to perform the desired numerical calculation of  $dR/dE_R$ .

## E.2 Differential Number Density

The differential neutron number density  $dn/dE_i$  is obtained by transporting the spectrum of energies emitted by the  $^{252}\text{Cf}$  source through the CDMS shielding layers. There appears to be some uncertainty regarding the high-energy tail of this spectrum. The spectrum used for the GEANT3 simulations of the shallow-site shielding configuration is approximately given by

$$\frac{dn}{dE_i} \propto e^{-E_i/(1.42 \text{ MeV})} \sqrt{E_i}, \quad (\text{E.2.1})$$



## APPENDIX E. $^{252}\text{Cf}$ SPECTRAL SHAPES

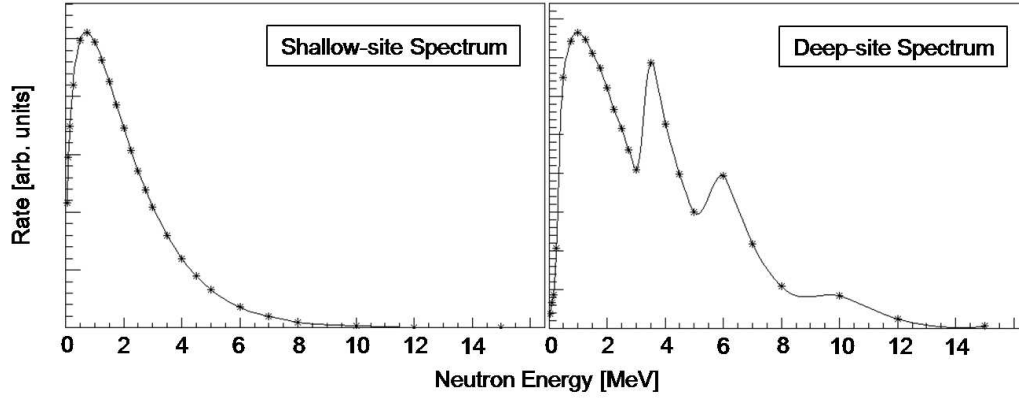


Figure E.1: The spectra used in simulations of the CDMS shallow- (left) and deep-site (right) shielding configurations to represent the range of neutron energies emitted by  $^{252}\text{Cf}$ . Figure adapted from [771].

and is shown in the left panel of Figure E.1. A more complicated multi-peaked spectrum was used for the GEANT4 simulations of the deep-site shielding configuration, and is shown in the right panel of Figure E.1. Fortunately, the presence (of lack) of the high-energy structure exhibited by the deep-site spectrum does not appear to significantly affect the ZIP detector’s nuclear-recoil response for recoil energies  $< 100$  keV. Neutrons that penetrate the shielding layers and deposit energy in the range of a few to 100 keV primarily come from the large peak centered at  $\sim 1$  MeV, which both spectra in Figure E.1 have in common.

The spectrum of neutron energies directly emitted by the source is not quite what is needed for the numerical calculation. CDMS neutron calibrations are typically conducted with the  $^{252}\text{Cf}$  source located such that the neutrons have to penetrate several layers of shielding in order to interact in a ZIP detector. Consequently, the emitted energy spectrum must be transported through the shielding layers to obtain the differential number density required by Equation E.1.5. I did not perform this part of the calculation. Instead, I used the results of a GEANT4 simulation of the deep-site detector geometry to obtain the

## E.2. DIFFERENTIAL NUMBER DENSITY

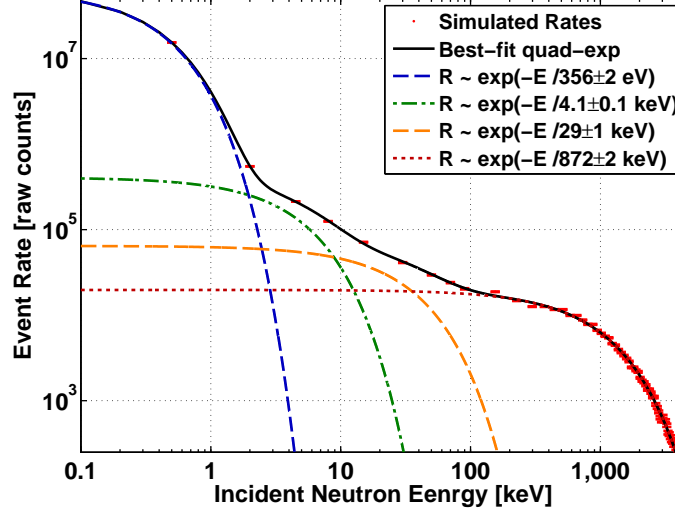


Figure E.2: Spectrum of  $^{252}\text{Cf}$  neutron energies (red dots with error bars) incident upon the ZIP detectors following simulated transport through the CDMS deep-site shielding layers. The multi-exponential fit (solid line) is used to evaluate  $dR/dE_R$  numerically.

spectrum of energies incident upon the ZIP detectors.<sup>1</sup> The simulation transported the multi-peaked spectrum in Figure E.1 through the various deep-site shielding layers, resulting in the spectrum of neutron energies shown in Figure E.2. For convenience, I parameterized the incident spectrum according to a multi-exponential fit. The exponential with the largest decay constant contributes most of the events observed in the ZIP detectors (for  $E_R < 100$  keV), and is given approximately by

$$\frac{dn}{dE_i} \propto e^{-E_i/872 \pm 2 \text{ keV}}. \quad (\text{E.2.2})$$

The best-fit eight-parameter (4 decay plus 4 normalization constants) multi-exponential indicated in Figure E.2 provides the first input needed to evaluate  $dR/dE_R$  numerically.

<sup>1</sup>The GEANT4 simulations of the CDMS deep-site shielding configuration are courtesy of work done by Scott Fallows—a graduate student (at the time of this writing) at the University of Minnesota—to better understand the ZIP detectors’ nuclear-recoil energy response.

### E.3 Elastic Scattering Cross Section

The differential cross section for neutrons to elastically scatter from nuclei is composed of two parts. The first part, denoted  $\sigma(E_i)$ , describes the dependence of the cross section on the incident neutron energy. GEANT4 uses nuclear cross-section data from the ENDF database [770] to model  $\sigma(E_i)$ . I had difficulty figuring out how the low-energy portion of  $\sigma(E_i)$  is stored in the ENDF database files. Consequently, what follows is actually based on data from the JENDL database [609]. For the five stable Ge isotopes, the JENDL and ENDF databases contain identical versions of  $\sigma(E_i)$  (up to 20 MeV), while for <sup>28</sup>Si there are some very slight differences for incident neutron energies greater than a few MeV. Additionally, the JENDL cross sections cut off at 20 MeV, while the ENDF cross sections extend to  $\sim 150$  MeV. None of these differences should significantly effect the recoil-energy spectra for  $E_R < 100$  keV. Most of the events in the recoil-energy range of interest correspond to incident neutron energies less than a few MeV.

Due to the 20 MeV limitation of the JENDL cross sections, I had to constrain the evaluation of Equation E.1.5 to incident energies less than 20 MeV. Consequently, relative to the Monte Carlo simulated recoil-energy spectra, the numerical calculation excludes a range of incident neutron energies between 20 and 150 MeV. The contribution to the differential event rate due to high-energy neutrons falls off exponentially with increasing energy, as does the number density per keV of incident energy. It is therefore unlikely that the excluded energies would significantly alter the recoil spectra presented below in Section E.5.

The JENDL database files are available as text files in which  $\sigma(E_i)$  is listed at several discrete energies between  $1 \times 10^{-5}$  eV and 20 MeV. In order to perform the numerical evaluation of  $dR/dE_R$  to the desired precision, it was necessary to interpolate between these discrete values such that  $\sigma(E_i)$  could be evaluated at arbitrary energies. The resulting interpolated cross sections for the most abundant Ge and Si isotopes are shown in Figure E.3.

#### E.4. ELASTIC SCATTERING ANGULAR PROBABILITY

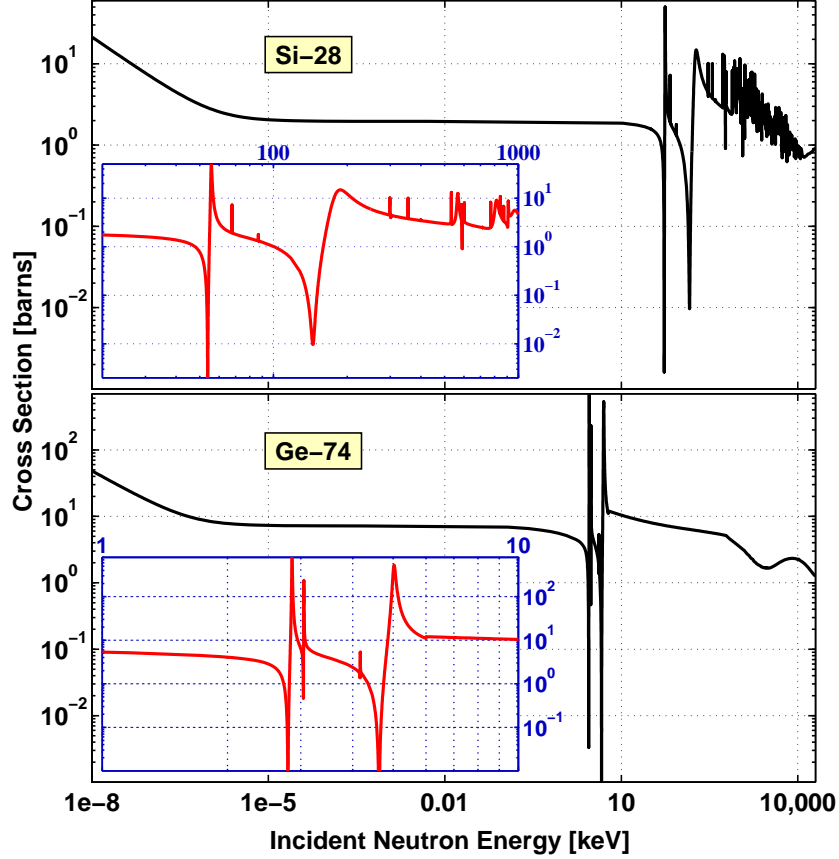


Figure E.3: Cross sections for neutrons to elastically scatter from Si (top) and Ge (bottom) nuclear targets as a function of incident neutron energy. Zoomed in views of the resonant regions are provided in the insets. Cross sections interpolated from data found in the JENDL database [609].

### E.4 Elastic Scattering Angular Probability

The second part of the differential cross section, denoted  $P(\cos \theta^* | E_i)$ , is the probability for a neutron of a given incident energy to scatter with a particular center-of-mass scattering angle. These angular probabilities are stored in the ENDF database files as Legendre polynomial coefficients. Coefficients are

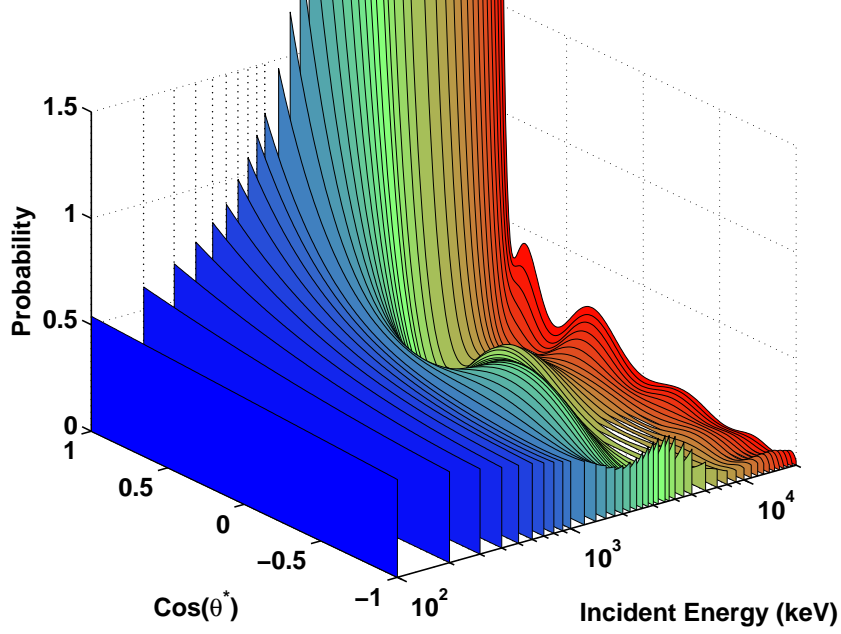


Figure E.4: Angular probability for neutrons to scatter from  $^{74}\text{Ge}$  for several slices of incident neutron energy. Angular data taken from the ENDF database [770].

provided at several discrete energies between  $1 \times 10^{-5}$  eV and 150 MeV, and can be used to construct the angular probabilities according to

$$P(\cos \theta^* | E_i) = \frac{1}{2} + \sum_{l=1}^N \frac{2l+1}{2} a_l(E_i) \mathcal{P}_l(\cos \theta^*), \quad (\text{E.4.1})$$

where  $\mathcal{P}_l$  is the  $l^{\text{th}}$  Legendre polynomial,  $a_l(E_i)$  is the  $l^{\text{th}}$  coefficient for incident energy  $E_i$ , and the sum runs from  $l = 1$  to the highest-order nonzero term. If there are no nonzero coefficients at a given incident energy, the cross section is isotropic (*i.e.*, all angles are equally likely). Similar to  $\sigma(E_i)$ , interpolation was used to obtain the angular probabilities at arbitrary energies.

A representation of the three-dimensional angular probability for  $^{74}\text{Ge}$  is shown in Figure E.4. For a given slice in incident neutron energy, the  $z$  axis represents the angular scattering probability density as a function of  $\cos \theta^*$ . As

## E.5. RECOIL SPECTRA

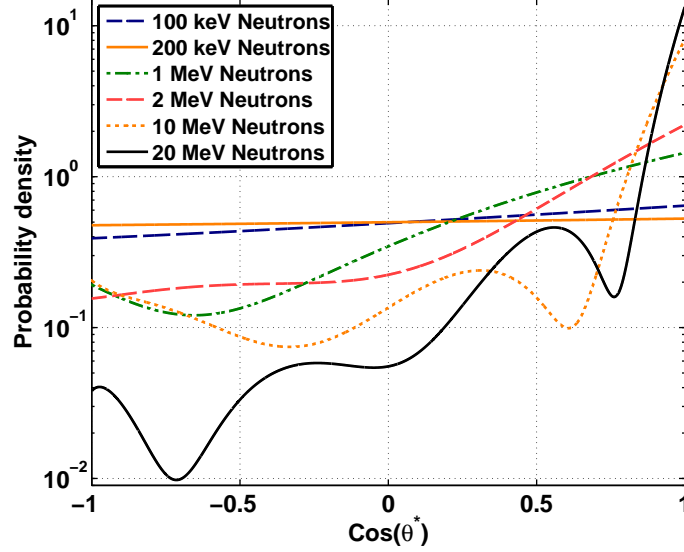


Figure E.5: Angular probability for neutrons to scatter from  $^{28}\text{Si}$  for several values of incident neutron energy. Angular data taken from the ENDF database [770].

the incident neutron energy increases, forward scattering ( $\cos \theta^* = 1$ ) becomes increasingly likely. There are also total backscattering ( $\cos \theta^* = -1$ ) peaks that correspond to resonances in  $\sigma(E_i)$ . An alternative view of the angular probabilities is provided for  $^{28}\text{Si}$  in Figure E.5.

## E.5 Recoil Spectra

With the differential number density and cross-section data specified as described above, Equation E.1.5 was evaluated for six nuclear targets ( $^{28}\text{Si}$  and the five stable Ge isotopes) for  $E_R = 1\text{--}100\text{ keV}$  in steps of  $0.1\text{ keV}$ . At each recoil energy considered, a range of incident neutron energies was calculated (between  $(1 + A)^2 E_R / 4A$  and  $20\text{ MeV}$ ) based on a range of  $\cos \theta^*$  values from  $-1$  to  $0.9999$  (in steps of  $0.0001$ ). The three inputs described above were either evaluated or interpolated at each incident neutron energy, and the results were

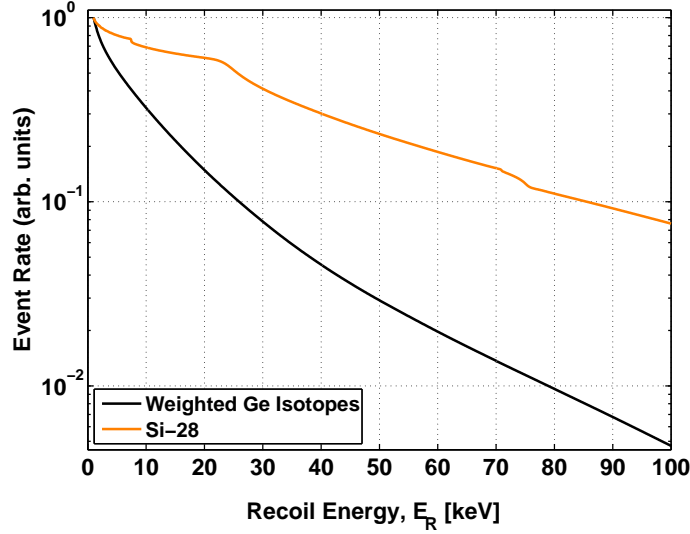


Figure E.6: Numerically calculated  $^{252}\text{Cf}$  differential event rates for Ge (dark/black line) and Si (light/orange line) nuclear targets.

multiplied (including the factors of  $\sqrt{E_i}$  and  $dE_i$ ) and summed appropriately to yield the differential event rate at the given value of  $E_R$ . The resulting Ge spectra were averaged according to their isotopic abundances to obtain a single Ge spectrum. The results are shown in Figure E.6, where the spectra have been normalized to 1 at 1 keV.

