A Dark Matter Search Using the Final CDMS-II Data and 100 mm SuperCDMS Germanium Detector Ionization Test

A DISSERTATION

SUBMITTED TO THE FACULTY OF

UNIVERSITY OF MINNESOTA

BY

Jianjie Zhang

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Vuk Mandic

July 2014

Acknowledgements

This thesis would not have been possible without the support of many people. I would like to first thank my advisor, Vuk Mandic, for giving me the opportunity to work with him and start the new lab. His insights into research topics and breadth of physics knowledge set an example that I hope to follow. I especially appreciate the freedom to work independently he has given me with great trust and patience as well as constant encouragement.

I received numerous help from many people in setting up the new lab. Paul Crowell taught me how to operate the dilution fridge and generously lent us the electronics room of his group. Nader Mirabolfathi showed me the CDMS test facility at Berkeley and how to do various measurements such as the critical temperature and the I–V curve of the phonon sensors. Dennis Seitz was a hardware and electronics guru of CDMS. My first impressions of the CDMS hardware and electronics were from him, which were invaluable in designing the UMN test facility. I also learned from him how to build things the right way for them to survive the hash environment in a dilution fridge. Jon Kilgore, Peter Ness, Ronald Bystrom, and a few other members of the Physics Machine Shop provided extremely helpful guidance on my designs and carefully turned them into reality. They also taught me a great deal on machining

skills and helped tremendously moving things into and out of the lab. Sten Hansen provided multiple versions of the DCRC boards. Jeter Hall wrote the Labview interface for the DCRC's. Lauren Hsu and Bruno Serfass provided the CDMS data processing package `CDMSBats` and detailed instructions on how to set it up. The visits of the four of them to the lab significantly sped up the integration of the data chain. Thanks also go to Grace Elwell and Ashik Idrisy who led the efforts of building the big and the small cleanrooms for the lab, respectively.

I have greatly enjoyed and benefited from working with the team of the K100 fridge. Hassan Chagani is due special thanks for sharing and then taking over the responsibility of running the lab, without which the detector testing would have been much more stressful and I would not be able to focus on data analysis later on. Inspired by the atmosphere of the Physics Machine Shop and the CDMS group at Berkeley, I have been trying to make the lab an enjoyable workplace. We have been lucky to have Hassan in this regard, who is a master to keep a sense of humor in the lab. Thanks to Roxanne Radpour for the long hours of detector work in the cleanroom, installing the tower, innumerable cryogen transfers, designing and building the neutron shield, and also for organizing the softball games so I could walk out of the subbasement and enjoy the beautiful summer afternoons. Thanks also go to Allison Kennedy and Alex Codoreanu for their extensive detector and fridge work, and designing and constructing the gamma shield. I also thank everyone else who worked in the lab and made various contributions to the experiments.

My thanks also go to the team which made the ionization measurements of the 100 mm diameter ionization test devices possible. Paul Brink fabricated the detectors. Astrid Tomada did the detector wire bonding and packaging. Ben Shank, Jeff Yen,

Betty Young, and Matt Cherry screened the detectors in the Stanford ^{3}He fridge and handled the shipping of the detectors from California to Minnesota. Dennis Seitz proposed how the 100 mm detectors could be mounted on a CDMS-II 3-inch tower. Gary Godfrey designed the 100 mm detector housing, the SCHUBA extender, and the detector shipping container. The K100 fridge crews took care of the fridge work, data acquisition and processing. Hassan, Daniel Brandt, and I had many useful discussions about the data analysis. Vuk Mandic, Bernard Sadoulet, Blas Cabrera, Dan Bauer, Nader Mirabolfathi, Richard Partridge, and Eduardo do Couto e Silva coordinated the whole effort including the germanium crystal procurement. Various other people—Jeter Hall, Sten Hansen, and Rudy Resch to name a few—also made their unique contributions.

Vuk introduced me to CDMS data analysis and taught me how to use the CDMS data analysis package CAP. Cathy Bailey, Oleg Kamaev, Jodi Cooley, Lauren Hsu, and Bruno Serfass provided helpful advice to my early data analysis tasks. I also owe a debt of gratitude to the c58R analysis team. Lauren Hsu, Bruno Serfass, and Mark Kos prepared a special version of `CDMSBats` for the reprocessing of the c58 data and provided detailed guidance on submitting jobs to the FermiGrid. Thanks to Tom Hofer and Rito Basu Thakur for sharing the reprocessing workload. Thanks also to Anthony Villano, Tom Hofer, Joseph Kiveni, Dennile Speller, Mark Pepin, Angie Reisetter, and Prisca Cushman for sharing the c58R analysis tasks. Special thanks to Richard Schnee, Steve Yellin, and Dave Moore for providing helpful insights and guidance in the c58R analysis. The CDMS experiment which made the c58R analysis possible is the work of many collaborators. I thank them all for their efforts.

My officemates of the UMN CDMS group—Anthony Villano, Tom Hofer, Scott

Fallows, Mark Pepin, and a few others—have provided an enjoyable work atmosphere. The weekday lunch excursions have always been enjoyable and sources of interesting conversations. I also thank my Soudan shift partners with whom I shared the cold dark winter and lovely summer of northern Minnesota: Vuk Mandic, Wolfgang Rau, Kyle Sundqvist, Martin Huber, Anthony Villano, Lauren Hsu, Jeff Yen, Yohan Ricci, Todd Doughty, Brad Welliver, Ray Bunker, Gary Godfrey, Ben Shank, Kunj Prasad, Elias Lopez-Asamar, Dennile Speller, Bedile Kara, Silvia Scorza, and a few others. Special thanks to Dan Bauer, Jim Beaty, Dan DeVaney, Denise Osterholm, Joel Sander, and Angie Reisetter for providing guidance and advice during my Soudan shifts.

Finally, I thank my parents for giving me the freedom of choosing my own way, and my sisters for taking care of my parents when I was away from home.

Dedication

To my family

Abstract

Astrophysical observations indicate that approximately 85 % of the matter in the universe is nonluminous, nonbaryonic, and nonrelativistic (cold) dark matter. Weakly Interacting Massive Particles (WIMPs) are a particularly well motivated dark matter particle candidate. They would be thermally produced in the early universe and their relics account for the current dark matter abundance. WIMP candidate particles are naturally provided by extensions to the Standard Model of particle physics, such as supersymmetry. The Cryogenic Dark Matter Search (CDMS) experiment operates cryogenic germanium and silicon particle detectors in the low-background environment of the Soudan Underground Laboratory in northern Minnesota to search for WIMP-nucleus scatters while rejecting electron-recoil background. The detectors simultaneously measure the ionization and phonon energies of each scattering event. The difference in ionization yield (ratio of ionization energy to recoil energy) discriminates nuclear recoils from the electron-recoil background efficiently.

More sensitive detectors are required to probe the WIMP parameter space with lower WIMP-nucleon scattering cross sections. To support the R&D effort especially the detector R&D and characterization of the SuperCDMS experiment, a new CDMS test facility has been developed on the University of Minnesota campus. This thesis documents the test facility and the work involved in its development. In the test

facility, we performed the first ionization collection efficiency measurements of the ionization test devices. The test devices are fabricated with detector-grade germanium crystals that are 100 mm in diameter, which is the largest available, and 33 mm in thickness. The measured efficiencies are consistent with the earlier measurements conducted with smaller Ge crystals, demonstrating that these 100 mm crystals can be used for development of the next generation dark matter detectors.

Improvements of data analysis methods can also potentially improve the sensitivity of an experiment. The data taken during the last four runs of CDMS II with total raw exposure 612 kg · day were reprocessed with improved ionization pulse reconstruction algorithm. We present the classic timing analysis with the reprocessed data in this thesis. For the four runs combined, this analysis resulted in a new WIMP-nucleon cross section $4.4 \times 10^{-44}\text{cm}^2$ for a WIMP mass of 70 GeV, which is a factor of 1.6 improvement compared to the original c58 classic timing analysis.

Contents

List of Tables

List of Figures

Chapter 1

The Dark Matter Problem

It is well established that the majority of the matter content in the universe is in the form of cold dark matter, which, by definition, does not interact electromagnetically. The existence of dark matter can only be inferred by its gravitational attraction to the ordinary matter surrounding it or the photons passing by. While the astrophysical and cosmological observations supporting the existence of dark matter can hardly be interpreted otherwise, the particle nature of dark matter still reminds a mystery. Dark matter particles are massive, nonrelativistic, and interacting with a force at the level of the weak force or even weaker with ordinary matter as well as among themselves, which is almost all we know about them from the indirect evidence. None of the known Standard Model particles fits this portrait. A series of dark mater search campaigns, including direct searches, indirect searches, and collider searches, give conflicting information.

This chapter gives an introduction to the dark matter problem. The first few sections review the observational evidence supporting the existence of dark matter,

the standard halo model, and possible dark matter particle candidates. The rest of this chapter presents an overview of dark matter detection with the focus on direct searches.

1.1 The standard cosmology

It is convenient to discuss the dark matter problem in the framework of the standard cosmology. In the standard cosmology, the energy content (or mass by Einstein's mass-energy equivalence principle) of the universe consists of three forms based on their scaling properties during the expansion of the universe: radiation (photons), matter (including protons, neutrons, electrons, and the conjectured dark matter etc.), and dark energy (or the cosmological constant). The dynamics of the universe is governed by the Einstein field equation of the General relativity. We follow the sign conventions and part of the notations in [1] in introducing the framework.

The observed universe is isotropic and homogeneous at large scales. It is natural to assume the geometry of the universe is also isotropic and homogeneous. To account for our ignorance about the geometry of the universe in its largest scale, we introduce a forth fictitious spatial dimension w and imagine that our 3D universe is in the surface of a 4D sphere:

$$K^2 = x^2 + y^2 + z^2 + w^2. \tag{1.1}$$

The differential spatial separation in this 4D Euclidean space is

$$\mathrm{d}l^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 + \mathrm{d}w^2. \tag{1.2}$$

Using Eq. (1.1) to eleminate dw in Eq. (1.2), we have the spatial metric expressed in the 3D spherical coordinates:

$$\mathrm{d}l^2 = |K^2|\left[\frac{\mathrm{d}r^2}{1-kr^2} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2)\right], \tag{1.3}$$

where (r, θ, ϕ) are the comoving spherical coordinates, which are defined on the 3D unit sphere and do not scale with K; and

$$k \equiv \frac{|K^2|}{K^2} \in \{-1, 0, 1\} \tag{1.4}$$

is the normalized curvature. The values -1, 0, and $+1$ describe an open, flat, and closed universe, respectively.

With the time component added, the full Friedman-Robertson-Walker (FRW) metric is

$$\mathrm{d}s^2 = \mathrm{d}t^2 - R^2(t)\left[\frac{\mathrm{d}r^2}{1-kr^2} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2)\right] = g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu, \tag{1.5}$$

in which $R^2(t) \equiv |K^2|$ is the scale factor describing the expansion of the universe, the speed of light is $c = 1$, and the Minkowski metric is $\eta_{\mu\nu} = (+,-,-,-)$.

The dynamics of the geometry (or metric) is governed by the Einstein field equation

$$\boldsymbol{R}_{\mu\nu} - \frac{1}{2}g_{\mu\nu}\boldsymbol{R} - g_{\mu\nu}\Lambda = 8\pi G T_{\mu\nu}, \tag{1.6}$$

where $\boldsymbol{R}_{\mu\nu}$ and $\boldsymbol{R}$ are the Ricci tensor and scalar, respectively, $g_{\mu\nu}$ is the spacetime metric, Λ is the cosmological constant, G is the gravitational constant, and $T_{\mu\nu}$ is the

stress-energy tensor. With the FRW metric, the non-zero elements of the Ricci tensor are

$$R_{00} = -3\frac{\ddot{R}}{R}, \tag{1.7}$$

$$R_{ij} = -\left[\frac{\ddot{R}}{R} + 2\left(\frac{\dot{R}}{R}\right)^2 + \frac{2k}{R^2}\right] g_{ij}, \tag{1.8}$$

and the Ricci scalar is

$$R = -6\left[\frac{\ddot{R}}{R} + \left(\frac{\dot{R}}{R}\right)^2 + \frac{k}{R^2}\right]. \tag{1.9}$$

Assume the energy in the universe is in the form of perfect fluid, then the stress-energy tensor is

$$T_{\mu\nu} = (\rho + p)u_\mu u_\nu - p g_{\mu\nu}, \tag{1.10}$$

where ρ and p are the energy density and pressure of the fluid, respectively; and in the rest frame of the liquid $u^\mu = (1, 0, 0, 0)$.

The 00 and ii components of Eq. (1.6) are

$$\left(\frac{\dot{R}}{R}\right)^2 + \frac{k}{R^2} - \frac{\Lambda}{3} = \frac{8\pi G\rho}{3}, \tag{1.11}$$

$$2\frac{\ddot{R}}{R} + \left(\frac{\dot{R}}{R}\right)^2 + \frac{k}{R^2} - \Lambda = -8\pi G p. \tag{1.12}$$

Equation (1.11) is also referred to as the Friedmann equation. The energy density ρ in Eq. (1.11) can be in the form of radiation ρ_r and matter ρ_m. They scale differently as the universe expands:

$$\rho_r = \rho_{r,0}\left(\frac{R}{R_0}\right)^{-4}, \qquad \rho_m = \rho_{m,0}\left(\frac{R}{R_0}\right)^{-3}, \tag{1.13}$$

where the subscript 0 means the values are taken at the present time $t = t_0$.

Define the Hubble parameter

$$H \equiv \frac{\dot{R}}{R} \tag{1.14}$$

and the critical density

$$\rho_c \equiv \frac{3H^2}{8\pi G}. \tag{1.15}$$

Then Eq. (1.11) can be written as

$$\begin{aligned} \rho_c &= \rho_r + \rho_m + \rho_k + \rho_\Lambda \\ &= \rho_{r,0}\left(\frac{R}{R_0}\right)^{-4} + \rho_{m,0}\left(\frac{R}{R_0}\right)^{-3} + \rho_{k,0}\left(\frac{R}{R_0}\right)^{-2} + \rho_\Lambda, \end{aligned} \tag{1.16}$$

where

$$\rho_{k,0} = -\frac{3k}{8\pi G R_0^2}, \quad \rho_\Lambda = \frac{\Lambda}{8\pi G}. \tag{1.17}$$

With both sides divided by ρ_c, Eq. (1.16) becomes

$$1 = \Omega_r + \Omega_m + \Omega_k + \Omega_\Lambda, \tag{1.18}$$

where $\Omega_x = \rho_x/\rho_c$ is the density fraction of component x. We see that the energy content of the universe $\Omega_t = \Omega_r + \Omega_m + \Omega_\Lambda$ determines its curvature:

- $\Omega_t = 1 \Rightarrow k = 0$, flat,
- $\Omega_t < 1 \Rightarrow k < 0$, open,
- $\Omega_t > 1 \Rightarrow k > 0$, closed.

In a flat universe, Ω_t is always unity during the evolution of the universe. Because of the different scaling properties (see Eq. (1.16)) for each component, radiation

dominates in the early stage of the universe, then the matter, then the curvature if it exists, eventually the dark energy. Our very universe is found surprisingly flat with $100\,\Omega_k = -0.10^{+0.62}_{-0.65}$ (95 % CL) [2]. The dominant components are dark energy $\Omega_\Lambda = 0.685^{+0.018}_{-0.016}$ (68 % CL) and matter $\Omega_m = 0.315^{+0.016}_{-0.018}$ [2]. As part of the matter, the baryon has a fraction of $\Omega_b h^2 = 0.02205 \pm 0.00028$ (with $h = 0.673$) [2]; almost all the rest of the matter is dark matter. Today's universe is dark energy dominated, the second dominant component would be the curvature term by the above argument, which contradicts the observation. An epoch of inflation in the early universe was proposed to solve the flatness problem.

Equations (1.11) and (1.12) together determine the scale of the universe at any moment during its evolution. As the universe expands, the wavelength of the photons stretches accordingly, red shifting the photons. The redshift of observed photons reveals the scale, and thus the age, of the universe when they were emitted. The redshift z is defined by

$$1 + z = \frac{R_0}{R}. \tag{1.19}$$

It is a single-valued function of R and thus of time t.

The energy contents of the universe, including radiation and matter, went through a number of phases during the evolution of the universe and form everything we see today. Detailed discussions about this can be found in standard texts such as [1].

1.2 Evidence for dark matter

The existence of dark matter is supported by a wide variety of observational evidence ranging from the galaxy scale to the Cosmic Microwave Background (CMB) scale and

from the early universe to today. Dark matter is now an indispensable ingredient in describing the dynamics and evolution of the universe.

1.2.1 Dark matter in galaxies

The most intuitive argument for the existence of dark matter is from the rotation curves of spiral galaxies. Figure 1.1 shows the rotation curve for the spiral galaxy NGC 6503 [3], which plots the circular rotation speed as a function of the radius from the galactic center. At radii beyond the extent of the luminous component, Newtonian gravity predicts that the circular speed falls off as $1/\sqrt{r}$. However, the measured rotation speed stays roughly constant at large radii, which indicates that there is additional invisible matter with density $\rho(r) \sim 1/r^2$ to provide additional gravitational attraction. The flat rotation curves are observed on a large number of spiral galaxies including our Milky Way [3–5]. Observations of the HI gas beyond the luminous component reveal that the halo is nonrotating and spheroidal [6, 7]. In recent years, studies of the stellar kinematics of the dwarf galaxies indicate that they are the most dark matter dominated known galaxies [8–10]. The satellite dwarf galaxies of the Milky Way can be valuable sources of dark matter annihilation signals in indirect dark matter search. For elliptical galaxies, notably NGC 4636 [11–13], the observations of the X-ray emitting gas, stellar velocities, and the velocities of the nearby globular clusters show that the mass profile of the galaxy is consistent with the presence of a dark matter component with Navarro–Frenk–White (NFW) or cored profile.

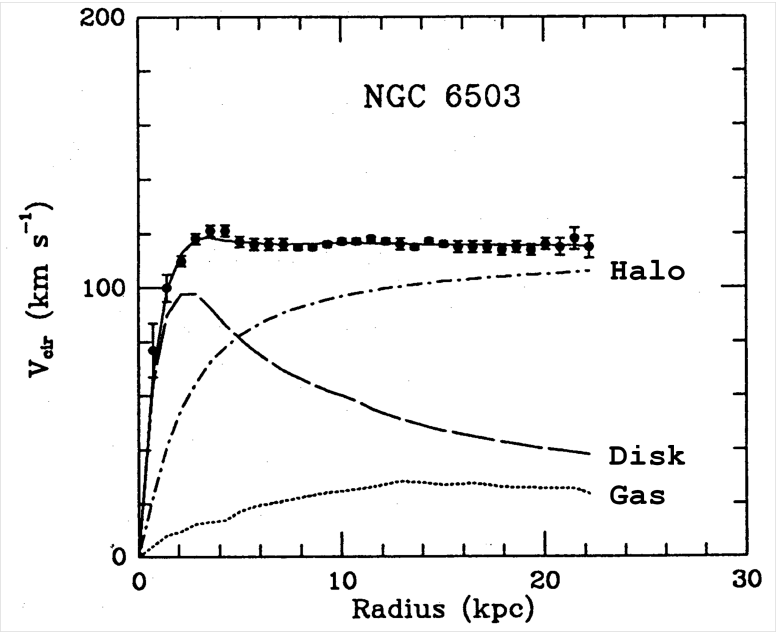


Figure 1.1: Rotation curve for the spiral galaxy NGC 6503. The gravitational attraction from the gas and galaxy disk can not explain the observed flat rotation curve in black squares. A dark matter halo is needed to make the theory compatible with the observation. *Figure from Begeman* et al. *[3].*

1.2.2 Dark matter in clusters of galaxies

A galaxy cluster consists of hundreds to thousands of galaxies. They are the largest gravitationally bound objects in the universe, large enough to represent a sample of the average composition of the universe. The masses of the galaxy clusters are commonly estimated in three ways: galactic motion, X-ray emission, and gravitational lensing. On average, stars make up $\sim$1 % of the cluster mass, hot gas $\sim$10 %, and the rest is dark matter [14–16].

Estimating the cluster mass using the galactic velocity dispersion is the earliest approach. In its most crude form, the cluster is assumed to be virialized within the virial radius $R_{\rm vir}$, and the dispersion of the velocity is $\langle v^2 \rangle = G\frac{M_{\rm vir}}{R_{\rm vir}}$, where G is the gravitational constant and $M_{\rm vir}$ is the virial mass. The virial radius $R_{\rm vir}$ can be determined by photometric measurements. In the early 1930s, Zwicky used this method and found that invisible matter is needed to explain the observed velocities of the galaxies in the Coma Cluster. Nowadays, when high resolution observations of the tracers are available, the Jeans analysis is used to infer the density distribution [10]. There are severe limitations to this method: (a) only the line-of-sight velocity can be measured, true velocity is inferred by deprojection, which may suffer from large systematic uncertainties; (b) the system is assumed at dynamical equilibrium which may not be true.

Since most of the baryonic matter of a cluster is in the form of hot X-ray emitting gas, X-ray emission observation provides a more direct way of estimating its mass [14]. In this approach, the gas density is inferred from the brightness of the X-ray, and the temperature of the gas measures the gravitational potential. The mass density can be obtained by combining the two. Figure 1.2 shows a composite image of the

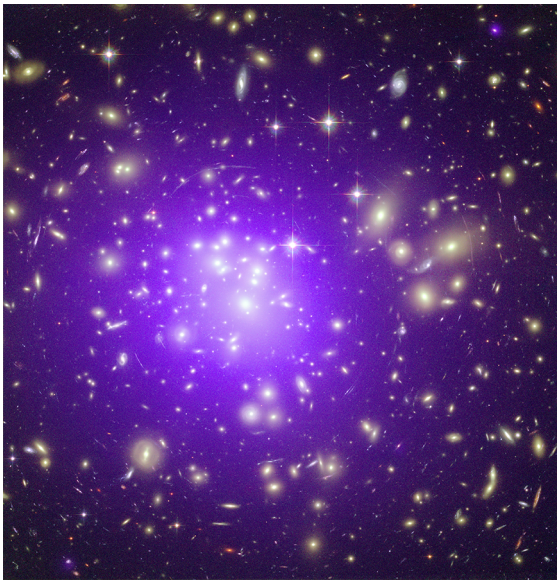
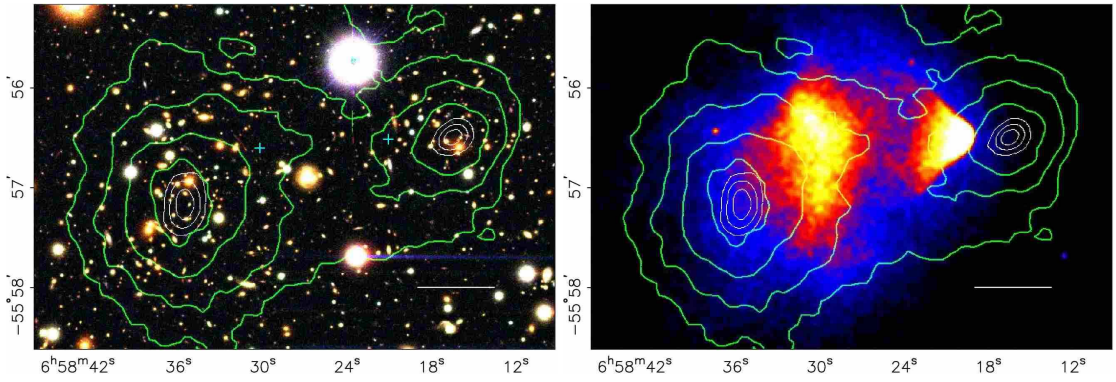


Figure 1.2: Composite image of Abell 1689. X-ray observation of the hot gas is colored purple and optical observation of the galaxies is colored yellow. Abell 1689 is a strong gravitational lens. The long arcs in the optical image are caused by gravitational lensing of background galaxies by matter in the galaxy cluster. *Figure from Chandra X-ray Observatory [17].*

cluster Abell 1689, where the X-ray observation of the hot gas is colored purple and the optical observation of the galaxies is colored yellow. The intracluster hot gas has an smooth appearance and fills the space between the galaxies. In the Chandra observations of the nearby relaxed galaxy clusters, the fraction of the hot gas in the total mass was found to be $\sim$4 to 11 % [14].

Gravitational lensing is the most reliable way to estimate the mass of a cluster. The lensing effect only depends on the mass distribution of the cluster, and is independent of the underlying dynamics or state of the cluster, i.e., whether it is in equilibrium or not. Much like a convex optical lens, lights passing through a cluster bend toward the mass center of the cluster because of the gravitational potential they feel. In strong lensing, as show in Fig. 1.2, multiple images and long arcs of the same galaxy



(a) (b)

Figure 1.3: The Bullet Cluster (1E 0657-558). (a): Optical image and the mass contours inferred from gravitational lensing. (b): Mass contours from gravitational lensing overlaid with X-ray image. The separation of the center of the total mass and that of the baryonic mass is clearly shown. The right subcluster is moving to the right as indicated by the shock front. *Figure from Clowe* et al. *[15]*.

behind the cluster can show up. And the lensing effects of each object have to be reconstructed separately. However, in weak lensing, the background galaxies are only distorted slightly. The distortions can only be detected statistically by comparing lensed and unlensed samples. The mass distribution of the cluster can be obtained by reconstructing the lensing effects. Figure 1.3(a) shows the lensing effects of the Bullet Cluster (1E 0657-558) and the derived mass contours.

All three approaches obtain exceedingly high mass-to-light ratio for the clusters, indirectly supporting the existence of dark matter. However, the combination of X-ray observation and gravitational lensing is even more powerful. The striking discovery of the Bullet Cluster, in which two subclusters collide and pass through each other, clearly shows the separation of the baryonic matter and the gravitational potential (Fig. 1.3(b)), providing strong support to the existence of the invisible dark matter and revealing its (nearly) collisionless nature. Alternatives such as modified gravity

would have difficulty to explain such separation.

1.2.3 Big Bang nucleosynthesis

As the universe expands, it cools off and the matter content evolves into the structures we see today. During the period of the Big Bang nucleosynthesis (BBN), which occurred ∼1–180 s after the Big Bang [18], heavier nuclei were synthesized from protons and neutrons as the temperature of the universe dropped below their binding energy. The abundances of the primordial elements such as helium and deuterium place tight constraint on the density of the baryonic matter in today's universe, which indicates the bulk of the matter in the universe is not baryonic.

At ∼1 s after the Big Bang, the temperature of universe dropped to $T \sim 1\,\mathrm{MeV}$. Before this moment, protons and neutrons were in thermal equilibrium with the reactions:

$$p + e^- \rightleftharpoons n + \nu,$$
$$n + e^+ \rightleftharpoons p + \bar{\nu},$$

and the ratio of the neutron to proton number is determined by the Boltzmann statistics: $n/p = e^{-Q/T}$, where $Q = 1.293\,\mathrm{MeV}$ is the neutron-proton mass difference. The above reactions stopped shortly after, as the expansion rate of the universe exceeded the neutron-proton inter-conversion rate and the neutron-proton ratio was frozen at $n/p \simeq 1/6$. The neutrons were then free to β-decay

$$n \to p + e^- + \bar{\nu}$$

and the neutron to proton ratio dropped to $\simeq 1/7$ by the time nuclear reactions began.

The nucleosynthesis started with the formation of deuteron (D) $\sim$100 s after the Big Bang when the temperature dropped to ~ 0.1 MeV,

$$p + n \rightleftharpoons D + \gamma.$$

The binding energy of deuteron is 2.23 MeV. The above reaction did not proceed substantially to the right until the temperature was much lower than the binding energy is because the photon density was billion times higher than that of the nucleon. The formation of deuteron preserves neutrons in the universe, otherwise all neutrons would decay into protons and the early universe would be left only with hydrogen.

Heavier nuclides ^{3}H, ^{3}He, and ^{4}He were then formed through deuteron in reactions [19]:

$$D + n \rightarrow {}^3\mathrm{H} + \gamma,$$
$$D + p \rightarrow {}^3\mathrm{He} + \gamma,$$
$${}^3\mathrm{H} + p \rightarrow {}^4\mathrm{He} + \gamma,$$
$${}^3\mathrm{He} + n \rightarrow {}^4\mathrm{He} + \gamma,$$

which involve photon emission, and reactions [19]:

$$D + D \rightarrow {}^3\mathrm{He} + n,$$

$$D + D \rightarrow {}^3\mathrm{H} + p,$$

$${}^3\mathrm{H} + D \rightarrow {}^4\mathrm{He} + n,$$

$${}^3\mathrm{He} + D \rightarrow {}^4\mathrm{He} + p,$$

$${}^3\mathrm{He} + n \rightarrow {}^3\mathrm{H} + p,$$

$${}^3\mathrm{He} + {}^3\mathrm{He} \rightarrow {}^4\mathrm{He} + 2p,$$

which usually go faster since they do not involve the relatively slow photon emission. Only two-body reactions are important in the BBN because the density has become rather low by this time [18]. Almost all neutrons ended up in ^{4}He at the end of the BBN. The combined reaction is effectively

$$D + D \rightarrow {}^4\mathrm{He} + \gamma.$$

This gives the mass fraction of ^{4}He $Y_p \simeq 0.25$. Trace amount of ^{7}Be and ^{7}Li were also produced in the BBN. The unstable nuclides ^{3}H and ^{7}Be decayed into ^{3}He and ^{7}Li, respectively, and did not survive. Because of the absence of stable nuclei with mass number 5 or 8, the BBN chain did not extended beyond ^{7}Be.

The whole BBN reaction chain is controlled by the baryon (nucleon) and the photon densities. Reaction rates and the densities of the nuclei as a function of time can be obtained by solving the system of equations for the reactions. The observations of the primordial abundances of the light elements place tight constraints

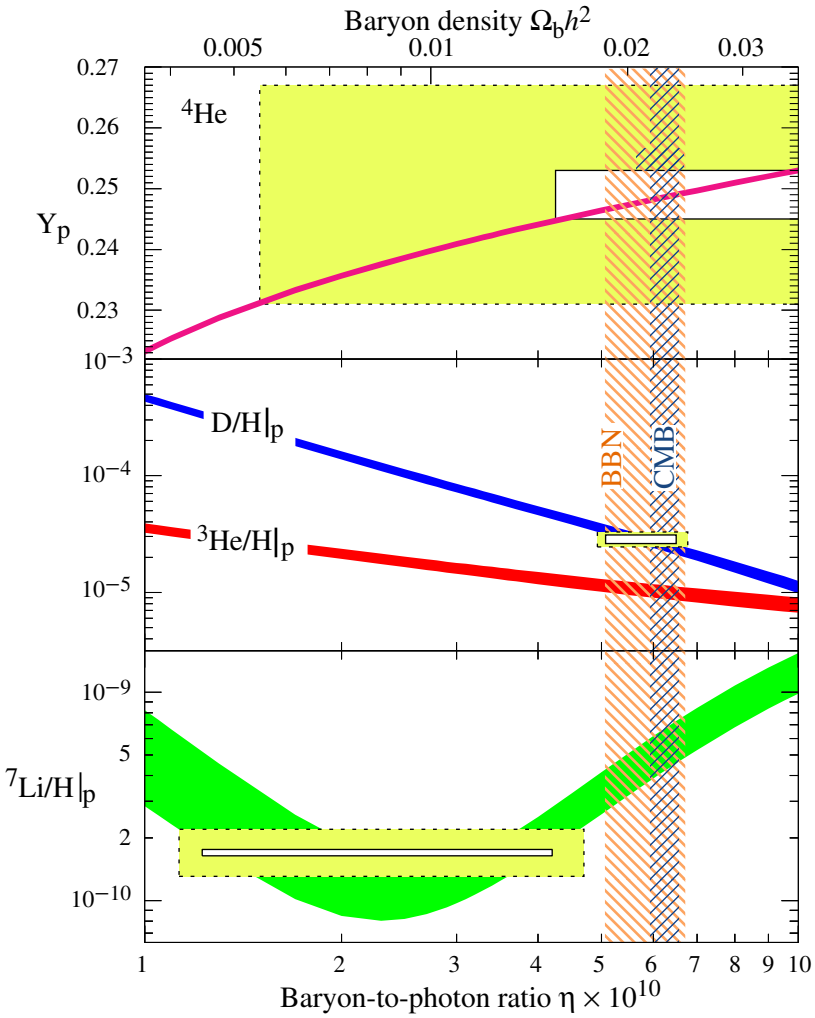


Figure 1.4: Primordial abundances of the light elements as a function of the baryon-to-photon ratio predicted by the Big Bang nucleosynthesis. The abundance of ^{4}He is in units of mass fraction and all the others are expressed in relative number to H. The bands show the 95 % CL range. Observed abundances and $\pm 2\sigma$ uncertainties are shown as horizontal boxes (statical errors are shown in small boxes). The vertical stripes are the CMB measurement and the BBN concordance ranges of the baryon-to-photon ratio (95 % CL). *Figure from Beringer* et al. *[18].*

on the the baryon density, which is usually expressed in the relic photon density as $\eta \equiv n_b/n_\gamma \equiv \eta_{10} \times 10^{-10}$. Figure 1.4 shows the baryon density constrained by the observed primordial abundances of the light elements. Deuteron gives the tightest constraint as it is believed there are no astrophysical sources of deuteron. The predicted abundance of ^{7}Li is a few times higher than the observation. The discrepancy is still unresolved [18]. ^{3}He can be easily destroyed and produced in stars and it is not suitable as a baryonmeter. The baryon density measured by the CMB observation is also plotted in Fig. 1.4. With the photon density measured by the CMB, the baryonic density is $\Omega_b h^2 = 0.0225 \pm 0.006$ (with $h = 0.704 \pm 0.025$) [18], which clearly indicates non-baryonic matter dominates in the universe.

1.2.4 The cosmic microwave background

During a long time after the BBN, nuclei and electrons were in the state of hot plasma. The universe was opaque to photons because of their short mean free path caused by frequent scattering. Roughly 380,000 years after the Big Bang, the universe became transparent to photons as it cooled sufficiently for the neutral hydrogen to form (formation of neutral helium happened slightly earlier). Decoupled from baryons, the photons free stream in the universe, becoming the nowadays observed cosmic microwave background (CMB) with red shift $z \sim 1100$. The spectrum of the CMB is well described by the blackbody radiation with temperature $T = 2.7255(6)\,\mathrm{K}$ [20]. The temperature of the CMB is highly isotropic over the entire sky with anisotropy at the level of 10^{-5}. Measurements of the anisotropy provide us rich information about the universe, precisely constraining the densities of the baryonic matter and dark matter.

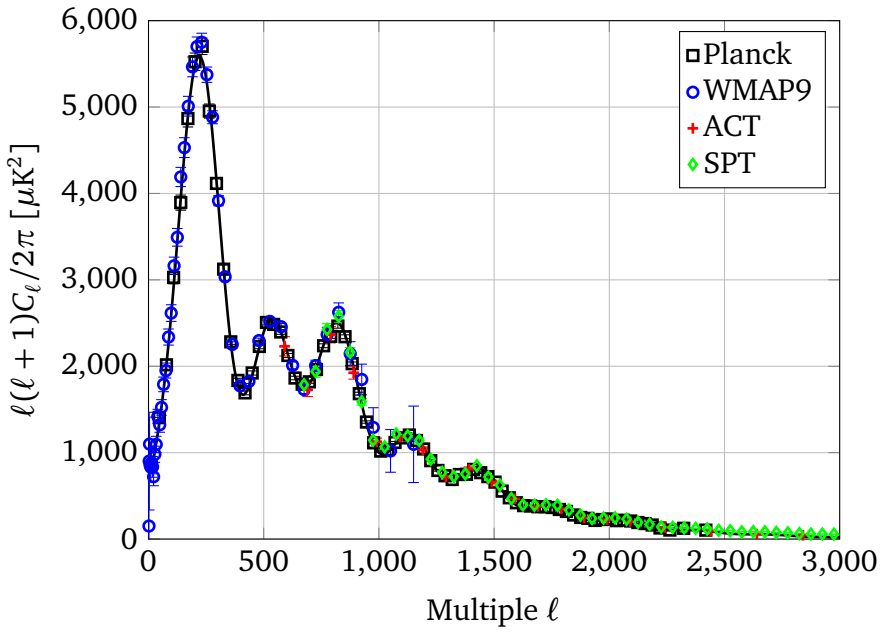


Figure 1.5: Angular power spectrum (TT) of the CMB temperature anisotropies measured by Planck, WMAP, ACT, and SPT. The data points are plotted with data from [21]. Only high ℓ data from Plank are plotted. The solid line is calculated using `CAMB` with parameters from [2].

The CMB temperature $T(\theta, \phi)$ can be expressed in spherical harmonics:

$$T(\theta, \phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\theta, \phi). \tag{1.20}$$

The angular power spectrum is then

$$C_\ell = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2. \tag{1.21}$$

However, the temperature anisotropy of the CMB is usually expressed in "band power" $\ell(\ell+1)C_\ell/2\pi$ as shown in Fig. 1.5. The peaks in Fig. 1.5 reflect the baryon acoustic oscillation before the photon decoupling. Prior to the era of recombination, photons and baryons were tightly coupled. The over density regions tended to accrete more baryons with gravitational attraction and become denser, however, the infall of baryons were impeded by the increased radiation pressure from the photons. The two competing effects caused the baryon acoustic oscillation at all scales that were causally connected. The collisionless cold dark matter, which does not feel the radiation pressure, started gravitational collapse much earlier, providing gravitational wells for the infall of the baryons. The system can be viewed as a harmonic oscillator in an extremely simplified picture, in which the baryon, radiation, and dark matter densities represent the mass, feedback strength, and zero-point of the oscillator, respectively. Indeed, this oscillation was observed and the measurements of the angular power spectrum have provided the most precise constraints on the baryonic and dark matter densities (Fig. 1.6).

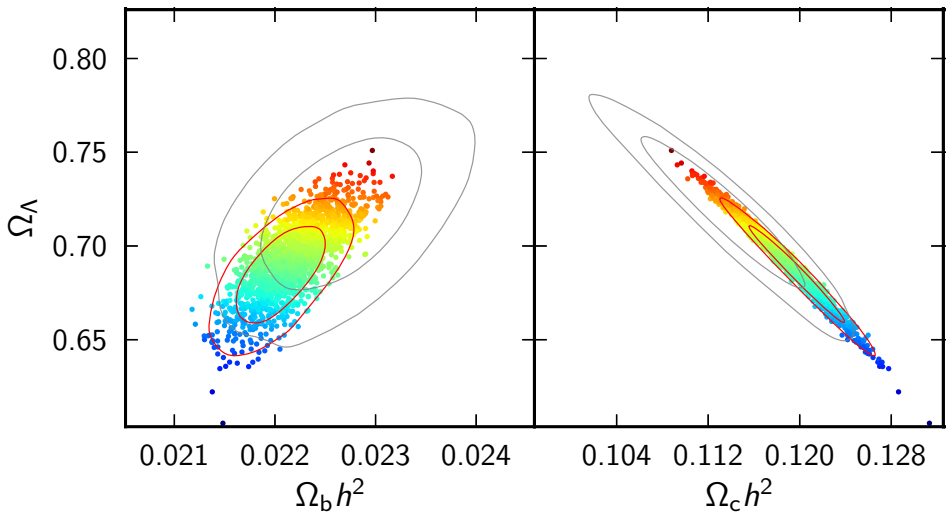


Figure 1.6: Likelihood contours (68 % and 95 %) of the densities of dark energy, baryonic matter, and cold dark matter constrained by WMAP-9 data (gray) and Planck+WMAP low ℓ polarization (red). *Figure from Ade* et al. *[2]*.

1.2.5 Large scale structure formation

The structures we see today such as galaxies and galaxy clusters are believed to have grown from the over densities in the early universe. However, the observed mass distribution fluctuation today is orders of magnitude larger than can be explained by the inhomogeneity of the baryons at the era of recombination inferred from the CMB anisotropy. Dark matter is believed to provide the additional inhomogeneities needed for the structure formation.

In the radiation dominated era, relativistic particles did not collapse under gravity and form over densities. Structure formation could only start until matter radiation equality, after which dark matter began to collapse under gravity with the density fluctuations seeded by the inflation. The collapse of the baryonic matter started after recombination when their gravitational collapse was not countered by the photons anymore. The gravitational potential wells formed by the dark matter accelerated

the collapse of the baryonic matter, giving it enough time to form the structures we see today. It is apparent the dark matter must be cold (non-relativistic) by the era of matter domination, otherwise it would not undergo gravitational collapse.

The density inhomogeneities observed today are evolved from the density fluctuations imprinted by the inflation, the power spectrum of which has a rather simple form $P(k) \propto k$, where k is the wave vector. The growth of the overdensities can be rather nonlinear at small spatial scales (large k), and detailed studies usually involve N-body simulations. However, at large scales, the growth is in the linear regime and can be calculated perturbatively. Measurements of the density power spectrum provide constraints on the dark matter and baryonic matter densities. Figure 1.7 shows the galaxy power spectrum measured by the 2dF Galaxy Redshift Survey (2dFGRS) [22], which gives the ratio of baryonic matter to dark matter $\Omega_b/\Omega_m = 0.17$. The SDSS experiment reported similar results [23].

1.3 The dark halo

The dark halo provides the gravitational potential which gives a flat rotation curve for spiral galaxies. Observations of the atomic hydrogen gas beyond the luminous galactic disk suggest that the dark halo is spherical and non-rotating [6, 7]. While there are sophisticated models to describe the dark matter distribution in the halo [24, 25], the isothermal sphere model is commonly adopted by the dark matter community for its simplicity.

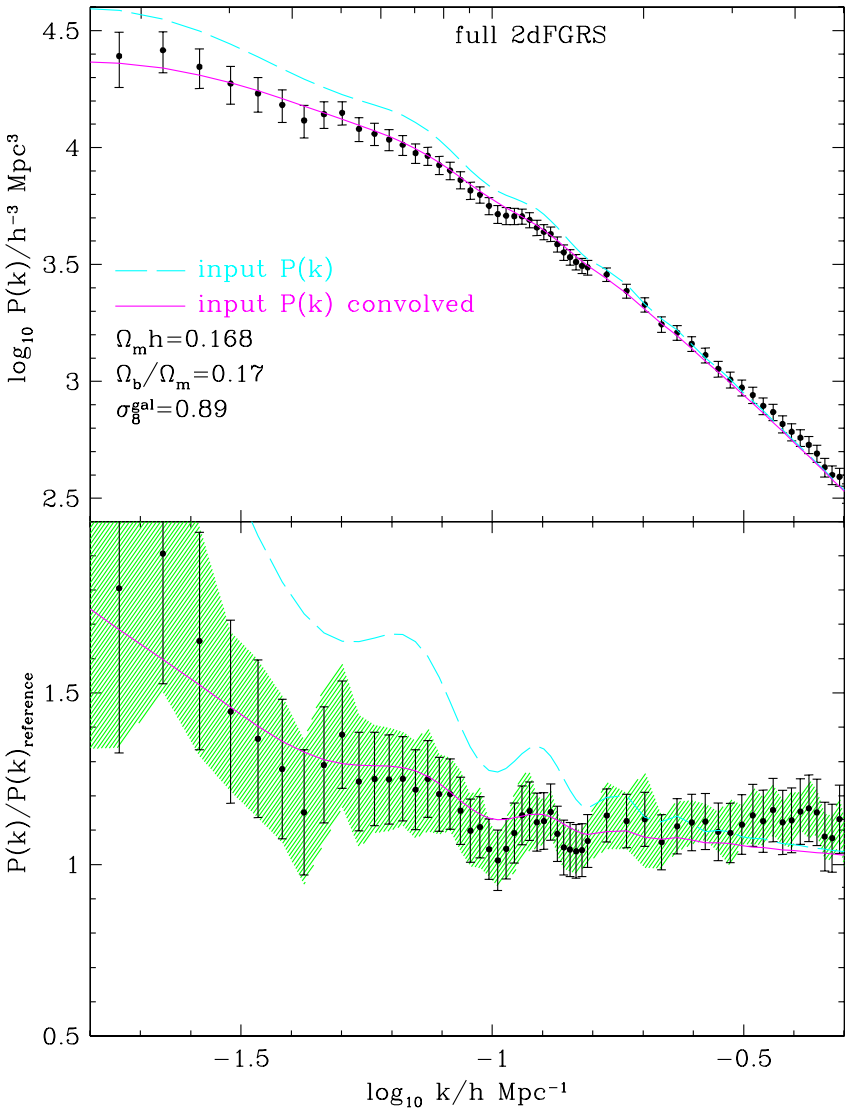


Figure 1.7: Galaxy power spectrum measured by the 2DFGRS survey. (Top): The black dots are the data, the dashed line is the theoretical model, and the solid line is the theoretical model convolved with the window function. (Bottom): Power spectrum normalized by an unrealistic reference model. *Figure from Cole* et al. *[22].*

1.3.1 Isothermal sphere model

The equation of state of the isothermal gas is $p = K\rho = \frac{k_B T}{m}\rho$. When the halo is in hydrostatic equilibrium, we have

$$\frac{\mathrm{d}p}{\mathrm{d}r} = \frac{k_B T}{m}\frac{\mathrm{d}\rho}{\mathrm{d}r} = -\rho\frac{\mathrm{d}\Phi}{\mathrm{d}r} = -\rho\frac{GM(r)}{r^2}, \tag{1.22}$$

where k_B is Boltzmann's constant, G is the gravitational constant, p and T are the pressure and temperature of the gas, respectively, m is mass of the dark matter particle, Φ is the galactic potential, and $M(r)$ is the halo mass inside radius r. Multiplying both sides of Eq. (1.22) by $r^2 m/\rho k_B T$ and differentiating with respect to r, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}\ln\rho}{\mathrm{d}r}\right) = -\frac{4\pi G}{\sigma^2}r^2\rho, \tag{1.23}$$

where the relationship $\mathrm{d}M/\mathrm{d}r = 4\pi r^2\rho$ has been used and σ is defined by

$$\sigma^2 \equiv \frac{k_B T}{m}. \tag{1.24}$$

The solution of Eq. (1.23) is

$$\rho = \frac{\sigma^2}{2\pi G r^2}. \tag{1.25}$$

Then the enclosed mass and the circular rotation speed at radius r are

$$M(r) = \frac{2\sigma^2 r}{G} \text{ and } v_c(r) = \sqrt{2}\sigma. \tag{1.26}$$

The Poisson's equation for the halo is

$$\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}\Phi}{\mathrm{d}r}\right) = 4\pi G\rho. \tag{1.27}$$

Solving Eq. (1.27), we obtain the galactic potential

$$\Phi = 2\sigma^2 \ln r + c, \tag{1.28}$$

where c is an arbitrary constant.

The corresponding velocity distribution with respect to the galactic potential given by Eq. (1.28) is

$$f(\mathrm{v}, \mathbf{r}) = \frac{1}{4\pi R\mathbf{r}^2}\frac{1}{(2\pi\sigma^2)^{3/2}}\exp\left(-\frac{\mathbf{v}^2}{2\sigma^2}\right), \tag{1.29}$$

where the normalization is within the sphere $r \leq R$. The distribution is pure Maxwellian at a given location.

The halo model given above has infinite density at $r = 0$. Since we are more interested in its behavior from the galactic center outward, we will regularize Eq. (1.23) using the conventions in [26], so that the model is well behaved at the center. Define

$$\tilde{\rho} \equiv \frac{\rho}{\rho(0)} \text{ and } \tilde{r} \equiv \frac{r}{r_0}, \tag{1.30}$$

where

$$r_0 \equiv \sqrt{\frac{9\sigma^2}{4\pi G\rho(0)}}. \tag{1.31}$$

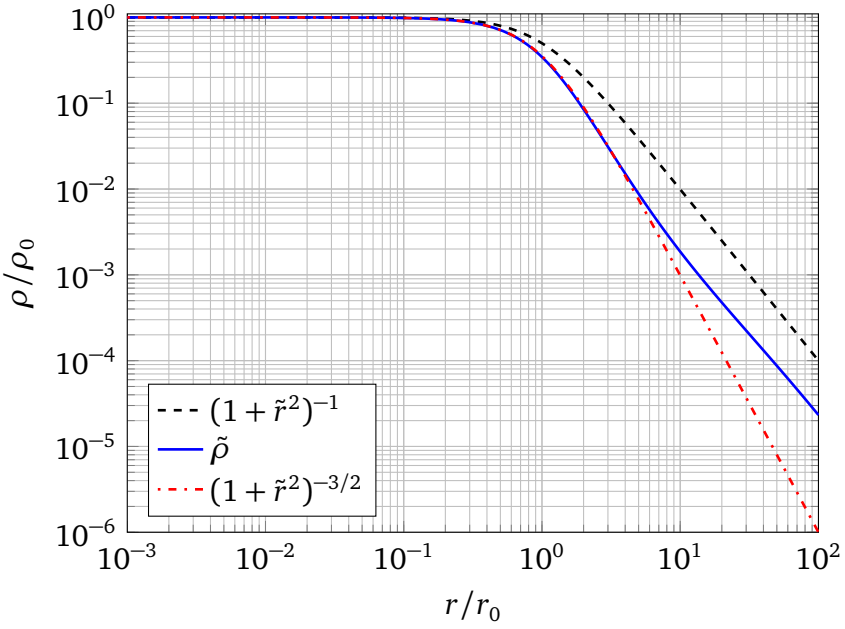


Figure 1.8: Density profile of the isothermal sphere dark halo. Two approximations, the cored isothermal sphere $(1+\tilde{r}^2)^{-1}$ and the modified Hubble model $(1+\tilde{r}^2)^{-3/2}$, are also plotted for comparison.

Then Eq. (1.23) can be written as

$$\frac{\mathrm{d}}{\mathrm{d}\tilde{r}}\left(\tilde{r}^2\frac{\mathrm{d}\ln\tilde{\rho}}{\mathrm{d}\tilde{r}}\right)=-9\tilde{r}^2\tilde{\rho}. \tag{1.32}$$

We numerically integrate Eq. (1.32) with the initial conditions $\tilde{\rho}|_{\tilde{r}=0}=1$ and $\frac{\mathrm{d}\tilde{\rho}}{\mathrm{d}\tilde{r}}|_{\tilde{r}=0}=0$. The resulting $\tilde{\rho}(\tilde{r})$ is plotted in Fig. 1.8. Two approximations of the density profile, the cored isothermal sphere $(1+\tilde{r}^2)^{-1}$ and the modified Hubble model $(1+\tilde{r}^2)^{-3/2}$ are also plotted. The modified Hubble model agrees with $\tilde{\rho}$ in a larger range of radii than the cored isothermal sphere, but its asymptotic slope at large radii is -3 instead of -2 of the isothermal sphere which is the correct trend.

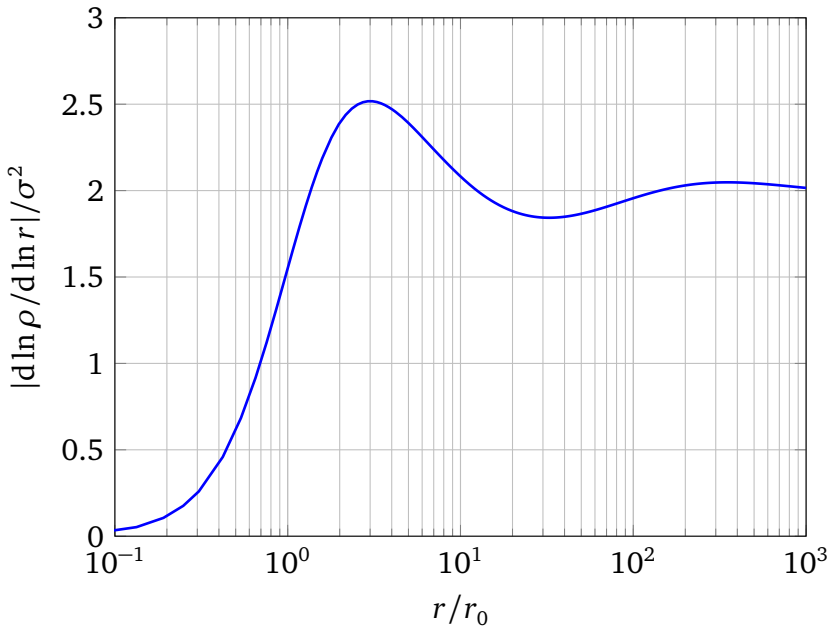


Figure 1.9: Rotation speed squared for the isothermal sphere. The rotation speed oscillates at large radii and approaches $\sqrt{2}\sigma$.

By integrating Eq. (1.22), we obtain the square of the circular rotation speed

$$v_c^2 = \frac{GM(r)}{r} = -\sigma^2 \frac{\mathrm{d}\ln\rho}{\mathrm{d}\ln r} = -\sigma^2 \frac{r\rho'(r)}{\rho}. \tag{1.33}$$

The numerical solution of v_c^2 is plotted in Fig. 1.9. The circular rotation speed oscillates and approaches $\sqrt{2}\sigma$ at large radii. The oscillation is caused by the deviation of $\tilde{\rho}$ from $\tilde{r}^{-2}$. The isothermal sphere model gives the correct overall trend of the rotation curves for the spiral galaxies shown in Sec. 1.2. The deviations are due to not including the contributions from the galactic disk and the possible complex structures of the halo itself, which is largely unknown at the moment.

Nice as the isothermal sphere model is, however, it is unphysical. It is unbounded spatially and in mass. Its mass increases $\sim r$ at large radii. Nothing can escape from

this infinite potential well. If we impose a finite radius R to the sphere and let it be vacuum outside the sphere, then the escape speed at the edge of the sphere is

$$v_e^2(R) = \frac{2GM(R)}{R} = 4\sigma^2 = 2v_c^2. \tag{1.34}$$

At radius r inside the sphere, the local escape velocity satisfies

$$\frac{1}{2}v_e^2(r) - \frac{1}{2}v_e^2(R) = \Phi(R) - \Phi(r). \tag{1.35}$$

Using Eqs. (1.28) and (1.34), we can write Eq. (1.35) as

$$v_e^2(r) = 2v_c^2\left(\ln\frac{R}{r} + 1\right). \tag{1.36}$$

At large radii ($r \leq R$), the escape speed changes slowly with respect to the changes in r.

1.3.2 The standard halo model

The standard halo model (SHM) [27–29] is actually the above isothermal sphere model with slight modifications to make it "more physical". The cuspy halo model predicted by simulations is disfavored due to lack of observational support. Since we only conduct the direct dark matter search experiments on our own planet Earth, by Eq. (1.29), the velocity distribution is Maxwellian. For the spatially finite Milky Way, however, particles with high enough speed will escape the galaxy. And thus we

impose a cutoff velocity, the local escape speed $v_{\rm esc}$, to the velocity distribution:

$$f(\mathbf{v}) = \begin{cases} \frac{1}{N_{\rm esc}} \frac{1}{(\pi v_0^2)^{3/2}} \exp\left(-\frac{\mathbf{v}^2}{v_0^2}\right) & |\mathbf{v}| \leq v_{\rm esc}, \\ 0 & |\mathbf{v}| > v_{\rm esc}, \end{cases} \tag{1.37}$$

where v_0 is the characteristic velocity:

$$v_0 \equiv \sqrt{2}\sigma = v_c(\infty), \tag{1.38}$$

and the normalization factor $N_{\rm esc}$ is given by

$$N_{\rm esc} = \mathrm{erf}\left(\frac{v_{\rm esc}}{v_0}\right) - \frac{2}{\sqrt{\pi}} \frac{v_{\rm esc}}{v_0} e^{-v_{esc}^2/v_0^2}. \tag{1.39}$$

The characteristic speed v_0 is the circular rotation speed at radii far beyond the extent of the luminous Galactic disk, where the gravitational force are mainly provided by the dark halo. Fortunately, the Sun is in the flat region of the Galactic rotation curve, and v_0 can be chosen as the circular speed of the local standard of rest around the Sun, which largely avoid the difficulty of measuring the rotation speed of distant objects in the Galaxy. We use the canonical value $v_0 = 220\,\mathrm{km\,s^{-1}}$ [7] in our analysis unless otherwise noted. The dispersion of the speed distribution without the $v_{\rm esc}$ cutoff is $\sqrt{\langle v^2 \rangle} \approx 270\,\mathrm{km\,s^{-1}}$.

A lower limit of the escape speed $v_{\rm esc}$ can be established by measuring the upper limit of the stellar speeds. We use $v_{\rm esc} = 544\,\mathrm{km\,s^{-1}}$ in our analysis, which is the central value of the 90 % CL interval, $498 < v_{\rm esc} < 608\,\mathrm{km\,s^{-1}}$, given by the RAVE survey [30].

1.3.3 Local dark matter density

It is important to know the dark matter abundance in the vicinity of the Earth in planning the direct dark matter search experiments. We give an estimate of the local dark matter density in this section following [7].

The circular speed of an object in the Galaxy can be decomposed into two components:

$$v_c^2(r) = v_d^2(r) + v_h^2(r), \tag{1.40}$$

where v_d and v_h are the contributions from the Galactic disk and the dark halo, respectively.

1.3.3.1 Halo contribution

For simplicity, we assume the halo is a cored isothermal sphere,

$$\rho(r) = \rho_0 \frac{a^2 + r_0^2}{a^2 + r^2}, \tag{1.41}$$

where $r_0 = 8.5\,\mathrm{kpc}$, which is the radius of the Sun from the Galaxy center. The halo contribution to the rotation speed:

$$v_h^2(r) = 4\pi G \rho_0 (r_0^2 + a^2)\left(1 - \frac{a}{r}\tan^{-1}\frac{r}{a}\right). \tag{1.42}$$

When $r \to \infty$, the contribution from the dark halo dominates:

$$v_\infty^2 = v_c^2(\infty) = v_h^2(\infty) = 4\pi G \rho_0 (r_0^2 + a^2). \tag{1.43}$$

Then the local halo density can be written as

$$\begin{aligned}\rho_0 &= \frac{v_c^2(\infty)}{4\pi G(r_0^2+a^2)} \\ &= 0.47\,\mathrm{GeV\,cm^{-3}}\frac{(v_\infty/220\,\mathrm{km\,s^{-1}})^2}{(r_0/8.5\,\mathrm{kpc})^2\,[1+(a/r_0)^2]}.\end{aligned} \tag{1.44}$$

The core radius a can be determined from the rotation speed of the Sun:

$$\frac{v_h^2(r_0)}{v_\infty^2} = 1 - \frac{a}{r_0}\tan^{-1}\frac{r_0}{a}. \tag{1.45}$$

1.3.3.2 Galactic disk contribution

Assume the surface density of the Galactic disk has an exponential form:

$$\Sigma(r) = \Sigma_0 \exp\left(-\frac{r-r_0}{H}\right), \tag{1.46}$$

where $H \approx 3.5\,\mathrm{kpc}$ and Σ_0 is the local surface density at the radius of the Sun. The circular speed at radius r in the Galactic disk is [26]

$$\begin{aligned}v_d^2(r) &= 4\pi G\Sigma_0 e^{r_0/H} H y^2\,[I_0(y)K_0(y) - I_1(y)K_1(y)] \\ &\approx (143\,\mathrm{km\,s^{-1}})^2 \frac{\Sigma_0}{50\,M_\odot\,\mathrm{pc^{-2}}},\ \text{for } r = r_0,\end{aligned} \tag{1.47}$$

where $y \equiv r/(2H)$, $M_\odot$ is the solar mass, and I's and K's are the modified Bessel functions.

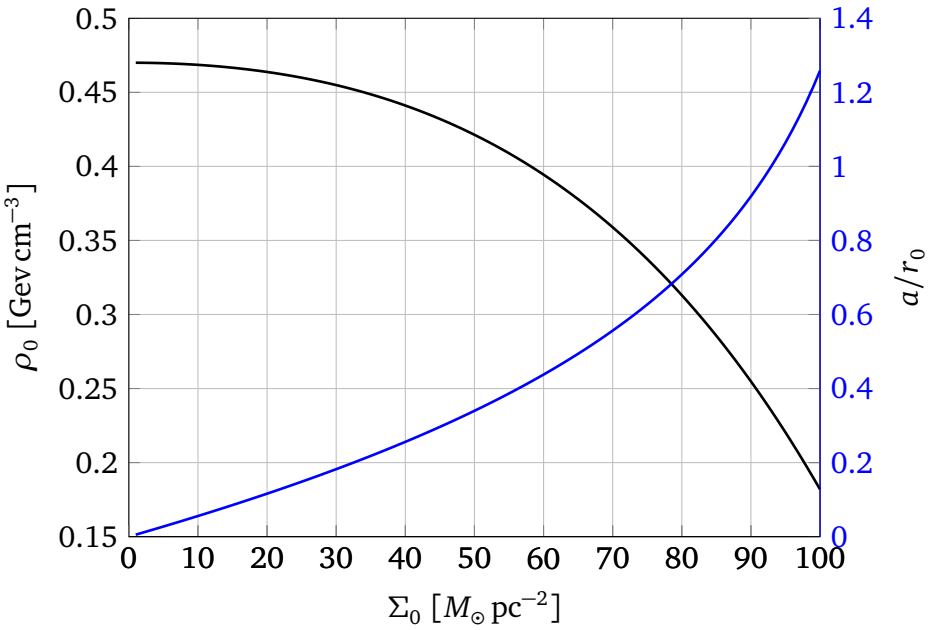


Figure 1.10: Local halo density of the Galaxy (left) and the radius of the dark halo core (right) vs. the local Galactic disk density.

1.3.3.3 Summary

At $r = r_0$, the local halo density is determined by the local Galactic surface density Σ_0 through Eqs. (1.47), (1.40), (1.43), and (1.45). Figure 1.10 shows the dependence of the local halo density ρ_0 and the size of the dark halo core a/r_0 on the local Galactic disk density Σ_0. There are more discussions about the measurements of Σ_0, their uncertainties, and the resulting ρ_0 in Ref. [7] and references therein. We adopt the canonical value $\rho_0 = 0.3\,\mathrm{GeV\,cm^{-3}}$ in this thesis, which corresponds to $\Sigma_0 = 82\,M_\odot\,\mathrm{pc^{-2}}$ and $a = 6.4\,\mathrm{kpc}$. The resulting contributions to the rotation speed from the halo and disk are $121\,\mathrm{km\,s^{-1}}$ and $184\,\mathrm{km\,s^{-1}}$, respectively.

1.4 Dark matter candidates

1.4.1 Basic properties

A variety of evidence supporting the existence of dark matter at different spatial and temporal scales of the universe have been presented in earlier sections. We can infer the basic properties that any dark matter candidate must have to fit all the scenarios. Some of the known Standard Model (SM) matter in the universe can escape our detection because of their low detectability, such as the massive compact halo objects (MACHOs) and neutrinos. However, current constraints show that their contributions are too small to account for the abundance of the dark matter in the universe. The majority of the dark matter must be some yet to be discovered (likely) non-SM particles. We list their properties below.

- *Interacting at the weak scale or weaker*: The no-detection by current astrophysical and terrestrial experiments indicate that the candidate does not participate strongly in electromagnetic or strong interactions. Besides the gravitational interaction, if other interactions between dark matter particle themselves or between dark matter and ordinary matter are present, they must be very weak, which explains the collisionless behavior shown in the Bullet Cluster.
- *Massive and cold*: The candidate must be cold (non-relativistic) and undergo gravitational collapse in the early universe. And it must be massive to account for the dark matter abundance in the universe.
- *Non-baryonic*: With the constraints shown by BBN, CMB, and large scale structure formation, the majority of dark matter must be non-baryonic.
- *Stable*: These particles were produced in the early universe. Their lifetime must

be sufficiently long to account for the current abundance.

Because of the properties listed above, the dark matter candidates fall under this category acquired a generic name: Weakly Interacting Massive Particles (WIMPs). We discuss more details about WIMPs and other possible candidates in the following sections.

1.4.2 WIMPs

Weakly Interacting Massive Particles (WIMPs) are the most studied and currently the primary focus of the direct and indirect dark matter search experiments. Besides their apparent properties which enlist them as a candidate for dark matter, their mass and annihilation cross section are in coincidence with the lightest stable particles predicted by a number of theories beyond the Standard Model, naturally bridging cosmology and particle physics.

Like ordinary matter, WIMPs were produced thermally in the early universe and in thermal equilibrium with other particles. The abundance of the WIMP was determined by the Boltzmann statistics. As the universe cools off and the temperature was no longer higher than the WIMP mass m_χ, the equilibrium favors the annihilation direction with decreasing temperature. The density of the WIMP is thus given by $n \propto n_0 e^{-m_\chi/T}$, where n_0 is the equilibrium density at $T > m_\chi$. Had the WIMPs always been in thermal equilibrium, they would have disappeared completely by now. Fortunately, at some point in the early universe, the Hubble expansion rate exceeded the WIMP annihilation rate. The annihilation stopped and the comoving density of the WIMP remains constant since. A higher WIMP annihilation cross section would delay the freeze out to a later time and result in a lower comoving density, and the

opposite would happen for a lower cross section. Therefore, the WIMP annihilation cross section can be constrained by the their present relic density in the universe. Detailed calculations show that [7, 18]

$$\Omega_\chi h^2 \simeq \frac{0.1\,\mathrm{pb}\,c}{\langle \sigma_A v \rangle}, \tag{1.48}$$

where c is the light speed, σ_A is the WIMP annihilation cross section, and v is the WIMP relative velocity. For not so unrealistic estimates $\Omega_\chi h^2 \sim 1$ and $v \lesssim c$, we have $\sigma_A \sim 0.1\,\mathrm{pb}$. The conclusion above is from thermodynamic arguments and is independent of the WIMP mass.

From the particle physics side, if a new particle with weak-scale interactions exists, then its annihilation cross section would be $\sigma \sim G_F^2 (100\,\mathrm{GeV})^2 \approx 0.01\,\mathrm{pb}$. This remarkable coincidence motivated a lot of theorists and experimentalists to search for this yet unknown particle. Aside from providing a dark matter particle for the cosmology, the theoretical motivation to search for new physics beyond the SM is to cure the "hierarchy problem" in particle physics, i.e., the delicate cancellation to the quadratic divergent radiative corrections to the Higgs scalar mass. Numerous models have been proposed by different authors, and the most studied framework is supersymmetry which we briefly touch upon.

1.4.2.1 Supersymmetry

Supersymmetry is a geometrical symmetry of spacetime like the Poincaré symmetry in Minkowski spacetime. It extends the Poincaré algebra with supercharges, the generators of operators which transform a fermion field into a boson field with

the same mass and vise versa. One can construct any theory to be supersymmetry invariant in superspace.

One appealing feature of the supersymmetric quantum field theory is that the quadratic divergent radiative corrections to the mass of the scalar fields vanish in all orders because of the boson-fermion mass degeneracy, which introduces opposite signs to the corrections. This resolves the "hierarchy problem" in SM naturally, which is arguably the most important motivation underlying the supersymmetric extension of the Standard Model. No known particles possess the properties of the supermultiplets, for the same argument, supersymmetry must be a broken symmetry (at some energy level, if it exists), thus super partners to the SM particles have to be added in the supersymmetric extension of the SM. The extension with the minimum number of super partners introduced is referred to as the Minimal Supersymmetric Standard Model (MSSM).

In the MSSM framework, before supersymmetry breaking, each spin-1/2 matter field (leptons and quarks) is paired with two spin-0 spartners, one for each chirality, called sfermions (sleptons and squarks). Each spin-1 gauge field is paired with a gaugino (bino, winos, zino, and gluinos). An additional Higgs doublet is added for the requirement of the theory as well as the spartners of all the Higgs, higgsinos. Electroweak symmetry breaking mixes the gauginos and higgsinos and forms the physical charginos and neutralinos. More than 100 parameters are introduced to the MSSM to "soft" break the supersymmetry. Phenomenological studies usually focus on constrained models with a manageable number of parameters. A new symmetry called "R parity", defined as $R = (-1)^{2s+3B+L}$, where s, B, and L are the particle's spin, baryon, and lepton number, respectively, is introduced. SM particles all have $R = 1$

and the s-particles have $R = -1$. If R parity is conserved, the lightest supersymmetric particle (LSP) would be stable and serve as a natural candidate for dark matter. The neutralino is the most studied candidate for dark matter, though other possibilities exist for different mixings. More details about MSSM and their relevance to dark matter can be found in [7, 31, 32].

The recent discovery of the Higgs boson of mass around 125 GeV at the LHC put stringent constraints on the allowed parameter space of MSSM. Though MSSM is not completely ruled out as a viable SM extension, the required fine tuning to accommodate the Higgs mass weakens its original motivation to solve the hierarchy problem. And there is more interest shifting to the next-to-minimal supersymmetric extension to the SM [33].

1.4.3 Axions

Axion is the pseudo Nambu-Goldstone boson resulting from the Peccei-Quinn symmetry breaking, proposed to preserve the CP symmetry of the QCD Lagrangian. The production of axion is non-thermal through the misalignment mechanism, i.e., they would be non-relativistic even if produced in the hot early universe. This makes them a suitable cold dark matter candidate.

The axion mass is determined by the Peccei-Quinn symmetry breaking scale f_a:

$$m_a \approx \frac{6\,\text{eV}}{f_a/10^6\,\text{GeV}}. \tag{1.49}$$

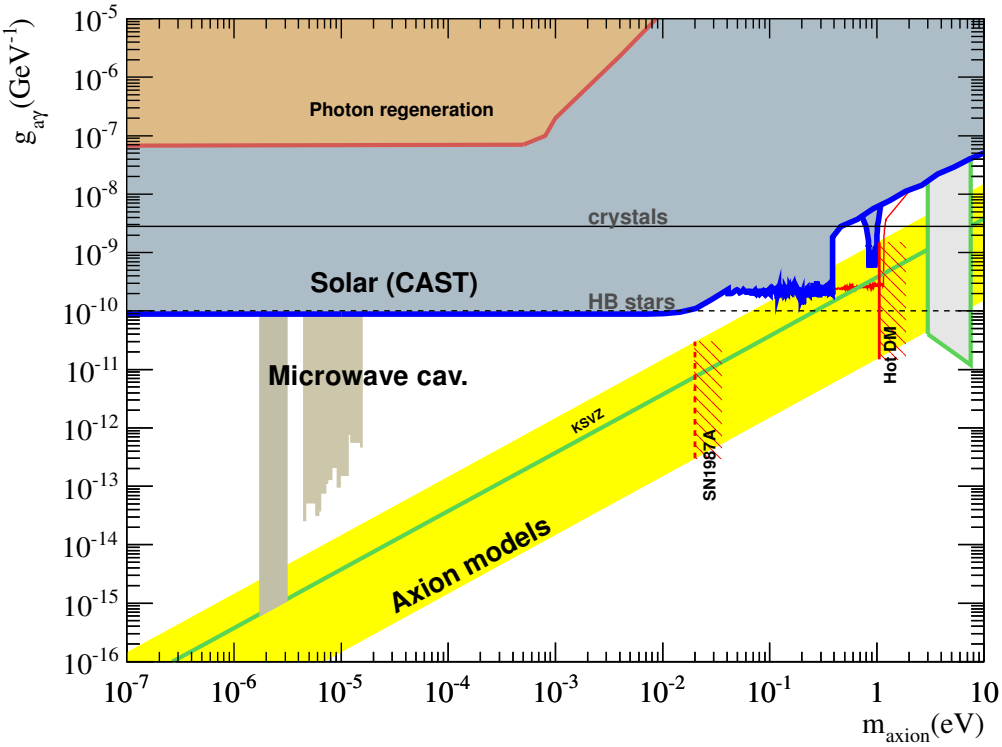


Figure 1.11: Experimental constraints on the axion mass and axion-photon coupling strength. The green line shows the prediction of the KSVZ model. The yellow band shows the range of the somewhat realistic models. The microwave cavity experiments are starting to constrain the cosmologically interesting region. *Figure from Irastorza* et al. *[34].*

Their relic density is [18]

$$\Omega_a h^2 = \kappa_a \left(\frac{f_a}{10^{12}\,\mathrm{GeV}} \right)^{1.175} \theta_i^2, \tag{1.50}$$

where κ_a is a constant between 0.5 and a few. If $\theta_i \sim O(1)$, the abundance of the non-baryonic dark matter is saturated by the axion at $f_a \sim 10^{11}\,\mathrm{GeV}$. This gives a axion mass $m_a \sim 60\,\mu\mathrm{eV}$. If the abundance of the axion is smaller, the mass scales accordingly. The most interesting range generally lies between $\sim$1 µeV and $\sim$1 meV.

There are a few experiments actively search for the axion (or axion like particles, ALP), notably, CERN Axion Solar Telescope (CAST) and Axion Dark Matter Searches (ADMX). These experiments rely on the coupling $g_{a\gamma}\mathbf{E}\cdot\mathbf{B}\,a$, which converts an axion into a photon in strong magnetic field. The produced photon can then be detected, for example, in a microwave cavity. Figure 1.11 shows the relatively recent status of the axion search. The green line shows the prediction of the KSVZ model. The yellow band shows the range of the somewhat realistic models. The microwave cavity experiments are starting to constrain the cosmologically interesting region.

