Dark Matter Physics with P-type Point-contact Germanium Detectors: Extending the Physics Reach of the MAJORANA Experiment

Michael G. Marino

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

University of Washington

2010

Program Authorized to Offer Degree: Physics

University of Washington
Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

Michael G. Marino

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Chair of the Supervisory Committee:

_____

John Wilkerson

Reading Committee:

_____

Hamish Robertson

_____

Leslie Rosenberg

_____

John Wilkerson

Date: _____

University of Washington

Abstract

Dark Matter Physics with P-type Point-contact Germanium Detectors: Extending the Physics Reach of the Majorana Experiment

Michael G. Marino

Chair of the Supervisory Committee:
Professor John Wilkerson
Physics

P-type point-contact (P-PC) germanium detectors present an exciting detector technology, yielding sub-keV thresholds and intrinsically low electronic noise. Characteristics of the detectors enhance their background-rejection capabilities for experiments searching for neutrinoless double-beta decay in ^{76}Ge and, as such, the Majorana experiment will deploy a Demonstrator module with arrayed P-PCs. In addition, these same qualities make the detectors sensitive to direct dark matter detection. The consecutive deployment of two P-PC detectors underground at Soudan Underground Laboratory is presented, providing results and conclusions about low-energy backgrounds and data-acquisition requirements at low energies. Data from the lower-background detector is used to generate limits on the spin-independent nuclear recoil of low-mass ($\lesssim 10$ GeV) WIMPs as well as on the strength of the axion-electron coupling. Finally, a contextual discussion of these results is given, focusing on estimating the sensitivity of the Majorana Demonstrator to detect dark matter.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

$0\nu\beta\beta$: Neutrinoless double-beta decay.

BEGe: Broad-energy germanium detector. Also refers to a low-background P-PC detector deployed underground at Soudan Underground Laboratory.

DAQ: Data acquisition.

P-PC: P-type point-contact germanium detector.

P-PC2: P-PC detector deployed underground at Soudan Underground Laboratory.

WIMP: Weakly-Interacting Massive Particle

ACKNOWLEDGMENTS

There are but few moments in one's life when one can truly ponder the road that has lead to that point. Looking back one can choose to wonder at the failures narrowly averted, but how much more meaningful to realize the gracious helping hands that made possible the way. I devote this section to the latter, in the hopes of conveying my gratitude to those from whom I benefited so greatly.

My advisor and mentor, John Wilkerson, I thank for the support, the encouragement, the living example, and the sometimes prodding it took to help mold me into a scientist. The experience as his student has given me invaluable resources from which to draw as I embark upon my career. I look forward to any collaborative opportunity the future might hold.

From the University of Notre Dame, I must thank the late Larry Lamm for introducing me to the nitty-gritty of experimental physics and for helping me to realize I could contribute to the field. My career as a physicist would have likely ended after my undergraduate studies had it not been for Alejandro Garcia, who fostered me through the earliest moments of my research career in his previous post at Notre Dame.

Juan Collar provided the brilliance, many ideas, and passion which made most of my thesis work possible, and for that I am grateful. Working with Juan there is truly never a dull moment and I hope to be able to inject such enthusiasm and skill into my future work. As well I must thank the staff at Soudan Underground Laboratory, including Jim Beaty, Dave Saranen, Jerry Meier, and Curt Lerol among others who made working underground a real joy.

Hamish Robertson I thank for the many insightful comments throughout my time at CENPA that helped focus my research. To Wick Haxton I am grateful especially for overseeing my foray into Nuclear Matrix Elements. A thank you to Leslie Rosenberg for discus-

sions on Axions and their relevance to my work. I must thank Michael Miller for being an advisor to matters-not-only-physics and for being a friend. Among many other things, he taught me that sometimes the most valuable thing to do to ensure research progress is to go skiing.

The staff at CENPA has been so instrumental in supporting my time as a graduate student. I thank especially the front office, Victoria Clarkson and Kate Higgins, for tackling the inevitable administrative issues I had and for helping my life at the lab run so smoothly. Dick Seymour was always eager to help me solve a hard-to-find hardware or software issue. His replacement, Gary Holman, has already demonstrated his great value to the lab and I am disappointed to have overlapped with him for only so short a time. To Mark Howe I am indebted for the always interesting discussions on hardware, software, or whatever, though I greatly missed their regularity after his move to North Carolina. To others at CENPA, I am grateful for the collegial atmosphere they generated that made it truly a special place.

Jason Detwiler, my first postdoc mentor, I cannot thank enough for providing endless resources and guidance. His selflessness in helping others and me in particular speaks to his character, skill and ability, and his example challenges me to do likewise as I transition to a postdoctoral position.

I am so lucky to have shared much of the past six years with my officemates, Alexis Schubert and Rob Johnson. It is difficult to summarize the office rapport, but I will not forget the help and frank discussions that formed much-anticipated interruptions to the work day. I know that Alexis will continue crafting beautiful work and that Rob's memory will keep rivaling that of an elephant. A special thank you to the rest of the EWI group.

I thank the MAJORANA collaboration, especially Reyco Henning, John Orrell, and David Radford for aiding my development as a physicist by casting a critical eye to my work.

For my friends with whom I lived, played soccer, enjoyed bike rides, played late night board games and took advantage of the outdoors (they know who they are): I am grateful for the necessary balance that they brought to my life.

Finally, I must thank my family without whom this would not have been possible: to

my grandparents for giving me the gift of education and to my parents who taught me the meaning of values and the value of meaning. And to Verena, meine Österreicherin, I thank not only for having been on this road with me, but also for being my companion as I walk ahead.

DEDICATION

To my family

Chapter 1

INTRODUCTION

The MAJORANA experiment experiment will search for the rare process of neutrinoless double-beta decay ($0\nu\beta\beta$) to determine characteristics of the neutrino. The choice of detector technology, p-type point-contact germanium detectors, will also allow the experiment to search for dark matter. The following discusses the reasons behind using this technology with a focus on the additional physics questions beyond those related to neutrinos which may be answered with these detectors.

1.1 Neutrinos and Neutrinoless Double-beta Decay ($0\nu\beta\beta$)

Neutrinos have provided evidence for physics beyond the Standard Model: neutrino oscillation experiments have shown that neutrinos are massive but cannot determine the absolute neutrino mass scale or the nature of the neutrino mass (see, e.g. [1]). Neutrinoless double-beta decay ($0\nu\beta\beta$) is a process which, if discovered, would imply that the neutrino is its own antiparticle (a so-called 'Majorana' particle) and could help define the neutrino mass scale. $0\nu\beta\beta$ could occur in several even-even nuclei (e.g. ^{48}Ca, ^{76}Ge, ^{130}Te, ^{136}Xe, etc.) otherwise stable against single-beta decay and is characterized by the emission of two electrons with total kinetic energy equal to the Q-value of the reaction, $Q_{0\nu\beta\beta}$. For example:

$$(Z, A) \rightarrow (Z + 2, A) + e^- + e^- \tag{1.1}$$

where the Q-value is the mass difference $M(Z, A) - M(Z+2, A) - 2m_e$. The signal, therefore, is the deposition of energy in a detector equal to $Q_{0\nu\beta\beta}$. A schematic of a possible lepton-number-violating mechanism for neutrinoless double-beta decay is shown in Figure 1.1. Assuming left-handed currents and that the decay is dominated by the exchange of a light massive Majorana particle, the half-life of $0\nu\beta\beta$ can be related to the effective mass of the neutrino by:

$$\left(T_{1/2}^{0\nu}\right)^{-1} = G^{0\nu}|M^{0\nu}|^2\langle m_{\nu_{\beta\beta}}\rangle^2 \tag{1.2}$$

where $G^{0\nu}$ refers to an exactly calculable phase-space integral, $|M^{0\nu}|^2$ is a nuclear matrix element, and $\langle m_{\nu_{\beta\beta}}\rangle$ is the effective neutrino mass. ^{76}Ge currently holds the best limit for the $0\nu\beta\beta$ half-life: $T_{1/2}^{0\nu} \geq 1.6 \times 10^{25}$ [2]. Several review articles exist outlining the current theoretical and experimental landscape (see e.g. [3, 4]). Two-neutrino double-beta decay ($2\nu\beta\beta$) is a related decay that can exist in the same nuclei allowed to undergo $0\nu\beta\beta$:

$$(Z, A) \rightarrow (Z + 2, A) + e^- + e^- + \bar{\nu}_e + \bar{\nu}_e \tag{1.3}$$

The $2\nu\beta\beta$ process has been seen in several nuclei including ^{76}Ge and has a half-life for these nuclei around $T_{1/2}^{2\nu} \sim 10^{20}$ yrs.

1.2 The MAJORANA Experiment

The search for a rare process such as $0\nu\beta\beta$ necessarily involves maximizing the magnitude of the expected signal while simultaneously reducing backgrounds that may mimic the sought-after process. The MAJORANA experiment proposes to search for $0\nu\beta\beta$ in ^{76}Ge using high-purity germanium as both source and detector, thereby maximizing a possible signal from $0\nu\beta\beta$. The first stage of this experiment will involve the deployment of 20-40 kilograms of germanium in an arrayed fashion in a DEMONSTRATOR module with the goal to determine the feasibility of scaling up to the 1-tonne scale. To achieve this, the MAJORANA experiment seeks to demonstrate less than 1 background count per year per tonne of enriched material in the region-of-interest, a $\sim$ 4 keV window around the $\beta\beta$-decay Q-value of ^{76}Ge (2039 keV). Low-capacitance, low-noise p-type point-contact (P-PC) detectors will be deployed in these modules to take advantage of characteristics which make them beneficial for rare-process searches in general and for looking for $0\nu\beta\beta$ in particular. These characteristics and other details about P-PC detectors will be discussed in Section 1.3.

To achieve its background goals, the MAJORANA experiment will employ standard techniques for background reduction including: creation of detector mounts, cryostats, and other components close to the detectors using ultra-clean electroformed copper and other radiop-

(a) $2\nu\beta\beta$

(b) $0\nu\beta\beta$
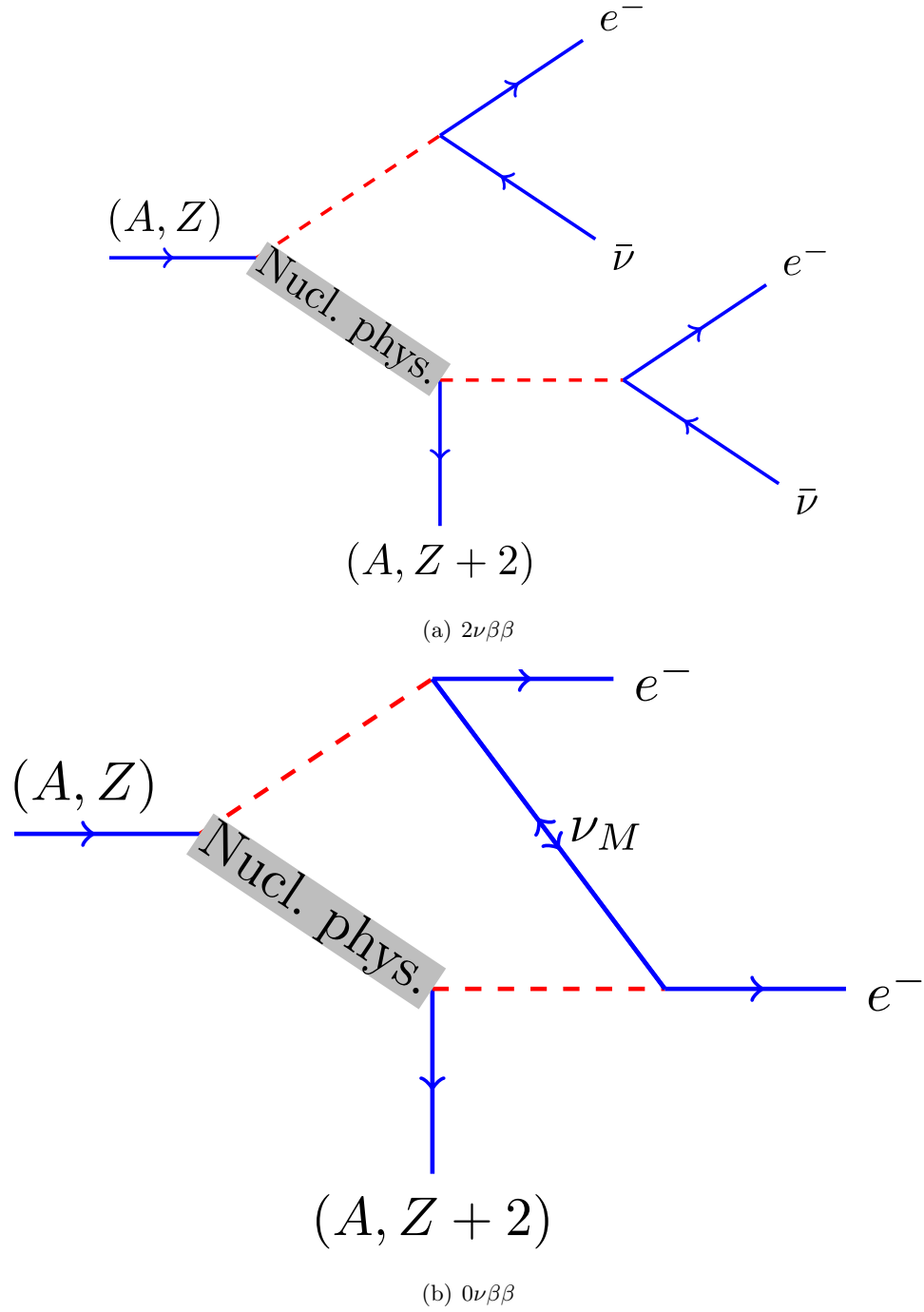
Figure 1.1: Schematic figure of the two types of double-beta decay, the 2ν mode and the 0ν mode. In this schematic, the 0ν mode is mediated through the exchange of a virtual Majorana neutrino, ν_M.

ure materials; use of passive shielding against external radiation including a lead shield for gamma radiation and a borated polyethylene shield for the moderation of cosmic-ray-induced neutrons; use of active shielding (vetos) against cosmic-ray muons; and deployment of the module underground in the Sanford Underground Lab at Homestake, South Dakota, for moderation of cosmic-ray muons. Engineering drawings given in Figures 1.3 and 1.2 show both the cryostat design and the expected shield construction, and indicate the modular design of the experiment; additional, independent cryostats may be deployed within the same shield geometry. The expected sensitivity of this first stage, assuming 3 years with 30 kg of enriched material or 90 kg-yr of ^{76}Ge exposure, is $T_{1/2} \geq 10^{26}$.

1.3   P-type Point-contact (P-PC) Germanium Detectors

P-type point-contact (P-PC) germanium detectors are an exciting detector technology whose characteristics provide powerful tools in the search for $0\nu\beta\beta$ and enable searches for other rare physics processes at low energy, e.g. dark matter. The electrode geometry of P-PC detectors significantly reduces the capacitance, reducing the energy threshold and enhancing the detector's ability to detect low-energy ($O(100 \text{ eV})$) processes. Figure 1.4 shows a picture of the geometry of this detector in comparison to a standard semi-coaxial crystal. In addition to improving the electronic response of the detector, the crystal geometry also yields a weighting field strongly peaked at the readout electrode (see e.g. Figure 1 in [6]). This means that as charge drifts to the p contact after being created from energy deposition in the crystal (e.g. from a physics interaction), no signal will be induced in the contact until the drifting charge is very near (within ~ 1 cm of) the contact. Additionally, charge drift times are increased by the longer drift paths. These two characteristics coupled together improve the ability to distinguish between charge originating from one point in the crystal or from multiple sites in the detector.

An n-type detector with a point-contact geometry was developed by Luke et al. in 1989, demonstrating detector capacitance on the order of 1 pF [7]. With such a low capacitance and therefore a low-energy threshold this detector was seen as a potential tool for dark matter detection, but this detector exhibited poor energy resolution due to incomplete charge collection attributed to charge-trapping effects. In 2007, Barbeau et al. presented a
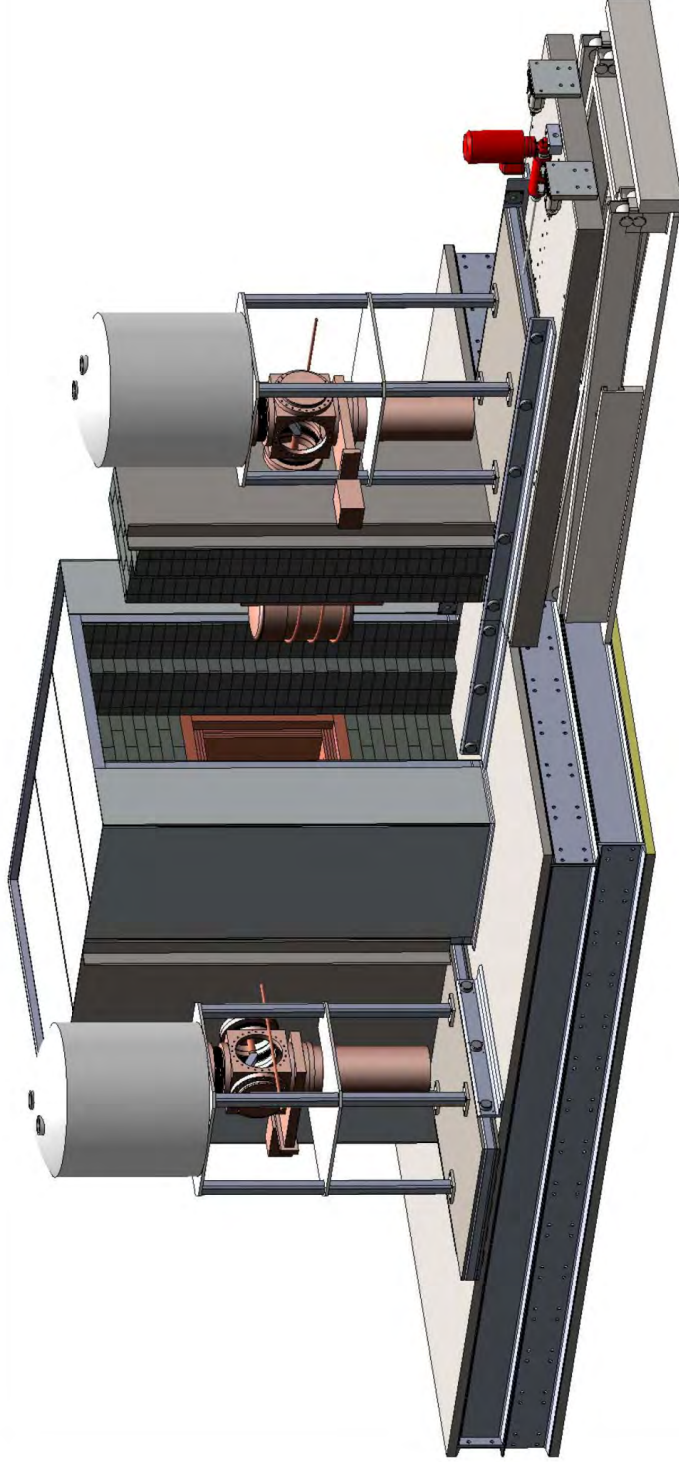
Figure 1.2: MAJORANA DEMONSTRATOR shield geometry. The modular design of the shield will enable a phased deployment of cryostats, allowing sets of detectors to be easily added after initial commissioning of the experiment. From the MAJORANA collaboration [5].
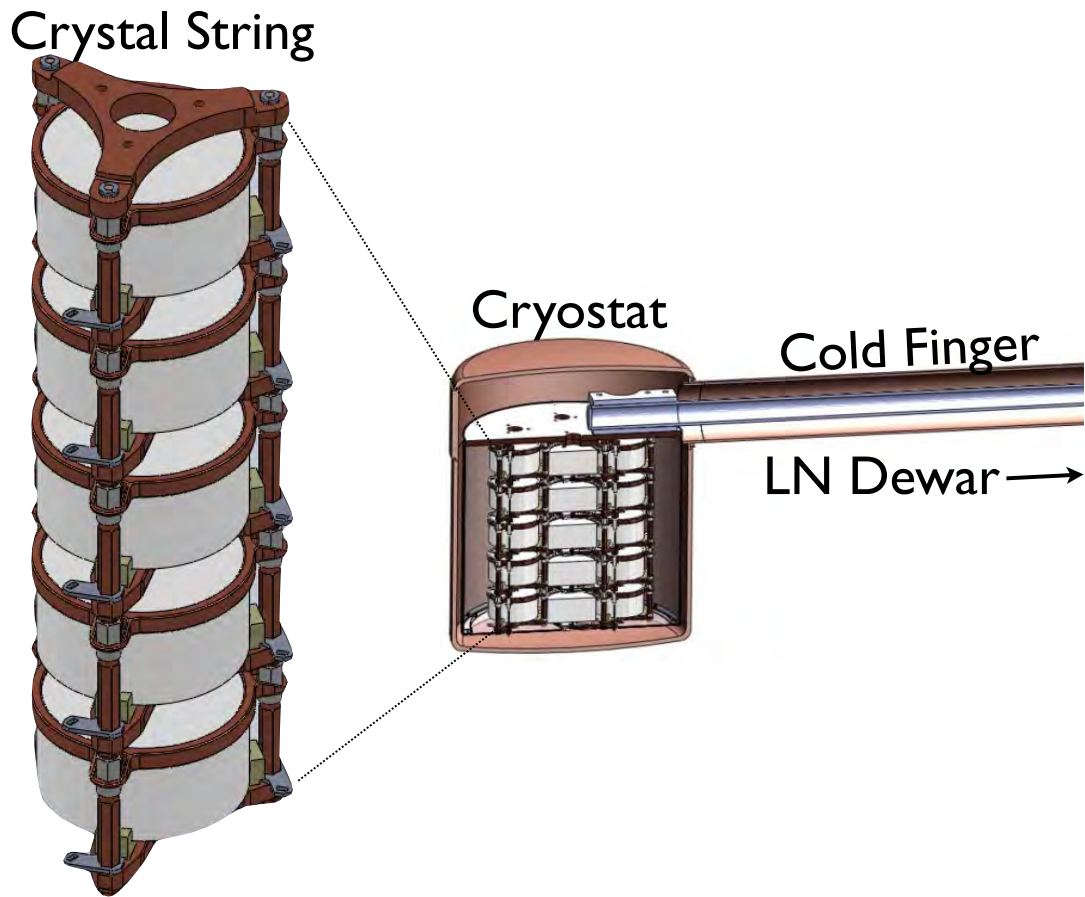
Figure 1.3: MAJORANA DEMONSTRATOR cryostat and crystal string geometry, each crystal string houses 5 p-type point-contact germanium detectors. From the MAJORANA collaboration [5].

(a) Point-contact geometry.   (b) Semi-coaxial geometry.
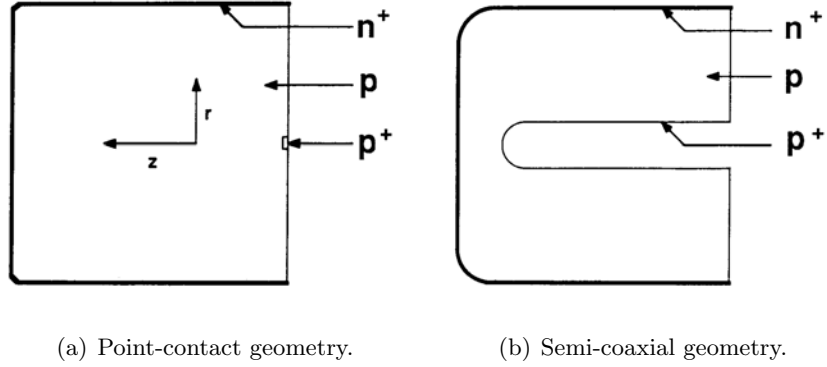
Figure 1.4: The left figure depicts the point-contact geometry, the right figure a conventional semi-coaxial geometry. Figure adapted from [7]. The small diameter of the p contact in the P-PC detector reduces the capacitance of the detector improving the intrinsic noise characteristics.

new detector with a point-contact geometry, departing from previous convention by creating it out of p-type material [8]. The reason for this change was to take advantage of the reduced sensitivity of p-type crystals to electron trapping in germanium. Because p-type crystals collect holes instead of electrons at the p contact of the crystal, they are less likely to see a degradation of signal from electron trapping [9]. With these changes, Barbeau et al. demonstrated a resolution comparable to conventional semi-coaxial germanium detectors and an energy threshold of 330 eV making them an excellent candidate for double-beta decay searches. Other work with these detectors has supported these conclusions (see e.g. [10, 11]).

1.4 P-PC Detectors for the MAJORANA Experiment

For the proposed MAJORANA experiment, the characteristics of the P-PC detector will aid in the reduction of backgrounds and in enhancing the physics reach of the experiment. Since $0\nu\beta\beta$ is an inherently single-site event, cutting out multi-site events during data analysis provides a powerful background reduction tool. For conventional semi-coaxial detectors, various tools have been developed to tag multi-site events based upon the shape of the measured pulse (see, e.g. [12]). The expansion in time of charge collection in P-PC detec-

tors makes distinguishing multi-site events easier and improves the efficiency for single-site acceptance vs. multi-site rejection of a pulse-shape analysis routine (see [13, 6]).

To improve the bandwidth of the detector readout it is necessary to place front-end preamplification electronics (e.g. field-effect transistors) nearby the detectors inside the detector cryostat. As with any low-background experiment the increase of material (and especially material close to sensitive detectors) introduces potential radioactivity that could increase background. The single-contact nature of the P-PC requires only one front-end per crystal, thus reducing the material inside the cryostat and possibly softening radiopurity requirements.

Cosmogenically-produced ^{68}Ge is a significant background source for the MAJORANA experiment, decaying first via electron capture to ^{68}Ga ($T_{1/2} = 271$ d) and then to ^{68}Zn (Q=2921 keV, $T_{1/2} = 68$ m). The initial decay is characterized by the emission of Auger electrons and/or soft x-rays summing to the orbital energy of the captured electron (e.g. 1.3 keV for L-capture, 10.3 keV for K-capture) and it is possible to use these to tag the ^{68}Ge decay. A time cut can then be introduced to veto events occurring within a few ^{68}Ga lifetimes to mitigate background from that decay without seriously affecting detector live time. P-PC detectors have an improved sensitivity to low-energy physics because of their low noise and low-energy threshold and so should significantly enhance the tagging of the ^{68}Ge decay.

Though characteristics of P-PC detectors will prove useful in searching for $0\nu\beta\beta$, the low-energy threshold will also expand the physics reach of the MAJORANA experiment. Access to low energies will make the DEMONSTRATOR sensitive to low-energy nuclear recoils and enhance its capacity as a dark matter detector. The functionality and usefulness of the P-PC detector as a tool for dark matter searches is considered in much more detail throughout the remainder of this dissertation.

1.5 Searching for Dark Matter: Low-energy Physics with P-PC Detectors

1.5.1 Evidence for Dark Matter

From cosmological observation there exists significant evidence that the matter density of the universe is mostly composed of non-luminous, gravitationally-interacting particles referred to as 'dark matter'. Perhaps the most convincing and intuitive empirical indications arise from measurements of galactic rotational curves. Astronomers have measured the rotational velocity of stars in galaxies and determined the relation of this velocity versus the radial distance from galactic center. When the rotational velocity is compared to that expected given the observed radial distribution of mass, it is found that galaxies are rotating much more quickly than they should be. Essentially, there is simply not enough observable mass to explain why the galaxies' rotational velocities do not tear them apart. This realization indicated that some other mass must exist to hold the galaxies together, and that this mass was hidden due to its weak or nonexistent coupling with photons. Such a weak interaction with electromagnetism would make the dark matter unsusceptible to frictional forces and also would make it non-luminous or 'dark.' By including a mass density of dark matter in a 'dark matter halo' around galaxies, astronomers could successfully reproduce observed rotation curves. An example of such a rotational curve and a three-component fit (visible matter (e.g. stars), gas, and dark halo) has been included from Reference [14] in Figure 1.5.

Other evidence of dark matter comes from measurements of large-scale structure through the analysis of the anisotropy of the cosmic microwave background (CMB) (see e.g. [15]). The extraction of cosmological parameters (e.g. total matter density, baryon density, Hubble parameter, etc.) occurs by fitting the CMB data using model-based assumptions. In particular, the parameters-of-interest are the total matter density, Ω_m, and the baryon density, Ω_b, which have been measured to be (in units of critical density):

$$\Omega_m = 0.256 \pm 0.02$$

$$\Omega_b = 0.044 \pm 0.003 \ [15]$$

If baryonic matter composed the majority of the mass in the universe, one would expect these two quantities to be equivalent. Instead, the significant discrepancy between the two
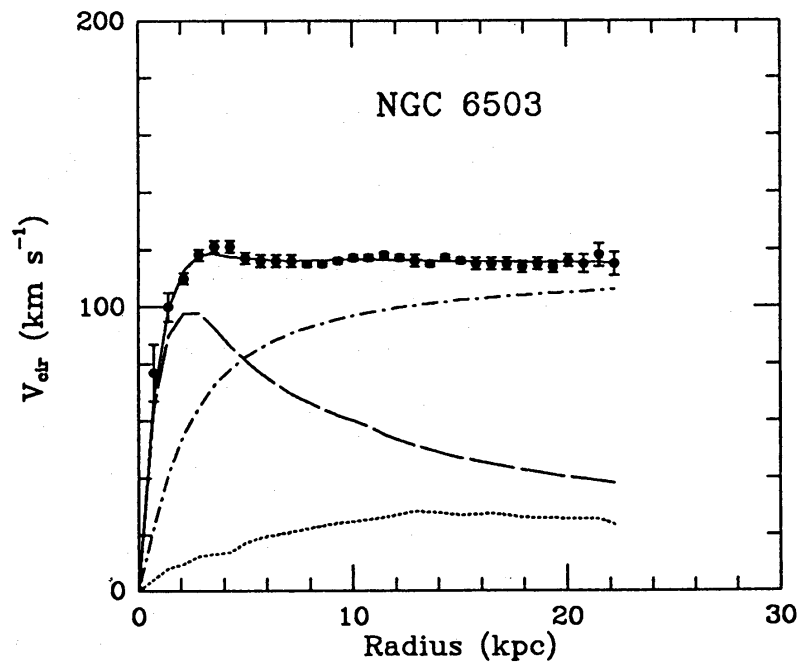
Figure 1.5: Galactic rotational velocity curve of NGC 6503 from [14]. The solid line is a three-component fit to the data with components from: visible mass (dashed), gas (dotted) and the dark halo (dash-dot).

indicates the presence of unaccounted for material. Therefore, the dark matter component is the difference between these two components: $\Omega_{dm} = \Omega_m - \Omega_b = 0.21 \pm 0.02$, indicating a large proportion of dark matter in the universe on the order of 20%.

Perhaps the most visually stunning evidence for dark matter comes from the 'bullet cluster' (galactic cluster 1E 0657-56) [16]. In this particular example two subclusters have collided with one another ~100 Myr ago, subjecting the contents of each to frictional forces. The stars and galactic components of the subclusters were scarcely affected due to their relatively wide spacing, passing by one another and interacting largely only gravitationally. However, the trajectories of the hot gas making up the majority of the subclusters' baryonic mass density were significantly altered. The matter density of the gas component was measured using images from the Chandra x-ray observatory and the total mass density of the cluster was measured using gravitational lensing. When plots of these measurements are superimposed (see Figure 1.6 from [16]), it is clear that the two centers (of each subcluster) of the total mass density are offset from the centers of the gas-plasma mass density. This indicates the presence of non-luminous (dark) matter that interacts weakly with both the normal baryonic matter and with itself.

1.5.2 Dark Matter Candidates

Whereas there are many particle candidates for dark matter, they share several general qualities:

1. Stable, having lifetimes at least on the order of the age of the universe. Otherwise, dark matter would have decayed away.

2. Poor or no interaction with electromagnetism, making them non-luminous or 'dark'.

3. Non-relativistic at the time of galaxy formation: 'hot' dark matter would move too quickly to clump, not allowing it to affect large-scale structure formation as required by cosmological observation.

Additionally, dark matter will compose a halo in our galaxy which the earth passes through as it transits around the sun. Therefore, depending on the velocity dependence of
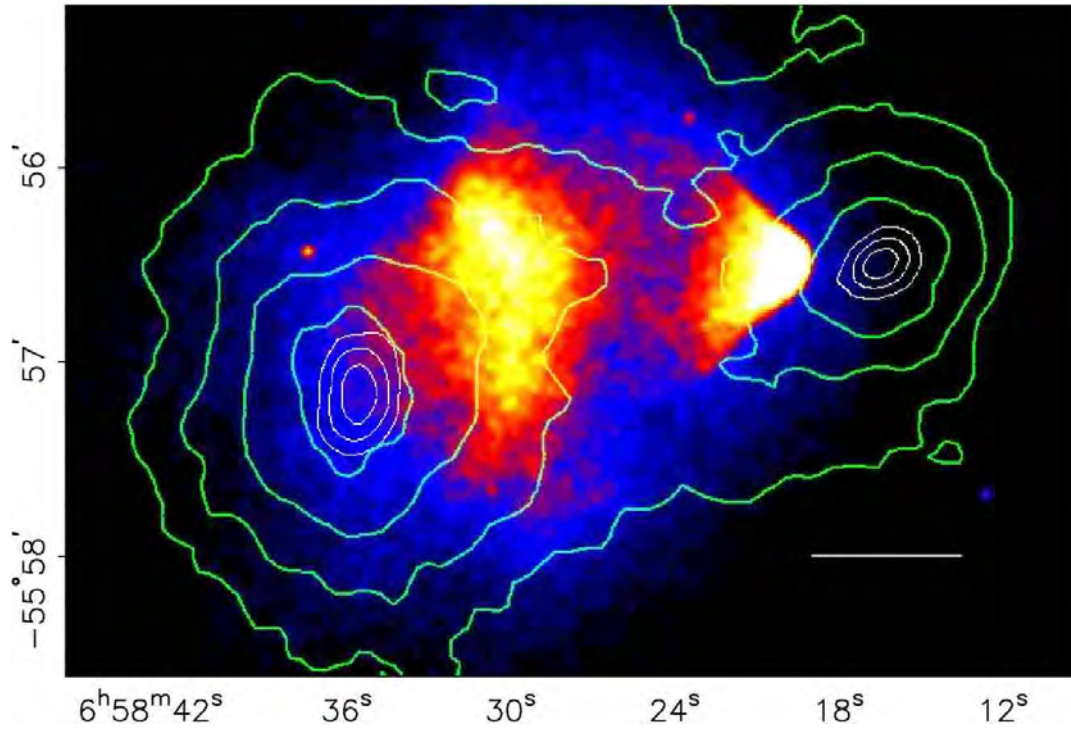
Figure 1.6: Mass and x-ray image of galactic cluster 1E 0657-56, the 'bullet cluster'. The total mass, denoted by contour lines, has been measured using gravitational lensing; the density of the gas, which composes the majority of the baryonic matter density in the cluster, has been imaged using the Chandra x-ray observatory. Figure taken from reference [16].

the dark matter interaction and the velocity distribution of WIMPs in the galactic halo, it is possible that the earth's orbit can generate an annual modulation in the dark matter signal, providing an additional distinguishing characteristic. In the following, we will consider two potential species of dark matter in particular: WIMPs and axions.

WIMPs

WIMPs, or Weakly-Interacting Massive Particles, (denoted from here on by χ) are more of a class of candidate particles than one in particular, defined by the fact that these types of particles predominantly (or exclusively) interact via the weak force. Several reviews outline in more detail the characteristics of WIMP particles, see e.g. [17, 18]. An argument for the existence of WIMPs relates to how they can 'naturally' explain the dark matter relic abundance. For example, during the period right after the big bang as the temperature of the universe was still quite hot, WIMPs would have been in equilibrium with other fermions, undergoing annihilation and generation according to the two-way reaction: $\chi\bar{\chi} \leftrightarrow f\bar{f}$ (see e.g. Figure 1.7(a)). As the temperature of the universe cooled below the mass of χ, eventually this two-way process would shut off in one direction allowing only the interaction $\chi\bar{\chi} \rightarrow f\bar{f}$ to remain. At this point in time, the rate of decay of χ has been calculated to be $\Gamma_\chi = \langle \sigma_A v \rangle^{-1} n_\chi$ [17], where σ_A is the average cross-section for χ to annihilate into lighter fermions, v is the average particle velocity, and n_χ is the density of WIMP particles. The expansion of the universe would then reduce the density of χs, effectively shutting off this annihilation process and 'freezing out' a so-called relic abundance of the particles.

An estimate of the density of this abundance remaining today times the Hubble constant squared, $h^2 \sim 0.5$, can be calculated (see, e.g. [17]):

$$\Omega_\chi h^2 = (2.6 \times 10^{-10} \text{ GeV}^{-2}) \langle \sigma_A v \rangle^{-1} \tag{1.4}$$

$$\tag{1.5}$$

and if we assume that this interaction follows the weak force, we can estimate the average

cross section times the velocity as:

$$\langle \sigma_A v \rangle^{-1} \quad \sim \quad \frac{\alpha^2}{M_{\text{weak}}^2} \sim 10^{-9} \text{ GeV}^{-2} \tag{1.6}$$

$$\tag{1.7}$$

which leaves us with an estimate of $\Omega_\chi h^2 \sim 0.2$, close to the observed cold-dark-matter density. This realization has not been interpreted as a 'coincidence' and instead is suggested as compelling support for the possibility of WIMP existence [19].

The direct detection of WIMPs is possible via their elastic scattering off fermions, see e.g. Figure 1.7(b). In creating a detector, we are of course generally limited to composing it out of electrons, neutrons, and protons. In principle, WIMPs could scatter off any fermion sharing a coupling to a mediating boson, but the kinematics of such a non-relativistic recoil maximize the energy transfer when the target particle has mass comparable to that of the incident particle. In the case of WIMPs with hypothesized masses $\sim 1 \rightarrow 1000$ GeV, an energy transfer is maximized during coherent recoil off a nucleus having similar or equal mass. Additionally, WIMP scattering off bound electrons is suppressed due the mass ratio m_e/m_χ, as well as the small size of the electronic wave function (see e.g. [20]) making the nuclear recoil the expected primary method of detection. The standard form of the energy spectrum for a WIMP-induced nuclear recoil is outlined later in Section 4.1.

Axions

Axions are light, pseudoscalar bosons that were originally suggested as a mechanism to solve the CP problem of Quantum Chromodynamics (QCD) [21]. In particular, the QCD Lagrangian includes a CP-violating term:

$$\mathcal{L}_{\bar{\Theta}} = \bar{\Theta} \frac{\alpha_s}{8\pi} G^{\mu\nu a} \tilde{G}_{\mu\nu}^a$$

where $\bar{\Theta}$ is an angle with possible value between $-\pi$ and π. The lack of experimental evidence for CP violation in QCD suggests that this value is very small or 0, an 'unnatural' solution given its range of possible values. The Peccei-Quinn mechanism was proposed [21] to explain this issue by adding an additional $U(1)_{PQ}$ symmetry to the Lagrangian. The

(a) WIMP annihilation and production   (b) WIMP elastic scattering
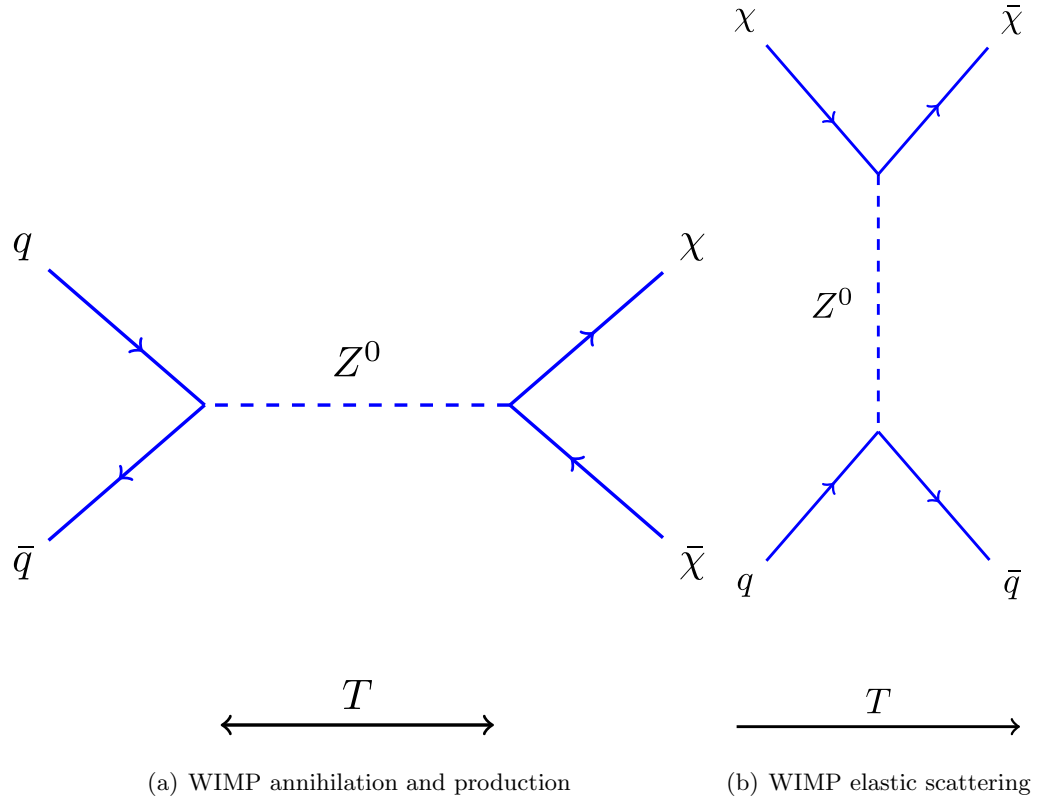
Figure 1.7: Example WIMP-quark interactions mediated by the exchange of a Z^0 boson with the direction of time, t, indicated for each at bottom. The left figure is applicable for WIMPs undergoing annihilation and generation during the equilibrium phase at the beginning of the universe, when the temperature of the universe was much greater than the mass of χ. The right figure denotes a mathematically equivalent interaction, an elastic-scattering process of χs off quarks becoming relevant as the temperature of the universe reduced below the mass of χ.

breaking of this symmetry results in a Nambu-Goldstone boson – the axion – updating the CP-violating portion of the QCD Lagrangian to be:

$$\mathcal{L}_{\bar{\Theta}} = \left(\bar{\Theta} - \frac{\phi_A}{f_A} \right) \frac{\alpha_s}{8\pi} G^{\mu\nu a} \tilde{G}^a_{\mu\nu}$$

with ϕ_A the axion field and f_A the axion decay constant [15]. The expectation value of the axion is found to be $\phi_A = f_A \bar{\Theta}$ thereby directly canceling the CP violating term. Further detail regarding axions can be found in several current reviews, including [15, 22].

Bounds on axion properties can derive from cosmological observations as well as direct experimental searches. For example, cosmological data can constrain axion models by looking for the evidence of additional unexplained energy loss in stars (see, e.g. [23, 24]). Direct detection measurements may take advantage of axions coupling with 2 photons, looking for their interaction in high-intensity magnetic fields, as is the case with the CAST experiment [25] searching for axions produced in the sun, or searching for an axion interaction in a resonant microwave cavity within a magnetic field, as does the ADMX experiment [26]. Germanium detectors are not ideal for searching for such a signal, but can put bounds on the strength of the axion's coupling with electrons by searching for the inelastic scattering of an axion off an electron. This interaction and the observed signal is discussed in more detail in Section 5.1.1.

1.5.3   P-PC Detectors and Dark Matter

P-PC detectors can provide excellent dark matter detectors due to their low-energy threshold ($O(100$ eV$)$) and excellent resolution down to threshold. This can be made more clear if we consider the signal that WIMPs and axions can leave in a germanium detector. These signals and their exact spectral form are discussed in more detail later for WIMPs (see Section 4.1) and axions (see Section 5.1.1), but we consider their general properties here. For a more complete reference on the direct detection of dark matter, see [27].