The  $^{28}\text{Si}$  spectrum exhibits three bumps that are presumably due to the three most prominent resonances in the  $^{28}\text{Si}$  cross section (at  $E_i \simeq 55$ , 200 and 550 keV—see inset in upper panel of Figure E.3). The decaying exponential hypothesis used to fit the  $^{252}\text{Cf}$  nuclear-recoil spectrum for the Si detector ensemble in Section 5.4.2 is clearly an inadequate description of the true shape. However, in the case of low statistics, it is easy to imagine how fluctuations (and energy resolution to a lesser extent) could wash out the resonances, and cause the spectrum to appear approximately exponential. As demonstrated in Figure E.7, with the numerically calculated shape to guide the eye, the bump at  $\gtrsim 20$  keV can just be discerned in the  $^{252}\text{Cf}$  spectrum observed for the Si detector ensemble at the CDMS shallow site. The discrepancy at low energy

### E.5. RECOIL SPECTRA

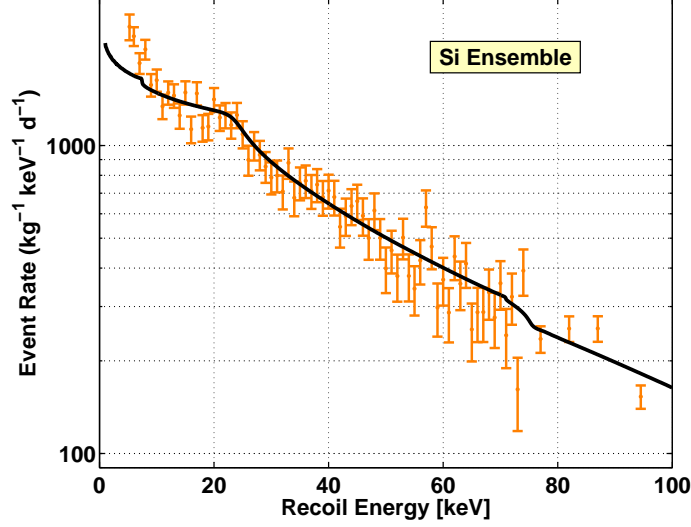


Figure E.7: Comparison of  $^{252}\text{Cf}$  event rate observed for the Si detector ensemble at the CDMS shallow site (error bars) to a scaled version of the numerically calculated shape (solid line). The observed event rate reflects the best-fit two-parameter energy-scale correction discussed in Section 6.4.2.

( $E_R \lesssim 7 \text{ keV}$ ) might be related to the tendency for neutrons to multiply scatter, which was not taken into account by the numerical calculation.

The numerically calculated Ge spectrum closely resembles the dual-exponential hypothesis used in Section 5.4.2, but fails to follow the shape exactly for the lowest and highest energies considered. The shapes calculated for the individual Ge isotopes are virtually indistinguishable, and none have any defining features due to cross-section resonances. The isotopically averaged shape and the spectrum observed for the Ge detector ensemble at the CDMS shallow site are compared in Figure E.8. Similar to Figure E.7, the agreement is nearly perfect except below  $\sim 7 \text{ keV}$ .



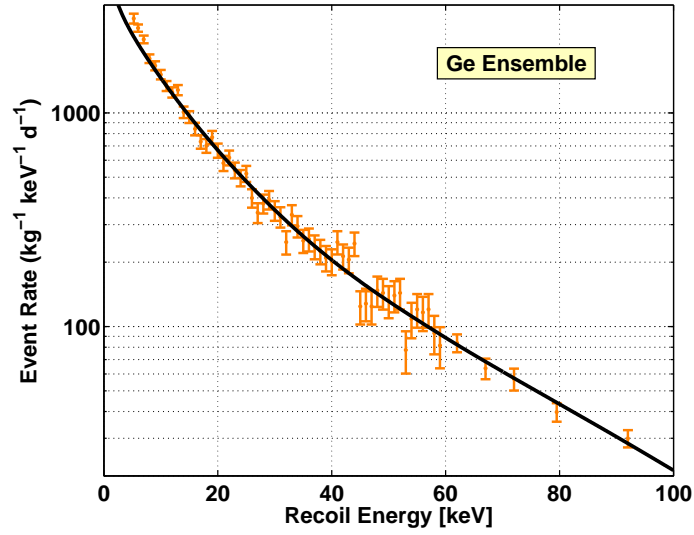


Figure E.8: Comparison of  $^{252}\text{Cf}$  event rate observed for the Ge detector ensemble at the CDMS shallow site (error bars) to a scaled version of the numerically calculated shape (solid line). The observed event rate reflects the best-fit two-parameter energy-scale correction discussed in Section 6.4.2.

# Appendix F

## Detection Efficiencies

This appendix features the combined detection efficiencies not shown in Section 5.5 (equivalent to Figure 5.28). See the main text and figure caption in that section for further details.

Additionally, the averaged efficiencies for the Ge and Si detector ensembles are provided in Figures F.10 and F.11, respectively. To calculate these, each detector's threshold efficiency is first converted to  $Q$ -corrected recoil energy via a Gaussian convolution, where the Gaussian has a  $1\sigma$  width equal to the quadrature difference between the  $Q$ - and  $Y_{\text{NR}}$ -corrected recoil-energy resolutions (see Table 5.6). This procedure adds ionization noise to the threshold efficiencies. Each detector's total detection efficiency is then derived from the product of its analysis-cuts efficiency and (smeared) threshold efficiency. These are averaged over the 3V and 6V runs for Z2, Z3 and Z5 (Z4 and Z6) to obtain an average total efficiency as a function of  $Q$ -corrected recoil energy for the Ge (Si) detector ensemble.

## APPENDIX F. DETECTION EFFICIENCIES

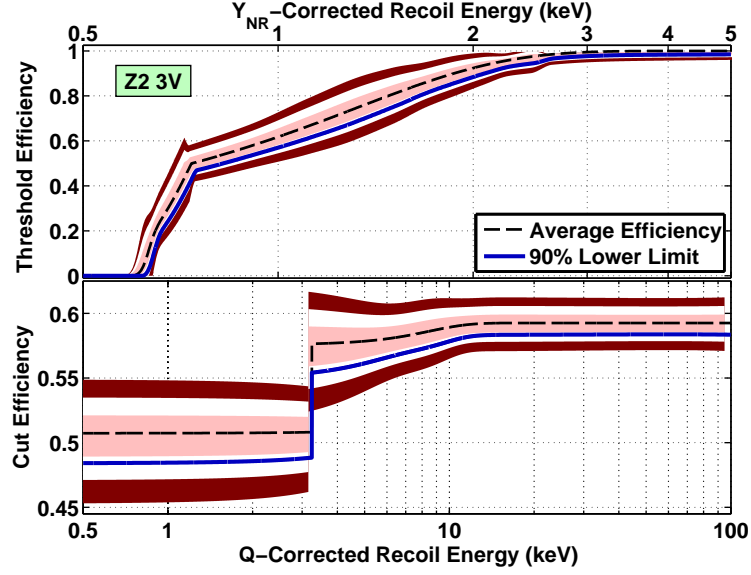


Figure F.1: Combined detection efficiencies for the Z2 3V data.

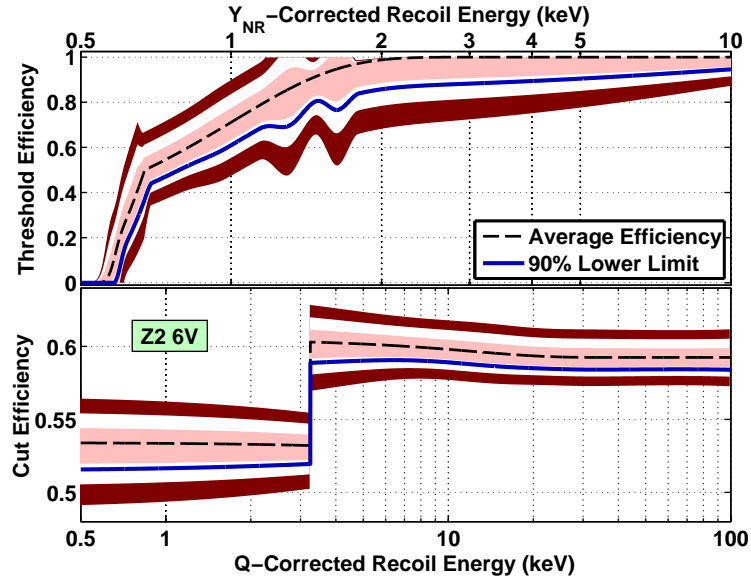


Figure F.2: Combined detection efficiencies for the Z2 6V data.

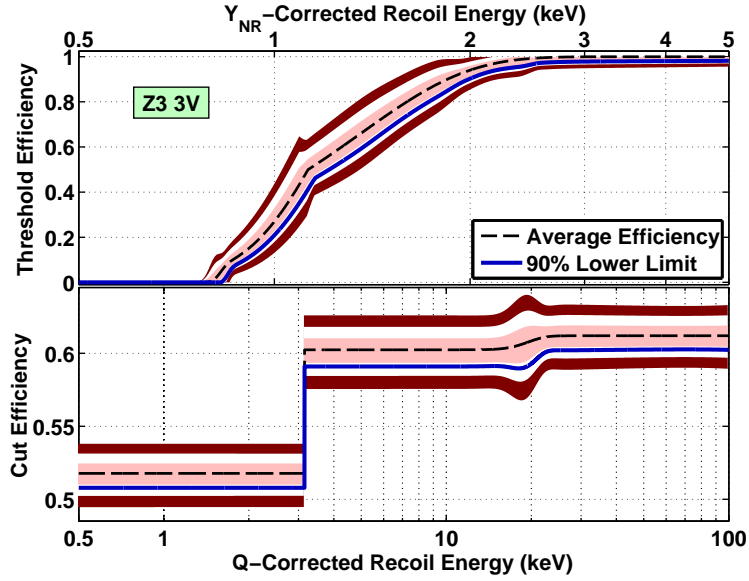


Figure F.3: Combined detection efficiencies for the Z3 3V data.

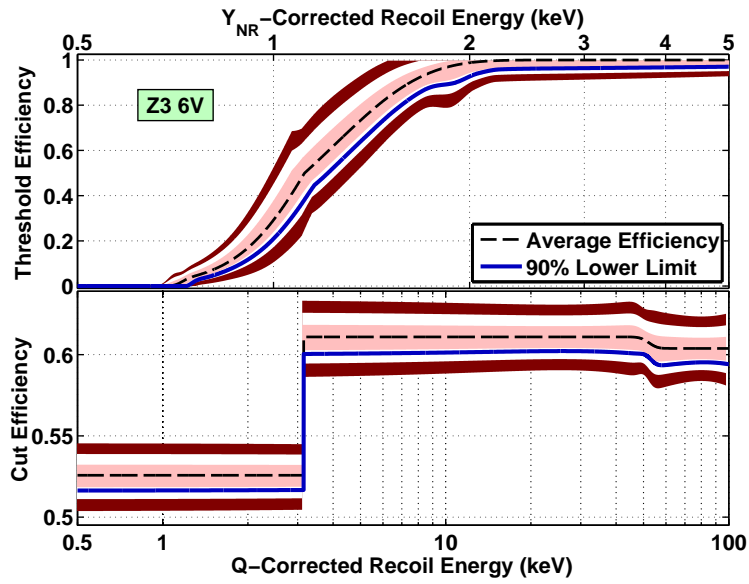


Figure F.4: Combined detection efficiencies for the Z3 6V data.

## APPENDIX F. DETECTION EFFICIENCIES

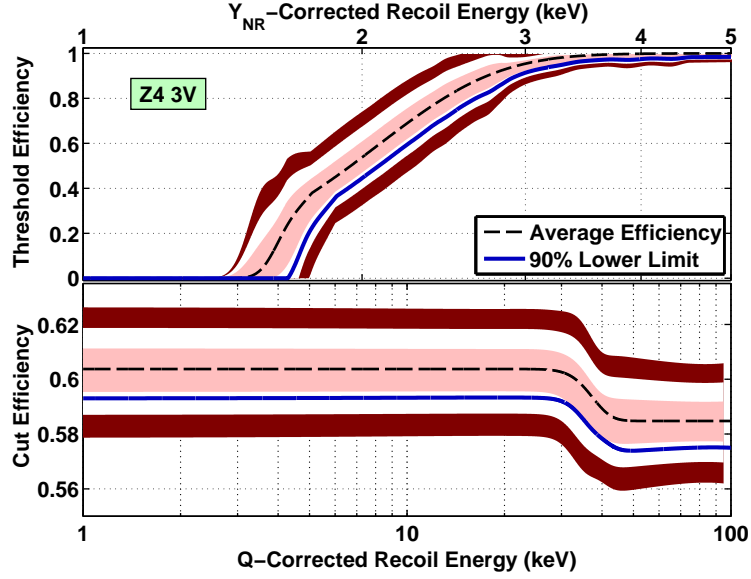


Figure F.5: Combined detection efficiencies for the Z4 3V data.

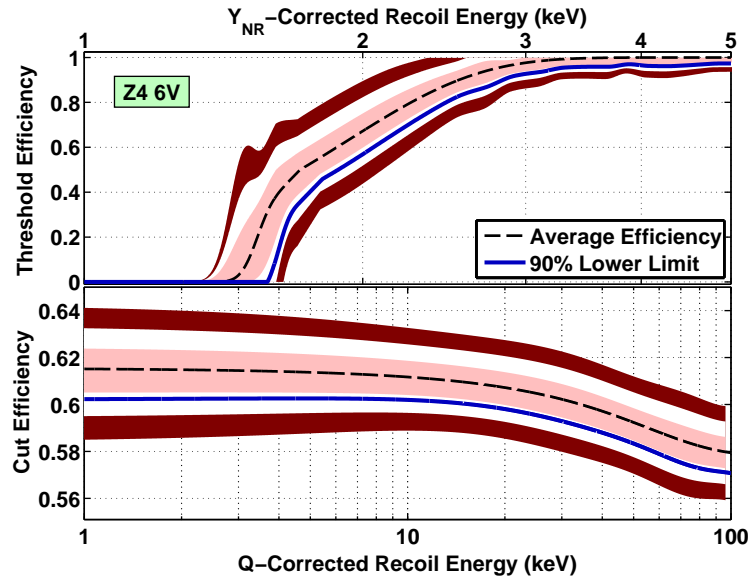


Figure F.6: Combined detection efficiencies for the Z4 6V data.

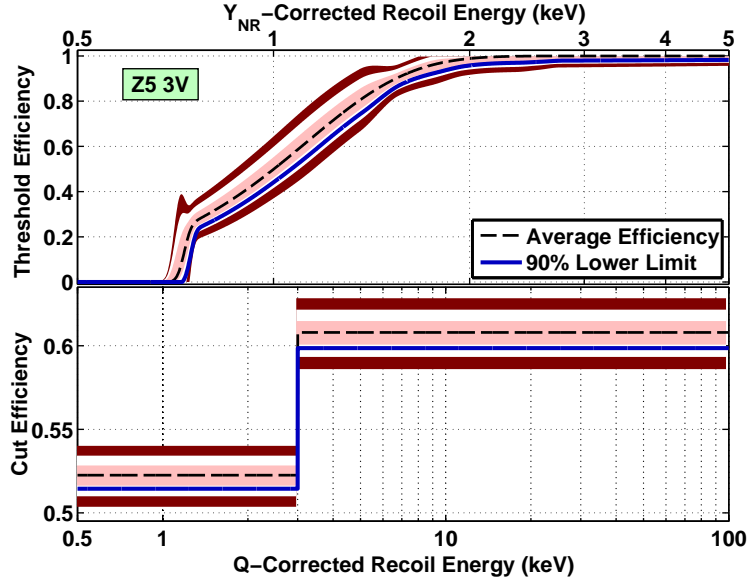


Figure F.7: Combined detection efficiencies for the Z5 3V data.

