A SEARCH FOR PARTICLE DARK MATTER
USING CRYOGENIC GERMANIUM AND SILICON DETECTORS
IN THE ONE- AND TWO-TOWER RUNS
OF CDMS-II AT SOUDAN

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF PHYSICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Reuben Walter Ogburn IV
June 2008

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

_____

Blas Cabrera   Principal Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

_____

Patricia Burchat

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

_____

Robert Wagoner

Approved for the Stanford University Committee on Graduate Studies.

Abstract

Images of the Bullet Cluster of galaxies in visible light, X-rays, and through gravitational lensing confirm that most of the matter in the universe is not composed of any known form of matter. The combined evidence from the dynamics of galaxies and clusters of galaxies, the cosmic microwave background, big bang nucleosynthesis, and other observations indicates that 80% of the universe's matter is dark, nearly collisionless, and cold. The identity of the dark matter remains unknown, but weakly interacting massive particles (WIMPs) are a very good candidate. They are a natural part of many supersymmetric extensions to the standard model, and could be produced as a nonrelativistic, thermal relic in the early universe with about the right density to account for the missing mass. The dark matter of a galaxy should exist as a spherical or ellipsoidal cloud, called a "halo" because it extends well past the edge of the visible galaxy.

The Cryogenic Dark Matter Search (CDMS) seeks to directly detect interactions between WIMPs in the Milky Way's galactic dark matter halo using crystals of germanium and silicon. Our Z-sensitive ionization and phonon ("ZIP") detectors simultaneously measure both phonons and ionization produced by particle interactions. In order to find very rare, low-energy WIMP interactions, we must identify and reject background events caused by environmental radioactivity, radioactive contaminants on the detectors, and cosmic rays. In particular, sophisticated analysis of the timing of phonon signals is needed to eliminate signals caused by beta decays at the detector surfaces.

This thesis presents the first two dark matter data sets from the deep underground experimental site at the Soudan Underground Laboratory in Minnesota. These are

known as "Run 118," with six detectors (1 kg Ge, 65.2 live days before cuts) and "Run 119," with twelve detectors (1.5 kg Ge, 74.5 live days before cuts). We have analyzed all data from the two runs together in a single, combined analysis, with sensitivity to lower-energy interactions, careful control of data quality and stability, and further development of techniques for reconstructing event location and rejecting near-surface interactions from beta decays.

We also present a revision to the previously published Run 119 analysis, a demonstration of the feasibility of a low-threshold (1 or 2 keV) analysis of Soudan data, and a review of the literature on charge generation and quenching relevant to our ionization signal.

Acknowledgements

I would like to thank Blas Cabrera, my Ph.D. advisor. Blas is an extremely insightful and very patient planner and advisor, but always an experimentalist first and foremost. I learned a lot from our time working together in Soudan, especially his approach to problem-solving. Blas has taught me the value of taking time to consider the right approach and the choose the right tools; and also the need to take big risks sometimes. He always treats his students as valued colleagues, and I appreciate the opportunities he has given me to set my own research goals and to decide for myself when a project is done.

It has been wonderful to be part of the CDMS collaboration, and it would be impossible to give proper thanks to everyone whose support and instruction I have enjoyed. I would especially like to thank Clarence Chang, who has been a faithful friend and mentor since I first wandered into the control room seven years ago. Thanks to the more senior students who showed me the way, Tarek, Vuk, Ray, and Joel. Betty, Dan A., Paul, and Rupak, thanks for all your encouragement at times when I really needed it. I am privileged to have been at Soudan with Dan B. for the successful slow transfer and the genesis of underground ping-pong. Dan and Long, I bow before your skills. CDMS is indebted to Jim Beaty for working weekends, late nights, and through his vacation days to make an underground mine into a science lab, and especially for keeping an eye on the biggest safety hazard in the lab (ladders).

Thanks and much sympathy to Jodi for taking Dark Pipe off my hands! Thanks to Matt and Kyle for many stimulating discussions and a lot of education about the gory details of phonon physics, SQUID amplifiers, and anisotropic charge mobilities. Thanks to Steve Yellin for keeping the bits flowing. Thanks to Pat, Jen, Steve L.,

Contents

Bibliography   **299**

List of Tables

List of Figures

Introduction

This dissertation flows from my graduate research in the search for particle dark matter as a collaborator in the Cryogenic Dark Matter Search experiment. CDMS is a collaborative experiment, currently including over 50 scientists at 16 institutions. It is fitting to recognize from the start that the results presented in this thesis therefore rest on the effort of a large number of people, including both the current members of the collaboration, and also many who laid the foundations for the experiment. This introduction serves to introduce the contents of the dissertation, and to clearly identify the parts that are uniquely my own work. The section "Author's involvement in CDMS" gives a different perspective, telling the story in chronological order and including the hard work that went into commissioning and operating the experiment in the deep site.

The fundamental goal of CDMS is to find out what the universe is made of. We believe that most of the matter in the universe is unseen, cold, and composed of something other than atoms. Chapter 1 reviews the astrophysical observations that support this idea, and introduces one of the leading candidates for the identity of dark matter: the weakly interacting massive particle or WIMP. The CDMS experiment has been designed to directly detect WIMPs through their rare, low-energy interactions with a target on Earth. This requires very sensitive detectors with the ability to find a small signal in the midst of a huge background. The group at Stanford University has developed the ZIP detector with these goals in mind. The ZIP is a low-temperature device made of a crystal of silicon or germanium, able to measure the energy released as phonons and as ionization when a particle interacts inside the crystal. Chapter 2 describes the design, fabrication, and principles of operation of the ZIP detector,

and Chapter 3 is devoted to the cryogenic systems, data acquisition electronics and hardware, shielding, and experimental facilities that transform the bare detectors into a working dark matter experiment.

The next three chapters cover a topic that has been a particular focus of my data analysis in CDMS. Chapter 4 describes the position sensitivity of the ZIP detector phonon signal, and several approaches to treating this position sensitivity in the data analysis. Chapter 5 discusses position dependences in the ionization channel. It presents my simulations of charge collection for problematic regions in the outer part of the detector, and the correction I have developed for the gradient in ionization signal pulse height with position. Chapter 6 presents a new technique for locating particle interactions within the detector, using a functional form-based model of the detector response.

The final section of the thesis is devoted to the combined analysis of the first two data sets from the deep underground site in the Soudan Mine, with six and twelve detectors. I have reanalyzed these two data sets together, starting from raw data. Chapter 7 describes the history and performance of the two runs, and lists the periods of low-background data that have been excluded in this analysis to ensure good data quality. Chapter 8 gives the data quality cuts and physics cuts used to select candidate WIMP scatter events, along with the cut efficiencies, but not including the timing cuts. These timing cuts, designed to reject interactions that occur near the detector surfaces, are the subject of Chapter 9. Finally, the WIMP-search results of this combined analysis are presented in Chapter 10. The final results are a slight improvement in efficiency over the previously published analyses in Ge detectors, and a slightly less stringent result in Si detectors. A signal indicative of WIMP detection is not seen.

Author's involvement in CDMS

When I first joined CDMS, in September 2001, the cryostat in the shallow underground site at Stanford had just been cooled to base temperature with a cargo of six ZIP detectors. This was Run 21, the first science run with ZIPs. The transition to

ZIP detectors from the earlier BLIP (Berkeley large ionization and phonon) design marked the beginning of CDMS-II. Run 21 was my introduction to several of the areas that would be central to this thesis work: data analysis, cryogenics, and CDMS software. I was trained in data analysis through work on the muon veto performance and background studies (including the first study of ^{210}Po alphas, which are closely related to our dominant surface-scatter background of ^{210}Pb; and identifying a class of high-energy gamma rays from cosmogenic activation). I began to study the position sensitivity of ZIPs, (reviewed in Chapter 4), and made a first attempt to construct a set of unambiguous coordinates for the location of each event. This early position reconstruction is described in Sec. 4.6.

Run 21 was also my introduction to the operation of the experiment, including cryogenics. I learned to transfer liquid helium and nitrogen, and made many weekend visits to the Stanford Underground Facility, less formally known as the "tunnel," for cryogen fills.[1] At SUF I was trained to identify changes in the noise performance of incoming data and to resolve the problems that came up. Run 21 was a successful demonstration of the capabilities of the CDMS detectors and experimental design, and we were able to achieve sensitivity to previously untested WIMP scattering cross sections.

However, the SUF site was shallow enough that after a year of running, the WIMP search was limited by the cosmogenic neutron background. For this reason, the Stanford site was only a temporary home for CDMS-II. At the same time as Run 21 data was flowing in, the Fermilab group was constructing a second facility in the Soudan Underground Mine in Soudan, Minnesota. A second experiment was built from the floor up, including the RF-shielded experimental clean room, a new cryogenic detector enclosure, a larger shield and muon veto, improved data acquisition electronics, and a second dilution refrigerator was purchased from Oxford Instruments. Most of these components were directly copied from the SUF design with only minor modifications. This duplicate facility was intended to allow the detectors themselves to be immediately transferred from Stanford to Soudan at the end of Run 21 with the

---

[1]During a few of these weekend fills, the dilution refrigerator required more than the standard daily cryogen fill. One of these incidents led to my first experience of a fridge crash, in June 2002.

smallest possible interruption to the WIMP search.

Unfortunately, the construction and commissioning of physics experiments seldom conform to pre-set schedules. This is especially true when a dilution refrigerator is involved. By late 2001, several runs of the Soudan fridge to base temperature on its own (disconnected from the detector enclosure) had failed to reach base temperature. Low-temperature experts from the entire collaboration were called to Soudan to repair the facility as quickly as possible. My first visit to Soudan was during a collaboration meeting in Ely, Minnesota, in March 2002. At this time I was a cryogenic novice, with experience only of maintaining the generally stable SUF fridge, which had been commissioned and cooled down for Run 21 before my arrival. The Soudan cryostat, in contrast, suffered leaks between the liquid helium bath and the insulating vacuum, and between the ^{3}He dilution unit and the vacuum; both of these appeared when cold, but disappeared at room temperature, making them very difficult to pin down. The helium bath itself was partly crumpled from a much earlier incident, with the strong possibility of failure at any time.

However, progress remained slow, and as Run 21 wound down in the middle of 2002, the Stanford group began to take frequent shifts in Soudan. My first shift was in June 2002, and starting in August I began to take regular two- and three-week shifts, spending half of my time in Minnesota and half at Stanford. During the next few months, I served as an assistant and trainee while the fridge was disassembled and reassembled, the leaks were found and repaired, and the cooldowns began to be successful. By good fortune in scheduling, I was able to cool the fridge from 4 K to base temperature with Betty Young in August 2002, for the first successful and stable run of the dilution unit in a dewar borrowed from Blas Cabrera at Stanford. Around this time I took responsibility for testing and repairing the cryogenic thermometry wiring, and revised and updated the cryogenic instruction documents. I took part in the first runs of the Soudan fridge in its own dewar (September 2002), the first cooldown with the icebox attached (December 2002), the insertion of the striplines (January 2003), and vast amounts of leak-checking. In March 2003 the six ZIP detectors from Stanford were installed ("Tower 1"), along with six new detectors ("Tower 2"). We first cooled down the two towers in April 2003, and began to work

out the remaining obstacles to smooth, stable running.

In May 2003, I experienced my second fridge crash, when a small remaining leak between the helium bath and the insulating vacuum suddenly increased in size, causing the (already low) liquid helium reservoir to quickly boil away. This intermittent leak would remain during the coming runs, causing periodic spikes in the rate of cryogen boiloff, and a stair-step pattern in the helium level. The phenomenon also decreased the hold time to the degree that two cryogen transfers were required every day.

In June 2003 we cooled the detectors again, and this time, we were able to stably remain at base temperature. I participated in various cryogenic upgrades, the detection and elimination of sources of noise on the detector readout channels, and the commissioning of the detectors. From October 2003 through August 2004 we continued in a mode of operation with much lower stress levels than were known in 2002: stable running. We used only the time-tested Tower 1 through January 2004, and began to configure Tower 2 from late January through early March. In August 2004 we deliberately ended the run, warming up to room temperature to prepare for the installation of an additional 18 detectors and a mechanical cryocooler (to improve the cryogen hold time).

At this point my involvement in day-to-day operation of the experiment decreased. I made several trips to help with the installation of the additional detectors and the cryocooler, and took charge of planning and executing the insertion of the striplines for the additional towers. My last trip to Soudan, in February 2005, was the (unsuccessful) first attempt at a cooldown with all five towers.[2] Between 2002 and 2005, I took 25 shifts at Soudan, for a total of 255 days underground.

While the first Soudan runs were taking data, we were also hard at work on analyzing the data. I studied the small variation in ionization pulse heights across the crystal, as well as the correction of crosstalk between the ionization channels. These are two of the issues discussed in Chapter 5, which is devoted to understanding the

---

[2]It would be more than another year before the full 30 detectors reached base temperature. The series of chance setbacks, acoustic noise problems, and imperfect thermal connections that had to be overcome for the success of the five-tower run were after the end of my travel to Soudan, and will be described in the dissertations of Jeff Filippini, Cathy Bailey, and others.

details, and especially the position dependences, in the ionization signal. I constructed a correction for the ionization pulse height gradient and for the small amount of crosstalk that was not accounted for in the initial pulse fitting. In the initial analysis of the two-tower data taken from March-August 2004, I constructed the blinding cut used to mask WIMP-search events that might lie in the final signal region, and performed a number of checks on data quality.

Continuing work on the reconstruction of event position, I used data from the 2003 and 2004 runs at Soudan to put together a functional form-based position analysis. This approach successfully unfolds the degeneracies in radial position for all detectors. It is detailed in Chapter 6.

After the publication of the two-tower result, I began to work on a combined analysis of the one- and two-tower runs. Alongside the initial analyses, I had also worked to prepare the data processing code to be compiled into a stand-alone executable and run outside of MATLAB. This allowed our pipeline to run on the computing farm at Fermilab, instead of small clusters of machines at Stanford and Soudan as had been previously done. At the same time, I made a number of other fixes and improvements to the data processing code. For the combined analysis, all of the data from the 2003 and 2004 runs at Soudan was reprocessed on several Fermilab farms using the new, stand-alone version of the analysis pipeline. Chapter 7 describes the data sets and excluded periods as used in the combined analysis.

I led the combined analysis effort, with the contributions of many others in setting cuts. Bruno Serfass's effort in defining the improved look-up table should especially be recognized. I performed the ionization channel calibration, crosstalk correction, and position correction; developed cuts to guard against a number of data quality problems; selected the good data sets for the WIMP search; defined the inner fiducial volume cut; selected the surface beta calibration sample and defined all of the timing cuts used; calcualted all efficiencies and leakages; and put all of the parts together to make a WIMP-search result. All cuts besides the timing cuts are described here in Chapter 8, the timing cuts for rejecting surface events are in Chapter 9; and the WIMP-search results are presented in Chapter 10. In this analysis I discovered an error in the previously published two-tower data, which is described in Appendix A.

Chapter 1

The dark matter mystery

In astrophysics, things that cannot be seen directly have sometimes been found through their effects on visible objects. The first planet discovered since classical times was Uranus, found by Sir William Herschel in 1781. Herschel identified Uranus as a planet because he was able to resolve it as a small disk, while stars appeared as perfect points. Over the next decades it was found that the new planet's motions could not be precisely accounted for by any one set of orbital elements. There always remained residual deviations that could not be reconciled with any single Keplerian orbit around the sun. In the 1840s, LeVerrier of France and Adams of England used observations of Uranus over the preceding century to deduce the properties of an additional planet capable of explaining the residuals.[1] The new planet was observed by Galle in 1846, confirming the calculations of LeVerrier and Adams, and Neptune thus became the first object to be seen indirectly by its gravitational effects before it could be directly seen [2].

Since then, hundreds of extrasolar planets have been discovered in radial velocity surveys, which use Doppler shifts in the spectra of stars to infer the existence of planets. Several of the extrasolar planets have been confirmed when they transit their stars, such as the "hot Jupiter" around HD 209458, which was discovered by the radial velocity method in 1999 and observed in transit in the same year [3].

---

[1]Uranus had been known at that time for 60 years, or only about three-quarters of one orbital period; however, earlier data points were actually available. It had previously been observed and recorded several times, but misidentified as a fixed star.

More similar to the methods used to discover Neptune are astrometric searches, which use precise measurements of the positions of stars as they orbit the common center of mass of their planetary systems. A sensitive survey of this type will be carried out by the NASA Space Interferometry Mission, to be launched in 2015 [4]. SIM will be a Michelson interferometer with 9-m baseline, giving parallaxes at 4 micro-arc second resolution, or differential resolution of 0.6 μas for bright targets. SIM PlanetQuest will operate in a targeted mode, making astrometric observations for stars of interest, including an "Earth Analog Survey" of 129 stars for which detection of terrestrial planets will be possible.

A complementary approach is taken by the NASA Terrestrial Planet Finder and ESA's Darwin, which are designed to image the starlight reflected from extrasolar planets. TPF uses two complementary approaches to solve the problem of seeing a relatively small, dim planet very near a much brighter parent star. The first is interferometry with *nulling*, in which the star's light interferes destructively to effectively disappear, leaving any planets visible. The second is a visible light coronagraph, with optics to block the light of the star. Darwin shares the interferometric approach. By measuring the spectrum of light from a planet, these missions will be able to determine temperature and detect the presence of water, oxygen, and organics [5]. Unfortunately, funding for TPF has been placed on hold since 2006. Darwin is currently in the assessment phase, and will launch no sooner than 2015.

1.1   Dark matter history

One of the most appealing reasons to search for extrasolar planets is to discover analogs of our own Earth - wet, terrestrial planets where life and even intelligent life could exist. However, the results have more often been surprising and inhospitable new categories like the "hot Jupiters," giant planets very close to their stars, with orbital periods of only a few days. On larger scales, too, gravity has revealed that the most abundant material in galaxies, clusters, and the universe is something unfamiliar and unexpected. Their dynamics are dominated by a dark component that has not been directly observed, but whose properties can be deduced from the behavior of

the visible matter.

The first signs of dark matter came in 1933, three years after the discovery of Pluto, when Fritz Zwicky and Walter Baade measured the velocities of galaxies in the Coma Cluster. From the observed dynamics, they calculated the cluster mass. The velocity distributions yielded the startling result that the galaxies were moving in the gravitational potential of a total mass about 160 times greater than expected based on the luminous matter in the cluster! On the basis of this observation, Zwicky postulated the existence of an unseen component of the cluster that contributed most of the mass, and coined the term "dark matter." This dark matter problem did not receive much attention for several decades. In 1970, Vera Rubin and Kent Ford brought renewed attention to dark matter when they published rotation curves for the Andromeda Galaxy M31. Again, the orbital velocities of objects in the galaxy were much faster than expected, revealing a much larger mass than indicated by the luminosity of the galaxy. In addition, the rotation curves were found to rise in the inner region, but to flatten near the edge of the disk and remain nearly constant out to high radius. This was a stark contrast to the expected Keplerian behavior with $v \propto r^{-1/2}$. The flat rotation curve implies a distribution of matter that extends to higher radius than the visible disk of the galaxy and contains much more mass.

In the following decades, evidence for the existence of dark matter was collected on a wide variety of length scales. In the late 1990s dark matter became one part of a "concordance cosmology" known as ΛCDM. The Λ (taken from Einstein's cosmological constant) stands for dark energy, which makes up 73% of the energy density of the universe and drives accelerating expansion of space. Cold dark matter (CDM) contributes another 23%, and is composed of an unknown, non-baryonic form of matter. The remaining 4% baryons are such a small component that they are left out of the acronym entirely! The dark matter must be "cold," in that it is non-relativistic at the time of decoupling, so that dark matter can drive the formation of structure in the universe [6]. It must be non-baryonic in order to avoid skewing the formation of light elements in big bang nucleosynthesis [7]. The following sections will give a brief overview of the dark matter problem in the context of galaxies and clusters, with particular attention to the halo properties and scattering rates that are relevant to

direct detection. A full review may be found in a recent article by Bertone, Hooper, and Silk [1].

1.2   Spiral galaxies

The dynamics of spiral galaxies, including the Milky Way, remain a key avenue for understanding the properties of dark matter. Rubin and Ford in 1970 used Hα and other spectral lines to study individual ionized atomic hydrogen (H II) clouds in the Andromeda galaxy [8, 9]. By combining spectral redshifts from several different types of objects it is possible to determine the distribution of mass at the center of the bulge, through the disk, and in the outer halo [10]. Observations have used various optical lines, 21 cm radio, CO rotational transitions, maser lines, and for stars near the center of the Milky Way galaxy, proper motion. Figure 1.1(a) shows such a composite picture of the rotational velocity of the Milky Way, on logarithmic axes. In the central parsecs the super-massive black hole near Sagittarius A* dominates, and the velocity falls off with radius. Moving outward, velocities then increase as the mass of the bulge is enclosed, out to a few hundred parsecs. Beyond this, including the disk and outer halo, the rotational velocity is nearly flat. Figure 1.1(b) shows the profiles for a number of spiral galaxies, with substantial variation in the inner region, but a consistent flattening beyond the visible disk. For any specific spiral galaxy it is possible to decompose the rotation curve into several components, although the decomposition is not necessarily unique. Figure 1.2 shows such a decomposition for the Milky Way, from Olling and Merrifield [12]. It is impossible to match the observed rotation curve either in magnitude or in shape using only the visible components of the galaxy; an extended, massive halo of dark matter is required. In this analysis, the positions and motions of objects within the Milky Way are observed relative to Earth, and must be translated into the coordinates of the galactic center. Therefore, uncertainties about the distance R_0 of the sun from the center of the galaxy and the sun's orbital velocity v_0 affect the data points as well as the model curves., These assumptions affect the data points, and not only the models, because the positions and motions of objects within the Milky Way are observed relative to Earth, and

(a)

(b)
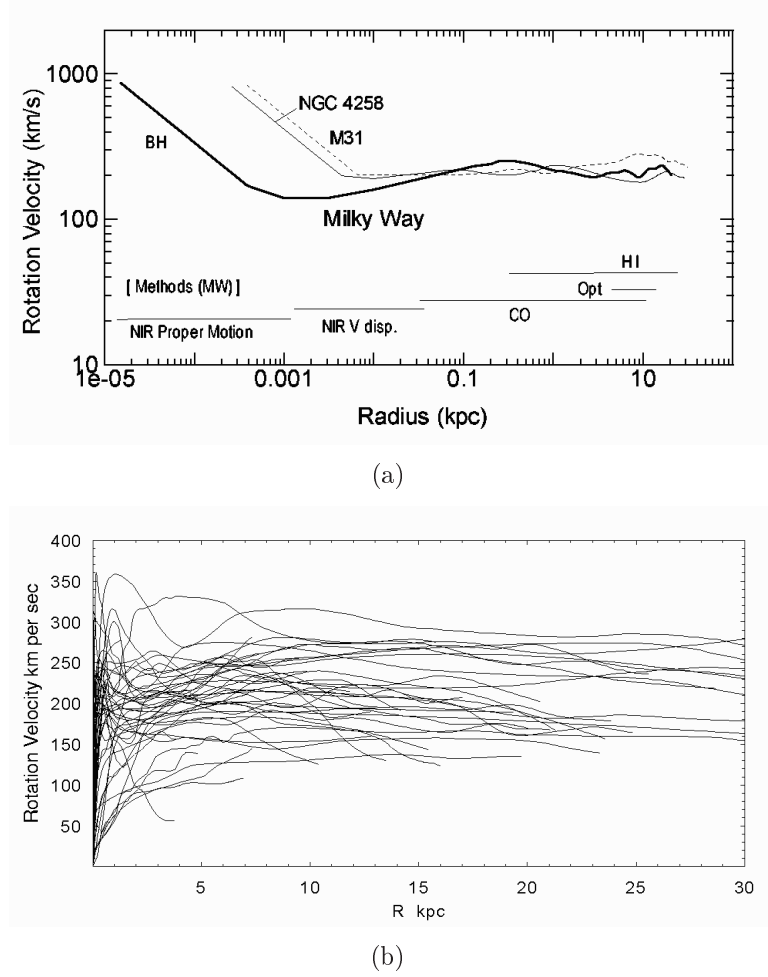
Figure 1.1: Rotation curves for (a) the Milky Way galaxy, the Andromeda Galaxy, and NGC 4258; and (b) a collection of spiral galaxies. Source: Sofue and Rubin [10]. In (a), the several sources of data for different radii are indicated, and the logarithmic scale shows that velocities in the inner region do not approach zero, but rise in Keplerian form as the central supermassive black hole becomes dominant. Subfigure (b) originally comes from Sofue et al. [11] and superimposes a number of spiral galaxies. The behavior in the inner regions varies substantially, but there is a consistent flattening at high radii - including the region beyond the disk.
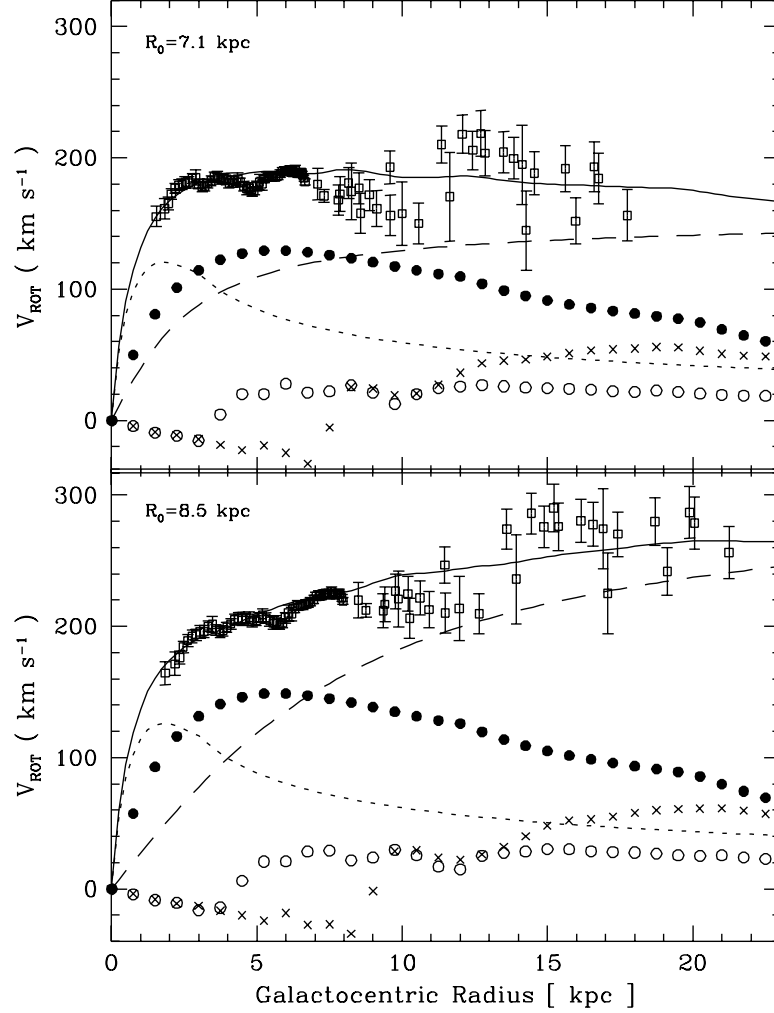
Figure 1.2: Milky Way velocity profile with decomposition by Olling and Merrifield [12]. The two plots differ in their assumptions: the upper panel uses $R_0 = 7.1$ kpc, $v_0 = 185$ km/s, and the lower panel uses $R_0 = 8.5$ kpc, $v_0 = 220$ km/s. The individual components are: bulge (small dashes), stellar disk (filled circles), H I layer (crosses), H_2 layer (open circles), and dark matter halo (long dashes). The solid curve is the sum of the components. The data points come from observation of hydrogen gas regions., with larger error bars in the outer regions because of the increased difficulty in constraining distances outside the Solar System's galactic radius.

must be translated into the coordinates of the galactic center. Figure 1.2 illustrates the effect of these uncertainties by using different values of R_0 and v_0 in the top and bottom panels.

Based on observed rotation curves, theory, and modeling, several forms have been suggested for the distribution of dark matter within a galaxy's halo. Given spherical symmetry, flat rotation curves imply $\rho(r) \propto r^{-2}$, with a mass $M(r) \propto r$ contained within radius r. In order to avoid a singularity at the center, the density can be limited at some inner radius R to give a flattened isothermal sphere:

$$\rho(r) = \frac{\rho_0}{1 + (\frac{r}{R})^2}. \tag{1.1}$$

However, the total mass in this profile grows arbitrarily large out to unlimited radius. Another commonly used form, designed to match simulations of structure formation in dark matter-dominated galaxies, is the Navarro-Frenk-White profile [13],

$$\rho(r) = \frac{\rho_0}{\frac{r}{R}(1 + \frac{r}{R})^2}. \tag{1.2}$$

This approximates r^{-2} at intermediate radius, has a somewhat cuspy central behavior, and falls off as r^{-3} at large radius. (This asymptotic r^{-3} behavior gives a total mass that is logarithmically divergent.) The isothermal sphere is an example of a "cored" inner profile - with constant density at the center - while the NFW profile is an example of a "cusp." This distinction has strong implications for indirect detection. Since both models have divergent total mass, some form of truncation is needed to keep the total halo mass finite.

1.3   Galaxy clusters

Galaxy clusters are the largest bound structures in the universe, containing not only visible galaxies, but also a much larger mass of hot intracluster medium (ICM) and a halo of dark matter. Fritz Zwicky's ground-breaking work measured the redshift of eight galaxies in the Coma cluster [14]. The galaxies were identified as part of the

same cluster by their similar redshifts (and by Hubble's law, similar distances from Earth). But each had a slightly different redshift, indicating a spread in velocities as the galaxies orbit within the cluster. The velocity dispersion was found to be large, 1019±360 km/s. Zwicky combined this with Hubble's estimate of the mass of a typical spiral galaxy to get the total kinetic energy of all the galaxies in the cluster. From the mean velocity (again, from the common redshift) and Hubble's law, he calculated the distance to the cluster. The best value of the Hubble parameter available in 1933 was ~ 550 km/s/Mpc, and it would take another 65 years for H_0 to be fully pinned down![2] A measurement of the cluster's angular diameter on the sky could then be translated into a physical dimension - although the too-large value of H_0 gave a size that was about seven times too small.

Since the cluster has been bound for a very long time, the virial theorem can be used to relate the kinetic and potential energies:

$$2\langle T\rangle = -\langle V\rangle \sim \frac{GM_{cluster}^2}{2\langle r\rangle}. \qquad (1.3)$$

The surprising result was that the gravitational mass was higher than expected based on the luminosity of the galaxies by a factor of 400! This is usually expressed as a mass-to-light ratio, M/L, normalized to 1 for the mass and luminosity of the sun, with M inferred from gravitational dynamics using the virial theorem, and L from the visible light emitted by the cluster galaxies. Zwicky's result was inflated by the overestimate of H_0, but modern estimates of M/L remain similarly high. A study of 29 clusters from the ESO Nearby Abell Cluster Survey (ENACS) found an average value of $\frac{M}{L} = 454h\frac{M_\odot}{L_\odot}$ [16, 17] (where h is the Hubble constant in units of 100 km/s/Mpc). The denominator in the mass-to-light ratios only take into account the galaxies, and not the hot intracluster medium, which makes up the bulk of the baryons, so M/L does not directly give the dark matter fraction. Related but more sophisticated techniques can reveal the distribution of all mass within the cluster. The Jeans equation relates the velocity dispersion as a function of cluster-centric radius to the underlying mass distribution. By subtracting the contributions of the observed

---

[2]As of 2007, the three-year WMAP data favor $H_0 = 73.2 \pm 3.2$ km/s/Mpc [15].
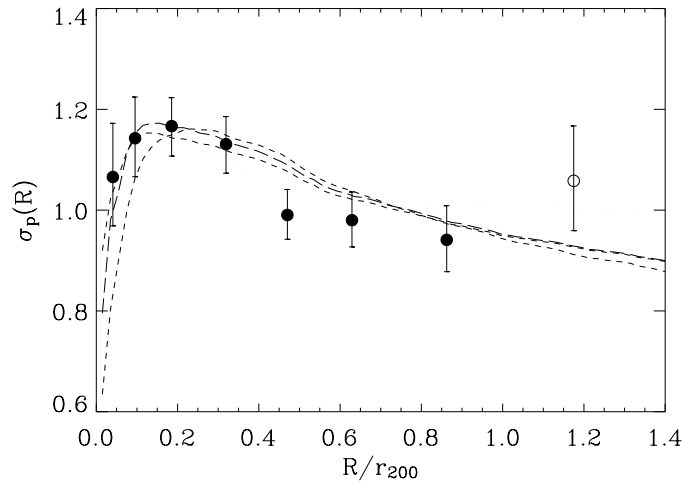
Figure 1.3: Cluster velocity dispersions from the ESO Nearby Abell Cluster Survey [17]. Circles with error bars are from early-type galaxies (last data point not used). The curves represent several halo models: isothermal (dotted), NFW (long dashes), M99 (short dashes), and Burkert (dot-dash). Both axes are normalized to give relative, dimensionless quantities. The cluster-centric radius on the x axis has been normalized to r_{200}, the radius at which the cluster density is 200 times greater than the cosmological critical density. The velocity dispersion on the y axis is normalized to the global value σ_p for each cluster considered as a whiole.

galaxies and gas from the total gravitational mass profile $M(r)$, the dark matter profile can be inferred. This approach has been used to test halo models on 59 clusters in the same ESO survey [17]. The results so far are compatible with cuspy profiles like NFW as well as cored profiles like the isothermal sphere, although any cores must have a very small radius in order to remain consistent with observation. Figure 1.3 shows the velocity dispersions from the data (averaged over clusters), compared with selected halo profiles.

The hot intracluster gas can also be used as a probe of dark matter. The gas is heated to 10-100 million K as it falls into the center of the cluster's gravitational well, and emits X-rays by thermal bremsstrahlung; the emission spectrum depends on the depth of that well. Therefore, observations of the X-rays from the ICM can reveal not only the mass of the ICM, but also the dark matter mass and distribution. Lewis *et al.* have used Chandra data for the cluster A2029 to reconstruct the density and

temperature of the gas as a function of cluster-centric distance [18]. This study gives an especially detailed probe of the ICM, and therefore also the halo profile, near the center of the cluster. They find a mass distribution consistent with cuspy models and inconsistent with a cored profile, unless two core sizes are used - one large and one small. The deduced mass fraction of the ICM (out of a total mass including galaxies, ICM, and dark matter) is below 3% at the center of the cluster, and up to 13.9% at larger radii. The dark matter halo accounts for 90-95% of the total mass.

While optical and X-ray studies of clusters have been underway for many years, it has more recently become possible to use gravitational lensing as a probe of mass. Strong gravitational lensing produces arcs, Einstein rings, and multiple images from background objects, and weak lensing gives a slight distortion to the shapes of distant galaxies. Since most galaxies are not intrinsically spherical, the lensing effect can only be detected statistically, using a large number of galaxies. High-resolution CCD images and computerized image processing make this feasible. Johnston et al. [19] have compiled weak lensing profiles for 130,000 groups and clusters ($0.1 \leq z \leq 0.3$) from the Sloan Digital Sky Survey, all of which have identifiable, central "brightest cluster galaxies" (BCGs). The BCGs are typically very large, diffuse elliptical galaxies that have grown by the merger or cannibalization of a large number of progenitors, and are massive enough to contribute significantly to the lensing profile in the central region. The observed lensing profiles have been averaged over many clusters in bins of luminosity or mass, and decomposed into several components as shown in Fig. 1.4. The largest contribution is from the dark matter halo, with an inferred profile consistent with NFW.

The "bullet cluster"

All three of these approaches to measuring the masses of clusters (orbital velocities of galaxies, X-ray temperature, and gravitational lensing) have recently been combined for the extraordinary case of the galaxy cluster 1E 0657-56, better known as the "bullet cluster." It is made up of two subclusters in the process of merging, which appear to have passed directly through one another. The space between galaxies is large
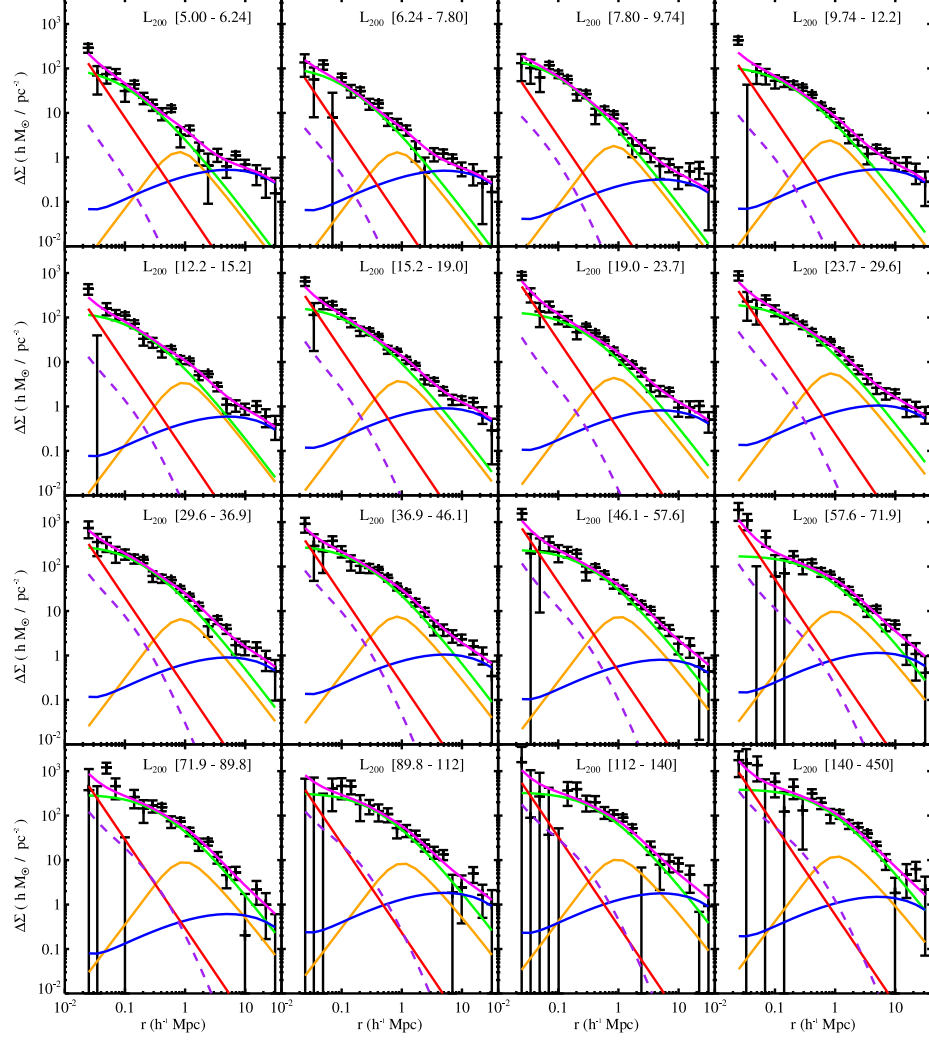
Figure 1.4: Weak lensing profiles averaged over SDSS clusters in bins of luminosity, from [19]. The vertical axis $\Delta\Sigma$ is a surface mass density for the cluster, projected along the line of sight, as inferred from lensing. The contributions are: NFW dark matter halo (green), mis-centered halo component (orange), brightest cluster galaxy (red), neighboring halos (blue), and non-linear contribution (purple, dashed). The combined best fit model is shown in magenta. The total cluster luminosity L_{200} is the total luminosity within r_{200} (as inferred from galaxies), in units of $10^{12}hL_{\odot}$.

enough that the galaxies in each subcluster have passed through the other subcluster without collisions. The ICM, on the other hand, has been shocked and heated by the interaction, and remains concentrated closer to the collision point. This sets the stage for a very beautiful test of the dark matter model: if the mass of the clusters is dominated by collisionless dark matter, then weak lensing will show two centers near the concentrations of galaxies. If, on the other hand, the mass is mainly in the ICM (with the high M/L ratios of galaxies and clusters explained by modified gravity), then weak lensing should show one center, near the gas. Clowe et al. have found that the lensing signal is indeed strongest very near the main cluster and the smaller cluster, and significantly separated from the highest intensity of X-rays, consistent with the dark matter model [20].

1.4   WIMPs

Although the composition of dark matter remains unknown, one natural candidate is a weakly interacting, massive particle. These "WIMPs" have several advantages: they arise naturally in supersymmetric models and as Kaluza-Klein particles in theories with extra dimensions; they are cold, as required for structure formation; and they can be generated as a thermal relic of the big bang, with approximately the required density. The transition from thermal equilibrium to a fixed relic density is known as "freeze-out," and is illustrated in Fig. 1.5. A relic particle is initially in equilibrium with all other matter and energy in the early universe, with a density given by a relativistic thermal distribution, $n_{eq} \propto T^3$, and with roughly as many WIMPs as photons. This holds as long as the temperature is high compared to the particle mass, so that the WIMPs can both be produced and annihilate. When $k_B T$ falls below the WIMP mass M_χ, production is suppressed, but annihilations continue. If this equilibrium went on until the present day, the density of WIMPs would be very low, $n_{eq} \approx (M_\chi T/2\pi)^{3/2} e^{-M_\chi/T}$. However, the WIMPs do not decay individually, but only annihilate in pairs. Once the expansion of the universe dilutes their numbers sufficiently that WIMP-WIMP interactions become very rare, their comoving density is fixed. The final abundance is set by the cross-section for annihilation:
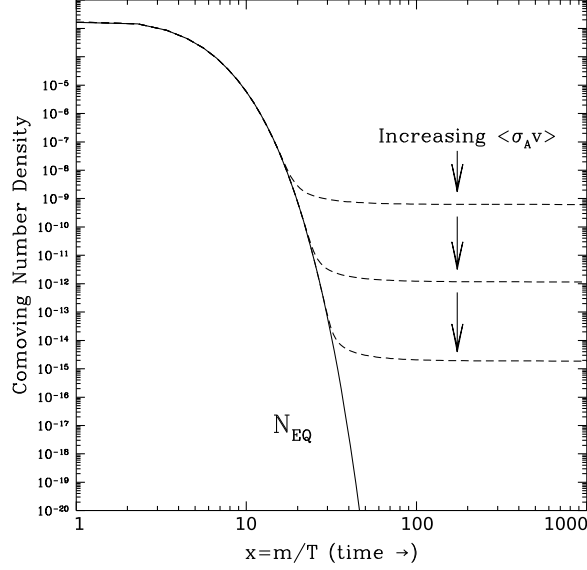
Figure 1.5: Abundance of a thermal relic particle before, during, and after freeze-out. The solid line shows the thermal equilibrium abundance, and the dashed lines show the relic density as it depends on the cross-section for annihilation. (The normalization of the vertical scale is unknown.) Figure from [21].

$$\Omega_\chi h^2 = \frac{3 \times 10^{-27} \text{ cm}^3\text{s}^{-1}}{\langle \sigma_A v \rangle}. \tag{1.4}$$

In this expression, $\langle \sigma_A v \rangle$ is the thermally averaged cross-section for annihilation times the relative velocity of the WIMPs. For a single species of particle with weak-scale interactions, we can expect $\Omega_\chi \sim 0.3$, to match the cosmological dark matter density [22].

The mechanism above is generic to thermal relic particles, but we can consider several specific particle physics scenarios that give rise to WIMPs. The first is super-symmetry. Although SUSY was originally conceived to stabilize the mass of the Higgs boson against loop corrections, it can also provide convenient dark matter candidates. In supersymmetry, each particle in the standard model has a supersymmetric partner with identical gauge charges, higher mass, and complementary statistics: bosons have fermionic partners, and fermions have bosonic partners. In weak-scale supersymmetry with R-parity conservation, the lightest supersymmetric particle (LSP) is stable and has mass on the order of 100 GeV. The LSP is often a neutralino: a mixture of

the superpartners of the B and neutral W, and the two lightest Higgsinos (where B and W are the uncharged vector bosons of $U(1)$ and $SU(2)$ respectively in unbroken electroweak theory). A thorough review of supersymmetric dark matter may be found in Jungman, Kamionkowski, and Griest [22]. Similarly, in theories with extra dimensions, a stable lightest Kaluza-Klein state (LKP) can have the properties of a WIMP, especially if it is a KK-photon [23].

1.5   WIMP scattering

The halo profile for the Milky Way predicts a dark matter density of $\rho_0 \approx 0.3$ GeV/cm^3 in the vicinity of Earth. For the case of 100-GeV WIMPs, we can expect a flux of

$$\Phi = \frac{\rho_0 v_0}{M_\chi} \approx 7 \times 10^4 \text{ cm}^{-2}\text{s}^{-1} \tag{1.5}$$

and a density of about three particles per liter at any given time! In this picture there are an abundance of particles to be detected, if only the cross-sections are high enough. The possibility of direct detection was suggested by Goodman and Witten, who pointed out the importance of coherent interactions between dark matter particles and nuclei [24].

Since relic WIMPs have weak-scale cross-sections for annihilating to standard model particles, they may also be expected to have weak-scale cross-sections for elastic scattering on standard model particles. The form and magnitude of these interactions depends on the precise model considered (even within the constrained minimally supersymmetric standard model or CMSSM, there are many adjustable parameters affecting the cross-sections). However, we can write a general framework for WIMP-detector scattering, with two classes of interactions: *spin-independent* interactions which couple to the entire nucleus; and *spin-dependent* interactions which couple to unpaired spins. We will focus on the cross-sections and rates for nuclei of interest as targets in dark matter detection experiments: in particular, the materials used by CDMS, Ge and Si, along with Xe, which is used by several competing experiments.

1.5.1 Spin-independent scattering

The spin-independent or scalar interactions have differential cross-section [22]

$$\frac{\mathrm{d}\sigma}{\mathrm{d}|\mathbf{q}|} = \frac{1}{\pi v^2} \left[Zf_p + (A - Z)f_n\right]^2 F^2(q). \tag{1.6}$$

Here A and Z are the atomic mass and atomic number of the target material, v is the WIMP velocity, $\mathbf{q}$ is the momentum transferred, and $F(q)$ is a form factor that accounts for the momentum distribution of nucleons in the target. The couplings f_p and f_n to protons and neutrons are typically similar, and depend on the scattering processes available to a given type of WIMP. The cross-section does not scale linearly with the size of the nucleus, but quadratically. This arises because the neutralino can interact coherently with the entire nucleus rather than with a single nucleon, as long as the momentum transferred is not too large. When q corresponds to a wavelength much larger than the nucleus, the scattering is fully coherent. As q rises, the nuclear structure becomes important, with the exact dependence encoded in $F(q)$. There are several commonly used models of $F(q)$. We adopt the form factor of Lewin and Smith [25]. This is a Helm form factor of the form

$$F(q) = \frac{3j_1(qR_1)}{qR_1}e^{-q^2s^2/2}, \tag{1.7}$$

where $j_1(x) = sin(x)/x^2 - cos(x)/x$ is a spherical Bessel function of the first kind, R_1 is an effective nuclear radius, and s is a skin depth reflecting the sharpness of the edge of the nucleus. Lewin and Smith give the following formulas for R_1 and s, applicable to any nuclide:

$$
\begin{aligned}
s &= 0.9 \text{ fm} \\
c &= \left(1.23A^{1/3} - 0.60\right) \text{ fm} \\
a &= 0.52 \text{ fm} \\
R_1 &= \sqrt{c^2 + 7\pi a^2/3 - 5s^2}.
\end{aligned}
$$

These values have been chosen to make Eq. 1.7 approximately equal to the Fourier transform of the Woods-Saxon formula, which is a two-parameter Fermi distribution giving the spatial distribution of charges in the nucleus. Our selected $F(q)$ therefore amounts to a "fit of a fit." Although more sophisticated approaches are available, based on electron elastic scattering data rather than model-dependent parameterizations, the Lewin & Smith formula has been found to be an acceptably close match [26] and has been adopted as a working standard by many direct-detection experiments. The Helm form factor, with parameters as in Lewin & Smith, is shown in Fig. 1.6(a) for Si, Ge, and Xe.

It is convenient to express cross-sections in terms of the cross-section for scattering off a pointlike nucleus, at zero momentum transfer:

$$\sigma_0^{SI} = \frac{4\mu^2}{\pi}[Zf_p + (A - Z)f_n]^2. \tag{1.8}$$

Here μ is the reduced mass of the WIMP-nucleus system.

1.5.2   Spin-dependent interactions

For spin-dependent interactions, on the other hand,

$$\frac{\mathrm{d}\sigma}{\mathrm{d}|\mathbf{q}|} = \frac{8G_F^2}{(2J + 1)v^2}S(q), \tag{1.9}$$

where J is the total nuclear spin, and $S(q)$ is a "spin structure function" given by

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

The spin-dependent couplings of the WIMP to neutrons and protons enter as a_n and a_p, with $a_0 = a_n + a_p$ and $a_1 = a_p - a_n$. The $S_{ij}(q)$ describe the distribution of spins within the target nucleus, analogously to $F(Q)$ in the spin-independent case.

Although neutralinos often have intrinsically larger spin-dependent than spin-independent couplings to nucleons, Eq. 1.6 shows that scalar interactions are enhanced by the square of the target nuclear mass, while the spin-dependent cross-section in Eq.

1.9 does not increase with A. Therefore, direct detection through scalar interactions is favored in most cases. We will focus on the spin-independent case and note that CDMS has also presented analyses for the case of dominant spin-dependent scattering [27].

1.5.3 Event rates

The differential rate for scalar interactions can be written in terms of σ_0^{SI} from Eq. 1.8:

$$\frac{\mathrm{d}R}{\mathrm{d}E} = \frac{\rho \sigma_0^{SI} |F(q)|^2}{2M_\chi \mu^2} \int_{v>q/2\mu} \frac{f(\vec{v}, t)}{v} \mathrm{d}^3 v. \tag{1.10}$$

The reduced mass is μ, and the momentum transferred is q. The lower limit of integration is the minimum WIMP velocity required in order for it to be kinematically possible for an energy E to be transferred to the nucleus. For the velocity profile we assume a Maxwellian distribution truncated at the galactic escape velocity v_{esc}; however, v_{esc} is large enough that it has little effect on the calculation, and so we have omitted it here. The Maxwellian distribution is

$$f(v) \mathrm{d}^3 v = \frac{1}{v_0^3 \pi^{3/2}} e^{-v^2/v_0^2} \mathrm{d}^3 v. \tag{1.11}$$

Substituting this form into Eq. 1.10 gives an energy spectrum that is a falling exponential modified by $F(q)$.

$$\frac{\mathrm{d}R}{\mathrm{d}E} = \frac{\rho_0 \sigma_0^{SI} |F(E)|^2}{\sqrt{\pi} v_0 M_\chi \mu^2} \exp\left(-\frac{EM_N}{2\mu^2 v_0^2}\right). \tag{1.12}$$

Here we have not included the small corrections for the galactic escape velocity, the Earth's motion around the Sun, and the Sun's motion around the galactic center. A treatment including these effects can be found in Lewin and Smith. In order to compare results across experiments, it is useful to express rates in terms of the cross-section for scattering on a single nucleon, $\sigma_{0,n}^{SI}$. The cross-section for a nuclear species i is $\sigma_{0,i}^{SI} = \sigma_{0,n}^{SI} \frac{\mu_i^2}{\mu_n^2} A^2$, where μ_i and μ_n are the reduced masses for the WIMP-nucleus

(a)

(b)

(c)
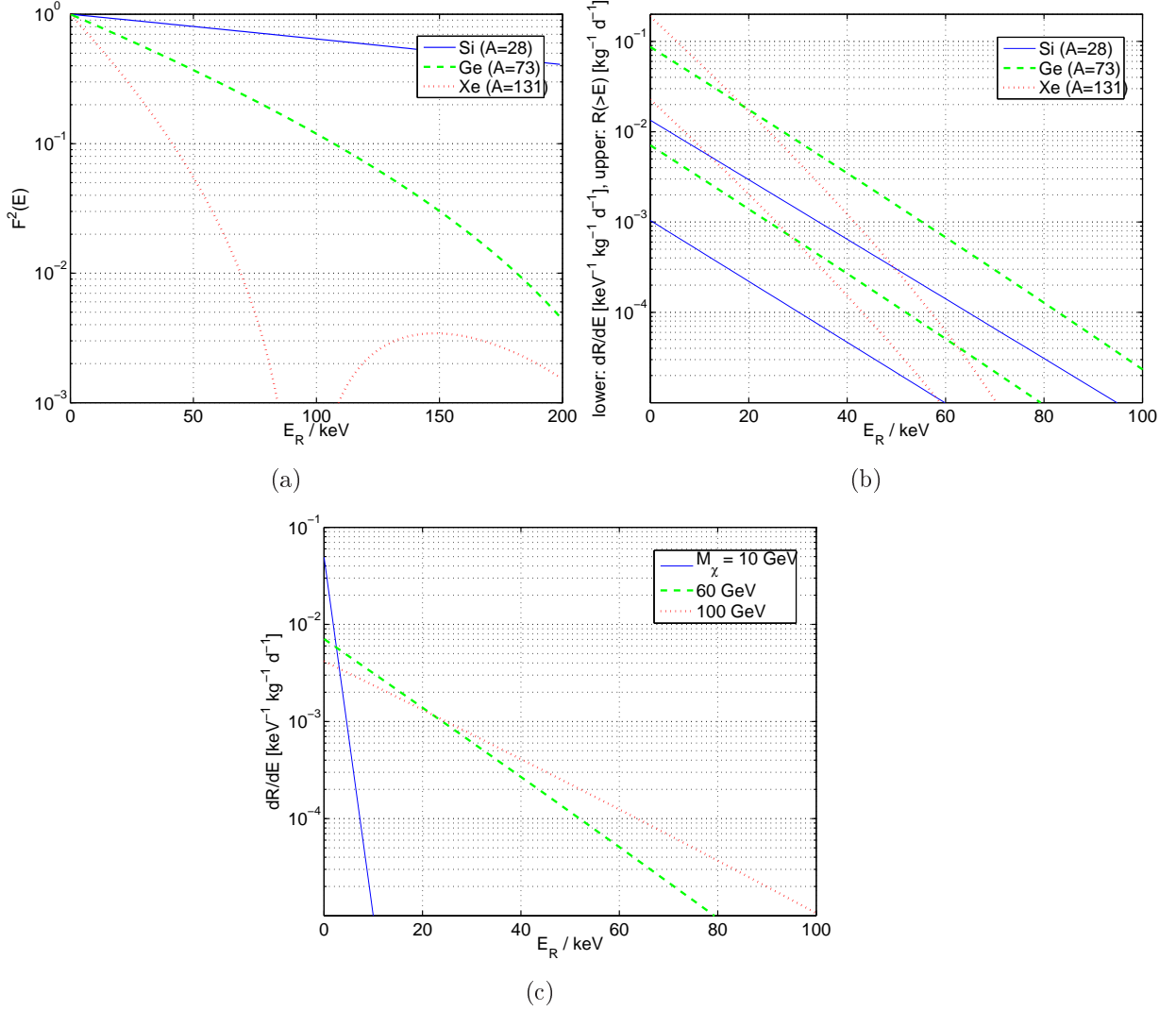
Figure 1.6: Elastic scattering form factors and rates as a function of recoil energy, target nucleus, and WIMP mass. In (a), the form factors $F^2(E)$ for Si, Ge, and Xe. Xenon loses coherence at lower energies, and reaches the first diffraction minimum at 100 keV. In (b), the differential and integrated rates for a 60-GeV WIMP with $\sigma_{0,n}^{SI} = 2 \times 10^{-43}$ cm^2. Characteristic rates for Ge and 10-keV threshold are ~ 0.04 per kg per day. In (c), the dependence of rate and energy spectrum on WIMP mass for Ge. All figures use the Helm form factor after Lewin & Smith.

and WIMP-nucleon systems respectively. Putting it all together,

$$\frac{\mathrm{d}R_i}{\mathrm{d}E} = \frac{\rho_0 \sigma_{0,n}^{SI} A^2 |F(E)|^2}{\sqrt{\pi} v_0 M_\chi \mu_n^2} \exp\left(-\frac{E M_N}{2\mu_i^2 v_0^2}\right). \tag{1.13}$$

The rate and spectrum are shown in Fig. 1.6(b) and (c) for targets and WIMP properties of interest. For this calculation we have set $\sigma_{0,n}^{SI} = 2 \times 10^{-43}$ cm^2 and used standard halo properties $\rho_0 = 0.3$ GeV/cm^3, $v_0 = 220$ km/s. In (b), the target is varied while the WIMP mass is held fixed at 60 GeV. In (c), the target is Ge and the three curves show the spectrum and rate for different values of the WIMP mass.

The expected scattering rates depend on several assumptions about the precise properties of the Milky Way galaxy's dark matter halo. Since the dark matter is cold, it should have velocities similar to other objects orbiting the galaxy. In the vicinity of Earth, this is $\langle v_0 \rangle \approx 220$ km/s. In order to compare experiments with each other and with theory, we will adopt a set of "standard halo" assumptions. The standard halo has the following local properties: $v_0 = 220$ km/s, a density $\rho_0 = 0.3$ GeV/cm^3, and a Maxwellian velocity distribution. These assumptions are astrophysically reasonable, but subject to some uncertainty. The true value of v_0 may range from $170 - 270$ km/s, and ρ_0 from $0.2 - 0.4$ GeV/cm^3. Different choices of parameters can change the expected rates by a significant amount, up to factors of a few [28, 29]. This is acceptable, as the theoretical properties of WIMPs are subject to even larger uncertainties. Comparisons between the sensitivities of competing direct detection experiments are not strongly sensitive to halo uncertainties. The standard halo will be used for the calculation of scattering rates in Sec. 1.5 and the exclusion limits in Chapter 10.

1.6 Indirect detection

The thermal relic mechanism suggests a possible indirect signal of particle dark matter. If particles and antiparticles are equally common, or if they are Majorana (as in the case of SUSY WIMPs), they can annihilate wherever the dark matter density is high enough. Annihilation requires two particles, so that the signal scales with

the square of the WIMP density. WIMPdensity may be high at the center of the galaxy, depending on whether the density profile exhibits cusps or cores at the center of the halo; this can also be enhanced if there are subhalos or "clumps." Dark matter particles may also be trapped at the center of the Earth and the Sun.

Annihilations at the galactic center can produce high-energy gamma rays with characteristic spectra, as well as an antimatter component in cosmic rays. Indirect detection experiments focused on the Earth and the Sun search for very high-energy neutrinos. A recent review of indirect detection experiments may be found in Carr *et al.* [30]. Three results will be briefly mentioned here: the gamma ray signal from the galactic center reported by EGRET and others; the limits on annihilations in the sun from Super-Kamiokande; and the intriguing "WMAP haze."

Gamma rays from the galactic center

Gamma ray observatories have yielded several excesses signals from the galactic signal which may be interpreted as a result of WIMP annihilation. To date, however, none of these has been positively confirmed as a dark matter signal. The excesses at energies from 3 MeV to 100 TeV are shown together in Figure 1.7.

EGRET is the Energetic Gamma Ray Experiment Telescope on the Compton Gamma Ray Observatory. In 1997, Mayer-Hasselwander et al. reported an excess of energetic gamma rays above 500 MeV [31]. This emission was localized to within 85 parsecs of the galactic center, and can be interpreted as the result of WIMP annihilation in a cusp or spike at the center of the dark matter halo. The EGRET spectrum can be matched well by self-annihilating neutralinos of mass 10-100 GeV [32].

Excesses have also been detected by the ground-based instruments HESS, CANGAROO-II, and VERITAS, but at higher energies of 0.1-10 TeV. These results can also be explained in terms of WIMP annihilation. However, they require heavier WIMPs; the best-fit model gives a mass of 19 TeV with HESS data, or 1 TeV for CANGAROO data [33]. These masses are higher than those usually favored for WIMPs, as WIMPs more massive than $\sim$1 TeV tend to have thermal relic densities that are too high.
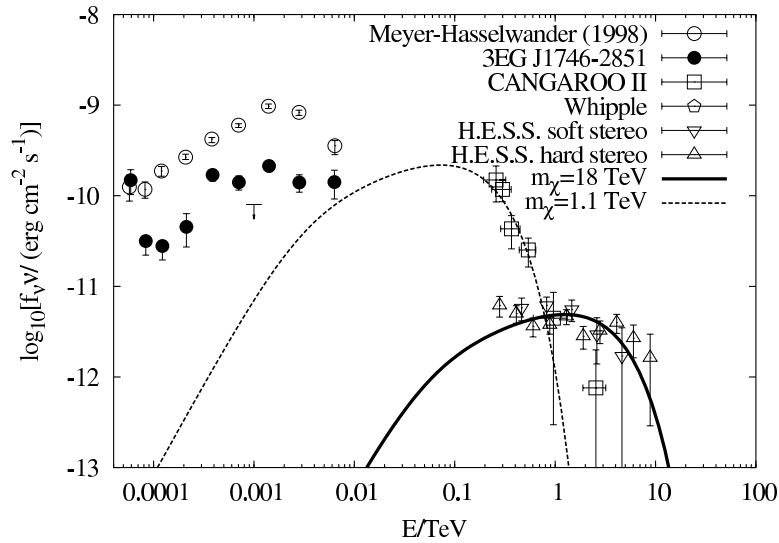
Figure 1.7: Excess γ flux from the galactic center, from D. Horns [33]. Open and filled circles represent data from EGRET, consistent with a possible WIMP of mass 10-100 GeV; squares are CANGAROO data, matched well by a 1.1 TeV WIMP; and triangles are HESS data, matched well by a 21.1 TeV WIMP.

This problem can be avoided in some models, but the interpretation as a dark matter annihilation signal can be ruled out in some of the simplest supersymmetric models. It has been shown that the HESS data cannot be explained entirely as WIMP annihilation under the assumptions of a thermal relic neutralino in minimal supergravity (mSUGRA) [34].

The INTEGRAL satellite has shown evidence for electron-positron annihilation at the galactic center in 511 keV gamma rays. This can also be a signature of dark matter annihilations, but very light particles are required, with masses $\sim$1-100 MeV [35]. These fall into a separate category of dark matter candidate from WIMPs, and must have exotic properties, such as annihilation mediated by new gauge forces, or unusual self-interactions.

The EGRET signal will be explored in more detail and with better angular resolution by GLAST [36], while the higher-energy region will be studied by the ground-based CANGAROO III [37] and MAGIC [38] telescopes. Given the abundance of signals that are poorly understood at present, and that cannot be collectively explained

by any single, straightforward dark matter annihilation model, it seems certain that some of the features will be explained by astrophysical sources not related to dark matter. The next-generation experiments will go a long way towards improving our understanding of the physics of the galactic center, and will have great potential for discovering or constraining a WIMP annihilation signal.

Neutrinos

The SuperKamiokande neutrino detector can set limits on the rate of WIMP annihilations in the sun. As particles of dark matter travel through the solar system, some will elastically scatter inside the sun. If they lose enough energy to become gravitationally bound, they will eventually settle into the center of the sun, where the WIMP density can grow high enough for annihilations to occur. Because of the abundance of hydrogen in the sun, the interaction rate is especially sensitive to WIMPs with spin-dependent interactions on unpaired protons. Among other products, the annihilation can yield neutrinos. When very high-energy neutrinos (about 5 GeV to 5 TeV) interact in the rock near the detector they can produce muons that propagate all the way through the SuperK chamber. Muons traveling downward may be confused with atmospheric muons, but there is a very low background for upward-going muons in this energy range. The SuperK data set from 1996 through 2001 recorded 1892 such events over 1679.6 live days. By counting the events within a few degrees of the directions of the center of the Earth, the Sun, and the galactic center, an upper limit can be placed on the rate of neutralino annihilations in each of these regions. These limits on rate are shown, and compared with similar results from other neutrino experiments, in Figure 1.8.

The SuperK upper limits can be compared to direct detection results using certain model assumptions. These include astrophysical assumptions about the distribution of WIMPs in the halo, as well as particle physics assumptions about nuclear form factors and scattering channels. In general, the SuperK analysis has used conservative assumptions. The combined results from the direction of the center of the Earth and the Sun have been used to give upper limits (see Fig. 1.9) on the cross-section of
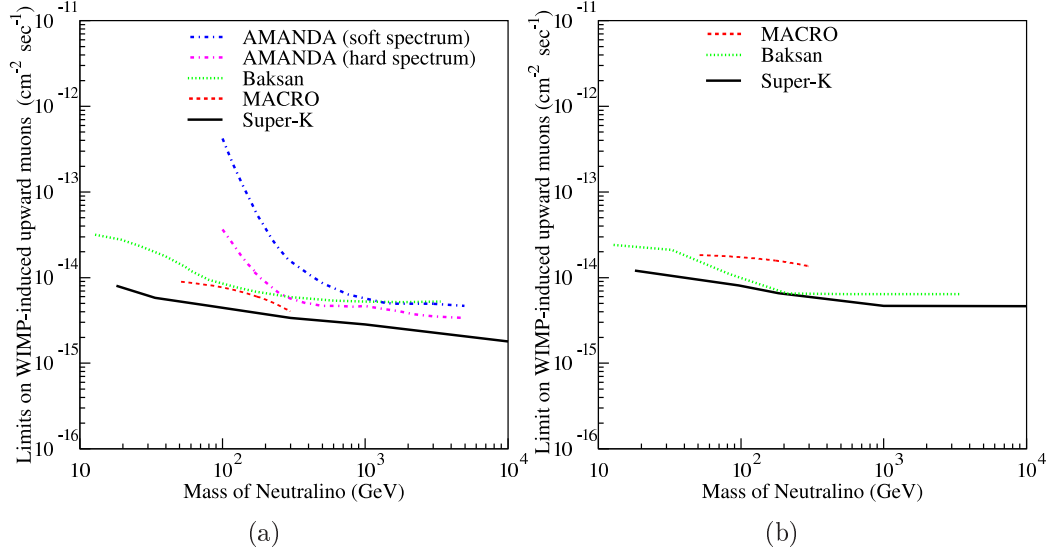
Figure 1.8: Limits from SuperKamiokande on neutralino annihilation rate in the Earth (a) and the Sun (b). Also shown are comparable limits from AMANDA, MACRO, and Baksan. Source: Desai *et al.* [39].
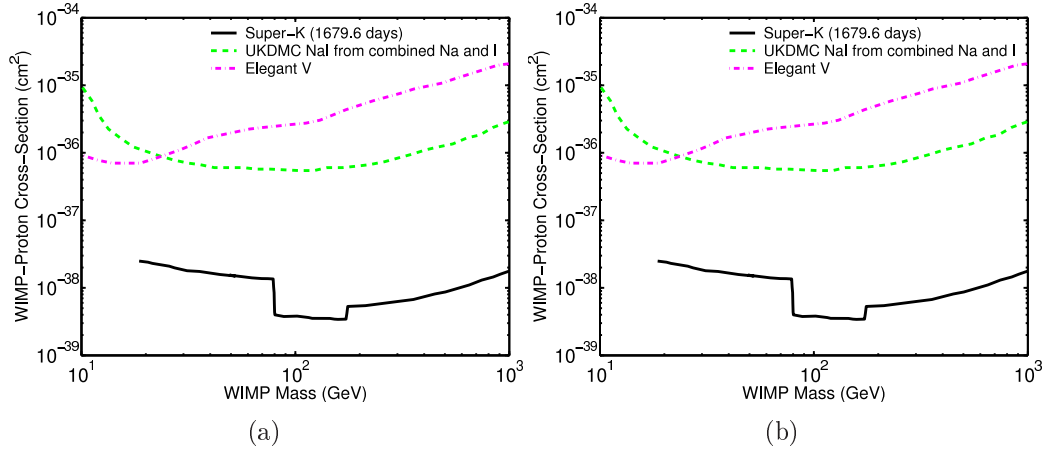
Figure 1.9: Neutralino exclusion limits from SuperKamiokande given spin-independent couplings (a) and spin-dependent couplings on protons (b). Also shown for comparison are (somewhat dated) limits from direct detection experiments. Source: Desai *et al.* [39].

WIMPs with scalar interactions, and the results from the direction of the Sun give upper limits on WIMPs that scatter on unpaired protons. The published analysis considers only WIMPs of mass at least 18 GeV. This restriction is not a fundamental limit on the sensitivity, but is for convenience in analysis; the muons corresponding to annihilation of lighter WIMPs will frequently stop within the detector volume, so extending the work to lower WIMP masses would require more sophisticated event selection [39].

The good indirect detection sensitivity of SuperK to WIMPs with spin-dependent coupling to unpaired protons is especially convenient, as it explores the one type of coupling that is not well covered by direct detection experiments with Ge and Si targets, such as the one described in the body of this thesis.

WMAP haze

The Wilkinson Microwave Anisotropy Probe (WMAP) has observed an unknown source of microwaves within 20° of the galactic center, known as the "WMAP haze." This excess signal is not associated with any known source, and is not well explained by standard types of microwave emission in the interstellar medium, such as dust or synchrotron radiation of electrons and positrons from supernovae. The haze has been interpreted as synchrotron radiation by electrons and positrons from WIMP annihilation in the galactic center [40].

Hooper, Finkbeiner, and Dobler [41] have found a good match between the spatial localization of the haze and WIMP annihilation for halo models with a steeper inner slope than the Navarro-Frenk-White profile, $\rho(r) \propto r^{-1.2}$. The strength of the haze signal in two frequency bands, at 22 and 33 GHz, allows an estimation of the haze spectrum and thereby the energy distribution of the emitting particles. It is found that $I_\nu \propto \nu^{-0.4}$, although the exponent is sensitive to the method of background subtraction. This is a significantly harder spectrum than other classes of synchrotron radiation, and implies a source of higher-energy electrons and positrons. The intensity and spectrum can be explained by annihilation of WIMPs heavier than ~ 50 GeV and with cross-sections on the order expected for a weakly interacting thermal relic,

$\sigma v \sim 3 \times 10^{-26}$ cm^{3}s^{-1}. The comparison of WMAP data with a WIMP annihilation model is shown in Fig. 1.10. If the haze is truly the result of WIMP annihilation, there is a good chance that GLAST will find a confirming gamma-ray signal for WIMP annihilation.

1.7 Direct detection

By contrast with the low elastic scattering rates for WIMPs calculated above, 0.1 kg^{-1}d^{-1}, background radiation abounds in a natural environment. There are 200,000 decays of ^{40}K per minute in a human body [42]; about one cosmic-ray muon per cm^2 per minute at sea level; and 100 or more decays of radon per second in 1 m^3 of indoor air. These event rates from the natural radioactivity of ordinary materials and from cosmic rays at the surface are vastly larger than the expected signal from WIMPs. The dark matter signal is also highest at relatively low energy, less than 100 keV. Therefore, the requirements for a direct detection experiment are to be sensitive to a WIMP signal with a low energy threshold and to somehow eliminate the large backgrounds. A number of approaches have been developed, most using some form of discrimination of electron recoil from nuclear recoil events. While WIMPs elastically scatter from nuclei in the target, almost all background interactions are with electrons.

Ionization and phonons

The experiment described in this thesis work, the Cryogenic Dark Matter Search or CDMS, rejects electromagnetic backgrounds using the different quenching factors for electron and nuclear interactions. For every event we measure a phonon signal, which gives the total energy, and the ionization signal, which is quenched by a factor of $2-3$ for nuclear recoils relative to electron recoils. In ionization yield, defined as the ratio of the ionization to phonon signals, and normalized to $y = 1$ for gamma scatters, electron recoil events (such as gammas) and nuclear recoil events (such as neutrons) form two distinct bands. The clear separation of the electron and nuclear recoil bands is shown in Fig. 1.11. By cutting on y we can reject $> 99.99\%$ of electron recoil events
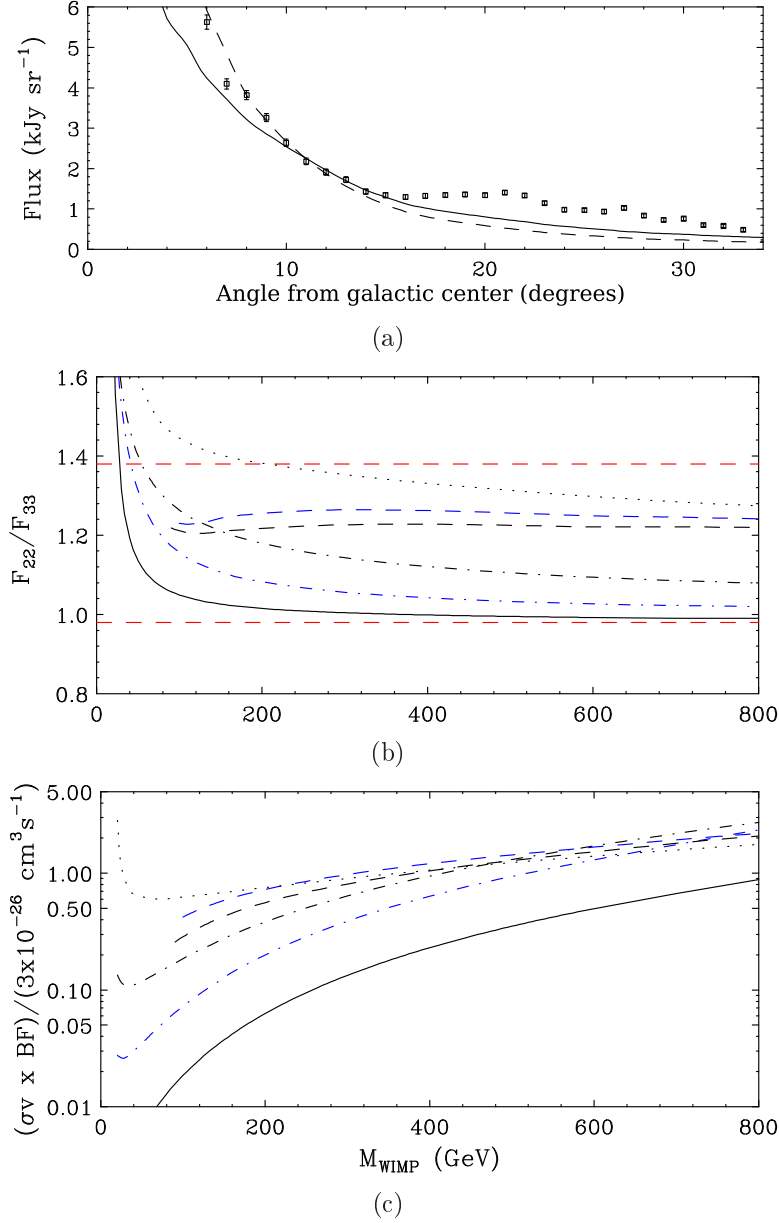
Figure 1.10: WIMP annihilation interpretation of the WMAP haze. In (a), flux at 22 GHz compared to models with NFW profile (solid) and steeper $\rho(r) \propto r^{-1.2}$ profile. In (b), ratio of 22 and 33 GHz fluxes compared with models. Horizontal dashed (red) lines indicate measured ratio. In (c), cross-sections for annihilation required to match observed flux. In (b) and (c), channels are: e^+e^- (solid), $\mu^+\mu^-$ (blue dot-dash), $\tau^+\tau^-$ (dot-dash), W^+W^- (dashed), ZZ (blue dashed), and $b\bar{b}$ (dotted). Source: Hooper, Finkbeiner, and Dobler [41].
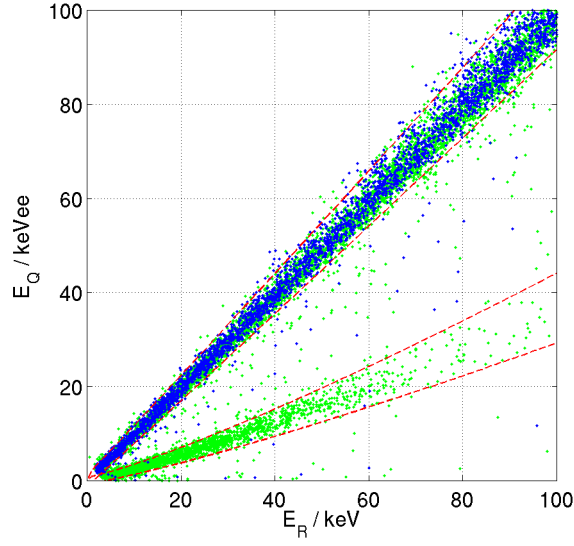
Figure 1.11: Discrimination of nuclear and electron recoils using ionization and phonons. The upper band contains scatters of gammas from ^{133}Ba, while the lower band contains neutrons from ^{252}Cf. The x axis is the total recoil energy calculated from the phonon signal, and the y axis is the ionization signal, calibrated to match the recoil energy for scatters on electrons.

in the bulk of the crystal. Chapter 2 describes the technical details of the ionization and phonon measurement using CDMS-II ZIP detectors, and Chapters 8-10 describe the analysis of the data taken in 2003 and 2004 in the deep underground site in Minnesota, including the analysis machinery for yield-based background rejection.

Chapter 2

The ZIP detector

As detailed in the previous chapter, the challenges for a dark matter direct-detection experiment are to achieve large mass, low energy threshold, and excellent background rejection in order to achieve good sensitivity to WIMP recoils. In order to meet these goals, the Cryogenic Dark Matter Search (CDMS) has developed the Z-sensitive ionization and phonon (ZIP) detector. The phonon signal produced by each interaction in a ZIP detector gives a sensitive recoil energy measurement with thresholds as low as 1 keV, and the simultaneous measurement of ionization allows discrimination against electromagnetic backgrounds. The advantage of ZIP detectors is the ability to collect fast, athermal phonons, which still retain information about where and when the initial particle interaction took place. Each ZIP detector has four phonon channels on its top surface, and the relative magnitude and timing of the four pulses, as well as their absolute delay relative to the very fast ionization signal, provide information about the physical location of the event. In particular, problematic shallow interactions at the detector surfaces can be effectively rejected.

CDMS-I used Berkeley large ionization and phonon (BLIP) detectors, using neutron-doped thermistors for the phonon measurement [43]. These were calorimetric detectors, in which position and timing information were lost as the recoil energy thermalized. The ZIP detectors of CDMS-II were developed as an improvement on BLIPs with better potential for surface event rejection.
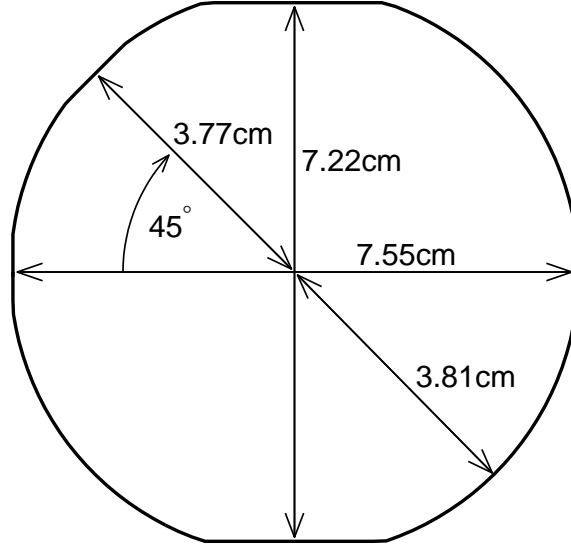
Figure 2.1: Geometry of a ZIP detector substrate as seen from the top, showing all flats. The major flats are at north and south, and the minor flats are at east and west. The small fifth flat is at 45° north of the west position, indicating a ⟨100⟩ crystal axis.