The nuclear recoil signal from a WIMP is essentially described as an exponential in energy space truncated by the finite escape velocity of the WIMP from the halo (see, e.g. [17, 18]). The exponential constant is determined by the mass of the WIMP and the spectrum becomes steeper as the mass gets smaller. This can be visualized in Figure 1.8. It is clear
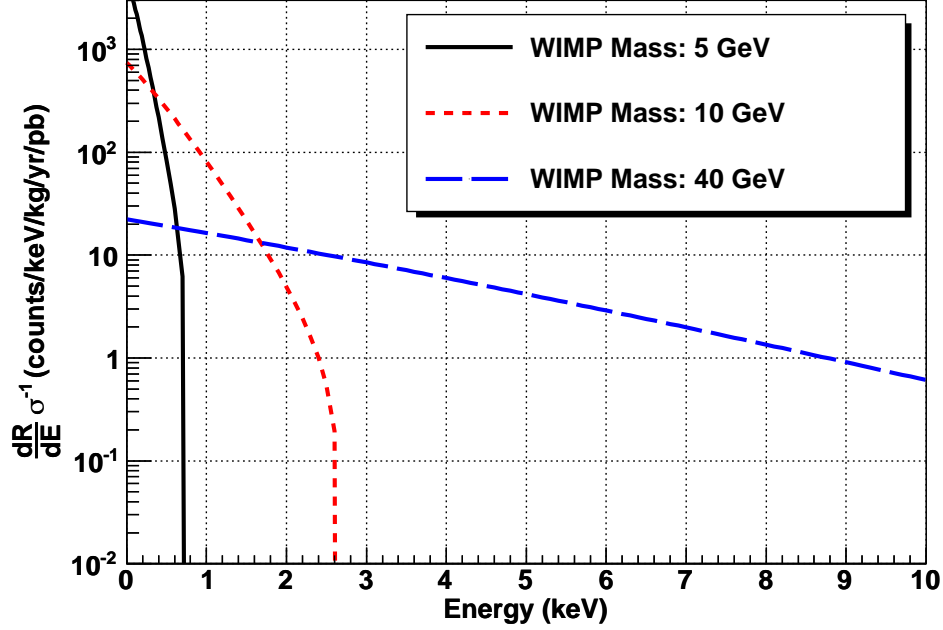
Figure 1.8: Ionization spectrum (differential rate vs. energy) of a WIMP nuclear-recoil in a Ge detector normalized by the interaction cross section. The signal becomes steeper with lower WIMP mass and the truncation due to the escape velocity is more apparent at low mass. Because of the characteristics of the signal, for example, a germanium detector with threshold greater than 1 keV could not detect a recoiling WIMP of mass 5 GeV, underscoring the need for low thresholds to detect low-mass WIMPs.

that lowering the threshold of the germanium detector is critical to detecting low-mass WIMPs, due to the steepening of the curves and the truncation of the signal. The sub-keV thresholds of P-PC detectors therefore enables sensitivity to these low masses. Most WIMP experiments rely on the ability to distinguish between electron and nuclear recoils for background reduction since the vast majority of backgrounds (e.g. gammas, alphas, and betas) interact solely with electrons. P-PC detectors will be unable to distinguish between these two types of events, instead relying on the passive reduction of radioactivity through the choice of radiopure materials and appropriate shielding.
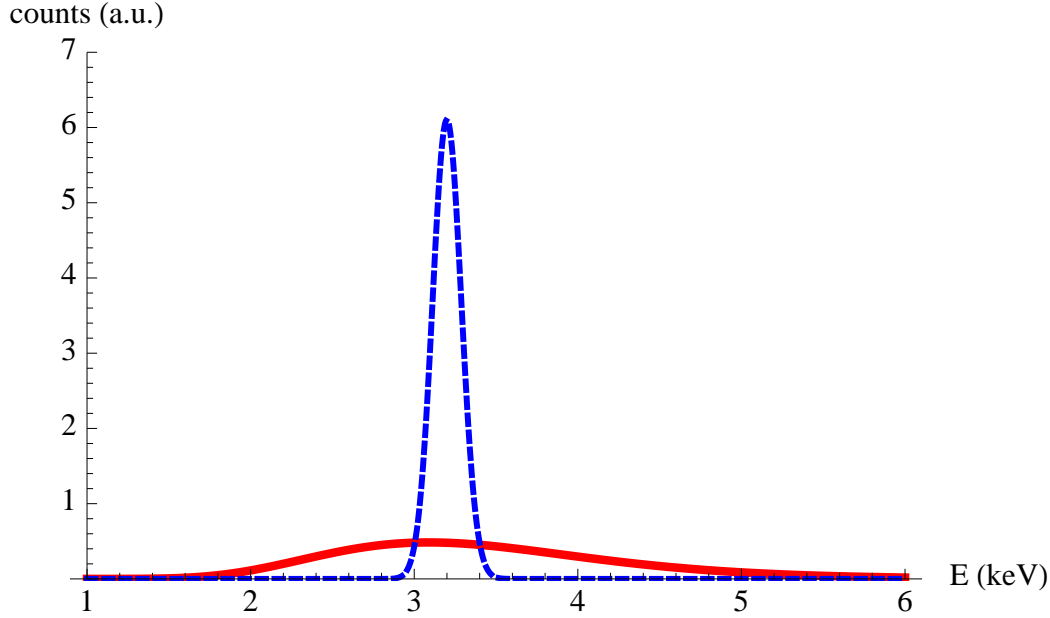
Figure 1.9: Axioelectric signal at $m_a = 3.2$ keV comparing the detector responses of germanium (blue, dashed) and NaI (red, solid) assuming some axion-electron coupling $g_{a\bar{e}e}$. The resolution functions $(\sigma(E))$ for Ge and NaI are $\sqrt{(0.0697)^2 + (0.3)(2.96 \times 10^{-3})E}$ and $(0.448/\sqrt{E} + 0.0091)E$ with E in keV, from [11] and [29] respectively.

The signal for a non-relativistic axion interacting via the axioelectric effect has been derived in [28]. This particular inelastic interaction involves the deposition of the *complete* energy of the axion, yielding an electron-recoil signal centered at the mass of the axion. The low noise of P-PC detectors yields excellent sensitivity to non-relativistic axions with masses ≤ 10 keV due their sharp resolution at low energies. This enhanced performance is especially clear when compared to the capabilities of NaI scintillation detectors. A comparison of an axioelectric signal at a defined axion-electron coupling $(g_{a\bar{e}e})$ is provided in Figure 1.9 for characteristic resolutions of NaI and germanium detectors. In this plot, it is clear that the improved resolution of the germanium detector allows less smearing of the signal, yielding more counts in a narrower peak. This enhanced resolution enables superior distinction of the signal in the presence of background.

It is important to note that P-PC detectors do not search for any specific signal other

than energy deposition in the crystal. Because of this, they may be sensitive to interactions of particles not yet theorized or not yet recognized as theoretically well motivated. Therefore, these detectors provide a necessary complementarity to those experiments looking for particular types of signals and will widen the experimental landscape essential for the development and constraint of theories pertaining to dark matter.

1.6   Outline of this Dissertation

This dissertation will focus on understanding and exploiting the low-energy capabilities of P-PC detectors. It will begin by describing the application of a digital data acquisition system to a P-PC detector deployed underground at Soudan Underground Laboratory in Soudan, Minnesota, and outline conclusions from this initial study. Results and knowledge gained from this initial detector were then applied to the deployment of another, lower-background P-PC at Soudan. The analysis of this detector, including the generation of limits for WIMP and axion dark matter, comprises the bulk of the thesis. Finally, the work ends with some contextual discussion of these results, focusing on estimating the sensitivity of the Majorana Demonstrator to detect dark matter. Appendices detailing the development and use of DAQ hardware and DAQ and analysis software for the results of this dissertation are included for reference.

Chapter 2

DEPLOYMENT OF A DIGITAL DATA ACQUISITION SYSTEM (DAQ) FOR A P-PC DETECTOR AT THE SOUDAN UNDERGROUND LABORATORY

2.1 Introduction

A P-PC detector (henceforth referred to as P-PC2) was procured by Pacific Northwest National Laboratory in January 2008 to be used in studies aimed at more clearly understanding the detector technology [30]. Initial tests took place in the laboratory, focusing on basic operation and pulse-shape analysis techniques relevant for background reduction in the $0\nu\beta\beta$ region-of-interest and for dark matter searches [31]. In November 2008, P-PC2 was deployed underground at Soudan Underground Laboratory at a depth of 2100 m.w.e. The goals of this deployment were two-fold: (1) demonstrate the reliability and operation of a scaled-down MAJORANA-like digital DAQ system in an underground environment, and (2) investigate the possibility of obtaining dark matter exclusion data. The latter goal was uncertain due to the lack of knowledge of the intrinsic backgrounds of the detector and its cryostat.

Underground, the detector was placed inside 8 inches of Pb with a 2-inch inner shield composed of OFHC Cu. This shield was surrounded by a sealed radon-exclusion box over-pressurized with nitrogen gas from LN boil-off to limit the introduction of Rn gas. Borated-polyethylene planks were built up surrounding the outer lead for moderation of neutrons generated by interactions of cosmic-ray muons in the rock walls. An engineering drawing of the shield can be seen with all components in Figure 2.1. Relevant characteristics of the crystal are given in Table 2.1.

2.2 Description of DAQ System

The DAQ system was designed as a single-channel prototype for the MAJORANA DEMON-STRATOR. Additionally, a significant amount of effort was put in to ensure that the setup
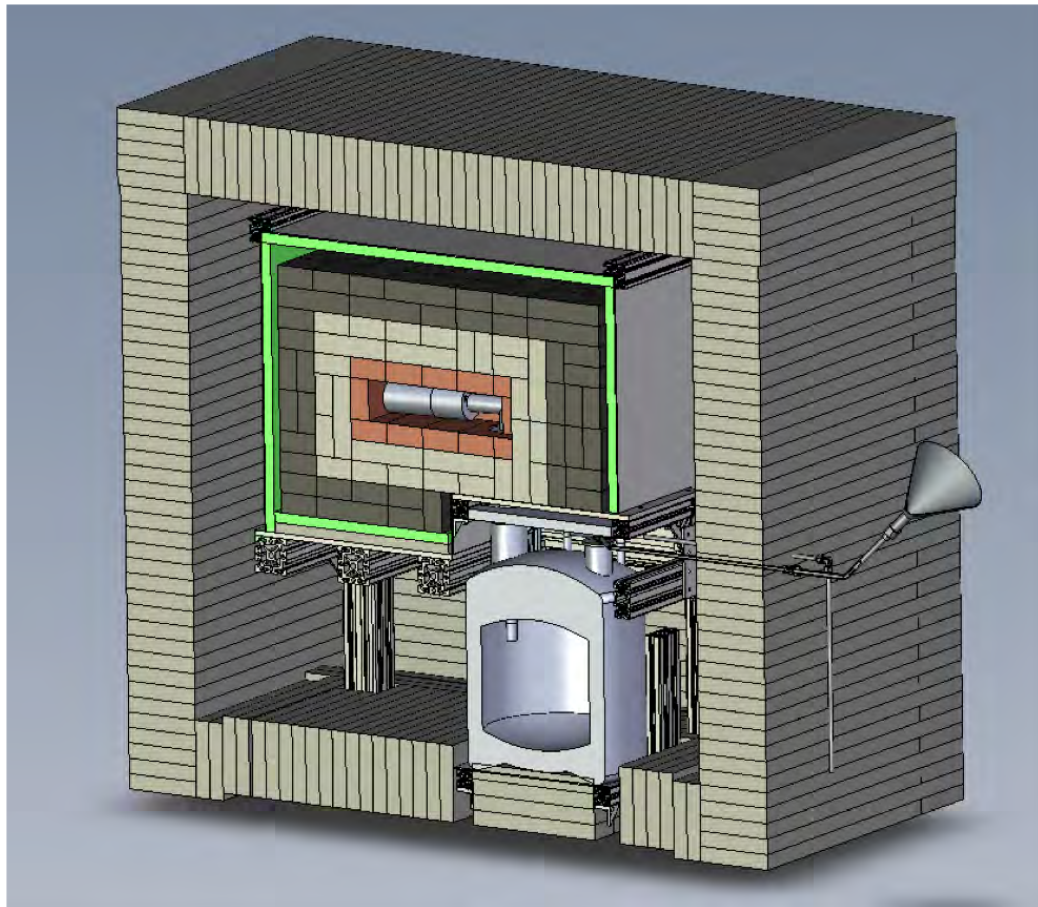
Figure 2.1: Engineering drawing of the setup including inner and outer Cu and Pb shields, Rn exclusion box, and outer neutron moderator. The two colors for the lead outside of the copper inner shield were an internal designation specifying 'cleaner' lead (with less oxidization, better cosmetic appearance, etc.) to be used closest to the inner shield. Graphic courtesy of J. Orrell and E. Fuller.

Table 2.1: P-PC2 detector characteristics. The diameter of the p contact was proprietary manufacturer information and so was unknown.

Property	Value
Manufacturer	Canberra
Mass	528 g
Outer diameter	50.1 mm
Length	50.3 mm
Useful volume	89.4 cm^3
Dead layer	0.5 mm

was remotely configurable to enable the modification of experimental parameters by experimenters at the University of Washington and PNNL. Signals from the detector were read out using a Gretina Mark IV digitizer designed by the GRETA collaboration [32]. This VME64x-based ADC digitized 10 independent channels at 100 Ms/s with a resolution of 14 bits. The card was read out on the VME backplane using a Concurrent Technologies VX 407/042 single-board computer controlled via gigabit ethernet by the ORCA DAQ and slow controls software [33] running on an Apple Mac Mini computer. For more details regarding the ORCA DAQ software and the development of its components relevant to this hardware configuration, please see Appendix C. Also, additional details regarding the development and investigation of this DAQ system are given in Appendix A. The high voltage on P-PC2 was applied with a Canberra LYNX system which allowed for remote powering control of the detector bias. The Liquid Nitrogen (LN) dewar was filled manually by the laboratory staff so no explicit tag was set in the data stream during these fills. Techniques to identify and cut out these events are discussed in Section 2.3.4.

As with many low-noise designs, the P-PC2 preamp employed a pulse-reset circuit to avoid the Johnson noise associated with a feedback resistor. The preamp had 3 outputs: 2 signal outputs and 1 inhibit output which would fire when the reset circuitry of the preamp was active. Additionally, there was a test input for pulser signals used in electronic testing.

The 2 signal outputs were AC-coupled to remove their DC components using independent capacitors of ~ 500 nF to yield a roughly 50 μs decay time of the preamp pulses. One of these channels was input directly into the digitizer; the other was routed through a Phillips Scientific 777 fast amplifier (DC$\rightarrow$200 MHz bandwidth) with roughly 10$\times$ gain to obtain better resolution near threshold from 0.5 $\rightarrow$ 80 keV. The signal was split after the 777, with one output being input into a digitizer channel in the Gretina card and the other routed through a conventional spectroscopy amplifier (Ortec 667) to generate the trigger for the high-gain channel. Figure 2.2 provides a simplified visualization of the DAQ setup including signal paths. See Section A.4.1 for more details regarding the high-gain trigger and its development. The inhibit output was sent directly into the Gretina digitizer card to record the timing information of the reset circuitry.

The test input of the preamp was used for pulser signals to measure the efficiency of the trigger. The pulser used was an Agilent 33220A waveform generator controlled via ethernet by ORCA running on the Mac mini. The pulser output passed through 2 computer-controlled attenuators (HP 8494/5G, 0$\rightarrow$4 GHz bandwidth) providing up to 71 dB of attenuation in increments of 1 dB. These attenuators were actuated using a constructed driver box controlled with an InterPak 408 digital I/O card on the VME bus. For more details on the pulser test setup, see Section A.4.1. A synchronization pulse from the Agilent was input into a digitizer channel for timing information. The pulser was also used during counting runs, inputting a signal of low amplitude ($\sim$2.5 keV) into the DAQ system at low rate (~ 1 Hz) to measure any deviations in the electronics over time.

Finally, an input channel on the Gretina digitizer was designated as an input for a muon veto that was installed after initial deployment of the DAQ system. However, the veto was unused in the analysis and so details about its operation are omitted. A summary of the input channels on the Gretina card including the type of data saved for each is given in Table 2.2. The DAQ system was fully deployed underground at the beginning of April 2009.

2.3   Analysis and Results

After a few months of commissioning and calibration tests, counting runs were begun on 3 May 2009. Runs were cycled every hour, and data were automatically synchronized with a

24
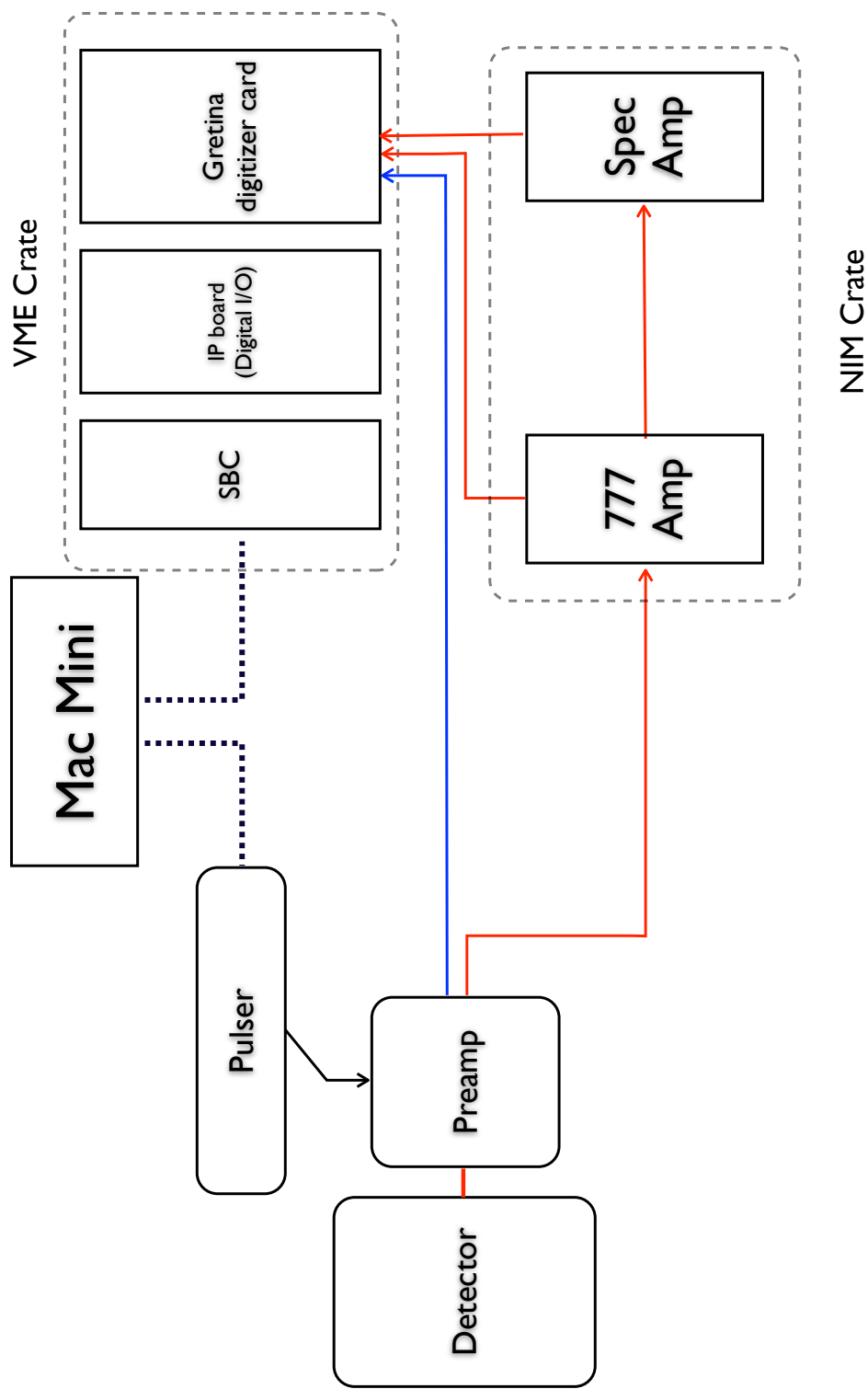
Figure 2.2: Simplified schematic of DAQ setup. Lines with arrows denote signal flow.

Table 2.2: Channel summary for the Gretina digitizer card.

' Type	Channel	Info
Reset inhibit	9	Only timing
Low-gain signal	8	10 μs trace (1022 samples)
Pulser sync	7	Only timing
Muon veto	2	Only timing, unused in analysis
High-gain trigger	1	Only timing
High-gain signal	0	10 μs trace (1022 samples)

server at the University of Washington. On 10 June 2009, automatic trigger efficiency tests were begun in between runs to monitor the stability of the trigger over time. Counting runs ended on 24 July 2009, yielding a total live-time of 70.4 days. A CouchDB database [34] was used to store both metadata (e.g. timing, configuration, file names, etc.) as well as analysis parameters (e.g. average baseline, trigger efficiency, measured electronic noise) from each run. This database was also used to facilitate the analysis chain described in the following section.

2.3.1 Data Processing

All data were handled in a tiered reduction scheme. The analysis process was automated and made significant use of the ROOT analysis framework [35] as well as waveform analysis tools (MGDO, see Appendix B for more information) developed for the MAJORANA experiment. An outline of the tiered data-flow process is given:

Tier 0: Raw binary data – from the ORCA DAQ system

Tier 1: ROOTified data – raw data converted to MGDO objects and stored in ROOT TFiles

Tier 2: Waveform processed data – extraction of waveform characteristics using MGDO Transforms

Tier 3: Reduced data and split into sub-components (folders) according to different channels

This process was tracked using a CouchDB database. The Tier $0 \to 1$ conversion process took the raw binary data and converted it into MGDO C++ objects which could then be serialized to disk in ROOT TFiles. The Tier $1 \to 2$ processing calculated several parameters from the waveform data, including baseline, extrema (maximum and minimum), and rise-time and determined the amplitude of the waveform using a trapezoidal filter[1] [36]. More information regarding the baseline and rise-time calculations are given in Sections 2.3.4 and 2.3.5. Additionally, at this stage coincidences between events in different channels were calculated using the timing information from the digitizer card, defining events within 1 μs of each other as concurrent. Finally, the Tier $2 \to 3$ processing reduced the data size, removing noise events (i.e. events well below energy threshold) and splitting the two signal channels 0 and 8 into separate file groups. The Tier 3 files were used directly in analysis and made available to others in the MAJORANA collaboration through a cloud-based file system (Dropbox[2]).

2.3.2   Initial Calibration

The system was initially calibrated with a ^{133}Ba source using peaks detailed in Table 2.3. The high- and low-gain channels were calibrated independently and spectra for each channel are shown in Figure 2.3. The two channels exhibited an energy range up to $\sim$ 100 keV and $\sim$ 2.6 MeV, respectively, though the range of the low-gain channel is discussed in more detail in Section 2.3.6 . These initial calibrations were used for commissioning tests such as trigger efficiency and electronic noise measurements and were later refined using prominent peaks from backgrounds e.g. in the U and Th chains, ^{40}K, and x-ray lines from ^{68}Ge and ^{65}Zn electron capture decays (see Section 2.3.6). During the counting runs, regular calibration runs were not made; instead we monitored the behavior of the detector using an injected pulser.

---

[1]The Gretina digitizer included an on-board trapezoidal filter, but it was chosen to use an offline filter for greater flexibility.

[2]See the Dropbox website: `https://www.dropbox.com` for more information.
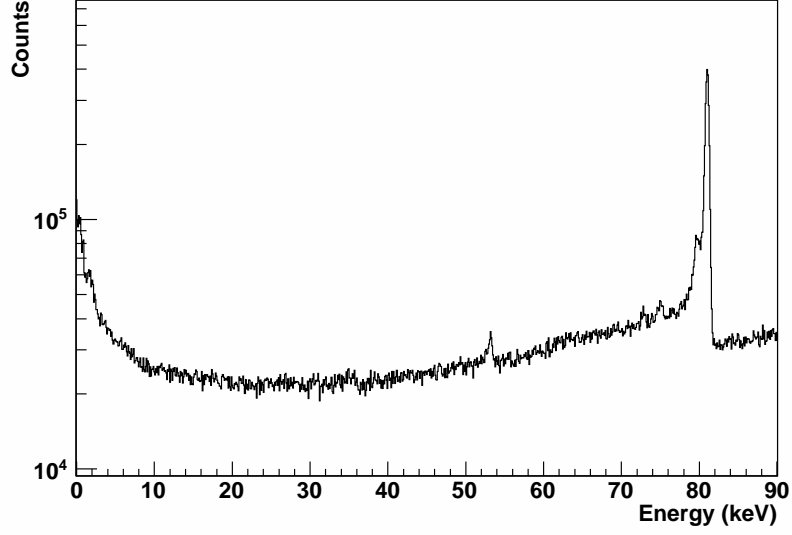
(a) High-gain channel
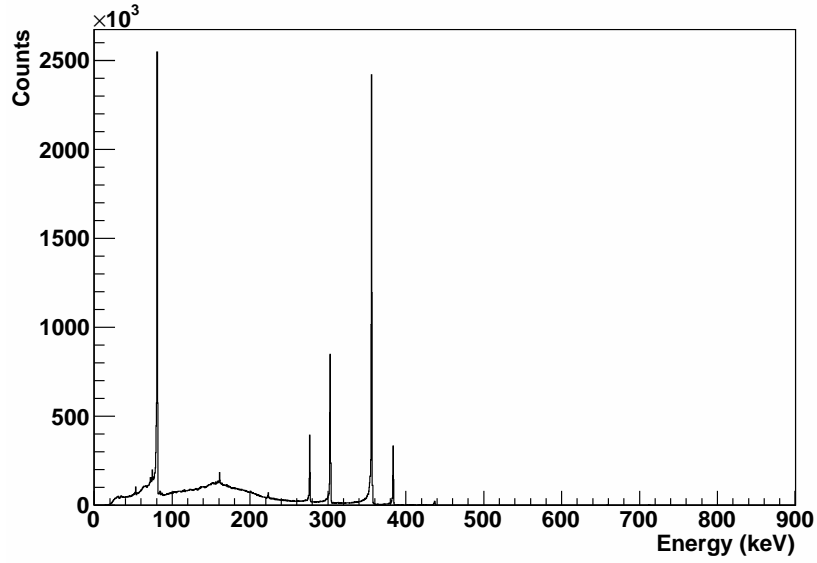
(b) Low-gain channel

Figure 2.3: Calibration of the high- and low-gain channels. The high-gain channel was amplified at roughly 10×. Only a limited range of the low-gain channel is shown since this is the region populated by gammas from the ^{133}Ba source. However, the range of the low-gain channel went out to roughly 2 MeV (see Section 2.3.6).

Table 2.3: Selected gamma lines used for calibration in the ^{133}Ba spectrum, data adapted from [37].

Energy (keV)	Intensity
53.16	2.2 %
79.61	2.62 %
80.997	34.1 %
160.61	0.65 %
223.24	0.45 %
276.4	7.16 %
302.85	18.33 %
356.01	62.05 %
383.85	8.94 %

2.3.3   Detector Parameters vs. Time

Several parameters of the system were tracked over time, including the trigger efficiency and the electronic noise of the system. The former was tested by performing efficiency tests every hour in between run cycles; the latter by looking at the recorded data of the input pulser pulses (1 Hz) versus time. Several other parameters, including the baseline of the waveforms and the number of events in a certain energy ranges, were tracked as well: these results are presented in Section 2.3.4. In addition to these specific parameters being tracked, the rates of each channel were tracked. While tracking these rates, it was determined that the reset rate of the preamp increased towards the end of an LN fill cycle. Higher values of inhibit rate (lower baseline values) occurred as LN in the dewar boiled off, with sharp returns when the dewar was refilled. The interpretation of this was that as the level of LN in the dewar fell, the temperature of the front-end detector electronics (and possibly the crystal) would increase which would increase the detector leakage current. An increase in leakage current would cause the preamp to reset more often to clear the charge more quickly

accumulating on the feedback capacitor. It was not determined precisely what caused this temperature sensitivity; however, it was likely due to insufficient thermal coupling between the detector and the liquid nitrogen. During installation of the detector within the Pb shield, an extension had to be added to the dewar end of the cold finger. It is possible that this extension didn't properly thermally couple to the cold finger and thus didn't cool the detector and/or the internal electronics as effectively. The change of the reset rate over time had implications for other parameters as discussed in the following sections. A summary of comparisons between detector parameters is shown later in the section in Figure 2.5. In this figure, LN fills occurred around runs 1590, 1665, and 1750. The average values of the parameters in the following sections is given in Table 2.4.

Trigger Efficiency

Trigger efficiency was measured as described in Section A.4.1. Initial tests were performed during the commissioning period (i.e. before counting runs began) and automatic tests were run after every hour-long cycle beginning 38 days into the counting run period. The goal of these tests was to investigate how the trigger threshold behaved over time, tracking any changes that might occur due to shifts in environmental conditions or because of other unforeseen phenomena. To parameterize the efficiency, the data were fit to an error function, $f_{eff}(E)$, in the form:

$$f_{eff}(E) = \frac{1}{2} - \frac{1}{2} \operatorname{erf}\left(s(E - \rho)\right) \tag{2.1}$$

with energy, E, in keV; scaling parameter, s, in keV^{-1}; and offset, ρ, in keV. An example of a fit to this data is given in Figure 2.4. Some results of the trigger tests over time are given in Figure 2.5, comparing them to other detector parameters. The offset parameter, ρ, was largely stable but increased before the transition from I→II and again during region III. ρ returned to a lower value during region II. The scaling, s, was largely insensitive to any change in reset rate, though it exhibited slight deviations low in transitions between regions I and II.
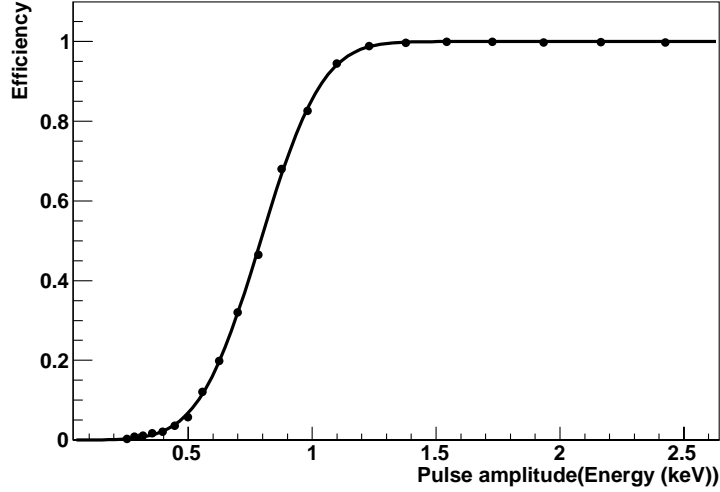
Figure 2.4: Triggering efficiency measurement example. Line is a fit to Equation 2.1.

Pulser Data

Pulses from the Agilent waveform generator were input into the test input at a rate of 1 Hz and at low energy: ~ 2.5 keV. These events underwent analysis equivalent to every other waveform and were selected by finding all waveforms which arrived in coincidence with the pulser synchronization signal. Each hour-long run yielded a data set of ~ 3600 events; the calculated amplitudes (energies) of these events were histogrammed and fit to a Gaussian to extract the mean, μ, and the sigma, σ, of the pulser signal for each run. Results of these fits are plotted versus run number in Figure 2.5. It is clear that the width of the noise (σ) did not significantly alter for reset rates below 15 Hz. However, the mean (μ) of the pulser demonstrated three clear transition regions an example of which has been labeled 'I', 'II', and 'III' in Figure 2.5: from reset rates 3→7 Hz (region I labeled in the plot), the mean increased, transitioned to lower values for reset rates 7→11 HZ (region II), and then returned for rates above 11 Hz (region III). The absolute cause of this behavior was not determined, but it is possibly due to a modification of operating parameters of the front-end crystal electronics (e.g. the FET) by the increasing leakage current.
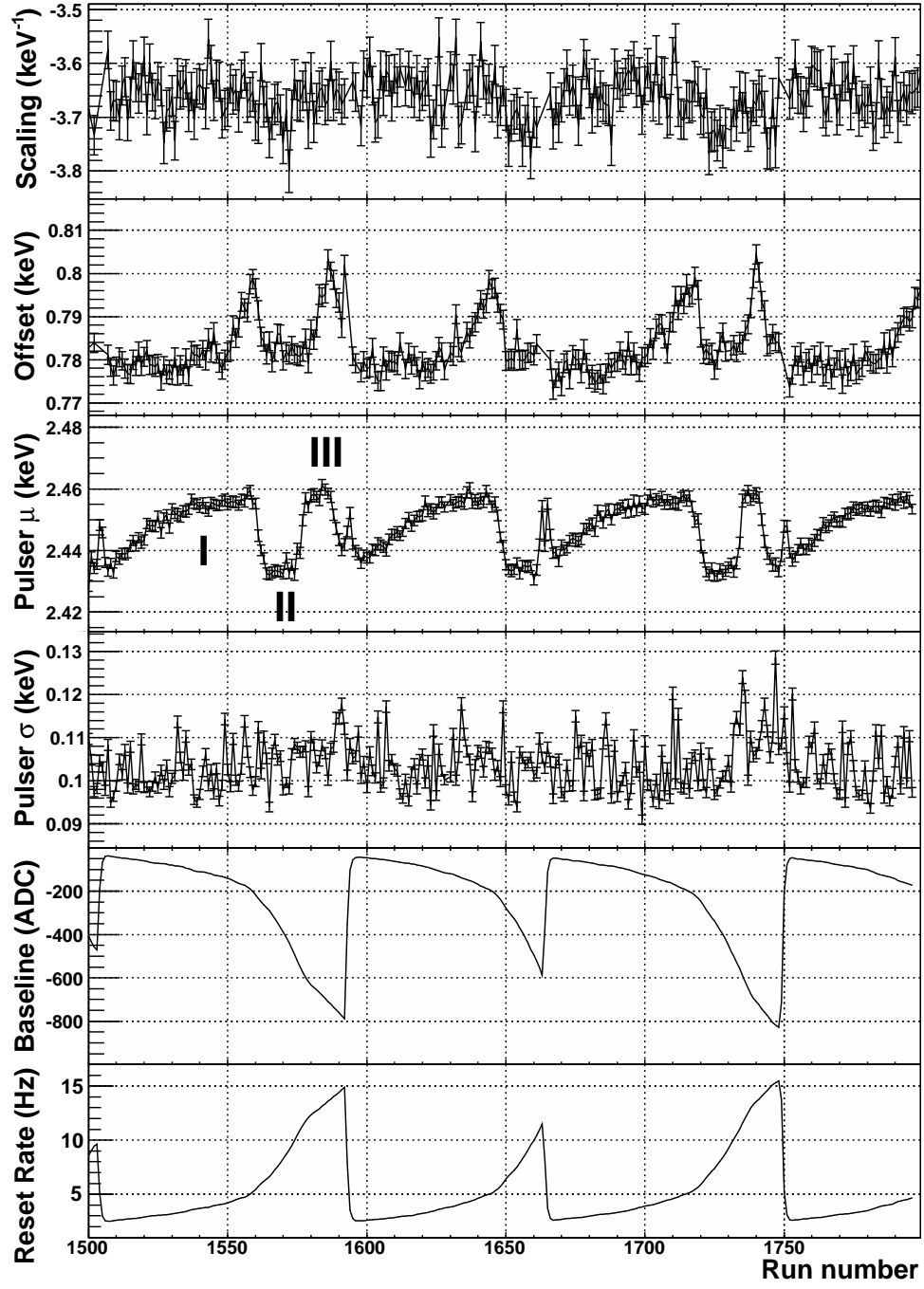
Figure 2.5: Comparison of various parameters versus time for a subset of runs. See text for details.

Table 2.4: Average parameters for trigger efficiency and electronic noise.

Parameter	Value
Pulser Data	
Mean, μ	2.45 ± 0.009 (keV)
Sigma, σ	0.100 ± 0.0046 (keV)
Trigger Efficiency	
Offset, ρ	0.784 ± 0.006 (keV)
Scaling, s	-3.666 ± 0.044 (keV^{-1})

Conclusions

The results of this section underscore the need to track detector parameters over time to monitor changing environmental conditions and other unanticipated behavior. The gain and offset shifts evident in the pulser and trigger efficiency results do not significantly affect these data because their magnitudes (<25 eV) are much less than the nominal noise FWHM ($\sim$235 eV) and trigger threshold (90% efficient at 1 keV) (see average values in Table 2.4). However, these deviations could begin to have a larger effect with improved systems with reduced noise and threshold. Additionally, when extracting information on physics signals in regions close to threshold (such as dark matter), it is critical to study parameters affecting this region over time to understand systematics important for limits or claims.

2.3.4 Cuts

Cuts were introduced to clean the data with the goal of removing spurious events from, for example, noise, microphonics and reset-pulse events. These consisted of (1) a timing cut for removing reset pulse events, and (2) a baseline cut and (3) an integrated-counts-per-run cut for removing noise and microphonics-induced events. The first two cuts were applied on an event-by-event basis, whereas the last cut removed an entire hour-long run.

Reset-pulse Timing Cut

Reset pulse waveforms had distinct qualities (positive clipping whereas energy pulses were negative-going) which allowed the waveforms to be removed online from the data stream while retaining the timing information of the event. The offline timing cut introduced a veto window of 1 ms following a reset event. Given the average reset rate ($\sim$10 Hz), this introduced a $\sim$1% correction to the live-time.

Baseline Cut

The baselines of the traces were monitored on an event-by-event basis by calculating a 1st-degree polynomial fit to the first 3.5 μs of each preamp trace. From this information, an average baseline was calculated for each run to generate a cut based upon significant deviations from these average values. In a stable system, the baseline will be Gaussian distributed around a central value according to the magnitude of the noise. The integrated-counts-per-run in particular energy windows (0.6-10 keV and 10-70 keV) were monitored as well to observe deviations from expectation assuming Poisson statistics. This second cut relied upon the fact that spurious events from increases in environmental noise and microphonics arrive in bursts, see e.g. [38]).

The average baseline per run shifted in a way systematically correlated with a shift in the measured inhibit rate, shown in Figure 2.5. As the leakage current increased, the inhibit rate went up and the measured baseline dropped due to additional charge collection on the capacitors used for AC coupling the signal to the digitizer. In general, this downward shift in baseline could affect the dynamic (energy) range of a digitizer channel because it would allow fewer ADC values for the (negative-going) pulses before clipping. To generate a cut on this value, the values of the baseline for each event with an energy greater than 0.4 keV during a run were histogrammed and fit to a Gaussian. (The integrated counts in this energy range for an hour-long run was $\sim$4000.) A cut was then generated based upon the results of this fit: rejecting events outside the range $\mu \pm 3\sigma$ for a 99.7% acceptance. An example of one of these fits is shown in Figure 2.6. During LN fills, the baseline would return to its maximum value making a cut based upon the average baseline value impossible. Runs
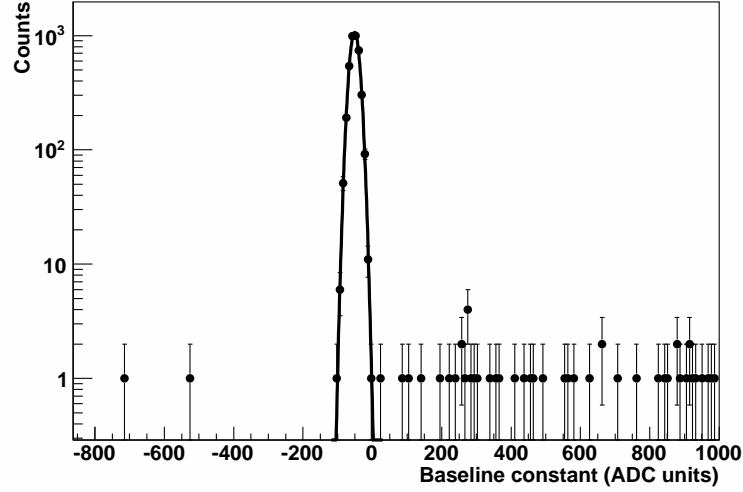
Figure 2.6: An example of a fit to the baseline for one hour-long run. Line is a Gaussian fit to the data.

during these transition times were cut from the data set.

Integrated-counts-per-run Cut

To determine the integrated-counts-per-run cut, counts in a designated energy window for each run were calculated and histogrammed. The resulting data were then fit to a Poisson distribution and a cut was generated based upon 99.99% acceptance using the fit distribution. Results of this calculation are shown in Figure 2.7. All runs with ≤ 29 counts between 0.6 and 10 keV were accepted.

All of these cuts were applied to the data to obtain a cleaned data set and all results described in the following sections use data after cuts.

2.3.5   Rise-time Calculations

The rise-time of each pulse was calculated using a simple algorithm:

1. Subtract the baseline, b, and determine the amplitude, a.
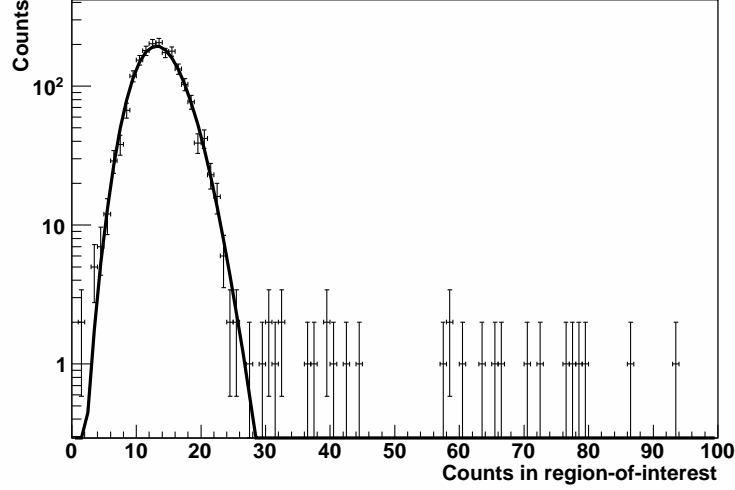
Figure 2.7: Calculation of cut based upon the number of events in the energy window 0.6-10 keV per run. The line is a fit to a Poisson distribution.

2. Forward search from beginning of the waveform to determine when the waveform reaches $\frac{1}{2}a$ to find the middle of the pulse, m.

3. From m, search backwards (forwards) to determine start (end) of the rise of the waveform. The start and stop of the waveform are defined when the pulse reaches 10% and 90% of the amplitude.

The rise-time versus energy is plotted for each channel in Figure 2.8. The low-gain channel demonstrates a sharp intensity band centered at $\sim$0.3 μs rise-time indicating that most pulses have this fast rise-time. A number of gamma lines at higher energy are apparent in the data, including ^{40}K (1460.8 keV), ^{214}Bi (1764.5 keV), and ^{208}Tl (2614.5 keV)[3], with rise-time greater than 0.3 μs. Because a large portion of events in these lines are composed of Compton-scatter interactions which deposit energy at multiple sites in the detector, this suggests multi-site events are being observed as slow-rise-time events.

---

[3]The thallium line is shifted in down in energy to $\sim$2550 keV. An explanation of this is provided in Section 2.3.6

At low energy in the high-gain channel, one does not expect to see a significant multi-site population because Compton scatters summing to such a small energy are very improbable. However, the rise-time distribution for the high-gain channel yields some interesting results: the strong line at 0.3 μs still exists and there is a distribution near threshold ($\lesssim 1$ keV) where it is clear the rise-time calculation algorithm breaks down, but there is also a strong population of slow-rise-time events at low energy, $1 \leq E \leq 20$ keV. Closer inspection reveals a significant visual difference between 'fast' (~ 0.3 μs) pulses and 'slow' (> 1 μs) pulses as can been seen in Figure 2.9. A similar result was found for this detector by J. Orrell [39]. Since these slow-rise-time events are predominately at low energy, they can compose a background to any signal near threshold. See Section 3.4.3 for discussion on improved identification and removal of these events including a discussion on their potential origins.

2.3.6   Energy Spectra

Studies on the high-energy channel were performed by A. Schubert to identify peaks and study the resolution of the detector with higher energy gamma lines [40]. The spectrum was calibrated using prominent peaks in the spectra. The resolution was fit to a second-order polynomial function of energy and determined to be:

$$\sigma = \sqrt{(0.242 \pm 0.016)^2 + E(4.04 \pm 0.1 \times 10^{-3})^2 + E^2(9.2 \pm 0.0004 \times 10^{-4})^2} \qquad (2.2)$$

with E, energy in keV. It was found that the 2614.5 keV peak of the ^{208}Tl decay was shifted down in energy to 2561 keV and had a larger-than-expected sigma (given Equation 2.2) of 8.6 keV. This was explained by noting that large energy depositions in the crystal could saturate the pulse-reset preamp introducing either clipped or distorted pulses into the digitizer. A selection of a few prominent lines fit in [40] is given in Table 2.5. Results from this suggest that the linearity of the high-energy channel can not be trusted past the 1764 keV line of ^{214}Bi. This conclusion highlights a concern regarding the difficulty of using low-noise preamps with reset electronics while simultaneously looking at high-energy data >3-4 MeV as the MAJORANA DEMONSTRATOR will do. Preamp solutions will have to address the optimization problem of ensuring enough dynamic range to search for $0\nu\beta\beta$ while retaining low noise and threshold.
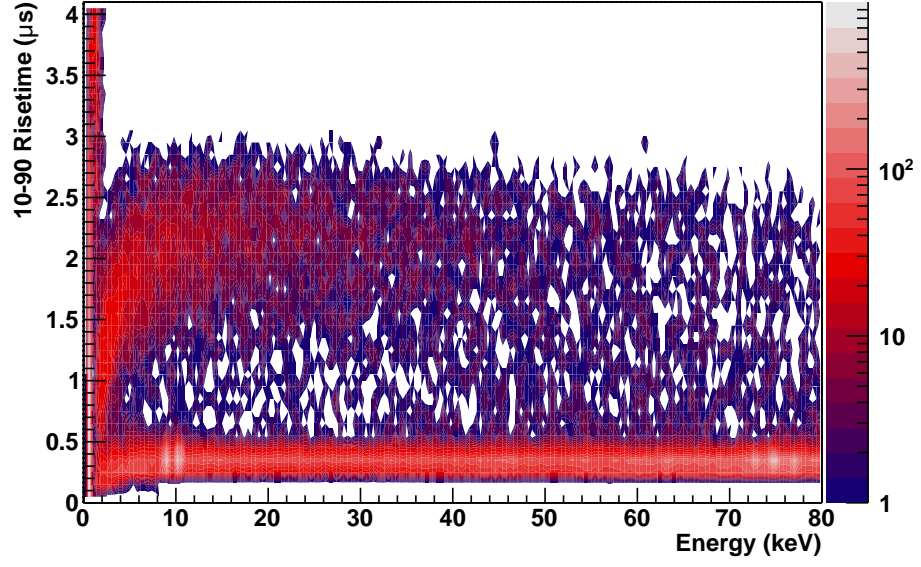
(a) High-gain channel.
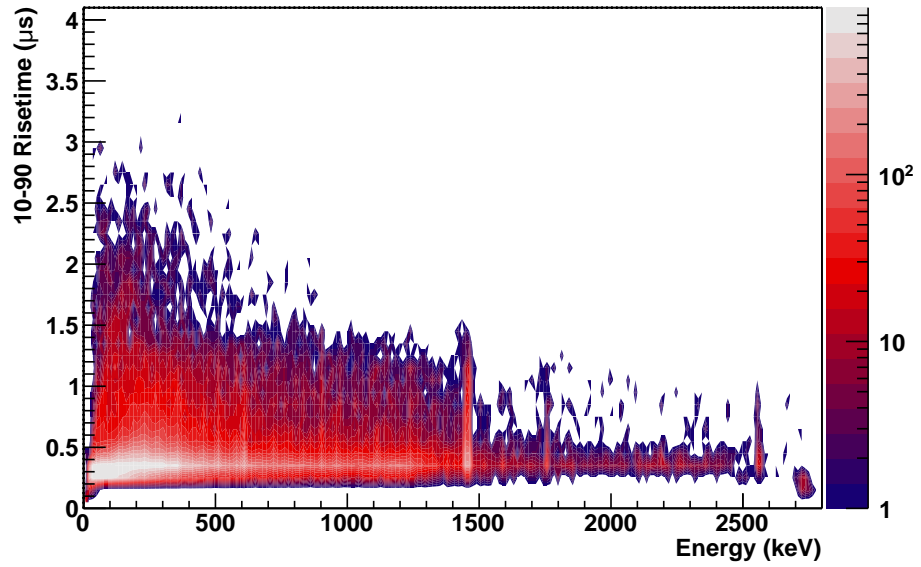
(b) Low-gain channel.

