A DARK-MATTER SEARCH USING THE FINAL CDMS II DATASET AND A NOVEL DETECTOR OF SURFACE RADIOCONTAMINATION

Thesis by
Zeeshan Ahmed

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

2012
(Defended July 14, 2011)
To my family
Acknowledgments

These past six years have been the most instructional, challenging, and fun-filled years of my life. I owe thanks to many people for teaching and supporting me along the way. As much as I’d like to recall and write down every name from every professional and personal interaction I’ve had, every CDMS meeting, insightful discussion, and inspirational pep talk since 2005, my memory will fail me, and someone will be missed — for this, I apologize; you have all contributed to an excellent educational journey.

I would like to first thank my advisor, Prof. Sunil Golwala. Before I started at Caltech, I was told that the relationship between a Ph.D. advisor and a graduate student was like a marriage; it required not only matching (research) interests, but the ability to set reasonable expectations of each other, communicate well, and be patient with each other through ups and downs. I was lucky to have Sunil offer me space in his group as soon as I joined Caltech, and I dare say the match couldn’t have been better. Sunil’s brilliance, rational thinking, systematic problem solving, work ethic, and exceptional ability to multitask and refocus at will inspired me, and challenged me to adopt and hone some of these skills. I learned from Sunil how to really deconstruct a problem and get to core interesting scientific questions. He hammered into me how to think about noise and optimal filters, something that, in turn, the graduate students I now work with at Stanford have to suffer through! Through all of this, he gave me a lot of freedom and was extremely patient with me, as I stumbled along and learned. Oh, and thanks for writing that awesome limerick about me, Sunil!

Next, I owe a great deal to Jeff Filippini, Matt Pyle and Bruno Serfass. The decision, early in my graduate career, to regularly visit UC Berkeley paid off well, for I picked up a great set of mentors and friends. These three taught me how to build ZIPs, test them and analyze data from them. Jeff in particular, upon moving to Caltech as a postdoctoral fellow, was virtually a live-in consultant to the analysis team for the CDMS II final dataset (Runs 125–128). Matt and Bruno taught me how to calibrate phonon-pulse shapes, mostly so that they could pawn it off on me and get off the hook for Runs 125–128! But, I appreciate it nonetheless. It provided me with the opportunity to make significant contributions to that analysis. Also, thanks to them, I am acquainted with well over 20 different coffee shops in San Francisco, our preferred workspace for turning caffeine and Wi-Fi internet into limits on WIMP-nucleon interaction cross sections.
The analysis of Runs 125–128 was a mammoth task. Unlike the spin-off low-energy threshold analysis and exotic dark matter searches, everything for the primary analysis starts from scratch — from first-tier data processing, to selecting good datasets, calibrating the full complement of detectors over several cryogenic runs, setting physics cuts and battling surface events. We had an excellent team to do this. It was a pleasure coordinating this analysis with Prof. Jodi Cooley-Sekula, who reassured me at every step that I would get to sleep some day. Kevin McCarthy and Scott Hertel were excellent and eager students who spent an entire month at Caltech, supporting the phonon position and energy calibration. Lauren Hsu was instrumental in processing all the data, playing devil’s advocate, and triple-checking that every ‘i’ was dotted and every ‘t’ crossed. Tobi Bruch and Seb Arrenberg handled complicated cuts and calculations by coordinating with me from Zürich over Skype, often in the wee hours of the morning. Matt Fritts and Oleg Kamaev were critical to furthering our understanding of surface event leakage and designing new rejection cuts. Thanks are also due to Scott Fallows, Joseph Kiveni, Mark Kos and Jeter Hall for taking on important studies during the analysis.

There’s one name I didn’t include in the last paragraph, and that’s because Dave Moore gets his own paragraph. It has been my honor and pleasure to work with and share an office with Dave over the past few years. Dave is unbelievably brilliant, ridiculously fast and humble to boot. He makes physics look effortless. Dave was a fount of ideas during the analysis of Runs 125–128, and had a hand in every major study for that analysis, including the final surface-event leakage calculation. He also subsequently mastered the low-energy threshold analysis, disproved some interpretations of the recent CoGeNT dark matter detection [1], and invented MKID-based dark matter detectors. Dave has been an excellent sounding board for my ideas and research problems, and I have learned all manner of amazing analysis tricks from him. Ask him how to compare efficiently lists of Series and Event numbers in MATLAB; your jaw will drop to the floor — I kid you not.

Over the past few years, I’ve had the honor of working with some very intelligent and accomplished scientists in CDMS — Blas Cabrera, Bernard Sadoulet, Dan Bauer, Richard Schnee, Steve Yellin, Prisca Cushman, Rupak Mahapatra, Tarek Saab, Tali Figueroa, Walt Ogburn, and Kyle Sundqvist, to name a few. I have always gained great insights from conversations with all of you. A special thanks to Dennis Seitz for teaching Sunil how to solder “correctly”, so that he could then enslave me till I wired up the BetaCage “correctly”. Jim Beaty, I don’t know how CDMS would work without you and all the gizmos you have invented to fix little problems and avert catastrophes. That and Scotch-Brite.

Working on the BetaCage was challenging and frustrating at first. The project presented a tractable problem, but we underestimated the number of technical quagmires along the way, and were initially short-staffed. But, we overcame those issues. We really picked up momentum when John Hanson and, eventually, Bob Nelson joined us. I thank both of them for their insights and con-
tributions to the design of the prototype BetaCage. I especially thank Bob for providing leadership in construction and testing of the prototype as I left Caltech.

During my time at Caltech, we lost Daphne Chang and Andrew Lange, both stellar physicists, at very different stages in their careers — one with a bright future cut short, and one with a glorious past of discovery that only seemed like the beginning. You are role models for me, and I miss you both. We'll keep fighting the good fight for you.

Special thanks also go to our administrative assistants and my friends Kathy Deniston and Barbara Wertz. I can’t count how many times they’ve saved me from administrative and purchasing debacles. They’ve given me pep talks when nothing seemed to work. Quite frequently, if I were working too late, Barbara would stop by and drop off some candy for me to snack on, before she left for the day.

I conclude my professional thanks by blaming Prof. Gene Bickers, my undergraduate advisor and Vice Provost at USC, for tricking me into liking physics so much that I gave up a lucrative and sane engineering career in favor of an overworked life, answering the fundamental questions of nature. Not the worst trade off though, right?

Outside the realm of research, the past six years went from being tolerable to being extremely fun because of great friendships I was fortunate to enjoy.

At Caltech, I owe the most to Pinkesh Patel and Dave Doll — Pinkesh for his support of my quest to figure out why the heck I was doing all of this; Dave for reminding me to have fun along the way, and for dragging me to happy hours at Whole Foods. Yes, you read that correctly. Justus Brevik and Randol Aikin were great friends more than coworkers. Mike Siddiqui was such an awesome and fun roommate. You’re welcome for the free parking space, son. Varun Bhalerao and Shriharsh Tendulkar helped me develop a healthy disregard for safety rules, and taught me some astronomy along the way. Kaveh Pahlevan occasionally traded me food for the latest gossip in physics. The incoming physics class of 2005 was a great support group, especially Riccardo Schmidt, Matt Matheny, Yousi Ma, and Nicole Czakon. The Resident Associates of Caltech, and the staffs at Student Affairs, ISP, the Counseling Center and Institute Housing were great teams to work with while serving the student population here. My life at Caltech was also enriched by many active members of the graduate student community, including those involved with the GSC, the Caltech Y, and OASIS.

Outside Caltech, my friends from college and in LA kept assuring me that it was awesome to be a particle astrophysicist, and let me escape to their world whenever I needed to. Thanks are due to Aditya Varma, Nitin Sharma, Nilay Vora, Jessica Marek, Anish Prasad, Kamila Sikora, Rahul Dutta, Kaleb Keller, Morgan Hendry, Kerri Anderson, Anand Ramasswamy, Jenny Lorch, Kimi Porter and Kirstine Odegard. You’re all correct. We are really good at life.

Finally, Dad, Mom, Zahra, Zoha, and Zeba, I’m nowhere without you. Thank you so much. This, truly, is for you.
Abstract

Substantial evidence from galaxies, galaxy clusters, and cosmological scales suggests that \( \sim 85\% \) of the matter of our universe is invisible. The missing matter, or "dark matter," is likely composed of non-relativistic, non-baryonic particles, which have very rare interactions with baryonic matter and with one another. Among dark matter candidates, Weakly Interacting Massive Particles (WIMPs) are particularly well motivated. In the early universe, thermally produced particles with weak-scale mass and interactions would ‘freeze out at the correct density to be dark matter today. Extensions to the Standard Model of particle physics, such as Supersymmetry, which solve gauge hierarchy and coupling unification problems, naturally provide such particles.

Interactions of WIMPs with baryons are expected to be rare, but might be detectable in low-noise detectors. The Cryogenic Dark Matter Search (CDMS) experiment uses ionization- and phonon-sensitive germanium particle detectors to search for such interactions. CDMS detectors are operated at the Soudan Underground Laboratory in Minnesota, within a shielded environment to lower cosmogenic and radioactive background. The combination of phonon and ionization signatures from the detectors provides excellent residual-background rejection.

This dissertation presents improved techniques for phonon calibration of CDMS II detectors and the analysis of the final CDMS II dataset with 612 kg-days of exposure. We set an upper limit of \( 3.8 \times 10^{-44} \text{ cm}^2 \) on WIMP-nucleon spin-independent scattering cross section for a WIMP mass of 70 GeV/c\(^2\). At the time this analysis was published, these data presented the most stringent limits on WIMP scattering for WIMP masses over 42 GeV/c\(^2\), ruling out previously unexplored parameter space.

Next-generation rare-event searches such as SuperCDMS, COUPP, and CLEAN will be limited in sensitivity, unless they achieve stringent control of the surface radioactive contamination on their detectors. Low-penetrating radiation, such as alpha and beta particles, will mimic signal in these experiments. This dissertation also presents the design and prototyping of a novel detector for surface radiocontaminants, called the BetaCage — a neon-gas time projection chamber built from radiopure materials and operated underground with shielding similar to CDMS II. The BetaCage will enable beta screening of materials at world-best sensitivity of \( 10^{-5} / \text{cm}^2/\text{keV/day} \), providing a valuable tool to the physics community.
Contents

Acknowledgments iv

Abstract viii

1 Dark Matter: Evidence and Candidates 1

1.1 Our ΛCDM Universe . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1

1.2 Evidence for Dark Matter ........................................... 3

1.2.1 Galaxy rotation curves ....................................... 3

1.2.2 Galaxy clusters ................................................. 5

1.2.3 Modern cosmology .............................................. 6

1.3 Dark Matter Candidates ........................................... 10

1.3.1 Modified gravity .............................................. 10

1.3.2 Neutrinos ....................................................... 11

1.3.3 Axions ........................................................... 11

1.3.4 Weakly Interacting Massive Particles (WIMPs) ......... 12

1.4 Supersymmetric WIMPs .......................................... 14

1.4.1 Framework and consequences ............................... 14

1.4.2 Minimal Supersymmetric Standard Model ................. 14

1.4.3 Dark matter from CMSSM .................................. 16

2 Detecting WIMPs ..................................................... 18

2.1 Production in Particle Colliders ................................ 18

2.2 Indirect Astrophysical Detection ................................ 19

2.2.1 Gamma rays .................................................... 21

2.2.2 Antimatter ...................................................... 22

2.2.3 Neutrinos ....................................................... 23

2.3 Direct Detection ................................................... 24

2.3.1 Rates and recoil spectra .................................... 25

2.3.2 WIMP-nucleon elastic scattering .......................... 27
6.4.2 Phonon pre-pulse baseline ................................. 107
6.4.3 Charge pre-pulse baseline ................................. 108
6.4.4 Phonon start time ........................................... 109
6.4.5 Phonon manifold distance ................................. 111

6.5 Physics Cuts ...................................................... 111
6.5.1 Scintillator veto activity ................................... 111
6.5.2 Single scattering ............................................. 112
6.5.3 Ionization yield ............................................. 113
6.5.4 Fiducial volume ............................................. 115
6.5.5 Thresholds .................................................. 116

6.6 Surface-Event Rejection ....................................... 118
6.6.1 Calibration data ............................................. 118
6.6.2 Past techniques and lessons ............................... 118
6.6.3 Systematic differences between calibration and WIMP-search surface events ............................... 119
6.6.4 Estimating leakage, accounting for systematic differences ............................................ 121
6.6.5 Cut setting algorithm ....................................... 122
6.6.6 Surface-event rejection cut for Runs 125–128 ................. 125
6.6.7 Other candidate cuts ....................................... 125

6.7 WIMP-search Exposure ......................................... 127
6.7.1 Exposure .................................................. 128
6.7.2 Efficiency .................................................. 128

7 WIMP-search Results ............................................... 130
7.1 Expected Neutron Background ................................. 130
7.1.1 Radiogenic neutrons ....................................... 130
7.1.2 Cosmogenic neutrons ...................................... 131

7.2 Expected Photon Background .................................. 132

7.3 Expected Surface-Event Background ............................ 132
7.3.1 Premise .................................................. 133
7.3.2 Poisson and Bayesian estimates ......................... 134
7.3.3 Method 1: Using WIMP-search multiples in the NR band .... 134
7.3.4 Method 2: Using WIMP-search multiples outside the NR band ......................... 135
7.3.5 Method 3: Using calibration data ....................... 136
7.3.6 Combined estimate ....................................... 137
7.3.7 Systematic errors ....................................... 137

7.4 Unblinding WIMP Candidates .................................. 138
7.5 Understanding the Two Observed Events ................................. 142
	7.5.1 Basic checks ................................................. 142
	7.5.2 Surface-event rejection close to threshold .......................... 142
	7.5.3 Charge start-time reconstruction ................................. 144
	7.5.4 Quantifying background/WIMP likelihood of candidate events ... 147

7.6 Constraints on WIMP-Parameter Space ................................. 150

8 Characterization of Surface Radioccontaminants ......................... 153
	8.1 Identification of Surface-Event Contributors in CDMS II .......... 154
	8.2 Surface-event Reduction Requirements for SuperCDMS ............... 156
	8.3 Surface Radioccontamination Goals ................................ 157
	8.4 Current Assay Techniques and Shortcomings ......................... 157
	8.5 Need for Dedicated Beta- and Alpha-screening ..................... 158

9 BetaCage Design ..................................................................... 160
	9.1 Basic Design and Operational Principles ............................. 160
	9.2 Multi-wire Proportional Counter (MWPC) Units ...................... 162
		9.2.1 Particle detection using MWPCs ............................. 162
		9.2.2 MWPCs in BetaCage ......................................... 162
	9.3 Gas, Handling and Containment ........................................ 165
	9.4 Drift Volume and Field Cage .......................................... 166
	9.5 Voltage and Gain Requirements ...................................... 166
	9.6 Design Tolerances .................................................... 169
		9.6.1 HV supply .................................................... 169
		9.6.2 Wire diameter variation ........................................ 169
		9.6.3 MWPC wire positioning ...................................... 170
		9.6.4 MWPC wire tensioning ....................................... 171
		9.6.5 Gas purity ..................................................... 171
	9.7 Expected Signal in Bulk MWPC ...................................... 172
		9.7.1 Approximate pulse shape ...................................... 172
		9.7.2 Signal attenuation in network ................................... 175
		9.7.3 System optimization based on features of unamplified signal ... 179
	9.8 Readout Electronics .................................................... 180
		9.8.1 Signal amplification ............................................ 180
		9.8.2 Digitization ..................................................... 181
	9.9 Electronics Noise ........................................................ 182
		9.9.1 Thermal noise from network .................................... 182
9.9.2 Amplifier noise .................................................. 183
9.9.3 Total electronic noise .......................................... 184
9.10 Energy Resolution ................................................. 184
9.10.1 Electron-ion pair generation and avalanche statistics . 184
9.10.2 Readout resolution ............................................. 185
9.10.3 Digitization noise .............................................. 186
9.10.4 Summed energy resolution ................................... 187
9.11 Backgrounds and Expected Sensitivity ...................... 187

10 Protocage: a BetaCage prototype ............................... 189
10.1 Thin-frame Layered MWPC Units ............................. 189
10.1.1 UHMWPE ...................................................... 190
10.1.2 Delrin 570 frames ............................................ 192
10.2 Four-part MWPC units .......................................... 196
10.2.1 Design features ................................................ 197
10.2.2 Assembly and construction ................................... 202
10.3 Other ProtoCage Components .................................. 205
10.3.1 Field cage ...................................................... 205
10.3.2 DAQ, electronics and HV .................................... 205
10.3.3 Gas handling .................................................. 206
10.4 Test MWPC with Anode Plane ................................ 206
10.5 Future plans ....................................................... 208

A Phonon-Pulse Shape Correction for Runs 125–128 .......... 209
A.1 First Pass Correction .............................................. 209
A.1.1 Selecting photon samples .................................... 210
A.1.2 Lookup table generation .................................... 210
A.1.3 Phonon-pulse shape correction ............................. 211
A.2 Final Correction ..................................................... 212
A.2.1 Refining the photon sample selection ..................... 212
A.2.2 Lookup table generation ................................... 212
A.2.3 Correction using PipeCleaner .............................. 214

B Detector Development for SuperCDMS ......................... 216
B.1 mZIP .......................................................... 216
B.1.1 Design ........................................................ 216
B.1.2 mZIP testing and characterization ......................... 218
Chapter 1

Dark Matter: Evidence and Candidates

1.1 Our ΛCDM Universe

We live in interesting times in an interesting universe — only over the past decade have we narrowed down fundamental cosmological parameters to few-percent levels. Our observations and inferences show that our universe started and expanded from a singularity, that it is 13.7 billion years old, and its content is the right amount to make it obey Euclidean geometry on large scales. During its history, the universe underwent phases of radiation dominance (before redshift, $z \sim 3000^1$), matter dominance (before $z \sim 0.4$) and is now dominated by energy. The left pane of Figure 1.1 shows a pie chart of the universe’s content budget as it stands today and at $z \sim 1100$, deduced from a variety of cosmological observations. Note that the two largest energy and matter components are labeled ‘dark’. This is because they seem invisible; they have eluded detection via conventional observation. We infer their abundances from various cosmological observations as shown in the right pane of Figure 1.1, but we do not understand the nature of these components. Thus, we have the ΛCDM model of the universe, dominated by dark energy (denoted Λ) and cold dark matter (CDM; cold, for reasons I explain in Section 1.2.3.3). Both components are subjects of active research in physics and cosmology. This dissertation focuses on a search for a class of dark matter candidates called Weakly Interacting Massive Particles (WIMPs). In this first chapter, I build a case for dark matter, explain how we came to deduce its abundance, and show how WIMPs are an attractive candidate to be the dark matter of our universe. I also describe Supersymmetry, a popular particle-physics framework for WIMPs.

---

1 $z$ denotes redshift, or how much longer the wavelength of a photon is than when it was released. $z$ is related to $a$, the scale factor of the universe, by $a = \frac{1}{1+z}$. A time when the universe was $\sim 3000$ times smaller than its present size is denoted by $z \sim 3000$. 

Figure 1.1: Left: Content budget of universe today (top) and at $z \sim 1100$ (bottom). Today, the universe is dominated by dark matter and dark energy, whereas it was dominated by dark matter at $z \sim 1100$, with notable contributions from baryons, photons, and neutrinos. *Pie chart from: WMAP five-year data release* [2]. Right: Abundance of dark energy ($\Omega_\Lambda$) vs. abundance of matter ($\Omega_M$), where the abundances are normalized by the energy density required to make the universe flat. Three datasets — Type Ia supernovae (SNe), the cosmic microwave background (CMB), and baryon acoustic oscillations (BAO) are shown with 1σ, 2σ, and 3σ confidence regions. *Plot from: [3]*.
1.2 Evidence for Dark Matter

The evidence for dark matter comes from various types of measurements and observations. I broadly organize these in order of the length scale in the universe over which these observations are made.

1.2.1 Galaxy rotation curves

I start with the galactic scale, the smallest scale over which a coherent and reliable picture of dark matter has emerged through observations of hundreds of spiral and elliptical galaxies. The motion of stars, dwarf satellite galaxies, and other objects in a galaxy enables one to infer the mass distribution present in it, through very simple arguments.

In the case of spiral galaxies, the rotation speed of objects as a function of galactic radius is obtained by measuring the Doppler shift of spectral features such as the electronic and hyperfine spectral lines of neutral and ionized atomic hydrogen, and the rotational transitions of CO, along the line-of-sight. \[4\]. The matter, \(M(r)\), contained within galactic radius \(r\) is related to these measured rotational speeds, \(v(r)\), using a simple Newtonian argument.

\[
v^2(r) = \frac{GM(r)}{r}
\]

(1.1)

where \(G\) is the universal gravitational constant.

If the matter distribution in these galaxies is dominated by visible matter, then the rotational speeds of objects beyond the visible disc of spiral galaxies should fall off as \(\sim 1/\sqrt{r}\). However a wealth of data show that generally between a factor of 1 and 2 of the visible radii of galaxies, the rotation curves are flat, implying \(M(r) \propto r\) \[4\]. This requires the addition of an invisible halo of matter to the mass profiles of galaxies, extending beyond the visible component. Today, this is generally attributed to dark matter. A decomposition of a rotation curve into components attributed to visible matter, gas, and a dark halo is shown in Figure 1.2. Figure 1.3 shows the rotation curves of several spiral galaxies on the same axes.

For elliptical galaxies, rotation curves, and hence the galactic mass as a function of galactic radius, cannot be directly measured, since its constituent objects do not have well-defined directionality. However, there is another trick that enables inference of the masses of elliptical galaxies. Gas falling into the gravitational well of the galaxy, is heated, emitting brehmsstrahlung X-rays. Assuming the galaxy has reached a relaxed state, this temperature is related to the depth of the gravitational well because of equipartition and virial theorem. The velocity dispersion of stars within the galaxy, \(\sigma\), is also related to the galaxy’s mass by virial theorem. In general, \(T \propto \sigma^2\), for a gravitational well provided by the stars alone. It is observed however that \(T \propto \sigma^{1.45}\) \[6\]. This is best explained by dark matter dominating the composition of elliptical galaxies \[7\].
Figure 1.2: Rotational speed vs. galactic radius for NGC 6503. The data are shown with black points. A solid curve is fit through the data by varying the three constituent components — the dashed curve is the visible component, the dotted curve is attributed to gas, and the dashed-dotted curve represents dark matter. Plot from [5].

Figure 1.3: Rotation curves of several spiral galaxies plotted on the same axes. Plot from [4].
1.2.2 Galaxy clusters

Galaxy clusters are the largest gravitationally bound structures in the universe. They consist of of $\sim O(10–1000)$ visible galaxies, a more massive component of hot gas, called intracluster medium, and finally mostly of dark matter. This composition has been inferred using a combination of several different types of observations, outlined below.

1.2.2.1 Velocity dispersion of galaxies in clusters

The earliest argument for missing matter in galaxy clusters was made by Fritz Zwicky in 1933 using observations of the Coma cluster [8], [9]. He measured the velocity dispersion of eight galaxies in that cluster, and deduced their total kinetic energy by assuming an estimated spiral galaxy mass. His estimate of the gravitational potential energy relied on an incorrect value of the Hubble parameter, which inflated the inferred mass-to-light ratio. Even after correcting for this, the mass-to-light ratio remains higher than expected from visible matter alone. More recent studies of clusters, such as the ESO Nearby Abell Cluster Survey reveal similarly high mean mass-to-light ratios of $M/L = 454 h M_\odot/L_\odot^2$ for 29 clusters [10]. This indicates that visible galaxies comprise only a small fraction of galaxy clusters.

1.2.2.2 X-ray emission of intracluster medium

The intracluster medium (ICM), composed of gas, is also a tracer of the gravitational field of clusters. Just as for elliptical galaxies, gas falling into a cluster’s gravitational well is pressure heated to temperatures of $10^7–10^8$ K — temperatures deduced from the observation of X-ray bremsstrahlung emission. This, in turn, enables deduction of mass by virial theorem. Most surveys have shown that the ICM, while much more massive than the stars in a cluster, does not account for a cluster’s remaining mass. A survey of 13 low-redshift clusters from Chandra X-ray Observatory data has been used to determine the average mass fraction of the intracluster medium to be $\sim 13\%$ [11], consistent with other similar surveys.

1.2.2.3 Gravitational lensing

Finally, the third technique of inferring cluster masses is by observing the light deflection field of background galaxies and other objects, generated by a cluster in the foreground [12]. In the weak-field limit, the refractive index of a gravitational lens is directly proportional to its gravitational field. This enables images of lensed configurations to determine a true mass density map, including dark matter, without any assumptions of virial equilibrium. Figure 1.4 shows an image of galaxy cluster Abell 2218; lensing arcs are clearly visible in the image and have enabled determination of

\(2h\) is the Hubble constant in units of 100 km/s/Mpc.
mass-to-light ratios of $80–180 M_\odot/L_\odot$ for this cluster \cite{13}. A larger weak-lensing survey, the Sloan Digital Sky Survey, has used $\sim 130,000$ clusters in a stacking analysis to determine mass-to-light ratios, in general agreement with other techniques \cite{14}.

Figure 1.4: Image of cluster Abell 2218 from the Hubble Space Telescope. Prominent lensed arcs are visible from gravitational lensing. *Image from: NASA/HST.*

1.2.2.4 Bullet Cluster

The observation of cluster 1E 0657-56 (Bullet Cluster) serves as a poster child for the case for dark matter. A study using all of the techniques mentioned above for cluster characterization has presented a picture of a merger of two galaxy subclusters \cite{15}. Figure 1.5 shows two images from \cite{15} that indicate a large separation between the subcluster mass densities inferred by gravitational lensing but a smaller separation and a bow shock between clumps of baryonic gas, inferred by X-ray imaging. This suggests that a majority of the matter content of the cluster is non-baryonic and collisionless, enabling the mass-centers of the subclusters to pass through each other during the collision, while the interaction of gas from the subclusters caused a bow shock and heating.

1.2.3 Modern cosmology

I now move on to evidence from cosmological observations made at the largest scales that further solidifies the footing on which the dark matter hypothesis stands. We see evidence at these scales requiring a large fraction of dark matter to be cold (non-relativistic) in addition to being non-baryonic.
Figure 1.5: Left: Optical image of Bullet Cluster from the Hubble Space Telescope with weak lensing contours indicating mass density distribution. Right: X-ray image of Bullet Cluster from the Chandra X-ray Observatory, again with weak lensing contours indicating mass density distribution. Image from: [15].

1.2.3.1 Big-Bang Nucleosynthesis

In the first few minutes after the Big Bang, conditions were briefly suitable for free neutrons and protons to undergo fusion and form the light elements — $^2$H, $^3$He, $^4$He, $^3$Li. The time duration of this process is strictly controlled by the expansion rate of the universe, since the process can only occur while the universe is still sufficiently hot. It turns out that this is regulated strictly by the baryon-to-photon ratio $\eta = n_B/n_\gamma$ at the time. More baryons for a given photon density implies faster fusion of $^2$H and $^3$He into $^4$He, leaving smaller residual quantities of the former. Measurements of the absolute intensity of the Cosmic Microwave Background (Section 1.2.3.2) constrain the photon density very precisely, so the relic densities of the light elements set stringent constraints on the baryon fraction of the universe [16].

The best constraints are set by $^2$H, which has no mechanism for production after nucleosynthesis. As shown in Figure 1.6, deuterium abundance constrains $\eta \approx 5.5 \times 10^{-10}$ [17], implying a baryon density $\Omega_B \approx 0.04$. Thus a large fraction of matter in the universe has to be non-baryonic.

1.2.3.2 Cosmic Microwave Background

The Cosmic Microwave Background (CMB) is the relic blackbody radiation of the early universe. It constitutes a snapshot of the universe at the point when the photon-matter plasma had cooled sufficiently to enable most electrons and protons to combine into neutral hydrogen, and for photons to decouple and stream freely. This happened $\sim 380,000$ years after the Big Bang ($z \sim 1100$), and the CMB is observed today as a redshifted spectrum with blackbody temperature of 2.725 K, and peak frequency of $\sim 160$ GHz. The CMB and its anisotropies provide standard measures to characterize various cosmological parameters and are a critical component of the precision cosmology of today [18].
Figure 20.1: The abundances of $^4$He, D, $^3$He, and $^7$Li as predicted by the standard model of Big-Bang nucleosynthesis [11] — the bands show the 95% CL range. Boxes indicate the observed light element abundances (smaller boxes: $\pm 2\sigma$ statistical errors; larger boxes: $\pm 2\sigma$ statistical and systematic errors). The narrow vertical band indicates the CMB measure of the cosmic baryon density, while the wider band indicates the BBN concordance range (both at 95% CL). Color version at end of book.

Figure 1.6: Abundances of light elements as predicted by Big-Bang Nucleosynthesis, reported as a fraction of hydrogen abundance, vs. baryon-to-photon ratio. Bands show 95% CL. Boxes indicate observed abundances with $\pm 2\sigma$ error bars — the small boxes show statistical errors, and the big ones show those and systematic errors. Plot from [17].
CMB temperature primary anisotropies are sourced by the acoustic oscillations of the photon-baryon plasma before decoupling. Baryons tended to gravitationally collapse and form overdensities, whereas the photon pressure countered these overdensities, causing oscillations. At the time of decoupling, the CMB recorded the state of these oscillations. The key is that not all matter participated in these acoustic oscillations — the dark matter was collisionless and immune to resistance by photon pressure. The relative amplitude of the peaks of the CMB angular power spectrum, shown in Figure 1.7, constrain baryon density and non-baryonic dark matter density, when combined with other cosmological observations [19].

Figure 1.7: Temperature power spectrum as a function of multipole moment (l), from WMAP seven-year data. Also included are data points from ACBAR and QUAD for l ≥ 690. The line through the data points is a fit to WMAP data alone. Plot from [19].

1.2.3.3 Large-Scale Structure formation

The final piece of evidence for dark matter, I report here, is the observed large-scale structure today, which is inconsistent with a universe comprised only of baryonic matter [20]. From the CMB, we have deduced the matter-density-fluctuation power spectrum at the time of photon-baryon decoupling (z ∼ 1100). While density fluctuations at scales shorter than the horizon scale remain frozen until matter-radiation equality (z ∼ 3000), they are free to grow afterwards. This is not true, however, for baryonic matter, which continues to oscillate in a photon-baryon fluid until decoupling. Baryonic overdensities can grow only after z ∼ 1100. Thus, the overdensities from baryonic matter alone are insufficient to seed large-scale structure formation in the timespan that we start observing galaxies, clusters etc. This problem is solved if a component of matter such as dark matter, uncoupled to...
photons, is allowed to grow in overdensities between matter-radiation equality and baryon-photon decoupling. After decoupling, baryons track these pre-existing overdensities, leading to a start of structure formation on the correct timescale. However, this also imposes the requirement that dark matter be non-relativistic at the time of matter-radiation equality to prevent it from free-streaming away from overdensities.

1.2.3.4 Combined constraints

Observations of CMB, BBN, and large-scale structure paint a picture of a universe with a sizable fraction of the mass-energy budget contained in cold, non-baryonic dark matter. The WMAP seven-year data release uses a combination of datasets from all the cosmological observations mentioned above to constrain cosmological parameters of the ΛCDM model. Their reported density of cold dark matter is $\Omega_{CDM} h^2 = 0.1126 \pm 0.0036$ and that of baryonic matter density is $\Omega_b h^2 = 0.02255 \pm 0.00054$. Using the current best fit value of $h = 0.702 \pm 0.014$, $\Omega_{CDM} = 0.228 \pm 0.016$ and $\Omega_b = 0.0457 \pm 0.0027$.

1.3 Dark Matter Candidates

Now I provide a brief overview of some candidates for dark matter. I emphasize Weakly Interacting Massive Particles (WIMPs), the subject of the search described in this dissertation. In Section 1.4 I describe a popular particle physics framework for WIMPs called Supersymmetry. In Chapter 2 I describe the general strategies employed to search for WIMPs.

1.3.1 Modified gravity

The simplest and most naïve explanation offered to explain the discrepancy of galactic rotation curves was to modify the theory of Newtonian gravity over long distances. This required no dark matter, and was done through simple one-parameter fits [21]. Such modified Newtonian dynamics (MOND) failed to explain gravitational lensing. Extended relativistic theories of gravity such as TeVeS [22] use MOND-like phenomenology to overcome some of these issues, but by and large fail at explaining clusters and the observed matter power spectrum. Proponents of modified gravity typically invoke a combination of neutrinos and TeVeS-like theories to match observations such as the Bullet Cluster, where the separation of mass density and gas is difficult to explain through modified gravity alone [23]. Regardless, the consensus in the cosmology and particle astrophysics communities is that modified gravity theories are not consistent with all cosmological observations.

---

$h$ is the Hubble constant in units of 100 km/s/Mpc
1.3.2 Neutrinos

Standard-Model neutrinos almost certainly contribute to some fraction of non-baryonic dark matter in the universe. Early in the universe’s history, at sufficiently high temperatures, neutrinos are in thermal and chemical equilibrium with other particle species, with balance between rates of production and annihilation. As the universe expands and cools down, annihilation becomes inefficient and the population of neutrinos “freezes-out.” This point occurs when the annihilation rate $\Gamma_A < H$, where $H$ is the Hubble expansion rate. There are no known mechanisms to enhance or reduce the number of neutrinos in the universe after freeze-out, so their energy density declines with scale $a$ as $\rho_\nu \propto a^{-3}$. Assuming Majorana neutrinos, the relic abundance can be worked out [20] to be:

$$\Omega_\nu h^2 = \sum_i \frac{m_i}{94 \text{eV}/c^2}$$

where $m_i$ is the mass of an individual neutrino species $i$. For all of dark matter to be attributed to neutrinos, we would require $\sum_i m_i \sim 10 \text{eV}$. Current cosmological constraints on this number are $\sum_i m_i < 0.44 \text{eV}$ [19], limiting the fraction of neutrinos to a few percent of dark matter at best. Additionally, a neutrino abundance too high at matter-radiation equality would reduce the required matter overdensities that seed structure formation at small scales, leading to a matter power spectrum inconsistent with what we observe today.

1.3.3 Axions

Another attractive dark matter candidate is a one originally proposed to fix the “strong CP problem” — a pseudo-Goldstone boson, which ends up being potentially cosmologically significant. The QCD Lagrangian includes a CP-violating term, which would force neutrons to have an electric-dipole moment. Experimental constraints on neutron electric-dipole moment show that the coefficient of the CP-violating term is very small, $\Theta < 10^{-10}$ [17], although it is normally expected to be $\sim O(1)$. An attractive solution to the problem is to minimize the CP-violation by invoking the symmetry breaking of a U(1) global symmetry, called the Peccei-Quinn symmetry [24]. This generates a pseudo-Goldstone boson, called the axion. In terms of the QCD Lagrangian, the CP-violating term is driven to zero by adding $-m_a/f_a$ to its coefficient, where $m_a$ is the axion mass and $f_a$ is a coupling constant. Originally, it was assumed that $f_a$ was around the electroweak symmetry-breaking scale, but such “standard” axions have been experimentally ruled out [17]. However, any $m_a$ and $f_a$ solve the strong CP problem, as long as

$$f_a m_a = C f_\pi m_\pi$$

where $C$ is a constant of $\sim O(1)$ and $f_\pi$ and $m_\pi$ are the pion decay constant and mass, respectively. Thus the search for “invisible” axions with $f_a >>$ the electroweak symmetry-breaking scale remains
interesting. The production of axions in the early universe, even at QCD energies, occurs as a non-relativistic Bose-Einstein condensate \[25\]. This makes them an interesting cold dark matter candidate.

The expected relic density of axions, derived in \[20\], is

\[
\Omega_a h^2 \approx 0.3 \times \left( \frac{f_a}{10^{12} \text{ GeV}} \right)^{7/6}
\]

making \(m_a \sim 10 \mu\text{eV}\) viable for dark matter. Generally, \(10^{-3} < m_a < 10^{-6} \text{ eV}\) is the interesting range for axion masses based on two bounds. The first is a constraint of overclosure of the universe for axions that are too light. This second is that the observed length and intensity of the neutrino burst from supernova 1987A would be too high, if axions are too heavy. Searches for axion dark matter are based on axion to photon decay or on the Primakoff process, which converts axions to photons in the presence of strong magnetic fields \[26\]. These searches use microwave cavity resonators and look for photon excesses; limits they have set are shown in Figure 1.8. Note that in the process of acquiring sensitivity to interesting axion models, these experiments also search for axion-like particles that do not necessarily satisfy Equation 1.3, i.e., the areas outside the colored axion model band.

Figure 1.8: Squared axion-photon coupling vs. axion mass for cavity searches. Axion models are shown with the blue band. The vertical bands show parameter space ruled out by searches. *Plot from: \[17\].*

1.3.4 Weakly Interacting Massive Particles (WIMPs)

The arguably most appealing class of candidates for dark matter is Weakly Interacting Massive Particles (WIMPs). They are named so because they interact with other particle species only
through the gravitational force and a weak force. WIMPs are a natural choice for dark matter because of independent reasons in both cosmology and particle physics that I outline below.

1.3.4.1 Thermal production arguments

Just like neutrinos, WIMPs are hypothesized to be thermally produced in the early history of the universe. WIMP particles, $\chi$, remain in thermal and chemical equilibrium with other particle species until the temperature of the universe, $T$, drops below $M_\chi$, the mass of the WIMP. At this point, production ceases, but annihilation continues. Annihilation maintains equilibrium until its rate, $\Gamma_{\chi\chi}$, falls below $H$, the expansion rate of the universe. At that point, WIMPs freeze out, and their number density is approximately given by that for a massive non-relativistic particle [20]:

$$n_\chi \approx 4 \left( \frac{M_\chi k T_f}{2 \pi \hbar^2} \right)^{3/2} \exp \left( -\frac{M_\chi c^2}{k T_f} \right)$$

where $T_f$ is the freeze-out temperature for WIMPs. For $M_\chi \sim 10$ GeV, the relic abundance can be shown to be [27]:

$$\Omega_\chi h^2 \approx 3 \times 10^{-28} \frac{\text{cm}^3/\text{s}}{(\sigma_{\chi\chi} v)}$$

For $\Omega_\chi h^2 \sim 0.1$, we learn from this equation that the annihilation cross section is characteristic of weak-scale interactions. This is because the order of magnitude of the matrix element governing the annihilation process matches weak force coupling, $G_F$. Thus, simple cosmological arguments for a thermal relic of suitable abundance to be dark matter choose a weakly interacting particle in $\sim$ GeV mass range.

1.3.4.2 Particle physics arguments

On the particle physics front, it is now apparent that the Standard Model picture is unsettlingly incomplete. We understand how electroweak symmetry breaking (EWSB) generates the $W^\pm$ and $Z^0$ bosons, but a scalar Higgs field is required to explain their masses. The mass of the as-yet undiscovered Higgs boson, $M_H$, is constrained experimentally to be $\sim O(100)$ GeV/$c^2$. It is unclear why quadratically divergent radiative corrections to $M_H$ do not pull it towards the Planck scale, $M_{\text{Planck}}$. Fermions and gauge bosons are protected from radiative corrections because of chiral and gauge symmetries, respectively, but the Higgs, as a scalar, is not. This is called the “hierarchy problem.” The problem might be resolved if new, undiscovered physics at the weak scale naturally suppressed these radiative corrections and kept the Higgs mass from requiring fine tuning. Such physics will be associated with new particles of masses $\sim O(10-100)$ GeV/$c^2$; one of these may be stable, providing a natural dark matter candidate.

Remarkably, cosmology and particle physics both point to the weak scale and the same approx-
imate mass range to provide a viable dark matter candidate with the correct relic density.

1.4 Supersymmetric WIMPs

One of the most promising and well-motivated extensions to the Standard Model is a framework called Supersymmetry (SUSY). Many models within this framework easily provide a WIMP dark matter candidate. Exhaustive literature exists on the subject [28, 29, 30], and I only provide a brief overview here.

1.4.1 Framework and consequences

Supersymmetry, for the purposes of this discussion, requires that, if supersymmetry holds, the Lagrangian of the full theory is invariant under transformations of fermions into bosons and vice versa. This implies fermions and bosons must be partnered. The existing particles cannot be partnered with each other due to mismatched quantum numbers under other symmetries, so new partner particles must be introduced.

As a consequence, all radiative corrections to the Higgs mass are offset by equal corrections from the superpartners, but with opposite sign. Since superpartners with masses of Standard Model particles have not been observed, it is safe to say that if SUSY is a fundamental symmetry of nature, it has been broken. Thus superpartner masses will be different from the Standard-Model particles. Regardless, the effect of radiative corrections appears as a difference in masses times the logarithm of the energy cut-off, which still prevents the Higgs mass from running away to $M_{\text{Planck}}$.

Another problem addressed by SUSY is that of coupling unification at high energy scales. Coupling constants, which govern the strength of interactions at vertices, change as a function of interaction energy, as shown in Figure 1.9. It is believed that the $U(1)$ hypercharge, the $SU(2)$ electroweak and the $SU(3)$ strong force couplings all unify at some Grand Unified Theory (GUT) scale. However, under the Standard Model, the force couplings almost unify at $\sim 10^{14}$ GeV, but only within a few orders of magnitude in energy. The introduction of new particles in SUSY models tends to fix this problem — the running of coupling constants is modified to incorporate correction terms from the superpartners, thereby providing better unification, as seen in Figure 1.9.

1.4.2 Minimal Supersymmetric Standard Model

While SUSY provides the framework for extending the Standard Model, specific theories or models are required to fill in the details of particle masses, trilinear couplings, mixing angles, etc., since supersymmetry would be broken in nature. The simplest of these is called the Minimal Supersymmetric Standard Model (MSSM) [31], which adds the bare minimum particles and parameters to the Standard Model:
While the apparent unification of gauge couplings at some high energy scale naturally accommodates gauge coupling unification below the electroweak scale as a strong hint in favor of a grand unified theory (GUT), the specific mass thresholds are varied between the Standard Model (dashed lines) and the MSSM (solid lines). Figure 5.8 compares the RG evolution of the inverse gauge couplings in the Standard Model (dashed lines) and in MSSM (Section 1.4.2). The 1-loop RGEs for the Standard Model include only SU(2) and SU(3) couplings and soft masses as well.

The particles above do not necessarily represent mass eigenstates. In particular, the electroweak Majorana fermions mix to give the following particles:

1. Majorana neutralinos: $\chi^0_1, \chi^0_2, \chi^0_3, \chi^0_4$, each a mixture of $\tilde{b}^0, \tilde{w}^0$, and $\tilde{h}^0_{1,2}$.
2. Majorana charginos: $\chi^+_1, \chi^+_2$, each a mixture of $\tilde{w}^±$ and $\tilde{h}^0_{1,2}$.

For SUSY to be actually broken in MSSM, “soft breaking terms” are introduced in the Lagrangian with no assumptions about the mechanism causing the breaking. It is assumed that all SUSY breaking is diagonal in flavor space and that new CP-violating phases vanish.

Even with bare minimum additions, MSSM contains ~100 new parameters. Further simplifications are made to produce what is called the Constrained MSSM (CMSSM) [32], which naturally lead to GUTs. The resulting scheme is characterized by just five parameters:

1. Scalar mass $m_0$: The scalar masses all unify at the GUT scale to this mass.
2. Gaugino mass $m_{1/2}$: The gaugino masses all unify at the GUT scale to this mass.
3. Trilinear coupling $A_0$: All trilinear couplings unify at the GUT scale to this value.
4. $\tan \beta$: The ratio of the two neutral non-pseudoscalar Higgs’ vacuum expectation values.
5. $\text{sign}(\mu)$: The sign of the Higgs mass parameter.

This scheme enables the study of the SUSY parameter space using Monte Carlo simulation algorithms, without necessarily being exactly representative of nature.

### 1.4.3 Dark matter from CMSSM

There is one last issue that needs to be addressed before dark matter naturally emerges from SUSY. SUSY in its naïve form introduces some baryon- and lepton-number violating couplings. These need to be prohibited because baryon- and lepton-number violation would make the proton lifetime short ($\sim 10^{-2} \text{s}$), whereas stringent experimental constraints confirm that the proton lifetime is $> 10^{31} \text{y}$. A new discrete symmetry called R-parity is introduced to resolve this issue and is defined as $R = (-1)^{2s+3B+L}$, where $s$, $B$, and $L$ are the particle’s spin, baryon number, and lepton number, respectively. All standard particles and the Higgs bosons have $R = 1$ while all the superpartners have $R = -1$. Conservation of R-parity thus forces the proton to be stable and the lightest supersymmetric particle (LSP) to be stable as well. This makes the LSP of any supersymmetric model a viable dark matter candidate.

The lightest neutralino, $\chi^0_1$, is typically an LSP in several SUSY models and one of the most common dark matter candidates. For the CMSSM, one can compute the relic abundance of LSP neutralino dark matter as a function of model parameters. Four regions of parameter space have been identified, where the relic density is compatible with cosmology and particle physics constraints. These are shown for $A_0 = 0$, $\tan \beta = 45$, and $\mu < 0$ in Figure 1.10 for mSUGRA (minimal Supergravity), a popular CMSSM:

1. Bulk region: A region of p-wave neutralino to fermion annihilation, at low $m_0$ and low $m_{1/2}$, but a large part of it has been excluded by LEP2 bounds on Higgs and sfermion masses.
2. Hyperbolic branch/Focus point: At large $m_0$, there are heavy scalars and a heavy neutralino ($\gtrsim 1 \text{TeV}/c^2$), which leads to s-wave annihilation channels to gauge bosons.
3. Co-annihilation region: In this region, at low $m_0$ and most values of $m_{1/2}$, the LSP and the next-to-lightest SP (usually the stau) are almost degenerate in mass. During the late stages of LSP annihilation in the early universe, before freeze-out, the LSP and NLSP are able to interconvert and enhance overall annihilation rates for both. This allows the LSP to attain the correct relic density.
4. A-annihilation funnel: At large $\tan \beta$, the lightest neutralino is approximately half the mass of a particle such as $A^0$, allowing resonant annihilation.
Plot 2.9 in the next chapter, shows two interpretations of theoretically preferred MSSM space in terms of parameters constrained by direct searches for WIMPs.

Figure 1.10: Plot of constraints in $m_0$ vs. $m_{1/2}$ for mSUGRA parameter space. The green contours show WMAP five-year data release constraints for relic dark matter density in this parameter space. The red contour is the LEP-2 $m_h = 114.1$ GeV Higgs bound. Contours of muon anomalous magnetic moment (blue), $b \rightarrow s\gamma$ branching fraction ($\times 10^4$) (magenta), and $B_s \rightarrow \mu^+\mu^-$ branching fraction (cyan) are also indicated. Regions completely excluded are filled with solid color. Plot from: 33.
Chapter 2

Detecting WIMPs

As explained in the previous chapter, Weakly Interacting Massive Particles (WIMPs) are a leading class of candidates to be the dark matter of the universe. In this chapter I review the experimental approaches to search for WIMPs — production in particle colliders, astrophysical searches for WIMP-annihilation products and direct searches using sensitive, low-background particle detectors. For the first two techniques, I provide only a brief overview of the state of these vast fields as relevant to the search for WIMPs. For direct detection, the subject of this dissertation, I provide a more detailed framework to motivate its plausibility and the strategies used. It is worth noting that all three techniques provide different information about WIMPs and are all required to understand the full picture of WIMP dark matter; I comment on this at the end of the chapter.

2.1 Production in Particle Colliders

The search for new particles is traditionally the purview of particle colliders. Although WIMPs are likely an old relic, they will be produced in colliders with sufficient energy reach. If WIMPs are part of supersymmetry or other such proposed extensions of the Standard Model, then they might be within the reach of the Large Hadron Collider (LHC) at CERN in the next few years or within the reach of the proposed International Linear Collider (ILC) in the next couple of decades [34].

The evidence for WIMPs at particle colliders would come though the observation of missing energy and momentum in particle collisions, because their scattering cross sections are too small to be directly detected. This signal is easily reconstructed in linear lepton colliders through various decay channels where the collision energy and momentum are precisely known. It is much harder in hadron colliders where the collisions are those of composite particles and the exact longitudinal energy and momenta of individual quarks are not known. Thus, at hadron colliders, the properties of new particles with small interaction cross sections must be inferred statistically from the “missing transverse momentum” of a large number of collisions. Many Standard-Model extensions predict new colored particles with mass similar to what is expected for WIMPs. Such particles would be
produced in abundance at the LHC if their mass is below $\sim 2\text{ TeV}/c^2$ [34]. These would decay to LSPs (potentially WIMP dark matter) and partons, producing a characteristic signature of hard jets and missing transverse momentum.

In the meanwhile, model-dependent limits have been set by various colliders. LEP searches for supersymmetry have set a lower limit on $M_\chi > 45\text{ GeV}/c^2$ [35]. For CMSSM, a plot of this lower limit combined with other constraints is shown in Figure 2.1. However, without the assumption of gaugino-mass unification, which limits the field content and masses of neutralinos, this limit is relaxed. Searches for charginos and neutralinos with D0 and CDF at the Tevatron have excluded some parameter space in the CMSSM with $\sim 2\text{ fb}^{-1}$ of integrated luminosity, with no excess events observed above background [36, 37]. A plot of these results is shown in Figure 2.2. Current constraints on CMSSM in the $m_{1/2}$ vs. $m_0$ plane from CMS with $\sim 35\text{ pb}^{-1}$ of integrated luminosity are shown in in Figure 2.3. Here too, no significant evidence for SUSY or WIMPs is seen yet.