2.1 Crystal substrates

Each ZIP detector is a large, cylindrical crystal of germanium or silicon. The substrate is 1 cm thick and 3 inches in diameter (3.81 cm radius). The outer edge of the crystal is not a perfect circle, but has five "flats" in order to facilitate alignment and handling (as in Fig. 2.1): two major flats at the north and south positions, separated by 7.22 cm; two minor flats at east and west, separated by 7.55 cm; and a small fifth flat at the northwest, with a distance of 3.77 cm to the center. The exact position of the fifth flat indicates the orientation of the crystal axis. All ZIP detectors in use for the WIMP search so far have the top surface aligned with a ⟨100⟩ axis, and the fifth flat centered at 45° north of the west position. Any future detectors made with ⟨111⟩ substrates will have the fifth flat centered instead at 30° north of west. This is simply for identification; the fifth flat itself is not aligned with any particular crystal axis.

ZIP detectors can be made from either n-type or p-type substrates. All the detectors of Tower 1 and Tower 2 have been made with n-type crystals.

Table 2.1: Isotopic composition of natural Si and Ge.

Isotope	Abundance	p	n	Notes
^{28}Si	92.23%	14	14	
^{29}Si	4.67%	14	15	unpaired neutron
^{30}Si	3.10%	14	16	
^{70}Ge	21.23%	32	38	
^{72}Ge	27.66%	32	40	
^{73}Ge	7.73%	32	41	unpaired neutron
^{74}Ge	35.94%	32	42	
^{76}Ge	7.44%	32	44	

Differences between the physical properties of silicon and germanium lead to significant differences between the two types of detectors. The density of silicon is 2.33 g/cm^3, while germanium is much heavier at 5.34 g/cm^3. Each Ge ZIP detector has a mass of 250 g, while each Si detector weighs only 100 g. The A^2 enhancement of coherent scattering on nuclei means that Ge, with mean atomic mass 72.6, is strongly favored over Si (atomic mass 28.1) for spin-independent cross-sections of all but very WIMP light masses. The substrates are natural silicon and germanium, with isotopic compositions as shown in Table 2.1. All isotopes have even numbers of protons and neutrons except ^{29}Si (4.67%) and ^{73}Ge (7.73%), each of which has an unpaired neutron. These give CDMS-II its sensitivity to WIMPs with spin-dependent interactions.

The speed of sound is significantly faster in silicon than in germanium. This gives shorter phonon pulse delays and faster rise times in Si detectors. The band gaps and the average energy required to create an electron-hole pair are also different in the two materials. These differences will also have effects on the use of Si and Ge detectors in the dark matter analysis.

2.2   Ionization signal

Particle interactions in the crystal dissipate some of their energy by creating electron-hole pairs. This ionization process proceeds in competition with phonon emission; for electron-recoil events, as created by incident betas and gammas, the partition of

energy is characterized by a quantity ϵ, equal to the average incident energy required to create one electron-hole pair. This is larger than the band-gap E_g because some of the energy is shed as optical and acoustic phonons. These values inform the choice of bias voltage for our ionization channel, and are important for correct calculation of the recoil energy scale. See Appendix C for a review of current knowledge on the energy partition and epsilon numbers. We adopt $\epsilon_{Ge} = 3.0$ eV and $\epsilon_{Si} = 3.82$ eV, independent of energy over the range of interest to CDMS.

For scatters off of nuclei, as experienced by neutrons or WIMPs, ionization is smaller by a factor of ∼2-3. This quenching factor for nuclear recoils is energy-dependent and different for silicon and germanium targets.

The charge signal can be suppressed by a number of other effects, discussed in more detail in Chapter 5. In particular, the electrons and holes in the initial cloud can recombine if they are not separated quickly enough. This happens when the drift field is too low. As the field is increased, the charge signal asymptotically approaches a "full collection" value with no recombination.

2.2.1   Charge transport and Luke phonons

In the absence of any outside influence, the electrons and holes would gradually recombine. We apply a bias voltage of -3 V (Ge) or -4 V (Si) to create a drift field, under which the charges separate and migrate to opposite faces of the crystal. When the charges exceed the sound speed in the crystal they begin to shed energy to phonons, known as *Neganov-Trofimov-Luke* phonons, or simply Luke phonons.

Charge transport

In general, charge transport in pure crystals at low temperature is a complicated and incompletely understood phenomenon. A group at the University of Modena has collected experimental data and constructed a Monte Carlo simulation for electrons and holes in silicon and germanium, published over six years in a series of four monumental papers [44, 45, 46, 47]. They find drift velocities on the order 10 km/s at 3 V/cm, somewhat higher in Si than in Ge. Drift velocity generally increases at lower

temperatures and higher fields, although there are regions of negative differential mobility. Transport is strongly anisotropic, as discovered by Sasaki and Shibuya [48]. In particular, conduction of electrons occurs in band minima at the edges of the Brillouin zone, along the $\langle 111 \rangle$ direction in Ge and the $\langle 100 \rangle$ direction in Si. Somewhat counterintuitively, this results in larger mobility in $\langle 100 \rangle$ in Ge and $\langle 111 \rangle$ in Si. In the case of holes, the band minima are all in the center of the Brillouin zone, but the effective mass is not isotropic, with larger mobility along $\langle 100 \rangle$ in both cases. These results, however, go to temperatures no lower than 8 K in most cases, and are affected by both phonon emission and scattering on thermal phonons. In the case of CDMS at 50 mK, the thermal phonons effectively disappear, and only phonon emission and charge-charge interactions are important. A Monte Carlo simulation appropriate to these conditions is currently being developed by Kyle Sundqvist. Preliminary results show good agreement with the Modena group at 8 K, and in some cases predict faster drift velocities at 40 mK [49]. This simulation will allow the Luke phonon spectrum to be calculated as well.

Current ZIP detectors are made with $\langle 100 \rangle$ crystals for both silicon and germanium, so that the charges are drifted along the fast axis for Ge and an axis that is faster for holes, but slower for electrons, in Si. Some future test detectors will be fabricated on $\langle 111 \rangle$ crystals, with the expectation that this may change the distribution of Luke phonons as well as the focusing of primary phonons.

Luke phonons and Luke correction

Luke phonons can be emitted when the charges travel faster than the sound speed in the crystal; they are analogous to Čerenkov radiation or a sonic boom [50, 51]. As the charges travel, work is done on them by the drift field, and the energy is entirely shed as Luke phonons:

$$E_{Luke} = Q \times V = N_Q eV. \tag{2.1}$$

The observed charge signal is calibrated in units of energy "electron equivalent," that is, the energy of an electron recoil that would generate the same amount of ionization. For nuclear recoil events, this is 1/3-1/2 the true recoil energy. This is known in CDMS

as the ionization energy, $E_Q = N_Q\epsilon$. The Luke contribution to the phonon signal can therefore be rewritten

$$E_{Luke} = \frac{E_Q eV}{\epsilon}. \tag{2.2}$$

This remains true even when some charges are trapped before reaching the surface. As given in more detail in 5.1, the measured charge signal is proportional to the number of charges and the Ramo potential that they drift through. When some charges stop within the crystal, the loss to the ionization signal is in the same proportion as the loss to Luke phonons. Therefore, Eq. 2.2 is applicable to all classes of events.

2.2.2 Charge measurement

The top and bottom surfaces of each detector serve as electrodes for the charge circuit. The top surface is grounded[1], while the bottom is biased to a few volts. As the charges drift across the crystal, image charges are induced on the metallized top and bottom surfaces. According to Ramo's theorem, the charge induced is equal to the charge drifted, so that the integrated current flowing to either surface is a measurement of the ionization produced by the initial interaction. This is discussed more fully in Section 5.1.

The ionization side is divided into two separate electrodes: a disk-shaped inner electrode (85% by area) and a thin, annular outer electrode (15% by area). These are read out separately. The inner electrode defines a fiducial volume in which charge collection is good, and the outer electrode forms a "guard ring" to reject events that may suffer from degraded ionization signals. On each surface of the detector is a layer of amorphous silicon, which has a different band structure than the Ge and Si crystals. This suppresses the diffusion of charges of the wrong sign to the electrode for events occurring near the surface, ameliorating the surface beta problem.

Each ionization channel is biased through a large resistance, and capacitively coupled to a FET amplifier. This ensures that the electrodes are well coupled to the readout for fast pulses, and coupled only to the bias circuit over longer time periods.

---

[1]The top surface also contains the voltage-biased phonon sensors, but their V_{bias} is on the order of μV and is entirely negligible here.
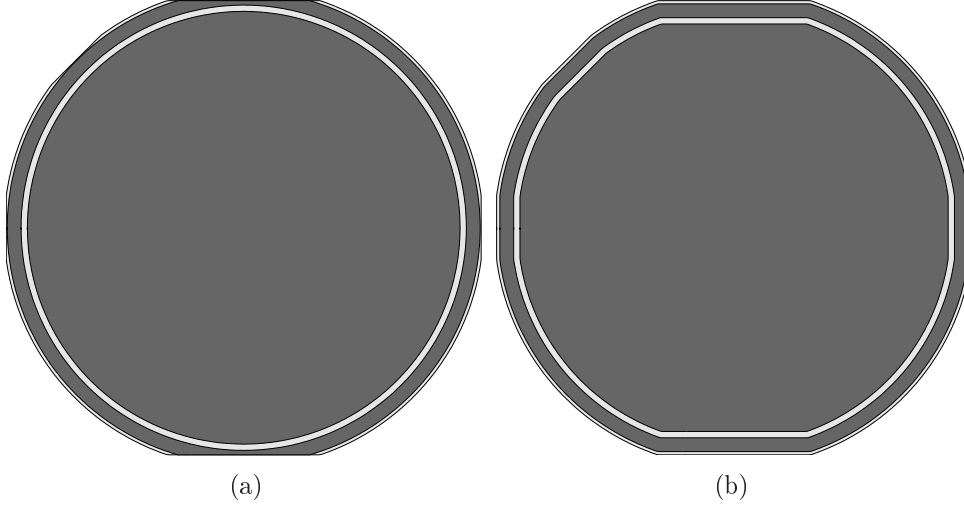
Figure 2.2: ZIP detector ionization side: old design (left) and current design (right). The old design has a very narrow outer electrode near the major flats. If the alignment of the photolithography is not perfect, there may be gaps in the outer electrode. In the new design, the electrode boundaries follow the flats, so that the outer electrode has a constant width. The old design is used in T1 Z1, Z4, and Z6. The new design is used in all others.

The bias and readout are thus decoupled. The ionization circuit is schematically shown in Fig. 2.3.

The charge bias line passes through a large (40 kΩ) bias resistor, which does not dissipate much charge over the time scale of a pulse. Each electrode is connected to a 300 pF coupling capacitor, which passes the fast charge pulses, but keeps the bias and readout circuits relatively floating. The amplifier contains a feedback capacitor of about 1 pF, which has the same charge induced on it as the feedback capacitor. The FET amplifier has its second input grounded, and adjusts the output voltage to remove the potential difference between its inputs. Thus, the very fast generation of image charges at the electrode is converted into a fast step in the output voltage of the amplifier. The feedback capacitor is in parallel with a feedback resistor of 40 kΩ, which slowly bleeds off the charge. The combination of feedback resistor and capacitor acts as a high-pass filter with time constant $\sim 40\mu$s. The ionization pulses therefore have a very fast rise, with rise time set by the amplifier response and the
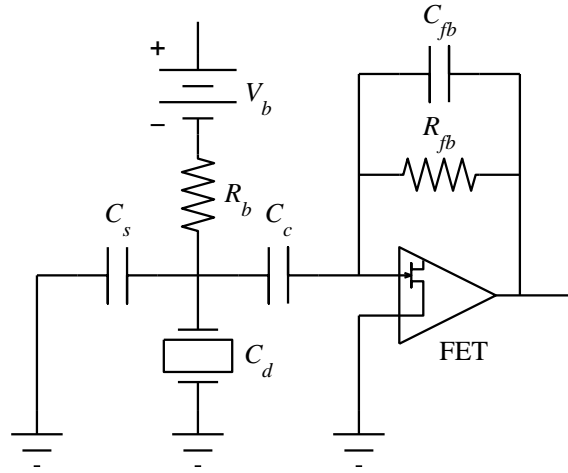
Figure 2.3: Ionization readout electronics. The ZIP detector has capacitance $C_d \sim 50$ pF, and is coupled to the amplifier through a coupling capacitor $C_c = 300$ pF. As an image charge is induced on the detector surface, the same charge appears at the feedback capacitor $C_{fb} \approx 1$ pF. The FET amplifier adjusts the output voltage to keep zero potential difference between its two inputs. The charge is bled off through the feedback resistor $R_{fb} = 40$ kΩ. There are also stray capacitances $C_s \sim 100$ pF. The detector is biased with a voltage V_b, through a bias resistor $R_b = 40$ kΩ.

antialiasing filter in the RTF boards, and a 40μs fall time set by $R_{fb}C_{fb}$.

A fraction of the charge signal is lost in the capacitive coupling. The charge induced at the feedback capacitor is less than the induced image charge by a factor $\frac{C_c}{C_c+C_d+C_s}$, where $C_c =$300 pF and $C_d \approx$50 pF – depending on the dielectric constant, $\epsilon =$11.8 (16.0) in Si (Ge), and on the geometry (with lower capacitance for QO than QI). The stray capacitance C_s incorporates several different components, and is harder to calculate. This includes the relative capacitance of QI and QO, capacitance to neighboring detectors, and to the detector housing. The reduction of the charge signal is acceptable, no worse than $\sim 50\%$.

2.2.3   Noise

Theoretical ionization noise is dominated by the voltage noise of the JFET amplifier and by Johnson noise in the feedback and bias resistors. In practice, there is also a very large contribution from pickup of ambient noise. This has been reduced by the use of an RF-shielded experimental enclosure (Sec. 3.1) and a careful grounding scheme, but it remains significant. Fortunately, the bulk of the noise pickup is in large peaks at frequencies above 100 kHz, while the ionization pulse have most of their signal power below 10 kHz. Therefore, the electronic noise can be very effectively filtered out, so that the performance of the ionization channel is limited only by its intrinsic signal-to-noise. Chapter 7 gives the history of observed noise levels during Soudan Run 118 and Run 119, along with a demonstration of the insensitivity of our event reconstruction to external noise.

An expression for the total voltage noise of the circuit as shown in Fig. 2.3 has been derived by Shutt [52]:

$$e_0^2 = |A(f)|^2 \left\{ e_{FET}^2 \left[(C_d + C_s + C_{fb})^2 (2\pi f)^2 + \left(\frac{1}{R_{fb}} + \frac{1}{R_b} \right)^2 \right] \right.$$

$$\left. + 4k_B T \left(\frac{1}{R_{fb}} + \frac{1}{R_b} \right) \right\}. \tag{2.3}$$

(a)   (b)
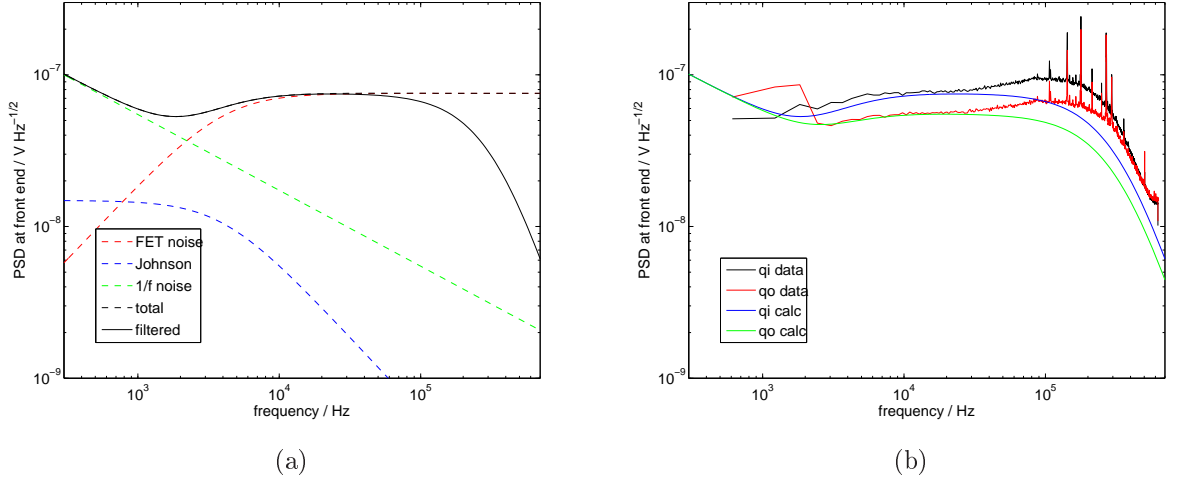
Figure 2.4: Predicted and observed noise spectrum of the ionization channel. In (a), the noise model of Eq. 2.3. All resistances and capacitances are as given in Sec. 2.2.2. A $1/f$ noise component has been added, with arbitrary amplitude. The filter in the final curve is a 336-kHz second-order Butterworth filter in the RTF board (See Sec. 3.6). In (b), a comparison with a typical data set for a Si ZIP detector. The outer channel has lower noise because of its smaller detector capacitance. The $1/f$ has been scaled to approximately match the data. Strong peaks from electromagnetic pickup are visible above 100 kHz.

The "signal shape" $A(f)$ is given by

$$A(f) = \frac{R_{fb}}{1 + 2\pi i f R_{fb} C_{fb}}.$$ (2.4)

The temperature for the Johnson noise term is that of the bias and feedback resistors, which are on the side coax board at 50 mK. The voltage noise of the JFET has been measured by Mandic to be typically 0.5 nV$/\sqrt{\text{Hz}}$ at 50 kHz [53].

2.3   Athermal phonon signal

When a particle interacts in the detector, athermal phonons are generated through several distinct processes with very different energy spectra and transport properties. The initially displaced nucleus or electrons shed their kinetic energy in the form of *primary phonons* at the interaction site. Drifting charges shed *Luke phonons* (Sec. 2.2.1). At the surface, the charges are trapped by a layer of amorphous silicon. They slowly tunnel across, at which time *recombination phonons* are emitted. The apportionment of energy among the various processes for electron recoils is given by a model due to Claude Klein, discussed in detail in Appendix C. Using C.3, we can predict the fraction of energy in each phonon population.

The primary phonons consist of high-energy, low-momentum optical phonons, and lower-energy acoustic phonons that come from the residual kinetic energy of the charges. The energies from primary and recombination phonons are

$$E_{optical} = e N_Q \bar{E}_R = \frac{E \bar{E}_R}{\epsilon}$$ (2.5)

$$E_{acoustic} = e N_Q \times \frac{9}{5} E_g = \frac{9 E E_g}{5\epsilon}$$ (2.6)

$$E_{recombination} = e N_Q \times E_g = \frac{E E_g}{\epsilon}.$$ (2.7)

Using appropriate values for Ge (Si), this gives 30.6% (14.6%) of the recoil energy in primary optical phonons, 44.6% (54.8%) in primary acoustic phonons, and 24.8% (30.5%) in recombination phonons. These sum to 100%, as all of the initial recoil

energy is in principle recoverable in phonons.

Luke phonons dissipate energy that was not deposited in the initial scatter, but has been added by the work that the drift field does on the moving charges. The contribution to the total phonon energy is given by Eq. 2.2. At a bias of 3V in Ge, or 3.82V in Si, the Luke phonon contribution is equal to the primary and recombination phonons. This informs our choice of 3V bias in Ge and 4V bias in Si.

The optical phonons have frequency 7-9 THz in Ge, and 12-16 THz in Si, while the primary acoustic phonons have energies of a few THz [54]. Luke phonons in Si have an average frequency of about 250 GHz under a drift field of ~ 4V/cm [55]. The case in Ge is similar.

2.3.1   Phonon transport

At 50 mK, phonon-phonon interactions are negligible. The athermal phonons are subject to two types of scattering, both of which become stronger with increasing frequency. Therefore, lower-energy phonons propagate more freely in the crystal.

Isotope scattering is Rayleigh scattering on isotopic impurities, with a ν^4 frequency dependence. It is a form of elastic scattering. For silicon and germanium, isotope scattering is independent of direction and phonon mode [56], with scattering time [57]

$$\tau^{-1} = A_0 \nu^4 \tag{2.8}$$

$$A_o \approx ca^3 \left(\frac{\Delta m}{\bar{m}}\right)^2 \frac{(2\pi)^4}{4\pi v^3}. \tag{2.9}$$

Here c is the concentration of mass impurities, a is the lattice spacing, $\Delta m/\bar{m}$ is the fractional mass difference, and v is the sound speed. Reported values of A_0 are 3.67×10^{-41}s^3 for Ge [58] and 2.43×10^{-42} s^3 for Si [56], corresponding to a mean free path of ~ 0.15 mm and 3.5 mm at 1 THz.

Anharmonic decay occurs when a higher-energy phonon splits into two lower-energy phonons. The frequency dependence is $\tau^{-1} \propto \nu^5$. This scattering process occurs primarily for longitudinal acoustic phonon modes, rather than slow or fast transverse modes. The rates for transverse phonons are very anisotropic, and can

approach the rates for longitudinal phonons only in certain directions . However, this does not mean that phonons in emitted in some transverse modes avoid down-conversion. It has been found that the frequent elastic scattering of phonons at frequencies capable of down-conversion mixes the several modes and randomizes directions sufficiently well that all modes and directions undergo anharmonic decay [59]. The rates for longitudinal acoustic phonons at low temperatures (3 K) and 1 THz are 1.23×10^5 s^{-1} and 1.62×10^6s^{-1} in Si and Ge [60].

For phonons below about 350 GHz (Ge) or 750 GHz (Si), the path length is as long as the thickness of the crystal, and the phonons can reach the surface without scattering [61]. These low-energy modes propagate ballistically at the speed of sound. Very high energy phonons, on the other hand, scatter very frequently. They quickly down-convert until they reach a "bottleneck frequency" around ~ 1.6 THz, in which isotope scattering dominates. In this regime, the phonons undergo many isotope scatterings and occasional anharmonic decays, so that the motion is a random walk with increasing step size. This mode of propagation is known as "quasi-diffusion." While ordinary diffusion is characterized by a travel distance proportional to the square root of time, in quasi-diffusion the increasing mean free path causes a ball of phonons to expand nearly linearly with time. The characteristic speed is about 0.14 times the sound speed in Si and 0.07 times the sound speed in Ge [62].

Ballistic phonons are the most prompt component of the phonon pulse, and are subject to phonon focusing [63]. The quasi-diffusive phonons form a long-tailed distribution and do not exhibit phonon focusing. There are also intermediate phonons with intermediate energies, below the "bottleneck frequency" but still scattering several times before reaching a surface. For electron recoils, the aggregate behavior of all these populations gives a pulse with the highest energy deposition at about 1/3 the sound speed in Si [64] and presumably somewhat later in Ge.

Only a very small fraction of primary phonons travel ballistically. For electron recoil events, the ballistic fraction has been bounded $\leq 25\%$, and Monte Carlo simulations predict $\sim 0.9\%$. For nuclear recoils, there is a possibility that phonon "hot spots" or hot electron-hole "droplets" could be formed because of the very dense energy deposition at the interaction site. These very localized hot regions would emit

thermal phonons with a significant fraction having low enough energies to travel ballistically, and high enough to be detected. The existence of a ballistic enhancement has not been established so far. The enhancement of ballistic phonons for nuclear recoils is in any case not large, $2 \pm 5\%$ of the total phonon energy [65]. Therefore, primary phonons are considered mainly diffusive for both nuclear recoils and electron recoils in the bulk of the crystal.

Recombination phonons are emitted at high energy, but at the surface of the detector. Because they can interact with the metal of the phonon and ionization sensors, they can quickly shed energy and are considered prompt when emitted. Therefore, recombination phonons and Luke phonons are considered ballistic.

2.3.2   Phonon measurement with TESs

The top surface of each ZIP detector is patterned with tungsten transition edge sensors (TESs) for collecting and measuring phonons. A TES is a small strip of superconducting material kept at the temperature of its superconducting-normal transition. Because this transition is very steep, a small change in the temperature of the metal causes a sharp change in its resistance. A TES is thus a very sensitive device for detecting any energy deposited into its electron system. Our TESs are tungsten (W) thin films with a T_c of 80-90 mK.

In order for a TES to be stably operated, it must somehow be kept at its transition temperature. This could be achieved by very precisely adjusting the temperature of the substrate, but this requires fine control, and only works if every TES on the same crystal has the same transition temperature. TESs have been successfully operated with a current bias [66]. the substrate is kept near the transition temperature. Because the resistance R is a steeply rising function of temperature, Joule heating $P = I^2 R$ causes a warmer portion of the TES to be heated even more. Therefore, in the current-biased regime, Joule heating must be kept small, and the energy deposited in the TES by an event is removed by thermal conduction. The interplay between the temperature of the TES and the applied power is known as electrothermal feedback (ETF) - in this case, unstable, positive ETF.
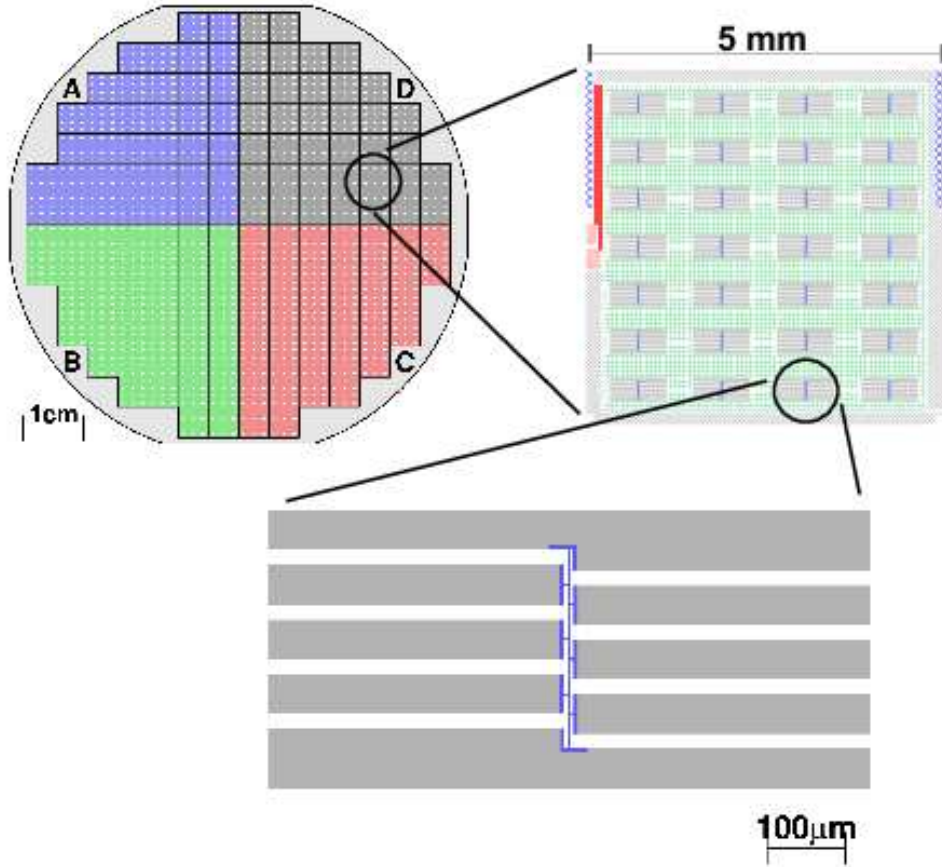
Figure 2.5: QETs on a ZIP detector. At upper left, the entire phonon side: each quadrant has 37 5mm×5mm dies. The area outside the stairstep boundary of the phonon sensors is covered with metal grid, and is electrically isolated in some detectors. At upper right, a single die with 28 TESs surrounded by Al grid (green crosshatch). The horizontal stripes of grid bias the TESs, with one stripe at the bias voltage and the next at ground. Around the periphery are bias rails (grey diagonal hatch), wire bonding pads (red horizontal hatch), and alignment marks (blue diagonal crosshatch). Bottom: a single QET, showing the Al absorber fins around the thin W TES.

A much more convenient scheme is the voltage-biased TES with negative ETF [67]. The power applied by Joule heating is now $P = V^2/R$. If the temperature rises, the increase in resistance causes the power to fall, returning to equilibrium. If the temperature falls, the power increases, and the TES again returns to the same equilibrium. Voltage-biased TESs are stable, and a large number of individual TESs can be biased in parallel with a single bias voltage. A TES with a higher T_c will automatically draw a larger current, and one with a lower T_c will draw a smaller current, so that both are stably maintained within their transitions. This is the approach adopted by CDMS-II.

The heat flow from the tungsten into the crystal substrate is given by $P = \kappa\left(T^n - T_s^n\right)$, where n has been measured to be 5 in similar devices. The steepness of the superconducting-normal transition is characterized by a dimensionless quantity $\alpha = \frac{T}{R}\frac{dR}{dT}$. The full energy of a particle interaction is pulled out of the crystal through the deficit in Joule heating and the increase in energy radiated into the substrate:

$$E_T = \int \left(\Delta P_J + \Delta P_S\right) dt. \tag{2.10}$$

If α is large, the W temperature remains almost constant during a pulse; and the crystal substrate is large, so that T_s is also nearly constant. Then, $\Delta P_s = 0$, and the full energy of the event is equal to the deficit in Joule heating.

Strong ETF makes the phonon pulses much faster: the fall time of a pulse is sped up from $\tau_{thermal} = \frac{C}{G}$ to $\tau_{ETF} = \frac{1}{\alpha}\frac{C}{G}$. For the tungsten films used in ZIP detectors, the transition width is as narrow as 1 mK at 70 mK, for $\alpha \sim 1/100$.

2.3.3 Phonon collection

Phonons in the crystal do not couple directly to the W films, but to an absorber, which in this case consists of fins of superconducting aluminum radiating outward from each TES (See Fig. 2.5, lower). The aluminum has thickness 3500 and an energy gap $2\Delta_{Al} = 0.36$ meV. A phonon reaching the Al with more than this energy can break a Cooper pair in the Al, creating quasiparticles. As the quasiparticles relax to the gap energy, they can break other pairs as well as emitting sub-gap phonon
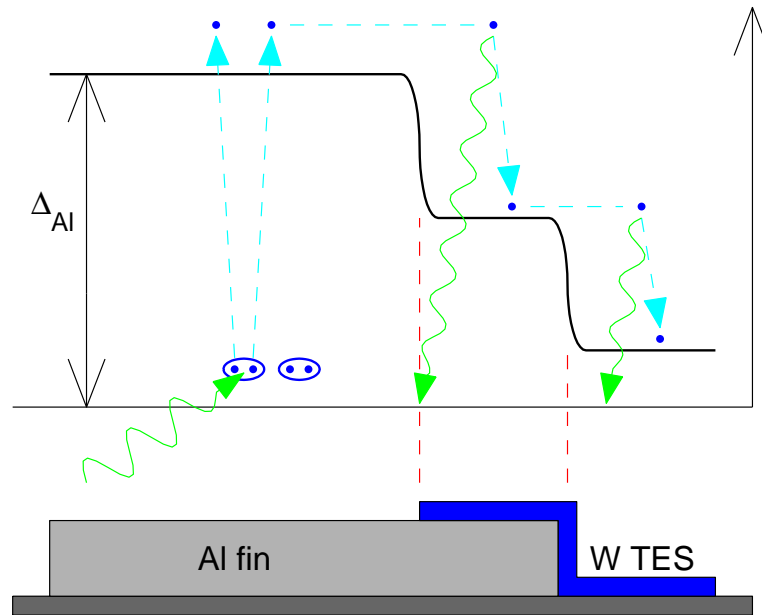
Figure 2.6: Quasiparticle-assisted phonon measurement. An incident phonon breaks a Cooper pair, creating quasiparticles above the Al gap. These diffuse into the W-Al overlap region, and emit phonons to relax to the smaller gap. They no longer have enough energy to migrate back into the Al fin and are trapped. The quasiparticles finally diffuse into the W TES and contribute to the observed pulse.

back into the crystal. The quasiparticles diffuse throughout the Al, until they reach the end near the TES. There is a zone of overlap between the Al and W, with a superconducting gap smaller than that of Al, but larger than that of W. This ensures that any quasiparticles that diffuse into the overlap are trapped and can only continue into the W, rather than going back into the Al. As the quasiparticles enter the W they raise its electron temperature, increasing the resistance. This overall design, using voltage-biased tungsten TESs with Al absorber fins, is known as the quasiparticle-assisted electrothermal feedback transition edge sensor (QPA ETF TES, or *QET* - an acronym of acronyms!). The process is illustrated in Fig. 2.6.

At 50 mK, thermal phonons have energy on the order of $k_B T = 4\mu$eV, or gigahertz frequencies. The thermal phonons cannot break Cooper pairs in the superconducting films as long as $E_{phonon} < 2\Delta$, where 2Δ is the superconducting energy gap of the film. According to BCS theory, $2\Delta \approx \frac{7}{2}k_B T_c$, which is 360μeV for Al and $\sim 24\mu$eV for our Tungsten films. The thermal phonons have too little energy to destroy Cooper pairs (but in the case of the W films, not by much). Indeed, this is precisely why Al and the W films are superconducting at base temperature, with the W not far from its transition. Thus, the QETs are not sensitive to thermal phonons in the crystal, in contrast to the thermistors used with BLIP detectors. ZIP detectors are instead sensitive to higher-energy, athermal phonons. The energy gap of the Al absorber fins allows phonons of energy ≥ 88 GHz to contribute to the observed phonon signal.

The QETs of a ZIP detector are arranged in four quadrants, each consisting of 1036 TESs wired in parallel. For convenience of fabrication, the phonon sensors are formed in square dies 5 mm on a side, with 28 TESs per die and 37 dies per quadrant. The sensor geometry is shown in Fig. 2.5. Each aluminum fin is 380 μm long, chosen to match the diffusion length of quasiparticles in our aluminum films.

2.3.4   TES bias and SQUID readout

The TES readout must convert a small current through the TES into a voltage that can be digitized. This is accomplished with SQUID amplifiers as shown in Fig. 2.7. Each quadrant is in series with a 250 nH inductor called the input coil, which generates

a magnetic field proportional to the current. The input coil sits on a washer that also houses a DC SQUID array. The voltage across the SQUID depends on the flux through it, in a way that is periodic and nonlinear. The SQUID's output voltage is amplified by an integrating amplifier, and returned to a second inductor on the same washer, the feedback coil. The feedback coil has only one turn while the input coil has ten turns, so that $L_i/L_{fb} = 10$. The amplifier adjusts its output voltage to cancel the flux through the SQUID, giving a closed-loop gain of 10. This amplifier is fast compared to the time scale of detector phonon pulses, so that the SQUID remains locked at the same flux, and the nonlinearity of its flux-voltage characteristic does not have a large effect.

A feedback resistor $R_{fb} = 1\mathrm{k}\Omega$ in series with the feedback coil converts the current signal to a voltage, so that $\Delta V_{out} = \Delta I_s \times 10 \times R_{fb}$. The input coil and the sensor resistance form a low-pass filter with time constant $\tau = L_i/R_s \approx 2.5\mu\mathrm{s}$.

2.3.5   Noise

The largest sources of noise in ZIP phonon channels are the Johnson noise of the TES and other resistive elements in the QET circuit, and the phonon noise in the TES. Johnson noise is caused by the thermal fluctuations of charge carriers in a resistor at finite temperature. For an ordinary resistor, the Johnson noise can be considered a voltage noise per unit bandwidth $v_n = \sqrt{4k_B T R}$ in series with the resistor, or equivalently, as a current noise per unit bandwidth $i_n = \sqrt{\frac{4k_B T}{R}}$ in parallel with the resistor. These expressions hold, with no frequency dependence, for all frequencies $f \ll \frac{k_B T}{h}$. This frequency is 1 GHz for a resistor at 48 mK, colder than any resistive element in ZIP detectors. Since CDMS digitizes at only 1.25 MHz, this cutoff can safely be ignored.

A TES in negative electrothermal feedback, on the other hand, must be treated in a different way. The TES bias current performs work on the voltage noise source (or equivalently, the TES bias voltage performs work on the current noise source), which must be included in the power load of the device. The result is that the Johnson noise of a TES is suppressed at frequencies below the ETF frequency. This was first
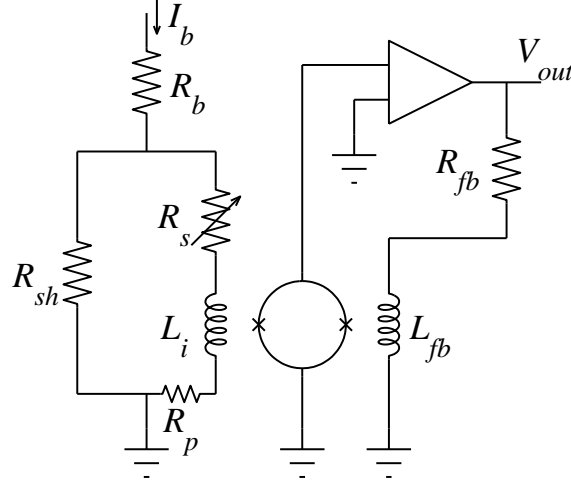
Figure 2.7: Phonon channel bias and readout electronics for one quadrant. The TES is shown as a variable resistance R_s ~100 mΩ. It is effectively voltage biased by being placed in parallel with a shunt resistor $R_s \approx 25$ mΩ. The parasitic resistance R_p is small, a few mΩ or less. This bias condition creates stable, negative electrothermal feedback. The current source is provided by a large resistor $R_b = 1$kΩ at room temperature. The input coil $L_i \approx 250$nH couples the TES current to the SQUID. The integrating amplifier adjusts its output voltage to cancel the flux through the SQUID, resulting in an amplification given by the turn ratio of the input and feedback coils, L_i/L_{fb}=10. The feedback resistor R_{fb}=1kΩ converts the current signal to a voltage.

pointed out for bolometers by Mather [68], and the expression for noise in the case
of a voltage-biased TES was derived by Irwin [67] in the form

$$I_{TES}^2 = \frac{4k_BT}{R_0}\frac{n^2/\alpha^2 + \omega^2\tau_{eff}^2}{1 + \omega^2\tau_{eff}^2} + \frac{4k_BT}{R_0}\frac{n/2}{1 + \omega^2\tau_{eff}^2}. \tag{2.11}$$

In this expression, the first term is the Johnson noise, and the second term is the
phonon noise (i.e. thermal fluctuations in the thermal link between the tungsten
electron system and the substrate).

The noise at the input coil resulting from the Johnson noise of the parasitic resis-
tance and the bias and shunt resistors is given by the following sum:

$$i_R^2 = \frac{4k_BT_p}{R_p}\left(\frac{R_p}{R_s + R_p + R_{sh}\left(1 + \frac{R_{sh}}{R_B}\right)^{-1}}\right)^2 + \frac{4k_BT_{sh}}{R_{sh}}\left(\frac{R_{sh}}{R_{sh} + (R_s + R_p)\left(1 + \frac{R_{sh}}{R_B}\right)^{-1}}\right)^2$$

$$+\frac{4k_BT_B}{R_B}\left(\frac{R_B}{R_B + (R_s + R_p)\left(1 + \frac{R_B}{R_{sh}}\right)^{-1}}\right)^2. \tag{2.12}$$

This expression can be simplified using the fact that the bias resistor is very large
compared to the other resistances, $R_B \gg R_s + R_p + R_{sh}$.

$$i_R^2 = \frac{4k_B}{(R_s + R_p + R_{sh})^2}\left(T_pR_p + T_{sh}R_{sh} + T_B\frac{R_{sh}^2}{R_B}\right). \tag{2.13}$$

The total current noise at the input coil can then be written down:

$$i_L^2 = \frac{4k_B}{(R_s + R_p + R_{sh})^2}\left(T_pR_p + T_{sh}R_{sh} + T_B\frac{R_{sh}^2}{R_B}\right) + \frac{4k_BT_{TES}}{R_s}\frac{n^2/\alpha^2 + \omega^2\tau_{eff}^2}{1 + \omega^2\tau_{eff}^2}$$

$$+\frac{4k_BT_{TES}}{R_s}\frac{n/2}{1 + \omega^2\tau_{eff}^2}. \tag{2.14}$$

This entire expression is rolled off at high frequencies by the self-inductance of the
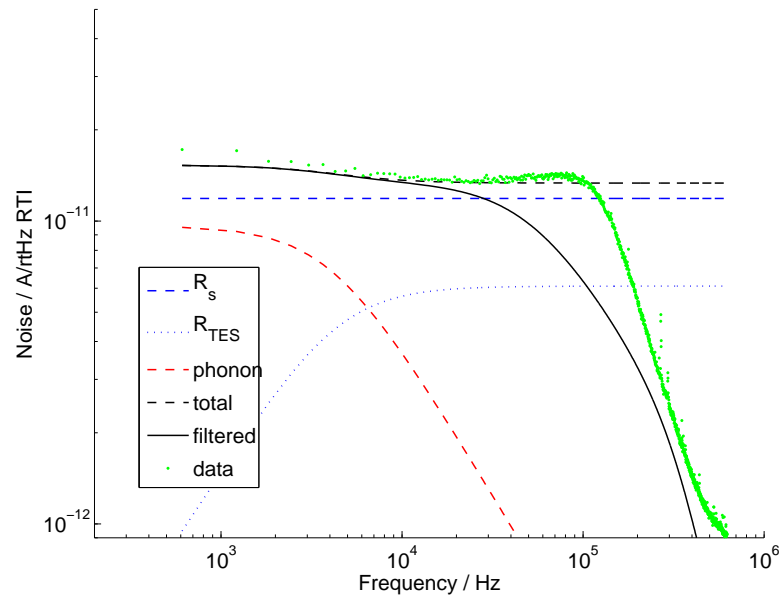
Figure 2.8: Predicted and observed noise in a ZIP detector phonon channel. The resistances in Eq. 2.14 are as measured by Dennis Seitz during SQUET testing, or in the case of R_s, as inferred using saturating pulses from muon events. The discrepancy at around 100 kHz indicates an unexpected resonant behavior of the SQUID amplifier, which extends the bandwidth beyond the L/R cutoff.

input coil, with a time constant L/R_s. The action of the RL circuit is a single-pole low-pass filter, with a gain

$$G_L(\omega) = \frac{\omega L}{\sqrt{R_s^2 + (\omega L)^2}}.$$ (2.15)

In the case of CDMS, some of the quantities in this expression are known well, while others can only be estimated. The sensor resistance is typically $R_s \sim 250$ mΩ in the biased state, with $T_s \approx T_c \approx 70 - 90$mK. These vary by detector and channel, depending on differences in fabrication, and on the bias current settings for each QET. The shunt resistor, $R_{sh} = 20$ mΩ, is at the 600 mK stage, while the bias resistor, $R_B = 1.2$ kΩ, is at room temperature (300K). The parasitic resistance is attributed mainly to pressure contacts (Millmax pins) in the cryogenic wiring, with a resistance of $R_p \sim 4$ mΩ at 4K. This gives a frequency-independent i_R^2 of 4.5 pA/$\sqrt{\text{Hz}}$. The contributions from the parasitic resistance and shunt resistor are comparable, while the contribution of the large bias resistor is negligible.

For thin-film tungsten TESs, $n = 3$, α is as high as 100 (depending on fabrication and bias point), and τ_{etf} is 25-140 μs.

2.3.6   Observed phonon noise

For comparison with observed noise spectra, it is convenient to group the total current noise at the input coil into two terms: one that is frequency-independent, and one that rolls off for low frequencies (slow compared to the ETF time).

$$i_L^2 = A + \frac{B}{1 + \omega^2 \tau_{eff}^2}$$ (2.16)

where

$$A = \frac{4k_B}{(R_s + R_p + R_{sh})^2} \left(T_p R_p + T_{sh} R_{sh} + T_B \frac{R_{sh}^2}{R_B} \right) + \frac{4k_B T_{TES}}{R_s},$$ (2.17)

$$B = \frac{4k_B T_{TES}}{R_s} \left(\frac{n^2}{\alpha^2} + \frac{n}{2} - 1 \right).$$ (2.18)
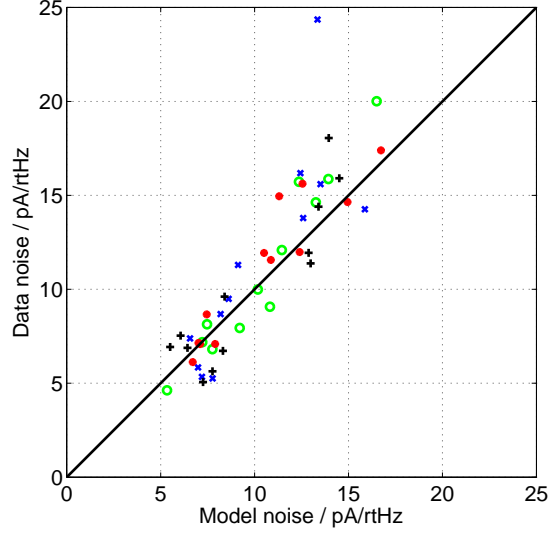
Figure 2.9: Correlation of predicted and observed phonon noise at 20 kHz. Each symbol represents one channel of one detector: PA (x), PB (o), PC (*), PD (+). All twelve detectors are shown, using noise traces from the data set 140808_1920.

Then, the difference between the noise at low and high frequencies is a measurement of B, and can be used to extract information about the bias condition of the QET.

The noise predicted by 2.14, rolled off by the L/R low-pass filter 2.15, may be compared to the observed noise power spectral density. The calculation and comparison are shown for one channel in Fig. 2.8. The agreement at low frequencies, where term B dominates, is not good, since the value of α for each sensor is not well known. The agreement at middle frequencies around 10 kHz is quite good. At higher frequencies, the observed noise spectrum typically does not roll off at the L/R frequency, but at a significantly higher point. In some channels, there is even a noticeable bump in the noise spectrum just below the roll-off. This indicates the presence of resonant behavior in the SQUID amplifier.

We have used the cold hardware resistances, measured by Dennis Seitz for each SQUET card, and the transition temperatures measured by the Berkeley and Case Western test facilities in calculating the expected phonon noise. The biased-state sensor resistance R_s has been calculated from saturating muon pulses [69]. We find

very good agreement between the predicted and observed noise at 20 kHz, as shown in Fig. 2.8.

2.3.7   Bandwidth

The intrinsic rise time and delay of a phonon pulse are modified by the bandwidth of the QET circuit and the SQUID amplifier. The current through the QET is inductively coupled to the SQUID, giving an intrinsic L/R time constant, where R is the resistance of the QET. Nominally, $L = 0.25~\mu$H, but this is modified by the coupling of the feedback coil to the input coil, so that a value of $0.35~\mu$H is seen in practice. Additionally, the RTF board includes an antialiasing filter immediately before the signal is digitized. This filter acts as a second-order Butterworth filter at 335 kHz. The actual bandwidth of the readout circuit is heavily affected by resonant behavior in the SQUID amplifier, so that the minimum rise time is typically faster than L/R ($\sim 1.5~\mu$s), but still slower than the frequency of the antialiasing filter.

2.3.8   Falling edge

When energy is quickly deposited into a TES, it must cool back to its equilibrium temperature by shedding heat into the substrate. The power dissipated from the TES is approximately constant. In equilibrium, it exactly balances Joule heating; when the TES warms up, the Joule heating is greatly reduced, so that the TES cools. This is the effect of negative electrothermal feedback (ETF). The thermal time constant of the TES is $\tau_0 = C/g$, while the ETF time is much faster [67]:

$$\tau_{etf} = \frac{\tau_0}{1 + \alpha/n}. \tag{2.19}$$

However, we find that the falling edges of pulses created by particle interactions have much longer fall times. This indicates that the physical arrival of energy at the TES is not all fast: there are long tails, either because some phonons continue to be absorbed in the aluminum fins at late times, or because some quasiparticles are trapped in the aluminum until late times.
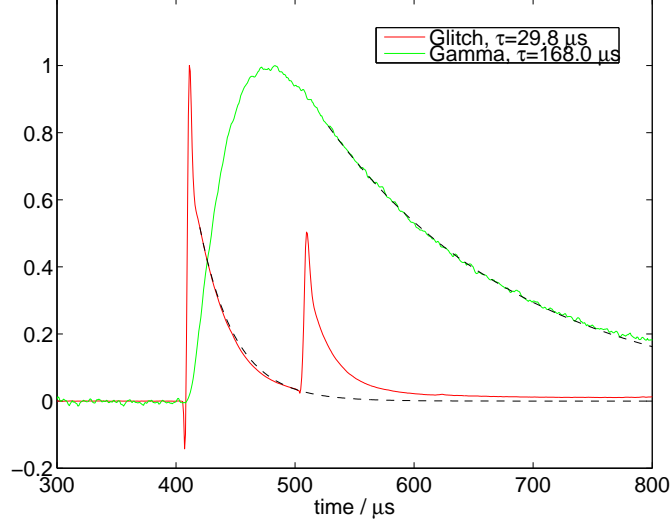
Figure 2.10: Pulse fall time and bare ETF time in phonon channel. The glitch event has two sharp, delta-function depositions of energy into the QET, and reveals the bare ETF time ~ 30 μs. The gamma event shows a slower rise and fall. Pulses shown are from quadrant A of T1 Z5.

The bare ETF time, unaffected by the long tails in energy arrival, may be estimated using noise glitch events. During Run 118, there were frequent noise events seen on all phonon channels of all detectors. These have a characteristic, very repeatable double-pulse shape. A first, very fast delta-function deposition of energy heats the TESs, after which they recover with the ETF time constant, until a second, somewhat smaller burst of energy occurs 100 μs later. The glitch is believed to come from external electronics, and heat the TES through Joule heating. There are very small additional depositions of energy between the two large spikes, and after the second spike, but these have a negligible effect on the fall time.

By averaging over many of these noise glitch traces and fitting exponentials to the falling tails, the true ETF time of each sensor of Tower 1 has been found. Because the glitch events were eliminated by improved grounding and electronic isolation in December of 2003, before the start of Run 119, they are not available for determining the ETF times of Tower 2 detectors. The fall time of a real gamma event and an averaged glitch are shown in Fig. 2.10 along with the time constants. The long tails

from late phonons or from quasiparticle trapping cause the fall time of a true pulse to be long, $\sim 150\ \mu s$, much longer than the $\sim 30\ \mu s$ ETF time. Understanding the long tails is one goal of the detector physics simulations currently in progress and to appear in the doctoral thesis of Matt Pyle.

2.4 T_c tuning

The transition temperature T_c of the TESs depends on the thickness and phase mixture of the tungsten films. Two different phases commonly appear in tungsten thin films: an α phase with body centered cubic structure, and a β phase with more complicated and less well understood A15 structure [70]. Although α-W has a low T_c of 15.5 mK, the long-range structure of β-W allows it to superconduct at higher temperatures, up to 4 K. Tungsten TESs operated at $\sim$100 mK must have a mixture of the two phases. This requires careful control of the metallization process and wafer preparation according to a recipe developed at Stanford [71].

The properties of the films often vary across the crystal, leading to a T_c gradient. The most common pattern is a T_c that is highest at the $+y$ edge and lowest at the $-y$ edge, with little variation in x.

After each detector is fabricated, it is cooled to below 50 mK in one of two CDMS test facilities, at Berkeley and Case Western. Each quadrant is characterized by measuring its $I_b - I_s$ characteristics: the variation of sensor current as a function of bias current at a fixed substrate temperature. This is equivalent to measuring the sensor resistance as a function of bias current.

At low bias, the sensor is fully superconducting, and there remains only a small parasitic resistance R_p. At sufficiently high bias, the sensor has its full normal-state resistance R_n. Between these two limits is the biased region, in which the resistance is intermediate and the TESs are held within their transition. A representative set of $I_b - I_s$ data is shown in Fig. 2.11. Based on the critical-current and $I_b - I_s$ data a model is built for the distribution of T_c among the dies in each quadrant. The comparison of data and model for a good Ge detector is shown in Fig. 2.12.

Fine control over T_c is achieved by implanting magnetic $^{56}Fe^+$ ions, which suppress
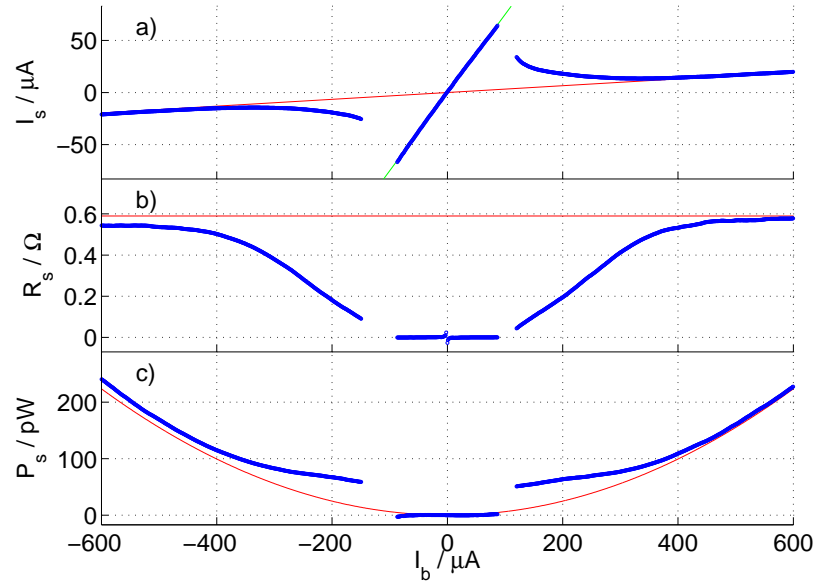
Figure 2.11: $I_b - I_s$ data for a silicon detector, S09 channel D. In (a) is the sensor current as measured. In (b) is the calculated resistance of the sensor. In (c) is the calculated Joule heating power of the sensor. The QET is in the superconducting state at low current, appearing as a diagonal line in (a) and a flat line near zero in (b) and (c). At high current, the tungsten is driven normal, with a constant normal-state resistance. This appears as a diagonal line with smaller slope in (a), a horizontal line in (b), and a parabola in (c). In between is the biased region. The shape of the power curve in the biased region indicates the T_c distribution of the sensor.
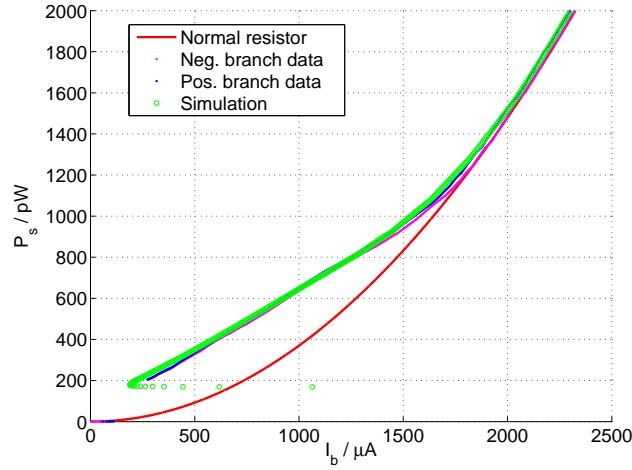
(a)
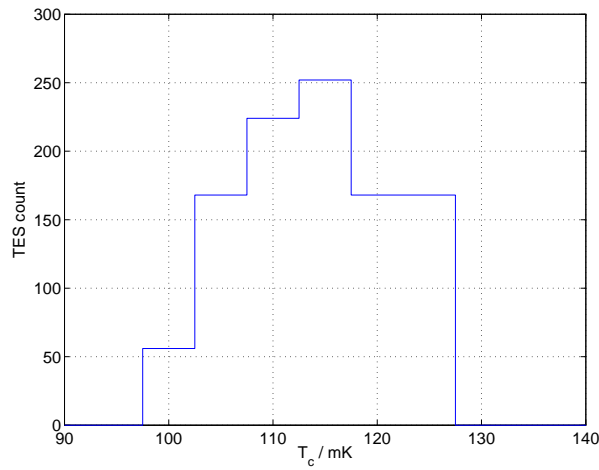
(b)

Figure 2.12: Sensor T_c distribution inferred from $I_b - I_s$ data for detector G33C channel C, taken at 60 mK before ion implantation. In (a), the output of a simulation is overlaid on the data; as I_b is swept from negative to positive, the QET passes through the bias region twice. Because the substrate temperature is well below T_c, there is substantial phase separation. In (b), the distribution of T_c for the TESs in the model. A later ion implantation reduced the T_c to a more uniform ~ 80 mK.

the transition temperature [72]. The implantation is performed in stripes oriented nearly parallel to the x axis, to match the T_c gradient and give uniform performance. A second run at one of the test facilities verifies that the post-implantation detector characteristics are as intended.

2.5 Calibration

The energy response of ZIP detectors is determined *in situ* using radioactive sources. We use a ^{133}Ba gamma source with strong lines at 275, 303, 356, and 384 keV to calibrate the ionization energy in Ge detectors. Because the Si detectors rarely contain the full energy of a gamma at these energies, the lines do not appear in Si ZIPs. Instead, we calibrate Si detectors using shared events: if a 356-keV gamma distributes its energy between a Si detector and an adjacent Ge detector, the sum should match the full 356 keV. Each of the Ba lines is therefore visible as a diagonal line with slope of -1 in a scatter plot showing the energy in each detector. We adjust the calibration in the Si detector to correctly place the shared energy line. The peaks are also visible in the phonon spectrum for Ge detectors, but with coarser resolution at high energy. Therefore, instead of calibrating the phonon response directly, we set ionization yield (charge / phonons) to one for gamma scatters. The spectra from the calibration source are also simulated in Monte Carlo, and give very good agreement with the data, as seen in Fig. 2.13.

The recoil energy scale for nuclear recoils should also be unquenched, but it is conceivable that the different phonon populations for electron and nuclear recoils cause a systematic difference in calibration. We can check this by comparing the spectrum from calibration with a ^{252}Cf neutron source to the predictions from a Monte Carlo simulation. The Monte Carlo has been performed in the doctoral dissertation of Sharmila Kamat [73], which is shown in Fig. 2.14 along with the Run 118 neutron calibration spectra. After calibration, the gamma band sits at $y = 1$ by definition, and the neutron band is at $y \approx 1/3$ (Ge, as in Fig. 2.15) or $y \approx 1/2$ (Si).

(a)

(b)   (c)
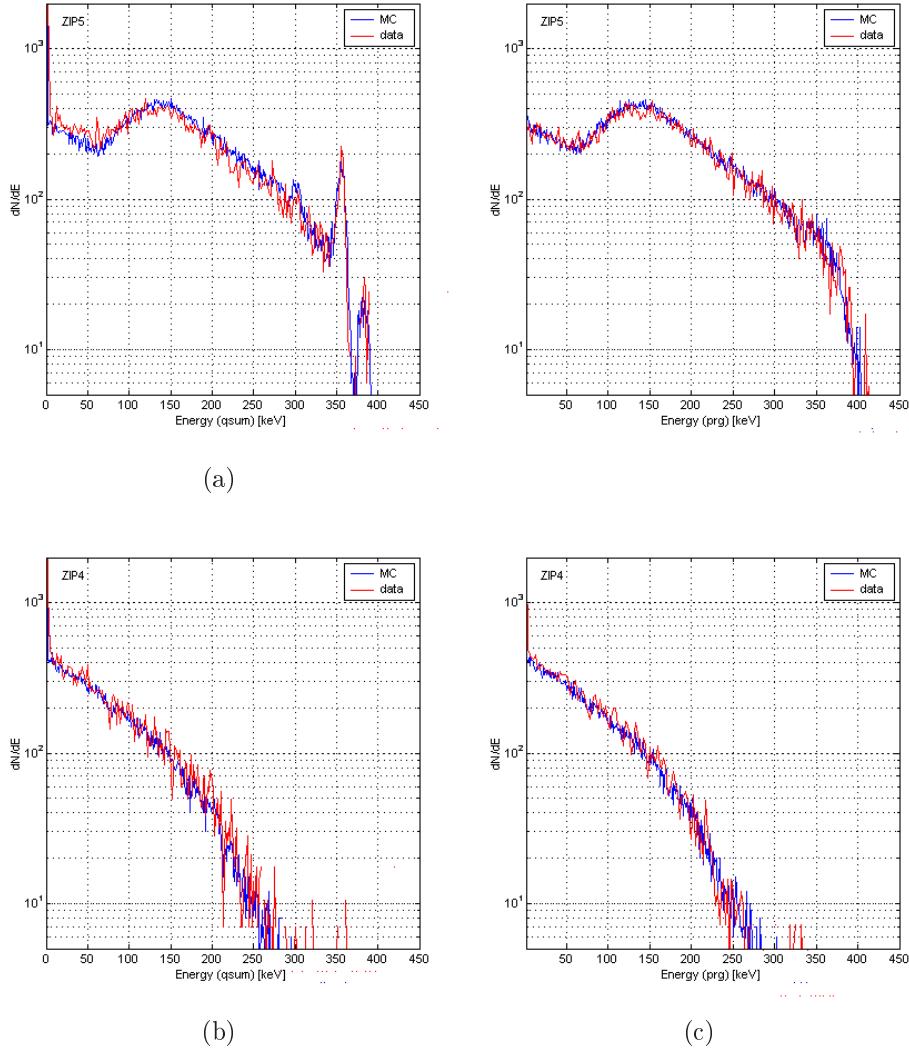
Figure 2.13: ^{133}Ba gamma spectra in Run 118 compared with Monte Carlo estimates. The top plots (a) and (b) show a Ge detector, T1 Z5, with prominent lines at 356 and 384 keV in ionization energy. The bottom plots (c) and (d) show a Si detector, T1 Z4, with no visible lines. The plots on the left, (a) and (c), are ionization energy, while (b) and (d) show phonon recoil energy. The agreement between data and Monte Carlo is very good. Source: L. Baudis.
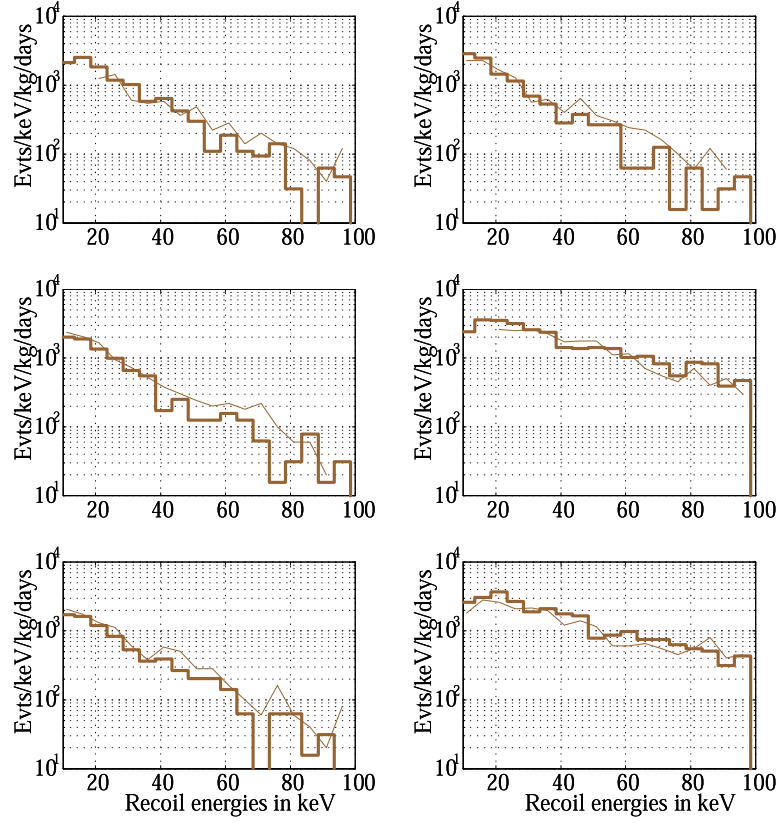
Figure 2.14: Comparison of observed ^{252}Cf neutron spectrum (heavy stairstep) with Monte Carlo (lighter line) for Tower 1 detectors in Run 118. The detectors are (top row) T1 Z1 and Z2, (middle row) Z3 and Z4, and (bottom row) Z5 and Z6. The good agreement demonstrates that the energy scale for nuclear recoils is not incorrect. Source: S. Kamat [73].
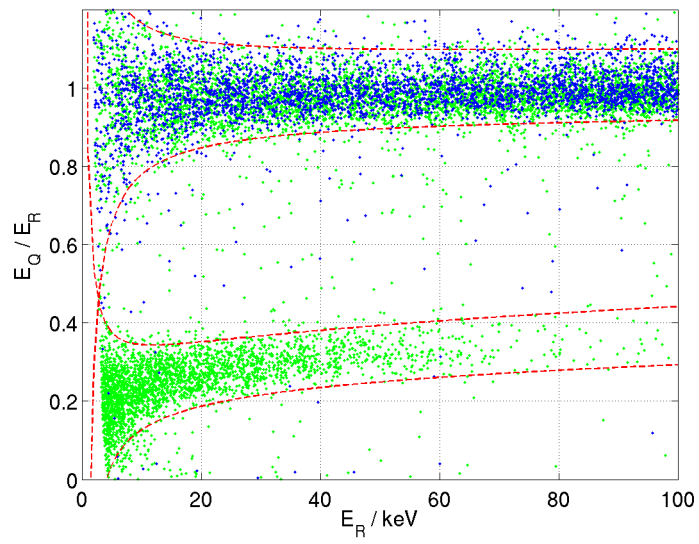
Figure 2.15: Gamma and nuclear recoil bands for a Ge ZIP detector.

Chapter 3

The CDMS-II experiment

3.1 Experimental sites

The first experimental site for CDMS was the Stanford Underground Facility, at a depth of 11 m [74]. The first version of the icebox and refrigerator, along with all passive shielding and the muon veto, were used for several runs with BLIP detectors through 1999 [43, 75, 76]. The first WIMP-search data run with ZIP detectors was Run 21, from August 2001 through July 2002 [77, 74, 78]. The shallow depth of SUF was sufficient to stop the hadronic component of cosmic rays, but there remained a high flux of muons (6 kHz through the veto shield). Most of the neutrons produced by muons were vetoed, but occasionally, a muon interacted in the walls of the cavern to produce an "external neutron" that passed through the shield and interacted in the detectors. These external neutrons constitute an irreducible background for SUF, as they cannot be distinguished from WIMP scatters on an event-by-event basis. However, they can be statistically subtracted, since some of them scatter in multiple detectors, and they produce more frequent above-threshold events in Si than in Ge (unlike all but the lightest WIMPs).

The second CDMS-II installation is in the Soudan Mine of northern Minnesota's Iron Range. The mine was Minnesota's oldest, deepest, and richest underground iron mine, operated by U.S. Steel until its closure in 1962. It is currently owned by the state Department of Natural Resources, and is open to the public as a state park.

The DNR and the University of Minnesota have formed a partnership to make part of the mine available to physics experiments as the Soudan Underground Laboratory. Current and past projects include the Soudan and Soudan II proton decay detectors [79], the far detector of the MINOS neutrino oscillation experiment [80], and low-background counting facilities. The laboratory is in Level 27, the deepest level of the mine at 2341 feet (714 m) below the surface. This corresponds to an effective depth of 2100 mwe. Access to the laboratory requires accompaniment by a qualified member of the laboratory staff, as well as an operator for the electric mine hoist. Although the DNR and the laboratory staff are very helpful in accommodating the exigencies of a cryogenic experiment, there is no regular access at night or on weekends. Outside of working hours, emergency access can be arranged within half an hour or one hour. Every effort has therefore been taken to make the operation of the experiment robust, automatic, and remotely controllable.

3.2   Cryogenics

ZIP detectors operate in the strong ETF regime, with substrate temperatures well below the tungsten T_c. Therefore, they must be enclosed in a cryostat that can achieve 50 mK or better over long periods of continuous operation. Cooling is provided by an Oxford Instruments 400S dilution refrigerator, which can apply 400 μW of cooling power at 100 mK. In order to keep the detectors in a radio-pure environment, the fridge is attached to an external "icebox" made out of counted copper (Fig. 3.5). The icebox consists of six nested copper cans, at room temperature, 77K, 4K, 600mK, 50mK, and 10mK nominal temperatures. These correspond to thermal stages of the fridge, and each can is attached to the appropriate fridge layer via a cold stem of nested copper tubes known as the cold stem or "C-stem." The innermost can is sunk to the mixing chamber, and achieves temperatures around 50mK when running with twelve detectors in the icebox.

The detector readout is through an electronics stem or "E-stem," opposite the C-stem. The E-stem brings the signal striplines from 4K to room temperature with heat sinking at 4K and 77K. It also provides gamma shielding with a copper block

that clamps around the striplines (without making thermal contact) as the E-stem
penetrates the lead shield. The end of the E-stem is a breakout box with 48 vacuum-
tight D connectors.

This design has modified the Stanford cryogenic setup as little as possible. The
one major change is that the E-stem and C-stem are attached to opposite sides of the
Icebox, while at SUF the E-box was immediately next to the fridge.

3.3   Cryogenic control

The cryogenic control systems allow the experiment to run in a stable mode within
the access constraints of the underground laboratory. Although the icebox and refrig-
erator closely match the SUF installation, extensive electronic control and monitoring
have been added. These are provided by several discrete systems that work together
to cover most aspects of cryogenic operation, as shown in Fig. 3.1.

The ^{3}He/^{4}He mixture circulation loop is controlled by an Oxford Intelligent Gas
Handling unit (IGH). The IGH communicates across a modified RS-232 serial link
with a Macintosh computer running virtual instruments in LabVIEW, partly from
Oxford and partly specially written for CDMS. The IGH Mac is equipped with remote
desktop software that allows it to be accessed from the surface building. The mixture
can be diverted into or taken out of the storage keg, and the circulation pumps, as
well as the pump on the 1K pot, can be turned on or off in this way. Heaters on
the mixing chamber, still, and helium bath are also controlled in this way (although
they are unused during ordinary running). The IGH also allows monitoring, but
not control, of fridge temperatures and cryogen levels. The IGH Mac publishes the
current values, trending plots, and long-term historical data to a web page (Fig. 3.2)
that can be monitored from anywhere.

Much of the cryogenic system communicates with a Moore APACS industrial
control unit. It continuously monitors temperatures in the refrigerator and icebox,
cryogen levels, pressures and pump speeds, and can automatically open and close
valves, initiate or stop cryogen transfers, and turn off pumps to prevent damage in
runaway situations. As long as dewars are attached to the transfer stations, APACS
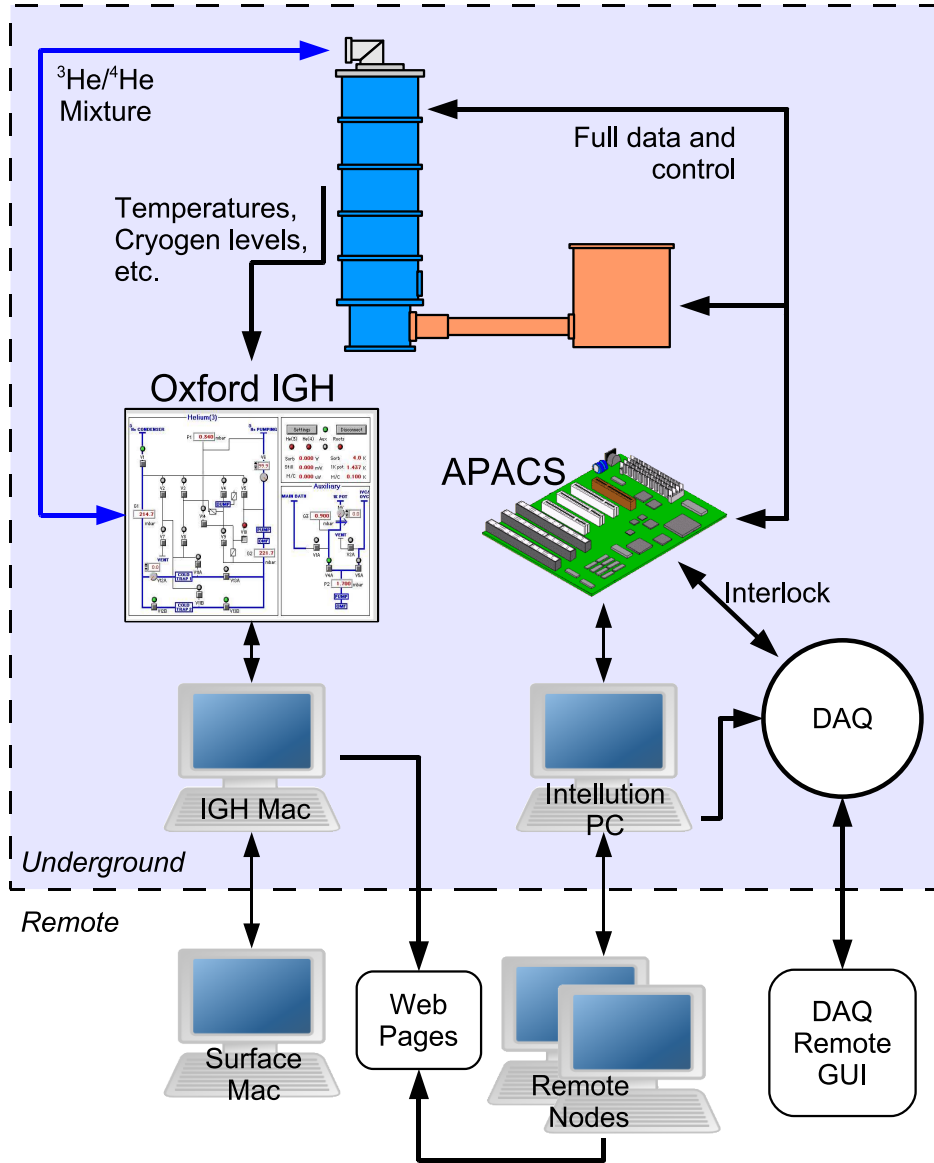
Figure 3.1: Cryogenic control system at the Soudan Underground Lab. Components inside the dashed box are physically inside the mine. Two-way arrows represent control and monitoring, while one-way arrows represent monitoring only. The remote Intellution nodes are the primary means of remote control, with the surface IGH Mac for controlling mixture circulation only. The web pages are the primary means of remote monitoring.
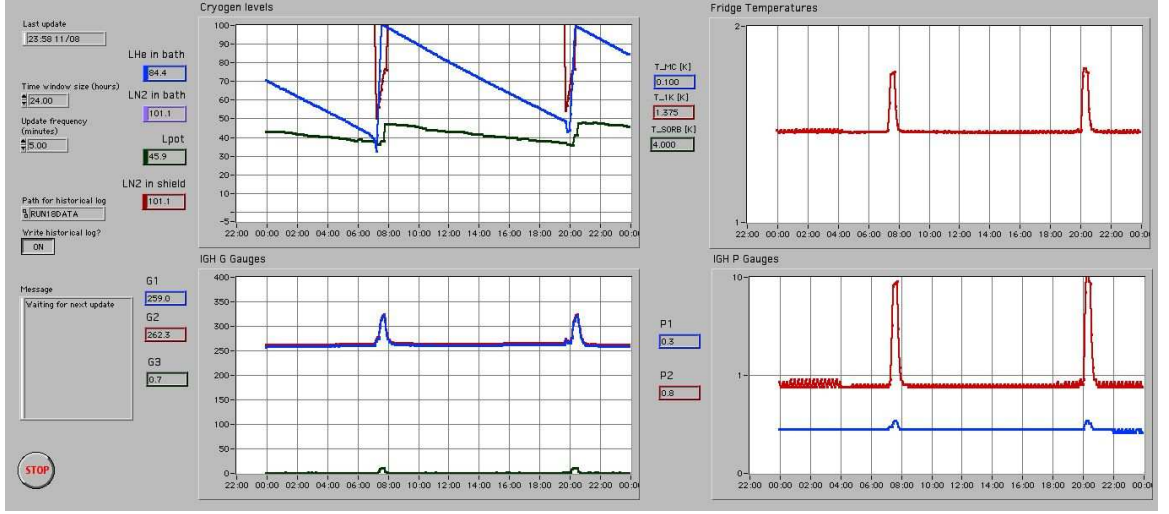
Figure 3.2: Trending monitoring plots from 2003, from IGH web page. Note two cryogen transfers daily.

can keep the helium and nitrogen reservoirs and the cold trap dewars full. These operations can be triggered by minimum level thresholds, or on a pre-set daily schedule. APACS can also adjust a needle valve to control the filling rate of the 1K pot. This is a coarse setting, used as a backup for the finely adjustable manual needle valve. These capabilities make it possible for the fridge to run over weekends with no human intervention, and to reach a safe state if anything unexpected should happen overnight.

A PC running Fanuc Intellution software is the interface to APACS, allowing manual control from the local PC or from Intellution nodes at the surface and at Fermilab. Monitoring (but not control) is supported through a web page. The Intellution PC also aggregates historical data and sends it to the DAQ for use in the DAQ's remote monitoring GUI.

The APACS hardware has its own channel of communication with the DAQ, used as an interlock. Before starting a cryogen transfer, APACS notifies the DAQ and waits for data taking to stop. After the transfer has finished, APACS signals the DAQ that it is safe to start taking data again.

The mine is subject to occasional power outages, especially on summer nights when

Minnesota has thunderstorms. Sudden loss of power could easily cause a fridge crash as the vacuums soften, the 1K pot warms up, and circulation stops; it is important to ensure that the automatic and remote control systems, as well as the remote monitoring, can survive an outage long enough for the crew to take action. The APACS cabinet, IGH, and the Intellution and IGH computers are all powered through uninterruptable power supply (UPS) units. The circulation pumps, vacuum pumps, and 1K pot pump have also been connected to UPSs since the beginning of Run 118.

3.4 Cold hardware

The icebox contains a set of cold hardware needed to mount and operate the detectors. These fixtures have been designed to minimize the heat conducted and radiated into the cold stages, and to allow the FETs to operate with suitably low noise. They have been constructed from screened low-background materials, in particular copper, kapton, and a custom low-activity solder. The design and construction of the cold hardware for CDMS-II have been overseen and documented with great care by Dennis Seitz of the UC Berkeley group.

The full cold hardware assembly is shown in Fig. 3.3. The detectors hang in the middle of the icebox can, in a "stack" of six; their electrical interface is through the detector interface board (DIB) and side coax board. Above the stack sits a heat-sinking assembly known as the "tower," and on top of the tower sit the cold hardware electronics on a FET board and a SQUID board. The signals are carried to room temperature in copper-kapton striplines, not shown in the figure. These components are described in more detail in the sections below.