1.4.4 Others

There are countless other dark matter candidates proposed by different authors. We list a few here.

MOND: Modified Newtonian dynamics. This was originally proposed to explain the flat rotation curves observed in spiral galaxies and it more or less served its purpose. However, as more evidence supporting the dark matter hypothesis come up, MOND has difficulties to provide satisfactory explanations. For example, in the case of the Bullet Cluster, there is clear separation between gravitational potential and the visible matter. MOND itself can not explain the observation without bringing in additional mass to the system such as neutrinos. Moreover, MOND is not relativistic, it can not explain gravitational lensing, and it is difficult to apply MOND to cosmological scales such as in the CMB. While it is an interesting idea, it is less appealing as the replacement for the dark matter hypothesis.

SuperWIMPs: While this category of candidates may have different names, the essence is that they only have the gravitational interaction or other interactions which

are very weak, much weaker than the weak sale. Gravitino falls under this category.

Asymmetric dark matter: Like ordinary matter, dark matter and anti-dark-matter were both produced in the early universe. Because of the asymmetry between them, only dark matter survived. The annihilation cross section between dark matter could be much smaller than between dark matter and anti-dark-matter. This is in line with the absent of dark matter signals in the indirect detection experiments.

1.5 Dark matter detection

Despite of the overwhelming astrophysical and cosmological evidence for the existence of dark matter, the particle nature of the dark matter still remains unknown. Detecting dark matter particles posits one of the biggest challenges in the current physics research. The schematic in Fig. 1.12 depicts the processes involved in the possible detection schemes: direct detection, indirect detection, and collider production. In direct detection, the dark matter particles interact with the detectors directly and the latter record the signals registered by the scattering events. The indirect detection searches for the annihilation products of the dark matter, mainly gamma rays, neutrinos, and antimatter. Lastly, the dark matter particles may be produced in a particle collider like LHC if the collisions have enough energy. The signature of production of dark matter would show up as missing energy and momentum in a collision event. The three different detection approaches are complimentary and each probes a different category of reactions as depicted in Fig. 1.12.

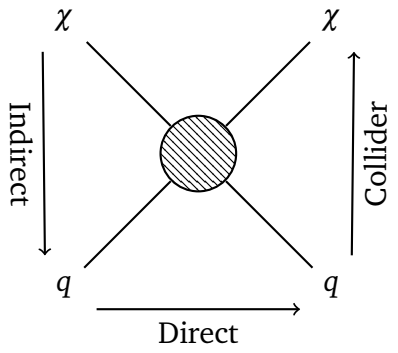


Figure 1.12: Dark matter particle detection schemes. χ and q represent dark matter and standard model particles, respectively. The direct detection utilizes the process $\chi q \rightarrow \chi q$. The annihilation $\chi\chi \rightarrow qq$ and and production $qq \rightarrow \chi\chi$ are for the indirect detection and the collider searches, respectively. *Figure adapted from Moore [35].*

1.5.1 Direct detection

As was discussed in Sec. 1.3, the dark matter particles in the halo of our Milky Way are non-relativistic. They are believed to scatter elastically off the target nuclei when interacting with the dark matter detector. Scenarios involving inelastic scatters have also been studied by some authors [36–41]. However, we will focus on the elastic case.

1.5.1.1 Event rates and energy spectra

In the simplest case, where a beam of WIMPs with constant number density n and constant velocity v collide with a target of N nuclei at rest, the scattering rate between them is $nv\sigma N$, where σ is the WIMP-nucleus scattering cross section. Considering the WIMP velocity distribution, the recoil energy dependence of the cross section,

and the motion of the Earth, the differential scattering rate can be written as

$$\frac{\mathrm{d}R}{\mathrm{d}E} = nN \int_S v \frac{\mathrm{d}\sigma}{\mathrm{d}E} f(\mathbf{v}, \mathbf{v}_E)\, \mathrm{d}^3\mathbf{v}, \tag{1.51}$$

where R is the event rate, $\mathbf{v}$ is the relative velocity of a WIMP with respect to the Earth, $\mathbf{v}_E$ is the Earth velocity in the Galactic frame, $f(\mathbf{v}, \mathbf{v}_E)$ is the WIMP velocity distribution, and E is the energy transfer (or the recoil energy of the recoiling nucleus in the lab frame). The integration is carried out in the manifold S, which is discussed in more detail below. The details of the WIMP-nucleus scattering are encapsulated in the differential cross section $\mathrm{d}\sigma/\mathrm{d}E$, which may have complex dependence on the momentum transfer. Here we assume σ is isotropic and only depends on the recoil energy E. The dependence of σ on E largely dictates the selection of target material, which we will discuss in more detail later.

We did not limit the range of E in Eq. (1.51). In practice, however, a dark matter detector can only detect the scattering events with recoil energies above a certain minimum value. This detector energy threshold E_{th} limits the amount of WIMPs the detector can "see". For WIMPs with a certain mass, those ones which are too slow will not be detected by the detector. Similarly, the finite escape speed v_{esc} means the WIMPs that are too light will escape the detection as well. In a word, a dark matter detector with finite energy threshold will miss the lower tail of the velocity or mass spectra for the WIMPs.

With the simple hard-ball model, the energy transfer of a WIMP with mass m and

initial velocity v colliding into a target nucleus M at rest is

$$E = \frac{2mM}{(m+M)^2}(1-\cos\theta)\frac{1}{2}mv^2 = \frac{r}{4}mv^2(1-\cos\theta), \tag{1.52}$$

where θ is the scattering angle in the center of momentum (COM) frame and $r \equiv \frac{4mM}{(m+M)^2}$. The requirement $E > E_{\text{th}}$ gives

$$v > v_{\text{min}} = \sqrt{\frac{2E_{\text{th}}}{mr}} \tag{1.53}$$

for a fixed WIMP mass m, and

$$m > \frac{M}{\sqrt{\frac{2Mv_{\text{max}}^2}{E_{\text{th}}}} - 1} \tag{1.54}$$

for a given cutoff velocity v_{max}. The latter constraint is important in designing experiments for detecting light WIMPs. Nevertheless, in the discussions below, we assume the WIMPs are heavy enough to give us a detectable signal.

Based on Eq. (1.52), we can have some rough estimates about the parameter range we are trying to probe. A 50 GeV WIMP moving at the typical velocity $v_0 = 220\,\text{km}\,\text{s}^{-1}$ has kinetic energy 13 keV. The mass of the Ge target used in CDMS is 73 GeV. This gives $r = 0.97$. So the recoil energy will be between 0 and 13 keV*. For the same Ge target with a 2 keV threshold, the minimum relative velocity is $v_{\text{min}} = 86\,\text{km}\,\text{s}^{-1}$. The highest possible relative velocity is $v_{\text{max}} = v_{\text{esc}} + v_0 = 764\,\text{km}\,\text{s}^{-1}$ (we assume the Earth moves at the circular velocity v_0). Together with the given threshold, it gives

*The recoil energy of an electron scattering off a WIMP is typically less than 1 eV, which is far below any dark matter detector energy threshold.

the minimum detectable WIMP mass 3.5 GeV. Note that if light WIMPs do exist, it will be more challenging than detecting "regular" WIMPs with mass ∼100 GeV, since the detector is only sensitive to WIMPs on the high velocity tail and the deposited energy is close to the detector threshold. Both the event rates and the signal-to-noise ratio are reduced.

Under the assumption that the scattering is isotropic (see Eq. (1.52)), the differential cross section is also independent of the energy transfer. This gives the differential cross section a simple form in the entire recoil energy range

$$\frac{\mathrm{d}\sigma}{\mathrm{d}E} = \frac{\sigma_0}{mv^2 r/2}, \tag{1.55}$$

where σ_0 is the total cross section.

The same Maxwellian velocity distribution in Eq. (1.37) is used in Eq. (1.51). Explicitly,

$$f(\mathbf{v}, \mathbf{v}_E) = k \exp\left(-\frac{(\mathbf{v}+\mathbf{v}_E)^2}{v_0^2}\right) \tag{1.56}$$

and k is a normalization factor.

The integral manifold S in Eq. (1.51) is constrained by

$$S = \{\mathbf{v} : v_{\min} < |\mathbf{v}| \ \wedge \ |\mathbf{v}+\mathbf{v}_E| < v_{\mathrm{esc}}\}. \tag{1.57}$$

Putting things together, we can write Eq. (1.52) as

$$\frac{\mathrm{d}R}{\mathrm{d}E} = \frac{2nNk\sigma_0}{mr} \int_S \frac{1}{v} \exp\left(-\frac{(\mathbf{v}+\mathbf{v}_E)^2}{v_0^2}\right) \mathrm{d}^3\mathbf{v}. \tag{1.58}$$

The above integral can be integrated out analytically. And the result [28, 38, 42] is,

for $0 < v_{\rm min} < v_{\rm esc} - v_E$,

$$\begin{aligned}\frac{\mathrm{d}R}{\mathrm{d}E} =& \frac{2nNk\sigma_0}{mr} \times 2\pi v_0^2 \times \\ & \left\{ \frac{\sqrt{\pi}}{4} \frac{v_0}{v_E} \left[\mathrm{erf}\left(\frac{v_{min} + v_E}{v_0} \right) - \mathrm{erf}\left(\frac{v_{min} - v_E}{v_0} \right) \right] - e^{-v_{esc}^2/v_0^2} \right\},\end{aligned} \tag{1.59}$$

and for $v_{\rm esc} - v_E < v_{\rm min} < v_{\rm esc} + v_E$,

$$\begin{aligned}\frac{\mathrm{d}R}{\mathrm{d}E} =& \frac{2nNk\sigma_0}{mr} \times 2\pi v_0^2 \times \\ & \left\{ \frac{\sqrt{\pi}}{4} \frac{v_0}{v_E} \left[\mathrm{erf}\left(\frac{v_{\rm esc}}{v_0} \right) - \mathrm{erf}\left(\frac{v_{min} - v_E}{v_0} \right) \right] - \frac{v_{\rm esc} + v_E - v_{min}}{2v_E} e^{-v_{esc}^2/v_0^2} \right\}.\end{aligned} \tag{1.60}$$

In the limit $v_E \to 0$ and $v_{\rm esc} \to \infty$, the differential event rate can be integrated from Eq. (1.58), which is

$$\begin{aligned}\frac{\mathrm{d}R}{\mathrm{d}E} &= \frac{2nNk\sigma_0}{mr} 2\pi v_0^2 \exp(-v_{min}^2/v_0^2) \\ &= \frac{2nNk\sigma_0}{mr} 2\pi v_0^2 \exp\left(-\frac{E}{rmv_0^2/2} \right),\end{aligned} \tag{1.61}$$

where we have used Eq. (1.53) to translate v_{min} to the corresponding recoil energy. Clearly, in this limit, the energy spectrum is exponential. And the characteristic energy is the maximum energy transfer from a WIMP with the characteristic velocity v_0.

The differential event rate is commonly expressed in units of $\mathrm{kg^{-1}\,keV^{-1}\,d^{-1}}$, i.e., counts per kilogram target material per keV per day. Under this convention, the number of target nuclei is

$$N = \frac{N_A}{\mu}, \tag{1.62}$$

where μ is the molar mass of the target and N_A is the Avogadro number. The dark

matter number density is

$$n = \frac{\rho_0}{m}, \tag{1.63}$$

where $\rho_0 = 0.3\,\mathrm{GeV\,cm^{-3}}$, is the local dark matter density.

1.5.1.2 Nuclear form factors

In the derivations above, we assumed that the WIMP-nucleus scattering cross section σ does not depend on the energy transfer of the collision. This is generally not true. Especially when the wavelength of the incident particle approaches the size of the nucleus, the cross section drops as the energy of the incident particle goes up.

Without knowing the interaction details, empirically, the dependence of σ on the momentum transfer $q = \sqrt{2ME}$ can be written as [7, 28, 43]

$$\frac{\mathrm{d}\sigma}{\mathrm{d}q^2} = \frac{\sigma_0}{4m_r^2 v^2} F^2(q), \tag{1.64}$$

where $m_r = \frac{mM}{m+M}$ is the reduced mass, v is the relative velocity between the incident particle and the target, and F is the nuclear form factor, which encodes the momentum transfer dependence of the differential cross section. The normalization is that $F(0) = 1$, so the cross section at zero momentum transfer is recovered.

In the first Born approximation, the form factor is the Fourier transform of the density distribution of the scattering centers. For spin independent (SI) interaction, F is the Fourier transform of the nucleon density in the nucleus. It was found the form factors for different nuclei can be expressed in a universal form for the spin independent interaction. The one commonly used is the Helm form factor [28]. When expressed as a function of the dimensionless quantity $qr_n/\hbar$ (r_n is the effective nuclear

radius and $\hbar = 1$ here), it is [28]

$$
\begin{aligned}
F(qr_n) &= 3\frac{j_1(qr_n)}{qr_n} e^{-(qs)^2/2} \\
&= 3\frac{\sin(qr_n) - qr_n\cos(qr_n)}{(qr_n)^3} e^{-(qs)^2/2}, \\
r_n^2 &= c^2 + \frac{7}{3}\pi^2 a^2 - 5s^2, \\
c &= 1.23A^{1/3} - 0.60\,\text{fm}, \\
a &= 0.52\,\text{fm}, \\
s &= 0.9\,\text{fm},
\end{aligned}
\tag{1.65}
$$

where j_1 is the Bessel function and A is the mass number of the nucleus. The form factors for a few commonly used target nuclei in direct dark matter searches are plotted in Fig. 1.13.

For the spin dependent (SD) case, however, there is no such universal analytical form for the form factor. Detailed calculations must be performed for each nuclear structure. Conventionally, the form factor is expressed as [7, 28]

$$F^2(q) = S(q)/S(0), \tag{1.66}$$

and

$$S(q) = a_0^2 S_{00}(q) + a_0 a_1 S_{01}(q) + a_1^2 S_{11}(q), \tag{1.67}$$

$$a_0 = a_p + a_n, \quad a_1 = a_p - a_n, \tag{1.68}$$

where a_p and a_n are the spin coupling constants for protons and neutrons in a nucleus,

respectively. The three terms in $S(q)$ describe the contributions from isoscalar $(p+n)$, isovector $(p-n)$, and the interference between them.

To compare the results or assess the performances of different target materials, it is desirable to normalize the cross section to that of the nucleons, i.e., protons and neutrons. This involves the details of the WIMP-nucleon interaction and is model dependent. We will limit our discussion within the supersymmetric dark matter. Through squark or Higgs exchange with quarks, WIMPs could have a scalar interaction with nuclei. The total cross section for SI interaction at zero momentum transfer is [7, 28, 43, 44]

$$\sigma_0^{\mathrm{SI}} = \frac{4}{\pi} m_{\chi N}^2 \left[Z f_p + (A-Z) f_n \right]^2, \tag{1.69}$$

where $m_{\chi N}$ is the reduced mass of a WIMP and a nucleus, Z is the atomic number, and f_p and f_n are the scalar coupling constants for protons and neutrons, respectively. For the lightest neutralino WIMP, we have $f_p \simeq f_n$ [7, 43], which gives $\sigma_0^{\mathrm{SI}} \propto A^2$. WIMPs see no difference between protons and neutrons in scalar interaction. All scattering amplitudes add coherently which gives the A^2 enhancement in the cross section. Thus we have

$$\sigma_0^{\mathrm{SI}} \simeq \frac{4}{\pi} m_{\chi N}^2 A^2 f_p^2. \tag{1.70}$$

Assuming Eq. (1.70) still holds for a single nucleon, say, a proton, similarly, we have

$$\sigma_{\chi p}^{\mathrm{SI}} \simeq \frac{4}{\pi} m_{\chi p}^2 f_p^2. \tag{1.71}$$

Then we can define the scaling between the WIMP-nucleus and WIMP-nucleon cross

sections as

$$\sigma_0^{\mathrm{SI}} = A^2 \left(\frac{m_{\chi N}}{m_{\chi p}} \right)^2 \sigma_{\chi p}^{\mathrm{SI}}. \tag{1.72}$$

WIMPs could also have spin dependent interaction with nuclei through squark or Z boson exchange with quarks. The cross section for the SD interaction is [7, 43, 44]

$$\sigma_0^{\mathrm{SD}} = \frac{32}{\pi} G_F^2 m_{\chi N}^2 \frac{J+1}{J} \left[a_p \langle S_p \rangle + a_n \langle S_n \rangle \right]^2, \tag{1.73}$$

where G_F is the Fermi coupling constant, J is the nuclear spin, $\langle S_p \rangle = \langle N | S_p | N \rangle$ is the expectation value of the spin of the proton group in the nucleus, a_p is the spin coupling constant for protons, which we introduced in Eq. (1.68), and similarly for $\langle S_n \rangle$ and $\langle a_n \rangle$. The spin dependent scattering amplitudes from a pair of nucleons with opposite spins cancel out. Only the unpaired nucleon spin in the nucleus, either a proton or a neutron, contributes effectively. Though calculations show that $\langle S_n \rangle$ may be comparable to $\langle S_p \rangle$ for some proton-spin dominant nuclides [7], and vice verse, it is usually a good approximation to assume only the unpaired spin contributes substantially. Unlike the spin independent case, a_p and a_n are generally different as well as WIMP type dependent in the spin dependent interaction [45]. Nevertheless, we can simplify Eq. (1.73) when one species dominates:

$$\sigma_0^{\mathrm{SD}} = \frac{32}{\pi} G_F^2 m_{\chi N}^2 \frac{J+1}{J} \langle S_{(p,n)} \rangle^2 a_{(p,n)}^2. \tag{1.74}$$

For a single proton or neutron, $J = 1/2$ and $\langle S_{(p,n)} \rangle = 1/2$, so we obtain

$$\sigma_0^{\mathrm{SD}} = \frac{4}{3} \frac{J+1}{J} \langle S_{(p,n)} \rangle^2 \sigma_{\chi(p,n)}^{\mathrm{SD}}. \tag{1.75}$$

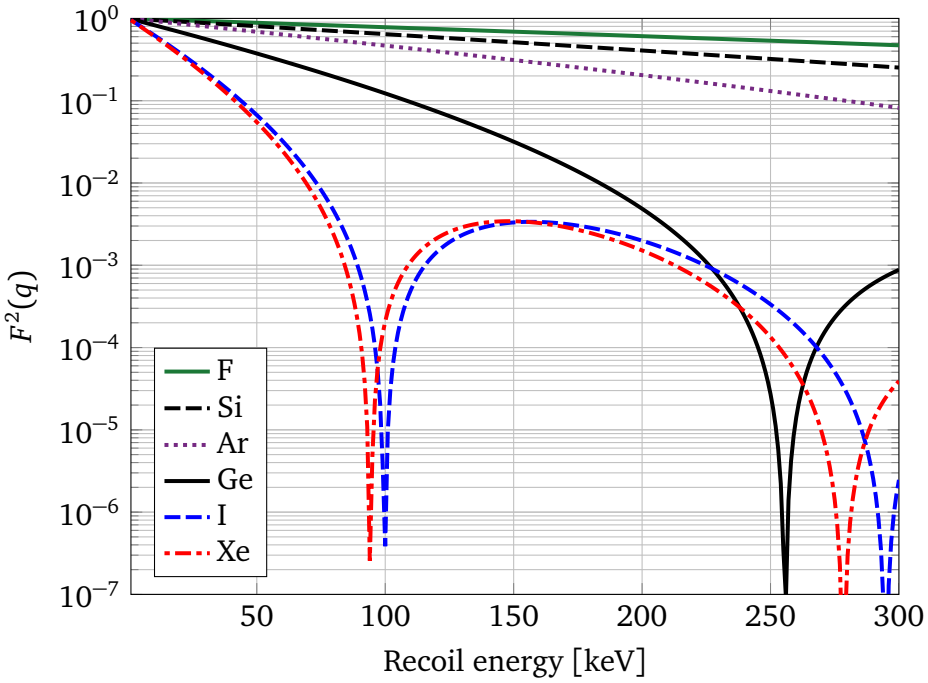


Figure 1.13: Helm nuclear form factors for SI cross section of few commonly used nuclei in direct detection experiments. Note that the first minimum happens earlier for heavier nuclei.

The cross sections measured in experiments include contributions from both SI and SD (if nuclear spin is nonzero) interactions. Because of the A^2 enhancement to the SI cross section, it is preferable to use heavy nuclei as the target material. However, since the SD cross section is only sensitive to the unpaired single proton or neutron in the nucleus, light nuclei with large nuclear spin are preferred in SD detection, such that higher nuclear spin can be obtained for unit mass. It is nontrivial to optimize the sensitivities to both SI and SD interactions for a single detector. The current direct dark matter detection experiments mostly focus on the SI cross section. There are other complications brought in by the nuclear form factors in selection of the target materials. As shown in Fig. 1.13, the Helm nuclear form factor for the SI

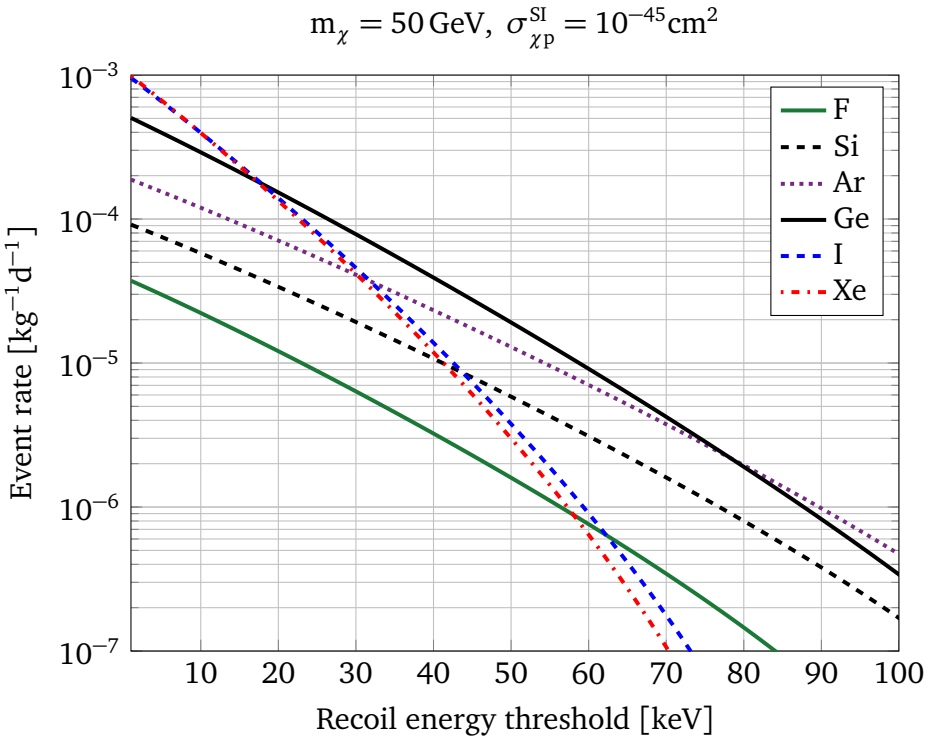


Figure 1.14: Integrated event rate vs. detector energy threshold for a few commonly used target nuclei interacting with 50 GeV WIMPs. The zero momentum transfer WIMP-nucleon cross section is chose as $10^{-45}\mathrm{cm}^2$; Solar circular speed is $v_0 = 220\,\mathrm{km\,s^{-1}}$; and the escape velocity is $v_{esc} = 544\,\mathrm{km\,s^{-1}}$. Detector is assumed 100 % efficient above the threshold. Below ∼18 keV, Xe and I have the highest integrated rates. Ge has the highest rate between roughly 18 and 80 keV.

cross section has an oscillatory behavior. Heavier nuclei reach the first minimum at lower recoils energies. For Xe and I, the drop in $F^2(q)$ before 100 keV, the energy range we are most interested in, largely impairs the advantage of the A^2 enhancement for detecting reasonably heavy WIMPs. Ge is arguably the optimal target material choice in the recoil energy range 0 to 100 keV, especially considering the availability of high purity Ge from the semiconductor industry.

Figure 1.14 shows the integrated event rate in units of $\mathrm{kg^{-1}d^{-1}}$ as a function of

the detector recoil energy threshold. The heavy nuclei Xe and I have the highest event rates at low thresholds. Ge performs better at thresholds 18 to 80 keV. It is crucial to lower the detector thresholds in direct dark matter detection experiments, while at the same time, the detector would likely be more susceptible to backgrounds, which we will discuss in more detail in Sec. 1.5.1.4.

1.5.1.3 Annual modulation

The relative speed between the Earth and the dark halo modulates sinusoidally because of the orbital motion of the Earth around the Sun. The speed modulation translates into the modulation of WIMP flux in the lab frame, which gives an annual modulation of the WIMP event rate in the dark matter detector.

In the Galactic coordinate system, the rotation velocity of the local standard of rest (LSR) of the Sun is $(0, 220, 0)$ km s^{-1} [27]. And the peculiar velocity of the Sun with respect to the LSR is $(9, 12, 7)$ km s^{-1} [28]. The Earth orbits the Sun in a nearly circular orbit whose plane has an angle 60° with respect to the velocity of the Sun. The average orbital speed of the Earth in the rest frame of the Sun is $\sim$30 km s^{-1}. The maximum and minimum projections of the ecliptic velocity of the Earth in the direction of the Galactic velocity of the Sun are $\pm$15 km s^{-1}, respectively, which gives the modulation in the direction of the Sun's motion:

$$u_g \approx 232 + 15\cos\left(2\pi\frac{t - t_0}{365.25}\right), \tag{1.76}$$

where 232 km s^{-1} is the speed of the Sun in the Galaxy, t is time in Julian days with t_0 around June 2. This 6 % speed modulation gives $\sim$3 % modulation in event rate [28].

A more detailed derivation of u_g is given in Appendix A.

1.5.1.4 Backgrounds

While a dark matter detector is designed for searching for dark matter particles, unwanted signals originating from the interactions between the detector and other particles, notably gammas, betas (e^-), neutrons, and α particles, are also present. The sources, impacts, and mitigations of these backgrounds are unique to each experiment. A thorough discussion is impossible without referring to a specific experiment. In this section, we have a general discussion of the statistical impacts of the presence of the backgrounds, possible sources, and approaches for mitigations.

Direct dark matter search results are usually reported in two ways (steps):

(1) Calculate the discovery potential by comparing the observed events to a background model. If the events can not be explained by possible background fluctuations at a certain significance level, e.g., 5σ, it is an indication of a discovery of new particles. Regardless of whether a claim of discovery can be made, the likelihood that the signals are compatible with the background model is reported.

(2) If there is no indication of a discovery, an upper limit of the WIMP-nucleon scattering cross section is calculated from the observed events and the background model. Since the limits are directly related to the *expected* signal event counts, the treatment of the background in the *observed* events can vastly change the result. In the CDMS experiment, all observed events are treated as WIMPs in setting the limits.

Before discussing the impacts of the backgrounds to the results, we need to clarify

a few terms we will use. The aforementioned WIMP-nucleon cross section limits in Step (2) is also referred to as the sensitivity, which is defined as the *average* upper bound that would be obtained by an ensemble of experiments with the expected background and no true signal [46, 47]. This is the definition used in the CDMS data analysis as well as this work. Here we show this definition explicitly.

Suppose in an experiment, the signal and the background are Poisson processes with mean event counts s and b, respectively, the sum of the signal and the background is a Poisson process with mean event count $\lambda = s + b$. If an observation of k events is obtained, estimates of λ itself $\lambda(k)$ and its upper bound $\lambda_\alpha(k)$ at confidence level (CL) α can be calculated. We know $\lambda(k)$ is just k but we will keep the notation for clarity. $\lambda_\alpha(k)$ can be computed by different methods such as the Poisson confidence interval (CI), the Bayesian interval, or the Feldman-Cousins interval [18]. Without a signal, $b = \lambda(k)$, and then the sensitivity is

$$\lambda_\alpha^s = \sum_{i=0}^{\infty} \frac{\lambda(k)^i e^{-\lambda(k)}}{i!} \lambda_\alpha(i). \tag{1.77}$$

More often than not, the upper bound λ_α is referred to as the "upper limit" in literature, which is very confusing when talking about the dark matter search limits at the same time. However, there is another level of inconsistency and confusion in the definition of the sensitivity among the authors [46, 48–50]. Some argued that the sensitivity should be chosen based on the outcome of hypothesis tests on $\lambda(k)$. If suitable tests were chosen, $\lambda_\alpha(k)$ would be chosen as the sensitivity. We will not expand on this and stick with the definition in Eq. (1.77).

When no events are observed, no discovery can be made. An upper limit of the

WIMP-nucleon cross section based on the sensitivity $\lambda^s_\alpha = \lambda_\alpha(0)$ can be reported. The sensitivity λ^s_α is the mean *expected* signal upper bounds at CL α given the observation, which satisfies $\lambda^s_\alpha \propto \sigma^s_0 M t$, where M and t are the target mass and the experiment running time, respectively. So the upper limit of the cross section $\sigma^s_0 \propto \frac{\lambda^s_0}{Mt}$, which is inversely proportional to the exposure Mt.

When there are a substantial number of, say, k, events observed and there is potential for a discovery, we construct a hypothesis test with the null hypothesis H_0 : $\mu \leq b$[†], where μ is the Poisson parameter for the observed events. At the pre-chosen significance level β, if $P(x > k; b) = \sum_{k+1}^{\infty} \frac{b^i e^{-b}}{i!} \leq \beta$, then H_0 is rejected and there is indication of new events showing up. For the common choice of $5\,\sigma$ significance, $\beta = 2.87 \times 10^{-7}$, a single observation would incur a discovery if $b \leq 0.012$.

The experiments run continuously, and the expected background events accumulate with time. Suppose the event rates for the signal, the background, and the observation are r_s, r_b, and r_o, respectively, at the experiment running time t, the distribution of the expected background events can be approximated by the normal distribution $N(r_b t, r_b t)$ (assuming $r_b t$ is not too small). The criterion of a discovery can be satisfied when $r_o t = r_b t + Z\sqrt{r_b t}$, where $Z = \Phi^{-1}(1-\beta)$ and Φ is the cumulative distribution of the Standard Gaussian. This gives the required running time $t = \frac{Z^2 r_b}{(r_o - r_b)^2}$ and $\frac{\delta t}{t} = \frac{r_o + r_b}{r_o - r_b}\frac{\delta r_b}{r_b}$. The relative uncertainties on the experimental runtime is always larger than that of the background estimate.

There is this intermediate region, where some events are observed but not significant to claim a discovery. Then the sensitivity at CL α is reported. Assuming the

[†]The null hypothesis can be chosen as $H_0 : \mu = b$, however, since we only care about whether the observed events excess the background, the left tail does not concern us. It is possible in extreme cases the observed events are lower than the estimated background, then both the experiment and the background estimate need be examined carefully to rule out possible errors.

number of the observed events is not too small, the distribution can be approximated by $N(r_o t, r_o t)$. The probability in Eq. (1.77) falls off quickly when i deviates the observed counts $r_o t$, so we have $\lambda_\alpha^s \approx \lambda_\alpha(r_o t) = r_o t + Z_\alpha\sqrt{r_o t}$. Now we need to decide how to deal with the background:

(1) No background subtraction. This means $\sigma_0^s M t \propto r_o t + Z_\alpha\sqrt{r_o t}$, which gives $\sigma_0^s \propto \frac{1}{M}(r_o + Z_\alpha\sqrt{r_o}/\sqrt{t})$. It tells us the upper limit on the cross section goes down as the runtime increases but very slowly, and as $t \to \infty$, it measures the cross section of the signal plus background. In current direct dark matter search experiments $s \ll b$, the reachable cross section is background limited.

(2) With background subtraction. Properly subtracting the background requires the distribution of the difference of two Poisson processes. We treat both the observation and the background as Gaussian variables. The difference is another Gaussian with mean $r_o t - r_b t$ and variance $r_o t + r_b t$. Then we have $\sigma_0^s \propto \frac{1}{M}(r_o - r_b + Z_\alpha\sqrt{r_o + r_b}/\sqrt{t})$. Again we assume $s \ll b$, (a) if the background is known and well characterized, the cross section will converge to that of the signal after sufficiently long runtime, while at the earlier stage of the experiment the upper limit on cross section drops as $1/\sqrt{t}$; (b) if background is underestimated, this case is between no background subtraction and full background subtraction, the specific behavior depends on how large the residual background is, but it's possible the residual background is still large compared to the signal and still dominates the cross section; (c) the background is overestimated, with the assumption $s \ll b$, an overestimated background could drive the sensitivity negative, which is an indication that either the experiment or the background estimation is in error.

The largest impacts of underestimate or overestimate of the background may be their effects on the discovery potential. The former could lead to false discovery and the latter may miss a potential discovery by treating it as background. The low event rate makes it extremely difficult to characterize the background precisely. The best solution is to reduce the background as much as possible and, ideally, run the experiment at the background free regime.

There are two types of background events that could be misidentified as dark matter signals, electron recoils and nuclear recoils. Electron recoils can be from gammas and betas. Ambient gammas and betas can be effectively shielded by high-Z materials such as Lead. Gammas and betas from the radioactive contaminants in the materials around the detector and the detector itself must be reduced as low as possible by using radiopure materials. However, the residual electron recoil background, which is predominantly caused by the residual gammas, is usually not a big concern since most direct dark matter detection experiments have strong discrimination against it. These experiments measure both the heat and charge (Ge), or heat and light ($CaWO_4$), or light and charge (Xe, charge by secondary scintillation) of a scattering event. Electron recoils can be identified with high probability because of their different signatures from that of the nuclear recoils. Some experiments (C_3FI) are intrinsically insensitive to electron recoils and they only see the nuclear recoils. Though electron recoils do not constitute a major part of the background, the radioactive contamination level is still estimated by Monte Carlo simulations of the measured gamma spectrum, which provides an independent handle on the radiopurity of the materials used.

Nuclear recoils are the dominant and most dangerous background to WIMP search.

They can be from neutrons, α-particles, and recoiling heavy daughter nuclei in α decays. α's and the heavy nuclei can not penetrate more than a few microns of the shielding material. They can only be from the α decays of the radio contaminants in the detector or on the surfaces of the surrounding materials exposed to the detector. The recoiling α's from the U and Th chains typically have a few MeV kinetic energy, which is way above the recoil energy range of a WIMP. However, the recoiling heavy daughter nuclei in α decays have kinetic energy $\sim$100 keV, which can mimic a WIMP if the accompanied α particle escapes the detection. The neutron background is the most dominant one among all. This includes cosmogenic neutrons and radiogenic neutrons. The cosmogenic neutron flux depends on the muon flux and the surrounding rock environment of the apparatus. At the Soudan Underground Lab, which is 714 m (2090 meters water equivalent) below the surface, the cosmogenic neutron background is comparable to that of the radiogenic neutron in the CDMS II experiment. At deeper sites, the radiogenic neurons are expected to be the dominant background. Low energy cosmogenic neutrons can be moderated effectively by hydrogen rich materials such as water or polyethylene, and those accompanied by muons or secondary showers can be tagged by active vetos with high efficiency. Radiogenic neutrons produced in spontaneous fissions and $(\alpha,\ n)$ reactions in the U and Th chains around the detector can only be lowered by using radiopure materials. Compared to WIMPs, neutrons have much larger cross sections interacting with nuclei and they tend to scatter multiple times within the detector, which provides another handle to estimate the neutron backgrounds. While the electron recoil background can be characterized by taking gamma calibration data, the neutron background can only be estimated by Monte Carlo simulations.

In summary, a direct WIMP search experiment is preferably running at the background free regime rather than dealing with various backgrounds. The dominant background for the coming generation of direct search experiments is expected to be radiogenic neutrons. Acquisition and characterization of radiopure materials are therefore crucial for the experiment design.

1.5.2 Indirect detection

WIMPs ceased annihilation in the early universe when their density went too low to maintain the reaction. However, after gravitational collapse, the WIMP density at the center of the dark matter halos may be high enough to resume the annihilation. In this case, WIMPs may be indirectly detected by observing their annihilation products, including neutrinos, gamma rays, positrons, and anti-protons.

The dark matter density is expected to be high at the galactic centers and the core of the Sun. Besides the greater gravitational potential they feel, WIMPs also have a higher total scattering rate in such locations, which would slow down, capture, and trap a fair amount of dark matter. The difficulty of detecting WIMP annihilation signals is to discriminate them from the various astrophysical backgrounds. Nearby dark matter halos are preferred, since they suffer less from solid angle suppression. The closest galactic halo core is expected to be in the center of our Milky Way. Due to its high background and lack of thorough understanding of the astrophysical activities, the Milky Way center is less favored as a signal source for the experiments dedicated for dark matter search. The dwarf galaxies in the Milky Way are promising sources in this regard due to their dark matter domination and simpler astrophysical activities. The galaxy clusters are also believed to be dark matter dominated and are

interesting signal sources. Simulations show that the dark halos at the cluster scale have substructures. The sub-halos bring in uncertainties in estimation of the WIMP density and thus their self-annihilation cross section. The dark matter density at the core of the Sun, which is expected to be dependent on the capture process, is model dependent.

Unlike in the direct dark matter search experiments, which rely on the WIMP-nucleus elastic scattering, the potential signals, as well as the annihilation modes and their branching ratios, for the indirect searches are model dependent. This complicates the interpretation and comparison of results. Regardless, within a specific model, all the processes can be calculated in detail. Another challenge for the indirect searches is the astrophysical backgrounds, which are usually not well understood or well characterized. Certain assumptions about the backgrounds have to be made in the data analysis.

1.5.2.1 Gamma rays

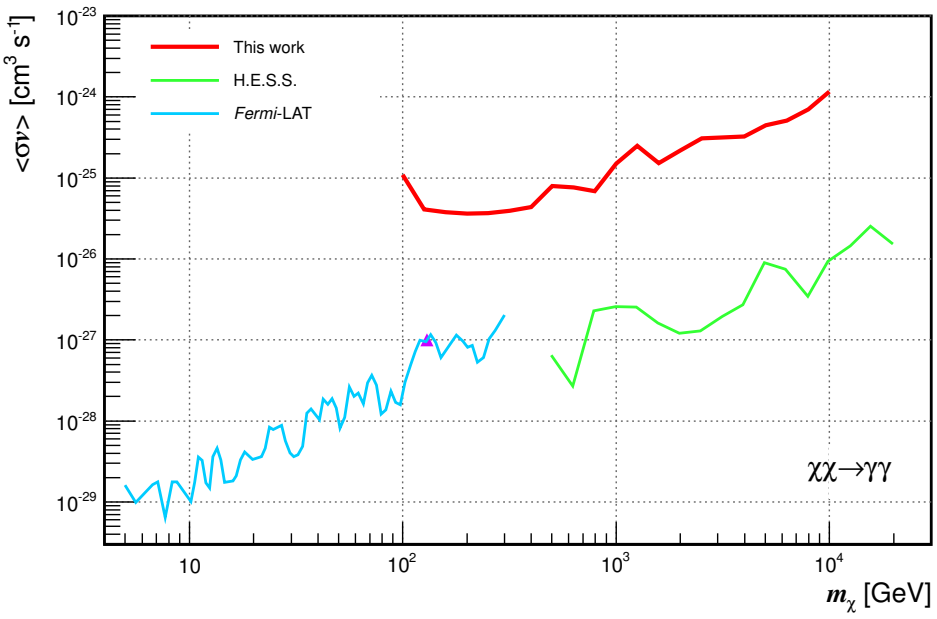
Gamma rays as a dark matter tracer have the advantages of good directivity and (relatively) easy detectability. Monoenergetic gamma lines from the WIMP self-annihilation channels with two-body final states $\gamma\gamma$ and γZ are the most convincing signatures in gamma ray detection. With $m_\chi \gg m_Z$, the energy of the resulting photon is close (or equal for $\gamma\gamma$) to the the WIMP mass $E_\gamma \simeq m_\chi$. However, since WIMP does not interact with photons directly by definition, the predicted branching ratios are very low due to loop suppression. The photons that are not produced in the two-body final states result in a continuum spectrum, which can also be utilized but is more susceptible to astrophysical backgrounds.

There are a few major gamma ray detection experiments currently in operation: Fermi/LAT, VERITAS, H.E.S.S., and MAGIC. Fermi/LAT is a satellite based experiment utilizing pair production, silicon particle tracker, and CsI calorimeter to capture and measure gammas. The latter three are all ground based imaging atmospheric Cherenkov telescopes (IACTs) which measure the Cherenkov radiation produced when gammas traverse the atmosphere. The sensitive energy range of Fermi/LAT is 30 MeV to 300 GeV and the typical energy range for the IACTs is ∼100 GeV to ∼30 TeV. Accordingly, they are sensitive to WIMPs with mass in the same range, which are complementary to the terrestrial experiments.

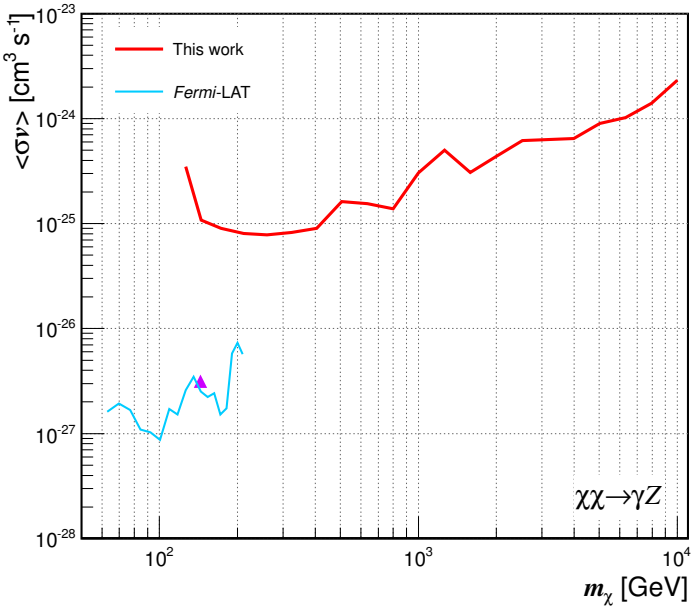
No significant gamma ray emission above the astrophysical background has been detected so far, instead, upper limits of the WIMP annihilation cross section were reported by these experiments. Figure 1.15 shows the recent gamma line search result reported by MAGIC [52] together with those from H.E.S.S. and Fermi/LAT. The reported 130 GeV gamma line with $4.6\,\sigma$ local significance and $3.2\,\sigma$ global significance from the vicinity of the Galactic center [51] is also shown in purple triangle. These experiments are approaching or are starting to constrain the canonical WIMP annihilation cross section $\langle \sigma v \rangle = 3 \times 10^{-26}\,\mathrm{cm}^3\,\mathrm{s}^{-1}$. With more data taken, the tentative signal may be conformed or some theoretical models can be ruled out. More about these results can be found in Refs [51–57].

1.5.2.2 Neutrinos

Neutrino is another leading dark matter tracer. WIMPs have a finite probability to scatter with ordinary matter and lose their energy. By repeating this process, they can be captured and trapped in the core of the Sun and the Earth. Equilibrium between



(a)



(b)

Figure 1.15: Gamma line search results from MAGIC, H.E.S.S., and Fermi/LAT. The purple triangle is the tentative 130 GeV gamma line from Weniger [51]. "This work" means MAGIC in the figure. *Figure from Aleksić* et al. *[52].*

capture and annihilation can be established in a time scale shorter than the age of the solar system. The annihilation rate is eventually determined by the capture rate. Energetic neutrinos produced in WIMP annihilation can penetrate the Sun and the Earth and thus be detected by terrestrial detectors. Typical energy of the neutrinos produced in WIMP annihilation is 1/3 to 1/2 of the WIMP mass, which is much higher than that of the solar neutrinos (MeV scale). Directional information can distinguish the WIMP annihilation signal from the atmospheric neutrino background. WIMP annihilation neutrinos from the Sun or the Earth, if present, are a distinctive signal that can be clearly identified. More detailed discussions can be found in [7]. Clearly, detectors constructed for detecting signals from the Sun or the Earth can also be used to target other sources like galaxies or clusters.

The IceCube experiment has performed such searches targeting the Sun, and nearby galaxies and clusters. Measurement of the neutrinos from the Sun resulted in more stringent spin-dependent WIMP-proton cross section limits than direct searches above 35 GeV [59]. Figure 1.16 shows their recent dark matter limits on galaxy and cluster searches for annihilation modes $\chi\chi \rightarrow \mu^+\mu^-$ and $\chi\chi \rightarrow \tau^+\tau^-$ in solid lines [58]. Also plotted in dotted and dashed lines are the limits set by the gamma ray detection experiments. The green and yellow regions are the PAMELA and Fermi/LAT positron excess interpreted as dark matter signal, respectively. Though still not reaching the canonical annihilation cross section, shown as the gray band in Fig. 1.16, these searches are starting to constrain the signal interpretations from peer experiments. IceCube can detect neutrinos between 100 GeV and a few PeV, which covers a fairly large range of possible WIMP mass, providing complementary information to other dark matter detection techniques.

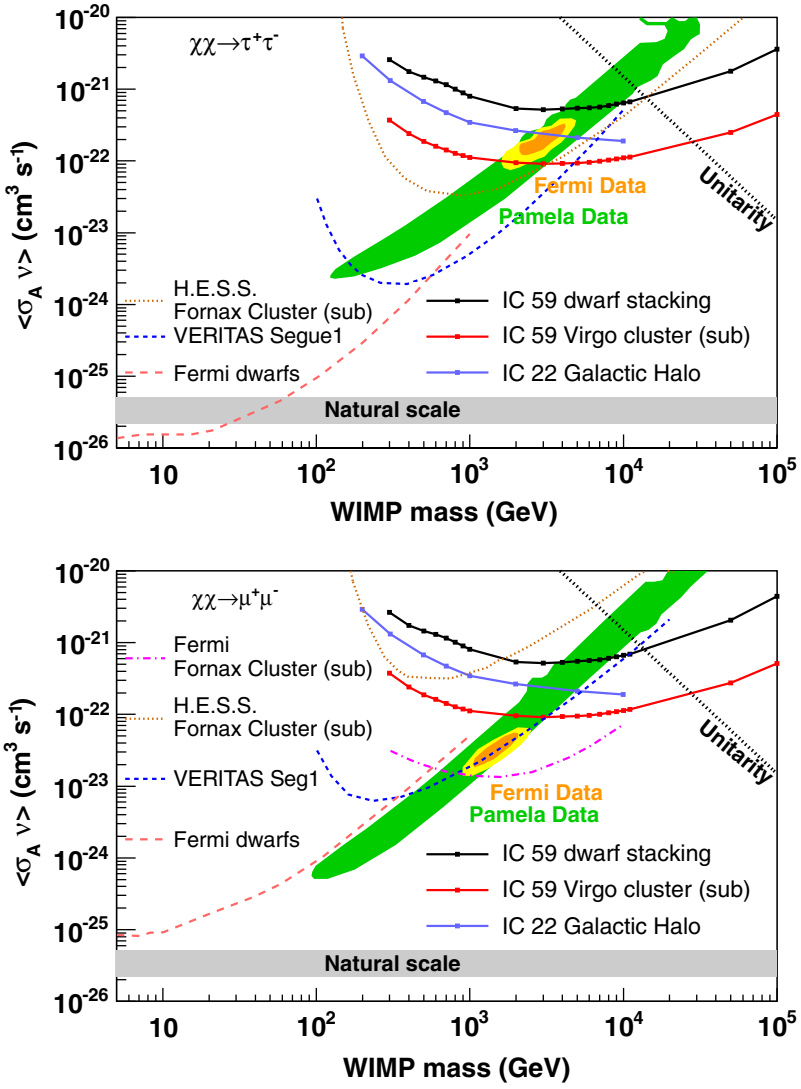


Figure 1.16: WIMP limits set by IceCube in the $\tau^+\tau^-$ and $\mu^+\mu^-$ annihilation channels (solid lines). Dotted and dashed lines show the limits from the gamma ray detection experiments. The green and yellow regions are the positron excess in cosmic rays interpreted as WIMP signals for PAMELA and Fermi/LAT, respectively. The gray band shows the canonical cross section. Unitarity limits the cross section at high mass. *Figure from Aartsen* et al. *[58].*

1.5.2.3 Antimatter

Antimatter produced in WIMP annihilation, mainly positrons, antiprotons, or even anti-nuclei, can, though difficult, be used to trace dark matter. The majority of cosmic rays are protons, and the rest are primarily heavier nuclei. The fraction of antiproton is very small. The flux of antiprotons in cosmic rays is expected to fall drastically at energies below $\sim$1 GeV [7]. In contrast, the flux of the low energy antiprotons produced in WIMP annihilation should not have such decrease. Observation of excess of low energy antiprotons would help to identify dark matter. The situation for positions is more challenging because of the large uncertainties in understanding of the background. Positrons produced in WIMP annihilation have a broad energy distribution, which would be difficult to distinguish from the background.

Figure 1.17 shows the recent measurements of the positron fraction in cosmic rays. The positron fraction rise above 10 GeV was interpreted as possible dark matter signals [60]. However, this is already in tension with the constraints set by gamma ray and neutrino detections (see Fig. 1.16), which raises the question whether positrons are still a viable dark matter tracer [62]. As for antiprotons, it was shown by Cirelli and Giesen [62] that there is possibility for AMS-02 to reach the canonical thermal cross section but large uncertainties exist in the calculation. Nevertheless, it is difficult for antimatter alone to identify dark matter.

1.5.3 Collider searches

Dark matter may be produced in high-energy particle collisions if their mass is not too high and their couplings to the SM matter is sufficiently strong. If produced, they

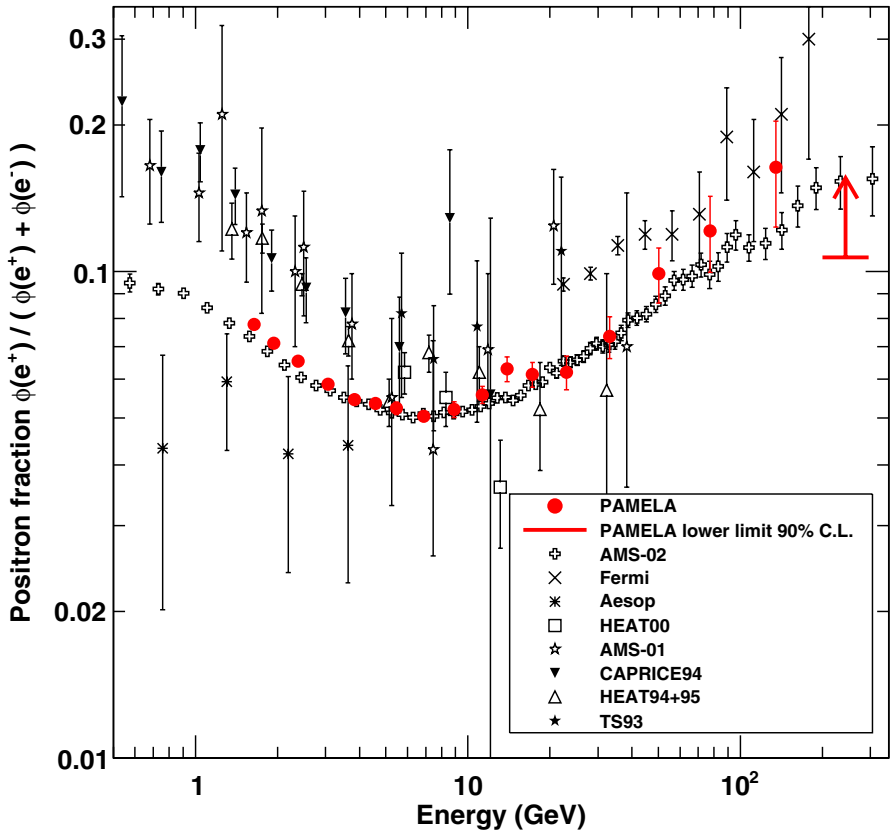


Figure 1.17: Cosmic ray positron fraction vs. energy. PAMELA, Fermi, AMS-01, and AMS-02 are space-borne experiments. The positron fraction rise above 10 GeV was interpreted as possible dark matter signal [60] (also see Fig. 1.16), but alternative explanation exists. *Figure from Adriani* et al. *[61].*

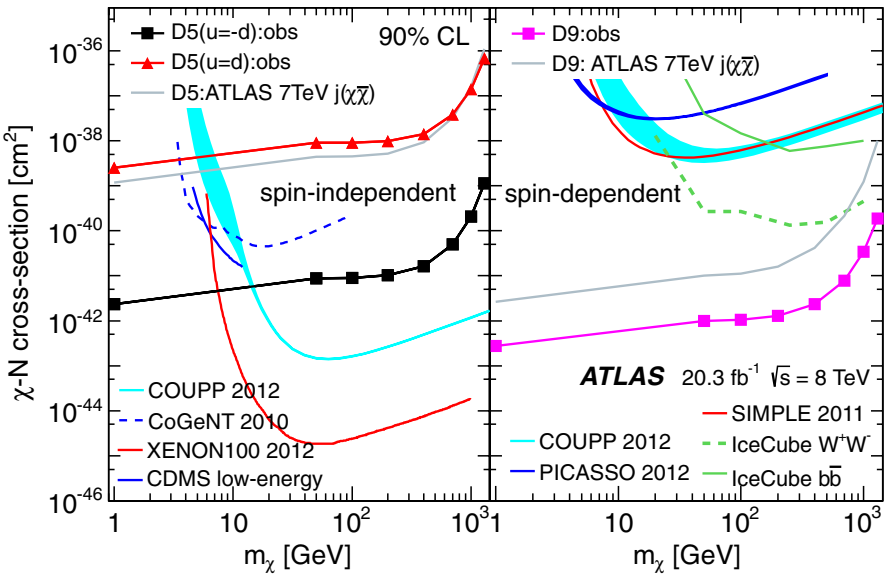


Figure 1.18: Spin-independent and spin-dependent WIMP-nucleon cross section limits of the ATLAS dark matter search. Limits were calculated in the effective field theory framework with WIMP production modes $pp \to \chi\bar{\chi}W$ and $pp \to \chi\bar{\chi}Z$. Black squares are the limit assuming u quark and d quark couplings to WIMPs have different signs. The limit is much lower in this case because of enhanced W production. *Figure from Aad* et al. *[63]*.

are likely to escape the direct detection and result in missing momentum and energy in the collision event. While the direct and indirect searches probe the cross section downward in a certain WIMP mass range, the collider searches work orthogonally, probing all masses up to its energy limit with the cross section floor determined by the integrated luminosity. Because light WIMPs are easier to produce at a given colliding energy, collider searches tend to be more sensitive to low-mass WIMPs.

Collider searches have all been consistent with the SM predictions so far. The recent searches on the LHC conducted by the CMS [64, 65] and the ATLAS [63, 66, 67] collaborations focused on the production mode $pp \to \chi\bar{\chi} + X$, where X is a hadronic jet or a photon. Analyses showed that monojet events tend to give stronger limits [63].

The WIMP-nucleon cross sections were calculated in the framework of an effective field theory [68–70], where WIMP-quark and WIMP-gluon interactions are four-point contact interactions mediated by a heavy particle M_*. The coupling strengths between u and d quarks to WIMPs are also assumed to be equal.

Figure 1.18 shows the recent results from a dark matter search with the ATLAS detector [63]. Events with a W or a Z boson (subsequently decays into quark pairs) and missing transverse momentum in pp collisions at $\sqrt{s} = 8\,\text{TeV}$ were selected in the analysis. The strongest limits for the spin-dependent WIMP-nucleon cross section (purple squares) at WIMP mass from 1 GeV to 1 TeV and the spin-independent cross section (red triangles) at WIMP mass less than 5 GeV were set. The black squares are the limit assuming the interactions of u and d quarks to WIMPs have different signs, which is three orders of magnitude lower than the $u = d$ case due to enhanced W production. It is apparent that collider searches have already provided strong constraints in the WIMP parameter space. However, it is less clear whether the effective field theory is still a good approximation at large momentum transfer [71]. Due to the strong model dependence of the collider search limits, an accurate comparison to the direct searches may need more detailed studies.

1.6 Direct dark matter search experiments

1.6.1 The current experiments

The direct dark matter search experiments are now really diverse compared with the early days. Generally speaking, they fall under the following categories by different criteria:

By type of dark matter signal: The direct search experiments either discriminate the background on an event by event basis or by the statistical signature of the dark matter signal. In the latter case, an experiment can measure the event rate annual modulation, WIMP direction diurnal modulation, or the energy spectrum of the excess events.

By target material/technology: Commonly used targets include cryogenic detectors (Ge, Si), noble liquids (Xe, Ar), scintillators (NaI, $CaWO_4$), and hot droplets/bubble chambers (CF_3I, C_3F_8).

By readout signals: The event by event discrimination experiments typically measure a combination of the effects produced in WIMP-nucleus scattering: charge, heat, and light. A primary discrimination parameter is defined usually as the ratio of two measured quantities. The event location and timing (or pulse shape) information, if available, are also recorded and utilized to improve the discrimination power. The experiments targeting the statistical characteristic of the WIMP signal typically measure only one effect to simplify the experiment design. However, the former class of experiments can also perform single parameter analysis at the cost of reduced discrimination power.

The classification above is not comprehensive or rigorous, however, for the the majority of the direct search experiments, it is usually an adequate description. Below we briefly review the major direct dark matter search experiments.

DAMA/LIBRA: The DAMA/LIBRA experiment measures the WIMP event rate annual modulation with ∼250 kg highly radiopure NaI(Tl) at the Laboratori Nazionali del Gran Sasso in Italy. Only the scintillation light from the crystal is measured and the experiment can not discriminate nuclear recoils from electron recoils. Recently

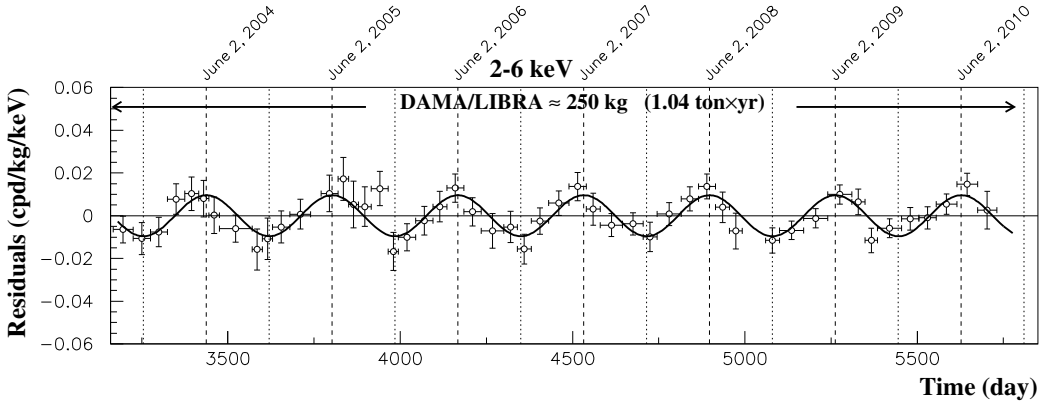


Figure 1.19: WIMP rate annual modulation signal from the 7-year DAMA/LIBRA–phase1 data in the energy range 2-6 keV. *Figure from Bernabei* et al. *[72].*

they published the 7-year (2003-2010) DAMA/LIBRA–phase1 result with exposure 1.04 ton $\cdot$ yr [72]. In the energy range 2-6 keV, the DAMA/LIBRA–phase1 data gave a modulation signal with the expected phase and period at 7.5 σ CL (Fig. 1.19), the significance increased to 9.3 σ with the 1.33 ton $\cdot$ yr DAMA/NaI data included [72]. Though the dark matter interpretation of the modulation signal, first reported in 1998 [73], is in tension with a number of other constraints, no alternative physical sources have been found [72]. DAMA/LIBRA–phase2 started from the end of 2010 with improved sensitivity.

CoGeNT, C-4: CoGeNT operates ∼400 g p-type point-contact germanium (PPC) detectors at the Soudan Underground Laboratory aiming to search for the event rate annual modulation of low-mass ($m_\chi < 10$ GeV) WIMPs. The PPC detector measures only the ionization. The point contact reduces the capacitance of the ionization readout, enabling a low threshold of 0.5 keV (for electron recoils). A nuclear recoil (assuming a quenching factor of 1/3) from a 10 GeV WIMP above the threshold requires a minimum WIMP speed 252 km s^{-1}, larger than the characteristic speed

$v_0 = 220\,\mathrm{km\,s^{-1}}$. The experiment is sensitive to the high velocity tail of the WIMP speed distribution but will miss the peak. The C-4 (CoGeNT-4) experiment is a scaled-up version of CoGeNT with ×12 target mass and reduced energy threshold. CoGeNT reported a modulation signal at $2.8\,\sigma$ significance level compatible with dark halo origin in 2011 [74]. In their more recent result [75], the signal was confirmed again at $2.2\,\sigma$, but the modulation amplitude was ∼4–7 larger than predicted with dark halo origin.

EDELWEISS-II: EDELWEISS is an event by event discrimination experiment utilizing cryogenic Ge detectors operated at temperatures ∼20 mK. The discrimination parameter is the ionization yield defined as the ratio of the ionization energy to the recoil energy. Unlike CDMS, they measure the temperature of the whole detector substrate (thermal phonons) with Ge neutron transmutation doping (NTD) sensors. To better reject surface events, they developed concentric ring electrodes on the top and bottom detector surfaces symmetrically, biased at alternating voltages, typically 6 V to 8 V across the detector [76]. Ring electrodes are also deposited on the cylindrical side walls to increase the uniformity of the axial fields. In the previous phase during 2009-2010 [77, 78], 10×400 g target mass was deployed. In a status report in 2013 [79], 40×800 g detectors with improved performance were being installed. The experiment is located at Laboratoire Souterrain de Modane (France).

CRESST-II: The CRESST experiment, located at the Gran Sasso underground facility, operates 18 (up to 33) $CaWO_4$ detector modules (∼300 g each) running at ∼10 mK to search for the WIMP signal. Each detector module is instrumented with two tungsten superconducting transition sensors (TES), one for measuring the temperature of the scintillating $CaWO_4$ crystal (thermal phonons) and the other for

monitoring the temperature of the Si light absorber. Both temperature signals are read out with SQUID electronics, and the ratio of scintillation light to phonon energy is used as the discrimination parameter. In their most recent result published in 2012 [80], 67 events were found in the signal region. A maximum likelihood analysis disfavored these events being caused by background at significance level above 4σ. Two possible WIMP signal regions were reported (Fig. 1.20).

Noble liquids: This category includes a large number of experiments, notably LUX (370 kg Xe), XENON100 (160 kg Xe), ZEPLIN-III (12 kg Xe), and DarkSide-50 (50 kg Ar). The noble liquid detectors are typically in the form of dual-phase time projection chambers (TPC), with noble liquids sitting at the bottom and vapor at the top. Particle interactions with the liquid excite and ionize the target atoms. The primary scintillation (S_1) produced in the liquid can be detected by the photomultiplier tubes (PMT), the ionized electrons are drifted out of the liquid by the voltage applied across the detector and then produce the secondary scintillation (S_2) in the gas. The ratio between S_1 and S_2 can be used to discriminate nuclear recoils from the backgrounds. The time delay between S_1 and S_2 gives the location information of the event, which can be used to define the fiducial volume. Generally, ionizing radiation in noble liquids lead to the formation of dimers in singlet or triplet state [81], with characteristic decay times τ_1 and τ_3, respectively. Nuclear recoils and electron recoils generate different mixes of the two states, leading to different S_1 pulse shapes. This is a powerful discriminator for Ar because of the large difference between τ_1 and τ_3 ($\tau_1 = 6\,\text{ns}$, $\tau_3 = 1.59\,\mu\text{s}$), but substantially difficult for Xe ($\tau_1 = 2.2\,\text{ns}$, $\tau_3 = 21\,\text{ns}$).

The background control is relatively easier for Xe since it does not have long-lived radioactive isotopes. The major radio contaminants ^{85}Kr can be filtered out by

distillation. For Ar, the radioactive ^{39}Ar is relatively long lived ($\tau_{1/2} = 269\,\text{yr}$), which could saturate the trigger in a large scale experiment. The discovery of underground Ar with low ^{39}Ar alleviated this drawback somewhat. One big advantage of Xe is its self-shielding against backgrounds because of its high mass number. A few centimeters into the inner region from the surface is relatively background free [82]. This greatly simplifies shielding of the experiment as well as data analysis. The scintillation of Xe (175 nm) can be detected directly by PMTs while Ar (128 nm) needs wavelength shifters.

Though the few leading noble liquid experiments are all dual-phase (liquid and gas), single phase (liquid only) experiments do exist for their simpler design and better scintillation collection efficiency. They rely on shielding (e.g., Xe) and pulse shape discrimination (e.g., Ar) to suppress the backgrounds. Noble liquid detectors are highly scalable, they are currently leading the direct dark matter search limits, particularly for high-mass WIMPs. The recent limit from LUX [82], also shown in Fig. 1.20, currently holds the leading position.

Superheated liquids: Superheated liquids are liquids heated above their boiling point without boiling because of lack of bubble nucleation in the liquid. Liquids start boiling when bubbles inside them grow infinitely. Due to the existence of surface tension, the pressure difference between inside and outside of the bubble is inversely proportional to the radius of the bubble, $\Delta p = \frac{2\gamma}{R_b}$, where γ is the surface tension and R_b is the bubble radius. A small bubble would need to overcome the enormous differential pressure to grow. However, if there is sufficient energy supply to the bubble, the growth can be achieved and the liquid will start boiling. Recoiling nuclei and electrons can provide this energy. Since recoiling nuclei typically lose

more energy than recoiling electrons in the same track length, a superheated-liquid detector is intrinsically insensitive to gamma or muon backgrounds. The desired background rejection power against electron recoils can be reached by tuning the detector threshold through varying fluid pressure and temperature. Neutrons, α-particles, and recoiling heavy nuclei, however, can show up as backgrounds. The acoustic signals during bubble formation were reported to have good discrimination power to distinguish neutrons and other recoils.

Superheated-liquid detectors are threshold based, the energy of an event can not be directly obtained. Mapping out the energy spectrum requires running the detector at a series of thresholds. The target liquid can be easily changed, making the detector quite versatile in probing the WIMP properties. Due to the increased interest in low-mass WIMPs recently, target liquids such as C_3F_8 and C_4F_{10} gained more attention than the earlier CF_3I. These liquids are also powerful in probing the spin-dependent cross section owning to the large nuclear spin of fluorine (^{19}F).

The experiments COUPP, PICASSO, and SIMPLE are pursuing this route.

1.6.2 The current status and future

Figure 1.20 shows the WIMP parameter space constrained by the current direct DM search experiments and the projected sensitivities for the future experiments for the spin-independent interaction. The results from DAMA, CoGeNT, CRESST, and CDMS Si showed excesses of events at WIMP mass $\sim$10 GeV and $\sigma^{\rm SI}_{\chi p} \sim 10^{-41}{\rm cm}^2$. These signal regions, however, are ruled out by the CDMS Ge, EDELWEISS, COUPP, ZEPLIN-III, and Xenon100 results either partially or completely. This controversy is likely to be resolved by the future generation direct DM search experiments since

the region is well within the projected science reaches. Theoretical models such as asymmetric DM and Magnetic DM predicted the WIMP signal region to be $m_\chi \sim 8\,\mathrm{GeV}$ and $\sigma^{\mathrm{SI}}_{\chi p} \sim 10^{-42}\mathrm{cm}^2$. The current experiments have already excluded some of the predicted signal regions. Completely constraining these models will need the next generation direct DM search experiments. Ge based experiments such as SuperCDMS SNOLAB have strong potential at the low-mass region owing to the low detector threshold and relatively heavy target nuclei. The neutrino background at $m_\chi \leq 10\,\mathrm{GeV}$ is $\sim 10^{-44}\mathrm{cm}^2$. The future generation experiments will hit this background.

At high-mass region, the current experiments are probing $\sigma^{\mathrm{SI}}_{\chi p} \sim 10^{-45}\mathrm{cm}^2$ at their most sensitive WIMP mass $\sim 50\,\mathrm{GeV}$. Generally, WIMPs with mass less than a few hundred GeV are less favored by the constraints from the collider and indirect DM searches. The SUSY MSSM and Extra Dimension models predicted WIMPs with mass $\gtrsim 1\,\mathrm{TeV}$. The current experiments are approaching the predicted signal regions but are still far away from the most probable parameter space. The next generation direct DM search experiments will greatly if not completely constrain these theoretical models. The neutrino background is less constraining to the direct DM search experiments at the high-mass region, which is likely not to limit the next generation experiments. Xe based experiments are competitive in the high-mass region because of the heavy target nuclei.

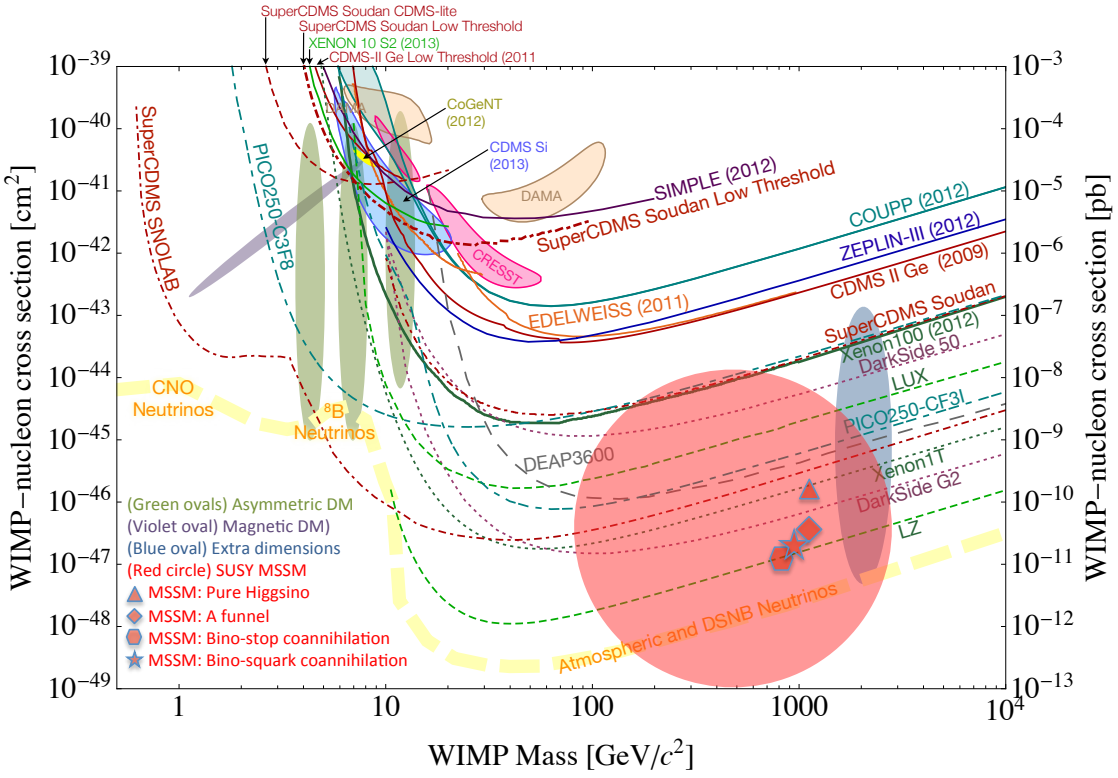


Figure 1.20: Measured and projected spin-independent WIMP-nucleon cross section limits from direct dark matter searches. The results from DAMA, CoGeNT, CRESST, and CDMS Si show excesses of events at WIMP mass ∼10 GeV. The same region is excluded by CDMS Ge, EDELWEISS, COUPP, ZEPLIN-III, and Xenon100 at varying degrees. Low-mass WIMP region predicted by the Asymmetric DM and Magnetic DM models are partially excluded by the current experiments. Heavy WIMPs predicted in SUSY MSSM and Extra Dimension models have mass ∼1 TeV. The current experiments are starting to constrain the parameter space for the heavy WIMPs. *Figure from Galbiati [83].*

Chapter 2

The CDMS Experiment

Cryogenic Dark Matter Search (CDMS) is a pioneer and had been a long time leading experiment in direct dark matter search. At the heart of the experiment are the cryogenic Ge and Si detectors, interacting with the hypothesized WIMPs, converting the kinetic energy transferred to the detector to charges and heat, and collecting the ionization and phonon energies. Nuclear recoils and electron recoils are identified by their different ratios of ionization energy to recoil energy, i.e., the ionization yield.

Conceptually simple, however, it is quite a challenge experimentally. Take the phonon signal as an example, if the kinetic energy transferred to the Ge crystal is 10 keV in an collision event, the temperature rise will be 5 μK for one mole of Ge (73 g, Debye temperature taken as 374 K) at an operation temperature 20 mK. Such a small temperature rise would be extremely difficult to measure even though theoretically possible. Fortunately, the propagation of the phonons is ballistic rather than diffusive for the most part, so that a fraction of the phonons can be collected by the phonon bolometers without being dissipated by the crystal. Furthermore, the

phonon arrival time delays between different channels provide the depth and lateral position information about the event, which turned out to be essential to discriminate surface electron recoils. These events can mimic nuclear recoils because of their reduced ionization yields.