![Figure 2.1: Lightest neutralino mass vs. CMSSM tan$\beta$, from searches at LEP for charginos, sleptons, and neutral Higgs bosons. The green hatched area is excluded by a combination of constraints from different channels marked by the various colored solid lines. Plot taken from: [35].](image)

**2.2 Indirect Astrophysical Detection**

An alternative to producing WIMPs in colliders is to try to detect the ones already in the universe. One way of doing this is to search for WIMP-annihilation signatures in regions of high dark mat-
Figure 2.2: Constraints from D0 and CDF on CMSSM in the $m_{1/2}$ vs. $m_0$ plane for $\tan \beta = 3$, $A_0 = 0$, and $\text{sign}(\mu) > 0$. The grey regions are excluded by LEP and the green regions are excluded by D0. The black contours indicate the region excluded by CDF. *Plot taken from: [38]*.

Figure 2.3: Constraints from CMS on CMSSM in the $m_{1/2}$ vs. $m_0$ plane for $\tan \beta = 3$, $A_0 = 0$, and $\text{sign}(\mu) > 0$. Colored areas indicate regions of parameter space already excluded by other experiments. CMS’s measured exclusion contours at 95% CL are plotted in dot-dashed green and solid red for leading-order and next-to-leading-order cross sections obtained from simulation. The dashed blue curve indicates the expected limit. Area below the curves is excluded. Also plotted are contours of constant squark and gluino masses. *Plot taken from: [39]*.
ter density \[39, 27\]. WIMP-annihilation rates scale as the square of the WIMP number density \(N_{\chi} \propto \Gamma_{\chi\chi \rightarrow X} \propto n_{\chi}^2\), so the annihilation processes that produced the WIMP relic density should still be active to some extent in overdense regions. The prime candidate regions to seek annihilation signatures are galactic halo centers \[40\]. WIMPs are also expected to accumulate in the centers of large celestial objects like the Sun and undergo annihilation \[41, 42\].

While the total annihilation cross section required to produce the WIMP relic density is known, the exact processes involved, branching fractions, and individual annihilation cross sections are not known. They are highly model-dependent and require assumptions to calculate expected rates. Also, a number of astrophysical assumptions need to be made for parameters that are not well known experimentally. This includes the shape of galactic dark matter halo profiles and their “clumpiness,” which can lead to vast differences in WIMP-density assumptions. These translate to orders of magnitude in uncertainty on expected rates of annihilation products \[34\]. Moreover, the rates of background processes, above which an annihilation signal might be claimed, have order-of-magnitude uncertainties associated with them in some cases.

Below, I provide a brief review of a few of the common indirect search techniques, organized by the annihilation product sought.

### 2.2.1 Gamma rays

Gamma rays might be an excellent tracer for WIMP annihilation in the galactic halo. Such gammas might be monoenergetic or part of a continuous spectrum \[43\]. If WIMPs can annihilate to a photon-containing two-body final state, then the resulting photons will be monoenergetic and easily distinguishable from continuum background. The \(\gamma\gamma\) or \(\gamma Z^0\) channels are the preferred mechanisms for this \[44\]. Unfortunately, the branching ratios of such modes are generally quite low. Also, the resolution of gamma ray telescopes is generally insufficient to distinguish lines from continuum. Continuous spectra, on the other hand, may result from the annihilation processes involving charged final states such as \(W^+W^-\) or bremsstrahlung from charged intermediate states \[45\]. These would be brighter than monoenergetic gammas but would be harder to distinguish from astrophysical processes.

Atmospheric Cherenkov telescopes located on the Earth, such as MAGIC, VERITAS, and H.E.S.S., search for gammas with energies > 30 GeV, i.e., gammas that can penetrate the atmosphere without being lost to electron-positron pair production \[46\]. These experiments have not observed significant excesses of gammas that could be attributed unambiguously to WIMP annihilation. H.E.S.S. reported an excess of \(\sim\) TeV gammas from a point-like source at the galactic center, which was interpreted as the annihilation signature of a 10 GeV/c^2 WIMP \[47\]. However, the observed spectrum is rather easily explained by a power law indicative of an astrophysical source.

Space-based telescopes are able to observe gammas with lower thresholds (tens of MeV as op-
posed to tens of GeV) and with better resolution than ground-based instruments. Of note are EGRET aboard the Compton Gamma Ray Observatory (CGRO) and LAT aboard the Fermi satellite. EGRET collaborators claimed a gamma ray excess in their observed gamma spectrum and invoked an $\sim 80\text{GeV}/c^2$ WIMP hypothesis to explain it, by assuming an unusually clumpy dark matter halo \cite{48}. An analysis of LAT data for the same energy range has attributed the spectrum to secondary production processes, ruling out a WIMP explanation for the EGRET excess \cite{49}. Also, these data have been interpreted under a WIMP hypothesis to set upper limits on WIMP-annihilation cross section in the quark, lepton, and photon channels. The limits set for $b\bar{b}$ annihilation for an MSSM WIMP using the continuous gamma spectrum are shown as an example in Figure 2.4.

![Figure 2.4: Cross section $\langle \sigma v \rangle$ vs. WIMP mass for WIMP annihilation into $b\bar{b}$ states from the Fermi-LAT data. The lines represent limits set at different confidence levels (marked as percentage on the line) and under different annihilation enhancement scenarios (type of line) described in detail in \cite{49}. The blue area marks the exclusion region under one of these scenarios. Finally, the grey shaded regions represent parameter space where the branching ratio for annihilation into $b\bar{b}$ states is greater than 80% under the MSSM model used by the authors. Plot from: \cite{49}.](image)

### 2.2.2 Antimatter

Another category of indirect WIMP-annihilation signatures is that from antimatter — primarily positrons, antiprotons, and antideuterons \cite{27}.

WIMP annihilation is expected to produce equal numbers of electrons and positrons, but would change the observed electron-positron ratio from theoretical expectation — there would be an excess of high-energy positrons in place of a declining power law from background processes \cite{27}. HEAT observed a positron excess at 8 GeV, which has been interpreted as a possible WIMP signature \cite{50}.
PAMELA has also measured an excess in positron fraction in the 10–100 GeV range [51], plotted in Figure 2.5. The total electron+positron spectrum measured by ATIC, Fermi-LAT, and HESS also exceeds simple secondary production model predictions. However in all these cases, the secondary production models contain over an order-of-magnitude uncertainty in predicted spectra based on choice of model parameters [52]. Thus it is plausible for these observations to be explained by secondary production processes. Nonetheless, speculation abounds on dark matter interpretations [53, 54, 55].

![Figure 2.5: Positron fraction vs. energy measured by PAMELA and other experiments.](image-url)

The antiproton signal is expected to be distinguishable from background only at low energies. PAMELA’s [56] and BESS’s [57] measurements of the antiproton spectrum have shown no excess so far.

Antideutrons would provide a strong signal for WIMP annihilation, especially because of low backgrounds in the signal region, but are expected to be generated far more rarely. Sufficient data have not been acquired yet to allow any firm conclusions on this front, though AMS-02, launched recently on one of the final US space shuttle missions, and GAPS, an experiment to be flown on a 100-day Antarctic balloon flight in 2014, will probe this potential signature [58, 59].

### 2.2.3 Neutrinos

WIMPs are expected to be attracted to large gravitational wells such as that of the Sun. Despite the fact that they rarely interact with matter, some WIMPs might lose energy through occasional elastic scatters in the Sun and become gravitationally bound. Over a time shorter than the age of

\[ \text{ratio of positrons to sum of positrons and electrons} \]
the solar system, these WIMPs would interact more frequently and settle to the center leading to an equilibrium WIMP overdensity balanced by captures and annihilation. While most annihilation products would not escape the Sun, some muon neutrinos would travel unimpeded and might be detected in large neutrino telescopes on Earth [41, 42]. The same processes would also occur for the Earth, but with much less efficiency and far fewer annihilations.

High-energy neutrinos (GeV–TeV range) from WIMP annihilation in the sun may interact with material surrounding neutrino detectors, generating muons, which constitute the signal these detectors seek. These muons are distinguished from those generated by other neutrino sources by looking for upward-traveling muons. Such muons would be generated by neutrinos with much higher energies, traveling all the way through the Earth before reaching the detector, than typical solar neutrinos (keV–MeV range).

High-energy neutrino detectors are large Cherenkov detectors such as Super Kamiokande (SuperK), AMANDA, and IceCube. SuperK, a 50,000 m$^3$ water detector, has used its $\sim 1700$-liveday dataset and searched for events within a few degrees of the line of sight to the Sun’s center. It has set best upper limits on muons of $\sim O(1000) / \text{km}^2 / \text{year}$ [60]. AMANDA and now IceCube are continually instrumenting the Antarctic ice with photomultiplier tubes (PMTs) to have ever-larger volumes to conduct high-energy neutrino searches. Their current limit on high-energy neutrinos from the Sun is a factor of a few better than SuperK’s [61, 62].

WIMP-annihilation neutrino production rates in the Sun are governed by WIMP capture and hence by WIMP-nucleon scattering cross section. Thus the results of high-energy neutrino searches for dark matter can be interpreted as limits on the WIMP-scattering cross section, by making model assumptions about the WIMP halo, WIMP capture, and WIMP-annihilation branching fractions. Generally, these are not competitive with direct search limits on spin-independent WIMP-nucleon cross section, but provide excellent upper limits on the spin-dependent WIMP-proton cross section, because of the hydrogen content of the Sun. These are plotted in Figure 2.6.

### 2.3 Direct Detection

Finally, instead of producing WIMPs, or seeking their annihilation signature, we can directly search for them through their interactions in terrestrial detectors. As the Earth moves through the Milky Way’s dark matter halo of density $\rho_x \approx 0.4 \text{ GeV}/c^2/\text{cm}^3$, at $v_E \approx 250 \text{ km/s}$ [63], we would experience WIMP flux of $\sim 10^5 / \text{cm}^2 / \text{s}$ from a $100 \text{ GeV}/c^2$ WIMP! Yet, WIMPs are hard to detect because of their vanishingly small interaction cross section with regular matter.

The idea for experimental efforts to directly set limits on dark matter interaction was first proposed by Goodman and Witten. They realized that non-relativistic weak dark matter particles would coherently scatter off nuclei, producing detectable recoils in sufficiently sensitive detectors [64].
A strong motivation to believe that WIMPs might not be completely non-interacting is the "crossing argument," suggesting similar cross sections for WIMP annihilation and scattering processes \cite{27}. In Section 1.3.4 I showed that WIMPs would have an annihilation cross section of $\sim 1$ pb to set the correct dark matter relic density. Since the WIMP scattering and annihilation processes share the same matrix element, it is not unreasonable to expect the scattering cross section to be similar to the weak-scale annihilation cross section. Note that this is mostly a qualitative argument to suggest a starting point to probe scattering cross sections. The precise relationship between these two cross sections is model-dependent. Assuming for the moment that the WIMP-nucleon scattering cross section is indeed $\sigma \sim 1$ pb, a 100 GeV/c$^2$ WIMP would produce a few scatters per day in a kilogram of hydrogen target mass. This is a sufficiently rare rate to require carefully built low-background experiments for this pursuit.

In this section, I provide a quantitative argument for the expected WIMP-scatter rate and details of the nature of WIMP-nucleon couplings. Then I comment on the effects of background on sensitivity and the general strategies employed by direct searches. I conclude with a very brief look at the current state of the field.

### 2.3.1 Rates and recoil spectra

Even without WIMP model-specific assumptions, one can construct general arguments for rates of interaction of WIMPs in terrestrial detectors and the recoil energy spectra one would expect in such detectors. These are explained in detail in \cite{64} by Lewin and Smith, so I just summarize these arguments here.

We start by assuming that WIMPs populate an isothermal halo in the galactic rest frame fol-
lowing a Maxwellian phase-space distribution of velocities, given by:

\[ f(v, v_E) \, d^3v = e^{-\left(\frac{v+v_E}{v_0^2}\right)^2} \, d^3v \]

(2.1)

where \( v_0 \) is the characteristic or most probable WIMP velocity with respect to the Earth and \( v_E \) is the Earth’s velocity with respect to the galactic rest frame. We use \( v_0 \approx 220 \text{ km/s} \) in the solar neighborhood [65]. For the Earth’s velocity, we use \( v_E \approx 244 + 15 \cos (2\pi t) \text{ km/s} \), where \( t \) is the time measured in years since the maximum velocity near June 2 [65]. This sinusoidal functional form approximately captures the motion of the Earth relative to the Sun, as the solar system orbits the Milky Way.

The WIMP phase-space distribution is related to the number density of WIMPs, \( n_0 \) by a simple differential relationship:

\[ dn = n_0 \frac{f(v, v_E) \, d^3v}{\int_0^{v_{esc}} f(v, v_E) \, d^3v} \]

(2.2)

where \( v_{esc} \) is the WIMP escape velocity, i.e., the velocity beyond which WIMPs would be ejected from the halo. We assume \( v_{esc} \approx 544 \text{ km/s} \) [66]. The denominator of the above expression provides appropriate normalization.

With these expressions in hand, we can write down a naïve differential rate of WIMP scattering with a target nucleus of mass atomic mass \( A \):

\[ dR = \frac{N_0}{A} \sigma \, v \, dn \]

\[ = \frac{0.932N_0}{M_T} \sigma \, n_0 \, \frac{v \, f(v, v_E) \, d^3v}{\int_0^{v_{esc}} f(v, v_E) \, d^3v} \]

(2.3)

where \( N_0 \) denotes the Avogadro number, \( \sigma \) is the WIMP-nucleus scattering cross section, and, in the second step, atomic mass \( A \) is converted to \( M_T \) in GeV/c^2, and Equation 2.2 in used. Next, we rewrite this expression using Equation 2.1. Also, the WIMP number density is rewritten as \( n_0 = \rho_\chi/M_\chi \), where \( \rho_\chi \) is the WIMP mass density (taken to be, \( \rho_\chi = 0.3 \text{ GeV}/\text{cm}^2 \) in the solar system neighborhood [67] and \( M_\chi \) is the WIMP mass. Thus we get the WIMP-nuclear-scattering differential rate in terms of WIMP velocity:

\[ dR = \frac{0.932N_0}{M_T} \sigma \, \rho_\chi \, \frac{v \, e^{-\left(\frac{v+v_E}{v_0^2}\right)^2} \, d^3v}{\int_0^{v_{esc}} e^{-\left(\frac{v+v_E}{v_0^2}\right)^2} \, d^3v} \]

(2.4)

Note that the simple astrophysical assumptions made so far provide values for everything in this expression except the WIMP properties \( \sigma \) and \( M_\chi \). However, we already see that for a specified value of \( M_\chi \), a measurement of WIMP-scattering rate enables us to probe WIMP-scattering cross section. This is what provides direct search experiments sensitivity to WIMP-scattering cross section. The differential expression in Equation 2.4 is not yet useful because we will not measure \( R \) for a slice \( d^3v \).
in phase space. The differential rate of interest is that with respect to recoil energy measured in a particle detector, $E_R$.

If a WIMP of energy $E = \frac{1}{2} M_\chi v^2$ scatters off a target nucleus, the recoil energy of that nucleus will be:

$$E_R = \frac{1}{2} E \cdot r (1 - \cos \theta) \quad (2.5)$$

where $r \equiv 4M_T M_\chi / (M_T + M_\chi)^2$. Assuming isotropic scattering, i.e., $0 \leq E_R \leq E \cdot r$ with uniform distribution, we can write the differential rate with recoil energy:

$$\frac{dR}{dE_R} = \int_{E_{\min}}^{E_{\max}} \frac{dR(E)}{E \cdot r} \quad (2.6)$$

Converting this integral in WIMP energy to one in WIMP velocity and invoking Equation 2.4, we get:

$$\frac{dR}{dE_R} = \frac{1.864N_0}{M_T} \frac{\rho_\chi}{r M_\chi^2} \int_{E_{\min}}^{E_{\max}} \frac{1}{v} e^{-(v + v_E)^2/v_0^2} \frac{d^3v}{v^2} \quad (2.7)$$

For the toy case of $v_E = 0$ and $v_{esc} = \infty$ this reduces to:

$$\frac{dR}{dE_R} = \frac{R_0}{E_0 r} e^{-E_R/E_0 r} \quad (2.8)$$

where $R_0$ and $E_0$ are the characteristic total scattering rate and WIMP energy, respectively. Thus we expect a featureless recoil spectrum exponentially decreasing with recoil energy.

### 2.3.2 WIMP-nucleon elastic scattering

So far, the event rate and spectra were computed without any input from WIMP-scattering physics. Here, I review expectations of WIMP-scattering cross section. Kurylov and Kamionkowski have shown that the WIMP-quark scattering amplitude is generally dominated by scalar and axial-vector terms in the the non-relativistic limit [68]. Converting these to WIMP-nucleon scattering cross section is dominated by uncertainties in the quark-flavor and spin content of nucleons [27]. Regardless, with some assumptions, the WIMP coupling to nucleons can be studied separately for spin-independent and spin-dependent cases.

#### 2.3.2.1 Spin-independent scattering

Scalar WIMP-nucleon interactions are characterized by coupling constants $f_p$ for protons and $f_n$ for neutrons. In the zero-momentum transfer limit, scattering amplitudes with individual nucleons add coherently, giving the following spin-independent WIMP-nucleon cross section [68]:

$$\sigma_{SI}^0 = \frac{4}{\pi} \mu^2_{\chi N} [Z f_p + (A - Z) f_n]^2 \quad (2.9)$$
where $\mu_{\chi N} = M_\chi M_N/(M_\chi + M_N)$ is the reduced mass of the WIMP and the nucleon, and $A$ and $Z$ are the atomic mass and number, respectively. Most supersymmetric models predict $f_p \approx f_n$ \cite{27}, implying $\sigma_{SI}^0 \propto A^2$. Thus WIMP detectors made of heavy nuclei such as germanium see a large coherent enhancement and are more sensitive to spin-independent interactions than those made of lighter nuclei.

In reality, collisions cause some momentum transfer $q = \sqrt{2M_T E_R}$. At sufficiently high $q$, the WIMP wavelength ($\hbar/q$) starts probing the structure of the nucleus, leading to imperfect interference of scattering amplitudes, lowering the cross section. The differential cross section acquires a multiplicative correction term $F^2(q)$ \cite{68}:

$$\frac{d\sigma_{SI}}{dq^2} = \frac{4}{\pi} \mu^2_{\chi N} |Z f_p + (A - Z) f_n|^2 F^2(q) \tag{2.10}$$

Several forms exist for $F^2(q)$, but for the work presented in this dissertation, we utilize the commonly used Helm form factor \cite{65}:

$$F(q) = 3 \frac{j_1(q r_n)}{q r_n} e^{-q^2 s^2/2} \tag{2.11}$$

where $r_n$ is the nuclear radius and $s$ is the nuclear-skin thickness. We use the following expressions for $r_n$ and $s$, also from \cite{65}:

$$r_n^2 = c^2 + \frac{7}{3} \pi^2 a^2 - 5 s^2$$

$$c = 1.23 A^{1/3} - 0.60 \text{ fm}$$

$$a = 0.52 \text{ fm}$$

$$s = 0.9 \text{ fm} \tag{2.12}$$

The form factor of Equation 2.11 is plotted for germanium and xenon in the left pane of Figure 2.7 as a function of recoil energy. This form factor is easily added as a correction to Equation 2.7. Integrating the resulting differential rate above a recoil energy threshold gives the total expected rate. This is done for a WIMP of mass 70 GeV/c$^2$ and $\sigma_{SI} = 10^{-44} \text{ cm}^2$ in the right pane of Figure 2.7.

It is worth noting that the multiplicative form factor, $F^2(q)$, is accounted for in the calculation of scattering rates when direct-search experiments set exclusion limits. The limits are therefore quoted for zero-momentum-transfer cross section and are normalized by the number of nucleons in the experiment’s target to enable comparison of results between different experiments.
2.3.2.2 Spin-dependent scattering

The spin-dependent WIMP-nucleon scattering amplitude is proportional to the inner product of the WIMP and nucleon spins, and its strength is governed by coupling constants $a_p$ and $a_n$ for protons and neutrons, respectively. The interaction amplitude switches signs when a nucleon spin is flipped, causing destructive interference between contributions of opposite-spin nucleons. Since nucleons align into spin-singlet pairs in a nucleus, the spin-dependent cross section depends only on unpaired nucleons. Thus spinless nuclei are completely insensitive to spin-dependent scattering and are undesired in a WIMP-search experiment seeking these interactions. As opposed to spin-independent searches, these searches prefer the use of light odd-nucleon targets to maximize nuclear spin per unit mass rather than target mass per unit volume.

The zero-momentum transfer limit WIMP-nucleon spin-dependent cross section is given by [68]:

$$\sigma_{SD}^0 = \frac{32(J + 1)}{\pi J} G_F^2 \mu_N^2 [a_p \langle S_p \rangle + a_n \langle S_n \rangle]^2$$

(2.13)

where $J$ is the nuclear spin and $\langle S_p \rangle$ and $\langle S_n \rangle$ are the expectation values of the proton and neutron spin, respectively, and must be obtained from nuclear structure calculations.

For spin-dependent scattering, $a_p \approx a_n$ in general since their signs and magnitudes vary with choice of WIMP model. This implies that a finite momentum transfer correction cannot be factored out of the cross section in a model-independent way, unlike the case for spin-independent interactions. The cross section is therefore written in the following form [69]:

$$\frac{d\sigma_{SD}}{dq^2} = \frac{8G_F^2}{(2J + 1)c^2} S(q)$$

(2.14)
where $v$ is the WIMP velocity and

$$S(q) \equiv a_0^2 S_{00}(q) + a_1^2 S_{11}(q) + a_0 a_1 S_{01}(q)$$

(2.15)

where $a_0 \equiv a_p + a_n$ and $a_1 \equiv a_p - a_n$. $S_{ij}(q)$ include the effects of finite momentum transfer as well as proton and neutron spin expectations, and must be computed separately for every constituent nuclide of a target mass. For the analysis presented in this dissertation, they are taken from [70].

2.3.3 Backgrounds

Direct search experiments seek to constrain properties of WIMPs by characterizing the scattering rates and spectral shapes of a handful of WIMP-nucleon interactions amidst a sea of background interactions with other particles. They do this by operating detectors with the innate capability of identifying and discriminating against background on an event-by-event basis or looking for astrophysical modulation signatures “on top” of their well-known backgrounds. However, backgrounds, if not identified as such, can contaminate signal rates and lead to incorrect conclusions about WIMP characteristics. Thus it is worthwhile understanding how unidentified background affects the sensitivity of a WIMP search. I follow the arguments made in [71] and [72]. If an experiment acquires exposure $MT$ (mass times exposure time, typically measured in kilogram-days) and expects to observe $B$ background events indistinguishable from WIMPs, it can operate in one of three background regimes:

1. **Background-free regime:** If $B << 1$ event, any observed signal candidates are evidence for WIMPs. In the absence of signal candidates, the experiment can set a 90% CL Poisson upper limit on the WIMP-scattering rate of $R = 2.3/MT$ events per kilogram-day. If $B << 1$ continues to remain true with increasing exposure, then the sensitivity improves proportionally to $MT$.

2. **Background subtraction:** If $B$ is non-negligible but is well characterized with negligible systematic errors, then it may be subtracted from the count of signal candidates. The statistical uncertainty on this subtraction is governed by Poisson statistics; $\sigma_B = \sqrt{B}$. Thus the residual count of signal candidates needs to be larger by some factor, say five times $\sigma_B$ to present evidence for WIMPs. As the exposure increases, we expect $B \propto MT$ and $\sigma_B \propto \sqrt{MT}$. The number of background-subtracted candidates needed to have sufficient evidence for WIMPs scales as $\sigma_B$, so sensitivity scales as $\sqrt{MT}$.

3. **Background-limited regime:** It is usually not possible to characterize $B$ with negligible systematic error. Even if it is for small exposures, systematic errors in estimating $B$ typically scale proportionally to $B$, whereas the statistical error scales with $\sqrt{B}$. For sufficiently large exposures, systematic errors become larger than statistical ones and if they truly scale as the
exposure $MT$, then no increased exposure will enhance the sensitivity of the experiment until improvements are made in background levels or background rejection.

Thus it is most desirable for direct search experiments to operate as close to the background-free regime as possible.

### 2.3.4 Search strategies and current status

All direct search experiments try to maximize passive shielding against backgrounds by locating themselves deep underground in clean low-radioactivity facilities. Beyond this, there are two general strategies employed by direct search experiments to search for WIMP scatters over their residual backgrounds.

#### 2.3.4.1 Modulation signatures

Some experiments search for WIMPs by seeking their astrophysical annual or diurnal modulation signature. The former is caused by the revolution of the Earth around the Sun as the solar system tracks its path through the Milky Way. This leads to a modulation in the Earth’s velocity with respect to the galactic frame and hence a detectable modulation of WIMP flux. The latter relies on a change in the WIMP flux as the Earth rotates on its axis, causing a change in the direction of WIMP-induced recoils. The detection of modulation-based evidence of WIMPs typically requires a very large sample of WIMP recoils, and extremely well-characterized and stable backgrounds.

The DAMA collaboration has claimed a $>8\sigma$ detection of WIMPs based on over a decade of annual modulation observed in the recoil spectrum of its NaI crystals [73]. A plot of their signal rate variation with time is shown in Figure 2.8. While they have taken exquisite care to minimize their background levels, it is uncertain whether the modulation they observe is a WIMP modulation, not just a seasonally correlated background modulation. A standard WIMP interpretation of DAMA’s modulation signal has been ruled out by several other direct search experiments. Recently, the CoGeNT collaboration has also claimed a $2.8\sigma$ annual modulation signal based on 1 year of observation with a Ge p-type point-contact bolometer, but with phase different from the expected modulation phase by $4\sigma$ [74].

Diurnal modulation searches for WIMPs are slowly getting off the ground. It is particularly difficult to measure the direction of a target nucleus recoil, unless the detector medium is a gas, in which case the target masses are usually light and small. With a careful choice of target gas, this makes them suited to detection of spin-dependent scattering signatures of WIMPs. The DMTPC [75] and DRIFT [76] collaborations operate TPC experiments to search for a diurnal signature of WIMPs.
Several analyses on the model-independent investigation of the DM annual modulation signature have been performed as previously done in Ref. [73]. The curve over the data points is a best fit sinusoidal with fixed period and phase to match the expected dark matter modulation signal.

### 2.3.4.2 Event-by-event discrimination

A large number of direct searches for WIMPs use active discrimination technologies on an event-by-event basis to pick out WIMP candidates from backgrounds in such a way that their residual background is negligible or small. They typically rely on the fact that WIMPs are expected to deposit energy in short dense tracks that produce less ionization and scintillation for a given recoil energy compared to electromagnetic backgrounds. Active discrimination experiments thus use some combination of phonons, ionization, and scintillation to measure recoil energy and identify the type of interaction.

There are a large number of existing and proposed experiments with active event-by-event discrimination. They use solid-state crystals (e.g., CDMS II [77], EDELWEISS-II [78], CRESST-II [79]) and liquid nobles xenon and argon (e.g., XENON 100 [80], WARP [81]) as targets to search for WIMP recoils. Another promising variant of the event-by-event discrimination technique is the COUPP experiment, which uses a superheated bubble chamber target tuned in pressure and temperature to be completely insensitive to electron-recoil background [82]. All these experiments have varying levels and modes of background contamination and strategies to mitigate them. It is worth noting that no experiment of this class has produced significant evidence for WIMPs yet. Most of them have set upper limits on the WIMP-nucleon scattering cross section as a function of WIMP mass based on a small number of observed events consistent with expected background or no events. In the low-mass WIMP regime, there is some controversy however. The CoGeNT experiment has claimed a detection [83], while CDMS II [1] and XENON 10 [84] have ruled out CoGeNT’s preferred WIMP model via low-energy-threshold reanalyses of their datasets. A plot of the current upper limits for spin-independent WIMP-nucleon cross section is shown in Figure 2.9. As we see in this plot, direct detection experiments are broaching interesting parameter space for WIMP-scattering models.

For the rest of this dissertation, I focus on the Cryogenic Dark Matter Search (CDMS) experi-
ment, one of the leading experiments in the field, and the one used for the WIMP search presented in this dissertation.

![Figure 2.9: WIMP-nucleon spin-independent cross section vs. WIMP mass. Limits are shown from XENON 100 (blue) [80], CDMS II (red, work described in this dissertation), EDELWEISS-II (dashed green) [78], and WARP (magenta dotted line) [81]. The regions enclosed within pink lines denote the DAMA/LIBRA 3σ allowed parameter space as interpreted in [85]. The region enclosed with a green line denotes the 90% CL boundaries of a CoGeNT-compatible WIMP model, determined by a DC-rate analysis, [83], though the region is statistically ill-defined. Finally the light shaded regions represent theoretically allowed MSSM space interpreted in grey by [88] and in green by [87]. Plot generated by: [88].](image)

### 2.4 Complementary nature of the different approaches to WIMP search

In closing this chapter, it is important to note that collider, indirect, and direct searches for WIMPs constitute complementary probes of WIMP dark matter. Colliders search for new particles probing physics that may be related to WIMPs. Indirect searches probe the annihilation processes that produced the WIMP relic density we observe today. Direct searches seek signals of WIMP-scattering off hadronic matter. In effect, these three probes ask different questions about WIMPs and provide three different handles on understanding them.
Also, the three different approaches are sensitive to different regions of WIMP parameter space, as is easily demonstrated with the CMSSM \[32\]. Current hadron colliders are most sensitive to the bulk region of CMSSM with its abundance of squarks and gluinos, while they have poorer sensitivity to the focus point. Indirect searches are most sensitive to the focus point and funnel regions. They have poorer sensitivity to bulk and coannihilation regions, the WIMPs from which were produced by temperature-dependent annihilation mechanisms and NLSP-coannihilation, respectively — both inactive in today’s universe. Finally, direct detection is most sensitive in the bulk and focus point regions as well because of a greater adherence to crossing arguments in these regions. Thus a combination of all the three approaches is required for sensitivity to all parts of CMSSM parameter space.

Finally, some WIMP dark matter properties will be resolved only with complementary information from different techniques. Colliders might be able to produce LSPs but will be unable to measure its lifetime, its cross sections, etc., and will be unable to constrain the relic density of that LSP. Only direct or indirect detection of particles with the same properties as those seen in colliders will confirm that it is dark matter. Some degeneracies in WIMP models, especially if it is a wino-like WIMP, may also be resolved only if a direct detection measurement provides an order of magnitude estimate of the scattering cross section \[34\]. Finally, a gamma line detection from WIMP annihilation will precisely peg the WIMP mass, which will be useful to narrow searches in both collider and direct searches.
Chapter 3

ZIP Detectors

As I have described in the previous chapter, an effective search for WIMPs requires sensitivity to an extremely small signal rate at low energies while suppressing backgrounds. The success of CDMS as a leading search for particle dark matter stems from its detector technology. CDMS II uses Z-dependent Ionization and Phonon (ZIP) detectors, made of germanium and silicon and operated at 40 millikelvin. The charge and phonon energy from a particle interaction are both measured in these detectors, providing excellent discrimination between electron recoils and WIMP-like nuclear recoils. Events occurring close to the surfaces of solid-state detectors mimic nuclear recoils, but the depth information obtained from the shape of phonons pulses in ZIPs enables rejection of this background as well. Thus ZIPs have enabled CDMS to maintain high sensitivity in WIMP searches.

Germanium ZIPs are expected to provide greater sensitivity to WIMPs compared to the silicon ones, because of the former’s larger atomic mass and thus larger cross section for WIMP-nucleon scattering. However, silicon ZIPs are capable of identifying neutron background based on differences in rates between the two detector types. At the CDMS II installation in Soudan, neutrons have not proved to be a dominant background (see Section 7.1).

3.1 Physical Description

ZIP detectors are 7.6-cm-diameter, 1-cm-thick cylindrical semiconductor substrates of germanium or silicon. On average, the germanium detectors have mass of $\sim$ 230 g and the silicon ones have mass of $\sim$ 110 g. The dimensions, and hence masses, of all detectors are slightly different because of variations in the degree of polishing of individual crystals, but they have been carefully recorded for every detector. The crystals have five “flats” on the cylindrical wall to help position them in their copper housings and to mark the lattice orientation. A picture of a ZIP is shown in Figure 3.1 and its dimensional layout is shown in Figure 3.2.

The flat faces of the detector are photolithographically patterned with sensors that provide the signal used to identify and characterize particle interactions in the substrate. One of the faces has
Figure 3.1: A Z-dependent Ionization and Phonon (ZIP) detector. The squares visible on the surface are photolithographically deposited groups of phonon sensors. *Courtesy:* The CDMS Test Facility at UC Berkeley.

Figure 3.2: Geometry of ZIP detector as seen from top. *Taken from:* [89].
four quadrant-shaped phonon sensors and is called the “phonon side” or “phonon face.” This face is visible in Figure 3.1. Each phonon sensor consists of 37 repeatedly tiled networks of 28 phonon sensing elements each. The other flat face, called the “charge side” or “charge face,” consists of a thin-film aluminum grid, divided into two concentric ionization electrodes. One electrode is a disc covering \( \sim 85\% \) of the area of the face (“fiducial”) and the other is a ring around the first electrode (“outer or guard ring”). The layout of the ionization electrodes is shown in Figure 3.3. Thus, particle interactions in ZIPs are characterized by the energy and timing of signals they generate in the phonon and ionization sensors.

![Figure 3.3: Layout of thin-film aluminum grid for the fiducial and outer ring charge electrodes.](Taken from: [89]).

For the purposes of discussions in this dissertation, the flat faces of a ZIP contain the x-y coordinate plane when referring to locations in or on the detector. The z axis is the cylinder axis, perpendicular to the two flat faces. Thus the z coordinate is used to refer to the depth of interactions in the ZIP.

### 3.2 Ionization

The ionization readout for ZIPs is modeled on that of a traditional Si charge tracker or high-purity Ge detector. Subtleties arise, however, because of subkelvin operation temperature and the application of low-strength electric fields. Extensive information on the charge readout in CDMS is provided in [89] and [71].
3.2.1 Physics

A particle impact in a ZIP deposits energy in the electron system of the semiconductor, liberating valence electrons into the conduction band along the particle track. Depending on the amount of energy deposited, the primary electrons generated in this process can cause secondary ionizations as well. This results in a population of electrons and holes along the particle track, proportional to the energy deposited. On average, 3 eV of energy is required to generate an electron-hole pair in germanium and 3.8 eV in silicon. This is higher than the band gap for Ge (0.74 eV) or Si (1.1 eV) at 0 K, because a large fraction of the energy is dissipated into the phonon system during electron scattering. Thus, ~ 300 electron-hole pairs are generated for every keV of energy deposited in the electron system of the Ge crystal.

An electric field is generated between the flat faces of the ZIP to collect the electrons and holes generated by a particle interaction. This is done by applying a small voltage bias (+3 V for Ge, +4 V for Si, in CDMS II) to the ionization electrodes on the charge face with respect to the sensors on the phonon face. In the absence of this field, the charge carriers would just diffuse until they recombine with one another or get “trapped” in localized states in the band gap. The drifting electrons and holes induce image charges on the electrodes, which are read out using charge sensitive amplifiers described below. If the electrons and holes drift all the way to their respective collection electrodes, the charge read out by the amplifier is equal to the charge created by the event.

If a carrier drifts only partway across the detector, the observed charge will be reduced. This can occur because of recombination or trapping, the likelihood of which is controlled in part by the magnitude of the electric field. The fields used in CDMS II ZIPS are well above the minimum to ensure complete charge collection for well “neutralized” crystals with low trap concentrations.  

The crystals used in CDMS have donor/acceptor impurity concentrations of ~ 10^{11} impurities/cm^3. However, even at these low impurity concentrations, charged trapping centers have high trapping cross sections as they remain ionized in the absence of free charge at the millikelvin temperatures of ZIPs. This is remedied by grounding the ZIPS and exposing them to gamma radiation from radioactive sources and to infrared light from LEDs mounted in the detector casing during cool down at the start of a cryogenic run. A large number of free charge carriers is generated in this process, some of which combine with the trapping sites, reducing their trapping cross section. The “neutralized” configuration remains stable only for time periods of ~ O(10) hours, as drifting charge from particle interactions reionizes the traps. After the initial neutralization during cool down, data taking is periodically paused to expose the detectors to LED photons and maintain the neutralized state, well before charge degradation occurs. This process reduces data-taking duty cycle by only ~ 5–10%, and ensures data-taking with full charge collection.

\footnote{Missing valence electrons}
3.2.2 Charge amplification and readout

The signals produced on the fiducial and outer-charge electrodes are read out by a custom-designed low-noise transimpedance amplifier set up, described in great detail in [90], and pictured in Figure 3.4. The amplifier is configured in negative feedback mode. Thus, charge collected at the detector produces a voltage signal at the amplifier output through the feedback circuit. A coupling capacitor separates the DC voltage biasing of the detector from the AC image current induced on the detector by charge transport. The charge collection process takes under \(\sim 1\ \mu s\). This leads to a virtually instantaneous risetime seen by the digitization electronics of CDMS II (0.8 \(\mu s\) per sample), but the pulse falltime is \(\sim 40\ \mu s\), set by the feedback resistor, \(R_{fb} = 40\ \text{M}\Omega\) and its parasitic capacitance, \(C_{fb} = 1\ \text{pF}\). Thus the charge pulses from this system have a fixed shape, and vary in amplitude proportional to the charge collected by the electrodes. The transimpedance of this system is given by:

\[
A(\omega) = \frac{R_{fb}}{1 + j\omega R_{fb} C_{fb}}
\]

where \(\omega\) is the frequency of a Fourier component of the signal.

The voltage noise in this system comes primarily from the amplifier’s first-stage JFET and from the feedback and bias resistors. Other sources of noise include the current noise from the JFET, detector leakage current and microphonic effects in wiring. A theoretical discussion is presented in [90]; here I only reproduce the measured noise spectrum for a CDMS II ZIP in Figure 3.5, which matches the theoretical prediction of the summed JFET noise and dissipative noise. The model is dominated by \(0.5\ \text{nV}/\sqrt{\text{Hz}}\) noise from the JFET. Peaks are observed in measured spectrum because of electrical pickup and microphonic resonances, which are worse for some detectors than others.
Figure 3.5: Fiducial charge noise spectrum for a typical ZIP in CDMS II, referred to the FET gate and overlaid with noise models for the JFET and dissipative Johnson noise. The observed spectrum matches the model well. Plot from: [72].

3.2.3 Event reconstruction

Charge pulses are digitized and recorded by the CDMS II DAQ, described in Chapter [5]. Since a pulse has a fixed shape set by the electronics and the noise is easily characterized by recording randomly triggered pulse-free traces, optimal filtering provides an unbiased estimate of the true signal amplitude in a noisy trace. The optimal filtering algorithm used in CDMS is described in detail in [71] and [72]. The charge pulses from both the fiducial and charge electrodes are simultaneously fit to templates that estimate and remove the cross talk ($\sim 6\%$) between the two electrodes. In addition to providing estimates for pulse height, the optimal filter fits provide the start time for the pulses. This process is carried out during first-tier data processing, as described in Chapter [6].

3.2.4 Energy calibration

After optimal filtering, the pulse height estimates are stored in arbitrary units. In second-tier processing, these are converted to units of energy (keV). This calibration is performed using large $^{133}$Ba datasets with several million events per detector. The process is described in detail in [93], but I summarize it here. First, any residual cross talk ($\sim 1\%$) between the two ionization electrodes is removed by diagonalizing their correlation matrix. Next, a position dependence of the charge signal collection across the crystal is removed by a linearizing transformation. Figure 3.6 shows this dependence in the 356-keV spectral line in $^{133}$Ba-calibration data, plotted as a function of $y_{del}$, a $y$-event coordinate estimator described in Section 4.1.4. The exact cause for this position
dependence is not understood quantitatively, but is found to be correlated with the ion-implantation gradient across the detector. Next, the overall energy scale for the fiducial electrode is set for each detector by the 356-keV spectral line from $^{133}$Ba, as shown in Figure 3.7. Finally, the calibration of the outer electrode is set using 356-keV events shared by it and the fiducial electrode. There are insufficient high-energy events contained entirely in the outer electrode to do this calibration using outer-electrode events alone. Note that for silicon detectors, the 356-keV spectral line is not visible because of lower photon-scattering cross section. We calibrate these detectors by using 356-keV events shared between them and neighboring germanium detectors.

![Figure 3.6: Optimal-filter peak-height estimate (arbitrary units) for fiducial charge energy vs. y-coordinate estimator ($\mu$s) for $^{133}$Ba-calibration events for a typical germanium detector. The 356-keV spectral line is fit to a polynomial to remove the position dependence. Courtesy: Kyle Sundqvist.](image)

### 3.3 Phonons

ZIPs rely on superconducting thin-film sensors to collect and measure the initial wave of fast athermal phonons after a particle hit, and before that energy is lost to the thermal bath of the detector. In addition to the reconstruction of event recoil energy, athermal phonons contain information about position of events within the detector. This information provides background rejection power for the ZIPs and cannot be acquired from thermal phonons. An exhaustive overview of athermal phonon measurement in CDMS will be provided in Matt Pyle’s dissertation, which is currently under preparation. [94]
3.3.1 Physics

Particle interactions in a ZIP generate three different kinds of athermal phonons, each attributed to a different process. These are explained below:

1. **Primary phonons**: Recoiling nuclei and electrons from a particle interaction dissipate part of their recoil energy $E_R$ in optical and acoustic phonons in the detector. These primary phonons of energy $E_{prim}$ contain information about position, energy, and timing.

2. **Relaxation phonons**: The remaining fraction of $E_R$ drives $N_Q$ electron-hole pairs across the energy gap $E_{gap}$. The electrons and holes hold this energy as they drift across the detector until they reach the electrodes and relax to the Fermi level. This energy is then released into relaxation phonons. The energy of relaxation phonons is $E_{relax} = N_Q E_{gap}$.

3. **Neganov-Trofimov-Luke Phonons**: The work done by a ZIP’s electric field to drift an electron-hole pair across it is dissipated into the athermal phonon system as Neganov-Trofimov-Luke phonons [95, 96]. These are analogous to Cherenkov radiation because they are generated by charge carriers moving faster than the speed of sound in the crystal. Their energy, $E_{luke} = eV_b N_Q$, where $V_b$ is the voltage bias across the detector.
Thus the total energy available in the athermal phonon system, $E_p$, is given by the following:

$$E_p = E_{prim} + N_Q E_{gap} + eV_b N_Q = E_R + eV_b N_Q$$  \hspace{1cm} (3.2)$$

After a particle interaction, nuclei and electrons recoil through the detector crystal and first generate optical phonons at the Debye frequency of $\sim 10$ THz. These phonons can now undergo two processes: elastic scattering off isotopic impurities in the crystal, and spontaneous anharmonic decay into lower frequency phonons. Phonon-phonon scattering and phonon-carrier scattering are absent at the millikelvin temperatures of ZIPs because of a paucity of free carriers and high-energy phonons. Isotopic-scattering cross section $\sigma \propto \nu^{-4}$ and anharmonic-decay cross section $\sigma \propto \nu^{-5}$. This high frequency dependence keeps the mean free path of the initial high-energy phonons small, localizing them to the interaction site. Anharmonic decay rapidly reduces these to 1.6-THz phonons at which point there is a downconversion “bottleneck” — isotopic scattering dominates and the phonons become “quasi-diffusive.” After a few microseconds of slower downconversion, the mean free path of the phonons becomes comparable to the detector size, making them “ballistic,” i.e., they become free to travel through the crystal. This is when primary phonons can first reach the sensors on the detector flat surface. Luke phonons are expected to be generated at ballistic frequencies, while recombination phonons are generated at high frequencies, but rapidly down-convert by interacting with the metal films at the detector flat surfaces. The difference in arrival time of the different kinds of athermal phonons has important implications for characterizing event position in the detector.

### 3.3.2 Phonon detection

As mentioned earlier, one of the flat faces of a ZIP is photolithographically patterned with four quadrant-shaped phonon sensors, labeled A, B, C, and D in clockwise order, starting with the upper left quadrant. Figure 3.8 shows a schematic of the patterning of phonon sensors on a detector face. Each sensor consists of 1036 superconducting tungsten thin-film sensors called Transition-Edge Sensors (TES), wired in parallel, but divided into 37 tiles of 28 TESs each. Each TES is fed by superconducting aluminum fins that collect phonon energy and concentrate it in the much smaller TESs. The aluminum absorbers and TESs are together called “QETs”: Quasiparticle-trap-assisted Electrothermal-feedback Transition Edge Sensors. These are subsequently read out by SQUID-array amplifiers as explained in the next section.

#### 3.3.2.1 Absorber fins

Phonons are collected in absorber fins made of aluminum, 350 $\mu$m long, 50 $\mu$m wide, and 300 nm thick. Since the crystal temperature is well below superconducting transition of 1.2 K for aluminum, the energy gap to break cooper pairs into quasiparticles is high, $2\Delta_{Al} = 360 \mu$eV. Athermal phonons
Figure 3.8: QET layout on a ZIP. Upper left: Phonon side of ZIP showing four quadrant-shaped phonon sensors and, their labels A, B, C and, D. Each phonon sensor is fabricated by tiling 37 5 mm × 5 mm templates (upper right), each with 28 QETs in grey and surrounding aluminum grid in green. Bottom: Zoom in of a single QET with aluminum absorbers in grey color and tungsten TES in blue. *Figure taken from:* [89].
are typically sufficiently energetic to do this and deposit their energy in the quasiparticle system, whereas thermal phonons (energy \( \sim kT = 3.4 \mu eV \)) are unable to do so. The quasiparticles generated by the phonons diffuse through the aluminum until they recombine or find their way into the tungsten TESs. There is a region of overlap between the Al and W, where the superconducting gap transitions from that of Al to that of W. Tungsten has a lower critical temperature and hence a lower energy gap for cooper-pair breaking, \( 2\Delta_W \approx 24 \mu eV \). Quasiparticles entering the W TESs quickly lose energy and fall below the high energy gap of Al, becoming unable to diffuse back into the Al. Thus the interface of the Al absorber and the W TES serves as a quasiparticle trap, concentrating energy from a large collection area into the TES. Despite micron-scale features, 28 QETs alone are therefore able to collect phonon energy from a 5 mm \( \times \) 5 mm area of the detector flat surface. A schematic of the phonon absorption and quasiparticle trapping in the QETs is shown in Figure 3.9.

![Schematic of QET](https://example.com/schematic.png)

Figure 3.9: Schematic of QET. Phonons entering the aluminum absorber break cooper pairs, generating quasiparticles. The quasiparticles diffuse through the aluminum and are trapped in the tungsten TESs.

### 3.3.2.2 Transition-Edge Sensors

Superconducting Transition-Edge Sensors (TESs) are very sensitive, high-bandwidth variable resistors that change resistance with temperature, and can have extremely small feature size — 1 \( \mu \)m wide, 250 \( \mu \)m long and 35 \( \text{nm} \) thick in CDMS II ZIPs. They are operated at their critical temperature \( T_c \), partway through their superconducting transition. This transition is typically very sharp, causing a large change in resistance for a small change in temperature, as long as the TES equilibrium temperature is within the narrow window of the transition. This makes the TES a precision measurement tool for small inputs of energy. A detailed review of TESs is provided in [97].

TESs need to be held very close to \( T_c \) and need to repeatably return to the initial bias point after temperature excursions caused by energy deposition. In CDMS, this is accomplished by voltage biasing the TESs. The power flowing into a TES’s electron system at temperature \( T_c \) is a combination of Joule heating by current flowing through it (\( P_J = V_{bias}^2 / R_{TES} \)), and any external power loading, i.e., quasiparticle energy introduced in it the absorbers, \( P_{ext} \). The power flowing out is through the weak thermal link \( G_{cp} \) between the TES electron system and the TES phonon system. The latter
of these is tightly coupled to the experiment’s heat bath at temperature $T_0$ via heat sinking of the
detectors to the dilution fridge. Thus the following relation holds:

$$\frac{V_{bias}^2}{R_{TES}} + P_{ext} \simeq G_{ep}(T_e - T_0) \quad (3.3)$$

The key for proper functionality of the TES is to have a suitably low $G_{ep}$ and a carefully selected
$V_{bias}$ such that the TES electron system self-heats to an equilibrium temperature equal to the
superconducting transition point. Then, upward fluctuations in $P_{ext}$ would increase $T_e$ and hence
$R_{TES}$, but would cause a drop in Joule heating and eventually restore the TES back to its equilibrium
temperature. This is called negative electrothermal feedback, and enables the TESs to be operated
stably. This scheme works not only for single TESs, but for the parallel TES arrays of a ZIP phonon
sensor. Even if there are slight variations in $T_c$ of individual TESs, a single voltage bias allows all
TESs to self-heat to appropriate equilibrium temperatures within the transition. This does, however,
soften the sharpness of the transition and hence the resolution of the measurement.

Note that CDMS TES films are deposited with a target $T_c \sim 120$ mK but end up with 10–20%
variations. The TESs are implanted with Fe ions after initial fabrication to uniformly tune the $T_c$
closer to 80 mK [98].