3.4.1 Detector housing

Each ZIP detector is mounted in a hexagonal copper housing, held in place with three cirlex clamps on the top face and three on the bottom face. In contrast to a "coffin" design that would maximally isolate each detector, the ZIP detector housings are open at the top and bottom. There is a vacuum gap of 1 mm between adjacent
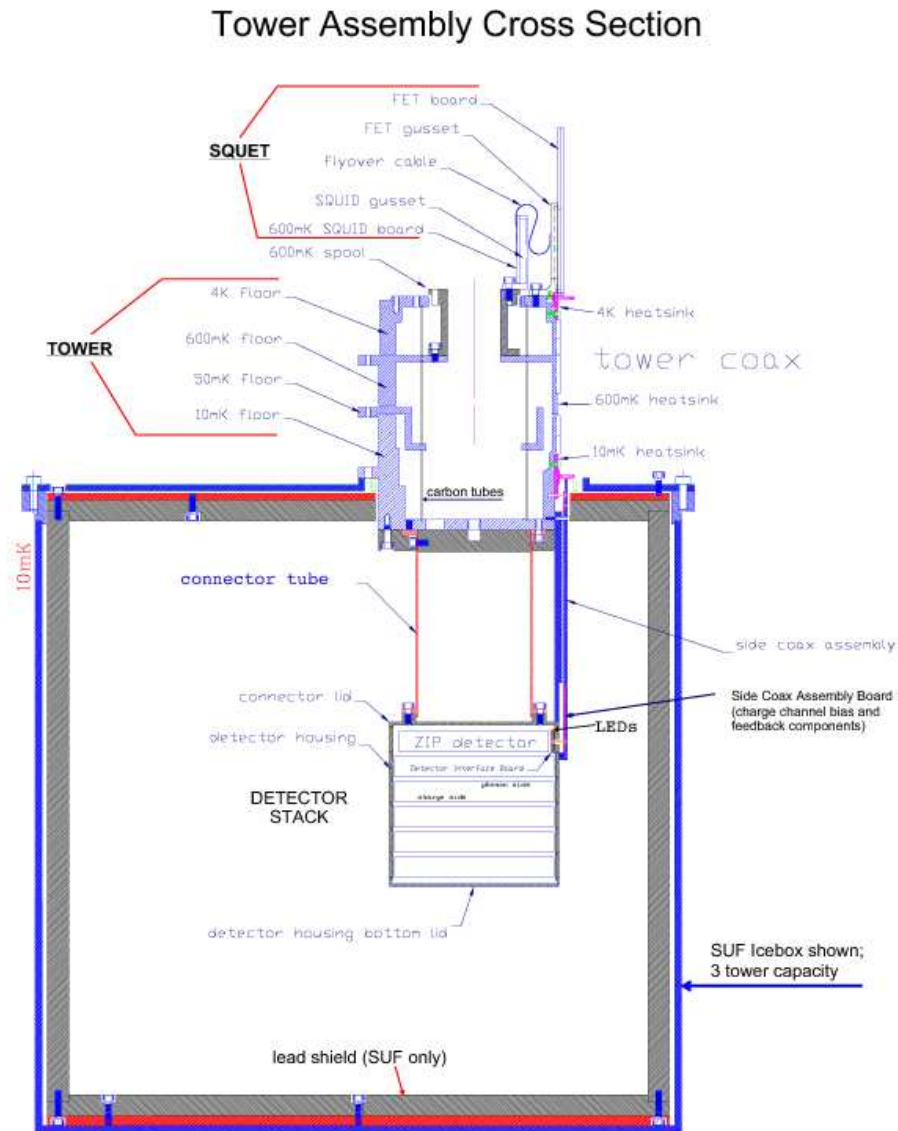
Figure 3.3: Overview of cold hardware, detector stack, and icebox can. The configuration shown is for the Stanford icebox. At Soudan the inner lead shield is absent, and the towers are placed closer to the edge of the can (with one exactly in the center). Diagram by Dennis Seitz.

detectors. This allows particles to multiply scatter in several detectors, which can be a powerful tool for studying and rejecting surface contaminant backgrounds. Lids are present only below the bottom and above the top detector of a stack.

The detector substrate's flats leave space between the housing and the edge of the crystal. The gap at one of the major flats is used to house the detector interface board. The DIB is mounted to the detector housing with two 0-80 screws, one at each end of the board. In order for the DIB to mate with the mill-max pins on the side coax board, a special slot is cut in the side of the detector housing at this position.

3.4.2   Detector Interface Board

The immediate connection to the ZIP detector is through the detector interface board, or DIB, shown in Fig. 3.4 (a). It is a printed circuit board that sits at the major flat in the $+y$ direction of each detector (as shown in Fig. 3.3 for the topmost ZIP). The DIB is wire-bonded to the detector and mates with the side coax board through mill-max connectors.

The DIB also contains the LEDs used to neutralize the detectors, shown in Fig. 3.4 (b). The position and orientation of the LEDs is important for understanding the effectiveness of neutralization. Because the LEDs are both located on top of pin 10 of the DIB, they are near the top surface (phonon side) of the detector, 6.2 mm to the right of the center. This corresponds to an angle of 10° clockwise from $+y$. One LED is placed physically in front of the other, so that the front LED looks out directly at the detector, while the back LED is partly shadowed. The intended convention for wire bonding is that the front LED be wired to pin 9 (LED 1), and the back LED be wired to pin 11 (LED 2). However, this has not been uniformly followed. DIBs that were wire bonded at Berkeley have LED 1 in front of LED 2, but DIBs that were wire bonded at Stanford have LED 2 in front of LED 1. Figure 3.4 (a) shows the "Berkeley convention." The LED assignment will be standardized for future cold hardware assembly, with LED 1 always closest to the detector.

The LEDs are Photonic Detectors PDI-E940 GaAs near-infrared emitters, with a peak wavelength of 940 nm. CDMS uses the bare, 0.4-mm cubical LEDs rather than
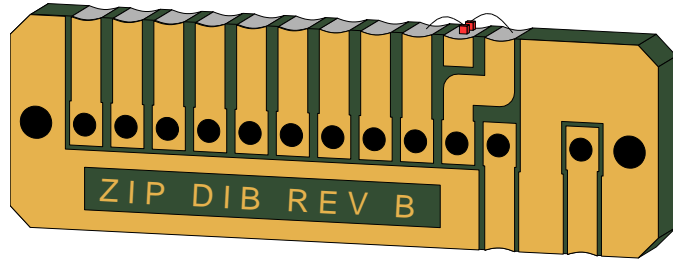
(a)
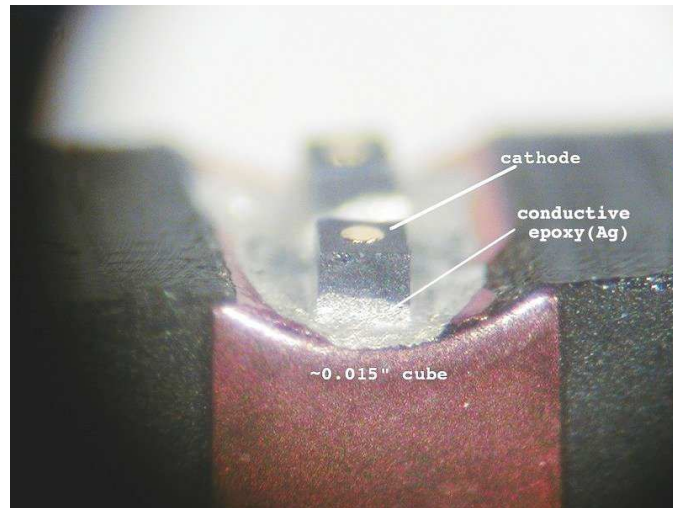
(b)

Figure 3.4: The DIB and LEDs. In (a), an entire DIB, with the LEDs above the tenth trace. The small holes are mill-max connectors that mate with the side coax board, while the two larger holes on either end mount the DIB to the detector housing. The tinned channels at the top and bottom are wire-bonded to bonding pads on the detector itself. In (b), a closer view of the LEDs (photo by Dennis Seitz). The three channels at the top right are used for the LEDs: both anodes are epoxied to the central pin, and the cathodes are wire-bonded to the pins on either side. One LED sits directly in front of the other.

the packaged surface-mount version. Emission is through the top surface (outside the circular gold cathode, visible in 3.4 (b)) and partly through the side surfaces. The emission angle is 70° half-angle about the *vertical*: the light is preferentially directed upwards, away from the detector itself, and towards the detector above. Therefore, the exact geometry of the LEDs, and the pin assignment, should have little influence on the effectiveness of neutralization. The standard approach to neutralization uses only LED 1, which is intended to be the nearest LED to the detector. Although some detectors have the two LEDs reversed, the direction of illumination means that there should be little difference between baking and flashing with the front or back LED.

In a few cases, the wiring or behavior of the LEDs are different from usual. The Si detector T1 Z6 has its LED 1 upside-down, with the cathode epoxied to the DIB instead of the anode. It can still be operated by using a negative voltage bias, but the light from the diode may shine mainly into the epoxy. This detector is especially difficult to neutralize.

3.4.3   Side Coax and Tower

Six detectors, along with their housings and DIBs, are assembled to make a "stack." Above the stack sits the "tower," a multi-stage wiring and heat-sinking construction that brings the detector electronics from the base temperature stage to the 4K stage of the icebox. Although the term "tower" strictly refers only to the heat-sinking assembly, the entire combination of six detectors plus cold hardware is often referred to as a "tower" as well. On top of the tower are mounts for the SQUET cards and SQUID boards that contain the cold electronics for each detector.

The connection between the mill-max sockets on each detector's DIB and the tower is made by the side coax board, which holds sixteen coaxial lines and a set of mill-max connector pins at each end. The side coaxes come in six different lengths, to attach to the six detector positions within the stack. The feedback and bias resistors for the ionization readout, as well as the coupling capacitor, are also housed on the side coax board, near the detector end.

3.4.4 Cold electronics

Some of the components of the phonon and ionization readout circuits, shown in Figs. 2.7 and 2.3, must be operated at cryogenic temperatures. The four SQUIDs sit on a small board at 600 mK. The FET and additional components are on a larger board connected to the 4K stage; the FET itself is only weakly heat-sunk, and self-heats to a toasty $\sim$130 K for optimal noise performance. The two boards are connected by a flyover cable, and the whole assembly is collectively known as a "SQUET."

3.4.5 Striplines

The signals are carried to room temperature through flexible, 2.5-cm-wide, 3-meter-long copper-kapton striplines. These are heat-sunk at 4 K and 77 K, and pass through a copper radiation shield as they proceed through the electronics stem and into a breakout box.

The cold end of a stripline attaches to a SQUET card, and the warm end attaches to a vacuum D connector in the E-stem breakout box. The striplines are flexible enough to turn over a radius of a few inches, but cannot easily cross over each other. Therefore, it is necessary to arrange and stagger the striplines with the correct lengths to reach their target SQUET cards with no crossings. This arrangement is fixed before the striplines are inserted, as the striplines are clamped into the thermal intercepts and a gamma radiation intercept before the entire bundle is pulled through the electronics stem.

In preparation for the five-tower run, the arrangement of the striplines was pre-planned with a full-scale paper and cardboard mock-up. This should be re-made before installing any additional detector towers.

3.5 Shield

To reduce backgrounds to an acceptably low rate, the icebox is surrounded by several layers of passive and active shielding, as shown in Fig. 3.5. Innermost is the copper of the detector tower and housing itself. This copper is thin ($\sim$few cm on average),

but is sufficient to stop alpha and beta radiation. Because betas are an especially problematic background, it is important that the detectors are exposed to betas only from sources on the detectors themselves and a small amount of immediately surrounding, clean material.

Outside the icebox is a lead shield to stop environmental gammas. The outer 18 cm are ordinary lead, and the inner 4.5 cm are ancient lead with no remaining ^{210}Pb. Outermost is 40 cm of polyethylene neutron moderator. The lead and polyethylene shields are shaped similarly to the icebox, with a cylindrical form and a removable lid. An active muon veto is used to reject (rare) events caused by cosmic ray muons passing through or near the icebox. The veto is composed of 40 panels of scintillating plastic, with photomultiplier tubes attached to them. WIMP candidate events are required to be in anticoincidence with veto hits. A full description of the Soudan veto can be found in the doctoral thesis of its principal architect, Raymond Bunker [81].

3.6   Data acquisition hardware

The interface to the detectors at room temperature is through the ZIP front end board (FEB). This is a 9U rack-mounted board designed and built at Fermilab. It applies the charge and phonon bias voltages; biases and locks the SQUIDs; zaps the SQUIDs to remove trapped flux when necessary; runs the FET heaters; and controls the flashing of the LEDs for space charge neutralization. The front end board contains the integrating amplifier that keeps the SQUID locked, and a front end amplifier for the ionization circuit. The final stage is an output driver on each channel, with a high pass filter at $\sim$ 1 Hz. A second "slow" output for each phonon channel is low-pass filtered instead, and retains only the offsets. The FEBs are housed in an electronics rack immediately next to the electronics stem.

From the FEB the signals are passed to a receiver-trigger-filter (RTF) board in the electronics room. The receiver adjusts all signals to be referenced to the RTF board ground, and creates a summed pulse for the four phonon channels, and one for the two charge channels. The RTF board uses the phonon and / or charge summed pulses to generate triggers according to adjustable thresholds. There are five distinct triggers
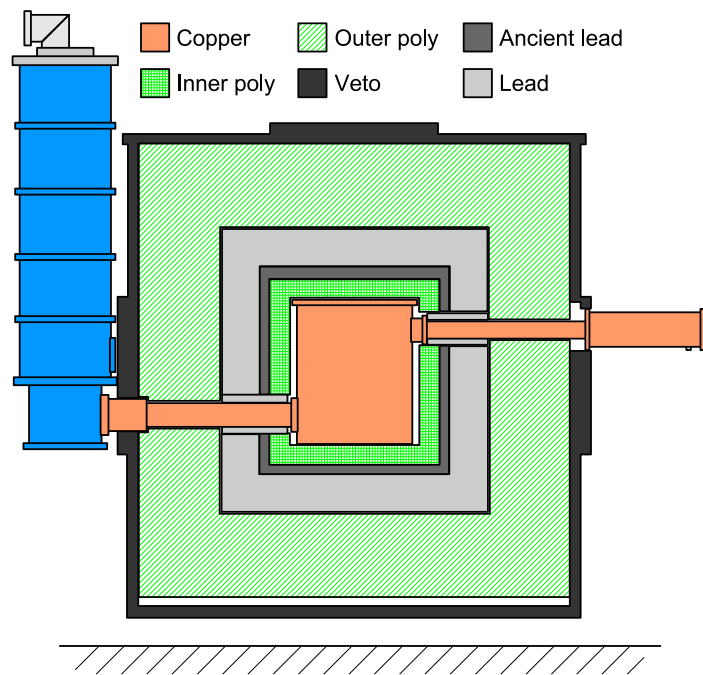
Figure 3.5: CDMS-II shield. The copper icebox is at the center, surrounded by inner polyethylene, ancient lead, outer lead, outer polyethylene, and the muon veto. The fridge is to the left, and the electronics stem is to the right.

per detector: a high and low threshold for phonons, a high and low threshold for charge, and a very low "whisper" trigger for phonons. The signals from the individual channels proceed from the receiver to the physics filter. This adjusts the offset of the signals in order to control the baseline voltage, applies a high-pass filter at 150 Hz to the charge channels only, and finally passes all signals through an antialiasing filter. This is the last stage before digitization. The antialiasing filter is nominally a 500 kHz low-pass filter, but simulation using Spice software shows it actually rolls off at 336 kHz.

The phonon and charge channels are recorded with Struck SIS 3301 digitizers. These have a 100 MHz sample rate, but perform on-board averaging in blocks of 80 samples to produce a signal at 1.25 MHz with true 14-bit precision. When a trigger is received, the digitizers record 2048 samples (1.6384 ms) for each channel. The trigger is at sample number 512, so that we acquire the trace for 410 μs before and 1229 μs after the trigger time. The slow offset signals are monitored and recorded only once per trigger.

A trigger logic board (TLB) aggregates the trigger signals from the RTF boards, and generates a global trigger whenever certain programmable conditions are satisfied. Usually, a phonon low trigger on any detector gives a global trigger. Global triggers can also be caused by hits in multiple veto paddles or by the DAQ software ("random triggers") to sample noise. The global trigger instructs the DAQ computers to read out the digitizers and trigger information. All triggers are stored in a trigger history buffer (SIS 3400) with 1 μs timing resolution, including those triggers that do not cause a global trigger. The original history buffer was capable of filling up and overflowing if it accumulated too many hits before a global trigger. In that case triggering information would be lost. In early 2004 it was upgraded to a ring buffer, which is not vulnerable to this problem.

The veto paddle traces are recorded by Joerger VTR812 digitizers, which take 1024 samples and operate at 5 MHz. Veto information is both stored as raw traces and run through a comparator to generate discrete hits, which are saved in a veto history buffer. The DAQ can be configured to trigger on simultaneous hits in multiple veto paddles (in the absence of any detector activity), for use in background studies

with the veto itself.

3.7 Data acquisition software

The CDMS-II DAQ software has two components: the DAQ servers, which run in the mine; and the DAQ graphical user interface (GUI), which can be used anywhere. Both software packages have been developed by the UC Santa Barbara group of CDMS, and are detailed in the doctoral dissertation of Joel Sander [82].

The local DAQ server software runs on six dedicated computers in the electronics room, running Fermi Scientific Linux: *tower1, tower2, vetocrate, builder, control,* and *monitor*. The first three machines receive the data stream from the digitizers for the detectors of Tower 1, Tower 2, and the muon veto respectively. *Control* communicates with the DAQ hardware across a GPIB interface, sending commands and settings to the FEBs, RTF boards, TLBs, and TCBs. *Builder* receives all event data from the hardware, including digitized traces and the history buffers, and forms CDMS-II raw event data structures that are written to disk. *Monitor* ensures that the DAQ hardware and other software components are working correctly, and communicates with the GUI.

An auxiliary computer in the mine, *datasrv*, runs a web server to publish information from the DAQ. The archived cryogenic monitoring information from the Intellution PC is kept here, as well as a table of completed data sets with operator comments and information about the run mode and number of events, a record of the experimental configuration, and noise spectra from each run. Once the data has been written to disk in the mine, *datasrv* writes a copy to digital tape, and sends a copy to the Soudan Analysis Cluster in the surface building. A second tape archive is made at the surface. Because the internet connection from the SAC to the outside world has limited bandwidth, raw data is sent to other institutions for future reprocessing in the form of tape archives. The raw data is processed and reduced with Dark Pipe at the SAC and other institutions (see Sec. 8.1).

The DAQ GUI is a self-contained program that can be run anywhere. It has been written in Java, and is compatible with any computer architecture and operating

system that has a Java virtual machine. Ordinarily, the GUI allows users to see the DAQ state, monitoring information, and trending cryogenic data. Users in this "observer" mode cannot control anything, beyond posting messages that are visible to all users. The DAQ GUI can also be run in a "trusted" mode from a few locations in the mine and at the surface. This allows runs to be started and stopped, and all run configurations to be changed. Operators taking shifts at Soudan can fully control the experiment through the GUI, which communicates with the servers using Java remote method invocation (RMI) and CORBA protocols.

Routine operations of the DAQ are automated, and require little or no operator intervention. Data taking is automatically stopped before a cryogen transfer begins, using the signal received from APACs, and restarted after the end of the transfer. The detectors and DAQ hardware are configured at the start of each run using stored settings. LED flashes are also performed periodically, in order to maintain good neutralization and ensure full charge collection. Shift takers use the GUI to change the experimental mode when taking calibration data with a radioactive source, to take test data, and to diagnose or fix problems with the experimental state.

Extensive remote monitoring is possible through the DAQ GUI. A noise viewer shows the noise power spectral density on each channel, using the 500 noise traces recorded at the beginning of every data set. This is routinely checked by the shift takers in order to make sure that the detectors are properly biased, and that no new sources of noise pickup have appeared (for instance, from a piece of electronic equipment that's not normally in use during data taking). There is also an automatic check of the noise spectra, used as a backup for direct human supervision. The DAQ also monitors the rate of triggers on each detector, the rate of veto hits, and the offsets on the phonon channels. Any change in these generates a warning, and may indicate that a channel is not configured properly, has a loose electrical connection, or has lost SQUID lock and needs to be reset. In most cases, such problems can be resolved by stopping the data run and restarting, since this repeats the entire detector and DAQ configuration procedure.

Besides the DAQ monitoring tools, there is a separate fast analysis program, written in Root, which generates additional diagnostics. These include box and delay

plots, energy spectra, and yield bands, and are often helpful to the shift takers in diagnosing and fixing any problems.

The CDMS-II DAQ at Soudan can take data at a high rate, with a dead time as low as 26 ms (38 Hz). The trigger rate during low-background data is low, less than 1 Hz, but in calibration mode we take data at 10 Hz or above. This is useful for accumulating large statistics in the gamma calibration, but is most important for *in situ* neutron calibration. Exposure to neutrons activates the Ge detectors, transmuting some ^{70}Ge atoms into unstable ^{71}Ge. We therefore desire to minimize the time spent with the neutron source inserted, and to take neutron calibration data at as high a rate as possible.

3.8 Commissioning

The CDMS-II experiment at Soudan reached stable operation through the concerted effort of many people. A large part of this doctoral thesis work was spent in the commissioning effort from 2002 through 2005. The author's first shift at Soudan was in June 2002, with 25 trips between then and February 2005. A very complete record of the commissioning process may be found in the thesis of another graduate student who was instrumental in making the experiment work, Vuk Mandic [53].

Chapter 4

ZIP position sensitivity

ZIP detectors were designed to provide a very rich stream of information for every interaction. The quasiparticle-assisted TESs respond quickly to athermal phonons, with rise times on the order of tens of microseconds. The sound speed is about 5 mm/μs in Ge and 8 mm/μs in Si[1], so that the shapes and start times of the phonon pulses contain information about the types of phonons absorbed and the distances they traveled before reaching the sensors. Because the phonon sensors are segmented into four quadrants, the partition of energy among the four channels and their relative delays allow the x-y position of events to be estimated in two independent ways. The good performance of the straightforward position reconstructions in the inner few cm of the detector has been demonstrated using collimated sources [84], but here we focus on the badly-behaved regions where reconstruction becomes very difficult.

This and subsequent chapters will use a coordinate system fixed to the ZIP detector, with its geometry shown in Fig. 2.5. The depth z is defined to be zero at the top (phonon) surface, and 1 cm at the bottom (charge) surface. The positive x axis is defined by the line from the center of the detector along the boundary between quadrants C and D, and the positive y axis is the boundary between quadrants A and D. These will often be converted into polar coordinates r and θ where convenient. These detector coordinates do not have the same true orientation in space for

---

[1] At 20 K. The values at 50 mK are similar, with different group velocities for the transverse and longitudinal modes.
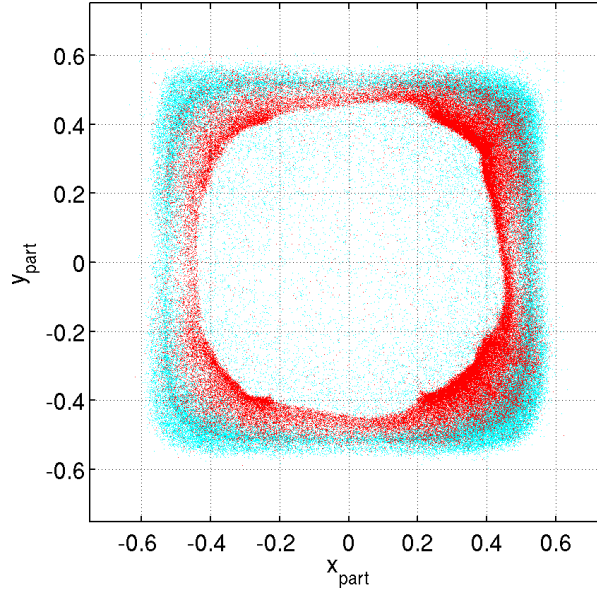
Figure 4.1: Box plot for T1 Z5. Cyan (light) points are inner electrode events, while red (dark) points are outer and shared events.

all detectors. Starting from the topmost detector of a tower and going down, each detector is rotated by 60° relative to the one above it.

4.1   *x-y* position

The first position reconstruction comes from the partition of energy among the four phonon quadrants. The two coordinates are

$$xppart \quad = \quad \frac{pc + pd - pa - pb}{pt}$$
$$yppart \quad = \quad \frac{pa + pd - pb - pc}{pt}$$

They can also be expressed as polar coordinates, useful for the cylindrical ZIP

detectors.

$$
\begin{aligned}
rppart &= \sqrt{xppart^2 + yppart^2} \\
thetappart &= \arctan^{-1}\left(\frac{yppart}{xppart}\right)
\end{aligned}
$$

A plot of these two partition coordinates is also known as a "box plot" because of its characteristically square shape. Partition is a good estimator of position near the center of the detector, but farther out it begins to undergo some pathologies. The events closest to the outer edge of the detector, with the highest physical radius, do not have the highest values of $rppart$, but instead "fold back" to smaller values. The density of points is very high at the outer edge of the plot, representing a second pathology: "pile-up" where $rppart$ is not very sensitive to changes in physical radius. The foldback and pile-up are clearly visible in Fig. 4.1, where the outer-electrode events occur closer to the center of the plot than many inner-electrode events.

A second position reconstruction comes from the relative delays of the pulses in the four phonon quadrants. We use a start time defined by the point at which the pulse in each quadrant reaches 20% of its peak value, $P?r20$. The two coordinates are

$$
\begin{aligned}
xdel &= \begin{cases} PAr20 - PDr20, & \text{for events in quadrants A or D} \\ PBr20 - PCr20, & \text{for events in quadrants B or C} \end{cases} \\
ydel &= \begin{cases} PBr20 - PAr20, & \text{for events in quadrants A or B} \\ PCr20 - PDr20, & \text{for events in quadrants C or D} \end{cases}
\end{aligned}
$$

The polar coordinates $rdel$ and $thetadel$ are defined in the same way as their partition equivalents.

A plot of these two delay coordinates is known as a "delay plot." Like partition, delay has a very clean response near the center of the detector, but begins to pile up and fold back somewhat at large radii. Fig. 4.2 illustrates the properties using outer-electrode events; it also shows the difference in delay caused by the choice of crystal substrate. Delay plots in Si (right) are much smaller than those in Ge (left) because
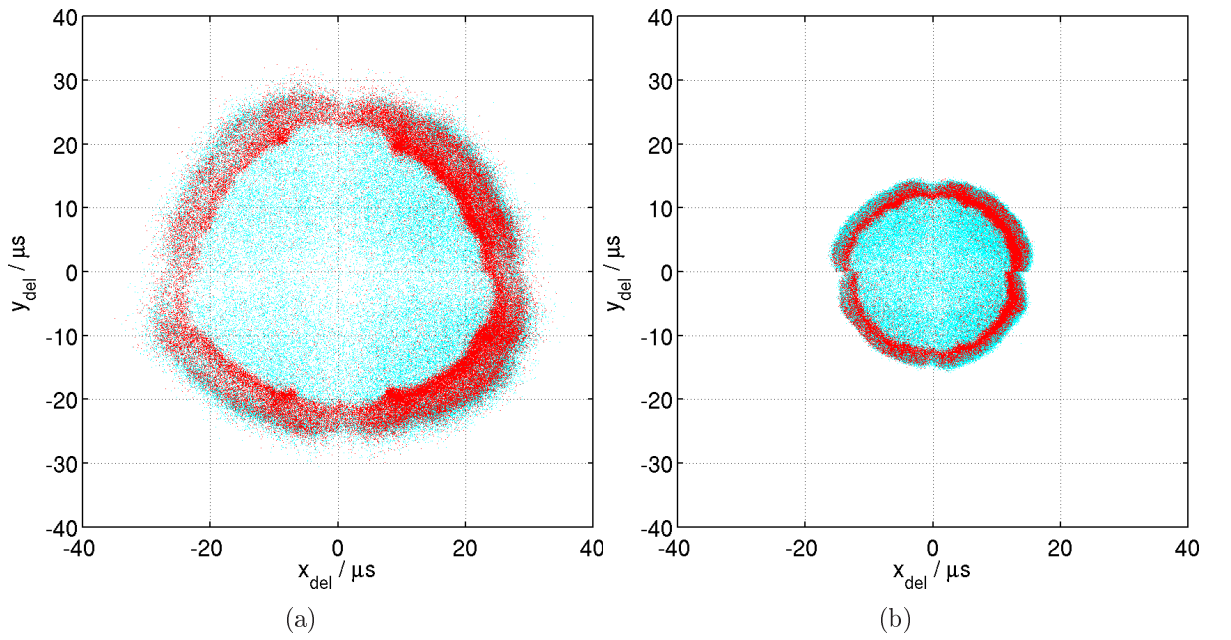
Figure 4.2: Delay plots for T1 Z5 (a) and T2 Z2 (b). Cyan (light) points are inner electrode events, while red (dark) points are outer and shared events. The delay plot on the right is much smaller because of the faster speed of sound in Si than in Ge.

of the faster sound speed in Si. The physical outer edge of the crystal actually corresponds to the inner edge of the ring of QO events. This radial degeneracy is unfortunately generic to ZIP detector position estimators, because of decreased resolution far from the quadrant boundaries, and because of increased reflection from the outer detector edge. Fortunately, delay and partition do not undergo foldback at exactly the same physical radius, so one can be used to break the degeneracy of the others. This will be explored later in this chapter.

4.2   Depth (z position)

The original design goal of ZIP detectors was to use the fast, athermal phonon signals to measure the depth of each event. Because the phonons take several microseconds to propagate across the 1-cm thickness of the crystals, while the ionization signal is very fast, it was expected that the delay of the phonon pulse relative to ionization would be directly proportional interaction depth. An interaction near the top (phonon sensor side) would produce a phonon pulse starting almost immediately, with a very small delay. For an interaction near the bottom (ionization sensor side), the phonons would have to propagate across the full 1 cm before being absorbed, and would therefore have a large delay. Events in the bulk of the crystal would have intermediate phonon start times, faster or slower depending on the z coordinates of the interactions. This depth sensitivity would be used to reject events near either surface, where the ionization signal can be suppressed and yield-based discrimination is compromised.

In practice, this continuous depth information is overwhelmed by several other effects. The first is the large contribution of the fast Neganov-Trofimov-Luke phonons to the rising edge of the phonon pulses. Because the Luke phonons have relatively low energies (THz frequencies), they scatter very little and propagate ballistically. The number of Luke phonons is directly proportional to the bias of the ionization circuit. In the deep site at Soudan, and for half of the shallow SUF run, the bias is -3V for Ge detectors and -4V for Si detectors. At this bias, the phonon signal is composed of equal parts primary and Luke phonons. The primary phonons are initially high-energy, and undergo frequent scattering (elastic Rayleigh scattering) and down-conversion. They

propagate as a diffuse cloud until they are absorbed or reach a low enough energy that their mean free path is long compared to the size of the crystal and they propagate ballistically. The primary phonons produced by an interaction therefore convey depth information that can be extracted by studying phonon pulse delay. The Luke phonons, on the other hand, are always fast, propagating ballistically from the start. They tend to dominate the rising edge of the pulse, and to wash out the depth information.[2]

If the Luke phonons are tightly beamed in the direction of travel of the charges, the situation may be even worse: they can actively cancel out some of the position information of the primary phonons. If an interaction takes place near the top (phonon side) of the detector, the charges (ordinarily, holes) moving toward the bottom (ionization side) travel through a potential difference of nearly the full 3V, while the charges (ordinarily, electrons) moving toward the phonon side experience only a small change in potential. The majority of the Luke phonons are then emitted by the holes and beamed downward, away from the phonon sensor. In the opposite case, an interaction near the bottom of the detector will produce Luke phonons predominantly beamed toward the phonon sensor. Thus, the start of the phonon pulse becomes faster for events far from the phonon side, and slower for events near the phonon side. This runs exactly counter to the depth sensitivity of the primary phonons, and may partially negate it.

The second effect that makes depth reconstruction using phonon start time difficult is a surface effect. When an interaction takes place near either the top or the bottom surface, the phonon pulses become significantly faster than usual, both in start time and rise time. This effect is attributed to the interaction of primary phonons with the metal films on the surfaces, allowing them to down-convert to lower energies and become ballistic. Because the effect is similar on both sides of the detector, it is not simply proximity to the phonon sensors that is important, but to any metallized surface. The faster timing for surface events (and relatively slower timing for neutrons) is shown in Fig. 4.3.

In summary, the expected direct mapping between z and phonon pulse start time

---

[2]At low ionization bias this effect disappears, and the dispersion in phonon delay for gamma scatters increases. This is an indication of partially recovered depth sensitivity.
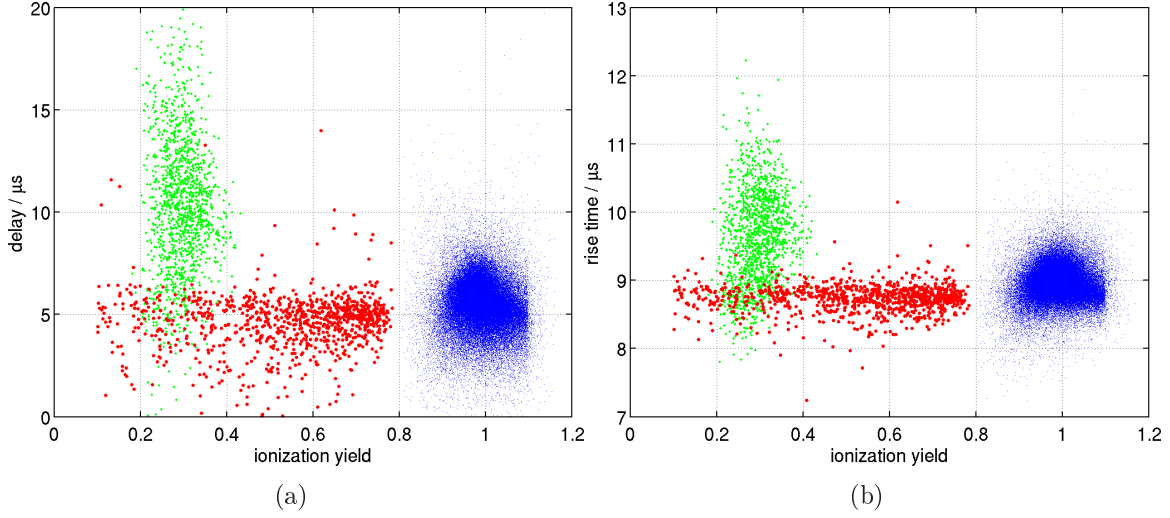
Figure 4.3: Delay (a) and rise time (b) of surface events in T1 Z5. The small (blue) points are ^{133}Ba calibration gamma scatters, the light (green) points are ^{252}Cf neutrons, and the dark (red) points are ejectrons from the barium calibration set. All events are $10 - 100$ keV in recoil energy.

is weak, but there is a significant surface effect that gives events very near the surfaces faster start times and rise times. Although this is not the form of depth information originally envisioned from ZIP detectors, it is now the standard way of defining a fiducial volume and rejecting surface events in CDMS-II. We wish to remove near-surface events that may have reduced ionization yield, and keep events in the bulk. For this application, analog depth information is not needed, and the faster rise and delay of surface scatters provide enough information.

4.3   Complications of ZIP detector analysis

The simplest analysis of ZIP detector data is a direct application of the techniques developed for an earlier generation of CDMS detectors, the Berkeley Large Ionization and Phonon detectors ("BLIPs") [43]. The BLIPs were true calorimeters, collecting thermal phonons from the crystal using neutron transmutation-doped thermistors. As such, they provided no information about the physical location of an interaction,

with the exception that events occurring inside the thermistors themselves were iden-
tified and rejected. The slow, thermal phonon pulses had typical rise times of several
milliseconds, and fall times of tens of milliseconds; the shapes and start times of these
pulses did not allow discrimination of surface recoils from bulk recoils. Furthermore,
the calorimetric measurement had no significant position-dependent effects, and nat-
urally gave a linear response with energy, so that there was little need for complex
corrections or reconstructions. The phonon pulses were fit in the time domain, with
a fixed template for each channel. The phonon response of the BLIPs did suffer from
a dependence on the crystal substrate temperature, but this was effectively removed
using a time-varying calibration based on daily heat-pulse calibrations and continuous
measurements of the NTD resistances. The ionization pulses were fit using a fixed
signal template for each channel, but in the frequency domain with an optimal filter
generated by the template and the known noise power spectrum.

Therefore, it was possible to effectively extract and use all of the information
provided by BLIP detectors in a straightforward way. Electromagnetic background
rejection was accomplished by a cut on ionization yield, assuming Gaussian variance
on the energy estimates and a simple energy-dependent form for the mean yield of
nuclear recoils: $Y_{NR} = cE_R^d$ and $\sigma_{NR} = aE_R + b$. The four parameters a, b, c, and d
were fitted to the observed yield bands of neutrons from *in situ* ^{252}Cf calibrations. Be-
sides the main yield-based background rejection cut, a fiducial volume cut was applied
based on the amount of energy in the outer and inner electrodes, and data quality
cuts rejected events with excessive electronic noise, event pile-up, or coincidence with
muons in the veto [75, 76].

This analysis was carried over to ZIP detectors with a few changes: a different
energy estimation algorithm, and new cuts to make use of the new position informa-
tion generated by ZIPs. Because the ZIP phonon pulses are much faster, and because
their shapes are not fixed, they are fit in the frequency domain with an optimal fil-
ter (Sec. 8.1.1) instead of a time-domain template. This decision to apply optimal
filtering to the phonon pulses had significant consequences for the uniformity of the
phonon energy calibration, as will be seen. The ZIP detectors were designed to take
advantage of phonon timing information, so an additional cut was developed to use

rise time to reject events near the crystal surfaces, where yield-based discrimination can fail.

Surface event rejection can be accomplished by a fixed cut on 10-40% rise time. Surface events have faster rise times and are rejected, while most bulk nuclear and electron recoils have sufficiently slow rise times to be accepted. This rise time cut, combined with the standard ionization yield cut, constitutes the most straightforward ZIP analysis, and the one closest to the analysis techniques used for BLIPs. During the first run of six ZIP detectors at SUF, however, it quickly became apparent that this simple approach was inadequate. The phonon pulse shape, including the rise time, has a strong intrinsic position dependence, and the phonon energy calibration may also vary with position for detectors with a significant T_c gradient, as described above 4.1. Furthermore, the use of optimal filtering on the phonon pulses introduced new complications. The optimal filter was calculated from a template with a fixed pulse shape for each channel, but ZIP detectors (by design) have pulse shapes that vary with depth, Luke phonon fraction, and x-y position. Optimal filtering folds this pulse shape variation into the phonon energy calculation: two pulses with the same total energy, but different shapes, generally have different estimated energies. In particular, the pulse whose shape is more similar to the template shape will have a higher estimated energy. Only one of the original six ZIP detectors, Z1, had a very large T_c gradient, but all suffered from optimal filter-induced variation in the phonon energy calibration, and all required a way to separate the depth information in pulse shapes from the intrinsic variation with x-y position in order to reject surface events while keeping bulk events across the crystal.

Chapter 9 is devoted to the construction of timing cuts used to reject surface events for the one- and two-tower WIMP-search runs from Soudan.

4.4   Look-up table correction

These issues have been addressed by corrections applied before making any cuts. The intent was to create relative quantities that are uniform with x-y position and energy, and as far as possible, consistent across detectors. The corrected phonon

recoil energy is the best estimate of the physical energy deposited by an interaction. The corrected timing quantities, on the other hand, no longer correspond directly to the physical properties of the phonon pulse, but instead combine several different kinds of information into a timing parameter suitable for cutting on. For example, the corrected rise time of a pulse may be quite different from the physical rise time of that pulse, but corrected rise time retains depth information while evening out x-y position and energy dependences. The correction algorithms described below were developed by Blas Cabrera and Clarence Chang, and are detailed in the latter's doctoral dissertation [83].

The first pass at a position correction was a matrix made by gridding the detector radially and azimuthally on a delay plot. The correction value for a given quantity at a given grid cell was found by averaging that same quantity for a sample of "good" calibration gamma events in that cell. The grid of correction values formed a look-up table: the correction was applied by interpolating between the nearby grid points to find the local correction value, and dividing it out. This technique was applied directly to timing quantities like rise time, and phonon energies were corrected by applying the correction to ionization yield.[3] Near the center of the detector, and for the best detectors, this approach worked well. However, phonon delay is not an unambiguous measure of event location because of pile-up and foldback. In addition, the gridding in polar coordinates was somewhat undesirable, since the size of each cell was not very well matched to the density of calibration points available to calculate the correction value for that cell.

The current look-up table substantially improves on the earlier version. We use a matrix of calibration gamma events in three dimensions, with no binning. The x and y coordinates are the two axes of a partition plot, *xppart* and *yppart*. The third coordinate, forming the z axis, is the delay plot radius *rdel*. In this space, the locus of bulk gamma events takes the form of a "basket" that curls in around the top, as both partition and delay become badly behaved and fold back. The combination of delay and partition is sufficient to break the degeneracies in each and achieve a

---

[3]Assuming that the ionization signal has no position dependence. See Sec. 5.9 for a small gradient in charge calibration.
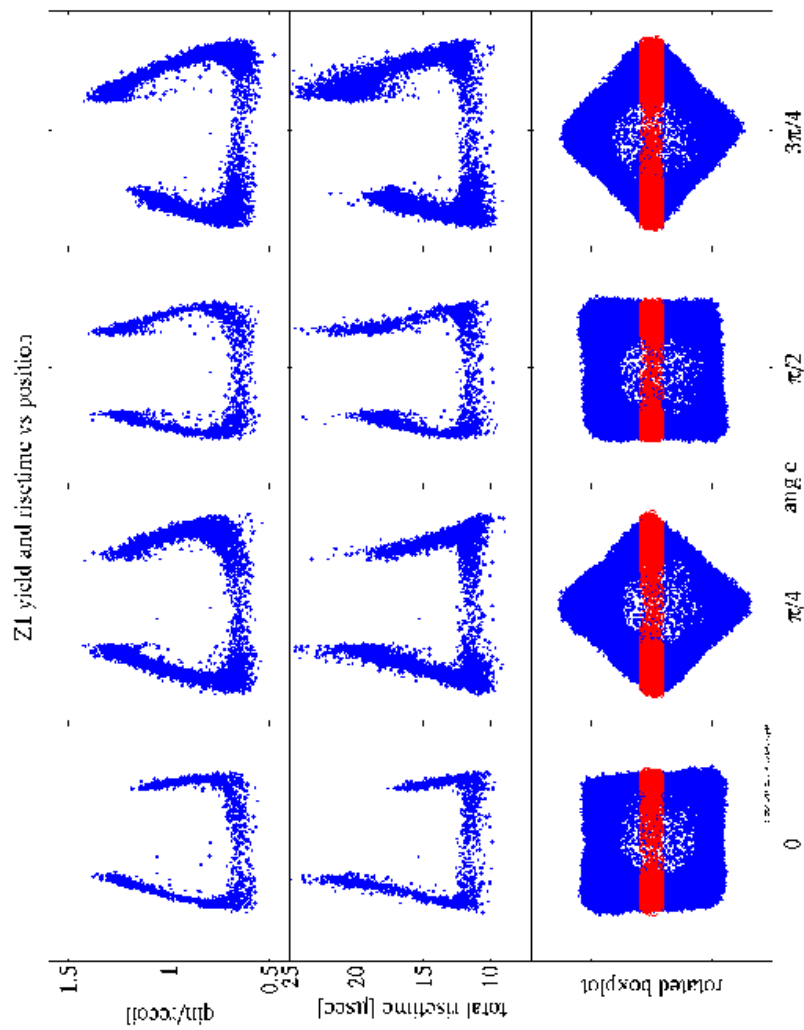
Figure 4.4: Look-up table correction of T1 Z1, showing uncorrected yield (top row) and rise time (middle row). The columns represent four different slices through the 3-d manifold, as indicated in the bottom row. Source: C. Chang [83].
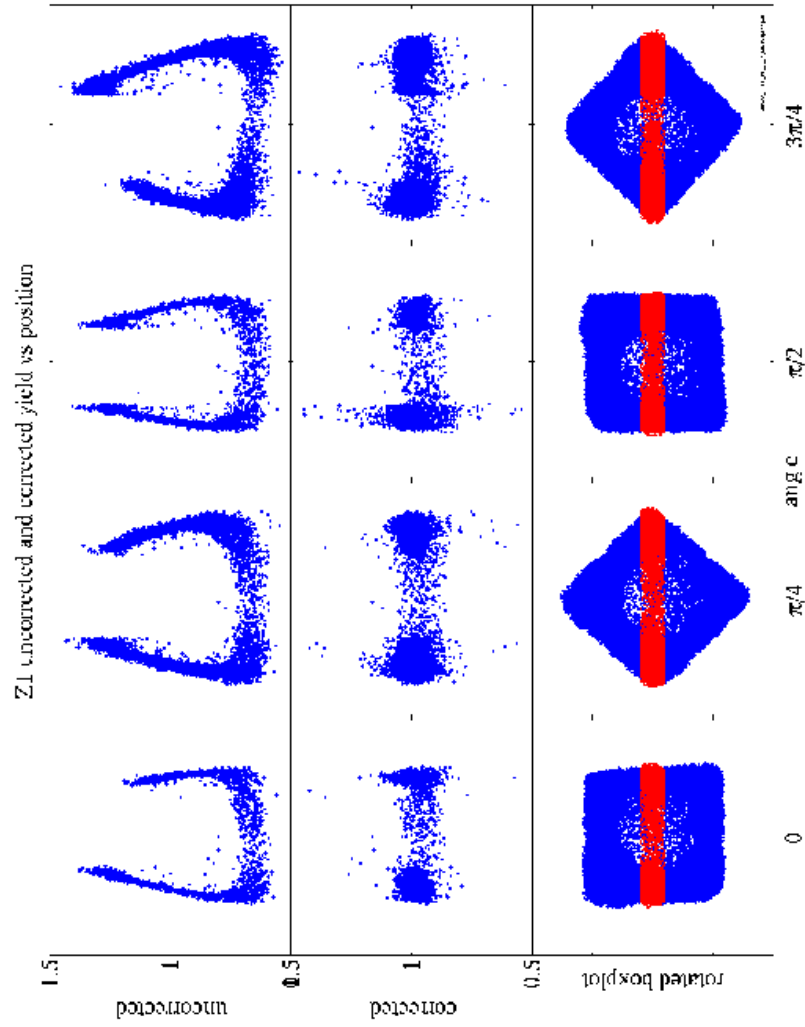
Figure 4.5: Look-up table correction of T1 Z1, showing yield before and after correction. The four columns are four slices through the 3-d correction manifold, as indicated in the bottom row. The top row shows uncorrected ionization yield, and the middle row shows corrected yield. Source: C. Chang [83].

good correction. The metric of this three-dimensional space is Euclidean, with a fixed scaling applied to the delay of each detector so that the basket is about as tall as it is wide. Each look-up table point is created by averaging the yield, rise time, etc. of a number N_{avg} of calibration gammas. To apply the correction to a given data event, its values of *xppart*, *yppart*, and *rdel* determine its location in the 3-d space of the manifold. The correction factors used for the event are the weighted average of the $N_{neighbor}$ nearest look-up table points, where $N_{neighbor}$ is optimized to be large enough to smooth out any random scatter in the look-up table points, but small enough to preserve real position variation. The manifold for the look-up table correction is shown in Figs. 4.4 and 4.5. In each case, we show four slices through the manifold, as indicated on the box plot in the bottom row of the figure. Figure 4.4 shows uncorrected ionization yield and phonon rise time, and Fig. 4.5 shows the effectiveness of the correction in removing the position correction in ionization yield.

In addition to position dependence, the timing of ZIP detector phonon pulses varies with energy. The phonon energy calibration also varies somewhat with energy, because of nonlinearities in the TES responses. Therefore, the current phonon correction also incorporates an energy correction, determined by a simple functional form with a few parameters. The total phonon energy is linearized relative to the ionization energy (assuming that the ionization signal has no saturation effects or other nonlinearities), using the functional form $p' = \tau \left(e^{\frac{p}{\tau}} - 1 \right)$. The timing parameters, such as rise time and phonon delay, have been corrected using the form $p' = \frac{p}{a + cE_R^b}$. In the most recent analysis of Soudan data, some timing parameters are instead energy-corrected with the functional form $p' = \frac{p}{a + c\tanh(bE_R)}$, but the idea is unchanged. The coefficients are chosen so that the energy dependence of each timing quantity is removed for a sample of good bulk calibration gamma events.

The look-up table correction, including the energy linearization, has been a very successful and integral part of all published CDMS results from ZIP detectors [85, 86, 87, 27].

It is important to note that the look-up table correction breaks the degeneracy in the partition and delay plots in order to group events that are physically near each other in *x*-*y* position, but it does not provide an estimated x and y. Such a position

reconstruction would also be quite useful in several ways. First, it is somewhat difficult to study the effects of the current look-up table correction in the three-dimensional manifold space. Second, a physical radius can be used to help classify multiply-scattering events: low-yield surface events in two detectors should be found at almost the same location in both detectors, but doubly-scattering neutrons travel farther between interactions and are likely to have moved farther in x-y space. Third, surface contaminants may be localized on or near the detectors, so that a position reconstruction can help identify them. Finally, the effectiveness of the timing and yield discrimination appear to be dominated currently by systematics associated with remaining position dependence, especially in the outer parts of the detector, where delay and partition are worst-behaved and the manifold tends to "bunch up." A position reconstruction would make these problems more tractable.

As a more general consideration, the analyses using the look-up table correction are forced to choose some measured quantities as position estimators (relative partition and delay of phonon energy among the four quadrants), and others as discrimination parameters (rise time and delay of phonons relative to ionization). In reality, all of these quantities are sensitive to position, recoil type, and depth, to varying degrees. (This is true to an even greater degree for the new generation of one-inch-thick detectors, now under development). The partition and delay used to create the correction manifold do not have identical distributions for bulk electron recoils, surface electron recoils, and nuclear recoils, so that a look-up table created from a gamma calibration cannot perfectly correct neutrons or surface betas. It would be desirable to put all of the measured quantities on an even footing, taking advantage of the information about x-y position, recoil type, and depth in every observable in a consistent way.

4.5 Delay-based position reconstruction

The problem of reconstructing event location has also been addressed by Gensheng Wang, as detailed in his dissertation [88]. His approach uses physical calculations and empirical fits to data to construct a mapping from event location to measured

phonon delay. Inverting this map allows the location of an event to be estimated in x, y, and z. The delays used are the start times of the phonon signals in all four phonon sensors relative to the start of the ionization signal.

The calculation begins with simulated events distributed throughout the crystal in a uniform grid of x, y, and z coordinates. For each starting location, it is assumed that the phonons propagate ballistically until they are absorbed by one of the phonon sensors and contribute to the observed pulse in that sensor; or run into another surface, and are lost. The result is a calculated delay between each interaction and the start of the phonon pulse in each of the four quadrants. This calculated result is matched to the data by allowing the delays in each quadrant to be scaled by a constant factor, and demanding that the delay plot have the correct size. This gives a formula for delay as a function of position, with a calculated functional form and four free parameters. The inverse function is represented by a fifth-order polynomial, and can be used as a position reconstruction.

This reconstruction does largely eliminate the fold-back in delay radius for most detectors, although it leaves a region of pile-up near the edge of the detector. Because the reconstruction is defined for bulk gammas, other types of events (neutrons and surface betas, which have very different delays) are assigned reconstructed positions that are not physical, but that may still be useful. For example, a significant fraction of nuclear recoils have reconstructed position with radius > 3.81 cm (the physical radius of the detector substrate), allowing the reconstructed radius to be used as a discrimination parameter. Interpretation of the reconstructed z coordinate is complex. Like the phonon delay time itself, the z coordinate constructed in this analysis has weak sensitivity to depth within the bulk of the crystal, and significant sensitivity to surface effects, so that events near either surface have low reconstructed z, and events in the bulk of the crystal have a distribution of z values. This has also successfully been used by Wang as a cut parameter for surface event rejection.

4.6 Parametrizing the look-up table manifold

Since the space of the look-up table does a reasonably good job of breaking the degeneracies present in the delay and partition reconstructions individually, this same space could form the basis of a position reconstruction. All that would be needed is a way of parametrizing the two-dimensional manifold within the three-dimensional space of $xppart$, $yppart$, and $rdel$. A first attempt along these lines was made in 2003, during the analysis of the one-tower run in the Stanford Underground Facility, but with $rdel$ replaced by another position-sensitive quantity: the sum of the peak-to-20% fall times in the four quadrants, $ftsum$.

The approach in this analysis was completely empirical, and was only applied to one Ge detector (Z5). First, the detector was divided into twenty narrow pie-slice wedges using the partition angle. Within each slice, the locus of points in the two-dimensional space of $rppart$ and $ftsum$ formed a shape resembling a shrimp, with the tail at the center of the detector, the back curving along the region in which the two parameters fold back, and the head made up of points near the edge, in the outer electrode. The points can be ordered by following the shrimp from its tail ($r = 0$) to its head ($r = 3.81$ cm), and this ordering gives a physical radius estimate when combined with the constraint that gamma events in the low-background data set should have a uniform physical density. The reconstruction was performed iteratively. The "zeroth step" ordered the points according to the sum of $rppart$ and $ftsum$, appropriately weighted, and the ordering was scaled to give a radius parameter between 0 and 1 for each data point. The first step fit a second-order polynomial to $rppart$ and to $ftsum$ as a function of the radial parameter. These functional-form fits effectively parametrized the manifold. Every data point was matched to the closest point on the parametrized manifold, giving a new ordering and a new radial parameter. The second step fit new polynomials to $rppart$ and $ftsum$ as functions of the radial parameter from the first step, and so on through four iterations. All iterations after the first used third-order polynomials. The final step was to convert the final ordering of points into a physical radius estimate using the known uniform distribution of low-background events.

(a)   (b)
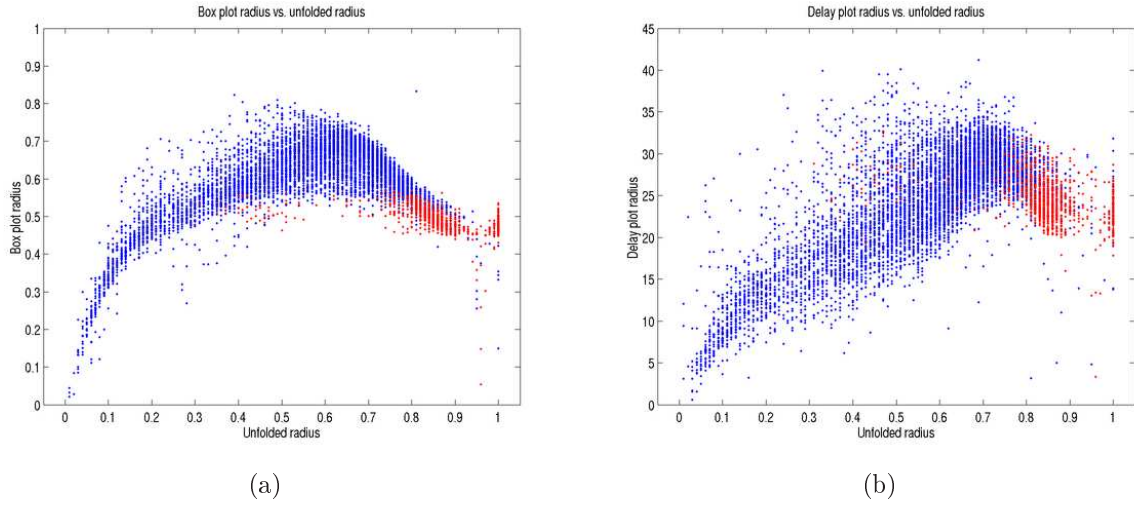
Figure 4.6: Partition radius *rppart* (a) and delay radius *rdel* (b) vs. radius recon-
structed according to the method of section 4.6. Blue points are events primarily in
the inner electrode, while red points are events primarily in the outer electrode.

This procedure was very successful, constructing a monotonic radius estimate for
gamma events in a single detector, shown in Fig. 4.6.

However, this approach had several limits on its applicability. First, it required
a significant amount of human intervention to build the model. The relative weight-
ing of *rppart* and *ftsum* were set by hand, and the degrees of the polynomials in
each step were carefully chosen so that the iterative process would not fall into any
pathologies, such as over-fitting. Even so, the ordering of points produced by this
process was not perfect; near the edge of the detector, the fits sometimes doubled
back on themselves, producing non-physical and incorrect radial estimates. There-
fore, applying this technique to multiple detectors would be labor-intensive. Second,
the fits in each pie-slice of the detector were completely independent of each other.
Thus, it could not be guaranteed that the parameterizations in neighboring angular
bins would match up well, and the fitting process had to be supervised to ensure
this. Third, the fall-time sum proved only to be a useful position parameter for some
detectors. In the silicon detectors, it does not convey any useful position information
at all. This was also true in one of the germanium detectors, Z2.

Therefore, this reconstruction can be regarded as a promising and interesting approach, but in need of further development. The ZIP functional-form-based position correction developed in Chapter 6 of this thesis will improve and generalize this approach into a full and robust analysis of ZIP detector data.

Chapter 5

Position dependence in ionization

Although the ionization signal is not as strongly position dependent as the phonon signal, several types of position dependence are seen. There are strongly localized features associated with fringing effects and changes in grounding near the outer edge of each detector. When a detector is not well neutralized, its charge collection efficiency may vary with event location. Finally, even in the case of well-neutralized detectors, we consistently see a residual gradient in charge collection, manifested as a gradient along the detector's y axis. Several possible causes of this effect will be discussed.

These effects are dealt with in the analysis in different ways. The discrete, near-edge effects are cut out by the fiducial volume cut. Neutralization problems are avoided through an adequate LED flashing routine and monitoring, with any suspect data sets removed from the analysis. The residual gradient is corrected using a simple functional-form adjustment as described below.

5.1   Theory of ionization readout

Whenever a particle interaction creates ionization in the crystal, the charges initially appear as a small cloud of electrons and holes. Some of these quickly recombine, but most of them are separated in the drift field and move across the crystal. Since we ordinarily operate with a charge bias of -3 V in Ge or -4 V in Si, the drift field is

(a)   (b)
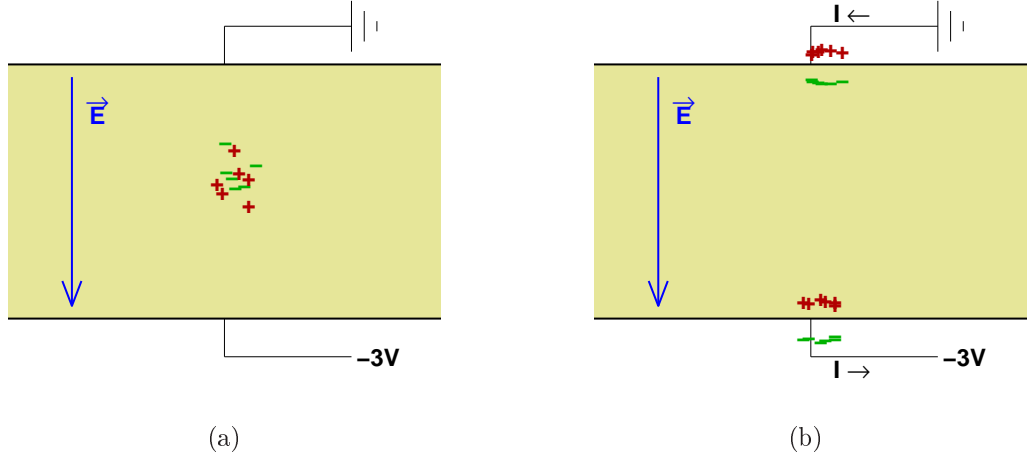
Figure 5.1: Mechanism of ionization signal. The initial interaction produces a cloud of electrons and holes (left). As they drift across the crystal, current flows to form image charges on the electrodes (right).

$3 - 4$ V/cm, with positive holes attracted towards the electrode on the ionization side of the detector, and negative electrons attracted towards the ground plane defined by the phonon side. The charges are continuously accelerated by the drift field, and they shed energy by emitting Neganov-Trofimov-Luke phonons ("Luke phonons") until they reach a constant drift velocity (approximately the speed of sound). At the surface, the charges are trapped at the boundary between the crystal and the amorphous silicon layer, where they remain until they slowly tunnel across this barrier into the electrode. This tunneling is very slow compared to the charge readout, so that it allows surface charges to dissipate without affecting the ionization signal.

Although the charges do not physically enter the electrodes over the time scale of a pulse, the motion of the charges through the crystal does induce external currents. The physical charges cause image charges to arise on the conductors as shown in Fig. 5.1, and these image charges are actually measured. Ramo's theorem provides a convenient method of calculating the induced currents.

The bottom face of each ZIP detector is segmented into an inner disc-shaped electrode and an outer annular electrode. Any motion of charges within the crystal induces currents on the electrodes. In general, the currents can be induced on the

inner and outer electrode, ground, and the electrodes of neighboring detectors; all of these possibilities must be considered. For the usual, full-collection case (Sec. 5.1.2), it will be shown that the situation is simple, and there is no crosstalk.

5.1.1   Ramo's theorem

Ramo's theorem gives the current induced by the motion of charges in space, on a conductor held at constant potential. The theorem can be stated as

$$I_j = -q\vec{v} \cdot \vec{\nabla} V_j \tag{5.1}$$

where I_j is the current induced on the jth conductor by the motion of a charge q with a velocity $\vec{v}$. The Ramo potential V_j is the potential that results if the conductor j is held at 1 volt, all other conductors are grounded, and any other space or surface charges are removed. However, the measured pulse height depends only on the total induced charge, not on the instantaneous currents. The integral of the current has the very convenient property that it only depends on the Ramo potentials at the stopping and starting points, and not on the path that the charges follow. The induced charge is

$$Q_j = -q \left(V_j \left(\vec{x}_f \right) - V_j \left(\vec{x}_i \right) \right) \tag{5.2}$$

when a charge q travels over a path from $\vec{x}_i$ to $\vec{x}_f$. In our detectors, the ordinary case is that a number n of electrons and n holes are created at an initial interaction site $\vec{x}_i$, the electrons all drift until they reach a surface at $\vec{x}_e$, and the holes reach a surface at $\vec{x}_h$. The induced charge is then

$$Q_j = ne \left(V_j \left(\vec{x}_e \right) - V_j \left(\vec{x}_h \right) \right) \tag{5.3}$$

depending only on the final destination of the electrons and the holes.

The Ramo fields of the inner and outer electrodes, along with the full drift field, can be calculated by relaxation. The CDMS collaboration has performed several such computations to study different aspects of the ionization readout. One calculation has been done by Blas Cabrera using a finite element method in MATLAB's PDE

toolbox. Another calculation, taking into account the floating metal surfaces in the outer region of the detector, has been performed by the author using a fixed grid in MATLAB. The results of this calculation are shown in Fig. 5.5, and will be used to understand various charge-channel phenomena in the following sections. It is assumed that all charges drift along field lines until they encounter an electrode, the edge of the detector, or a trapping site. This assumption is not perfect, however, since electrons and holes in cryogenic Si and Ge have anisotropic masses. However, deviations from this law are unimportant whenever the field coincides with a $\langle 100 \rangle$ or $\langle 111 \rangle$ crystal axis, as is the usual case for ZIP detectors.

Special care has been taken in handling the floating metal surfaces in the outer region of the phonon surface. These patches are conductors, so there can be no potential difference across them; however, the potential is not fixed, and is not known a priori. We assume that the floating regions are coupled to the ground of the phonon side through a large resistance, so that they are at 0 V before the beginning of a pulse. As the charges drift, the total charge on each floating patch must be kept constant. This condition is imposed using the principle of superposition. We first perform a relaxation to find the Ramo field when the floating metal is held at 1 V, and then a second time at 0 V. These calculations give two sets of Ramo potentials, $V_j^{1\text{ V}}(\vec{x})$ and $V_j^0(\vec{x})$. In each case, we then use Gauss's law to find the induced charge on the floating metal. Let this be denoted $Q_{fl}^{1\text{ V}}$ and Q_{fl}^0 for the two voltages. The true Ramo potential is

$$V_j(\vec{x}) = V_j^0(\vec{x}) - \frac{Q_{fl}^0}{Q_{fl}^{1\text{V}}} V_j^{1\text{ V}}(\vec{x}). \tag{5.4}$$

This gives the correct Ramo potential of any point in the detector (including the floating metal surface itself) relative to the conductor j, in the presence of the nongrounded conductor.

5.1.2 Full charge collection

Full charge collection occurs when the electrons migrate to the grounded surface of the detector's phonon side, and the holes migrate all the way to the inner or

outer electrode. For full-charge-collection events in the inner region, $V_{inner}(\vec{x}_e) = V_{outer}(\vec{x}_e) = V_{outer}(\vec{x}_h) = 0$, and $V_{inner}(\vec{x}_h) = 1$, so that $Q_{inner} = ne$ and $Q_{outer} = 0$. Thus, Ramo's theorem shows that there is no intrinsic crosstalk between the inner and outer channels, and that the induced charge is equal to the amount of ionization produced. This case corresponds to the main QI population of events (Fig. 5.2 A), and the secondary QO population of events (Fig. 5.2 I).

Although Ramo's theorem predicts no dependence on the value of the bias voltage, the field must be large enough to separate the electrons and the holes in the initial charge cloud. If the voltage is set to zero, many of the charges recombine, and charge collection is zero or small, depending on the residual stray fields. As the voltage is increased, charge collection increases, approaching a constant, full collection at a few volts.

5.1.3   Suppressed charge collection

Whenever the charges fail to reach a surface with Ramo potential of zero or one, the induced charge on the electrode will be less than ne. This can happen for a number of reasons. The drifting charges may be trapped at charged impurity sites; they may intersect the outer cylindrical wall of the detector; they may reach a surface that is electrically floating; or they may fail to be separated effectively and recombine. These cases are summarized schematically in Fig. 5.2 and discussed in the following sections.

5.2   Charge trapping and neutralization

Charge trapping occurs when the electrons or holes encounter charged defect sites in the crystal. When the detectors are initially cooled to operating temperature, there are many such sites, so that the drift length is short. Because the charges are stopped before drifting across the full thickness of the crystal, the induced charge is suppressed, and yield-based discrimination may be compromised. In order to achieve a good ionization signal, the detector is illuminated with LEDs while the electrodes

are grounded. The near-infrared photons excite electrons and holes that are trapped at the charged impurities, neutralizing them. As the drift length approaches the thickness of the detector, full charge collection is restored. Extended "LED baking" is performed at the beginning of each fridge run, until the trapping sites are neutralized and full charge collection is achieved. Historically, the initial, high-trapping state has been known as Mode I, while the neutralized, high-charge-collection state has been known as Mode II [89]. Under normal data-taking conditions, every particle interaction creates mobile charges in the crystal. These electrons and holes re-charge the trapping sites, so that extended data taking (especially high-rate calibration data with radioactive sources) gradually de-neutralizes the crystals. Therefore, care must be taken to perform intermittent LED flashing during ordinary data taking, frequently enough to maintain good neutralization.

The Ge crystals are much easier to neutralize than Si, so extra care must be taken in baking the Si detectors. This is especially true for the Si detectors in the bottom position in Tower 1 and Tower 2, since they can be neutralized only by their own LEDs. Detectors in other positions can receive photons from their own LEDs as well as those of the detectors below them. The Si detector at the bottom of Tower 1, T1 Z6, is especially difficult to neutralize, as its LED is oriented in the wrong direction (see Sec. 3.4.2).

5.3 Surface events

The top and bottom surfaces of the detector form an ionization "dead layer" with a thickness of a few microns. Recoils occurring within the dead layer do not have full charge collection, because some of the charges in the initial cloud diffuse to the wrong electrode before the electrons and holes are separated in the drift field. Therefore, ionization yield is suppressed. The problem is ameliorated by a layer of amorphous silicon, which has a different band gap and presents a barrier to prevent this mode of trapping [89].

Surface events are shown as population C in Fig. 5.2. Electron recoil events very close to the surface can have yield low enough to fall within the nuclear recoil band. In
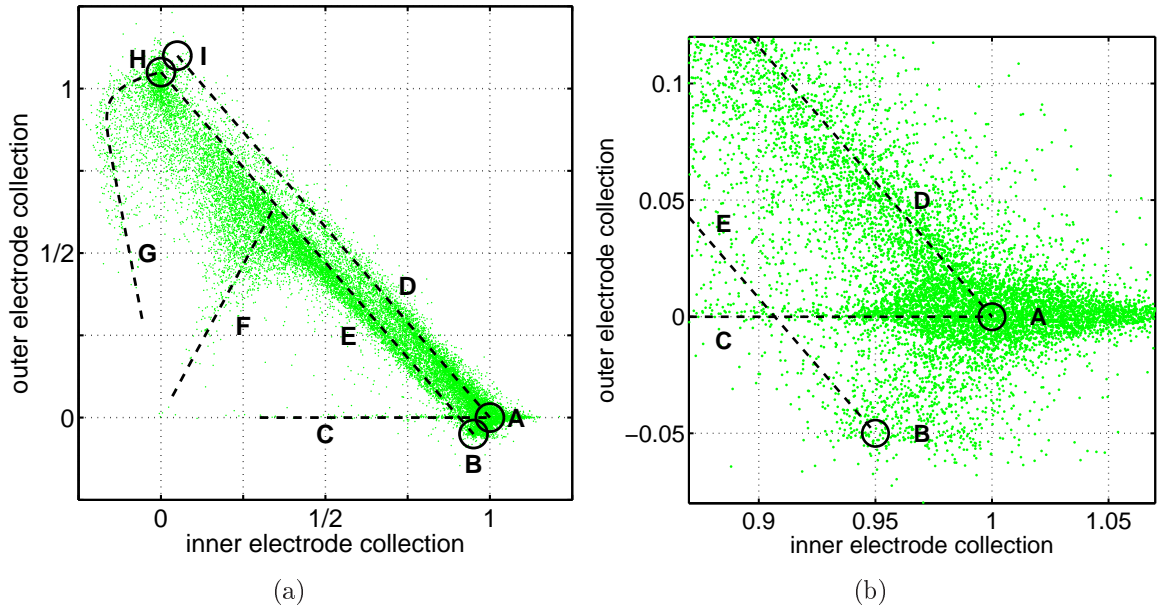
Figure 5.2: Ionization collection in the inner and outer electrode of a Ge ZIP. Right plot is zoomed in around inner electrode events. The features are (A) full collection in inner electrode; (B) "holiday islands"; (C) inner-electrode surface events; (D) full-collection shared; (E) floating-ground shared; (F) "funnel"; (G) "ear"; (H) floating-ground outer electrode; (I) full-collection outer electrode.

low-background running, the surface events primarily consist of electrons from beta-emitting radioactive contaminants on the detector. Calibration with a Ba source also gives a population of surface events: when a photon Compton scatters near the surface of one detector, it may eject an electron into a neighboring detector. The electron then has a small penetration depth in the neighboring detector, and appears as a surface event. This population of photon-induced surface events is known as the "ejectrons," and is used to set cuts for surface event rejection in Chapter 9.

5.4 Discontinuities

Three of the detectors in Tower 1 were fabricated using an older design for the electrode geometry. In this design, the outer electrode became very narrow near the flats, and sometimes left gaps (See Sec. 2.2.2 and particularly Fig. 2.2). These gaps compromise the effectiveness of the fiducial volume cut in these detectors. Affected detectors are T1 Z1, Z4, and Z6. In all subsequent detectors, the shape of the outer electrode follows the contours of the flats, so that QO has a constant width. The gap can be clearly seen in the position-reconstructed plots for T1 Z1, Fig. 6.7.

5.5 Floating metal surfaces

The phonon sensors of a ZIP detector do not extend all the way to the edge of the crystal, but have a stair-step outer boundary a few mm to one cm from the edge. Outside of this boundary are inactive metal features: this is ordinarily an aluminum grid pattern, but occasionally may be solid metal, depending on the effectiveness of mask exposure near the wafer edge. Conduction through the amorphous silicon allows these features to stay at 0V over sufficiently long time scales, but over the duration of an ionization signal they are electrically floating. This is seen as a slight suppression in the ionization signal for these events. Furthermore, because the Ramo potential at the floating surfaces is nonzero relative to both the inner and outer electrodes, a small signal is also seen in the "wrong" electrode; effectively, events near floating patches on the ground side have extra QI-QO crosstalk. These effects are primarily
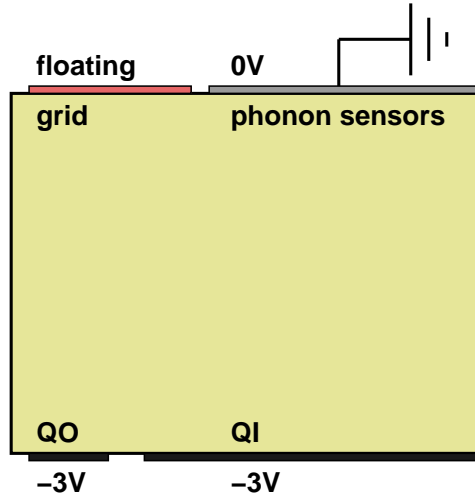
Figure 5.3: Layout of electrodes and ground plane near outer edge of ZIP detector. A vertical cross-section through the detector is shown. The "holiday islands" are where QI overlaps with the floating grid.

seen in outer-electrode events. However, because the boundary is a stairstep instead of a circle, it does not follow the division between the inner and outer electrodes perfectly. Some inner-electrode events in the regions known as "holiday islands" are associated with floating patches on the ground side, and some outer-electrode events reach true ground. The "holiday island" events are found in region B of Fig. 5.2, while the grounded QO events are found in region I.

The configuration of biased, grounded, and floating surfaces is shown in vertical section in Fig. 5.3. The layout along the detector's top surface is shown in data and in a schematic in Fig. 5.4. Charges that come to rest at the floating metal do not induce the same signal as charges that come to rest at the grounded phonon sensors. This can be taken into account by finding the Ramo potential at the floating outer grid. The Ramo potential relative to each surface as calculated by the author in a model with cylindrical symmetry is shown in Fig. 5.5.

Because the detectors are in a close-packed configuration, the phonon side of one detector (which serves as the ground of the ionization readout circuit) is very close to the ionization side of the neighboring detector. Although the floating metal patches cannot draw current to form image charges, the nearby ionization electrode
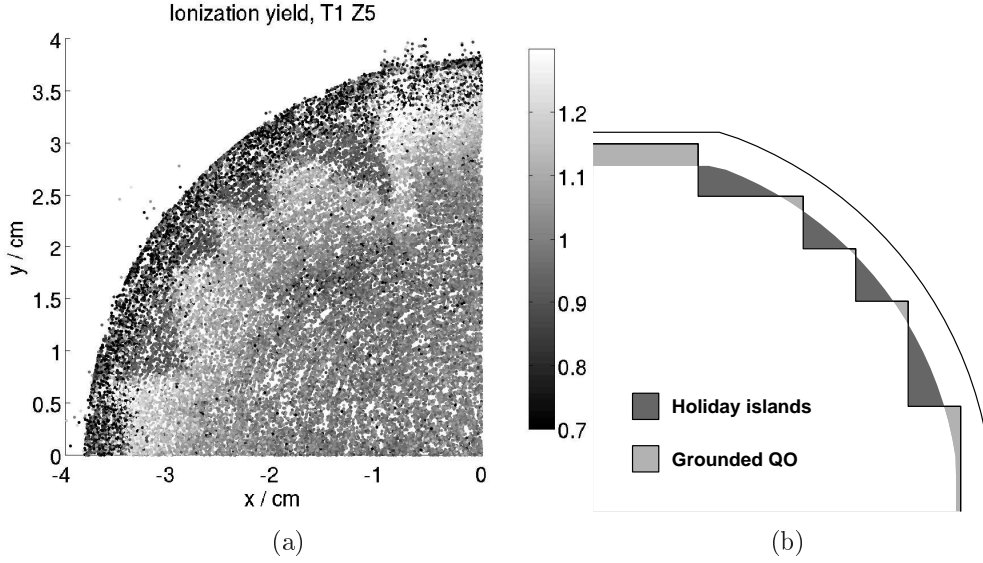
Figure 5.4: The "holiday islands" shown in data (a) and in a schematic (b).

can do so. This manifests itself as a between-detector crosstalk specifically associated with the outer regions of the detector. Under ordinary operating conditions the detectors have their ionization sensors biased at a negative voltage, so that electrons drift upwards to the phonon side, while holes drift downwards to the charge side. The detector-detector crosstalk involves a positive image charge formed on the charge side of the next detector, while the charge side ordinarily forms a negative image charge. Therefore, the detector-detector crosstalk appears as negative-going ionization pulses, that is, a negative ionization energy.

Study of these events has led to an improvement in later ZIP detectors. The detectors of Tower 3 and later have the outer metal surfaces on the phonon side connected directly to ground, so that there are no longer any floating patches. This eliminates the anomalies in yield and crosstalk caused by floating patches in Tower 1 and Tower 2 detectors.

However, stray capacitances and the non-zero resistance to ground prevent the grounding of the phonon side from being perfect even in the absence of floating metal. Therefore, a very small irreducible crosstalk effect occurs even in detectors without the "holiday island" problem. This is shown in Fig. 5.6 for T2 Z3, a detector of Tower

(a)   (b)

(c)   (d)
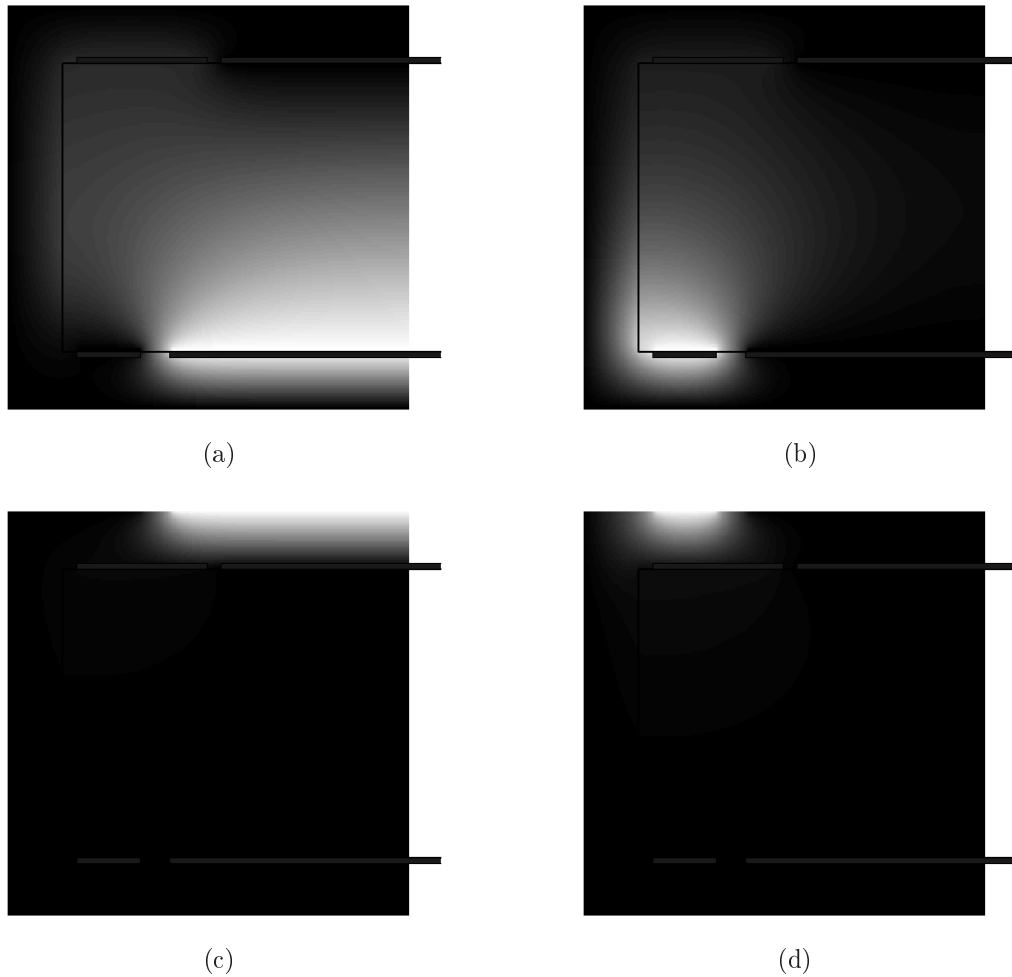
Figure 5.5: Ramo potentials in the outer part of a Ge ZIP detector. The outer and inner electrodes are on the bottom face; the metal on the top face is grounded in the central (right) part, and floating in the outermost portion. The Ramo potential is shown relative to the inner electrode (a), outer electrode (b), inner electrode of the detector above (c), and outer electrode of the detector above (d). The nonzero potential in (c) and (d) generate detector-detector crosstalk.
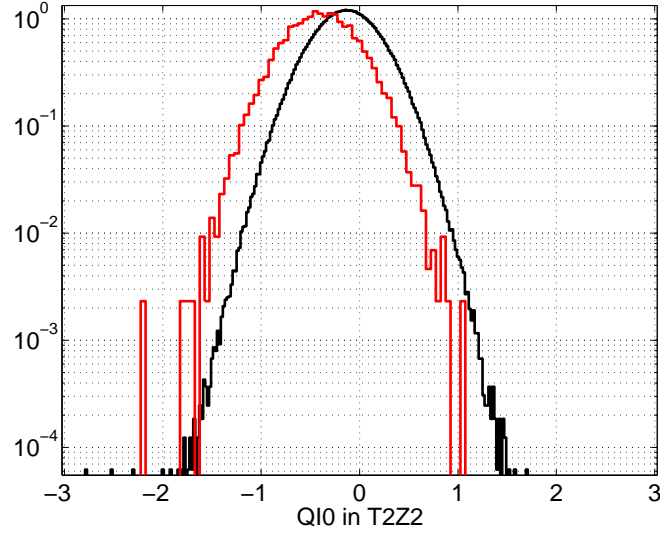
Figure 5.6: Residual crosstalk in a detector with no "holiday island" problem. The horizontal axis shows energy in the inner ionization channel in the detector T2 Z2 in keV, with the pulse start time forced to zero in order to guarantee that negative-going pulses are found. The dark (black) histogram represents events with 0-50 keV of energy in the next detector, T2 Z3, while the light (red) histogram shows events with 200-500 keV in T2 Z3. All events are required to pass the inner-electrode cut in T2 Z3. The shift between the two histograms shows a small residual crosstalk of order 0.2%.