Figure 2.8: Rise-time measurements for both high- and low-gain channels given in time from 10%→90% pulse amplitude.

(a) Fast pulse.

(b) Slow pulse.
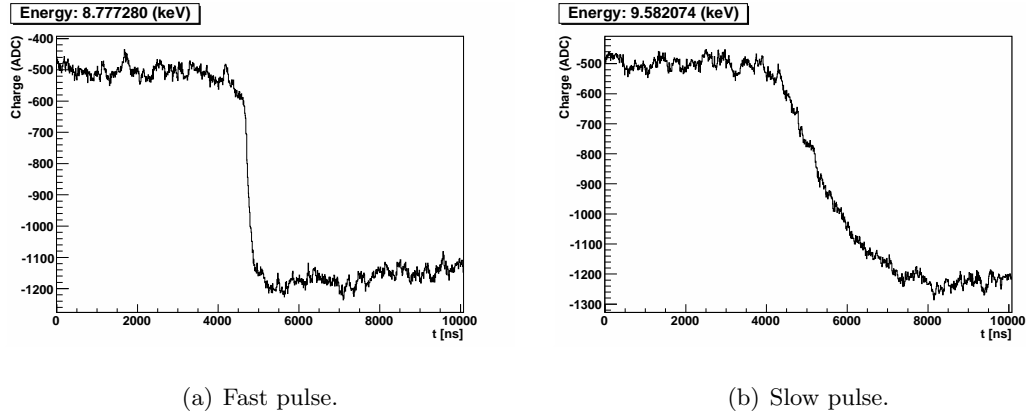
Figure 2.9: Examples of slow and fast rise-time pulses.

Table 2.5: Selection of gamma lines in the high-energy channel of P-PC2. Table courtesy A. Schubert.

Source	Energy [keV]	Fit centroid [keV]	Sigma [keV]
^{226}Ra	186.2	185.7 ± 0.04	0.37 ± 0.03
^{212}Pb	238.6	238.5 ± 0.04	0.32 ± 0.01
^{214}Pb	242.0	241.9 ± 0.04	0.30 ± 0.06
^{214}Pb	295.2	295.1 ± 0.04	0.41 ± 0.02
^{228}Ac	338.3	338.1 ± 0.04	0.42 ± 0.04
^{214}Pb	351.9	351.7 ± 0.04	0.38 ± 0.01
^{208}Tl	583.2	583.5 ± 0.04	0.58 ± 0.03
^{214}Bi	609.3	609.6 ± 0.04	0.60 ± 0.02
^{228}Ac	911.2	911.6 ± 0.1	0.71 ± 0.06
^{228}Ac	969.0	969.3 ± 0.1	0.79 ± 0.10
^{214}Bi	1120.3	1120.5 ± 0.1	1.03 ± 0.11
^{40}K	1460.8	1460.7 ± 0.04	1.33 ± 0.02
^{214}Bi	1764.5	1764.3 ± 0.2	1.72 ± 0.22
^{214}Bi	2204.2	2201.5 ± 0.5	1.94 ± 0.36
^{208}Tl	2614.5	2560.6 ± 0.8	8.59 ± 0.70
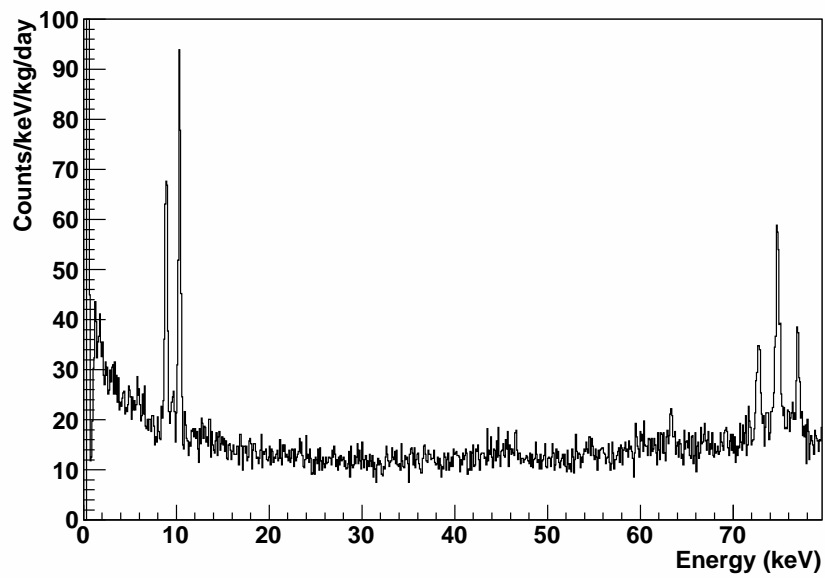
(a) High-gain channel.
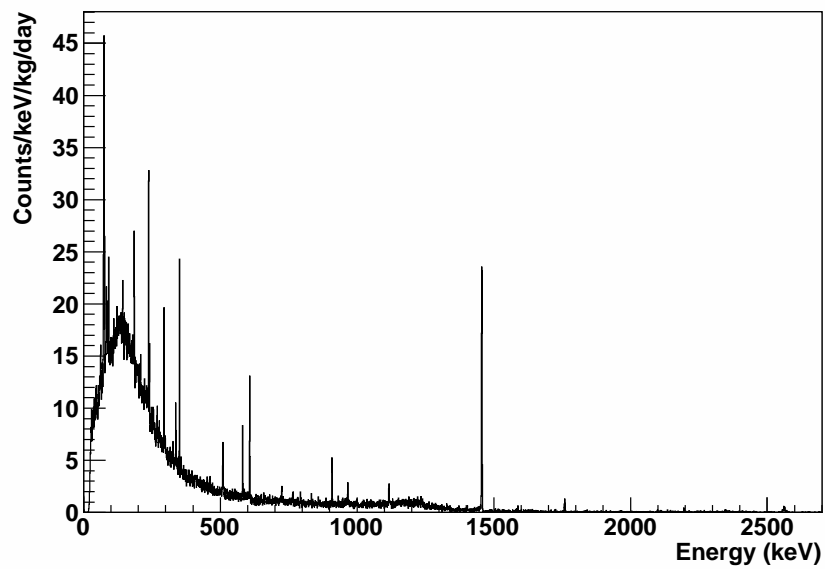
(b) Low-gain channel.

Figure 2.10: Energy spectra of P-PC2.

Table 2.6: Lines in the low-energy channel of P-PC2.

Source	Energy (keV)	Fit Mean (keV)	Fit Sigma (eV)
^{65}Zn K-capture	8.979	$8.95 \pm 7.6 \times 10^{-3}$	139 ± 7.3
^{68}Ge K-capture	10.367	$10.394 \pm 5.8 \times 10^{-3}$	119 ± 7.3
$^{(208)}$Pb K$_{\alpha2}$	72.805	$72.86 \pm 2.1 \times 10^{-2}$	188 ± 19.7
$^{(208)}$Pb K$_{\alpha1}$, $^{(214)}$Bi K$_{\alpha2}$	74.969, 74.815	$74.92 \pm 1.3 \times 10^{-2}$	234 ± 19.7
$^{(214)}$Bi K$_{\alpha1}$	77.107	$77.11 \pm 1.9 \times 10^{-2}$	163 ± 19.7

The calibration of the low-energy channel was refined using the ^{68}Ge and ^{65}Zn K-capture lines plus the non-overlapping x-ray lines from Pb and Bi (see the observed energy spectra in Figure 2.10). The final maximum likelihood fits to these lines, means and sigmas and their expected value, are given in Table 2.6. The x-ray lines from Pb and Bi come from the β-decays of ^{208}Tl and ^{214}Pb in the ^{232}Th and ^{238}U decay chains. The two lines at $\sim$73 keV were fit with one Gaussian which generated a wider width.

2.4 Estimate of ^{68}Ge Surface Production

^{68}Ge is a cosmogenically-produced isotope inside germanium detectors which is a well-known background to $0\nu\beta\beta$ (see e.g. [41]) due to the fact that the decays of its ^{68}Ga daughters can deposit energy in the $0\nu\beta\beta$ region-of-interest. A schematic of the decay chain is shown in Figure 2.11. The decay of ^{68}Ge proceeds via electron-capture with a half-life of 271 days to ^{68}Ga, which subsequently decays to ^{68}Zn with a half-life of 68 minutes. The latter decay has a Q-value of 2921 keV which could deposit energy at 2039 keV, the $0\nu\beta\beta$ Q-value. This background can be reduced two ways: (1) an analysis cut which tags the initial decay of ^{68}Ge and vetoes the detector for a specified multiple of ^{68}Ga half-lives, and/or (2) reduce the intrinsic ^{68}Ge content in germanium detectors by either storing them in a low-cosmic-ray-flux environment (i.e. underground) and waiting or by actively removing ^{68}Ge through isotopic depletion. Because ^{68}Ge decays inside the detector, the total energy emitted through Auger electrons and x-rays is summed to lines of the K-, L-, or M-electron binding energy of the

Table 2.7: Decay information of ^{68}Ge. Capture probabilities are adapted from reference [45], characteristic energies from [46].

^{68}Ge decays	Percent	Energy (keV)
K-capture	$86.25 \pm 0.22\%$	10.367
L-capture	$11.45 \pm 0.20\ \%$	1.299
M-capture	$1.92 \pm 0.08\%$	~ 0.1

$$^{68}\text{Ge} \xrightarrow{\epsilon,\ \text{T}_{1/2}=271\ \text{d}} {}^{68}\text{Ga} \xrightarrow{\epsilon,\ \text{T}_{1/2}=68\ \text{m, Q}=2921\ \text{keV}} {}^{68}\text{Zn}$$
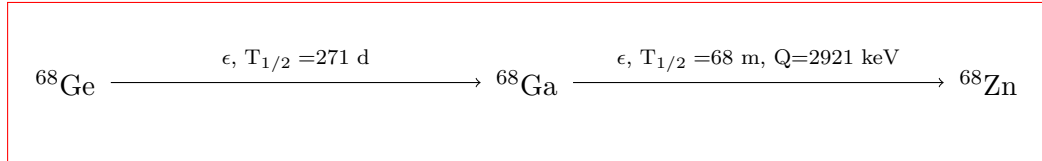
Figure 2.11: Decay chain of ^{68}Ge.

daughter nuclei (see Table 2.7). These lines can then be used to tag events likely to belong to a ^{68}Ge decay. Tagging the ^{68}Ge decays to reduce the ^{68}Ga background is dependent on the decay rate of ^{68}Ge, $\lambda^{(^{68}Ge)}$, and is not a viable solution until $\lambda^{(^{68}Ge)}\tau_{1/2}^{(^{68}Ga)} \ll 1$. Estimations of the ^{68}Ge activation rate at the earth's surface in both natural and enriched (86% ^{76}Ge) have been performed previously (see e.g. [42, 43, 44]).

The P-PC2 detector provided an excellent opportunity to measure the ^{68}Ge activation rate at the surface. Though the background rate at the L-capture energy of ^{68}Ge was too high to resolve any peak, the K-capture line at 10.367 keV provided a means to measure the activation rate. At the surface, the detector was assumed to be at equilibrium: that is, the decay rate of the cosmogenic isotopes were equal to the population rates. This was a good approximation since the detector had been kept at the surface for a period on the order of the lifetime of ^{68}Ge. Once the detector is brought underground, the cosmogenic isotopes are being populated at a rate much lower than that at the surface and it is a safe to assume at the depths of Soudan ($\sim$ 2000 m.w.e.) that this rate is 0. Therefore, an underground measurement of the decay rate of ^{68}Ge in counts/day can be used to determine the activation

rate in counts/day at the surface. The measurement was performed beginning 7 months after the detector had been deployed underground and after ^{71}Ge, a short-lived cosmogenic ($\tau_{1/2} = 11.4$ days) which also populates the same capture peaks, had decayed.

To perform the measurement, an energy window was selected around the K-capture peak from $10 \rightarrow 10.73$ keV as the "signal" region – chosen as symmetric around the expected signal mean – and from $11 \rightarrow 15$ as the "background" region, nearby and without any x-ray lines. The average rate, in counts/keV/kg/day, was computed for each 72 hours of live-time in both regions and fit simultaneously to the function:

$$f(E,t) = \begin{cases} a + be^{-(\log 2)t/\tau_{1/2}} & 10 \leq E \leq 10.73 \text{ (keV)} \\ \\ a & 11 \leq E \leq 15 \text{ (keV)} \end{cases} \tag{2.3}$$

with the free parameters a (flat background rate), b (^{68}Ge rate), and $\tau_{1/2}$ (^{68}Ge half-life). Fit results of this are summarized in Table 2.8 and plotted in Figure 2.12. At the beginning of the counting runs, the detector had already been underground 212 days, meaning that the estimated decay rate, b, must be multiplied by the value $e^{(\log(2)\ 212d/270.95d)} = 1.72$ as well as the energy window (0.73 keV) to get the counts-per-day-per-kilogram the day the detector was brought underground. It was necessary to estimate the acceptance of the energy window using parameters (μ and σ) found from a maximum likelihood fit to the data (see, e.g. Section 2.3.6); this acceptance was measured as 0.99797 ± 0.001. Finally, since only the K-capture line was used in this estimation, the percentage of decays into this channel must be taken into account (see Table 2.7). The final result was then determined to be 53.3 ± 4.3 decays/kg/day which can also be interpreted as the activation rate for ^{68}Ge, assuming the detector was at saturation at the surface. A comparison of this value to previous results is given in Table 2.9. It should be noted that this detector was fabricated in Europe and then shipped via airplane to Richland, WA, USA so that additional activation would have occurred during transportation. Therefore, the calculated activation rate of 53.3 ± 4.3 ^{68}Ge atoms/kg/day should be considered conservative since the detector would not have reached equilibrium from its original shipment date until its deployment underground (~ 1 year).

Table 2.8: Fit information for ^{68}Ge analysis.

Parameter	Fit Value
a	16.166 ± 0.33 (counts/keV/kg/d)
b	36.54 ± 2.94 (counts/keV/kg/d)
$\tau_{1/2}$	192 ± 96 (days)
χ^2/NDF	$32.3/45$
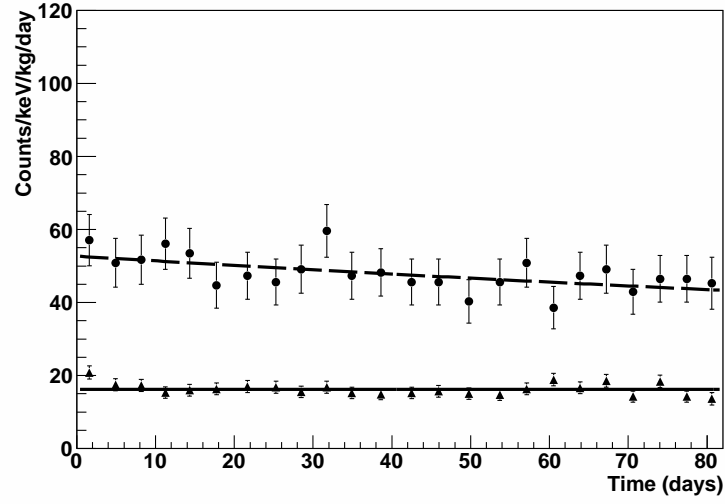
Figure 2.12: Estimation of ^{68}Ge production at the surface. Data and fit for two regions: (1) $10 \to 10.73$ keV, fit to exponential + constant (circles); and (2) $11 \to 14$ keV, fit to constant (triangles). See text for fit details.

Table 2.9: A summary of previous estimates of ^{68}Ge surface activation rates in natural Ge, adapted from Table 1 in [43].

Rate (^{68}Ge atoms/kg/day)	Reference
53.3 ± 4.3	This calculation
29.6 (Calculation)	[42]
30 ± 7 (Experiment)	[42]
58.4	[47]
82.8	[48]
89	[41]
45.8	[49]
41.3	[44]

2.5 Conclusions

A MAJORANA-like data acquisition system operated stably in a deployed environment for over a half year, demonstrating software and hardware robustness required for the MAJORANA DEMONSTRATOR. A tiered analysis framework was constructed that allowed automated processing and management of data from the deployed system. This flexible framework was applied to data from two other detectors, including one analyzed in Chapter 3. Several parameters of the data were tracked and exhibited deviations over time, underscoring the need to monitor channel rates, electronic noise, and trigger efficiency throughout the life of a deployed system. Although the backgrounds seen in the detector were too large to provide competitive limits on dark matter signals, two important results from the data were made including a measurement of rise-time versus energy for low-energy pulses and an estimation of ^{68}Ge surface activation in natural germanium. The measurement of the former determined a potential source of background for applications such as the MAJORANA DEMONSTRATOR using P-PCs for direct dark matter detection. Results from this deployment reaffirmed conclusions made in Appendix A, namely that the trigger efficiency

and electronic noise measured with this system were insufficient for dark matter searches or for sensitivity to the L-capture line of ^{68}Ge and would therefore require an update in hardware for improvement.

These conclusions helped lead to the deployment of another lower-background P-PC with improved triggering capabilities and the ability to measure the rise-time of preamplifier traces. As well, the tools developed for the analysis of P-PC2 would be directly applicable to the new detector. The deployment and analysis of data from this low-background P-PC are the subject of Chapters 3, 4 and 5.

Chapter 3

DEPLOYMENT AND ANALYSIS OF A LOW-BACKGROUND MODIFIED-BEGe DETECTOR

3.1 Experimental Setup

A 440 g P-PC modified **B**road-**E**nergy **G**ermanium (BEGe) detector manufactured by Canberra Industries was deployed underground on 24 August 2009 in the Soudan Underground Laboratory in Soudan, MN. The purpose of the deployment of this detector was to obtain low-threshold, low-background counting data to seek improved limits on WIMPs in the low-mass range ($\lesssim 10$ GeV). The design of this detector included several custom modifications to address backgrounds seen with a previous P-PC used to obtain dark matter limits [11], including several internal mounting components replaced by ones constructed with cleaner materials by J. Collar at the University of Chicago. A summary of relevant geometrical aspects of the detector is provided in Table 3.1. The detector was placed in the same shield as described in Section 2.1 with a few small changes. The inner OFHC copper shield was replaced by old Pb bricks from the University of Chicago and a different PMT-based muon veto was placed around the external borated polyethylene shield.

This chapter presents results from data taken with the modified-BEGe detector underground from the time period beginning 4 December 2009 until 15 May 2010 for a total live-time of 150.6 days. During this time, there were two main gaps in run time, including a roughly 5-day shutdown period during a planned power outage at the lab and another 6-day period during which parameter settings for the DAQ were modified. Results from a parallel analysis of a previous subset of the data occupying about 2 months of live-time have been presented in [50].

Table 3.1: Modified-BEGe detector characteristics.

Property	Value
Manufacturer	Canberra
Mass	440 g
Outer diameter	60.5 mm
Length	31 mm
Active Volume	83.4 cc
Capacitance	1.8 pF (@3000 V bias)

3.2 Data Acquisition and Processing

The DAQ system used was similar to the one used in [11]. The preamp of the modified BEGe included 2 signal outputs, an inhibit output generating a logic signal when the reset circuitry of the preamp was active, and a test input for waveform generator pulses. The initial system was set up without the ability to digitize raw preamp traces, but a later upgrade on 1 Dec 2009 introduced additional hardware to take this data. The readout system was a PCI-based National Instruments digitizer with 6 channels, sampling at 20 MS/s with a resolution of 8 bits. This digitizer allowed selectable voltage ranges, enabling higher resolution at lower energies. The acquisition software used was a Windows- and Labview-based program designed by J. Collar.

One of the preamp signal outputs was run into an analog spectroscopy amplifier with a 10μs shaping time. Two outputs from this amplifier were input into the digitizer with different voltage ranges: -0.05→0.05 V (high gain) and -0.25→0.25 V (low gain). The second preamp output was AC-coupled to a Phillips Scientific 777 fast (DC→200 MHz) amplifier using a capacitor to yield a ∼50 μs decay time. This stage provided a ∼ 1× gain and was used as a fan-out for two outputs to the next stage: (1) one output was input into a spectroscopy amplifier with 6 μs shaping time and then into the digitizer; (2) the other output was sent to another gain stage of the 777. The two outputs from the fan-out of this stage were then input into two channels of the digitizer with different voltage ranges:

Table 3.2: DAQ Channel readout characteristics.

Channel	Characteristics
0	High-gain (0-3.5 keV), $6\mu s$ shaping, triggering
1	High-gain (0-3.5 keV), $10\mu s$ shaping
2	Low-gain (0-14 keV), $10\mu s$ shaping
3	Muon-veto
4	High-gain, AC-coupled preamp trace (unshaped)
5	Low-gain, AC-coupled preamp trace (unshaped)

-0.05→0.05 V (high gain) and -0.25→0.25 V (low gain). The 777 allowed adjustment of the output offset, and so the baseline was set to be close to the upper maximum of the high-gain channel. This was to maximize the dynamic range for digitizing the negative-going preamplifier pulses.

The muon veto was composed of 10 flat panels situated around the outside of the polyethylene shield, with 6 on the sides and 4 covering the top. The outputs of all the panels were coupled together and reduced to one channel, effectively OR-ing all the PMTs. This single channel was then input into a discriminator with a threshold set to output logic pulses when any of the PMTs registered a single photon event and the logic output was input into the last remaining channel of the digitizer. The outputs read into the DAQ system are detailed in Table 3.2. All channels were digitized at 20 MHz with 8 bits, and each trace length was 8000 samples (400 μs) long. The level-sensitive trigger was generated by the high-gain, 6 μs-shaped, channel 0. When a pulse exceeded the trigger level in channel 0, all 6 channels were read out.

The inhibit output from the preamp was used as an online veto by splitting the inhibit signal and using the 'INHIBIT' inputs on the spectroscopy amplifiers. When the logic is active on this INHIBIT input, the spectroscopy amplifier maintains the baseline of its output, effectively removing any trigger generated by a reset of the preamp. It should be noted that, in contrast to the DAQ system described in Section 2.2, no timing informa-

tion of the inhibit pulse was retained to perform a later estimate of dead-time. Instead, during detector deployment it was found that the reset pulse rate was stable at $\lesssim 3$ Hz, suggesting that differences in temperature did not affect this setup as profoundly as P-PC2 (see Section 2.3.4) or that the detector cold finger was more properly coupled to the liquid nitrogen. This stability was further confirmed by a measurement of the baseline vs. time detailed in Section 3.5.1. A 1 ms veto following the pulse reset then produces a less than 0.3% reduction in live-time, which is ignored.

The digitizer card maintained an internal buffer to store a set of events. After 20 events were stored, data from the digitizer buffer were written to disk and file names were cycled (open file closed, saved, and new file opened) every 3 hours. An automatic data management chain was setup similar to that described in Section 2.3.1 with a run database used to facilitate the data processing. Files were synchronized back to a server at the University of Washington where they were persisted on RAID-ed disks. Once a file appeared on the UW server, a corresponding record was introduced into the database and this record was used to control and track further processing. The processing progressed in a tiered fashion as follows:

Tier 0: Raw data from the Labview DAQ system

Tier 1: ROOTified data - raw data converted to MGDO objects and stored in ROOT TFiles

Tier 2: Waveform processed data - extraction of waveform characteristics using MGDO Transforms

Details on the waveform processing are given in the following Sections: 3.2.1, 3.2.2, and 3.2.3. For more details regarding the MGDO objects and the framework of the analysis chain, see Appendix B.

3.2.1 Shaped Channel Processing

The shaped channels (0, 1, and 2) were processed to extract amplitude information of the pulses. These traces were first run through a 100 kHz low-bandpass filter to remove high-frequency noise and artifacts from the limited bit-depth of the digitizer. Both extrema

values, maximum and minimum, of each waveform were recorded and the baseline of each waveform was calculated by averaging the first 280 μs (5600 samples). Maxima were calculated both before and after the bandpass filter and both values saved; the minimum value was found on the unfiltered waveform. The amplitude could then be calculated as the difference between the baseline and the maximum measured on the filtered waveform. The minimum was saved to analyze pulse health later.

3.2.2 Unshaped Channel Processing

The unshaped channels (4 and 5) were processed to extract information on the characteristics of the waveform, including baseline, extrema values, and rise-time information. The baseline was calculated using the first 5600 samples and the maximum and minimum were found for each pulse. Calculating the rise-time of each pulse required de-noising using a time-invariant stationary wavelet transformation (SWT). This process is described later in Section 3.4.3. Extracting the extrema values – the maximum and minimum of the waveform – required running the waveform through a 100 kHz low-bandpass filter. The filter was run on an unmodified waveform and *before* any other de-noising was done. After the application of this filter, the maximum and minimum values and positions were recorded.

3.2.3 Muon-Veto Channel

The muon-veto channel digitized the logic output from a muon veto. To process these waveforms, all the regions of each trace were found when the veto was logic positive. Saving this information enabled any cut based upon the muon veto to be performed further down the analysis chain, retaining flexibility in determining parameters for such a cut. However, due to the high count rate of the muon veto and the reduction in live-time any cut from this veto would create, it was decided to not use any cuts based upon results from this channel in the analysis. For example, the muon veto fired at a rate of 6200 Hz – the threshold was set to trigger on single photon events – and would therefore generate a reduction in live-time given by $1 - \exp(-[6200 \text{ Hz}]\tau_{veto})$ where τ_{veto} is the length of time to reject after a muon veto. Assuming $\tau_{veto} = 100$ μs, the reduction would be 46%.

3.3 Data Analysis

The results of several key measurements were used throughout the analysis process, in particular energy calibration, the resolution of the detector versus energy, and the efficiency of the trigger. Additionally, the fitting of the energy spectrum – fitting function described in following section – was key to the extraction of several parameters of interest. These results and procedures are detailed in the following section and referred to throughout the remainder of the chapter.

3.3.1 Fitting Energy Spectra

Many of the results required fitting the energy spectrum of the data to measure a particular parameter or set of values. To accomplish this, the RooFit toolkit [51] was used to develop a general PDF which could be consistently used. The RooFit toolkit has the advantage of allowing the creation of general PDFs which may be fit to data using both binned and unbinned maximum likelihood fits, or by the minimization of the χ^2 function. The spectral fitting was performed using either binned or unbinned maximum likelihood. The general PDF for the data was constructed:

$$b_1 \exp\left(c_1 E\right) + b_2 + \frac{b_3}{2} \operatorname{erfc}\left(\frac{E - \mu_{Ge}}{\sqrt{2}\sigma_{Ge}}\right) + \sum_i^{xrays} \frac{a_i}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{(E - \mu_i)^2}{2\sigma_i^2}\right) \tag{3.1}$$

where b_1, b_2 and b_3 are the exponential, flat, and low-energy (erf) flat background amplitudes respectively, c_1 is the exponential constant and the sum is over the x-ray lines present in the fit (see Table 3.3 for a list). The amplitudes of all components (a_i, b_1, b_2, b_3) and the exponential constant were allowed to float independently. This equation does not explicitly include normalization terms, but each component was appropriately normalized automatically by RooFit so that the amplitudes corresponded to counts present in each fit component simplifying the interpretation of fit results. The parameters (μ_i and σ_i) of the x-ray lines were allowed to float in a small range around their theoretical values. The float range was determined so that the final fit value was not near a range boundary.

The error function parameterizes the 'plateau' below the Ge K-capture peak present in

52

the data due to partial charge collection as indicated by previous results [52]. To summarize, another measurement using a P-PC observed the decay of ^{71}Ge (11.4 day half-life) in three separate regions: the Ge K-capture line (10.367 keV), the Ge L-capture line (1.3 keV), and the flat region in between ($2 \to 6$ keV). The results found that the decay of count rate in the flat region matched the decays of count rates in the L- and K-capture regions, suggesting partial energy deposition from the ^{71}Ge decay below 10.367 keV. In general, sets of these error functions should be included for each prominent x-ray line, but since the Ge K-capture line dominates, only one error function centered on the Ge K-capture line was included. The parameters for the error function, μ_{Ge} and σ_{Ge}, were defined as 10.367 keV and 0.1 keV, respectively.

The exponential function was included in the fit function because the data exhibited a general exponential shape. Essentially, this shape parameterizes the measurement of counts near threshold. A discussion of the origin of this shape in the data follows in Section 3.4.5.

Since some applications of this PDF would not require a fit over the entire entire range, the function was updated appropriately, adding or removing components as necessary. For example, a fit over the range $0.5 \to 3.5$ keV would only include the exponential- and flat-background components plus the two L-line components. A corresponding fit over a larger range (e.g. up to 10 keV) would also include the x-ray lines within that range. Additionally, the state of a variable could change, for example, by either being set to constant or allowed to float over a wider range than normal. Any deviations from the prescription outlined here used in a particular analysis will be noted.

3.3.2 Triggering Efficiency

The trigger efficiency of the DAQ system was measured by scanning a pulser of known amplitude across the threshold. This measurement determined the ability of the DAQ electronics to trigger at certain signal amplitude given the noise characteristics of the preamplifier and the readout electronics. The data were then fit to the function:

$$1 - \frac{\mathrm{erf}\left(s(V - \sigma)\right)}{2} \tag{3.2}$$

Table 3.3: Summary of prominent x-ray lines in the data set.

Isotope	Energy
^{65}Zn L-capture	1.1 keV
^{68,71}Ge L-capture	1.299 keV
^{65}Zn K-capture	8.979 keV
^{68,71}Ga K-capture	9.659 keV
^{68,71}Ge K-capture	10.367 keV
^{73,74}As K-capture	11.103 keV

where V is in volts, s is a scaling parameter and σ is a position parameter. This yielded the results: $\sigma = 7.137 \times 10^{-3} \pm 8 \times 10^{-6}$ V and $s = -1.772 \times 10^{3} \pm 28$ V^{-1}. Results are shown in Figure 3.1. This study was done during initial deployment and not throughout the experiment due to concerns that noise or stray signals from the pulser could negatively affect the performance of the detector. However, results from the previous deployment of a P-PC detector and the monitoring of detector parameters versus time (Section 2.3.3) suggested that other parameters could provide information to probe whether or not the trigger efficiency changed over time. Such parameters include waveform characteristics (baseline, extrema) and triggering rates and these are discussed in Section 3.5.

3.3.3 Energy Calibration

Energy calibration at low energies is complicated by several factors. In particular, low-energy x-ray peaks (< 10 keV) from any source will be heavily attenuated by the source itself as well as materials between the source and the crystal, including the outer cryostat, mounting components, and the crystal dead layer. Higher energy x-rays have a larger probability to interact in the crystal, but these are unsuitable for calibration because of their distance from the signal energy region which, in the case of dark matter, is close to threshold. However, internal cosmogenic isotopes provide excellent candidates for calibration since there are several with characteristic lines near or below 11 keV. Therefore, the energy
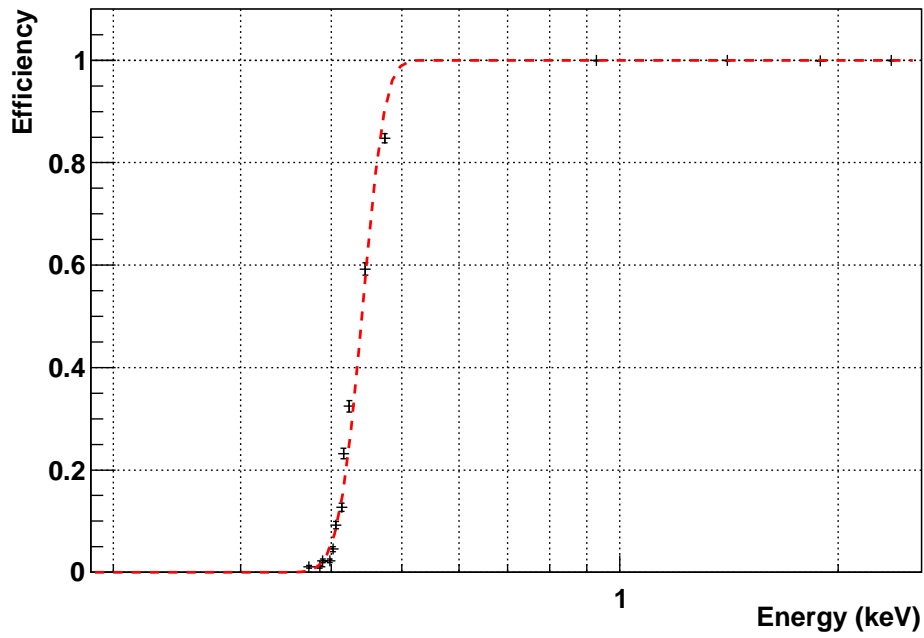
Figure 3.1: Modified-BEGe triggering efficiency measured with a pulser. Error bars are binomial and the fit is to the error function in Equation 3.2.

calibration was determined by fitting the low-gain channel simultaneously to the peaks listed in Table 3.3 yielding a linear equation:

$$E_{ion}(keV) = aV + b$$

with $a = 63.81 \pm 0.25$ (keV/V) and $b = -0.014551 \pm 0.016$ keV. These results were applied to both the low- and low-gain channels. An example of a typical fit spectrum is shown in Figure 3.2 and in other figures throughout the remainder of the chapter.

3.3.4   Resolution of Results

The resolution of the detector was determined by measuring the intrinsic electronic noise using a pulser and by measuring the widths of x-ray lines. These widths are measured by performing a simultaneous fit to prominent x-ray lines. The fitting function is a combination of Gaussians for each x-ray line, a flat background component, and a second 'plateau' background component below the x-ray lines parameterized by an error function equivalent to Equation 3.1 described earlier in Section 3.3.1. The fit is shown in Figure 3.2. The results of these measurements were folded into the following equation:

$$\sigma = \sqrt{\sigma_{elec}^2 + E\eta F} \tag{3.3}$$

to determine the intrinsic electronic noise σ_{elec} and estimate the Fano factor F. E is the energy in keV, η the amount of energy required to generate an electron-hole pair (2.96 eV). σ_{elec} was found to be 70.48 ± 0.54 eV and F was estimated as 0.241 ± 0.013. Results are shown in Figure 3.3. As expected, the electronic noise measured with a pulser is independent of energy. The scatter of the x-ray measurements is larger than statistically expected, in particular the ^{65}Zn and ^{68}Ga exhibit σs well away from the best-fit line. This is most likely due to insufficient bit depth of the digitizer and that this resolution is not properly included in the error bars[1]. The low-gain channel has an intrinsic resolution of 1.95 mV/ADC, or ~ 125 eV/ADC. The actual resolution of the measurements estimating the amplitudes of

---

[1]The resolution due to the bit depth of the digitizer (i.e. keV/ADC count) should be small compared to the resolution of the measured process and so a future DAQ upgrade would benefit from such an improvement.
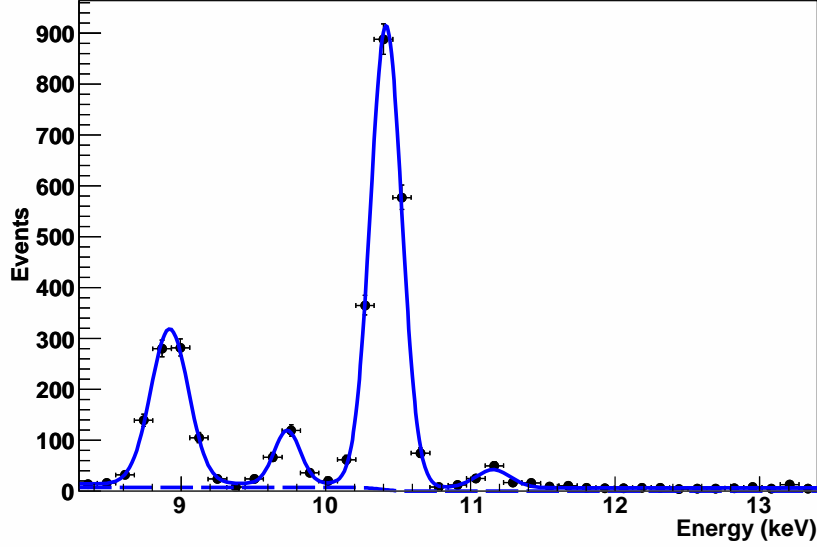
Figure 3.2: Fit to estimate the resolution of the modified BEGe, including Gaussian fits to some of the prominent lines listed in Table 3.3.

pulses is better due to the fact that many samples are averaged together for that calculation. However, the main purpose of this equation is to estimate the resolution of the x-ray lines at low-energy, around 1 keV, and it is clear that small deviations to this measurement do not significantly affect results in this region.

3.4 Data Cleaning and Cuts

Several cuts were employed to clean the data, removing spurious events from electronics noise or microphonics. Additionally, other cuts were employed to remove events with slow rise-time. This section discusses the implementation and results of *all* cuts performed on the data set.

3.4.1 Microphonics and Noise Cuts

Vibrations in detector components at different electric potentials, such as the cryostat, cold-finger, or crystal mount, can induce electronic signals due to the changing capacitance
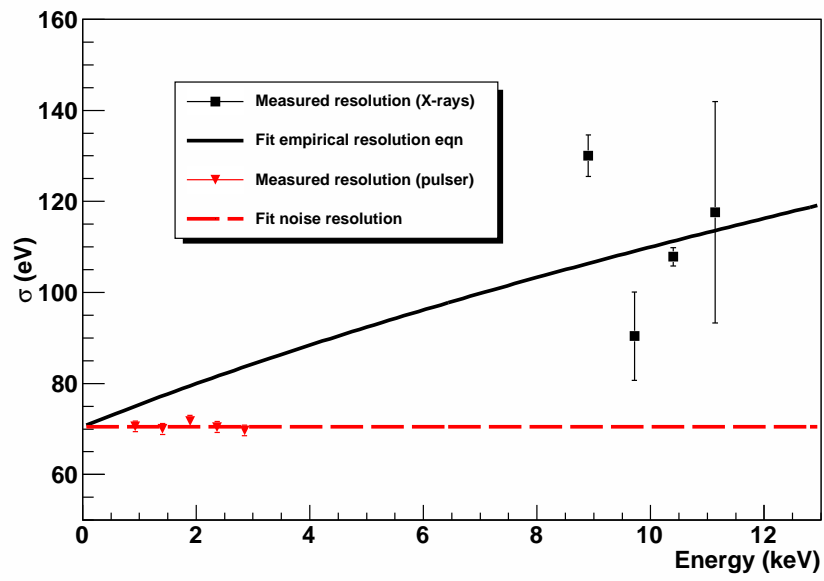
Figure 3.3: Modified BEGe resolution versus energy. Red dashed line is measured and fit intrinsic electronic noise, black points are measured resolution from x-ray lines, black line is a fit to Eqn. 3.3.
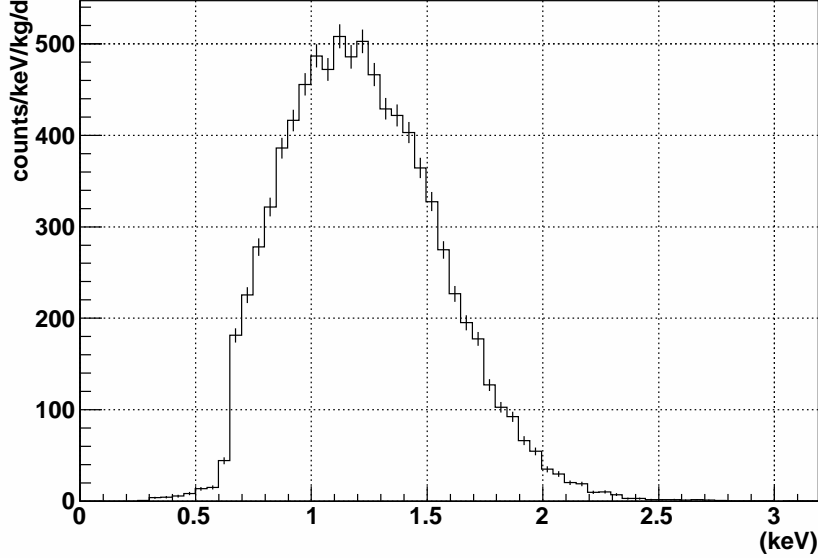
Figure 3.4: Energy spectrum during LN fills. No microphonics cuts were applied to these data.

created by the movement. These electronic deviations can generate extra noise at low energies (threshold to a few keV) possibly obscuring any signals in that region. During LN filling of the dewar, vibrations due to the filling generate a significant amount of additional microphonics noise from vibration caused by the influx of new liquid nitrogen. See Figure 3.4 for an energy spectrum of events occurring during an LN fill. Therefore, during these filling periods, a flag was set in the data to enable later removal of events occurring during a fill. This flag was then lowered 5 minutes after the fill completed, resulting in a veto time of ∼15 minutes. LN fills occurred every 2 days, meaning that they occupied 0.5% of the run time.

For microphonics induced from sources other than LN fills, another cut was made. Morales et al. developed a technique to mitigate this class of events, taking advantage of the fact that microphonics tend to have characteristics (e.g. rise-time, fall-time, baseline shift) significantly different from events arising from charge collection in the crystal [38]. This procedure analyzes the ratio of amplitudes from two signal channels with different

shaping times and accepts or rejects events based upon their deviation from the expected ratio. This expectation can be determined by using a source or a pulser at low amplitudes; in a setup where the amplification of the two shaped channels is nearly equivalent, the expectation value of the ratio is close to 1.

In this application, a pulser was used to train the microphonics cut by taking high-statistics pulser runs at several discrete amplitudes near the threshold. At these discrete amplitudes, the ratios of the two channels were histogrammed and fit to a Gaussian. The cut points were then determined by taking the values $\mu \pm 3\sigma$ to accept 99.7% of the pulser events. Since these cut points were estimated at discrete amplitudes, a 4th-order polynomial was fit to them to interpolate between the points. The results of this are shown in Figure 3.5. To avoid errors from the interpolation, the cut was softened forcing the upper limit to be greater than 1, and the lower limit to be no less than 0.8. The final cut is then shown in Figure 3.6 which shows an overlay of the calculated cut on data from a scanned pulser run (Figure 3.6(a) and from the run data (Figure 3.6(b)). Interesting to note is how the run data and the scanned pulser data follow a different distribution at energies above 1 keV, in particular that the scanned pulser data is centered around a ratio of 0.9 and the run data is centered at a lower value. This suggests that the difference between a pulser and an actual crystal event become more significant at these energies: future calibrations should use a source instead of a pulser to ensure complete consistency. Interestingly, it appears that there may exist a crescent-shaped distribution of events that runs through the low-energy acceptance region (beginning below threshold at high ratios and ending from $1 \rightarrow 2$ keV at low ratios). Whether or not such a distribution really exists has not been determined, but will be considered as a possible background in this low-energy region.

3.4.2   Electronics Cuts

This class of cuts was intended to remove events coming from spurious electronic events which were distinctly not signal events. Only two electronics cuts were applied: the first to remove negative-going pulses including uncaught reset events that don't trigger the active veto, the second to remove noise pulses that were found to occur during a particular interval
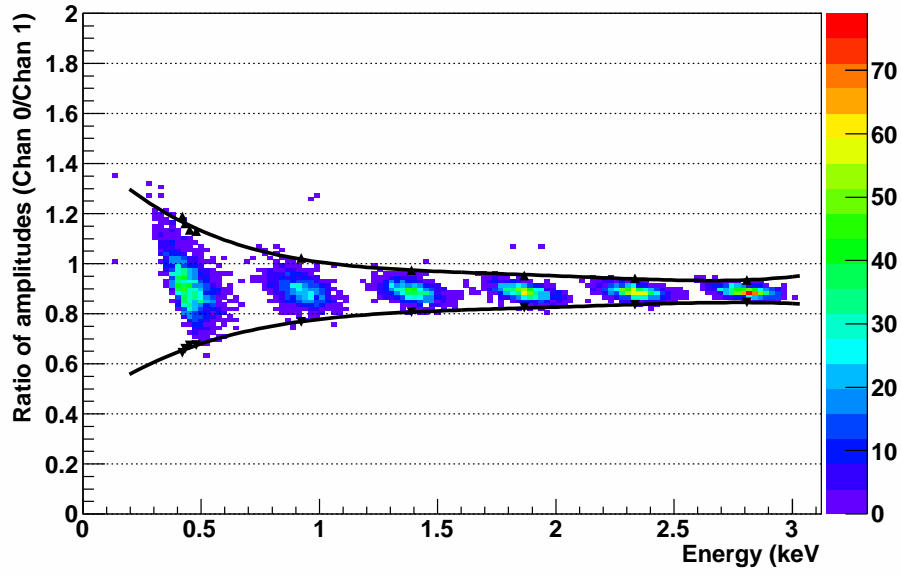
Figure 3.5: Calculation of microphonics cuts using the ratio of outputs from two different spectroscopy amplifiers with an input pulser at discrete amplitudes. Line is an estimate of the cut, see text for details.
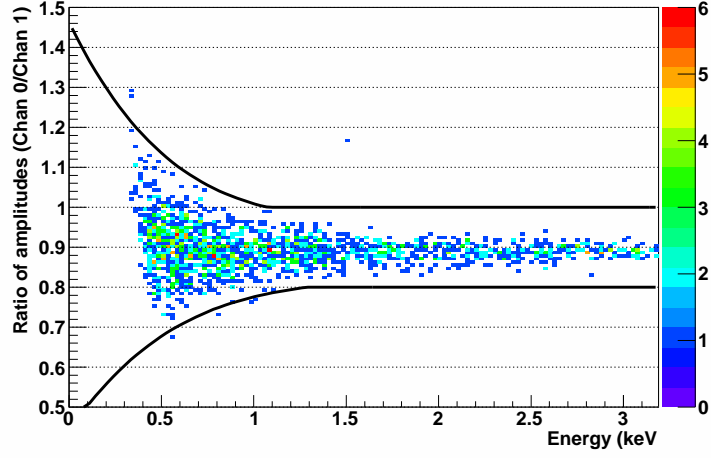
(a) Scanned pulser run
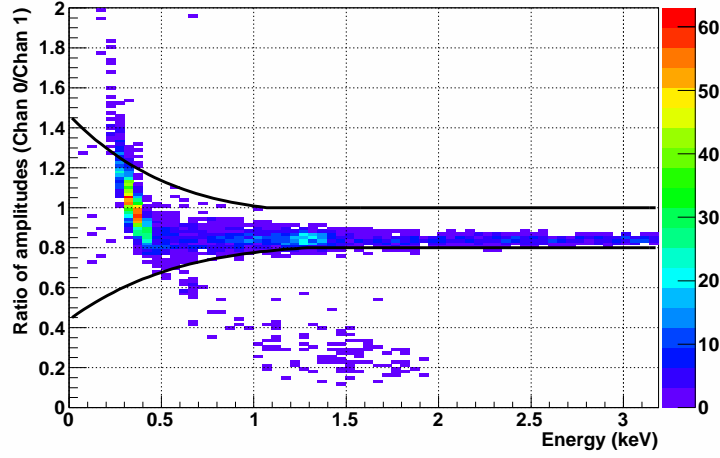
(b) Data

Figure 3.6: Ratio of two shaped channels versus energy for a scanned pulser run and for run data. The drawn line is the cut at 99.7% efficiency, calculated using high-statistics pulser runs.

of the run time between 15 February 2010 and 15 March 2010.