The evolution of the series of CDMS experiments is driven by the desire to reach lower backgrounds, higher background discrimination power, and larger detector mass. To reduce the cosmic muon flux, starting from CDMS-II, the experimental site transitioned from the Stanford Underground Facility (SUF) to the deeper Soudan Underground Laboratory in northern Minnesota. The proposed SuperCDMS SNOLAB, if eventually happens, will locate the experiment in the even deeper SNOLAB underground facility in Canada. The detector designs too have gone through several rounds of evolution. The biggest leap is probably the transition from thermal phonon measurement with NTD thermistors to athermal phonon detection using superconducting transition edge sensors (TES). A few generations have been developed in the latter category, each with improved surface event rejection power and better scalability.

This chapter gives a brief overview of the CDMS experiment with the focus on CDMS-II, during which the data of this analysis was taken. More detailed descriptions can be found in [84–86]. The detector development for the SuperCDMS SNOLAB is discussed in Chapter 4.

2.1 ZIP detectors

Measurement of the arrival time delay between the ionization signal and the athermal phonon signal enables the determination of the depth of an event, thus the detectors

were given the name Z-sensitive Ionization and Phonon (ZIP) detectors. A few generations of ZIP detectors have been developed and deployed in science runs since their appearance. The oldest ones are the CDMS-II type detectors (hence the name oZIP) as shown in Fig. 2.1(a) and (b). The oZIPs are fabricated from cylindrical Ge or Si substrates of diameter 76 mm and thickness 10 mm. The masses of the Ge and Si oZIPs are 250 g and 100 g, respectively. Four phonon sensors each covering equal area are lithographically patterned on the top surface of the substrate (Fig. 2.1(b)). The start time differences and the energy partition of the phonon signals can determine the lateral position of an event. Two concentric charge electrodes are patterned at the bottom surface with the inner electrode covering ∼85 % of the bottom face. The outer electrode is used to veto events which deposit a large fraction of its energy on the outer ring. The Si detectors were originally used to identify the neutron backgrounds in the SUF shallow site, which is not necessary for the deeper Soudan site. However, the lighter Si detectors can be used to detect light WIMPs.

It was found in the CDMS-II data analysis that the event radii constructed from the phonon signal delays or energy partitions had degeneracies at large radii, which degrade the phonon energy resolution and electron recoil rejection power [87]. In a modified version of the ZIP, which is referred to as mZIP, the four-quadrant phonon sensor layout was replaced by the one shown in Fig. 2.1(d). The outer ring phonon channel is used to lift the radial degeneracy. In addition, the thickness of mZIP was increased to 25 mm. However, the surface electron recoil rejection power of the mZIP was not as good as expected.

To better reject surface events, interleaved ZIPs (iZIP) were then designed to increase the asymmetry between the surface events and the bulk events. As shown in

Fig. 2.1(c) and (d), ionization and phonon sensors are patterned on both the top and bottom surfaces. The charge and phonon rails in the same plane are interleaved and biased at different voltages to further increase the asymmetry. The fourth version (Fig. 2.1(d)) iZIP design was found to have the performance to support an experiment at ∼100 kg scale. More information about mZIPs and iZIPs can be found in the CDMS theses [88–90].

2.1.1 Ionization signal

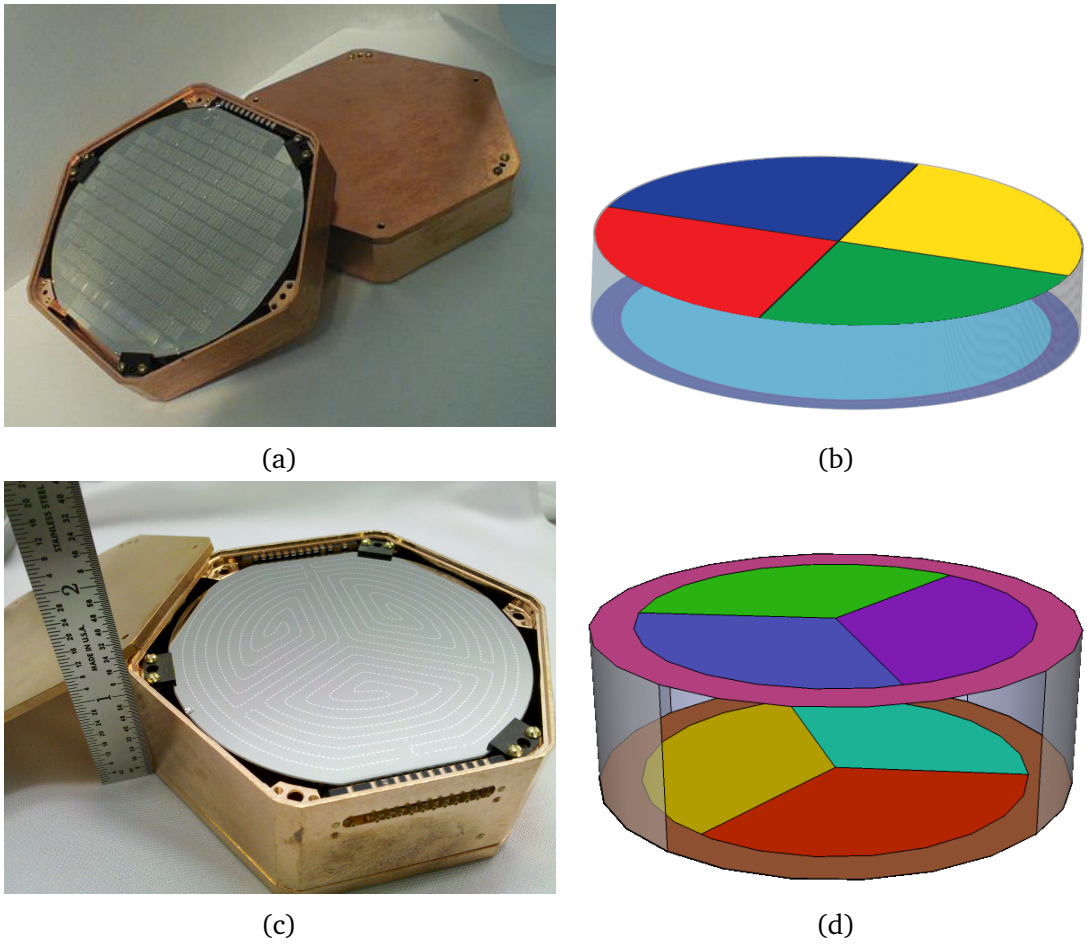
2.1.1.1 Charge generation

Energetic photons, if with energy above the band gap of the detector substrate ($E_g = 0.78$ eV for Ge and 1.17 eV for Si at millikelvin temperature), can be absorbed and create electron-hole pairs in the crystal. The number of charge pairs created is given by

$$N_Q = \frac{E_r}{\epsilon}, \tag{2.1}$$

where E_r is the recoil energy and ϵ is the average energy needed to create an electron-hole pair. For Ge, $\epsilon = 3.0$ eV and for Si, $\epsilon = 3.8$ eV [84, 86, 91]. More than ∼70 % of the recoil energy is converted into optical phonons during charge pair creation [89, 92]. It is much easier for a recoiling nucleus to lose energy in the form of phonons than a recoiling electron, thus less efficient in creating charge pairs. The reduction of the efficiency is usually referred to as the "quenching factor".

A small bias, usually −3 V for Ge and −4 V for Si, is applied to the charge electrodes while keeping the phonon sensors at 0 V. Free electrons and holes generated in the collision drift in opposite directions in the electric field, and image charges are induced



(a) (b)

(c) (d)

Figure 2.1: CDMS ZIP detectors and electrode layout. (a) Image of a CDMS-II ZIP. (b) CDMS-II ZIP electrode layout with phonon sensors on the top surface and charge electrodes at the bottom. The inner charge electrode covers ∼85 % of the bottom surface. (c) Image of a SuperCDMS Soudan iZIP (Version 4). (d) iZIP4 sensor layout. Each solid color region represents a phonon sensor covering equal area. Phonon sensor rails and charge electrode rails are interleaved. The inner charge electrode on each face covers the same area as the three inner phonon sensors. The outer electrodes are located at the two outer rings.

on the electrodes according to the Ramo theorem [93]:

$$Q = -q\varphi_0(\mathbf{x}), \tag{2.2}$$

where Q is the induced charge, q is the charge of the drifting carrier, and $\varphi_0(\mathbf{x})$ is the weighting potential at position $\mathbf{x}$, i.e., the potential generated by the electrode under consideration with this electrode at unity potential while keeping all the other electrodes at zero potential and all charges removed. It is easy to see the total charge induced by an electron-hole pair which arrive at their target electrodes at the end of their drift:

$$\begin{aligned}\Delta Q &= e[\varphi_0(\mathbf{x}_f^e) - \varphi_0(\mathbf{x}_i)] - e[\varphi_0(\mathbf{x}_f^h) - \varphi_0(\mathbf{x}_i)] \\ &= e[\varphi_0(\mathbf{x}_f^e) - \varphi_0(\mathbf{x}_f^h)] \\ &= e,\end{aligned} \tag{2.3}$$

where we have used the assumption that $\varphi_0(\mathbf{x}_f^e) = 1$ and $\varphi_0(\mathbf{x}_f^h) = 0$. It is interesting to see that the total induced charge on an electrode is equal to the charge of the carrier drifting to the electrode with opposite sign. In a non-rigorous sense, this is often referred to as "collecting electrons" or "collecting holes" depending on the polarity of the charge bias. However, it is really "collecting" image charges with the opposite sign.

In Eq. (2.3), we made the assumption that both the electrons and holes successfully drift to the ends of the detector, which is not always true. Some of the electrons or holes are trapped by the neutral or charged impurities in the crystal. Accordingly,

the collected charge will be reduced. It is crucial to keep the detectors neutral to reduce the trapping effect after running the detector for some time under bias. We use infrared LEDs to flash the detectors while keeping the electrodes grounded. The electron-hole pairs created by the infrared light neutralize the charged trapping centers and recover the performance of the detector. The energy of the infrared photons is peaked at 1.31 eV (940 nm), which is barely higher than the absorption edge of Si but much higher than that of Ge [92]. The absorption of Ge only allows the photons to penetrate a thin layer. The bulk of the crystal can only be neutralized by carriers propagating into the middle. Considering the fast rise time of the charge signal, the short penetration depth may not be a limiting factor to neutralization.

If the electron-hole pairs are created close to the surface of the detector, these hot carriers may propagate against the electric force and be collected by the "wrong" electrode, thus reducing the ionization signal. It was found depositing a thin layer of amorphous silicon can reduce the back diffusion of the carriers because of the large band gap of Si.

2.1.1.2 Ionization signal readout

Figure 2.2 shows the schematic of the preamplifier for the charge readout. The detector is denoted by the capacitor C_d. Similar to what happens in the detector, the induced charges on the charge electrodes further induce image charges in the capacitor network. Eventually the same amount of charge shows up on the feedback capacitor C_f. We neglect the effect of R_b and R_f during the short rising edge. The amplitude of the signal is then given by $V_o = Q/C_f$. Take a 60 keV photon and Ge detector for example, there will be 20,000 electron-hole pairs created, which will

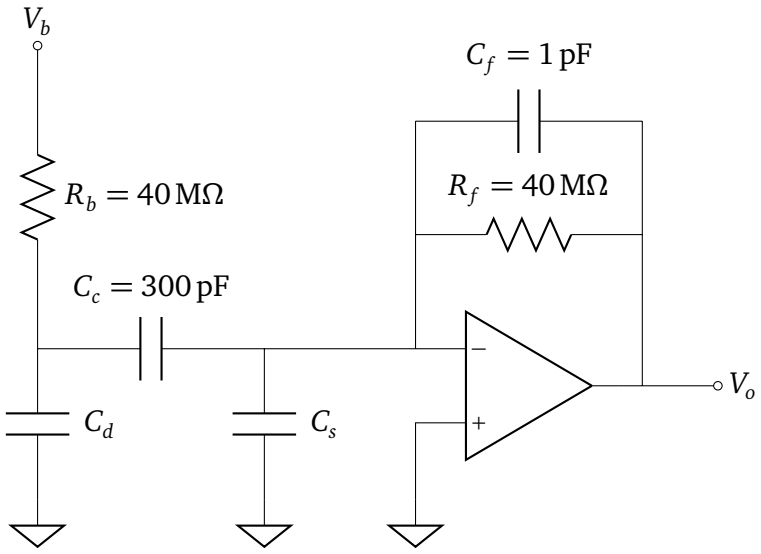


Figure 2.2: Schematic of the CDMS charge readout circuit. The detector is represented by the capacitor C_d. V_b is the detector bias. C_s is the stray capacitance. Typical values are $V_b = -3\,\mathrm{V}$, $C_d = 50\,\mathrm{pF}$, and $C_s = 100\,\mathrm{pF}$.

generate a signal of amplitude 3 mV. Then C_f will discharge through R_f with the time constant $\tau = R_f C_f \approx 40\,\mu\mathrm{s}$. The whole charge pulse rises to it maximum within $1\,\mu\mathrm{s}$ and then decays with time constant $40\,\mu\mathrm{s}$.

The noise performance of the preamplifier is determined by the thermal noise of the resistors and the voltage and current noises of the JFET. A complete noise model is derived in [84]. Here we cite the result in a slightly different form [91]

$$e_0^2 = |A(f)|^2 \left\{ e_{\mathrm{FET}}^2 \left[(C_d + C_f + C_s)^2 (2\pi f)^2 + \left(\frac{1}{R_f} + \frac{1}{R_b} \right)^2 \right] + 4k_B T \left(\frac{1}{R_f} + \frac{1}{R_b} \right) + i_{\mathrm{FET}}^2 + i_d^2 + i_\mu^2 \right\}, \tag{2.4}$$

where e_{FET} and i_{FET} are the voltage and current noises of the JFET, respectively, i_d is the detector current noise, i_u is the current noise due to microphonics, and $A(f)$ is

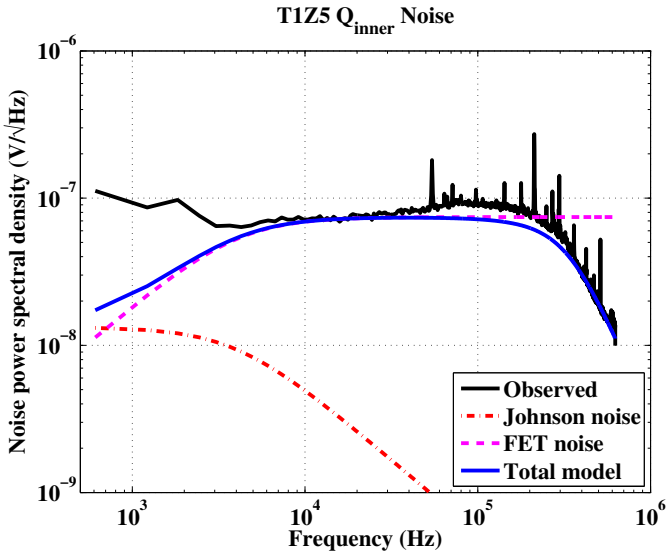


Figure 2.3: Predicted and measured charge noise power spectrum referred to the output of the charge preamplifier. The modeled noise agrees well with the data between 10 kHz and 100 kHz, where the JFET noise is the dominant source. There is a fair amount $1/f$ noise below 10 kHz. *Figure from Filippini [91].*

the transresistance of the amplifier

$$A(f) = \frac{R_f}{1 + 2\pi i f R_f C_f}. \tag{2.5}$$

The JFET voltage noise e_{FET} was measured to be $0.5\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$ at 50 kHz [86], which is the dominant noise at frequency above ∼1 kHz. At low frequencies $1/f$ noise dominates. However, microphonic noise peaks can show up anywhere in the passband and are detector specific. They are especially harmful if show up in the frequency range between 10 kHz and 100 kHz where most of the signal power resides. Figure 2.3 shows the comparison of the above noise model and the measured noise. The model agrees well with the data in the frequency range we are interested in. With an estimated noise floor $80\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$ below frequency 160 kHz, the rms noise is 32 μV.

At 2σ level, the detector threshold is 1.3 keV, which is consistent with the result from data analysis.

2.1.2 Phonon signal

2.1.2.1 Phonon generation

As was mentioned in Sec. 2.1.1.1, a large fraction of the energy deposited by an incoming particle is dissipated as phonons while the electron-hole pairs are generated. These phonons are initially generated around the Debye frequency (a few THz). Within a few μs, through anharmonic decay and isotopic scattering, the initial high frequency phonons down convert into phonons of frequency ∼1 THz, whose mean free paths are then comparable to the size of the detector substrate. The lower frequency phonons propagate ballistically in the crystal and can be detected by the phonon sensors when they reach the surface. Their energy is $E_r - E_g N_Q$.

The second source is the Neganov-Luke phonons emitted by the electrons and holes drifting in the electric field. With complete charge collection, the Luke phonon energy equals the work done by the electric field to drift the charge, which is $eN_Q V_b$, where V_b is the detector bias voltage. The frequency of these phonons is in the range 0.1–0.8 THz [94], which is above the energy required to break a Cooper pair in Al, $2\Delta = 0.084$ THz (0.34 meV).

When the charge carriers arrive at the surface of the detector, they relax to the Fermi level and release their remaining energy by emitting phonons. Since the band gaps of Ge and Si are much larger than the pair-breaking energy in Al, these phonons are sufficiently energetic to create quasiparticles in Al and be detected. Their

contribution to the phonon energy is $E_g N_Q$. The total phonon energy is the sum of the three sources:

$$E_P = E_r + eN_Q V_b \tag{2.6}$$

2.1.2.2 Phonon detection

The ballistic phonons can be detected by the phonon sensors once they reach the surface. Figure 2.4 shows the phonon sensor layout of the CDMS-II detectors. There are four phonon channels for each detector, each occupies a quadrant of the phonon face and consists of 37 tiles. Each tile includes 28 quasiparticle-trapping-assisted electrothermal-feedback transition-edge-sensors (QETs). Each QET has long thin W wire running across the Al fins. There are 1036 QETs in total lithographically patterned on the phonon face for each channel.

Figure 2.5 depicts how the QETs work in detecting phonons conceptually. While in operation, the detectors are cooled to ∼40 mK. Before the occurrence of a scattering event, the Al fins ($T_c = 1.2$ K) are in superconducting state and the W TES (T_c∼80 mK) are voltage biased in the superconducting-normal transition region, which is very sensitive to temperature change. Within a few μs following the collision of an incoming particle and the detector, the ballistic phonons of frequency ∼1 THz reach the detector surface. They are collected and absorbed by the Al fins, creating a cascade of quasiparticles by breaking the Cooper pairs in the Al. The "hot" quasiparticles then diffuse toward the W TES, relaxing to lower energy states by emitting phonons and creating another cascade of quasiparticles in the W element. The quasiparticles which have relaxed to the lower energy states can not diffuse back because of the potential barrier, making the Al fins collectors for the phonons. However, the phonons with

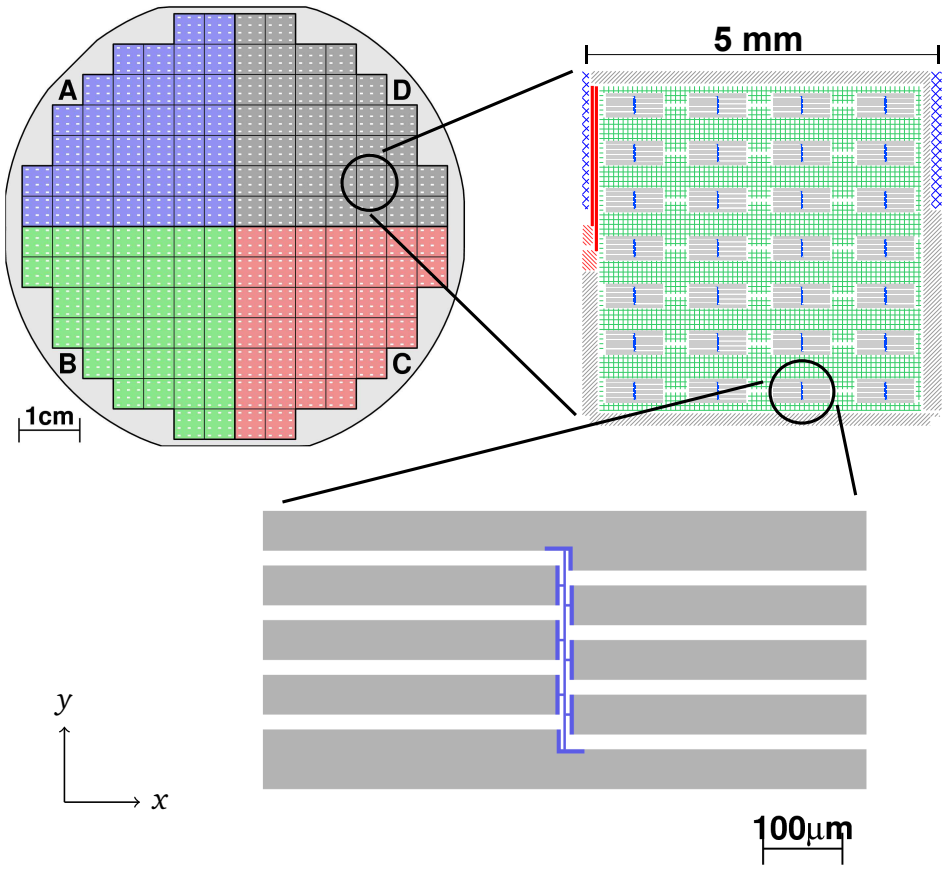


Figure 2.4: Phonon sensor layout of the CDMS-II detectors. (Top left) Phonon face of the detector is divided into A, B, C, and D four quadrants (channels). Each quadrant consists of 37 tiles. (Top right) Each tile includes 28 QET sensors. (Bottom) A single QET sensor with the 250 μm × 1 μm × 35 nm W wire in the middle and Al collecting fins (gray) connected. *Figure from Ogburn [95].*

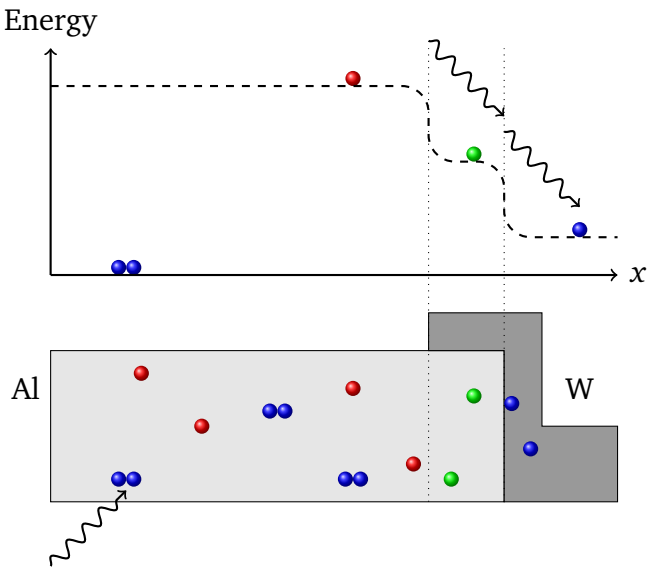


Figure 2.5: Conceptual schematic of quasiparticle trapping. Energetic phonons coming from the detector substrate create a cascade of quasiparticles in Al by breaking the Cooper pairs. These quasiparticles diffuse to the W element and create another cascade of quasiparticles. Colors represent their energy, with red the highest and blue the lowest.

energy less than 2Δ in Al or W can not be detected and the energy they have will be lost.

2.1.2.3 Electrothermal feedback

The W TES elements are voltage biased in their superconducting-normal transition region. The cascade of quasiparticles created by the phonons will raise the temperature and thus the electrical resistance of the TES elements, which will subsequently reduce the Joule heating and assist the TES to recover to its original operating point. This negative feedback is referred to as the electrothermal feedback (ETF). The steepness of this feedback is characterized by the ratio of fractional resistance change to

fractional temperature change at the operating point, i.e.,

$$\alpha = \left.\frac{\mathrm{d}R/R}{\mathrm{d}T/T}\right|_{R=R_0} = \left.\frac{T}{R}\frac{\mathrm{d}R}{\mathrm{d}T}\right|_{R=R_0}, \tag{2.7}$$

where we have used the resistance R_0 to denote the operating point of the TES. The CDMS detectors are usually biased at R_0 between 100 mΩ and 200 mΩ.

To gain some insight about the working mechanism of the TES, we consider a simplified initial condition problem, in which the temperature of the TES is held δT above its operating point T_0 at $t = 0$ by an instant energy deposition. The evolution of δT is governed by

$$C\frac{\mathrm{d}\delta T}{\mathrm{d}t} = \frac{\delta P}{\delta T}\delta T - G\delta T, \tag{2.8}$$

where C is the heat capacity of the TES, $P = \frac{V_b^2}{R}$ is the Joule heating power, and G is the thermal conductance. All quantities are held at their operating points except δT. It is apparent Eq. (2.8) describes an exponential decay of δT. A simple manipulation gives the time constant of the decay

$$\tau = \frac{C}{\alpha\frac{P_0}{T_0} + G} = \frac{C}{\alpha G_0 + G}, \tag{2.9}$$

where we have defined $G_0 \equiv \frac{P_0}{T_0}$. For the CDMS II detectors τ is typically around 200 μs. Equation (2.9) is in close analogy to the time constant of an electrical low-pass RC circuit, for which the 3 dB roll-off frequency is given by $\omega = 1/\tau = \frac{\alpha G_0 + G}{C}$. So α controls the bandwidth of the TES. We expect α to be large, however, since a phonon channel consists many TESs in parallel, the broadening of the transition temperature reduces α and thus the bandwidth. Uniformity of the critical temperature is one of

the central goals of the fabrication of TES. The TESs are initially fabricated with a target $T_c \approx 120\,\mathrm{mK}$ and then fine tuned to around 90 mK.

2.1.2.4 SQUID amplifier

The Joule heating power of a TES under DC bias is $P = IV$. As shown in Eq. (2.8), the change of the Joule heating power δP is proportional to δT, which gives the change of current $\delta I = -I_0 \alpha \frac{\delta T}{T_0}$. This gives us another reason to expect α to be large, i.e., it is the gain of the TES. As was mentioned at the beginning of this chapter, the temperature rise of the detector or its component, say, the TES, due to an event is tiny. And therefore SQUID amplifiers are used to readout this signal.

Figure 2.6 shows the schematic of the CDMS II SQUID amplifier. It is used in negative feedback mode to provide stable gain and adequate bandwidth. The SQUID array consists of 100 SQUIDs in series, each provides a gain of $5 \times 10^4\,\Omega$ to the input coil current I_s (see Appendix B). The total open loop gain is $5 \times 10^6\,\Omega$. The non-inverting op-amp at the upper right corner in Fig. 2.6 multiplies another factor of 20 to the open loop gain, which results in a total open loop gain of the amplifier $5 \times 10^8\,\Omega$. However, this analysis assumes all the SQUIDs work in perfect condition, which is not achievable in real application. Nevertheless, the open loop gain is high enough to enable the characteristic of the amplifier to be determined by the feedback loop. The negative feed keeps the magnetic flux in the SQUID loop as a constant, which indicates $I_f = 10 I_s$ based on the inductance ratio between the input and the feedback coils. Multiplying the value of the feedback resistor R_f gives the gain the the amplifier 12 kΩ referred to I_s.

The bandwidth of the amplifier is determined by the time constant of the input

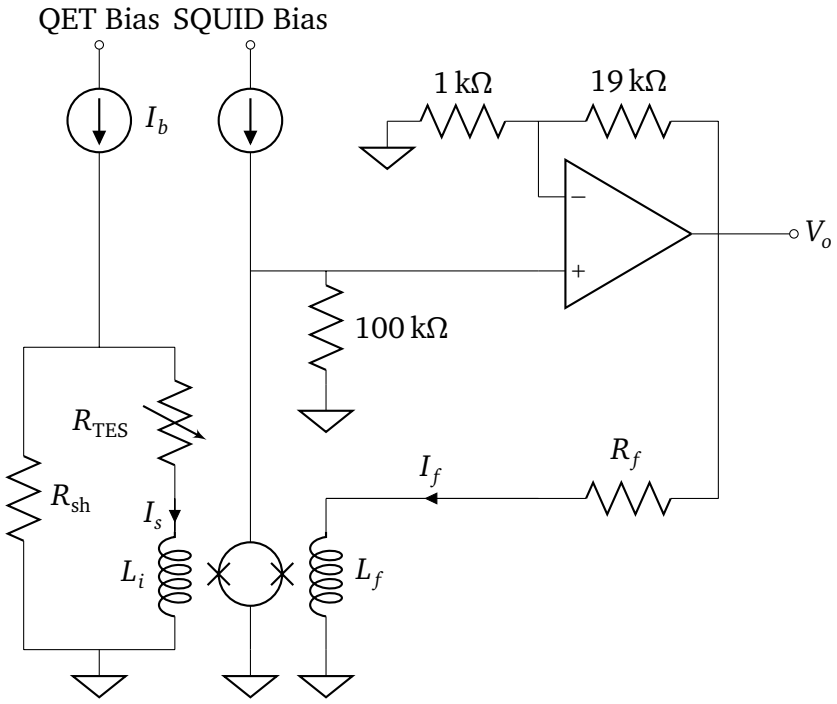


Figure 2.6: Schematic of the CDMS phonon readout circuit. The DC resistance of the phonon sensor is represented by the variable resistance R_{TES}. I_b and R_{sh} are the QET bias and the shunt resistance, respectively. L_i, L_f, and R_f are the input coil inductance, feedback coil inductance, and the feedback resistance, respectively. Typical values are $R_{\mathrm{sh}} = 25\,\mathrm{m\Omega}$, $L_i = 250\,\mathrm{nH}$, $L_f = 25\,\mathrm{nH}$, and $R_f = 1.2\,\mathrm{k\Omega}$. R_{TES} is usually between 100 mΩ and 200 mΩ when biased and around 1 Ω at normal. The modulation depth (peak to peak amplitude of V–Φ curve) of the SQUID is 5 mV and Φ_0 corresponds to $I_s = 25\,\mu\mathrm{A}$. The noise current of the SQUID is typically $2\,\mathrm{pA}/\sqrt{\mathrm{Hz}}$ at 1 kHz referred to the input current I_s.

end, which gives a -3 dB roll-off frequency $\frac{R_{\mathrm{TES}}}{2\pi L_i} \approx 100\,\mathrm{kHz}$ for $R_{\mathrm{TES}} \approx 150\,\mathrm{m\Omega}$. The feedback branch is also a R-L circuit which has similar frequency characteristic but the cutoff frequency is much higher and is not a limiting factor. For typical phonon pulses with rise time $\sim$10 µs and fall time $\sim$200 µs, the bandwidth is well adequate.

The noises in the phonon signal come from three categories of sources, the phonon noise from the TES, the thermal noise from the resistors at the input end, and the SQUID current noise. The thermal noise sources include R_{sh}, the parasitic resistance R_p (on the I_s branch, not shown in Fig. 2.6), the bias resistor R_b (on the I_b branch, not shown in Fig. 2.6), and R_{TES}. All thermal noises are wide band except that of R_{TES} which is rolled off by the electrothermal feedback. The typical SQUID noise is $2\,\mathrm{pA}/\sqrt{\mathrm{Hz}}$ referred to the current of the input coil. The amplifier is designed such that the total electronic noise does not outweigh the phonon noise. In the signal passband, both the phonon noise and the summed electronic noise are around $7\,\mathrm{pA}/\sqrt{\mathrm{Hz}}$, giving a total noise around $10\,\mathrm{pA}/\sqrt{\mathrm{Hz}}$ referred to I_s [86, 90, 96, 97]. There is also $1/f$ noise at low frequencies.

Besides reading out the phonon signals, the SQUID amplifier is also used to measure the T_c's and I–V curves of the TES sensors in situ. In measurement of the T_c's, a small triangle wave bias I_b is sent down the QET bias line while the TES is superconducting, resulting in a small but clear triangle wave output which can be monitored on an oscilloscope. The triangle wave output becomes distorted when the temperature of the TES reaches the superconducting-normal transition region, and then disappears completely or becomes much smaller than its original amplitude when the TES becomes normal. Ideally T_c should be the transition temperature at zero bias, however, rigorousness is less a concern than consistency among the test

facilities for the CDMS application. Practically the smallest bias can be used is limited by the apparatus and the noise environment. A peak-to-peak bias current of 5 μA is normally used in the CDMS measurements (referred to I_s).

The I–V curve (often referred to as I_b–I_s curve) of the TES is important in choosing its bias point. A simple analysis gives

$$I_s = I_b \frac{R_{\rm sh}}{R_{\rm sh} + R_p + R_{\rm TES}} \tag{2.10}$$

$$V_o = \frac{L_i}{L_f} R_f I_s. \tag{2.11}$$

The parasitic resistance R_p can be obtained from the I–V measurement while the TES is superconducting. Then $R_{\rm TES}$ as a function of I_b can be determined.

2.1.3 Background discrimination

Once the charge and phonon pulses as well as their timing are recorded, interested parameters can be extracted through reconstruction of the pulses (see Sec. 5.3.1 for more details). Then the given event can be identified as a nuclear recoil or background event. In the CDMS-II data analysis, three major quantities are used to reject the backgrounds: ionization yield, event position, and event timing.

2.1.3.1 Event position

The position of an event itself may be used to discriminate against certain types of backgrounds, e.g., less penetrating backgrounds are likely to occur at the detector surface. This is actually a powerful quantity in liquid Xe based experiments because of its self-shielding. However, event position is mainly used to improve data quality

in the CDMS-II analysis, i.e., either remove bad regions of the detector or make corrections to them based on their location.

Event position can be defined with the ionization signals from the inner and outer electrodes. At the digitizing rate 2.5 MHz, the rising edges of the ionization pulses are typically no more than two bins, which is not sufficient to provide useful information about the relative delay or pulse shape difference between the inner and outer electrodes. However, event positions can be defined by their energy distribution between the two electrodes. If an event deposits a large fraction of its energy into the outer electrode, it is likely this event is close to the edge of the detector, which may suffer from incomplete ionization collection. Indeed, a `Qinner` cut is defined in the CDMS-II data analysis to reject such events.

In the inner region of the detector, because of the non-uniformity of the crystal itself or the phonon sensors, there is position dependence of the detector response, which broadens the distribution of the background as well as the signal events and hence increases the background leakage. Position correction to the parameters extracted from the signal pulses are performed in the CDMS-II data analysis with respect to a preselected sample of events with known energy and distribution. The 356 keV gamma events from ^{133}Ba sources are used as sample events in position correction (see Sec. 5.3.1.4 and Ref. [87]), which populate the crystal uniformly.

Two types of event positions based on phonons *phonon partition* and *phonon delay* are defined in the CDMS-II data analysis. Phonon partitions can be viewed as the

"center of mass" of the phonon energy distribution:

$$x_P = \frac{P_D + P_C - P_A - P_B}{P_A + P_B + P_C + P_D}, \tag{2.12}$$

$$y_P = \frac{P_A + P_D - P_B - P_C}{P_A + P_B + P_C + P_D}, \tag{2.13}$$

where P_i is the phonon energy for Sensor i. The radius of the phonon partition is defined as

$$r_P = \sqrt{x_P^2 + y_P^2}. \tag{2.14}$$

Phonon delays are the phonon arrival time delays with respect to the primary senor for its adjacent neighbors in the x and y directions. The primary sensor is defined as the one with the earliest phonon start time (or the maximum phonon energy). Take an event with Sensor A (Fig. 2.4) as the primary sensor as an example, its x and y delays defined with the phonon start times are $t_D - t_A$ and $t_B - t_A$, respectively, both of which are positive by definition. Then the delays are mapped to the quadrant of the primary sensor they belong to. For Sensor A, we flip the sign of its x delay because A is in the quadrant $(-,+)$. Putting things together, we have the definition for the phonon delays:

$$x_d = t_{x-} - t_{x+}, \tag{2.15}$$

$$y_d = t_{y-} - t_{y+}, \tag{2.16}$$

where the primary sensor picks one t_x and one t_y based on its quadrant and leave the rest to its nearest neighbors. This way we can easily identify the primary sensor

of an event based on its phonon delay. The phonon delay radius is defined similarly:

$$r_d = \sqrt{x_d^2 + y_d^2}. \tag{2.17}$$

Because of phonon reflection near the side wall of the detector, there are degeneracies in the phonon partition or phonon delay. This degeneracy can be lifted if both positions are used together. This is what was done in the position correction, more details about which can be found in Refs. [87] and [88]. The distributions of the parameters are much narrower after the position correction, improving the separation between the signal and the background.

2.1.3.2 Ionization yield

As was discussed in Sec. 2.1.1.1, an electron recoil is more effective in creating charges than a nuclear recoil with the same energy. For electron and nuclear recoils in the bulk, of which the charge collections are nearly complete, this distinction is reflected in the ionization yield, which is defined as the ratio of the ionization energy to the recoil energy,

$$y \equiv \frac{E_Q}{E_r} = \frac{E_Q}{E_P - eN_Q V_b} = \frac{E_Q}{E_P - \frac{eV_b}{\epsilon} E_Q}. \tag{2.18}$$

The normalization is chosen such that $y = 1$ for electron recoils. The yield for nuclear recoils is around $1/3$ under the same normalization. Because of the difference in ionization yield, the same measured ionization energy corresponds to vastly different recoil energies for different types of recoils. It is sometimes necessary to specify the type of recoils for the recoil energy. A common practice is to use units such as keVee and keVnr to denote electron recoils and nuclear recoils, respectively.

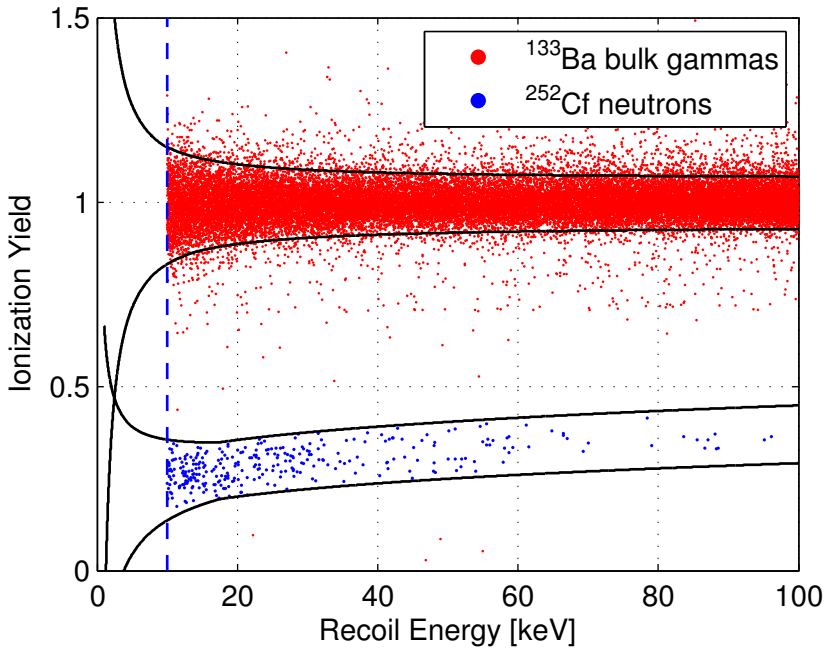


Figure 2.7: Ionization yield vs. recoil energy for ^{133}Ba gamma calibration events and ^{252}Cf neutron calibration events. The bulk electron recoil band is centered at yield = 1. The yield for the nuclear recoil band is centered $\sim$1/3 with moderate recoil energy dependence. The solid lines are the $\pm 2\,\sigma$ boundaries of the yield bands. The electron recoil and nuclear recoil bands are well separated down to below 5 keV. The blue dashed line shows the typical analysis threshold 10 keV. Data from detector T1Z5.

Figure 2.7 shows the electron recoil band from the ^{133}Ba gamma events and the nuclear recoil band from the ^{252}Cf neutron events. The two bands are well separated down to below 5 keV, providing strong discrimination against the bulk electron recoil background.

2.1.3.3 Timing

The bulk electron recoils are well separated from the nuclear recoil band. However, for electron recoils that occurred close to the surface of the detector, the charge collection may not be complete. This reduced ionization yield makes them difficult to be identified as electron recoil background. Fortunately, the phonon signals for these surface events are more prompt than for the bulk nuclear recoil events. The surface event phonon pulses tend to have shorter delay relative to their charge pulses as well as shorter rise time. A timing parameter is defined for an event as the sum of the phonon pulse delay and the phonon pulse rise time for the primary phonon channel, which has been proven to have good discrimination power against surface events.

Figure 2.8 shows the combined discrimination with ionization yield and phonon timing against electron recoil backgrounds. The two discriminators work orthogonally providing $\sim 10^{-6}$ rejection to bulk electron recoils and $\sim \frac{1}{200}$ rejection to surface events [87].

2.2 Experimental installation

It took great effort to put the CDMS experimental setup in the Soudan mine and make it work. There are quite a few excellent theses [47, 86, 91, 95] documenting the

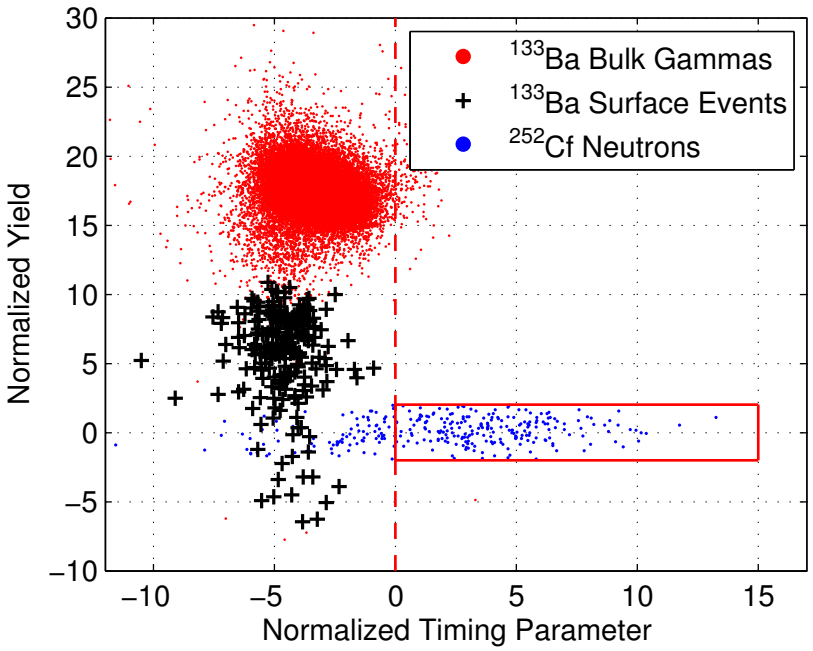


Figure 2.8: Combined discrimination with ionization yield and phonon timing against electron recoil backgrounds. The ionization yield separates the backgrounds and the nuclear recoils vertically, and the timing parameter further separates them in the orthogonal direction. Both the yield and timing parameter in the plot are normalized by their standard deviations. The vertical dashed line shows the position of the timing cut and the red box is the signal acceptance region. *Figure from Ahmed* et al. *[98].*

early days and explaining the science behind the apparatus. Since there has not been major changes to the experimental installation, a brief introduction will be given in this section focusing on the parts of direct consequence to the science reach of the experiment. More details can be found in the aforementioned references.

2.2.1 Soudan Underground Laboratory

As we discussed before, it is essential to reduce the neutron background to an acceptable level in a direct dark matter search experiment. Figure 2.9 shows the relative importance of various neutron sources as a function of depth. At depths greater than $\sim$100 m.w.e., the neutron production is dominated by the natural radioactivity in rocks and the rate for the fast muon induced neutrons is a few orders of magnitude smaller. However, the radiogenic neutrons produced in the former process are of low energy (typically < 8 MeV) and can be shielded relatively easily. The fast muon induced neutrons have a very hard spectrum extending to several GeV and can penetrate a significant depth of surrounding rocks and detector shielding materials. Locating the experiment in a lower depth is the only viable way to reduce this fast neutron background.

The Soudan Underground Laboratory (SUL) is on the 27th level of the decommissioned Soudan Mine in northern Minnesota, which is 714 m below the surface. The rock overburden provides shielding against cosmic rays equivalent to 2090 meters of water, reducing the surface muon flux by a factor of 5×10^4 [100] (see Fig. 2.10). The neutron background of the entire CDMS-II run at the SUL is $\sim$0.1 events [98], which is not a limiting factor to the experiment, especially compared with the much larger surface event background. However, it will become necessary to move to a

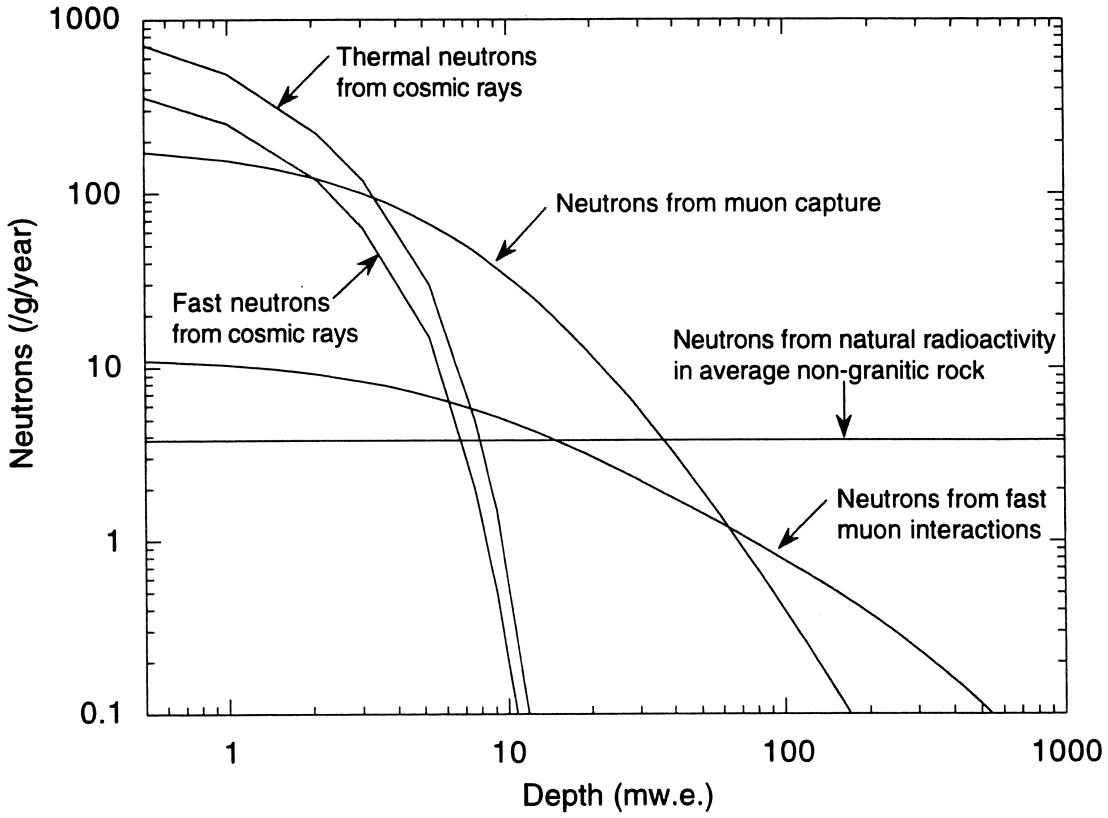


Figure 2.9: Neutron production for non-granitic rock vs. depth. *Figure from Silva [99].*

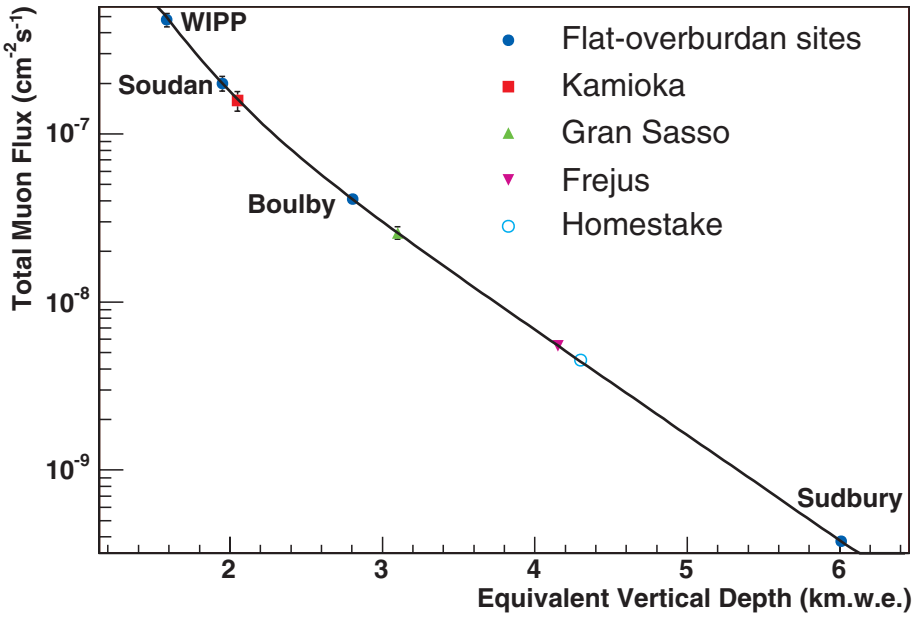


Figure 2.10: Measured total muon flux for various underground sites. The solid line is $I_\mu(h_0) = (67.97e^{-h_0/0.285} + 2.071e^{-h_0/0.698}) \cdot 10^{-6}$, where h_0 is the equivalent vertical depth in the range ∼1 to 8 km.w.e. The muon flux at the sea level from Rossi [101] is $J_2 = 1.68 \cdot 10^{-2}\text{cm}^{-2}\text{s}^{-1}$. The flux at Soudan $2 \cdot 10^{-7}\text{cm}^{-2}\text{s}^{-1}$ corresponds to a reduction factor of $\approx 8 \cdot 10^4$. The flux at Sudbury is $3.77 \cdot 10^{-10}\text{cm}^{-2}\text{s}^{-1}$ which corresponds to a reduction factor of $\approx 4 \cdot 10^7$. *Figure from Mei and Hime [102].*

deeper site and reduce the neutron background even more for the next generation ∼100 kg scale experiment.

2.2.2 Shielding

Multiple layers of shielding are used to reduce the backgrounds, as shown in Fig. 2.11. The outermost layer is the active muon veto consisting of 40 panels of 5 cm thick plastic scintillators. Both muons and ambient gammas can trigger the veto system. The total trigger rate is ∼600 Hz, of which the muon rate is one per minute on average. The veto system can distinguish muons from ambient gammas by their energy deposition.

Muons typically deposit 10 MeV onto the veto system while the highest energy of the photons from natural radioactivity is 2.6 MeV. The muon detection efficiency is nearly 100 % [100, 103]. Veto traces for each panel are recorded in a time window −185 μs before and 20 μs after the event trigger from a detector. Events with any veto activity within 50 μs before the event trigger are rejected in analysis.

The second layer from outside is the 40 cm thick polythene neutron moderator. Ambient gammas are blocked by the 22.5 cm thick lead shield inside the outer polyethylene. The inner 4.5 cm of the lead shield is made out of ancient lead with low radioactivity. Inside the lead shield is the 10 cm thick inner polyethylene neutron moderator primarily to stop the neutrons from the lead shield. A μ-metal shield is installed between the radiation shield and the Icebox to shield the SQUIDs from the Earth magnetic field. The space between the μ-metal shield and the Icebox is injected with pure N_2 gas from liquid N_2 boil-off to reduce the radon level. The copper cans of the Icebox at various temperate stages provide 3 cm of shielding directly surrounding the detectors. The coppers used are counted to ensure low radioactivity. The shield covers 99 % of the solid angle outside the Icebox with exceptions at the cold stem and the electrical stem.

To reduce the radioactive contamination during installation and handling of the detectors, the fridge and Icebox are enclosed in a Class 10,000 clean room, which is also a Faraday cage shielding the electrical noise from outside.

2.2.3 Cryogenics, hardware, and electronics

The detectors are cooled to a base temperature ∼40 mK while in operation by an Oxford Kelvinox 400 dilution fridge. Additional GM cryocoolers are installed to

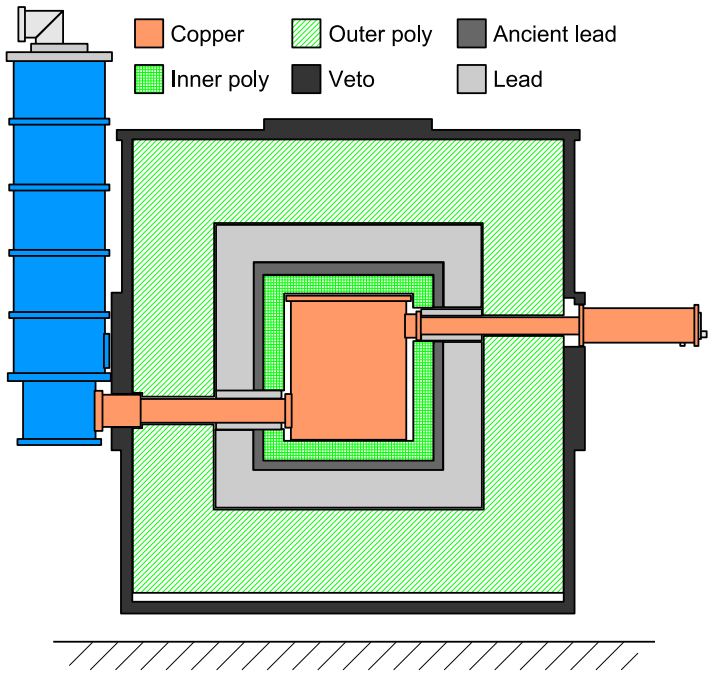


Figure 2.11: Shielding around the icebox. The blue cylinder at the left is the dilution refrigerator, which is connected to the Icebox via the "cold stems" at its tail. The copper can in the center of the shield is the Icebox. The shield consists of, from outside to inside, 5 cm thick plastic scintillator muon veto shield, 40 cm thick polyethylene neutron moderator, 22.5 cm thick lead gamma shield, of which the inner 4.5 cm is ancient lead, and 10 cm thick polyethylene neutron moderator. There is a μ-metal shield outside the Icebox (not shown in the figure). Pure N_2 gas is injected into the space between the μ-metal shield and the Icebox to reduce the Radon level. The high purity copper cans outside the detectors accumulate to a thickness of 3 cm. *Figure from Ogburn [95].*

provide extra cooling power to the 77 K and 4 K stages.

Ge and Si detectors are packaged in copper detector housings and stacked in groups of six onto towers. Thirty detectors in five towers were installed for the full scale CDMS-II runs. Each tower consists of four thermal stages connected by low thermal conductivity graphite tubes to provide sufficient thermal isolation. The thermal stages are heat sunk to the mixing chamber (∼40 mK), cold plate (∼150 mK), still (∼600 mK), and inner vacuum can (IVC, at 4 K) in the order of increasing temperature. For each detector, charge and phonon signals are brought out by the wire bonds to the detector interface board (DIB) on one face of the detector housing. A vacuum sidecoax bridges the signals from the DIB to the 40 mK stage of the tower. The signals are then amplified by the phonon amplifiers installed on the 600 mK stage and charge amplifiers at the 4 K stage. The amplified signals are then brought out by striplines to the electronic box (EBox) at room temperature.

Room temperature electronics further amplify, trigger, and digitize the signals. Recorded raw data are transferred to the Soudan Analysis Cluster on the surface, where the data are reduced to reduced quantities for data quality check and detector performance assessment. Primary data processing is carried out on the FermiGrid at Fermilab. The processed data can then be copied to individual institutions and be analyzed locally. The CDMS collaboration also have computing resources at the Stanford Underground Facility to centralize the storage, distribution, and analysis of the processed data as well as publishing the analysis results.

Chapter 3

The CDMS R&D Facility at UMN

The decision to develop a new CDMS R&D facility on the University of Minnesota campus was made in the summer of 2008. The main objectives of the facility, as part of the CDMS collaboration, were to explore solutions to the problems in scaling up the CDMS-II experiment to a second generation direct dark matter search experiment with detector payload in the order of 100 kg and to characterize the detectors before their final deployment in an underground lab. The facility would also be an independent research lab, where interesting ideas could be pursued and new frontiers of physics could be explored. In a word, the new facility would be an important part of the CDMS experiment in pushing the dark matter sensitivity to newer levels and, at the same time, would bring in a whole new scope of possibilities to the existing cryogenic particle detector research.

Following the decision, with the newly acquired lab space at that time, going through lab space planning, infrastructure construction, reengineering and installation of a Kelvinox 100 dilution fridge, integration of the fridge with the CDMS cold

hardware and cold/warm electronics, and integration of the fridge control/monitoring and detector control/readout software, the new facility was up and running by October 2009, when we did the very first full scale test with a 3 inch diameter CDMS-II detector and saw our first phonon pulses. A few months earlier, in the summer of 2009, two soft wall cleanrooms were built with major efforts from two students in the summer Research Experience for Undergraduates program, providing clean environments for the detector and cold hardware work. Shortly after the first runs, in order to test the 100 mm diameter ionization/phonon devices, the innermost copper can of the fridge was upgraded from 3 inch (76 mm) to 4 inch (100 mm). The high event rate of a large detector can overwhelm the data acquisition system. With the effort primarily from Roxanne Radpour, Alexander Codoreanu, Allison Kenedy, and Sean Vigg, a polyethylene neutron shield and a lead gamma shield were constructed to reduce the ambient radiation illuminating the detectors.

Since our first run, the following categories of research activities had been performed in the facility:

- Charactering a subset of the SuperCDMS Soudan iZIP's before their deployment at the Soudan mine.
- Testing and characterizing the 100 mm diameter ionization and phonon test devices.
- Testing the Detector Control and Readout Card (DCRC) designed for the next generation CDMS experiment.
- Testing, debugging, and improving the detector control and readout software working with the DCRC.
- Exploring new designs of the charge readout circuit with the intention to simplify

the detector tower design.

The above is a brief history of the test facility. This chapter presents a somewhat detailed introduction to the facility and its operation, focusing on the big picture and some of the details which may be of importance for the future of the facility.

3.1 Overview of the facility and its operation

3.1.1 Tower assembling and installation

Figure 3.1 shows the top view of the lab while in operation. The orientation of the schematic follows the standard map convention with north up. The dilution fridge (F) is at the northeast corner of the main room (S34). The rectangle surrounding the fridge is the supporting structure with four pillars (see Fig. 3.2(a)) at the four corners and an aluminum platform on top. The dilution unit is seated on top of the aluminum platform. A cylindrical pit was dug in the ground straightly below the dilution unit. The fridge Dewar can be lowered into and raised out of the pit. It is stored in the pit when the fridge is open to install/remove the detectors (Fig. 3.2(c)).

The tower assembly (see Fig. 3.11), which typically consists of one or two detectors, calibration radioactive sources, a tower, and necessary vacuum sidecoaxes, is assembled in the large cleanroom (Fig. 3.3(b)) at the southeast corner (Fig. 3.1) of the main room. The detectors and cold hardware such as detector housings, radioactive source holders, and vacuum sidecoaxes, are stored in the boiloff N_2 purged cabinets in the large cleanroom to avoid radon plate out. There is a cleanbench and a regular bench in the cleanroom, both are covered with electrostatic discharge safe mats. The tower assembly is assembled/disassembled on the cleanbench, whereas other general

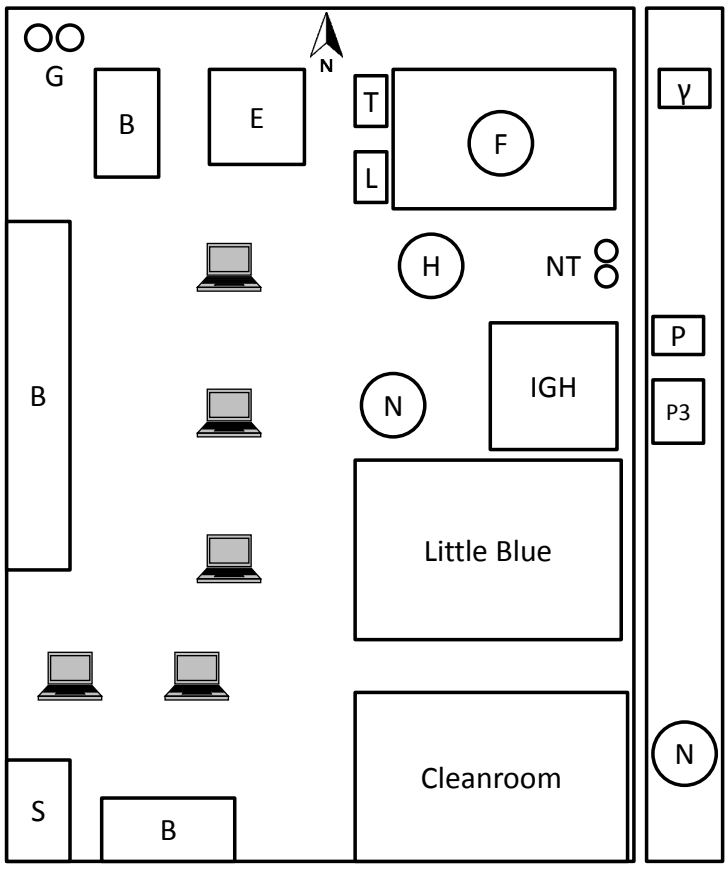


Figure 3.1: Schematic of the top view of the UMN R&D facility located at Room S34 in the subbasement of Tate Lab while in operation (not to scale). The facility consists of a larger room on the left and a long narrow corridor on the right. The radioactive source safe (γ), 1 K pot pump (P), the circulation pump (P3), and the liquid N_2 tank (N) are in the corridor. Boil off N_2 gas from the tank is used to purge the storage cabinets in the cleanroom. The compressed N_2 and ^{4}He gas cylinders (G), workbenches (B), electronics rack and DAQ computer (E), turbo pumping station (T), helium leak detector (L), Kelvinox 100 dilution fridge (F), N_2 cold traps (NT), dilution fridge gas handling unit (IGH), and liquid ^{4}He (H) and N_2 (N) storage Dewars are in the main room. The big cleanroom is in the southeast corner of the main room. The rest of the space is used by another smaller dilution fridge Little Blue and five work stations (the computer symbols). At the southwest corner is the sink (S).

work is performed on the regular bench. After passing the electrical checkout, the completed tower assembly is mounted upside down (with the detector stack on top) on a tower insertion tool and double bagged in clear polyethylene cleanroom bags.

The tower assembly is then transferred to the small clean-tent (Fig. 3.3(a)) and installed into the fridge. It is good practice to do a second round of electrical checkout after the tower is raised and bolted to the 30 mK can. After installations of all the radiation shields and thermal links, the fridge inner vacuum can (IVC) is sealed if the the tower passes the electrical checkout conducted from the breakout box (Fig. 3.6(a)). Then the fridge Dewar is raised.

3.1.2 Fridge cooldown

The cooldown of the fridge follows the procedure for a regular dilution fridge. The IVC is evacuated at room temperature and leak checked. If no leak is found, N_2 heat exchange gas is then added into the IVC and liquid N_2 is transfered into the Dewar to cool the cryostat to 77 K. When the tower temperature gets close enough to 77 K, the N_2 exchange gas is pumped out. Compressed ^{4}He gas is then used to push out the residual liquid N_2. At the same time, a helium leak detector is hooked up to the IVC to detect any cryoleak. If no leak is detected, liquid ^{4}He is then transferred into the fridge Dewar with ^{4}He heat exchange gas added to the IVC. When tower temperature approaches 4.2 K, the ^{4}He heat exchange gas is pumped out, and the process of condensing and circulating the mixture is initiated to cool the cryostat to the base temperature.

At each temperature stage, especially at room temperature and 77 K, an electrical checkout is conducted and the result is recorded. Any fatal failure found would halt

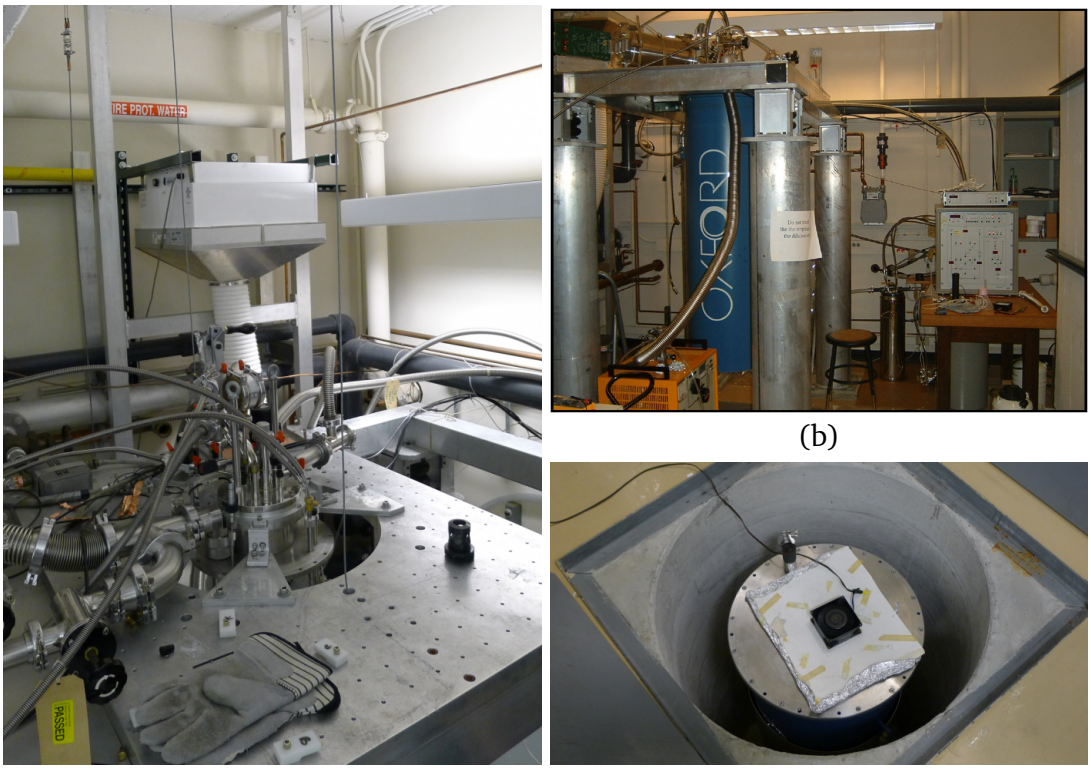


Figure 3.2: Kelvinox 100 dilution fridge, the supporting structure, and the pit. (a): The dilution unit held by the three aluminum brackets bolted to the aluminum plate. The HEPA filter unit hanging on the aluminum rack behind is for the small cleanroom surrounding the IVC of the fridge (Fig. 3.3(a)). (b): The dilution fridge with the Dewar (big blue cylinder) raised. The cube on the bench on the left is the gas handling unit (IGH) of the fridge. The gray cylinder under the bench is the ^{3}He -^{4}He mixture storage keg. (c): The fridge Dewar lowered into the pit. The box fan on top of the Dewar blows warm air into the Dewar to speed up the warm-up and remove the moisture from the Dewar.

(a) (b)

Figure 3.3: (a): Clean tent encloses the lower part of the IVC. Filtered air from the HEPA filter (Fig. 3.2(a)) maintains a positive pressure in the cleanroom, which keeps the dust particles out. The cleanliness was measured to be of the order of Class 100. (b): The large Class 10000 cleanroom at the southeast corner of Room S34. There is a cleanbench on the left side. At the right side, the boiloff N_2 purged storage cabinets are at the inner corner, outside is a regular bench. All benchtops are covered by electrostatic discharge safe mats (blue).

the cooldown process until the problem is fixed, which usually requires warm-up of the cryostat. The volume of the modified IVC is quite large compared with that of the original one, which requires longer than normal pumping time. The timing of pumping out the heat exchange gas is critical, especially for the ^{4}He exchange gas. Superfluid helium kills the phonon signal and thermally shorts different temperature stages. Residual ^{4}He heat exchange gas in the IVC could ruin the entire run. Our experience was that a half hour pumping on the IVC before the tower temperature reaches 4 K would provide enough isolation to stop the tower from reaching 4 K. To make sure the ^{4}He heat exchange gas is pumped out, we usually pump the IVC overnight after liquid helium transfer. ^{3}He was used as heat exchange gas in the early days of the facility operation, which does not have the liquefaction problem at 4 K. The increasing difficulty to obtain ^{3}He deprived us of this luxury. Residual N_2

exchange gas is normally not a threat to the cooldown since it can always be pumped out when the Dewar is warmed up above 77 K after the liquid N_2 is removed.

The lowest base temperature obtained with a CDMS-II detector payload was 25 mK. Because of heat load variation and possible heat leaks, the base temperature of the fridge varied from 50 mK to 80 mK. A base temperature of 50 mK to 60 mK was usually reachable without too much optimization, which was adequate for the CDMS detector operation. The critical temperature of the phonon sensors was tuned targeting ∼90 mK, though it could be as low as 70 mK. Nevertheless, there was still room for optimization to obtain a lower base temperature if needed.

3.1.3 Detector neutralization and tuning

When a detector first cooled to the base temperature from room temperature, it could be in a bad state in which no ionization signal could be observed if biased. This was because the impurity centers ionized at room temperature were not neutralized during cooldown and the electron-hole pairs generated in an event were all trapped. Flashing the detector with the infrared LED or exposing the detector to strong gamma radiation usually neutralizes the detector and fixes the problem [84]. What we normally did was to flash the detector with the infrared LED overnight.

Detector tuning follows the LED flashing. The ionization channel tuning only requires biasing the charge electrodes to the desired voltage. The phonon channel tuning is more involving. First we tune the SQUID amplifiers alone to obtain the best responsivity by varying the SQUID bias and lock point. A zapping process to get rid of the trapped magnetic flux in the SQUID loop is usually needed before tuning the SQUIDs. To reduce the magnetic field surrounding the SQUIDs, the fridge is shielded

by a cylindrical μ-metal shield. The SQUID amplifier is switched to closed-loop mode once the best responsivity is obtained. Next we measure the I-V curve of the phonon sensors with the tuned SQUID amplifier and then bias the TES at a DC resistance of $\sim$200 mΩ in the transition region.

3.1.4 Calibration sources and background radiation

Radioactive sources are used to study the response of the detector to radiations with different characteristics. The facility has encapsulated ^{241}Am, ^{133}Ba, and ^{252}Cf sources stored in a safe in the corridor besides the main room (Fig. 3.1). The ^{241}Am sources are collimated to provide beams of 60 keV photons to investigate the localized bulk electron recoils of the detector. The 356 keV line from the ^{133}Ba source is used to calibrate the ionization energy scale. The ^{252}Cf source provides low energy neutrons, which can be used to study the nuclear recoils of the detector.