2 which is found not to have the large crosstalk caused by floating regions. Even without the floating metal problem, large ionization signals cause a small crosstalk in the detector above.

5.6   Bare outer surface

The cylindrical outer wall of each ZIP detector is unpolished, bare crystal. Whenever the fringing fields of the ionization bias intersect this surface, the drifting charges may stop and become trapped, as shown in Fig. 5.7. This outer wall has a Ramo potential relative to either electrode that is generally between 0 and 1, so that ionization collection will be nonzero but incomplete. The "ear" events in Fig. 5.2(G) occur in this region. These events are well removed by the fiducial volume cut, so they are not
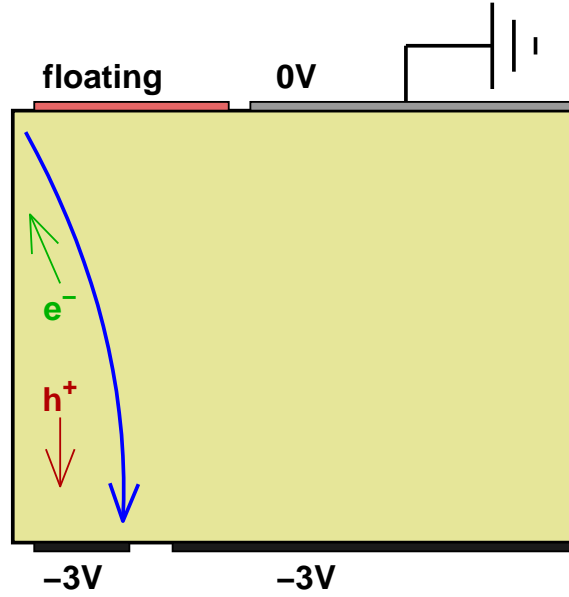
Figure 5.7: Schematic of "ear" events. Holes drift to the bottom electrode, but electrons are trapped on the cylindrical outer surface because of fringing fields.

a significant source of leakage into the nuclear recoil band.

5.7   The funnel

The shared region between the inner and outer electrodes has a peculiar population of events known as the "funnel," region F in Fig. 5.2. The funnel events have yield values from near zero up to full charge collection, and a fixed ratio of inner to outer electrode signal. The effect is seen in all Ge detectors in Run 118 and Run 119, although it was not seen in previous data sets from the shallow Stanford site. There is no clear funnel effect in Si detectors.

The partition of charge between the inner and outer electrodes is not perfectly constant, but varies with position. All funnel events are in the outer part of the detector, within the outer electrode or the outer edge of the inner electrode. Charge partition has a clear dependence on azimuthal angle θ: it changes dramatically near the flats in the substrate, as is shown in Fig. 5.8. Since the outer electrode changes shape to keep a constant width near the flats, it may be that the funnel effect is
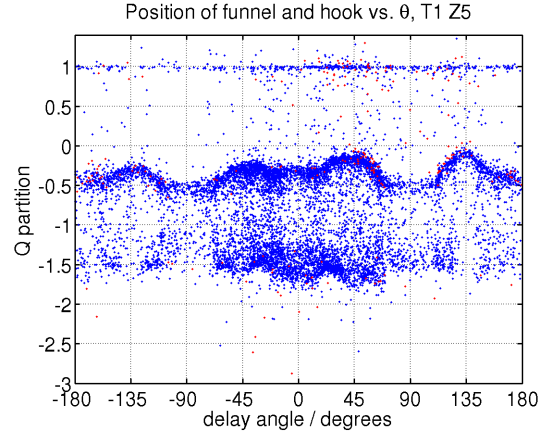
Figure 5.8: Position dependence of the funnel in T1 Z5. Events shown are ^{133}Ba calibration events with low yield between 0.3 and 0.75. Full collection inner electrode events are defined to have partition +1, and outer electrode events have partition −1. The band of events near −1/2 are funnel events, and the events near −1.5 are "ear" events. For funnel events, the partition favors QO more near the flats (major flats at ±90°, minor flats at 0 and 180°), and QI more away from the flats. The "ear" events form a mirror image, with the opposite position dependence.

sensitive to the geometry of the outer electrode. The physical origin of the funnel events remains poorly understood, but they indicate charge trapping in the bulk in the outer region, which is surprisingly difficult to eliminate through LED flashing. A similar phenomenon has been studied in the Edelweiss experiment, which uses the different propagation times of electrons and holes in Ge crystals to estimate interaction depth [90, 91, 92, 93].

Funnel events are ordinarily rejected by the fiducial volume cut, since they have a signal in the outer electrode. However, when the ionization signal is small, noise in the QI signal may cause some of them to leak into the nuclear recoil acceptance region. In the analysis of this thesis, we have made a special effort to avoid such events in setting the timing cuts; see Sec. 9.4.1.
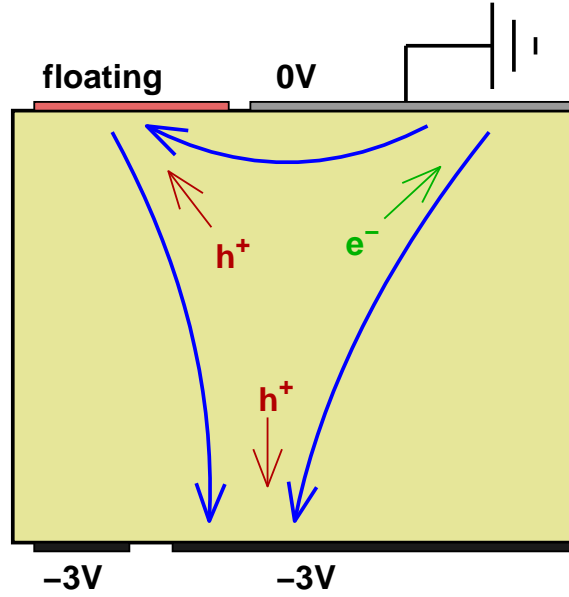
Figure 5.9: Electric fields (heavy arrows) and motion of charges when tangential fields exist.

5.8   Tangential fields

In some cases there may be electric fields parallel instead of perpendicular to the detector surface. The gap between the inner and outer electrodes, with bare, non-conductive crystal surface between, implies the existence of small tangential fields at the charge side. Additionally, if conduction through the amorphous silicon is not able to hold the floating metal surfaces, or bare unmetallized surfaces, on the phonon side to 0V, tangential fields will form between these areas and the grounded phonon sensors. These tangential fields imply that there must be locations of null field, where charges can recombine before they are effectively separated. If the tangential fields are sufficiently strong, interactions near the null region may have their electrons drift to the grounded phonon-side grid, and their holes drift to the slightly negatively charged outer regions. This effect is shown schematically in Fig. 5.9. Such events form a negative image charge on the ionization side of the detector above, so that they create a positive-going crosstalk signal. Because the charges drift through only a small Ramo potential difference, the ionization signal is heavily suppressed. Therefore, events
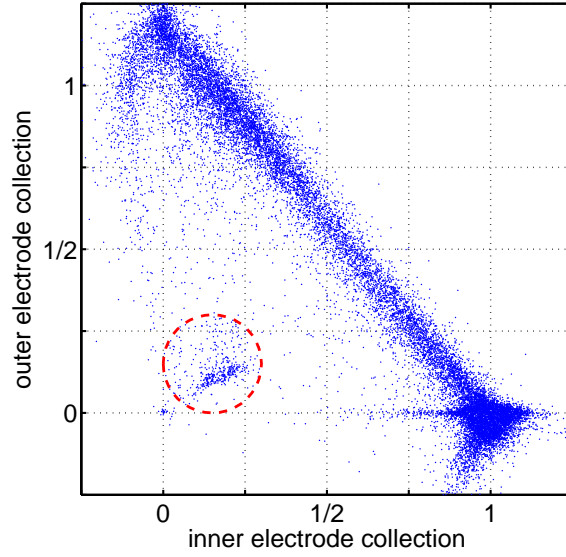
Figure 5.10: Low-yield events explained by tangential fields in T1 Z4. Points inside the circle are strongly clustered near the $-y$ side of the detector, and show large crosstalk to the Z3 ionization channels, unlike other events.

associated with tangential fields at the phonon side should have highly suppressed, tightly clustered ionization yields with energy shared between the inner and outer electrodes, and positive-going crosstalk to the detector above.

Such tangential fields should be extremely rare, since conduction through the amorphous silicon is ordinarily sufficient to prevent significant voltage differences along the surface. Evidence of tangential fields is seen in the two silicon detectors of Tower 1, Z4 and Z6, in the area near the major flats (see Sec. 2.1 and particularly Fig. 2.1). This is illustrated for T1 Z4 in Fig. 5.10. It is likely that post-fabrication detector surgery intended to remove shorts between two phonon channels cut through the metal and amorphous silicon layers, creating small islands on the phonon side with no electrical connection to the rest of the detector.

These events have ionization shared between the two electrodes, so they are generally removed by the fiducial volume cut. However, since they have low ionization, random noise on the QO measurement could cause them to leak into the nuclear recoil acceptance region. Improvements in the fabrication process should prevent the occurrence of tangential fields in all detectors beyond Tower 1.
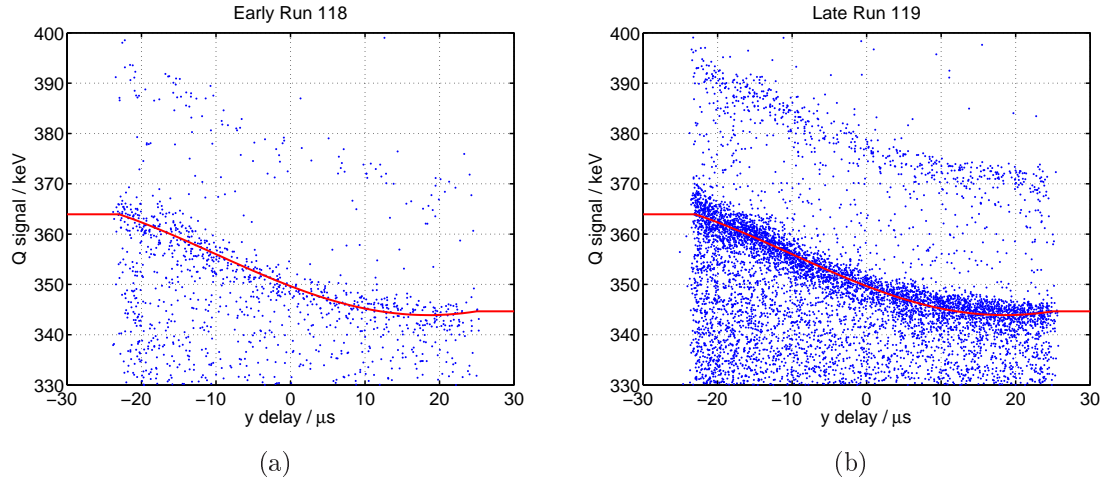
Figure 5.11: Stability of the ionization gradient in T1 Z5. Early Run 118 (October-December 2003) data is shown in (a), and late Run 119 (July-early August 2004) data is shown in (b). The 356 and 384 keV ^{133}Ba lines are clearly visible, with the same position dependence in both time periods. The solid red curve shows the form of the correction used for this detector in the combined Run 118 / Run 119 analysis.

5.9   Ionization gradient

In addition to the effects described in the preceding sections, there is a residual position dependence of the ionization signal even for bulk electron recoils well within the area of the inner electrode. This "charge gradient" is seen most clearly as a variation in the pulse height of the large ^{133}Ba peak at 356 keV in Ge detectors. Because the Si detectors very rarely absorb the full energy of 356-keV gamma rays, they do not show strong lines in the Ba calibration, and any position dependence of the ionization signal is therefore undetectable. In Ge, the charge gradient is quite stable with time, and has a generally consistent form from detector to detector. Stability is shown by comparing data from the beginning of Run 118 with data from the end of Run 119. There is no appreciable difference in the form or size of the signal, as shown in Fig. 5.11.

The gradient is oriented along each detector's own y axis. For a given gamma line in a Ge detector, the smallest ionization signal is consistently seen in the $+y$ direction, nearest the DIB and the LEDs; the largest signal is in the $-y$ direction, farthest from
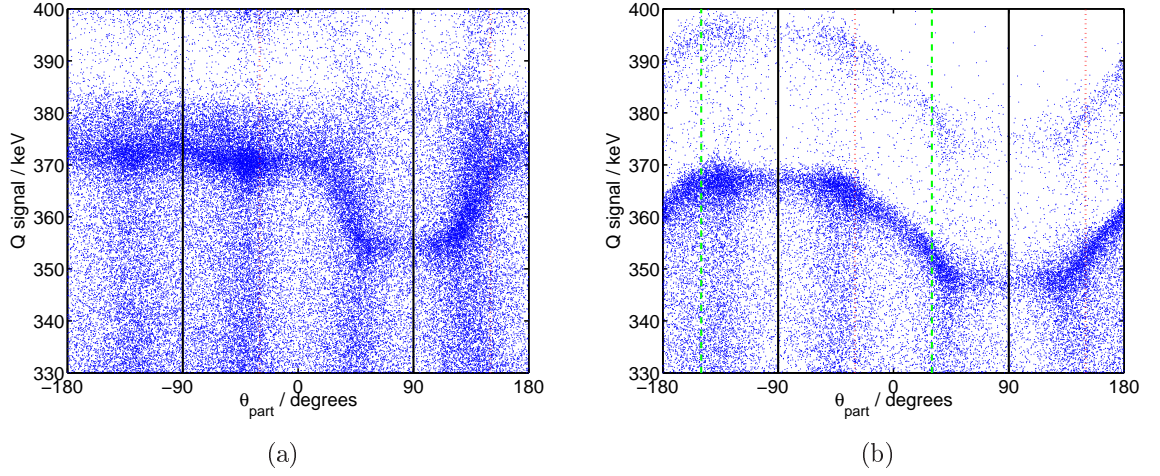
Figure 5.12: Orientation of the position dependence in T1 Z1 (a) and T1 Z2 (b). The 356 and 384 keV lines vary as a function of azimuthal angle of the scatter location within the detector. The solid black lines show the direction of the detector's own LED ($+90°$) and the opposite direction ($-90°$); the dotted lines show the direction of the LED on the detector below ($+150°$) and the opposite direction; and the dashed green lines show the direction of the LED on the detector above ($+30°$) and the opposite direction. (There is no detector above Z1). The ionization signal is symmetric about the detector's on y axis, and not skewed by exposure to its neighbors' LEDs.

the DIB. This suggests that the gradient may be caused by charge trapping in the regions of the crystal near the LEDs. This is the opposite of the expected behavior, since the very purpose of the LED flashing is to neutralize all defect sites and prevent trapping. However, it is possible that not all radiation from the LEDs has the desired effect. In addition to the near infrared light at 940 nm, the LEDs self-heat to an unknown temperature of tens or even hundreds of kelvin, and emit thermal radiation. As this heats the crystal, thermally excited charges can de-neutralize the trapping sites.

The LEDs are mounted on top of the DIB, which sits at the $+y$ major flat of each ZIP detector.[1] From this position it illuminates not only its own detector, but also the detector above. Thus, it is conceivable that any gradient in neutralization

---

[1]As described in Section 3.4.2, the LEDs themselves are offset from the $+y$ direction by about $10°$. This could potentially resolve the source of de-neutralizing thermal radiation as coming from the LEDs alone, or the entire DIB.
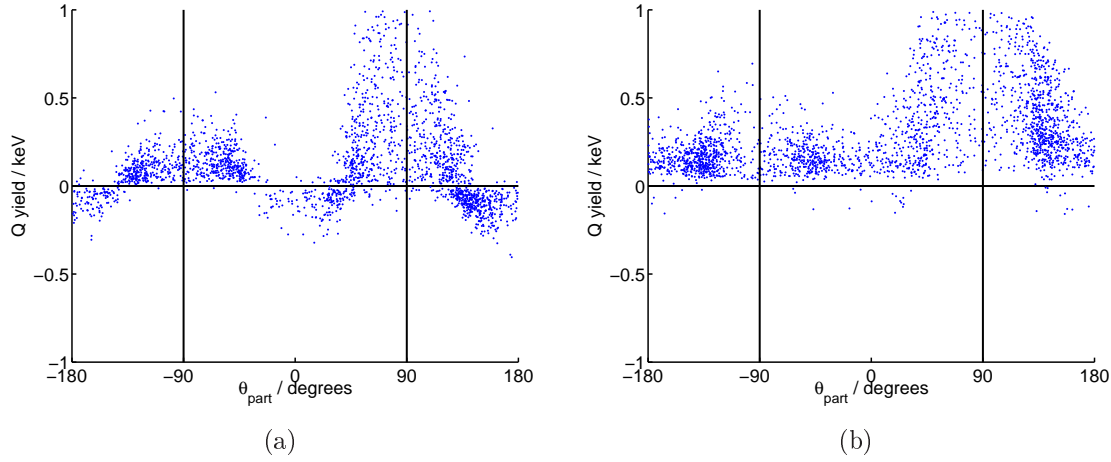
Figure 5.13: Position dependence at 0V bias in T1 Z1 (a) and T1 Z2 (b). The horizontal axis is the same as in Fig. 5.12, but the vertical axis is ionization yield, the ratio of energies inferred from charge and phonons.

could be correlated with the location of the detector's own LED, and also that of the detector below it. This can easily be tested. Each detector in a tower is oriented differently, with a 60° offset between adjacent slots. Therefore, the gradient would be somewhat skewed relative to the detector's own axis. This is not observed. In all Ge detectors, the gradient is strongly symmetric about the detector's own axis (Fig. 5.12).

A similar position dependence is seen when a detector is operated with its electrodes grounded. In this case, there is no applied drift field, so the charges can only diffuse randomly. If there are any space charges in the crystal, however, they can generate residual fields that separate the electrons and holes and pull them towards the electrodes. Even small residual fields as low as 0.1 V/cm can achieve charge collection that is a significant fraction of the full-bias case. Several 0V bias data sets were taken at the end of Run 119, and they show substantial charge collection in parts of the Ge detectors. The ionization signal is typically largest in the $+y$ direction, and close to zero in the $-y$ direction, as in Fig. 5.13.

The coincidence between the regions where charge collection is suppressed and regions where there are residual fields at zero applied voltage strongly suggests that

the two effects are related. The LED flashing must heat the crystals enough near the DIB that neutralization is somewhat compromised, and the Ge detectors reach an equilibrium state where charge collection does not improve after continued baking or flashing. This may be caused by direct heating by infrared photons from the LED itself, or it may heat the DIB enough that the entire assembly radiates into the crystal. In order to generate stray fields with a consistent polarity at zero applied voltage, the space charge must lie predominantly near one face of the crystal. This fact is consistent with direct heating by the LED, which is placed near the edge of the detector's top surface. It would also be unlikely for the DIB to reach a temperature much higher than that of the detector housing, since it is mounted to the housing with two 0-80 screws, which should provide adequate heat-sinking.

This model for the origin of the ionization gradient can be tested by carefully flashing a detector using only its neighbor's LEDs. This can be done in the Icebox in Soudan, or in a test facility run with multiple detectors. After repeatedly flashing the LEDs of the detector below, the detector above would be expected to achieve a more uniform ionization response, or in a more pessimistic case, to show a reorientation of its gradient by 60°. If the test confirms such an origin for the gradient, the current LED flashing scheme should be modified to proceed from the top to the bottom of each tower. As each detector's LED flashes the crystal will be somewhat de-neutralized towards the DIB, but the subsequent flashing of the LEDs below it should serve to undo the effect.

5.10 Ionization position correction

The gentle gradients described in the previous section have been corrected in a simple way to give a uniform calibration of the ionization response within the inner electrode region. This position correction is integrated with the calibration and removal of crosstalk. The raw output of the optimal filter is in QIOF and QOOF, in units of volts and accounting for crosstalk in the templates. The corrected quantities are qi and qo, in units of keV (electron equivalent). The position of the interaction within the detector is indicated by x_{del} and y_{del}, the delay plot coordinates. The x and y

dependences are considered independent because the primary effect is consistently in the y direction, and only a very small correction is ever required in the x direction. The form of the correction is as follows:

$$
\begin{aligned}
qi &= \frac{A \times (QIOF + B \times QOOF)}{\left[1 + C \sin\left(\frac{x_{del} - E}{F}\right)\right] \times \left[1 + G \sin\left(\frac{y_{del} - H}{J}\right)\right]} \\
qo &= K \times (QOOF + L \times QIOF).
\end{aligned}
$$

The parameters A and K are calibration factors converting volts at the front-end amplifier to keV (electron equivalent) of the interaction. Any residual crosstalk is removed by B and L. The coefficients C and G represent the magnitude of the gradient in the x and y directions, respectively; C is consistently small. The shape of the gradient is represented by a simple sinusoid with adjustable period (F, J) and offset (E, H).

Because the ionization response is flat in the outer regions of the detector, this correction is only applied within an inner window. For locations outside this window, the correction factors are held constant:

$$
x_{del} \leftarrow
\begin{cases}
x_{del}^{min}, & x_{del} < x_{del}^{min} \\
x_{del}, & x_{del}^{min} \leq x_{del} \leq x_{del}^{max} \\
x_{del}^{max}, & x_{del} > x_{del}^{max}
\end{cases}
$$

$$
y_{del} \leftarrow
\begin{cases}
y_{del}^{min}, & y_{del} < y_{del}^{min} \\
y_{del}, & y_{del}^{min} \leq y_{del} \leq y_{del}^{max} \\
y_{del}^{max}, & y_{del} > y_{del}^{max}
\end{cases} .
$$

All adjustable parameters are matched to the observed behavior of the ^{133}Ba 356-keV line in calibration data. The shape of the correction for T1 Z5 is shown in Fig. 5.11.
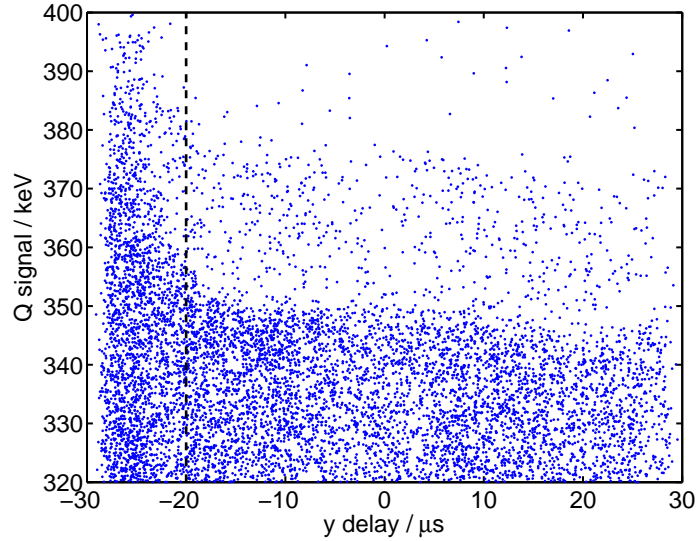
Figure 5.14: "Cliff" phenomenon in T2 Z5. The vertical line at $y_{del} = -20$ μs shows the "cliff cut" used to remove the problematic region.

5.11   "Cliff" phenomenon

In one detector, T2 Z5, there is a feature of the ionization response that is similar to the usual gradient, but more localized. In the extreme $-y$ region, where phonon quadrants B and C meet, the ionization yield sharply *increases* compared to the rest of the detector (see Fig. 5.14). This sudden shift has been labeled the "cliff" effect. This region has been found to contain a disproportionate number of events with low yield that are difficult to remove with the timing cuts. These "outlier betas" are the most problematic background for CDMS. The ionization position correction is used to remove this effect as well as possible, but it is not clear that the form of the correction is a good match to the phenomenon. The phonon look-up table correction should remove the effect in ionization yield, but it may not adequately separate events on the two sides of the sharp transition. Some events therefore have (after the correction) reduced yield and bulk gamma-like timing. In Sec. 9.3 we impose a "cliff cut" on this detector to remove the problematic region from the WIMP-search analysis.

Chapter 6

Functional-form position reconstruction

In this chapter we develop a position correction that combines a number of observable quantities to produce a single, best-estimate physical position for every event. Although each of the observables individually may have regions of foldback, pile-up, or other pathologies, a well-behaved quantity can be created by using all of them together. The approach is to make an empirical model of the detector response as a function of position - primarily from data, but with guidance from detector Monte Carlo work. The model itself is a set of functions taking physical coordinates as inputs and giving the values of observable quantities as outputs: that is, it predicts the response of the ZIP detector to a physical interaction at a given location. In the basic position reconstruction, the physical coordinates are r and θ. This framework can be generalized to include other physical coordinates: depth z, recoil energy E_R, and event type (nuclear recoil vs. electron recoil, bulk vs. surface event). The observable quantities are selected from the usual reduced quantities of CDMS ZIP detector analysis: relative partition and delay of the energy deposited in the four phonon sensors, delay of the phonon signal relative to ionization, and measures of phonon pulse shape such as rise time and fall time.

The model can then be applied in the opposite direction as well: given the set of measured quantities for one event, determine the corresponding physical coordinates

(for instance, by varying the physical coordinates to minimize a chi-squared). This works as a reconstruction of position.

Let q_i be the m observable quantities, and x_j be the n physical coordinates. Then the model is defined by a set of functions $f_i(x_1, x_2, ..., x_n; \alpha_i)$, where α_i are the free parameters of the model, fitted to the data. The chi-squared to be minimized in reconstructing the physical coordinates for an event is calculated from the difference of the observed quantities q_i and the model predictions f_i:[1]

$$\chi^2 = \sum_{i=1}^{n} \frac{\left| q_i - f_i\left(x_1, x_2, ..., x_n; \alpha_i\right)\right|^2}{\sigma_i^2}. \tag{6.1}$$

We adjust x_i for each event to minimize the χ^2. This approach has several advantages. First, given a suitable choice of functional forms for the model, the numerical model parameters can be fitted automatically. This allows the detector models to be generated for each detector with little or no human supervision. Second, it gives a natural interface with detector simulations. Third, it can naturally be generalized to a full fiducial-volume and recoil-type discrimination analysis.

6.1 Model construction

6.1.1 Expectation-Maximization

For real events in the calibration data set, the physical coordinates are in general unknown: they are "latent variables." The process of constructing the model must somehow recover the values of the latent variables. In a 2-d position reconstruction, the physical azimuthal angle θ is very close to the angle on a box or delay plot. The radial position r is more complicated. We make the assumption that each of the

---

[1]This formula neglects correlations. If correlations are nonzero but fixed for all events, then let C_{ij}^{-1} be the inverse of the covariance matrix, and

$$\chi^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} C_{ij}^{-1} \left(q_i - f_i\left(x_1, x_2, ..., x_n; \alpha_i\right)\right) \left(q_j - f_j\left(x_1, ...; \alpha_j\right)\right).$$

observable quantities varies with r in a continuous and sufficiently smooth way. We can then use the algorithm of expectation-maximization (EM), a convenient technique for fitting the parameters in models with latent variables [94].

Expectation-maximization is an iterative process with two alternating steps. The *estimation* step of EM estimates the values of the latent variables, given fixed model parameters. The *maximization* step chooses new values for the model parameters, given the current estimates of the latent variables. EM alternates these two steps to give a self-consistent model that fits the data well after several iterations. In the application to ZIP detector data, the E-step estimates the physical coordinates x_j for the calibration events. The M-step fits new parameters α_i to make the model functional forms f_i match the data. The process can be started by making initial guesses at the α_i based on easily determined information like the width of the box and delay plots. For the 2-d position reconstruction, θ can be calculated first in a straightforward way. Then, the E-step consists of estimating the radial position r for each event, and the M-step consists of calculating the parameters that govern the behavior of the observable quantities in r and θ. Each iteration of E and M improves the model and the reconstructed coordinates.

6.1.2   First-guess model

The expectation-maximization process must be started with an initial guess that is sufficiently good to ensure that iteration will converge to a correct model without any pathologies. It is convenient to let this initial guess be an M-step: we need an initial value for each of the model parameters α_i. We choose a functional form for each observable quantity, informed by results from Chapter 4, up to an overall scaling. The scaling is then set for each detector by taking the median value of the quantity from calibration data.

6.2   2-d position reconstruction

This section is a first reconstruction of the position of events from the Run 118 and Run 119 two-tower data in two dimensions. The model is constructed for electron recoil events. With this well-behaved r coordinate in hand, the initial model can be a stepping stone for the study of the detector response to other coordinates such as depth and recoil type.

The initial model takes the azimuthal angle θ from a corrected partition angle θ_{part}, and constructs r as a latent variable using EM based on on three phonon pulse parameters: *rppart*, *rdel*, and the delay *pdel*. The earlier position reconstructions described in 4.6 and 4.5 provide guidance about which functional forms to choose.

6.2.1   Theta reconstruction

Determining the azimuthal angle θ is a logical first step because the partition-plot and delay-plot angles are already monotonic and well behaved. When these azimuthal coordinates are compared for the same events in Fig. 6.1 they are mutually consistent, although each shows bunching near the centers of quadrants. Besides this consistency check, the correctness of delay and partition azimuthal angles can be verified using surface events shared between neighboring detectors. The most useful class of events for this test are those in which a radioactive contaminant (such as ^{210}Po, part of the radon decay chain) on a detector surface undergoes alpha decay. It is not uncommon for the energy of the recoiling atomic nucleus to be absorbed in one detector (possibly accompanied by a small part of the alpha energy), while the next detector absorbs the energy of the alpha particle. Since the detectors are in a close-packed configuration, separated by 3.5 mm with no intervening material, the interactions are constrained to be within a few mm of each other in x-y position. The good agreement of the delay-plot and partition-plot angles between the two detectors, shown in Figs. 6.2 and 6.3, confirms that these angles are physically meaningful.

To remove the bunching near the quadrant centers, a simple correction is applied to the partition angle to make the distribution of gamma events from low-background data uniform. We assume that each detector has perfect fourfold symmetry, e.g., that
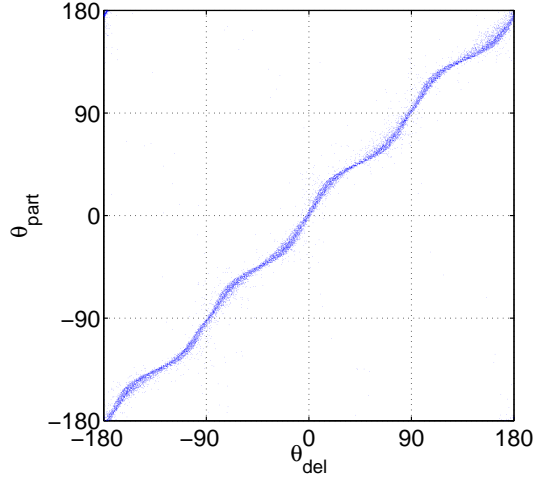
Figure 6.1: Delay-based and partition-based θ. The tight correlation indicates that each of these independent quantities is a good representation of physical θ.
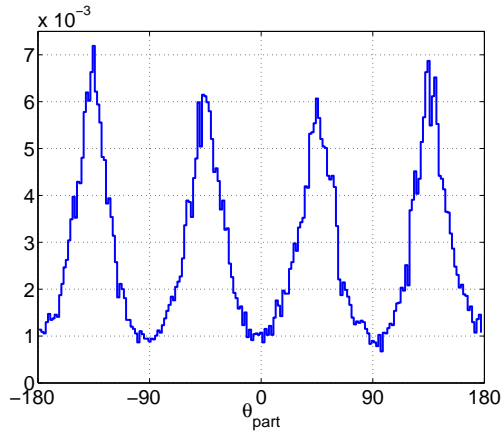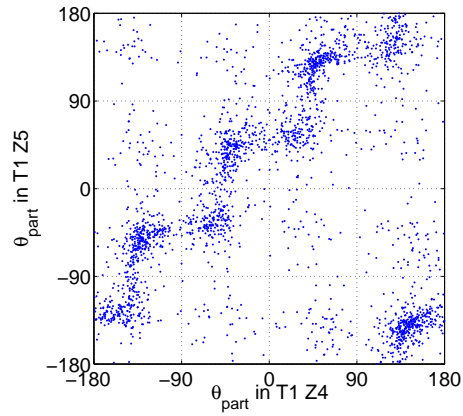
Figure 6.2: Partition-based θ. In (a), the distribution of low-background gamma events in a single detector, showing strong bunching at quadrant centers. In (b), the correlation between θ in events multiply scattering in two adjacent detectors. The stairstep pattern is a result of bunching.

(a)

(b)
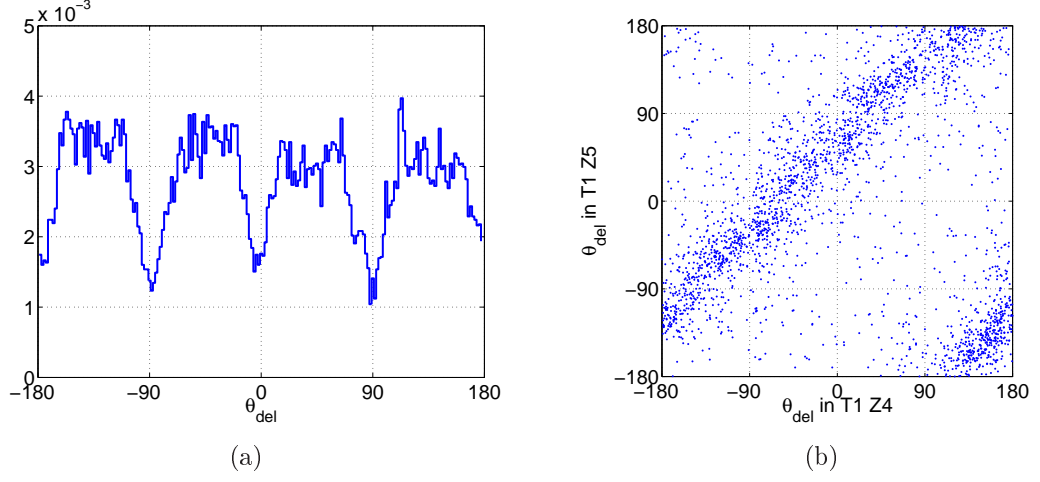
Figure 6.3: Delay-based θ. In (a), the distribution of low-background gamma events in a single detector, showing sparsity around quadrant edges. In (b), the correlation between θ in events multiply scattering in two adjacent detectors.

there is no variation between the four quadrants. The term in $\sin(4\theta)$ corrects for bunching at the center of each quadrant, while the term in $\sin(8\theta)$ is simply the next harmonic, included because the shape of the bunching is not perfectly matched by a single sinusoid.

$$\theta = \theta_{part} - a_1 \sin\left(4\theta_{part}\right) - a_2 \sin\left(8\theta_{part}\right). \qquad (6.2)$$

If the distribution function for physical angles is given by $\rho(\theta)$, it can be transformed to give the distribution of events in observed θ_{part}:

$$\rho'(\theta_{part}) = \rho(\theta) \times \left(1 - 4a_1 \cos\left(4\theta_{part}\right) - 8a_2 \cos\left(8\theta_{part}\right)\right). \qquad (6.3)$$

The parameters a_1 and a_2 are adjusted to make this distribution function match the observed distribution of points in the low-background data set, with $\rho(r) = \text{constant}$.

However, the assumption that θ can be reconstructed independently of r is not perfect. Very near the center of the detector, where the edges have little effect, *xppart* and *yppart* both respond linearly and independently, so that $\theta_{part} = \theta$ with no bunching. Around the corners of the box plot, there is considerable bunching,

apparently in both r and θ. This correlation is an added complexity, but it is not large in magnitude. We neglect it for the present, and will briefly consider it again in Sec. 6.3.1.

6.2.2   Functional forms and parameter seed values

For the initial position reconstruction, only three observables are used: partition radius *rppart*, delay radius *rdel*, and delay *pdel* of the summed phonon signal relative to ionization. The model is very simple, with only a few fitted parameters. The azimuthal dependence is limited to a single sinusoidal form with fixed period of $\frac{\pi}{2}$. This forces the detector model to have exact fourfold radial symmetry. The functional form for *rppart* is suggested by Fig. 4.6(a): a parabola passing through the origin, with its apex at $r \approx 2.5\,\mathrm{cm}$. Similarly, the functional form for *rdel* is suggested by Fig. 4.6(b): a kinked line passing through the origin, rising linearly until $r \approx 3.0$ cm and then linearly decreasing again. The final functional form, that for *pdel*, is a flat, constant value below $r \approx 3.0$ cm; outside this radius, it rises linearly. The sharp bends are smoothed over a radius $\tau = 0.1$ cm in order to avoid singularities in the minimization routines. The approximate formulas below make use of the Heaviside step function $H(x)$, replaced by a smooth sum of exponentials in the exact formulas. The radial dependences (shown in Fig. 6.4(a-c)) are given by

$$
\begin{aligned}
rppart \sim q_1 &= b_{11} r (1 - b_{12} r) & \text{(6.4)} \\
rdel \sim q_2 &= b_{21}\left(b_{22} - \tau \log\left(1 + e^{\frac{r - b_{22}}{\tau}}\right) - \tau \log\left(1 + e^{-\frac{r - b_{22}}{\tau}}\right)\right) & \text{(6.5)} \\
&\approx b_{21} r + b_{21}(b_{22} - 2r) H(r - b_{22}) \\
pdel \sim q_3 &= b_{31} + \tau b_{32} \log\left(1 + e^{\frac{r - b_{33}}{\tau}}\right) & \text{(6.6)} \\
&\approx b_{31} + b_{32}(r - b_{33}) H(r - b_{33}).
\end{aligned}
$$

Each of the parameters b_{ij} is permitted to have an azimuthal dependence of the form $b_{ij} = a_{ij} + c_{ij} \cos(4\theta)$. The only exception is b_{31}, which cannot vary with θ since q_3 must be single-valued at $r = 0$.

As a starting point for the EM process, reasonable initial values are needed for
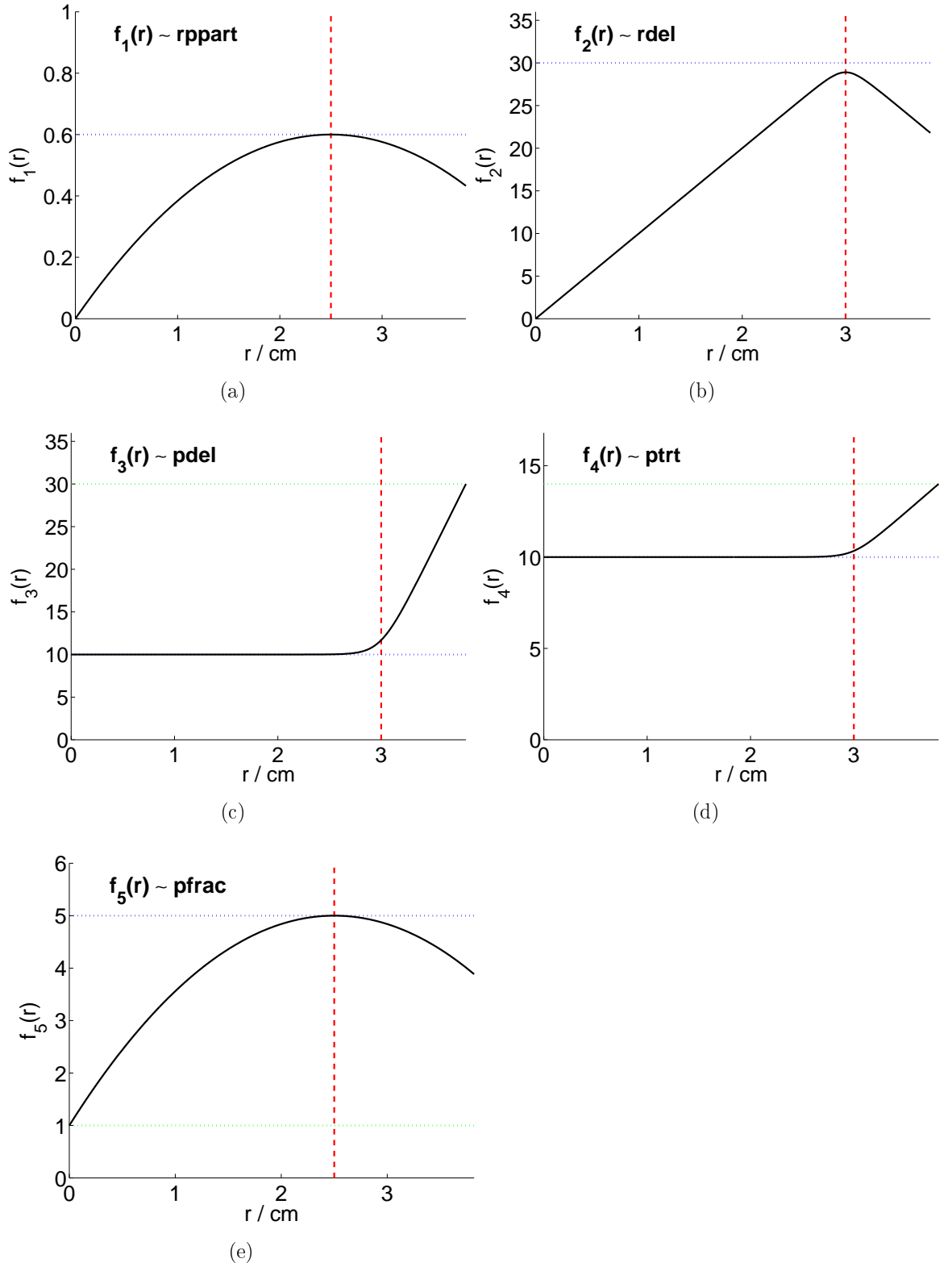
Figure 6.4: Functional form chosen for *rppart*, *rdel*, *pdel*, *ptrt*, and *pfrac* in the first-pass ZRM analysis, as given by Eq. 6.4-6.9. The curves shown use typical values of the free parameters for Ge detectors. The dashed (red) vertical lines show the radius at which each functional form typically folds back or becomes constant. The dotted (blue and green) horizontal lines indicate typical minimum and maximum values for each quantity.

all the free parameters in the model. We base these choices on typical ZIP detector behavior and histograms of the measured quantities for each detector. No azimuthal dependence is allowed in the first M-step:

$$c_{ij} = 0 \; \forall i, j. \tag{6.7}$$

The radial dependences are set in the following way:

$$
\begin{aligned}
a_{11} &= (\text{mean of } rppart + 1\sigma) \times 2/2.5 \\
a_{12} &= 1/(2 \times 2.5) \\
a_{21} &= (\text{mean of } rdel + 1\sigma)/3.0 \\
a_{22} &= 3.0 \\
a_{31} &= \text{mean of } pdel \text{ near center of detector} \\
a_{32} &= (\text{mean of } pdel \text{ in outer electrode} - a_{31})/(3.81 - a_{32}) \\
a_{33} &= 3.0
\end{aligned}
$$

The standard deviations used in the first E-step are similarly reasonable, representative values: $\sigma_{rppart} = 0.7$, $\sigma_{rdel} = 35$, and $\sigma_{pdel} = 1.5 \times 10^{-5}$. The first E-step then proceeds by minimizing the chi-squared defined in Eq. 6.1 to find the best-fit values of the latent variable r for each event.

6.2.3   Add ptrt, pfrac

After this initial round we add two more observable quantities: the rise time *ptrt* of the summed phonon pulse, and the ratio *pfrac* of phonon energy in the event's local quadrant to energy in the opposite quadrant. The rise time *ptrt* is modeled by a constant value near the center of the detector and a linear rise near the edge, with the kink smoothed by an exponential (exactly the same functional form as for *pdel*). The phonon fraction *pfrac* is modeled by a parabola with its maximum value near $r = 2.5 \, \text{cm}$, like *rppart*, but constrained to pass through 1 instead of 0 at the origin.

These choices are shown in Fig. 6.4(d, e).

$$ptrt \sim q_4 \quad = \quad b_{41} + \tau b_{42} \log \left(1 + e^{\frac{r - b_{43}}{\tau}}\right) \tag{6.8}$$

$$\approx \quad b_{41} + b_{42}(r - b_{43})H(r - b_{43})$$

$$pfrac \sim q_5 \quad = \quad 1 + b_{51}r(1 - b_{52}r) \tag{6.9}$$

The standard deviations are again fixed by hand to reasonable values: $\sigma_{ptrt} = \sigma_{pfrac} = 5$.

6.2.4 "R step" - make gammas uniform

The r produced in the E step maps onto physical radius, but it requires a calibration. This requires more physical input: we need a sample of events whose true distribution in space is known, to serve as a standard. There are several possible choices. The gammas from the ^{133}Ba calibration favor the side of the detector closest to the source. Results from the Monte Carlo simulation for the barium data set give the true event density and high statistics [95]. Here, however, we use a simpler approach. The gammas from low-background data should be uniformly distributed in x and y, and for sufficiently high energies, also in z.[2] We define a calibration by dividing the detector into 100 bins in θ, and assigning the low-background events in each bin a value of $r_{physical}$ that preserves the ordering from the E-step, but matches the expected distribution. The same correction can then be applied to the calibration and other data sets. This constitutes an additional "R step," performed immediately after each E step.

6.2.5 Results

The recipe described above successfully generates a well-behaved model for all twelve ZIP detectors of Towers 1 and 2, with no human intervention. This produces a good

---

[2]This is only approximately true. Radiation impinging on a detector from surrounding materials is more likely to come from the sides, and has therefore a slight bias towards the outer regions of the crystal.

radius estimator for bulk gamma events, without foldback. The results are shown here for several example detectors: T1 Z1, T1 Z5, and T2 Z2.

The Ge detector T1 Z1 has strong position dependences caused by a T_c gradient and the application of an early ion implantation scheme. The valid reconstruction of physical radius for this detector is therefore a good test of the ability of this analysis to work in the face of difficult detector behavior. Figure 6.5 shows the results of this reconstruction. In (a)-(e) the horizontal axis is the reconstructed physical radius after making the low-background gammas uniformly distributed. The vertical axes show the observed and model values of the five input observables as a function of radius and azimuthal angle. The overall behavior is correct, although there remain features that are not reproduced by the model. Subfigure (f) shows the shape and orientation of the azimuthal slices relative to the ZIP detector quadrants. Figure 6.7 shows x-y plots for the same detector using reconstructed r and θ. Color shows the variation in each quantity with position across the detector. Panel (a) shows partition of charge between the inner and outer electrode, indicating that the outer electrode has been correctly reconstructed to be around the outer edge of the detector. Panels (b) and (c) show two of the commonly used position indicators, both of which peak at middle radius and fold back towards the outer edge. Panel (d) shows the ionization yield, which is lowest in the center and rises at middle radius. This is caused by the concentration of biased and working phonon sensors in the middle of the detector: outside of the central region, the TESs have T_c too low to bias, and so the phonon sensors are not active. This causes the ratio of ionization to phonons signals to increase.

By contrast, Figure 6.8 shows the beautiful results for a good Ge detector, T1 Z5, and Figs. 6.6 and 6.9 for a good Si detector, T2 Z2. The x-y plots of ionization yield reveal the stairstep pattern at the edges of the phonon sensors in crisp detail, and the symmetric delay and partition plots show that the phonon sensors are all working well.

This version of the position reconstruction, using five observables, has been run on all WIMP-search and calibration data in the Run 118+Run 119 combined data set. It will be used in some of the WIMP-search analysis, particularly some of the timing
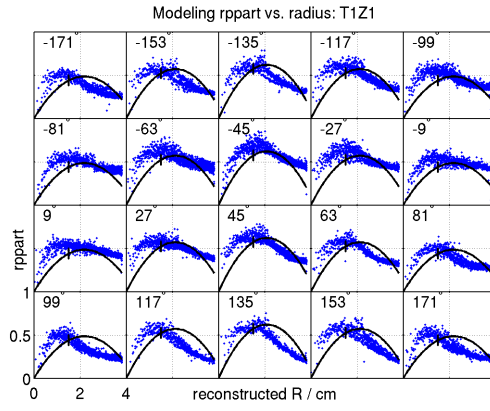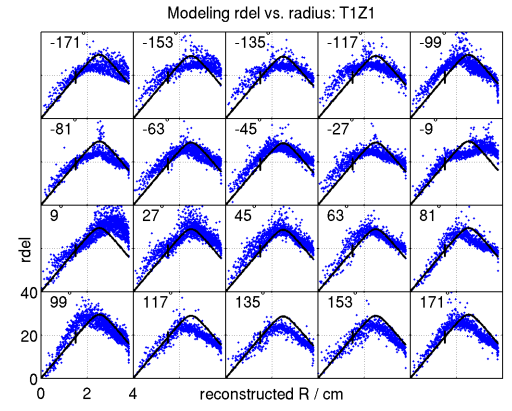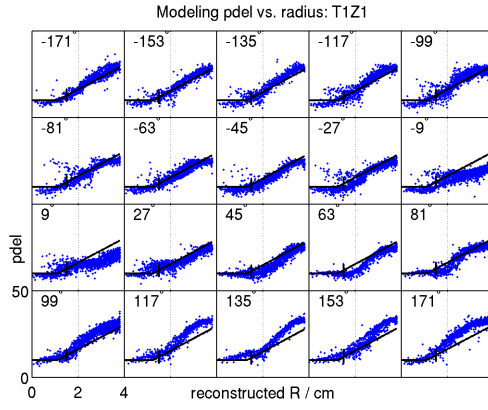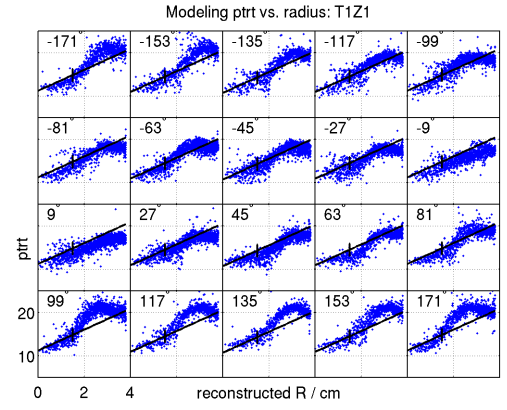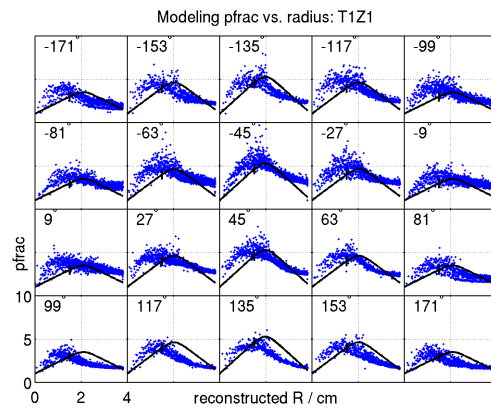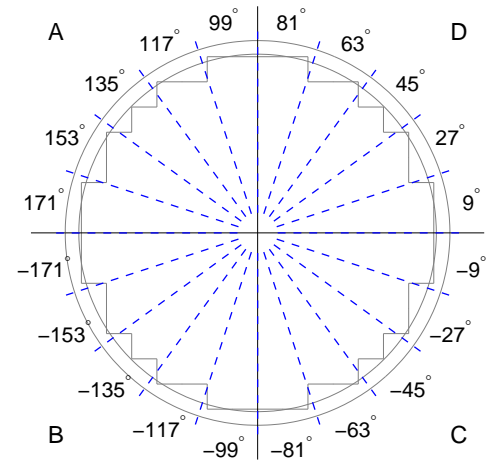
Figure 6.5: Fitted functional forms and data for detector T1Z1. The data and fitted forms are shown in 18° azimuthal bins, although the fitting is unbinned. The locations of the azimuthal bins are shown in (f).
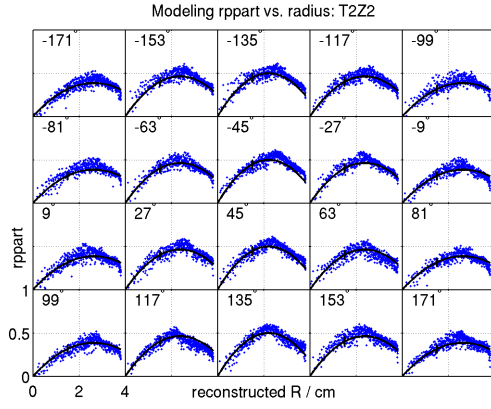
(a)
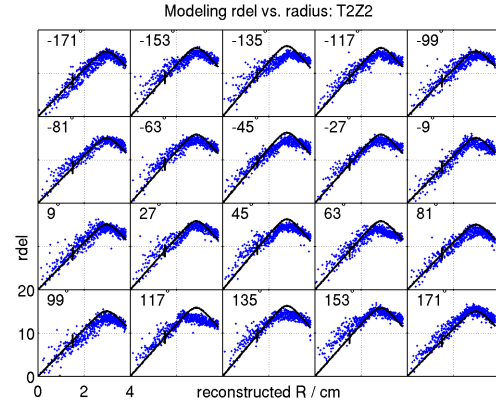
(b)
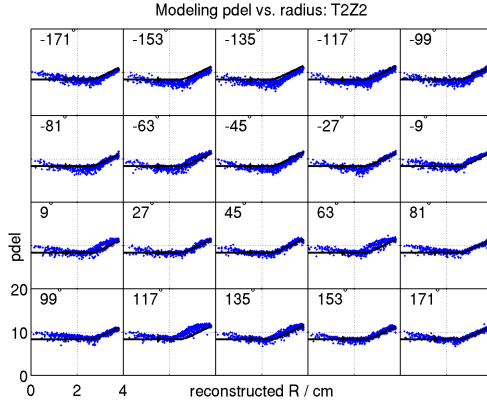
(c)
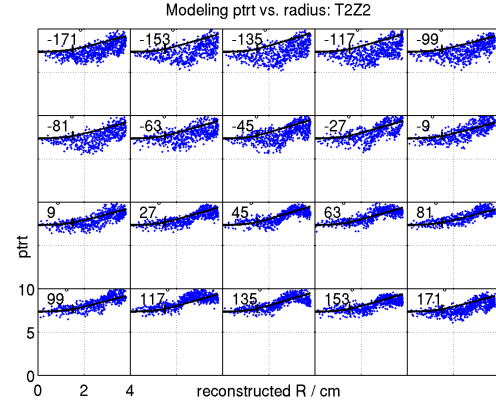
(d)
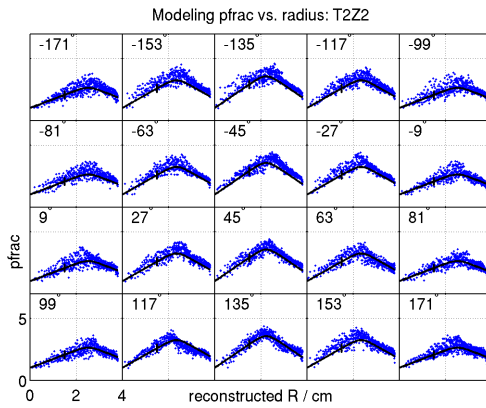
(e)
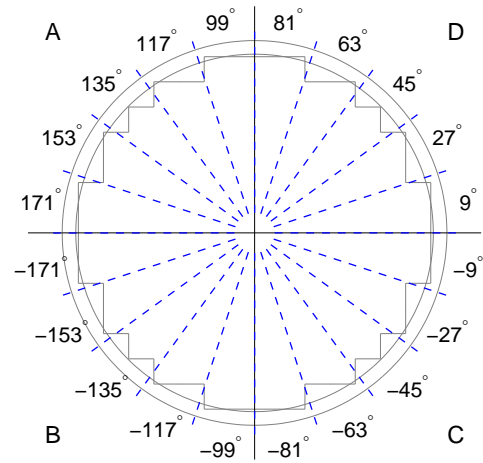
(f)

Figure 6.6: Fitted functional forms and data for detector T2Z2. The data and fitted forms are shown in 18° azimuthal bins, although the fitting is unbinned. The locations of the azimuthal bins are shown in (f).

(a)   (b)

(c)   (d)
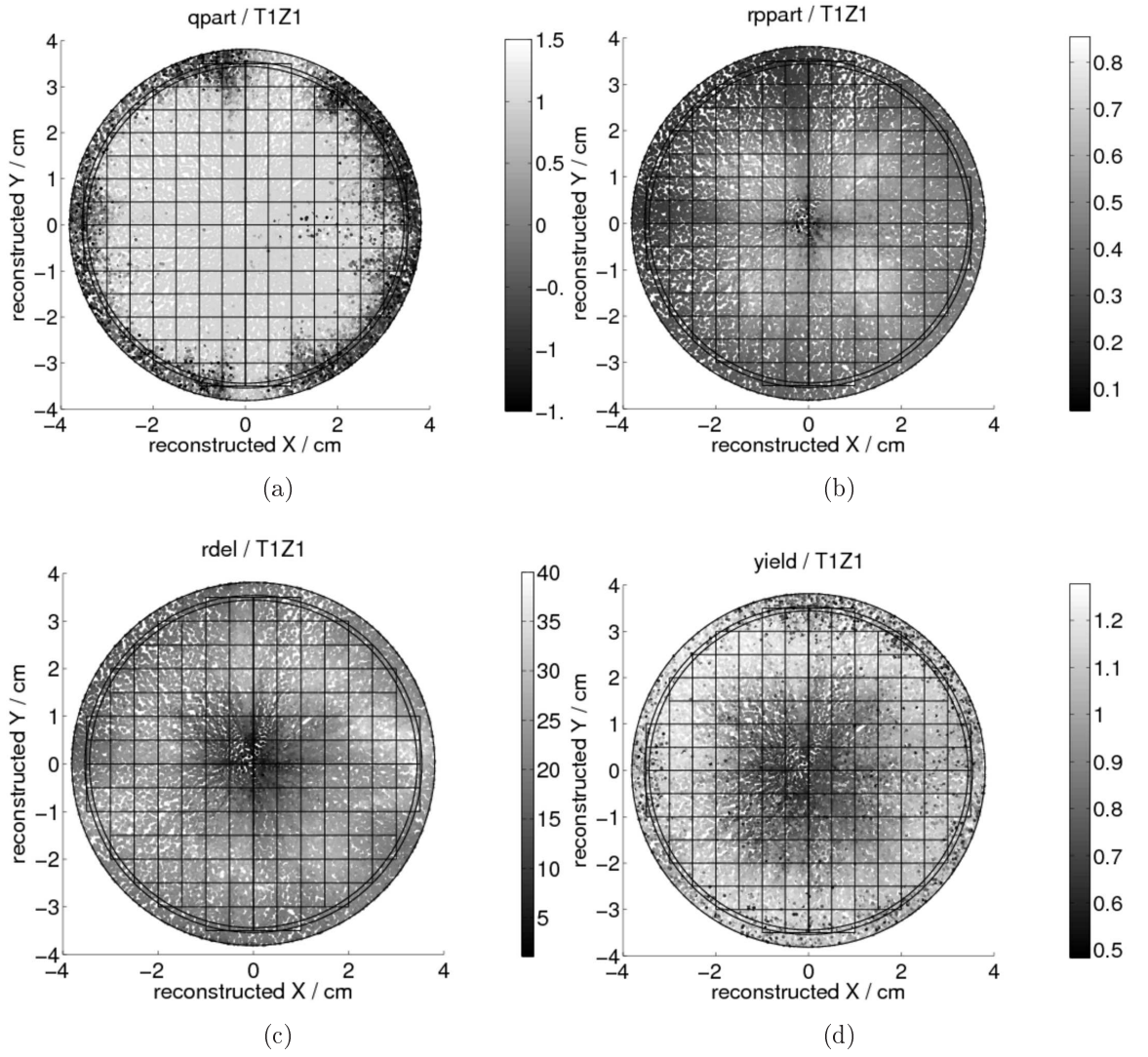
Figure 6.7: Results of the position reconstruction for detector T1Z1. Reconstructed x and y are calculated from corrected θ, and r after the last R-step.
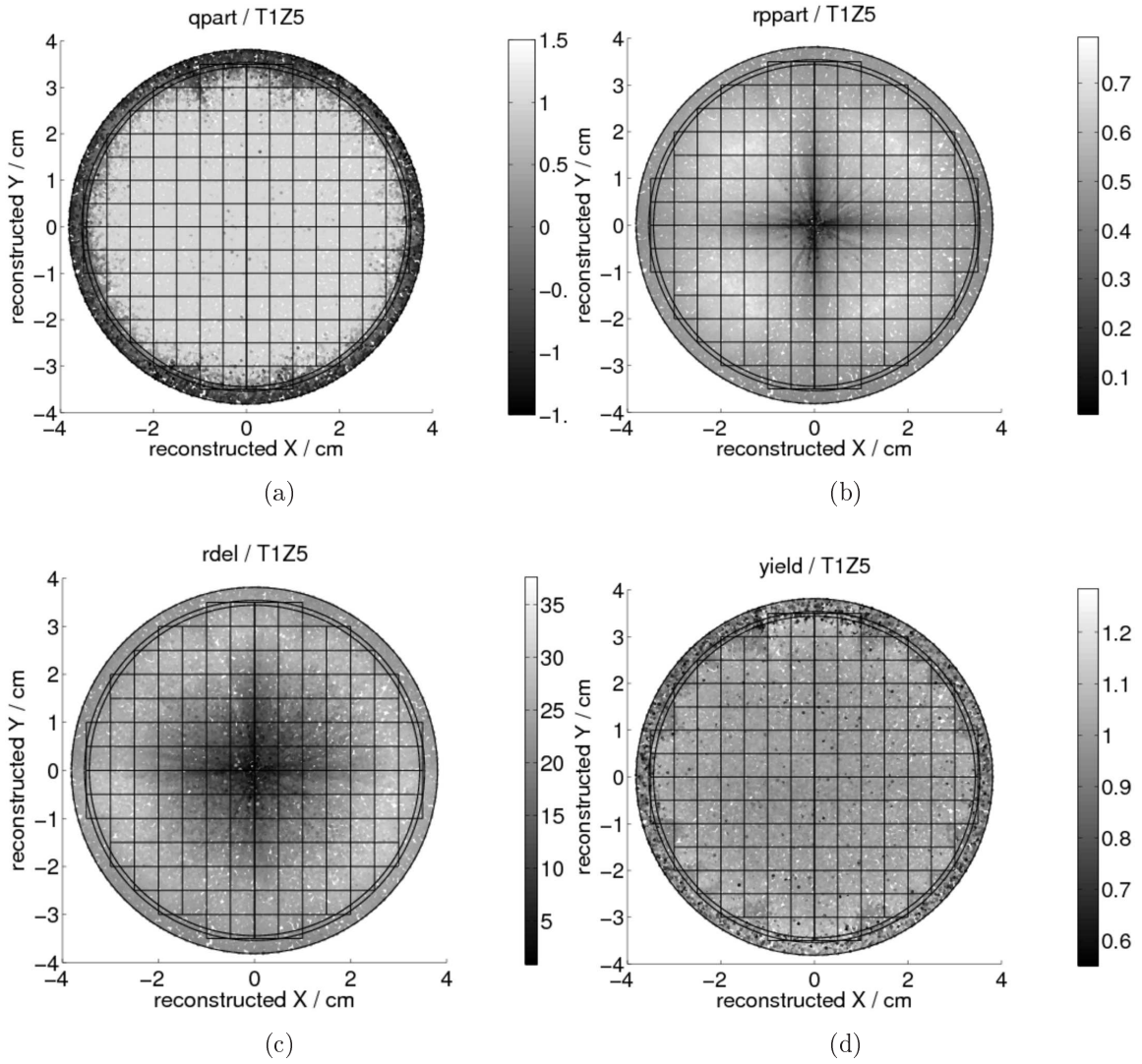
Figure 6.8: Results of the position reconstruction for detector T1Z5. Reconstructed x and y are calculated from corrected θ, and r after the last R-step.
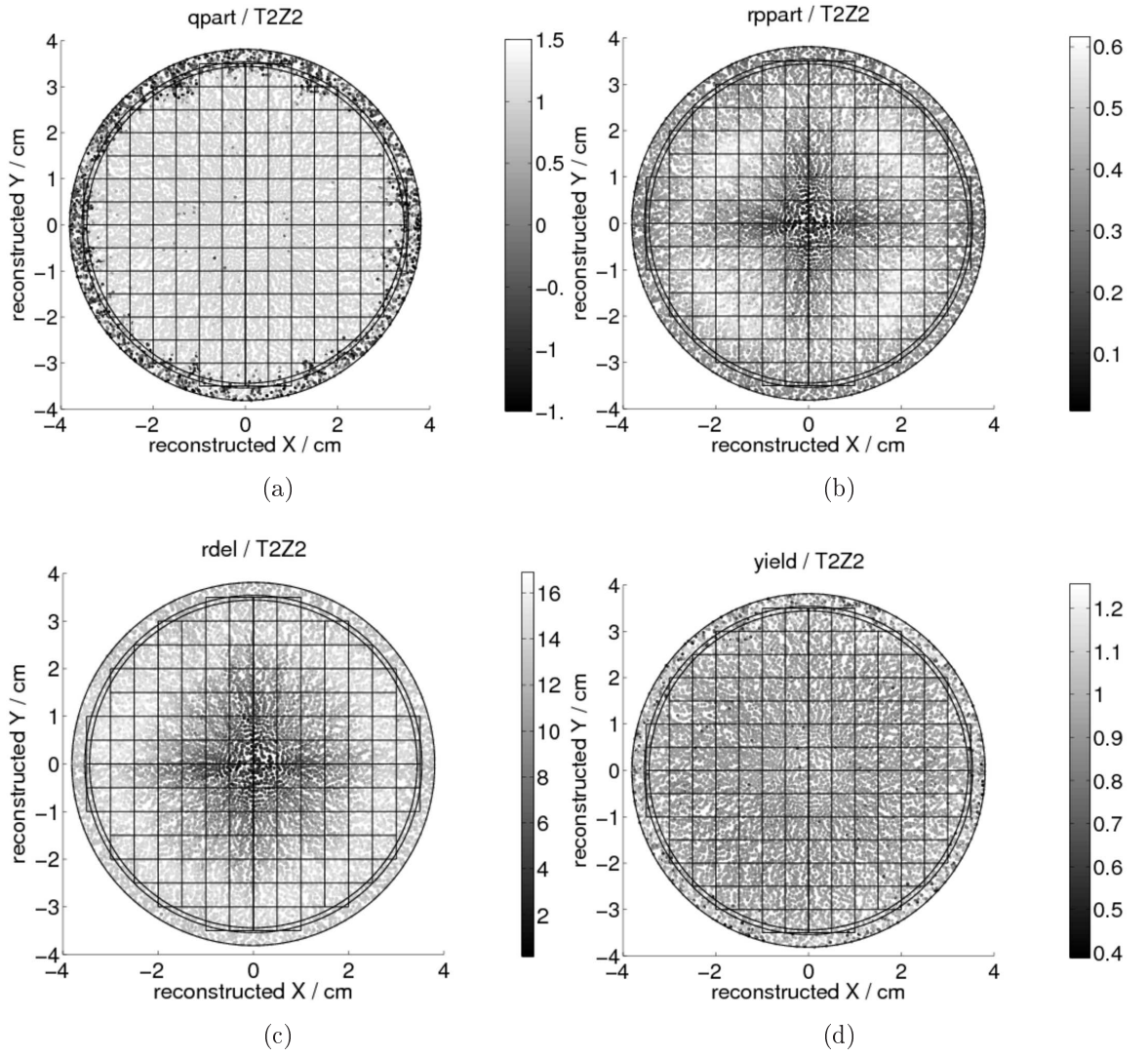
Figure 6.9: Results of the position reconstruction for detector T2Z2. Reconstructed x and y are calculated from corrected θ, and r after the last R-step.

cuts in Chap. 9. In order to make the calculation of the E step run in a reasonable amount of time, it was necessary to write custom optimization code. I have written a vectorized Nelder-Mead minimization routine which is much faster than the standard MATLAB function for this application. It may be found in Appendix D.1.

6.3   Follow-up to initial position reconstruction

With the initial position reconstruction in hand, a number of useful and interesting studies of the detectors become possible. Two of them are briefly described here, with first glances at the results, and finally, we give an outline for the further development of the ZIP model-based analysis.

6.3.1   Disentangling r and θ

The conversion of radial and azimuthal parameters into physical coordinates depended on the assumption that the bunching in partition angle is independent of radius. However, this assumption is not expected to hold at all radii. The actual behavior can be checked by using multiple-scatter events, and taking advantage of some quirks of the ZIP Tower geometry.

The ZIP detectors in a tower are packed closely, with a 3.5-mm separation and no intervening material. Therefore, near-surface events in one detector are very often seen in the next detector, with a very nearby physical position. Given a multiply scattering event with a position (r_1, θ_1) in one detector, the most likely position of the scatter in the next detector is the point directly above or below it. However, the tower assembly rotates each detector by 60° relative to its neighbors. Therefore, the most likely point in the detector coordinates of the neighboring detector will be $(r_2, \theta_2) = (r_1, \theta_1 \pm \frac{\pi}{3})$. This is extremely convenient because it is not commensurable with the quadrant geometry of a ZIP detector, so that the degeneracies in reconstructing positions are broken. We can take advantage of this to construct a better θ correction that allows the amount of bunching to vary with r, and does not rely on the density of events. Fig. 6.10 shows a first version of such an analysis. We select multiple

(a)   (b)
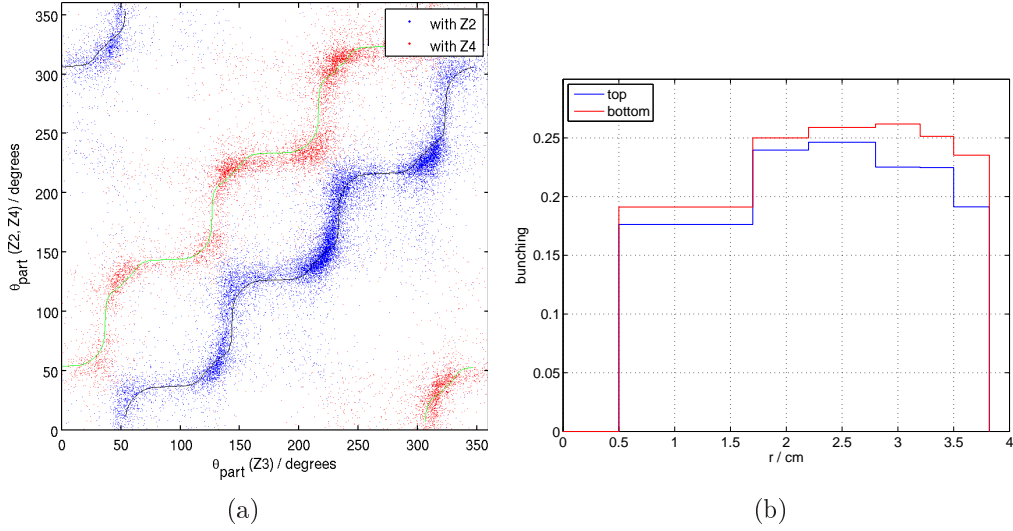
Figure 6.10: Radial dependence of the bunching in θ. In (a), the value of θ_{part} in T1 Z3 is correlated with the values in adjacent detectors for multiple scatter events. The curves show the best-fit corrections from Eq. 6.2 with $a_2 = 0$ (to fit only the leading-order effect). In (b), the "bunching factor" a_1 as a function of radius.

scatter events between the detector T1 Z3, and its neighbors Z2 and Z4. Several bins are defined in reconstructed radius r from the preceding sections. Then, within each radial bin, we adjust the parameters of the θ correction in Eq. 6.2 to minimize the squared-sum distance between the interactions in neighboring detectors. The fits in a single radial bin, 2-3 cm, are shown in Fig. 6.10(a), while Fig. 6.10(b) shows how the amount of bunching depends on radius. We find the least bunching at large and small radius, and more at intermediate r. Unfortunately, the low statistics in the central region make it difficult to determine whether or not the bunching entirely disappears at the center of the detector.

6.3.2 Depth information

There are several ways in which depth information may be encoded in the phonon signals. The ZIP detector was originally designed under the assumption that depth would be directly reflected in the overall delay of the phonon signal: events near the phonon side would be seen immediately, events near the charge side would start later

because of the travel time of the phonons, and there would be a continuous gradient in between. In practice, surface effects caused the phonon pulses to have substantially shorter delays near both surfaces, with longer delays in the bulk. There is also a substantial dependence of delay on x-y position. These effects make it difficult to extract any depth information from delay except in the surface regions.

Phonon partition may also be expected to reflect depth. Near the phonon side, a larger fraction of the initial phonons is absorbed very locally, and a smaller fraction is left to propagate across the crystal. Therefore, phonon-side events will have larger phonon partition radius, or equivalently, larger values of $pfrac$. This effect is in fact observed in the Ge detectors. Events at or within a few millimeters of the charge side have lower partition, with no depth effect; a few mm into the crystal, partition begins to increase with increasing z; and surface events at the phonon side can have very large partition. This is demonstrated in Figs. 6.11 and 6.12. Partition radius is the vertical axis in the first plot and the horizontal axis in the second plot; uncorrected ionization yield is indicated by the coloring, which can be used to distinguish different populations of events. In each case, it is apparent that the foldback happens in several different places. The innermost foldback, with a relatively larger number of events, is the maximum partition for events near the ionization side (bottom) of the detector and through about half of the depth. The outermost foldback is the maximum partition for events near the phonon side (top) of the detector. As a function of depth, we propose that the value of foldback at partition is qualitatively similar to

$$
r_{part}^{max}(\theta, z) = r_{part}^{max}(\theta) \times \begin{cases} 1, & z \leq 1/2 \\ 1/3 \times (z - 1/2), & z > 1/2 \\ \gtrsim 1.2, & \text{top surface events} \end{cases} \quad . \qquad (6.10)
$$

A similar dependence of partition on depth appears to hold in other locations as well, and not only at foldback. In one-inch-thick detectors, the depth dependence of partition appears to be further enhanced. Therefore, it may be possible in the new detectors to extract much more analog depth information in this way.
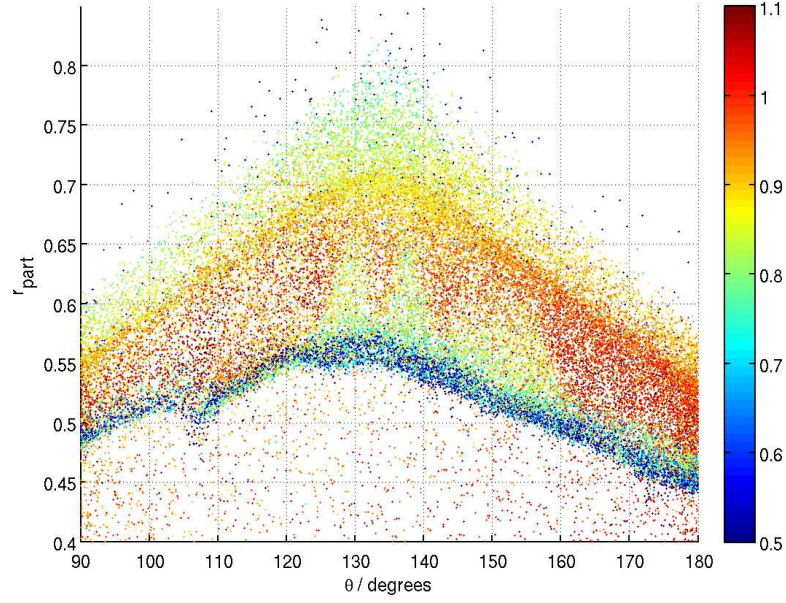
(a)
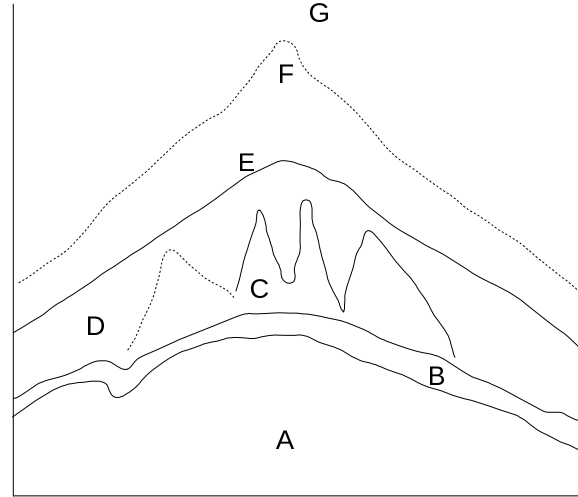
(b)

Figure 6.11: Depth sensitivity through phonon partition, in partition radius and angle. The $x - y$ coordinates are similar to one corner of the partition or "box" plot, zoomed in to the top left corner to show only Quadrant A. In this view, the box plot is conveniently shown in r and θ. Color in (a) is uncorrected ionization yield, and is used to identify the populations of events labeled in (b). The plot covers quadrant A of T1 Z5, in r and θ from the box plot. The regions are: A, events very near the center of the detector; B, outer electrode events; C, "holiday island" events; D, bulk inner electrode events; E, the foldback point for Q-side events; F, the foldback point for P-side events; and G, P-side surface events.

(a)

(b)
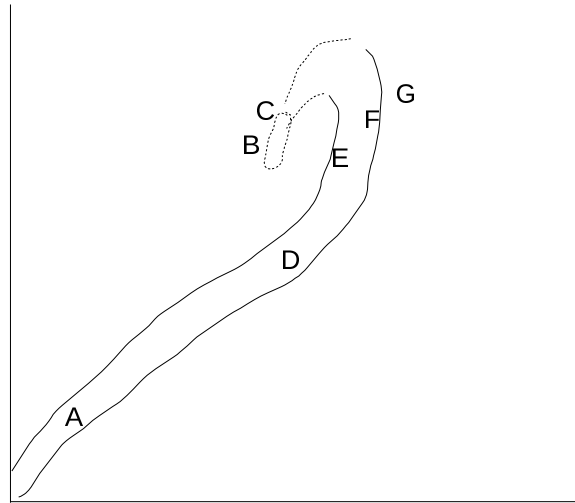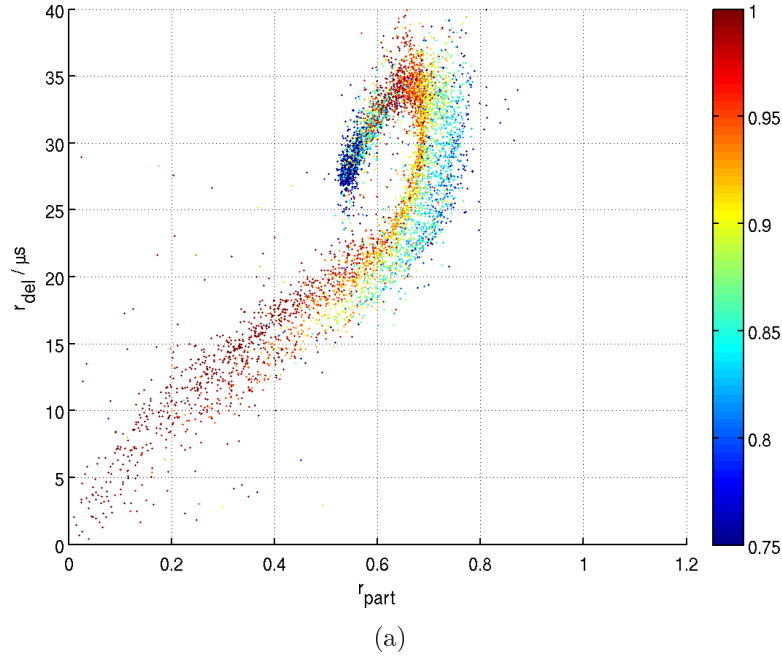
Figure 6.12: Depth sensitivity through phonon partition, in partition and delay radius. Color in (a) is uncorrected ionization yield, and is used to identify the populations of events labeled in (b). The plot covers a slice through the middle of quadrant A of T1 Z5, in r from the box plot and delay plot. The regions are: A, events very near the center of the detector; B, outer electrode events; C, "holiday island" events; D, bulk inner electrode events; E, the foldback point for Q-side events; F, the foldback point for P-side events; and G, P-side surface events.