The first cut removed events with a minimum in the shape channel below a certain voltage value chosen to be -0.02 V. The population of events excluded by this cut included uncaught reset pulse events, negative-going events, and large energy depositions resulting in significant undershoot on the baseline. Most of these events also saturated the digitizer positively, but many of the negative-going events did not. A full explanation for these events was not determined though it is possible that some may arise from breakdowns of the crystal voltage. From this cut, only 184 events were removed over the 150.375 day run period, generating a negligible reduction in live-time.

While running tests analyzing the rates of events in different energy regions (see Section 3.5.2), a class of events was found that occurred only during a sub-interval of the run time from 15 February 2010 until 15 March 2010. These events were distinctive from true events, but were problematic because they populated the spectrum near threshold around 0.5 keV. An example of such a pulse is given in Figure 3.7. These were distinguishable by comparing the difference between the extrema (maximum - minimum) of the unshaped waveforms. A plot of this parameter versus energy is given in Figure 3.8 where the dashed line denotes the region of parameter space populated by these pulses.

To study the origin of the pulses, a rate analysis similar to that performed in Section 3.5.2 was applied to the subset of events. The time between these events was histogrammed and fit to an exponential; the results are included in Figure 3.9. This revealed that the events arrived not with a defined frequency, but rather in a Poisson fashion with a rate estimated from the fit as 2.97 ± 0.1 counts/hour. This can be compared to the normal rate for this energy region $0.5 \rightarrow 1$ keV of ~ 0.124 counts/hour (see Section 3.5.2). The structure of the pulse suggests an induced signal possibly coming from a noise source capacitively-coupled to the signal lines. It is likely that these pulses arose from some cross-talk issues from a separate, independent DAQ system that was connected to the detector in parallel at the time since the removal of this parallel DAQ system corresponded exactly to when these pulses stopped occurring. This class of events was entirely removed using the discriminating line in Figure 3.8.
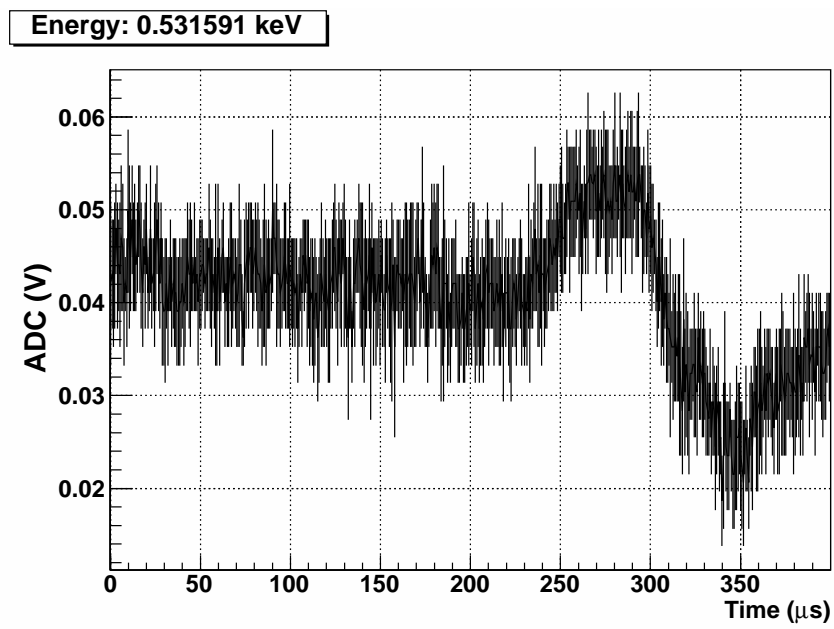
Figure 3.7: Example of noise pulse of unknown origin. The energy of the pulse is specified in the figure: 532 eV. This indicates that these events can populate the spectrum near threshold.
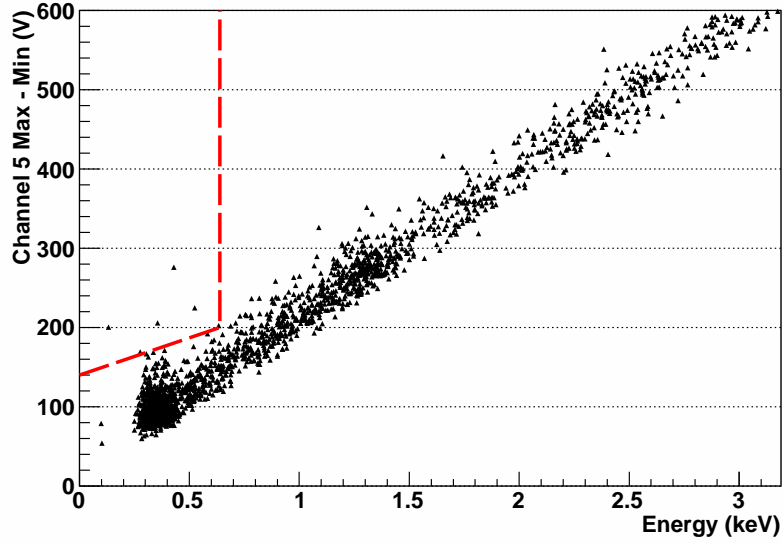
(a) Data around time of interest
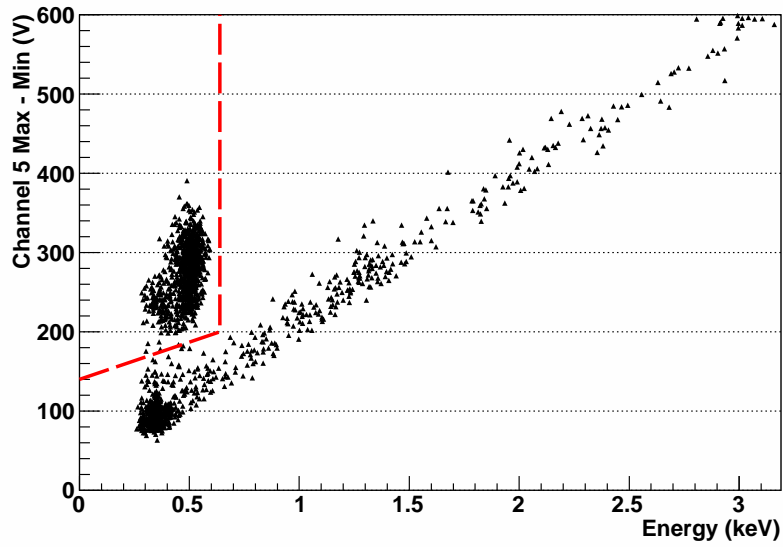
(b) Data during time of interest

Figure 3.8: Difference in unshaped waveform extrema (maximum - minimum) versus energy. The dashed line is an indication of the cut used to exclude the noise events removing all events within the selected region. A plot of the data around the time of interest show that only $O(1)$ 'true' events are removed by this cut.
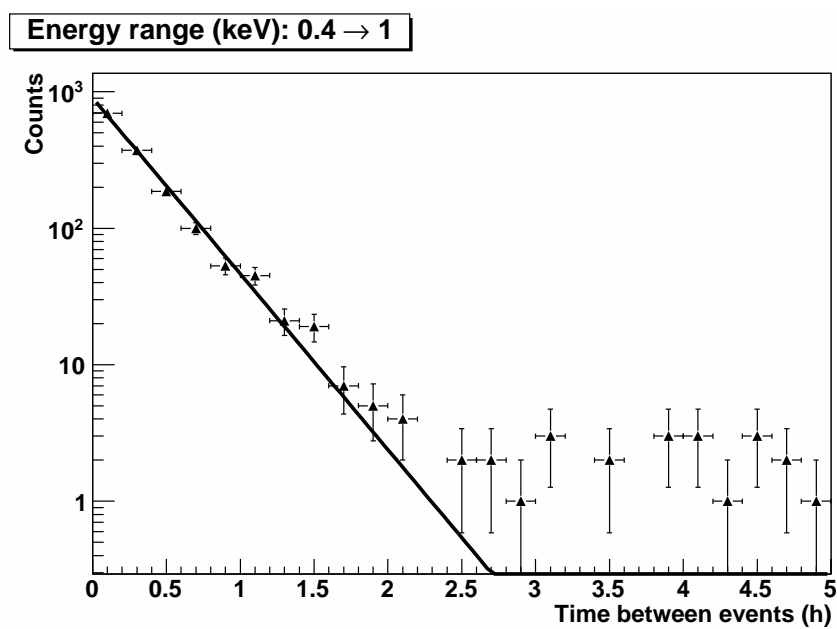
Figure 3.9: Time between events in the energy region $0.4 \rightarrow 1$ keV including noise pulses described in the text. The line is a fit to an exponential. The events at longer time intervals are 'contaminations' from non-noise events arriving at a much slower rate.

3.4.3 Rise-time Cuts

Pulses of slow rise-time were found during the deployment of P-PC2 (see Section 2.3.5). These pulses were predominately at low energy near threshold which meant that they could compose a possible background to any signal in this region. This section studies the methods developed to measure the rise-times at low signal-to-noise ratios against threshold and explores systematics related to a cut on this quantity. Additionally, the likely origin of these pulses and possible tests for further exploration are discussed.

Wavelet De-noising

As the amplitude-to-noise ratio of waveforms shrinks at low energies, calculating the rise-time becomes more sensitive to the magnitude of the noise (i.e. the electronic fluctuation of the signal) due to the reduction in signal-to-noise. De-noising via a simple bandpass filter is undesirable when both signal and noise are distributed across similar frequency bands as the signal-to-noise ratio will not be enhanced by removing particular frequencies. When calculating the rise-time of a pulse, a bandpass filter can greatly attenuate the high frequencies present in the rising edge of the pulse. Wavelet shrinkage provides methodology to reduce noise on a generic function (see e.g. [53, 54]) when the function and noise occupy the same frequency space. The algorithm follows:

1. Choose a wavelet basis.

2. Perform a wavelet transformation using the chosen basis to a level n, obtaining n sets of detail and approximation coefficients.

3. Apply thresholding to the detail coefficients.

4. Perform an inverse transformation.

In this particular application it is necessary to use a translation-invariant version of the wavelet transformation called a Stationary Wavelet Transformation (SWT) (see [55, 56]). The SWT performs transformations at all possible translations for a given data set and basis
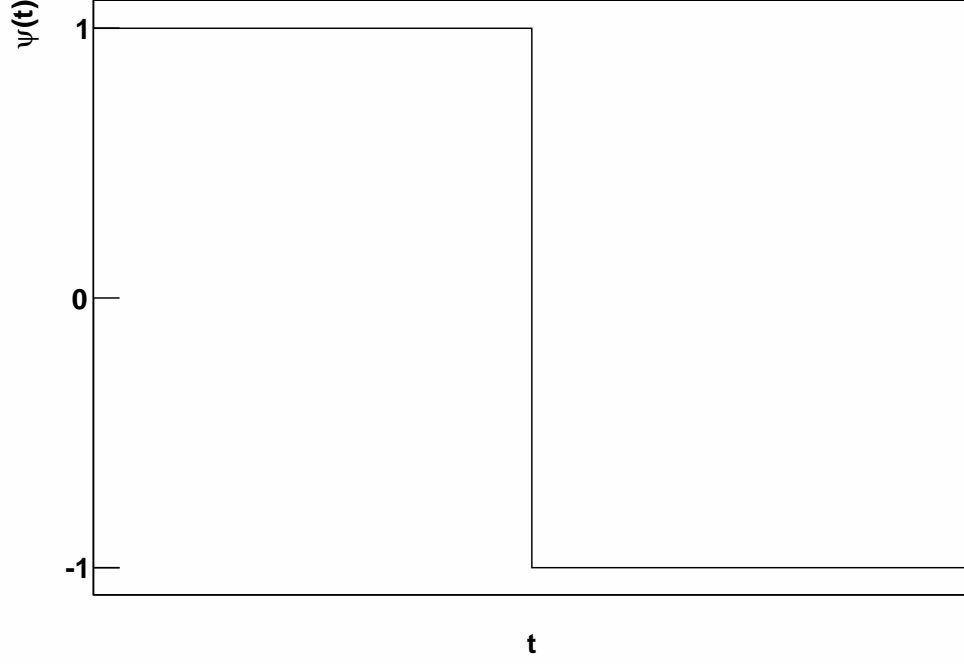
Figure 3.10: An example of a Haar wavelet. This base wavelet may be scaled and translationally shifted to compose a complete basis set.

wavelet. A subsequent inverse SWT effectively averages these together, avoiding artifacts induced by any chosen time origin of the waveform. Examples of artifacts induced by using a origin-dependent transformation instead can be found in [55, 56].

For this wavelet analysis, the python package PyWavelets [57] was used. Since an implementation of the inverse SWT was missing from this distribution, the necessary extension to the package was written (see Section B.1). A Haar wavelet (see Figure 3.10) was chosen as a basis wavelet due to its simplicity and asymmetry. In this case, the asymmetry of the wavelet is desirable because the signal (an unshaped preamplifier trace) exhibits the same characteristic. Thresholds for each set of detail coefficients, $D_n^{(i)}$, were calculated using a pure-noise waveform training set of length j. For each noise waveform $x^{(i)}$, a 6-level SWT was used to generate $D_n^{(i)}$ and the thresholds were calculated for each level n according to

the equation proposed by Donoho and Johnston [58]:

$$\tau_n^{(i)} = \sigma_n^{(i)} \sqrt{2 \log N^{(i)}} \tag{3.4}$$

$$\sigma_n^{(i)} = \frac{\text{MAD}\left(D_n^{(i)}\right)}{0.6745}$$

with $N^{(i)}$ the length of waveform, $x^{(i)}$, and MAD is the median average deviation. The threshold at a level n was then defined as $\tau_n = 0.8 \max(\tau_n^{(0)}, ..., \tau_n^{(j)})$. An example of the coefficients calculated using a 6-level SWT is shown in Figure 3.11. This figure also includes the thresholds calculated at each level, denoted by dashed lines.

Noise reduction was implemented by applying hard thresholding to each set of detail coefficients. In this technique, all coefficients D_n with an absolute value less than the threshold τ_n were set to zero. Coefficients above this threshold value were unchanged. The resultant coefficients were then used in an inverse SWT to produce a de-noised waveform. An example of the wavelet de-noising is shown in the top figure of Figure 3.12.

Rise-time Calculation

The de-noising process ensured that the waveforms be ready for rise-time calculations. After de-noising, the smoothed derivative of the waveform was generated using a Savitzky-Golay derivative filter [59]. The extremum of the derivative - in this case the minimum since the pulse was negative-going - was then found and used to determine the middle of the rising edge, p_m. The full-width at half maximum (FWHM) was calculated and used to estimate the beginning and end of the rise of the waveform: the beginning, $p_b = p_m - 1.5 \times \text{FWHM}$ and the end, $p_e = p_m + 1.5 \times \text{FWHM}$. The baseline and amplitude of the pulse were each found by averaging over 1 μs (20 samples) beginning at $p_b - (1~\mu\text{s})$ and p_e, respectively. These values were used to estimate the amount of time it took the pulse to rise 10%$\rightarrow$90% in amplitude. Linear interpolation was used to refine the time values which came between digitization points. An example of this calculation, employing the same pulse used to generate the wavelet decomposition in Figure 3.11, is shown in Figure 3.12.
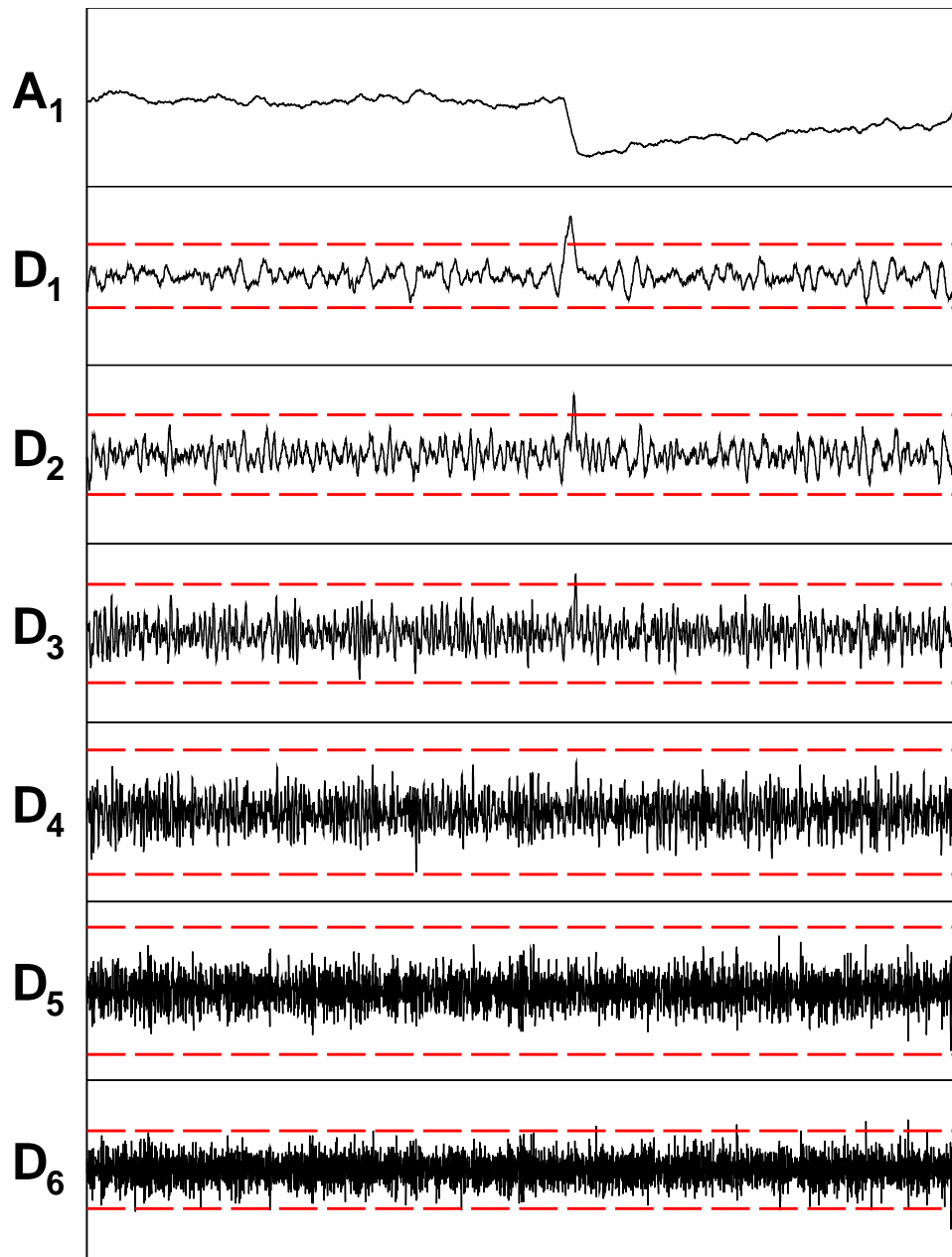
Figure 3.11: Example wavelet decomposition of pulse in Figure 3.12. A_1 denotes the first-level approximation coefficients, the D_n denote the detail coefficients at the nth level of the 6-level stationary wavelet transformation. Dashed lines indicate the thresholding to be used for each set of detail coefficients.
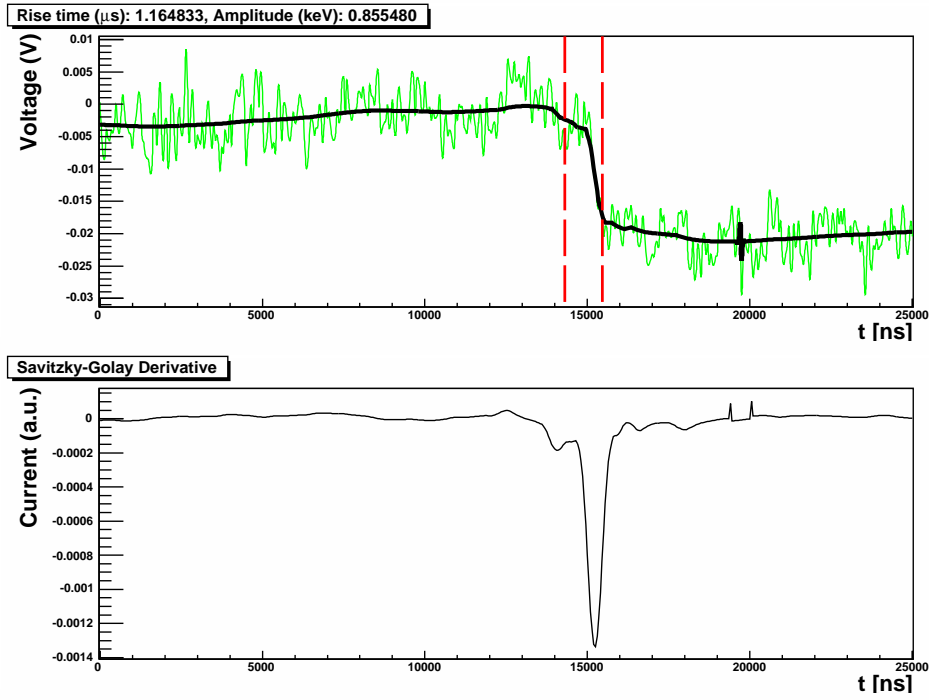
Figure 3.12: Example of rise-time calculation technique applied to a preamp trace. The top shows the raw (green) and de-noised waveforms (black), and the vertical dashed lines represent the result of the rise-time calculation. The bottom plot shows the smoothed derivative of the trace calculated using a Savitzky-Golay filter [59] of degree 2, width 6.
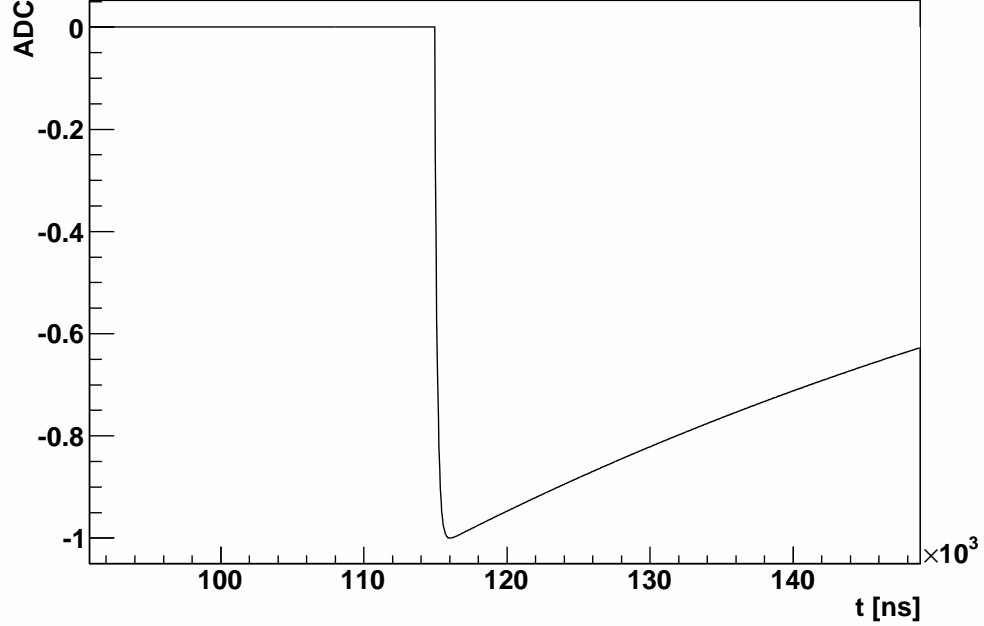
Figure 3.13: Example of a simulated pulse before the addition of noise.

Rise-time Simulation

To investigate how a cut based upon rise-time affected the spectrum, it was necessary to perform a simulation of the rise-time calculation on waveforms similar to the data. The idea was to produce waveforms with similar characteristics (i.e. rise-time, noise) as seen in the detector and run them through the same algorithm used to analyze the detector data. The waveform was generated by taking a tail pulse of 0 rise-time and running it through a digital low-pass RC filter. The RC constant in the filter was tuned to reproduce the rise-time of "fast" pulses seen in the data. In general this would not precisely reproduce all the characteristics of the detected pulses, but since the only parameter of interest was the rise-time it was reasonable to choose such a simple pulse construction. An example of a simulated pulse before noise was added is given in Figure 3.13

The electronic noise of the detector was measured by looking at the baseline of all pulses and taking the average of power spectra for each preamp trace channel. Once the

average power spectrum was determined, it was possible to use this to add noise to the simulated pulse through the techniques outlined in [60] by Wan Chan Tseung. Essentially, a measurement of an average power spectrum, $\Omega = X^2 + Y^2$ where X and Y are the real and imaginary components of the Fourier Transform, gives you an average value μ_i at a frequency bin i. If there is no phase information in the noise (i.e. $\tan^{-1}(X_i/Y_i)$ is flatly distributed), X_i and Y_i are Gaussian-distributed variables around 0 with the same standard deviation, σ_i, which is related to the average value of bin i via $\mu_i = 2\sigma_i^2$. Therefore, for each simulated pulse, a noise waveform was generated in frequency space, transformed to the time domain using a discrete inverse Fourier Transform, and added to the original simulated pulse.

Since the energy of each event was determined using the amplitude of shaped pulses, it was necessary to determine a relationship between amplitudes of the shaped and unshaped low-gain channels (channel 2 and channel 5) and the amplitudes of the shaped and unshaped high-gain channels (channel 1 and channel 4). This was done by fitting the relationship from data, an example of which is shown in Figure 3.14. Additionally, the waveforms' position in the trace window exhibited a dependence on the energy of the event: for events of smaller amplitude the trigger tended to arrive later, so that the waveform moved left in the trace window. This dependence was measured by tracking the start of the pulse in the trace window versus amplitude and fitting it to an empirical polynomial (see Figure 3.15). This information was folded back into the simulation to control the starting point of the pulse given its amplitude.

Once the simulated pulses were generated, they were run through the same analysis chain as the waveforms from the detector, in particular through the rise-time calculation algorithms described earlier in this section. The amplitudes of the pulses were sampled according to the triggering efficiency measured in Section 3.3.2. Two simulations were run, one for each the high- and low-gain set of channels, for $\sim$5M events. These results were then used to calculate contours of particular acceptances, 20, 30, 40, 50, 60, 70, 80, 90, 95, and 99%. To calculate the contour, the data were binned in a 2-dimensional histogram with energy bin sizes for the low- and high-gain channels, respectively, 1.6 eV and 0.4 eV, and time bin sizes of 5 ns. Slices of the histogram were then taken at each energy bin and the
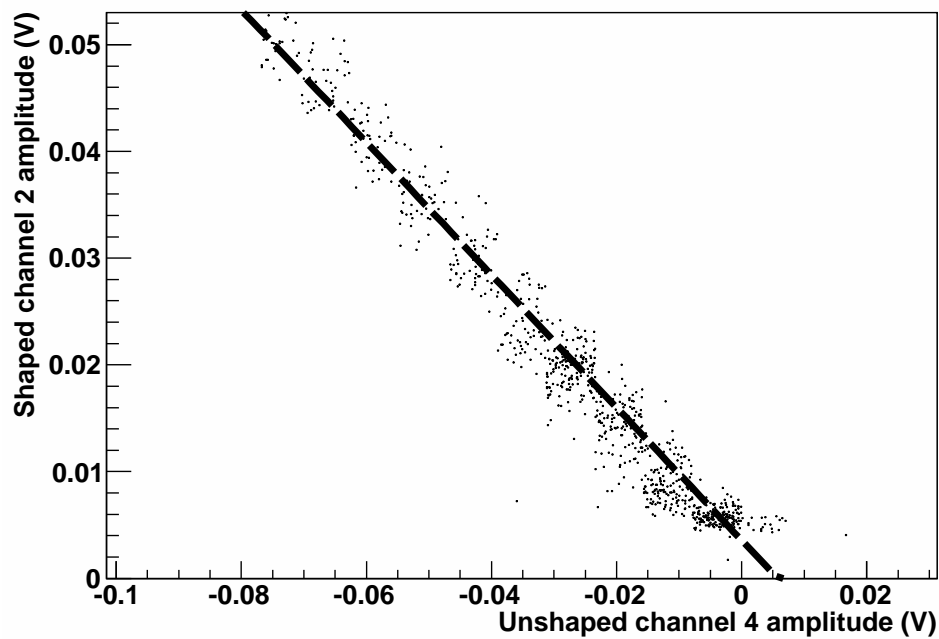
Figure 3.14: Comparison between the amplitudes in the unshaped channel 4 and shaped channel 1. The line is a linear fit to the data. The spectroscopy amplifier output positive-going pulses whereas the preamplifier output negative-going pulses which explains why the slope is negative. The voltage ranges on both the x and y axes have been made large enough to include events from the entire energy range.

.
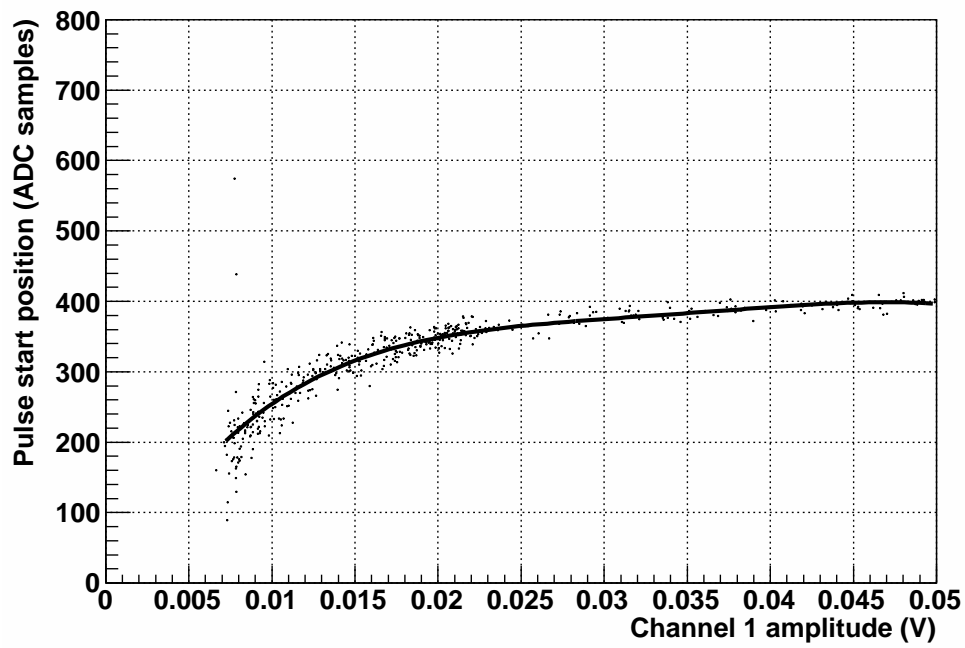
Figure 3.15: Comparison between the start of the preamp pulse rise-time and the amplitude of the shaped channel. The line is a fourth-order polynomial fit which is used to parameterize this relationship, the order of the polynomial was chosen 'by eye'.

.
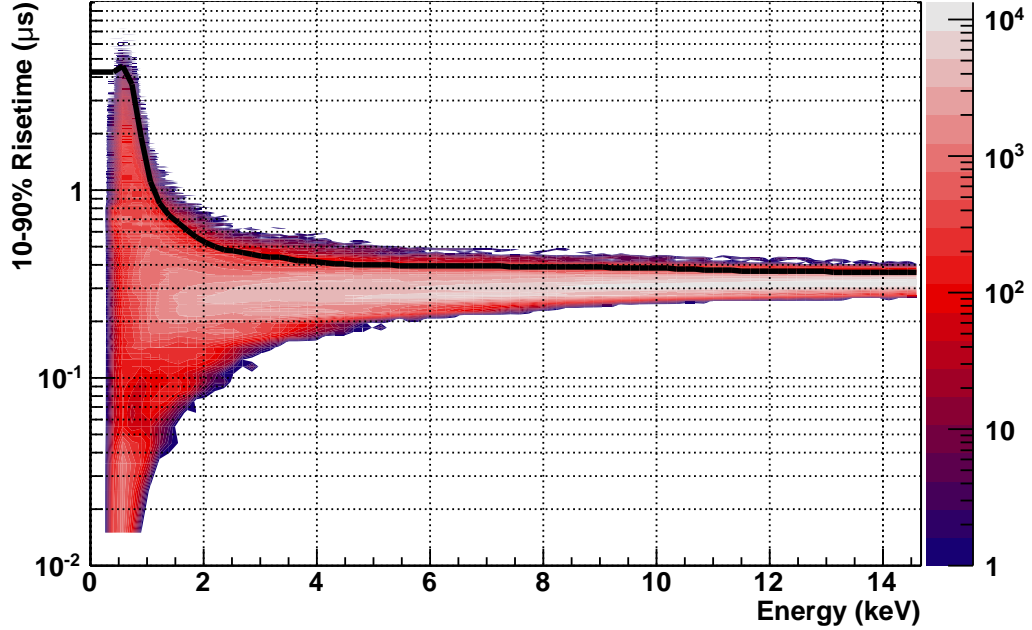
Figure 3.16: Simulation of calculated rise-time for the low-gain channel of the modified BEGe. The line is the calculated 99% contour, which includes 99% of events below the line.

upper limit of time calculated which accepted the required percentage of events. Results of the simulation for the low-gain channel are shown in Figure 3.16 along with a calculated 99% acceptance contour. The acceptance region is defined as all waveforms having a calculated rise-time less than or equal to the calculated contour. A comparison of the calculated line to data for both high- and low-gain channels appears in Figure 3.17. The acceptance lines shown in both of these plots exhibit an increase as the energy nears threshold due to the reduction of the signal-to-noise making it more difficult to reliably extract the true rise-time.

Rise-time Systematic Tests

Systematic tests of the cuts were performed to verify the simulation performed adequately and that the cuts behaved as expected. To do this, cuts of different acceptance were applied to the data set, and unbinned extended maximum likelihood fits to the cut data sets were
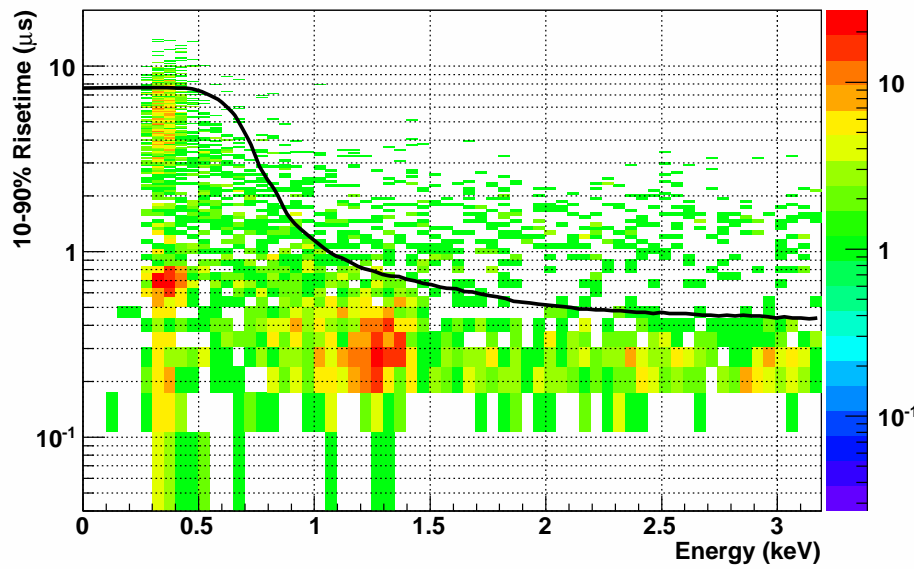
(a) High-gain channel
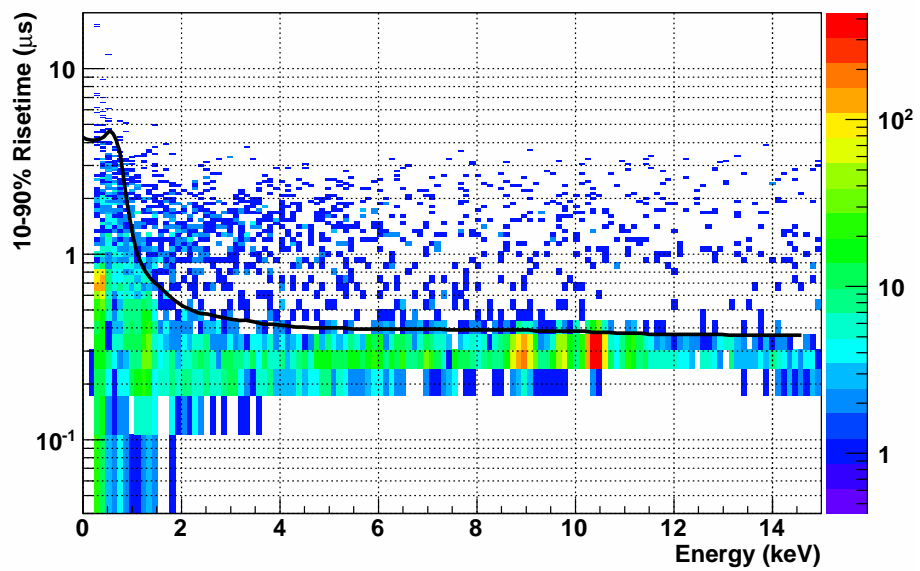
(b) Low-gain channel

Figure 3.17: Measured rise-time versus energy. The line is a calculated rise-time cut of 99% acceptance, removing events that fall above the line.

performed to measure how the cut affected different features in the data. The fit function used in this analysis is described earlier in Section 3.3.1. These fits were performed for data sets with different cuts applied as noted:

1. only LN fills

2. LN fills and microphonics cuts

3. LN fills, microphonics cuts, and rise-time cuts varying between 20% and 99%.

In the following, fits with only LN and microphonics cuts applied are labeled as 'LN + micro' and those with a rise-time cut applied are labeled using the rise-time acceptance percentage. Fits were performed for both low- and high-gain channels, an example of a fit for each channel is given in Figure 3.18. A calculation of the relative percentage remaining for each component was made using the cut LN-fills-plus-microphonics as a reference. This calculation gives a metric for determining the validity of each cut; for example, a rise-time cut with a calculated 70% acceptance efficiency should retain 70% of the counts in the x-ray lines and remove an unknown but larger percentage of the counts in the background components (exponential plus two flat components). To make this simpler to compare to the expected value, the expected efficiency (e.g. 0.7 in the case of a 70% rise-time acceptance cut) was divided from the relative percentage. Therefore, the presentation of the results of this section is as follows: for each component of the fit, the quantity (relative efficiency)/(expected efficiency) – denoted as ρ_{eff}– is calculated and plotted versus the cut applied to the data. Additionally, the extracted fit value for the exponential constant, c_1, of the background provides a test of the cut model and can act as a probe of the shape of the background distribution.

Low-gain channel results

The results for the low-gain channel are given in Figure 3.19 and in Table 3.4. The higher-energy x-rays (> 8 keV) behave well at the higher acceptance percentages ($\geq 90\%$), though
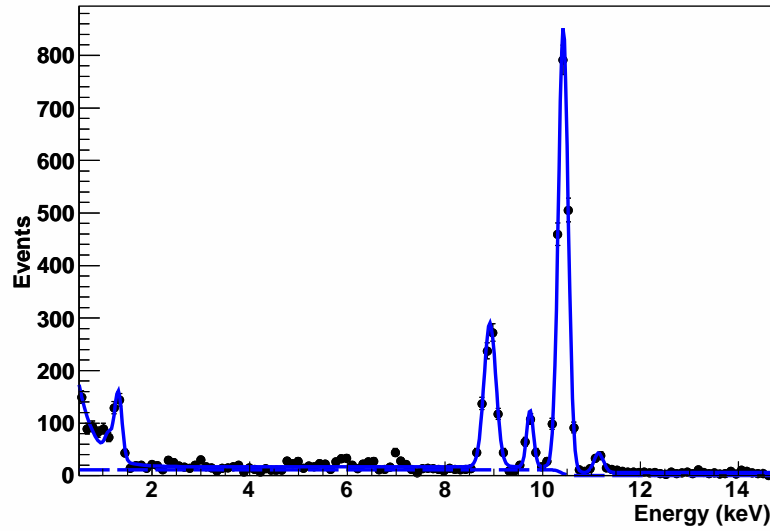
(a) Low-gain channel
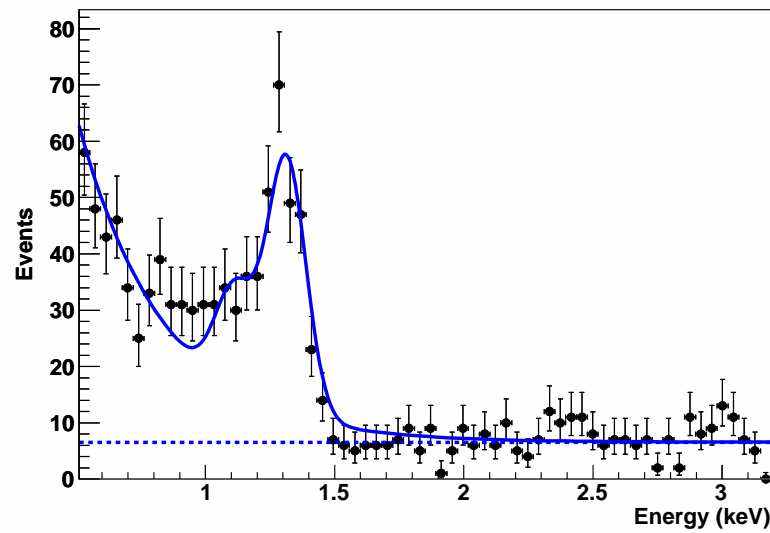
(b) High-gain channel

Figure 3.18: Example of fits performed to estimate amount cut by rise-time cuts. These data had the 99% rise-time cut applied. The l.ow-gain channel plot includes the error function component of the fit, seen as a dashed line below the x-ray lines.

the agreement becomes worse at lower acceptances as the lines tend to retain a larger percentage of counts than expected. This is evident in the figure as the points begin to increase for rise-time cuts of lower efficiency. The lower-energy L-capture line from ^{68}Ge behaves similarly to the higher-energy lines, but it is clear the amplitude of ^{65}Zn L-capture line fluctuates significantly. This is due to the lack of resolution at low energy for this channel, and so the high-gain channel is required to test this cut more precisely.

The amplitudes of the background components behave generally as expected, showing a reduction significantly greater than 99% for a 99%-acceptance rise-time cut. This suggests that slow pulses comprise a significant portion of the background in the exponential and two flat components. The slopes of these lines can give us an idea of the relative composition of slow pulse and fast pulse in each component. For example, if slow pulses are completely removed using the 99% acceptance cut, then one expects that the quantity ρ_{eff} should remain roughly the same for cuts of lower acceptance. This is true for the flat components ('Flat' and 'Erfc') for larger acceptance percentages ($\geq 90\%$) but the value starts increasing similarly to the results from the high-energy x-ray lines. This is expected, since the results from the x-ray lines indicate that the cuts retain a larger percentage of events than expected. However, the two flat components do not behave consistently at lower rise-time cut acceptances, though this is due to a strong correlation between the error function and flat backgrounds. For example, for low-acceptance very few events remain in the data set and the error function and flat function become indistinguishable. Despite this, the fact that the ρ_{eff} is statistically equivalent for the two background components at higher rise-time acceptances suggests that the cut behaves correctly there.

The exponential background component exhibits the opposite behavior from the flat background components, decreasing as the rise-time acceptance percentage reduces. This suggests that slow pulses still remain in this region after rise-time cuts are applied. However, as noted before, the resolution of this channel is insufficient to test this properly and so this must be verified with the high-gain channel. Finally, the fit exponential constant, c_1, demonstrates a marked shift when the rise-time cut is initially implemented, but is consistent between the different rise-time cuts (see Figure 3.19(d)). The major shift arises as the 99% rise-time cut removes many events near threshold and the fact that the value

of c_1 remains consistent as further rise-time cuts are applied suggests that the calculated acceptance contours are consistent near threshold.

It is possible that the same mechanism which generates the partial charge collection in the 'plateau' discussed in Section 3.3.1 is related to or the same as that which gives rise to slow-rise-time pulses. If this were true, we would expect the rise-time cut to significantly reduce or completely remove the contributions to the background from the error function. The rise-time cut leaves a remaining contribution in the plateau, but this could also be due to deviations between the fit PDF and the data. One can clearly see this in Figure 3.20 comparing two fits to data with a rise-time cut of 95% and another set with only microphonics cuts. The significant deviation from the best-fit line of data in the 'plateau' region of the spectrum indicates the presence of unaccounted-for backgrounds, in particular possibly other electron-capture cosmogenic species[2]. The presence of this background does pull the contribution from the error function up.

Other possible sources for the additional background in the flat component have been investigated, in particular:

- Contributions from the Lorentzian tails from the finite widths of the K-, L-, and M-vacancy states

- Continuum from internal bremsstrahlung electron capture (IBEC)

- Overlap/exchange corrections due to the inexact overlap between the initial and final atomic wavefunctions.