3.3.3 Phonon amplification and readout by SQUIDs

Since the TESs are voltage biased, a change in their resistance caused by energy deposition in the
detector results in a small change in current in the biasing circuit. Figure 3.10 shows a schematic of
the phonon readout scheme used in CDMS II. An inductive coil in series with the TESs of a phonon
sensor translates changes in current to changes in magnetic flux. This coil, called the input coil,
is coupled to a DC SQUID array, essentially a very sensitive magnetometer. The SQUID array is
connected to a voltage amplifier whose low-pass filtered output drives a second inductor in negative
feedback mode. Changes in flux in the SQUID are countered by an opposing flux in the second coil,
called the feedback coil. The feedback coil has 1/10 the number of turns as the input coil, requiring
the amplifier to respond to any changes in input coil current with a 10x larger current.

A full treatment of TES and SQUID noise is beyond the scope of the discussion here; details are
provided in [97]. For our purposes, the noise in the phonon readout is dominated by the Johnson
noise of the shunt resistor, which is located at a temperature stage of 600 mK$^2$. The TES and SQUID
contributions are suppressed in comparison. The noise spectral density for a phonon channel of a
typical ZIP is shown in figure 3.11 and is consistent with the expectation of 15 pA/$\sqrt{\text{Hz}}$ shunt-
resistor-dominated noise.

$^2$The shunt resistor could be located at 40 mK to lower its Johnson noise, but the thermal load on the dilution
fridge becomes too high for this configuration
Figure 3.10: CDMS phonon amplifier schematic for ZIPs. $R_{TES} \sim 200 \text{ m}\Omega$, $R_{sh} = 25 \text{ m}\Omega$, $R_{fb} = 1200\Omega$, and $L_i = 250 \text{ nH} = 100L_{fb}$. Figure taken from: [72].

Figure 3.11: Phonon noise spectrum for channel A of a germanium ZIP, referred to the amplifier input coil in Figure 3.10. Plot from: [72].
3.3.4 Event reconstruction and energy calibration

Unlike the charge pulse, the phonon pulse does not have a fixed shape. Athermal phonons in ZIPs preserve the position and time-of-arrival information from a particle interaction, in addition to information of recoil energy. All this information is entangled and appears as significant variation in the amplitude and shape of phonon pulses with event energy and position. Thus, the energy of the event is not accurately extracted from a pulse by simply applying an optimal filter to it; this would lead to systematic errors in energy estimation. We are also interested in extracting the event-depth information, as this enables discrimination of bulk events from surface events. As explained in Section 3.4.3, surface events mimic nuclear recoils and are the largest source of undesirable background in ZIP detectors. While position information is contained in the ZIP’s athermal phonon pulses, this information is a complicated function of three-dimensional position information. Extracting the depth information alone is not trivial. Finally, the same issues impact phonon energy-scale calibration. The optimal-filter energy estimate from phonon pulses results in a non-linear energy scale with position dependences.

The complications caused by phonon-pulse-shape variation are mitigated by an empirical correction. We first make naïve estimates of energy, pulse-arrival time, and pulse-shape characteristics using optimal-filter and time-walk algorithms. We then use a large $^{133}$Ba-calibration dataset, with tens of millions of photon events, to characterize the phonon-pulse-shape response detector-by-detector, as a function of the naïve energy and position estimates. The undesired variation observed in x-y position and energy is subtracted out, leaving energy and event-depth estimators with better resolution. I dedicate all of Chapter 4 to explaining the technical details of phonon event reconstruction, energy calibration and empirical phonon-pulse-shape correction.

3.4 Background Discrimination in ZIPs

As discussed in Section 2.3.3 of Chapter 2, a well-designed rare-event search should have the ability to discriminate background from signal with high significance. Residual backgrounds limit sensitivity. In CDMS, the ZIP detector was designed with this in mind. A measurement of ionization and phonon signatures of particle interactions allows background electron recoils to be discriminated from nuclear recoils. The position information encoded in the athermal phonon pulse enables the rejection of surface electron recoils, which can mimic nuclear recoils.

3.4.1 Primary discrimination with Ionization Yield

Particles interact with electrons and nuclei of the medium in which they traverse, and lose energy during these interactions. The rate of energy loss to electrons and that to nuclei depends on the
charge, mass, and energy of the incident particle. Light, fast particles such as electrons shed most of their energy to the electron system in long tracks with sparse interactions. Slow, heavy particles such as nuclei are less efficient at ionization. In addition, they have shorter tracks with greater energy-deposition density. This results in more energy deposition in the phonon system of the crystal lattice as well as more electron-hole recombination at the interaction site, moving energy from the electron-hole system to the phonon system. We have exploited this difference in energy-loss rate to provide discrimination in ZIPS. Photon or electron backgrounds are expected to deposit more energy in the electron system of the crystal lattice while neutrons or WIMPs are expected to deposit less energy in the electron system. In other words, electron recoils will have a larger ionization energy compared to nuclear recoils, for the same recoil energy. Thus, we use ionization yield as our primary discrimination parameter to reject electron recoil background:

\[ y = \frac{E_Q}{E_R} = \frac{E_Q}{E_P - \frac{\epsilon}{\epsilon}E_Q} \]

where \( \epsilon \) is the average electron-recoil energy needed to generate an electron-hole pair, \( E_Q \) is the “charge energy” such that \( y = 1 \) for electron recoils. Figure 3.12 shows the discrimination power of ionization yield. Yield “bands” for electron recoils and nuclear recoils are defined using a sample of photons from \(^{133}\)Ba-calibration dataset and a sample of neutrons from a \(^{252}\)Cf-calibration dataset. The edges are \( \pm 2\sigma \) around the mean ionization yield for that sample. The residual leakage of bulk photons and electrons into the nuclear-recoil band is better than \( 10^{-4} \) per detector per nuclear recoil at as low a threshold as 5 keV.

### 3.4.2 Surface Events in dead layer

There is a population of electron recoils visible in Figure 3.12 that has ionization yield lower than that expected for electron recoils. About 10% of these events appear in the nuclear-recoil band and pose the most significant background risk for CDMS. These events are attributed to electron recoils that occur in the “dead layer,” the first \( \sim 10 - \mu m \)-thick slice of the top and bottom flat faces of a ZIP. Dedicated measurements have shown that ionization yield is suppressed for these “surface events” [99, 100]. It is believed that charge carriers can back-diffuse into the electrode closest to them, regardless of polarity, if the particle interaction occurs very close to an electrode. Carriers diffusing into the “wrong” electrode relax to the Fermi level before they can be influenced by the drift field across the detector, leading to a loss of collected charge at the correct electrode. It is also possible that ion implantation of the TESs or other surface treatments during the fabrication process introduce a large number of defects and trapping sites, capable of capturing charge carriers from surface events.

The dead-layer problem is mitigated to an extent by depositing a layer of amorphous silicon on
Figure 3.12: Ionization yield as a function of recoil energy for calibration data for a typical Ge ZIP. The blue points are nuclear recoils from $^{252}$Cf neutrons. The red points are photons and Compton-scattered electrons from a $^{133}$Ba source. The dashed lines represent the $\pm 2\sigma$ yield “bands” for electron recoils (top) and nuclear recoils (bottom). There is a 15\sigma separation between the mean ionization yields of nuclear recoils and electron recoils.

The crystal surface before the aluminum for the electrodes is deposited [101]. This introduces a larger energy gap between the electrode and the diffusing charge particle of wrong polarity, decreasing the likelihood of back-diffusion. However, the dead layer is not eliminated entirely, leading to the residual low-yield surface events seen in Figure 3.12.

3.4.3 Surface-event rejection

An alternate and final line of defense against surface events is the timing and position information from ZIPs afforded by the athermal phonon signal. Recall that high-frequency phonons from particle interactions in the bulk of the detector suffer from a downconversion bottleneck at 1.6 THz, when isotopic scattering keeps them in a quasi-diffusive state. Thus it takes on the order of $\sim 5\text{–}6\mu$s before they downconvert sufficiently to become ballistic and propagate to the phonon sensors. On the other hand, particle interactions close to the flat surfaces of a ZIP downconvert by interactions with the metal thin-films, shortcircuiting the bottleneck. This gives surface event phonon pulses faster arrival times and slopes.

A surface event discrimination parameter (or timing parameter) is constructed from phonon-pulse timing characteristics, usually as the sum of the arrival time of the tallest phonon pulse (out of 4 pulses for an event) and the time for that pulse to rise from 10% to 40% of its maximum amplitude$^3$. Before this parameter is useful for discrimination, it must be corrected to remove

$^3$The technical definition of this parameter is provided in Section 4.1.3.
position dependences that tend to wash out intrinsic discrimination. This is discussed at length in Chapter 4. Histograms of the distributions of the corrected timing parameter are shown for $^{133}$Ba-calibration-induced surface events and for $^{252}$Cf-neutron-induced nuclear recoils in Figure 3.13. A typical cut based on this timing parameter might be tuned for a residual surface event leakage of $5 \times 10^{-2}$ per detector per nuclear recoil above 7–10 keV.

![Histograms of timing parameter distributions for $^{133}$Ba-induced surface events (SE) and neutrons from $^{252}$Cf (NR) after removal of position and energy dependences.](image)

Figure 3.13: Histograms of timing parameter distributions for $^{133}$Ba-induced surface events (SE) and neutrons from $^{252}$Cf (NR) after removal of position and energy dependences.

### 3.4.4 Combined discrimination

The ionization-yield-based discrimination and phonon-timing-based discrimination are combined to provide maximal rejection of electron-recoil background. A lower limit on the residual bulk-electron-recoil rate achieved in ZIPs is $\sim 10^{-6}$ per detector per nuclear recoil. Typical surface event residual rates are $\sim 5 \times 10^{-3}$ per detector per nuclear recoil. Figure 3.14 shows the separation achieved between nuclear recoils and electron recoils for calibration data. This promising ZIP-based discrimination, combined with an underground experimental site and passive shielding (see Chapter 5) makes CDMS a competitive experiment in the search for WIMPs.
Figure 3.14: Ionization yield vs. timing parameter for calibration data, with recoil energy 10–100 keV, taken with a typical Ge detector. Bulk electron recoils from a $^{133}\text{Ba}$ source are marked with red points. $^{133}\text{Ba}$-source-induced surface events are marked with black crosses. Nuclear recoils from $^{252}\text{Cf}$ neutrons are marked with blue points. Using the combined discrimination of ionization yield and phonon-pulse timing, we can define a nuclear-recoil acceptance region with low electron-recoil leakage.
Chapter 4

Position & Energy Calibration of Phonon signal

The athermal phonons generated by a particle recoil in a ZIP detector depend not only on the energy of the recoil, but on the event’s position in the detector. Unfortunately, the ZIP’s one-sided four-channel phonon readout is insufficient to deconvolve all this information, resulting in pulse shapes that vary with both event energy and position. Energy estimation with a fixed-pulse-shape optimal filter carries systematic errors. Also, event depth information, capable of providing surface-event discrimination, is mingled with x-y position information, diminishing discrimination power.

In this chapter, I explore these issues and explain how they are resolved using an empirical correction of information derived from phonon-pulse shapes. The first section shows how preliminary estimates of pulse energy, timing, and event position are made using naïve methods. The second section describes the shortcomings of these naïve estimates. The third section of the chapter discusses correction of the estimates, which exploits the fact that local regions of a ZIP have similar pulse-shape response and variations from region to region can be subtracted out. In the last section, I cover improvements I made to the correction technique, which made possible the sensitivity achieved in the WIMP-search analysis presented in this dissertation.

4.1 Preliminary Event Reconstruction

Preliminary estimates of phonon energy are made using a fixed-pulse-shape optimal filter and those of pulse shape or timing using a rising-edge “walking” algorithm. These two techniques independently provide preliminary estimates with lowest measurement noise for the two types of measurement, but with significant systematic errors. Nonetheless, the estimates enable preliminary position reconstruction of events. The systematic errors are dealt with in the empirical correction described later.
4.1.1 Energy measurement

The phonon energy deposited by a particle in a ZIP detector is estimated by measuring the height of the resulting pulses from the four phonon channels. For every event in a detector, the pulse height is estimated in first-tier data processing using a fixed-template optimal filter for each quadrant [102]. The template pulse uses a two-exponential functional form, \( A(t) = A_0(1-e^{-t/\tau_1})e^{-t/\tau_2} \), where \( \tau_1 \) and \( \tau_2 \) are characteristic risetimes and falltimes, estimated from several good pulses. Figure 4.1 shows raw traces for an event in a detector with templates overlaid. We have attempted to fit the phonon pulses to functional forms in the time-domain in order to estimate pulse energy [103, 104]. These fits have typically had fit-convergence problems and other irreducible systematic errors, leading to degraded energy resolution compared to that of the fixed-shape optimal filter described above [105, 106]. In addition to the optimal-filter pulse-height estimate, the area of the raw pulse is also integrated and saved as a reduced quantity available for estimating event energy [107]. However, the optimal-filter estimate has better fractional resolution than the integral estimate because it is able to suitably de-weight noisy low-frequency components of the acquired traces. With increasing energy, the integral estimate improves as the signal-to-noise increases, whereas the optimal-filter estimate starts to suffer from systematic errors because of mismatches in true pulse shape and template shape [108].

4.1.2 Preliminary energy-scale calibration

Phonon pulses are first recorded by the DAQ in arbitrary digitizer units. After first-tier data processing, the optimal filter pulse height and phonon-pulse area are still stored in these arbitrary units. In second-tier data processing, these are converted to units of energy (keV), by calibrating with the pre-calibrated ionization energy scale. We use a subset of low-energy \(^{133}\)Ba calibration events, consisting of over 99% low-energy gammas, to determine least-squared scaling factors for each channel, such that the summed phonon energy of all channels, including the contribution from Neganov-Luke phonons, is on average equal to twice the ionization energy \(^{11}\) [109]. The resultant phonon energy for a detector, T1Z5 (Runs 125–128), is plotted in Figure 4.2 as a function of the charge energy. Then, these scaling factors are adjusted so that the distribution of energy partitioning between channels is roughly identical, i.e., the contribution of no one channel can dominate the summed phonon energy. This is shown in Figure 4.3. Note that this process only provides a preliminary calibration at the 10% level, and does not ensure linearity of the scale.

At this point, it is worth clarifying the nomenclature of phonon energy measurements in CDMS analysis. After preliminary energy-scale calibration, we call the total phonon energy measured in a detector, \( p_t \). This is the sum of the phonon energy measurements, \( p_a, p_b, p_c, \) and \( p_d \), corresponding

\[ E_R = E_{\text{Luke}} = E_Q \] for photons. See next paragraph for clarification.
Figure 4.1: Unfiltered event traces for the four phonon channels of detector T4Z2 (digitizer readings vs. time samples). Channel A, with the tallest phonon pulse has 12.3 keV recoil energy. A two-exponential template is used for the signal shape in the optimal filter, and is overlaid on the traces, multiplied by the amplitude estimate from the algorithm. Courtesy: Lauren Hsu.
Figure 4.2: Phonon energy as a function of calibrated charge energy for a random sample of calibration gammas in detector T1Z5 after calibrating their mean ionization yield to 1. The total phonon energy (including that from Neganov-Luke phonons) is twice the charge energy. For WIMP-search Runs 125–128, this preliminary calibration is done only for photons with charge energy between 65 and 100 keV as indicated by the vertical dashed lines. Thus non-linearities are still apparent, and are not corrected until later in the calibration process.

Figure 4.3: Relative calibration of phonon sensors is accomplished by aligning the phonon partition histograms for the four sensors to roughly ~10%. The plot above uses four different colors for the four phonon sensors of T1Z2.
to the four phonon channels. As explained in Section 3.3.1 and summarized in Equation 3.2, the measured phonon energy is a sum of the recoil energy \( E_R \) and the contribution by Neganov-Luke phonons \( E_{\text{luke}} \). This can be written in analysis variables as follows:

\[
pt = pr + q \frac{V_b}{E_{\text{charge}}} \quad (4.1)
\]

\[
\Rightarrow pr = pt - q \frac{V_b}{E_{\text{charge}}} \quad (4.2)
\]

where \( q \) is the calibrated charge energy measured in an event and \( E_{\text{charge}} \) is the energy per coulomb. If a charge fiducial volume cut is applied, then events passing the cut have outer-charge-electrode signal consistent with noise. For such events, the total phonon energy can be written as:

\[
pr_i = pt - qi \frac{V_b}{E_{\text{charge}}} \quad (4.3)
\]

For photons, ionization yield, \( y = 1 \), according to Lindhard theory \[110\], implying \( q = pr \) or \( qi = pr \), in the case of a charge-fiducial-volume cut. We define a quantity called photon-equivalent recoil energy, \( pr_g \), defined only using \( pt \). This quantity has the advantage of being a less noisy estimator of recoil energy for photons than \( pr \). It can be written as follows:

\[
pr_g = pt - pr_g \frac{V_b}{E_{\text{charge}}} \quad (4.4)
\]

\[
\Rightarrow pr_g = \frac{pt}{1 + \frac{V_b}{E_{\text{charge}}}} \quad (4.5)
\]

Photon-equivalent yields can also be defined: \( yg = q / pr_g \) and \( ygi = qi / pr_g \). These photon-equivalent quantities turn out to be useful for phonon-pulse-shape correction, since they are not affected by charge noise.

### 4.1.3 Timing measurement

Phonon-pulse-shape characteristics help distinguish background surface events from nuclear-recoil signal. Figures of merit for characterizing the pulse shape are constructed during first-tier processing in the following way. Phonon pulses are first low-pass filtered using a 50 kHz Butterworth filter. Then an algorithm “walks” down the rising edge of the filtered pulse, and records the time corresponding to the first-crossing point in pulse amplitude at some fraction of the pulse height, say 20% \[111\]. These are called “risetimes.” The same algorithm is also applied on the falling edge of the pulse to provide “falltimes,” but falltimes do not provide surface event discrimination in ZIP detectors. The risetimes can be compared with the charge-signal start time, from the ionization optimal filter, to provide a phonon-pulse arrival “delay.” A schematic of the walk algorithm applied to traces from a calibration event is provided in Figure 4.4.
Figure 4.4: Application of walk algorithm. The first row displays the charge trace and marks its start time. The second through fifth rows show the phonon traces for the four channels in black, and the 50-kHz-Butterworth filtered traces in blue. The second column provides a zoomed-in view of the traces displayed in the first column. For each phonon trace, the 20% risetime determined by the walk algorithm is marked with a red cross. At low energy, this algorithm suffers from poor signal-to-noise, leading to misestimated timing parameters and poor resolution. This is seen for channel A in the trace above. Courtesy: Scott Hertel.
Risetimes and delays can be constructed for all phonon quadrants, but those for the primary phonon quadrant, i.e., the one with maximum pulse height, and hence maximum signal-to-noise among the four channels are particularly important:

1. **Primary Phonon Risetime**: The difference of the 40% and 10% risetimes of the primary phonon pulse is called primary phonon risetime, $p_{\text{minrt}}$, and is a powerful discriminator between surface events and nuclear recoils.

2. **Primary Phonon Delay**: The difference in the charge signal start time as computed by the charge optimal filter and the 20% risetime of the primary phonon pulse, is called primary phonon delay, $p_{\text{del}}$, and is also a key discriminator against surface events in WIMP-search running.

Both primary risetime and primary delay are marked on a typical event trace in Figure 4.5. Several other timing discriminators have been constructed and tested but the two above have provided most surface-event rejection in past WIMP searches and were used in this search as well.

![Figure 4.5](image-url)

**Figure 4.5**: Traces from a high-energy event for all four phonon channels and the two charge channels are shown as a function of time. The primary channel, based on pulse amplitude, is A. Thus, its 10%–40% risetime is the primary phonon risetime, and the difference of its 20% risetime and the charge-pulse start time is the primary phonon delay. *Courtesy: Jeff Filippini.*

The walk algorithm has the advantage of being model-independent, but is susceptible to noise at recoil energies lower than 10–15 keV on most ZIPs, as seen for a trace (channel A) in Figure 4.4. An experimental algorithm was tested that varied the Butterworth filter frequency to match pulse size. Smaller pulses were filtered with a lower roll-off to mitigate the effects of lower signal-to-noise.
Unfortunately, this introduced a systematic increase in risetime at low energies that could not be corrected trivially. Thus the modified walk algorithm was not used for the WIMP-search analysis. Time-domain fits using simple pulse shapes are comparable in performance to “walked” timing quantities, and trade low-noise performance for systematic errors. Two pulse-shape kernels, called PipeFit and WedgeFit, were included in the first-tier data processing package to produce fit-based risetimes and delays in addition to the ones from the walk algorithm. These algorithms are described in Section 6.2.6 of [113].

### 4.1.4 Position reconstruction

The x-y position of an event and its depth information are estimated using different handles, which I describe below separately.

#### 4.1.4.1 X-Y event position

Using energy and timing information from the four phonon quadrants, we can glean approximate x-y position information for an event. We construct two sets of position estimators, one using the energy information, and one using timing.

1. **Partition:** The phonon energy distribution between channels is expected to depend on event location. More energy deposited in a quadrant compared to another makes it more likely for the event to have occurred closer to the former. The x position is estimated by comparing the summed phonon energy in quadrants A and B with that of C and D, and y position is estimated by comparing the summed phonon energy in quadrants A and D with that of B and C. This can be written as follows:

   \[
   \begin{align*}
   x_{\text{part}} &= \frac{(pc + pd) - (pa + pb)}{pa + pb + pc + pd} \\
   y_{\text{part}} &= \frac{(pa + pd) - (pb + pc)}{pa + pb + pc + pd}
   \end{align*}
   \]

   When these estimators are plotted, they acquire a square shape [115]; this figure is commonly known to CDMS collaborators as a “box plot.” It is shown in [4.6] for a subset of calibration gammas from a $^{133}$Ba source. Despite the reconstructed shape, a partition radius can be defined as $r_{\text{part}} = \sqrt{x_{\text{part}}^2 + y_{\text{part}}^2}$.

2. **Delay:** The phonon pulse is expected to arrive sooner at a quadrant that the event is close to rather than at one it is farther away from. Thus, we can estimate a delay-based position for an event by taking the risetime (we use 20% risetime) of the largest of the four phonon
pulses, say in channel A, and subtracting it from the risetime of the two adjacent quadrants as follows:

\[ x_{\text{del}A} = \text{PArt20} - \text{PDrt20} \]
\[ y_{\text{del}A} = \text{PBrt20} - \text{PArt20} \]

The resulting “delay plot” is shown in Figure 4.6 for the same subset of calibration events used above. In analogy to the partition radius, a delay radius can be defined as \( r_{\text{del}} = \sqrt{x_{\text{del}}^2 + y_{\text{del}}^2} \).

Figure 4.6: Left: Event coordinate reconstruction using partitioning of energy between phonon sensors. Right: Reconstruction using relative arrival times of phonon pulses in the four sensors. In both plots, events with energy in the outer-charge electrode are marked with red crosses. They show up at smaller radius than expected, indicating a reconstruction degeneracy. This is discussed in Section 4.1.5.

4.1.4.2 Z-position information

The success of ZIPs lies in the ability to reconstruct event depth information from phonon pulse timing. The quantities \( p_{\text{minrt}} \) and \( p_{\text{del}} \) show discrimination between surface and bulk events. As explained in Section 3.4.3, the difference in timing between surface and bulk events occurs because of faster downconversion of high-frequency phonons for events in contact with the metal films of the surface. This was first shown in a prototype silicon ZIP, where non-penetrating betas from a $^{14}$C source were found to have faster phonon risetimes than 60-keV bulk gammas from a $^{241}$Am source and neutrons from a $^{252}$Cf source [116]. Subsequently, one of the first germanium ZIPs, called G31, was irradiated at UC Berkeley with non-penetrating electrons at 62 and 84 keV from a collimated
$^{109}$Cd source. The observed events had faster timing than neutrons from a $^{252}$Cf source, as shown in Figure 4.7. 

![Figure 4.7](image)

Figure 4.7: Primary phonon risetime (left) and primary phonon delay (right) vs. ionization yield for calibration data taken at UC Berkeley for detector G31. Both parameters show discrimination between $^{252}$Cf-induced neutrons (blue dots) and $^{109}$Cd surface events (green dots). Courtesy: Bruno Serfass and Jeff Filippini.

Event depth information is also encoded to some extent in the energy partition. Particle interactions close to the instrumented side of the detector will deposit a larger fraction of their energy in one sensor than those deeper within the detector. This information is, unfortunately, not as easily extracted from the energy partition, but can be qualitatively demonstrated after a discussion in the next section of the “phonon manifold” and reconstruction degeneracies.

### 4.1.5 Reconstruction degeneracies

If the events in the outer charge channel are tagged in the box and delay plots, as shown in Figure 4.6, one sees that the position reconstruction explained above does not estimate event radius monotonically. Events of low and high radius can have degenerate reconstructed radii. Fortunately, the two x-y reconstruction techniques can be combined to break this degeneracy, because the location of the degeneracy is different for partition compared to delay. In Figure 4.8, we combine the x- and y-partition coordinates with delay radius to generate a three-dimensional bowl-shaped “phonon position manifold.” One can divide the bowl-shaped manifold into pie-like slices in azimuth. Such a slice can then be projected on a coordinate plane of partition radius and delay radius, and is shown in Figure 4.9. The true radius of an event can thus be estimated by walking along the shrimp-like shape (“shrimp”) in that figure.
Figure 4.8: The x-y reconstruction degeneracies are broken by combining information from both the partition and delay information. We take the box plot (x and y axes) and project the delay radius on the z axis. The resulting 3-D manifold is a bowl-shaped structure with an curved lip.

Figure 4.9: A radially-symmetric 20° slice through the 3-D phonon position manifold of Figure 4.8 yields a “shrimp”-shaped Figure on a delay radius (µs) vs. partition radius plot. Walking along the body of the shrimp indicates the radius of an event in the detector.
As mentioned earlier, the depth information is harder to extract. The G31 calibration dataset from UC Berkeley showed that changing the detector face exposed to the $^{109}$Cd source, altered the location of the surface-event population on the shrimp plot. Specifically, the population moved in partition, as expected. Figure 4.10 shows the full 360° span of the 3-D phonon-position manifold, projected on partition-radius and delay-radius axes for the two source configurations.

Figure 4.10: Phonon position manifold for G31 calibration data in radial coordinates: delay radius ($\mu$s) vs. partition radius. The $^{109}$Cd source is placed on the phonon-sensor-instrumented side on left pane and on the the uninstrumented side on the right pane. The location of the surface-event population, marked with green points, is different in partition depending on the detector face that is exposed to the source. Photons and neutrons are marked with red and blue points, respectively.

Looking at timing also leads to similar ambiguity. If we take the shrimp plot for calibration gammas and color its data points with pulse timing, say primary phonon risetime, as in Figure 4.11, we observe z-dependence across partition at low radius. The faster events (deeper blue on the plot), closer to the detector faces constitute the edges of the shrimp and the slower events (cyan) form the center of the shrimp. At high x-y radius, the shrimp shows a systematic shift to slower timing, and washes out timing-based depth deconvolution. This is an artifact of phonon-pulse-shape variation, which if left uncorrected would substantially diminish the ability of ZIPs to distinguish bulk events from surface events. The rest of this chapter is devoted to developing an understanding of this problem and discussing empirical methods to correct for it.

### 4.2 Phonon-Pulse-Shape Variation with Position and Energy

The phonon pulse shape varies with position of events for several reasons [117, 118]. The largest x-y variations in pulse shape occur because phonon reflections off detector edges and walls alter the
distribution and arrival times of phonons between sensors. High-radius events suffer more reflections before being absorbed by QETs. Also, QET coverage is better at low radius than at high radius, leading to slower reconstructed times for high-radius events as noted in Figure 4.11. There is some variation of phonon pulse azimuthally as well, because QET coverage varies azimuthally. The variation of pulse shapes with depth is a design feature, exploiting faster phonon downconversion for events close to the detector’s metalized surfaces.

The variation in pulse shape due to depth is hard to deconvolve from the x-y position-induced variation because ZIPs have only four phonon sensors and only on one face of the detector. Furthermore, they are divided azimuthally in quadrants, with no radial division to aid the breaking of radial reconstruction degeneracies. Imaging of events is constrained by limited information, where some but not all degeneracies can be broken. Future generations of ZIPs, discussed in Appendix B, have resolved these issues, but they remain in CDMS II ZIPs.

The phonon pulse shape varies slightly with energy too. In Figure 4.12 we see a systematic increase of the mean ionization yield of photons with energy, implying that we systematically underestimate the fixed-template optimal-filter pulse height with increasing energy. This is attributed to TES saturation effects. As the amount of energy in a particle interaction is increased, more TESs directly above the interaction site saturate, causing the signal to be dominated by TESs farther away from the event, slowing down pulse timing.
Second-order variations in pulse shape occur because our photolithography process has not provided consistent TES transition temperatures across all TESs in a channel, or between channels. This leads to variability in phonon pulse shapes and heights for identical events in different detectors, or events of same energy, depth, and radius in the same detector.

Figure 4.12: Plot of preliminary yield vs. preliminary recoil energy for $^{133}$Ba calibration gammas in a detector. The definition of recoil energy used in this plot assumes that the events are photons and so their ionization energy should be equal to their recoil energy. Note that the ionization yield increases with recoil energy, instead of staying constant at 1, implying an undermeasurement of recoil energy with increasing pulse height. The fitted mean of the photons is marked with a red dashed line. For comparison a green solid line is shown at ionization yield of 1.

4.3 Empirical Correction of Pulse-Shape Variation

In the absence of a quantitative model for phonon propagation and collection in ZIPs, the only way to correct for pulse shape variations is to do so empirically. By recording the response of the detector to known particle interactions with a range of energies and position, pulse shape variation can be calibrated out.

4.3.1 Basic principle

Despite the variation of phonon pulse shapes with energy and event position in a detector, it turns out that phonon pulses for events of similar energy and of a particular interaction type (electron recoil or nuclear recoil) remain similar on adequately small physical scales in the detector. This means that for a local x-y region on the detector, bulk events and surface events of similar recoil energy have sufficiently different pulse shapes. Only over large physical regions and large spreads in
energy do timing parameters lose discrimination power. Thus phonon pulse shape variations over long x-y scales can be measured and subtracted out. Similarly, by characterizing the pulse shapes in small energy bins, but over a large energy range, variations due to energy can be subtracted out.

The gist of the method is as follows. First, a sample of photons populating the entire detector is acquired in the energy range of interest. Then the variation of the ionization yield and phonon-pulse-shape parameters such as $p_{\text{minrt}}$ and $p_{\text{del}}$ is characterized as a function of physical event location in the detector and the true recoil energy of the event. Then linearizing transformations are made to these functions, providing corrected pulse-shape parameters. In reality, it is very challenging to ascertain accurately the true physical location and recoil energies of events without elaborate expansions to the experimental setup. The practical solution is to use as proxy for these quantities their estimators from preliminary reconstruction, i.e., the partition- and delay-based coordinates explained in Section 4.1.4 and the fixed-template optimal filter pulse height.

### 4.3.2 Early correction techniques for ZIPs

The simplest correction is to fit the parameters of interest ($p_{\text{minrt}}$, $p_{\text{del}}$ etc.) from a gamma calibration sample as a function of the preliminary reconstruction estimators, and subtract out the non-linearity. This is easily done to remove energy dependence of timing parameters, as described in section 4.1.2 of [118]. This correction of energy dependencies was used for CDMS analyses from 1999 until 2006.

To remove position dependence, the delay plot for the same set of photons was divided up into 2-dimensional bins of equal size. Next, average values of the parameters of interest ($p_{\text{minrt}}$, $p_{\text{del}}$ etc.) were computed bin-by-bin and recorded on a map. These were then subtracted out for all events in all datasets (calibration and WIMP-search) according to the bin they occurred in. Obviously, the degeneracy of the delay parameter (Figure 4.6) severely degraded the efficacy of this technique. Even if the 2-dimensional bins were replaced with three-dimensional bins in the phonon-position manifold (Figure 4.8), the distribution of events in those bins was non-uniform and the estimators were sufficiently non-linear in true physical space, to make the correction ineffective in certain regions of the detector [118].

### 4.3.3 The lookup table

In 2002, a new position-dependence-removal algorithm was implemented that used varying bin size instead of bins of fixed size [118]. Each bin comprised of $N$ events in the three-dimensional phonon manifold of photons. Also, a bin was defined at each photon in the manifold, enabling finer and smoother characterization of phonon response throughout the detector. To select $N$ nearest neigh-
bors to form a bin, a distance metric was defined on the manifold,

\[ d = \sqrt{\Delta x_{\text{ppart}}^2 + \Delta y_{\text{ppart}}^2 + \frac{\Delta r_{\text{del}}^2}{L_{\text{del}}}} \]  

(4.6)

where \( L_{\text{del}} \) is a normalization factor chosen to weight the delay radius approximately the same as the partition radius. \( N \) was chosen approximately to represent the physical scale of phonon response variation in the detector. See the top pane of Figure 4.13 for a schematic representation of the lookup-table algorithm.

For each bin (or equivalently, photon,) average values of ionization yield and timing parameters were computed and compiled in a lookup table, after the energy dependence correction. Then, for all events in all datasets, these parameters were “position corrected” using the following formula

\[ \text{par}_{\text{corr}} = \frac{\langle \text{par} \rangle_{\text{bin}}}{\langle \text{par} \rangle_{\text{global}}} \]  

(4.7)

where \( \langle \text{par} \rangle_{\text{bin}} \) is the bin average for the parameter such as \( p_{\text{minrt}} \) or \( p_{\text{del}} \), as recorded in the lookup table, and \( \langle \text{par} \rangle_{\text{global}} \) is either a detector average or an arbitrary scaling factor. The bottom pane of Figure 4.13 shows a schematic representation of position dependence removal using the lookup table algorithm. To remove position dependence from the phonon energy estimate, a clever trick was devised. Instead of directly correcting energy, the ionization yield was corrected instead with \( \langle yg \rangle_{\text{global}} = 1 \), which is true for photons. Then, using the charge energy and corrected ionization yield, a new phonon energy was assigned.

Overall, this method showed marked improvement over a fixed-size-bin correction, as evidenced by better discrimination between surface events and bulk events in Run 118 and Run 119 compared to Run 21. [120, 121]. However, the correction of the position dependence of pulse shape parameters, separately from the energy dependence correction, reintroduced energy dependence [122]. This is shown in Figure 4.14. Switching the order of operations reintroduced position dependence. In retrospect, this is not surprising since energy and position dependencies are correlated in a partition-based manifold.

### 4.3.4 Combined removal of position and energy dependencies

In 2006, a combined correction of position and energy dependences was implemented [123]. The lookup-table metric was modified to include energy, and eliminated separate linearization of the energy dependence of parameters. The metric was modified to the following:

\[ d = \sqrt{\Delta x_{\text{ppart}}^2 + \Delta y_{\text{ppart}}^2 + \frac{\Delta r_{\text{del}}^2}{L_{\text{del}}} + \frac{\Delta pr_{\text{g}}^2}{L_{E}}} \]  

(4.8)
Figure 4.13: Top: Schematic representation of lookup-table creation. Bottom: Schematic representation of position-dependence removal using the lookup table.
where \( prg \) is the phonon recoil energy computed assuming the event is a photon. \( L_{del} \) and \( L_E \) are weighting parameters for delay coordinates and the phonon energy estimate, respectively. Note that this change in the metric effectively made the phonon manifold 4-dimensional, obfuscating visual representation. Also, physical length scales represented by the averaging bin were made longer because bins were now required to contain photons close together in energy. The strictness of this requirement is controlled by \( L_E \). Regardless, this change enabled empirical determination and removal of energy and position correlations in pulse-shape parameters across the detector.

### 4.3.5 Optimization of technique

A systematic study was undertaken to determine optimal values for \( N, L_{del} \) and \( L_E \) \[123\]. The optimal number of photons, \( N \), in a nearest-neighbor bin balances the statistical error on the average of the parameter of interest, and the physical scale of variations in phonon response on the detector. This is verified in Figure 4.15 which plots the standard deviation of the distribution of bin-wise standard deviations of \( p_{del} \). The standard deviation of \( p_{del} \) computed for each bin, \( \sigma_{p_{del}} \), is an indicator of the spread of \( p_{del} \) in that bin. The spread of \( \sigma_{p_{del}} \) for all bins is a measure of the variability of \( \sigma_{p_{del}} \) over the detector. At low \( N \), we see the spread decline with increasing \( N \), because of decreasing statistical error. After a point, the trend turns over because of increased variation due to true position dependence, as larger differences in phonon response are probed. Thus, we seek
the minimum of this curve. The optimal range for $N$ was observed to be $\sim O(100)$ for a sample of 400,000 photons in the detector. For $L_{del}$ and $L_E$, values were sought that maximized discrimination between surface and bulk events in calibration datasets, improved phonon energy resolution, and maximally removed position and energy effects. These were $\sim O(100)$ for $L_{del}$ and $\sim O(1000)$ for $L_E$, respectively. The efficacy of the correction was observed to be stable for a wide range of values in these optimal ranges.

![Figure 4.15](image)

Figure 4.15: Variation of bin size vs. $N$. The optimal $N$ minimizes this variation and selects the natural physical scale for phonon-pulse-shape variation in the detector. This plot was originally produced by Matt Pyle and Bruno Serfass. I used the calibration dataset from Runs 125–128 to reproduce it and verify their findings.

### 4.4 Improvements to Calibration Techniques for Current WIMP-search Analysis

In CDMS II, the reach of standard WIMP-search analyses has been limited by exposure time and expected residual surface events after rejection cuts. Since surface-event discrimination in ZIPs is, in turn, primarily limited by pulse shape variation, we always seek ways to improve our calibration and phonon-pulse-shape-correction techniques. In 2009, I undertook a full review of phonon calibration techniques for CDMS II and, over the course of a year, introduced enhancements intended to enable the experiment to reach projected goals for its final dataset. In this section, I outline these improvements.
4.4.1 Phonon-pulse-shape correction for events in outer charge electrode

During the first WIMP-search analysis with all five towers of detectors of CDMS II (Runs 123–4), it was observed in calibration data that the empirical correction algorithm occasionally miscorrected phonon timing of surface events to look more nuclear-recoil-like than the uncorrected phonon timing parameters [123]. The projected incidence rate of this phenomenon in WIMP-search data for that exposure was greater than one event, jeopardizing WIMP-search reach. A large fraction of these “outliers” in calibration data were observed to have high radius and slow timing in the phonon manifold. A careful study showed that these events were being corrected using a lookup table bin of low radius and fast timing events, pushing the corrected timing values to slow, nuclear-recoil-like timing (e.g. Figure 4.16, top pane). The quick fix proposed to mitigate the problem for that WIMP-search analysis was a “manifold cut” that removed all events within a certain distance of high-radius, outer-charge-electrode events on the phonon manifold. This worked, but at the price of a decline in WIMP-search efficiency [125].

While the manifold cut was being implemented, Bruno Serfass suggested a possible way to eliminate such events using the lookup table. Most CDMS analyses discard events with signal in the outer-charge electrode, because of the charge collection pathologies for such events. This argument was traditionally applied to lookup tables as well since the phonon recoil energy estimate required knowledge of the charge collection. Serfass proposed keeping the outer-charge-electrode events in the lookup table to tag calibration or WIMP-search events too close to them.

In preparation for the phonon-pulse-shape correction for Runs 125–128, we learned that the inclusion of outer-charge-electrode events in the lookup table actually prevented timing miscorrection of high-radius events [126]. As shown on the top pane of Figure 4.16, the event marked with the star has slow timing and is close to the charge boundary of the phonon manifold. The manifold was defined with a charge fiducial volume cut applied to its photons. The starred event’s nearest-neighbors bin is marked with bold-outlined markers, and consists primarily of fast timing events. This lowers the average timing for the bin, and thus makes the event seem slower and more nuclear-recoil-like. On the bottom pane, the manifold includes “good” photons from the entire detector. Consequently, the nearest-neighbors bin for the same starred event includes slower timing events from the outer-charge electrode. This marks the average phonon timing of that spot in the manifold as slow. Thus, the event now looks normal for its locality and is corrected to average photon timing as it should have been. Figure 4.17 shows the timing discrimination between calibration $^{252}\text{Cf}$ neutrons and $^{133}\text{Ba}$-induced surface events with and without the inclusion of outer-charge-electrode events in the manifold. For comparison, the effect of the manifold cut of Runs 123–4 is also shown. This change in the lookup table led to $\sim 80\%$ reduction in high-radius miscorrection of phonon timing and improved WIMP-search efficiency by $\sim 10\%$. 

Figure 4.16: Top: A photon event with high radius and slow timing in the phonon-position manifold is marked with a star. Its averaging bin, highlighted with black borders, consists of events with faster timing. Bottom: If the lookup-table manifold includes events in the outer-charge electrode, the averaging bin for the same event now includes better matched high-radius events with slower timing. This makes the starred event “normal” for its neighborhood.
Figure 4.17: Scatter plot of corrected event timing vs. recoil energy for $^{252}$Cf neutrons (green) and $^{133}$Ba-induced surface events (red). The top pane uses timing parameters corrected using a lookup table without outer-charge-electrode events. The bottom pane uses the new lookup table. Both show a timing cut at 80% neutron acceptance (horizontal blue line), without the effect of the manifold cut used in Runs 123–124. On the top pane, “outliers” identified and removed by the manifold cut are circled (surface events) or dotted (neutrons) in black. The loss in neutron efficiency because of the manifold cut is 10%. On the bottom pane, all but one of the true outliers, i.e., the circled events identified previously by the manifold cut, are corrected to faster timing and lie below the 80%-neutron-efficiency cut line. No manifold cut is used.
4.4.2 Delay in the lookup-table metric

Sometimes the two different ways of primary channel selection for events can lead to inconsistent results. The left pane of Figure 4.18 shows a delay plot with calibration gammas that have highest pulse height in phonon channel A, and the right pane shows a partition plot with calibration gammas that have the fastest pulse delay in phonon channel A. Neither selection leads to a sample of events within the expected physical boundaries of channel A. The misplaced events fall under different channels based on primary delay and primary partition. They are conjectured to be a combination of internal-multiple-scattering gammas, i.e., gammas with more than one center of energy deposition in the same detector, and events whose physical radius is sufficiently low to prevent the center of energy deposition being clearly identified in one channel.

![Figure 4.18](image)

Figure 4.18: Left: Delay plot with events that have highest phonon pulse height in channel A. Right: Partition plot with events that have fastest phonon pulse arrival in channel A. The spill of events over channel boundaries poses problems for identifying nearest neighbors for an averaging bin.

These events have problematic nearest-neighbor bin selection. An example is shown in Figure 4.19: the nearest-neighbors averaging bin for this event is selected within a local partition neighborhood, but with delay in the opposite channel. This leads to incorrect estimation of local phonon response.

Two methods were implemented to deal with this problem. First, a primary-channel consistency cut was imposed for selection of lookup table photons. Only events that had the same delay- and partition-based primary channel were included in the lookup table. Second, the distance metric in the lookup-table-generation code was modified to replace the delay radius, $r_{del}$, with $x_{del}$ and $y_{del}$. The new metric is as follows:

$$d = \sqrt{\Delta x_{pp}^2 + \Delta y_{pp}^2 + \frac{\Delta x_{del}^2}{L_{del}} + \frac{\Delta y_{del}^2}{L_{del}} + \frac{\Delta prg}{L_E}}$$

(4.9)

\[^3\text{either tallest phonon pulse, or fastest arriving phonon pulse among the four quadrants}\]
Figure 4.19: Delay, partition, and shrimp plots for an event with different primary partition channel and primary delay channel. It is marked with a pink dot on the delay and box plots and with an “X” on the shrimp plot. Its nearest neighbors are marked with green dots on the box and delay plots, and with colored dots on the shrimp plot, with \( \text{pdel} \) for the color scale. On the partition plot, the event is surrounded by its averaging bin. On the delay plot, we notice that the averaging bin is in the opposite quadrant from the event. This happens because the distance metric forces events in the averaging bin to be of similar delay radius, not similar delay x-y neighborhood.
Thus it was no longer sufficient for the bin to have events of the same delay radius; the bin had to be compact on the delay plot. For the event shown in Figure 4.19, the nearest-neighbors bin with the new metric was revised as shown in Figure 4.20. A casual glance makes the situation seem worse. But in reality, the unusually larger bin enables clear identification of the event as an “off-manifold” outlier. It is easily rejected by a cut on bin size $[127]$. Almost all of these events are also removed by a cut discussed in the next section.

Figure 4.20: Same event as in Figure 4.19 but with averaging bin identified by the metric in Equation 4.9. Now, the metric forces the events of the averaging bin to be in the same x-y partition neighborhood as well as x-y delay neighborhood as the event in question. This photon is now easily identified and removed from the lookup table.

### 4.4.3 Distance of event to phonon manifold

The study of new detectors for SuperCDMS Soudan identified a new technique for discriminating between surface electrons and bulk nuclear recoils — by comparing events directly with phonon manifolds of photons and neutrons. A phonon manifold with timing parameters and yield in addition to position and energy estimators was observed to be different for electron recoils and nuclear recoils.
The difference in chi-squared distances of an event to these two manifolds compared its characteristics to electron recoils and nuclear recoils in the same physical region of the detector, as demonstrated by the cartoon in Figure 4.21. This idea was imported into the lookup-table-generation code for CDMS II ZIPs, because the computation of a chi-squared distance to the manifold was computationally cheap. However, a lack of sufficient neutron calibration data, as well as a smaller difference between the neutron and photons manifolds for ZIPs, hampered the success of discrimination using this technique. Instead, we devised other uses for this chi-squared distance — to create quality cuts on the phonon-pulse-shape correction itself. A cut was developed to remove photons insufficiently “similar” to their nearest-neighbor averaging bins from the lookup tables for Runs 125–128 [128]. This also removed a majority of photons with large nearest-neighbor bin size. Another cut was developed for WIMP-search data that rejected events corrected using lookup-table bins too distant from the photon manifold, and is explained in Section 6.4.5 [129].

Figure 4.21: Schematic of difference between electron-recoil and nuclear-recoil phonon manifolds in several parameters including energy, position, and timing. The chi-squared distance of an event to the two manifolds would serve to classify the event as one or the other.

4.4.4 Optimization of source code and processing

Finally, we gutted the old lookup-table-generation code and the pulse-shape correction code that had been passed from generation to generation, having acquired makeshift patches along the way. I rewrote from scratch a streamlined MATLAB package called CorrTools and implemented version control for it using CVS. The new package took advantage of advances in parallel computing and was adapted to handle large quantities of data. The runtime for lookup-table generation and pulse-shape correction was reduced from several days to hours.
4.5 Results using calibration data

The effectiveness of the empirical correction technique can be seen by comparing uncorrected phonon-pulse-shape parameters with their corrected counterparts. Figure 4.22 shows the uncorrected primary phonon risetime as a function of uncorrected recoil energy for a ZIP using calibration gammas on the left pane. There is a clear energy dependence as well as a non-gaussian tail towards slower risetimes. The corrected risetime is displayed on the right pane of the figure. The units of the risetime are normalized by subtracting the mean in both plots to facilitate comparison between the two.

In Figure 4.23 I plot a slice of the phonon position manifold and show uncorrected and corrected primary risetime, once again, normalized by subtracting the mean. The x-y position dependence seen on the left pane is not seen on the right pane. Similarly, in Figures 4.24 and 4.25 the energy and x-y position dependencies are compared for the uncorrected and corrected ionization yield.

![Figure 4.22](image_url)

Figure 4.22: Normalized $p_{\text{minrt}}$ vs. photon-equivalent recoil energy for $^{133}$Ba calibration gammas before and after removal of position and energy dependencies.

![Figure 4.23](image_url)

Figure 4.23: Normalized $p_{\text{minrt}}$-colored shrimp plot for $^{133}$Ba calibration gammas before and after removal of position and energy dependencies.

![Figure 4.26](image_url)

Figure 4.26 shows histograms of the summed timing discriminator ($p_{\text{minrt}} + p_{\text{del}}$) for $^{133}$Ba-calibration-induced surface events and $^{252}$Cf neutrons with and without the lookup table correction.
Figure 4.24: Ionization yield vs. photon-equivalent recoil energy for $^{133}$Ba calibration gammas before and after removal of position and energy dependencies.

Figure 4.25: Ionization yield vs. photon-equivalent recoil energy for $^{133}$Ba calibration gammas before and after removal of position and energy dependencies.
applied to them. The correction sharpens both surface event and neutron timing distributions and shortens the surface-event-timing tail. Finally, in Figure 4.27 we see quantitatively the decrease in surface-event contamination at any desired signal acceptance due to correction of pulse shape timing. At 60% nuclear-recoil acceptance, the empirically corrected phonon timing provides 20× better rejection of surface events. As I will discuss in Chapter 6, the true surface-event leakage rate in WIMP-search data is worse than that seen for calibration data in Figure 4.27. This is due to systematic differences in the distribution of surface events between calibration data and WIMP-search data, and can be taken into account while setting surface-event rejection cuts. Regardless, the empirical phonon-pulse-shape correction provides a significant improvement in surface-event rejection and improves WIMP-search reach. For the analysis of WIMP-search Runs 125–128 presented in this dissertation, the required improvement in surface-event rejection was a factor of 2x over that of Runs 123–124, to take advantage of increased exposure. This was accomplished without any loss in signal-acceptance efficiency compared to Runs 123–124 (see Section 6.7), implying that the improvements in phonon-pulse-shape correction provided the necessary increased surface-event rejection power.

