Advanced Analysis and Background Techniques for the Cryogenic Dark Matter Search

A DISSERTATION

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL

OF THE UNIVERSITY OF MINNESOTA

BY

Xinjie Qiu

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

Doctor Of Philosophy

December, 2009

To my parents

and

To my wife Wenhua Meng

Contents

List of Tables

List of Figures

Preface

"Rome wasn't built in one day", and neither was the CDMS experiment. The CDMS experiment is a collaborative effort of many institutions, scientists and engineers. Not all the work described in this thesis is the author's own. This preface summarizes my own major contribution to the CDMS collaboration.

Soudan operations

I began taking shifts at Soudan Underground Laboratory during the installation and commissioning of the 5-Tower data runs from the summer in 2005. Most of my cryogenics and DAQ system knowledge were learned for the first time while there from numerous experienced and knowledgeable scientists and engineers. This valuable experience trained me to be an independent shift leader for the later CDMS runs at Soudan. Since then, I was one of the people who spent most time at Soudan to operate the experiment each year, due to my willingness to learn more, and due to my proximity to the Soudan Lab. I spent Christmas Eve and New Years during the winter of 2006, 2007, 2008 at Soudan to maintain uninterrupted system running, and enjoyed the peace and quiet of the small town as well.

Monte Carlo Simulation

Monte Carlo simulation based on the Geant4 package is a critical study required to characterize the response and performance of our detectors. I continued in this work throughout of my Ph.D. period. The Minnesota group has focused on the neutron simulations. My Monte Carlo simulation work included the study of nuclear recoil spectra, nuclear recoil band corrections due to the inelastic scattering, fission neutrons in the shielding, fiducial volume cut error estimation, and the modeling of nuclear recoil energy scale.

Data analysis

My major focus with respect to data analysis was the Run 119 data set (two tower run), the combined Run 118/119 data set (one tower + two tower run), and the most recently released Run 123/124 data set (5 tower run). I took a major role in the Run 123/124, in data collection, quality control, data analysis development and coordination. I developed the classic style 2D discrimination technique, as described in [**?**]. This classic 2D surface event cut was the primary analysis for the Run 123/124 data, and the result was published in Physical Review Letters [5]. I was the lead author on the corresponding publication. I also developed the improved the neural network discrimination technique, as described in [**?**], which proved to be useful for the future analysis.

Detector Surface Characterization

To maintain a background-free experiment, CDMS takes the improved understanding of background seriously. To determine the source of the surface contamination on CDMS ZIP detectors, Minnesota group utilized Ion Beam Analysis(IBA) and particle

induced X-ray emission(PIXE) at the Characterization Facility at University of Minnesota.

Chapter 1

Dark Matter Evidence and Candidates

The existence of dark matter is well established based on the cosmological and astrophysical observations over the last few decades. The theory of dark matter is widely accepted by the scientific community due to its successful explanation of the observed galactic rotation, structure formation and galaxy evolution, gravitational lensing by galaxy clusters, and the effects on the anisotropy of the cosmic microwave background, etc [6, 7]. Since the Standard Model of particle physics does not provide any satisfactory dark matter candidate, the search for physics beyond the Standard Model is one of the well motivated reasons to propose promising dark matter candidates.

1.1 Evidence for dark matter

1.1.1 Galactic rotation curves

The early and probably still the most convincing evidence for the dark matter is from the observations of galactic rotation curves, which is the velocity of rotation versus the

distance from the center of galaxy.

From Newtonian dynamics, rotational velocity can be predicted by setting gravitational force equals to centripetal force due to the mass inside the Keperian orbit with radius r around a galaxy.

$$\frac{GmM(r)}{r^2} = \frac{mv^2}{r} \tag{1.1}$$

The circular velocity as a function of radius is expected to be

$$v(r) = \sqrt{\frac{GM(r)}{r}} \tag{1.2}$$

where $M(r) = 4\pi \int \rho(r) r^2 dr$.

If r is outside of the visible part of the galaxy, and light traces the galaxy's matter distribution, velocities of stars should decrease with radius as $v(r) \propto \frac{1}{\sqrt{r}}$. However, experiments found the velocities $v(r)$ in most galaxies are approximately constant. Figure 1.1 shows that observational rotation curves of six spiral galaxies extend to large radii [1], with much larger velocity than expected from the above equations. This cannot be explained by the visible matter inside the galaxies, thus implying the existence of a centrally-concentrated spherically-symmetric dark halo with $M(r) \propto r$, and $\rho(r) \propto \frac{1}{r^2}$, with luminous matter in a disc at the center.

1.1.2 Galaxy clusters

The observations of clusters of galaxies indicate the dark matter affects galaxy clusters as well.

The first evidence of the existence of dark matter is from the measurement of the velocity dispersion in the Coma galaxy clusters by Fritz Zwicky [8, 9]. The mass-to-light ratio is measured by Doppler effect to be around 400 for solar masses per solar luminosity, using the virial theorem. Based on these measurements, Zwicky suggested there must be some unseen, invisible "missing mass" providing enough mass and gravity

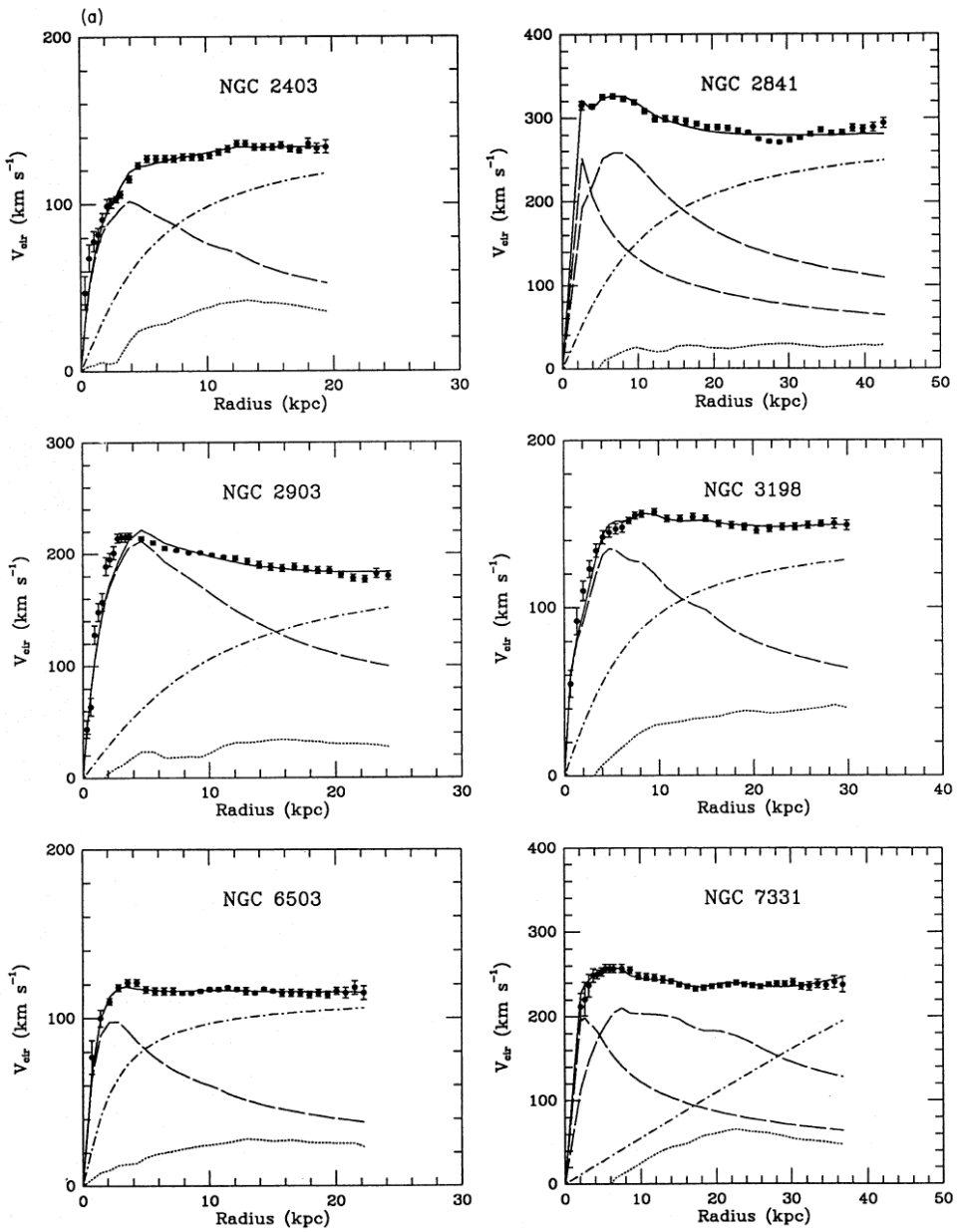


Figure 1.1: Rotation curves of six spiral galaxies. Different curves represents contributions from different components. The dashed values are for the visible components, the dotted curves for the gas, and the dash-dotted curves for the dark matter halo. Figure from [1].

to hold the cluster together. Modern measurements have confirmed the idea that galaxy clusters are dominated by dark matter.

Since then, more measurements from the observation of clusters of galaxies allow us to directly estimate the gravitational mass based on the light from galaxies. X-ray temperature measurements of hot intracluster gas inside the cluster estimate the mass-to-light ratios for large clusters to be about 10 to 1. This independent determination of the mass of clusters gives a result consistent with other methods, such as cluster dynamics and gravitational lensing.

1.1.3 Gravitational lensing

Study of the gravitational lensing of background galaxies by clusters is another important tool [10]. Lensing uses the equations of general relativity to predict masses, which is an independent method to measure dark matter. There are two kind of lensing effects. When mass density of the lens is large, "strong lensing" results in a bending and distortion of background galaxies behind the cluster into giant arcs surrounding the brightest galaxy, when light from these galaxies passes through the gravitational lens. It has been observed for a few distant clusters including Abell 2218 in Figure 1.2.

Figure 1.2: Image taken with the Hubble Space Telecope (HST) of a cluster "Abell 2218" shows an example of a strong lensing by clusters.

The weak gravitational lensing, on the other hand, looks at subtle lensing distortions of observed galaxies in galactic surveys using statistical analyses to reconstruct the distribution of dark matter and to find mass-to-light ratios [11].

The separation between dark matter and baryonic matter in the recent weak-lensing observations of Bullet Cluster 1E 0657-558 [2], provides evidence of dark matter. A collision between two galaxy clusters were observed in both X-rays and lensing techniques after they merged. In Figure 1.3, gravitational lensing maps the distribution of dark matter galaxies, rather than the dominant baryonic mass component (measured by X-ray emitting plasma). Since dark matter haloes do not interact by electromagnetic forces, this component of the two clusters passes through each other, while the gas particles slow down near the collision interaction region.

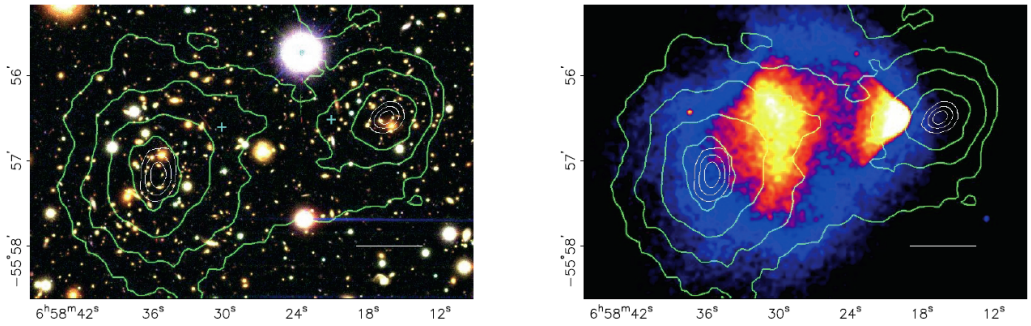


Figure 1.3: Image of the Bullet Cluster (1E0657-558) from the Hubble Space Telescope with total mass contours (dominated by dark matter) from a lensing analysis overlaid. Green contours in both panels are the weak-lensing reconstructions. The blue plus signs show the mass centers of the plasma clouds. Left: Optical photograph from the Hubble Space Telescope. Right: X-ray image from the Chandra X-Ray Observatory. Figure from [2].

Both strong and weak gravitational lensing provides strong evidences that the majority component of the universe consists of invisible dark matter.

1.1.4 Cosmic microwave background

The Big Bang theory of the universe proposes that the source of the CMB is from blackbody radiation left over from the Big Bang. In the early universe, photons, electrons and baryons filled the universe in a small hot condensed state. They frequently interacted with each other. As the universe expanded, the plasma cooled until electrons combined with protons and formed hydrogen atoms. This recombination happened around 379,000 years ("time of last scattering") after the Big Bang, when the thermal radiation was no longer absorbed by atoms; they were decoupled from atoms and began to propagate freely through space ever since. As the universe continued to expand, the temperature of the radiation photons continued to cool down, enabled more complex structures to form, like stars, planets and galaxies.

Although the cosmic background radiation in general is quite uniform and isotropic at the 10^{-5} level at the current temperature of 2.725 K as measured by NASA Cosmic Background Explorer (COBE) satellite, precise measurements show small temperature fluctuations superimposed on this average [12], shown in Figure 1.4.

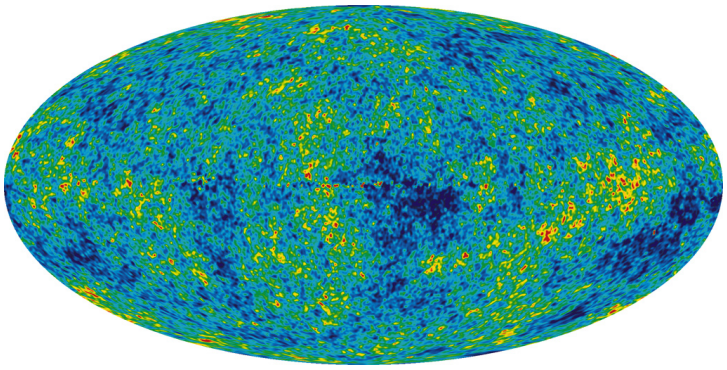


Figure 1.4: WMAP image of the CMB temperature anisotropy from the 5-year WMAP data seen over the full sky. The color shows ~ 10 μK temperature fluctuations on the 2.725 K background radiation. Figure from http://map.gsfc.nasa.gov/.

The almost perfect black body spectrum and its detailed prediction of the anisotropies

in the CMB is expected from Big Bang theory. By taking Fourier transform to decompose the CMB map into spherical harmonics, the information in CMB maps can be compressed into an angular power spectrum of the anisotropy. It is obtained and plotted in Figures 1.5, measured by the recent Wilkinson Microwave Anisotropy Probe (WMAP) [13] and several other experiments.

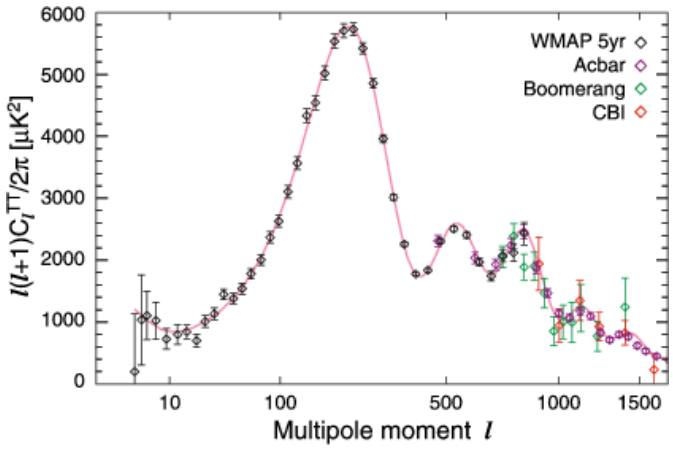


Figure 1.5: The WMAP 5-year power spectrum of primary CMB anisotropies, along with other recent experiment results.

The peaks of CMB power spectrum contain important information, enabling us to accurately test cosmological models, and make estimates of the parameters in the standard ΛCDM cosmological model to answer fundamental questions about the Universe, such as the shape and contents of the Universe. The first peak determines the curvature of the Universe. The second peak determines the reduced baryon density. The third peak can be used to extract information about the dark matter density.

This data is perfectly consistent with a flat universe modern cosmological model, where dark energy provides 72% of the total density of the Universe, 23% comes from dark matter, and only 4.6% from baryonic ordinary matter.

1.2 Dark matter candidates

The evidence from different aspects mentioned in the last section confines the properties of non-baryonic dark matter that candidates for non-baryonic dark matter must satisfy several conditions: (1) they must be cold, (2) they must be stable, and (3) they must interact at the electroweak scale. Being "cold" means that they are non-relativistic during the galaxy formation. This is necessary to explain the large-scale structure of the universe. They need to be stable because they are still around today. Their weak interaction property seems plausible since the electroweak scale cross section fits so well with early Universe thermal equilibrium conditions (details in section 1.2.3) Certainly they cannot interact electromagnetically or strongly as they are not undetectable by the emitted radiation. Some candidates were proposed to be the possible dark matter, which are listed below.

1.2.1 Neutrinos

Neutrinos are an obvious candidate for non-baryonic dark matter, since they are weakly interacting particles with confirmed observations of masses from solar mixing [14] and atmospheric mixing [15]. Thus neutrinos must constitute some of the universe's non-baryonic dark matter. However, the experiment constraint on neutrino masses implies an upper bound on the total neutrino density of

$$\Omega_\mu h^2 \lesssim 0.07 \tag{1.3}$$

which means neutrinos are not abundant enough to account for the dominant component of dark matter.

In addtion, the relativistic collisionless property of neutrinos makes it difficult to explain the observed large scale structure, since "hot" neutrinos would freely move from high to low density regions and remove density fluctuations.

1.2.2 Axions

A hypothetical elementary particle called the axion was introduced in particle physics in an attempt to prevent strong CP violation in quantum chromodynamics [16]. the axion is the result of the spontaneously broken of Peccei-Quinn symmetry, which leaves a pseudo-Nambu-Goldstone boson field.

The nonthermal axion production mechanism in the early universe provides a cold dark matter candidate despite their low masses. Astrophysical observations are currently the best strategy to search for axions [17]. Observations to detect axions are not yet sensitive enough to probe the mass regions where axion mass should be if they are supposed to be dark matter particles.

Recently CDMS collaboration reported first axion search results from Soudan 5-tower run [18]. No indication of solar axion conversion to photons was observed, which set a world leading upper limit on the coupling constant $g_{\alpha\gamma\gamma}$.

1.2.3 Weakly interacting massive particles

Weakly Interacting Massive Particles(WIMPs) are particles with mass roughly between 10 GeV/c^2 and a few TeV/c^2 range. These particles interact through the weak force and gravity. It's a class of particle candidates that were in thermal equilibrium in the early universe. WIMPs can be produced through reactions from the standard model particles, and also pair-annihilated into these particles. As long as $T > m_\chi$, the WIMP number density is comparable to the number density of electrons, positrons, and photons. However, when the unverse expanded, the temperature dropped below m_χ, the WIMPs dropped out of thermal equilibrium, causing WIMP abundance to drop. When the WIMP number density is so low that the WIMP mean free time for annihilation exceeds the age of the universe, then "freeze-out" occurs at a density determined by the

WIMP annihilation cross-section and implies that

$$\Omega_\chi h^2 \simeq \frac{0.1pb \cdot c}{< \sigma_A v >} \tag{1.4}$$

where σ_A is the total annihilation cross section of a pair WIMPs into standard model particls, and v is the relative velocity between two WIMPs.

The current most popular and motivated WIMP candidate is the lightest super-particle (LSP) in supersymmetric models [19]. Supersymmetry is an extension of the standard model of particle physics. As a new symmetry in supersymmetry, R parity is preserved, the LSP cannot decay. It can imply the existence of a new stable particle, "neutralino".

The problem of searching for WIMPs is that they interact with ordinary matter rarely, which makes them difficult to detect. In the rest of this thesis, I will discuss how CDMS uses state-of-the-art detectors to probe the existence of WIMP dark matter particles.

Chapter 2

CDMS Detectors

The Cryogenic Dark Matter Search (CDMS) uses Z-sensitive Ionization and Phonon (ZIP) solid state semiconductor detectors, operating at low temperature ($\sim$ 40mK), to simultaneously measure both phonon and ionization energy. This technique allows event-by-event discrimination between electron recoil background events and nuclear recoil signal events.

2.1 Particle interactions

A particle interacts with ZIP detector either with electrons in the crystal by Compton scattering, K-capture, phonoelectric("electron recoil") or with a nucleus ("nuclear recoil"). Any interaction deposits energy in the crystal through charge excitations and lattice vibrations. Depending on the material and the type of recoil, 6% - 33% of the recoil energy is first converted into ionization before subsequent conversion to phonons. Charge carriers are drifted by an electric field (3V/cm for Ge, 4V/cm for Si) applied on the electrodes across the crystal, and collected on two concentric electrodes on the bottom surface. Athermal phonons are collected on the top surface. A CDMS ZIP

detector measures both ionization and phonon energy for every event.

2.2 Zip detectors

Each CDMS ZIP detector is a cylindrical high-purity Ge or Si crystal, 7.6 cm in diameter and 1 cm thick. Each single Ge (Si) ZIP has a mass of 250 g (100 g). A tower is composed of six ZIP detectors. Within a tower, the six ZIP detectors are stacked 2 mm apart with no intervening material. A CDMS ZIP detector with phonon sensor shown is in Figure 2.1.

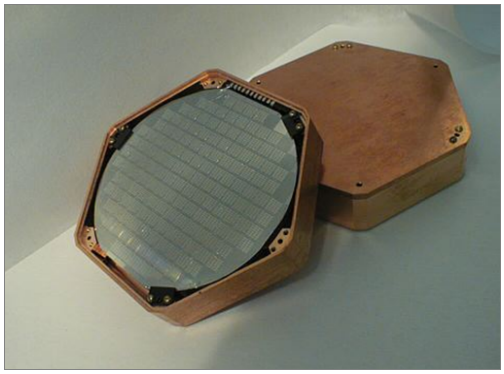


Figure 2.1: A ZIP detector with its hexagonal copper housing and photolithographic phonon sensors.

2.3 Ionization signal and measurement

Each ZIP detector has two ionization channels, called the inner and outer electrodes. These two concentric ionization electrodes and four independent phonon sensors are photolithographically patterned onto each side of crystals. The disk-shaped inner electrode defines an inner fiducial region, which covers $\sim 85\%$ of the detector bottom surface, and is shielded from low energy electron sources on surrounding surfaces. The outer electrode ring serves as a "guard" electrode to reject events occurring near the unpolished

edges of the crystal. These near-edge events are rejected due to the following reasons: (1) the ionization signal may be degraded since the electric field is not uniform in this region, (2) the phonon response is worse since this region is not well covered by the phonon sensors, and (3) background interactions are more likely due to the absence of self-shielding by other detectors in the stack.

On average, one electron-hole pair is produced for every $\epsilon \approx 3.0$ eV (3.8 eV) of energy deposited by an electron recoil in Ge (in Si). The ionization energy E_Q is defined for convenience as the recoil energy inferred from the detected number of charge pairs, N_Q, by assuming that the event is an electron recoil with 100% charge-collection efficiency:

$$E_Q = \epsilon N_Q \tag{2.1}$$

Ionization energy is usually reported in units of "keVee", keV of the equivalent electron recoil determined from electron-recoil calibration measurements.

Electrons and holes produced by an interaction are drifted to the detector faces under the applied electric field. The number of charges collected on the electrodes, N_Q, is proportional to the number of electron-hole pairs produced by the interaction weighted by the drift distance of each of the charges [20]. In the case where all of the electrons and holes drift across the crystal, N_Q (see Eq. 2.1) equals the total ionization. The prompt recombination of the drifted electrons and holes in the electrodes releases its energy from the electron system into the phonon system.

Periodic LED flashing is required to neutralize the crystals, since impurity sites can be left with a net charge at low temperatures. These charged impurity sites can trap the drifting electrons or holes from a recoil event, resulting in underestimates of the total ionization. Neutralizing the Si detectors takes more flashing cycles than for Ge, thus, some Si crystals suffer incomplete charge collection due to their more impurities sites than the Ge.

For events close to one of the charge electrodes, the relatively low applied electric field

and self-screening effect make some of the electrons or holes drift into the "incorrect" electrode, which is called "back-diffusion". This can be substantially reduced [21, 22] by depositing a thin ($\sim$ 40 nm) layer of lightly doped amorphous silicon between each electrode and the detector surface.

Charge collection is also reduced for interactions within the 10 μm "dead layer". These events are called "surface events". Timing and energy information from the athermal phonon measurement (see next section) makes it possible to identify and reject these surface events. The surface event rejection analysis is discussed in detail in section 6.4.6 and advanced analysis in Chapter 5.

2.4 Phonon signal and measurement

There are two different mechanisms for producing phonons: primary phonons directly from the particle interaction and Luke phonons from electron and hole charge carriers drifting through the crystal.

- Particle interaction phonons :

  Most of the energy is deposited into the phonon system as the scattered nuclei and electrons return to rest. Phonons move quasi-diffusively and this type of phonon is dominated by two processes: (1) anharmonic decay and (2) isotopic scattering [23, 24]. In the anharmonic decay, a phonon decays into two phonons of smaller energy. This process occurs at the rate of 5th power of frequency:

  $$\tau_d \propto (\frac{1THz}{\nu})^5 \tag{2.2}$$

  where τ_d is the anharmonic decay time constant and ν is the phonon frequency.

Isotopic scattering occurs when phonons elastically scatter with isotopic impurities. It occurs with a rate proportional to the 4th power of frequency:

$$\tau_I \propto (\frac{1THz}{\nu})^4 \tag{2.3}$$

where τ_I is the isotopic scattering time constant.

- Luke phonons :

  As electrons and holes drift across the crystal due to an applied electric field, some of their energy is dissipated into the crystal as phonon radiation. This effect was measured by Neganov-Trofimov [25] and Luke [26]. This additional phonon energy by Luke phonons is

$$E_{Luke} = eV_b \sum_i \frac{d_i}{d} \tag{2.4}$$

  where V_b is the bias voltage across the detector, d is the detector thickness, d_i is the drifted distance for each charge i.

The total phonon energy from all three contribution is therefore:

$$E_P = E_R + E_{Luke} = E_R + eV_b \sum_i \frac{d_i}{d} \tag{2.5}$$

where E_R is the phonon recoil energy including both primary and carrier relaxation phonon energies. The summation in Eq. 2.5 is the number of charges weighted by their drift distances and is equal to the charge collected on the electrodes, N_Q. The definition of ionization energy given by Eq. 2.1 means that

$$E_P = E_R + eV_b N_Q = E_R + \frac{eV_b}{\epsilon} E_Q \tag{2.6}$$

Equation 2.6 is valid even for events with incomplete charge collection. Since we calibrate electron recoils with full charge collection to have $E_Q = E_R$, then we find

$E_P = (1 + eV_b/\epsilon)E_R$ for these events. In practice, the recoil energy E_R of an event is inferred from measurements of the phonon and ionization energies:

$$E_R = E_P - \frac{eV_b}{\epsilon}E_Q \tag{2.7}$$

Phonon sensors cover the top surface of the detector, using four circuits of superconducting thin films in quadrant pattern, which are composed of 4144 Quasiparticle-assisted Electrothermal feedback Transition-edge sensors (QETs) [27], operating in parallel, 1036 in each quadrant (see Figure 2.2). Each QET consisting of a 1 μm-wide, 350 μm-long tungsten strip is connected to eight aluminum quasiparticle trap fins. Once the phonon energy collected by the phonon TES tungsten sensor, it raises the temperature of the film, increasing its resistance and reducing the current, causing the electrothermal feedback which maintains its operating point within the superconducting to normal transition. The measurement of athermal phonons makes it possible to reject events in the dead layer by comparing phonon to ionization signals (section 2.5) and by measureing the risetime of the phonon pulse (section 2.6).

