A Search for WIMP Dark Matter Using the First Five-Tower Run of the Cryogenic Dark Matter Search

by

Jeffrey Peter Filippini

A.B. (Harvard University) 2002
M.A. (University of California, Berkeley) 2004

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

in

Physics

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:
Professor Bernard Sadoulet, Chair
Professor Hitoshi Murayama
Professor Leo Blitz

Fall 2008

The dissertation of Jeffrey Peter Filippini is approved:

Chair Date

Date

Date

University of California, Berkeley

Fall 2008

A Search for WIMP Dark Matter Using the First Five-Tower Run of the Cryogenic Dark Matter Search

Jeffrey Peter Filippini

Abstract

A Search for WIMP Dark Matter Using the First Five-Tower Run of the Cryogenic Dark Matter Search

by

Jeffrey Peter Filippini

Doctor of Philosophy in Physics

University of California, Berkeley

Professor Bernard Sadoulet, Chair

In recent decades astronomers and physicists have accumulated a vast array of evidence that the bulk of the universe's matter is in some non-baryonic form that remains undetected by electromagnetic means. This "dark matter" resides in diffuse halos surrounding galaxies and other cosmic structures. Particle theorists have proposed a wide array of candidates for its nature. One particularly promising class of candidates are Weakly Interacting Massive Particles (WIMPs): quanta with masses of order 100 GeV/c^2 and interactions characteristic of the weak nuclear force.

The Cryogenic Dark Matter Search (CDMS) experiment seeks to directly detect the rare elastic interactions of galactic WIMPs with terrestrial nuclei. To this end, CDMS operates an array of crystalline Ge and Si particle detectors in Soudan Underground Laboratory in northern Minnesota. These crystals are operated at millikelvin temperatures and instrumented to measure the ionization and athermal phonons generated by each particle interaction. This combination provides a powerful two-fold discrimination against the interactions of particles generated by radioactive decay and cosmogenic showers.

This dissertation describes the commissioning, analysis, and results of the first WIMP-search data runs of the CDMS experiment with its full complement of 5 "Towers" of detectors. These data represent a substantial increase in target mass and exposure over previous CDMS results. The results of this work place the most stringent limits yet set upon the WIMP-nucleon spin-independent cross section for WIMP masses above ~ 44 GeV/c^2, as well as setting competitive limits on spin-dependent WIMP-nucleon interactions. This

work also outlines the larger context of this and other probes of the WIMP theory of dark matter, as well as some current development efforts toward a larger cryogenic experiment.

Professor Bernard Sadoulet
Dissertation Committee Chair

To my parents,

and

To Heather

Contents

List of Figures

List of Tables

Acknowledgments

This page is wholly inadequate to express my gratitude to all those who deserve it: my colleagues, family, teachers, and friends over the past several years. Listing everyone is hopeless so I won't even try, but I hope you all know that you have my deepest appreciation. Nonetheless, there are a few people I would like to thank specifically - my sincere apologies for any omissions.

Bernard Sadoulet, my dissertation advisor, has provided invaluable support and guidance throughout my years at Berkeley. He sets an inspiring example of both breadth and depth, identifying the key aspects of the task at hand while maintaining a close eye on the big picture beyond the laboratory walls. Bernard also sets the tone that has helped to make his group a good place to do science: wide-ranging interests and hard work, combined with a sense of humor and an appreciation of balance in life.

My time in the Berkeley CDMS group has been a pleasure and a privilege. My lab-mates have not only been excellent scientific colleagues, but have also become good friends both inside and outside of the lab. I'd like to give particular thanks to Dennis, both for all of the help gave me and for the honest discussions he was always ready to have. For the good, the trips to Strada, and sometimes just for shooting the breeze, many thanks to all those who made the Berkeley group so enjoyable: Elizabeth Arscott, Miguel Daal, Nathan Earnest, Caryl Esteves, Mary MacCready, Al Lu, Vuk Mandic, Dan'l Martinez, Patrizia Meunier, Nader Mirabolfathi, Keith Patterson, Arran Phipps, Dennis Seitz, Bruno Serfass, Danielle Speller, Kyle Sundqvist, Rachel Winheld, Jeff Yen, and Yung-Ruey Yen.

CDMS II is the work of countless people whose names do not appear on this dissertation, and I owe them all a debt of gratitude. My thanks to all those who shared their time with me and taught me so much, especially when that time was shared in the cold northland of Minnesota. My admiration and thanks to Dan Bauer and Jim Beaty for managing to hold the experiment together through countless trials and tribulations. Thanks also to Richard Schnee and Steve Yellin for their frequent insights into analysis and statistics, and for always keeping me honest (or at least trying to...). My special thanks also to Cathy Bailey, Jodi Cooley, Matt Pyle, and Zeesh Ahmed: honorary Berkeley group members one and all, and the source of countless good conversations.

My lasting gratitude to my parents: without your love, support, and encouragement (and your patience with my years of babbling about science) I would never have

gotten here. Finally, my warm thanks and love to my wife, Heather: without your support I wouldn't have made it, and with it the future looks bright.

Preface

The CDMS II experiment is not the work of any one person, but a collaborative effort by more than fifty scientists and engineers spanning more than a decade. Not all (or even most) of the work described in this dissertation is thus the author's own. For clarification, this preface summarizes the major threads of my own involvement with the CDMS collaboration.

Throughout the text I have tried to properly attribute figures and key insights generated by others. I generally avoid citations to internal notes that are not publicly available. I apologize in advance to my colleagues and to the reader for any errors or oversights.

Detector testing and characterization

Upon becoming a member of the Sadoulet group in January 2003, I joined the detector testing and characterization efforts at the Berkeley test facility. As part of the detector team at Berkeley I worked extensively with the Berkeley 75 μW dilution refrigerator, carrying out the detector and Tower checkouts described in Section 5.1 over several dozen cryogenic runs. I took part in the testing of essentially all of the 18 new ZIP detectors added to the Soudan installation for this data run, which now fill Towers 3, 4 and 5. I also took part in choosing the stacking order of the detectors in Towers 4 and 5, in accordance with their radon exposures and performance at the test facilities. I also shared in the analysis of data from an extensive characterization of detector G31 (later identified as T2Z5) with a ^{109}Cd source in 2002. This study, described in [1, 2], made major contributions to our understanding of the response of ZIP detectors to surface electron recoils, our primary WIMP-search background.

After the completion of the CDMS II detector complement, I shared in the analysis

of data from early prototype detectors for SuperCDMS. I was part of the characterization team for tests of the first interleaved ZIP detector and the first one-inch-thick Si ZIP detector at Berkeley, and I led the initial characterization work on the first one-inch-thick Ge ZIP. As part of the latter, I demonstrated the increased power of phonon energy distribution parameters (originally suggested by Vuk Mandic) for the identification of surface events in large-format detectors.

Soudan operations

I began taking shifts at Soudan Underground Laboratory just before the commencement of the first successful data run at Soudan (Run 118), and worked at the mine for several weeks during the 2-Tower data run (Run 119) in 2004. There I learned the ropes of the cryogenic and data acquisition systems, knowledge which proved very useful during later CDMS runs at Soudan.

Over numerous shifts at Soudan over the next few years I participated in the installation and commissioning of the 5-Tower data runs. I took part in a broad range of the activities described in Chapter 5, including fridge testing, the reassembly of the fridge tails, both installations of the cold hardware and stripline wiring, and the biasing of the SQUID and TES channels.

Data analysis

I took part in the analysis of the Run 119 data set, focusing on advanced techniques for rejection of surface events. I developed a discrimination technique based on χ^2 tests both without (in parallel with collaborators at UCSB) and with an energy-correction, as described in Appendix B. This energy-corrected surface event cut was the primary analysis for the silicon ZIPs [3].

I played a leadership role in the analysis of the Run 123 and 124 data described in Chapter 6. I took part in much of the organization and analysis development for this run, with particular emphasis on the following portions:

- Overall organization of the data quality cuts and definition of the WIMP-search signal region, as well as maintenance of the official list of "good" data series for this analysis.

- Identification of poorly-performing detector regions and development of a cut to reject them (Section 6.4.1).
- Characterization and exclusion of periods of elevated cryocooler noise (Section 6.4.4).
- Discovery of cross-detector pileup as a source of mis-reconstructed events and development of algorithms and cuts to defend against it (Section 6.4.5).
- Development (in collaboration with Cathy Bailey) of methods to characterize neutralization and cuts to remove poorly-neutralized Ba calibration data (Section 6.3.6).
- Development of a substantially-enhanced multiplicity cut (Section 6.5.2).
- Defined ionization yield bands for nuclear and electron recoils (Section 6.5.4), as well as the cut selecting the surface event sample used in timing cut characterization.
- Discovered the GPS timing discrepancy which afflicted the experiment in Fall 2006 (Section 6.3.7).
- Significant characterization work on timing cuts for this run.

At the conclusion of this analysis I developed the surface event background estimate discussed in Section 7.2 and Appendix C, unmasked the blind analysis as described in Section 7.3, and computed the experimental limits in Section 7.4. The main results of this dissertation will appear in Physical Review Letters in early 2009 [4].

Scientific context

I played a substantial role in the collaboration's connection with the broader scientific context of our work. I computed our collaboration's first limits on spin-dependent WIMP-nucleon interactions, and was the lead author on the corresponding publication [5]. I computed the spin-independent and spin-dependent experimental limits from the work described in this dissertation. I have also participated in dialogues with members of other experiments and (with M. Battaglia) investigated the complementary relationships between direct, indirect, and collider investigations of WIMP dark matter (Section 2.5).

Chapter 1

The Dark Matter Problem

1.1 Introduction

What is the universe made of? has been a central question of science since antiquity. In the late 19th century, Dmitri Mendeleev's periodic table established a pattern in the dizzying array of chemical elements known at the time, as well as predicting the existence and properties of new elements. The early 20th century saw the elucidation of the electronic structure of matter and the fundamental mechanisms behind chemistry. Ensuing decades revealed the structure of the nucleus, the workings of the stars, and the structure of galaxies. By the 1970's science could claim a solid understanding of atomic matter, from its basic constituents to some of the largest structures it forms in the universe.

This complacency has been upended in the last three decades, however, not by laboratory investigations but largely through observations of the most distant parts of our universe. Many independent lines of evidence now lead to the conclusion that the visible, well-studied universe of atoms and molecules is far outmassed by a "dark universe," whose contents remain mysterious. Figure 1.1 shows the contents of the cosmos as inferred from current observations: a universe dominated by "dark matter" and "dark energy," with the familiar protons, neutrons, and electrons of chemistry supplying less than 5% of its total energy density. Discovering the nature of these dark components is the most pressing question of modern cosmology.

This dissertation describes a search for the constituents of one of these mystifying pie pieces: dark matter. "Dark matter" is the name given to matter which clumps together under gravity and whose existence we infer solely from the effects of its gravity. Dark matter

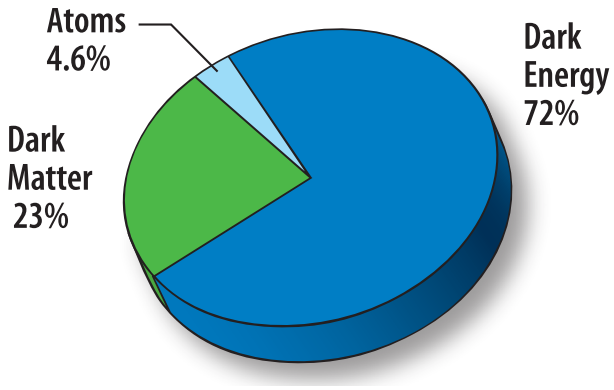


Figure 1.1: The present contents of the universe, as inferred from current cosmological measurements. Plot from the NASA/WMAP science team, 2008.

may be baryonic (*i.e.* built from the protons, neutrons, and electrons of ordinary atomic matter) or non-baryonic in nature, but the vast majority of it appears to be the latter. The details of this search will be described in later chapters. In this chapter I give a very brief introduction to the motivation for dark matter, attempting to answer a simple question: why has something invisible become an accepted part of modern astrophysics?

1.2 ΛCDM Cosmology

Before proceeding, it is useful to review the broad outline of our current picture of modern cosmology. A full discussion would take us very far afield of our main topic, so I will merely focus on introducing some of the terminology used in later sections. For more details on cosmological theory and the evidence that supports it, see any of several modern textbooks [6, 7, 8, 9, 10].

1.2.1 Basic framework

Despite the obvious variations in the world we see around us, current evidence suggests that the universe is homogeneous and isotropic on its very largest scales. This is perhaps the ultimate extension of the Copernican principle that we do not live in a special place in the cosmos. When averaged over a box gigaparsecs (Gpc) in size, one part of the universe looks much like any other.

Though the universe does not change (on average) as we move through space, it

does look very different from different vantage points in time. The best evidence for the universe's evolution is its expansion, described in the 1920s by Edwin Hubble and others. Inspired by this, we can look for the most general possible description of the universe's large-scale evolution. Applying the conditions of homogeneity and isotropy to the spacetime metric of general relativity gives the Friedmann-Robertson-Walker (FRW) metric,

$$ds^2 = c^2 dt^2 - R^2(t) \left(\frac{dr^2}{1 - kr^2} + r^2 (d\theta^2 + \sin^2 \theta d\phi^2) \right). \tag{1.1}$$

In this expression, t is cosmic time, (r, θ, ϕ) are dimensionless comoving spherical coordinates, k identifies the spatial curvature ($k = 0, -1, +1$ for flat, negatively-curved, or positively-curved), and $R(t)$ is a scale factor with dimensions of length which may evolve over time. The choice of time variable is in general ambiguous in general relativity, but the conditions of spatial homogeneity and isotropy choose a preferred time axis in this case.

We can see how this metric evolves with time by applying the Einstein field equations. In accordance with general relativity, the metric's evolution depends upon the universe's contents. Three classes of matter and energy are generally considered, each differentiated by its relative values of energy density ρ and relativistic pressure p:

1. **Matter**: non-relativistic material with negligible relativistic pressure ($p = 0$). The energy density of matter dilutes as $\sim 1/R^3$ as the universe expands.

2. **Radiation**: photons and relativistic matter (*e.g.* neutrinos) with positive relativistic pressure ($p = +\rho c^2/3$). Radiation's energy density dilutes yet faster than that of matter ($\sim 1/R^4$), due to relativistic effects on the radiation's momentum.

3. **Vacuum energy** (Λ): energy density associated with space itself, which has negative pressure ($p = -\rho c^2$) and does not dilute as the universe expands.

Defining a dimensionless scale factor $a \equiv R/R_0$, where R_0 is the scale factor at the present day, the Einstein equations become the Friedmann equations:

$$\left(\frac{\dot{a}}{a} \right)^2 = \frac{8\pi G}{3} \rho - \frac{kc^2}{a^2} \tag{1.2}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3} \left(\rho + \frac{3p}{c^2} \right). \tag{1.3}$$

Defining the Hubble parameter $H \equiv \frac{d \log R}{dt}$, we can see from Equation 1.2 that the universe will have zero curvature ($k = 0$) if the energy density ρ equals a critical density $\rho_c \equiv \frac{3H^2}{8\pi G}$.

Replacing densities and pressures with those of individual components and expressing all densities as fractions of the critical density: $\Omega_x \equiv \rho_x/\rho_c$, we get the following more useful forms of the Friedmann equations:

$$\Omega - 1 = \frac{kc^2}{a^2H^2} \tag{1.4}$$

$$\dot{H} = -H^2\left(1 + \frac{1}{2}\Omega_m + \Omega_r - \Omega_\Lambda\right). \tag{1.5}$$

These equations, in combination with the scaling laws listed above (*e.g.* $\Omega_m \propto a^{-3}$), may be solved for the evolution of the universe's scale factor ($a(t)$) given boundary conditions at the present day indicated by "0" subscripts: $a_0 = 1$, H_0, $\Omega_{0,m}$, $\Omega_{0,r}$, $\Omega_{0,\Lambda}$. These values constitute the basic parameters of our cosmological model.

A few general comments are in order about the effects of these various components:

- If $\Omega = 1$ today, then $k = 0$ and $\Omega = 1$ throughout the universe's history. If not, $\Omega \to 1$ as time advances.

- The contributions of the three components depends differently upon a, so the balance between them varies with time. Radiation dominates the universe's evolution at very early times, while vacuum energy (if it exists) dominates at very late times.

- When positive vacuum energy dominates the universe's total energy density, the scale factor increases exponentially with time. The scale factor's rate of change (the Hubble parameter) decreases with time when other components dominate.

Within this framework, it is often convenient to measure cosmological time in terms of the scale factor. This is useful because many attributes of the universe – *e.g.* its temperature and density – depend directly upon the scale factor, while varying the contents of our model of the universe will change the correspondence $a(t)$. The scale factor is most easily observed through its effects on electromagnetic radiation: light emitted with a wavelength λ_1 at a time when the scale factor was a_1 will be observed at a wavelength $\lambda_2 = \lambda_1\frac{a_2}{a_1}$ when the scale factor reaches a_2. This relation is often expressed in terms of the cosmological redshift, $z \equiv a - 1$, ranging from $z = 0$ at the present day back to $z \to \infty$ at the Big Bang. Cosmological redshift can be determined directly from the observed frequencies of features in the spectra of astrophysical objects. Such measurements are the primary method by which very distant objects are located in space and time.

1.2.2 Contents

The modern concordance model of cosmology is known as ΛCDM cosmology, for "Λ cold dark matter". Here Λ symbolizes the cosmological constant, the simplest form of vacuum energy or "dark energy." This model describes a universe with $\Omega \approx 1$, strongly dominated today by contributions $\Omega_{0,m}$ and $\Omega_{0,\Lambda}$. The term "cold" indicates that the bulk of the universe's matter was non-relativistic during the formation of large-scale structure.

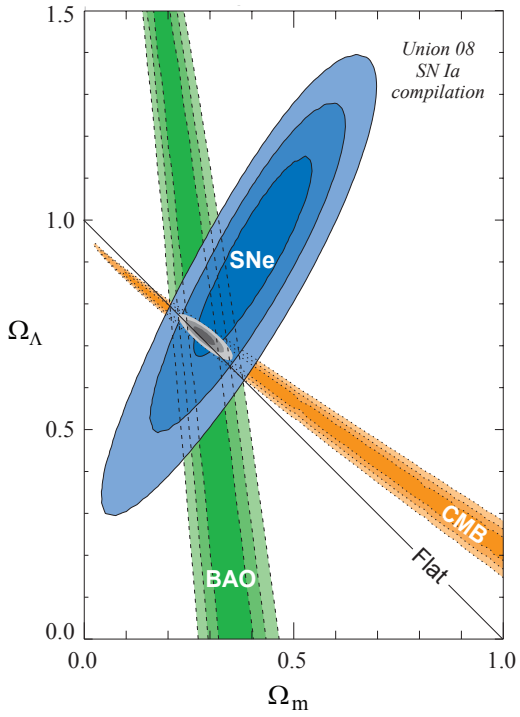


Figure 1.2: Current constraints on the normalized matter density ($\Omega_m \equiv \rho_m/\rho_c$, where ρ_m is the matter density and ρ_c is the critical density) and normalized dark energy density (Ω_Λ) of the universe. Colored regions represent 1σ, 2σ, and 3σ confidence regions from current measurements of Type Ia supernovae (SNe), the cosmic microwave background (CMB), and baryon acoustic oscillations (BAO). Figure from [11].

This framework is supported by a wide variety of measurements. Figure 1.2 shows a recent illustration of the agreement between three very different classes of measurement: studies of the cosmic microwave background, observations of baryon acoustic oscillations in the galaxy cluster distribution, and observations of distant Type Ia supernovae. Taken togethe, these observations describe a consistent set of cosmological parameters: $\Omega_{0,b} = 0.0456 \pm 0.0015$ (baryonic matter), $\Omega_{0,cdm} = 0.228 \pm 0.013$ (cold, non-baryonic dark matter),

and $\Omega_{0,\Lambda} = 0.726 \pm 0.015$ (dark energy) [12].

1.2.3 Thermal history

Within this model, cosmologists have developed a detailed understanding of the universe's thermal history, from the Big Bang to the present day. I list a few highlights below:

1. The universe's current expansion seems to have originated in the **Big Bang**, an event some 13.7 billion years ago. The entire visible universe is believed to have originated in a tiny ($a \approx 0$), hot region.

2. Some 10^{-34} seconds after the Big Bang, the universe is thought to have undergone a period of rapid expansion known as **inflation** [13]. The universe expanded rapidly by a factor of at least e^{50} during this period, smoothing out any initial variations and leaving the flat ($\Omega = 1$), homogeneous, and isotropic universe we see today. This expansion is believed to have come from a temporary dominance by the vacuum energy of some quantum field (the "inflaton"), though the detailed phenomenology of this field is not known. After the end of inflation, the decay of the inflaton reheated the universe to a hot state. Quantum fluctuations of the inflaton field gave rise to the small inhomogeneities that generated today's large-scale structure.

3. After ~ 1 second, the universe cooled enough for the **decoupling of neutrinos** from the primordial thermal bath. This occurs when the expansion rate of the universe rendered neutrino annihilation inefficient, freezing the comoving neutrino density at a value far above its thermal relic density. The modern cosmological neutrino density (still undetected) was set at this time, determined by the thermal density at the time when

4. Some $\sim$ 1 minute after the Big Bang, the universe became cool enough to allow the efficient fusion of free protons and neutrons into light nuclei, notably deuterium, helium, and lithium. The bulk of the universe's supplies of these elements was created at this time, in the process of **Big Bang nucleosynthesis**.

5. Almost $\sim 400,000$ years ($z \approx 1000$), the cosmos cooled enough to allow atomic hydrogen to form from free protons and electrons. This drastically reduced the universe's

optical depth, rendering it essentially transparent until the present day. The blackbody radiation which last scattered in this era, now redshifted into the microwave band, is visible today as the cosmic microwave background. The contribution of matter to the universe's energy density began to dominate over that of radiation around this same time.

6. Over the course of the past few billion years, the cosmic structures we see today formed from the **gravitational collapse** of slight over-densities in the universe's primordial energy distribution. Non-baryonic dark matter began its collapse early on, but radiation pressure prevented baryonic matter from joining the process significantly until after the decoupling of radiation from matter in the epoch of last scattering. Structure formation is hierarchical in the ΛCDM cosmology: small structures form early on, and larger structures form only later on.

7. About 4 billion year ago ($z \approx 0.4$), the contribution of vacuum energy to the universe's energy density began to exceed that of matter. The expansion of the universe began to accelerate, a process which continues to the present day.

1.3 Evidence for dark matter

The greatest strength of the dark matter hypothesis is that it is supported by a variety of disparate lines of evidence. These arguments fall into three broad categories:

1. **The modern universe**: The present distribution of matter – visible or invisible – can be mapped through its effects on the motions of visible objects and of light rays. Observations on galactic and cluster scales reveal that the universe's structures are much more massive than can be accounted for by visible matter alone.

2. **The primitive universe**: In addition to its signatures in the modern universe, dark matter's presence has also affected the universe's evolution. Observations of the early universe, both indirectly through light-element abundances and more directly through images of the cosmic microwave background, confirm the presence of dark matter and indicate that it is primarily non-baryonic in nature.

3. **The growth of structure**: The universe's composition affect the evolution of structure within it. The additional gravitation of dark matter is needed to explain the

rapid formation history of galaxies and larger structures.

At their roots, all of these lines of reasoning are based upon gravitational effects, but each probes very different regimes of scale and time through very different signatures. Though loopholes may be found in any one of these, thus far no theory has satisfactorily addressed them all without resorting to non-baryonic dark matter. I cannot do justice to the web of observations supporting the current cosmological model in this dissertation, but I sketch some of the more prominent supports of the dark matter hypothesis below.

1.3.1 Galactic dynamics

We can infer the total mass content of a region of space (or its excess above a uniform background density) by observing the motions of nearby visible objects. The first great success of this method came in 1846, with the discovery of the planet Neptune by astronomer Johann Galle. This first observation of Neptune was directed by calculations of Urbain Le Verrier, who had inferred its existence and position based upon observed perturbations in the orbit of Uranus. Related techniques are used today in searches for extrasolar planets [14], which are often found through perturbations in the motions of their parent stars. This same general principle can be applied on a cosmological scale: the motions of stars and galaxies carry information about the surrounding mass distributions.

Some of the clearest evidence for the existence of dark matter comes from the rotational dynamics of spiral galaxies. A spiral galaxy such as our own consists of a central bright bulge surrounded by a rotating disk of younger stars. Measurements of the galactic rotation speed as a function of radius have now been conducted for many spiral galaxies. Such measurements are based upon observed Doppler shifts in various spectral features [15], notably the visible hydrogen spectral lines, the rotational transitions of CO, and the 21cm HI line; the latter is particularly useful, as it is detectable at radii well outside the visible galactic disk. Only line-of-sight velocities can be measured directly, but true velocities can be estimated from knowledge of the disk tilt and proper motion.

If light traces the matter distribution, we would expect the orbital velocity of galactic material to decrease as $\sim r^{-1/2}$ at radii outside the visible disk. We see in Figure 1.3, however, that observed rotation curves are remarkably flat at large radii. From Newtonian mechanics, this behavior is consistent with a roughly spherically-symmetric mass distribution with a total enclosed mass proportional to radius ($M(\leq r) \propto r$). These data

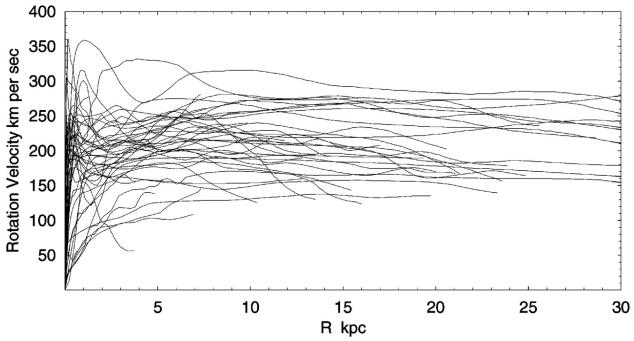


Figure 1.3: A superposition of rotation curves from an array of spiral galaxies, taken from [15]. These curves are measured with a variety of techniques at different radii: CO in the central region, optical in the disk, HI in the outer disk and halo.

suggest that the visible galaxies are embedded in substantially larger "dark matter halos," which extend well beyond the edge of the visible disk. Without the gravitational mass of this halo, the disk would fly apart.

Similar arguments may be made for the more amorphous structures of dwarf and elliptical galaxies. Rather than a rotation curve, we can infer the total mass M of these objects (or any gravitationally-relaxed system) from the velocity dispersion of the stars and gas composing them using the virial theorem:

$$2\langle T\rangle = -\langle V\rangle \quad \Rightarrow \quad \langle v^2\rangle \sim GM\Big\langle\frac{1}{r}\Big\rangle. \tag{1.6}$$

Here, too, dark matter is generally required to keep these structures gravitationally-bound, despite the high velocities of their constituents. Studies of the dwarf satellites surrounding our galaxy, such as that by Simon and Geha [16], place them among the most dark-matter-dominated objects in the known universe, with mass-to-light ratios in excess of several hundred the solar value. These high values are generally explained by inefficient star-formation in these tenuous small systems, which are easily stripped of their gas by feedback mechanisms based upon supernovae or stellar winds.

1.3.2 Galaxy clusters

Galaxy clusters, the largest bound structures in our universe, are also an excellent laboratory in which to observe the effects of unseen matter. Clusters are also a useful path to the cosmological value of $\Omega_{0,m}$: since they form from the gravitational collapse of

relatively large regions of the early universe, clusters may be more representative of the average composition of the universe than smaller structures.

Cluster dynamics

Dark matter at the cluster scale was first postulated by Fritz Zwicky in 1933 [17, 18]. Zwicky used the Doppler effect to measure the line-of-sight velocities of galaxies in the Coma cluster. Based upon the galaxies' velocity dispersion and an estimate of the cluster's size based upon its apparent angular size, net Doppler motion, and a (very inaccurate) value of Hubble's constant, Zwicky inferred a mass-to-light ratio of several hundred using the virial theorem. Zwicky's result was very crude and numerically-inaccurate, but modern calculations with better data and more sophisticated techniques support the same basic conclusion: galaxy clusters are strongly dominated by unseen matter.

X-ray emission

We can also map the gravitational potential well of clusters through emission from hot gas trapped within that well. Only $\sim 10\%$ of baryons within clusters are actually gathered into stars; the vast majority remain in within the intergalactic medium, captured in the cluster's gravitational well. In accordance with the virial theorem and considerations of hydrostatic equilibrium, this gas is elevated to $> 10^6$ K and emits in X-rays. Maps of this X-ray emission by instruments like the Chandra and XMM-Newton space telescopes indicate cluster masses consistent with those from galaxy motions and gravitational lensing.

According to numerical simulations, the bulk of the universe's baryonic matter is not contained within clusters. Instead, much of it is bound to the "cosmic web", the tenuous network of dark matter threads which link galaxy clusters to one another. This "warm-hot intergalactic medium" (WHIM) is expected to have very low densities and relatively low temperatures ($10^5 - 10^7$ K), making it difficult to detect in X-ray observations. In 2008, however, Werner and collaborators [19] announced an observation of such a gas trail between clusters Abell 222 and Abell 223 using the XMM-Newton X-ray telescope. Assuming that other clusters are linked by similar threads, this observation explains the gap between the inventory of observed baryons in the universe and that inferred from nucleosynthesis and the microwave background (see below). This further confirmation of the cosmic web also supports the predictions of ΛCDM models of structure formation.

Gravitational lensing

Cluster masses can also be inferred from their effect upon light rays from distant sources [20, 21]. In the weak-field, low-velocity limit of general relativity (applicable to essentially all astrophysical lens configurations), a region with Newtonian gravitational potential $\Phi(x)$ (with boundary conditions $\Phi \to 0$ as $x \to \infty$) has an effective refraction index of

$$n(x) = 1 + \frac{2}{c^2} \left| \Phi(x) \right| . \tag{1.7}$$

The gravitational potential well of a galaxy cluster or other massive object thus acts as a complex lens, distorting the images of more distant galaxies or quasars. The major advantage of this approach is that it provides a true map of mass overdensity, independent of assumptions of virial equilibrium. Mass estimates from lensing cannot be biased by incomplete gravitational collapse, unlike estimates based on galaxy motions or X-ray emissions.

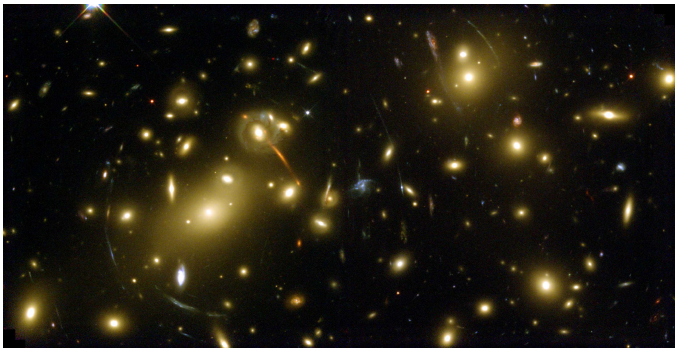
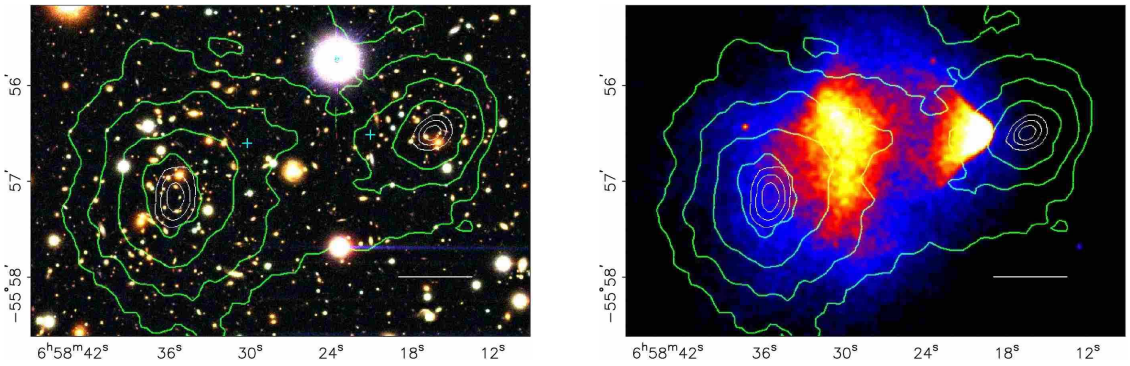


Figure 1.4: Hubble Space Telescope image of the cluster Abell 2218, showing prominent arcs from strong gravitational lensing. Image from the Space Telescope Science Institute.

Lensing studies occur in two major regimes. If the projected mass density $\Sigma(\theta, \phi) = \int \rho(\theta, \phi, z)\, dz$ of the lens is sufficiently large ("strong lensing"), distant sources may be multiply-imaged or distorted into obvious arcs or rings. Figure 1.4 shows Abell 2218, a rich galaxy cluster surrounded by prominent lensing arcs. Kneib and collaborators [22, 23], for instance, have used these lensing arcs to estimate a mass-to-light ratio of 80-180 in solar units, depending upon the portion of the cluster under consideration. Subtler lensing distortions are often difficult to identify, since the lensed galaxies themselves may be elongated in an unknown manner. Nonetheless, the mass distribution of such a system may be reconstructed using statistical correlations among the visible shapes of background sources ("weak lensing"). Johnston and coauthors [24], for example, have performed weak

lensing analyses of 130,000 galaxy clusters and groups identified in the Sloan Digital Sky Survey; they find general agreement with virial cluster masses and with the overall ΛCDM cosmology, suggesting a total matter density (with respect to the critical density) of $\Omega_{0,m} \approx 0.2 - 0.3$.

Colliding clusters



(a) (b)

Figure 1.5: Images of the Bullet cluster (1E0657-558) overlaid with weak lensing mass contours. (a) Optical photograph from the Hubble Space Telescope. (b) X-ray image from the Chandra X-Ray Observatory. Taken from [25].

One particularly dramatic illustration is given by the recent observations of the "Bullet cluster" 1E057-558, a pair of colliding clusters visible shortly after their first intersection. Through a multi-wavelength observational program, Clowe and coauthors [25] have reconstructed the distribution of hot gas and of total mass in this system. Figures 1.5(a) and 1.5(b) from their 2006 paper show the lensing mass contours overlaid on the optical and X-ray images of the structure. The stark segregation between total matter (traced by lensing) and the smaller proportion of baryonic matter (traced by the X-ray gas) is completely consistent with the dark matter model: the collisionless dark matter halos (as well as the galaxies themselves) have passed through one another essentially unchanged, while the two populations of hot gas have been shocked by interactions during the collision and remain between the two colliding structures. This segregation provides dramatic confirmation that the bulk of a cluster's mass is not merely gas, but is in some essentially non-collisional form. Bradac and coauthors characterized a second such system (MACS J0025.4-1222) in 2008 [26], confirming this interpretation.

1.3.3 Big bang nucleosynthesis

To study nature on its largest scales, we turn to observations related to the early universe. The first of these to make a big impact was the study of the abundances of light elements, notably deuterium, helium-3, helium-4, and lithium [27, 28]. These elements are synthesized from free protons and neutrons in the hot early universe, during the first few minutes after the Big Bang. Their primordial abundances depend critically upon the conditions during the period when such fusion was possible, and in particular on the baryon-to-photon ratio $\eta \equiv n_B/n_\gamma \times 10^{10}$. The phonon density sets the length of nucleosynthesis through its control of the universe's expansion rate; given equal photon densities, a greater baryon density leads to a faster rate of fusion to ^{4}He and fewer nucleons left over in ^{2}H and ^{3}He. Since the photon density is well-known from the microwave background's temperature, light element ratios thus provide an excellent "baryometer" with which to measure the abundance of ordinary matter.

Deuterium is the most powerful of these baryometers. Deuterium is a delicate nucleus, easily destroyed within stars and therefore no longer created in the modern universe. Since the deuterium abundance is sensitive to η and decreases with time, any measurement of deuterium yields a direct upper limit on η. Current deuterium measurements from quasar absorption lines [30] indicate $\eta \approx 5.5$ [29], implying a modern baryon density of $\Omega_B \approx 0.04$. Ordinary, baryonic matter thus *cannot* constitute more than a small fraction of the universe's total matter density. $\sim 80\%$ of dark matter must therefore be non-baryonic in nature.

Nucleosynthesis data for other light nuclides are generally more difficult to interpret. The primordial abundance of ^{3}He cannot be observed because it can be produced within stars, rendering it unsuitable as a baryometer. The lithium abundance is observable in low-metallicity stars, but has been problematic to match to nucleosynthesis models. The production and destruction of lithium in stars and cosmic ray collisions is not completely understood, contributing to the significant systematic errors in Figure 1.6.

1.3.4 The cosmic microwave background

Further support for the standard cosmology, including the non-baryonic nature of dark matter, comes from the cosmic microwave background (CMB). The CMB is the universe's primordial blackbody radiation, last scattered when the universe became transparent

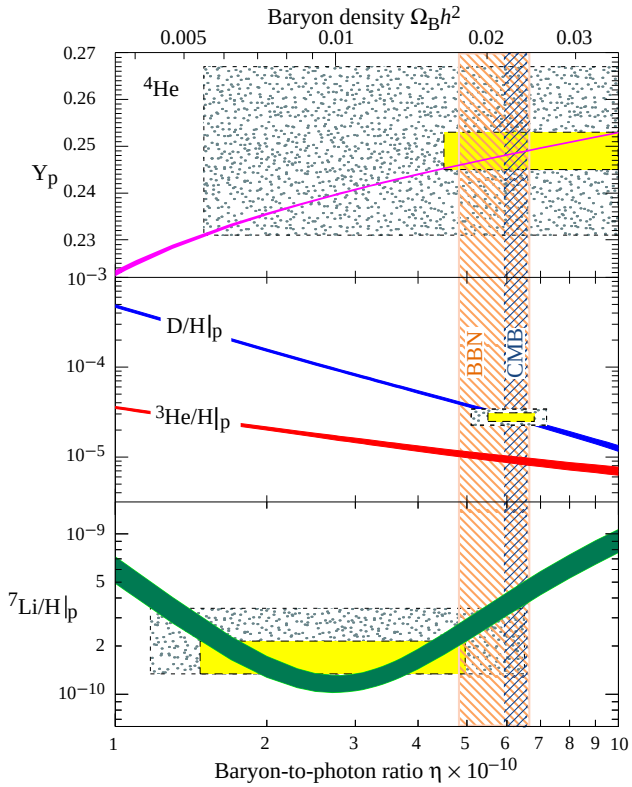


Figure 1.6: Summary figure of primordial light-element abundances, taken from [29]. Bands show 95% confidence level predictions from the standard cosmological model of the primordial abundances of ^{4}He, D, ^{3}He, and ^{7}Li as functions of the baryon-to-photon ratio during nucleosynthesis. All abundances are number ratios to hydrogen except ^{4}He, which is a ratio by mass (Y_P). Note the use of a linear scale for Y_P, rather than the logarithmic scale used for the other abundances. Small boxes indicate 2σ observational ranges based upon statistical errors, while large boxes incorporate systematic errors. Vertical bands indicate 95% confidence level ranges from the CMB and combined nucleosynthesis results.

approximately 400,000 years after the Big Bang. Superimposed on its blackbody spectrum at 2.73 K, the CMB has a anisotropies in temperature [31] and polarization [32] at the $\sim 10\mu$K and $\sim 0.1\mu$K levels, respectively. Figures 1.7 and 1.8 shows the power spectrum of CMB fluctuations in temperature and E-mode (curl-free) polarization, as measured by the Wilkinson Microwave Anisotropy Probe (WMAP) and other recent experiments [33, 12].

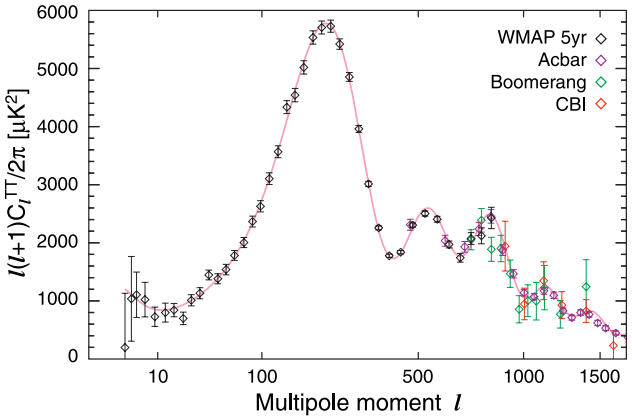


Figure 1.7: Temperature (TT) power spectrum of primary CMB anisotropies, as estimated from the WMAP 5-year data and an assortment of observations at small angular scales. The curve indicates the prediction of the best-fit ΛCDM cosmology. Figure taken from [33].

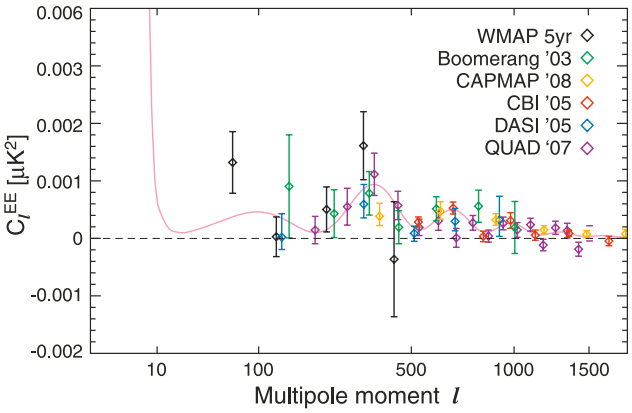


Figure 1.8: E-mode polarization (EE) power spectrum of primary CMB anisotropies, as estimated from the WMAP 5-year data and an assortment of other observations. The curve indicates the prediction of the best-fit ΛCDM cosmology. Figure taken from [33].

The microwave background's temperature fluctuations record inhomogeneities in the photon-baryon fluid at the era of last-scattering. These inhomogeneities can be seen as incoherent acoustic waves in the photon-baryon fluid of the last-scattering surface, and the

densities of baryonic and non-baryonic matter have strong effects upon these oscillations: baryons increase the inertia of the oscillating photon-baryon fluid, while dark matter reduces the driving effect of those oscillations upon the gravitational potential. The angular power spectrum of the anisotropies is thus a sensitive probe of cosmological parameters, encoded in the positions and heights of the various spectral peaks. When combined with other cosmological measurements (to break degeneracies), current CMB data provides exceedingly precise confirmation of the modern cosmological model.

1.3.5 Structure formation

A final argument for the existence of non-baryonic dark matter comes from attempts to link our observations of the primordial universe with those of its modern configuration. We know that the cosmic microwave background shows anisotropies at the 10^{-5} level, as described in Section 1.3.4. Since photons and baryons were tightly coupled at that era, this amounts to a snapshot of the level of baryon inhomogeneity when the universe was only a few hundred thousand years old ($z \approx 1000$). The scale factor a of the universe has grown by a factor of ~ 1000 since then, and today we see a wealth of galaxies and galaxy clusters. Bright galaxies are visible back to redshifts of at least $z \approx 7.6$ [34], indicating that large structures formed in the first few hundred million years after the Big Bang. Given these snapshots of baryonic structure, we can ask: has there been sufficient time for gravitational collapse to bring this change about?

In the standard theory of structure formation, the structures we see in the universe today originate from quantum fluctuations in the local energy density of the inflaton field. Regions with slightly higher energy densities than their surroundings collapsed under their own gravity, eventually forming the seeds of galaxies and galaxy clusters. These over- and under-densities are characterized by the density contrast field

$$\delta(\vec{x}, t) \equiv \frac{\rho(\vec{x}, t) - \bar{\rho}(t)}{\bar{\rho}(t)}, \tag{1.8}$$

where $\bar{\rho}(t)$ is the universe's mean energy density at time t. In practice, it is most convenient to consider the statistical distribution of the density contrast at different length scales, given by the power spectrum of δ:

$$\delta(\vec{k}, t) \equiv \frac{1}{(2\pi)^{3/2}} \int \delta(\vec{x}, t) e^{-i\vec{k}\cdot\vec{x}} \, d^3\vec{x} \tag{1.9}$$

For small density contrasts ($\delta \ll 1$) we can solve for the evolution of $\delta(\vec{k}, t)$ using linear perturbation theory. The various values of $\vec{k}$ decouple from one another in this regime, which lasts until $\delta \approx 1$. Propagation to later epochs requires numerical simulations. The full relativistic theory of gravitational collapse in an expanding universe is beyond the scope of this work; for a full discussion, see [6, 13, 8]. Here I give only some of the main results:

1. Radiation pressure prevents perturbations in the photon-baryon fluid from growing significantly after their entry into the causal horizon (when $ckt = 1$), if they enter during the radiation-dominated era. The photon-baryon fluid undergoes acoustic oscillations during this period, which lasts until the baryons decouple from the photon field at the epoch of last scattering ($z \sim 1000$).

2. The density contrast in pressureless matter, *e.g.* dark matter, grow only logarithmically during radiation domination.

3. The dark matter density contrast begins to grow as $\delta \sim a$ once the universe becomes matter-dominated. The photon-baryon fluid continues to oscillate until the baryons decouple, at which point it too can grow in proportion to a.

4. After decoupling, the baryons quickly (*i.e.*within a few expansion times) fall into the existing potential wells of the dark matter perturbations.

This last step is a crucial part of the argument for dark matter. The observation of nonlinear structure at the present day demands $\delta \gtrsim 10^{-3}$ at $z \sim 1000$, far greater than that observed. The small anisotropies observed in the microwave background at small scales ($\ell \sim 1000$ corresponds to the scale of observed galaxy clusters) have not had time to grow into the structures we see in the modern universe. Non-baryonic cold dark matter solves this problem very neatly by decoupling the microwave background from the total matter density. The smooth baryon distribution seen in the small-scale CMB anisotropies rapidly coalesces onto the larger dark matter over-densities, which had more time in which to grow. Accurate solutions to the linear perturbation problem in a universe with baryonic matter, non-baryonic matter, and radiation demand numerical integration, but the basic story remains.

1.4 Candidates for dark matter

Our current model of dark matter encompasses four main facts about its nature:

1. Dark matter is predominantly **non-baryonic** in nature. This is supported by evidence from nucleosynthesis and the microwave background, as described above.

2. Non-baryonic dark matter is also **nearly non-interacting**. This is supported by upper limits on self-interaction from astrophysical observations (*e.g.* the Bullet cluster) and terrestrial searches.

3. Dark matter must be essentially **stable**, or at least have a lifetime long compared to the present age of the universe.

4. Finally, dark matter is apparently **cold** – that is, slow-moving. More precisely, a cold dark matter candidate must be non-relativistic throughout the formation of large-scale structure. The large random velocities of relativistic dark matter particles (*e.g.* light neutrinos) would disrupt the hierarchical formation of large-scale structure.

Beyond these three conditions, very little about dark matter is known. In particular, the identity of the particle or particles which constitute it remain a mystery. Numerous candidates have been proposed; below I give a brief listing of some of the more prominent candidates.

1.4.1 MACHOs

Unobserved ordinary matter is the simplest candidate for dark matter. In particular, it has been suggested that compact objects of various sorts, most commonly cold white dwarfs or black holes, could have remained undetected. These are generically referred to as "MACHOs", for "massive compact halo objects". MACHOs as initially proposed are assumed to be composed of baryonic matter, which runs afoul of constraints on the baryonic matter density from nucleosynthesis and microwave background observations. Nonetheless, searches for such objects serve as a consistency check on current models and a limit on the contribution from massive, non-baryonic structures.

Gravitational microlensing is the primary search mode for such objects. A microlensing event takes the form of the temporary brightening of distant stars due to the passage of a compact gravitational lens between source and observer. Several collaborations

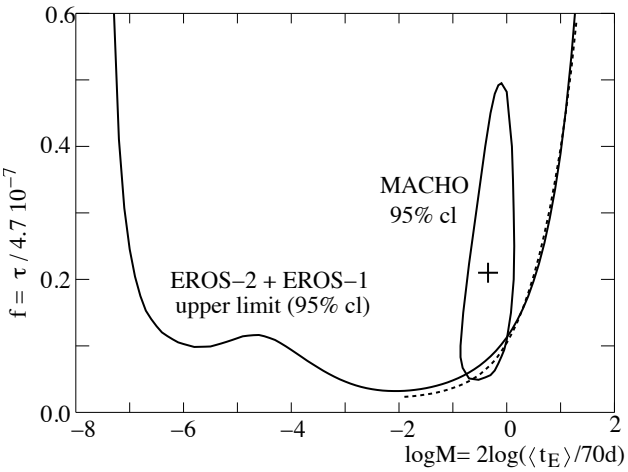


Figure 1.9: 95% confidence upper limits on the halo mass fraction f of MACHOs as a function of $\log_{10} M_{MACHO}/M_\odot$ from the EROS-1 and EROS-2 experiments. The closed contour represents the 2000 signal claim of the MACHO collaboration [35], and all curves are drawn under the assumptions of the S-model from that reference. Figure from [36].

have pursued this strategy, generally through long observations of stars in the nearby Magellanic Clouds. Most recently, the EROS-2 microlensing survey searched for such lensing events among 7 million bright stars for over a large solid angle (93 deg^2) toward the two Magellanic Clouds. Their 2007 results [36] found one candidate lensing event in 6.7 years, well below the ~ 39 expected from a halo composed of $\sim 0.4 M_\odot$ objects. This result essentially rules out MACHOs or other compact objects of mass $0.6 \times 10^{-7} < M/M_\odot < 15$ as primary constituents of the Milky Way halo, as illustrated in Figure 1.9. Compact objects with masses outside of this range are still allowed: lighter objects produce smaller lensing magnifications, while lensing events from heavier objects are too rare to be noticed.

1.4.2 Axions

The Lagrangian of quantum chromodynamics includes a CP-violating term which would, if present, endow the neutron with an electric dipole moment. Experimental limits on this dipole moment [29] force the coefficient Θ of this term to be extremely small ($\Theta \lesssim 10^{-10}$), when one might otherwise expect a value of order unity. In 1977 Peccei and Quinn proposed what is still the standard solution to this "strong CP problem" [37]: a global $U(1)$ symmetry whose spontaneous breaking forces QCD to the CP-preserving minimum. The axion is the pseudo-Nambu-Goldstone boson resulting from the breaking of this symmetry,

and is a popular dark matter candidate. The axion is characterized by a coupling constant f_a with units of energy, and its mass is given by [29]

$$m_a \simeq \frac{600\,\mu\,\mathrm{eV}}{f_a/10^{10}\,\mathrm{GeV}}. \tag{1.10}$$

Axions may be created in the early universe as a non-relativistic condensate through the "mis-alignment" mechanism [38]. The average axion relic density is

$$\Omega_a h^2 \sim \left(\frac{f_a}{10^{12}\ \mathrm{GeV}}\right)^{7/6}, \tag{1.11}$$

where h is the current Hubble parameter in units of 100 km/s/Mpc. This implies that an axion of mass $m_a \approx 10\mu$eV can constitute the universe's dark matter, while lighter axions are cosmologically excluded. Note that axions are produced non-thermally, and so are cold (non-relativistic) dark matter despite their very low masses.

Numerous astrophysical and laboratory constraints upon the axion have been devised, ranging from limits on their effects on the cooling of hot bodies (*e.g.* stars and supernovae) to searches for their effects on light propagation. The approximate reaches of various techniques are shown in Figure 1.10. The searches most relevant to axion dark matter are based on the Primakoff process [39], the conversion of axions to photons within a strong magnetic field. Such experiments attempt to detect excess photons in resonant microwave cavities within strong magnetic fields [40]. The leading such experiment at μeV masses is ADMX [41], which has already begun to probe cosmologically-interesting axion masses (see Figure 1.11).

1.4.3 Light neutrinos

Standard-Model neutrinos are an obvious candidate for non-baryonic dark matter, not least because (unlike many dark matter candidates) their existence has been experimentally verified. Observations of solar [42], atmospheric [43], and terrestrial [44, 45] neutrino oscillations have confirmed that at least some neutrino flavors have mass, with the heaviest having $M_\nu > 0.04$ eV [29]. Neutrinos must thus constitute at least some of the universe's non-baryonic dark matter.

The relic density of light neutrinos left behind from the early universe is determined by a process called thermal freeze-out, discussed in many standard texts [6, 7]. At sufficiently high temperatures, neutrinos and other particles are held in thermal and chemical equilibrium with the surrounding bath by a balance between the rates of annihilation

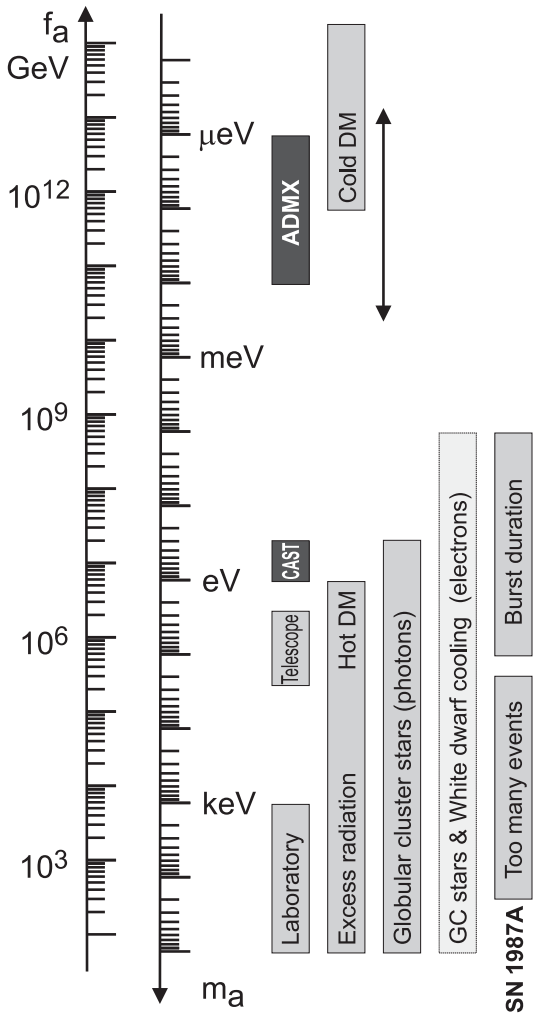


Figure 1.10: Approximate regions of axion parameter space excluded by various astronomical and laboratory constraints (*light rectangles*) and accessible to current and upcoming experiments (*dark rectangles*). The vertical arrow indicates the approximate range of axions with cosmologically-interesting relic densities. Figure from [29].

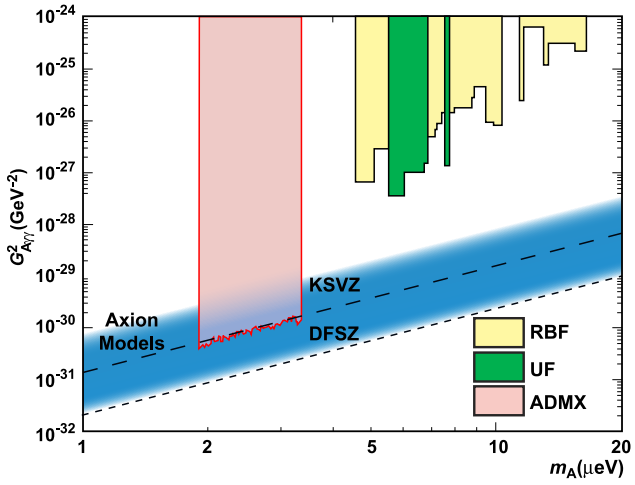


Figure 1.11: Current exclusion limits on the axion-photon-photon coupling from leading experiments in the region of axion dark matter candidates. The blue band indicates the approximate range of model-dependence for this coupling. The dashed and dotted lines indicate two common axion benchmark models. Figure from [29].

and of thermal production. The number density of light neutrino species i is just that given by Fermi-Dirac statistics, with negligible mass and vanishing chemical potential:

$$n_{\nu_i}^{(eq)}(T) = \frac{3g_i}{4\pi^2}\zeta(3)\left(\frac{kT}{\hbar c}\right)^3, \tag{1.12}$$

where g_i is the number of internal states available to neutrino species i ($g_i = 2$ for a Majorana neutrino and $g_i = 4$ for a Dirac neutrino). Equilibrium is enforced by neutrino annihilation, for which the rate (per particle) is given by

$$\Gamma_{\nu\nu} = n_\nu \langle \sigma_{\nu\nu} v \rangle, \tag{1.13}$$

where $\sigma_{\nu\nu}$ is the neutrino annihilation cross section, v is the relative velocity of the two neutrinos, and the angle brackets indicate an average over the thermal ensemble. Thermal equilibrium is only enforced so long as annihilation remains efficient within an expanding universe, which continues so long as

$$\Gamma_{\nu\nu} \gtrsim H \tag{1.14}$$

As the universe expands, this condition is violated at some "freeze-out" temperature T_f. Annihilation processes are inefficient after this point, leaving the number of neutrinos essentially unchanged from freeze-out until the present day. The expansion of the universe continues to dilute the neutrino energy density, however, as the neutrino density declines as

$n \propto a^{-3}$ and their individual momenta decline as $p \propto a^{-1}$. Note that this argument applies to any particle that was relativistic when it decoupled from the thermal bath.

Carrying through this calculation for neutrino annihilation gives a present-day relic density of

$$\Omega_{0,\nu} h^2 = \sum_{i=1}^{3} \frac{(g_i/2) m_i}{94\,\mathrm{eV}/c^2}. \tag{1.15}$$

Note that at least some neutrino flavors are non-relativistic today ($k_B T \sim 10^{-4}$ eV), but because their (comoving) phase space distribution was fixed at freeze-out this is essentially irrelevant to our computation. To account for the entirety of the observed non-baryonic dark matter, we would require $\sum_i m_i \sim 10$ eV, at the margin of consistency with direct bounds of $m_{\bar{\nu}_e} < 2$ eV from observations of tritium decay [29]. Light neutrinos are also hot dark matter, rather than cold, so they should have substantial effects on the formation of large-scale structure. Cosmological observations constrain $\sum_i m_i < 0.67$ eV [12], thus limiting neutrinos to no more than a few percent of the total population of non-baryonic dark matter.

1.4.4 WIMPs

Weakly interacting massive particles (WIMPs) are a generic class of dark matter candidate, widely regarded as among the most promising. A WIMP χ is a hypothetical particle with mass $10\,\mathrm{GeV} \lesssim M_\chi c^2 \lesssim 10\,\mathrm{TeV}$ and coupling strengths characteristic of the weak interactions. The strength of the WIMP hypothesis comes from a confluence between very general considerations of cosmology and particle physics. Both lines of reasoning point independently to the existence of such particles, as described below.

Consider a massive particle χ, in thermal and chemical equilibrium with the universe's thermal bath soon after inflation. Following the reasoning of Section 1.4.3, such a particle will remain in thermal equilibrium until the universe's expansion renders annihilation inefficient ($\Gamma_{\chi\chi} \lesssim H$). A massive particle χ may be non-relativistic when this occurs, causing its equilibrium density to be suppressed exponentially at low temperatures:

$$n_\chi^{(eq)}(T) \simeq g \left(\frac{M_\chi k T}{2\pi\hbar^2} \right)^{3/2} e^{-\frac{M_\chi c^2}{kT}}. \tag{1.16}$$

It is relatively simple to compute the relic density of a particle which is either highly relativistic or highly non-relativistic during freeze-out, but the calculation is more complex

in the transition region. Accurate calculations require numerical solution of the Boltzmann equation in an expanding universe

$$\frac{dn_\chi}{dt} + 3Hn_\chi = -\langle\sigma_{\chi\chi}v\rangle\left[(n_\chi)^2 - (n_\chi^{(eq)})^2\right]^2, \tag{1.17}$$

but an approximate solution for the relic density in the relevant regime is [46]

$$\Omega_\chi h^2 \simeq \frac{0.1\text{ pb}\cdot c}{\langle\sigma_{\chi\chi}v\rangle}. \tag{1.18}$$

Figure 1.12 indicates the general character of this solution: decoupling occurs at $M_\chi/T_f \approx 20$, with only logarithmic corrections from other particle properties, while the relic density is determined by the thermally-average annihilation cross section at decoupling. A particle with a larger annihilation cross section can remain in equilibrium slightly longer, leading to further Boltzmann-suppression and a substantially lower relic density.

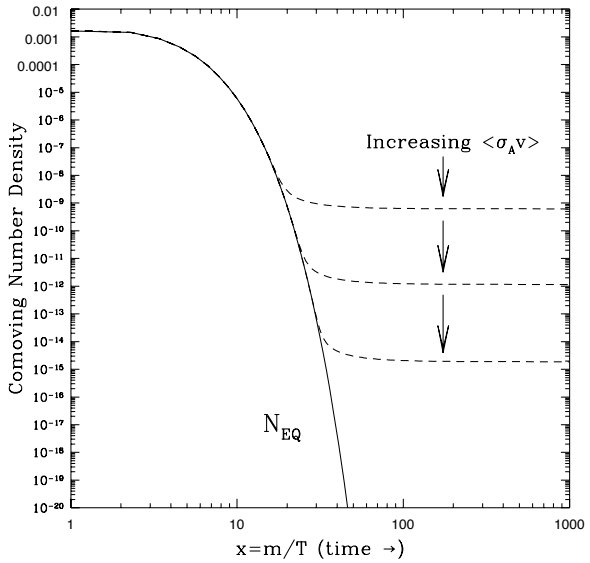


Figure 1.12: Evolution of a WIMP's comoving number density as a function of temperature (in units of the WIMP mass M) in the early Universe. The solid curve represents the equilibrium abundance, while the dashed curves represent actual abundances for various choices of velocity-weighted annihilation cross section. Figure from [6].

Equation 1.18 suggests that χ will have a relic density comparable to that observed for non-baryonic dark matter ($\Omega_{CDM}h^2 \sim 0.1$) if it has an annihilation cross section of order a fraction of a picobarn, a typical order of magnitude for the weak cross sections. By dimensional analysis, such a particle should have a mass of order $(Mc^2)^2 \sim \sigma/\alpha^2 \sim (10\,\text{GeV})^2$.

This is quite remarkable: with no detailed input from particle physics, cosmological data points to a possible connection between dark matter and the weak nuclear force. Any stable, massive (GeV-TeV) particle interacting through the weak nuclear force would quite naturally be the observed dark matter. Such a particle would also satisfy constraints from large-scale structure observations, since it is cold enough not to disrupt structure formation.

For completely independent reasons, our understanding of the Standard Model of particle physics also argues for the existence of new particles at the weak scale (~ 100 GeV). Electroweak symmetry breaking demands new processes at the energy scale of the W and Z bosons, generally assumed to be a scalar (spinless) Higgs boson of $M_H \sim 100$ GeV/c^2. The mass of a scalar particle is subject to radiative corrections [47], however, unless yet more particles are present which cancel those corrections. This instability is often called the "hierarchy problem": why is $M_Z \ll M_{Planck}$, when radiative corrections will naturally drive the Higgs mass (and thus the weak scale) upwards? Solutions to this problem (*e.g.* supersymmetry) generally require new particles at the weak scale to limit these corrections. If one of these new particles is stable, as it is in many models, it is a natural dark matter candidate. This is the strength of the WIMP hypothesis: observations of the universe on its largest and smallest scales point independently toward the WIMP theory of dark matter.

Chapter 2 discusses WIMPs and their experimental signatures in greater detail.

1.4.5 Other possibilities

In addition to the major candidates described above, countless other possibilities have been proposed over the years to explain the apparent need for dark matter. Each clearly represents physics beyond the Standard Model. I briefly describe a few of these other options below, though this list is far from exhaustive.

SuperWIMPs

If we allow dark matter to be created non-thermally, then we are free to consider candidates with much lower annihilation cross sections. Feng has dubbed such candidates "SuperWIMPs" [48]. The gravitino – the supersymmetric partner of the graviton and the lightest supersymmetric partner in some models – is the leading example of a SuperWIMP. If it exists, the gravitino is a spin-3/2 particle with interactions of roughly gravitational strength. Though the gravitino shares many attributes with the WIMPs described in Sec-

tion 1.4.4, its "weaker than weak" interactions distinguish it from that class.

Though not a WIMP itself, in some models the gravitino can nonetheless benefit indirectly from the same thermal production mechanism. Feng and collaborators have proposed models in which the next-to-lightest supersymmetric partner is a neutralino or sfermion which decays to the gravitino well after thermal freeze-out [49]. This decay process may disrupt the predictions of Big Bang nucleosynthesis, but need not if the decay occurs late enough ($10^4 - 10^8$ s in the Feng *et al.* models). The final gravitino density is reduced from that of the WIMP progenitor by a factor of $M_{\tilde{G}}/M_{NLSP}$, retaining the WIMP miracle while leaving dark matter composed of a much-less-detectable constituent. Such models have interesting collider phenomenology, particularly in cases where a charged NLSP is stable enough to exit a detector.

Super-heavy relics

Various theories have also been proposed which can produce dark matter candidates of relatively large masses, *e.g.* $M_{GUT} \sim 10^{15}$ GeV/c^2 or even $M_{Planck} \sim 10^{19}$ GeV/c^2. Examples of such particles include gravitationally-generated "WIMPzillas" [50] and non-topological solitons such as "Q-balls" [51]. Due to their large individual mass these candidates would be vastly rarer than standard WIMPs, which may help them avoid detection. This class of candidates remains very speculative, though some astrophysical bounds exist (*e.g.* [52]).

Modified gravity

Finally, perhaps the most naively-appealing explanation for these phenomena involves no dark matter at all. Milgrom has suggested [53] that a modification to Newtonian dynamics ("MOND") at very low accelerations might replicate the observed rotation curves of galaxies with no need for non-baryonic dark matter. In this view, dark matter is a complication of our cosmological model to account for our ignorance of gravity at large scales. MOND was initially an *ad hoc* one-parameter model, designed to fit the observed flat rotation curve but providing no predictions for gravitational lensing or early-universe signatures. Bekenstein has attempted to incorporate MOND-like phenomenology into a fully relativistic tensor-vector-scalar (TeVeS) theory with additional predictive power [54].

Despite its intuitive appeal, MOND suffers difficulties at the scale of galactic clus-

ters. It is particularly difficult to reconcile with the separation between visible and dark matter seen in the Bullet cluster; in the MOND paradigm, baryonic mass and lensing are expected to trace one another. Several solutions have been proposed to support MOND-like modified gravity on the cluster scale, with particular reference to the Bullet cluster. Angus and coauthors [55], for example, have discussed a solution in which MOND-like modified gravity is combined with neutrino dark matter. Given a neutrino mass just low below current upper limits (which may be addressable with the upcoming KATRIN neutrino experiment), they argue that MOND can replicate observations of this sort without exotic dark matter. Nonetheless, there is a strong consensus in the astrophysics community that non-baryonic cold dark matter best explains the wealth of cosmological observations available to us, and is a real component of the universe in which we live.

Chapter 2

Signatures of WIMP dark matter

As described in the previous Chapter, Weakly Interacting Massive Particles (WIMPs) are the leading class of candidates for the constituents of dark matter. Experiments are needed to test the WIMP hypothesis and to determine which (if any) flavor of WIMP makes up the universe's missing mass. After outlining some of the most prominent WIMP candidates, this chapter discusses three major classes of experimental signatures: collider production of WIMPs, indirect detection of WIMP annihilation, and direct detection of WIMP scattering. I discuss the first two briefly and more fully describe the third, which is the focus of this dissertation. I also touch upon the complementary relationships between the various detection techniques.

In the following discussion I avoid grammatical contortions by assuming that WIMP dark matter exists. I also frequently specialize to the case of supersymmetric dark matter for clarity, though most of the arguments carry over to other WIMP candidates. In particular, I will often use χ to symbolize a generic WIMP rather than just a SUSY neutralino. In all Feynman diagrams, time progresses from left to right.

2.1 WIMP candidates

The general idea of WIMP dark matter has been described in Section 1.4.4: a particle with mass and annihilation cross section characteristic of the weak interactions will naturally be produced in the right quantity to account for our observations of dark matter. The process of thermal freeze-out is well understood, and can be modeled for a variety of candidates using various software packages, *e.g.* DarkSUSY [56, 57] and MicrOMEGAs

[58, 59]. The confluence of these relic density considerations with the need for new particles at the TeV scale is a powerful motivation for the WIMP theory of dark matter.

Within the generic WIMP framework (if it is correct), we still need to know how the dark matter candidate fits into the larger framework of particle physics. Countless specific dark matter candidates have been proposed over the years, each associated with a particular theory of physics beyond the Standard Model. Below I give a brief outline of three of the most popular frameworks for WIMP candidates: supersymmetric theories, theories with extra spatial dimensions, and "little Higgs" theories. Though these candidates differ substantially in many of these WIMP candidates shares the same generic experimental signatures described later in this Chapter.

2.1.1 The hierarchy problem

In the highly successful Standard Model of particle physics [47, 60], the interactions among the fermionic fields of nature (quarks and leptons) are mediated by spin-1 gauge fields. Unlike the other gauge interactions of the Standard Model, the weak interactions are mediated by massive bosons: the $W^{\pm}$ ($m_W = 83$ GeV/c^2) and Z^0 ($m_Z = 91$ GeV/c^2). In the very successful Glashow-Weinberg-Salam theory, these particles are rendered massive by the spontaneous breaking of an SU(2)×U(1) gauge symmetry to the simpler U(1) of electromagnetism. This symmetry is broken by the formation of a spinless, SU(2)-doublet condensate, but the condensing field itself has thus far remained undiscovered in accelerator experiments. The condensing field is generally taken to be a fundamental scalar field – a Higgs boson, with mass comparable to that of the $W^{\pm}$ and Z^0 – but theories also exist in which the role of the Higgs is played by a composite particle or other construct. With the addition of this condensate (and of neutrino mass terms), the Standard Model is in excellent agreement with essentially all data from particle accelerators to date.

Despite its successes, the Standard Model has several problems which suggest that it is not a final theory of nature. Perhaps the most prominent of these is the hierarchy problem: new physics is needed at the TeV scale to stabilize the large hierarchy between the weak (10^3 GeV) and Planck (10^{19} GeV) scales without extreme fine-tuning. The masses of the Standard Model particles are subject to radiative corrections due to loop diagrams. The contributions of these diagrams are formally divergent, but they become finite if we assume that new physics will appear to regulate these loops at some high energy scale Λ

(*e.g.* the Planck scale). Such corrections are not overly upsetting as long as their magnitude depends only logarithmically upon Λ, but if they are large (*e.g.* proportional to a positive power of Λ) we expect them to drive the relevant masses toward Λ. Fermion and gauge boson masses are protected from such large corrections by chiral and gauge symmetries, respectively, but scalar masses are unprotected by any symmetry. The Standard Model must thus be fine-tuned to prevent the Higgs mass (and thus the W and Z masses) from sitting near the Planck scale.

The new weak-scale particles needed to solve the hierarchy problem are natural WIMP candidates. In fact, it is very common for the lightest new particle in such a model to be stable, and thus an ideal constituent of dark matter. The remainder of this Section surveys three examples of this general framework: supersymmetry, extra dimensions, and little Higgs theories.

2.1.2 Supersymmetry

Supersymmetry (SUSY) has emerged over the past few decades as the leading candidate for physics beyond the Standard Model. Supersymmetric extensions to the Standard Model also often provide very natural candidates for WIMP dark matter which are often accessible to direct detection. Below I describe the basics of supersymmetry as it pertains to dark matter. For further details of the theory and phenomenology of supersymmetric theories, see for example [61, 62, 63, 64]. For a full discussion of supersymmetric dark matter and its experimental signatures, see [46].

Superpartners

Supersymmetry is a proposed symmetry of nature between fermionic and bosonic fields. Supersymmetric models rest upon the observation that radiative corrections due to fermion and boson loops have opposite signs. If each fermionic degree of freedom in the Standard Model (1 per Majorana fermion, 2 per Dirac fermion) is matched to a corresponding "sfermion" of the same mass and gauge charges, their respective radiative contributions (*e.g.* from the two Feynman diagrams in Figure 2.1) cancel exactly. Similar cancellations occur if each gauge boson degree of freedom is matched to a fermion "gaugino" of the same mass and charges.

Supersymmetry clearly must be a broken symmetry of nature, however, since no

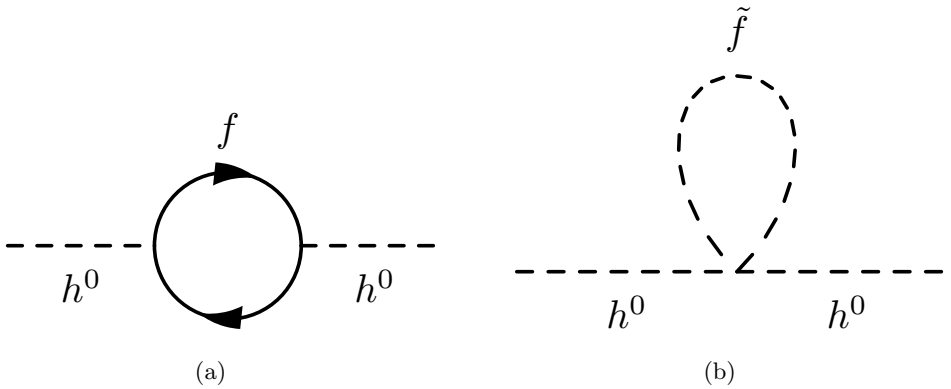


Figure 2.1: Loop diagrams giving radiative corrections to the Higgs boson mass from (a) a Standard Model fermion, and (b) its scalar supersymmetric partner.

such particles have been observed. This is not a major problem, however: the "superpartners" cancel out the worst of the radiative corrections even if they do not exactly match the masses of their partners, leaving the only a logarithmic dependence on the mass difference between them. If the new partner particles have masses near the TeV scale, radiative corrections are limited to this level and the weak scale is naturally stabilized.

Though the above justification is phenomenological, supersymmetry is not simply an *ad hoc* addition of particles; it has deep formal justification as a symmetry of nature. Supersymmetry can be thought of as the introduction of one or more "fermionic" extra dimensions of space whose coordinates are complex Grassmann (anticommuting) numbers. A quantum field theory developed in this "superspace" will naturally encompass bosonic and fermionic fields with interactions constrained to avoid divergent radiative corrections. Standard ($N = 1$) supersymmetry incorporates one such extra dimension, but more complex theories with several distinct fermionic dimensions ($N > 1$) have also been proposed.

Supersymmetry can also resolve a long-standing difficult with formulating Grand Unified Theories (GUTs), which unify the Standard Model gauge interactions at high energies as part of some larger gauge group [65]. For this to be true, we expect that radiative corrections will drive the various gauge couplings to a common value at some high energy scale. These corrections depend upon the masses and couplings of the various particles in the theory, and for the Standard Model particles the couplings almost unify at $\sim 10^{15}$ GeV, but not quite (see Figure 2.2). If we believe in GUTs, this suggests that new particles are needed to alter these radiative corrections and force the couplings into agreement. The addition of supersymmetric partners to the Standard Model particles at TeV masses alters

the way these coupling constants run, allowing them to unify near 10^{16} GeV. This higher unification scale also raises the proton lifetime sufficiently in some SUSY GUT models to evade experimental lower limits [66, 67, 68].

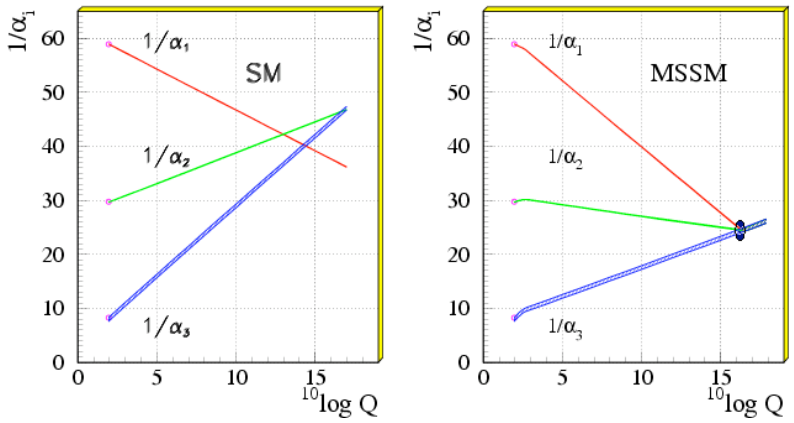


Figure 2.2: Gauge coupling unification in grand unified theories (GUTs) without (*left*) and with (*right*) TeV-scale supersymmetry. In both panels, the y-axis represents the inverse of each gauge group's dimensionless coupling constant ($U(1)_Y$, $SU(2)_L$, $SU(3)_c$) and the x-axis represents the base-10 logarithm of the renormalization energy scale. Figure taken from [69].

Supersymmetry can also be promoted to a local (rather than global) symmetry of spacetime [61], in analogy with the local gauge symmetries of the Standard Model. Local SUSY is known as "supergravity" (SUGRA) because it naturally incorporates general coordinate invariance, and thus contains general relativity in its low-energy limit. $N = 1$ SUGRA incorporates a spin-2 graviton and its spin-3/2 partner, the gravitino.

The MSSM

The simplest extension of particle physics which incorporates these ideas is the Minimal Supersymmetric Standard Model (MSSM) [70, 71]. The MSSM adds the minimal number of new particles necessary to incorporate the Standard Model into a supersymmetric theory:

1. A Majorana fermion for each Standard Model gauge boson before electroweak symmetry breaking: $\tilde{b}^0$, $\tilde{w}^0$, $\tilde{w}^{\pm}$, $\tilde{g}_i^0$.

2. Two scalars for each Dirac fermion, representing left and right chiralities: $\tilde{e}_{L,R}$, $\tilde{u}_{L,R}$, etc.

3. A second complex Higgs doublet (in addition to the one from the Standard Model) is required to give masses to both up-type and down-type quarks. After electroweak symmetry breaking, this introduces a rich Higgs sector: a light and heavy CP-even neutral Higgs (h^0 and H^0), a CP-odd neutral Higgs (A^0), and a pair of charged Higgs scalars ($H^{\pm}$).

4. Four Majorana fermion higgsinos, two charged and two neutral: $\tilde{h}^{\pm}$,$\tilde{h}^0_{1,2}$.

Groups of these particles with identical quantum numbers can mix to produce new physical mass eigenstates. In the MSSM, such mixing occurs for the electroweak Majorana fermions:

- Four Majorana neutralinos χ^0_i ($i = 1..4$), each a mixture of $\tilde{b}^0$, $\tilde{w}^0$, and $\tilde{h}^0_{1,2}$.
- Two pairs of Majorana charginos $\chi^{\pm}_j$ ($j = 1, 2$), each a mixture of $\tilde{w}^{\pm}$ and $\tilde{h}^{\pm}$.

R-parity

If we write down the most general supersymmetric Lagrangian with this particle content which is invariant under the Standard Model gauge group, we obtain the usual gauge and Yukawa interactions of the Standard Model fields, as well as a new mass parameter μ which gives mass to the Higgs bosons. We also find several addition terms which mediate non-Standard-Model interactions breaking baryon or lepton conservation. The naive MSSM contains a tree-level vertex allowing an up-quark to change into a down-quark by emitting a spositron ($u \rightarrow d\tilde{e}^+$), for example, rather than a W^+ boson; this interaction is perfectly consistent with gauge charge conservation, but the analogous Standard Model process ($u \nrightarrow de^+$) violates angular momentum conservation. Taken together, these new interactions would permit tree-level proton decay through processes like those shown in Figure 2.3. Such processes are suppressed only by the TeV-scale masses of the superpartners, leading to proton decay rates far in excess of current experimental limits. To satisfy experimental constraints, a realistic theory of low-energy supersymmetry must thus remove (or at least strongly suppress) these new processes.

Nearly all practical supersymmetry models eliminate rapid proton decay by introducing a new discrete symmetry of nature called R-parity. An MSSM particle's R-parity can be expressed as:

$$R = (-1)^{2j+3B+L}, \tag{2.1}$$

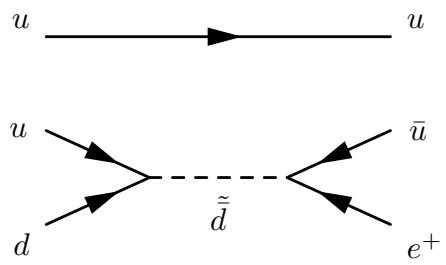


Figure 2.3: Example of a tree-level proton decay process $p \to \pi^0 e^+$ in the naive MSSM. Both vertices in this diagram violate R-parity (see text), so this process is forbidden in realistic supersymmetric models.

where j is the particle spin, B is its baryon number (1/3 for quarks and squarks, -1/3 for their antiparticles, 0 for others) and L is its lepton number (+1 for leptons and sleptons, -1 for their antiparticles, 0 for others). All Standard Model particles and Higgs bosons have an R-parity of +1, while their supersymmetric partners have an R-parity of -1. If R-parity is conserved, all of the MSSM's B- and L-violating couplings (and the associated rapid proton decay processes) are eliminated. Furthermore, this renders the lightest supersymmetric partner (LSP) absolutely stable, since the lightest $R = -1$ particle cannot decay without violating R-conservation.

Supersymmetry breaking

In order to match experiment, we need to break supersymmetry in the model above. In the MSSM we do not attempt to explain SUSY-breaking from first principles, but merely to describe the low-energy effective field theory after it is broken. The MSSM thus augments the supersymmetric Lagrangian with the most general set of interactions which break supersymmetry without re-introducing the Λ^n divergences that supersymmetry was introduced to eradicate in the first place. These "soft" SUSY-breaking terms encompass gaugino masses, scalar masses, and trilinear couplings among the various scalar fields. All told, the full MSSM contains more than 100 new free parameters – masses, trilinear couplings, various mixing angles, and CP-violating phases – beyond those of the Standard Model. Most of these parameters are in the soft terms, and so are presumably inter-related by the detailed physics of SUSY-breaking. Without an understanding of the SUSY-breaking mechanism, however, we must treat these as free parameters.