Figure 4.26: Histograms of normalized $p_{\text{minrt}} + p_{\text{del}}$ distributions for $^{133}$Ba-induced surface events (SE) and neutrons from $^{252}$Cf (NR) before and after removal of position and energy dependencies.
Figure 4.27: Surface-event leakage rate vs. nuclear-recoil acceptance efficiency for $^{133}$Ba-induced surface events and neutrons from $^{252}$Cf. At 60% signal acceptance, the empirically corrected phonon timing provides 20x better rejection of surface events.
Chapter 5

The CDMS II Experiment

5.1 Soudan Underground Laboratory

The CDMS II experiment is situated in the Soudan Underground Laboratory, a research facility in the Soudan Underground Mine in northern Minnesota. The mine was active and a source of iron ore for US Steel until 1962. After it was decommissioned, it was acquired by Minnesota’s Department of Natural Resources as part of a state park. Its primary use now is as a mining museum and as a physics research laboratory. The lab is located 713 m (2341 ft) below the surface, at the 27th level of the mine. CDMS II is housed in the cavern previously used for the Soudan 2 experiment and shares resources with a few other experiments, most notably the MINOS experiment far detector located in an adjacent cavern. The underground location was picked for the shielding provided against atmospheric muons by its rock overburden. Muons are capable of interacting in materials surrounding the experimental setup and generating neutrons that could mimic WIMP signal. At 2090 meters water equivalent, the depth of Soudan Lab suppresses muon flux to $2 \times 10^{-5}$ of the level observed at the surface, substantially reducing the cosmogenic neutron background rate (see Section 7.1) and increasing sensitivity to WIMPs. A picture of the Soudan mine headframe and the Soudan 2 cavern are shown in Figure 5.1.

5.2 Infrastructure

The Soudan 2 cavern was fitted with special infrastructure to support CDMS II, a cryogenic experiment with stringent requirements for low background and noise. A schematic cross section of the installation space is provided in Figure 5.2. The detectors are housed in a cold volume called the “icebox,” connected to a dilution fridge via a “cold stem” (c-stem) and to a front-end electronics system via an “electronics stem” (e-stem). All of these sit in a class-10000 clean room with RF shielding to provide a low-electrical-noise environment. The cryogenic support equipment and pumps to run the dilution fridge are housed on the “cryopad,” adjacent to the RF-shielded clean
Figure 5.1: Top: Soudan mine framehead. Bottom: Picture of Soudan 2 cavern taken from the mezzanine level. The HVAC and top half of the RF room are visible on the left, and the CDMS II office space on the right.
room (RF room), and have dedicated feedthroughs through the wall dividing them. The entrance to the RF room is through an anteroom to enable preparation and cleaning of materials being brought in, as well as to allow personnel to don clean-room apparel before entering. Finally, a mezzanine level houses the clean-room HVAC system, the data acquisition (DAQ) system for the experiment. Feedthroughs through the wall dividing the front-end electronics and the electronics room carry signal wiring to the DAQ. There is also some office space for CDMS II scientists and engineers on the mezzanine level.

Figure 5.2: Schematic cross section of Soudan 2 cavern space used by CDMS II. Courtesy: Dan Bauer.

5.3 CDMS II shield

The icebox housing the ZIPs is surrounded by a layer of passive and active shielding to reduce the ambient photon and radiogenic neutron background, and to discriminate against muon showers potentially containing cosmogenic neutrons. A 3-D CAD rendering of the icebox, the shielding, and the dilution fridge is shown in Figure 5.3 and a labeled cross-sectional schematic of this setup is shown in Figure 5.4.

The outermost layer of the CDMS II shield is an active veto, consisting of 40 BICRON BC-408 plastic scintillator panels that produce scintillation light from electromagnetic particle interactions. The panels overlap each other to leave no exposed sections except for the room required for the c-stem and e-stem. The scintillator panels are connected to Hamamatsu R329-02 photomultiplier tubes by acrylic light guides. Optical fibers connected to LEDs on one end and fed to the veto provide the ability to calibrate PMT response.

The inner layers of the shield provide passive protection against background. The outermost of these is a 40-cm-thick cylindrical shell made of custom-cut polyethylene bricks. This layer moderates low-energy neutrons from external sources to well below the experiment’s energy threshold. Inside
this polyethylene shield is a 22.5-cm-thick shell of lead to moderate the photon background. This is divided into a 17.8-cm-thick shell of lead bricks, and a further 4.4-cm-thick shell of ancient low-activity lead\(^1\) to shield from the intrinsic radioactivity of the regular lead. Inside this, another 10-cm-thick shell of polyethylene provides neutron moderation for high-energy neutrons that make it through the outermost polyethylene and lead. In the absence of the inner polyethylene shell, neutrons penetrating the lead shield would continue to elastically scatter off the much heavier lead nuclei and get trapped in the volume enclosed by the lead. Finally, a mu-metal can surrounds the icebox, not for the purpose of reducing particle backgrounds, but to reduce the external magnetic field. Stray variable magnetic fields can cause flux jumps in the SQUIDs used to read out the ZIP phonon sensors, resulting in non-stationary noise. The mu-metal shield surrounding the CDMS icebox provides a factor 100x reduction in ambient magnetic field strength.

An added measure of passive background suppression is flushing the space between the vacuum-sealed icebox and the shielding layers with boil-off gas from liquid nitrogen. This displaces the high radon concentrations found in Soudan mine air (\(~700\text{ Bq/m}^3\) from that gap and leads to a drop in ambient photon rates by a factor 4x and beta rates by 2x.

\(^1\)The ancient lead is void of radioisotope \(^{210}\text{Pb}\), which has a 22.3-year half life.
Figure 5.4: Cross-sectional view of the CDMS II cryostat showing the active and passive shielding layers employed in the experiment, along with annotations. The top panel shows a side view, and the bottom panel shows a top view. Original CAD: Susanne Kyre.
5.4 Cryogenics

As mentioned already, the innermost sanctum of the experiment where the detectors are housed is the icebox. The CDMS II icebox is almost an exact copy of the CDMS I icebox described in [130]. Cooling is provided to the icebox by an Oxford Instruments 400S $^{3}\text{He-}^{4}\text{He}$ dilution refrigerator (Figures 5.3 and 5.4), with a cooling power of 400 $\mu$W at 100 mK and a base temperature of 10 mK without a thermal load. The fridge itself is not made of radiopure materials and has high contamination levels of U, Th, and $^{60}\text{Co}$. Thus it is mounted outside the shielding and off to a side to prevent a direct line of sight of its materials to the detectors. The thermal coupling to the dilution fridge is provided by a set of concentric copper pipes, called the “cold stem” (c-stem), which mate each of the fridge’s temperature stages to concentric low-activity copper cans inside the icebox. The dilution fridge is supported by vacuum pumps, plumbing, and cryogens (liquid helium and nitrogen) located on the cryopad. The purity of the $^{3}\text{He-}^{4}\text{He}$ gas mixture in the circulation loop is maintained by running it through a series of regularly cleaned cold traps. The loop is controlled and monitored by an Oxford Instruments Intelligent Gas Handling (IGH) unit, which can be operated remotely via the internet.

In 2005, the payload of CDMS II was expanded from 12 ZIPs to 30 ZIPs, thereby increasing the heat load on the cryogenic system. Modeling showed that without additional cooling power, the cryogen consumption rate would double, requiring more frequent refill cycles. This would reduce the livetime of the experiment. Also, the base temperature would be high enough to prevent TESs on some detectors from achieving superconductivity. Thus a Gifford-McMahon cryocooler was added to the system to provide additional cooling. It provides 1.5 W of cooling at 4 K, and 40 W at 77 K and is mounted to those temperature stages of the system at the e-stem as shown in Figure 5.3. This was sufficient to keep the cryogen hold time for the dilution fridge at $\sim 24$ hours. While the cryocooler added cooling power, it was unfortunately a source of microphonic noise for the remainder of CDMS II’s data taking. This is discussed in Section 6.4.3.

The full cryogenic system is monitored and controlled by a Moore APACS industrial control system, located on the cryopad. It records more than 150 parameters of the system and enables electronic control of most of them, on site or over the internet. Its most important function is to automate the daily refilling of the dilution fridge’s liquid He and liquid N baths. It does this in concert with the DAQ system, which pauses data acquisition and engages veto calibration and detector neutralization during a fill.
5.5 Cold Hardware

The detectors, their housing, stacking infrastructure, first-stage amplifiers, and electrical connectors housed inside the icebox are collectively called the cold hardware and are described in detail in [131].

5.5.1 Detector housing

Each ZIP is housed in a hexagonal high-purity, low-radioactivity copper container and is held in place in the container with Cirlex clamps. A picture is shown in Figure 5.5. The housings have top and bottom caps but these are not used except for storage or transportation of individual ZIPS. Detectors in use are stacked on top of one another with an unobstructed view of their flat faces separated by 2 mm. This enables tagging and rejection of events with low-penetration depth that scatter in neighboring detectors. A small detector interface board (DIB) provides electrical connections between the detector and the signal amplification chain downstream; it is shown in Figure 5.5. The DIB also houses the infrared LEDs that are used to neutralize detectors. Connections between the detector and DIB are made by wirebonds. The DIB is connected to a removable side coax carrying signals via mill-max sockets exposed through a cutout in one of the housing faces. This enables up to six detectors to be stacked on top of each other with one face of the stack occupied by the side coax for one detector.

Figure 5.5: Left: ZIP housings pictured with metallic placeholders of the mass (Ge) and outer shape of ZIPS. Right: Detector Interface Board. Courtesy: Dennis Seitz.

5.5.2 Tower

The tower (Figure 5.6) is a hexagonal copper structure that holds a stack of six ZIPS on one end and six SQUET cards (Section 5.5.4) that provide first stage amplification of ZIP signals on the other end. It also provides electrical connections between the removable side coaxes and the SQUET cards. A tower consists of four thermally isolated sections separated and supported by graphite tubes. The
four sections provide heat-sinking to four different temperature (4 K, 600 mK, 50 mK, and 10 mK) stages by mating to different copper cans of the icebox. The SQUET cards reside at the 4 K stage at the top of the tower. The wires connecting the SQUETs and the side coaxes reside in vacuum coax channels, are tensioned to reduce microphonic noise, and are heat-sunk along the way to reduce thermal load on the side coaxes at base temperature.

Figure 5.6: A CDMS II tower, consisting of four isolated metal stages, and stretched wires on the sides to provide connections between SQUET cards on top and the side coax mill-max connectors on the bottom. *Courtesy:* Dennis Seitz.

The term “tower” colloquially refers to the tower, the SQUETs, and the detector stack as a single assembly. CDMS II thus had five towers of six ZIPs each. The detectors are identified by their tower number and their position in the stack. Thus the fifth detector in the fourth stack is called “T4Z5.” A schematic layout of Si and Ge ZIPs in their towers is provided in Figure 5.7.

### 5.5.3 Side coax

Detectors are electrically connected to the base temperature stage of their tower by a custom connector card called a side coax. A side coax also houses the coupling capacitors and bias and feedback resistors for the charge channels of a detector, to minimize Johnson noise. Side coaxes for CDMS II were made in six different lengths to reach each of the detectors from the base temperature stage of the tower. A side coax is pictured in Figure 5.8.
Figure 2.3: Physical arrangement of the ZIPs in the final CDMS II configuration. The detectors are labeled by tower (T1–T5) and position within tower (Z1–Z6) (e.g. T4Z5). ZIPs are color-coded by type in the diagram: beige for silicon, aqua for germanium. The ZIPs are arranged in stacks of six inside a hexagonal copper housing above which are situated the cards containing the sensor circuit elements; this assembly is referred to as a “tower”. There are five towers and thus 30 ZIPs, 19 composed of germanium and 11 of silicon. Due to detector failures or poor performance some ZIPs are not used for WIMP search, but all 30 serve as coincidence detectors. The arrangement is shown in figure 2.3.

2.2 Shielding and Cosmic Ray Screening

The ZIP towers are inside the innermost of six nested copper cans, each connected to a different temperature stage of the cryogenic system by concentric copper tubes extending through the layers of shielding. The wires from the internal ZIP electronics are run through a second, similar set of tubes. Low radioactivity material is used for the cans and other internal hardware. The cans and housings provide an average of 3 cm of copper shielding, effectively shielding the ZIPs from alpha and beta radiation external to the cans. The space between the cans is under vacuum.

Figure 5.7: Arrangement of ZIPS in CDMS II. The detectors are labeled by tower (T1–T5) and position within tower (Z1–Z6) (e.g. T4Z5). ZIPs are color-coded by type in the diagram: beige for silicon, aqua for germanium. This arrangement of towers into two rows was exactly how the towers were positioned in the icebox. The bottom of the picture points North (cf. Figure 5.3). Courtesy: Matt Fritts.

Figure 5.8: A CDMS II sidecoax, which carries signals from the base of the tower to the appropriate detector. Courtesy: Dennis Seitz.
5.5.4 SQUET card

A SQUET card (Figure 5.9) is a two-piece electronics module that houses the SQUID arrays for phonon signal amplification and the first stage FETs for charge signal amplification for the channels of a single detector. Thus each of the six SQUET cards sit atop the tower, aligned with the face that houses the side coax for its detector. The two cards are connected by a “flyover” cable. The primary card of a SQUET houses the FETs, and a smaller secondary card contains the SQUIDs, shunt resistors, and the input and feedback coils of the phonon readout circuit. The FETs are suspended on a Kapton membrane and inside a copper gusset on the primary card. This enables the FETs to self-heat to \( \sim 140 \text{ K} \) for nominal operation, while still being mounted to the 4K stage. The SQUID card is heat sunk to the 600 mK stage for optimal performance.

![CDMS SQUET Assembly](image)

Figure 5.9: Photograph (left) and circuit board layout (right) of a SQUET card. Courtesy: Dennis Seitz.

5.5.5 Stripline

Flexible striplines bring electrical signals from the SQUETs at 4K, through the e-stem, to the electronics box (e-box) at room temperature. They are 3-m-long flat ribbons that sandwich copper traces between two ground planes and are enclosed in insulating kapton. Each stripline connects one SQUET and hence one ZIP to a 50-pin connector on the e-box. Along the way, it is heat sunk in two locations to reduce the thermal loading from the outside world at the 4K stage. A picture of

\(^2\)The name SQUET is an abbreviation derived from SQUID and FET.
the top of four towers with SQUET cards and striplines is shown in Figure 5.10.

Figure 5.10: The top of the open icebox, showing tops of towers with SQUET cards and their respective striplines winding around and exiting the icebox through the e-stem.

A CAD rendering of the tower and a cross-sectional view are provided in Figure 5.11.

Figure 5.11: Left: CAD rendering of a tower, ZIPs and the SQUETS when assembled. Right: Cross-sectional view of the same assembly, with one of the ZIPs pulled out. Courtesy: Patrick Wikus.

5.6 Warm Electronics

Beyond the e-box, the electronics and data acquisition chain sit at room temperature and are called the warm electronics.
5.6.1 Front-end boards

The front-end boards (FEBs) are custom 9U PCBs that house the remainder of the signal amplification chain and readout circuits for the charge and phonon channels; there is one FEB per detector. The FEBs also contain circuits to control biasing of the detectors, cold amplifiers, SQUIDs, LEDs, etc. Signals are brought to them from the e-box by 50-wire cables with detachable connectors. These processed signals are then sent to the electronics room for triggering and digitization. The FEBs are operated from the electronics room by a fiber-optic-linked GPIB controller. A picture of a 9U rack of FEBs is shown in Figure 5.12.

![9U rack of FEBs in the RF room.](image)

Figure 5.12: 9U rack of FEBs in the RF room.

5.6.2 Receiver-trigger-filter boards

The receiver-trigger-filter (RTF) boards receive signals from the FEBs over cables fed through the RF room wall. Signals are first conditioned on the RTF boards by adjusting their baselines and by applying a 336-kHz 2-pole Butterworth anti-aliasing filter. The filtered signals are then used to generate five types of triggers for the digitizers. The primary trigger, $P_{lo}$, is issued when a comparator determines that the sum of the four phonon pulses exceeds a pre-defined amount set in software. $P_{hi}$ is similar but has a larger threshold, while $P_{whisper}$ has a lower threshold. Finally, $Q_{hi}$ and $Q_{lo}$ are analogs of $P_{hi}$ and $P_{lo}$ that used the summed charge signal to generate triggers.
5.6.3 Veto triggers and signal conditioning

Signals from the veto panels are transported to the electronics room, just as those for the ZIPs. LeCroy discriminators compare the PMT pulse heights to a software-defined threshold and issue triggers for veto panels whose signal exceeds threshold. Unlike the RTFs for the ZIPs, the LeCroys have only one threshold per panel. Separately, the PMT signals are conditioned by stretching them from $\sim$ ns scales to $\sim$ $\mu$s scales by a special filter network before being sent to the DAQ hardware.

5.6.4 Data acquisition (DAQ) hardware

Triggers from the RTF boards and the LeCroy discriminators are received and processed by a trigger logic board (TLB), which issues a “global” trigger to record ZIP and veto traces to disk if any of the following conditions are met:

1. The $P_{lo}$ trigger for any ZIP is issued by an RTF board.
2. The LeCroy discriminators issue triggers for two or more veto panels.
3. The DAQ software issues a “random trigger” for purposes of recording and monitoring noise levels.

When a global trigger is received during WIMP-search or $^{252}$Cf-calibration running, digitized traces from all channels of all ZIPs and all veto panels are recorded to disk. During $^{133}$Ba-calibration running, digitized traces are recorded from all channels of ZIPs that had $P_{lo}$ triggers, while veto traces are not recorded.

ZIP signals are digitized by Struck SIS 3301 ADCs. They sample at 80 MSa/s, but a 64-sample running average produces an effective sampling rate of 1.25 MSa/s with 14 effective number of bits. Once a global trigger is issued and a ZIP channel needs to be recorded, an ADC records 2048 samples ($\sim$ 1640$\mu$s) of that channel to disk in a $[-512, 1535]$ sample window around the global trigger. Veto signals after being conditioned are recorded to disk by 12-bit Joerger VTR812 digitizers, operating at 5 MSa/s and taking 1024 samples.

The time stamps for all triggers are monitored and kept in a circular buffer by a set of Struck SIS 2400 TDCs, with 1 $\mu$s resolution. When an event is recorded to disk, a delayed snapshot of this buffer is also recorded to disk to provide information of trigger times preceding and following the one being recorded.

Finally, a “slow DAQ” monitors all ZIP and veto channels, recording signal offsets and mean trigger rates once a minute. The former helps track loss of SQUID flux lock, and the latter helps monitoring of changes in detector noise.
5.7 Data Acquisition Software

The CDMS-II DAQ is controlled by a custom-written software package of programs in Java and C++, run on a small cluster of computers in the electronics room. Different components of the DAQ are controlled by dedicated pieces run on their own servers, but with great flexibility in adding or removing components. This is facilitated by the CORBA network messaging framework and Java Remote Method Invocation for communication between modules. All modules have user interfaces for control by the experiment operator, but a main interface called RunControl enables control and monitoring oversight. RunControl is a Java cross-platform network app that can control the entire experiment over the internet. The DAQ of course limits the locations and users from which it accepts RunControl commands to the mine, the surface building at Soudan, and a few other “super users” located anywhere.

The DAQ deadtime for acquiring traces from the digitizers and storing them after a global trigger is \( \sim 50 \text{ ms} \), allowing a maximum event rate of 20 Hz, far lower than the \( \sim 0.3 \text{ Hz} \) background rate during WIMP-search running. In selective-readout mode for \(^{133}\text{Ba}\) calibration, this deadtime falls to \( \sim 15 \text{ ms} \), enabling event rates as high as 70 Hz.

5.8 Data Storage and Processing

Raw event data acquired by the DAQ software is first stored on local disks on the DAQ cluster. This is automatically backed up to tape as well as transferred by a dedicated gigabit link to larger storage devices at the surface building. A cluster of computers at the surface building processes a copy of this raw data, with very rough calibration, to provide preliminary reduced datasets that the operators can verify experiment performance with. Separately, the raw data are transferred over the internet to Fermilab for complete data processing on the FermiGrid cluster, to generate reduced datasets that end users can analyze to search for WIMPs. This is described in the next chapter.
Chapter 6

WIMP-search Analysis Pipeline

In this chapter, I review the analysis pipeline for WIMP-search data acquired in CDMS-II germanium detectors between July 2007 and September 2008. These data were taken in four discrete runs of the dilution fridge, numbered 125, 126, 127, and 128. Interruptions between runs were for cryogenic maintenance. Care was taken to account for small differences in run conditions between the four data-taking periods. Data from Runs 123 and 124, the first data with the full complement of 5 towers of ZIP detectors were analyzed and presented in [72] and [132].

I first briefly recap the first- and second-tier data processing mentioned in previous chapters, in the context of the analysis pipeline. Then, I outline the process of blinding the signal region, selecting “good” WIMP-search data and setting reconstruction and physics cuts. I devote a complete section to the implementation of the surface-event rejection cut, in which I played a key role. Finally I review the efficiency and WIMP-search exposure for this analysis. The results of the analysis are the subject of the next chapter. The work of a large number of collaborators made this analysis pipeline possible, including that of former CDMS graduate students and postdocs. As much as possible, I try to directly cite CDMS internal notes and call out names of key contributors to ideas or work that were not my own, so that readers have a paper trail for the evolution of this pipeline.

6.1 Data-Processing Pipeline

After acquisition using the DAQ, the data undergo several stages of data processing on Fermilab’s computing cluster before it is used for analysis [133].

1. **First-tier processing**: During first-tier processing, raw data, consisting of detector and trigger settings, event traces, veto activity, etc., are converted into reduced quantities (RQs). These have unphysical, uncalibrated units until the second-tier processing is performed. For Runs 125–128, a first-tier processing package in MATLAB called *DarkPipe* was retired in favor of a streamlined, C++ package called *BatRoot*. BatRoot combined all the old algorithms with new time-domain fitting routines from an alternate pipeline called PipeFitter. The advantage
of the new package was its modularity in adding new experimental algorithms, and its ability to produce RQs in ROOT Ntuple format [134]. The latter was done with the objective of eventually transitioning CDMS data storage format to ntuples, a better standard for handling large datasets than MATLAB’s data format. However, the full transition would have required conversion of all data-processing software to support this format, which could not be achieved in time for the analysis of Runs 125–128.

2. **Second-tier processing:** During second-tier processing, the charge and phonon RQs undergo calibration resulting in physically meaningful quantities called relational RQs (RRQs). The position dependence in charge is removed at this stage, making it ready for use. The phonon calibration is only preliminary at this stage. Also, the blinding procedure, described in Section 6.2, is performed for WIMP-search data. For Runs 125–128, a MATLAB package called `PipeCleaner` was used for these calibrations, since the C++ version called `BatCalib` was not ready for use. SuperCDMS Soudan has already transitioned to an integrated C++ data-reduction pipeline using BatRoot and BatCalib.

3. **Lookup table generation:** This process is described at length in Chapter 4 and Appendix A. Lookup tables for Runs 125–128 were generated using `CorrTools`, a MATLAB package that I wrote.

4. **Application of lookup table correction:** The phonon energy and timing parameters of events in all datasets underwent phonon-pulse-shape correction using the lookup tables for Runs 125–128. This was done with a CorrTools plugin for PipeCleaner.

5. **Assembly of calibrated, corrected dataset for end users:** The pipeline, as it stood at the end of Runs 125–128, produced a combination of MATLAB and ROOT datafiles, which were then converted to independent datasets on both platforms to allow users freedom in choosing an analysis platform. Additionally, as cuts were developed, smaller skimmed datasets were produced for both platforms to allow faster loading and shorter operation times. Cuts were produced in MATLAB and ported to ROOT using a standardized boolean format. In the future, there will be a common standard to define cuts on both platforms.

A schematic of this data-processing pipeline is provided in Figure 6.1. Eventually, all of its components will be transitioned to C++ packages that produce ROOT datasets. The standard set of MATLAB tools used for analysis of CDMS datasets, called CDMS Analysis Package (CAP) [135], has already been modified to directly read ROOT n-tuples for SuperCDMS Soudan.

### 6.2 Blinding WIMP Candidates

After the first- and second-tier data processing, events satisfying liberally defined criteria to be potential WIMP candidates are removed from all distributions of WIMP-search data until the entire
analysis is finalized [136]. All cuts for the analysis are then defined using calibration data and data outside the masked signal region. The original dataset containing these blinded events is made available only for study of signal sidebands and subsequent unblinding, towards the very end of the analysis. The motivation for this is to prevent bias in defining or tuning the selection criteria for WIMP candidates, which could influence the outcome of the search. Note that the blinding process is designed to “overblind.” As the analysis progresses and relevant cuts are rigorously defined and frozen, events failing those cuts are returned to the unblinded dataset.

A logical AND of the following criteria is used for blinding:

1. **Nuclear recoil**: Events within $3\sigma$ of the mean ionization yield as a function of recoil energy, for calibration neutrons from a $^{252}$Cf source. This is based on preliminary phonon energy calibration.

2. **Veto-anticoincidence**: No activity in the scintillator veto panels surrounding the apparatus for a 50 $\mu$s window before a global trigger.

3. **Energy range**: Events with recoil energy between 2 keV or the $6\sigma$ noise threshold for that detector (whichever is higher) and 130 keV.

4. **Passing charge fiducial volume**: Charge energy below 5 keV in the outer charge electrode, i.e., events contained mostly in the fiducial electrode.

5. **Single scatter**: Energy deposition above 2 keV or the $6\sigma$ noise threshold (whichever is higher) for only one detector.
As a matter of unfortunate tradition, the blinding was done for Runs 125–128 during second-tier processing, instead of after phonon-pulse-shape correction. While the criteria listed above tend to overblind and are typically insensitive to changes in recoil energy caused by the correction, we note that some events that were originally not blinded did move into the final nuclear recoil band. We believe that the spirit of blinding was not violated since the properties of these events were not studied, and an updated blinding cut removed these events from WIMP-search data until the analysis was finalized. For future WIMP searches, I encourage analyzers to run the blinding process after phonon-pulse-shape correction. Note that the phonon-pulse-shape correction uses calibration data, not WIMP-search data, to measure detector response for the lookup tables. Thus running correction before blinding WIMP candidates will not bias the WIMP search.

6.3 Data Selection and Data Quality Cuts

The first set of cuts is made on data quality. We require data-taking conditions to be consistent with standards defined in this section; we remove data series taken in suboptimal operating conditions for the experiment or for particular detectors within it.

6.3.1 Operational detectors

First, we used only the germanium detectors for this WIMP search because their sensitivity is far greater than that of the silicon detectors. However, all detectors except T5Z2 were used to enforce the single-scatter requirement for WIMP search. Of the 19 germanium detectors, 5 were unusable for WIMP search for the entire duration of Runs 125–128 for various reasons, and are listed below.

1. **T1Z1**: Only one working phonon channel, and no working LEDs for neutralization. Problem since Run 123.
2. **T1Z3**: Outer-charge channel disconnected from amplifier, making a charge fiducial volume cut impossible. Problem since Run 123.
3. **T5Z1**: Phonon response for channel D was distorted compared to other channels, because of unknown reasons, making phonon-pulse-shape correction difficult. Also large periods of glitches in T5Z2 prevent its use as a multiple-scatter veto increasing background rate on T5Z1.
4. **T5Z2**: Fiducial charge electrode intermittently shorted to ground. Also, extensive glitch events. Problem since Run 123.
5. **T5Z6**: Extremely poor signal-to-noise on phonon channels, making its analysis threshold very high, and contribution to WIMP-search reach low. Problem since Run 123.

Some germanium detectors were unusable for specific runs and are listed below.

\(^4\) except T5Z2 in Run 127
1. **T1Z5**: Phonon channel B was lost during the partial warmup between Runs 126 and Runs 127, due to a superconducting short. Thus it was excluded from Runs 127 and 128.

2. **T2Z3**: Phonon channel C was lost before R128 for similar reasons as above. Thus the detector was excluded for R128.

3. **T2Z2, T4Z2**: Excluded for lack of sufficient calibration statistics in Run 128 to generate lookup tables for phonon-pulse-shape correction. (See end of Section 6.3.2)

The detector selection was enforced by a cut called $c_{\text{GoodDet}}$.

It is worth noting that a final cryogenic run, Run 129, took place, but had several more detector failures. Most of the new failures were due to poor connections after the installation of a new electronics vacuum feedthrough box after Run 128. Due to changes in dimensions, extenders were attached to the striplines and these extenders caused unforeseen problems with the electronics. Also, the functioning detectors suffered reduced phonon pulse amplitudes, presumably because of superconducting helium films. Thus very little WIMP-search livetime was obtained before it was decided to scrap the run in favor of upgrading the experiment to SuperCDMS Soudan.

### 6.3.2 Data quality assurance using KS tests

During stable data taking, WIMP-search datasets were usually acquired in 11-hour or shorter segments called series. Interruptions occurred primarily for automatic LED flashing to keep the detectors neutralized, for cryogenic fills, or for a couple of hours of $^{133}$Ba-source calibration, three times a week. Source calibrations using $^{252}$Cf were far more infrequent. Possible variations in detector conditions between these several-hour-long chunks, as well as changes in run conditions between cryogenic runs necessitates verification of data quality for both calibration and WIMP-search datasets. This was done using Kolmogorov-Smirnov (KS) tests. A KS test compares a 1-D empirical distribution with the cumulative distribution of a reference set and assigns a distance between them called the KS statistic. Under the null hypothesis, both sets are derived from the same parent distribution.

KS tests were performed for every data series in a run against a randomly selected golden set of $\sim 30$ data series (of the same run) with average livetime. This was done separately for calibration data series and for WIMP-search data series. The KS tests were performed on ionization yield, primary risetime, primary falltime, delay radius, partition radius, charge partition distribution (between fiducial and outer electrode), and the chi-squared goodness-of-fit parameter for the charge optimal filter. For each data series, the $\sim 30$ KS statistic values were averaged for each of the seven RRQs and then assembled into distributions. Data series that had average KS statistic values below $2\sigma$ of the mean KS value for any of the tested RRQs were removed by cuts $c_{\text{BadDet, bg}}$ for WIMP-search data and $c_{\text{BadDet, ba}}$ for $^{133}$Ba data. The code also generated a colored matrix...
plot summarizing results for every RRQ test for every detector, making it easy for an analyzer to visually inspect the results. Figure 6.2 shows a sample portion of such a matrix where it is obvious that the beginning of the cryogenic run did not have stable detector conditions.

Additionally, KS tests were used to compare dataset characteristics between Runs 127 and 128\cite{footnote}. Conservatively, we wanted to generate separate lookup tables for the four cryogenic runs. However, the $^{133}$Ba-calibration statistics for Run 128 were sufficiently low to degrade lookup table quality. Qualitative comparisons of RRQ distributions had already shown similarity between Runs 125 and 126 and separately for Runs 127 and 128. KS tests between data series of Run 127 and Run 128 rigorously showed similarity between the two runs in all but two detectors — T2Z5 and T4Z2. This enabled $^{133}$Ba-calibration data to be combined for Runs 127 and 128, except for these detectors, for which Run 128 WIMP-search data was not used.

### 6.3.3 Detector neutralization

Maintaining detector neutralization is critical for a WIMP search. Loss in neutralization reduces charge collection and increases the low-yield event rate, which can potentially introduce spurious WIMP candidates. As explained in Section 3.2.1, the detectors are grounded and bombarded with LED photons and radioactive-source gammas, at the start of a run and then periodically, to maintain neutralization. Periodic neutralization is performed every 11 hours during WIMP-search running, and after any $^{133}$Ba-calibration data series longer than 30 minutes. Note that this frequency is set to be shorter than the natural timescale for neutralization loss. As an additional precaution, however, all datasets are evaluated for neutralization state based on criteria listed below. Por-
tions of datasets deemed to have been acquired during poor detector neutralization are removed by cuts \texttt{cBadNeut\_bg\_c58} and \texttt{cBadNeut\_ba\_c58} for WIMP-search data and $^{133}$Ba-calibration data, respectively [143, 144].

The low-yield event fraction (ionization yield<0.8) is measured for chunks of calibration data with 100,000 events each, or for each WIMP-search data series. A chunk or series was removed if any of the following criteria were satisfied:

1. The low-yield fraction was greater than $2\sigma$ above the mean for the cryogenic run based on error bars for the chunk or series using binomial confidence intervals.
2. If the probability of seeing the actually observed number of low-yield events or more was lower than $10^{-10}$, given the mean low-yield fraction for the run.
3. For WIMP-search data, if the low-yield fraction was a $5\sigma$ outlier measured in units of average standard deviation from all data series.

A plot of low-yield fraction for a detector as a function of time is shown in Figure 6.3. Series discarded by the cut are marked.

![Figure 6.3: Low-yield fraction vs. time for WIMP-search data in T4Z2 in Run 125. Data series passing \texttt{cBadNeut\_bg\_c58} are marked with black points. Data series $2\sigma$ above the run mean are marked with red points and are removed. Series that fail the probability condition and the $2\sigma$ condition are marked with cyan points and are also removed. The $5\sigma$ run-averaged outlier level is marked with a blue dashed line and all series with low-yield fraction above this are removed as well. Courtesy: Oleg Kamaev.](image)

### 6.3.4 Noise and resolution

Sometimes the experiment saw elevated levels of noise, leading to degraded resolution and hence poor reconstruction of event timing and energy. The optimal filtering algorithm for both charge and phonons produces the inferred amplitude and timing resolution for all detector channels, for every data series. These are stored as RQs and can be used to set a series-dependent cut for the
entire run. Series with either amplitude or timing resolution in any of the channels greater than 25% over the mean for the entire cryogenic run were removed from the WIMP-search dataset using cut `cBadResTight`. A plot of phonon amplitude resolution as a function of time for data series is given for a detector in Figure 6.4.

Figure 6.4: Optimal-filter phonon-amplitude resolution normalized by run mean for datasets vs. time for T2Z3 in Run 125. We see that towards the end of Run 125, the amplitude resolution for a few data series crosses the 25% threshold (black dashed line) defined by `cBadResTight`. Courtesy: Sebastian Arrenberg.

Additionally, another observed pathology with possible repercussions for event reconstruction quality was periods of unusually high charge noise. These were identified by selecting all phonon noise events, and setting a 6σ cut on the 20-event moving average of summed charge energy. Periods of 10 minutes surrounding outliers were removed by cut `cHighQNoise`. Additionally, any series with a 20-series moving average of summed charge energy for these events above 4σ of the run mean were also excluded. A plot of charge noise events as a function of time is given in Figure 6.5.

Another mode of noise affecting charge channels on some detectors in the experiment is cryocooler noise, which rejects periods of high noise on an event-by-event basis, and will be covered under reconstruction quality cuts in Section 6.4.3.

### 6.3.5 Triggering and glitches

All events eligible to be WIMP candidates must have proper triggering, i.e., they should be recorded after a global trigger was issued because of the phonon trigger on a detector. Events that have problems with their trigger information are cut from WIMP search by `cErrMask`. Triggers randomly generated by the DAQ for the purpose of monitoring noise performance of the detector channels are also removed during WIMP-search analysis. Sometimes, glitches lead to excessive
simultaneous triggers on multiple phonon channels and are removed by the cut \texttt{cGlitch.c58} \cite{148}.

Trigger thresholds are also required to be stable for the entire data-taking period. Brief periods of elevated triggering are sometimes caused by electronics problems — this includes periods with overall trigger rate $> 0.7\,\text{Hz}$, and also individually identified periods of elevated triggering during Runs 125 and 127. These periods were removed by cut \texttt{cTrigBurst.c58} \cite{149}.

6.3.6 Miscellaneous data quality cuts

I note the following miscellaneous data quality cuts that are also applied to WIMP-search data:

1. **Stable detector tuning**: \texttt{cStabTuning.c58} ensures that only data series with correct QET biases and charge biases are selected \cite{150}. Very few data series, mostly experimental ones, are discarded by this cut.

2. **Bad detector regions**: \texttt{cBadDetRegions.c58} removes quadrants C and D for T3Z2 \cite{151}. T3Z2 demonstrates excessive low-yield events in those quadrants, because of a disconnected region of the outer-charge electrode in that half of the detector, leading to reduced charge collection in that area. In the absence of a clear way to control the low-yield events that mimic nuclear recoils, this half of the detector is excluded.

3. **Helium films**: Runs 125 and 127 were ended because a gradually increasing number of
detectors suffered from a high-trigger-rate pathology characterized by phonon pulses in the absence of charge pulses. It is believed that over long running times, the imperfectly sealed electronics box introduced helium into the CDMS-II cryostat at a very slow rate. Eventually, the helium condenses in superconducting thin films on the detectors and is accompanied by a decline in phonon-pulse amplitudes, and an increase in very-low-ionization recoils. The trigger rate eventually cripples livetime. The problem goes away only after cryogenic recycling. Cut cHeFilm_c58 removes periods of data-taking after the onset of helium films, on a detector-by-detector basis [152]. The criterion used for making the cut is a sustained increase in the rate of events with ionization yield < 0.1.

4. Neutron activation: The detectors are exposed to neutrons from a $^{252}\text{Cf}$ source a few times every run to provide nuclear recoils to define the signal acceptance region in ionization yield. This leads to activation of the copper surrounding the detectors and an increased ambient gamma rate by a factor of a few. While the half life of this elevated gamma rate is measured to be $\sim 12.7$ hours ($^{64}\text{Cu}$-decay dominant), all data taken for the first 48 hours after a $^{252}\text{Cf}$-source run is removed from the WIMP-search dataset, to reject increased surface-event rate from gamma comptons [153].

5. MINOS neutrino beam: A beam of GeV muon neutrinos is received from Fermilab at the MINOS far detector at Soudan, located in the cavern next to CDMS-II. Neutrinos may interact in CDMS’s shield or surrounding rock to produce neutrons. The residual background from this is expected to be negligible, but out of an abundance of caution, all events within a 60 $\mu$s window of beam firing are removed from WIMP-search data by cNuMI_c58 [154].

### 6.4 Reconstruction Quality Cuts

With all “good” WIMP-search data series identified, we place cuts on the reconstruction quality of individual events in the eligible data series. This is done for calibration data as well.

#### 6.4.1 Event reconstruction

The RQ $\text{QSOFchisq}$, which is the $\chi^2$ goodness-of-fit parameter from the charge optimal filter, is used to assess charge reconstruction quality on an event-by-event basis. This is done primarily to reject “pileups”, the occurrence of more than one particle hit in a single trace length of $\sim 1640 \mu$s. Such traces cannot be used to correctly estimate energy or timing. Pileups have large optimal filter $\chi^2$ values and can be removed by setting a cut on this parameter, as a function of energy. Intuitively, an energy dependence is not expected but is seen as shown in Figure 6.6 for a couple of reasons [71]. First, the template is not perfect — it is obtained by averaging several pulses, and at large energies,
the noise in the template makes the goodness-of-fit worse. Second, the charge pulse start time is
determined only to the nearest time bin, which is 0.8 \mu s. Sub-bin interpolation has been shown to
reduce the energy dependence of the \chi^2 value, but was not implemented in BatRoot in time for
Runs 125–128.

![Figure 6.6: QSOFchisq vs. summed charge energy for \(^{133}\text{Ba}\)-calibration data in T1Z2 in Run 125. The energy bin centers used for fitting are marked with crosses. The cut is marked by the line through the crosses. Data points under the line, marked in blue, are accepted and red ones, above the line, are rejected. Courtesy: Sebastian Arrenberg.](image)

The cut \texttt{cChiSq.c58} was defined in the following way. The \chi^2 distributions for each detector in
each run were binned into four energy intervals. Then the \chi^2 values in each bin were fit to a gaussian.
The 3.5\sigma \chi^2 values from each of the bins were fit to a quadratic to provide the energy-dependent
cut \cite{155}. The cut is shown in Figure 6.6 for detector T1Z2 in Run 125.

Note that the primary WIMP-search analysis for Runs 125–128 does not employ a phonon
reconstruction cut based on the phonon optimal filter goodness-of-fit parameter. This is because of
inherent phonon-pulse-shape variation, leading to large variation in \chi^2 values. In the case of the
charge optimal filter, the \chi^2 parameter can be used for a reconstruction quality cut because the
charge pulse is truly expected to have a known shape set by the electronics. However, I note that
analyses using the time-domain-fitting algorithms PipeFit and WedgeFit, described in Section 6.2.6
of \cite{104} use goodness-of-fit cuts \texttt{cPFGoodFit.c58} and \texttt{cWFGoodFit.c58} \cite{156}.

### 6.4.2 Phonon pre-pulse baseline

The second line of defense against pileup events is a cut on the standard deviation of pre-pulse
baseline of the phonon traces. This helps to identify traces that may contain residual long phonon
tails of events preceding the global trigger. Such traces need to be discarded for the same reason as
before — ambiguity in determining energy and timing information. Note that this pathology is not
tackled by the charge goodness-of-fit cut because ionization pulses have much shorter falltimes than
phonon pulses.

The cut \(C_{P\text{std}}\) rejects phonon-tail pileup events. It was made by fitting a gaussian to the
distribution of the standard deviation of the first \(\sim 400\,\mu s\) (stored as \(RQ\,P*\text{std}\)) of all traces of a
data series and then rejecting \(5\sigma\) outliers on a series-by-series basis \[157\,158\]. In past analyses, the
fits were made to the baseline standard deviation for the entire run, but the new technique prevents
entire data series from being penalized by shifts in noise, and hence baseline standard deviation, over
long time periods. Figure 6.7 shows \(P*\text{std}\) distributions for detector T1Z2 and the final cut for Run
125. The efficiency of this cut is \(\sim 99\%\) for most detectors and is accounted for detector-by-detector
and run-by-run in the final efficiency calculation.

6.4.3 Charge pre-pulse baseline

The charge pre-pulse baseline cut, \(C_{Q\text{std}}\), is similar in implementation to the phonon pre-
pulse baseline cut. However, the purpose of this cut is to reject events with high noise caused
by microphonic pickup from cryocooler mechanical vibrations. The cryocooler, attached at the e-
stem of the experiment, undergoes mechanical compression cycles with frequency 0.8 Hz, ringing
the system at all frequencies with every cycle. The phonon channels and most charge channels are immune to this issue, whereas the fiducial charge channels of detectors T1Z4, T3Z2, and T3Z5 and the outer charge channels of T1Z4, T2Z1, T2Z3, T2Z5, T3Z2, and T3Z5 are susceptible. Figure 6.8 shows noise spectra for T1Z4 during Runs 123–124 with and without cryocooler operation. The frequency and amplitude of cryocooler cycles are sufficiently varied for a blanket livetime cut at known intervals. Thus, the charge pre-pulse baseline, an indicator of cryocooler-induced noise, is the best identifier to tag and cut events affected by cryocooler cycling.

![Noise power spectra for T1Z4 in Runs 123–124 with (red) and without (black) cryocooler operation. The caption indicates the r.m.s. resolution in the two cases. Courtesy: Jeff Filippini.](image)

Similar to cPstd_c58, gaussians are fitted to the distribution of Q*std, the charge pre-pulse baseline for all traces of a data series, and then traces with Q*std > 4σ are removed on a series-by-series basis [157, 158]. Once again, the series variation of the threshold allows noise to drift over long time scales during a run, without penalizing entire series. Series with unusually high noise were already removed using the cuts described in Section 6.3.4. For detectors T2Z3 (fiducial), T3Z5 (fiducial and outer) and T4Z5 (fiducial), the thresholds were adjusted to 3σ, 3σ, and 2.5σ, respectively, to account for extended periods of high-frequency noise, which artificially inflate the Q*std distribution. The efficiency of this cut is ~ 99% for most detectors is accounted for detector-by-detector and run-by-run in the final efficiency calculation.

### 6.4.4 Phonon start time

A different pileup pathology is that of “cross-detector” pileup in high-event-rate calibration data, when a particle hit in a detector other than the triggering detector appears outside the optimal-
filter-search window. When the DAQ receives a global trigger, traces from all detectors are recorded within $\sim [-400, +1200] \mu s$ of the global trigger. During first-tier data processing, the reconstruction algorithms search for an ionization (phonon) pulse in all recorded traces only inside a window of $[-100, +10] \mu s$ ($[-50, +200] \mu s$) around the global trigger. An unrelated event on a non-triggering detector may not fall in the search window of the triggering event. The “walked” phonon timing and optimal-filter phonon energy for such a non-triggering event is typically reconstructed correctly because of a relatively large search window. However, the optimal filter ionization energy and start time might be misreconstructed, leading to anomalously long phonon delays and lower yields. Such events can be immune to other event reconstruction cuts. Figure 6.9 shows an example of an event where the charge pulse does not lie in the search window defined by the triggering detector.

![Figure 6.9: Charge and phonon traces of an event in detector T1Z3 where the trigger was issued by another detector. The charge pulse lies outside the search window. The true charge amplitude is 4.6 keV, but is reconstructed as 1.6 keV from the noise in the search window. Also the charge start time is assigned to be 416.8 $\mu s$. The reconstructed charge fit has a sufficiently low $\text{QSOFchisq}$ to pass the event reconstruction cut. Courtesy: Jeff Filippini.](image)

The cut $c\text{GoodPStartTime}$ rejects events that do not lie in the overlap of the charge and phonon-pulse-search windows, i.e., $[-50, +10] \mu s$. With few exceptions, the efficiency of this cut is 100%. Note that increasing the charge-pulse-search window in first-tier processing would be computationally expensive, and enforcing the overlap criterion of this cut during reconstruction would produce erroneous walk algorithm times. Thus far, the best solution seems to be enforcing the requirement after data processing. Also, this is not a problem for WIMP-search data with low event rates.

---

2 all detectors in WIMP-search mode; only triggering detectors in calibration mode, because of high data rates
6.4.5 Phonon manifold distance

As explained in Section 4.4.3, the distance of all events to the lookup table averaging bin used for their correction is recorded. In particular, two distances are measured by CorrTools during phonon-pulse-shape correction and recorded by PipeCalib as RRQs during data processing.

1. Distance to position manifold ($p\text{chisq}_{m, \text{rtft}}$):
   \[ d = \sqrt{\Delta x_{\text{part}}^2 + \Delta y_{\text{part}}^2 + \frac{\Delta x_{\text{del}}^2}{L_{\text{det}}} + \frac{\Delta y_{\text{del}}^2}{L_{\text{det}}}} \]

2. Distance to timing manifold ($p\text{chisq}_{t, \text{rtft}}$):
   \[ d = \sqrt{\Delta p_{\text{min}}^2 + \Delta p_{\text{del}}^2} \]

Note that these two distance measures are not used to determine the choice of nearest neighbors in the averaging bin. That is done by Equation 4.9 as described in Chapter 4. These two distance measures are used as a check on the correction quality because we found timing outliers to be correlated with large distances from their lookup table bins. Thus, we devised a cut $c\text{GoodRTFTManifold}_\text{c58}$ to reject the outlier tail in the sum of these two parameters with 97% efficiency for nuclear recoils [129]. The analog cut for PipeFit timing quantities is called $c\text{GoodPFManifold}_\text{c58}$, and uses distance analogs for PipeFit quantities.

6.5 Physics Cuts

This section includes cuts that restrict the pool of WIMP candidates because of physics arguments and define the parameter space within which we search for WIMPs. Note that one of these cuts, the surface-event-rejection cut is described separately in its own section (6.6).

6.5.1 Scintillator veto activity

Plastic scintillator veto paddles surrounding the experiment are used for identifying muons and secondary showers that can interact near the apparatus and the shielding to generate cosmogenic neutrons. Information from them is recorded in two ways. First, the time of any hit with energy deposition greater than a hardware threshold ($\sim 2\text{ MeV}$) in a panel is marked in the history buffer. Additionally, during WIMP-search runs, full traces from all panels are recorded in a $[-185, +25]\mu s$ window around the global trigger, regardless of whether it was a veto trigger or a ZIP trigger.

The expected muon rate at Soudan is about one per minute. However, the veto trigger rate is closer to 400Hz because of ambient gamma background. The ZIPs do not see such a high rate because of passive shielding around them, from which the veto panels do not benefit. If complete veto anti-coincidence were required for selection of WIMP candidates, we would lose a significant fraction of livetime, diminishing the experiment’s sensitivity. Instead, the veto coincidence cut, $c\text{VTStrict}_\text{c58}$, selects events that satisfy either of the following criteria [161]:
1. The maximum amplitude of the PMT trace from a veto panel lies above a calibrated threshold, in any of the panels in the full $[-185, 25] \mu s$ window around a global trigger in the experiment. The thresholds are tuned to reject gamma background, but accept muons and muon-related showers. The average threshold is $\sim 3.8 \text{ MeV}$.

2. The hardware threshold is crossed on a panel in a $[-50, 0] \mu s$ window of a global trigger.

Figure 6.10 shows an energy histogram of events registered in a top veto panel, along with the hardware and trace-amplitude thresholds. The efficiency of this cut, as computed from random triggers, is 97.964%. 2.036% of random triggers in the efficiency study were tagged as veto-coincident because of the large $50 \mu s$ window before the hardware trigger and the lower 2 MeV discrimination threshold that is susceptible to ambient photons.