2.5 Yield discrimination

The dimensionless ionization yield parameter, y, is the ratio of ionization energy to true recoil energy,

$$y = \frac{E_Q}{E_R} \tag{2.8}$$

where E_R is the full recoil energy for the event. This definition of yield gives unity for electron recoil events with complete charge collection. Nuclear recoils produce fewer charge pairs than electron recoils of the same energy, and hence less ionization energy E_Q, (see Figure 2.3). The ionization yield y for nuclear recoil events depends on both the material and the recoil energy, with y $\sim$ 0.3 (y $\sim$ 0.25) in Ge (in Si) for $E_R \geq 20$

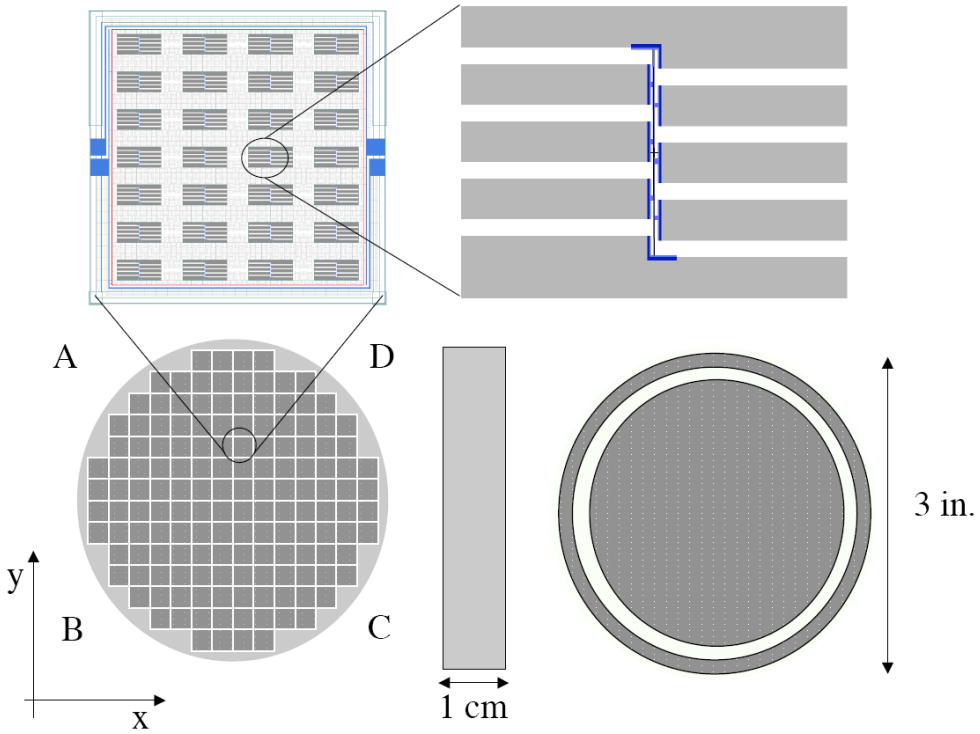
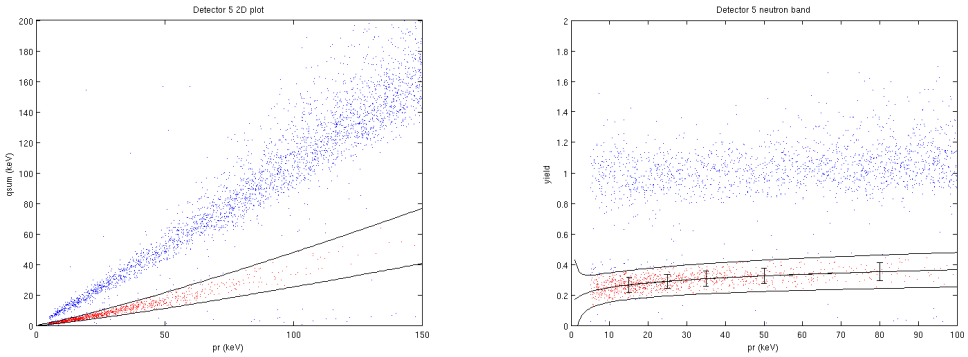


Figure 2.2: Schematic of the CDMS-II ZIP detector configuration. *Bottom right*: ionization side with inner and outer electrodes. *Bottom left*: phonon side with four sensor quadrants. *Top left*: One of the 37 photolithography dies that make up each quadrant of the phonon surface. Each die consists of 28 individual QETs. *Top right*: A single QET, consisting of a 1 μm-by-by 250 μm tungsten TES connected to aluminum collector fins.

keV. The simultaneous measurement of ionization and recoil energy therefore makes it possible to identify and reject most electron-recoil background events.



(a) Ionization energy versus recoil energy for detector T1Z5 (Ge). Blue dots correspond to electron recoil and red dots correspond to nuclear recoil events. Both are from a ^{252}Cf source which produces both gammas and neutrons.

(b) The same data plotted as yield versus recoil energy. The upper distribution of events represents the electron recoil band. The lower distribution of events define the nuclear recoil band.

Figure 2.3: Separation of nuclear and electron recoil using ionization and phonon information for detector T1Z5

2.6 Timing discrimination

This primary phonon component has both high-energy optical phonons and lower-energy acoustic phonons. High-frequency phonons from elastic scattering and anharmonic decay propagate quasi-diffusively at approximately one third the speed of sound through the crystal [28, 29]. The quasi-diffuse propagation increases the total number of phonons while reducing the average phonon frequency. When the frequency of a phonon is sufficiently low ($<$ 1 THz), the mean free path becomes comparable to the size of the detector. Such phonons are "ballistic" balls and travel at the speed of sound through the crystal. Eventually the quasi-diffuse propagation converts all of the initial high-frequency phonons into ballistic phonons, which then eventually thermalize.

In additional to the above ballistic phonons from quasi-diffuse propagation, ballistic Neganov-Trofimov-Luke phonons are produced when electron-hole pairs drift across the entire crystal under an electric field, and ballistic relaxation phonons upon relaxation at the electrodes [30].

Since an electron recoil produces more charges than a nuclear recoil, these processes lead to a sharper initial population of ballistic phonons for electron recoils. The difference in the speed of propagation leads to a faster phonon leading edge for electron recoils than for nuclear recoils because of the larger ballistic fraction. It also results in a shorter delay time between charge signal and the beginning of the phonon pulse.

Chapter 3

The Cryogenic Dark Matter Search at Soudan

3.1 Soudan Underground Laboratory

The Cryogenic Dark Matter Search (CDMS) experiment is located at 27th level, 2341 feet (713 meters, about 2090 m water equivalent depth) below the surface in the Soudan Underground Laboratory at the Soudan Underground Mine State Park. The laboratory is one of the leading deep underground science and engineering laboratories in the United States, due to its very low rate of cosmic rays in the deep underground site. The choice of deep underground mine makes it possible to reduce the cosmic ray muon flux (by a factor of $\sim$ 50,000), and the muon-induced neutron, which in turn increases the experiment sensitivity.

3.2 The cryogenic dark matter search experiment

Before moving to the Soudan mine, CDMS ZIP detectors were operated from August 2001 through July 2002 at the Stanford Underground Facility (SUF) shallow site [31] at a depth of 11 meters (17 m water equivalent depth). The first CDMS-II dark matter run with the same six detectors at the Soudan mine, occurred from October 2003 through January 2004. The second CDMS-II run with six additional detectors ran from March 2004 through August 2004.

The five-tower run, which is the subject of this thesis, ran from October 2006 until July 2007 with a total of 30 detectors (19 Ge and 11 Si) in five towers. The data set started with Run 123, which acquired data for more than 148 calendar days (107 live days of WIMP search running) between October 2006 and March 2007. An ice plug in the dilution refrigerator plumbing required a warm-up in late March. Soudan WIMP search Run 124 resumed in April 2007 until July 2007 for approximately 83 calendar days or 56 live days. Run 124 was ended for the same reason as Run 123, due to problems with a blockage in the dilution refrigerator. There were some substantial neutralization difficulties for several of the previously good detectors, due to changes in LED settings.

The total WIMP search exposure was substantially increased. The bar chart in Figure 3.1 summarizes the detector masses operated during different Soudan runs, with Ge detector mass in green and Si detector mass in yellow. The raw livetime is shown in blue bars.

3.3 Cryogenics upgrade

The CDMS-II ZIP detector phonon signal measurement with TES technology requires detectors operating at millikelvin temperature. This ultracold temperature is achieved by the use of an Oxford Instruments Kelvinox 400μW dilution refrigerator.

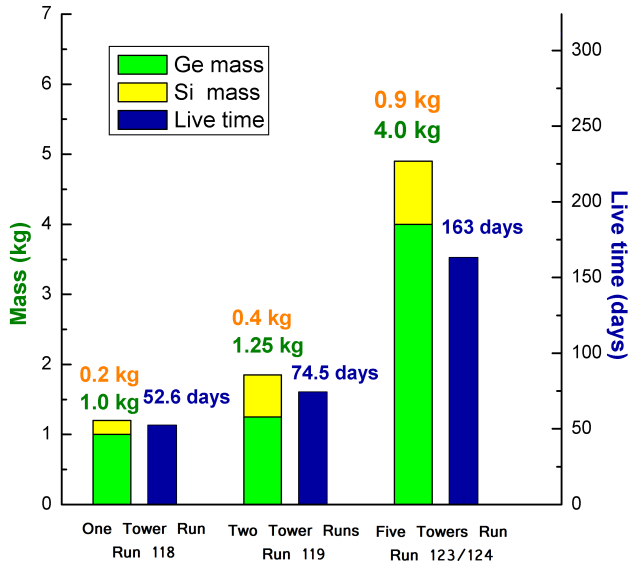


Figure 3.1:

Detectors operate inside a specially made cold volume (called "icebox") attached via a set of nested concentric extensions of the vacuum, helium and nitrogen vessels to the dilution refrigerator.

It's a challenge to have such a large detector mass operated under stable conditions for long periods of time. Addition of three new towers to the original two required a number of custom modifications to be made to the refrigerator. During the system upgrade in 2005 between the two-tower and five-tower runs, a Gifford-McMahon cryocooler (Figure 3.2) was added as a secondary refrigeration system to mount on the E-stem. The cryocooler was put on to reduce our load on helium usage and increase the cryogenic hold time. The intention was to require only one cryogen liquid helium and nitrogen baths fill per day, substantially increase the experimental live time and reduce the cost on the expensive liquid helium. We could't cool the five-tower detector down

to < 50 mK however after it was attached. Later we figured out that microphonics from the cryocooler was the reason it did not cool down. We had to replace the solid copper with copper braids to reduce the mechanical coupling. Interesting enough, it still did not get cold enough due to the loose screws on the lid. After all these were fixed, we finally made all 5 Towers operate at the desired low temperature.

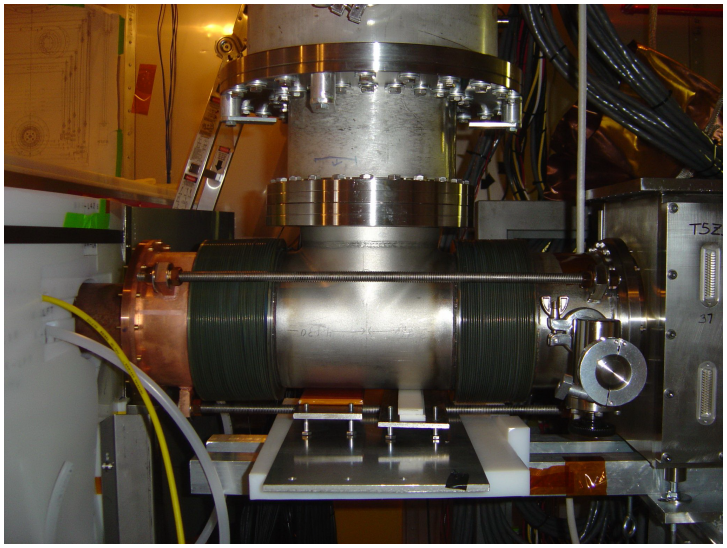


Figure 3.2: Cryocooler head (the double cylinder) attached to the E-stem.

3.4 DAQ upgrade

The CDMS Soudan data acquisition (DAQ) system controls the acquisition and storage of events. The DAQ system was written extensively in C++ and Java. It could be controlled locally from underground, or remotely from the control room at the Soudan surface building, as well as providing online diagnostic and monitoring from anywhere in the world. The user-friendly interface also featured an error reporting system to promptly send messages to the operator's attention with different emergency level tags.

Raw data was stored on the local tape in the mine and also transferred to surface cluster of computers (the Soudan Analysis Cluster or "SAC"), where the data is written to a second copy of tape for backup. The tape was later brought to the Fermilab computer cluster for final data processing. The raw data on the SAC was initially processed for the data quality monitoring purposes. A direct ethernet link between the Soudan underground mine and Fermilab was established during the later data run for faster and reliable raw data transfer. No tape backup was required after that.

To accommodate the five tower high-rate gamma calibrations from Ba source, selective readout was enabled to record data only these detectors for which triggers are issued. This allowed for a sustained acquisition throughput of greater than 50 Hz, greatly increasing the speed of detector calibration, reducing the load on the DAQ and raw data processing.

Chapter 4

CDMS Background

In event-counting experiments like CDMS, the sensitivity is approximately proportional to the exposure (the target mass and the live time). However, if there is background, then the result is proportional to the square root of the exposure with background substraction and is eventually limited by any systematic errors in the background estimate.

CDMS aims to be a background-free experiment. Therefore, it is critical for CDMS to understand the sources, to characterize the properties, to improve the rejection techniques, and to estimate the remaining level of the background. In this and the following chapters, different sources of background will be explored. In this chapter, I will describe the manner in which the background limits the signal detection sensitivity, the sources of background, and how it is reduced and the amount quantified.

4.1 How background limits WIMP detection sensitivity

To understand how the background rates in a rare-event counting experiment limit the signal sensitivity, consider an experiment with total exposure of MT (in units of

kg-day). We would like to determine the rate of signal events S and background B. Both B and S are measured in events/kg/day.

In CDMS, additional parameters, such as ionization yield (section 6.4.3) and phonon timing(section 6.4.6) provide information about the nature of the interaction in order to discriminate between nuclear recoil signal, and background. If α represents the fraction of nuclear recoil events passing the discrimination cut, and β represents the background fraction passing the same cut, then an ideal experiment would have $\alpha = 1$ for perfect signal acceptance and $\beta = 0$ for perfect background rejection.

The systematic uncertainty (4.1) and statistic uncertainty (4.2) can be written [32] as:

$$\delta S_{syst} = \left(\frac{\beta}{1-\beta}\right)\left(\frac{\delta\beta}{\beta}\right)_{syst} B \tag{4.1}$$

$$\delta S_{stat} = \sqrt{Q}\sqrt{\frac{B}{MT}} \tag{4.2}$$

The statistical uncertainty determines a statistical upper limit on the rate of WIMP search events, where

$$Q = \frac{\beta(1-\beta)}{(\alpha-\beta)^2}$$

is the background rejection quality factor, which is determined by the specific signal/background discrimination technique. The smaller the Q, the better the sensitivity.

If no WIMP candidate events are observed, the experiment can set a 90% confidence level Poisson upper limit on the WIMP recoil rate of

$$S = \frac{2.3}{MT}$$

events per kg-day.

Thus, the final estimate of the signal event rate S, if no WIMP events are observed

is

$$S = \underbrace{\frac{2.3}{MT}}_{\substack{\text{signal}\\ \text{rate}}} \pm \underbrace{\sqrt{\frac{\beta(1-\beta)}{(\alpha-\beta)^2}}\sqrt{\frac{B}{MT}}}_{\substack{\text{statistical}\\ \text{error}}} \pm \underbrace{\left(\frac{\beta}{1-\beta}\right)\left(\frac{\delta\beta}{\beta}\right)B}_{\substack{\text{systematic}\\ \text{error}}} \tag{4.3}$$

If $B << 1$, then the experiment's sensitivity improves linearly with the exposure MT, as the first term in equation 4.3. However, if the background rate B is not negligible, but is well-characterized with negligible systematic uncertainty, it may be subtracted from the observed count rate, and an excess of events above that background is considered as signal rate, resulting in sensitivity increase as the square root of exposure MT (the 2nd term in equation 4.3).

Uncertainties in determination of β and the background rate B, both contribute to the systematic uncertainties, while the contribution from α is heavily suppressed when no candidate events are observed. One systematic uncertainty comes from the difference between the β determined in the calibration run and the distribution of background events in the WIMP search run. With increasing exposure, the systematic uncertainties (equation 4.1) will begin to exceed statistical uncertainties at sufficiently large exposures, so the experimental sensitivity ceases to improve with exposure. The experiment has to either seek improved background rejection techniques to better characterize the background (small Q), or reduce the background rate B itself by improved shielding and material-handling techniques. Smaller background level results in smaller systematics.

Figure 4.1 illustrates how the backgrounds affect sensitivity: the red dotted straight line corresponds to the linear improvement in sensitivity for a zero-background experiment and the two curved lines show the effect of backgrounds. The background rates calculated for the CDMS 5 tower run are 4.2×10^{-4} events/kg-d for gammas with 2×10^{-6} rejection efficiency, 1.2×10^{-3} events/kg-d for betas with 2.1×10^{-3} rejection efficiency, and 1.5×10^{-4} events/kg-d for neutrons with no rejection power. These numbers correspond to expected exposures for future CDMS runs.

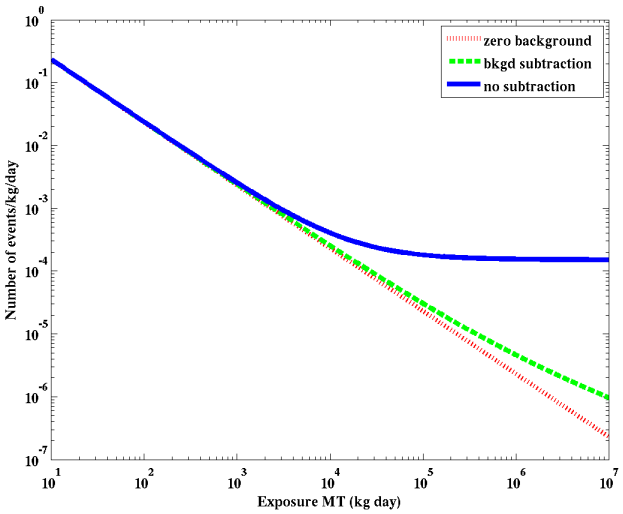


Figure 4.1: Sensitivity as a function of effective exposure (kg-d) for CDMS. The red dotted line indicates the sensitivity for a zero-background experiment, blue solid curves indicate the sensitivity without background subtraction, and green dashed curves indicate the sensitivity including subtraction of the one or two background events expected.

CDMS takes full advantage of the background-free strategy, where even for the whole projected exposures for the CDMS-II Soudan scientific data run, less than one background event would be expected.

4.2 Source of the background

To identify dark matter particles with confidence, backgrounds must be reduced at least a million-fold. The sources of background are as follows:

1. Radioactivity from cavern itself

   Photons from the uranium and thorium decay series and from ^{40}K decay, coming from naturally occurring radioisotopes in the cavern rocks. The uranium and thorium decay chains are the dominant alpha sources in rocks.

2. Contamination in detector and shielding material

Due to its high atomic number, reasonable cost, and mechanical properties, lead is the typical shielding material. However, its intrinsic radioactivity is generally not negligible. The major radio-impurities are usually uranium and thorium. At deep locations, the neutron background arising from environmental radioactivity, primary by spontaneous fission and (α, n) reactions from the uranium/thorium decay chain, exceeds the neutrons produced by muons in a massive lead shield. Electrons are emitted by radioactive contaminants on detector or nearby surfaces. Surface contamination depends on the detailed fabrication and handling history of the detectors.

3. Radon activity

   Our detectors are exposed to radon during fabrication, mounting and testing, during which time the adhesion of radon daughters on detector surfaces (radon plateout) occurs. ^{222}Rn decays by alpha-decay with $T_{1/2} = 3.82$ days, which itself is a decay product of ^{238}U. One of the dangerous daughters ^{210}Pb has a long half life of 22.3 years and decays by beta-decay. The gamma ray emitted by radon gas inside the lead shield could be another source of background, but purging with nitrogen has reduced this to a negligible amount.

4. Cosmogenic background

   Primary cosmic ray particles interact with atmospheric atoms to generate secondary elementary particles. Electrons, photons and protons are easily absorbed by overburden rock surface or by the lead shield, so only high energy muons make it through the rock. Such energetic cosmic-ray-induced muons are the primary source for cosmogenic neutrons inside the rock and shielding material. A principal motivation for WIMP searches at deep underground sites is the reduction

of ambient cosmic-ray-induced neutron flux. At the Soudan's depth of approximately 2000 mwe, the remaining muon flux is about 1 per minute in the CDMS veto shield.

4.3 Classification of background

The above background sources generate four primary types of backgrounds for CDMS: photons, electrons, neutrons, and alphas. Photons and electrons are electromagnetic backgrounds, since both produce electron recoils. The majority of these backgrounds can be rejected with the information obtained from simultaneous measurement of charge and phonon deposition energy of the interaction (chapter 3). Neutrons produce nuclear recoils, which cannot be rejected on an event-by-event basis unless they scatter in multiple detectors. Events produced by the recoiling nuclei of alpha decays can be mistaken for WIMPs if the emitted alpha particles are not detected.

4.3.1 Electromagnetic backgrounds

Depending on the depth of the interaction, electromagnetic backgrounds are classified as "bulk" or "surface". Bulk electromagnetic background rejection is extremely good due to our primary yield discrimination method. However, within 10 μm of the detector surface, electron recoils have poor charge collection and suffer ionization-yield suppression. If the ionization yield is sufficiently low, they will be misidentified as nuclear recoils. This misidentified electromagnetic background contributes to the major background to the WIMP search. This thin surface region in the ZIP detectors is called the "dead layer". We rely on the phonon timing information to reject these surface events. Detailed discussion of dead layer, surface events, yield rejection and phonon timing rejection is deferred to Chapter 5.

The majority of the electron recoils in the bulk of the detector where charge collection

is complete. They are identified empirically as events with unity normalized ionization yield. Bulk electromagnetic backgrounds consist primarily of high-energy photons, as well as electrons emitted from radioisotopes in the bulk of the detectors. Because the electron component is usually negligible in comparison to the photon component, bulk events are generically referred to as "photons".

Betas arising from contamination on the detector surface are the most difficult background to characterize. There are a number of possible beta emitters such as ^{40}K, ^{14}C, and ^{210}Pb. Surface events from radioactive contamination on the detector surfaces are generically referred to as "betas", reflecting the likelihood that most of these events arise from beta decay. They are identified empirically as all events with ionization yield below the 5σ below the "bulk" event band. They are the dominant background in CDMS at Soudan.

4.3.2 Neutron backgrounds

Neutrons produce nuclear recoils, which mimic WIMP signal, and cannot be rejected on an event-by-event basis unless they scatter in more than one detector, thus are the most dangerous background to the search of WIMP dark matter experiment. Two major sources contribute to neutron background: neutrons from cosmogenic background interactions, neutrons from radioactive decay.

The dominant source of neutrons that cause nuclear recoils in detectors, and are anticoincident with the veto, come from showers produced by muons that penetrate and interact in the rock that surrounds the CDMS experimental hall. This muon-induced neutron background was estimated independently using two Monte Carlo techniques (GEANT4 [33, 34] and FLUKA [35, 36]) and checked for consistency. The simulation performance was verified by comparing to the handful of nuclear recoils in the detectors coincident with tagged muon veto hits.

We can estimate the expected cosmogenic neutron background to the WIMP search at Soudan based on the simulation to be expect $(1.7 \pm 1.7) \times 10^{-3}((4.3 \pm 1.8) \times 10^{-2})$ events/kg-year of Ge (Si) exposure, leading to the conclusion that the cosmogenic neutron background is negligible for the first 5-tower run (Run 123/124). For this analysis we have raw exposures of 421.43 kg-d for Ge and 77.09 kg-d for Si.

Radioactive processes is primarily through (α, n) reactions and spontaneous fission, from uranium and thorium decay chains in natural rock and in many shielding materials.

Table 4.1: The concentration of uranium and thorium used in the simulation materials

Component	U - concentration [ppb]	Th - concentration [ppb]	Mass (kg)
Cu Cans	0.0797	0.2989	260
Cu Tower Guts	0.2	1	14.95
Inner Poly Shield	0.12	0.12	120
Inner (ancient) Pb	< 0.05	< 0.05	1917
Outer (Doe Run) Pb	< 0.05	< 0.05	12190

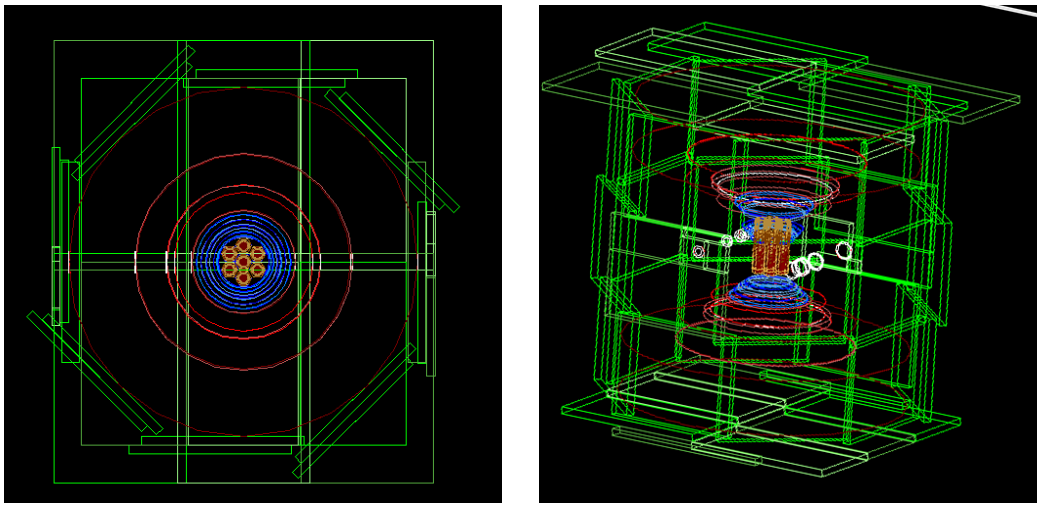
A Geant4 Monte Carlo simulation was performed to estimate the neutron production from (α, n) reactions and spontaneous fission in the copper cold hardware, the copper icebox cans, and the inner polyethylene shield. Assuming approximately 100 live days the result is 0.016 single nuclear recoils in Ge and 0.009 single nuclear recoils in Si, with negligible statistical errors.

The rate of neutrons emitted from 1 kg of natural uranium by spontaneous fission is estimated to be 16 neutrons/kg/s by several authors [37]. This corresponds to rate of 500 neutrons/kg/yr for a uranium concentration of 1 ppm in rock. The longer half-life of ^{232}Th makes neutron production from this chain negligible compared to that of uranium.

Samples of lead used in the CDMS-II shield which were counted at the Oroville facility show no uranium at the 50 ppt (g/g) level. This limit is consistent with measurements by G. Heusser and coworkers, who have conducted the best measurements to

date for uranium and thorium levels in lead. They find no detectable down to 20 ppt level.

The contribution of fission neutrons from the lead shield was separately simulated with Geant4. (α, n) processes are less significant in lead due to the absence of light target nuclei for the alpha to disintegrate. Two simulations were performed with the same Soudan shielding and veto geometry, but on different detector configurations, one with the 1 cm CDMS II ZIP detector in 5-Towers for the Soudan five tower run analysis, the other with the 1 inch SuperCDMS iZIP detector in 7-SuperTowers for the CDMS SNOLAB proposal (figure 4.2). At the time these two simulation were performed, they were for the purpose of comparing the self-shielding factor due to different detector configurations. Later we found they were also perfect for the background production in the detectors from the fission neutron in the lead shield.



(a) (b)

Figure 4.2: Monte Carlo simulation geometry of 7 super towers Soudan setup

In the simulation, neutrons were generated from an imaginary ^{252}Cf source interspersed uniformly throughout the Pb shield with the assumption that the ^{252}Cf and

^{238}U neutron spectra are similar (figure 4.3), and the multiplicity difference does not make much difference for self shielding factor. The generated neutrons were then propagated through both detector layouts, and recorded once they registered recoil energy in detectors.