There is a fair amount of ambient radiation at the depth of the lab which easily saturates the DAQ for a large radiation detector such as the 100 mm diameter test devices discussed in Chapter 4. When taking phonon traces which typically last for a few ms, pileup pulses would make the data completely unusable. A polyethylene neutron shield (Fig. 3.4) with 8 inch polyethylene panels each side and a Pb gamma shield (Fig. 3.5) with 0.5 inch thick walls were built to attenuate the ambient radiation reaching the detector. The Pb shield blocks line of sight photons and reduces the total incident flux substantially. In order to study the surface event rejection capability of the detector, it is necessary to reduce the ambient neutrons bombarding the detector. Neutrons typically scatter multiple times in the material and they do not travel in straight lines. Thus we put as much polyethylene as we could inside and outside the

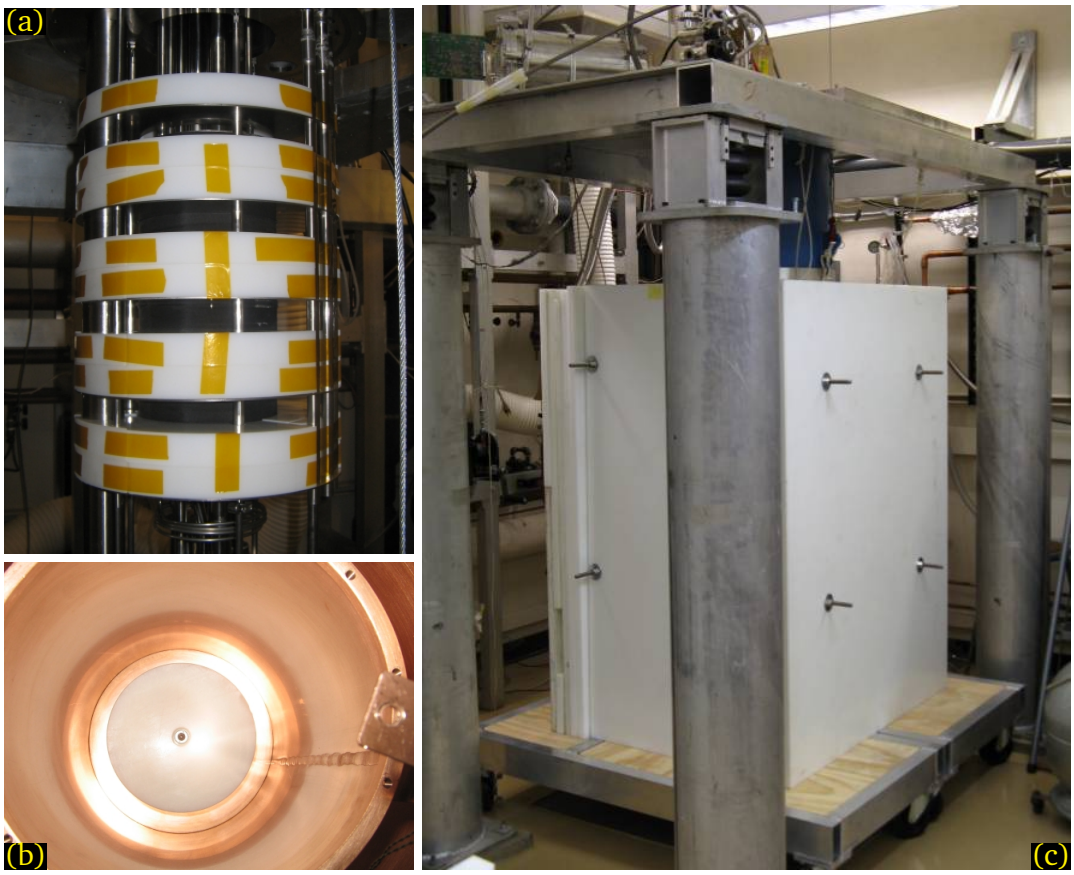
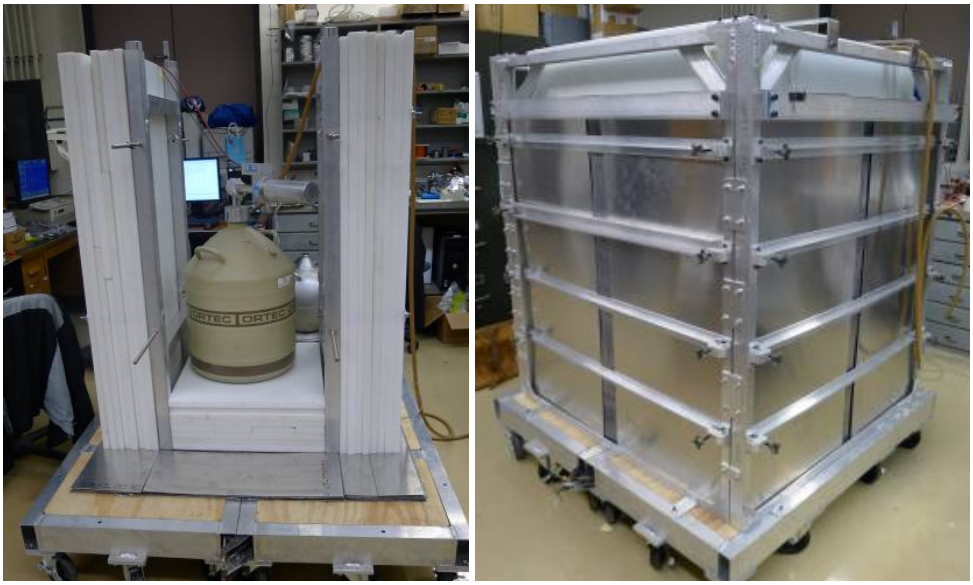


Figure 3.4: Neutron shields in the facility. (a): Polyethylene sheets between the radiation baffle plates above the IVC. (b): Polyethylene block on the ceiling of the innermost copper can which holds the tower. (c): Assembled polyethylene neutron shield surrounding the fridge. The walls of the shield are 8 inch thick each side. There are also 8 inch of polyethylene right underneath the fridge Dewar (Fig. 3.5(a)).



(a) (b)

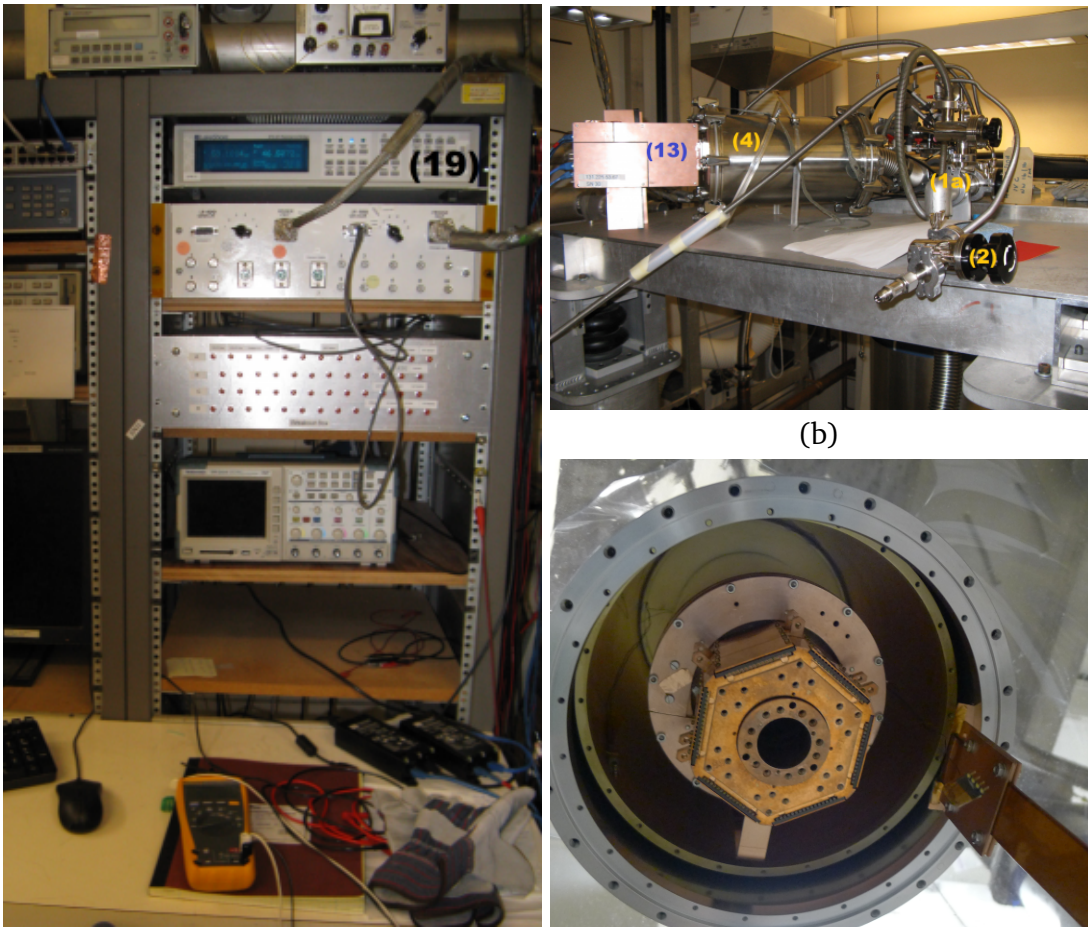
Figure 3.5: Gamma shield in the facility. (a) High-purity Ge (HPG) gamma ray spectrometer sitting inside the shield. There is 0.5 inch of Pb underneath the polyethylene neutron shield. (b): Completely assembled gamma shield.

fridge and in all directions of the detector (Fig. 3.4) to reduce the neutrons sneaking in. With all the measures taken, the neutron flux can be reduced by a factor of ∼10 [104].

3.1.5 Data acquisition and processing

To better understand the background radiation in the lab and the detectors, a high-purity germanium (HPGe) gamma spectrometer was used to characterize the ambient radiation with different shield configurations (Fig. 3.5). A Monte Carlo simulation based on the geometry of the fridge was also formulated, though still at the beginning stage of the development.

Ionization and phonon signals are amplified by the cold electronics first and then



(a) (b) (c)

Figure 3.6: Electronics in the facility. (a): From top to bottom, Lake Shore Model 370 AC Resistance Bridge (19), fridge thermometry breakout box, detector signal breakout box, and oscilloscope on the electronics rack. The DAQ computer is on the lower left corner of the rack (partially shown). (b): CDMS warm electronics. The large stainless cylinder is the EBox (4). Six 50-pin D connectors are welded on to the blank flange on the left of the EBox. A DCRC (13) is plugged into one D connector. The Ethernet cable connecting the left end of the DCRC provides power to the board and brings the digitized signal to the DAQ computer. (c): Tower installed inside the fridge. The striplines are on the right.

are brought out by the striplines (Fig. 3.6(c)) and sent to the EBox (Fig. 3.6(b)). The DCRCs (Fig. 3.6(b)) plugged into the EBox perform all the functionalities of the Soudan warm electronics, including signal amplifying, filtering, digitizing, and triggering. The Labview detector control and DAQ software running on a Windows PC communicates with the DCRCs via the standard TCP/IP protocol through the Ethernet cable. The DCRCs are powered by Power Over Ethernet (POE) injectors between the PC and the DCRC. One DCRC can only control a single CDMS-II type detector which has two charge channels and four phonon channels. If more channels need to be controlled/read out, multiple DCRCs need be used and time synced to work together. The data traffic is then handled by an Ethernet switch between the PC and the DCRCs.

The raw ionization and phonon pulses requested by the DAQ software based on the trigger information provided by the DCRCs are written to raw data files on the harddrive of the DAQ computer. The raw data files are transferred to the departmental Linux cluster and processed with the CDMS data processing package `CDMSBats`. The Condor system is utilized to process the data parallelly. The reduced quantities are stored on the departmental Linux cluster and can be analyzed with the CDMS Analysis Package (CAP) or ROOT.

The above is a synopsis of the typical operations of the facility when radiation-detector payloads are used. For some of the research activities such as testing of new charge readout circuits, the procedures are modified accordingly.

3.2 Fridge modification

The fridge used in the facility was modified from an Oxford Instruments Kelvinox 100 dilution fridge. It was named "K100" after it came online. The dilution unit was not changed. The sliding seal and the superconducting magnet assembly were removed to make full use of the space inside the fridge Dewar. A flanged collar was fabricated to connect the dilution unit and the Dewar. The IVC was enlarged to accommodate the CDMS tower assembly and cold hardware/electronics. Necessary inner structures such as radiation shields and thermal links etc. were added inside the IVC to attach the tower assembly and cold electronics. A 1.5 inch diameter wiring port was added for the striplines connecting the room temperature and the 4 K electronics.

The modification was a typical cryogenic engineering process. In this section I discuss some general considerations in reengineering the fridge and show a few specific examples with the emphasis on the potential future needs and some subtleties that need special attention.

3.2.1 General considerations in reengineering the fridge

The main objective of reengineering the fridge was to explore the feasibility of using large diameter Ge or Si crystals as cryogenic dark matter detectors. The largest crystal in mind before we started was 6 inch diameter Si crystals. The largest detector grade high purity ($(N_d - N_a) \sim 10^{10}\,\mathrm{cm}^{-3}$) Ge crystal turned out to be smaller, the diameter of which was around 4 inch as of the time of writing. The Dewar of the Oxford Instruments Kelvinox 100 dilution fridge we obtained had a inner diameter of 10 inch, which is the maximally allowed diameter of the IVC inside the Dewar. An indium seal

flange would easily take 0.5 inch each side leaving a usable diameter of 9 inch. The detector housing of the 6 inch diameter crystal may take another 1 inch in diameter. The rest usable space is about 1 inch each side to fit all the radiation shields, thermal links, and necessary fixtures. The space was pretty tight. Nevertheless, there was enough space to achieve the objective if carefully engineered.

The second question was whether the fridge had enough cooling power for the thermal load of a CDMS tower assembly. The dilution fridge of the CDMS test facility at UC Berkeley had a nominal cooling power of 75 μW at 100 mK and could reach a base temperature of ≈ 50 mK. The Soudan dilution fridge had a nominal cooling power of 400 μW at 100 mK and could reach a base temperature of ≈ 40 mK. By a simple comparison, we can easily see that the Kelvinox 100 fridge, which had 100 μW nominal cooling power at 100 mK, had enough cooling power for the desired thermal load.

The next step was to design, fabricate, and test the peripheral structures surrounding the dilution fridge until a satisfactory solution was reached. Different components normally have very different critical requirements and need to be treated separately. However, generally speaking, these requirements belong to the categories of geometrical, mechanical, thermal, electromagnetic, and vacuum constraints.

3.2.1.1 Material selection

The suitable material for a component is determined by its function requirements in various aspects such as thermal, mechanical, and electromagnetic etc. Usually one or two of these requirements are critical which have to be met, whereas others can be relaxed to some extent. The critical requirements usually confines the material in a

fairly narrow range. The detailed designs and the corresponding fabrication process for a component to perform the same functions with different materials could be very different. Before considering the design details, the material needs to be selected with the qualitative requirements.

Unless extreme performance is needed, commonly used materials for a component performing similar functions to the one to be designed would usually provide satisfactory performance. If the cryostat under modification had similar components, then the same material could be a good choice. Otherwise, existing solutions may exist in literature such as Pobell [105], Betts [106], and White and Meeson [107]. The machinists of the physics machineshop had extensive experience in cryogenic engineering. They could usually provide useful information on material selection and fabrication. The machineshop had a local storage of commonly used materials, which could save turn around time and cost if utilized.

For the Kelvinox 100 modification, we mainly used 304 stainless steel for load-bearing structures and the walls of vacuum cans, Oxygen-free high thermal conductivity (OFHC) copper for radiation shields and thermal links, and G10 fiberglass for insulating components.

Unlike general condensed matter low temperature experiments, the low background counting experiments such as CDMS require radio-pure materials surrounding the detector. Though it was not a concern for us as a R&D facility, it may limit the choices of available materials in certain occasions. Magnetic properties of the materials deserve special attention as well in our cryostat because of the presence of SQUID amplifier. Any magnetic material should be avoided around the SQUIDs. Stainless steel, though generally considered as nonmagnetic, may become somewhat magnetic

because of the iron phase transition in the alloy and should also be avoided.

3.2.1.2 Geometry and thermal expansion

The geometry of a component can be relatively easily calculated and modeled in a computer program. Besides the necessary clearances at room temperature, the design should also consider the thermal expansion of the component itself and the neighboring components to avoid undesired interference or excessive strain. The thermal expansion coefficients of some typical materials can be found in Pobell [105] and other literature. Note that the thermal expansion coefficient of the material being used may be different from the documented values for the material with the same name. It is good practice to have a higher safety margin built into the design if possible. For critical components, the thermal expansion coefficient of the material may need be measured in the desired temperature range. For noncritical components, if the thermal expansion coefficient is unknown, a reasonable upper limit may be assumed.

3.2.1.3 Thermal conductance and mechanical strength

For a given material, the larger the cross-sectional area, the higher the thermal conductance and the mechanical strength. If a high thermal conductance is desired, it does not impose any constraints on the mechanical strength. A solution satisfying both requirements can always be obtained. However, if a low thermal conductance is desired, too small a cross-sectional area may not provide enough mechanical strength. In this case, if a geometry satisfying both requirements does not exist, a stronger (per unit thermal conductance) material should be used.

For load-bearing structures, regardless of the thermal conductance requirements, a stress analysis must be performed to make sure the structure has enough mechanical strength. There are finite element stress analysis software packages that can be used for this purpose. The 3D design software Autodesk Inventor also has finite element stress analysis functionality built in, which is very convenient. Young's modulus and the shear modulus of a material typically increase as the temperature goes down. Thus the material is stronger at lower temperature. If the structure passes the stress test at room temperature, it is safe to assume it will work at low temperatures.

3.2.1.4 Surface and interface

Joints between different components are unavoidable. At low temperatures, the thermal boundary resistance between two metals makes it difficult to make a good thermal contact. Generally, firmly pressed contacts with clean surfaces have small thermal resistance [105]. Some of the parts may need be joined together by welding, silver soldering, or soft soldering. Welding and silver soldering provide good thermal contact. The disadvantage is the high temperature softens and deforms the structure. If the precise shape of a part needs to be kept, e.g., a flange, soft soldering maybe preferred. Most of the soft solder become superconducting below 1 K which degrades the thermal performance. The non-superconducting soft solder can be used in stringent requirements.

Surface condition affects the thermal performance not only through the quality of the joints but also through the ability of receiving and emitting energy via radiation. Unless a surface is intended to absorb electromagnetic radiation of a certain characteristic, it is always good to have the surface clean and polished in order not to

receive radiation from other sources. This is especially true for the radiation shields.

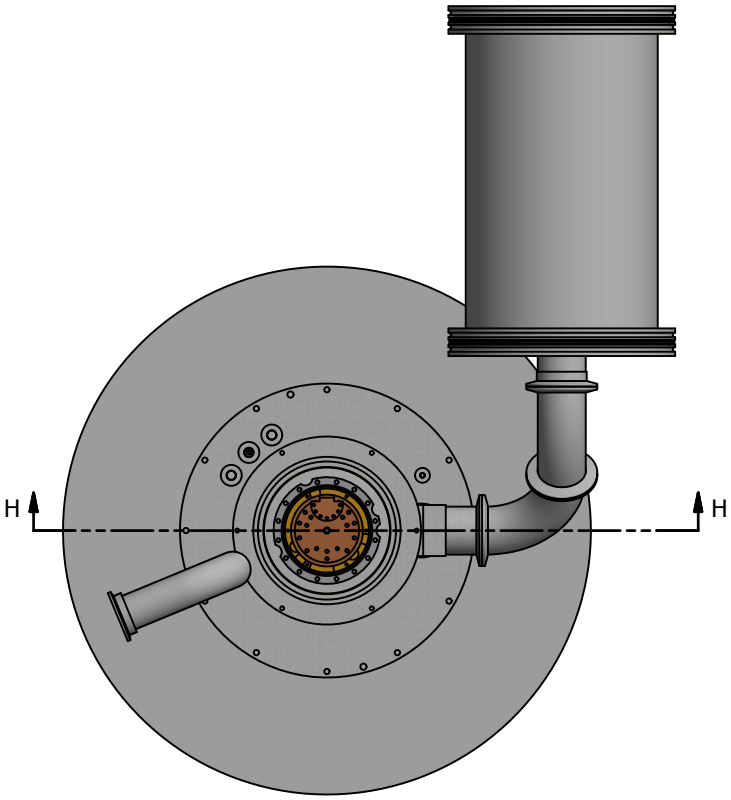
3.2.1.5 Vacuum

The vacuum level in the IVC of the fridge ($\sim 1 \times 10^{-6}$ mbar) does not impose stringent requirements on the materials used inside IVC or the surface treatment. At low temperatures, the cryo-pumping of the cold surface helps to maintain good vacuum inside IVC. However, proper selection of the materials and design of the structure inside IVC would help reduce the pumping time. Thus the guidelines for high vacuum systems still apply to the dilution fridge. Porous material or material with high vapor pressure should be avoided as much as possible inside the IVC. And the obstruction to the air flow should be as little as possible. There should be venting holes for any otherwise enclosed volume.

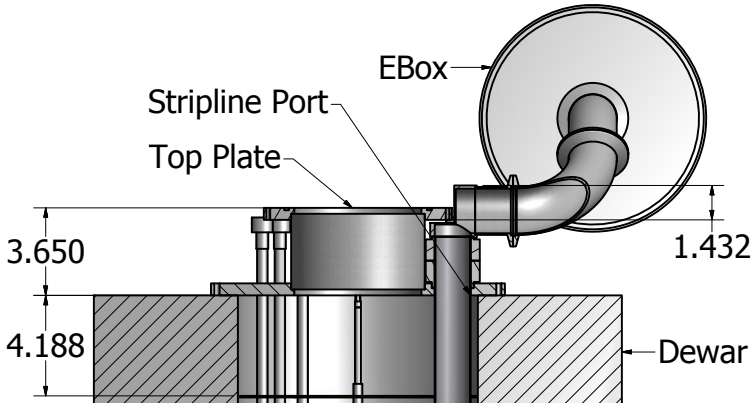
3.2.2 The top plate

The top plate (Fig. 3.7) consists of two concentric flanges connected by a stainless pipe between them. There are a number of ports for liquid N_2 and ^{4}He transfer as well as exhaust, liquid helium level meter, and the striplines on the lower (larger) flange. The top plate connects the fridge and the Dewar and seals the Dewar when the two o-ring flanges are tightened. Though a simple structure, a finite element analysis was still performed to make sure it could bear the load.

There is the need to have as many facilities as possible to test the detectors and towers for the SuperCDMS SNOLAB experiment because of the large number of detectors that will be deployed. In the current configuration, the IVC of the K100 fridge is a few inches short to host a SuperCDMS SNOLAB tower with a full set of



(a) Top view of the fridge



(b) Cross section of view of H-H

Figure 3.7: Top and cross-sectional view of the fridge top plate.

detectors loaded. The height of the IVC can be extended either by using a taller Dewar or raising the dilution unit. In the latter case, a taller top plate can be used, extending the height 3.650 inch by a few ($<$ 4.188, Fig. 3.7) inches without changing the radiation baffle shield on the dilution unit.

An additional benefit of extending the top plate is there would be more room for the tubing containing the striplines. And more striplines may be installed. As shown in Fig. 3.7(b), the obstruction of the top flange can be eliminated if the flange could be raised by a minimum height of 1.432 inch.

3.2.3 IVC

The design of the IVC (Fig. 3.8) was with the intention to maximize the usable space for hosting the CDMS detectors, tower, and cold electronics. To avoid possible interference between the IVC and the wall of the fridge Dewar, the diameter of the bottom indium seal flange was reduced and the indium seal was placed inside the wall of the middle section of the IVC. Because additional structures (a flange) are needed inside the IVC to attach the 4 K radiation shields and to anchor the striplines, the reduction of the flange size did not reduce the usable space inside IVC. However, it brought in some inconvenience that the bottom lid has to be bolted from below the flange.

The spacing of the bolts for the bottom and middle indium seal flanges was $\sim$1 inch, similar to that of the existing flanges. When the IVC was in vacuum and the fridge Dewar filled with liquid N_2, the walls of the IVC would subject to a differential pressure of 1 atm plus the pressure of the liquid N_2. A finite element stress analysis was performed to select the correct thickness of the side walls and the bottom lid.

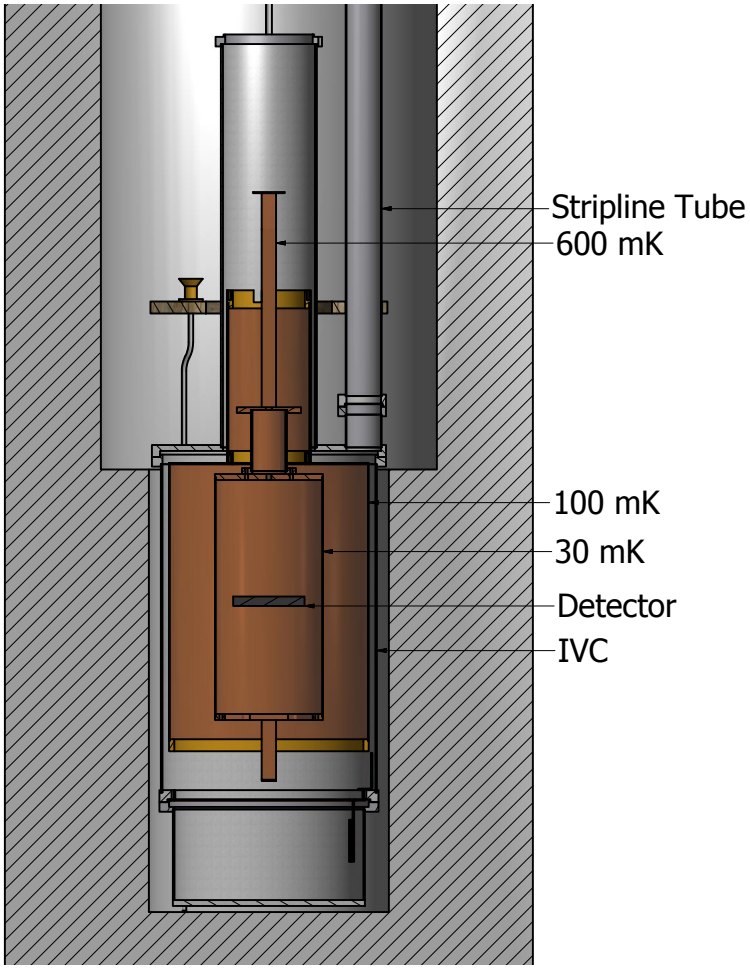


Figure 3.8: Cross-sectional view of the IVC (the stainless can inside the Dewar) and the copper radiation shields. The dark hatched block shows the size and the position of a 3 inch diameter CDMS-II type detector.

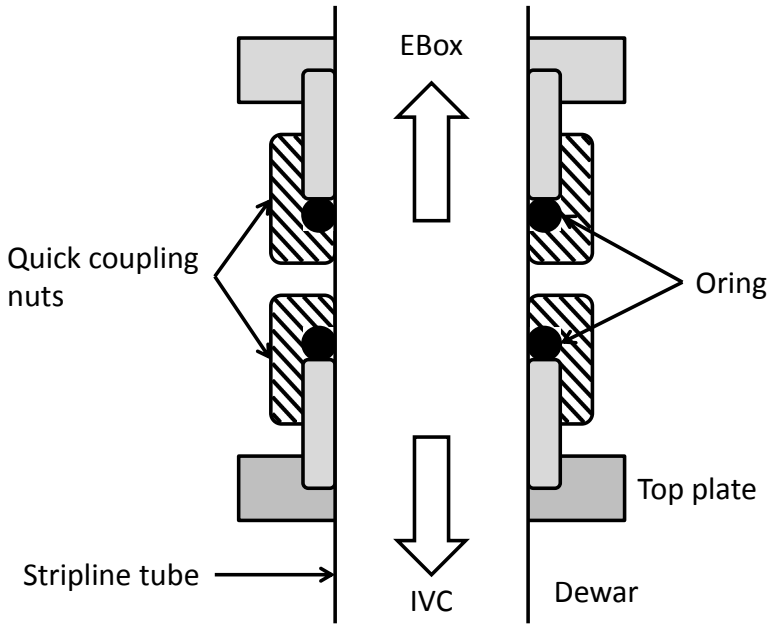


Figure 3.9: Cross-sectional schematic of the stripline port at the top of the fridge (not to scale). The top o-ring seal connects the tube from the EBox and the bottom one seals the fridge Dewar.

3.2.4 Stripline port

The width of the CDMS stripline is 1 inch, which is wider than all the original ports in the dilution unit. An additional port of outer diameter 1.5 inch was added on the edge of the IVC (Fig. 3.7 and Fig. 3.8). At the bottom, the stripline tube is connected to the IVC with an indium seal (Fig. 3.8); at the top, the tube went through two quick coupling o-ring seals which seal the Dewar and the Ebox simultaneously (Fig. 3.9).

The purpose of the lower o-ring seal in the design in Fig. 3.9 is to provide a sliding seal for the stainless tube going through it so that the stress induced by the thermal expansion/contraction of the tube can be relieved. The volume for cryogens (liquid N_2 and ^{4}He) of the fridge Dewar does not require high vacuum. The highest vacuum needed is $\sim$1 mbar when pumping the residual liquid N_2 out of the Dewar at 77 K.

The o-ring seal is more than sufficient for the desired vacuum level. The IVC and the EBox do require high vacuum, which the upper o-ring seal takes care of. Other designs of the stripline tube were also considered. Since the section of the dilution unit parallel with the stripline tube is mainly constructed in stainless steel, a rigid connection between the IVC and the top plate may also be acceptable if no significant stress builds up when assembling the stripline port at room temperature. This presents stringent requirements on the accuracy of the design, fabrication, and assembling. It is more practical to use a flexible connection such as an o-ring seal. Metal bellows were also considered besides the rigid stainless tube. Though the bellows are most flexible when the two ends are not constrained, it is not the case when the two ends are fixed. It is possible to tune the bellows to have zero stress induced by differential pressure and thermal expansion at a given vacuum level and temperature. However, at all the other differential pressures and temperatures, the fridge will be under stress due to either expansion or contraction, which is not desirable.

Thermal loads introduced through the stripline port and the striplines need be dealt with carefully not to impair the performance of the fridge. Assuming the warm (300 K) and cold (4 K) ends of the stripline port are two black body disks of diameter 1.5 inch and the inner wall the stripline tube reflects all the radiation, the thermal load introduced by radiation is 0.48 W. If the inner wall of the stripline port is blackened and absorbs all radiation intercepts with it, for a 40 inch long tube, the radiation that reaches the cold end is 17 μW. Blackening the inner wall of the tube could reduce the thermal load induced by radiation significantly. However, we did not have an easy-to-apply recipe to blacken the inner wall of the tube, which absorbs the radiation from a 300 K surface efficiently. Instead, we decided to use conventional radiation

baffles in the stripline tube to reflect out the radiation. The baffle structure is shown in Fig. 3.10. Striplines are sandwiched between two half-circled baffle plates and rectangular spacers, all of which are made of OFHC copper. Seven such assemblies were installed in the stripline tube between 300 K and 4 K, providing fourteen layers of radiation shields, whereas in a helium Dewar, five layers of radiation shields were considered to be sufficient in literature, as was the case for the fridge Dewar.

3.2.5 Tower installation and heat sinking

The tower is installed upside down in the fridge as shown in Fig. 3.11, with the 30 mK stage directly bolted to the 30 mK can of the fridge (see Fig. 3.8). The detector stack is at the top of the installed tower, enclosed by the 30 mK can. The 100 mK and 4 K floors are thermally sunk to the 100 mK can and the IVC, respectively. Ring shaped planar radiation shields are attached to the bottom flange of the 100 mK can and the ears of the 100 mK tower floor, sealing the opening between the can and the tower while letting the tower pass through. Similar shields are used to connect the IVC and the 4 K tower floor. Larger opening holes on the radiation shields at each temperature stage are sealed using copper tape to minimize the infrared radiation leakage to the detector, which is believed to degrade the detector neutralization. However, there is a fair amount of air enclosed inside the 30 mK can and the 100 mK can, hermetically sealing the radiation shields would prolong the pumping time at the least, or worse yet, impede the cooldown to the desired temperature completely. If ^{4}He exchange gas is used during the cooldown from 77 K to 4 K, the residual ^{4}He exchange gas would become superfluid below 2.2 K, plate out on the cold surfaces, and thermally short different temperature stages and the phonon sensors. Keeping a good venting path

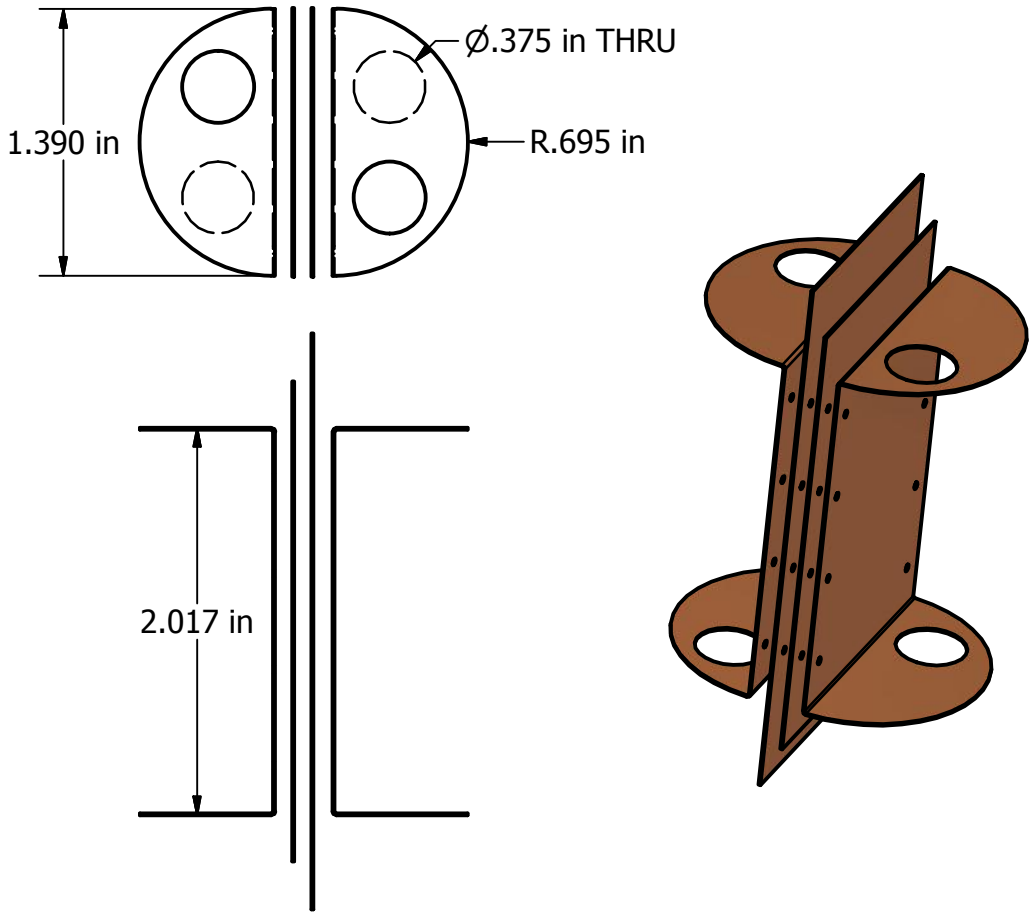


Figure 3.10: Stripline baffles. Top: Top view of the baffles. The striplines are sandwiched between the half-circle baffle plates and the rectangular spacers. Two $\frac{3}{8}$ inch through holes are located on each half-circle baffle plate with no overlapping in the axial direction to ensure good pumping speed while blocking line-of-sight radiation. Bottom left: Front view of the baffle assembly. Right: 3D mock up of the baffle assembly.

for the inner volume is equally important to sealing the IR radiation. Nevertheless, the configuration we have been using works reasonably well in pumping efficiency and IR blockage, though in more demanding situations more careful balance between the two may be needed.

Thermal links with two lugs at the ends and high thermal conductivity copper wires in between are installed between the tower and the cans at the 100 mK, 600 mK, and 4 K stages to enhance the thermal connection. Initially the 600 mK strap heat sunk to the still was not installed, since earlier measurements [108] showed that the graphite tubes joining the tower floors could provide enough thermal isolation. However, we could not cool the detector below the T_c's of the phonon sensors with a much lower base temperature at the mixing chamber in our first runs and the strap was then added. The majority of this heat load is from the JFETs which are mounted at the 4 K stage but work at ∼140 K to minimize the noise. Each JFET dumps ∼5 mW to the 4 K stage, which is not acceptable for a larger scale experiment. Integrating the high-electron-mobility transistor (HEMT) into the CDMS charge readout circuit is currently under investigation at Berkeley, which would vastly reduce the heat load.

The SQUET cards are installed at the bottom (Fig. 3.11) of the tower module. The JFET charge amplifier is mounted at the 4 K floor while the SQUID amplifier is mounted at 600 mK stage, both for minimizing the noise and improving the SNR. Aside of the heat load from the JFET by heat conduction, as the hottest component in the IVC, the JFET working at ∼140 K emits IR radiation peaked at ∼34 meV, which is energetic enough to ionize the shallow impurity levels of the crystal. The JFET is shield by an IR absorber on the SQUET card, however, to further prevent possible IR leaks, two L-shaped IR absorbers are added on the 4 K radiation shield inside the IVC.

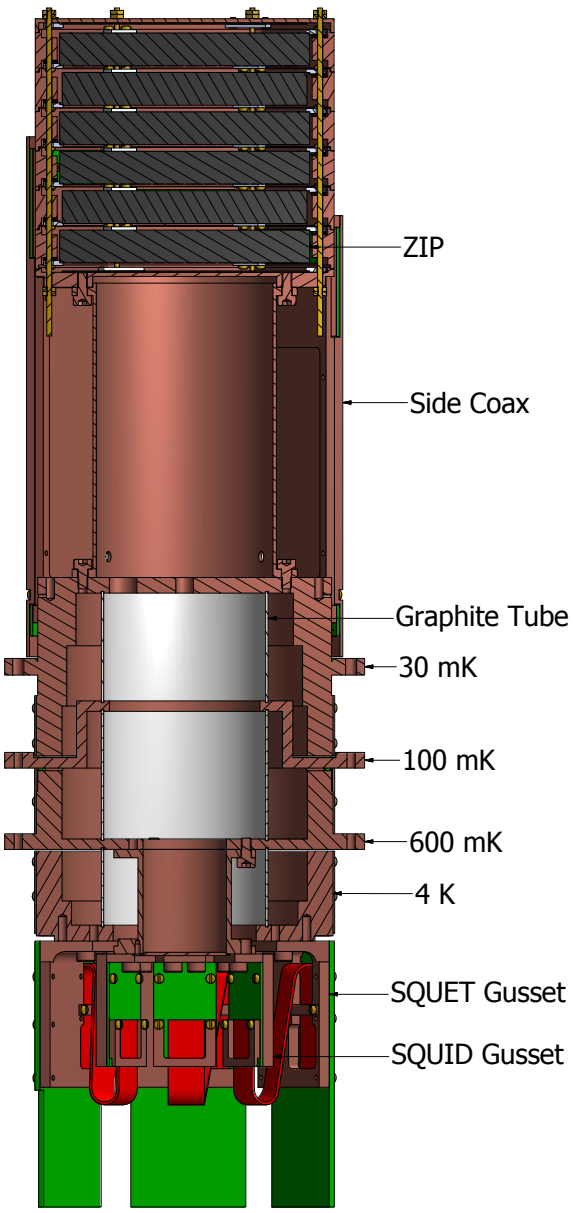


Figure 3.11: Cross-sectional view of the CDMS-II tower. The tower is positioned upside down as how it is mounted in the CDMS test facilities. The dark disks at the top are the ZIP detectors. Four temperature stages at the bottom are separated by the low thermal conductivity graphite tubes. The SQUET card is mounted at the very bottom with the charge amplifier at the 4 K stage and the SQUID amplifier at the 600 mK stage. *Figure adapted from SolidWorks 3D model by P. Wikus.*

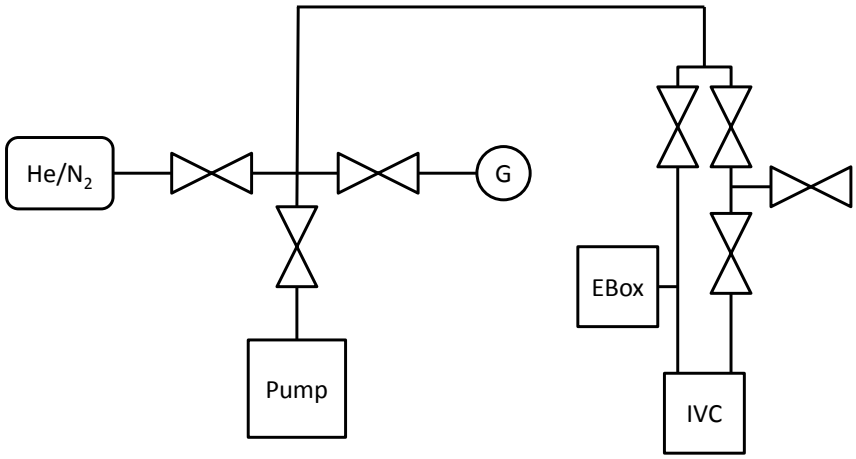


Figure 3.12: Schematic of the vacuum system of the K100 fridge. Valves, vacuum gauge, and various other components are shown.

3.3 Gas handling system

3.3.1 Fridge vacuum system

Figure 3.12 shows the schematic of the vacuum system of the fridge. Because of the much enlarged IVC, the added stripline port, and the EBox, there was the concern that pumping out the gas inside the IVC might take an unacceptably long time. This section presents the calculation that was performed for selecting the vacuum components especially the vacuum pump. The vacuum system model is highly simplified, however, the results should still reflect the pumping efficiency within a factor of a few if the abstraction is reasonable.

3.3.1.1 Volumes that need to be pumped out

Table 3.1 lists the dimensions and volumes of the components that need to be pumped to high vacuum. The majority of the volume is from the IVC, stripline tube, and

Table 3.1: Various components that need to be pumped out in the K100 fridge.

Part	Diameter (inch)	Height (inch)	Volume (in^3)	Volume (L)
IVC Low	8.87	18.7	1155.5	18.9
IVC High	3.83	17	195.9	3.2
Stripline Tube	1.44	42	68.4	1.1
Half Nipple, QF50	2	1.58	5.0	0.08
Elbow, QF50, 45°	2	4	12.6	0.2
Bellow, QF50	2	4	12.6	0.2
EBox	8	13.22	664.5	10.9

the EBox (see Fig. 3.8 and Fig. 3.7). The rest is from the necessary pieces to make connections. Adding the volume together gives the total volume 34.6 L.

3.3.1.2 Pumping speeds at various pressures

In the original configuration, the pumping line inside the fridge is a wire tube with ID (inner diameter) 0.5" and length 30". Outside the fridge are a ∼0.5 m long ID 0.5" hose and a 1.5 m long ID 1" hose connected in series. We first calculate the conductance of a tube for viscous and molecular flows.

The flow rate of viscous flow through a long straight tube is given by the Poiseuille equation

$$Q = \frac{\pi d^4}{128\eta l}\frac{(P_1+P_2)}{2}(P_1-P_2) = C(P_1-P_2), \tag{3.1}$$

where Q is the flow rate in unit of pressure times volume, d is the diameter of the tube, l is the length of the tube, η is the dynamic viscosity of the fluid, P_1 and P_2 are the pressures at the two ends of the tube, and C is the conductance. The conductance for air at 0°C is

$$C(\mathrm{L/s}) = 1.38\times 10^6 \frac{d^4}{l}\frac{(P_1+P_2)}{2}. \tag{3.2}$$

The conductance for molecular flow inside an infinitely long circular tube is

$$C = \frac{\pi}{12} v \frac{d^3}{l}, \tag{3.3}$$

where v is a constant. For air at 22°C this becomes

$$C(\mathrm{m}^3/\mathrm{s}) = 121 \frac{d^3}{l}, \tag{3.4}$$

which is independent of pressure. For tubes of arbitrary length, Clausing's solution gives

$$C = \frac{a' v}{4} A, \tag{3.5}$$

where A is the cross-sectional area. For air at 22°C,

$$C(\mathrm{L}/\mathrm{s}) = 1.16 \times 10^5 a' A. \tag{3.6}$$

The ratio of the mean free path to a characteristic dimension of the system, say, the diameter of a pipe is called Knudsen's number K_n:

$$K_n = \frac{\lambda}{d}. \tag{3.7}$$

When $K_n < 0.01$, the flow is viscous; when $K_n > 1$, the flow is molecular flow; when $0.01 < K_n < 1$, the flow is called transition flow and is not well understood.

For most gasses, the mean free path is about 1 cm at the pressure 1 Pa and 1×10^{-2} cm at pressure 60 Pa. Assume the pressure at the inlet of the turbo pump is very low and all the pressure drop are along the pumping hoses, and assume the

Table 3.2: Conductance of a 1 m long and 1.25 cm diameter tube at various pressures.

d=1.25 cm, l=1 m	$\overline{P}$(1 Bar)	$\overline{P}$(0.3 mBar)	$\overline{P}$(0.005 mBar)	$\overline{P}$(10^{-6} mBar)
C(L/s)	3369	1	0.24	0.24
Flow	viscous	viscous	molecular	molecular

thinner tube in the pumping line is 1 m long with an ID 0.5", then the conductance is given in Table 3.2.

3.3.1.3 Pumping time

The equation describing the pressure in the chamber is

$$SP - Q = -V\frac{\mathrm{d}P}{\mathrm{d}t}, \tag{3.8}$$

where S is the pumping speed, P is the pressure, Q is the gas flow entering the chamber via outgassing or leak, and V is the volume of the chamber. The solution is

$$P = P_0 e^{-St/V} + \frac{Q}{S}. \tag{3.9}$$

If we assume there is no leak or outgassing, namely $Q = 0$, then the pumping time is

$$t = \frac{V}{S}\ln\frac{P_0}{P}. \tag{3.10}$$

With the conductance in Table 3.2, we have the pumping time in viscous flow range

$$t_v = \frac{34.6}{1}\ln(\frac{10^5\mathrm{Pa}}{10\mathrm{Pa}}) = 276\,\mathrm{s}, \tag{3.11}$$

and in the molecular flow range

$$t_m = \frac{34.6}{0.24} \ln(\frac{1\text{Pa}}{10^{-4}\text{Pa}}) = 1151\,\text{s}. \tag{3.12}$$

In the pressure range 1 Pa to 10 Pa we take the conductance as 0.5 L/s, the pumping time is

$$t_t = 34.6 \times 2 \times \ln 10 = 159\,\text{s}. \tag{3.13}$$

The total pumping time is

$$t_v + t_t + t_m = 1586\,\text{s} = 26\,\text{min}. \tag{3.14}$$

3.3.1.4 Conclusion

At all pressures except the starting moment of the pump, the bottleneck of the pumping speed lies in the thin wire tube in the fridge through which the IVC is pumped. A pumping time of 0.5 hours will be sufficient to pump the IVC to the working pressure assuming there is no leak or outgassing. If the IVC is pumped from the larger stripline tube with an ID of 1.5", the pumping time will be 10 times ($3^3 = 27$) less. But the baffles in the tube decrease the pumping speed. To measure the pressure in IVC, it may be necessary to install a pressure gauge near the EBox. Any turbo pumping station with a pumping speed of several tens L/s and the ultimate pressure below 10^{-6} mBar (Torr) would work fine.

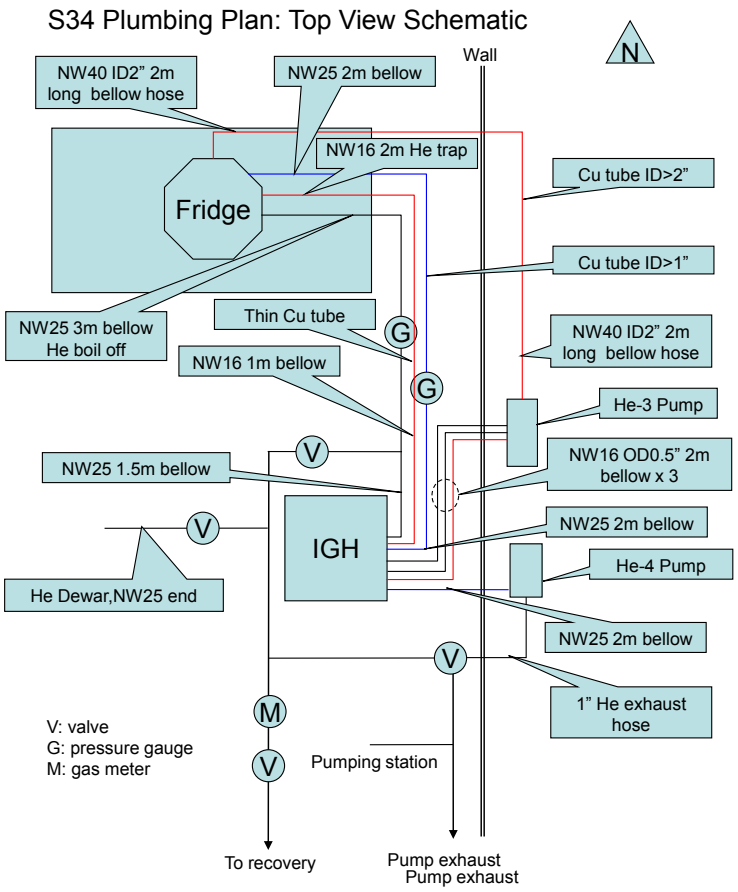


Figure 3.13: Schematic of the top view of the gas handling plumbing in the lab.

3.3.2 Fridge gas handling plumbing

A similar pumping speed analysis may be performed for the fridge gas handling system shown in Fig. 3.13. However, the fridge was used in a more spread out configuration before being moved into S34, and we were certain that the compact configuration in S34 would work fine. So the analysis was omitted.

3.4 Electronics and data handling

Figure 3.14 shows the signal and data flow in the UMN test facility. All the cold electronics (below room temperature) are the same as the ones used in the CDMS Soudan experiment. The warm electronics used in Soudan are substituted with a compact single-board module called Detector Control and Readout Card (DCRC) designed and fabricated at Fermilab. DCRCs are plugged into the electronics box (EBox) directly without any cables between them, minimizing the signal attenuation and electromagnetic pickup. Ethernet cables are used to connect the DCRCs and the DAQ computer. A power over Ethernet (POE) injector is inserted between each DCRC and the DAQ computer to provide power to the DCRC. One DCRC can only control and read out a single CDMS-II type detector, which has two charge channels and four phonon channels. DCRCs are daisy chained by a timing link to work synchronously for reading out more channels. A Ethernet switch was then inserted between the POE and the DAQ computer to route the communication.

The DCRC has all the functionalities of the Soudan warm electronics including detector biasing, and signal amplifying, digitizing, and triggering. A Labview software package developed at Fermilab is used to control the DCRC and read out data. The analog signals are digitized continuously by the DCRC and stored in a circular buffer which can hold 2 s of data. Addresses of the triggers in the buffer are stored in a trigger queue. The Labview DAQ software running on a PC first queries and reads out the trigger queue. For a given trigger, the address range of the whole pulse is then calculated by adding the number of data samples before and after the trigger address. Then the DAQ software requests the DCRC for the portion of data in the

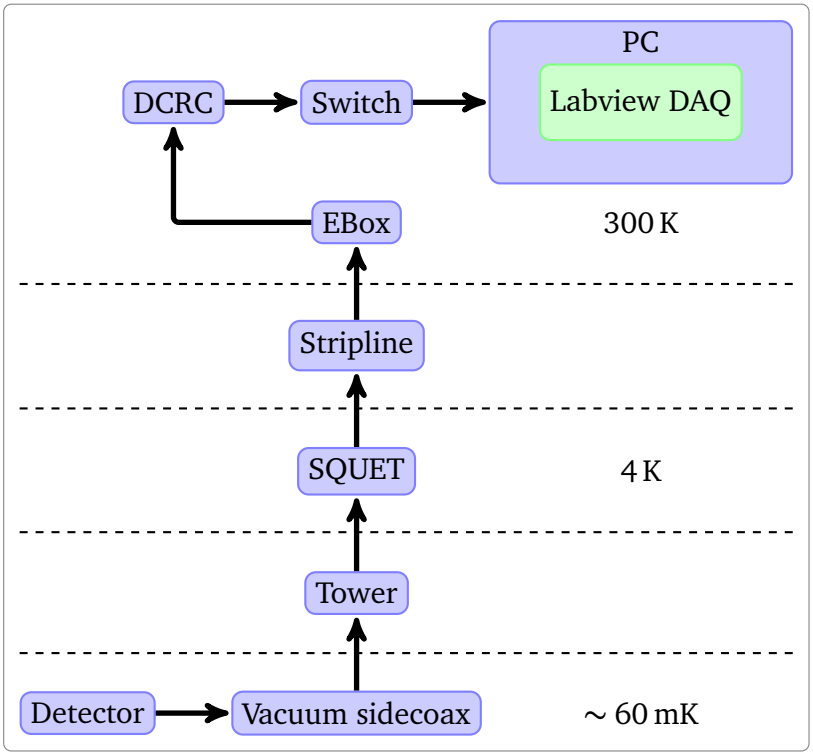


Figure 3.14: Signal/data flow in the UMN test facility.

address range of the pulse. The readout pulses are then combined with the header information containing the channel index, the trace length, and the time stamp etc. and written to the raw data file on the harddrive of the DAQ computer. The event trigger rate in DCRC is usually higher than what the DAQ software can handle, and a fraction of triggers are dropped.

The raw data files are transfered to the Linux cluster of the physics department and processed using the CDMS data processing package `CDMSBats` following the pipeline described in Sec. 5.3.1. The data processing is parallelized using the Condor system on the cluster.

Chapter 4

100 mm Ge Crystal Ionization Test

To reach a ten times smaller WIMP-nucleon scattering cross section than the current limit, the SuperCDMS SNOLAB experiment will deploy ∼200 kg of germanium detectors. It is then necessary to use larger detectors to reduce the number of detectors needed, and thus reduce the overhead of detector fabrication, testing, and commissioning. A smaller number of detectors also reduces the effort of data analysis. In addition, larger detectors provide better surface event rejection capability because of their larger volume to surface ratio.

Due to the difficulty of quality control at large radius, the largest detector grade high purity germanium crystals that can be grown are of diameter 100 mm at the moment. Before using them to develop the next generation detectors for the SuperCDMS experiment, a series of ionization tests were performed to make sure they meet the requirements. In this chapter, we present the very first ionization tests with two four-ionization-electrode 100 mm diameter test devices. These tests showed that the charge collection efficiencies of the two test devices were consistent with those

Figure 4.1: Side view of a 100 mm diameter Ge crystal. *Image from Paul Brink.*

measured with CDMS-II detectors, and these crystals can be used to develop the next generation SuperCDMS detectors.

4.1 100 mm ionization test devices

The first ionization tests were performed with two test devices G101a and G102 fabricated at the Stanford Linear Accelerator Center (SLAC). Both were fabricated with [100] crystals grown by Umicore. Figure 4.1 shows a picture of the Ge crystal as it arrived. The physical properties of the two crystals are listed in Table 4.1. After shaping and polishing, the dimension of the crystal was 100 mm diameter and 33.3 mm thick. Figure 4.2 depicts the top view dimensions of the 100 mm [100] Ge crystals.

Four concentric electrodes each covering equal surface area were patterned by photolithography on the top surface. A single ground plane grid was patterned on the bottom side. The electrode layout and structure of the grid is shown in Fig. 4.3. Figure 4.4 shows the patterning process of the electrodes and the ground plane:

Table 4.1: Physical properties of the first two 100 mm diameter Ge crystals tested.

			$N_d - N_a$ (10^{10} cm^{-3})		Mobility (10^3 cm^{2}V^{-1}s^{-1})		Dislocation (cm^{-2})	
	Orientation	Type	Head	Tail	Head	Tail	Head	Tail
G101	[100]	N	0.55	0.90	20.2	22.0	2400	3000
G102	[100]	N	0.90	1.50	22.0	24.4	3000	3816

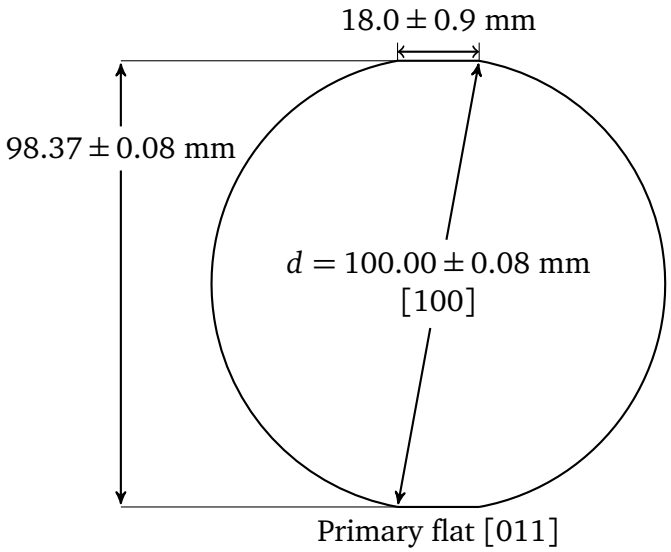


Figure 4.2: Detailed dimensions of the 100 mm diameter Ge crystals. *Figure adapted from the work of Paul Brink.*

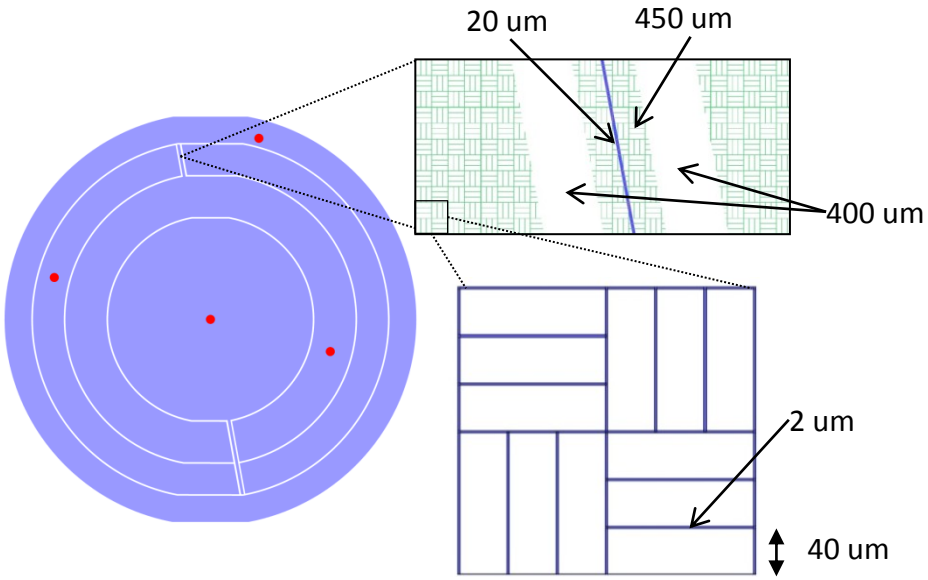


Figure 4.3: Charge electrode layout and grid structure of G101a and G102. Both the top and the bottom surfaces of the crystal are covered with the grid of the pattern shown at the lower right corner of the figure, which has 2 μm wide traces with a 40 μm pitch. The top surface is divided into four concentric electrodes with equal area by the 400 μm wide trenches. The radii of the four electrodes from the center of the crystal to the center of the 400 μm wide trench are 25, 35.4, 43.3, and 50 mm. The electrodes are named Q_1-Q_4 from the center to the edge. The signals from Q_1 and Q_2 are brought out with the 20 μm wide solid traces on top of the 450 μm wide grid channels through the gaps on Q_2 and Q_3. The red dots depict the locations and area illuminated by the four ^{241}Am sources each collimated at approximately 20 Bq. *Figure adapted from the work of Astrid Tomada.*

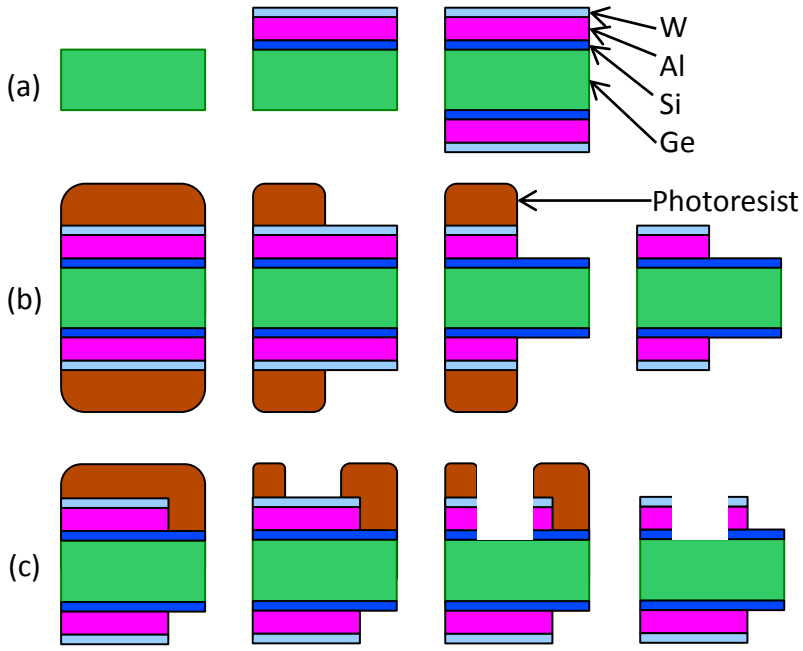


Figure 4.4: Photolithography steps to define the charge electrodes and the ground plane for the 100 mm ionization devices. (a): Deposition of 40 nm armorphos Si, 20 nm Al, and 40 nm W. (b): Defining and wet etching the grid on the top and bottom surface. (c): Defining and dry etching the trenches separating the electrodes on the top surface. *Figure adapted from the work of Paul Brink.*

(a) Thin films of 40 nm amorphous Si, 20 nm Al, and 40 nm W are deposited on the top and bottom surfaces of the Ge substrate after cleaning. The thickness of tungsten is thick enough for wirebonding but still sufficiently thin for attaching thermistors to measure the thermal phonons.

(b) Both sides are coated with photoresist, exposed to UV light with the grid pattern mask, developed, and then wet etched to remove the W and Al layer. The photoresist is stripped at the last step.

(c) The top side is coated with photoresist, exposed to UV light with the trench mask, developed, and dry etched to remove the W, Al, and the conductive amorphous Si layer. To make sure the electrodes are electrically isolated. a few hundred nanometers of Ge substrate is removed by the dry etching. The photoresist is stripped at the last step.

The solid traces bringing the signals of the inner electrodes out for wire bonding can be deposited either before or after Step (c) in the above process.

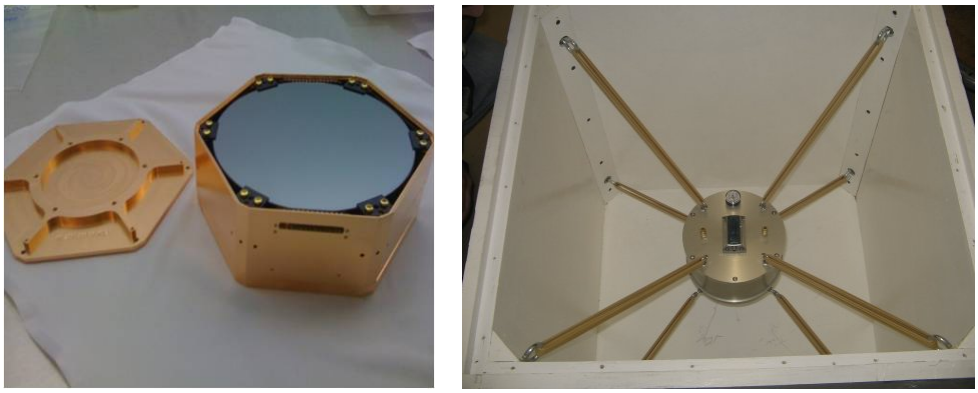
The ionization test devices were mounted in the 4 inch Cu housings designed and fabricated in SLAC, held by the Cirlex clamps from both sides. The 4 inch detector housings were scaled up from the 3 inch ones used in the CDMS II experiment. The electrodes and the ground plane were wire bonded to the detector interface board (DIB). Because the 100 mm ionization devices doubled the number of charge channels on a CDMS II detector, two DIBs were mounted on opposite faces of the 4 inch detector housing. Table 4.2 shows the electrode connection of G101a and G102. Note that the phonon channels on DIB2 of G101a were not grounded. We suspected this resulted in the fact that the signals on Q_1 and Q_3 were much weaker than those on Q_2 and Q_4. But this may not be conclusive since G101a also had other

Table 4.2: Electrode connections of ionization test devices G101a and G102. The ground plane of G101a only connected to the phonon Channel A on DIB1. Both DIBs on G102 were grounded through phonon Channels B, C, and D.

	G101a		G102	
Pin	DIB1	DIB2	DIB1	DIB2
A-	GND			
A	GND			
B-			GND	
B			GND	
C-				GND
C				GND
D-				GND
D				GND
LED1	LED1	LED1	LED1	LED1
LED2	LED2	LED2	LED2	LED2
Qo	Q_4	Q_3	Q_4	Q_3
Qi	Q_2	Q_1	Q_2	Q_1

problems. This issue was corrected on G102 and the phonon channels on both DIBs were grounded. Figure 4.5(a) shows a wire bonded detector sitting in the copper housing.

After an electrical check at temperature around 400 mK in the Stanford ^{3}He fridge, the 100 mm ionization device were then sent to the UMN test facility in a hermetic shipping vessel filled with over pressured N_2 to prevent radon contamination. The shipping vessel was suspended using rubber bands in a wooden shipping container to protect it from shock damage. Figure 4.5(b) shows the shipping vessel and the shipping container.



(a) (b)

Figure 4.5: (a): A 100 mm ionization test device in the copper housing. The DIB is behind the slit in the front face. (b): A 100 mm detector shipping vessel held by rubber bands in the shipping container. The ionization test device is inside the shipping vessel. A pressure gauge and a GPS module are mounted on top of the shipping vessel.

4.2 ^{3}He fridge screening

As what was done to normal CDMS detectors, the 100 mm ionization devices were screened in the Stanford ^{3}He fridge for potential crystal, fabrication, or wiring problems before they were shipped out for further tests.

There was a visible scratch on G101a caused by the OGP scanner while inspecting the device, along the radial direction stretching from the center of Q_2 to the edge of crystal. The depth of the scratch was measured around 20 μm. Although the scratch did cut through the outer three electrodes and damage their integrity, it was not expected to cause electrical problems. However, in the ^{3}He fridge run [109], it was found that the outer three electrodes had much lower freeze-out temperature, ∼20 K, across the substrate than the innermost channel, whose was ∼60 K. A resistance of 500 MΩ across the Ge substrate was chosen ad hoc as the mark for freeze-out as it

was the maximum resistance the meter used could measure. In addition, substantial surface current leakages between the three outer electrodes were observed at the base temperature ∼400 mK. The resistance of Q_1-Q_2 was measured 59 GΩ whereas that of Q_2-Q_3 was 5.3 MΩ and Q_3-Q_4 was 3.3 MΩ, respectively. The scratch may have to do with the surface current leakage, but the larger leakage across the crystal for the outer channels can not be explained by the scratch alone. Because the Ge substrate is conductive at room temperature, these small leakages can only be measured at low temperatures, which makes it quite involving to diagnose and fix these problems. Despite the undesired leakages on G101a, it was tested with the K100 cryostat.

G102 behaved much more as expected than G101a [110]. All through-crystal and cross-surface resistances were at or above the 100 GΩ meter limit at temperatures below 100 K. At base temperature (∼465 mK), all through-crystal and cross-surface resistances were ∼1 TΩ.

4.3 Cold hardware and ^{241}Am source holder

Existing CDMS three-inch tower and vacuum sidecoaxes were used in the ionization tests. The four-inch detector housing is wider than the three-inch tower by half an inch on each side. Two SCHUBA extenders with Mill-Max pins designed and fabricated at SLAC were used to interconnect the tower and the sidecoaxes as shown in Fig. 4.6(b). A larger copper can holding the tower and the accompanied adapter flange connecting the tower and the can were designed and fabricated at UMN to accommodate the larger tower assembly. The completely assembled tower is depicted in Fig. 4.6(d). The ^{241}Am sources are in the detector housing on the top facing the

four ionization channels of the device below. Two vacuum sidecoaxes are attached outside the detector housing containing the ionization device, bringing the signals to the 10 mK stage of the CDMS three-inch tower, which is under the lower detector housing and inside the adapter plate.

60 keV gammas from four ^{241}Am sources each collimated at $\approx$ 20 Hz were used as the ionization excitation source. The penetration depth of the 60 keV photons in Ge crystal is $\approx$ 1 mm, which generates electron recoils in the detector bulk while still localized to the the nearest ionization electrode, and makes it possible to probe the position-dependent ionization collection efficiency of the Ge crystal. Figure 4.6(a) shows the four ^{241}Am sources in copper housings with the lead collimators facing up. The copper housings were flipped to have the collimators facing the copper plate providing a photon beam directed toward the ionization electrodes. Figure 4.6(c) shows the position of the photon beams before the source holder was mounted on top of the detector. The illuminated areas of the collimated photon beams 1 mm beneath the crystal surface is illustrated as red dots in Fig. 4.3 with the correct scale. The four sources were well separated spatially in order not to introduce additional correlation between the electrodes.

The assembled tower including the 100 mm ionization test device and the ^{241}Am sources was installed in the dilution fridge in the orientation shown in Fig. 4.6(d), and then cooled to a base temperature between 60 mK and 80 mK.

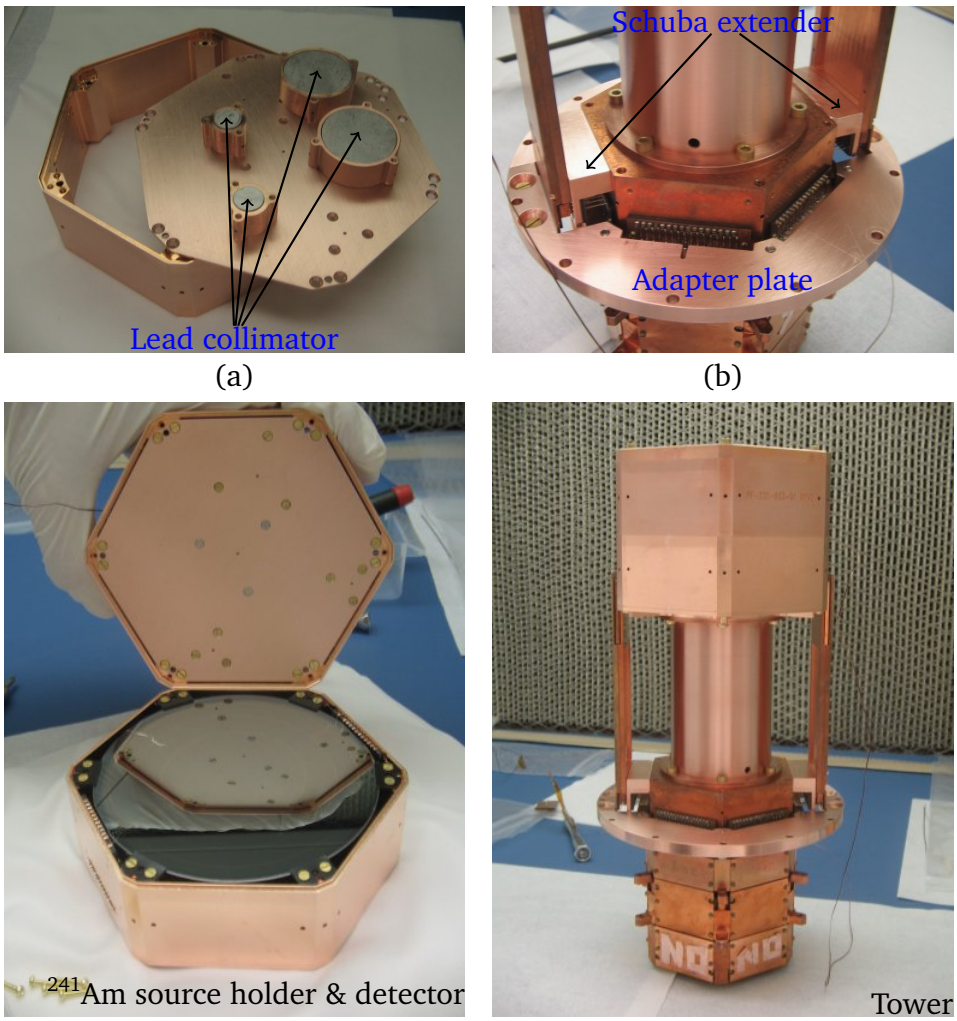


Figure 4.6: (a): ^{241}Am sources in copper housings. (b): Closeup of the SCHUBA extenders and the four-inch adapter plate. (c): Assembled ^{241}Am source holder is being mounted on top of a detector. (d): Assembled tower.

4.4 Data acquisition and processing

The ionization pulses had a very sharp rise time less than 1 μs and a fall time of ∼40 μs determined by the time constant of the integrator. The ionization channel digitization rate was 2.5 MHz (0.4 μs per sample). A trace length of 512 samples with 125 samples (100 μs) before the trigger and 387 samples (154.8 μs) after the trigger was used to take data. The timespan of the trace was long enough to contain the whole charge pulse. The number of samples was integer powers of 2 to utilize the FFT algorithm in data processing. The relative short trace length also reduces the chance of pulse pileups.

The digitized ionization pulses were reconstructed with the charge optimal filter algorithm in data processing. Due to the the capacitive couplings between the electrodes, an ionization signal in one channel would induce similar but smaller signals in other channels. The amplitudes of the induced pulses were proportional to the strength of the coupling between the electrodes and the amplitude of the original signal. This crosstalk effect was corrected to the first order in the Soudan data processing by incorporating a crosstalk template, which was an induced pulse by a signal with unity amplitude, into the charge optimal filter. However, these crosstalk templates were constructed by averaging real pulses selected in a first pass of data processing, where the crosstalk templates were not used and all channels were reconstructed independently. Theoretically, we could use this iterative data processing scheme for the 100 mm ionization test devices. The difficulty was that the crosstalk correction in the charge optimal filter in `CDMSBats` was only for two channels, i.e., Qi and Qo on a CDMS-II detector. A new implementation of the optimal filter would

be needed for the new layout of the ionization electrodes. Instead of rewriting the CDMS data processing package, all channels were reconstructed independently and the crosstalk correction was performed in data analysis.

The four electrodes were wired as two CDMS-II detectors in the way shown in Table 4.2. Q_1 and Q_3 were wired as Qi and Qo of the first detector (DIB1); Q_2 and Q_4 were Qi and Qo of DIB2. The naming of reduced quantities (RQ's) after data processing also followed the wiring convention.

4.5 Ionization signal

The collected data may contain corrupted or degraded ionization pulses due to fluctuations of noise or degradation of detector performance. A number of data quality cuts were defined to remove these bad events and bad running periods of the detector before looking at the ionization signal.