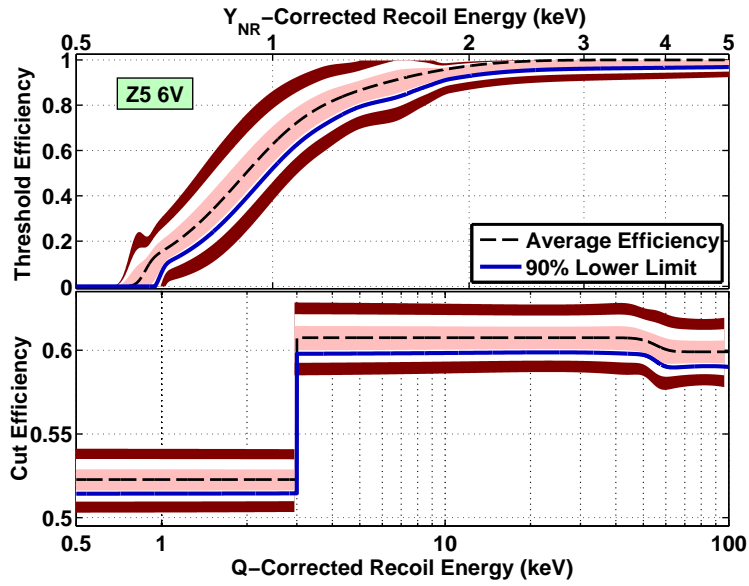


Figure F.8: Combined detection efficiencies for the Z5 6V data.

## APPENDIX F. DETECTION EFFICIENCIES

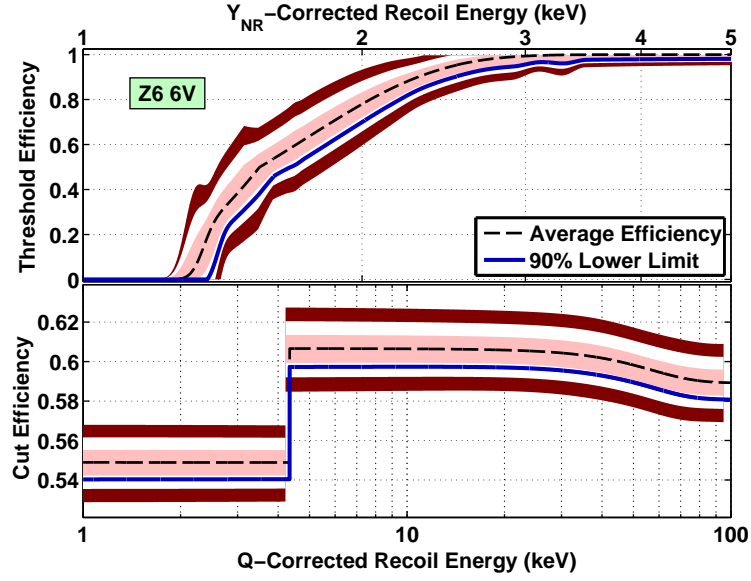


Figure F.9: Combined detection efficiencies for the Z6 6V data.

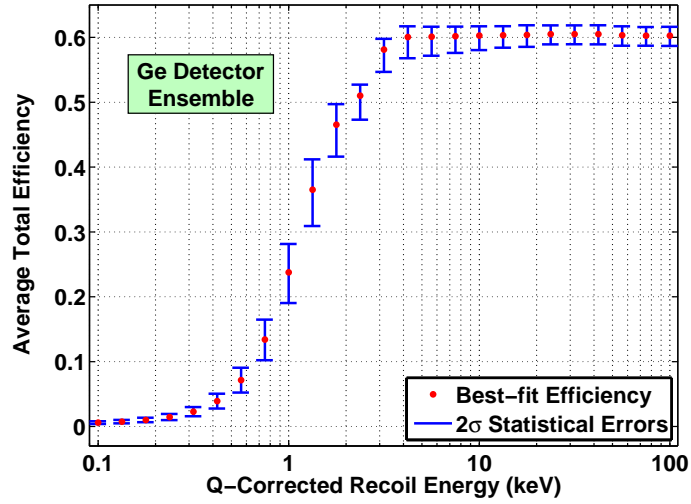


Figure F.10: Average total detection efficiency for the Ge detector ensemble.

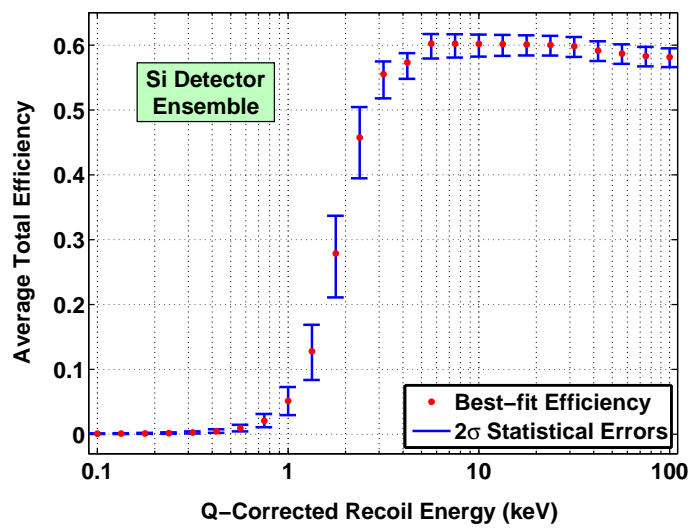


Figure F.11: Average total detection efficiency for the Si detector ensemble.





# Appendix G

## WIMP Candidates

This appendix includes additional scatter plots of WIMP-candidate ionization yield versus recoil energy for the detectors and bias voltages not included in Section 6.2. See the main text and figure captions in that section for further details.

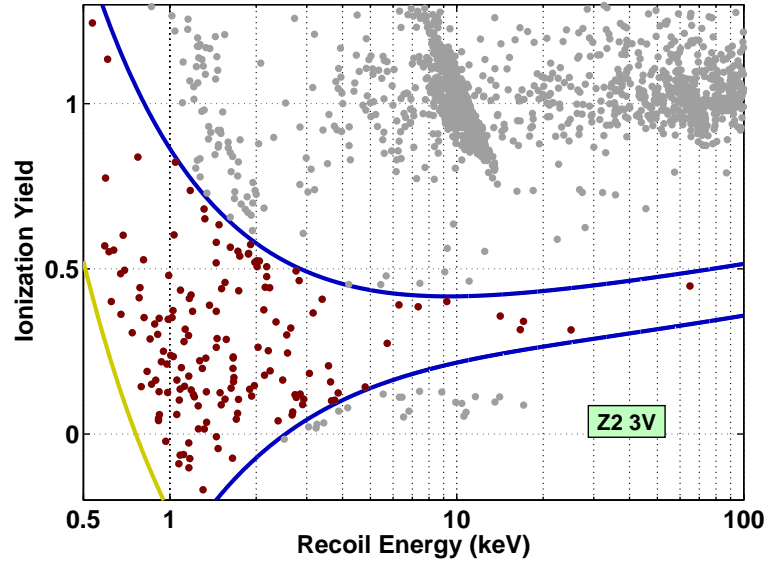


Figure G.1: Ionization yield vs. recoil energy for the Z2 3V WIMP candidates

## APPENDIX G. WIMP CANDIDATES

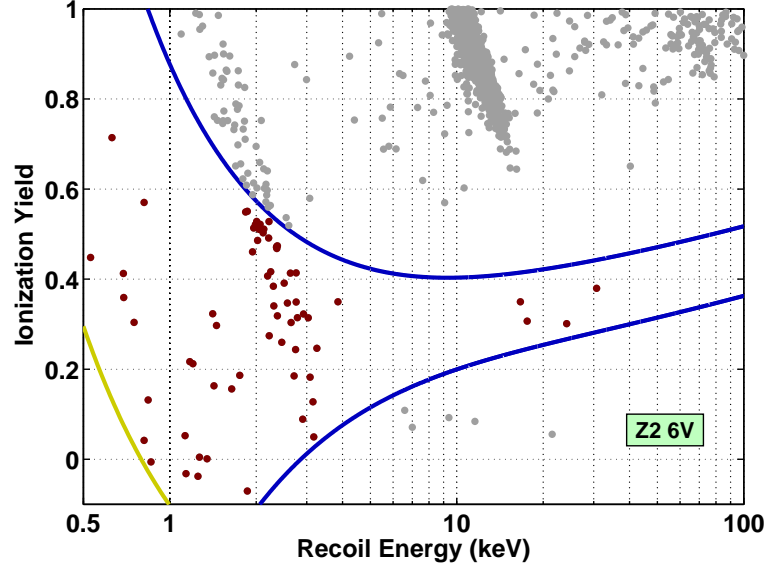


Figure G.2: Ionization yield vs. recoil energy for the Z2 6V WIMP candidates

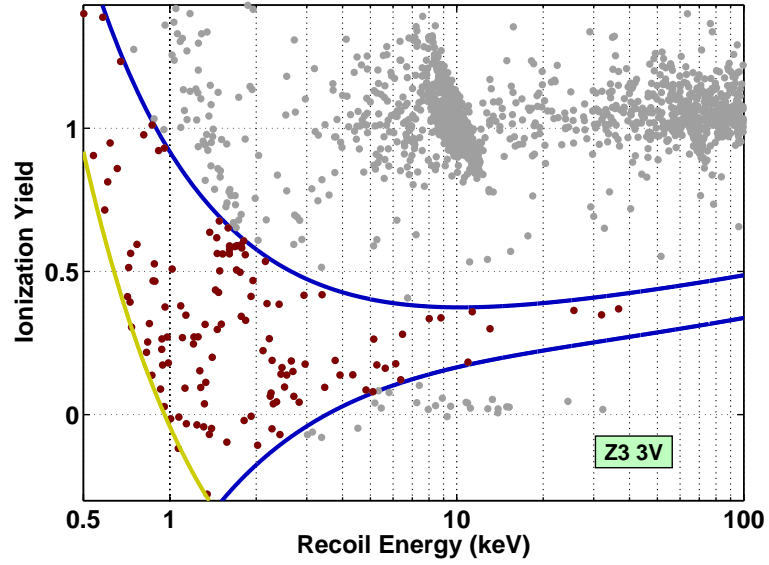


Figure G.3: Ionization yield vs. recoil energy for the Z3 3V WIMP candidates

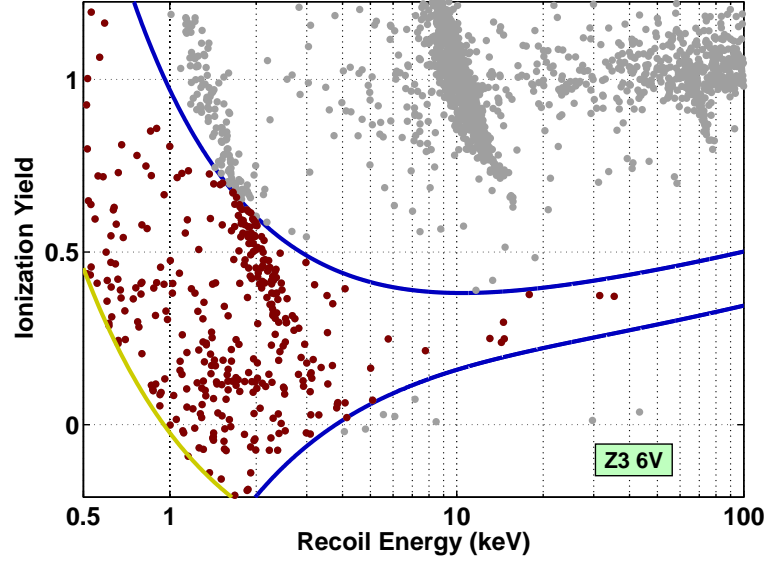


Figure G.4: Ionization yield vs. recoil energy for the Z3 6V WIMP candidates

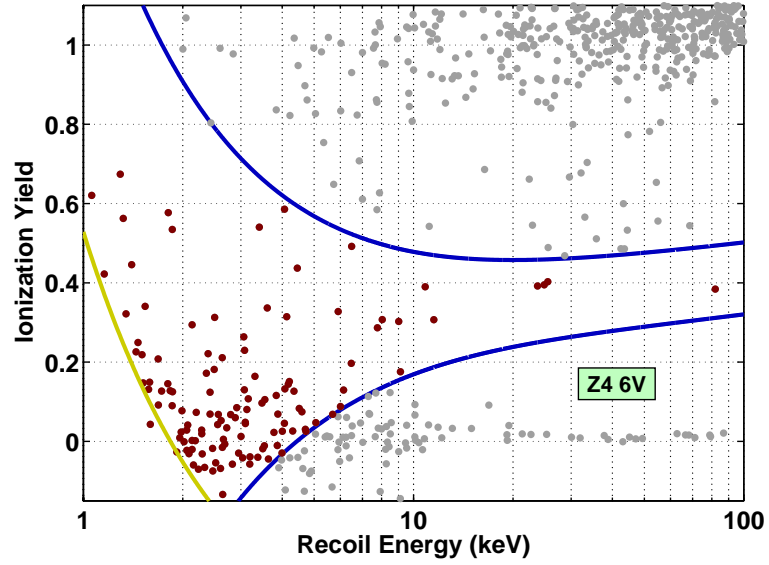


Figure G.5: Ionization yield vs. recoil energy for the Z4 6V WIMP candidates

# APPENDIX G. WIMP CANDIDATES

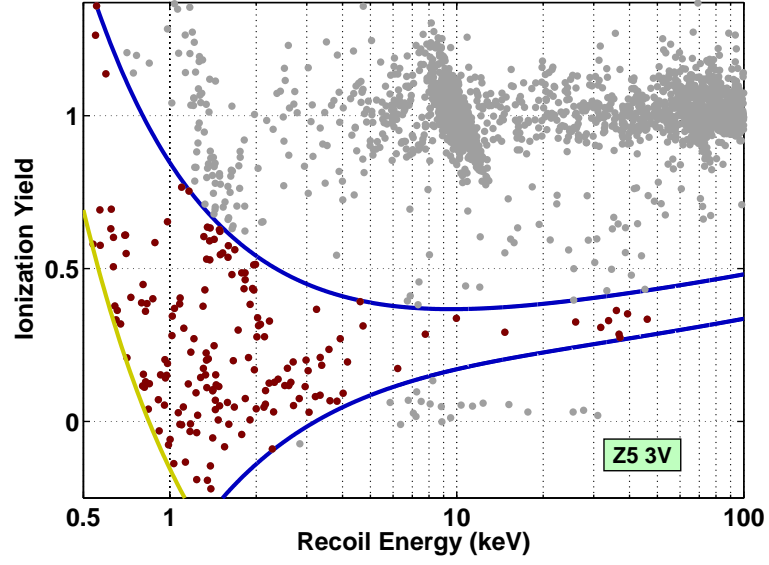


Figure G.6: Ionization yield vs. recoil energy for the Z5 3V WIMP candidates

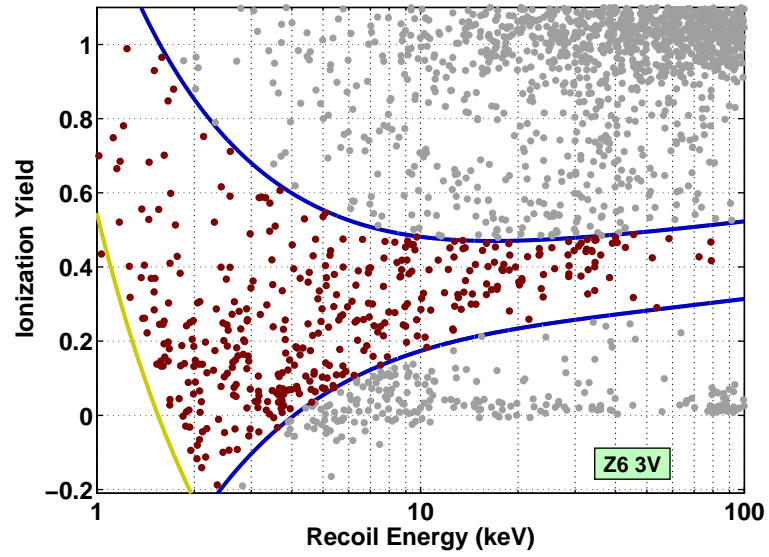


Figure G.7: Ionization yield vs. recoil energy for the Z6 3V WIMP candidates

# Bibliography

- [1] L. Bergström and A. Goobar, *Cosmology and Particle Astrophysics*, 2nd ed. (Springer Praxis Books, Chichester, UK, 2006).
- [2] G. Bertone, D. Hooper, and J. Silk, Phys. Rept. **405**, 279 (2005).
- [3] C. Amsler *et al.*, Phys. Lett. **B667**, 1 (2008).
- [4] F. Englert and R. Brout, Phys. Rev. Lett. **13**, 321 (1964).
- [5] P.W. Higgs, Phys. Lett. **12**, 132 (1964).
- [6] P.W. Higgs, Phys. Rev. Lett. **13**, 508 (1964).
- [7] G. Aad *et al.*, arXiv:0901.0512v4 (2008).
- [8] R. Adolphi *et al.*, JINST **3**, S08004 (2008).
- [9] J. Hosaka *et al.*, Phys. Rev. **D74**, 032002 (2006).
- [10] J. Schechter and J.W.F. Valle, Phys. Rev. **D22**, 2227 (1980).
- [11] A. de Gouvea and J.W.F. Valle, Phys. Lett. **B501**, 115 (2001).
- [12] A. Einstein, Sitz. Preuss. Akad. Wiss. **1**, 142 (1917).
- [13] E. Wright *et al.*, Astrophys. J. **396**, L13 (1992).
- [14] G. Efstathiou, J.R. Bond, and S.D.M. White, Mon. Not. Roy. Astron. Soc. **258**, 1P (1992).
- [15] L.M. Krauss and M.S. Turner, Gen. Rel. Grav. **27**, 1137 (1995).