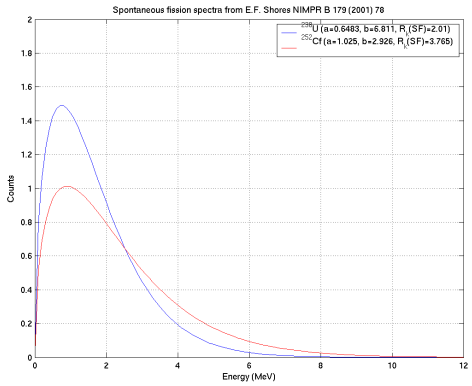
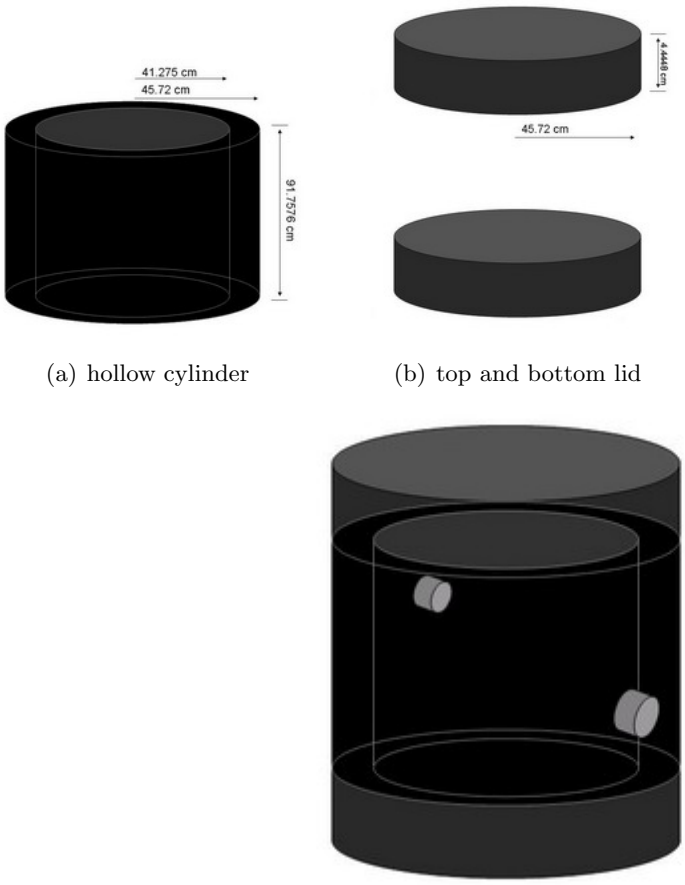


Figure 4.3: Spontaneous Fission spectra of ^{238}U and ^{252}Cf, cited from [3]

Since Geant4 General Particle Source only accepts simple geometry, three separate simulations were carried out with the ^{252}Cf source, one for a hollow cylinder, one for the top lid, and another for the bottom lid, thus completing the whole lead shield (figure 4.4). The 5-Towers simulation was done only for the inner lead shield, and the 7-SuperTowers simulation was for both inner and outer lead shield.

In the 5-Tower simulation, 4,999,995 neutrons million neutrons were generated in all three parts of the shield. The number of neutrons generated in each part of the geometry depends on the volume of each piece, presumably the source is uniformly distributed throughout the lead, which leads to 3,281,475 from the hollow cylinder, 859,260 from the top lid and 859,260 from the bottom lid.

In the analysis of the Monte Carlo simulation result, which records all the interaction inside the detectors, the following type of events are of main interest.



(a) hollow cylinder

(b) top and bottom lid

(c) e-stem and c-stem

(d) inner lead shield constructed with hollow cylinder, top and bottom lid, subtracting e-stem and c-stem

Figure 4.4: Cf source throughout the inner lead shield, subtract stems (Outer lead shield is similar)

Hits: Events with total energy between 10 keV and 100 keV in at least one detector. These are events with either nuclear recoil, electron recoil, or both.

Recoils: The subset of hits with less than 5 keV of electron recoil energy in the hit detector. These are events with mainly nuclear recoil events.

Singles: The subset of recoils with no other detector having greater than 2 keV total energy. This set is equivalent to the nuclear recoil background from the lead contamination that could mimic the WIMP signal.

For the above simulation, self-shielding factors for both 5-Towers and 7-SuperTowers are calculated and summarized in table 4.2. The self-shielding factor is the ratio of single-scatters to to fission neutrons in the shielding.

Table 4.2:

Tower Layout	Fission neutrons	Hits	Recoils	Singles	Self-shielding factor
5-Tower inner lead	4,999,995	35932	21816	12578	0.00252
7-SuperTower inner lead	1,000,000			6857	0.00686
7-SuperTower outer lead	1,000,000			2796	0.00280

The resulting rate was normalized by mass and assumed contamination levels of table 4.1, the ^{238}U decay rate (1 mBq/kg for 81 ppt U [38]), and the branching ratio for spontaneous fission of ^{238}U (5.45×10^7). This simulation did not separately tag the numbers of recoils in Ge and Si, but the ratio of Ge and Si event rates per detector is extrapolated from a previous simulation of the 2-Tower geometry. Details are described in table 4.3.

Summing these together, the expected total number of radiogenic background events from different materials surrounding the detectors is < 0.04 (< 0.1) in Ge (Si) for the current CDMS Run 123/124 exposure.

Table 4.3: Hits, Recoils, Singles, self-shielding factor, and fission neutron background from lead shield

	inner Pb(Ancient Pb)					outer Pb (Doe Run Pb)				
	top lid	bottom lid	side	c-stem	e-stem	top lid	bottom lid	side	c-stem	e-stem
Volume(cm^3)	29188.7	29188.7	111470	-563.079	-256.553	228623	228623	619384	-1126.16	-513.108
Total Volume(cm^3)			169028					1074990		
Volume Percentage	0.172686	0.172686	0.659478	-0.00333128	-0.00151781	0.212674	0.212674	0.576176	-0.0010476	-0.000477314
Hits	3025	3368	32090	15780	6612	865	1087	15716	7556	2705
Recoils	1381	1399	24198	10581	3268	310	364	11457	4863	1052
Singles	575	617	10112	4505	1423	135	163	4746	2049	464
Normalized Total Singles			6857					2796		
Self-shield factor	0.00686 (for 7 Super towers from MC)					0.00280 (for 7 Super towers from MC)				
Contamination (mBq/kg)	1					1				
Mass (kg)	1917					12190				
Decays(/s)	1.92					12.2				
% fissions	5.45E-07					5.45E-07				
Neutrons per fission	2.07					2.07				
Neutrons(/yr)	68.3					434				
Self Shielding Factor	0.00252 (for 5 CDMS II towers from MC)					0.00103 (for 5 CDMS II towers from normalization)				
100 days single neutron hit in all detectors	0.0472					0.122				

4.3.3 Alpha-induced backgrounds

In addition to emitting betas, contaminants on the detector surface can also emit alphas. Alphas have large recoil energies (> 1 MeV) and low ionization energies (< 1 MeVee), making them easy to identify. Thus, alphas themselves are not a background to WIMP interactions. However, the recoiling nucleus of a typical alpha decay in our detector produces a nuclear recoil with energy of order 20 - 100 keV. Events produced by the recoiling nuclei of alpha decays can be mistaken for WIMPs if the emitted alpha particles are not detected. Our detection efficiency for alphas is high, so we expect this background to be negligible.

4.4 CDMS efforts to reduce backgrounds

Most of the background is removed by the yield cuts and timing cuts shown in chapter 6 and chapter 5. The major remaining backgrounds are (a) surface betas with incomplete ionization collection, which are unable to be removed by either the yield cuts or the timing cuts and (b) neutrons which have no accompanying signal in the muon veto shield.

In general, reduced cosmogenic neutron backgrounds can be obtained by going deeper underground. Radiogenic neutron backgrounds from the surrounding rock can be reduced to a negligible level by appropriate shielding design. The biggest challenge is to obtain sufficiently low U and Th levels in the cryostat and shielding materials to ensure their contribution to the radiogenic neutron background is subdominant.

Many of the shielding and material-handling techniques are described in a review article by Heusser [38]. Listed below are some techniques used by CDMS.

4.4.1 Background shielding

CDMS uses several layers of active and passive shielding to reduce the overall background rate and minimize the number of interactions that might mimic the nuclear recoils.

The outermost layer of the CDMS shield is 713.5 meters (2,341 feet) of rock above the Soudan Underground Laboratory, constituting an overburden of 2090 meters water equivalent (m.w.e.). This overburden reduces the surface muon flux by a factor of 5×10^4.

Multiple layers of shielding are then arranged concentrically around the icebox. Outermost is an active muon veto system, composed of 40 overlapping scintillator panels. It tags the remaining incoming muons that initiate particle showers flux in the shielding. The muon identification efficiency is more than $99.98 \pm 0.02\%$, which reduces through-going muon background to insignificance.

Inside the scintillator shield is a 40 cm thick of cylindrical outer polyethylene layer to moderate low-energy neutrons from radioactive decays to below-threshold energies. Within the outer polyethylene is a 22.5 cm thick cylindrical lead shield, of which the inner 4.5 cm thickness consists of ancient lead for protection against external gamma rays, without further contamination from the ^{210}Pb chain. Another 10 cm of cylindrical polyethylene inside the inner lead provide further neutron moderation, particularly the neutrons by fission and (α,n) within the lead itself. The shield is about 99% complete coverage, with the only penetrations being the cold stem and the electrical stem.

The detector housings and the bulk of the cold hardware within the icebox are also made from high-purity copper, which limits radioactive contamination near the detectors. The total thickness of copper surrounding the detectors is a few cm, sufficient to stop alpha and beta radiation from outside the cans. The copper cans are also surrounded with a mu-metal magnetic shield to protect the SQUIDs and TES's from

external magnetic fields.

4.4.2 Radon reduction

The radon concentration level in the Soudan mine air is high. Compared to typical 40 Bq/m^3 in indoor air, and 10 Bq/m^3 for outside air, the radon concentration in the Soudan air has a seasonal variation, with peaks of 600-700 Bq/m^3 in the summer and valleys of 100-200 Bq/m^3 in the winter. Although the lead blocks most of the gamma rays emitted by radon decays outside of the shield, radon inside the lead shield can still create a significant gamma rate. The solution is to run all detectors and cold electronics under purge. During the early Soudan data run (most of Run 118, and all of run 119), medical grade breathing air was stored for at least two weeks to allow radon decay (the radon decay half life is 3.82 days) before being introduced under pressure into the air volume (roughly 100 liters) between the outside of the copper icebox and the inside of the mu-mental magnetic shield through a single polyethylene line. Studies of the rate before and after the purge indicates this practice of purging air reduced the gamma rate by more than a factor of 4, and beta rate by a factor of 2. Later for the five tower run, dry nitrogen gas was used instead of "old" air, with three additional purge lines for better gas flow. The purge flow rate was monitored on the daily basis.

4.4.3 Surface characterization

Betas from contamination on the detector surface, such as ^{40}K, ^{14}C, and ^{210}Pb, are the most difficult background to reject, and are the dominant background in CDMS. ^{14}C is produced by cosmic-ray induced reactions. Natural carbon could be introduced during processing. Another beta emitter candidate ^{40}K is also produced in the atmosphere by cosmic-ray interactions. Natural potassium may be deposited through human contact. The majority of detector surface background comes from the decay of ^{210}Pb, which has

a half life of 22.3 years. They are the decay products from radon, to which detectors were exposed during fabrication, testing and installation.

^{40}K and ^{14}C can be identified using standard surface characterization techniques (PIXE, RBS, SIMS, ICPMS), whereas ^{210}Pb was identified by determining the number of α's in the data and correlating them with the 45.6 keV β peak from ^{210}Bi in Rn decay chain. We have performed surface tests in the Shepard Lab surface characterization facility on the Minnesota campus to determine the source of the contamination on the surface of CDMS ZIP detectors. Thin test wafers used for the analysis were produced at intermediate and final stages of processing along with the CDMS ZIP detectors at the Stanford Nanofabrication Facility. These witness samples were tested to look for trace amounts of possible beta emitters that may cause the beta background in the data.

Ion beam analysis (IBA) is a nondestructive technique to quantify the absolute atomic ratios in compounds or mixtures. The IBA accelerator at the Characterization Facility at the University of Minnesota is a MAS 1700 pelletron tandem ion accelerator equipped with charge exchange RF plasma source by National Electrostatics Corporation (NEC). It can produce MeV ion beams with maximum energies of 3.4 MeV (H^+, He^+), and 5.1 MeV (He^{++}). It can probe elemental composition as a function of depth to several microns with a typical depth resolution of 100-200 angstroms. The analytical endstation (RBS 400) and control software (HYPRA) are by Charles Evans & Associates. HYPRA is also used as the specialized simulation and data analysis software.

IBA involves several specific techniques: Rutherford backscattering spectrometry (RBS) is sensitive to heavy elements in a light matrix, it can determine the film thickness (or density). Particle induced X-ray emission (PIXE) analysis utilizes distinctive characteristic X-rays emitted from the different target elements upon the beam bombardment to give the trace and minor elemental composition, and ensure the accurate identification of similar mass elements. We found that RBS combined with PIXE, done

simultaneously, provided the most sensitive result for the presence of light elements such as potassium.

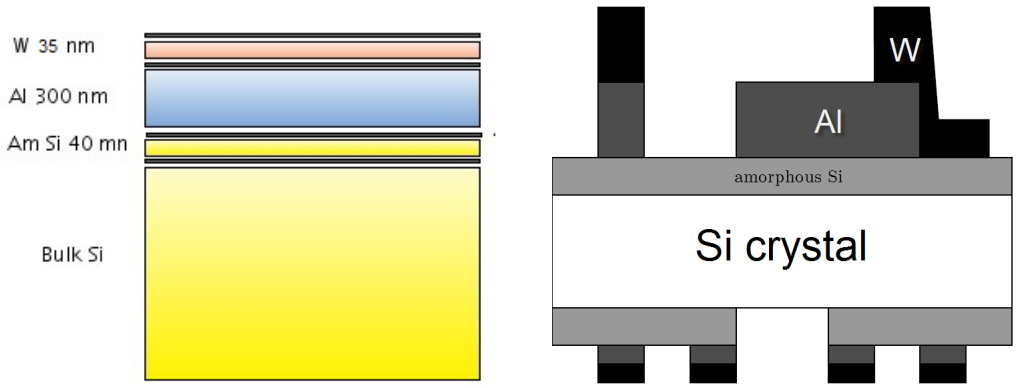
RBS measures the energy of the backscattered ions, thus effectively separating the elements, with the heaviest to the right and the lightest to the left on the spectrum. In addition, for each given known element, the RBS spectrum provides detailed depth information; as the atoms are deeper buried, the backscattered ions have lower energies, hence shift to the left on the spectrum.

If it is not known whether the contaminants are on the surface or buried, this introduces an ambiguity. The DOS program HYPRA provided by the facility does the analysis with Monte Carlo to figure out whether an RBS peak is a lighter element on the surface or a heavier element that has been buried. HYPRA also allows the user to model layer profiles and compare to data, as well as measure abundances in the peaks, as long as the element is known. The PIXE gives elemental x-ray lines and is used to resolve ambiguities.

Surface analysis of silicon wafers was performed, with a trilayer of amorphous silicon (α-Si), aluminium (Al) and tungsten (W) on top of bulk Si substrates. The trilayer sample is what we have just before etching a real detector, with thickness of each layer deposited on the phonon side of the crystal supposed to be W (35 nm), Al (300 nm), α-Si (40 nm), as illustrated in figure 4.5.

The use of amorphous Si film is to provide a blocking layer against charge carriers collected in the adjacent ionization electrode and improve the charge collection efficiency for the surface events. After the actual detector fabrication, a 1 μm-wide tungsten (W) transition-edge sensor (TES), fed by a set of 350 μm-long aluminum (Al) collector fins forms a single "QET", standing for quasiparticle-trap-assisted electrothermal-feedback transition-edge sensors. The TES serves as a sensitive thermometer and the fins act as antennas, concentrating phonon energy from a wide area onto the tiny TES. See [39,

40, 41] for more information on the ZIP detector phonon sensor design. More than a thousand (1036) of such active phonon-sensing TES, wired in parallel, form the four phonon sensor arrays. On the other side of crystal are deposited with a similar grid of W and Al which are wire bonded such that they create an inner disk and an outer ring electrode.



(a) Cartoon of Si substrate sample with three films on top (amorphous silicon, aluminium and tungsten)

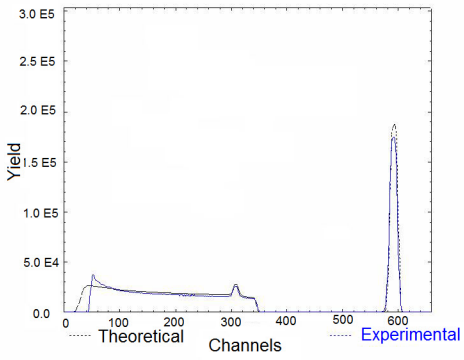
(b) Schematic picture of etched detectors. Si crystal is shown in white, amorphous silicon in lighter gray, aluminium in darker gray and tungsten in black.

Figure 4.5: Trilayer Si before (left) and after (right) etching

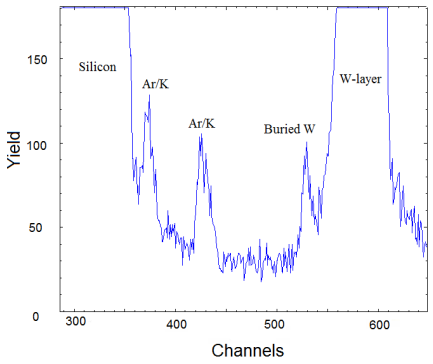
To match experimentally collected RBS spectrum in figure 4.6(a), thicknesses of each layer were found empirically to be W (35 nm), Al(250 nm), amorphous Si and bulk Si together (1700 nm). These values are close to the detector design.

Using the simulation, thin film layers of K or Ar are hypothetically inserted between each layer or on the top surface of W to best match experimental peaks around channel 380 and 420. The best fit gave two 2 $\mathring{A}$ of Ar inserted between W and Al, and between Al and amorphous Si (figure 4.6(b)). These two peaks may indicate the possible contamination of Ar in the sample.

This model can also put an upper limit on the density of element peaks. If the peak around channel 375 is Ar, the upper limit of Ar on the surface of amorphous silicon will



(a) Overview of the RBS spectrum. Wide silicon substrate on the left edge with aluminium on the top, huge tungsten peak on the right.



(b) A Zoomview of the gap between silicon edge to the left and tungsten layer to the right, showing two small buried Ar/K contaminant peaks and one W contaminant peak

Figure 4.6: Matrix fit of each layer and possible contaminations

be 9.0×10^{14} atoms/cm^2. If the peak around channel 425 is Ar, the upper limit of Ar between W and Al is 1.2×10^{15} atoms/cm^2.

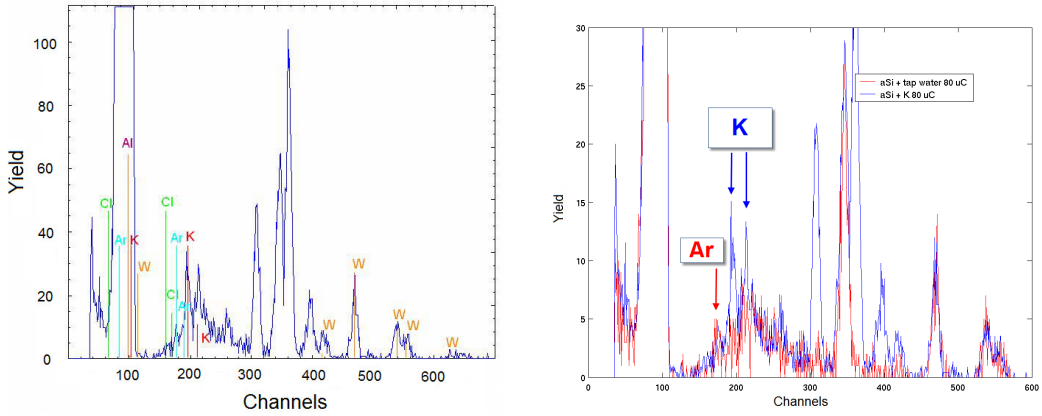
Correlating these result with previous Minnesota group measurement of other samples, we conclude that every time a film is deposited in fabrication, it starts the deposition with some contaminating Ar and W, which dies out as the deposition continues, resulting in a thin layer of Ar/W between each deposited layer.

While RBS result can help us quantify the absolute atomic ratios in compounds and determine the film thickness (or density), looking only at the RBS, we could not tell if the lighter element was Ar or K. Ar would be unimportant but K would be dangerous. PIXE analysis ensures the accurate identification of ambiguous mass elements.

Calibration of PIXE before the analysis is necessary for accurate identification. Calibration is performed by the energies (channel) and relative intensities (peak height) of each element-specific X-ray line in PIXE. In figure 4.7(a), W peaks are used for calibration. The relative height of peaks fits well with the result, indicating the correct energy-channel scale.

To distinguish Ar from K, we compared a clean amorphous Si wafer sample with another amorphous Si wafer sample where we artificially enhanced the K component by dipping it into the salt substitute (KCl) dissolved in tap water. The α-Si film was coated in three different ways: (1) An original clean Si substrate with amorphous Si film, (2) A second piece of dipped in tap water and let dry, (3) A third piece in in KCl dissolved in tap water.

We ran the prepared samples for 4 hours each. This enabled us to finally separate clearly the PIXE peaks of Ar and K. Figure 4.7(a) is an example of the x-ray spectra seen by PIXE for sample Si with salt to identify the elements to peaks. It illustrates where the K lines are. Figure 4.7(b) shows a closeup of Si with and without salt. You can see the small Ar line always there to the left.



(a) PIXE of amorphous Si wafer dipped in KCl

(b) PIXE of the Si sample with salt substitute added (blue) compared to the same sample before dipping in salt solution

Figure 4.7:

The above IBA technique with RBS and PIXE has convinced us we can put an upper limit on the buried and surface natural K layers as small as 1 - 2 $\times 10^{14}$ atoms/cm^2. This contamination level would yield a beta rate of about 1-2/detector/day over all energies. Simulations indicate that about 9% of these betas produce energy depositions in the

10 - 40 keV range and 15% of those (1.4% overall) are single scatters, resulting in 0.02 - 0.04 single scatters/detector/day. This could explain a portion of the subdominant non-^{210}Pb component. Ahtough we know the small peak we see is Ar, the existence of this peak right at the position in which K could be means that this limit is not as good as it would be if we had no interfering Ar in our background subtraction. We could avoid such confusion by replacing the Ar atmosphere used during metal deposition with He or Kr, enabling us to detect buried and surface K layers as small as 10^{13} atoms/cm^2.

The surface screening continues to be useful, as it provides quality control for the future new CDMS detectors, as well as quantification of new processes and new substances that may be introduced into the fabrication.

4.5 Background estimate

The surface event leakage was estimated based on the observed numbers of single and multiple scatter events in each detector within and surrounding the 2σ nuclear recoil region. The expected background due to surface interactions in this Run123/124 WIMP search analysis is 0.6 ± 0.5 events.

Neutrons induced by radioactive processes or by cosmic ray muons interacting near the apparatus can generate nuclear recoil events that cannot be distinguished from possible dark matter interactions on an event-by-event basis. Monte Carlo simulations of the cosmic ray muons and subsequent neutron production and transport have been conducted with FLUKA [35, 36], MCNPX [42] and GEANT4 [33, 34] to estimate this cosmogenic neutron background. Normalizing the results to the observed veto-coincident multiple scatter nuclear recoil rate leads to a conservative upper limit on this background of < 0.1 events in our WIMP-search data.

Additional Monte Carlo simulations of neutrons induced by nuclear decay were based on gamma ray measurements of daughter products of U and Th in the materials of our

experimental setup and the assumption of decay equilibrium. The respective background estimate is < 0.1 event, dominated by the upper limit of U in the Pb shield. Direct measurements of U in Pb [43] from the same source as the Pb used in our shield suggest a considerably lower contamination.

Chapter 5

Surface Event Rejection

The CDMS detectors discriminate nuclear recoils from electron recoils caused by background betas and gammas by measuring simultaneously the ionization and phonon energy. The ratio of ionization to phonon energy or "yield" enables a determination of the recoil type.

The principle is that the total stopping power of a target material atom consists of both an electronic and a nuclear component. The electronic stopping power is the recoiling nucleus energy loss due to electronic excitation and ionization. The nuclear stopping power is the energy loss due to atomic collisions that contribute to the kinetic energy (thermal motion) of the atoms. These two stopping powers have different dependencies on the recoil energy of the nucleus. Nuclear recoils are not stopped efficiently by electrons and so deposit most of their energy through interactions with the target's nuclei, giving a reduced yield for nuclear recoils. In the energy range of WIMP-nucleus elastic scattering, the nuclear stopping power plays a significant role in the energy loss of the recoiling nucleus. The phenomenon of reduced ionization yield of a nuclear recoil varies with respect to electron recoil, quantitatively explained by Lindhard *et al.* [44], is the key discrimination power of CDMS and many other direct detection experiments.

However, surface events in the dead layer [45] from radioactive contamination on the detector surfaces, or as a result of external gamma ray interactions releasing low-energy electrons from surfaces near the detectors have reduced ionization yield. A correlation analysis between alpha decay and surface event rates provides evidence that ^{210}Pb (a daughter-product of ^{222}Rn) is a major component of our surface event background. If the yield is significantly low, they can resemble nuclear recoil yield. They are the primary background for CDMS, as discussed in Chapter 4.

The quicker response of the phonon sensors to "athermal" phonons allows us to resolve subtle differences in the phonon pulse timing and spatial distribution from different types of particle events, including surface vs bulk electron recoils. This rich information of pulse timing, amplitude, and position on an event by event basis provides CDMS with discrimination power to reject more than 99% of surface events, while still maintaining substantial acceptance of nuclear recoils.

5.1 Summary of discrimination techniques

Numerous parameters were found to have discrimination power from different pulse reconstruction. With such a wealth of information, different sophisticated statistical discrimination techniques have been developed with these discrimination parameters.

1. **Discrimination parameters :** phonon timing and energy information

   (a) **Primary phonon rise time :** the difference between the 10% and 40% of the primary phonon pulse rising edge, as shown on figure 5.1. Phonons from a surface event have a shaper rise time due to the down-conversion process, whereas energetic primary phonons become lower-energy ballistic phonons.

   (b) **Primary phonon delay time :** the difference between the start of the ionization pulse and the 20% of the primary phonon pulse rising edge 5.1.

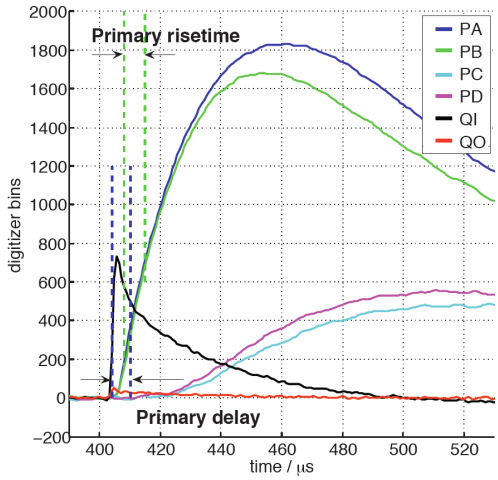


Figure 5.1: Illustration of phonon timing parameters of digitized phonon and charge trace. The primary (largest) phonon pulse of this event occurs in channel A. The vertical dashed lines indicate the event's primary rise time (time between the 10% and 40% of the primary phonon pulse rising edge) and primary phonon delay time (time between the charge start time and the 20% of the primary phonon pulse rising edge).

Primary phonon rise time and primary phonon delay time are strongly correlated, as shown on figure 5.2. Surface events which have a quicker phonon rise time typically also have shorter delay between phonon and charge pulses.

(c) **Phonon energy partition :** the ratio of the amplitude of the largest phonon pulse to the amplitude of the pulse in the opposite phonon channel, as shown on figure 5.3, was demonstrated to have discrimination power (Figure 5.4(a)).

(d) **Phonon timing partition :** the ratio of opposite to primary phonon delay time is higher for a surface event than for a bulk event (shown in figure 5.4(b)), since phonons from a surface event have to first reflect off the bottom (or the top) surfaces and side surfaces to get back to the opposite sensor.