The expected shape of the energy spectrum from the electron-capture decay of a species internal to the detector is given by [61]:

$$\frac{dW}{dE} = M^2 (Q_{EC} - E)^2 \sum_H \psi_H^2 \frac{\Gamma_H}{(E - E_H)^2 + \Gamma_H^2/4} \tag{3.5}$$

where M is the transition matrix element, Q_{EC} is the total available decay energy, E is the summed energy deposited in the detector, ψ_H is the overlap of the wavefunction of

---

[2]Contributions from the continuum of higher-energy processes should already be taken into account by the flat (non-error function) background.
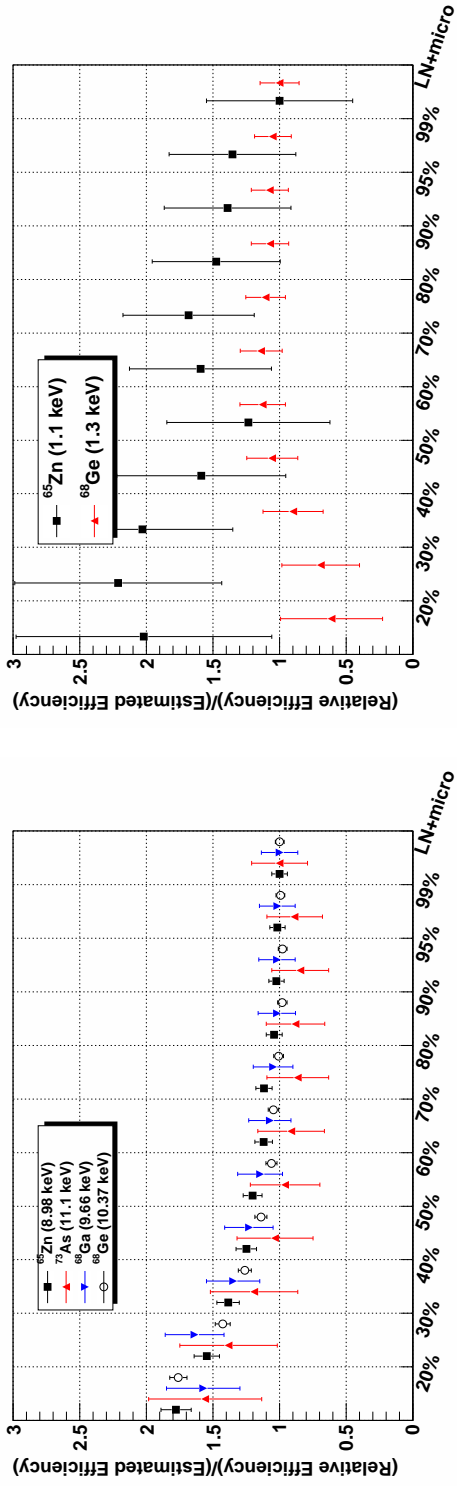
(a) X-rays
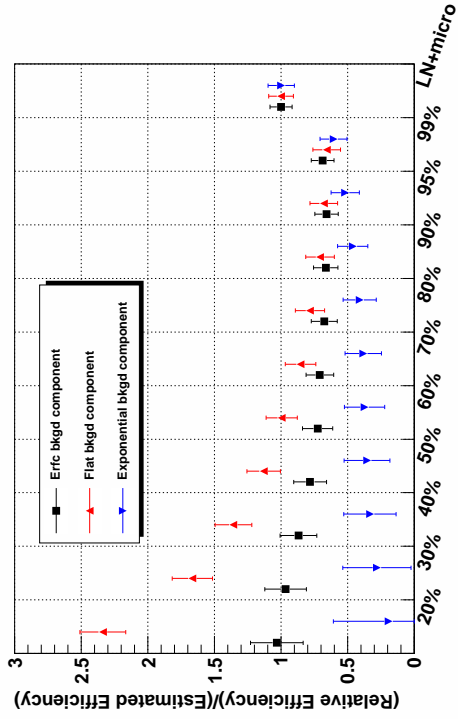
(b) Low x-rays

(c) Background components
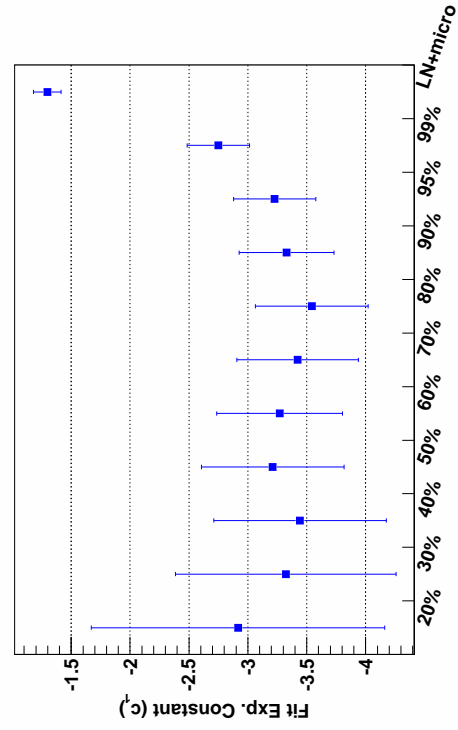
(d) Exponential Constant, c_1

Figure 3.19: Behavior of fit components after cuts for low-gain channel. 'LN+micro' refers to data with only LN and micro-phonics cuts applied. The percentages refer to data with a rise-time cut applied with the designated efficiency.
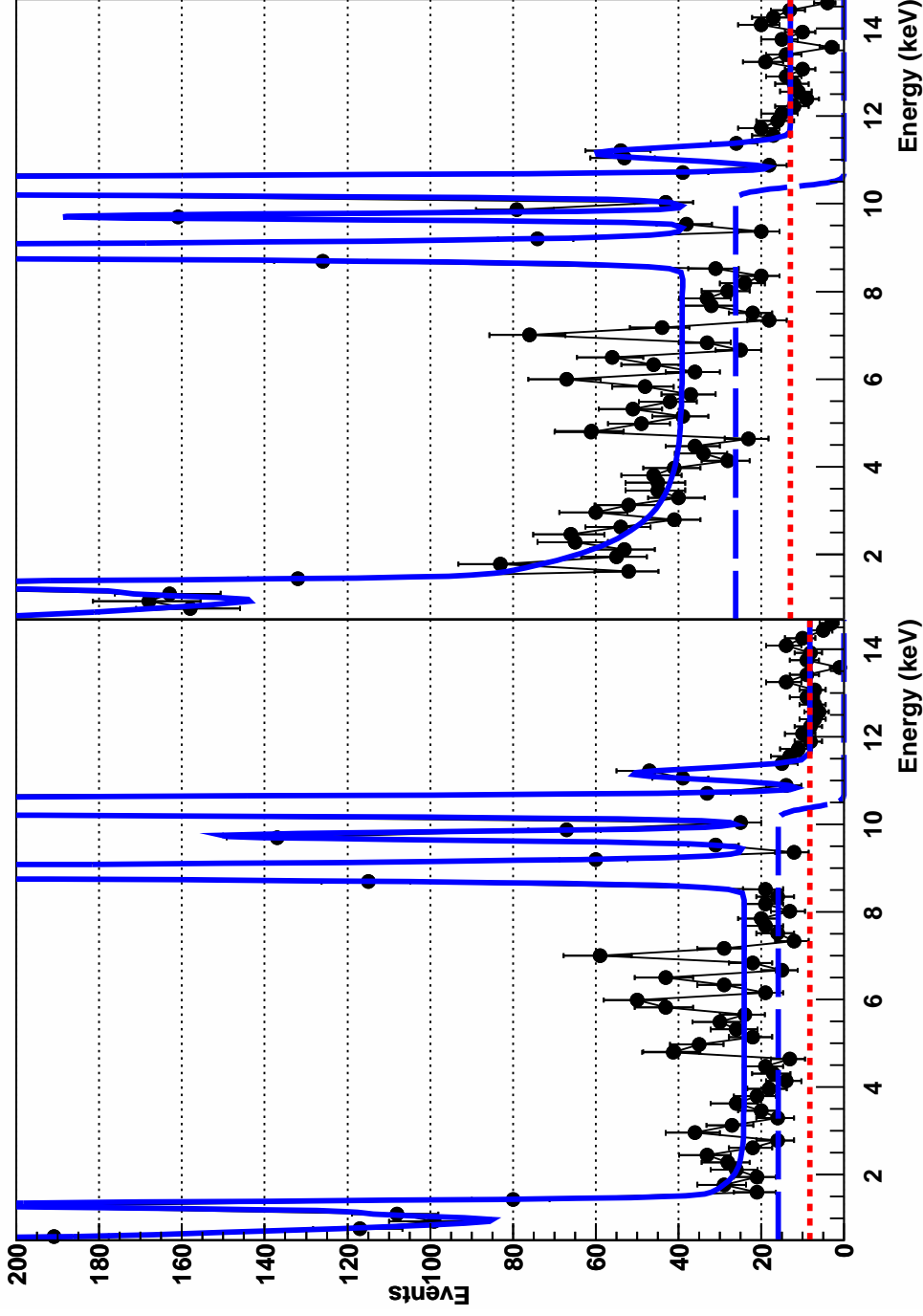
Figure 3.20: Comparison of fit to data with 99% rise-time cut applied (left) and data with only microphonics cuts applied (right). Components from the flat background (red, dotted) and the error function background (blue, dashed) to the overall fit (blue, solid) are shown. To note is the reduction in flat regions including the 'plateau' below the Ge K-capture line. However, several significant deviations of data from fit remain, suggesting the presence of additional background possible from other electron capture species. See text for details.

Table 3.4: Behavior of fit components after cuts for low-gain channel. Components are given in total counts.

Cut	Components (counts)						
	^{65}Zn (8.98 keV)	^{68}Ga (9.66 keV)	^{68}Ge (10.37 keV)	^{73}As (11.1 keV)	Erf bkgd	Exp bkgd	Flat bkgd
Rise-time 20%	264.9 ± 17.9	52.8 ± 8.9	681.3 ± 26.8	29.9 ± 7.9	302.9 ± 44.7	31.8 ± 12.1	520.1 ± 47.8
Rise-time 30%	345.9 ± 20.3	82.4 ± 10.9	828.9 ± 29.5	39.8 ± 9.8	425.5 ± 48.8	69.1 ± 15.3	555.8 ± 51.7
Rise-time 40%	413.2 ± 22.1	90.4 ± 11.5	976.0 ± 32.0	45.7 ± 10.5	510.3 ± 51.9	109.1 ± 17.7	604.1 ± 54.6
Rise-time 50%	465.6 ± 23.4	103.2 ± 12.2	1102.8 ± 34.0	49.6 ± 9.5	573.9 ± 52.1	145.3 ± 20.5	628.9 ± 52.7
Rise-time 60%	537.4 ± 25.0	115.2 ± 13.0	1232.4 ± 35.9	55.1 ± 9.8	638.9 ± 53.8	183.9 ± 22.3	665.0 ± 53.8
Rise-time 70%	583.8 ± 26.1	125.9 ± 13.5	1417.7 ± 38.4	61.3 ± 10.7	728.9 ± 55.6	219.5 ± 23.8	664.9 ± 55.3
Rise-time 80%	666.0 ± 27.7	140.7 ± 14.2	1559.8 ± 40.3	66.2 ± 10.8	793.1 ± 57.2	268.5 ± 25.6	697.3 ± 56.3
Rise-time 90%	698.2 ± 28.4	153.9 ± 14.8	1706.7 ± 42.1	76.0 ± 11.7	877.5 ± 59.2	340.1 ± 29.1	707.8 ± 57.7
Rise-time 95%	724.7 ± 28.9	162.5 ± 15.1	1796.9 ± 43.2	77.0 ± 11.6	919.0 ± 59.6	403.3 ± 31.2	718.2 ± 57.3
Rise-time 99%	749.8 ± 29.5	168.7 ± 15.5	1903.4 ± 44.4	84.2 ± 12.2	999.9 ± 61.9	491.2 ± 35.5	724.2 ± 58.6
LN + micro	745.1 ± 30.4	167.6 ± 16.3	1935.5 ± 45.2	95.9 ± 14.2	1467.7 ± 85.7	818.1 ± 57.2	1112.4 ± 72.9

Table 3.5: Decay widths for K-, L-, and M-vacancy states of ^{68}Ge. If the tabulation in the references did not explicitly include the Ga atomic state, the value of the width was estimated using adjacent atoms.

^{68}Ge decays	Decay widths (eV)	Reference
K-capture	2	[62]
L-capture	9	[63]
M-capture	5	[64]

orbital H at the nucleus, E_H is the binding energy of state H, and Γ_H the width of the state. Relevant values for ^{68}Ge are given in Tables 2.7 and 3.5. Relative intensity spectra are shown in Figure 3.21 for both the raw calculation and for it folded with the resolution function (Gaussian with σ defined as in Equation 3.3). It is clear that the Lorentzian tails contribute very little between the K- and L-capture lines.

IBEC is a second-order process by which the final state of the decay includes one or more radiative photons. These photons share energy with the emitted neutrino and so exhibit a continuous spectrum up to $Q_{EC} - E_H$, where E_H is the binding energy of the vacant orbital and Q_{EC} is the total energy available to the electron-capture decay. The relative intensity of this process has been found to be $\frac{\alpha}{12\pi}(\frac{Q_{EC}-E_H}{m_e c^2})^2$ IBEC-decays-per-K-capture-decays [65], or $\sim 10^{-5}$ for ^{68}Ge. The energy from these emitted photons would add to the energy released by the filling of the vacant orbital, yielding a continuum to the high-energy side of the K-, L-, and M-capture lines (see e.g. Figure 17 in reference [66]). If this process were to contribute significantly, a step up in the spectrum as one traverses the K-capture line from low to high energy should be observed; this is opposite to what is seen.

Corrections arising from the incomplete overlap of the initial and final atomic wavefunctions have been calculated in the literature, see e.g. [67]. These calculations largely only affect the decay spectrum by adjusting the relative intensities of the K-, L-, and M-decays by a few percent. Because of this, it is not expected that these corrections should lead to a significant contribution in the energy region between the L- and K-capture lines.
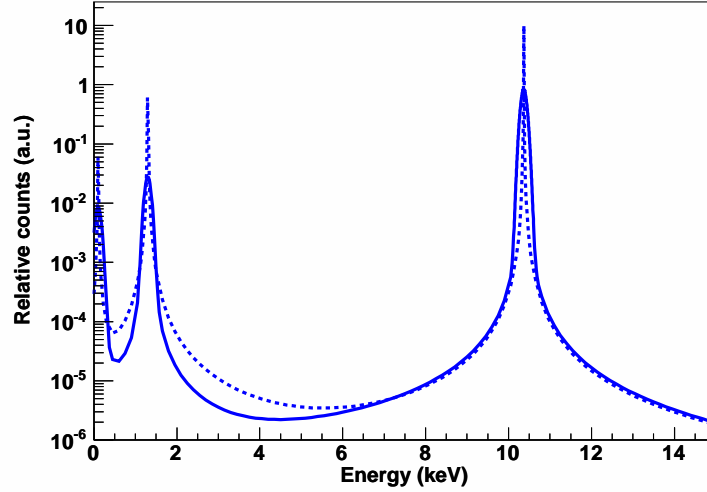
Figure 3.21: Calculated relative intensity spectrum for ^{68}Ge for the raw calculation (dashed) and for the raw calculation folded with the detector resolution (solid) using Equation 3.3.

High-gain channel results

Results for the high-gain channel are shown in Figure 3.22. The conclusions for this channel set are similar to the low-gain channel set. The enhanced resolution of the high-gain channel has reduced the size of the error bars for the L-capture lines and ρ_{eff} behaves very similarly to the low-gain channel for higher rise-time acceptances. Additionally, the behavior of the exponential constant is consistent in this channel as with the low-gain channel except with the 20% rise-time cut data. In this particular instance, the fit actually fails and runs up against the parameter limit for the exponential constant. This failure occurred due to insufficient counts remaining in the data set resulting in the two background components looking too much alike.

In this channel, only two background components, flat and exponential, remained since the energy range did not exceed 3.5 keV. ρ_{eff} with 20% rise-time cut data is distorted because the two background components look alike, rendering results from this data set incomparable to the other sets. The ρ_{eff} of the flat component behaves well at higher

rise-time acceptances, remaining flat for acceptances $\geq 80\%$. However, for the exponential component, ρ_{eff} reduces with decreasing rise-time acceptance, possibly indicating the cut does not *fully* remove background from slow-rise-time pulses. If one considers the measured rise-time distribution for the high-gain channel in Figure 3.17(a) with an included 99% cut line, it is clear that the cut begins to break down near threshold. This is not unexpected, since as the signal-to-noise of the pulse reduces de-noising will have less and less of an effect and the ability to measure the rise-time of pulses effectively will diminish.

Conclusions and Discussion

Evidence from past results [68, 69] and recent results presented in [50] suggest that slow pulses can arise due to poor charge collection from weak electric fields in the dead layer of the n^+ Li contact. These results also suggest that energy information of the initiating event is lost so that these slow pulses can comprise a background at low energy, though one of not-yet-precisely-determined distribution. The conclusions from these references were determined by scanning a source of low-energy gammas (^{241}Am, 59.5 keV gamma) around the outside n^+ contact of the detector (the detector geometries varied in the references) and determining the distribution of rise-times for recorded events. The low energy of the americium source ensured that a majority of the gammas deposited energy in or near the outer contact. The resulting rise-time measurements demonstrated a substantial distribution of slow rise-time pulses, supporting the conclusion that at least a component of these slow-rise-time events come from energy deposition at or near the outer contact.

Another measurement performed using the different, but similar P-PC detector [52] measured the rate versus time at low energies in several energy regions, watching the decay of ^{71}Ge (11.4 day half-life). This measurement focused on three main regions: the Ge K-capture line (10.367 keV), the Ge L-capture line (1.3 keV), and the flat region in between ($2 \rightarrow 6$ keV). The results found that the decay of counts in the flat region matched the decays of counts in the L- and K-capture regions, suggesting partial energy deposition from the ^{71}Ge decay below 10.367 keV. It is possible that this result is related to the slow-rise-time pulses, in particular that decays of ^{71}Ge near the outer contact resulted in both slow
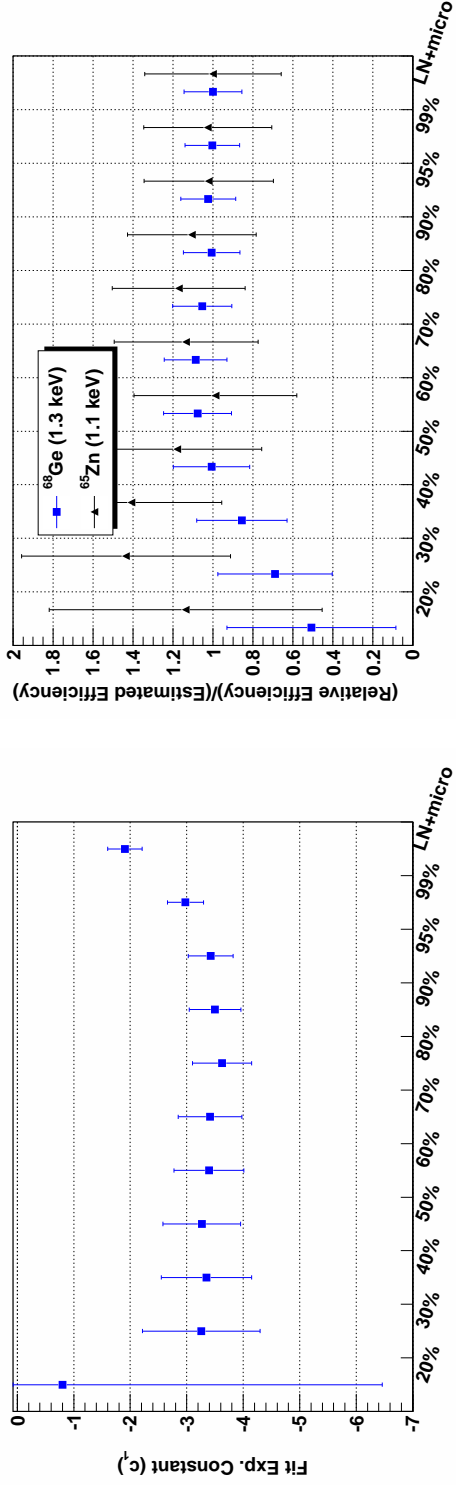
(a) Exponential Constant, c_1

(b) X-rays
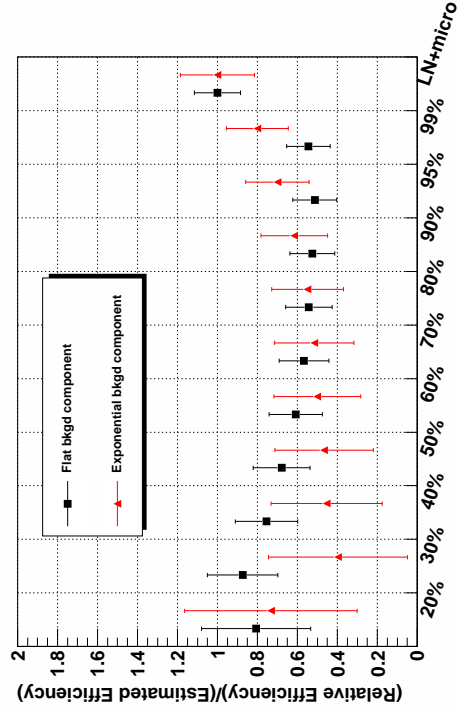
(c) Background components

Figure 3.22: Behavior of fit components after cuts for high-gain channel. 'LN+micro' refers to data with only LN and microphonics cuts applied. The percentages refer to data with an rise-time cut applied with the designated efficiency.

Table 3.6: Behavior of fit components after cuts for high-gain channel. Components are given in total counts.

Cut	^{65}Zn (1.1 keV)	^{68}Ge (1.3 keV)	Exp bkgd	Flat bkgd
Rise-time 20%	18.3 ± 7.7	21.0 ± 7.4	83.2 ± 26.4	125.5 ± 25.3
Rise-time 30%	34.7 ± 9.8	42.8 ± 9.3	67.7 ± 17.1	203.8 ± 19.1
Rise-time 40%	45.4 ± 11.0	70.7 ± 11.1	103.3 ± 19.1	234.5 ± 20.5
Rise-time 50%	47.5 ± 11.9	104.1 ± 12.9	132.5 ± 21.5	263.9 ± 21.9
Rise-time 60%	47.7 ± 12.5	133.6 ± 14.3	170.8 ± 23.2	283.8 ± 22.7
Rise-time 70%	64.0 ± 13.8	157.2 ± 15.4	205.1 ± 25.0	308.7 ± 23.8
Rise-time 80%	75.5 ± 14.7	174.3 ± 16.2	249.7 ± 26.2	337.5 ± 24.5
Rise-time 90%	80.2 ± 15.8	187.1 ± 17.1	315.0 ± 29.5	367.7 ± 26.1
Rise-time 95%	78.2 ± 16.2	201.1 ± 17.7	377.7 ± 31.7	379.0 ± 26.7
Rise-time 99%	81.9 ± 17.2	205.4 ± 18.4	450.5 ± 36.5	419.3 ± 29.6
LN + micro	80.6 ± 19.4	206.7 ± 21.2	568.4 ± 74.5	777.7 ± 63.7

pulses and partial charge collection, but no such data were taken during this measurement. To firmly establish the equivalency of these results, it is essential to measure the rise-time of the preamplifier traces, remove slow pulses, and produce the resulting counts vs. time for the region between the L- and K-capture lines. If partial energy deposition is arising in the same process producing the slow pulses, the removal of slow-pulse events should result in the flattening of the counts vs. time curve for the energy region between the capture lines. Unfortunately, for this analysis preamplifier traces were only recorded following the decay of the majority of ^{71}Ge and the slow decay of ^{68}Ge (271 days) yields insufficient statistics over a short period of time to measure this well. The relative content of ^{71}Ge in the detector could be increased through activation by either bringing the detector to the surface or introducing a neutron source, but such a test cannot be performed until the conclusion of the counting runs.

Whereas it is likely that a component of these slow pulses is arising from charge deposition near the crystal outer contact, it is important to determine other possible sources of these events. For example, surface channel effects, i.e. distortion of crystal fields through charge collection on the passivated surface, might produce similar results. One could use the extreme sensitivity of the surface channel to temperature detailed in [70] to determine if the population of slow-rise-time events changes with temperature. Additionally, characterization measurements using a low-energy gamma source must be performed for several other P-PC detectors to determine how sensitive the distribution of slow-rise-time events is to manufacturing and geometry differences.

Whatever the origin of these slow-pulse events, it is clear that they can produce a background to low-energy signals. If the majority of these slow pulses arise from energy deposition near the outer surface, a cut on rise-time would effectively generate a cut on the fiducial volume and mass of the detector. Understanding the active fiducial volume of the detector is critical for experiments sensitive to mass exposure times, such as those searching for $0\nu\beta\beta$ or dark matter. For double-beta decay, even though $Q_{0\nu\beta\beta}$ is around 2 MeV, the decay is internal to the detector and energy is deposited by two electrons summing to $Q_{0\nu\beta\beta}$. The range of electrons at this energy is on the order of 1 mm, so it is expected that poor charge collection in an outer layer of the detector could still have an effect on a fiducial

mass for $0\nu\beta\beta$ results.

3.4.4 Behavior of Energy Spectrum with Cuts Applied

It is useful to look at the behavior of the energy spectrum as cuts are subsequently applied to check for deviations or irregularities (e.g. strong peaks or valleys) that may be indicative of a problem with the analysis. The results of this are shown in Figure 3.23 where in all cases the base or reference data set was with LN fill events and electronic noise events (see Section 3.4.2) removed. Figure 3.23(a) shows the energy spectrum as subsequent cuts are applied. A complementary plot is provided in Figure 3.23(b) which shows the energy spectrum of events *removed* by the microphonics cuts and 95% rise-time cuts. It is interesting to note that the number of removed events diminishes as the energy moves to threshold (0.5 keV). It is expected that this arises from the reduced resolution of the rise-time calculation as the signal-to-noise ratio diminishes.

3.4.5 Extracting or Looking for Signals in Cut Data

The ultimate goal of introducing a cut to a data set is to remove events to enhance the signal-to-noise ratio of the data set. In this particular analysis, work focused on removing 'noise' events that might populate the energy region near threshold because this energy region is critical for deriving exclusion limits for dark matter. With many analyses, it is common to study cuts by choosing in parameter space a background region or set of regions and a background+signal region, preferably bookending the background+signal area with background-only regions to avoid any sort of boundary effects. This analysis explored rise-time cuts in a similar manner: calculate what you should see when you take a cut and see if the result matches your expectations. For higher-rise-time-cut acceptances, the calculation proved a sufficient model.

Understanding the data at low energies proves difficult for several reasons: (1) we do not know the exact energy distribution of slow pulses; (2) the most critical region, energies near threshold, is inherently a boundary region; and (3) a complete interpretation of slow pulses and their origin in this particular detector is missing. Addressing these points is important
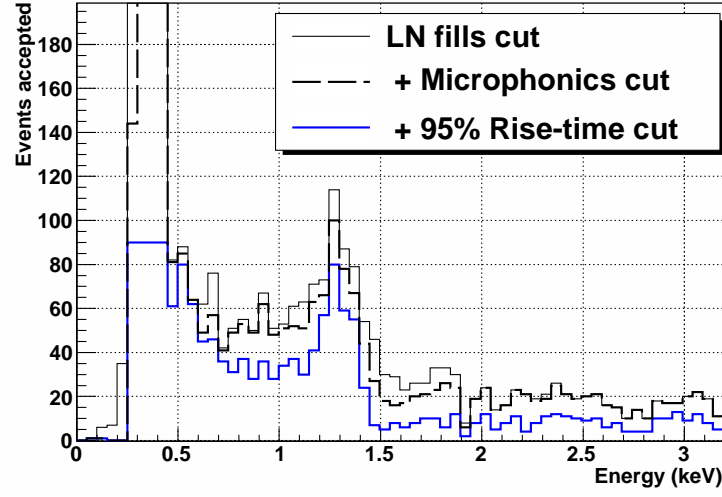
(a) Energy spectrum as subsequent cuts are applied
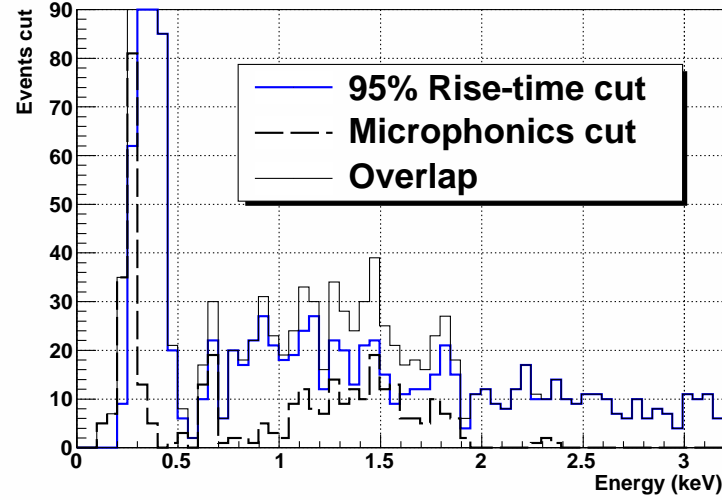
(b) Energy spectrum of events *removed* by microphonics and 95% rise-time cuts. The overlap of these two is also included.

Figure 3.23: Behavior of analysis cuts on the high-gain energy spectrum. In all cases, the base data set was with LN fill events and electronic noise events (see Section 3.4.2) applied. Events below 0.5 keV were not used to extract signal, but this energy region is still shown.

to using this data for physical interpretation:

1. Whereas the exact energy distribution of slow pulses for this detector is unknown, we can infer something about the population through the systematic results. In particular, the flat background components of the data essentially parameterizing the range $1.5 \rightarrow 14.5$ keV, behaved as one would expect if *all* slow pulses were cut using the 99% rise-time cut. In contrast the exponential component, parameterizing the population from threshold to ∼1.5 keV, behaved as if the 99% rise-time cut did not remove all slow-pulse events.

2. Near threshold, the two x-ray lines at 1.1 and 1.3 keV provide a test of the validity of the model, and these lines behaved well with the cuts. Additionally, the insensitivity of the exponential constant to different rise-time cuts implies that the calculated cut lines at different acceptance percentages were consistent until the threshold. However, it does *not* necessarily imply that slow-rise pulses were fully removed. The same result can be expected if slow and fast pulses occupy a similar rise-time distribution in this energy region due to limitations of the rise-time calculation at low signal-to-noise ratios.

3. Evidence suggests that at least a portion of these slow pulses arise from energy deposition near the outer contact and therefore a rise-time cut would require a reduction in fiducial mass. Even if not all slow pulses are generated by this mechanism, the assumption that they are is considered conservative .

Points 1 and 2 relate to one another and both suggest that the region near threshold could still be contaminated with slow-rise-time pulses. The presence of additional uncut events here could explain the general exponential shape seen in the data especially considering that the calculated acceptance line necessarily increases as it nears threshold (see Figure 3.17). Other explanations of this feature could be considered as well: (1) a population of events arising from a modulation of noise; (2) uncut microphonics; and (3) presence of an exponential signal from, for example, a WIMP source or some other unknown background. The first

two possibilities should exhibit characteristic time signatures in events near threshold; in particular, they should arrive with a rate that varies with time over short time scales. The presence of time variation in the data is explored in the following section (Section 3.5). It is also possible to measure the noise profile over time by generating random triggers, but this was not performed for this detector. The third explanation coupled with the understanding that slow-rise-time pulses could still populate near threshold suggests that any analysis of this near-threshold region must treat it with care, conservatively foregoing any claims other than exclusion limits. The extraction of a limit on a signal in the presence of unknown background is explored in the following chapter, in particular Section 4.3.2.

Following point 3, for this analysis and for the purpose of estimating the mass of the detector we assume that either all or a significant majority of slow-rise-time events are coming from energy deposition near the surface. We therefore assume that the detector includes an outer dead layer within the outer Li contact, followed by a transition layer with weak electric fields. We can calculate the width of this transition region through two methods and compare, in particular using the results of the rise-time systematic tests described earlier as well as another calculation presented in [50]. If we assume a majority of the slow-rise events in the 'plateau' beneath the Ge K-capture line come from partial charge collection of the energy deposited from these electron-capture decays, then the percentage of volume of the detector in which these slow-pulse events occur may be estimated as follows: calculate the sum of events in the prominent x-ray lines at high energy, $y = \sum_i N_i$ where i includes ^{68}Ge, ^{68}Ga, ^{65}Zn, and ^{73}As, and then calculate the difference in counts, Δx, between events in the error function distribution before and after a 99% rise-time cut, properly adjusting for the expected efficiency of the cut (99%). The value, $\frac{\Delta x}{\Delta x + y}$ is then the fraction of volume affected by slow-pulse generation with respect to the total active volume of the detector. The values Δx and y were measured as 457.7 ± 105.9 and 3401.8 ± 121.3, respectively, yielding an affected fraction of active volume of 0.135 ± 0.031. The active volume can be estimated by assuming an outer dead layer of 1 mm, reducing the 0.44 kg total mass to an active mass of 0.4 kg. Therefore, the final fiducial mass after rise-time cuts would be assumed to be 0.346 ± 0.0124 kg. This estimation ignores any effect from the passivated surface of the detector.

We may compare this to a separate estimation presented in [50]. This particular measurement coupled a source scan with ^{241}Am with a simulation to estimate the dead layer plus transition region in a different but similar P-PC detector. The results of this measurement suggested an outer dead layer of 1 mm, plus a transition layer of 1 mm yielding a total active volume of 0.33 kg. This measurement agrees with the above estimation and to remain conservative the smaller value will be used in the later presented analyses.

When fitting to cut data, it is important to take into account the efficiencies of each cut. The total energy-dependent efficiency can be defined by the equation:

$$f_{eff}(E) = f_{Trigger}(E) \prod_{i \ \in \ \{cuts\}} f_i(E) \tag{3.6}$$

where $f_{Trigger}(E)$ is the trigger efficiency defined in Equation 3.2, and the product is over all cuts taken. The efficiency of the microphonics cut was estimated as $f_{micro} = 0.997$ and the rise-time efficiencies are estimated as equivalent to the acceptance percentage. This assumption is valid for higher rise-time acceptances, $\geq 80\%$, and conservative for lower percentages. The efficiency function is used to determine the weight of an event at a particular energy, given by $w(E) = 1/f_{eff}(E)$.

3.5 Time Correlations and Systematics

Searching for time dependence of parameters and any possible time correlation of events provides a powerful systematic test of the data. Small deviations in detector health (e.g. noise, baseline shifting, etc.) can have a more marked effect on events near threshold. For example, a change in noise or a shift in baseline could affect the triggering efficiency, modifying the collection probability for events near threshold. Additionally, searching for time correlations provides a metric for ensuring that data collection is behaving appropriately and as expected. In addition to tracking detector health parameters, it is important to analyze event rates and times between events to look for any deviation from Poisson statistics. Such deviations could arise from environmental changes or events. For example, nearby work in the experimental hall generating additional vibration could increase microphonics events, or an introduction of a hot source in the shared lab space could increase count rates. Finally, it is possible to perform tests in combined energy-time space looking to see if classes of

events (e.g. events in a certain energy range) have correlations *outside* their class region. An obvious example of this is a decay chain, in which an initial energy deposition from a parent is correlated in time with another event at different energies from the decay of the daughter. Whereas these types of correlations between decay chains should show up in the data, the discovery of other correlations of unknown origin could point to a potential source of event contamination.

3.5.1 Detector Health versus Time

In contrast to the previous data run with P-PC2, fewer parameters were tracked during the lifetime of the experiment (see Section 2.3.3). In particular, due to limited channels in the DAQ electronics the trigger efficiency was not monitored with automated tests nor was the rate of the reset electronics recorded. In general, both these parameters are important to catalogue since they could affect the signal region near threshold. However, due to the strong correlation of the reset rate on the measured baseline in unshaped pulses seen in the results of Section 2.3.3, the baseline can be used as a proxy to ensure that the reset rate did not change significantly. Similarly, it was possible to monitor the noise by using the measured RMS of the baseline to determine whether it changed significantly over time.

The baselines of two channels were tracked, channel 1 (shaped) and channel 5 (unshaped preamp trace). For every day of run-time, the calculated baselines were binned and fit to a Gaussian. The binned baseline data included outliers in one direction which were attributed to having come from the return of the baseline to its nominal value following a reset event in the preamplifier. Since these outliers were a low percentage of the total events, only one Gaussian was fit to the binned baseline data. From these fits, the mean and σ values were extracted and the relationship of these parameters versus time are plotted in Figure 3.24. An inspection of the results reveals that the σ values from both unshaped and shaped channels were stable with time. In contrast, the mean value of the baseline in the unshaped channel demonstrates variation, but this is not mirrored exactly in the mean of the baseline calculated from the shaped channel. This is likely due to baseline restoration electronics in the spectroscopy amplifier which render the baseline of the output less sensitive to baseline

modulation of the input. Regardless, it is evident that the noise of the baseline does not change significantly in either the shaped or unshaped channels over the time range of the experiment. This suggests that the trigger efficiency and electronic noise remained similarly unchanged throughout the time of the experiment.

3.5.2   Rates in Energy Regions

Several tests can be performed to analyze the Poisson nature of the data: (1) event rate in an energy region; (2) time difference between events in an energy region; and (3) forward and backward time difference between events occurring in an energy region and *any* other events. All of these tests should be made with different cuts applied (see Section 3.4 for a description of the cuts) since an analysis cut may preferentially remove a portion or an entire population of correlated events. Therefore, for these tests several cuts were used including the minimal cut (LN fill cut), microphonics cuts and rise-time cuts of varying acceptances. To perform all these tests, events were selected in particular energy regions: $0.45 \to 0.55$ keV, $0.5 \to 1$ keV, $1 \to 2$ keV, $3 \to 8$ keV, and $10 \to 10.76$ keV. The first two regions tested around and just above threshold, the next two tested regions not significantly populated with ϵ decays, and the final examined a region symmetric around the mean of the Ge K-capture x-ray (10.367 keV).

During LN fills, all events are vetoed and this could, in principle, generate a significant distortion to any calculated Poisson rates. However, since the amount of time vetoed during an LN fill is small compared to the measurement time this has little effect on the results. This can be seen by considering that the probability that one or more events will happen in a certain time period, τ, is $1 - \exp(-\lambda_\tau)$, where λ_τ is the average count rate over that time period. If we consider that the time vetoed during an LN fill is ~ 15 minutes, and these fills happen once every two days, the expected number of times one or more events occurs during the roughly 160 day run period is $\sim 80(1 - \exp(-\lambda_\tau))$. Additionally, the percentage of events affected (events affected/events expected) is given by $\frac{1}{192\lambda}(1 - \exp(-\lambda))$. The maximum total rate for any of these regions is less than 1 count per 15 minutes, meaning that the percentage of events affected is less than 0.5%. Since all other cuts (microphonics
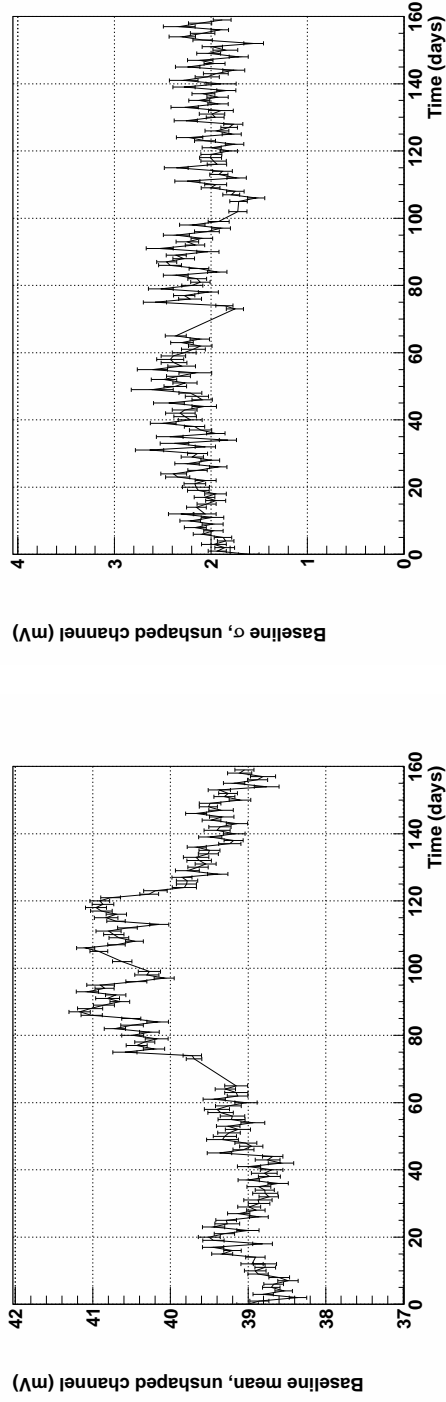
(a) Unshaped channel, mean

(b) Unshaped channel, σ

(c) Shaped channel, mean
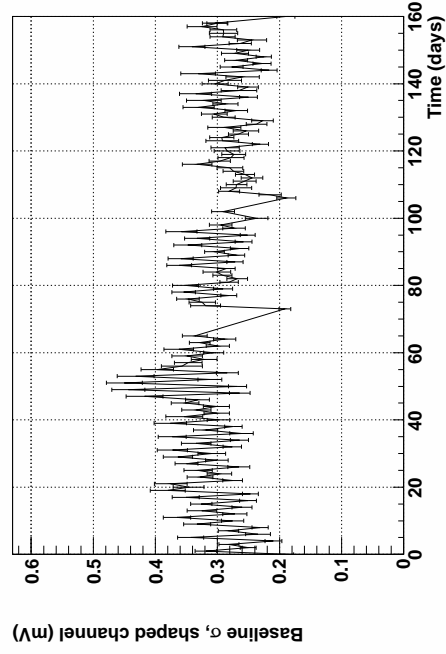
(d) Shaped channel, σ

Figure 3.24: Baseline mean and σ versus run time for both shaped and unshaped preamp traces.

and rise-time cuts) should have no time dependence (assuming that slow rise-time events and microphonics events occur randomly and independently), it is expected that rates calculated from data with these cuts applied should be slower, but still Poisson in nature.

The first test involves analyzing the event rates in a given energy window. The data were windowed in 8-hour bins and the numbers of counts in the different energy regions were calculated for each of these 8 hours. Any 8-hour time periods with less than 100% live-time (i.e. 8 hour runs during which an LN fill occurred) were removed from the data set. The number-of-counts-seen was histogrammed and fit with a Poisson distribution to extract the rate. Examples of these fits are shown in Figure 3.25 and a summary of the fits, including extracted rate and goodness-of-fit, is given in Table 3.7. These results indicate that the regions selected behave as expected, demonstrating count rates that follow Poisson distributions. Additionally, as expected the count rates were reduced with the introduction of cuts while retaining a Poisson shape.

The times between adjacent events were calculated for events in selected energy regions. The differences were then binned and the result fit to an exponential to extract the decay parameter, λ, which is equivalent to the parameter extracted in the previous test. Because of this, analyzing the time-difference also tests the Poisson nature of the data, but it can have different sensitivities since it is not confined to a particular time range. In particular, the previous test is not sensitive to correlations beyond 8 hours, whereas a time-difference test is not limited to a specific time range. Results of the time-difference test are presented as before, with examples of fits to LN-cut data given in Figure 3.26 and extracted parameters given in Table 3.8. It is clear that the P-value of the fit for $10 \rightarrow 10.76$ keV region is low, or rather close to one of the bounds (0 or 1) and, in contrast to the other regions, does not significantly change as one would expect with the application of different cuts. No corrections for LN fills have been introduced for these tests because doing so would require assuming an underlying Poisson distribution. However, since the count rate in the ^{68}Ge region is higher than in the other regions, it is possible that this has distorted the distribution enough to increase the χ^2. It is not known if this will explain the apparent 'bump' at 9-10 hours visible in Figure 3.26(d). Despite this, all the rate values calculated using this method are consistent with the rates calculated using the previous method, lending

Table 3.7: Count rates in 8-hour time periods for selected energy ranges. Results from data with LN, LN+microphonics, and LN+microphonics+99% rise-time cut are included. λ is normalized to counts per hour. The P-Value is the probability of observing this χ^2 value or greater given the numbers-of-degrees-of-freedom (NDF).