4.5.1 Reconstruction quality

The shape of the ionization pulse was mainly determined by the FET amplifier. For well behaved pulses from the 100 mm device, we didn't expect any difference in pulse reconstruction with respect to the standard ionization pulses from a CDMS-II detector. However, as noise interference and pulse pileups could change the shape of a pulse, the χ^2 of the pulse reconstruction was examined and a charge χ^2 cut was defined to remove poorly reconstructed events. Figure 4.7 shows the charge χ^2 versus ionization energy for the four channels on G101a at -6 V bias. All events above the parabolic cut lines were excluded as badly reconstructed events.

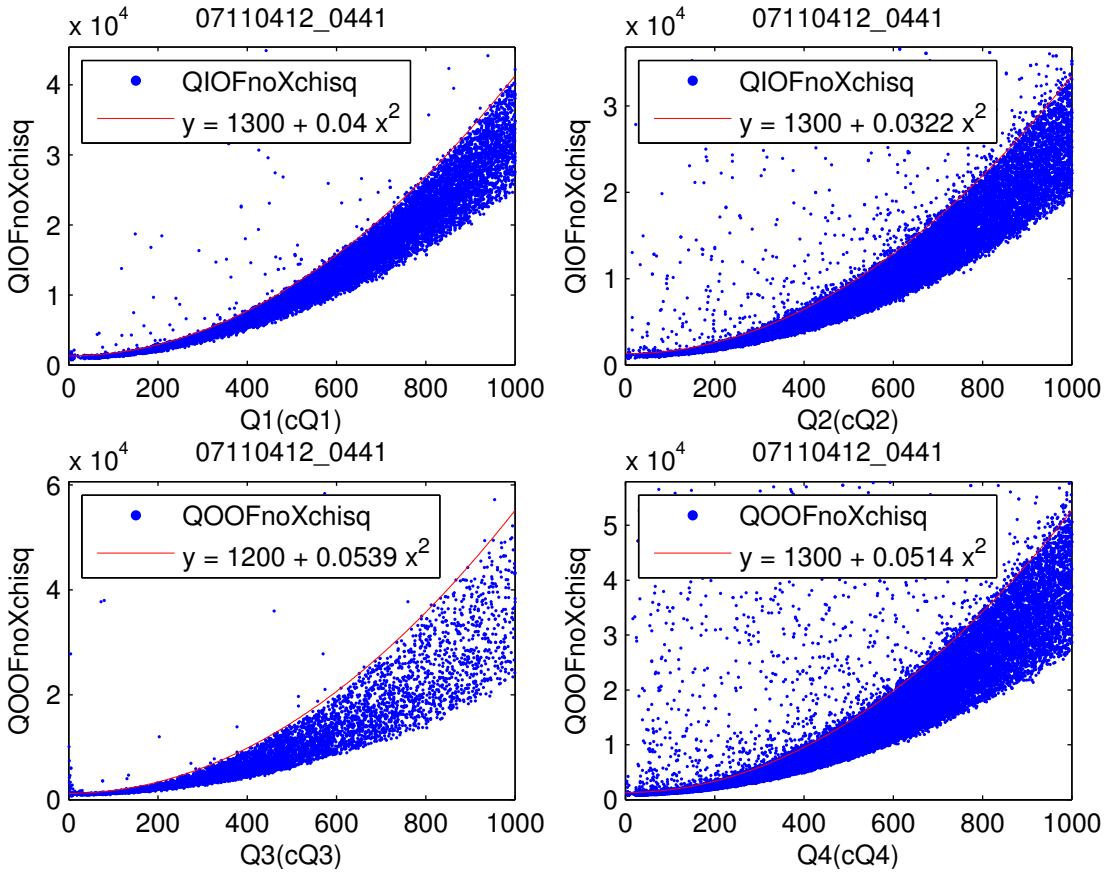


Figure 4.7: Charge χ^2 vs. ionization energy for G101a at -6 V. Only primary events on each channel selected by the cuts cQ# are shown. The ionization energy is in arbitrary units where 60 keV is roughly around 200. The y-axis label `QIOFnoXchisq` and `QOOFnoXchisq` are the charge χ^2 without crosstalk between Qi and Qo. The parabolic cut lines (red) are also shown. The events below the cut lines are selected as good events.

Besides the pulse shape distortion, noise fluctuations while taking data could change the charge χ^2 systematically. The standard deviation of the baseline of the charge pulse was plotted against time. Periods with high noise could also be excluded. This was usually not much of a problem because time scale where the test devices could maintain their performance was relatively short.

4.5.2 Signal stability

The crystal was neutralized by infrared LED flashing before taking data. While taking data, ionized electrons and holes were drifted out of the crystal, leaving the impurity centers charged. Electrons and holes are more easily trapped by the charged impurity centers, decreasing the the amount of charges collected by the electrodes. This detector performance degradation due to charge trapping is called loss of neutralization in CDMS. The detailed mechanism of the loss of neutralization is not fully understood yet. But it is clear that both bulk trapping and near-surface effects contribute. The latter seems to be the dominant cause for the loss of neutralization whereas the former maybe important in the long term. However, we only needed to know when the detector maintained good performance and the detailed physics behind it did not concern us here.

Figure 4.8, 4.9, and 4.10 show the charge signal versus the time the detector was under bias for the four channels on G101 for biases of $-6\,\mathrm{V}$, $-4\,\mathrm{V}$, and $-2\,\mathrm{V}$, respectively. Holes were collected on the electrodes under negative biases. It is clear that the good neutralization time is longer at higher (more negative) biases. They increase from a few minutes at $-2\,\mathrm{V}$ to about half an hour at $-6\,\mathrm{V}$. Charge drifting is more efficient and trapping is less effective at higher biases. These observations agree

well with previous studies [111]. As shown in the plots, the 60 keV line is missing on Q_3 and the event rates for Q_2 and Q_4 are much higher than those of Q_1 and Q_3. More discussion about these issues is in Sec. 4.5.3.

The neutralization plots for G102 at biases of −4 V and 4 V are shown in Fig. 4.11 and 4.12. The good neutralization time of G102 at −4 V is about 0.2 hours, similar to that of G101 at the same bias. At 4 V bias, however, the good neutralization time is much longer. There is no significant signal degradation in the first hour. It is apparent that holes have a much larger trapping rate than electrons.

4.5.3 Ionization spectra

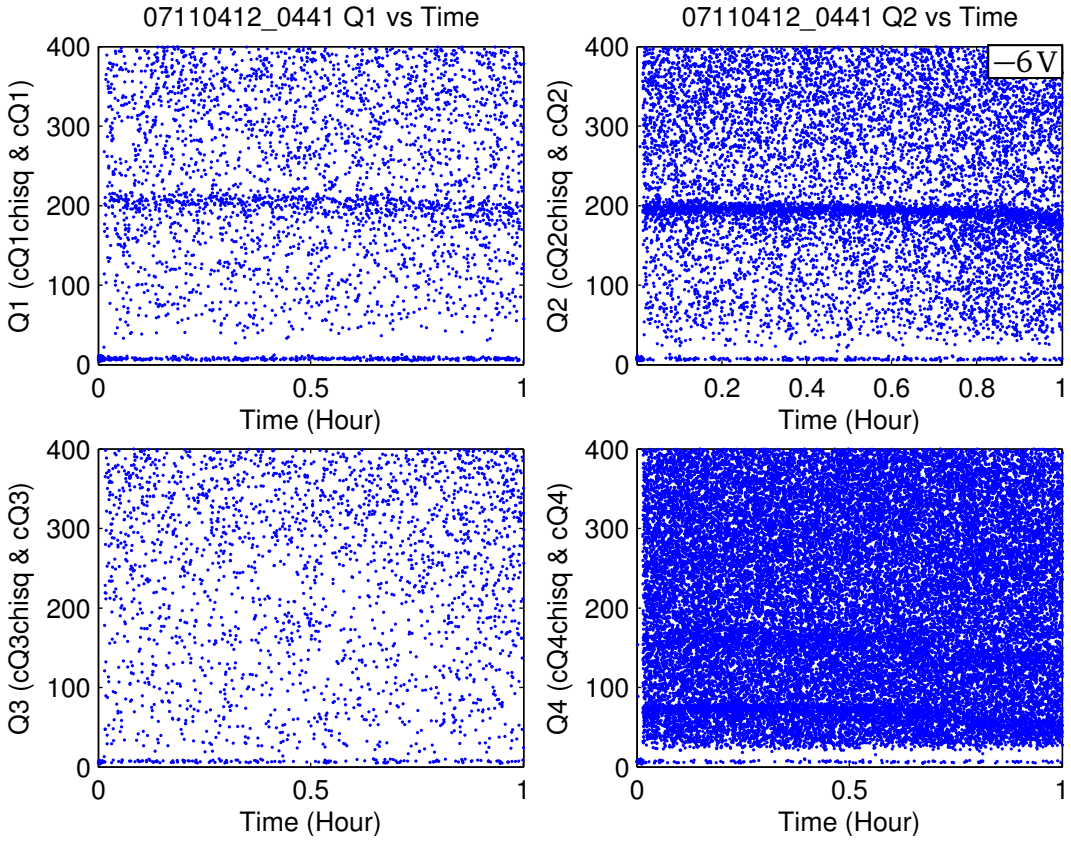


Figure 4.8: Ionization energy vs. bias time for the four charge electrodes on G101a at −6 V bias. The ionization energy is in arbitrary units. The horizontal bands around 200 are the ^{241}Am 60 keV line. The charge χ^2 cut and primary channel cut are applied. The degradation of the signal in the first half hour is not significant.

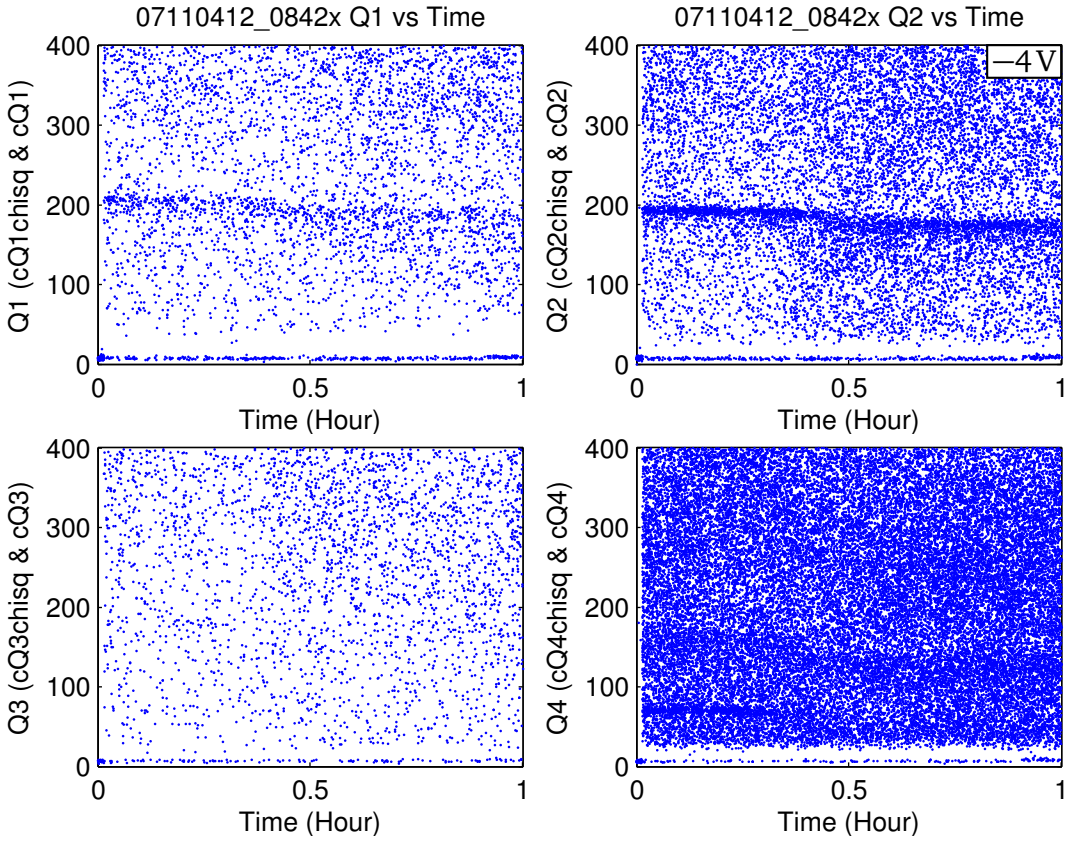


Figure 4.9: Ionization energy vs. bias time for the four charge electrodes on G101a at −4 V bias. The ionization energy is in arbitrary units. The horizontal bands around 200 are the ^{241}Am 60 keV line. The charge χ^2 cut and primary channel cut are applied. G101a can maintain neutralization for about a quarter hour at this bias.

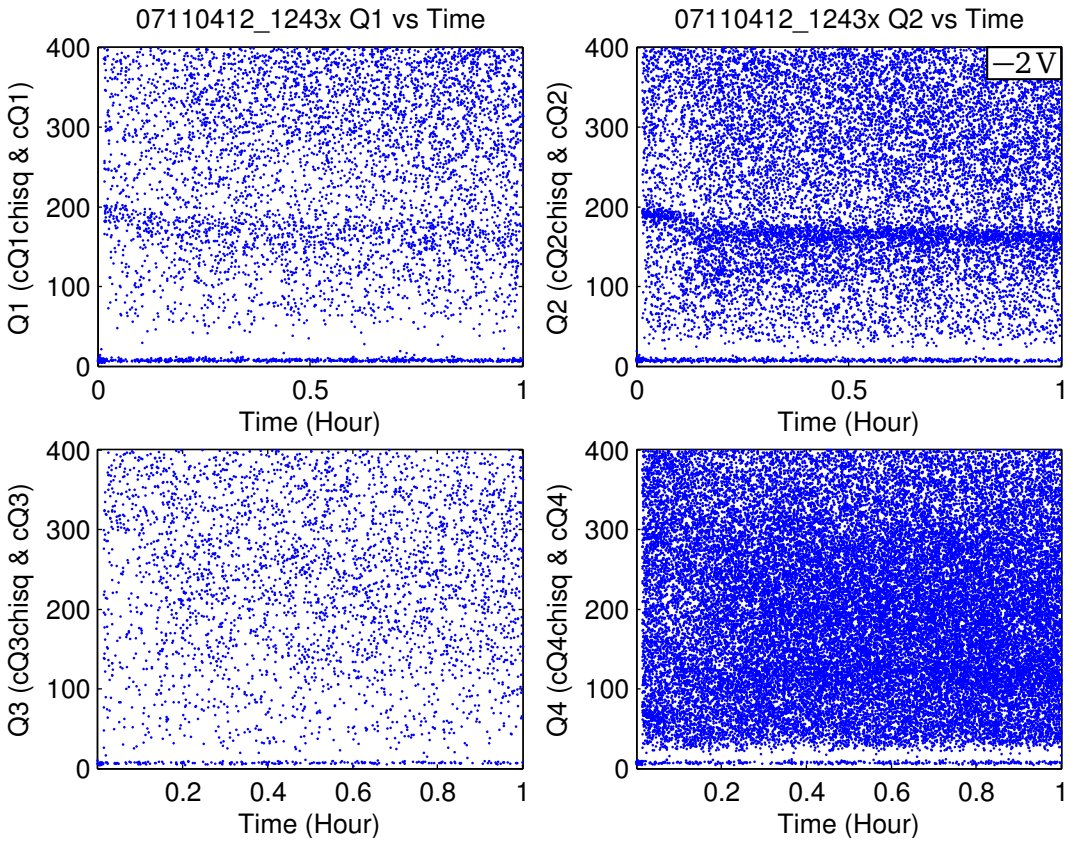


Figure 4.10: Ionization energy vs. bias time for the four charge electrodes on G101a at −2 V bias. The ionization energy is in arbitrary units. The horizontal bands around 200 are the ^{241}Am 60 keV line. The charge χ^2 cut and primary channel cut are applied. G101a loses neutralization after the first few minutes.

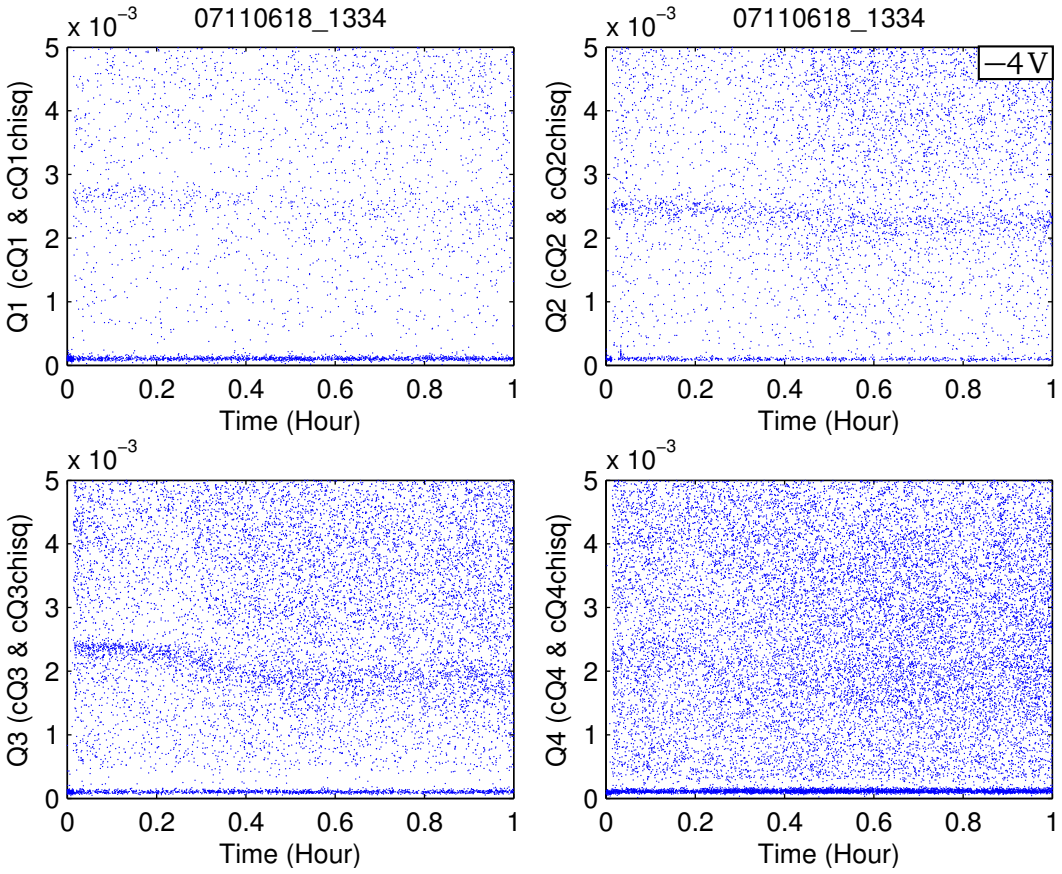


Figure 4.11: Ionization energy vs. bias time for the four charge electrodes on G102 at −4 V bias. The ionization energy is in arbitrary units. The horizontal bands are the ^{241}Am 60 keV line. The charge χ^2 cut and primary channel cut are applied. The neutralization time is about 0.2 hours.

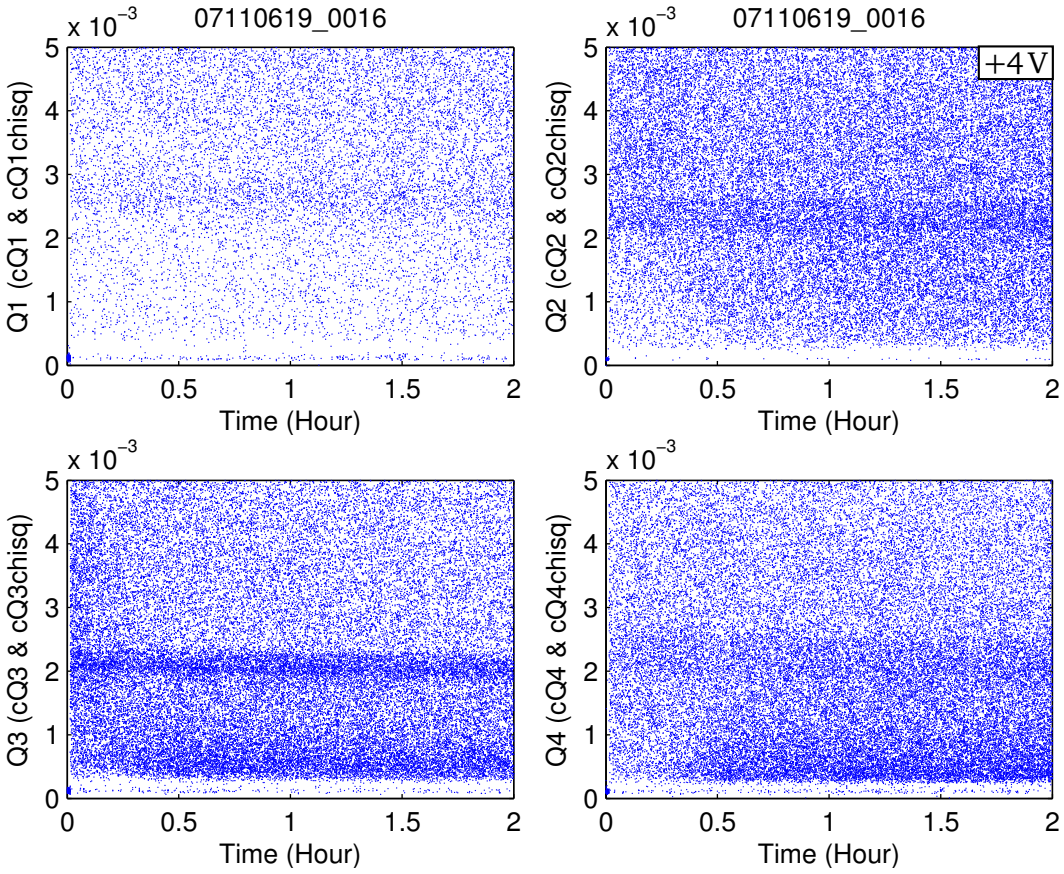
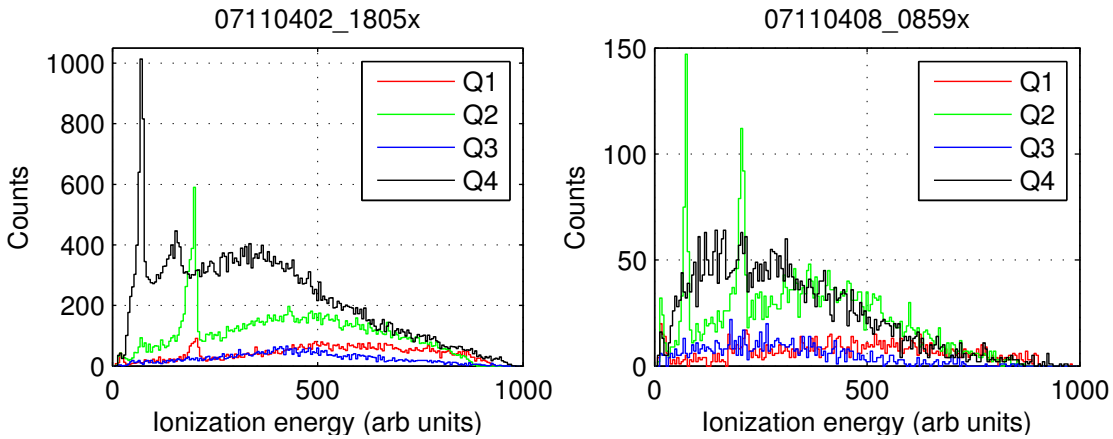


Figure 4.12: Ionization energy vs. bias time for the four charge electrodes on G102 at 4 V bias. The ionization energy is in arbitrary units. The horizontal bands are the ^{241}Am 60 keV line. The charge χ^2 cut and primary channel cut are applied. The signal degradation in the first two hours is not significant.



(a) (b)

Figure 4.13: Uncalibrated ionization spectrum for G101a at biases of −4 V (a) and 4 V (b). The ionization energy is in arbitrary units. The primary channel cut, energy threshold cut, and charge χ^2 cut are applied. The sharp peaks on the plots are from the 60 keV gamma of ^{241}Am.

With the good events selected, we had a look at the ionization spectra of the 100 mm devices.

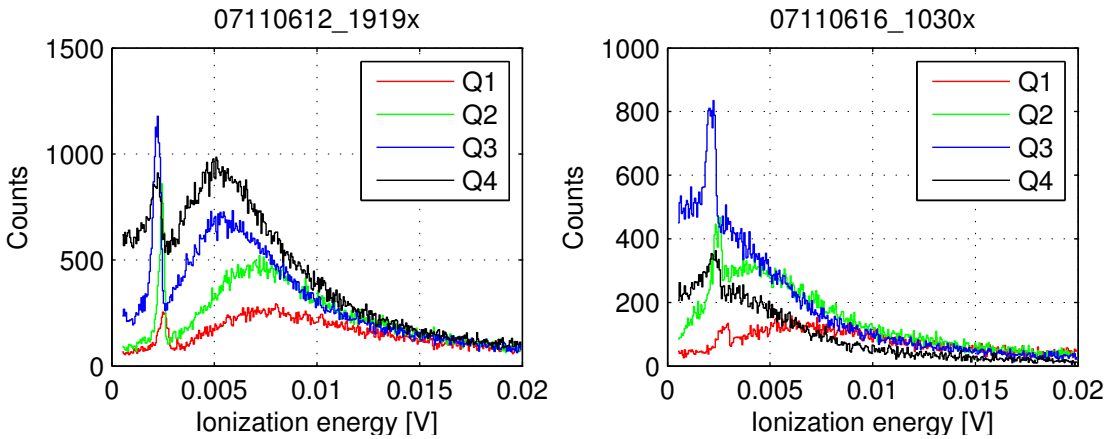
4.5.3.1 G101a

Figure 4.13(a) and (b) show the uncalibrated ionization spectra for G101a at biases of −4 V and 4 V, respectively. The 60 keV ^{241}Am lines can be clearly seen on Q_2 at both polarities. However, they can only be seen on Q_1 and Q_4 at negative bias. It is surprising that no signature of the 60 keV line can be seen on Q_3 regardless of the polarity. There are two obvious issues with the data collected on G101a. First, the event rates on Q_1 and Q_3 are much smaller than Q_2 and Q_4 at both polarities. Second, Q_2 has two peaks from the 60 keV photons.

We tried different configurations in order to mitigate the problem without success. First we checked the noise of each channel and the trigger settings and they seemed

to be reasonable. Then we suspected that the electric field under Q_1 and Q_3 might be weaker than Q_2 and Q_4. We calibrated the output voltage of the DCRC's to within a few mV uncertainty and got the same result. We then swapped the two DCRC's on the two DIB's and still observed the same behavior. The scratch on the device may have to do with the pathology of the event rates. But since Q_1 was found well isolated in the ^{3}He fridge screening, it should not be affected by the scratch. Since Q_1 and Q_3 were on the same DIB, this may be correlated with the fact that the charge return line for DIB2 (for Q_1 and Q_3) was not connected (see Table 4.2). In this case, the charge return line on DIB1 (for Q_2 and Q_4) also served as the return line for Q_1 and Q_3. Thus the parasitic capacitance to ground for Q_1 and Q_3 would be higher, reducing the overall gain of the charge amplifier. But the 60 keV line on Q_1 is roughly at the same position as on Q_2. There did not seem to be a penalty on the gain of the charge amplifier.

When we disassembled the tower, a screw for the Q_3 ^{241}Am source was found fallen onto Q_2. The screw was tall enough to short Q_2, which we did not observe in electrical checks at room temperature and 77 K. At lower temperatures, if it did fall out and short Q_2, we would observe loss of events on Q_2, which did not happen. It would be consistent with the data if the screw shorted Q_3. However, the screw hole was on top of Q_2. Since G101a was a defective detector, we decided not to spend too much time to understand the problem and moved on to G102. The fact that we observed 60 keV lines on Q_1, Q_2, and Q_4 was still very encouraging.



(a) (b)

Figure 4.14: Uncalibrated ionization spectra for G102 at biases of −4 V (a) and 4 V (b). The ionization energy is in arbitrary units. The primary channel cut, energy threshold cut, and charge χ^2 cut are applied. Only the first 12 minutes of data in each series are used. The sharp peaks on the plots are from the 60 keV gammas of ^{241}Am.

4.5.3.2 G102

Figure 4.14(a) and (b) show the uncalibrated ionization spectra for G102 with the ^{241}Am sources at biases of −4 V and 4 V, respectively. These spectra do not have the pathologies of G101a. At negative biases, holes are collected on the electrodes, which travel along the direction of the field lines. High energy ambient photons can penetrate the crystal and the resulting events populate the crystal uniformly. Since the four electrodes cover the same detector volume, they see statistically the same number of events at a given time interval, as shown in the high energy tails where the ionization energy is larger than 0.01 V in Fig. 4.14(a).

For photons with ionization energy between 0.005 V and 0.01 V, the penetration depths are small compared with the crystal dimension. Thus the inner volume of the crystal is shield by the crystal surrounding it. This self-shielding effect results in

lower event rates on the inner electrodes.

The four ^{241}Am sources were each collimated at 20 Hz according to a simple calculation assuming a point source at the center of the ^{241}Am capsule. However, the radioactive material was distributed in a small cavity inside the capsule. The nonuniformity of the distribution was likely to introduce systematic difference in the collimated rates between the sources. The event counts difference in the 60 keV peak for the four electrodes in Fig. 4.14(a) were likely caused by the difference in the collimated source rates and the transport of holes in the crystal. The relative strength of the 60 keV peaks were similar to that in Fig. 4.14(a) at −6 V and −8 V bias, which suggests the 60 keV event rate difference is more likely due to the difference of the collimated rates.

When the electrodes collect electrons under positive bias (Fig. 4.14(b)), the behavior of the spectra can also be understood by similar arguments as the case for collecting holes. Because the electrons propagate obliquely in a [100] Ge crystal, their distribution is more complex than that of the holes. A quantitative description of the spectra in Fig. 4.14(b) would require a simulation, which is not the topic of this thesis. More information about charge Monte Carlo simulation of CDMS detectors can be found in [89].

A ^{133}Ba source was also used to take data with G102. An example of the ionization spectra with the ^{133}Ba source at −6 V bias is shown in Fig. 4.15. The four channels were plotted separately to avoid overlap.

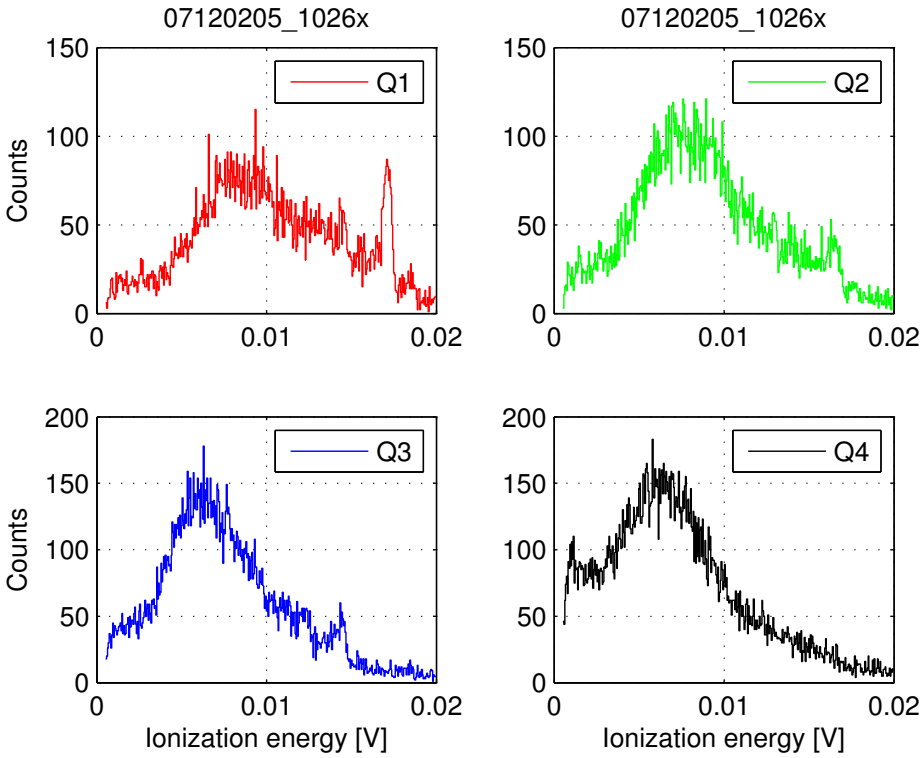


Figure 4.15: Uncalibrated ionization spectra for G102 with ^{133}Ba source at −6 V bias. The primary channel cut, energy threshold cut, and charge χ^2 cut are applied. Only the first 12 minutes of data in each series are used. They are plotted in separated panels to show the ^{133}Ba 356 keV peaks between 0.01 and 0.02 V.

4.6 Charge calibration

The energy of the events presented in earlier sections were all in some arbitrary units output by the data processing package `CDMSBats` or voltage after hardware gain correction. It would be more convenient to use the physical units such as keV to characterize the energy. Thus we transform the energy of the events from the voltage space to the energy space.

If the four channels are all independent, only four scaling factors are needed to scale the voltage to energy. However, the capacitive crosstalk correlates all four channels. The energy measured in one channel had to do with the energies measured in all the other three channels. For an event, in the most general form,

$$\hat{w}_i = \sum_{j=1}^{4} v_j a_{ji}, \quad i = 1, \cdots, 4, \tag{4.1}$$

where $\hat{w}_i$ is the estimator of the energy deposition of the event onto electrode i, a_{ji} are the constants to be determined, v_j is the voltage measured on electrode j. There are 16 a_{ji}'s to be determined. In the four dimensional space, the transformation matrix a_{ji} can be completely determined by the transformation of four linearly independent vectors (events).

Usually there are many more events taken and Eq. (4.1) is over-constrained. Assume there are N events (spanning the whole four dimensional space) with coordinates in both the energy space and the voltage space known. For event n, following

Eq. (4.1), we have

$$\hat{E}_{ni} = \sum_{j=1}^{4} V_{nj} a_{ji}, \quad i = 1, \ldots, 4 \text{ and } n = 1, \cdots, N, \tag{4.2}$$

where $\hat{E}_{ni}$ is the estimator of the energy of event n deposited onto electrode i and V_{nj} is the voltage of event n measured on electrode j. Equation (4.2) can be written in a matrix form

$$\hat{E}_{N\times 4} = V_{N\times 4} a_{4\times 4}. \tag{4.3}$$

If the likelihood function of the energy for each of the N events are known, a maximum likelihood estimation of a can be formulated. Suppose the energy deposition onto the electrodes follows the Gaussian distribution:

$$E_{ni} \sim \mathcal{N}(\mu_{ni}, \sigma_{ni}). \tag{4.4}$$

The likelihood function for the energies measured on electrode i is

$$L_i = \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi}\sigma_{ni}} \exp\left[-\frac{(E_{ni} - \hat{E}_{ni})^2}{2\sigma_{ni}^2}\right]. \tag{4.5}$$

Since the electrodes are independent in energy space, their likelihood functions can be maximized separately. If σ_{ni} is energy independent, maximizing Eq. (4.5) gives the least squares estimator (LSE) of the i^{th} column of a:

$$a_{(i)} = (V'V)^{-1} V' E_{(i)}, \tag{4.6}$$

where V' is the transpose of V and $E_{(i)}$ is the i^{th} column of E. Collecting all the $a_{(i)}$

and $E_{(i)}$, we have

$$\hat{a} = [a_{(1)}, a_{(2)}, \cdots] = (V'V)^{-1}V'E. \tag{4.7}$$

There is one special case when only four calibration events are used. In this case, V and E are square matrices. Equation (4.7) can be simplified as

$$\hat{a} = V_{4\times4}^{-1}E_{4\times4}. \tag{4.8}$$

The 60 keV events are natural choices for calibration since their coordinates in the energy space and the voltage space are both known. To simplify the calculation and also keep the accuracy of the charge calibration, the centers of the four 60 keV event peaks for the four electrodes were used as the four calibration events. Let γ be the energy of the calibration events (60 keV), then $E = \gamma I$, where I is the 4×4 unity matrix, and Eq. (4.8) can be written as

$$\hat{a} = \gamma V^{-1}, \tag{4.9}$$

where V_{ij} is the voltage of the 60 keV event blob for electrode i measured on electrode j.

Figure 4.16 shows the ionization energy of Q_2 versus Q_1 for G102 at -4 V bias. The white blue band along the Q_1 axis are the events that deposited most of their energy onto the Q_1 electrode, which are also referred to as Q_1 primary events. Similarly the band along the Q_2 axis corresponds to the Q_2 primary events. It is apparent that there is crosstalk between Q_1 and Q_2 since the two bands are not perpendicular to each other. The two event blobs above 2×10^{-3} on the two bands are due to the

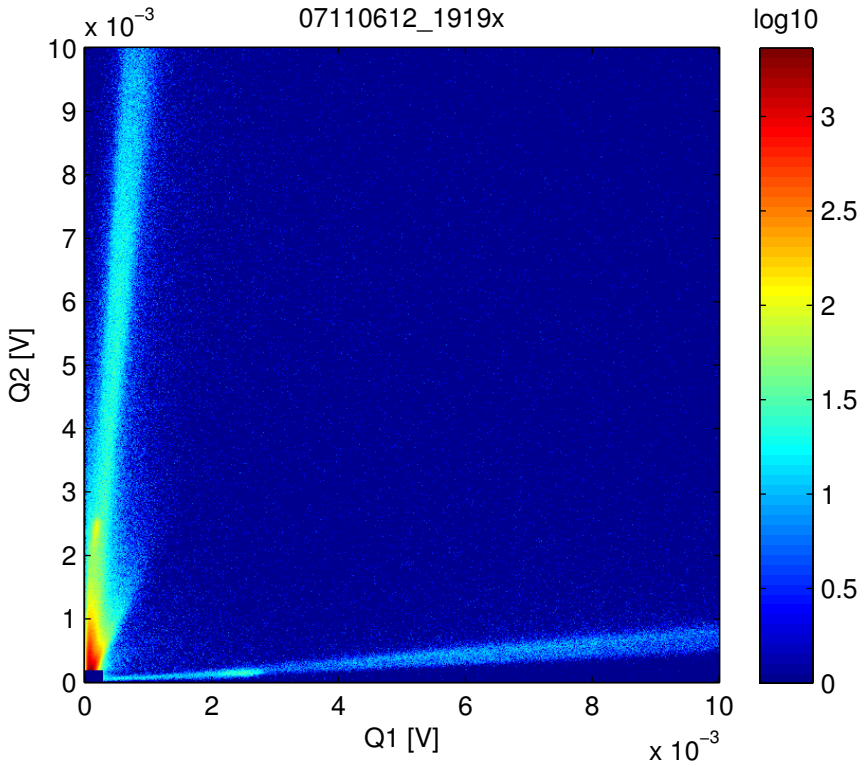


Figure 4.16: Q_2 vs. Q_1 for G102 at -4 V bias. The blob between 2×10^{-3} and 4×10^{-3} on the Q_1 band and the blob between 2×10^{-3} and 3×10^{-3} on the Q_2 band both are events from the 60 keV gammas. Point density is colored in $\log_{10}$ scale. The noise blob at the origin is cut out to show the less dense area with the full color map.

60 keV gamma events. The projection of the Q_1 blob, say S_1, onto the Q_1 axis is V_{11} in Eq. (4.9). Similarly the projection of S_1 onto the Q_2 axis is V_{12}. Once all the elements of V are obtained, the transformation matrix a is determined.

However, the projection of S_1 onto Q_2 is rather small as can be seen in Fig. 4.16; the projections onto Q_3 and Q_4 are even smaller. Projecting the whole Q_1 band onto a rather small interval in Q_2 doesn't give the blob S_1 much signal to noise. Observing that all the 60 keV blobs are on their corresponding bands which are straight lines (to the first order) going through the origin, we can decompose the projections V_{ij} into an amplitude term g_{ii} and a crosstalk term k_{ij}:

$$V_{ij} = g_{ii} k_{ij}, \tag{4.10}$$

where g_{ii} is the projection of S_i onto the axis of Q_i, and the events on the Q_i band satisfy

$$Q_j = Q_i k_{ij}. \tag{4.11}$$

g_{ii} is just the position of the S_i (60 keV) events on the Q_i axis. It can be determined relatively easily from the ionization spectra shown in Fig. 4.14. In data analysis, we fitted the 60 keV peak with a Gaussian plus a linear or quadratic background, and the mean of the Gaussian was chosen as g_{ii}. k_{ij} is the slope of the Q_i band in the Q_i-Q_j plane. It is obvious $k_{ii} = 1$. Since there are many more events in the Q_i band than in the S_i blob, the slope can be determined with much smaller statistical uncertainty.

Although the slope k_{ij} can be obtained by a fit to the Q_i band in the Q_i-Q_j plane, in analysis, instead, we projected the 2D distribution onto a 1D parameter to simplify

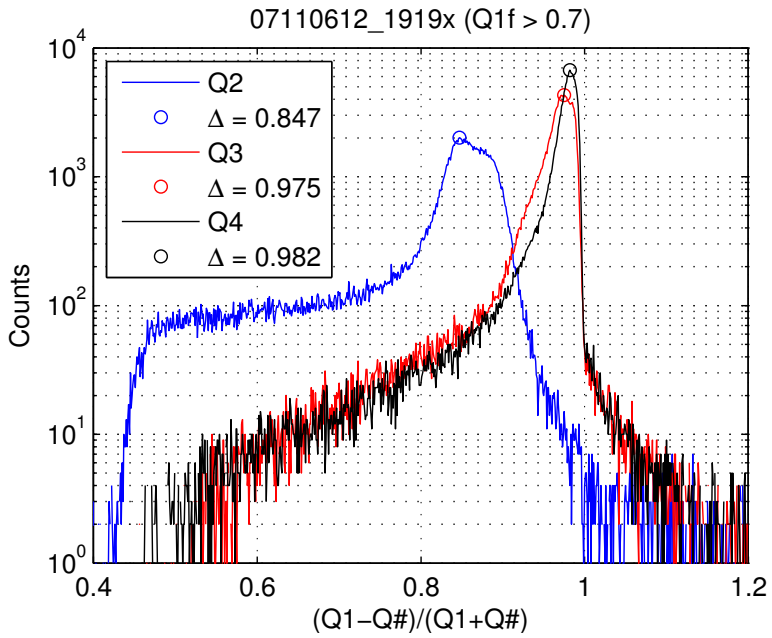


Figure 4.17: Q partition between Q_1 and the other three channels for G102 at bias −4 V. The circles indicate the peak position. Only events which deposited more than 70 % of their energy onto Q_1 are plotted.

the calculation. Define the partition between two channels:

$$P_{ij} = \frac{Q_i - Q_j}{Q_i + Q_j}. \tag{4.12}$$

The Q_i band is then projected to a peak in the parameter P_{ij} distribution. Assume the peak has its maximum height at $P_{ij} = \Delta_{ij}$, combining Eq. (4.11) and (4.12) gives

$$k_{ij} = \frac{1 - \Delta_{ij}}{1 + \Delta_{ij}}. \tag{4.13}$$

Figure 4.17 shows the charge partition between Q_1 and the rest three channels for G102 at bias -4 V. The crests of the peaks are indicated with the circles. The slopes k_{ij} can be calculated with the observed Δ's.

With all the transformation matrix a determined, Eq. (4.1) gives the calibrated energies of the events. Figure 4.18 shows the calibrated charge spectra for G102 at the bias of -4 V. Note all the ^{241}Am peaks are aligned at 60 keV.

In all the spectra shown above, we were focused on the ionization energy measured by the individual channels. This is reasonable for the events depositing most or all of their energy into one channel. For the events depositing their energies into more than one channels, the sum of the calibrated ionization energy (`qsum`) on all electrodes should be used. Figure 4.19 shows the spectra of `qsum` for the four electrodes on G102 at -4 V. The 60 keV peaks of the three inner channels are at the correct energy. The 60 keV peak of Q_4 shifts up somewhat. Hence the ionization energy of the 60 keV events was fully collected on the three inner electrodes; Q_4 collected most of the ionization energy of the 60 keV events and the rest was picked up by the three inner channels. This is not unexpected since Q_4 is the outermost electrode and the electric

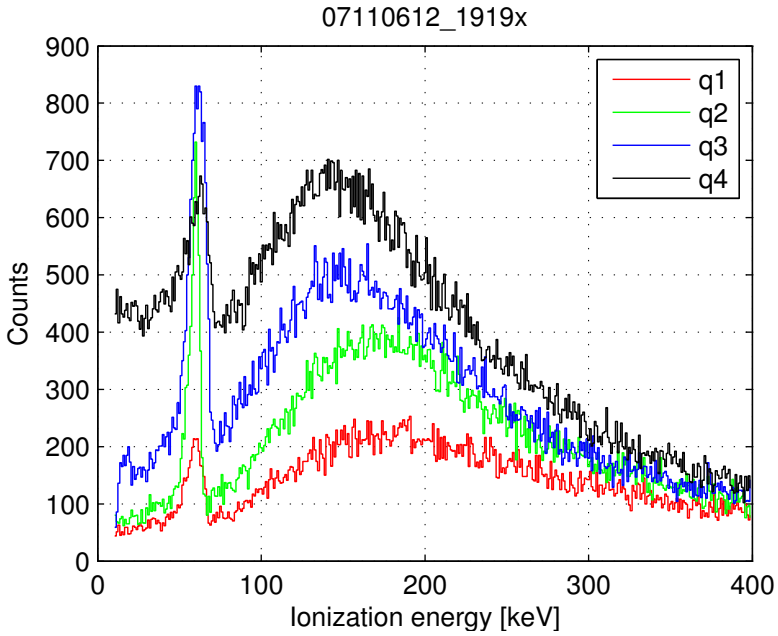


Figure 4.18: Calibrated ionization spectra of the primary events for Q_1 to Q_4 on G102 at −4 V bias. The primary channel cut and charge χ^2 cut are applied to each channel. Only the first 0.2 hours of data at the beginning of each series are used.

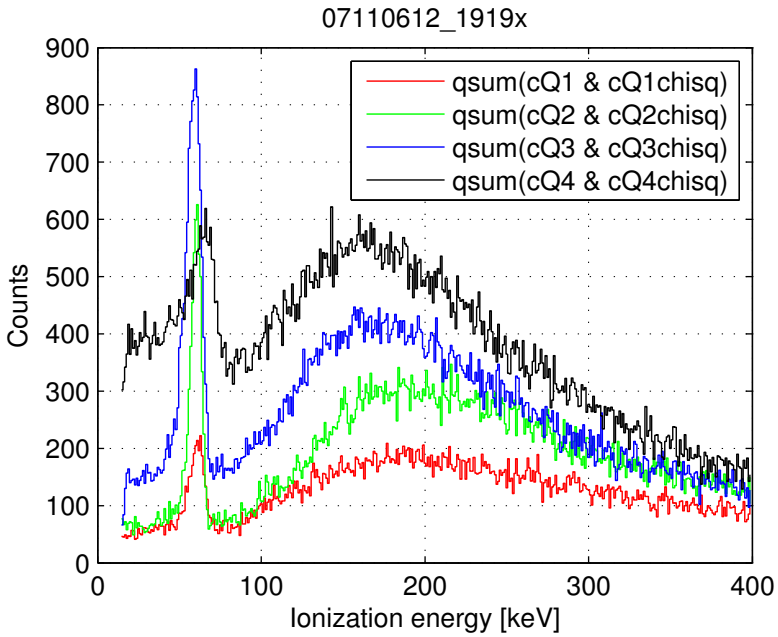


Figure 4.19: Calibrated ionization spectra of the primary events for Q_1 to Q_4 based on sum of the ionization energy on G102 at −4 V bias. The primary channel cut and charge χ^2 cut are applied to each channel. Only the first 0.2 hours of data at the beginning of each series are used.

field deviates from uniform close to the side wall of the crystal, which decreases the ionization energy collection efficiency. In WIMP search analysis, these outer channel events are used to define a fiducial volume inside which the ionized charges are fully collected.

Since the capacitive coupling between two channels is symmetric, the crosstalk matrix k_{ij} is symmetric. We didn't use this symmetry in the charge calibration, instead calculated each term independently from the data.

4.7 Charge collection efficiency

4.7.1 Definition

The spectra in Fig. 4.18 show promising results about the crystals: the 100 mm ionization devices were able to identify the 60 keV gammas at the radii of the four electrodes. The next question was how efficient the test devices were in recovering the ionization energy of a scattering event. An inefficient detector in dark matter search would increase the energy threshold of the detection and reduce the event rates, lowering the sensitivity of the experiment. We characterized the performance of the test devices in recovering the ionization energy with the *charge collection efficiency*, defined as the ratio of the number of electrons (holes) collected on the electrodes to the total number of electron-hole pairs created in a scattering event. Since the number of electron-hole pairs generated in an event is proportional to the ionization energy, the charge collection efficiency is also referred to as the ionization collection

efficiency. So we have the relationship

$$\begin{aligned}\text{charge collection efficiency} &= \frac{\#\text{ of e (h) collected}}{\#\text{ of e-h pairs creaeted}} \\ &= \frac{\text{ionization energy recovered}}{\text{initial ionizaiton energy}}.\end{aligned} \tag{4.14}$$

In deriving the transformation matrix a in Sec. 4.6, we assumed full collection efficiency of the electrodes for the 60 keV blobs, which was rarely true. If the gain of the charge readout circuit was precisely known, the calculated energy of the 60 keV blobs would probably be smaller than 60 keV due to incomplete charge collection. Unfortunately, the gain of the FET amplifier can not be measured accurately [84] because of the parasitic effects and temperature dependence of the capacitances and resistances. Thus we need a gamma line with known energy in the data to determine the correct energy scale.

Nevertheless, suppose in an ideal configuration, say under a relatively high bias voltage, the charges can be fully collected, and we have all the voltage projections of the 60 keV blobs following Eq. (4.10):

$$V_{in}^0 = g_{ii}^0 k_{ij}. \tag{4.15}$$

Then Eq. (4.9) can be written as

$$a^0 = \gamma (g^0 k)^{-1} = \gamma k^{-1} (g^0)^{-1}. \tag{4.16}$$

Feeding Eq. (4.16) into Eq. (4.1) gives

$$w = va^0 = v\gamma k^{-1}(g^0)^{-1}. \tag{4.17}$$

And rearranging the terms, we have

$$v = \frac{1}{\gamma} w g^0 k. \tag{4.18}$$

For a calibration event on Q_i with energy u, the row vector w is

$$w_j = u\,\delta_{ij}, \quad j = 1, \cdots, 4. \tag{4.19}$$

Combing Eqs. (4.10), (4.18), and (4.19), we obtain

$$v_j = g_{ii} k_{ij} = \sum_{lm} \frac{1}{\gamma} u\,\delta_{il}\, g^0_{lm}\, k_{mj} = u \frac{g^0_{ii}}{\gamma} k_{ij}, \quad j = 1, \cdots, 4. \tag{4.20}$$

Since g_{ii} is the voltage measured on electrode i, we see that $\frac{g^0_{ii}}{\gamma}$ is the induced voltage on electrode i by unit energy deposition on the same electrode, which is the gain.

Now we define the charge collection efficiency for an arbitrary event induced by the complete absorption of a photon with energy u and with measured voltages v. The total energy recovered by the detector (four electrodes) is

$$\begin{aligned} \sum_i w_i &= \gamma \sum_{ijk} v_i (k^{-1})_{ij} \left[(g^0)^{-1}\right]_{jk} \\ &= \sum_{ij} v_i (k^{-1})_{ij} \frac{\gamma}{g^0_{jj}} \end{aligned} \tag{4.21}$$

And the charge collection efficiency is

$$\eta = \frac{\sum_i w_i}{u} = \frac{1}{u} \sum_{ij} v_i \, (k^{-1})_{ij} \, \frac{\gamma}{g^0_{jj}}. \tag{4.22}$$

Generally speaking, the charge collection efficiency for an arbitrary event is an nontrivial combination of all the voltages and gains of the four electrodes.

However, for events localized on the electrodes, Eq. (4.22) can be simplified. Suppose an event is localized on electrode m, then

$$v_i = g_{mm} \, k_{mi}, \quad i = 1, \cdots, 4. \tag{4.23}$$

Putting v_i into Eq. (4.22), we obtain

$$\begin{aligned} \eta_m &= \frac{1}{u} \sum_{ij} g_{mm} \, k_{mi} \, (k^{-1})_{ij} \, \frac{\gamma}{g^0_{jj}} \\ &= \frac{1}{u} \sum_{j} g_{mm} \, \delta_{mj} \, \frac{\gamma}{g^0_{jj}} \\ &= \frac{\gamma}{u} \, \frac{g_{mm}}{g^0_{mm}}. \end{aligned} \tag{4.24}$$

If the charge collection efficiency is measured using photons with the same energy but different experimental conditions, i.e., $u = \gamma$, then Eq. (4.24) can be written as

$$\eta_i = \frac{g_{ii}}{g^0_{ii}}, \quad i = 1, \cdots, 4. \tag{4.25}$$

Equation (4.25) says the charge collection efficiency for events localized on electrode i is proportional to the voltage measured on the same electrode. This is not surprising

since we assume the events localized on one electrode only deposit energy to the same electrode, then the voltage measured by the electrode is a measure of the collection efficiency of the whole detector. Equation (4.25) is not only rather convenient, but also provides a way to measure the position (radial) dependence of the charge collection efficiency.

In practice, the direct sum of the measured voltages (`Qsum`) was often times used as a measure of the energy of an event. We see from Eq. (4.21) that, the direct sum of the voltage does not usually make a meaningful quantity because the the crosstalk and gain difference between the channels. However, in the cases where the crosstalk is small, e.g., the crosstalk is corrected in data processing, and all channels have roughly the same gain, which is not unreasonable, then Eq. (4.21) can be written as

$$\sum_i w_i = \sum_{ij} v_i (k^{-1})_{ij} \frac{\gamma}{g^0_{jj}} \approx \sum_{ij} v_i \delta_{ij} \frac{\gamma}{g^0_{jj}} = \frac{\gamma}{g^0_{ii}} \sum_i v_i. \tag{4.26}$$

Thus we have

$$\sum_i v_i = \frac{g^0_{ii}}{\gamma} \sum_i w_i. \tag{4.27}$$

The direct sum of the voltage is proportional to the energy of the event. This is a convenient approximation to estimate the energy of an event without going through the detailed calibration.

As a special case, for the events localized on an electrode Q_i,

$$\sum_j v_j = \sum_j g_{ii} k_{ij} = g_{ii} \sum_j k_{ij}. \tag{4.28}$$

Combined with Eq. (4.24), Eq. (4.28) gives

$$\begin{aligned}\sum_j v_j &= \eta_i u \frac{g^0_{ii}}{\gamma} \sum_j k_{ij} \\ &= \eta_i \frac{u}{\gamma} \sum_j v^0_j,\end{aligned} \tag{4.29}$$

where v^0_j are the voltages measured at full charge collection efficiency. We see that the sum of the voltage is proportional to the total energy deposition ($\eta_i u$). If u is a constant, the sum of the voltage is then proportional to the charge collection efficiency η_i.

In the CDMS experiment, the outermost electrode was used as a guard ring to ensure the charge propagation was confined in the fiducial volume. Only the events whose energy deposition in the outermost ring was consistent with noise were selected for WIMP search analysis. The outermost electrode of the 100 mm devices was also used as a guard ring in calculating the charge collection efficiency. Though the derivation above is based on a four-electrode detector, the conclusions are still valid for a detector with three electrodes.

4.7.2 Measured charge collection efficiency

4.7.2.1 ^{241}Am on electrode side

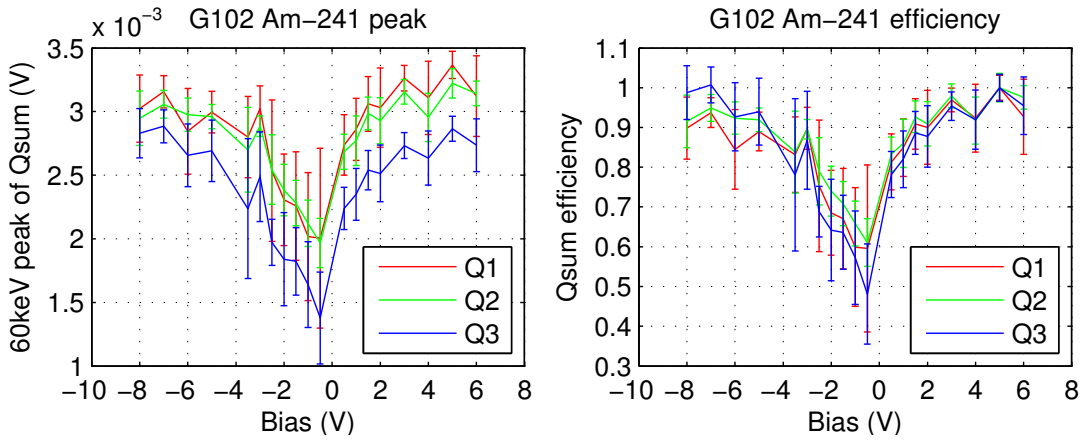
We took data with the same gamma source (^{241}Am and ^{133}Ba) configuration at various bias voltages. Good events were selected with the neutralization cut, the charge χ^2 cut, and a simple charge fiducial volume cut `cQin` using Q_4 as the veto. A properly defined `cQin` requires the calibrated ionization energy on Q_4 is consistent with noise

while some distance away above the noise floor on at least one of the other channels, which needs the full charge calibration. However, by Eqs. (4.24) and (4.29), the charge collection efficiency based on the localized calibration events is independent of the crosstalk. A full charge calibration is not necessary. Instead we defined the `cQin` on the uncalibrated voltages by requiring

$$\texttt{cQin} : 0.8 < \frac{v_1 + v_2 + v_3 - v_4}{v_1 + v_2 + v_3 + v_4} \leq 1. \tag{4.30}$$

Because of the crosstalk between Q_4 and the other three channels, a decent signal on the inner electrodes induces a visible signal on Q_4. And thus the requirement that v_4 is consistent with noise can not be imposed. The largest crosstalk between the three inner channels and Q_4 is Q_3-Q_4, which is $\approx 10\,\%$. This gives the 0.8 on the left side of the inequality in Eq. (4.30). Our calculation of the charge collection efficiency was based on the direct sum of the measured voltages for the primary events on the three inner channels. In the worst scenario, which is very unlikely to happen, the `cQin` in Eq. (4.30) would introduce an upward bias of 10 % to the charge collection efficiency, which is not unacceptable. From Fig. 4.19 we see that the localized calibration events on the three inner electrodes had negligible energy leakage to the other electrodes even without `cQin` applied. So the simple version of the `cQin` should suffice for our purpose.

With good events selected, the histograms of `Qsum` of the primary events for the three inner channels were plotted. The 60 keV peaks were fitted to a Gaussian and the mean and sigma were obtained. The mean and $\pm 1\,\sigma$ of the 60 keV peak versus bias is shown in Fig. 4.20(a). The energy resolution is $\sim$20 % (FWHM) except for



(a) (b)

Figure 4.20: (a): Position of the 60 keV peak vs. bias for the three inner electrodes on G102. The error bars are the $\pm 1\,\sigma$ width of the 60 keV peaks. (b): Charge collection efficiency vs. bias for G102. The charge collection efficiency at 5 V is chosen as unity for all the three channels.

the biases close to 0 V, where it gets worse. The errors of the fits for the mean of the Gaussian are significantly smaller and they are not shown in the plot. The 60 keV positions of Q_1 and Q_2 track each other very well, but those of Q_3 are systematically lower. We chose the charge collection efficiency at 5 V as unity for each channel, and scaled all the others accordingly. The charge collection efficiency versus bias is plotted in Fig. 4.20(b). We see that the efficiency curves for the three channels resemble one another. The observation that the 60 keV peaks on Q_3 are systematically lower than Q_1 and Q_2 is likely due to a lower gain of the charge readout circuit on Q_3. The efficiency curves in Fig. 4.20(b) are roughly symmetric about the lowest point. The minimum efficiency is not at 0 V but at a small negative bias. This is because hole trapping is more effective than electron trapping in the crystals we tested. A higher bias voltage (in amplitude) is needed to achieve the same collection efficiency at small biases. When collecting holes (under negative bias), full charge collection can

be achieved at biases $\gtrsim 6\,\mathrm{V}$ ($1.8\,\mathrm{V\,cm^{-1}}$). This is consistent with the field strength used by the CDMS-II detectors ($3\,\mathrm{V\,cm^{-1}}$) and the earlier measurements [84, 111].

4.7.2.2 ^{241}Am souce on the ground plane side

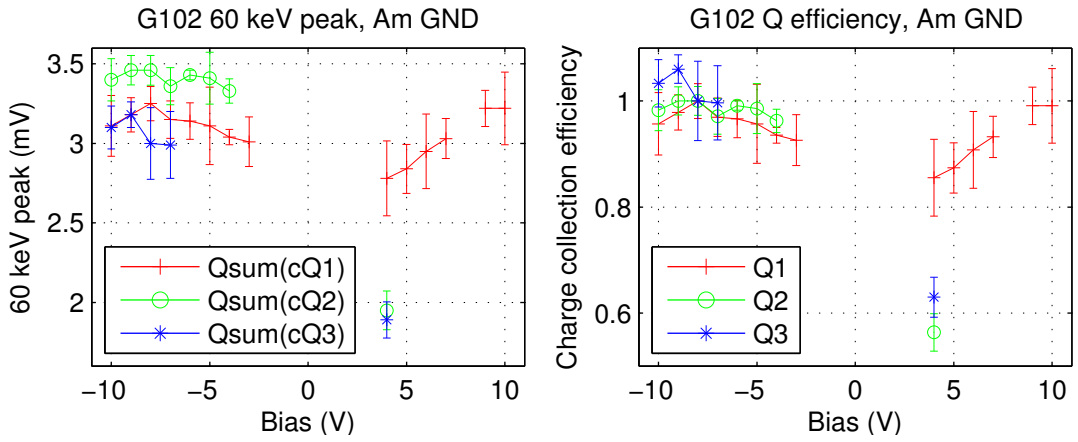
We also took data with a single ^{241}Am source on the ground plane side facing Q_2 to investigate the transport of hot charge carriers in the Ge crystal.

The anisotropy of the velocity of holes in Ge is small and the transport of the holes are commonly treated as isotropic [112, 113]. The acceleration of holes is in the direction of the electric field. Because of the scattering during drifting, the motion of the holes in transverse direction is diffusive. The standard deviation of the transverse straggle of a hole traveling through the crystal is [113]

$$\sigma_\perp \propto \sqrt{x}(E)^{-1/10}, \tag{4.31}$$

where x is the thickness of the crystal and E is the electric field strength. Scaling from the numbers in [113], for the $x = 33.3\,\mathrm{mm}$ thick and 100 mm diameter crystal at a bias of 8 V, $\sigma_\perp = 1.7\,\mathrm{mm}$. The collimated ^{241}Am source illuminates an area of a disk of diameter 1 mm at a depth of 1 mm underneath the crystal surface. Thus the $2\,\sigma$ area collecting holes on Q_2 is of diameter 4.4 mm. The width of Q_2 is 10 mm. This means almost all holes would be collected on Q_2.

However, the 60 keV peak showed up in all three inner electrodes at relatively large negative biases, which is shown in Fig. 4.21(a). The amplitudes of the 60 keV peaks in Fig. 4.21(a) are comparable to those in Fig. 4.20(a), which means the 60 keV events were fully collected (upto the efficiency loss at a certain bias) on the electrodes



(a) (b)

Figure 4.21: (a): Position of the 60 keV peak vs. bias for the three inner electrodes on G102. The ^{241}Am source is on the ground plane side facing Q_2. The error bars are the $\pm 1\,\sigma$ width of the 60 keV peaks. (b): Charge collection efficiency vs. bias for G102. The charge collection efficiencies are normalized to unity at −8 V.

even though the ^{241}Am source was at the opposite side of the crystal. There was suspicion that the electric field in the crystal had a radial component which caused wider spread of the holes when reaching the other side of the crystal. It is possible that the electric field in the crystal had a transverse component which deflected the path of the holes. But this can not explain why the holes end up reaching all three inner electrodes and were fully collected on each. The observation is not fully understood in the simple transport model of the holes.

As discussed earlier in this chapter, the position of the 60 keV peak characterizes the charge collection efficiency. The close-to-fully collection shown in Fig. 4.21(a) shows that the holes were not significantly trapped when traveling through the crystal. This transparency is important in operation of the detector for dark matter search.

Contrary to that of holes, the transport of electrons in Ge is highly anisotropic due to the anisotropy of the electron effective mass and the multivalleyed conduction band

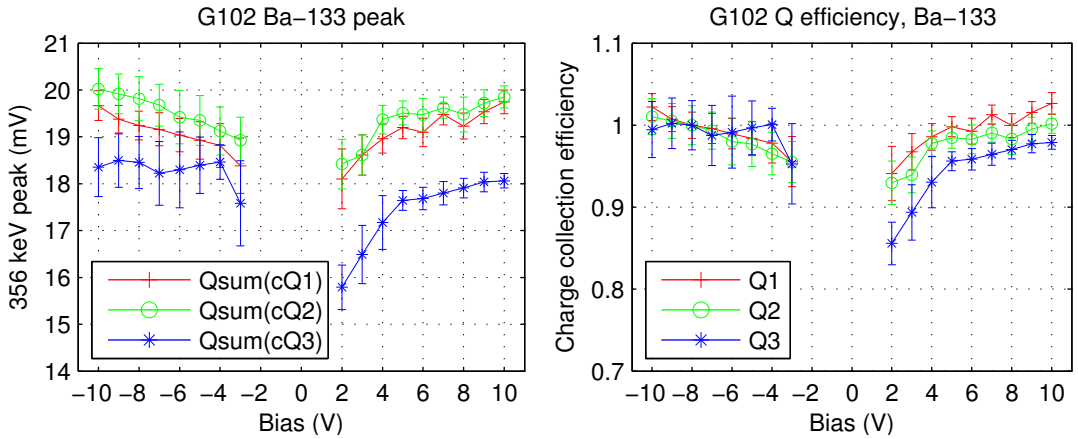
structure. The lowest points of the conduction band for electrons in Ge are the four L valleys along the $\langle 111 \rangle$ directions. At temperatures below $\sim$1 K and electric fields less than $\sim$5 V cm^{-1}, the inter-valley transition is energetically suppressed and the electrons remain in the same valley while drifting in the crystal [113]. The electrons in the four valleys propagate at an angle of $\sim$30 degrees from the z axis. The number of electrons resulting from a 60 keV event reaching Q_2 would be negligible. The data is compatible with this prediction for most of the positive biases as shown in Fig. 4.21(a). However, the 60 keV peak did show up on Q_2 and Q_3 at 4 V. As for the case of holes, a simple model of the electron transport may not explain the data well. A thorough understanding of the data would need a detailed Monte Carlo simulation.

At biases close to 0 V, the 60 keV peak could not be resolved in the data, which is shown as the gap in Fig. 4.21. This is likely due to charge carrier trapping at low biases.

4.7.2.3 ^{133}Ba source

Data with an external ^{133}Ba source was taken as well to investigate the response of the ionization device to the 356 keV photons. The attenuation length of the 356 keV photons in Ge is $\sim$20 mm. The ^{241}Am source was $\sim$1 m above the detector and 30 cm off the axis. There was a certain level of nonuniformity due to the nonuniform illumination to the ionization device by the source. Again, detailed understanding of the charge collection efficiency with the ^{133}Ba source shown in Fig. 4.22 would require Monte Carlo simulations. The 356 keV peaks on Q_3 are consistently lower than those on Q_1 and Q_2, which may have to do with the gain of this channel.

As for the case of the ^{241}Am source on the ground side, possibly due to charge



(a) (b)

Figure 4.22: (a): Position of the 356 keV peak vs. bias for the three inner electrodes on G102. Error bars are the $\pm 1\,\sigma$ width of the 356 keV peaks (b): Charge collection efficiency vs. bias for G102. The charge collection efficiency is normalized to unity at -8 V.

carrier trapping, the 356 keV peak could not be identified in the data at low biases.

4.7.2.4 Comparison and conclusion

The charge collection efficiencies versus bias for the three source configurations are plotted in Fig. 4.23. Overall, the three configurations give consistent charge collection efficiencies. The crystal shows significant signal loss at low biases if the events occurred far from the electrodes, which may be due to charge carrier trapping. At the operating electric field strength of the CDMS-II ZIPs ($3\,\mathrm{V\,cm^{-1}}$) or SuperCDMS iZIPs ($1.6\,\mathrm{V\,cm^{-1}}$), the crystal showed good collection efficiency.

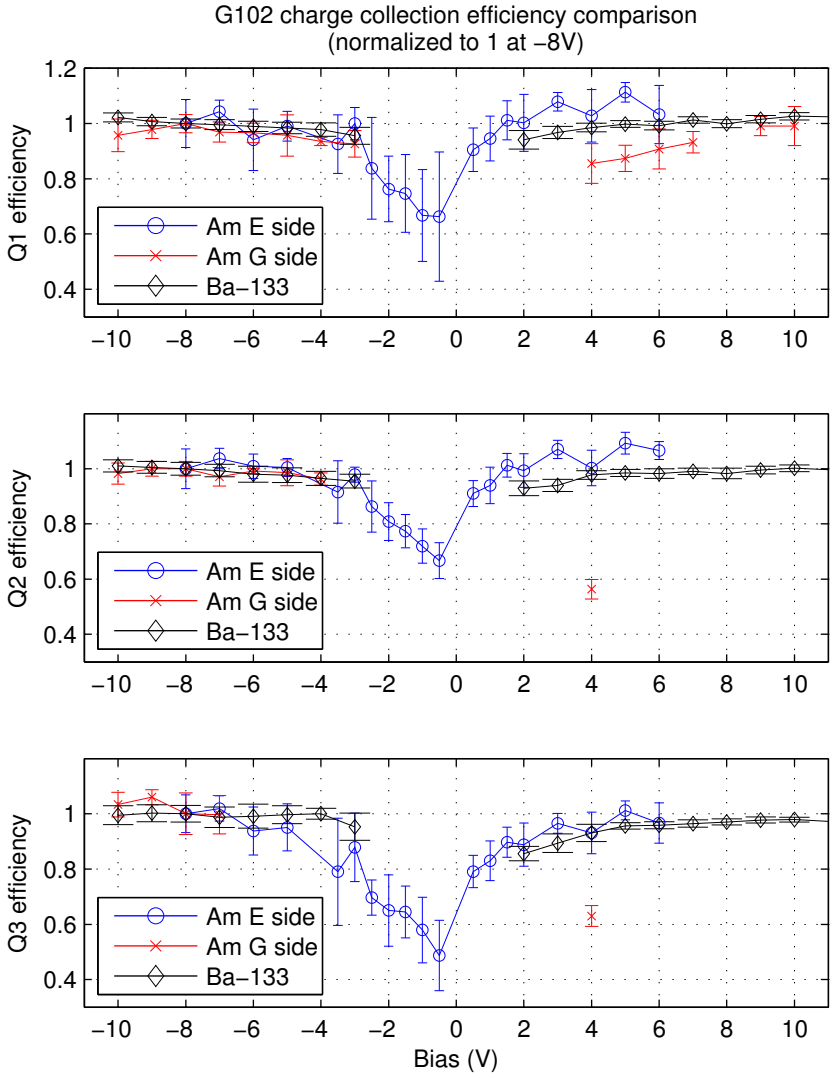


Figure 4.23: Charge collection efficiency comparison with the ^{241}Am source on the electrode side, on the ground side, and with the ^{133}Ba source for the three inner electrodes on G102.

Chapter 5

Reanalysis of the CDMS II Runs 125-128 Data

5.1 Detector and data selection

This reanalysis used the same detector set and data set that were used in the previous analysis. The data were taken with the five tower installation between July 2007 and September 2008 in four discrete runs 125-128. The previous analysis was thus given the name "C58 analysis", and we call this reanalysis as the "C58R analysis". Table 5.1 shows the detector stacks in the five towers during the Runs 125-128. Out of the 19 Ge detectors, 5 were not used for WIMP search analysis because of broken channels or high phonon noise. Fourteen detector were used for WIMP-search analysis, four of them had problems in one or more runs because of broken phonon channels or low event rates. The details are summarized in Table 5.2.

All good series of data were selected by the cut `cGoodSeries_c58` defined in the

Table 5.1: CDMS-II detector stacks in five towers. “T” represents a tower and ”Z” labels a detector. The shaded ones are Ge detectors and the rest are Si detectors. The 10 detectors shaded in green were good for all runs. The 4 shaded with light-gray had problems for some of the runs. The 4 shaded with red had problems for all four runs.

	T1	T2	T3	T4	T5
Z1	Ge	Si	Si	Si	Ge
Z2	Ge	Si	Ge	Ge	Ge
Z3	Ge	Ge	Si	Si	Si
Z4	Si	Si	Ge	Ge	Ge
Z5	Ge	Ge	Ge	Ge	Ge
Z6	Si	Si	Ge	Ge	Ge

C58 analysis. We used the same list of series in the reanalysis.