2. **Event reconstruction**

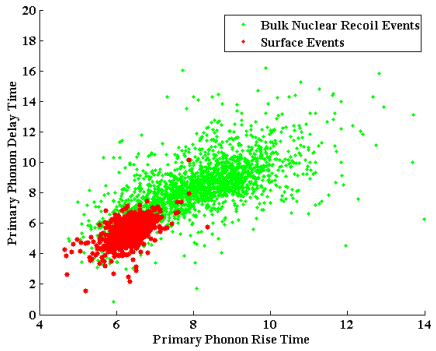


Figure 5.2: Scatter plot of primary phonon rise time and delay for calibration events on detector T2Z3, shows strong correlation between primary phonon rise time and delay time. Bulk nuclear recoil events are in green, surface events in red.

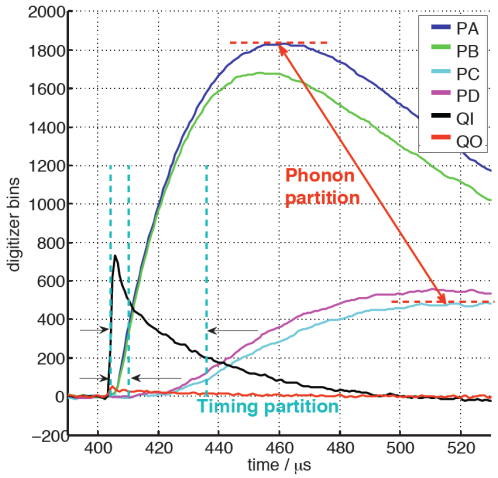
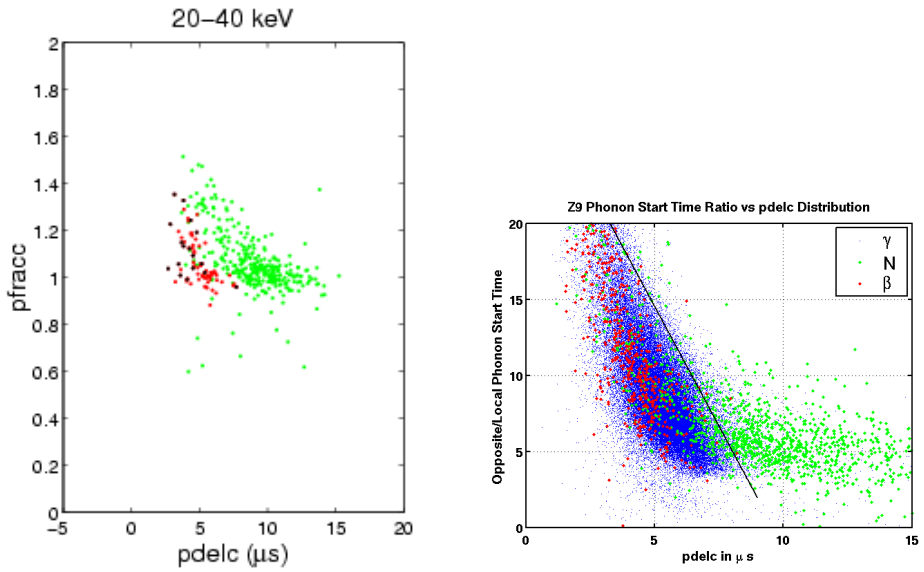


Figure 5.3: Illustration of phonon partition parameters of digitized phonon trace. The primary (largest) phonon pulse of this event occurs in channel A. The opposite phonon pulse occurs in channel C. The horizontal dashed lines indicate the event's primary phonon amplitude and opposite phonon channel amplitude.



(a) Scatter plot of phonon timing partition, the ratio of primary sensor amplitude to opposite phonon sensor amplitude. Bulk nuclear recoil events are in green, and surface events in red.

(b) Scatter plot of phonon energy partition, the ratio of opposite to primary phonon delay time. Bulk nuclear recoil events are in green, bulk electron recoil events in blue, and surface events in red.

Figure 5.4: Scatter plot of two phonon partition parameters.

(a) **Optimal Filtering :** The above phonon pulse amplitudes (energies) are estimated using an "optimal filter" algorithm, a frequency domain fitting. Optimal filtering is the optimal technique for amplitude estimation with knowledge of the characteristics of signal (known pulse shape) and background (known spectrum of Gaussian noise). The fit of both phonon pulse template and noise in the frequency domain using the Fourier transforms of the trace will account for the dependence of variation in noise with frequency. With the existence of non-white noise, optimal filtering gives optimal resolution.

(b) **Walking algorithm :** The above timing parameters come from the time-domain "walking" algorithm, computed using a time-domain walk algorithm, applied independently after the filtering. It finds the maximum of a pulse, then "walks" down along the rising edge to find the times of 10%, 20%, and 40% of its maximum amplitude.

(c) **Time domain phonon pulse shape fitting :** The optimal filter approach uses template a pulse shape to reconstruct traces. However, the pulse shape also depends on the position of the events, which requires the reconstructed events to be position-corrected as a function of detector radius. It is natural to use a χ^2 minimization fit in the time domain to match to a parameterized function form. This algorithm can give a more accurate determination of the true start time of the phonon pulse by extrapolating the functional form to a zero crossing. The time domain phonon pulse fitting improves the energy resolution by using the actual noise background instead of an average taken during each earlier run, and also provides us deep understanding of underlying phonon physics of the detector.

3. **Discrimination techniques**

(a) **Simple timing cut :** the simple timing cut defines a rectangular signal region in the parameter space, by a series of cut values of these various parameters. This technique was used in the Soudan one-tower data run (Run 118) [46]. A nuclear recoil candidate is required to have rise time and delay time above some limiting valuesto be accepted as WIMP candidate event. The requirement is intended to reduce the surface event background to an acceptable level.

(b) **Correlated timing cut :** the 2-tower and the current 5-tower analysis take advantage of the correlation between two timing parameters, the primary phonon rise time and delay time, seen in figure 5.2. The rotated axes of (delay time + rise time) and (delay time - rise time) are used to make cuts. The cut value is set along the line of (delay time + rise time), where a certain percent of surface events is allowed to leak into the nuclear recoil signal region. Consistency cut between upper and lower limit of (delay time - rise time) rejects events far away from nuclear recoil population.

(c) χ^2 **analysis :** Since the likelihood distributions are complicated or unknown, we make simplifying assumption that the signal and background populations follow multivariate Gaussian distributions (rise time, delay time, phonon energy partition, or more), the likelihood χ^2 of each event with respect to the two distributions is used as discrimination parameter.

(d) **2D gaussian density model :** The author of this dissertation developed a 2D gaussian density model (section 5.3) to take advantage of the correlated properties of rise time and delay time, assuming the bivariate gaussian distribution of these two parameters. The idea is *similar* to the χ^2 analysis.

(e) **Neural networks :** Neural network techniques use a set of discrimination parameters such as phonon rise time, delay time, partition, etc. to train the

networks. The network output is then used to distinguish nuclear recoils from surface events.

5.2 Classic 2D analysis

The analyses of previous runs demonstrated that the use of two phonon timing parameters: phonon delay time and phonon rise time, provides good rejection of surface electron recoils while retaining reasonable acceptance of nuclear recoils. A simple analysis technique, summing the phonon delay time and rise time quantities to form a timing parameter, was applied in the last Run 119 analysis. This classic 2D timing parameter analysis is simple and robust. Hence, this technique was used as primary analysis to report the results for Run 119 data.

Due to the simplicity and robustness, the primary surface event rejection cut for the current analysis follows the same scheme developed for the Run 119 analysis. This cut is based upon two similar quantities, the primary phonon sensor rise time (pminrtc) and the primary phonon pulse delay with respect to the ionization pulse (pdelc), position and energy corrected (indicated by the letter"c" at the end of the name), using the phonon correction table. Previously, further empirical energy corrections to each parameter were used to flatten the upper edge of the surface event distribution. This step is not required in this run due to the improvements to the look-up table.

A consistency cut for signal candidates demands that the difference between these parameters (pdelc - pminrtc) be consistent with the calibration neutron population, which is defined to only accept events within 4σ of the mean neutron value for the chosen detector, perpendicular to timing parameters in the 2D delay vs rise time graph. This consistency cut essentially rejects incorrectly reconstructed timing parameters.

Figure 5.5 shows the signal region accepted by the timing cut is between the two consistency cuts and above the timing cut threshold line. This removes the surface beta

events in red and isolates the nuclear recoil neutron events in green. Figure 5.7 shows the recoil energy dependence of the timing parameter.

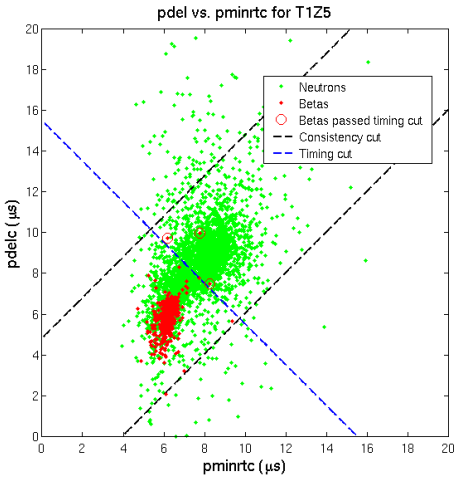


Figure 5.5: Consistency cut region (between dashed black lines) and timing parameter cut (dashed blue line) in the 2D delay time vs rise time graph. Surface beta events in red with red "leakage" events circled, nuclear recoil neutron events in green.

Timing cuts were set based on leakages from a subset of ^{133}Ba calibrations, which satisfied the cBeta cut selection: 1)Between a yield of 0.1 and 0.7, 2) below 5 σ electron recoil band, and 3) no detector-detector crosstalk due to the "holiday island" events in towers 1 and 2. "Holiday islands" are regions where inner-electrode has floating patches on the ground side due to the stairstep bounday instead of circle for the inner-electrode. The minimum value (threshold) is determined on a detector-by-detector basis by setting the allowed leakage (events passing the timing cut) passage fraction to be about 0.3% for both Ge and Si detectors in the odd event-numbered subset of ^{133}Ba calibrations, and slightly adjusted to allow optimal WIMP signal efficiency over all detectors. The performance of the timing cuts is then checked on the even event-numbered subset of ^{133}Ba calibrations. The left plot in figure 5.8 shows the surface event leakage fraction for each detector used in the current WIMP search run, the right plot in figure 5.8 shows

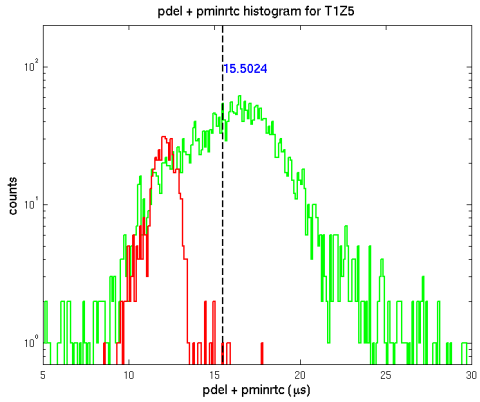


Figure 5.6: Neutron and beta timing parameter histogram. Surface beta events in red and nuclear recoil neutron events in green. The dashed line demonstrates the possible timing cut threshold, which rejects events to the left side (surface events with quick phonon pulses) and accept events to the right side (bulk nuclear recoil events with slow phonon pulses)

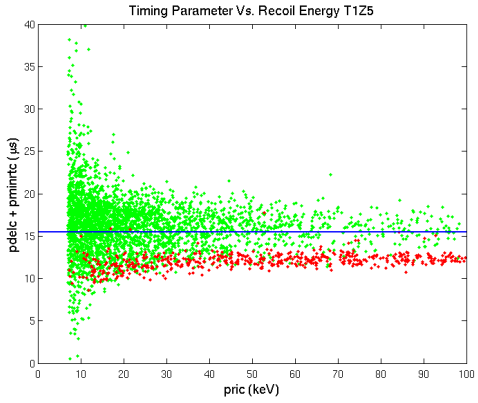
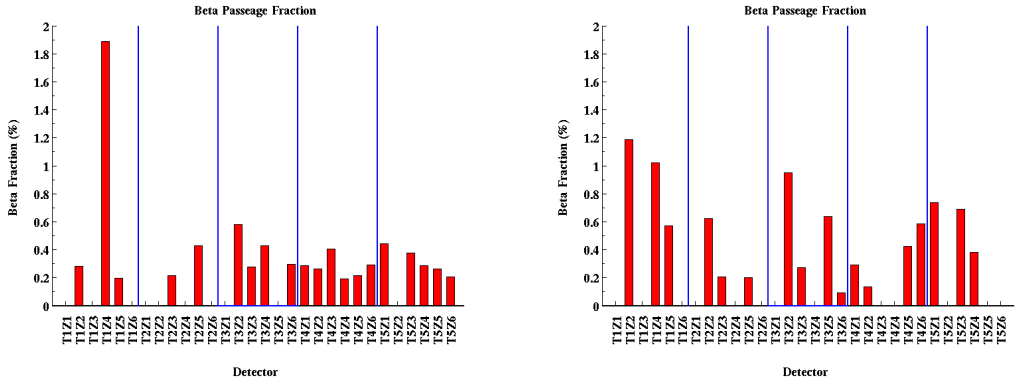


Figure 5.7: Neutron and beta timing parameter vs. the phonon recoil energy. Surface beta events in red and nuclear recoil neutron events in green.

the performance on the even event-numbered subset of ^{133}Ba calibrations. The detailed leakage event counts and fraction is listed in table 5.1 and table 5.2.



(a) Odd beta passage fraction, the average is 0.29 ± 0.02% (0.33 ± 0.06%) for Ge (Si).

(b) Even beta passage fraction, the average is 0.33 ± 0.02% (0.32 ± 0.05%) for Ge (Si).

Figure 5.8: Odd and Even Beta Passage Fraction

Table 5.1: (allowed leakage/all beta) leakage fraction on odd Ba

	T1	T2	T3	T4	T5
Z1		(0/206) 0.00%		(1/349) 0.29%	(2/450) 0.44%
Z2	(1/353) 0.28%	(0/125) 0.00%	(2/344) 0.58%	(2/767) 0.26%	
Z3		(1/468) 0.21%	(1/362) 0.28%	(1/246) 0.41%	(1/264) 0.38%
Z4	(2/106) 1.89%	(0/167) 0.00%	(2/467) 0.43%	(1/527) 0.19%	(2/701) 0.29%
Z5	(1/505) 0.20%	(2/466) 0.43%	(0/287) 0.00%	(1/460) 0.22%	(1/381) 0.26%
Z6			(3/1006) 0.30%	(2/684) 0.29%	(1/485) 0.21%

Co-added Passage Fraction:
Ge: (24/8351) (0.29±0.02)%
Si: (6/1825) (0.33±0.06)%

The WIMP signal efficiency is computed based on neutron events in the 2 σ nuclear recoil band in ^{252}Cf calibrations, and the timing cut threshold. The fraction of nuclear recoil neutron events inside the signal region accepted by the timing cut and between the two consistent cut are computed as the timing cut WIMP signal efficiency. The Table 5.3 and figure 5.9 shows the WIMP efficiency for each detector. The signal

Table 5.2: (leakage number/ all beta) leakage fraction on even Ba

	T1	T2	T3	T4	T5
Z1		(0/202) 0.00%		(1/342) 0.29%	(3/407) 0.74%
Z2	(4/337) 1.19%	(1/160) 0.62%	(3/316) 0.95%	(1/749) 0.13%	
Z3		(1/481) 0.21%	(1/368) 0.27%	(0/249) 0.00%	(2/290) 0.69%
Z4	(1/98) 1.02%	(0/166) 0.00%	(0/515) 0.00%	(0/577) 0.00%	(3/782) 0.38%
Z5	(3/525) 0.57%	(1/497) 0.20%	(2/314) 0.64%	(2/470) 0.43%	(0/402) 0.00%
Z6			(1/1058) 0.09%	(4/685) 0.58%	(0/447) 0.00%

Co-added Passage Fraction:
Ge: (28/8562) (0.33±0.02)%
Si: (6/1875) (0.32±0.05)%

acceptance of phonon timing cuts is roughly 50%, which imposes the highest costs in signal acceptance.

Table 5.3: Classic 2D Timing Cut Neutron Efficiency

	T1	T2	T3	T4	T5
Z1		55.0%		56.2%	56.2%
Z2	69.9%	64.7%	57.1%	41.6%	
Z3		67.3%	3.2%	67.9%	50.8%
Z4	46.9%	73.5%	60.9%	57.6%	16.9%
Z5	50.8%	56.4%	58.4%	75.3%	52.2%
Z6			13.8%	38.2%	62.9%

There are challenges to characterizing the timing cut leakage on the WIMP search data, since no accurate model describes the detector response to the surface nuclear recoils, particularly the significant variation among detectors. Since there are significant systematic differences in the surface leakage between the ^{133}Ba calibration data sample and the WIMP-search data sample, the performance of our timing cut cannot be reliably determined solely from the calibration data. We must estimate from the observed event counts in the much poorer statistics WIMP-search data. A Bayesian statistical framework was used in a study to estimate the expected surface event background after

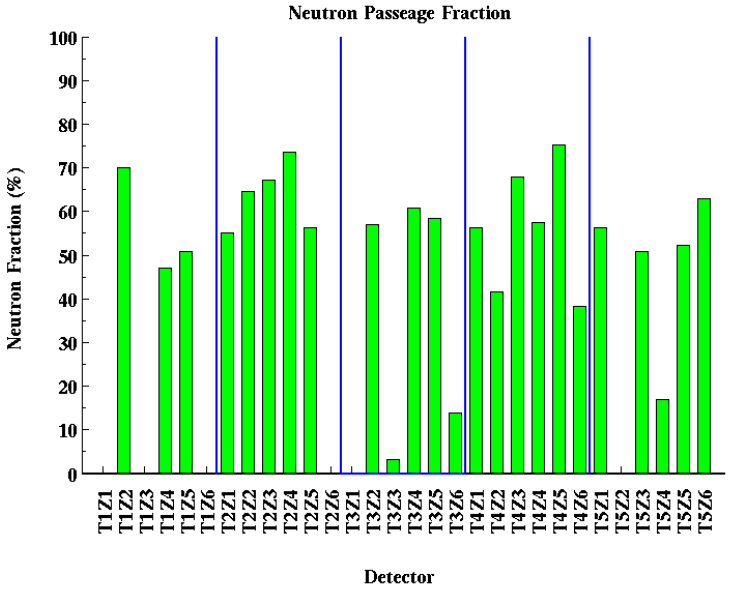


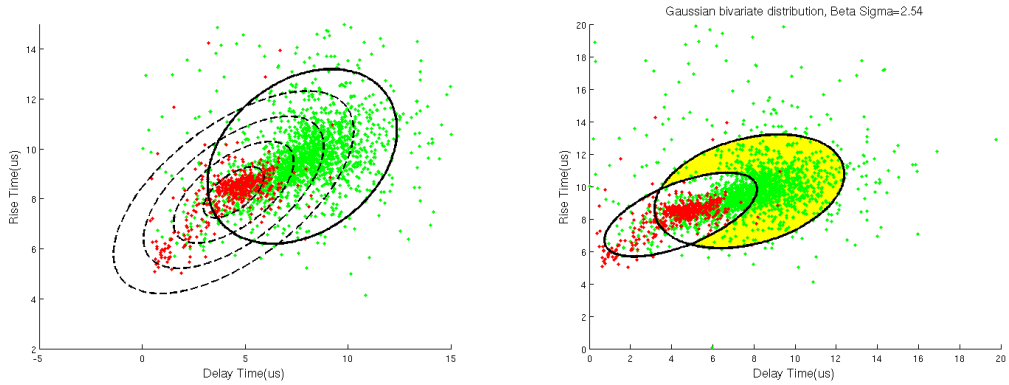
Figure 5.9: Classic 2D Timing Cut Neutron Efficiency

the timing cut. In the current WIMP search analysis, a total of 0.6 $\pm$ 0.5 events is expected to pass the timing cuts for the entire Run 123/124 on Ge data, according to this study.

5.3 2D Gaussian Timing Cut

The simple one parameter timing cut and the classic 2D timing cut reject surface events based on either phonon rise time or delay time or both, with the cut threshold as a straight line in the timing parameter space. A one dimensional Gaussian distribution is assumed for the main population of both neutron and surface beta events. The cut threshold is determined by the fraction of surface beta outliers. Since there are strong correlation between these two powerful discrimination parameters, it is natural to think of bivariate Gaussian distribution in two dimensions to confine two separate populations in their tilted ellipse regions. This would be determined by the resulting

covariance matrix (figure 5.10(a)). In this bivariate Gaussian model, neutron events and surface beta events are represented by two Gaussian distributions separately. To reject surface beta events we must accept events that are within a fixed number of σ of the neutron region (consistency cut), while outside of the beta region. The yellow region in figure 5.10(b) is a sample signal acceptance region.



(a) 2D Gaussian model with different standard derivations

(b) Gaussian bivariate distribution, signal region in yellow

Figure 5.10: 2D Gaussian Model

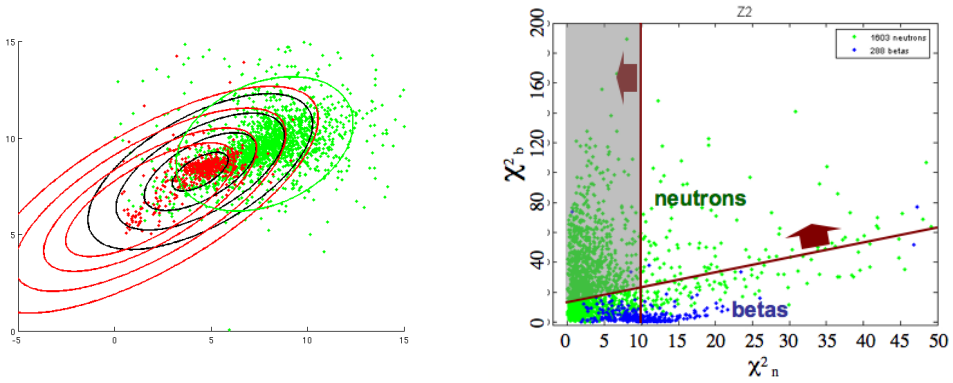
It turns out that this independently proposed timing cut based upon two dimensional Gaussian model is mathematically identical to the χ^2 timing cut. The number of σ in the 2D Gaussian model is the square root of χ^2. In the χ^2 analysis, χ^2 is calculated as

$$\frac{\sum (x_i - \mu_j)^2}{cov(i,j)} \tag{5.1}$$

The χ_b^2 is the distance from the beta hypothesis and the χ_n^2 is distance from the neutron hypothesis, both are used to define a likelihood of an event being of type beta or neutron. The neutrons have smaller distance from the neutron hypothesis and thus have smaller χ_n^2. Similarly, the betas have smaller χ_b^2. The correlations between the discriminators play a significant role and strongly enhance the combined discrimination power.

The difference is that in the 2D Gaussian model, a constant χ_b^2 threshold is set as the

ellipse around the beta region, while in the χ^2 timing cut, threshold is set on $\chi_b^2 - \chi_n^2$, which leads to a skewed ellipse, with the constant $\chi_b^2 - \chi_n^2$ lines pushed towards left bottom side. Figure 5.11 compares the 2D Gaussian timing cut with the χ^2 timing cut.



(a) constant χ_b^2 lines are in black, constant 1σ,2σ,3σ,4σ $(\chi_b^2 - \chi_n^2)$) lines in red, and constant χ_n^2 line in green

(b) χ^2 timing cut, grayed area is the signal acceptance region.

Figure 5.11: Comparison of 2D Gaussian model with χ^2 model

5.4 Feature Selection and Transformation

All the above surface rejection analyses only utilize a tiny fraction of the information obtained from a ZIP detector's phonon and charge pulses. In principle, we should be able to achieve even better discrimination by incorporating more information into our analysis.

One of the difficulties in a multivariate problem is to visualize the relationships between many variables in high dimensions. Another difficulty is the extra computing time required for solving high dimensionality data. One of the simplest solution for dimensionality reduction is to select a subset of the input variables, and to discard the rest. This is useful if (1) there are variables carry little useful information to discriminate different classes of events, or (2) there are very strong correlations between sets of

variables so that same information is repeated in several variables.

Since a large amount of information is ready to choose from, it would be useful to quantitatively identify which variables have more discrimination/classification power, extract the most of the information from them, and reduce the dimension space to a lower one. A couple of statistical techniques are available.

5.4.1 Correlation Matrix

The classic 2D timing cut has two components, showing the strong positive correlation between the phonon rise time and delay time, which is reasonable due to the overlap between these two pulse timing spans. However, CDMS has a larger number of timing, energy, position variables, making direct visualization of the correlation between all the variables difficult. There are ways to visualize high dimensional data. Viewing slices through lower dimensional subspaces is one way to partially work around the limitation of low dimensions. For example, we can display an array of all the bivariate scatter plots between variables, along with a univariate histogram for each variable on the bottom row in the matrix plot Figure 5.12. It illustrates the correlation between any of the two variables. The histogram also shows the relative classification power of each individual variable. The points in each scatter plot are color-coded by the type of events: red for surface beta events, green for nuclear recoil neutron events. This array of plots makes it easy to pick out patterns in the relationships between pairs of variables.

The variables in the matrix are:

1. position- and energy-corrected quality of primary phonon sensor rise time from 10% to 40% (pminrtc),

2. primary phonon sensor delay time (pdelc),

3. phonon partition (pfracc = primary sensor phonon amplitude / opposite one),

4. phonon recoil energy (pric),

5. ionization yield (yic = charge / phonon recoil energy)

6. two position related features reconstructed on phonon delay (rppart) and phonon energy (rdel).

We may also compute the following correlation matrix for these 7 variables. It's obvious that some pairs of these variables are positively correlated (pminrtc and pdelc, for example), some are negatively correlated (pminrtc and pfracc), and there are some very strong pairwise correlations, while others are weak.

The correlation matrix table 5.4 shows strong correlation between some of the features, like pminrtc and pdelc (0.78), yic and pric (0.60), pfracc and rppart(0.56), rppart and rdel (0.53). This means we might be able to reduce the dimension by appropriate matrix rotation.