Constrained models

Due to the vast number of free parameters in the MSSM, phenomenological studies of supersymmetry are generally carried out in restricted subspaces of this parameter space. The most popular of these is the constrained MSSM (CMSSM), sometimes called "minimal supergravity" (mSUGRA) [72]. This model imposes a series of strong constraints inspired by GUTs and theories of gravity-mediated SUSY-breaking:

- All scalar masses unify to a common value M_0 at the GUT scale. Radiative corrections separate these masses at lower energies.
- Similarly, all gaugino masses unify to a common value $M_{1/2}$ at the GUT scale. To lowest-order, this causes the gaugino masses to run with renormalization scale in proportion to their respective coupling constants:

$$\frac{M_1}{\alpha_1} = \frac{M_2}{\alpha_2} = \frac{M_3}{\alpha_3} \propto M_{1/2}. \tag{2.2}$$

  This generically forces the gluino to be the heaviest of the gauginos and the bino the lightest, with $M_1 : M_2 : M_3 \sim 1 : 2 : 7$.
- All trilinear couplings take a common value A_0 at the GUT scale.
- All CP-violating phases beyond those in the Standard Model vanish.
- Electroweak symmetry is broken by radiative corrections to the Higgs potential, following the Coleman-Weinberg mechanism [73].

After imposing these constraints, the entire SUSY spectrum can be specified using only five free parameters:

1. The unified scalar mass M_0
2. The unified gaugino mass $M_{1/2}$
3. The unified trilinear coupling A_0
4. The ratio of the two Higgs vacuum expectation values, specified as $\tan\beta \equiv \frac{v_2}{v_1}$
5. The sign of the Higgs mass parameter μ.

This space is small enough to be explored through direct scans or Markov chain Monte Carlo techniques [74, 75]. This simplified model is almost certainly incorrect as a true representation of nature, but hopefully encompasses most of the major mechanisms at work in the correct theory.

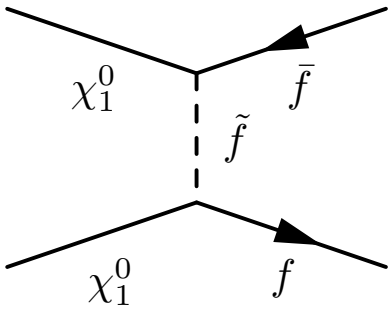
Dark matter

These simple SUSY models contain three neutral, colorless particles which are plausible candidates for LSP dark matter: the gravitino $\tilde{G}$, the lightest sneutrino $\tilde{\nu}$, and the lightest neutralino χ_1^0. More complex models can incorporate additional candidates (*e.g.* the axino). Of these three, the gravitino would be a non-thermal SuperWIMP (Section 1.4.5), while the sneutrino is essentially excluded by the null results of direct dark matter searches. The lightest neutralino is the most commonly considered SUSY dark matter candidate. It's basic phenomenology is now very well understood, as it's thermal relic density and experimental signatures have been studied extensively for more than 20 years (see *e.g.* [46]).

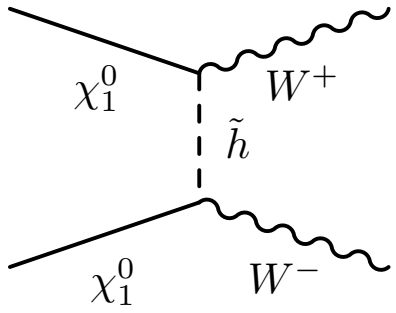
To be more concrete, we limit our discussion to neutralino dark matter within the CMSSM. The CMSSM generically has a $\tilde{b}$-like LSP due to unification constraints ($M_1 < M_2 < M_3$), though there can be a significant higgsino component in certain regions of parameter space. The small couplings of a nearly-pure $\tilde{b}$ give such an LSP a relatively low annihilation cross section, and thus a relic density somewhat greater than that implied by current cosmological observations [12]. Only in four "special" regions of parameter space with enhanced annihilation rates is the relic density low enough to be compatible with current constraints:

- The **bulk region**: In this region the SUSY spectrum is relatively light, so a variety of annihilation modes are kinematically accessible. Here, as elsewhere, neutralino annihilation generally takes place in the p-wave due to the Majorana nature of χ_1^0. This is the "generic" region in which SUSY was expected to lie, but much of it has been excluded by lower limits on Higgs and sfermion masses.

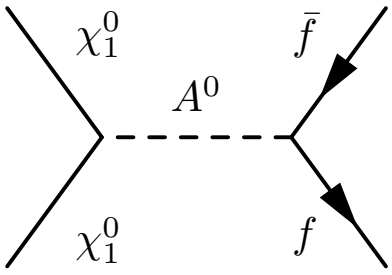
- The **focus point**, or hyperbolic branch [76, 77]: This is a region of heavy scalar particles, in which squarks and sleptons largely decouple from weak-scale physics. The lightest neutralino is also relatively heavy (often $\gtrsim 1$ TeV/c^2), opening up new



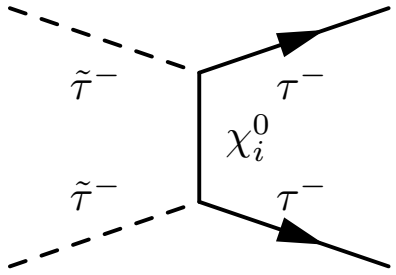
(a) Annihilation to Standard Model fermions, characteristic of the bulk region.



(b) Annihilation to gauge bosons, characteristic of the focus point.



(c) Resonant annihilation through A^0 in the rapid-annihilation funnel.



(d) A stau annihilation process relevant in the coannihilation region.

Figure 2.4: Feynman diagrams for representative annihilation processes in the four WMAP-allowed regions of the CMSSM.

s-wave annihilation channels to gauge bosons: $\chi\chi \to W^+W^-$, $\chi\chi \to Z^0Z^0$. Chargino coannihilation (see below) can also contribute significantly at high neutralino masses.

- The **rapid-annihilation funnel**: Here the lightest neutralino is approximately half the mass of some other particle in the theory (often the axial Higgs boson A^0), allowing it to annihilate rapidly on resonance.

- The **coannihilation tail**: If the lightest and next-to-lightest supersymmetric particles (LSP and NLSP) are nearly degenerate in mass, a substantial density of the NLSP (often a $\tilde{\tau}$) remains during the late stages of LSP annihilation in the early universe. This opens up pathways for "cooperative" annihilation: the LSP and NLSP are rapidly interconverted via $\chi\tau \to \gamma\tilde{\tau}$, and the annihilation rate of the charged NLSP is far greater than that of the neutral LSP. Such "coannihilation" processes result in lower relic densities for the LSP WIMP.

Figure 2.4 illustrates representative annihilation processes for the lightest neutralino in each of these four regions of parameter space. The experimental signatures of these various phenomenological regions are discussed in Section 2.5.1.

2.1.3 Extra spatial dimensions

Though supersymmetry is the most popular extension to the Standard Model, it is far from the only one. Much attention has also been focused in recent years on theories that postulate additional spatial dimensions to the universe beyond the three commonly observed. These extra dimensions may be either spatially flat or strongly curved. I describe each briefly below.

Flat extra dimensions

The simplest theories incorporating additional spatial dimensions are called Kaluza-Klein (KK) models. These theories augment the usual dimensions of spacetime (3 spatial and one time) with an additional d spatial dimensions, in the form of a spatially-flat compact manifold M. The simplest such model takes M to be a circle of radius R, but more complex manifolds are also possible. If R is sufficiently small, these extra dimensions may have minimal effects on macroscopic physics.

KK theories became prominent in recent years due to the work of Arkani-Hamed, Dimopoulos, and Dvali to address the hierarchy problem [78]. In a model with d extra dimensions of radius R, the gravitational force law changes with distance scale:

$$F_{grav} \propto \begin{cases} r^{-2}, & \text{for } r \gg R \\ r^{-(2+d)}, & \text{for } r \ll R. \end{cases} \tag{2.3}$$

Because of this faster decay law at short distances, gravity can appear to be much weaker at large distances than it actually is at high energies. If we augment spacetime with d universal, compact extra dimensions of total volume V_d, we find that the observed Planck scale M_{Pl} is related to the "fundamental" Planck scale $M_{Pl,0}$ by

$$M_{Pl}^2 = M_{Pl,0}^{d+2} V_d. \tag{2.4}$$

The hierarchy problem can thus be eliminated by the addition of two extra dimension with $R \sim 1\text{mm}$, which reduces the fundamental Planck scale to the TeV scale while maintaining the observed long-distance force law. Such theories are now strongly constrained by limits on variations in the gravitational force law at short distances, however, *e.g.* [79].

More recently, several authors have considered theories of one or more universal extra dimensions (UEDs) at the $R \sim 1\text{TeV}^{-1}$ scale [80], in which all particles of the Standard Model (rather than merely gravity) are free to propagate in the manifold M. Such models do not address the hierarchy problem, but they do predict new particles at the TeV-scale while remaining consistent with current observations. Momenta in the compact dimensions are quantized in units of $1/R$ (taking $\hbar = c = 1$) and manifest themselves as added mass; for each Standard Model particle of mass M_0, this gives rise to a tower of excited states with masses

$$M_n^2 = M_0^2 + \frac{n^2}{R^2} + \text{loop corrections}, \tag{2.5}$$

where n is a non-negative integer. For a single extra dimension, these KK-excitations have the same spins as their parent particles.

UED models have a natural analog to the R-parity of supersymmetry. Momentum conservation in the extra dimensions leads to a conservation of KK-index n in all particle interactions at tree level. Loop corrections break this down to a conserved KK-parity $K \equiv (-1)^n$. This limits the effects of the KK-excitations upon Standard Model processes to loop level, as well as rendering the lightest KK-excited particle (LKP) stable. If the LKP

is neutral and colorless (*e.g.* a KK-excitation of a graviton, neutrino, photon, Z^0, or Higgs), it is a natural WIMP candidate.

The phenomenology of LKP dark matter is qualitatively similar to that of SUSY's LSP dark matter. It's dark matter phenomenology is discussed extensively in a recent article by Arrenberg *et al.* [81], but is not pursued further in the present work.

Warped extra dimensions

Following a suggestion by Randall and Sundrum in 1999 [82], a strongly-curved extra dimension can also present a solution to the hierarchy problem. Their original setup proposed a 4+1 dimensional bulk spacetime bounded by two 3+1-dimensional spaces ("3-branes"). The extra dimension ϕ ($0 \leq \phi \leq \pi$) separating the branes is negatively curved, such that the overall metric becomes

$$ds^2 = e^{-2kR\phi}\eta_{\mu\nu}dx^\mu dx^\nu + R^2 d\phi^2, \tag{2.6}$$

where $eta_{\mu\nu}$ is the usual Minkowski metric in 3+1-dimensional space, R is the brane separation, and k is a number of order the Planck mass. This exponential factor causes the bulk graviton to be strongly concentrated away from the $\phi = \pi$ "infrared brane", exponentially suppressing the gravitational force strength observed by fields confined to this brane. In this way, the exponentially-small overlap of the graviton wavefunction with the infrared brane can cause a relatively small hierarchy to ($kR \sim 50$) to appear as a very large one ($M_{Pl}/M_H \sim 10^{15}$). Later work has established that this general picture holds even if essentially all Standard Model fields can propagate in the bulk, as long as the Higgs field is confined to the infrared brane. This picture can even explain the hierarchy of particle masses within the Standard Model through small changes in the profiles of the various fields in the extra dimension.

Theories of warped extra dimensions also predict a spectrum of new KK-excitations of the Standard Model particles with TeV-scale masses. Agashe and Servant [83] have proposed that such theories, when combined with grand unification considerations, can also predict a viable dark matter candidate. Just as in supersymmetric theories, warped GUT theories generically predict new baryon- and lepton-violating processes which contradict current bounds on the proton lifetime. In analogy with R-parity, the authors propose a new Z_3 parity as a proof-of-principle way to eliminate these interactions. This naturally renders the lightest Z_3-charged particle (LZP) a stable WIMP candidate, with phenomenology

similar to that of a SUSY neutralino and very promising signatures in direct-detection experiments.

2.1.4 Little Higgs

Little Higgs theories [84, 85] are a final example of the relationship between WIMP dark matter and solutions to the hierarchy problem. In analogy with the pions of the Standard Model, the Higgs is postulated to be a pseudo-Goldstone boson of some spontaneously-broken global symmetry of a new strongly-interacting sector. Several such symmetries are arranged to break in a collective manner, introducing new particle states at the TeV scale which conspire to cancel the quadratic divergences of the Higgs mass. Some tree-level corrections nonetheless remain in the simplest Little Higgs models which can spoil consistency with precision electroweak constraints. In exact analogy with SUSY's R-parity, one can impose a new "T-parity" under which all non-Standard-Model particles are odd and all tree-level corrections to the Standard Model are forbidden. The lightest T-odd particle (LTP) is a WIMP candidate [86].

2.2 Collider production

Particle accelerators are the traditional tools with which to investigate the phenomenology of new elementary particles. Hadron colliders such as Fermilab's Tevatron and CERN's Large Hadron Collider (LHC), as well as electron-positron colliders such as the proposed International Linear Collider (ILC), will push the energy frontier of the Standard Model. These instruments are designed to characterize particle phenomena at the TeV energy scale, the regime associated with electroweak symmetry breaking. If dark matter is related to these new phenomena, accelerators may produce WIMPs in the laboratory and illuminate the associated physical theory.

Due to their small interaction cross sections, WIMPs are expected to escape particle detectors unnoticed. The primary signatures of a WIMP are thus similar to those of a neutrino: missing energy and momentum from a particle collision. These signatures are relatively straightforward in a lepton collider where the collisional energy and momentum are well known. By scanning the energy of the colliding beams, the ILC or similar instrument can precisely measure the masses and cross sections of particles it produces [87]. A lepton collider with polarized beams can often also distinguish the spins of these new particles

through precision studies of the angular distributions of scattering products [88]. In principle, the precision measurements possible at a lepton collider can accurately reconstruct the spectrum and properties of nearly all particles particles at accessible masses (less than 1/2 the center-of-mass energy).

Despite the advantages of lepton colliders, hadron colliders currently populate the energy frontier and will continue to do so for at least the next decade. WIMP hunting at hadron colliders is complicated enormously, however, by the composite nature of hadrons. Even when the momenta of the colliding protons are well known, those of their parton constituents (quarks and gluons) are not. The net longitudinal (along the beam) momentum of an individual hard parton collision is unknown, so we cannot evaluate missing energy or momentum on an event-by-event basis. A collision's total momentum transverse to the beamline must be relatively small, however, so a "lopsided" collision with large net momentum perpendicular to the beamline suggests the ejection of an unseen particle. WIMP searches thus generally focus upon searches for missing transverse momentum. If a large number of such events can be reconstructed, the statistical distribution of transverse momentum can be used to estimate the properties of the interaction, *e.g.* through the shapes and positions of Jacobian peaks.

It is still an open question just how well these statistical techniques will be able to reconstruct the detailed phenomenology of new physics at the TeV scale. If a rich set of particles and processes are visible at LHC energies, as at the benchmark point SPS1a/LCC1, it may be possible to reconstruct much of the particle spectrum from kinematic constraints [90]. The precision of a linear collider may be needed when fewer processes are visible at the LHC, however, or when similar models are difficult to distinguish.

Though the analysis of data from a hadron collider is difficult, there is good reason to believe that it will be rewarding. Though the WIMP itself has very small cross sections for detection and production, many models of WIMP-related physics also predict new particles of similar mass with color (squarks and gluinos in supersymmetry, K-K quarks and gluons in UED theories, etc.). Hadron colliders are extremely good at pair-producing colored particles, with cross sections in the tens of picobarns at the LHC [87]. The new colored particle should decay down to the WIMP and a spray of partons, leading to a distinctive signature: hard jets plus missing transverse energy. As an example, Figure 2.5 shows the predicted reach of the LHC into supersymmetric parameter space through the production of squarks and gluinos. Either particle should be detectable at the LHC if its mass lies

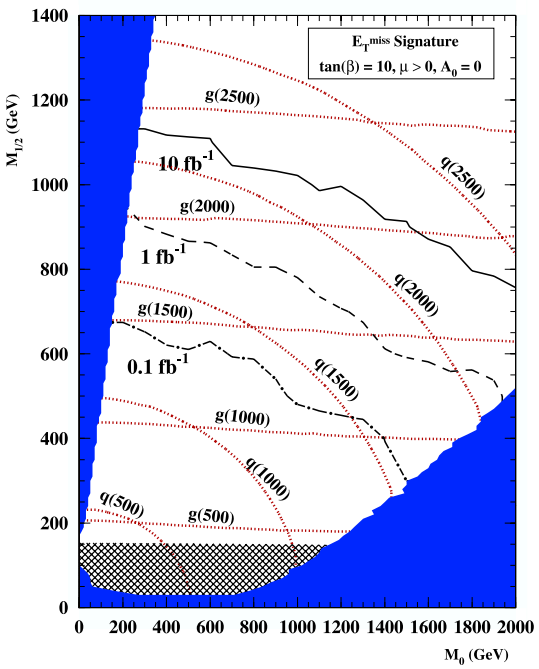


Figure 2.5: Projected reach of the LHC for colored SUSY particles through a missing-energy-plus-jets signature. Red dotted lines indicate various mass contours for gluons and squarks. Black curves indicate the reach from various exposure levels at the LHC (regions below the curves lead to discovery). For comparison, the LHC's first-year design luminosity is $\sim 10\text{fb}^{-1}$. Figure from [89].

below ~ 2 TeV/c^2, and their decays should produce many LSPs and other related particles.

Direct collider constraints upon the lightest neutralino or other WIMP candidates tend to be somewhat model-dependent, since the most probable WIMP-generating pathways are generally the decay chains of heavier charged or colored particles. LEP's tight constraint on the invisible width of the Z^0 [29] sets strong limits on the interactions of any WIMP with $M_\chi < M_Z/2 \approx 45$ GeV/c^2. This constraint can be avoided in some models, however: a sufficiently bino-like neutralino, for example, can have vanishingly small couplings to the Z^0. Lower limits of 18 GeV/c^2 [91] and 6 GeV/c2 [92] have been proposed for supersymmetric models based primarily upon LEP's lower limits on slepton masses, but even these bounds can be defeated in more general models [93]. Thus far no absolute lower bound on the WIMP mass exists from collider constraints, but accelerator constraints generally favor masses of order 100-1000 GeV/c^2.

2.3 Indirect detection

If the WIMP model of dark matter is correct, the same WIMP annihilation processes that determined the relic density in the early universe (Section 1.4.4) may continue today in regions of high dark matter density. Rather than attempt to produce WIMPs in a laboratory, we can thus make use of the ready supply already present in the cosmos. WIMP detection can thus be taken as a problem in observational astronomy: a search for signatures of WIMP annihilation in the sky. Such astrophysical searches are collectively known as "indirect detection" [46, 94].

The rate of WIMP annihilation is proportional to the square of the local WIMP density ($\Gamma_{\chi\chi\to X} \propto n_\chi^2$), so the observable flux is vastly enhanced in regions of high WIMP density. The most obvious targets are thus the central regions of dark matter halos, ranging from large galactic halos to the mini-halos of dwarf galaxies [95] or the substructure of our own galaxy [96]. Elastic scattering may also concentrate WIMPs in the centers of massive bodies such as the Sun and Earth, leading to enhanced annihilation there as well [97, 98].

Predictions of indirect detection signals suffer from substantial model-dependencies from both particle physics and astrophysics. Though the total WIMP annihilation cross section in the hot early universe is well-known from the WIMP's relic density, different annihilation processes may dominate in today's cold universe (see Section 2.5). Furthermore, only the inclusive annihilation cross section is known – the annihilation rate into a specific

detectable channel (*e.g.* $\chi\chi \rightarrow \gamma\gamma$) is model-dependent. Though they vary significantly between WIMP models, these quantities can be calculated precisely within a given model. More speculatively, Arkani-Hamed and coauthors have recently suggested [99] that a new force carrier in the dark sector may increase the WIMP annihilation rate in the cold modern universe substantially through Sommerfeld enhancement: an increase in the interaction rate between low-velocity particles due to attractive forces between them.

On the astrophysical side, there are large uncertainties surrounding the density profiles and uniformity of dark matter halos. Numerical simulations of cold dark matter structure formation tend to predict a sharp central cusp in the density profile [100, 101] and substantial small-scale substructure from dark matter clumps, but there is large variation between models. Many observations suggest a "cored" profile that is flatter near the center [102], suggesting that feedback from the baryon system may alter the halo shape. These halo uncertainties are often expressed in terms of two dimensionless quantities:

1. The normalized line-of-sight density

$$J(\Omega) \equiv \int \frac{dz}{r_0} \frac{\rho^2(z)}{\rho_0^2}, \tag{2.7}$$

   where $r_0 = 8.5$ kpc is the distance to the galactic center, $\rho_0 = 0.3$ GeV/cm^3 is the assumed local dark matter density, and the integral is over a line of sight. This factor depends upon the smooth structure of the halo, and is enhanced by sharp cusps at its center.

2. The boost factor $B \equiv \langle\rho^2\rangle/\langle\rho\rangle^2$. This factor is increased by clumpy substructure in the halo.

The WIMP annihilation rate is proportional to both $J(\Omega)$ and B, but these quantities are very poorly constrained from experimental data. Figure 2.6 illustrates the density profiles and density integrals predicted by three common halo models: an isothermal sphere, a Navarro-Frenk-White (NFW) cusped profile [100], and a steeper Moore profile [101]. Our lack of knowledge of these two factors translates into an uncertainty of several orders of magnitude in the annihilation rate, a severe limit on the accuracy of predictions. The recent Via Lactea II simulation of the dark matter halo of a galaxy similar to our own Milky Way [103, 104] suggests a density profile intermediate between the NFW and Moore profiles and a boost factor of order ~ 10 or less.

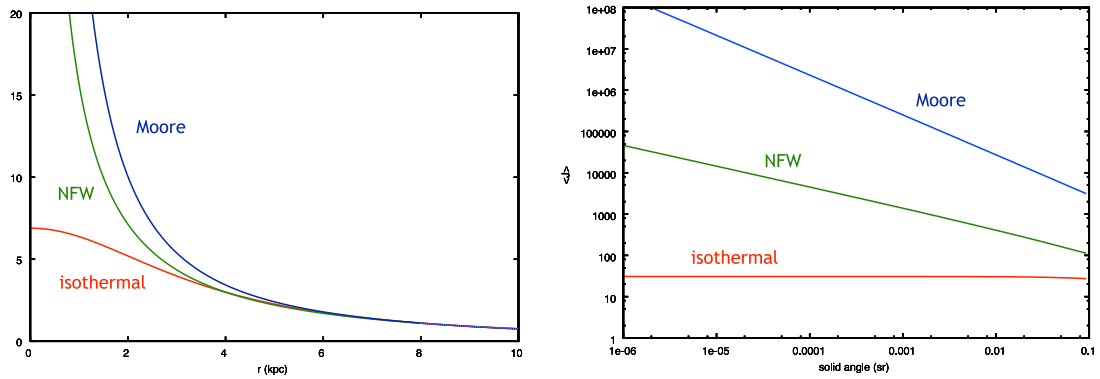


Figure 2.6: (a) Radial halo density profiles as functions of galactic radius near the galactic center, as suggested by three common models. (b) Corresponding values of $\langle J \rangle$ in these models, as functions of the angular resolution looking toward the galactic center. Both figures from [87].

Indirect WIMP searches encompass a wide variety of astrophysical techniques, depending on the choices of annihilation product and observational target. Several experiments have reported hints of possible signals, but thus far none has been widely accepted as evidence for dark matter. A full review of this large and growing field is far beyond the scope of this work, but I outline some of the major signatures below. The relationship between this and other WIMP search techniques is considered further in Section 2.5.

2.3.1 Gamma rays

Gamma rays are particularly promising as dark matter signatures because they are relatively easy to detect and retain their directionality over extragalactic distances [105]. In principle, a sensitive gamma ray detector may not only identify the annihilation signature of dark matter but map its distribution in the local universe. In optimistic models a gamma ray telescope could explore the distribution of dark matter in the center and periphery of the Milky Way's own halo, as well as numerous pinpricks from nearby substructure and the hazy extragalactic glow of unresolved halos throughout the cosmos [106].

Spectral features

If the WIMP can annihilate to a two-body final state containing a photon, this process will appear as a monoenergetic gamma ray line at GeV-TeV energies. This is an ideal "smoking gun" signature of WIMP annihilation: easily distinguishable from contin-

uum backgrounds and difficult to mimic with conventional astrophysical processes. The phenomenology of monoenergetic gamma ray peaks from supersymmetric WIMPs has been discussed extensively in the literature, *e.g.* [105, 107, 108]. Annihilation to $\gamma\gamma$ [109] and γZ^0 [110] are the most commonly-considered sources for such a line. Other final states may also be available depending on the WIMP model, such as γH (forbidden for Majorana WIMPs) or $\gamma\rho^0$. Since WIMPs are non-relativistic, the line corresponding to $\chi\chi \to \gamma X$ would be observed at $E_\gamma = (M_\chi^2 - M_X^2)/2M_X$. As expected, this expression simplifies to $E_\gamma = M_X$ for $\chi\chi \to \gamma\gamma$.

Unfortunately, the branching ratios for all such two-body annihilation modes are generally quite low. Since WIMPs are electrically neutral, there can be no tree-level diagrams for $\chi\chi \to \gamma X$; the amplitudes for these processes are suppressed by loop factors. Gamma ray telescopes have generally poor resolution, making It difficult to distinguish narrow, dim features above bright continuum backgrounds.

In addition to the line features described above, WIMP annihilation may also yield continuous gamma ray spectra. These result from annihilation processes involving charged final states, notably hadronic jets or W^+W^-, or bremsstrahlung from charged intermediate states. The latter can be particularly significant for Majorana WIMPs such as neutralinos, since it permits annihilation channels which would otherwise be forbidden in the S-wave configuration [111]. Continuous spectra are expected to be much brighter than isolated lines, but naturally much more difficult to interpret. We can nonetheless hope to distinguish some distinguishing feature of such a spectrum that separates it from astrophysical backgrounds, *e.g.* an unusually hard spectrum with an abrupt cutoff near the WIMP mass.

Observations

Electron-positron pair production renders Earth's atmosphere essentially opaque to gamma rays with energies $E \gg 1$ MeV. Gamma rays of sufficiently high energy ($E \gtrsim 30$ GeV) are nonetheless detectable using ground-based Air Cerenkov Telescopes (ACTs) [112]. An ACT is a large array of optical telescopes, each generally a large mirror focusing light upon an array of photomultiplier tubes (PMTs), used to monitor Cerenkov light emissions over a wide area of the sky. ACTs reconstruct the Cerenkov light patterns from particle showers produced in the atmosphere by high-energy gamma rays (and cosmic rays). Performance is roughly limited by the light-collection area of each telescope (larger dishes can

detect dimmer, low-energy events), the area of the array (larger arrays can collect more events in the same exposure time), and background rejection (more dishes and more pixels means better tracks, and thus easier rejection of cosmic rays and meteors). Prominent instruments in this category include CANGAROO, H.E.S.S., MAGIC, and VERITAS.

In 2006 the H.E.S.S. collaboration reported an excess of TeV gamma rays from the galactic center, consistent with a near-pointlike source [113]. The authors and others have proposed this as a possible signature of heavy (~ 10 TeV/c^2) WIMP annihilation. The observed spectrum is only marginally consistent with expectations from WIMP annihilation, however, and more closely resembles a power law typical of astrophysical sources (see Figure 2.7).

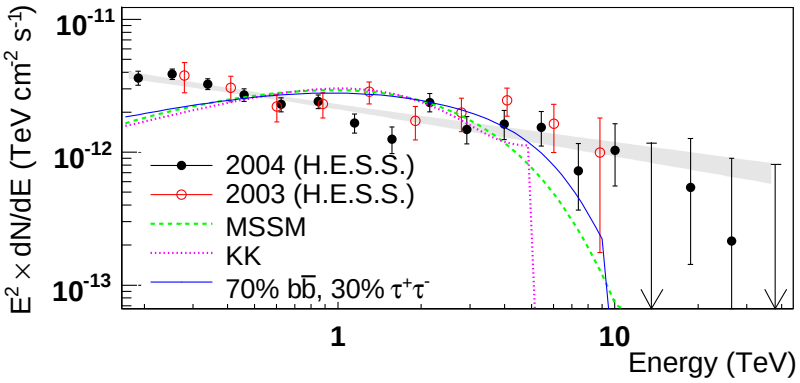


Figure 2.7: Excess of TeV gamma rays reported by the H.E.S.S. collaboration in 2006. The shaded band indicates a power law fit to the data, while other curves represent proposed fits WIMP fits to the data: a 14 TeV/c^2 neutralino (dashed), a 5 TeV/c^2 Kaluza-Klein particle (dotted) and a generic 10 TeV/c^2 with fine-tuned annihilation rates (solid). Plot taken from [113].

Space-based gamma ray telescopes permit full-sky mapping with superior energy resolution and lower energy thresholds, at the price of smaller detector areas and greater expense. The EGRET telescope aboard the Compton gamma ray observatory has produced the most extensive measurements to date on the gamma ray sky between ~ 20 MeV and 30 GeV. de Boer and collaborators have noted an excess of GeV-scale gamma rays in EGRET data and have proposed a WIMP interpretation of this excess [114]. They argue that this signal is consistent with annihilation radiation from a ~ 60 GeV/c^2 WIMP but not with expected galactic backgrounds, and have further used the EGRET sky map to estimate the WIMP distribution in the galactic halo under this assumption (see Figure 2.8). There

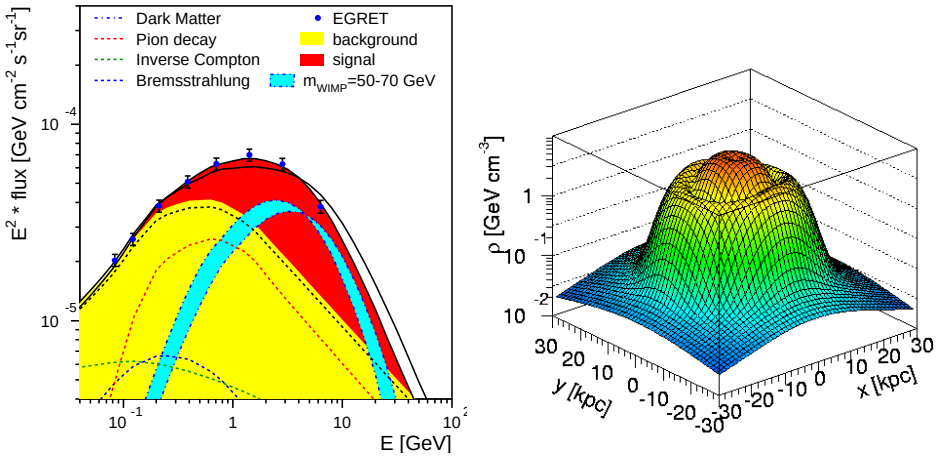


Figure 2.8: (a) Illustration of the proposed EGRET gamma ray excess (red), including the contribution from a ~ 60 GeV/c^2 neutralino model (cyan). (b) Proposed density profile of the galactic WIMP halo based upon the EGRET data. Plots taken from [114].

is substantial controversy about this claim, however. The observed gamma ray excess seems inconsistent with limits on antimatter production from WIMP annihilation[115]. The authors' proposed WIMP density profile traces the galaxy's visible matter more closely than expected for a WIMP halo, as well as demanding an anomalously-large boost factor of $B \sim 100$. Finally, several EGRET collaborators have proposed that instrumental calibration errors could explain the proposed signal [116].

EGRET data also indicates an excess of GeV gamma rays in the diffuse extragalactic background [118]. This spectrum has also been attributed to the annihilation of a WIMP of mass ~ 500 GeV/c^2, as shown in Figure 2.9 [117]. This signal is also consistent with generic models of blazars and other extragalactic sources, however, so the case for a dark matter origin is not airtight.

The Fermi Gamma-ray Space Telescope (formerly known as "GLAST"), launched on June 11, 2008, will provide an important leap in sensitivity for indirect searches in the next few years. Through its main instrument, a large silicon tracker known as the Large Area Telescope (LAT), Fermi promises unprecedented sensitivity to gamma rays in the 10 MeV - 300 GeV regime. The LAT improves upon EGRET with an increased field of view, greater target area, superior energy resolution, and an improved anti-coincidence shield to limit the self-vetoing which plagued EGRET. Fermi has already released its first image of the gamma ray sky (Figure 2.10) and promises to revolutionize gamma ray astronomy in

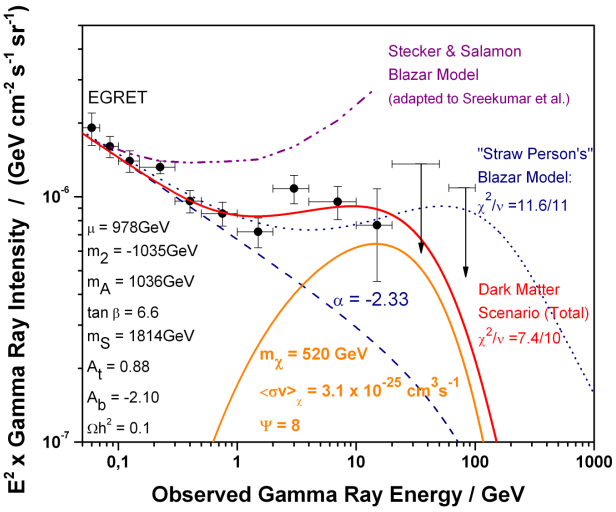


Figure 2.9: Extragalactic gamma ray observations from EGRET, showing an apparent excess above 1 GeV. Also shown are expectations incorporating a candidate WIMP model (red solid) or simple models of extragalactic blazars (dot, dash-dot). Figure taken from [117].

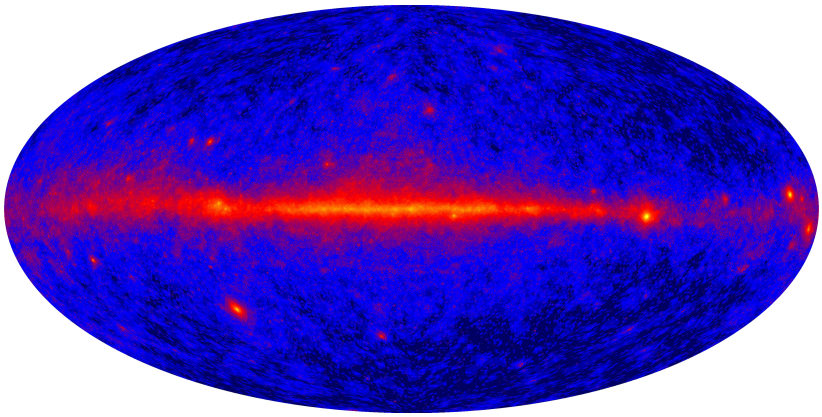


Figure 2.10: "First light" image of the gamma ray sky from the Fermi Gamma-ray Space Telescope. Image from [119].

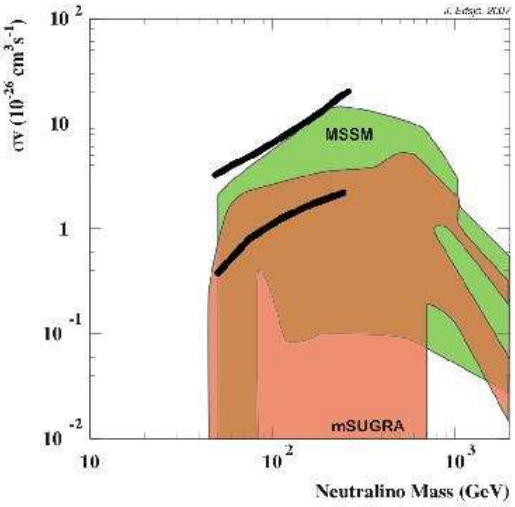


Figure 2.11: Projected 5-year reach of Fermi into supersymmetry parameter space, assuming an NFW halo profile [100]. The upper black line represents the estimated 5σ sensitivity of an analysis focused on the galactic center, the lower on a broader halo analysis. Plot taken from [120].

the coming years. Fermi observations of our own galactic halo are expected to probe the most optimistic neutralino models within the NFW halo profile (Figure 2.11), but plausible changes in the halo profile and boost factor (*e.g.* the steeper profile suggested by Via Lactea II) can improve this reach by up to two orders of magnitude [120, 121]. Searches for WIMP annihilation in galactic substructure, satellite galaxies, and extragalactic sources also hold great promise.

2.3.2 Neutrinos

Neutrinos are another leading tracer for WIMP annihilation. Like gamma rays, neutrinos retain directional and spectral information over cosmological distances. Their extremely low scattering cross section renders them vastly more difficult to detect, but also allows them to penetrate dense matter. This makes neutrinos ideal probes of WIMP annihilation in the cores of massive objects such as the Sun and Earth [97, 98]. Neutrinos are also a possible signature of WIMP annihilation in the galactic center, but the observable signal is greatly suppressed by the small solid angle of the Earth as seen from the galactic center.

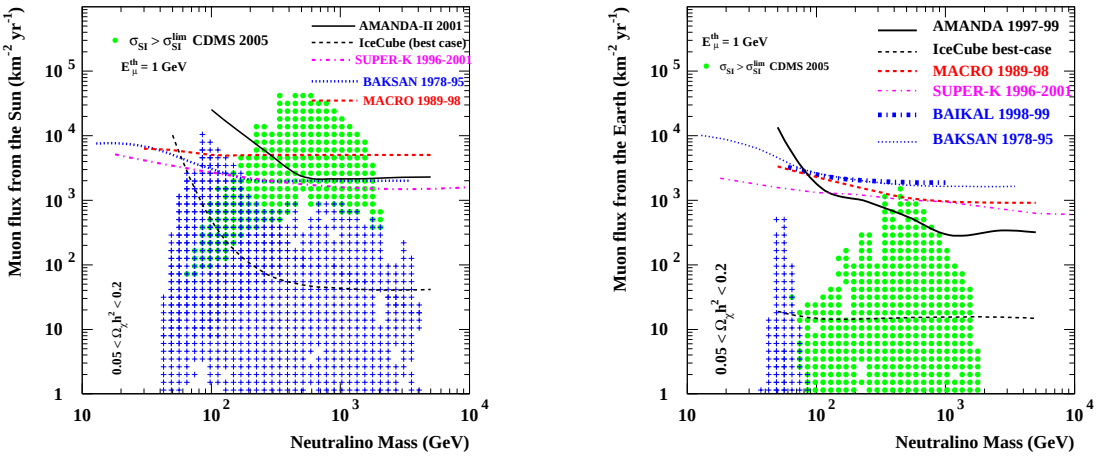
Though WIMPs generally travel through even massive objects unimpeded, a few

will scatter elastically and lose energy. A WIMP that loses enough energy will become gravitationally bound to the scattering object, leading it to scatter more frequently in the future. A WIMP captured by the Sun is expected to settle near the solar center over a time scale short compared to the age of the solar system, yielding a WIMP overdensity. If the capture rate is high enough, an equilibrium should have already been reached between WIMP capture and annihilation in the solar center. Most annihilation products are captured rapidly in the solar material, but many neutrinos can escape to be detected in terrestrial detectors. Capture in the Earth is also possible but much slower, and so generally less promising [122].

If a high-energy ($\gg 1$ GeV) muon neutrino reaches the Earth, it may scatter in material near the surface to produce an energetic muon. This daughter muon will be emitted in a direction well-aligned with that of the incident neutrino, thus retaining directional information about the annihilation source. Upward-going muons of this sort are easily distinguished from solar neutrinos by their far greater energies (solar neutrinos have energies characteristic of nuclear processes: keV-MeV) and from atmospheric neutrinos by their directional correlation with the Sun's position in the sky.

High energy neutrino instruments generally take the form of large Cerenkov detectors, in which an array of photomultipliers detect the Cerenkov light from energetic charged particles in a large volume of water (Super-Kamiokande, ANTARES) or Antarctic ice (AMANDA, IceCube). Target volumes range from Super-Kamiokande's 50,000 m^3 to IceCube's eventual goal of $\sim$ 1 km^3. Energy thresholds are near 5 MeV for Super-Kamiokande, but rise to $\gtrsim 10$ GeV for the larger detectors due to sparser instrumentation. Low energy thresholds are not necessarily crucial for indirect searches, however, since the expected signal is relatively high in energy itself. Moreover, the sensitivity of such experiments grows significantly at higher neutrino energies: the cross section for neutrino scattering scales as E_ν, as does the penetration depth of the resulting muon through rock (and hence the effective mass of the rock target). Conversely, however, the probability of a neutrino escaping the Sun without scattering declines exponentially with energy. This effect depletes the neutrino flux above ~ 200 GeV [94].

In many ways, neutrino-based WIMP searches are more similar to direct detection experiments (see Section 2.4) than to indirect searches with gamma rays. Under the assumption of capture-annihilation equilibrium, the annihilation rate in the solar center depends upon the cross section for WIMP-nucleon elastic scattering rather than the annihi-



(a) Neutrinos from the Sun's center.

(b) Neutrinos from the Earth's center.

Figure 2.12: Limits and projections for indirect WIMP detection with various neutrino detectors. Colored points indicate a sample of MSSM models. Green circles are excluded by CDMS's 2005 result[3], while blue crosses are not. Plots taken from [123].

lation cross section. Since the Sun is composed primarily of hydrogen, the spin-dependent WIMP-proton cross section generally dominates the capture rate for SUSY neutralinos. Spin-independent scattering also contributes to the capture rate, and may dominate for *e.g.* scalar (spinless) dark matter. The expected neutrino flux rate also depends upon the WIMP's annihilation branching fractions and somewhat upon the detailed model of WIMP capture.

Generally speaking, neutrino detectors are a less-promising avenue for WIMP observation than direct-detection experiments. Figure 2.13 indicates the relative WIMP-sensitivities of neutrino telescopes and direct-detection experiments, as computed by Kamionkowski *et al.* [124]. For comparison, the current Super-Kamiokande detector has an effective area of $\sim$ 1200 m^2 for a WIMP-search using upward-going muons [125]. Even small direct-detection experiments are clearly superior to large neutrino detectors for the detection of WIMPs with significant spin-independent couplings (true in most WIMP models), but neutrino detectors are generally dominant for WIMPs with large spin-dependent scattering cross sections. Figure 2.12 illustrates the approximate reach of current and future indirect searches with neutrino detectors. Note that neutrinos from the solar center (Figure 2.12(a)) are generally a more promising signal than those from the Earth's center (Figure 2.12(b)).

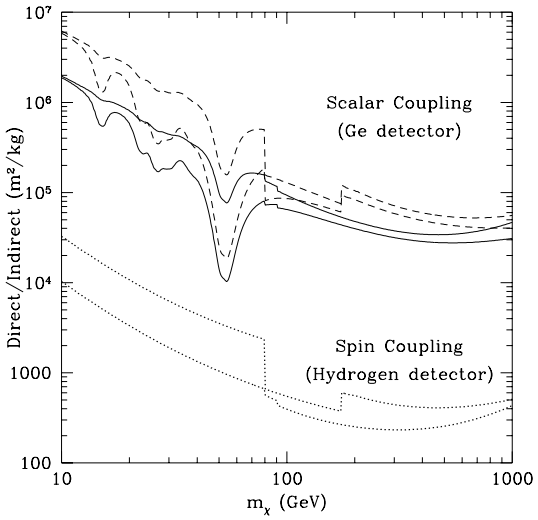
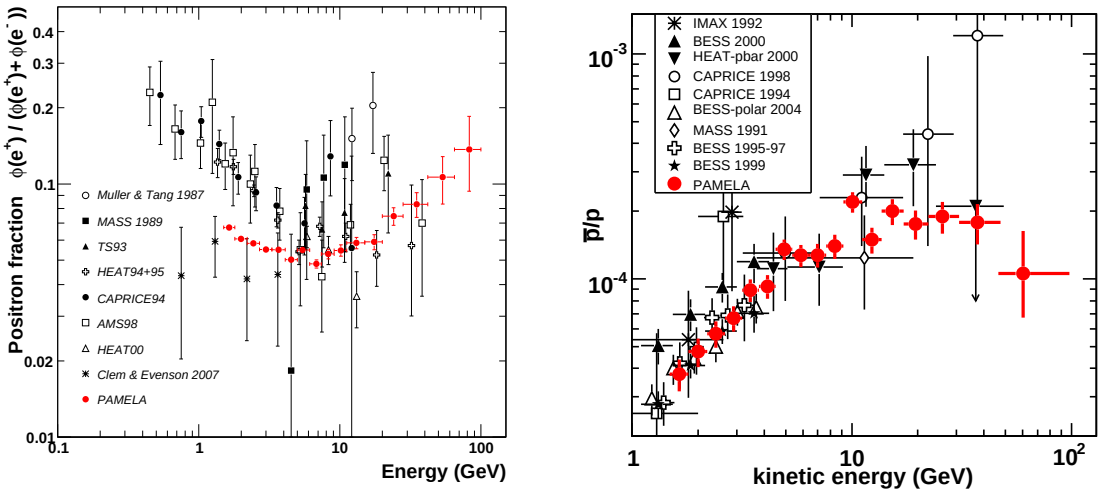


Figure 2.13: Relative sensitivities of neutrino detectors and direct-detection experiments to WIMP dark matter, taken from [124]. The y-axis indicates the logarithm of the relative reach of a 1-kg direct detection experiment with respect to that of a neutrino detector with 1 m^2 effective area, exposed for equal times. The upper (lower) group of lines represents a WIMP with spin-independent (spin-dependent) couplings and assumes a Ge (H) detector. The spread of lines within each group represents theoretical uncertainties, as described in the original paper.

2.3.3 Antimatter

Anomalous populations of antimatter in cosmic rays – notably positrons, antiprotons, and antideuterons – are also a possible signature of WIMP annihilation in the cosmos [46]. Antimatter particles are usually secondary annihilation products, so they generally lack strong spectral features. The propagation of antimatter through the galactic magnetic field is complex and diffusive, erasing directional information and producing large theoretical uncertainties in the expected antimatter flux. Antimatter can nonetheless be a prominent tracer for WIMP annihilation in energy regimes where astrophysical backgrounds are low.

WIMP annihilation can produce positrons either directly (though $\chi\chi \rightarrow e^+e^-$ is forbidden for Majorana WIMPs) or through the decay of muons, themselves generated directly or from the decay of pions, $W^{\pm}$, etc. The multi-body nature of this process produces a broad spectrum of hard positrons, though some models do predict a significant cutoff near $M_\chi/2$ or other spectral features [46]. WIMP annihilation is expected to produce an equal mixture of electrons and positrons, but this raises the (generally low) electron-positron ratio from more mundane background processes. WIMP annihilation should generally appear as



(a) (b)

Figure 2.14: Combined data on antimatter cosmic rays from recent observations by PAMELA (*red*) and other insruments. (a) Positrons as a fraction of the total electron-positron flux [126]. (b) Antiproton-proton ratio [127].

a broad excess of high-energy positrons, inconsistent with the declining power laws expected from background processes.

The PAMELA antimatter telescope [128], launched on June 15, 2006, has brought a great improvement in sensitivity to this field. The PAMELA detector array maintains $\sim$ 80% detection efficiency for positrons while rejecting protons from its positron sample at a rate of 10^5 : 1, making its data sample unusually free of background. PAMELA's first results [126], announced in 2008, show a large excess of positrons in the 10-100 GeV regime, as shown in Figure 2.14(a). The balloon-borne ATIC detector has recently reported an excess of cosmic ray electrons at energies of 300-800 GeV citeChang:2008zz, lending credence to PAMELA's observation. Numerous authors (*e.g.* [129, 130, 131]) have already suggested that these features may be signatures of dark matter annihilation in the galactic halo.

Hadronic jets from WIMP annihilation can generate antiprotons and even antinuclei, and these may also contribute to cosmic rays. For antiprotons, we may look for excess flux at high or low energies [46, 132]. Several experiments, including the PAMELA satellite, have mapped out the antiproton spectrum without finding evidence for an excess

(see Figure 2.14(b)). Antideuterons are produced much more rarely, but the astrophysical background is also vastly lower [133, 134]. Experiments such as GAPS and AMS-02 will probe the antideuteron flux in Earth's vicinity.

In a related vein, the SPI instrument aboard the INTEGRAL satellite has reported in excess of 511 keV gamma rays from non-relativistic positron annihilation in the galactic center [135]. Boehm et al. [136] have put this forward as yet another possible signature of dark matter annihilation, this time of particles with masses of $\sim$ 1-100 MeV/c^2. Other galactic sources of positrons are known, however, and this may easily have a more mundane astrophysical origin.

2.3.4 Synchrotron radiation

Finkbeiner [137] has recently noted a diffuse "haze" of microwaves concentrated in the inner 1-2 kpc of the galactic center, after subtracting the CMB and other known sources. The haze has an extremely hard spectrum, inconsistent with known astronomical processes in the central portion of the galaxy. This hard spectrum is consistent with synchrotron emission from a large population of ultra-relativistic e^+e^- concentrated at the galactic center. The annihilation of heavy WIMPs can produce just such a population [138], making the haze a very intriguing candidate for indirect detection.

Additional work by Hooper, Finkbeiner, and Dobler [139] argues that this signal is consistent with a WIMP halo profile of $\rho(r) \propto r^{-1.2}$, somewhat cuspier than the $\rho_{NFW}(r) \propto r^{-1}$ of an NFW profile [100] and consistent with predictions from Via Lactea II. The total power necessary to generate the haze is also broadly consistent with that expected for annihilation of a generic SUSY WIMP within this halo model. If confirmed, this sharp halo profile implies very good prospects of indirect searches with Fermi and other gamma ray telescopes [121].

2.4 Direct detection

Because of Earth's position within the dark matter halo of the Milky Way, WIMP interactions are not merely an astronomical phenomenon but are also a terrestrial one. Once in a great while (current upper limits are less than one scattering event per month in one kilogram of Ge) a WIMP should scatter elastically upon any material sample, and a sufficiently sensitive detector may be able to observe this interaction [140, 141]. This is

the goal of CDMS and a host of other direct detection experiments: to observe the rare interactions of galactic WIMPs in laboratory particle detectors.

The impetus to pursue direct detection comes from an argument of "crossing symmetry" between annihilation and scattering processes. The scattering process $\chi q \to \chi q$ and the annihilation process $\chi\chi \to q\bar{q}$ share the same matrix element, but have different kinematic factors. We have already argued (*cf.* Equation 1.18) that the WIMP self-annihilation cross section must be of order ~ 1 picobarn to set its relic density to the observed value. This is a cross section characteristic of the weak interactions, so it is plausible that other WIMP interaction cross sections may be of similar magnitude. It is important to note that this is not a quantitive prediction, but merely an argument for an approximate order of magnitude. Many model-dependent considerations can alter the relationship between scattering and annihilation cross sections, notably the $\chi\chi \to q\bar{q}$ branching fraction, resonant annihilation processes, or enhancements to the annihilation rate at high temperatures (*e.g.* coannihilation processes).

Based solely upon this simple argument, we can estimate the approximate event rate we might expect from a generic WIMP. I assume a WIMP-nucleon scattering cross section $\sigma_{\chi p} \sim 10^{-36}$ cm^2, a local WIMP density of 0.3 GeV/c^2/cm^3, and an incident WIMP velocity of $\sim 10^{-3}c$. For a 100 GeV/c^2 WIMP, this corresponds to a scattering rate of a few events per day in a one-kilogram hydrogen target mass. This event rate can be altered enormously by model-dependent considerations, but we are nonetheless led to a general experimental design: a search for rare WIMP impacts occurring once (or less) per day of exposure and kilogram of target.

The bulk of this dissertation describes the Cryogenic Dark Matter Search (CDMS), a world-leading experiment pursuing event-by-event discrimination against background events. I discuss the general issues of direct WIMP detection in this Section, returning to a brief overview of other leading technologies in Chapter 8.

2.4.1 WIMP elastic scattering

Even without knowing the details of any specific WIMP model, there are several general conclusions we can draw about WIMP-nucleus elastic scattering. WIMPs are bound within the galactic halo and thus should travel at typical galactic velocities in the solar neighborhood: $v \approx 270$ km/s $\sim 10^{-3}c$. A WIMP with mass $M = 100$ GeV/c^2 thus has a

kinetic energy of $E_\chi \approx 50$ keV. In an elastic collision with a stationary target of mass m, the mean kinetic energy of the recoiling target (after averaging over recoil angles) is

$$\frac{E_{recoil}}{E_\chi} = \frac{2\frac{m}{M}}{\left(1+\frac{m}{M}\right)^2}. \tag{2.8}$$

The maximum recoil energy is twice this amount. An elastic collision between a 100 GeV/c^2 galactic WIMP and a Ge nucleus ($m_{Ge}c^2 \approx 67$ GeV) thus spurs the nucleus to recoil with typical kinetic energy of $\sim$ 25 keV. This energy deposition is detectable in a variety of low-threshold particle detectors. A WIMP-electron collision, in contrast, typically endows the electron with < 1 eV of kinetic energy – not enough to ionize a single charge carrier, and thus a challenging target for any massive detector technology. Direct detection efforts thus focus upon WIMP-induced nuclear recoils.

The slow speed of an incident WIMP also has important implications for its scattering cross section. A 100 GeV/c^2 galactic WIMP has a deBroglie wavelength of $\lambda_\chi = \frac{h}{M_\chi v} \approx 12$ fm, the diameter of a large atomic nucleus. An incident WIMP thus interacts coherently with an entire atomic nucleus rather than scattering off of single nucleons. Cross section computations must account for constructive or destructive interference among the individual WIMP-nucleon scattering amplitudes. Roughly speaking, a WIMP's long wavelength means that it "sees" the nucleus at this scale of resolution and cannot resolve individual nucleons.

A priori, the WIMP-nucleon scattering amplitude can take a variety of different forms depending on the WIMP's spin and the symmetries of its couplings: scalar, pseudoscalar, vector, axial-vector, tensor, or pseudotensor. Kurylov and Kamionkowski have shown [142] quite generally, however, that only scalar and axial-vector terms survive in the extreme non-relativistic limit; the remaining terms are either suppressed by large factors or can be absorbed into these two. In this limit, the general interaction Lagrangian becomes

$$\mathcal{L}_{\chi N} = 4\chi^\dagger\chi\left(f_p\eta_p^\dagger\eta_p + f_n\eta_n^\dagger\eta_n\right) + 16\sqrt{2}G_F\chi^\dagger\frac{\vec{\sigma}}{2}\chi\cdot\left(a_p\eta_p^\dagger\frac{\vec{\sigma}}{2}\eta_p + a_n\eta_n^\dagger\frac{\vec{\sigma}}{2}\eta_n\right) + \mathcal{O}\left(\frac{q}{M_{p,\chi}}\right), \tag{2.9}$$

where χ is the WIMP wavefunction, $\eta_{p(n)}$ is the proton (neutron) Weyl spinor, $\vec{\sigma}$ is the spin operator (Pauli spin matrices), and q is the recoil energy. The WIMP-nucleon interaction is thus characterized by five parameters: the WIMP mass M_χ, the "spin-independent" (SI) couplings f_p and f_n, and the "spin-dependent" (SD) couplings a_p and a_n. I consider these two classes of interaction separately below.

Spin-independent scattering

A spin-independent WIMP-nucleon interaction corresponds to a coupling to the nucleon density operators, characterized by coupling constants f_p and f_n to protons and neutrons, respectively. In the limit of vanishing momentum transfer (*i.e.* an extremely soft collision) the scattering amplitudes with individual nucleons add coherently, giving a total scattering amplitude $\mathcal{M}_{SI} \propto (Zf_p + (A-Z)f_n)$ for a nucleus with atomic number Z and atomic mass A. Squaring the amplitude and adding in kinematic factors, we obtain the spin-independent WIMP-nucleus cross section

$$\sigma^0_{SI} = \frac{4}{\pi}\mu^2_{\chi N}\left[Zf_p + (A-Z)f_n\right]^2, \tag{2.10}$$

where $\mu_{\chi N} = \frac{M_\chi M_N}{M_\chi + M_N}$ is the collision's reduced mass. Most supersymmetric and similar models predict $f_p \approx f_n$ [46], giving a cross section $\sigma^0_{SI} \propto A^2$. This coherent enhancement can be enormous ($A^2_{Ge} \sim 5000$), so a WIMP target built of heavy nuclei is vastly more sensitive to WIMP interactions than one composed of lighter nuclei.

For real collisions with non-zero momentum transfer, the various WIMP-nucleon scattering amplitudes will not interfere perfectly constructively and the actual cross section will be somewhat less than that in Equation 2.10. This reduction is expressed as a multiplicative nuclear form factor $F^2(q)$, where q is the recoil momentum and $0 < F^2(q) < 1$ is the normalized Fourier transform of the nuclear density distribution [143]. The differential WIMP-nucleon cross section is then

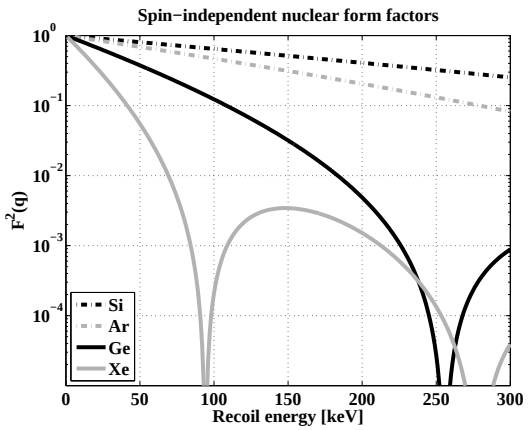
$$\frac{d\sigma_{SI}}{dq^2} = \frac{4}{\pi}\mu^2_{\chi N}\left[Zf_p + (A-Z)f_n\right]^2 F^2(q). \tag{2.11}$$

A convenient approximation to $F^2(q)$ for a wide range of nuclides (see [144] for recent commentary) under the assumption that protons and neutrons are similarly distributed and $f_p \approx f_n$ is the Helm form factor [143]:

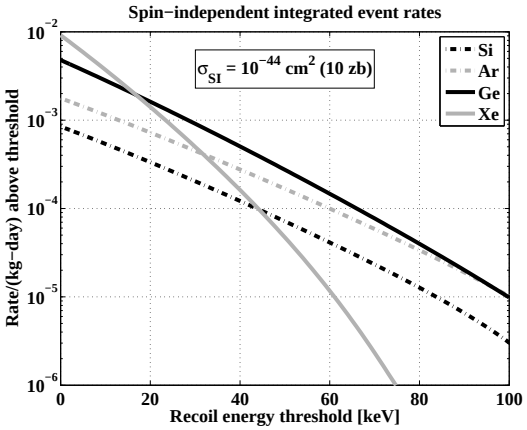
$$F(q) = 3\frac{j_1(qr_n)}{qr_n}e^{-(qs)^2/2}, \tag{2.12}$$

where r_n is the nuclear radius and s is the nuclear skin thickness. In this work I follow [143] and take $r_n = c^2 + \frac{7}{3}\pi^2a^2 - 5s^2$, $c = 1.23A^{1/3} - 0.60$ fm, $a = 0.52$ fm, and $s = 0.9$ fm.

Figure 2.15(a) plots this form factor for several choices of target nuclide. Note the periodic deep troughs in $F^2(q)$, corresponding to momentum transfers which give completely destructive interference across the nucleus. These dips in sensitivity occur at lower



(a) Helm nuclear form factor as a function of recoil energy for several nuclei relevant to WIMP direct detection. Note the minima at high recoil energies for heavy target nuclei.



(b) Integrated WIMP scattering rate as a function of recoil energy threshold for various choices of target nucleus. All curves for a WIMP with $M_\chi = 60$ GeV/c^2 and $\sigma_{SI} = 10^{-44}$ cm^2 (10 zeptobarns).

Figure 2.15: Effects of choice of target nucleus on spin-independent direct detection rates.

momentum transfers for heavier target nuclei, somewhat counteracting their A^2 sensitivity enhancement over light nuclides. As shown in Figure 2.15(b), for example, a target made of a heavy nuclide (*e.g.* Xe) must run at a low detector threshold to see the same event rate as a comparable mass of Ge. The event rate of a light nuclide (*e.g.* Si or Ar) experiment varies comparatively little with threshold.

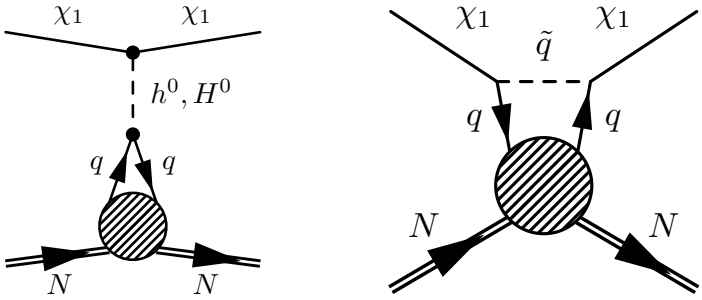


Figure 2.16: Major Feynman diagrams for spin-independent WIMP-nucleon scattering in supersymmetric models. Shaded blobs indicated strong processes within the nucleon.

For SUSY neutralinos, the dominant contributions to the WIMP-nucleon SI coupling usually come from the two classes of Feynman diagrams shown in Figure 2.16: t-channel exchange of a neutral Higgs boson and s-channel squark exchange. The former

is generally the dominant process, since squark-exchange diagrams are suppressed as $m_{\tilde{q}}^{-4}$ and accelerator limits constrain $m_{\tilde{q}} > 379$ GeV/c^2 under mild assumptions [29]. Note that the scattered quark in either process can be a strange quark or heavier flavor created by strong interactions in the gluon sea, not just an up or down quark. Strange quarks usually dominate Higgs exchange processes, in fact: up and down quarks are too light to couple significantly to the Higgs, while heavy flavors (c, b, t) are too rare in the nucleon to contribute substantially. WIMP-nucleon cross section computations thus depend critically upon the strange quark content of the nucleon, a relatively poorly-known quantity [145]. This also explains why $f_p \approx f_n$ in a wide range of models: Higgs-strange couplings are often the dominant part of scalar couplings, and the proton and neutron have similar strangeness contents.

Spin-dependent scattering

In the non-relativistic limit, axial-vector couplings (characterized by coefficients a_p and a_n) give amplitudes proportional to the inner product of the WIMP and nucleon spins. Since the interaction amplitude switches signs when the nucleon spin is flipped, a WIMP's spin-dependent (SD) interaction amplitudes with two nucleons of opposite spin will interfere *destructively* in the zero-momentum-transfer limit. Nucleons align into spin-singlet pairings within nuclei, so spin-dependent cross sections are dominated by unpaired nucleons and vanish entirely for spinless nuclides. This leads to very different constraints on experimental design: experiments targeting SD interactions generally use light odd-proton or odd-neutron nuclides to maximize the nuclear spin per unit mass, rather than the heavy nuclides preferred for SI-sensitive detectors.

At vanishing momentum transfer, the WIMP-nucleon spin-dependent interaction cross section is [146]

$$\sigma_{SD}^0 = \frac{32(J+1)}{\pi J} G_F^2 \mu_{\chi N}^2 \left[a_p \langle S_p \rangle + a_n \langle S_n \rangle\right]^2 , \tag{2.13}$$

where J is the nuclear spin and $\langle S_{p(n)} \rangle$ is the expectation value for the nucleus of the proton (neutron) spin. The spin expectation values must be obtained from detailed nuclear structure calculations [147, 148]. Such calculations also show that an odd-neutron nucleus may have a small but non-zero $\langle S_p \rangle$ (or vice versa) due to polarization effects within the nucleus.

In the case of spin-dependent scattering we *cannot* generally take $a_p \approx a_n$; their ratio may vary in sign and magnitude among neutralino models, for example [146]. This means that we must generally consider a three-parameter space: a_p, a_n, and M_χ. Furthermore, it means that the finite-momentum-transfer effects of the form factor cannot be factored out of the cross section in a model-independent fashion. For the extreme case of an SD WIMP that interacts solely with protons (neutrons), the form factor should depend on the distribution of proton (neutron) spin in the nucleus. The distributions of proton and neutron spin may be very different in a given nucleus, and so finite momentum effects may be very model-dependent. The preferred way to deal with this is to follow [149] by writing the WIMP-nucleus differential cross section in the form

$$\frac{d\sigma_{SD}}{dq^2} = \frac{8G_F^2}{(2J+1)v^2} S(q), \tag{2.14}$$

where v is the incident WIMP velocity and

$$S(q) \equiv a_0^2 S_{00}(q) + a_0 a_1 S_{01}(q) + a_1^2 S_{11}(q), \tag{2.15}$$

with $a_0 \equiv a_p + a_n$ and $a_1 \equiv a_p - a_n$. $S(q)$ encompasses the effects of finite momentum transfer, as well as values for the neutron and proton spin expectations.[1] There is no universal form of $S(q)$ - it must be computed separately for each nuclide using nuclear structure models [147, 148]. Figure 2.17 shows $S_{xy}(q)$ for ^{73}Ge used in this analysis, taken from [148].

In supersymmetric models, spin-dependent WIMP interactions are mediated by the two major diagram classes shown in Figure 2.18: t-channel Z^0 exchange and s-channel squark exchange. Since experimental limits demand that $M_{Z^0} \ll M_{\tilde{q}}$, Z^0 exchange generally dominates. Note that the tree level $\chi_1^0\chi_1^0 Z^0$ coupling shown in the diagram is only present for neutralinos with Higgsino components, since the Z^0 has no direct couplings to gauge bosons in the standard model.

In many models, spin-dependent WIMP-*nucleon* cross sections are often ~ 1000 times larger than analogous spin-independent cross sections. This is broadly true for Majorana WIMP candidates such as SUSY neutralinos, for which tree-level vector interactions

[1] There is a variation on this method advocated by some authors [46, 143] in which the scattering expressions are made to look more similar to those in the SI case by multiplying the above expression for σ_{SD} by a model-dependent form factor $F^2(q) \equiv S(q)/S(0)$. This is equivalent in principle, but the division by $S(0)$ may increase numerical errors in practice.

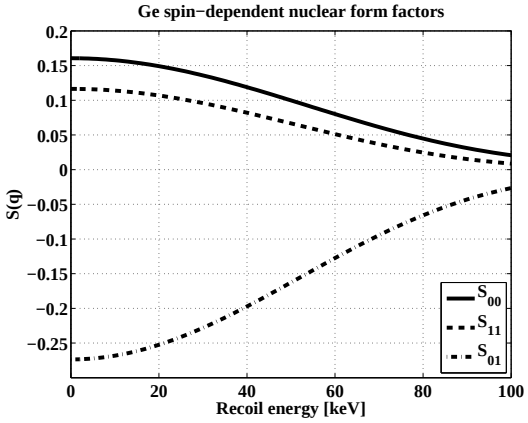


Figure 2.17: Spin-dependent WIMP-nucleus form factors for ^{73}Ge, using the parameterization computed in [148].

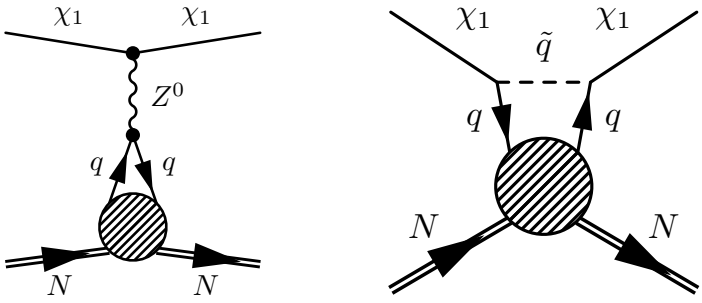


Figure 2.18: Major Feynman diagrams for spin-dependent WIMP-nucleon scattering. Shaded blobs indicate strong processes within the nucleon.

with the Z^0 are forbidden. Due to the great power of coherent enhancement, however, spin-independent WIMP-*nucleus* cross sections are generally much greater. Because of this, and because of the relative rarity of heavy spin-sensitive isotopes, spin-independent interactions are targeted by most leading direct searches. WIMP scattering is expected to be detected through spin-independent couplings first in nearly all SUSY models, but there are models in which scalar interactions are sufficiently suppressed to allow axial ones to dominate [150].

Halo model

In addition to the WIMP interaction cross sections described above, a model of the incident WIMP population is needed to determine the rate and spectrum of recoils expected in a given detector. Most experiments interpret their results using a standard halo model described by Lewin and Smith [143]. This model posits an isotropic Maxwellian WIMP velocity distribution ($f(v) = e^{-v^2/v_0^2}$) in the reset frame of the galactic halo, with characteristic velocity $v_0 = 230$ km/s in the solar neighborhood and truncated at a galactic escape velocity of $v_{esc} \approx 650$ km/s. The halo density is assumed to be $\rho = 0.3$ GeV/c^2/cm^3 near Earth. The Earth moves through this halo at a relative velocity v_E, the vector sum of the Earth's orbital velocity around the Sun and the Sun's orbital velocity around the galactic center. The Earth's net velocity through the halo thus varies sinusoidally during the course of its orbit, with approximate form

$$v_E \approx 244 + 15\cos(2\pi t)\text{km/s}, \tag{2.16}$$

where t is the time measured in years since the maximum velocity near June 2nd. This few-percent modulation in the incident velocity distribution is currently ignored by most experimental groups, since it can only be observed in an experiment which has detected thousands of WIMP-induced recoils. If we do acquire such large event samples, however, it will be an important signature of WIMP interactions (see below).

2.4.2 Backgrounds

Based on the discussion above, a leading direct detection experiment must be able identify a spectrum of $\mathcal{O}(10)$ keV nuclear recoils occurring less than once a month in each kilogram of target material. Unfortunately, this energy regime is awash with background events from natural radioactivity and cosmic rays. Direct WIMP detection is thus a low-background enterprise, demanding exquisite protection from these spurious events. The

ultimate sensitivity of a WIMP experiment is set by its rate of background events which are indistinguishable from WIMP candidates.

One can understand the important of background rates in a rare-event search by performing a simple thought experiment. Consider a very simplistic event-counting experiment that operates for an exposure of MT kilogram-days and expects to observe B background events in that exposure indistinguishable from WIMPs. This experiment can operate in three idealized modes, depending upon the magnitude of B:

1. **Background-free**: If $B \ll 1$ event, then any observed candidate events are evidence for WIMPs. If no candidate events are observed, the experiment can set a 90% confidence level Poisson upper limit on the WIMP recoil rate of $r = \frac{2.3}{MT}$ events per kilogram-day. As long as B remains much less than unity, the experiment's sensitivity improves in proportion to the exposure MT.

2. **Background subtraction**: If B is not negligible but is well-characterized from independent data (*i.e.* has negligible systematic errors), then we may subtract it from the observed count rate and look for an excess of events above that consistent with background. The accuracy with which this subtraction is performed is limited by Poisson statistics, giving an r.m.s. error on the subtraction of $\sigma_B = \sqrt{B}$. As we increase the exposure, we expect that $B \propto MT$ and thus $\sigma_B \propto \sqrt{MT}$. The number of excess candidate events needed to claim a detection (a 3σ or 5σ excess, for example) scales as σ_B, so the experiment's sensitivity grows only as $\sqrt{MT}$.

3. **Background-limited**: The systematic errors on the background count often grow in proportion to B, while the statistical errors grow only as $\sqrt{B}$. This happens because a systematic error is generally an uncertainty in the expected rate of the background per unit exposure; the statistical error, in contrast, is an uncertainty in determining that rate from the observed event count. Systematic errors therefore exceed statistical errors at sufficiently large exposures. At this point the sensitivity of the experiment no longer increases at all with increasing exposure, since the systematic error on the background rate constitutes an irreducible lower limit upon the rate of candidate events.

Due to these differences in power laws, it is extremely advantageous for a WIMP-detection experiment to operate as close to the background-free regime as possible. Any experimental

technology naturally progresses down this list as its exposure is increased, operating in a background-free mode for short exposures and reaching limiting sensitivity for sufficiently long exposures. Once an experiment moves down this list, only a change in the rejection or characterization of its background can move it back upward. Note that a simple event-counting experiment assuming no knowledge of its confounding backgrounds skips the background subtraction stage, as it has an effectively infinite systematic error upon its background rate.

Radioactive and cosmogenic background rates may be reduced by a variety of shielding and material-handling techniques, many of which are described in a review article by Heusser [151]. I discuss several of the techniques in use by CDMS in Chapter 4 and elsewhere in this work. Some rate of background events invariably remains, however, and the most competitive experiments generally pursue some scheme to separate the WIMP signal from these background events. These techniques take two general forms: statistical signatures of WIMP interactions or event-by-event identification of WIMP recoils. Below I briefly touch upon each of these in turn.

Statistical signatures

Even if individual WIMP events cannot be identified amidst a lingering background, the WIMP population may possess statistical signatures which can identify its presence. The two most such signatures discussed most prominently in the literature are the annual modulation in the WIMP recoil spectrum and the daily modulation of the incident WIMP direction. Both signatures take advantage of the extraterrestrial origin of the WIMP signal, and so should provide strong evidence that an observed signal is indeed related to dark matter. As statistical signatures, however, both require an enormous sample of WIMP recoils. If WIMPs eventually are detected, both will be key components of "WIMP astronomy": the determination of the local properties of the galactic halo through large studies of WIMP recoils.

As noted in Section 2.4.1, the Earth's motion about the Sun introduces an annual modulation in the velocity distribution of WIMPs incident upon a terrestrial detector. In the absence of large WIMP streams in the Earth's vicinity, Earth's net velocity through the galactic halo is greatest in June and lowest in December. A WIMP detector should thus observe a slightly higher rate of candidate events above its energy threshold in June

than in December, generally by a few percent [152, 153]. This signature has been pursued for more than a decade by the DAMA collaboration, which has monitored a very strong such modulation over 11 annual cycles. Other effects may also modulate the background rate with a similar period and phase, notably changes in the temperature of the readout electronics and modulations of the cosmogenic muon rate due to atmospheric conditions, but the collaboration has gone to great efforts to reduce these. It is worth noting that both of these effects correlate with temperature, so it would be interesting to conduct a similar experiment constructed in the southern hemisphere. A sufficiently sensitive experiment might also be able to detect a daily modulation in recoil energy spectra due to the velocity of the Earth's rotation [154].

WIMPs should also show a diurnal modulation in recoil energy direction with the Earth's daily rotation [155]. Due to the orientation of the Earth's axis with respect to our motion through the halo, this corresponds to a shift of $\sim 90^\circ$ in the mean recoil direction every 12 sidereal (not solar) hours. This effect should be very strong, but its detection is complicated by the enormous difficulty of measuring the direction of the short track produced by a low-energy nuclear recoil. Several collaborations – notably DRIFT, NewAge, and DMTPC – are currently pursuing this signature.

Event-by-event discrimination

Rather than looking for statistical signatures among a large sample of candidate events, we may instead try to identify individual WIMP-induced recoils on an event-by-event basis. A sufficiently accurate discrimination technique can reduce an experiment's background rate low enough to maintain operation in the zero-background regime, even at very large WIMP-search exposures. This requires a powerful particle identification technique that operates at low energies, but the benefits of low-background operation allow one to extract enormous sensitivity from relatively small detector exposures in this way.

Most modern direct-detection experiments use event-by-event discrimination techniques to identify nuclear recoils from WIMPs (or neutrons) amidst a far larger rate of electron recoils from radioactive decay and cosmogenic processes. Some experiments (notably bubble chambers and other phase-transition detectors) use the difference in dE/dx between electron- and nuclear-recoil tracks to render their detectors unresponsive to electron recoils, thus achieving a particularly simple sort of event-by-event discrimination. Discrimination is

more commonly accomplished by measuring each event in two (or more) distinct detection channels and using their ratio to identify the recoil type. The most common such channels are heat (phonons), ionization, and scintillation light. These channels differ enormously in the mean nuclear recoil energy needed to create the individual quanta: generally a few meV per phonon, ~ 10 eV per charge carrier, and ~ 100 eV per scintillation photon. The choice of detection channels carries tradeoffs in readout technology and experimental resolution.

The rest of this dissertation discusses the techniques for event-by-event discrimination used by the CDMS experiment, based upon a simultaneous measurement of ionization and out-of-equilibrium phonons. Chapter 8 returns briefly to some of the other discrimination techniques in use in the community.

2.5 Complementary contributions

Collider, indirect, and direct investigations of WIMPs provide three complementary probes of the nature of dark matter. Each of the three asks a fundamentally distinct question about the properties of the dark matter particle.

1. **Collider production**: *Do new colored or charged particles related to WIMP physics appear at accessible energies?* The production rate and detectability of WIMPs at colliders depends primarily upon the masses of the WIMP and related new particles: if the particles are too heavy, they cannot be produced.

2. **Indirect detection**: *Do the WIMP annihilation processes of the hot early universe still occur in today's colder universe?* The primary annihilation processes that set the relic density in some WIMP models are temperature-dependent, *e.g.* coannihilation or p-wave annihilation. The annihilation rates of such models are greatly reduced in the cold modern universe, making indirect detection less promising given equivalent halo models.

3. **Direct detection**: *Is there a relationship between WIMP annihilation and scattering cross sections?* The "crossing symmetry" argument posits that the dominant WIMP annihilation and scattering processes have comparable amplitudes. Direct detection is less promising if this relationship is broken, as when annihilation occurs on resonance.

Different WIMP models answer these questions differently, and thus have different detection prospects. Some models generate robust signals for all three search strategies, others in only a subset. These three investigations can also provide complementary information about the same WIMP, opening up new opportunities for elucidating its nature.

2.5.1 Complementarity in the CMSSM

One can get a flavor for the interplay between these three questions by considering a representative family of WIMP models. Perhaps the best-studied such family is the Constrained Minimal Supersymmetric Standard Model (CMSSM), also sometimes called "minimal supergravity" (mSUGRA). This set of models imposes strong constraints inspired by Grand Unified Theories (GUTs), reducing the number of free parameters from the MSSM's > 100 to a mere five: the unified mass m_0 of all scalars, the unified mass $m_{1/2}$ of all fermions, the unified trilinear scalar coupling A_0, the ratio $\tan\beta$ of Higgs vacuum expectation values, and the sign of the Higgs mass parameter μ. This far smaller parameter space is manageable enough to permit numerical investigation (*e.g.* [74, 156, 87, 157] and many others) while remaining complex enough to illustrate a variety of different phenomenologies.

Detection prospects

WIMP detection prospects are very model-dependent, varying enormously over the allowed CMSSM parameter space. We can nonetheless make a few qualitative observations about each of these annihilation mechanisms based upon the considerations of Section 2.5:

1. **Collider production:** The bulk region is generally a favorable target at colliders, due to the wealth of light charged and colored sfermions. Prospects are much poorer in the focus point region and variable in the other two domains.

2. **Indirect detection:** The bulk and coannihilation regions involve temperature-dependent annihilation mechanisms, so such models are expected to give poor signals in indirect searches. The p-wave annihilation processes of the bulk region shut down at low WIMP velocities, while coannihilation processes suffer from a lack of NLSPs in the modern era. The focus point and funnel regions, in contrast, should give relatively robust annihilation signatures.

3. **Direct detection:** The argument of crossing symmetry fails in the funnel and the coannihilation tail, since in both cases the WIMP benefits from a "spectral coincidence" which greatly enhances its annihilation rate without altering its other couplings. Direct detection is more promising in the bulk and focus point regions, where WIMP annihilation is representative of other WIMP interactions in magnitude. Note that the highest-mass ($M_\chi \gtrsim \text{TeV/c}^2$) region of the focus point is dominated by chargino coannihilation and so is also less promising. Figure 2.19 illustrates the detectability of these classes of model using representative scans of CMSSM parameter space using DarkSUSY.

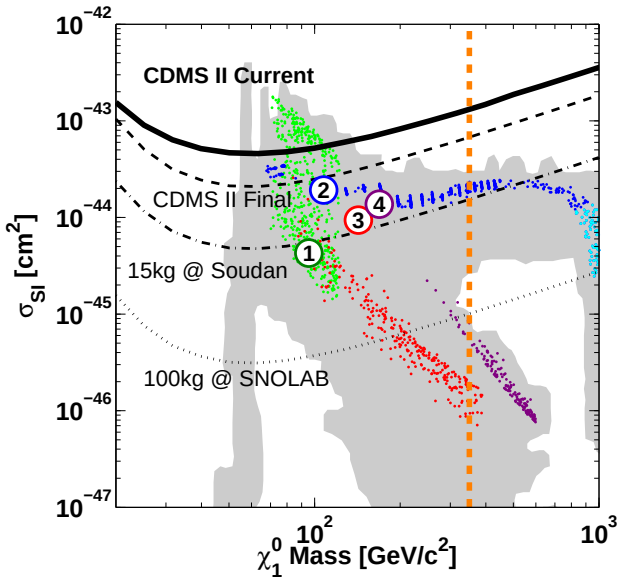


Figure 2.19: Masses and spin-indepedent WIMP-nucleon cross sections within the CMSSM. The horizontal curves represent upper limits from this work and from proposed extensions to CDMS technology. The vertical dotted line represents the approximate upper limit of the LHC's reach in neutralino mass. The filled (*gray*) region is a broad scan of CMSSM parameter space by Baltz and Gondolo [74], while the colored points indicate scans of this parameter space by the author using DarkSUSY. The colors represent representative portions of the qualitative regions described in Section 2.1.2: bulk (*green*), focus point (*blue*), Higgs funnel (*purple*), and coannihilation (*red*). The highest-mass portion of the focus point, in which chargino coannihilation becomes significant, is colored separately (*cyan*). Numbered points indicate the four Linear Collider Cosmology (LCC) benchmark models from [87].

This variation among models allows for many narratives of discovery. Generally speaking, accelerators are mass-limited while astrophysical searches are limited by the mag-

nitudes of various interaction cross sections. If a model in the bulk region were the true description of the universe, for example, we can look forward to a wealth of collider information and a signal in current or near-future direct searches. If the true model is within the focus point, in contrast, collider prospects are dim but both indirect and direct detection signals may be just around the corner.

The CMSSM scans from Figure 2.19 can also be used to compare the reaches of direct and indirect probes of spin-independent and spin-dependent WIMP interactions. Figure 2.20 plots the same CMSSM theory points in the $\sigma_{SI} - \sigma_{SD}$ plane, overlaid with approximate constraints from a leading direct search, a leading indirect neutrino search, and a projected large direct detection experiment. Note that the spin-dependent sensitivities of direct and indirect searches are currently comparable, but the spin-independent sensitivity of the direct searches is much closer to the bulk of CMSSM models.