Range (keV)	λ (cts / hr)	χ^2	NDF	P-Value
LN				
0.45 → 0.55	0.046 ± 0.004	6.08296	2	0.0477643
0.5 → 1	0.158 ± 0.006	2.73425	5	0.740876
1 → 2	0.262 ± 0.010	4.34057	7	0.73982
3 → 8	0.373 ± 0.010	4.94593	9	0.838996
10 → 10.76	0.609 ± 0.016	15.5406	12	0.213196
LN + Microphonics				
0.45 → 0.55	0.041 ± 0.004	1.78836	2	0.408943
0.5 → 1	0.143 ± 0.006	3.82657	5	0.574646
1 → 2	0.204 ± 0.008	1.40094	6	0.965801
3 → 8	0.341 ± 0.010	7.37705	8	0.496551
10 → 10.76	0.563 ± 0.016	14.9226	11	0.186068
LN + Microphonics + 99% Rise-time cuts				
0.45 → 0.55	0.038 ± 0.004	2.57558	2	0.27588
0.5 → 1	0.124 ± 0.006	6.25875	5	0.281849
1 → 2	0.139 ± 0.006	3.6481	5	0.601106
3 → 8	0.221 ± 0.008	2.73918	6	0.840798
10 → 10.76	0.544 ± 0.015	15.0561	11	0.179943

100

(a) 0.45 → 0.55 keV

(b) 0.5 → 1 keV

(c) 3 → 8 keV

(d) 10 → 10.76 keV
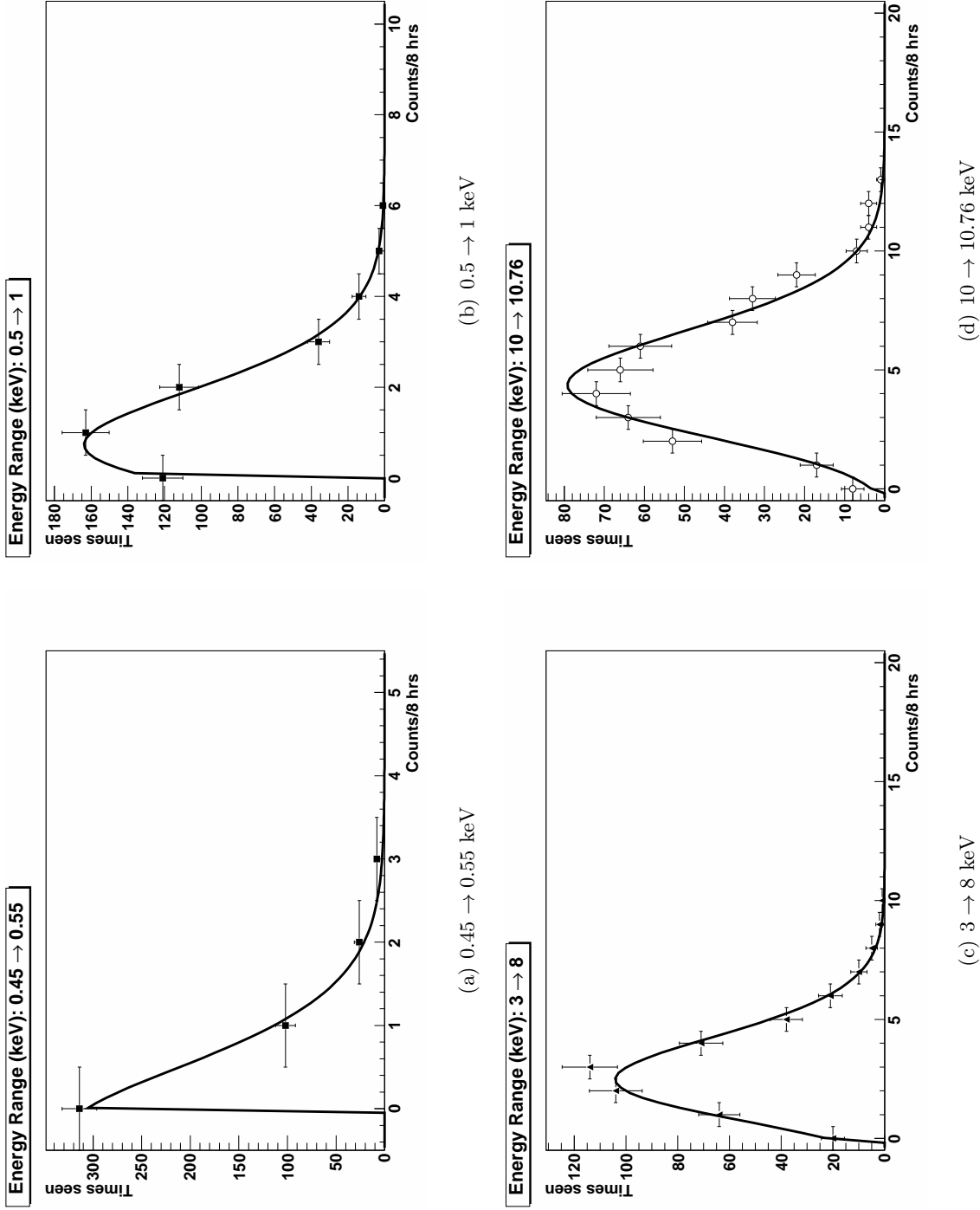
Figure 3.25: An example of fits to count rates in 8-hour time periods for selected energy ranges. LN cuts were applied to the data. Range 1 → 2 keV has been omitted.

confidence to the conclusion that the count rates are behaving as Poisson variables.

The final test described in this section involves selecting events in a given energy region and calculating the time differences to the events at *any* energy (below 14 keV) both immediately previous and immediately following. These time differences are then binned and fit to an exponential to extract the rate. The purpose of this test is to search for correlations between events occurring in the selected region and any other events, looking for any possible 'memory' either forwards or backwards in time. In other words, the rates calculated for both forward and backward events and for different energy regions should all be consistent and the calculated rate should be the average data rate given a set of cuts.

Also, by applying it to data with different cuts, it is possible to test the gross efficiency of these cuts since it is essentially a measurement of the count rate. For example, the rate calculated for data with 70% rise-time cut applied should be 0.7/0.99 times the rate calculated for a 99% rise-time cut, assuming that the majority of background from slow pulses has been removed. Results from this test are shown in Figure 3.27. The line in the figure is an expectation of the calculated rate value, determined by taking the average, ρ, of the forward and backward rates for the $2 \rightarrow 8$ keV region for the 99% rise-time-cut data and extrapolating to the lower efficiencies. That is, the values at other rise-time efficiencies, ϵ, are calculated by $\frac{\epsilon}{0.99}\rho$. The results indicate that the forward and backward rates are consistent amongst the three regions tested and that the reduction in rate follows the correct trend. The scatter of the rate measured for the $0.5 \rightarrow 1$ keV is higher due to the fewer triggers seen in this region.

3.5.3   Time-energy Correlations

Two-dimensional energy-time correlations were calculated by analyzing a time window around events occurring in a selected region. As before, a set of events were selected within an energy range to be used as 'trigger' events. Given the time of a trigger event as a reference point, a two-dimensional energy vs. time histogram was generated for all events occurring within a time window forwards and backwards around the reference. The 2-d histograms for each trigger were then added together, producing a final intensity plot. Any

Table 3.8: Results of fits to time differences between events in selected energy ranges. Results from data with LN, LN+microphonics, and LN+microphonics+99% rise-time cut are included. λ is measured in counts per *hour*, in contrast to results in Table 3.7. The P-Value is the probability of observing this χ^2 value or greater given the NDF.

Range (keV)	λ (cts)	χ^2	NDF	P-Value
LN				
0.45 → 0.55	0.061 ± 0.008	23.9883	27	0.630963
0.5 → 1	0.172 ± 0.007	22.3103	16	0.133452
1 → 2	0.266 ± 0.009	5.42462	13	0.964619
3 → 8	0.380 ± 0.010	13.4823	15	0.5651
10 → 10.76	0.638 ± 0.016	24.603	12	0.0168202
LN + Microphonics				
0.45 → 0.55	0.049 ± 0.008	20.021	27	0.829872
0.5 → 1	0.152 ± 0.007	15.6124	18	0.61958
1 → 2	0.207 ± 0.008	7.90211	13	0.849919
3 → 8	0.353 ± 0.010	12.2548	19	0.874449
10 → 10.76	0.593 ± 0.015	24.9871	13	0.0231737
LN + Microphonics + 99% Rise-time cuts				
0.45 → 0.55	0.045 ± 0.009	23.0667	26	0.62917
0.5 → 1	0.133 ± 0.007	14.4484	19	0.756952
1 → 2	0.139 ± 0.008	16.5919	16	0.412466
3 → 8	0.231 ± 0.010	12.3232	20	0.904504
10 → 10.76	0.577 ± 0.015	27.4457	14	0.0168382
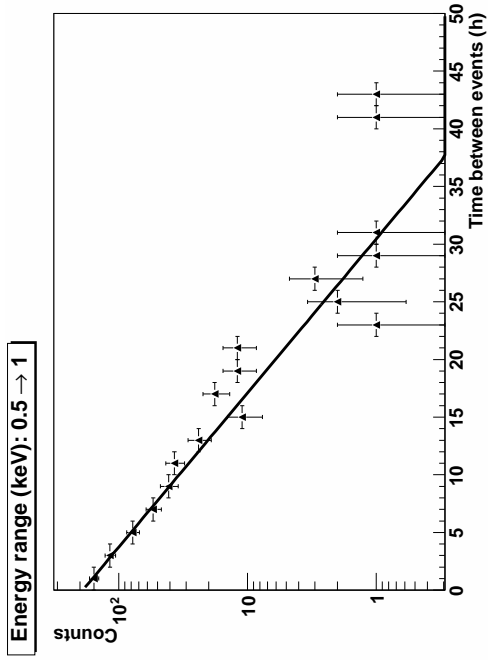
(a) $0.45 \rightarrow 0.55$ keV

(b) $0.5 \rightarrow 1$ keV
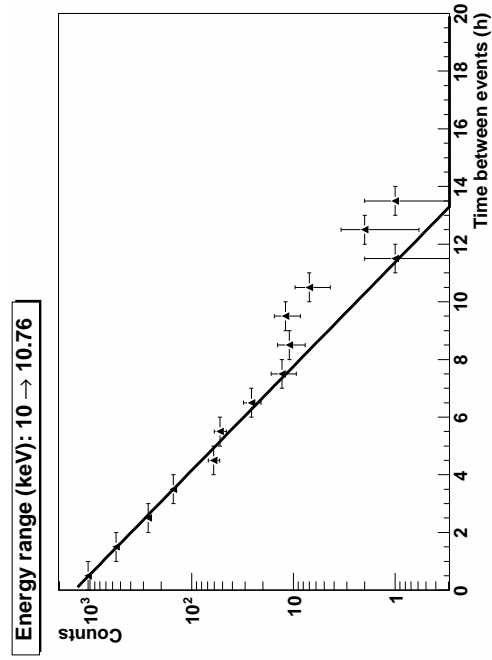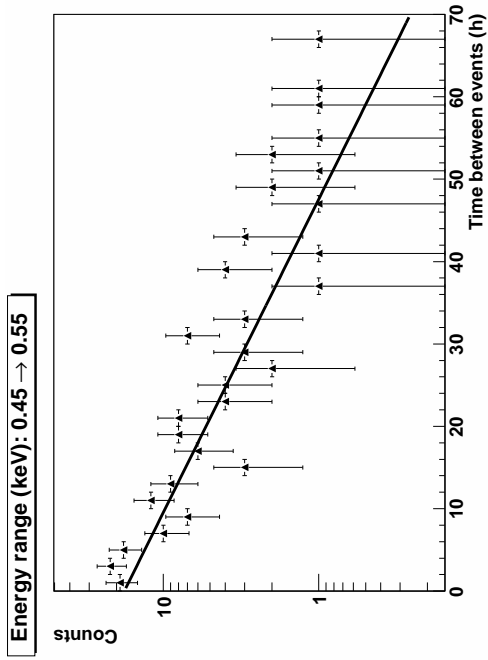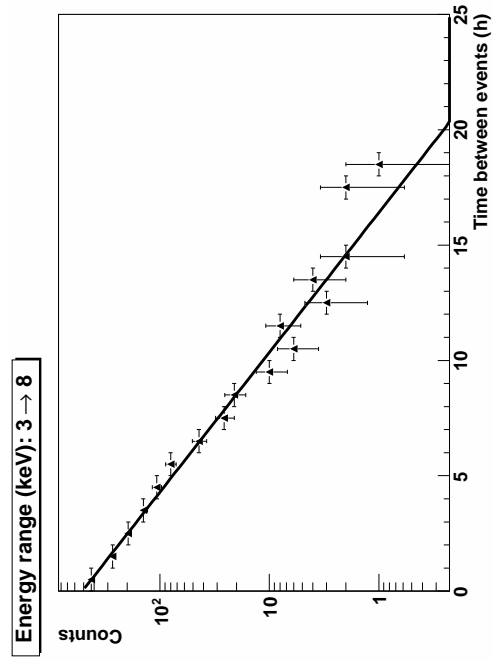
(c) $3 \rightarrow 8$ keV

(d) $10 \rightarrow 10.76$ keV

Figure 3.26: An example of fits to time differences between events in selected energy ranges. LN cuts were applied to the data. Range $1 \rightarrow 2$ keV has been omitted.
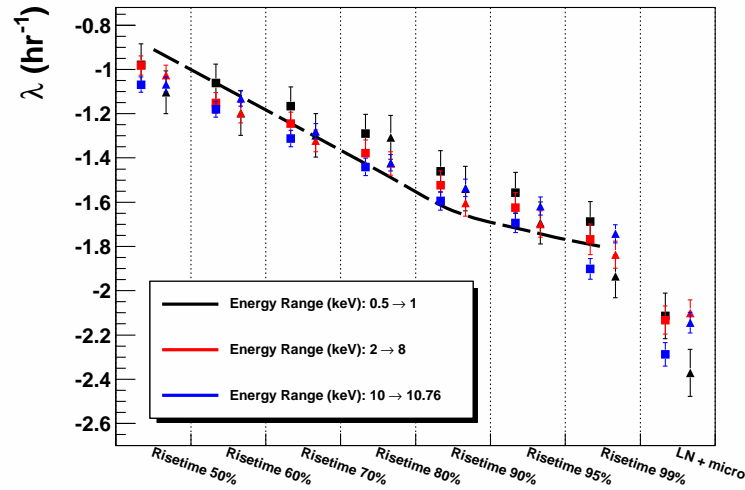
Figure 3.27: Rates calculated from the time difference between events in selected regions and events outside that region. λ is the raw parameter in the exponential fit $(\exp(\lambda t))$. The rates were calculated both forward in time (squares) and backwards in time (triangles) for different cut types. The line drawn is an estimate of the expected efficiency of the rise-time cuts using the $2 \rightarrow 8$ keV region, see text for details.

Table 3.9: Results of fits to time differences between events in selected energy ranges with different cuts. The P-Value is the probability of observing this χ^2 value or greater given the NDF. 'B' and 'F' correspond to backwards and forwards time differences, respectively.

Cut	Range (keV)	Triggers	Direction	Rate (cts/hr)	χ^2	NDF	P-value
	$0.5 \rightarrow 1$	166	B	-1.10 ± 0.10	4.77	7	0.689
			F	-0.98 ± 0.09	2.81	7	0.902
Rise-time 50%	$2 \rightarrow 8$	651	B	-1.03 ± 0.05	19.10	18	0.386
			F	-0.98 ± 0.04	13.48	20	0.856
	$10 \rightarrow 10.76$	1158	B	-1.07 ± 0.03	39.84	45	0.690
			F	-1.07 ± 0.03	26.00	39	0.945
	$0.5 \rightarrow 1$	205	B	-1.20 ± 0.10	7.52	6	0.275
			F	-1.06 ± 0.09	2.80	7	0.902
Rise-time 60%	$2 \rightarrow 8$	711	B	-1.20 ± 0.05	23.24	18	0.182
			F	-1.15 ± 0.05	30.70	19	0.044
	$10 \rightarrow 10.76$	1292	B	-1.13 ± 0.03	28.50	39	0.892
			F	-1.18 ± 0.04	19.85	36	0.987
	$0.5 \rightarrow 1$	242	B	-1.30 ± 0.10	7.04	6	0.317
			F	-1.17 ± 0.09	3.23	7	0.863
Rise-time 70%	$2 \rightarrow 8$	769	B	-1.32 ± 0.05	14.98	17	0.597
			F	-1.24 ± 0.05	24.93	18	0.127
	$10 \rightarrow 10.76$	1482	B	-1.28 ± 0.03	31.32	36	0.691
			F	-1.31 ± 0.04	28.13	36	0.823
	$0.5 \rightarrow 1$	293	B	-1.31 ± 0.10	3.56	5	0.615
			F	-1.29 ± 0.09	4.38	8	0.821
Rise-time 80%	$2 \rightarrow 8$	824	B	-1.43 ± 0.05	11.13	15	0.744
			F	-1.38 ± 0.06	22.71	16	0.122
	$10 \rightarrow 10.76$	1628	B	-1.42 ± 0.04	37.31	34	0.319
			F	-1.44 ± 0.04	26.58	34	0.814
	$0.5 \rightarrow 1$	350	B	-1.54 ± 0.10	1.89	4	0.756
			F	-1.46 ± 0.09	4.15	7	0.762
Rise-time 90%	$2 \rightarrow 8$	884	B	-1.60 ± 0.06	9.95	15	0.823
			F	-1.52 ± 0.06	27.90	15	0.022
	$10 \rightarrow 10.76$	1779	B	-1.54 ± 0.04	33.90	32	0.376
			F	-1.59 ± 0.04	24.89	30	0.730
	$0.5 \rightarrow 1$	402	B	-1.69 ± 0.09	2.74	4	0.602
			F	-1.56 ± 0.09	3.80	6	0.704
Rise-time 95%	$2 \rightarrow 8$	914	B	-1.70 ± 0.06	10.25	14	0.744
			F	-1.62 ± 0.07	30.54	15	0.010
	$10 \rightarrow 10.76$	1872	B	-1.62 ± 0.04	28.90	32	0.624
			F	-1.69 ± 0.04	23.45	27	0.661
	$0.5 \rightarrow 1$	457	B	-1.94 ± 0.10	5.89	4	0.207
			F	-1.69 ± 0.09	1.94	5	0.858
Rise-time 99%	$2 \rightarrow 8$	976	B	-1.84 ± 0.06	11.25	13	0.590
			F	-1.77 ± 0.07	24.01	14	0.046
	$10 \rightarrow 10.76$	1982	B	-1.74 ± 0.04	24.28	28	0.666
			F	-1.90 ± 0.05	32.06	27	0.230
	$0.5 \rightarrow 1$	526	B	-2.37 ± 0.11	8.35	4	0.080
			F	-2.11 ± 0.10	1.92	4	0.751
LN + micro	$2 \rightarrow 8$	1564	B	-2.10 ± 0.06	11.87	11	0.374
			F	-2.13 ± 0.06	23.39	10	0.009
	$10 \rightarrow 10.76$	2054	B	-2.14 ± 0.05	26.70	24	0.319
			F	-2.29 ± 0.05	29.52	22	0.131
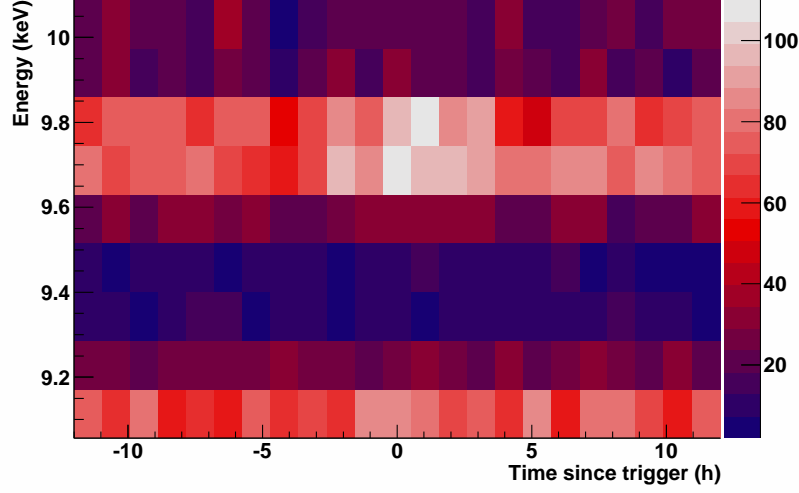
Figure 3.28: Time-energy correlation triggering on ^{68}Ge decays (2211 triggers), z axis is in counts. The region intensity centered at 9.7 keV between 0 and 2 hours post trigger arises from the K-capture decay of the ^{68}Ga daughter.

strong correlation between events in energy-time space should show up as having significant intensity in these figures. The data used for this analysis had only LN fill cuts applied to avoid removing correlating events from the data set.

To give an example of how this calculation works, the time-energy correlation between the ^{68}Ge and ^{68}Ga decays was analyzed. The details of this decay chain are discussed in Section 2.4, but a summary of the relevant points is given: ^{68}Ge internal to the crystal decays to ^{68}Ga, depositing 10.36 keV of energy $\sim$ 90% of the time via x-rays and Auger electrons. ^{68}Ga has a 68 minute half-life and decays via electron-capture $\sim$ 10% of the time; 90% of these decays are K-line captures and deposit 9.7 keV into the crystal. We can therefore look for a time correlation by choosing the trigger region 10 $\rightarrow$ 10.76 keV and looking forwards and backward in time around the 9.7 keV decay region of ^{68}Ga [45]. Results for 2211 triggers are shown in Figure 3.28 where the intensity of the daughter decay is clear between 0 and 2 hours post trigger. The rough expected count rate for the Ga K-capture decay is about $0.09 \times 2211 \sim 200$ which agrees with the results of the plot.

This analysis was applied to other trigger regions, including a region against threshold, $0.5 \to 1$ keV, and away from cosmogenic peaks, $2 \to 6$ keV. The results are presented in Figure 3.29 for an energy window from 0.5 to 8 keV, avoiding higher energies where the count rates from the ^{65}Zn and ^{68}Ge dominate the intensities. The only apparent features in these plots are intensities at around 1.3 keV, which likely correspond to random coincidences with the ^{68}Ge L-capture decay. It is possible to integrate these 2-dimensional histograms along the time axis, producing an average energy spectrum around a trigger in the designated region. By dividing by the number of triggers and the time of integration, one can compare the spectra from different trigger regions to identify any particular discrepancies. This was done for all three regions presented in this section, integrating over ± 12 hours for the lower energy regions and ± 3 hours for the ^{68}Ge K-capture region. The results of this integration are shown in Figure 3.30. All of the trigger regions are consistent, except that the $10 \to 10.76$ keV trigger region demonstrates an enhancement in the ^{68}Ga K-capture line. This is to be expected since the average rate of the Ga decay should be higher centered on a trigger from the parent decay.

3.5.4   Conclusions

The search for time correlations in the data yielded results consistent with Poisson statistics, suggesting that the regions of the data sets tested are clean from events arriving in bursts or with defined frequency. Additionally, the electronic noise of the detector remained constant throughout the counting runs. The one significant deviation found resulted in the discovery of a class of events detailed in Section 3.4.2, but the lack of other environmental information made it impossible to determine their origin. For long-running experiments such as the Majorana Demonstrator, the monitoring and control of the environment is critical for tracking down and explaining events that may appear in the data. As the abilities of future experiments expand to include the unexplored regions enabled by ultra-low-energy thresholds, so too must the effort to ensure that these areas are fully understood.
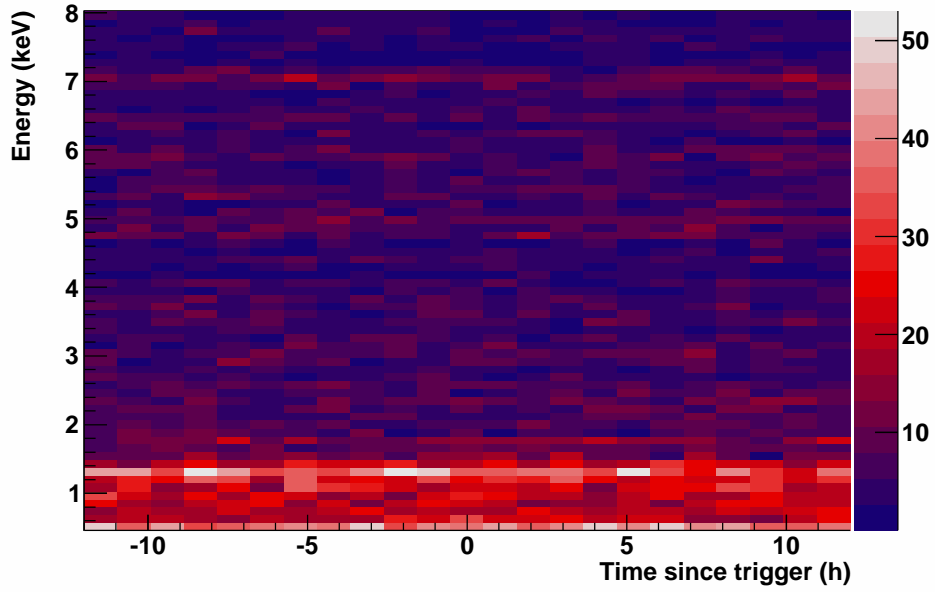
(a) $0.5 \rightarrow 1$ keV
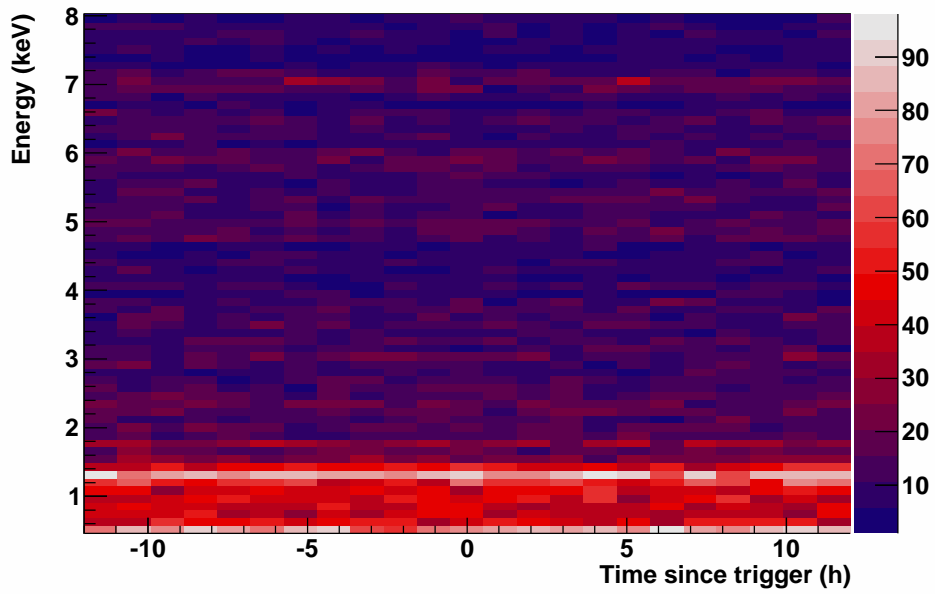
(b) $2 \rightarrow 6$ keV

Figure 3.29: Time-energy correlations in selected regions with 575 triggers in the $0.5 \rightarrow$ 1 keV region, 1200 in $2 \rightarrow 6$ keV. The z-axes are in counts.

(a) Low energy   (b) High energy
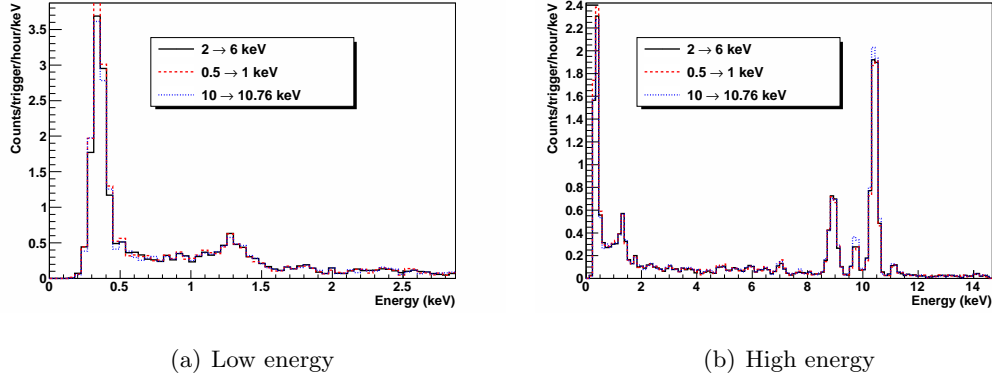
Figure 3.30: Average energy spectra around triggers in selected energy regions. The $0.5 \rightarrow 1$ and $2 \rightarrow 6$ keV regions were integrated ± 12 hours around a trigger, the $10 \rightarrow 10.76$ keV region ± 3 hours. As expected, there is a slight enhancement in the ^{68}Ga K-capture peak for triggers in the ^{68}Ge region. The source of other enhancements at low energies has not been determined.

3.6 Systematic Error Summary

The previous sections have discussed possible systematic errors in detail. To summarize, systematic errors may arise from the following:

- Threshold drift

- Contributions at low energy from non-Poisson processes (environmental noise)

- Cut efficiencies (errors in the estimation of microphonics and rise-time cut efficiencies)

- Estimation of detector fiducial mass

The threshold drift can be estimated by considering that the σ of the baseline (Section 3.5.1) only varies by 0.1 mV. Since the magnitude of the baseline is subtracted for every events, drifts over long time scales shouldn't affect the estimation of the energy on an event-by-event basis. Therefore, we expect that the drift in threshold can be estimated using the σ of the baseline, which corresponds to roughly 7 eV in energy units.

An upper limit on the contribution of a non-Poisson process to the count rate at $0.5 \rightarrow 1$ keV can be made by conservatively assuming that it must be less than or equal to the error on the measured rate in that region. That is, the rate (after LN cuts) is measured as 0.172 ± 0.007 counts/hour in that region. Therefore, a non-Poisson process could contribute $\leq 4\%$ to events in this region.

The *acceptance* efficiencies for the rise-time cuts have been estimated by performing a systematic analysis on the behavior of the cuts on peaks in the data. At higher acceptance ($\geq 70\%$), the estimated efficiencies are correct within 10%. This error is not explicitly taken into account, but the effect of the different cuts is studied during limit calculations (Section 4.4.4). The error is biased since the efficiencies are always underestimated, meaning that limit estimations will be conservative. The detector fiducial mass has been estimated as 0.346 ± 0.0124 kg, but the value of this has been conservatively assumed to be at the lower end of this error: 0.33 kg. As with the rise-time efficiency underestimation, this will make limit calculations more conservative since a decrease in assumed mass reduces the sensitivity of a detector.

3.7 *Conclusions*

During the counting runs presented here, the detector and electronics demonstrated good stability, running without significant intervention for almost half a year. At the time of writing, data collection continues. A robust analysis chain developed for a previous P-PC detector at Soudan was applied to the distinct DAQ setup used for the modified BEGe, demonstrating the adaptability required for use with the MAJORANA DEMONSTRATOR. Tools were developed to measure the rise-time of preamplifier traces with small signal-to-noise ratios, and the results can be applied to future detectors with similar or better capabilities including P-PC detectors in the DEMONSTRATOR. Future DAQ systems for the DEMONSTRATOR will likely forego the use of spectroscopic amplifiers, opting for the simplicity afforded by directly digitizing the preamplifier output. Therefore, the study of microphonics cuts based upon preamplifier traces must progress, though it is expected that the ability to differentiate between 'real' events and those vibrationally induced will improve because the deviations in preamplifier traces should manifest more significantly.

Table 3.10: Summary of estimation of systematic errors.

Source	Estimate	Comments
Threshold drift	$\lesssim 7$ eV	Estimated from baseline versus time measurements, negligible/ignored
Non-Poisson process contribution	$\leq 4\%$ of counts	Possible 'unknown background' contribution to $0.5 \rightarrow 1$ keV region.
Rise-time cut efficiency error	$\lesssim 10\%$	For efficiencies $\geq 70\%$; effects of different cuts studied during limit calculations (Section 4.4.4)
Fiducial mass error	4%	Error ignored, but mass conservatively assumed to be 0.33 kg

The study of the rise-time cuts demonstrated the ability to cut slow-rise-time pulses, but also implied that further work must progress to establish the robustness of this technique near threshold. Additionally, determining if other origins of slow pulses exist and how the distribution of these pulses depends on geometry and manufacturing requires further quantitative tests with other similar detectors. The cuts and analysis of this data have readied it for further interpretation: possible sensitivities and exclusion limits are discussed in the following chapters.

Chapter 4

LIMITS ON LIGHT WEAKLY-INTERACTING MASSIVE PARTICLES (WIMPS)

As discussed in Chapter 1, p-type point-contact detectors have sensitivity to low-mass WIMPs given their intrinsic low noise and low-energy thresholds. This chapter presents the framework and methodology required for deriving limits on these particles given the data taken with the modified-BEGe detector at Soudan Underground Laboratory. Emphasis is given on handling the difficulties described in the previous chapter related to unknown backgrounds at low energy.

4.1 Signal from WIMP Dark Matter

A WIMP from the dark matter halo of the galaxy may interact with matter by recoiling off nuclei. This interaction is in general dependent upon properties of the halo – e.g. density, escape velocity – and characteristics of the detector – e.g. detector mass, atomic number. The mathematical form has been derived in several reviews [18, 17]. The WIMP signal used in this analysis was the most generic parameterized form from Lewin and Smith [18] given as a differential rate per recoil energy of the nucleus:

$$
\begin{aligned}
\frac{dR(v_E, v_{esc})}{dE_{rec}} = & \\
\frac{k_0}{k_1} \frac{R_0}{E_0 r} & \left\{ \frac{\pi^{1/2}}{4} \frac{v_0}{v_E(t)} \left[\mathrm{erf}\left(\frac{v_{min} + v_E(t)}{v_0} \right) - \mathrm{erf}\left(\frac{v_{min} - v_E(t)}{v_0} \right) \right] - e^{-v_{esc}^2/v_0^2} \right\}
\end{aligned} \tag{4.1}
$$

with

$$
v_E(t) = v_{E_0} + v_{E_1} \sin(2\pi t)
$$

$$
\frac{k_0}{k_1} = \left[\mathrm{erf}\left(\frac{v_{esc}}{v_0} \right) - \frac{2}{\sqrt{\pi}} \frac{v_{esc}}{v_0} e^{-v_{esc}^2/v_0^2} \right]^{-1}
$$

R_0 is related to the WIMP-nucleus cross section, σ_{nucl} by:

$$R_0 = \frac{2}{\sqrt{\pi}} \frac{N_A}{A} \frac{\rho_D}{M_W} \sigma_{nucl} v_0$$

The above equations have a number of parameters which are described as follows:

E_{rec} – Recoil energy of the nucleus.

E_0 – Energy of WIMP moving with velocity v_0

r – Dimensionless reduced mass given by $\frac{4 M_W M_T}{(M_W + M_T)^2}$

v_0 – Velocity parameter in the Maxwellian velocity distribution of the dark matter halo

v_{min} – Minimum WIMP velocity which can give a recoil energy E_{rec}

v_{esc} – Dark matter halo escape velocity

v_{E_0}, v_{E_1} – Parameters describing the velocity of the earth

M_W – Mass of a WIMP

M_T – Mass of a nucleus in the detector material

A – Atomic number of a nucleus in the detector material

ρ_D – Dark matter halo density

N_A – Avogadro's number

σ_{nucl} – WIMP-nucleus cross section at zero velocity

This distribution is generally approximated using a single exponential in energy [18]:

$$\frac{dR(v_E, v_{esc})}{dE_{rec}} = \frac{k_0}{k_1} \frac{R_0}{E_0 r} \left(c_1 e^{-c_2 E_R / E_0 r} - e^{-v_{esc}^2 / v_0^2} \right) \tag{4.2}$$

with constants, c_1, c_2 slightly varying with time. The complete form (Equation 4.1) was chosen to allow eventually performing two-dimensional fits over both time and energy. For these results the time dependence of the WIMP interaction is ignored and $t \to 0$, which is an excellent approximation of the average of Equation 4.1.

The WIMP-nucleus cross section is valid at small momentum transfers, $q = \sqrt{2M_T E_{rec}}$, when h/q is large compared to the size of the nucleus. As the momentum transfer increases, the cross section for a coherent-nuclear-recoil interaction decreases. This reduction in cross section is parameterized with a form factor to account for the dependence on momentum transfer: $\sigma_{nucl} \to \sigma_{nucl} F^2(qr_n)$ where r_n is the effective nuclear radius. In this analysis, the Woods-Saxon/Helm form factor [71] was used:

$$F(qr_n) = 3\frac{j_1(qr_n)}{qr_n}e^{-(qs)^2/s} \tag{4.3}$$

with (from [18])

$$qr_n = \frac{\sqrt{2M_T E_{rec}}}{197.3}1.14A^{1/3}$$

and s is the nuclear skin thickness. In practice, this correction is implemented by replacing all σ_{nucl} with $\sigma_{nucl}F^2$ or, equivalently, by multiplying the differential rate by the form factor squared. A plot of this form factor for germanium is shown in Figure 4.1, indicating that the form factor can provide a significant correction to the cross-section for ionization energies above ~ 10 keV and a $\sim 10\%$ correction in the 0$\to$3 keV region.

The WIMP-nucleus cross section, σ_{nucl}, is defined for a particular nucleus. However, it is common to scale this parameter to a WIMP-nucleon cross section, σ_{W-n}, to compare between experiments using different detectors and therefore distinct nuclei. The relationship from [72] is given as:

$$\sigma_{W-n} = \left(\frac{\mu_1}{\mu_A}\right)^2 \frac{1}{A^2}\sigma_{nucl} \tag{4.4}$$

where $\mu_A = \frac{M_W M_T}{M_W + M_T}$ and μ_1 is defined at $A = 1$. All exclusion plots will be presented in terms of σ_{W-n}, but all other results will be presented in terms of σ_{nucl}.
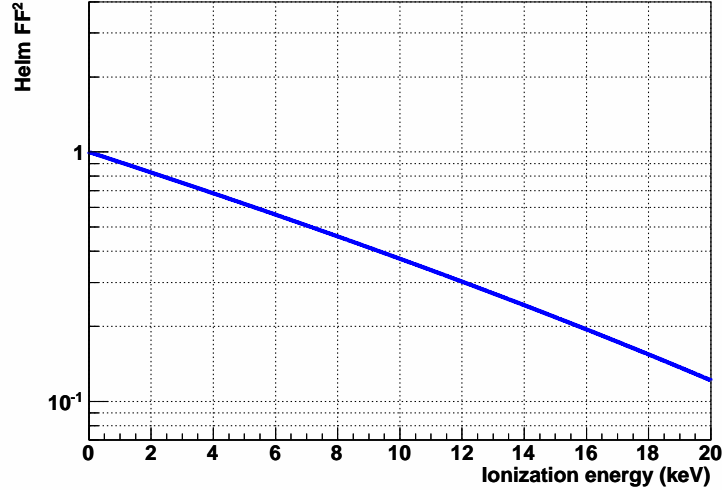
Figure 4.1: A plot of the Ge form factor versus ionization energy. For a discussion of the conversion from nuclear recoil energy to ionization energy, see Section 4.2.

.

4.2 Quenching

In ionization detectors, the calibration relating the energy deposited per amount of charge collected is generally performed by using gamma lines from a known source. However, since photons interact by recoiling off electrons in the detector, this calibration is only valid for other processes which deposit energy in a similar manner. In contrast, a particle that deposits energy in a detector by recoiling off a nucleus (such as a WIMP) does not generate the same amount of charge per amount of energy deposited as an electron-recoil interaction. The relationship between the 'seen' ionization from a nuclear recoil and the 'seen' ionization from an electron recoil at the same energy is referred to as quenching. For ionization detectors, a theory has been developed by Lindhard et al. [73] to parameterize this relationship:

$$E_{ion} = \frac{kE_{rec}g(\epsilon)}{1 + kg(\epsilon)} \qquad (4.5)$$

with

$$g(\epsilon) \;=\; 3\epsilon^{0.15} + 0.7\epsilon^{0.6} + \epsilon$$

$$\epsilon \;=\; 11.5 E_{rec} Z^{-7/3}$$

For germanium detectors, the constant, k, has been recently estimated as 0.2 by Barbeau, et al. [52] in the energy region $\lesssim$ 4 keV. In a single-mode readout detector (e.g. reading out *only* ionization), nuclear recoil and electron recoil events are indistinguishable and so it is necessary to transform all theoretical spectra defined in recoil energy to ionization energy. Unfortunately, accomplishing this requires inverting the Lindhard equation which is impossible to do analytically and so it essential to either obtain it numerically or to estimate this function with another function. The function $E_{ion} = \alpha E_{rec}^{\beta}$ provides an excellent approximation to this equation and can be easily inverted and differentiated. Results from a fit of this function to the Lindhard equation (Eqn. 4.5) for germanium detectors are shown in Figure 4.2. Fit values for α and β are 0.204971 and 1.13615, respectively.

Combining this all together to determine the WIMP time and energy spectrum in terms of ionization energy yields the final differential rate:

$$\frac{dR}{dE_{ion}} = \left(\frac{dR}{dE_{rec}} \right) \left(\frac{dE_{rec}}{dE_{ion}} \right) F^2 \tag{4.6}$$

A summary of the parameters used is given in Table 4.1. The effect of including the escape velocity and the form factor for a low-mass (10 GeV) WIMP is shown in Figure 4.3 where $t = 0$. This plot demonstrates the importance of a low threshold for detecting low-mass WIMPs since the escape velocity effectively truncates the WIMP spectrum above a certain energy. Therefore, count rates for low-mass WIMPs become negligible for detectors with high enough thresholds. The small correction from the form factor can be seen as a reduction in the count rate. An example two-dimensional (energy and time) WIMP spectrum is shown in Figure 4.4. In both of these plots, the differential rate is divided by the WIMP-nucleus cross-section.

Table 4.1: Summary of parameters used in the WIMP exclusion fits.

Parameter	Value (unit)
Fit parameters	
Atomic Mass	72.96 (amu)
ρ_D	0.4 (GeV cm^{-3})
Average velocity, v_0	230 (km s^{-1})
DM escape velocity, v_{esc}	600 (km s^{-1})
Velocity of the earth, v_{E_0}	244 (km s^{-1})
Detector mass	0.33, 0.4[a](kg)
Nuclear skin thickness s	0.9 (fm)
Data set limits	
Threshold	0.5 (keV)
Max energy	3.5 (keV)

[a] The smaller mass value is used when rise-time cuts are applied, see Section 3.4.3.
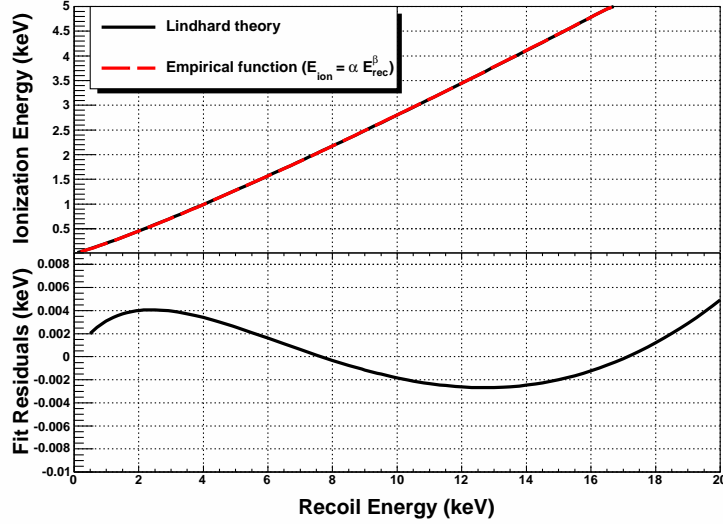
Figure 4.2: Lindhard function with a fit to $E_{ion} = \alpha E_{rec}^{\beta}$. The bottom plot displays the residuals, indicating that the estimating function is good to better than 1% in the ionization energy range 0.5→3.5 keV.

4.3   Fit Methodology

The analysis of the modified-BEGe data presented in the previous chapter, in particular Section 3.4.5, uncovered several difficulties which must be addressed when determining limits with the data. This section discusses the extraction of a signal limit using the maximum-likelihood method and explores the necessary mechanisms for dealing with unknown backgrounds.

4.3.1   Obtaining Limits Using Maximum Likelihood

A maximum-likelihood analysis provides a mechanism for estimating the parameters of a model given a set of data. There also exist standard techniques within the maximum-likelihood framework for hypothesis testing or searching for and generating limits on a signal that may exist in the data set. In contrast to using χ^2-minimization for parameter estimation, maximum likelihood does not require the binning of data and it does not provide
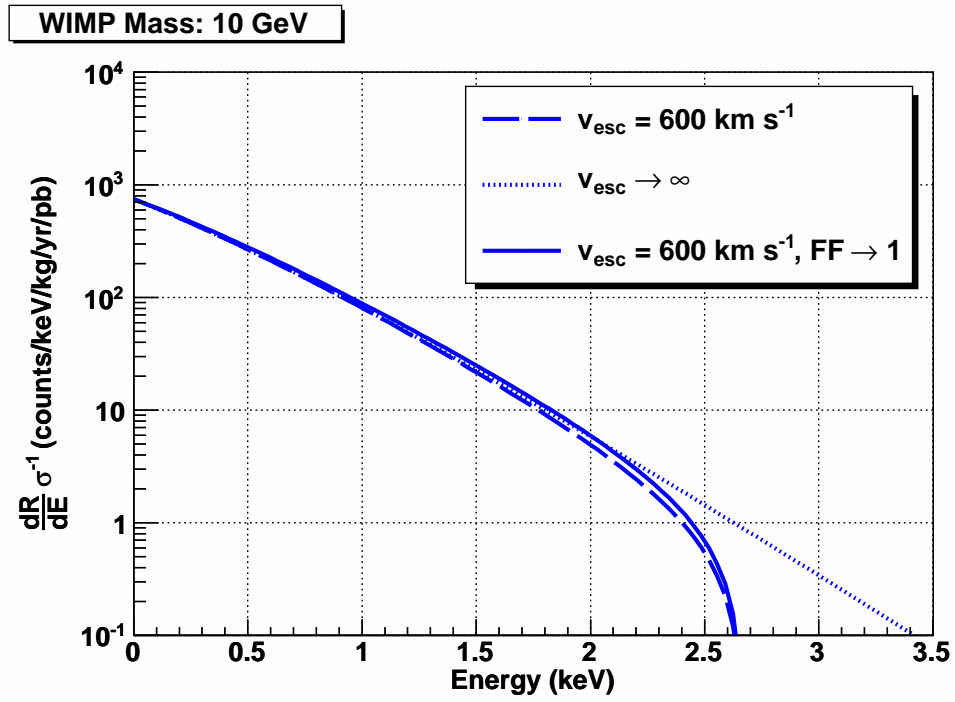
Figure 4.3: Count rate/cross section vs. ionization energy for a 10 GeV WIMP for 3 different model variations: (1) Finite escape velocity and Helm form factor, (2) infinite escape velocity, and (3) finite escape velocity and a form factor of 1.
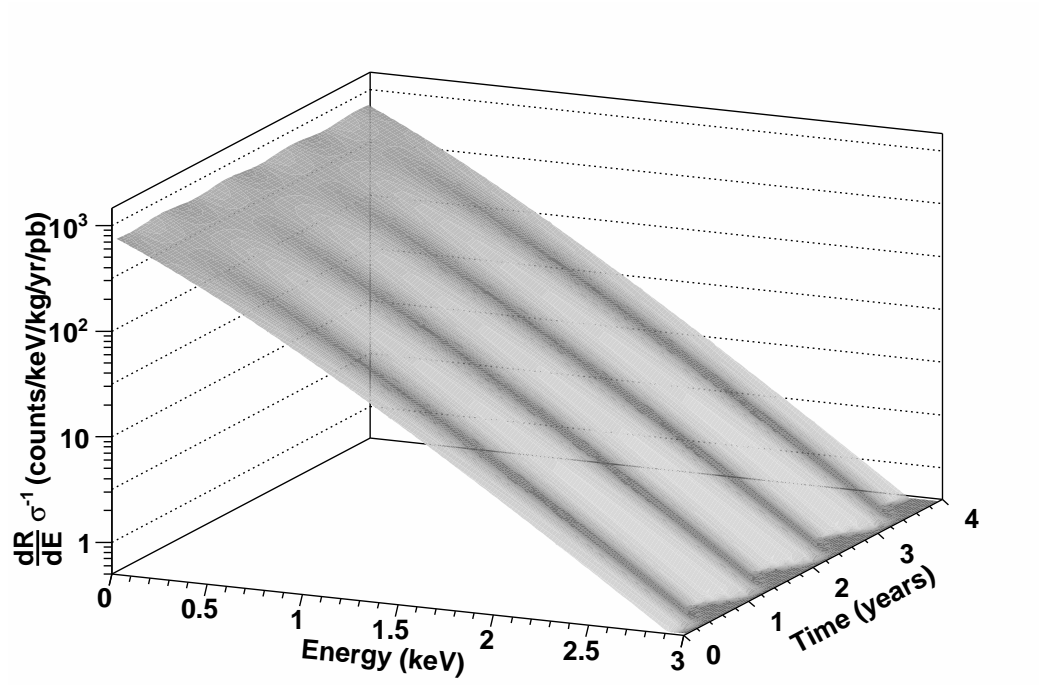
Figure 4.4: 2-dimensional WIMP dark matter signal, $M_{WIMP} = 10$ GeV.

a mechanism for estimating the goodness-of-fit. The procedure for determining the set of parameters which best describes a data set involves constructing a likelihood function as the following:

$$L = \prod^{n} f(\boldsymbol{\theta}, x_n) \tag{4.7}$$

where $f(\theta, x)$ is a probability distribution function (pdf) describing the data, x_n, with parameter set $\boldsymbol{\theta}$. The true values of the parameters, $\boldsymbol{\theta_T}$, are then estimated by maximizing L, or, equivalently, by minimizing $-\log L$. For some more-trivial pdfs, it is possible to find an analytical solution to the equation set: $-\frac{\partial(\log L)}{\partial \boldsymbol{\theta}} = 0$, but normally the minimization of the function is handled numerically.