6.4   Further directions

Much work remains to be done in the position reconstruction, and in moving beyond location to a full functional-form event reconstruction. This section lists some of the next tasks as well as longer-term goals.

- Calculate true standard deviations for the observable inputs. These would replace the current values, which are put in by hand to match the average thickness of the various quantities over all detectors. Instead, σ should be calculated from the part of the dispersion in each RQ that is not explained by position. The contribution to σ from baseline noise can be directly calculated, using the same simulation developed in Sec. 8.10.

- Adopt more physically meaningful observables as inputs. This will give the functional form model a more direct interpretation for comparison with the detector Monte Carlo and detector physics calculations. For example, the partition radius can be replaced with $P?int/PTint$, the fraction of total phonon energy (as calculated using the integrated pulse) found in one phonon sensor. Here the question mark represents quadrant A, B, C, or D. The phonon integral is to be preferred over the optimal filter because it does not depend on pulse shape; and raw quantities are preferred over corrected quantities because we wish to retain all position dependence.

- Add depth as one of the coordinates. Depth information comes mainly from partition; although we have some physical understanding of this, it complicates model building and has been left out so far.

- Adopt less idealized functional forms that allow for asymmetry and imperfections in the detector. This may benefit from input from the detector Monte Carlo effort.

- Perform a more sophisticated θ correction that takes into account the different amounts of bunching in the inner and outer parts of the detector.

- Calculate model detector responses for neutrons and surface events as well as bulk gammas, using calibration samples.

- Construct timing cuts based on the reconstructed properties of real events as compared to the model gamma, neutron, and beta responses.

- Tarek Saab at the University of Florida has plans to acquire a scanning source for use in CDMS detector testing. This would allow the detector response functions $f_i(x_j; \alpha)$ to be directly measured!

Chapter 7

Soudan data sets

The data set analyzed in this work consists of two "runs," Run 118 and Run 119, made up of about 1600 "series numbers." These first data runs followed several test and commissioning runs of the Soudan refrigerator and icebox. Once the cryogenic system in the mine was operational, Towers 1 and 2 were shipped to the site and installed in the icebox. Installation was completed in March 2003. It was decided to initially commission and run only Tower 1, consisting of the same six detectors that had previously been run in the Stanford shallow site, and only then to turn on the new detectors of Tower 2. The first Tower 1 commissioning data was taken on August 13, and stable running conditions were reached in early October. This period of data taking with one tower continued through January 11, 2004, and is known as Run 118. Commissioning of Tower 2 began with the end of Run 118. This did not require a warmup and cooldown, so the two-tower data was taken during the same physical run of the dilution refrigerator. However, for convenience, the two-tower data set was given the designation Run 119. These run numbers correspond roughly to cooldowns of the dilution unit of the Soudan refrigerator, plus an offset of 100 to distinguish Soudan runs from SUF runs.

Each time the DAQ is configured and started, a new series number is assigned to the incoming data. These series numbers are constructed from the date and time when the DAQ was started, as follows: XYMMDD_hhmmss, where Y is the last digit of the year, MM is the month, DD the day, and hhmmss is the local time in

hours, minutes, and seconds. The first digit X is a location code: 1 represents the Soudan deep site, 0 represents the Stanford shallow site, and 2 and 3 indicate the test facilities at Berkeley and Case Western. Since they were taken at Soudan in 2003 and 2004, all of the Run 118 and Run 119 data sets have series numbers beginning with 13 or 14.

Within each data set, the events are identified by event numbers of the form FFFFEEEE. The first part, FFFF, is the file number, and the second part, EEEE, is the sequential event number within the file. Each file has 500 events. The first event of each data set is therefore event number 10001, the first event of the second file is 20001, and so on.

7.1 Run 118: One tower

Run 118 was the first science run at the Soudan deep site, using Tower 1 only. The analysis of Run 118 was the subject of doctoral dissertations by Clarence Chang [83] and Vuk Mandic [53]. The WIMP-search results as published in a preliminary short paper [85] and a full publication [86] set exclusion limits that led the field using the germanium data. The blinded analysis found no candidate events, and a subsequent analysis using a different ionization energy estimator found a single event in Z5 at 64 keV. No Si results were published from Run 118 alone, since only one good Si detector was active.

WIMP-search data

The primary running mode of CDMS is WIMP-search data taking, also known as low-background data. This means that no radioactive calibration sources are inserted through the shielding, and in fact the source storage box should be removed to a separate cavern, behind the MINOS experiment. Low-background data has a low event rate, < 1 Hz, and each series number typically covers at least an hour of continuous running, and up to about twelve hours (limited by the need to transfer cryogens twice daily).

The first low-background data was taken on August 29. The period of stable WIMP-search running began on October 2 and ended on January 11 with the beginning of Tower 2 commissioning. Run 118 accumulated 65.2 live days before cuts, with a total of 1 kg of Ge and 200 g of Si target mass.

Calibration data

Soudan data taking includes *in situ* calibration with gamma and neutron sources. During Run 118, several different gamma sources were used. The initial commissioning data in August and early September of 2003 used a 1 mCi^{137}Cs source with a gamma line at 662 keV, and a 1 mCi ^{133}Ba source with lines at 276, 303, 356, and 384 keV. With the sources inserted all the way through the lead shield, to give the best resolution of spectral lines, these sources gave trigger rates of > 2 kHz. Because each event is digitized over 1.6 ms, such a high event rate leads to substantial pileup. Therefore, we purchased a weaker, 1-μCi Ba source and used it beginning on December 1, 2003. This source allowed calibration with as little intervening material as possible, while keeping the trigger rate reasonably low, and it therefore became the primary gamma calibration source for Run 118 and Run 119. Barium calibrations were taken intermittently throughout the stable running period.

Neutron calibration was performed with a 5 μCi 252Cf source. Because neutron capture activates the Ge detectors, neutron calibration is performed only rarely, and only after stable running is achieved. Run 118 included three neutron calibration runs, on November 25, December 19, and January 5 (series numbers 131125_1538, 131219_1447, and 140105_1556).

Noise performance

During the commissioning of the Soudan facility, a large effort was put into reducing the ambient electronic noise levels, especially by removing noise sources inside the RF-shielded clean room. Some components of the cryogenic monitoring and control system, such as the liquid nitrogen level meter and the base temperature thermometry, were found to be problematic and were disabled during data taking. All plumbing

lines entering the room had dielectric breaks installed, and the penetrations were insulated so that the plumbing did not touch the Faraday cage. Tracking down and eliminating noise sources was a large project conducted mainly during spring and summer of 2003, led by Vuk Mandic with the effort of the author and many others. In addition, the mains power supply to the RF room was carefully evaluated, and reworked as necessary, by Stan Orr. However, Run 118 still had substantial external noise, particularly on the ionization channels. There was a forest of prominent, distinct lines from pickup at high frequencies. Several sources of this noise were discovered during the course of Run 118 and eliminated during the commissioning of the two-tower run in early 2004.

Grounding

The grounding scheme used during Run 118 was susceptible to pickup of external electronic noise. In particular, the detector readout cables running between the front end boards (inside the RF-shielded room) and the RTF boards (in the electronics room) compromised the grounding scheme: one pin of each cable connected the electrical ground inside and outside the RF-shielded room. This connection both created ground loops, with one new loop for each detector, and allowed noise picked up in the electronics room to propagate to the ground inside the RF-shielded room. In diagnostics between Run 118 and Run 119, we verified that disconnecting Pin 20 of each detector cable at the RF room wall dramatically reduced noise pickup on the ionization channels.

As part of a general improvement to grounding, these connections were broken during January of 2004, before the start of Run 119. The improvement in noise levels can be seen in Fig. 7.1. Because the noise pickup primarily appeared at frequencies of a few hundred kHz, which is rejected well by the optimal filter used for pulse height estimation, these changes have a small effect on energy resolution. However, they are significant for pulse timing reconstruction, especially at low energies.
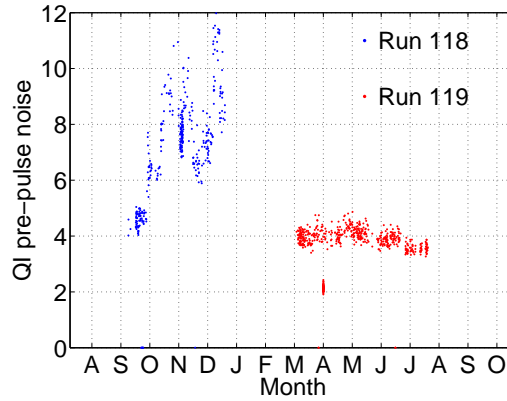
Figure 7.1: Ionization channel noise stability in Run 118 and Run 119. The standard deviation of the pre-pulse ionization baseline is shown for events in Run 118 (blue) and Run 119 (red). The noise level is lower and more stable after the grounding changes.

Sliding seal heater

The top of the helium bath is closed by a sliding O-ring seal between the fridge dewar and the insert. This O-ring was warmed by a resistive heater tape to keep it from freezing. The sliding seal heater was controlled by a thermostat, and came on for about a minute out of every 5-10 minutes, depending on the bath boiloff rate. During Run 118 the AC power supply to the sliding seal heater was picked up on the detectors' ionization channel readout, causing elevated noise during the heater's active phase. Although this extra noise did not significantly impact the quality of pulse fitting, it did cause elevated chi-squared values for ionization pulses, causing two bands of chi-squared values (see also Sec. 8.3.8). In order to simplify the analysis, it was decided to accept only the lower-noise band, effectively excluding the periods when the sliding seal heater turned on.

The heater tape was replaced on March 5, 2004 with a heater and fan, partially to eliminate this noise source. Independently, the correction of the grounding issues described above better isolated the ionization channels, so that the signal of the old heater tape no longer had any effect on noise.

Publications

The Run 118 results have been presented in an initial, short paper [85] and a longer, full treatment [86] as well as the doctoral dissertations of Vuk Mandic and Clarence Chang [53, 83]. The spin-independent WIMP exclusion limits derived from this data set were the most sensitive available at the time, and the first to fully cover the region corresponding to the interpretation of the Dama annual modulation signal as a result of WIMP scattering.

7.2   Run 119: Two towers

The analysis of this data set was the focus of the doctoral dissertations of Angela Reisetter [96], Mike Attisha [97], and Joel Sander [82]. A short paper on the initial analysis included a blinded Ge result with 10-keV threshold; a Si result with 7-keV threshold; and limits combining the Run 119 result with the existing Run 118 analysis [87]. This combined limit set the leading exclusion limits for the WIMP-nucleon spin-independent cross section at the time. The data set was also analyzed for the case of spin-dependent interactions, setting the leading limits for scattering on neutrons [27]. There was a single Run 119 candidate event, at 10.5 keV in T2 Z5, which was found to have occurred during a series number with badly impaired charge neutralization. Although this event can be confidently interpreted as a bulk electron recoil with suppressed yield because of the degraded charge collection, it was included as a WIMP candidate for all physics results.

WIMP-search data

The first low-background data set of Run 119 was taken on March 6, 2004, and stable running was achieved on March 25. The WIMP-search run ended on August 2. Run 119 took data over 74.5 live days before any cuts, for a total exposure of 111.8 kg-days of Ge data or 44.7 kg-days of Si.

Calibration data

Run 119 calibration used the same 1 μCi ^{133}Ba source and 5 μCi ^{252}Cf source as Run 118. In order to better characterize any drift or variation in detector performance over time, and to collect the best possible statistics on surface electron recoils, it was decided to take barium calibration data for approximately four hours each day.

Neutron calibrations were on May 7, June 15, July 12, and August 2 (series numbers 140507_0906, 140615_0949, 140615_1014, 140712_0849, and 140802_1341).

When a series number is started, 500 noise traces are digitized right away. These are intended to sample the noise profile in the absence of any pulses. Therefore, we ordinarily remove the source when the DAQ is restarted by pulling the source back through the source tube until it is outside the lead shield. Once the DAQ has taken 500 events, the source is reinserted. Occasionally, the barium source was left fully inserted during the noise traces, usually in order to do overnight calibration running when no the operators were not physically in the mine.

Noise performance

Between Run 118 and Run 119 the sliding seal heater was replaced and the grounding scheme was substantially improved. Both of these upgrades led to smaller noise pickup on all ionization channels during Run 119. Because the noise pickup was at high frequencies above 100 kHz, well away from the frequencies of interest for our signal pulses, this had only a very small effect on energy resolution. The improvement is shown in Fig. 7.2.

Publications

The Run 119 results have already been presented in a short publication [87], and will be included in a longer paper on the reanalysis of the Run 118 and Run 119 data sets performed in this work. In addition, an analysis in terms of spin-dependent couplings has been performed for the first time [27].

(a)   (b)
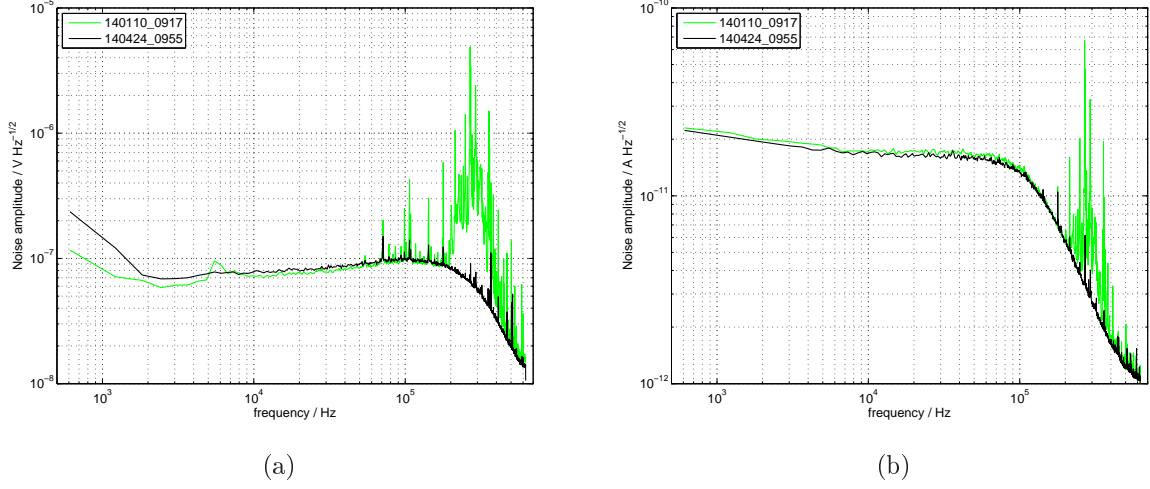
Figure 7.2: Noise spectra from Run 118 and Run 119 in ionization and phonons. The data sets compared are 140110_0917, a low-background data set from near the end of Run 118 with high noise pickup, and 140424_0955, a low-background data set from Run 119. In (a), ionization noise on QI of T1 Z1. In (b), phonon noise on PC of T1 Z1.

7.3 Excluded periods

Certain data sets contain periods in which the data quality was potentially compromised by elevated noise levels or other problems. Although these data sets are otherwise usable in the WIMP-search or calibration analysis, the problematic parts must be removed by a bad data cut. The specific reasons and ranges of excluded events are listed below.

7.3.1 Transfers during data sets

Three low-background data sets, 131011_2115, 131014_1845, and 14508_1035, continued to run during cryogen transfers. Because transfers cause elevated noise, these runs showed abnormally high trigger rates after the transfers started. The high-trigger periods were excluded by cutting out 131011_2115 after event #360222, 131014_1845 after #280027, and 140508_1035 after #100000.

During one barium calibration data set, 131211_0920, the level sensor for the

liquid nitrogen bath turned on. It is ordinarily inactive except during transfers, precisely because it causes elevated noise on the phonon channels. The bad period was excluded by cutting out events #1200500-1320500.

7.3.2 LED flashes during data sets

Two barium sets, 140630_0917 and 140707_1014, as well as the low-background data set 140525_1552, continued to record data during LED flashing for neutralization. LED flashing causes elevated temperatures and degraded performance. These periods are excluded by cutting 140630_0917 events #2680376 and 2680377; 140707_1014 events #2500411-2510095, and 140525_1552 events #60123-60128.

7.3.3 Spurious events during Cf data

The californium source was inserted after the first 500 events (used for noise analysis) of each data set. The operator's notes for 140105_1556 note that there were spurious events during the insertion of the source. Therefore, the second 500 events, corresponding to event numbers 20000-20501, are excluded.

7.3.4 Partial loss of power

During the low-background data set 140407_0913, the circulation pumps turned off, and the SQUIDs lost lock. Therefore, all events after #70155 are excluded. During the low-background data set 140611_1708, a power outage caused some pumps and other electronics to turn off. Events after #40428 are excluded.

7.3.5 Z6 not working

The phonon channels of T1 Z6 were not working during the barium calibration sets 140513_0919, 140525_1338, 140608_0932, and 140622_0938, although the detector was configured and set to trigger. These data sets are excluded for T1 Z6 only.

In addition, the early part of 140622_0938 was dominated by noise triggers from the non-functioning detector T1 Z6. Therefore, events up to #530003 are excluded

altogether.

7.3.6   Detectors turned off

During some runs, not all detectors were active. This was deliberate in some Run 118 barium calibrations, when separate runs were taken for each detector individually. At other times, some detectors had problems with front end electronics or SQUID stability and were turned off for affected runs. This was not infrequent for the detector T1 Z6 during Run 118. These data sets are cut out for the inactive detectors.

7.3.7   Wrong or no triggers

The ordinary trigger mode for CDMS-II is the "phonon low" trigger. This means that the signal on at least one phonon channel must exceed a certain low threshold in order to trigger an event. However, several other trigger modes are also supported. The "phonon high" trigger demands a much higher energy, around 100 keV, to trigger. There are also "ionization high" and "ionization low" triggers, as well as "gated triggers," which require the pulse height to lie between the low and high thresholds. A few data sets do not have the phonon low trigger set for all detectors. In some cases this was caused by a desire to test the other trigger modes. In other cases, the DAQ was inadvertently configured to use a different trigger mode. In a very few cases, the DAQ apparently did not work properly, and used a different trigger mode than the one selected by the operator.

Data taken with different trigger modes may still be used in some cases. For barium and californium data, either the phonon low or ionization low trigger is acceptable. For WIMP-search data, a good understanding of the threshold is very important. Therefore, we require that the DAQ be configured for phonon low triggers only, and exclude any data sets in which the DAQ actually used a different trigger. These data sets are retained for any detectors that were properly triggering on phonon low.

7.3.8 Crash after flashing

The low-background run 140422_1951 was paused at midnight on April 23 for LED flashing. After the flashing, the experiment was automatically resumed, but the detectors were not correctly configured again. The SQUIDs were not re-locked, so that the DAQ recorded uninterrupted, random noise with continual triggers. Therefore, the part of this run taken after midnight is excluded.

7.3.9 Trigger burst periods

During several parts of Run 118 and Run 119, there were intermittent bursts of rapid triggers on the noise in a single detector. These do not cause spurious WIMP candidates, since the traces typically have only brief excursions just above the trigger threshold. The fitted energy is then far below the analysis threshold. However, these periods do complicate analysis, in particular the calculation of cut efficiencies. Their usefulness as WIMP search data is minimal because the summed live time of a trigger burst period is near zero, even though it may contain many events. Therefore, they are removed.

One type of trigger burst was seen in November and December 2003, when spike-shaped glitches appeared on the phonon channels. Channel C of T1 Z4 was most severely affected. These periods can be identified by studying the pre-pulse standard deviation of the voltage output on this channel: many of the traces have spikes in the pre-pulse period, and therefore have very high standard deviation values. The affected runs have been cleaned by keeping only periods of time in which at least 75 consecutive events had pre-pulse noise values consistent with the absence of any noise bursts.

The second type of trigger burst occurred intermittently, mainly in Run 119. This was not caused by any noise pickup or electronic glitches, but by changes in the true baseline noise level on one detector. When the noise on a phonon channel became large enough to frequently reach the trigger threshold, a noise burst was the result. The most frequently affected detector was T2 Z2, although T1 Z1 also experienced a few short trigger bursts. The bursts are identified as periods during which at least

30% of all events triggered on T2 Z2 (or 40% on T1 Z1), with less than 1 keV in the triggering detector. Whenever this condition is satisfied over 100 or more events, the period is rejected for all detectors.

The two trigger burst cuts have a combined effect of reducing the total live time of the WIMP search data set by 1.04%, and the total number of events by 21.44%. The cuts have been set in the manner described above, with no explicit reference to the live time of the events removed.

7.3.10   Bad neutralization in Ba set

The operation logs for April 2004 note that on April 14, LED flashes during cryogen transfers were found not to be working correctly. Subsequent investigation of the record shows that there were LED flashes during all transfers except the one beginning at 6:05 am on April 14. Therefore, data taken between this transfer and the next set of corrective flashes may not have good neutralization. The only affected data set is the Ba calibration run 140414_0757, which was part of the "closed Ba set" for the initial Run 119 analysis.

The problem was rediscovered during the combined analysis, when the single data set 140414_0757 was found to dominate the timing outlier low-yield events for T2 Z6. Subsequent investigation showed that the data set has compromised neutralization for T1 Z6 and for T2 Z1, Z5, and Z6. It has therefore been excluded for the combined analysis. Although systematic KS tests have not been performed for Ba data sets, any similarly degraded periods of Ba data would also be found through their contribution of too many timing outliers to the low-yield population in affected detectors. In the absence of any such remaining effect, we can be confident that no similarly bad Ba data sets remain in the selected data. For WIMP search data, the KS tests have already removed any such problematic data sets.

Chapter 8

Combined analysis

In the combined analysis of this thesis work, the Run 118 and Run 119 data have been analyzed anew, starting again with raw data. The goal was to treat the two runs identically: with the same offline processing, the same cuts, and the same criteria for good data selection. This chapter begins by describing the significant changes made to offline processing, and especially to the farming of the data, in Sec. 8.1. Section 8.3 describes the data quality cuts, and Sections 8.4 through 8.10 describe the physics cuts *except* the surface-event rejection cuts. Finally, Sec. 8.11 calculates the efficiencies of all cuts except the surface-event rejection cuts. The timing cuts themselves are complex enough to merit their own chapter, Chap. 9.

8.1 Offline processing

The raw data taken by CDMS-II is turned into calibrated quantities suitable for humans to study by a software package collectively known as "CDMS Pipe." The first stage is "Dark Pipe," which takes as its input raw data files and auxiliary files written by the DAQ, and produces reduced quantities, or RQs, in MATLAB file format. The RQs include pulse heights, start and rise times, and information extracted from the trigger and veto records. The second stage is "Pipe Cleaner," which calibrates the RQs, removes position and energy dependences, and performs diagnostic tests. The output of Pipe Cleaner is a set of special RQ files known as the relative reduced

quantities, or RRQs.

Dark Pipe was originally written for use with BLIP detectors, and was ported to run on the analysis machines at Stanford and the test facilities. Pipe Cleaner was initially added to give a convenient set of calibrated and reduced quantities, and during the analysis of SUF Run 21, was expanded to perform the position correction. In preparation for the move to the deep site in Soudan, all the distinct versions were merged into a common code base stored in a central code repository, and the remaining BLIP code was removed. During the completion of this thesis work, a significant amount of effort has gone into upgrading, testing, correcting, and maintaining CDMS Pipe. The entire analysis chain has been put into a form that can be compiled into a stand-alone executable and run on the Fermilab computer farm. Previously, CDMS Pipe was limited to small clusters of computers under our direct management and with one MATLAB license required for each CPU. The new farming capability was first demonstrated in the reprocessing of the combined data set.

8.1.1   DarkPipe

This section is a summary of the analysis performed by Dark Pipe, the first-pass automated analysis software of CDMS and CDMS-II. A thorough treatment may be found in the doctoral dissertation of Michael Attisha [97], who shepherded the software through the first two years of data taking in Soudan. At the heart of Dark Pipe lie four different algorithms for estimating pulse heights, and one robust routine for extracting timing information from raw pulses.

Optimal filter

Ionization and phonon pulse heights are calculated in the frequency domain using optimal filtering. In an optimal filter, frequency components which are represented in the signal are weighted more heavily, and frequency components that are represented in the noise are de-weighted. The trace for each event is integrated against the filter, with an appropriately aligned start time. To construct the filters, a noise sample is taken as part of each series number. Five hundred random trigger events are collected

immediately, at the beginning of each data set. After this initial sample, random triggers are interspersed with ordinary events to make up 5% of the total. These noise traces are used in both ionization and phonons. The signal pulse in phonons is represented by a fixed, functional-form template with an adjustable rise time and fall time for each detector. The template in ionization is constructed by averaging over many real pulses in each detector.

In the past, only the initial 500 random triggers have been used to generate the noise spectra for the optimal filter. In this analysis, for the first time, random triggers taken during the whole data set can be used. We have chosen to use all random triggers for low-background data sets. This allows the optimal filter to more accurately reflect conditions during data taking, including any changes over long low-background data sets, and possible systematic differences between the beginning of a run and subsequent periods. For calibration data, we continue to use the first 500 noise traces in the interest of speedy processing.

The random triggered events are intended to sample baseline noise in the absence of an event. If there happens to be a real pulse within the time window of a random triggered event, that event must not be used in optimal filter generation. Whenever real pulses are included in the noise sample, the optimal filter will be tuned to filter out the frequencies of real pulses, which should be retained by the filter. Therefore, it is very important that the filter generation code be able to consistently weed out any accidental pulses in the noise traces. The analysis code that performs this function has been broken for previous Run 118 and Run 119 analyses. Some photon calibration runs in Run 118 and Run 119 were taken with the ^{133}Ba source inserted during all of the random triggers, so that a non-negligible number of the noise traces contained gamma scatter events. As these were not rejected, the optimal filters were compromised, leading to defective pulse height estimation and calibration offsets in some photon calibration data in the published Run 118 and Run 119 analyses. In the published Run 119 result, this was addressed using an energy rescaling on affected barium calibration data. A similar rescaling was also applied to some WIMP-search data sets, with a significant impact on WIMP-search results that has not previously been recognized or documented. A first study of this effect and a revision of the

published Run 119 results are presented in Appendix A.

For the analysis in this thesis, these problems have been studied in detail, and the code has been corrected and improved. All data from Run 118 and Run 119 has been reprocessed with correct optimal filters, and stray events during random triggers can no longer compromise the optimal filter. There is no need for any rescaling, and the problems described in Appendix A do not apply to the analysis and results presented here.

Fixed template fit "F5"

The ionization channel is also fit with an alternative, time-domain routine called "F5." The fit for each detector uses four templates: one each for the inner and outer electrodes, and two to represent crosstalk between them. These are the same templates used for the optimal filter on the ionization channel. The fit has five parameters: the amplitudes of the pulses in the inner and outer channels; the pre-pulse baseline in the two channels; and a common start time for both pulses. Of these five, three are not actually free parameters of the fit, but are calculated beforehand: the pre-pulse baselines are found by simply averaging the part of the trace before the pulse starts; and the start time is taken from the optimal filter algorithm.[1]

Pulse integral

The phonon channels also have a secondary energy estimator, the phonon integral. This is found by first subtracting the pre-pulse baseline, then rolling off the trace with a high-pass filter at 100 Hz and integrating. The integral provides a more linear energy estimator, less strongly affected by position dependence and saturation effects than the optimal filter. However, it is more sensitive to noise, and has worse energy resolution for small pulses. The integral is used for analysis of events at very high energies, such as the 5.4-MeV alpha particles emitted by ^{210}Po, and also to study the

---

[1]Saturating ionization pulses (above 1 MeV) are analyzed only with the F5 algorithm, and not the optimal filter. Therefore, another start time must be used. For these pulses, the start time input to the template fit is found by walking backwards from the first saturated sample of the trace to find the last sample that's within five standard deviations of the pre-pulse baseline.

position dependence of the optimal filter.

Timing walk

The start time and shape of each pulse are analyzed by the *rtftWalk* algorithm originally created by Paul Brink. The pulses are first smoothed with a second-order Butterworth low-pass filter at 22.5 kHz (Ge phonons), 90 kHz (Si phonons), or 200 kHz (ionization). The *rtftWalk* code begins at the highest point of the filtered pulse, which is defined as the 100% point on both the rising and falling edges. It then walks down the rising edge to find the rise time points of interest, and similarly down the falling edge for fall time points. For each point of interest it walks down the slope until it encounters the first sample that is below the desired threshold, and then interpolates between the samples just above and just below the threshold[2]. This interpolation gives an estimate of timing with resolution potentially finer than the digitizer sample time of 0.8 μs. The points of interest on the phonon pulses are at every 10% along the rising edge, and at 80%, 40%, and 20% on the falling edge.

The 22.5 kHz low-pass filter applied to the phonon pulses of germanium detectors appears to be overly severe. It cuts significantly into the bandwidth of the pulse itself, stretching out the rise time and artificially pushing the start time earlier, as in Fig. 8.1. The main effect is that rise times are made more uniform, while delay becomes a more powerful discriminator. On the whole, we find that timing discrimination is not significantly impaired. Beginning with the five-tower Run 123, this filter has been relaxed to 50 kHz. For the Run 118+Run 119 combined analysis in this thesis, the Butterworth filter remains at 22.5 kHz for Ge phonon pulses.

---

[2]In the phonon *rtftWalk* algorithm, this interpolation currently has a flaw when operating on the falling edge. Instead of interpolating between the samples just above and just below the desired threshold, it actually extrapolates using the two samples just above threshold. Although this error was detected during the transition to compiled Dark Pipe, it introduces only a slight increase in noise for the fall time quantities, which are peripheral to the analysis. The code has been kept as it was for consistency within the data set. A correction will be made for processing future runs.
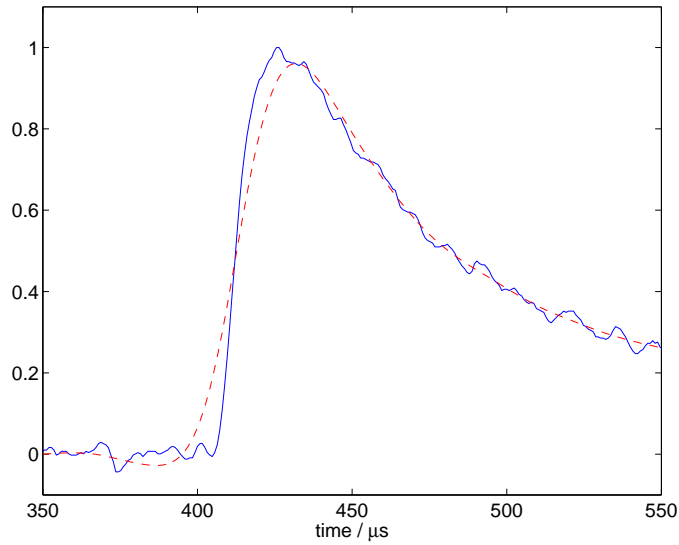
Figure 8.1: Surface event phonon trace before (blue, solid) and after (red, dashed) the Butterworth filter.

8.1.2   PipeCleaner

PipeCleaner further processes the RQs produced by Dark Pipe. It applies calibrations to all energy and timing quantities. It also performs a simple functional-form position correction on the ionization energy estimates and a sophisticated look-up table position correction of phonon quantities (see Sec. 5.10 and Sec. 4.4 respectively). Although some data quality cuts use Dark Pipe output, all physics cuts are based on quantities from Pipe Cleaner.

8.1.3   Stand-alone version

Since CDMS Pipe is written for MATLAB, it has in the past always been hosted on dedicated clusters of computers owned and controlled by CDMS member institutions, with one MATLAB license per CPU. The test facilities at U.C. Berkeley and Case Western run CDMS Pipe locally on their own machines. The Stanford Underground Facility has a cluster of seven processor Pentium III machines (mostly dual-processor, for 13 CPUs) purchased for the analysis of the shallow site data of CDMS-I and CDMS-II at Stanford. All of these use data acquisition systems built in LabVIEW,

with a relatively low maximum data rate of about one event per second.

For CDMS-II at Soudan, an entirely new data acquisition package was written by the U.C. Santa Barbara group, based on Java and C++ (see Sec. 3.7). This was designed for much faster throughput, allowing event rates above 50 Hz. The data rate is now limited by pile-up of multiple events, rather than the capability of the DAQ. Although the data rate during ordinary low-background running is much slower than this (less than 1 Hz), it is desirable to take large ^{133}Ba data sets daily to collect a sample of as many near-surface electron recoils as possible. During Run 118 and Run 119, all data taken at Soudan has been initially processed on-site in the Soudan Analysis Cluster (SAC, up to 34 machines and 65 CPUs), and all reprocessing at SUF. With the beginning of five-tower data taking, however, the data sets are too large to conveniently process all data at the SAC or SUF. Therefore, CDMS Pipe has been modified to create a stand-alone version that can be run on the Fermilab farm without requiring that each machine have access to MATLAB.

The process of compiling CDMS Pipe has required a substantial effort, including the following sub-projects:

- A new utility *get_ events* has been written to read in raw CDMS data. This replaces the older *readEvents*. Advantages of *get_ events* include compatibility across computing platforms, flexibility in the source of the data (across a local network, from local disk, or across the Internet), and incorporation of several updates to the data format. The older *readEvents* was limited to reading events across a local network and required a special server, *netServer*, to be running on each machine. The new *get_ events* has the ability to read from uncompressed and compressed local files, and works with all combinations of endianness in the files and on the local machine. In addition, it provides the new capability of extracting events across an SSH connection, so that users at any institution have easy access to events from the complete raw data repository on disk at Fermilab. Compatibility has been maintained with the optional use of *netServer*, and with the slightly different file formats used by the SUF and test facility data acquisition systems.

- The MATLAB compiler supports a substantially restricted subset of the MATLAB programming language. To compound matters, the compilers for Version 6 and Version 7 of MATLAB are entirely different, and have significantly different requirements for the M-file programs they take as input. Workarounds have been developed for those MATLAB features that cannot be compiled into stand-alone programs, such as script M-files, the *eval* statement, and many others.

- A preparser for MATLAB code has been written. This allows the necessary modifications to be made to the software in order for it be compiled, without interfering with the existing code. The preparser reads special directives in the code beginning with the sequence *%//*, and outputs a compiler-ready version. For example, wherever a script M-file is invoked, it can be marked with the directive *%//script*. This instructs the preparser to directly include the text of the script M-file. More details, along with the code of the preparser, may be found in D.2.

- A reduced but compatible version of the CDMS Analysis Package ("CAP", documented in [43]) has been written. Known as "miniCAP," it is needed for compiling PipeCleaner. CAP is a set of tools for working with RQ data in MATLAB, handling the extraction of quantities from file, buffering, and the definition of derived quantities and cuts. MiniCAP performs the same functions, but without buffering and with the ability to work with only one series number at a time.

- Scripts have been written to manage the pre-parsing, compiling, and packaging needed to generate a stand-alone CDMS Pipe package. The three user commands are *Compile_Dark_Pipe*, *Compile_Pipe_Cleaner*, and *Make_DP_Package*. The resulting package includes all needed libraries, as well as wrapper scripts to conveniently run Dark Pipe and Pipe Cleaner.

8.1.4 Reprocessing at Fermilab

The combined Soudan data set has been fully reprocessed at Fermilab, using the QCD cluster and the main Fermilab Farm. In 2005 the CDMS group at Fermilab purchased a fast machine with dual Intel Xeon processors and 8 terabytes of disk, known as *cdmstiny*. This machine has a single MATLAB license, and keeps all raw and processed data on disk. In October 2005 CDMS Pipe was first compiled on *cdmstiny* using the then-new MATLAB Version 7. The Run 118 + Run 119 data set was processed with Dark Pipe and Pipe Cleaner in November and December of 2005. The Pipe Cleaner portion of the processing has been re-run in March and September of 2006, to take advantage of improved calibrations. This reanalysis of the one- and two-tower Soudan data has exercised CDMS Pipe on the Fermilab farm. This is the primary method for processing the full calibration data set in the five-tower runs Run 123, Run 124, and further runs currently in progress.

After processing at Fermilab, all data sets have been subdivided into two halves. This split is done event-by-event, based on sequential event numbers: there is an "odd" sample and an "even" sample. The even-odd split is a separate process after Dark Pipe and Pipe Cleaner, and in this analysis, has been performed on the computer cluster at Stanford rather than at Fermilab. The even and odd portions are important for the ^{133}Ba calibration data, where the odd set is used for defining cuts, and the even set provides an independent sample for calculating leakages and efficiencies.

8.2 Trigger efficiency

CDMS is a triggered experiment, so a particle interaction must trigger the DAQ if it is to become an observed event. The CDMS trigger is performed in the RTF boards, and ordinarily operates as a user-adjustable voltage threshold on the phonon signal of each detector, summed over the four channels (see also Sec. 3.6). When this threshold is reached, a "global trigger" is issued and the signals in all channels of all detectors are digitized and saved. The threshold is typically set just above the noise, so that WIMP sensitivity extends to the lowest possible recoil energies without the rate of

triggers on noise becoming too high.

WIMPs are expected to cause only single scatter events, so we are interested only in the individual trigger efficiencies for the detectors considered singly. However, calculating the trigger efficiency requires a sample of events that do *not* trigger a given detector, which necessitates consideration of other detectors. In all past CDMS analyses, we have selected events in which any other detector generated the first, "global" trigger, and looked for a subsequent trigger in the detector of interest. The fraction with a subsequent trigger is computed as a function of the energy deposited in the detector of interest to give an efficiency function. The trigger efficiency for detector i, as a function of energy E observed in i, is

$$\epsilon_i(E) = \frac{N(\text{global not in } i, \ E \text{ in } i, \ \text{trigger in } i)}{N(\text{global not in } i, \ E \text{ in } i)}. \tag{8.1}$$

We have demanded that the global trigger arise from another detector in order to avoid being biased by events that were recorded only because of the trigger in detector i. These count as successful triggers in the numerator, but had the trigger failed in detector i, they would never have been digitized at all, and would have been lost to the denominator as well. Therefore, they should be excluded. This formula would be exactly correct if we also required that the pulse in detector i begin after the global trigger. However, we have not ordinarily imposed this condition, so there is a small bias towards making the efficiency estimate conservative. To see this, consider a pulse in which the signal for detector i crosses the trigger threshold first, followed by detector j. A trigger is generated for j. If i also successfully triggers, then i will be the global trigger, so this event is excluded from the numerator and the denominator. If, on the other hand, i fails to trigger, then j will be the global trigger, so the event is included in the denominator only. Thus, there is a class of events that can count as failures, but never successes. The effect of this discrepancy is very small, and it gives a conservative estimate.

In this analysis, I introduce a modified definition for the trigger efficiency that avoids this issue. Instead of requiring that detector i not generate the global trigger, I require only that an enabled trigger appear for at least one other detector. An
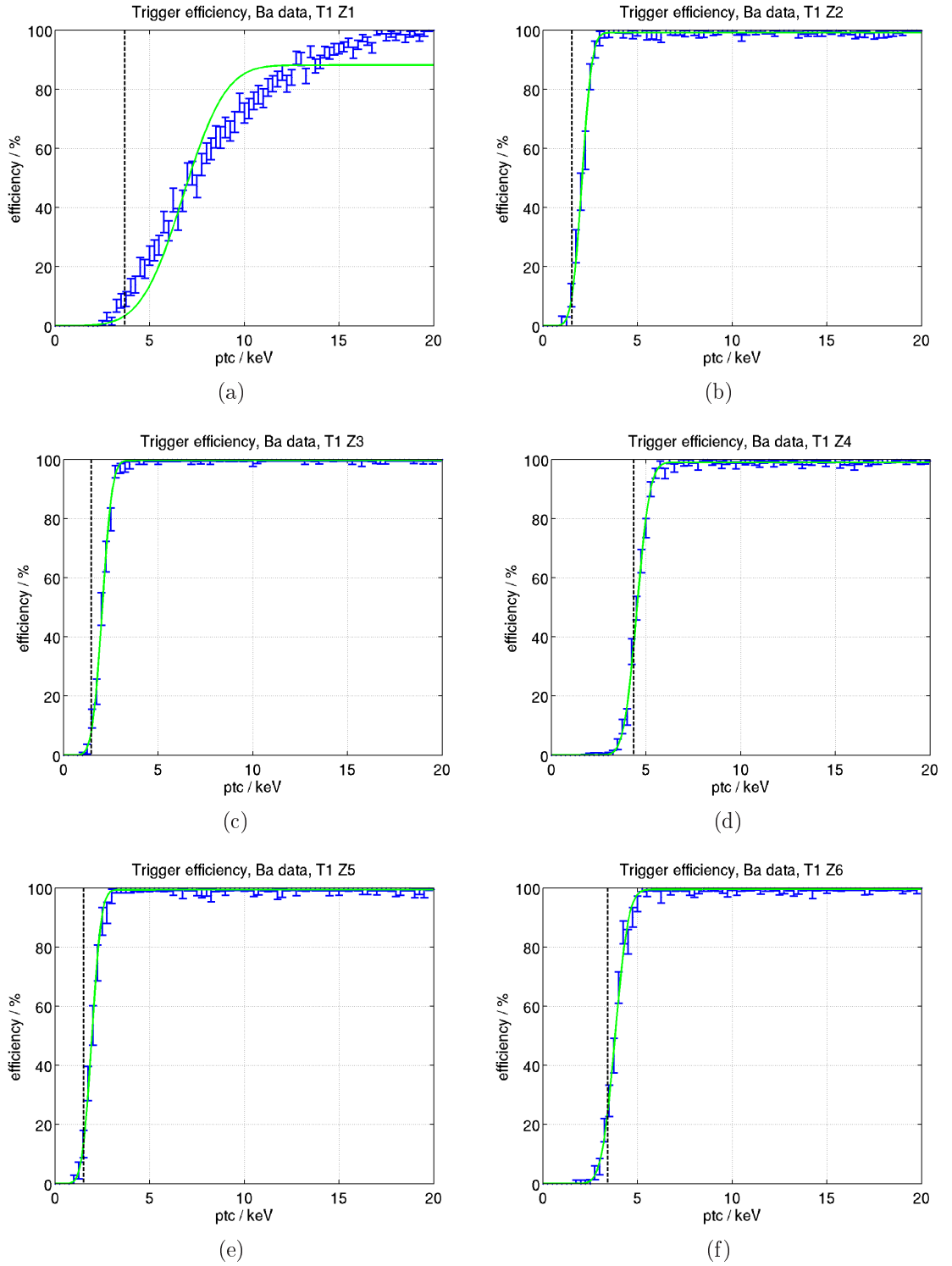
Figure 8.2: Trigger efficiency functions for Tower 1 detectors.
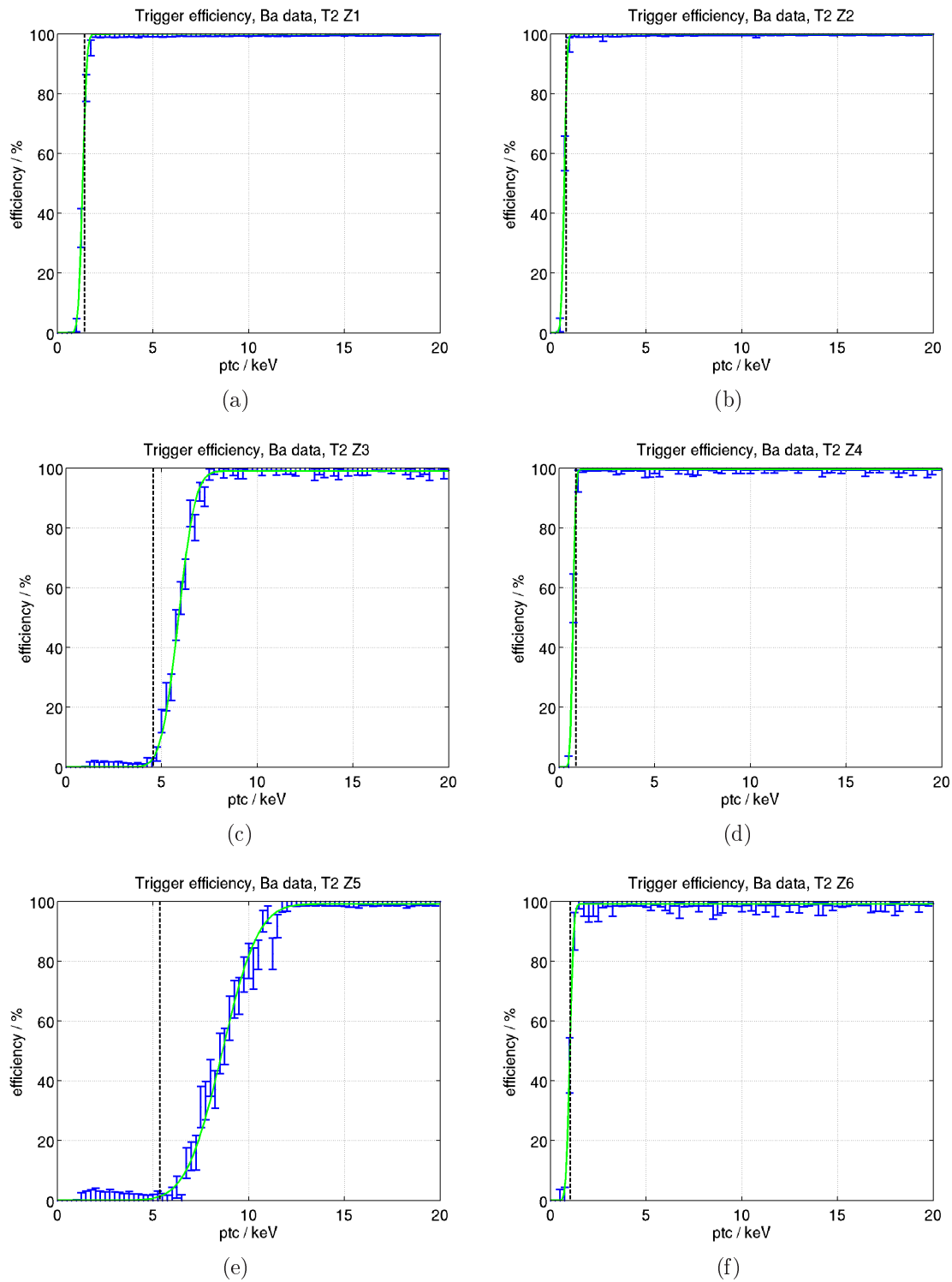
Figure 8.3: Trigger efficiency functions for Tower 2 detectors.

Table 8.1: Fitted parameters for trigger efficiency in Tower 1 (left) and Tower 2 (right) detectors.

T1	a	b	c×100
Z1	6.78	2.48	12
Z2	2.07	0.59	0.84
Z3	2.05	0.59	0.34
Z4	4.52	0.77	0.98
Z5	1.95	0.56	0.62
Z6	3.81	0.79	0.53

T2	a	b	c×100
Z1	1.32	0.24	0.08
Z2	0.72	0.16	0.06
Z3	5.88	0.99	0.99
Z4	0.73	0.15	0.47
Z5	8.61	2.09	0.85
Z6	1.02	0.22	0.84

enabled trigger means one that will generate a global trigger if one hasn't already been issued. This is not the case when one detector's triggers are suppressed in order to take data with high statistics on another detector, or to exclude a detector that is temporarily not working properly. As long as an event has an enabled trigger in at least one detector other than i, it is definitely recorded, regardless of whether i fails to generate a trigger, generates the global trigger, or generates another trigger. This then gives a truly unbiased efficiency estimate:

$$\epsilon_i = \frac{N(\text{enabled trigger in at least one other det, } E \text{ in } i, \text{ trigger in } i)}{N(\text{enabled trigger in at least one other det, } E \text{ in } i)}. \qquad (8.2)$$

This formula keeps the pre-selection criterion strictly independent of triggering in detector i, for a truly unbiased sample.

The CDMS trigger information is stored in a history buffer, recording the time of each trigger event along with a mask indicating which triggers were active. The triggers at the time of the global trigger are recorded as TGMask20 and TGTime20. Up to four preceding triggers are retained as TGMask16-19 and TGTime16-19. Up to five following triggers are retained as TGMask21-25 and TGTime21-25. A trigger relevant for the efficiency calculation may appear in any of the history buffer slots 20-25. It can in theory never appear in any of the preceding slots, because the first enabled trigger generates the global trigger and appears in slot 20. Therefore, the efficiency calculation must search TGMask20-25 for relevant triggers within a 50 μs time window. In some cases there are enough triggers that TGTime25 is within the 50 μs window. It is possible that extra triggers were present, but are not easily available

for analysis. Therefore, we exclude events for which TGTime25 $< 50\mu$s. The detailed procedure for calculating ϵ_i is as follows:

1. For each detector, select all data sets in which the phonon low trigger is enabled to generate global triggers.

2. Exclude all events for which TGTime25 $< 25\mu$s.

3. Select events with an enabled trigger in any of TGMask20-25 in at least one *other* detector.

4. Histogram the events selected in step 2 over energy E observed in detector i.

5. Select all events that additionally have an enabled trigger in any of TGMask20-25 in detector i.

6. Histogram the events selected in step 4 over energy E observed in detector i.

7. $\epsilon_i(E)$ is the ratio of the histograms from steps 6 and 4.

The trigger threshold can also be estimated by converting the voltage threshold setting into a pulse height in keV. This does not give a detailed efficiency function, but only an approximate threshold level that can be used as a cross-check on the calculation. The results of the full calculation from data, along with the estimates from voltage level settings, are shown in Fig. 8.2 (Tower 1) and Fig. 8.3 (Tower 2). The data points are from Ba calibration data, with statistical error bars. They have been fit to an error function with three parameters: central energy threshold, steepness, and efficiency at high energy. The functional form is

$$\epsilon_i(E) = \frac{1-c}{2} \times \left(1 + \mathrm{erf}\left(\frac{E-a}{b} \right) \right), \tag{8.3}$$

with the fitted parameters given in Table 8.1. This is a good fit for all detectors except T1 Z1.

The two germanium detectors of Tower 2, T2 Z3 and Z5, have rather high trigger thresholds. The detector performance and noise levels would have been sufficiently

good to allow lower thresholds, but it was not realized until the end of the run that the thresholds were so high. T1 1 Z1 suffers from a large variation in T_c with position: the central TESs are well biased and working, while many of the outer TESs are too warm, and give no signal or a small signal to noise ratio. Therefore, interactions of the same energy at different locations cause signals of different sizes, and the trigger efficiency is a strong function of position. This smears the trigger turn-on out to higher energies. It is not necessary to perform a more sophisticated efficiency fit for T1 Z1, since it is excluded as a WIMP search detector and used only as a veto. The three remaining T1 Ge detectors, Z2, Z3, and Z5, have sharp and low trigger thresholds. The two silicon detectors T1 Z4 and Z6 are from an older design, with smaller signal, and similarly have higher thresholds. The four silicon detectors of Tower 2 have very low and sharp thresholds, in two cases below 1 keV.

The detectors with higher trigger thresholds cannot detect single scatters at low energy, so the phonon analysis thresholds are adjusted accordingly (see Sec. 8.10).

8.3 Data quality cuts

A set of data quality cuts exclude any periods of time or individual events for which the performance of the detectors or the DAQ may be impaired. In most cases, these cuts are targeted at specific, known problems that occurred during the run. There are also several more general cuts that remove events for which the noise is larger than normal, or several pulses may have piled up within the digitization window.

8.3.1 Bad data cut

The first cut applied is a bad data cut to remove the problem periods described in section 7.3 and those flagged by automated Kolmogorov-Smirnov tests. It simply removes the ranges of events (or detectors) affected for these data sets. Since this cut removes parts of data sets declared a priori to be unreliable, no efficiency is calculated for this cut. Instead, it is accounted for in the pre-cut live time for each detector. The live times after the bad data cut are given in Table 8.2. The main reason for

Table 8.2: Live times after bad data cut, in live days.

Detector	T1 R118	T1 R119	T2 R119
Z1	38.196	74.473	73.331
Z2	34.725	73.714	71.965
Z3	38.898	74.473	73.674
Z4	36.354	74.429	74.473
Z5	38.678	74.473	73.953
Z6	36.262	73.957	74.114

differences between the live times of different detectors is that data sets sometimes fail the KS data quality checks for only one or a few detectors.

Kolmogorov-Smirnov tests

We use automated checks of data quality as an aid in constructing the list of good data sets. These checks compare the one- and two-dimensional distributions of various quantities in each series number to template data sets with known good performance, applying a Kolmogorov-Smirnov test to determine whether the distributions are consistent. In the original Run 119 analysis and the Run 118+Run 119 combined analysis, the KS tests have been used for low background data only, and the calibration data sets (of which there are fewer) are selected by hand. In subsequent runs, the KS tests have been applied to calibration data as well (with suitable templates), and review of the KS values has been organized as rotating data quality task, assigned in shifts of one or two weeks to CDMS member institutions with the provision that data quality shifts and on-site shifts in Soudan should not overlap. The KS tests have been developed by Long Duong, who also performed the initial round of WIMP-search data selection for this analysis.

Whenever a series number differs from the templates by a sufficiently significant degree ($p \leq 10^{-3}$), a warning flag is generated and human attention is required to determine whether or not the data set can be used. In many cases, failing KS tests are caused by shifts in detector performance that are statistically detectable, but do not impair data quality. For example, small changes in base temperature cause small shifts in phonon calibration and phonon energy resolution. In other cases, the

KS tests may fail because a phonon channel is not properly biased, or the charge neutralization is degraded.

8.3.2 Error mask cut

The error mask is a set of flags recorded by the DAQ to indicate any problems with the trigger or veto history buffer. A few events have bad values for the error mask, and are cut. Only events with error mask values indicating no problems (ErrorMask=0 or 128) are retained. One common type of problem is that some events are missing the global trigger, resulting in an error mask value of 160. These events also flag the global trigger cut, and are discussed in more detail in 8.3.3.

In the Run 118 WIMP-search data, the most common type of error flagged is that the veto history buffer overflows. This happens when there is a very long time between events. This does not happen in calibration data, when the event rate is high. Between Run 118 and Run 119, the history buffer was changed to a ring buffer, which does not suffer overflows. Therefore, the history buffer overflow events stopped after Run 118. A cut on ErrorMask was not imposed in the original Run 118 analysis, and these events were kept. However, it was imposed in the published Run 119 analysis, where very few events were affected. The combined analysis has uniformly imposed the error mask cut, thereby losing a fraction of the Run 118 exposure, but simplifying the analysis. The loss in exposure is 58,075 out of 1,003,860 events, or 5.79% of the Run 118 data.

8.3.3 Global trigger cut

A global trigger is the signal for the DAQ to start recording an event. Depending on the configuration of the trigger logic, phonon high and low triggers, ionization triggers, random triggers, and veto triggers all may or may not generate global triggers. Since the DAQ does not record an event until a global trigger is issued, every event should have a global trigger in its history buffer data. However, a few rare events are missing their global triggers because of an unknown DAQ error. This most often happens when a random trigger (used to sample the baseline noise for generating the optimal

Table 8.3: Events without global trigger.

	No global	All events	Fraction
Ba randoms	192	545,345	3.5×10^{-4}
Ba not randoms	21	9,443,568	2×10^{-6}
Cf randoms	7	19,650	3.6×10^{-4}
Cf not randoms	0	290,647	0
bg randoms	155	222,297	7.0×10^{-4}
bg not randoms	2	1,580,671	1×10^{-6}

filter) causes the DAQ to begin digitizing, but the history buffer fails to record the global trigger. This happens for about one in every 2,000 random triggers. More rarely, a detector-triggered event has no global trigger in the history buffer. This happens for about one in every 200,000 events.

The affected detector-trigger events are found only in a few data sets. Both affected WIMP-search events are in 131014_1817. The Ba data sets are 131009_1122 (20 events) and 140413_0856 (1 event). The first of these, 131009_1122, is from very early in Run 118; it is in fact the first "good" barium run of the entire combined data set. One other data set, 131211_0920, has 22 detector-triggered events missing the global trigger, but only during the period of time during which the nitrogen shield level meter was active. This portion of the run is already excluded from the good data set (see 7.3.1). In all of these cases, inspection of the raw data reveals that the DAQ did in fact record a global trigger, but that spurious times were also recorded in the history buffer, so that the trigger records could not be parsed correctly. Therefore, these events should be cut. The event from 140413_0856 is caused by a history buffer overflow. This has since been fixed, and is in any case already tagged and cut out by the ErrorMask cut. We can conclude that the global trigger is always present for detector-triggered events, but that very rare problems with the history buffer time records can hinder its identification. The fraction of detector-triggered events removed by the global trigger cut is very small, $< 10^{-5}$, and its effect on efficiency may be neglected.

The random-trigger events with missing global triggers are interspersed evenly throughout the running period, both in Run 118 and Run 119. Their global triggers

Table 8.4: Events with global trigger only.

	Trigger mode	Global only	All events	Fraction
131029_1628	Q-low	228,324	300,501	76%
131119_1110	Q-gated	22,907	32,847	70%
131205_1220	Q-gated	41,834	60,001	70%
All other Ba	P-low	75	4,813,903	2×10^{-5}

are missing in the raw data, and it appears that there remains a DAQ bug that drops the global trigger for about one in 2,000 random triggers. This is unrelated to the small number of detector-trigger events discussed above.

There are also some events that have a global trigger, but nothing else. A check through the raw data indicates that the specific trigger information was not written by the DAQ (it s not a Dark Pipe problem). These are almost entirely restricted to a few runs during Run 118 that were taken with non-standard trigger modes. Several barium calibration data sets in Run 118 used an ionization low trigger instead of the standard phonon low trigger. Some data sets also imposed a maximum value, so that an event is recorded only if its pulse height falls within a certain range: the "Q gated" trigger. Table 8.4 is taken only from "odd" Ba calibration data.

There remains a very small number of events in phonon-low triggered data sets with only a global trigger. Even in the absence of useful history buffer information, it is possible to check whether they are random trigger events: the DAQ software writes a special "event category" flag when it initiates a random trigger event. Inspection of this flag indicates that all 75 affected events are detector triggers. There is no evidence of any other problem in the global-trigger-only events; they seem to be missing only trigger history buffer information, and it is possible to analyze them as usual. The low-background data also has a negligibly small number of events with *apparently* only a global trigger. These events do in fact have other triggers in the history buffer, but at a slightly different time from the global trigger. These are good events in every way, and are included in the analysis.

8.3.4   Charge bias cut

Very rarely, the DAQ may record a different value of ionization voltage bias than the usual -3V for Ge and -4V for Si. The most common problem is a 0V reading for the last event (or last few events) of a data set. It seems that the DAQ may continue to digitize one or several extra events after it has begun to de-configure the detectors at the end of running. To eliminate these events, the charge bias cut demands that the absolute value of the inner-electrode bias be between 2.5V and 4.5V. This cut removes only one otherwise good event in the low-background data set: the last event (#140435) of 140325_1750, for T1 Z1 only.

The barium calibration data set also has a few events removed by this cut. The last event of 131010_1233, #360001, has zero bias for T1 Z1-Z5, and correct bias only for Z6. In addition, the last several events of 131010_1313 have zero bias on T1 Z1, as does the final event of 140720_1518. The data set 131029_1628 has an error code for the bias of T1 Z5 and Z6, but is already excluded from the good data list for these detectors because they were not configured at the time. No californium events are affected.

Although the cut is somewhat loose, allowing anything from 2.5V to 4.5V,all passing events have exactly QIbias=2.998V (Ge) or 3.999V (Si). These are slightly different from 3V and 4V only because of the finite numerical precision of the DAQ.

Since the charge bias cut removes at most 6 events out of the full data set, the loss of efficiency is minuscule and can be neglected.

8.3.5   Anti-glitch cuts

Occasional bursts of electronic noise from various sources can be picked up and cause detector triggers. These glitch events were especially common during Run 118, and much rarer in Run 119, after improvements to the grounding scheme. The glitch events do not resemble real detector pulses. The glitch events typically appear on several detectors at once, and cause more phonon triggers than ionization triggers. They are removed using a cut on the pattern of triggers characteristic to glitches: at least two detectors have phonon triggers, and the number of phonon triggers must be
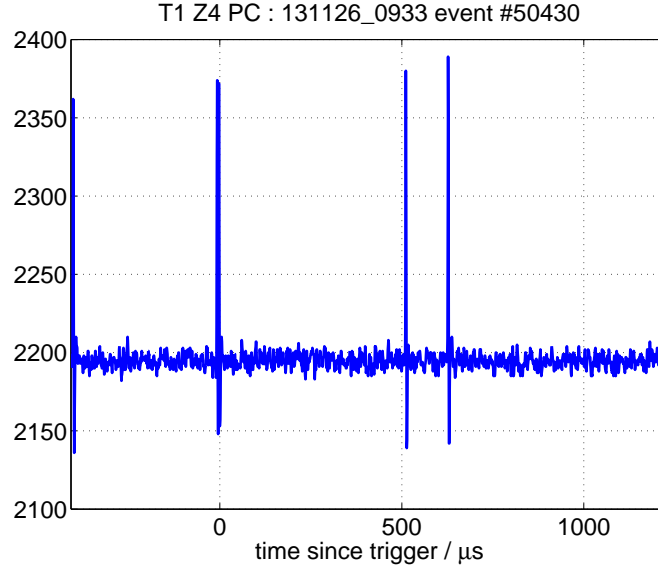
Figure 8.4: An event removed by the fall time-based anti-glitch cut. This is an event from the Run 118 low-background data set, from a run in which T1 Z1 channel C had a large number of noise triggers.

greater than the number of charge triggers by at least two.

Because this cut removes only events triggering multiple detectors, the efficiency for WIMP candidate events is 100%. The rate of accidental coincidences is already taken into account in the efficiency of the single-scatter cut. Although the anti-glitch cut has no direct effect on the acceptance fraction for WIMP candidate events, it does remove a large number of Run 118 glitch events that could otherwise skew the efficiency calculations for cuts such as the pre-pulse noise cut and the charge chi-squared cut. Therefore, the anti-glitch cut is applied before making all efficiency calculations in the following sections.

A second type of glitch event is found in bursts in certain phonon channels. Because these are not cut out by the cut above, a second anti-glitch cut is needed. These events take the form of very quick spikes, with rise and fall times much faster than typical events, as shown in Fig. 8.4. Their fall times are $< 20\mu s$, while the minimum possible fall time for a real pulse is the electrothermal feedback time (23-150 μs, typically about 50 μs). In practice, real pulses have considerably slower fall times, because

a significant amount of phonon energy continues to arrive at late times. Therefore, a cut at 20μs is effective for removing spike-like glitches, while preserving all real pulses. The efficiency of the cut is calculated by applying it to neutron events; it is found to be 99.95±0.01%, close enough to 100% to represent a negligible loss in efficiency.

8.3.6   Phonon pre-pulse noise cut

Each event is digitized back to 400 μs before the trigger. The pre-trigger part of the trace is used to evaluate the baseline noise levels and stability of the phonon channels during the event, and to reject pile-up. In particular, the standard deviation of the pre-pulse samples provides a way to cut out events in which the baseline is changing or too noisy. This cut requires that each phonon channel have a suitably low pre-pulse standard deviation, with the cut level set independently for each channel, and sometimes adjusted for data sets with different noise levels.

The cut is set at at $\pm 5\sigma$ of the distribution of random trigger events (excluding those in which real pulses or noise bursts were coincidentally found in the digitized trace). In order to retain as much WIMP-search exposure as possible, the cut is specially tuned for some of the WIMP-search data sets in which the pre-pulse noise values drifted from the usual mean or standard deviation. There is no special tuning for any calibration data sets; the overall values from the low-background data set are used. The standard values are given in Table 8.5.

The pre-pulse standard deviation is closely related to the noise spectrum. Under the assumption of stationary noise, the standard deviation may be found by adding in quadrature the contribution from each frequency. If i_j is the current noise at the input coil in amps per root-Hertz at frequency f_j, Δf is the frequency interval (digitization rate divided by number of samples), and g is the size of a digitizer bin in amps, then $\sigma^2 = \frac{\Delta f}{g^2} \sum_{f_j \neq 0} i_j^2$. Anything that changes the noise spectrum will also change the distribution of σ: the Johnson noise of the shunt and parasitic resistances are sensitive to the temperature of the 600 mK and 4K stages, respectively. These may be elevated by cryogen transfers, and take some time to recover. Accordingly, some channels have a gradual shift in phonon pre-pulse standard deviation over the

Table 8.5: Cut ranges for phonon pre-pulse noise.

Detector	PA	PB	PC	PC
T1 Z1	3.410-8.010	2.425-8.015	2.235-6.105	2.450-5.690
T1 Z2	1.680-3.740	1.670-3.770	1.880-4.360	2.220-4.500
T1 Z3	1.820-4.260	2.215-5.525	2.615-5.565	1.750-3.910
T1 Z4	2.455-5.965	2.500-5.380	2.225-5.675	1.650-3.450
T1 Z5	3.040-5.980	2.990-6.390	3.460-7.100	2.845-7.655
T1 Z6	2.525-5.675	2.615-5.384	2.405-5.295	2.605-5.895
T2 Z1	1.535-3.245	2.120-4.100	2.380-4.700	1.660-3.180
T2 Z2	1.360-2.480	1.615-3.305	1.490-3.070	1.425-2.935
T2 Z3	2.675-4.845	2.290-3.850	2.180-3.880	2.540-5.020
T2 Z4	1.415-2.485	1.400-2.340	1.620-2.740	1.345-2.315
T2 Z5	1.400-2.700	1.510-2.770	1.450-2.770	1.495-2.665
T2 Z6	1.425-2.515	1.680-2.940	1.755-3.085	1.870-3.150
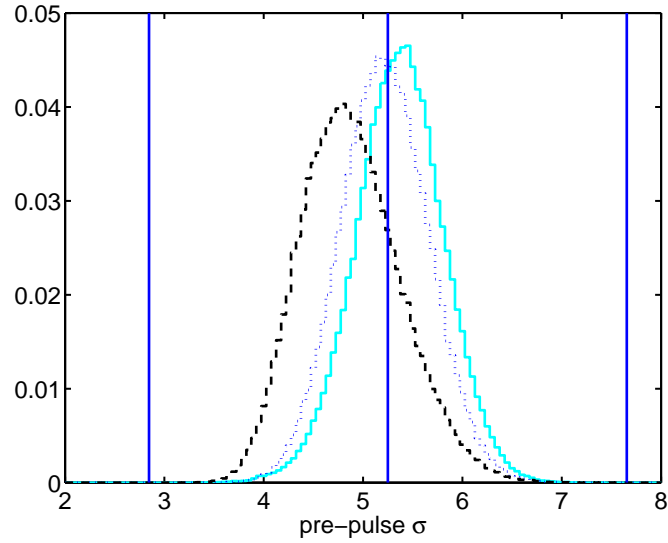
Figure 8.5: Drift in phonon pre-pulse noise with time since run start. This effect is largest in T1 Z5, channel D, the channel shown. The event samples are the initial 500 random trigger events (light blue, solid); 0.1-1.1 hours (dotted, blue); and 5-6 hours (dashed, black). The 5σ cut values and nominal mean are shown as solid vertical lines.
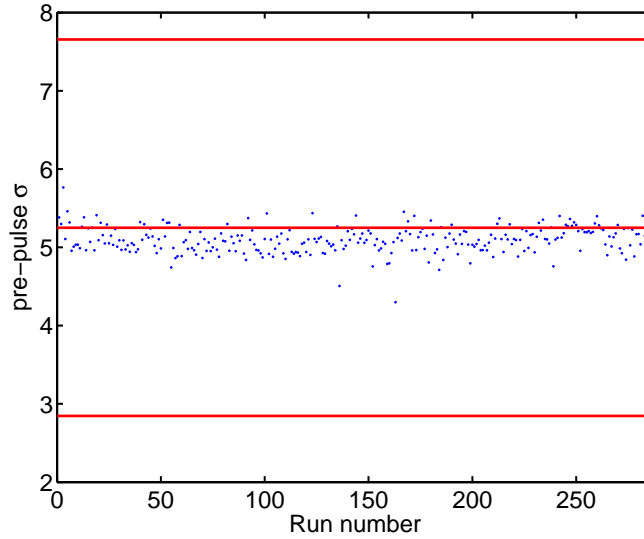
Figure 8.6: Stability of pre-pulse noise over the entire R118 + R119 combined WIMP-search data set. The horizontal axis is run number within the full data set; the vertical axis is pre-pulse standard deviation in T1 Z5 channel D. Two runs near 150 have lower pre-pulse noise because changes in SQUID behavior have reduced their bandwidth. These runs have special values of the phonon pre-pulse noise cut.

course of a run. This leads to small differences between the nominal values used to set the cut and the actual distribution of pre-pulse standard deviation in a given data set (See Fig. 8.5). The small shifts lead to a loss of efficiency relative to the nominal two-sided 5σ cut efficiency of $1 - 6 \times 10^{-7}$. Fig. 8.6 shows that stability over long time periods is generally good, but some data sets have significant shifts in bandwidth because of changes in the anomalous SQUID behavior near rolloff. This cut applies specially tailored values for low-background data sets with such shifts.

Since this cut evaluates only the part of the digitized trace before the beginning of the signal pulse, it is called a "pre-trigger" cut. Pre-trigger cuts have efficiencies uncorrelated with any properties of the event itself, such as energy, position, or recoil type. The overall efficiency on WIMP-search data is given in Table 8.6. Both of the anti-glitch cuts are applied before calculating these efficiencies.

Table 8.6: Events lost in pre-pulse phonon noise cut.

Detector	T1	T2
Z1	0.7%	0.02%
Z2	0.02%	0.8%
Z3	0.009%	0.05%
Z4	0.08%	0.04%
Z5	0.02%	0.05%
Z6	2.0%	0.03%

Table 8.7: Loss in efficiency from charge pre-pulse cut.

Detector	T1 R118	T1 R119	T2 R119
Z1	7.5%	-	-
Z2	0.018%	-	-
Z3	-	-	-
Z4	0.003%	-	0.16%
Z5	-	-	-
Z6	-	-	0.33%

8.3.7 Charge pre-pulse noise cut

The ionization pre-pulse noise cut is similar to the phonon pre-pulse noise cut. The cut is made on the standard deviation of the baseline samples before the beginning of the pulse, with a cut level set by hand along with the charge chi-squared cut. It removes very few events, since the event analysis is relatively insensitive to high-frequency noise. It does remove a significant number of events for T1 Z1 during Run 118. This detector's ionization channels were unusually sensitive to noise pickup before the grounding corrections in early 2004. The loss in live time for T1 Z1 is about 7%, but this is irrelevant since the sole affected detector is already excluded from the WIMP search. If a future analysis were to reinstate T1 Z1, this cut could be safely relaxed.

The efficiency of the cut is estimated from random trigger events in the WIMP-search data. The effect on efficiency is given in Table 8.7. Where the fraction of events lost is 10^{-5} or smaller, it is simply neglected.

8.3.8   Charge chi-squared cut

The chi-squared cut on the ionization channels rejects events for which the observed pulses differ too much from template pulses. This rejects events caused by electronic glitches rather than particle interactions, pile-up events, traces with shifting baselines, and some events with excess noise. The "chi-squared" value is not normalized as a true chi-squared, and is energy dependent. Therefore, the cut is implemented as a quadratic in energy, as shown in Fig. 8.7b. For Run 118, the charge chi-squared cut removes events during which the sliding seal heater was active. These form a second, higher band of chi-squared values, above the cut (See Sec. 7.1). Events taken while the heater was on are believed to be usable for the WIMP search, but are excluded in order to simplify the chi-squared cut. This represents an acceptably small loss of live time.

The efficiency of the cut is calculated from gamma events in the low-background data. Run 118 and Run 119 must be handled separately because the main source of events rejected by the cut is the sliding-seal heater effect, which only occurred in Run 118. There is also an energy dependence, which is fit using an error function in ionization energy. The efficiency calculation for T1 Z5 is shown in Fig. 8.7b.

8.3.9   Negative phonon cut

The negative phonon cut is present primarily for safety; there should be no way to generate significant negative-going pulses in the phonon channels. The cut is set at N standard deviations below the noise blob, and is imposed on each channel independently. For this analysis, N=6 and the efficiency for real events is taken to be 100%.

8.3.10   Phonon saturation cut

This cut demands that no phonon channel have any saturating samples. Because saturation happens only at very high energies > 1 MeV, saturating events are well outside the usual region of interest. The phonon saturation cut is used simply to guard against any pathologies in fitting the energies of saturating pulses, which could

(a)

(b)
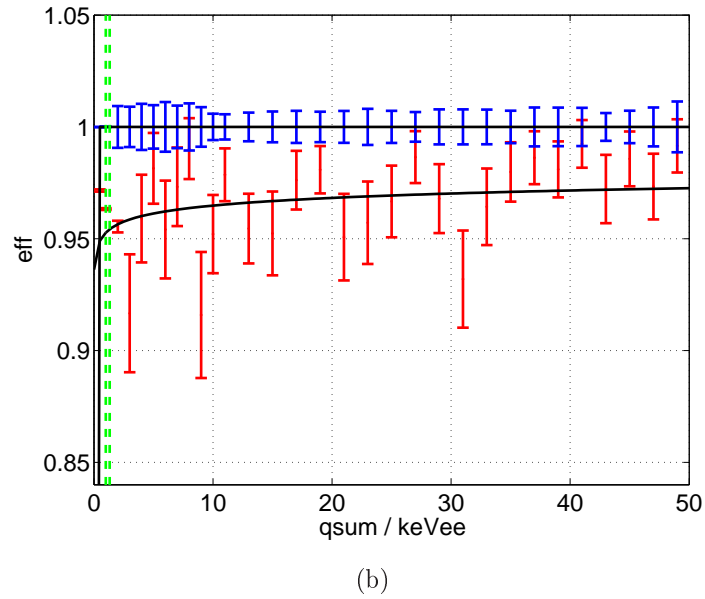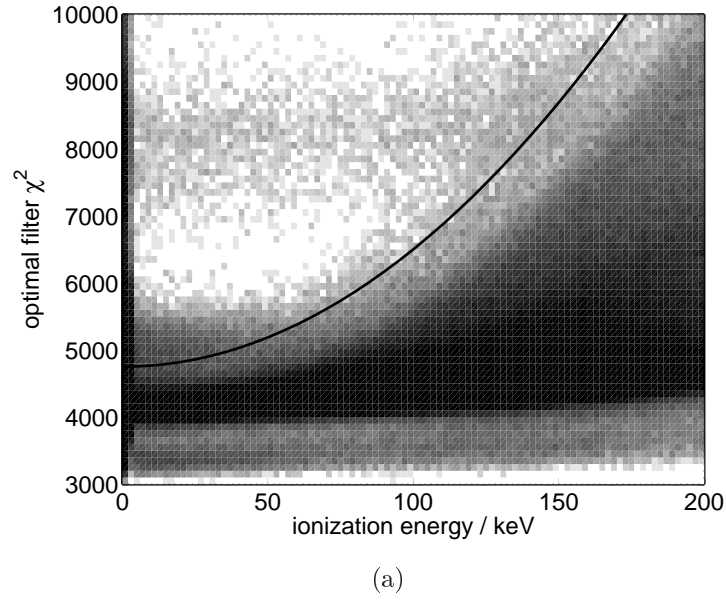
Figure 8.7: In (a), the ionization channel chi-squared cut in T1 Z5. The data shown are barium calibration events. The events along the left edge of the plot are electronic glitches, which can make large signals but don't resemble real pulses. The faint second band at χ^2 values around 8000 comes from the activity of the sliding seal heater. In (b), the efficiency of the chi-squared cut in T1 Z5. The upper (blue) error bars are from Run 119, with nearly 100% efficiency. The lower (red) error bars are from Run 118, with efficiency reduced by a few percent because of the sliding seal heater noise.

conceivably make them appear at low energies. The efficiency for 10-100 keV WIMP candidates is 100%.

8.4   Single-detector cut

WIMPs are expected to have a scattering length of light years, so any WIMPs passing through the CDMS icebox should scatter in at most one ZIP detector. The single-scatter cut rejects events that have energy significantly above the baseline noise level in more than one detector. The efficiency for WIMP candidates is very high, with the possibility of rejection only in the event that the noise rises above the threshold, or for accidental coincidences.

The cut is defined as a threshold on total phonon energy pt in each detector. For an event to be selected as a single scatter in a given detector, its phonon energy must exceed the mean of the random-triggered noise traces by at least 6 standard deviations in that detector; and the phonon energies in all other detectors must be within 3.5 standard deviations of the noise mean. An event can fail by having pt more than 3.5σ above *or below* the noise. The cut threshold for one detector, using WIMP-search random trigger events, is shown in Fig. 8.8b.

The efficiency of the singles cut is calculated from the random-triggered noise traces. All events in the energy range of interest pass the first part of the cut, since any event above the 5-keV phonon analysis threshold is automatically more than 6 standard deviations above the noise blob. For each detector, the efficiency of the singles cut is the fraction of noise traces for which no other detector has pt outside the 3.5σ acceptance window. Because the random triggers are taken from low-background data, this procedure accounts both for events that fail the cut because of noise excursions, and events that fail the cut because of accidental coincidences. The random trigger events are required to pass the good data selection cut for each detector, and to have energy in the detector of interest consistent with noise. The results of this calculation are tabulated in Table 8.8. Although there is a small variation by detector, the efficiency is taken to be a constant 98.9%, averaged over all twelve ZIPs.

Table 8.8: Singles cut efficiency calculation.

T1	Randoms	Fail cut	Eff / %		T2	Randoms	Fail cut	Eff / %
Z1	256836	2654	98.967		Z1	163474	1839	98.875
Z2	248282	2547	98.974		Z2	160470	1430	99.109
Z3	258610	2662	98.971		Z3	163695	2053	98.746
Z4	253933	2610	98.972		Z4	165966	2179	98.687
Z5	258255	2664	98.968		Z5	164609	2173	98.680
Z6	252171	2514	99.003		Z6	165138	2131	98.710

In the original Run 119 analysis, the singles threshold was 6σ instead of 3.5σ, resulting in a slightly higher 99.8% nominal efficiency. However, an error in a correction to phonon energies in T2 Z1 caused a 17% loss in real efficiency, documented in Appendix A.

8.5 Fiducial inner volume cut

The ZIP detector's charge sensor is segmented, with a disk-shaped inner electrode and a ring-shaped outer electrode, in order to define an inner fiducial volume. Near the outer wall of the detector, several types of problems can reduce charge collection so that gamma events fall at low yield. In some areas the field lines intersect the outer edge of the crystal, so that charges do not drift all the way to the electrodes or ground. The continuity of the electrode and grid and the uniformity of the fields are not as good in the outer region as the inner region. In the first two towers of detectors, the grid on the phonon side outside the phonon sensors is floated instead of grounded, leading to several types of anomalous behavior and reduced yield (see Chapter 5). Therefore, only events contained within the inner electrode area are retained for the WIMP search analysis.