Table 5.4: correlation matrix of 7 selected variables

	pminrtc	pdelc	pfracc	pric	yic	rppart	rdel
pminrtc	1.0000	0.7841	-0.3488	-0.1422	-0.2700	-0.1781	0.0285
pdelc	0.7841	1.0000	-0.2606	-0.2300	-0.3857	-0.1689	0.1211
pfracc	-0.3488	-0.2606	1.0000	-0.1300	-0.2198	0.5567	0.3069
pric	-0.1422	-0.2300	-0.1300	1.0000	0.5960	-0.0913	-0.1990
yic	-0.2700	-0.3857	-0.2198	0.5960	1.0000	-0.0950	-0.1801
rppart	-0.1781	-0.1689	0.5567	-0.0913	-0.0950	1.0000	0.5316
rdel	0.0285	0.1211	0.3069	-0.1990	-0.1801	0.5316	1.0000

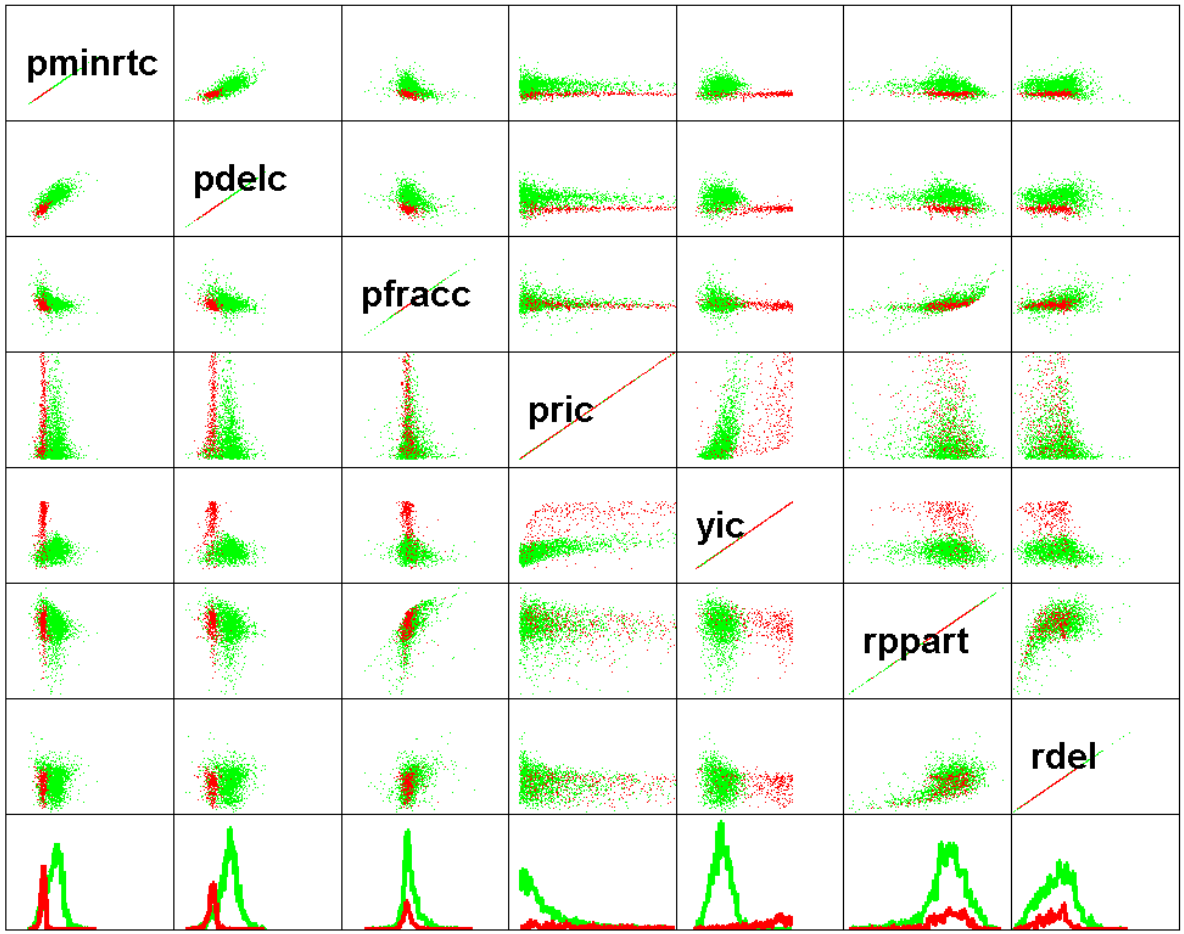


Figure 5.12: Correlation Matrix Plot

5.4.2 Principal Components Analysis

With this many variable, it is not easy to get a good feel for the relationships among all of them. There may be important patterns in higher dimensions, and those are not easy to recognize in matrix correlation plot. Since more than one variable might be measuring the same underlying behavior, groups of variables are often correlated and move together. We can take advantage of this information redundancy in order to simplify the problem. One can replace the detector-based variables with new variables by mapping the data from a space of higher dimensionality into a space of lower one. In general, such a reduction procedure will lose some of the information. The goal in dimensionality reduction is therefore to preserve as much of the relevant information as possible.

The principal components analysis (PCA) provides such a mechanism [47, 48]. PCA is a quantitatively method of linear dimensionality reduction procedure. This method identifies the components underlying the correlated variables of vectors $\overrightarrow{\mathbf{x}}$ in d-dimensional space $(x_1, ..., x_d)$ through its correlation structure, then generates a new set of variables of vectors $\overrightarrow{\mathbf{u}}$ in a m-dimensional space $(u_1, ..., u_m)$, called principal components, where $m < d$. Each principal component is a linear combination of the original variables.

$$\mathbf{x} = \sum_{i=1}^{d} z_i \mathbf{u}_i \tag{5.2}$$

where the vectors $\overrightarrow{\mathbf{u}}$ are orthogonal to each other

$$\mathbf{u}_i^T \mathbf{u}_j = \delta_{ij} \tag{5.3}$$

so there is no redundant information. The principal components as a whole form an orthogonal basis in the new data space, which essentially can be regarded as a simple rotation of the coordinate system from the original x array to a new set of coordinates u around the mean of the data points. This linear transformation moves as much of

the variance as possible into the first few dimensions. The values in the remaining dimensions, therefore, tend to be highly correlated and may be dropped with minimal loss of information. The technique is illustrated schematically in Figure 5.13 for the case of two dimensional space.

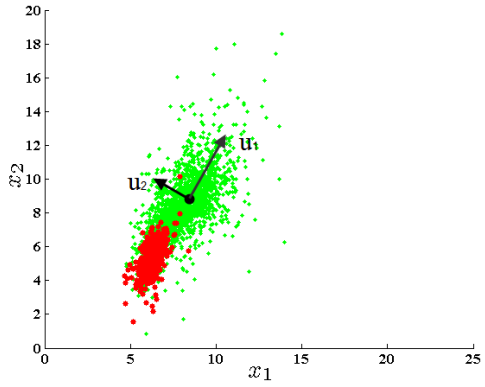


Figure 5.13: Schematic illustration of principal component analysis applied to data in two dimensions.

In practice, the algorithm first computes the mean of the vectors $\overrightarrow{\mathbf{x}}$ and then subtracts it off. The covariance matrix is then calculated and its eigenvectors and eigenvalues are found. The eigenvectors corresponding to the m largest eigenvalues are retained and the original vectors $\overrightarrow{\mathbf{x}}$ are projected onto the eigenvectors to give the components of the transformed vectors $\overrightarrow{\mathbf{u}}$ in the m-dimensional space.

The first principal component is a single axis in space. When you project each observation on that axis, the resulting values form a new variable. And the variance of this variable is the maximum among all possible choices of the first axis. The second principal component is another axis in space, perpendicular to the first. Projecting the observations on this axis generates another new variable. The variance of this variable is the maximum among all possible choices of this second axis, and so on. The full set of principal components is as large as the original set of variables, *e.g.* $m <= d$.

But it is common for the sum of the variances of the first few principal components to exceed 80% of the total variance of the original data. By examining plots of these few new variables, we may develop a deeper understanding of the driving mechanism that generated the original data.

The table below is the component matrix. Each of the columns represents one of the components. Each cell's entries represent the partial correlation between the variable and the component. The first principal component may be viewed as the single best summary of the correlations among the variables. This particular linear combination of the variables accounts for more variability than that of any other conceivable linear combination. As we expected, pminrtc and pdelc accounts for the most weight for the first principle component, 0.56 and 0.57 respectively.

Table 5.5: Component Coefficients

component	1st	2nd	3rd	4th	5th	6th	7th
pminrtc	0.5607	-0.0364	0.3570	-0.1968	-0.2685	-0.1380	0.6534
pdelc	0.5741	-0.1153	0.2785	-0.1813	0.0852	-0.1867	-0.7103
pfracc	-0.3281	-0.4517	-0.0897	-0.5790	0.1780	-0.5529	0.0874
pric	-0.2448	0.3950	0.5317	-0.4897	0.3561	0.3666	0.0178
yic	-0.3141	0.4276	0.4097	0.2644	-0.2966	-0.6182	-0.1032
rppart	-0.2906	-0.4638	0.3307	-0.0886	-0.6541	0.3517	-0.1782
rdel	-0.0725	-0.4769	0.4765	0.5249	0.4964	-0.0133	0.1347

Each eigenvalue indicates the portion of the variance that is correlated with each eigenvector. Thus, the sum of all the eigenvalues is equal to the sum squared distance of the points with their mean divided by the number of dimensions. Component variances are the variance explained by the corresponding principal component, and plotted in Figure 5.14. The eigenvalues of each components are 2.32, 2.20, 1.04, 0.55, 0.39, 0.31, 0.18 (dimensionless). An eigenvalue of "1" would then mean that the component would explain about "one variable's worth" of the variability. The Figure 5.14 shows that the amount of variance accounted for by each component is clearly separated between the

second and third components.

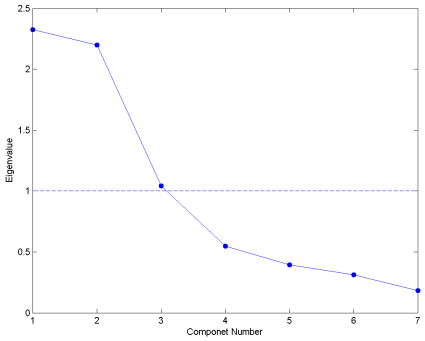


Figure 5.14: the variance explained by the corresponding principal component

The percent of the total variability represented by each principal component is also calculated. A simple line segment plot showing the fraction of total variance in the data as represented by each principal component, is in Figure 5.15. The first two components represent only about 64.57% of the variance, so more components are probably needed. The first three components represent 79.45% of the variability in the variables, so that might be a reasonable dimension to reduce data to.

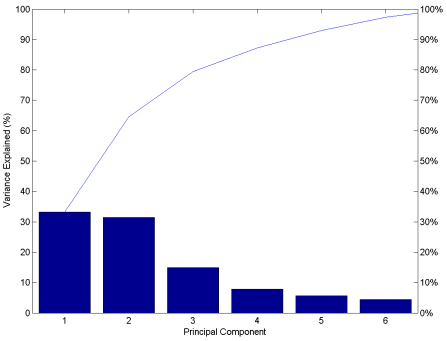


Figure 5.15: Scree plot of the percent variability represented by each principal component is 33.19%, 31.38%, 14.88%, 7.84%, 5.63%, 4.46%, 2.61%

5.5 Neutral network analysis

In our particular problem, the main population of bulk nuclear recoil events and the main population of surface electron recoil events are roughly gaussian distributions in many of the powerful discrimination parameters (e.g. primary phonon rise time, primary phonon delay time). However, the behavior and relationship of outliers from the main distribution are barely understood. Gaussian distribution is strongly assumed in the previously described methods. No such assumption is made in the correlated classic 2D timing cut analysis, but information such as correlations between outliers is not fully utilized. Artificial neural networks analysis is another statistical method to provide powerful solutions to overcome the above limitations and problems.

We may extract an extremely large number of variables from the collected raw data through different analysis algorithms, many of which have power to discriminate between types of events. The correlation matrix and the principal components analysis are two popular statistic methods used to find the correlation relation ship between variables, and reduce them to a suitable lower dimensional space for the neural networks. They are also an important data visualization and pattern detection method.

5.5.1 The basics of neural networks

An neural network is an interconnected group of nodes (or "neurons", "units"). Simple nodes are connected together in various ways to form a complex network. The complexity of the network behavior is determined by the neurons, the structures between neurons, and the connection parameters. The training algorithms are designed to adjust the weights of the connections in the network to produce a desired signal flow.

Neural networks are typically organized in layers. Two or more of the neurons can be combined in a layer, which includes the combination of the weights $\mathbf{W}$, the bias $\mathbf{b}$, the multiplication and summing operation Σ, and the transfer function (or "activation

function") f. A network could contain one or more such layers. Each input element x is connected to each neuron through the weight matrix $\mathbf{W}$. Each neuron has a summing operation that gathers its weighted inputs $\mathbf{x}$ and bias $\mathbf{b}$ to form its own output. This output then passes through the transfer function f to get to the end of this layer. After that it can be either the input of next layer, or the remain as the final output.

There are three most commonly used transfer functions. The first is hard-limit transfer function (a step function). The output is either 0 for input smaller than 0, or 1, for input greater than or equal to 0. These kinds of functions are useful for binary classification. The second is linear transfer function. The third one is a log-sigmoid transfer, also known as a logistic function, $y = \dfrac{1}{1 + exp(-a)}$. It compresses values between $\pm\infty$ into an output between 0 and 1.

5.5.2 Neural networks training and improving generalization

The possibility of learning is the most attractive and central idea in neural networks. During the training, neural networks are trained and adjust themselves with a new set of weight or bias parameters. They create in situ a model for the underlying generator of the data to make the best possible predictions when new input data is presented.

Neural networks training is usually a optimization problem of minimization using a smooth error function. The above network training with parameter adjustment is based on maximum likelihood which is achieved by minimizing the performance function (error function). The performance function is normally chosen to be the mean sum of squares of the network errors (mse) on the training set. Most of the algorithms have shown good performance in training neural networks in a wide range of problems. However there is no universally optimal algorithm. One of the challenges of actually using a neural

network in your project is to find the appropriate algorithm for these optimizations.

$$mse = \frac{1}{N}\sum_{i=1}^{N} e_i^2 = \frac{1}{N}\sum_{i=1}^{N}(t_i - a_i)^2 \tag{5.4}$$

Over-fitting is one of the problems during neural network training. The error of the performance function on the training set is trained to have a very small value, but the error is large when the network is presented with new data. This memorization of training examples is the over-fitting, since the network has not learned to generalize to new situations. This comes from using one of the maximum likelihood method to fit models to data. It is intuitive to allow the best set of parameters to match the dataset we see. However we should emphasize that the goal of neural network training is not to learn an exact representation of the training data itself, but to build a statistical model that generates the data. When the new inputs are presented to the network, it should make a good prediction. A network with too few parameters gives poor predictions, i.e. poor generalization, since it has too little flexibility. On the other hand, a network with too many adoptive parameters to adjust cannot distinguish relevant parameters and also fits too much noise on the training data. We must find a balance between these two to optimize the complexity of the model and thus achieve the best generalization.

One way to find the optimal balance is to vary the complexity of the network by changing the number of parameters in the network. We may change the number of hidden units, or start with relatively large network and gradually remove the least significant connections, or start with small network and add units during learning.

Regularization is another approach for improving generalization. This involves modifying the performance function. An additional penalty term is added that consists of the mean of the sum of squares of the network weights and biases (msw) to the performance function. Using this new performance function causes the network to have smaller weights and biases, and this forces the network response to be smoother and

less likely to over-fit.

$$msereg = \gamma mse + (1 - \gamma)msw \tag{5.5}$$

where γ is the performance ratio, and

$$msw = \frac{1}{n}\sum_{j=1}^{n} w_j^2 \tag{5.6}$$

The problem with structural stabilization or regularization is that it is difficult to determine the optimum value for the size of the adjustable parameters(for the case of structural stabilization), or performance ratio parameter(for the case of regularization). It is desirable to determine the optimal regularization parameters in an automated fashion. One approach to this process is the Bayesian framework. In this framework, the weights and biases of the network are assumed to be random variables with specified distributions. The regularization parameters are related to the unknown variances associated with these distributions. These parameters can then be estimated using statistical techniques. Bayesian methods allow the training done on the entire training data, without separating it into training and validation data. The relative importance of different inputs can be determined using the Bayesian technique of *automatic relevance determination*[49, 50]. If the particular coefficient is a large value, this indicates the corresponding inputs is irrelevant and can be eliminated.

The integral evaluation of the prediction distribution is difficult for most practical problems however, because they are analytically solvable for only a small class of likelihood distributions. Simple numerical integration algorithms are most likely impossible due to the high dimension of network parameters. The evidence procedure with approximation method to the posterior, or Monte Carlo method with Markov Chain sampling to evaluating integral numerically are extensively used with Bayesian methods with neural networks.

The evidence procedure is an iterative algorithm for determining optimal weights

and hyperparameters (the parameter of a prior distribution). Because it searches for optimal hyperparameters instead of integrating them out over all unknown parameters, it is not a fully Bayesian approach. Nevertheless it has given good results on many applications, and is less computationally costly than full Bayesian procedures such as Monte Carlo integration.

5.5.3 The implementations of neural networks

As indicated in the Matlab Neural Network Toolbox 6 User's Guide[51], the form of Bayesian regularization implemented in the Matlab neural network toolbox does not perform as well on pattern recognition problems as it does on function approximation problems.

Ian Nabney provides another neural network package in his book "NETLAB: algorithms for pattern recognition"[52] with relevant training algorithms more suitable for our classification purpose. The software package in NETLAB toolbox, which works with Matlab, is available to download for free [53].

In the CDMS neural network analysis where we classify events as part of a parent distribution of NR bulk versus ER surface events, efforts are devoted to finding a suitable training algorithm for the neural network method. The neural network is constructed with a simple single hidden layer, a log-sigmoid output backpropagation network with 30 neurons with 2 inputs, primary phonon delay time (pdelc) and primary phonon rise time (pminrtc). Input vectors and the corresponding target vectors are used to train a network until it can classify input vectors to the specific output vectors. Each input parameter is normalized to have mean zero, and unit standard variation. The transfer function logsig generates outputs between 0 and 1 as the neuron's net input goes from negative to positive infinity. The weights and biases of the network are assumed to be random variables with specified distributions. The regularization parameters are

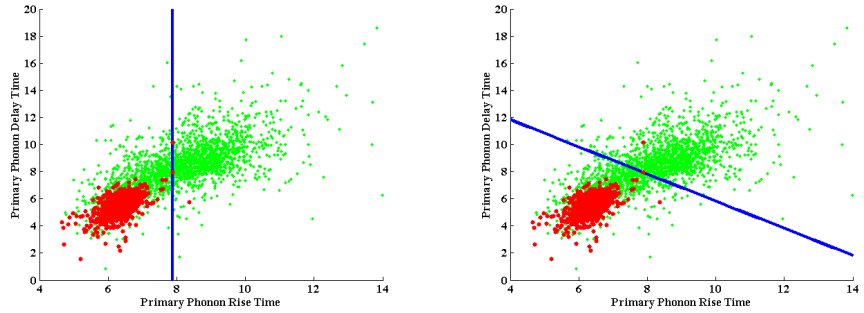
related to the unknown variances associated with these distributions. These parameters can then be estimated using statistical techniques. Once the network is trained by on the input data by a Bayesian technique with evidence procedure, it then decides the probability of classifying events as neutrons or betas for the newly presented data, on each position in the input parameter space.

5.5.4 Analytical calculation of the Bayesian decision boundary

Before we present the result from neural network analysis, we need to ask ourselves a question; how well can we trust the classification decision made by the trained network? This is a reasonable question to ask, given "traditional" analysis either provides us with a binary straight line cut in the one or two parameter space (5.16(a), 5.16(b)), or ellipse in the two parameter space (5.16(c)), while typically the neural network result presents nonlinear boundaries (5.16(d)).

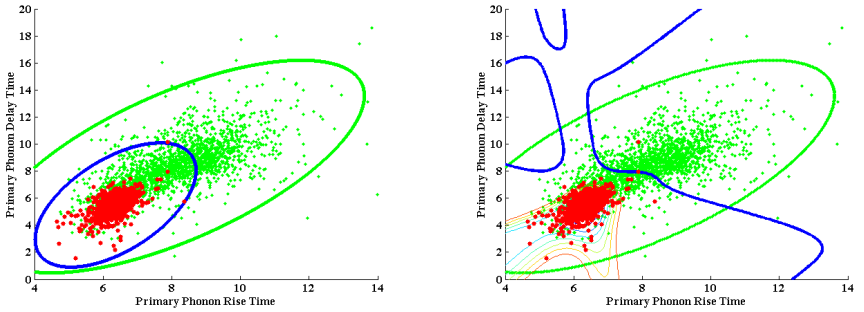
Although the nonlinear decision boundaries features seem non-intuitive at the first thought, it takes the advantage of not-quite-understood behavior and information based on the Gaussian distribution of outlier events for the surface event population, without the empirical requirement of any function form (e.g. polynomial) extrapolation from the main population.

The Gaussian mixture model mentioned in section 5.3 is a good way to benchmark the neural network output result, since the former result can be analytically obtained from the Bayesian theorem. Using the CDMS neutron and beta events for example, the goal is to classify them in such a way as to minimize the probability of misclassification. If we had collected a large number of such events from the Cf calibration and Ba calibration, we could find the fractions which belong in each of the two classes C_k. The *prior* probabilities (or mixing coefficients) $P(C_k)$ is the fraction of each type of events in each class. Thus, if in a particular detector, we collected 500 beta events, and 4500



(a) Straight line cut on phonon rise time

(b) Straight line cut on phonon rise time and delay time



(c) Ellipse cut based on 2D gaussian or χ^2

(d) Nonlinear neural network boundary

Figure 5.16: Different decision boundaries (blue lines) for different cuts, only one event is allowed to be accepted in the signal region

neutron events, we would have $P(C_1) = 0.1$ and $P(C_2) = 0.9$.

The probability of observing a value X with respect to the whole data set $P(X)$, irrespective of the class membership, is the linear combination of basis functions,

$$P(X) = \sum_{k=1}^{M} P(X|C_k)P(C_k) \tag{5.7}$$

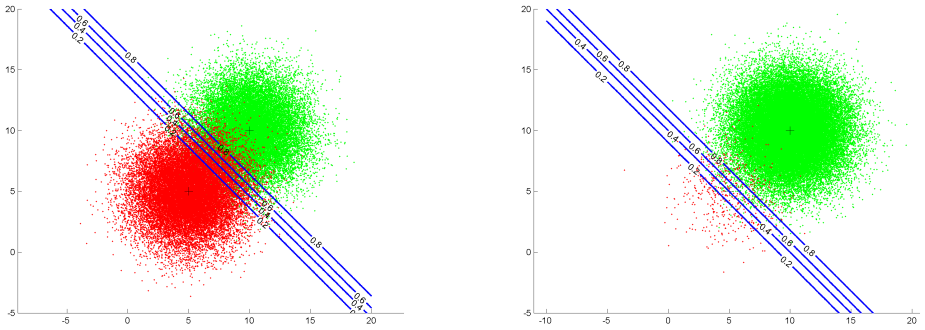
where $P(X|C_k)$, the component density functions, is the conditional probability specifies the probability for observation of X given that it belongs to class C_k.

For a given measurement of X, the probability that it belongs to class C_k is the *posterior* probabilities, which can be expressed in terms of the prior probability $P(C_k)$ from Bayes' theorem (named after British mathematician Thomas Bayes)

$$P(C_k|X) = \frac{P(X|C_k)P(C_k)}{P(X)} \tag{5.8}$$

Let's make some simple toy models to explicitly calculate the probability decision boundaries with Bayes' theorem. Start with two classes of events of multivariate Gaussian distribution, centered at position (5,5) and (10,10) respectively, but having the same standard deviation. The first case is shown in Figure 5.17(a), two classes have equal prior $P(C_1) = P(C_2) = 0.5$ resulting in the calculated probability decision boundaries as straight lines with 0.5 probability exactly in between the two classes. The second case is shown in Figure 5.17(b), one class has a greater prior $P(C_1) = 0.99$ and the other has a smaller value $P(C_2) = 0.01$. The probability decision boundaries are still straight, but now they are shifted towards the population of events with smaller prior. This shift biases the decision, so it's important to have event classes with the same prior during the neural networks training.

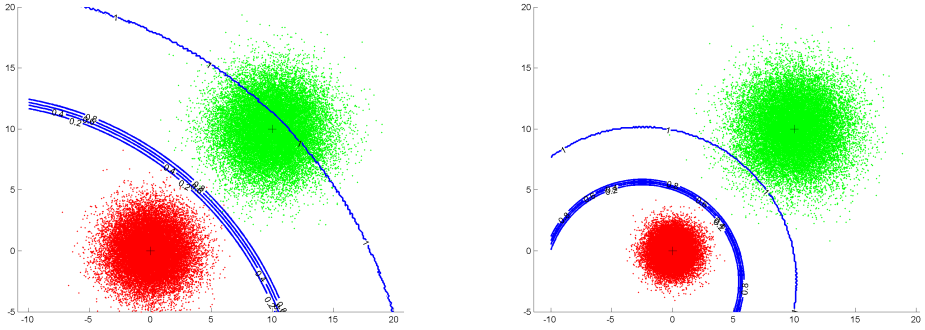
Now changing the lower left population to a smaller standard deviation (visually smaller radius), the probability decision boundaries become curve around the events with smaller standard deviation (Figure 5.18(a). Even smaller standard deviation results in smaller curvature of decision boundaries in Figure 5.18(b).



(a) Equal prior probabilities 50% neutrons vs. 50% betas

(b) Unbalanced prior probabilities 99% neutrons vs. 1% betas

Figure 5.17: Analytically calculated decision boundary for two classes with same standard deviation

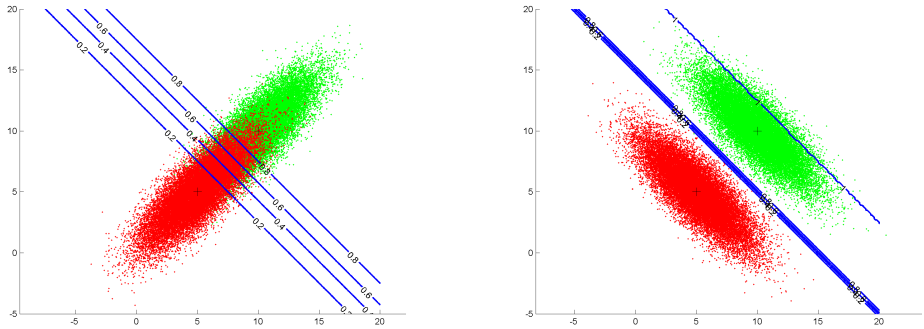


(a) Small betas radius

(b) Smaller betas radius

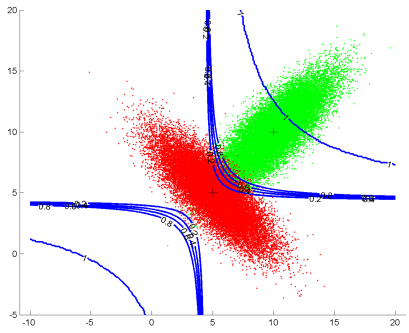
Figure 5.18: Analytically calculated decision boundary for two classes with different standard deviation

Next, we may change the shapes and orientations. Figure 5.19(a) has two ellipses with long axis both along the top right, bottom left direction. Figure 5.19(b) has two ellipses with long axis both along the top left, bottom right direction. Since the two ellipses are parallel, decision boundaries are still straight. Things become interesting when the long axis of two ellipses are perpendicular in Figure 5.19(c) and Figure 5.19(d). The calculated decision boundaries are curved in directions base on their posterior probability.

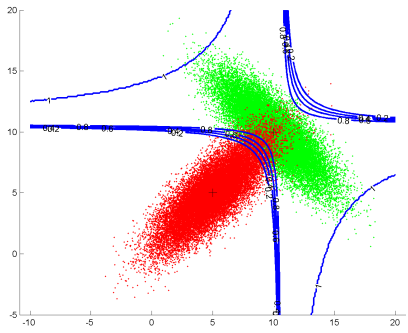


(a) Long axis parallel

(b) Long axis parallel



(c) Long axis perpendicular

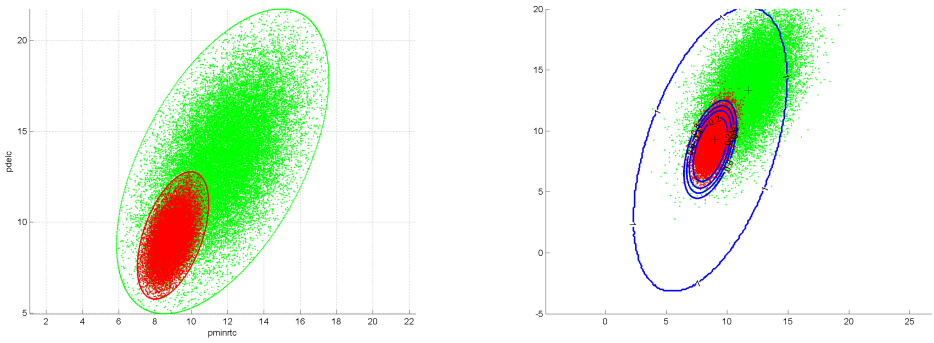


(d) Long axis perpendicular

Figure 5.19: Analytically calculated decision boundary for two classes with non-circle shape and different orientation

Now we are ready to understand the *real* data collected in the CDMS experiment.

For detector T1Z5, we calculated the mean and covariance matrix in two-dimensional space of phonon delay time and phonon rise time for neutron events and beta events separately, excluding some extreme outlier events beyond 6 σ, since values from these events are probably due to the improper event reconstruction. With the mean and covariance matrix, events are resampled (Figure 5.20(a)) for two reasons (1) to gain sufficient statistics, and (2) to generate two classes of events with same prior probabilities. Figure 5.20(b) is the analytical result of the resampled data.