5.2 Motivation for the reanalysis

Two WIMP candidate events were observed in the C58 analysis for the Ge detectors. Investigations of the two events revealed that the charge reconstruction algorithm did not choose the global minimum of the χ^2 of the the charge pulse fit for one of the events. This resulted in an earlier start time for the fitted charge pulse by $4\,\mu s$, and increased the phonon pulse delay by the same amount. If the start time corresponding to the global minimum of χ^2 of the fit had been chosen, this event would not have passed the timing cut. Figure 5.1 shows the raw charge pulse of the WIMP candidate event in T3Z4 overlaid with the pulse fitted by the optimal filter, which suggests that if the fitted pulse were shifted to the right by a few bins, a better fit may be obtained. The flaw of the charge reconstruction algorithm was caused

Table 5.2: Ge detectors with various issues in Runs 125-128.

	R125	R126	R127	R128
T1Z1	Three P channel broken, no LED			
T1Z3	Q outer broken			
T1Z5			Phonon B broken	
T2Z3				Phonon C broken
T2Z5				Low ^{133}Ba statistics
T4Z2				Low ^{133}Ba statistics
T5Z1	Distorted phonon pulse			
T5Z2	Intermittent Q outer short, glitch events			
T5Z6	High phonon noise			

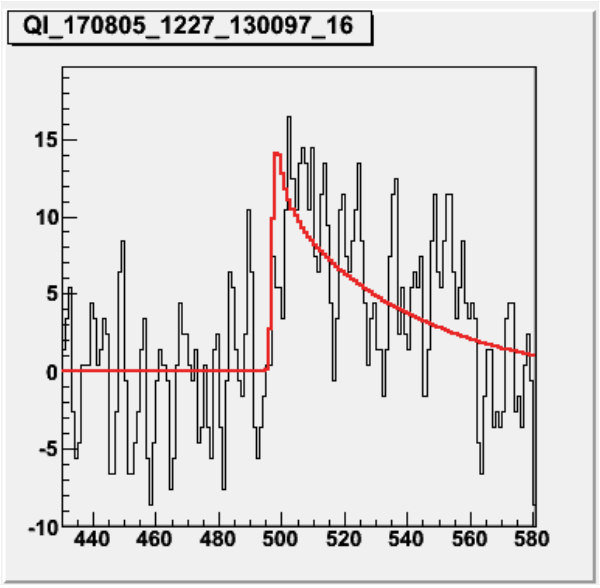


Figure 5.1: Raw charge pulse for the WIMP candidate event in T3Z4 with the optimal filter fitted pulse overlaid. Both the vertical and horizontal axes are in units of ADC bins. A later start time for the fitted pulse seems to give a better fit. *Figure from Lauren Hsu.*

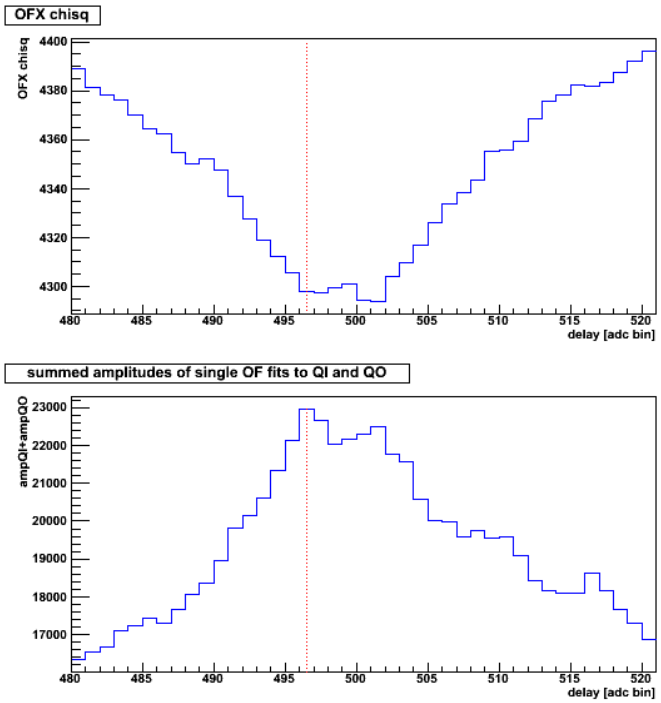


Figure 5.2: Top: Charge optimal filter χ^2 vs. fitted charge pulse start time with respect to the beginning of the pulse in ADC bins. The start time indicated by the vertical dotted line was picked by the charge optimal filter in the C58 analysis, which corresponds to the maximum of the sum of the amplitudes for Qi and Qo. The global minimum of the χ^2 is a few bins to the right, which corresponds to a local maximum of the summed amplitudes of Qi and Qo . Bottom: The sum of the fitted pulse amplitudes for Qi and Qo vs. the charge pulse start time. *Figure from Hsu [114].*

by choosing the maximum of the summed amplitudes of Qi and Qo as the objective function instead of the minimum of the χ^2 of the fit, in order to speed up the data processing in the C58 analysis. We can see this behavior in Fig. 5.2, which shows the optimal filter χ^2 of the fit and the summed amplitudes of Qi and Qo versus the start time of the fitted pulse for the WIMP candidate event in T3Z4.

Besides the two WIMP candidate events, there are two near-miss events passing the timing cuts in detectors T4Z2 and T4Z6, which are shown in Fig. 5.3. These two events were just above the 2σ nuclear recoil band. Other such events which were close to the WIMP signal region but somewhat further away were also present in a few detectors, e.g., T1Z5 and T4Z2. As we can see from Fig. 5.2, choosing a different minimum of χ^2 would change the start time and the amplitude of the ionization signal simultaneously, which in turn changes both the timing parameter and the ionization yield of the event. This would possibly drive the event to move closer to or further away from the WIMP signal region and change the result of the WIMP search.

Additionally, further checks in the C58 analysis on the ^{133}Ba calibration data found the events in the vicinity of the ionization energy threshold had a higher chance to pass the timing cuts for some of the detectors. Figure 5.4 shows such an example. The slow timing of the low charge energy events around the ionization energy threshold may be the symptom of the pathology of the charge reconstruction algorithm. A simulation of low energy charge pulses, constructed by adding scaled charge pulse templates and real noise traces from the data, was run through the charge optimal filter. It was found that there was a noticeable bias towards earlier start times at energies below the analysis threshold (1 to 2 keV), but not evident at the energy of the WIMP candidate events ($\sim$4 keV) [114]. Given the start time resolution of

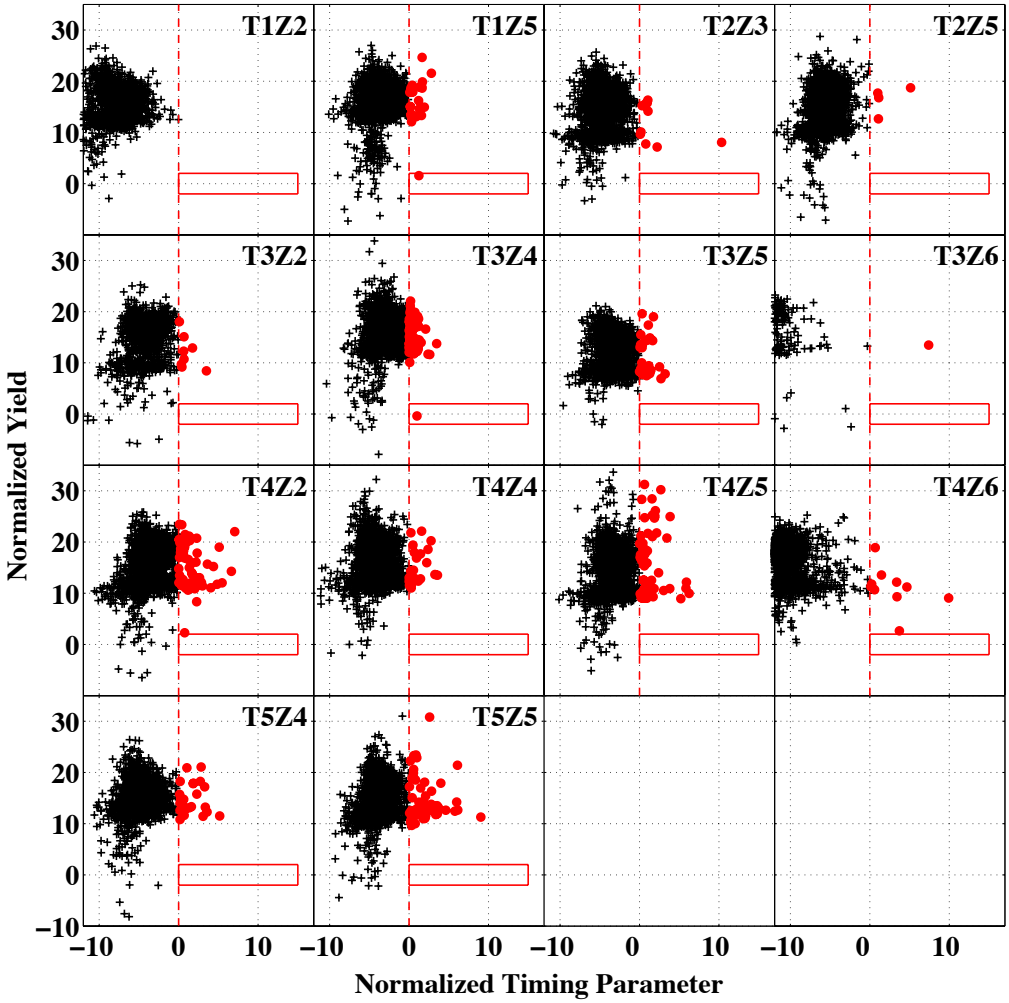


Figure 5.3: WIMP search events on the normalized ionization yield vs. normalized timing parameter plane. The vertical dashed lines show where the timing cuts are placed. The rectangular boxes to the right of the timing cuts define the WIMP signal regions. *Figure from Ahmed* et al. *[98].*

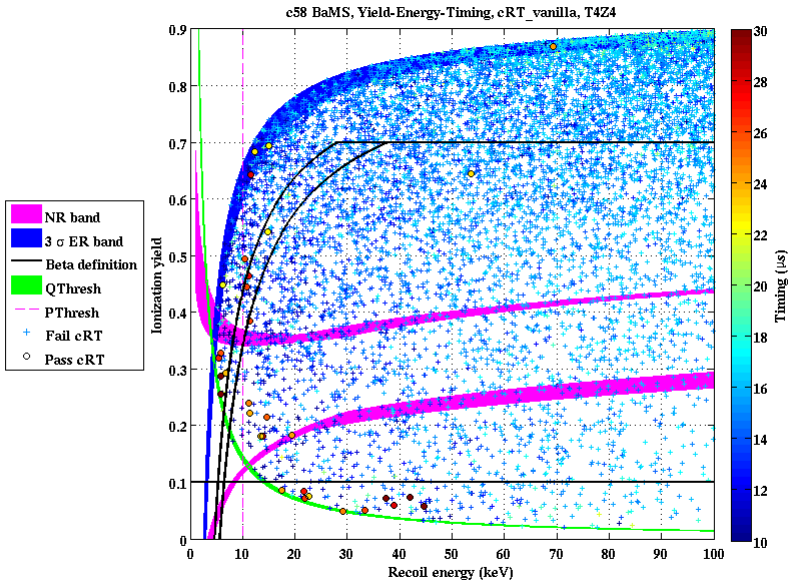


Figure 5.4: Multiple scatter events of the ^{133}Ba calibration data for the detector T4Z4. The events are colored by their timing parameter. The circles are the events passing the timing cut while the crosses are not. The 2σ electron recoil band lower bound (blue line), the 2σ nuclear recoil band edges (magenta line), the ionization energy threshold (green line), the phonon energy threshold (vertical dashed magenta line), and the surface events selection region boundaries (black line) are shown. The thicknesses of the lines and the gap between the curved black lines indicate the run-to-run variation. *Figure from Ahmed [87].*

the charge reconstruction algorithm measured on the simulated charge pulses, the probability of start time fluctuations by 4 μs or higher was at the ∼ 1% level for charge pulses with ionization energy below 6 keV [87, 114]. The same simulated pulses were then run through the optimal filter which minimizes the χ^2 of the fit rather than maximizes the summed pulse amplitudes. The resulted start time resolution was found about a factor of two better than in the previous case [87, 114]. The pulse amplitude resolution did not appear to improve with the change of the algorithm. However, interestingly, the χ^2 minimization gave slightly higher pulse amplitudes than the amplitude maximization.

Due to the extremely low WIMP-nucleon scattering cross section, direct dark matter search experiments like CDMS need to use very pure materials for detectors and shields to achieve very low radio backgrounds, and run the experiment for a long time in stable operation to accumulate exposure. This poses many challenges to the experiments and makes the experiments very expensive in cost, time, and manpower. We can be almost certain that, by removing the smearing of the timing parameter due to the pathology of the charge reconstruction algorithm, a higher exposure for the WIMPs can be obtained at the same background level from the existing CDMS-II data. A lower limit on the WIMP-nucleon cross section and/or a higher confidence level in claiming seeing a WIMP signal may be obtained. This is nothing less than building a better experiment and running it longer.

The charge reconstruction algorithm has significant impacts on the result of the WIMP search analysis. By switching to χ^2 minimization, the analysis would be able to benefit from better reconstructions of timing and ionization energy related parameters, and may be able to obtain more WIMP search exposure and improve

the WIMP search result. To properly account for the improvement to the charge reconstruction altorithm, the CDMS-II data taken during Runs 125-128 needs to be reprocessed and reanalyzed. The flaw of the charge reconstruction algorithm was found at the end of the C58 analysis, due to the limited time available before releasing the result, the decision of reprocessing and reanalyzing the data was deferred, and instead the surface event background was adjusted upward to reflect the impaired performance of the charge optimal filter.

Three timing analyses were used in the C58R analysis, which are the classic timing presented in this thesis, neural network timing analysis by Tommy Hofer, and the 5D χ^2 timing analysis [115] by Joseph Kiveni.

5.3 CDMS-II data analysis pipeline

The CDMS-II data analysis pipeline was a collaborative effort with contributions from many members of the collaboration over the years. General discussions about the pipeline and thorough coverage on various topics can be found in past CDMS publications and theses. Here we present a brief overview of the pipeline used in the C58 analysis, highlighting the sections which had changes in the reanalysis.

5.3.1 Data processing and calibration

The charge and phonon pulses are digitized and saved in the raw data files as well as the detector and hardware configuration settings. The energy and pulse shape information of an event must be extracted and properly calibrated, before the data can be used to do WIMP search analysis. The whole data processing and calibration

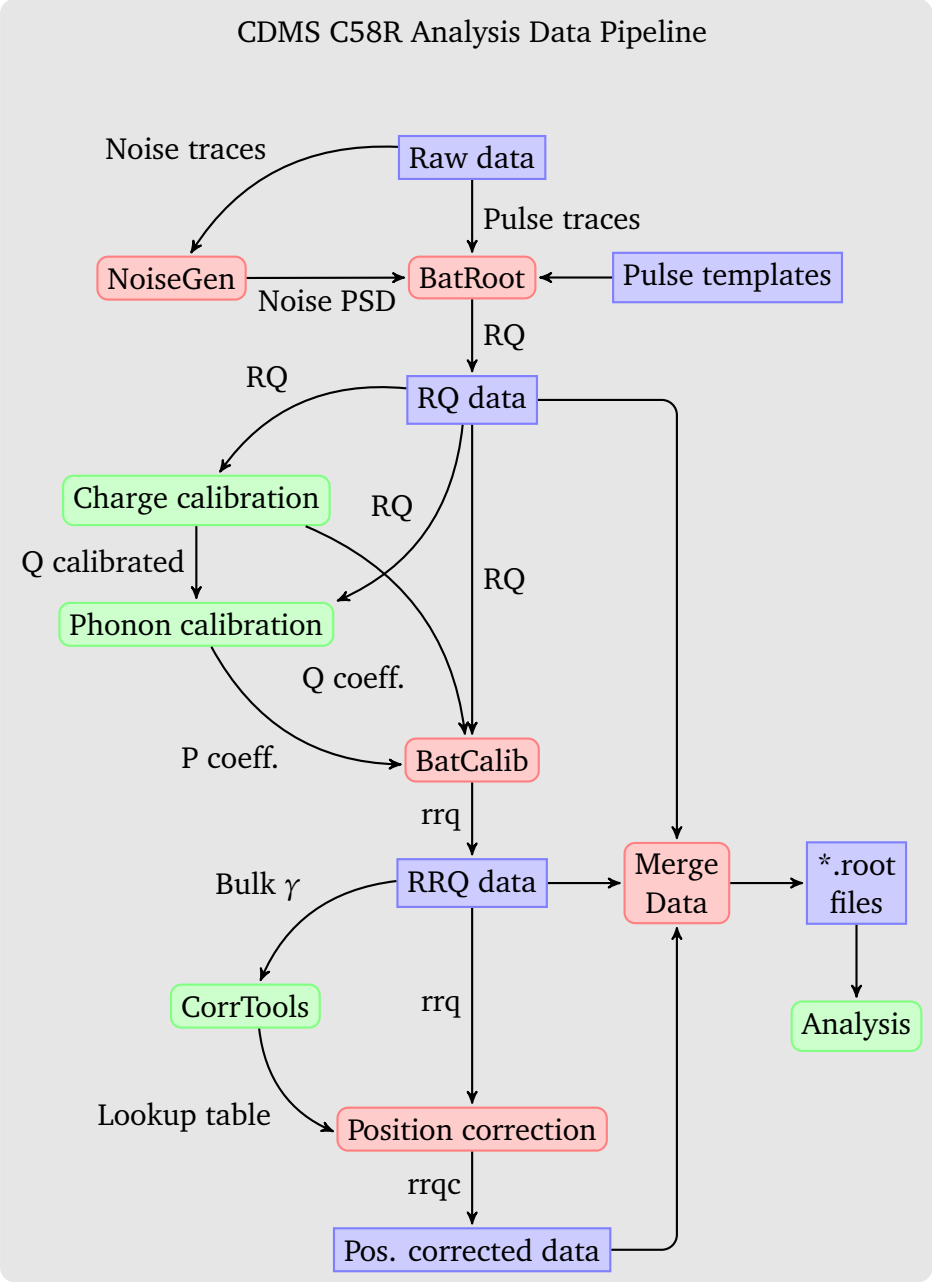


Figure 5.5: C58R analysis data pipeline.

pipeline for the C58R analysis is shown in Fig. 5.5. There are three stages in the whole process: data reduction, charge and phonon calibration, and position correction. The first two stages were performed with the CDMS data processing package "cdmsbats" (NoiseGen, BatRoot, and BatCalib) and the position correction was done in MATLAB with the lookup table generated with the MATLAB routine CorrTools. BatCalib was not ready for use in the C58 analysis and the calibration was done with the MATLAB package PipeCleaner. At the end, all data from the three stages were merged together into ROOT format files and got ready for WIMP search analysis.

5.3.1.1 Data reduction and optimal filtering

In the first stage, performed by NoiseGen and BatRoot, the digitized charge and phonon pulses are reconstructed by the pulse fitting algorithms and the energy and pulse shape information is extracted, generating the reduced quantities (RQ's). Detector and hardware configuration settings are also applied at this step, normalizing the charge and phonon energies to the output of their first stage amplifiers, in units of volts and amperes, respectively. Cdmsbats was designed modular. Several pulse reconstruction algorithms were implemented and applied to the data processing. We are concerned about the optimal filtering, which had the best performance and gave the RQ's used in this reanalysis as well as the C58 analysis. General theoretical discussions can be found in the appendices of Sunil Golwala's [85] and Filippini Filippini's [91] theses, the latter also discussed the implementation of the optimal filter for the CDMS-II data processing. I will cite the results from Filippini's thesis directly and point out one possible issue which might have caused the pathology of the charge reconstruction algorithm for the C58 analysis, and then give some

thoughts on how a proper χ^2 can be defined for a multi-channel system.

A recap to the optimal filtering

The idea behind the optimal filter used in CDMS pulse reconstruction is that, for a linear system, the noise fluctuations are Gaussian and the components at different frequencies are independent [85]. A maximum likelihood fit to the pulse in frequency domain simplifies to a χ^2 fit weighted by the noise power at each frequency. Given a digitized pulse

$$S_i = aA_i + n_i, \quad i = 0, \cdots, N-1, \tag{5.1}$$

where $S(t)$ is the real pulse, $A(t)$ is the expected pulse shape with unity amplitude, $n(t)$ is random noise with known power spectral density $J(f)$, digitized at times $t = i\Delta t$ during the time interval $T = N\Delta t$, the best estimate of the amplitude a is given by minimizing

$$\chi^2(a) = \sum_{n=0}^{N-1} \frac{|\tilde{S}_n - a\tilde{A}_n|^2}{J_n}, \tag{5.2}$$

where $\tilde{S}_n$ and $\tilde{A}_n$ are the discrete Fourier transforms (DFT) of time series S_i and A_i, respectively, for example,

$$\tilde{S}_n = \frac{1}{N}\sum_{k=0}^{N-1} S_k e^{-\frac{2\pi i n k}{N}}. \tag{5.3}$$

If the pulse template $A(t)$ needs to be shifted by t_0 to match the measured pulse, a phase factor $e^{2\pi i t_0 f}$ is introduced by the Fourier transform. The best estimates of the amplitude a and the start time t_0 are then given by minimizing

$$\chi^2(a, t_0) = \sum_{n=0}^{N-1} \frac{|\tilde{S}_n - ae^{-2\pi i t_0 f_n}\tilde{A}_n|^2}{J_n}, \tag{5.4}$$

where $f_n = \frac{n}{T}$. For a given t_0, the best estimator of a can be found by solving

$$\frac{\partial \chi^2(a, t_0)}{\partial a} = 0, \tag{5.5}$$

which gives

$$\hat{a}(t_0) = \frac{\sum_{n=0}^{N-1} e^{2\pi i t_0 f_n} \frac{\tilde{A}_n^* \tilde{S}_n}{J_n}}{\sum_{n=0}^{N-1} \frac{|\tilde{A}_n|^2}{J_n}}. \tag{5.6}$$

However, the best estimator for t_0 can not be obtained in a similar way because the equation

$$0 = \frac{\partial \chi^2}{\partial t_0} = -2a \sum_{n=0}^{N-1} 2\pi i f_n e^{2\pi i f_n t_0} \frac{\tilde{A}_n^* \tilde{S}_n}{J_n} \tag{5.7}$$

is nonlinear.

By noting that

$$\frac{\partial \hat{a}(t_0)}{\partial t_0} = \frac{\sum_n 2\pi i f_n e^{2\pi i t_0 f_n} \frac{\tilde{A}_n^* \tilde{S}_n}{J_n}}{\sum_n \frac{|\tilde{A}_n|^2}{J_n}} \propto \frac{\partial \chi^2}{\partial t_0}, \tag{5.8}$$

and checking the second order derivative, we find the value of t_0 which minimizes χ^2 also maximizes $\hat{a}$. The numerator of Eq. (5.6) is the inverse DFT of $\frac{\tilde{A}_n^* \tilde{S}_n}{J_n}$. Its computation speed can be optimized using the FFT algorithm, while this is not possible for Eq. (5.4).

The C58 data processing took a shortcut by maximizing Eq. (5.6) to find the best estimates of t_0 and a, which were then used to compute the χ^2 using Eq. (5.4). This is not a problem if only one pulse needs to be reconstructed at a time. Since χ^2 depends on a quadratically, for well behaved pulses, the only extremum is when a is

maximized and χ^2 is minimized simultaneously. However, if multiple pulses need to be fitted simultaneously, we will see that this could cause problems.

Extension for two dimensional problems

For the CDMS-II detectors having two charge channels, in the most general form, the signals Qi and Qo can be expressed with four pulse amplitudes a_{ij}, four pulse templates A_{ij} including two crosstalk terms, and two noise terms n_i, in the form below:

$$S_1 = a_{11}A_{11} + a_{12}A_{12} + n_1 = s_1 + n_1, \tag{5.9}$$

$$S_2 = a_{21}A_{21} + a_{22}A_{22} + n_2 = s_2 + n_2, \tag{5.10}$$

where s_i are the expected signals without noise. The pulse template A_{ij} is the induced pulse in Channel i by the unit amplitude pulse A_{jj} in Channel j. The crosstalk terms are linear in the first order, which means we only need to have one template for each channel, and the crosstalk terms can be factored into a product of the pulse of a neighboring channel and a coupling constant which can be absorbed into the amplitudes a_{ij}. In practice, if the crosstalk is not strictly linear, it would be beneficial or necessary to specify all the crosstalk templates. The explicit expressions of S_1 and S_2 are only used to assist our discussions below. Details of the realization do not matter here.

Then following the idea in Filippini's thesis [91], the χ^2 is constructed by fitting

the two signals simultaneously:

$$\chi^2 = \sum_n \left(\frac{|\tilde{S}_1 - \tilde{s}_1|^2}{J_1} + \frac{|\tilde{S}_2 - \tilde{s}_2|^2}{J_2} \right). \tag{5.11}$$

CDMS implementation and issues

The implementation of the optimal filter for charge reconstruction in the CDMS-II data processing is as follows:

1. The amplitudes of Qi and Qo are estimated using single pulse optimal filter independently for a range of start times. The start time which gives the maximum summed amplitudes is chosen as the true start time.
2. The χ^2 with crosstalk accounted for are then calculated with the amplitudes and the start time from the first step.

For the first step, Filippini pointed out that there was an ambiguity in choosing which amplitude to maximize and the summed amplitudes was an reasonable choice. Since the summed amplitudes characterize the total ionization energy of an event, it seems there is no better alternative unless we have a reason to weigh one channel more than the other. However, the assumption that the two pulses have the same start time may not describe the data very well especially for the low charge pulses, where a clear fast rising edge does not exist because of noise. The real start time for these pulses can not be precisely determined and strongly depends on the noise. If the start times of the two pulses identified by the optimal filter are different and the offset is large enough, there would be double maxima for the summed amplitudes. This is exactly what we saw in Fig. 5.2. The multiple maxima issue is just intrinsic to the multi-dimentional problems. There is not much we can do besides check each

maxima and pick the one satisfying our criteria.

As for using the minimum χ^2 as the criterion to choose the optimal fit, which was the case for the reprocessing, besides a large penalty in computing speed, the multi-minima issue still exists. Then the question would be which minimum should we choose, or which pulse should we weigh more if the chosen minimum does not optimize each pulse at the same level? In the scenario where there are two traces but only one contains the signal, it is reasonable to only fit the one with the signal and discard the other one instead of doing a fit and then carrying its χ^2 forward. Along this line, it seems reasonable to use the energy fraction to weigh the χ^2 of each pulse in Eq. (5.11). So we can redefine the χ^2 as

$$\chi^2 = \sum_n \left(\frac{a_1}{a_1 + a_2} \frac{|\tilde{S}_1 - \tilde{s}_1|^2}{J_1} + \frac{a_2}{a_1 + a_2} \frac{|\tilde{S}_2 - \tilde{s}_2|^2}{J_2} \right) \tag{5.12}$$

to reflect their relative importance.

Impact of time jitter

For a discrete pulse, the start time t_0 can only be determined with the accuracy of one digitizing bin $\Delta t = \frac{T}{N}$. An offset to the real start time underestimates the amplitude and overestimates the χ^2. For any specific pulse, the start time offset lies between 0 and Δt. Statistically, the rms offset is $\delta_t = \frac{\Delta t}{\sqrt{12}}$. This uncertainty introduces the

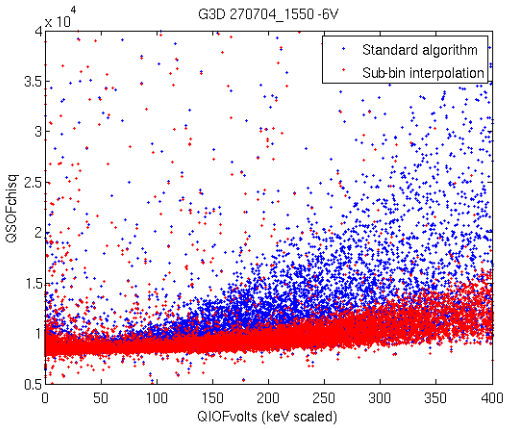


Figure 5.6: Optimal filter χ^2 vs. ionization energy. The distribution of the χ^2 with charge pulse start time sub-bin interpolation (red) is much narrower than without (blue). *Figure from Filippini [91].*

second order uncertainties to the amplitude a and the χ^2:

$$\delta a \approx \frac{1}{2}\frac{\partial^2 a}{\partial t_0^2}\delta_t^2 = -a\delta_t^2 \frac{\sum_n (2\pi i f_n)^2 \frac{|\tilde{A}_n|^2}{J_n}}{\sum_n \frac{|\tilde{A}_n|^2}{J_n}}, \tag{5.13}$$

$$\delta\chi^2 \approx \frac{1}{2}\frac{\partial^2 \chi^2}{\partial t_0^2}\delta_t^2 = \frac{1}{2}a^2\delta_t^2 \sum_n (2\pi i f_n)^2 \frac{|\tilde{A}_n|^2}{J_n}. \tag{5.14}$$

The errors on amplitude are quite small, ∼0.15 % on average and ∼0.5 % at the most. The errors for the χ^2 scale with the square of the pulse amplitude and can be quite large at higher energies. Filippini [91] developed a method to better estimate the start time by interpolating the three bins around the minimum of the χ^2 quadratically and finding the time corresponding to the interpolated minimum. Figure 5.6 shows the improvement of this technique to χ^2.

The effect of switching from maximizing summed charge pulse amplitudes to minimizing the χ^2 in data processing is equivalent to that of a start time offset. The changes induced by the algorithm change to different quantities are estimated to be:

charge pulse start time $\sim \frac{1}{\sqrt{12}}\Delta t$ (0.23 μs), ionization energy $< 1\,\%$, and χ^2 ∼100 %.

5.3.1.2 Charge and phonon calibration

Charge and phonon energies in the reduced quantities produced by BatRoot are in units of volts and amperes. They need to be calibrated to a known energy scale in order to show the energy deposition of the scattering event correctly. The ionization energy is calibrated with the 356 keV line in the ^{133}Ba calibration data. Phonon energies are calibrated with respect to charge. The calibration constants derived from the ^{133}Ba calibration data are then applied to all data sets including the ^{252}Cf and the WIMP search data.

Charge calibration

Charge calibration performs three corrections to the uncalibrated ionization energies of the inner and outer charge channels, `QIOFvolts` and `QOOFvolts`, and generates the corrected quantities, `qi` and `qo`, in units of keV.

The first step is the overall energy scale calibration, in which the 356 keV peak of the spectrum for the summed energy of Qi and Qo in the ^{133}Ba calibration data is fitted to a Gaussian, and the mean of the Gaussian is then scaled to 356 keV. Figure 5.7 shows the ^{133}Ba calibration data charge spectrum after overall calibration.

The second step corrects the residual crosstalk between the two charge channels. The most part of the crosstalk is taken care of by the optimal filter in the data processing. However, the data still contains residual crosstalk of the second order. A linear transformation is performed on Qi and Qo to make them orthogonal while keeping the sum of them unchanged. Figure 5.8 shows the plots of `QOOFvolts` vs.

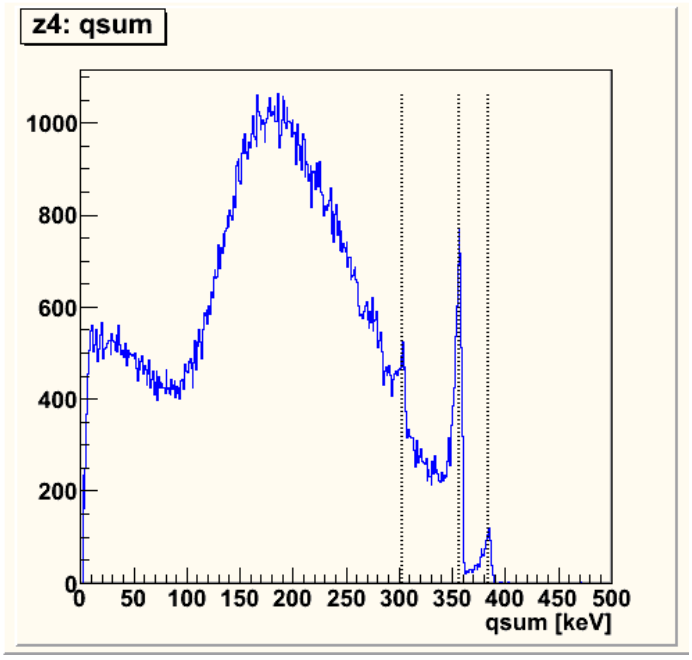


Figure 5.7: Calibrated energy spectrum of the summed energy of Qi and Qo in the ^{133}Ba calibration data. The 356 keV line is calibrated to the correct position. The energy lines at 303 keV and 384 keV show that the ionization signal has good linearity with respect to the photon energy. *Figure from Hall and Hsu [116].*

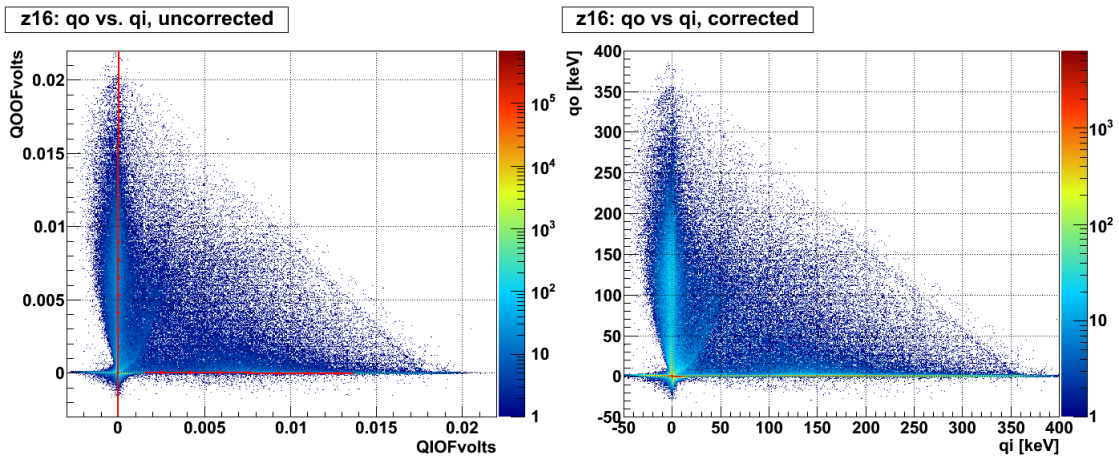


Figure 5.8: Residual crosstalk correction in charge calibration. The left figure shows `QIOFvolts` vs. `QOOFvolts` before correction and the right figure after the correction. Note that the overall energy scale is also corrected in the right figure. The two figures are colored by point density. *Figure from Fallows [117].*

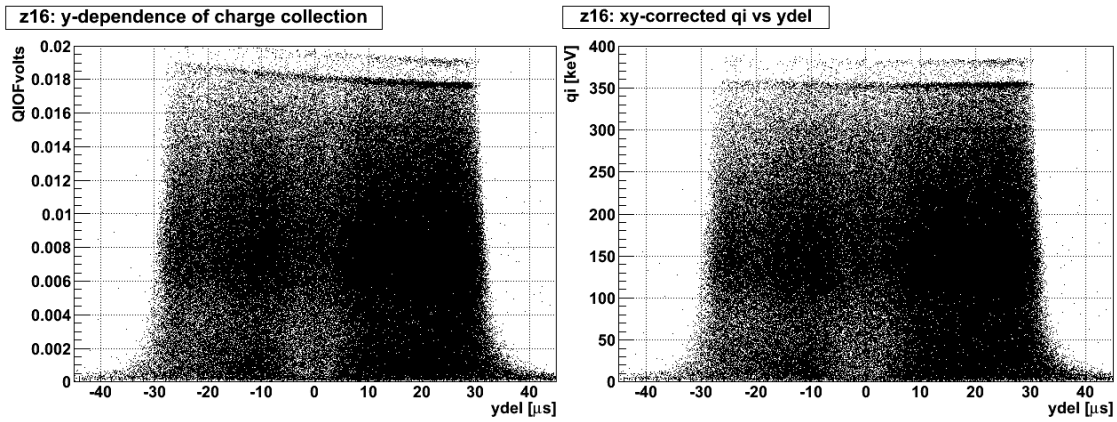


Figure 5.9: Ionization energy of Qi vs. phonon relative delay in y-direction (`ydel`) before (left) and after (right) position correction. Note the overall energy scale is also corrected in the right figure. *Figure from Fallows [117].*

`QIOFvolts` before and after crosstalk correction.

Position dependence of charge signal in the x-y plane is corrected in the third step. Due to the nonuniformity of the detector, the magnitude of the ionization signal from events with the same energy varies with the position of the event in the x-y plane. We use the relative delays of two neighboring phonon channels with respect to the primary channel in x and y directions, `xdel` and `ydel` to characterize the event position. Figure 5.9 shows the position dependence of the ionization energy of the inner channel of a detector in the y-direction before and after correction. The 356 keV line is fit to a fourth order polynomial, and then all events are normalized by the fitted function with their `ydel`. The position dependence in the x-direction is treated in the same way.

At the end, the calibration constants in the three steps are combined together get the final values, which are then applied to all data sets through BatCalib.

5.3.1.3 Phonon calibration

There are no clearly identifiable ^{133}Ba lines in the phonon energy spectrum that can be used for calibration, due to their much worse energy resolution. Instead, the phonon energies are calibrated with respect to calibrated charge by normalizing the ionization yield of gamma induced electron recoils to 1. The ionization yield of an event is defined as the ratio between the ionization energy and the recoil energy:

$$y = \frac{E_q}{E_r}. \tag{5.15}$$

The yield for bulk electron recoils is conveniently chosen to be unity. The recoil energy E_r and the work done to drift the charge out of the crystal are both dissipated as phonons eventually. The measured total phonon energy is the sum of the two:

$$P_t = E_r + \frac{eV_b}{\epsilon} E_q, \tag{5.16}$$

where e is the electric charge, V_b is the voltage bias across the detector, and ϵ is the average energy needed to produce one electron-hole pair. The phonons generated in drifting charge are referred to as the Neganov-Luke phonons (or more commonly Luke phonons in the US). The gamma-equivalent recoil energy for all events is then:

$$P_{rg} = \frac{P_t}{1 + \frac{eV_b}{\epsilon}}, \tag{5.17}$$

and the gamma-equivalent yield

$$y_g = \frac{E_q}{P_{rg}}, \tag{5.18}$$

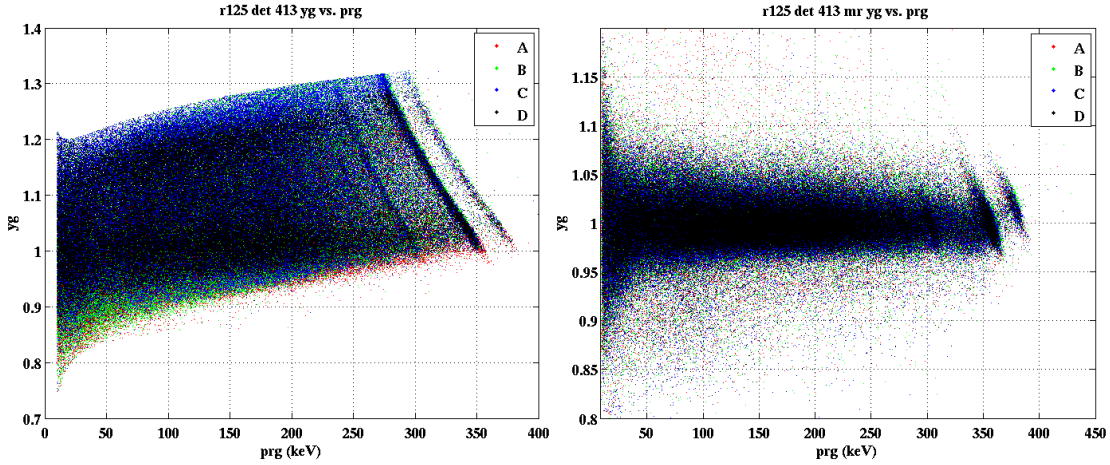


Figure 5.10: Gamma-equivalent ionization yield vs. gamma-equivalent recoil energy after phonon overall energy calibration (left) and phonon position correction (right). *Figure from Hertel and Ahmed [118].*

which is defined to be 1. The overall phonon energy calibration is done by minimizing the χ^2 defined as

$$\chi^2 = \sum_n (E_q - P_{rg})_n^2 \tag{5.19}$$

for all events. The left figure in Fig. 5.10 shows the gamma-equivalent yield versus gamma-equivalent energy after the overall phonon energy calibration. We can see that yg has significant recoil energy dependence and the distribution of yg at a given prg is very broad. This will be corrected in the phonon position correction discussed below.

The performances of the four phonon sensors on one detector are not identical even though great care was taken to make them as uniform as possible. A second step called phonon relative calibration is then performed after the overall energy calibration to normalize the phonon sensor response. Four weighting factors whose sum is unity are each assigned to a phonon channel. They are then adjusted to make

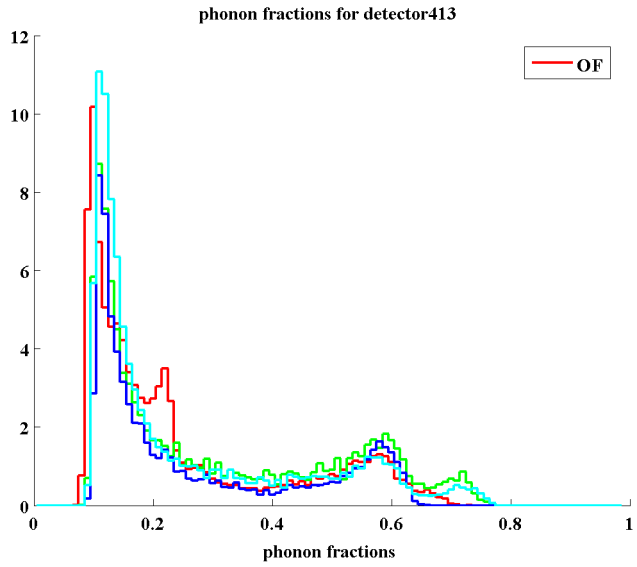


Figure 5.11: Phonon energy fraction distributions of the four phonon channels on a detector after the relative calibration. The four histograms represent the four phonon channels. *Figure from McCarty [119].*

the distributions of the phonon energy fraction of all channels match as much as possible. Figure 5.11 shows the phonon energy fraction distribution after relative calibration.

5.3.1.4 Phonon position correction

The phonon relative calibration removes much of the non-uniformity at the sensor level, however, as we can see in Fig. 5.10, there is still significant non-uniformity in phonon signal at finer scales associate with geometry and phonon propagation details. We take this position and energy dependence out by a procedure caled phonon position correction. The idea is, first select an ensemble of events to characterize the position dependent response of the detector in the bulk, then all measured quantities are normalized by the local gain at where the event occurred. Note that this procedure

only removes the nonuniformity of the detector response in the bulk and along the x-y directions. We don't normalize the detector response along the z direction, which would impair the detector performance to reject surface events [88].

First a lookup table is created with bulk gammas within the interested energy range. The position of an event in the manifold of the lookup table is given by

$$\vec{r} = (\texttt{xpart},\ \texttt{ypart},\ \frac{\texttt{xdel}}{L_d},\ \frac{\texttt{ydel}}{L_d},\ \frac{\texttt{prg}}{L_e}), \tag{5.20}$$

where L_d and L_e are normalization constants to make each term take roughly equal weight. The distance is defined with the metric:

$$r = \sqrt{\texttt{xpart}^2 + \texttt{ypart}^2 + \frac{\texttt{xdel}^2}{L_d^2} + \frac{\texttt{ydel}^2}{L_d^2} + \frac{\texttt{prg}^2}{L_e^2}}. \tag{5.21}$$

Any quantity α_i that needs to be corrected is then normalized by the local average (gain) of its nearest neighbors:

$$\alpha_{ic} = \alpha_i \frac{\langle\alpha\rangle_g}{\langle\alpha\rangle_n}, \tag{5.22}$$

where α_{ic} is the corrected quantity, $\langle\alpha\rangle_g$ is the global average of α, and $\langle\alpha\rangle_n$ is the average of the nearest neighbors of α_i. Because the event distribution is discrete, the quantity under correction may not be right in the center of its nearest neighbor cluster. To account for this effect, a gradient correction method was also introducted, which is

$$\alpha_{ic} = \langle\alpha\rangle_g \left(\frac{\alpha_i}{\langle\alpha\rangle_n} - \frac{\alpha_i}{\langle\alpha\rangle_n^2} (\vec{r} - \langle\vec{r}\rangle_n) \cdot \frac{\partial\alpha}{\partial\vec{r}}|_{\vec{r}\approx\langle\vec{r}\rangle_n} \right), \tag{5.23}$$

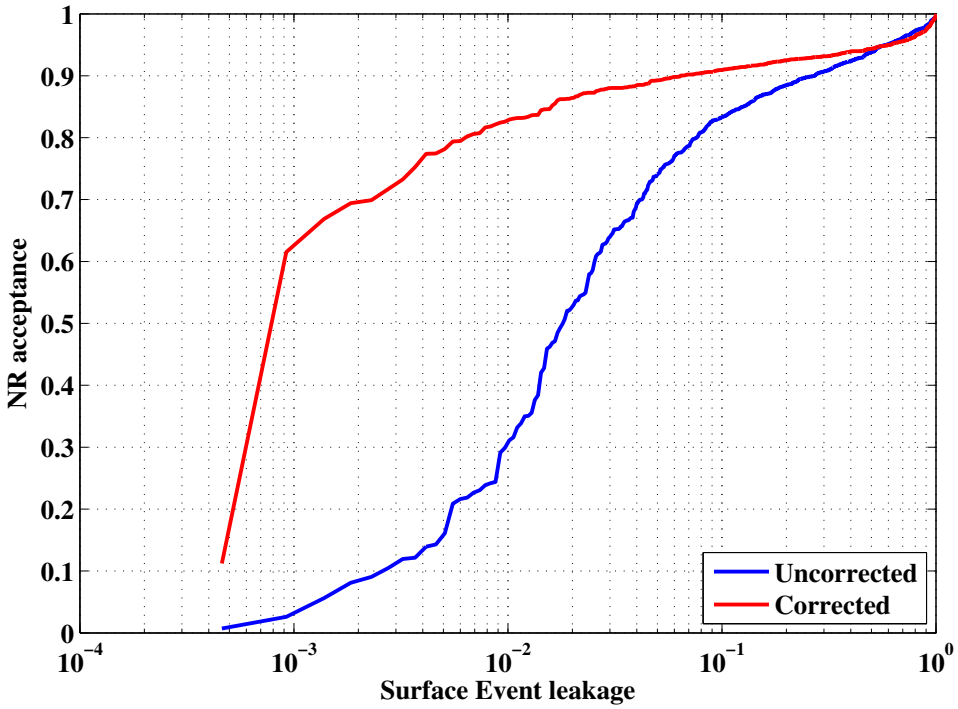


Figure 5.12: Neutron acceptance efficiency measured on ^{252}Cf data vs. surface event rejection efficiency measured on surface events from the ^{133}Ba calibration data. The phonon position corrected quantities provide much better surface event rejection performance. *Figure from Ahmed [87].*

where $\langle \vec{r} \rangle_n$ is the geometrical center for the nearest neighbor cluster of α_i. The number of nearest neighbors used in the average and the scales of L_d and L_e are all optimized to have small statistical errors while still reflect the gain change on small scales.

The right figure of Fig. 5.9 shows the photon-equivalent ionization yield versus photon-equivalent recoil energy. Both the spread of yg and the energy dependence were reduced by position correction. Figure 5.12 shows the improvement of surface event discrimination power with phonon position correction on calibration data sets.

Phonon position correction takes a lot of computing power and thus only a

selection of yield and timing quantities were corrected. Discussions in greater depth can be found in the theses of Ahmed [87], Hertel [88], and McCarthy [89].

5.3.2 Blinding

The CDMS WIMP search analysis including previous analyses as well as this reanalysis followed a philosophy called "blind analysis". The expected signal region in the WIMP search data is masked during the analysis. All cuts are tuned using the WIMP search data outside the signal region and calibration data. The signal region is then unmasked once all the cuts are finalized. The blind analysis minimizes potential human bias introduced by conditional cut tuning to obtain the desired result, and also simplifies the statistical characterization of the WIMP signal.

Defining the WIMP signal region in the WIMP search data for the blinding cut is a simplified version of the WIMP search analysis. In the analyses of Runs 123-124 and Runs 125-128, the blinding cuts were developed using the quantities before the phonon position correction [87, 91]. Due to the poorer resolution of the uncorrected phonon quantities, the cuts were designed to cover a wider region (or over-blind) to ensure the coverage [91]. The blinding cut in this reanalysis was updated on top of the final version of the blinding cut of the C58 analysis. This blinding cut covers a tighter region compared to the initial version. The criteria used in selecting the WIMP signals for blinding are

1. Veto-anticoincident: the veto shield should stay quiet during the 50 μs window before the event trigger.
2. Energy range: recoil energy should be above the higher value of the detector threshold and 2 keV, and below 130 keV.

3. Nuclear recoil like: ionization yield should be within the 2σ nuclear recoil band.
4. Inner charge electrode events: the event deposits all or most of its energy in the inner charge channel, the energy deposited in the outer charge channel is within 2σ of the noise.
5. Single scattering events: the total phonon energy of the event is 6σ above its noise floor for one detector and within 4σ for all the other detectors.

5.3.3 Data selection cuts

A number of cuts are defined to select good data to be used in the analysis. They can be classified into two categories. The first one is called data quality cuts, which identify and exclude time periods when there were hardware failures or the detectors had poorer than nominal performance. The second category is referred to as reconstruction quality cuts and they remove the events which are not well reconstructed in the data processing. A brief discussion of these cuts is presented in this section. More details can be found in the theses of Ahmed [87] and Filippini [91].

5.3.3.1 Data quality cuts

Table 5.3 lists the data quality cuts used in the C58 analysis, divided into a number of subsets according to the similarity of the issues they address. These cuts ensure the data were taken with working detectors, nominal charge and phonon biases, correct trigger settings and trigger information, low charge and phonon noise, and sufficient LED flashing to neutralize the detector. Data with inter-detector crosstalk, bad phonon regions, helium film plating the phonon sensors, and taken during other

Table 5.3: Data quality cuts used in the C58 analysis.

cGoodDet_c58	Selects working detectors for each run
cQnormbias_c58	Select events with nominal charge bias
cStabTuning_c58	Selects periods with stable phonon sensor bias
cErrMask_c58	Select events with erroneous trigger bits
cGlitch_c58	Selects events caused by electronics glitches
cTrigBurst_c58	Selects periods with abnormally high trigger rates
cPlo_Disabled_c58	Selects periods when plo trigger is disabled
cHighQNoise_c58	Selects periods with high charge noise
cBadResLoose_c58, cBadResTight_c58	Select periods with poor phonon optimal filter amplitude and start time resolution
cBadDet_Ba_c58, cBadDet_bg_c58	Selects datasets with low KS statistics
cBadNeutBa_c58, cBadNeutBg_c58	Select datasets (or subsets) with high fraction of low yield events (poor neutralization)
cBadFlash_c58	Select datasets with insufficient LED flashing to neutralize the detector
cNoDDxtalk_c58	Reject events with inter-detector charge crosstalk in Towers 1 and 2
cBadDetRegions_c58	Selects events with bad phonon regions on some detectors
cHeFilm_c58	Selects periods when phonon sensors plated with helium films
cBadTimes_c58	Selects bad time periods, primarily after power outages
cNuMI_c58	Selects times when the MINOS neutrino beam is active

operator labeled bad times such as power outage are excluded. Kolmogorov–Smirnov (KS) tests are also performed on all datasets with 30 selected reference datasets. Data with low average KS statistics are also excluded.

Note that the cuts `cBadDet_bg_c58` and `cNoDDxtalk_c58`, which depend on the ionization energy and yield were updated in the C58R analysis. The neutralization cuts, which count the fraction of events with ionization yield between 0 an 0.8 in the energy range of `pt` $> 10\,$keV and `qi` $< 500\,$keV , also depends on ionization yield. However, since the optimal filter change in the reprocessing only affects the reconstructed energy for a small fraction of the low charge events at the 1 % level, it is unlikely the neutralization cut will change. So the neutralization cuts were kept unchanged for the Ge reanalysis.

5.3.3.2 Reconstruction quality cuts

The charge and phonon optimal filter pulse reconstruction algorithms are based on the assumptions that both the signal and noise are stationary. The corrupted pulses due to events pile up or noise fluctuations will not be reconstructed correctly. A number of reconstruction quality cuts are developed to deal with these corrupted pulses. They are listed in Table 5.4.

In principle, these events with corrupted pulses can be cut out by a cut placed on the χ^2. This is the case for the charge pulses, which don't have much pulse shape variation, and a cut based on the charge χ^2 was developed. The large variation of phonon pulse shape, however, makes it impractical for the phonons. Instead a cut on phonon pre-pulse baseline `Pstd_c58` was developed to select the phonon pulses with pileups in the pre-pulse. Phonon pulses with high baseline noise are also cut by this

Table 5.4: Reconstruction quality cuts used in the C58R analysis.

cChiSq_c58	Select events with good charge χ^2
cQstd_c58	Cryocooler cut using charge baselines
cGoodPStartTime	Enforces OF search window
cPstd_c58	Phonon pre-pulse baseline
cGoodRTFTManifold_c58	Selects good position corrected pipefitter timing quantities

cut. A similar cut for the charge Qstd_c58 was also developed to select charge pulses with excessive baseline noise.

Another class of events which can not be constructed correctly are the inter-detector pileups, i.e., when two events happen closely in time on two detectors. The first event issues a global trigger and the second event is also read out. The optimal filter has a start time search window of $[-100,\ +10]\,\mu$s for the charge pulses and $[-50,\ +200]\,\mu$s for the phonons. Sometimes the phonon pulse of the second event is reconstructed correctly because of the larger start time search window but the charge pulse is not. The cut cGoodPStartTime was developed to enforce the phonon start time is within the overlap of the charge and phonon start time search window [87, 120, 121].

There are outlier events in phonon position which are anomalously far from their nearest neighbors in the lookup table measured on two projected distances on subspaces (pminrt, pdel) and (xpart, ypart, xdel, ydel) [87, 122]. The cut cGoodRTFTManifold_c58 was developed to exclude these outliers.

Table 5.5: Physics cuts used in the C58 analysis.

`cVTStrict_c58`	Selects events with activity in the veto
`cSingle_c58`	Selects single scattering events
`cER_c58`	2σ electron recoil band defined with ^{133}Ba data
`cBelowER_c58`	Events below the 3σ electron recoil band
`cNR_c58`	2σ nuclear recoil band defined with ^{252}Cf data
`cQin_c58`	Charge fiducial volume
`cQThresh_c58`	Ionization energy threshold
`cRT_vanilla_c58`	Surface event rejection using `pminrtCFc + pdelCFc` as a discriminator

5.3.4 Physics cuts

With all good events selected, we define the criteria to select the WIMP signals. Table 5.5 lists the majority of the physics cuts used in the C58 analysis. The conservative versions of these criteria for the blinding cut were touched upon in Sec. 5.3.2 except the surface event rejection cut. This section presents the final version of these cuts and the surface event rejection cut.

5.3.4.1 Scintillator veto

A muon veto shield consisting of 40 plastic scintillator panels was built surrounding the passive Pb/polyethylene shield of the experiment. Ambient radiation passing through the scintillator panels is recorded. The cut `cVTStrict_c58` was defined so that any WIMP-like events from the ZIPs coinciding with the veto events would be excluded.

The veto activities are recorded in two ways. First, the time when the PMT output from any panel is above the hardware threshold is recorded in resolution of $1\,\mu$s. Second, the veto trace in a time window of $[-185,\ 25]\,\mu$s around a ZIP event trigger

during WIMP search data run is digitized and recorded. The veto trigger rate in the Runs 125-128 was around 400 Hz (10 Hz per panel) at hardware thresholds ∼2 MeV. Most of the triggers were caused by photons whose deposited energy in the panel is less than 3 MeV. The hardware thresholds were chosen to stop most of the ambient photons but not all of them. The muons penetrating the 713 m rock overburden deposit approximately 10 MeV when traveling through a panel. Panel specific veto trace amplitude thresholds between 3 and 5 MeV are used in `cVTStrict_c58` to identify high energy muons.

Two criteria are devised to define the veto-coincident events based on energy or time:

1. The amplitude of the recorded veto trace is above the amplitude threshold.
2. The veto output is above the hardware threshold in the 50 μs window proceeding a ZIP event trigger.

The efficiency of this cut measured on random triggers was 97.88 %, 2.12 % of the random triggers were tagged as veto-coincident events [103, 123]. More details about this cut can be found in the thesis of Fritts [103].

5.3.4.2 Single scatters

Due to the extremely low WIMP-nucleon scattering cross section, a WIMP is expected to scatter only once with the entire detector set of the experiment. Background particles like neutrons and gammas have much larger scattering cross section and commonly scatter multiple times with the detector set. The single scatter event cut `cSingle_c58` was defined to reject the multiple scatter backgrounds.

A single scatter event requires the total phonon energy `pt` of an event is 6σ above

the average pt of random triggers for one detector and within 4σ of the average pt of random triggers for all the other detectors. The criteria were loosened for some of the detectors in the C58 analysis to increase the efficiency. The efficiency of cSingle_c58 in the C58 analysis was ∼99 %. More details can be found in Refs [87, 124].

5.3.4.3 Ionization yield

The WIMPs are expected to interact with the nuclei of the detector material. The ionization yield of such events are required to satisfy:

1. Within 2σ of the nuclear recoil band.
2. At least 3σ below the electron recoil band.

The nuclear recoil band is defined by the cut cNR_c58 calculated with ^{252}Cf calibration data. The electron recoil band cER_c58 was calculated with ^{133}Ba calibration data. The 3σ lower edge of the electron recoil band is defined by the cut cBelowER_c58.

The efficiency of the cut was measured on ^{252}Cf data and taken as the ratio between the number of events in the 2σ nuclear recoil band and that within the 4σ nuclear recoil band. Leakages from the electron recoil band was also corrected. The final efficiency was ∼94 %.

5.3.4.4 Fiducial volume

Electron recoil events occurring close to the side wall of a detector may suffer from incomplete charge collection and result in lower ionization yield. These events may fall into the nuclear recoil band and become a background to the expected WIMP signal. The charge fiducial volume cut cQin_c58 rejects these events by requiring that the ionization energy of an event deposited in the outer charge channel is within

$\pm 2\,\sigma$ of the noise, i.e., the Qo signal is noise like.

The efficiency of this cut was $\sim$75 % at low energies and drops slightly as energy increases because of increased probablity of energy deposition on Qo. More details can be found in Ref. [125].

5.3.4.5 Energy thresholds

The ionization energy thresholds are set $4.5\,\sigma$ above the mean ionization energy of the charge noise traces for each series and detector, rejecting all events with noise-like charge signals from the analysis.

The recoil energy threshold was set at 10 keV to ensure:

1. Phonon timing parameter is still effective to discriminate surface events from nuclear recoils.
2. There is clear separation between the nuclear recoil band and the electron recoil band and there is still sufficient population of events in the nuclear recoil sidebands to tune the timing cut.
3. Event leakage from electron recoil band to nuclear recoil band must be small.

The ionization threshold cut was redefined in the reanalysis but the recoil energy threshold was kept unchanged.

5.3.4.6 Surface event rejection cut

With all the above cuts defined applied to the WIMP search data, the expected WIMP signal, surface event background, and neutron background will all show up in the signal region. A surface event rejection cut utilizing phonon timing information is defined to reject the dominant background due to surface events.

The timing cut `cRT_vanilla_c58` in the C58 analysis was defined on the timing parameter (`pminrt` + `pdel`). Several other forms of timing cuts were also developed in the C58 analysis. Nevertheless, all of them were linear combinations of different versions of `pminrt` and `pdel`. So the essential information is encoded in these two timing quantities.

Bulk nuclear recoil events tend to have longer delay time and rise time relative to the surface events. By requiring that the timing parameter exceeds a certain threshold, most of the surface events can be rejected while keeping a large fraction of the nuclear recoil events. In the C58 analysis, a timing cut was set for each detector, and the cuts were optimized by maximizing neutron acceptance at a total surface background of 0.5 events [87, 98, 103].

The efficiency of `cRT_vanilla_c58` was ∼60 % for interior detectors and as low as ∼10 % for endcap detectors because their exterior face could not reject multiple scatters.

5.3.5 WIMP search exposure and sensitivity

With all data quality cuts, and WIMP signal criteria and their efficiencies calculated, the spectrum averaged exposure (SAE) τ for WIMPs with given mass is

$$\tau = M\,T\,\frac{\int_{E_l}^{E_u} dE\,\epsilon(E)\,\frac{dR}{dE}}{\int_{E_l}^{E_u} dE\,\frac{dR}{dE}}, \tag{5.24}$$

where M is the target mass, T is the livetime, $[E_l,\ E_u]$ is the recoil energy range, $\epsilon(E)$ is the detector efficiency, and $\frac{dR}{dE}$ is the differential recoil spectrum for a WIMP. The detector efficiency $\epsilon(E)$ depends on the timing cut positions, which are determined

for a given target background level B. So the SAE is a function of the number of expected background events, and it can be written as $\tau(B)$.

With the expected background B, the upper limit at confidence level α on the number of observed background events can be calculated. There are three commonly used methods to do this. The first one is the standard Poisson method, which is used in CDMS to set the sensitivity when tuning the timing cuts. The second is the Feldman and Cousins approach [46], which gives a more proper and more conservative frequentist confidence interval. The third is Yellin's optimal interval method [48, 126], which gives stronger (lower) limits if unknown backgrounds can not be ruled out from the observed events. It is used in CDMS to generate the WIMP-nucleon cross section limit after unblinding the data. There are no existing analytical forms for the latter two methods and pre-calculated tables must be used. The distincitons between the three methods is not essential in tuning the timing cuts because we only want to know what timing cut positions give the best sensitivity to WIMPs. Using the Poisson method, the upper limit of number of expected background events N_α is

$$N_\alpha = \sum_{k=0}^{\infty} \frac{e^{-B} B^k}{k!} \frac{1}{2} F^{-1}_{\chi^2}(\alpha; 2(k+1)), \tag{5.25}$$

where $\frac{1}{2}F^{-1}(.)$ is the upper bound of the Poisson interval at CL α with k events observed. If no background events show up in the final result of the experiment due to statistical fluctuation, N_α is the minimum number of WIMPs needed such that the WIMPs can be detected by an experiment with probability α on average.

The sensitivity of the experiment to WIMP-nucleon cross section can be calculated once τ and N_α are known.

5.4 Cuts retuned in reanalysis

The analysis on the optimal filter in Sec. 5.3.1.1 shows that the change of the charge reconstruction algorithm from maximizing amplitude to minimizing χ^2 only affects low charge events. Three quantities, the χ^2, charge start time, and charge amplitudes, are affected directly by this change. The changes on the charge start time and charge amplitude will propagate to phonon delay time and ionization yield, both are important discriminators against backgrounds. Based on these changes, the cuts `cChiSq_c58R`, `cQThresh_c58R`, `cQin_c58R` [127], `cSingle_c58R` [128], `cNR_c58R` [129], and `cRT_c58R` were created in the reanalysis. Most of these cuts were defined in the same way as in the C58 analysis. Some changes are presented in this section.

The data quality cuts rejecting time periods with environmental interference to the experiment do not need retuning. The rest of the data quality cuts based on statistics of the data, like KS tests and neutralization checks, were not expected to change, because the low charge events are only a small fraction of the data, and the change on them should not change the statistical tests on the entire data sets, which are dominated by high energy events. However, we did retune the neutralization cut on ^{133}Ba data before phonon position correction and found that a few series rejected by the previous neutralization cut in the C58 analysis were accepted by the new cut [130].

5.4.1 Charge χ^2 cut

Figure 5.13 shows the charge χ^2 versus the total ionization energy for the reprocessed ^{133}Ba data. The red dashed line is the functional form for the previous charge χ^2 cut

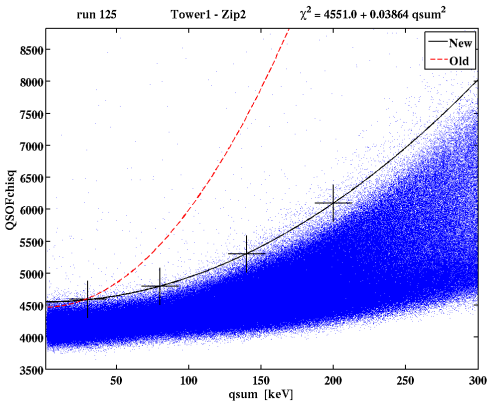


Figure 5.13: Charge χ^2 vs. total ionization energy. The red dashed line is the χ^2 cut in the C58 analysis. The black solid line is the new cut. The crosses are the 3σ edges for the corresponding qsum bins.

in the C58 analysis; the black line is the new one. At low energies, the values of χ^2 almost do not change. As energy goes higher, the χ^2 value difference between the initial processing and the reprocessing goes up quadratically, which agrees well with Eq. (5.14). The improvement to the charge reconstruction algorithm does reduce the χ^2 significantly. This also indicates there are charge start time changes even at high energies, necessitating a new timing analysis.

The functional form of the charge χ^2 cut is an empirical parabola of the form

$$y = a + b\,\mathrm{qsum}^2 \tag{5.26}$$

going through the χ^2 threshold of each energy slice. Let pdf $f(\chi^2)$ be the χ^2 distribution in a qsum slice, the threshold χ^2_t satisfies

$$\frac{f(\chi^2)_{\max}}{f(\chi^2_t)} = \frac{g(\mu;\sigma)}{g(\mu+3\sigma;\sigma)} = 90.017, \tag{5.27}$$

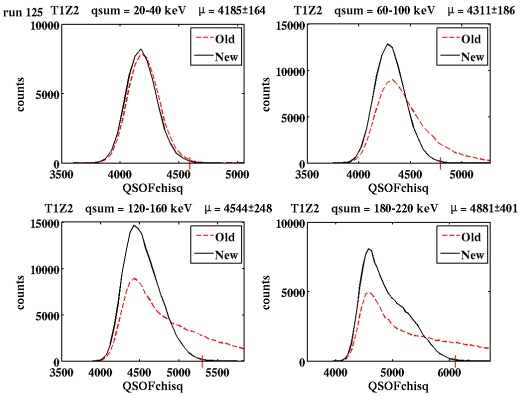


Figure 5.14: Histograms of QSOFchisq for four qsum bins: 20 to 40 keV (top left), 60 to 100 keV (top right), 120 to 160 keV (bottom left), 180 to 220 keV (bottom right). The red dashed line is the value in the C58 analysis and the black line is that after reprocessing. At low energies the distributions are close to Gaussian while at high energies they have a long tail at the upper end. The distribution of QSOFchisq is narrower for the reprocessed data at higher energies, indicating the improvement made to the charge optimal filter.

where g is the Gaussian pdf.

Specifically, four slices of χ^2 in qsum bins 20 to 40 keV, 60 to 100 keV, 120 to 160 keV, and 180 to 220 keV are chosen. The histograms of χ^2 for the four slices are shown in Fig. 5.14. At low energies the distribution of χ^2 is close to Gaussian. The threshold χ^2_t determined by Eq. (5.27) is just at the $3\,\sigma$ upper limit. However, the distributions of χ^2 have a long tail at high energies. In the C58 analysis, χ^2_t was initially chosen as $3\,\sigma$ above the mean for each bin and then was manually adjusted higher for high energy bins, which involved tedious and subjective fine tuning to the cuts. χ^2_t determined by Eq. (5.27) automatically chooses a higher threshold for distributions with a long tail and avoids the fine tuning.

As a side remark, determining χ^2_t based on accumulative probability does not work well, especially for WIMP search data. This is because there are not too many

events in each qsum slice, so enforcing a certain CDF value often includes too many events on the tail and sets the threshold too high.

The efficiency of the charge χ^2 cut is $> 98\,\%$ for most detectors.

5.4.2 Charge threshold cut

The new charge threshold cut was defined in the same way as in the C58 analysis, $4.5\,\sigma$ above the noise floor for each series. However, due to the change in the charge reconstruction algorithm, the noise blob was very different from that in the C58 analysis.

Figure 5.15 shows the noise blob for the C58 analysis (top) and this analysis (bottom). The amplitude maximization used previously biases the amplitude positively and gives a positive mean for the noise traces [91]. The χ^2 minimization does not have a preference on the sign of the amplitude and the biases are in both directions. Zero amplitude is unfavored because the charge templates and the noise traces contain the same noise. A non-zero amplitude would reduce the χ^2 in the fitting. Another feature in the bottom figure is that the donut shaped noise blob has four wings along the x and y axes. It means the χ^2 minimization weighs the higher-amplitude pulse more and tends to settle on it.

5.4.3 Surface event rejection

5.4.3.1 Timing cut definition

As discussed in Sec. 5.3.4.6, the surface event rejection cuts identify and reject the surface events by their faster timing than neutrons. Specifically, the classic timing cut

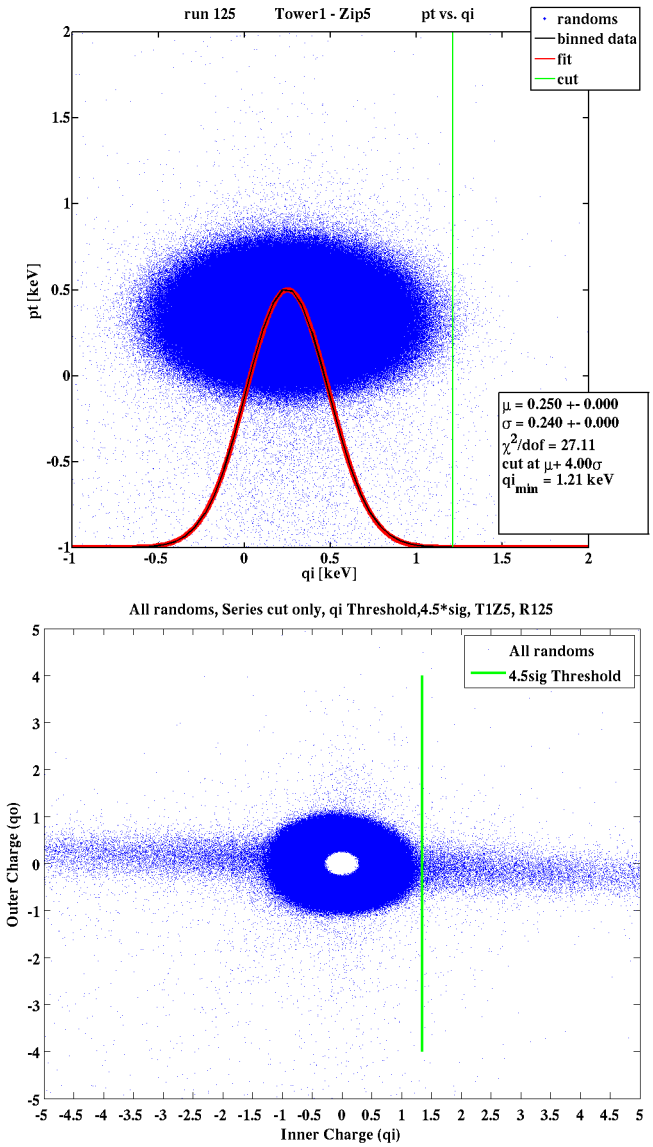


Figure 5.15: Noise blob of T1Z5 with Run 125 data before (top, *Figure from Arrenberg [131].*) and after (bottom, *Figure from Speller [132].*)reprocessing. In the initial processing, the charge optimal filter maximizes the summed amplitude of Qi and Qo. This biases the amplitude positively. In the reprocessing, χ^2 minimization biases the amplitude in both positive and negative directions. Note the y-axis for the two plots are different.