Several types of fiducial volume cut have been used in the past. The simplest is a partition-based cut on the quantity $qpart = \frac{QI-QO}{QI+QO}$. For events fully in the inner electrode, $qpart = +1$, and for events fully contained in the outer electrode, $qpart = -1$. Inner electrode events can be selected with the broad cut $0.8 \leq qpart \leq 1.2$, which was used in the analysis of the -3V bias data from SUF.

(a)   (b)
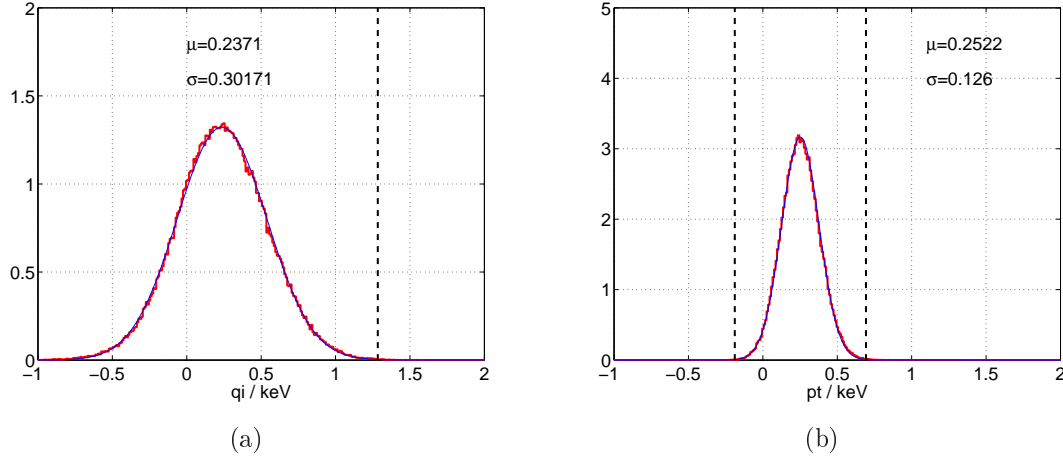
Figure 8.8: In (a), the ionization threshold cut for T1 Z3 in Run 119. The noise blob is formed from random trigger events in the WIMP search data set. The distribution of pulse heights is Gaussian, and the threshold cut is set at 3.47σ above the mean in QI. In (b), the singles cut thresholds for T1 Z3. The noise blobs are again used, with the cut set at 3.5σ above and below the mean.

A second type of fiducial volume cut is defined by demanding that the signal in the outer electrode be consistent with noise. Instead of cutting on a ratio of inner and outer energies, this imposes $|QO| \leq n\sigma_{QO}$, where σ_{QO} is the root-mean-square energy of noise events in the outer channel and we use $n = 2$. In practice the cut is somewhat more complex, since the population of noise events is not centered at QO=0, and because the cut is often made somewhat looser at higher energies. This style of cut was used in the analysis of the -6V bias data from SUF and the published analyses of Soudan Run 118 and Run 119.

Each of these cuts has some advantages and disadvantages. The qpart-based cut requires an energy-dependent fraction of the total energy in the inner electrode, so that in the absence of noise the physical volume of the cut should be fixed. In the presence of noise, however, this cut has an efficiency that falls steeply at low energies, as the noise becomes large enough to push most events outside of the cut boundaries. The noise-based cut has an acceptance fraction for good inner-volume events that is approximately constant with energy. Its weakness is that at low energies, it allows in more and more events that are not truly contained in the inner electrode.
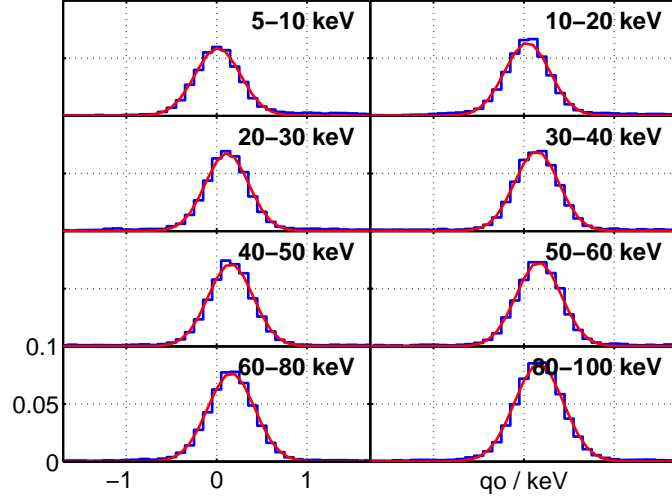
(a)
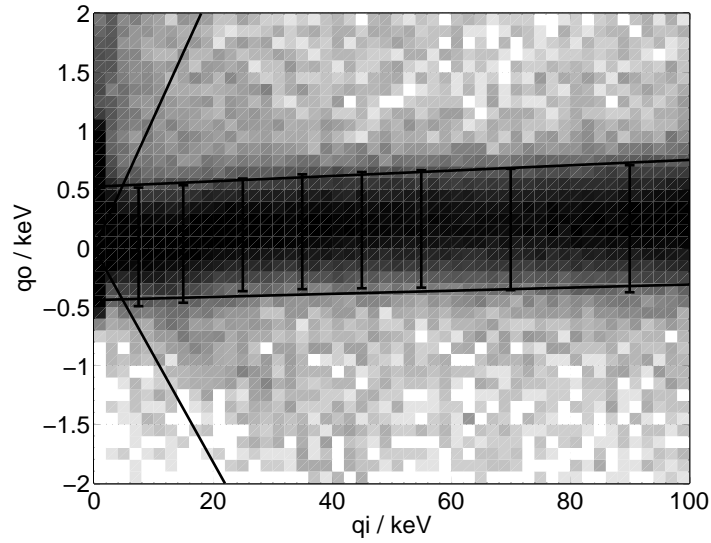
(b)

Figure 8.9: Setting the inner-electrode fiducial volume cut. In (a), histograms of the QO pulse height for events fully contained in QI, with Gaussian fits. In (b), color represents the density of points. The vertical error bars show the 2-σ Gaussian widths by energy bin, and the near-horizontal lines are the linear cut boundaries. The diagonal lines represent a partition-based cut imposed for beta selection only, at low energies. Both plots are for T1 Z5 in Run 119.

At low energies, the noise-based cut can also bias the yield towards lower values. If the yield is defined using the inner electrode signal only, $yic = \frac{qi}{pric}$, then events with some energy sharing between the two electrodes will have lower yic than $yc = \frac{qsum}{prc}$. The noise-based cut allows in such events at low energies, causing some gammas to leak into the beta region, and possibly the nuclear recoil region.

It is necessary to balance the advantages and disadvantages of these two types of inner electrode cut. For selecting WIMP candidates, the goal is to maximize exposure while allowing an acceptably low level of leakage. It is especially important to maintain high acceptance at low energies, where the WIMP recoil spectrum is expected to be strongest. The noise-based cut gives significantly higher efficiency for nuclear recoils at low energies, without many high-radius gammas that might leak into the nuclear recoil band. Therefore, it is acceptable for WIMP candidate selection in the low-background data.

On the other hand, the barium calibration data has many more gamma recoil events. If the noise-based inner electrode cut is used, a significant number of these may leak into the beta region, contaminating the surface event population. This sample of surface scatters is used to set the timing cut, and even a small number of bulk gamma events leaking in from high radius may poison the timing cut, forcing it to be very strict and resulting in a very low WIMP acceptance. Therefore, a stricter cut is needed for the selection of betas in the calibration data. For this purpose, the intersection of the partition-based and noise-based cuts is used: events must have outer electrode energy within two standard deviations of the noise, *and* have partition between 0.8 and 1.2 (see Sec. 9.4.1). The noise-based cut has effect at higher energies, and the partition-based cut at lower energies.

The cut-setting procedure, along with both 2-σ and partition-based cuts, is shown in Fig. 8.9.

Based on the geometry of the inner and outer electrodes, the fiducial volume cut should have an efficiency of about 85%. This assumes that there is a sharp boundary between inner and outer events, running right down the middle of the trench between the two electrodes. However, the reality is more complicated. The electric field lines are not exactly vertical, so that the inner region is not a perfect cylinder. The charges
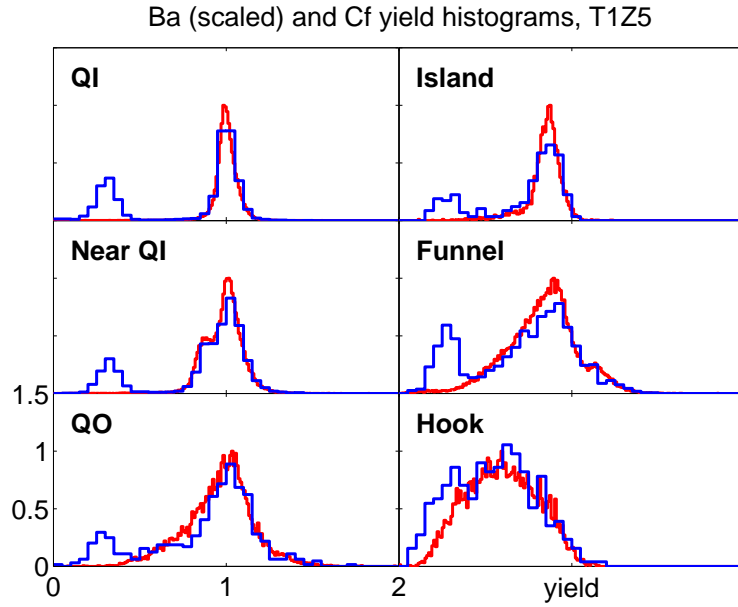
Figure 8.10: Scaling barium to californium gammas for nuclear recoil band calculation in T1 Z5. The barium data (finer binning, red) includes only electron recoil events, while the californium data (coarser binning, blue) includes both gammas near yield of 1 and neutrons at yield of about 1/3. The barium gammas in each region have been scaled to match the californium gammas, so that the low-yield portion can be subtracted.

spread out into a ball with finite size, so that events sufficiently close to the boundary have ionization shared between the two electrodes. The "holiday island" effect also changes the partition of charge. Furthermore, for particles that may multiply scatter, such as gammas and neutrons, the event will fail the inner electrode cut if any of the scatters deposits significant energy into the outer region.

We calculate the efficiency from neutrons in the californium data set. This gives a conservative estimate, since it includes some multiply-scattering neutrons (which are more likely to fail the cut). The fraction of neutron events that include multiple interactions is estimated to be ~1/3. The calculation is complicated by the existence of low-yield gammas, which may leak into the neutron sample. In particular, the "funnel" and "ear" populations discussed in Chapter 5 have suppressed yield and fail the inner electrode cut, giving an unreasonably low efficiency estimate. These gammas must be accounted for. It's possible to cut them out using a timing cut, but this is unreliable for several reasons. First, the timing cuts are tuned to distinguish neutrons from surface scatters, rather than bulk electron recoils, and do not always have strong rejection for bulk gammas. Second, the timing cuts are intended for use only with events inside the fiducial volume, and do not account for changes in timing quantities in the outer region.

Therefore, we choose to statistically subtract the low-yield gammas, using the barium data set as a reference sample for the distribution of yield and charge partition. We define samples of neutrons and gammas in six energy bins and six detector regions, scale the gammas from the barium sample to match the event rate in the californium data, and subtract the scaled Ba from the Cf data in each bin to get the corrected neutron counts.

Neutrons are selected using a special band cut on the total yield $yc = \frac{qsum}{prc}$ (unlike the ordinary band cut, which is valid only for inner events and uses the inner yield yic). The cut used here goes to three sigma above the nuclear recoil band (a wide yield cut), and four sigma below the nuclear recoil band. This asymmetric cut was chosen because neutrons in the outer regions of the detector may tend to fall below the main yield band, but in no region are they likely to fall much above it. Extending the cut further to high yields would allow more leakage of electron recoils, and is not
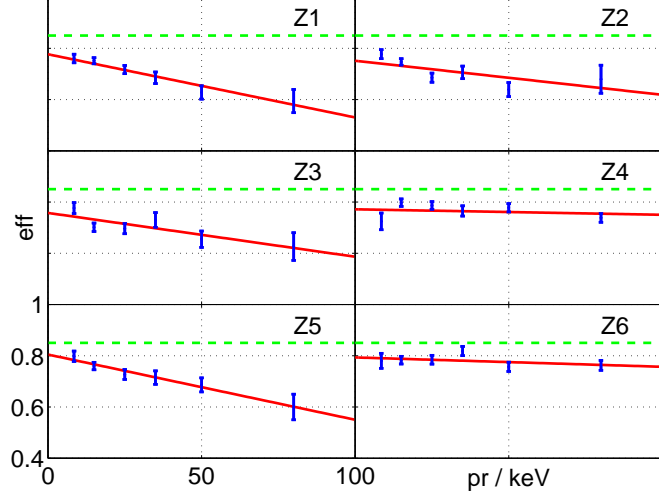
Figure 8.11: Efficiency of fiducial volume cut in Tower 1 detectors. The dashed (green) line shows the nominal 85% passage fraction, and the solid (red) line shows a linear fit to the measured efficiency.

be necessary to get a good sample of neutrons.

The six energy bins used are 7-10, 10-20, 20-30, 30-40, 40-60, and 60-100 keV in phonon recoil energy. The six regions are the "ear" (-2.5≤$qpart$<-1.1), "outer" (-1.1≤$qpart$<-0.8), "funnel" (-0.8≤$qpart$<0), "near QI" (0≤$qpart$<1.05, but failing the fiducial volume cut), "inner" (passing the fiducial volume cut), and "holiday is-land" (1.05≤$qpart$<2). In each of these bins, we normalize the spectrum of barium-calibration gammas to match the number of events in the gamma band in the cali-fornium data. This is shown in Fig. 8.10, summed over the energy bins. The number of gammas leaking into the wide nuclear recoil band is then subtracted from the raw neutron count. In calculating the efficiency, the numerator is the total corrected neu-tron count in the QI region, and the denominator is the total corrected neutron count from all regions. The statistical error comes from Poisson statistics on the number of events in the nuclear recoil band in each bin in Ba and Cf.

The resulting efficiencies are given in Fig. 8.11 and Fig. 8.12. They are generically lower than the nominal 85%, and decrease with energy. This decrease is attributed
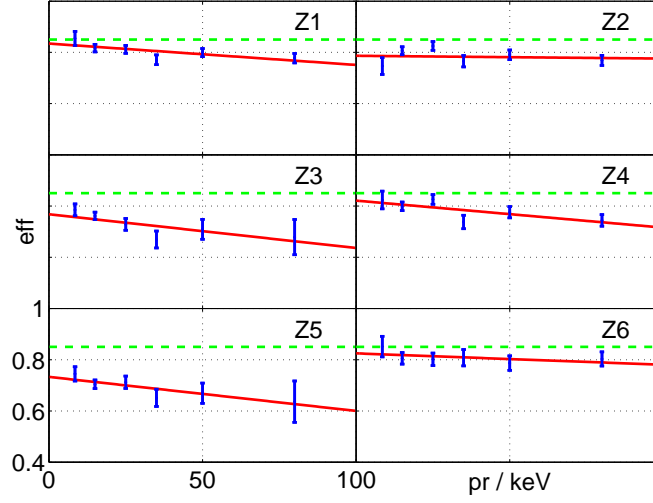
Figure 8.12: Efficiency of fiducial volume cut in Tower 2 detectors. The dashed (green) line shows the nominal 85% passage fraction, and the solid (red) line shows a linear fit to the measured efficiency.

to the increased probability of multiple scatters at higher energy. It is slight in Si detectors, and more pronounced in Ge. We fit a straight line to the efficiency as a function of energy for each detector, and include this linear functional fit in the overall efficiency calculation.

8.6   Ionization threshold

When the signal-to-noise of the ionization channel becomes too low, electron recoil and nuclear recoil events can no longer be reliably distinguished, and the fiducial volume cut becomes compromised. The ionization threshold cut is used to select only events with ionization energy sufficiently far out of the noise to give good discrimination. This cut is set at a fixed number standard deviations above the mean of the population of WIMP-search noise events, with the exact number chosen to give <0.1 noise events leaking above the cut in the WIMP-search data set. This yields a cut at 3.50σ for Run 118 and 3.48σ for Run 119. The case for one detector is shown in Fig. 8.8a.

The efficiency of the ionization threshold cut is calculated together with the below-gamma-band cut in Sec. 8.7.

8.7   Below gamma band cut

Besides the standard nuclear recoil band cut, we also impose a cut to reject events that are too close to the gamma band. This is necessary because at low energies the band merge, so that the tails of the gamma distribution may begin to be accepted as neutrons. We impose this cut at 3.5σ from the mean of the gamma band, in order to ensure that all detectors have less than 0.02 events expected leakage of bulk gammas into the nuclear recoil band before timing cuts.

The efficiency of this cut is calculated together with the ionization threshold cut. These two cuts carve into the nuclear recoil band at low energy, coming from opposite directions. The efficiency is the fraction of nuclear recoils within the 2σ band that are not cut out by the ionization threshold cut or the low-yield cut. This is calculated algebraically, under an assumption of Gaussianity, rather than from data. Except at very low energies, the combined efficiency is 100%. An error function is fit to the combined efficiency for each detector. Because the ionization threshold for Tower 1 is set separately for the Run 118 and Run 119 segments of the data set, the efficiencies are done separately as well. Figure 8.13 shows the acceptance region and efficiency function for T1 in Run 119. The calculation for T1 in Run 118 and for T2 in Run 119 are very similar and are not shown here.

8.8   Nuclear recoil band

The primary background discrimination of CDMS comes from *ionization yield*, the ratio of ionization to phonons. The ionization signal is calibrated such that electron recoil events have yield of 1. Nuclear recoil events have ionization suppressed according to Lindhard theory, and have yield $\sim \frac{1}{3}$. The electron recoil band is constructed from gamma events from ^{133}Ba data, while the nuclear recoil band is constructed from neutron events from ^{252}Cf data.
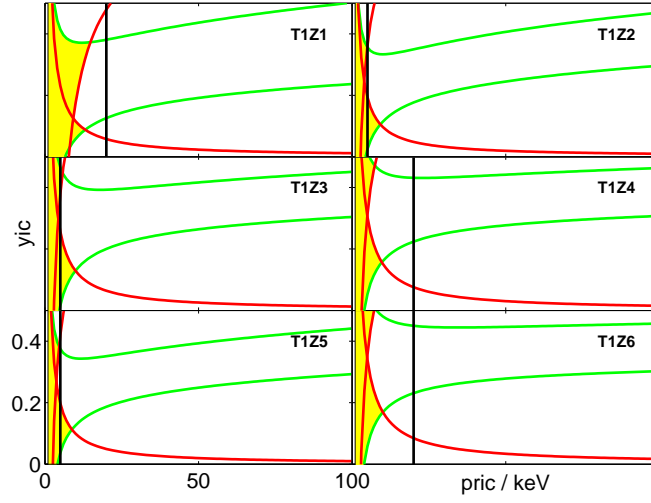
(a)
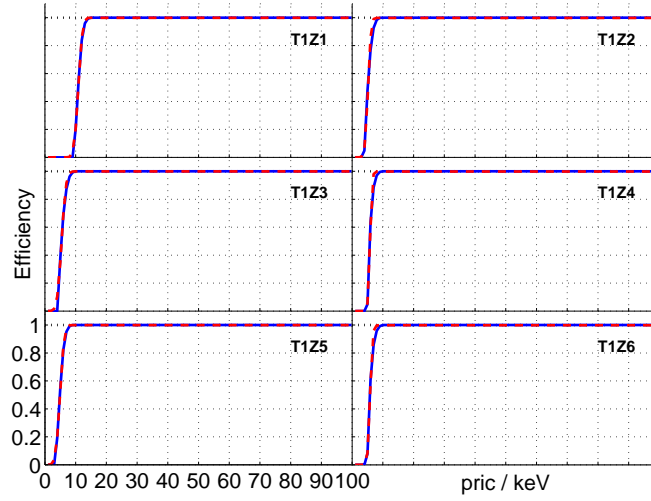
(b)

Figure 8.13: Acceptance region and efficiency calculation for ionization threshold cut and below-gamma-band cut for Tower 1 detectors in Run 119. In (a), the acceptance regions are shown. The shaded region is the part of the nuclear recoil band excluded by the cuts. The nuclear recoil band is in green, the ionization threshold cut is a hyperbola in red, and the below-er cut is a red curve sloping upward. The vertical lines indicate the phonon analysis threshold. In (b), the combined efficiencies (blue) and fits to an error function (red, dashed).
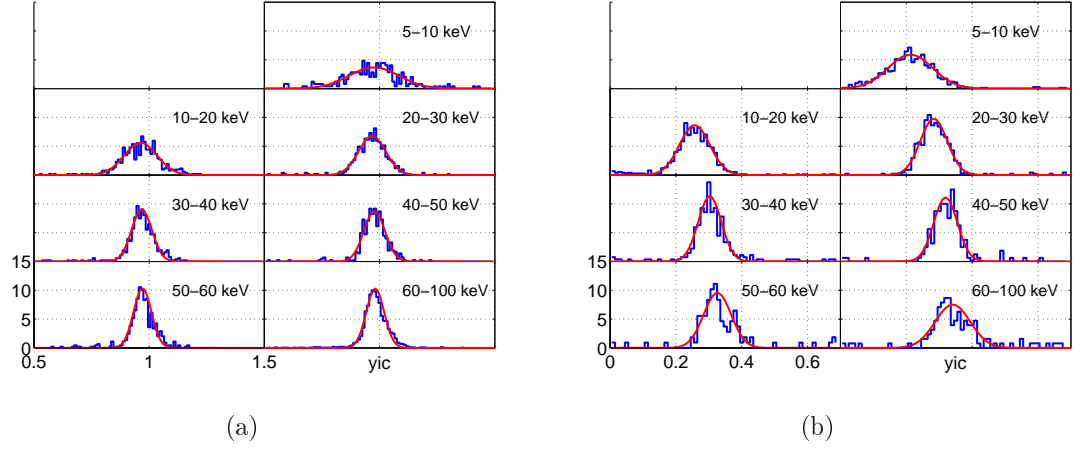
Figure 8.14: Gaussian fits for yield bands of barium electron recoils (a) and californium nuclear recoils (b) in T1 Z5.
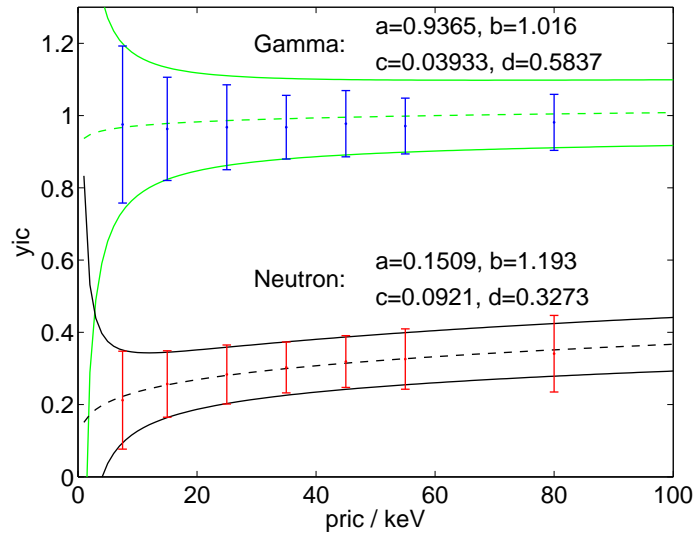
Figure 8.15: Functional form fits for electron recoil and nuclear recoil bands for T1 Z5. The error bars and band widths are shown at 2σ.
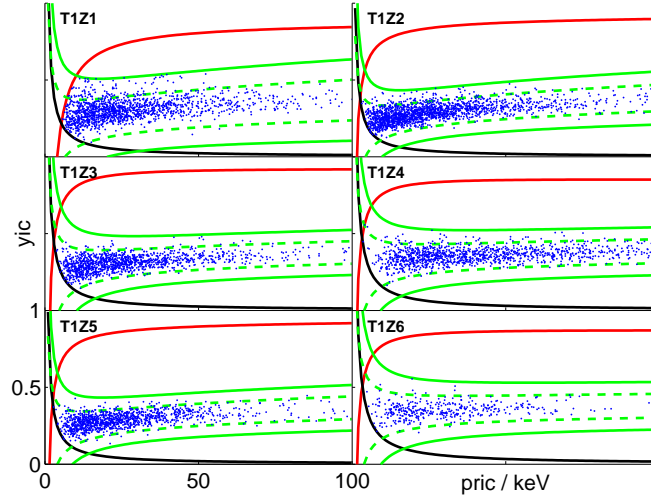
Figure 8.16: Selection of events for nuclear recoil band efficiency calculation in Tower 1.

The nuclear recoil band cut is defined in terms of the corrected, inner-electrode recoil energy quantity $pric = ptc - \lambda qi$, where ptc is the position-corrected and linearized total phonon energy, qi is the energy in the inner electrode, and λ is the Luke phonon factor. Because this is defined in terms of the inner-electrode ionization qi instead of the total ionization energy qsum, it is correct only for inner-electrode events. The recoil energy $prc = ptc - \lambda qsum$ could be used instead, but it has the additional statistical noise of qo. Therefore, pric is preferred.

The band cut accepts events within two standard deviations of the energy-dependent fitted mean yield for neutrons. The standard procedure selects neutrons in several energy bins in phonon recoil energy: 5-10, 10-20, 20-30, 30-40, 40-50, 50-60, and 60-100 keV. A Gaussian is fit to the distribution in each bin, as shown in Fig. 8.14. Finally, one functional form is fit to the Gaussian means as a function of energy, and a second functional form is fit to the standard deviations. The functional forms used
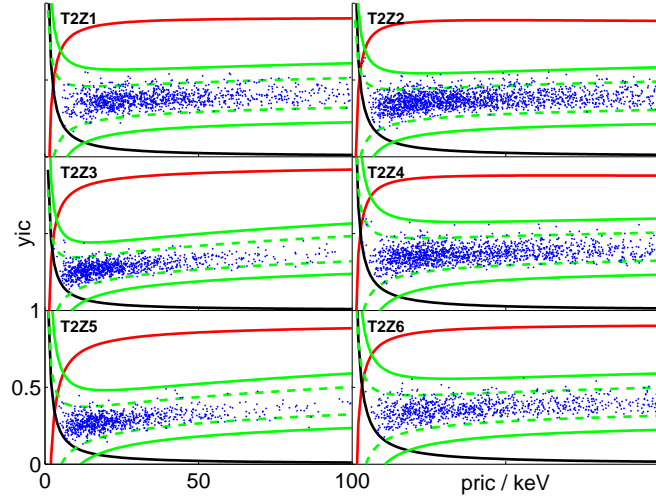
Figure 8.17: Selection of events for nuclear recoil band efficiency calculation in Tower 2.

are based on Lindhard, but with adjustable parameters:

$$
\begin{aligned}
\mu_y(E_r) &= a \times E_r^{b-1} \\
\sigma_y(E_r) &= c \times \mu_y + d/E_r.
\end{aligned}
$$

The term in d represents the energy-independent baseline noise in the ionization channels, and the term in c represents noise that scales with energy (coming from position dependence and other sources other than baseline noise). The same procedure and functional forms are used for neutrons and gammas. The resulting bands, along with the Gaussian error bars from each energy bin, are shown in Fig. 8.15.

The efficiency of the band cut is calculated using californium calibration neutrons. They are selected using a wider, four-sigma band, and required to pass a simple timing cut. The neutrons are also required to be at least three standard deviations away from the gamma band mean, a requirement which has effect only at the lowest energies. The efficiency is then the fraction of selected neutrons that also pass the two-sigma yield band cut. This is calculated within energy bins, and then the efficiency as a
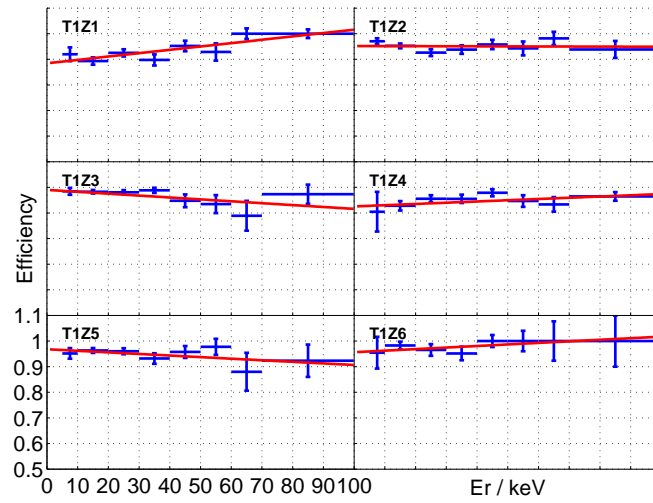
Figure 8.18: Nuclear recoil band cut efficiencies, with linear fits, in Tower 1.
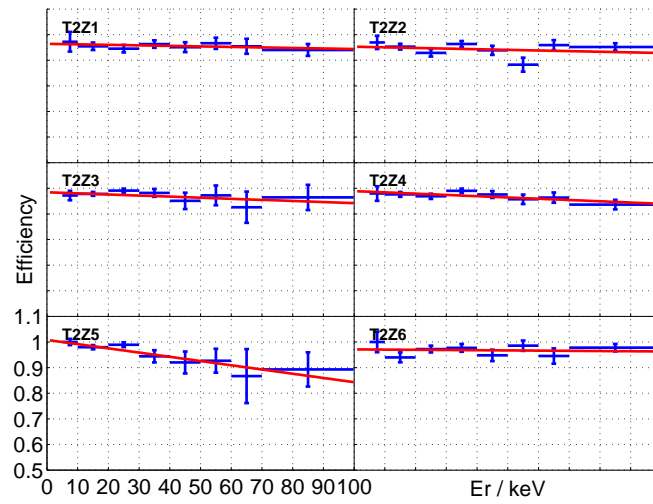
Figure 8.19: Nuclear recoil band cut efficiencies, with linear fits, in Tower 2.

function of energy is fit to a straight line with a slope. The selection of neutrons within the four-sigma band is shown in Figs. 8.16 and 8.17, and the resulting efficiencies, along with linear fits, are shown in Figs. 8.18 and 8.19.

8.9 Detector selection

Three detectors are excluded from the WIMP-search analysis. The germanium detector T1 Z1 has a strong T_c gradient and was ion-implanted using an older scheme, before the typical geometry of the position dependence was characterized. This resulted in the phonon sensors being active near the center of the detector, but largely dead in the outer regions. It also suffers from gaps in the outer electrode (Sec. 5.4), so that outer events with impaired charge collection may not be fully rejected. Finally, in its position at the top of Tower 1, surface events on this detector can only be vetoed by multiple scatter on one side. (Indeed, it was placed in this position largely to serve as a veto for other detectors, since its relatively poor performance was already well known at the time of tower assembly).

Two silicon detectors are also excluded. T1 Z6 is similarly in an end position within a tower, and is known to be contaminated with beta-emitting ^{14}C as a result of indirect exposure to a leaking calibration source during a test run. The high rate of surface scatters on this detector, along with the ability to veto them only on one side, make it unsuitable for use as a WIMP search detector. As these issues were known at the time of tower assembly, it was deliberately placed in the end position to serve as a veto, like T1 Z1. The final excluded detector is the Si ZIP T2 Z1. Although its performance is good in many respects, it has a low T_c and a significant gradient, with lowest T_c in the $-y$ region. This causes the phonon sensors at the bottom of quadrants B and C to be largely inoperative.

All of the excluded detectors are carried through all of the cut definitions and the full analysis, but their exposure is not included in calculating exclusion limits, and any candidate events in them are ignored. The decision to exclude these detectors was made in a blinded fashion.

8.10   Phonon analysis threshold

Because discrimination against gamma and beta backgrounds becomes more difficult at low energies, the analysis is cut off at a minimum phonon recoil energy. This is the analysis threshold. The value of the threshold is typically well above the minimum observable energy observable; it does not reflect the ultimate capability of the detectors, but rather the minimum energy for which a particular analysis has been performed, and WIMP acceptance and background rejection are well understood. A typical analysis threshold has been 10 keV recoil; this was used for the initial Run 118 analysis and the official blinded analysis of Run 119. The Run 118 analysis in the doctoral thesis of Clarence Chang used a 5-keV threshold. The initial publication on Run 119 also included a 7-keV analysis, which was prepared under the same blinding scheme; however, the decision to present the 7-keV result was made after unblinding. The Stanford shallow-site Run 21 used 5 keV in the first published analyses, but subsequent work by Raymond Bunker has produced an analysis (now nearing completion) down to 1 keV recoil energy! SUF Run 19 with BLIP detectors also used a 10 keV threshold. The target energy threshold for the Run 118 + Run 119 combined analysis is 5 keV recoil.

The 7-keV low energy threshold in Run 119 was chosen based on calculations which indicated that below this energy, noise on the ionization channel would begin to cause gammas to fall at low yield. According to this result, these events would be problematic outliers, not consistent with Gaussian tails. In order to understand this issue for the case of Run 118, with its large electronic noise pickup, I have performed a realistic calculation of the ionization channel resolution with noise. (Other effects such as position dependence are not included). First, realistic noise samples are generated to match the noise spectrum of random trigger events from data. This varies somewhat over time, and especially between Run 118 and Run 119. The simulated noise traces can be made to match the real noise from any data set, but for this calculation, I have chosen as representative samples a very noisy data set from Run 118, 140110_0917, and a clean data set from Run 119, 140424_0955. In both cases I study the inner electrode of T1 Z1, a detector especially susceptible to
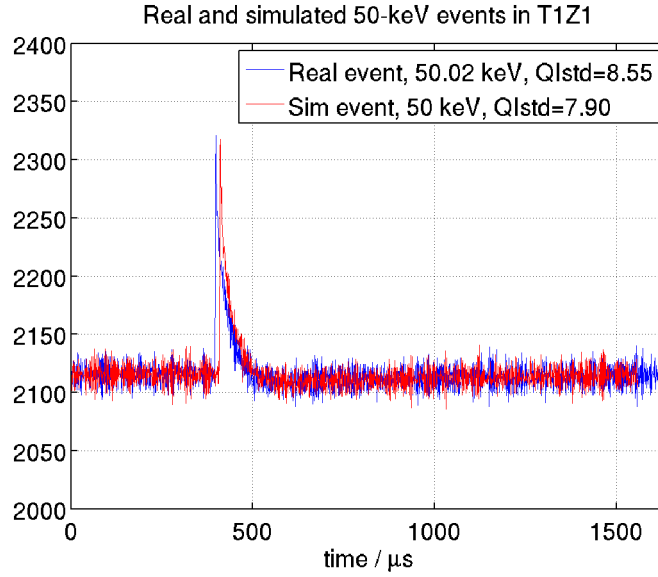
Figure 8.20: Simulated and real 50-keV events in T1 Z1, QI.

environmental noise and with relatively small signal-to-noise, in order to obtain a conservative worst-case picture.

Simulated ionization signals are constructed from the standard templates in the Dark Pipe optimal filter and F5 analysis. CDMS Pipe includes one such template for every channel; they are simply reused here. To create simulated events, the template is scaled to the desired energy using calibration constants from Pipe Cleaner, and shifted in start time. Cubic spline interpolation is used to allow the start time to vary by increments smaller than the digitization tick time, 0.8 μs. The pulses are then added to the generated noise samples and rounded off to give traces with noise, signals, and digitization identical to real data.

The simulated pulses are finally analyzed using the same optimal filter algorithm applied to all CDMS data in Dark Pipe, and with the filter parameters originally used for the same data sets. The ratio of fitted to nominal energy gives a "yield" for the optimal filter analysis in the presence of noise. The results from 200,000 simulated events under "noisy" Run 118 conditions are shown in Fig. 8.21. The yields were very consistent with a Gaussian distribution, and no fits gave $Y \leq 0.7$ at recoil energies above 3.8 keV. In addition, it was found that the "noisy" Run 118 conditions and
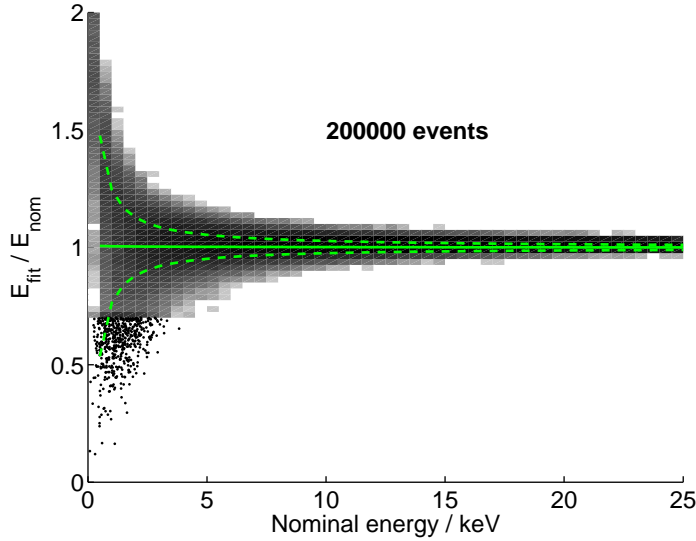
Figure 8.21: Optimal filter performance in the presence of noise. The vertical axis is the ratio of fitted to nominal ionization energy for simulated pulses for T1 Z1 QI in 140110_0917. On the horizontal axis is Luke-corrected recoil energy. Dangerous events with $Y \leq 0.5$ are shown as discrete points.

"quiet" Run 119 conditions resulted in essentially identical energy resolutions.

This test demonstrates that gamma leakage is not dangerous given the 5 keV energy threshold for the combined analysis, and that gamma rejection is not adversely impacted by the high noise levels of Run 118. At energies below 4 keV the gamma distribution flares out to high and low yield, but its spread is found to remain Gaussian, and our standard yield-band cuts are well suited to reject these events in a straightforward way. This result stands in contrast to the pessimistic projections used in the original Run 119 analysis, and favors a 5 keV analysis.

The 5 keV threshold is used for all detectors except T1 Z4 and T2 Z5. The silicon detector T1 Z4 is of an older design with larger energy dependence than the new T2 silicon detectors, and worse surface event rejection, especially at low energy. Therefore, its analysis threshold is set to 20 keV as in the Run 119 analysis. The germanium detector T2 Z5 has adequate performance at low energy, but its trigger threshold was unfortunately set high enough that it has no sensitivity to single scatter events below about 7 keV (see Sec. 8.2). Therefore, a 7 keV threshold is imposed.

Of the excluded detectors, T1 Z1 and T1 Z6 are analyzed with a 20 keV threshold, and T2 Z1 has a 5 keV threshold. These are not used for the WIMP result, but are included for background studies.

8.11 Efficiencies

In this section the combined efficiencies are calculated *before application of the timing cuts*. The timing cuts are handled separately in Chapter 9 along with their efficiencies. All cuts described in Chapter 8 are included in the efficiency functions here. The good data selection described in Sec. 7.3 is counted as a reduction of live time, rather than as part of the efficiency. Because of the slightly different noise performance and cuts, the efficiencies for the Tower 1 detectors are estimated separately in Run 118 and Run 119. The efficiencies are given for each detector in Figs. 8.22 and 8.23.
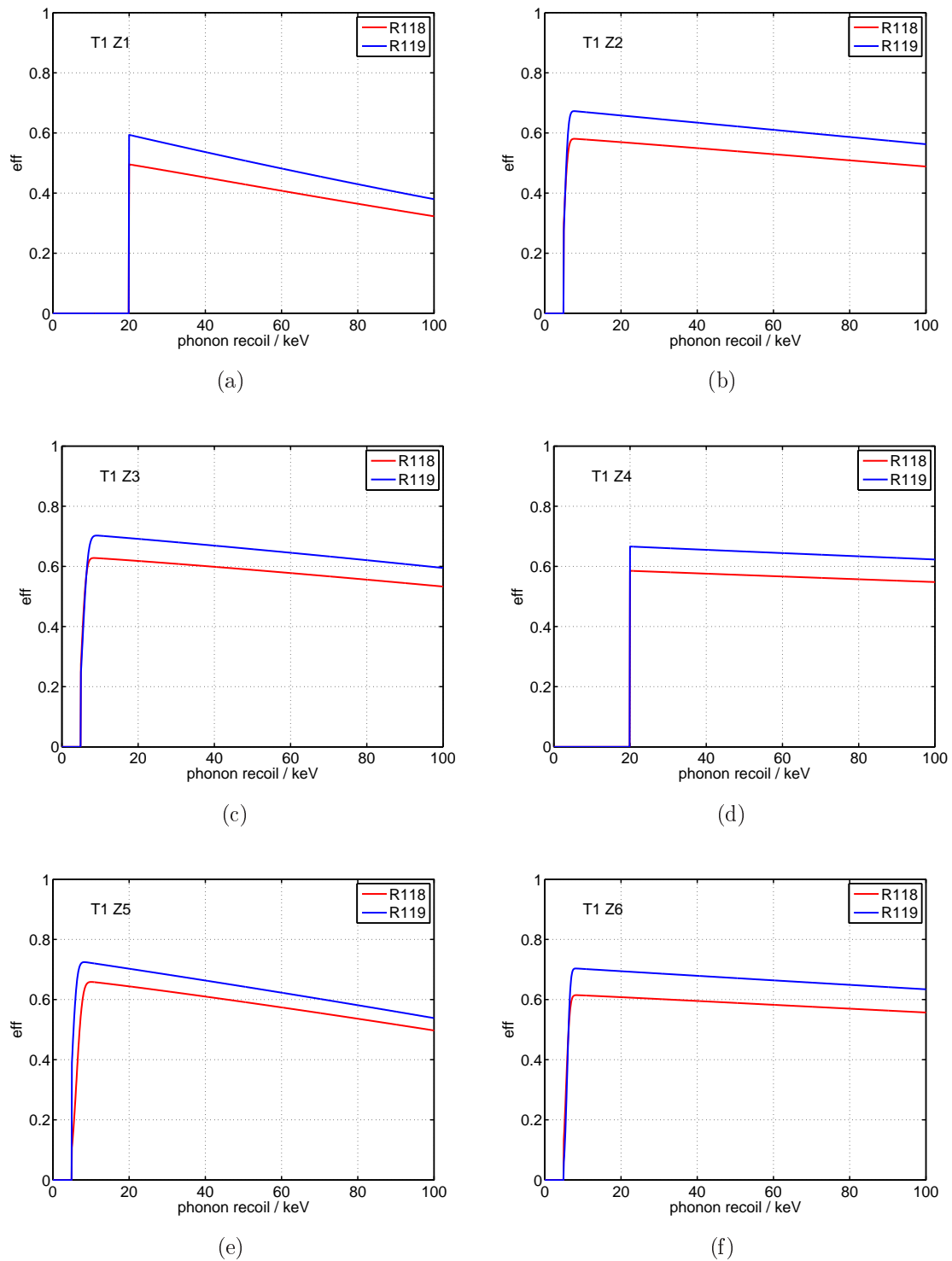
Figure 8.22: Overall efficiencies before timing cuts for Tower 1 detectors. The upper (blue) curve shows efficiencies for Run 119, and the lower (red) curve for Run 118. The Run 118 values are lower because of the sliding seal heater noise and the veto history buffer overflows.
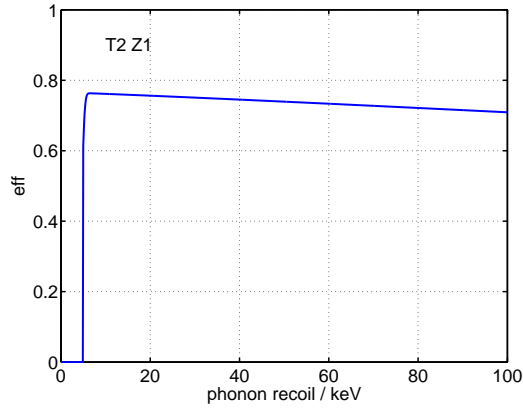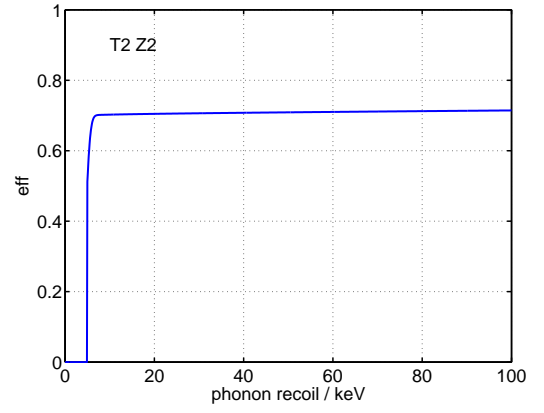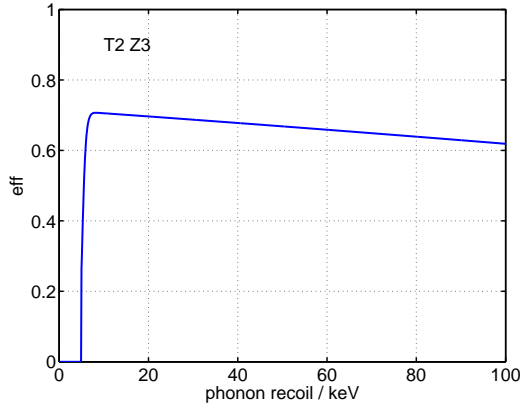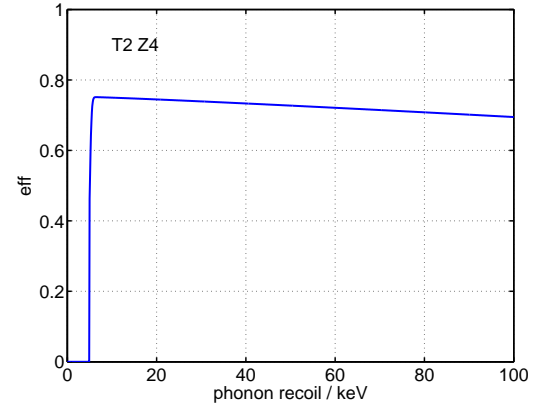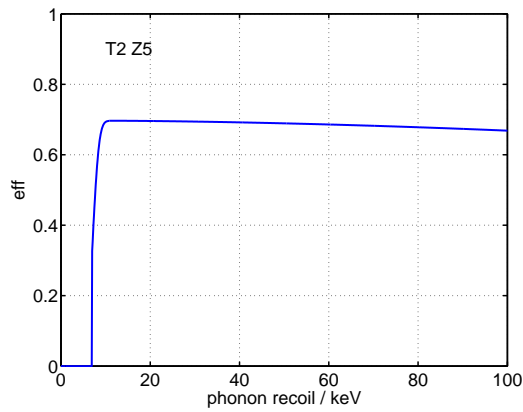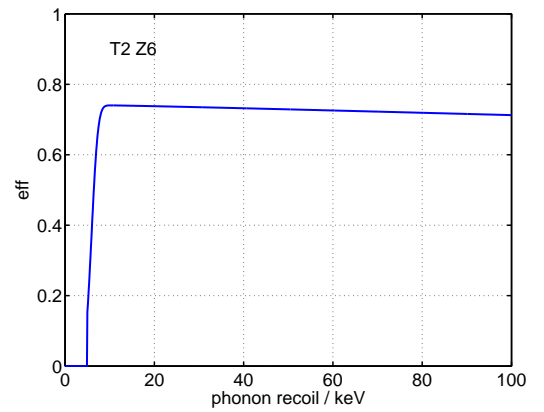
Figure 8.23: Overall efficiencies before timing cuts for Tower 2 detectors.

Chapter 9

Surface event rejection

The primary background for CDMS in the shallow underground site at Stanford (described in Sec. 3.1) consisted of neutrons produced by cosmic rays. Since unvetoed, single-scatter neutrons in germanium cannot be individually distinguished from WIMP interactions, the large flux of muons resulted in an irreducible neutron background that could only be statistically subtracted. The solution is to minimize the rate of cosmic-ray muons by moving the experiment to a deeper site. In the Soudan Underground Laboratory, the rate of muons is reduced to about one and a half per human hand per day, with an expected rate of cosmogenic single-scatter neutrons of 0.05 events in Ge and 0.05 in Si in the Run 119 data set .

In the absence of neutrons, surface events become the most important source of background (see also Sec. 5.3). When an electron recoil happens very close to the top or bottom surface of the detector, some charges of the wrong polarity may become trapped on the nearby electrode, so that the ionization signal is suppressed. Within a very thin "dead layer" this suppression can be severe enough to make electron recoil events fall into the nuclear recoil band. Surface events in the WIMP-search data set are primarily produced by the beta decay or very low-energy gamma decay of radioactive contaminants on the outer layers of the detectors themselves. A second source of surface events, more prominent in gamma calibration data, consists of secondary electrons produced by a gamma scatter near the surface of one detector and ejected towards the neighboring detector. These interactions are informally termed
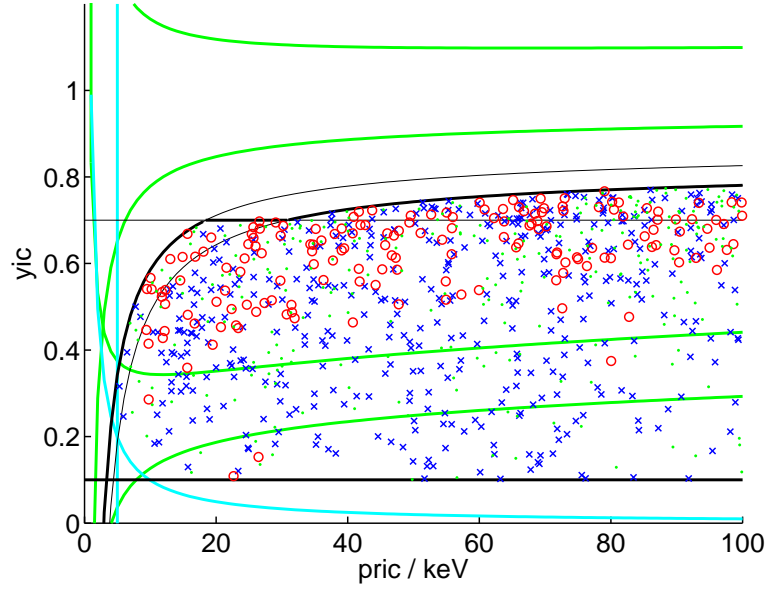
Figure 9.1: Yield of betas on the phonon side (blue X) and ionization side (red O) of T1 Z5. Betas that could not be tagged by side are shown as small green points.

"ejectrons," and can be tagged by selecting events with a bulk electron recoil in one detector and a surface event in its neighbor.

9.1 Discriminators for surface event rejection

9.1.1 Ionization yield

Although surface events have reduced ionization relative to bulk electron recoils, ionization yield remains effective in rejecting most of them. At energies greater than $\sim$ 60 keV the "beta band" merges with the main "gamma band." This is expected, since electrons of higher energy become more penetrating, so that their energy deposition is no longer contained within the dead region. At low energy the beta band droops, and has greater overlap with the nuclear recoil band.

The ionization yield cut is not usually considered part of the surface event rejection scheme, but it has a very significant impact on the surface event cuts. Because the high-yield betas are already rejected, the timing cuts only need to discriminate against

low-yield betas, which have timing and other properties most strongly influenced by surface effects. In particular, the yield cut and timing cut are complementary for events on the two sides of a ZIP detector. Surface events on the top (phonon side) have more strongly suppressed yield, but also have faster timing; events on the bottom (charge side) have higher yield, but timing properties more like bulk gammas (Fig. 9.1). Therefore, phonon-side surface events can more easily be rejected by the timing cut, while charge-side events can more easily be rejected by the ionization yield cut. Because many surface events are multiple scatters, including the "ejectrons" and a fraction of beta decays, we can often tag which side an interaction happened on. Although this is not possible for WIMP candidates, which are single scatters, we account for the difference between phonon-side and charge-side surface events in the calibration data sample and in the WIMP-search multiple scatters when calculating the expected beta leakage.

9.1.2 Phonon rise time

The original parameter used by CDMS for rejecting surface events is the rise time of the phonon pulse. When an event occurs near the surface of the detector, many of the phonons quickly reach the metal films of the QETs or electrodes. Interactions with the metal allow the phonons to quickly down-convert to lower frequency, faster-traveling modes (see also Sec. 2.3). As a result, the phonon energy reaches the aluminum fins faster and causes a faster-rising pulse. This is especially true on the phonon side of the detector, for which the phonons are generated very close to the absorber fins, but is also true for events on the charge side. A fraction of the phonons down-convert in the metal of the electrode on the charge side, and are then able to travel ballistically (and quickly) across the 1-cm thickness of the detector to be absorbed at the phonon side. The enhanced ballistic component is especially evident near the beginning of the rising edge of the phonon pulse.

In the analysis of SUF Run 21, the first CDMS result using ZIP detectors, the surface event analysis consisted of a single energy-dependent cut on rise time alone. The
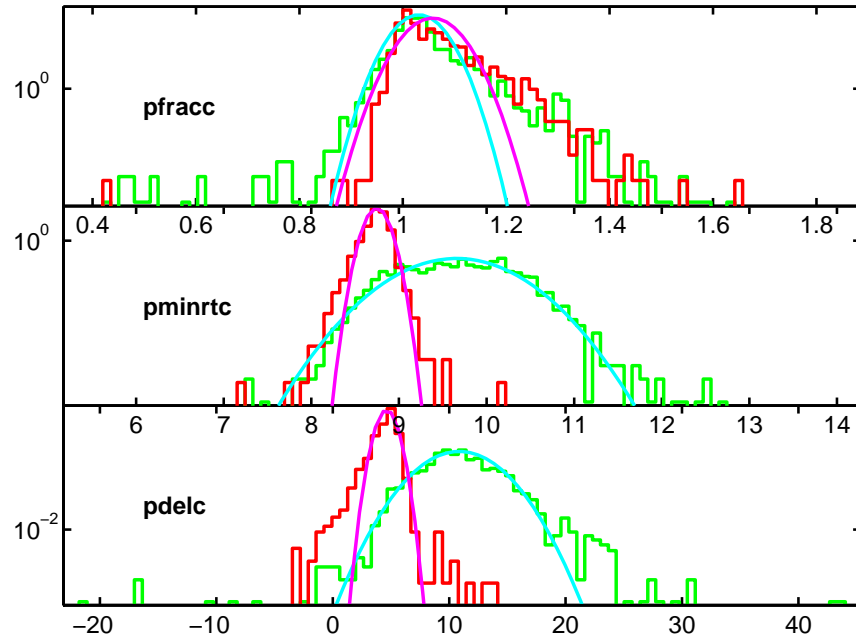
Figure 9.2: Timing discriminators in a Ge detector, T1 Z5. The histograms of the three parameters (phonon fraction, delay, rise time) for neutrons (green, lighter) and betas (red, darker) are overlaid with Gaussian fits, which will be used in defining the χ^2 timing cuts.

rise time was defined as the time for the phonon pulse, summed over the four quadrants, to rise from 10% to 40% of its peak value. The Run 21 analysis also introduced the phonon look-up table correction, initially devised to remove position dependence of the ionization yield. However, the rise time also has a position dependence, so the look-up table correction was used to make the rise time uniform across the detector as well. In the analysis of Soudan data, the rise time parameter has been redefined in terms of the phonon pulse in the primary quadrant only, rather than the sum of the four quadrants. This definition has a larger intrinsic position dependence, because events near the edge of a quadrant behave differently from events in the center of the quadrant. However, the look-up table correction is very effective in removing this effect to create a uniform discrimination parameter. Therefore, we choose to maximize the signal-to-noise by using the rise time from the primary quadrant, which retains most of the signal while experiencing the noise from only one of the four quadrants (on the average, a factor of two improvement in noise). This rise-time in the local quadrant is known as *pdel*, or *pdelc* after position correction. The distributions of betas and neutrons in *pdelc* are shown in the lower pane of Fig. 9.2.

9.1.3   Phonon pulse delay

The faster down-conversion of phonons makes the phonon pulse for surface events not only rise more steeply, but also arrive sooner relative to the (immediate) ionization signal. Therefore, delay can also provide surface event rejection. The phonon delay is defined as the time elapsed between the 20% rise of the ionization pulse and the 20% rise of the phonon pulse in the primary quadrant. This is position-corrected using the look-up table, in the same way as rise-time.

Besides the physical difference in delay between surface and bulk events, there is also a degree of discrimination introduced by our timing analysis. As described in Sec. 8.1.1, pulse timing is calculated using a "risetime-falltime walk" algorithm which first applies a second-order Butterworth filter. In Ge detectors, this filter has been set at 22.5 kHz, which is a low enough frequency to affect the shape of the pulse rising edge. This causes the leading edge of fast-rising pulses to be made less steep,

which moves the start time earlier. The net effect is that rise time is made more uniform, and timing discrimination is shifted into delay instead. For Ge detectors in this analysis, delay is the most powerful discriminator for this reason. The delay of phonons relative to ionization is called *pdel*, or *pdelc* after position correction. The distributions of betas and neutrons in *pminrtc* are shown in the middle pane of Fig. 9.2.

9.1.4 Phonon partition

When interactions occur very close to the phonon side of the detector, a large fraction of the total phonon signal can be absorbed by fins close to the interaction point. This causes the partition of phonon energies to favor the local quadrant. A bias towards the local quadrant can be quantified in several different ways, but we have chosen to use the *phonon fraction*, defined as the phonon energy in the primary quadrant divided by the phonon energy in the diagonally opposite quadrant. (It has been shown that *partition radius*, the distance from the center on a box plot, has similar properties and would work as well.)

These partition quantities are primarily *x-y* position indicators, and only secondarily give information about interaction depth. The primary *x-y* information must be removed well before partition can become a useful discrimination parameter. This has been done using the same look-up table correction as for ionization yield, delay, and rise time, to give a *corrected phonon fraction*. This corrected quantity has an arbitrary normalization to 1 for bulk events, and is no longer directly physically meaningful as a measure of the distribution of energy among the phonon sensors. Instead, it is a relative quantity that measures the degree to which the phonon energy for an event is more locally concentrated than for other nearby events. In many of the Run 119 and later analyses, this phonon fraction has been used as a major discrimination parameter along with rise time and delay, as a third "timing" quantity. It is especially useful for rejecting surface events on the phonon side, where ionization yield is most strongly suppressed. The usefulness of *pfracc* comes primarily from its correlation with rise time and delay, which the chi-squared timing cuts are able to take advantage

of (see Fig. 9.8). The phonon fraction is abbreviated $pfrac$, or $pfracc$ after position correction. The distributions of betas and neutrons in $pfracc$ are shown in the top pane of Fig. 9.2.

9.2   Types of surface event cuts

Rise time only

The first SUF Run 21 analysis used a simple cut on position-corrected rise time, with the cut position set independently in each of several energy bins for each detector. This was sufficient in the SUF case, since cosmogenic neutrons were the dominant background rather than surface events.

Yield-rise time maximum likelihood

In the analysis of the full 3V and 6V data sets from SUF Run 21, Don Driscoll fit functional forms to the distributions of surface events and neutrons in the yield-rise time plane [74]. Neutron and beta likelihoods were calculated from these functions, and combined to make a likelihood ratio upon which the cut was defined.

Although this approach can be effective, subsequent analysis has taken the approach of defining independent yield and timing cuts. This allows the problems of gamma rejection and beta rejection to be separated for conceptual simplicity. The chi-squared analyses described below could naturally be generalized to make a combined yield-timing cut by adding yield as an extra dimension in the chi-squared.

Rise time and delay

The simplest way to use both rise time and delay in a single cut is to combine them into a single parameter with the best discrimination between neutrons and betas. A consistency cut is defined in the perpendicular direction, rejecting outliers whose rise time and delay are not well matched.

Position-tuned cut

In his doctoral dissertation, Gensheng Wang defined a timing cut with different thresholds for events near the center of the detector ($r \leq 3$ cm) and events in the outer region ($r > 3$ cm) [88]. Although this precise approach has not been repeated here, we follow in the same spirit by defining a radial correction of the timing parameters. This is used in two cuts, the r-corrected simple timing cut and the r-corrected chi-squared timing cut.

Chi-squared

A cut can be defined to take advantage of any number of discrimination parameters by defining a chi-squared for each event relative to a neutron model and a beta model. First, the distributions of neutrons and betas in each parameter are modeled by gaussians, taking into account the correlations between parameters. Each event then has a χ^2_n and a χ^2_β based on how well it matches these model distributions. The cut is defined in terms of the difference $\chi^2_\beta - \chi^2_n$, with consistency and quality cuts to remove outliers.

The Run 119 blinded analysis developed two distinct chi-squared cuts: one using a single neutron and a single beta model, and one with energy-dependent models. Each analysis was defined in terms of the three parameters rise time, delay, and phonon fraction. We define several χ^2 cuts in this analysis, including the primary timing cuts for Ge and Si.

Neural network

Several neural network analyses were developed for Run 119 by Michael Attisha. One of these was defined in terms of delay, rise time, and phonon fraction, and was completed during the Run 119 blind analysis. A later neural network analysis operated on wavelet coefficients calculated from raw event traces, rather than the standard reduced quantities [97]. Although a neural network timing analysis is being carried out for the current combined analysis, it has not been brought to completion in time for this thesis.

Linear discriminants

Preliminary work has been done by Jodi Cooley and Rupak Mahapatra on using Fisher discriminants or linear discriminant analysis (LDA) for a timing analysis [98]. This is a family of related techniques, similar to principal component analysis, which prescribe a recipe for constructing a linear combination of timing parameters to use as a discriminant. Although detailed comparisons have not yet been performed, this may be expected to generate discriminators very similar to the simple "rise time and delay" cut traditionally used by CDMS. So far, the LDA approach has been applied to the Run 119 data set, but not to the reprocessed, combined analysis data set.

Projection and correlation approximation

Darren Grant has applied the projection and correlation approximation (PCA[1]), a cut-free signal extraction technique, to CDMS timing analysis [99]. The approach is to apply a coordinate transformation to the input timing parameters, in which the distribution of neutrons and betas are generally not Gaussian, to construct a space in which they are Gaussian. Correlations and likelihood functions are calculated in this space, giving an estimate for the probability that any given candidate event comes from the neutron or beta population. Although the PCA analysis has been applied to the combined Run 118 + Run 119 data set, it will not be covered in detail in this thesis.

9.3   T2 Z5 "cliff cut"

One germanium detector, T2 Z5, has a distinct area of altered ionization yield near the edge of the detector, between quadrants B and C (see Sec. 5.11 and particularly Fig. 5.14). The look-up table position correction removes the effect in yield reasonably well, but we find that the affected region still contains a disproportionate number of low-yield timing outlier events. It appears that the look-up table does not work well

---

[1]Not to be confused with principal component analysis, the Presbyterian Church in America, or El Partido Comunista de la Argentina, with which it shares an acronym.
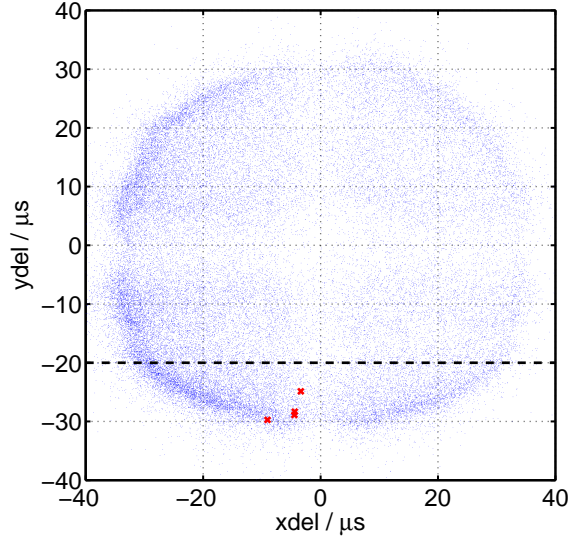
Figure 9.3: "Cliff cut" and timing outliers in T2 Z5. The blue points are electron recoils from ^{133}Ba calibration, and the four red crosses are calibration betas that pass the primary timing cut in Ge. All four are cut out by the "cliff cut."

in the presence of the very sharp transition, causing some bulk electron recoils to be corrected to low yield. Therefore, we use a "cliff cut" to remove the affected region. This is not strictly a surface-event rejection cut, but it is treated as part of the timing cuts for several reasons. First, the cut is used to remove a set of "outlier" events that are difficult for the timing cuts to deal with, but that would otherwise be accepted. Second, the outlier events in question have timing properties similar to bulk electron recoils: they have delay and rise time faster than those of nuclear recoil events, but slower than surface events. We tune our timing cuts to remove surface events and keep neutrons, while the effect on bulk electron recoils is unimportant. Given the diversity of timing analyses, some tend to keep events with bulk gamma-like timing, while others have good efficiency for rejecting these events. Therefore, in the Run 119 analysis it was left up to the developers of each timing cut to decide whether to impose the cliff cut, losing a small amount of exposure but eliminating a class of problematic events, or to leave it out. We continue to classify it along with the timing cuts. The cliff cut is imposed for all cuts in the current combined analysis.

We considered several forms for the cliff cut, but settled on a very simple position-based cut. The cliff cut, identical to the one used in the original Run 119 analysis, rejects events with $ydel < -20 \mu$ s and keeps all others. The loss in exposure in T2 Z5 is about 25%, and is factored into the efficiency of each timing cut. This cut is shown in Fig. 9.3, along with the betas passing the primary Ge timing cut.

9.4   Defining timing cuts

The surface event rejection cuts have been defined blind, using only the calibration data. A standard framework has been developed through the initial Run 118 and Run 119 analyses, and further developed in the combined analysis.

9.4.1   Beta selection cut

The traditional CDMS definition of a "beta" has been anything below the electron recoil band and above the nuclear recoil band. However, this suffers from several difficulties. First, the beta region becomes narrow and disappears at low energies, as the nuclear recoil and gamma bands widen and eventually merge. Second, it does not take advantage of the very low-yield betas from ^{133}Ba calibration data, some of which are around or inside the nuclear recoil band. Third, many of the events in this region are more gamma-like than beta-like in timing and pulse shape, so that a tighter yield-based cut may be needed. Finally, it does not take advantage of information from other detectors in multiple scatter events.

An important complication is the interaction of detector-detector crosstalk with beta selection. Many of the betas in the barium calibration data are multiple scatters. If there is negative-going crosstalk from one detector to another in the ionization channels (as detailed in Sec. 5.5), events can be artificially pushed down out of the gamma band, and shared events can be pushed into the inner electrode region. These appear to have the yield of betas, but their timing properties are those of gammas. This creates a class of outlier betas with relatively slow timing, which are hard to cut out. They can never become WIMP candidates, since they are multiple scatters;
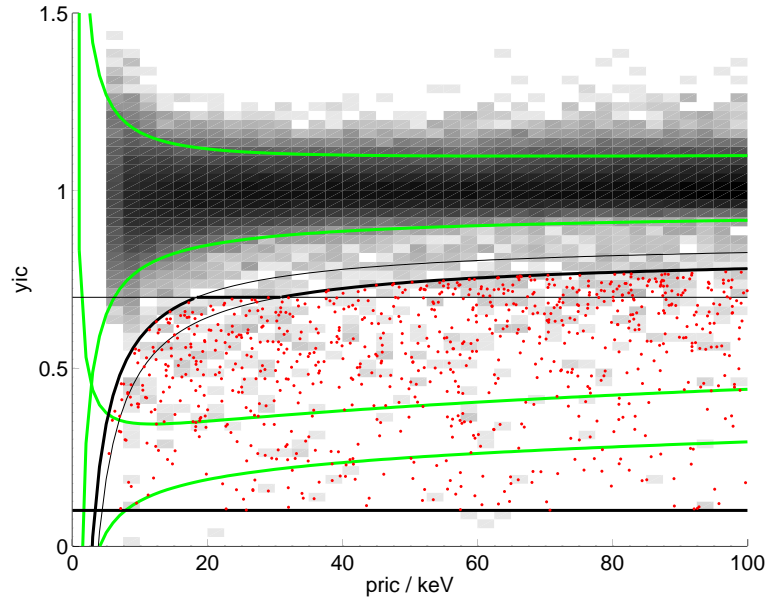
Figure 9.4: Beta selection in the yield-recoil plane, for T1 Z5. The green bands are the electron recoil and nuclear recoil yield bands. Barium calibration data are shown: shading indicates the density of events, and red points are selected betas. The lower horizontal line is at yield=0.1, the lower limit of the beta region. The two dark curved lines are at 4σ and 5σ below the gamma band, forming the upper edge of the wide beta band at low and high energies respectively. The definition crosses over at yield of 0.7, indicated by the upper horizontal line.

but by leaking into the beta sample for setting the timing cuts, they can force the cuts to become unnecessarily strict, and artificially inflate the leakage estimates. In order to avoid this, the Run 118 + Run 119 combined analysis adds several new requirements to the beta selection cut. First, any betas in which the detector below has a significant amount of energy ($\geq$5 keV) must also pass the inner electrode cut in that detector. This should eliminate most of the crosstalk events, which are primarily crosstalk from the outer electrode. Some are nominally within the inner electrode, but should fail the fiducial volume cut because of the holiday island effect. Second, the detector below must not have more than a certain maximum energy (200 keV). In order to avoid contamination with gammas near the QO boundary, the beta selection cut uses a stricter version of the inner electrode cut defined in Sec. 8.5), demanding both that *qo* be within 2σ of the noise, and that *qpart* be between 0.8 and 1.2.

The final Run 118 + Run 119 beta selection cut (shown in Fig. 9.4) is constructed to choose events that:

- Are at least 5σ below the electron recoil band, *or* have yield below 0.7. (This takes effect at higher energies).

- Are at least $N_{LE}\sigma$ below the electron recoil band, where $N_{LE} = 4$ in most detectors. (This takes effect at lower energies). The exceptions are: $N_{LE} = 4.5$ in T1Z1; 4.2 in T1 Z2 and Z6; and 4.4 in T1 Z4.

- Have yield of at least 0.1.

- Anti-crosstalk cut: the detector below must have $qsum \leq 3$ keV, or must have $qsum \leq 3$ keV *and* pass the tight QI cut. (This is not imposed for several of the cut definitions, as described below).

The transition from a 4σ cut at low energies to a 5σ cut at high energies is designed to avoid accepting too many bulk electron recoils from the tail of the gamma band, while keeping enough low-energy betas.

9.4.2   Surface event samples

Surface events are selected from the ^{133}Ba calibration data using the beta selection cut of Sec. 9.4.1. In Run 119, the very large barium calibration data set allowed two independent subsets to be defined: the "open" and "closed" barium data. These were interleaved run-by-run, so that one data set (typically a few hours) was "open" and the next was "closed." The open data was used to set cut positions, and the closed data was used to calculate the leakages and efficiencies of cuts. In this way, independent samples were used for defining the cuts and for evaluating their performance. The alternation between open and closed sets was intended to ensure that any changes over time were equally represented in each sample.

For the Run 118 + Run 119 combined analysis, this has been refined to event-by-event alternation. The events with *odd* event number are used for defining all cuts, and the events with *even* event number are used for calculating leakages and

efficiencies. This ensures that even temporary variations in performance affecting only a short period of time or a single data set are identically represented in both samples. The independence of the two samples avoids the possibility that the calculation of cut performance is biased by using the same events for cut definition and characterization. Details of the generation of these two halves are contained in Sec. 8.1.4.

The WIMP-search data itself provides a final sample of surface events. Of course, the WIMP-search singles in the nuclear recoil band, which may eventually become WIMP candidates, are blinded and cannot be used. However, we can study three beta samples in the background data: wide beta multiple scatters; multiple scatters in the nuclear recoil band; and single scatter wide betas outside the nuclear recoil band. These form a better model for the true beta background than source calibration events do, because surface scatters from radioactive contamination on the detectors have somewhat different properties from "ejectrons." However, statistics are lower. We use these low-background samples outside the WIMP search signal region in calculating the expected leakages of the timing cuts.

9.4.3   Energy correction

In Run 119, an additional energy correction (beyond the energy linearization in Pipe Cleaner) was performed on the rise time and delay quantities used for the primary timing cut. A similar correction is explored here, and a well-defined procedure is described for performing the correction. However, it is ultimately left out of the timing cuts used here.

The primary timing cut of the initial Run 119 analysis used the two timing parameters rise time and delay. These are represented by the calculated quantities *pminrtc* and *pdelc* respectively - the position-corrected versions of the raw pulse rise time and delays. In order to construct a cut that is flat with energy, an additional energy correction was performed. The correction factors use the functional form $f(E) = C + \sqrt{A + \frac{B}{E^2}}$, where E is the phonon recoil energy, and A, B, and C are chosen by hand to make $f(E)$ visually match the upper edge of the beta distribution.

Note that the position correction used to generate *pminrtc* and *pdelc* also includes
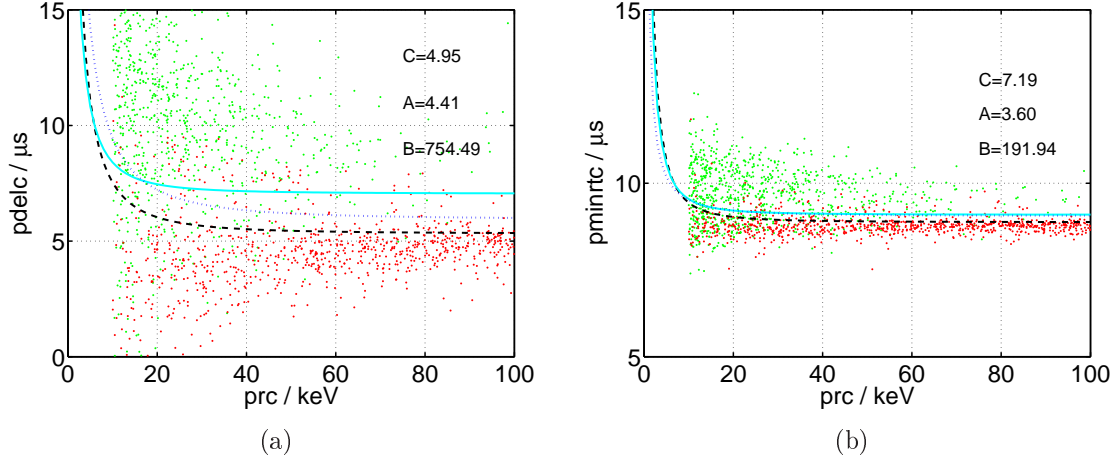
Figure 9.5: Beta timing vs. energy and band edges in data from published Run 119 analysis. Phonon pulse delay is shown in (a), and rise time in (b). The green (light) events are selected neutrons, and red (dark) are surface events from photon calibration data. The curves are as follows: blue, dotted - chosen by eye in published Run 119 timing cut; black, dashed - fitted here using 80% point; cyan, solid - fitted here using 90% point. A single detector, T1 Z5 (Ge), is shown.

an energy correction, for which the criterion is to make the mean of the bulk electron recoil distribution flat with energy: the center of the gamma band is horizontal. The energy flattening discussed here is a distinct additional step with a different criterion. It is defined with reference to the surface rather than bulk electron recoils, and it it attempts to make the upper edge constant rather than the center. This criterion is more directly applicable to defining a surface event rejection cut than the simple energy correction already performed.

For the combined analysis, the same correction has been revisited but with a more systematic approach. Instead of matching the upper edge of the beta population by eye, the parameters are found by making the curve $f(E)$ fall above a fixed fraction (80% or 90%) of the betas. This fraction is to be independent of energy as far as possible. A new fitting routine created for this purpose defines a residual that is large whenever the fraction of betas below the curve deviates from the specified value, and finds the values of A, B, and C that keep the residual as small as possible. The fitting routine *bandedge* is described in Appendix D.3.
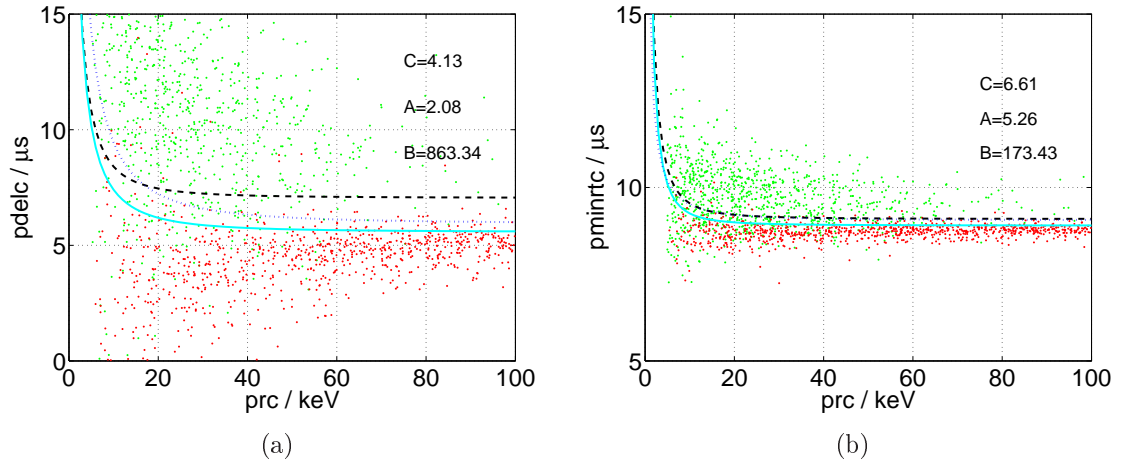
Figure 9.6: Beta timing vs. energy and band edges in Run 118+Run119 combined analysis. Phonon pulse delay is shown in (a), and rise time in (b). The green events are selected neutrons, and red are surface events from photon calibration data. The curves are as follows: blue, dotted - chosen by eye in published Run 119 timing cut; black, dashed - fitted on initial analysis Run 119 data; cyan, solid - fitted here on Run 118+R119 combined analysis data. The curves shown are fit to the 90% band edge in all cases.

Table 9.1: Energy flattening parameters for rise time.

T1	A	B	C		T2	A	B	C
Z1	4.33	265.2	7.02		Z1	5.16	169.9	7.07
Z2	6.41	59.4	6.27		Z2	5.92	215.1	6.29
Z3	5.21	191.3	6.62		Z3	4.82	207.8	6.60
Z4	11.84	617.5	6.02		Z4	5.35	179.3	6.50
Z5	5.26	173.4	6.61		Z5	5.29	216.7	6.61
Z6	4.13	292.5	7.51		Z6	5.14	181.9	6.76

The first step is to revisit the Run 119 data in the form used for our initial publication. Since this earlier analysis, the look-up table correction has been redone with higher-statistics tables, different functional forms, and more careful control over the selection of reference events. Therefore, the results for the initial Run 119 analysis and the current Run 118 + Run 119 combined analysis may differ. The initial Run 119 data is revisited here in order to compare the results of the systematic band edge fit with the correction curves chosen by eye in the earlier Run 119 timing analysis. The band edge has been defined by both an 80% and a 90% edge for comparison.

The results of the fit in Fig. 9.5 generally show a smaller energy dependence than the functional forms used in the initial Run 119 analysis. The 80% curves are much flatter, and the 90% curves are similar but still indicate a smaller correction than that used in the published Run 119 result.

The Run 118 + Run 119 combined analysis shows even less need for an additional energy flattening, as shown in Fig. 9.6. This is likely thanks to the improved form and more careful construction of the correction tables for the combined analysis. Below 10 keV, however, there is still a significant trend towards widening of the betas. We choose not to apply this correction for the Run 118 + Run 119 timing cuts.

Subsequent CDMS runs have modified the timing parameters: they use Butterworth filters at higher cutoff frequencies, and calculate the ionization pulse start time from the optimal filter rather than a time-domain walk algorithm. Therefore, the question of whether or not an additional energy flattening for betas is needed for Run 123 and later will have to be revisited in subsequent analyses.