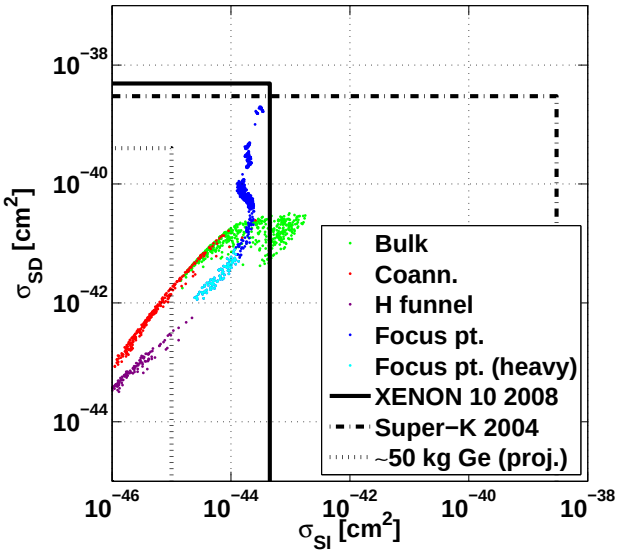
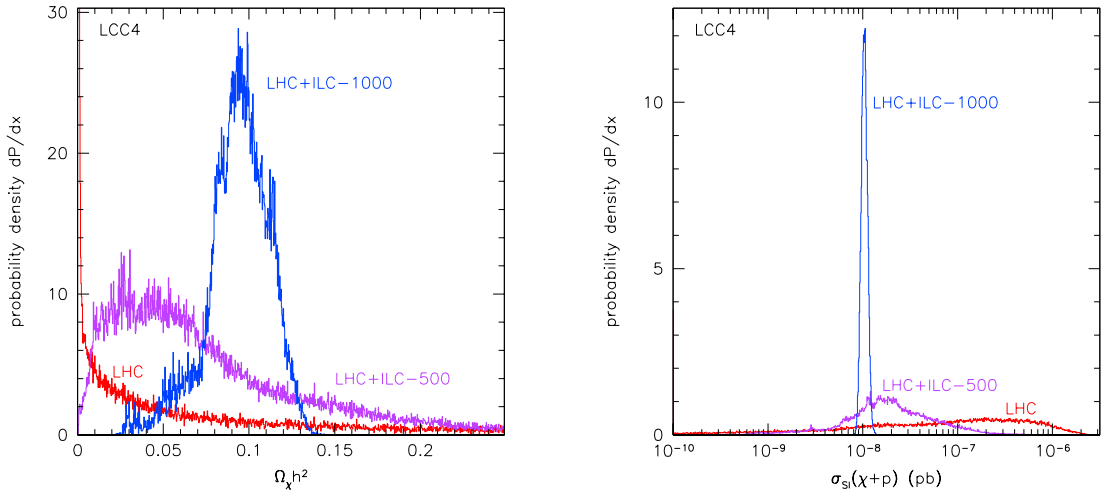


Figure 2.20: Comparison of spin-independent and spin-dependent WIMP interaction cross sections and experimental sensitivities. Colored points represent a selection of CMSSM models from the regions described in the text, following the color conventions of Figure 2.19. Lines represent approximate minimum upper limits on spin-independent and spin-dependent WIMP-nucleon scattering cross sections from various leading experiments, under the assumption that $a_p \approx a_n$. Boxes represent recent results from a leading direct search (XENON 10 [158, 159]), a leading indirect neutrino search (Super-K [125]), and the projected reach of a hypothetical large Ge detector.

Complementary information

Since each experimental search strategy asks different questions about the nature of dark matter, the combination of multiple signals can be more informative than any one alone. Though colliders can measure many aspects of the low-energy supersymmetric spectrum, for example, they cannot directly identify any of the observed (or unobserved) particles as a dark matter constituent. We cannot merely assume *a priori* that the LSP constitutes the entirety of the dark matter, and collider data alone may not be able to reconstruct the LSP's relic density if it depends upon the masses and couplings of inaccessible particles. Only through a consistent combination of collider and astrophysical data can we confirm the LSP's identity as dark matter.



(a) Neutralino relic density

(b) WIMP-nucleon spin-independent cross section

Figure 2.21: Projected collider constraints possible from the LHC and two versions of the ILC (500 or 1000 GeV center of mass energy) for the model LCC4, a stau-coannihilation benchmark model described in [87]. Plots taken from [87].

Astrophysical data can break degeneracies among families of models with equivalent collider signatures. Baltz *et al.* [87] have considered the complementary nature of collider data and astrophysical WIMP searches in great detail for a set of four "benchmark" points, one from each of these regions. Figure 2.21 illustrates an example of degeneracy-breaking for LCC4, a benchmark point in the coannihilation tail. The LHC cannot measure enough of the supersymmetric spectrum to accurately determine the relic density, as shown

in Figure 2.21(a). In particular, the LHC alone cannot exclude models with a wino-like neutralino, indicated by a peak at very low relic density. This range of models also predicts a very wide range of values for the direct detection cross section (Figure 2.21(b)), controlled mostly by variations in the mass of the heavy neutral Higgs H^0. Even an order-of-magnitude measurement of the direct detection cross section would be sufficient to eliminate the wino-like solutions and provide evidence for a bino-like neutralino with a relic density near that observed. A similar constraint could not be achieved at colliders alone without data from a high-energy ILC.

Direct and indirect detection can also provide independent information on the WIMP mass, a value which may be poorly constrained at a hadron collider. A gamma ray line from $\chi\chi \to \gamma\gamma$, for example, can provide a very precise measurement of the WIMP mass which may be used to interpret collider data. Measurements of the WIMP mass with direct detection are also possible [160, 143, 161, 162], though hundreds of observed events can be required for precise measurements. Even a weak mass constraint can nonetheless provide a useful consistency check, linking astrophysical observations to collider measurements.

Chapter 3

ZIP detectors

As described in the previous Chapter, the direct detection of dark matter presents enormous technical challenges. A competitive experiment must be able to identify a handful of WIMP-induced scattering events amongst a far larger population of background interactions. This identification technique must perform effectively with keV-scale energy depositions in kilogram-scale detectors, all while maintaining stable performance over months of live time. The heart of any direct detection experiment is its answer to these challenges: its detector technology.

CDMS II's solution to the challenges of direct detection is the Z-sensitive Ionization and Phonon (ZIP) detector, a low-temperature semiconductor device for simultaneously measuring the charge carriers and phonons released by a particle impact. Taken together, the phonon and ionization signals from each scattering event can be used to measure energy deposition with a resolution better than 500 eV (rms) and to discriminate nuclear recoils from electromagnetic backgrounds at energies well below 5 keV. Furthermore, the detailed timing and distribution properties of the phonon signal identify an event's location within the detector. By combining this wealth of information about each particle event, these detectors provide the sensitivity and discrimination power needed to maintain a leading search for WIMPs.

3.1 Physical characteristics

Each ZIP (pictured in Figure 3.1) is a cylindrical semiconductor crystal 7.6 cm in diameter and 1 cm thick. The substrate material is Ge or Si, prepared with low impurity

Figure 3.1: A Z-dependent Ionization and Phonon (ZIP) detector, showing the photolithographic patterning of the phonon sensors.

and dislocation concentrations ($\sim 10^{11}$ impurities/cm^3 and < 5000 dislocations/cm^2 for Ge) to ensure good charge transport. The cylinder axis of each detector is oriented along a $\langle 100 \rangle$ crystal axis, while a small flat region along the crystal edge indicates a $\langle 011 \rangle$ direction. The edge of each detector also has larger flat regions aligned with the $\langle 010 \rangle$ and $\langle 001 \rangle$ directions to aid in alignment. Most detectors in this data run were made from n-type boules, with the exception of 8 p-type Ge ZIPs in Towers 3, 4, and 5.

Due to its larger larger atomic mass ($A_{Ge} = 72.64$, $A_{Si} = 28.09$), a Ge atom has a larger cross section for coherent WIMP-nuclues interactions than a Si atom. For all but the lightest WIMP masses, Ge is thus expected to be more sensitive to spin-independent WIMP interactions than an equal mass of Si. Each Ge ZIP (250g) is also substantially more massive than its Si neighbors (100g). The use of a mixture of Ge and Si detectors is thus a hedge against a possible neutron background: a comparison of nuclear recoil rates in both detector substrates can statistically differentiate a WIMP signal from a neutron background. This scheme was used to great effect in early runs at the Stanford Underground Facility [163], but should be unnecessary at deeper sites.

Each of the two detector faces is photolithographically patterned with sensor wiring, as shown schematically in Figure 3.2. The bottom face (as detectors are commonly oriented at Soudan) is divided into two concentric ionization electrodes: an inner electrode ("Qinner") covering $\sim 85\%$ of the detector surface and an outer ring 3 mm in width ("Qouter"). The top face of each detector is occupied by four phonon sensors, each

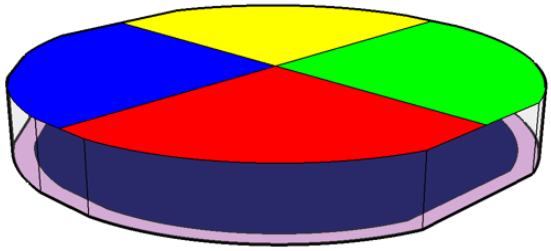


Figure 3.2: Schematic sensor configuration of a ZIP detector, showing four phonon sensor quadrants on the top face and two concentric charge electrodes on the bottom. Detector edge flats are also shown.

composed of 1036 tungsten transition-edge sensors (TESs) wired in parallel and fed by an array of aluminum quasiparticle traps. The charge and phonon sensors are described in more detail in Sections 3.2 and 3.3, respectively.

3.2 Ionization

The ionization readout is perhaps the most familiar element of the ZIP detector. The basic principle is essentially that of a standard Si particle tracker or high-purity Ge detector [164]: electrons and holes generated along a particle track are drifted to collection electrodes using an electric field. The very unusual operating regime of a ZIP – millikelvin temperatures and V/cm electric fields – leads to significant subtleties, however. Below I review the fundamentals of this process, referring the reader to the dissertations of Tom Shutt [165] and Sunil Golwala [166] for further details.

3.2.1 Carrier generation

A particle interacting in a semiconductor lattice can transfer a portion of its deposited energy to the electronic system of the crystal. If the energy deposited is much greater than the semiconductor's band gap (generally $\sim O(1)$ eV), the particle can liberate valence electrons from their bound states into the conduction band. These primary electrons may be sufficiently high in momentum to liberate other electrons, producing a cascade of charge carriers as the energy of the primaries dissipates into the crystal. The process culminates in a population of lower-momentum electrons and holes in the vicinity of the particle track.

Though the band gaps of Ge and Si are 0.743 eV and 1.12 eV at 0 K, respectively, the average deposited energy required to generate an electron-hold pair is substantially greater: 3.0 eV in Ge (3.8 eV in Si) in the millikelvin regime. In other words, a 10 keV photon fully absorbed in Ge generates $10^4/3 \approx 3000$ electron-hole pairs. Roughly 25-30% of the deposited energy is thus consumed in driving charge carriers across the band gap, while the remainder is dissipated in the form of phonons. This partitioning originates in the need to conserve momentum through phonon emission during each electron scattering event, reducing the amount of energy that can be used to create carriers [167].

3.2.2 Charge collection

This initial cascade produces a gas of hot electrons and holes along convoluted tracks around the initial particle impact. These carriers may initially be protected from the influence of the surrounding drift field by self-shielding and propagate quasi-diffusively. If left to themselves, these electrons and holes will eventually either recombine with one another across the gap or fall into localized energy states within the gap ("traps"). Either way, the gap energy contained in these carriers is released to the phonon system.

CDMS uses an applied electric field to draw these charge carriers away from the particle track before they can recombine or diffuse to a trap. A small voltage bias (-3V for Ge ZIPs in the current experiment, -4V for Si) is applied to the ionization electrodes with respect to the phonon sensors on the opposite. Since all charge carriers are frozen out at millikelvin temperatures, no "depletion zone" exists and the resulting drift field is nearly uniform. These carriers induce corresponding image charges on the conducting electrodes as they drift across the crystal which may be detected with a sensitive charge amplifier. Since a drifting hole is equivalent to an electron drifting along the same track in the opposite direction, a fully-collected electron-hole pair produces the same image charge expected from a single electron drifting all the way across the detector. In the limit of perfect charge collection and a fast charge amplifier, the total image charge collected on the ionization electrodes is equal in magnitude to that generated by the particle interaction, independent of the interaction's depth within the crystal.

If a carrier drifts only partway across the crystal, the observed charge signal will be reduced. The efficiency of the ionization readout channel depends on how likely carriers are to cross the crystal without recombining or becoming trapped. Both processes are

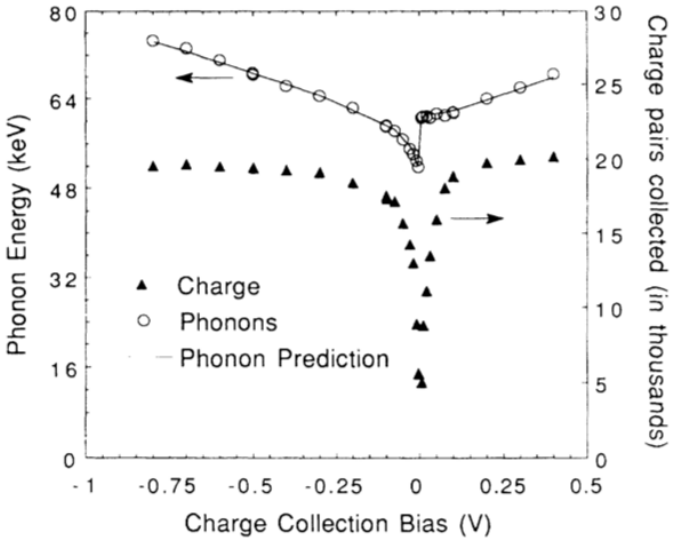


Figure 3.3: Charge collection efficiency for 59.5-keV photons as a function of charge bias, taken from [168]. These data were taken with a Ge device 1 cm thick, so equivalent voltages on current ZIPs would produce equivalent fields. Observed phonon amplitudes (*circles*) match the predictions (*curve*) from Neganov-Luke production (see text).

more likely at very low electric fields, which allow more opportunity for carrier diffusion before reaching the electrodes. The field-dependence of charge collection in Ge at millikelvin temperatures was studied by Shutt *et al.* [168], culminating in the "seagull plot" shown in Figure 3.3. At the low trap concentrations achieved by these authors, complete charge collection is possible in Ge with applied fields as low as $\sim$ 200 mV/cm, well below those applied in this work.

Charge carriers can still be lost during drift if the crystal has too high a density of trapping centers. The trapping rate depends not only upon the density of trapping centers but upon their character; neutral impurity atoms and defects are expected to have trapping cross sections orders of magnitude lower than those of charged centers. For this reason, CDMS detectors are manufactured from highly pure substrate materials. Even this is not enough, however: donors and acceptors are ionized in equilibrium at low temperatures, providing enough charged traps to thwart charge collection even at these purities.

Fortunately, these and other charged trapping centers can be neutralized by exposure to a strong radioactive source or to infrared LED light [168, 165]. These generate a large population of free charge carriers, and if no electric field is present these will quickly recombine. A small fraction of these carriers, however, will diffuse to a charged trapping

center and bind to it, neutralizing the center and reducing its trapping cross section. This neutralized configuration is not the equilibrium ground state, but at millikelvin temperatures it is stable over practical time scales.

The charge contacts themselves take the form of Al grids covering the two detector faces, each separated from the detector bulk by a $\sim$ 40 nm layer of amorphous Si. The top electrode is the phonon sensor grid described in Section 3.3, held at electrical ground by the low-impedance phonon wiring and covering 32% of the detector face. The bottom electrode is an Al grid covering 16% of the face and connected to a low-noise charge amplifier (described below). Grid coverage is deliberately kept low to avoid unnecessary phonon absorption far from the TES sensors. The amorphous Si layer greatly improves charge collection for near-surface events by blocking the back-diffusion of carriers into the nearby electrode [169]; surface events are discussed further in Section 3.5.

3.2.3 The charge amplifier

Figure 3.4 shows a simplified schematic of the ZIP charge amplifier. Charge collected across the detector crystal quickly transfers to the feedback capacitor C_{fb}, producing a voltage spike at V_{out}. This pulse decays back to quiescence as this capacitor drains through the feedback resistor, producing an exponential falling edge to the voltage pulse. In a standard ZIP amplifier, $R_{fb} = 40$ MΩ and $C_{fb} = 1$ pF (the parasitic capacitance of the resistor), giving a falltime of $\tau_{fall} = 40\mu$s. Since the actual charge collection process takes less than a microsecond, the ionization pulse shape is set solely by the electronics; all charge pulses have the same shape, up to possible effects from electronics saturation. A faster charge amplifier could measure the sub-μs structure of the charge pulse rising edge [170], but it is difficult to attain sufficient bandwidth with our existing high-capacitance wiring.

The theoretical noise performance of this charge amplifier is set by the noise of the amplifier's first-stage JFET, the Johnson noise of the feedback and bias resistors, and the total capacitance at the amplifier input. The JFET and resistors are dissipative circuit elements, and so add temperature-dependent noise fluctuations to the amplifier. The input capacitance is non-dissipative and so generates no noise itself, but it does determine the gain of the JFET noise; for total input capacitance C_i, the RMS voltage noise e_{FET} at the FET input appears equivalent to a charge noise of $C_i e_{FET}$ at the amplifier input, rolled off at low frequencies by current flow through the bias and feedback resistors. There are also

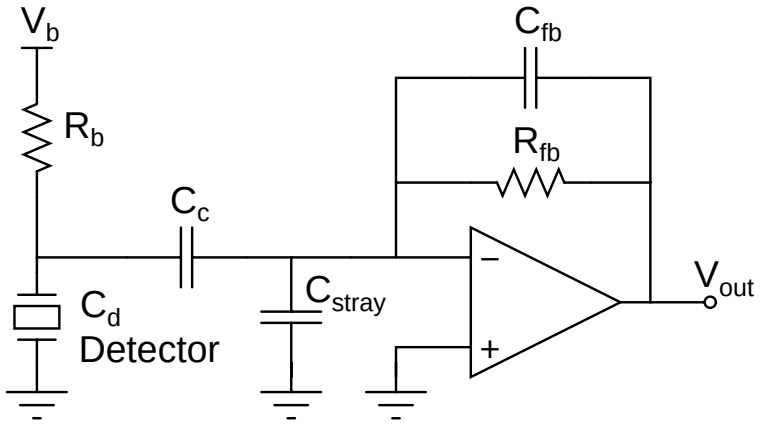


Figure 3.4: Simplified schematic of the CDMS charge amplifier. Currently, $R_b = R_{fb} = 40$ MΩ, $C_{fb} = 1$ pF, $C_d = 93$ pF for the inner electrode (36 pF for the outer), $C_{stray} \approx 75$ pF, and $C_c = 300$ pF. Capacitance values taken from [171].

current noise contributions from the FET input (i_{FET}), the detector leakage current (i_d), and any microphonic effects on the detector or its wiring (i_μ). Shutt [165] has considered the performance of this amplifier in detail, deriving the following expression for its output voltage noise as a function of frequency f:

$$e_o^2 = |A(f)|^2 \Bigg\{ e_{FET}^2 \left[(C_d + C_{fb} + C_{stray})^2 (2\pi f)^2 + \left(\frac{1}{R_{fb}} + \frac{1}{R_b} \right)^2 \right] + 4k_B T \left(\frac{1}{R_{fb}} + \frac{1}{R_b} \right) + i_{FET}^2 + i_d^2 + i_\mu^2 \Bigg\}. \tag{3.1}$$

In this expression, T is the resistor temperature and $A(f)$ is the amplifier transresistance

$$A(f) = \frac{R_{fb}}{1 + 2\pi i f R_{fb} C_{fb}}. \tag{3.2}$$

The resistors are mounted at the detector base temperature, giving $T \sim 40$ mK in the current data runs. The voltage noise at the JFET input is ~ 0.5 nV/$\sqrt{\text{Hz}}$, as measured by Mandic [1]. The FET and detector current noises are believed to be negligible at these temperatures, but microphonic currents contribute significantly to ionization noise at very low frequencies. Furthermore, some detectors show resonant microphonic peaks at a few kHz, possibly due to reduced tension in the tower wiring.

Figure 3.5 shows a representative charge noise spectrum from the Soudan installation, overlaid with contributions from the model described above. A set of high-frequency peaks is clearly visible above the noise baseline due to various forms of electronic pickup. A

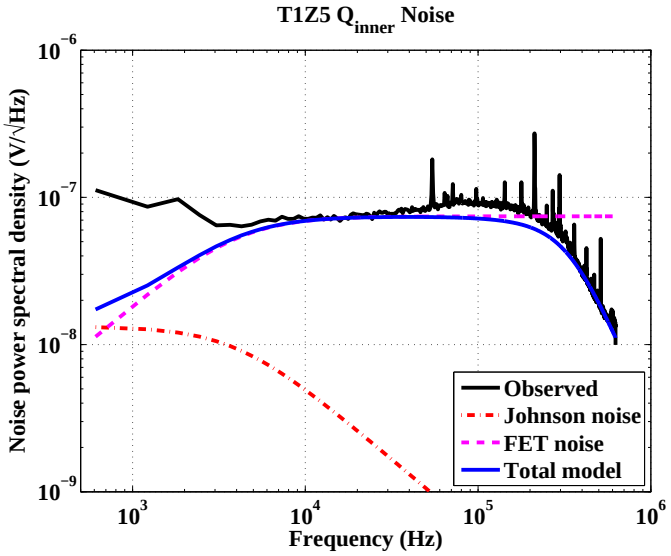


Figure 3.5: Observed ionization noise spectrum for T1Z5 in Soudan Run 123, referred to the FET gate and overlaid with model predictions following Equation 3.1. The FET noise is normalized to 0.44 nV/$\sqrt{\text{Hz}}$ to match the data. The noise spectrum is also rolled off at high frequencies by the antialiasing filter and other poles of the warm electronics. A charge pulse (signal) has a power spectrum matching the shape of the Johnson noise contribution.

substantial excess noise is also apparent at low frequencies; this is ascribed to a combination of electronic pickup (including 60-Hz harmonics), microphonic pickup of mechanical vibrations, and $1/f$ noise contributions from the JFET. The power spectrum of a charge pulse matches that of the Johnson noise in Figure 3.5; note that the JFET noise rolls off within the signal bandwidth, leaving low-frequency noise as the dominant contribution to experimental resolution. CDMS achieves an RMS resolution of $\sigma_Q \approx\sim 250eV$ on each charge channel through use of an optimal filter pulse reconstruction (see Appendix A), though a few channels with great microphonic sensitivity perform worse.

3.2.4 Charge reconstruction

The CDMS data reduction package (DarkPipe) converts each event's digitized ionization traces into more physical quantities – energy, start time, etc. – for later analysis. This package is discussed further in Chapter 6. I briefly outline the major reconstruction algorithms used for the ionization channels.

Optimal filtering

The primary charge-reconstruction algorithm used by the CDMS collaboration is optimal filtering. This technique takes advantage of the special characteristics of ZIP ionization signals: each pulse from a given ZIP detector is of essentially fixed shape, with time constants determined by the electronics, and the noise is predominantly gaussian in nature. This makes them ideal candidates for optimal filtering: a frequency-domain fit of a fixed template to the observed trace, accounting for variations in the noise power with frequency. This algorithm is described futher in Appendix A of this work and in Appendix B of Sunil Golwala's dissertation [166].

In this analysis, we fit the two ionization pulses from each event (inner and outer electrodes) simultaneously to a set of templates to find the charge deposited in each channel. Figure 3.6 shows a representative set of ionization templates from this run. The four template traces represent the pulse shapes of the two channels and the $\sim 6\%$ capacitive crosstalk between them. To account for tiny variations between channels, the templates are generated by averaging the traces from a selection of well-collected events from each detector. For well-designed templates, the χ^2 returned by the optimal filter algorithm is an excellent goodness-of-fit parameter with which to reject malformed pulses.

Time-domain fits

DarkPipe also incorporates a simple time-domain fit to the two ionization pulses. This "F5" fit has substantially worse noise performance than the optimal filter but (unlike the optimal filter) can easily be applied to large ionization pulses which saturate the digitizers for some portion of their duration. This algorithm is used primarily for background studies at MeV-energies and will not be discussed further in this work.

3.2.5 Ionization calibration

The ionization channels of each detector are calibrated using extensive (several million events) calibration sets with ^{133}Ba sources. The primary benchmark is a 356-keV spectral line, easily visible in Ge detectors.

The ionization sensor calibration of a Ge detector has 4 major components:

1. A correction for **residual crosstalk** between the two ionization channels. We use two linear correction factors, generally $< 1\%$ in magnitude.

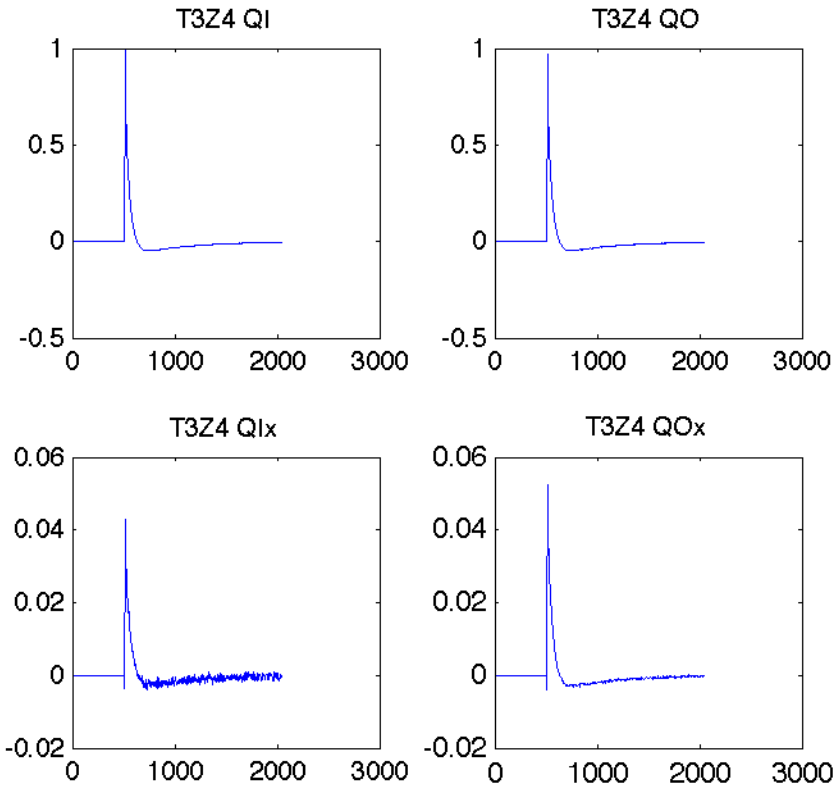


Figure 3.6: Run 123 ionization templates for detector T3Z4. The QI and QO templates represent the pulses generated in these channels by events under these same electrodes, while the QIx (QOx) template represents the crosstalk signal observed in QI (QO) due to a pulse of unit amplitude in the opposite channel.

2. A correction for the **position dependence** of charge collection across the crystal. Figure 3.7 shows the significant ($\sim 5\%$) dependence of the ionization amplitude of the 356 keV ^{133}Ba line as a function of detector position, and particularly of `ydel`. This is not understood quantitatively, but may result from the y-dependence of TES ion implantation or neutralization effects due to varying distance from the LEDs. Beginning with this analysis, we correct `qi` using polynomials fitted to the centroid of the 356-keV line as functions of `xdel` and `ydel`, as shown in Figure 3.7. This correction improves the ionization resolution significantly at high pulse energies.

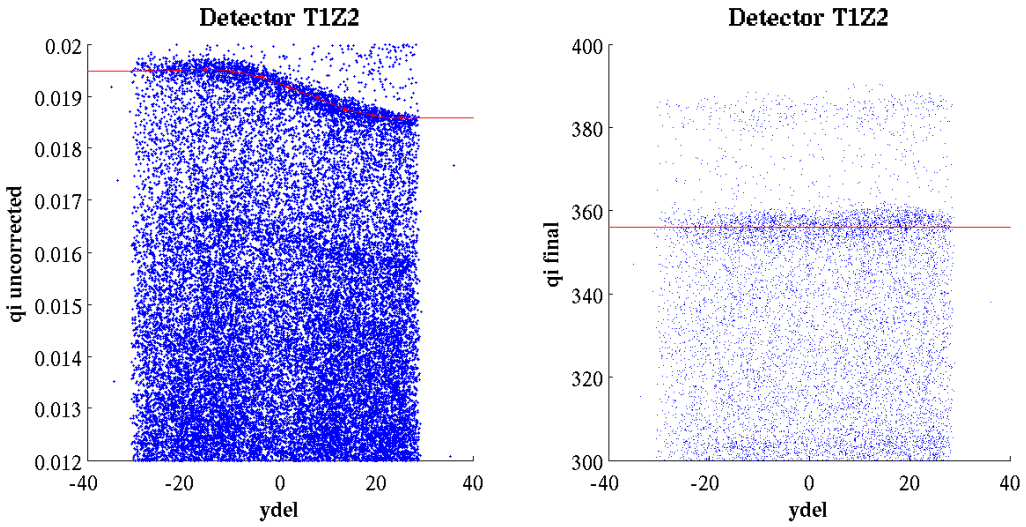
3. Overall calibration of the inner electrode amplitude using ^{133}Ba spectral lines, as shown in Figure 3.8.

4. Cross calibration of the inner and outer electrode amplitudes using shared events. The calibration of Qouter generally differs from that of Qinner by a few percent. Very few high-energy events deposit their energy solely in Qouter, so we correct for this discrepancy using the small population of 356 keV Ba events shared between the two electrodes as shown in Figure 3.9. Due to low statistics this calibration is less precise than that for Qinner, but since `qo` does not directly figure into the WIMP-search analysis this is more than adequate.

The energetic ^{133}Ba lines are not always visible in Si ZIPs due to their lower stopping power, necessitating more indirect calibration methods. We calibration `qi` for Si detectors using 356 keV events shared with neighboring ZIPs, in direct analogy to the `qi-qo` calibration in Figure 3.9. This is most effective if the neighboring detector is a working Ge ZIP but works adequately even for T2Z1, the only Si detector with no Ge neighbor. In the absence of a charge line we take the Qouter and Qinner calibrations of Si detectors to be equal and do not perform a charge position correction.

3.3 Athermal phonons

ZIP detectors also measure the initial wave of phonons released by a particle interaction, before the detector mass has had time to come to thermal equilibrium. These out-of-equilibrium phonons are often called "athermal phonons." In addition to indicating the particle energy, this prompt phonon signal retains a great deal of information about



(a) `qi` versus `ydel` before correction

(b) `qi` versus `ydel` after `ydel` correction

(c) `qi` versus `xdel` after `ydel` correction

(d) `qi` versus `xdel` after both corrections

Figure 3.7: Example of ionization position correction for a Ge ZIP (T1Z2). Red lines in each figure show the fitted contour of the 356 keV ^{133}Ba line, both before and after polynomial correction. Note that the overall energy calibration is incorporated between Figures 3.7(a) and 3.7(b). Figures by K. Sundqvist.

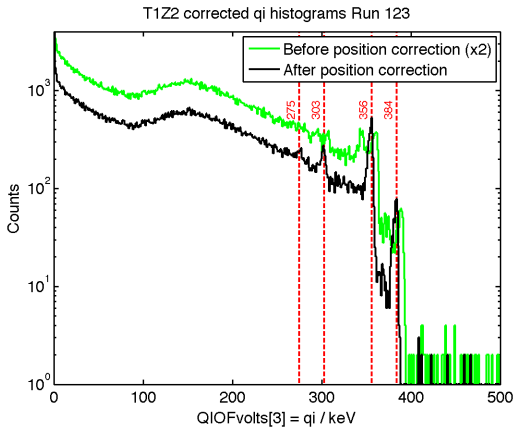


Figure 3.8: Histogram of `qi` for good ^{133}Ba calibration events, with several spectral lines indicated in red. The green histogram in the background illustrates the spectrum before charge position correction.

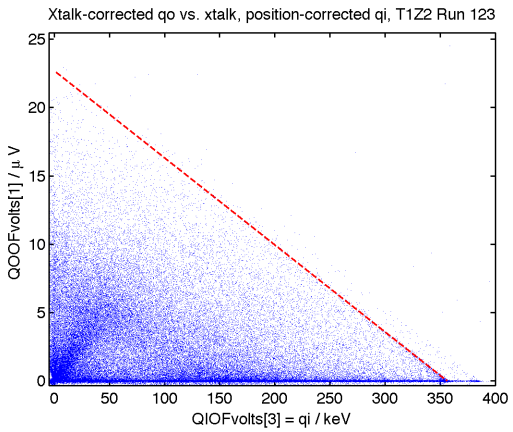


Figure 3.9: `qi` and `qo` amplitudes for ^{133}Ba events in T1Z2, showing events shared between the electrodes. The red line indicates `qi` + `qo` = 356 keV.

the location, timing, and character of the recoil itself. This additional information (lost in an equilibrium phonon detector) is a vital contributor to CDMS's event reconstruction and excellent background rejection.

3.3.1 Phonon generation

The phonons precipitated by a particle impact originate from three distinct generation mechanisms: primary, carrier relaxation, and Luke phonons.

Primary phonons

Most of the energy deposited by a particle impact is dissipated into the phonon system as the scattered nuclei and electrons return to rest. A recoil of energy E_r deposits energy $E_{prim} = E_r - N_Q E_{gap}$ as *primary phonons*, where N_Q is the number of charge carriers produced and E_{gap} is the energy gap of the semiconductor. This component, comprising both high-energy optical phonons and lower-energy acoustic phonons, carries information about the position, timing, and other characteristics of the particle event.

Carrier relaxation phonons

The remaining energy of the initial particle impact drives electrons across the semiconductor band gap to produce charge carriers. An energy E_{gap} is stored in each electron-hole pair. The carriers eventually restore this energy to the phonon system by relaxing to the Fermi level at the detector surface, producing $E_{relax} = N_Q E_{gap}$ of *recombination phonons*. Shutt *et al.* have shown [168] that these phonons contribute to the slow thermal signal measured by an NTD attached to a Ge detector. It has not been shown explicitly that relaxation phonons contribute to a ZIP's faster athermal phonon signal, but the voltage-invariance of ZIP response to surface events and nuclear recoils suggests that they do. In this work, I assume that relaxation phonons are collected at the same rate as primary phonons.

Luke phonons

The imposed electric field does work eV on each electron-hole pair drawn to the detector surfaces. This energy is dissipated to the crystal as phonon radiation [172], analogously to the emission of Cerenkov radiation by rapidly-moving charges in dielectrics. This

mechanism, proposed by Neganov and Trofimov [173] and by Luke[174], contributes additional energy to the crystal beyond that in the initial interaction. These Neganov-Trofimov-Luke phonons (commonly known to American researchers simply as "Luke phonons") contribute an energy equal to the field's work on each charge:

$$E_{Luke} = \sum_i q_i \int_{\text{path}_i} \vec{E} \cdot d\vec{x}, \tag{3.3}$$

where the sum is over each charge q_i and the integral is over the path taken by that charge. Since the electric field is essentially uniform within the crystal and the carriers possess a charge of magnitude e, we can simplify this integral to a sum:

$$E_{Luke} = eV_b \sum_i \frac{d_i}{d}. \tag{3.4}$$

In the above expression V_b is the detector voltage bias, d is the detector thickness, and d_i is the distance traveled by charge i. In the case of full charge collection, each electron-hole pair covers a total distance d and this expression becomes $E_{Luke} = eV_bN_Q$.

Assuming equal collection rates of all three of these populations, the total phonon signal is thus given by

$$E_P = E_r + eV_b \sum_i \frac{d_i}{d}. \tag{3.5}$$

3.3.2 Phonon transport

Phonon propagation across a semiconductor crystal can be impeded by several processes [175, 176]: phonon-carrier scattering, phonon-phonon scattering, phonon-impurity scattering, or anharmonic phonon decay. The first two processes are essentially insignificant at millikelvin temperatures, as no free carriers or energetic phonons are available outside of the initial interaction "fireball." Impurity scattering is elastic scattering from atoms of unusual mass, predominantly isotopic impurities in our high-purity crystals; the cross section for isotope scattering scales as $\sigma \propto \nu^{-4}$, where v is the phonon frequency. Anharmonic phonon decay is the spontaneous decay of a phonon to two, each generally carrying away approximately half of the original energy on average; the anharmonic decay lifetime scales as $\tau \propto \nu^{-5}$.

Since both processes described above are strongly frequency-dependent, the mean free path of low-frequency phonons is much longer than that of their higher-frequency brethren. The vast majority of primary phonons are generated at multi-THz frequencies

and thus have sub-mm mean free paths. Anharmonic decay rapidly reduces these to a "bottleneck" frequency near 1.6 THz at which isotope scattering begins to dominate. After a few microseconds of quasi-diffusive propagation, these phonons decay sufficiently to bring their mean free path above the crystal dimension. These "ballistic" phonons can finally travel to the detector surface without significant further scattering.

Since phonons must reach the detector surface to be detected, the observed phonon signals in ZIP detectors are dominated by ballistic phonons. Primary phonon detection is thus delayed by the process described above: quasi-diffusion followed by ballistic propagation. We expect most Luke phonons to be generated at ballistic frequencies [172]. Recombination phonons are generated at energies of a few tenths of an electron volt but down-convert rapidly through interaction with the metal films at the detector surface. This down-conversion mechanism has important implications for particle interactions near the detector surface, as discussed in Section 3.6.

3.3.3 Phonon detection

Configuration

The top face of each ZIP detector is patterned with four phonon sensor arrays, each filling a quadrant of the detector's surface. Each of the four arrays – identified as sensors A, B, C, and D – consists of 1036 active phonon-sensing elements, wired in parallel to act in concert as a single sensor. Figure 3.10 depicts the patterning of phonon sensors across the detector surface. Each sensitive element comprises a superconducting tungsten (W) transition-edge sensor (TES) 1μm in width, fed by a set of 350 $-$ μm-long aluminum (Al) collector fins. The TES serves as a sensitive thermometer and the fins act as antennas, concentrating phonon energy from a wide area onto the tiny TES. Each TES and its surrounding fins are together referred to as a "QET," standing for *quasiparticle-trap-assisted electrothermal-feedback transition-edge sensors.*

Quasiparticle trapping

Rather than measuring the detector substrate temperature directly, our TES's are configured to detect phonons collected in an array of aluminum fins arranged around each TES. If a phonon has sufficient energy (*i.e.* at least twice the superconducting gap energy in Al: $2\Delta_{\mathrm{A}\,l} = 0.36$ meV) it can break a Cooper pair in the Al fins to generate

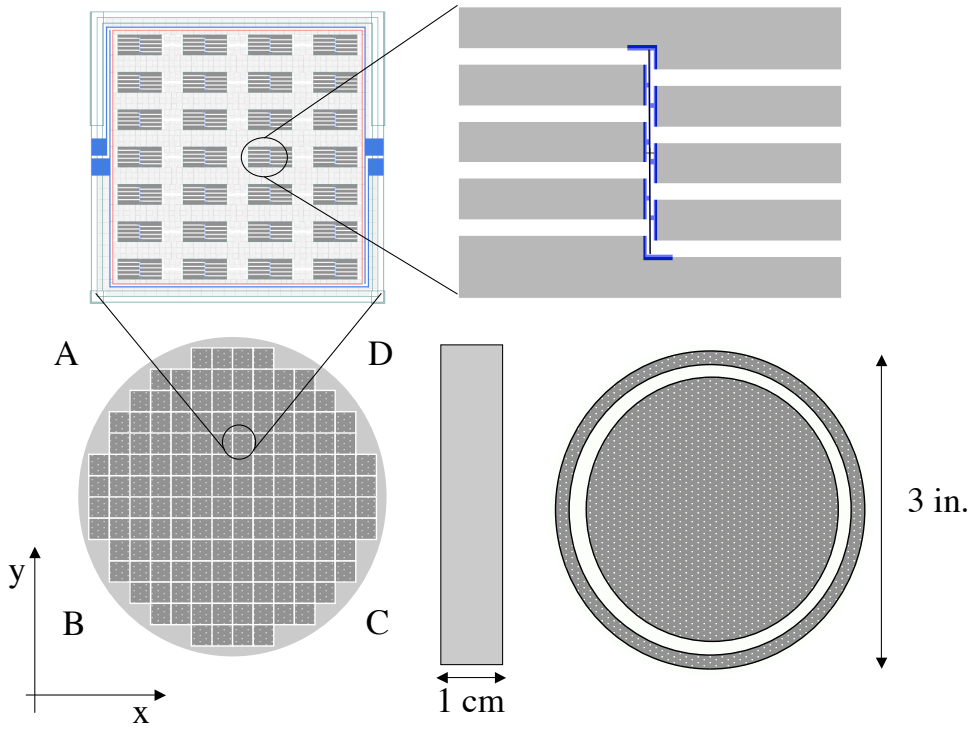


Figure 3.10: Composite cartoon of the layout of a ZIP detector, taken from [177]. *Bottom right:* bottom (ionization) face, showing the inner and outer electrodes. *Bottom left:* top (phonon) face, illustrating the nomenclature four the four sensor quadrants (A-D) and the $x - y$ coordinate system (z is depth within the substrate). *Top left*: One of the 37 photolithography dies that make up each quadrant of the phonon surface, illustrating the 28 individual QETs surrounded by the ionization grid. *Top right*: A single QET, consisting of a 1μm-wide W TES connected to Al collector fins.

quasiparticles. Note that this is much greater than the typical energy $k_B T = 3.4\mu$eV of thermal phonons at 40 mK; only energetic phonons far from equilibrium can create Al quasiparticles. These primary quasiparticles can lose excess energy by creating additional quasiparticles and by radiating sub-gap phonons, eventually reaching energies near the minimal 2Δ. The quasiparticles diffuse throughout the fins, eventually recombining after a time scale set by the cleanliness of the Al film.

These quasiparticles are collected from the Al films into the W TES following a mechanism suggested by N. Booth [178]. Since our W TES films have T_c's far lower than those of the surrounding Al films ($\sim$ 80 mK versus 1.2 K), the minimum quasiparticle energy in the TES ($2\Delta_{\mathrm{W}} \approx 25\mu$eV) is correspondingly lower than that in the fins (though still larger than the thermal phonon energy). An Al quasiparticle that diffuses to the TES edge can thus enter the W, but quickly loses enough energy to prevent its return to the Al. A slight overlap between the Al and W films produces a border region of intermediate gap through the proximity effect, reducing the probability of reflection at the interface. The Al-W interface thus acts as a one-way door, concentrating quasiparticle energy in the Al fins into the TES. This has the effect of amplifying our sensitivity to athermal phonons enormously: phonon energy from a wide area is concentrated into a TES of low heat capacity, producing a much larger temperature change than would be experienced with the TES alone.

Transition-edge sensors

In order to detect variations in the initial fast wave of athermal phonons, we require phonon sensors with high bandwidth ($\sim$ 100 kHz) and extremely low noise. CDMS achieves these goals using arrays of superconducting transition-edge sensors (TESs), the central sensitive elements of the ZIP phonon channels. A TES is a thin film of superconductor maintained at its transition temperature T_c, part-way through its transition from superconductivity to normal conduction. Since this transition is extremely sharp, a TES can serve as an extremely sensitive thermometer: a tiny temperature change causes a macroscopic change in resistance which may be detected with a sensitive amplifier. The sensitivity of a TES is characterized by the dimensionless derivative of the transition at the bias point (quiescent temperature T_0, resistance R_0):

$$\alpha \equiv \frac{d\log R}{d\log T} = \frac{T_0}{R_0}\frac{dR}{dT}. \tag{3.6}$$

Thermal fluctuation processes set the width of the superconducting transition [179] and thus the sensitivity of an ideal TES. Single TESs can attain values of $\alpha > 1000$, with transitions much less than 1 mK wide.

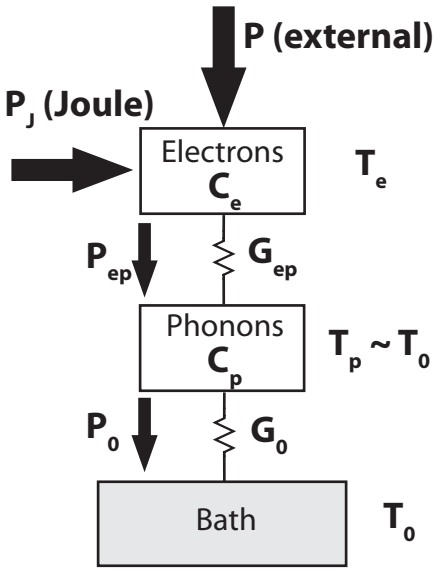


Figure 3.11: Cartoon of the thermal circuit of a TES based upon electron-phonon decoupling within the TES, *i.e.* $G_{ep} \ll G_0$.

Figure 3.11 depicts the basic thermal configuration of one of our tungsten TES's. The crystal lattice of the tungsten film (phonon system) is heat-sunk to the crystal substrate, which is cooled by a dilution refrigerator (thermal bath) to $T_p \approx T_0 \approx 40$ mK. The electron system of the tungsten film is heated by the Joule power of the electric currrent flowing through the TES, as well as by any external power from incoming quasiparticles. Electron-phonon interactions are very weak at these low temperatures, so the thermal link between the two systems is weak ($G_{ep} \propto T^5$, $G_{ep} \ll G_0$) and the electron system can self-heat to a temperature substantially higher than that of the substrate. In equilibrium the electron temperature T_e finds a value such that the heat flow entering from Joule heating ($P_J = IV = V^2/R$) and any constant ambient external power matches that flowing away to the bath: $P_J + P = P_{ep} = G_{ep}(T_e - T_p) \simeq G_{ep}(T_e - T_0)$. Since the TES's temperature affects its resistance and thus the Joule power, the exact choice of bias point is set by the feedback between this thermal configuration and the electrical bias circuit. The full theory of TES

behavior in realistic circuits is described in a review article by Irwin and Hilton [180]; here I limit myself to the simplest conclusions in the case of .

Stable operation can be achieved by operating a TES in a voltage-biased configuration with negative electrothermal feedback (ETF) [181]. If the voltage V_b across the TES is held constant by the external circuit, then any increase in TES temperature (and thus resistance) leads to a decrease in the Joule power:

$$\frac{dP_J}{dT} = \frac{\partial(V_b^2/R)}{\partial R}\frac{dR}{dT} = -\frac{V_b^2}{R_0^2}\frac{dR}{dT} = -\alpha\frac{P_{J,0}}{T_0}. \tag{3.7}$$

This reduction in Joule power acts as negative feedback, allowing the TES to quickly cool to its quiescent temperature after an energy deposition. Electrothermal feedback is most effective if the TES is self-heated well above the temperature of the local thermal bath ($T_e \gg T_0$).

ETF also greatly increases the operational bandwidth of a TES sensor. Consider a self-heated, voltage-biased TES of heat capacity C connected to a thermal bath T_0 by a link of conductivity G (assumed constant for small temperature excursions). If its temperature deviates from equilibrium by an amount $\delta t \equiv (T - T_0)$, the system will recover with a characteristic timescale of $\tau_{thermal} = C/G$. For a TES we must further account for the contribution of Joule heating to the heat flow balance:

$$C\frac{d\delta T}{dt} = \delta P_J - G\delta T = -\left(\frac{\alpha P_{J,0}}{T_0} + G\right)\delta T. \tag{3.8}$$

For the rapid spike in temperature from a particle event, the solution to this equation is an exponential decay with a new time constant. Defining the dimensionless gain factor $\mathcal{L} \equiv \frac{\alpha P_{J,0}}{GT_0}$ for convenience, we can write this new recovery time constant as

$$\tau_{ETF} = \frac{GT_0 + \alpha P_{J,0}}{CT_0} = \frac{\tau_{thermal}}{1+\mathcal{L}}. \tag{3.9}$$

In the limit of extreme electrothermal feedback, $\mathcal{L}$ can be much greater than unity and the recovery time can be sped up enormously. Our TES's exhibit an ETF recovery time of order $\tau_{ETF} \sim 200\mu$s.

It is interesting to note that the calibration of a TES in strong electrothermal feedback is particularly simple. In the limit of extremely strong ETF, the temperature change in the TES resulting from an external power input goes to zero. This means that the change in Joule power must exactly balance the incident power. This conclusion survives

in the limit of a perfectly linear TES transition, and in most practical cases we have

$$\Delta E = \int dt\, P_{ext} = -\int dt\, \delta P_J = -V_b \int dt\, \delta I. \tag{3.10}$$

In practice, however, we generally calibrate our phonon amplitudes empirically and little use is made of this relationship.

The four phonon sensors on the top surface of a ZIP detector are each composed of 1036 tungsten (W) TES's wired in parallel. Voltage biasing allows for stable operation of such an array as if it were a single TES, even if the individual sensors have slightly different transition temperatures: self-heating will naturally bias each TES within its transition. This array configuration allows TESs to be distributed across the entire detector surface without the complexity of reading each out separately. The price is a broadening of the transition (and thus a reduction in α to values below 100) due to variations between individual TES's.

Each TES is a tungsten film 1μm wide, 250μm long, and 35 nm thick. The normal-state resistance of a CDMS TES is ~ 1.25 kΩ, giving each sensor array a resistance of order $\sim 1\Omega$. The TES films are deposited by a Balzers sputterer as a mixture of α-phase and β-phase tungsten, targeted for a transition temperature of $T_c \sim 120$ mK. Following a technique developed by B. Young *et al.* [182, 183], we use implanted Fe ions to tune the transition temperatures of each sensor to produce a uniform T_c near 80 mK across the detector.

3.3.4 The phonon amplifier

Figure 3.12 is a simplified schematic of a ZIP phonon channel amplifier. A full analysis of the transfer function of this amplifier, accounting for finite gain and the finite impedance of the voltage biasing circuit, would take us afield of our main narrative, but I refer the reader to the aforementioned review article [180]. The TES (actually a parallel TES array) is kept voltage-biased in its transition by the current bias I_b and the small shunt resistor $R_{sh} \ll R_{TES}$. Any change in TES resistance produces a corresponding change in the input coil current $\Delta I_i \approx -I_i \frac{\Delta R_{TES}}{R_{TES}}$. A small input coil translates this into a change in magnetic flux through a sensitive magnetometer: a series array of DC SQUIDs. The resulting change in SQUID voltage is amplified and fed back to a feedback coil, arranged in negative feedback to keep the flux threading the SQUID constant. The feedback coil has one-tenth as many turns as the input coil, so the amplifier must drive a current $10\Delta I_i$ to cancel the change in SQUID flux. The overall amplifier thus has an approximate transresistance

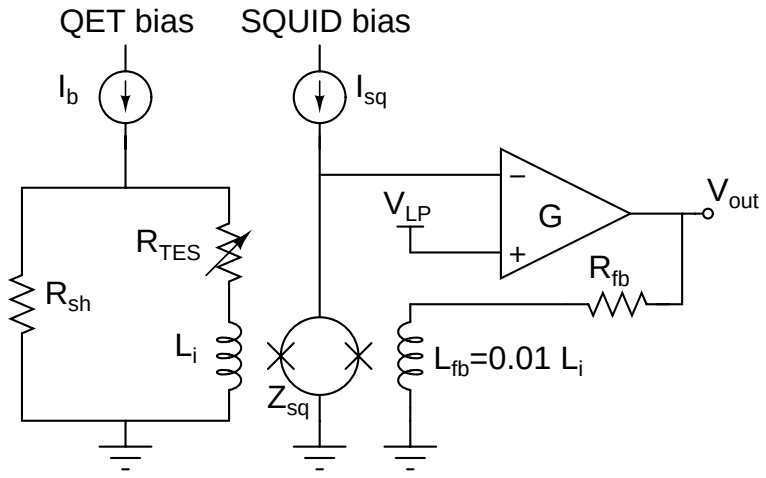


Figure 3.12: Simplified schematic of the CDMS phonon amplifier. In the current CDMS circuit, $R_{TES} \sim 200$ mΩ, $R_{sh} = 25$ mΩ, $R_{fb} = 1200\Omega$, and $L_i = 250$ nH $= 100 L_{fb}$.

of $10R_{fb} = 12$ kΩ referred to currents in the input coil. This gain is rolled off by the input coil's self-inductance at $f_{3dB} = \frac{Z_{TES}}{2\pi L_i} \approx 80(150)$ kHz for Ge (Si) ZIPs.

The noise performance of the phonon channel is set by fluctuations in the system's three main dissipative components – the TES's electrical resistance, the shunt resistor, and the electron-phonon conductance G within the TES – and by the noise current of the SQUID array. The full theory of TES noise is quite intricate due to the complex interplay between the electrical and thermal circuits; the full result is discussed in [180]. In our case, it is a good approximation to say that the Johnson noise of the shunt resistor (located with the SQUID at ~ 1 Kelvin, rather than with the TES at 40 mK) is the dominant contribution to the current noise in the input coil: $i_{sh} = \frac{\sqrt{4k_B T_{sh} R_{sh}}}{R_{TES}} \approx \frac{\sqrt{4k_B(1\text{K})(0.025\Omega)}}{(0.25\Omega)} \approx 15$ pA$/\sqrt{\text{Hz}}$. The TES contributions (Johnson noise and thermal fluctuation noise) are suppressed relative to this by the lower TES temperature, while the SQUID itself contributes only a few pA$/\sqrt{\text{Hz}}$. Like the signal gain, these fluctuations are rolled off by the L/R time constant formed by the input coil and TES resistance.

As a final note, it is possible to use the electrical response of the TES circuit – *i.e.* its complex impedance – to probe the electrothermal characteristics of the TES itself. The TES's response to an electrical stimulus of its bias line is similar (though not identical) in form to its response to added thermal power, and some of the relevant time constants may be read off of this. Figure 3.14 shows a recent attempt to measure the frequency-dependence of the TES circuit's amplitude response. The L/R pole is clearly visible near ~ 80 kHz,

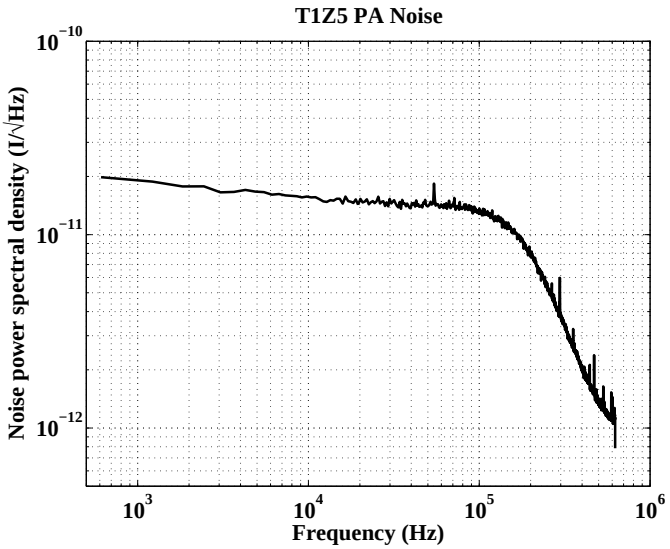


Figure 3.13: Observed phonon noise spectrum for T1Z5 channel A in Soudan Run 123. Spectrum is referred to the SQUID amplifier input coil.

but the undulation at lower frequencies is not completely understood. Matt Pyle has shown that these features are likely to be indications of phase separation, a condition in which a narrow TES cannot maintain a uniform temperature along its length and partitions into regions of different resistance. This does not have immediate implications for this WIMP-search analysis, but does suggest modifications to our TES design that will make our phonon sensors easier to characterize.

3.3.5 Phonon reconstruction

Significant variations in pulse shape with amplitude, position and recoil type render phonon pulse reconstruction significantly more complex than charge reconstruction. No single filter configuration is optimal for all aspects of this process. The current data reduction package draws a sharp distinction between two aspects of phonon pulse analysis: energy reconstruction and pulse shape characterization. Each has different requirements and applies different algorithms, as described below.

Energy reconstruction

Phonon pulse energies are computed using an optimal filter algorithm, as described in Section 3.2.4. A start time and amplitude are fit separately for each phonon sensor, within

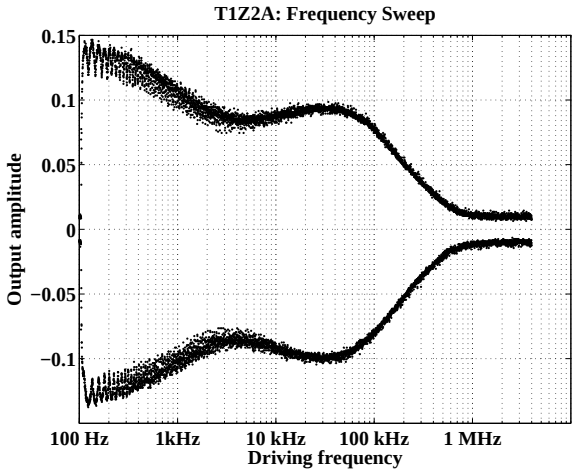


Figure 3.14: Amplitude response of phonon channel T1Z2A to electrical stimulus, measured by driving the TES bias line with a sine wave frequency sweep and observing the output. Amplitude units are arbitrary for our purposes. The "neck" at the plot's right extremity is an artifact of experimental noise, not a true plateau in response.

start time windows of $[-50, +200]\mu$s for Ge and $[-25, +100]\mu$s for Si. Each detector uses a simple template based on a two-exponential functional form, characterized by a risetime and falltime: $A(t) = A_0(1 - e^{-t/\tau_1})e^{-t/\tau_2}$. Since phonon pulse shapes vary substantially with position and energy while the template remains static, the resulting amplitude estimates have significant position dependence and energy nonlinearity which must be corrected for. This variation in pulse shape limits the usefulness of the phonon χ^2 and start time, so they are generally ignored during the analysis.

DarkPipe also computes an alternate estimate of each phonon pulse energy using the integral of each trace. This estimate has smaller position and energy variation than the optimal filter result, but due to its poorer noise performance it is only used for high-energy events. These integral estimates play no role in the present WIMP-search analysis.

Pulse shape characterization

The phonon pulse shape parameters currently in use by CDMS take the form of first-crossing times for particular thresholds. The primary phonon risetime `pminrt`, for example, is defined as (`PXr40-PXr10`), the difference between the times at which the primary phonon pulse last exceeds 10% and 40% of its maximum amplitude along its rising edge. DarkPipe computes these crossing times using a time-domain walk algorithm, applied after

filtering the pulse with a 50 kHz (90 kHz for Si) zero-delay low-pass Butterworth filter to reduce noise jitter.

A primary advantage of these walked timing parameters is that they do not depend on any particular model of the phonon response of the detector, which is relatively complex. The price of that ignorance of the signal, however, is poorer discrimination against noise. This is generally not a major concern at practical energies for the primary (largest) phonon pulse for each event, but it can be more problematic for the much smaller and slower opposite (smallest) phonon pulse. As shown in the left plot of Figure 3.15, the distribution of walked opposite phonon pulse delay can be distorted at low pulse amplitudes. Less distortion is apparent in a simple time-domain fit of the phonon pulse, shown in the right plot. Further development of such fits is in progress, but due to time constraints they were not used in the present analysis.

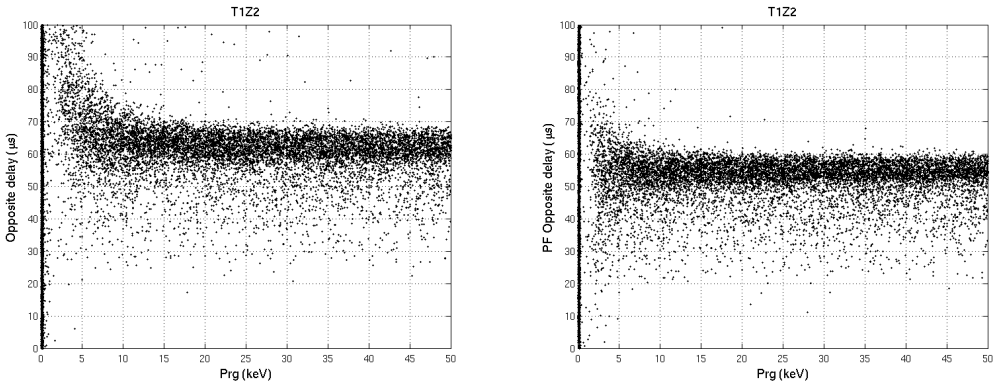


Figure 3.15: Time delay of the *smallest* (opposite) phonon pulse for a given event as a function of total event energy, computed for *Left:* DarkPipe's walk algorithm and *Right:* PipeFitter's time-domain fitting algorithm. No correction for position dependence has been performed. Note the greater stability of PipeFitter's fitted delay at low energies, where noise fluctuations can thwart a simple walking algorithm.

Position reconstruction

The approximate position of each event within the detector may be reconstructed by comparing the phonon amplitude and shape parameters described above among the four phonon sensor arrays. CDMS uses two distinct position-reconstruction techniques based upon these parameters:

1. **Phonon partition (`xppart, yppart`)**: An event should deposit more of its energy in the nearest phonon sensor than in its farther neighbors. We define two position-reconstruction parameters based upon the partitioning of phonon energy among the four sensors:

$$\texttt{xppart} \equiv \frac{(\texttt{pc}+\texttt{pd})-(\texttt{pa}+\texttt{pb})}{\texttt{pa}+\texttt{pb}+\texttt{pc}+\texttt{pd}}$$

$$\texttt{yppart} \equiv \frac{(\texttt{pa}+\texttt{pd})-(\texttt{pb}+\texttt{pc})}{\texttt{pa}+\texttt{pb}+\texttt{pc}+\texttt{pd}}.$$

   Where needed, we can also define a radial parameter $\texttt{rpart} \equiv \sqrt{\texttt{xppart}^2+\texttt{yppart}^2}$. These parameters produce the position manifold shown in Figure 3.16(a), known internally as the "box plot." Partition parameters map the circular ZIP detector to an approximate square, but give the best position resolution due to the excellent amplitude resolution of the optimal filter algorithm.

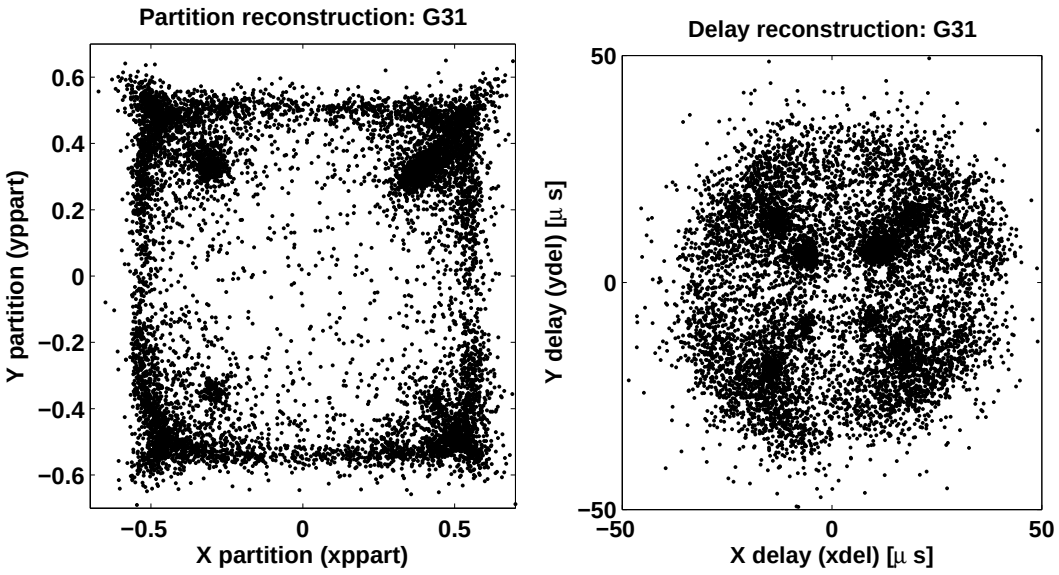
2. **Phonon delay (`xdel, ydel`)**: An event's phonon energy should also reach the nearest phonon sensors before the farther ones. We define two position-reconstruction parameters based upon the relative start times (technically the 20% rising edge points, `PXr20`) of the largest phonon pulse and its two adjacent neighbors. For an event under sensor A, these quantities become:

$$\texttt{xdel}^{(A)} \equiv \texttt{PAr20}-\texttt{PDr20}$$

$$\texttt{ydel}^{(A)} \equiv \texttt{PBr20}-\texttt{PAr20}.$$

   Where needed, we can also define a radial parameter $\texttt{rdel} \equiv \sqrt{\texttt{xdel}^2+\texttt{ydel}^2}$. This yields the more circular "delay plot" shown in Figure 3.16(b).

Each of these position reconstructions suffers from problems of degeneracy: two events at different radii may show up at similar places in the box or delay plot. This is primarily due to phonon reflections, which distribute the energy of an event near the crystal edge to the more distant TES's. Figure 3.17 illustrates this effect and its solution: both partition and delay reconstructions fold back differently at high event radii, so their combination can be used to break the degeneracy. This combined position reconstruction is an important component of the phonon correction manifold described in the next Section. It is important to note, however, that reconstruction degeneracies may still remain near the outer edge.



(a) (b)

Figure 3.16: Position reconstruction plots for data from the G31 ^{109}Cd calibration using (a) phonon partition and (b) phonon delay. The dark splotches indicate event concentrations from the eight narrow holes in the source collimator, in this case placed on the detector's phonon face.

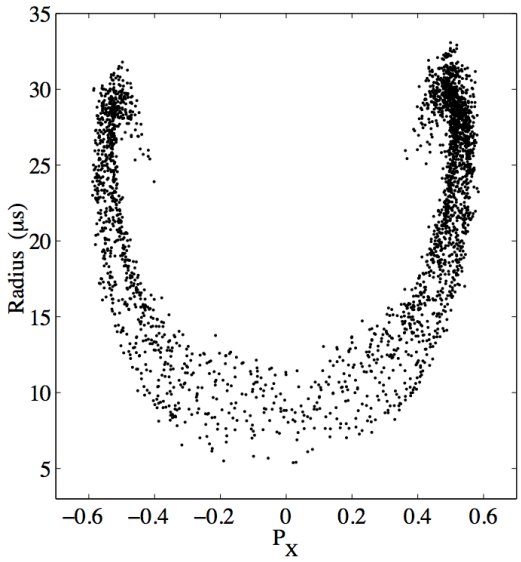


Figure 3.17: Plot of the two major position reconstruction parameters for events along a narrow slice across a ZIP detector (small `yppart`). The x-axis is the partition parameter `xppart`, while the y-axis is the radial delay parameter `rdel`. Note that `xppart` is less position-sensitive at high radii, but this degeneracy is greatly reduced by the combination of the two parameters.

3.3.6 Phonon calibration

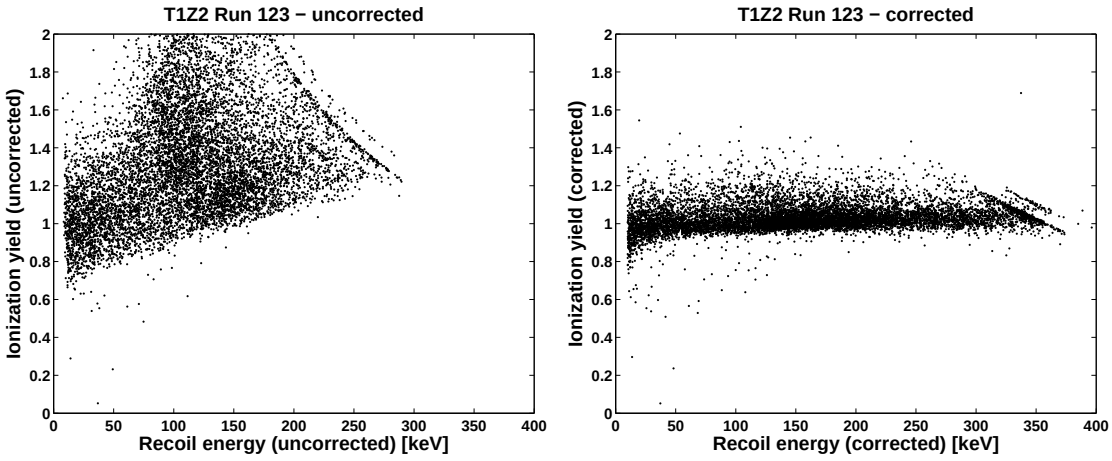
Basic calibrations

Within a single detector, the sensitivities of the four phonon sensors often differ by $\sim 10 - 15\%$. To make the phonon response more uniform across the detector, the four sensor amplitudes are scaled by four gain factors chosen to minimize the width of the ionization yield band. After applying these relative gains, the phonon energy is calibrated with reference to the ionization signal: the summed phonon amplitude is multiplied by an overall gain factor chosen to ensure ionization yields near unity at low energies.

Figure 3.18(a) is a plot of ionization yield versus recoil energy after these calibrations have been applied. The yield band is very diffuse and its mean value depends strongly upon event energy. Both of these effects are due to the strong position- and energy-dependence of athermal phonon signals. The optimal filter algorithm assumes a fixed pulse shape, and so assigns slightly different amplitude values for two pulses of the same energy but different shapes. Variations in pulse shape with event energy (due to TES nonlinearity) and position (due to geometric effects) thus appear as variations in phonon

energy sensitivity. These same variations in pulse shape cause walked timing parameters (*e.g.* risetime) to vary with event position and energy. Both of these effects are made yet worse by variations in TES properties within and between each phonon sensor array.

Phonon correction table



(a) (b)

Figure 3.18: Ionization yield as a function of recoil energy for ^{133}Ba calibration data (a) before and (b) after applying the phonon correction table.

These position- and energy-dependencies are reduced through the use of a phonon correction table, a technique initially developed by Blas Cabrera and Clarence Chang and described in [177]. The basic principle of this method is to compare each event's phonon parameter values against a look-up table describing how the mean values of these parameters vary with event location and energy. By factoring out the bulk trends expressed in the look-up table, we obtain "corrected" parameters with more uniform response and superior resolutions.

The look-up table of correction factors is derived empirically from the large sample of photon events acquired during the ^{133}Ba calibration runs. A separate table is defined for each ZIP detector and for events beneath each of the four phonon sensors within that detector. Each event is located in a four-dimensional manifold defined by three position variables (`xppart`, `yppart`, `rdel`) and one energy variable (`prg`). The addition of phonon energy to the table definition is a new innovation for this run by Bruno Serfass, which achieves superior parameter resolution and obviates the need for a separate correction for

energy nonlinearities. Each event is associated with a set of "nearest-neighbor" events within this manifold, evaluated using a metric on this parameter space chosen empirically to achieve good correction performance. The table takes the form of a mapping between each of these calibration photon events and the mean values of the desired phonon parameter across its set of nearest-neighbor events. The number of nearest neighbors affects the performance of the correction algorithm: a table with too few neighbors is limited by statistical noise, while one with too many neighbors averages over too large a region of parameter space and misses local trends.

For each WIMP-search or calibration event we wish to correct, we find its nearest neighbor from among the lookup table gammas according to the same metric. The lookup table provides an associated mean value over the set of neighbor events. For a parameter "`par`", the corrected parameter value "`parc`" is given by

$$\texttt{parc} = \texttt{par} \cdot \frac{\langle \texttt{par} \rangle_{\text{ all gammas}}}{\langle \texttt{par} \rangle_{\text{lookup table neighbors}}}$$

The correction table algorithm is vulnerable to degeneracies in the 4D manifold on which the lookup table is based. If two separated portions of the detector give similar values of `xppart`, `yppart` and `rdel` the set of nearest neighbors for a given event may include events from both detector regions. If these regions have different intrinsic distributions of the parameter under consideration, erroneous corrections can occur. ZIP detectors unfortunately show substantial degeneracies between outer and inner detector radii in all of these parameters. Though the combination of the three breaks most of these degeneracies, troublesome regions of miscorrection remain.

3.4 Ionization yield discrimination

A charged particle moving through a crystal lattice can lose energy to the lattice in two major ways: by liberating valence electrons or by imparting kinetic energy to entire atomic cores. The relative rates of energy loss through these two pathways ($\frac{dE_e}{dx}$, $\frac{dE_a}{dx}$) depend upon the target material and on the projectile's mass, charge, and kinetic energy. Fast, lightweight projectiles (*e.g.* electrons) deposit almost all of their energy in the electron system, while slower, heavier projectiles (*e.g.* recoiling atomic nuclei) can transfer a significant proportion to atomic cores.

This division of energy has important implications for particle detection. Both charge carriers and scintillation photons are generated from excitations of the electronic system, so the ionization and scintillation signals observed from a nuclear recoil event is reduced relative to an electron recoil event of the same energy. This effect must be accounted for in the analysis of data from any WIMP-detection experiment. Moreover, a properly-designed experiment may distinguish electron recoils from nuclear recoils by measuring two or more attributes of each particle event.

Following this principle, CDMS discriminates electron recoil events (most radioactive backgrounds) from nuclear recoil events (WIMPs and neutrons) by comparing the magnitudes of each event's ionization and phonon signals. The ionization signal of nuclear recoil events is reduced as described above, while the phonon signal is proportional to an event's total energy deposition. CDMS's primary discrimination parameter is ionization yield, defined as

$$y \equiv \frac{E_Q}{E_r} = \frac{E_Q}{E_P - \frac{eV_b}{\epsilon} E_Q}. \tag{3.11}$$

In this expression, ϵ is the mean electron-recoil energy required to generate a single electron-hole pair and the charge energy E_Q is normalized so that $y = 1$ for electron recoils. Figure 3.19 shows the enormous discrimination power of this parameter: nuclear recoils are clearly separated from electron recoils at recoil energies above ~ 3 keV. Fewer than one electron recoil in 10^4 above ~ 10 keV leaks into the $\pm 2\sigma$ nuclear recoil band, as defined by neutron calibration data.

Lindhard and coauthors [184] have modeled the energy-dependent ionization yield expected from nuclear recoils in various crystals. The position and shape of the nuclear recoil band in Ge ZIPs, shown in Figure 3.19, is in excellent agreement with that predicted by Lindhard's model. As Figure 3.20 shows, the agreement is less impressive for Si: the observed nuclear recoil band is uniformly lower than predicted at high recoil energies. In this Figure, the variations visible among detectors and the poor agreement with the model at very low energies ($\lesssim 10$ keV) are consequences of calibration variations and the simple fitting function; the large discrepancy with the model at high energies is not. This discrepancy is believed to be a consequence of reduced ionization collection in Si at low electric fields (4 V/cm, in this case), but the detailed mechanism is not understood. This theory is supported by data from a Si ionization detector taken at much higher fields (2000 V/cm) by Dougherty *et al.* , which agrees well with the Lindhard prediction. This discrepancy is

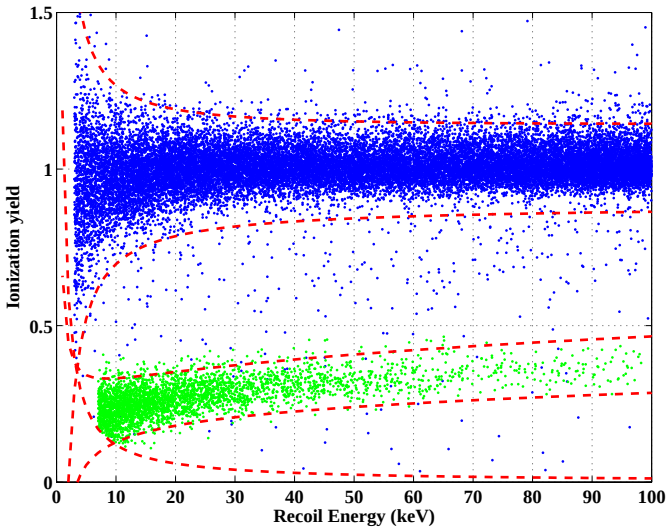


Figure 3.19: Ionization yield as a function of recoil energy for the Ge ZIP T1Z5. The blue points (*upper band*) represent electron-recoil events from a calibration with a ^{133}Ba source. The green points (*lower band*) represent nuclear-recoil events from a ^{255}Cf neutron source. Dashed lines indicate the approximate $\pm 2\sigma$ boundaries of the bands, as determined by empirical fits.

not expected to impact our data analysis, since our yield bands are fit empirically and any localized differences in charge recombination should not affect the accuracy of our recoil energy scale.

3.5 The dead layer

As described above, ionization yield is an extremely powerful discriminant against electron recoils in the bulk of the target crystal. Unfortunately, these are not the only background faced by CDMS. Electron recoils which occur in a $\sim 10\mu$m-thick "dead layer" near the detector surface show suppressed ionization collection, occasionally giving ionization yields low enough for such surface events to be mistaken for nuclear recoils. This effect was observed in early work with Ge ionization/phonon detectors by Shutt and collaborators in 1992 [186]. Surface events can arise from incident electrons or low-energy X-rays, which have very short penetration depths, or from Compton scattering within the near-surface layer. It is these surface events, rather than bulk electron recoils, which constitute the dominant background of the CDMS experiment (and the related EDELWEISS experiment). The successful rejection of surface events at ever-higher exposures is crucial to the

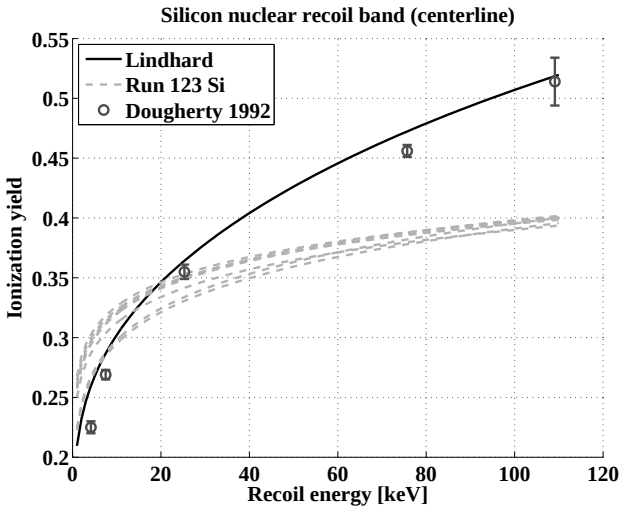


Figure 3.20: Comparison between the fitted nuclear recoil band centerlines for the various Si detectors in this data run (*dashed curves*) with the prediction of Lindhard theory (*solid curve*). Also shown are data points taken with a Si ionization-only detector at high field by Dougherty *et al.* [185].

future of semiconductor detector technology.

3.5.1 Origins of the dead layer

To understand the origin of the surface event problem, consider the case of an electron recoil within a few microns of one of the surface electrodes. For definiteness, suppose that the interaction occurs near the more-negative electrode (the ionization electrode face, given the usual CDMS bias polarity); an event near the opposite electrode shows the same behavior, with electrons and holes swapped. Holes from this event drift only a short distance, so the ionization signal is determined by the number of electrons which drift across the crystal to the opposite electrode.

In principle, there are at least three mechanisms by which an event's proximity to the surface electrode could reduce the ionization signal:

1. **Back-diffusion**: Some charge carriers will reach the surface electrode before the bias field draws them fully free of the initial charge droplet. In particular, a substantial number of opposite-sign charge carriers (electrons, in this case) will back-diffuse into the electrode. Any carrier reaching the metal electrode is expected to relax to the Fermi level almost immediately and not drift across the crystal. This mechanism is

illustrated schematically in Figure 3.21.

2. **Enhanced trap density**: The crystal surface could contain a higher density of defects and other trapping sites due to damage from the fabrication process (*e.g.* from ion implantation) or chemical reactions with air. These traps could capture a disproportionate number of carriers for near-surface events, preventing drift across the crystal.

3. **Electrostatic induction**: An electric charge near a conducting surface is attracted to that surface due to its polarizing effect upon the conductor. For an isolated electron, this force exceeds that of a 3 V/cm drift field when the electron approaches within $\sim 100\mu$m of the conducting electrode. This distance scale is clearly far too large, however, since we do not actually have isolated charges – a realistic model (outside the scope of this work) must account for the polarization effects of both carrier types in the initial charge droplet, which drastically reduces the net attractive force. Nonetheless, this mechanism may affect the motions and trapping of charge carriers near a detector surface.

Back-diffusion is believed to be the dominant source of the dead-layer effect, a theory supported by work on charge contacts by Shutt and collaborators [169]. The authors found that an amorphous Si layer between the bulk crystal and the electrode drastically reduced the thickness of the dead layer, consistent with a blocking of back-diffusion by the large pseudo-band-gap of the amorphous material. A dead layer effect still remains, however, due either to imperfect blocking or to the other two mechanisms. N. Mirabolfathi and the Berkeley group are currently attempting to study the source of the dead layer effect using Ge test devices with charge contacts suspended above the detector surface.

3.5.2 Characterization of the dead layer

To study the dead layer in more detail, the Berkeley CDMS group undertook an extensive calibration study of a Ge ZIP detector (designated "G31," now installed at Soudan as T2Z5) with a collimated ^{109}Cd β-source [2]. This calibration spanned two runs of the Berkeley dilution refrigerator in late 2002, one with the source illuminating the detector's charge face, the other illuminating its phonon face. The analysis of these data is described more fully in V. Mandic's dissertation [1]. Figure 3.22 illustrates the recoil energies and

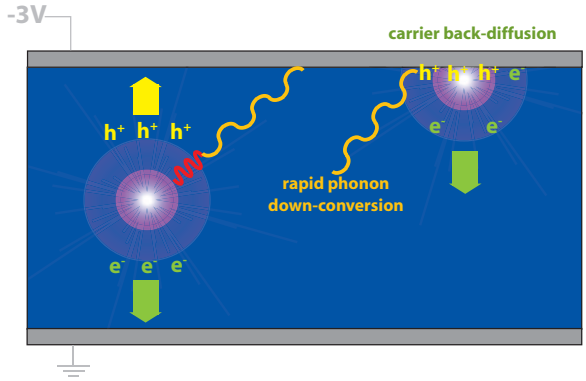


Figure 3.21: Cartoon of the dead layer effect in a ZIP detector. Carriers (h^+, e^-) from an event in the detector bulk (*left*) are fully collected, while those from a surface event (*right*) may be lost to diffusion into the nearby electrode. The metal electrode also mediates a more rapid conversion of high-frequency phonons (narrow sine) to low-frequency ballistic phonons (wide sine).

ionization yields of events from these calibrations, including the prominent 22-keV X-ray line and low-yield surface events of ^{109}Cd decay.