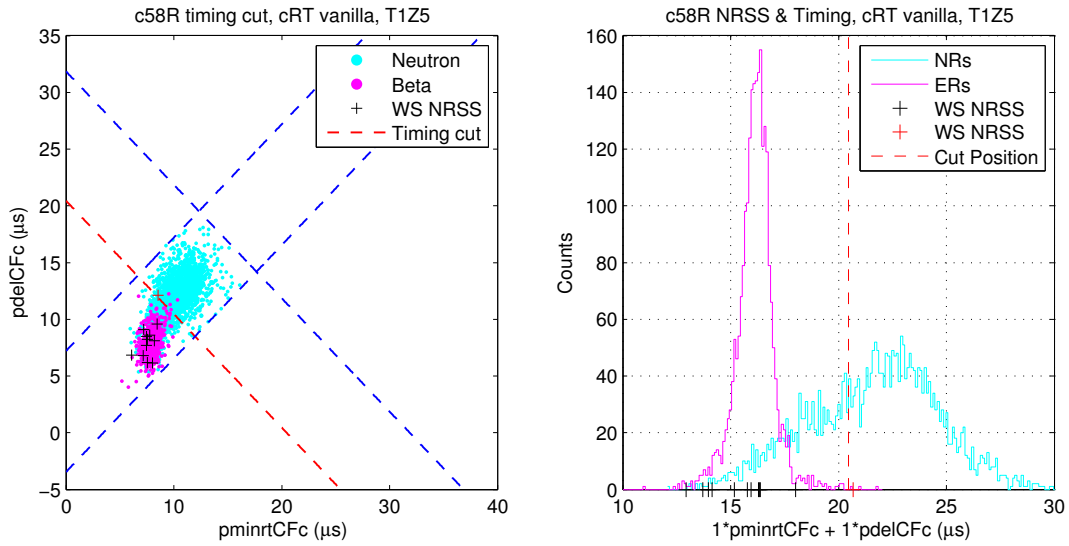


Figure 5.16: Classic timing cut in `pdel` vs. `pminrt` plane (left) and in one-dimensional histogram (right). The red dashed lines indicate where the timing cut is placed. The magenta points and histogram are surface events with shorter timing; and the cyan points and histogram are neutrons with slower timing. The crosses are the nuclear recoil single scatters in the WS data with red crosses passing the timing cut.

uses the sum of pdel and pminrt as the discriminator. The right figure of Fig. 5.16 shows the timing parameter distribution of surface events and neutron samples and the position of the timing cut. Timing cut position can be varied to have lower or higher surface event leakage with the neutron acceptance moving in the opposite direction. The left figure shows the timing cut in the pdel versus pminrt plane. Besides the timing cut which is drawn in red dashed line, there are three consistency cuts shown in blue dashed lines to remove the outliers in timing parameter.

5.4.3.2 Surface event leakage estimation

Once the timing cut position is chosen, the leakage of surface events into the neutron region can be estimated. The signal region, which contains the nuclear recoil single scatters, is masked by the blinding cut. We use two other event samples to estimate the leakage: the multiple scatters in the nuclear recoil band and the events outside but around the nuclear recoil band, which is referred to as the beta band in CDMS. The beta band is defined as ionization yield below the lower value of 0.7 and 5σ below the the electron recoil band mean, and above 0.1, and outside the nuclear recoil band. Two methods are used each with a different event sample to estimate the leakage.

Method 1

Method 1 uses the timing cut pass/fail ratio of the nuclear recoil multiple scatters multiplying the expected number of nuclear recoil singles to estimate the surface event leakage:

$$n = \sum_z N_z \frac{b_z}{B_z}, \tag{5.28}$$

where n is the total surface event leakage, z is detector index, N_z is the expected number of nuclear recoil single scatters failing the timing cut on detector z, b_z and B_z are the number of nuclear recoil multiple scatters passing and failing the timing cut, respectively. The sum is over all interior detectors. Method 1 can not account for the endcap detectors because their exterior face can not tag multiple scatters. Due to the limited statistics, only one bin is used.

Method 2++

Method 2++ uses the timing cut pass/fail ratio of the beta multiples multiplying the expected number of nuclear recoil singles to estimate the leakage:

$$n = \sum_{z,e,f} N_z s_{z,e,f} \frac{m_{z,e,f}}{M_{z,e,f}}, \tag{5.29}$$

where n, z, and N_z are the same as in Method 1, e and f are the energy and face bin indexes, respectively, $s_{z,e,f}$ is the fraction of N_z in Bin (e, f) on Detector z, m and M are the number of beta multiples passing and failing the timing cuts, respectively. To reduce the systematic difference between events in the nuclear recoil band and the beta band, three energy bins 10 to 20 keV, 20 to 30 keV, 30 to 100 keV, and two face bins, p-side and q-side, are used in the calculation. The s factors describe the distribution of N_z in the six bins.

The initial version of Method 2++ only used two face bins and one energy bin, which was called Method 2. Subsequently two energy bins were used to reduce the systematic effects and it was named Method 2+. Method 2++ is Method 2 with three energy bins. Method 1 is expected to have the least errors due to the systematic

Table 5.6: Expected number of nuclear-recoil singles failing the timing cut for the C58R analysis. The errors are simple Poisson errors.

	N(1)	N(2)	N(3)
T1Z2	15.30 ± 4.40	33.00 ± 14.83	11.48 ± 2.65
T1Z5	7.70 ± 2.10	12.89 ± 5.31	22.08 ± 3.90
T2Z3	7.50 ± 2.10	13.42 ± 7.00	17.29 ± 3.98
T2Z5	8.60 ± 2.40	4.40 ± 3.24	19.64 ± 3.83
T3Z2	2.20 ± 0.50	1.85 ± 1.90	12.98 ± 3.35
T3Z4	6.50 ± 1.30	7.31 ± 4.03	12.04 ± 2.42
T3Z5	4.00 ± 0.80	1.72 ± 1.80	4.83 ± 1.53
T3Z6	20.90 ± 4.80	35.44 ± 17.24	95.11 ± 20.95
T4Z2	5.40 ± 1.10	6.53 ± 3.60	12.46 ± 2.74
T4Z4	6.30 ± 1.30	2.49 ± 1.55	7.11 ± 1.84
T4Z5	5.80 ± 1.20	6.49 ± 3.23	12.80 ± 3.03
T4Z6	20.70 ± 4.70	24.14 ± 8.57	80.56 ± 15.07
T5Z4	6.10 ± 1.20	5.45 ± 3.31	18.73 ± 3.66
T5Z5	6.10 ± 1.20	7.31 ± 4.03	9.35 ± 2.34
Sum	123.10 ± 9.43	162.45 ± 27.46	336.47 ± 27.88

difference between single and multiple scattering surface events. But because of the low statistics, Method 2++ is used in the timing cut tuning.

Without unblinding the data, the number of nuclear singles failing the timing cut in the WIMP-search data can be estimated in three different ways:

1. Scaling the observed c34 NRSS counts by the c58:c34 livetime ratio.
2. Multiplying the observed c34 NRSS:NRMS ratio by observed c58 NRMS count.
3. Multiplying the c58 wide beta NRSS:NRMS ratio by observed c58 NRMS count.

and the values are listed in Table 5.6. N(1) is expected to have the smallest systematic error and it was used to calculate the leakage when tuning the timing cuts.

5.4.3.3 Timing cut optimization

Procedure

Now that the timing parameter distributions for the surface events and neutrons as well as the leakage estimation methods are in place, we can set the timing cuts for the detectors. Equal leakage fraction and equal number of leaked events for all detectors under the constraints of pre-chosen total leakage were used in the past to set the timing cuts [133]. Given that there are systematic differences between the detectors, these two methods may not give the best limit on WIMP-nucleon scattering cross section. In the C58 analysis, an optimization technique was used to optimize the SAE with a given total leakage.

The objective function to minimize is

$$f(t_1, t_2, \ldots) = \left(1 - \frac{\sum_z \tau_z(t_z)}{\tau_{\max}}\right)^2 + 10\left(1 - \frac{\sum_z n_z(t_z)}{n_{\text{target}}}\right)^2, \tag{5.30}$$

where t_z is the timing cut position, τ_z is the SAE for Detector z, $\tau_{\max}$ is the maximum achievable total SAE given t_z in a pre-selected range, n_z is the leakage on Detector z, and n_{target} is the desired total leakage. More exposure means less neutrons are cut out. So $\tau_{\max}$ is chosen as the exposure when the timing cuts are placed at the lower bounds of the intersections of the timing parameters of betas and neutrons

$$\tau_{\max} = \sum_z \tau_z(\min(t_z)), \quad \text{given } t_z \in (t_z^\beta \cap t_z^n). \tag{5.31}$$

A series of target leakages in the range between 0.1 and 1 events are chosen to map out the dependence of total SAE on the desired leakage. The factor 10 is chosen to

constrain the tolerances on the leakages of individual detectors in the optimization.

The optimization is done using the MATLAB function `patternsearch` for a 60 GeV WIMP mass. The functions $\tau_z(t_z)$ and $n_z(t_z)$ are both smoothed before fed into the optimization. A neutron consistency cut is also imposed, which selects neutrons between 0.5 % and 99.5 % quantiles of the timing distribution. At the beginning, all detectors are assigned the average leakage. For some high leakage detectors with broad beta timing distributions, even the initial leakage pushed the timing cut all the way to the upper bound of the neutron timing distribution, thus loosing all their exposure. These detectors are excluded from the optimization with the timing cut set at the edge of the neutron distribution. The right figure of Fig. 5.17 shows such a situation for the detector T4Z6. The left figure shows the results for T1Z5, a typical interior detector.

Exposure and leakage

The SAE and 90 % Poisson upper limit on the WIMP rate versus expected WIMP-search nuclear singles leakage are shown in Fig. 5.18. The optimized results in solid lines are globally better than the results based on equal leakage per detector for this analysis (red) and the C58 analysis (magenta). For comparison, the R123-124 result based on equal leakage fraction is also shown. It is worse than the two aforementioned methods.

Comparing the right and left figures in Fig. 5.18, we see that when expected leakage goes above 0.5 events, though the SAE still increases with leakage, the upper limit on WIMP rate stays roughly constant and goes up a little when leakage approaches 1 event. This is because the increase in SAE is balanced by the increase

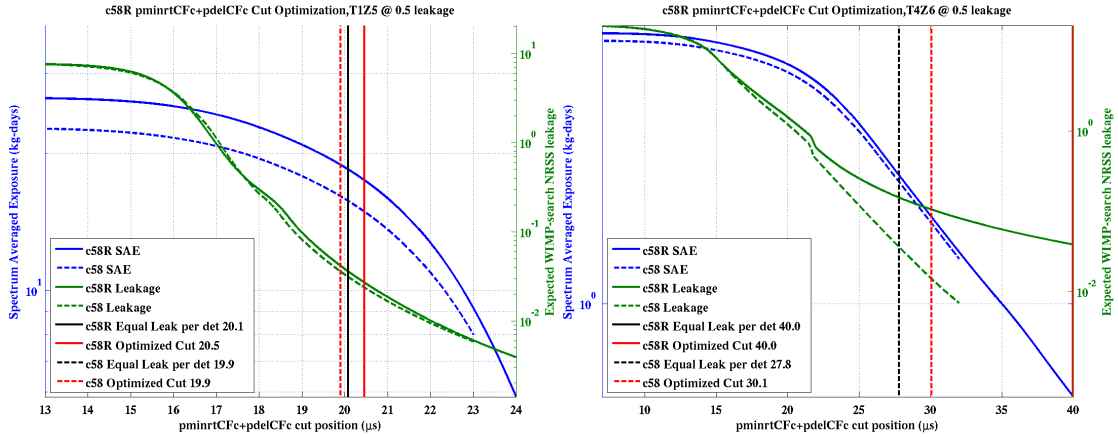


Figure 5.17: SAE (blue, left y-axis) and leakage (green, right y-axis) vs. sum of `pdel` and `pminrt`. The solid lines are for this analysis and the dashed lines are for the C58 analysis. The black vertical lines are the initial positions of the timing cut based on equal leakage. The red vertical lines are the optimized cut positions. The left figure shows a typical interior detector while the right figure shows the high leakage detector T4Z6. Note that the timing cuts for T4Z6 are pushed to the right end because of the high beta tail.

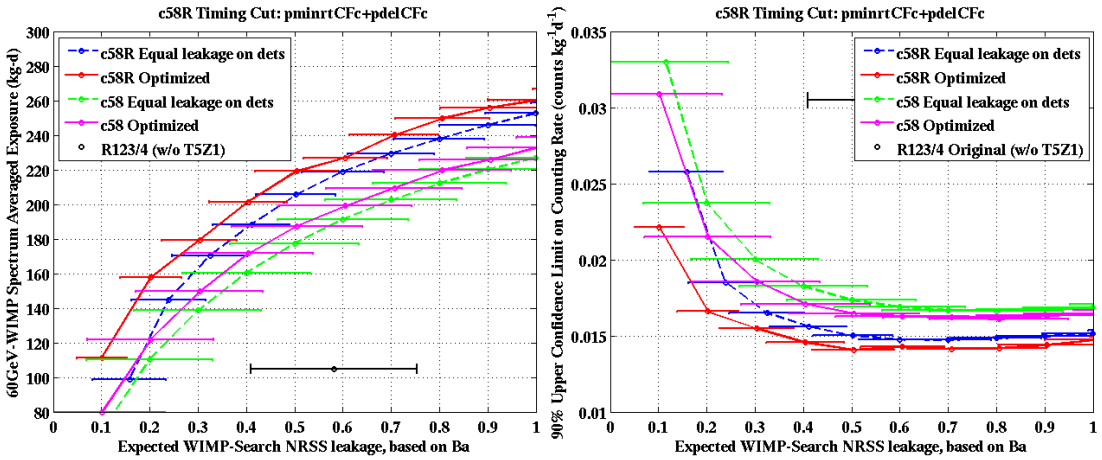


Figure 5.18: Spectrum averaged exposure for a 60 GeV WIMP mass (left) and the 90 % Poisson upper limit on WIMP rate assuming no background (right) vs. expected WIMP-search nuclear recoil singles leakage. The results for this analysis and the C58 analysis are both shown. Dashed lines are calculated with equal leakage on each detector; and solid lines show the results based on global optimization. The R123-124 results based on equal leakage fraction are also shown.

of expected background events at leakage above 0.5 events; and the WIMP sensitivity does not increase anymore. This is a warning sign for the CDMS II experiment. With the existing data, a fraction of the exposure is lost because of increased background. For the longer term, if the same detectors were run longer and the surface events are the only background, the WIMP sensitivity would inversely scale with the square root of running time in this best scenario. Adding the not quite well characterized neuron background and possibly some other unknown backgrounds, the WIMP sensitivity may not increase much with a longer running time. It is clear that the CDMS-II experiment is background limited and lower background detectors are needed to push the WIMP sensitivity to lower cross sections.

From the right figure in Fig. 5.18, the best WIMP sensitivity lies in the leakage range between 0.5 and 0.7 events. We choose 0.5 events to be our target leakage in this reanalysis to be conservative. The same number was also used in the C58 analysis. The corresponding SAE increased from 187.3 kg · day in the C58 analysis to 219.1 kg · day in this analysis, a 17 % increase. Fig. 5.19 shows detector breakdown of this SAE change. There is some amount of increase for most of the detectors except the two endcaps which lost almost all their exposure.

The estimated leakages with different methods on the WIMP-search and ^{133}Ba calibration data are listed in Table 5.7. A best fit of different methods following Yellin's method (see supplemental material of Ref [134]) are also given as well as the χ^2 of the fit and the probability of having higher χ^2 values in the parentheses.

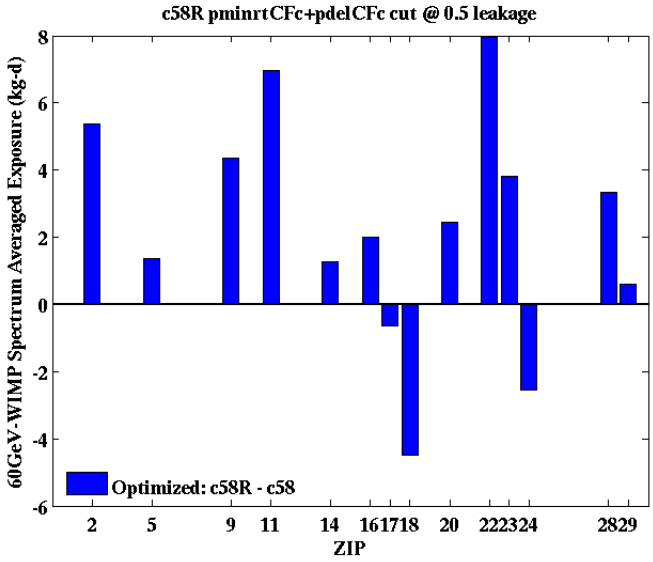


Figure 5.19: Spectrum averaged exposure increase for a 60 GeV WIMP mass relative to the C58 analysis.

Table 5.7: Estimated leakage with the optimized timing cut on WIMP-search and ^{133}Ba calibration data. The best fit value, the χ^2 of the fit, and the probability of having higher χ^2 are also listed. Endcap detectors are not used in Method 1.

Method	N(1)	N(2)
Method 1 (No EC):	1.357 ± 0.741	1.140 ± 1.066
Method 2:	0.129 ± 0.293	0.141 ± 0.448
Method 2+:	0.144 ± 0.075	0.161 ± 0.099
Method 2++:	0.159 ± 1.074	0.180 ± 1.586
Method 2++ Ba:	0.504 ± 0.085	0.563 ± 0.151
Best fit 1, 2++, 2++ Ba:	0.507 ± 0.085	0.565 ± 0.150
χ^2 (prob)	0.762 (0.683)	0.191 (0.909)

Timing cut efficiency

The efficiency of the timing cut were measured on the ^{252}Cf calibration data. For most detectors the efficiencies vary between 50 % and 75 % and decreases with energy. The two endcap detectors T3Z6 and T4Z6 lost most of their exposure and the efficiency is close to zero.

5.4.3.4 Other timing cuts

Besides the "classic" timing parameter, which is the sum of the `pdel` and `pminrt`, two other linear combinations of `pdel` and `pminrt` were also optimized. One of them is `pdel` + 0.75 `pminrt`, and the other one had varying weights for different detectors. Both of them show very close but slightly worse performance than the classic combination.

5.5 WIMP search exposure

5.5.1 Exposure

In the timing cut optimization, Eq. (5.30) maximizes the spectrum averaged exposure defined in Eq. (5.24). The detector mass and livetime are constants in the SAE. We actually optimized the spectrum averaged detector efficiency for a 60 GeV WIMP mass. Although the timing cut was chosen based on the response from a 60 GeV WIMP mass, its efficiency was measured on data and is independent of the WIMP recoil spectrum. Once the detector efficiency is known, the sensitivity of the experiment to WIMPs with different masses can be computed.

Table 5.8: Detector mass in grams. Data from Arrenberg and Bruch [125].

	T1	T2	T3	T4	T5
Z1	230.5	101.4	104.6	101.4	224.5
Z2	227.6	104.6	231.2	238.9	229.5
Z3	219.3	219.3	104.6	101.4	101.4
Z4	104.6	104.6	238.9	234.6	224.5
Z5	219.3	238.9	238.9	231.9	234.8
Z6	104.6	104.6	231.7	238.9	223.6

Table 5.9: WIMP-search exposure for the reanalysis of the Runs 125-128 data by detectors. The unit is $\mathrm{kg} \cdot \mathrm{day}$.

	T1	T2	T3	T4	T5
Z1					
Z2	48.10		17.73	44.45	
Z3		34.68			
Z4			55.28	51.03	47.90
Z5	36.49	43.47	34.90	46.85	49.33
Z6			50.61	51.36	

Table 5.8 lists the masses of all the detectors calculated with their exact dimensions. With all bad periods removed, the total exposure of the Runs 125-128 data in the reanalysis is $612.17\,\mathrm{kg} \cdot \mathrm{day}$, almost the same as $612.13\,\mathrm{kg} \cdot \mathrm{day}$ in the C58 analysis. Table 5.9 shows the exposure for all the Ge detectors used in the reanalysis.

5.5.2 Analysis efficiency

The total cut efficiency for a detector is the product of the efficiencies of the data reconstruction quality cuts, physics cuts, and the surface event rejection cuts. The exposure averaged total cut efficiency is the efficiency for the whole experiment. Fig. 5.20 shows the experiment efficiency with data reconstruction quality cuts,

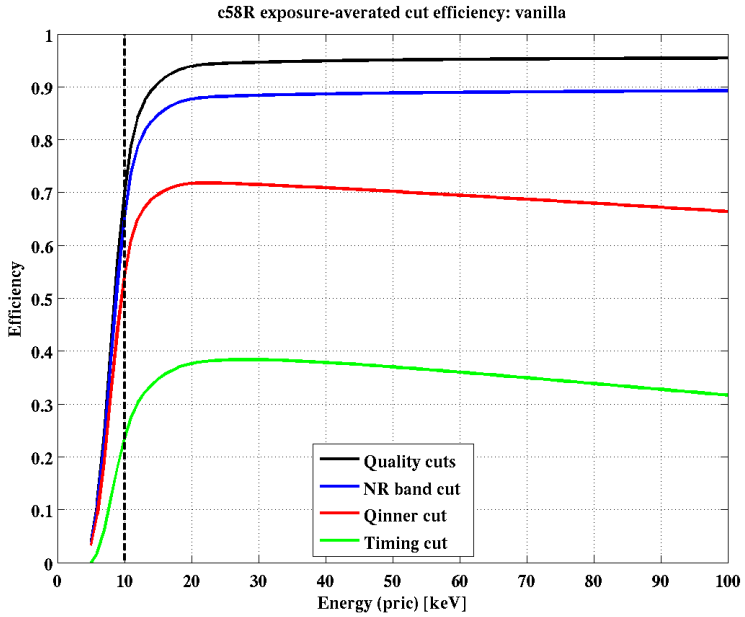


Figure 5.20: Exposure averaged cut efficiency with reconstruction quality cuts, nuclear recoil band cuts, charge fiducial volume cuts, and timing cuts applied accumulatively (from top to bottom).

nuclear recoil band cuts, charge fiducial volume cuts, and surface event rejection cuts applied accumulatively. We can see that most of the exposure loss are due to the charge fiducial cut and the surface event rejection cut. The energy dependence of the total efficiency is introduced by the charge fiducial volume cut. The 60 GeV WIMP spectrum averaged efficiency for the WIMP-search data is 35.6 %. The corresponding SAE is 217.7 kg · day, which is very close to what we got in the timing cut optimization, 219.1 kg · day.

5.6 WIMP search result

With all cuts defined, all efficiencies and exposures calculated, the WIMP signal region in the WIMP-search data is ready to be unmasked. Besides the potential WIMP signals, background events induced by neutrons, gammas, and beta particles can also show up in the signal region. A thorough understanding of the background is critical to interpret the events that could show up in the signal region. Although the background events can be characterized to a certain confidence level, they are not subtracted when constraining the WIMP parameter space. Instead, all events showing up in the signal region are treated as WIMPs and conservative limits on WIMP-nucleon cross section are set.

5.6.1 Backgrounds

Any events other than WIMPs passing all criteria for WIMPs would be indistinguishable from real WIMP signals, and constitute backgrounds to WIMP search. A background event can be an electron recoil or a nuclear recoil. The former is from photons or beta particles scattering off electrons in the target atom and the latter is primarily from neutrons scattering with the target nuclei.

Most of the photons interacting with the detector produce bulk electron recoils. The majority of these events belong in the electron recoil band and are well separated from the nuclear recoil band. By statistical fluctuations, however, some of these events could fall into the nuclear recoil band and be misidentified as WIMPs, forming the gamma background. In particular, both the electron and the nuclear recoil bands flare as recoil energy goes lower, reducing the separation between the two bands and

increasing the probability of gamma leakage. A simulation was performed in the C58 analysis to assess this gamma background and found it was negligible at the analysis threshold of 10 keV [135].

Photons depositing energy close to the detector surface and incident β particles produce near surface electron recoils. This surface event background is the dominant background to CDMS WIMP-search. Nuclear recoil background contributes a non-negotiable amount of events though small compared with the surface events. These two background will be discussed further in the sections below.

5.6.1.1 Nuclear recoil backgrounds

The nuclear recoil background is mainly induced by neutrons, and α particles and recoiling daughter nuclei in spontaneous fission of U/Th decay chains. At the depth of the Soudan site, 713 m (2090 m.w.e.) below the surface, the dominant neutron sources are the cosmogenic neutrons from fast muon spallation and the radiogenic neutrons from spontaneous fission of U/Th decay chains and the $(\alpha,\ n)$ reaction. The cosmogenic neutrons with accompany electromagnetic showers and/or the parent muon(s) intersect the muon veto can be tagged at nearly 100% efficiency. Radiogenic neutrons produced in the cavern walls can not penetrate the passive shield. Both of them can be excluded as background. However, the cosmogenic neutrons not vetoed by the muon veto and the radiogenic neutrons produced in the materials surrounding the detector which are inside the passive shield can reach the ZIPs and register nuclear recoil events. The multiple scatters of these events can be excluded as WIMP signals by their multiplicity, leaving the single scatters as the nuclear recoil background. Due to the limited statistics and mixing with potential WIMP signals, this background can

not be estimated from the data directly and Monte Carlo simulations must be used.

A summary of the nuclear recoil background is presented in the following sections. More details about the simulations and estimation methods can be found in [103, 136–141].

Cosmogenic neutrons

The nuclear recoil background due to the cosmogenic neutrons is estimated as the product of the number of vetoed neutrons in the WIMP-search data and the ratio of the number of unvetoed neutrons to that of vetoed in the Monte Carlo simulation. Both single and multiple scatter neutrons are included to reduce the statistical errors.

Muons were generated using simulation code with the correct spectra and angular distribution for Soudan. These muons were then used as the input of a second GEANT4 simulation, which spread the muons at and propagated them through the 10 m of cavern rock surrounding the CDMS experiment. A complete mock-up of the CDMS-II experiment including shieldings and detector installations was used in the simulation. Muons and their secondaries were all tracked through the experiment.

In the simulated data which corresponds to 69.8 years of exposure, the cosmogenic neutron background for the classic timing cut was found to be $0.21^{+0.014}_{-0.010}$ events [142].

Radiogenic neutrons

Radiogenic neutrons result from the radioactive contaminants in the materials surrounding the detectors, primarily the U/Th decay chains. The contamination levels of the radioactive isotopes were determined by matching the simulated electron recoil ionization energy spectrum with that from the WIMP-search data, both with

appropriate cuts applied. A second neutron simulation was performed throwing neutrons from each shielding element (poly, Pb, Cu) with the neutron spectra from U/Th chains and propagating the neutrons though all materials and detectors. Single and multiple nuclear recoil rates were recorded for each contaminant. The total simulated radiogenic neutron rate of the experiment was then the sum of the neutron rates in the neutron simulation weighted by the contamination levels determined in the gamma spectrum simulation. For the 14 Ge detectors used in the reanalysis, the rate of the single nuclear recoils was found to be $(1.15 \pm 0.14) \times 10^{-4}$ kg^{-1} $\cdot$ day^{-1} in the analysis energy range 10 to 100 keV. Multiplying the exposure 612.17 kg $\cdot$ day and the averaged analysis efficiency for the classic timing cut 35.4 %, the expected radiogenic neutron background was 0.025 ± 0.003 events [142].

α-induced nuclear recoils

Recoiling α particles and the daughter nuclei in spontaneous fission of the U/Th decay chains can also induce nuclear recoils in the detectors. However, α particles typically deposit a few MeV, which is far beyond the analysis threshold. Additionally, the ionization yield of the α induced nuclear recoils falls below the nuclear recoil band. It is very unlikely the α's could be mistaken as WIMPs.

The emitted daughter nuclei in fission of the U/Th decay chains typically have energy $\lesssim$ 100 keV, which is in the analysis energy range. The nuclear recoils induced by them are usually α coincident unless the contaminant is on the outward face of an endcap detector, in which case the α could escape. A typical example is ^{210}Po, a radon daughter. However, the ionizing and penetrating power are greatly suppressed because of their large mass. This kind of background is expected to be

negligible [91, 143].

5.6.1.2 Surface events

There are three main sources for surface events:

1. β particles from β emitters such as ^{210}Pb and ^{40}K on the detector and materials around them.
2. Electrons emitted by surrounding materials through photoelectric effect.
3. Photons which deposit energy close to the surface of the detector.

All these events behave similarly in ionization yield and timing. A Bayesian framework introduced in J. Filippini's thesis [91] was used to estimate the background from the surface events. The detailed procedure and calculation can be found in Appendix C of this thesis. The total surface background was found to be

$$n = 0.55^{+0.17}_{-0.14}. \tag{5.32}$$

5.6.1.3 Combined background

Figure 5.21 shows the WIMP-search backgrounds from the interior, endcap, and all detectors, with both neutron and surface event backgrounds included. The total background to WIMP-search is

$$n = 0.64^{+0.17}_{-0.15}. \tag{5.33}$$

5.6.2 Unbliding WIMP signal region

Two events were observed for the classic timing cut, one for the neural network timing analysis [144, 145], and zero for the 5D χ^2 timing analysis [115]. For the classic

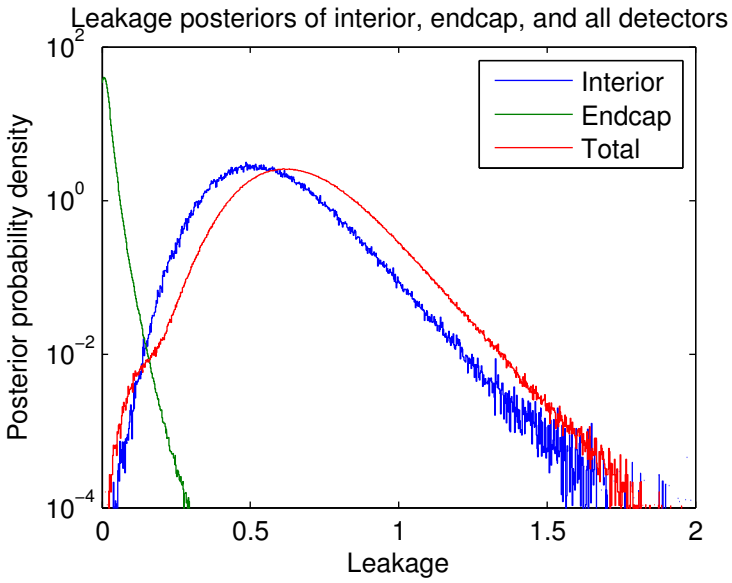


Figure 5.21: WIMP-search background from the interior, endcap, and all detectors. Both neutron and surface event backgrounds are included.

timing analysis, one event showed up on detector T1Z5 and the other on T3Z2.

Figure 5.22 shows these two events in the planes of ionization yield versus recoil energy, normalized ionization yield versus normalized timing parameter, and phonon delay time versus phonon rise time. Both events are close to the recoil energy thresholds and the timing cut lines. Apart from that, they did not show signs of suffering from bad data quality, bad reconstruction quality, or bad event positions in the detector in detailed further checks. Table 5.10 lists the detailed information about the two WIMP candidate events.

With the estimated background of 0.64 events, the probability of observing two or more background events is 13.5 %.

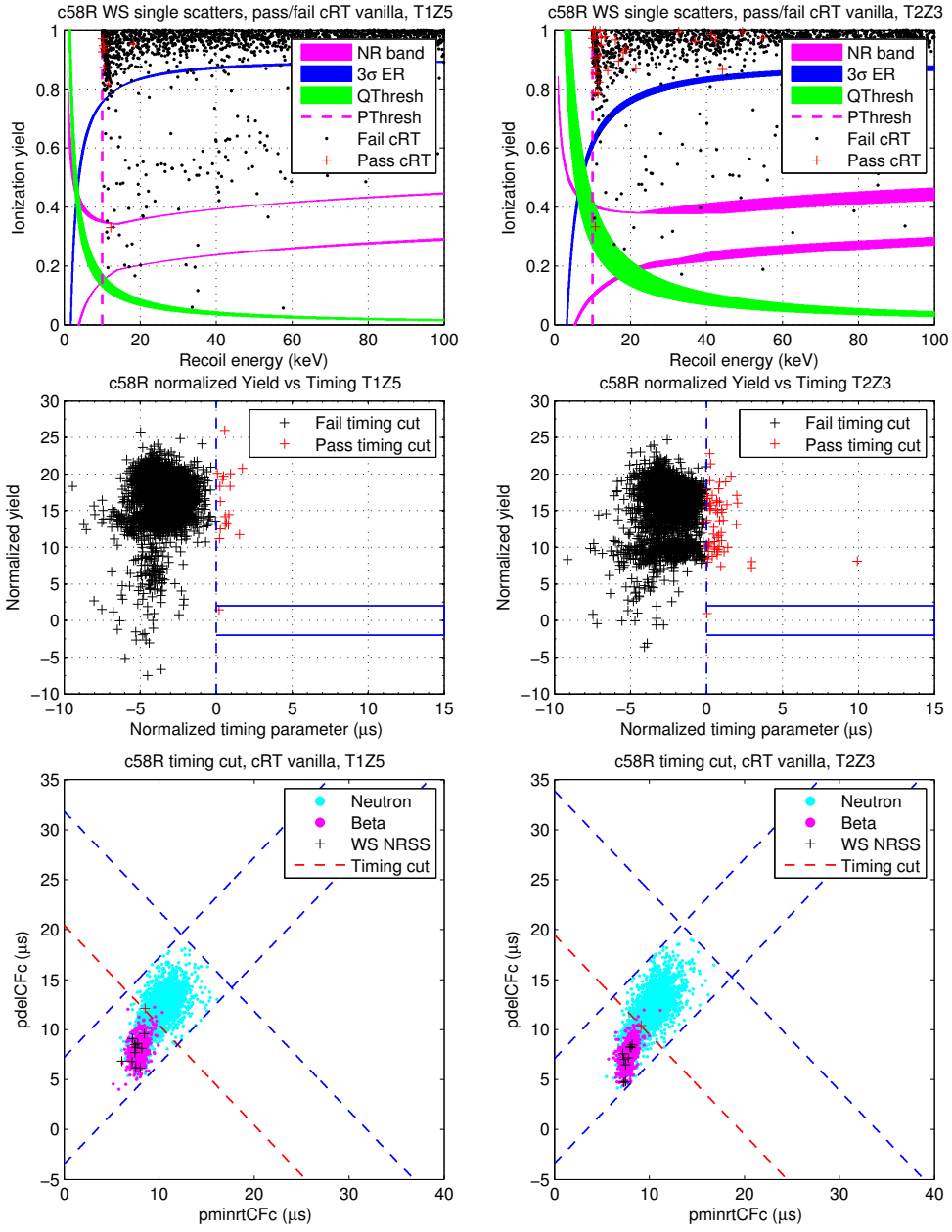


Figure 5.22: WIMP candidate events on T1Z5 and T2Z3 in the C58R classic timing analysis. Left: The WIMP candidate event on T1Z5 in the inoization yield vs. recoil energy (top), normalized yield vs. normalized timing parameter (middle), and phonon delay time vs. phonon rise time (bottom) planes. Right: The WIMP candidate event on T2Z3.

Table 5.10: WIMP candidate events for the C58R classic timing analysis.

Detector	T1Z5	T2Z3
SeriesNumber	1710271127	1805301534
EventNumber	50408	140126
pric	12.30	10.81
yic	0.330	0.332
pminrtCF	5.35	6.40
pdelCFc	12.11	10.39
qsum	4.0	4.0
QSOFchisq	3996.4	4152.1

5.6.3 Constraints on WIMP parameter space

The observed number of WIMP candidate events in this analysis can be explained as background fluctuations with a relatively large probability. A discovery claim can not be made in this case, instead we set upper limits on the WIMP-nucleon cross section. The limits were set conservatively assuming all observed WIMP candidate events are WIMPs, which also avoided the difficulty in background subtraction. Unlike in the timing cut optimization, the optimal interval method [48, 126, 146] was used to calculate the upper limits, which tends to give lower limits than the Poisson method. The standard halo model was assumed with the characteristic velocity $v_0 = 220\,\mathrm{km\,s^{-1}}$, escape velocity $v_{\mathrm{esc}} = 544\,\mathrm{km\,s^{-1}}$, and local dark matter density $\rho_0 = 0.3\,\mathrm{GeV\,cm^{-3}}$. The Helm nuclear form factor was used for the spin-independent interaction. Spectrum averaged exposures at target WIMP masses were used in limit calculation, taking care of the WIMP spectrum and the detector efficiency.

Figure 5.23 compares the 90 % CL upper limits on the spin-independent WIMP-nucleon cross section set by this analysis and the C58 analysis. The black dotted line shows the limit curve set by the C58 analysis with the CDMS II data taken during

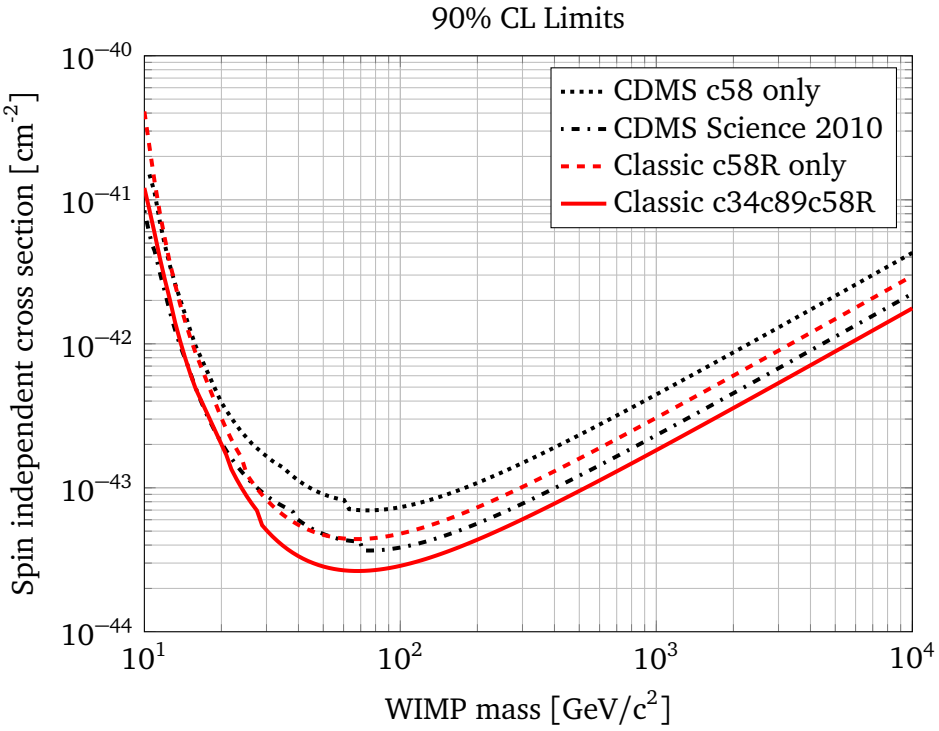


Figure 5.23: 90 % CL upper limits on spin-independent WIMP-nucleon cross section versus WIMP mass. The black dotted line is the limit set by the C58 analysis with only the Runs 125-128 data. The red dashed line is the limit set by this analysis with Runs 125-128 data. The black dash-dotted line and the red solid line are the limits set by the C58 analysis and this analysis with all CDMS II data. The standard halo model is assumed in calculation of the limits.

Runs 125-128, whereas the limit set by this analysis with the same data is shown in the red dashed line. This analysis resulted in a lower limit at WIMP mass $\gtrsim 13\,\mathrm{GeV}$. At the most sensitive WIMP mass of 70 GeV, this analysis gives a WIMP-nucleon cross section $4.4 \times 10^{-44}\mathrm{cm}^2$ compared to $7.0 \times 10^{-44}\mathrm{cm}^2$ given by the previous analysis. This improvement is largely due to the elimination of the candidate event shown in T3Z4 in the C58 analysis with energy $\texttt{pric} = 15.45\,\mathrm{keV}$, though the 17 % increase in exposure also contributes. At WIMP mass $\lesssim 13\,\mathrm{GeV}$, despite the increase in exposure, this analysis shows a higher limit because of the presence of the candidate event in T2Z3 with $\texttt{pric} = 10.81\,\mathrm{keV}$. The two analyses share one common candidate event in T1Z5 with $\texttt{pric} = 12.30\,\mathrm{keV}$, where the two limits are close. The combined limits with the rest of the CDMS II data have similar behavior.

In conclusion, the reprocessing increased the SAE by 17 % for a WIMP with mass 60 GeV. One candidate event of the previous analysis in T3Z4 with $\texttt{pric} = 15.45\,\mathrm{keV}$ disappeared in this analysis; and a new candidate event in T2Z3 with $\texttt{pric} = 10.81\,\mathrm{keV}$ showed up. The new analysis resulted in a lower WIMP-nucleon cross section at WIMP mass $\gtrsim 13\,\mathrm{GeV}$, but the limit is worsened for low WIMP masses.

Chapter 6

Conclusions and Future Work

6.1 UMN R&D facility

We have developed a new CDMS R&D facility at the UMN campus starting from scratch in a little over a year. An Oxford Instruments Kelvin 100 dilution fridge was modified and integrated with the CDMS cold hardware and electronics. Since the first run, we have characterized a subset of the SuperCDMS Soudan iZIPs that are currently running at SUL and measured the charge collection efficiency of the 100 mm diameter Ge crystals which will be used to develop dark matter detectors for the SuperCDMS SNOLAB experiment. The test facility is also used as a test bed for the newly designed Detector Control and Readout Cards, the new DAQ software, and potentially the future development of the CDMS data processing package `CDMSBats`, all aiming for the future generations of CDMS including SuperCDMS SNOLAB and possibly beyond. Besides the major projects, efforts were also made improving the existing CDMS charge readout circuit with the intention to simplify the SuperCDMS

SNOLAB tower design without losing performance.

The test facility has been working reasonably well and being upgraded from time to time according to the requirements of the projects under investigation. For the near future of testing the SuperCDMS SNOLAB detectors, the current configuration provides enough space to host multiple detectors in a single run. The number of detectors will need be reduced if internal gamma or beta sources are used. However, the detector volume is not high enough to take the whole detector stack as it will be installed in the SuperCDMS SNOLAB cryostat. It is likely that the larger cryostats at UC Berkeley, University of Florida, and Queen's University will come online by the time the whole tower testing is needed. The lack of whole tower testing capability in the facility does not significantly limit its science reach. The capability, if desired, can be achieved with a taller fridge Dewar or a taller top plate.

The preliminary version of the SuperCDMS SNOLAB detectors (Fig. 6.1) currently being tested has 4 charge channels and 12 phonon channels on one detector. The charge channels are read out by the JFET based preamplifers at the moment, each dumps 5 mW heating power to the 4 K stage of the fridge. The JFET based preamplifer will be replaced by a version based on HEMT which only generates 100 μW heating power each. This will reduce the heat load of the cold electronics vastly. The added wire traces for the increased number of channels from room temperature to the 4 K stage of the fridge will introduce more heat load, which, however, can be handled with carefully choosing the material, geometry, and heat sinking of the cable. All considered, the K100 fridge will have sufficient cooling power for the testing of the large detectors.

The target T_c of the phonon sensors of the SuperCDMS SNOLAB detectors is around

70 mK and the detectors are intended to be running at a substrate temperature of 50 mK. In comparison, the target T_c of the SuperCDMS Soudan iZIP's is around 90 mK. The lower detector running temperature presents a mild challenge to the K100 fridge whose base temperature has been around 70 mK for the most part. A lower base temperature can be achieved with moderate optimization to the fridge. The addition of the 600 mK thermal link between the still and the 600 mK stage of the tower lowered the base temperature of the detector substantially (from $\sim$ 70 mK to $\sim$ 30 mK) in the past, which shows that a rather fraction of the heat load on the mixing chamber is conducted from the 4 K stage of the tower where the JFET's are installed. The low thermal conductivity thin graphite tubes isolating the tower stages do not actually provide sufficient thermal isolation as intended. The elevated base temperature since the addition of the 600 mK thermal link could be due to degradation of the thermal isolation of the tower or a touch in the fridge which reduces the effectiveness of the 600 mK thermal link. The real culprit can be identified with a series of test runs of the fridge with parts removed incrementally. Nonetheless, introduction of the SuperCDMS SNOLAB tower and HEMT based charge readouts may solve the problem easily or at least help to do so because of improved thermal isolation and decreased heating power directly dumped to the tower.

The current version of the SuperCDMS SNOLAB detector, iZIP6, is specially designed to be used with the CDMS II phonon readouts whose nominal phonon sensor operating resistance is 200 mΩ. The new SQUID array for SuperCDMS SNOLAB allows an operating resistance of 50 mΩ. It is foreseeable that the entire CDMS-II cold hardware and electronics assembly will be replaced by the version for SuperCDMS SNOLAB in the test facility. In addition, the cabling, EBox, and DCRC's will be

upgraded accordingly. This will be a major change to the current configuration of the K100 fridge, which will require redesign of most if not all the radiation shields and thermal connections. The structure of the IVC may possibly need be modified as well.

The changes in detector design and electronics will eventually necessitate a whole new data chain in the test facility. Currently the detector control and DAQ software are both written in Labview and talking directly to the DCRC's. The SuperCDMS SNOLAB DAQ software is MIDAS based and currently under development with the effort led by the UBC group. The new detector control software is possibly still written in Labview though. The new detector design will require a new data format and new data processing algorithms. The future development of `CDMSBats` can theoretically be based on mockup data sets. However, the data from real detectors may be helpful to the developers in identifying potential issues that may show up in the production data. The data analysis pipeline is also going to change with the new data format.

6.2 SuperCDMS SNOLAB detectors

We have measured the ionization energy collection efficiencies of two 100 mm diameter and 33 mm thick Ge test devices. The measured efficiencies are consistent with the earlier measurements performed with smaller Ge crystals. The measurements have demonstrated that such 100 mm diameter Ge crystals can be used to develop dark matter detectors for the SuperCDMS SNOLAB experiment.

SuperCDMS Soudan has demonstrated that the iZIP technology has sufficient surface event rejection power to ensure less than 1 background event in the SuperCDMS SNOLAB experiment. The development of the SuperCDMS SNOLAB iZIP's is

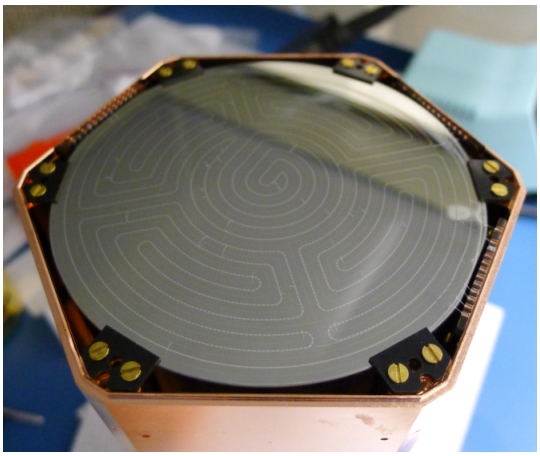


Figure 6.1: Image of an iZIP6 in detector housing. This is a preliminary version of the SuperCDMS SNOLAB detector. The visible lines on the top surface of the substrate are the phonon rails. There are six phonon channels on each face of the detector: one outer ring, one inner disk, and the rest four sandwiched between them. Each face also has two charge electrodes: the inner disk and the outer concentric ring. One detector has four charge channels and twelve phonon channels in total.

well underway. Because the SuperCDMS SNOLAB electronics are still in the R&D phase, a preliminary version of iZIP, iZIP6, is specially created to work with the CDMS-II electronics. Multiple iZIP6 detector are being tested at the UMN facility at the moment. In the meantime, another version of iZIP designed to the SuperCDMS SNOLAB specifications, iZIP7, has also been developed. The test of iZIP7 and the new SQUID array is currently carried out at MIT.

The background rejection capability of the new iZIP's has yet to be studied before the full scale implementation of SuperCDMS SNOLAB. Because of the superseded surface event rejection power of the iZIP technology, the study will need be carried out at a low neutron-background environment at SUL or SNOLAB. The SuperCDMS SNOLAB experiment will shed new light on the WIMP hypothesis in the coming future, especially the low-mass region where CDMS detectors have the unique advantage of

low energy threshold and relatively high WIMP-nucleus cross section.

6.3 Classic timing analysis

The classic timing analysis of the reprocessed c58 data yielded a new WIMP-nucleon cross section of $4.4 \times 10^{-44}\text{cm}^2$ for a WIMP mass of 70 GeV under the standard halo model, which is a factor of 1.6 improvement over the original c58 analysis. The same number of two events showed up in the WIMP signal region as the original c58 analysis. The estimated background was 0.64 events, giving the probability of observing two or more background events 13.5 %. One WIMP candidate event in the original c58 analysis remained in the signal region while the other one with higher recoil energy disappeared and a new event with the lowest energy of the three showed up. Despite a 17 % increase in spectrum averaged exposure after timing cut optimization, the new analysis yield a higher WIMP-nucleon cross section at WIMP mass $\lesssim 13$ GeV because of the appearance of the new event in the low energy region.

Overall, the c58R classic timing analysis did not probe new WIMP parameter space that has not been excluded by other results. The timing cut analysis showed its effectiveness in rejecting surface events, however, in the meantime, also showed its limitation which in turn more or less declared the end of the CDMS-II technology. As indicated by the original c58 analysis as well as this analysis, surface events are the dominant background and are starting to limit the sensitivity of CDMS II. The SuperCDMS Soudan iZIP's currently running are instrumented with charge electrodes and phonon sensors on both faces of the detectors. The partition of ionization energy between the two faces provides precise depth information of an event, giving a

misidentification rate of surface events $< 1.7 \times 10^{-5}$ with ionization alone. The detector performance is sufficient to ensure less than 1 background event in the SuperCDMS SNOLAB experiment. Partition of phonon energy between the two faces provides less stringent but complimentary surface event misidentification rate 4.5×10^{-4}. Theoretically the timing information could also incorporated into the data analysis in a similar fashion as what was done in the c58 analysis, however, the powerful discrimination provided by the Z charge partition makes the timing analysis less appealing or unnecessary to some extent. It is possible that timing analysis will not play an important role in future analyses any more after its decade long development.

Bibliography

[1] E. Kolb and M. Turner, *The Early Universe (Frontiers in Physics)* (Westview Press, 1994).

[2] P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, *et al.* (The Planck Collaboration), (2013), arXiv:1303.5076 .

[3] K. Begeman, A. Broeils, and R. Sanders, Mon. Not. R. Astron. Soc. **249**, 523 (1991).

[4] M. Persic, P. Salucci, and F. Stel, Mon. Not. R. Astron. Soc. **281**, 27 (1996), arXiv:astro-ph/9506004 .

[5] Y. Sofue and V. Rubin, Annu. Rev. Astron. Astrophys. **39**, 137 (2001).

[6] A. K. Drukier, K. Freese, and D. N. Spergel, Phys. Rev. D **33**, 3495 (1986).

[7] G. Jungman, M. Kamionkowski, and K. Griest, Phys. Rep. **267**, 195 (1996).

[8] L. E. Strigari, J. S. Bullock, M. Kaplinghat, A. V. Kravtsov, O. Y. Gnedin, K. Abazajian, and A. A. Klypin, Astrophys. J. **652**, 306 (2006).

[9] J. D. Simon and M. Geha, Astrophys. J. **670**, 313 (2007).

[10] L. E. Strigari, J. S. Bullock, M. Kaplinghat, J. D. Simon, M. Geha, B. Willman, and M. G. Walker, Nature **454**, 1096 (2008).

[11] R. Johnson, D. Chakrabarty, E. O'Sullivan, and S. Raychaudhury, Astrophys. J. **706**, 980 (2009).

[12] Y. Schuberth, T. Richtler, and M. Hilker, in *Globular Clusters - Guides to Galaxies*, Eso Astrophysics Symposia, edited by T. Richtler and S. Larsen (Springer Berlin Heidelberg, 2009) pp. 445–448.

[13] Y. Schuberth, T. Richtler, M. Hilker, R. Salinas, B. Dirsch, and S. S. Larsen, Astron. Astrophys. **544**, A115 (2012).

[14] A. Vikhlinin, A. Kravtsov, W. Forman, C. Jones, M. Markevitch, S. S. Murray, and L. Van Speybroeck, Astrophys. J. **640**, 691 (2006).

[15] D. Clowe, M. Bradač, A. H. Gonzalez, M. Markevitch, S. W. Randall, C. Jones, and D. Zaritsky, Astrophys. J. Lett. **648**, L109 (2006).

[16] M. Bradač, S. W. Allen, T. Treu, H. Ebeling, R. Massey, R. G. Morris, A. von der Linden, and D. Applegate, Astrophys. J. **687**, 959 (2008).

[17] Chandra X-ray Observatory, "Abell 1689: A galaxy cluster makes its mark," (2008), http://chandra.harvard.edu/photo/2008/a1689/.

[18] J. Beringer *et al.* (Particle Data Group), Phys. Rev. D **86**, 010001 (2012).

[19] R. Esmailzadeh, G. D. Starkman, and S. Dimopoulos, Astrophys. J. **378**, 504 (1991).

[20] D. J. Fixsen, Astrophys. J. **707**, 916 (2009).

[21] NASA, "Lambda – data products," (2014), http://lambda.gsfc.nasa.gov/product/.

[22] S. Cole, W. J. Percival, J. A. Peacock, P. Norberg, C. M. Baugh, C. S. Frenk, I. Baldry, J. Bland-Hawthorn, T. Bridges, R. Cannon, *et al.*, Mon. Not. R. Astron. Soc. **362**, 505 (2005), arXiv:astro-ph/0501174 .

[23] W. J. Percival, R. C. Nichol, D. J. Eisenstein, J. A. Frieman, M. Fukugita, J. Loveday, A. C. Pope, D. P. Schneider, A. S. Szalay, M. Tegmark, *et al.*, Astrophys. J. **657**, 645 (2007).

[24] A. A. Klypin, S. Trujillo-Gomez, and J. Primack, Astrophys. J. **740**, 102 (2011).

[25] F. Prada, A. A. Klypin, A. J. Cuesta, J. E. Betancort-Rijo, and J. Primack, (2011), arXiv:1104.5130 .

[26] J. Binney and S. Tremaine, *Galactic Dynamics: (Second Edition) (Princeton Series in Astrophysics)*, 2nd ed. (Princeton University Press, 2008).

[27] K. Freese, J. Frieman, and A. Gould, Phys. Rev. D **37**, 3388 (1988).

[28] J. Lewin and P. Smith, Astropart. Phys. **6**, 87 (1996).

[29] C. Savage, K. Freese, and P. Gondolo, Phys. Rev. D **74**, 043531 (2006).

[30] M. C. Smith, G. R. Ruchti, A. Helmi, R. F. G. Wyse, J. P. Fulbright, K. C. Freeman, J. F. Navarro, G. M. Seabroke, M. Steinmetz, M. Williams, *et al.*, Mon. Not. R. Astron. Soc. **379**, 755 (2007).

[31] S. K. Vempati, (2012), arXiv:1201.0334 .

[32] H. Murayama, (2007), arXiv:hep-ph/0002232 .

[33] H. Baer and J. List, Phys. Rev. D **88**, 055004 (2013).

[34] I. Irastorza, F. Avignone, S. Caspi, J. Carmona, T. Dafni, M. Davenport, A. Dudarev, G. Fanourakis, E. Ferrer-Ribas, J. Galán, *et al.*, J. Cosmol. Astropart. Phys. **2011**, 013 (2011).

[35] D. Moore, *A Search for Low-Mass Dark Matter with the Cryogenic Dark Matter Search and the Development of Highly Multiplexed Phonon-Mediated Particle Detectors*, Ph.d. thesis, California Institute of Technology (2012).

[36] D. Smith and N. Weiner, Phys. Rev. D **64**, 043502 (2001).

[37] S. Chang, G. D. Kribs, D. Tucker-Smith, and N. Weiner, Phys. Rev. D **79**, 043513 (2009).

[38] C. McCabe, Phys. Rev. D **82**, 023530 (2010).

[39] Z. Ahmed, D. S. Akerib, S. Arrenberg, C. N. Bailey, D. Balakishiyeva, L. Baudis, D. A. Bauer, P. L. Brink, T. Bruch, R. Bunker, *et al.* (CDMS Collaboration), Phys. Rev. Lett. **106**, 131302 (2011).

[40] L. Baudis, G. Kessler, P. Klos, R. F. Lang, J. Menéndez, S. Reichard, and A. Schwenk, Phys. Rev. D **88**, 115014 (2013).

[41] N. Bozorgnia, J. Herrero-Garcia, T. Schwetz, and J. Zupan, J. Cosmol. Astropart. Phys. **2013**, 049 (2013).

[42] C. Savage, G. Gelmini, P. Gondolo, and K. Freese, J. Cosmol. Astropart. Phys. **2009**, 010 (2009).

[43] G. Bertone, D. Hooper, and J. Silk, Phys. Rep. **405**, 279 (2005).

[44] C.-L. Shan, J. Cosmol. Astropart. Phys. **2011**, 005 (2011).

[45] D. Tovey, R. Gaitskell, P. Gondolo, Y. Ramachers, and L. Roszkowski, Phys. Lett. B **488**, 17 (2000).

[46] G. J. Feldman and R. D. Cousins, Phys. Rev. D **57**, 3873 (1998).

[47] J. Sander, *Results from the Crogenic Dark Matter Search Using a Chi Squared Analysis*, Ph.d. thesis, University of California, Santa Barbara (2007).

[48] S. Yellin, Phys. Rev. D **66**, 032005 (2002).

[49] V. L. Kashyap, D. A. van Dyk, A. Connors, P. Freeman, A. Siemiginowska, J. X u, and A. Zezas, (2010), arXiv:1006.4334 .

[50] G. Punzi, (2003), arXiv:physics/0308063 .

[51] C. Weniger, J. Cosmol. Astropart. Phys. **2012**, 007 (2012).

[52] J. Aleksć *et al.*, J. Cosmol. Astropart. Phys. **2014**, 008 (2014).

[53] M. Ackermann, M. Ajello, W. B. Atwood, L. Baldini, G. Barbiellini, D. Bastieri, K. Bechtol, R. Bellazzini, R. D. Blandford, E. D. Bloom, *et al.* (Fermi-LAT Collaboration), Astrophys. J. **761**, 91 (2012).

[54] M. Ackermann, M. Ajello, A. Albert, A. Allafort, L. Baldini, G. Barbiellini, D. Bastieri, K. Bechtol, R. Bellazzini, E. Bissaldi, *et al.* (Fermi-LAT Collaboration), Phys. Rev. D **88**, 082002 (2013).

[55] M. Ackermann, A. Albert, B. Anderson, L. Baldini, J. Ballet, G. Barbiellini, D. Bastieri, K. Bechtol, R. Bellazzini, E. Bissaldi, *et al.* (Fermi-LAT Collaboration), Phys. Rev. D **89**, 042001 (2014).

[56] T. Arlen, T. Aune, M. Beilicke, W. Benbow, A. Bouvier, J. H. Buckley, V. Bugaev, K. Byrum, A. Cannon, A. Cesarini, *et al.*, Astrophys. J. **757**, 123 (2012).

[57] A. Abramowski, F. Acero, F. Aharonian, A. G. Akhperjanian, G. Anton, S. Balenderan, A. Balzer, A. Barnacka, Y. Becherini, J. Becker Tjus, *et al.* (H.E.S.S. Collaboration), Phys. Rev. Lett. **110**, 041301 (2013).

[58] M. G. Aartsen *et al.* (IceCube Collaboration), Phys. Rev. D **88**, 122001 (2013).

[59] M. G. Aartsen *et al.* (IceCube Collaboration), Phys. Rev. Lett. **110**, 131302 (2013).

[60] P. Meade, M. Papucci, A. Strumia, and T. Volansky, Nucl. Phys. B **831**, 178 (2010).

[61] O. Adriani, G. C. Barbarino, G. A. Bazilevskaya, R. Bellotti, A. Bianco, M. Boezio, E. A. Bogomolov, M. Bongi, V. Bonvicini, S. Bottai, *et al.*, Phys. Rev. Lett. **111**, 081102 (2013).

[62] M. Cirelli and G. Giesen, J. Cosmol. Astropart. Phys. **2013**, 015 (2013).

[63] G. Aad *et al.* (ATLAS Collaboration), Phys. Rev. Lett. **112**, 041802 (2014).

[64] S. Chatrchyan *et al.* (CMS Collaboration), J. High Energ. Phys. **2012**, 1 (2012).

[65] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Rev. Lett. **108**, 261803 (2012).

[66] G. Aad *et al.* (ATLAS Collaboration), J. High Energ. Phys. **2013**, 1 (2013).

[67] G. Aad *et al.* (ATLAS Collaboration), Phys. Rev. Lett. **110**, 011802 (2013).

[68] J. Goodman, M. Ibe, A. Rajaraman, W. Shepherd, T. M. P. Tait, and H.-B. Yu, Phys. Rev. D **82**, 116010 (2010).

[69] L. M. Carpenter, A. Nelson, C. Shimmin, T. M. P. Tait, and D. Whiteson, Phys. Rev. D **87**, 074005 (2013).

[70] Y. Bai and T. M. Tait, Phys. Lett. B **723**, 384 (2013).

[71] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **108**, 211804 (2012).

[72] R. Bernabei, P. Belli, F. Cappella, V. Caracciolo, S. Castellano, R. Cerulli, C. Dai, A. d'Angelo, S. d'Angelo, A. Marco, *et al.* (DAMA/LIBRA Collaboration), EPJC **73**, 1 (2013).

[73] R. Bernabei, P. Belli, F. Montecchia, W. D. Nicolantonio, A. Incicchitti, D. Prosperi, C. Bacci, C. Dai, L. Ding, H. Kuang, *et al.*, Phys. Lett. B **424**, 195 (1998).

[74] C. E. Aalseth, P. S. Barbeau, N. S. Bowden, B. Cabrera-Palmer, J. Colaresi, J. I. Collar, S. Dazeley, P. de Lurgio, J. E. Fast, N. Fields, *et al.* (CoGeNT Collaboration), Phys. Rev. Lett. **106**, 131301 (2011).

[75] C. E. Aalseth, P. S. Barbeau, J. Colaresi, J. I. Collar, J. D. Leon, J. E. Fast, N. E. Fields, T. W. Hossbach, A. Knecht, M. S. Kos, *et al.* (CoGeNT Collaboration), (2014), arXiv:1401.3295 .

[76] A. Broniatowski, X. Defay, E. Armengaud, L. Bergé, A. Benoit, O. Besida, J. Blümer, A. Chantelauze, M. Chapellier, G. Chardin, *et al.*, Phys. Lett. B **681**, 305 (2009).

[77] E. Armengaud, C. Augier, A. Benoît, L. Bergé, J. Blümer, A. Broniatowski, V. Brudanin, B. Censier, G. Chardin, M. Chapellier, *et al.*, Phys. Lett. B **702**, 329 (2011).

[78] E. Armengaud *et al.* (EDELWEISS Collaboration), Phys. Rev. D **86**, 051701 (2012).

[79] V. Y. Kozlov (EDELWEISS collaboration), (2013), arXiv:1305.2808 .

[80] G. Angloher, M. Bauer, I. Bavykina, A. Bento, C. Bucci, C. Ciemniak, G. Deuter, F. Feilitzsch, D. Hauff, P. Huff, *et al.*, EPJC **72**, 1 (2012).

[81] M. Boulay and A. Hime, Astropart. Phys. **25**, 179 (2006).

[82] D. S. Akerib, H. M. Araujo, X. Bai, A. J. Bailey, J. Balajthy, S. Bedikian, E. Bernard, A. Bernstein, A. Bolozdynya, A. Bradley, *et al.* (LUX Collaboration), (2013), doi:10.1103/PhysRevLett.112.091303, arXiv:1310.8214 .

[83] C. Galbiati, "Cf1 direct dark matter detection summary," (2013), snowmass on the Mississippi.

[84] T. Shutt, *A Dark Matter Detector Based on the Simultaneous Measurement of Phonons and Ionization at 20 mK*, Ph.d. thesis, University of California, Berkeley (1993).

[85] S. Golwala, *Exclusion Limits on WIMP-Nucleon Elastic Scattering Cross-Section from the Cryogenic Dark Matter Search*, Ph.d. thesis, University of California, Berkeley (2000).

[86] V. Mandic, *First Results from the Cryogenic Dark Matter Search Experiment at the Deep Site*, Ph.d. thesis, University of California, Berkeley (2004).

[87] Z. Ahmed, *A Dark-Matter Search Using the Final CDMS II Datasetand a Novel Detector of Surface Radiocontamination*, Ph.d. thesis, California Institute of Technology (2012).

[88] S. Hertel, *Advancing the Search for Dark Matter: from CDMS II to SuperCDMS*, Ph.d. thesis, Massachusetts Institute of Technology (2012).

[89] K. McCarthy, *Detector Simulation and WIMP Search Analysis for the Cryogenic Dark Matter Search Experiment*, Ph.d. thesis, Massachusetts Institute of Technology (2012).

[90] M. Pyle, *Optimizing the Design and Analysis of Cryogenic Semiconductor Dark Matter Detectors forMaximum Sensitivity*, Ph.d. thesis, Stanford University (2012).

[91] J. Filippini, *A Search for WIMP Dark Matter Using the First Five-Tower Run of the Cryogenic Dark Matter Search*, Ph.d. thesis, University of California, Berkeley (2008).

[92] K. Sundqvist, *Carrier Transport and Related Effects in Detectors of the Cryogenic Dark Matter Search*, Ph.d. thesis, University of California, Berkeley (2012).

[93] Z. He, Nucl. Instr. Meth. Phys. Res. A **463**, 250 (2001).

[94] G. Wang, J. Appl. Phys. **107**, 094504 (2010).

[95] R. Ogburn, *A search for particle dark matter using cryogenic germanium and silicon detectors in the one- and two-towerruns of CDMS-II at Soudan*, Ph.d. thesis, Stanford University (2008).

[96] T. Saab, *Search for Weakly Interacting Massive Particles with the Cryogenic Dark Matter Search Experiment*, Ph.d. thesis, Stanford University (2002).

[97] S. Nam, *Development of Phonon-Mediated Cryogenic Particle Detectors with Electron and Nuclear Recoil Discrimination*, Ph.d. thesis, Stanford University (1998).

[98] Z. Ahmed *et al.* (CDMS-II Collaboration), Science **327**, 1619 (2010), arXiv:0912.3592 .

[99] A. D. Silva, *Development of a Low Background Environment for the Cryogenic Dark Matter Search*, Ph.d. thesis, University British Columbia (1996).

[100] D. S. Akerib, M. S. Armel-Funkhouser, M. J. Attisha, C. N. Bailey, L. Baudis, D. A. Bauer, P. L. Brink, R. Bunker, B. Cabrera, D. O. Caldwell, *et al.* (CDMS Collaboration), Phys. Rev. D **72**, 052009 (2005).

[101] B. Rossi, Rev. Mod. Phys. **20**, 537 (1948).

[102] D.-M. Mei and A. Hime, Phys. Rev. D **73**, 053004 (2006).

[103] M. Fritts, *Background Characterization and Discrimination in the Final Analysis of the CDMS II Phase of the Cryogenic Dark Matter Search*, Ph.d. thesis, University of Minnesota (2011).

[104] R. Radpour, *Design, Construction, and Assessment of a Neutron Shield for CDMS Test Facilities*, Master's thesis, University of Minnesota (2011).

[105] F. Pobell, *Matter and Methods at Low Temperatures*, 3rd ed. (Springer, 2007).

[106] D. S. Betts, *An Introduction to Millikelvin Technology (Cambridge Studies in Low Temperature Physics)* (Cambridge University Press, 2005).

[107] G. K. White and P. Meeson, *Experimental Techniques in Low-Temperature Physics (Monographs on the Physics and Chemistry of Materials, 59)*, 4th ed. (Oxford University Press, USA, 2002).

[108] D. Akerib, "Thermal model for graphite towr supports," (1993), http://cdms.berkeley.edu/cdms_restricted/coldhardware/Tower/Thermal/GraphiteThermalTo

[109] B. Young, CDMS ELog **Detector history**, 1342 (2011).

[110] B. Shank, CDMS ELog **Detector history**, 1356 (2011).

[111] C. Bailey, *The Cryogenic Dark Matter Search: First 5-Tower Data and Improved Understanding of Ionization Collection*, Ph.d. thesis, Case Western Reserve University (2009).

[112] A. Broniatowski, J. Low Temp. Phys. **167**, 1069 (2012).

[113] B. Cabrera, M. Pyle, R. Moffatt, K. Sundqvist, and B. Sadoulet, (2010), arXiv:1004.1233 .

[114] L. Hsu, CDMS Analysis Notes **R125-128**, 162 (2009).

[115] J. Kiveni, *A Search for WIMP Dark Matter using an Optimized Chi-square Technique on the Final Data from the Cryogenic Dark Matter Search Experiment (CDMS II)*, Ph.D. thesis, Syracuse University (2012).

[116] J. Hall and L. Hsu, CDMS Analysis Notes **R130**, 5 (2009).

[117] S. Fallows, CDMS Analysis Notes **R130**, 7 (2009).

[118] S. Hertel and Z. Ahmed, CDMS Analysis Notes **R125-128**, 78 (2009).

[119] K. McCarty, CDMS Analysis Notes **R125-128**, 26 (2009).

[120] K. McCarthy, CDMS Analysis Notes **R125-128**, 79 (2009).

[121] J. Filippini, CDMS Analysis Notes **R123**, 181 (2008).

[122] J. Hall, CDMS Analysis Notes **R125-128**, 112 (2009).

[123] F. Fallows, CDMS Analysis Notes **R125-128**, 83 (2009).

[124] T. Bruch, CDMS Analysis Notes **R125-128**, 58 (2009).

[125] S. Arrenberg and T. Bruch, CDMS Analysis Notes **R125-128**, 115 (2009).

[126] S. Yellin, (2007), arXiv:0709.2701 .

[127] D. Speller, CDMS Analysis Notes **R125-128**, 215 (2012).

[128] D. Speller, CDMS Analysis Notes **R125-128**, 215c (2012).

[129] A. Villano, CDMS Analysis Notes **R125-128**, 205a (2011).

[130] A. Villano, CDMS Analysis Notes **R125-128**, 205 (2011).

[131] S. Arrenberg, CDMS Analysis Notes **R125-128**, 92 (2009).

[132] D. Speller, CDMS Analysis Notes **R125-128**, 214 (2012).

[133] J. Filippini, CDMS Analysis Notes **R123**, 187 (2008).

[134] M. Ahmed, CDMS Analysis Notes **R125-128**, 142 (2009).

[135] T. Bruch, CDMS Analysis Notes **R125-128**, 101 (2009).

[136] T. Bruch, *A Search for Weakly Interacting Particles with the Cryogenic Dark Matter Search Experiment*, Ph.d. thesis, University of Zurich (2010).

[137] T. Hofer, CDMS Analysis Notes **R125-128**, 349 (2013).

[138] P. Cushman, CDMS Analysis Notes **R125-128**, 351 (2013).

[139] M. Pepin, CDMS Backgrounds Notes , 218 (2002).

[140] M. Pepin, CDMS Backgrounds Notes , 221 (2004).

[141] M. Pepin, CDMS Backgrounds Notes , 222 (2004).

[142] R. Agnes *et al.* (SuperCDMS Collaboration), "Improved wimp-search reach of the cdms ii germanium data," Unpublished.

[143] R. Agnese, Z. Ahmed, A. J. Anderson, S. Arrenberg, D. Balakishiyeva, R. Basu Thakur, D. A. Bauer, A. Borgland, D. Brandt, P. L. Brink, *et al.* (CDMS Collaboration), Phys. Rev. D **88**, 031104 (2013).

[144] T. Hofer, CDMS Analysis Notes **R125-128**, 267 (2012).

[145] T. Hofer, CDMS Analysis Notes **R125-128**, 373 (2013).

[146] S. Yellin, (2011), arXiv:1105.2928 .

[147] J.-C. Liu, Z. Zhu, and H. Zhang, (2010), arXiv:1010.3773 .

[148] H. Goldstein, C. P. Poole, Jr., and J. L. Safko, *Classical Mechanics (3rd Edition)*, 3rd ed. (Addison-Wesley, 2001).

[149] M. E. Huber, "A practical introduction to dc squids and squid series array amplifiers," (2005), http://cdms.berkeley.edu/cdms_restricted/cdmsnotes/0408/040831/Intro_to_SQUIDs.pdf.

[150] D. Moore, CDMS Analysis Notes **R125-128**, 164 (2009).

[151] D. Moore, CDMS Analysis Notes **R125-128**, 102 (2009).

[152] J. Filippini, CDMS Analysis Notes **R123**, 236 (2008).

Appendix A

Annual Modulation of Earth Speed

In the Galactic coordinate system, the circular speed of the local standard of rest (LSR) of the Sun is $(0, 220, 0)$ km s^{-1} [27]. And the peculiar motion of the Sun with respect to the LSR is $(9, 12, 7)$ km s^{-1} [28]. The origin of the Galactic coordinate is at the center of the Sun. The x-axis is from the Sun to the center of the Galaxy. The z-axis is pointing to the Galactic north pole. In the J2000.0 equatorial system, the Galaxy center is in the direction RA $= 12^{\mathrm{h}}51^{\mathrm{m}}26.27549^{\mathrm{s}}$, Dec $= 27°07'41.7043''$, and the Galactic north pole is at $17^{\mathrm{h}}45^{\mathrm{m}}37.19910^{\mathrm{s}}$, $-28°56'10.2207''$ [147].

It is most convenient to describe the Earth's orbital motion around the Sun in the ecliptic coordinate system. We neglect the shifts of the ecliptic coordinates due to the precession and nutation of the Earth. In the Epoch J2000.0, the x-axis is at $0^{\mathrm{h}}0^{\mathrm{m}}0^{\mathrm{s}}$, $0°0'0''$ by definition; and the z-axis is pointing to the ecliptic north pole which is $18^{\mathrm{h}}0^{\mathrm{m}}0^{\mathrm{s}}$, $66°33'38.55''$.