## BIBLIOGRAPHY

- [16] J.P. Ostriker and P.J. Steinhardt, *Nature* **377**, 600 (1995).
- [17] R.M. Wald, *General Relativity* (The University of Chicago Press, Chicago, Illinois, 1984).
- [18] P.J.E. Peebles, in *Principles of Physical Cosmology*, edited by P.W. Anderson, A.S. Wightman, and S.B. Treiman (Princeton University Press, Princeton, New Jersey, 1993).
- [19] H.P. Robertson, *Proc. N.A.S.* **15**, 822 (1935).
- [20] A.G. Walker, *Proc. London Math. Soc.* **42**, 90 (1936).
- [21] A. Friedmann, *Z. Phys.* **10**, 377 (1922).
- [22] A. Friedmann, *Z. Phys.* **21**, 326 (1924).
- [23] L. Bergström, *Rept. Prog. Phys.* **63**, 793 (2000).
- [24] E. Komatsu *et al.*, *Astrophys. J. Suppl.* **192**, 18 (2011).
- [25] A.G. Riess *et al.*, *Astrophys. J.* **730**, 119 (2011).
- [26] A.G. Riess *et al.*, *Astrophys. J.* **699**, 539 (2009).
- [27] W.L. Freedman *et al.*, *Astrophys. J.* **553**, 47 (2001).
- [28] P.J.E. Peebles and B. Ratra, *Rev. Mod. Phys.* **75**, 559 (2003).
- [29] A. Albrecht *et al.*, *arXiv:astro-ph/0609591v1* (2006).
- [30] R.R. Caldwell, R. Dave, and P.J. Steinhardt, *Phys. Rev. Lett.* **80**, 1582 (1998).
- [31] J. Frieman, M.S Turner, and D. Huterer, *Ann. Rev. Astron. Astrophys.* **46**, 385 (2008).
- [32] A. Guth, *Phys. Rev.* **D23**, 347 (1981).
- [33] E. Gaswiser and J. Silk, *Phys. Rept.* **333**, 245 (2000).
- [34] J.C. Baez and J. Huerta, *arXiv:0904.1556v2* (2010).

## BIBLIOGRAPHY

- [35] S.L. Glashow, in *Fifth Workshop on Grand Unification* (World Scientific, Singapore, 1984), p. 88, edited by K. Kang, H. Fried, and F. Frampton.
- [36] I. Affleck and M. Dine, Nucl. Phys. **B249**, 361 (1985).
- [37] A. Riotto and M. Trodden, Ann. Rev. Nucl. Part. Sci. **49**, 35 (1999).
- [38] M. Longair, *Galaxy Formation*, 2nd ed. (Springer, Berlin, 2006).
- [39] T. Padmanabhan, *Structure Formation in the Universe* (Cambridge University Press, Cambridge, UK, 1993).
- [40] E. Hubble, Proc. N.A.S. **15**, 168 (1929).
- [41] E.W. Kolb and M.S. Turner, *The Early Universe* (Addison-Wesley, Reading, MA, 1990).
- [42] S. Sarkar, Rept. Prog. Phys. **59**, 1493 (1996).
- [43] V. Mukhanov, Int. J. Theor. Phys. **43**, 669 (2004).
- [44] M. Pospelov and J. Pradler, Ann. Rev. Nucl. Part. Sci. **60**, 539 (2010).
- [45] K.A. Olive, CERN Yellow Report **CERN-2010-002**, 149 (2010).
- [46] V.F. Shvartsman, JETP Lett. **9**, 184 (1969).
- [47] G. Steigman, D.N. Schramm, and J.E. Gunn, Phys. Lett. **B66**, 202 (1977).
- [48] D. Buskulic *et al.*, Phys. Lett. **B313**, 520 (1993).
- [49] R. Akers *et al.*, Z. Phys. **C65**, 47 (1995).
- [50] P. Abreu *et al.*, Z. Phys. **C74**, 577 (1997).
- [51] M. Acciarri *et al.*, Phys. Lett. **B431**, 199 (1998).
- [52] E. Komatsu *et al.*, Astrophys. J. Suppl. **180**, 330 (2009).
- [53] K.A. Olive and E.D. Skillman, Astrophys. J. **617**, 29 (2004).



## BIBLIOGRAPHY

- [54] E. Vangioni-Flam, K.A. Olive, B.D. Fields, and M. Cassé, *Astrophys. J.* **585**, 611 (2003).
- [55] S.G. Ryan, T.C. Beers, K.A. Olive, B.D. Fields, and J.E. Norris, *Astrophys. J.* **530**, L57 (2000).
- [56] R.H. Cyburt, B.D. Fields, and K.A. Olive, *New. Aston.* **6**, 215 (2001).
- [57] M. Pettini, B.J. Zych, M.T. Murphy, A. Lewis, and C.C. Steidel, *Mon. Not. Roy. Astron. Soc.* **391**, 1499 (2008).
- [58] R.H. Cyburt, B.D. Fields, and K.A. Olive, *JCAP* **0811**, 012 (2008).
- [59] B.D. Fields and K.A. Olive, *Astrophys. J.* **516**, 797 (1999).
- [60] F. Spite and M. Spite, *Astron. Astrophys.* **115**, 357 (1982).
- [61] J. Melendez and I. Ramirez, *Astrophys. J.* **615**, L33 (2004).
- [62] A. Hosford, S.G. Ryan, A.E. Garcia Perez, J.E. Norris, and K.A. Olive, *arXiv:0811.2506v1* (2008).
- [63] O. Richard, G. Michaud, and J. Richer, *Astrophys. J.* **619**, 538 (2005).
- [64] R.H. Cyburt *et al.*, *JCAP* **0910**, 021 (2009).
- [65] A.A. Penzias and R.W. Wilson, *Astrophys. J.* **142**, 419 (1965).
- [66] G.F. Smoot *et al.*, *Astrophys. J.* **396**, L1 (1992).
- [67] J.B. Peterson *et al.*, *Astrophys. J.* **532**, L83 (2000).
- [68] A.D. Miller *et al.*, *Astrophys. J. Suppl.* **140**, 115 (2002).
- [69] S. Padin *et al.*, *Astrophys. J.* **549**, L1 (2001).
- [70] E.M. Leitch *et al.*, *Astrophys. J.* **532**, 37 (2000).
- [71] S. Hanany *et al.*, *Astrophys. J.* **545**, L5 (2000).
- [72] B.P. Crill *et al.*, *Astrophys. J. Suppl.* **148**, 527 (2002).
- [73] C.L. Bennet *et al.*, *Astrophys. J. Suppl.* **148**, 1 (2003).

## BIBLIOGRAPHY

- [74] N. Jarosik *et al.*, *Astrophys. J. Suppl.* **192**, 14 (2011).
- [75] J.L. Weiland *et al.*, *Astrophys. J. Suppl.* **192**, 19 (2011).
- [76] D. Larson *et al.*, *Astrophys. J. Suppl.* **192**, 16 (2011).
- [77] B. Gold *et al.*, *Astrophys. J. Suppl.* **192**, 15 (2011).
- [78] C. Bennet *et al.*, *Astrophys. J. Suppl.* **192**, 17 (2011).
- [79] L. Page *et al.*, *Astrophys. J.* **585**, 566 (2003).
- [80] G. Hinshaw *et al.*, *Astrophys. J. Suppl.* **170**, 288 (2007).
- [81] G. Hinshaw *et al.*, *Astrophys. J. Suppl.* **180**, 225 (2009).
- [82] E.L. Wright, G. Hinshaw, and C.L. Bennett, *Astrophys. J.* **458**, L53 (1996).
- [83] C. Bennett *et al.*, *Astrophys. J. Suppl.* **148**, 97 (2003).
- [84] C.L. Reichardt *et al.*, *Astrophys. J.* **694**, 1200 (2009).
- [85] M.L. Brown *et al.*, *Astrophys. J.* **705**, 978 (2009).
- [86] R.K. Sachs and A.M. Wolfe, *Astrophys. J.* **147**, 73 (1967).
- [87] J. Silk, *Astrophys. J.* **151**, 459 (1968).
- [88] K. Nakamura *et al.*, *J. Phys. G* **37**, 075021 (2010), with 2011 partial update for the 2012 edition.
- [89] U. Seljak and M. Zaldarriaga, *Astrophys. J.* **469**, 437 (1996).
- [90] A. Lewis, A. Challinor, and A. Lasenby, *Astrophys. J.* **538**, 473 (2000).
- [91] D. Scott and A. Moss, *Mon. Not. Roy. Astron. Soc.* **397**, 445 (2009).
- [92] L.A.L. da Silva, *Astrophys. Space Sci.* **202**, 215 (1993).
- [93] S. Chandrasekhar, *Mon. Not. Roy. Astron. Soc.* **95**, 207 (1935).
- [94] M.M. Phillips, *Astrophys. J.* **413**, L105 (1993).

## BIBLIOGRAPHY

- [95] A.G. Riess *et al.*, Astron. J. **116**, 1009 (1998).
- [96] S. Perlmutter *et al.*, Astrophys. J. **517**, 565 (1999).
- [97] C. Contreras, Astron. J. **139**, 519 (2010).
- [98] M. Hicken *et al.*, Astrophys. J. **700**, 1097 (2009).
- [99] M. Kowalski *et al.*, Astrophys. J. **686**, 749 (2008).
- [100] S. Jha *et al.*, Astron. J. **131**, 527 (2006).
- [101] A.G. Riess *et al.*, Astron. J. **117**, 707 (1999).
- [102] K. Krisciunas *et al.*, Astron. J. **130**, 2453 (2005).
- [103] M. Hamuy *et al.*, Astron. J. **112**, 2391 (1996).
- [104] G. Miknaitis, Astrophys. J. **666**, 674 (2007).
- [105] P. Astier, Astron. Astrophys. **447**, 31 (2006).
- [106] R.A. Knop *et al.*, Astrophys. J. **598**, 102 (2003).
- [107] R. Amanullah *et al.*, Astron. Astrophys. **486**, 375 (2008).
- [108] B.J. Barris *et al.*, Astrophys. J. **602**, 571 (2004).
- [109] J.A. Holtzman *et al.*, Astron. J. **136**, 2306 (2008).
- [110] N. Suzuki *et al.*, arXiv:1105:3470v1 (2011), submitted to Astrophys. J.
- [111] R. Amanullah *et al.*, Astrophys. J. **716**, 712 (2010).
- [112] A.G. Riess *et al.*, Astrophys. J. **659**, 98 (2007).
- [113] J.L. Tonry *et al.*, Astrophys. J. **594**, 1 (2003).
- [114] V.M. Wood-Vasey, Astrophys. J. **666**, 694 (2007).
- [115] R. Kesler *et al.*, Astrophys. J. Suppl. **185**, 32 (2009).
- [116] J. Guy *et al.*, arXiv:1010.4743v1 (2010), accepted to Astron. Astrophys.

## BIBLIOGRAPHY

- [117] E.V. Linder, Phys. Rev. Lett. **90**, 091301 (2003).
- [118] W.J. Percival *et al.*, Mon. Not. Roy. Astron. Soc. **401**, 2148 (2010).
- [119] D.J. Eisenstein and M. White, Phys. Rev. **D70**, 103523 (2004).
- [120] H.-J. Seo and D.J. Eisenstein, Astrophys. J. **598**, 720 (2003).
- [121] C. Alcock and B. Paczyński, Nature **281**, 358 (1979).
- [122] P.J. Outram *et al.*, Mon. Not. Roy. Astron. Soc. **342**, 483 (2003).
- [123] S.M. Croom *et al.*, Mon. Not. Roy. Astron. Soc. **349**, 1397 (2004).
- [124] T.J. Broadhurst, R.S. Ellis, D.C. Koo, and A.S. Szalay, Nature **343**, 726 (1990).
- [125] W.J. Percival *et al.*, Mon. Not. Roy. Astron. Soc. **327**, 1297 (2001).
- [126] C.J. Miller and D.J. Batuski, Astrophys. J. **551**, 635 (2001).
- [127] W.J. Percival *et al.*, Mon. Not. Roy. Astron. Soc. **353**, 1201 (2004).
- [128] M. Tegmark *et al.*, Astrophys. J. **606**, 702 (2004).
- [129] D.J. Eisenstein *et al.*, Astrophys. J. **633**, 560 (2005).
- [130] D.G. York *et al.*, Astron. J. **120**, 1579 (2000).
- [131] M. Colless *et al.*, arXiv:astro-ph/0306581v1 (2003).
- [132] M.A. Strauss *et al.*, Astron. J. **124**, 1810 (2002).
- [133] K.N. Abazajian *et al.*, Astrophys. J. Suppl. **182**, 543 (2009).
- [134] J.P. Ostriker and P.J. Steinhardt, arXiv:astro-ph/9505066v1 (1995).
- [135] F. Zwicky, Helv. Phys. Acta **6**, 110 (1933).
- [136] F. Zwicky, Astrophys. J. **86**, 217 (1937).
- [137] G. Steigman and M.S. Turner, Nucl. Phys. **B253**, 375 (1985).

## BIBLIOGRAPHY

- [138] G. Jungman, M. Kamionkowski, and K. Griest, Phys. Rept. **267**, 195 (1996).
- [139] B. Lee and S. Weinberg, Phys. Rev. Lett. **39**, 165 (1977).
- [140] S.M. Faber and J.S. Gallagher, Ann. Rev. Astron. Astrophys. **17**, 135 (1979).
- [141] M.R. Blanton *et al.*, Astrophys. J. **592**, 819 (2003).
- [142] J.H. Oort, in *La Structure et L'Évolution de L'Univers*, edited by R. Stoops (Onzième Conseil de Physique, Solvay, Bruxelles, 1958), p. 163.
- [143] J.R. Gott and E.L. Turner, Astrophys. J. **209**, 1 (1976).
- [144] K. Kuijken and G. Gilmore, Astrophys. J. **367**, L9 (1991).
- [145] J.N. Bahcall, C. Flynn, and A. Gould, Astrophys. J. **389**, 234 (1992).
- [146] C. Flynn and B. Fuchs, Mon. Not. Roy. Astron. Soc. **270**, 471 (1994).
- [147] M.A.C. Perryman *et al.*, Astron. Astrophys. **258**, 7 (1992).
- [148] J. Holmberg and C. Flynn, Mon. Not. Roy. Astron. Soc. **352**, 440 (2004).
- [149] J. Holmberg and C. Flynn, Mon. Not. Roy. Astron. Soc. **313**, 209 (2000).
- [150] C. Flynn *et al.*, Mon. Not. Roy. Astron. Soc. **372**, 1149 (2006).
- [151] K.A. Olive, arXiv:astro-ph/0301505v2 (2003).
- [152] V.C. Rubin and W.K. Ford, Astrophys. J. **159**, 379 (1970).
- [153] V.C. Rubin and W.K. Ford, Astrophys. J. **170**, 25 (1971).
- [154] P. Salucci and M. Persic, in *the Sesto DM1996 Conference*, edited by M. Persic and P. Salucci (ASP Proceedings Conference Series, Chicago, 1997), Vol. 117, p. 1, available at arXiv:astro-ph/9703027v1.
- [155] B.F. Burke, K.C. Turner, and M.A. Tuve, Carnegie Yrb. **63**, 341 (1964).
- [156] M.S. Roberts, Astrophys. J. **144**, 639 (1966).

## BIBLIOGRAPHY

- [157] V.C. Rubin, N. Thonnard, and W.K. Ford, *Astrophys. J.* **238**, 471 (1980).
- [158] S. Casertano and J.H. Van Gorkom, *Astrophys. J.* **101**, 1231 (1991).
- [159] K.G. Begeman, *Astron. Astrophys.* **223**, 47 (1989).
- [160] M. Persic and P. Salucci, arXiv:astro-ph/9502091v1 (1995), submitted to *Astrophys. J. Suppl.*
- [161] B. Catinella, R. Giovanelli, and M.P. Haynes, *Astrophys. J.* **640**, 751 (2006).
- [162] R.H. Sanders and M.A.W. Verheijen, *Astrophys. J.* **503**, 97 (1998).
- [163] M. Persic, P. Salucci, and F. Stel, *Mon. Not. Roy. Astron. Soc.* **281**, 27 (1996).
- [164] Y. Sofue, M. Honma, and T. Omodaka, arXiv:0811.0859v2 (2008), submitted to *Publ. Astron. Soc. Jpn.*
- [165] H.T. Freudenreich, *Astrophys. J.* **492**, 495 (1998).
- [166] V.C. Rubin, D. Burstein, W.K. Ford, and N. Thonnard, *Astrophys. J.* **289**, 81 (1985).
- [167] D. Burstein and V.C. Rubin, *Astrophys. J.* **297**, 423 (1985).
- [168] M. Persic and P. Salucci, *Astrophys. J.* **368**, 60 (1991).
- [169] W.B. Burton and M.A. Gordon, *Astron. Astrophys.* **63**, 7 (1978).
- [170] D.P. Clemens, *Astrophys. J.* **295**, 422 (1985).
- [171] M. Fich, L. Blitz, and A.A. Stark, *Astrophys. J.* **342**, 272 (1989).
- [172] L. Blitz, M. Fich, and A.A. Stark, *Astrophys. J. Suppl.* **49**, 183 (1982).
- [173] S. Demers and P. Battinelli, *Astron. Astrophys.* **473**, 143 (2007).
- [174] M. Honma and Y. Sofue, *Publ. Astron. Soc. Jpn.* **49**, 453 (1997).
- [175] M. Honma and Y. Sofue, *Publ. Astron. Soc. Jpn.* **49**, 539 (1997).