The G31 data analysis revealed a great deal of qualitative and quantitative information about the dead layer. Most prominently, these data revealed a stark asymmetry between surface events on the two detector faces. This difference is clearly visible in Figure 3.22: the intermediate-yield band of ^{109}Cd electron events appears at higher ionization yield when the source is on the charge face than when it is on the phonon face. Figure 3.23 illustrates the same effect in surface events from ^{133}Ba calibration data taken at Soudan. The origin of this phenomenon is not fully understood, but may be due to asymmetric Schottky-type effects at the charge contacts. This effect has important implications for the detector's response to surface events: the vast majority of surface events which can mimic nuclear recoils in yield occur on the detector's phonon face.

This observation can also be described in terms of variations in the dead layer's thickness. These data were used to estimate the dead layer's thickness using a technique due to T. Shutt: under the assumption that yield is a monotonic function of depth, we can use the known (exponential) depth distribution of the 22-keV X-ray photons and the observed distribution of ionization yield to infer the dependence of yield upon depth. Figure 3.24 shows the result of such an analysis, based upon events from collimated sources (shielded with foil to block surface electrons). Finite resolution effects determine the slope of this

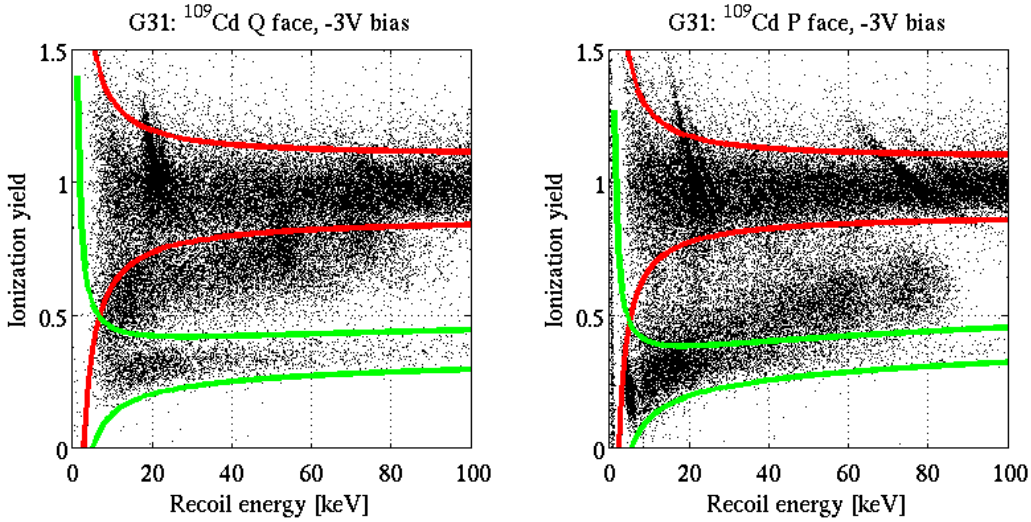


Figure 3.22: Ionization yield bands from calibrations of detector G31 with a ^{109}Cd source on the charge face (*left*) and phonon face (*right*), after correction for energy nonlinearity and position dependence. The solid lines indicate the $\pm 2\sigma$ electron recoil (*top*) and nuclear recoil (*bottom*) bands. Surface β events appear between the two bands, with different yields and reconstructed energies on the two faces. Both data sets include calibration runs with a ^{255}Cf source to illustrate the location of the nuclear recoil band.

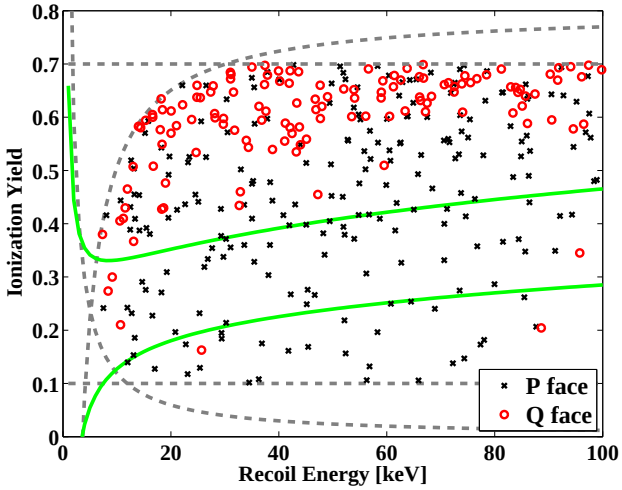


Figure 3.23: Distribution of surface events from ^{133}Ba calibrations of detector T1Z5 during this data run. The gray dotted lines define the various bounds of the surface event selection used here. The solid lines indicate the bounds of the $\pm 2\sigma$ nuclear recoil band. Many fewer events from the charge surface ($\circ$) than from the phonon surface ($\times$) appear within the nuclear recoil band.

plot at large depths, but the sharp drop in yield at shallower depths corresponds to the actual dead layer. In addition to confirming the dead layer's general scale ($\lesssim 10\mu$m), these data suggest a much thicker dead layer on the detector's phonon face that on its charge face. Comparisons of the dead layer thickness at different voltage biases also reveal that the dead layer narrows at high biases (as expected) and is much thicker at positive biases than negative.

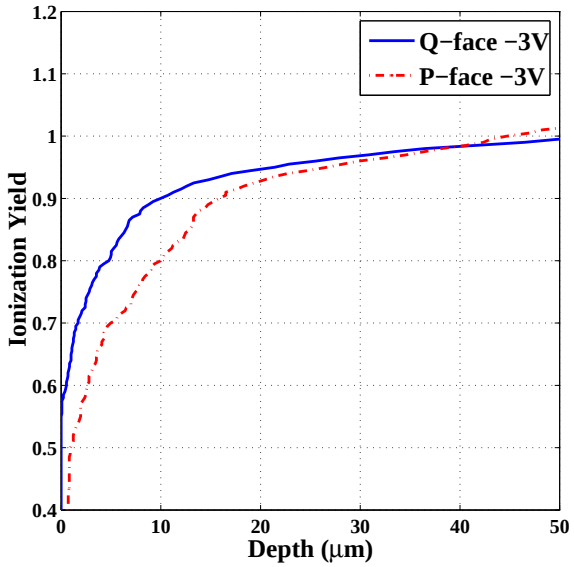


Figure 3.24: Estimated relationship between ionization yield and event depth, based upon 22-keV X-ray photons from the ^{109}Cd calibration of G31. The two curves represent data taken with the source on the detector's charge face and phonon face, indicating the asymmetry in response between the two faces. Plot modified from work by V. Mandic.

3.6 Surface event discrimination

The advantage of ZIPs over earlier semiconductor detectors technologies lies in their ability to distinguish surface electron recoils from bulk nuclear recoils, despite their similar ionization yields. The key to this discrimination power is the rapid response of the TES arrays to out-of-equilibrium ("athermal") phonons. This faster response permits us to resolve subtle differences in the temporal and spatial structure of the initial wave of phonons from a particle event. In this way, the CDMS phonon sensors behave as bolometers (power-sensing devices) rather that merely as calorimeters (energy-sensing devices), detecting subtle

differences in the "sounds" made by different particle impacts. These differences in phonon pulse timing and spatial distribution allow the CDMS experiment to reject more than 99% of surface events, while still maintaining substantial acceptance of nuclear recoils.

3.6.1 Discrimination parameters

Phonon timing

The CDMS experiment identifies surface electron recoils based primarily upon the rapid arrival of their athermal phonon signals. Relative to a bulk event of similar energy and location, phonon pulses from a surface event have a sharper rising edge and arrive more quickly after the charge pulse. Figure 3.25 illustrates the substantial difference between the shapes of 60-keV bulk and surface events in the same region of the detector. This difference in phonon timing is believed to originate in the down-conversion process whereby energetic primary phonons become lower-energy ballistic phonons. Primary phonons reaching a metal film are down-converted rapidly, short-circuiting the quasi-diffusive regime and resulting in a more rapid release of ballistic phonons. Moreover, the phonon pulses of bulk nuclear recoils are observed to be yet slower than those of bulk electron recoils. This effect may be due to differences in the Luke phonon contribution (bulk electron and nuclear recoils have different ionization yields) or in the initial phonon population. Either way, this phenomenon further improves our ability to discriminate nuclear recoils from surface events.

In this analysis we characterize the speed of the phonon signal using the shape of each event's primary (largest) phonon pulse. We define two major parameters based on the "walked" timing quantities described above:

1. The **primary phonon delay (`pdel`)**, defined as the difference between the start of the ionization pulse (as determined by the optimal filter algorithm) and the 20% point along the primary phonon pulse's rising edge (`PXr20`).

2. The **primary phonon risetime (`pminrt`)**, defined as the difference between the 10% and 40% points along the primary phonon pulse's rising edge (`PXr40-PXr10`).

Figure 3.26 illustrates the definitions of these parameters.

The raw timing parameters described above vary substantially with event energy and position due to TES nonlinearities and the finite propagation speed of ballistic phonons across the detector. A surface event near the detector's outer rim and a bulk event near it's

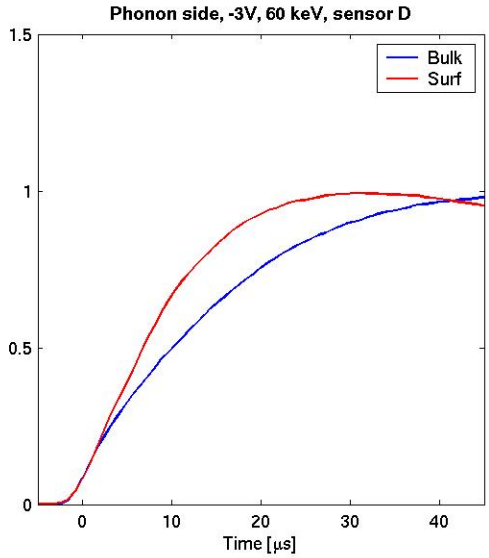


Figure 3.25: Comparison of the leading edges of 60-keV electron recoils in the bulk and (phonon) surface of a detector. Each pulse is an average of several similar events in a localized detector region of the G31 calibration data. Plot due to P. Meunier and B. Serfass.

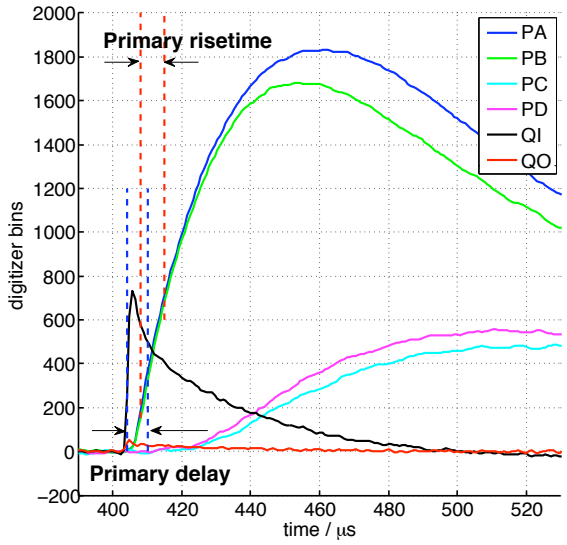


Figure 3.26: Plot of a digitized event trace from a ZIP detector, zoomed in on the leading edge to illustrate phonon timing parameters. The "primary" (largest) phonon pulse of this event occurs in channel A. The vertical dashed lines indicate the event's primary risetime (time between the 10% and 40% points of the primary phonon pulse's rising edge) and primary delay (time between the charge start time and the 20% point of the primary phonon pulse's rising edge).

center, for example, may have similar values of `pdel`. To improve surface event discrimination, this analysis uses phonon parameters corrected for energy- and position-dependencies using the lookup-table described in Section 3.3.6. Figure 3.27 compares the risetimes and delays of nuclear recoils and surface events after this correction: surface events are confined to a relatively narrow range of low values (fast pulses relative to their neighbors), while nuclear recoils extend to much higher timing values. A combination of a lower limit on these timing parameters with an upper limit in ionization yield will accept nuclear recoils while excluding surface events.

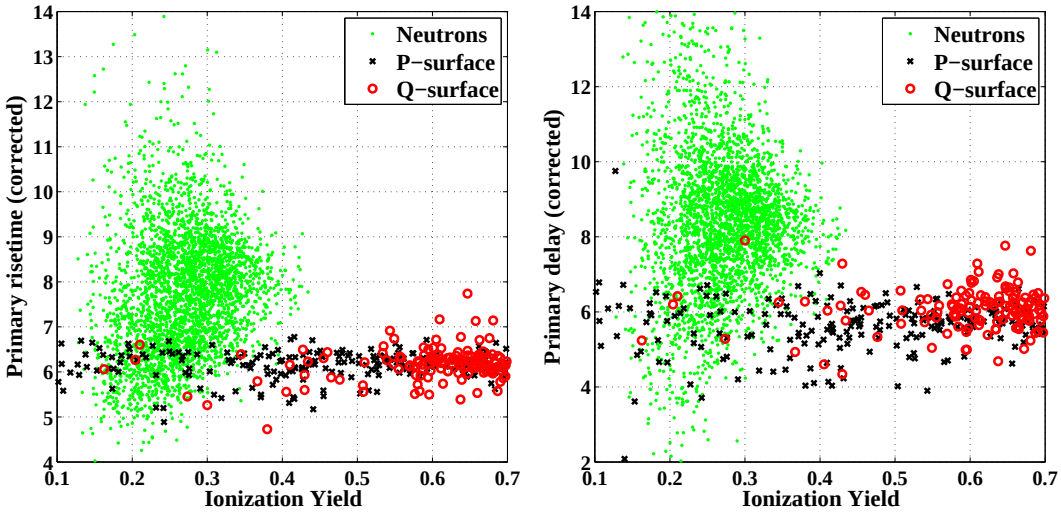


Figure 3.27: Ionization yield versus phonon timing parameters for T1Z5 calibration data from this data run. The left pane shows the primary risetime, while the right shows the primary delay; both timing parameters and yield are shown after application of the look-up-table correction.

As shown in Figure 3.28, primary risetime and primary delay are strongly correlated with one another. This correlation is not surprising: the two parameters depend upon similar processes, and their definitions include overlapping portions of the primary pulse (10%-20%). The various phonon shape and distribution parameters are generally correlated with one another fairly strongly. This is an important consideration when developing criteria to discriminate nuclear recoils from surface events.

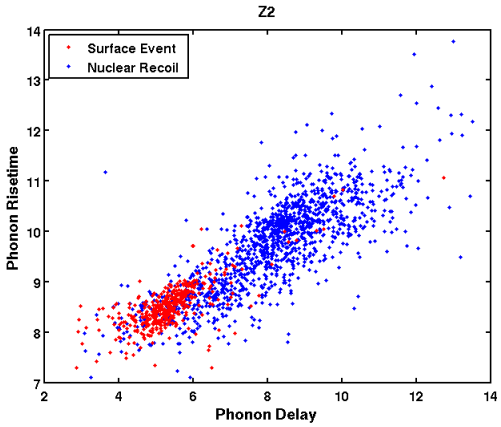


Figure 3.28: Scatter plot of primary risetime and delay (after correction) for calibration events on detector T1Z2.

Phonon partition

Surface and bulk events also differ in the distribution of collected energy among the four phonon sensor arrays. This is primarily a consequence of geometry: from the perspective of an event near to the phonon surface of the detector, the primary (nearest) sensor quadrant subtends a large solid angle. A near-surface event on a detector's phonon face will thus deposit more of its energy in the primary phonon sensor than would an event in the detector's bulk or near its charge face. Phonon-surface events can thus be distinguished from bulk events (and from charge-surface events) by an unusual concentration of phonon energy in a single sensor.

Figure 3.29 illustrates the power of phonon partition discrimination using data from the G31 calibration runs at Berkeley. The $x-$ and $y-$ axes of each panel represent approximate radial parameters based upon relative phonon amplitudes (`rpart`) and relative phonon timing (`rdel`), respectively. Surface events from the collimated sources are visible as green blobs atop the main distribution of photons and neutrons. Surface events on the phonon face are clearly shifted with respect to the bulk-event distribution, occurring at somewhat higher `rpart` and lower `rdel`.

This difference in phonon distribution has been used to construct two new parameters with which to distinguish surface events:

1. `pfracc` ("phonon fraction"): the ratio of the amplitudes of the largest phonon pulse and the pulse in the opposite channel, corrected for dependence on position and energy.

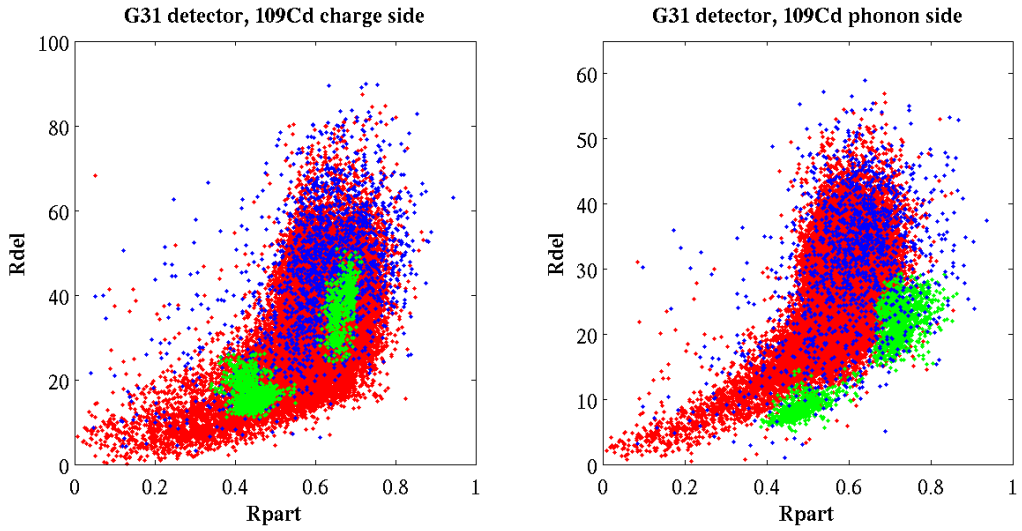


Figure 3.29: Illustration of the role of phonon delay and partition parameters in discrimination against surface electron recoils, based on data from the G31 calibration at Berkeley. In both plots above, the axes are approximate radius parameters generated from the relative amplitudes ("Rpart") and relative start times ("Rdel") of each event's phonon pulses. Each plot shows bulk electron recoils (*red*), bulk nuclear recoils (*blue*), and surface electron recoils localized near the collimated ^{109}Cd sources (*green*). The sources are in the same $x-y$ positions in both panels, but have larger Rpart and smaller Rdel relative to the bulk event population when the source lies on the phonon surface. Plots from B. Serfass.

For an event beneath sensor A, for example, `pfrac` = `pa`/`pc` and `pfracc` is the version after application of the look-up table correction.

2. `ppartc` ("phonon partition"): the box plot radius `rpart`, corrected using the look-up table to remove most dependence on position and energy.

Both of these parameters have more discrimination power at higher detector radii, where phonon distribution depends less upon event position. Studies of alternate surface event discrimination techniques in this and previous data runs have used `pfracc`, but the poor signal-to-noise of the opposite phonon pulse renders this parameter poorly-behaved at low energies. Future analyses are likely to use `ppartc` (or some variant thereof), which expresses similar information with better noise performance.

3.6.2 Discrimination techniques

With these phonon parameters in place, we must decide how to use them to discriminate surface events from bulk nuclear recoils. We may view this problem in terms of a multi-dimensional space defined by the various event parameters. The signal (nuclear recoil) and background (surface event) populations occupy particular portions of this parameter space. As in the rest of our data analysis (described further in Chapter 6), we seek to define the boundaries of a "signal region" which contains most nuclear recoils but which is essentially free of surface events.

Simple timing cuts

The simplest signal region to define is a rectangular solid in this parameter space, defined by a series of limiting values of the various parameters. This is essentially the technique used in the first data run at Soudan (Run 118) [187, 1]: we demand that a nuclear recoil candidate have risetimes and delays above some lower limits, defined to reduce the surface event background to acceptable levels (generally $\lesssim 1$ event in the associated WIMP-search data). We can also impose upper limits on these parameters to reject misreconstructed events which do not match either hypothesis (nuclear recoil or surface event). This is a simple solution, but may not be optimal: it does not account for the correlations among various discrimination parameters, and we must tune ever more free parameters (the cut levels along each axis) if we wish to incorporate new discrimination parameters into our analysis.

Correlated timing variables

The 2-Tower analysis (Run 119) and the current analysis both use a slight modification of this technique. Instead of making independent cuts along the `pminrtc` and `pdelc` axes, we take advantage of the strong correlation shown in Figure 3.28 by making cuts along rotated axes: (`pdelc`+`pminrtc`) and (`pdelc`-`pminrtc`). The Run 119 analysis implemented an additional correction for the energy-dependence of these parameters, but improvements to the look-up table rendered this unnecessary in the current analysis. A lower limit in (`pdelc`+`pminrtc`) serves as a *discrimination cut*, positioned to accept nuclear recoils while rejecting surface events. An upper and lower limit in (`pdelc`-`pminrtc`) acts as a *consistency cut*, rejecting events which lie far from the nuclear recoil population. The Run 119 implementation of this technique is described in [188], while the timing cut for the current analysis is discussed in Section 6.5.5.

Multi-parameter techniques

Primary risetime and delay represent only a tiny fraction of the information contained in a ZIP detector's athermal phonon pulses. In principle, we should be able to achieve better discrimination by incorporating more of this information into our analysis. This generally amounts to assembling several timing and distribution quantities into one or more composite discrimination parameters, then defining cuts on these discrimination parameters that accept nuclear recoils while excluding surface events. A variety of such multi-parameter techniques exist, and several have been applied to data from CDMS II's runs at Soudan.

- If we assume that the signal and background populations follow multivariate gaussian distributions in the space of phonon pulse parameters, we can use the χ^2 of each event with respect to the two distributions as discrimination parameters. Two versions of this technique, sometimes called "quadratic discriminant analysis," were implemented for the 2-Tower data run using `pminrtc`, `pdelc`, and `pfracc`: a straightforward version by R. Mahapatra and J. Sander [189], and one with a simple energy correction by me. The latter was used for the published analysis of the Si detectors from that run. This technique is described further in Appendix B.

- Artificial neural networks may also be used to discriminate multi-parameter event

populations. M. Attisha tested this technique during the 2-Tower run [190], incorporating information from a wavelet decomposition of the phonon pulses.

These techniques have proven promising, but thus far none has proven to be greatly superior to the simpler `pminrtc/pdelc` analysis described above. We believe that these techniques are limited by systematic variations in the parameters themselves with position and energy (even after correction), as well as the poor performance of ratio-based parameters such as `pfracc`. Future work may be able to account for these effects with improved correction techniques and choices of phonon parameters, but for the current analysis we have settled upon the simpler two-parameter discrimination technique.

Cut-free techniques

Rather than using cuts to define a background-free region of parameter space, we can instead attempt to estimate the total number of signal events present in the entire parameter space using a cut-free analysis. V. Mandic and coauthors [191] have described a method for doing this, with reference to a CDMS data set taken at the Stanford Underground Facility (Run 21). Given known distributions $f_m(\vec{\theta})$ for M classes of events (signal and $(M-1)$ types of background) in our multi-parameter space ($\vec{\theta}$ includes energy, ionization yield, and various phonon parameters), we can express the total event distribution as a sum:

$$f(\vec{\theta}) = \sum_{m=1}^{M} N_m f_m(\vec{\theta}). \tag{3.12}$$

Given an observed set of events $\vec{\theta_i}$ ($i = 1..N$), we can select the values of N_m which maximize the extended likelihood function:

$$L = \frac{e^{-(\sum N_m)}}{N!} \prod_{i=1}^{N} \left[\sum_{m=1}^{M} N_m f_m(\vec{\theta_i}) \right]. \tag{3.13}$$

We can estimate the various contributions of each species of events by maximizing this function over the various N_m. When applied to the Run 21 data, the authors found a $\sim 17\%$ improvement in sensitivity over the standard cut-based analysis.

This method is extremely powerful in principle, but due to lack of time has not been implemented for the current data set. The primary drawback of this method for CDMS is its dependence upon knowledge of the distribution functions $f_m(\vec{\theta})$. Without a quantitative understanding of each detector's phonon response, these functions must be fit

empirically from calibration data. Systematic errors in the determination of these functions limit the discovery potential of this technique, though they were not the dominant errors in the neutron-limited Run 21 analysis. Nonetheless, as knowledge of the detector response improves this will be an important direction for later analyses.

Chapter 4

The CDMS II Experiment

4.1 Introduction

The CDMS II installation at Soudan was initially commissioned from 2001-2003. This was a lengthy process, fraught with numerous cryogenic difficulties and hindered by the unique challenges of limited access at a remote experimental location. This initial commissioning is described extensively elsewhere [1].

The experiment was successfully commissioned for initial data runs with 1-Tower (Run 118) and later 2-Towers (Run 119), each leading to world-leading limits on WIMP-nucleon interactions [187, 192, 3, 5].

4.2 Soudan Underground Laboratory

Soudan Underground Laboratory is a subterranean physics laboratory located in Soudan Underground Mine, part of a state park of the same name in the Iron Range region of northern Minnesota. Soudan Mine was Minnesota's first iron mine, a rich source of ore from its founding in 1882 until its decommissioning in 1962. The laboratory is located on the 27th level, 2341 feet (713 m) below the surface. The CDMS installation is located in the cavern once used by the Soudan 2 proton decay experiment. The far detector for the MINOS neutrino oscillation experiment is located in an adjacent excavation.

The great depth of the Soudan site makes it extremely attractive to many forms of low-background experiment, particularly searches for rare events. Its depth of 2090 meters water equivalent (m.w.e.) reduces the cosmic ray muon flux by a factor of ∼50,000 from its

Figure 4.1: Entrance to the headframe of the Soudan Underground Laboratory.

value at the surface. This reduction in muons is accompanied by a corresponding reduction in muon-induced particle showers, particularly the neutrons generated in such showers. Since neutron-induced nuclear recoils cannot be distinguished from their WIMP-induced counterparts, direct detection experiments at depth have much lower background rates (and hence greater sensitivities) than similar experiments at the surface. CDMS itself provides a prime example of this benefit: the sensitivity of the first CDMS run at Soudan in 2003 (Run 118) was an order of magnitude greater than that of the last run at Stanford Underground Facility (17 m.w.e.) in 2001. Both runs used the same detectors and had similar exposures, but the shallower run observed 20 candidate events (consistent with expected cosmogenic neutrons) [193] while the deeper saw one (consistent with misidentified surface events) [192].

In addition to its scientific merits, Soudan also provides unique opportunities for education and outreach. The Minnesota Department of Natural Resources (DNR) runs several tours of the mine each day throughout the summer, including two daily tours of the physics laboratory. Since the "Boundary Waters" region is a very popular fishing and vacation locale, thousands of tourists tour the mine (and laboratory) each year. Many walk away having learned a little about our experiment and gotten a brief taste of the excitement surrounding our endeavor.

4.3 Experimental installation

The design of a low-noise, low-background cryogenic installation presents a number of special challenges. Millikelvin cryogenics demands a wide array of specialized support equipment: pumps, plumbing, thermometry, electronic control systems, etc. Much of this support equipment produces a great deal of electrical noise, and so must be isolated from the detector electronics as much possible. Moreover, traditional low-temperature and high-vacuum components are generally not made from low-activity materials. A dilution refrigerator itself is built primarily of steel, generally with an unacceptable level of uranium, thorium, and ^{60}Co contamination. The CDMS installation must therefore not only provide a complex cryogenic and electrical infrastructure; it must also protect the detectors from that infrastructure!

These challenges are exacerbated by the installation's remote, underground location. Physicists only have regular access to the mine from 7:30 a.m. to 5:30 p.m. each weekday, though emergency off-hours access is possible. Data acquisition and cryogenic operations continue day and night, however. This necessitates a complex system of automated monitors and control systems to keep the fridge operating without human intervention and allow it to be controlled by remote users, as well as a system of automated alarms to notify experts if something goes wrong. After some initial growing pains during the earliest runs at Soudan, this framework has operated nearly perfectly during the 5-Tower data runs.

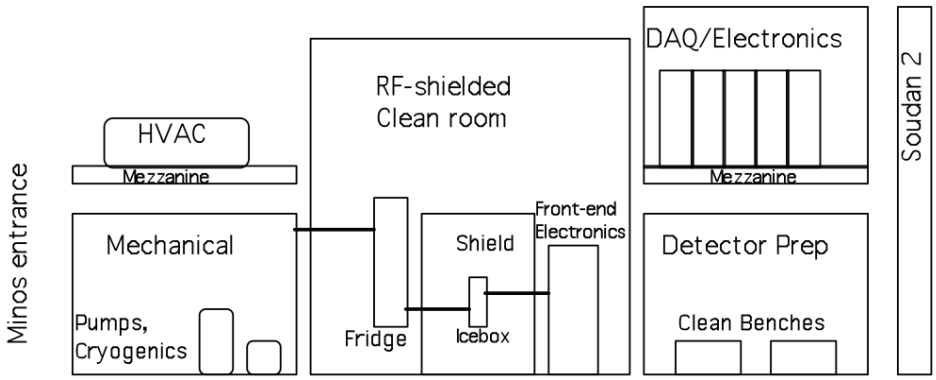


Figure 4.2: Schematic arrangment of the CDMS installation at Soudan. The mechanical support area ("cryopad") and detector preparation room lie on the main lab level, while the electronics room, air handling equipment, and offices reside on a second, "mezzanine" level. North is approximately to the left of this diagram.

The lab layout is illustrated schematically in Figure 4.2. The detectors, front-end

electronics, and primary cryogenics of the CDMS experiment are located in an RF-shielded, class-10,000 clean room. Pumps and other cryogenic support equipment are located on the "cryopad" mechanical support area outside the north wall of the RF room. Cryogen plumbing and readout wiring pass through this wall to control systems on the cryopad. Electrical connections for readout and electrical power pass through the RF room's south wall to the electronics room on the second "mezzanine" level. The mezzanine also carries the air handlers for the RF room, office space for CDMS physicists and engineers, and a ping-pong table to aid in the production of quality research.

4.4 Cryogenics

The detectors are cooled by an Oxford Instruments 400S ^{3}He-^{4}He dilution refrigerator, with a rated cooling power of 400 μW at 100 mK and a base temperature below 10 mK with no external load. The refrigerator is mounted atop a custom set of "tails", connecting its various temperature stages to a horizontally-displaced cold volume known as the "icebox". The icebox and tails are connected through the icebox's surrounding shielding (Section 4.5) by a concentric set of copper cylinders called the "fridge stem" or "F-stem". A matching "E-stem" extends through the shield opposite the F-stem. The E-stem carries the detector stripline wiring from the icebox to the "E-box'," a box-shaped structure which connects the striplines to external cabling through an array of vacuum-sealed D-connectors.

Most of the dilution refrigerator's support plumbing resides on the cryopad, outside the RF room. The circulation of the ^{3}He-^{4}He mixture is driven by a Pfieffer rotary pump and a Roots blower connected in series. Before entering the fridge the mixture is circulated through a series of three cold traps to clean it of impurities: a liquid nitrogen cold trap on the cryopad, an external liquid helium cold trap on the cryopad, and an internal liquid helium trap within the helium bath of the refrigerator itself. All three of these traps are cleaned regularly during normal operation, generally on a monthly interval. An Oxford Instruments Intelligent Gas Handling (IGH) unit controls and monitors the ^{3}He-^{4}He circulation loop. The IGH regularly reports pressures, valve settings, and cryogen levels to a dedicated Macintosh computer, which publishes this information to a web site for remote monitoring.

A Moore APACS automated industrial control system located on the cryopad monitors and directs numerous aspects of cryogenic operations, notably the daily refilling

of the refrigerator's liquid helium and nitrogen baths. Refills occur automatically at a user-set time each day, but APACS (or the user) can override this schedule to initiate a fill when the bath levels become low. APACS is interlocked to the data acquisition system, pausing the DAQ APACS also regularly records more than 150 temperature, pressures, flow rates, and other quantities; this provides a complete historical record of cryogenic operations that can be accessed from anywhere over the internet through custom Java GUIs.

The addition of Towers 3, 4, and 5 substantially increased the heat load on the icebox. The vast majority of this extra heat is dumped at 4K by radiation from 36 additional FETs and conduction through 19 new striplines (18 for detectors and one for icebox thermometry). Thermal modeling suggested that this heat load would be unacceptable. The increased boil-off rate of liquid helium would have necessitated more than two helium transfers each day, significantly reducing experimental live time and incurring a prohibitive cost in liquid helium. Worse yet, heat leakage through the Towers would have been sufficient to raise the detectors above the transition temperatures of at least some TES's, rendering many readout channels inoperable.

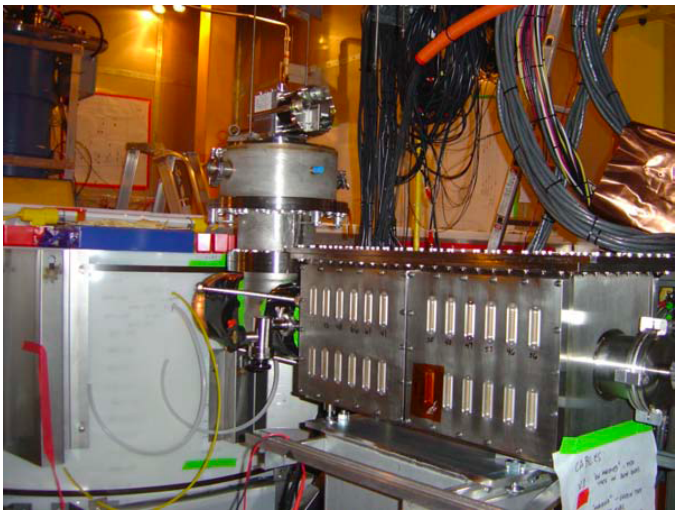


Figure 4.3: View of the CDMS II shield, E-stem and E-box (left to right), with detector cables unplugged. The double cylinder attached vertically to the E-stem is the cryocooler head, currently disconnected from its He flow lines. Bellows to aid in vibration isolation are visible along the E-stem (wide, dark segments of the E-stem.

A Gifford-McMahon cryocooler from Sumitomo Heavy Industries was added as a secondary refrigeration system during the 2005 system upgrade. The cryocooler is a closed-circuit helium refrigerator that provides 1.5 Watts of cooling power at 4K and 40

Watts at 77K without the need for external cryogens. The cryocooler consists of a two-stage drive head mounted on the E-stem as shown in Figure 4.3, driven by an external compressor outside the RF room. Helium gas is driven to the head at high pressure (~ 20 atmospheres) and compressed/expanded by a piston, providing cooling power. The unit has two cooling stages, cooling both the 77K and 4K stages of the E-stem. From its position opposite the main cryostat, the cryocooler intercepts the heat load of the striplines and prevents large temperature gradients across the 4K layer. With the cryocooler installed and all 5 Towers activated the cryogenic hold time is slightly greater than 24 hours, thus requiring only one cryogen fill per day. The cryocooler cold fingers are connected to the E-stem through flexible copper couplings designed to limit the transmission of mechanical vibration while maintaining excellent thermal conductivitiy. Section 5.2.1 gives further details of the cryocooler installation.

4.5 Shielding

The detector cold volume ("icebox") is surrounded by a multi-layer arrangement of active and passive shielding, depicted in Figure 4.4. This shield greatly reduces the rate of cosmogenic and radioactive particles – neutrons, photons, electrons, and alpha particles – within the detector array, thus reducing the "weight" placed upon data analysis to discriminate against these backgrounds.

The outermost layer of the CDMS shield is an active veto to tag muons and particle showers from outside the cryostat. The shield is formed by 40 overlapping scintillator panels in a hermetic configuration, punctured only by the cryogenic and electrical feed stems. Each panel consists of a 5-cm-thick sheet of plastic scintillator viewed by one or two 2-inch R329-02 photomultiplier tubes. The light response of each photomultiplier is checked daily during cryogen fills using LED pulses fed into each panel through optical fibers.

Within the scintillator shield, the detector cold volume is surrounded by several layers of passive shielding. 40 cm of polyethylene bricks form the outermost layer of the shield; this material serves to moderate incident low-energy neutrons sufficiently to prevent nuclear recoils above the detector energy threshold. Within the polyethylene lies 22.5 cm of lead, an effective shield against external gamma rays. Since lead's own radioactivity is significant, the inner 4.5 cm of this shield is composed of low-activity ancient lead, recovered from the ballast of a sunken ship near France. A further 10 cm of polyethylene within the

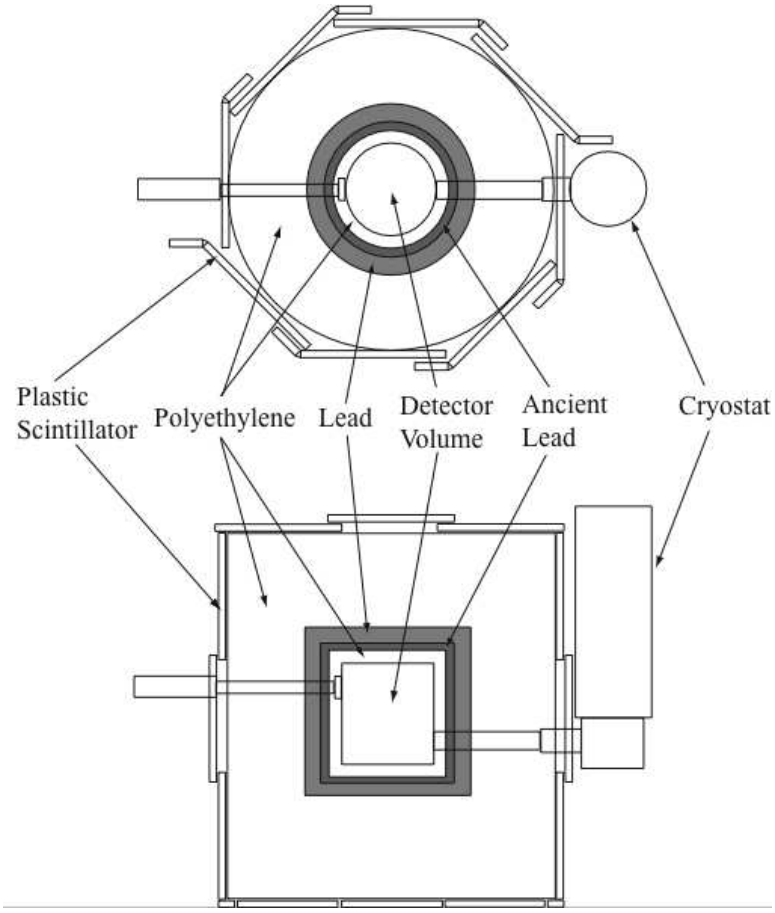


Figure 4.4: Cross-sectional diagram of the Soudan icebox, showing the configuration of the passive and active shielding surrounding the "icebox" detector volume. Also shown schematically are the E-box (*left*) and dilution refrigerator (*right*). Figure from [192].

lead gives additional neutron moderation, particularly against neutrons generated by fission and (α,n) processes within the lead itself. The upper portion of the entire shield may be removed by crane for access to the icebox.

Within the shield, the icebox itself consists of a set of concentric copper cans heat-sunk to the various temperature stages of the dilution refrigerator: mixing chamber ($\sim$ 40 mK), cold plate ($\sim$ 130 mK), still ($\sim$ 1 K), helium bath (4 K), nitrogen shield (77 K), and room temperature (300 K). The detector housings and the bulk of the cold hardware within the icebox are also made from high-purity copper due to its excellent radiopurity, which limits radioactive contamination near the detectors. The total thickness of copper surrounding the detectors is a few cm, sufficient to stop alpha and beta radiation from outside the cans. The copper cans are also surrounded with a mu-metal magnetic shield to protect the SQUIDs and TES's from external magnetic fields.

Early in the first Soudan data run in 2003, radon inside of the Pb shield was found to be a significant source of gamma ray background. The detectors themselves are kept under vacuum, but the radon level in the mine air is relatively high ($\sim$ 700 Bq/m^3). The lead shield provides adequate protection from air outside the shield, but any substantial volume of air within the shield can yield a significant gamma rate. A continuous purge of aged air to exclude mine air from the volume between the outer copper can and the mu-metal shield was begun in November 2003 [192]. This purge reduced the gamma rate within the detectors by a factor of > 4, as well as halving the observed beta rate. A similar purge has been used in the 5-Tower data runs, with additional injection tubes for better gas flow and using dry nitrogen rather than aged air as the purge gas.

4.6 Cold hardware

The ZIP detector array is connected to a system of cryogenic amplifiers, wiring, and support structures known collectively as "cold hardware." Figure 4.5 illustrates the cross section of the Soudan icebox and cold hardware. I describe each of the primary cold hardware components briefly below.

4.6.1 Detector housing

Each ZIP is housed within a hexagonal ring-shaped housing made of high-purity copper. The detector is mounted within its housing by a set of six flat Cirlex clamps. De-

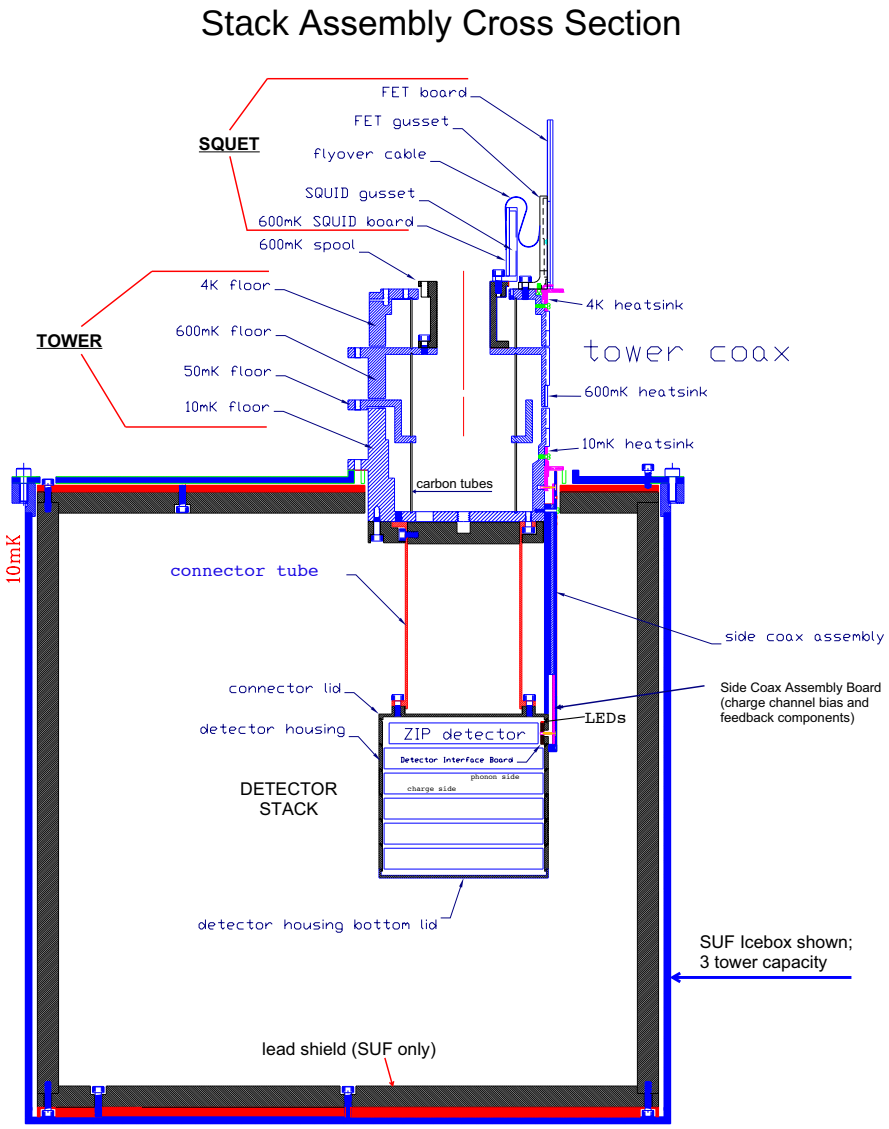


Figure 4.5: Cross section of the cold hardware within the innermost icebox can. Note that the icebox depicted is that of the Stanford facility previously used by CDMS; the Soudan icebox can house seven Towers and has no inner lead shield.

tectors are mounted together in stacks of six, each attached to a Tower support structure (see below). The housings within the Tower have no top or bottom lids, giving each detector face a full view of its neighbor's face 3.5 mm away. This unobstructed arrangement greatly increases the proportion of events, particularly events due to radioactive surface contamination, which scatter in two neighboring detectors; such events are easily rejected by an event multiplicity cut.

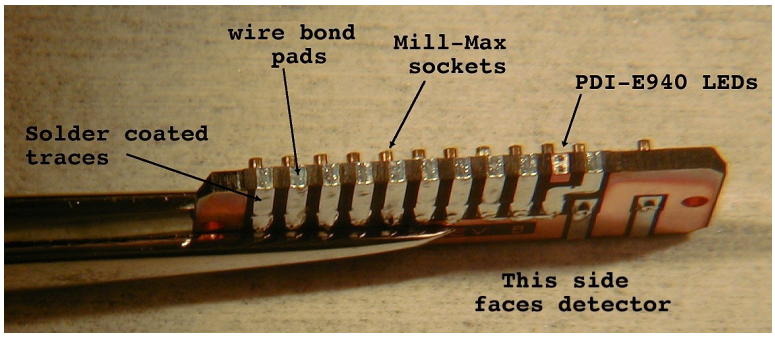


Figure 4.6: Photograph of a detector interface board (DIB), showing connection pads and LEDs. Picture and annotations by Dennis Seitz.

Each detector is linked to the outside world through a small detector interface board (DIB) located on one edge of its housing, mounted adjacent to the boundary between sensors A and D. The DIB plugs into one end of a side coax and couples to the detector through ten Al wirebond connections. The DIB also houses two infrared LEDs for clearing trapped charge in the detectors. Figure 4.6 shows a photograph of a DIB, with mounting pads, sockets, and LEDs marked. The LED shines on the phonon face of its own detector and the charge face of the neighboring detector.

4.6.2 Tower

The Tower, pictured in Figure 4.7(a) and shown in cross section in Figure 4.5, is a hexagonal copper structure that supports the detector stack and connects it electrically to the SQUET cards. The Tower consists of four copper stages, each heat-sunk to one of the icebox cans. A central graphite cylinder supports the four stages, holding them separate without excessive heat conduction between them. Each face of the tower carries the bias and signal wires of one detector, connecting the side coax at the base temperature stage to the SQUET card mounted at 4K. The wires are heat-sunk to three of the temperature

stages (base, still, and 4K) and held under tension in vacuum channels along the Tower faces to limit their sensitivity to mechanical vibration.

Note that the term "Tower" is often used to refer to the Tower, connector tube, and detector stack assembly as a unit. I follow this convention throughout much of this dissertation. In particular, I identify individual ZIPs based upon their Tower and stack position, *e.g.* "T3Z5" is the fifth detector (counting from the top) attached to Tower 3.

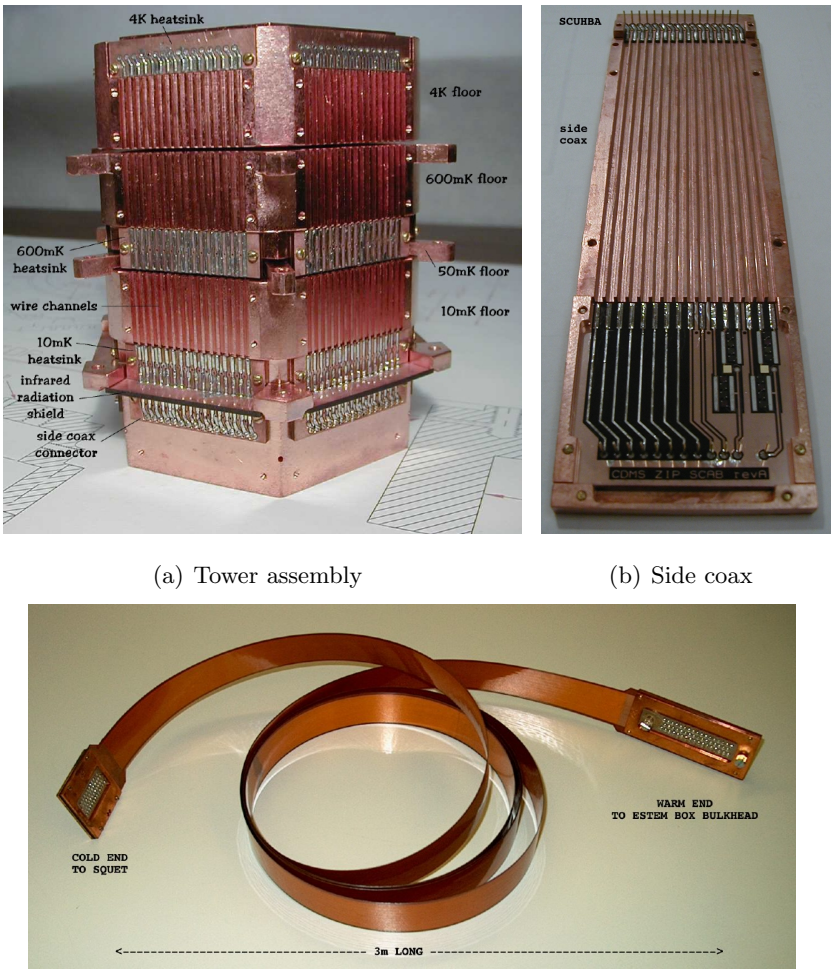


Figure 4.7: Photographs of assorted cold hardware components. Pictures and annotations by Dennis Seitz.

4.6.3 Side coax

Each detector is connected to the base temperature stage of its respective Tower face by a connector card called a side coax, shown in Figure 4.7(b). The side coax also houses the coupling capacitors and bias and feedback resistors of the detector's two charge amplifiers; housing these resistors at base temperature lessens their Johnson noise contributions to the ionization noise. Side coaxes are made in six distinct lengths so as to reach from the Tower to each of the six detector positions in the stack.

4.6.4 SQUET card

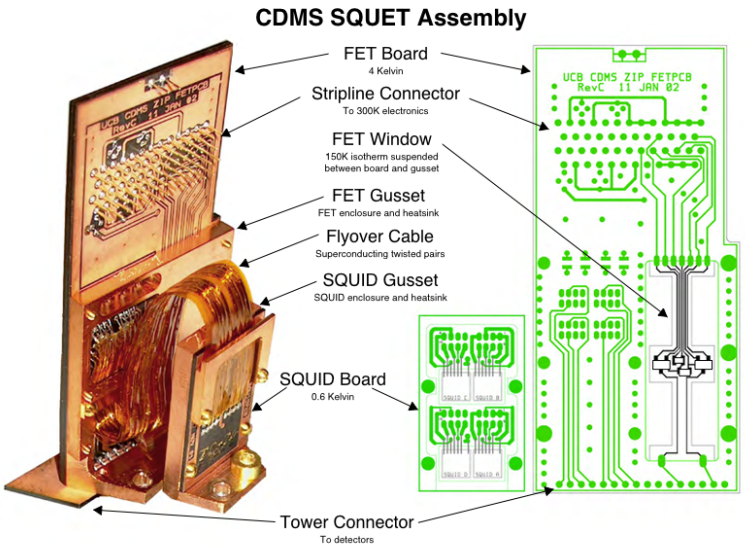


Figure 4.8: Photograph (*left*) and circuit board layout (*right*) of a SQUET card. Pictures and annotations by Dennis Seitz.

Each Tower face is topped by a SQUET ("SQUID and FET") card, shown in Figure 4.8. The SQUET is a combination of two printed circuit boards: a primary card housing the charge amplifiers' FETs and a subsidiary card containing the SQUIDs, shunt resistors, and input and feedback coils of the phonon amplifiers. The FETs are supported on a Kapton membrane within a copper gusset on the primary card. This arrangement enables the FETs to self-heat to $\sim$ 140 K (the carriers within them would be frozen out at 4K) while protecting the rest of the 4K stage from excessive infrared radiation from the hot FETs. The subsidiary SQUID card is heat-sunk to an extension of the still layer

of the Tower to improve SQUID performance and reduce the Johnson noise of the shunt resistors. The two cards are joined by a flexible "flyover" cable composed of twisted-pair superconducting niobium wires sandwiched between layers of Kapton tape.

4.6.5 Stripline

The warm and cold electronics are linked by a set of flat electrical cables called striplines, illustrated in Figure 4.7(c). Each is a 3 m flexible circuit made of Kapton and copper, consisting of 50 copper traces sandwiched between two copper ground planes and surrounded by kapton insulation. Each stripline connects the SQUET card of one detector through the E-stem to an E-box connector several meters away. The stripline bundle is heat-sunk at two thermal intercepts within the E-stem to limit heat flow between room temperature and the 4K stage. Once inside the icebox, the striplines fan out to reach the various SQUETs in a carefully-chosen arrangement shown in Figure 4.9.

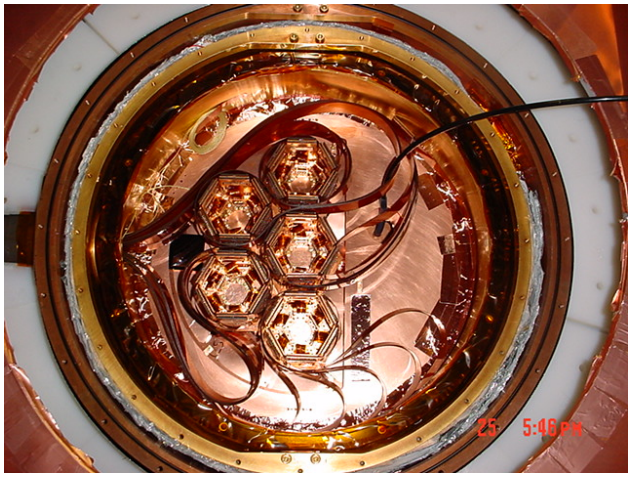


Figure 4.9: Final configuration of the Soudan icebox before the 5-Tower runs, showing the SQUETs and striplines atop the 4K lid. The E-stem and warm electronics lie to the left of the photo, while the dilution refrigerator lies to the right.

4.6.6 Icebox thermometry

The 5-Tower runs use an extensive set of thermometers to monitor the temperature profile within the icebox. Lakeshore thermometer diodes are mounted on the 4K stages of each of the five Towers, as well as on the 4K lid. Ruthenium oxide thermistors are mounted on each of the inner cans, as well as on each temperature stage of Tower 5. All of these

signals are read out to APACS through a single extra stripline from a custom connector within the icebox.

4.7 Warm electronics

4.7.1 Front-end boards

Signals from the E-box pass through 50-wire cables to a rack of front-end boards (FEBs) in the RF room, one for each detector. Each FEB is a 9U custom printed circuit board carrying the second- and later-stage amplifiers and other components of the charge and phonon readout circuits, as well as control circuits for the detector LEDs, charge biasing, SQUID biasing, etc. The FEB also contains driver amplifiers for each readout channel to buffer and further amplify these signals before digitization. The FEBs are controlled through a GPIB interface, which is linked from the electronics room through fiber optics.

4.7.2 Receiver-trigger-filter boards

Amplified signals from each FEB travel through the RF room wall to a corresponding receiver-trigger-filter (RTF) board in the electronics room. The RTF board conditions the FEB signals for digitization by adjusting their baselines and applying a 336-kHz 2-pole Butterworth anti-aliasing filter.

Each RTF board also generates five logic-level trigger signals that are used to actuate the digitizers. The five signals are denoted Qlo, Qhi, Pwhisper, Plo, and Phi. Plo is the primary event acquisition trigger, issued by a comparator which tests whether the sum of the four phonon signals exceeds its baseline by a software-defined amount, generally 3-4 mV. Phi is a similar test with a much higher threshold ($\sim$ 500 mV), intended for use in gated trigger configurations to reject high-energy events. Pwhisper is a threshold similar to Plo but often somewhat lower, intended for tagging multiply-scattered events. Qlo and Qhi are analogues for Plo and Phi for the summed ionization channels.

4.7.3 Data acquisition hardware

These trigger signals and comparable ones from the veto panels pass to a trigger logic board (TLB) which determines when to issue a "global" trigger, *i.e.* a signal to initiate event acquisition. The criteria for issuing such a trigger are user-defined, but include:

- A "detector trigger", initiated by a Plo trigger issued by any ZIP (except T1Z1, which has damaged phonon sensors)
- A "veto multiplicity" trigger initiated by a simultaneous hit on multiple veto panels (used for muon background characterization)
- A "random" trigger issued intermittently by the data acquisition system to measure the experimental noise.

Upon receipt of a global trigger, the six readout channels of each ZIP are recorded by an array of 14-bit Struck SIS 3301 analog-digital converters and written to disk. These digitizers operate at a 100 MHz sampling rate, but 80-sample sequences are internally averaged to yield an output rate of 1.25 MHz with a corresponding reduction in digitizer noise. When triggered, each digitizer records a 2048-sample (1.6 ms) trace. The trigger itself occurs in the 512th bin, so each trace includes 409 μs preceding the trigger.

Photomultiplier signals from the 40 scintillator veto panels are processed by an analogous set of boards. The raw veto signals are extremely short in duration ($\sim$ 10s of ns); these signals are reshaped by a pulse-stretching filter network before digitization. The reshaped veto pulses are recorded by an array of 12-bit Joerger VTR812 ADCs, each acquiring 1024 samples at 5 MHz. A set of comparators also issues trigger signals from the veto shield; these are FED to the trigger logic board for evaluation of global triggers.

All of the above trigger signals – phonon, charge, and veto – are also recorded by a set of Struck SIS 2400 time-to-digital convertors (TDC's). These boards record the time stamps of each trigger in a circular buffer with 1 μs resolution. A portion of this trigger history buffer is recorded along with each event written to disk, giving a record of the several triggers immediately preceding and following the recorded trigger.

In addition to the event-specific data record described above, various experimental quantities are monitored by a slower data-acquisition system. This "slow DAQ" records the average values of the various signal channels and the mean trigger rates once each minute; the former is monitored for evidence of a SQUID losing flux lock, the latter for sudden changes that indicate changes in detector noise. All of these quantities are written to a separate file on disk for use in later analysis.

4.8 Data acquisition software

The Soudan data acquisition system (DAQ) is controlled by an interlocking system of Java and C++ programs run on a small cluster of computers in the Soudan electronics room. The DAQ software was written and maintained primarily by collaborators at the University of California at Santa Barbara. The system's overall design and various software components are described in detail by J. Sander [189].

The DAQ software is highly modular, with each major function (GPIB communication, slow DAQ monitoring, high voltage control, etc.) handled by a dedicated server program. The various specialized servers communicate with one another through the CORBA object-oriented network messaging framework. Each major function is accessible through a graphical user interface (GUI), each in turn accessed from a "master GUI" called RunControl. RunControl is a cross-platform network application written in Java; Figure 4.10 shows its primary control window. When run from the mine or surface building, RunControl allows full control of the data acquisition system. Experimental control from other remote locations is generally not allowed, but the RunControl GUI may be used remotely to monitor the experiment's status messages and settings.

As configured for the current data run, the data acquisition system can acquire event traces from all detectors at $\sim$ 20 Hz, far in excess of the $\sim$ 0.3 Hz event rate during WIMP-search operation. Calibration data may be taken in a new "selective readout" configuration at rates of up to $\sim$ 70 Hz, permitting shorter calibration data runs and more WIMP-search live time. The selective readout upgrade is discussed further in Section 5.2.1.

4.9 Data handling

Raw event data acquired by the DAQ software is stored to local disks in the mine. As each raw event file is completed, a set of Perl scripts compresses the file, writes a copy of the compressed file to a digital backup tape in the mine, and transfers it to the surface. Another Perl script maintains adequate disk space in the mine by automatically deleting old data files which have been copied to the surface and to tape.

At the surface, the data are stored on disk at the Soudan Analysis Cluster (SAC), a cluster of 30 1-GHz Pentium dual-CPU computers running Linux. The SAC processes the incoming raw data through our data reduction package to produce a preliminary set

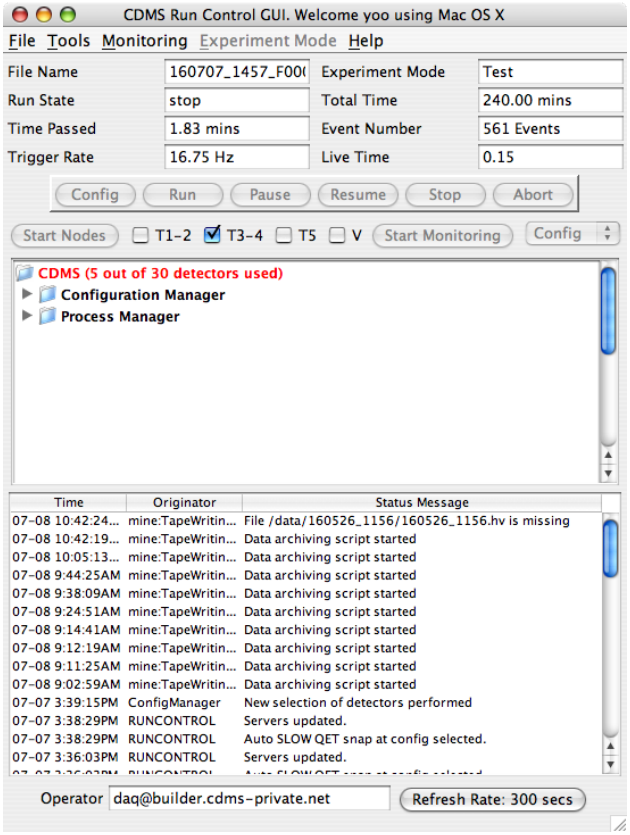


Figure 4.10: Graphical user interface for RunControl, the primary control program for the Soudan data acquisition software.

of reduced quantities. These preliminary quantities are used by automated data quality checks (*e.g.* KS tests) and for any *ad hoc* manual checks of the experimental performance. The SAC also records the raw data to tape for additional security.

Due to the increased demands of the 5-Tower data set, primary data processing occurs at Fermilab on the FermiGrid computing cluster. Raw data was transferred to the cluster from the mine on digital tape during Runs 123 and 124; the raw data is transferred over the internet in recent runs. The final quantities used in this analysis were generated on this cluster and processed with ~ 100 CPUs at a time.

Chapter 5

Commissioning the 5-Tower Run

After the completion of the 1- and 2-Tower runs, the collaboration made numerous changes and upgrades to the Soudan experimental installation. This Chapter summarizes the commissioning process of the 5-Tower data runs at CDMS. I focus primarily on changes from previous data runs; for details on the initial commissioning of the Soudan site, see [192, 177, 1].

5.1 Detector testing

The 5-Tower runs incorporate three new Towers containing a total of 18 new ZIP detectors, added to the 12 detectors in place during Runs 118 and 119. Each of these detectors underwent an iterative process of testing and performance characterization at one of two "test facilities" (Berkeley and Case Western) before being checked out for installation at Soudan [194]. The timely fabrication and testing of these devices was a very labor-intensive process, entailing heroic efforts from the groups at Stanford, Berkeley, and Case Western during 2002-4.

In particular, my own involvement in CDMS began nearly in coincidence with the most intense period of detector testing work at Berkeley. The testing and commissioning of Towers 3-5, encompassing more than 30 dilution refrigerator runs at Berkeley, was my first introduction to CDMS and the major task of the early portion of my graduate career. The work described below was completed as part of the Berkeley test facility team, in close collaboration with the test facility team at Case Western and the fabrication group at Stanford. Many of the statistics describing the detector characterization runs are adapted

from an internal summary prepared by Nader Mirabolfathi.

5.1.1 Electrical checks

Each detector underwent a series of electrical continuity checks before installation in a dilution refrigerator. Dilution refrigerator runs are extremely costly in time, manpower and cryogens, so there is a great advantage to catching electrical failures in the detectors or cold hardware before a full cool-down.

A program of continuity checks on each detector was performed at various temperatures and stages of cool-down. Note that each of the basic detector tests labeled "Berkeley" below could also be performed in a similar manner at Case Western Reserve University, the other CDMS II test facility.

- **Stanford:** Before distribution to a test facility, each detector was checked for electrical continuity at room temperature and at 77K in a liquid nitrogen test dewar ("Genietto"). Many gross failures (e.g. open connections or damaged LEDs) are apparent at room temperature. The test at 77K is significant because of the low resistance of Ge at room temperature. The four phonon channels are electrically connected to one another through the crystal at room temperature, so many electrical shorts are only apparent at lower temperatures. Unfortunately, many common shorts took the form of thin flakes of Al from photolithography failures. Such Al shorts have large resistances at room temperature, and are only detectable below the superconducting transition temperature of aluminum ($T_c = 1.2$K).

- **Berkeley - bench:** Before installation in the dilution refrigerator, we performed a bench test of electrical conductivity on each detector to guard against damage in transit. Each detector was also checked again after assembly into a stack, ensuring good connections to and within the side coaxes and tower.

- **Berkeley - fridge:** After installation in the dilution refrigerator, each detector was tested at four points during the cooldown process: 300K, 77K, 4K and base temperature ($\sim 10mK$). These tests identified poor connections to or within the cryostat wiring, notably intermittent failures in the striplines which came and went with varying temperature.

Al shorts were a common failure mode of CDMS II ZIP detectors, unfortunately, affecting more than half of all detectors. Though they were generally reparable through manual surgery at Stanford, such shorts were the most frequent cause of aborted dilution refrigerator runs. Such failures have been reduced by a switch from step-and-repeat to whole-field photolithography for SuperCDMS (see Section 8.4.1), but this change was not implemented until after the conclusion of CDMS II fabrication.

5.1.2 TES characterization

Once a detector passed electrical checks and cooled to millikelvin temperatures, it was subjected to a series of tests to characterize its phonon sensor response. Each ZIP went through at least two rounds of testing in two distinct fridge runs. The first round characterized the TES sensors through the series of tests described below. The detector then returned to Stanford for ion implantation (see below) to re-tune the TES transitions as needed. A second round of tests in a dilution refrigerator either confirmed the sensor performance or (in some cases) informed a second round of re-implantation.

T_c measurement

The performance of a ZIP detector depends crucially upon the superconducting transition characteristics of its transition edge sensors (TESs). The transition temperature (T_c) of each phonon sensor should be $\sim 70 - 80$ mK in order to ensure good electrothermal feedback. Higher transition temperatures yield poorer signal-to-noise characteristics, while lower transition temperatures lead to deviations from the strong-ETF regime. These transition temperatures should also be relatively uniform across the detector surface. Variations in T_c across a single sensor quadrant can widen the effective transition width of that sensor's array, leading to reduced sensor responsivity (lower α) and thus deviations from the strong-ETF regime. Variations in T_c between sensor quadrants lead to nonuniform energy and pulse shape response for events at different detector positions, demanding more from the phonon correction table.

In addition to the transition edge criteria described above, we also demand that the normal resistance of each phonon sensor (its electrical resistance when superconductivity is destroyed by elevated temperature or electrical current) be within a nominal range. For our nominal TES configuration (1036 TES's in parallel, each $\sim$ 1.25 kΩ), each sensor

should have a normal resistance near 1.2 Ω. A substantially higher resistance may indicate poor TES deposition, in the form of "holes" in the TES array or narrowed TESs through over-etching. Since such problems are generally impossible to repair without repolishing the detector, some detectors with significant "bad patches" were eventually installed at Soudan.

To perform the measurement itself, we drive a small (few mV), low frequency ($\sim$ 100 Hz) triangle wave down the TES current bias line and measure the resulting input coil current using the SQUID amplifier. This gives enough information to compute the sensor resistance R_s ($R_{sh} = 25$ mΩ, $R_b = 1200\Omega$, $R_s \approx 0.2 - 1\Omega$):

$$I_s = \frac{V_s}{I_s} = \frac{V_b}{R_b}\frac{R_{sh}}{R_s}\left(1 + \frac{R_{sh}}{R_s} + \frac{R_{sh}}{R_b}\right)^{-1} \approx \frac{V_b}{R_b}\frac{R_{sh}}{R_s}.$$

To map out the resistance as a function of temperature we perform a series of such measurements at different detector temperatures, controlled using a heating resistor on the refrigerator's mixing chamber. Due to the substantial thermal time constants of the fridge, a full scan of detector temperatures required approximately 2 hours during this round of detector testing. As an example, Figure 5.1 shows the measured T_c curves of detector G26 (later installed in position T5Z4 at Soudan) before and after successful ion implantation.

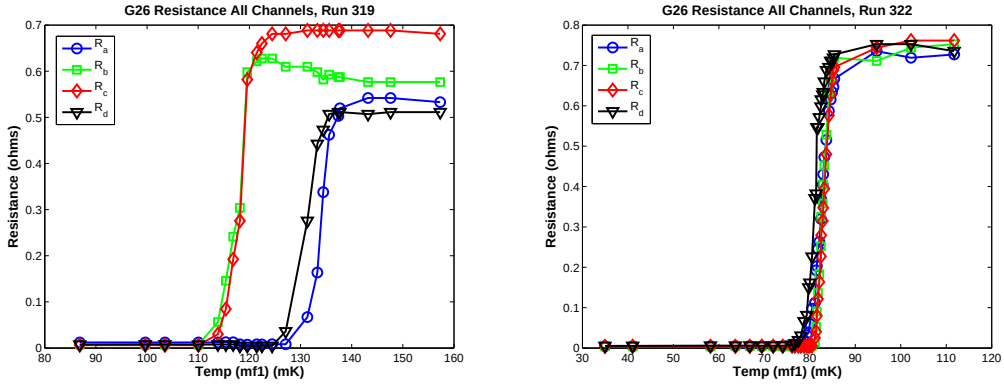


Figure 5.1: Sensor resistance as a function of temperature from Berkeley test runs of detector G26 (T5Z4 at Soudan). *Left:* Before ion implantation, from UCB Run 319 (June 2004). *Right:* After successful ion implantation, from UCB Run 322 (July 2004).

The Berkeley cryostat has since been upgraded to use a heater resistor at the tower level. This allows for much finer and faster control of the detector temperature, reducing hysteresis and greatly improving the temperature scan time. This has made later detector tests much faster, but was not available in time for this round.

I_c measurement

The T_c measurement described above is not always adequate to determine the transition characteristics of a parallel TES array. If a portion of a sensor's TESs have superconducting transitions significantly lower in temperature than those of the rest of the array, for example, their transition to normal conduction may not be apparent on a resistance curve while the rest of the array remains superconducting. A measurement of the critical current of each sensor as a function of temperature gives complementary information which can be critical to defining a robust ion implantation mask.

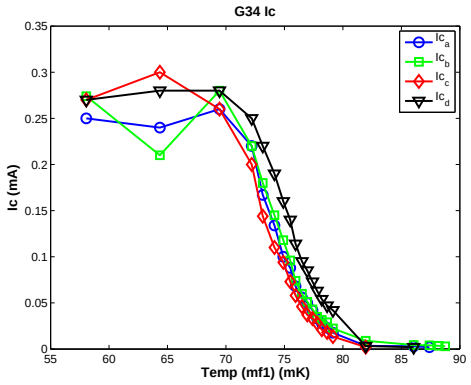


Figure 5.2: Sensor critical current (I_c) as a function of temperature for detector G34 (later installed at Soudan as T4Z5), taken during UCB Run 316 (April 2004).

The I_c measurement is accomplished by determining the largest triangle wave which can be driven down the sensor bias line without distortion for various detector temperatures. An example plot for T4Z5 is shown in Figure 5.2.

$I_b - I_s$ measurement

The final measurement performed on each sensor array is a DC measurement of the sensor current (I_s) as a function of the bias current (I_b). From these data the full current-voltage relation may be reconstructed for each phonon sensor, providing further input into the implantation mask. This technique has been described in Chapter 4 of Tarek Saab's dissertation [195]. Several hours are generally required to perform this measurement due to the long time constants involved, but substantial automation work by Kyle Sundqvist and the introduction of a Tower-level heater has substantially improved the throughput. Figure 5.3 shows an example from detector G26 (T5Z4).

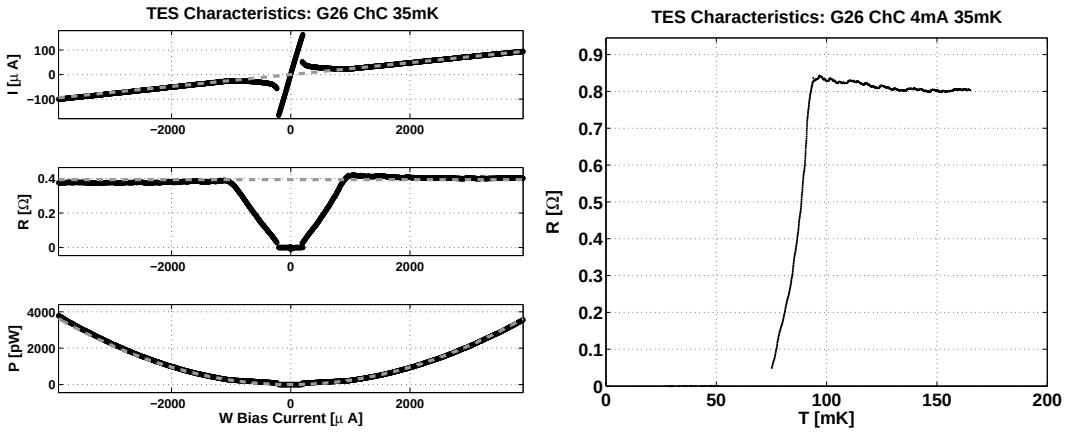


Figure 5.3: Results of the $I_b - I_s$ analysis of phonon sensor C of detector G26 (T5Z4). *Left:* From top to bottom, input coil current, TES resistance, and TES power dissipation as functions of the channel bias current. Light, dotted lines show the fits used in sensor modeling. *Right:* Reconstructed relationship between resistance and TES temperature.

Ion implantation

The results of the tests above were used to develop a map of the transition temperatures of the TESs across the surface of the detector. Based upon this map, the detector is then implanted with Fe ions in order to correct the TES transitions. After ion implantation each detector underwent a second round of the above tests to verify its performance (or occasionally to characterize it for further implantation or repair). Figure 5.1 shows the result of a successful implantation: uniform, sharp transitions near 80 mK.

Some detectors required only one round of the testing above, but many detectors required a second round of ion implantation and testing to tune the T_c's to the desired level. Detectors that completed the testing program successfully were stored in a radon-purged cart at the Stanford Underground Facility. Those deemed inadequate for Soudan were repolished for another round of photolithography, beginning this process all over again.

5.1.3 Tower selection

Towers 3, 4, and 5 were constructed at Stanford in February, July, and September of 2004, respectively. The detectors composing each Tower were the best-performing which had completed testing at the time. Two major factors were considered in determining the stacking order of the detectors within the Tower:

1. The expected *radon exposure* of each detector. Detectors with high radon exposures are predicted to have high beta rates due to ^{210}Pb plate-out on detector surfaces.

2. The observed performance of each detector's phonon sensors, as evaluated using T_c and I_c measurements and short data runs at the test facilities.

Detectors performing poorly by either metric – high contamination or poor phonon performance – were relegated to the Z1 and Z6 positions of the Tower whenever possible.

The approximate radon exposure of each detector was evaluated based upon records of its location history throughout the fabrication and testing process. Each detector's radon exposure are assumed to be reset to zero at its last repolishing, and radon levels at each of 15 locations are assumed to match on-site measurents taken *circa* 2001. The total exposure is converted to a predicted beta rate using a simplified model developed by L. Baudis, based upon a 0.5 m column height of air and a 1% plateout rate of radon decays within that volume. Figure 5.4 shows the resulting rate estimates, following calculations by L. Baudis and me. The general downward trend in exposure for the later detectors is primarily due to a change in handling procedure: a polyethylene-shielded purge cart was constructed in late 2001 for long-term storage of detectors at Stanford under a nitrogen atmosphere, rather than room air. These estimates suffer from enormous uncertainties, primarily due to the crude plateout model and the poorly-characterized environment at the repolishing facility. The predicted surface event rate is roughly an order of magnitude higher than that observed in this WIMP-search data run, described in Section 7.2.1. The qualitative trend is nonetheless in reasonable agreement: the surface event rates of Towers 3-5 are substantially lower than those of Towers 1 and 2, confirming the importance of the purge cart.

The stacking orders of Towers 3 and 4 were chosen primarily based upon these very rough estimates. In each case we endeavored to keep the "cleanest" detectors in the interior portions of the Tower, as is apparent in Figure 5.4. Tower 5 did not follow this pattern; all Tower 5 detectors were expected to have low radon rates, but several showed relatively poor phonon sensor performance. Tower 5's stacking order was thus set primarily based upon sensor quality, placing the problematic G07 and G24 in the Z1 and Z6 positions, respectively.

The 5-Tower data runs described in this dissertation yield an excellent sample of data with which to evaluate this contamination model. The estimated ^{210}Pb contamination

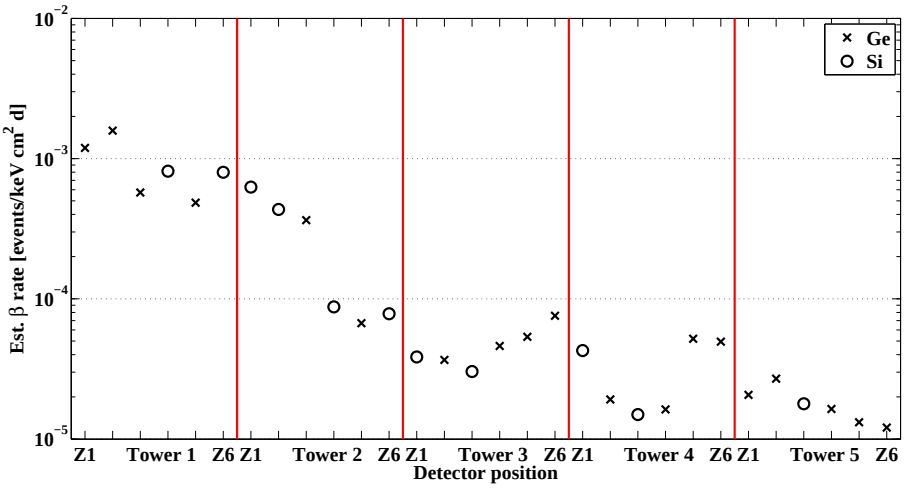


Figure 5.4: Comparison of expected surface beta rates on each detector, based upon their estimated Rn exposures and a crude model of Rn-daughter deposition from the air.

rates of each detector are discussed in Section 7.2.1. Some general trends in contamination do match the expectations of this exposure model, notably the reduced contamination of Towers 3-5, but the data do not strongly support the pattern of smaller variations predicted between detectors.

5.1.4 Tower checkout

After construction, each Tower underwent a series of system-level checks. The Tower was mated to its corresponding set of SQUET cards and mounted in a dilution refrigerator at one of the two test facilities. A short data run was taken with each of the six detectors to evaluate its performance and that of its cold hardware. Any faulty cold hardware components were swapped out for replacements. This process required at least three refrigerator runs, since neither facility had enough working striplines to test all six detectors at once.

The last of the new Towers was given a clean bill of health by November, 2004. When completed, the Towers were shipped to Soudan by truck to await final installation in the icebox.

5.2 Soudan commissioning

After the completion of the initial 1- and 2-Tower runs at Soudan in August, 2004, the collaboration began a series of electronic and cryogenic upgrades to prepare the Soudan installation for the 5-Tower runs. Much like the initial commissioning of the Soudan site (lovingly chronicled by Vuk Mandic [1]), this re-commissioning proved far more difficult than anticipated. All told, the process consumed more than two years, from the conclusion of the last 2-Tower data set on August 9, 2004, to the commencement of the first data run of this analysis on October 21, 2006.

This Section summarizes the central trials and tribulations of these two years, culminating in the successful data runs described in this dissertation. I was personally inspired by the perseverance, dedication, skill, and insight shown by my collaborators during this difficult time for our experiment. The success of this re-commissioning (and thus of the results of this dissertation) was a true team effort, brought about through countless hours of work at the mine and at all of the CDMS institutions.

5.2.1 Infrastructure upgrades

The addition of Towers 3, 4, and 5 to the icebox more than doubled the target mass of the CDMS experiment. This new, larger experiment places proportionally greater demands upon the Soudan infrastructure. 18 new detectors meant 108 new readout channels, including 72 new SQUIDs and TESs to tune. The 18 additional striplines and 36 additional FETs more than double the two dominant heat loads on the 4K stage of the icebox, adding stress to the cryogenic system. The additional channels also increase the data size of a single event to $\sim$ 800 kilobytes, increasing DAQ dead time and demanding increased bandwidth and storage.

Below I outline several of the upgrades implemented for the 5-Tower runs. These were planned upgrades to address the demands of increased target mass and improve cryogenic stability. Some of the relevant systems are described further in Chapter 4. Later sections detail several additional enhancements implemented as part of troubleshooting.

OVC vacuum improvements

The 2-Tower runs suffered from generally poor vacuum quality in the outer vacuum chamber (OVC) due to a known leak to the LHe bath. Even with continuous pumping

on the OVC, the experiment was prone to occasional "burps": catastrophic softenings of the vacuum due to runaway boiling of helium leaked into the OVC from the bath. The helium bath boiled away dramatically during a burp, warming the detectors and boiling the dilution refrigerator's ^{3}He-^{4}He mixture; recovery generally took several hours. Burps occurred every 3-4 weeks, generally during cryogen fills. With two cryogen fills each day, they added substantial stress to the lives of the scientists on shift.

To improve vacuum quality, Fermilab engineers modified the dilution refrigerator tails to attach a large (6-inch diameter) pump-out line to the OVC. The new line traveled ~ 2 meters through the RF room wall to a dedicated turbo pump. This line greatly increased the pumping speed available on the OVC. The seals at the OVC-bath juncture were also tightened, reducing the leak itself.