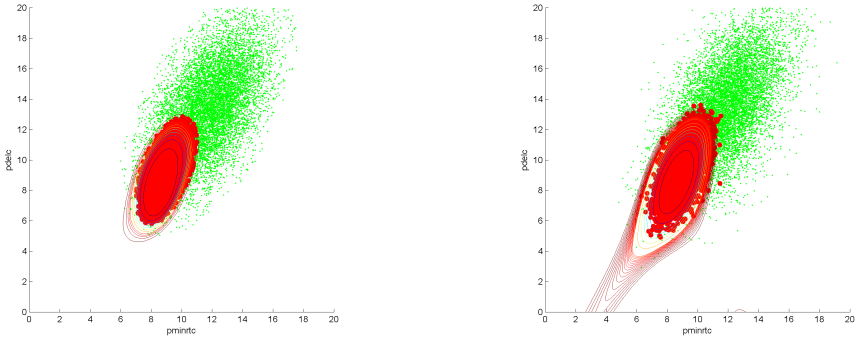


(a) Resampled neutrons and betas population based on same centers and covariant matrices, assuming 2D Gaussian distribution for both populations

(b) Analytically calculated decision boundary on resampled data with same centers and covariant matrices, assuming 2D Gaussian distribution for both populations

Figure 5.20: Analytically calculated decision boundary on resampled data

The neural network is trained on the same resampled data. The output provides the probabilities to be classified as each class on every position. The contour lines give the probability decision boundaries. The data in Figure 5.21(a) removed outlier events outside 4 σ, makeing the network result match closely the analytically calculated decision lines, while Figure 5.21(b) did not remove outliers. Now the contour lines are less like perfect ellipse shapes. This study provided confidence in the neutral network software and our implementation of it.



(a) neural network result training on data without outliers

(b) neural network result training on data with outliers

Figure 5.21: Neural network output Bayesian decision boundary

5.5.5 Neural network surface rejection cut

Figure 5.22 is a typical neural network result for one of the CDMS detector (T2Z2) with probability contour lines of high neutron probability colored warmer, and lower probability colored colder. Contour lines are only plotted through positions where beta events exist. Which probability line to choose as the cut is determined by how many beta leakage events are allowed in our signal region shown by the area to the right of the contour and to the left of the green ellipse. The green ellipse is a 4σ consistent cut line. The straight blue line in this figure is the classic 2D timing cut. It's obvious that the nonlinear neural network cut takes full advantage of the neighboring outlier events, with the lost of small neutron concentration in region B to trade for the high neutron concentration in region A, which improves the signal efficiency dramatically.

Not all detectors shown convincing efficiency gain though. Figure 5.23 shows detector T3Z3. If one or two beta leakage events are allowed, the neural network cut will bypass a large number of outliers at the top right corner, and grab the beta-events-almost-free region in the center of neutron population. We would argue this is probably not a problem for these reasons, (1) this only happens when the outlier situation is

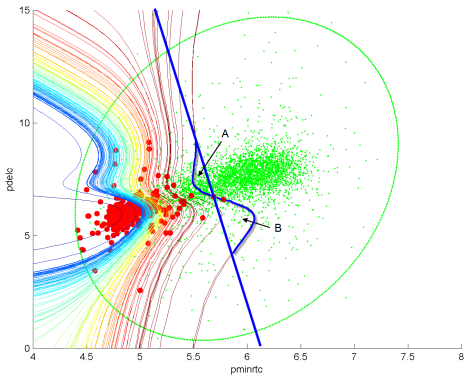


Figure 5.22: Comparison of neural network cut with 2D classic cut. Region B lost by neutral network to trade for region A. The total efficiency is gained.

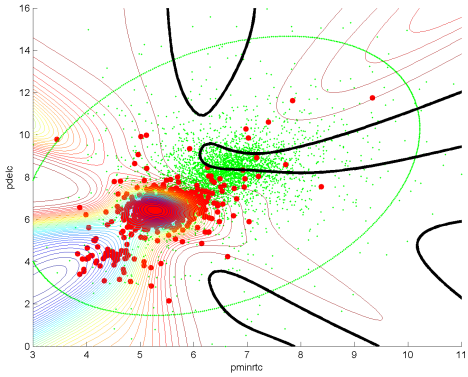


Figure 5.23: Neural network cut bypasses a large number of outliers at the top right corner, gains efficiency at the beta-events-almost-free region in the center of neutron population

already poor such that the classic cut is not able to give us much neutron efficiency, (2) the additional efficiency gained is from the correlation among the outlier events which is not possible get from the classic method, so the validation from classic cut is not a true argument.

Figure 5.25 tabulates the result of neural network analysis for all the working detectors in Run 123/124 analysis. The networks were trained with 2 input parameters. The analysis to follow used three: primary phonon delay time (pdelc) and primary phonon rise time (pminrtc) and the ratio of opposite sensor phonon amplitude (pfracc). The three dimensional network output contour line which allows one leakage event is plotted in Figure 5.24.

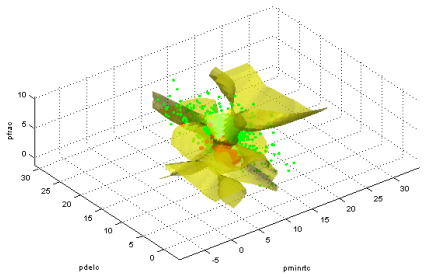


Figure 5.24: 3D inputs trained neural network output Bayesian decision boundary contour lines

To compare the surface event cuts neutron efficiency of classic 2D, 2-input neutral network, and 3-input neural network, the efficiency vs phonon recoil energy is tabulated in Figure 5.26, with green for the classic 2D cut, cyan for 2-input neutral network cut, and red for 3-input neural network. In general, 2-input neutral network cut is better than classic 2D cut in terms of efficiency, and the 3-input neural network cut is the best of the three.

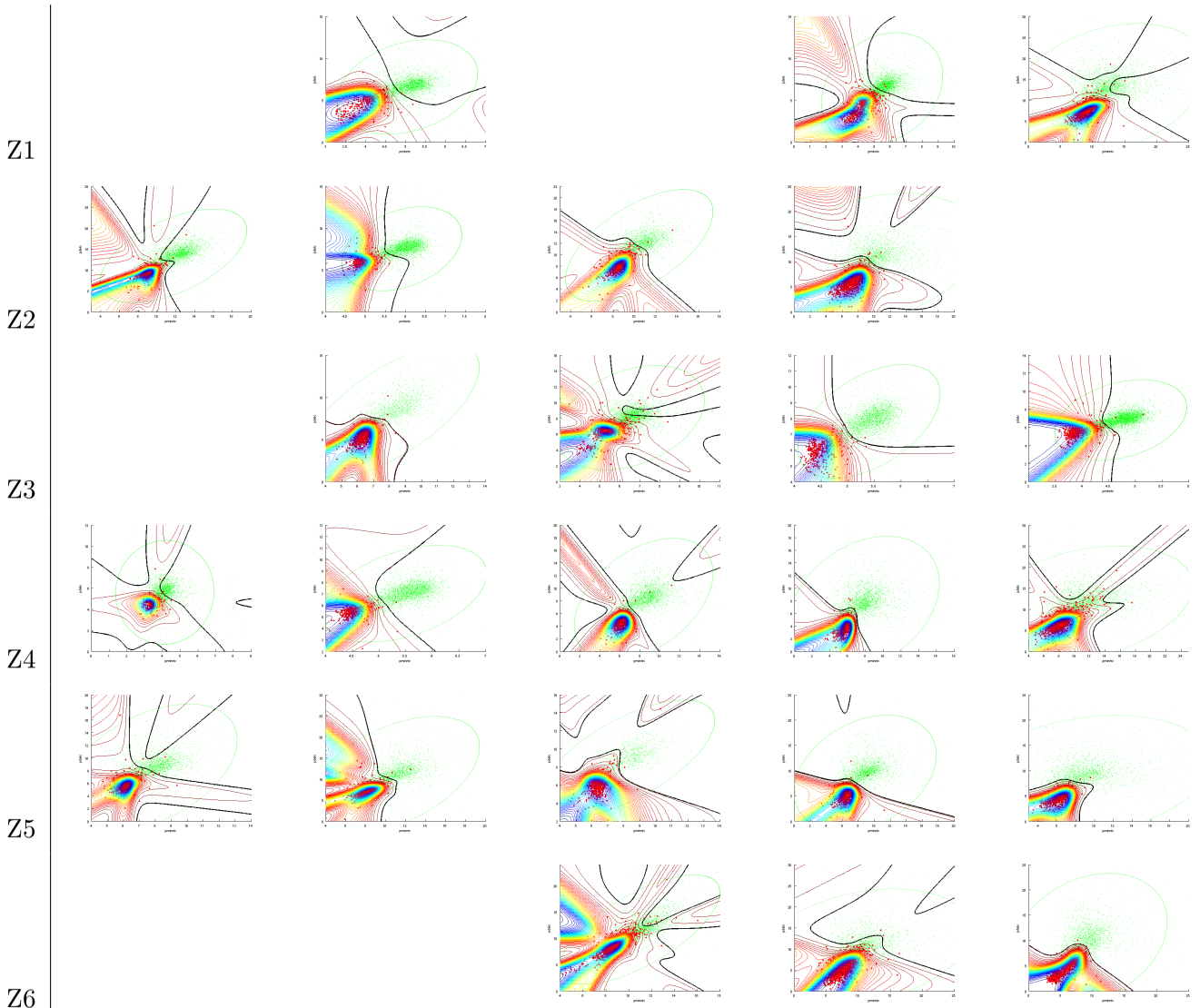


Figure 5.25: Trained Neural network output with equal probability spacing (e.g. 2%) decision boundary contour lines. The color of the contour lines indicate the classification probability. Solid black lines are where the set surface cuts

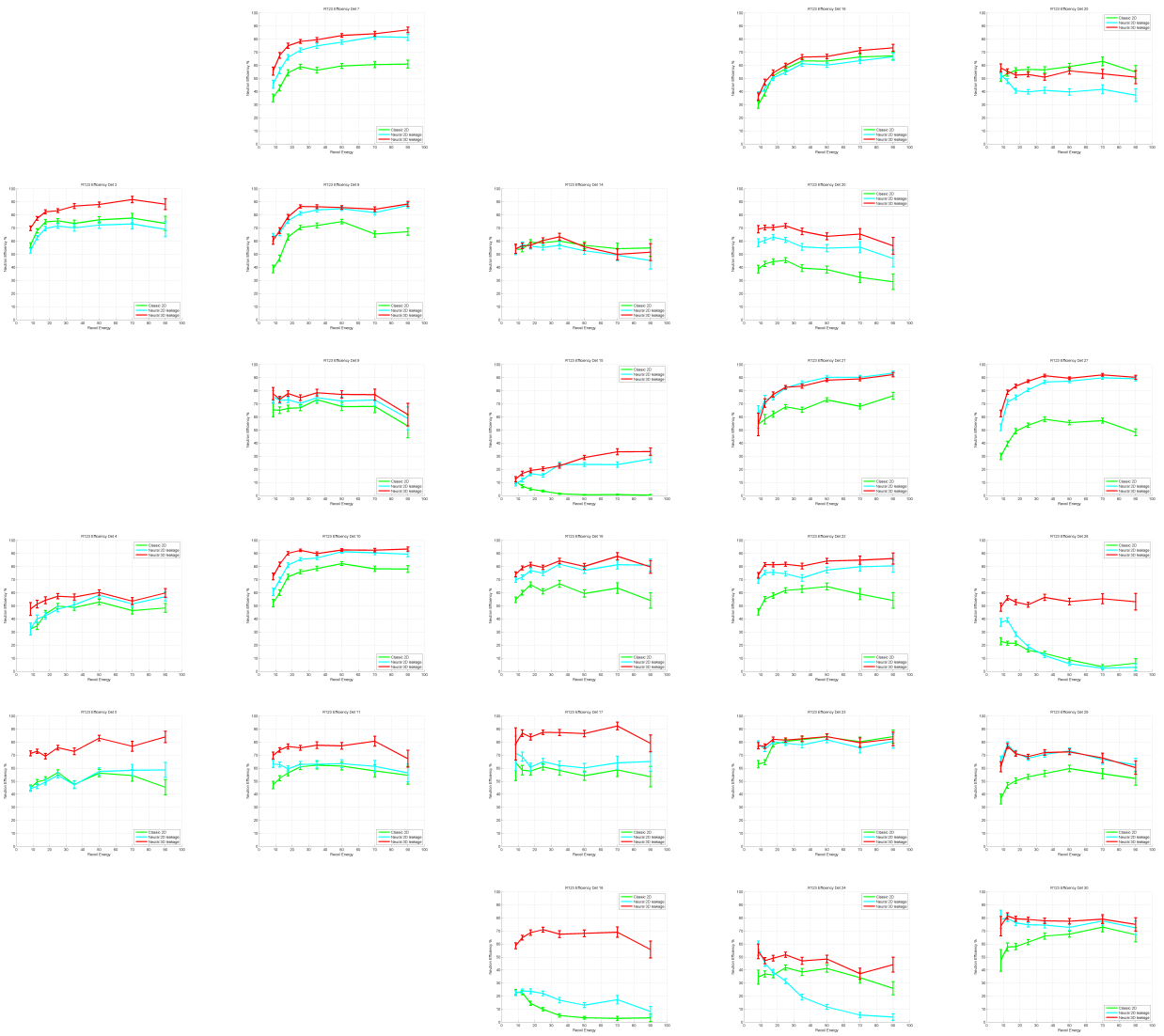


Figure 5.26: Neural network surface rejection cut neutron efficiencies

5.5.6 Neural network analysis result

With the increased neutron efficiency compare with the classic 2D timing cut, the total exposure after all the quality and analysis cuts will also be increased. Figure 5.27 is the Nuclear recoil signal acceptance efficiency of Run 123/124 WIMP search as a function of recoil energy. The first three curves represent the cumulative efficiencies at various stages during the analysis, with gray line after quality, singles, veto cut, dash dotted gray line after fiducial volume cut, black dash line after nuclear recoil band cut. The phonon timing cut imposes the highest costs in signal acceptance. The final efficiencies after two different timing cuts are shown with dash dotted red line after the classic 2D timing cut, and solid red line after the neural network timing cut.

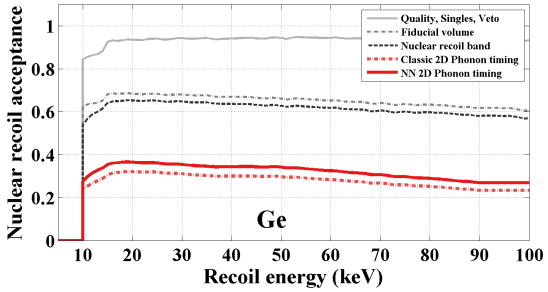


Figure 5.27: Nuclear recoil signal acceptance efficiency of Run 123/124 WIMP search as a function of recoil energy, averaged over all detectors used in the analysis, weighted by their individual livetimes.

Once the neural network is trained based on the neutron calibration and beta calibration data, and the phonon timing cut is determined by the desired total number of allowed background leakage, the signal region is unmasked. Nuclear recoil single events are presented to the timing cut to see whether they fail or pass this cut. For Run123/124 Ge analysis, with the neural network surface rejection timing cut assigned to have the

same beta leakage fraction, there are a total of two candidate events pass the timing cut. These two events are both in detector T4Z6, and correspond to the exact same two events passed the Run 123/124 energy dependent χ^2 cut, the secondary timing cut analysis. Their information is listed below in table 5.6.

Table 5.6: Selected parameters of the two candidate events passing the neural network timing cut.

Series Number	Event Number	pric	yic	xdel	ydel	pminrtc	pdelc	pfracc	tfracc
1610300203	50403	16.4	0.22	1.98	-76.0	7.87	11.25	3.62	15.7
1702172047	90095	14.3	0.24	-59.8	-2.73	5.59	12.78	3.11	7.70

Figure 5.28 shows where these two candidate events are located in the phonon rise time vs delay time plot, with both classic 2D timing cut and neural network timing cut displayed. Blue dashed straight line is the classic 2D timing cut, two dashed black lines are consistent cuts. The signal region is within the consistent cuts while greater than 2D timing cut. The neural network cut is the thin black curved line, while the green ellipse is the consistent cut. Nuclear recoil single events in neutron calibration are plotted in green, Beta surface events in Beta calibration are plotted in red. The masked nuclear recoil events are in blue dots, with two events passing the neural network cut circled. These two events are close to the lower energy threshold, and also near the timing cut threshold, consistent with the possible background leakage into the signal region.

Figure 5.29 shows the upper limits on WIMP-nucleon cross sections calculated from the Ge neural network analysis reported in this section using standard assumptions for the galactic halo [54], and assuming they are background events. Also shown is the limit obtained from previous Soudan one and two tower re-analysis [4] and our published PRL paper with classic 2D timing cut [5]. Due to the presence of these two events, neural network timing cut analysis has a weaker limits to constraint on WIMP-nucleon cross

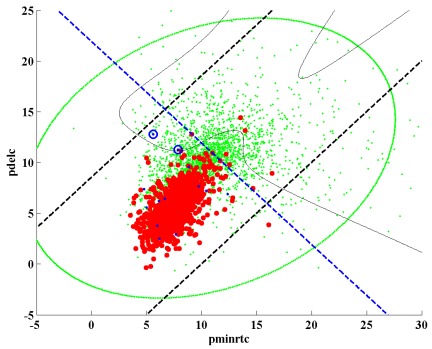


Figure 5.28: Two NR single events with phonon recoil energy 14.3 keV and 16.4 keV in T4Z6 pass neural network timing cut.

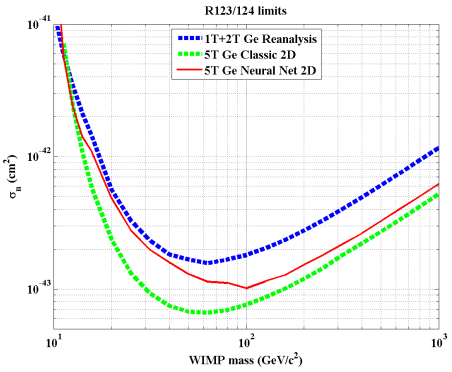


Figure 5.29: WIMP-nucleon cross section upper limits (90% C.L.) versus WIMP mass. The upper curve (dotted) is the result of a re-analysis [4] of our previously published data. The lower curve (dotted) is the result of classic 2D timing cut. The middle solid red line is the limit from this neural network analysis.

sections, even with greater signal acceptance fraction.

5.5.7 Neural networks limitations

The neural network method can be extremely robust if the model, cost function and training algorithm are appropriately selected. But the choice of model depends on the data representation and the application; selecting and tuning a training algorithm requires a significant amount of trial-and-error experimentation. It is essential to have a good understanding of the underlying physics.

There are also some other limitations for the neural networks method: (1) instability of training results when a small change in the training data leads to a large change in the networks, both in its structure and parameters, and (2) the temptation to use it as a black boxes. Apart from defining the general architecture of a network and perhaps initially seeding it with a random numbers, the user has no other role than to feed it input and watch it train and wait for the output. Some software does allow the user to sample the networks 'progress' at regular time intervals, but the learning itself progresses on its own. The final product of this activity is a trained network that provides no equations or coefficients defining a relationship beyond it's own internal mathematics.

Chapter 6

Five Tower WIMP Search Analysis and Results

The first five tower WIMP search runs at Soudan underground mine is a combination of two data runs from October 2006 to July 2007, separated by a short period of cryogenic maintenance. Run 123 accumulated data from October 21, 2006 through March 20, 2007; Run 124 is from April 20 to July 16, 2007.

WIMP-candidate events are identified after a series of processing, event reconstruction, calibration, position correction, data quality cuts, and physics cuts. All the analysis was performed blindly without looking at the WIMP signal region before all the cuts are finalized. The analysis pipeline and cuts are summarized in this chapter, with the estimates of the exposure and the results of the analysis.

CDMS achieved the world leading WIMP-nucleon spin-independent interaction limits of 6.6×10^{44} cm^2 on WIMP dark matter detection for WIMP masses above 44 GeV/c^2, with zero background.

6.1 Data acquisition and processing

After the DAQ recorded the raw data on the disk and backed up on tapes, the DarkPipe package (compiled executables written in Matlab) started processing the raw data. It performed a series of analysis on the data in order to extract useful quantities, such as pulse risetimes, noise information, event triggers, etc. It then created a matrix collection of "Reduced Quantities" (RQs) which store values calculated from the raw event information. Triggering and veto information were calculated in this step, as was the optimal filters, the pulse-walking algorithm. PipeCleaner (also compiled executables written in Matlab) is the name of a second analysis package afterwards, which applied calibrations and computed cuts to the RQs generated from DarkPipe, PipeCleaner's output consisted of "relative reduced quantities" (RRQs) in more directly physical units. It also produced diagnostic plots for fast feedback. Details of DarkPipe and PipeCleaner can be found in [4].

Both RQs and RRQs can be accessed through the CDMS Analysis Package (CAP) within Matlab. CAP provides a fast and flexible interface to the CDMS data, due to the advanced management of combination of RAM and disk retrieval and buffering of these quantities through specialized functions.

All the data was initially processed on the Soudan Analysis Cluster (SAC) to produce both RQs and RRQs for data quality monitoring purpose only. The official data was reprocessed later on Fermilab computer farm with new optimal filter templates, energy calibration, and new phonon correction table.

6.2 Blinding procedure

A blind procedure was performed to avoid bias. The blinding procedure detailing the WIMP search data blinding criteria and the actual blinding cut for the Run 123

and 124 analysis was proposed and accepted by the collaboration in early 2006 prior to beginning data analysis.

The blinding cut was defined during initial data reduction to remove events inside and close to the signal region from WIMP search data sets. The defined the mask cut was using Barium and Californium calibration data taken between October 2006 and January 2007 and initially processed on the SAC, and applied later to the final processing of the WIMP search data at Fermilab. Two versions of official data distribution were released to the collaboration, one with the masked events removed from the WIMP search data, the other contained masked events.

When developing the analysis, all cuts are set using calibration data and WIMP search data outside of the masked region, so no bias could influence the result. After WIMP selection criteria were finalized and approved by the collaboration, the masking was then relaxed in two stages. The first step was to release the actual signal region for events that fail the timing cut and checked for statistical consistency with our background expectations in an effort to find failures of background estimation or data quality. If the expected background was low enough, we proceed with the second step to reveal the events passing the timing cuts. Otherwise, we would pause until an appropriate action being selected.

Events masked by blinding cut represents a simplified and conservative version of the WIMP-search analysis. It satisfies the following criteria,

1. **Veto anti-coincident:** no signal in the veto panels immediately preceding the global event trigger

2. **Inner electrode:** energy in the outer ionization electrode is consistent with noise

3. **Energy range:** phonon recoil energy is between the detector's analysis threshold and 130 keV

4. **Nuclear recoil band:** within the 3σ nuclear recoil band defined from neutron calibration data with a ^{252}Cf source.

5. **Single event:** phonon energy within the 6σ noise in all but one detector, since we expect WIMP candidate events only interact once with one ZIP detector

The above blinding procedure was implemented during the final re-processing of the WIMP search data at Fermilab by the PipeCleaner package. The official distribution of analysis data was with masked events removed.

6.3 Quality cuts

To ensure data quality and to allow only well-reconstructed events to be used in the analysis, a series of quality cuts were applied. There are two major reasons for poor data quality: one is poor detector performance, the other is events that are poorly reconstructed.

6.3.1 Detector selection

The data taken during runs 123 and 124 were from 30 detectors (19 germanium and 11 silicon) which were organized into 5 towers of 6 detectors each. Out of these 30 detectors, only 23 detectors were used for the WIMP search analysis during Run 123. The other 7 detectors are known to have poor performance for the reasons listed below.

- T1Z1: Unrecoverable broken phonon channels B, C, and D, and no working LEDs
- T1Z3: Disconnected Qouter amplifier means that the Qouter electrode cannot be biased
- T1Z6: Poor neutralization due to no working LEDs

- T2Z6: Poor neutralization due to no working LEDs
- T3Z1: Phonon channel C has unstable behavior which makes the calibration data sample insufficient for further analysis
- T5Z2: Qinner electrode is shorted to chassis meaning that the Qinner electrode is not biased
- T5Z6: Very poor signal-to-noise for the phonon sensors. Decision to exclude this detector was made before the unblinding since we saw higher than expected leakages on this detector.

In addition to the above-mentioned problematic detectors, in Run 124, T2Z5 and T3Z6 were not used as WIMP search detectors due to poor neutralization. In order to get a result in a timely manner, we decided not to incorporate the shorter Run 124 data into Run 123 analysis unless the data from each detector was similar between the two runs. This avoided the effort of regenerating the phonon position correction table, or making an entire set of new cuts, as would be required if there were a change of bias and neutralization status between two runs. In order to determine which detectors in Run 124 were in a state very similar to what they were in during Run 123, a number of Kolmogorov-Smirnov Tests (KS-tests) were used to evaluate each Run 124 detector data with Run 123 templates. Detectors listed below were used for the WIMP search for Run 124: T1Z2, T1Z5, T2Z4, T4Z3, T4Z4, T4Z5, T4Z6, T5Z4, T5Z5. Detectors not selected were not necessarily bad, and were still used to determine whether an event was a single or multiple.

6.3.2 Stable detector tuning history

Detector tuning allow us to gain the maximum signal to noise without losing the timing information for surface event rejection. Uniformity of the timing parameters

among the 4 phonon sensors for each detector is avhieved during the tuning process. The SQUID must be tuned well with desired bandwidth, and the QET must be tuned to maintain phonon timing response balanced across each detector. Most detectors had achieved stable tuning prior to the official R123 WIMP search start date of October 21, 2006. Exceptions are:

- T3Z4 (October 26, 2006)
- T4Z6 (October 27, 2006)
- T5Z3 (October 31, 2006)

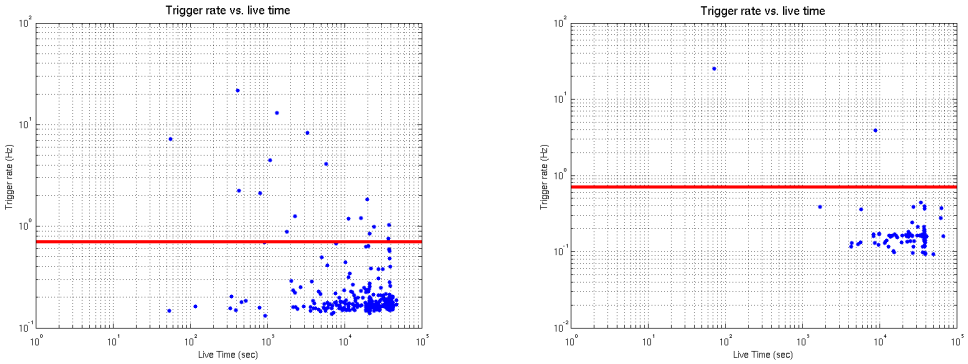
The cut cStabTuning_123 is detector dependent and removes all datasets for each detector prior to the date its tuning stopped changing in the early stage of Run 123. The use of this cut allows the Run 123 start date to be moved prior to the date where the tuning of every detector was complete.

6.3.3 Trigger stability

WIMP search runs with abnormally high trigger rates usually indicates electronics problems or increases in phonon noise. The cut cTrigBurst_123 removes trigger burst periods from the WIMP search data set. Two types of trigger bursts are excluded. The first class of trigger burst excludes periods during which at least 100 consecutive non-random events have trigger rate > 0.7 Hz. Events with very short periods between trigger bursts (less than or equal to two events in between) are also removed. Figure 6.1 identifies and flags questionable data sets with higher rigger rates (> 0.7 Hz).

The total Run123 (Run124) live time was reduced by 0.0103% (0.00614%) while the total number of events was reduced by 7.31% (5.57%).