In addition, we neglect the small ellipticity of the Earth orbit and assume the Earth has a constant mean circular speed $u_E = 29.79$ km s^{-1} [28]. We choose the

starting time t_0 at the vernal equinox, which is around March 20. In the ecliptic coordinates, the velocity of the earth is

$$\mathbf{u}(\lambda) = u_E(-\sin\lambda, \cos\lambda, 0), \tag{A.1}$$

where λ is the ecliptic longitude and $\lambda(t_0) = \pi$.

If the coordinate bases transformation between the ecliptic and equatorial systems has the form

$$(\mathbf{e}_{s,x}, \mathbf{e}_{s,y}, \mathbf{e}_{s,z}) = (\mathbf{e}_{e,x}, \mathbf{e}_{e,y}, \mathbf{e}_{e,z})A, \tag{A.2}$$

then the corresponding coordinate transformation is

$$\mathbf{x}_e = A\mathbf{x}_s, \tag{A.3}$$

where subscript s and e denote the ecliptic and the equatorial systems, respectively. Since $(\mathbf{e}_{e,x}, \mathbf{e}_{e,y}, \mathbf{e}_{e,z})$ is a unit matrix in its own basis, we have

$$A = (\mathbf{e}_{s,x}, \mathbf{e}_{s,y}, \mathbf{e}_{s,z})_e. \tag{A.4}$$

Similarly, between the Galactic and the equatorial systems, we have

$$B = (\mathbf{e}_{g,x}, \mathbf{e}_{g,y}, \mathbf{e}_{g,z})_e. \tag{A.5}$$

With the two transformations combined, between the galactic and the ecliptic system, there is

$$\mathbf{x}_g = B^{-1}A\mathbf{x}_s = B^T A\mathbf{x}_s = C\,\mathbf{x}_s. \tag{A.6}$$

With the aforementioned equatorial coordinates for the axes of the ecliptic and the Galactic systems, the numerical values for the transformation matrices are

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.9175 & -0.3978 \\ 0 & 0.3978 & 0.9175 \end{pmatrix}, \tag{A.7}$$

$$B = \begin{pmatrix} -0.0549 & 0.4941 & -0.8677 \\ -0.8734 & -0.4448 & -0.1981 \\ -0.4838 & 0.7470 & 0.4560 \end{pmatrix}, \tag{A.8}$$

and

$$C = \begin{pmatrix} -0.0549 & -0.9938 & -0.0965 \\ 0.4941 & -0.1110 & 0.8623 \\ -0.8677 & -0.0004 & 0.4971 \end{pmatrix}. \tag{A.9}$$

The net velocity of the Sun in the Galaxy is (9, 232, 7) km s^{-1}, and the z-axis of the ecliptic coordinates transformed into the Galactic coordinates is just the third column of C. Thus we obtain the angle between the velocity of the Sun and the ecliptic plane is 61°. Combining Eqs. (A.1) and (A.6), as well as the proper and peculiar motions of the Sun, we obtain the velocity of the Earth in the Galactic coordinates

$$\begin{aligned} \mathbf{u}_g = C\mathbf{u} = (&1.63547 \sin\lambda - 29.6053 \cos\lambda + 9, \\ &- 14.7192 \sin\lambda - 3.30669 \cos\lambda + 232, \\ &25.8488 \sin\lambda - 0.011916 \cos\lambda + 7). \end{aligned} \tag{A.10}$$

Only keep the $\sin\lambda$ term to the first order, the speed of the Earth is then

$$u_g = |\mathbf{u}_g| \approx 234.183 + 14.4378\sin(\lambda + \pi + 0.3107). \tag{A.11}$$

The relative speed u_g reaches its maximum value when $\lambda + 0.3107 = 3\pi/2$, i.e., $t - t_0 = 73.25$ days, which is around June 2 with $t_0 \sim$ March 20. Equation (A.11) is slightly different from Eq. (1.76) due to the perpendicular component of the ecliptic velocity of the Earth with respect to the Galactic velocity of the Sun and the $\sim 1^\circ$ angle difference.

Appendix B

Josephson Junction and DC SQUID

B.1 Josephson junction

When an extremely thin (∼15 Å) layer of insulator is sandwiched between two superconductors, the Cooper pairs can tunnel through the barrier and form supercurrents. If we connect the junction to a DC current source and crank up the current from zero, we shall find the current flowing through the junction can reach up to a critical current I_c without developing a voltage drop across the junction. At the instant the current passes over I_c, the junction becomes resistive because of the participation of single electron tunneling. However, the tunneling current is far smaller than the critical currents of the two superconductors and they still remain superconducting. Things are different if we sweep the current back to zero from $I \geq I_c$, the junction does not become superconducting at $I = I_c$, but remains resistive and become superconducting until much later. The behaviors of the *supercurrents* are described by the Josephson

equations

$$I = I_c \sin\varphi, \tag{B.1}$$

$$\frac{\mathrm{d}\varphi}{\mathrm{d}t} = \frac{2e}{\hbar} V, \tag{B.2}$$

where φ is the phase* difference between the two superconductors, V is the voltage across the junction, and e is the electric charge. It is easy to see that when $V = 0$, Eq. (B.1) gives a constant supercurrent, because of which it is referred to as the DC Josephson effect. Correspondingly, the case with $V \neq 0$ is referred to as the AC Josephson effect, where the supercurrent oscillates at frequency 483.6 MHz/μV.

In this short note, we vaguely call the superconductor-insulator-superconductor (SIS) junctions which have the Josephson effects as Josephson junctions. Please note that the naming may be slightly different in different literature. Because of the hysteresis behavior described above, Josephson junctions are usually parallelized with a shunt resistor (the self resistance may be sufficient for some types of junctions) in real applications, as shown in the equivalent circuit digram in Fig. B.1. Thin film junctions may also have non-negligible capacitances. An inductor could also be added in series with the diagram in Fig. B.1, however, the inductances are normally small. We also include a noise current I_N in the circuit diagram. The cross denotes an ideal Josephson junction which is described by the Josephson equations. With the junction

* Here "phase" is the phase of the wave function of the Cooper pairs. For a plane wave $\psi \propto e^{-i\mathbf{p}\cdot\mathbf{x}/\hbar}$, the phase change along any path C is $\Delta\varphi_0 = \frac{1}{\hbar}\int_C \mathbf{p}\cdot\mathrm{d}\mathbf{l}$. The canonical momentum of a charged particle in electromagnetic field is $\mathbf{P} = m\mathbf{v} - q\mathbf{A}$. The phase change along a closed path C is $\Delta\varphi = \frac{1}{\hbar}\oint_C \mathbf{P}\cdot\mathrm{d}\mathbf{l} = \Delta\varphi_0 - \frac{q}{\hbar}\oint_C \mathbf{A}\cdot\mathrm{d}\mathbf{l} = \Delta\varphi_0 - \frac{q}{\hbar}\Phi$, where Φ is the magnetic flux.

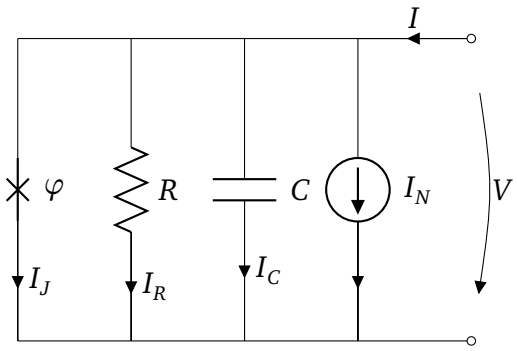


Figure B.1: Equivalent circuit of the Josephson junction. The junction at the left denotes an ideal Josephson junction which only has supercurrent. R and C are the resistance and the capacitance, respectively. I_N is the noise current. The junction is biased under a current source.

biased under a current source, the current is given by

$$I = I_c \sin\varphi + \frac{V}{R} + C\frac{\mathrm{d}V}{\mathrm{d}t} + I_N. \tag{B.3}$$

Substituting Eq. (B.2) into the above equation, we obtain

$$\frac{\hbar C}{2e}\frac{\mathrm{d}^2\varphi}{\mathrm{d}^2 t} + \frac{\hbar}{2eR}\frac{\mathrm{d}\varphi}{\mathrm{d}t} + I_c \sin\varphi = I - I_N. \tag{B.4}$$

Equation (B.4) has the exact form of the equation of motion for a damped pendulum. There is no analytical solution in its most general form. For a more detailed discussion about the analogy to a pendulum and the chaotic behavior of the equation under certain conditions, the reader can refer to [148].

Below we discuss an important case for application in which the capacitance is small. We also neglect the noise current to keep things simple. In this case, Eq. (B.4)

becomes

$$\frac{\hbar}{2eR}\frac{\mathrm{d}\varphi}{\mathrm{d}t} + I_c \sin\varphi = I. \tag{B.5}$$

It is apparent that $\varphi = \arcsin\frac{I}{I_c}$ is a solution for $I < I_c$, which gives the DC constant supercurrent and zero voltage drop. For $I \geq I_c$, φ is a periodic function of time and hence the voltage V. Equation (B.5) can be solved analytically, however, we are interested in the I–V characteristic of the junction, which can be obtained in a simpler approach. The DC voltage of the junction is the time average of V in a full cycle,

$$\bar{V} = \frac{1}{T}\int_0^T V\,dt = \frac{\hbar}{2eT}\int_0^T \frac{\mathrm{d}\varphi}{\mathrm{d}t}\mathrm{d}t = \frac{\hbar}{2eT}\int_0^{2\pi}\mathrm{d}\varphi = \frac{h}{2eT}, \tag{B.6}$$

where T is the period. And

$$T = \int_0^T \mathrm{d}t = \int_0^{2\pi}\left(\frac{\mathrm{d}\varphi}{\mathrm{d}t}\right)^{-1}\mathrm{d}\varphi = \frac{\hbar}{2eR}\int_0^{2\pi}\frac{\mathrm{d}\varphi}{I - I_c\sin\varphi} = \frac{h}{2eR\sqrt{I^2 - I_c^2}}. \tag{B.7}$$

We have used Eq. (B.5) in the above derivation. Combining Eqs. (B.6) and (B.7) gives

$$\bar{V} = R\sqrt{I^2 - I_c^2}. \tag{B.8}$$

The complete I–V curve can be expressed as

$$\bar{V} = \begin{cases} 0, & I < I_c, \\ R\sqrt{I^2 - I_c^2}, & I \geq I_c. \end{cases} \tag{B.9}$$

B.2 DC SQUID

As shown in Fig. B.2, a DC SQUID consists of two Josephson junctions in a superconducting loop, which are parallely current biased. Now consider the phase difference between Point B and Point A from two branches

$$\varphi_a + \frac{2e}{\hbar}\int_a \mathbf{A}\cdot d\mathbf{l} = \varphi_b + \frac{2e}{\hbar}\int_b \mathbf{A}\cdot d\mathbf{l}, \tag{B.10}$$

in which the fact the charge for a Cooper pair $-2e$ has been used. Then we have

$$\varphi_a - \varphi_b = -\frac{2e}{\hbar}\int_a \mathbf{A}\cdot d\mathbf{l} + \frac{2e}{\hbar}\int_b \mathbf{A}\cdot d\mathbf{l} = -\frac{2e}{\hbar}\oint \mathbf{A}\cdot d\mathbf{l} = -\frac{2e}{\hbar}\Phi = -2\pi\frac{\Phi}{\Phi_0}, \tag{B.11}$$

where Φ is the magnetic flux through the loop pointing out of the paper[†] and $\Phi_0 \equiv \frac{h}{2e}$ is the flux quantum.

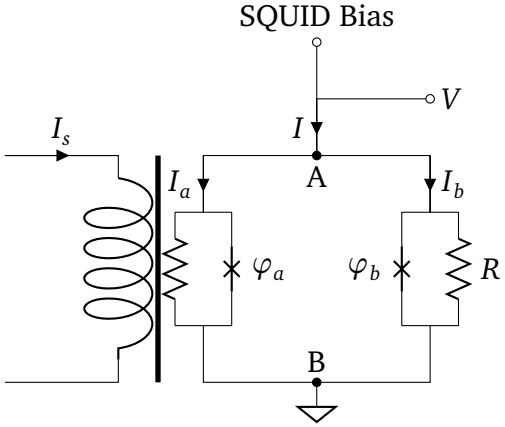


Figure B.2: Circuit diagram of the DC SQUID and the inductive coupling to the input coil.

The difference in φ in Eq. (B.11) gives an interference pattern of the supercurrents

[†]The direction of the flux follows the convention of φ_a in this case.

for the two branches in a similar fashion to the double-slit interference, i.e.,

$$\begin{aligned} I &= I_a + I_b = I_c \sin\varphi_a + I_c \sin\varphi_b = 2I_c \sin\frac{\varphi_a+\varphi_b}{2}\cos\frac{\varphi_a-\varphi_b}{2} \\ &= 2I_c \cos\left(\frac{\Phi\pi}{\Phi_0}\right)\sin\left(\varphi_a + \frac{\Phi\pi}{\Phi_0}\right). \end{aligned} \tag{B.12}$$

By comparing Eq. (B.12) with Eq. (B.1), we see that two Josephson junctions in parallel behave much like a single junction with a flux modulated critical supercurrent

$$I_{c,s} = 2I_c \left|\cos\left(\frac{\Phi\pi}{\Phi_0}\right)\right|. \tag{B.13}$$

Similarly, the I–V curves are

$$\bar{V} = \begin{cases} 0, & I < I_{c,s}, \\ \frac{R}{2}\sqrt{I^2 - 4I_c^2\cos^2\left(\frac{\Phi\pi}{\Phi_0}\right)}, & I \geq I_{c,s}. \end{cases} \tag{B.14}$$

For $I \geq 2I_c$, $\bar{V}$ is a periodic function of Φ, which is also referred to as the V–Φ curve. The magnitude of $\bar{V}$ falls in the range $[\frac{R}{2}\sqrt{I^2-4I_c^2},\ \frac{IR}{2}]$.

The V–Φ curve can be used to measure magnetic flux and hence the current generating the field. Because $\Phi_0 = 2.07 \times 10^{-15}$ Wb is such a small value, the DC SQUID is tremendously sensitive to magnetic field. However, we are more interested in the magnetic flux change due to the change of input current rather than the absolute flux. In this case, the SQUID can be locked at a given flux so that a small variation due to input current climbs up and down along the V–Φ curve from the lock point. It can be shown from Eq. (B.14) that the maximum slope on the V–Φ curve is $\pi R I_c/\Phi_0$

when $I = 2I_c$. For $I > 2I_c$, the gain will be reduced, for the sake of completeness, it is

$$\left.\frac{d\bar{V}}{d\Phi}\right|_{\max} = \frac{\pi R I_c}{\Phi_0}\frac{\sqrt{y(1-y)}}{\sqrt{k^2-y}}, \tag{B.15}$$

where

$$y = k^2 - k\sqrt{k^2-1}, \tag{B.16}$$

$$k = \frac{I}{2I_c}. \tag{B.17}$$

In the derivations above the SQUID was assumed to be ideal, in reality, $I_{c,s}$ can not swing from 0 and $2I_c$ and the gain is also reduced. The SQUID amplifier also has to be tuned to avoid resonances. For a detailed analysis of the CDMS SQUID amplifier, the reader can refer to [92].

At the end, we list some numbers for the CDMS SQUID circuit which may be useful for calculations. Each SQUID chip has 100 units as shown in Fig. B.2 connected in series. The parameters for each unit at the operating point are: $I_s \approx 1\,\text{nA}$, $I \approx 100\,\mu\text{A}$, $R \approx 2\,\Omega$, $V_o \approx 50\,\mu\text{V}$, $Z_{\text{out}} \approx 1\,\Omega$, and the bandwidth is between DC and 1–10 GHz. These numbers are taken from [149].

Appendix C

Surface Event Leakage Estimation with a Bayesian Approach

This chapter presents the method used to estimate the post-unblinding surface event leakage for the C58R analysis. This work is based on the Bayesian framework introduced in Appendix C of the thesis by Filippini [91] and the note by Moore [150], with changes made to better suit the data in the C58R analysis.

The expected number of surface events misidentified by the timing cuts as WIMP signals is the product of the expected number of nuclear recoil single scatters (NRSS) failing the timing cut and the expected timing cut pass/fail ratio of the NRSS. For the final leakage estimate after unblinding, the observed numbers of NRSS in the WIMP-search (WS) data were used (see Sec. 5.4.3.2 for before unblinding). The NRSS timing cut pass/fail ratios were measured on the nuclear recoil multiple scatters (NRMS) in the WS data and the wide beta band multiples (WBMS) in the WS and ^{133}Ba calibration data. Three data samples were used to reduce the potential systematic

errors in each of them.

The interior detectors and the endcap detectors were treated differently in leakage calculations when the WBMS data were used because the endcaps did not have multiples at their exterior face. For the same reason and lack of statistics, the endcaps were not used with the WS NRMS. In total, there were five individual leakages in the whole procedure and they were combined at the end to obtain the final estimate.

C.1 Leakage estimation methods

C.1.1 Method 1

Method 1 uses the WS NRMS timing cut pass/fail ratios to estimate those of the WS NRSS. The leakage can be written as

$$n = \sum_z N_z \frac{b_z}{B_z}, \tag{C.1}$$

where n is the total leakage, z is the detector index, N_z is the number of WS NRSS failing the timing cut for Detector z, b_z and B_z are the number of WS NRMS passing and failing the timing cuts, respectively. The sum is over all interior detectors used in the WIMP-search analysis.

Method 1 is expected to have the least systematic errors since both event samples are in the WS NR band. As mentioned before, endcaps were not used in Method 1 because there were no multiples tagged on the exterior face of these detectors.

C.1.2 Method 2++

The timing cut pass/fail ratios of the WS NRSS were measured on the WS WBMS and Ba WBMS in Method 2++. The use of WBMS could bring in larger systematic errors compared with Method 1. At least it was found there was systematic difference between the WS WBMS and the Ba WBMS [151]. But there were more statistics in the WBMS which could reduce the statistical errors. To reduce the systematic errors due to the difference between the NRSS and the WBMS, the events were divided into six energy/face bins. Three energy bins, 10 to 20 keV, 20 to 30 keV, and 30 to 100 keV, and two face bins, p-side and q-side, were used.

The expression for the total leakage is

$$n = \sum_{z,e,f} N_z s_{z,e,f} \frac{m_{z,e,f}}{M_{z,e,f}}, \tag{C.2}$$

where n, z, and N_z are the same as in Method 1, e labels the low, medium, or high energy bin, f labels the charge or phonon face, $s_{z,e,f}$ is the fraction of NRSS in (energy, face) bin (e,f) measured on WS NRMS, $m_{e,f}$ and $M_{e,f}$ are the number of WBMS passing and failing the timing cuts, respectively.

The leakages on the interior detectors and the endcaps were calculated separately in Method 2++ because the latter needed additional corrections to Eq. (C.2).

C.2 Bayesian theorem

A Bayesian approach was used to estimate the leakages and their errors in C34 and C58 analyses [91, 150] because of its superior performance and simplicity in implementation compared to an earlier frequentist framework [91]. The same Bayesian approach was used in the C58R analysis as well.

The Bayes' theorem states that

$$p(\theta|x) = \frac{p(x|\theta)\,p(\theta)}{\int p(x|\theta)\,p(\theta)\,d\theta}, \tag{C.3}$$

where

- x is the observed data. It is a scalar here, but it can also be a vector representing a set of observations.
- θ is an unknown parameter we wish to determine. It can also be a vector.
- $p(\theta)$ is the prior p.d.f. of the parameter θ, which represents our knowledge about θ before the experiment.
- $p(x|\theta)$ is the conditional p.d.f. or likelihood function of x given θ.
- $p(\theta|x)$ is the posterior p.d.f. of θ given the observation x.
- The denominator on the right hand side is a normalization factor which makes $p(\theta|x)$ a proper p.d.f..

The posterior p.d.f. $p(\theta|x)$, which summarizes all knowledge about θ, represents our degree of belief of the distribution of θ after the experiment. The point estimators for θ can be derived from its posterior. The median was used as the estimator for θ in this analysis as what was done in the C58 and C34 analyses.

C.3 Choice of prior

The prior $p(\theta)$ represents our historical knowledge about θ. It is quite often that we know nothing about θ before the experiment. In these cases, a class of uninformative or objective priors are preferred in seeking objectivity.

C.3.1 Poisson distribution

The likelihood function of the Poisson distribution is

$$p(k|\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}. \tag{C.4}$$

The class of priors of the form $p(\lambda) \propto \lambda^c$ are considered as uninformative for the Poisson distribution.

Comparing Eq.(C.3) and Eq.(C.4), we can see that the posterior p.d.f. also has the form λ^c. If the posterior and the prior both are drawn from the same family of distributions, the prior is referred to as a *conjugate piror*. The conjugate prior of the Poisson distribution is the Gamma distribution

$$p_\gamma(x|\alpha,\theta) = \frac{x^{\alpha-1}e^{-x/\theta}}{\Gamma(\alpha)\,\theta^\alpha}. \tag{C.5}$$

With an observation of k events from a Poisson process, the posterior is another Gamma distribution with $\alpha \to \alpha + k$ and $1/\theta \to 1/\theta + 1$. The priors of the form x^c are limits of the Gamma distribution with $\alpha = c + 1$ and $\theta \to \infty$. The corresponding posterior is then

$$p_\gamma(x|k + c + 1, 1) = \frac{x^{k+c}e^{-x}}{\Gamma(c+1)}. \tag{C.6}$$

The Jaynes prior ($c = -1$) and the Jeffreys prior ($c = -1/2$) are the two commonly used priors among the uninformative class. We prefer the Jaynes prior because the mean of a Poisson sample is invariant under binning.

C.3.2 Binomial distribution

The probability of having k successes in n trials with the probability of success x for each trial follows the binomial distribution. The likelihood function is

$$p(k|n,x) = \binom{n}{k} x^k (1-x)^{n-k}. \tag{C.7}$$

The conjugate prior of the binomial distribution is the beta distribution

$$p_\beta(x|\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}. \tag{C.8}$$

A binomial process can be viewed as two Poisson processes with successes k and failures $(n-k)$. If the priors x^{c_k} and $x^{c_{n-k}}$ are chosen for the two Poisson processes, the equivalent prior for the binomial process is the beta distribution with $\alpha = c_k + 1$ and $\beta = c_{n-k} + 1$.

With an observation of k successes in n trials, the posterior is another beta distribution with $\alpha \to \alpha + k$ and $\beta \to \beta + (n-k)$. There are three obvious choices of uninformative priors for the binomial distribution:

1. $p(x) \propto x^{-1}(1-x)^{-1}$. This is a limit of the beta distribution with $\alpha, \beta \to 0$, which corresponds to the Jaynes prior with $c_k = c_{n-k} = -1$.
2. $p(x) \propto x^{-1/2}(1-x)^{-1/2}$. This is a beta distribution with $\alpha = \beta = 1/2$, which

corresponds to the Jeffreys prior with $c_k = c_{n-k} = -1/2$.

3. $p(x) = 1$, or the uniform prior. This corresponds to $\alpha = \beta = 1$ for the beta distribution or $c_k = c_{n-k} = 0$ for the Gamma distribution.

The uniform prior was used in the C58 analysis when generating the fractions of single scatters in the nuclear recoil band, NRSS/(NRSS + NRMS), with the MATLAB function `betarnd`.

C.3.3 Multinomial distribution

Multinomial distribution is an extension of the binomial distribution where more than two categories of events can be classified in each trial. The probability mass function is

$$p(n_1, \cdots, n_K | n, x_1, \cdots, x_K) = \frac{n!}{n_1! \cdots n_K!} x_1^{n_1} \cdots x_K^{n_K}, \tag{C.9}$$

where

$$\sum_{i=1}^{K} n_i = n, \tag{C.10}$$

$$\sum_{i=1}^{K} x_i = 1. \tag{C.11}$$

The case with $K = 2$ is the binomial distribution.

The conjugate prior of the multinomial distribution is the Dirichlet distribution:

$$p_D(x_1, \ldots, x_K | \alpha_1, \ldots, \alpha_K) = \frac{\Gamma(\sum_{i=1}^{K} \alpha_i)}{\prod_{i=1}^{k} \Gamma(\alpha_i)} \prod_{i=1}^{K} x_i^{\alpha_i - 1}. \tag{C.12}$$

The uninformative priors can be chosen similarly as for the binomial case:

1. $p(x_1, \ldots, x_K) \propto \prod_{i=1}^{K} x_i^{-1}$, or $\alpha_i \to 0$. This corresponds to the Jaynes prior for the Poisson distribution.
2. $p(x_1, \ldots, x_K) \propto \prod_{i=1}^{K} x_i^{-1/2}$, or $\alpha_i = 1/2$. This corresponds to the Jeffreys prior for the Poisson distribution.
3. $p(x_1, \ldots, x_K) = 1$, or $\alpha_i = 1$, which is the uniform prior.

The uniform prior was used in the C58 analysis when generating $s_{e,f}$, the fractions of NRSS in each energy/face bin.

C.4 Leakage estimation procedure

Systematic errors in leakages due to different sources were calculated independently and then added together by quadrature. The total systematic error was then added with a standard Gaussian profile to the posterior of corresponding leakage. The estimates of each leakage and its statistical error are the median and the 68% central confidence interval of the posterior, respectively.

To account for the effects of prior dependence, we also included an additional systematic error due to choice of prior. A list of priors was applied to a number of test cases which were chosen by varying the observed data. Bias and coverage of the nominal leakage of each test case for each prior were then calculated. A reference test case was chosen for each observation. For the reference test case, the best prior was then chosen where the bias was minimized and the coverage was acceptable. The spread of bias at the optimal prior was assigned to a systematic error. The procedure is detailed below.

1. Based on the observed result of the experiment, a number of, say T, possible

outcomes of the experiment were chosen as the test cases by varying the observed event counts on the detectors. A few typical test cases were:

(a) All the interior (endcap) detectors had the same leakage and the total leakage was equal to the observed total leakage n_i (n_e) on interior (endcap) detectors.

(b) All the interior (endcap) detectors had the same leakage and the total leakage was $n_i \pm \sigma_i$ ($n_e \pm \sigma_e$), where n_i (n_e) was the observed total leakage on interior (endcap) detectors and σ_i (σ_e) was the Poisson error.

(c) All the observed leakage on interior (endcap) detectors n_i (n_e) were from one (or two) of the detectors.

(d) A finite but small number of leaked events were assigned to a detector when zero were observed.

2. For each test case, use the assigned values of N_z, b_z, B_z, s_{zef}, m_{zef}, and M_{zef} as the mean to generate G sets of Poisson distributed random numbers as G trial experiments. This can be done using the matlab function `poissrnd`.

3. Choose a number of Jaynes priors. For each trial experiment and prior combination, generate L random numbers following the posterior distribution. The conjugate prior of the Poisson distribution is the Gamma distribution. The random numbers can be generated using the matlab function `gamrnd`. The fractions of NRSS in each energy/face bin s_{zef} can be generated using the Dirichlet distribution with a uniform prior.

4. The leakage is calculated for each set of random numbers. The estimates of leakage and its statistical error are chosen as the median and the 68% central confidence interval of the leakage posterior for each trial experiment.

5. Now we have a set of G leakages $\{\Lambda_i\}$ and the corresponding 68% CI's $\{\Delta_i\}$ for each prior and test case. The estimated leakage Λ_p and statistical error Δ_p for the prior and test case combination are then chosen as the *median* of $\{\Lambda_i\}$ and $\{\Delta_i\}$, respectively. The coverage is defined as the fraction of Δ_i's which covers the test case nominal leakage Λ_t. The bias is the difference $(\Lambda_p - \Lambda_t)$.
6. A reference test case is chosen based on which the optimal prior will be determined. This can be the observed case if the total leakage is not zero.
7. For the reference test case, the optimal prior is chosen as the one which minimizes the bias and has roughly 68% coverage. The estimated leakage Λ_p and statistical error Δ_p are then determined based on the chosen prior.
8. The spread of bias at the optimal prior over the test cases is defined as the systematic error.
9. If the observed total leakage on the interior (endcap) detectors is zero, the reference test case can not be chosen as the observed case because the nominal leakage is zero and the resulting biases are always positive and coverages are always zero except $c = -1$. In this case, to be conservative, the 90% upper limit of the number of expected leaked events is assigned to the detector to choose the optimal prior and systematic error due to choice of prior. Once the optimal prior is chosen, the estimated leakage and statistical error still uses the observed counts. This treatment was applied to the endcaps in this analysis because zero leakage were observed.

C.5 Leakage estimates

C.5.1 Method 1

As discussed in Refs. [152] and [150], there are two possible systematics in Method 1:

1. Difference in timing cut pass/fail ratio between NRSS and NRMS
2. Bias or under coverage due to choice of prior.

C.5.1.1 Systematic difference between singles and multiples

Method 1 assumes that the WS NRSS have the same timing cut pass/fail ratio as the WS NRMS. Unfortunately, this can not be tested with the observed data. First, the statistics of the NRSS and NRMS passing the timing cut are very low; the resulting statistical errors are likely larger than the systematics, which makes the test practically impossible. Second, the NRSS events passing the timing cut contain possible WIMP signals, which biases the NRSS timing cut pass/fail ratio upward. A possible systematic difference in the timing cut pass/fail ratio between the NRSS and NRMS would lead to an error on the leakage estimate. Here we estimate this systematic difference using the WBSS and WBMS in the WS data.

Table C.1 lists the number of WS WBSS and WBMS passing/failing the timing cut in the third and fourth columns for the three analyses. Define r as the ratio of the timing cut pass/fail ratio of WBSS to that of WBMS

$$r = (s/S)/(m/M), \tag{C.13}$$

Table C.1: Systematic difference of timing cut pass/fail ratio between singles and multiples in the WIMP-search wide beta band.

					84%/90%[b] upper bounds of r				
		Pass	Fail	r_o[a]	α_{up}(%)	$c:-0.5$	$c:-0.9$	$c:-1$	Sys. (%)[c]
C58R	S.	1	530	0.42	84	1.17	0.93	0.87	NA
	M.	9	2024		90	1.45	<u>1.19</u>	<u>1.12</u>	15
C58	S.	1	473	0.55	84	1.54	<u>1.25</u>	<u>1.16</u>	20[d]
	M.	7	1828		90	1.93	1.60	1.51	NA
C34	S.	1	251	0.36	84	1.00	0.80	0.74	NA
	M.	9	821		90	<u>1.24</u>	1.02	<u>0.96</u>	10[e]

[a.] Ratio r with observed data.
[b.] The confidence level of each row is indicated by α_{up}.
[c.] Systematic difference of pass/fail ratio between singles and multiples.
[d.] From Ref. [150]. 20% is derived from the mean of the two underlined values. If the values at $c = -0.5$ and -1 were chosen, the systematic difference would be 64%.
[e.] From Ref. [152]. 10% is derived from the mean of the two underlined values.

which characterizes the difference in timing cut pass/fail ratio between WBSS and WBMS. The fifth column lists the values of r calculated using the observed counts on the left.

With a Jaynes prior chosen for the Poisson counts s, S, m, and M, the posterior of r can be obtained through Eq. (C.13). An upper bound of r from its posterior at a certain CL is a characterization of the systematic difference in timing cut pass/fail ratio between WBSS and WBMS. Columns 6-9 list the CL's and the corresponding upper bounds for three different Jaynes priors.

In previous analyses, averages of two upper bounds (underlined) of r at the same CL (α) but for different priors, $\bar{r}_\alpha$, were calculated, and $(\bar{r}_\alpha - 1)$ was then taken as the systematic difference in timing cut pass/fail ratio between WBSS and WBMS. In C34 analysis, r_{90}'s for $c = -1/2$ and $c = -1$ were chosen and a systematic error of 10% was obtained [152]. In C58 analysis, r_{84}'s for $c = -0.9$ and $c = -1$ were chosen and the systematic error was 20% [150].

In C58R analysis, the average of r_{90} for priors with $c = -0.9$ and $c = -1$ were chosen and a systematic error of 15% was obtained. There were two considerations to make this choice. First, the 90% CL was chosen to give a conservative estimate. Second, the range of c should be narrower rather than wider and close to the optimal priors chosen in the leakage estimations to avoid additional systematic effects due to averaging.

Strictly speaking, these choices seem rather subjective and arbitrary. Even with the upper bounds of r listed in the table, multiple combinations can be chosen to calculate the averages, which give quite different values of the systematic error, some are even negative. This is not surprising though if we compare the upper bounds of

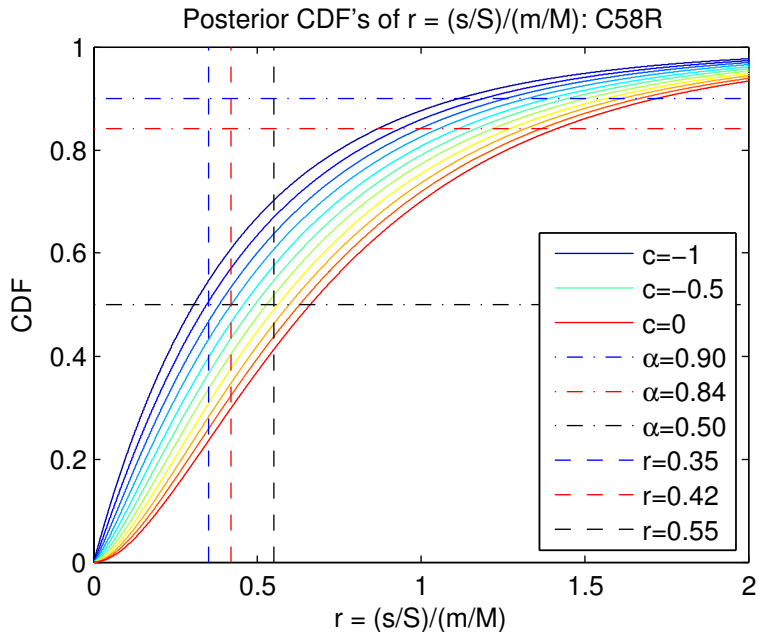


Figure C.1: Prior dependence of the cumulative probability for the posterior distributions of $r = (s/S)/(m/M)$ for WIMP-search data of C58R.

r and its observed values, which are $\ll 1$. It is possible that even with significant upward fluctuation, the value of r is still < 1, which means there is a high probability that WBMS is overestimating the timing cut pass/fail ratio of the WBSS. A reasonable choice may be to lower the timing cut pass/fail ratio of the WBMS to get closer to that of the WBSS. But by doing this we will lose our conservativeness in leakage estimation. Furthermore, there is possible systematic difference between NR and WB events. So it is difficult to quantify the over- or under- estimation. If there were enough statistics, these questions can be answered. With no more information available, the choice of a systematic error of 15% seems reasonable with the two considerations given above.

It is interesting to know whether there is a favored prior and CL to choose an upper bound of r. The cumulative probability functions (CDF) of the posteriors of r with priors from $c = -1$ to $c = 0$ in steps of 0.1 are shown in Fig. C.1. For a

rather large range of CL, the upper bounds of r linearly depend on c. Apparently, no particular c is favored over others.

As was discussed earlier in Sec. C.4, the bias of the median to the nominal value can be used to choose a prior. We can apply the same method to choose a prior for determining the upper bound of r. The three vertical lines in Fig. C.1 are the observed values of r for the C34, C58R, and C58 analyses from left to right, respectively. The posterior whose median intersects a vertical line is the posterior with minimized bias, and can be chosen as the best posterior for r of an analysis. The priors determined this way for the three analyses are very different. The exponent c increases as the observed r increases. More importantly, these priors are all outside the [-1, -0.9] range of c, and may not be good choices to characterize the systematic difference between WBSS and WBMS.

C.5.1.2 Choice of prior

The second systematic error is from the choice of prior in computing the leakage. The posterior of leakage is very sensitive to the choice of prior due to the very small number of events passing the timing cut. Figure C.2 shows the posteriors of the leakage for a few prior exponent c.

The systematic error due to choice of prior was calculated following the procedure described in Sec. C.4. Figure C.3 shows the biases and the coverages of the 68% central confidence intervals for six test cases (only interior detectors used):

1. $\triangle$: All detector with observed leakage. The observed N_z, b_z, and B_z are listed in Table C.2. The total leakage was $n_o = 2.27$ events.
2. $\bigcirc$: All detectors with averaged N_z, b_z, and B_z, i.e., they were assumed to be

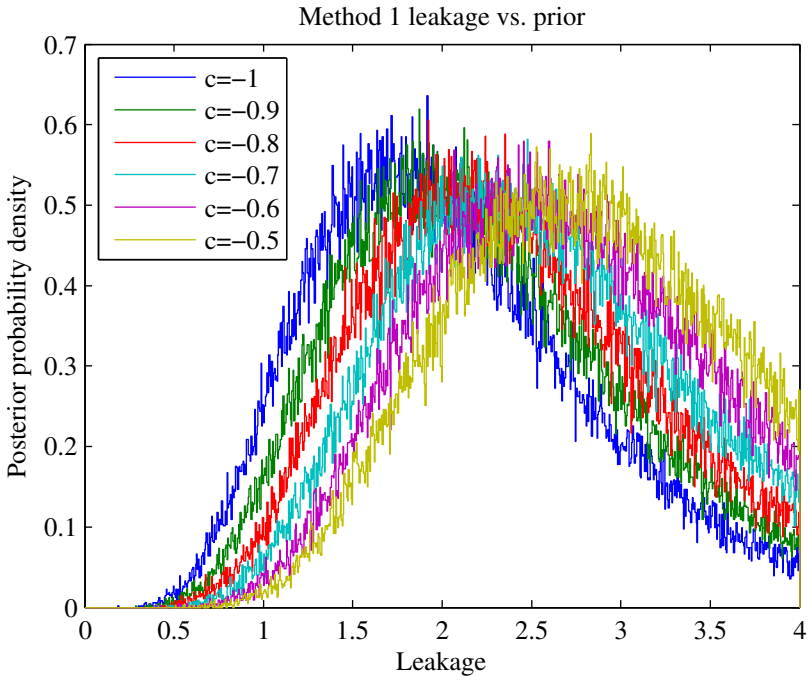
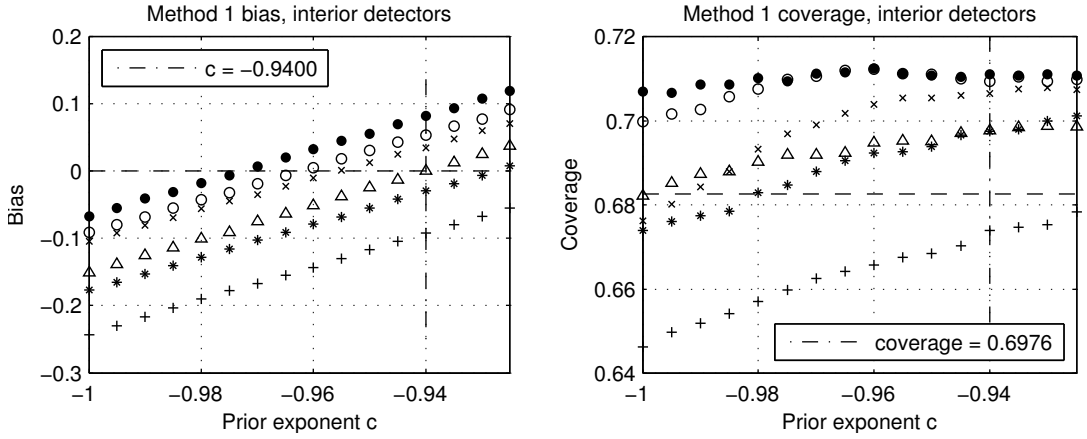


Figure C.2: Leakge posterior vs. prior for Method 1 in the C58R classic timing analysis. Systematic errors are not included.



(a) Bias vs. prior exponent c.

(b) Coverage vs. prior exponent c.

Figure C.3: Bias and coverage vs. prior exponent c for Method 1.

identical. The total leakage was $n = 1.81$ events.

3. ●: All detectors with the same leakage. Total leakage was $n + \sigma = 2.64$ events.
4. ×: All detectors with the same leakage. Total leakage was $n - \sigma = 0.98$ events.
5. +: All leakage on one detector. Total leakage was $n = 1.81$ events.
6. $\ast$: All leakage on two detectors. Total leakage $n = 1.81$ events.

The first test case, which had the observed counts, was used to choose the optimal prior. The bias reaches zero at $c = -0.9400$ and the coverage is 0.6976 which is slightly higher than the desired value. The spread of the bias at $c = -0.9400$ is from -0.09 to 0.08. The half width of this interval 0.09 is assigned as the systematic error due to choice of prior. Adding by quadrature, the total systematic error of Method 1 is

$$\sqrt{(2.27 \times 15\%)^2 + 0.09^2} = 0.35. \tag{C.14}$$

And the total leakage of Method 1 is

$$n_1 = 2.27^{+1.06}_{-0.76}\,(\text{stat.}) \pm 0.35\,(\text{sys.}). \tag{C.15}$$

C.5.2 Method 2++: Interior detectors

Method 2++ uses the timing cut pass/fail ratio of the WBMS and the energy/face distribution of the WS NRMS to calculate those of the WS NRSS. The possible systematic errors for the interior detectors are:

1. Timing cut pass/fail ratio difference between singles and multiples. This was calculated in Method 1 above. Here we use the same value of 15%.
2. Systematic error due to incomplete weighting of low energy events.

3. Systematic error due to choice of prior.

C.5.2.1 Incomplete reweighting of low energy events

Method 2++ assumes the WS NRSS follow the same energy/face distribution as the WS NRMS and have the same pass/fail ratio as the WBMS in each energy/face bin. As the timing resolution goes poorer when energy goes lower, the lower energy events tend to have higher timing cut pass/fail ratios. Binning the energy inadequately would underestimate the pass/ratio and thus underestimate the leakage.

Three energy bins were used in Method 2++ in order to reduce the error due to incomplete reweighting of low energy events. However, it may be possible to reduce this error even further. As pointed out in C58 analysis note [150], the systematic error due to under binning in energy may not be estimated using finer binning because the large statistical error would possibly outweigh the systematic difference.

Fundamentally, this error is caused by the systematic difference between WS NRSS and (WS or Ba) WBMS. Since there is least systematic difference between the WS NRSS and WS NRMS, along this line, maybe a method which uses both WS NRMS and (WS or Ba) WBMS would be able to have both advantages of small systematic difference and large statistics.

In C58 analysis, the primary motivation to bring up the inadequate energy binning error was to account for the flaw of the charge reconstruction algorithm which caused loss of timing resolution at low energies. For the C58R analysis, since this problem was fixed, though theoretically there is still this type of error, but before a better handle can be found, it may be reasonable to assume dividing the data into six energy/face bins accounted for this error adequately.

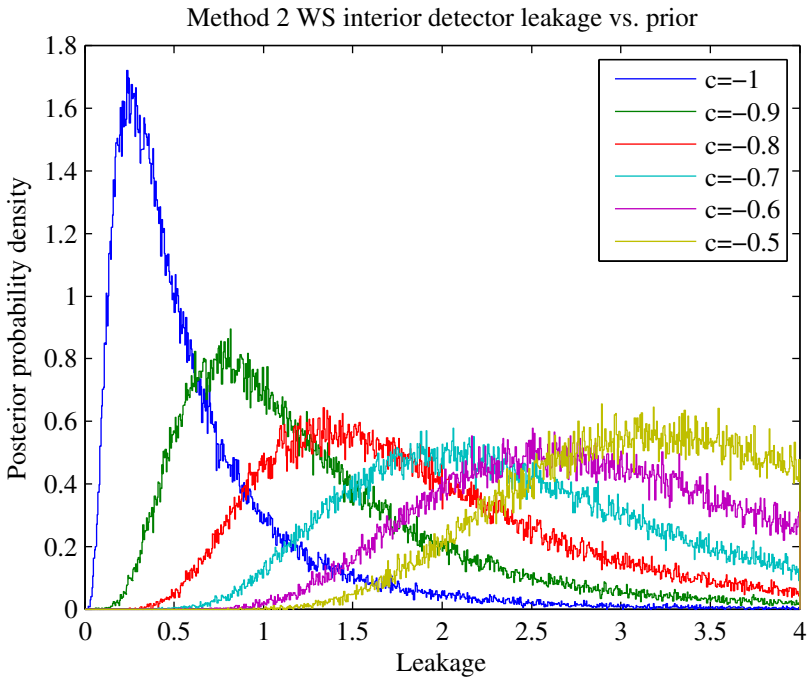


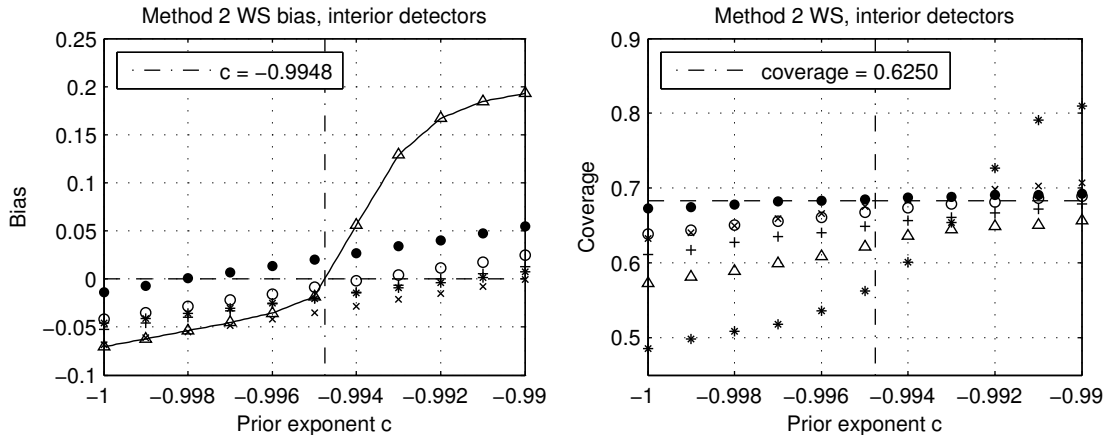
Figure C.4: Leakage posterior vs. prior for the interior detectors in Method 2++ with timing cut pass/fail ratios measured on WIMP-search wide band beta multiples in the C58R classic timing analysis. Systematic errors are not included.

C.5.2.2 Choice of prior

WS data

Figure C.4 shows the posterior of the leakage from Method 2++ with the timing cut pass/fail ratio measured on WS WBMS. Figure C.5 shows the bias and coverage for the test cases below (only for interior detectors):

1. $\triangle$: All detector with observed leakage. The total leakage was $n_o = 0.41$ events. N_z's are listed in Table C.2. The s factors are calculated using the WS NRMS counts in Table C.3. The WS WBMS m and M are listed in Table C.4.
2. $\bigcirc$: All detectors with averaged event counts of N_z, s, m and M over interior



(a) Bias vs. prior exponent c. (b) Coverage vs. prior exponent c.

Figure C.5: Leakage posterior bias and coverage for interior detectors in Method 2++ with timing cut pass/fail ratio measured on WS wide band betas.

detectors. Total leakage $n = 0.36$ events.

3. ●: All detector with the same leakage. Total leakage $n + \sigma = 0.61$ events.
4. ×: All leakage on one detector. Total leakage $n = 0.36$ events.
5. +: All detectors with averaged event counts and the true weighting factors s_{ef}. Total leakage $n = 0.36$ events.
6. $\ast$: All interior detectors with equal leakage. Total leakage $n - \sigma = 0.11$ events.

The bias is minimized at $c = -0.9948$, and the corresponding coverage is 0.7098. The spread of the bias at the chosen prior is [-0.03, 0.02], which gives a systematic error of 0.03. Combined with the 15% systematic error between singles and multiples, the leakage from the interior detectors with WS data is

$$n_2^{\mathrm{WS}} = 0.41^{+0.39}_{-0.21}\,(\mathrm{stat.}) \pm 0.07\,(\mathrm{sys.}). \tag{C.16}$$

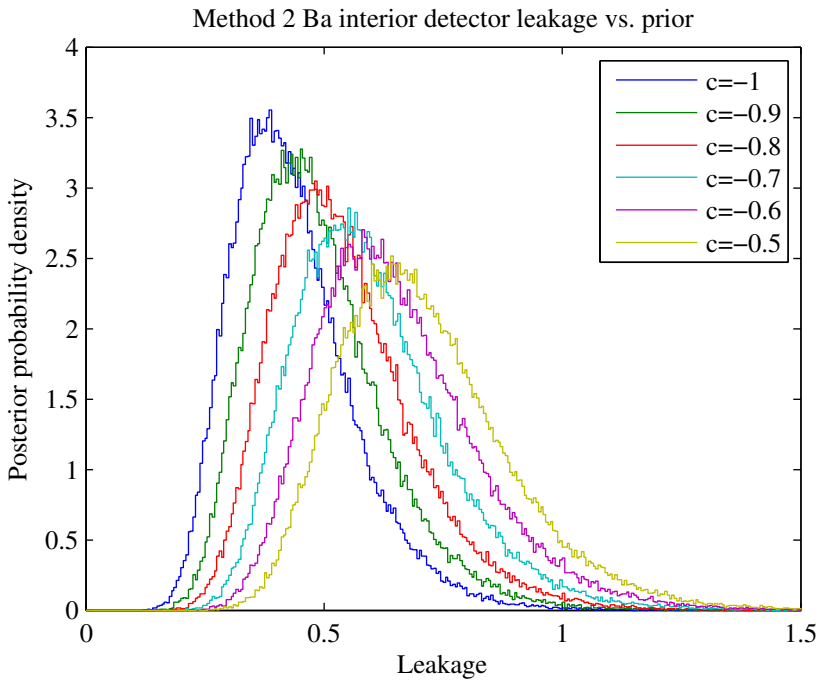
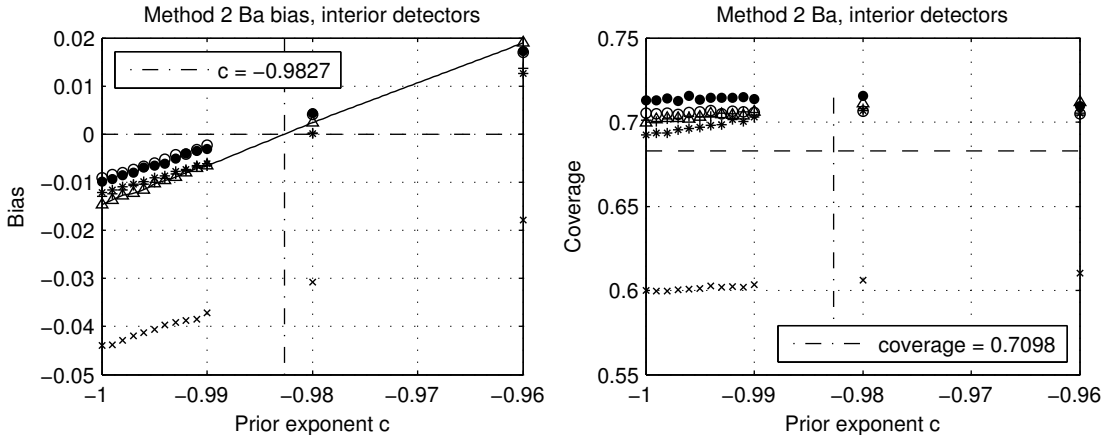


Figure C.6: Leakage posterior vs. prior for the interior detectors in Method 2++ with timing cut pass/fail ratio measured on ^{133}Ba wide band band multiples in the C58R classic timing analysis. Systematic errors are not included.



(a) Bias vs. prior exponent c

(b) Coverage vs. prior exponent c

Figure C.7: Leakage posterior bias and coverage for interior detectors in Method 2++ with timing cut pass/fail ratio measured on ^{133}Ba wide band betas.

Ba data

With the timing cut pass/fail ratio measured on Ba WBMS (see Table C.5), the leakage posterior versus prior is shown in Figure C.6. Figure C.7 shows the bias and coverage for the leakage posteriors with the test case chosen the same way as for the WS data. The prior exponent c is not evenly spaced. Since the dependence of bias on c is mostly linear, the nonuniform spacing of c should not affect the result. The bias is minimized at $c = -0.9827$. The coverage at the chosen prior is 0.6250, which is slightly undercovered. The spread of bias is $[-0.03, 0.00]$, which gives a systematic error of 0.02. The total leakage with the 15% difference between singles and multiples incorporated is

$$n_2^{\mathrm{Ba}} = 0.48^{+0.14}_{-0.11}\,(\mathrm{stat.}) \pm 0.07\,(\mathrm{sys.}). \tag{C.17}$$

C.5.3 Method 2++: Endcap detectors

C.5.3.1 Event distribution on endcaps

Two bottom endcap detectors T3Z6 and T4Z6 were used in the C58 and this analysis. The charge faces of these two detectors are exterior faces and have no tagging efficiency. This introduces two more complications.

In the C58 and this analysis, all untagged events were assumed to be at the charge side. The fractions of the NRSS in each face/energy bin can not be calculated the same way as for the interior detectors. A method detailed in Ref [151] was used to make corrections to the s factors for endcaps. We present a summary of the results below.

Let $n_{p/q}^{t/u}$ be the number of events on the phonon/charge side which would be

tagged/untagged (multiple/single) with good tagging efficiency, then the NRSS fractions for the two faces at a given energy bin are

$$f_p^{\mathrm{SS}} = \frac{n_p^u}{n_p^u + n_q^u + n_q^t} = \frac{n_p^u}{n_p^u + \frac{n_q^u}{\Gamma}} = \frac{1}{1 + \frac{f_q}{\Gamma f_p}}, \tag{C.18}$$

$$f_q^{\mathrm{SS}} = \frac{n_q^u + n_q^t}{n_p^u + n_q^u + n_q^t} = \frac{\frac{n_q^u}{\Gamma}}{n_p^u + \frac{n_q^u}{\Gamma}} = \frac{1}{1 + \frac{\Gamma f_p}{f_q}}, \tag{C.19}$$

where the would be q-side multiples n_q^t now appear as q-side singles for the q-side endcaps. f_p and f_q are the "raw" NRSS fractions for the p- and q-side, respectively. They are calculated using the observed counts:

$$f_p \equiv \frac{n_p^u}{n_p^u + n_q^u} \approx \frac{M_p}{M_p + S}, \quad f_q \equiv \frac{n_q^u}{n_p^u + n_q^u} \approx \frac{S}{M_p + S}, \tag{C.20}$$

where M_p is the number of phonon side multiples, and S is the rest and all assumed on the charge side. S contains the would be tagged charge side multiples and all the singles. Figure 18 in Ref [151] shows that the fraction of NR singles, single/(single + multiple), for an interior detector is around 0.1. So the fraction of singles in S may be small. This justifies the approximation made in Eq. (C.20). However, contamination of the (would be tagged) phonon side singles to S does introduce a systematic error which must be accounted for. The Γ in Eqs. (C.18) and (C.19) is the fraction of singles in the NR band:

$$\Gamma = n_q^u/(n_q^u + n_q^t).$$

The average value of Γ of interior detectors was used in calculating f_p^{SS} and f_q^{SS}. Averaging over interior detectors and extrapolating to endcaps introduces a

systematic error. The standard deviation of Γ across the interior detectors is assigned as the systematic error. The statistical error is a binomial error assuming all interior detectors combined into a single detector.

$$\Gamma = \left(\frac{S}{S+M}\right)_{\mathrm{NR}} = 0.15^{+0.015}_{-0.014}\,(\mathrm{stat.}) \pm 0.04\,(\mathrm{sys.}). \tag{C.21}$$

A second systematic error is due to the contamination of p-side singles to S. Average p-side singles fraction on interior detectors was assumed for endcaps. And the contamination was subtracted.

In the C58 analysis, because there were nonzero leakage on the endcaps, the above two corrections were both nonzero. However, zero leaked events were observed in the C58R classic timing analysis, so neither correction had an effect on the leakage.

C.5.3.2 Choice of prior

Zero leaked events were observed on the endcaps in the C58R classic timing analysis. Following Step 9 of the procedure described in Sec. C.4, three* events were assigned to each of the endcaps and further divided evenly into 6 face/energy bins. So the test case used to choose the prior had 0.5 leaked events per bin.

C.5.3.3 WS data

The leakage posteriors with priors from $c = -1$ to $c = -0.5$ in steps of 0.1 are shown in Fig. C.8. The posterior with $c = -1$ is sharply peaked at leakage equals zero and

*The observed pass/fail WS WBMS events are 0/398 for T3Z6 and 0/282 for T4Z6 (see Table C.4). The 90% upper limits on the number of leaked events are 2.3 for both detectors. 3 was chosen to make the it more like a real observation.

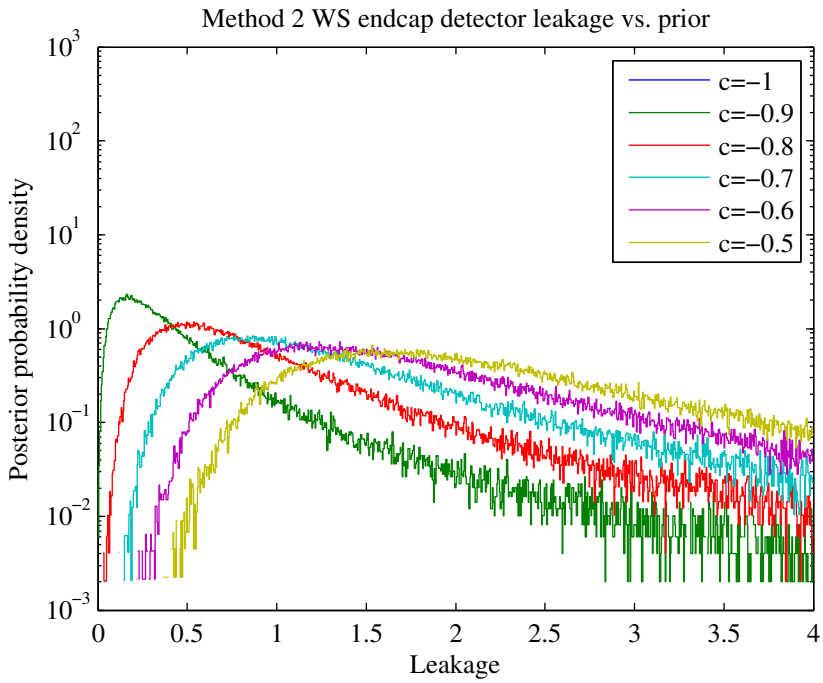
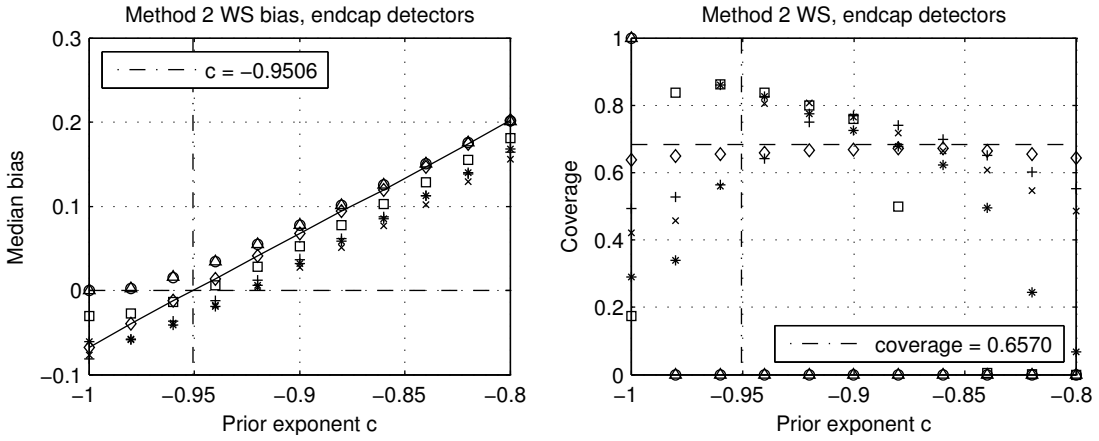


Figure C.8: Leakage posterior vs. prior for the endcap detectors in Method 2++ with timing cut pass/fail ratio measured on WIMP-search wide band beta multiples. Systematic errors are not included.



(a) Bias vs. prior exponent c

(b) Coverage vs. prior exponent c

Figure C.9: Leakage posterior bias and coverage for endcap detectors in Method 2++ with timing cut pass/fail ratio measured on WIMP-search wide band betas.

can not be seen on the figure.

Figure C.9 shows the bias and coverage for the test cases below (only for endcaps):

1. $\triangle$: All endcaps had the observed leakage, which is 0.
2. $\bigcirc$: All endcaps equal, 0.025 events were assigned if 0 observed. Total leakage n = 0.031 events.
3. $\square$: All endcaps equal, 0.05 events were assigned if 0 observed. Total leakage n = 0.061 events.
4. $\ast$: All endcaps equal, 0.1 events were assigned if 0 observed. Total leakage n = 0.12 events.
5. $\times$: All endcaps equal, 0.15 events were assigned if 0 observed. Total leakage n = 0.18 events.
6. $+$: All endcaps equal, 0.5 events were assigned if 0 observed. Total leakage n = 0.61 events.

The bias is minimized at $c = -0.9506$ and the corresponding coverage is 0.6570. The spread of bias at the chosen prior is $[-0.03,\ 0.02]$. A systematic error of 0.03 is assigned. Once the optimal prior was chosen, the actual leakage was then determined with the observed counts, which was

$$0.0242^{+0.0976}_{-0.0214}.$$

Combined with the two systematic errors, the final leakage is

$$n_2^{\mathrm{WSEC}} = 0.0242^{+0.0976}_{-0.0214}(\mathrm{stat.}) \pm 0.03\,(\mathrm{sys.}) \tag{C.22}$$

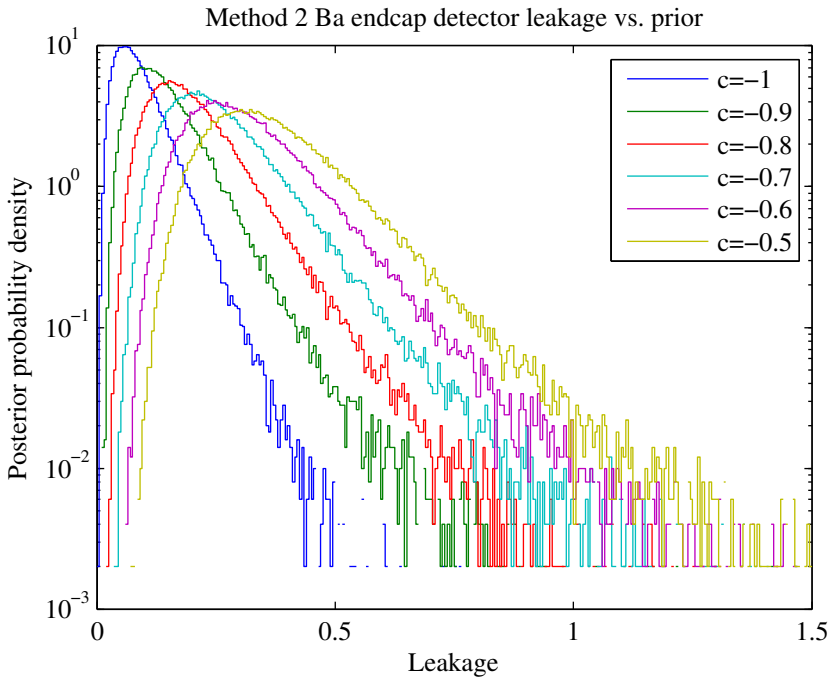
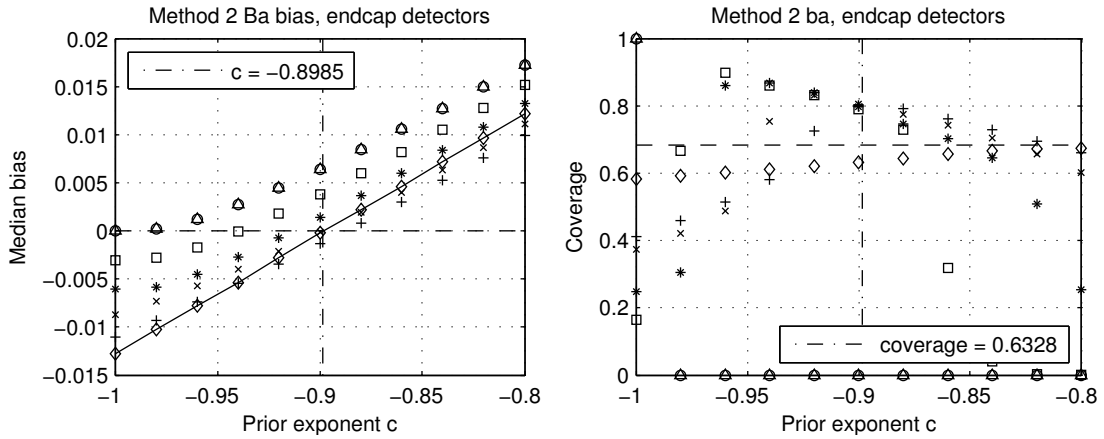


Figure C.10: Leakage posterior vs. prior for the endcap detectors in Method 2++ with timing cut pass/fail ratio measured on ^{133}Ba wide band beta multiples. Systematic errors are not included.

C.5.3.4 Ba data

The same test cases were chosen for the Ba data as for the WS data. Figure C.11 shows the bias and coverage. The case with 0.5 events assigned was chosen to estimate the systematic error. The bias is minimized at $c = -0.8985$ and the corresponding coverage is 0.6328. The spread of bias at the optimal prior is $[-0.00,\ 0.01]$ and a systematic error of 0.01 is assigned. The final leakage is

$$n_2^{\mathrm{WSBa}} = 0.001^{+0.016}_{-0.005}(\mathrm{stat.}) \pm 0.010\,(\mathrm{sys.}) \tag{C.23}$$



(a) Bias vs. prior exponent *c* (b) Coverage vs. prior exponent *c*

Figure C.11: Leakage posterior bias and coverage for endcap detectors in Method 2++ with timing cut pass/fail ratio measured on ^{133}Ba wide band betas.

C.5.4 Combined surface event background

Three methods (1, 2++, and 2++ Ba) were used to estimate the leakage on the interior detectors. At a given leakage, the three posteriors are independent estimates and the joint posterior was used as the final leakage on the interior detectors to reduce systematic errors. The leakage on the endcaps were treated similarly. Then the leakage posteriors for the interior and endcap detectors were randomly sampled and added together to get the total surface leakage.

Figure C.12 shows the final surface event leakage posteriors for the interior, endcap, and all detectors. Zero leakage were observed on the endcaps, the finite leakage on the figure are purely because of choice of the prior. The surface event

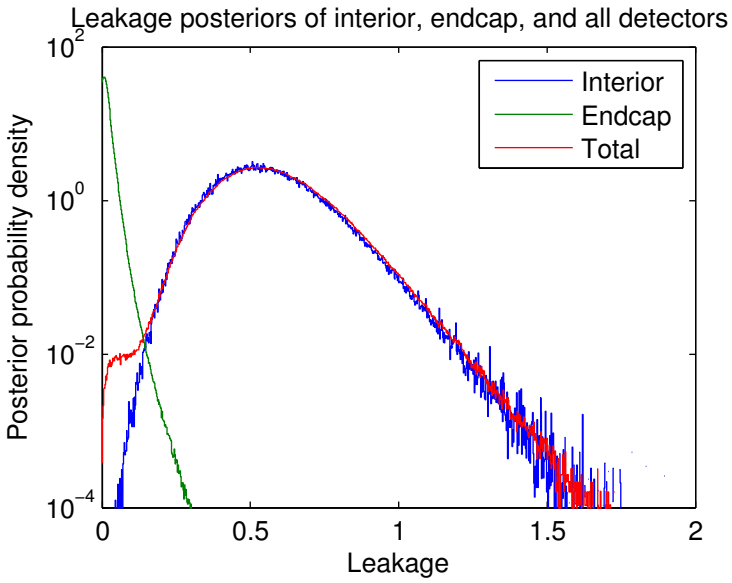


Figure C.12: Surface event leakage posteriors for the interior, endcap, and all detectors in the C58R classic timing analysis.

leakage estimates with all errors incorporated are

$$n^{\mathrm{Int}} = 0.53^{+0.16}_{-0.14} \tag{C.24}$$

$$n^{\mathrm{EC}} = 0.011^{+0.015}_{-0.009} \tag{C.25}$$

$$n^{\mathrm{Tot}} = 0.55^{+0.17}_{-0.14} \tag{C.26}$$

C.5.5 Data used in leakage estimation

Table C.2: WS NRSS failing the timing cut and WS NRMS passing and failing the timing cut for the C58R classic timing analysis.

Detector	N_z	b_z	B_z
T1Z2	5	0	36
T1Z5	11	1	58
T2Z3	9	0	51
T2Z5	9	0	77
T3Z2	7	0	50
T3Z4	12	1	53
T3Z5	4	0	31
T3Z6	54	0	29
T4Z2	12	2	49
T4Z4	7	4	34
T4Z5	14	1	48
T4Z6	12	0	50
T5Z4	9	0	80
T5Z5	12	1	53

Table C.3: C58R WIMP-search multiples inside the nuclear recoil band passing and failing the timing cut.

Det.	b_{ql}	b_{qm}	b_{qh}	b_{pl}	b_{pm}	b_{ph}	B_{ql}	B_{qm}	B_{qh}	B_{pl}	B_{pm}	B_{ph}
T1Z2	0	0	0	0	0	0	11	0	1	9	5	5
T1Z5	1	0	0	0	0	0	10	5	0	13	9	18
T2Z3	0	0	0	0	0	0	15	3	3	11	10	6
T2Z5	0	0	0	0	0	0	23	6	17	10	4	13
T3Z2	0	0	0	0	0	0	8	0	1	15	8	15
T3Z4	0	0	0	0	0	0	3	1	0	18	13	16
T3Z5	0	0	0	0	0	0	3	0	0	6	6	13
T3Z6	0	0	0	0	0	0	23	7	26	16	3	8
T4Z2	1	0	0	0	0	0	13	1	2	10	5	16
T4Z4	0	0	0	2	0	0	2	2	0	9	2	17
T4Z5	0	0	1	0	0	0	8	2	0	10	4	20
T4Z6	0	0	0	0	0	0	9	2	4	17	4	27
T5Z4	0	0	0	0	0	0	1	3	1	24	17	32
T5Z5	1	0	0	0	0	0	1	0	0	14	6	29

Table C.4: C58R WIMP-search wide beta band multiples passing and failing the timing cut.

Det.	m_{ql}	m_{qm}	m_{qh}	m_{pl}	m_{pm}	m_{ph}	M_{ql}	M_{qm}	M_{qh}	M_{pl}	M_{pm}	M_{ph}
T1Z2	0	0	0	0	0	0	20	30	43	15	19	22
T1Z5	0	0	0	0	0	0	43	39	80	24	18	44
T2Z3	0	0	0	0	0	0	4	11	43	2	10	39
T2Z5	0	0	0	0	0	0	24	33	68	9	13	36
T3Z2	0	1	2	0	0	0	7	9	18	12	12	35
T3Z4	1	0	0	0	0	0	26	37	51	28	20	70
T3Z5	0	0	0	0	0	0	3	12	47	5	12	28
T3Z6	0	0	0	0	0	0	63	44	215	22	17	37
T4Z2	0	1	0	0	0	1	10	15	65	7	11	58
T4Z4	0	0	0	0	0	0	11	13	30	9	10	60
T4Z5	1	0	0	0	0	0	7	9	38	6	4	57
T4Z6	0	0	0	0	0	0	25	20	138	8	15	76
T5Z4	0	0	1	0	0	0	7	14	58	19	18	80
T5Z5	0	0	0	0	0	0	3	6	48	3	12	62

Table C.5: C58R ^{133}Ba wide beta band multiples passing and failing the timing cut.

Det.	m_{ql}	m_{qm}	m_{qh}	m_{pl}	m_{pm}	m_{ph}	M_{ql}	M_{qm}	M_{qh}	M_{pl}	M_{pm}	M_{ph}
T1Z2	1	0	0	0	0	0	114	94	410	58	55	242
T1Z5	1	0	0	0	0	0	131	128	635	103	93	497
T2Z3	0	0	0	0	0	1	26	70	422	52	122	772
T2Z5	1	0	1	1	0	0	90	154	747	74	116	723
T3Z2	2	0	3	0	1	1	54	91	520	42	69	512
T3Z4	2	0	1	4	0	1	237	275	1364	373	428	2546
T3Z5	0	0	0	0	0	2	40	131	903	53	192	1108
T3Z6	0	0	0	0	0	0	617	722	3649	253	200	590
T4Z2	3	2	0	1	2	0	96	203	794	171	269	1632
T4Z4	1	0	0	2	0	1	52	140	599	169	234	1722
T4Z5	0	0	0	0	1	0	40	82	627	102	204	1182
T4Z6	0	0	0	0	0	0	114	232	1377	124	225	1216
T5Z4	4	0	8	0	2	1	129	246	1592	217	318	1839
T5Z5	1	0	0	2	1	1	25	75	966	138	316	2182