Gifford-McMahon cryocooler

During earlier runs at Soudan the dilution refrigerator suffered from a relatively poor cryogenic hold time, requiring refills of liquid helium and nitrogen every ~ 12 hours. This high cryogen boil-off rate resulted from the substantial heat load at 4K from FET heating and stripline conduction. The frequent cryogen fills made the maintenance of the Soudan cryogenic system significantly more laborious and expensive. More importantly, thermal model estimates suggested that the additional heat load of three new Towers and the associated wiring would be too great for the existing cryogenic system to handle.

In order to maintain adequate hold time, a Gifford-McMahon cryocooler from Sumitomo Heavy Industries was installed on the E-stem in 2005. This system is described further in Sections 4.4 and 5.2.3. The price of this additional cooling power is the large amount of mechanical vibration introduced by the cryocooler piston. The piston action is very noticeable within the RF room, roughly akin to the impact of a large hammer on the E-stem every 1.2 seconds. The detector configuration is relatively immune to vibrational noise, but some significant effects of these vibrations are noted below.[1]

New E-box

Due to leakage issues with the previous electronic feedthrough box ("E-box"), a new E-box was fabricated at Fermilab for use during the 5-Tower runs. This initially

[1]There is no definitive evidence that the cryocooler's loud, regular "tweeting" sound induces mental instability in nearby scientists.

showed superior leak performance than the previous E-box, but significant leakage rates were soon seen in this one as well. As with the previous E-box, we sealed many of the more troublesome seams on the new E-box with Apiezon Compound Q vacuum sealant to limit leakage. In order to simplify pumping out the E-box and measuring pressures near the leaky box, an auxiliary pumpout line and new temperature gauges were attached to the IVC and OVC at the E-box.

GPS time

The NuMI neutrino beam from Fermilab to Soudan was activated on March 4, 2005, marking the beginning of the MINOS neutrino oscillation experiment [196]. The MINOS experiment uses two iron calorimeters – a smaller "near" detector at Fermilab and a larger "far" detector in the Soudan mine – to monitor the composition of this beam over a 735 km baseline. The beam is generated in 10 μs pulses every 2.5 seconds, for a total duty cycle of 4×10^{-6}. Event timings at the two locations are synchronized using the Global Position System (GPS) time signal.

In principle, the NuMI beam can slightly increase CDMS's background rate. Incident muon neutrinos from the beam can generate muons in the cavern walls, which may in turn spallate neutrons from the rock and other surrounding materials. The vast majority of these neutrons will not produce events in our detectors, and $> 90\%$ of those that do will be coincident with activity in the scintillator veto shield. Dan Bauer has estimated that CDMS will see only one veto-coincident event in ~ 3 years of exposure.

Despite this very low rate, the 5-Tower DAQ system was upgraded to monitor the timing of events with respect to the NuMI beam spills. To do this, we transmit time signals from a GPS receiver in the MINOS cavern to the CDMS electronics room over 250m of fiber optic line. The measured one-way time delay across this link is ~ 1450 ns, consistent with time-of-flight and small delays in the LuxLink BNC-fiber convertors. This time signal is interpreted by a Symmetricom BC635 VME module, and the DAQ system records a GPS time stamp from this module with each event acquired. Section 6.3.7 describes the data quality cut created from this time stamp.

Computing infrastructure

The Run 123 and 124 data sets – calibration and WIMP-search – total 5.3 terabytes (TB) of raw data and 1.1 TB of processed "reduced quantities." Several changes were made to the computing infrastructure at Soudan (and elsewhere) to better handle this data stream:

- The mine's network infrastructure was upgraded substantially to handle the increased demands of MINOS and CDMS. The mine-surface connection was upgraded to gigabit ethernet, the surface-internet connection to a T3 line.
- Raw data is now compressed before transfer to the surface.
- The SAC surface computing cluster was upgraded to an array of 30 dual-CPU machines, recently retired from the CDF computing cluster at Fermilab.
- Primary data processing was moved to Fermilab's FermiGrid computing cluster, leaving the SAC dedicated to real-time evaluations of data quality. This greatly increased the rate of data reprocessing when the final calibrations became available.
- Selective readout was implemented for data acquisition during high-rate calibrations (see below).

Selective readout

The digital record of a 5-Tower event is roughly 2.5× larger than that of a 2-Tower event, and so requires $\sim 2.5\times$ longer to transmit over the network and write to disk. This increases the DAQ's dead time – the time required to reset the system to accept a new global trigger – and thus reduces the maximum rate at which the DAQ can acquire events. This limitation is unimportant for WIMP-search data runs, which have an event rate of only ~ 0.2 Hz, but is a significant restriction upon calibration data runs. A typical calibration run with a ^{133}Ba source can have a recoil rate of $\gtrsim 200$ Hz, 10× the rate at which the DAQ can actually acquire data from 30 detectors. This limitation requires us to take calibration data for a longer period of time (and thus lose more WIMP-search live time) in order to acquire a given number of calibration events.

To ameliorate this, the CDMS triggering system was upgraded to allow "selective readout" for this data run. In this configuration the DAQ only records the digitizer output for those detectors which issued phonon triggers for a given event, rather than for all

detectors and the veto shield as is usual. Since only a few detectors ($\sim$ 2 on average) trigger for each event, this greatly reduces the amount of data required to record an event. This not only saves disk space but also reduces the DAQ dead time associated with data transfer, allowing for much faster data acquisition rates at the same total data throughput. Ba calibration data sets are now taken in selective readout mode, enabling acquisition rates of $\sim$ 70 Hz (up from $<$ 20 Hz in Run 119) and reducing the amount of live time spent on calibration. To be conservative, WIMP search and Cf data sets are still acquired in standard readout mode.

5.2.2 Initial installation and fridge blockage

Towers 3, 4, and 5 were installed in the icebox during October and November, 2004. After a few initial difficulties with damage due to electrostatic discharges while working atop the polyethylene shield, the SQUETs and striplines were installed in late January, 2005, and the fridge began cooling in February.

The February cooldown was aborted due to a severe blockage in the circulation loop. After several failed attempts to clear the plumbing with gas flows, the fridge was warmed to room temperature and disassembled for repairs. The cryostat from its tails and opened to bare the dilution unit plumbing, illustrated in Figure 5.5. We remained in close contact with representatives of Oxford Instruments during the dissasembly, who on March 16th disclosed to us the presence of a small copper sinter between the 1K pot and the primary impedance line, also shown in Figure 5.5. The next day, Betty Young and Dan Bauer found and extracted the minute flake of sinter which had come loose and blocked the primary impedance line. The reassembled cryostat was run through a succesful thermal cycle without the icebox in April.

The re-mating of the fridge to its tails consumed the next month of work at the mine. Tight tolerances render this operation extremely delicate. Since the icebox and associated stems span several meters, thermal contraction shift the cans by several millimeters during cooldown. The system is designed to take up this slack, but due to the long lever arms involved the fridge must be aligned with each of the concentric icebox tails to few-mil accuracy. After many days of painstaking work with micrometers and dial indicators, the system was ready for another attempt by mid-summer.

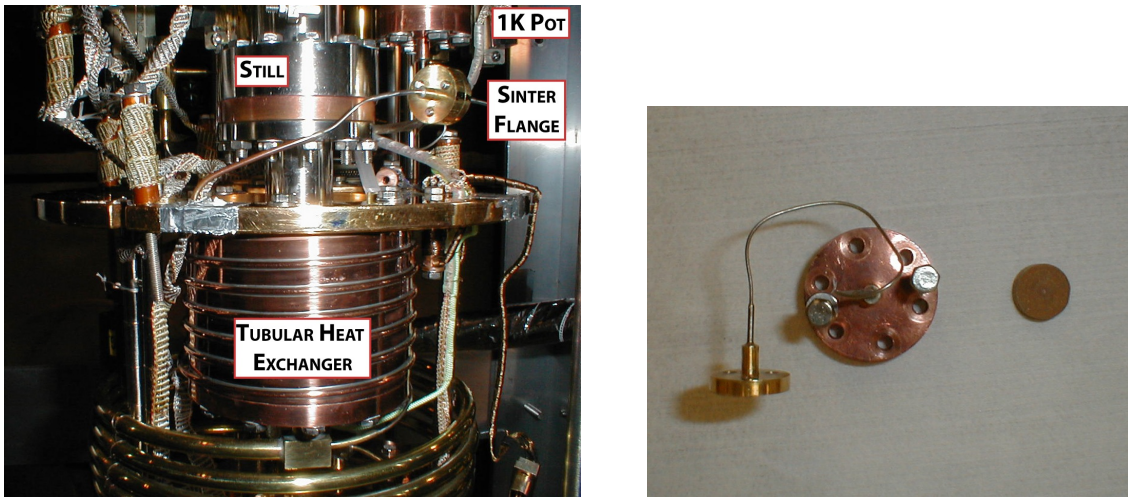


Figure 5.5: *Left:* Annotated photograph of the still region of the Soudan dilution refrigerator. The blocked primary impedance connects the sinter flange to the bottom of the 1K pot. *Right:* Close-up of the impedance tube connecting the brass sinter flange to the copper floor of the 1K pot. The copper disk is a sinter housed within the sinter flange.

5.2.3 First icebox run

The first successful cooling of the refrigerator including the 5-Tower icebox lasted from July through September of 2005. The dilution refrigerator's mixing chamber achieved a temperature of < 25mK and thermometry on the attached icebox can reached 40mK. After initial celebration, however, the crew was faced with two major problems:

1. The noise on several of the ionization channels was prohibitively high. This excess noise vanished when the cryocooler was turned off, as shown in Figure 5.6. The blame was soon assigned to the cryocooler's massive mechanical vibrations, presumably picked up microphonically through the capacitors and stretched wires of the cold hardware.

2. Even more seriously, none of the phonon sensors showed any event pulses. Despite the low temperature of the surrounding icebox can, the sensors responded in all ways as if they were normal (non-superconducting). Based on studies of the phonon noise spectra, Kyle Sundqvist and Bernard Sadoulet concluded that the TES temperature was somehow elevated to 200-300mK.

After a variety of thermal tests, we warmed the cryostat to fix a vacuum leak and replace the cryocooler coupling with something less vibrationally-prone. The hard

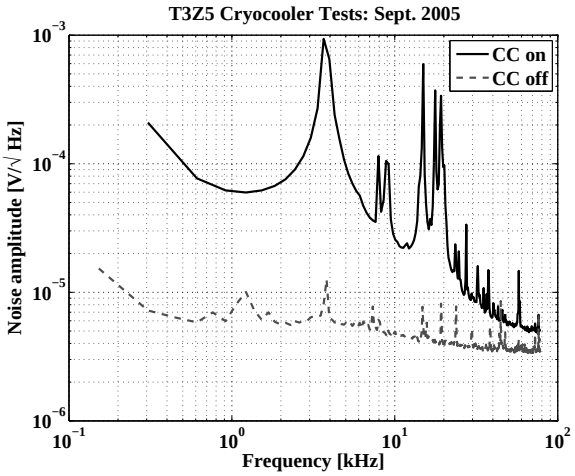


Figure 5.6: High-resolution Qinner noise spectra from T3Z5 from September 7, 2005. The two traces represent two noise runs: one with the cryocooler on (top), one with it off (bottom).

couplings between the cryocooler drive head and the E-stem were replaced with annealed copper braids, as shown in Figure 5.7. Based upon accelerometer and conductivity tests at Berkeley, these new couplings were expected to provide adequate thermal conductivity while greatly reducing the transmission of vibration.

5.2.4 Second icebox run

With the new cryocooler coupling installed, a second icebox run took place in February, 2006. Charge microphonics were much improved, as expected, but the TES's remained unresponsive. A wide variety of theories were put forward to explain the normal TESs, including:

- Heating due to cryocooler vibrations or electrical noise
- Improper heat-sinking of the stripline bundle within the E-stem
- Heat conduction through superfluid helium films
- Infrared leaks from warmer temperatures
- Thermal contact between two icebox cans
- Evil spirits

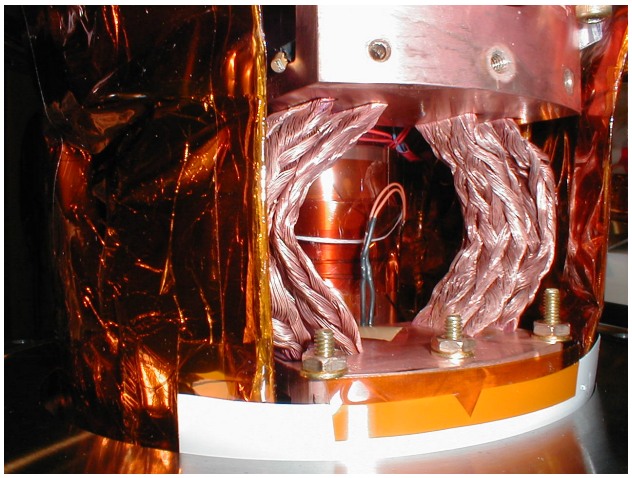
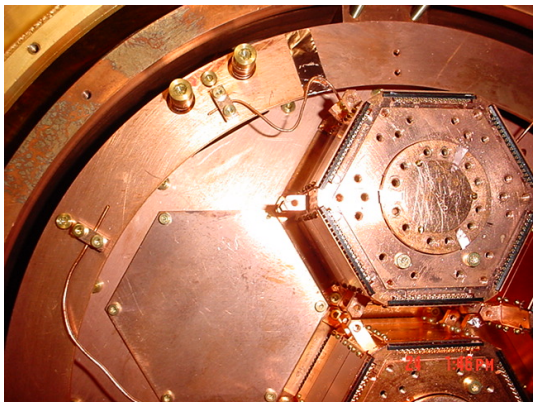
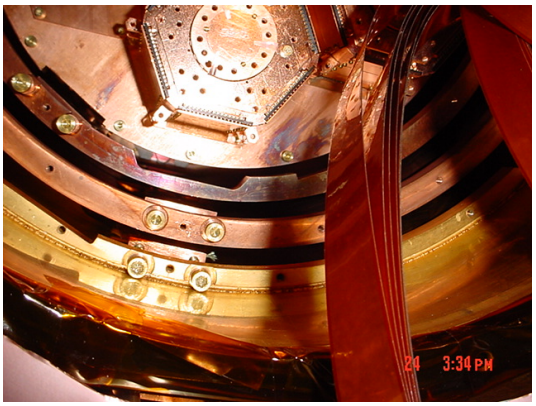


Figure 5.7: Side view of the coupling of the cryocooler to the E-stem prior to the installation of the vacuum jacket. The flexible copper braids form a flexible thermal connection to the 77K stage. The central vertical cylinder is the cryocooler's 4K drive head, which is connected to the E-stem by a similar braid (not visible).

After much discussion and calculation, two developments drew attention toward the performance of thermal connections within the icebox. At Stanford, Blas Cabrera and Walter Ogburn updated and improved upon a thermal model of the Soudan cryostat originally developed by Ron Ross. Their simulations suggested that a combination of poor conductivity through copper-copper contacts and higher-than-expected conductivity through the Tower graphite tubes could replicate the observed profile. Simultaneously, a program of electrical resistance measurements by Kyle Sundqvist at Berkeley showed that a joint between tarnished copper surfaces has an electrical impedance (and hence, by the Wiedemann-Franz relationship [197], a thermal impedance) more than an order of magnitude greater than that of a joint between pristine copper surfaces. The experiment was warmed to room temperature yet again to examine these connections and install additional thermometry. Upon opening the icebox, many of the thermal joints within the icebox were indeed found to have substantial tarnish and corrosion. Figure 5.2.4 shows example photographs of the 50 mK and 600 mK can rims; note the dark tarnish visible in numerous locations.



(a) Top view showing heat sink wires running to the cold plate arc. Tarnish is visible on the surrounding still flange.



(b) Top view after removing the cold plate arc, showing extensive tarnish beneath it.

Figure 5.8: Tarnish visible on the Soudan icebox can flanges in April 2006.

5.2.5 Final configuration and cooldown

With a consistent thermal model now in hand, the collaboration began a new round of changes to improve the thermal performance of the icebox:

1. All copper-copper connections to the Towers and icebox lids were polished with Scotch-Brite to expose fresh, untarnished copper.

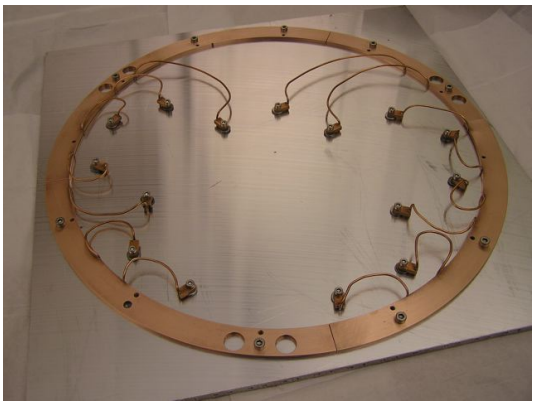
2. The heat conduction wire arrangement that links the Towers to the icebox can lids was completely revamped. The new lids and wires are shown in Figure 5.9. Rather than bolting the heat wires to the lids and connection lugs, the new wires are laser-welded to improve thermal connections. The wires are also annealed for improved conductivity.

3. A segmented heat shield was added to the 600mK layer, blocking infrared leakage from the 4K layer to the base temperature can. This shield is shown *in situ* in Figure 5.10(a).

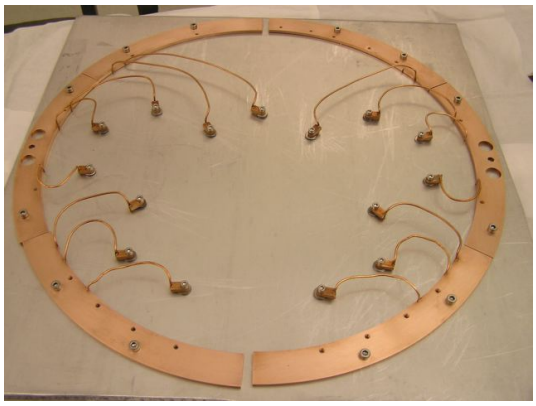
After a series of small on-site modifications to make things fit, Figure 5.10 illustrates the new icebox lids and heat wires in place.

With these changes made, the cryostat reached base temperature on June 26, 2006. The first phonon pulses were seen later the same day, marking the successful beginning of

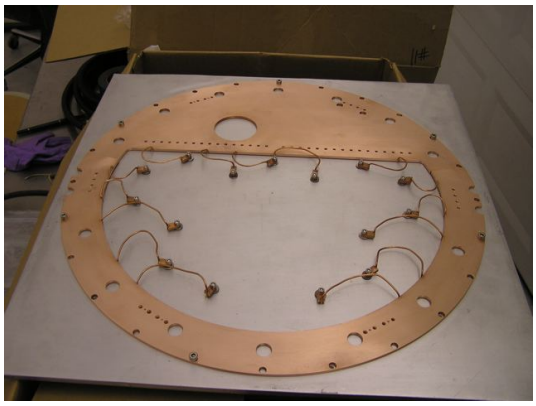
Run 123.



(a) Cold plate lid segments before annealing.



(b) Still lid segments before annealing.



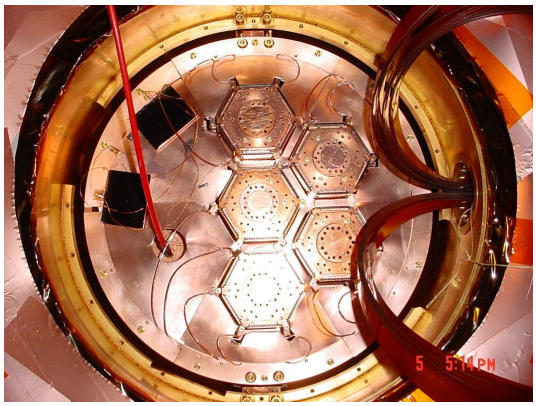
(c) 4 K lid before annealing.

Figure 5.9: New can lids installed in June 2006, including laser-welded heat conduction wiring.

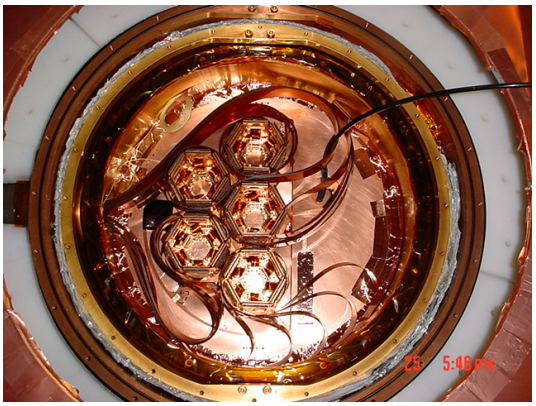
5.3 Detector tuning

Once the detectors reached based temperature, we began the lengthy process of choosing bias points for all 120 SQUIDs and TESs. We made a substantial effort to systematize the tuning process in Run 123, due to the enormous number of channels involved. After a bit of discussion and iteration, we implemented the standard set of procedures described below.

The tuning process lasted from June through September, in parallel with various electronics upgrades and noise troubleshooting. Run 123 data acquisition began in earnest



(a) Completed 600 mK layer, showing new heat wires, IR shield, and IR absorbers.



(b) Completed 4K lid, showing heat wires and IR shield.

Figure 5.10: New thermalization hardware installed in the Soudan icebox in June 2006, shown just before the Run 123 cooldown.

in October 2006, and the last change in sensor tuning (T5Z3 sensor A) was made on October 31.

5.3.1 SQUID tuning

The lock point along each SQUID's modulation curve was chosen according to two main conditions:

1. Each SQUID must be locked at a point away from any resonances, visible as kinks on the modulation curve.

2. The SQUID must have a bandwidth of $\sim$ 350 kHz when the TESs are normal. This value was chosen to lie comfortably above the TES bandwidth but low enough to avoid stability problems.

The SQUID tuning process was conducted using LabView software developed by Dennis Seitz.

5.3.2 TES tuning

The bias current of each TES array was chosen to reconcile two competing constraints. A low sensor resistance R_{TES} gives the TES array a relatively high sensitivity ($\alpha = \frac{T}{R_{TES}} \frac{dR_{TES}}{dT}$ but a low bandwidth ($BW \approx \frac{R_{TES}}{2\pi L_i}$). Raising R_{TES} increases sensor noise

($\sigma \propto \sqrt{BW}$), however. In order to achieve the best signal-to-noise ratio while maintaining risetime discrimination, we bias our TES channels to have noise bandwidths just high enough to encompass the observed risetime information: 80 kHz for Ge and 150 kHz for Si.

Once each SQUID channel was tuned, the bias of the corresponding TES array was set following the above recipe. The In some cases, the four TES channels on a given detector were adjusted slightly for uniform response across the detector. If needed, the TES biases were adjusted to align the histograms of phonon delay (the time delay between phonon and charge pulses).

Chapter 6

WIMP Search Analysis

The first 5-Tower WIMP-search data runs of CDMS II were conducted from Fall 2006 through Spring 2007. Run 123 acquired data from October 21, 2006 through March 20, 2007; the shorter Run 124 commenced after a brief period of cryogenic maintenance, then ran from April 20 through July 16, 2007.

After applying the event reconstruction and calibration algorithms described in Chapter 3, the bulk of the data analysis pipeline consists of defining a series of "cuts" upon these calibrated data. Each cut is a selection criterion that defines one aspect of the region of parameter space in which we look for WIMP-candidate events. The analysis pipeline and relevant cut definitions for the Run 123/124 WIMP search analysis are summarized below, followed by estimates of the exposure and expected signal acceptance of this analysis. Note that the pipeline described below is a team effort, drawing upon the contributions of numerous past and present members of the CDMS collaboration.

6.1 Data reduction pipeline

The primary data acquisition and reduction system for CDMS II consists of three major software packages:

1. The **data acquisition (DAQ) package**, discussed in Section 4.8, controls the acquisition and storage of particle events from the detectors, as well as providing a convenient user interface for local and remote control of the acquisition hardware.

2. **DarkPipe** is the primary data reduction package. This is a software package written

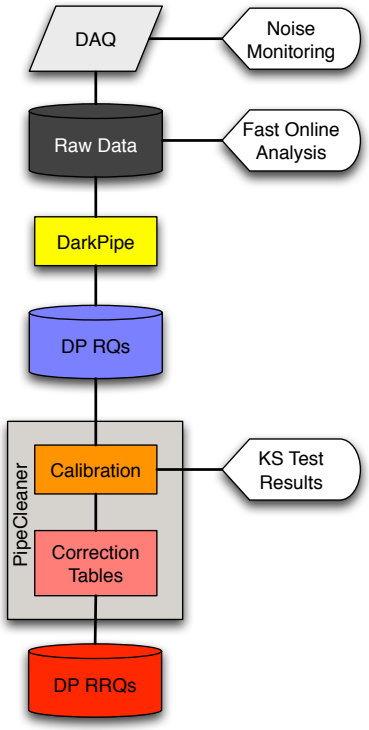


Figure 6.1: Schematic of the CDMS analysis pipeline, indicating major processing modules and real-time monitoring systems.

in the Matlab programming language which converts the raw data stream produced by the DAQ into arrays of Reduced Quantities (RQs) characterizing each event. DarkPipe implements the optimal filters, the pulse-walking algorithm, and a variety of triggering and veto calculations.

3. **PipeCleaner**, also a Matlab package, performs data calibration tasks. PipeCleaner acts upon the RQs generated by DarkPipe, augmenting them with an array of "RRQs" implementing detector corrections and calibrated to physical units (*e.g.* keV rather than voltage at the digitizer).

DarkPipe and PipeCleaner are run consecutively in form of compiled Matlab executables, as described in [198]. Each data set is automatically processed on the Soudan Analysis Cluster (SAC) within hours of its acquisition to produced preliminary RQs and RRQs. These are used for several manual and automated data quality checks, including an array of diagnostic plots and Kolmogorov-Smirnov tests. These are monitored by personnel on shift on-site and off-site, who flag data quality and processing failures soon after they occur.

Later reprocessings of these data took advantage of the greater computing power of the Fermilab Lattice QCD computing cluster. The entire Run 123 data set was reprocessed through DarkPipe and PipeCleaner during the spring of 2007, after the generation of new optimal filter templates and preliminary energy calibrations. PipeCleaner was re-run in June after the phonon correction table was finalized, producing the final set of analysis quantities. Run 124 data were reprocessed similarly during the summer of 2007.

6.2 Blindness mask

This data analysis was performed "blindly," in that the definitions of the various analysis criteria were set without knowledge about the detailed characteristics of any candidate or near-candidate events which might be present in the data. This helps to keep the WIMP-search analysis unbiased by its eventual result, which greatly simplifies the statistical issues associated with reporting that result. Blindness is maintained in the CDMS experiment through the use of a blinding mask. The blinding mask is a cut that hides the signal region of the data stream, removing it from study until after the WIMP signal criteria have been defined. All cuts are set using data from source calibrations and WIMP-search data outside of the masked region, so that there can be no fine-tuning of cut placement to

achieve a desired result.

The blinding mask defines a simplified and conservative version of the WIMP-search analysis (excepting the timing cut), selecting events satisfying all of the following criteria:

1. Veto-anticoincident: no signal in the veto counter in the 50μs preceding the global trigger.

2. Inner charge electrode: amplitude of the outer electrode energy is small ($|\texttt{qo}| < 5$ keV)

3. Energy range: recoil energy less than 130 keV and greater than some minimum threshold (generally 2 keV, sometimes higher for noisier detectors)

4. Ionization yield: within an approximate 3σ nuclear recoil band

5. Multiplicity: Phonon signal within the 6σ noise for all but one ZIP.

These cuts are defined before the application of the correction table, using hard-coded calibration factors copied from PipeCleaner. Due to the wider cut definitions and the poorer resolution of uncorrected phonon variables, these cuts are designed to substantially over-cover the true WIMP signal region. The final version of this cut which was designated `cBlind070123_123`.

The blinding cut was applied by the PipeCleaner package during the final stage of data reduction. The events tagged by this cut were excluded from the RQ data files used in later analysis. Full versions of the RQ files including all tagged events were stored separately on disk and only accessed for final leakage computations after all analysis cuts were fixed.

During the analysis one minor failure of this blinding cut was uncovered. This error was traced back to the effect of variations in noise amplitude during the run on the single-scatter cut, as described in Section 6.5.2. Particularly noisy periods on one or more detectors led many events to be falsely rejected by the single scatter cut, allowing some true single-scatter events to fail the blinding cut and remain unmasked. This failure was discovered very early on in the study of the WIMP search data, and no timing quantities of these events were studied. For security I created a "backup" version of the blinding cut including the final version of the singles cut. This backup cut was applied in addition to the

standard blinding mask for the remainder of the WIMP search analysis. We do not believe that the blinding of the analysis was compromised by this issue.

6.3 Data quality cuts

The combined Run 123 and 124 data sets encompass more than 4 million triggered events. We developed a series of software criteria ("cuts") to select WIMP candidates from among this large background population. Consistent with our blinding philosophy, each of these cuts was set based on calibration data and on WIMP search data outside the blinding mask. The various cuts are divided into three rough categories:

1. **Data quality cuts**, which identify periods of poor detector performance and faulty event acquisitions.

2. **Reconstruction quality cuts**, which identify events that are insufficiently well-reconstructed to be considered as WIMPs.

3. **Physics cuts**, which define the ultra-low-background region of parameter space in which we will look for WIMP candidate events.

Once the various event parameters are reconstructed, several data quality cuts are applied to exclude various categories of "bad" events from further consideration. These cuts take two major forms: those that identify periods of poor detector performance, and those that identify individual events which may be poorly reconstructed.

6.3.1 Trigger criteria

All WIMP search candidate events should be acquired in the standard operating mode of the data acquisition system: a global trigger induced by a detector Plo trigger. We enforce this through the use of several cuts on the trigger information of a given event. These cuts exclude events which have an error in their trigger information, were acquired as "random" triggers by the DAQ (acquired to characterize event noise), or were triggered by simultaneous hits in multiple scintillator paddles (acquired to monitor veto counter performance).

6.3.2 Detector selection

As described previously, not all detectors performed equally well during the Run 123 and Run 124 analyses. Only a subset of detectors were used for WIMP search during each run, with the rest used only for the multiplicity cut. The good detector list for each run was enforced by the cut `cGoodDet`.

Seven detectors were excluded from both data runs for a variety of reasons related to cold hardware and detector failures:

- T1Z1: disconnected phonon channels due to damaged tower wiring
- T1Z3: disconnected Qouter amplifier
- T1Z6, T2Z6: poor neutralization due to broken LED wiring
- T3Z1: bistable performance of phonon sensor C. T3Z1C frequently switched between a "good" state and a "lazy" state, in which its phonon pulses were much smaller and slower. The switch seemed sensitive to event rate, so almost all calibration data was acquired in the lazy state while the WIMP search data was more mixed. Without a proper calibration sample, detailed analysis was impossible.
- T5Z2: disconnected Qinner amplifier
- T5Z6: very poor signal-to-noise for the phonon sensors

Part-way through the Run 123 analysis the decision was made to incorporate as much of the briefer Run 124 as possible into that analysis. This raised concerns about data uniformity, since some detectors had changed bias and neutralization states subtly (or unsubtly) during the intervening warm-up of the cryostat. Strict criteria based on KS tests and other data comparisons between the runs were set for the inclusion of Run 124 detectors into the Run 123 analysis. The following list of detectors were eventually excluded for Run 124, in addition to the seven listed above: T1Z4, T2Z1, T2Z2, T2Z3, T2Z5, T3Z2-6, T4Z2, T5Z3, T5Z4. The criteria which led to this good data set were extremely conservative, so it is very likely that a later analysis could make use of most of the excluded detectors.

6.3.3 Stable TES tuning

The QET biases of each of the 120 phonon channels of CDMS II must be tuned to attain the desired bandwidth and maintain balanced phonon timing response across each

	T1	T2	T3	T4	T5
Z1	12 Sep 09:47	02 Oct 16:51	27 Sep 11:46	09 Oct 10:49	01 Aug 12:08
Z2	12 Oct 01:44	02 Oct 16:51	29 Sep 13:21	02 Oct 16:51	29 Sep 14:11
Z3	12 Aug 14:25	22 Sep 09:58	29 Sep 13:21	15 Aug 20:37	**31 Oct 12:51**
Z4	29 Sep 17:47	29 Sep 14:11	**26 Oct 15:34**	27 Sep 21:33	01 Aug 16:37
Z5	02 Oct 16:51	21 Sep 10:59	27 Sep 21:33	29 Sep 13:21	12 Aug 13:10
Z6	29 Sep 14:11	24 Sep 12:47	12 Oct 01:44	**27 Oct 17:23**	27 Sep 21:33

Table 6.1: Starting times for the first data sets taken after the final QET tuning of each detector for Run 123. WIMP candidates are not accepted from each detector from earlier data runs. In practice this cut only affects the three **boldface** detectors.

detector. The tuning process was largely complete before the official commencement of the Run 123 WIMP search (October 21, 2006), but a few channels were re-tuned as late as October 31. The cut `cStabTuning_123` enforces that no WIMP candidate is accepted from a detector before its date of last QET tuning, given in Table 6.1.

6.3.4 Trigger stability

Some detectors show elevated trigger rates for brief periods of time, which may indicate electronics problems or increases in phonon noise. The `cTrigBurst_123` cut excludes two classes of such periods:

1. Periods of WIMP search data with an overall trigger rate > 0.7 Hz over at least 100 consecutive non-random events. Excluding these events reduces the overall R123 (R124) live time by only 0.01% (0.006%) while removing more than 5% of events.

2. Periods of elevated trigger rates on T1Z4 and T2Z2 specifically in Run 123, identified by an unusually high rate of events with negligible phonon energy (< 1 keV) in a 99-event window. These cuts reduce the livetime on T1Z4 (T2Z2) by 0.06% (0.9%), but have no effect on the Ge analysis.

6.3.5 Anomalous noise spectra

The cuts `cHighQNoise_123` and `cHighQNoise_124` remove occasional time spans of elevated charge noise within a data set. These periods are identified using a running average of charge energy (`qsum`) for events with negligible phonon energy, assumed to be noise events. These cuts remove periods during which the running charge average exceeds

the mean value by 1.5σ, including one event on either side of the deviating window and disregarding extreme outlier events. The livetime cost of these cuts is 1.2% in Run 123 and 0.4% in Run 124.

The various phonon noise spectra were generally fairly consistent throughout the WIMP search and calibration data. Most variations in resolution or bandwidth were small enough to have negligible effect on background discrimination, but some detectors showed variations in noise amplitude or bandwidth large enough to be of concern. `cBadResTight_123` identifies series with particularly poor sensor resolution which might degrade yield and timing discrimination, most prominently extended periods of elevated low frequency noise on T1Z4A/C/D and T2Z4C. The resolution shifts on T1Z4 are also visible in Figure 6.12. `cBadRolloff_123` removes series with phonon noise bandwidths significantly different from normal, notably significant periods of low noise bandwidth on T2Z4 and T3Z3 which may alter the distribution of timing parameteres.

6.3.6 Detector neutralization

If a millikelvin semiconductor detector is left under voltage bias for an extended period of time, accumulated ionized trapping centers eventually compromise charge collection. Our understanding of this process of "neutralization loss" is not yet fully quantitative, but the rate of degradation is expected to be proportional to the total deposited charge energy. These trapping centers may be cleared using periodic irradiation with LEDs on the detector housings.

The cuts `cBadNeut_123` and `cBadNeut_124` remove periods in which the detector neutralization is suspect. These cuts combine information from three studies of neutralization loss during the two runs:

1. A quantitative characterization of the yield distributions of the various Ba data sets, based on the fraction of events at low ionization yield ($0.1 <$ yield < 0.8, yield below the 2σ electron recoil band). For each Ba data set and for each segment of 10^5 events within a Ba set, I computed the fraction of events passing basic quality cuts which also lay at low ionization yield. Any data set more than 2σ above the overall mean (marked in red in the left portion of Figure 6.2) was marked for manual inspection, as was any series in which one or more segments were 2σ above the mean (a less sensitive criterion, given the larger error bars from the lower statistics of a segment).

The former were generally excluded from use in surface event studies, as were the latter portions of some data sets from either category in which a clear upward trend in low yield fraction was visible (right portion of Giure 6.2). In a few cases where the low yield fraction remained high for several consecutive Ba sets, the intervening WIMP search sets were also excluded.

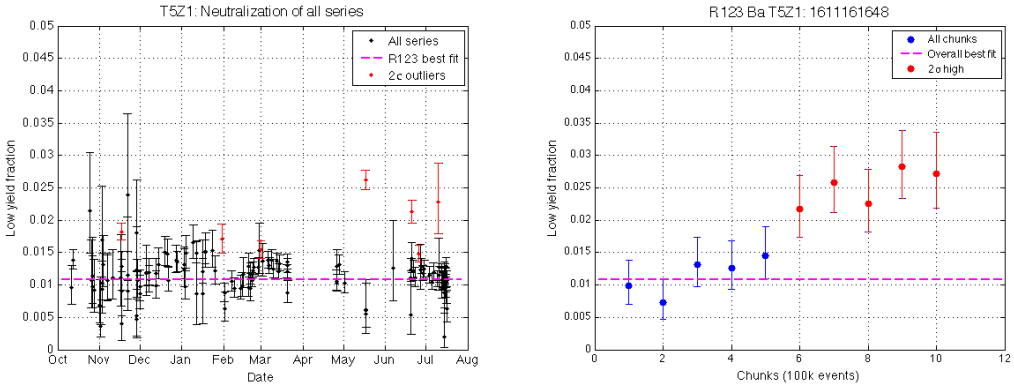


Figure 6.2: *Left:* Low-yield fraction computed for each data series of Runs 123 and 124. Red symbols indicate series exceeding the mean fraction by more than 2σ. *Right:* Example of apparent neutralization loss in T5Z1, series 161116_1648. Error bars indicate low yield fractions in 100,000 event bins, illustrating a clear rise late in the run. Red symbols indicate periods identified as 2σ above the mean.

2. A characterization of the WIMP-search data sets by Cathy Bailey using similar methods. Due to much lower statistics, this study could not subdivide series into smaller segments and was generally less sensitive to variations in yield distribution. This study nonetheless provided the most direct indication of low-yield fraction for WIMP-search data runs, and led to the exclusion of several data series. We also excluded any WIMP search series following a Ba series excluded by the first study without an intervening LED flash to re-neutralize the crystals.

3. An accounting for each event of the length of time the detectors spent under bias since the previous LED flash, performed by Wolfgang Rau using code developed by Joel Sander for Run 119. A few Ba data sets were not followed by an LED flash due to bugs in the DAQ's automated flashing code. If a Ba data set shows evidence of neutralization loss on a detector, all data from that detector before the next LED flash are excluded from the WIMP-search analysis.

All told, these cuts exclude approximately 1-7% of the experiment's raw live time on most detectors. Longer time periods are excluded for three detectors (T1Z4, T2Z2, and T3Z6) due to extended periods of poor neutralization. These cuts should be relatively conservative, but it is not impossible for a series with degraded neutralization to slip past, particularly if the degradation is localized to a region within a detector. Work on quantitative modeling of neutralization loss mechanisms is in progress, and may inform future analyses.

6.3.7 NuMI neutrino beam

As part of the MINOS neutrino oscillation experiment, an intense beam of muon neutrinos at GeV energies is currently projected from the Main Injector facility at Fermilab to the MINOS far detector at Soudan. Neutrinos from this beam may interact in the shield or cavern rock to generate muons and associated neutrons, which may be detectable in CDMS II. These are not expected to constitute a significant source of unvetoed background to the experiment, but we nonetheless attempt to exclude events in coincidence with the beam. The MINOS beam is active for only 10 μs every 2.5 seconds, so the cost of such a cut in live time is negligible.

The cut `cNuMI_R123` selects events in coincidence with the MINOS beam based upon their GPS time signatures. The GPS time stamp of each event is recorded to the nearest 100 ns using a fiber optic link to the MINOS GPS receiver. After correcting for the measured fiber optic delay time (1.45 μs), these time stamps are compared against the time record of each NuMI beam dump. The cut excludes detector-triggered events within 60 μs of a beam spill and shield-triggered events within 10 μs of a spill. The cut is only functional after series 161214_1221, when a failure of the fiber optic connection during the fall of 2006 was repaired. A total of 20 events are excluded from this analysis based on these criteria, none of which would have been WIMP candidates in any case.

6.4 Reconstruction quality

6.4.1 Poor detector regions

Two Ge detectors showed very anomalous behavior localized to some portion of their layout. In each case these regions produced an excess of events at low ionization yield

and slow phonon timing, and thus a large population of timing outliers.

1. **T3Z2 C/D:** As shown in the left plot of Figure 6.3, quadrants C and D of T3Z2 show a huge population ($\sim$ 1000) of low-yield events, the vast majority of which are also outliers in phonon timing. The right plot shows that this population lies along an apparent gap in Qouter, an arc in which events under the outer ionization electrode deposit negligible charge in that electrode. This is likely the result of a region of disconnected TESs, a patch of floating metal which distorts the local electric field and prevents Qouter charge collection. To avoid subtle effects of this field distortion on the phonon correction table, all events for which channels C or D are the primary sensors are excluded from WIMP-search analysis.

2. **T5Z1 D:** A "curl" distortion is visible in quadrant D of T5Z1's delay and box plots. This is probably a symptom of a malformed region of the TES pattern. This region exhibits phonon manifold degeneracy and an excess of timing outliers, so quadrant D is excluded from WIMP-search analysis.

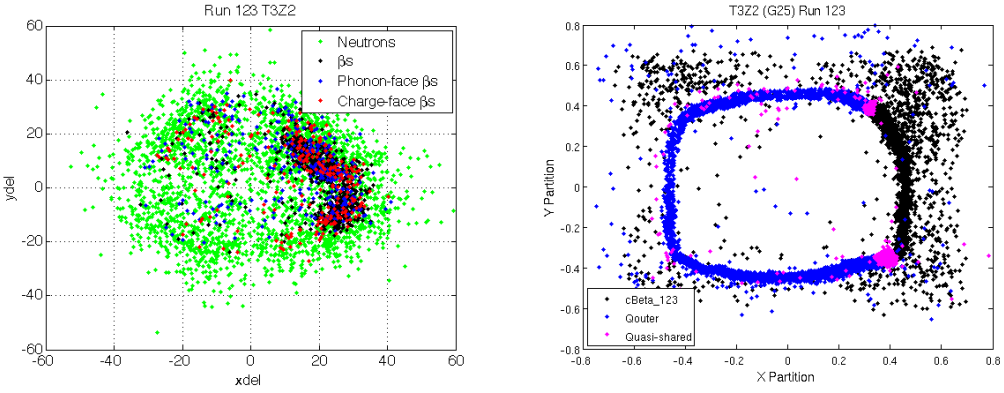


Figure 6.3: *Left:* Delay plot for T3Z2, highlighting neutrons (green) and low-yield events (black, with blue and red indicating phonon- and charge-face selections respectively). Note the excess of low-yield events in the right portion of the detector. *Right:* Phonon partition ("box") plot for T3Z2, illustrating a bad detector region. Black indicates Qinner events (*i.e.* negligible Qouter), blue indicates Qouter events, and magenta indicates shared events. The right half of this detector is excluded from analysis.

The cut `cBadDetRegions_123` selects these three detector quadrants based on their values of `xdel` and `ydel`. This cut removes 50% of T3Z2's exposure and 25% of T5Z1's, with no effect on the other detectors.

6.4.2 Ionization goodness-of-fit

If two particle events occur sufficiently closely together in time (*i.e.* separated by less than the post-trigger digitization window, $\sim 1200\mu$s) in a single detector, the pulses from the two events can overlap in a digitizer trace. Such events are very rare at the low trigger rates of WIMP search data, but are very common in calibration sets. The distorted pulses from such "pileup" events can be reconstructed incorrectly, and so should be removed from consideration as WIMP candidates or for detector characterization.

Pileup events in the ionization channels are poorly fit by the usual optimal filter templates, and so possess high values of optimal filter χ^2 (`QSOFchisq`). `cChiSq_123` enforces an upper limit on `QSOFchisq`, consistent with that of an isolated event free of pileup. This cut takes the form $a + b \cdot Q^2$, with a generally near 5000 (the expected χ^2 is 4800) and b set empirically to match the upward flaring of the χ^2 distribution for large pulses, shown in Figure 6.4.

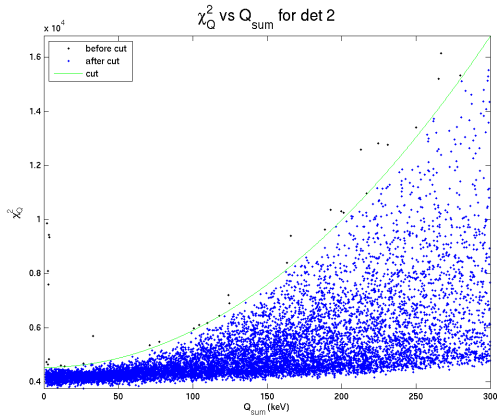


Figure 6.4: Charge optimal filter χ^2 (`QSOFchisq`) versus summed charge amplitude (`qsum`) for T1Z2 WIMP search data. The green line indicates the position of the cut. Black dots indicate events failing the cut. Figure by D. Moore.

The curvature and broadening of the χ^2 band at high ionization amplitudes visible in Figure 6.4 is due to two major factors. The templates retain some noise since they are produced by averaging a finite number of real pulses. This noise enhances χ^2 slightly at high pulse amplitudes, leading to the slight curvature of the distribution's lower edge. Furthermore, DarkPipe only determines the start time of the ionization pulses to the nearest digitizer bin (0.8μs). This inaccuracy in the start time determination broadens the difference

in χ^2 between events with slightly different true start times. This broadening reduces the effectiveness of `cChiSq_123` slightly, though pileup rejection remains very powerful. This will be addressed in future analyses through sub-bin interpolation of the ionization start time and χ^2 (Appendix A.4).

6.4.3 Pre-pulse phonon noise

The ionization χ^2 cut is extremely effective at identifying pileup events, but is not always sufficient to exclude all such events. The cut is less powerful at higher energies due to the broadening described above, and also has limited power for the tiny ionization pulses of nuclear recoils near threshold. Phonon pulses also have substantially longer falltimes than ionization pulses, so a trace can contain a phonon pulse tail from a previous trace but not an ionization pulse tail. Ideally we would impose a similar cut on the χ^2 of the much larger phonon pulses, but this is ineffective due to the large variations in phonon pulse shape with amplitude and position.

`cPstd_123` rejects events for which the pre-pulse baselines (the first 400 digitizer samples) of any phonon trace has an unusually high standard deviation (RQs `P*std`). Since these samples are well before the phonon trigger, they should just sample the overall noise of the phonon channels and should have a gaussian distribution that is essentially constant over time. The cut level on each channel is set at the 5σ level based upon a gaussian fit to the distributions of `P*std` for a large sample of data. This cut serves as an additional protection against pileup, as well as rejecting events with unusual phonon noise in general.

6.4.4 Pre-pulse charge noise

In analogy with `cPstd_123`, this analysis imposes a new cut `cQstd_123` on the pre-pulse standard deviations of the ionization channels. This cut does have rejection power for pileup events, but in that context is essentially redundant to `cChiSq_123` and `cPstd_123`. Its primary purpose is to reject events taken during periods of unusual ionization noise, specifically those during the more microphonic periods of the cryocooler duty cycle.

As described in Section 5.2.3, the E-stem cryocooler transmits substantial vibrations to the icebox during portions of its 1.2 second refrigeration cycle. The phonon channels and most ionization channels are essentially immune to these vibrations, but a few ionization channels show significant cryocooler-induced microphonic pickup. The largest effects

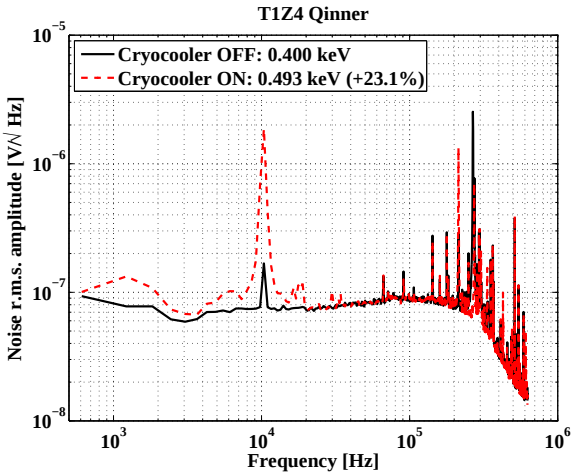


Figure 6.5: Noise power spectrum for T1Z4 Qinner, showing cryocooler-induced microphonic peaks near and below 10 kHz. The solid spectrum shows the situation with the cryocooler temporarily deactivated, while the dashed shows the increased microphonics with the cryocooler in normal operation. The caption indicates the r.m.s. resolution in each configuration.

are seen on the Qinner channels of T1Z4, T3Z2 and T3Z5 and the Qouter channels of T1Z4, T2Z1, T2Z3, T2Z5, T3Z2 and T3Z5. Lesser effects are visible on the Qinner channels of T2Z3 and T2Z5 and the T4Z2 Qouter. Figure 6.5 shows the difference in the average noise spectrum of T1Z4's random triggers with the cryocooler on and off.

The cryocooler vibrations vary greatly in amplitude during the refrigeration cycle, and this variation is clearly visible in real time on the microphonically-sensitive charge channels. The random triggers from the noisiest cryocooler periods are generally excluded from DarkPipe's noise analysis by the automatic pulse-rejection algorithm, and the noise spectrum used in the optimal filter calculation represents an average over the quieter and intermediate portions of the cryocooler cycle. The microphonic peaks in Figure 6.5 are thus larger for some fraction of events and smaller for others. The signal-to-noise performances of these charge channels thus vary at different portions of the cryocooler cycle, suggesting that variations in the analysis with cryocooler phase may also be beneficial.

Since the timing of the cryocooler cycle is mechanical and derived from the 60 Hz power line frequency, it is relatively unstable over long periods of time longer than a few seconds. As shown in Figure 6.6, the standard deviation of the pre-pulse baseline (`QIstd` and `QOstd`) on microphonically-sensitive charge channels is a useful proxy for the phase

of the cryocooler cycle. It is convenient to measure in normalized units, based upon the approximate mean and standard deviation of each parameter distribution. To limit the effects of the long tails of the `Q*std` distributions, the mean and standard deviation are determined through an iterative gaussian fit: the fit is performed several times, throwing away outlying events at each iteration until stability is achieved. As shown in Figure 6.7, cryocooler-sensitive detectors have a significant population of events at high baseline noise. These events have degraded ionization standard deviation, as illustrated in Figure 6.8. Neither of these effects is visible in an insensitive detector.

The cut `cQstd_123` demands that `QIstd` and `QOstd` each lie below $+4\sigma$ from the mean in these normalized units. This cutoff is indicated by the blue vertical lines in Figures 6.7 and 6.8; the approximate efficiencies of these cuts are shown beside the lines in the former. The cut is set at somewhat lower values on T2Z3Qi, T3Z2Qo, T3Z5Qi, T3Z5Qo, and T4Z5Qi to avoid significant changes in measured charge resolution within the accepted region. Cut efficiencies are essentially 100% in most detectors, dropping as low as 69% in the most microphonic detector, T3Z5.

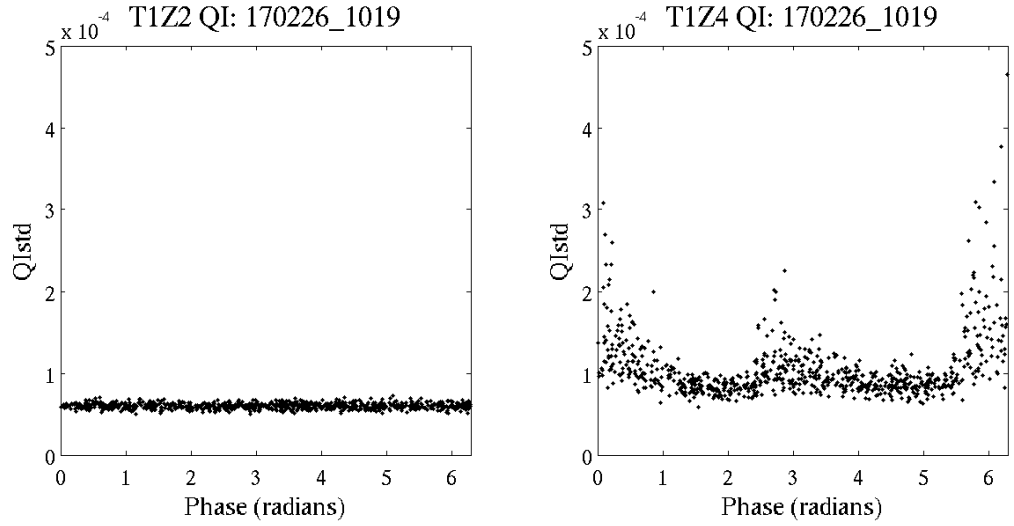


Figure 6.6: Relation between approximate cryocooler phase and pre-pulse ionization standard deviation `QIstd`. The left plot shows T1Z2, a detector relatively insensitive to cryocooler microphonics, while the right plot shows the much more susceptible T1Z4. The plotted cryocooler phase is an approximation based upon the measured oscillation period over a short span of data.

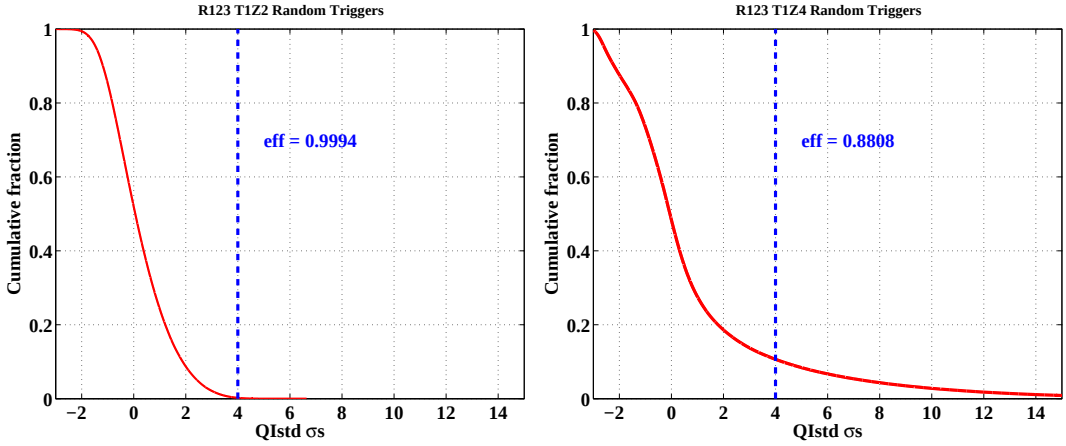


Figure 6.7: Distribution of `QIstd` for a two detectors, one microphonically insensitive (T1Z2, *left*) and one sensitive (T1Z4, *right*). The x-axes are measured in normalized units of standard deviations from the mean, based upon gaussian fits to the observed distributions. The curves indicate the fraction of random trigger events with `QIstd` greater than the given value. Note the larger tails to high numbers of σ in the sensitive detector.

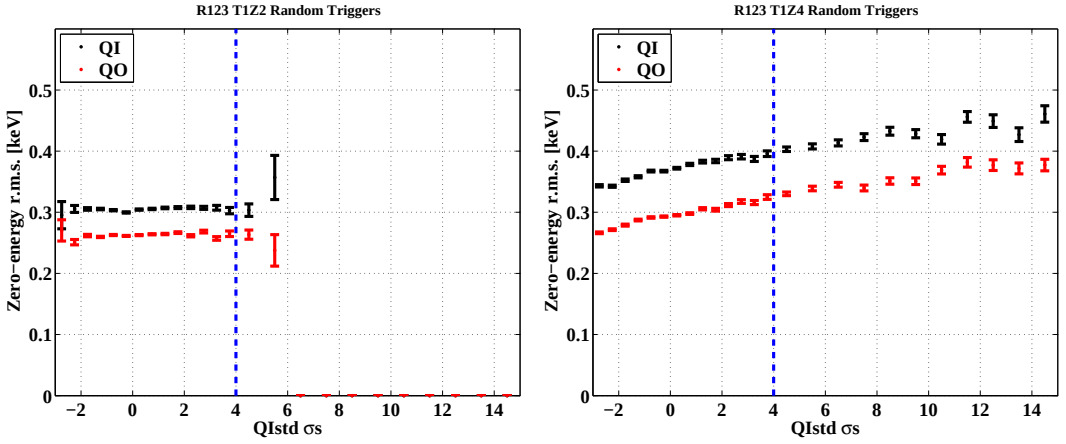


Figure 6.8: Ionization resolution (standard deviation) as a function of `QIstd`, measured in normalized units. The left plot shows a microphonically insensitive detector (T1Z2), the right a microphonically sensitive one (T1Z4).

6.4.5 Phonon start time

In addition to the pileup events described above, a multi-detector configuration can suffer from an additional reconstruction pathology called "cross-detector pileup". In this class of events a particle event triggers the data acquisition system and a second, unrelated event occurs on another detector within the digitization time window. There is no overlap within a single detector so each pulse is properly formed, but the second pulse occurs abnormally late within its digitization window. An example from Run 119 is shown in Figure 6.9. This is very rare at the low acquisition rates of WIMP-search data but relatively common during high-rate calibration.

Cross-detector pileup is problematic because the pulses of the second particle event may occur outside of DarkPipe's optimal filter search window. DarkPipe only searches for the start time of a Ge ionization (phonon) pulse within a window of [-100,+10] ([-50,+200]) μs around the trigger time, which occurs 512 bins ($\sim$ 400 μs) into the digitized trace. Due to the faster phonon response of Si detectors, the ionization and phonon windows for Si events are modified to [-50,+10] and [-25,+200] μs, respectively. If a pulse lies outside this window the optimal filter will select an incorrect start time within the window, resulting in anomalously low amplitudes. Widening this window would slow processing time significantly, and pulses well outside the window may be truncated and thus not well-reconstructed in any case.

Such cross-detector pileup events have poor ionization fits and so are almost always rejected by `cChiSq_123`. In rare cases, however, a low-energy gamma can sneak past this cut and be mis-reconstructed to low ionization yield, mimicing a nuclear recoil. Since the phonon timing quantities are determined by a walking algorithm independently of the optimal filter, the primary phonon risetime is generally correct (*e.g.* gamma-like) for such events while the primary phonon delay is anomalously long (the ionization start time is based upon the optimal filter, and thus incorrect). In previous analyses the charge and phonon timing quantities were all determined using the walking algorithm, so the timing discriminants were essentially correct even for these events. These often passed the timing cut and appeared as WIMP candidates, since they have the slower timing of gammas.

`cGoodPStartTime` rejects cross-detector pileup by enforcing that the walked start time of the primary phonon pulse lie within the optimal filter window for the ionization channels, *i.e.* that `PXr20` $\in [-50, +10]\mu$s for Ge ($[-25, +10]\mu$s for Si). This corresponds to

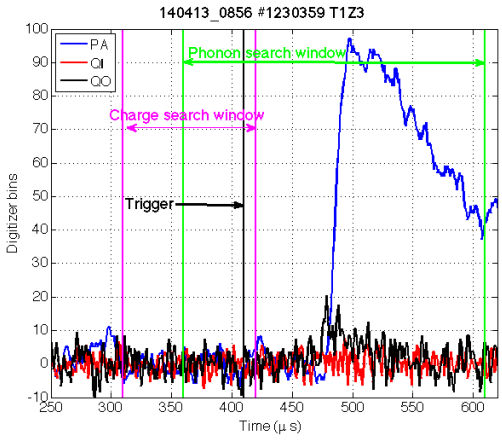


Figure 6.9: Example of a cross-detector pileup event in T1Z3 during Run 119. The initial trigger was on an unrelated event in T1Z1. Raw traces for the ionization and primary phonon sensors are shown, as well as vertical lines indicating the trigger time and optimal filter search windows. Note that the ionization pulse occurs well outside the appropriate search window, leading to an incorrect value for `pdel` (63 μs).

the overlap between the charge and phonon optimal filter search windows shown in Figure 6.9. Signal efficiency is nearly perfect for WIMP-search events passing other quality cuts, but there is a small decrease in efficiency at very low energies. This loss occurs because the trigger threshold is crossed later for a smaller phonon pulse, shifting the entire event earlier in the digitized trace with respect to the fixed window.

6.5 Physics cuts

6.5.1 Scintillator veto

The veto cut `cVTStrict_123` selects detector events coincident with activity in any of the 40 scintillator panels of the surrounding muon veto counter. The events passing this cut are rejected from consideration as WIMP candidates.

The primary obstacle to setting this cut is the high ($\sim$ 600 Hz) trigger rate in the veto counter from ambient gamma rays. The detectors are protected from these events by substantial shielding and powerful discrimination, but these gammas would lead to unacceptable dead time if they were allowed to trigger the veto cut. Events passing the veto cut must thus satisfy either (or both) of the following two conditions:

1. The photoelectron signal lies above a panel-specific threshold anywhere within a veto

window of [-180,+20] μs around the global trigger for any of the veto panels. This threshold is set high enough to exclude essentially all gamma events, generally between 30-50 pC. The cut position for one panel is plotted in Figure 6.10, indicating a threshold near the upper edge of the gamma distribution but well below most muons and showers.

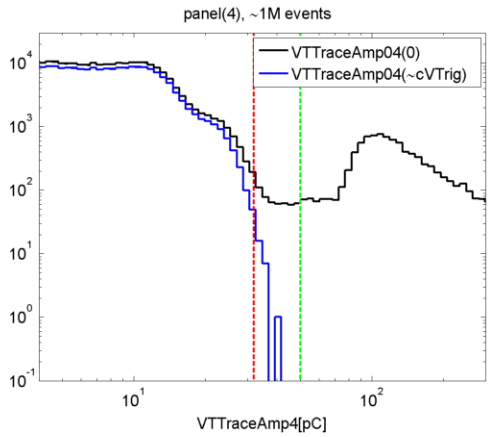


Figure 6.10: Histogram of WIMP search veto amplitudes for veto panel #4. The black line indicates all veto triggers. The blue indicates events not satisfying the veto multiplicity trigger, predominantly gamma events rather than muons or showers. The green vertical line indicates the 50 pC fixed threshold of the Run 119 analysis, while the red vertical line marks the threshold for this panel during the current analysis. Figure by R. Hennings-Yeomans

2. The largest photoelectron signal within the veto window lies within [-50,0] μs relative to the global trigger.

Though these cuts eliminate the vast majority of false gamma events, some dead time is nonetheless introduced into the analysis due to the shield event rate. The WIMP search efficiency is calculated to be 97.6% in Run 123 and 96.6% in Run 124, based upon the fraction of random triggers which pass the veto cut.

6.5.2 Event multiplicity

Due to their extremely small interaction cross sections, WIMPs are not expected to scatter multiple times within any experimental apparatus of practical size. Multiple scatters are quite common for neutrons and electromagnetic backgrounds, however, and therefore we demand that any WIMP candidate event scatter in only a single detector. We classify an

event as a single scatter (and thus a possible WIMP candidate) if it shows significant energy deposition in one and only one detector. More specifically, the cut `cSingle_123` requires that the uncorrected phonon energy `prg` exceed the 6σ upper edge of one detector's phonon noise blob but fall below the 4σ upper edge for all other detectors. The background rejection power of this cut naturally increases as the number of detectors grows, and false rejections of true single scatter events should be very rare at the low (< 1 Hz) trigger rates of WIMP search data. In the idealized case of static gaussian noise, the rate of false positives should be only $(30-1)\frac{1}{2}\mathrm{erfc}\left(\frac{4}{\sqrt{2}}\right) = 0.09\%$ (*i.e.* 99.91% signal efficiency).

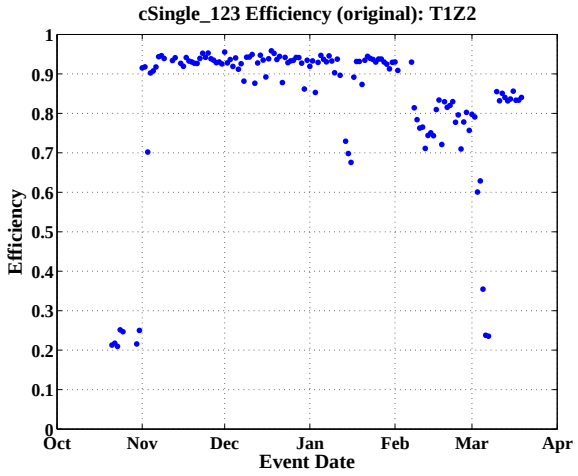


Figure 6.11: Initial signal efficiency of `cSingle_123`, measured as a function of series date.

A naive application of this scheme has substantially worse efficiency than expected, however, and that efficiency fluctuates strongly from run to run as shown in Figure 6.11. The bulk of the observed poor performance is due to the significant variations in noise baseline seen in some detectors over the run, illustrated in Figure 6.12 for Tower 1. These variations do not generally have a dramatic effect on WIMP search performance, and time period of especially bad performance on a given detector are removed from consideration by quality cuts. Changes in resolution can have a dramatic effect on the multplicity cut, however. A detector experiencing poorer resolution will pollute the singles cut with frequent false vetoes, as noise fluctuations drive `prg` intermittently over its 4σ threshold. A detector may issue such false vetoes even while excluded from the WIMP search, drastically reducing the efficiency of the singles cut and thus our effective livetime for true WIMP recoils. Each loss in efficiency in Figure 6.11 can be traced back to degraded performance by some detector,

most notably T5Z3A (for which QET tuning was not completed until the end of October) and T1Z4C (known to have poor noise performance in the spring).

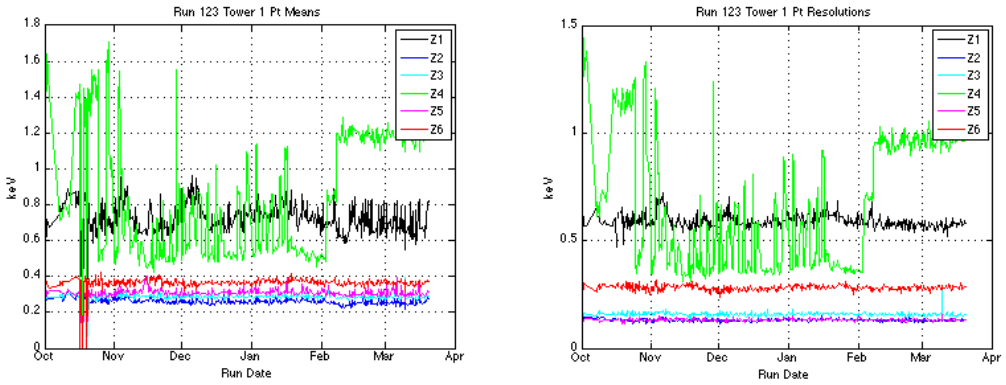


Figure 6.12: Variation in mean *(left)* and standard deviation *(right)* for phonon noise of Tower 1 ZIPs during Run 123, as measured using gaussian fits to `pt` for each run's random triggers. Note the large variations for T1Z4, a particularly prominent source of false vetoes in a naive singles cut.

Excluding all periods of sub-optimal noise performance would significantly reduce the power of the multiplicity cut. The final form of `cSingle_123` instead implements series-dependent values for the various phonon noise means and resolutions, each computed by an iterative gaussian fit to the random triggers of a given data series. `cSingle_123` sets the high and low thresholds on a run-by-run basis using these fitted quantities. No detector is ever fully excluded from vetoing a prospective single, and the threshold is adaptively placed as close to the noise as is practical. In addition to these phonon thresholds, four detectors with relatively poor phonon sensor performance (T1Z1, T3Z1, T5Z5, T5Z6) are also checked for ionization signals above analogous 4σ thresholds.

Efficiency is also reduced in some time spans by intermittent false multiples due to a series of glitches on T5Z2A . An example is shown in Figure 6.13. Such events are not a concern elsewhere in the analysis, since T5Z2 is not a WIMP-search detector. For simplicity T5Z2A is excluded from the singles cut entirely; the phonon noise cut is applied to the three sum of the remaining three channels.

The efficiency of the final version of `cSingle_123` is shown in Figure 6.14 for T1Z2, as measured using random trigger events. The mean efficiency of 99.4% is slightly lower than that expected for static gaussian noise, presumably due to the finite error bars on the run-specific noise parameters.

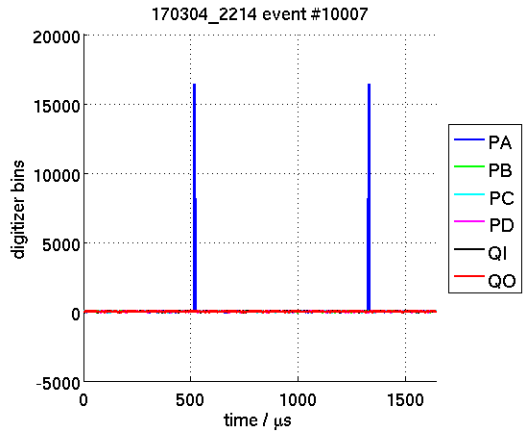


Figure 6.13: Representative phonon glitch event on T5Z2. All channels except Phonon A have low amplitudes, consistent with noise.

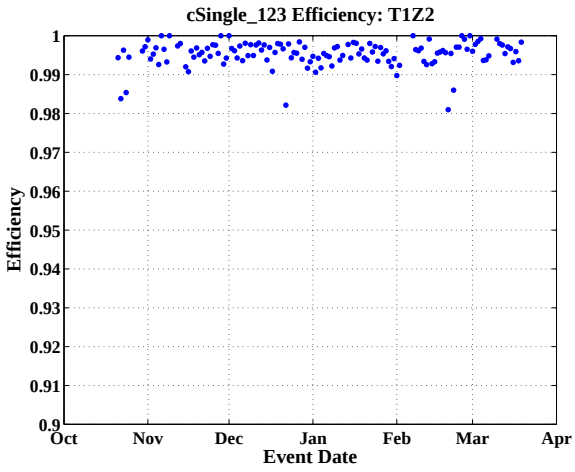


Figure 6.14: Final signal efficiency of `cSingle_123`, measured as a function of series date. Note the shorter y-axis range compared to Figure 6.11.

Due to the very high event rate during calibration with radioactive sources, overlapping events in different detectors ("cross detector pileup", see Section 6.5.2) is relatively common and the singles cut naturally becomes less useful. Such coincidences are essentially nonexistent in WIMP search data runs due to the far lower event rate (well below 1 Hz).

6.5.3 Fiducial volume

The fiducial volume cut rejects events at very large radius, near the outer rim of the detector. These edge events may be corrupted by distortions of the electric field configuration, poor phonon sensor coverage, and charge-deficit effects along the outer surface. This cut is based upon the relationship between the inner- and outer-electrode ionization signals. `cQin_123` accepts events for which the reconstructed Qouter energy (`qo`) is consistent with noise at the $\pm 2\sigma$ level. Note that this cut is extremely strict at high energy but becomes relatively loose near threshold: an event with 4 keV of ionization, for example, can drop half of its energy into the outer charge channel without failing this cut.

The signal efficiency of the fiducial volume cut was measured using neutron-induced nuclear recoils from the ^{255}Cf calibration runs, following a method developed for Run 119. Efficiencies are generally near 80% near threshold, near the nominal value of 85% based upon the area of the outer charge electrode, dropping to 70-75% near 100 keV. This energy-dependence is a natural consequence of two effects: the cut definition is stricter at low energies (as described above), and neutron multiple-scatter events can fail the cut by depositing some of their energy in the outer electrode. Note that the latter effect does not operate for WIMPs, making this measurement of the fiducial volume a slight underestimate. Preliminary Monte Carlo simulations by Steve Yellin and Xinjie Qiu suggest that the effect of this effect is likely to be of order $\sim 5\%$ in scale. I revisit this systematic error briefly at the end of this Chapter.

6.5.4 Ionization yield

This analysis demands that a WIMP candidate event satisfy two ionization yield criteria:

1. Ionization yield within the 2σ nuclear recoil band
2. Ionization yield less than the 3σ lower edge of the electron recoil band

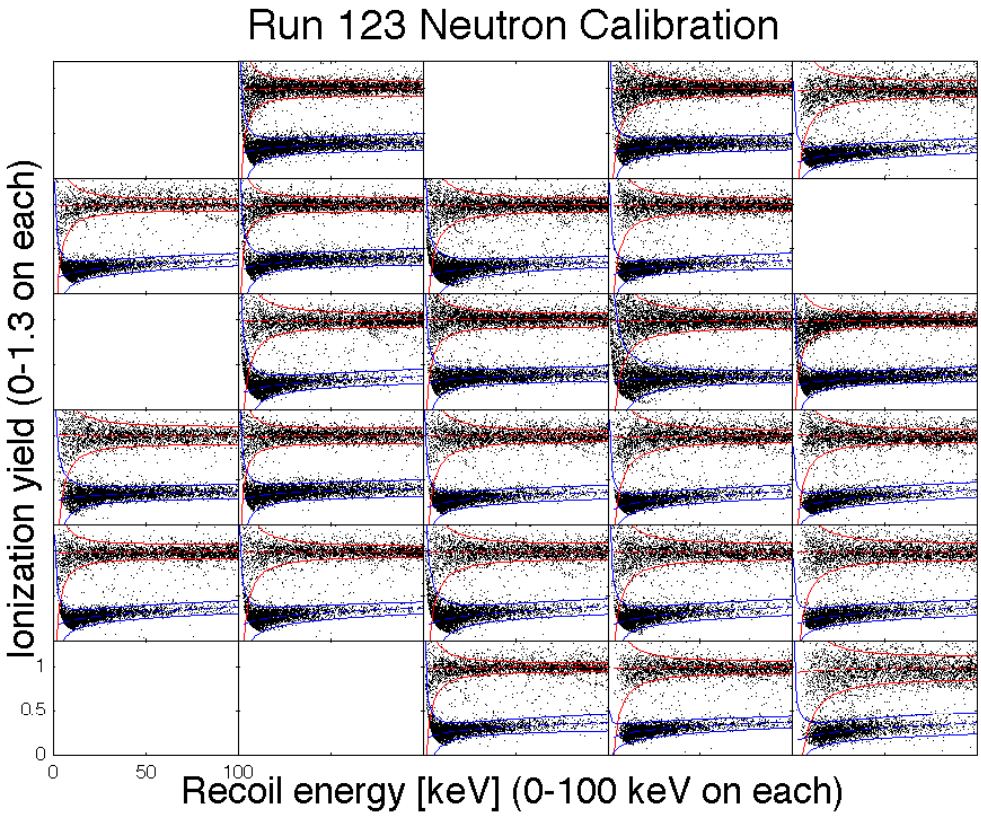


Figure 6.15: Ionization bands from Run 123 ^{252}Cf calibration data for all working ZIP detectors. Colored curves indicate 2σ edges of the electron recoil (red) and nuclear recoil (blue) bands.

The yield bands for the electron and nuclear recoil populations are defined using calibration data sets. In each case we divide the data into energy bins and fit a gaussian to the corrected yield (`yic`) profile in each bin. Simple functional forms are then fitted to the fitted means and standard deviations as functions of recoil energy (`pric`): $\mu_{\texttt{yic}}(\texttt{pric}) = a_1 \cdot \texttt{pric}^{a_2}$, $\sigma_{\texttt{yic}}(\texttt{pric}) = (b_1^2 \cdot \texttt{pric}^{b_2} + b_3^2)/\texttt{pric}$. The nuclear recoil band is fitted using the Run 123 ^{252}Cf sets, as is a preliminary version of the electron recoil band. The final electron recoil band is fit to a sample of the Run 123 ^{133}Ba data. The fitted bands for all good detectors are shown in Figure 6.15. The 2σ neutron band generally has a signal acceptance above 90%.

6.5.5 Surface event rejection

The primary surface event rejection cut for this analysis (`cRT_123`) follows the scheme developed for the Run 119 analysis [188]. This cut is based upon the primary phonon sensor risetime (`pminrtc` = `PXr40-PXr10`) and the delay of the primary phonon pulse with respect to the ionization pulse (`pdelc` = `PXr20-QSOFdelay`+constant), each corrected using the usual phonon correction table. As before, a further empirical correction was applied to each parameter to flatten the upper edge of the surface event distribution.

`cRT_123` has two components, each exploiting the strong positive correlation between `pminrtc` and `pdelc`. The "consistency cut" demands that the difference between these parameters is consistent with the calibration neutron population, *i.e.* within 4σ of the mean neutron value for the chosen detector. This cut has substantial discrimination power itself on some detectors, but is primarily intended to reject "outlier" events for which one or both timing parameters is incorrectly reconstructed.

The bulk of the cut's overall power generally comes from the "discrimination cut", which demands that the sum of the two corrected timing parameters is above a chosen threshold. This serves to separate the slower nuclear recoils from the faster surface event background. The threshold is set separately for each detector to allow a passage fraction of no more than 0.3% for surface events from the Ba calibration. The actual power of the timing cut to reject surface events in WIMP-search data is discussed as part of the surface event background estimate in Section 7.2.

6.5.6 Phonon position reconstruction

Even after all of the cuts described above, a small non-gaussian tail of timing outlier events still extends to low yield and slow timing in the Ba calibration data. These have been reduced by cuts on detector-detector crosstalk and cross-detector pileup (neither of which affect WIMP search data appreciably), but a few such events remain.

Degeneracies in the phonon position reconstruction are one source of outlier events. The partition and delay parameters used in the construction of the phonon correction table both have "foldback" in their radial dependence, *i.e.* the associated radial coordinates are not monotonic. Combining the two parameters eliminates much of the foldback [177], but Figure 6.16 shows that some degeneracy can remain at high radii. These degeneracies can introduce mis-corrections of timing parameters, which generally become slower at large detector radii. As long as high-radius events are corrected using high-radius neighbors there is no problem - the table corrects for the overall trend in timing parameters. Occasionally, however, an event in the outer portion of the detector ($\star$ in Figure 6.16) may be corrected using a set of nearest neighbors ($\circ$ in Figure 6.16) which includes events at lower radii. The slower timing of the outer event will then appear as a deviation from the faster neighbors, pushing the corrected timing parameters to anomalously high values. This mechanism can allow the occasional surface event near the detector rim to pass the phonon timing cut.

Matt Pyle devised a simple cut to exclude the bulk of these degenerate events based upon the Qouter event population, which provides a tracer for the outer detector region. We plot the Qouter events on top of the usual phonon manifold, then exclude from further consideration all events within some radius R from the Qouter events. This acts as a sort of extension to the fiducial volume cut, removing events whose positions are reconstructed dangerously close to the degeneracies of the outer rim. The radius R is set on a detector-by-detector basis to remove a large fraction of outliers at minimal cost in neutron efficiency. The cut has an efficiency of 90-95% on nearly all detectors, with T4Z2 (80%) and T5Z4 (75%) somewhat lower to encompass larger populations of outliers.

This cut is somewhat *ad hoc*, but future analyses can improve upon it substantially. Bruno Serfass has suggested including Qouter events in the construction of the correction table. We can then exclude any event for which a significant number of its nearest neighbors were from the Qouter population. This scheme could not be applied to this analysis due to lack of time.

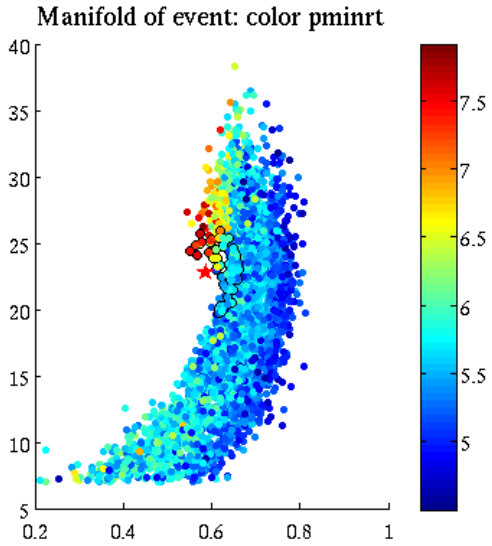


Figure 6.16: Slice of the phonon position correction manifold for T1Z5, plotted in box plot radius (x-axis) versus delay plot radius (y-axis, μs). Colors indicate values of the primary phonon risetime `pminrt`. The star ($\star$) indicates a timing "outlier" event, while the black circles indicate the nearest neighbors of this event used in the correction table. Figure by M. Pyle.

6.5.7 Thresholds

Ionization threshold

The charge threshold cut `cQThresh_123` rejects events for which the ionization signal is consistent with noise. This population has poorly-defined ionization yield and may be contaminated by anomalous events.

`cQThresh_123` enforces that the inner electrode ionization signal `qi` exceeds the noise plateau by at least 4σ, based upon a gaussian fit to the ionization amplitudes for random triggers. To avoid leakage during series with poorer ionization resolution, we measure the mean and width of the qi noise blob independently for each series in R123 and R124, as for the multiplicity cut (see Section 6.5.2). The cut is set for each run at the larger of this run-dependent threshold and the mean threshold from a gaussian fit to the noise blob for the entire run, thus allowing only upward fluctuations in the threshold from a base value.

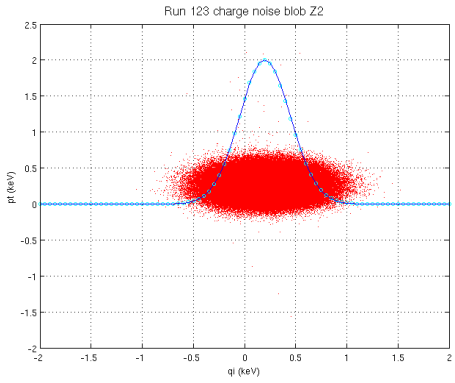


Figure 6.17: Gaussian fit to the `qi` noise blob for T1Z2 used in the ionization threshold cut, superimposed on the noise blob in `qi` and `pt` for this detector. Figure. by X. Qiu.