Table 9.2: Energy flattening parameters for delay.

T1	A	B	C		T2	A	B	C
Z1	2.03	550.4	4.59		Z1	1.49	714.5	5.22
Z2	5.55	521.6	3.32		Z2	1.88	591.0	4.77
Z3	2.18	1304.2	4.13		Z3	4.40	3627.9	2.90
Z4	0.43	1726.6	5.39		Z4	2.02	550.6	4.67
Z5	2.08	863.3	4.13		Z5	0.53	928.4	4.89
Z6	3.59	771.0	4.57		Z6	2.07	496.0	4.72

9.4.4 Radial correction

The position reconstruction described in Chapter 6 lends itself naturally to a full analysis, making use of all available measured quantities to form the best possible estimates of event location and recoil type. Such an analysis is desirable, but it must unfortunately lie outside the scope of this thesis. Since the position reconstruction is available now, it is still worthwhile to consider its application to the improvement of our timing cuts.

The most direct approach to adding radial position information to the timing cuts is with a new correction, patterned on the energy flattening described in Section 9.4.3. The pulse delay and rise time typically rise near the outer edge of the detector, even after the look-up-table correction has been performed. There are several reasons for this. First, the look-up table itself becomes more ambiguous near the outer edge, where both delay and rise time experience severe foldback. Second, the betas and neutrons tend to fall far from the main gamma manifold in this outer region, so that they are corrected relative to gamma events that may not have occurred in exactly the same physical region of the detector. Third, any outlier events that leak through the fiducial volume cut may have low yield and gamma-like timing, and they are associated with the outer regions of the detector. Therefore, the expected phenomenology of the timing parameters is to have a constant value in the inner several centimeters, and then to rise near the detector edge. There should be no strong dependence on the azimuthal angle θ.

This is very clearly demonstrated in some detectors, such as T1 Z5 (Ge). A natural

functional form for the new flattening is

$$
\begin{aligned}
f(r) &= a, r \leq r_0 \\
f(r) &= a + b(r - r_0), r > r_0
\end{aligned}
$$

The sharp corner at $r = r_0$ is smoothed out at a small scale using exponential functions. There are several ways to fit the parameters of this functional form to the data. One approach is to use the function *bandedge* (Appendix D.3), exactly as in the energy correction. A second approach is to explicitly maximize the ratio S^2/B, where S is the number of signal events (neutrons) accepted, and B is the number of background events (betas) leaked. This suffers from rather sensitive dependence on the exact positions of beta outliers, since it is dominated by the few most extreme events in B. A third, very simple, approach is to simply fit the mean of the betas as a function of radius and use this as the correction value. This simple approach is the most stable against excessive perturbation by outliers, and is used here. The first two corrections have also been explored, but do not perform better and introduce additional complication. The correction function is simply subtracted from the reduced quantities *pdelc* and *pminrtc*.

9.5   Cuts for Run 118 + 119 combined analysis

Several types of cuts have been defined for the Run 118 + Run 119 combined analysis, with further developments. All of them include the T2 Z5 "cliff cut" described in Sec. 9.3. The first four cuts described below (the simplest cut, radius-flattened cut, simple chi-squared cut, and radius-flattened chi-squared cut) are defined in pairs for comparison, to determine whether or not the radial flattening using the reconstructed position quantities developed in Chap. 6 improves the performance of the timing cuts. For consistency in this comparison, these four cuts reject all events that do not have reconstructed radius $0 \leq r \leq 4$ cm. This removes events for which the position reconstruction failed, and has a very high efficiency on neutrons, $\sim 99\%$ in most detectors. The other cuts do not use the position reconstruction, so this condition is

not imposed.

Each cut has been tuned to allow a fixed number a of betas from the odd Ba calibration sample to pass the cut in each detector. For the ordinary cuts, $a = 8$ in Ge detectors and $a = 4$ in Si detectors. For the "tight" cuts, $a = 3$ in Ge and $a = 2$ in Si. The energy-dependent chi-squared cuts also have "coadded" versions, in which a single cut level is set for all Ge detectors, and a second cut level for all good Si detectors. In this case, the cut is set to leak a times the number of good detectors for Ge, and separately in the same way for Si.

9.5.1   Simplest cut

The simplest type of timing cut is defined on rise time *pminrtc* and delay *pdelc*, combined into a single timing parameter. First, the median value of *pminrtc* and *pdelc* is found for the beta and neutron samples. The median is used rather than the mean because of its insensitivity to outliers. The axis of the combined timing parameter is the line joining the two medians in the rise time-delay plane. It is defined to take the value -1 at the beta median and +1 at the neutron median point. The timing cut is defined in terms of this parameter.

There is also a consistency cut, imposed in the perpendicular direction. The consistency cut is placed at $\pm 2\sigma$ of the neutron distribution. In effect, it demands that the rise time and delay of each event match each other. An event with fast rise time but slow delay, or vice versa, may have suffered problems in pulse reconstruction, or have other artifacts such as multiple interactions, and is rejected. The construction of the timing parameter and consistency cut are shown in Fig. 9.7. This procedure is repeated for each detector independently.

9.5.2   Radius-flattened cut

A slight variation on the simplest cut is to use radius-flattened rise time and delay as defined in Section 9.4.4 instead of *pminrtc* and *pdelc*. The construction of a single timing parameter and a consistency cut are otherwise identical to Sec. 9.5.1.

(a)

(b)
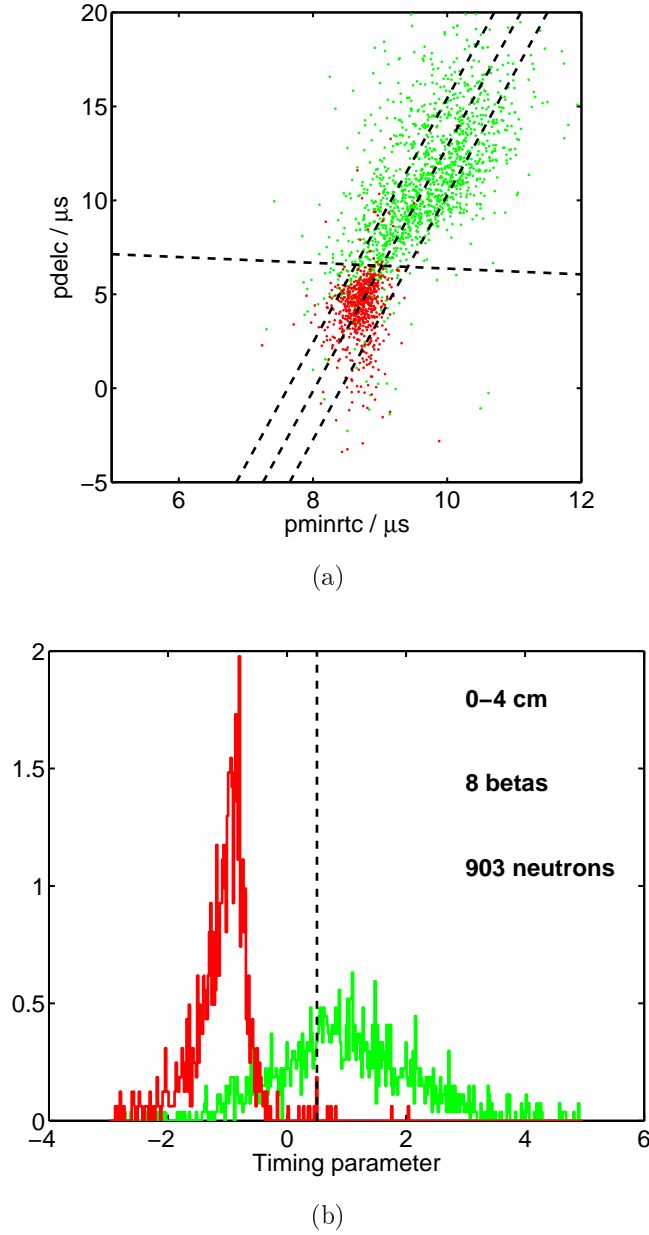
Figure 9.7: Simple timing parameter cut with parameter definition and cut position in T1 Z5 (Ge). Panel (a) shows the construction of the cut in the delay-rise time plane. Neutrons are in green (on the right), and betas in red (on the left). The three diagonal lines show the boundaries of the consistency cut (outer lines) and the axis of the combined timing parameter (inner line). The nearly-horizontal line is a possible beta rejection cut. The beta rejection is mainly achieved with delay instead of rise time because of the effects of the Butterworth filter. (Note that the consistency cut and the discrimination cut are exactly perpendicular. They appear to meet at an acute angle only because of the different axes scales for delay and rise time). Panel (b) shows the setting of the cut level to keep exactly 8 betas. The majority of the neutrons pass the cut.

9.5.3 Simple chi-squared cut

A simple chi-squared cut has been defined based on the energy-independent chi-squared analysis from Run 119. First, a Gaussian is fit to the neutrons and betas in *pdelc*, *pminrtc*, and *pfracc* (Fig. 9.2). This defines the mean μ and width σ of each population in each dimension. Next, the covariances C_n and C_b of the three parameters are calculated. The covariances are shown as ellipses in Fig. 9.8. The chi-squared for candidate events is then given by

$$\chi_n^2 = \sum_{i,j} (x_i - \mu_i^n) \left(C_n^{-1} \right)_{ij} (x_j - \mu_j^n), \tag{9.1}$$

where i and j index over the three timing parameters, and n indicates the chi-squared relative to neutrons. The chi-squared relative to betas is defined analogously.

An effective timing cut requires both that candidate events look more like neutrons than betas, and that they match the neutron distribution itself sufficiently well. In this way it is possible to effectively reject betas, and to avoid accepting outliers that are unlike either neutrons or betas. The outliers have been rejected by demanding $\frac{\chi_n^2}{\text{d.o.f.}} \leq 4$, where d.o.f. is the number of timing parameters used, in this case 3. Betas are rejected by demanding $\chi_b^2 - \chi_n^2 \geq d$, with the value of d tuned to allow the desired leakage of betas from the odd barium calibration data set. This tuning is performed independently for each detector. As with the simple timing cuts, we use $a = 8$ for Ge detectors, and $a = 4$ for Si. The position of the cuts for one germanium detector is shown in Fig. 9.9.

9.5.4 Radius-flattened chi-squared cut

A similar chi-squared cut has been defined using the radius-flattened quantities from Sec. 9.4.4 rather than *pdelc* and *pminrtc*. The phonon fraction *pfracc* is left unmodified.

(a)

(b)

(c)
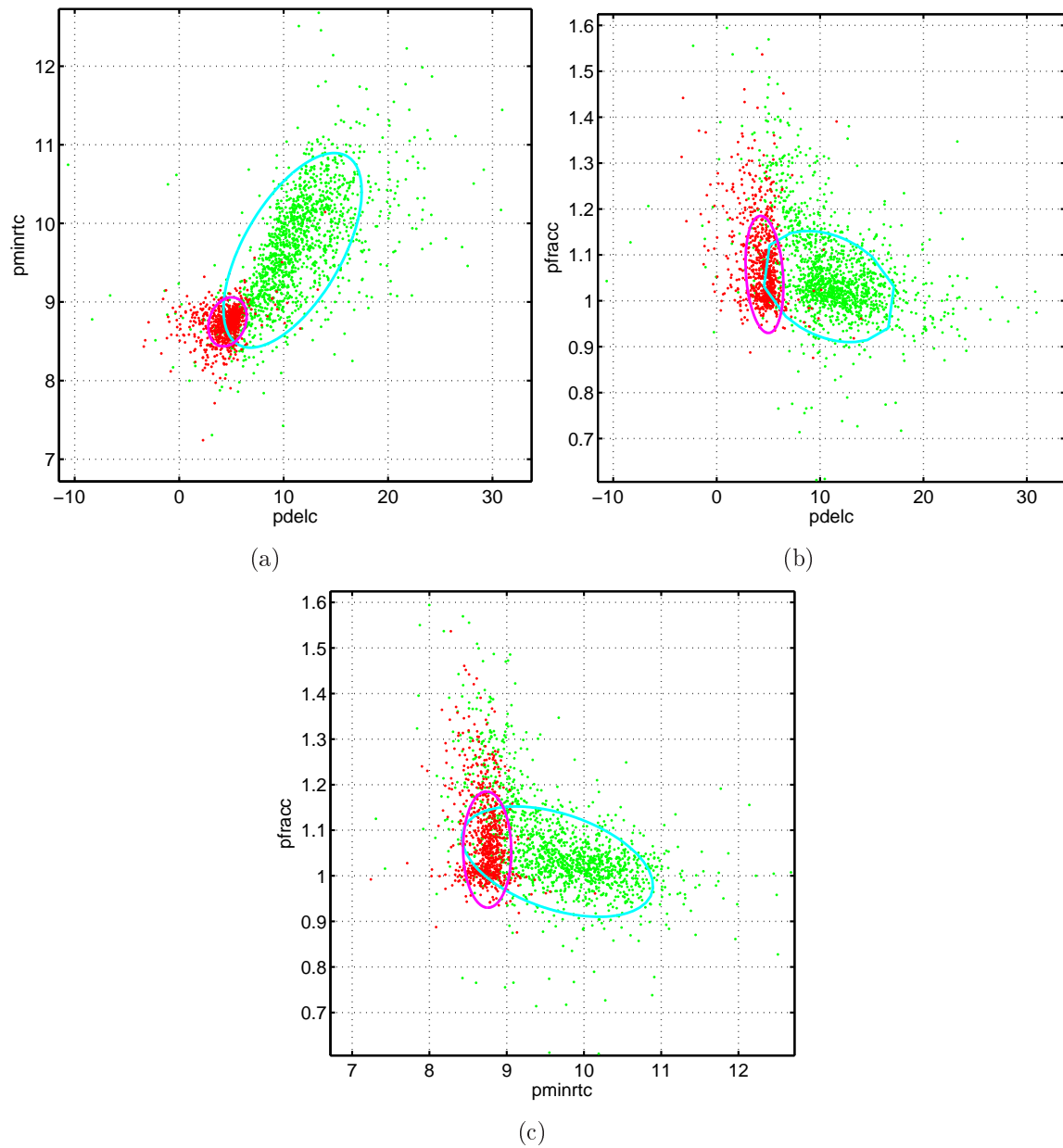
Figure 9.8: Correlations between the three parameters defining the simple chi-squared cut. The neutrons (green) and betas (red) are surrounded by ellipses representing the 2σ contours from the covariance matrix. The neutrons have a broader distribution in delay and rise time, with strong correlation, because of their more pronounced sensitivity to depth.
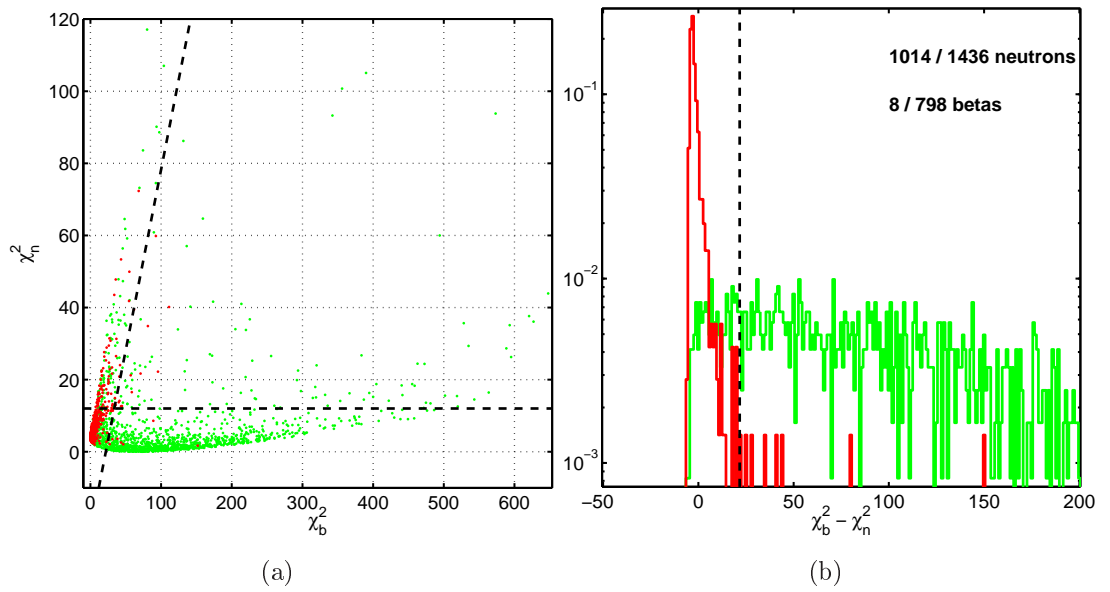
Figure 9.9: Setting the simple chi-squared cut in T1 Z5. The location of betas (red) and neutrons (green) in the $\chi_b^2 - \chi_n^2$ plane is shown in (a) along with the position of the neutron consistency cut ($\chi_n^2 \leq 12$) and the principal beta rejection cut ($\chi_b^2 - \chi_n^2 \geq 21.69$). The setting of the cut value is as shown in (b), to accept 8 betas while retaining a large majority of neutrons.
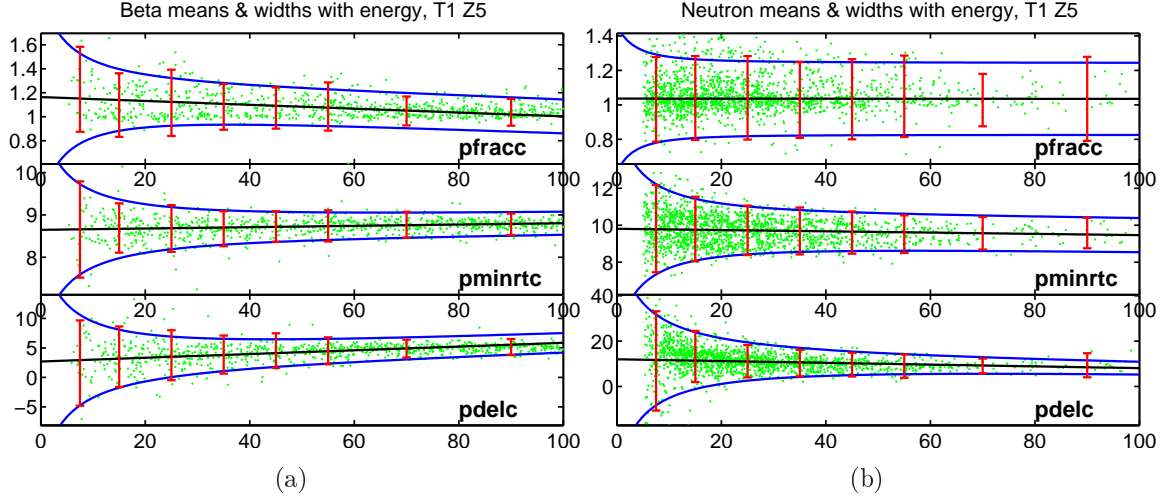
Figure 9.10: Energy-dependent means and widths of timing parameters for chi-squared timing cut in T1 Z5. Betas are shown in (a), and neutrons in (b).

9.5.5   Energy-dependent chi-squared cut

A more sophisticated chi-squared cut allows the means, standard deviations, and covariance to vary with energy. A cut of this form was first demonstrated in the Run 119 analysis, and has been slightly modified here. The cut parameter is a 3-d chi-squared for which $\mu_i(E)$, $\sigma_i(E)$, and $C_{ij}(E)$ are functions of energy, with appropriately selected functional forms and fitted parameters. (In the Run 119 analysis, the means μ_i were not allowed to vary with energy). The energy dependence has been implemented in the following way:

$$
\begin{aligned}
\mu_i(E) &= A_i + B_i \times E \\
\sigma_i^2(E) &= D_i + \frac{F_i}{(E + 5\ \mathrm{keV})^2} \\
C_{ij}(E) &= G_{ij}\sigma_i(E)\sigma_j(E).
\end{aligned}
$$

Fits to these forms for the neutrons and betas in one detector are shown in Fig. 9.10. The 5-keV offset in the denominator of the variance ensures that the acceptances do not "blow up," approaching infinite width at low energies. The expression for covariance is equivalent to assuming a constant, energy-independent correlation. The
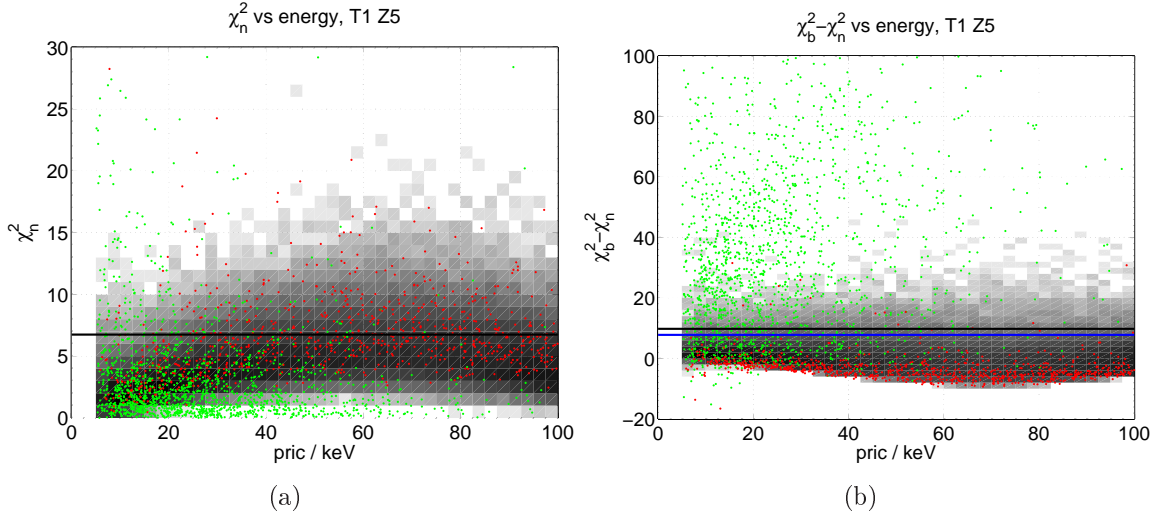
Figure 9.11: Performance of chi-squared timing parameters over energy in T1 Z5. The chi-squared relative to neutrons is shown in plot (a), and the chi-squared difference in plot (b). Bulk gamma events are indicated by the grayscale histogram, neutrons by green (light) points, and betas by red (dark) points. The horizontal line in (a) shows the neutron consistency cut level, set at 6.75σ. The upper horizontal line in (b) represents the cut defined to accept exactly 8 betas in this detector, while the lower (blue) horizontal line is the "coadded" cut level.

energy dependence of μ_i is generally very small, but non-negligible in some cases. Once the energy-dependent means and covariances have been defined, the chi-squared cut is defined exactly as in Eq. 9.1.

The cut level can be set detector-by-detector, as in the simple timing cut, but the energy-dependent version makes the performance of the cut consistent enough across detectors that it is also possible to set a single, common cut level. This variation will be referred to as a "coadded" cut. All parameters and the chi-squared values themselves are identical for the detector-by-detector and coadded cuts, so that they differ only in the cut thresholds. In the coadded cuts, it remains necessary to treat silicon and germanium separately, so we use one coadded cut threshold for all good Si detectors, and another for all good Ge detectors. The neutron consistency cut is set at $\frac{\chi_n^2}{\text{d.o.f.}} \leq 2.25$, a tighter cut than in the simple chi-squared cut, but sufficient to keep the vast majority of the neutrons. The performance of the χ_n^2 and $\chi_b^2 - \chi_n^2$

parameters, along with the coadded and detector-by-detector cut values, are shown in Fig. 9.11.

Since the energy-dependent chi-squared cut gives very good neutron acceptance while maintaining beta rejection, we have chosen to use this type of cut for our primary timing cut. In order to appropriately tune the expected leakage, we have varied the cut in three independent ways, giving a total of eight energy-dependent chi-squared timing cuts.

Det-by-det and coadded

The cuts can be defined detector-by-detector, or coadded over all Ge and over all Si detectors. The coadded version is generally less sensitive to the positions of the last few outliers than the detector-by-detector version, and less fine-tuned to the odd Ba calibration data set. Therefore, the coadded approach is preferred for the final timing cut.

Tight and loose

We have defined the cuts with the looser leakage allowance ($a = 8$ for Ge, $a = 4$ for Si) as well as a tighter variant ($a = 3$ for Ge, $a = 2$ for Si). The final decision over which version to use is made after the expected leakages and neutron acceptances are calculated in the following sections.

"Classic" beta definition

The beta selection cut defined in Sec. 9.4.1 includes a special crosstalk-event rejection cut that is new to this analysis. For comparison, we have also defined "classic" cuts using the same beta selection in yield, but without the crosstalk rejection. Since this includes some events with slower, bulk-like timing, the result is generally a tighter timing cut that has a lower acceptance of gammas. The decision about whether to use the new or "classic" beta definition is postponed until after the expected leakage and neutron efficiency calculations of the following sections.

9.6 Surface event passage fractions

There are several independent samples for estimating the fraction of surface events that will leak past the timing cuts. One sample is the "even" barium calibration set, as described in Sec. 9.4.2, which was not used for setting the cuts. The second sample comes from multiple scatter events in the WIMP search data, inside the nuclear recoil yield band. The third sample is WIMP-search multiple scatters in the wide beta region.[2] The passage fractions have been calculated separately for events on the phonon side and the charge side, tagged by energy in the neighboring detectors. The results are given in Table 9.3 (Ge) and Table 9.4 (Si).

9.7 Phonon-side fraction

Betas on the top (phonon) side of the detector and on the bottom (charge) side have very different properties in ionization yield and in timing. In Ge detectors, the betas on the phonon side have lower yield and are more likely to leak into the nuclear recoil band, but are much easier to reject with the timing cut. Betas on the charge side, on the other hand, have higher yield and rarely show up in the nuclear recoil band, but have phonon pulse timing less distinct from bulk events. (In Si detectors, the difference is less pronounced). These two populations are therefore studied separately for some purposes, such as leakage estimates. The "phonon-side fraction" f_P is the fraction of surface events *in the nuclear recoil band* that are on the phonon side. This is different from the fraction of all surface scatters that are on the phonon side, because very few surface events on the charge side fall within the nuclear recoil band.

There are several ways to go about calculating f_P. One way would be to select low-yield nearest-neighbor multiple-scatter events in low background or barium calibration data, imposing the standard yield-band cut, and find the fraction that are on the phonon side. However, this estimate can be skewed because a significant number

---

[2]Because this includes the nuclear recoil region, the second and third samples are not entirely statistically independent. However, because most of the wide betas in the WIMP search data are outside the nuclear recoil band, the correlation is small.

Cut type	W.S. mult. wide β		W.S. mult. NR band		Even Ba	
	P side	Q side	P side	Q side	P side	Q side
2d simple	0.52 ± 0.45	5.00 ± 1.30	0.77 ± 1.08	16.36 ± 5.22	0.17 ± 0.11	1.38 ± 0.51
2d r-corr	0.79 ± 0.52	6.33 ± 1.44	1.54 ± 1.32	12.73 ± 4.77	0.28 ± 0.14	1.03 ± 0.45
χ^2 simple	0.79 ± 0.52	6.00 ± 1.41	1.54 ± 1.32	18.18 ± 5.42	0.45 ± 0.17	0.45 ± 0.17
χ^2 r-corr	1.57 ± 0.69	6.67 ± 1.47	2.31 ± 1.52	18.18 ± 5.42	1.11 ± 0.25	1.90 ± 0.59
χ^2 E-dep	0.00 ± 0.26	4.67 ± 1.26	0.00 ± 0.77	10.91 ± 4.51	0.67 ± 0.20	1.03 ± 0.45
" coadded	0.52 ± 0.45	4.67 ± 1.26	0.77 ± 1.08	7.27 ± 3.89	0.56 ± 0.18	1.03 ± 0.45
χ^2 E-dep tight	0.00 ± 0.26	2.33 ± 0.93	0.00 ± 0.77	3.64 ± 3.07	0.22 ± 0.12	0.86 ± 0.42
" coadded	0.00 ± 0.26	1.33 ± 0.74	0.00 ± 0.77	0.00 ± 1.82	0.17 ± 0.11	0.69 ± 0.38
χ^2 E-dep classic	0.00 ± 0.26	2.67 ± 0.99	0.00 ± 0.77	7.27 ± 3.89	0.33 ± 0.15	0.69 ± 0.38
" coadded	*0.00 ± 0.26*	*1.33 ± 0.74*	*0.00 ± 0.77*	*1.82 ± 2.54*	*0.17 ± 0.11*	*0.34 ± 0.30*
χ^2 E-dep cl. tight	0.00 ± 0.26	2.00 ± 0.87	0.00 ± 0.77	1.82 ± 2.54	0.22 ± 0.12	0.00 ± 0.17
" coadded	0.00 ± 0.26	0.33 ± 0.47	0.00 ± 0.77	0.00 ± 1.82	0.11 ± 0.10	0.00 ± 0.17

Table 9.3: Passage fraction estimates for coadded Ge detectors, in percent. The uncertainties are Poisson errors given the statistics of each beta sample. The primary cut for Ge is the energy-dependent "classic" coadded cut, shown in italics.

Cut type	W.S. mult. wide β		W.S. mult. NR band		Even Ba	
	P side	Q side	P side	Q side	P side	Q side
2d simple	2.76 ± 1.33	0.00 ± 1.61	0.00 ± 1.61	0.00 ± 2.08	4.49 ± 1.13	0.41 ± 0.58
2d r-corr	2.76 ± 1.33	0.00 ± 1.61	3.51 ± 2.97	0.00 ± 2.08	3.37 ± 0.99	0.82 ± 0.71
χ^2 simple	4.97 ± 1.70	0.00 ± 1.61	8.77 ± 4.08	0.00 ± 2.08	3.65 ± 1.03	0.41 ± 0.58
χ^2 r-corr	4.42 ± 1.62	0.00 ± 1.61	8.77 ± 4.08	2.08 ± 2.90	2.81 ± 0.92	0.82 ± 0.71
χ^2 E-dep	4.97 ± 1.70	0.00 ± 1.61	3.51 ± 2.97	0.00 ± 2.08	4.21 ± 1.10	0.41 ± 0.58
” coadded	3.87 ± 1.53	0.00 ± 1.61	1.75 ± 2.45	0.00 ± 2.08	3.93 ± 1.06	0.41 ± 0.58
χ^2 E-dep tight	3.87 ± 1.53	0.00 ± 1.61	3.51 ± 2.97	0.00 ± 2.08	2.53 ± 0.88	0.41 ± 0.58
” coadded	1.66 ± 1.09	0.00 ± 1.61	0.00 ± 1.75	0.00 ± 2.08	2.53 ± 0.88	0.41 ± 0.58
χ^2 E-dep classic	3.87 ± 1.53	0.00 ± 1.61	1.75 ± 2.45	0.00 ± 2.08	3.65 ± 1.03	0.41 ± 0.58
” coadded	1.66 ± 1.09	0.00 ± 1.61	0.00 ± 1.75	0.00 ± 2.08	3.37 ± 0.99	0.41 ± 0.58
χ^2 E-dep cl. tight	1.66 ± 1.09	0.00 ± 1.61	0.00 ± 1.75	0.00 ± 2.08	1.97 ± 0.79	0.41 ± 0.58
” coadded	*1.66 ± 1.09*	*0.00 ± 1.61*	*0.00 ± 1.75*	*0.00 ± 2.08*	*2.25 ± 0.83*	*0.41 ± 0.58*

Table 9.4: Passage fraction estimates for coadded Si detectors, in percent. The uncertainties are Poisson errors given the statistics of each beta sample. The primary cut for Si is the energy-dependent coadded "classic" tight cut, shown in italics.

of the charge-side multiples with very low yield may be compromised by detector-detector crosstalk effects, as described in Secs. 5.5 and 9.4.1. It is possible to remove these events with the anti-crosstalk cut, but this also removes some real betas on the charge side, again artificially shifting the result. Therefore, we take a very different approach. We assume that the total number of surface scatters on each face of the crystal is the same. Then, we find the fraction of the betas on each face that fall within the nuclear recoil yield band, designated PNR_P and PNR_Q for the phonon side and the charge side respectively. The selection of betas on the charge side includes the anti-crosstalk cut, which reduces the overall statistics slightly, but should allow us to determine the true value of PNR_Q. The number of betas falling within the nuclear recoil band on each side and the phonon-side fraction are given by

$$
\begin{aligned}
NNR_P &= N\beta_P \times PNR_P \\
NNR_Q &= N\beta_Q \times PNR_Q \\
f_P &= NNR_P/(NNR_P + NNR_Q).
\end{aligned}
$$

With the assumption that the intrinsic beta rate on each side is the same, $N\beta_P = N\beta_Q$, this gives

$$
f_P = \frac{1}{1 + PNR_Q/PNR_P}. \tag{9.2}
$$

This calculation cannot be performed for top or bottom detectors. (The only good detector in the top or bottom position within a tower is T2 Z6.) The beta sample used is that of the even Ba data set with standard quality and beta selection cuts. The results are shown in Table 9.5.

Detector	P side		Q side		f_P
T1 Z1	-		29/93	31.2%	-
T1 Z2	40/306	13.1%	1/68	1.5%	0.8989±0.13
T1 Z3	95/464	20.5%	2/138	1.4%	0.9339±0.05
T1 Z4	32/182	17.6%	26/98	26.5%	0.3986±0.06
T1 Z5	79/420	18.8%	2/166	1.2%	0.9398±0.05
T1 Z6	65/337	19.3%	-		-
T2 Z1	-		25/174	14.4%	-
T2 Z2	13/59	22.0%	15/98	15.3%	0.5901±0.09
T2 Z3	84/415	20.2%	3/127	2.4%	0.8955±0.06
T2 Z4	3/44	6.8%	15/48	31.3%	0.1791±0.10
T2 Z5	47/192	24.5%	7/81	8.6%	0.7391±0.08
T2 Z6	7/71	9.9%	-		-
Coadd good Ge	345/1797	19.2±0.9%	15/580	2.6±0.7%	0.8813±0.028
Coadd good Si	55/356	15.5±1.9%	56/244	23.0±2.7%	0.4023±0.041

Table 9.5: Calculation of f_P from even Ba data.

9.8 Expected beta leakage rate

The timing-cut passage fractions are calculated separately for events on the phonon side and the charge side. We define

$$
\begin{aligned}
f_P &= \text{of events in nuclear recoil band, fraction that are on the P side} \\
P_P &= \text{timing cut passage fraction for events on the P side} \\
P_Q &= \text{timing cut passage fraction for events on the Q side} \\
N_{NRSS} &= \text{number of NR-band single scatters before timing cuts} \\
N_{leak} &= \text{expected leakage after timing cuts.}
\end{aligned}
$$

The passage fractions P_P and P_Q have been calculated from three different samples of events in 9.6, and the phonon-side fraction f_P has been calculated in Sec. 9.7. These two can be combined to estimate the overall passage fraction,

$$N_{leak}/N_{NRSS} = f_P P_P + (1 - f_P)P_Q. \tag{9.3}$$

Cut type	W.S. mult. β	W.S. mult. NR	even Ba *beta*
2d simple	1.06 ± 1.22	2.62 ± 4.72	0.31 ± 0.46
2d r-corr	1.45 ± 1.36	2.87 ± 4.37	0.37 ± 0.42
χ^2 simple	1.41 ± 1.33	3.52 ± 4.94	0.58 ± 0.50
χ^2 r-corr	2.18 ± 1.44	4.20 ± 4.98	1.21 ± 0.57
χ^2 E-dep	0.56 ± 1.14	1.30 ± 4.04	0.71 ± 0.44
" coadded	1.02 ± 1.18	1.54 ± 3.56	0.61 ± 0.43
χ^2 E-dep tight	0.28 ± 0.85	0.43 ± 2.79	0.30 ± 0.39
" coadded	0.16 ± 0.69	0.00 ± 1.74	0.23 ± 0.35
χ^2 E-dep classic	0.32 ± 0.90	0.87 ± 3.50	0.38 ± 0.36
" coadded	*0.16 ± 0.69*	*0.22 ± 2.34*	*0.19 ± 0.28*
χ^2 E-dep cl. tight	0.24 ± 0.80	0.22 ± 2.34	0.20 ± 0.19
" coadded	0.04 ± 0.47	0.00 ± 1.74	0.10 ± 0.17

Table 9.6: Overall passage fraction for coadded Ge detectors. Passage fractions are in percent, or equivalently, expected leakage events per 100 candidates pre-timing cuts. The primary cut for Ge is the energy-dependent "classic" chi-squared cut, shown in italics.

Cut type	W.S. mult. β	W.S. mult. NR	even Ba $beta$
2d simple	1.11 ± 0.85	2.12 ± 1.62	2.05 ± 0.54
2d r-corr	1.11 ± 0.85	1.41 ± 1.46	1.85 ± 0.50
χ^2 simple	2.00 ± 0.96	3.53 ± 1.88	1.71 ± 0.49
χ^2 r-corr	1.78 ± 0.94	4.77 ± 2.03	1.62 ± 0.47
χ^2 E-dep	2.00 ± 0.96	1.41 ± 1.46	1.94 ± 0.52
” coadded	1.55 ± 0.91	0.71 ± 1.29	1.83 ± 0.51
χ^2 E-dep tight	1.55 ± 0.91	1.41 ± 1.46	1.26 ± 0.43
” coadded	0.67 ± 0.79	0.00 ± 1.09	1.26 ± 0.43
χ^2 E-dep classic	1.55 ± 0.91	0.71 ± 1.29	1.71 ± 0.49
” coadded	0.67 ± 0.79	0.00 ± 1.09	1.60 ± 0.48
χ^2 E-dep cl. tight	0.67 ± 0.79	0.00 ± 1.09	1.04 ± 0.40
” coadded	*0.67 ± 0.79*	*0.00 ± 1.09*	*1.15 ± 0.41*

Table 9.7: Overall passage fraction for coadded Si detectors. Passage fractions are in percent, or equivalently, expected leakage events per 100 candidates pre-timing cuts. The primary cut for Si is the energy-dependent "classic" tight chi-squared cut, shown in italics.

The results of this calculation are given in Tables 9.6 and 9.7 for Ge and Si, respectively. They are listed as leakage fractions in percent, or equivalently, the expected number of events leaked per 100 candidates pre-timing cuts.

In the entire WIMP-search data set, we expect on the order of 50 events in Ge and 50 in Si before timing cuts. We designate as *primary* the cuts for Ge and Si that give the best exposure while still keeping well under one event expected leakage. We choose for Ge the energy-dependent, "classic" chi-squared cut, and for Si the energy-dependent, tight, "classic" chi-squared timing cut. We estimate systematic errors on the passage fractions by finding the scatter between three variations on the passage rate: those for the coadded and detector-by-detector versions of the same timing cut, in WIMP-search multiple betas, WIMP-search multiple nuclear-recoil band events, and even Ba betas. The final result is a predicted leakage of $0.16 \pm 0.7 \pm 0.3$ per 100 in Ge and $0.67 \pm 0.8 \pm 0.5$ per 100 in Si.

9.9 Cut efficiency

We find the efficiencies of the timing cuts from the acceptance fraction of californium calibration neutrons. All timing cuts used here apply the "cliff cut" for T2 Z5, and it is included when calculating the efficiencies. The first four timing cuts also include the position reconstruction consistency cut, $0 \leq r \leq 4$ cm. The efficiencies are calculated in bins by energy, and then fitted to the functional form

$$\epsilon(E) = a \times \left(1 - \frac{b}{(E - c)^d} \right) \times (1 + fE).$$
(9.4)

The first factor in parentheses parameterizes the rise in efficiency from zero at the lowest energies to near maximum efficiency by about 40 keV. The second factor in parentheses allows for a small fall-off in efficiency at high energies in some detectors, in which case $f < 0$. This is the result of narrowing acceptance bands in the case of the energy-dependent chi-squared cuts; in other cuts it may occur because the timing quantities for neutrons shift slightly between low energies, where most of the neutron sample is, and high energies, where the statistics are low.   The efficiencies and fits for
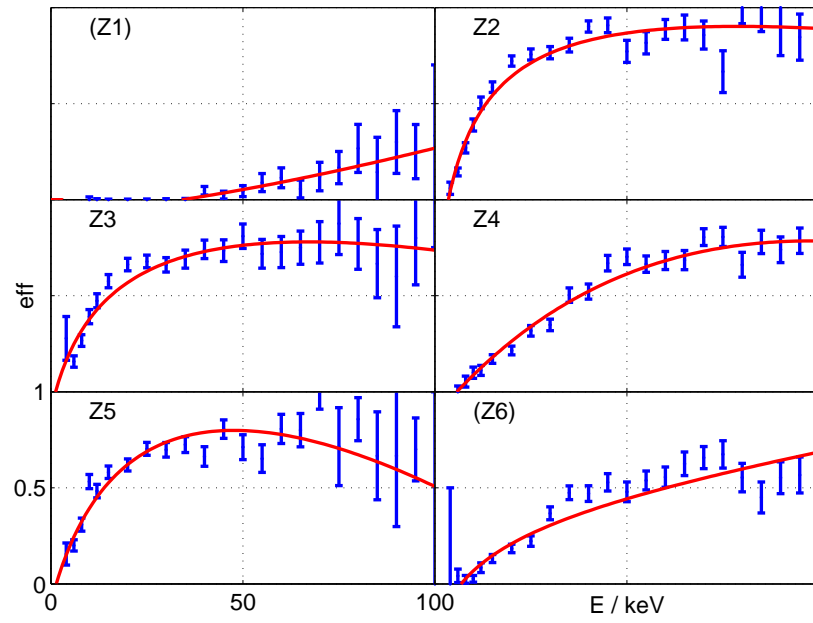
Figure 9.12: Efficiency on neutrons of the energy-dependent, coadded, "classic" chi-squared timing cut, for T1 detectors. Z1 and Z6 are not used, and Z4 has a lower efficiency at low energies.
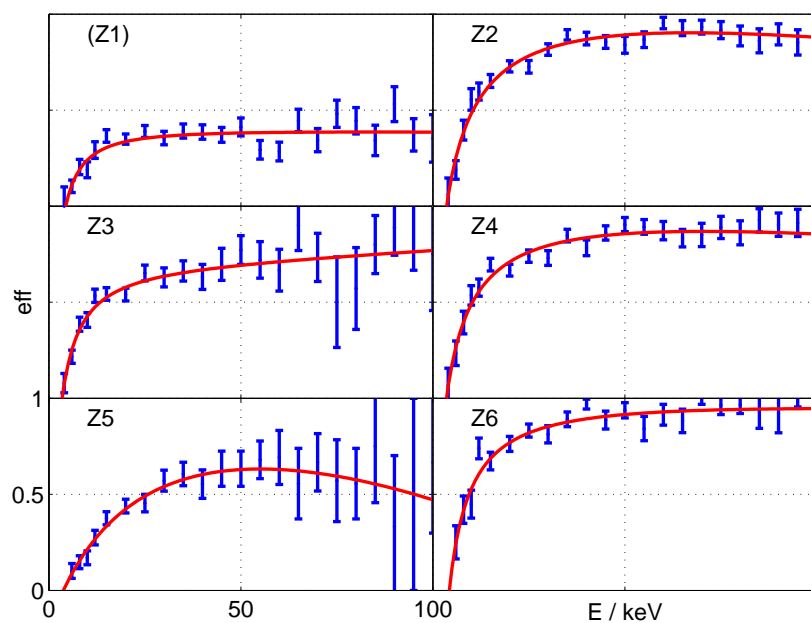
Figure 9.13: Efficiency on neutrons of the energy-dependent, coadded, "classic" chi-squared timing cut, for T2 detectors. Z1 is not used.
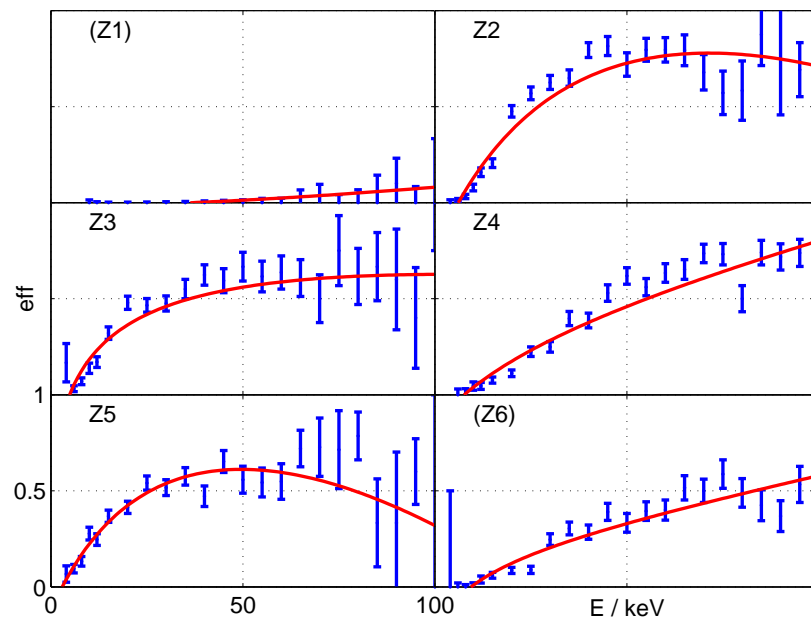
Figure 9.14: Efficiency on neutrons of the tight energy-dependent, coadded, "classic" chi-squared timing cut, for T1 detectors. Z1 and Z6 are not used, and Z4 has a lower efficiency at low energies.
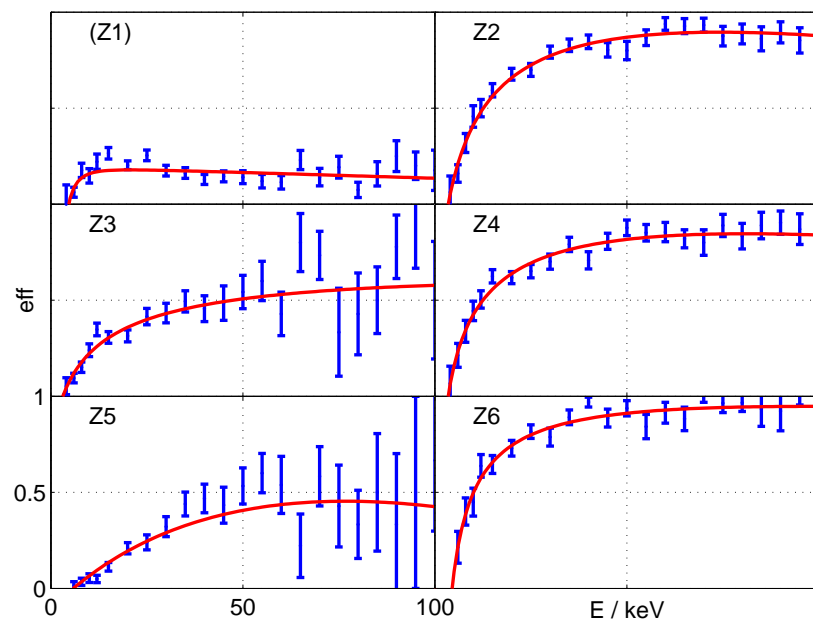
Figure 9.15: Efficiency on neutrons of the tight energy-dependent, coadded, "classic" chi-squared timing cut, for T2 detectors. Z1 is not used.

two of the cuts are shown here for both towers. These are both coadded chi-squared timing cuts, so the cut levels have been set once for good Ge and once for good Si detectors. These cut levels do not usually apply well to the excluded detectors, which behave differently and would require different cut levels if they were to be used. Therefore, the low efficiencies for the excluded detectors are not directly meaningful. Figures 9.12 and 9.13 show the efficiency of the coadded, "classic" energy-dependent chi-squared cut (the primary cut for Ge) on Tower 1 and Tower 2, respectively. Figures 9.14 and 9.15 show the efficiency of the *tight* coadded, "classic" energy-dependent chi-squared cut (the primary cut for Si). These efficiencies are used in the total exposure calculation in the next chapter.

Chapter 10

WIMP-search results

The base analysis described in Chapter 8, along with the choice of surface-event rejection cuts from Chapter 9, constitute a complete prescription for selecting WIMP candidate events and rejecting backgrounds. The entire work up to this point has been carried out exactly as if blinded, although it was not formally a blind analysis since it is not the first analysis of the Run 118 and Run 119 data sets. This chapter "opens the box" to reveal the candidate events and calculate dark matter limits.

10.1 Expected beta leakage

The beta leakage estimates are calculated as a final step just before unblinding. Although preliminary calculations are made using WIMP-search events outside the signal region, the final estimates make use of the number of events inside the signal region before the application of timing cuts. This includes any events that will become WIMP candidates, but is dominated by surface events that do not pass the timing cuts. Therefore, it does not significantly reveal information about the WIMP candidates. Because they are included, however, this step is not performed until all cuts are completely frozen, and just before unblinding. The number of candidate events N_{NRSS} (pre-timing cuts) is multiplied by the estimated timing-cut passage

248

	N_{NRSS}	expected leakage
Ge, 5-10 keV	30	$0.05 \pm 0.2 \pm 0.09$
Ge, 10-100 keV	26	$0.04 \pm 0.2 \pm 0.08$
Si, 5-10 keV	24	$0.2 \pm 0.2 \pm 0.1$
Si, 10-100 keV	40	$0.3 \pm 0.3 \pm 0.2$

Table 10.1: Expected beta leakage for Ge and Si, in the low-energy range and the 10-100 keV range.

fraction P_T to give the expected leakage:

$$N_{leak} = N_{NRSS} \times P_T,$$

where P_T is taken from Table 9.6 and 9.7. We calculate the expected leakage separately for four cases: Ge, 10-100 keV (the case of greatest interest for WIMP detection); Ge, 5-10 keV; Si, 10-100 keV; and Si, 5-10 keV. The raw counts before timing cuts and the inferred expected leakages are shown in Table 10.1 for the primary timing cuts in Si and Ge. These expected leakages are acceptably low, predicting a very likely outcome of zero background events altogether.

This approach depends on the assumption that the sample of events used to calculate the passage fractions is representative of the WIMP-search singles events. This assumption fails whenever the candidate events preferentially occur in detectors or energy ranges where rejection is less effective. In that case, the passage fraction and event count must be calculated as a function of energy, or separately for each detector. As will be seen after unblinding, the assumption was not viable for this data set, and the number of final candidate events was not consistent with the expected leakages.

10.2   Exposure

The WIMP-search exposure is calculated from the mass and live time for each detector, the efficiency of all cuts up to timing cuts, and the timing cut acceptance:
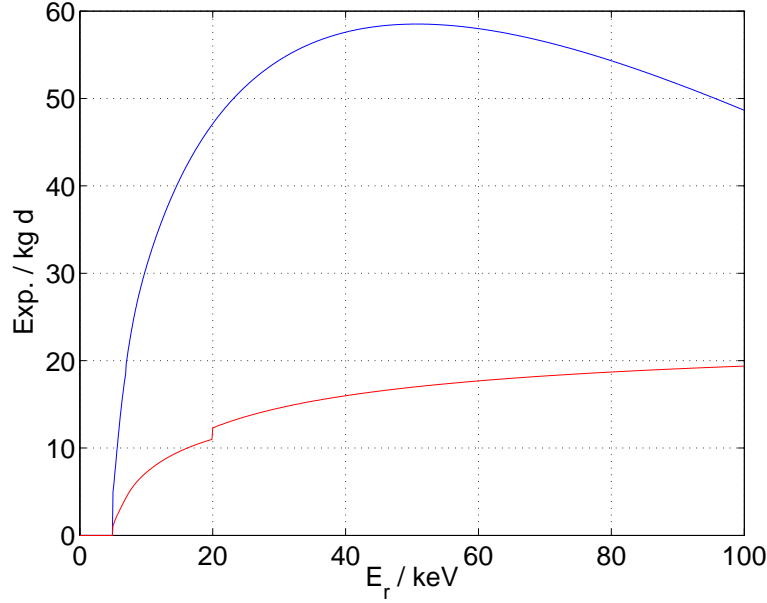
Figure 10.1: Total exposure from Run 118 + Run 119 combined data set after all cuts. The blue (upper) curve shows Ge exposure, and the lower (red) curve shows Si exposure.

$$\text{Exposure} = \sum_i L_i M_i \epsilon_i \epsilon_i^t. \tag{10.1}$$

The live times for each detector are given in Table 8.2. The mass M_i is 250 g per Ge detector, or 100 g per Si detector. The pre-timing efficiencies ϵ_i are as shown in Figs. 8.22 and 8.23, and the timing cut acceptances ϵ_i^t are as in Figs. 9.12 and 9.13 (primary cut for Ge), or Figs. 9.14 and 9.15 (primary cut for Si). We calculate a summed exposure for all good Ge detectors and a second exposure for Si detectors, as shown in Fig. 10.1.

10.3   Candidate events

Tables 10.2 and 10.3 give the number of events remaining after each cut is applied to the Ge and Si WIMP-search data, respectively; the last several steps are also shown on a yield-energy plot in Figs. 10.2 (Ge) and 10.3 (Si). After all cuts we find four
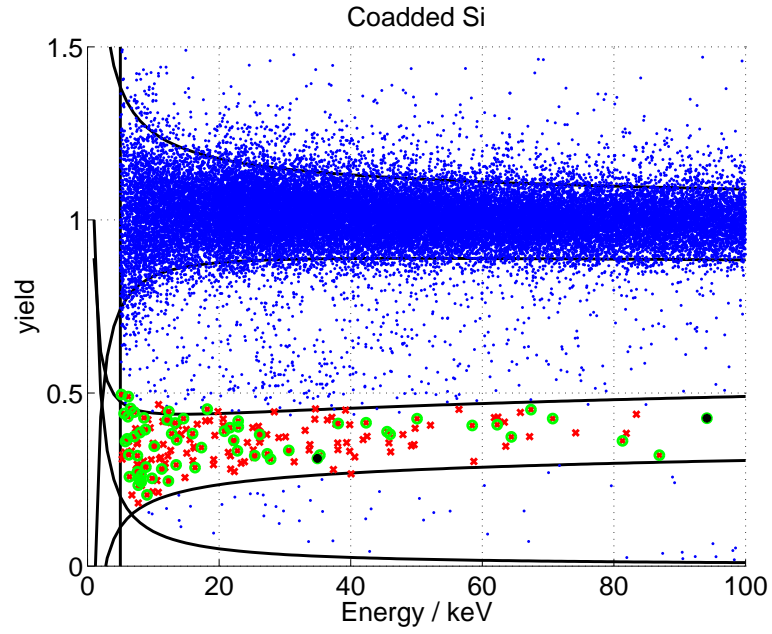
Figure 10.2: Unblinding the Ge data set. The symbols show the effect of imposing successive analysis cuts: blue points are events that pass good event selection, fiducial volume, and energy threshold cuts; red Xs also pass the nuclear recoil band cut; green circles also pass the single-scatter cut; and black dots are the two candidate events passing all cuts.
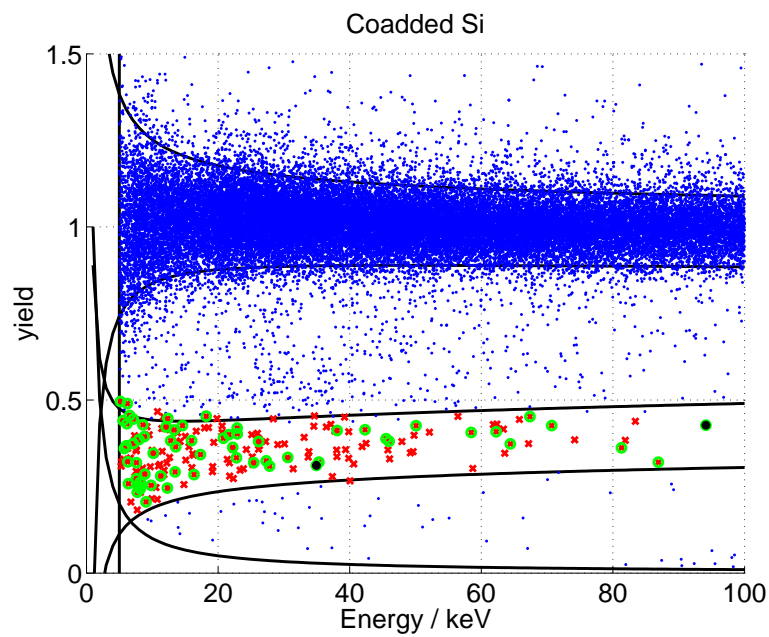
Figure 10.3: Unblinding the Si data set. The symbols show the effect of imposing successive analysis cuts: blue points are events that pass good event selection, fiducial volume, and energy threshold cuts; red Xs also pass the nuclear recoil band cut; green circles also pass the single-scatter cut; and black dots are the two candidate events passing all cuts.

Cut	Events selected
Total events	2,547,622
Good data	1,790,392
Det trigger	1,054,440
Noise cuts	1,047,807
Veto cut	1,014,181
Thresholds	72,284
Inner electrode	49,014
Yield band	278
Singles	56
Timing	2

Table 10.2: Events remaining after cuts are progressively applied for all Ge detectors.

candidate events, two in Ge and two in Si. The two in Ge are at very low energy, 5.3 and 7.3 keV, and do not have a large effect on the resulting limits since both are near threshold. The two in Si are at higher energy and cause the WIMP exclusion limits to be somewhat weakened. The four events have several interesting properties in common. First, all are from the Run 119 portion of the data set, with none from Run 118. This is somewhat surprising, since noise performance was better during Run 119, and we have generally regarded it as a better-controlled running period where data quality is concerned. It is less surprising because Run 119 has more exposure than Run 118, by a large factor in the case of Si. The second surprising feature is that none of the four candidate events was a candidate *before the application of timing cuts* in the published Run 119 analysis. Each of the four was rejected for another reason, rather than by the surface event rejection analysis.

10.3.1   T1 Z2 (Ge), 5.3 keV

The first candidate event (Fig. 10.4) is just above threshold in Tower 1 ZIP 2 (Ge), at 5.3 keV. This event was not a candidate in the original published Run 119 analysis because it is below the 10-keV primary analysis threshold, and even below the 7-keV threshold of the low-energy analysis. It has series number 140801_0038 and event number 110053.

Cut	Events selected
Total events	2,547,622
Good data	1,750,248
Det trigger	1,022,494
Noise cuts	1,000,495
Veto cut	968,390
Thresholds	46,067
Inner electrode	34,161
Yield band	181
Singles	64
Timing	2

Table 10.3: Events remaining after cuts are progressively applied for all Si detectors.
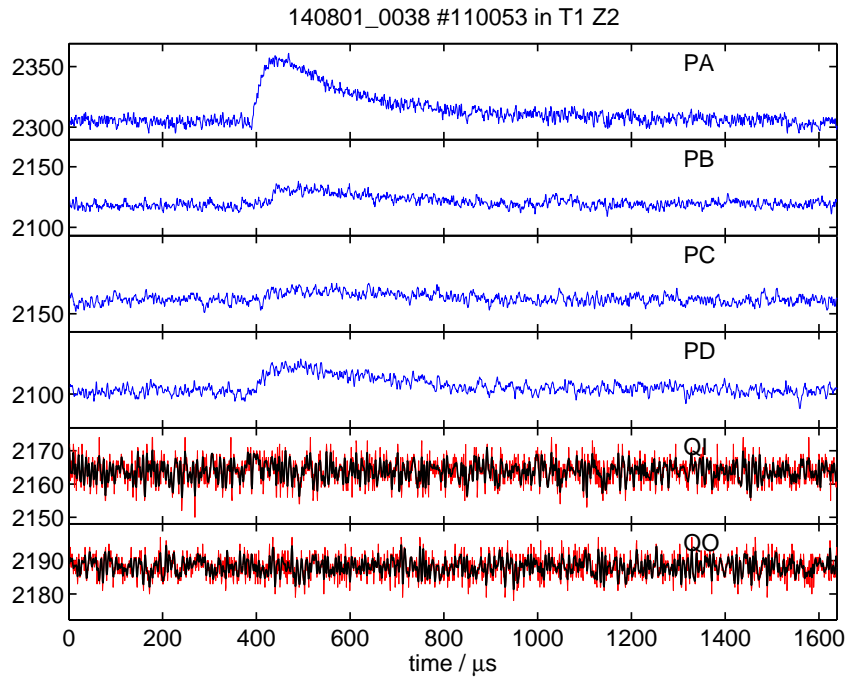
Figure 10.4: Candidate event at 5.3 keV in T1 Z2. In the charge traces, red (light) is the raw trace, and black (dark) shows the trace as filtered for timing. Although the noise in the ionization channel is large, the optimal filter algorithm is still effective in identifying the pulse, since most of the noise is at higher frequencies.
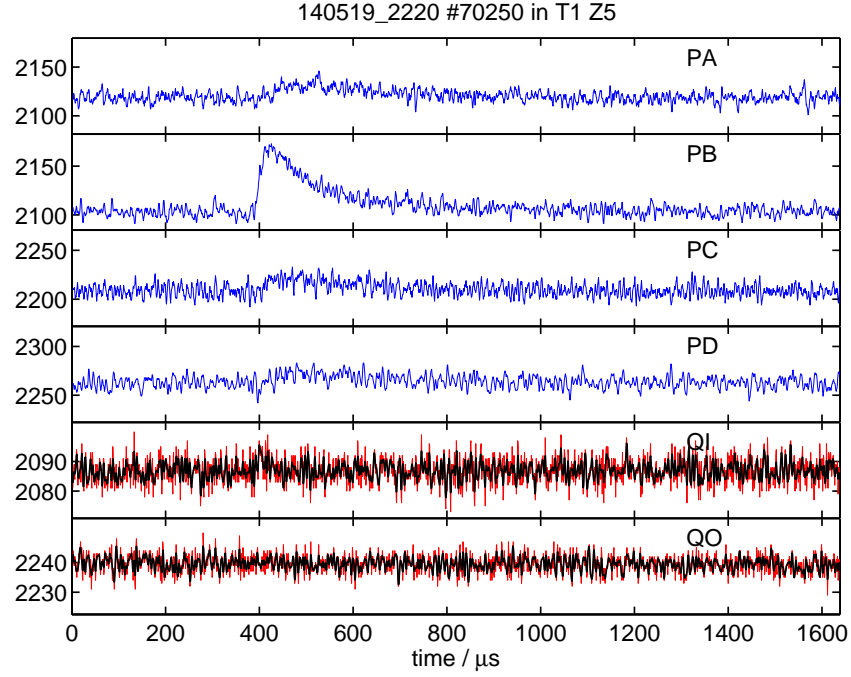
Figure 10.5: Candidate event at 7.3 keV in T1 Z5. In the charge traces, red (light) is the raw trace, and black (dark) shows the trace as filtered for timing.

10.3.2   T1 Z5 (Ge), 7.3 keV

The second candidate is also at low energy, at 7.3 keV in Tower 1 ZIP 5 (Ge). This event failed the ionization chi-squared data quality cut in the original Run 119 analysis. It has series number 140519_2220 and event number 70250. On inspection, it becomes apparent that this candidate should have been rejected as a bad event. Its phonon delay is actually negative! It was able to pass the energy-dependent chi-squared timing cut because of a pathology in the cut definition at very low energies. In hindsight, it is easy to see how to design a quality cut to protect against such events. The timing of this event is shown in Fig. 10.6 along with the inner ionization channel raw trace and the largest phonon trace.
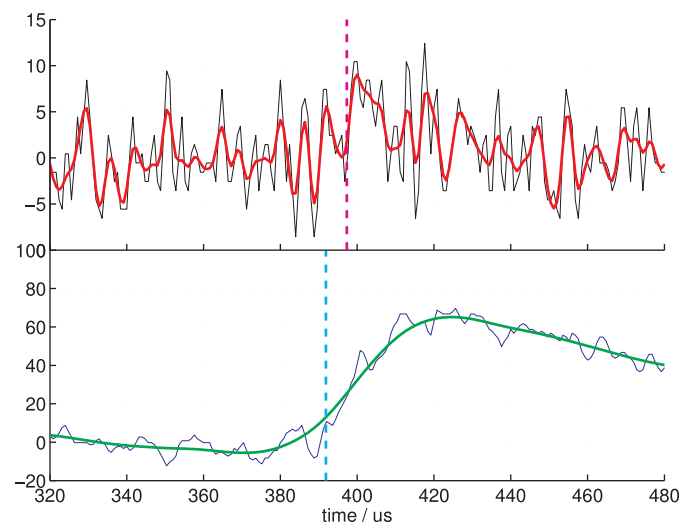
Figure 10.6: Phonon and ionization start times of T1 Z5 candidate event. The phonon start time (bottom pane) is reconstructed well, and is earlier than the ionization start time (top pane). This should never be the case for physical events. The ionization signal is quite small compared to the noise, so it is likely that the calculated start time is not physically meaningful.
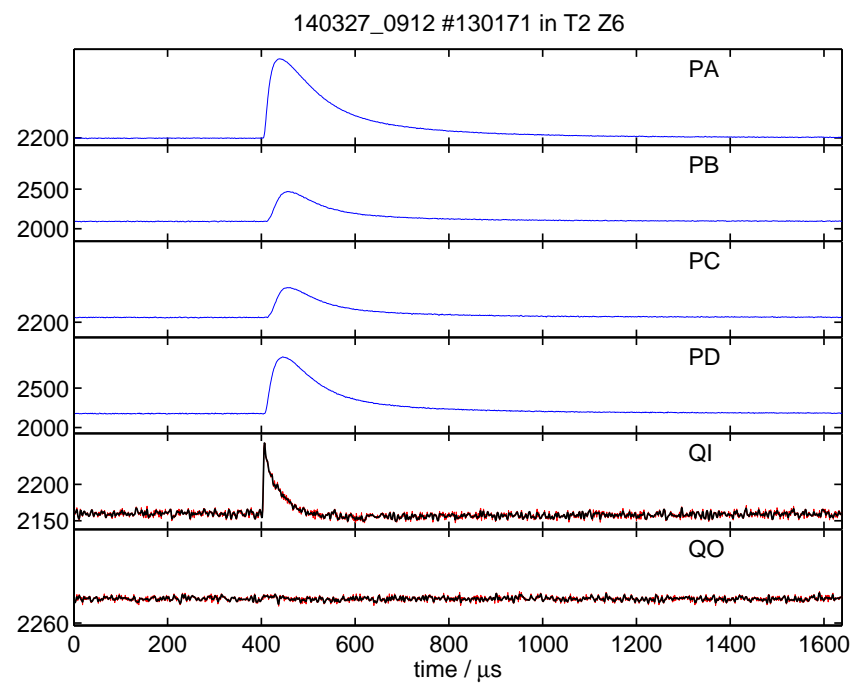
Figure 10.7: Candidate event at 34.9 keV in T2 Z6. In the charge traces, red (light) is the raw trace, and black (dark) shows the trace as filtered for timing.
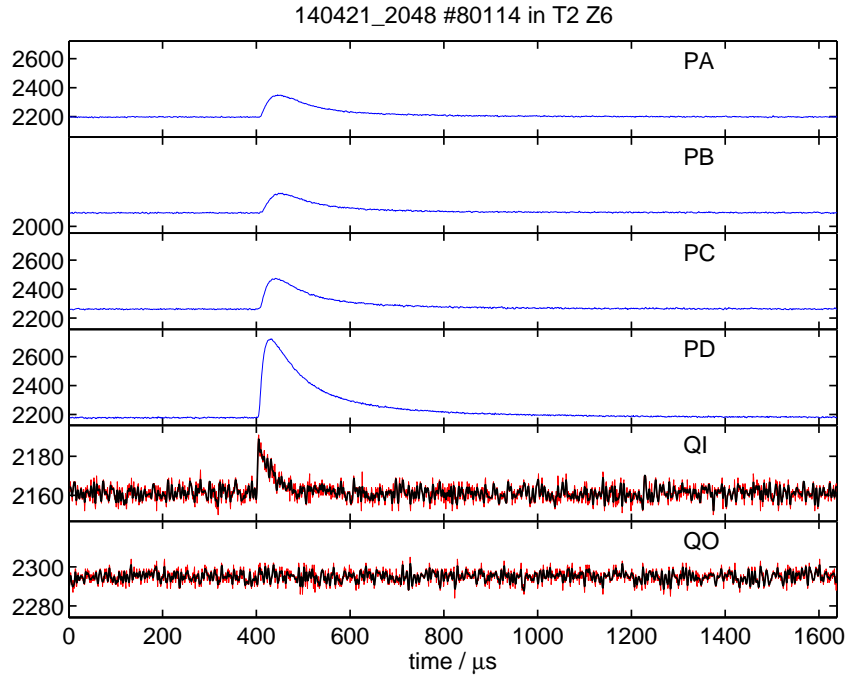
Figure 10.8: Candidate event at 94.1keV in T2 Z6. In the charge traces, red (light) is the raw trace, and black (dark) shows the trace as filtered for timing.

10.3.3   T2 Z6 (Si), 34.9 keV

This event has series number 140421_2048 and event number 80114. It appears at an intermediate energy, 34.9 keV, in the silicon detector T2 Z6. Because it is in the middle of the energy range of interest, this event does have an effect on the WIMP exclusion limits. This candidate failed the single-scatter cut in the original Run 119 analysis, despite the fact that no interaction was seen in any other detector. This strange fact is explained by an error in the published Run 119 analysis, as a result of which a significant number of good events were improperly and unintentionally removed by the single-scatter cut. The origins and effects of this error, which had not been detected previously, are the subject of Appendix A.

10.3.4   T2 Z6 (Si), 94.1 keV

The final candidate is also in T2 Z6 (Si), at an energy of 94.1 keV, just below the 100-keV cutoff. It has series number 140327_0912 and event number 130171. Because this event is so close to the upper end of the energy range, it has little effect on the WIMP search results. It was excluded in the original Run 119 analysis because of a problem with triggers in another detector, T2 Z3. In the current analysis this problem was also caught, but the data set was retained. It was considered a bad run for T2 Z3 only, since there was no evidence that any other detector was affected. The trigger problem in T2 Z3 appears to be entirely unrelated to the appearance of a candidate event in T2 Z6.

10.4   Interpretation

All four of the candidate events are assumed to be true WIMP scatters for purposes of setting exclusion limits, in accordance with the principles of blind analysis. However, it is also valuable to further study the events in hindsight, to understand what they are and why the actual number of events is higher than the expected leakage.

One of the events, the 7.3-keV event in T1 Z5 (Ge), is clearly a bad event and cannot be interpreted as a true WIMP candidate. Because the pathology exemplified by this event is not very common, but increasingly common at lower energies, it was unlikely to be duplicated in the beta samples. The second Ge event is also at very low energy. The expected leakages were calculated by multiplying the expected number of candidate events before timing cuts by the timing-cut passage fraction of surface events. This formula is correct if the two numbers are uncorrelated. In this case, there is a correlation between energy and passage fraction, since more events will leak through the timing cuts at low energy, and there are also more candidates at low energy. The concentration of events just above threshold in Ge is clearly visible in Fig. 10.2. For this reason, the expected leakage was unrealistically low.

In silicon, neither candidate event is obviously a bad event. Here, too, there is a correlation that was unaccounted for in the expected leakage calculation. A single

detector, T2 Z6, accounted for almost half of the events before the timing cut, and also suffered from a higher beta passage fraction than the other silicon detectors. Because the events were more likely to be in a detector where rejection was worse, which was not accounted for in the expected leakage calculation, the expected leakage was again unrealistically low. Since the end of Run 119, the backgrounds working group of CDMS has studied the correlation between rates of 5.4-MeV alpha particles and surface betas. It appears very likely that the beta rate in WIMP-search running is dominated by ^{210}Pb contamination on the detector surfaces; ^{210}Pb has a 22-year half-life and undergoes beta decay. Its product, ^{210}Bi, has only a 5-day half life, and also beta decays, producing ^{210}Po. This has a 137-day half-life, and produces 5.4-MeV alphas; its decay product is stable ^{206}Pb. The background correlation analysis, counting alphas in one detector with coincident nuclear recoil in the next, finds the detector pair T2 Z5-Z6 to be the hottest pair [100]! Therefore, it is likely that charge-side betas on T2 Z6, which could not be accounted for in the P-side/Q-side analysis of beta leakage rates because of the lack of a neighboring detector, contributed an unusually large beta rate.

10.5   WIMP exclusion limits

Given the energy-dependent overall exposures for the germanium and silicon detectors, and the energies of the four candidate events, we can calculate WIMP exclusion limits (Fig. 10.9). This involves several astrophysical and particle-physics assumptions: we assume a standard isothermal WIMP halo with mean velocity 220 km/s in the vicinity of Earth, and WIMPs with purely spin-independent elastic scattering. The limits are calculated using the optimal interval method developed by Steve Yellin [101]. The germanium limits are stronger than the limits produced by combining the previous Run 118 and Run 119 limits. The silicon limits are weaker at high WIMP mass because of the presence of the candidate event at 34.9 keV. The Run 119 limits shown here for comparison are not identical to those previously published. The sensitivities have been reduced by 17.3% in accordance with the exposure correction detailed in Appendix A.
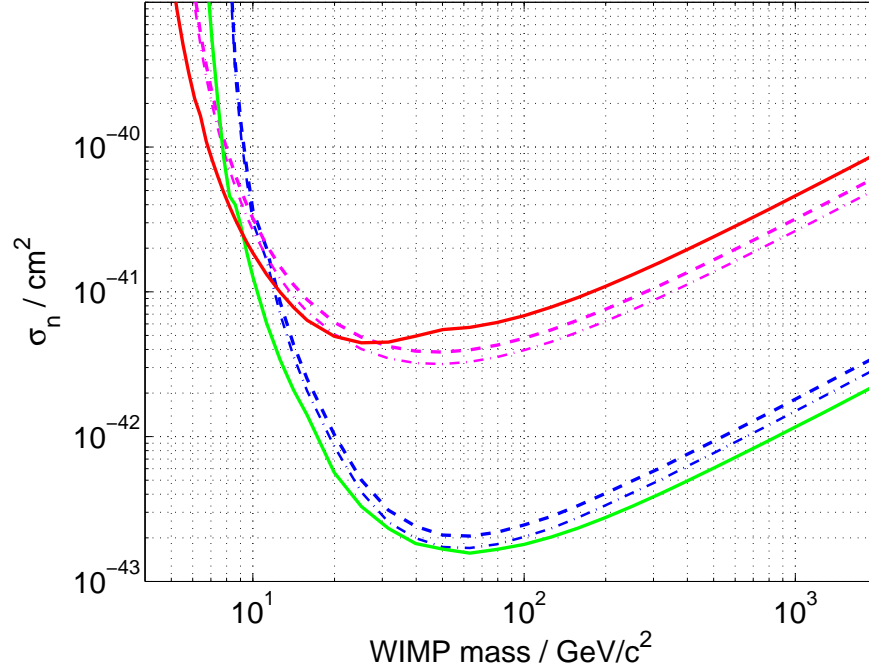
Figure 10.9: New WIMP exclusion limits for the combined analysis. The Ge result (green solid curve) has slightly better sensitivity than the previous analyses, while the Si result (red solid curve) sets a stronger limit only for WIMPs of mass below 30 GeV. For Si, we show the published Run 119 result (dot-dash magenta) and the version corrected for the 17% efficiency overestimate. For Ge, we have not yet calculated the precise magnitude of the correction to the previously published Run 118 + Run 119 combined limits. Therefore, we show the published curve (dot-dash blue) and a second curve uniformly moved up by 17% (dashed blue). The true limit of the previous analysis lies somewhere between these.

10.6   Further developments and future prospects

This reanalysis of the one- and two-tower runs will be published as part of a detailed paper on Run 119. CDMS has not stood still since the end of Run 119, and currently has the leading WIMP sensitivity in the mass range of greatest interest for supersymmetric models. We continue to record data at Soudan, and longer-term planning is underway for SuperCDMS. We briefly mention a few developments:

- Three additional towers of ZIPs were installed in the Soudan icebox in fall of 2004, and after substantial improvements and repairs were made to the cryogenic system, the five towers became fully operational in October 2006. This raises the count to 19 Ge and 11 Si detectors (4.75 kg Ge, 1.1 kg Si). Five-tower data-taking has been entirely within a single cryogenic cooldown, but the data set has been divided into several logical runs. They are separated by short periods of cryogenic maintenance when the detectors warmed above their usual base temperature, but remained well below room temperature.

- The first 5-tower run, Run 123 (October 2006-March 2007) corresponds to 430 kg-d of raw exposure.

- Run 124 (April-July 2007) corresponds to 224 kg-d raw exposure.

- Run 125 (July 2007-January 2008) corresponds to 465 kg-d raw exposure.

- Run 126 (January 2008-end 2008) is ongoing.

- A new WIMP search analysis has been performed on Run 123 and Run 124. This is being prepared for publication, and is now available as a pre-print [102]. The blinded analysis found zero events in Ge 10-100 keV, and sets a limit of $\sigma_n = 6.6 \times 10^{-44}$ cm^2 at $M_{WIMP} = 60$ GeV/c^2. This result is the most sensitive limit for WIMP masses above 42 GeV/c^2.

- Analysis of Run 125 and Run 126, now in progress, will more than double the currently available exposure.

- The SuperCDMS detector development program has successfully demonstrated one-inch-thick detectors, interdigitated double-sided detectors [103], and optimized phonon sensor design for improved timing, energy, and position resolution.

- The initial stage of SuperCDMS is to be a two-supertower demonstration program in the current Soudan facility. We have proposed the construction of a larger Icebox in SNOLab, Sudbury, Ontario, for a 25-kg SuperCDMS Phase A [104]. Further expansion to 125 kg and one ton are envisioned as Phase B and C of SuperCDMS.

10.7   Conclusions

The reanalysis of the one- and two-tower data sets has improved the sensitivity slightly for Ge detectors, improved the analysis techniques and data processing, and made a significant correction to the exposure of the previous two-tower result. The new analysis begins to explore parameter space of interest for supersymmetric WIMP models; this is being improved upon by continuing data-taking with additional detectors, with stiff competition from the Xenon experiment. A current snapshot of the race toward sensitivity to lower cross-sections is shown in Fig. 10.10. The uppermost (blue) curve is the Ge exclusion limit from this thesis analysis. The green curve, which shows the best sensitivity for low WIMP masses, is the first result from Xenon-10, using 10 kg of liquid xenon with nuclear recoil discrimination from ionization and scintillation [106]. Subsequent CDMS results regain the lead for higher WIMP masses: the red curve is the first publication of five-tower data [102], and the lowest (black) curve is the result of combining the five-tower result with the one- and two-tower reanalysis.