Hypothesis testing defines a mechanism by which one compares two models to the same data to find the favored model. It is useful when searching for limits on a signal since one can compare the null hypothesis (no signal) versus the model with signal included. Hypothesis testing using likelihood functions is performed by using a profile-likelihood test which involves the construction of a profile-likelihood function over a parameter or set of parameters, θ_0:

$$\lambda(\theta_0) = \log L_{max} - \log L_{max}(\theta_0, \boldsymbol{\theta_n}) \tag{4.8}$$

assuming that θ_0 is a subset of $\boldsymbol{\theta}$, or $\boldsymbol{\theta} = \{\theta_0, \boldsymbol{\theta_n}\}$. L_{max} is the likelihood function maximized for all parameters, and $L_{max}(\theta_0, \boldsymbol{\theta_n})$ is the likelihood function at parameter value θ_0 maximized over all other parameters except θ_0. The profile-likelihood function, λ is asymptotically distributed according to a χ^2 distribution with number-of-degrees-of-freedom (NDF) equal to the dimensionality of θ_0, so that:

$$2\lambda(\theta_0) \sim \chi^2_{dim(\theta_0)} \tag{4.9}$$

This relationship can be used to determine a confidence interval for θ_0.

To demonstrate this methodology, consider the search for a signal, S, in some background with distribution, B, so that the log-likelihood function, following Equation 4.7, is:

$$-\log L(\alpha, \boldsymbol{\beta}) = -\sum^{n} \log(\alpha S(x_n) + B(\boldsymbol{\beta}, x_n)) \tag{4.10}$$

where we have assumed that the only variable parameterizing the signal, S, is its amplitude, α. The true values of the model parameters are estimated by maximizing L to obtain α_0

and $\boldsymbol{\beta_0}$. The upper and lower limits of α are then determined by finding the values α_{upper} and α_{lower} so that $\lambda(\alpha_0 + \alpha_{upper}) = P_{\chi_1^2}(CL)/2 = \lambda(\alpha_0 - \alpha_{lower})$. $P_{\chi_1^2}(CL)$ is the χ^2 quantile for 1 degree-of-freedom at the CL (confidence level) percentage. For example, if we assume a confidence level of 90%, we solve the equations $\lambda(\alpha_0 + \alpha_{upper}), \lambda(\alpha_0 - \alpha_{lower}) = 2.71/2$ to find α_{lower} and α_{upper}. Typically, these solutions are found by scanning the parameters of interest (α in this case) in the profile-likelihood function around the global extremum on a grid size determined by the desired precision. More information regarding the profile-likelihood method and calculating confidence intervals with it can be found in [74].

4.3.2   Obtaining Limits in the Presence of Unknown Backgrounds

In some experiments it is sometimes difficult or impossible to determine precisely the background contamination in a region-of-interest for a signal. This could arise for several reasons: e.g. (1) poorly understood systematics from simulation, (2) overlapping signal and background in the region of interest, or (3) the inability to perform independent measurements of background by, for example, performing background-only data runs. Especially in dark matter experiments, some or all of these could present an issue when attempting to understand data and generate limits on potential signals. This is largely due to several factors, including that detectors looking for dark matter might operate at or close to their noise limits in poorly understood regions (e.g. near threshold), signal can not be 'turned off' to get a pure-background measurement, and simulations in low-energy regions might not offer precise estimates of backgrounds. Several methods have been proposed to deal with poorly understood or unknown backgrounds, including the Maximum-gap method described in [75] and a method proposed by Rolke et al. [76] to treat the uncertainties in the background as statistical errors. These methods are generally applicable to experiments with very low count rates in the signal region. A method used by the CRESST experiment [77] proceeded by initially fitting the spectrum to an empirically determined function using maximum likelihood and then using a profile-likelihood-like technique to determine the exclusion on a WIMP signal. Another method similar to the CRESST method but with deeper numerical study has been proposed by Rolke et al. to handle uncertainties in background by using a

profile-likelihood technique [78]. The Rolke method is used in this analysis.

The Rolke method treats the background as a nuisance parameter in the likelihood analysis, essentially using the data to estimate both the background and the signal simultaneously. The technique itself is roughly equivalent to the profile-likelihood method, but accommodates boundary cases such as when the best estimate of a parameter constrained to be greater than 0 (e.g. as in the case of a cross-section or number of counts) is less than 0, or when the likelihood function has no maximum as is the case when the signal and background look very similar or the number of data seen is less than that expected in background. Rolke et al. presented two techniques for dealing with these issues: an unbounded-likelihood method and a bounded-likelihood method; this analysis makes use of the latter. The bounded-likelihood method is no different than the profile-likelihood method if the best fit of the parameter is in a physical or 'acceptable' region. However, if the best-fit parameter moves into an unphysical region, then the parameter is forced to its nearest physical boundary and the maximum likelihood is set as the value of the profile likelihood at that boundary.

This can be made clearer by considering the example in Section 4.3.1: In this example, α must be greater than or equal to 0 since a negative signal has no physical meaning. However, it is possible that a best fit can yield an estimate for α of $\alpha_0 < 0$. In this case, α_0 would be quoted as 0 and, for purposes of calculating an upper limit, L_{max} would be replaced with the likelihood function L maximized at $\alpha = 0$. This can be summarized with the following equations:

$$
\begin{aligned}
\alpha_0' &= \alpha_0 H(\alpha_0) \\
\lambda'(\alpha) &= \lambda(\alpha) - \lambda(\alpha \to 0)(1 - H(\alpha_0))
\end{aligned}
\tag{4.11}
$$

where the primed values are used in the bounded-likelihood method and $H(x)$ is the Heaviside step function. In this formalism, it is clear that the primed variables become the original unprimed variables for $\alpha_0 \geq 0$. The upper limit on α is calculated similarly as before by increasing α until the profile likelihood, $\lambda'(\alpha)$, reaches the value of the desired χ^2 quantile.

Especially when the parameters of the fit are near or at their boundaries, it is not ensured that λ' follows a χ^2 distribution and so it is essential to perform Monte Carlo studies to

investigate the coverage of this technique for a particular model. This can be done, for example, by the following:

- Scan the allowed parameter space of a model, generating a data set at each point in parameter space via Monte Carlo

- Perform a likelihood analysis using the Rolke method, calculating limits at an assumed confidence level, CL

- Determine if the calculated limits include the 'true' value(s) (i.e. those values of the model used to generate the data set) of the parameter(s) of interest

- Repeat many times to measure how often the calculated limits include the true values

In essence, this is a guess-and-check test: one initially *assumes* that the profile likelihood follows a χ^2 distribution and then performs tests to ensure that it in fact does. A model is over covered if the percentage of times the calculated limits include the true values is greater than CL and is under covered if the opposite is true. Rolke et al. provide several concrete models where they demonstrate that their technique properly covers the model parameter space, i.e. that the calculated coverage is at least the expected CL. In practice, determining and scanning the allowed parameter space for coverage tests is difficult: for this analysis, this is discussed in Section 4.4.3.

4.4 Dark Matter Limits Using Data from a Low-background Modified-BE-Ge Detector at Soudan Underground Laboratory

The data described and analyzed in Chapter 3 were used to calculate limits on light dark matter following the procedures outlined in previous sections. This section describes the data and model used during fitting and explores difficulties that arose during generation of these results. Several systematic tests were performed to investigate how details of the fitting procedure might affect the results. For systematic tests, a smaller subset of the data with ∼2 months of live-time was used, however the conclusions derived from this limited data set remain valid when considered for the entire data set.

4.4.1 Data and Model

The likelihood fitting was performed using the RooFit [1] framework developed by Wouter Verkerke and David Kirkby within the ROOT analysis package [35]. RooFit provides an abstract framework to perform different types of fitting, including both binned and unbinned maximum likelihood as well as chi-square minimization. The package includes a set of pdfs from which one can construct more complex models and enables the user to generate his or her own models which could not have otherwise been constructed from the provided building blocks. A toolkit extension was built within the RooFit framework including several pdfs of WIMP interactions. More information on this constructed WIMP fitting framework can be found in Section B.3. RooFit takes the built pdf models, automatically normalizes them, and constructs the appropriate negative log-likelihood or χ^2 function which can then be minimized using numerical algorithms (i.e. Minuit [79]).

The likelihood function for the fit was constructed from the following pdfs (normalization not shown):

- Flat background – $f_{flat}(E) = 1$

- Exponential background – $f_{exp}(E) = \exp\left(c_1 E\right)$

- Ge L-capture line – $f_{Ge}(E) = \frac{1}{\sigma_{Ge}\sqrt{2\pi}} \exp\left(-\frac{(E-\mu_{Ge})^2}{2\sigma_{Ge}^2}\right)$

- Zn L-capture line – $f_{Zn}(E) = \frac{1}{\sigma_{Zn}\sqrt{2\pi}} \exp\left(-\frac{(E-\mu_{Zn})^2}{2\sigma_{Zn}^2}\right)$

- WIMP pdf – $f_{WIMP}(E) =$ Function 4.6

These pdfs were combined additively, each with an associated parameter – N_{flat}, N_{exp}, N_{Ge}, N_{Zn}, N_{WIMP} – measuring the number of events in each distribution. The sum of these parameters, $\sum N_x$, was included in the likelihood function so that the final extended likelihood function (with each pdf, f_x, automatically normalized over the range of the fit)

---

[1]See http://roofit.sourceforge.net/.

was:

$$-\log L = -\log\left(f_{Pois}\left(\sum_x N_x, N_{obs}\right)\right) - \sum_i w_i \log\left(\frac{1}{\sum_x N_x}\left(\sum_x N_x f_x(E_i)\right)\right) \quad (4.12)$$

where the sums over x are over the pdfs in the distribution, the sum over i is over each data point or bin, and w_i is the weighting for a particular data point with $N_{obs} = \sum_i w_i$. The weight of each event, i, is determined by the inverse of the total efficiency function (see Section 3.4.5) at the energy of the event, E_i. The formulation of the WIMP pdf in Equation 4.6 actually provides counts/kg/keV/day and so the number of counts, N_{WIMP}, in this pdf is not an independent parameter but rather equal to the integration of f_{WIMP} over the fit energy range multiplied by the time of the experiment and the mass of the detector. This meant that N_{WIMP} was proportional to the WIMP-nucleus cross section, σ_{nucl}, and so constraints on the number of WIMP interactions directly related to a limit on the cross section. Therefore, the profile-likelihood function for σ_{nucl}, $\lambda(\sigma_{nucl})$, was calculated directly. The mass of the WIMP, M_W is also a free parameter which defines the shape of the distribution. To determine limits on σ_{nucl} for a range of WIMP mass values, M_W was stepped through a range from $\sim 4 \rightarrow 100$ GeV, calculating limits on σ_{nucl} at each mass value.

The mean values of the L-capture lines (μ_{Zn}, μ_{Ge}) were fixed (1.1 and 1.299 keV) and the sigmas (σ_{Zn}, σ_{Ge}) were fixed according to the empirically determined resolution function from Equation 3.3. All other parameters were allowed to float, though the numbers of the flat and exponential background, N_{exp}, N_{flat}, were constrained greater than or equal to 0 and the shape parameter, c_1, of the exponential was allowed to float only slightly positive. A summary of the parameters and their ranges is given in Table 4.2. The inclusion of the exponential background in the null model was not due to any *a priori* assumption or background simulation, but instead was a mechanism to quantify our agnosticism as to the source of this shape in the data. The cause could be from several factors in indeterminate combination: (1) noise fluctuations – deviations in noise could manifest as a widening of the noise pedestal (Gaussian) which would appear exponential; (2) untagged microphonics – microphonics can generate an exponential at low energies [38]; (3) slow-rise-time-event contamination – an incomplete rejection of slow events could induce this shape, see Sec-

Table 4.2: Allowed ranges and values of parameters used in the WIMP fit. Limits were chosen to avoid obtaining best-fit parameters near limit boundaries. Therefore, some parameters – e.g. cross sections – were allowed to float to unphysical regions.

Parameter	Range	Unit
Ge L-capture mean, μ_{Ge}	1.299 (fixed)	keV
Ge sigma, σ_{Ge}	7.55×10^{-2} (fixed)	keV
Zn L-capture mean, μ_{Ge}	1.1 (fixed)	keV
Zn sigma, σ_{Ge}	7.48×10^{-2} (fixed)	keV
WIMP-nucleus cross section, σ_{nucl}	$-10 \rightarrow 100$[a]	pb
Exponential shape parameter, c_1	$-100 \rightarrow 5$	keV^{-1}
N_{flat}	$0 \rightarrow 10^5$	counts
N_{exp}	$0 \rightarrow 10^5$	counts
N_{Ge}[b]	$0 \rightarrow 10^5$	counts
N_{Zn}[b]	$0 \rightarrow 10^5$	counts

[a] The upper limit was expanded dynamically to ensure that $\lambda(\sigma_{nucl})$ would exceed the desired χ^2-quantile value.

[b] For fits when the relative amplitude of the two L-lines was constrained, these two parameters became a single parameter with the same limits.

tion 3.4.3; (4) other unknown environmental (electronic and/or temperature) variations; (5) low-mass WIMP interactions. The source of this shape is discussed in more detail in Section 3.4.5. Only after all sources of background have been independently and conclusively measured can one determine a more precise background model to feed back into the likelihood function. To generate limits using this data given our current state of knowledge, it must suffice to treat the shape and amplitude of this background as nuisance parameters and follow the prescription of Rolke et al. [78] as described in Section 4.3.2. However, since the WIMP signal is almost indistinguishable from an exponential in energy space, the inclusion of an unconstrained exponential posed some difficulties during the fits; these and other difficulties are outlined in Section 4.4.2. Fits with different constraints and methodologies were performed to investigate systematics of the fit and to determine how different constraints affected the calculated limits. Three main sets of fits were performed:

- Unbinned fit – No additional constraints on parameters

- Binned fit – A particular binning was chosen

- Fixed relative amplitudes of the Ge and Zn L-capture lines (with both binned and unbinned fits)

In each of these fit sets, exclusions were calculated using data with different cuts applied, including rise-time cuts of varying acceptance efficiency and microphonics cuts:

1. Microphonics and LN-fill cuts

2. Microphonics, LN-fill cuts *and* one of: 40%, 50%, 60%, 70%, 80%, 90%, 95%, and 99% acceptance rise-time cuts.

These cuts and associated data are described in detail in Sections 3.4.1 and 3.4.3. The results of the different fit sets are outlined in Sections 4.4.4 and 4.4.5.

4.4.2   Fit Difficulties and Likelihood-function Pathologies

While performing the exclusion fits, a number of issues were encountered with the log-likelihood and profile-likelihood functions. In particular, the profile likelihood did not always exhibit a parabolic shape due to three main reasons: (1) similarity of signal and background – background (both exponential and flat) could be similar to a WIMP signal at certain WIMP masses; (2) parameters at bounds during the calculation of $\lambda(\sigma_{nucl})$; and (3) unconstrained background exponential shape, leading to local minima away from the global minima. Examples of profile-likelihood functions are given in Figure 4.5 and will be referred to throughout this section. These functions were generated with data with 99% rise-time cuts applied and unbinned fits performed with constraints on the relative amplitudes of the Ge and Zn L-capture lines (see Section 4.4.5), but are exemplary of $\lambda(\sigma_{nucl})$ for all types of fits performed. This data set and model combination has been used to generate all the example results in this section.

Signal, Background Similarity

For certain values of WIMP masses, the background and signal can appear very similar. This can be seen, for example, in Figure 4.6 which is a plot displaying WIMP signals of $M_W = 8.75$ and 100 GeV together with an exponential spectrum ($e^{-3.3E}$, the best-fit shape parameter for the model without included WIMP signal to data with a 99% rise-time cut applied) and a flat spectrum both with arbitrary normalization adjusted for comparison. Because $\lambda(\sigma_{nucl})$ includes an implicit maximization of the likelihood for different values of σ_{nucl}, it is instructive to consider how the components of those fits depend on σ_{nucl}. Figure 4.7 includes two plots of counts in background components (flat, exponential and L-line amplitudes) versus σ_{nucl}, where the background components are the parameters which maximize the likelihood for a given value of σ_{nucl}. If we first consider Figure 4.7(a) (M_W = 100 GeV), it is clear that L-line background components vary only slightly with σ_{nucl} whereas the exponential and flat components vary almost linearly with the flat component changing more quickly than than the exponential component. When the limit of the flat background is reached at 0, it generates a kink in the likelihood function visible in
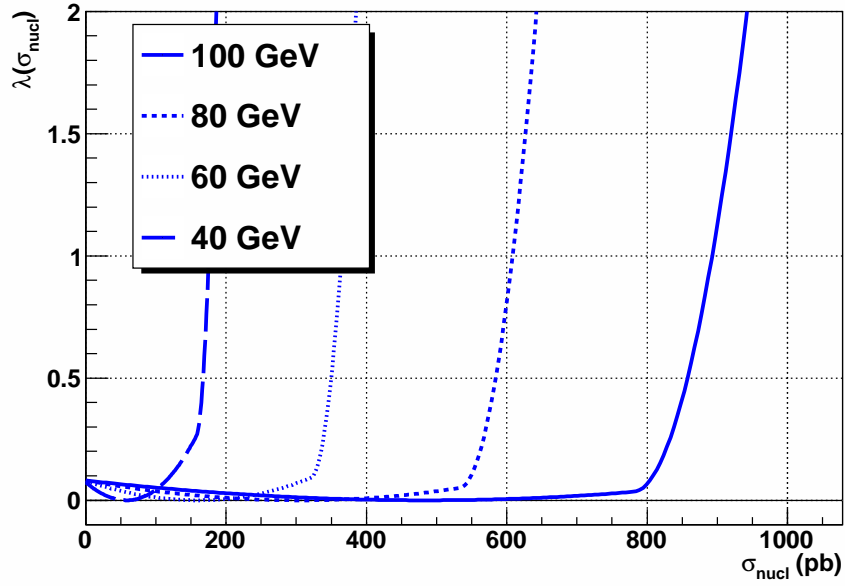
(a) WIMP masses 40→100 GeV
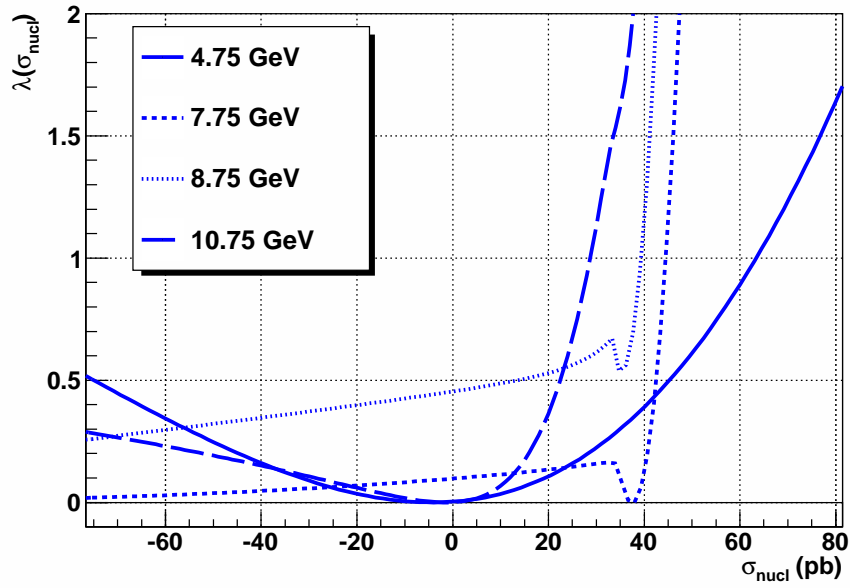
(b) WIMP masses 4→11 GeV

Figure 4.5: $\lambda(\sigma_{nucl})$ for a range of WIMP Masses. See text for discussion of the shapes seen.
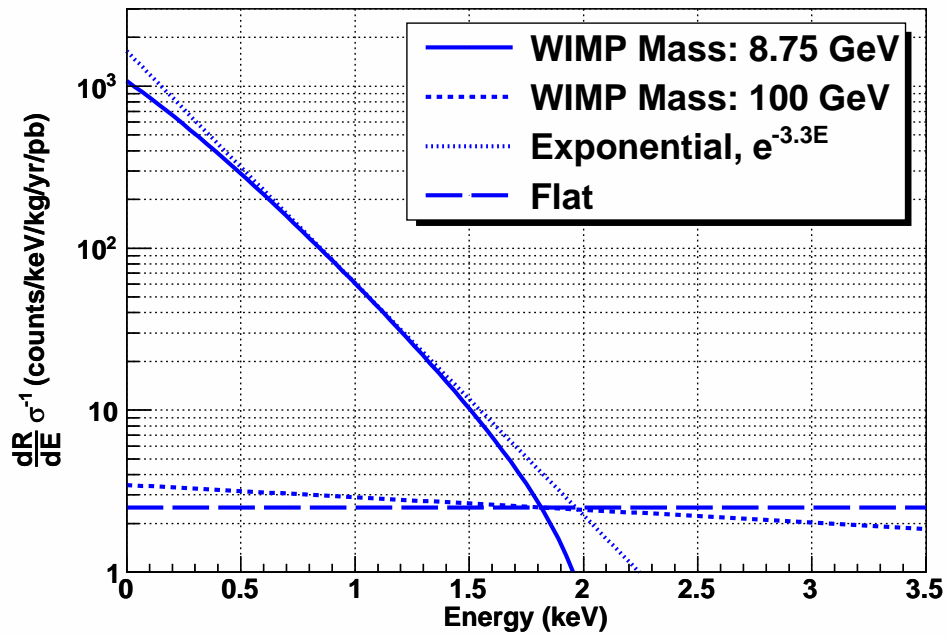
Figure 4.6: Similarity of signal and background, comparing the exponential and flat components of the background with a WIMP signal. The shape parameter of the exponential function is set to -3.3 keV^{-1}, the best-fit value from data. Flat and exponential backgrounds have arbitrary normalizations for comparison to the WIMP signals.
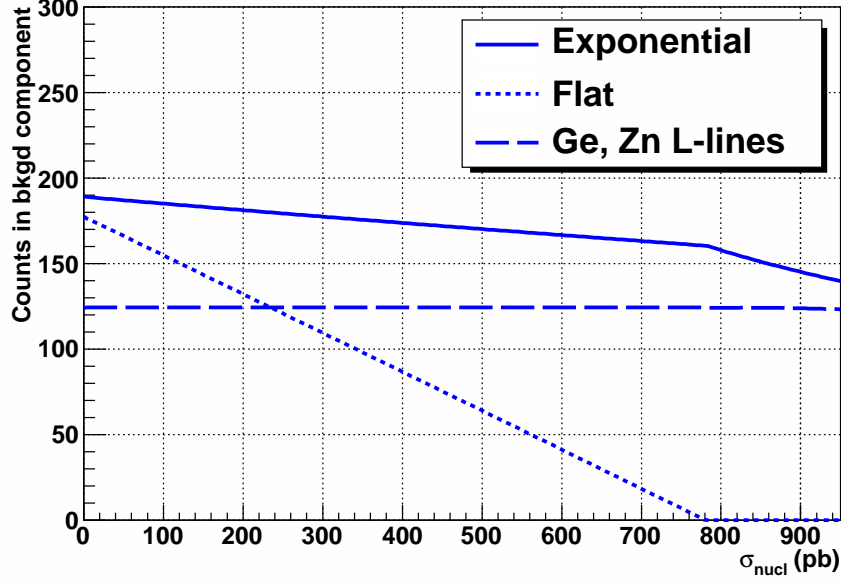
(a) WIMP mass 100 GeV
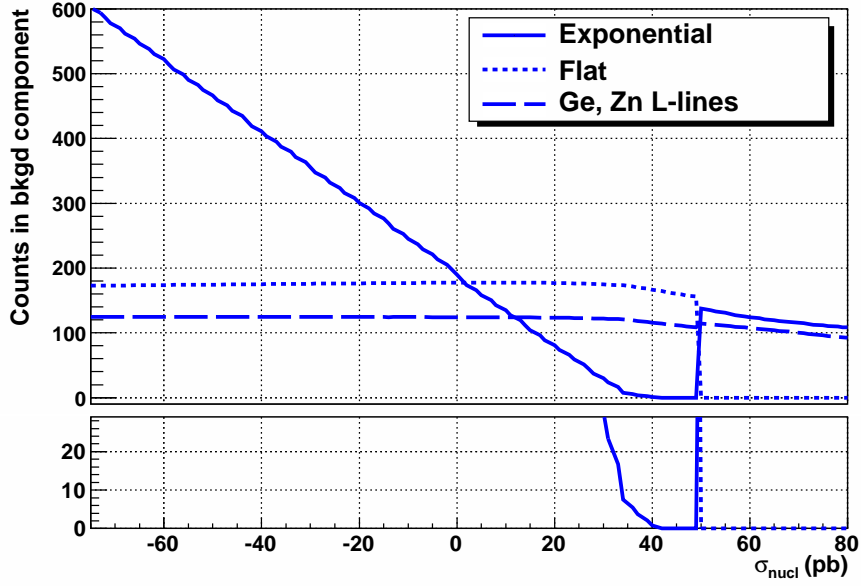
(b) WIMP mass 8.75 GeV, with zoom in on count region $0 \rightarrow 30$.

Figure 4.7: Counts in background components versus σ_{nucl}. The values of the components are determined during a profile-likelihood scan of σ_{nucl}.

Figure 4.5(a). Generally, parameters are allowed to float beyond their physically-allowed values to avoid such features in the likelihood function. However, due to the similarity of the signal and background it was found that the flat background component would continue decreasing if it were allowed to float below zero leading to a very flat profile-likelihood function. The linear variation of the flat and exponential background components with σ_{nucl} coupled with the relative independence of the other parameter indicates that the variation of $\lambda(\sigma_{nucl})$ is dominated by the extended parameter[2]: $-\log\left(f_{Pois}\left(\sum_x N_x, N_{obs}\right)\right)$. Therefore, the upper limit on σ_{nucl} calculated using the profile-likelihood method is essentially equivalent to calculating a Poisson upper limit on true counts given number of counts seen, and assuming that all these counts come from signal.

We can then consider lower WIMP mass (8.75 GeV) where the shape of the WIMP signal is roughly equivalent to the shape of the exponential background component in the data. The relevant plots for this are Figure 4.7(b) and Figure 4.5(b), focusing on M_W=8.75 GeV in the latter. It is important to note the $\lambda(\sigma_{nucl})$ for this mass does not intersect 0 in Figure 4.5(b); this is because no minimum of the $-\log L$ was found in the scanned region for σ_{nucl} and was therefore defined as the value of $-\log L$ at the lowest scanned value of σ_{nucl} (-100 pb). To define an upper limit with this function, the Rolke prescription outlined in Section 4.3.2 would be followed, defining the minimum of $-\log L$ to be at $\sigma_{nucl} = 0$. It is clear that, in contrast to the previous case with signal of a WIMP at M_W=100 GeV, the flat and L-line components of the background are largely independent of σ_{nucl}, whereas the exponential component varies almost linearly below 35 pb and reaches its lower bound of 0 around 40 pb. (The transition region between 35 and 40 pb which manifests a local minimum will be discussed in the following section.) As before, the linear variation indicates that the likelihood is dominated by the extended term because the signal and background have a roughly equivalent shape. The transition at higher σ_{nucl} (50 pb) is due to the exponential shape constant floating to ~ 0 and taking over the contributions from the flat component. This can be seen in Figure 4.8.

---

[2]It can be shown that a minimization of this term leads to a linear dependence between the counts and σ_{nucl}. If other terms in $\lambda(\sigma_{nucl})$ contribute, then the relationship becomes non-linear.
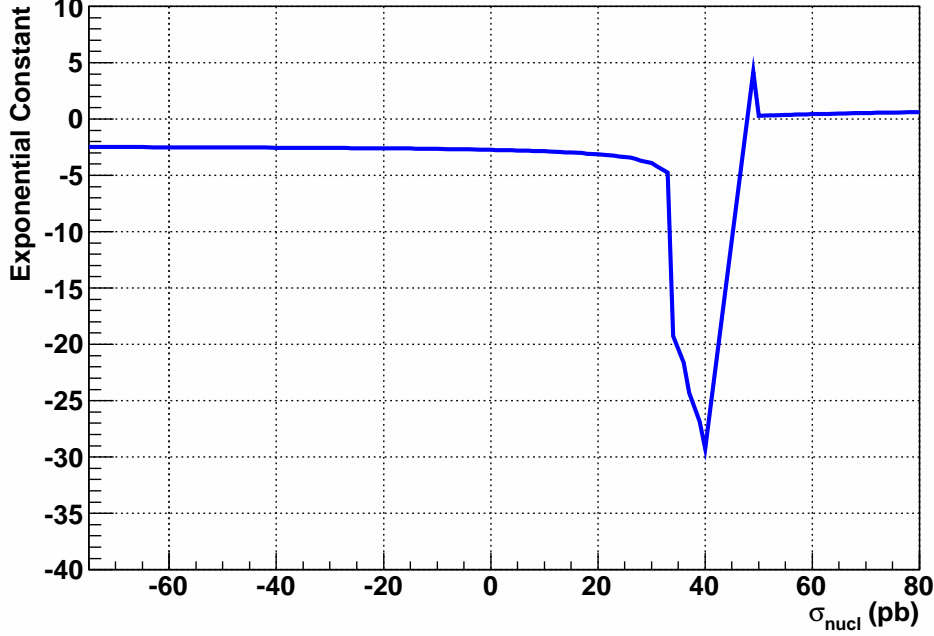
Figure 4.8: Exponential constant (shape parameter) versus σ_{nucl} for WIMP mass 8.5 GeV. The value of the exponential constant is determined during a profile-likelihood scan of σ_{nucl}.

Local Minima of $\lambda(\sigma_{nucl})$

The $\lambda(\sigma_{nucl})$ for $M_W = 8.75$ GeV in Figure 4.5(b) includes a local minimum contained in the range $34 \leq \sigma_{nucl} \leq 40$ pb. From Figure 4.7(b), it is clear this feature does not arise from the exponential background being at its limit since the amplitude does not reach its lower bound until $\sigma_{nucl} > 40$ pb. Instead, this characteristic is due to the fact that the exponential shape parameter is allowed to float to very negative numbers, producing a background function sharply decreasing with energy. (The lower limits of this parameter were kept low enough so that no fit would push the parameter to its bound.) This can be seen clearly in Figure 4.8 where the exponential shape parameter decreases sharply around 30 pb after remaining largely constant for lower values of σ_{nucl}. A complementary set of plots in Figure 4.9 provide a visualization of this abrupt change. This issue occurred when the shape of the exponential background was very similar to the shape of the applied
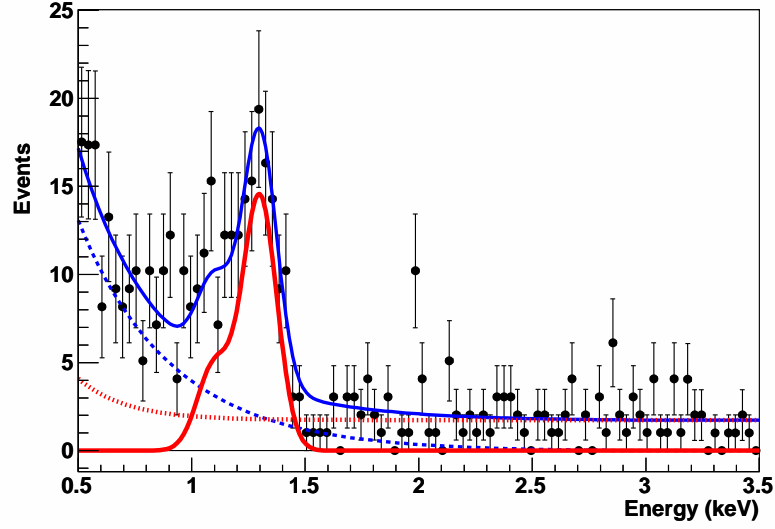
(a) $\sigma_{nucl} = 33.3$ pb
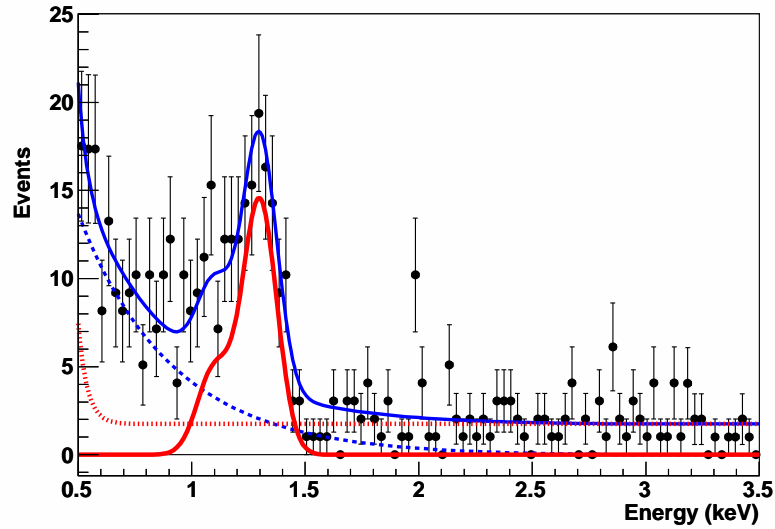
(b) $\sigma_{nucl} = 34.8$ pb

Figure 4.9: Example of how the exponential background shape changes abruptly during a profile likelihood scan for a WIMP Mass of 8.75 GeV. The separate components of the fit are shown: WIMP signal (dashed), L-capture lines (solid), exponential + flat background (dotted), and the sum (solid).

WIMP signal which meant, for example, that the affected WIMP mass range was different for fits generated with data with rise-time cuts applied and for fits made on data with only microphonics cuts applied. In particular, the rise-time-cut data exhibited a sharper exponential decline (more negative exponential shape constant) than the microphonics-cut data yielding an affected range of $\sim$6-9 Gev versus 7-10.5 GeV, respectively. The abrupt variations in the exponential shape parameter also induced features in exclusion plots for σ_{W-n}; these are discussed in Sections 4.4.4 and 4.4.5.

Conclusions

The realization of several non-parabolic features underscores the care which must be taken when deriving limits from models with similar background and signal. Some methods (e.g. the HESSE functionality in MINUIT [79]) estimate the error on the parameter by assuming a parabolic shape and extrapolating $\lambda(\sigma_{nucl})$ around its minimum by using the measured second derivative at the extremum. For non-parabolic $\lambda(\sigma_{nucl})$ functions it is clear this won't work, and could possibly yield inappropriate bounds on σ_{nucl}. For example, in Figure 4.5(b) the $\lambda(\sigma_{nucl})$ for $M_W = 7.75$ GeV indicates a minimum at $\sim$37 pb: a parabolic estimate of the lower bound would be significantly non-zero whereas an estimate using the Rolke method would define the lower bound as 0! Of course, as with all Frequentist methods it is essential to estimate the performance of the method using Monte Carlo tests; results of such an investigation are presented in Section 4.4.3.

4.4.3   Coverage Tests

Rolke et al. emphasize that correct application of their method [78] must be accompanied by a coverage test analyzing its effectiveness. The protocol for this test is outlined in Section 4.3.2 where it is made clear that the *entire* range of possible parameter space must be scanned. For all but the most simple models, this is essentially impossible and so it is necessary to reduce the parameter space to a manageable size by selecting those parameter combinations which are the most 'likely'. Since we are primarily concerned with the coverage of the model versus the amount of included signal (proportional to σ_{nucl}), it
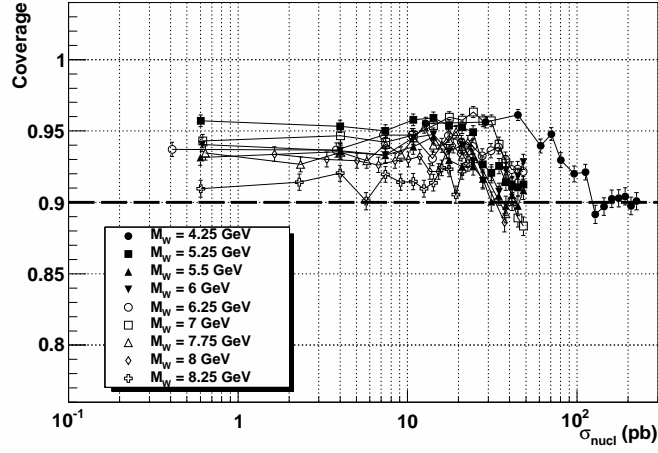
(a) 4.25→8.25 GeV
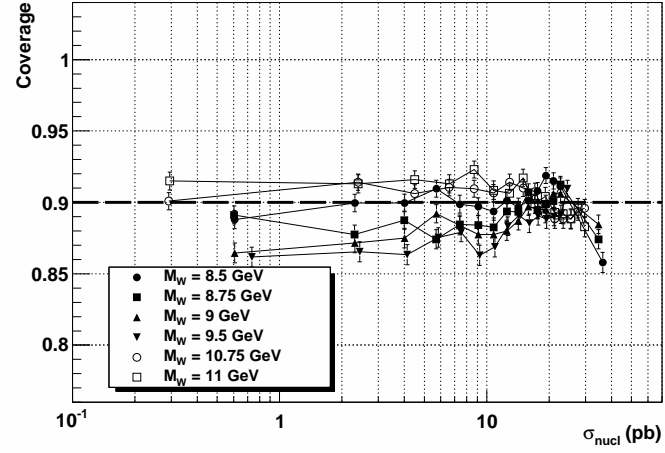
(b) 8.5→11 GeV
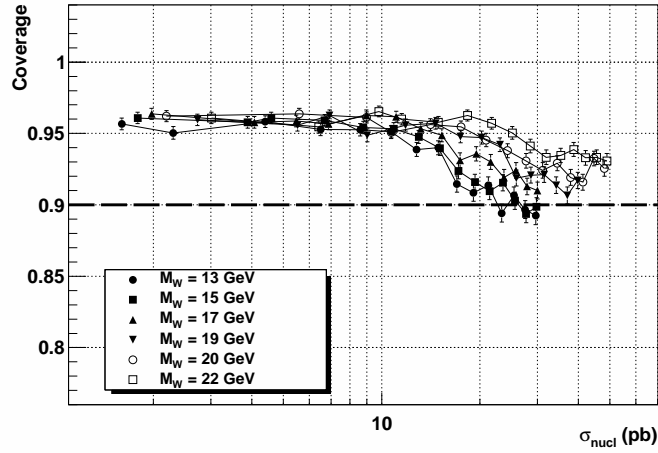
(c) 13→22 GeV
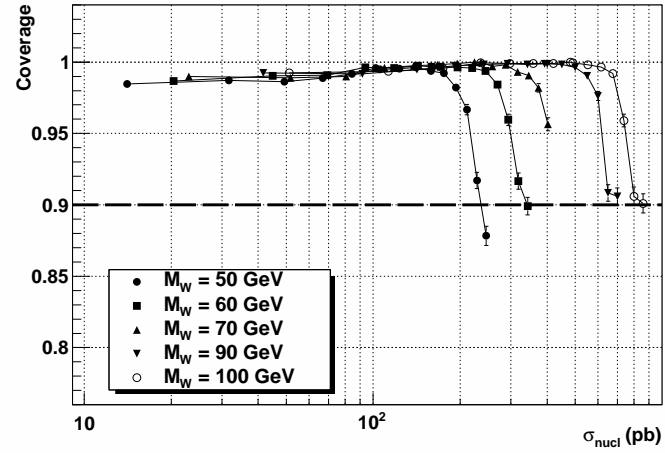
(d) 50→100 GeV

Figure 4.10: Coverage test results for WIMP mass range 4.25→100 GeV, see text for details. The range of the x axes (σ_{nucl}) differs since the coverage scans only cover values of the profile likelihood that satisfy $\lambda(\sigma_{nucl}) \leq 2$.

is reasonable to use $\lambda(\sigma_{nucl})$ to define the considered parameter space. $\lambda(\sigma_{nucl})$ provides a 1-dimensional path parameterized by σ_{nucl} along an extremum in likelihood space, basically allowing the data to define which parameters of the model are the most likely. It also reduces the scanned space to one dimension[3] thereby reducing the numerical problem to something tractable. To perform these tests, a model and a set of data were chosen to be the same as used throughout this section: the model used constrained the relative amplitudes of the Ge and Zn L-capture lines (see Section 4.4.5) and the data used was that with a 99% rise-time cut applied. Each fit during this procedure was unbinned. For each WIMP mass, a profile-likelihood function was calculated for σ_{nucl} by calculating the maximum likelihood along a grid in σ_{nucl}-space. At each point on this grid, the results of the fit (i.e. parameter values maximizing L) were saved for later use. Points were then selected from the $\lambda(\sigma_{nucl})$, first taking the best-fit result ($\lambda(\sigma_{nucl}) = 0$) and then sampling the remainder of the space $\sigma_{nucl} > 0$ and $\lambda(\sigma_{nucl}) \leq 2$ for 14 more points. For each of these 15 points, 2400 toy simulations were run[4] generating events according to the distributions defined by the parameters. Since the models used an extended-likelihood formalism, the number of events generated for each simulation was Poisson distributed according to model parameters. For each simulation, limits were calculated at 90% confidence level using the Rolke technique. The $\sim$ 1M simulations took roughly 1 year of 2.5 GHz CPU clock time, running on the Athena cluster at the University of Washington [80].

Results of these simulations are shown in Figure 4.10, where the coverage is defined as the percentage of time the calculated limits included the 'true' parameter value of σ_{nucl}, the value used to simulate the data set. The results from the range of WIMP masses are split into four different plots and grouped together according to similar mass. Each plot includes a line at 0.9, designating the expected coverage percentage as defined by the input confidence level. These results indicate in general good coverage with under- and over-coverage mostly limited to within 5% of the expected value. The significant over-coverage for larger WIMP mass suggests that over the lower range values of σ_{nucl}, the signal is

---

[3]In practice, the dimension is still 2 taking into account the variation of the WIMP mass.

[4]For WIMP masses 5.5, 11, 70, and 100 GeV only 2000 toy simulations were run. This was solely due to scheduling issues on the computer cluster used.

completely indistinguishable from background. The coverage does tend back to 90% at larger values of σ_{nucl}. The under-coverage in Figure 4.10(b) is likely due to the similarity of exponential background and WIMP signal and could also be affected by the profile-likelihood features discussed in Section 4.4.2. It is encouraging, however, that despite the fit difficulties and likelihood features discussed in Section 4.4.2, the coverage remains close to 90%. This suggests that the Rolke method as applied to this model and data set is robust and that exclusions determined at 90% confidence level are valid.

4.4.4   Limits from Unbinned and Binned ML Fits.

Limits were calculated using data described in Chapter 3 and the model in Section 4.4.1. The RooFit framework includes an abstracted interface for all data sets, making a transition between using a binned and unbinned analysis very simple and enabling a comparison between the two types of results. Therefore, both binned and unbinned limits were calculated, selecting a bin size of 23.4 eV over the energy range (0.5→3.5 keV, bin number: 128). Fits were performed for chosen values of M_W in the range $3.75 \rightarrow 100$ GeV, sampling different ranges of the mass with a variable step size, Δ: $3.75 \rightarrow 11$, $\Delta = 0.25$ GeV; $11 \rightarrow 25$, $\Delta = 1$ GeV, $30 \rightarrow 100$, $\Delta = 10$ GeV. Given the large amount of fits required to generate, the calculation method was parallelized for submission to the Athena cluster [80] at the University of Washington.

Fit Components versus σ_{nucl}

When running a large number of fits, it is challenging to provide effective quality control to ensure that the limit calculation has proceeded correctly. To check this, the parameters of the values were tracked at their best-fit value (defined as the minimum $-\log L$ in the region $\sigma_{nucl} \geq 0$) and at the 90% exclusion limit of σ_{nucl}. Examples of two of these plots are given in Figures 4.11 and 4.12 which are for unbinned and binned results, respectively. The top three plots in each of these figures includes the individual components of the background events: the counts-per-kilogram-per-day in the exponential, flat, and separate L-line backgrounds. The bottom plot includes the sum of the flat and exponential background and the sum

of the Ge and Zn L-capture lines. The parameters are generally smoothly varying versus σ_{nucl} except in the region $\sim 6 \rightarrow 10$ GeV where a significant amount of oscillation can be seen in the amplitudes of the flat and exponential components. This oscillation is due mainly to the fact that the exponential shape constant is allowed to float positive and therefore can become indistinguishable from the flat background. This is also confirmed by the largely-smooth behavior of the sum of the exponential and flat background components. An example of the variation of the exponential shape parameter is discussed in Section 4.4.2 and shown in Figure 4.8. However, some large variations do remain in the best-fit values of the sum components; these are due to the appearance in this WIMP mass region of local extrema away from the global minimum (see Figure 4.5(b)). Since the global extremum may be in the disallowed region (i.e. $\sigma_{nucl} < 0$), it is possible that the local minima are then interpreted as the global minima (i.e. the best fit), but this is dependent on the granularity of the scanned σ_{nucl} space and the size of the local minimum since it is possible to not sample fully the local minimum if it has a small width. In other words, it is possible to 'skip over' the local minimum and therefore interpret $\sigma_{nucl} = 0$ as the best fit. However, despite the sharp features in the best-fit values of the exponential and flat components it is clear that the value of the sum of these components at the 90% exclusion of σ_{nucl} varies smoothly.