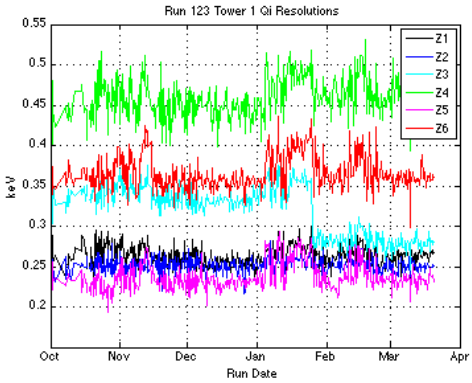


Figure 6.18: Variation in Tower 1 `qi` resolutions during the course of Run 123, as measured using gaussian fits to random trigger events during each data series.

Recoil energy threshold

The phonon threshold is set as a cut on recoil energy (pric), rejecting events at low energies. This population has poorly-defined charge yield and noisy timing parameters, leading to poor discrimination. No explicit cut is placed on phonon amplitude alone, which would generally be weaker than the charge threshold cut. Though our standard analysis can be extended down to at least 5 keV [193, 198] and modified analyses to 1 keV or less, the first analysis of these data for publication used a conservative threshold of no less than 10 keV in recoil energy (`pric`) for all Ge detectors (7 keV for Si). Future analysis may use lower threshold energies.

The recoil thresholds for this analysis are shown in Table 6.2. Note that several detectors have thresholds somewhat higher than the minimum values given above. These thresholds were generally increased to ensure that the following conditions are satisfied:

1. The surface event rejection cut should have significant efficiency.
2. The charge threshold cut and the edge of the electron recoil band must not exclude the nuclear recoil band.
3. There must be a significant population of calibration surface events with which to characterize the cut, which generally means that the 5 sigma lower edge of the electron recoil band must not encompass the nuclear recoil band.

Table 6.2: Recoil energy thresholds in keV for the Run 123/124 blind analysis. *Italics* indicate Si detectors.

	T1	T2	T3	T4	T5
Z1		*7*		*10*	10
Z2	10	*7*	10	10	
Z3		10	*7*	*15*	*7*
Z4	*10*	*7*	10	10	10
Z5	10	10	15	10	10
Z6			10	10	20

6.6 WIMP search exposure

The power of a WIMP search analysis is characterized by its exposure (the sensitivity to observation of WIMP recoils) and its expected background rate (which determines

the significance of any observed events). This section estimates the exposure of this analysis. The expected background is described in Section 7.2.

The WIMP search exposure itself can be divided into two components: a live time of the target mass and an energy-dependent signal efficiency within that live time. The exact dividing line between these two is somewhat arbitrary - only their product is physically meaningful. I generally assign cuts excluding specific time periods or detector mass to the computation of the live time, relegating event-specific (and particularly energy-dependent) cuts to the efficiency.

6.6.1 Live time

Table 6.3: Live times (days) for Run 123 in this analysis, after cuts described in the text. Italics indicate Si detectors.

	T1	T2	T3	T4	T5
Z1		*87.299*		*96.560*	72.464
Z2	91.313	*81.014*	37.918	96.295	
Z3		81.229	*52.723*	*93.124*	*85.781*
Z4	*48.948*	*65.967*	86.521	97.493	98.011
Z5	83.678	95.929	63.177	89.856	97.625
Z6			68.015	89.394	

Table 6.4: Live times (days) for Run 124 in this analysis, after cuts described in the text. Italics indicate Si detectors.

	T1	T2	T3	T4	T5
Z1				*52.348*	
Z2	45.024				
Z3				*53.741*	
Z4		*53.402*		53.608	47.369
Z5	51.148			45.501	53.262
Z6				46.207	

Tables 6.3 and 6.4 list the livetimes computed for each ZIP detector in this analysis. These quantities include all good data series after the following cuts:

- All data quality cuts in Section 6.3.
- ∼`cBadDetRegions_123`, which removes three poorly-performing detector quadrants (see Section 6.4).

- `cQstd_123`, which primarily removes poor time periods of the cryocooler cycle on certain detectors (see Section 6.4).

Summed over all good detectors, the total live time of this analysis is:

- **Ge:** 312.23 + 85.53 = 397.76 kg-d
- **Si:** 61.14 + 15.95 = 77.09 kg-d (53.47 kg-d excluding T2Z1 and T4Z1, which have not been unmasked at the time of this writing.)

6.6.2 Signal efficiency

The efficiencies of the remaining event reconstruction and physics cuts for nuclear recoils are combined into the signal efficiency. The efficiency of each cut is determined for each detector by the methods described above. The total experimental efficiency is a weighted average of these detector-specific efficiencies, weighted by their relative live times from Tables 6.3 and 6.4. Figure 6.19 plots this total efficiency as a function of recoil energy, indicating the cumulative contributions of various cuts. The net efficiency is approximately 30%, somewhat lower than that achieved in the Run 119 analysis but falling off less rapidly at low recoil energies. As in previous analyses, the largest costs in WIMP efficiency come from the fiducial volume and surface event rejection cuts.

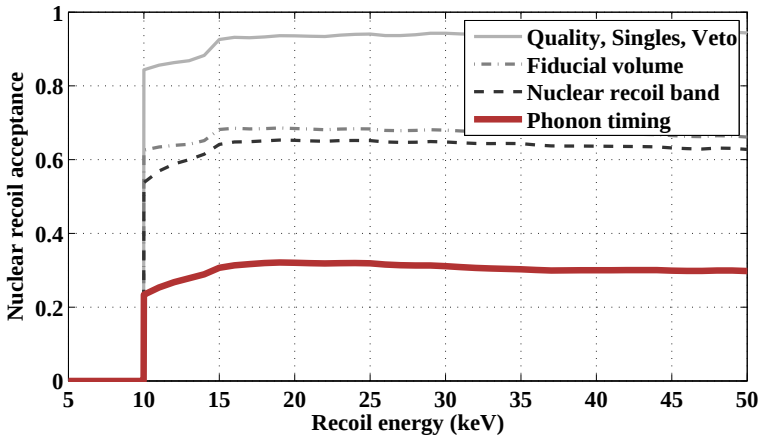


Figure 6.19: Efficiency of the Run 123/124 WIMP search analysis as a function of recoil energy. Curves represent total efficiency after the subsets of cuts described in the caption, with each curve representing the efficiency after adding one set of cuts to those of the curve above.

6.6.3 WIMP spectrum-averaged exposure

It is difficult to fairly compare two experiments with different exposures and efficiency functions, since the shape of this function affects the sensitivity to WIMPs at a given mass. One useful measure of the total exposure of a given experiment is the WIMP spectrum-averaged exposure for a given WIMP mass, defined as the exposure of a hypothetical experiment with 100% efficiency with equivalent sensitivity at a given WIMP mass. This value is given by:

$$\epsilon_{WSA,M} = LT \cdot \frac{\int_{E_{min}}^{E_{max}} dE\, \epsilon(E)\eta_M(E)}{\int_{E_{min}}^{E_{max}} dE\, \eta_M(E)}$$

, where LT is the livetime, E is the recoil energy, $\epsilon(E)$ is the efficiency function, $[E_{min}, E_{max}]$ is the energy range under analysis, and η_M is the expected recoil energy spectrum for a WIMP of mass M.

For this analysis, the spectrum-averaged exposure for a 60 GeV/c^2 WIMP is 121.3 kg-d for Ge, 17.9 kg-d for Si. For comparison, the equivalent numbers for Run 119 were 34 kg-d and 12 kg-d, respectively. Near the minimum of its limit curve, this Ge exposure is thus roughly 3.5$\times$ greater than any previous single WIMP exposure of CDMS.

6.6.4 Systematic errors

This estimate of the WIMP-search exposure of this analysis is subject to several sources of systematic error. One such source is the presence of internal multiple-scatter events in the neutron sample: neutrons which deposit energy in two or more separated locations within a single ZIP. Monte Carlo simulations indicate that $\sim$ 15% of nuclear recoils from neutron calibrations will have this topology. WIMPs should never scatter twice within the same detector, so these events constitute a systematic difference between the calibration data and the signal population. Simulations indicate that the presence of these events in the calibration data leads us to underestimate the efficiency of the fiducial volume cut and to slightly mis-fit the nuclear recoil band with negligible effect on signal acceptance. We do not yet have a quantitative understanding of the effect these events have on the performance of the timing cut: some internal multiples should appear with unusually slow or fast timing parameters, some should fail the consistency cut, and some should be mis-corrected by the lookup table, but the magnitude of the net effect is unclear.

A second, very different source of systematic error comes from inaccuracies in the actual masses of the detector crystals. Our collaboration has traditionally taken the mass of a Ge detector to be 250g and that of a Si detector to be 100g. These are only approximations, however. A more detailed analysis of the detector geometries suggests that a Ge ZIP has a mass closer to 239g (4% lower) and a Si ZIP closer to 106g (6% higher). Even these numbers are only approximate, however, since individual detector thicknesses vary by several tenths of a millimeter (several percent) due to differences in polishing history.

Given our incomplete understanding of some of these effects, it is difficult to give a quantitative estimate of our systematic error. For Ge detectors, the inaccuracies in detector mass and the systematic bias of the fiducial volume cut's efficiency seem to roughly cancel; for Si detectors they add, suggesting that our exposure numbers may be conservative by $\mathcal{O}(10\%)$ for Si. Other effects of internal multiple-scatter events are not well quantified. Altogether, an estimated systematic error of $\mathcal{O}(10\%)$ on the determination of the WIMP-search exposure seems plausible. Further simulation work is needed to characterize these effects more accurately.

Chapter 7

WIMP Search Results

This chapter presents the results of the blind analysis described in the previous Chapter, beginning with a discussion of the expected non-WIMP background rate and culminating in new limits on the rate of WIMP-nucleon elastic scattering. Strictly speaking, the actual scientific result is independent of the expected background: in order to remain conservatively independent of any uncertainties in the background model, the WIMP-search analysis is performed without background subtraction. Nonetheless, the background estimate constitutes a critical evaluation of the performance of this analysis and of the experiment's likely performance in future, longer data runs.

7.1 Nuclear recoil background

Non-WIMP nuclear recoils are the most dangerous background to an event-by-event search for WIMP dark matter, as they cannot be distinguished from WIMPs individually. Such nuclear recoils arise from three major sources: neutrons from cosmic ray interactions, neutrons from radioactive decay, and heavy nuclei recoiling from alpha decays near the detectors. The CDMS installation (described in Chapter 4 is designed to protect the detectors from all three. This section summarizes the expected rate from each of these three sources during the current data run. For further details on simulations of the neutron background, I refer the reader to several CDMS dissertations [199, 200, 201] and to CDMS internal documentation [202].

7.1.1 Cosmogenic neutrons

Neutrons at MeV energies can produce keV nuclear recoils in our detectors which may be mistaken for WIMP interactions. The primary source for such neutrons, both above and below ground, is the interaction of energetic cosmic-ray-induced muons. Such muons can generate neutrons through spallation (muon-induced nuclear disintegration) or through various secondary processes within their consequent hadronic and electromagnetic showers. Since these muons are too penetrating to be stopped by practical amounts of detector shielding, the primary defense against cosmogenic neutrons is experimental depth. At a depth of 2090 meters water equivalent, CDMS II's location at the Soudan Underground Laboratory reduces its incident muon flux by a factor of $\sim 5 \times 10^4$ with respect to the surface value. Nonetheless, the remaining cosmic ray flux ($\sim$ 1 muon per minute in the CDMS veto shield) produces a lingering neutron background which must be evaluated.

Large Monte Carlo simulations are used to model the complexities of neutron production and propagation throughout the cavern rock and the CDMS shield. A series of such simulations were performed for this run using two major software packages for the propagation of particles through matter: GEANT 4 [203, 204] and FLUKA [205, 206], the latter mated to a modified version of MCNPX [207] to model low-energy neutron scattering. The GEANT 4 and FLUKA simulations were carried out by Angela Reisetter and Raul Hennings-Yeomans [201], respectively. Initial runs of the two simulation codes used slightly different initial conditions and different criteria for event identification, but the final simulations described here standardized upon a common set of inputs and definitions.

Both simulations began with an initial sample of muon energies and incidence angles appropriate to the Soudan depth, equivalent to 6.34 live years of exposure underground. This sample was generated using the MUSIC simulation package [208], based upon the geology and geometry of the rock overburden and observations of the muon angular distribution in the mine. The main simulation codes propagated these muons through a 10m-thick shell of rock surrounding the CDMS cavern to produce a series of particle showers, recording a snapshot of the particles which entered the cavern from each incident muon. These showers were then thrown at a simulation of the CDMS installation (detectors, shield, and veto counter) 20 times using different random number seeds, producing the equivalent of 126.8 live years of exposure.

The vast majority of cosmogenic nuclear recoils in this simulation are accompanied

by energy deposits in another active element of the CDMS apparatus (another ZIP or the veto counter), and so would not be identified as WIMP candidates. Some nuclear recoils will not be identified as such because they are coincident with significant electromagnetic energy in the same detector. Of those nuclear recoils which are identifiable, most are part of events which deposit energy in other detectors and so fail the multiplicity cut. Finally, the veto counter tags $\sim 98.8\%$ of all events containing an identifiable nuclear recoil.

Table 7.1 lists the observed event counts from the GEANT 4 simulation; the FLUKA simulation gave consistent results (not listed). We can verify the simulation's performance by testing its predictions of veto-coincident neutrons against experimental data. To maximize the sensitivity to multiple-scatter events, these studies are performed with slightly loosened quality cuts and no application of the timing cut. Scaling down the counts from the Table, the model predicts 0.5 (1.4) veto-coincident neutron singles per kg-year of Ge (Si) exposure, or a total of 0.4 in the current exposure. No such events are observed, in accordance with these expectations. The simulation also predicts 2.3 (7.9) nuclear recoils within multiple-scatter events, or a total of 1.8 in the current exposure. We observe 5 apparent nuclear recoils within veto-coincident multiple-scatter events, but estimate that ~ 2 could be due to surface events leaking into the nuclear recoil band. Again, the observations are consistent with the simulation.

With our confidence in the simulation thus supported, we can estimate the expected cosmogenic neutron background to the present WIMP search. The parenthesized numbers in the last column of Table 7.1 represent the true WIMP candidates observed in the simulation: single nuclear recoils not coincident with any activity in the veto counter or in the other detectors. Scaling down these numbers and including appropriate Poisson errors, we expect $(1.7 \pm 1.7) \times 10^{-3}$ ($(4.3 \pm 1.8) \times 10^{-2}$) candidate events per kg-year of Ge (Si) exposure. This represents $(5.5 \pm 5.5) \times 10^{-4}$ ($(6.3 \pm 2.6) \times 10^{-3}$) expected neutron background events in the present data analysis. Even assigning conservative systematic errors of $\sim 100\%$ to these rates, the cosmogenic neutron background should be essentially negligible for the current run.

It is worth noting that this cosmogenic background rate is substantially lower than that found by most previous simulations of neutron backgrounds underground. Mei and Hime [209] have argued for a background level of ~ 0.5 candidates/kg-year for the CDMS II facility at Soudan, for example, similar to the 0.9 ± 0.4 candidates/kg-year computed in Sharmila Kamat's dissertation [200]. These studies were based upon somewhat different

Table 7.1: Counts of identifiable nuclear recoils (NRs) from the 126.8-year GEANT 4 simulation of the CDMS II neutron background. In events containing multiple NRs, each individual recoil contributes to the count. Parenthesized numbers in the last column are events not tagged by the veto counter. Numbers from an internal note by Angela Reisetter.

	NRs	NRs in multiple-NR events	NRs in other multiples	Single NRs
Ge	1700	675	728	297 (1)
Si	1295	542	561	192 (6)

assumptions about the muon flux at depth, and did not account for the full effects of coincident electromagnetic energy: many cosmogenic neutrons which would otherwise be WIMP candidates may be tagged by shower energy deposited in the veto counter or in other detectors. Based on these new simulations, cosmogenic neutrons should not be a significant background at Soudan until we reach much larger exposures. An array of 42 detectors totaling 25 kg of Ge would begin to see cosmogenic neutrons at an exposure of order ~ 100 kg-years, corresponding to WIMP sensitivities near 10^{-46} cm^2. With a well-designed shield and a partitioned detector, an extremely sensitive WIMP-detection experiment may thus be operated at a relatively shallow depth.

7.1.2 Radiogenic neutrons

Radioactive processes in and around the CDMS installation can also generate MeV neutrons, primarily through $(\alpha\, n)$ reactions and spontaneous fission. The uranium and thorium decay chains are the dominant alpha sources in natural rock and in many refined materials. Uranium is also the dominant source of neutrons from spontaneous fission, while the fission rate of thorium is essentially negligible. The expected rate of radiogenic neutrons thus depends upon the contamination levels of various materials in and around the icebox, as well as geometric factors concerning the rate at which radiogenic neutrons reach the detectors.

Table 7.2 lists the uranium and thorium contamination levels we assume for materials within the outer polyethylene shield. Radiogenic neutrons generated outside the outer polyethylene shield are moderated enough not to pose a significant danger, while the detectors substrates themselves are pure enough ($\ll 1$ ppt U/Th) to contribute negligibly. These concentrations are chosen to lie on the conservative side of gamma-counting measurements at the Oroville counting facility and elsewhere. Note that the extrapolation from gamma

rates to contamination levels often involves assumptions of secular equilibrium among the various elements and isotopes in a decay chain. Chemical and physical processes during the smelting process probably violate these assumptions, introducing a substantial systematic error. Our collaboration has not attempted to quantify this error, and until more information is available I conservatively assume $\sim 100\%$ systematic error bars on the resulting neutron rates.

Table 7.2: U/Th contamination levels assumed in radiogenic neutron simulations. All concentrations are reported by mass: 1 ppb = 10^{-9} g/g.

Component	U conc. (ppb)	Th conc. (ppb)	Mass (kg)
Cu cold hardware	0.2	1	14.95
Cu icebox cans	0.0797	0.299	260
Inner polyethylene	0.12	0.12	120
Inner (ancient) Pb	< 0.05	< 0.05	1917
Outer Pb	< 0.05	< 0.05	12190

The figures for lead contamination in Table 7.2 bear special mention, as they have a substantial bearing on our background from spontaneous fission. The Table lists an upper limit of 50 ppt on uranium and thorium based upon null observations from gamma counts of our lead at the Oroville counting facility. We believe this upper limit to be very conservative: the EXO collaboration has set limits of < 10 ppt U/Th in lead from the same sources [210].

Michael Tarka has performed a Monte Carlo simulation of neutron production from $(\alpha\, n)$ and spontaneous fission processes in the copper cold hardware, the copper icebox cans, and the inner polyethylene shield. Neutron rates and spectra were generated using the SOURCES4A simulation package [211], scaled by the contamination levels and masses in Table 7.2, and propagated through a GEANT 4 simulation of the 5-Tower geometry.

Xinjie Qiu has simulated the contribution of fission neutrons from the Pb shield; $(\alpha\, n)$ processes are less significant in lead due to the absence of light target nuclei for the *alpha* to disintegrate. This simulation was carried out entirely within GEANT 4: neutrons were generated from an imaginary ^{252}Cf source interspersed throughout the Pb shield (under the assumption that the Cf and U neutron spectra are similar) and propagated through the 5-Tower geometry. The resulting rate was normalized by the masses and assumed contamination levels of Table 7.2, the ^{238}U decay rate (1 mBq/kg for 81 ppt U [151]), and the branching ratio for spontaneous fission of ^{238}U (5.45×10^{-7}). This simulation did not separately tag the numbers of recoils in Ge and Si, but the ratio of Ge and Si event rates per detector is extrapolated from a previous simulation of the 2-Tower geometry.

Table 7.3: Expected rates of single-scatter, 10-100 keV nuclear recoils in the detectors from radioactive contamination within the various icebox components, based upon the contamination levels from Table 7.2. All statistical errors are negligible due to the large simulation sizes.

Component	Ge rate (n/kg-yr)	Si rate (n/kg-yr)	Ge 123/4 (n)	Si 123/4 (n)
Cu cold hardware	9.06×10^{-3}	2.59×10^{-2}		
Cu icebox cans	1.83×10^{-2}	4.76×10^{-2}		
Inner polyethylene	5.02×10^{-3}	1.30×10^{-2}		
Total Cu/poly	3.2×10^{-2}	8.7×10^{-2}	1.1×10^{-2}	1.8×10^{-2}
Inner (ancient) Pb	$< 8.0 \times 10^{-3}$	$< 6.3 \times 10^{-2}$		
Outer Pb	$< 2.1 \times 10^{-2}$	$< 1.6 \times 10^{-1}$		
Total Pb	$< 3 \times 10^{-2}$	$< 2 \times 10^{-1}$	$< 1 \times 10^{-2}$	$< 5 \times 10^{-2}$

Table 7.3 summarizes the expected rates of single-scatter nuclear recoils within the 10-100 keV energy range due to these various radiogenic sources. The right columns show the expected number of radiogenic background events in the current CDMS run. Even adding in $\sim 100\%$ systematic uncertainties, we can be confident of having < 0.04 (< 0.1) radiogenic neutrons in Ge (Si) during the current run. Additional Monte Carlo work and counting measurements are underway to improve this model for future data runs at Soudan.

7.1.3 α-induced nuclear recoils

When an atom emits an α-particle, the remaining nucleus recoils in the opposite direction with a kinetic energy of $E_N \simeq \frac{4}{A}Q$, where A is the atomic mass of the initial nucleus and Q is the mass difference between the decay products and the original nucleus. In principle, such an event could appear as a WIMP candidate if this recoiling nucleus deposits its energy within a ZIP detector while the α-particle escapes detection. This process is essentially only possible for α-emitting surface contaminants, since both heavy nuclei and α-particles have extremely short penetration depths through matter ($\sim 0.1\mu$m and $\sim 10\mu$m, respectively). The primary such contaminant is ^{210}Po, a radon daughter with $Q = 5.41$ MeV that yields a recoiling ^{206}Pb nucleus with a kinetic energy of 103 keV.

The ionization yield of such a recoiling nucleus is suppressed below that of a standard nuclear recoil by three distinct effects. A Pb nucleus is much heavier than a Ge or Si nucleus, making it even less ionizing ($\frac{dE}{dx}_e \propto v \propto \sqrt{E/M}$) than a lattice recoil of similar energy. This gives the heavy ion recoil a lower ionization yield than that of a standard WIMP- or neutron-induced recoil. The recoiling nucleus is also a very-near-surface event,

which further suppresses its ionization yield. Finally, a Pb nucleus originating within a few nm of the detector surface or incident from a decay on a neighboring surface (detector or copper housing) deposits much all of its energy in the 40nm amorphous Si layer, from which no ionization is expected to escape. All told, the ionization yield of an α-induced Pb recoil is often low enough to fall below the ionization threshold.

No quantitative estimate of the background contribution from this process was made for this first 5-Tower analysis, but there is good reason to believe that this is not a limiting background within the signal region. Pb recoils visible in coincidence with α-decays are generally found at recoil energies of a few tens of keV and extremely low ionization yields, usually low enough to fail the ionization threshold cut. It is also relatively rare for such an event to pass the multiplicity cut: a particle leaving a detector surface which faces a neighboring detector (i.e. not in the outer surfaces of the Z1 or Z6 positions) within the inner electrode region has an $\sim 84\%$ chance to hit that neighboring detector. This naturally does not hold for particles leaving the outward faces of Z1 or Z6 detectors, so this background may be more significant in these cases.

7.2 Surface event background

Near-surface electron recoils suffer from suppressed ionization yield, as described in Section 3.5, and we expect our timing cut to misclassify some fraction of these surface events as nuclear recoils. These classification errors constitute the major background to this WIMP search. Though the misclassification rate for ^{133}Ba calibration data was studied during the definition of the timing cut (Section 6.5.5), systematic differences between the calibration and WIMP-search samples – notably in face distribution and energy spectrum – affect this rate significantly. In order to be conservative, this analysis bases its surface event background estimate upon minimally-biased samples of surface events chosen from the WIMP-search data sets.

7.2.1 Sources of surface events

Surface electron recoils can originate from several sources:

1. β-particles (electrons) from β-emitters contaminating the detector and icebox surfaces, notably ^{210}Pb (a radon daughter) and ^{40}K.

2. Electrons ejected from neighboring material by photons through the photoelectric effect.

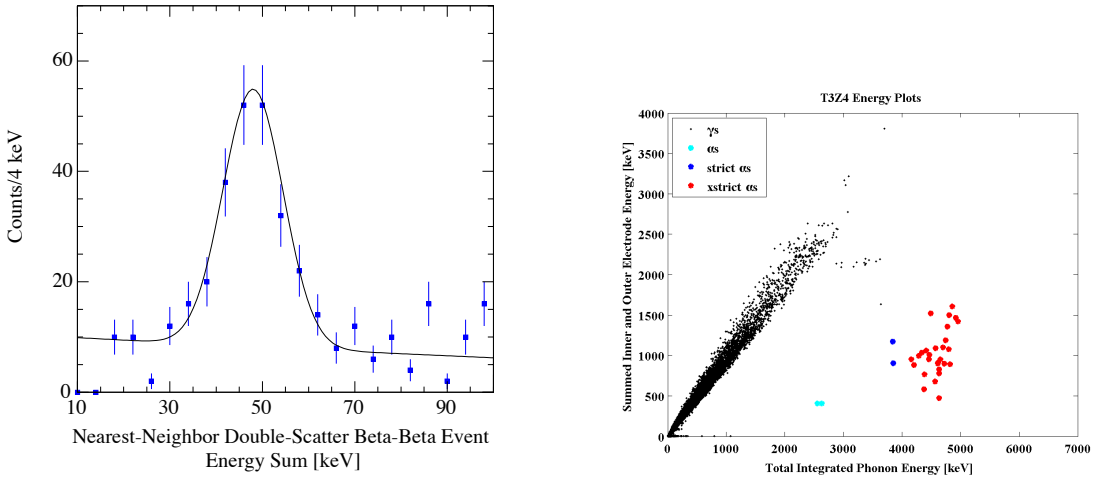
3. Photon interactions which deposit most of their energy within a few microns of a detector surface. Such events can arise from low-energy X-rays with short ranges in Ge or from Compton scattering of higher-energy photons.

The latter two processes scale with the photon background rate in the vicinity of the detectors, while the former depends upon the preparation and handling of our detector and cold hardware surfaces. All of these event categories are expected to be present in the WIMP-search data upon which our leakage estimate is based, and so are incorporated into that single estimate.

Recent work has shown that the dominant sources of single-scatter surface events at Soudan are ^{210}Pb decay and photon-induced backgrounds. Figure 7.1(a) illustrates the primary evidence for ^{210}Pb contamination: a broad $\sim$ 45-keV peak in the summed energies of double-scatter events which register low-yield hits in two nearest-neighbor detectors. A ^{210}Pb nucleus decaying on a detector surface releases $\sim$ 46.5 keV of X-rays, internal-conversion electrons, and Auger electrons; if this surface faces another detector, this energy should be shared among the two ZIPs. The observed peak is roughly consistent with our expectations for this process. A population of low-yield α-particles is also visible near 5 MeV in Figure 7.1(b), consistent with those from the decay of the daughter nucleus ^{210}Po. These two processes should occur at the same rate, up to the effects of geometry and cut acceptance, and Figure 7.2 shows that they are indeed strongly correlated. Based upon a multi-parameter contamination model fitted to the observed event rates in Runs 119 and 123, these data are consistent with approximately equal contributions from ^{210}Pb and from photon-induced backgrounds. It is noteworth that Towers 3-5 have significantly lower fitted contaminations, consistent with improved radon abatement during fabrication and testing.

7.2.2 Estimation framework

An estimate of the surface event background of this data analysis presents several challenges. We have no model of the detector's response at the level of accuracy needed to characterize timing cut leakage, particularly given the significant variation in timing response among detectors. This means that we cannot make full use of the our plentiful ^{133}Ba calibration data when determining the performance of our timing cut, since we cannot



(a) (b)

Figure 7.1: Signatures of ^{210}Pb contamination in Soudan WIMP-search data. (a) Histogram of the summed energies of surface-event/surface-event coincidences between neighboring detectors, showing a broad peak consistent with ^{210}Pb contamination. (b) Scatter plot of phonon and ionization energies for T4Z4, illustrating a population of ~ 5 MeV α-particles.

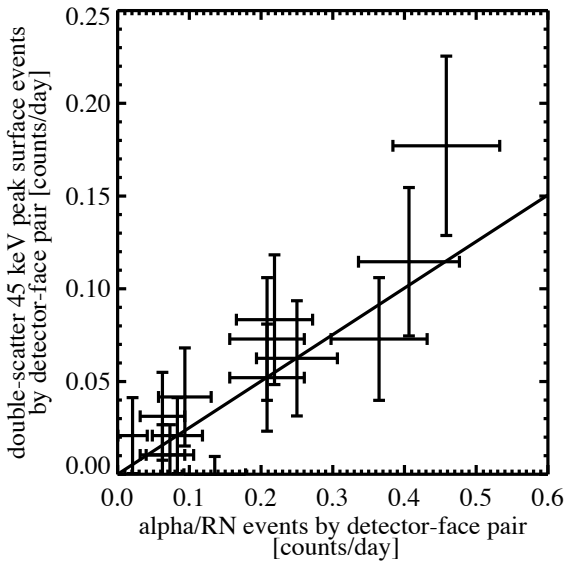


Figure 7.2: Correlation between the event rates of surface-event/surface event double scatters within the 45-keV peak (*cf.* Figure 7.1(a)) and α-particles (*cf.* Figures 7.1(b)) during Run 123. The overlaid line is the best-fit contamination model based on Run 119 and Run 123 data.

account for the systematic differences between this sample and the WIMP-search data reliably. We must instead base our estimate upon observed event counts from the WIMP-search data set, which has much poorer statistics. Worse yet, in order to achieve sufficient rejection power we must set our cut far our in the tail of the surface event distribution where statistics are very poor and empirical fits are unreliable. Finally, we must account for the systematic differences in performance between endcap detectors – those in the Z1 or Z6 position, with a face that does not point toward another detector – and those interior to a Tower; the outwardly-directed face of an endcap detector will observe many more single-scatter events than an interior detector surface.

The surface event background estimate used in this analysis is a conservative response to these challenges, based solely upon events observed outside the signal region in the WIMP-search data set. This Section describes the model used in the estimate, while the statistical technique is described in Appendix C.

Terminology

In developing the surface event leakage model, I consider several classes of events in and around the WIMP-search signal region. I list the most important of these classes below for later reference, as well as the symbols I use to identify the associated event counts. Quantities identified by uppercase (lowercase) letters represent numbers of surface events failing (passing) the timing cut. The subscript i is an index that identifies the detector.

1. **Single-scatter events within the nuclear recoil band** (n_i, N_i): These represent events within the signal region, concealed by the blinding mask until all event-selection cuts are defined. We seek to estimate $\langle n \rangle = \sum_i \langle n_i \rangle$, the expected number of surface events within the WIMP-search signal region.

2. **Multiple-scatter events within the nuclear recoil band** (b_i, B_i): This population should be very similar the single-scatter population above, up to possible systematic differences between single- and multiple-scatter events.

3. **Multiple-scatter events outside the nuclear recoil band** (m_i, M_i): This population has higher statistics than the last, but is less similar to the actual background event population (n_i) due to differences in ionization yield, energy spectrum, and multiplicity.

4. **Single-scatter events outside the nuclear recoil band** (s_i, S_i): This sample may suffer from yield- and energy-related systematics, but not from multiplicity-related ones. This analysis employs this sample only for the endcap detectors, for which multiple-scatter events do not represent single-scatter events well.

I also refer to surface events which are known to occur on one or the other of the two detector faces. I assign an event to the phonon (charge) face if it is part of a multiple-scatter event with its nearest neighbor detector with lower (higher) detector number. This method does not identify the faces of all surface events, and in particular cannot identify the face of single-scatter events or events on outwardly-directed faces. A surface event in T1Z5 that also deposits energy in T1Z4 is assigned to T1Z5's phonon face, for example, while an event depositing energy in neither T1Z4 nor T1Z6 (or in both) is not assigned to either face. Where necessary, I will add subscripts P and Q to the quantities in the list above to indicate events from these categories tagged as occurring on the phonon-sensor or charge-sensor face of a detector, respectively.

Basic model

The expected total number $\langle n \rangle$ of single-scatter surface events in the nuclear recoil band passing the timing cut is a sum of contributions from individual detectors:

$$\langle n \rangle = \sum_i \langle n_i \rangle = \sum_i \langle N_i \rangle r_i. \tag{7.1}$$

In this expression, the sum is over all good detectors of a given substrate type (Ge or Si), labeled by an index i. r_i is the expected ratio of events passing the timing cut to those failing, estimated from a separate event sample. In other words, in order to estimate the total number of single-scatter surface events within the nuclear recoil band passing the timing cut, we estimate the number of such events failing the timing cut and multiply by a pass/fail ratio.

A few comments on Equation 7.1 are in order. By using the number N_i of events *failing* the timing cut on the right-hand side, rather than the total number of events *before* the timing cut (i.e. $N_i + n_i$), we render the estimate statistically independent of the actual number of WIMP candidates. Note also that we allow r_i to vary between detectors – that is, we do not assume that all detectors show the same timing cut performance. This choice carries a statistical penalty but avoids systematic biases. Finally, note that the use of

expectation values in this expression reflects the fact that the *expected* surface event leakage depends upon the *expected* event rates. The observed counts have Poisson noise on top of this "true" value, a distinction that is important when computing errors on our estimate.

The primary challenge comes in selecting a sample of surface events from which to estimate r_i. This sample should represent the population of background events – single-scatter surface events within the nuclear recoil band – with minimal bias. In the current analysis we base our estimate on three such samples chosen from the masked WIMP search data: multiple-scatter events within the nuclear recoil band (b_i, B_i), multiple-scatter events outside the nuclear recoil band (m_i, M_i), and single-scatter events outside the nuclear recoil band (s_i, S_i). The techniques used to estimate r_i for interior and endcap detectors are described separately in Sections 7.2.3 and 7.2.4.

Statistical technique

With a leakage model in place, we are left with a statistical problem: how to compute the expected total leakage and associated error bar. The error bar is the primary challenge here: no events pass the timing cut on most detectors, calling into question the application of many classical statistical techniques based on asymptotic assumptions (*e.g.* profile likelihood intervals [212]). This framework also contains numerous nuisance parameters (*e.g.* the r_i's, which are not themselves of interest) and correlations among the estimates for the various detectors, all of which complicate a classical computation.

In order to account for these nuisance parameters and correlations in a straightforward manner, I pursue a Bayesian estimate of the expected background. This estimate is based upon a Monte Carlo simulation, thrown in accordance with the observed event counts and a choice of prior distributions. We take the median of the total leakage from the simulated ensemble as our parameter estimate, and its [15.87%, 84.13%] percentile interval as our 68% credibility interval. I choose the Bayesian prior to ensure good classical properties for the resulting estimator, approximately minimizing its bias and ensuring adequate coverage. The sections that follow list the estimators use for the r_i's and the resulting estimates. Details of the statistical technique are relegated to Appendix C.

7.2.3 Interior Ge detectors

We use two distinct techniques to measure the pass/fail ratios r_i for each Ge detector interior to a Tower. The final estimate of each r_i is a statistical combination of $r_i^{(1)}$ and $r_i^{(2)}$, made using a Bayesian technique described Appendix C.

First estimate

The simplest estimate of the background for each interior detector uses the multiple-scatter events within the nuclear recoil band to measure the pass/fail ratios r_i from Equation 7.1:

$$r_i^{(1)} = \frac{\langle b_i \rangle}{\langle B_i \rangle}. \tag{7.2}$$

These events should have the same ionization yield and face distribution as events within the signal region, so the bias of this estimate should be minimal. The Achilles' heel of this method is its low statistics: B_i is of order ~ 30 and $b_i = 0$ for most interior detectors. Equation 7.2 thus gives little bias but large statistical errors.

To estimate a systematic error upon this estimate, I consider contributions from two effects. The first is from bias or undercoverage of the estimator itself. These classical characteristics are not strong concerns in most Bayesian work, but they provide a useful benchmark when we have little prior information. As described in Appendix C, I assign a systematic error of ± 0.05 to account for the likely range of variation in bias and undercoverage. I also consider the contribution of possible variations in timing performance between the WIMP-search single- and multiple-scatter events. Observed event counts outside the nuclear recoil band give weak evidence that the timing cut may actually pass somewhat fewer single-scatter surface events than multiple-scatter surface events ($0.39 \pm 0.40\%$ versus $1.08 \pm 0.36\%$), which would lead this technique to overestimate the expected leakage. I assign a 10% systematic error to the expected leakage due to this effect, since even this weak evidence makes it unlikely that the timing cut could be much more permissive for single-scatter events.

All told, this technique predicts a total surface event background across the 12 interior detectors of

$$\langle \hat{n}^{(1)}_{int\ Ge} \rangle = 0.59^{+0.54}_{-0.30}(stat.) \pm 0.10(syst.) \text{ events}. \tag{7.3}$$

The estimate is essentially identical to that obtained with asymptotic approximations to classical techniques.

Second estimate

To improve our statistics, we can also consider the timing cut performance of multiple-scatter events outside of the nuclear recoil band. These events differ from their counterparts in the nuclear recoil band, however, in two ways we must correct for:

1. Events outside the nuclear recoil band are partitioned differently between the two detector faces. In Ge ZIPs, charge-face events have higher ionization yields and slower timing than phonon-face events, so this higher-yield sample will pass the timing cut more often than our true background sample will.

2. There may also be additional systematic biases not accounted for by face distribution. Bruno Serfass has shown that such differences do exist, primarily as a result of variations of energy spectrum with yield.

To correct for these two systematic biases, we use a somewhat more complex expression for r_i:

$$\begin{aligned} r_i^{(2)} &= \Phi_i r_{Pi}^{(2)} + (1-\Phi_i) r_{Qi}^{(2)} \\ &= \Phi_i \left(\alpha_P \frac{\langle m_{Pi} \rangle}{\langle M_{Pi} \rangle} \right) + (1-\Phi_i) \left(\alpha_Q \frac{\langle m_{Qi} \rangle}{\langle M_{Qi} \rangle} \right). \end{aligned} \tag{7.4}$$

To account for the difference in face distribution, we compute separate pass/fail ratios $r_{P/Qi}^{(2)}$ for the two detector faces, based upon face-tagged multiple-scatter events outside the nuclear recoil band. The effective pass/fail ratio for all events is a weighted sum of these two quantities, weighted according to the fraction $\Phi_i \equiv \langle N_{Pi} \rangle / (\langle N_{Pi} \rangle + \langle N_{Qi} \rangle)$ of events within the nuclear recoil band which occur on the phonon face. The correction factors α_P and α_Q account for any remaining differences in timing performance between surface events within and outside the nuclear recoil band. These correction factors are defined by $\alpha_{Pi} \equiv \frac{\langle n_{Pi} \rangle / \langle N_{Pi} \rangle}{\langle m_{Pi} \rangle / \langle M_{Pi} \rangle}$.

Due to blindness constraints and very limited statistics, these new quantities Φ_i and $\alpha_{P,Qi}$ cannot be computed directly as described above. To compute Φ_i, we assume that multiple-scatter events within the nuclear recoil band have the same face distribution as

their single-scatter brethren, allowing us to compute

$$\Phi_i \equiv \frac{\langle N_{Pi} \rangle}{\langle N_{Pi} \rangle + \langle N_{Qi} \rangle} \simeq \frac{\langle B_{Pi} \rangle}{\langle B_{Pi} \rangle + \langle B_{Qi} \rangle}. \tag{7.5}$$

For the correction factors, even the replacements $n \to b$, $N \to B$ do not provide sufficient statistics to compute these quantities for individual detectors. Instead, on the assumption that event energy spectra do not vary much between detectors, I compute common values for all ZIPs of the same substrate material by summing the statistics using a slightly-loosened version of the timing cut (quantities with tildes below):

$$\alpha_{P,Q} = \left(\frac{\sum_i \tilde{b}_{P,Qi}}{\sum_i \tilde{B}_{P,Qi}} \right) \Big/ \left(\frac{\sum_i \tilde{m}_{P,Qi}}{\sum_i \tilde{M}_{P,Qi}} \right). \tag{7.6}$$

To compensate for this assumption I impose a large systematic error, equal in size to the large statistical error. For the good interior Ge detectors, the correction factors are then:

$$r_P = 0.39^{+0.34}_{-0.20}(stat.)^{+0.34}_{-0.20}(syst.) \tag{7.7}$$

$$r_Q = 1.46^{+0.64}_{-0.46}(stat.)^{+0.64}_{-0.46}(syst.) \tag{7.8}$$

This method gives a somewhat lower estimate of the expected leakage, with better statistics but larger systematic errors:

$$\langle \hat{n}^{(2)}_{int\ Ge} \rangle = 0.23^{+0.20}_{-0.11}(stat.)^{+0.15}_{-0.10}(syst.) \text{ events.} \tag{7.9}$$

The systematic error above contains all contributions from the first estimate, as well as the systematic errors on $\alpha_{P,Q}$ given above.

Combined estimate

The two estimates above cannot be combined naively, since the presence of the N_i in both introduces correlations among them. These correlations are accounted for by computing both estimates jointly as part of a single Monte Carlo simulation, as described in Section C.5. Combining these two and accounting for the systematic errors listed above (some correlated between the two estimates), we obtain a total estimate for the interior Ge detectors of

$$\langle \hat{n}_{int\ Ge} \rangle = 0.26^{+0.18}_{-0.11}(stat.)^{+0.11}_{-0.10}(syst.) \text{ events.} \tag{7.10}$$

7.2.4 Endcap Ge detectors

The endcap detectors present a further complication problem to the leakage estimate, since single- and multiple-scatter events are distributed differently across the two detector faces. Our only sample of surface events on the exterior face (i.e. the phonon face of a Z1 detector or the charge face of a Z6 detector) are single-scatter events outside the nuclear recoil band, but we cannot determine upon which face these single-scatter events occur.

I follow the general format of Equation 7.4, writing r as a sum of contributions from the two faces, weighted by the inferred division between the two faces within the nuclear recoil band (Φ_{Z1} or Φ_{Z6}). We use multiples within the nuclear recoil band for the interior (charge) face's contribution, as in Equation 7.2. The vast majority of single-scatter Ge events occur on the exterior (phonon) face, and so we expect that the pass/fail ratio of these events is slightly higher than that of a pure sample of phonon-face events. We can conservatively take the phonon-face pass/fail ratio to be that of the single-scatter events, giving the following expression for Z1 detectors:

$$r_i^{(Z1)} = \Phi_{Z1}\left(\alpha_P \frac{\langle s_i\rangle}{\langle S_i\rangle}\right) + (1-\Phi_{Z1})\frac{\langle b_{Qi}\rangle}{\langle B_{Qi}\rangle}. \tag{7.11}$$

We do much the same for Z6 detectors, now inferring the exterior (charge) face performance from the singles. For the exterior face, we assign a fraction ϕ_{Z6} of failing singles to the phonon face, where ϕ is the analogue of Φ for events outside the nuclear recoil band. To be conservative, I assume that all single-scatter events passing the timing cut are on the exterior face:

$$r_i^{(Z6)} = \Phi_{Z6}\frac{\langle b_{Pi}\rangle}{\langle B_{Pi}\rangle} + (1-\Phi_{Z6})\left(\alpha_Q \frac{\langle s_i\rangle}{(1-\phi_{Z6})\langle S_i\rangle}\right). \tag{7.12}$$

Equations 7.11 and 7.12 contain new quantities $\Phi_{Z1,Z6}$ and $\phi_{Z1,Z6}$ which describe the face distribution of single-scatter surface events on the endcap detectors. To infer these values, I assume that endcap detectors have the same intrinsic distribution of events across the two faces but that all events on the outward face are classified as single-scatters. More explicitly, we define

- $\Sigma \equiv \frac{\langle N_i\rangle}{\langle B_i\rangle}$, the ratio of single- to multiple-scatter events within the nuclear recoil band, and
- $\Gamma \equiv \frac{\langle B_{Pi}\rangle + \langle B_{Qi}\rangle}{\langle B_i\rangle}$, the fraction of events tagged as nearest-neighbor multiple-scatters,

and use mean values of Σ and Φ for all interior Ge detectors. With these expressions (and analogous ones for ϕ, σ, and γ with $B \to M$), we obtain

$$\Phi_{Z1} = \frac{\Phi(\Sigma + \Gamma)}{\Sigma + \Gamma\Phi} = 0.94 \pm 0.01(stat.) \pm 0.04(syst.) \tag{7.13}$$

$$\Phi_{Z6} = \frac{\Sigma\Phi}{\Sigma + \Gamma(1 - \Phi)} = 0.59 \pm 0.04(stat.) \pm 0.14(syst.) \tag{7.14}$$

$$\phi_{Z6} = \frac{\sigma\phi}{\sigma + \gamma(1 - \phi)} = 0.49 \pm 0.02(stat.) \pm 0.14(syst.). \tag{7.15}$$

I assign large systematic errors based loosely upon the spread in values of Φ and Σ among the interior detectors.

Based on the observed event counts, these expressions give the following expected background from the three good Ge endcap detectors (T5Z1, T3Z6, T4Z6):

$$\langle \hat{n}_{endcap\ Ge} \rangle = 0.29^{+0.41}_{-0.20}(stat.)^{+0.3}_{-0.1}(syst.) \text{ events.} \tag{7.16}$$

Note that this background value is almost solely due to the contribution of T4Z6, with $s_{24} = 2$.

Total Ge surface event background

The leakage estimates for the interior and endcap detectors can be combined by using a single Monte Carlo to compute both at once, thus properly accounting for all correlations between them. Combining statistical and systematic errors (with care not to count any systematic error twice), we obtain the following estimate of the total surface event background:

$$\langle \hat{n}_{Ge} \rangle = 0.59^{+0.49}_{-0.27}(stat.)^{+0.32}_{-0.15}(syst.). \tag{7.17}$$

7.2.5 Si detectors

The same methods may also be applied to estimate the surface event background on the Si detectors. Due to the relatively poor timing performance of T2Z1 and T4Z1 (especially the former), at the time of this writing the collaboration has only proceeded with the unblinding of the six interior Si detectors: T1Z4, T2Z2, T2Z4, T3Z3, T4Z2, and T5Z3. I discuss the leakage estimate for these below, following the methods of Section 7.2.3.

First method

Just as for the internal Ge detectors, the simplest background estimate for the internal Si detectors comes from multiple-scatter events within the nuclear recoil band. Following Equation 7.2 and incorporating similar systematic error contributions, we obtain a leakage estimate of

$$\langle \hat{n}^{(1)}_{int\ Si} \rangle = 1.08^{+0.93}_{-0.56}(stat.) \pm 0.12(syst.) \text{ events.} \tag{7.18}$$

Second method

In order to proceed with a second estimate using multiple-scatter events outside the nuclear recoil band, we need an estimate of the factors α_P and α_Q used in Equation 7.4. Unfortunately, similar estimates of these quantities for Si detectors suggest values of order $\sim 5 - 10$ (with very large errors), rather than of order unity. This indicates that multiple-scatters outside of the nuclear recoil band are not a reliable proxy for events within the signal region, and thus that Equation 7.4 is not useful. This large discrepancy is likely to be due to the low energy threshold of the Si detectors: the Si timing cut shows its worst performance at low energies, but because of the curvature of the `cBelowER_123` cut there are very few events from 7-10 keV satisfying the surface event definition and outside the nuclear recoil band. We therefore take Equation 7.18 as our surface event background estimate for the Si detectors:

$$\langle \hat{n}_{Si} \rangle = 1.08^{+0.93}_{-0.56}(stat.) \pm 0.12(syst.) \text{ events.}$$

7.3 Unblinding the WIMP search data

7.3.1 Ge unblinding

The WIMP search data from this analysis were unmasked on the evening of February 5, 2008, for the "good" Ge detectors listed in the previous chapter. Before the final unmasking, we removed the blinding cut and counted the total number of single-scatter events *failing* the timing cut. The resulting event counts are shown in Table 7.4. The observed total of 97 such events is consistent with expectations of 96 ± 15 events, based on the inferred ratio of singles/multiples for low-yield Ge events. Note the disproportionate number of singles on the endcap detectors, as expected from the reduced power of the multiplicity cut for these detectors.

Table 7.4: Number of events in the unmasked Run 123/124 good Ge and Si detectors *failing* the surface event cut but passing all other cuts. These are the N_i values used in the surface event leakage estimate of Section 7.2.

	T1	T2	T3	T4	T5
Z1					30
Z2	9	*6*	1	4	
Z3		5	*9*	*2*	*10*
Z4	*7*	*14*	4	3	3
Z5	8	2	0	5	4
Z6			6	13	

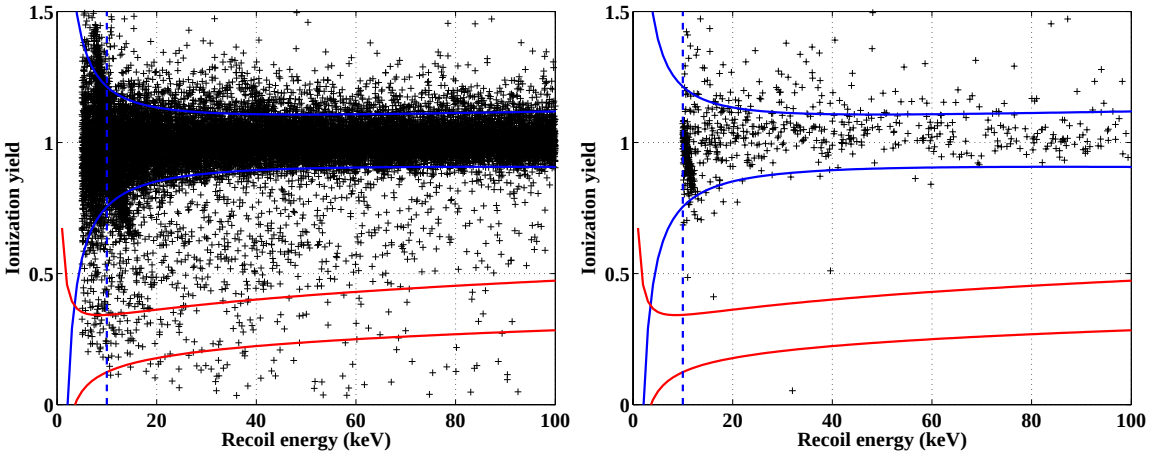


Figure 7.3: Ionization yield versus recoil energy in all Ge detectors included in this analysis for events passing all cuts except (*left*) and including (*right*) the timing and yield cuts. The data are overlaid with the 2σ bands and energy threshold from T1Z2; note that these are merely representative – the actual curves vary between detectors.

Since no obvious anomalies were observed, we unmasked the signal box to complete the blind analysis. Figure 7.3 overlays the ionization and yield values of all Ge events from this WIMP-search run before (*left*) and after (*right*) applying the timing cut. No candidate events were observed satisfying all cuts, and so no dark matter signal has been observed in this exposure.

As a check of our leakage estimate, we can consider in more detail the four events near the nuclear recoil band passing the timing cut shown in the right pane of Figure 7.3.

- One event occured on T1Z5, an interior detector. Following the methods of Section 7.2, the expected number of surface events in interior detectors passing the timing cut and outside the nuclear recoil band is $2.1^{+0.8}_{-0.6}(stat.)^{+0.3}_{-0.2}(syst.)$.

- Two events occured on T4Z6, an endcap detector with an open charge face. These are used as input to the leakage estimate, leading T4Z6 to contribute almost half of the total expected leakage of this run.

- The last event occured on T5Z1, an endcap detector with an open phonon face. Due to its intermediate energy and extremely low yield (`yic`=0.05), this may be a recoiling Pb nucleus from an α-decay. This population is not expected to be a significant background in the nuclear recoil band during this run.

This set of events is thus consistent with our model of the surface event background rate, though not in an especially strong sense – two of the four are inputs to the model, while one is outside of the model's scope.

7.3.2 Si unblinding

The Si detectors were not unmasked at the time of the Ge unblinding due to their higher expected background from surface events. In the interests of preparing a prompt report of the stronger Ge result, the Si detectors were left masked for later analysis.

We revisited the Si detectors in November, 2008. The endcap detectors T2Z1 and T4Z1 were discarded from further study due to poor timing cut performance (for T2Z1 alone, 9 low-energy surface events outside the signal region passed the timing cut) and difficulties with the leakage estimate. Xinjie Qiu attempted to tighten the timing cut slightly on the remaining detectors to lower the background rate, but found that poor statistics in the beta

calibration rendered this difficult. In the end, the collaboration decided to accept the higher expected leakage and proceed to unblind the six interior Si detectors.

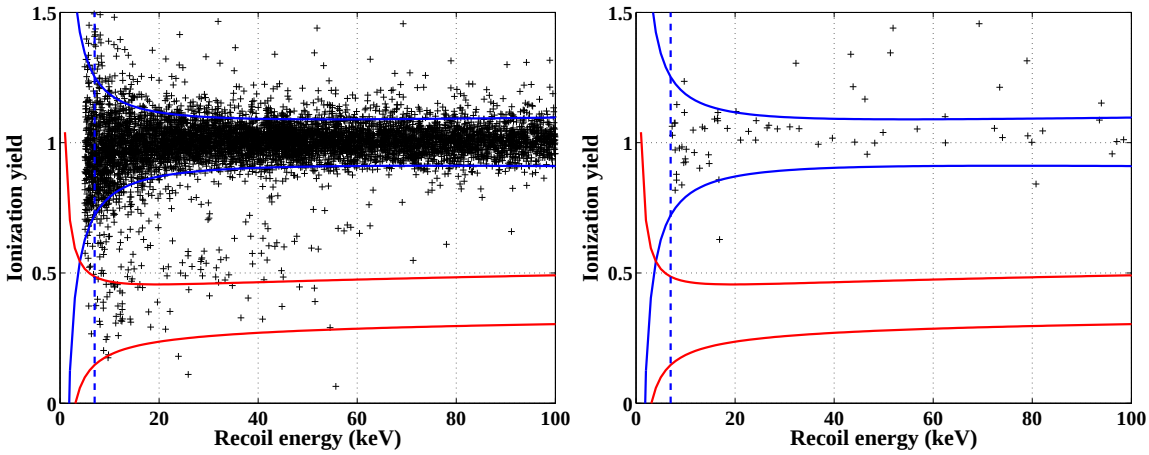


Figure 7.4: Ionization yield versus recoil energy in all Si detectors included in this analysis for events passing all cuts except (*left*) and including (*right*) the timing and yield cuts. The data are overlaid with the 2σ bands and energy threshold from T2Z4; note that these are merely representative – the actual curves vary between detectors.

The Si detectors were unmasked on December 3, 2008. Figure 7.4 shows the distribution of event energies and yields before and after the timing cut. Once again, no candidate events were observed passing all cuts.

7.4 Limits on WIMP-nucleon interactions

These null results set upper limits on the rate of WIMP-nucleus elastic scattering, and thus on the properties of WIMPs. Below I interpret these results as limits on the WIMP-nucleon spin-independent and spin-dependent scattering cross sections as functions of the WIMP mass. For easy comparison with other experiments, I follow the framework described by Lewin and Smith [143] and assume the "standard" halo model described therein: a Maxwellian velocity distribution with a characteristic velocity of $v_0 = 270$ km/s and a local density of $\rho_0 = 0.3$ GeV/c^2. When computing the expected recoil spectra I use the Helm nuclear form factor from that reference for spin-independent scattering and the analogous spin structure functions from Ressell *et al.* [147] (Si) and Dimitrov *et al.* [148] (Ge) for spin-dependent scattering.

The CDMS limit computations are performed using a custom Matlab package de-

veloped over the past several years by Richard Schnee, Steve Yellin, and me. The computation accounts for the experimental live times and the energy-dependent efficiency functions from Section 6.6, as well as the energies of any observed candidate events. The finite energy resolution of the experiment is not currently taken into account in these limit computations, but previous work has shown that this makes no significant difference for WIMP masses above a few GeV/c^2. To be conservative, no background subtraction is attempted – all candidate events are assumed to be WIMPs for the purposes of setting an upper limit. All CDMS limits are calculated using Yellin's optimum interval method [213, 214], a statistical technique for setting an upper limit upon the normalization of a signal of known spectrum amidst an unknown background. The optimum interval method generally sets a stronger limit than that from Poisson statistics in the presence of candidate events, but reduces to the usual Poisson-statistics limit in the zero-background case. Note that the optimum interval method sets one-sided limits, and so cannot be used to detect a signal. Future analyses using more detailed models of surface event backgrounds should allow the use of a Feldman-Cousins technique [215] to set two-sided limits, and thus more easily claim a discovery.

7.4.1 Spin-independent interactions

Figure 7.5 illustrates 90% confidence level upper limits on the spin-independent WIMP-nucleon scattering cross section from current CDMS data, computed within the framework described above. The dashed curves give the results of the present data analysis for the Ge (*black*) and Si (*gray*) detectors. Dashed-dotted curves illustrate the combined upper limits from the previous two Soudan data runs (118 and 119), based upon a recent reanalysis by R.W. Ogburn [198]. Note that this previous analysis corrected an error in the exposure calculation for Run 119, but was not background-free: the Ge analysis observed two events near its energy threshold (5.3 and 7.3 keV), while the Si observed two at intermediate energies (34.9 and 94.1 keV, both in T2Z6), consistent with expected surface event backgrounds. The solid curves give combined results from the entire CDMS II exposure at Soudan, a naive combination of the exposures and observed events of the two separate analyses. The Ge combined result is 3.4× stronger than that of the 118/119 reanalysis at the minimum of its limit curve. The Si result from the current analysis is 3× stronger than that of the 118/119 analysis at its minimum, but the naive combined limit is actually

somewhat weaker than that from the present data set alone. This strange situation is due to the observed Si events from the 118/119 analysis; we do not believe that these are WIMP candidates, but we nonetheless follow our usual conservative procedure by assuming that they are.

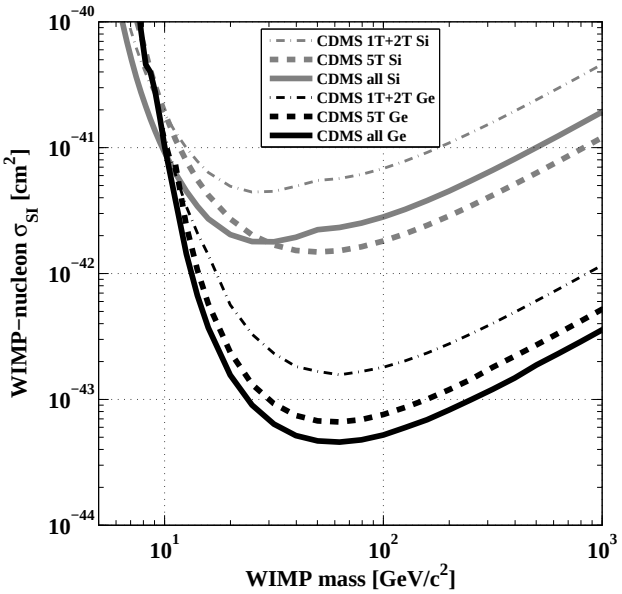


Figure 7.5: Upper limits (90% C.L.) from CDMS II data sets on the spin-independent WIMP-nucleon cross-section as a function of WIMP mass. The figures shows three curves each for Ge and Si: the result of a re-analysis [198] of the previous data from Soudan (Runs 118 and 119), the current data set, and the naive combination of the two.

Figure 7.6 compares the combined upper limits from Figure 7.5 against the results of other leading experiments. The other plotted curves show recent limits from EDELWEISS I[216] (Ge), WARP [217] (Ar), ZEPLIN III [218] (Xe), and XENON10 [158] (Xe). Note that considerable uncertainty remains in the scintillation yield of low-energy nuclear recoils in liquid xenon, and thus in the interpretation of data from liquid xenon detectors; the dotted line in Figure 7.6 represents the effect of recent measurements of scintillation efficiency [219, 220] upon the published XENON10 limit. The combined CDMS upper limit is the strongest yet reported at masses $M_\chi \gtrsim 44$ GeV/c^2.

To put these results in context, Figure 7.6 also indicates three regions of WIMP parameter space of special theoretical and experimental interest. The largest (*dark gray*) filled region indicates the range of parameter space predicted by models in the LEEST (low-energy effective supersymmetry theory) framework, a broad selection of supersymmetry

models described by Ellis *et al.* [221]. This region is representative of the predictions of relatively unconstrained models; note that the allowed region extends down to $\sigma_{SI} \sim 10^{-48}$ cm^2 (10^{-3} zb), well below the figure's lower limit. The green and blue regions indicate parameter space regions suggested by more constrained CMSSM models, identified by Roszkowski *et al.* [157] to be consistent with a variety of experimental results at the 90% and 68% levels, respectively. In both the unconstrained and constrained cases, current experiments are beginning to probe the most optimistic models of supersymmetry. The light gray region at higher cross sections is consistent at the 3σ level with the DAMA/LIBRA annual modulation signal [222], as interpreted within the standard halo model by Savage *et al.* [223]. This region is inconsistent with the null results of all other experiments within this model framework.

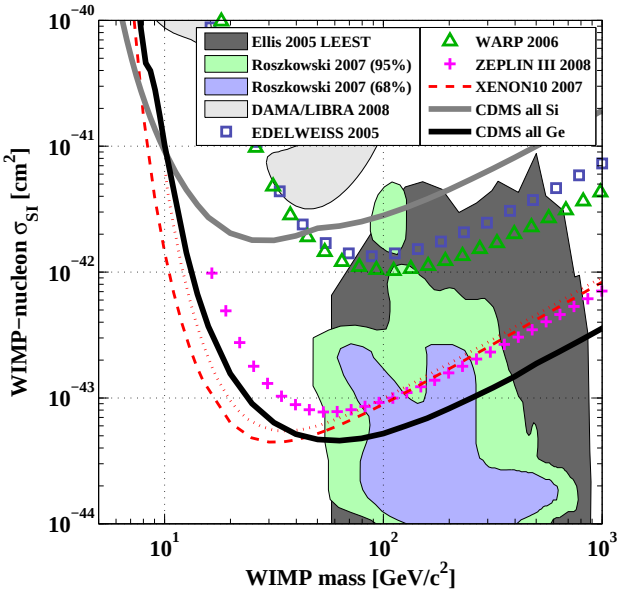


Figure 7.6: Upper limits (90% C.L.) from CDMS II and other leading experiments on the spin-independent WIMP-nucleon cross-section as a function of WIMP mass. The solid lines represent the combined limits from CDMS Soudan data. Other curves give recent results from other leading experiments (references in text). The dotted red curve indicates the effect of the newest scintillation yield measurements upon the XENON10 limit. Filled regions represent parameter ranges expected from the DAMA/LIBRA result and from a broad range of MSSM and CMSSM models.

7.4.2 Spin-dependent interactions

The two panels of Figure 7.7 interpret the same CDMS data in terms of spin-dependent interactions, based upon scattering with the relatively rare isotopes ^{29}Si and ^{73}Ge. Following the convention established by Tovey and coauthors [146], I illustrate limits assuming coupling to neutrons only ($a_p = 0$, *left*) and to protons only ($a_n = 0$, *right*). These figures also show the region of CMSSM parameter space for spin-dependent interactions computed by Roszkowski *et al.* [157] and limits from recent leading experiments. The CDMS limit is weaker than that from the XENON10 experiment [159] in both cases, due to the greater abundance of odd-nucleon isotopes (^{129}Xe and ^{131}Xe) in Xe. CDMS and XENON10 provide the leading constraints on WIMP-neutron scattering, though they have not yet reached the bounds of CMSSM parameter space. Limits on WIMP-proton scattering are dominated by dedicated experiments using odd-proton nuclides, notably COUPP [224] (CF_3I) and KIMS [225] (CsI), as well as limits from an indirect search for neutrino annihilation in the sun using Super-Kamiokande [125].

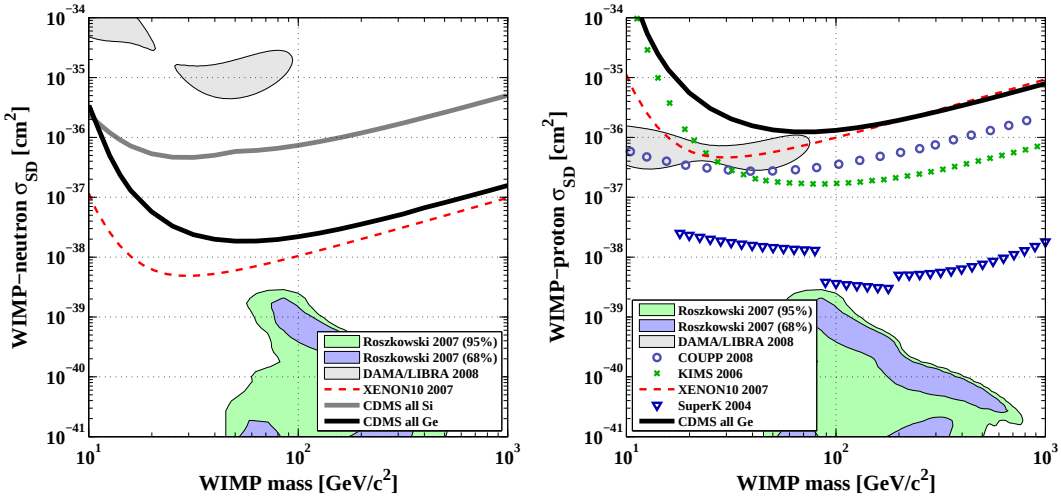


Figure 7.7: Upper limits (90% C.L.) on the spin-dependent WIMP-neutron (*left*) and WIMP-proton (*right*) cross-sections as functions of WIMP mass. The solid lines represent the combined limits from CDMS Soudan data. Other curves give recent results from other leading experiments (references in text). The dotted red curve indicates the effect of the newest scintillation yield measurements upon the XENON10 limit. Filled regions represent parameter ranges expected from the DAMA/LIBRA result and from CMSSM models.

It is important to note that the two cases plotted in Figure 7.7 are not especially

realistic. Most WIMP models predict spin-dependent cross sections upon protons and neutrons to be of similar magnitude, *i.e.* $|a_p| \sim |a_n|$. The relative strengths of the various experimental constraints vary substantially with a_p/a_n, and in most realistic cases the XENON10 and Super-Kamiokande data provide comparable constraints. Figure 7.8 gives a different view upon this parameter space, plotting a_p versus a_n for WIMP mass $M_\chi = 60$ GeV/c^2. Each experimental limit appears as an allowed ellipse in this plane, with an orientation and aspect ratio determined by the spin structure functions of the target isotopes. The value of experiments with different target isotopes is clear: by combining two such ellipses at different orientations, we greatly limit the allowed region of parameter space. Such combinations also limit our sensitivity to the poorly-modeled cancellations which occur along the long axes of these ellipses. The gray filled regions in this plot indicate the 3σ best-fit regions consistent with the DAMA/LIBRA result at this mass, as interpreted by Savage and coauthors [223]. Note that the combination of current constraints from direct and indirect detection exclude a spin-dependent interpretation of the DAMA/LIBRA signal for this WIMP mass (and, indeed, for all $M_\chi \gtrsim 60$ GeV/c^2).

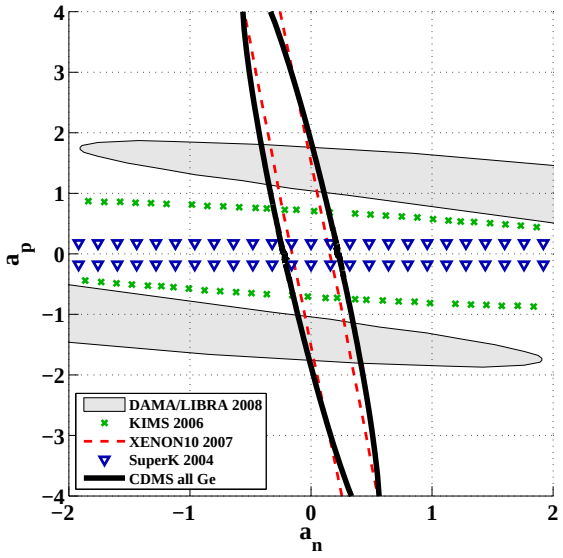


Figure 7.8: Recent experimental results regarding spin-dependent WIMP-nucleon interactions, plotted in the $a_p - a_n$ plane for a WIMP of mass $M_\chi = 60$ GeV/c^2. All curves are 90% confidence upper limits, *i.e.* the regions outside the relevant ellipses are excluded. The filled regions are consistent with the DAMA/LIBRA result at 3σ, as interpreted by Savage and coauthors. No region of parameter space is consistent with DAMA/LIBRA and with all experimental limits at this WIMP mass.

7.4.3 Additional comments

Low WIMP masses

In this analysis, we have only considered WIMP masses $M_\chi \gtrsim 10$ GeV/c^2. This is a consequence of our choice of detector threshold: in order to produce a 10-keV Ge recoil, a 10-GeV/c^2 WIMP must be incident at $v \gtrsim 630$ km/s – far out in the tail of the standard halo model's velocity distribution. The situation is slightly better for Si: a 10-GeV/c^2 WIMP need only be incident at $v \gtrsim 400$ km/s to produce a 7-keV Si recoil. In genereal, however, our limits at low WIMP masses thus depend strongly upon the poorly-understood tails of the halo velocity distribution in the solar neighborhood, and so suffer from substantial systematic errors. For our current choices of target nuclide and threshold, we thus avoid detailed conclusions about these low WIMP masses.

It is important to note, however, that our current choices of energy threshold are very conservative. Most of our detectors have recoil energy resolutions of < 1 keV r.m.s. and ionization or phonon resolutions of < 300 keV r.m.s., so ionization yield discrimination often remains useful well below 5 keV. The primary limiting factor is our desire for a zero-background analysis: the phonon timing parameters have poor signal-to-noise at these low recoil energies, so we cannot reject our beta background while maintaining significant sensitivity to nuclear recoils. If we relax this requirement, the we can use lower energy thresholds and set much stronger limits on low-mass WIMPs. Such an analysis can be carried out with or without background-subtraction, depending upon how much work we wish to put into our background model. Ray Bunker has carried out such an analysis for the Run 21 data set at Stanford Underground Facility (paper in preparation), while other collaborators are currently working on a similar analysis of the Soudan data set.

Model dependence

To conclude this Chapter, it's useful to briefly review the model assumptions that go into the interpretation of results from this and similar dark matter searches. These have become particularly interesting in light of the recent reiteration of the DAMA collaboration's claim of an annual modulation signal.

1. **WIMP halos:** The rate of WIMP recoils is affected by the local WIMP density and velocity distribution. The local WIMP density affects all direct-detection experiments

equally, but it does alter comparisons between experiment and theory. The effects of changes in the velocity distribution are more complex, and can alter the relative reaches of different experiments.

2. **Dark matter interactions:** These results assume that dark matter particles scatter elastically from atomic nuclei. If dark matter consists of particles which interact primarily electromagnetically (*e.g.* axions) or through inelastic scattering, it may avoid detection in this class of experiment. This does not invalidate any of the limits described above, but it may affect the interpretation of detected signals.

3. **Nuclear physics:** Uncertainties in nuclear structure have substantial effects upon the WIMP interaction rate [145]. Changes in the nucleus's strangeness content or spin structure affect the comparison between theory and experiment, but generally affect all experiments equally.

4. **Detector physics:** The interpretation of these results could also be affected by unusual detector processes not accounted for in our analysis. A prime example is ion channeling [226], which may alter the ionization and scintillation yields of nuclear recoils in crystal detectors [227].

Chapter 8

Conclusions and future directions

8.1 Status and future of CDMS II

The CDMS II experiment currently sets world-leading limits on WIMP-nucleon interactions, based upon the two data runs described in this dissertation. CDMS II continues to collect data, however, and will continue to increase its reach. At the time of this writing, CDMS is undergoing cryogenic maintenance after its sixth successful data run at Soudan, Run 128. Figure 8.1 illustrates the progression of total live time thus far. The combined data set of CDMS II currently exceeds the exposure used in this dissertation by a factor of three. By its conclusion in early 2009, CDMS II is expected to reach a sensitivity of $\sim 2 \times 10^{-44}$ cm^2 (20 zeptobarns) for WIMP masses near 60 GeV/c^2.

8.2 Direct detection: a status report

Recent years have seen enormous advances in WIMP direct detection. A field that was once dominated by cryogenic detectors now hosts several competitive technologies. To help put CDMS's result in context, I give a brief rundown below of some recent highlights of the field.

8.2.1 Solid-state detectors

Solid-state detectors remain leading players in the field due to their excellent energy resolution and powerful discrimination against electron recoils. CDMS is currently the leader in this category, but at least three other techniques bear special mention.

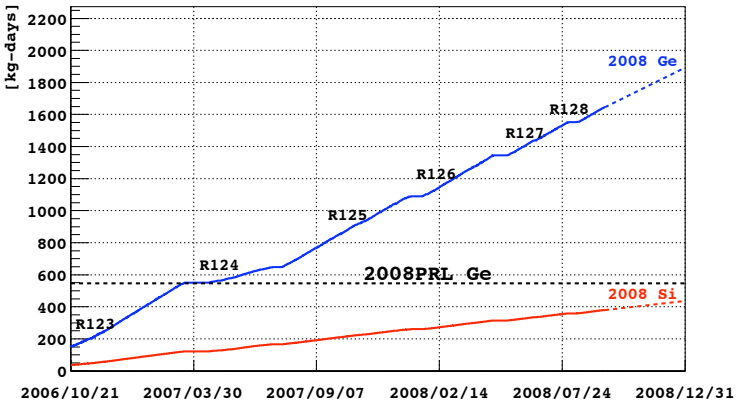


Figure 8.1: Exposure of the CDMS experiment (before cuts) as a function of calendar time. Each data run is labeled, interspersed with periods of cryogenic maintenance. The horizontal dotted line indicates the Ge exposure from this analysis before cuts, after removing unused detectors.

EDELWEISS

The EDELWEISS collaboration employs a technology very similar to that of CDMS: ionization and phonon detection with Ge detectors. Rather than using TESs to detect athermal phonons far from equilibrium, however, EDELWEISS has focused on the use of NTD Ge thermistors to measure the near-equilibrium temperate changes of the entire crystal substrate. The primary discrimination variable is ionization yield, just as it is for CDMS. Without the extra discrimination power provided by phonon timing, however, surface event backgrounds have restricted this approach to limits at the $\sim 10^{-42}$-cm^2 level [216]. A new cryostat is currently operating for the larger EDELWEISS II experiment.

The EDELWEISS collaboration is also heavily engaged in detector development work for large-scale Ge detectors. Extending a design developed by our collaboration in 2005 [228], EDELWEISS has recently focused its efforts upon an interleaved charge electrode design [229]. In this scheme, each detector face is patterned with two interleaved ionization grids at different voltage biases. The CDMS test device was a 1-cm Si ZIP patterned with a meandering ground electrode and a single biased electrode (+3V on one face, -3V on the other) connected to a charge amplifier on each face. Charge carriers within the crystal bulk drift according to the average field (3V/cm), producing equal and opposite pulses on the two readout channels. Events near the surface, in contrast, are collected between the ground

and biased electrodes on that face, produce an asymmetric signal. In this way the ionization channels alone can be used to identify surface events with great accuracy. This design was poorly matched to the TES electronics used in our athermal phonon readout, but recent work [230] has shown that it gives excellent performance when coupled to EDELWEISS's thermal phonon readout. Since such detectors have excellent background rejection power and can be mass-produced with relative ease, they are the now dominant technology for constructing a ton-scale EDELWEISS installation.

CRESST

The CRESST collaboration also detects the thermal phonon signal from a particle impact in a cryogenic substrate, but matches it with a measurement of scintillation light rather than charge. A CRESST detector is a $CaWO_4$ crystal instrumented with two TES thermometers: one to measure the crystal temperature, the other to measure the temperature of an associated light-absorbing Si wafer. CRESST has set limits on spin-independent WIMP-nucleon interactions at the 5×10^{-43}-cm^2 level [231], and is also proceeding with a larger experimental installation (CRESST II).

CRESST and EDELWEISS have joined the EURECA project, a pan-European effort to develop a next generation cryogenic WIMP search experiment.

Low-threshold Ge

The recently-renewed claim of an annual modulation signal by the DAMA collaboration (Section 8.2.4) has led to new interest in WIMPs of very low mass ($0.1 < M_\chi < 10$ GeV/c^2). This region of parameter space represents the only remaining small window of consistency between the DAMA result and the null results of other experiments within the standard WIMP framework. Such WIMPs would produce recoils with energies well below 10 keV, where the background-discrimination power of many technologies begins to break down.

Juan Collar and collaborators have recently proposed the use of ultra-low-noise germanium detectors for dark matter searches [233]. These authors use a low-capacitance, point-cointact charge electrode configuration to achieve a zero-energy ionization resolution of 70 eV r.m.s., 1/3 the level achieved by the higher-capacitance CDMS detectors. These detectors measure ionization only, and so cannot distinguish between different event types.

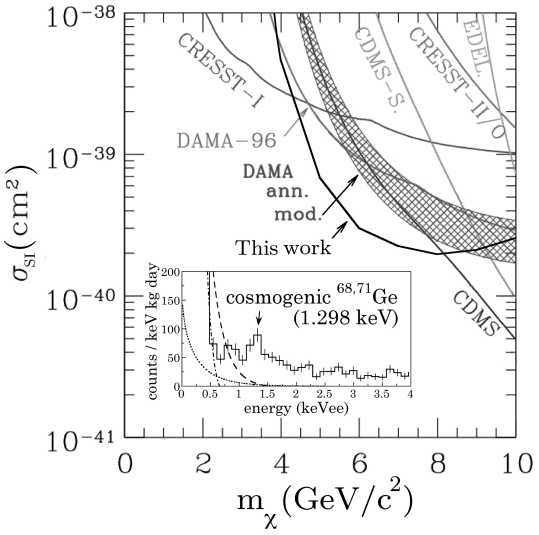


Figure 8.2: Recent limits on spin-independent WIMP-nucleon interactions at low WIMP masses, taken from [232]. "This work" refers to the first limits from CoGeNT from that paper. The shaded region represents a low-mass interpretation of the DAMA/NaI annual modulation signal. The inset shows the observed event spectrum at low energies, overlaid with the spectra of some reference WIMP candidates described in the original reference.

The first run of the CoGeNT experiment acquired an 8.4-kg-day exposure with such a detector in a shielded environment within the Chicago water system (330 m.w.e.). The observed event rate sets strong limits upon the interactions of low-mass WIMPs, sufficient to exclude one of the last regions of WIMP parameter space consistent with the DAMA result (excluding possible considerations of ion channeling). CoGeNT plans to scale this experiment to larger exposures in a new cryostat in the near future to increase its sensitivity.

Other collaborations are also pursuing similar techniques. The TEXONO collaboration has also reported limits with 0.34 kg-days of exposure with Ge detectors with a threshold of 220 eV [234], though some authors have raised concerns about the validity of this analysis [235]. The CDMS collaboration has also conducted a low-threshold analysis of its data runs at Stanford, with a publication in preparation.

8.2.2 Noble liquids

The past few years have seen dark matter detectors based on the scintillation of liquefied noble gases leap to the cutting edge of the field. Such instruments have already set competitive limits on WIMP dark matter, and they promise easy scalability to large

detector masses.

Basic principles

Liquid noble detectors are large tanks of liquefied neon, argon, or xenon observed by photomultiplier tubes (PMTs). Table 8.1 shows some of the properties of noble liquids relevant to direct detection. Noble liquids are excellent scintillators, with light yields approaching that of NaI(Tl) and long ($>$ 1m) attenuation lengths for that light. All except xenon scintillate in the vacuum ultraviolet, necessitating the use of fluorescent wavelength-shifting absorbers alongside the PMTs. Their relatively high boiling points make the cryogenic requirements of a noble liquid detector far simpler than those of an experiment like CDMS. Their chemical inactivity allows them to be purified to very high levels, though the Rn continuously released by U/Th contamination in the PMTs remains a challenge. Most importantly, noble liquid detectors hold the promise of large target masses: for a bigger detector, just build a bigger tank.

Table 8.1: Some properties of noble liquids relevant to WIMP detection. Numbers taken from [236, 237].

	A	Density [g/cm^3]	Boiling Point [K]	Yield [ph/keV]	τ_1/τ_3 [ns]	λ [nm]	Long-Lived Radioisotopes
Ne	20	1.2	27.1	15	2.2 / 2900	77	*none*
Ar	40	1.4	87.3	40	6 / 1590	128	^{39}Ar ($\sim$ 1 Bq/kg)
Kr	84	2.4	120			147	^{85}Kr ($\sim 10^5$ Bq/kg)
Xe	131	3.0	165	42	2.2 / 21	175	*none*

Nuclear and electron recoils in a noble liquid detector can be distinguished in two distinct ways:

1. Nuclear and electron recoils produce different ionization and scintillation yields in a noble liquid. The primary scintillation signal may be measured with phototubes. The charge carriers can be drifted using a high-voltage electrical grid and extracted into a gas volume, where they cause secondary proportional scintillation. A "dual-phase" (liquid-gas) detector of this type thus observed two scintillation pulses for each particle event. The amplitudes of the two light signals identify the energy and recoil type of each event, while the time delay between them identifies the event's position along the drift axis.

2. The pulse shape of the primary scintillation signal is also a useful discriminant between

electron and nuclear recoils. Each noble liquid has two scintillation time constants (τ_1 and τ_3 in Table 8.1), corresponding to the de-excitation rates of singlet and triplet excimer states within the liquid. Nuclear recoils populate the singlet state preferentially, giving more prompt light in the primary scintillation signal.

A single detector may use both of these discrimination techniques, but there are design tensions between the two: pulse shape discrimination favors a spherical geometry for optimal light collection, while charge extraction favors a cylindrical configuration for uniform charge drift.

Large WIMP detectors have been proposed based upon neon, argon, and xenon. I touch briefly on the progress of each of these materials below, from heaviest to lightest.

Xenon

Liquid xenon is the most obviously promising of the noble liquid targets. It has the highest boiling point, the largest light yield, no long-lived radioisotopes, and scintillation light which can be detected using ordinary PMTs without wavelength-shifting. It's large atomic mass gives it a large cross section for spin-independent interactions (though form factor effects can counteract this at high energy thresholds). Its high density allows for compact detectors and makes for very effective self-shielding: most background events cannot penetrate more than a few cm into the detector volume. Xenon is not well-suited to pulse shape discrimination, however, due to its extremely short scintillation times ($\tau_1 =$ 2.2 ns, $\tau_3 = 21$ ns).