![Energy Histogram](image)

Figure 6.10: Histogram of maximum amplitude of veto traces from a top panel during Runs 125–128 in WIMP-search mode. The green vertical dashed line indicates the hardware threshold, and the red vertical dashed line indicates the trace maximum amplitude threshold. The characteristic 10 MeV muon bump is visible. Courtesy: Matt Fritts.

### 6.5.2 Single scattering

WIMP candidates are expected to interact only once each in the experiment, given their vanishing interaction cross section. Background particles typically scatter more than once, making multiplicity tagging an excellent way to reject background. An event is classified as a single scatter if it deposits energy above threshold in only one detector, while the signal recorded in all other detectors is consistent with noise. This is enforced by cut $c_{\text{Single c58}}$.

For every data series, WIMP search or calibration, the optimal filter amplitude is determined for random trigger traces for all phonon and charge channels to determine noise distributions for every channel, i.e., the channel-wise distribution of energy estimates in the absence of particle hits. These
distributions are fit to gaussians, and a mean noise level and standard deviation is determined for each channel, for all data series in all four runs. An event is selected as a single scatter in a detector if the phonon energy of that event is $6\sigma$ above the series-wise noise distribution mean for that detector (summing the four phonon channels), and the phonon energy in all other detectors is within $4\sigma$ of the series-wise noise distribution means for the respective detectors. Checking whether other detectors have energy greater than $4\sigma$ above the noise mean is called “multiples tagging.”

The efficiency for this cut is measured using a clever technique. For a given detector, we take random-trigger events and measure the fraction with reconstructed energy greater than $4\sigma$ above the series-wise noise distribution mean in one or more other detectors. This gives the efficiency with which $c_{Single,c58}$ would mistag true single scatters. The series-dependent threshold allows events from series with slightly higher or lower noise than the run average to be included in the WIMP-candidate pool by adjusting thresholds accordingly. Despite this, the efficiency of $c_{Single,c58}$ was very poor for some data series and for some detectors, especially towards the end of Runs 125 and 127. So, the following following changes were made:

1. Detectors T1Z1, T5Z5, and T5Z6 used charge-based multiples tagging in addition to phonon-based tagging; i.e. $4\sigma$ above phonon OR ionization noise means on these detectors, in coincidence with a $6\sigma$ hit in another detector, also qualified an event as a multiple scatter.
2. For Run 125, T5Z1 used only charge-based multiples tagging and T2Z2 had its multiples threshold raised from $4\sigma$ to $5\sigma$.
3. For Run 127, T5Z2 is not used for multiples tagging.
4. For Run 127, detectors T2Z1, T5Z3, and T5Z4 showed very noisy periods towards the end of the run, where the noise distribution means for some series are $25\%$ higher than the run mean. For these series, if a phonon channel exhibited high noise for multiples tagging, it was replaced by charge-based multiples tagging. If more than two phonon channels exhibited high noise for multiples tagging, all phonon-based multiples tagging was replaced by charge-based multiples tagging.

After these corrections, the efficiency of $c_{Single,c58}$ is $> 99\%$ for all detectors in all runs except Run 127, where the efficiencies hover around $\sim 98\%$. The efficiency is plotted as a function of data series in Figure 6.11 for detector T1Z5 in Run 125.

### 6.5.3 Ionization yield

Nuclear recoils like WIMPs or neutrons are expected to have suppressed ionization yield compared to electron recoils. The signal acceptance region in ionization yield is defined by the cut $c_{NR,c58}$ as a $\pm 2\sigma$ band about the mean ionization yield as a function of recoil energy, for neutrons from
Figure 6.11: Efficiency of cSingle_c58 vs. time for detectorT1Z5 in Run 125. Courtesy: Tobias Bruch.

\(^{252}\)Cf calibration. This was implemented by binning the ionization yield of the calibration neutrons in recoil energy, fitting each to a gaussian to determine the bin-wise mean and standard deviation, and then fitting recoil-energy-dependent polynomials to the mean and width of the band \([164]\). This cut is slightly different for each of the runs because of slight changes in phonon response between cryogenic runs. The nuclear recoil cut for a typical detector is shown in Figure 6.12. The efficiency of the cut is measured by taking the fraction of \(^{252}\)Cf-calibration neutrons contained in a ±2σ band to that contained in a ±4σ band, and is \(\sim 94\%\). Also, based on expected gamma leakage into this band (see Section 7.2), it is required that WIMP candidate events be below \(-3\sigma\) of the mean electron-recoil ionization yield as a function of energy. The efficiency of this additional requirement is taken into account for computing exposure. The electron-recoil band is also determined in a manner similar to the nuclear-recoil band, using \(^{133}\)Ba-calibration gammas.

It turns out that this band definition is not optimal to maximize signal acceptance while minimizing surface event background rejection. After the release of results, Joseph Kiveni showed that the optimal nuclear-recoil acceptance band for Runs 125–128 was \([-1.9\sigma, +1.5\sigma]\) about the mean ionization yield \([165]\). This was computed using techniques I developed and present in Section 6.6 for optimizing signal acceptance in phonon pulse timing.
115

Figure 6.12: Ionization yield vs. recoil energy (keV) for $^{252}$Cf events (green dots) for T1Z5 in Run 125. The solid red lines depict the $\pm 2\sigma$ band about the mean ionization yield for neutrons, which is marked with a red dotted line. The blue line depicts the $-3\sigma$ edge of the electron recoil band. Note that energy threshold cuts, discussed in Section 6.5.5, are not applied in this plot.

6.5.4 Fiducial volume

The charge fiducial volume cut, $c_{Qin \_c58}$, rejects events with signal in the outer electrode, which may suffer incomplete charge collection because of fringing fields. This is done by requiring that candidate events have energy in the outer-charge electrode of each detector consistent with $\pm 2\sigma$ of the series-specific noise for the electrode $^{[166]}$. Outer-charge electrode energy is binned by fiducial charge energy and gaussians are fitted to each of the bins to determine the bin-wise mean and standard deviation. Then, the outer-charge noise band is defined by fitting polynomials to the bin-wise mean and standard deviation. A sample fiducial volume cut in the outer-charge vs. fiducial-charge plane is shown in Figure 6.13.

In practice, this cut becomes restrictive at high charge energies because the probability of multiple scattering within a detector, and hence distributing charge between the fiducial and outer electrodes, increases with energy. Thus we see that the efficiency of $c_{Qin \_c58}$ is $\sim 75\%$ at low energies, but decreases with energy. This results in a systematic error in computing the fiducial volume efficiency for WIMPs; $^{252}$Cf-induced neutrons can multiple scatter in the detector and share energy between the outer and fiducial electrodes, whereas WIMPs would uniformly single scatter in a detector. A GEANT4 Monte Carlo simulation showed that the fiducial volume measured by neutrons in our calibration configuration is $95.48 \pm 0.82$(stat)$^{+0.65}_{-0.45}$(sys)$\%$ of the true WIMP fiducial volume, averaged over 10-100 keV and over the germanium detectors in Runs 125–128 $^{[167]}$. This is taken into account as a multiplicative factor in the computation of the analysis efficiency.
6.5.5 Thresholds

Two types of energy thresholds are imposed on potential WIMP candidates and are described below.

6.5.5.1 Ionization energy threshold

Events with charge signal in the fiducial electrode consistent with noise are excluded from the WIMP-search dataset. The data-series-wise noise distributions computed for \( c_{\text{Single}} \) in the ionization channel are used for the computation of charge threshold as well. The cut \( c_{Q_{\text{Thresh}}} \) requires events to have fiducial charge energy > 4.5\( \sigma \) above the noise distribution mean for the series the event belongs to. The series-wise noise thresholds protect the search from worse resolution in noisier runs, by automatically adjusting to a higher threshold. However, a broader requirement for the fiducial-electrode energy to be > 4.5\( \sigma \) of the noise distribution mean for the run is also imposed, and the greater of the two thresholds is used on a series-by-series basis for conservative selection [28]. The number 4.5\( \sigma \) was chosen to limit noise leakage to below 0.01 events per detector per cryogenic run.

6.5.5.2 Recoil energy range

A threshold is also set on the recoil energy. In this case, it is not sufficient to set a cut based solely on the phonon noise distributions. The threshold for this analysis was set to 10 keV, for all
14 germanium detectors used, for the following reasons:

1. At low recoil energy, surface-event discrimination and ionization-yield discrimination are diminished because timing information is derived from noisy quantities. Additionally, the phonon-pulse-shape correction becomes dominated by noise below \( \sim 8\text{--}10 \text{keV} \).

2. Below 10 keV, the \(-5\sigma\) edge of the electron-recoil band impinges on the \(\pm 2\sigma\) nuclear-recoil band, reducing the available statistics of \(^{133}\)Ba-calibration-induced surface events to tune surface event rejection cuts. At 10 keV, there is sufficient statistics on all detectors.

3. The gamma leakage into the nuclear-recoil band is \(< 0.001\) total events on all detectors if the threshold is set to 10 keV [169].

The selection of events is truncated at 100 keV because of diminished expected contribution to the event spectrum from standard WIMPs and because of insufficient calibration data at high energies. As an aside, the low-energy threshold analysis of Runs 123–128 does not employ surface event discrimination, and uses the phonon noise distributions to set the threshold around 2 keV [1]. Also, a special analysis for inelastic dark matter models permits WIMP candidates up to 150 keV because of the higher energy of the expected spectrum [170].

A scatter plot of ionization yield as a function of recoil energy is shown in Figure 6.14, showing all physics cuts applied to \(^{252}\)Cf-calibration neutrons for a detector.

![Figure 6.14: Ionization yield vs. recoil energy for \(^{252}\)Cf neutrons in T3Z2 for Run 125. The nuclear-recoil band is marked in red. The \(-3\sigma\) electron recoil edge is marked in blue. The charge threshold is marked with a green line, and the recoil energy threshold is marked with a magenta line. Events passing physics cuts are marked with black points. Note that the surface-event rejection cut has not been applied here.](image-url)
6.6 Surface-Event Rejection

Surface events are the dominant misidentified background for ZIPs because of their suppressed ionization yield and must be excluded from the WIMP-candidate set. As I showed in Chapter 4, phonon timing shows discrimination between surface events and bulk events, even more so after the application of phonon-pulse-shape correction. In this section, I review the successes and drawbacks of the techniques used to set the surface-event rejection cut in past analyses, and explain the improvements made and techniques employed for Runs 125–128.

6.6.1 Calibration data

First, I define the selection of calibration nuclear recoils and surface events typically used in any CDMS analysis, including that for Runs 125–128. All the data quality measures, reconstruction cuts, and physics cuts listed in this chapter provide the framework to do this. The two event groups are selected in the following way:

1. **Nuclear-recoil sample**: Events passing all reconstruction cuts and all physics cuts from $^{252}$Cf data series that pass data quality cuts provide the nuclear recoil sample.

2. **Surface-event sample**: Events passing all reconstruction cuts; below the $-5\sigma$ edge of the electron recoil band and with ionization yield between 0.1 and 0.7; within fiducial volume; above energy thresholds; and from $^{133}$Ba data series that pass data quality cuts constitute the surface event sample.

Note that since ZIPs are closely stacked in towers, surface events on the phonon-sensor-instrumented side (simply called phonon side), and the uninstrumented side (charge side), are easily distinguished by looking for a coincident event in the detector above (phonon side) or the detector below (charge side). Two-thirds of all calibration surface events can be identified as fitting one category or the other by this nearest-neighbor tagging.

6.6.2 Past techniques and lessons

Previous analyses [121, 132] have succesfully used the simple sum of $p_{\text{min}} + p_{\text{del}}$, or a chi-squared distance constructed from these two variables and a partition-derived RRQ, as the primary discrimination variable. The cuts were set on this variable by requiring a target leakage rate, say 0.005 surface events per nuclear recoil per detector, on half the calibration data (selected by picking every other event sequentially from the entire run). The other half was used to provide a statistical check against overtuning the cut. There were a few problems with this scheme:

1. There were never sufficient surface events in half the calibration dataset to set the cut to required leakage level. For most detectors, the cut was simply set at the timing value that
excluded the slowest calibration surface event. Subsequently, the leakage on the second half of the dataset was consistent with the first or showed higher leakage.

2. The leakage measured on WIMP-search sidebands — single scatters outside the nuclear-recoil band, or multiple scatters inside the nuclear-recoil band — always showed a factor of a few higher surface-event leakage than the calibration data on which the cut was set. Thus, the cut always required tightening, but with no scale for the amount of retuning required, since there were no calibration surface events left to measure leakage on. The retuning was typically done by tightening the cut by an arbitrary factor, say 1.5x or 2x.

Rudimentary studies done during Run 119 and Runs 123–124 suggested that systematic differences in the surface-event distributions of WIMP-search data and calibration data cause the difference in measured leakage rates. In this analysis, we studied these differences further and attempted to correct for them before setting the surface-event rejection cut. Also, in order to avoid the problem of low surface-event statistics in the calibration data, we abandoned the division of the calibration dataset into two, and fit the tails of the timing distributions, enabling us to estimate leakage as a continuous function of a phonon timing variable.

### 6.6.3 Systematic differences between calibration and WIMP-search surface events

Since there were no changes to background rates between Runs 123–124 and Runs 125–128, we could use surface events from the entire unblinded R123-4 WIMP-search dataset and directly compare with calibration surface events to quantify systematic differences. Two studies have shown systematic differences on the same two fronts \[171, 172\]:

1. **Energy**: The most important systematic difference between calibration and WIMP-search surface events is energy distribution. As shown in Figure 6.15, WIMP-search data has a larger fraction of events at low energy, compared to calibration data. This is important because low-energy surface events have a longer slow-timing tail than high-energy events, as shown in Figure 6.16.

2. **Ionization yield**: ZIPs have a slightly different response to phonon-side surface events compared to charge-side surface events. This was shown for phonon-energy partition in Figure 4.7, but is also true for phonon timing and ionization yield. Charge-side surface events have higher ionization yield, but a longer slow-timing tail than phonon-side surface events, as shown in Figure 6.17 for calibration data. Thus the other important systematic difference between WIMP-search data and calibration data arises from the difference in numbers of phonon-side and charge-side events between the two datasets. The ratio of charge-side to phonon-side
Figure 6.15: Left: Recoil energy spectrum of surface events from the 12 inner germanium detectors used for WIMP search in Runs 125–128. The blue curve shows calibration surface events and the red curve shows WIMP-search surface events within the nuclear-recoil band for Runs 123–124.

Figure 6.16: Tail distribution of corrected $p_{\text{minr}} + p_{\text{del}}$ normalized by subtracting the mean for calibration surface events. The blue line is the tail for low-energy events (10–30 keV), and the red line is the tail for higher-energy events (30–100 keV).
surface events in the WIMP-search nuclear-recoil band is larger compared to the ratio of charge-side to phonon-side surface events in calibration data as seen by comparing figures 6.17 and 6.18.

![Figure 6.17](image1.png)

**Figure 6.17:** Left: Yield distribution (normalized by nuclear-recoil-band standard deviation), for charge-side (red) and phonon-side (black) surface events in calibration data. Right: Tail distribution of corrected $p_{\text{minrt}} + p_{\text{del}}$ normalized by subtracting the mean for charge-side (red) and phonon-side (black) calibration surface events. Both distributions are summed over the 12 internal germanium detectors used for this WIMP search.

![Figure 6.18](image2.png)

**Figure 6.18:** Yield distribution (normalized by nuclear-recoil-band standard deviation), for charge-side (red) and phonon-side (black) surface events in WIMP-search data from Runs 123–124, summed over the 12 internal germanium detectors used for WIMP search in Runs 125–128. The ratio of charge-side to phonon-side surface events is different in this distribution compared with that of calibration data displayed in Figure 6.17.

### 6.6.4 Estimating leakage, accounting for systematic differences

To account for the systematic differences between WIMP-search and calibration data explained above, I proposed the following estimator for surface-event leakage on a detector, $n_i$, after consulting
with Jeff Filippini and Sunil Golwala:

\[ n_i = N_i \sum_{e,f} b_{e,f}^{(i)} s_{e,f}^{(i)} \]  

(6.1)

where \( N_i \) is the expected number of single-scatter surface events on the \( i \)-th detector, \( b_{e,f}^{(i)} \) is the detector-specific passage fraction measured on calibration data for energy bin \( e \) and detector face \( f \), and \( s_{e,f}^{(i)} \) is the detector-specific measured fraction of events of that class in WIMP-search data. In effect, this formula takes the calibration-measured leakage in each of the bins and reweights it in accordance with the occurrence of that type of event in WIMP-search data. The surface events of the WIMP-search data from Runs 123–124 were taken and divided into four bins to compute \( s_{e,f}^{(i)} \):
- Charge-side with energy 10–30 keV; charge-side with energy 30–100 keV; phonon-side with energy 10–30 keV; phonon-side with energy 30–100 keV [172]. Since there were no WIMP candidates in Runs 123–124, the count of events passing WIMP-candidate cuts except for the timing cut provided a direct measurement of the background rates of the detectors. That scaled by the livetime for Runs 125–128 provided \( N_i \). Thus, by measuring \( b_{e,f} \) on calibration data as a function of timing discriminator such as \( \text{pminrt+pdel} \), one could set \( n_i \) for each detector to the desired level.

This last step turned out to be a little complicated. Even with the large statistics of the undivided \(^{133}\text{Ba}\)-calibration dataset, some categories of surface events such as low-energy charge-side events had poor statistics. So, to estimate \( b_{e,f} \) in far-out regions of the tails, we decided to fit the distribution tails. It had been shown that the tails could be fit fairly well using generalized Pareto distributions [173]. Thus a Pareto fit was performed on the last \( \sim 10 \) events of the CDF of the distribution, and the remainder was smoothed using a gaussian kernel. Then the two pieces were matched using a cubic-spline interpolation. The resulting fits were better than anyone expected. A sample is pictured in Figure 6.19. We then verified that by applying this scheme to calibration data for Runs 123–124, we could correctly match the WIMP-search leakage prediction from calibration data to that from WIMP-search sidebands for Runs 123–124 [174].

6.6.5 Cut setting algorithm

The next question was how to optimally set the maximum allowed leakage \( n_i \) on each detector. The easy solution was to require equal leakage on all detectors, say 0.07 events per detector, for 14 detectors to give a total of 1 leaked surface event. This could be implemented with a minimization algorithm to manipulate the timing cut and set \( n_i \) to 0.07 events. However this scheme did not take into account the WIMP-search reach of the individual detectors. Some detectors, if allowed slightly more surface-event leakage, could increase exposure much more by subtracting that leakage elsewhere on a more contaminated detector with less sensitivity. In the past, schemes had been proposed to optimize the total leakage and maximize exposure over all detectors [173, 175]. Combining their
Figure 6.19: Top pane shows the fit of the calibration surface-event $p_{\text{minrt}}$ distribution for low-energy, charge-side events in T1Z5 for Runs 125–128. The bottom left pane shows a closeup of the cubic-spline matching of the gaussian-smoothed CDF and the Pareto-tail fit. The bottom right pane shows a close up of the Pareto fit to the tail. Data are blue (along with error bars), the fit is shown in red before interpolation, and green afterwards. The cyan vertical line indicates the position where the smoothed CDF and fit are joined. The $\pm 1\sigma$ error bars on the fit are shown in dashed green. 

*Courtesy:* David Moore.
ideas with our new leakage estimation scheme, I developed a new package [176] in MATLAB and Matt Fritts developed one [177] in ROOT to accomplish this. The basic scheme in MATLAB is as follows:

1. Fit calibration surface-event timing CDFs as described in the previous section providing \( b_{e,f} \).

2. Estimate \( N_i \) from Run 123–124, by scaling observed surface-event counts on the detectors by Run 125–128 livetime.

3. Compute kernel-smoothed Spectrum-Averaged Exposure, \( SAE_i \), for each detector for a 60-GeV/c\(^2\) WIMP, as a function of timing cut position. This quantity is defined in Equation 6.3 in Section 6.7 and is a measure of the sensitivity of a detector.

4. First find the cut positions that equalize leakage on all detectors. This is done with the \texttt{fminsearch} function in MATLAB. This provides the starting point for optimizing the \( \sum SAE_i \) while constraining \( \sum n_i \). Also, a provision was made for detectors with starting cuts already set to the maximum allowed by the tail fit and contributing greater than 15% of the total leakage, to be immediately removed. No detectors were removed in this step, but a significant fraction of the exposure of the endcap detectors, T3Z6 and T4Z6, was discarded in the last step.

5. Using the objective function below, a numerical optimization algorithm in MATLAB called \texttt{patternsearch} was used to maximize \( \sum SAE_i \) while constraining \( \sum n_i \) to 0.1, 0.2, 0.3...2 events. The cut positions from the previous step are used as the starting condition.

\[
f(t) = \left( 1 - \frac{\sum_i SAE_i(t)}{\left( \sum_i SAE_i \right)_{max}} \right)^2 + \left( 1 - \frac{\sum_i n_i(t)}{\left( \sum_i n_i \right)_{target}} \right)^2
\]  
(6.2)

The first term seeks maximum summed exposure and second term seeks the target leakage. In practice a weighting factor is added to the second term to set a tolerance on achieving the target \( \sum n_i \).

Finally, to compare the performance of candidate timing cuts, one could compare the final spectrum-averaged exposure, the 90\% CL limit on the WIMP Poisson rate, or the median Optimum Interval expected sensitivity to a 60-GeV/c\(^2\) WIMP. This also automatically provided the target leakage for best WIMP-search reach. I did this for WIMP-search data from Runs 123–124 to verify the technique and showed that a 15\% improvement in limits could have been obtained over the simple timing cut used for that analysis [176] (see Figure 6.20). We also note that 0.5–0.6 is the target surface-event leakage for CDMS II for maximal sensitivity, ignoring contributions from other backgrounds.
Figure 6.20: Left: 60-GeV/c^2 WIMP spectrum-averaged exposure vs. target leakage in optimization. Blue curve shows equal contributing surface-event leakage from each detector, and red curve shows optimized result. The original result from Runs 123–124 is marked with a black data point. Right: 90% CL Upper limit on WIMP rate for the same two scenarios as well as a comparison with the original result from R123–124. Note that the best limit is obtained for target leakage of ~ 0.5–0.6 events.

6.6.6 Surface-event rejection cut for Runs 125–128

Using the techniques outlined above, I prepared a cut using $p_{\text{minrt} + p_{\text{del}}}$, called cRT_{vanilla,c58}, with a target total leakage of 0.5 events [178]. This cut was chosen as the primary surface-event rejection cut for Runs 125–128. As I explain in the next chapter, the cut required one more tuning based on a proposed change in the energy bins of Equation 6.1. In addition to requiring $p_{\text{minrt} + p_{\text{del}}}$ to be greater than some detector-wise threshold, there is also a requirement for it to be $< 99.5$th percentile timing value for the calibration neutron distribution. Also, a nuclear-recoil “consistency” requirement on $p_{\text{minrt} - p_{\text{del}}}$ enforces this value to be between the 0.5th and 99.5th percentile for calibration neutrons. The cut acceptance region is plotted for all 14 detectors in Figure 6.21. Finally, the efficiency of this cut is typically $\sim 60\%$ on interior detectors, but as low as $\sim 10\%$ on the endcaps (T3Z6, T4Z6), because of their inability to reject multiple scatters on their charge faces.

6.6.7 Other candidate cuts

Instead of using the simple linear sum of a primary risetime estimator and a primary delay estimator, one can also construct for any event a difference of chi-squared distances to the surface-event distribution and the nuclear-recoil distribution using calibration data. A cut on this parameter can be made by using the leakage estimation and the exposure optimization schemes outlined earlier. The drawback of this technique is that it assumes a gaussian distribution for the risetime and delay parameters for both surface events and nuclear recoils, which is incorrect. The chi-squared cut I attempted to make had too many outliers in the chi-squared distance tails, making the cut performance much worse than cRT_{vanilla,c58}. 
Figure 6.21: Normalized $p_{\text{minrt}}$ vs. normalized $p_{\text{del}}$ for all 14 germanium detectors used for WIMP-search in Runs 125–128, using calibration events. Neutrons are plotted with green points, charge-side surface events with red points, phonon-side surface events with blue points, and untagged surface events with black points. The area included within the four dashed back lines on each subplot denote the acceptance region for $cRT_{\text{vanilla,c58}}$. Surface events in the acceptance region are marked with magenta circles. Note that the cut position in $p_{\text{minrt}} + p_{\text{del}}$ is set after reweighting systematic differences between calibration data and WIMP-search data and maximizing the exposure on detectors with low inherent surface background. Thus the number of events seen passing the cut here is not representative of the true surface-event leakage.
It is worth noting that Matt Fritts and Lauren Hsu led an effort to investigate the discrimination performance of primary risetime and primary delay obtained from PipeFit and WedgeFit as well. They investigated various weighted linear combinations of these and walked timing parameters on a detector-by-detector basis [177, 112, 104]. We compared the performance of the six most promising cuts using simulated datasets from the CDF fits. As shown in Figure 6.22 the performance was within a few percent of each other [179].

Figure 6.22: Left: Estimated total leakage for six candidate timing cuts optimized with 0.5 events of target total leakage. Right: 60-GeV/c^2 WIMP spectrum-averaged exposure events for the same cut. Blue points use the real calibration dataset, while the red points are simulated datasets, generated using the fitted CDFs. Courtesy: David Moore.

### 6.7 WIMP-search Exposure

To compare the efficiency of different analyses or even different WIMP searches, one needs a standard way of quoting the WIMP-search exposure. One way is to multiply the exposure, i.e., the detector mass times the livetime, by the WIMP-recoil-spectrum weighted efficiency of the experiment or analysis, for a WIMP of a certain mass. This is called the spectrum-averaged exposure and can be written as follows:

\[
SAE = MT \int_{E_{\text{min}}}^{E_{\text{max}}} dE \epsilon(E) \frac{dR}{dE}
\]

Here \(MT\) is the target mass times the livetime, i.e., the exposure, \(\epsilon(E)\) is the energy-dependent analysis efficiency, and \(dR/dE\) is the differential recoil spectrum for a WIMP. The integrals run over the energy range of the analysis. Note that the expression for spectrum-average exposure in Equation 6.3 was used in the timing cut optimization technique of Section 6.6.5 with approximations for \(M, T\), and \(\epsilon(E)\). In this section, the spectrum-averaged exposure for a 60-GeV/c^2 WIMP is reviewed, to measure the performance of this analysis.
6.7.1 Exposure

The average mass of germanium ZIPs in CDMS is \( \sim 230 \) grams. The exact masses are calculated based on dimensional measurements of each of the detector crystals made prior to installation in CDMS II [180]. Next, we add up the livetime for all runs for each detector separately. This is done by summing over all events in the entire WIMP-search dataset an RQ called \textit{LiveTime} that measures the time the DAQ was live, waiting for that event hit. Note that prior to computing livetime, all data quality cuts (Section 6.3) and the phonon and charge baseline cuts (Sections 6.4.2 and 6.4.3) are applied to the WIMP-search data. The product of the detector masses and their individual livetimes gives the exposure and is listed in Table 6.1. The total exposure for Runs 125–128 was 612.13 kg-days [181]. Without any data quality cuts, this number is \( \sim 833 \) kg-days. The biggest losses in livetime occur because of cBadDet\_bg\_c58, cHeFilm\_c58, and cQstd\_c58 [182].

<table>
<thead>
<tr>
<th></th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z2</td>
<td>48.11</td>
<td></td>
<td>17.73</td>
<td>44.45</td>
<td></td>
</tr>
<tr>
<td>Z3</td>
<td></td>
<td></td>
<td>34.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z4</td>
<td></td>
<td></td>
<td>55.27</td>
<td>51.02</td>
<td>47.91</td>
</tr>
<tr>
<td>Z5</td>
<td>36.49</td>
<td>43.46</td>
<td>34.89</td>
<td>46.85</td>
<td>49.33</td>
</tr>
<tr>
<td>Z6</td>
<td></td>
<td></td>
<td></td>
<td>50.60</td>
<td>51.35</td>
</tr>
</tbody>
</table>

Table 6.1: Exposure (kg-d) for Runs 125–128 in germanium detectors used for WIMP search.

6.7.2 Efficiency

The efficiency functions of the remaining reconstruction quality cuts (Section 6.4), physics cuts (Section 6.5), and the surface-event rejection cut (Section 6.6) are computed and sequentially multiplied to give the analysis efficiency as a function of recoil energy, for every detector [181]. Where possible, this is done using functional fits of measured cut efficiencies as a function of recoil energy. Run-by-run efficiency functions are weighted by livetime for each run. The net analysis efficiency is computed by combining the efficiency functions of all 14 detectors weighted by exposure, and is plotted as the green line in Figure 6.23. The 60GeV/c^2-WIMP-spectrum-averaged exposure for Runs 125–128 is 185.3 kg-days.
Figure 6.23: Analysis efficiency vs. recoil energy for Runs 125–128. Reconstruction quality cuts have the efficiency given by the black line. That combined with the nuclear-recoil-band efficiency give the blue line. These two combined with the fiducial-volume-cut efficiency give the red line. All of these combined with the surface-event-rejection cut efficiency give the analysis efficiency, i.e., the green line. Error bars on the efficiency curve fit are depicted by magenta errorbars. For comparison, the analysis efficiency of Runs 123–124 is depicted by the dashed black line.
Chapter 7

WIMP-search Results

In this chapter, I review the expected background leakage from neutrons, photons, and surface events, and go over the results of unblinding the signal region. Two events are observed in the signal region, but do not constitute statistically significant evidence for WIMPs. Upper limits on WIMP-nucleon scattering cross section are presented with the conservative assumption that the events are WIMPs, i.e. without any background subtraction. Note that throughout this chapter, “WIMP-search events,” refer to veto-anticoincident events in WIMP-search datasets that pass data quality cuts, are within the detector fiducial volume and exceed recoil- and ionization-energy thresholds. Thus signal-candidate events are called WIMP-search nuclear-recoil single scatters passing surface-event rejection.

7.1 Expected Neutron Background

Neutrons cannot be distinguished from WIMPs in ZIP detectors, making them a background. However, as explained in Chapter 5, passive shielding moderates neutrons originating from outside the experiment, and the scintillator veto tags most of the ones that make it through. This significantly reduces the rate of misidentified neutrons in the ZIP detectors. Nonetheless, we carefully study and document the residual neutron background rate for its two sources — radiogenic neutrons and cosmogenic neutrons. Detailed accounts are provided in [183, 184, 104]; here I give an overview.

7.1.1 Radiogenic neutrons

Radiogenic neutrons, i.e ones generated by radioactive processes in materials surrounding the experiment, have several sources. Neutrons are generated by natural fission of U and Th nuclei or by ($\alpha$, n) processes. Outside the experiment apparatus, U and Th are found with natural abundance in the cavern rock surrounding it. Typical neutron energies from these sources are as high as $\sim 6 \text{ MeV}$, but are moderated by the passive polyethylene shield around the experiment to energies far below that required to generate recoils above the recoil energy threshold of the experiment. Thus the
largest source of radiogenic neutrons is radioactive processes occurring in materials within the shield and surrounding the ZIPs. Note that the U/Th impurity level in the ZIPs themselves is negligible compared to other sources (<< 1 ppt level), leading to a negligible contribution to the expected residual neutron rate.

The natural abundance of U and Th in materials around the ZIPs has been ascertained using two techniques. The first is radiopurity screening of materials using High-Purity Ge (HPGe) and other gamma counters. The second is by matching the gamma spectra observed in the ZIPs to those obtained from Monte Carlo simulations of decay chain propagation in the experimental setup, with free parameters for U, Th, K and Co contamination levels on the icebox cans and cold hardware (Section 8.1.1 of [186]). Note that both these techniques assume secular equilibrium between isotopes we actually measure and the ones earlier in the decay chain. Our best estimates of U/Th contamination for materials surrounding the experiment are listed in Table 7.1.

<table>
<thead>
<tr>
<th>Component</th>
<th>U (ppb)</th>
<th>Th (ppb)</th>
<th>Mass (kg)</th>
<th>Rate in Ge (kg-y)^{-1}</th>
<th>Expected count in R125-8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Copper cold hardware</td>
<td>0.16</td>
<td>0.25</td>
<td>14.95</td>
<td>3.6 × 10^{-3}</td>
<td>4.2 × 10^{-3}</td>
</tr>
<tr>
<td>Copper icebox cans</td>
<td>0.18</td>
<td>0.56</td>
<td>260</td>
<td>2.3 × 10^{-2}</td>
<td>2.7 × 10^{-2}</td>
</tr>
<tr>
<td>Inner polyethylene</td>
<td>&lt; 0.12</td>
<td>&lt; 0.12</td>
<td>120</td>
<td>&lt; 6 × 10^{-3}</td>
<td>&lt; 7 × 10^{-3}</td>
</tr>
<tr>
<td>Inner (ancient) lead</td>
<td>&lt; 0.05</td>
<td>&lt; 0.2</td>
<td>1917</td>
<td>&lt; 3 × 10^{-3}</td>
<td>&lt; 3.5 × 10^{-3}</td>
</tr>
<tr>
<td>Outer lead</td>
<td>&lt; 0.05</td>
<td>&lt; 0.4</td>
<td>12190</td>
<td>&lt; 7 × 10^{-3}</td>
<td>&lt; 8.2 × 10^{-3}</td>
</tr>
</tbody>
</table>

A second set of simulations provides the expected nuclear-recoil single-scatter rate in the ZIPs in our experimental setup, given the U/Th abundance rates in materials surrounding the experiment. For the copper components and the inner polyethylene, neutron rates were computed in SOURCES4A [187], and propagated in ZIPs using GEANT4 [188]. For the lead shield surrounding the experiment, a simulation was performed entirely in GEANT4, using a 252Cf neutron spectrum as an approximation for fission neutrons from U/Th [189]. The results are listed in the last two columns of Table 7.1. Multiplying the rates by the exposure of Runs 125–128, 612 kg-days, and the neutron efficiency in analysis cuts, we get the total expected count of 0.04$^{+0.00}_{-0.02}$.

### 7.1.2 Cosmogenic neutrons

Cosmogenic neutrons result from muon spallation or muon-related particle showers in the materials surrounding the experiment. At the depth of the Soudan Underground Laboratory, the muon flux is suppressed by a factor of $\sim 2 \times 10^{-5}$ compared to the surface, which provides the first line of protection against cosmogenic neutrons. An active scintillator veto surrounding the experiment is able to tag residual through-going muons with > 99.9% efficiency [102]. However, not all muons deposit energy in the veto. Also, particle showers coincident with cosmogenic neutrons might not
always interact in the veto or other ZIPs in the experiment.

The rate of untagged cosmogenic nuclear recoils in the ZIPs was studied using several Monte Carlo simulations. Two different muon simulation codes were used to produce muons with the correct flux and angular distribution for Soudan [190, 191]. These muon distributions were then thrown at 10 m of cavern rock surrounding the experiment to generate three sets of simulated particle showers (two sets with GEANT4 and one with FLUKA-MCNPX) with different lengths of simulated livetime ranging between ∼ O(10) and ∼ O(100) years [192, 193, 184]. From the simulations, we recorded the counts of veto-anticoincident nuclear-recoil single scatters, i.e. nuclear recoils without interactions in other ZIPs or the veto. We also recorded nuclear recoils in coincidence with veto and other ZIP hits. This gave estimates of the mean ratio of unvetoed to vetoed cosmogenic nuclear recoils, 0/9, 2/36, and 33/561 (first two from GEANT4, third from FLUKA-MCNPX). The estimates were consistent with one another and were combined to mitigate systematic differences between the simulations and enhance statistics, giving 35/606. To avoid systematic errors in using muon and particle shower flux rates from the simulation, the ratio is normalized by the observed count of vetoed nuclear recoils from data. We use the livetime from Runs 123–128, 1110 kg-days, in which three vetoed cosmogenic neutrons were observed. This gives a raw unvetoed single-scatter cosmogenic neutron rate of $3/1110 \times 35/606 = 1.6 \times 10^{-4}$ events per kg-day. Scaling this by the livetime for Runs 125–128, 612 kg-days, multiplying by the approximate analysis nuclear-recoil efficiency (∼ 50%) and taking into account statistical uncertainties, we get $0.04^{+0.04}_{-0.03}$ events as expected leakage.

7.2 Expected Photon Background

It is possible that some non-negligible fraction of photons may leak into the ±2σ nuclear-recoil band. Since photons typically have bulk-event-like timing and > 12σ separation from nuclear recoils in ionization yield, this effect is expected to be purely statistical, and should occur only at low energies because of the flaring of the electron-recoil and nuclear-recoil bands. A Monte Carlo simulation was done to produce $10^8$ photons, spectrum matched to the WIMP-search gamma spectrum [169]. The resolution of the 10.36 keV line was used to smear the phonon and ionization energies. From the simulation one could directly determine the leakage into the nuclear-recoil band as a function of recoil-energy threshold and electron-recoil-band cut, as shown for a detector in Figure 7.1. This number was computed to be < 0.01 events for Runs 125–128.

7.3 Expected Surface-Event Background

While cRT_vanilla_c58 was tuned using the expected surface-event leakage on calibration data, the performance of the cut can also be independently measured on two datasets — WIMP-search
multiple scatters inside the nuclear-recoil band, and WIMP-search multiple scatters outside the nuclear-recoil band. A combination of all three estimates of surface-event leakage provides the final number. Note that another possible category, WIMP-search single scatters outside the nuclear-recoil band, suffers from low statistics in addition to energy and face distribution systematic differences from WIMP-search nuclear-recoil single scatters. Thus this category is not used to produce an estimate. The following describes the procedure that David Moore and I undertook to compute this leakage estimate.

### 7.3.1 Premise

The basic premise of measuring surface-event leakage detector is given by the simple formula:

$$n = \sum_i N_i \, r_i$$  \hspace{1cm} (7.1)\

where $n$ is the expected number of WIMP-search nuclear-recoil single scatters passing surface-event rejection, summed over all detectors, just as before; $N_i$ is now the number of nuclear-recoil single scatters failing surface-event rejection on detector $i$; $r_i$ is an estimator for the cut performance in terms of a pass-fail ratio on that detector. This formulation statistically decouples the number of nuclear-recoil singles passing or failing surface-event rejection from the total number of WIMP-search
events. Note that using the count of WIMP-search nuclear-recoil single scatters failing surface-event rejection did not violate blinding since WIMP-search nuclear-recoil single scatters passing the surface-event rejection cut were not counted till the analysis was finished. The pass-fail ratios $r_i$, and thus $n$ were estimated using three different methods, each using one of the three datasets mentioned earlier. In practice, we estimate $n$ as the sum of the leakage contribution from endcap detectors, $n_{ec}$, and that from the interior detectors, $n_{int}$.

### 7.3.2 Poisson and Bayesian estimates

The leakage estimates for all three methods were made in two stages. The first stage used values of $N_i$ extrapolated from Runs 123–124, instead of using the true count of WIMP-search nuclear-recoil single scatters failing surface-event rejection in Runs 125–128. This enabled a check of the surface-event rejection cut, before applying the cut to WIMP-search data. Also, poisson error bars were propagated for estimates of $r_i$ and $N_i$, leading to large uncertainties on the value of $n$, because of low statistics. No systematic errors were estimated. This was deemed sufficient for a check of surface-event rejection performance.

After the first stage, the surface-event rejection cut was applied to determine the number of nuclear-recoil single scatters failing surface-event rejection, providing the true $N_i$ for this analysis of Runs 125–128 (see Section 7.4). In the interest of reduced uncertainties on the leakage estimate, the second stage used the true $N_i$, and a Bayesian prescription developed by Jeff Filippini for Runs 123-4, the details of which are presented in Appendix C of [72]. The estimate is made by Monte Carlo simulation of event counts using $\sim 100,000$ throws. This has the advantage of marginalizing over nuisance parameters such as pass-fail ratios and reducing complications due to correlated variables. The final estimate of expected leakage was taken to be the median value of the simulated distribution with a 68% credibility interval between [15.87, 84.13] percentile values. Priors for the poisson distribution governing event counts and an appropriate multinomial distribution governing detector energy and face distribution of events are picked from prior “conjugate” families to these distributions. The details of the implementation of second-stage estimates are documented in [195].

### 7.3.3 Method 1: Using WIMP-search multiples in the NR band

Multiple scatters inside the nuclear-recoil band should have the least systematic differences from nuclear-recoil single scatters, assuming no systematic difference between single and multiple scatter $r_i$. Thus testing the cut on this population provides an estimate of $r_i$ for nuclear-recoil single scatters.

$$ r_i = \frac{b_i}{B_i} $$

(7.2)
where \( b_i \) and \( B_i \) are counts of WIMP-search nuclear-recoil multiple scatters passing and failing the surface-event rejection cut, respectively, on the \( i \)-th detector.

The first estimate of \( n_{int} \) yielded \( 1.312 \pm 0.475 \) events, much greater than, but within error bars of the 0.5 event target leakage for the surface-event rejection cut. This number also disagreed with the first-stage estimate from Method 2, using WIMP-search singles outside the nuclear-recoil band. A survey of events failing timing cuts showed that these occurred at the lowest recoil energies. This prompted a retuning of \( \text{cRT\_vanilla\_c58} \), using revised energy binning, as described under Method 2. After the retuning, the first-stage estimate of \( n_{int} \) was \( 0.513 \pm 0.272 \) and agreed with the estimate from calibration data.

The second-stage estimate gave a final expected leakage summed over interior detectors of

\[
  n_{int} = 0.65^{+0.46}_{-0.29}\text{(stat.)} \pm 0.13\text{(syst.)} \quad (7.3)
\]

A number using this method was not computed for the endcap detectors (T3Z6 and T4Z6) because the systematic error would be too large in the absence of multiple scatters on their exterior faces. Method 1 provides smallest systematic errors, but greatest statistical errors.

### 7.3.4 Method 2: Using WIMP-search multiples outside the NR band

WIMP-search multiples outside the nuclear-recoil band avoid any possible systematic differences between single and multiple scatters, but have systematic differences from nuclear-recoil single scatters in energy and ionization yield. The systematics reweighting scheme from Equation [6.1] can be employed here to account for these. \( r_i \) is thus estimated by weighting measurements of the pass-fail ratio in detector face and energy bins in the following way:

\[
  r_i = \sum_{e,f} s_{e,f}^{(i)} \frac{m_{e,f}^{(i)}}{M_{e,f}^{(i)}} \quad (7.4)
\]

where \( m_{e,f}^{(i)} \) and \( M_{e,f}^{(i)} \) are counts of WIMP-search multiple scatters outside the nuclear-recoil band, failing and passing the surface-event rejection cut, respectively, on detector \( i \). \( s_{e,f}^{(i)} \) are bin-wise expected fractions of total WIMP-search nuclear-recoil single scatter count, estimated from WIMP-search nuclear-recoil multiple scatters for Runs 125-128. Recall, that to tune the surface-event rejection cut, we used \( s_{e,f}^{(i)} \) estimated from Run 123–124 WIMP-search nuclear-recoil surface events. The bins were at first the same: Low energy (10–30 keV) charge-side, low energy (10–30 keV) phonon-side, high energy (30–100 keV) charge-side, high energy (30–100 keV) phonon-side. Note that in the absence of tagged charge-side events for endcaps \(^1\) \( m_{e,\text{charge}}^{(\text{endcap})} \) and \( M_{e,\text{charge}}^{(\text{endcap})} \) were conservatively estimated using all singles on that endcap. \( s_{e,f}^{(\text{endcap})} \) were corrected by reassigning a fraction of

---

\(^1\)This analysis did not include end cap detectors at the top of tower, i.e., ones with an outward facing phonon face.
untagged events as phonon- and charge-side events. The reassigned fraction was determined using the average singles-to-multiples ratio for internal detectors. The resulting count was split between the two faces using the average phonon-to-charge count ratio.

The first-stage estimate yielded $n_{\text{int}} = 0.298 \pm 0.117$ and $n_{\text{ec}} = 0.283 \pm 0.298$. Based on disagreement with the Method 1 estimate, and the observed surface-event energy distribution for multiple scatters inside the nuclear-recoil band, the energy binning in the systematics reweighting scheme was revised to be 10–20 keV, 20–30 keV and 30–100 keV [196]. This made for a total of six bins (charge-side and phonon-side for each energy bin.) The cut, cRT\textunderscore vanilla\textunderscore c58 was therefore re-optimized for target leakage of 0.5 events on calibration data using the new bins. After the retuning, the first-stage estimate of $n_{\text{int}}$ was 0.287 ± 0.165 and for $n_{\text{ec}}$ was 0.529 ± 0.563 for endcaps. The estimate did not change very much, but was in better agreement with Method 1. The final Method 2 leakage estimate, using the Bayesian technique, was:

$$n_{\text{int}} = 0.72^{+0.56}_{-0.25}\,(\text{stat.}) \pm 0.18(\text{syst.})$$

$$n_{\text{ec}} = 0.29^{+0.33}_{-0.14}\,(\text{stat.}) \pm 0.05(\text{syst.})$$

(7.5)

(7.6)

Method 2 provides larger systematic errors than Method 1, but smaller statistical errors.

### 7.3.5 Method 3: Using calibration data

Method 3 uses exactly the same formulation as Method 2, except that $m_{e,f}^{(i)}$ and $M_{e,f}^{(i)}$ are counted for calibration data inside and outside the nuclear-recoil band. Using the revised six-bin systematic reweighting scheme, the first estimate of leakage was $n_{\text{int}} = 0.934 \pm 0.115$ and $n_{\text{ec}} = 0.205 \pm 0.098$, higher than the original 0.5 target. Following retuning of the surface-event rejection cut, these estimates were $n_{\text{int}} = 0.448 \pm 0.094$ and $n_{\text{ec}} = 0.057 \pm 0.098$. The final Method 3 leakage estimate, using the Bayesian technique, was:

$$n_{\text{int}} = 0.69^{+0.12}_{-0.10}\,(\text{stat.}) \pm 0.19(\text{syst.})$$

$$n_{\text{ec}} = 0.14^{+0.06}_{-0.04}\,(\text{stat.}) \pm 0.06(\text{syst.})$$

(7.7)

(7.8)
7.3.6 Combined estimate

The correlated estimates from all three methods were combined in the Bayesian Monte Carlo simulation to provide the following combined leakage estimate:

\[
\begin{align*}
n_{\text{int}} &= 0.66^{+0.11}_{-0.09} \text{(stat.)}^{+0.16}_{-0.15} \text{(syst.)} \\
n_{\text{ec}} &= 0.16^{+0.06}_{-0.04} \text{(stat.)}^{+0.05}_{-0.05} \text{(syst.)} \\
n_{\text{tot}} &= 0.82^{+0.12}_{-0.16} \text{(stat.)}^{+0.20}_{-0.19} \text{(syst.)}
\end{align*}
\]

The distributions obtained from the simulation for \( n_{\text{int}}, n_{\text{ec}}, \) and \( n_{\text{tot}} \) are plotted in Figure 7.2.

Figure 7.2: Surface-event leakage posterior distributions obtained from a Bayesian Monte Carlo simulation for Runs 125–128. Courtesy: David Moore.

7.3.7 Systematic errors

Systematic errors quoted above were computed separately from statistical ones, using the Bayesian technique. The following sources of systematic errors were accounted for:

1. **Single scatters vs. multiple scatters**: The computations for surface-event leakage in Method 1 assume no difference in the pass-fail ratios of the two. In the absence of sufficient statistics in the nuclear-recoil band to quantify a possible systematic difference, we use all surface events. In Runs 123–124, the ratio of pass-fail ratios between single and multiple scatters was measured to be \( \sim 0.8 \), whereas in Runs 125–128, it was measured to be \( \sim 1.24 \). Thus a 20% systematic error is assigned.
2. **Choice of prior**: The values of $n_i$ turn out to be sensitive to the choice of prior in some cases. Using 20,000 mock datasets, the prior parameters are picked to approximately minimize bias and ensure good coverage around the “true” values of $n_i$ in these mock sets. A 3–6% systematic error is assigned to the chosen prior based on the spread in the optimal parameters for different “true” leakages.

3. **Endcap systematics**: The reassignment of untagged events to charge-side and phonon-side to get correct $s^{(\text{endcap})}_{e,\text{charge}}$ is based on the average singles-to-multiples ratio for internal detectors. The systematic error assigned to this is the standard deviation in computing this average and is $\sim 5\%$. Also, the counts for $m^{(\text{endcap})}_{e,f}$ and $M^{(\text{endcap})}_{e,f}$ included all available untagged events. If the same correction used for $s^{(\text{endcap})}_{e,\text{charge}}$ is used here too, the systematic error is under a percent.

4. **Errors in reweighting for low energy events**: After unblinding, we discovered a possible residual systematic error due to lack of events with very low charge energies in the calibration sample. This is better explained after introducing the problem in Section 7.5.3. For now, I just quote the additional expected leakage (already accounted for in the combined estimate above): 0.19 and 0.04 events for interior and endcap detectors, respectively, for both Methods 1 and 2, and 0.23 and 0.06 events for interior and endcap detectors, respectively, for Method 3. A 50% systematic error is assigned to this component of leakage to account for possible errors in its estimation.

### 7.4 Unblinding WIMP Candidates

As explained in Section 7.3.2, the surface-event rejection cut is applied to WIMP-search nuclear-recoil data to count events failing the cut, $N_i$. This provides the last sanity check before unblinding. The counts of these events are listed in Table 7.2 alongside expected values based on a livetime scaling of values from Runs 123–124 \[197\].