The second class of trigger burst excludes periods of triggers on baseline noise two



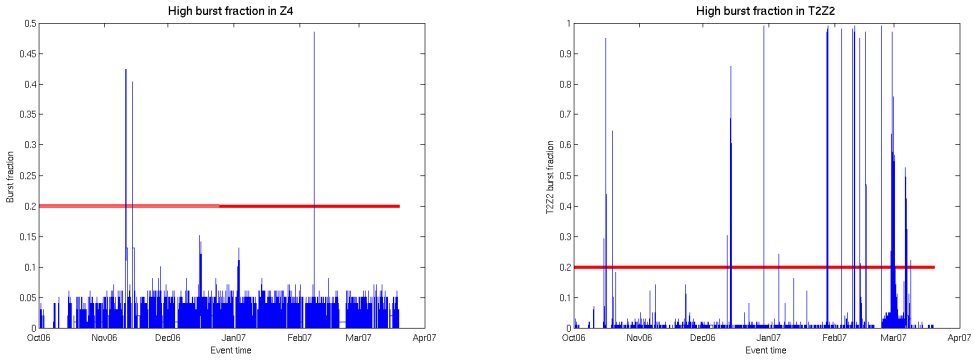
(a) Run 123 (b) Run 124

Figure 6.1: Trigger rate vs. Live time for both Run 123 and Run 124

specific detectors T1Z4 (rare) and T2Z2 (common). The periods are identified by fraction of events triggered with negligible phonon energy (prg $<$ 1 keV) in these two detectors. The fraction was calculated as an un-binned statistic: for any given event, count the number of detector specific triggers with prg $<$ 1 keV in the next 49 and previous 49 events, as well as the event itself. Exclude this period if the fraction is greater than 20%. Figure 6.2 shows the periods with high trigger fraction by these two detectors. This class of removing trigger burst reduced the livetime on T1Z4 (T2Z2) by 0.0647% (0.921%), reduced events by 0.0934% (11.9%).

6.3.4 Negative phonon events

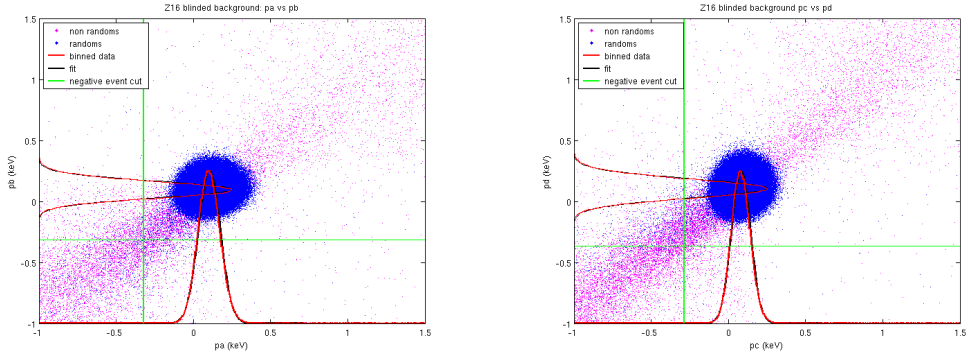
A number of events with negative phonon energy in one of the phonon channels exist. To remove these negative going phonon pulses, a detector dependent cut, cNegPhononPulse_123, was defined to reject events with a phonon energy well below the noise blob (6 sigma below a channel's mean) in a given detector's channel. DarkPipe does not search for negative-going phonon pulses, so negative phonon energies are typically a mis-constructed event from shifted noise baselines. Figure 6.3 is a typical noise blob on detector T3Z4, all events are plotted in magenta, random triggers in blue, data



(a) T1Z4 (b) T2Z2

Figure 6.2: Trigger fraction by specific detector of T1Z4 and T2Z2

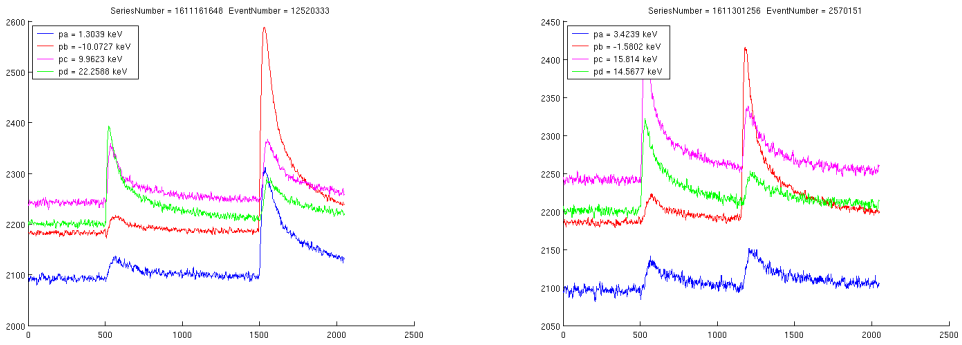
histogram in red, gaussian fit in black, the cuts ($\mu - 6\sigma$) are in green lines.



(a) Phonon Channel A and Channel B Noise (b) Phonon Channel C and Channel D Noise

Figure 6.3: T3Z4 phonon noise blob on each channel

Generally it is redundant to other data quality cuts, since other cuts are highly efficient at removing events with negative phonon energy in at least one phonon channel. Therefor this cut removes very few events. In Run 123, two events (in Figure 6.4) are clearly due to pile up of the high trigger rate during the Ba run, but they were not rejected by other quality cuts, so this specific cut on removing negative phonon pulses and pileup events is good to have.



(a) Trace 1 (b) Trace 2

Figure 6.4: Pile up events in Ba calibration data not rejected by other data quality but negative phonon cut

6.3.5 Other detector performance and reconstruction quality cuts

A number of other studies were performed to ensure the stability of detector performance and event reconstruction.

1. Charge noise stability

   cHighQNoise is detector dependent and select time intervals with high charge noise.

2. Charge crosstalk

   The vast majority of detector-to-detector crosstalk seems to be caused by floating grid sections on the phonon side of tower 1 and tower 2 detectors. The charge crosstalk cut, cNoDDxtalk_123, removes detector-to-detector crosstalk and is only applied to towers 1 and 2.

3. Electronic glitches

   Electronics glitches typically cause multiple phonon triggers with accompanying charge triggers of fewer multiplicity. cGlitch_123, selects events caused by trigger characteristics of these electronics glitches. We thus use the following logic to

select them in Run 123/124. An event is a glitch event if and only if the following two logical conditions hold: (1) number of phonon triggers greater than or equal to 4, and (2) the difference between the number of phonon triggers and the number of ionization triggers is greater than or equal to 4. These conditions were optimized for maximal removal of glitch events while retaining the majority of real events.

4. Detector neutralization

   Loss of neutralization can be dangerous to our results because it causes electron recoil events to droop down in yield, potentially into the nuclear recoil band. During this analysis there was a lot of work done to identify periods of decreased neutralization. There were three specific studies done which led to cuts attempting to remove these periods of bad neutralization.

5. NuMI neutrino beam

   The NuMi neutrino beam from Fermilab to the MINOS detector in the Soudan mine may interact in the shield or rock to generate muons and associated neutrons. Although it is not expected to form any significant observable unvetoed background for CDMS, in an abundance of caution, cNuMI_R123 vetoes any events arriving within a window of 10 microseconds for the neutrino beam with a period of 2.5 seconds using the MINOS GPS time.

6. Ionization goodness-of-fit

   Primarily designed to reject pileup events, cChiSq_123 removes events with poorly-fitted charge pulses on a given detector, as measured using the optimal filter χ^2 quantity.

7. Pre-pulse phonon noise

   Also primarily designed to reject pileup events, cPstd_123 removes events with noisy pre-pulse phonon baselines on a given detector.

8. Pre-pulse charge noise

   Primarily removes events during noisy periods of the cryocooler cycle on sensitive detectors, as well as some pileup events, cQstd_123 removes events with noisy pre-pulse charge baselines on a given detectors.

9. Phonon start time

   cGoodPStartTime rejects cross-detector pileup events for which the walked phonon start times on a given detector lie outside the search window used by the optimal filter algorithm to find the charge pulse. Such events will generally have charge pulses which begin outside the search window and thus give mis-fitted charge amplitudes.

6.4 Background cuts

After the application of data quality cuts, event characteristics were examined to define a parameter space in which only WIMPs are to be found. These isolate the veto-anticoincident, bulk, signal nuclear recoil scatters that correspnd to the signal expected for a WIMP.

6.4.1 Muon veto cut

The CDMS experiment is located half a mile (2090 meters of water equivalent (m.w.e.)) underground, which reduces the surface muon flux by a factor of 5×10^{-4}. The remaining comic-ray muons (at the rate of 1 muons per minutes in the veto pannels) can still interact with the shielding materials and generate low-energy neutrons that cause nuclear recoils in detectors. To reduce these comic-ray-induced nuclear recoils, the lead (to shield gammas) and polyethylene (to shield neutrons) passive shielding surrounding the detectors are covered with forty scintillator panels that compose the active muon

veto system. Figure 6.5 shows a schematic drawing of the CDMS veto system.

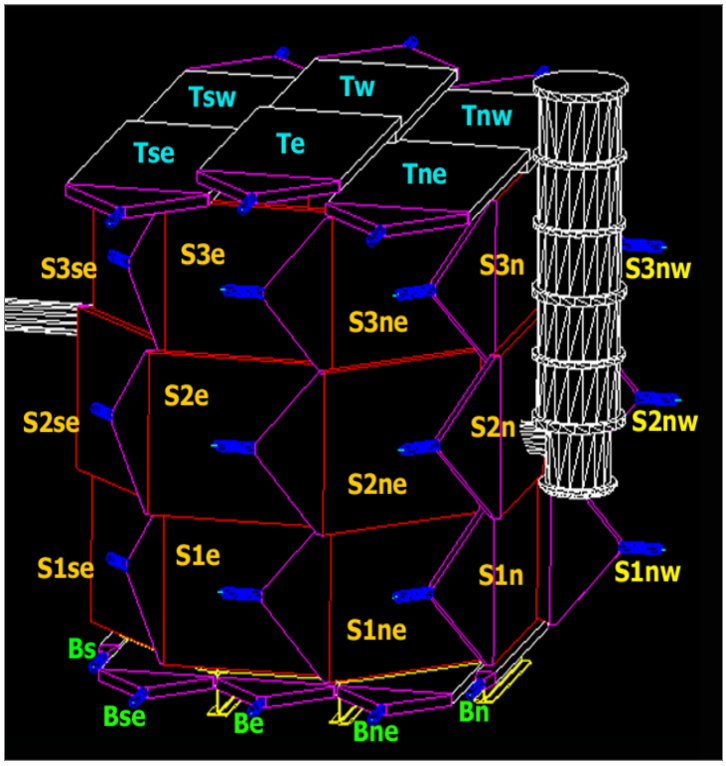


Figure 6.5: schematic drawing of the CDMS-II veto system

Signal events must not satisfy the following two condistions:

- a veto trigger within 50 μs before the detector trigger
- A reconstructed energy in the veto above a set threshold anywhere in the -180 to +20 μs window around the detector trigger for which veto pulses are recorded.

The energy threshold used in the cut for this run was set on a panel-by-panel basis to maximize efficiency to detect muons without too much dead time, compared to the previous run where a constant threshold for all panels at 50 pC was set for all panels. Figure 6.6 shows the energy deposition histogram for panel 4, one of the top panels, in units of pC, of 1 Million events recorded in Run 123. The dashed green vertical line is

the 50 pC mark were the threshold was previously set for an old version of this cut and the red dashed is were it is now set for cVTStrict_123.

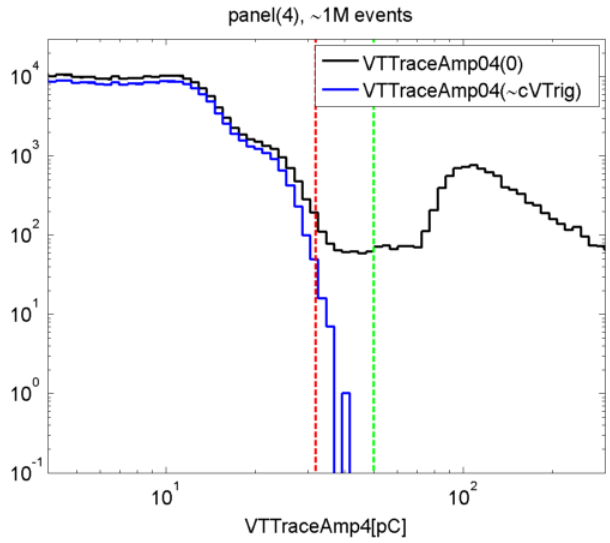


Figure 6.6: Histogram of deposited charge energy by largest veto hit before global trigger for veto panel 4(one of the top veto panels). The black line are all the veto trigger events, the blue line are events for events that are not caused by the veto multiplicity trigger, mostly induced by ambient gammas.

6.4.2 Energy thresholds

Events at low energies sometimes get mis-reconstruction due to the noise. Such performance degradation for the low pulse amplitudes can be avoided by applying thresholds on the phonon and charge signals in the analysis.

Charge threshold

The charge threshold cut (cQThresh_123) rejects events for which the ionization signal is consistent with noise. The cut threshold is set detector dependently at the mean + 4σ upper edge of the Q_{inner} noise blob, based on a gaussian fit to DAQ random trigger events in the WIMP search run. Figure 6.7 shows an example fit for T1Z2.

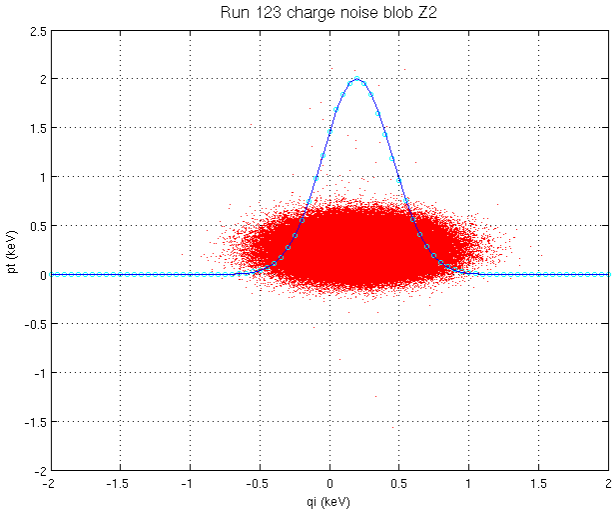


Figure 6.7: Gaussian fit to the Q_{inner} noise blob for T1Z2 used in the charge threshold cut, superimposed on the noise blob in Q_{inner} and pt for this detector.

There were variations in the ionization resolution with time however (see for example in Figure 6.8 for Tower 1 in Run 123), which could lead to increase the leakage during the poor ionization resolution period. The noise blobs in each data series were then independently measured for the mean and width, the final thresholds were set at the maximun of the overall averaged 4 σ threshold and the run-dependent 4 σ upper edge of the Qinner noise blob.

Phonon threshold

The phonon threshold is cut set on phonon recoil energy. Low energy events have poor discrimination due to poorly-defined yield and noisy timing parameters. The threshold is also detector dependent to satisfy three conditions:

- The surface event rejection cut must have significant efficiency. This was the primary consideration in setting the previous R119 threshold to 7 keV.

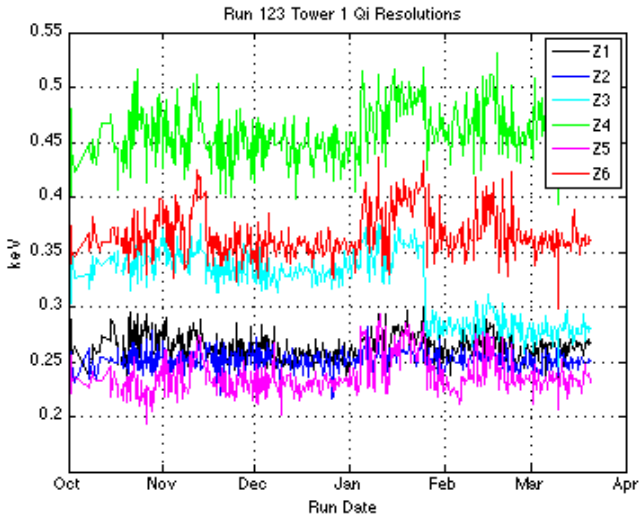


Figure 6.8: Variation in Tower 1 Qi resolutions during the course of Run 123, as measured using gaussian fits to random trigger events during each data series. Figure. by Jeff Filippini

- The charge threshold cut and the edge of the electron recoil band must not eliminate the nuclear recoil band.
- There must be a significant population of surface events on which to characterize the cut, particularly the systematic error in the background leakage

Table 6.1 shows the thresholds suggested by the latter two of the above conditions, imposing a mandatory minimum of 7 keV on Si and 10 keV on Ge, and are the current recommended analysis thresholds.

6.4.3 Ionization yield bands

The reduced ionization yield of nuclear recoils with respect to electron recoils is the strong and primary discrimination power in CDMS to reject bulk electron background

Table 6.1: Phonon recoil energy thresholds (in keV) for the Run 123 and 124 analysis. Silicon detectors are in brakets.

Detector	T1	T2	T3	T4	T5
Z1		(7)		(10)	10
Z2	10	(7)	10	10	
Z3		10	(7)	(15)	(7)
Z4	(10)	(7)	10	10	10
Z5	10	10	15	10	10
Z6			10	10	20

events. It's defined as the ratio of the ionization energy to the phonon recoil energy,

$$y = \frac{Q}{E_r} \tag{6.1}$$

where Q is the ionization energy collected on the charge channels, and E_r is the phonon recoil energy on the phonon channels corrected with the Luke phonon contribution. The ionization yield of a nuclear recoil varies with energy and various with crystals (atomic number and atomic mass). Lindhard *et al.* provided explanation of this phenomenon. Lindhard *et al.* also computed quantitatively the energy dependence of this yield as a function of atomic number and atomic mass in 1936 [44]. The Run 123 band fits are compared with Lindhard theory in Figure 6.9. The agreement is generally very good for Ge detectors, but less good for Si detectors. The source of the Si discrepancy is unknown, but we use bands that are fitted to our calibration data, not to the theoretical curve.

There are a couple of reasons we need to define ionization yield band in the analysis,

1. WIMP search signal region

   The nuclear recoil band is needed as part of our signal selection box for the WIMP-search data, and to calculate the signal efficiency.

2. Low yield surface events (betas) definition

   A sample of low-yield surface events is need to define the surface event discrimination cuts. Selecting this sample requires a careful study of the low-yield edge

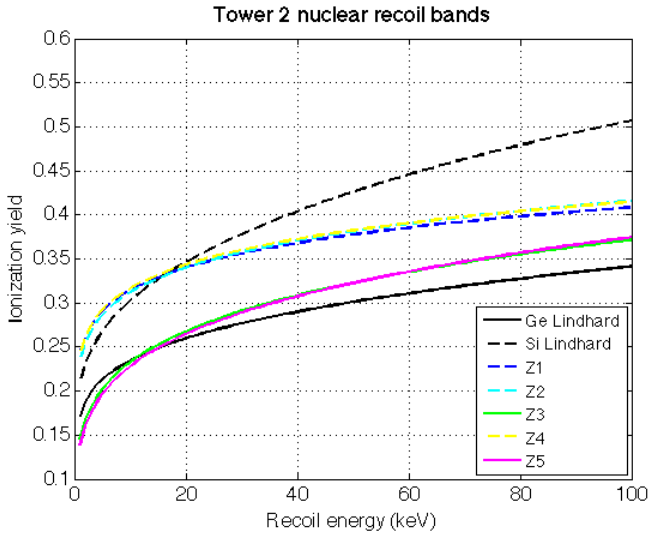


Figure 6.9: Comparison of Lindhard bands (black) with fitted nuclear recoil band medians (colored) for Tower 1 Ge (solid) and Si (dashed) ZIPs.

of the gamma sample.

To calculate the electron recoil bands and the nuclear recoil bands, we first fit Gaussians to the distributions of ionization yield y for both type of recoil events. The electron recoil events are from the ^{133}Ba calibration gammas and the nuclear recoil events are from $^{2}52$Cf calibration neutrons. The means and standard deviations are fitted versus yield in several recoil energy bins. The means and standard deviations in each bin are then fitted as a function of phonon recoil energy to define the recoil bands. The means are fitted to

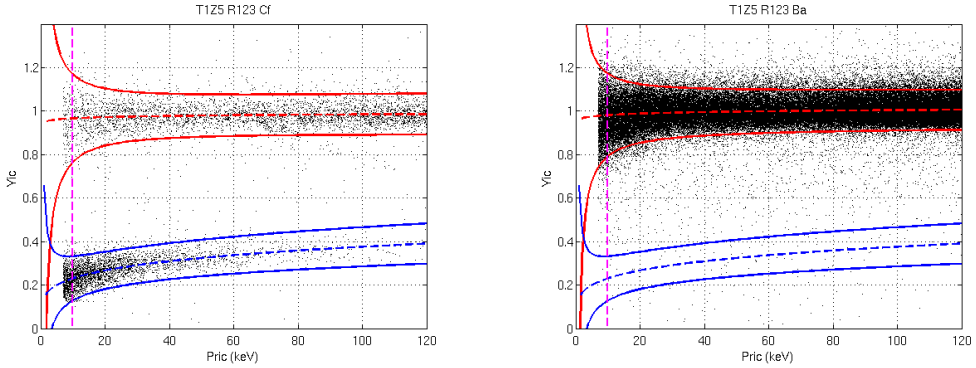
$$\mu_y = \alpha_1 \cdot E_r^{\alpha_2} \tag{6.2}$$

The standard deviations are fitted to

$$\sigma_y = \frac{\beta_1^2 \cdot E_r^{\beta_2} + \beta_3^2}{E_r} \tag{6.3}$$

The bands are then taken to be $\pm\, 2\sigma$ for both electron recoil bands and the nuclear recoil bands (Figure 6.10). The efficiency of the nuclear recoil band cut is determined

by the nuclear recoil events in the neutron calibration. Figure



(a) Events from ^{252}Cf neutron calibration data (b) Events from ^{133}Ba gamma calibration data

Figure 6.10: Ionization yield band for detector T1Z5 as a function of phonon recoil energy. Red lines are calculated electron recoil bands and blue lines are nuclear recoil bands. The dashed lines are the center of the bands, the solid lines are the $\pm$ 2σ edges.

A WIMP candidate event must satisfy two criteria based on ionization yield determinations: (1) the event's ionization yield must lie within the $\pm$ 2σ nuclear recoil band, (2) the event's ionization yield must be less than the lower 3σ edge of the electron recoil band.

6.4.4 Fiducial volume cut

Events near the edge of detectors are rejected for three reasons: (1) degraded ionization signal due to the non-uniform or even distorted electric field, (2) worse phonon response due to the poor phonon sensor coverage, and (3) high probability of background interactions due to the absence of self-shielding.

Since we don't have a very accurate event location sensitivity, the fiducial volume cut is based on their signals energy on the outer ionization electrode (or "outer guard ring") to reject events that take place close to the edge of the crystal.

It is implemented in the following steps: (1) iteratively Gaussian fit to the Q_{outer}

distribution for good Ba Q_{inner} events in numerous bins of charge energy, commonly from 0-300 keV in Q_{inner}. (2) Line fit to the means and a quadratic to the widths of these Gaussians as functions of energy. (3) Based on these fits, the cut is set at $\pm$ 2σ away from the mean. This cut makes sure the outer electrode shows an ionization signal consistent with noise. figures 6.11 shows a representative plot in detector T1Z5 for such fits.

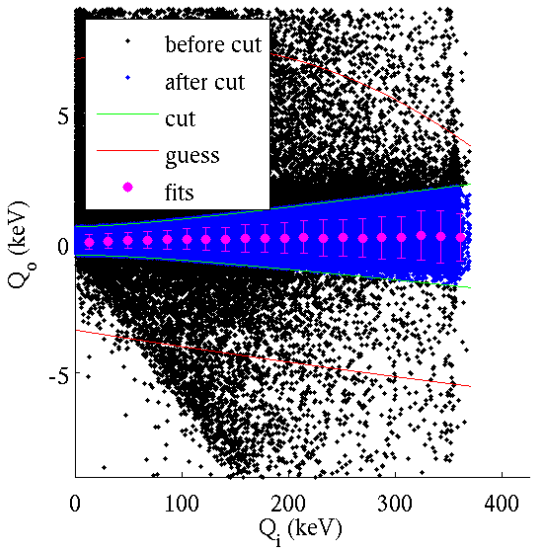


Figure 6.11: Q_{inner} vs. Q_{outer} for ^{133}Ba calibration events in detector T1Z5

The efficiency of fiducial volume cut may be under-estimated due to the neutron self-shielding effect. Because of the finite attenuation length of neutrons, detector regions outside the fiducial volume are more likely to interact with neutrons than inside. Therefore, the neutron interaction fraction inside the fiducial volume is smaller than the true fiducial volume, while WIMPs have uniform chances to interact throughout each region of detector.

Since the fiducial volume cut is defined from neutron calibration, the cut efficiency is effectively smaller for that should be for WIMPs. Monte Carlo simulation gives a

rough estimate of this effect to be about 5%.

6.4.5 Singles cut

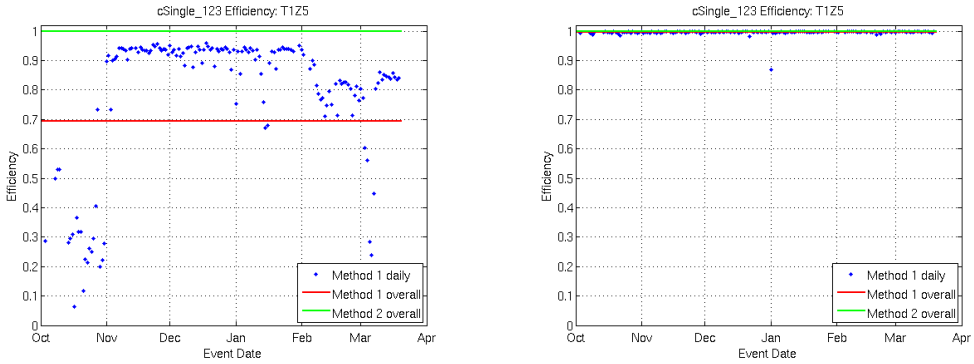
WIMPs are not expected to interact multiple times in different detectors due to their extremely small interaction cross sections. On the other hand, neutron and electromagnetic backgrounds may have multiple interactions. Therefore it is natural to require WIMP candidate events only scatter in one single detector. This requirement is implemented by defining a single scatter as having significant energy deposition greater than 6σ of the upper edge of one detector's phonon noise blob, but below 4σ of the upper edge of phonon noise blob in all other detectors.

These thresholds should give us an efficiency in excess of 99%, as it is rare that the noise in any detector will exceed the cut threshold which is coincidence with a true signal pulse in any other detectors. However, after the initial singles cut, it has been found, due to variations in phonon resolution and noise (Figure 6.13), the cut efficiencies had significant fluctuations with time and were generally much lower than expected (Figure 6.12(a)).

The final cSingle_123 cut takes into account the variations in phonon resolution, calculates values for the various phonon noise means and resolutions for each series independently, and sets the high and low thresholds on a run-by-run basis. Four detectors (T1Z1, T3Z1, T5Z5, T5Z6) with relatively poor phonon sensor performance are also set ionization signals 4σ thresholds, along with their phonon thresholds. The efficiency of the final version of singles cut is shown in Figure 6.12(b), with the efficiency above 99%.

6.4.6 Surface event rejection cut

Electron recoil events near the detector surface are the primary backgrounds for CDMS due to their reduced ionization yield, which results in misidentification as signal



(a) Initial singles cut efficiency as a function of series date

(b) Final singles cut efficiency as a function of series date

Figure 6.12: Singles cut efficiencies on detector T1Z5 as a function of series date

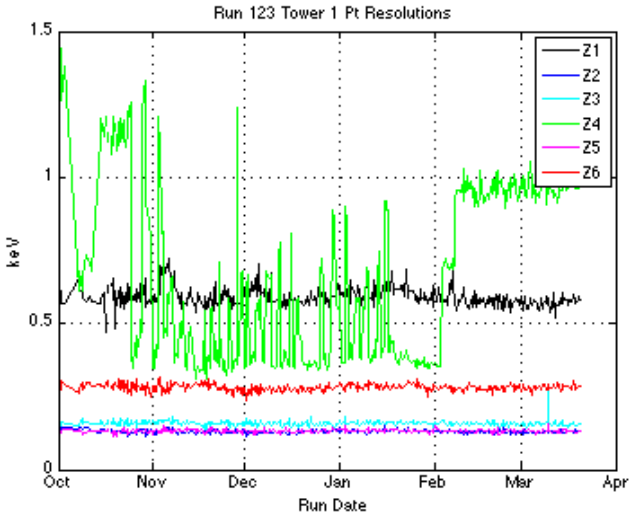


Figure 6.13: The resolution of the phonon energy noise blob for Tower 1 as a function of time during Run 123.

events if only ionization yield is used. These low yield events are subsequently rejected by the phonon pulse timing distributions. Phonon pulses from these surface recoils are faster than those from recoils in the detector bulk. Two timing parameters in the phonon quadrant with the largest phonon amplitude are particularly powerful: the time delay of the phonon signal (20% amplitude) relative to the fast ionization signal, and the phonon pulse rise time (10% to 40% amplitude).