## BIBLIOGRAPHY

- [176] M. Honma *et al.*, Publ. Astron. Soc. Jpn. **59**, 897 (2007).
- [177] M. Honma, N. Kawaguchi, and T. Sasao, Proc. SPIE **4015**, 624 (2000), edited by H.R. Buthcer.
- [178] G. de Vaucouleurs, Astrophys. J. **128**, 465 (1958).
- [179] M.I. Wilkinson and N.W. Evans, Mon. Not. Roy. Astron. Soc. **310**, 645 (1999).
- [180] T. Sakamoto, M. Chiba, and T.C. Beers, Astron. Astrophys. **397**, 899 (2003).
- [181] M. Steinmetz *et al.*, Astron. J. **132**, 1645 (2006).
- [182] T.C. Beers *et al.*, Astron. J. **199**, 2866 (2000).
- [183] M.C. Smith *et al.*, Mon. Not. Roy. Astron. Soc. **379**, 755 (2007).
- [184] F.J. Kerr and D. Lynden-Bell, Mon. Not. Roy. Astron. Soc. **221**, 1023 (1986).
- [185] H.J. Mo, S. Mao, and S.D.M. White, Mon. Not. Roy. Astron. Soc. **295**, 319 (1998).
- [186] R.B. Tully and J.R. Fisher, Astron. Astrophys. **54**, 661 (1977).
- [187] J. Pizagno *et al.*, Astrophys. J. **633**, 844 (2005).
- [188] J.L. S ersic, Bolet n de la Asociaci n Argentina de Astronom a **6**, 41 (1963).
- [189] L. Ciotti and G. Bertin, Astron. Astrophys. **352**, 447 (1999).
- [190] D. Merritt, J.F. Navarro, A. Ludlow, and A. Jenkins, Astrophys. J. **624**, L85 (2005).
- [191] E. Memola, P. Salucci, and A. Babic, arXiv:1106.5133v1 (2011), accepted to Astron. Astrophys.
- [192] F. Bertola, A. Pizzella, M. Persic, and P. Salucci, Astrophys. J. **416**, L45 (1993).

## BIBLIOGRAPHY

- [193] S. Ettori and A.C. Fabian, *Mon. Not. Roy. Astron. Soc.* **369**, L42 (2006).
- [194] A. Cavaliere and R. Fusco-Femiano, *Astron. Astrophys.* **9**, 137 (1976).
- [195] P. Salucci and A. Burkert, *Astrophys. J.* **537**, L9 (2000).
- [196] M. Milgrom, *Astrophys. J.* **270**, 365 (1983).
- [197] M. Milgrom, *Astrophys. J.* **270**, 371 (1983).
- [198] D.S. Davis and R.E. White, *Astrophys. J.* **470**, L35 (1996).
- [199] M. Loewenstein and R.E. White, *Astrophys. J.* **518**, 50 (1999).
- [200] Y. Fukazawa, J.G. Botoya-Nonesca, J. Pu, A. Ohto, and N. Kawano, *Astrophys. J.* **636**, 698 (2006).
- [201] M.C. Weisskopf, H.D. Tananbaum, L.P. Van Speybroeck, and S.L. O'Dell, *Proc. SPIE* **4012**, 2 (2000).
- [202] F. Jansen *et al.*, *Astron. Astrophys.* **365**, L1 (2001).
- [203] R. Nagino and K. Matsushita, *Astron. Astrophys.* **501**, 157 (2009).
- [204] D.A. Buote and P.J. Humphrey, arXiv:1104.0012v1 (2011).
- [205] D. Makarov and I. Karachentsev, arXiv:1011.6277v1 (2010), accepted to *Mon. Not. Roy. Astron. Soc.*
- [206] S. van den Bergh, *Astron. Astrophys. Rev.* **9**, 273 (1999).
- [207] R.G. Carlberg *et al.*, *Astrophys. J.* **552**, 427 (2001).
- [208] A. Biviano, in *Tracing Cosmic Evolution with Galaxy Clusters*, edited by S. Borgani, M. Mezzetti, and R. Valdarnini (ASP Conference Series, San Francisco, 2002), Vol. 268, p. 127.
- [209] G. Giuricin *et al.*, *Astron. Astrophys.* **199**, 85 (1988).
- [210] A. Dressler, *Astrophys. J.* **236**, 351 (1980).



## BIBLIOGRAPHY

- [211] A. Biviano, in *Constructing the Universe with Clusters of Galaxies*, edited by F. Durret and D. Gerbal (IAP meeting, Paris, France, 2000), available at arXiv:astro-ph/0010409v1.
- [212] R.A. Sunyaev and Ya.B. Zeldovich, *Comments Astrophys. Space Phys.* **4**, 173 (1972).
- [213] D. Layzer, *Astrophys. J.* **138**, 174 (1963).
- [214] S. Smith, *Astrophys. J.* **83**, 23 (1936).
- [215] J. Binney and S. Tremaine, *Galactic Dynamics* (Princeton University Press, Princeton, New Jersey, 1987).
- [216] A. Biviano, in *Lectures at International School of Physics, Enrico Fermi* (Società Italiana di Fisica, Varenna, Italy, 2008), available at arXiv:0911.3535v1.
- [217] N.H. Abel, *J. Reine u. Angew. Math.* **1**, 153 (1826).
- [218] G.A. Mamon and G. Boué, *Mon. Not. Roy. Astron. Soc.* **401**, 2433 (2010).
- [219] M. Hénon, *Astron. Astrophys.* **114**, 211 (1982).
- [220] R.G. Carlberg, *Astrophys. J.* **485**, L13 (1997).
- [221] R.P. van der Marel, J. Magorrian, R.G. Carlberg, H.K.C. Yee, and E. Ellingson, *Astron. J.* **199**, 2038 (2000).
- [222] M. Girardi, G. Giuricin, F. Mardirossian, M. Mezzetti, and W. Boschin, *Astrophys. J.* **505**, 74 (1998).
- [223] L.S. The and S.D.M. White, *Astron. J.* **92**, 1248 (1986).
- [224] A. Biviano, P. Katgert, and T. Thomas, *Astron. Astrophys.* **387**, 8 (2002).
- [225] P. Popesso, A. Biviano, H. Bhoeringer, M. Romaniello, and W. Voges, *Astron. Astrophys.* **433**, 431 (2005).
- [226] G. Giuricin, C. Marinoni, L. Ceriani, and A. Pisani, *Astrophys. J.* **543**, 178 (2000).

## BIBLIOGRAPHY

- [227] M. Girardi, P. Manzato, M. Mezzetti, G. Giuricin, and F. Limboz, *Astrophys. J.* **569**, 720 (2002).
- [228] D. Yentis *et al.*, in *Digitized Optical Sky Survey*, edited by H. MacGillivray and E. Thomson (Kluwer, Dordrecht, 1992), p. 67.
- [229] R.L. Pennington *et al.*, *Publ. Astron. Soc. Pac.* **105**, 521 (1993).
- [230] M.J. Ledlow *et al.*, *Astron. J.* **112**, 388 (1996).
- [231] A. Zabludoff and J.S. Mulchaey, *Astrophys. J.* **496**, 39 (1998).
- [232] G. Paturel *et al.*, *Astron. Astrophys. Suppl.* **124**, 109 (1997).
- [233] M. Girardi, S. Borgani, G. Giuricin, F. Mardirossian, and M. Mezzetti, *Astrophys. J.* **530**, 62 (2000).
- [234] M. Girardi and G. Giuricin, *Astrophys. J.* **540**, 45 (2000).
- [235] P. Norberg *et al.*, *Mon. Not. Roy. Astron. Soc.* **336**, 907 (2002).
- [236] M.R. Blanton *et al.*, *Astron. J.* **121**, 2358 (2001).
- [237] J. Liske *et al.*, *Mon. Not. Roy. Astron. Soc.* **344**, 307 (2003).
- [238] S.P. Driver *et al.*, *Mon. Not. Roy. Astron. Soc.* **360**, 81 (2005).
- [239] G. Kauffmann, J.M. Colberg, A. Kiaferio, and S.D.M. White, *Mon. Not. Roy. Astron. Soc.* **303**, 188 (1999).
- [240] A.J. Benson, S. Cole, C.S. Frenk, C.M. Baugh, and C.G. Lacey, *Mon. Not. Roy. Astron. Soc.* **311**, 793 (2000).
- [241] G. Efstathiou, R.S. Ellis, and B.A. Peterson, *Mon. Not. Roy. Astron. Soc.* **232**, 431 (1988).
- [242] J. Trümper, *Adv. Space. Res.* **2**, 241 (1983).
- [243] J. Trümper, *Q. J. R. Astron. Soc.* **33**, 165 (1992).
- [244] J. Trümper, *Science* **260**, 1769 (1993).

## BIBLIOGRAPHY

- [245] G.W. Pratt, J.H. Croston, M. Arnaud, and H. Böhringer, *Astron. Astrophys.* **498**, 361 (2009).
- [246] A. Ortiz-Gil, L. Guzzo, P. Schuecker, H. Böhringer, and C.A. Collins, *Mon. Not. Roy. Astron. Soc.* **348**, 325 (2004).
- [247] Y.-J. Xue and X.-P. Wu, *Astrophys. J.* **539**, 65 (2000).
- [248] S. Borgani, in *Lectures for 2005 Guillermo Haro Summer School on Clusters* (Springer, Tonantzintla, Mexico, 2006).
- [249] N. Kaiser, *Mon. Not. Roy. Astron. Soc.* **222**, 323 (1986).
- [250] G.L. Bryan and M.L. Norman, *Astrophys. J.* **495**, 80 (1998).
- [251] M. Arnaud, E. Pointecouteau, and G.W. Pratt, *Astron. Astrophys.* **441**, 893 (2005).
- [252] O. Kotov and A. Vikhlinin, *Astrophys. J.* **633**, 781 (2005).
- [253] J.P.F. Osmond and T.J. Ponman, *Mon. Not. Roy. Astron. Soc.* **350**, 1511 (2004).
- [254] T.H. Reiprich and H. Böhringer, *Astrophys. J.* **567**, 716 (2002).
- [255] B.J. Maughan, *Astrophys. J.* **668**, 772 (2007).
- [256] J.F. Navarro, C.S. Frenk, and S.D.M. White, *Mon. Not. Roy. Astron. Soc.* **275**, 720 (1995).
- [257] A.E. Evrard and J.P. Henry, *Astrophys. J.* **383**, 95 (1991).
- [258] G.M. Voit, *Rev. Mod. Phys.* **77**, 207 (2005).
- [259] R.A. Burenin *et al.*, *Astrophys. J. Suppl.* **172**, 561 (2007).
- [260] A. Vikhlinin *et al.*, *Astrophys. J.* **692**, 1033 (2009).
- [261] A. Vikhlinin *et al.*, *Astrophys. J.* **692**, 1060 (2009).
- [262] A.V. Kravtsov, A. Vikhlinin, and D. Nagai, *Astrophys. J.* **650**, 128 (2006).

## BIBLIOGRAPHY

- [263] J. Tinker *et al.*, *Astrophys. J.* **688**, 709 (2008).
- [264] C.R. Mullis *et al.*, *Astrophys. J.* **607**, 175 (2004).
- [265] J.P. Henry, *Astrophys. J.* **609**, 603 (2004).
- [266] O. Chwolson, *Astron. Nachr.* **221**, 329 (1924).
- [267] A. Einstein, *Science* **84**, 506 (1936).
- [268] J.A. Tyson, R.A. Wenk, and F. Valdes, *Astrophys. J.* **349**, L1 (1990).
- [269] G. Fahlman, N. Kaiser, G. Squires, and D. Woods, *Astrophys. J.* **437**, 56 (1994).
- [270] G.A. Luppino and N. Kaiser, *Astrophys. J.* **475**, 20 (1997).
- [271] A. Zitrin *et al.*, arXiv:1103.5618v2 (2011), submitted to *Astrophys. J.*
- [272] R. Link and M.J. Pierce, *Astrophys. J.* **502**, 63 (1998).
- [273] J. Gilmore and P. Natarajan, *Mon. Not. Roy. Astron. Soc.* **396**, 354 (2009).
- [274] E. Jullo, *Science* **329**, 924 (2010).
- [275] K. Umetsu, T. Broadhurst, A. Zitrin, E. Medezinski, and L.-Y. Hsu, arXiv:1011.3044v2 (2010), accepted to *Astrophys. J.*
- [276] M.B. Bayliss *et al.*, *Astrophys. J. Suppl.* **193**, 8 (2011).
- [277] G.M. Bernstein and M. Jarvis, *Astron. J.* **123**, 583 (2002).
- [278] M. White, L. van Waerbeke, and J. Mackey, *Astrophys. J.* **575**, 640 (2002).
- [279] C.A. Metzler, M. White, and C. Loken, *Astrophys. J.* **547**, 560 (2001).
- [280] M. Bartelmann and P. Schneider, *Phys. Rept.* **340**, 291 (2001).
- [281] E.S. Sheldon *et al.*, *Astron. J.* **127**, 2544 (2004).
- [282] E.S. Sheldon *et al.*, *Astrophys. J.* **703**, 2217 (2009).

## BIBLIOGRAPHY

- [283] E.S. Sheldon *et al.*, *Astrophys. J.* **703**, 2232 (2009).
- [284] J.L. Tinker, D.H. Weinberg, Z. Zheng, and I. Zehavi, *Astrophys. J.* **631**, 41 (2005).
- [285] D.E. Johnston *et al.*, arXiv:0709.1159v1 (2007).
- [286] M. Arnaud *et al.*, arXiv:0902.4890v1 (2009), white paper submitted to the Astro2010 Decadal Survey.
- [287] M. Markevitch *et al.*, *Astrophys. J.* **567**, L27 (2002).
- [288] D. Clowe *et al.*, *Astrophys. J.* **648**, L109 (2006).
- [289] D. Clowe, A. Gonzalez, and M. Markevitch, *Astrophys. J.* **604**, 596 (2004).
- [290] Y. Rephaeli, *Annu. Rev. Astron. Astrophys.* **33**, 541 (1995).
- [291] M. Birkinshaw, *Phys. Rept.* **310**, 97 (1999).
- [292] M. Zemcov *et al.*, arXiv:1005.3824v1 (2010), accepted to *Astron. Astrophys.*
- [293] S. Nozawa, N. Itoh, and Y. Kohyama, *Astrophys. J.* **508**, 17 (1998).
- [294] S.H. Hansen, *New. Astron.* **9**, 279 (2004).
- [295] A.D. Dolgov, S.H. Hansen, S. Pastor, and D.V. Semikoz, *Astrophys. J.* **554**, 74 (2001).
- [296] J. Bernstein and S. Dodelson, *Phys. Rev.* **D41**, 354 (1990).
- [297] R. Stanek, E. Rasia, A.E. Evrard, F. Pearce, and L. Gazzola, *Astrophys. J.* **715**, 1508 (2010).
- [298] S. Wang, J. Khoury, Z. Haiman, and M. May, *Phys. Rev.* **D70**, 123008 (2004).
- [299] L. Ferramacho and A. Blanchard, arXiv:1105.3680v1 (2011), accepted to *Astron. Astrophys.*

## BIBLIOGRAPHY

- [300] M. Shimon, S. Sadeh, and Y. Rephaeli, arXiv:1009.4110v1 (2010).
- [301] M. Roncarelli *et al.*, Mon. Not. Roy. Astron. Soc. **402**, 923 (2010).
- [302] F.K. Hansen *et al.*, Mon. Not. Roy. Astron. Soc. **361**, 753 (2005).
- [303] S. Joudaki, J. Smidt, A. Amblard, and A. Cooray, JCAP **1008**, 027 (2010).
- [304] M. Bonamente *et al.*, Astrophys. J. **675**, 106 (2008).
- [305] J.-B. Melin *et al.*, Astron. Astrophys. **525**, A139 (2011).
- [306] J.E. Carlstrom *et al.*, arXiv:0907.4445v2 (2011), accepted to Publ. Astron. Soc. Pac.
- [307] D.S. Swetz *et al.*, Astrophys. J. Suppl. **194**, 41 (2011).
- [308] Z. Staniszewski *et al.*, Astrophys. J. **701**, 32 (2009).
- [309] K. Story *et al.*, arXiv:1102.2189v1 (2011), submitted to Astrophys. J.
- [310] T.A. Marriage *et al.*, arXiv:1010.1065v2 (2011), accepted to Astrophys. J.
- [311] K. Vanderlinde *et al.*, Astrophys. J. **722**, 1180 (2010).
- [312] F. Menanteau *et al.*, arXiv:1006.5126v2 (2010), accepted to Astrophys. J.
- [313] N. Sehgal *et al.*, Astrophys. J. **732**, 44 (2011).
- [314] K. Andersson *et al.*, arXiv:1006.3068v1 (2010), submitted to Astrophys. J.
- [315] T. Plagge *et al.*, Astrophys. J. **716**, 1118 (2010).
- [316] J. Ellis, Phys. Scripta **T85**, 221 (2000).
- [317] M. Taoso, G. Bertone, and A. Masiero, JCAP **0803**, 022 (2008).
- [318] R.D. Peccei and H.R. Quinn, Phys. Rev. Lett. **37**, 1440 (1977).