I computed likelihoods for the expected values based on observed counts in Runs 125–128 and deduced p-values for each detector. The p-values are easily combined using a Fisher statistic, yielding 0.48. Next, I studied the distribution of events in time. I compared the CDFs of cumulative livetime for the events with a uniformly distributed CDF, detector by detector and together, producing a KS test statistic and a p-value. The cumulative p-value was 0.75, indicating no abnormalities in the distribution of these events over time. The only discrepancy was in the counts $N_i$ for detector T3Z6, which were abnormally high. The time distribution of events in T3Z6 produced a reasonable p-value of 0.4. Comparing the singles-to-multiples ratio and the phonon-side event fraction for that detector between Runs 123–124 and Runs 125–128, it was evident that phonon-side multiples tagging efficiency had declined in the latter. This produced an abnormally high number of single scatters in that detector. The sum of nuclear-recoil single and multiple scatters failing the surface-event
Table 7.2: Counts of WIMP-search nuclear-recoil single scatters rejected by the surface-event cut, along with predictions from Runs 123–124.

<table>
<thead>
<tr>
<th>Detector</th>
<th>Actual $N_i$</th>
<th>Predicted $N_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1Z2</td>
<td>5</td>
<td>$13.95 \pm 4.65$</td>
</tr>
<tr>
<td>T1Z5</td>
<td>9</td>
<td>$9.87 \pm 3.49$</td>
</tr>
<tr>
<td>T2Z3</td>
<td>9</td>
<td>$9.73 \pm 4.35$</td>
</tr>
<tr>
<td>T2Z5</td>
<td>7</td>
<td>$3.79 \pm 2.68$</td>
</tr>
<tr>
<td>T3Z2</td>
<td>6</td>
<td>$2.02 \pm 2.02$</td>
</tr>
<tr>
<td>T3Z4</td>
<td>8</td>
<td>$10.70 \pm 5.35$</td>
</tr>
<tr>
<td>T3Z5</td>
<td>4</td>
<td>$0.00 \pm 0.00$</td>
</tr>
<tr>
<td>T3Z6</td>
<td>52</td>
<td>$19.27 \pm 7.87$</td>
</tr>
<tr>
<td>T4Z2</td>
<td>10</td>
<td>$7.73 \pm 3.86$</td>
</tr>
<tr>
<td>T4Z4</td>
<td>5</td>
<td>$4.32 \pm 2.49$</td>
</tr>
<tr>
<td>T4Z5</td>
<td>10</td>
<td>$7.46 \pm 3.34$</td>
</tr>
<tr>
<td>T4Z6</td>
<td>10</td>
<td>$20.61 \pm 5.72$</td>
</tr>
<tr>
<td>T5Z4</td>
<td>6</td>
<td>$4.40 \pm 2.54$</td>
</tr>
<tr>
<td>T5Z5</td>
<td>9</td>
<td>$5.57 \pm 2.79$</td>
</tr>
<tr>
<td>Sum:</td>
<td>150</td>
<td>$119.43 \pm 15.29$</td>
</tr>
</tbody>
</table>

The surface-event rejection cut was consistent between Runs 123–124 and Runs 125–128 after appropriate scaling. Thus it was concluded that the abnormality did not pose a risk and any increases in expected surface-event leakage would be minimal.

With the analysis finalized, and all possible checks completed, we were ready to unblind. On November 5, 2009, Matt Fritts and I were given the go-ahead by the CDMS II collaboration to unblind the WIMP-search data for Runs 125–128, in the ROOT and MATLAB platforms, respectively. We ran scripts to find and count WIMP-search nuclear-recoil single scatters, passing the surface-event rejection cut \cite{198, 199}.

We observed two events at recoil energies of 12.3 keV and 15.5 keV on T1Z5 and T3Z4, respectively. Their characteristics are shown in Figure 7.3. Additionally, there were two near-miss events on T4Z2 and T4Z6, which passed the surface-event rejection cut, but were just outside the ±2σ nuclear-recoil band. Figure 7.4 shows the yield and timing distributions of all WIMP-search single scatters. Generally the low-yield single-scatter populations are well separated from the signal box, except for a few detectors where they are close.

Combining the background leakage estimates for surface events and neutrons (photons had a negligible contribution), and integrating over the posterior distribution, we compute that the probability of observing two or more background events is 23.3%. Therefore, an observation of two events does not constitute statistically significant evidence for WIMPs.
Figure 7.3: Left: Ionization yield vs. recoil energy for WIMP-search single scatters in T1Z5 (top) and T3Z4 (bottom). The vertical blue dashed line and magenta dashed line are the recoil energy and ionization thresholds, respectively. The red lines show the $\pm 2\sigma$ nuclear-recoil acceptance region. Right: Yield (normalized by nuclear-recoil band standard deviation) vs. $p_{\text{minrt}} + p_{\text{del}}$ (normalized to show signal acceptance between 0 and 15) for WIMP-search single scatters in T1Z5 (top) and T3Z4 (bottom). Histograms showing normalized yield and timing distributions for neutrons are shown in blue for comparison. Red box indicates signal acceptance region. In all plots, events that pass the surface-event rejection cut are marked in red.
Figure 7.4: Yield (normalized by nuclear-recoil band standard deviation) vs. $p_{\text{minrt}} + p_{\text{del}}$ (normalized to show signal acceptance between 0 and 15) for WIMP-search single scatters for all 14 Ge detectors. Histograms showing normalized timing distributions for neutrons are shown in blue for comparison.
7.5 Understanding the Two Observed Events

Next, we conducted a series of studies to understand the two signal-candidate events.

7.5.1 Basic checks

First we conducted basic checks of the experimental conditions under which these events were recorded to see if either event was included in the final WIMP-search datasets by accident. The following is a list of checks and outcomes.

1. The events seemed to be adequately spaced in time. They occurred on August 5, 2007, and October 27, 2007 on two different detectors in different towers.
2. Their traces do not exhibit any obvious abnormalities, as shown in Figure 7.3. The charge-pulse fits are investigated separately in Section 7.5.3.
3. The detectors appear to be well neutralized during the data series in which these events occur. The fraction of low-yield events is below the run-average value for both.
4. The background gamma and surface-event rates for both series were consistent with the averages for all four runs.
5. The data series have average KS test statistics for all seven tested RRQ distributions. Additionally, the number of events in these data series were sufficient to provide sensitivity in series-wise tests.
6. Studying the activity in all detectors for these events reveals that these were clearly single scatters — except for one detector each with a high phonon-threshold crossing, the energy in all other detectors was well below threshold. Also, the phonon and charge noise levels were normal for the data series containing the events.
7. Veto activity for either event is not atypical for a veto-anticoincident event.
8. The charge and phonon pre-pulse baselines look normal for both events.
9. The location of the events on the phonon-position manifold was consistent with having average detector radius. There were also no peculiarities with the nearest-neighbor averaging bins surrounding the event.

There are no obvious bad-data-taking conditions that would prompt the exclusion of the signal-candidate events.

7.5.2 Surface-event rejection close to threshold

The yield plot for the event in T1Z5 (Figure 7.3) indicated a band of surface events at low ionization yield, impinging on the nuclear-recoil band at low energies. The observed event was part of this band. This prompted me to study the performance of the surface-event rejection cut just below the
Figure 7.5: Top: Raw traces from all six channels of T1Z5 for the observed signal candidate. Bottom: Same, for candidate in T3Z4
recoil threshold and above the charge threshold on all detectors to try and see if any trends existed. I did this for both $^{133}$Ba-calibration data and WIMP-search data. I summarize important trends here, without being quantitative:

1. As shown in Figure 7.6, two events below the recoil threshold and part of the T1Z5 low-yield surface-event band pass the surface-event rejection cut. Their proximity to the candidate event in T1Z5 seems to suggest that timing flares more rapidly at low recoil energies than accounted for by the reweighting used to tune $cRT_{\text{vanilla}}c58$.

2. Figure 7.7 for $^{133}$Ba-calibration events on a detector shows that events just passing the charge-threshold cut have slower than normal surface-event timing and have a higher passage rate. This seems to indicate a timing reconstruction problem for events with low charge energy. This trend was observed for several detectors, but not before this study. It was not quantified until a detailed study of charge time reconstruction was conducted, explained next, in Section 7.5.3. The plot also shows the calibration surface-event selection boundaries, which incidentally excluded most events of this pathology. This prevented the setting of the surface-event rejection cut from accounting for flared event timing at low energy. The correct way to deal with this problem, from a cut-setting perspective, would have been to expand the calibration sample region to include the problematic events, and add charge-energy-based bins to the systematic reweighting scheme used to set the surface-event rejection cut.

### 7.5.3 Charge start-time reconstruction

Hints of charge timing reconstruction issues at low-charge energies were also observed during pulse reconstruction studies of the two signal-candidate events. In particular, the charge-optimal-filter fitted pulse for the signal candidate in T3Z4 seemed not to match the unfiltered charge pulse, as shown in Figure 7.8. The pulse start time selected by the optimal filter was earlier than what one would guess by eye. To check whether the best fit was picked by the optimal filter, Hsu studied the $\chi^2$ goodness-of-fit parameter distribution for the trace. The CDMS II charge optimal filter uses a shortcut — instead of picking the fit with the best $\chi^2$ value, it picks the fit that maximizes pulse amplitude, saving a factor of 2–3x in computation time. The study of the $\chi^2$ for the trace of the T3Z4 event indicated however that the $\chi^2$ that maximized pulse amplitude was not the one that had the lowest overall $\chi^2$, as seen in Figure 7.9. The fit with the lowest $\chi^2$ would have produced a later charge start time and hence a $pdel$ smaller by 4µs. The event with the correct $pdel$ would not have passed the surface-event rejection cut.

A simulation of low-charge energy pulses with noise added from real noise traces was run through the charge optimal filter. The optimal filter time resolution for maximum-amplitude fits was observed
Figure 7.6: Ionization yield vs. recoil energy for WIMP-search single scatters in T1Z5. Events are colored by $p_{\text{min}} + p_{\text{del}}$. Events that are circles instead of crosses are ones that pass the surface event rejection cut. In addition to the candidate event seen above threshold, two more events pass the cut right below threshold, and all appear to be part of a surface-event band.

Figure 7.7: Ionization yield vs. recoil energy for $^{133}$Ba-calibration events in T4Z4, below the $-3.5\sigma$ electron-recoil edge and passing the charge-threshold cut. Events are colored by $p_{\text{min}} + p_{\text{del}}$. Events that are circles instead of crosses are ones that pass the surface-event rejection cut. The electron-recoil band edge is shown as a blue band, indicating the run-to-run variation of that edge. Similarly, the nuclear-recoil band and the charge-threshold cut are indicated with magenta and green bands. The recoil-energy threshold is indicated with a vertical magenta-dashed line. The black curved lines show the run-to-run maximum variation of the $-5\sigma$ electron-recoil edge used for selecting calibration surface events. The horizontal black line at $y_{\text{ic}} = 0.1$ is the lower edge of the selection area. We note a large incidence of slow-timing events along the charge-threshold cut, indicating a charge timing reconstruction pathology for events with low charge energy. These events were mostly excluded in the calibration surface-event sample.
Figure 7.8: Raw fiducial charge trace for the signal candidate in T3Z4, with the optimal filtered pulse overlaid. By eye, it seems that the pulse might be better shifted down a few bins. Courtesy: Lauren Hsu.

Figure 7.9: Top: Charge optimal filter $\chi^2$ vs. charge start time in ADC bin. The $\chi^2$ for the fit picked by the optimal filter algorithm is marked with a red vertical dotted line, and is based on maximizing the pulse amplitude. Clearly this is not the minimum $\chi^2$. Bottom: Charge optimal filter $\chi^2$ vs. charge pulse amplitude. Here it is clear that the fit is picked to maximize the summed charge amplitude. Courtesy: Lauren Hsu.
to be greater than minimum $\chi^2$ by up to a factor of 2 for pulses with $< 6$ keV of ionization energy. Also, the rate of a 4-$\mu$s fluctuation in the optimal filter charge start time compared to the true start time was 1% for pulses of charge energy $< 6$ keV.

Using this information, we were able to correct the expected surface-event leakage estimate. We took samples of surface events (WIMP-search multiple-scatter and calibration surface events) with low recoil energy but sufficiently high charge energy ($> 6$ keV) and smeared their values of $p_{\text{del}}$ with the spread in reconstructed charge time for the simulated pulses. This provided a sample for measuring the systematic increase in surface-event leakage because of charge-start-time reconstruction errors at low charge energy. We then measured the charge energy spectrum for WIMP-search nuclear-recoil single scatters and for WIMP-search multiple scatters (or calibration surface events, depending on whether the estimate was for Method 2 or 3). The difference of these two spectra, convolved with the difference in surface-event leakage rate between the smeared and unsmeared surface-event samples provides the leakage unaccounted for in the original estimation of surface-event leakage.

It was against the spirit of a blinded WIMP-search analysis to go back and correct the $p_{\text{del}}$ value for this event and remove it as a signal-candidate event. Moreover, changing the charge-optimal-filter algorithm to use minimum $\chi^2$ instead of the $\chi^2$ that maximized signal amplitude would affect the charge reconstruction of all events. A correction would require reprocessing of the entire dataset starting with first-tier data processing. Also, this would have repercussions on phonon-pulse shape correction as well as setting the surface-event rejection cut. Thus we decided to leave such a tedious reprocessing for a later reanalysis of the entire CDMS-II dataset. The charge-optimal-filter algorithms for SuperCDMS have been updated to take the hit in computation time in favor of better charge start-time resolution.

### 7.5.4 Quantifying background/WIMP likelihood of candidate events

With the complete distribution of WIMP-search surface events in Runs 125–128 now unblinded and available to us, we were able to conduct basic likelihood studies for the two signal-candidate events. We made three estimates for the probability density functions for surface events and nuclear recoils in T1Z5 and T3Z4. One estimate used a non-parametric kernel-smoothing technique, and the other two used generalized lambda distributions. In all cases, the surface-event and neutron samples used for the computations adhered to the definitions in Section 6.6.1 with energy ranges as indicated below. Note that while statistical errors are taken into account in the following, systematic errors were not quantified.

1. A 3-D gaussian-kernel-smoothed density estimate of the two population distributions was created using recoil energy (between 10 keV and 100 keV), ionization yield, and $p_{\text{minrt}} + p_{\text{del}}$. The recoil energy and yield density functions were estimated as one 2-D density function us-
ing WIMP-search data. Separately, the recoil energy and the timing density function were estimated as one 2-D density function using $^{133}$Ba-calibration surface events. The two were combined with appropriate normalization, assuming little correlation between ionization yield and timing. This is not always a valid assumption, and is a caveat for any of these density estimates.

2. A 2-D density function was created by combining 1-D generalized lambda distributions for ionization yield and $p_{\text{minrt}} + p_{\text{dell}}$, for events with recoil energy between 5 keV and 20 keV\[203\]. Again it was assumed that there was little correlation between ionization yield and timing.

3. Another 3-D density function was created by combining the 1-D generalized lambda distributions for ionization yield, $p_{\text{minrt}}$ and $p_{\text{dell}}$ for events with recoil energy between 10 keV and 30 keV\[204\]. The correlation between $p_{\text{minrt}}$ and $p_{\text{dell}}$ was removed by diagonalizing the covariance matrix and transforming to an orthogonal basis for the two timing parameters.

These probability density functions were used to generate $\sim O(10^5)$ simulated Monte Carlo datasets, which were in turn used to measure log-likelihood ratios of the nuclear-recoil hypothesis to the surface-event hypothesis, given event properties. Since the signal-candidate events clearly satisfied criteria to be classified as nuclear recoils, and had low surface-event likelihood, Steve Yellin posed the following questions to seek meaningful information from these likelihood studies.

1. **What is the probability of observing a true surface event more nuclear-recoil-like than a signal-candidate event, i.e., with greater likelihood ratio?**

   First we compute this probability for one trial, $p$, in the table below.

<table>
<thead>
<tr>
<th>Detector</th>
<th>3D KDE</th>
<th>2D Lambda</th>
<th>3D Lambda</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1Z5</td>
<td>0.28 ± 0.04%</td>
<td>0.60 ± 0.08%</td>
<td>0.362 ± 0.009%</td>
</tr>
<tr>
<td>T3Z4</td>
<td>0.07 ± 0.03%</td>
<td>0.27 ± 0.05%</td>
<td>0.207 ± 0.005%</td>
</tr>
</tbody>
</table>

   Then we compute the binomial probabilities, $P_B = 1 - (1 - p)^N$, of observing a true surface event, out of the full sample of size $N$, to have a likelihood ratio greater than that of the signal candidate in T1Z5 or T3Z4. These are listed in table below,and are sufficiently large to treat a nuclear-event hypothesis for the signal-candidate events with reservation.

<table>
<thead>
<tr>
<th>Detector</th>
<th>3D KDE</th>
<th>2D Lambda</th>
<th>3D Lambda</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1Z5</td>
<td>24.23 $^{+5.31}_{-4.84}$%</td>
<td>11.9 ± 2.2%</td>
<td>12.2 ± 1.9%</td>
</tr>
<tr>
<td>T3Z4</td>
<td>4.23 $^{+2.88}_{-1.62}$%</td>
<td>4.6 ± 1.3%</td>
<td>5.1 ± 1.0%</td>
</tr>
</tbody>
</table>

2. **What is the probability of observing a true nuclear recoil, with energy equal to a candidate event and within the signal acceptance region, that is more surface-event-like than the candidates?**
This measures the probability with which true nuclear recoils would occur as close or closer to the yield and timing cut boundaries as the candidates. The probabilities are listed in the table below and are generally low, indicating that the events show poor consistency with being nuclear recoils.

<table>
<thead>
<tr>
<th>Detector</th>
<th>3D KDE</th>
<th>3D Lambda</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1Z5</td>
<td>1.18%</td>
<td>3.21 ± 0.06%</td>
</tr>
<tr>
<td>T3Z4</td>
<td>11.5%</td>
<td>1.88 ± 0.04%</td>
</tr>
</tbody>
</table>

3. **What is the probability of observing a true surface event to be more nuclear-recoil-like than the candidates in the signal acceptance region in these detectors?**

Note that this is different from the first question, because it imposes the boundaries of the signal acceptance region. These probabilities are listed in the table below and are high enough to encourage a surface-event hypothesis for the candidate events.

<table>
<thead>
<tr>
<th>Detector</th>
<th>3D KDE</th>
<th>3D Lambda</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1Z5</td>
<td>83.17%</td>
<td>27.7 ± 0.6%</td>
</tr>
<tr>
<td>T3Z4</td>
<td>54.52%</td>
<td>33.9 ± 0.7%</td>
</tr>
</tbody>
</table>
These studies indicate that the observed signal-candidate events are most likely background. A full reprocessing of the CDMS-II dataset with the corrected charge reconstruction algorithm will make it worthwhile to investigate a cut-free profile likelihood reanalysis of the entire dataset. The kernel density and generalized lambda distribution estimates of the background probability density functions will require scrutiny to ensure minimal systematic errors. Sebastian Arrenberg is currently preparing the framework for such an analysis.

7.6 Constraints on WIMP-Parameter Space

The results of this WIMP-search analysis can be interpreted in terms of constraints on the WIMP-nucleon elastic-scattering cross section as a function of WIMP mass. An upper limit was computed on both the spin-independent and spin-dependent cross sections using a MATLAB package developed over the years within the CDMS collaboration [205]. The statistical technique used by the package is the optimum interval method [200], which constrains normalization of spectra of known shape in the presence of unknown background. As explained in Chapter 2, we assume a dark matter halo with a Maxwellian velocity distribution of characteristic velocity $v_0 = 220 \text{ km/s}$, an escape velocity of 544 km/s and a local WIMP density of $\rho_0 = 0.3 \text{ GeV/cm}^3$. While scanning across WIMP mass, the package uses the Helm nuclear form factor (see Equation 2.11) to compute expected recoil spectra for the spin-independent case and the spin-structure functions in [70] for the spin-dependent case. Additionally, the analysis-efficiency-weighted exposure presented in Section 6.7 is taken into account. Note that the two observed events are conservatively assumed to be WIMPs under this treatment, without any attempted background subtraction.

The 90% CL upper limit on the spin-independent WIMP-nucleon elastic-scattering cross section, determined by the procedure above, is shown in Figure 7.10. The upper limit obtained from this analysis is shown as the red curve in the plot. It is a weaker limit compared to what would have been obtained with zero observed events and weaker than the limit from Runs 123–124, shown with a dashed-dotted blue curve. The kinks observed in the curve occur because of switching of optimum intervals around the two events for the computation of the limit at certain WIMP masses. Simply combining all CDMS II exposure with all events observed, i.e., the two events observed here and the events at 5.3 keV and 7.3 keV in Runs 118-119, yields the black curve. The expected median sensitivity of this analysis, based on the final background estimate, is shown in combination with previous CDMS II results as the black dotted curve. For comparison, competitive upper limits produced by experiments of similar scale are shown. The plot also shows theoretically interesting regions of parameter space predicted by models in the LEEST (low-energy effective supersymmetry theory) framework, described in [86], and regions suggested by more constrained CMSSM models in [87]. At the time this result was announced, it set the most stringent constraints on spin-independent
WIMP-nucleon scattering for WIMP masses $> 44\text{GeV}/c^2$.

Figure 7.10: 90% CL upper limit on the spin-independent WIMP-nucleon elastic-scattering cross section vs. WIMP mass. The red line is the limit obtained from this analysis. The solid black line is the limit obtained from all CDMS II exposure. The dotted black line represents expected sensitivity based on the final background estimate, combined with previous exposure in CDMS II. Prior results from CDMS II [132], EDELWEISS II [78], XENON10 [207], and ZEPLIN III [208] are shown for comparison. The shaded regions indicate allowed parameter space calculated from certain Minimal Supersymmetric Models. Courtesy: Jeff Filippini.

Figure 7.11 shows the 90% CL upper limit on spin-dependent WIMP-nucleon scattering cross-section, assuming coupling only to neutrons. This limit is weaker than that from WIMP searches with other targets which typically have a higher abundance of odd-nucleon isotopes. Our Ge crystals boast only 7.73% of $^{73}\text{Ge}$. Regardless, the sensitivity of none of the WIMP searches to date have reached interesting parameter space and are still more than an order of magnitude away. The scattering cross section limits for WIMP-proton spin coupling are not competitive with those from odd-proton targets such as CF$_3$I (from COUPP) or CsI (from KIMS), or from indirect search limits.
Figure 7.11: 90% CL upper limit on spin-dependent WIMP-neutron cross section vs. WIMP mass. The solid red line represents the limit set by this analysis. The solid black curve shows the limit set by the combined exposure of CDMS II. Previous results from CDMS II [132], XENON10 [207], and ZEPLIN III [208] are also shown for comparison.
Chapter 8

Characterization of Surface Radiocontaminants

The analysis of the final CDMS II dataset presented in this dissertation excluded WIMP parameter space just shy of projected sensitivity, because of the observation of events at low energy, likely misidentified surface events. Since the dominant expected background was indeed surface events, with only a minor contribution from cosmogenic and radiogenic neutrons, it is conceivable to continue a CDMS-style experiment using the existing infrastructure at Soudan, with improvements in surface background rejection, or with lowered surface backgrounds. As explained in Section 2.3.3, operation in a near-zero-background regime is essential for maximum sensitivity to WIMPs.

The CDMS collaboration is using a multi-faceted approach to the problem of surface background reduction for the SuperCDMS experiment at Soudan. Improvements in the ZIP design, matched with corresponding analysis improvements, are being explored to enable better discrimination against surface events. I explain some of this work in Appendix A. Alongside, we are making efforts to understand the sources of surface events to try to reduce their raw rates before surface-event rejection techniques are applied in analysis. It turns out that at least half of the surface events seen in ZIPS are caused by shallow-penetrating electrons emitted by radiocontaminants, as I will explain below.

In this short chapter, I first briefly review an analysis that measured the contribution of $^{210}$Pb-decay betas to be at least half of the surface events seen in ZIPS. Then, I highlight the required levels of surface contaminant reduction to maintain zero backgrounds for future Ge-based direct searches for WIMPs. Existing screening technologies are not well suited for measuring and characterizing the levels of surface radiocontaminants required for CDMS detector substrates and other detector materials. This is a problem faced not only by CDMS, but several other rare-event search experiments. This motivated us to start an effort to develop a dedicated screener for surface alpha- and beta-emitting radiocontaminants. I describe this effort in the remainder of this dissertation.
8.1 Identification of Surface-Event Contributors in CDMS II

Since surface events in ZIPs are caused by particle interactions within a 10-µm-thick dead layer, our naïve expectation is that they are particles with low penetration depth, i.e., low-energy electrons. Low-energy photons with penetration depths smaller than 10 µm fall below typical ZIP recoil-energy thresholds and are thus not a contributor to surface events. Low-energy electrons can be easily ejected from a detector’s neighboring materials by the photoelectric effect of ambient photons, or they can be beta particles emitted by radioisotope contamination on the detector and its copper housing.

The photon-induced surface events are expected to scale with ambient photon rate. During $^{133}$Ba calibrations we flood the ZIPs with X-rays and gammas, and accompanying photon-induced electrons, enabling a measurement of the ratio of surface events to bulk electron recoils. This is measured to be $8 \pm 4 \times 10^{-4}$ \cite{209}. We then use this ratio and the WIMP-search low-background ambient photon rate, $295 \pm 135$/kg/day \cite{210} to deduce the photon-induced surface-event rate, $0.22 \pm 0.10$/kg/day. The total surface-event rate measured in CDMS II ZIPs is $0.37 \pm 0.18$/kg/day \cite{210}. Thus if the estimate of photon-induced surface-event rate is correct, a contribution at least equal to it should be expected from beta emitters.

The dominant long-half-life beta emitting contaminants on the ZIPs or their copper housings are expected to be $^{210}$Pb, $^{40}$K, and $^{14}$C — the first caused by plate-out from radioactive decay of $^{222}$Rn present in the air, and the latter two from processing and human contact.

$^{210}$Pb contamination has some unique signatures that are unambiguously observed in ZIPs. 84% of all $^{210}$Pb decays are associated with a 17-keV beta end point and the emission of a 46.5-keV photon or an internal conversion electron followed by Auger-electron or X-ray emission \cite{211}. Because of the multiple particles involved and the possibility of backscattering, all this energy does not appear clearly in one detector, but shows expected characteristics when the energies of two nearest-neighbor detectors are summed \cite{212}. This is shown in the left pane of Figure 8.1. As the $^{210}$Pb nucleus proceeds down its decay chain through $^{210}$Bi and $^{210}$Po, it reaches stable $^{206}$Pb by the emission of a 5.3-MeV alpha particle \cite{211}. In this final decay, the $^{206}$Pb nucleus recoils with kinetic energy of 103 keV. This signature is clearly visible in the ZIPs too, although the $^{206}$Pb recoil is missed about half the time, because of its extremely small penetration depth (tens of nm). A plot of charge energy vs. recoil energy clearly shows the high-energy alphas in the right pane of Figure 8.1. Rates of both signatures differ among ZIPs, but are correlated, confirming that the source of these signatures is indeed the $^{210}$Pb contamination. This correlation is shown in Figure 8.2. A maximum-likelihood model was constructed to estimate the expected $^{210}$Pb-induced surface-event rates given the observed rates of 46.5 keV summed nearest-neighbor surface events, the observed alpha rates, and the total

\footnote{All rates measured using data from Runs 119 and 123 and for recoils of 10–100 keV}
Figure 8.1: Left: Counts of summed energy for a nearest-neighbor detector pair where events on both detectors are surface scatters. The 46.5-keV gamma peak from $^{210}\text{Pb}$ decay is visible, smeared by several electrons and x-rays around the same energy. *Courtesy: Rupak Mahapatra.* Right: Charge energy vs. phonon energy recorded for a Ge ZIP during Run 123 WIMP search. At the energy scales shown here, 5.3-MeV alphas are clearly visible. *Courtesy: Jodi Cooley.*

Surface scatter rate $^{[213]}$. It used input from a GEANT4 Monte Carlo for the rates of the observed processes per $^{210}\text{Pb}$ decay and for geometric efficiency factors. The resulting best-fit $^{210}\text{Pb}$-induced surface-event rate averaged over all Ge detectors was $0.24 \pm 0.18/\text{kg/day}$. It was found that the rates for ZIPs in Towers 1 and 2 were a factor of 2–3 higher than those for Towers 3, 4, or 5. This is consistent with reduced exposure to air for the later towers as our handling techniques during fabrication improved over time.

$^{14}\text{C}$ and $^{40}\text{K}$ contamination cannot be directly measured using ZIPs themselves because of a lack of clear monoenergetic signatures. Some limits have been obtained on these contaminants through test wafer assays. The $^{14}\text{C}$ contamination was measured by Rutherford backscattering (RBS) and Auger spectroscopy and showed 2–3 monolayers of C atoms on wafer surfaces, consistent with exposure to air. The resultant surface-event rate is expected to be $0.006–0.06/\text{kg/day} (10–40\text{ keV})$ $^{[210]}$. The $^{40}\text{K}$ contamination was measured by RBS and particle-induced X-ray emission, but was limited in sensitivity to about the total observed surface scatter rate ($\sim 0.3/\text{kg/day}$) because of confusion with argon $^{[210]}$. Argon is used as an inert atmosphere in the processing of the ZIPs. In the future, replacement of this gas with He or Kr, at least for test wafers, will enable better sensitivity by an order of magnitude.

The $^{210}\text{Pb}$ likelihood analysis, together with information about the $^{14}\text{C}$-, $^{40}\text{K}$- and ambient-photon-induced surface-event rates showed that the total observed surface-event rates could be explained by equal contributions from photons and $^{210}\text{Pb}$ contamination, though subdominant contributions from emitters like $^{14}\text{C}$ and $^{40}\text{K}$ are likely. Thus, improvements in background control for future CDMS-style experiments will require equal efforts on reducing the photon background and...
Figure 8.2: Correlation between rates of 46.5-keV nearest-neighbor surface scatters and alpha+recoiling $^{206}$Pb nucleus rates for Run 123. Each data point represents a nearest-neighbor detector pair. The line is based on the best-fit $^{210}$Pb contamination model. *Courtesy: Sunil Golwala.*

8.2 Surface-event Reduction Requirements for SuperCDMS

SuperCDMS Soudan has a target sensitivity of $5 \times 10^{-45}$ cm$^2$ for a 60 GeV/c$^2$ WIMP, approximately a factor of 8x lower than the final CDMS II sensitivity [210]. Since this experiment will operate at Soudan using pre-existing infrastructure, the ambient photon background will remain unchanged. The desired 8x increase in exposure will be accomplished by increasing the total target mass to 15 kg (possibly as low as 9 kg with fiducial volume improvements). A corresponding 8x decrease in residual surface-event rate will enable the experiment to operate nearly background free. Each detector will be 2.5x thicker than CDMS II ZIPs, leading to a factor of 2.5 geometric reduction in surface-event rates per detector, as the electrode-instrumented area will remain the same. The remaining reduction in residual surface background will require adhering to the $^{222}$Rn exposure levels of Towers 3, 4, and 5, (a factor of 1.6) as well as improvements in phonon and ionization sensor design (a factor of 2). After this exposure, Soudan backgrounds will be dominated by cosmogenic neutrons, necessitating a deeper site for continued background-free operation.

The next phase, SuperCDMS SNOLAB, will use 100 kg of Ge target mass to achieve a sensitivity of $3 \times 10^{-46}$ cm$^2$ for a 60 GeV/c$^2$ WIMP, approximately a factor of 100x lower than the final CDMS II sensitivity. Installation at a new facility will permit the construction of an improved shield with at least a factor of 2 reduction in ambient photon backgrounds. This has already been
demonstrated in CDMS I with an inner lead shield between the inner polyethylene and the copper cans, but was not used at Soudan to the detriment of the ambient photon rate. The radiocontamination rates of detectors need at least a corresponding factor of 2 improvement to keep up with the reduction in photon-induced surface-event rate. We believe that the remaining reduction in residual surface background will be accomplished through continuing detector and analysis improvements (See Appendix B).

8.3 Surface Radiocontamination Goals

As stated above, the nominal goal for reduction of radiocontamination on detectors and their housings is a factor of 2 for SuperCDMS SNOLAB. Conservatively, this may need to be as high as 5x to account for any shortfalls in detector improvements. The current estimated $^{210}\text{Pb}$ contamination on ZIPs causes a raw surface-event rate of 5/kg/day $^{210}$, which translates to $1.1 \times 10^{-4}$/cm$^2$/keV/day of surface activity. A factor of 5x reduction requires a target rate of $\sim 2 \times 10^{-5}$/cm$^2$/keV/day.

To accomplish this reduction, we will first need to figure out which treatments and processes during detector fabrication introduce radiocontaminants. This is best done by exposing witness samples to different stages of the fabrication process and assaying them to quantify the contamination introduced at various stages. The most contaminating steps can be studied further to explore mitigation techniques. For the $^{210}$Pb contamination, which constitutes a majority of the surface radiocontamination, this means understanding which fabrication steps lead to the largest exposures to air and $^{222}\text{Rn}$, and modifying them to minimize such exposure. Unfortunately, current assay techniques are not very conducive to measuring surface radiocontaminant levels rapidly or cheaply, as I explain in the next section.

8.4 Current Assay Techniques and Shortcomings

Techniques for characterization of trace radioactivity can be broadly categorized into three groups and all three are limited in characterization of surface radioactivity:

1. **Semiconductor detectors**: High-purity germanium detectors (HPGe), operated at 85–100 K, measure gamma spectra with $\sim$ keV resolution and are used to identify contaminants based on specific photon spectral features they produce $^{[214]}$. Their energy thresholds are typically few tens of keV because they are separated from samples by a vacuum window, and also suffer from p-n junction dead-layer effects much worse than ZIPs $^{[215]}$. This makes them less ideal for detection of contaminants that emit low-energy electrons, more so if the process does not have any associated gamma spectral lines. For detectors with thin dead layers such as lithium-drifted silicon detectors and boron-implanted HPGe, the vacuum window is still a
sufficient hindrance to low thresholds. Additionally, these detectors have very small counting areas, requiring integration times on the order of weeks to achieve sensitivity to low-activity samples such as the ones we would screen for CDMS.

2. **Mass spectrometry:** Mass spectrometers characterize the elemental composition of a sample by ionizing liquid or vapor samples and then measuring the mass-to-charge ratio of different ion species in the sample \[210]. A large number of beta emitters can be identified even at low radioactivity because of ppb and ppt sensitivities achieved by mass spectroscopy. However, its main drawbacks are the processing required to convert a solid sample into liquid, which may introduce other contaminants, and the expense of the equipment and running assays using this technique.

3. **Alpha counters:** Direct counting of alpha emissions in a gas is a very promising technique for characterization of surface radioactivity by contaminants such as \(^{210}\)Pb, which produce alpha particles. Traditionally, sensitivity has not been sufficient for use in low-radioactivity counting, but this is changing now. A company called XIA\[^2\] has designed and prototyped an ultra-low background alpha-particle gas detector with 1800 cm\(^2\) counting area and claimed sensitivities of \(2.4 \times 10^{-3}/\text{cm}^2/\text{day}\) \[217\]. We have acquired a prototype for evaluation, which might be useful for SuperCDMS Soudan to screen samples for \(^{210}\)Pb contamination using the \(^{210}\)Po alpha decay. The likelihood analysis mentioned earlier determined the \(^{210}\)Po-alpha-decay rate for Towers 3, 4, 5 in CDMS II to be 0.32/detector/day, or \(4.6 \times 10^{-3}/\text{cm}^2/\text{day}\). As mentioned earlier, this rate cannot be exceeded for SuperCDMS Soudan detectors, and can be screened using the XIA counter if its advertised sensitivity is achieved. Unfortunately, beta emitters like \(^{14}\)C and \(^{40}\)K that do not have associated alphas in their decay chains cannot be characterized by alpha counters like XIA’s. Certainly, even the XIA counter will be insufficient for the screening needs of SuperCDMS SNOLAB.

8.5 **Need for Dedicated Beta- and Alpha-screening**

The lack of low-background direct counters for beta-emitting isotopes, and for low-activity surface contamination in general is a hindrance not only for CDMS, but for several other experiments in the physics community. EDELWEISS uses Ge detector technology similar to CDMS’ for dark matter searches, and suffers from the same low-energy surface electron problem as CDMS \[78\]. While its new interleaved-electrode detectors have provided significant improvement in surface event rejection \[218\], radiocntamination needs to be reduced further to achieve sensitivity increases for the proposed EURECA experiment \[219\]. Another promising dark matter search, COUPP, is limited

\[^2\]http://www.xia.com/Alpha products.html
not by electromagnetic backgrounds, but by alpha-induced recoils from the materials of its bubble chamber [82]. They have accomplished some discrimination against alphas, but continue to seek ways to reduce their contamination with alpha-emitting species. Similarly, proposed dark matter experiments DEAP-3600 [220] and miniCLEAN [221] will also face surface background issues. They wish to suppress backgrounds to a level where alphas ejected from detector walls would rarely cross into a fiducial volume boundary to mimic signal. In the future, the successor to miniCLEAN, a $pp$ neutrino detector called CLEAN, will face these challenges to a larger extent [222]. All these experiments would benefit from the development of a dedicated screener for low-background surface radioactivity. Motivated by this, the Caltech and Syracuse collaborators on CDMS decided to dedicate resources to solving this problem. In the next two chapters, I provide details of a new screener called BetaCage, which will fulfill this need.
Chapter 9
BetaCage Design

In the previous chapter, I presented the limitations of materials assay technology for surface radioactivity screening and provided motivation for a new low-energy beta and alpha counter. Such a counter would possess the following characteristics:

1. It would be optimized to count betas with energy $\sim 0$–200 keV.
2. It would have inherently low radioactive background to achieve high sensitivity. In practice this requires that the detector have minimal surface area and be shielded.
3. It would contain minimal stopping material to reduce backscattering of track electrons and to maintain high counting efficiency.
4. It would allow sufficient spatial reconstruction of tracks to enable rejection of events not originating from the sample.
5. It would be capable of accommodating large-area samples to permit short integration and analysis times.

In this chapter, I present the design for BetaCage, a counter optimized for low-energy betas and alphas, designed to meet these objectives.

9.1 Basic Design and Operational Principles

BetaCage will be an underground, shielded time-projection chamber, constructed with low-background materials. It will use quenched neon gas at STP as the stopping material for particle tracks.

A design schematic of BetaCage is shown in Figure 9.1. It will consist of three multi-wire proportional counters (MWPC) with keV-level energy resolution. A sample to be counted will be placed directly in gas underneath the “trigger” MWPC. A particle ejected from the sample will ionize gas along its trajectory through the drift region above the trigger MWPC. The height of this drift region is set by the practical range in neon of betas and alphas of interest. During the ionization process, the trigger MWPC will record a signal marking the event start time. The
ionization track will drift under the influence of an electric field to the “bulk” MWPC, which will record the x-y position, counts, and arrival times of electrons. Thus, information combined from the two MWPCs will allow energy estimation and position reconstruction of tracks. We will reject events not originating from the sample using several techniques. First, we will define a fiducial acceptance area in the trigger and bulk MWPCs by requiring no signal in wires that monitor the area outside the fiducial area. Second, we will require events to deposit energy in the bulk MWPC within a time window after depositing energy in the trigger MWPC, based on known drift times for electron clusters. The third, “veto,” MWPC below the sample will help reject through-going particles such as muons and other minimum ionizing particles, or any particles that have the correct timing profile in the Trigger and Bulk MWPCs, but originate from underneath the sample.

A gaseous time-projection chamber format naturally lends itself to a low-background design. Suitably screened stock of low-radioactivity materials can be used for the construction of structural and sensing elements of the counter. The gas chosen for operation, neon, has no long-lived radioactive isotopes. The construction itself can be done in a radon-suppressed environment and the counter can be kept under nitrogen purge during transportation to reduce radon plate-out. By locating the counter at an underground site such as Soudan Underground Laboratory or SNOLAB, and providing copper or lead shielding, a Ge equivalent rate of $\sim 1\text{keV}^{-1}\text{kg}^{-1}\text{day}^{-1}$ can be achieved easily.

Figure 9.1: Schematic of proposed BetaCage counter (not to scale). A track originating from a sample (shown in brown), between the trigger and veto MWPCs ranges out in the drift region. The track drifts to the bulk MWPC, which records position and energy information. The system is enclosed in a gas containment vessel and inside passive shielding to mitigate photon background.
9.2 Multi-wire Proportional Counter (MWPC) Units

9.2.1 Particle detection using MWPCs

The term MWPC typically stands for Multi-wire Proportional Chamber, referring to an enclosed particle detector unit containing gas and a plane of parallel high-voltage wires (anode) sandwiched between two grounded plates (cathodes). BetaCage MWPCs will have two key differences — instead of ground plates, we will use wire planes at some potential lower than the central plane of wires; also three such units share the same gas volume in our design, instead of being hermetically isolated from one another. These two changes will enable ionization tracks to freely move through the apparatus with minimal backscattering.

The principle of particle detection in an MWPC is based on proportional amplification of ionization deposited in them, or in the case of the BetaCage, drifted into them. Electrons belonging to an ionization track move along field lines to anode wires closest to them. Since the wires have a non-uniform, monotonically increasing field with wire radius, the electrons gain energy as they get closer to these anode wires. By design, the anode wire radius is small enough and the gas ionization potential is low enough, that initial track electrons gain sufficient energy to start ionization of gas molecules within a few wire radii. This leads to a cascade or an avalanche of electron-ion pair generation, with gains achievable up to \( \sim 10^3 \text{–} 10^6 \) [223, 224]. The avalanche of electrons is absorbed by the anode wire almost instantaneously (within a fraction of a nanosecond, typically), but the resulting ions have to travel to a cathode before being absorbed. This motion of ions registers image currents in all electrodes in the system. Recording these currents helps determine the initial electron count and hence the energy of the ionizing particle. In particular, the signal has greatest amplitude on the avalanche-causing electrode because of its proximity to the ions, and decreases on other electrodes proportionally to the ratio of the cross-capacitance of the measurement and avalanche electrodes to the self-capacitance of the avalanche electrode [224]. If the cathode plane is made of strips of metal or wires that are perpendicular to the direction of the anode wires, signals from both anode and cathode elements can be recorded to reconstruct x-y position of electron clusters. I present a detailed account of signal development in Section 9.7.

9.2.2 MWPCs in BetaCage

In BetaCage, each MWPC unit will consist of a central plane of meter-long 25-\( \mu \)m-diameter anode wires, parallel to each other, at 5-mm pitch. Above and below the anode wire plane, at 5-mm spacing, there will be a plane each of meter-long 125-\( \mu \)m-diameter cathode wires at 5-mm pitch. Each wire plane will have 200 wires making for a total 1-m\(^2\) counting area. The cathode wires will run perpendicular in direction to the anode wires, providing each MWPC with event position reconstruction capability. Figure 9.2 shows a schematic of the arrangement of wires in a BetaCage
Figure 9.2: Schematic of MWPC wire arrangement. The black sets of parallel wires are the cathode wires, and the perpendicular set of red wires is the anode. The green box demonstrates definition of a fiducial volume using a subset of wires of the MWPC.

MWPC.

The wires will be stainless steel to allow a combination of radiopurity and mechanical strength. Structural support for the suspension of wires will be provided by a custom-machined Noryl frame. Noryl was picked for its low radioactivity, ease of precision machining, and material stability. The wires themselves will be positioned using precision brass feedthroughs with copper ferrules. Mechanical and electrical connection with the wires is made by crimping the wire in the copper ferrule. We will mount low-radioactivity Circlex PCBs on the side of each MWPC to make electrical connections to the ferrules. The R&D leading to this design and an overview of the assembly process of MWPCs is described in detail in the next chapter.

9.2.2.1 Wire ganging and Channels

With three MWPCs, the BetaCage will have 1800 wires, making readout of individual wires complicated as well as prohibitively expensive for a project of small scale. Thus we will gang wires on the Circlex PCBs in the following way to extract necessary information without overburdening the readout electronics or the construction budget.

For the veto and trigger MWPCs, the anode wires and only one plane of cathode wires will be instrumented with readout electronics. The second cathode plane will be held at ground like the instrumented cathode plane, but will not be read out. Since we need timing information and not much position information about events in the veto and trigger MWPCs, each of the instrumented planes will have only two channels. One will consist of 150 wires ganged together in the central region, and the other will contain the remaining wires of the outer region (25 on either side of the central region). The central 150 wires from the anode and those from the cathode will be perpendicular to each other and together will define a “fiducial area” as demonstrated in Figure 9.2. Events passing the fiducial area cut will be required to have signal consistent with noise in the outer channels. This is because events close to the edges of the MWPC will encounter fringing electric fields that can cause non-uniform gain and hence degrade energy and timing resolution. Thus the trigger and veto MWPCs will have four channels each.

For the bulk MPWC, the anode wires and only one plane of cathode wires will be instrumented
with readout electronics, just like the other two MWPCs. However, we need to record energy, timing, and position information from tracks so two channels each will be insufficient. Outer channels consisting of 50 wires will be ganged into five channels of 10 wires each, and will collectively help reject events too close to the edges of the MWPC. The inner 150 wires will be ganged as 15 channels of 10 wires each, defining the bulk fiducial region. This will make for a total of 40 channels for the bulk MWPC. Note that repeating wire-gang patterns will produce degenerate solutions in position reconstruction. Thus BetaCage wire gangs will contain randomized sets of non-adjacent wires.

Therefore, the BetaCage will have a total of 48 channels of signal readout.

9.2.2.2 MWPC electrostatics

It is possible to write down analytically the electric potential function for an MWPC with parallel wires because the problem can be reduced to two dimensions. BetaCage MWPCs have perpendicular wire planes, making it complicated to write down an analytical solution in three dimensions. Therefore, I used COMSOL Multiphysics software to simulate the electrostatics of BetaCage MWPCs. I first used the symmetry of repeating unit cells to downsize the scope of the numerical solver. I define mirror boundary conditions on four sides of the cell and obtain the field configuration shown in Figure 9.3 for the bulk MWPC. The solution roughly agrees with an analytic treatment where the cathode wire planes are approximated as plates [224] and with a numerical solution by Bob Nelson using GARFIELD [225].

![Electric equipotentials (V) for a unit cell of the bulk MWPC. The spatial axis scale is in units of \( \mu m \). The two cathode wires are held at 2100 V and the anode wire is held at 4028 V.](image)

We also need to obtain a couple of components of the MWPC capacitance matrix in order to compute the charge density and hence the gain at an anode wire in Section 9.5 — the self-
capacitance of the anode wire, and its cross-capacitance with all other electrodes. Since anode and cathode wires are set at different voltages, I separated the cross-capacitance term into an anode and a cathode component. To obtain these numbers, I simulated \(N\)-wire MWPCs for \(N = 1, 3, 5, \ldots, 11\) to determine the scaling relation of these elements with \(N\). The resulting capacitances per unit length are summarized in Table 9.1.

<table>
<thead>
<tr>
<th>Matrix element</th>
<th>Capacitance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anode-self</td>
<td>9.0 pF/m</td>
</tr>
<tr>
<td>Anode-all other anodes</td>
<td>-6.7 pF/m</td>
</tr>
<tr>
<td>Anode-all cathodes</td>
<td>-2.3 pF/m</td>
</tr>
</tbody>
</table>

9.3 Gas, Handling and Containment

The most common gas used in drift chambers is argon, because it is cheap, has low ionization potential, and excellent drift properties. Assuming a 1-m\(^3\) volume for the containment vessel, the rate of decays from naturally abundant, radioactive \(^{39}\)Ar in the volume will be \(\sim 1.6\) Bq \([226]\). This will necessitate isotopic separation to enrich the non-radioactive component and reduce the background rate from the gas. For simplicity, we will instead use neon, which has no long-lived radioactive isotopes and exhibits good properties for use as a drift chamber gas. Since neon is more expensive than argon, we will purify and recirculate the required quantity.

In practice, every drift chamber gas is supplemented with a small quantity of “quench” gas to prevent runaway avalanches. Gains larger than \(10^3–10^4\) cannot be obtained in pure noble gases without breakdown \([223]\). This is because the ionization process in noble gases typically generates de-excitation photons, in addition to electron-ion pairs. These photons are energetic enough to extract photoelectrons from the cathode, inducing gas breakdown and electrical discharge. However, addition of a polyatomic organic compound helps absorb these photons through rotational and vibrational modes, thus preventing discharge. For BetaCage, we have decided to use methane for quenching, since its properties are well documented and it is sufficiently cheap. The two gases will nominally be combined in a 90\%-10\% CH\(_4\) mixture at STP. Increasing the CH\(_4\) content beyond 10\% will increase the drift velocity and reduce diffusion, both good things for the performance of the BetaCage, but will increase the radioactive carbon in the volume of the chamber. Lower quench gas fraction will degrade performance, but it might be possible to operate with a 95\%-5\% CH\(_4\) mixture. This will be determined empirically after building a prototype BetaCage.

The entire BetaCage setup will be enclosed in a gas-tight plastic pressure vessel. Gas will be introduced, maintained, and cycled using a dedicated gas handling system. Cylinders of high-purity
neon and methane will be connected via stainless steel tubing to mass flow control transducers that will mix the two components at the required ratio. Downstream of the controller transducers, a 0.2-µm membrane filter followed by a 0.01-µm membrane filter will remove particulate contamination before introducing the mixture into the vessel. A circulation pump will recover this mixture at 1.3–1.5 lpm and pass it through a SAES Pure Gas PS11-MC1-CH getter-stabilized zeolite gas purifier. This will remove water, CO₂, O₂, non-methane hydrocarbons, and particulates greater than 0.003 µm in size to < 1 ppb before recirculation.