The analysis of previous runs shows these two phonon timing parameters provides good rejection of surface electron recoils while retaining reasonable acceptance efficiency of nuclear recoil events [55]. WIMP nuclear recoil events are required to exceed a minimum value for this timing parameter (Figure 6.14). These cuts threshold were set based on leakages from a subset of ^{133}Ba calibrations, which demands around 0.3% leakage surface events on each individual detector from the Ba calibration. A consistency cut for signal candidates is also required to only accept events consistent with the nuclear recoil event distribution (i.e., within 4σ for the difference between delay and rise time from the neutron distribution mean). The simplicity and robustness of this classic 2D timing parameter analysis made it the primary analysis for Run 123/124. Other studies of advanced surface rejection techniques are further discussed in chapter 5.

The ionization yield discriminates nuclear recoils from electron recoils with a rejection factor of greater than 10^4. Phonon timing provides effective discrimination against surface electron recoil events, improving our overall rejection to more than 10^6. Figure 6.15 shows the discrimination power of the combination of these two. The signal region is defined by gamma events from ^{133}Ba and neutron events from ^{252}Cf calibration.

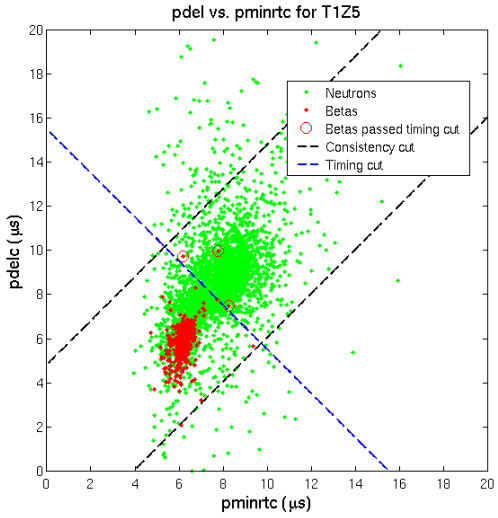


Figure 6.14: Simple timing cut based on phonon delay time and rise time. Blue dashed line is the cut minimum threshold. Surface beta events in red with leakage events circled, nuclear recoil neutron events in green.

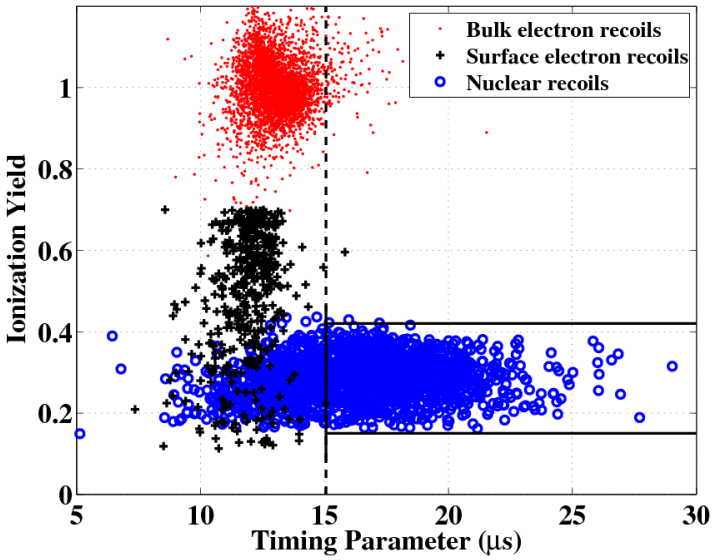


Figure 6.15: Ionization yield versus phonon rise time + delay time for calibration data in Ge detector T2Z3. Red dots are from ^{133}Ba gamma rays. Black crosses are low yield ^{133}Ba events. Blue circles are ^{252}Cf. The vertical dashed line indicates the minimum timing parameter allowed for candidate WIMP events in this detector, and the box shows the approximate signal region, which is in fact weakly energy dependent.

6.5 WIMP search results

6.5.1 WIMP search exposure

Only twenty-one detectors (15 Ge and 6 Si) were used for the WIMP search analysis during Run 123 and only nine detectors (7 Ge and 2 Si) during Run 124 were found to be sufficiently identical in behavior as to merit easy addition to the data.

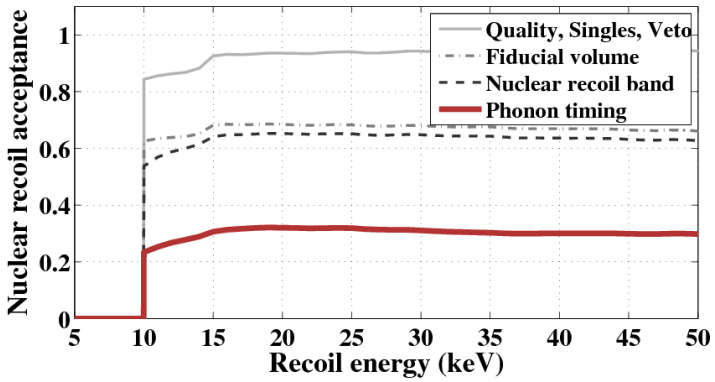
The total live time of the Run 123/124 WIMP search analysis for the data acquired between October 2006 and July 2007, after all data quality cuts and after multiplying by the detector masses was: 335.90 + 85.53 = 421.43 kg-d for Ge and 61.14 + 15.95 = 77.09 kg-d for Si (53.47 kg-d if we exclude two unmasked detectors T2Z1 and T4Z1 due to their unusual high expected leakage).

6.5.2 WIMP search efficiency

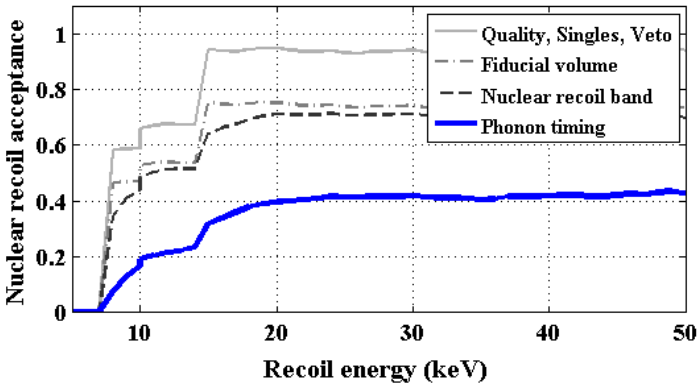
The overall nuclear recoil efficiency as a function of phonon recoil energy for Run123/124 is shown in Figure 6.16(a) for Ge and Figure 6.16(b) for Si. Four curves represent various stages during the analysis, with quality, singles and veto cuts applied for the top curve (reduce efficiency to $\sim$ 95%), fiducial volume cut for the second one (reduce to $\sim$ 70%) , nuclear recoil band cut for the third one (reduce to $\sim$ 65%), and the phonon timing cut for the bottom final efficiency. The net total efficiency is $\sim$ 30% for Ge detectors, and $\sim$ 40% for Si detectors above 20 keV phonon recoil energy.

6.5.3 Unblinding the WIMP search data

After all analysis cuts were finalized and leakage was estimated (in section 4.5), two stage unblinding strategy was followed. We first released the blinding cut and counted the total number of single events failing the timing cut. The counts are listed in table 6.2.



(a) Germanium



(b) Silicon

Figure 6.16: Nuclear-recoil acceptance efficiency of the Run 123/124 WIMP search analysis as a function of recoil energy.

Figure 6.17: Events in the unblinded Run 123/124 good Ge detectors failing the timing cut but passing all other cuts. Vertical magenta lines are phonon energy threshold, green lines are charge energy threshold, blue lines are the upper edge of yield for surface events. Nuclear recoil bands are in magenta.

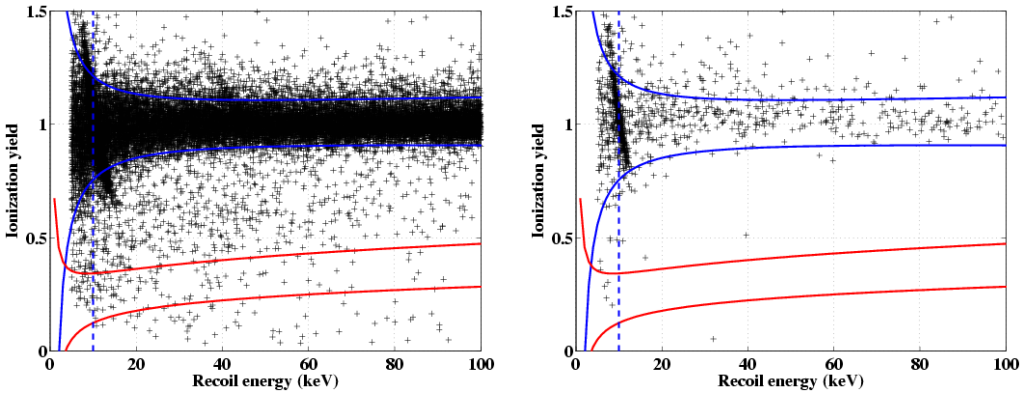
Table 6.2: Number of events in the unblinded Run 123/124 good Ge detectors failing the timing cut but passing all other cuts.

Detector	T1	T2	T3	T4	T5
Z1					30
Z2	9		1	4	
Z3		5			
Z4			4	3	3
Z5	8	2	0	5	4
Z6			6	13	

A total of 97 of such events were observed, which is consistent with expected 96 $\pm$ 15 events, calculated from the ratio of singles vs multiples for low-yield Ge events. These events are plotted for each detectors in Figure 6.17.

The final step of unblinding of the WIMP search signal region was completed on February 4, 2008. No candidate events were observed within the signal region. Figure 6.18 includes observed events from the sum of all WIMP search Ge detectors used in the analysis. The left figure shows the ionization yield distribution versus phonon recoil energy for singles events passing all cuts except the timing cut and yield cut. The right figure shows four events passing the timing cut which are outside the 2σ nuclear recoil region. The 2σ electron recoil bands (in blue), nuclear recoil bands (in red) and 10 keV energy threshold (blue dotted line) from T1Z2 are plotted to guide eyes. These curves are actually different from detector to detector.

Another illustrations which shows the power of our analysis is plotted in Figure 6.19. Again, no events are in our WIMP empty signal box region. The x-axis of the signal box is phonon timing rise time + delay time in μs from the cut position. The y-axis of the signal box is the ionization yield, in the unit of σ away from the center of nuclear recoil band. The signal region is between -2σ and +2σ. Note again, the signal box varies slightly between detectors due to the energy dependence of the nuclear recoil bands and the variation of timing cut positions.



(a) Before timing cut (b) After timing cut

Figure 6.18: Ionization yield versus recoil energy of all observed events in all good Ge detectors included in this analysis before (left) and after (right) the timing and yield cuts

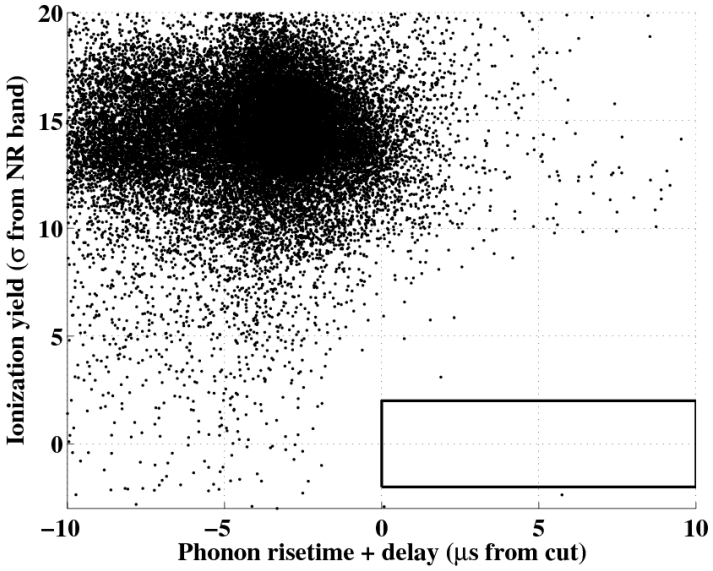


Figure 6.19: Empty box of WIMP signal region. The box is enclosed by -2σ and +2σ ionization yield from the nuclear recoil band mean, and the cut for the phonon timing parameters.

The Si detectors were not unblinded at the same time as the unblinding of Ge detectors due to the higher than expected surface background. They were unblinded later on December 3, 2008 with the same traditional 2D timing cut as used for Ge detectors, except that two detectors T2Z1 and T4Z1 due to their poor timing cut performance. No WIMP-candidate events were observed in silicon detectors either.

6.5.4 Limits on WIMP-nucleon interactions

Under the assumptions about the standard galactic halo (as described in [54]), no observation of candidate events in the Run 123/124 data sets the upper limit on the spin-independent WIMP-nucleon elastic-scattering cross-section. Calculated from Yellin's Optimal Interval Method [56], Figure 6.20 shows the upper limit of this analysis in solid red line. The minimum lies at 6.6×10^{44} cm^2 for a 60 GeV/c^2 WIMP at the Poisson 90% confidence level (C.L.). The reanalyzed Soudan data yield an improved sensitivity [4] than previous published Soudan results [46, 57, 58, 59], shown in dashed blue lines. When combined with all previous CDMS Soudan data, the upper limit in black solid line reaches 4.6×10^{44} cm^2, which is a factor of 3 lower than our previous published limit. Our previous Si result is also shown in green dashed line, along with the combined result is in diamond line.

The Soudan data represents the best WIMP sensitivity for spin-independent WIMP-nucleon interactions over a wide range of WIMP masses. Figure 6.21 compares the combined upper limits from Figure 6.20 to the results of other leading WIMP direct detection experiments. The upper dash dot curve is the Run 118/119 Soudan re-analysis result [4]. Our result has the same minimum cross-section as XENON10 result [60] (in orange dashed), but has better sensitivity at higher masses.

Our new limits exclude significant previously unexplored parameter space regions

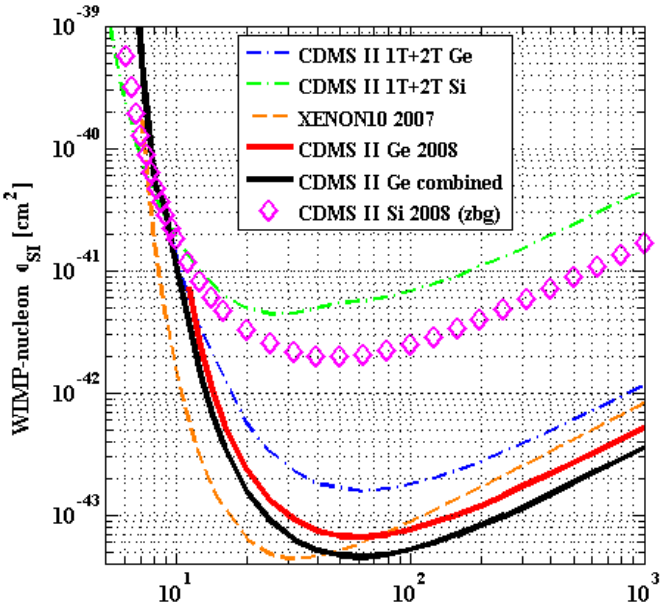


Figure 6.20: Spin-independent WIMP-nucleon cross-section upper limits (90versus WIMP mass for CDMS results.

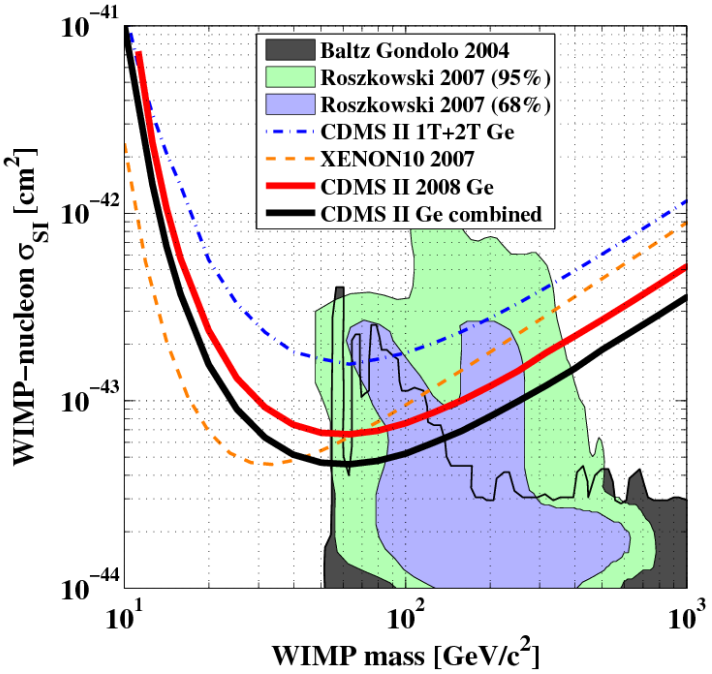


Figure 6.21: Spin-independent WIMP-nucleon cross-section upper limits (90% C.L.) versus WIMP mass for direct detection experiment and SUSY model.

predicted by supersymmetry models [61, 62]. Parameter ranges expected from supersymmetric models described by E. A. Baltz *et al.* are in grey, by L. Roszkowski *et al.* are shown (95% and 68% confidence levels in green and blue, respectively. With less than one expected event, and zero observed background in the signal region, CDMS maintains its high dark matter discovery potential.

References

[1] K.G. Begeman, A.H. Broeils, and R.H. Sanders. Extended rotation curves of spiral galaxies: dark haloes and modified dynamics. *Monthly Notices of the Royal Astronomical Society*, 249:523, 1991.

[2] Douglas Clowe *et al.* A direct empirical proof of the existence of dark matter. *The Astrophysical Journal Letters*, 648(2):L109–L113, 2006.

[3] E. F. Shores. Data updates for the SOURCES-4A computer code. *Nucl. Instrum. Meth., B*, 179(1):78–82, 2001.

[4] Reuben Walter Ogburn. *A search for particle dark matter using cryogenic germanium and silicon detectors in the one- and two-tower runs of CDMS-II at Soudan.* PhD thesis, Stanford University, 2008.

[5] Z. Ahmed *et al.* Search for weakly interacting massive particles with the first five-tower data from the cryogenic dark matter search at the soudan underground laboratory. *Phys. Rev. Lett.*, 102(1):011301, 2009.

[6] Gianfranco Bertone, Dan Hooper, and Joseph Silk. Particle dark matter: evidence, candidates and constraints. *Phys.Rept.*, 405:279–390, 2005.

[7] C. Amsler *et al.* (Particle Data Group). The Review of Particle Physics. *Physics Letters B*, 667:241, 2008.

[8] F. Zwicky. Spectral displacement of extra galactic nebulae. *Helvetica Physica Acta.*, 6:110–127, 1933.

[9] F. Zwicky. On the masses of nebulae and of clusters of nebulae. *Astrophysical Journal*, 86:217–246, 1937.

[10] P. Schneider, J. Ehlers, and E. E. Falco. *Gravitational Lenses*. Springer-Verlag, 1992.

[11] A. Refregier. Weak gravitational lensing by large-scale structure. *Annual Review of Astronomy and Astrophysics*, 41:645–668, 2003.

[12] W. Hu and S. Dodelson. Cosmic Microwave Background Anisotropies. *Annual Review of Astronomy and Astrophysics*, 40:171–216, 2002.

[13] M. R. Nolta *et al.* Five-year wilkinson microwave anisotropy probe observations: Angular power spectra. *The Astrophysical Journal Supplement Series*, 180(2):296–305, 2009.

[14] S. N. *et al.* Ahmed. Measurement of the total active $b8$ solar neutrino flux at the sudbury neutrino observatory with enhanced neutral current sensitivity. *Phys. Rev. Lett.*, 92(18):181301, May 2004.

[15] Y. *et al.* Ashie. Measurement of atmospheric neutrino oscillation parameters by super-kamiokande i. *Phys. Rev. D*, 71(11):112005, Jun 2005.

[16] R. D. Peccei and Helen R. Quinn. *cp* conservation in the presence of pseudoparticles. *Phys. Rev. Lett.*, 38(25):1440–1443, Jun 1977.

[17] P. Sikivie. Experimental tests of the "invisible" axion. *Phys. Rev. Lett.*, 51(16):1415–1417, Oct 1983.

[18] Z. Ahmed *et al.* Search for axions with the cdms experiment. *Physical Review Letters*, 103(14):141802, 2009.

[19] Kim Griest, Marc Kamionkowski, and Michael S. Turner. Supersymmetric dark matter above the w mass. *Phys. Rev. D*, 41(12):3565–3582, Jun 1990.

[20] D. *et al.* Abrams. Exclusion limits on the wimp-nucleon cross section from the cryogenic dark matter search. *Phys. Rev. D*, 66(12):122003, Dec 2002.

[21] T. Shutt *et al. Low Temperature Detectors*, page 615. American Institute of Physics, Melville, N.Y, 2002.

[22] T. Shutt, J. Emes, E. E. Haller, J. Hellmig, B. Sadoulet, D. Seitz, B. A. Young, and S. White. A solution to the dead-layer problem in ionization and phonon-based dark matter detectors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 444(1-2):340 – 344, 2000.

[23] Humphrey J. Maris. Phonon propagation with isotope scattering and spontaneous anharmonic decay. *Phys. Rev. B*, 41(14):9736–9743, May 1990.

[24] Humphrey J. Maris and Shin-ichiro Tamura. Anharmonic decay and the propagation of phonons in an isotopically pure crystal at low temperatures: Application to dark-matter detection. *Phys. Rev. B*, 47(2):727–739, Jan 1993.

[25] B. S. Neganov and V. N. Trofimov. Possibility of producing a bulky supersensitive thermal detector at a temperature close to absolute zero. *J. Exp. Theor. Phys.*, 28:328, 1978.

[26] P. N. Luke. Voltage-assisted calorimetric ionization detector. *Journal of Applied Physics*, 64(12):6858–6860, 1988.

[27] K.D. Irwin, S.W. Nam, B. Cabrera, B. Chugg, and B. Young. A quasiparticle-trap-assisted transition-edge sensor for phonon-mediated particle detection. *Review of Scientific Instruments*, 66:5322–5326, 1995.

[28] Shin-ichiro Tamura. Numerical evidence for the bottleneck frequency of quasidiffusive acoustic phonons. *Phys. Rev. B*, 56(21):13630–13633, Dec 1997.

[29] S. Tamura and H. J. Maris. Spontaneous decay of TA phonons. *Phys. Rev. B*, 31(4):2595–2598, Feb 1985.

[30] R. M. Clarke, P. L. Brink, B. Cabrera, P. Colling, M. B. Crisler, A. K. Davies, S. Eichblatt, R. J. Gaitskell, J. Hellmig, J. M. Martinis, S. W. Nam, T. Saab, and B. A. Young. Enhanced ballistic phonon production for surface events in cryogenic silicon detector. *Applied Physics Letters*, 76(20):2958–2960, 2000.

[31] D. S. Akerib *et al.* New results from the cryogenic dark matter search experiment. *Phys. Rev. D*, 68(8):082002, Oct 2003.

[32] R. J. Gaitskell, P. D. Barnes, A. DaSilva, B. Sadoulet, and T. Shutt. The statistics of background rejection in direct detection experiments for dark matter. *Nuclear Physics B (Proceedings Supplements)*, 51(2):279–283, 1996.

[33] S. Agostinelli *et al.* Geant4: A simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.

[34] J. Allison *et al.* Geant4 developments and applications. *IEEE Trans. Nucl. Sc.*, 53:270–278, 2006.

[35] A. Fasso *et al.* The physics models of fluka: Status and recent development. *arXiv:hep-ph/0306267*, 2003.

[36] A. Fasso, A. Ferrari, J. Ranft, and P. R. Sala. Fluka: A multi-particle transport code. *Nucl. Instrum. Meth.*, CERN-2005-10:INFN/TC 05/11, SLAC–R–773, 2005.

[37] T. Florkowski, L. Morawska, and K. Rozanski. Low-radioactivity background techniques. *Nucl. Geophys.*, 2(1):1–14, 1988.

[38] G. Heusser. Low-radioactivity background techniques. *Annual Review of Nuclear and Particle Science*, 45(1):543–590, 1995.

[39] T. Saab, R. M. Clarke, B. Cabrera, R. A. Abusaidi, and R. Gaitskell. Design of qet phonon sensors for the cdms zip detectors. *Nucl. Instrum. Meth. A*, 444:300, 2000.

[40] B. Cabrera. Design considerations for tes and qet sensors. *Nucl. Instrum. Meth. A*, 444:304, 2000.

[41] J. Hellmig, R. J. Gaitskell, R. Abusaidi, B. Cabrera, R. M. Clarke, J. Emes, S. W. Nam, T. Saab, B. Sadoulet, D. Seitz, and B. A. Young. The cdms ii z-sensitive ionization and phonon germanium detector. *Nucl. Instrum. Meth. A*, 444:308, 2000.

[42] J. S. Hendricks *et al.* Mcnpx 2.7.b extensions. Technical report. LA-UR-09-04150 available from http://mcnpx.lanl.gov/.

[43] D. S. Leonard *et al.* Systematic study of trace radioactive impurities in candidate construction materials for exo-200. *Nucl. Instrum. Meth. A*, 591:490–509, 2008.

[44] Lindhard *et al.* Integral Equations Governing Radiation Effects. *Mat. Phys. Medd. Dan. Vid. Sesk.*, 33(10), 1963.

[45] V. Mandic *et al.* Study of the dead layer in germanium for the CDMS detectors. *Nucl. Instrum. Meth., A*, 520(1-3):171–174, 2004.

[46] D. S. Akerib *et al.* First results from the cryogenic dark matter search in the soudan underground laboratory. *Phys. Rev. Lett.*, 93(21):211301, Nov 2004.

[47] I.T. Jollife. *Principal Component Analysis*. Springer, 2nd edition, 1986.

[48] Christopher M. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, College Station, Texas, 1995.

[49] MacKay D.J.C. *Bayesian methods for backpropagation networks*, chapter 6. Springer-Verlag, 1995.

[50] MacKay D.J.C. *Maximum Entropy and Bayesian Methods*. Springer, 2004.

[51] Howard Demuth, Mark Beale, and Martin Hagan. *Neural Network Toolbox 6 User's Guide*. The MathWorks, Inc., 2009.

[52] Ian Nabney. *NETLAB: algorithms for pattern recognition*. Springer, 2002.

[53] Ian Nabney. Software for netlab: algorithms for pattern recognition. Website. `http://www.ncrg.aston.ac.uk/netlab/down.php`.

[54] J. D. Lewin and P. F. Smith. Review of mathematics, numerical factors, and corrections for dark matter experiments based on elastic nuclear recoil. *Astroparticle Physics*, 6(1):87–112, 1996.

[55] Angela Reisetter. *Results from the Two-Tower Run of the Cryogenic Dark Matter Search*. PhD thesis, University of Minnesota, 2005.

[56] S. Yellin. Finding an upper limit in the presence of an unknown background. *Phys. Rev. D*, 66(3):032005, Aug 2002.

[57] 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 laboratory. *Phys. Rev. D*, 72(5):052009, Sep 2005.

[58] D. S. Akerib *et al.* CDMS Collaboration. Limits on spin-independent interactions of weakly interacting massive particles with nucleons from the two-tower run of the cryogenic dark matter search. *Physical Review Letters*, 96(1):011302, 2006.

[59] D. S. Akerib *et al.* CDMS Collaboration. Limits on spin-dependent wimp-nucleon interactions from the cryogenic dark matter search. *Physical Review D (Particles and Fields)*, 73(1):011102, 2006.

[60] J. Angle *et al.* XENON Collaboration. First results from the xenon10 dark matter experiment at the gran sasso national laboratory. *Physical Review Letters*, 100(2):021303, 2008.

[61] Edward A. Baltz and Paolo Gondolo. Markov chain monte carlo exploration of minimal supergravity with implications for dark matter. *Journal of High Energy Physics*, 2004(10):052, 2004.

[62] Leszek Roszkowski, Roberto Ruiz de Austri, and Roberto Trotta. Implications for the constrained mssm from a new prediction for b → sγ. *Journal of High Energy Physics*, 2007(07):075, 2007.