## BIBLIOGRAPHY

- [319] S. Weinberg, Phys. Rev. Lett. **40**, 233 (1978).
- [320] F. Wilczek, Phys. Rev. Lett. **40**, 279 (1978).
- [321] P. Sikivie, Phys. Rev. Lett. **51**, 1415 (1983).
- [322] H. Peng *et al.*, Nucl. Instr. Meth. **A444**, 569 (2000).
- [323] M. Shibata *et al.*, J. Low Temp. Phys. **151**, 1043 (2008).
- [324] M. Milgrom, Astrophys. J. **270**, 384 (1983).
- [325] M. Milgrom, in *The Invisible Universe International Conference* (AIP Conf. Proc., Paris, 2010), Vol. 1241, pp. 139–153.
- [326] J.W. Moffat, JCAP **0603**, 004 (2006).
- [327] J.W. Moffat and V.T. Toth, arXiv:1104.2957v1 (2011), submitted to JCAP.
- [328] I. Kobzarev, L. Okun, and I. Pomeranchuk, Sov. J. Nucl. Phys. **3**, 837 (1966).
- [329] M. Pavsic, Int. J. Theor. Phys. **9**, 229 (1974).
- [330] R. Foot, H. Lew, and R.R. Volkas, Phys. Lett. **B272**, 67 (1991).
- [331] R. Foot and X.-G. He, Phys. Lett. **B267**, 509 (1991).
- [332] R.N. Mohapatra, S. Nussinov, and V.L. Teplitz, Phys. Rev. **D66**, 063002 (2002).
- [333] R. Foot, Int. J. Mod. Phys. **D13**, 2161 (2004).
- [334] R. Schabinger and J.D. Wells, Phys. Rev. **D72**, 093007 (2005).
- [335] B. Patt and F. Wilczek, arXiv:hep-ph/0605188v1 (2006), report no. MIT-CTP-3745.
- [336] J. Kumar and J.D. Wells, Phys. Rev. **D74**, 115017 (2006).
- [337] S.M. Barr, Phys. Rev. **D44**, 3062 (1991).

## BIBLIOGRAPHY

- [338] V.A. Kuzmin, Phys. Part. Nucl. **29**, 257 (1998).
- [339] D.E. Kaplan, M.A. Luty, and K.M. Zurek, Phys. Rev. **D79**, 115016 (2009).
- [340] D.J.H. Chung, E.W. Kolb, and A. Riotto, Phys. Rev. **D59**, 023501 (1998).
- [341] V.A. Kuzmin and I.I. Tkachev, JETP Lett. **68**, 271 (1998).
- [342] D.J.H. Chung, Phys. Rev. **D67**, 083514 (2003).
- [343] H. Ziaeeepour, Grav. Cosmol. Suppl. **6**, 128 (2000).
- [344] T. Kaluza, Sitzungsber. Preuss. Akad. Wiss. **H1**, 966 (1921).
- [345] O. Klein, Zeitschrift für Physik **37**, 895 (1926).
- [346] N. Arkani-Hamed, S. Dimopoulos, and G.R. Dvali, Phys. Lett. **B429**, 263 (1998).
- [347] N. Arkani-Hamed, S. Dimopoulos, and G.R. Dvali, Phys. Rev. **D59**, 086004 (1999).
- [348] L. Randall and R. Sundrum, Phys. Rev. Lett. **83**, 3370 (1999).
- [349] T. Appelquist, H.-C. Cheng, and B.A. Dobrescu, Phys. Rev. **D64**, 035002 (2001).
- [350] D. Hooper and S. Profumo, Phys. Rept. **453**, 29 (2007).
- [351] G. Servant and T.M.P. Tait, Nucl. Phys. **B650**, 391 (2003).
- [352] S. Arrenberg, L. Baudis, K. Kong, K.T. Matchev, and J. Yoo, Phys. Rev. **D78**, 056002 (2008).
- [353] S. Coleman, Nucl. Phys. **B262**, 263 (1985).
- [354] A. Kusenko and M. Shaposhnikov, Phys. Lett. **B418**, 46 (1998).
- [355] A. Kusenko, V. Kuzmin, M. Shaposhnikov, and P.G. Tinyakov, Phys. Rev. Lett. **80**, 3185 (1998).



## BIBLIOGRAPHY

- [356] J.L. Feng, A. Rajaraman, and F. Takayama, Phys. Rev. Lett. **91**, 011302 (2003).
- [357] G.G. Raffelt, *Stars as Laboratories for Fundamental Physics* (University of Chicago Press, Chicago, 1996).
- [358] S. Davidson, S. Hannestad, and G. Raffelt, JHEP **0005**, 003 (2000).
- [359] M.L. Perl, E.R. Lee, and D. Loomba, Annu. Rev. Nucl. Part. Sci. **59**, 47 (2009).
- [360] G.D. Starkman, A. Gould, R. Esmailzadeh, and S. Dimopoulos, Phys. Rev. **D41**, 3594 (1990).
- [361] D.O. Caldwell, J. Phys. **G17**, S325 (1991).
- [362] R.N. Mohapatra, F. Olness, R. Stroynowski, and V.L. Teplitz, Phys. Rev. **D60**, 115013 (1999).
- [363] S. Andreon, arXiv:1004.2785v1 (2010), accepted to Mon. Not. Roy. Astron. Soc.
- [364] M. Sun *et al.*, Astrophys. J. **693**, 1142 (2009).
- [365] A. Vikhlinin *et al.*, Astrophys. J. **640**, 691 (2006).
- [366] J. Silk, Astrophys. J. **256**, 514 (1982).
- [367] J.H. Jeans, Phil. Trans. R. Soc. London **A199**, 1 (1902).
- [368] W.J. Borucki *et al.*, Astrophys. J. **728**, 117 (2011).
- [369] W.J. Borucki *et al.*, arXiv:1102.0541v2 (2011), submitted to Astrophys. J.
- [370] J.G. Hills, Astron. J. **92**, 595 (1986).
- [371] D.J. Hegyi and K.A. Olive, Phys. Lett. **B126**, 28 (1983).
- [372] D.J. Hegyi and K.A. Olive, Astrophys. J. **303**, 56 (1986).
- [373] S.J. Blanksby and G.B. Ellison, Acc. Chem. Res. **36**, 255 (2003).

## BIBLIOGRAPHY

- [374] F. Donato, N. Fornengo, and S. Scopel, *Astropart. Phys.* **9**, 247 (1998).
- [375] P.J. McMillan, arXiv:1102.4340v1 (2011), accepted to *Mon. Not. Roy. Astron. Soc.*
- [376] L. Bergström, P. Ullio, and J.H. Buckley, *Astropart. Phys.* **9**, 137 (1998).
- [377] A. Frebel, *Astrophys. J.* **660**, L117 (2007).
- [378] M. Wardle and M. Walker, *Astrophys. J.* **527**, L109 (1999).
- [379] S. Malhotra, *Astrophys. J.* **448**, 138 (1995).
- [380] P. Richter, B.P. Wakker, and B.D. Savage, *Astrophys. J.* **586**, 230 (2003).
- [381] A. Marasco and F. Fraternali, arXiv:1010.3563v1 (2010), accepted to *Astron. Astrophys.*
- [382] T.L. Wilson and R. Mauersberger, *Astron. Astrophys.* **282**, L41 (1994).
- [383] J.N. Bregman, *Annu. Rev. Astron. Astrophys.* **45**, 221 (2007).
- [384] R.P. Olling and M.R. Merrifield, arXiv:astro-ph/9907353v1 (1999), accepted to *Mon. Not. Roy. Astron. Soc.*
- [385] K. Griest, *Astrophys. J.* **366**, 412 (1991).
- [386] A. De Rújula, P. Jetzer, and E. Massó, *Astron. Astrophys.* **254**, 99 (1992).
- [387] S.S. Kumar, *Astrophys. J.* **137**, 1121 (1963).
- [388] S.S. Kumar, *Astrophys. J.* **137**, 1126 (1963).
- [389] B.D. Fields, K. Freese, and D.S. Graff, *Astrophys. J.* **534**, 265 (2000).
- [390] G. Chabrier, *Astrophys. J.* **513**, L103 (1999).
- [391] K. Freese, *Phys. Rept.* **333–334**, 183 (2000).
- [392] W.W. Ober, M.F. El Eid, and K.J. Fricke, *Astron. Astrophys.* **119**, 61 (1983).
- [393] B.J. Carr, J.R. Bond, and W.D. Arnett, *Astrophys. J.* **227**, 445 (1984).

## BIBLIOGRAPHY

- [394] J.R. Bond, W.D. Arnett, and B.J. Carr, *Astrophys. J.* **280**, 825 (1984).
- [395] J. Bachall, C. Flynn, A. Gould, and S. Kirhakos, *Astrophys. J.* **435**, L51 (1994).
- [396] P.C. Boeshaar, J.A. Tyson, and G.M. Bernstein, *BAAS* **185**, 22.02 (1994).
- [397] D.S. Graff and K. Freese, *Astrophys. J.* **456**, L49 (1996).
- [398] S. Torres, J. Camacho, J. Isern, and E. García-Berro, *Astron. Astrophys.* **486**, 427 (2008).
- [399] E.Zackrisson and C. Flynn, *Astrophys. J.* **687**, 242 (2008).
- [400] B. Paczyński, *Astrophys. J.* **304**, 1 (1986).
- [401] P. Tisserand *et al.*, *Astron. Astrophys.* **469**, 387 (2007).
- [402] C. Alcock *et al.*, *Astrophys. J.* **542**, 281 (2000).
- [403] L. Wyrzykowski *et al.*, *Mon. Not. Roy. Astron. Soc.* **413**, 493 (2011).
- [404] C. Alcock *et al.*, *Astrophys. J.* **550**, L169 (2001).
- [405] C. Renault *et al.*, *Astron. Astrophys.* **324**, L69 (1997).
- [406] L. Mancini, S. Calchi Novati, Ph. Jetzer, and G. Scarpetta, *Astron. Astrophys.* **427**, 61 (2004).
- [407] L. Wyrzykowski *et al.*, arXiv:1106.2925v1 (2011), accepted to *Mon. Not. Roy. Astron. Soc.*
- [408] D.P. Bennett, *Astrophys. J.* **633**, 906 (2005).
- [409] C.A. Nelson *et al.*, arXiv:0902.2213v1 (2009), submitted to *Astrophys. J.*
- [410] D.N. Schramm and G. Steigman, *Astrophys. J.* **243**, 1 (1981).
- [411] K.A. Olive, *Adv. Space Res.* **42**, 581 (2008).
- [412] R.B. Firestone, V.S. Shirley, C.M. Baglin, S.Y.F. Chu, and J. Zipkin, *Table of Isotopes*, 8th ed. (John Wiley and Sons, Inc., New York, 1999).

## BIBLIOGRAPHY

- [413] V.M. Lobashev *et al.*, Phys. Lett. **B460**, 227 (1999).
- [414] Ch. Kraus *et al.*, Eur. Phys. J. **C40**, 447 (2005).
- [415] M. Beck *et al.*, J. Phys. Conf. Ser. **203**, 012097 (2010).
- [416] Z. Maki, M. Nakagawa, and S. Sakata, Prog. Theor. Phys. **28**, 870 (1962).
- [417] P. Pontecorvo, JETP **26**, 984 (1968).
- [418] B.T. Cleveland *et al.*, Astrophys. J. **496**, 505 (1998).
- [419] B. Aharmim *et al.*, Phys. Rev. **C81**, 055504 (2010).
- [420] Y. Ashie *et al.*, Phys. Rev. **D71**, 112005 (2005).
- [421] P. Adamson *et al.*, Phys. Rev. Lett. **101**, 131802 (2008).
- [422] K. Abe *et al.*, Phys. Rev. Lett. **107**, 041801 (2011).
- [423] B.W. Lee and S. Weinberg, Phys. Rev. Lett. **39**, 165 (1977).
- [424] C.S. Frenk, S.D.M. White, and M. Davis, Astrophys. J. **271**, 417 (1983).
- [425] J.R. Bond, G. Efstathiou, and J. Silk, Phys. Rev. Lett. **45**, 1980 (1980).
- [426] J. Lesgourgues and S. Pastor, Phys. Rept. **429**, 307 (2006).
- [427] H. Wayne and D.J. Eisenstein, Astrophys. J. **498**, 497 (1998).
- [428] S.D.M. White, C.S. Frenk, and M. Davis, Astrophys. J. **274**, L1 (1983).
- [429] J.R. Bond, J. Centrella, A.S. Szalay, and J. Wilson, in *Formation and Evolution of Galaxies and Large Structures in the Universe*, edited by J. Andouze and J. Tran Thanh Van (Reidel, Dordrecht, 1983), p. 87.
- [430] M.D. Lehnert *et al.*, Nature **467**, 940 (2010).
- [431] G.R. Blumenthal, S.M. Faber, J.R. Primack, and M.J. Rees, Nature **311**, 517 (1984).
- [432] K. Ichiki, M. Takada, and T. Takahashi, Phys. Rev. **D79**, 023520 (2009).

## BIBLIOGRAPHY

- [433] S.A. Thomas, F.B. Abdalla, and O. Lahav, Phys. Rev. Lett. **105**, 031301 (2010).
- [434] M. Srednicki, R. Watkins, and K.A. Olive, Nucl. Phys. **B310**, 693 (1988).
- [435] E. Gildener and S. Weinberg, Phys. Rev. **D13**, 3333 (1976).
- [436] L. Susskind, Phys. Rev. **D20**, 273 (1979).
- [437] A. Dighe and S. Ray, Phys. Rev. **D76**, 113001 (2007).
- [438] A. Donini, M. Maltoni, D. Meloni, P. Migliozi, and F. Terranova, JHEP **12**, 013 (2007).
- [439] B. Holdom *et al.*, PMC Phys. **A3**, 4 (2009).
- [440] L.M. Carpenter and A. Rajaraman, Phys. Rev. **D82**, 114019 (2010).
- [441] H.P. Niles, Phys. Rept. **110**, 1 (1984).
- [442] M.F. Sohnius, Phys. Rept. **128**, 39 (1985).
- [443] Y.A. Gol’fand and E.P. Likhtman, JETP Lett. **13**, 323 (1971).
- [444] D.V. Volkov and V.P. Akulov, JETP Lett. **16**, 438 (1972).
- [445] J. Wess and B. Zumino, Nucl. Phys. **B70**, 39 (1974).
- [446] H.E. Haber and G.L. Kane, Phys. Rept. **117**, 75 (1985).
- [447] H. Baer and X. Tata, *Weak Scale Supersymmetry: From Superfields to Scattering Events* (Cambridge University Press, United Kingdom, 2006).
- [448] S. Bethke, J. Phys. **G26**, R27 (2000).
- [449] G. ’t Hooft, Nucl. Phys. **B61**, 455 (1973).
- [450] W.A. Bardeen, A. Buras, D. Duke, and T. Muta, Phys. Rev. **D18**, 3998 (1978).
- [451] D.I. Kazakov, in *Lectures given at the European School for High Energy Physics* (Caramulo, Portugal, 2000), available at arXiv:hep-ph/0012288v2.

## BIBLIOGRAPHY

- [452] R.N. Mohapatra, *Unification and Supersymmetry: the Frontiers of Quark-Lepton Physics* (Springer, New York, 1986).
- [453] S. Wolfram, Phys. Lett. **B82**, 65 (1979).
- [454] P.F. Smith and J.R.J. Bennett, Nucl. Phys. **B149**, 525 (1979).
- [455] J. Ellis, J.S. Hagelin, D.V. Nanopoulos, K. Olive, and M. Srednicki, Nucl. Phys. **B238**, 453 (1984).
- [456] T. Falk, K.A. Olive, and M. Srednicki, Phys. Lett. **B339**, 248 (1994).
- [457] E.A. Baltz and P. Gondolo, Phys. Rev. **D67**, 063503 (2003).
- [458] P. Gondolo *et al.*, JCAP **0407**, 008 (2004).
- [459] Z. Ahmed *et al.*, Science **327**, 1619 (2010), see also arXiv:0912.3592v1.
- [460] R. Gaitskell, V. Mandic, and J. Filippini, DMTools Limit Plot Generator (2011), <http://dmttools.brown.edu/>.
- [461] H. Baer, A. Mustafayev, E.-K. Park, and X. Tata, JCAP **0701**, 017 (2007).
- [462] J. Ellis, K.A. Olive, Y. Santoso, and V.C. Spanos, Phys. Rev. **D71**, 095007 (2005).
- [463] R. Trotta, F. Feroz, M.P. Hobson, L. Roszkowski, and R. Ruiz de Austri, JHEP **0812**, 024 (2008).
- [464] G.L. Kane, C. Kolda, L. Roszkowski, and J.D. Wells, Phys. Rev. **D49**, 6173 (1994).
- [465] E.A. Baltz and P. Gondolo, JHEP **0410**, 052 (2004).
- [466] A. Bottino, N. Fornengo, and S. Scopel, Phys. Rev. **D67**, 063519 (2003).
- [467] A. Bottino, F. Donato, N. Fornengo, and S. Scopel, Phys. Rev. **D68**, 043506 (2003).
- [468] A. Bottino, F. Donato, N. Fornengo, and S. Scopel, Phys. Rev. **D69**, 063501 (2004).