9.4 Drift Volume and Field Cage

Track energy deposition occurs in a large volume of gas between the trigger and bulk MWPCs. The practical range of 156 keV betas (¹⁴C endpoint) in a 90% Ne–10% CH₄ mixture is 30 cm, i.e., 90% of such betas should be contained in this height [227]. This will allow us to use a ¹⁴C source for calibration as long as the drift volume is at least 30-cm tall. For the BetaCage, we have set the height of the drift volume at 40 cm to enable calibration and contain beta endpoints and alphas from contaminants of interest, such as ¹⁴C and ²¹⁰Pb.

The drift volume is defined by the the “field cage,” the support structure that separates the trigger and bulk MWPCs. It will consist of 9-mm-thick rectangular acrylic frames separating 1-mm-thick rectangular copper frames at 1-cm intervals. The structure will have a 1-m² inner footprint to match that of the MWPCs. The copper frames, or “field shapers” will be held at uniformly increasing voltage from the trigger to bulk MWPC, to define a uniform electric drift field for track clusters. A schematic of the drift field region is shown in Figure 9.4. The biasing of the field cage is accomplished with a precision resistor chain attached to the field shapers with a 2000-V regulated high-voltage supply on one end and grounded on the other. With 40 copper field shapers, this sets a field of 50 V/cm. The interaction of the field cage with the containment vessel was modeled for a BetaCage prototype discussed in the next chapter and is plotted in Figure 9.5.

The RMS electron diffusion for this configuration is expected to be 470 µs/√cm [228]. This will lead to 0.3 cm of diffusion, less than the wire pitch of the MWPCs, for an electron traversing the full length of the drift region.

9.5 Voltage and Gain Requirements

Given gas properties, the mechanical specification of an MWPC, and the operating voltages, one can determine the gain or amplification provided by it. Conversely, given a desired gain, one can determine the bias voltages for anodes and cathodes. To calculate the expression for gain, we first
Figure 9.4: Schematic showing copper field shapers with resistor chain to provide electric field in the drift field region.

Figure 9.5: Electric potential (V) for a cross section of the drift cage, bookended by the trigger and bulk MWPCs, determined by Ansoft Maxwell simulation software. Spatial axes have units of meters. This assembly doesn’t contain a veto MWPC. It is placed in a cylindrical pressure vessel (grounded), oriented with its rotation axis pointing perpendicularly out of the page. This causes the circular equipotentials seen in the plot.
write down the line charge density of the \( n \)-th anode wire, \[
\lambda_n = \sum_{m=1}^{N} c_{nm} V_m = (c_{nn} + c_{na}) V_a + c_{nc} V_c
\] (9.1)

where \( c_{nn} \) is the self-capacitance of the anode wire, \( c_{na} \) is its cross-capacitance with all other anode wires, \( c_{nc} \) is its cross-capacitance with all cathode wires, and \( V_a \) and \( V_c \) are anode and cathode voltages, respectively. The capacitance values were computed in Section 9.2.2.2 and listed in Table 9.1.

The electric field at distance \( r \) around the \( n \)-th wire in the anode grid is approximated by the field of a line charge, assuming that \( r \) is much smaller than than the wire pitch or spacing between electrodes (true for BetaCage),
\[
E_n(r) = \frac{\lambda_n}{2\pi \epsilon_0 r}
\] (9.2)

Using Equation 9.2 we compute the potential difference between \( r = r_0 \), the wire radius and \( r = r_{\text{min}} \), the radius at which the electric field, \( E = E_{\text{min}} \), is sufficient to start an avalanche.
\[
\Phi(r_0) - \Phi(r_{\text{min}}) = \int_{r_0}^{r_{\text{min}}} E(r)dr = \frac{\lambda}{2\pi \epsilon_0} \ln \frac{r_{\text{min}}}{r_0} = \frac{\lambda}{2\pi \epsilon_0 r_0} \ln \frac{\lambda}{2\pi \epsilon_0 r_0 E_{\text{min}}}
\] (9.3)

If we divide this potential difference by the average potential difference through which an electron accelerates to ionize a gas molecule, \( \Delta V \), we get the number of generations of electron doubling in the avalanche process. Thus, the gain or the final number of electrons is, \( G = 2^{\Phi(r_0) - \Phi(r_{\text{min}})} / \Delta V \), and
\[
\ln G = \frac{\ln 2}{\Delta V} \frac{\lambda}{2\pi \epsilon_0} \ln \frac{\lambda}{2\pi \epsilon_0 r_0 E_{\text{min}}}
\] (9.4)

We assume \( E_{\text{min}} \approx 10 \text{kV/cm} \) and \( \Delta V \approx 41 \text{V} \), based on GARFIELD simulations of a 90% Ne–10% CH\(_4\) mixture. Equations 9.1 and 9.4 can be solved to obtain anode voltages for the case of \( G = 10^4 \) (bulk MWPC) and for \( G = 10^5 \) (trigger MWPC). The lower gain is acceptable at the bulk MWPC where all the track energy would be collected and the signal size would be appreciable. The higher gain will help amplify the trigger MWPC signal where only \( \sim \) keV-scale energy would be deposited. The trigger MWPC cathode will be grounded while the bulk MWPC cathode will be set at 2100 V to maintain continuity of the drift field. We obtain anode voltages of 1252 V for the trigger MWPC and 3217 V for the bulk MWPC. A plot of MWPC gain as a function of voltage is shown in Figure 9.6.
9.6 Design Tolerances

With fundamental design requirements in place, I now devote a section to discussing acceptable tolerances on implementing these requirements. The most important constraint is that of gas gain stability, because fluctuations in gain will degrade the energy and timing resolution. Fluctuations in gain are primarily caused by variation in charge density on a sense wire:

$$\frac{\partial G}{G} = \left[ \log G + \frac{\lambda \log 2}{2\pi \epsilon_0 \Delta V} \right] \frac{\partial \lambda}{\lambda}$$

(9.5)

Such variations could occur with time, because of HV supply fluctuations, or as a function of physical coordinates in the MWPC, because of variability in the mechanical suspension of the wires. The tolerance on gas purity is also important because electronegative impurities can absorb ionization deposited in the BetaCage causing signal deterioration.

9.6.1 HV supply

Equation 9.4 combined with Equation 9.1 gives the relationship of gain on the anode and cathode voltages. For the BetaCage, a 2% tolerance on gain fluctuation from the HV supply translates to a 1 V tolerance on the voltages, which is readily achieved with standard HV supplies.

9.6.2 Wire diameter variation

According to [229], single wire diameter variation causes line charge variation that is nearly independent of geometry

$$\frac{\partial \lambda}{\lambda} \approx 0.2 \frac{\partial r_0}{r_0}$$

(9.6)
Thus a 3% variation in anode wire diameter translates to \( \sim 10\% \) variation in gain.

### 9.6.3 MWPC wire positioning

A full analytical treatment of gain fluctuation caused by wire displacements is provided in [229], and here I just summarize the results. Displacement of a single wire, denoted 0, leads to changes in the line charge density of the displaced wire, as well as its neighboring wires, \( j \). These changes can be written as the sum of leading-order terms of a perturbative expansion in terms of displacements of the wire in its plane (\( \delta x \)), and perpendicular to the plane (\( \delta z \)).

\[
\frac{\partial \lambda_0}{\lambda} \approx a_2^{(0)} \left( \frac{\delta x}{s} \right)^2 + b_2^{(0)} \left( \frac{\delta z}{s} \right)^2 \\
\frac{\partial \lambda_j}{\lambda} \approx a_1^{(j)} \left( \frac{\delta x}{s} \right)^2 + b_2^{(j)} \left( \frac{\delta z}{s} \right)^2
\]  

(9.7)

Here, \( s = 0.5 \text{ mm} \) is the wire pitch, and the coefficients \( a_2^{(0)} = 0.2, b_2^{(0)} = 0.4, a_1^{(1)} = 0.11, \) and \( b_2^{(1)} = 0.06 \) are obtained from [229] for the wire radius, pitch, and plane spacing for the BetaCage. The resultant gain change is plotted in Figure 9.7 for bulk MWPC anode wires. Imposing a 5% gain tolerance requirement on wires, we require wire-centering error not to exceed 100 microns. This is achievable as described in the next chapter.

![Figure 9.7: Fractional gain change as a percentage vs. wire displacement. The solid curves show displacement within the plane and the dotted curves show displacement perpendicular to the plane. Wire 0 is the perturbed wire, and Wire 1 denotes the nearest neighbor wire. Plot taken from: [230].](image)
9.6.4 MWPC wire tensioning

Wires of an MWPC are subject to gravitational sag as well as repulsive electrostatic forces that displace them from their intended positions. This is countered by tensioning the wires. The sag produced in a wire of length $L$ and density $\rho$, tensioned with force $T$ is the following [224]:

$$\delta x = \frac{C}{k^2} \left( \frac{1}{\cos(kL/2)} - 1 \right)$$  (9.8)

where $C = \rho g \sigma / T$, $k^2 = 2\pi \epsilon_0 E_0^2(s)/T$ and $E_0(s)$ denotes the electric field of the wire at a distance of the wire pitch $s$. For the BetaCage design, this works out to displacements of under 30 $\mu$m for wire tensioning with 20-gram weights for the 25-$\mu$m-diameter wires and 200-gram weights for the 125-$\mu$m-diameter wires. This ensures that gain fluctuations due to electrostatic forces and gravitational sag will be under 1%, based on the discussion in Section 9.6.3.

9.6.5 Gas purity

With a 40-cm-tall drift region, electron loss due to attachment with electronegative impurities such as $O_2$ and $H_2O$ molecules in the gas is non-negligible. Too much electron attachment can severely degrade signal amplitude and resolution. However, with knowledge of gas mixture properties and drift cage design specifications, we can set constraints on the purity requirements for the gas mixture, following a description in [223].

The mean free path of an electron is given by

$$\lambda = \frac{1}{N\sigma(\epsilon)}$$  (9.9)

where $N = 2.69 \times 10^{19}$ cm$^{-3}$ is the molecular number density of any gas at STP and $\sigma$ is the collision cross section of electrons with gas molecules of that particular mixture. The cross section depends on $\epsilon$, the characteristic energy of electrons, which in turn depends on the drift field. For a drift field of 50V/cm in 90% Ne–10% CH$_4$, $\epsilon = 0.1$ eV [231] and the resulting cross section is $8.5 \times 10^{-17}$ cm$^2$ [231].

Now, the rate of collisions per unit time is just $u/\lambda$ where $u = \sqrt{2\epsilon/m}$ is the average instantaneous speed of electrons with mass $m$ and characteristic energy $\epsilon$. We multiply this gross collision rate with the impurity fraction, $p$, to get the collision rate of electrons with molecules of impurity, $pu/\lambda$.

Next, we introduce the attachment probability, $h$, an empirically determined quantity. For air ($O_2$, $H_2O$, etc.) at STP with 0.1 eV electrons in 50 V/cm drift field, $h = 3.3 \times 10^{-5}$ [232]. Multiplying the impurity collision rate with $h$ and dividing by the drift velocity of electrons $w$ (1.7 cm/µs for 50 V/cm in 90% Ne–10% CH$_4$ [231]), we get the average number of impurity attachments per drift
length. The inverse of that, i.e., the mean attachment drift length is thus given by

$$\lambda_c = \frac{\lambda w}{h \rho u} = \frac{w}{N\sigma(\epsilon)hp} \sqrt{\frac{m}{2\epsilon}}$$  \hspace{1cm} (9.10)

The electron loss fraction in a drift of length $l$ is $1 - \exp(-l/\lambda_c)$. In Figure 9.8 I plot the electron loss fraction as a function of the impurity fraction $p$. Based on this calculation, the impurity level in our 90% Ne–10% CH$_4$ mixture needs to be lower than 0.13% to keep electron loss under 5%. This is trivially achieved with the gas purification system mentioned in Section 9.3, since the ratio of materials outgassing rate in the pressure to the flow rate is $<< 0.1\%$.

![Graph](image.png)

Figure 9.8: Drift electron loss fraction vs. gas impurity fraction for 90% Ne–10% CH$_4$ mixture in 50 V/cm drift field.

### 9.7 Expected Signal in Bulk MWPC

In this section, I model the expected signal in the bulk MWPC. I use the conclusions from this section to motivate some of the decisions made with regard to readout electronics in the following section.

#### 9.7.1 Approximate pulse shape

As electrons from the ionization track, generated in the field cage, drift into the bulk MWPC, they follow field lines that terminate on the anode wires. The avalanche process occurs at a few microns from the anode surface, producing a large number of electron-ion pairs, given on average by the MWPC gain multiplied by the original number of electrons. Avalanche electrons and ions now drift
towards their respective electrodes; the motion of this charge produces image current, which we
record as signal, in the anode and cathode wires. The electrons travel only a few microns before
being absorbed by the anode and contribute $\sim 1\%$ to the signal. For a simple analytic calculation,
we can treat the entire signal as generated by the drift of ions to the cathode. We can compute the
shape and time evolution of the signal using Ramo’s Theorem \[233\].

Note that a precise description of the signal shape requires knowledge of the full MWPC potential,
the arrival times of track electrons at the MWPC anodes, and the exact path followed by avalanche ions to the cathode wires. In principle, an average pulse shape could thus be obtained by numerical simulation with GARFIELD. However, much can be gleaned with an analytical calculation, along the lines presented in \[224\], if one makes some simplifying assumptions.

Since the anode radius is much smaller than wire pitch or grid spacing, the electric field within
the avalanche region, i.e., a few anode radii of the wire, can be described by the electric field of a
single anode wire. Additionally, if the readout electronics only integrate the ion drift over the first few tens of anode radii, then the true signal shape would be closer to that presented in this simplified analysis.

We rewrite the electric field for an anode wire using Equations \[9.1\] and \[9.2\]

$$E_n(r) = \frac{\lambda_n}{2\pi \varepsilon_0 r} = \frac{1}{2\pi \varepsilon_0 r} \left( \sum_{m=1}^{N} c_{nm} V_m \right) \quad (9.11)$$

where the charge on the wire is determined by the capacitance matrix $c_{nm}$ and the potential $V_n$ on
a wire.

The velocity of the ions is $v = \mu E$ where $\mu$ is the ion mobility and is $3.7 \text{ cm}^2 \text{ V}^{-1} \text{s}^{-1}$ for a
90% Ne–10% CH$_4$ mixture$^{[224]}$. This gives for the ion trajectory,

$$\frac{d r(t)}{dt} = \frac{\mu}{2\pi \varepsilon_0 r(t)} \left( \sum_{m=1}^{N} c_{nm} V_m \right) \rightarrow r(t) = a \sqrt{1 + \frac{t}{t_0}} \quad (9.12)$$

where the characteristic time $t_0$ is given by

$$t_0 = \left[ \frac{\mu}{a^2 \pi \varepsilon_0 \left( \sum_{m=1}^{N} c_{nm} V_m \right)} \right]^{-1} \quad (9.13)$$

We apply Ramo’s Theorem \[233\] and ground all electrodes except the one at which the avalanche
occurs. The image charge induced on the electrode at an instant is given by

$$Q_n = -\frac{N_{tot}\varepsilon_0}{V_n} U(\vec{r})$$

$^{1}\mu_{Ne^{+}} = 4.14 \pm 0.2 \text{ cm}^2 \text{ V}^{-1} \text{s}^{-1}$ and $\mu_{CH_{4}^{+}} = 1.87 \text{ cm}^2 \text{ V}^{-1} \text{s}^{-1}$
Figure 9.9: The red dot-dashed line represents the signal injected on an anode wire, post-avalanche, from a 1-keV energy deposition in the drift cage, assuming instantaneous arrival of electrons at the wire. This $1/t$ signal is slower than an exponential of same characteristic time $t_0$, shown as a green dotted line. The solid blue line is the signal smeared with a gaussian shape to approximate the spread in arrival time of the electron cluster.

where $U(\vec{r})$ is the potential due to the wire at the location of the drifting ion and $N_{\text{tot}}$ is the total number of electrons after the avalanche. The induced signal on the wire is given by the derivative of the equation above,

$$I_{\text{sig}}(t) = -\frac{dQ_n}{dt} = -\frac{N_{\text{tot}} e_0}{V_n} E_n(\vec{r}) \frac{dr}{dt} = \frac{-N_{\text{tot}} e_0}{4\pi\epsilon_0} \frac{c_{nn}}{t + t_0}$$ (9.15)

Cleaning up factors, we rewrite the induced signal as

$$I_{n\text{sig}}(t) = \frac{I_0}{1 + \frac{t}{t_0}}$$ (9.16)

where

$$I_0 = \frac{-N_{\text{tot}} e_0 c_{nn}}{4\pi\epsilon_0} \frac{1}{t_0}$$ (9.17)

Since $I_0$ is proportional to number of charges generated by the avalanche, which in turn is proportional to the number of track electrons, the signal peak height can be used to measure energy deposited in an event. For the BetaCage, $c_{nn} \sim 9 \text{pF/m}$ (from Table 9.1) and $t_0 = 1.5 \text{ns}$. For a gain of $10^4$ and a point deposition of 1 keV, $I_0 \sim 3 \mu\text{A}$. Note that this $1/t$-shaped signal is slower than a decaying exponential of the same characteristic time.

We can approximate the spatial spread of the avalanche electrons by convolving the pulse shape in Equation (9.16) with a gaussian of mean $5\mu\text{s}$ and standard deviation of $2\mu\text{s}$. Both the original
pulse and the smeared pulse are plotted in Figure 9.9. Since this convolution is linear, the pulse peak height remains proportional to the energy deposited in the system.

### 9.7.2 Signal attenuation in network

The signal computed above is modified by the passive network of the HV biasing and filtering circuit and other sense wires, before it can be amplified and recorded. In this section, I model these effects and compute the network transfer function, following Walt Innes’ treatment of signal propagation in BaBar Drift Chamber system [234] as well as a general treatment of transmission lines in [235].

A sense wire in the MWPC can be modeled as a lossy transmission line, unterminated at one end, and terminated at the other by the impedance, in parallel, of other sense wires in its gang and that of the HV filter and amplifier. The characteristic impedance of a generic transmission line with a lossless dielectric is

$$Z_0(\omega) = \sqrt{\frac{R + j\omega L}{j\omega C}}$$  \hspace{1cm} (9.18)

where $R$, $C$, and $L$ are the resistance, capacitance, and inductance per unit length and $\omega$ is the frequency being considered. These parameters are listed in Table 9.2 for the sense wire and cabling used to connect components.

The characteristic propagation constant is a factor multiplying the position or time arguments of the exponential in the wave function on the transmission line. It is given by

$$\gamma(\omega) = \sqrt{(R + j\omega L) j\omega C}$$ \hspace{1cm} (9.19)

<table>
<thead>
<tr>
<th>Component</th>
<th>Resistance (Ω/m)</th>
<th>Capacitance (pF/m)</th>
<th>Inductance (µH/m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anode wire</td>
<td>1422</td>
<td>9.0</td>
<td>1.2</td>
</tr>
<tr>
<td>SHV cable</td>
<td>0.033</td>
<td>82</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Table 9.2: $R, C, L$ for anode wire and SHV cable in BetaCage

Using formulae 9.18 and 9.19 one can compute the impedance, looking into a transmission line of length $l$, terminated by load impedance $Z_L$. I quote the result here from [235].

$$Z_{in} = \frac{Z_L + jZ_0 \tanh \gamma l}{Z_0 + jZ_L \tanh \gamma l}$$  \hspace{1cm} (9.20)

Note that the characteristic impedance is not the actual impedance of the line, unless the load is “matched” ($Z_L = Z_0$). For a BetaCage sense wire, we learn from this expression that the avalanche signal encounters different impedances in opposing directions along the length of the wire, since the length to the end of the transmission line appears in the formula, as does the subsequent cascaded
impedance connecting the wire to the outside world. I use the block diagram in Figure 9.10 to represent the transmission line equivalent of a channel of the MWPC. The current $I_s$ represents the signal injected at location $l$ from the terminated end of a sense wire of length $L_w$. At this point it splits into $I_0$ and $I_1$.

$$I_s = I_0 + I_1$$  \hspace{1cm} (9.21)

Setting the load impedance, $Z_L = 0$ in Equation 9.20 gives the impedance to the unterminated end, $Z_w \coth \gamma_W (L_w - l)$. Thus, $V_1$ is given by

$$V_1 = Z_w \coth [\gamma_W (L_w - l)] I_0$$  \hspace{1cm} (9.22)

The section of the wire between the avalanche point and the terminated end can be modeled as the following two-port:

$$\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} Z_w \coth \gamma_W l & Z_w \csch \gamma_W l \\ Z_w \csch \gamma_W l & Z_w \coth \gamma_W l \end{pmatrix} \begin{pmatrix} I_1 \\ -I_k \end{pmatrix}$$  \hspace{1cm} (9.23)

Beyond this point, the current $I_k$ splits into a portion $I_l$ that flows into other sense wires in the same gang as the source wire, and $I_2$ that continues in readout cabling.

$$I_k = I_l + I_2$$  \hspace{1cm} (9.24)

The impedance of the sense wire, looking into it from the termination end is $Z_w \coth \gamma_W L_w$. Thus for a gang of $n$ wires, $n - 1$ wires, in parallel with the source wire, have impedance $Z_g = \frac{Z_w \coth \gamma_W L_w}{n - 1}$. The potential $V_2$ drops across this parallel set of wires, so the following relation holds:

$$V_2 = \frac{Z_w \coth \gamma_W L}{n - 1} I_l$$  \hspace{1cm} (9.25)
Current $I_2$ flows into standard coaxial cable, which is modeled with the following two-port:

\[
\begin{pmatrix}
V_2 \\
V_3
\end{pmatrix} = \begin{pmatrix}
Z_C \coth \gamma_C L_C & Z_C \csch \gamma_C L_C \\
Z_C \csch \gamma_C L_C & Z_C \coth \gamma_C L_C
\end{pmatrix}
\begin{pmatrix}
I_2 \\
-I_3
\end{pmatrix}
\tag{9.26}
\]

Finally, $I_3$ flows into the outside world, i.e., the amplification and HV biasing/filtering circuits. The combined impedance of this I write as $Z_{out}$. Thus, the following equation holds:

\[
V_3 = Z_{out} I_3
\tag{9.27}
\]

Equations 9.21 through 9.27 can be simultaneously solved to obtain the fraction of the input signal that makes it to the outside world, $I_3/I_s$.

\[
I_3/I_s = -(Z_{C2} Z_l Z_{W2} Z_g)/\left\{Z_{C2}^2[(Z_l + Z_{W1})(Z_g + Z_{W1}) - Z_{W2}^2]ight. \\
- (Z_{C1} + Z_{out})(Z_{C1}[(Z_l + Z_{W1})(Z_g + Z_{W1}) - Z_{W2}^2] \\
+ Z_g(Z_{W1}(Z_l + Z_{W1}) - Z_{W2}^2))\right\} \tag{9.28}
\]

where

\[
\begin{align*}
Z_{W1} &= Z_W \coth \gamma_W l \\
Z_{W2} &= Z_W \csch \gamma_W l \\
Z_{C1} &= Z_C \coth \gamma_C L_C \\
Z_{C2} &= Z_C \csch \gamma_C L_C \\
Z_l &= Z_W \coth[\gamma_W (L_W - l)]
\end{align*}
\]

Equation 9.28 is plotted as a function of frequency in Figure 9.11. The current $I_3$ is divided between the HV biasing/filtering circuit and the amplifier as shown in the full readout circuit diagram in Figure 9.12. The HV circuit can be represented by the following equivalent impedance:

\[
Z_{bias} = R_b + R_f + \frac{j \omega C_f}{1 + j \omega C_f} \tag{9.29}
\]

The impedance looking into the amplifier end, combining the decoupling capacitor $C_C$ and amplifier input impedance $R_{in}$ is

\[
Z_{in} = \frac{1}{j \omega C_C} + R_{in} \tag{9.30}
\]

These two impedances add in parallel to form $Z_{out}$, which was used in Equations 9.27 and 9.28.
Figure 9.11: Transfer function of the readout network vs. frequency. The different curves correspond to the avalanche at distance 0% through 100% of wire length from the amplifier, in steps of 10%. At high frequency, the farther the avalanche is from the amplifier, the faster is the roll-off.

Figure 9.12: Full readout circuit diagram. $Z_{\text{det}}$ is the equivalent impedance of an MWPC channel and cabling, as presented in the transmission line model of Figure 9.10. The current $I_3$ flows from the detector into the HV biasing/filter circuit and the amplifier.
Thus the current at the input of the amplifier $I_{in}$ is given by the following:

$$I_{in} = I_s \frac{I_3}{I_s} \frac{Z_{bias}}{Z_{bias} + Z_{in}}$$  \hspace{1cm} (9.31)

$H = I_{in}/I_s$ is the transfer function of the network and its amplitude is plotted in Figure 9.11. $I_{in}$ flows out through $R_{in}$ and produces voltage $V_{in}$ at the input of the amplifier. The voltage signal is plotted in the time domain in Figure 9.13.

Figure 9.13: Pulse at amplifier input vs. time. The different curves correspond to the avalanche at distance 0% through 100% of wire length from the amplifier, in steps of 10%. At high frequency, the farther the avalanche is from the amplifier, the smaller the pulse. For reference, the original signal injected at the avalanche site is displayed as a dotted-dashed magenta line.

### 9.7.3 System optimization based on features of unamplified signal

From the calculation in Section 9.7.2, we learn that, compared to the input pulse, the shape of the signal at the input of the amplifier is altered and its amplitude is diminished by the readout network. The former is controlled by the bandwidth of the network while the latter is affected by shunting of signal by wires, cabling, and electrical components. I note a few important effects here and comment on tweaking of network parameters to maximize output signal amplitude.

#### 9.7.3.1 Position dependence of pulse characteristics

As shown in Figure 9.13 we note a 13% change in output signal amplitude between the two ends of a sense wire because of series and shunt resistive losses. Without information on the location of the avalanche along the wire, this would lead to a corresponding degradation of energy resolution.
This is the motivation for orienting the cathode plane perpendicular to the anode plane and reading it out. Additionally, the pulse shape varies along the length of the wire because of changes seen by the avalanche signal in net capacitance to the unterminated end of the sense wire.

9.7.3.2 Effect of electrical component choice on pulse amplitude

From Equation 9.31 we see that $\frac{Z_{bias}}{Z_{bias} + Z_{in}}$ must be maximized. This is easily achieved in practice as bias resistors are much larger than amplifier impedances. Most of the signal loss in this circuit occurs because of the choice of lossy stainless-steel sense wire. However, this choice is dictated by radiopurity requirements and no good alternatives exist. Another source of signal loss is cabling between the MWPC and the amplifier. In the model above we assume 1 m of cabling between the MWPC and the first stage amplifier. Again, radiopurity requirements dictate that the amplifier circuitry be located outside the active measurement volume of the BetaCage. Thus, with careful design of the gas containment vessel, connection cabling length can be minimized.

9.7.3.3 Effect of electrical component choice on pulse shape

Signal bandwidth is controlled by various capacitances and resistances in the circuit. The transfer function roll-on, which determines the length of the pre-amplified pulse, is controlled by the time constant $\frac{1}{R_C C}$. In general, the roll-off, or equivalently the sharpness of the pulse, is determined by the time constant $\frac{1}{R_{in} C_d}$, where $C_d$ is a generic detector capacitance. In our case, this is the combined effective capacitance of the gang of wires and cabling. We note that changes in roll-off frequency by a factor of few has intangible effects on pulse shape and that the pre-amplifier bandwidth is sufficient to preserve pulse shape and timing information observed in the input pulse.

9.8 Readout Electronics

Based on pulse characteristics determined in the previous section, we can outline a design for readout electronics to record signals.

9.8.1 Signal amplification

First, the signals from all MWPCs, at the submillivolt level, will require amplification by a factor ~100 prior to recording. We require a cheap, high-bandwidth, low-noise amplifier that can be easily incorporated into our electronics fabrication. Several viable options are available in the market, and we pick one, MAR-6+, manufactured by Mini-Circuits Inc., based on expert recommendations. This is a 2-GHz-bandwidth surface-mount amplifier with a gain of 22 dB and a noise figure of 3 dB and costs $1.16 per piece. We will cascade two of these per channel to obtain required gain.
9.8.2 Digitization

As noted earlier, event energy can, in principle, be estimated by simply measuring pulse height or integrating the signal to obtain total charge deposited. However, with today’s cheap ADC technology, we can perform full waveform digitization and capture the entire shape of incoming pulses at comparable cost to charge integration. This would enable use of the entire pulse shape to aid position reconstruction and better energy estimation, while still ganging wires to keep the channel count low.

9.8.2.1 ROACH

We plan to use a readout system using the open-source ROACH architecture, encouraged by the success of our colleagues in sub-millimeter astronomy with such systems. ROACH (Reconfigurable Open Architecture Computing Hardware) is a standard, developed and supported by CASPER (Center for Astronomy Signal Processing and Electronics Research) at UC Berkeley, with a large user-base and excellent technical support. Using this standard, Digi-Com Electronics manufactures “ROACH boards,” which are Virtex-5 FPGA-based computers with on-board RAM and ethernet interfaces for communication with a data acquisition PC and with other ROACH boards. The ROACH boards use a Z-DOK daughter board interface to enable digitizer cards to be connected to them to accept data. A picture of a ROACH system with two digitizer boards is in Figure 9.14.

The FPGA on the ROACH board can be programmed using standard CASPER-supported tools and MathWorks Simulink to perform a host of signal processing tasks such as triggering, FFTs, downsampling, etc., allowing great flexibility in customizing hardware processing of signal pulses.

![ROACH board with two digitizer daughter boards](image)

Figure 9.14: ROACH board with two digitizer daughter boards. Courtesy: Digi-Com Electronics Inc.
9.8.2.2 Digitizer boards

Techne Instruments\(^2\) has designed the digitizer boards for BetaCage using ADS5400 1-GSa/s, 2-GHz bandwidth, 12-bit ADC chips from Texas Instruments. These boards will each record six channels. A set of eight boards (i.e., 48 channels) will be interconnected and will each be capable of triggering the entire system. Once triggered, all boards will dump traces via ethernet to a DAQ PC.

9.9 Electronics Noise

In this section, we compute the noise contribution to readout from the electronics chain, ignoring for the moment avalanche statistics, gain stability, and other sources of measurement uncertainty, which will be addressed in Section 9.10. The two main sources of electronics noise are the passive network which routes the signal to the amplifier, and the amplifier itself.

9.9.1 Thermal noise from network

Every resistive element generates a frequency-independent fluctuating voltage across itself, called thermal or Johnson noise, caused by Brownian motion of charge carriers \(^{236, 237}\). The square root of the thermal power noise spectrum is given by

\[
e_n = \sqrt{4kT R}
\]  

(9.32)

which has units of V/$\sqrt{\text{Hz}}$. Here, $k$ is the Boltzmann constant, $T$ is the temperature of the resistive element, and $R$ is the value of the resistance.

The resistive elements of our network are the bias and HV filtering resistors, and the resistance of the wire and cabling. Note that the input impedance of the amplifier is ignored here, but taken into account in the next subsection, where I calculate amplifier noise. While each of these elements generates thermal noise independent of one another, we only care about the summed voltage noise that appears at the amplifier input. According to Thevenin’s theorem, all voltage sources in a linear network can be replaced by a single impedance and a series voltage source \(^{238}\). Additionally, it can be shown that a passive network with a complex impedance $Z_{\text{net}}(f) = R_{\text{net}}(f) + jX_{\text{net}}(f)$ has thermal noise power spectrum

\[
e_{\text{net}}(f) = \sqrt{4kT \Re(Z_{\text{net}}(f))} = \sqrt{4kT R_{\text{net}}(f)}
\]  

(9.33)

We need to compute $Z_{\text{net}}$ from Figure 9.12. Working from the HV filter towards the amplifier, the

\(^2\)4920 Telegraph Ave #G Oakland California, 94609, USA rikraf@earthlink.net
summed impedance of the HV filter is given by

\[ Z_{HV} = \left( \frac{1}{R_f} + j\omega C_f \right)^{-1} \]  

(9.34)

To this we add the bias resistor, thus obtaining the same \( Z_{bias} \) as in Equation 9.29.

\[ Z_{bias} = Z_{HV} + R_b \]  

(9.35)

Then, we look at the impedance of a gang of sense wires and associated cabling, by invoking Equation 9.21. The impedance of the cable, terminated by \( n \) “unterminated” sense wires in parallel is given by

\[ Z_{MWPC} = Z_C \frac{Z_W \coth \gamma_W L_W + n j Z_C \tanh \gamma_C L_C}{n Z_C + j [Z_W \coth \gamma_W L_W] \tanh \gamma_C L_C} \]  

(9.36)

We add this and \( Z_{bias} \) in parallel,

\[ Z_{pre-amp} = \left( Z_{bias}^{-1} + Z_{MWPC}^{-1} \right)^{-1} \]  

(9.37)

Finally, we get

\[ Z_{net} = Z_{pre-amp} + \frac{1}{j\omega C_C} \]  

(9.38)

This is plugged into Equation 9.33 to obtain the thermal voltage noise contribution, and is plotted in Figure 9.15.

### 9.9.2 Amplifier noise

The amplifier has a quoted noise figure at input, \( NF \), which is converted into a noise temperature using the following formula:

\[ T_{amp} = (NF - 1) T_0 \]  

(9.39)

where \( T_0 \) is 290 K by convention. This temperature and the input impedance of the amplifier are plugged into Equation 9.32 to get the voltage noise contribution of the amplifier at input,

\[ \epsilon_{amp}(f) = \sqrt{4kT_{amp} R_{amp}} \]  

(9.40)

which is plotted in Figure 9.15. The noise from \( R_{in} \) is included in the noise figure.

We note that the amplifier noise is higher than the thermal noise. Ideally, we would like to match the two, since the latter represents a ceiling on the lowest noise floor achievable for the configuration presented here. In practice, as I show in the next section, the theoretical resolution achieved with this configuration is sufficient to obviate more expensive amplifiers with lower noise characteristics.
9.9.3 Total electronic noise

To compute this, we add the components in quadrature.

\[ \epsilon_{\text{tot}} = \sqrt{\epsilon_{\text{amp}}^2 + \epsilon_{\text{net}}^2} \]  

(9.41)

This quantity is plotted in Figure 9.15. Squaring it gives \( J(f) \), the noise power spectrum, quoted in units of \( V^2/\text{Hz} \).

![Figure 9.15: Noise (V/\sqrt{\text{Hz}}) vs. frequency. The dot-dashed green line represents the thermal noise of the passive readout network at the amplifier’s input. The dot-dashed blue line represents the amplifier noise for MAR-6+, also referenced at input. The solid red line is the sum of these two components. Finally, the solid teal line is the total noise at the amplifier’s output.](image)

9.10 Energy Resolution

Our ability to resolve spectral lines and beta endpoints to identify contamination of samples relies on reasonable energy resolution. In this section, I combine the effects of electron-ion pair and avalanche statistics with readout electronics noise and digitization noise to compute the expected energy resolution as a function of true energy of a track.

9.10.1 Electron-ion pair generation and avalanche statistics

Ionization along a particle trajectory is a random process governed by Landau statistics \[223\]. If \( \bar{n} \) is the mean number of electron-ion pairs generated by a particle of energy \( E \) in the BetaCage
drift chamber, then the variance, \( \sigma_n^2 \) on actual counts is \( f_i \bar{n} \). Here, \( f_i \) is a Fano factor that enables comparison with Poisson statistics.

The avalanche process at a sense wire is also a random process. For MWPCs, the probability distribution for an avalanche sourced by a single electron is closely approximated by a Polya function with mean equal to the mean number of electrons after the avalanche \( \bar{A} \) and variance \( \sigma_A^2 = f_A \bar{A} \), where \( f_A \) is the Fano factor for the avalanche process. To convert to a variance for \( \bar{n} \) original electrons, we multiply the single-electron variance by \( \bar{n} \), per central limit theorem.

Thus ionization and avalanche statistics combined follow a distribution with mean \( \bar{N} = \bar{n} \bar{A} \) and variance given by the following:

\[
\sigma_N^2 = \bar{n} \sigma_A^2 + \bar{A}^2 \sigma_n^2 = \bar{n} f_a \bar{A} + \bar{A}^2 f_i \bar{n}
\]

\[
\Rightarrow \sigma_N^2 \sim \bar{n} \bar{A}^2 f_i
\]

where the last approximation holds because \( f_a \ll \bar{A} f_i \). We can also rewrite Equation 9.43 in terms of the average measured event energy as

\[
\sigma_E^2 \sim k^2 \sigma_N^2 = k \bar{E} \bar{A} f_i
\]

where \( k \) is the energy per electron received at a sense wire, and using \( \bar{E} = k \bar{n} \bar{A} \).

For pure argon and neon, \( f_i \sim 0.17 \) \[240\]; for argon-methane mixtures, \( f_i \sim 0.2 \) \[241\]. Similarly, \( f_a \sim 0.6-0.8 \) for argon-methane mixtures \[239\]. I was unable to locate any predictions or data on \( f_i \) or \( f_a \) for neon-methane mixtures, so I conservatively assume \( f_i = 0.5 \) and \( f_a = 1 \). With average amplification per electron \( \bar{A} = 10^4 \) and average ionization energy \( W = 36 \text{ eV} \) \[224\], \( k = W/\bar{A} = 3.5 \text{ meV} \). Thus, the energy resolution from statistical fluctuation of pair creation and avalanche processes is \( \sigma_E \sim \sqrt{0.007 \bar{E}} \) for \( \bar{E} \) in keV. It is plotted as a function of energy in Figure 9.16.

### 9.10.2 Readout resolution

In Section 9.7.2 I computed signal amplitude at the amplifier input. Using that information and the total noise spectrum computed in the previous section, we can estimate the best possible energy resolution achieved using an optimal filter. Given an expected signal pulse shape and an estimate of the noise power spectrum, an optimal filter produces an unbiased estimate of the signal peak height and the one with best possible noise. Several good sources \[71\] exist on optimal filtering theory and implementation. Thus, I will skip details and simply state the expression for optimal
Figure 9.16: Contributions to single-channel predicted energy resolution from digitization noise (dotted red curve), readout electronics noise (dashed green curve), statistical fluctuation in pair creation and avalanching (dashed-dotted blue curve) and the total (solid teal curve) on a single channel. These are all plotted as a function of true energy deposited on a single wire in a single channel, under the assumption of perfect knowledge of avalanche positions.

The energy resolution per channel for the BetaCage is \( \sim 0.35 \text{ keV} \).

9.10.3 Digitization noise

The ADS5400 ADC chip has analog input range of \( 2V_{pk-pk} \) and ENOB (effective number of bits) \( \sim 9.3 \). Thus, the RMS digitization noise at output is given by dividing the voltage range by number of effective digitization bins and multiplying by \( q \), the energy per unit signal amplitude at the digitizer input.

\[
\sigma_{dig} = \frac{qV_{pk-pk}}{2\text{ENOB}}
\]  

(9.46)

For the BetaCage, \( q \sim 400 \text{ keV/V} \), after averaging over nominal position dependence effects. The RMS digitization noise is \( \sim 1.2 \text{ keV} \) per channel for the BetaCage, which is \( \sim 4 \) times higher than the amplifier noise and, thus, is an undesirable situation. ADC chips with higher resolution and
sufficient sampling rate were not available cheaply, so we decided to continue using the ADS5400. However, this problem is alleviated by adding a buffer with gain of \(\sim 4\) after the MAR-6+ amplifier chain. The resulting digitization noise is plotted in Figure [9.16](#).

### 9.10.4 Summed energy resolution

Adding in quadrature all the resolution terms discussed above, we get the solid teal curve in Figure [9.16](#). Note that this estimate is made assuming charge deposition on a single wire, and assuming the availability of perfect information of avalanche position so that a suitable optimal filter can be applied. Residual position dependence will add a linear term in energy to the resolution, which can be as high as \(0.1E\). Assuming that we reconstruct event energy by summing 15 fiducial channels, the total energy resolution is obtained by multiplying the single channel resolution by a factor of \(\sqrt{15}\).

### 9.11 Backgrounds and Expected Sensitivity

As noted throughout this chapter, special care has been taken to identify low-background materials for construction, ensuring that the background event rate from the counter itself is minimized. All of the readout electronics, including first-stage amplifiers will be located outside the containment vessel and the shield to minimize the gamma rate from their U/Th content. The backgrounds contribution of the components inside the detector are estimated in the following way. The U/Th activity of different materials was obtained in Bq/kg from the ILIAS materials radiopurity database [242], and in the case of the Noryl, used for structural support, was directly counted using the GOPHER detector [243]. Where reasonable, a geometry factor was estimated for each component based on the fraction of photons emitted by it that would result in triggers. For the Noryl structure, the geometry factor was obtained through a simulation in GEANT4 by Boqian Wang. Multiplying this number with the activity and mass of a material yielded its estimated background rate. These rates are listed in Table [9.3](#).

Table 9.3: Estimated rates of photon-induced triggers from background contamination of materials in BetaCage

<table>
<thead>
<tr>
<th>Material</th>
<th>Rate (cm(^{-2}) keV(^{-1}) day(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(^{14})C in 10% methane quench</td>
<td>(10^{-9})</td>
</tr>
<tr>
<td>Kr at 500 ppt in neon</td>
<td>(1 \times 10^{-7})</td>
</tr>
<tr>
<td>Stainless steel wires</td>
<td>(\approx 10^{-8})</td>
</tr>
<tr>
<td>Resistor chain</td>
<td>(\approx 10^{-6})</td>
</tr>
<tr>
<td>Noryl support structure</td>
<td>(2 \times 10^{-6})</td>
</tr>
</tbody>
</table>

However, the ultimate limiting background for the BetaCage is the ambient photon background.
in the operating environment. The nominal plan for the BetaCage is to be installed at the Low Background Counting Facility (LBCF) in Soudan Underground Laboratory, at a depth of 2030 meters of water equivalent. The LBCF has a 280-square-foot class-10,000 clean room with a 5'×8' experimental space reserved for BetaCage. The entire 35-m × 40-m × 100-m Soudan-2 detector hall, where LBCF is located, is surrounded by an active veto shield, with additional veto tubes hung above the clean room itself. This will enable rejection of events co-incident with cosmic shower activity. LBCF will also provide an 8" lead shield to enclose BetaCage with 10-cm of ultra-low-background electroformed copper as interior lining. This shield should reduce ambient background to a Ge equivalent rate of $\sim 1\text{ keV}^{-1}\text{kg}^{-1}\text{day}^{-1}$. A simulation by Luiz DeViveiros demonstrated that this would translate to a background rate of $6 \times 10^{-5}\text{cm}^{-2}\text{keV}^{-1}\text{day}^{-1}$ in the BetaCage [244].

To achieve sensitivity lower than this rate, we will subtract the background rate. For every run, we would first establish a background rate to some precision by counting for a length of time without the sample. Then we would introduce the sample and count for the amount of time expected to produce the target rate as a $3\sigma$ excess over the background rate established in the previous step. Figure 9.17 demonstrates required times for background and sample counting to reach a target sensitivity of $10^{-5}\text{cm}^2\text{keV}^{-1}\text{day}^{-1}$ for a 0.64 m$^2$ sample. Thus, by counting background for $\sim 1$ day and then the sample for $\sim 2$ days, this sensitivity can be achieved in minimal time.

Figure 9.17: Total counting time vs. background counting time for a 0.64 m$^2$ sample in the BetaCage to achieve a sensitivity of $10^{-5}\text{cm}^2\text{keV}^{-1}\text{day}^{-1}$. The minimum of the curve is at 1 day of background counting plus 2 days of sample counting.
Chapter 10

Protocage: a BetaCage prototype

Before committing resources to building a 1-m\(^2\) radiopure counter, we decided that we would construct a smaller, radio-impure prototype. This would allow us to focus on the challenges of mechanical design and DAQ development for the full BetaCage, while enabling us to make corrections for design features that may pose unforeseeable challenges.

The prototype, dubbed “ProtoCage,” will be different from BetaCage in the following ways. It will have a quarter of the counting area of the BetaCage, i.e., 0.25 m\(^2\). This will make feature sizes smaller and, as I report in this chapter, has already enabled rapid and cheaper prototyping of support structures. Instead of a neon-methane mixture, ProtoCage will use significantly cheaper but radio-impure P-10 (90\%Ar-10\%CH\(_4\)), which has very well-documented properties as a drift chamber gas [223]. The range of 156 keV \(^{14}\)C-endpoint betas in P-10 is \(\sim 18\) cm, enabling us to implement a shorter 20-cm-tall drift cage for this prototype [227]. The assembly will be housed in an aluminum vacuum chamber with ports for connections to signal amplifiers, HV supplies, and to the gas supply. Note that we have strived to utilize low-radioactivity materials for fabrication of the physical structures of ProtoCage knowing that radiopurity will be important for BetaCage, but we have not taken sufficient precautions against radon plate-out during the assembly process.

The focus of this chapter will be the development of BetaCage MWPCs since they were predicted to be, and have in fact turned out to be, the most challenging part of the fabrication process. I also describe, very briefly, the DAQ and gas handling for the Protocage and plans for construction and of the ProtoCage and the full BetaCage.

10.1 Thin-frame Layered MWPC Units

The physical support structure of an MWPC allows suspension of anode and cathode wires within required positioning and tensioning tolerances (see Section 9.6). It also lets electrical connections to be conveniently made to the wires. Since we require three wire planes layered on top of one another for each MWPC, my first mechanical design for BetaCage MWPCs took a modularized approach
to construction. Each wire plane was stretched on a 5-mm-thick support frame with 0.25-m² outer area; three frames were mated together with removable fasteners to assemble a complete MWPC, as shown in Figure 10.1. An inner square with an area of 0.16 m² was cut out on each frame, allowing 80 wires per plane. Electrical contact to the ends of the wires was made on PCBs mounted on two sides of the frame. Radiopurity and cost requirements dictated plastic frames, which nominally have lower radioactivity than the cleanest stainless steel [242]. The modular design would allow the orientation of wire planes to be changed at will and would enable easy repairs in case of wire failures. We used 125-µm gold-plated beryllium copper for the cathode planes and 25-µm gold-plated tungsten for the anode wires, the most commonly used types of wire by several experiments for their excellent electrical properties [223, 224]. In the next few subsections, I list the various iterations we explored for these layered MWPC units along with the challenges and improvements along each step.

Figure 10.1: Left: CAD rendering of three wire planes mounted on plastic frames. Right: CAD rendering of MWPC unit formed by mating three wire planes.

10.1.1 UHMWPE

10.1.1.1 Design and Construction

The first iteration for an MWPC frame was the simplest design one could sketch. The Caltech Physics Machine Shop fabricated a single frame out of UHMWPE (Ultra-high molecular weight polyethylene). They took a plastic sheet of appropriate size, cut out the central square and then skimmed the part on four sides to reach the desired dimensions. In parallel, I used PCB123 software [245] to design a 42-cm × 2.5-cm area PCB of 1.5-mm thickness to make soldered connections with the wires. Each PCB had 80 0.7-mm-diameter metalized vias at 5-mm pitch into which wires would be soldered at a right angle from the direction of suspension. Each via was further connected to a metal pad for wiring to the outside world. An electrical track ran along the length of the PCB to short all wires, but could be broken where required to create gangs of adjacent wires. Two PCBs were attached to the UHMWPE frame in recessed areas on opposite sides of the frame and the entire
frame was mounted on an assembly jig. A wire was threaded through vias that were across from each other on the two PCBs and suspended with weights on both ends. This would let the wire stretch to the correct tension and bend across the corner of the vias. After a wire was suspended, it was soldered into place at both ends, under a microscope. For this first test, we used the thicker 125-µm-diameter wire. A picture of an UHMWPE frame with the PCBs and two wires soldered in place is shown in Figure 10.2.

![Figure 10.2: A picture of an UHMWPE frame with the PCBs and two wires soldered in place](image)

10.1.1.2 Problems

There were a few major problems with the UHMWPE frames, which led to some improvements for the next iteration.

1. **Frame rigidity:** At the beginning of assembly, it became immediately clear that UHMWPE was too soft a material to work with for a part with such skewed aspect ratio (i.e., thickness << length or width). The material lacked sufficient shear modulus to prevent it from bending under its own gravitational moment if held at one end. Figure 10.3 shows the frame being tweaked easily with little force. This was a problem because taking the frame off the assembly jig and moving it around caused the wires to lose tension.

2. **Dimensional instability during machining:** Along similar lines it was noted that the material was too unstable to assure better than 5% accuracy on any dimension, which is a factor of ~25 worse than what has been achieved with other plastics.
Figure 10.3: UHMWPE was too soft for use as a structural support material for wire planes. It is easily tweaked by hand.

3. **Wire positioning tolerance:** Another apparent problem was that attaching the wires through electrical vias at a right angle to the direction of suspension did not ensure repeatable z-positioning of the ends of the wire.

At this point, a Caltech HEP engineer, John Hanson, with prior experience in construction of wire chambers, joined the effort and provided me with tremendous help in brainstorming ideas and troubleshooting problems with the layered MWPC frames.