The XENON10 collaboration announced its first WIMP-search results in April 2007 [158]. Using a 15-kg dual-phase detector located at Gran Sasso, the collaboration observed 10 candidate events (consistent with expected backgrounds) in a 136 kg-day exposure. These data set then-world-leading limits on WIMP-nucleon spin-independent couplings [158] and spin-dependent interactions with neutrons [159]. These limits have since been revised slightly upwards due to improved measurements of the scintillation efficiency of nuclear recoils in xenon [220, 219].

Several experiments are pushing xenon to yet higher target masses. The 100 kg XENON100 detector is currently being commissioned at Gran Sasso, scaling up the same basic design as XENON10. The 300 kg LUX experiment adds an active water shield to aid in neutron rejection; this instrument is under construction for deployment at the Sanford Un-

derground Science and Engineering Lab (SUSEL) in South Dakota. In Japan, the XMASS collaboration is developing an 800 kg single-phase detector, focusing on self-shielding and good position reconstruction to eliminate backgrounds. Each of these instruments has an expected sensitivity at or below 1 zeptobarn (10^{-45} cm^2), and designs for yet larger detectors are well underway.

Argon

Liquid argon combines the excellent light yield of xenon with extremely powerful pulse shape discrimination, made possible by argon's enormous difference in singlet and triplet decay rates [238]. Argon's lower atomic mass also demands larger detector masses than for an equivalent xenon detector, but it is also much cheaper than xenon or neon. Argon's main drawback is a large intrinsic trigger rate from radioactive ^{39}Ar, demanding enormous discrimination power ($\sim 10^8 : 1$) in a large-scale detector. Argon (like neon) also scintillates in the vacuum ultraviolet (VUV) band, giving its scintillation light a very short attenuation length in common gases. PMT readout of an argon target thus requires the use of a wavelength-shifting fluorescent material to convert this light to longer wavelengths, which unfortunately reduces light yield and complicates the propagation of light within the detector.

The WARP collaboration has recently published limits with a 2.3-L two-phase prototype detector at Gran Sasso [217], using both ionization and pulse-shape discrimination to set the (non-blind) upper limit shown in Figure 7.6. The WARP and ArDM collaborations are pursuing similar two-phase technology on a larger scale. In parallel, the DEAP collaboration is pursing a single-phase detector based solely on pulse-shape discrimination; a 7 kg prototype is currently taking data. Finally, recent discoveries of underground argon deposits low in ^{39}Ar [239] may open the door to multi-ton liquid argon detectors in the near future.

Neon

Neon retains (and improves upon) the excellent pulse shape discrimination of liquid argon and has no long-lived radioisotopes. Its low atomic mass demands multi-ton detector masses, however, and its cryogenics are significantly more demanding than those for liquid argon. McKinsey and coauthors have proposed a large-scale, single-phase neon experiment

for the detection of dark matter and solar neutrinos [240]. By switching argon for neon in the same detector, it may also be possible to test the A^2-dependence of a WIMP signal with excellent control of systematics.

8.2.3 Phase-transition detectors

CDMS's transition-edge sensors are an example of a very general rule: a system balanced at an abrupt transition makes an effective amplifier (or detector). Physical phase transitions, notably evaporation, can also form the basis of a detection technology. If a liquid is superheated to a metastable state above its boiling point, the energy of a particle impact can cause it to boil explosively. Boiling will only occur if enough energy is deposited in a small enough volume to nucleate a bubble, i.e. the particle track must have dE/dx greater than some threshold value. By tuning this threshold, a phase transition detector can be rendered insensitive to the sparse tracks of muons or electromagnetic backgrounds, while still remaining responsive to nuclear recoils. Figure 8.3 shows representative photographs of particle events from the COUPP experiment's CF_3I target, illustrating a cosmic ray event (only visible when the detector is tuned to a low dE/dx threshold), a neutron multiple-scatter event, and a neutron single-scatter event.

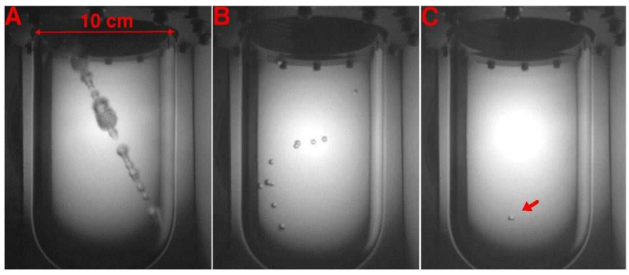


Figure 8.3: Three images of particle events in the COUPP 1.5-kg CF_3I bubble chamber: (A) cosmic ray event, visible in a run tuned to a very low threshold in dE/dx, (B) a neutron multiple-scatter event, and (C) a neutron single-scatter event. Plot taken from [224].

Phase transition detectors take a fundamentally different approach to event energy reconstruction than that of the other technologies I have discussed. It is very difficult to determine the energy deposited in the creation of any individual bubble. Bubble formation is thus an essentially binary detection technology: either enough energy was deposited to nucleate a bubble, or it wasn't. A bubble chamber or similar detector only measures the count rate above its energy threshold, and can only resolve energy information by acquiring

data at a range of different energy thresholds. A phase transition detector must thus acquire several long data runs at different detector tunings in order to distinguish WIMPs from a higher-energy background spectrum (*e.g.* that from α-decay) or to even partially resolve the WIMP recoil spectrum. Broadly speaking, phase transition detectors have excellent background rejection but produce little information about any observed events. The latter impedes the discovery potential of a future experiment, since such devices will have a difficult time characterizing the backgrounds (or signals) they observe.

Most phase transition detectors are relatively low-mass experiments with flourine-containing fluids, specialized to search for spin-dependent WIMP interactions. PICASSO and SIMPLE are two examples: both detectors are arrays of superheated droplets, and both have set leading limits on spin-dependent WIMP-nucleon interactions. The COUPP collaboration has recently expanded phase-transition technology to much larger masses by developing a bubble chamber with a long hold time and no sensitivity to electromagnetic backgrounds. COUPP has already set competitive limits on spin-dependent interactions using a 1.5-kg target of superheated CF_3I [224], and larger-mass detectors are under production. A ton-scale bubble chamber employing similar techniques could be a major competitor in search for spin-dependent and spin-independent WIMP interactions.

8.2.4 DAMA/LIBRA

The DAMA collaboration has operated low-background arrays of NaI(Tl) scintillator crystals at Gran Sasso National Laboratory in Italy since 1996. Particle events within these crystals are identified by their scintillation light alone. The scintillation yield of a nuclear recoil is reduced compared to that of an electron recoil of similar energy by a "quenching factor" Q: $Q_{Na} = 0.30$ for Na recoils, $Q_I = 0.09$ for I recoils [241]. Without an independent measurement with particle energy, however, this difference cannot be used for background rejection. The DAMA detectors are thus simple counters, with no event-by-event discrimination other than a multiplicity cut (a candidate event may not deposit energy in more than one detector module). The collaboration seeks to identify WIMPs on a statistical basis instead, using the expected annual modulation of the WIMP recoil spectrum due to the earth's motion around the sun. The rate of low-energy single-scatter counts in the DAMA detectors should vary during the year if a WIMP signal is present, while the rates of multiple-scatter events and higher-energy counts should remain constant.

Note that the identification of a few-percent annual modulation signature demands the detection of many thousands of WIMP recoils in a large exposure, but the relatively simplicity of NaI detectors has allowed the DAMA collaboration to acquire far larger exposures ($> 10^5$ kg-days) than any other experiment.

DAMA has fielded two large detector arrays at Italy's Gran Sasso National Laboratory. The 115-kg (87.3-kg fiducial) DAMA/NaI detector array initially reported evidence in 1997 for an annual modulation in its count rate in the 2-4 keV-electron-equivalent (keVee) energy range [242]. This corresponds to Na nucleus recoil energies of $7 - 13$ keV or I recoil energies of $22 - 44$ keV. The DAMA/NaI experiment continued to observe this signal over 7 annual cycles (0.29 ton-years) with this apparatus. This signal has since been reinforced by more than 3 years of data from the 250-kg DAMA/LIBRA installation [222], for a combined exposure of 0.87 ton-years. Figure 8.4 shows the combined modulation signal and single-scatter energy spectrum (in keV-electron-equivalent). The modulation amplitude is consistent with zero for multiple-scatter events and for single-scatter events above ~ 5 keVee, as expected from a WIMP signal.

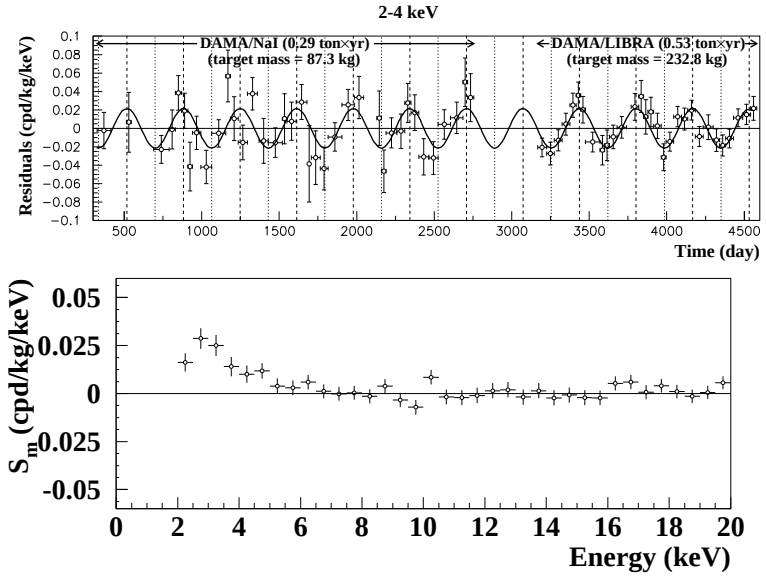


Figure 8.4: *Top*: Annual modulation (counts/day/kg/keVee) in the rate of 2-4 keVee single-scatter events observed by the DAMA/NaI and DAMA/LIBRA experiments. *Bottom*: Modulation amplitude observed by DAMA/LIBRA as a function of visible energy (keVee). Figures from [222].

The DAMA annual modulation signal has been a source of controversy in the com-

munity for more than decade. DAMA's observations satisfy the basic criteria for a WIMP signal, but have thus far not been confirmed by any other experiment. The collaboration initially interpreted its data within the standard halo model in terms of a ~ 60 GeV/c^2 WIMP with a cross section of $\sigma_{SI} \sim 10^{41-42}$ cm^2, but this model was quickly excluded by null results from other experiments: this model should have produced ~ 100 candidate events in the CDMS data run described in this thesis, for example, but none were observed. The vast majority of parameter space consistent with the DAMA/LIBRA result within the standard halo model is excluded by the null results of other experiments, though some authors [223] find a small amount of consistent parameter space at very low WIMP masses ($M_\chi < 10$ GeV/c^2). No direct cross-check of the annual modulation signature is possible, however, since no experiment of sufficient scale has yet been undertaken; the KIMS collaboration proposes to do so in the near future with 250-kg of CsI(Tl) scintillators [243].

If the DAMA signal has a dark matter origin, we need to explain why no signal has been observed in other experiments. Very low WIMP masses might explain the discrepancy, as might unusual halo models with enhanced modulation amplitudes. A non-WIMP dark matter candidate could also explain the difference if it deposits electromagnetic energy within a detector. Axion-like particles can deposit their energy electromagnetically, as can the inelastic scattering of a particle with very low-lying (< 100 keV) excited states. Such dark matter events would fail the nuclear recoil selection cuts of CDMS and most other experiments due to their elevated ionization and scintillation yields, but their modulation would still be visible in the non-discriminatory DAMA search. Each of these possibilities will be tested by CDMS in the near future: a low-threshold analysis can explore the parameter space of low-mass WIMPs and many unusual halo models, while the observed rate of low-energy electromagnetic events can set competitive limits upon axion-like particles and inelastic scattering.

We may also consider unusual detector physics which renders NaI more sensitive to WIMP interaction than Ge or other targets. Following a suggestion by Drobyshevskii [227], the DAMA authors [244] have recently argued that ion channeling effects may explain the discrepancy with other experiments. These authors argue that a fraction of recoiling nuclei in a crystal detector will propagate along "easy" directions in the anisotropic crystal lattice, causing them to deposit less energy in the crystal's nuclear system and appear at elevated ionization and scintillation yields. Such channeled recoils would be rejected by the discrimination cuts of CDMS, lowering our signal efficiencies slightly. DAMA, in contrast,

would observe them at higher scintillation energies, lowering its effective energy threshold. This phenomenon may open up new regions of consistent parameter space at low masses. The channeling model used in [244] is optimistic and not experimentally verified, however; further study of ion channeling (and blocking) effects is needed to evaluate their effects.

Finally, we must consider the possibility that the DAMA signal has a more mundane origin, unconnected with dark matter. An annually-modulated systematic effect on the electronics could duplicate the signal, but the authors have set stringent limits on many of the more obvious possibilities. The underground muon flux is known to be modulated over the year with roughly the correct phase [245], but the expected rate is quite low and it is difficult to understand how no modulation would appear in the multiple-scatter events. Interestingly, the modulation spectrum in Figure 8.4 may be interpreted as a peak near 3 keV, very close to the substantial 3.2-keV X-ray peak from ^{40}K decay. Any modulation in cut acceptance could lead to the observed signal, but no detailed mechanism has yet been identified. The origins of the DAMA effect – whether related to dark matter or otherwise – are still unknown.

8.2.5 Directional detection

A dark matter signal may also be identified by the directions of the resulting recoils, which should show a diurnal modulation with the earth's rotation. This modulation would be strong statistical evidence of a signal's extraterrestrial origin, but the short track length of a recoiling nucleus makes this extraordinarily difficult to detect. The DRIFT collaboration has pursued directional detection with low-pressure gas TPCs for several years [246]. DRIFT II has demonstrated significant head-teal discrimination of recoil direction in 40 Torr CS_2 gas with a density of 1.67×10^{-4} g/cm^3 [247]. The newer DMTPC collaboration [248] is actively developing a similar CF_4 detector at 50 Torr read out using an optical CCD, with plans to scale to a 1 m^3 detector and beyond. These technologies are very promising for verifying and characterizing an eventual WIMP signal, but their low density makes them unlikely discovery instruments: at 50 Torr, a 10-kg DMTPC-style detector would occupy 40 m^3.

8.3 Challenges for CDMS technology

As evinced by the result described above, CDMS has achieved great success in maintaining background-free operation over lengthy WIMP-search exposures. Following up on this success, it is natural to consider extensions of this technology to even larger masses and greater WIMP sensitivities. CDMS also does not operate in a vacuum; as described in the last Section, several technologies are vying for the lead at the zeptobarn scale and beyond. With this in mind, it is useful to consider some of the challenges facing CDMS technology in the race to scale to larger exposures. This section describes some of the most significant issues facing a larger version of CDMS.

8.3.1 Detector production

The fabrication and commissioning of ZIP detectors has thus far been a labor-intensive business. The photolithography of each detector at Soudan was performed using a "step-and-repeat" process, requiring 444 exposures to complete the patterning of each face. The various phonon sensors were then manually wired together using a lengthy series of wirebonds. Finally, each ZIP detector required at least three dilution refrigerator runs before departing for the mine. Most experienced short-circuits or other hardware failures along the way that required manual repair, while others needed to be repolished entirely. Even after this process, a few detectors appear to have regions of disconnected TES's, notably T3Z2 and T5Z1. The throughput (labor-per-kg) and must improve substantially before a larger experiment is possible.

CDMS also has room to improve the consistency of its fabrication process. Each CDMS ZIP is still somewhat "hand-crafted", and there are large variations in performance within the 5-Towers. The reduced phonon sensitivities of Tower 4 and 5 Ge detectors are the most prominent example of such variation, but there are also large variations in the position dependencies and timing performance of the phonon signals. This variability greatly increases the labor involved in calibration and data analysis. A larger-scale experiment with more detectors will demand greater consistency within its detector population.

8.3.2 Position reconstruction

Despite the successes of the lookup-table correction, our ability to reconstruct the position of an event within a ZIP detector remains limited. The look-up table is often

degenerate in the outer regions of the detector, leading to occasional mis-corrections of event parameters. This appears to be a fundamental characteristic of the 4-quadrant sensor configuration, and one that would become worse at larger diameters or thicknesses. This analysis already addressed such mis-corrected events with a specialized cut, and they could become a limiting background for larger exposures.

The existing position reconstruction algorithm is also unable to unambiguously identify internal multiple-scatters: events which deposit energy in two or more displaced vertices within a single detector. This class of event includes true multiple-scatters of a single particle (*e.g.* a neutron) and close coincidences (*e.g.* a neutron and a photon from the same cosmic ray shower). Some internal multiples fail selection cuts, but others can be mis-corrected into the signal region. This is particularly problematic for neutron internal multiples: internal multiple-scatters constitute a systematic difference between our calibration sample and our signal distribution, which contains only single-scatter events. The rate of such events grows in proportion to the mass of each detector, and each becomes more significant as we demand lower leakage rates into our signal region.

8.3.3 Surface events

The analysis described in this dissertation achieves $\gtrsim 30\%$ overall signal efficiency while maintaining $\sim$ 200:1 phonon-timing rejection of single-scatter surface events in interior detectors which pass all other cuts (0.25 expected candidates out of 48 single-scatter events in the nuclear recoil band). Without improved rejection techniques, larger exposures will demand lower signal efficiencies to maintain adequate background rejection. A large-scale CDMS experiment thus demands improvements in rejection power and/or reductions in surface event contamination in order to operate in the "zero-background" regime.

Endcap detectors – those in the Z1 and Z6 positions, at the top and bottom of a detector stack – are particularly troublesome in this regard. The close-packed detector arrangement of CDMS stacks is a major component of our rejection power against surface contamination ($\sim$ 86% of particles leaving the surface of an interior detector within the confines of the inner charge electrode will strike a neighboring ZIP). The background rejection of endcap detectors is undermined by an excess of single-scatter events and our inability to characterize their response to surface events on their exterior face. In the current configuration, we can expect endcap detectors to become less useful contributors to future

WIMP searches. The three endcap Ge ZIPs in this run (T3Z6, T4Z6, and T5Z1) already contributed half of the expected background to this WIMP-search.

Efforts to estimate the background level of the CDMS experiment are currently limited by the limited statistics of surface events in the WIMP-search data outside of the signal region. This could be alleviated somewhat if we could base our background estimate upon the more plentiful surface events from the ^{133}Ba calibration data. The systematic differences in energy and incidence angle between the calibration and WIMP-search populations currently make this dangerous, but a better model of the phonon timing in our detectors would allow us to correct for these differences and make full use of our calibration data.

8.4 Beyond CDMS II

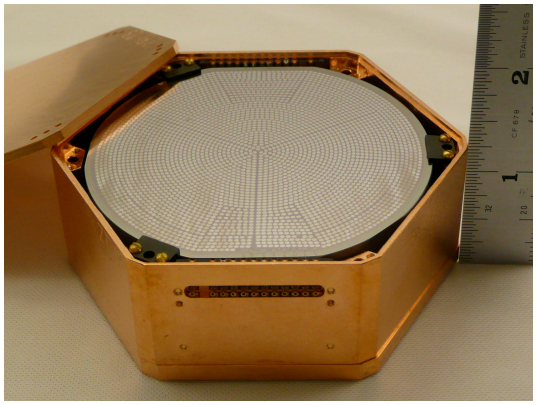
The CDMS collaboration is currently engaged in an ambitious program of research and development to address the challenges described above and proceed to larger detector masses. Proposals have been submitted for SuperCDMS [249], an extension of low-temperature WIMP detector technology to much larger masses. The eventual goal of this program is a ton-scale cryogenic experiment, competing with and complementary to the large noble-liquid instruments currently under discussion. A full discussion of these development efforts is far outside the scope of this dissertation, but I highlight a few recent developments below.

8.4.1 New ZIP design

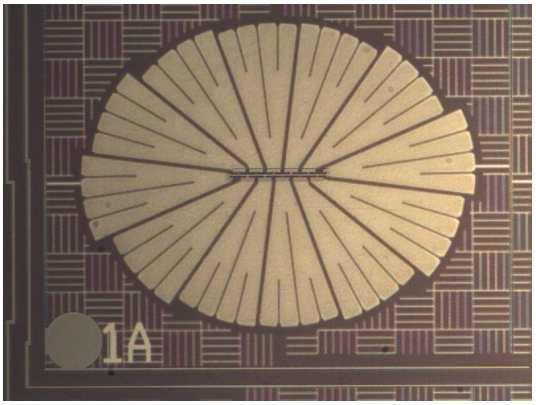
The CDMS collaboration has developed a next-generation ZIP detector design, currently under test at Berkeley and Case Western. Figure 8.5(a) shows one of the early prototypes of this configuration. These detectors incorporate several improvements in design, most of which trace back to two major fabrication upgrades:

- New ZIPs are fabricated on three-inch-diameter, one-inch-thick (620g) Ge substrates, 2.5× thicker than previous detectors.
- ZIPs are now fabricated using a whole-field mask process. This requires only two photolithographic exposures per detector, rather than the 444 required by the old "step-and-repeat" process.

Both of these changes have dramatically improved fabrication throughput and reliability. Larger substrates allow us to produce more target mass with each fabrication run, while the whole-field mask process greatly reduces the labor and complexity of each run. The whole-field process also allows much greater flexibility in detector patterning, since there is no need to divide the design into arrays of identical tiles. New ZIPs require far less wire-bonding than previous designs, for example, since all inter- and intra-channel wiring are now included in the photolithographic mask.



(a) A prototype SuperCDMS ZIP within its housing. The phonon face is visible, showing the tripartite sensor design.



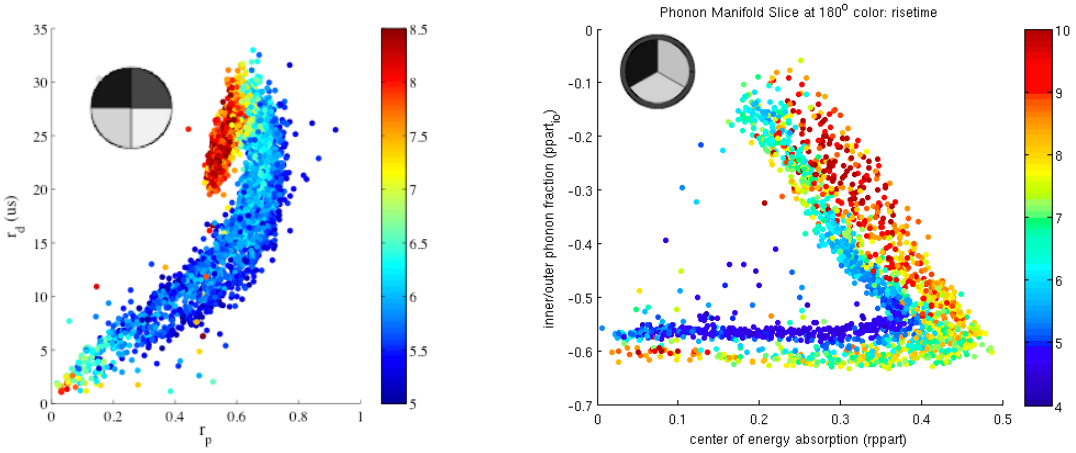
(b) Photomicrograph of a portion of the new CDMS phonon design, showing the "stadium" arrangement of slotted Al fins around a TES, as well as the surrounding charge collection grid.

Figure 8.5: Photographs of the current SuperCDMS ZIP design.

These new ZIPs are designed to have superior event reconstruction to that of current detectors. Rather than the four-quadrant phonon sensor design used in CDMS II, SuperCDMS ZIPs have three large inner sensors surrounded by an outer phonon guard ring. Monte Carlo simulations of phonon propagation within Ge showed this sensors design to have superior reconstruction of the radial positions of events, thus reducing the possibility of mis-corrections. This prediction has been borne out by initial tests: Figure 8.6 shows that the phonon parameters used to define event radius are less degenerate for new ZIPs than for old. The arrangement of aluminum collector fins around each TES has also been optimized, leading to an "stadium" arrangement (Figure 8.5(b)) that puts more aluminum within quasiparticle-diffusion range of the TES. The width of the outer charge electrode has also been increased to help reject events near the outer detector rim, where phonon

reconstruction can be poor.

SuperCDMS ZIPs are further expected to have significantly lower background rates than current detectors. The increased substrate thickness is already a significant advantage: a new ZIP has 2.5× more detector mass for the same inner-electrode surface area, thus reducing the rate of surface events at a given WIMP sensitivity by that same factor. New detectors are also expected to have a lower intrinsic beta rate due to improved handling to reduce radon exposure. Finally, each SuperCDMS Tower will only contain five one-inch ZIPs, surrounded by ionization-only veto detectors (1cm Ge) at the top and bottom of the stack. This will improve self-shielding and the power of the multiplicity cut, allowing all detectors to avoid the difficulties faced with endcaps in this run.



(a) A thin slice of the position reconstruction manifold from a CDMS II ZIP. The inset illustrates the arrangement of the four phonon sensors.

(b) A similar manifold slice for a SuperCDMS ZIP. Note the reduction in fold-back in the manifold and the smoother spatial variation in risetime. The tripartite sensor arrangement – three inner phonon sensors and one outer ring – is shown in the inset.

Figure 8.6: Comparison between the position reconstruction manifolds of CDMS II and SuperCDMS ZIPs. Each shows gamma events from a thin slice of a detector, plotting radial parameters reconstructed from phonon relative amplitudes (x-axis) and arrival times (y-axis) against one another. The detector center lies in the lower left corner of each pot. Color indicates primary phonon risetime, a surface event discrimination parameter.

As part of the SuperCDMS project (if it is approved), four Towers of these new detectors will be installed at the Soudan site alongside the best of the existing CDMS II detectors. This experiment will total ~ 15 kg of Ge target mass, with an expected reach of

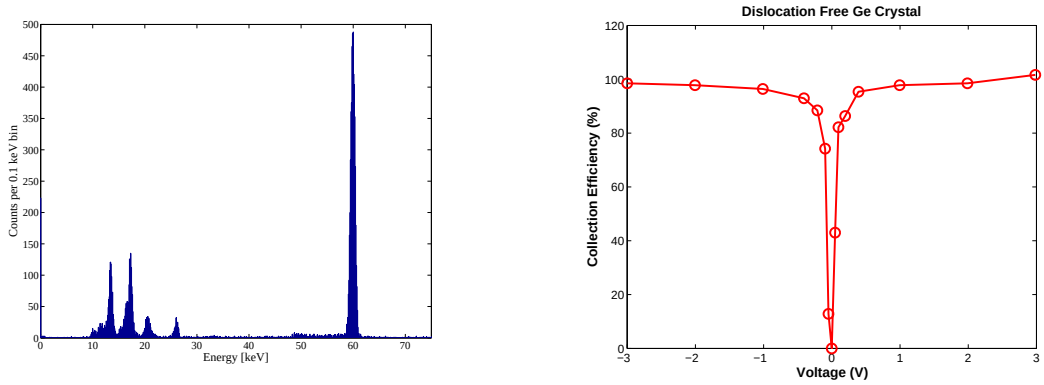
$\sigma_{SI} \sim 5 \times 10^{-45}$ cm^2. Based on this experiment and on continued research and development, we expect to propose a larger experiment of 100 kg target mass or more. This experiment would be located in an redesigned shield at the 6000 m.w.e. depth of SNOLAB [250] to reduce the neutron background to negligible levels.

8.4.2 Large detectors

For a 100-kg or ton-scale experiment, we would like to have detectors yet larger than the current prototypes. We cannot keep making detectors thicker indefinitely, however. A thicker detector requires a larger voltage bias to maintain the same electric field within its bulk; if the bias remains at 3V, we can expect an increase in low-yield backgrounds as dead-layer effects expand further into the detector. The contribution of Luke phonons, however, increases with bias voltage rather than field strength; as we increase the bias, these Luke phonons will swamp the primary phonons which provide our defense against surface events. We're thus in a bit of a bind: we can't make a detector arbitrarily thicker without endangering our discrimination power.

We are thus led to consider detectors of larger diameter, but here we find a problem as well. "Detector-quality" high-purity germanium must have a density of line-like crystal dislocations within a narrow range. Such dislocations help the crystal to relax away any vacancies which form during crystal growth, reducing the overall number of trapping sites in the crystal. A crystal with too high or low a dislocation density is unsuitable for ionization collection at 77 K. Thus far, however, it has proven very difficult to control the dislocation density sufficiently to grow substrates more than 3-4" in diameter. We thus seem limited in our ability to increase any of our detector dimensions.

Recent results from the Berkeley test facility, however, have reopened the possibility of increased detector diameters. A 1cm-thick ionization-only detector was constructed from a small sample of high-purity, dislocation-free germanium provided by Eugene Haller. The vacancies which thwart charge collection at 77K can be neutralized at millikelvin temperatures through illumination with an LED or a strong radioactive source. Figure 8.7(a) shows a clear spectrum of ^{241}Am obtained with this sample at a bias of 1V; Figure 8.7(b) shows that full charge collection persists to very low bias voltages, as with our usual detectors. Dislocation-free material is available in much larger diameters, and work is proceeding to acquire high-purity material for multi-kilogram detectors with 6-inch or larger diameters.



(a) Spectrum of ^{241}Am, measured in dislocation-free Ge at 1 V/cm.

(b) Charge collection efficiency as a function of voltage bias in dislocation-free Ge.

Figure 8.7: Sample plots of charge collection in a dislocation-free Ge sample, taken at Berkeley in February, 2008.

8.4.3 Alternate readout configurations

Symmetrical phonon readout

The asymmetric nature of our existing phonon readout scheme – only one face of each ZIP is instrumented with TES's – hampers our ability to reconstruct event position within our crystals. The cost of detector asymmetry is most apparent in the reconstruction of near-surface events: surface events on the charge face have slower phonon timing than those on the phonon face, and so are more difficult to discriminate from bulk recoils. More broadly, many phonon parameters are depth-dependent, particularly in thicker crystals; this effect is visible in the left-right color (risetime) gradients in Figure 8.6(b). The current one-sided TES configuration does not allow a clear depth parameter, which can thus translate into uncertainties in many aspects of event reconstruction.

A large ZIP detector would benefit from a symmetric phonon sensor configuration, with active elements on both faces of the detector. This configuration provides natural depth parameters: the relative arrival times and amplitudes of phonon pulses on each detector face. A start-time resolution of 0.5μs, for example, corresponds to a depth resolution of $\sim$ 2.7 mm in Ge; a surface-event cut based on timing alone in a one-inch ZIP could thus give signal efficiency comparable to that of the current timing cut. Relative pulse amplitudes can be even more powerful measurements of event depth, particularly at low pulse energies.

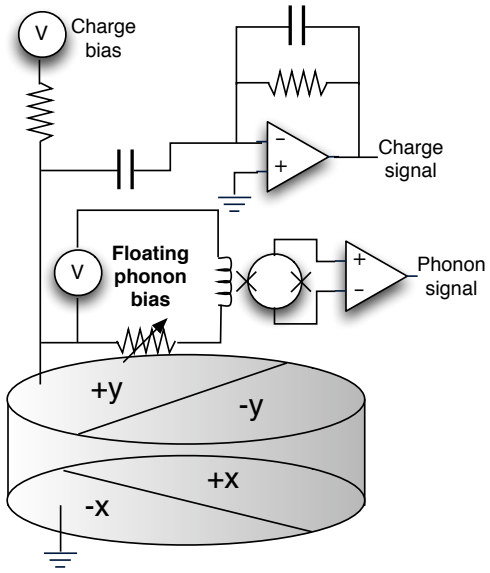


Figure 8.8: Schematic illustration of a floating-bias configuration for phonon readout on both detector faces. The bottom face has a standard non-floating phonon amplifier (not shown).

As an example, consider a 10 keV nuclear recoil (13 keV raw phonon energy) in a detector with a sensor configuration like that shown in Figure 8.8. With current energy resolutions ($\sim$ 100 eV/sensor), the difference in total energy deposited between the two faces could be measured to $\sim$ 3% accuracy. A true depth parameter constructed from timing and amplitude information would greatly improve position reconstruction and correction.

A symmetric phonon configuration is difficult to realize in practice because of a fundamental mismatch in the existing electronics. The TES biasing circuit is relatively low in impedance and high in capacitance, as it connects to multiple stripline traces (bias and return for each of four channels, plus extra returns for redundancy) through low-impedance components. The charge amplifier, in contrast, must have a high impedance at its input; any resistance or capacitance to ground will shunt away signal and increase noise. We thus cannot naively place a TES array on the same face as the charge amplifier without ruining our ionization signal-to-noise. Several configurations to address this are under investigation, most taking the general form shown in 8.8: biasing the TES array with some sort of low-capacitance floating bias circuit, providing up to 1 μV for the TES's without introducing more than a few pF of capacitance. The most promising solution thus far seems to be to

bias the TESs with a high frequency ($\sim$ 20 MHz) AC signal. Such a signal can still bias the TES effectively if it lies outside the ETF bandwidth, and it may be isolated from the charge amplifier by coupling through transformers or small capacitors.

Multiplexed phonon sensors

As detector diameters become larger, it's likely that our current phonon sensor arrangement – four sensors on one face – will become increasingly inadequate. Larger sensor arrays dilute localized effects of phonon timing and distribution, washing out information necessary for surface event rejection. Larger detector areas demand more sensor channels per surface, and each new channel brings new cold electronics, new wires, and new heat load to the icebox.

A multiplexed phonon readout system, in which several channels are read out using one amplifier and associated wiring, would provide the benefits of greater "pixelation" without the associated costs. The MIT group is pursuing a time-domain multiplexing strategy, in which a single SQUID amplifier switches between several TES input channels. Following upon development efforts for microwave and submillimeter astronomy [225], the Berkeley and Caltech groups are also pursuing a frequency-domain multiplexing scheme.

Kinetic inductance phonon readout

The Caltech and Berkeley groups are also currently pursuing kinetic inductance detector (KID) technology [251], an entirely different way to sense phonons in a crystal substrate. KIDs take advantage of the small change in a superconducting film's reactance with the addition of quasiparticles. If that superconducting film is part of a tank circuit driven near resonance, even a small deviation from resonance can be visible in the transmission properties of a coupled transmission line. KID readout is naturally multiplexable in the frequency domain, since each resonator can be tuned to a unique resonance frequency and the whole array driven by a single frequency comb. KID's require relatively low-resolution patterning of a single metal film, and so are easier to manufacture than high-resolution, multi-metal QET arrays. KID technology is also intrinsically low in noise, since the film remains essentially non-dissipative at all times; noise from surface states in the substrate can still remain, however, as does noise from the associated HEMT amplifier. Rapid progress is being made in this technology as a viable alternative for a large-scale WIMP-search

experiment.

8.5 Closing remarks

The Cryogenic Dark Matter Search is the world leader in the expanding field of WIMP direct detection. This dissertation has described the strongest upper limits yet set on WIMP-nucleon spin-independent scattering, and this reach is expected to grow by a factor of ~ 4 by 2009. CDMS is currently developing the next generation of cryogenic particle detector, leveraging the wealth of information provided about each particle event to extend our sensitivity to the zeptobarn-scale and beyond.

The larger community of dark matter searches is in the midst of a rapid expansion. Direct detection is becoming an ever-more-competive field, as new technologies vie with proven ones to leapfrog deeper into WIMP parameter space. Indirect detection is beginning to benefit from the unprecedented sensitivities of new instruments, led by GLAST and a new wave of Cherenkov telescopes. Finally, the Large Hadron Collider promises to revolutionize our understanding of particle physics at the TeV scale. Given these developments, there is a very real possibility that the next few years will finally see an answer to a question that has plagued science for so long: what is the universe made of?

Appendix A

Optimal filter charge reconstruction

This Appendix outlines the theory of optimal filtering, its particular implementation for the CDMS ionization channels, and some improvements to this implementation. I include this primarily for completeness. The 2-dimensional optimal filter used in the CDMS ionization pulse analysis is not documented elsewhere (though its use predates my involvement in the project), to my knowledge, nor is the sub-bin interpolation scheme of Section A.4. For further details on the general theory of noise and optimal filtering, please see the extensive discussion and derivation in Appendix B of Sunil Golwala's dissertation [166].

A.1 Notation and conventions

In this discussion I make frequent use of the Fourier transform, which maps a function $g(t)$ to its complex-valued counterpart $\tilde{g}(f)$ in the frequency domain. Among the several normalization conventions for the Fourier transform in broad use, I follow that advocated in [166] for the Fourier transform and its inverse:

$$\tilde{g}(f) = \int_{-\infty}^{+\infty} dt\, g(t) e^{-2\pi i f t} \tag{A.1}$$

$$g(f) = \int_{-\infty}^{+\infty} df\, \tilde{g}(f) e^{2\pi i f t}. \tag{A.2}$$

Other authors often write these formulae in terms of $\omega = 2\pi f$ rather than f itself, in which case $\int df \rightarrow \frac{1}{2\pi} \int d\omega$. I choose this convention as it allows the transform and its inverse to

be expressed in maximally symmetric notation, with no mysterious normalization factors. Note that the frequency variable f ranges over both positive and negative frequencies, and for a real-valued function $g(t)$ we have $\tilde{g}(-f) = \tilde{g}^*(f)$.

In practical situations we generally deal with discretely-sampled data, consisting of N observed values spaced uniformly over a time span $T = N\Delta t$. The most useful extension of the Fourier transform to this situation is the discrete Fourier transform (DFT), which maps a discretely-sampled series $\{g_k\}$ to a corresponding frequency-domain series $\{\tilde{g}_n\}$ of the same length. The DFT and its inverse are given by:

$$\tilde{g}_n = \frac{1}{N} \sum_k g_k e^{-\frac{2\pi i n k}{N}} \tag{A.3}$$

$$g_k = \sum_n \tilde{g}_n e^{+\frac{2\pi i n k}{N}} \tag{A.4}$$

These expressions can be put in closer correspondence with the continuous case by writing the phase factor's exponent in the form $\frac{2\pi i n k}{N} = 2\pi i f_n t_k$, where $t_k \equiv k\Delta t = k\frac{T}{N}$ is the time at which the sample g_k was taken ($g_k \equiv g(t_k)$) and $f_n \equiv n\Delta f = \frac{n}{T}$ is the frequency corresponding to the transform element $\tilde{g}_n$. With this definition, the DFT satisfies discrete analogues of Parseval's theorem, the convolution theorem, and other basic results of Fourier analysis. All modern implementations of the DFT are based upon the Fast Fourier Transform (FFT) algorithm, which employs trigonometric identifies to evaluate the DFT in $\mathcal{O}(N \log N)$ time with good numerical accuracy.

I have not yet specified the exact range of indices k (and n) for the N samples. Most authors take $k = 0 \ldots (N-1)$, but for close analogy with the continuous case it is convenient to number our time steps (and thus our frequencies) symmetrically about zero: $k = -\frac{N}{2} \ldots (\frac{N}{2} - 1)$ for even N, $k = -\frac{N-1}{2} \ldots \frac{N-1}{2}$ for odd N. For the rest of this discussion I assume that N is even, but these results hold with slight changes of indices for odd N.

With this choice, we can easily convert between integral and discrete formulations [166]:

$$\begin{aligned} g(t_k) &\leftrightarrow g_k \\ \tilde{s}(f_n) &\leftrightarrow \frac{s_n}{\Delta f} \\ \int_{-\infty}^{+\infty} df &\leftrightarrow \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} \Delta f \\ \int_{-\infty}^{+\infty} dt &\leftrightarrow \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} \Delta t \end{aligned} \tag{A.5}$$

With this correspondence table in place, I will sometimes use the continuous formulation where convenient in the rest of this treatment. The continuous expressions are often easier to manipulate, but may be converted to their discrete forms by referring to Equation A.5. Where indices are not specified, they should be taken to range over $-\frac{N}{2} \ldots (\frac{N}{2} - 1)$.

In my discussion of various estimation algorithms below, I will distinguish estimators from the parameters they estimate using a carat. In other words, the estimator for a model parameter a is denoted $\hat{a}$. $\hat{a}$ may be a function of the data and of other model parameters.

A.2 Optimal filter theory

The accuracy with which we can extract a signal from a background (*e.g.* an event pulse from random noise) depends strongly upon our knowledge of the characteristics of each. The more we know about the form of our signal and the properties of our background, the more precisely we can distinguish the two. Optimal filtering is a technique for addressing the signal extraction problem under particularly favorable circumstances: estimation of the amplitude of a signal of known shape $A(t)$ amidst a background of gaussian random noise of known power spectral density $J(f)$. In other words, our signal trace takes the form of $S(t) = aA(t) + n(t)$, where $A(t)$ is a known template and $n(t)$ is a noise realization with $J(f) = \left\langle \tilde{n}^2(f) \right\rangle$, and we wish to estimate the amplitude a. Optimal filtering is the optimal technique for amplitude estimation under these conditions, and often shows excellent performance even when these assumptions fail slightly.

A.2.1 Simple amplitude estimators

Before discussing the actual optimal filter algorithm , it is useful to consider two simpler estimators for a that one might consider. The simplest estimate of a is just the peak value of the signal trace: assuming that $\lim_{t\to\pm\infty} A(t) = 0$ for simplicity, this gives $\hat{a}_{peak} = \frac{\max S_k}{\max A_k}$. This estimator is quite poor, however, since it assumes no knowledge about the pulse shape or noise spectrum; it may be swayed substantially by noise fluctuations at a single sample along the trace.

A better solution is to perform a time-domain fit for a by minimizing the χ^2 between template and model:

$$\chi^2_{TD}(a) = \sum_n \frac{(S_k - aA_k)^2}{\langle n_k^2 \rangle}. \tag{A.6}$$

This expression may be minimized analytically by setting $\frac{d\chi^2}{da} = 0$ and solving for a:

$$\hat{a}_{TD} = \frac{\sum_k \frac{S_k A_k}{\langle n_k^2 \rangle}}{\sum_k \frac{A_k^2}{\langle n_k^2 \rangle}}. \tag{A.7}$$

Note that this expression allows the noise variance $\langle n_k^2 \rangle$ to vary with time; if the noise is time-invariant, it drops out of this expression:

$$\hat{a}_{TD} = \frac{\sum_k S_k A_k}{\sum_k A_k^2}. \tag{A.8}$$

The estimator in Equation A.7 makes full use of our knowledge of the signal, but rests upon a subtle assumption: Equation A.6 is only a proper χ^2 if the noise fluctuations in consecutive time bins are statistically independent. In the frequency domain, this is equivalent to a noise power spectrum $J(f)$ that is independent of f, commonly known as "white noise." Conversely, if the noise has a non-trivial power spectrum then noise fluctuations at different times are correlated and this estimator is no longer a proper maximum-likelihood estimator.

A.2.2 Optimal amplitude estimator

Since different frequency components are truly independent for gaussian random noise, a better solution is a one-parameter fit in the frequency domain. In analogy with

Equations A.6 and A.7, we can construct a frequency-domain χ^2 for such a fit

$$\chi^2(a) = \sum_n \frac{\left|\tilde{S}_n - a\tilde{A}_n\right|^2}{J_n}. \tag{A.9}$$

and minimize it as a function of a to find our parameter estimate

$$\hat{a} = \frac{\sum_n \frac{\tilde{A}_n^* \tilde{S}_n}{J_n}}{\sum_n \frac{|\tilde{A}_n|^2}{J_n}}. \tag{A.10}$$

This expression can also be viewed as the application of an (acausal) "optimal filter" ϕ, defined by $\phi \equiv \frac{\tilde{A}_n^*}{J_n}$. This interpretation in terms of a filter is unique to the linear case of estimating a pulse amplitude; more complex parameter estimation problems cannot generally be viewed in this light.

A.2.3 Estimating the time offset

The exact start time of the pulse within the digitizer window varies between traces, and so must be estimated as well. Translating the function $A(t)$ by a time offset t_0 corresponds to multiplying its Fourier transform by a phase factor $e^{2\pi i t_0 f}$. If a template feature occurs at time $t = 0$, the shifted template shows the same feature at $t = t_0$. Note that this scheme for time-translation grafts the end of the template onto its beginning (or vice-versa), so edge effects can result if $A(t) \neq 0$ for large $|t|$. Note also that t_0 must be an integral multiple of the digitizer bin width T/N in the discretized expressions, though this restriction is lifted for their continuous equivalents. We can thus shift the template in Equation A.9 by a time offset t_0, producing

$$\chi^2(a, t_0) = \sum_n \frac{|\tilde{S}_n - ae^{-2\pi i t_0 f_n}\tilde{A}_n|^2}{J_n} \tag{A.11}$$

To proceed, we should minimize this χ^2 to find the best estimates $\hat{a}$ and $\hat{t_0}$. Given an estimate $\hat{t_0}$, we can follow the derivation of Equation A.10 with a shifted template to obtain the optimal amplitude estimator:

$$\hat{a}(t_0) = \frac{\sum_n e^{2\pi i t_0 f_n} \frac{\tilde{A}_n^* \tilde{S}_n}{J_n}}{\sum_n \frac{|\tilde{A}_n|^2}{J_n}}. \tag{A.12}$$

We can similarly fine the value of t_0 which minimizes χ^2 by solving

$$0 = \frac{\partial \chi^2}{\partial t_0} = -2a \sum_n 2\pi i f_n e^{2\pi i f_n t_0} \frac{\tilde{A}_n^* \tilde{S}_n}{J_n}. \tag{A.13}$$

Unfortunately, this is a nonlinear equation that has no analytic solution. Instead, we take the derivative of Equation A.12 to find that

$$\frac{\partial \hat{a}}{\partial t_0} = \frac{\sum_n 2\pi i f_n e^{2\pi i \hat{t_0} f_n} \frac{\tilde{A}_n^* \tilde{S}_n}{J_n}}{\sum_n \frac{|\tilde{A}_n|^2}{J_n}}. \tag{A.14}$$

Dropping constants that do not depend on t_0, this shows that $\frac{\partial \chi^2}{t_0} \propto \frac{\partial \hat{a}}{\partial t_0}$. The value of t_0 which extremizes $\chi^2(a, t_0)$ is thus also the value which extremizes $\hat{a}(t_0)$. Checking the second derivatives, we find that these extrema are a minimum of χ^2 and a maximum of $\hat{a}$. The best-fit value of $\hat{t_0}$ is thus that value which maximizes the amplitude estimate $\hat{a}(t_0)$.

At first glance, this doesn't seem to have helped our computational situation. We need to compute $\hat{a}(t_0)$ (Equation A.12) for a range of values of t_0 in order to find the best-fit value, but we could just as well have computed $\chi^2(\hat{a}(t_0), t_0)$ (Equation A.11) for each value of t_0. The advantage of the former is that it is much faster to compute $\hat{a}$ for a range of values of t_0 than it is to compute χ^2 many times. The numerator of Equation A.12 (the only part that depends upon t_0) is just the inverse-DFT of $\tilde{A}_n^* \tilde{S}_n / J_n$, so we can compute the whole function $\hat{a}(t_0)$ very quickly using the FFT algorithm and just take its maximum value. No such optimization is possible for Equation A.11.

A.2.4 Performance

Resolution

The amplitude and start time resolutions of the optimal filter can be determined from the second derivatives of the χ^2 function, evaluated at the minimum where the first derivatives vanish:

$$\sigma_{\hat{a}}^2 = \left[\int_{-\infty}^{+\infty} df \, \frac{|\tilde{A}(f)|^2}{J(f)} \right]^{-1} \quad \text{or} \quad \sigma_{\hat{a}}^2 = \left[\sum_n \frac{|\tilde{A}_n|^2}{J_n} \right]^{-1} \tag{A.15}$$

$$\sigma_{\hat{t_0}}^2 = \left[\hat{a}^2 \int_{-\infty}^{+\infty} df \, (2\pi f)^2 \frac{|\tilde{A}(f)|^2}{J(f)} \right]^{-1} \quad \text{or} \quad \sigma_{\hat{t_0}}^2 = \left[\hat{a}^2 \sum_n (2\pi j f_n)^2 \frac{|\tilde{A}_n|^2}{J_n} \right]^{-1} \tag{A.16}$$

Note that $\sigma_{\hat{a}}$ is independent of pulse amplitude a, but timing resolution degrades at small pulse amplitudes ($\sigma_{\hat{t_0}} \propto a^{-1}$).

Digitization error

In the discrete case the algorithm above can only choose $\hat{t_0}$ with an accuracy of one digitizer bin ($\Delta t = T/N$). This limit can be defeated through the use of interpolated values for $\hat{a}(t_0)$, as described later in this Appendix. If no interpolation is performed, however, our choice of $\hat{t_0}$ does not actually lie at the minimum of χ^2 and we must add additional contributions to the uncertainties above. For a digitizer with sample spacing Δt, the r.m.s. timing uncertainty is $\delta_t = \frac{\Delta t}{\sqrt{12}}$. For an error in the start time of magnitude δt_0, we introduce additional errors of

$$\delta a \sim \frac{1}{2}\frac{\partial^2 a}{\partial t_0^2}\delta t_0^2 = -a\delta t_0^2 \frac{\sum_n (2\pi i f_n)^2 \frac{|\tilde{A}_n|^2}{J_n}}{\sum \frac{|\tilde{A}_n|^2}{J_n}} \tag{A.17}$$

$$\delta\chi^2 \sim \frac{1}{2}\frac{\partial^2 \chi^2}{\partial t_0^2}\delta t_0^2 = +\frac{1}{2}a^2\delta t_0^2 \sum_n (2\pi i f_n)^2 \frac{|\tilde{A}_n|^2}{J_n}. \tag{A.18}$$

Note that an error in start time always leads to an underestimate of the amplitude and an overestimate of χ^2, as expected. These errors in amplitude are generally quite small for the CDMS analysis: only $\sim 0.15\%$ on average, and at most $\sim 0.5\%$. The error on χ^2 scales as the square of the pulse amplitude, however, and widens the distribution of the ionization optimal filter χ^2 quantity (`QSOFchisq`) substantially at high pulse energies.

Noise blobs

If we apply this optimal filtering algorithm to a series of noise traces containing no actual pulses ($a = 0$), we might naively expect to obtain a distribution of estimates $\hat{a}$ with zero mean and a variance of $\sigma_{\hat{a}}^2$. In practice, however, we observe a "noise blob" with a somewhat narrower distribution of amplitude estimates and mean greater than zero. This upward bias is introduced by our algorithm for finding the time offset: by searching for the maximum value of $\hat{a}$, we introduce a small upward bias when there is no actual signal to find. The true experimental resolution can be measured by forcing $t_0 = 0$ for noise traces.

If we model the $\hat{a}(t_0)$ for each noise trace as a gaussian random variable (zero mean, variance $\sigma_{\hat{a}}^2$) and try N different values of t_0, our algorithm is equivalent to selecting the largest among N values drawn from that gaussian distribution. We can compute the expected probability density function $PDF_{max}(x, N, \sigma)$ for this maximal value x by noting that the cumulative distribution function (CDF) of the maximum of two variables is just

the product of their two individual CDFs. For our case, this yields:

$$PDF_{max}(x, N, \sigma) = \frac{d}{dx} CDF_{max}(x, N, \sigma) = \frac{d}{dx} \left(CDF_{normal}(x, \sigma)\right)^N \tag{A.19}$$

$$= \frac{N}{2^N \sigma} \sqrt{\frac{2}{\pi}} \left(1 + \mathrm{erf}\left(\frac{x}{\sigma\sqrt{2}}\right)\right)^{N-1} e^{-x^2/2\sigma^2}. \tag{A.20}$$

A different derivation of the same expression is due to Sunil Golwala and Rick Gaistkell [166]. As expected, this distribution is displaced from zero and slightly narrower than σ. It is difficult to match quantitatively to data, however, due to the deficiencies of this model: for a filter of finite bandwidth, $\hat{a}(t_0)$ and $\hat{a}(t_0 + \Delta t)$ are not completely uncorrelated. This reduces the effective N, and produces less upward bias in practice than this simple model would predict.

A.3 CDMS implementation

For posterity's sake I describe below to details of the DarkPipe implementation of this algorithm:

1. A set of search windows to limit the range of allowed start times
2. An extension of the basic algorithm to account for known crosstalk between the two ionization channels.

A.3.1 Search windows

If the pulse template is shifted too far toward the edge of the digitization window, edge effects can lessen the quality of the fit. DarkPipe requires that all ionization pulses occur within a window of $[-100, +10]\mu$s for Ge ($[-50, +10]\mu$s for Si) around the global trigger time. This condition should be satisfied for all normally-triggered events, but events can occasionally fall outside this window at high trigger rates. For details see the discussion of the cut `cGoodPStartTime` in Section 6.4.

A.3.2 Crosstalk

The two ionization channels on each detector have a small but noticeable ($\sim$ 6%) crosstalk between them, primarily due to the mutual capacitance of the electrodes. DarkPipe uses a generalization of the optimal filter algorithm to properly fit both amplitudes

and a single start time (the two ionization pulses are simultaneous at the precision of our digitizer resolution) self-consistently, with proper attention to inter-channel crosstalk.

To properly account for crosstalk we characterize our two-channel system by four templates : the Qinner signal from a Qinner event (A_I), the Qouter signal from a Qouter event (A_O), and two crosstalk templates giving the Qouter signal resulting from a Qinner event (A_{Ix}) and vice-versa (A_{Ox}). These templates are generated by averaging samples of several hundred calibration events. Examples from Run 123 are shown in Figure 3.6. Note that the A_{Ix} template generally has poor signal-to-noise due to the relative dearth of well-reconstructed Qouter events.

We generalize the optimal filter to reconstruct two amplitudes a_I and a_O simultaneously, given these four templates and two noise spectra, J_O and J_I. The equation to be solved now takes the form of a matrix expression:

$$\begin{pmatrix} S_I \\ S_O \end{pmatrix} = \begin{pmatrix} A_I & A_{Ix} \\ A_{Ox} & A_O \end{pmatrix} \begin{pmatrix} a_I \\ a_O \end{pmatrix} + \begin{pmatrix} n_I \\ n_O \end{pmatrix}$$

We can construct a χ^2 function in frequency space just as before, suppressing the subscripts n on S, A, and J for clarity:

$$\chi^2(a_I, a_O) = \sum_n \left(\frac{|\tilde{S}_I - (a_I \tilde{A}_I + a_O \tilde{A}_{Ix})|^2}{J_I} + \frac{|\tilde{S}_O - (a_I \tilde{A}_{Ox} + a_O \tilde{A}_O)|^2}{J_O} \right). \tag{A.21}$$

Solving for the minimum of χ^2 with respect to both amplitudes a_I and a_O, we obtain the estimator

$$\begin{pmatrix} \hat{a}_I \\ \hat{a}_O \end{pmatrix} = M^{-1} \begin{pmatrix} \sum_n \mathrm{Re}\left(\frac{\tilde{A}_I^*}{J_I} \tilde{S}_I + \frac{\tilde{A}_{Ox}^*}{J_O} \tilde{S}_O \right) \\ \sum_n \mathrm{Re}\left(\frac{\tilde{A}_{Ix}^*}{J_I} \tilde{S}_I + \frac{\tilde{A}_O^*}{J_O} \tilde{S}_O \right) \end{pmatrix}, \tag{A.22}$$

in which

$$M \equiv \begin{pmatrix} \sum_n \frac{|\tilde{A}_I|^2}{J_I} + \frac{|\tilde{A}_{Ox}|^2}{J_O} & \sum_n \mathrm{Re}\left(\frac{\tilde{A}_I^* \tilde{A}_{Ix}}{J_I} + \frac{\tilde{A}_O^* \tilde{A}_{Ox}}{J_O} \right) \\ \sum_n \mathrm{Re}\left(\frac{\tilde{A}_I^* \tilde{A}_{Ix}}{J_I} + \frac{\tilde{A}_O^* \tilde{A}_{Ox}}{J_O} \right) & \sum_n \frac{|\tilde{A}_O|^2}{J_O} + \frac{|\tilde{A}_{Ix}|^2}{J_I} \end{pmatrix}. \tag{A.23}$$

Note that most of the laborious computation in Equation A.22 lies in the calculation of the 2×2 matrix M and the ratios $\tilde{A}_I / J_I$, etc. These calculations only need to be performed once, rather than every time we process a trace. DarkPipe stores M^{-1} and the terms on the right hand side independent of $\tilde{S}_{I,O}$ in a file for use during data processing.

I have so far avoided discussion of the timing offset t_0 and how to estimate it (note that both channels have the same start time within our experimental resolution).

The procedure essentially follows that above – find the offset which gives the maximum amplitude estimate – but with a new ambiguity in the multi-dimensional case: which amplitude estimate should we use? A reasonable choice is to maximize $\hat{a}_I + \hat{a}_O$, though this can be problematic in rare cases in which the two pulses have similar magnitudes and opposite signs. In any case, the matrix expression in Equation A.22 is less amenable to being sped up with an inverse-DFT. For speed, DarkPipe chooses the start time which maximizes $\hat{b}_I + \hat{b}_O$, where these are the optimal-filter amplitudes determined from separate, crosstalk-free estimates of the two channels with their primary templates. Insofar as crosstalk is small and does not distort the pulse shapes too much, this should be sufficiently accurate for practical purposes. Once the start time is determined in this way, we can add an appropriate phase factor and proceed with the full matrix calculation.

A.4 Extension for improved timing resolution

Based upon Equation A.16, the expected timing resolution for a Run 123 ZIP detector is quite small: $\sigma_{\hat{t_0}} = 0.015\mu\text{s}/E$ for T2Z4, where E is the ionization energy in keVee. This is essentially insignificant compared to the digitizer bin width for nuclear recoils above ~ 10 keV. As discussed in Section A.2.4, the added uncertainty in pulse timing has negligible effect on our charge resolution ($\delta a/a < 1\%$). It does have two noteworthy effects on our analysis, however:

1. The digitizer error introduces an additional spread in the χ^2 distribution, widening it substantially at high pulse energies. This wide distribution forces us to loosen our cut on χ^2 to reject pileup events, possibly allowing small coincident pulses through into our analysis.

2. The phonon delay quantities used for surface event rejection and position reconstruction are measured from the ionization pulse start time, so this introduces additional uncertainty in to those measurements. This error is often significant compared to the errors in phonon pulse timing from the walk algorithm, especially for the faster phonon pulses seen in Si detectors.

Neither of these effects were substantial impediments to the current data analysis, but it could be useful to improve them for future analyses.

We can improve upon the limitations of digitizer error by interpolating the start time to values between adjacent digitizer bin edges. Unfortunately, such values of $t_0 \neq k\Delta t$ cannot be substituted directly into discrete expressions such as Equations A.11 or A.12. Instead, we approximate $\hat{a}(t_0)$ near its extremum by a parabola and interpolate between three neighboring times: $\hat{t_0} - \Delta t$, $\hat{t_0}$, and $\hat{t_0} + \Delta t$, where $\hat{t_0}$ is the discrete value maximizing $\hat{a}$. We fit a parabola to the values of $\hat{a}$ at these three points, find its maximum, and take the abscissa and ordinate of that maximum as our improved estimates of $\hat{t_0}$ and $\hat{a}$, respectively. All of these steps can be performed analytically for a parabola with little cost in computing time. The value of χ^2 at this new offset can be evaluated using a similar interpolation, but it is faster to approximate it based upon the closest discrete t_0 value and Equation A.18.

The situation is somewhat more complex for the two-channel algorithm described in Section A.3.2. The computational cost of the optimal filter goes up, and it is less obvious what quantity we should use to interpolate the maximum. Ideally we would compute the full χ^2 of Equation A.21 for all t_0 and interpolate near the maximal value, but this is computationally prohibitive. The fastest solution is to interpolate the start time using the simpler no-crosstalk fit we currently use to determine $\hat{t_0}$, then correct the fitted amplitudes and χ^2 values using first-order corrections analogous to those in Section A.2.4. This suffers from some systematic errors from the misuse of crosstalk templates, but should be a substantial improvement upon the non-interpolated algorithm. With a little more computational effort, we can compute χ^2 and $\hat{a}_{I,O}$ for the peak digitizer bin and its two neighbors, interpolate the maximum in χ^2, then apply this to interpolate the amplitudes. Various other combinations are possible. It is important to restrict the interpolation from choosing a start time outside of $\hat{t_0} \pm \Delta t$, which would indicate that the quadratic approximation is invalid; this should not happen for traces containing significant pulses, but may occur in some noise traces for certain interpolation schemes.

As a proof of principle, I have tested a conservative (*i.e.* relatively computationally-intensive) version of this algorithm upon a small selection of calibration data from the one-inch prototype detector G3D at Berkeley. The current code interpolates the best-fit time offset by maximizing $(\hat{a}_I + \hat{a}_O)$ (computed using the full crosstalk correction) over all possible start times and interpolates the χ^2 based upon this choice of offset. Figure A.1 shows the effects of this interpolation: a substantially narrower distribution of χ^2, even at ionization energies of several hundred keV. This conservative algorithm degraded the processing speed of DarkPipe by a factor of ~ 5, which is clearly unacceptable in practice.

Lauren Hsu and Bruno Serfass are implementing this algorithm into the new data analysis package which will replace DarkPipe, and will take advantage of one of the faster choices of interpolation scheme.

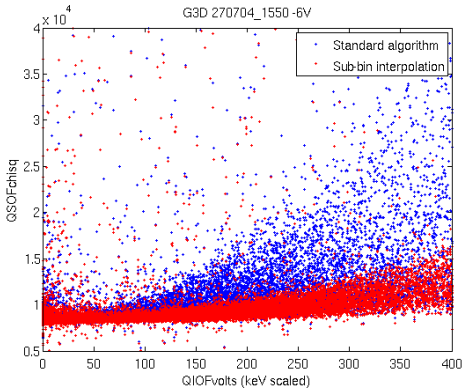


Figure A.1: Optimal filter χ^2 versus ionization energy for inner-electrode events in a G3D calibration run at UC Berkeley. The blue points show the results of the standard algorithm, while the red points show the (substantially narrower) distribution of χ^2 produced by the sub-bin interpolation code.

A.5 Multi-parameter fits

The formalism of optimal filtering is easily extended to fit more complex models with several free parameters. In this regime optimal filtering evolves into a general nonlinear fit in the frequency domain, in which we minimize an appropriate χ^2 function

$$\chi^2(\{\theta_i\}) = \sum_n \frac{|\tilde{S}_n - \tilde{A}_n(\{\theta_i\})|^2}{J_n} \tag{A.24}$$

to find the best-fit parameters $\{\hat{\theta}_i\}$. The mechanics are essentially identical to those of time-domain fitting algorithms, but a frequency-domain algorithm should show better performance in the presence of non-white noise.

Such an algorithm may be a useful approach in situations where pulse shape varies between events. Such fits are very computationally intensive, however, and may provide little benefit if the template does not match the pulses well or if the fitting problem is ill-conditioned. The fitting of phonon pulses with multi-exponential templates is both imperfect and ill-conditioned, so this avenue has not yet borne fruit for CDMS.

Appendix B

Likelihood ratio and χ^2 discrimination

In discriminating nuclear recoils from surface electron recoils we are faced with the problem of efficiently combining several observations to yield a single decision rule. This section describes a general method for making such decisions, a second order approximation to the general likelihood ratio test. We also describe some extrapolations from the simplest discrimination model.

B.1 Likelihood ratio tests

Suppose that we would like to classify events as members of either one of two populations (given the suggestive designations "signal" and "background") based on N observables $\vec{x}$. Suppose further that we can characterize the two populations by their probability densities ("likelihoods", in classical language) in these observables, $\mathcal{L}_{sig}(\vec{x})$ and $\mathcal{L}_{bkg}(\vec{x})$. If we are given an unknown event with observables $\vec{x}_0$ and wish to classify it with maximal accuracy, we intuitively would like to choose a signal acceptance region R_{sigl} in $\vec{x}$-space to minimize the misclassification rate ($\int_{R_{sig}} \mathcal{L}_{bkg}(\vec{x})$) at fixed signal acceptance ($\int_{R_{sig}} \mathcal{L}_{sig}(\vec{x})$). Following the argument in [252], this occurs when R_{sig} is defined by

$$\frac{\mathcal{L}_{sig}(\vec{x}_0)}{\mathcal{L}_{bkg}(\vec{x}_0)} > c, \tag{B.1}$$

where c is chosen to attain the desired acceptance of signal ("efficiency") and background ("leakage"). This is the standard *likelihood ratio test* discussed in statistics texts. By

reversing the "signal" and "background" labels, one can show that such a test also maximizes efficiency at given leakage. For later convenience we can also write the same relationship in terms of a difference of log-likelihoods:

$$\log \mathcal{L}_{sig}(\vec{x}_0) - \log \mathcal{L}_{bkg}(\vec{x}_0) > \log c \tag{B.2}$$

Note that the choice of classification threshold c should depend upon the goal of our test. A threshold of zero gives equal weight to false positives (classification of background as signal) and false negatives (classification of signal as background), but in practice we may be much more concerned about one kind of error than the other and so zero may not be the optimal choice. In a rare event search we are generally willing to allow many false negatives to avoid a single false positive, and the optimal choice of threshold will depend upon the expected rate of background events.

B.2 χ^2 tests

A full likelihood ratio test is appealing in principle, but may be difficult to implement if the likelihood distributions are complicated or unknown. To move forward, we make the simplifying assumption that each likelihood takes the form of a multivariate gaussian distribution. The resulting likelihood function is

$$\mathcal{L}(\vec{x}) = (2\pi \det |C|)^{-N/2}\, e^{-\frac{1}{2}(\vec{x}-\vec{\mu})^T \cdot C^{-1} \cdot (\vec{x}-\vec{\mu})}, \tag{B.3}$$

where $\vec{\mu}$ is the N-dimensional vector of parameter means and C is the $N \times N$ covariance matrix of the distribution:

$$C_{ij} \equiv \langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle . \tag{B.4}$$

The covariance matrix C encapsulates the variances σ_i^2 of each parameter ($C_{ii} = \sigma_i^2$) and the correlations $\rho_{ij} \in [-1, 1]$ between parameter pairs ($C_{i \neq j} = \rho_{ij}\sigma_i\sigma_j$). For a multivariate gaussian these fully characterize the distribution, but a general distribution will possess higher-order correlations not accounted for in this formalism.

In the gaussian approximation the likelihood ratio test takes a particularly simple form. Taking the logarithm of B.3 we obtain

$$\log \mathcal{L}(\vec{x}) = -N/2 \left(\log 2\pi + \log \det |C|\right) - \frac{1}{2}(\vec{x} - \vec{\mu})^T \cdot C^{-1} \cdot (\vec{x} - \vec{\mu}). \tag{B.5}$$

Substituting into Equation B.2,

$$\frac{N}{2}\log\frac{\det|C_{bkg}|}{\det|C_{sig}|}+\frac{1}{2}\left(\chi^2_{bkg}-\chi^2_{sig}\right)>c, \tag{B.6}$$

where χ^2 is defined as

$$\chi^2=(\vec{x}-\vec{\mu})^T\cdot C^{-1}\cdot(\vec{x}-\vec{\mu}). \tag{B.7}$$

In the simplest case the covariance matrices are constant, so we can lump the determinants in with c and the desired test reduces to

$$\chi^2_{bkg}-\chi^2_{sig}>\kappa. \tag{B.8}$$

The cut value κ can be chosen to incorporate both the covariance determinants and any desired bias between false positives and false negatives. For known gaussian distributions κ can be computed *a priori* using the above expressions, but in realistic situations with non-gaussian distributions the cut value is best chosen empirically to obtain the desired classification accuracy.

B.3 Extension for non-discriminatory parameters

Suppose that the means and variances of these distributions are functions of some additional parameter E which is not itself part of the main discrimination space. This is common in experimental situations - parameter resolutions often depend on event energy, for example. Here our dependence on the gaussianity assumption (particularly on the tails of the likelihood distribution) becomes more critical: non-gaussian tails can drastically change the expressions below. When in doubt we must return to the original likelihood ratio test (Equations B.1 or B.2), but the general discussion below can be more broadly informative.

In the E-dependent case the determinants in Equation B.6 cannot be treated as constants and will introduce dependencies of the test upon E. This could be treated as an E-dependence of κ, but to retain the intuitive meaning of χ^2 we lump these factors into its definition. This leaves Equation B.8 valid but replaces Equation B.7 with

$$\chi^2=(\vec{x}-\vec{\mu}(E))^T\cdot C^{-1}(E)\cdot(\vec{x}-\vec{\mu}(E))+N\log\det|C(E)|. \tag{B.9}$$

Even after this modification to χ^2 it is not necessarily desirable to apply Equation B.8 naively. This expression is the correct analog to the likelihood ratio test only if the

likelihood functions have no further dependence on E beyond the mean and covariance matrix. If the signal or background rate varies as a function of E (e.g. background rate varies with energy or position) then the likelihood ratio will further depend on the rate distributions $\rho_{sig}(E)$ and $\rho_{bkg}(E)$. In general this leads to a complex dependence of κ on $\rho_{sig}(E)/\rho_{bkg}(E)$, but given a likelihood model we can approximate $\kappa(E)$ using Monte Carlo.

B.4 Beyond the gaussian approximation

The derivation of the χ^2 test discussed above depends critically on the assumption of gaussianity, though the test itself (with κ set empirically) is often very productive even with significantly non-gaussian distributions. In principle we should perform a full fit to the likelihood functions and return to the raw likelihood ratio test when in doubt.

Short of that, it is often possible to retain the χ^2 formalism in non-gaussian situations through a "gaussianizing" parameter transformation. Given a variable x with known 1-dimensional distribution $\rho(x)$, then

$$y \equiv \Phi^{-1}\left(\int_{-\infty}^{x} \rho(x')dx'\right), \tag{B.10}$$

is a standard normally-distributed variable ($\mu = 0, \sigma = 1$) where $\Phi(z) = \frac{1}{2}(1+\mathrm{erf}(\frac{z}{\sqrt{2}}))$ is the normal cumulative distribution function (CDF). If we know the distribution of x *a priori*, this transformation is exact and the resulting χ^2 is the correct analogue of a likelihood. This analogy will be less precise if $\rho(x)$ must be determined empirically, but may nonetheless prove useful in removing lower-order distortions in $\rho(x)$.

In a multi-dimensional problem there is in general no simple way to gaussianize in all variables – or, rather, no way simpler than performing a likelihood ratio analysis using the full, multi-dimensional distribution functions. If the multi-dimensional distributions of each population are generally convex in shape, however, it may be useful (though not necessarily rigorously correct) to gaussianize independently along each parameter axes. This is unlikely to be useful for a highly non-convex distribution, *e.g.* a banana-shape in two dimensions.

Figure B.1 illustrates an example of this sort of analysis, applied to T1Z5 calibration data from Run 123. The left panel shows the distribution of `pfracc` for neutron events, showing the substantial asymmetry of that population. The overlaid curve indicates the probability density associated with a fitted CDF for this distribution. This fitted CDF was obtained in two steps: the inner 80% of events were used to compute a kernel-smoothed

empirical CDF, while the two outer tails were fitted to Pareto distributions. I performed similar fits for `pminrtc` and `pfracc` independently and carrier out a simple χ^2 analyses using the resulting transformed distributions. The right panel shows the clear separation between neutrons and surface events achieved using this technique. The separation between the two bulk populations is greater than that from a standard timing or χ^2 analysis. Unfortunately, its performance is limited at very low background levels by the same mis-corrected events which plague other timing analyses. This technique is promising, but was not pursued for the primary blind analysis described in this dissertation.

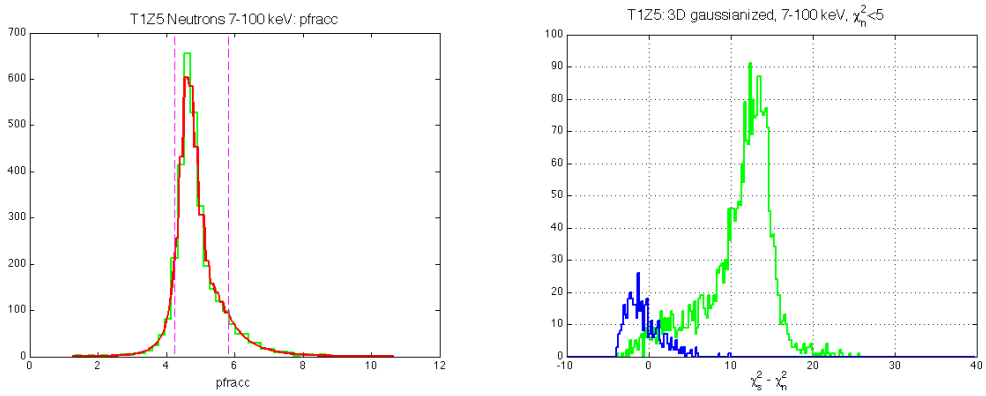


Figure B.1: *Left*: Distribution of `pfracc` for neutrons from the Run 123 calibrations of T1Z5, overlaid with the density function from the associated gaussianization transformation. The vertical dotted lines bracket the central 80% of events, which are fit with a smoothed empirical CDF. The outer tails are fitted to Pareto distributions. *Right*: Distribution of surface (*blue, dark*) and neutron (*green, light*) events in the $\Delta\chi^2$ discrimination parameter derived from the transformed parameter distributions.

B.5 Application to CDMS II data

These ideas were applied to CDMS II data during the analysis of the second data run at Soudan (Run 119). This work was part of a suite of "advanced" timing analyses from that run, developed to test alternative statistical techniques for surface event rejection. A simple χ^2 cut based upon the methods of Section B.2 was initially explored in parallel by Joel Sander and Rupak Mahapatra at UCSB and by me. The UCSB group continued to develop this analysis, which is discussed in detail in Joel Sander's dissertation [189]. Building upon this, I developed a similar analysis following Section B.3, attempting to account for the energy-dependence of the relevant covariance matrices. Both of these analyses were

conducted in accordance with our criteria for a blind analysis. The timing cut associated with this energy-dependent ("edep") χ^2 analysis was chosen before unblinding to be the primary cut for Si detectors in that analysis, and was used to establish the published WIMP-search limit from that data set [3, 5]. In this Section I give a brief description of the energy-dependent ("edep") χ^2 analysis initially carried out in that run.

This Section documents work from a previous data run, and is adapted from a brief internal note on the topic. This analysis was an initial attempt, and does not take advantage of several improvements in analysis techniques since Run 119.

B.5.1 Characterizing the populations

Event samples

The surface event and nuclear recoil event populations were characterized using "gold" samples of surface and neutron events from calibration data with radioactive sources. These events are assumed to accurately represent our surface event backgrounds and signal events, respectively. All gold events were chosen to pass the standard quality cuts of that run. The surface sample was composed of ^{133}Ba calibration events satisfying the quality cuts above and the standard Run 119 surface event cut, given by

- Charge yield below the 5σ gamma band, selected with `gband_qi_R119`
- Charge yield (`yic`) between 0.1 and a detector-specific maximum ([.6 .7 .85 .8 .85 .9 .75 .85 .85 .75 .8 .8] for Z1-Z12)

The neutron sample was composed of ^{252}Cf calibration events satisfying the quality cuts and

- `cNR_119`: Standard nuclear recoil band cut (set with `qi` quantities)

Parameter distributions

Like the basic χ^2 analysis, this analysis worked within a three-dimensional parameter space defined by `pminrtc` (primary phonon risetime), `pdelc` (primary phonon delay), and `pfracc` (the ratio of the largest and opposite phonon amplitudes). These parameters were corrected using a correction algorithm less advanced than that described in Section

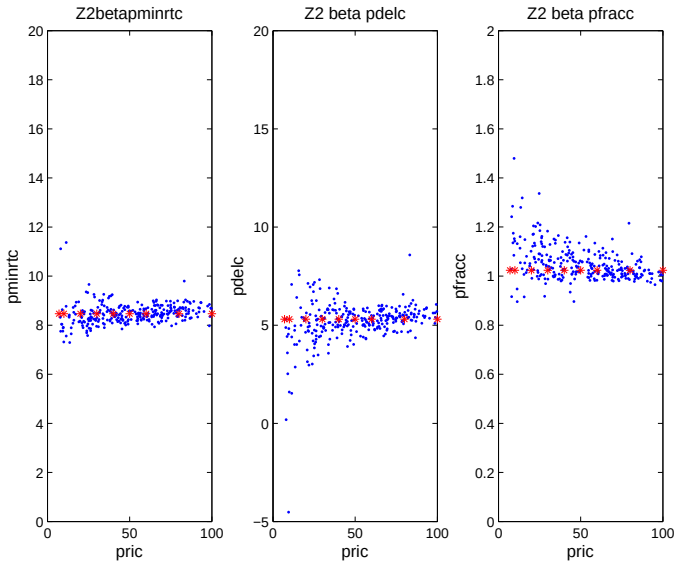


Figure B.2: Distributions of gold surface events from detector T1Z2 in the three phonon pulse shape parameters, each as a function of recoil energy (`pric`). Red stars indicate the fitted mean of each parameter, taken to be constant with energy in this analysis.

3.3.6: the energy dependences of the phonon parameters were corrected out using an approximate fit function, rather than including energy within the correction table. The Run 119 phonon parameters thus showed greater residual energy dependence than those used in the bulk of this dissertation.

Figures B.2 and B.3 show the distribution of the gold events in each of these three phonon pulse shape parameters. Events from the Ge detector T1Z2 are shown. It is apparent that the parameter variances grow substantially at low energies, particularly for neutron events. The parameter distributions are also noticeably asymmetrical (particularly in the case of `pfracc`). The discrimination power of `pfracc` is not obvious from these plots, but arises primarily from its strong anti-correlation with variations in the other two timing parameters.

Parameter Fits

For each detector and event type (surface or neutron), we compute the mean $\mu_{b(n)}$ of each event distribution using an iterative Gaussian fit in each parameter on events passing simple outlier cuts. This fit excludes outliers at each iteration to provide better stability.

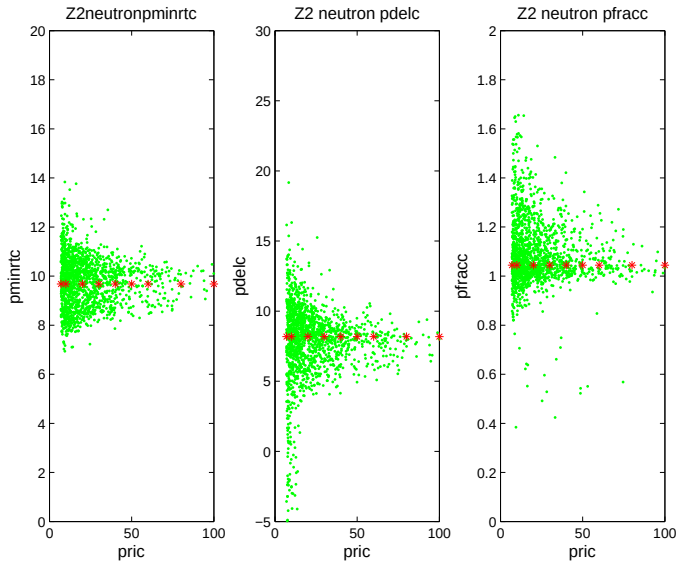


Figure B.3: Same as Figure B.2 for gold neutron events.

This mean is taken to be constant with energy for simplicity, though this is not precisely true.

We then divide the gold events into bins in recoil energy: 7-10, 10-20, 20-30, 30-40, 40-50, 60-80, and 80-100 keV. In each bin we compute the covariance matrix with respect to the fitted mean (not the actual mean within that bin, which has substantial statistical variation). Figures B.4 and B.5 show the variation of the six independent terms in the covariance matrix with energy, again for T1Z2. For the variances, we make a least-squares fit to a function of the form $\sigma^2(E) = a + b/E^2$, where E is the recoil energy (`pric`) and a and b are fitting parameters. This form is appropriate for the time at which a pulse crosses a fixed threshold in the presence of white noise: a term independent of energy added to a "time jitter" component whose r.m.s. amplitude scales inversely with the slope of the pulse edge (and thus with its energy). For simplicity, we make the approximation that the correlations are constant with energy and determine each using a least-squares fit. Altogether, this procedure fits 12 (3 means, 3 correlations, and 2 parameters fit to each of 3 variances) surface event parameters and 12 neutron parameters to each detector's gold events. All of these fits are shown as red curves in the figures.

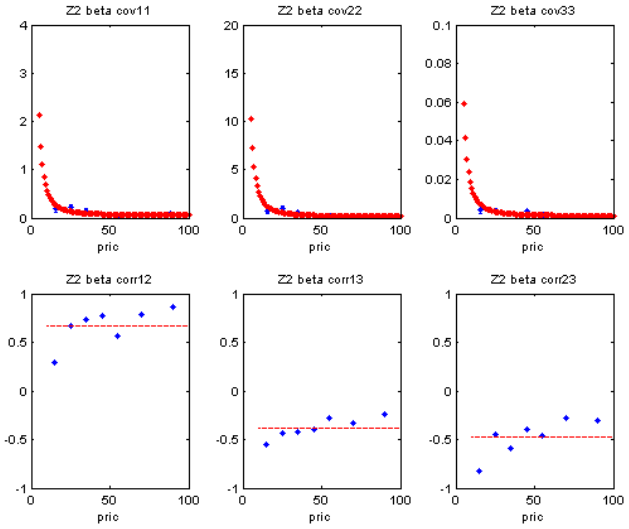


Figure B.4: Variation of covariance matrix terms with energy for gold surface events in T1Z2. Plots labeled "covii" (first row) represent diagonal elements C_{ii} of the covariance matrix, while those labeled "covij" (second row) show the normalized correlations $\rho_{ij} \equiv C_{ij}/\sqrt{C_{ii}C_{jj}}$. Red lines show the fits implemented in this analysis.

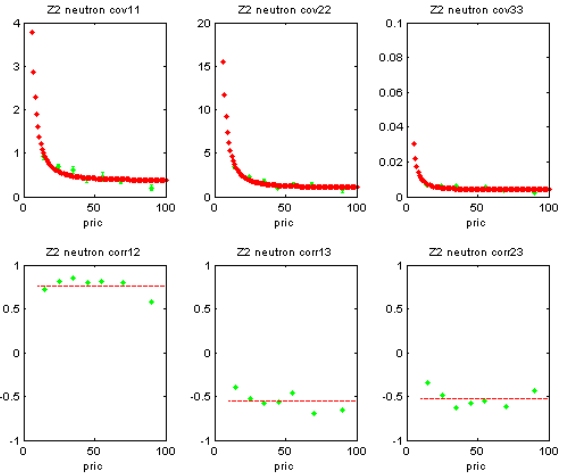


Figure B.5: Same as Figure B.4 for gold neutrons.