## BIBLIOGRAPHY

- [469] S. Meyers, in *The John Adams Memorial Lecture* (CERN, Geneva, 1990).
- [470] ALEPH Collaboration, Phys. Lett. **B526**, 206 (2002).
- [471] DELPHI Collaboration, DELPHI Note 2002-027-CONF-561 (2002).
- [472] L3 Collaboration, L3 Note 2707 (2001).
- [473] OPAL Collaboration, OPAL Physics Note PN470 (2002).
- [474] LEP Higgs Working Group, Note LHWG/2001-03, CERN-EP/2001-055 (2001), available at arXiv:hep-ex/0107029v4.
- [475] LEP Electroweak Working Group, Note LEPEWWG/2001-01 (2001).
- [476] LEP SUSY Working Group, Note LEPSUSYWG/02-06.2 (2002).
- [477] LEP SUSY Working Group, Note LEPSUSYWG/04-07.1 (2004).
- [478] A. Heister *et al.*, Phys. Lett. **B583**, 247 (2004).
- [479] J. Ellis, K. Enqvist, D.V. Nanopoulos, and K. Tamvakis, Phys. Lett. **B155**, 381 (1985).
- [480] G. Anderson, H. Baer, C. Chen, and X. Tata, Phys. Rev. **D61**, 095005 (2000).
- [481] M. Drees, Phys. Lett. **B158**, 409 (1985).
- [482] M. Drees, Phys. Rev. **D33**, 1468 (1986).
- [483] N. Fornengo, S. Scopel, and A. Bottino, Phys. Rev. **D83**, 015001 (2011).
- [484] G. Belanger, F. Boudjema, A. Cottrant, A. Pukhov, and S. Rosier-Lees, JHEP **0403**, 012 (2004).
- [485] L. Calibbi, T. Ota, and Y. Takanishi, JHEP **1107**, 013 (2011).
- [486] CMS Collaboration, Phys. Rev. Lett. **106**, 231801 (2011).
- [487] D. Hooper and T. Plehn, Phys. Lett. **B562**, 18 (2003).
- [488] H.K. Dreiner *et al.*, Eur. Phys. J. **C62**, 547 (2009).

## BIBLIOGRAPHY

- [489] J.A. Conley, H.K. Dreiner, and P. Wienemann, Phys. Rev. **D83**, 055018 (2011).
- [490] D.A. Vasquez, G. Belanger, C. Boehm, A. Pukhov, and J. Silk, Phys. Rev. **D82**, 115027 (2010).
- [491] J.F. Gunion, D. Hooper, and B. McElrath, Phys. Rev. **D73**, 015011 (2006).
- [492] M. Beech, *The Large Hadron Collider: Unravelling the mysteries of the universe* (Springer, New York, 2010).
- [493] S. Agostinelli *et al.*, Nucl. Instr. Meth. **A55**, 671 (2008).
- [494] CMS Collaboration, Phys. Lett. **B698**, 196 (2011).
- [495] CMS Collaboration, CMS Physics Analysis Summary **SUS-09-001**, (2009).
- [496] CMS Collaboration, CMS Physics Analysis Summary **SUS-08-005**, (2008).
- [497] D0 Collaboration, Phys. Lett. **B660**, 449 (2008).
- [498] D0 Collaboration, Phys. Lett. **B680**, 34 (2009).
- [499] CDF Collaboration, Phys. Rev. Lett. **102**, 121801 (2009).
- [500] D. Milstead, Nucl. Phys. **B210–211**, 289 (2011).
- [501] N. Bhattacharyya, A. Choudhury, and A. Datta, arXiv:1107.1997v3 (2011).
- [502] S. Choi, S. Scopel, N. Fornengo, and A. Bottino, arXiv:1108.2190v1 (2011).
- [503] S. Rudaz and F.W. Stecker, Astrophys. J. **368**, 406 (1991).
- [504] J. Silk, K. Olive, and M. Srednicki, Phys. Rev. Lett. **55**, 257 (1985).
- [505] K. Freese, Phys. Lett. **B167**, 295 (1986).



## BIBLIOGRAPHY

- [506] I.V. Moskalenko and A.W. Strong, *Astrophys. J.* **493**, 694 (1998).
- [507] P. Picozza *et al.*, *Astropart. Phys.* **27**, 296 (2007).
- [508] T.G. Guzik *et al.*, *Adv. Space Res.* **33**, 1763 (2004).
- [509] W.B. Atwood *et al.*, *Astrophys. J.* **697**, 1071 (2009).
- [510] O. Adriani *et al.*, *Nucl. Phys.* **B188**, 296 (2009).
- [511] D. Müller and K.K. Tang, *Astrophys. J.* **312**, 183 (1987).
- [512] R.L. Golden *et al.*, *Astrophys. J.* **436**, 769 (1994).
- [513] S.W. Barwick *et al.*, *Astrophys. J.* **482**, L191 (1997).
- [514] M. Boezio *et al.*, *Astrophys. J.* **532**, 653 (2000).
- [515] J. Alcaraz *et al.*, *Phys. Lett.* **B484**, 10 (2000).
- [516] H. Gast, J. Olzem, and S. Schael, in *XL1st Rencontres de Moriond on Electroweak Interactions and Unified Theories* (La Thuile, Italy, 2006), p. 421, available at arXiv:astro-ph/0605254.
- [517] J.J. Beatty *et al.*, *Phys. Rev. Lett.* **93**, 241102 (2004).
- [518] J. Clem and P. Evenson, in *30th International Cosmic Ray Conference* (Merida, Mexico, 2007).
- [519] O. Adriani *et al.*, *Nature* **458**, 607 (2009).
- [520] D. Hooper, A. Stebbins, and K.M. Zurek, *Phys. Rev.* **D79**, 103513 (2009).
- [521] M. Casolino *et al.*, arXiv:0810.4980v1 (2008), to appear in *J. Phys. Soc. Jpn.*
- [522] A. Yamamoto *et al.*, *Nucl. Phys.* **B166**, 62 (2007).
- [523] Y. Asaoka *et al.*, *Phys. Rev. Lett.* **88**, 051101 (2002).
- [524] G.L. Kane, L.-T. Wang, and J.D. Wells, *Phys. Rev.* **D65**, 057701 (2002).

## BIBLIOGRAPHY

- [525] E.A. Baltz, J. Edsjö, K. Freese, and P. Gondolo, Phys. Rev. **D65**, 063511 (2002).
- [526] L. Bergström, J. Edsjö, P. Gondolo, and P. Ullio, Phys. Rev. **D59**, 043506 (1999).
- [527] C. Grimani, Astron. Astrophys. **418**, 649 (2004).
- [528] X. Chi, K.S. Cheng, and E.C.M. Young, Astrophys. J. **459**, L83 (1996).
- [529] I. Büsching, O.C. de Jager, M.S. Potgieter, and C. Venter, Astrophys. J. **678**, L39 (2008).
- [530] O. Ganel *et al.*, Nucl. Instr. Meth. **A552**, 409 (2005).
- [531] J. Chang *et al.*, Nature **456**, 362 (2008).
- [532] J. Nishimura *et al.*, Astrophys. J. **238**, 394 (1980).
- [533] M.A. DuVernois *et al.*, Astrophys. J. **559**, 296 (2001).
- [534] S. Torii *et al.*, Astrophys. J. **559**, 973 (2001).
- [535] S. Torii *et al.*, Adv. Polar Upper Atmos. Res. **20**, 52 (2006).
- [536] H.C. Cheng, J.L. Feng, and K.T. Matchev, Phys. Rev. Lett. **89**, 211301 (2002).
- [537] A.W. Strong and I.V. Moskalenko, Adv. Space Res. **27**, 717 (2001).
- [538] S. Torii *et al.*, Nucl. Instr. Meth. **A452**, 81 (2000).
- [539] A. Kadokura *et al.*, Adv. Polar Upper Atmos. Res. **16**, 157 (2002).
- [540] A.A. Abdo *et al.*, Phys. Rev. Lett. **103**, 251101 (2009).
- [541] A.W. Strong, I.V. Moskalenko, and O. Reimer, Astrophys. J. **613**, 962 (2004).
- [542] F. Aharonian *et al.*, Astron. Astrophys. **508**, 561 (2009).
- [543] A.A. Abdo *et al.*, Phys. Rev. Lett. **102**, 181101 (2009).

## BIBLIOGRAPHY

- [544] M. Aguilar *et al.*, Phys. Rept. **366**, 331 (2002).
- [545] S. Torii *et al.*, arXiv:0809.0760v1 (2008), submitted to Astropart. Phys.
- [546] F. Aharonian *et al.*, Phys. Rev. Lett. **101**, 261104 (2008).
- [547] R.N. Manchester, G.B. Hobbs, A. Teoh, and M. Hobbs, Astron. J. **129**, 1993 (2005).
- [548] D. Grasso *et al.*, Astropart. Phys. **32**, 140 (2009).
- [549] W. Benbow *et al.*, in *High Energy Gamma-Ray Astronomy*, edited by F.A. Aharonian, H.J. Völk, and D. Horns (AIP Proc., Melville, New York, 2005), Vol. 745, p. 661.
- [550] J. Buckley, arXiv:0810.0444v1 (2008).
- [551] H.J. Voelk and K. Bernloehr, Exp. Astron. **25**, 173 (2008).
- [552] P.A. Caraveo, G. R. Mignani, and L.G. Taff, Astrophys. J. **461**, L91 (1996).
- [553] D. Hooper, P. Blasi, and P.D. Serpico, JCAP **0901**, 025 (2009).
- [554] A.A. Abdo *et al.*, Science **325**, 848 (2009).
- [555] D.J. Thompson *et al.*, Astrophys. J. Suppl. **86**, 629 (1993).
- [556] S.D. Hunter *et al.*, Astrophys. J. **481**, 205 (1997).
- [557] V.L. Ginzburg and S.I. Syrovatskii, *The Origin of Cosmic Rays* (Macmillan, New York, 1964).
- [558] F.W. Stecker, NASA Special Publication **249**, (1971).
- [559] W. de Boer, C. Sander, V. Zhukov, A.V. Gladyshev, and D.I. Kazakov, Astron. Astrophys. **444**, 51 (2005).
- [560] L. Bergström, J. Edsjo, M. Gustafsson, and P. Salati, JCAP **0605**, 006 (2006).
- [561] I. Gebauer, arXiv:0710.4966v2 (2007), submitted for the SUSY07 proceedings.

## BIBLIOGRAPHY

- [562] W. de Boer, I. Gebauer, C. Sander, M. Weber, and V. Zhukov, in *7th UCLA Symposium on Sources and Detection of Dark Matter and Dark Energy in the Universe* (AIP Conf. Proc., Marina del Rey, California, 2007), Vol. 903, p. 607.
- [563] I.V.Moskalenko *et al.*, Nucl. Phys. Proc. Suppl. **B173**, 44 (2007).
- [564] F.W. Stecker, S.D. Hunter, and D.A. Kniffen, Astropart. Phys. **29**, 25 (2008).
- [565] E.G. Berezhko and H.J. Völk, Astrophys. J. **611**, 12 (2004).
- [566] G. Zaharijas, A. Cuoco, Z. Yang, and J. Conrad, arXiv:1012.0588v1 (2010), contribution to IDM 2010 conference proceedings.
- [567] A.A. Abdo *et al.*, Astrophys. J. **712**, 147 (2010).
- [568] P. Scott *et al.*, JCAP **01**, 31 (2010).
- [569] M. Ackermann *et al.*, JCAP **05**, 025 (2010).
- [570] L. Dugger, T.E. Jeltema, and S. Profumo, JCAP **12**, 015 (2010).
- [571] A.A. Abdo *et al.*, Phys. Rev. Lett. **104**, 101101 (2010).
- [572] P. Sreekumar *et al.*, Astrophys. J. **494**, 523 (1998).
- [573] A.W. Strong, I.V. Moskalenko, and O. Reimer, Astrophys. J. **613**, 956 (2004).
- [574] A.A. Abdo *et al.*, arXiv:1002.4415v1 (2010), submitted to JCAP.
- [575] F. Calore, V. De Romeri, and F. Donato, arXiv:1105.4230v2 (2011), submitted to Phys. Rev. D.
- [576] M.R. Buckley and D. Hooper, arXiv:1004.1644v2 (2010).
- [577] D. Hooper and L. Goodenough, Phys. Lett. **B697**, 412 (2011).
- [578] F. Aharonian *et al.*, Astron. Astrophys. **425**, L13 (2004).
- [579] K. Kosack *et al.*, Astrophys. J. **608**, L97 (2004).

## BIBLIOGRAPHY

- [580] L. Goodenough and D. Hooper, arXiv:0910.2998v2 (2009).
- [581] A.A. Abdo *et al.*, *Astrophys. J. Suppl.* **188**, 405 (2010).
- [582] M.W. Goodman and E. Witten, *Phys. Rev.* **D31**, 3059 (1985).
- [583] R. Bernabei *et al.*, *Nucl. Instr. Meth.* **A592**, 297 (2008).
- [584] P.S. Barbeau, J.I. Collar, and O. Tench, *JCAP* **09**, 009 (2007).
- [585] K.S. Krane, *Introductory Nuclear Physics* (John Wiley & Sons, New York, 1988).
- [586] D.N. McKinsey, W.H. Lippincott, J. Nikkel, and W. Rellergert, *Phys. Rev. Lett.* **95**, 111101 (2005).
- [587] G. Angloher *et al.*, arXiv:1109.0702v1 (2011).
- [588] J. Lewin and P.F. Smith, *Astropart. Phys.* **6**, 87 (1996).
- [589] M. Kamionkowski and A. Kinkhabwala, *Phys. Rev.* **D57**, 3256 (1998).
- [590] J. Binney and M. Merrifield, *Galactic Astronomy* (Princeton University Press, Princeton, New Jersey, 1998).
- [591] G. Angloher *et al.*, *Astropart. Phys.* **31**, 270 (2009).
- [592] A. Bottino *et al.*, *Phys. Lett.* **B402**, 113 (1997).
- [593] G. Duda, A. Kemper, and P. Gondolo, *JCAP* **0704**, 012 (2007).
- [594] R.H. Helm, *Phys. Rev.* **D104**, 1466 (1956).
- [595] G. Fricke *et al.*, *Atomic Data and Nuclear Data Tables* **60**, 177 (1995).
- [596] D.S. Akerib *et al.*, *Phys. Rev.* **D82**, 122004 (2010).
- [597] G. Angloher *et al.*, *Astropart. Phys.* **18**, 43 (2002).
- [598] N. Ferrari *et al.*, *J. Phys. Conf. Ser.* **39**, 111 (2006).
- [599] S. Yellin, *Phys. Rev.* **D66**, 032005 (2002).

## BIBLIOGRAPHY

- [600] K. Ruddick, NuMI-L-210 (1996).
- [601] E.V. Bugaev *et al.*, Phys. Rev. **D58**, 054001 (1998).
- [602] J. Monroe and P. Fisher, Phys. Rev. **D76**, 033007 (2007).
- [603] A.J. Anderson, J.M. Conrad, E. Figueroa-Feliciano, K. Scholberg, and J. Spitz, Phys. Rev. **D84**, 013008 (2011).
- [604] E-I Esch, Detector Development for Dark Matter Research, Justus-Liebig University Giessen, 2001, inaugural dissertation.
- [605] G. Heusser, Annu. Rev. Nucl. Part. Sci. **45**, 543 (1995).
- [606] G. Heusser, M. Laubenstein, and H. Neder, Radioactiv. Environm. **8**, 495 (2006).
- [607] M. Laubenstein *et al.*, Appl. Rad. Isot. **61**, 167 (2004).
- [608] I. Barabanov *et al.*, Nucl. Instr. Meth. **B251**, 115 (2006).
- [609] K. Shibata *et al.*, J. Nucl. Sci. Technol. **48**, 1 (2011).
- [610] E. Fermi and E. Teller, Phys. Rev. **72**, 399 (1947).
- [611] J.S. Cohen, Phys. Rev. **A65**, 052714 (2002).
- [612] H. Primakoff, Rev. Mod. Phys. **31**, 802 (1959).
- [613] T. Suzuki, D.F. Measday, and J.P. Roalsvig, Phys. Rev. **C35**, 2212 (1987).
- [614] B. Pritychenko, E. Běták, M.A. Kellett, B. Singh, and J. Totans, Nucl. Instr. Meth. **A640**, 213 (2011).
- [615] E.F. Shores, Nucl. Instr. Meth. **B179**, 78 (2001).
- [616] A.J. Koning and D. Rochman, TENDL-2009: Consistent TALYS-based Evaluated Nuclear Data Library Including Covariance Data, [www.talys.eu/tendl-2009/](http://www.talys.eu/tendl-2009/), 2009.
- [617] D. Mei and A. Hime, Phys. Rev. **D73**, 053004 (2006).