### 10.1.2 Delrin 570 frames

#### 10.1.2.1 Design and construction

The UHMWPE experience immediately suggested a couple of strategy changes for future iterations. First, we needed plastic with better mechanical properties, machining stability and with low thermal expansion coefficient. Delrin was identified as a candidate meeting these requirements and was used for this iteration. The grade selected for the prototype was Delrin 570 with 30% glass-fill, which would be problematic for the radiopure version. We decided to go ahead with it, nonetheless, and resolve assembly issues first.

The second change was in the wire suspension technique. I redesigned the PCB to replace the vias with flat electrical pads. Wires would now be laid flat across the frame in the assembly jig, flush with the surface of the PCB. Tensioning weights would be suspended over pulleys attached to
the side of the frame. The disadvantage of this scheme was the absence of a natural physical guide (such as the vias of the previous iteration) to center wires to the correct pitch. As a compromise, I added 0.5-mm-wide channels to the frame in which the wires could run to aid the alignment process. Under a microscope, the wire could be aligned to \( \sim 100 \) microns by eye. A picture of the frame with its PCBs and one of the assembly process is shown in Figure 10.4. In the future, the flat suspension design would permit us to use wire winding machines to quickly wrap wire around a frame at the correct tension and would enable the technician (or student) to make mechanical and electrical connections fairly rapidly.

![Figure 10.4: Left: A Delrin 570 frame with PCBs for wire suspension is mounted on an assembly jig with small plastic pulleys. 0.5-mm-wide channels on the frame align with the PCB electrical pads and aid in centering the wire during soldering. Right: The wire is tensioned using a weight suspended over a sliding plastic pulley that aligns with the wire. Courtesy: John Hanson.](image)

10.1.2.2 Connections to the wire

In principle, solder presents an optimal way to provide both electrical and mechanical connections for wires. In practice, the process of soldering was tedious and time-consuming; one had to be careful not to disturb the alignment of the wire while soldering under a microscope. A couple of alternatives were explored and are listed below and pictured in Figure 10.5.

1. **Epoxy for mechanical connection:** We identified several epoxies with reasonable curing times, all of which enabled the wires to be secured in place accurately. Once this was done, the loose ends of the wire could be used to make electrical contact without attention to positioning accuracy. Hysol-1C, Scotchweld 2216 could be cured within 24 hours at room-temperature while Epo-Tek 353ND required oven curing at 80°C for an hour. Norland NEA-123HGA was the clear winner; it is a UV-activated epoxy with excellent properties that cures within 10 seconds of UV exposure. A close-up picture of bonds made using this epoxy is shown in Figure 10.6.
2. **Alternatives to soldering for electrical connection:** We explored the use of SPI conductive silver paint for electrical connections, as well as Circuit-Works 2400 conductive epoxy to make rapid electrical connections with the wires epoxied in place. Basic continuity checks showed a good connection. Our plan was to compare the noise characteristics of these two electrical joints with those of solder, but we abandoned it when the development of these layered MWPC units was superseded by the four-piece MPWC design discussed in Section 10.2.

![Figure 10.5: A sample Delrin 570 frame with sets of wires attached with different bonding techniques. From top to bottom: Two sets (four each of anode and cathode wires) using UV epoxy for mechanical connection and Circuit-Works 2400 conductive epoxy for electrical connection; Two sets (four each of anode and cathode wires) using UV epoxy for mechanical connection and SPI conductive silver paint for electrical connection; Two sets (four each of anode and cathode wires) soldered. Courtesy: John Hanson.](image)

### 10.1.2.3 Problems and possible mitigation

1. **Handling 25-µm-thickness wire:** We found it extremely difficult to handle and precisely align the thin 25-µm-thickness wire by eye under a microscope. Also, soldering to it was nearly impossible. Thus, any further progress with the layered MWPC frames would necessarily
require externally aided alignment of wires, using a combination of a wire winding machine and a precision transfer frame. Also, the conductive epoxy or paint reported above would be required for electrical contact instead of solder.

2. **Bowing/warping of frames**: Most plastic is industrially manufactured by extrusion, which leaves large residual internal stresses in the material. Machining of stock plastic typically causes imbalance in these stresses, which in turn leads to warpage and dimensional shift in parts over time. As it turns out, the MWPC frame discussed here is especially vulnerable to this problem because it is created by cutting out nearly 65% of the material from a Delrin stock piece with a skewed aspect ratio. If measured immediately after machining, the frames conform to required specifications; however, over the course of a few weeks, they start “potato-chipping” as shown in Figure 10.7. This is different from the issue of dimensional stability for UHMWPE where the material deformation occurs during the machining process due to thermal stresses. The bowing problem was first observed when frames were pulled off their assembly jigs and set aside to await other frames being layered on top. The problem is partially mitigated if the frame is always left attached to a stainless-steel supporting brace. We were able to mate frames together under this configuration and always leave one brace attached to the structure. However, in some cases, a subset of wires lost tension during the transfer process, rendering an entire frame unusable.
While the issues with the 25-µm-thickness wire seemed surmountable, attempts to solve the potato-chipping problem landed in road blocks. After exploring literature on the subject, we attempted, with some success, cyclic annealing and skimming of stock material \cite{246} to obtain the desired final shape — the bowing of the frame itself was minimized and frames could be transferred and stacked with the aid of a stainless-steel brace without a loss of wire tension. However, the 3-mm-thick ledges on the Delrin frame that housed the PCBs still acquire up to a 5% gradient in flatness, translating to a 10% smearing of energy resolution \cite{230}. Another possibility, in retrospect, would have been to custommold the rough shape of the frame with slightly larger dimensions, and then skim it down to the desired shape. This would have prevented internal stresses in the plastic at the casting phase, and thus would have obviated the problem of bowing in the frames.

### 10.2 Four-part MWPC units

The next redesign led to the current version of the BetaCage MWPCs. Its features are described in the first subsection. The implementation of the design, including solutions to problems along the way, is described in following subsection.
10.2.1 Design features

10.2.1.1 Bulkier plastic frame

The challenges presented by the layered MWPCs prompted me to rethink the MWPC design concept and change its unfavorable aspect ratio. We wanted the distance between wire planes to still be 5 mm, but wanted to increase the thickness of the frame. Thus I eliminate the layered modular scheme, in favor of using a monolithic part to hold all three wire planes. However, to avoid again having to cut out a large central area from a single plastic piece, I separated the frame into four “arms” that would mate to give the full MWPC, as pictured in Figure 10.8. Each arm would be machined from a stock piece with dimensions close to the final shape (i.e., from an extruded plastic bar instead of a sheet), thereby minimizing stress-induced deflection after machining. With precisely located fastening holes on the arms, the MWPC is easily assembled to ensure repeatable positioning of wire ends.

![CAD rendering of redesigned MWPC. Four arms mate together to produce a full MWPC.](image)

10.2.1.2 Materials

This redesign offered another opportunity to consider various plastics to meet our needs. We tested Delrin, Lucite, Acrylic, and Noryl for dimensional stability by subjecting rough-cut stock bars to
annealing cycles and then skimming down to the final shape of the MWPC arms. We concluded that Noryl was the best material to work with, and it also had sufficiently low radioactivity [238]. As a consequence of the process of testing materials, we included one annealing cycle in the fabrication process, performed after a rough cut of Noryl stock and before the final skim using a CNC mill. This enabled stresses in the plastic to be released before the final part was obtained. In parallel, we also switched the anode and cathode wire material to stainless steel, to ensure that any challenges with the material get worked out before construction of the full BetaCage.

10.2.1.3 Wire suspension

With monolithic Noryl MWPC arms, wire positioning would no longer rely on the flatness of a PCB mounting surface. We moved to a precision feedthrough system inspired by the one used in the BaBar drift chamber [247, 248, 249, 250]. The Noryl bars have countersunk feedthrough holes centered with 10-µm precision, into which feedthrough assemblies are inserted. For a given wire plane, the feedthrough assemblies are slightly different on the two ends of the wire.

On one end, the assembly consists of a hollow, 2.5”-long, 1/16”-outer-diameter, 0.03”-inner-diameter copper tube press-fit into a precision-cut brass piece that is itself press-fit into an MWPC countersunk hole. The brass piece has a 100-µm hole (175-µm for cathode wire) centered with 10-µm precision that guides and centers the wire. A wire inserted through the centering hole and out the other end can then be crimped in the copper tube, establishing a mechanical and electrical connection with the copper tube. We call this the “rigid” feedthrough. A picture is shown in Figure 10.9.

Figure 10.9: BetaCage “rigid” feedthrough.
On the other end, the wire is centered by a brass piece very similar to the one for the rigid feedthrough. This wire-centering piece has the same dimensions and centering hole, and is press-fit into an MWPC countersunk hole. The wire is crimped into a copper tube which has the same radial specifications as the one for the rigid feedthrough, but is 1.5” long. Here, the copper tube is not press-fit into the wire-centering piece. Instead, it is press-fit into a hollow brass plunger, which rests inside the wire-centering piece and can move in and out of it constrained by a spring, as seen in Figure 10.10. We call this assembly the “spring-loaded” feedthrough.

Figure 10.10: Top: Spring-loaded feedthrough assembly. Bottom: Compressed and uncompressed feedthrough assemblies.

A wire is suspended by threading its ends through the wire-centering pieces of the feedthroughs on both sides of the MWPC, then first crimping the spring-loaded feedthrough. The wire on the other end becomes free to be tugged on, but offers some resistance because of the spring. The free end is then tensioned by a weight and the rigid feedthrough assembly is crimped to complete wire suspension. Springs were chosen to provide the correct tension for anode and cathode wires with roughly 2–3 mm of compression. The advantage of a spring-loaded feedthrough assembly

\[\text{\textsuperscript{1}}\text{Century Spring Models 10068 (anode) and 70094S (cathode)}\]
is the smaller spring constant than that provided by the rigid MWPC frame. Thus small relative
movements between the MWPC arms because of residual stress release over time would not cause an
appreciable change in wire tension. We suspect that the system might be subject to increased noise
because of microphonic jitter from the springs. If this turns out to be the case, the spring-loaded
feedthrough assembly can be trivially replaced with rigid-feedthrough assembly.

Figure 10.11 shows a CAD rendered top-view of a corner of the MWPC, with the plastic set to
transparent, exposing the two types of feedthroughs described above.

![CAD rendered top-view of a corner of a four-part MWPC, with the plastic set to transparent, exposing the two types of feedthroughs described above.]

Figure 10.11: CAD rendered top-view of a corner of a four-part MWPC, with the plastic set to transparent, exposing the feedthroughs. One type of feedthrough just holds a wire in place by pinching it in a crimp in the copper tube and centers it with a precision drilled hole in its brass head. The other does this and provides tension by means of a spring.

10.2.1.4 PCBs

A PCB is mounted outside the MWPC frame next to the copper tubes, to gang signal from wires
into channels and to make electrical connections with the outside world. Fly wires with one bare end
are crimped into the open ends of the feedthrough copper tubes. The other end of the fly wires have
male quick-disconnect terminals that plug into receptacles on the PCB. Gangs of ten receptacles on
the board will be connected to an SHV cable that will interface with an SHV vacuum feedthrough
on the pressure vessel. Pictures of the PCB with and without a sample quick-disconnect connector
are shown in Figure 10.12. A picture of an MWPC with its wire feedthroughs connected to the PCB
via quick-disconnects is shown in Figure 10.13.
Figure 10.12: Left: Redesigned PCB mounts above or below a row of feedthrough copper tubes. Each tube will have a wire with a quick-disconnect connector crimped into its free end. These can then be plugged into the PCB. Right: PCB with a quick-disconnect connector plugged in.

Figure 10.13: A test MWPC (described in Section 10.4) with its wire plane connected to a PCB via quick-disconnect terminals.
10.2.2 Assembly and construction

The Protocage MWPCs will be assembled on an assembly jig custom designed by Bob Nelson and me. The assembly process itself is conducted in a small class-2000 clean room.

10.2.2.1 Assembly bench

The assembly setup is pictured in Figure 10.14. We used a Thorlab 4’×2’ optical breadboard with a movable stand as a working platform for assembly. The size was picked to permit comfortable access to the 0.25-m² footprint MWPC from all sides while two seated technicians worked on it. The MWPC frame was positioned at a height of 6” on the bench using stainless-steel rods. The feedthrough holes for the wire-plane being assembled were aligned with the long dimension of the breadboard. A few inches from one side of the frame, we placed a track with a movable cart on the bench, parallel to the side of the frame, and set it up at the height of the feedthrough holes in the frame. The movable cart housed a custom clamp for holding a spring-loaded feedthrough assembly as well as a cup to hold the crimper tool. A precision stage on the clamp enabled control of the position of the feedthrough assembly relative to the crimper jaws (see Figure 10.15). We call this side of the bench, the “spring side.” On the opposite side of the frame, the “tensioning side” a similar track and movable cart were placed on the bench, but located closer to the frame to enable crimping of copper tubes in place in the frame. A custom clamp on the cart could be precisely positioned to secure a copper tube in the frame for the crimping process. A second precision stage
was used to position the crimper tool relative to the copper tube in the MWPC frame (see Figure 10.15). Beyond this track, and away from the frame, we mounted a bar with three rows of 80 1/64” precision-drilled holes on a RTNA-4080-V low-friction slide from The Precision Alliance (TPA). The bar was positioned so that at least one row of holes lined up with the MWPC holes being used for assembly. This bar had a tensioning weight attached to it over a pulley at a small distance (see Figure 10.15).

Figure 10.15: Top left: Spring-side crimper jig with cup for a crimper tool and a clamp to secure feedthrough tubes during crimping. *Courtesy:* Bob Nelson. Top right: Tensioning-side crimper jig, specially designed to have bare minimum clearance between the MWPC frame and the crimper, and an additional stage to set the height of the crimper tool relative to the feedthrough tube. Bottom left: Precision-drilled bar with 1/64” holes to thread and secure sense wires, attached to a low-friction side. The bar is tied to a weight hung off a pulley, to tension wires secured to the bar. *Courtesy:* Bob Nelson. Bottom right: Close-up of clamp used to secure feedthrough tubes.
10.2.2.2 Feedthrough Crimping

Based on recommendations from wire-chamber assembly veterans [251], we used a Simonds EZ-Squeeze pneumatic crimper to crimp the stainless-steel sense wires in the feedthrough copper tubes. We bought blank soft jaws (Simonds model SJ-1A) and had them custom-machined to provide the desired crimp profile on our feedthrough copper tubes. The crimper and the custom-machined jaws are shown in Figure 10.16.

![Simonds EZ-Squeeze Pneumatic Crimper](image1)
![Close-up picture of custom-machined jaws](image2)

To ensure a secure and lasting electrical and mechanical connection to sense wires, we found optimal copper softness, crimping pressure, and crimper jaw shape through trial and error. The final working configuration involved annealing commercially available $\frac{1}{16}$-inch-diameter copper tubes for 1 hour at 800°F and cooling down back to room temperature over 10 hours. The crimper jaws had a 2 mm $\times$ 3 mm rectangular footprint with maximum closing distance of 0.6 mm at 40 psi. In this configuration, the copper is soft enough to deform around the 25-µm-diameter stainless-steel wires and prevent it from slipping. Unfortunately, it becomes too soft for the crimper to act on it without any support. In the absence of the clamps described in Section 10.2.2.1 and shown in the bottom right pane of Figure 10.15, the copper tube slips by an unpredictable amount (~ a few mm) between the jaws when it is compressed, leading to variation in crimp reproducibility. In fact, the clamps and precision stages were added to mitigate this problem. The 125-µm-diameter wires can be securely crimped without annealing the copper tubes.
10.2.2.3 Assembly procedure

A step-by-step assembly procedure for MWPC wire planes, along with pictures, is presented in Appendix C.

10.3 Other ProtoCage Components

10.3.1 Field cage

The prototype field cage will provide the same drift field (50 V/cm) as the final BetaCage, but will be only 20-cm tall, sufficient to contain the 156-keV beta endpoint for $^{14}$C in P-10 [227]. Rectangular copper rings for the field cage were fabricated from 1-mm-thick OHFC using water-jet cutting. The spacers between them were first fabricated as monolithic parts out of single sheets of stock Delrin. These potato-chipped over time, just like the layered MWPC frames and thus became useless. Therefore, I redesigned them to be similar in concept to the four-part MWPCs where four “arms” would together constitute a spacer between two copper rings. The redesigned spacers currently await fabrication at the University of Alberta Physics Machine Shop. A CAD rendering of the field cage is shown in Figure 10.17.

![Figure 10.17: CAD rendering of the 20-cm-tall prototype field cage with a 0.25-m² footprint to match the ProtoCage MWPC.](image)

10.3.2 DAQ, electronics and HV

The Protocage will use the same electronics system as the full BetaCage, described in Section 9.8. The digitizer boards for this readout system are currently under construction by Techne Instruments,
and will be available by August 2011. For tests in the meanwhile, we plan to use a cheap, commercial 60-MSa/s 12-bit, 8-channel digitizer from National Instruments, called NI-PCI-5105. This digitizer card can be plugged directly into a PC and readily works with LabView. The card will be sufficient to test individual MWPCs and will enable verification of the proposed trigger settings for the trigger and bulk MWPCs. The latter will require the wires of a wire plane to be ganged into two channels – fiducial and outer. Channels will be amplified using Cremat-111-INST-HV charge pre-amplifiers. Dan Sotolongo has prepared a DAQ module in LabView with basic triggering and data recording features for the initial tests. All HV will be provided to the system by NIM-format Bertan 375P HV supplies.

10.3.3 Gas handling

All test MWPCs and eventually the Protocage will be housed in a 24"-tall, 30"-diameter aluminum bell jar, custom-manufactured by Precision Cryogenic Systems. The bell jar is actually a vacuum chamber, but will be used as a pressure vessel for the 1-atm P-10 environment for the Protocage. It is mounted on a custom frame with a winch to lower and raise the bell jar off its aluminum base. A vacuum o-ring provides a seal between the bell jar and its aluminum base plate. The bottom of the baseplate is equipped with eight NW40 flanges, which we have used for gas circulation lines as well as SHV electrical connections to the inside. We have also attached a roughing pump to the bell jar to enable clearing of gas impurities before starting circulation of P-10 through the system. Since P-10 gas is fairly cheap, we do not intend to recirculate it. It will be introduced into the belljar at a modest flow rate of $\sim O(1)$ lpm and vented through an exhaust system. Pictures of the ProtoCage bell jar are shown in Figure 10.18.

10.4 Test MWPC with Anode Plane

Using the techniques discussed so far, Bob Nelson and Dan Sotolongo assembled an MWPC with a single plane of anode wires. A picture of this assembled anode plane is shown in Figure 10.19. Two copper-clad G10 plates will be used as cathode planes and mounted on the top and bottom of the MWPC. The use of cathode plates instead of wire planes saved effort in wire stringing, while enabling us to focus efforts on full system testing and testing stringing techniques for the delicate anode wires. A picture of the MWPC with just one cathode sheet in place is shown in Figure 10.20.

At the time of this writing, the ProtoCage system with the electronics chain, DAQ, gas handling, and the test anode plane was being prepared for a run. A $^{109}$Cd source will be mounted in the bell jar, next to the test MWPC to provide low-energy electrons for testing and calibration. The lack of a drift field cage will prevent all electrons from being stopped within the MWPC, but we might
Figure 10.18: Top: Picture of ProtoCage aluminum bell jar on its stand. Bottom: Picture of the bottom of the bell jar’s baseplate showing NW40 ports for SHV and gas connections.

Figure 10.19: ProtoCage MWPC with single anode wire plane and no cathodes. This picture was taken with a high-resolution SLR camera, manually focused on the wire plane to make the 25-micron thick wires visible. *Courtesy: Mayank Bakshi.*
Figure 10.20: Same MWPC as Figure 10.19 but with the wire feedthroughs connected to the PCB and with a copper-clad G10 plate used as a bottom cathode. A similar cathode plate will be attached to the top.

have sufficient statistics to reconstruct some features of the spectrum. The lessons from this test will either validate our plans for the ProtoCage, or provide feedback on elements of the system that require changes or redesign.

10.5 Future plans

Based on the outcome of the first live tests of the aforementioned MWPC, the full ProtoCage will be constructed and run. Its energy and pulse-shape response will be calibrated using radioactive sources, using the principles described in Chapter 9. If the ProtoCage functions successfully, it might be deployed at the Low-Background Counting Facility at Soudan Underground Laboratory, to perform screening of materials at moderate sensitivity, while the full BetaCage is constructed. Capital purchases for BetaCage construction have already started, with the hope of building and deploying it within a year from now.
Appendix A

Phonon-Pulse Shape Correction for Runs 125–128

Chapter 4 described the principles used to remove the effects of phonon-pulse-shape variation in energy and timing estimators across the detector. Here, I outline how this was accomplished in practice for Runs 125–128. The correction was performed in two stages, after basic data quality cuts and event reconstruction cuts had been applied. In the first stage, a sample of photons was selected for every operational detector in each Run to generate an interim lookup table of pulse shape estimators. These were used to eliminate outliers in timing, yield and manifold distance, in order to produce a cleaner photon sample. This sample was used to produce final lookup tables. The values of timing and energy estimators for all calibration and WIMP-search data series were corrected using the final lookup tables.

The code and scripts used to generate photon samples, lookup tables and the corrected RRQs are archived in the CDMS CVS repository under CorrTools v8. Also, it is important to note that corrections for Run 128 were performed using lookup tables that combined data from Runs 127 and 128. This improved statistics and correction quality.

A.1 First Pass Correction

The purpose of the first pass correction was to generate a clean sample of photons for characterizing the phonon response of the detectors. Without this process, position-dependent systematics in the phonon energy scale and ionization yield would lead to contamination of the photon sample with surface events. For Runs 125–128, the first pass correction also helped to remove phonon-manifold outliers and served as a quick check of changes in correction techniques and code before processing all data series.
A.1.1 Selecting photon samples

First pass photon samples were generated for each channel of a detector in a run, and for all runs by selecting all $^{133}$Ba-calibration events that passed the following criteria:

1. Data Quality cuts described in Section 6.3.
2. Event reconstruction cuts described in Section 6.4.
3. Uncorrected photon-equivalent recoil energy ($p_{rg}$) between 10 keV and 400 keV.
4. Inner charge electrode energy ($q_i$) greater than 3 keV.
5. Photon-equivalent ionization yield ($y_g$) within $\pm 2\sigma$ of the mean value as a function of $p_{rg}$.
6. Basic manifold cut defined by: $-0.7 < x_{part} < 0.7$; $-0.7 < y_{part} < 0.7$; and $rdel < 50$.
7. The primary channel defined according to phonon pulse amplitude and delay were the same.
   (See Section 4.4.2.)

As described in Section 4.4.1 a charge-fiducial-volume cut was not applied on any photon samples. This ensured better sampling of photons in the high-radius regions of the detectors. However, the samples included information about whether events passed the cut or not.

The photons passing the criteria listed above were randomly downsampled, channel-wise, to ensure that lookup tables were not dominated by events of a particular channel. A variation of up to 10% was allowed. Also, their energy spectra were flattened by bin-wise downsampling to ensure that events of a particular energy did not dominate the lookup tables. In particular, the number of events per 1-keV energy bin over the full $p_{rg}$ range was made to be consistent within 10% with the mean of event counts in 1-keV bins over 10–70 keV. On average, the first pass samples contained $\sim 100,000$ events per channel per detector per run.

A.1.2 Lookup table generation

First pass lookup tables were generated using the procedure described in Section 4.3.4, with the distance metric in Equation 4.9, reproduced here:

$$d = \sqrt{\Delta x_{part}^2 + \Delta y_{part}^2 + \frac{\Delta x_{del}^2}{L_{del}} + \frac{\Delta y_{del}^2}{L_{del}} + \frac{\Delta p_{rg}}{L_E}}$$

(A.1)

$L_{del}$ and $L_E$ were chosen to be 68 and 450, respectively, based on optimization studies described in Section 4.3.5. Also, the number of photons for nearest-neighbor bin selection, $N$, was set to 64 per 100,000 photons in a sample. The phonon-pulse shape parameters characterized in the first pass were $y_g$, $y_{gi}$, $p_{minrt}$, and $p_{del}$.

The lookup table was generated by code that performed the following procedure for each photon in the sample:
1. The distance of the photon to every other photon was computed using the metric described above.

2. Then the $N$ nearest photons, and, separately, the $N$ nearest photons passing the charge-fiducial-volume cut were picked to define the nearest-neighbor bins for that photon. The distinction was made so that estimators like $y_{gi}$ that were defined using the inner-charge-electrode energy were characterized using photons that passed the charge-fiducial-volume cut.

3. For the two sets of nearest-neighbor bins of a photon, the mean and standard deviations of pulse-shape parameters were computed with iterative outlier rejection. Photons outside three standard deviations of spread in the parameter were iteratively clipped till the mean and standard deviation were both stable to within 0.1% of the standard deviation for three consecutive iterations. This was done separately for each phonon-pulse-shape parameter.

4. For the nearest-neighbor bin without a charge fiducial volume cut, the covariance matrix for parameters $x_{del}$, $y_{del}$, $x_{ppart}$, $y_{ppart}$, $p_{minrt}$, and $p_{del}$ was computed and recorded with iterative outlier rejection. Photons with chi-squared values larger than the 99th percentile of a standard six-dimensional chi-squared distribution were iteratively clipped till the matrix was stable to within 0.1% for two consecutive iterations or if 10% of the data in that bin had already been discarded.

In addition, sample-wide averages for the phonon-pulse shape parameters were recorded to be used later for normalization.

### A.1.3 Phonon-pulse shape correction

For the first pass, I corrected pulse-shape parameters only for the photons in the samples. For each photon the corrected parameter was simply assigned using Equation 1.7, reproduced below:

$$par_{corr} = \frac{par}{\langle par \rangle_{bin}} \langle par \rangle_{global}$$  \hfill (A.2)

where $\langle par \rangle_{bin}$ is the bin mean for the parameter such as $p_{minrt}$ or $p_{del}$, as recorded in the lookup table, and $\langle par \rangle_{global}$ is a normalization factor, obtained by combining the sample averages recorded over channels and runs. For ionization yield ($y_{g}$ and $y_{gi}$), the normalization factor was set to one since all the events in the samples were photons. The corrected $y_{g}$ combined with the recorded ionization energy ($q_{sum}$) provided corrected $pr_{g}$.

The bin-wise covariance matrices were used to compute three chi-squared distances for each of the photons to their nearest-neighbor bins — one using the entire covariance matrix, one using only the metric position parameters $x_{del}$, $y_{del}$, $x_{ppart}$ and $y_{ppart}$, and one using only the timing parameters $p_{minrt}$ and $p_{del}$.
A.2 Final Correction

Next, we used the results of the first pass correction to refine photon samples. Also the lookup tables generation from these refined photon samples was optimized for separately for different sets of pulse-shape parameters to be corrected. Finally, the correction was carried out for all data series in Runs 125–128 using PipeCleaner.

A.2.1 Refining the photon sample selection

With corrected recoil energy and yield from the first pass correction, a new ±2σ yg band was defined. Photons from the first pass sample that fell within the new yield band were selected for second pass photon samples. To reduce miscorrections with the final lookup table, two additional criteria were imposed that removed most events outside the phonon manifold:

1. A cut was placed on the “size” of each nearest-neighbor bin, defined as the sum of the variance of \( \frac{x_{del}}{L_{del}}, \frac{y_{del}}{L_{del}}, x_{ppart} \) and \( y_{ppart} \) as recorded in the first pass lookup table [127]. The cut position was determined uniquely for each photon sample by studying the empirical distribution of bin sizes and ensuring that cut efficiency was >99.99%. An energy- and radius-dependent threshold was set to prevent preferential removal of low-energy and low-radius events.

2. Cuts were placed on the two chi-squared distances of each photon to its nearest-neighbor bin computed from the metric position parameters and the timing parameters. Photons with either of their chi-squared values greater than the 99th percentile of a standard chi-squared distribution (with four and two degrees of freedom for the position- and timing-based chi-squared value, respectively,) were removed from the photon sample.

The efficacy of these cuts in removing phonon-manifold outliers is demonstrated in Figure A.1.

Finally, the events in photon samples were sorted in order of ascending prg, to enable indexing in lookup tables and thus speeding up the correction process using PipeCleaner.

A.2.2 Lookup table generation

Multiple lookup tables were generated for a second pass sample, each optimized for a subset of phonon-pulse-shape parameters. One table, spanning 10–400 keV, was used primarily to correct recoil energy. The other three tables had a smaller prg range, but encompassed the WIMP-search recoil energy window, and were used to separately characterize timing parameters computed using the three different timing estimation kernels — RTFTwalk, PipeFit and WedgeFit. The metric for all these tables was the same as that for the first pass tables (Equation A.1). The relevant parameters for these lookup tables are listed in Table A.1.
Figure A.1: Phonon position manifold (Partition radius vs. Delay radius) for the first pass photon samples (all channels combined) of T3Z2 in Run 125. A recoil energy cut of 50–70 keV has been imposed. Photons are marked with black points. Events removed by the nearest-neighbor bin variance cut are marked in red, those removed by the chi-squared cut are marked in blue, and those removed by both are marked in magenta. Almost all manifold outliers are removed from the photon samples.

<table>
<thead>
<tr>
<th>Table</th>
<th>Recoil Energy range</th>
<th>$N$</th>
<th>$L_{\text{det}}$</th>
<th>$L_E$</th>
<th>Parameters to correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy Table</td>
<td>10–400 keV</td>
<td>64</td>
<td>68</td>
<td>450</td>
<td>yg, ygi</td>
</tr>
<tr>
<td>RTFTWalk Table</td>
<td>10–150 keV</td>
<td>128</td>
<td>60</td>
<td>500</td>
<td>pminrt, pdel, etc.</td>
</tr>
<tr>
<td>PipeFit Table</td>
<td>10–150 keV</td>
<td>128</td>
<td>60</td>
<td>500</td>
<td>prtPF1040, pdelPF20, etc.</td>
</tr>
<tr>
<td>WedgeFit Table</td>
<td>20–150 keV</td>
<td>128</td>
<td>60</td>
<td>500</td>
<td>prtwF020, pdelWF0, etc.</td>
</tr>
</tbody>
</table>

Table A.1: Second pass lookup table generation parameters
The code used to generate the final lookup tables was the same as that used for the first pass, except that it inserted into tables an indexed vector to point the location of the photon with closest round-number $\text{prg}$, i.e., the locations of photons closest to $[10, 11, 12, \ldots, 150]$ keV were marked in the RTFTwalk table, for instance.

### A.2.3 Correction using PipeCleaner

We used PipeCleaner to correct phonon-pulse-shape parameters for all calibration and WIMP-search data series in Runs 125–128. For every event, the closest indexed photon in recoil energy was located in each of the four tables. Then, photons with $\text{prg}$ within a 60-keV window of that indexed photon in a table were used to find the closest photon to the event by metric distance. Restricting the search to this energy window provided a dramatic speed up in correction processing, without compromising the correction results. The nearest-neighbor bin for the closest photon was assigned as the nearest-neighbor bin for the event. The average parameter value in that nearest-neighbor bin was then used to generate the corrected value for the parameter of the event under consideration, using Equation A.2.

Final corrected parameters did not overwrite the original measured values. They were stored under new variable names, labeled with the suffix, ‘c,’ after their original variable names. For instance, corrected $p_{\text{minrt}}$ was called $p_{\text{minrtc}}$.

It’s worth noting here how recoil energy was reconstructed from photon-equivalent recoil energy, after the correction process was completed. PipeCleaner defined a corrected recoil energy, $p_{\text{rc}}$, from corrected photon-equivalent recoil energy, $p_{\text{rgc}}$, as follows:

$$p_{\text{rc}} = p_{\text{rgc}}(1 + \frac{V_b}{E_{\text{charge}}}) - q_{\text{sum}} \frac{V_b}{E_{\text{charge}}}$$  \hspace{1cm} (A.3)

Similarly, events with a charge-fiducial volume cut use the recoil energy estimator $p_{\text{ric}}$:

$$p_{\text{ric}} = p_{\text{rgc}}(1 + \frac{V_b}{E_{\text{charge}}}) - q_i \frac{V_b}{E_{\text{charge}}}$$  \hspace{1cm} (A.4)

In addition to the corrected pulse-shape parameters, several correction-quality metrics were computed and written out to facilitate studies and diagnostics of the pulse-shape correction on an event-by-event basis, if needed. They were recorded separately for each of the four tables:

1. The index number of the closest photon to the event and that of the closest photon within the charge fiducial volume.
2. The difference in $\text{prg}$ between the event and its bin mean.
3. The position metric distance of the event to its nearest-neighbor bin.

---

1 The number 60 keV was determined empirically from a comparison of test corrections
4. Three chi-squared distances between the event and its bin — a position-metric-based chi-
squared value, a timing-based chi-squared value, and one combining both position metric and
timing.

5. The fraction of events in the nearest-neighbor bin that passed the charge-fiducial-volume cut.

These corrected parameter values and correction-quality metrics were then packaged by PipeCleaner
and added to the dataset as RRQs available for further analysis.
Appendix B

Detector Development for SuperCDMS

As noted in Chapter 8, future Ge-based dark matter searches by the CDMS collaboration will rely on increased target mass to accumulate large WIMP-search exposures and on detector redesigns to improve surface-event rejection. Here, I provide a brief overview of the evolution of CDMS detector technology beyond the ZIPS of CDMS II, which I will refer to as oZIPS (ordinary ZIPS). The first incremental step in detector improvements was the development of the mercedes ZIPS (mZIPS), which was SuperCDMS Soudan’s baseline design until 2010. Matt Pyle, Bruno Serfass, and I tested and developed new analyses for this detector, which I describe below in some detail. Subsequently, Matt Pyle designed an improved detector with interleaved charge and phonon sensors on both faces that improved surface-event rejection by an order of magnitude. This detector design, interleaved ZIP (iZIP), presents a clear path to achieving the surface-event rejection required for SuperCDMS SNOLAB. I highlight its main features and point out references for details.

B.1 mZIP

The lessons of CDMS II ZIPS guided the first redesigns of CDMS ZIP technology, which led to the mercedes ZIP (mZIP).[1]

B.1.1 Design

The mZIP design consists of 2.5-cm-thick, 7.6-cm-diameter germanium detectors that employ modified ionization and phonon sensors compared to the the CDMS-II design. Because of the increased thickness, they are 2.5 times more massive than CDMS-II detectors, at 0.64 kg each. Figure B.1 shows pictures of both an oZIP and an mZIP.

[1] Nader Mirabolfathi, Dennis Seitz, Kyle Sundqvist and Jeff Yen contributed to test run setup.

[2] So named because of the resemblance of its phonon sensor pattern to the Mercedes car company logo.
The first step towards improving discrimination between surface and bulk events was to shorten the phonon risetime and delay for surface events. In the new phonon sensor design, we increased the surface-area coverage of the phonon-trapping fins for each TES from 19% to 37%, to allow more phonons to be collected rapidly before reflections off detector surfaces. However, instead of just increasing the area of the oZIP QET, which would have increased aluminum fin lengths, we employed a “stadium” design (shown in an inset in Figure B.1) that left unchanged the distance that quasi-particles would have to traverse before diffusing into a TES. This was done to prevent a decrease of the phonon energy resolution due to increased quasi-particle recombination or trapping in impurities over a longer fin length. Athermal phonon response is essentially independent of volume, so no loss in energy resolution was expected from the increased detector thickness.

The next major design change was in the layout of the four phonon sensors of the ZIP. oZIP data pointed to an obvious area of surface-event rejection issues — high-radius events. The timing for these events was occasionally miscorrected during phonon-pulse-shape correction because of unresolved position degeneracies between high- and intermediate-radius events. This was especially true for events at the edge of the fiducial charge electrode. In CDMS II, we tackled this issue by incorporating events from both sides of the charge fiducial boundary in the phonon-pulse-shape correction (see Section 4.4.1). For SuperCDMS, we proposed a way to resolve this issue on a detector-design level. The four phonon readout channels were rearranged from a quadrant configuration to a combination of an outer annulus and three inner triangular regions. The new outer phonon channel would supplement the outer charge electrode, providing independent tagging of high-radius events. Note that moving to this layout eliminated the ease of repeated patterning with a single 5 mm × 5 mm tile of QETs over the entire phonon face. Instead the full mask had to be designed meticulously.
by hand. However, errors from repeated repositioning of the mask tile during fabrication were no longer an issue.

**B.1.2 mZIP testing and characterization**

We fabricated a prototype detector with these design modifications at the Stanford Nanofabrication Facility and performed a detailed characterization of it at the UC Berkeley Test Facility. The Berkeley facility has a $^3$He-$^4$He dilution fridge outfitted with electronics and DAQ similar to Soudan, but without the elaborate shielding and low-background protection of Soudan. The prototype was installed in this fridge along with two 1-cm-thick ionization-only oZIPs to serve as endcaps for tagging of charge-side and phonon-side events on the prototype mZIP. The first few datasets showed that the mZIP’s ambient event rate was $\sim 100$ Hz in the absence of radioactive sources, because of its increased mass and its installation at a surface facility subject to higher particle fluxes. This caused too many pile-up events, and pulses could not be properly reconstructed. The event rate was reduced by an order of magnitude by installing a 5-cm-thick lead shield around the dilution fridge. Then, we acquired electron-recoil-calibration data using $^{133}$Ba and $^{60}$Co, and nuclear-recoil-calibration data using $^{252}$Cf.

We calibrated the charge and phonon (preliminary) energy in the same way as we did for oZIPs. We observed that the phonon energy resolution of the mZIP prototype was similar to that of CDMS II detectors, as shown in Figure B.1.2 confirming that no loss mechanisms were introduced by modifying the phonon collection fins. We also observed that the increased detector thickness did not result in degradation of ionization energy resolution, indicating that trapping of drifting charges remained negligible.

Next, we constructed 3D phonon manifolds for mZIP events using $r_{del}$, $r_{part}$, and a new variable called $ppart_{io}$, which measured the ratio of summed phonon energy deposited in the inner sensors to that of the outer annular sensor. The shrimp-shaped manifold slice of the oZIP was replaced by one without radial foldback, as seen in Figure B.3. Also, we noted that the timing information of events for the mZIP became monotonic with depth — events closer to the charge side had faster timing than those near the phonon side, as seen in the mZIP manifold slice in Figure B.3. This was counterintuitive, but a direct effect of increased detector thickness, coupled with differences in aluminum coverage on the phonon and charge faces. Phonon-side events were likely undergoing slower phonon downconversion than charge-side events. An undesired dependence of timing on radius was still observed on moving from low to high $ppart_{io}$, which made it impossible to use the native timing for full depth reconstruction. Therefore, we resorted to our phonon-pulse-shape correction algorithm, incorporated the new manifold variable $ppart_{io}$ into its distance metric, and obtained corrected timing and energy for events.

We then used the corrected timing to test the prototype’s discrimination power between nuclear...
Figure B.2: Phonon baseline resolution compared between T3Z3 from CDMS II (blue) and the prototype mZIP (green). Courtesy: Matt Pyle.

Figure B.3: Left: Scatter plot of the phonon delay ($r_{del}$) vs. phonon energy partition radius ($r_{part}$) for oZIPs. The scatter plot is colored by $p_{minrt}$ in a very thin azimuthal slice of the detector. Right: Scatter plot of phonon energy partition between inner and outer sensors ($ppart_{io}$) vs. phonon energy partition radius constructed from the inner sensors ($r_{part}$). The scatter plot is colored by $p_{minrt}$ in a very thin azimuthal slice of the detector.
recoils and surface-event background. The addition of the lead shield unfortunately introduced a higher ambient neutron rate due to cosmogenic muon interactions, contaminating our surface event sample with neutrons. In the absence of a way to distinguish these neutrons from surface events in the nuclear-recoil band, we were limited to evaluating discrimination power using a subset of surface events between the electron- and nuclear-recoil bands, at the cost of an unmeasured systemic error. With this caveat, a simple $p_{\text{minrt}} + p_{\text{del}}$ timing cut rejects a large fraction of the tagged surface events while maintaining good acceptance of nuclear recoils, as shown in the left pane of Figure B.4. To reject the remaining surface events, we invoked a new style of event discrimination developed during this analysis. The various parameters measuring phonon-pulse shape such as partition variables $x_{\text{part}}, y_{\text{part}}, p_{\text{part,io}}$, delay variables $x_{\text{del}}$ and $y_{\text{del}}$, and timing variables $p_{\text{minrt}}$ and $p_{\text{del}}$ can be used to create a multi-dimensional manifold. These parameters are not only functions of the energy and event position, but also of the type of recoil; the characteristics of this manifold are different for electron recoils and nuclear recoils. So, for each event we constructed two $\chi^2$ parameters measuring the compatibility with a nuclear recoil hypothesis (for a manifold of calibration neutrons) and with an electron recoil hypothesis (for a manifold of calibration photons). The difference of the two $\chi^2$ values on an event-by-event basis provided a new discriminator, which we used to reject the remaining surface events as shown in the right pane of Figure B.4. This manifold-$\chi^2$ parameter was also adapted for CDMS II ZIP analysis as explained in Section 4.4.3. Finally, we put a 90% CL upper limit on the combined timing and $\chi^2$ rejection of 2.3/733 ($\sim 0.3\%$) at a nuclear recoil efficiency of 77.5% for the recoil energy range 10–100 keV. As a reminder, the CDMS II rejection was $\sim 0.6\%$ for the same nuclear recoil efficiency. These results sufficed in designating the design for this prototype as the baseline design for SuperCDMS Soudan detectors.

A batch of mZIPs was manufactured and installed at Soudan for engineering runs (Runs 130 and 131) to evaluate the performance of the mZIPs without the systematic errors of a surface test facility. The performance of these detectors was comparable to the the surface test, but the $\chi^2$ discriminator suffered because of limitations on the size of the $^{252}\text{Cf}$ neutron sample that could be acquired without risk of excessive copper activation. Overall, while the new phonon-sensor layout greatly ameliorated misreconstruction of events due to phonon manifold degeneracy, it was evident that the design did not fully exploit the monotonic dependence of timing information with depth. This would be possible only with time-of-flight measurements using phonon sensors on both faces of the ZIP. A second set of phonon sensors on the charge face would also help eliminate the residual radial dependence of timing parameters. This had always been considered difficult since the phonon sensors, which are biased by low-impedance circuits, would shunt charge signal if put on the same face as the charge electrode, degrading charge resolution. The next wave of SuperCDMS detectors solved this problem.
Figure B.4: Left: Histogram of $p_{\text{minrt}} + p_{\text{del}}$ for nuclear-recoil events (green) and tagged surface events (blue). The suggested cut is indicated by a dashed line. A simple timing parameter cut will remove almost all surface events. The same plot for CDMS II is shown in Figure 3.13. Right: Distribution of the surface events (blue) and nuclear recoil events (green) in the $\chi^2_{\text{ER}} - \chi^2_{\text{NR}}$ and energy parameter space. The events that pass the timing cut are circled in red but are rejected by the cut below the dashed line.

B.2 iZIP

The current baseline design for SuperCDMS Soudan and SNOLAB is a new style of 2.5-cm-thick ZIP that places phonon and charge sensors on both faces and accomplishes unoptimized discrimination of surface events from nuclear recoils at residual rates of $\sim O(0.1\%)$ from phonons and charge independently [252]. The placement of both types of sensors on one face, despite bias circuits of different impedance, is accomplished by interleaving of electrodes.

The idea of interleaved electrodes was first proposed in [253] as a means to improve resolution for ionization detectors in which only one type of charge carrier is collected efficiently. It was co-opted by CDMS to provide two-sided measurement of charge and phonons so that an asymmetry in pulse amplitudes and timing between the two sides would help identify surface events. A first attempt at such a detector was made even before mZIPs were conceived, but the prototype had an unusually high $T_c$ on one side and an electrical short on a channel, which made it difficult to measure performance [254]. Inspired by CDMS’ tests of this idea, the EDELWEISS collaboration developed a similar detector and demonstrated excellent performance [218]. This prompted CDMS to revisit the idea of interleaved electrodes. Matt Pyle, aided by collaborators at Stanford and Berkeley, subsequently designed and tested a working detector based on this principle, called the iZIP [252, 94].

Several iZIP designs by [255, 256] Matt Pyle and Scott Hertel have been tested at UC Berkeley as well as an iteration at Soudan (Run 132). The phonon and charge-electrode layout of the current iZIP design is shown in Figure B.5 along with the phonon sensor layout on both faces. Each face has two charge electrodes (fiducial and outer) and four phonon channels (similar to mZIP). The
phonon rails serve as a ground plane to the biased charged rails. The charge rails on opposite faces are biased with voltages of opposite polarity — $+2 \text{ V}$ and $-2 \text{ V}$ for instance (See Figure B.6 for electric field configuration). This produces a net $2 \text{ V}$ drop across the crystal for symmetric collection of ionization for bulk events on both sides, and very high fields at the surfaces between the charge and phonon electrodes to enable asymmetric collection of surface-event charge on one side compared to the other. This provides excellent ionization-yield-based discrimination, up to two orders of magnitude greater than the yield discrimination of oZIPs. For the phonons, the difference of delays between the two faces as well as energy partitioning between them provide two handles on surface-event rejection, together adding an order of magnitude in discrimination over that of oZIPs. I refer readers to the upcoming thesis work of Matt Pyle [94] for a more thorough treatment of iZIP design and performance.

Figure B.5: Top left: Schematic of phonon sensor layout for current configuration of iZIP. Top right: Phonon and charge electrode layout for one iZIP side. The charge rails are seen in grey and are interleaved with the phonon rails in blue. Bottom: A zoom in of the phonon rails shows QETs in parallel. The aluminum fins are marked in blue and the W TESs in red. *Courtesy:* Scott Hertel.
Figure B.6: Electric-field distribution for alternating charge and phonon electrodes of a iZIP in one half of a cross section perpendicular to the flat faces. The electrodes on the top and bottom surfaces are marked by regions where field lines accumulate to points. Every alternate electrode is a phonon rail and acts as ground. Thus the charge from a surface event is collected by a phonon rail and a charge electrode on the same detector face, while that of bulk event is collected by charge electrodes on opposite faces. Courtesy: Matt Pyle.
Appendix C

BetaCage wire-plane assembly procedure

Here I review the complete procedure for stringing a wire plane for a BetaCage MWPC. This procedure was developed by Bob Nelson, Dan Sotolongo, and me. The pictures below are courtesy of Bob Nelson.

1. All MWPC components, including plastic pieces, fasteners, copper tubes, brass-pieces, and springs are first cleaned in an ultrasonic cleaner and then stored in a purge box.

2. The MWPC frame is assembled from its constituent pieces in the clean room and mounted on its stand on the assembly bench.
3. A wire spool is positioned at the tensioning side of the assembly bench and a non-linting sponge soaked in alcohol cleans the wire as it is pulled from the spool.

4. A long, but small-diameter rod with a clip at the end is passed through an MWPC hole on the spring side and all the way through the corresponding hole on tensioning side of the MWPC. The wire from the spool is attached to the rod and is pulled all the way back through the MWPC holes.
5. The wire is threaded through a spring-loaded feedthrough assembly using the 100-µm-diameter hole (175 µm for cathode wire) and with the assistance of suction on the copper tube.

6. The brass head of the feedthrough and the spring is inserted into the MWPC.
7. The spring-side crimping cart is aligned with the hole in which the wire is being strung and then the brass plunger of spring feedthrough assembly is secured in the crimping jig clamp.

8. The crimper tool is then used to crimp the wire in the tube.
9. The crimped spring-loaded feedthrough is inserted into the hole and the forward most brass piece is press-fit into the frame.

10. On the other end of the MWPC, the wire is clipped with sufficient length, and threaded through a rigid feedthrough assembly, again with suction on the copper tube to aid threading.
11. This feedthrough assembly is then press-fit into the MWPC. Now, pulling the wire on the far side of the MWPC compresses the spring.

12. The tension-side crimping jig is aligned with the correct MWPC hole and the feedthrough copper tube is clamped. The crimper jig is locked in place once the feedthrough is secured.
13. The free end of the wire is inserted into the corresponding hole on the tensioning bar and is attached to the bar by electrical tape, while the low-friction slide is in locked position.

14. After securing the wire, the slide is gently released. The tensioning weight pulls the bar and hence tensions the wire.
15. The crimper is lowered into the jig and the wire is crimped in its copper tube. Excess wire is then clipped.

16. This concludes the stringing of a single wire. Steps 3 through 10 are repeated till all wires are suspended.

For material stress considerations, we chose the sequence of suspending wires to start with the outermost wires and end with the inner ones. With two technicians working at once, the stringing can be parallelized to an extent, but still takes ~8 minutes per wire. They can each work on one side of the bench and conduct the tasks for that side, i.e., one tech can pull wires from the spool through the MWPC and crimp spring feedthrough assemblies and seat them in the MWPC frame, while the other person can thread the wire through the rigid feedthrough assemblies, seat them in the frame and tension and crimp wires.
Bibliography


[188] M. Tarka. Expected neutron background from \((\alpha-n)\) reactions and spontaneous fission in Soudan as determined by Monte Carlo simulations. CDMS internal note, University of Zürich, 2008. URL\url{http://www.physik.unizh.ch/groups/groupbaudis/cdms/restricted/internet_notes/080111/nbackgrounds.html}