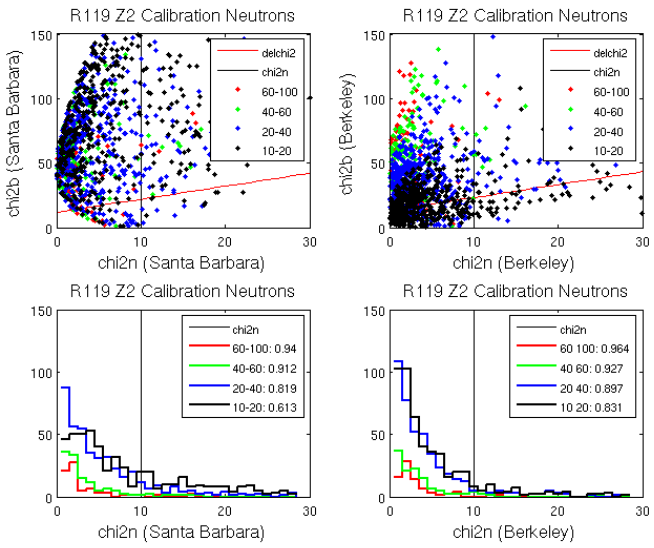


Figure B.6: Comparison between χ^2 parameters with ("Berkeley", *right*) and without ("Santa Barbara", *left*) energy fits for gold neutrons in detector T1Z2. The upper plots show the distributions of χ^2_n versus χ^2_b for neutrons of various energies (*colors*). also shown are the positions of the final cuts on neutron consistency (*black lines*) and surface discrimination (*red lines*), described below. The lower plots show the distributions of neutrons of various energies in χ^2_n. Lower legends show the consistency cut's passage fraction (fraction of neutrons with $\chi^2_n < 10$) as a function of neutron energy.

B.5.2 χ^2 Parameters

These covariance matrices allow us to compute two new RQs for each event: `chi2bedep_119` and `chi2nedep_119`, χ^2s with respect to the surface and neutron populations, respectively. The parameters themselves are computed with a custom MEX file, since the operation is difficult to "vector-ize" and the necessary loops are prohibitively slow in Matlab. The hope of this analysis was that these parameters will perform more consistently at low energies than those without an energy fit (`chi2b_119` and `chi2n_119`).

Figure B.6 compares the distributions of these parameters and those of the standard χ^2 analysis as functions of energy. The energy-fit parameters show less "spreading" at low energies than the standard parameters. In particular, the fraction of neutron events passing a "consistency cut" of the form $\chi^2_n < 10$ shows less variation with energy when implemented with the energy-fit parameters.

B.5.3 Defining the Cut

Our cut to identify nuclear recoil candidates has two parts:

- A **discrimination cut** which demands that $\chi^2_b - \chi^2_n > \kappa$. This is the analogue of a likelihood ratio test, enforcing that an event be more neutron-like than surface-event-like.
- A **consistency cut** which demands that $\chi^2_n < \lambda$. This cut asks that an event be neutron-like in absolute terms.

The consistency cut should pass the majority of neutron events, but should exclude events which are dissimilar from the calibration neutron sample. In this analysis the cut was set to $\chi^2_n < 10$, giving a neutron efficiency above 80%. Note that 90% of events should have $\chi^2 < 10$ in a three-variable gaussian distribution; the discrepancy is due to a combination of fitting errors and the non-gaussianity of the neutron calibration sample.

The discrimination cut is set to a common value $\kappa_{Ge,Si}$ for all detectors of a given substrate type, tuned on the "open" half of the Run 119 ^{133}Bs calibration data set. The cut values were chosen to pass a desired number of surface events in this calibration sample, chosen to give an expected leakage of ~ 1 event in the WIMP-search data set. In this case, we set $\kappa_{Ge} = 13.26$ (the position of the 11th surface event) to allow 10 total leakage events from 10-100 keV across all "good" Ge detectors (T1Z2, T1Z3, T1Z5, T2Z2, T2Z5) and set $\kappa_{Si} = 16.86$ to pass 13 events above 10 keV in the Si detectors (T1Z4, T1Z6, T2Z1, T2Z2, T2Z4, T2Z6).

B.5.4 Cut Performance

Signal efficiency and expected background

Figures B.7 and B.8 show the efficiency of the `cRTChi2Edep_119` cut on calibration neutrons as a function of energy for Ge and Si ZIPs, respectively. The large differences in efficiency between detectors have not been explained quantitatively, but this level of variation was common among the timing cuts in the Run 119 analysis.

Though all timing cuts for this analysis were set to have similar acceptance of surface events in the calibration data, they did not perform equally in the WIMP-search data set. The "edep" cut passed substantially fewer surface events than other timing analyses

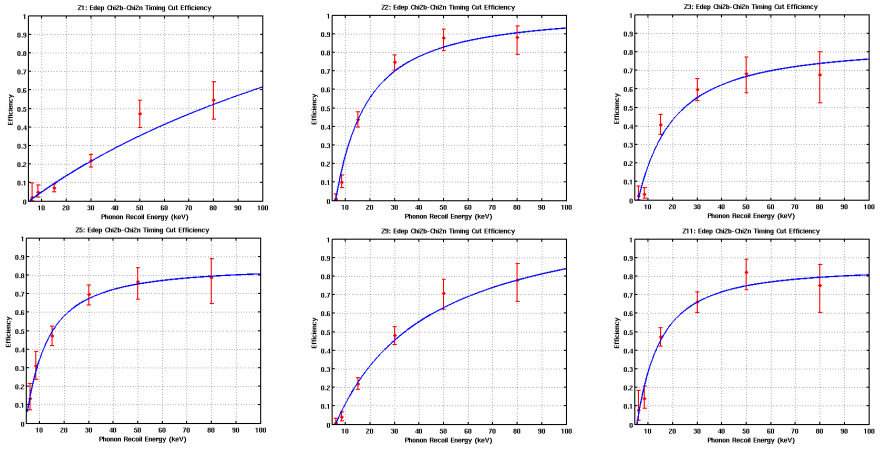


Figure B.7: Efficiency of `cRTChi2Edep_119` on calibration neutrons in Ge detectors passing other analysis cuts. Efficiencies computed and fit by Joel Sander.

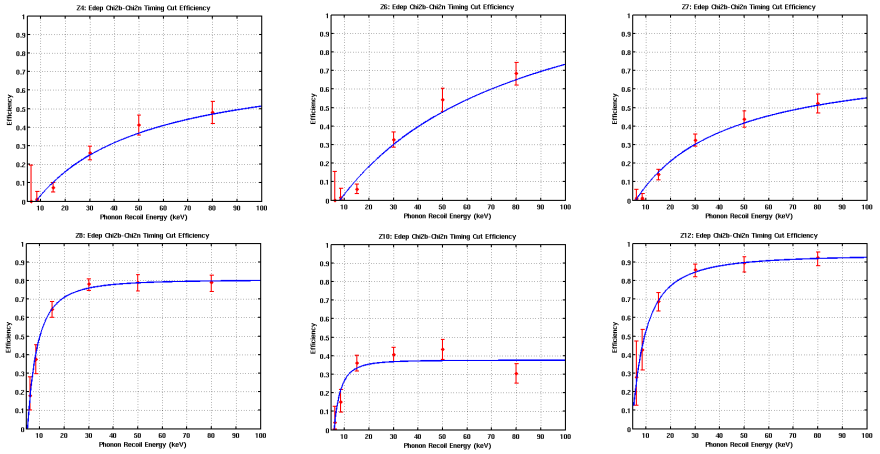


Figure B.8: Efficiency of `cRTChi2Edep_119` on calibration neutrons in Si detectors passing other analysis cuts. Efficiencies computed and fit by Joel Sander.

among the veto-anticoincident multiple-scatter events from the WIMP-search data sets: 5 Ge and 4 Si events from 10-100 keV, compared to 14 and 10 for the standard χ^2 analysis. Furthermore, these passing events were distributed fairly uniformly in recoil energy, while those passing the standard χ^2 analysis were concentrated at low energies (like a WIMP signal). Based on the performance of this cut on WIMP search multiple-scatter events and the expect singles/multiples ratio, we expected approximately 0.16 ± 0.07 (stat.) ± 0.23 (sys.) leakage events in Ge from 10-100 keV over the Run 119 exposure. We also expected 1.2 ± 0.6 (stat) ± 0.2 (syst) events between 7-100 keV in Si.

This analysis was found to have the best expected sensitivity for the Run 119 Si detectors, based upon the simple Monte Carlo study. A series of pseudo-experiments simulating the Run 119 analysis were thrown for each of the timing cuts, based upon each cut's measured efficiencies and an exponential fit to the spectrum of WIMP-search multiple-scatter events passing that cut. The analyses were judged based upon the median limit curve expected from these pseudo-experiments in the absence of a signal, and this analysis performs well by this metric - its lower efficiency at high recoil energies is compensated for by its leakage performance at low energies. On that basis, the limit curves of this analysis applied to the silicon ZIPs were chosen for presentation in [3, 5].

WIMP Search Data

When applied to the Run 119 WIMP-search data set, two events (at `pric` = 8.7 and 10.5 keV) in the Ge detectors passed all cuts in this analysis. Both were in the same detector (T2Z5) and data set (140616_1450). After unblinding, this data set was discovered to have occurred in a period of poor detector neutralization. These events are therefore not interpreted as true WIMP candidates, though they would be included in any WIMP limit if we had published one from this analysis.

No events passed all cuts in the Si detectors in this energy range. This zero-background result was used to set the WIMP limit from Si detectors in the R119 publications.

Appendix C

Bayesian Background Estimate

Section 7.2 described a model for estimating the expected surface event background for this analysis. Equations 7.1, 7.2, 7.4, 7.11, and 7.12 relate the expected number of events within the signal region to the rates of various classes of events outside the signal region. With a leakage model in place, we are left with a statistical problem: how to compute the expected total leakage and associated error bar. The error bar is the primary challenge here: no events pass the timing cut on most detectors in each of these event classes, calling into question the application of many statistical techniques based on asymptotic assumptions (e.g. profile likelihood intervals [212]). This framework also contains numerous nuisance parameters (*e.g.* the r_i's, which are not of interest themselves) and correlations between the detectors, all of which complicate a computation using classical statistics.

For these reasons, I follow a Bayesian statistical framework for this calculation. This Appendix outlines the statistical technique used in this analysis. Its results have been given in Section 7.2. This discussion follows the computation for the Ge detectors, but the Si estimate uses the same techniques.

C.1 Statistical background

C.1.1 Frequentist estimation

Most problems of parameter estimation in physics are handled using so-called frequentist (classical) statistical techniques [252]. Classical statistics conceives of probability in terms of frequencies of random processes. The most common tool for classical parameter

estimation is maximum likelihood analysis, a standard technique based upon the likelihood function,

$$\mathcal{L}(\theta) = P(x|\theta). \tag{C.1}$$

The likelihood function represents the probability of observing data set x given model parameter(s) θ. Our parameter estimate $\hat{\theta}$ is that choice of parameters which maximizes the likelihood of the observed data, *i.e.* the parameter values which render the date least "surprising." Parameter error bars take the form of "confidence intervals": parameter intervals constructed to contain the true parameter value some percentage of the time if the construction were repeated many times.

Maximum likelihood estimation is the most commonly-used scheme for physics data analysis, but it has two notable limitations for our purposes:

1. There is no standard way to deal with nuisance parameters: model parameters which are not part of our final estimate. Nuisance parameters greatly complicate the computation of confidence intervals, and many of the relevant schemes rely heavily on asymptotic approximations.

2. Some of the useful properties of maximum likelihood estimates (*e.g.* that they are unbiased, efficient and normally-distributed) are only asymptotically true themselves. This can cause issues at low statistics, especially for estimates near boundaries.

Unfortunately, our actual situation for leakage estimation involves numerous nuisance parameters (the leakage rates on the individual detectors) and is expected to be fairly far from the asymptotic regime. Statistics are generally very low on any individual detector. Many detectors show no passing events at all, leading to estimates on the boundary of the allowed region.

C.1.2 Bayesian estimation

An alternative scheme for parameter estimation comes from Bayesian statistics [253]. A Bayesian statistician treats probability as an extension of Boolean logic, reflecting a practitioner's degree of belief in a given proposition. Bayesian statistics agrees with frequentist reasoning in cases where random processes are involved, but its alternate foundations allow probabilities to be assigned to a wider variety of propositions. Bayesian reasoning makes no assumptions about asymptotic behavior, but its conclusions depend

to some extent upon the assumptions one puts in (the "priors"). Under weak assumptions, Bayesian estimators asymptotically satisfy many of the same properties as maximum likelihood estimates.

Bayes' theorem

A Bayesian calculation centers around Bayes' theorem, which describes how to update one's existing degree of belief in a proposition (the "prior") based upon the results of an experiment. For our purposes, we can write Bayes' theorem as follows:

$$P(\theta|x, I) = P(x|\theta, I)\frac{P(\theta|I)}{P(x|I)}. \tag{C.2}$$

- x is the observed data, which may consist of several measurements.
- I represents all of the information and assumptions we had going into the problem, i.e. our "prior knowlege". This may come from previous measurements or from vaguer preconceptions.
- $P(x|\theta, I)$ is the likelihood distribution, giving the probability of the observed data given the model.
- $P(\theta|x, I)$ is the posterior distribution, describing our degree of belief in the proposition given the data and our prior knowledge.
- $P(\theta|I)$ is the prior distribution, encoding the effects of our assumptions and prior information about the problem on the model parameters.
- $P(x|I)$ is essentially a normalization factor, ensuring that the posterior sums (or integrates) to unity.

We will usually suppress I from our discussions (as most authors do), but I include it here to illustrate that the prior is not qualitatively different from the other terms in this expression - it's just another conditional probability.

The primary controversy about Bayesian methods concerns the choice of prior distribution, which strikes many as inappropriately subjective. If we have a significant amount of data then a "reasonable" prior should make essentially no difference to a calculation's outcome. If we are very low on statistics or are using a particularly narrow prior, however, our choice of prior information will influence the result significantly (as it must). Various

forms of "uninformative" prior have been suggested to express a complete lack of knowledge and objectivity, but none are perfect. Ideally our priors should reflect our actual beliefs, but care must be taken that they do not take unwarranted control of our conclusions. For this analysis I attempt to sidestep the frequentist-Bayesian controversy by choosing the prior distributions in accordance with frequentist principles; this process is described further below.

Parameter estimates

A Bayesian parameter estimation problem produces the posterior probability distribution $P(\theta|x, I)$, expressing the practitioner's updated degree of belief in each choice of parameters. If we wish to extract a single parameter estimate from this function, we can choose its median or mean. The analogue of a confidence interval or error bar is a "credibility interval": a parameter volume containing a given fraction of the total probability. In more precise terms, a p-credibility interval is a volume V of parameter space satisfying $\int_V d\theta\, P(\theta|x, I) = p$. Like confidence intervals, credibility intervals are not uniquely defined.

One useful property of Bayesian reasoning is that it makes the inclusion of nuisance parameters essentially trivial. We simply integrate our posterior over the nuisance parameters to get the marginal distribution for the parameter(s) we want. If we have a nuisance parameter μ, for example, we proceed as follows:

$$P(\theta|x, I) = \int d\mu\, P(\theta, \mu|x, I). \tag{C.3}$$

In practice, Bayesian estimation thus amounts to a high-dimensional integration problem. Such integrals are difficult (or impossible) to perform analytically, but are often numerically tractable on modern computers.

A very useful concept in Bayesian methods is that of the "conjugate prior," a parameterized family of priors associated with a given likelihood function. If we choose our prior from the family of conjugate priors appropriate to our likelihood function, then the posterior will also be a member of that family (possibly with different parameters). Because of this great computational convenience, one generally chooses a prior distribution from among the family of conjugate priors whenever possible.

C.2 Useful Bayesian estimates

In this Section I outline a few useful Bayesian parameter estimation schemes. We will make use of these methods and techniques in the following Section.

C.2.1 Poisson distribution

Equation 7.1 and the various expressions for the pass/fail ratios r_i are functions of the expected counts of various event populations: $\langle b_i \rangle$, $\langle B_i \rangle$, etc. For each of these quantities we have a single observed count of events, from which we must estimate the expectation value. The core of this scheme is thus the Bayesian estimate of the rate parameter of a Poisson distribution. Given an observed count of n events from a Poisson distribution with expectation value ρ, the likelihood function is

$$P(n|\rho) = \frac{\rho^n e^{-\rho}}{n!}. \tag{C.4}$$

To compute a posterior distribution we need a choice of prior $P(\rho) \equiv P(\rho|I)$. To start with, we can seek objectivity through the use of "uninformative" (or at least minimally-informative) priors. One interesting class of priors are those of the form $P(\rho) \propto \rho^c$. We are free to choose other forms, of course, but any prior outside of this class seems relatively "informative." All members of this class are "improper priors": they cannot be normalized to have a unit integral. This isn't a problem as long as the resulting posterior distributions *can* be normalized. The scale-invariant choice $c = -1$ (the "Jaynes prior") is particularly appealing, since it is independent of our choice of time units for measuring the Poisson process. Steve Yellin has also argued in favor of this choice based upon ensuring the invariance of estimates when partitioning a Poisson time stream into several shorter streams.

If we use a prior from this class and observe n counts from a Poisson process, the resulting normalized posterior for the expected rate $\rho = \langle n \rangle$ is:

$$P(\rho|n,c) = \frac{P(n|\rho)P(\rho|c)}{\int d\rho\, P(n|\rho)P(\rho|c)} = \frac{\rho^{n+c}e^{-\rho}}{\Gamma(n+c+1)}. \tag{C.5}$$

Note that this posterior can only be normalized for all n if $c > -1$. The value $c = -1$ itself leads to a posterior that cannot be normalized in the case of zero observed events.

Taking the expectation value of the distribution from Equation C.5, we find

$$\langle \rho \rangle_c = \frac{\Gamma(n+c+2)}{\Gamma(n+c+1)} = n + c + 1. \tag{C.6}$$

This illustrates the practical meaning of c: higher values of c amount to greater expectation values for a given number of observed counts. The Jaynes prior is the most obviously unbiased choice, since then $\langle\rho\rangle = n\ \forall n$. Unfortunately, it achieves this by making a fairly strong assumption: asserting a certainty of zero count rate for zero observed counts. Higher values of c maintain normalization and avoid this effect at the price of increased apparent bias. For now, we show the effect of a range of choices of prior ($-1 < c < -1/2$) and make our selection later.

The conjugate prior of the Poisson distribution is the gamma distribution:

$$P_\gamma(x|\alpha,\theta) = \frac{x^{\alpha-1}e^{-x/\theta}}{\theta^\alpha\Gamma(\alpha)}. \tag{C.7}$$

Adding a single observation of N counts takes this gamma distribution to a new one, with $\alpha \to \alpha + N$ and $1/\theta \to 1/\theta + 1$. The x^c family of priors is a limiting form of the gamma distribution as $\theta \to \infty$ ($\alpha \to c+1$). This enables rapid sampling of Poisson posteriors using standard random number generators (e.g. Matlab's `gamrnd`).

C.2.2 Binomial distribution

Our leakage estimate also makes use of the binomial distribution in the definition of the phonon-face fractions Φ_i. Its likelihood function is

$$P(x|N,s) = \binom{N}{x} s^x(1-s)^{N-x}, \tag{C.8}$$

where s is the probability of "success" and we have observed x successes in N total trials. The conjugate prior of the binomial distribution is the beta distribution:

$$P_\beta(x|\alpha,\beta) = x^{\alpha-1}(1-x)^{\beta-1}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}. \tag{C.9}$$

Random numbers from this distribution are also easily thrown in Matlab using the `betarnd` function. An observation of x successes in N trials takes $\alpha \to \alpha + x$, $\beta \to \beta + (N - x)$.

There are at least two common choices of "uninformative" prior. The equivalent of the Jaynes prior (i.e. based on scale invariance) is $P(s) \propto s^{-1}(1-s)^{-1}$, or $\alpha, \beta \to 0$. We may also consider the constant prior $P(s) = 1$, corresponding to $\alpha = \beta = 0$. Note that the Jaynes prior favors $s = 0$ or $s = 1$ after a single observation and leads to an improper posterior in this case. The constant prior corresponds to the posterior from a Jaynes prior after observing a single success and a single failure.

C.3 Posterior distributions

Using the formalism developed above, we determine posterior distributions from the various background estimation formulas using a simple Monte Carlo simulation. The general scheme is to throw a large number of trials which follow the posterior distribution for the total expected background $\langle n \rangle$. It is not necessary, however, to draw these trials from the complicated posterior for $\langle n \rangle$ itself – it is completely equivalent to draw from the simpler posteriors for the component Poisson and binomial quantities. The distribution of the total background is identical in both cases, but many throws from a simple distribution can be orders of magnitude faster than one throw from a complicated distribution.

More explicitly, we proceed as follows:

1. Using the observed counts N_i, b_i, B_i , etc. and assumed priors, compute posterior distributions for the corresponding Poisson parameters: $\langle N_i \rangle$, $\langle b_i \rangle$, $\langle B_i \rangle$, etc. These posterior distributions take the form of gamma distributions and express our degree of belief in various values of the Poisson means. We treat the binomial parameters Φ_i and ϕ_i similarly, computing posterior beta distributions for them.

2. Throw some large number T of Monte Carlo trials based on the posteriors for these individual parameters. These trials map out the posteriors according to the information from the experiment (and the priors). This is easily done in Matlab using `gamrnd` and `betarnd`.

3. For each trial giving values of $\langle N_i \rangle$, etc. for each detector, we compute a trial value of $\langle n \rangle$ according to Equation 7.2 and its brethren. Note that this merely involves a couple of fast vector arithmetic operations. These trials will follow the same distribution we would have obtained by solving for for $P(\langle n \rangle | data, I)$, but are much faster to compute.

4. Finally, we take our parameter estimate $\langle \hat{n} \rangle$ to be the median of the trials of $\langle n \rangle$. The ends of the 1σ error bars are taken to be at the 15.86% and 84.14% points along the cumulative distribution this quantity. We could also have chosen our estimate as the mean of the posterior, but the median is invariant under parameter transformations and is less affected by long tails.

This whole scheme amounts to generating a large number of gamma- and beta-distributed random numbers and performing a few vectorized operations upon them.

This points to a general strategy for this sort of problem: avoid complicated integrals and throw your Monte Carlo at the earliest, simplest stage possible. The quality of the resulting approximation to the posterior will still improve as $1/\sqrt{N}$ for N complete trials, regardless of whether we draw from the simple initial distributions or the more complex later ones. Note also that the dimension of the problem (i.e. the number of detectors) only matters insofar as it makes each trial harder to generate, so the computation of the posterior scales only linearly with dimension.

Method 1 Figure C.1 shows the sampled posterior distributions for the Method 1 leakage (Equation 7.2) as determined by this scheme. Each full Monte Carlo throw includes 36 random numbers for each. We plot results for several choices of prior exponent c, with some *a priori* preference for the more negative values. The leakage posteriors are affected significantly by the choice of prior, as expected for such low statistics, but a limit is approached as $c \to -1$.

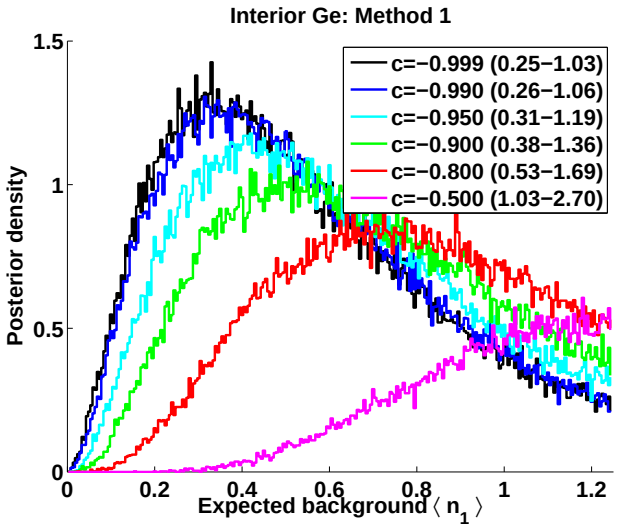


Figure C.1: Comparison of posterior distributions for the internal Ge detectors using "Method 1". The histograms represent posterior PDFs from 100,000 trials each, computed with a variety of Poisson priors c. The legend shows credible intervals for $\pm 1\sigma$ ($\pm 34.13\%$) from the median.

Method 2 We can do much the same for Equation 7.4. Each detector trial we throw now contains a total of 65 random numbers: all of the Poisson quantities in Equation

7.4, one binomial throw for each Φ_i, and throws of four Poisson rates (two numerators and two denominators) with which to construct α_P and α_Q. The correlations introduced among detectors by α_P and α_Q introduces a new wrinkle into our analysis. This would require significant care in a frequentist computation but is easily handled in a Monte Carlo implementation: just use the same values of the correlated parameters for all detectors within a single trial. Figure C.2 shows the posterior distributions for the internal detectors using Method 2 for several test priors c.

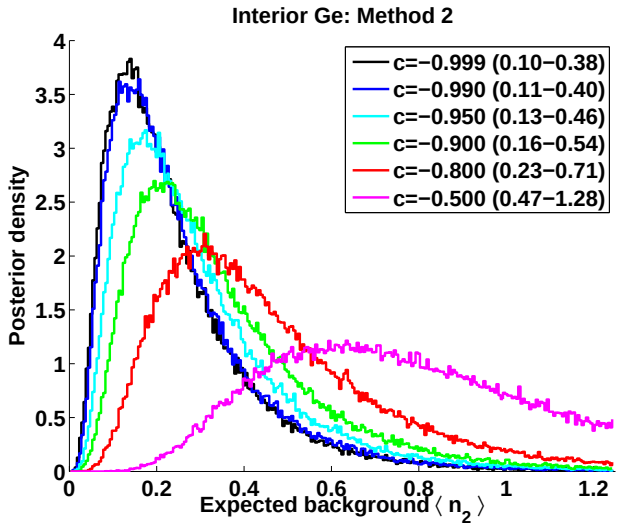


Figure C.2: Comparison of posterior distributions for the internal Ge detectors using "Method 2". The histograms represent posterior PDFs from 100,000 trials each, computed with a variety of Poisson priors c. The legend shows credible intervals for $\pm 1\sigma$ ($\pm 34.13\%$) from the median.

Endcaps We can apply the same methods trivially to the endcap detectors, using Equations 7.11 and 7.12. Note that we can easily generate the endcaps alongside the Monte Carlo for the interior detectors, thus automatically accounting for any correlation among their leakage estimates due to Φ_i, α_P or α_Q. Figure C.3 shows expected leakages from the three endcap detectors (T3Z6, T4Z6, and T5Z1) under our assortment of power-law priors.

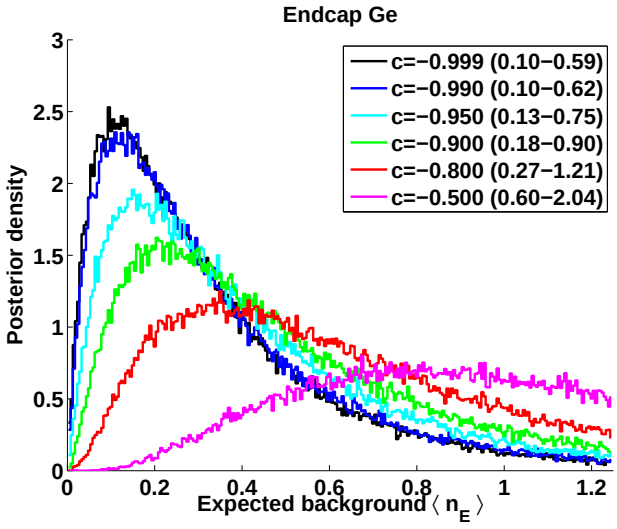


Figure C.3: Comparison of posterior distributions for the summed endcap Ge detectors using Equations 7.11 and 7.12. The histograms represent posterior PDFs from 100,000 trials each, computed with a variety of Poisson priors c. The legend shows credible intervals for $\pm 1\sigma$ ($\pm 34.13\%$) from the median.

C.4 Choice of priors

As shown in Figures C.1, C.2, and C.3, our choice of a value for c clearly affects our leakage estimate substantially. We don't have enough knowledge to make a truly informed choice of prior, but some choices would clearly bias our leakage estimate. To address this ambiguity, I choose these priors in accordance with classical statistical criteria. For each background estimation scheme, I choose a value of c which minimizes the mean bias of the corresponding estimate and ensures proper coverage for our error bars. Once these criteria are approximately met, any lingering bias or undercoverage can be added in as an additional systematic error.

I evaluate these classical attributes for each estimator through the use of a simple Monte Carlo:

1. Select values for the "true" expectation values of the various parameters in our model. These should be chosen close to the observed values, but with some variation to sample nearby parameter space.

2. Generate several (generally 25,000) mock data sets for our observed counts based upon these Poisson parameters.

3. For each mock data set, apply the background estimator to compute the leakage estimate and 1σ credible interval for each of several values of the prior parameter c.

4. Use these estimates to evaluate the bias (mean difference between the median inferred leakage and the true leakage) and coverage (percentage of mock sets for which the true parameter lies within the credibility interval) of our estimates.

5. Choose the value of c which approximately minimizes the bias. If the coverage of the method is low or if the bias varies substantially between several nearby "true" expectation values, add a systematic error to account for the bias or under-coverage.

Method 1 For Method 1, this Monte Carlo includes three choices of "true" Poisson expectation values which are close to the observed values:

1. All detectors assumed identical, all parameters set to their mean values across the detectors. This gives $\langle n \rangle = 0.41$, a bit lower than the estimate of 0.58 from the actual data at $c \sim -0.97$. The bias vanishes at $c = -0.97$, where the 1σ credible interval has 67% coverage.

2. All detectors assumed identical, with $\langle b_i \rangle$ tweaked upward to give a total leakage estimate of $\langle n \rangle = 0.58$. Again, the bias vanishes near $c = -0.97$ where the credible interval has $\sim 67\%$ coverage.

3. Same as above, but with the total background of $\langle n \rangle = 0.58$ lumped onto one detector. All the rest have $\langle b_i \rangle = 0$. Here the bias vanishes near $c = -0.96$ where the credible interval has 66% coverage.

The associated leakage and coverage curves are shown in Figure C.4. We take $c = -0.97$ as our bias-minimizing value, with a systematic error of ± 0.02 event to account for variation among these three choices.

Method 2 We do much the same for Method 2 with a somewhat different set of choices:

1. All detectors assumed identical, with all parameters set to their mean values across the detectors. This gives $\langle n \rangle = 0.32$.

2. All detectors assumed identical, with $\langle m_{P,i} \rangle$ and $\langle m_{Q,i} \rangle$ tripled to give $\langle n \rangle = 0.96$.

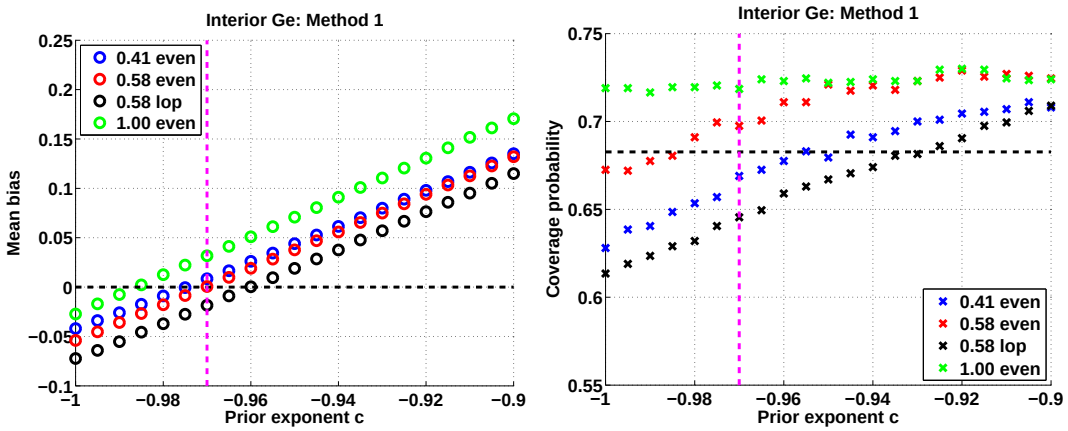


Figure C.4: *Left*: Median bias (median parameter estimate - true value) as a function of the prior exponent c for each of the Method 1 Ge trial cases described in the text. *Right*: Coverage (fraction of $\pm 1\sigma$ credibility intervals containing the true value) as a function of the prior exponent c for each of the three trial cases described in the text. In both plots, the horizontal dotted line indicates the metric's target value (no bias, exact coverage) and the vertical dotted line indicates our choice of c.

3. All leakage in one detector: same as #1, but with $12\times$ $\langle m_{P,i} \rangle$ and $\langle m_{Q,i} \rangle$ in one detector and zero in rest. This gives $\langle n \rangle = 0.32$.

4. All detectors assumed identical as in #1, but allowing Φ_i to vary between detectors to the observed values. This gives $\langle n \rangle = 0.30$.

The associated leakage and coverage curves are shown in Figure C.5. We take $c = -0.98$ as our bias-minimizing value, with a systematic error of ± 0.05 event to account for variation among these three choices.

Endcaps The procedure is much the same for the endcaps, but due to the small number of detectors we assign the various test cases somewhat differently:

1. All expected counts are taken to equal the observed counts, but we assign an expectation of 0.25 events in all cases where zero were observed. This increases the contributions of T3Z6 and T5Z1. Total leakage for this case is 0.71 events.

2. All expected counts are taken to equal the observed counts, but we assign an expectation of 0.1 events in all cases where zero were observed. Total leakage for this case is 0.50 events.

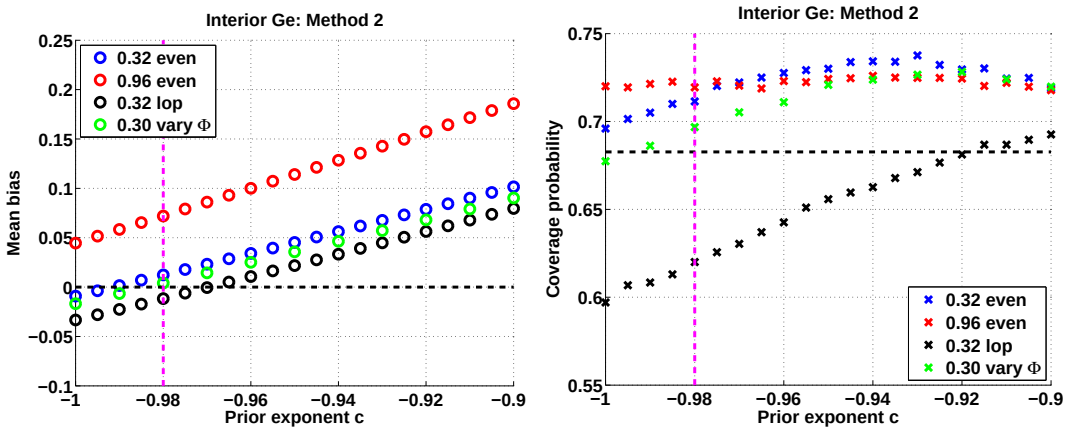


Figure C.5: Same as Figure C.4 but for the Method 2 estimate of the interior Ge detectors.

3. All expected counts are taken to equal the observed counts, but we halve the contribution of T3Z6 ($2 \rightarrow 1$) and assign an expectation of 0.03 events in all cases where zero were observed. Total leakage for this case is 0.22 events.

The associated leakage and coverage curves are shown in Figure C.6. We take $c = -0.95$ as our bias-minimizing value, with a systematic error of ± 0.05 event to account for variation among these three choices.

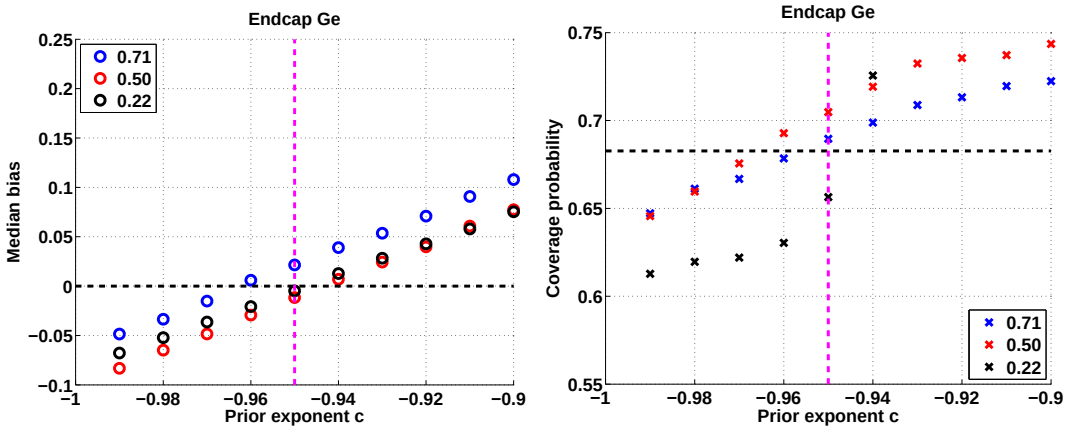


Figure C.6: Same as Figure C.4 but for the summed endcap Ge detectors. In this case I was forced to use the median bias rather than the mean, due to occasional division-by-zero issues.

C.5 Combined estimate

C.5.1 Internal Ge detectors

For the internal Ge detectors, we now have two different posterior distributions for r_i (Figures C.1 and C.2) that we would like to combine into a single distribution before computing the total leakage $\langle n_{int\ Ge} \rangle$. From the Bayesian point of view, the ideal solution is to take the posterior of one method as the prior for the next. This satisfies the expectation that each of these estimates represents a portion of a single chain of reasoning, but in practice would increase the problem's numerical complexity enormously.

In the case of uncorrelated estimates and consistent priors, a simple alternative is to multiply the resulting posterior distributions to produce a joint posterior. The subtlety here lies in the consistency of the priors: identical priors for the raw Poisson parameters may not correspond to identical states of knowledge about the leakages themselves. I press on in this direction nonetheless, trusting that this is unlikely to be a dominant effect upon our estimate or its error bars at the level of accuracy we require.

I cannot simply multiply the two posteriors for $\langle n^{(1)}_{int\ Ge} \rangle$ and $\langle n^{(1)}_{int\ Ge} \rangle$ because they are correlated by the common quantities $\langle N_i \rangle$. Instead, I follow a modified version of a Monte Carlo procedure suggested by Steve Yellin for efficiently combining such correlated estimates:

1. Throw a set of T iterations of the full set of correlating variables in accordance with their posterior probability distributions. In this case the correlating variables are the 12 values $\langle N_i \rangle$ and I take $T = 10^5$, so this is a 12×10^5 matrix of values. Each is thrown from the appropriate posterior (Equation C.5) for $c = -0.97$ and the observed event counts N_i.

2. For each of these T iterations of the correlating quantities, compute the posterior distributions for $\langle n^{(1,2)}_{int\ Ge} \rangle$. Multiply these and normalize to get the joint posterior distribution appropriate to that choice of correlating variables. This multiplication is acceptable because the two distributions are uncorrelated at *fixed* values of the correlating quantities. We now have a combined posterior distribution for each of the T iterations of the correlating quantities.

3. Sum over this set of posteriors and divide by T to get a single joint posterior for $\langle n_{int\ Ge} \rangle$. This step incorporates the distributions for the correlating variables.

Note that any systematic error not shared by both methods – *i.e.* the systematics on $\alpha_{P,Q}$, but not those related to the difference between single- and multiple-scatter events – should be included *before* carrying out this Monte Carlo. This method was used to generate the combined interior Ge result given in Equation 7.10.

C.5.2 Total Ge estimate

The procedure above extends easily to the calculation of the total (interior+endcap) Ge background estimate. The interior and endcap estimates are relatively independent: only α_P and α_Q are included as correlating variables. After combining the interior and endcap estimates using this Monte Carlo technique, we obtain the total Ge result given in Equation 7.17.

Bibliography

[1] V. Mandic. *First results from the Cryogenic Dark Matter Search experiment at the deep site*. PhD thesis, University of California, Berkeley, 2004.

[2] V. Mandic *et al.* Study of the dead layer in germanium for the CDMS detectors. *Nucl. Instrum. Meth.*, A520:171–174, 2004.

[3] D. S. Akerib *et al.* Limits on spin-independent WIMP nucleon interactions from the two-tower run of the Cryogenic Dark Matter Search. *Phys. Rev. Lett.*, 96:011302, 2006.

[4] Z. Ahmed *et al.* Search for Weakly Interacting Massive Particles with the First Five-Tower Data from the Cryogenic Dark Matter Search. *Phys. Rev. Lett.*, 102:011301, 2008.

[5] D. S. Akerib *et al.* Limits on spin-dependent WIMP nucleon interactions from the Cryogenic Dark Matter Search. *Phys. Rev.*, D73:011102, 2006.

[6] E. W. Kolb and M. Turner. *The Early Universe*. Frontiers in Physics. Westview Press, 1990.

[7] P. Peebles. *Principles of Physical Cosmology*. Princeton University Press, 1993.

[8] T. Padmanabhan. *Structure Formation in the Universe*. Cambridge University Press, 1993.

[9] J. A. Peacock. *Cosmological Physics*. Cambridge University Press, 1999.

[10] S. Dodelson. *Modern Cosmology*. Academic Press, 2003.

[11] M. Kowalski *et al.* Improved Cosmological Constraints from New, Old, and Combined Supernova Data Sets. *Astrophys. J.*, 686(2):749–778, 2008.

[12] E. Komatsu *et al.* Five-Year Wilkinson Microwave Anisotropy Probe WMAP Observations:Cosmological Interpretation. arXiv:0803.0547, 2008.

[13] A. R. Liddle and D. H. Lyth. *Cosmological inflation and large-scale structure.* Cambridge University Press, 2000.

[14] G. W. Marcy and R. P. Butler. Detection of Extrasolar Giant Planets. *Annu. Rev. Astron. Astrophys.*, 36:57–98, 1998.

[15] Y. Sofue and V. Rubin. Rotation Curves of Spiral Galaxies. *Annu. Rev. Astron. Astrophys.*, 39:137–174, 2001.

[16] J. D. Simon and M. Geha. The Kinematics of the Ultra-Faint Milky Way Satellites: Solving the Missing Satellite Problem. *Astrophys. J.*, 670:313–331, 2007.

[17] F. Zwicky. Spectral displacement of extra galactic nebulae. *Helv. Phys. Acta*, 6:110–127, 1933.

[18] F. Zwicky. On the Masses of Nebulae and of Clusters of Nebulae. *Astrophys. J.*, 86:217–246, 1937.

[19] N. Werner *et al.* Detection of hot gas in the filament connecting the clusters of galaxies Abell 222 and Abell 223. *Astron. Astrophys.*, 482:L29–L33, May 2008.

[20] P. Schneider, J. Ehlers, and E. E. Falco. *Gravitational Lenses.* Springer-Verlag, 1992.

[21] R. Narayan and M. Bartelmann. Lectures on gravitational lensing. arXiv:astro-ph/9606001, 1996.

[22] J. P. Kneib *et al.* Dynamics of Abell 2218 from optical and near-IR imagery of arc(let)s and the ROSAT/HRI X-ray map. *Astron. Astrophys.*, 303:27–40, November 1995.

[23] J.-P. Kneib *et al.* Hubble Space Telescope Observations of the Lensing Cluster Abell 2218. *Astrophys. J.*, 471:643–+, November 1996.

[24] D. E. Johnston *et al.* Cross-correlation Weak Lensing of SDSS galaxy Clusters II: Cluster Density Profiles and the Mass–Richness Relation. arXiv:0709.1159, 2007.

[25] D. Clowe *et al.* A direct empirical proof of the existence of dark matter. *Astrophys. J.*, 648:L109–L113, 2006.

[26] M. Bradac *et al.* Revealing the properties of dark matter in the merging cluster MACSJ0025.4-1222. *Astrophys. J.*, 687(2):959–967, 2008.

[27] R. H. Cyburt, B. D. Fields, and K. A. Olive. Primordial Nucleosynthesis in Light of WMAP. *Phys. Lett.*, B567:227–234, 2003.

[28] K. A. Olive, G. Steigman, and T. P. Walker. Primordial nucleosynthesis: Theory and observations. *Phys. Rept.*, 333:389–407, 2000.

[29] C. Amsler *et al.* Review of particle physics. *Phys. Lett.*, B667:1, 2008.

[30] J. Bechtold. Quasar Absorption Lines. arXiv:astro-ph/0112521, 2001.

[31] W. Hu and S. Dodelson. Cosmic Microwave Background Anisotropies. *Ann. Rev. Astron. Astrophys.*, 40:171–216, 2002.

[32] W. Hu and M. J. White. A CMB Polarization Primer. *New Astron.*, 2:323, 1997.

[33] M. R. Nolta *et al.* Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Angular Power Spectra. arXiv:0803.0593, 2008.

[34] L. D. Bradley *et al.* Discovery of a Very Bright Strongly Lensed Galaxy Candidate at z ˜ 7.6. *Astrophys. J.*, 678:647–654, May 2008.

[35] C. Alcock *et al.* The MACHO project: Microlensing results from 5.7 years of LMC observations. *Astrophys. J.*, 542:281–307, 2000.

[36] P. Tisserand *et al.* Limits on the Macho Content of the Galactic Halo from the EROS-2 Survey of the Magellanic Clouds. *Astron. Astrophys.*, 469:387–404, 2007.

[37] R. D. Peccei and H. R. Quinn. CP Conservation in the Presence of Instantons. *Phys. Rev. Lett.*, 38:1440–1443, 1977.

[38] J. Preskill, M. B. Wise, and F. Wilczek. Cosmology of the invisible axion. *Phys. Lett.*, B120:127–132, 1983.

[39] P. Sikivie. Experimental tests of the *invisible* axion. *Phys. Rev. Lett.*, 51:1415, 1983.

[40] R. Bradley *et al.* Microwave cavity searches for dark-matter axions. *Rev. Mod. Phys.*, 75:777–817, 2003.

[41] L. D. Duffy *et al.* A high resolution search for dark-matter axions. *Phys. Rev.*, D74:012006, 2006.

[42] S. N. Ahmed *et al.* Measurement of the total active B-8 solar neutrino flux at the Sudbury Neutrino Observatory with enhanced neutral current sensitivity. *Phys. Rev. Lett.*, 92:181301, 2004.

[43] Y. Ashie *et al.* A measurement of atmospheric neutrino oscillation parameters by Super-Kamiokande I. *Phys. Rev.*, D71:112005, 2005.

[44] D. G. Michael *et al.* Observation of muon neutrino disappearance with the MINOS detectors and the NuMI neutrino beam. *Phys. Rev. Lett.*, 97:191801, 2006.

[45] T. Araki *et al.* Measurement of neutrino oscillation with KamLAND: Evidence of spectral distortion. *Phys. Rev. Lett.*, 94:081801, 2005.

[46] G. Jungman, M. Kamionkowski, and K. Griest. Supersymmetric dark matter. *Phys. Rept.*, 267:195–373, 1996.

[47] M. E. Peskin and D. V. Schroeder. *An Introduction to quantum field theory.* Addison-Wesley, 1995.

[48] J. L. Feng, A. Rajaraman, and F. Takayama. Superweakly-interacting massive particles. *Phys. Rev. Lett.*, 91:011302, 2003.

[49] J. L. Feng, S. Su, and F. Takayama. Supergravity with a gravitino LSP. *Phys. Rev.*, D70:075019, 2004.

[50] D. J. H. Chung, E. W. Kolb, and A. Riotto. Nonthermal supermassive dark matter. *Phys. Rev. Lett.*, 81:4048–4051, 1998.

[51] A. Kusenko and M. E. Shaposhnikov. Supersymmetric Q-balls as dark matter. *Phys. Lett.*, B418:46–54, 1998.

[52] A. Kusenko, L. C. Loveridge, and M. Shaposhnikov. Astrophysical bounds on supersymmetric dark-matter Q- balls. *JCAP*, 0508:011, 2005.

[53] M. Milgrom. A Modification of the Newtonian dynamics as a possible alternative to the hidden mass hypothesis. *Astrophys. J.*, 270:365–370, 1983.

[54] J. D. Bekenstein. Relativistic gravitation theory for the MOND paradigm. *Phys. Rev.*, D70:083509, 2004.

[55] G. W. Angus *et al.* On the law of gravity, the mass of neutrinos and the proof of dark matter. *Astrophys. J.*, 654:L13–L16, 2007.

[56] P. Gondolo *et al.* DarkSUSY: Computing supersymmetric dark matter properties numerically. *JCAP*, 0407:008, 2004.

[57] P. Gondolo *et al.* http://www.physto.se/ edsjo/darksusy.

[58] G. Belanger, F. Boudjema, A. Pukhov, and A. Semenov. Dark matter direct detection rate in a generic model with micrOMEGAs2.1. arXiv:0803.2360, 2008.

[59] G. Belanger, F. Boudjema, A. Pukhov, and A. Semenov. micrOMEGAs2.0: A program to calculate the relic density of dark matter in a generic model. *Comput. Phys. Commun.*, 176:367–382, 2007.

[60] J. F. Donoghue, E. Golowich, and B. R. Holstein. *Dynamics of the Standard Model.* Cambridge University Press, 1992.

[61] J. Wess and J. Bagger. *Supersymmetry and supergravity.* Princeton University Press, 1992.

[62] M. Drees, R. Godbole, and P. Roy. *Theory and phenomenology of sparticles: An account of four- dimensional N=1 supersymmetry in high energy physics.* World Scientific, 2004.

[63] H. Baer and X. Tata. *Weak scale supersymmetry: From superfields to scattering events.* Cambridge University Press, 2006.

[64] P. Binetruy. *Supersymmetry: Theory, experiment and cosmology.* Oxford University Press, 2006.

[65] H. Georgi and S. L. Glashow. Unity of All Elementary Particle Forces. *Phys. Rev. Lett.*, 32:438–441, 1974.

[66] M. Shiozawa *et al.* Search for proton decay via $p \to e^+\pi^0$ in a large water Cherenkov detector. *Phys. Rev. Lett.*, 81:3319–3323, 1998.

[67] Y. Hayato *et al.* Search for proton decay through $p \to \bar{\nu}K^+$ in a large water Cherenkov detector. *Phys. Rev. Lett.*, 83:1529–1533, 1999.

[68] K. Kobayashi *et al.* Search for nucleon decay via modes favored by supersymmetric grand unification models in Super- Kamiokande-I. *Phys. Rev.*, D72:052007, 2005.

[69] D. I. Kazakov. Beyond the standard model (in search of supersymmetry). arXiv:hep-ph/0012288, 2000.

[70] S. Dimopoulos and H. Georgi. Softly Broken Supersymmetry and SU(5). *Nucl. Phys.*, B193:150, 1981.

[71] S. Dawson. The MSSM and why it works. arXiv:hep-ph/9712464, 1997.

[72] G. L. Kane, C. F. Kolda, L. Roszkowski, and J. D. Wells. Study of constrained minimal supersymmetry. *Phys. Rev.*, D49:6173–6210, 1994.

[73] S. R. Coleman and E. J. Weinberg. Radiative Corrections as the Origin of Spontaneous Symmetry Breaking. *Phys. Rev.*, D7:1888–1910, 1973.

[74] E. A. Baltz and P. Gondolo. Markov chain Monte Carlo exploration of minimal supergravity with implications for dark matter. *JHEP*, 10:052, 2004.

[75] R. R. de Austri, R. Trotta, and L. Roszkowski. A Markov chain Monte Carlo analysis of the CMSSM. *JHEP*, 05, 2006.

[76] J. L. Feng, K. T. Matchev, and T. Moroi. Multi-TeV scalars are natural in minimal supergravity. *Phys. Rev. Lett.*, 84:2322–2325, 2000.

[77] J. L. Feng and F. Wilczek. Advantages and distinguishing features of focus point supersymmetry. *Phys. Lett.*, B631:170–176, 2005.

[78] N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali. The hierarchy problem and new dimensions at a millimeter. *Phys. Lett.*, B429:263–272, 1998.

[79] D. J. Kapner *et al.* Tests of the gravitational inverse-square law below the dark-energy length scale. *Phys. Rev. Lett.*, 98:021101, 2007.

[80] I. Antoniadis. A Possible new dimension at a few TeV. *Phys. Lett.*, B246:377–384, 1990.

[81] S. Arrenberg, L. Baudis, K. Kong, K. T. Matchev, and J. Yoo. Kaluza-Klein Dark Matter: Direct Detection vis-a-vis LHC. arXiv:0805.4210, 2008.

[82] L. Randall and R. Sundrum. A large mass hierarchy from a small extra dimension. *Phys. Rev. Lett.*, 83:3370–3373, 1999.

[83] K. Agashe and G. Servant. Warped unification, proton stability and dark matter. *Phys. Rev. Lett.*, 93:231805, 2004.

[84] M. Schmaltz and D. Tucker-Smith. Little Higgs review. *Ann. Rev. Nucl. Part. Sci.*, 55:229–270, 2005.

[85] M. Perelstein. Little Higgs models and their phenomenology. *Prog. Part. Nucl. Phys.*, 58:247–291, 2007.

[86] A. Birkedal, A. Noble, M. Perelstein, and A. Spray. Little Higgs dark matter. *Phys. Rev.*, D74:035002, 2006.

[87] E. A. Baltz, M. Battaglia, M. E. Peskin, and T. Wizansky. Determination of dark matter properties at high-energy colliders. *Phys. Rev.*, D74:103521, 2006.

[88] M. Battaglia, A. Datta, A. De Roeck, K. Kong, and K. T. Matchev. Contrasting supersymmetry and universal extra dimensions at the CLIC multi-TeV e+ e- collider. *JHEP*, 07:033, 2005.

[89] D. R. Tovey. Inclusive SUSY searches and measurements at ATLAS. *Eur. Phys. J. direct*, C4:N4, 2002.

[90] G. Weiglein *et al.* Physics interplay of the LHC and the ILC. *Phys. Rept.*, 426:47–358, 2006.

[91] D. Hooper and T. Plehn. Supersymmetric dark matter - how light can the LSP be? *Phys. Lett.*, B562:18–27, 2003.

[92] A. Bottino, F. Donato, N. Fornengo, and S. Scopel. Zooming in on light relic neutralinos by direct detection and measurements of galactic antimatter. *Phys. Rev.*, D77:015002, 2008.

[93] S. Profumo. Hunting the lightest lightest neutralinos. *Phys. Rev.*, D78:023507, 2008.

[94] G. Bertone, D. Hooper, and J. Silk. Particle dark matter: Evidence, candidates and constraints. *Phys. Rept.*, 405:279–390, 2005.

[95] C. Tyler. Particle dark matter constraints from the Draco dwarf galaxy. *Phys. Rev. D*, 66(2):023509, Jul 2002.

[96] L. Bergstrom, J. Edsjo, P. Gondolo, and P. Ullio. Clumpy neutralino dark matter. *Phys. Rev.*, D59:043506, 1999.

[97] J. Silk, K. A. Olive, and M. Srednicki. The photino, the sun, and high-energy neutrinos. *Phys. Rev. Lett.*, 55:257–259, 1985.

[98] L. M. Krauss, M. Srednicki, and F. Wilczek. Solar System Constraints and Signatures for Dark Matter Candidates. *Phys. Rev.*, D33:2079–2083, 1986.

[99] N. Arkani-Hamed, D. P. Finkbeiner, T. Slatyer, and N. Weiner. A Theory of Dark Matter. arXiv:0810.0713, 2008.

[100] J. F. Navarro, C. S. Frenk, and S. D. M. White. A Universal Density Profile from Hierarchical Clustering. *Astrophys. J.*, 490:493–508, 1997.

[101] B. Moore, F. Governato, T. R. Quinn, J. Stadel, and G. Lake. Resolving the Structure of Cold Dark Matter Halos. *Astrophys. J.*, 499:L5, 1998.

[102] J. D. Simon, A. D. Bolatto, A. Leroy, and L. Blitz. High-Resolution Measurements of the Dark Matter Halo of NGC 2976: Evidence for a Shallow Density Profile. *Astrophys. J.*, 596:957–981, 2003.

[103] J. Diemand *et al.* Clumps and streams in the local dark matter distribution. *Nature*, 454:735–738, 2008.

[104] M. Kuhlen, J. Diemand, and P. Madau. The Dark Matter Annihilation Signal from Galactic Substructure: Predictions for GLAST. *Astrophys. J.*, 686(1):262–278, October 2008.

[105] L. Bergstrom, P. Ullio, and J. H. Buckley. Observability of gamma rays from dark matter neutralino annihilations in the Milky Way halo. *Astropart. Phys.*, 9:137–162, 1998.

[106] L. Bergstrom, J. Edsjo, and P. Ullio. Spectral gamma-ray signatures of cosmological dark matter annihilations. *Phys. Rev. Lett.*, 87:251301, 2001.

[107] L. Bergstrom and J. Kaplan. Gamma-ray lines from TeV dark matter. *Astropart. Phys.*, 2:261–268, 1994.

[108] L. Bergstrom, T. Bringmann, M. Eriksson, and M. Gustafsson. Two photon annihilation of Kaluza-Klein dark matter. *JCAP*, 0504:004, 2005.

[109] L. Bergstrom and P. Ullio. Full one-loop calculation of neutralino annihilation into two photons. *Nucl. Phys.*, B504:27–44, 1997.

[110] P. Ullio and L. Bergstrom. Neutralino annihilation into a photon and a Z boson. *Phys. Rev.*, D57:1962–1971, 1998.

[111] T. Bringmann, L. Bergstrom, and J. Edsjo. New Gamma-Ray Contributions to Supersymmetric Dark Matter Annihilation. *JHEP*, 01:049, 2008.

[112] F. A. Aharonian and C. W. Akerlof. Gamma-ray astronomy with imaging atmospheric cerenkov telescopes. *Ann. Rev. Nucl. Part. Sci.*, 47:273–314, 1997.

[113] F. Aharonian *et al.* HESS observations of the galactic center region and their possible dark matter interpretation. *Phys. Rev. Lett.*, 97:221102, 2006.

[114] W. de Boer, C. Sander, V. Zhukov, A. V. Gladyshev, and D. I. Kazakov. EGRET excess of diffuse galactic gamma rays interpreted as a signal of dark matter annihilation. *Phys. Rev. Lett.*, 95:209001, 2005.

[115] L. Bergstrom, J. Edsjo, M. Gustafsson, and P. Salati. Is the dark matter interpretation of the EGRET gamma excess compatible with antiproton measurements? *JCAP*, 0605:006, 2006.

[116] F. W. Stecker, S. D. Hunter, and D. A. Kniffen. The Likely Cause of the EGRET GeV Anomaly and its Implications. *Astropart. Phys.*, 29:25–29, 2008.

[117] D. Elsaesser and K. Mannheim. Supersymmetric dark matter and the extragalactic gamma ray background. *Phys. Rev. Lett.*, 94:171302, 2005.

[118] A. W. Strong, I. V. Moskalenko, and O. Reimer. A new determination of the extragalactic diffuse gamma-ray background from EGRET data. *Astrophys. J.*, 613:956–961, 2004.

[119] P. F. Michelson. Large area telescope first light. NASA teleconference, August 26, 2008.

[120] E. A. Baltz *et al.* Pre-launch estimates for GLAST sensitivity to Dark Matter annihilation signals. *JCAP*, 0807:013, 2008.

[121] D. Hooper, G. Zaharijas, D. P. Finkbeiner, and G. Dobler. Prospects For Detecting Dark Matter With GLAST In Light Of The WMAP Haze. *Phys. Rev.*, D77:043511, 2008.

[122] J. Lundberg and J. Edsjo. WIMP diffusion in the solar system including solar depletion and its effect on earth capture rates. *Phys. Rev.*, D69:123505, 2004.

[123] D. Hubert. Neutralino searches with AMANDA and IceCube: Past, present and future. In C. de los Heros, editor, *Proceedings of the first workshop on exotic physics with neutrino telescopes, EPNT06*, 2007.

[124] M. Kamionkowski, K. Griest, G. Jungman, and B. Sadoulet. Model independent comparison of direct versus indirect detection of supersymmetric dark matter. *Phys. Rev. Lett.*, 74:5174–5177, 1995.

[125] S. Desai *et al.* Search for dark matter WIMPs using upward through-going muons in Super-Kamiokande. *Phys. Rev.*, D70:083523, 2004.

[126] O. Adriani *et al.* Observation of an anomalous positron abundance in the cosmic radiation. arXiv:0810.4995, 2008.

[127] O. Adriani *et al.* A new measurement of the antiproton-to-proton flux ratio up to 100 GeV in the cosmic radiation. arXiv:0810.4994, 2008.

[128] P. Picozza *et al.* PAMELA: A payload for antimatter matter exploration and light-nuclei astrophysics. *Astropart. Phys.*, 27:296–315, 2007.

[129] V. Barger, W. Y. Keung, D. Marfatia, and G. Shaughnessy. PAMELA and dark matter. arXiv:0809.0162, 2008.

[130] L. Bergstrom, T. Bringmann, and J. Edsjo. New Positron Spectral Features from Supersymmetric Dark Matter - a Way to Explain the PAMELA Data? arXiv:0808.3725, 2008.

[131] M. Cirelli and A. Strumia. Minimal Dark Matter predictions and the PAMELA positron excess. arXiv:0808.3867, 2008.

[132] F. Donato, N. Fornengo, D. Maurin, and P. Salati. Antiprotons in cosmic rays from neutralino annihilation. *Phys. Rev.*, D69:063501, 2004.

[133] F. Donato, N. Fornengo, and P. Salati. Antideuterons as a signature of supersymmetric dark matter. *Phys. Rev.*, D62:043003, 2000.

[134] F. Donato, N. Fornengo, and D. Maurin. Antideuteron fluxes from dark matter annihilation in diffusion models. *Phys. Rev.*, D78:043506, 2008.

[135] P. Jean *et al.* Early SPI/INTEGRAL measurements of galactic 511 keV line emission from positron annihilation. *Astron. Astrophys.*, 407:L55, 2003.

[136] C. Boehm, D. Hooper, J. Silk, M. Casse, and J. Paul. MeV dark matter: Has it been detected? *Phys. Rev. Lett.*, 92:101301, 2004.

[137] D. P. Finkbeiner. Microwave ISM Emission Observed by WMAP. *Astrophys. J.*, 614:186–193, 2004.

[138] D. P. Finkbeiner. WMAP microwave emission interpreted as dark matter annihilation in the inner Galaxy. arXiv:astro-ph/0409027, 2004.

[139] D. Hooper, D. P. Finkbeiner, and G. Dobler. Evidence Of Dark Matter Annihilations In The WMAP Haze. *Phys. Rev.*, D76:083012, 2007.

[140] M. W. Goodman and E. Witten. Detectability of certain dark-matter candidates. *Phys. Rev.*, D31:3059, 1985.

[141] R. J. Gaitskell. Direct detection of dark matter. *Ann. Rev. Nucl. Part. Sci.*, 54:315–359, 2004.

[142] A. Kurylov and M. Kamionkowski. Generalized analysis of weakly-interacting massive particle searches. *Phys. Rev.*, D69:063503, 2004.

[143] J. D. Lewin and P. F. Smith. Review of mathematics, numerical factors, and corrections for dark matter experiments based on elastic nuclear recoil. *Astropart. Phys.*, 6:87–112, 1996.

[144] G. Duda, A. Kemper, and P. Gondolo. Model independent form factors for spin independent neutralino nucleon scattering from elastic electron scattering data. *JCAP*, 0704:012, 2007.

[145] J. Ellis, K. A. Olive, and C. Savage. Hadronic Uncertainties in the Elastic Scattering of Supersymmetric Dark Matter. *Phys. Rev.*, D77:065026, 2008.

[146] D. R. Tovey, R. J. Gaitskell, P. Gondolo, Y. Ramachers, and L. Roszkowski. A new model-independent method for extracting spin- dependent cross section limits from dark matter searches. *Phys. Lett.*, B488:17–26, 2000.

[147] M. T. Ressell *et al.* Nuclear shell model calculations of neutralino - nucleus cross-sections for Si-29 and Ge-73. *Phys. Rev.*, D48:5519–5535, 1993.

[148] V. Dimitrov, J. Engel, and S. Pittel. Scattering of weakly interacting massive particles from Ge-73. *Phys. Rev.*, D51:291–295, 1995.

[149] J. Engel. Nuclear form-factors for the scattering of weakly interacting massive particles. *Phys. Lett.*, B264:114–119, 1991.

[150] V. A. Bednyakov and H. V. Klapdor-Kleingrothaus. Update of the direct detection of dark matter and the role of the nuclear spin. *Phys. Rev.*, D63:095005, 2001.

[151] G. Heusser. Low-radioactivity background techniques. *Ann. Rev. Nucl. Part. Sci.*, 45:543–590, 1995.

[152] A. K. Drukier, K. Freese, and D. N. Spergel. Detecting Cold Dark Matter Candidates. *Phys. Rev.*, D33:3495–3508, 1986.

[153] K. Freese, J. A. Frieman, and A. Gould. Signal Modulation in Cold Dark Matter Detection. *Phys. Rev.*, D37:3388, 1988.

[154] J. I. Collar and F. T. Avignone. Diurnal modulation effects in cold dark matter experiments. *Phys. Lett.*, B275:181–185, 1992.

[155] D. N. Spergel. The Motion of the Earth and the Detection of WIMPs. *Phys. Rev.*, D37:1353, 1988.

[156] H. Baer, A. Belyaev, T. Krupovnickas, and J. O'Farrill. Indirect, direct and collider detection of neutralino dark matter. *JCAP*, 0408:005, 2004.

[157] L. Roszkowski, R. Ruiz de Austri, and R. Trotta. Implications for the constrained MSSM from a new prediction for $b \to s\gamma$. *JHEP*, 07:075, 2007.

[158] J. Angle *et al.* First Results from the XENON10 Dark Matter Experiment at the Gran Sasso National Laboratory. *Phys. Rev. Lett.*, 100:021303, 2008.

[159] J. Angle *et al.* Limits on spin-dependent WIMP-nucleon cross-sections from the XENON10 experiment. *Phys. Rev. Lett.*, 101:091301, 2008.

[160] J. R. Primack, D. Seckel, and B. Sadoulet. Detection of Cosmic Dark Matter. *Ann. Rev. Nucl. Part. Sci.*, 38:751–807, 1988.

[161] J. L. Bourjaily and G. L. Kane. What is the cosmological significance of a discovery of wimps at colliders or in direct experiments? arXiv:hep-ph/0501262, 2005.

[162] M. Drees and C.-L. Shan. Model-Independent Determination of the WIMP Mass from Direct Dark Matter Detection Data. *JCAP*, 0806:012, 2008.

[163] D. Abrams *et al.* Exclusion limits on the WIMP nucleon cross-section from the cryogenic dark matter search. *Phys. Rev.*, D66:122003, 2002.

[164] G. F. Knoll. *Radiation detection and measurement.* John Wiley & Sons, 3rd edition, 2000.

[165] T. Shutt. *A dark matter detector based on the simultaneous measurement of phonons and ionization at 20 mK.* PhD thesis, University of California, Berkeley, 1993.

[166] S. Golwala. *Exclusion limits on the WIMP-nucleon elastic-scattering cross section from the Cryogenic Dark Matter Search.* PhD thesis, University of California, Berkeley, 2000.

[167] R. Alig and S. Bloom. Electron-hole-pair creation energies in semiconductors. *Phys. Rev. Lett.*, 35(22):1522–1525, 1975.

[168] T. Shutt *et al.* Simultaneous high resolution measurement of phonons and ionization created by particle interactions in a 60g germanium crystal at 25-mk. *Phys. Rev. Lett.*, 69:3531, 1992.

[169] T. Shutt *et al.* A solution to the dead-layer problem in ionization and phonon-based dark matter detectors. *Nucl. Instrum. Meth.*, A444:340–344, 2000.

[170] A. Broniatowski, B. Censier, A. Juillard, and L. Berge. Cryogenic germanium detectors for dark matter search: Surface events rejection by charge measurements. *Nucl. Instrum. Meth.*, A559:378–380, 2006.

[171] A. Phipps. *The CDMS Ionization Readout Capacitance Budget.* Senior honors thesis, University of California, Berkeley, 2008.

[172] B. Cabrera. Notes on electron-phonon scattering. CDMS internal note, 1993.

[173] B. S. Neganov and V. N. Trofimov. Possibility of producing a bulky supersensitive thermal detector at a temperature close to absolute zero. *J. Exp. Theor. Phys.*, 28:328, 1978.

[174] P. N. Luke. Voltage-assisted calorimetric ionization detector. *Journal of Applied Physics*, 64(12):6858–6860, 1988.

[175] H. Maris. Phonon propagation with isotope scattering and spontaneous anharmonic decay. *Physical Review B*, Jan 1990.

[176] H. Maris and S. Tamura. Anharmonic decay and the propagation of phonons in an isotopically pure crystal at low temperatures: .... *Physical Review B*, Jan 1993.

[177] C. Chang. *The Cryogenic Dark Matter Search (CDMS-II) - first results from the Soudan Mine.* PhD thesis, Stanford University, 2004.

[178] N. E. Booth. Quasiparticle trapping and the quasiparticle multiplier. *Applied Physics Letters*, 50(5):293–295, 1987.

[179] M. Tinkham. *Introduction to Superconductivity.* Dover, 2 edition, 1996.

[180] K. D. Irwin and G. C. Hilton. Transition-edge sensors. In C. Enss, editor, *Cryogenic Particle Detection*, volume 99 of *Topics in Applied Physics*, pages 63–149. Springer-Verlag, 2005.

[181] K. D. Irwin. An application of electrothermal feedback for high resolution cryogenic particle detection. *Applied Physics Letters*, 66(15):1998–2000, 1995.

[182] B. A. Young, T. Saab, B. Cabrera, A. J. Miller, P. L. Brink, and J. P. Castle. Effect of implanted metal impurities on superconducting tungsten films. *Journal of Applied Physics*, 91(10):6516–6519, 2002.

[183] B. A. Young *et al.* Measurement of T_c suppression in tungsten using magnetic impurities. *Journal of Applied Physics*, 86(12):6975–6978, 1999.

[184] J. Lindhard, M. S. V. Nielsen, and P. Thomsen. Integral equations governing radiation effects. *Mat. Fys. Medd. Dan. Vid. Selsk.*, 33(10):1–42, 1963.

[185] B. L. Dougherty. Measurements of ionization produced in silicon crystals by low-energy silicon atoms. *Phys. Rev.*, A45(3):2104–2107, 1992.

[186] T. Shutt *et al.* Measurment of ionization and phonon production by nuclear recoils in a 60g crystal of germanium at 25-mk. *Phys. Rev. Lett.*, 69:3425, 1992.

[187] D. S. Akerib *et al.* First results from the Cryogenic Dark Matter Search in the Soudan Underground Lab. *Phys. Rev. Lett.*, 93:211301, 2004.

[188] A. Reisetter. *Results from the Two-Tower Run of the Cryogenic Dark Matter Search.* PhD thesis, University of Minnesota, 2005.

[189] J. Sander. *Results from the Cryogenic Dark Matter Search Using a Chi Squared Analysis.* PhD thesis, University of California, Santa Barbara, 2007.

[190] M. Attisha. *Cryogenic Dark Matter Search (CDMS-II) - Application of Neural Networks and Wavelets to Event Analysis.* PhD thesis, Brown University, 2006.

[191] V. Mandic, B. Sadoulet, and R. W. Schnee. Maximum-likelihood approach for signal estimation in direct detection experiments for dark matter. *Nucl. Instrum. Meth.*, A553:459–469, 2005.

[192] M. S. Armel-Funkhouser *et al.* Exclusion limits on the WIMP nucleon cross-section from the first run of the Cryogenic Dark Matter Search in the Soudan underground lab. *Phys. Rev.*, D72:052009, 2005.

[193] D. S. Akerib *et al.* New results from the Cryogenic Dark Matter Search. *Phys. Rev.*, D68(082002), 2003.

[194] N. Mirabolfathi *et al.* Detector commissioning for the CDMS-II final run at the Soudan Underground Laboratory. *Nucl. Instrum. Meth.*, A559:417–419, 2006.

[195] T. Saab. *Search for weakly interacting massive particles with the Cryogenic Dark Matter Search experiment.* PhD thesis, Stanford University, 2002.

[196] I. Ambats *et al.* The MINOS Detectors Technical Design Report. NUMI-L-337, 1998.

[197] C. Kittel. *Introduction to Solid State Physics.* Wiley, 2004.

[198] R. W. Ogburn. *A Search for Particle Dark Matter Using Cryogenic Germanium and Silicon Detectors in the One- and Two-Tower Runs of CDMS-II at Soudan.* PhD thesis, Stanford University, 2008.

[199] A. Da Silva. *Development of a Low Background Environment for the Cryogenic Dark Matter Search.* PhD thesis, University of British Columbia, 1996.

[200] S. Kamat. *Extending the Sensitivity to the Detection of WIMP Dark Matter with an Improved Understanding of the Limiting Neutron Backgrounds.* PhD thesis, Case Western Reserve University, 2005.

[201] R. Hennings-Yeomans. *First 5-Tower WIMP-Search Results from the Cryogenic Dark Matter Search with Improved Understanding of Neutron Background and Benchmarking.* PhD thesis, Case Western Reserve University, 2008.

[202] The CDMS Collaboration. The Run 123/124 analysis. Internal analysis document, February 2008.

[203] S. Agostinelli *et al.* GEANT4: A simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.

[204] J. Allison *et al.* Geant4 developments and applications. *IEEE Trans. Nucl. Sci.*, 53:270, 2006.

[205] A. Fasso *et al.* The physics models of FLUKA: Status and recent development. arXiv:hep-ph/0306267, 2003.

[206] A. Fasso, A. Ferrari, J. Ranft, and P. R. Sala. FLUKA: A multi-particle transport code. *CERN-2005-10, INFN/TC_05/11, SLAC-R-773*, 2005.

[207] http://mcnpx.lanl.gov/.

[208] V. A. Kudryavtsev. Muon simulation codes MUSIC and MUSUN for underground physics. arXiv:0810.4635, 2008.

[209] D. Mei and A. Hime. Muon-induced background study for underground laboratories. *Phys. Rev.*, D73:053004, 2006.

[210] D. Leonard *et al.* Systematic study of trace radioactive impurities in candidate construction materials for EXO-200. *Nucl. Instrum. Meth. A*, 591(3):490–509, 2008.

[211] W. B. Wilson, R. T. Perry, W. S. Charlton, T. A. Parish, and E. F. Shores. SOURCES: a code for calculating (alpha,n), spontaneous fission, and delayed neutron sources and spectra. *Radiat Prot Dosimetry*, 115(1-4):117–121, 2005.

[212] W. A. Rolke, A. M. Lopez, and J. Conrad. Confidence Intervals with Frequentist Treatment of Statistical and Systematic Uncertainties. *Nucl. Instrum. Meth.*, A551:493–503, 2005.

[213] S. Yellin. Finding an upper limit in the presence of unknown background. *Phys. Rev.*, D66:032005, 2002.

[214] S. Yellin. Extending the optimum interval method. arXiv:0709.2701, 2007.

[215] G. J. Feldman and R. D. Cousins. A Unified approach to the classical statistical analysis of small signals. *Phys. Rev.*, D57:3873–3889, 1998.

[216] V. Sanglard *et al.* Final results of the EDELWEISS-I dark matter search with cryogenic heat-and-ionization Ge detectors. *Phys. Rev.*, D71:122002, 2005.

[217] P. Benetti *et al.* First results from a dark matter search with liquid argon at 87-K in the Gran Sasso underground laboratory. *Astropart. Phys.*, 28:495–507, 2008.

[218] V. Lebedenko *et al.* Result from the First Science Run of the ZEPLIN-III Dark Matter Search Experiment. arXiv:0812.1150, 2008.

[219] P. Sorensen *et al.* Determination of the scintillation and ionization yield of liquid Xe from the XENON10 experiment. arXiv:0807.0459, 2008.

[220] E. Aprile *et al.* New Measurement of the Relative Scintillation Efficiency of Xenon Nuclear Recoils Below 10 keV. arXiv:0810.0274, 2008.

[221] J. R. Ellis, K. A. Olive, Y. Santoso, and V. C. Spanos. Update on the direct detection of supersymmetric dark matter. *Phys. Rev.*, D71:095007, 2005.

[222] R. Bernabei *et al.* First results from DAMA/LIBRA and the combined results with DAMA/NaI. *Eur. Phys. J.*, C56:333–355, 2008.

[223] C. Savage, G. Gelmini, P. Gondolo, and K. Freese. Compatibility of DAMA/LIBRA dark matter detection with other searches. arXiv:0808.3607, 2008.

[224] E. Behnke *et al.* Improved Spin-Dependent WIMP Limits from a Bubble Chamber. *Science*, 319:933–936, 2008.

[225] A. T. Lee. SQUID readout multiplexers for transition-edge sensor arrays. *Nucl. Instrum. Meth. A*, 559(2):786–789, 2006.

[226] D. S. Gemmell. Channeling and related effects in the motion of charged particles through crystals. *Rev. Mod. Phys.*, 46:129–227, 1974.

[227] E. M. Drobyshevski. Channeling Effect and Improvement of the Efficiency of Charged Particle Registration with Crystal Scintillators. *Mod. Phys. Lett. A*, 23(36):2077–3085, 2008.

[228] P. L. Brink *et al.* First test runs of a dark-matter detector with interleaved ionization electrodes and phonon sensors for surface-event rejection. *Nucl. Instrum. Meth.*, A559:414–416, 2006.

[229] A. Broniatowski *et al.* Cryogenic Ge Detectors with Interleaved Electrodes: Design and Modeling. *Journal of Low Temperature Physics*, 151:830–834, May 2008.

[230] X. DeFay. *Ph.D. defense talk*. PhD thesis, Université de Montpellier, Montpellier, France, 2008.

[231] G. Angloher *et al.* Commissioning Run of the CRESST-II Dark Matter Search. arXiv:0809.1829, 2008.

[232] C. E. Aalseth *et al.* Experimental constraints on a dark matter origin for the DAMA annual modulation effect. *to appear in Phys. Rev. Lett.*, 2008.

[233] P. S. Barbeau, J. I. Collar, and O. Tench. Large-Mass Ultra-Low Noise Germanium Detectors: Performance and Applications in Neutrino and Astroparticle Physics. *JCAP*, 0709:009, 2007.

[234] S. T. Lin *et al.* New limits on spin-independent couplings of low-mass WIMP dark matter with a germanium detector at a threshold of 200 eV. arXiv:0712.1645, 2007.

[235] I. Avignone, F. T., P. S. Barbeau, and J. I. Collar. Comments on New Limits on Spin-Independent Couplings of Low-Mass WIMP Dark Matter with a Germanium Detector at a Threshold of 200 eV. arXiv:0806.1341, 2008.

[236] C. Rubbia. Past and future of non-accelerator physics with cryogenic liquids. In *Cryogenic Liquid Detectors for Future Particle Physics*, 2006.

[237] M. G. Boulay and A. Hime. Direct WIMP detection using scintillation time discrimination in liquid argon. arXiv:astro-ph/0411358, 2004.

[238] M. G. Boulay and A. Hime. Technique for direct detection of weakly interacting massive particles using scintillation time discrimination in liquid argon. *Astropart. Phys.*, 25:179–182, 2006.

[239] C. Galbiati and R. Purtschert. Discovery of underground argon with low level of radioactive 39Ar and possible applications to WIMP dark matter detectors. *J. Phys. Conf. Ser.*, 120:042015, 2008.

[240] J. A. Nikkel, W. H. Lippincott, and D. N. McKinsey. Neutrino and dark matter detection with CLEAN: Preliminary studies. *Int. J. Mod. Phys.*, A20:3113–3114, 2005.

[241] R. Bernabei *et al.* New limits on WIMP search with large-mass low- radioactivity NaI(Tl) set-up at Gran Sasso. *Phys. Lett.*, B389:757–766, 1996.

[242] R. Bernabei *et al.* WIMPs search by scintillators: Possible strategy for annual modulation search with large-mass highly-radiopure NaI(Tl). *Nucl. Phys. Proc. Suppl.*, 70:79–84, 1999.

[243] H. S. Lee *et al.* First limit on WIMP cross section with low background CsI(Tl) crystal detector. *Phys. Lett. B*, 633:201–208, 2006.

[244] R. Bernabei *et al.* Possible implications of the channeling effect in NaI(Tl) crystals. *Eur. Phys. J.*, C53:205–213, 2008.

[245] M. Ambrosio *et al.* Seasonal variations in the underground muon intensity as seen by MACRO. *Astropart. Phys.*, 7:109–124, 1997.

[246] D. P. Snowden-Ifft, C. J. Martoff, and J. M. Burwell. Low pressure negative ion drift chamber for dark matter search. *Phys. Rev.*, D61:101301, 2000.

[247] N. J. C. Spooner and P. Majewski. The DRIFT Directional Dark Matter Detector and First Studies of the Head-Tail Effect. *to appear in the proceedings of the 6th International Heidelberg Conference on Dark Matter in Astro and Particle Physics, Sydney, Australia*, 2007.

[248] G. Sciolla *et al.* The DMTPC Detector. *to appear in the proceedings of IDM2008, Stockholm, Sweden*, 2008.

[249] R. W. Schnee *et al.* The SuperCDMS experiment. arXiv:astro-ph/0502435, 2005.

[250] A. J. Noble. SNOLab: An international facility for low background particle astrophysics experiments. *New Astron. Rev.*, 49:311–313, 2005.

[251] B. A. Mazin. *Microwave kinetic inductance detectors*. PhD thesis, California Institute of Technology, 2004.

[252] W. T. Eadie, D. Drijard, F. James, M. Roos, and B. Sadoulet. *Statistical Methods in Experimental Physics*. North-Holland, 1971.

[253] E. T. Jaynes. *Probability theory: the logic of science*. Cambridge University Press, 2003.