Superimposed on the plot are theoretical predictions of the scalar WIMP-nucleon cross-section from supersymmetry. The blue and green (darker) regions are from Rozskowski et al. [107], while the grey (lighter) region is from Baltz and Gondolo [108]. While all such predictions are heavily model-dependent, they serve to prove

(a)

DATA listed top to bottom on plot
CDMS 2004+2005 reanalysis Ge (5 keV threshold)
CDMS 2008 Ge
CDMS: 2004+2005 (reanalysis) +2008 Ge
XENON10 2007 (Net 136 kg-d)
Roszkowski/Ruiz de Austri/Trotta 2007, CMSSM Markov Chain Monte Carlos (
Roszkowski/Ruiz de Austri/Trotta 2007, CMSSM Markov Chain Monte Carlos (
Baltz and Gondolo, 2004, Markov Chain Monte Carlos
080402190201
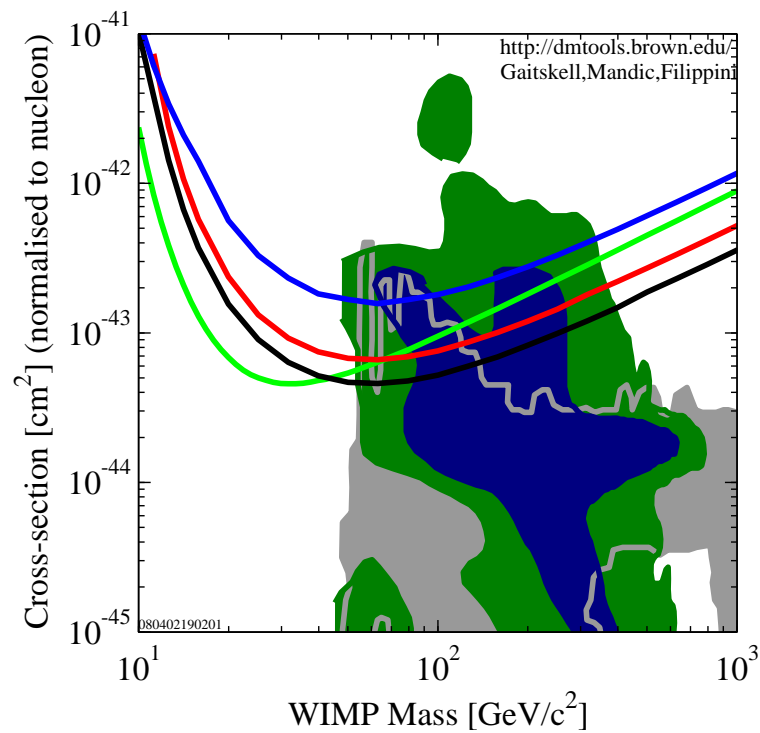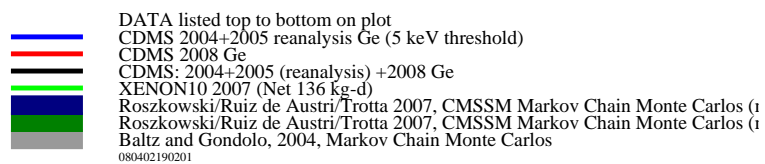
(b)

Figure 10.10: Recent WIMP exclusion limits from CDMS and Xenon experiments, with legend in (b). Figure courtesy of the DMTools limit plotter [105].

the point that CDMS has achieved sensitivity to interesting classes of WIMP models. Continued running of CDMS-II, along with the SuperCDMS program described above, provide a sensible approach to extending this reach through the most favored range of cross-sections for supersymmetric dark matter.

We have demonstrated the effectiveness of the functional-form-based position reconstruction for determining the interaction location for each event, without degeneracies. A first attempt to apply this as an aid in surface event rejection did not show significant improvement, but it is very likely that substantial additional information can be extracted using the same approach - especially by teasing out the depth of each event. Furthermore, such a reconstruction appears to be crucial for the analysis of thicker, one-inch detectors now being manufactured and demonstrated for Super-CDMS. Therefore, work on the position reconstruction will be carried forward to the five-tower runs and beyond.

Appendix A

Revision of published Run 119 results

During this work, an error was found affecting all previous Run 119 results. The true effective exposure was 17.3% lower than previously calculated, so that the true WIMP exclusion limits are a factor of 1.21 less stringent than the published regions. Corrected limit plots are presented here.

A.1   Compromised optimal filters

As described in Section 8.1.1, the optimal filter algorithm in Dark Pipe applies a filter to each trace, weighting frequencies characteristic of real pulses, and de-weighting frequencies characteristic of noise. The filter is constructed from a template pulse, which is a simple, double-exponential pulse with its time constants set appropriately. This template is fixed for each detector. The noise sample, on the other hand, is allowed to vary with changing noise conditions. The DAQ takes 500 noise traces immediately at the beginning of each data set, and additional "random trigger" noise traces are recorded, on average, as one in twenty events. The Fourier transform of the noise traces shows which frequencies should be de-weighted to filter out noise. In order for the optimal filter to work properly, this noise sample must sample only baseline noise, including the noise of the QETs and amplifiers, as well as noise pickup from external sources. It must not include any actual pulses. If an interaction happens to occur within the digitization window of a noise trace, it will cause the optimal

filter to partially de-weight the frequency range of the pulse, effectively filtering out pulses instead of passing them. In principle, Dark Pipe correctly accounts for the overlap between noise and pulses in Fourier space, so that the calibration should not be effected, and only the energy resolution will be impaired. If, however, several pulses, or large pulses, are included in the noise sample, there can also be random shifts in calibration.

For this reason, the noise analysis in Dark Pipe includes code to reject any real pulses from the noise sample. This is a simple check that the difference between the highest and lowest samples in each trace must be small enough to be consistent with noise. The actual threshold is set from the data, to keep a self-consistent sample from which any non-Gaussian tail has been removed. However, this code has been broken in versions of Dark Pipe used for the Run 118 and Run 119 analyses. Some barium calibration data sets of Run 119 were started with the barium source already inserted into the source tube, so that the initial 500 noise traces included a significant number of real pulses. Because these were not effectively rejected, these data sets were found to filter out a large part of the signal power in real pulses, a problem described as "low optimal filter power." These calibration data sets had increased noise and shifted energy calibrations on affected detectors. In the analysis of this thesis, Dark Pipe has been corrected and all data have been reprocessed, so that this problem is no longer seen, and future runs are similarly unaffected.

A.2   Rescaling phonon energies

In an attempt to correct these problems, the phonon energies of affected detectors in affected data sets were rescaled. The operation was not simply a linear re-calibration, but also allowed an offset. For each affected detector in an affected data set, a slope and intercept were chosen to minimize the width of the gamma band in yield and to place the mean at 1:

$$pt_{new} = A \times (pa + pb + pc + pd) + B.$$
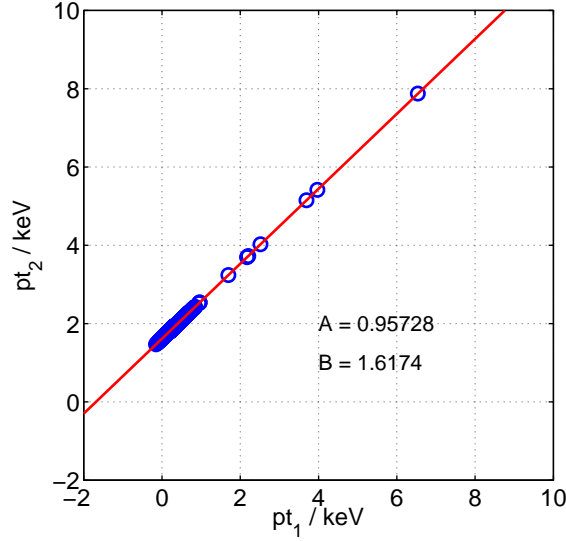
Figure A.1: Example of phonon energy rescaling for T2 Z1 in the WIMP-search data set 140421_2048. Although the energy scale is changed by only 4%, the 1.6-keV offset has a significant impact.

Since the rescaling was applied only to the total phonon energy, the phonon partition quantities of the box plot were not affected. For these data sets, the total phonon energy is no longer equal to the sum of the energies in the four quadrants. This rescaling was performed on data that had already been processed with Dark Pipe and Pipe Cleaner, as a separate and later correction to the RQ data.

The rescaling was originally and primarily restricted to the barium data sets for which the source was nto removed during the initial 500 random triggers. However, it was also found that some WIMP-search data sets had additional noise in the random triggers, with a corresponding calibration shift for the detectors T2 Z1, Z2, and Z3. This effect was largest in T2 Z1, and small in the other two detectors. Therefore, a similar correction was performed for these data sets, for T2 Z1 only. Because the rescaling values were not documented at the time, they have been fit directly from the data files of the original Run 119 analysis, as shown in Fig. A.1.

A.3   Effect on singles cut

The singles cut was defined to reject events as singles in one detector if any other detector had phonon energy more than six standard deviations above or below the mean of the noise. This was defined as a simple cut on *pt*. When the offset introduced in the rescaling was sufficiently large, it was capable of causing events with no phonon signal to fall outside the six-sigma window, so that some or all of the true single scatters in the data set would fail the singles cut. This is unimportant in the case of the barium data, where single scatters have no special significance, but critical for the 50 WIMP-search data sets in which T2 Z1 was rescaled. An efficiency adjustment for each of the rescaled WIMP-search runs has been defined by finding the fraction of random-trigger events that lie within the six-sigma *pt* window in T2 Z1. The efficiency is high for small offsets, but drops off sharply as the offset approaches ± 1 (see Fig. A.2). The total live time in affected data sets is 21.6 live days out of 77.4, or 27.9% of the overall Run 119 exposure. However, the true efficiency adjustment is smaller than this, since many of the affected data sets have small enough offsets to retain complete or partial acceptance of true single scatters. The actual loss of live time is found to be 13.4 live days, or 17.3% of the run.

A.4   Revised limits

The published Run 119 analysis does not take this effect into account, so the WIMP exclusion limits from the initial short paper [87] and several doctoral dissertations [97, 96, 82, 88] must be revised. The 17.3% reduction in exposure is independent of all other cuts, so it translates directly into a higher minimum cross-section, a factor of 1.21 higher than those previously presented. This is a small but perceptible shift when presented on a plot with logarithmic axes. Both the germanium and silicon limits are affected, as shown in Fig. A.3.
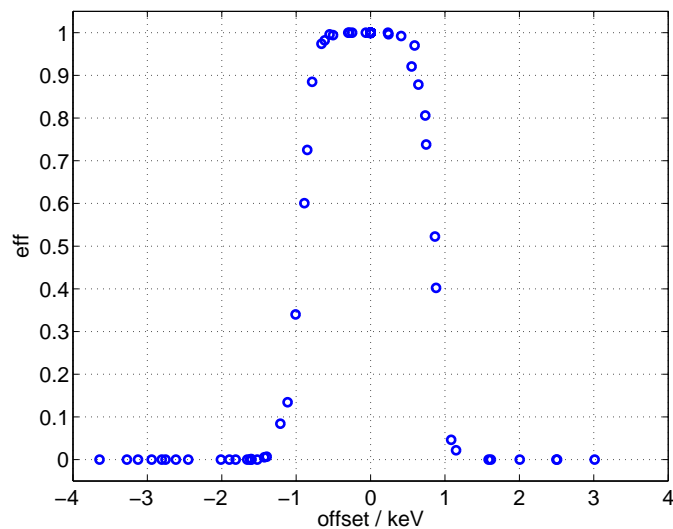
Figure A.2: Efficiency vs. rescaling offset for the 50 affected WIMP search data sets.
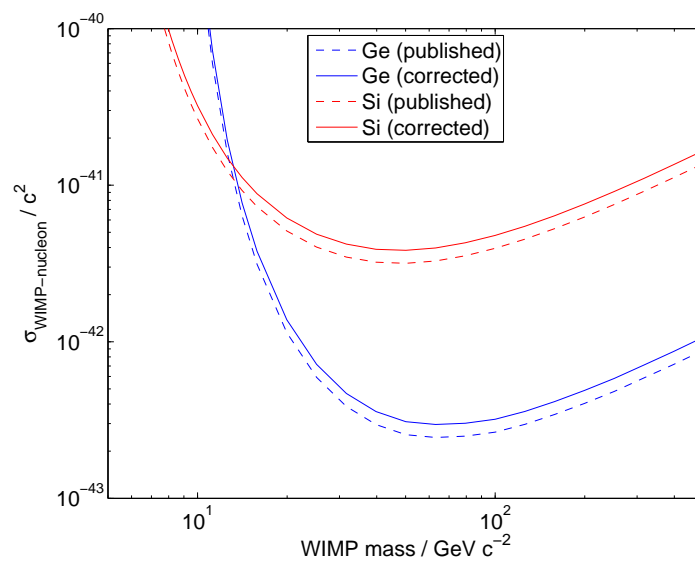
Figure A.3: Corrected Run 119 limits. The Ge limits are from the primary analysis with 10 keV threshold, and the Si limits are from the energy-dependent χ^2 analysis with 7 keV threshold. All other analyses in the previous work should be rescaled in the same way.

Appendix B

Low-threshold analysis

The preceding chapters have defined a set of cuts to allow a WIMP search in "background-free" mode, with an expected leakage of less than one event, down to an energy threshold as low as 5 keV. However, this is not a fundamental limit to the sensitivity of the detectors, which can detect interactions of 1 keV or lower. Instead, 5 keV is approximately the point at which the timing cut becomes less effective at rejecting surface events, and the yield bands widen enough that some bulk electron recoils may leak into the nuclear recoil region. Therefore, analysis below 5 keV shifts from a background-free mode to a background-limited mode in which a significant number of candidate events may be expected even in the absence of a WIMP signal. If the number of candidate events is low enough, this still permits an interesting dark matter result using low-energy data.

Although background rejection becomes less effective at low energies, many of the backgrounds are also less prominent. The beta contamination almost disappears, and gammas also decrease in rate down to a neutron activation line seen in Ge at 1.3 keV, below which they are extremely rare. At the same time, a population of "zero-ionization" events appears, caused by electronic glitches in the phonon readout and by low-energy surface scatters at the cylindrical outer surface of the detector.

The analysis of CDMS data at low threshold has been developed by Raymond Bunker for Run 21 at Stanford, as described in his doctoral dissertation [81] and a paper now in preparation. Run 21 has several advantages over Soudan at low energy:

the detectors were tuned to consistently trigger at a very low energy; overall noise levels were lower; and inner lead shielding inside the copper icebox lowered ambient gamma rates. However, the first runs at Soudan also have some advantages: the four Si detectors of Tower 2 have excellent performance compared to Tower 1 silicon detectors; activation of the Ge crystals through neutron exposure is reduced; and the irreducible neutron background of SUF has been removed. This chapter is a demonstration of the possibility of a low-threshold analysis of Soudan data, including recoil energies as low as 1 keV.

B.1   Backgrounds

The backgrounds in the 1-5 keV energy range are not identical to those at higher energies, and also differ in relative importance. At low energy the dominant backgrounds are bulk gamma recoils and zero-ionization events (including phonon-channel glitches). Because the timing cuts are not effective at such low energies, it is difficult to characterize remaining events as coming from the tails of the main gamma distribution, surface scatters, or neutrons.

Betas

At low energy, beta scatters become less prominent and nearly disappear. The surface contamination is believed to be dominated by ^{210}Pb, with beta energies of 16.96 and 63.5 keV [123]. However, photons with energy of a few keV are less penetrating, and may have their full energy absorbed within the dead layer, and can become surface events.

Gammas

At low energy the gamma and neutron bands flare in yield until they meet. Photons from the 1.3-keV line in Ge trail down well into the nuclear recoil band. To the extent that the low-energy gammas are caused by neutron activation, the gamma rate can be minimized by exposing the detectors to the ^{252}Cf neutron source as little as possible.

If necessary, periods of time immediately after the neutron calibration runs may be cut out.

Phonon-channel glitches

These are recorded when electronics noise causes a signal on the phonon channel with no corresponding ionization-channel signal. They appear to originate in the readout electronics rather than the detector itself, because the rise time and fall time are much faster than those of real events. They can be removed by a variety of cuts on energy partition, pulse shape, and proximity to periods of bad detector behavior.

Zero-ionization events

There may also be zero-ionization events caused by real particle interactions. When a particle scatters near the outer cylindrical surface of the detector, it will often fail to generate a detectable ionization-channel pulse. These must be rejected based on their event position and their extremely low ionization yield. In this initial look, we do not attempt to distinguish any real zero-ionization events from the very similar low-energy glitch events.

Neutrons

While neutrons can appear as WIMP candidates in the low-energy region as well as in the standard region of interest, they are unlikely to contribute more than a few events. Because the low-energy analysis has a significant leakage of electromagnetic and glitch events, the neutron background requirements are not as stringent as for the standard analysis. Existing CDMS Monte Carlo studies of the neutron background could be used to constrain the rate of low-energy neutron scatters.

B.2   Quick-and-dirty low energy analysis

To demonstrate the feasibility and reach of Soudan data at low threshold, we present
a "quick-and-dirty" analysis. We do not use the Ge detectors of Tower 2, since their
trigger thresholds were set too high to have any sensitivity below 5 keV. In Run 123
and beyond, they should be usable even at low energy. Tower 1 ZIP 1 similarly has
no sensitivity at low energy, and is excluded. This leaves the three Ge detectors
T1 Z2, Z3, and Z5. We use all four Si detectors of Tower 2, which have excellent
signal-to-noise and trigger thresholds < 1 keV. The two Si detectors of Tower 1 could
potentially be used at low energy (as in the SUF low threshold work), but this is not
necessary given the great performance advantage of the Tower 2 Si detectors.

Removing glitch events

The glitch events can be largely removed through cuts on bad data taking periods and
pulse shape. Some data sets have acceptable performance for the standard analysis,
but have trigger bursts at low energy. These can be removed by selecting only data
sets with an acceptably low event rate in 1-5 keV. The other anti-glitch cuts imposed
are:

- Fall time ≥ 50 μs. The glitches tend to have fast fall times, unlike real pulses.

- Rise time ≥ 7 μs (or 4 μs in some detectors). Glitches in the electronics often
  have fast rise times, while real pulses cannot rise faster than the L/R time.

- Localization on a box plot. Glitch events tend to have a fixed phonon partition,
  and appear in a tight region on the box plot.

- Phonon fraction $pfrac$. Although the difference is small, glitch events tend to
  be more localized in one quadrant than real events.

Two of the anti-glitch cuts, on partition and fall time, are shown in Fig. B.1.
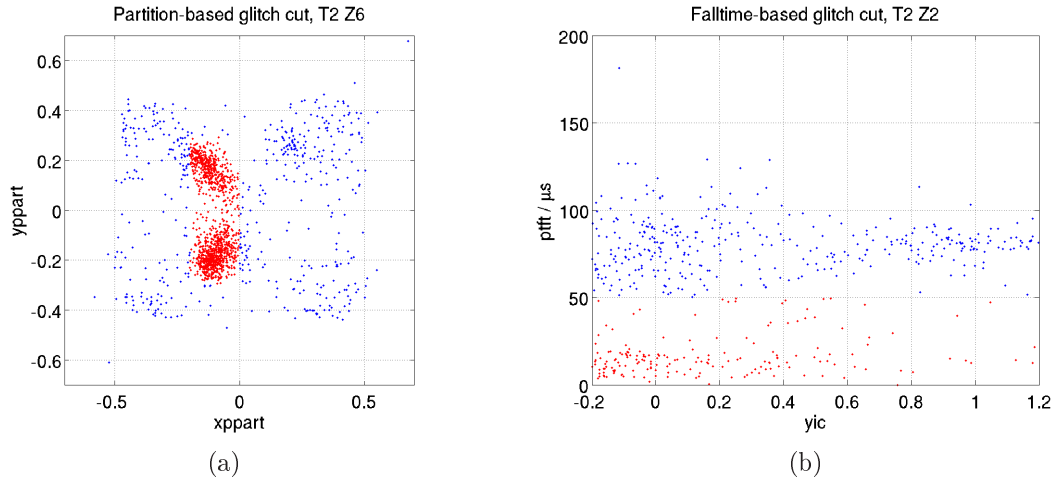
Figure B.1: Glitch removal cuts for the low-energy analysis. In (a), the box-plot cut removing most glitch events in T2 Z6. In (b), the fall-time cut removing many glitch events in T2 Z2.

Yield band cut

Setting a yield band cut at low energy is rendered more difficult by the scarcity of calibration neutrons, especially in silicon detectors. If the standard fits from high energy are extended below 5 keV, they flare out in yield much more quickly than the real neutrons do. We simply apply the standard yield band cuts, which are wide enough to clearly accept the neutrons, although a tightened yield cut for low energies could retain the neutrons while reducing the background leakage significantly.

Inner electrode cut

The inner electrode cut used is the standard two-sigma cut. This cut maintains high acceptance at low energy, at the expense of allowing some leakage of outer and shared events.

Single scatter cut

The standard single scatter cut is acceptable, as long as the thresholds for multiple scatters are well below the analysis threshold of the low-energy WIMP search.
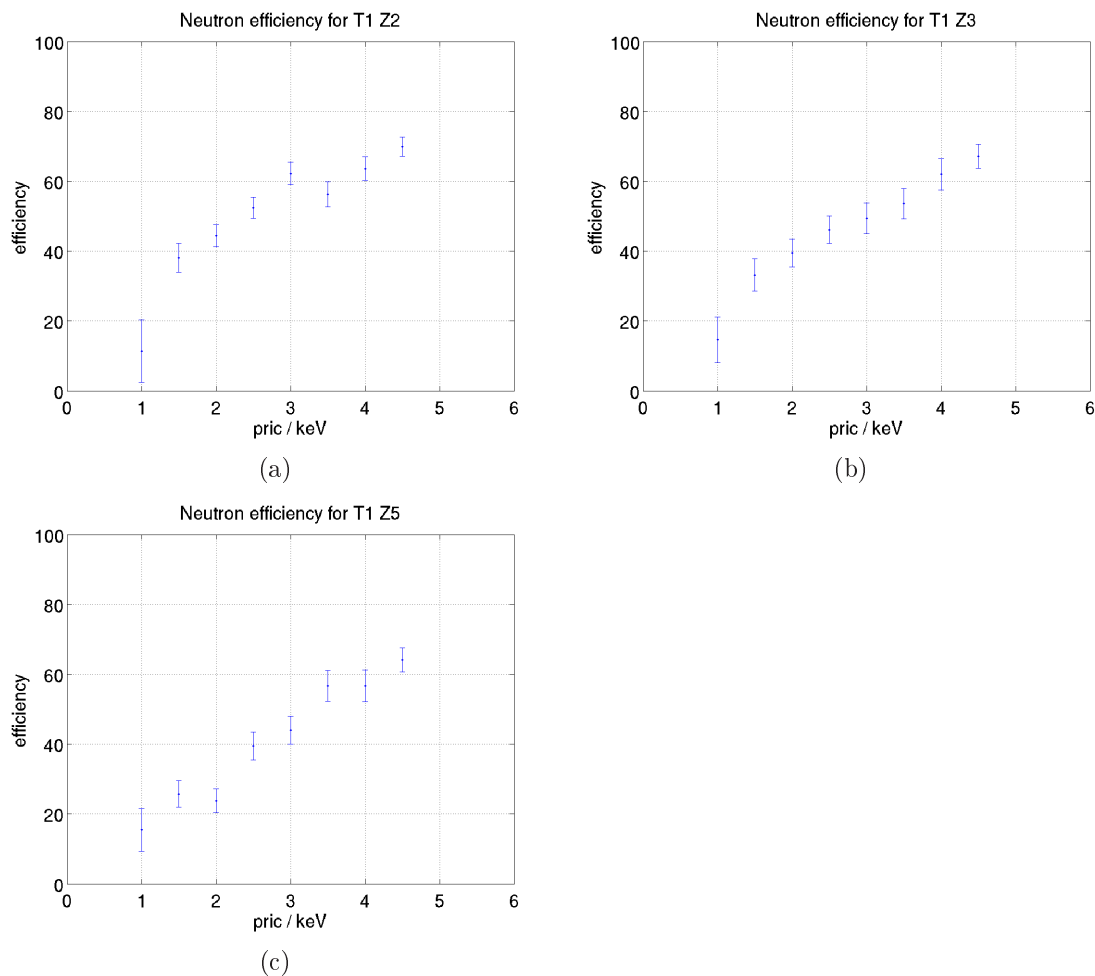
Figure B.2: Neutron acceptance at low energy in good Ge detectors.

Efficiencies

The cut efficiencies in Figs. B.2 (Ge) and B.3 (Si) are calculated directly on neutrons. They are typically around 50% over most of the range.

Remaining events

After applying the single scatter and anti-glitch cuts, there is evidence of the 1.3-keV neutron activation line in Ge detectors. Table B.1 (Ge) and Table B.2 (Si) show the number of candidate events and live time for each detector. These give an
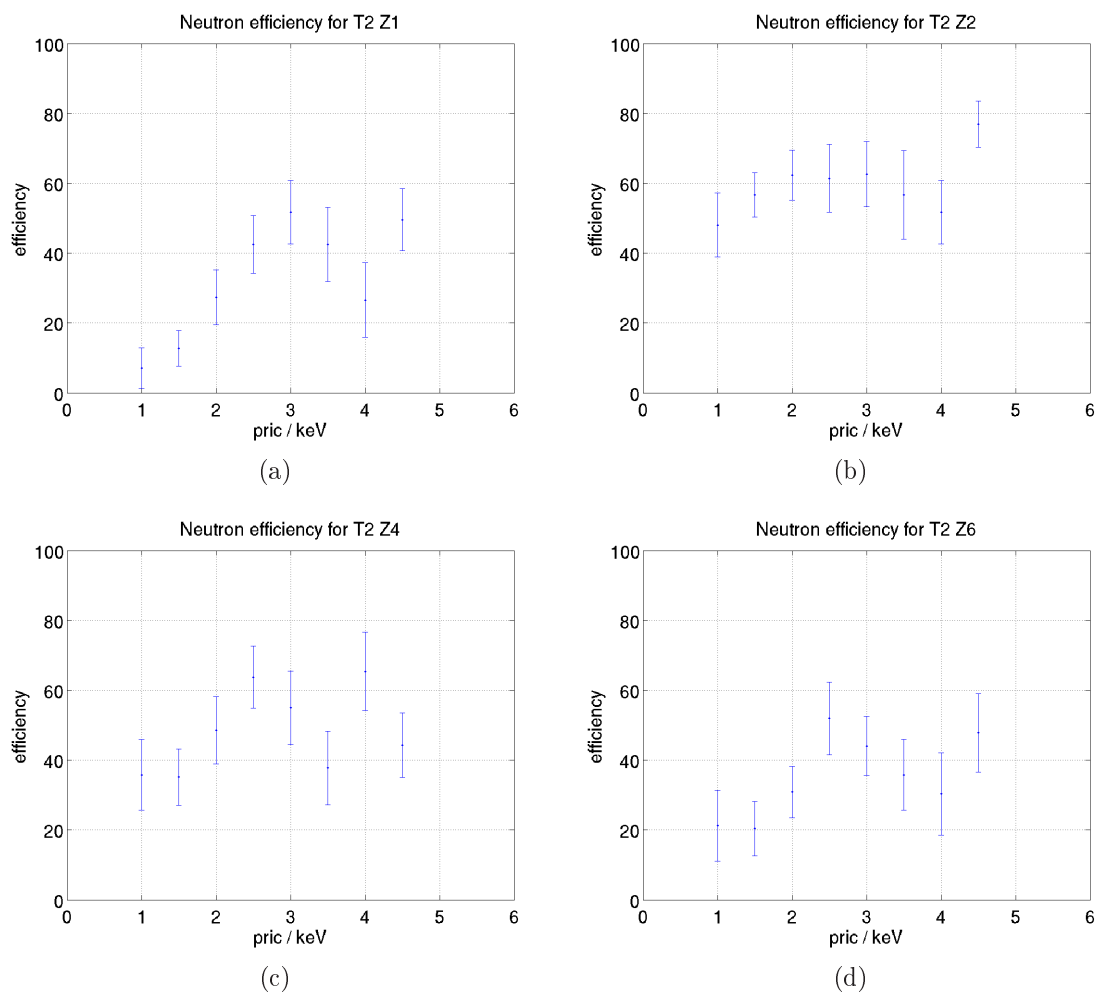
Figure B.3: Neutron acceptance at low energy in good Si detectors.

Table B.1: Events in 2-5 keV energy range, Ge detectors.

Detector	Live time / d	Counts
T1 Z2	104	49
T1 Z3	108	45
T1 Z5	110	59

Table B.2: Events in 1-5 keV energy range, Si detectors.

Detector	Live time / d	Counts
T2 Z1	72	18
T2 Z2	56	26
T2 Z4	72	25
T2 Z6	68	41

initial estimate of the background rate that can be extracted from Soudan data at low energies with a straightforward analysis. No WIMP exclusion limits have been calculated from these figures, but they are comparable to the background rate at SUF as found by Ray Bunker. Also, the thresholds used do not match perfectly. Given these caveats and coadding all used detectors, we find 153 Ge candidates 2-5 keV in a total of 322 detector-live days, or 0.5 d^{-1}. The Si results are similar : we find 110 events 1-5 keV in 268 detector-live days, or 0.4 d^{-1}. A more careful analysis should be capable of improving on these figures.

Appendix C

Charge generation

A particle interaction of a few keV or more creates a cloud of a large number of charges. The size of the ionization signal is given by the number of electron-hole pairs excited by an incident particle of a given energy. With the exception of ionization quenching for nuclear recoil events, this relation is linear in energy and does not depend on the identity of the initially scattering particle. Correct understanding of the fundamental charge yield in Ge and Si is crucial for CDMS, as the Luke phonon correction and the recoil energy scale for neutrons both depend on it. This section gives a brief review of the literature on charge generation in germanium and silicon.

The amount of energy needed to excite an electron-hole pair in a semiconductor is the band gap, which is very well known. The band gap has a small dependence on temperature, commonly approximated by the Varshni relation [111]

$$E_g(T) = E_g(0) - \frac{\alpha T^2}{T + \beta}. \tag{C.1}$$

The parameters for silicon and germanium are given in Table C.1. For both materials, the band gap at 50 mK does not deviate significantly from the 0K value, so we take E_g=1.170 eV (Si) and 0.7437 eV (Ge) for our detectors under operating conditions.

Most of the energy of any interaction, however, does not go into producing electron-hole pairs. The number of charge pairs produced by an electron recoil is
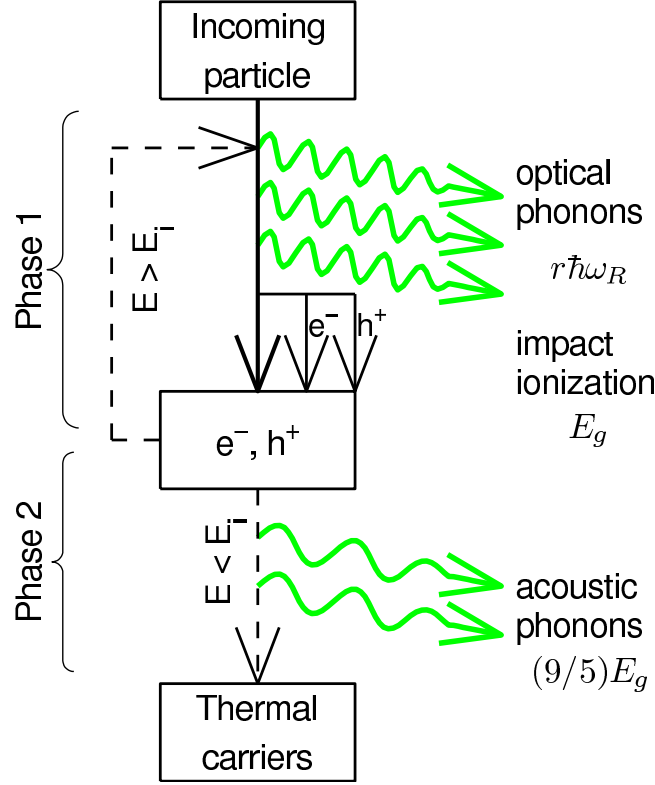
Figure C.1: Schematic of energy deposition in electron recoil events. In phase 1, a "hot" charged particle emits $\sim 10 - 20$ optical photons per ionization event. The excited daughter electrons and holes participate in the same scattering interactions until their energies are sufficiently low ($E < E_i$) that they can no longer ionize or couple to optical photon modes. At this point, the slower phase 2 begins, and the residual kinetic energy of the charges is dissipated as optical phonons. The energy accounting in Klein's model is shown to the right. (Note that this is not a full accounting of the *phonon* signal, as Neganov-Luke phonons and recombination phonons have not been included).

Table C.1: Parameters of the Varshni equation for band gap as a function of temperature, from Thurmond [112].

Material	$E_g(0)$	$\alpha/\ 10^{-4}$eV K^{-1}	$\beta/$K
Si	1.170 ± 0.001	4.730 ± 0.25	636 ± 50
Ge	0.7437 ± 0.001	4.774 ± 0.30	235 ± 40

$N_Q = \frac{E_R}{\epsilon}$, where ϵ is the average amount of energy required to create one electron-hole pair, taking into account the competition with other processes. It is usually assumed that the temperature dependence of ϵ comes only from small changes in band gap. Values of ϵ for Si and Ge are well known at 77K and above.

A model developed by Shockley [113] gives

$$\epsilon = 2.2E_i + rE_R. \tag{C.2}$$

The threshold energy for ionization, E_i [1], and the number of optical phonon scatters per ionizing event, r, are empirically derived from data; E_R is the Raman phonon energy (the energy of the highest energy optical phonons), found from neutron scattering data. The model may be understood as partitioning the total energy into three portions: the band-gap energy required to create the charge pairs; the kinetic energy of the charges; and energy used to create optical phonons. This model is compared favorably with experimental work in thin samples at very high field and at room temperature. The Shockley model was used down to 20K, still at high field, by Emery and Rabson [114], who found ϵ sharply increasing below 50K.

$$\epsilon_{Ge} = 2.2E_g(T) + 1.99E_g^{3/2}(T)exp(4.75E_g(T)/T)$$
$$\epsilon_{Si} = 2.2E_g(T) + 1.08E_g^{3/2}(T)exp(11.2E_g(T)/T)$$

The authors attributed the increase in ϵ without any corresponding change in E_g to an increase in r, the number of phonon-shedding interactions per ionizing interaction. However, the high values of ϵ may have been instead the result of trapping on impurities.

Klein has elucidated and improved upon the Shockley model, successfully explaining the Fano factors in various materials as well as ϵ [115, 116]. Energy deposition proceeds in two phases, shown schematically in Fig. C.1. The first, nearly instantaneous, phase is a cascade in which the high-energy incident particle knocks electrons and holes into the conduction band; these in turn create more charges, until the

---

[1] The threshold energy E_i differs from the band gap E_g because of the need to conserve momentum as well as energy. When ionization is assisted by phonons, it can occur at the gap energy [117].
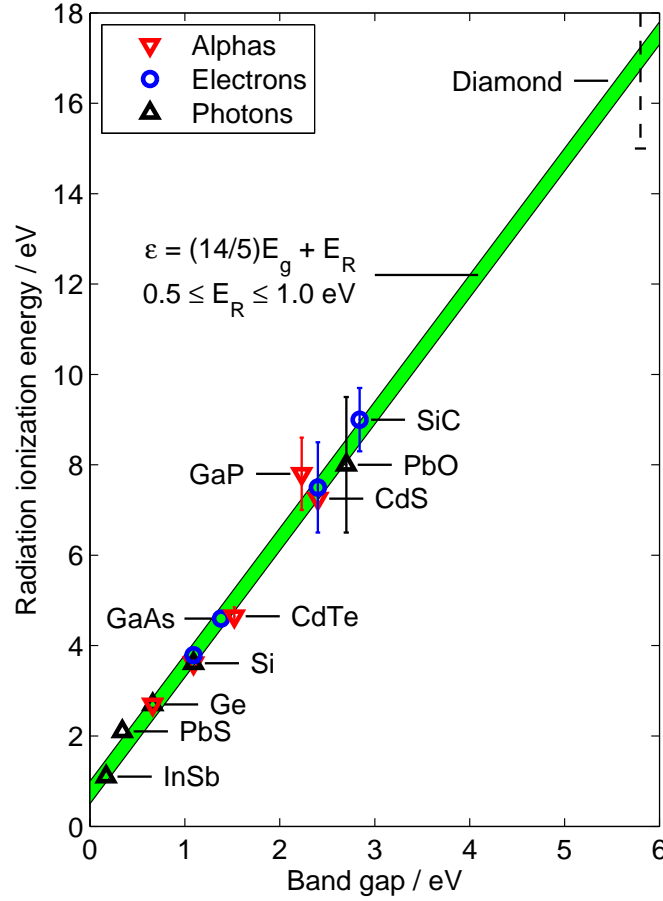
Figure C.2: Ionization energy ϵ vs. band gap energy for a variety of materials, from Fig. 3 of Klein's 1968 paper [115]. The green-shaded diagonal stripe shows the semi-empirical relation of ϵ to band gap, with optical phonon losses in the range 0.5-1.0 eV. The remarkable agreement for small-gap and large-gap semiconductors shows that optical phonon losses do not vary widely, while residual kinetic energy of the carriers is well modeled by $(9/5)E_g$. Note the discrepancy between alphas, betas, and photons in Si.

energy of the individual charge carriers is below the ionization threshold. During this process, the energetic electrons and holes also interact with the crystal lattice by emitting optical phonons (taken to be Raman phonons, but possibly including lower energies). The charge carriers now have kinetic energy below E_i, so that they cannot induce any further ionization, and no longer couple to optical phonon modes. The second phase is the dissipation of this residual kinetic energy to acoustic phonon modes over a time scale of $\sim 10^{-11}$ s.

The residual kinetic energy is related to the band gap by assuming that the carriers uniformly populate momentum space. This assumption gives a residual kinetic energy of $(9/5)E_g$, so that

$$\epsilon = \frac{14}{5}E_g + r\hbar\omega_R. \tag{C.3}$$

It is observed that for many semiconductors $r\hbar\omega_R$ is of order 0.5-1.0 eV, independent of band gap. According to this model, the temperature dependence of ϵ comes only from the change in E_g, with constant optical phonon losses. The relation between ionization energy and band gap is confirmed for a variety of semiconductor materials (See Fig. C.2), indicating that Klein's accounting is correct. The dependence of ϵ on temperature compares well with several data sets for Ge (Fig. C.3), as expected given the known temperature dependence of the band gap. A further success of this model is the good match to the Fano factor for germanium at 77 K.

We therefore take Eq. C.3 as correct, and calculate from it values of the ionization energy appropriate for CDMS-II ZIP detectors. For Ge, this is straightforward. In the case of Si, there is an additional complication in that ϵ does have a single value for all types of incident radiation: alphas apparently require *less* energy to create the same amount of ionization as a gamma or electron scatter. Figure C.2 shows the same value of ϵ for alphas and gammas, and a higher value for electrons. However, definitive measurements by Pehl et al. find $\epsilon_\alpha = 3.62 \pm 0.02$ eV and $\epsilon_\beta = \epsilon_\gamma = 3.67 \pm 0.02$ eV at room temperature: electrons and photons are consistent, and only alphas are anomalous. For CDMS, the value of interest is that for electrons and photons[2]. The

---

[2]Alpha-particle scatters are seen only on the surfaces of ZIP detectors, where the "dead region" effect greatly suppresses the ionization signal. They are not confused with WIMP events because of their large phonon signals.
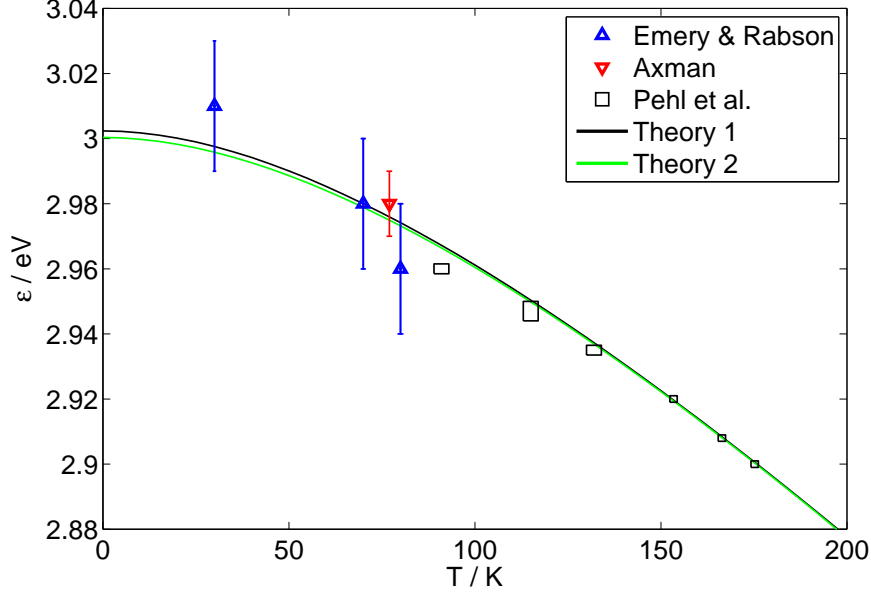
Figure C.3: Ionization energy vs. temperature for Ge, as in Figure 5 of Klein's paper [116]. The two "Theory" curves use Eq. C.1 for band gap and Eq. C.3 for ϵ. The black "Theory 1" curve is as in Klein, using parameters from Varshni [111] in Eq. C.1 and taking $r\hbar\omega_R = 0.927$ eV. The lighter, green "Theory 2" curve adopts more recent band-gap parameters from Thurmond [112] and adjusts the optical phonon losses to $r\hbar\omega_R = 0.918$ eV to maintain a good fit to the data.

third column of Table C.2 applies Eq. C.3 to the data of Pehl et al. for electrons and photons in Si at 300 K.

Key parameters for Si and Ge are summarized in Table C.2. From column 3 we adopt and $\bar{E}_R = 0.562$eV for Si at room temperature. Klein takes $r\hbar\omega_R = \bar{E}_R = 0.927$ eV for Ge at 77 K, but uses an older set of parameters for band gap. We adopt a more recent set of measurements for band gap in germanium, giving instead $\bar{E}_R = 0.918$ eV (see Fig. C.3). As the optical phonon losses are temperature-independent, the same values apply near 0 K. In conjunction with the band gap at base temperature as in Table C.1, we find that $\epsilon_{Ge} = 3.00$ eV and $\epsilon_{Si} = 3.84$ eV.

Table C.2: Calculating optical phonon losses. The column for Ge is taken directly from Table I of Klein [116]. The first column for Si is from the same source, although $\bar{E}_R$ is not explicitly given. The identity of the incident radiation is not specified, but may be inferred from the low value of ϵ to be α particles. The final column applies the same model to the ϵ values for β and γ radiation from Pehl et al. [118].

	Ge	Si (α)	Si (β, γ)
T/K	77	293	300
E_g/eV	0.735	1.1	1.11
ϵ/eV	2.975	3.6	3.67±0.02
Phonon losses (E_R/E_g)	1.26	0.47	0.506
$\bar{E}_R$/eV	0.927	0.52	0.562

Measurements have been done by the CDMS collaboration and others in calibrating discriminating dark matter bolometers. BLIP detectors used by CDMS-I include a charge "pulser" circuit that can be used to calculate an absolute calibration, including all capacitive effects. Shutt et al. [119, 52] do not explicitly calculate ϵ_{Ge} from their Ge BLIP data, but the consistency of their calibration with $\epsilon_{Ge} = 3.0$ eV confirms this to within a few percent. The Edelweiss calibration [120] gives 0.34 charges per eV, and their collaboration has adopted a value of 3.0 ± 0.1 eV.

Betty Young has operated and calibrated a 23-gram silicon BLIP detector, known as Si-1, with implanted charge collection contacts [121]. The ionization yields for gamma decay lines at high field are consistent with $\epsilon_{Si} = 3.8$ eV [122]. These results in Si and Ge confirm the applicability of the ϵ values adopted here to ZIP detectors at 50 mK.

The CDMS collaboration has adopted the values $\epsilon_{Ge} = 3.0$ eV and $\epsilon_{Si} = 3.82$ eV. These are consistent with the values found here to within present uncertainties, and are used in all data analysis in this work.

Appendix D

Software tools

The reduction and analysis of CDMS-II data are performed using several custom software packages in the MATLAB scientific computing environment. At several points in this thesis work it has become necessary to develop software tools that are needed for CDMS analysis, but that are not directly related to physics analysis and have a more general range of applicability. Two of these tools are documented and included in this appendix.

D.1   Vectorized Nelder-Mead minimization

Nelder-Mead minimization is a commonly used method for n-dimensional minimization [109]. Existing implementations of the Nelder-Mead algorithm, such as MATLAB's *fminsearch* and Nicholas Higham's *nmsmax* [110], typically use if-then logic to move and reshape the simplex and to determine when iteration should stop. Because of this, a single call to the minimization tool can only perform a single minimization; there is no possibility of vectorization. For applications in which the same function must be minimized a large number of times with different parameters, each minimization must be performed separately. A for-loop is required, and the number of calls to the function to be minimized scales linearly with the number of minimizations to be performed. In MATLAB , this is not efficient. A better approach is to use vectorized code in which each call to the function to be minimized actually evaluates the

function at a large number of points, taking advantage of MATLAB's very fast matrix libraries.

The function *nmvec* has been written with exactly the same Nelder-Mead algorithm, but it is implemented with mask-array logic instead of if-then logic. This allows a large number of minimizations to proceed in parallel, with substantial performance improvements. As the number of minimizations grows, the advantage of *nmvec* over *fminsearch* becomes quite large. For example, estimating the radial position of 30,000 events takes 480 seconds using repeated calls to *fminsearch*, but only 25 seconds using a single call to *nmvec*. [1]

D.1.1   Code listing for *nmvec.m*

```
function [X, FVAL] = nmvec (FUN, X0, OPTIONS, P, varargin)
% NMVEC Vectorized Nelder–Mead minimization — for when
% you want to solve the same problem many times with
% different parameters.
%
% X = NMVEC (FUN, X0) works like FMINSEARCH, except that
% each row of X0 represents a different set of starting
% guesses.  Multidimensional optimization may be done by
% using one column of X0 for each dimension.  The function
% FUN must accept inputs with arbitrary sizes in the first
% dimension.  For instance:
% FUN = inline ('sin (x(:,1)) + cos (x(:,2))');
% FUN may be a string, an inline function, or a function
% handle.
%
% X = NMVEC (FUN, X0, OPTIONS) accepts OPTIMSET options.
%
% X = NMVEC (FUN, X0, [], P) passes additional parameters P
% to FUN.  The rows of P should match up with the rows of
```

---

[1] MATLAB Version 7.1.0.183 (R14) Service Pack 3 on Linux, with a Pentium M 1500 MHz processor. Optimization parameters: maximum of 400 iterations, function tolerance 10^{-4}, coordinate tolerance 10^{-4}.

```
% X0, allowing fast minimization of a family of functions
% with extra parameters.
%
% [X, FVAL] = NMVEC (...) also returns the function value
% at the minimum.
%
% References:
% J.A. Nelder and R. Mead, A simplex method for function
%    minimization, Comput. J. 7, 308-313 (1965).
% N.J. Higham, The Matrix Computation Toolbox,
%    http://www.ma.man.ac.uk/¬higham/mctoolbox/
%
% RWO Sep 18-19 2006


% Notes:


% Note there's still lots of room for improvement!
%  - clean up comments left over from design
%  - add refs to papers, other code
%  - add feval counting
%  - force P to be an array, not a cell array
%  - ...then clean up fevals to only evaluate needed j values
%  - clean up preallocation of memory
%  - only calculate points if at least one j value needs them


% M = number of problems to solve
% N = dimensionality of optimization
% Data structures:
% cpos(i,j,k) - positions of corners
%    - k is index of corner, 1..N+1
%    - i is data sample index, 1..M
%    - j is dimension index, 1..N
%
% fval(i,k) - function val at corners
%
% Be willing to loop over k, but not i or j.
% (this is arbitrary; would be OK to loop over
```

```matlab
% j as well.  just don't want to loop over i
% under any circumstances!)

if ischar (FUN)
   FUN = inline(FUN);
end;
if (nargin < 4)
   P = {};
elseif ¬iscell (P)
   P = {P};
end;
[M N] = size(X0);
disp (['M=' num2str(M) ', N=' num2str(N) ...
   '.  Will work well for M>>N.']);
if (nargin < 3) | isempty (OPTIONS)
   OPTIONS = optimset ('fminsearch');
end;
MAXITER = optimget (OPTIONS, 'MaxIter');
FTOL = optimget (OPTIONS, 'TolFun');
SIZETOL = optimget (OPTIONS, 'TolX');

% 0. Prepare initial simplex
cpos = zeros(M,N,N+1);
fval = zeros(M,N+1);
% 0.a. one vertex is x0
cpos(:,:,1) = X0;
fval(:,1) = feval (FUN, X0, P{:}, varargin{:});
% 0.b. determine scale: 1 or max val in x0
scale = max (max (abs (X0), [], 2), 1);
% 0.c. set other vertices
%    0.c.i. equilateral − not coded
%    0.c.ii. right angles
for (k=2:(N+1))
   cpos(:,:,k) = X0;
   cpos(:,k−1,k) = X0(:,k−1) + scale;
   fval(:,k) = feval (FUN, cpos(:,:,k), P{:}, varargin{:});
end;
```

```
% 0.d. sort corners
[sval s] = sort (fval, 2);


% 1. Loop − while forever!
iter = 0;
while (1)
   iter = iter + 1;

   % 1.a. Check if done
   isdone1 = (sval(:,N+1) − sval(:,1) < FTOL);
   bestpos = zeros(M,N);
   for (k=1:(N+1))
      bestpos(s(:,1)==k,:) = cpos(s(:,1)==k,:,k);
   end;
   simplex_size = zeros(M,1);
   for (k=1:(N+1))
      simplex_size = max ([simplex_size ...
         sum(abs(cpos(:,:,k)−bestpos),2)], [], 2);
   end;
   % simplex_scaler = max (1, max (abs (bestpos), [], 2));
   simplex_scaler = 1;
   isdone2 = (simplex_size ./ simplex_scaler < SIZETOL);
   if sum (¬isdone1 & ¬isdone2) == 0
      break;
   end;


   if (iter > MAXITER)
      disp (['Reached ' num2str(MAXITER) ' iterations. stopping!']);
      break;
   end;


   % Find and store worst point
   worstpos = zeros(M,N);
   for (k=1:(N+1))
      worstpos(s(:,N+1)==k,:) = cpos(s(:,N+1)==k,:,k);
   end;
```

```
% 1.b. Find and store center, excluding worst point
cent = (sum (cpos, 3) − worstpos) / N;
% 1.c. Reflect worst corner through center
reflpos = 2*cent − worstpos;
% 1.d. Evaluate at new position
fvrefl = feval (FUN, reflpos, P{:}, varargin{:});
% 1.d.i.(a). if new best...
refl_is_best = (fvrefl < sval(:,1));
% ... expand 2x rel. center
expdpos = 2*reflpos − cent;
% eval fn
fvexpd = feval (FUN, expdpos, P{:}, varargin{:});
expd_is_best = refl_is_best & (fvexpd < sval(:,1));
% 1.d.i.(a). if refl & expd better than old best, keep expd
% 1.d.i.(b). if refl better than old best, keep refl
newpos = NaN * ones(M,N);
newpos(refl_is_best,:) = reflpos(refl_is_best,:);
newpos(expd_is_best,:) = expdpos(expd_is_best,:);


% 1.d.ii. If refl is better than worst, make new worst
refl_is_notworst = (fvrefl < sval(:,N+1));
worstpos(refl_is_notworst,:) = reflpos(refl_is_notworst,:);
% 1.d.iii. Move worst point halfway to center
contrpos = (worstpos + cent) / 2;
% eval fn
fvcontr = feval (FUN, contrpos, P{:}, varargin{:});
% 1.d.iii.(a). Better than second worst?
contr_is_not2worst = (fvcontr < sval(:,N));
newpos(¬refl_is_best & contr_is_not2worst,:) = ...
   contrpos(¬refl_is_best & contr_is_not2worst,:);
% 1.d.iii.(b). Else move all points halfway to best
bestpos = zeros(M,N);
for (k=1:(N+1))
   bestpos(s(:,1)==k,:) = cpos(s(:,1)==k,:,k);
end;
for (k=1:(N+1))
   cpos(¬refl_is_best & ¬contr_is_not2worst,:,k) ...
```

```
          = (cpos(¬refl_is_best & ¬contr_is_not2worst,:,k) ...
          +  bestpos(¬refl_is_best & ¬contr_is_not2worst,:)) ...
          / 2;
      fval(:,k) = feval (FUN, cpos(:,:,k), P{:}, varargin{:});
   end;
   use_newpos = refl_is_best | contr_is_not2worst;

   fvnew = feval (FUN, newpos, P{:}, varargin{:});
   % Replace old worst with whatever it's replaced with!
   for (k=1:(N+1))
      cpos(use_newpos & s(:,N+1)==k,:,k) ...
          = newpos(use_newpos & s(:,N+1)==k,:);
      fval(use_newpos & s(:,N+1)==k,k) ...
          = fvnew(use_newpos & s(:,N+1)==k);
   end;

   % 1.3 Sort corners
   [sval s] = sort (fval, 2);
end;


for (k=1:(N+1))
   X(s(:,1)==k,:) = cpos(s(:,1)==k,:,k);
end;
FVAL = sval(:,1);


if sum(isdone1)>0
   disp ([num2str(sum(isdone1)) ' reached function target']);
end;
if sum(isdone2)>0
   disp ([num2str(sum(isdone2)) ' reached simplex size limit']);
end;
if sum(isdone1&isdone2)>0
   disp ([num2str(sum(isdone1&isdone2)) ' did both']);
end;
if sum(¬isdone1&¬isdone2)>0
   disp ([num2str(sum(¬isdone1&¬isdone2)) ...
      ' did neither, max iterations']);
```

```
end;
disp (['After ' num2str(iter) ' iterations.']);
```

D.2 MATLAB preparser for compiling

MATLAB M-files can ordinarily run only inside the interpreted MATLAB environment (or a partially compatible package such as Octave). The MATLAB compiler toolbox may be used to compile MATLAB code into a stand-alone executable, with some limitations. In particular, all compiled code must consist of function M-files; script M-files are not allowed. Depending on the version of the MATLAB compiler, there are also other constraints in the use of the *eval* command to construct commands at runtime, execution of system shell commands, and other more minor items, as well as various small incompatibilities in the way some supported commands work inside the MATLAB environment and in a stand-alone program. Therefore, code that is intended to be compiled must be carefully developed to avoid these constructions.

The CDMS data reduction package *DarkPipe* was developed before the MATLAB compiler toolbox existed, and makes heavy use of features that do not compile well. In particular, it is mainly organized in the form of scripts that run in the same namespace and share variables. In order to generate a compiled, stand-alone version of *DarkPipe* that can be run on the QCD cluster and the main computing farm at Fermilab, therefore, a number of workarounds were required. Rather than rewriting all of *DarkPipe*, a MATLAB code preparser has been created. This preparser can be used to work around the prohibited constructions before compiling, so that the code remains unchanged when run inside of MATLAB , but can be compiled. For instance, the preparser can paste in the code of script M-files to replace calls to the scripts, and obeys a number of preparser tags that can be used to insert workaround code that is ignored when running inside of MATLAB , but included when compiling.

This preparser is quite general and powerful, allowing *DarkPipe* to be compiled using the MATLAB compiler version 3 (with MATLAB versions 6 and earlier). The MATLAB 4 compiler (with MATLAB versions 7 and later) eliminates some of the restrictions of the older compiler, so that only a subset of the preparser's features

(mainly the script inline replacement and path management) need be used.

D.2.1   Preparser tags

The preparser obeys a set of special tags that can be included in MATLAB M-files. These always begin with the sequence of characters %//, so that the interpreter treats them as comments and ignores them. In the following list, each entry that appears in single quotes must evaluate to a string. It may be a literal string or an expression.

- *%//script 'NAME'* - replaces any code on this line with the text of a script M-file NAME.

- *%//script* - looks for the name of the script M-file in the code before this tag and on the same line. For example, a line of code "*my_script %//script*" will be replaced with the text of my_script.m.

- *%//compile 'ON'/'OFF'* - indicate that the following code should be included in / omitted from the compiled version.

- *%//compile* - indicates that this line should be included in the compiled version, even if compile is 'OFF'.

- *%//compile 'SKIP'* - indicates that this line should be omitted from the compiled version, even if compile is 'ON'.

- *%//eval 'STATEMENT'* - evaluates STATEMENT in the workspace of the preparser (!).

- *%//insert 'CODE'* - inserts CODE into the code for compilation.

- *%//. CODE* - shorter form of insert, note no quotes.

- *%//replace 'CODE'* - replaces everything on the same line as this tag and preceding it with the code CODE. Like *%//insert*, but a replacement instead of an insertion.

- *%//linefeed 'ON'/'OFF'* - in the following lines of code, include / omit the line feed at the end of each line.

- *%//linefeed 'SKIP'*

- *%//linefeed* or *%//eol* - insert a line feed character here.

- *%//def 'NAME'* - sets a flag called NAME.

- *%//undef 'NAME'* - un-sets a flag called NAME. (Flags can be tested with *%//if 'def ("NAME")'*.)

- *%//if 'CONDITION'*

- *%//elseif 'CONDITION'*

- *%//else*

- *%//endif*

- *%//while 'CONDITION'*

- *%//endwhile*

- *%//repeat*

- *%//until 'CONDITION'*

- *%//for {'INIT', 'TEST', 'INCREMENT'}* - this is like a C for loop, but the initializer, test, and increment should be given as three strings in a cell array.

- *%//endfor*

- *%//stamp* - generates a comment line with a time stamp (the time of pre-parsing).

- *%//% COMMENTS...* - comments that will be stripped out by the pre-parser.

D.3   Fitting band edges

The function *bandedge* has been devised to fit a functional form to the edge of a band, rather than the mean. In the case of a CDMS yield band, the independent variable is energy E and the dependent variable is ionization yield y. A particular curve $f(E)$ is a good representation of the edge of the band if a fixed fraction p of the data points have $y < f(E)$, independent of E. In other words, the goal is to fit the location of a selected percentile of the events, rather than the mean. A band upper edge can be found with $p = 0.8$ or 0.9, and a median can be found with $p = 0.5$. This approach is more robust at finding the edge of non-Gaussian bands than others (such as fitting Gaussian curves to the distribution in several energy bins).

   The function *bandedge* minimizes a running residual w. This is initialized to $w = 0$. Next, all of the data points are stepped through in order of energy. For each event with $y < f(E)$, w is decremented by $p - 1$. For each event with $y > f(E)$, w is incremented by p. The abrupt step is smoothed by replacing it with a steep error function, in order to ensure good convergence. For a desired functional form with several parameters, the parameters are chosen to minimize the absolute value of the furthest excursion of w from zero. When this residual is minimized, the curve represents the p percentile of the band as well as possible given the chosen functional form.

D.3.1   Code listing for *bandedge.m*

```
function [A m w] = bandedge (x, y, f, A0, p)
% BANDEDGE fits the edge of a band.
%    It finds a curve with a fixed percentage of
%    data points below the curve, without making
%    assumptions about Gaussianity or using bins.
%
%    A = BANDEDGE (X, Y, F, A0, P)
%       X  = independent variable from data
%       Y  = dependent variable from data
```

```
%       F  = functional form for fit
%       A0 = initial parameter guess for F
%       P  = fraction of data points to fall
%            below the curve
%       A  = fitted final parameters
%
%    Example:
%       A = bandedge (prg, yc, ...
%            @(A,x) A(1) + A(2)*x.^2, ...
%            [1 0.01], 1/4);
%       fits a constrained quadratic to
%       the lower edge of the yield band.
%
%    Note that P=1/2 fits the median of the band.
%    This can give different results from fitting
%    the mean.
%
%    If P is not given, the default value 1 - 1/e
%    is used.
%
%    [A M] = BANDEDGE (...) also returns the value
%    of the objective function minimized to find A.
%
%    [A M W] = BANDEDGE (...) also returns the
%    residuals at each data point.
%
%    RWO 070612

if (nargin < 5) | isempty (p)
    p = 1 - 1/exp(1);
end;

[xx ii] = sort (x);
yy = y (ii);
A = fminsearch (@qualfunc, A0, [], xx, yy, f, p);

if (nargout > 2)
```

```
    [m w] = qualfunc (A, xx, yy, f, p);
elseif (nargout > 1)
    m = qualfunc (A, xx, yy, f, p);
end;



% The objective function.
%
% Note: data pairs (x, y) must already
% be sorted in x!

function [m wout] = qualfunc (A, x, y, f, p)
    smoothscale = (max(y) − min(y)) / 1000;
    scalefac = 1e4;
    yfit = f (A, x);
    % smoothscale = mean (abs (yfit − y));
    % w = p * double (y > yfit) + (p−1) * double (y < yfit);
    w = (p−1/2) + 1/2 * erf ((y − yfit) / smoothscale);
    w = cumsum (w);
    m = max (abs (w)) * scalefac;
    if (nargout > 1)
        wout = w;
    end;
```

Bibliography

[1] G. Bertone, D. Hooper, and J. Silk, Particle dark matter: Evidence, candidates and constraints. arXiv:hep-ph/0404175v1 (2004).

[2] J.J. O'Connor and E.F. Robertson, Mathematical discovery of planets. In the *MacTutor History of Mathematics Archive*, at `http://www-groups.dcs.st-and.ac.uk/~history/HistTopics/Neptune_and_Pluto.html` (1996).

[3] S. Jha, D. Charbonneau, P.M. Garnavich, D.J. Sullivan, T. Sullivan, T.M. Brown, and J.L. Tonry, Multicolor observations of a planetary transit of HD 209458. The Astrophysical Journal 540, L45-L48 (2000).

[4] S.C. Unwin *et al.* Taking the Measure of the Universe: Precision Astrometry with SIM PlanetQuest. Publications of the Astronomical Society of the Pacific 120, 38–88 (2008).

[5] P.R. Lawson, O.P. Lay, K.J. Johnston, and C.A. Beichman, Terrestrial Planet Finder Interferometer Science Working Group Report. `http://planetquest.jpl.nasa.gov/TPF/TPFIswgReport2007.pdf` (2007).

[6] G.R. Blumenthal, S.M. Faber, J.R. Primack, and M.J. Rees, Formation of galaxies and large-scale structure with cold dark matter. Nature 311, 517 (1984).

[7] D.N. Schramm, and M.S. Turner, Big Bang Nucleosynthesis enters the precision era. Rev. Mod. Phys. 70, 303 (1998).

[8] V. Rubin and W.K. Ford, Jr., Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. Astrophysical Journal 159, 379 (1970).

[9] V. Rubin, N. Thonnard, and W.K. Ford, Jr., Rotational properties of 21 Sc galaxies with a large rangee of luminosities and radii from NGC 4605 (R=4 kpc) to UGC 2885 (R=122 kpc). Astrophysical Journal 238, 471 (1980).

[10] Y. Sofue and V. Rubin, Rotation curves of spiral galaxies. Annu. Rev. Astron. Astrophys. 39, 137-74 (2001).

[11] Y. Sofue, A. Tomita, M. Honma, and Y. Tutui. Publ. Astron. Soc. Jpn. 51, 737 (1999).

[12] R.P. Olling and M.R. Merrifield, Two measures of the shape of the Milky Way's dark halo. Mon. Not. R. Astron. Soc. 311, 361-369 (2000).

[13] J.F. Navarro, C.S. Frenk, and S.D.M. White, The structure of cold dark matter halos. The Astrophysical Journal 463, 563 (1996).

[14] F. Zwicky, Die Rotverschiebung von extragalaktischen Nebeln. Helvetica Physical Acta 6, 110-127 (1933).

[15] D.N. Spergel *et al.,* Three-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Implications from cosmology. The Astrophysical Journal Supplement Series 170, 377-408 (2007).

[16] C. Adami, A. Mazure, P. Katgert, and A. Biviano, The ESO Nearby Abell Cluster Survey. VII. Galaxy density profiles of rich clusters of galaxies. Astronomy and Astrophysics 336, 63-82 (1998).

[17] P. Katgert, A. Biviano, and A. Mazure, The ESO Abell Cluster Survey. XII. The mass and mass-to-light ratio profiles of rich clusters. The Astrophysics Journal 600, 657-669 (2004).

[18] A.D. Lewis, D.A. Buote, and J.T. Stocke, *Chandra* observations of A2029: The dark matter profile down to below $0.01r_{vir}$ in an unusually relaxed cluster. The Astrophysical Journal 586, 135-142 (2003).

[19] D.E. Johnston *et al.,* Cross-correlation weak lensing of SDSS galaxy clusters II: cluster density profiles and the mass-richness relation. arXiv:0709.1159v1 [astro-ph] (2007).

[20] D. Clowe, M. Bradač, A.H. Gonzalez, M. Markevitch, S.W. Randall, C. Jones, and D. Zaritsky, A direct empirical proof of the existence of dark matter. The Astrophysical Journal 648, L109-L113 (2006).

[21] E. Kolb and M. Turner, The Early Universe. Menlo Park: Addison-Wesley (1989).

[22] G. Jungman, M. Kamionkowski, and K. Griest, Supersymmetric dark matter. Physics Reports 267, 195-373 (1996).

[23] G. Servant and T.M.P. Tait, Is the lightest Kaluza-Klein particle a viable dark matter candidate? Nuclear Physics B 650, 391 (2003).

[24] M.W. Goodman and E. Witten, Detectability of certain dark-matter candidates. Phys. Rev. D 31, 3059 (1985).

[25] 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).

[26] G. Dūda, A. Kemper, and P. Gondolo, Model independent form factors for spin independent neutralino-nucleon scattering from elastic electron scattering data. JCAP 2007, 12 (2007).

[27] D.S. Akerib et al., Limits on spin-dependent WIMP-nucleon interactions from the Cryogenic Dark Matter Search, Phys. Rev. D 73, 011102 (2006), astro-ph/0509269.

[28] A.M. Green, WIMP direct detection: astrophysical input. New Astronomy Reviews 49, 181 (2005).

[29] L.M. Krauss and C.J. Copi, Different halo models and Dama/CDMS/Edelweiss and future detectors. New Astronomy Reviews 49, 185 (2005).

[30] J. Carr, G. Lamanna, and J. Lavalle, Indirect detection of dark matter. Rep. Prog. Phys. 69, 2475-2512 (2006).

[31] H.A. Meyer-Hasselwander *et al.*, High-energy gamma-ray emission from the Galactic Center. Astron. Astrophys. 335, 161-172 (1998).

[32] A. Cesarini, F. Fucito, A. Lionetto, A. Morselli, and Piero Ullio, The galactic center as a dark matter gamma-ray source. Astropart. Phys. 21, 267-285 (2004).

[33] D. Horns, TeV γ-radiation from dark matter annihilation in the galactic center. Phys. Lett. B 607, 225-232 (2005).
erratum, Phys. Lett. B 611, 225 (2005).

[34] J. Ripken, D. Horns, L. Rolland, and J. Hinton. Dark matter annihilation as possible origin of the very high energy γ radiation from the galactic center measured by HESS. 29th International Cosmic Ray Conference in Pune 00, 101–106 (2005).

[35] M. Cassé, P. Fayet, S. Schanne, B. Cordier, and J. Paul. INTEGRAL and light dark matter. Proceedings of the 5th INTEGRAL Workshop, Feb. 16-20, 2004, Munich. astro-ph/0404490.

[36] J. Carson, GLAST: physics goals and instrument status. Journal of Physics: Conference Series 60, 115 (2007).

[37] H. Kubo *et al.,* Status of the CANGAROO-III project. New Astronomy Reviews 48, 323 (2004).

[38] D. Ferenc and the MAGIC collaboration, The MAGIC gamma-ray observatory. NIM A 553, 274 (2005).

[39] S. Desai *et al.,* Search for Dark Matter WIMPs using Upward Through-going Muons in Super-Kamiokande. Phys.Rev. D 70, 083523 (2004)
erratum, Phys. Rev. D 70 109901 (2004).

[40] D. Finkbeiner, WMAP Microwave Emission Interpreted as Dark Matter Annihilation in the Inner Galaxy. arXiv:astro-ph/0409027v2 (2004).

[41] D. Hooper, D. Finkbeiner, and G. Dobler, Evidence of dark matter annihilations in the WMAP haze. arXiv:0705.3655v1 [astro-ph] (2007).

[42] P. Frame. Environmental and background radiation - in the body. Health Physics Society. `http://www.hps.org/publicinformation/ate/cat10.html` (2002).

[43] S.R. Golwala, Ph.D. thesis, University of California, Berkeley, 2000. URL http://cosmology.berkeley.edu/preprints/cdms/golwalathesis/thesis.pdf

[44] C. Jacoboni, F. Nava, C. Canali, and G. Ottaviani, Electron drift velocity and diffusivity in germanium. Physical Review B 24, 1014 (1981).

[45] L. Reggianai, C. Canali, F. Nava, and G. Ottaviani, Hole drift velocity in germanium. Physical Review B 16, 2781 (1977).

[46] C. Canali, C. Jacoboni, F. Nava, G. Ottaviani, and A. Alberigi-Quaranta, Electron drift velocity in silicon. Physical Review B 12, 2265 (1975).

[47] G. Ottaviani, L. Reggiani, C. Canali, F. Nava, and A. Alberigi-Quaranta, Hole drift velocity in silicon. Physical Review B 12, 3318 (1975).

[48] W. Sasaki and M. Shibuya, Experimental evidence of the anisotropy of hot electrons in n-type germanium. J. Phys. Soc. Japan 11, 1202-1203 (1956).

[49] K.M. Sundqvist and B. Sadoulet. Detectors of the Cryogenic Dark Matter Search: Charge transport and phonon emission in Ge $\langle 100 \rangle$ crystals at 40 mK. To appear in proceedings of Low Temperature Detectors 12 (2007).

[50] B. Neganov and V. Trofimov. Otkrytiya, Izobreteniya 146, 215 (1985).

[51] P.L. Luke. J. Appl. Phys. 64, 6858 (1988).

[52] T. Shutt, Ph.D. thesis, University of California, Berkeley, 1993.

[53] V. Mandic, Ph.D. thesis, University of California at Berkeley, 2004.

[54] A.R. Al-Derzi, R.L. Johnston, J.N. Murrell, and J.A. Rodriguez-Ruiz. Potential energy functions for atomic solids III. Fitting phonon frequencies and elastic constants of diamond structures. Molecular Physics 73, 265 (1991).

[55] B. Cabrera, B.L. Dougherty, A.T. Lee, M.J. Penn, J.G. Pronko, and B.A. Young. Prompt phonon signals from particle interactions in Si crystals. Journal of Low Temperature Physics 93, 365 (1993).

[56] S. Tamura, J.A. Shields, and J.P. Wolfe. Lattice dynamics and elastic phonon scattering in silicon. Phys. Rev. B 44, 3001 (1991).

[57] J.A. Shields, S. Tamura, and J.P. Wolfe. Elastic scattering of acoustic phonons in Si. Phys. Rev. B 43, 4966 (1991).

[58] S. Tamura. Isotope scattering of dispersive phonons in Ge. Phys. Rev. B 27, 858 (1983).

[59] S. Tamura and H.J. Maris. Spontaneous decay of TA phonons. Phys. Rev. B 31, 2595 (1985).

[60] S. Tamura. Spontaneous decay rates of LA phonons in quasi-isotropic solids. Phys. Rev. B 31, 2574 (1985).

[61] B.A. Young, B. Cabrera, and A.T. Lee, Observation of ballistic phonons in silicon crystals induced by α particles. Phys. Rev. Lett. 64, 2795 (1990).

[62] H.J. Maris. Phonon propagation with isotope scattering and spontaneous anharmonic decay. Phys. Rev. B 41, 14 (1990).

[63] G.A. Northrop and J.P. Wolfe, Ballistic phonon imaging in germanium. Phys. Rev. B. 22, 6196 (1980).

[64] S. Tamura. Quasidiffusive propagation of phonons in silicon: Monte Carlo calculations. Phys. Rev. B 48, 18 (1993).

[65] A.T. Lee, B. Cabrera, B.L. Dougherty, M.J. Penn, J.G. Pronko, and S. Tamura. Measurements of the ballistic-phonon component resulting from nuclear and electron recoils in crystalline silicon. Phys. Rev. B 54, 3244 (1996).

[66] B.A. Young, B. Cabrera, A.T. Lee, and B.L. Dougherty, Detection of elementary particles using silicon crystal acoustic detectors with titanium transition edge phonon sensors. Nucl. Instrum. Method. A 311, 195 (1992).

[67] K.D. Irwin, An application of electrothermal feedback for high resolution cryogenic particle detection. Appl. Phys. Lett. 66 (15), 1998 (1995).

[68] J.C. Mather, Bolometer noise: nonequilibrium theory. Appl. Optics 21, No. 6, 1125 (1982).

[69] R.W. Ogburn, Quiescent resistance from saturated pulses. CDMS R118+R119 internal note #52 (2006).

[70] Y.G. Shen and Y.W. Mai, Structure and properties of stacking faulted A15 tungsten thin films. Journal of Materials Science 36, 93-98 (2001).

[71] B.A. Young, S.W. Nam, P.L. Brink, B. Cabrera, B. Chugg, R.M. Clarke, A.K. Davies, and K.D. Irwin, Technique for fabricating tungsten thin film sensors with $T_c \leq 100$mK on germanium and silicon substrates. IEEE Trans. Appl. Supercond. 7, 3367 (1997).

[72] B.A. Young, T. Saab, B. Cabrera, J.J. Cross, R.M. Clarke, R.A. Abusaidi, Measurement of T_c suppression in tungsten using magnetic impurities. Journal of Applied Physics 86, 6975-6978 (1999).

[73] S. Kamat, Ph.D. thesis, Case Western Reserve University, 2005.

[74] D. Driscoll, Ph.D. thesis, Case Western Reserve University, 2004.

[75] R. Abusaidi *et al.* Exclusion limits on the WIMP-nucleon cross section from the Cryogenic Dark Matter Search. Phys. Rev. Lett. 84, 5695 (2000).

[76] D. Abrams *et al.* Exclusion limits on the WIMP-nucleon cross section from the Cryogenic Dark Matter Search. Phys. Rev. D 66, 122003 (2002).

[77] T. Saab, Ph.D. thesis, Stanford University, 2002.

[78] Akerib *et al.* New results from the Cryogenic Dark Matter Search experiment. Phys. Rev. D 68, 082002 (2003).

[79] J.L. Thron, The Soudan-II proton decay experiment. Nucl. Instrum. Meth. A 283, 642 (1989).

[80] D.G. Michael *et al.,* Observation of muon neutrino disappearance with the MINOS detectors in the NuMI neutrino beam. Phys. Rev. Lett. 97, 191801 (2006).

[81] R. Bunker, Ph.D. thesis, University of California at Santa Barbara, 2008.

[82] J. Sander, Ph.D. thesis, University of California at Santa Barbara, 2007.

[83] C.L. Chang, Ph.D. thesis, Stanford University, 2005.

[84] V. Mandic *et al.* Position dependence in the CDMS II ZIP detectors. Low Temperature Detectors, Eds. F.S. Porter, D. McCammon, M. Galeazzi, and C.K. Stahle, AIP Conference Proceedings 605, 509 (2002).

[85] D.S. Akerib *et al.,* First Results from the Cryogenic Dark Matter Search in the Soudan Underground Lab, Phys. Rev. Lett. 93, 211301 (2004), astro-ph/0405033.

[86] D.S. Akerib *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), astro-ph/0507190.

[87] 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), astro-ph/0509259.

[88] G. Wang, Ph.D. thesis, Case Western Reserve University, 2004.

[89] R. Clarke, Ph.D. thesis, Stanford University, 1999.

[90] A. Broniatowski, N. Mirabolfathi, L. Dumoulin, A. Julliard, and L. Bergé, Electronic transport and event localization in germanium low-temperature detectors for the EDELWEISS dark matter experiment. Nuclear Instruments and Methods in Physics Research A 444, 327-330 (2000)

[91] A. Browniatowski, A. Julliard, L. Bergé, and L. Dumoulin, Pulse-shape analysis of ionization events in low-temperature germanium detectors. Low Temperature Detectors, ed. F. S. Porter et al., 2002.

[92] A. Browniatowski, B. Censier, A. Julliard, L. Bergé, and the Edelweiss collaboration, Cryogenic germanium detectors for dark matter search: Surface event rejection by charge measurements. Nuclear Instruments and Methods in Physics Research A 559, 378-380 (2006).

[93] N. Mirabolfathi, Ph.D. thesis, University of Paris / UFR Scientifique d'Orsay, 2002. `http://edelweiss.in2p3.fr/pub/these-NMirabolfathi.pdf`

[94] M. Sahani, Ph.D. thesis, Caltech, 1999. `http://www.gatsby.ucl.ac.uk/~maneesh/thesis/`

[95] S. Leclerq, private communication, 2006.

[96] A. Reisetter, Ph.D. thesis, University of Minnesota, 2005.

[97] M. Attisha, Ph.D. thesis, Brown University, 2006.

[98] J. Cooley and R. Mahapatra, Using Linear Discrmininant/Fisher Discriminant Analysis Technique in Timing Cut Definition. CDMS Run 119 internal note 339 (2006).

[99] D. Grant, Signal Extraction of WIMP Search Multiples; Calculated using PCA method. CDMS Run 118+Run 119 internal note 175 (2007).

[100] The CDMS collaboration, CDMS backgrounds and Monte Carlo studies. In preparation.

[101] S.J. Yellin, Finding an upper limit in the presence of an unknown background. arXiv: physics/0203002v2 (2002).

[102] CDMS collaboration, A search for WIMPs with the first five-tower data from CDMS. arXiv:0802.3530v1 [astro-ph] (2008).

[103] 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. A 559, 414 (2006).

[104] P.L. Brink *et al.,* Beyond the CDMS-II dark matter search: SuperCDMS. arXiv:astro-ph/0503583v1 (2005).

[105] R. Gaitskell, V. Mandic, and J. Filippini, Direct detection limit plotter. `http://dmtools.berkeley.edu/limitplots/`.

[106] J. Angle *et al.,* First results from the XENON10 dark matter experiment at the Gran Sasso National Laboratory. Phys. Rev. Lett. 100, 021303 (2008).

[107] L. Roszkowski, R. Ruiz de Austri, and R. Trotto. Implications for the constrained MSSM from a new prediction for b→sγ. Journal of High Energy Physics 200707, 075 (2007).

[108] E.A. Baltz and P. Gondolo, JHEP 200410, 052 (2004).

[109] J.A. Nelder and R. Mead, A simplex method for function minimization, Comput. J 7, 308-313 (1965).

[110] N.J. Higham. The Matrix Computation Toolbox. `http://www.ma.man.ac.uk/~higham/mctoolbox/`

[111] Y.P. Varshni. Temperature dependence of the energy gap in semiconductors. Physica 34, 149 (1967).

[112] C.D. Thurmond. The Standard Thermodynamic Functions for the Formation of Electrons and Holes in Ge, Si, GaAs, and GaP. J. Electrochem. Soc 122, 1133 (1975).

[113] W. Shockley, Problems related to p-n junctions in silicon. Solid State Electronics 2, 35 (1961).

[114] F.A. Emery and T.A. Rabson, Average energy expended per ionized electron-hole pair in silicon and germanium as a function of temperature. Physical Review 140, A2089 (1965).

[115] Claude A. Klein. Bandgap dependence and related features of radiation ionization energies in semiconductors. Journal of Applied Physics 39, 2029 (1968).

[116] Claude A. Klein. Semiconductor particle detectors: a reassessment of the Fano factor situation. IEEE Transactions Nucl. Sci. 15, 214 (1968).

[117] J.R. Hauser. Threshold energy for avalanche multiplication in semiconductors. J. Appl. Phys. 37, 507 (1966).

[118] R.H. Pehl, F.S. Goulding, D.A. Landis, and M. Lenzlinger. Accurate determination of the ionization energy in semiconductor detectors. Nucl. Instrum. and Methods 59, 45 (1968).

[119] T. Shutt et al., Simultaneous high resolution measurement of phonons and ionization created by particle interactions in a 60g germanium crystal at 25 mK. Physical Review Letters 69, 3531 (1992).

[120] X.-F. Navick, M. Chapellier, F. Déloit, S. Hervé, and L. Miramonti, 320 g ionization-heat bolometers design for the EDELWEISS experiment. Nuclear Instruments and Methods in Physics Research A 44, 361 (2000).

[121] B.A. Young and K.M. Yu, Ion-implanted charge collection contacts for high purity silicon detectors operated at 20 mK. Rev. Sci. Instrum. 66, 2625 (1995).

[122] Betty A. Young, personal communication, 2007.

[123] L.P. Ekström and R.B. Firestone, WWW Table of Radioactive Isotopes, database version 2/28/99 from URL `http://ie.lbl.gov/toi/index.htm`.