## BIBLIOGRAPHY

- [618] H.V. Klapdor-Kleingrothaus *et al.*, Astropart. Phys. **18**, 525 (2003).
- [619] V.A. Bednyakov and H.V. Klapdor-Kleingrothaus, Phys. Part. Nucl. **40**, 583 (2009).
- [620] W.J. Bolte *et al.*, Nucl. Instr. Meth. **A577**, 569 (2007).
- [621] M. Barnabé-Heider *et al.*, Nucl. Instr. Meth. **A555**, 184 (2005).
- [622] M. Felizardo *et al.*, Phys. Rev. Lett. **105**, 211301 (2010).
- [623] D.S. Akerib *et al.*, Phys. Rev. **D72**, 052009 (2005).
- [624] H. Lippincott, Searching for Dark Matter with COUPP, TAUP in Munich, Germany, 2011, conference presentation.
- [625] A. Brown, S. Henry, H. Kraus, and C. McCabe, arXiv:1109.2589v1 (2011).
- [626] A. Broniatowski *et al.*, Phys. Lett. **B681**, 305 (2009).
- [627] E. Armengaud *et al.*, Phys. Lett. **B702**, 329 (2011).
- [628] P. Benetti *et al.*, Astropart. Phys. **28**, 495 (2008).
- [629] E. Aprile *et al.*, arXiv:1107.2155v1 (2011).
- [630] E. Aprile *et al.*, arXiv:1104.2549v3 (2011).
- [631] D. Yu. Akimov *et al.*, Astropart. Phys. **27**, 46 (2007).
- [632] V.N. Lebedenko *et al.*, Phys. Rev. **D80**, 052010 (2009).
- [633] H.S. Lee *et al.*, Nucl. Instr. Meth. **A571**, 644 (2007).
- [634] H.S. Lee *et al.*, Phys. Rev. Lett. **99**, 091301 (2007).
- [635] M.-C. Piro *et al.*, arXiv:1005.5455v1 (2010).
- [636] M. Felizardo *et al.*, arXiv:1106.3014v2 (2011).
- [637] Z. Ahmed *et al.*, Phys. Rev. Lett. **102**, 011301 (2009).

## BIBLIOGRAPHY

- [638] J. Angle *et al.*, Phys. Rev. Lett. **100**, 021303 (2008).
- [639] J. Angle *et al.*, Phys. Rev. Lett. **101**, 091301 (2008).
- [640] V.N. Lebedenko *et al.*, Phys. Rev. Lett. **103**, 151302 (2009).
- [641] C.J. Copi, J. Heo, and L.M. Krauss, Phys. Lett. **B461**, 43 (1999).
- [642] G. Sciolla, Mod. Phys. Lett. **A24**, 1793 (2009).
- [643] P. Belli *et al.*, Phys. Lett. **B387**, 222 (1996).
- [644] A. Incicchitti *et al.*, SIF Conf. Proc. **57**, 147 (1997).
- [645] R. Bernabei *et al.*, Nucl. Phys. **A705**, 29 (2002).
- [646] R. Bernabei *et al.*, Phys. Lett. **B546**, 23 (2002).
- [647] R. Bernabei *et al.*, Phys. Lett. **B408**, 439 (1997).
- [648] R. Bernabei *et al.*, Ukr. J. Phys. **51**, 1037 (2006).
- [649] R. Bernabei *et al.*, Phys. Lett. **B389**, 757 (1996).
- [650] R. Bernabei *et al.*, Phys. Lett. **B424**, 195 (1998).
- [651] M. Moszyński *et al.*, Nucl. Instr. Meth. **A486**, 13 (2002).
- [652] D. Hooper, J.I. Collar, J. Hall, and D. McKinsey, Phys. Rev. **D82**, 123509 (2010).
- [653] R. Bernabei *et al.*, Phys. Lett. **B450**, 448 (1999).
- [654] P. Belli *et al.*, Phys. Rev. **D66**, 043503 (2002).
- [655] R. Bernabei *et al.*, Eur. Phys. J. **C53**, 205 (2007).
- [656] R. Bernabei *et al.*, Int. J. Mod. Phys. **D13**, 2127 (2004).
- [657] R. Bernabei *et al.*, Eur. Phys. J. **C67**, 39 (2010).
- [658] R. Bernabei *et al.*, Eur. Phys. J. **C56**, 333 (2008).



## BIBLIOGRAPHY

- [659] C. Savage, G. Gelmini, P. Gondolo, and K. Freese, *Phys. Rev.* **D83**, 055002 (2010).
- [660] C. Savage, G. Gelmini, P. Gondolo, and K. Freese, *JCAP* **0904**, 010 (2009).
- [661] C. Savage, K. Freese, P. Gondolo, and D. Spolyar, *JCAP* **0909**, 036 (2009).
- [662] P. Belli *et al.*, arXiv:1106.4667v2 (2011).
- [663] V.I. Tretyak, *Astropart. Phys.* **33**, 40 (2010).
- [664] R. Bernabei *et al.*, *Phys. Rev.* **D77**, 023506 (2008).
- [665] P.J. Fox and E. Poppitz, *Phys. Rev.* **D79**, 083528 (2009).
- [666] R. Bernabei *et al.*, *Eur. Phys. J.* **C23**, 61 (2002).
- [667] Z. Ahmed *et al.*, *Phys. Rev.* **D83**, 112002 (2011).
- [668] C.E. Aalseth *et al.*, *Phys. Rev. Lett.* **101**, 251301 (2008).
- [669] K.W. Jones and H.W. Kraner, *Phys. Rev.* **A11**, 1347 (1975).
- [670] C.E. Aalseth *et al.*, *Phys. Rev. Lett.* **106**, 131301 (2010).
- [671] A.L. Fitzpatrick, D. Hooper, and K.M. Zurek, arXiv:1003.0014v2 (2010).
- [672] J. Kopp, T. Schwetz, and J. Zupan, *JCAP* **1002**, 014 (2010).
- [673] S. Andreas, C. Arina, T. Hambye, F.-S. Ling, and M.H.G. Tytgat, arXiv:1003.2595v3 (2010).
- [674] S. Chang, J. Liu, A. Pierce, N. Weiner, and I. Yavin, arXiv:1004.0697v2 (2010).
- [675] R. Foot, *Phys. Lett.* **B692**, 65 (2010).
- [676] R. Essig, J. Kaplan, P. Schuster, and N. Toro, arXiv:1004.0691v2 (2010).
- [677] H. Sekiya *et al.*, *Phys. Lett.* **B571**, 132 (2003).

## BIBLIOGRAPHY

- [678] E.M. Drobyshevski, *Mod. Phys. Lett.* **A23**, 3077 (2008).
- [679] N. Bozorgnia, G.B. Gelmini, and P. Gondolo, *JCAP* **1011**, 019 (2010).
- [680] N. Bozorgnia, G.B. Gelmini, and P. Gondolo, *JCAP* **1011**, 028 (2010).
- [681] N. Wang, Ph.D. thesis, University of California Berkeley, 1991.
- [682] T.A. Shutt, Ph.D. thesis, University of California Berkeley, 1993.
- [683] W.K. Stockwell, Ph.D. thesis, University of California Berkeley, 1996.
- [684] A. Sonnenschein, Ph.D. thesis, University of California Santa Barbara, 1999.
- [685] S. Golwala, Ph.D. thesis, University of California Berkeley, 2000.
- [686] N. Wang *et al.*, in *IV International Conference of Low Temperature Dark Matter and Neutrino Detectors*, edited by N.E. Booth and G.L. Salmon (Editions Frontieres, Oxford, UK, 1992), p. 109.
- [687] D. Abrams *et al.*, *Phys. Rev.* **D66**, 122003 (2002).
- [688] K.D. Irwin, S.W. Nam, B. Cabrera, B. Chugg, and B.A. Young, *Rev. Sci. Instrum.* **66**, 5322 (1995).
- [689] R.M. Clarke *et al.*, in *Proceedings of the Second International Workshop on the Identification of Dark Matter*, edited by N.J.C. Spooner and V. Kudryavtsev (World Scientific, Singapore, 1999), pp. 353–358.
- [690] T. Saab *et al.*, in *Ninth International Workshop on Low Temperature Detectors* (AIP, New York, 2002), Vol. 605, pp. 497–500.
- [691] S.W. Nam, Ph.D. thesis, Stanford University, 1998.
- [692] A.J. Da Silva, Ph.D. thesis, University of British Columbia, 1996.
- [693] T. Perera, Ph.D. thesis, Case Western Reserve University, 2001.
- [694] D.S. Akerib *et al.*, *Phys. Rev.* **D68**, 082002 (2003).
- [695] T. Saab, Ph.D. thesis, Stanford University, 2002.

## BIBLIOGRAPHY

- [696] D.D. Driscoll, Ph.D. thesis, Case Western Reserve University, 2004.
- [697] R. Abusaidi *et al.*, Phys. Rev. Lett. **84**, 5699 (2000).
- [698] D.S. Akerib *et al.*, Phys. Rev. Lett. **93**, 211301 (2004).
- [699] V. Mandic, Ph.D. thesis, University of California Berkeley, 2004.
- [700] C.L. Chang, Ph.D. thesis, Stanford University, 2004.
- [701] G. Wang, Ph.D. thesis, Case Western Reserve University, 2005.
- [702] A.J. Reisetter, Ph.D. thesis, University of Minnesota, 2005.
- [703] J. Sander, Ph.D. thesis, University of California Santa Barbara, 2007.
- [704] R.W. Ogburn, Ph.D. thesis, Stanford University, 2008.
- [705] D.S. Akerib *et al.*, Phys. Rev. Lett. **96**, 011302 (2006).
- [706] D.S. Akerib *et al.*, Phys Rev. **D73**, 011102 (2006).
- [707] R. Hennings-Yeomans, Ph.D. thesis, Case Western Reserve University, 2008.
- [708] J.P. Filippini, Ph.D. thesis, University of California Berkeley, 2008.
- [709] C.N. Bailey, Ph.D. thesis, Case Western Reserve University, 2009.
- [710] T. Bruch, Ph.D. thesis, University of Zurich, 2010.
- [711] M.C. Fritts, Ph.D. thesis, University of Minnesota, 2011.
- [712] Z. Ahmed, Ph.D. thesis, California Institute of Technology, 2011.
- [713] Y.P. Varshni, Physica **34**, 149 (1967).
- [714] C.D. Thurmond, J. Electrochem. Soc. **122**, 1133 (1975).
- [715] C. Kittel, *Introduction to Solid State Physics*, 6th ed. (John Wiley, New York, 1986).
- [716] X.-F. Navick *et al.*, Nucl. Instr. Meth. **A444**, 361 (2000).

## BIBLIOGRAPHY

- [717] B.A. Young and K.M. Yu, Rev. Sci. Instrum. **66**, 2625 (1995).
- [718] C.A. Klein, IEEE Trans. Nucl. Sci. **15**, 214 (1968).
- [719] J. Lindhard *et al.*, K. Dan. Vidensk. Selsk., Mat.-Fys. Medd. **33**, No.10 (1963).
- [720] J. Lindhard *et al.*, K. Dan. Vidensk. Selsk., Mat.-Fys. Medd. **36**, No.10 (1968).
- [721] C.A. Klein, J. Appl. Phys. **39**, 2029 (1968).
- [722] R. Sarpeshkar, T. Delbruck, and C.A. Mead, IEEE Circuits Devices Mag. **9**, 23 (1993).
- [723] B. Neganov and V. Trofimov, Otkrytye. Izobreteniya. **146**, 215 (1985).
- [724] P.N. Luke, J. Appl. Phys. **64**, 6858 (1988).
- [725] T. Shutt *et al.*, Nucl. Instr. Meth. **A444**, 340 (2000).
- [726] B. Cabrera, Nucl. Instr. Meth. **A444**, 304 (2000).
- [727] T. Saab, R.M. Clarke, B. Cabrera, R.A. Abusaidi, and R. Gaitskell, Nucl. Instr. Meth. **A444**, 300 (2000).
- [728] B.A. Young, B. Cabrera, and A.T. Lee, Phys. Rev. Lett. **64**, 2795 (1990).
- [729] R. Clarke, Ph.D. thesis, Stanford University, 1999.
- [730] J.F. Cochran and D.E. Mapother, Phys. Rev. **111**, 132 (1958).
- [731] L.N. Cooper, Phys. Rev. **104**, 1189 (1956).
- [732] K.D. Irwin *et al.*, IEEE Trans. Appl. Supercond. **5**, 2690 (1995).
- [733] K.D. Irwin, Ph.D. thesis, Stanford University, 1995.
- [734] S. Basavaiah and S.R. Pollack, J. Appl. Phys. **39**, 5548 (1968).
- [735] B.A. Young, T. Saab, B. Cabrera, J.J. Cross, and R.A. Abusaidi, J. Appl. Phys. **86**, 6975 (1999).

## BIBLIOGRAPHY

- [736] B.A. Young, T. Saab, B. Cabrera, A.J. Miller, P.L. Brink, and J.P. Castle, J. Appl. Phys. **91**, 6516 (2002).
- [737] J. Johnson, Phys. Rev. **32**, 97 (1928).
- [738] H. Nyquist, Phys. Rev. **32**, 110 (1928).
- [739] R. Mancini *et al.*, Op Amps for Everyone, Texas Instruments Design Reference, 2002, application note SLOD006B.
- [740] R.P. Welty and J.M. Martinis, IEEE Trans. Appl. Supercond. **3**, 2605 (1993).
- [741] M.E. Huber *et al.*, IEEE Trans. Appl. Supercond. **11**, 4048 (2001).
- [742] C. Chasman, K.W. Jones, and R.A. Ristinen, Phys. Rev. Lett. **15**, 245 (1965).
- [743] C. Chasman *et al.*, Phys. Rev. **154**, 239 (1967).
- [744] C. Chasman *et al.*, Phys. Rev. Lett. **21**, 1430 (1968).
- [745] A.R. Sattler, F.L. Vook, and J.M. Palms, Phys. Rev. **143**, 588 (1966).
- [746] K.W. Jones and H.W. Kraner, Phys. Rev. **C4**, 125 (1971).
- [747] Y. Messous *et al.*, Astropart. Phys. **3**, 361 (1995).
- [748] L. Baudis *et al.*, Nucl. Instr. Meth. **A418**, 348 (1998).
- [749] E. Simon *et al.*, Nucl. Instr. Meth. **A507**, 643 (2003).
- [750] S.T. Lin *et al.*, Phys. Rev. **D79**, 061101 (2009), data in Figure 4.8 from arXiv:0712.1645v4.
- [751] P.S. Barbeau, Ph.D. thesis, University of Chicago, 2009.
- [752] A.R. Sattler, Phys. Rev. **138**, A1815 (1965).
- [753] P. Zecher *et al.*, Phys. Rev. **A41**, 4085 (1990).
- [754] G. Gerbier *et al.*, Phys. Rev. **D42**, 3211 (1990).

## BIBLIOGRAPHY

- [755] B.L. Dougherty, Phys. Rev. **A45**, 2104 (1992).
- [756] D. Moore, private communication, 2011.
- [757] Z. Ahmed *et al.*, Phys. Rev. Lett. **106**, 131302 (2011).
- [758] P.D. Barnes, Ph.D. thesis, University of California Berkeley, 1996.
- [759] A. Benoit *et al.*, Phys. Lett. **B513**, 15 (2001).
- [760] D.F. Bartlett, A.L. Duncan, and J.R. Elliott, Rev. Sci. Instrum. **52**, 265 (1981).
- [761] International Atomic Energy Agency, Database of Prompt Gamma Rays from Slow Neutron Capture for Elemental Analysis, Vienna, 2006, available at <http://www-nds.iaea.org/pgaa/>.
- [762] N. Craig and T. Lester, Hitchhiker's Guide to the Dilution Refrigerator, Marcus Lab, Harvard University, 2004, version 2.00.
- [763] X.-F. Navick *et al.*, Nucl. Instr. Meth. **A520**, 189 (2004).
- [764] D.S. Akerib *et al.*, Nucl. Instr. Meth. **A520**, 163 (2004).
- [765] U. Fano, Phys. Rev. **72**, 26 (1947).
- [766] R. Brun and F. Carminati, CERN Program Library Long Writeup **W5013**, (1993).
- [767] S. Kamat, Ph.D. thesis, Case Western Reserve University, 2005.
- [768] S. Yellin, arXiv:0709.2701v1 (2007).
- [769] S. Yellin, arXiv:1105.2928v2 (2011).
- [770] M.B. Chadwick *et al.*, Nucl. Data Sheets **107**, 2931 (2006).
- [771] S. Fallows, private communication, 2011.