The amplitudes of each L-line component vary quite little versus σ_{nucl}. Additionally, there is little to no difference between the amplitudes of these components at the global minimum and at the 90% exclusion of σ_{nucl}. This indicates that this portion of the background has little effect on the calculation of limits on a WIMP signal.

All of the same conclusions may be drawn from the binned analysis of the same data, results of which are shown in Figure 4.12. As with the unbinned data, it is clear that some sharp features exist in the best fit of the sum of the flat and exponential components. Regardless, the smoothness of the sum of the parameters at the 90% exclusion of σ_{nucl} implies that any sharp variation in the best-fit parameters are irrelevant. The similarity between binned and unbinned results indicates the robustness of the maximum-likelihood technique and also suggests that an unbinned analysis could be unnecessary for data with this number of counts.
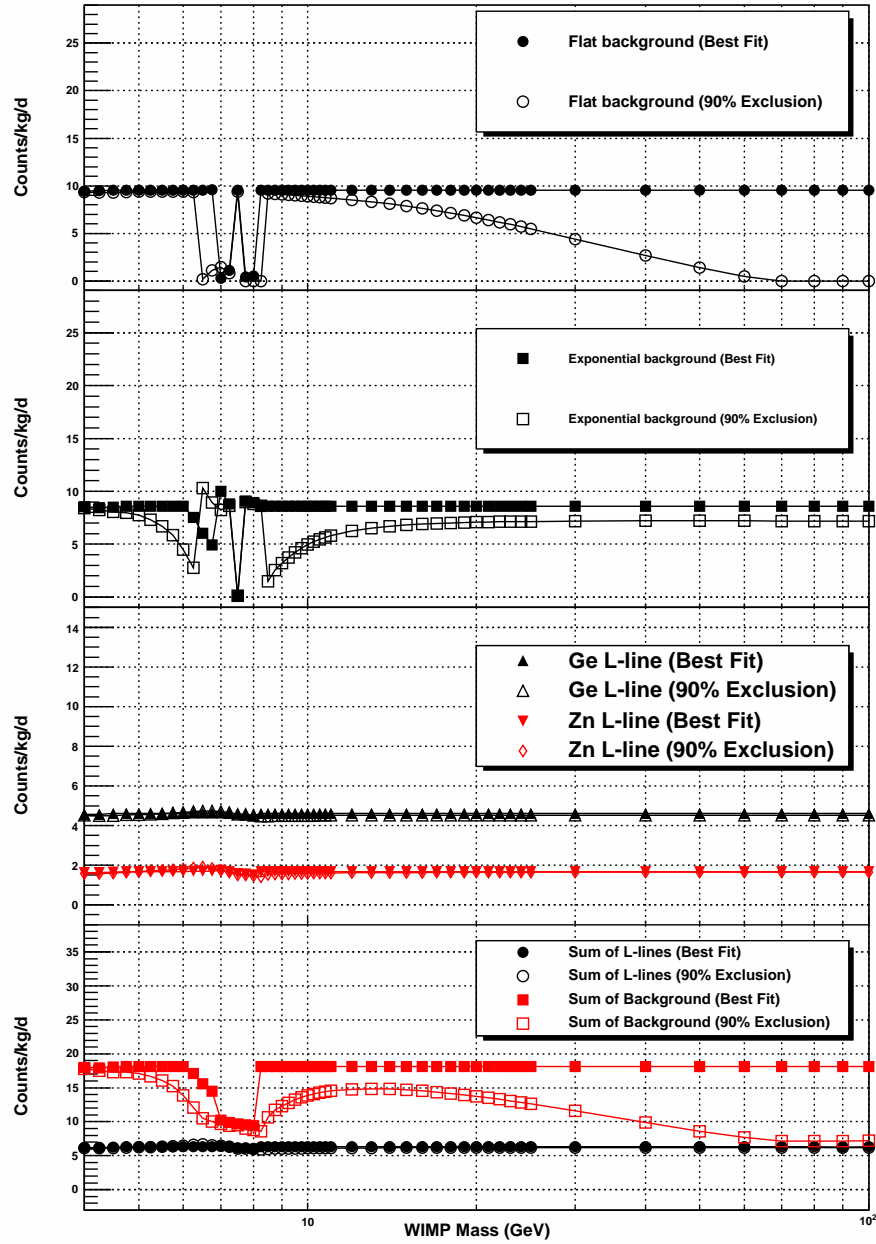
Figure 4.11: Results from an unbinned fit using data with 95% rise-time cut (+ microphonics cut) applied. The top three figures contain the variation of all independent parameters at their best-fit value and at the 90% exclusion limit of σ_{nucl}. The bottom figure contains a sum of the background components (flat + exponential) and the L-lines. See text for details. Lines between points are included to guide the eye.
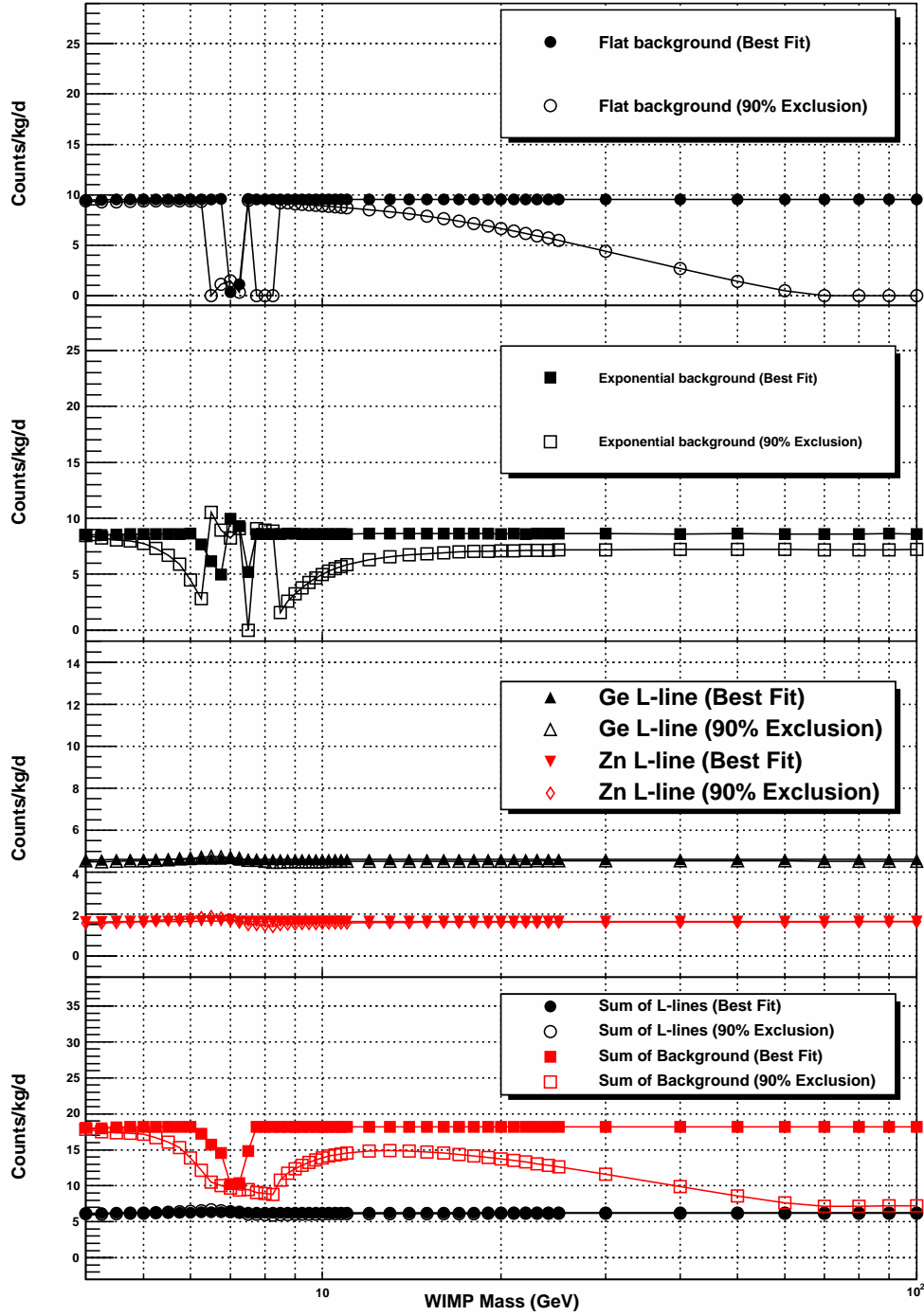
Figure 4.12: As Figure 4.11 but with binned data.
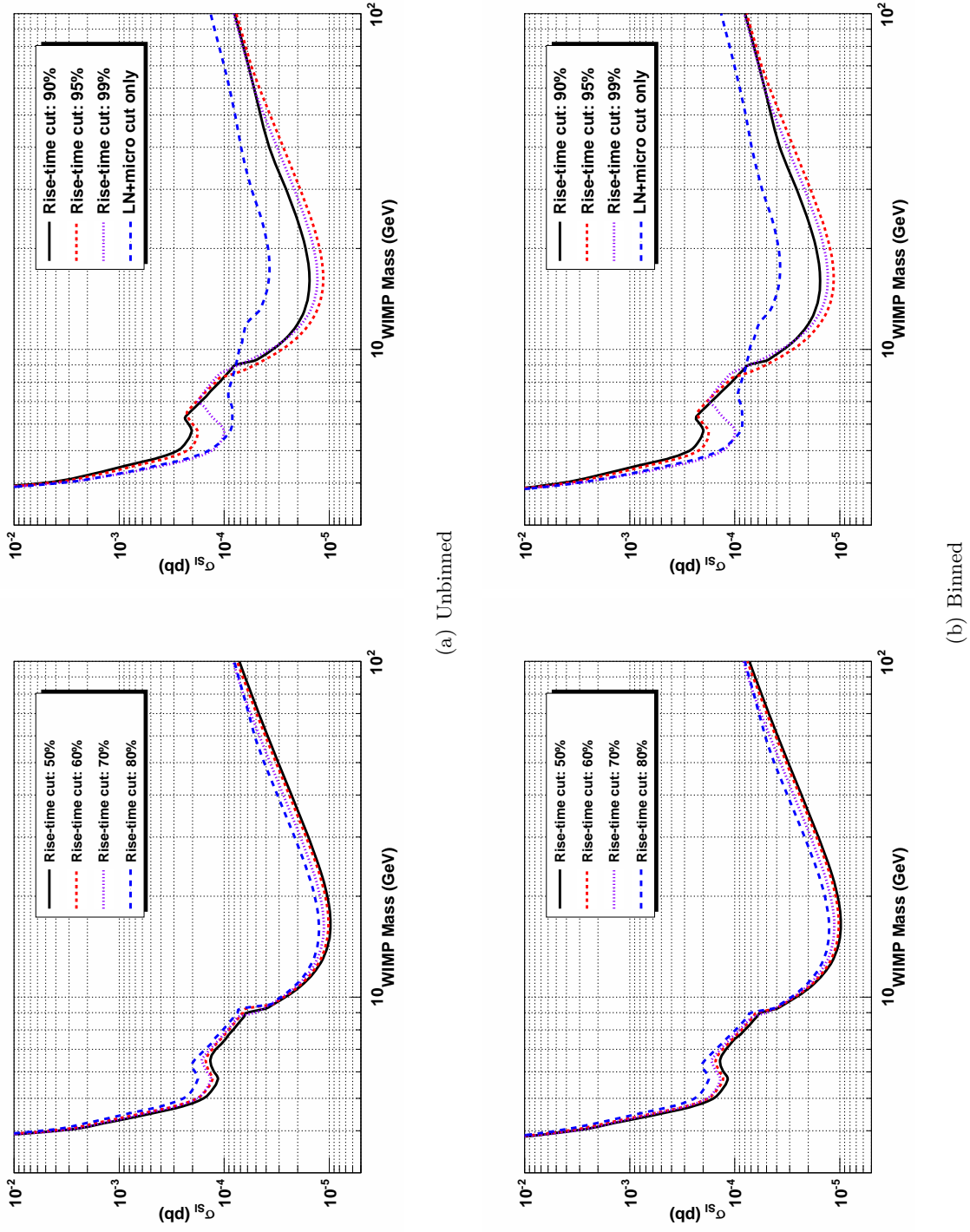
(a) Unbinned

(b) Binned

Figure 4.13: 90% CL limits on σ_{W-n} for various data sets.

144

Exclusion Limits

Exclusion limits at 90% confidence level for σ_{W-n} are shown in Figure 4.13, with σ_{W-n} calculated from σ_{nucl} given the relationship in Equation 4.4. Results for both binned and unbinned data are split into two plots each to allow clearer visualization of exclusions calculated for data sets with different cuts applied. It is clear that there exists very little distinction between limits on binned and unbinned data; however, several features are apparent among the different data sets. In the low-WIMP-mass range, 4-5.5 GeV, all the curves are very similar suggesting the limits in this small region are robust against cuts to the data. At higher WIMP mass, 20-100 GeV, the slope of the curves is very similar, though the normalizations are somewhat different. As expected, in this region the microphonics-cut data exhibits more conservative limits than the rise-time cut-data. The variation in the limits calculated from rise-time-cut data is expected to arise from systematic errors on the estimation of the efficiencies of these cuts as well as from the unknown ratio, (signal)/(signal + background), for each cut.

The mass range 6-9 GeV exhibits sharp features in the exclusion plots, existing in WIMP-mass regions where the shape of the signal is very similar assumed to the shape of the background. These features also appear in the microphonics-cut data, but at a slightly higher mass range from $\sim 7 \to 10.5$ GeV. These characteristics are essentially an indication that the calculated exclusion limit comes when the amplitude of the WIMP signal completely takes over the contribution from the exponential fit component. An example of this is given in Figure 4.14 which displays the fit at the 90% CL exclusion of σ_{W-n}. It is clear that at this value of σ_{W-n}, the WIMP signal fits to the low-energy data without any contribution from the exponential background. Close inspection also reveals that the exponential shape parameter has transitioned to a positive value, resulting in a mildly increasing background with energy.

At low-mass range the microphonics-cut data demonstrates a stronger limit than the 80%→99% rise-time-cut data. Initially this seems counter-intuitive, but can be explained due to the fact that the microphonics-cut data includes a larger contamination of slow-rise-time pulses which is not well fit with a single exponential. This indicates that the
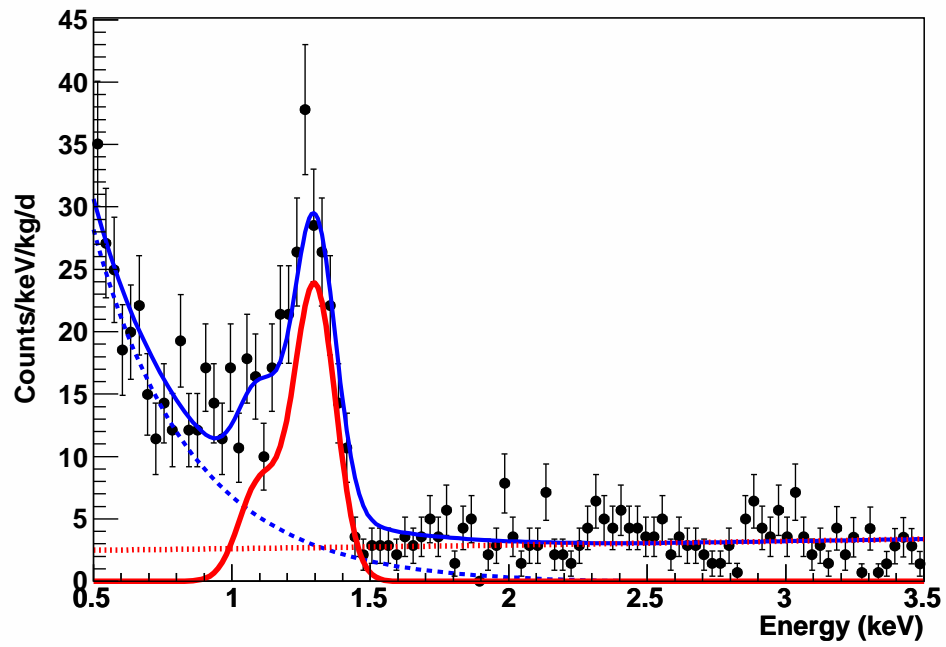
Figure 4.14: Fit at 90% CL exclusion, $\sigma_{W-n} = 1.2 \times 10^{-4}$ pb. WIMP signal is blue dashed, L-lines in solid red, flat + exponential background in dotted red, and sum of all the pdfs is blue solid. This is an unbinned fit with 95% rise-time cut data.

background distribution coming from slow rise-time pulses is not properly included in the set of background pdfs. Since this distribution is not known a-priori, the proper inclusion would require a source measurement to estimate its shape. It is interesting to note that the exclusion calculated using 99% rise-time cut data demonstrates a transition between the microphonics cut data and the 95% rise-time cut data around 6 GeV. The shape and coverage of the rest of the rise-time cut exclusions are similar, suggesting that the rise-time cut retains a consistent shape in the data as the acceptance is decreased from 99%. It also implies that some slow-rise-time-pulse background remains at 99% acceptance, consistent with results from Section 3.4.3 which demonstrated a slightly less-negative exponential constant with the 99% rise-time-cut data as compared to rise-time cuts with less acceptance. This final conclusion indicates that data with at least a 95% rise-time cut should be used to generate exclusions since these data will have as little slow-rise-time-pulse contamination as possible.

4.4.5 Constrained Ge and Zn Relative Amplitudes

The relative amplitudes of the Zn and Ge L-capture lines may be constrained by the K-capture lines of the same isotopes since the K- to L-capture ratio has been well understood through independent measurements [45, 81], see Table 4.3. The amplitudes (in counts) of the Ge and Zn K-lines were measured using an unbinned fit of the 99% rise-time-cut data[5] and determined to be 945.1±28.2 and 349.7±19.1, respectively[6]. Therefore, the expected ratio of the Ge and Zn lines (Zn/Ge) was 0.33±0.02. The relative amplitudes of these lines in the fitting model were then constrained by this value, reducing the number of parameters in the fit by one. In general, it would be more desirable to tie this ratio to the amplitudes of the K-capture lines and perform a simultaneous fit, but since the relative amplitudes of the lines did not vary significantly when allowed to float independently and since the measurements of each set of lines were in separate channels (high-, low-gain channels for L-,K-lines, respectively), it was not expected that proceeding in such a manner should have

---

[5]The 2 month live-time data was used to calculate this ratio.

[6]This measurement was performed using a smaller subset of data with 2 months of live-time.

Table 4.3: L/K capture ratios for ^{68}Ge and ^{65}Zn. The theory of Brysk and Rose [82] has corrections applied via Bahcall [67]. Useful tabulations for calculating L/K values are found in [83].

Atom	Value	Ref.
^{68}Ge	0.1328±0.002	[45]
^{65}Zn	0.119±0.007	[81]
^{68}Ge (theory, corrected)	0.126	[82, 67]
^{65}Zn (theory, corrected)	0.108	[82, 67]

a significant impact. Final results of the fit supported this initial assumption.

Results are presented as in Section 4.4.4, with a representative set of plots of parameters versus WIMP mass (Figures 4.15, 4.16) and exclusion limits in Figure 4.17. As before, it was determined that very few differences manifested between the binned and unbinned results. The amplitude of the sum of the L-lines appears very similar (~ 6 Counts/kg/d) as with the unconstrained fit (Figure 4.11). However, this comes as little surprise given the relative independence and lack of variation of these parameters in the unconstrained fits. Because of the clear similarity between the constrained and unconstrained results, the conclusions regarding this set are equivalent to the previous. However, the results of this section further emphasize that the amplitudes of the L-lines have little effect on the exclusion of a low-mass-WIMP signal since the sum of the L-lines varies so minimally with WIMP mass. Also, little difference exists between the amplitude of the L-line at best fit and at 90% exclusion, indicating that these parameters have very little sensitivity to σ_{nucl}.

4.5 Conclusions and Discussion

The data from the modified-BEGe detector deployed underground at Soudan Underground Laboratory have been used to generate limits on spin-independent WIMP interaction cross-section. A framework has been developed that should prove useful in generating limits using other R&D detectors and the MAJORANA DEMONSTRATOR. The results of the data
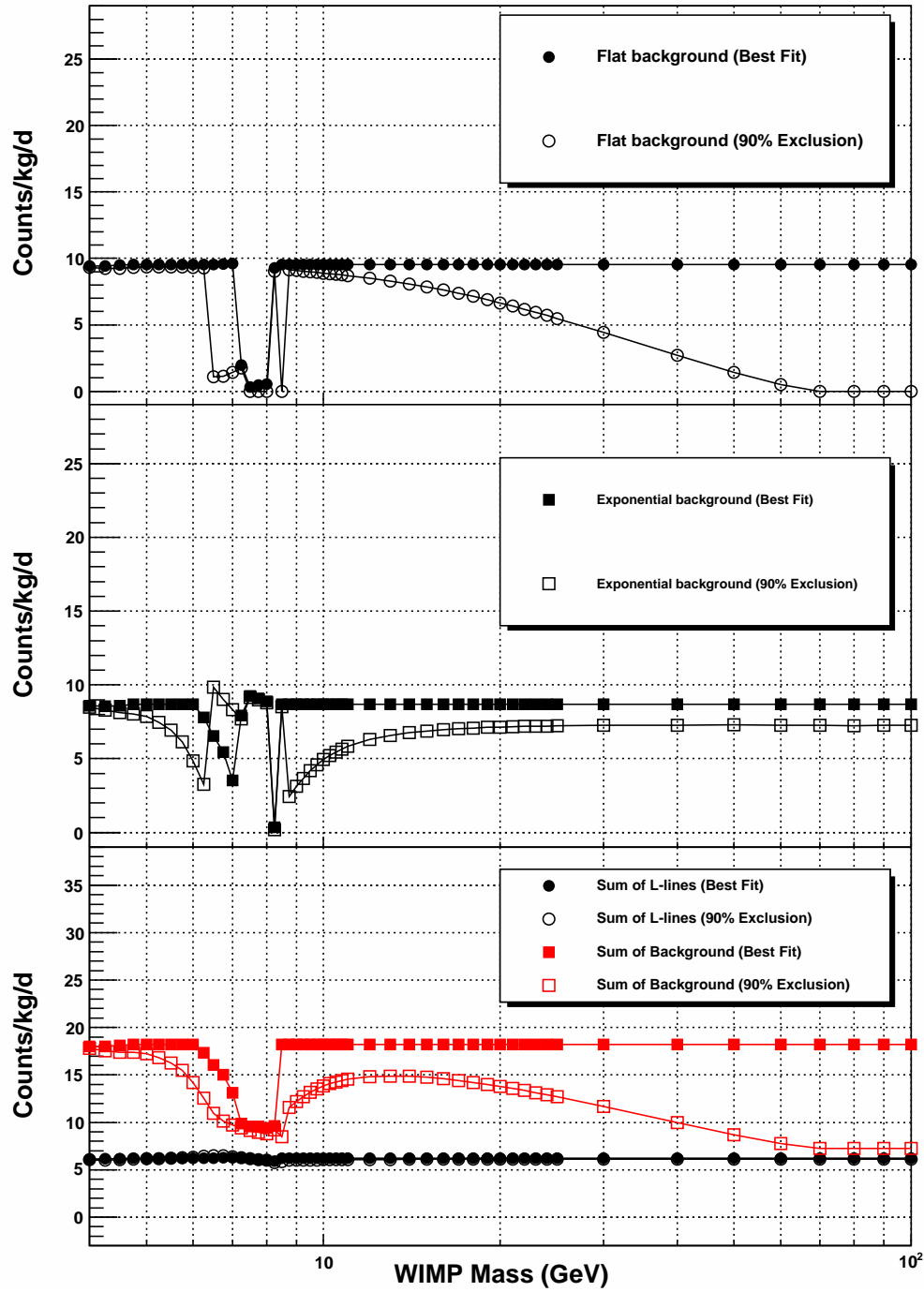
Figure 4.15: As Figure 4.11, but with a constraint on the relative amplitudes of the Ge and Zn lines. Therefore, only the sum of the L-lines (and not each contribution) is included.
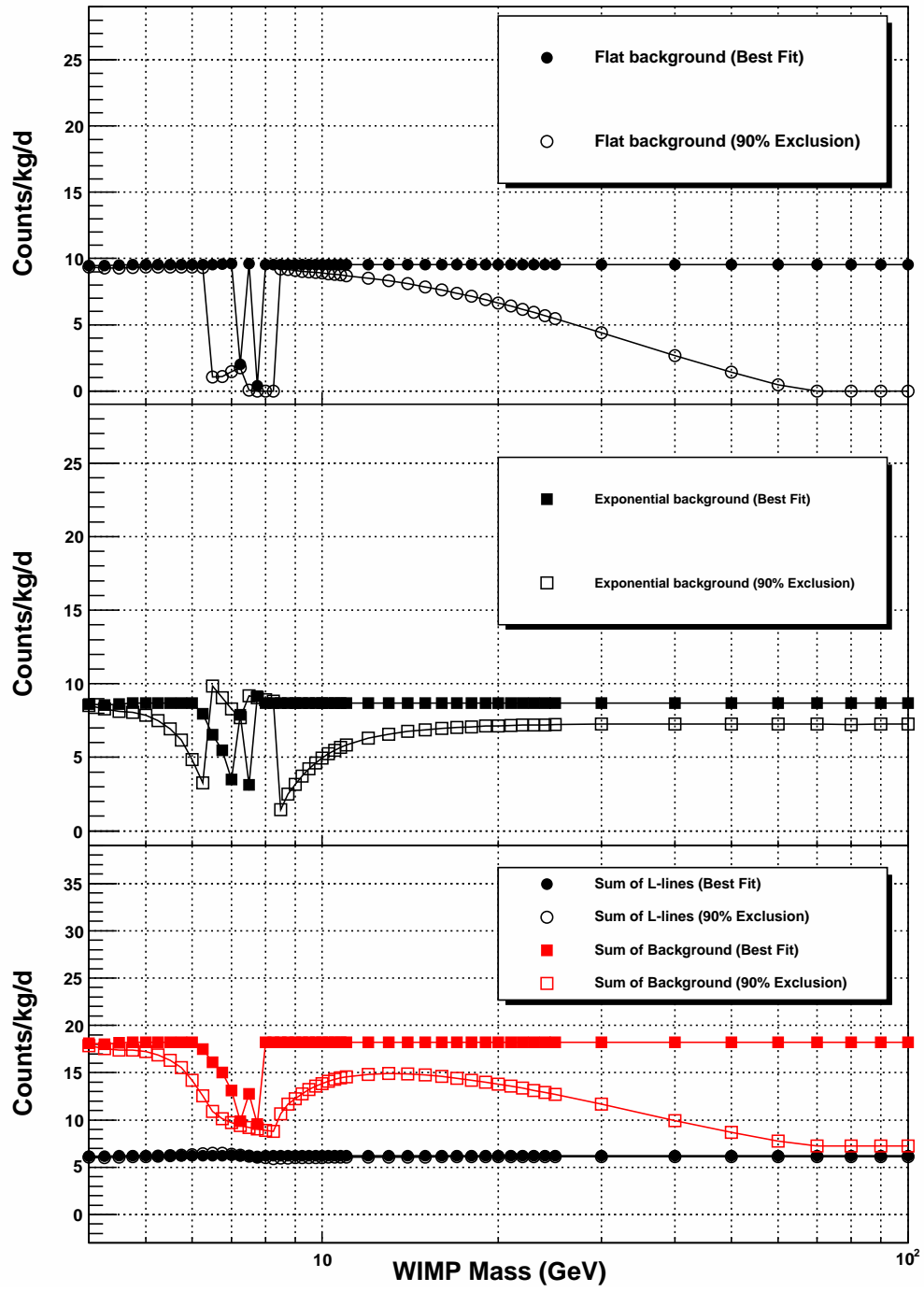
Figure 4.16: As Figure 4.15, but with binned data.

(a) Unbinned

(b) Binned
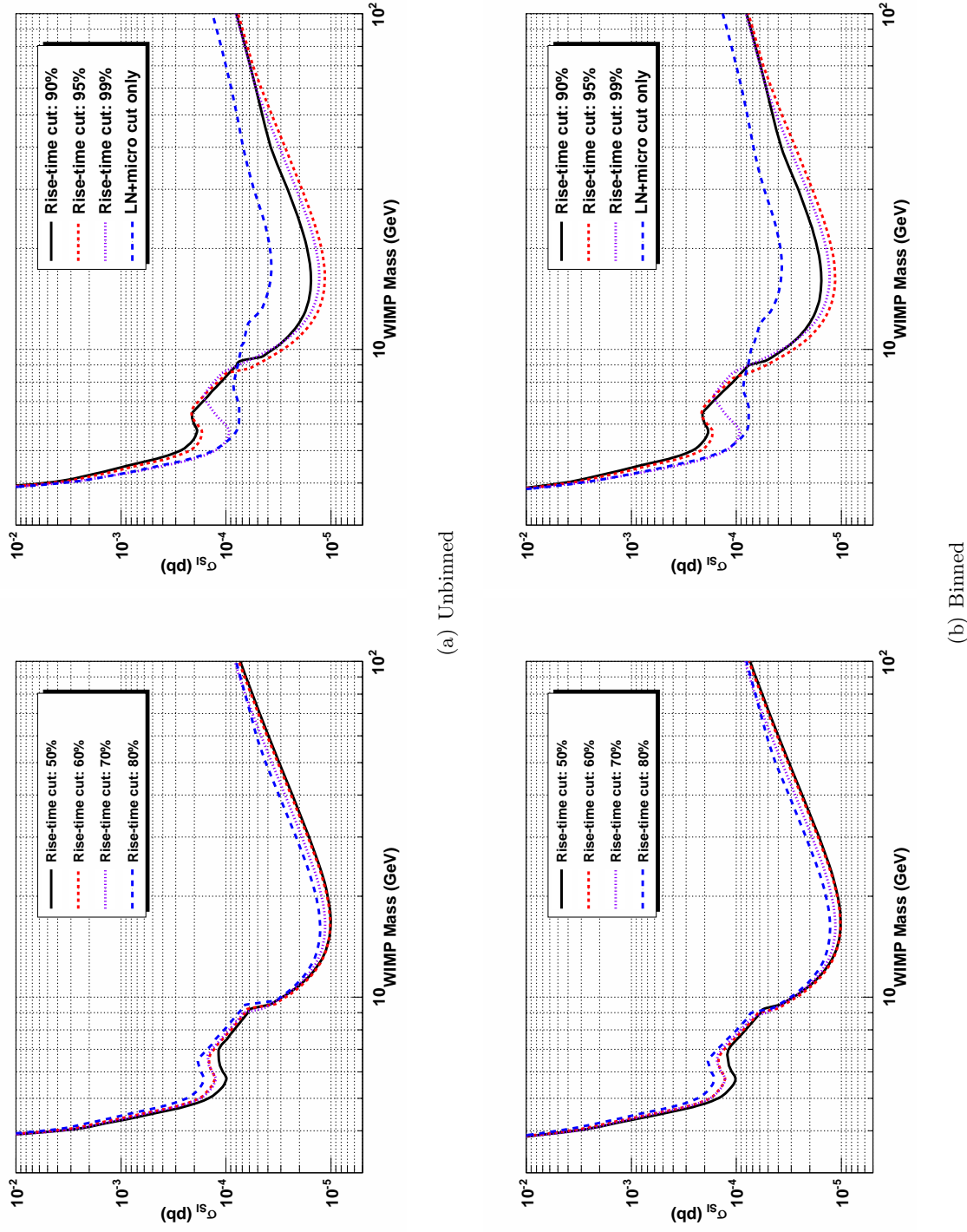
Figure 4.17: Limits on σ_{W-n} constraining the relative amplitudes of Ge and Zn L-lines.
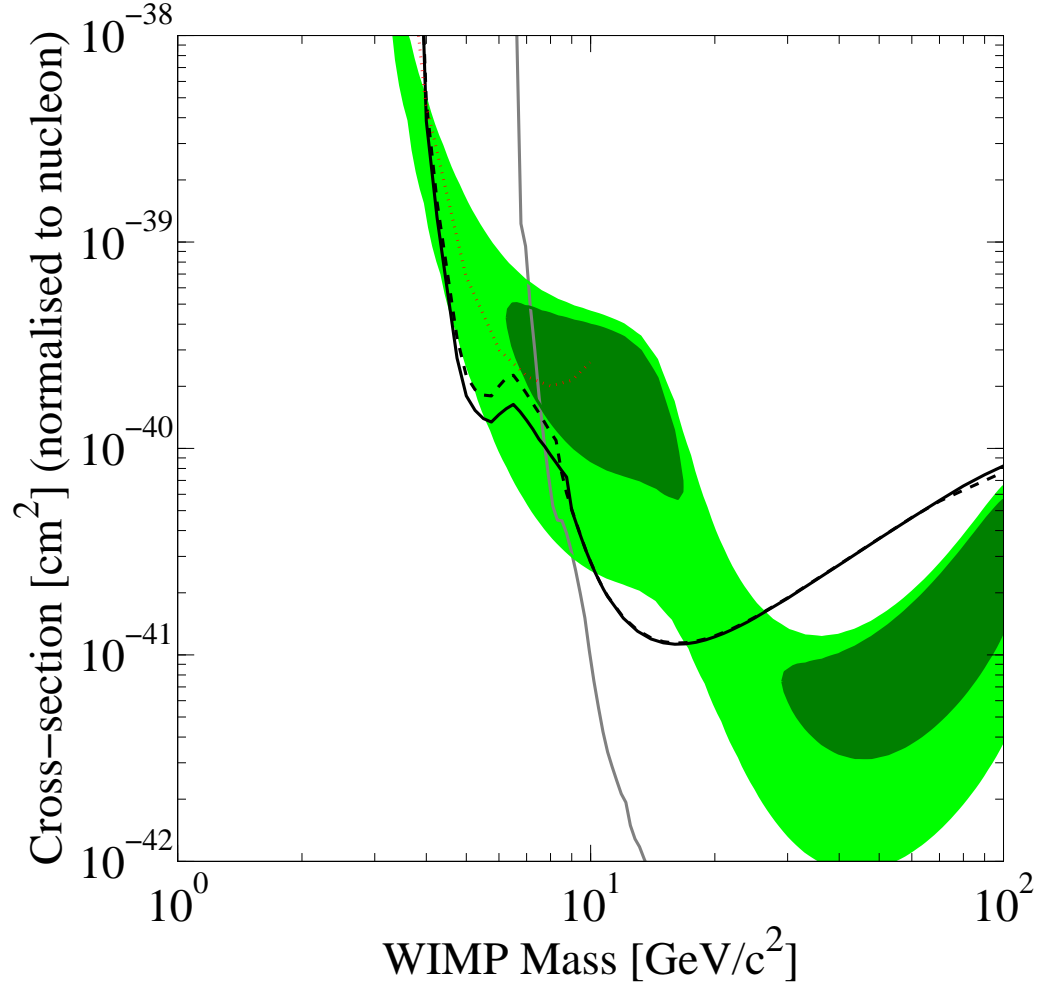
Figure 4.18: Comparison of results from this analysis to other results. Included lines are 90% CL exclusion limits using 95% (black dashed) and 70% (black solid) rise-time-cut data sets, results from the previous CoGeNT detector [11] (red dotted), most recent CDMS results [84] (gray solid), and acceptance regions from DAMA data as interpreted in [85] (3σ dark green region, 5σ light green region). Plot generated with dmtools.brown.edu [86].

from this detector may be compared to other experimental results as well. For example, in Figure 4.18, limits from these results are compared to previous results with a similar detector from CoGeNT [11], the most recent published data from the CDMS collaboration [84], and regions of acceptance for the DAMA data as interpreted in [85].

The DAMA experiment is a NaI-based experiment that has observed an annual modulation in their data. Interpreting this feature in their data as interactions of low-mass WIMPs yields the acceptance regions presented in Figure 4.18. The germanium-based CDMS experiment has the ability to distinguish between electron and nuclear recoils, enabling it to achieve far lower bounds on σ_{W-n}. However, the threshold of the CDMS detectors limits their sensitivities to WIMP masses above ~ 7 GeV. It is clear that results from this analysis fully exclude the DAMA 3σ acceptance region at low WIMP mass and remove almost all of the allowed space of the 5σ region save some space at mass below $M_W = 4$ GeV and some mass regions from $6 \rightarrow 10$ GeV. Accessing the region at low mass will require a further reduction of threshold below 0.3 keV. The exclusion of the still-accessible 5σ region $6 \rightarrow 10$ GeV will demand a clearer understanding of the unknown exponential background since the similarity between this background and a WIMP signal forces the limits high in this region. Rejecting this background as a possible WIMP signal can also come from looking at the time behavior of the data to see if it exhibits any annual oscillation characteristic of moving through the WIMP halo. The MAJORANA DEMONSTRATOR will have an enhanced sensitivity to a rate modulation given its larger mass and exposure time. The physics reach of the MAJORANA DEMONSTRATOR and additional analyses related to other dark matter candidates will be explored in the following chapter.

Chapter 5

OTHER LOW-ENERGY PHYSICS WITH P-TYPE POINT-CONTACT DETECTORS

The high resolution and low threshold of p-type point-contact detectors makes them sensitive to other dark matter signals in addition to those from WIMPs. This chapter explores this, deriving limits on the inelastic scattering of pseudoscalar dark matter candidates off electrons. The MAJORANA DEMONSTRATOR will deploy an array of P-PC detectors with a significantly lower background ($\gtrsim 10^3$ reduction in count rate) than seen in the data set presented in this thesis. The latter half of the chapter estimates the sensitivity the MAJORANA experiment to dark matter candidates making some conservative assumptions on the shape and magnitude of the expected background.

5.1 Other Dark Matter Candidates: keV-scale Bosons

The lack of understanding as to how a dark matter particle couples with normal 'Standard Model' matter has motivated wide theoretical investigation resulting in candidate particles in addition to WIMPs. As well, results from the NaI-based DAMA experiment [87] have demonstrated a clear annual modulation signal which could be interpreted as arising from the earth's movement through a galactic cloud of dark-matter particles. Because other experiments sensitive to WIMP interactions have failed to reproduce this result, the possibility remains that some non-WIMP-like process could underlie the cause. However, the results of the DAMA experiment are not the sole motivation for studying dark matter candidates beyond WIMPs. Whereas most dark matter experiments focus their sensitivity on looking for one type of signal (i.e. a nuclear recoil from a WIMP interaction), it is still essential to consider other possibilities.

Pospelov et al. have completed a study [28] analyzing the possibility of keV-mass bosons as dark matter candidates. In particular, they outline the expected interactions and rates for scalar, pseudoscalar, and vector bosons incident upon modern detectors. In general,

Pospelov et al. avoid in-depth discussions regarding theoretical motivation of each bosonic type, but the mathematics behind the pseudoscalar are the same as that for an axion, the particle responsible for preserving CP in QCD (see, e.g. [15]). Because of this equivalence, other work deriving axion processes or limits on such processes remains relevant. The following sections consider the coupling of a pseudoscalar to electrons in a detector and therefore refer to this process as an 'axioelectric' effect and the originating particle as an 'axion.'

5.1.1 Axioelectric Signal

The signal for a non-relativistic axion interacting via the axioelectric effect has been derived in [28]. This particular inelastic interaction involves the deposition of the *complete* energy of the axion, which in the non-relativistic case is essentially equal to the mass of the particle. Since the excitation of the electron via the axioelectric effect is similar to the process mediated by a photon in the photoelectric effect, the signal is a delta function centered at the mass of the axion, m_a. Convolved with the detector resolution, the signal would appear Gaussian with width exactly that of a gamma or x-ray of energy equivalent to m_a. The rate of this interaction has been estimated in [28] as:

$$R\left[\text{kg}^{-1}\text{day}^{-1}\right] \simeq \frac{1.2 \times 10^{19}}{A}g_{a\bar{e}e}^2 m_a \sigma_{photo} \tag{5.1}$$

where A is the atomic mass, m_a is the mass of the axion in keV, σ_{photo} is the measured photoelectric cross section in barns, and $g_{a\bar{e}e}$ is the dimensionless coupling constant related to the axion decay constant f_a (see Section 1.5.2) by $g_{a\bar{e}e} \equiv 2m_e/f_a$ (m_e is the mass of the electron). In the derivation of this result, the value of the density of dark matter was used: $\rho_D = 0.3$ GeV cm^{-3}. Also, Pospelov et al. noted that this rate should not exhibit a significant annual modulation from the earth's orbital velocity since the interaction does not include strong velocity dependence. Therefore no time dependence was included in the equation. The rate calculated for a germanium detector ($A = 72.96$) for an assumed coupling constant value, $g_{a\bar{e}e} = 10^{-11}$, is shown in Figure 5.1. The rate calculation used well-measured photoelectric cross sections obtained from the NIST FFAST database located
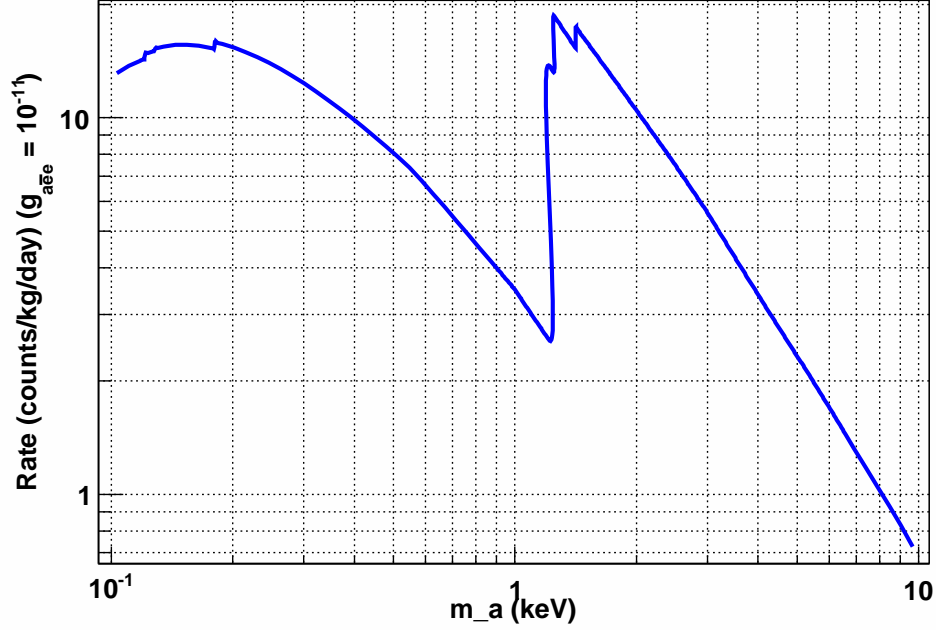
Figure 5.1: Non-relativistic axion axioelectric interaction rate in germanium. The photo-electric cross section for germanium was obtained from the NIST database [88].

online [88]. The sharp feature in the plot around $m_a \sim 1.3$ keV arises from the L-line edge at this energy. Other sharp features similarly relate to the energy levels of electrons in a germanium atom, but the K-line edge is notably absent due to the limited range of the plot (<10 keV).

5.1.2 Limits on the Axioelectric Effect

Limits were calculated using the profile-likelihood method described in Section 4.3.1. Fits were performed on the complete number of data sets described in Chapter 3, with a total live-time of 150.6 days. All data sets yielded similar results and so one data set was chosen with 95% rise-time acceptance cuts and microphonics cuts applied. As outlined before, assumptions about the source of slow-rise-time events reduced the fiducial mass to 0.33 kg (see Section 3.4.5). Additionally, results with unbinned and binned maximum-likelihood

fits were consistent and so the former was used for the final result. The limit calculation followed the same procedure as a peak search in the data and is outlined as follows:

- Define the Gaussian signal f_{axion}:

    - Choose mass m_a of the axion defining μ of Gaussian

    - Determine σ at $E = m_a$ using resolution in Equation 3.3.

- Fit to the function $B + N_{axion} f_{axion}$ where B is the background defined in Section 4.4.1 and determine the profile likelihood $\lambda(N_{axion})$.

- Determine the 90% upper limit on N_{axion} using $\lambda(N_{axion})$.

- Repeat for other values of m_a

During the fits, the μ and σ of the Gaussian signal, f_{axion}, were kept fixed and only the amplitude, N_{axion} was kept as a free parameter of the signal. For the background, the behavior of the parameters was the same as during WIMP exclusion fits (see Section 4.4.1) and the relative amplitude of the ^{68}Ge and ^{65}Zn L-lines was kept fixed as described in Section 4.4.5. Fixing this relative amplitude served to minimize the impact of the L-lines in the exclusion fits since a signal centered at 1.1 or 1.3 keV would look exactly like either L-capture line. The difficulties seen while determining limits on low-mass WIMPs did not appear in these calculations because the signal (Gaussian centered at m_a) was not similar to the background except for the case of the L-lines. The value of the axion mass was scanned from 0.1 keV to 7.8 keV in steps of 0.2 keV using both high- and low-gain channels: high-gain channel, $0.1 \rightarrow 2.9$ keV; low-gain channel, $3 \rightarrow 7.8$ keV. The axion mass was allowed to vary below threshold (0.5 keV) because the finite resolution of the detector would allow portions of the expected Gaussian signal to be detected above threshold. An example of an exclusion fit in the high-gain channel is shown in Figure 5.2 for $m_a = 3$ keV.

The 90% CL excluded rate in counts/kg/day, $R_{axion}(m_a)$, was determined from N_{axion} and from this value the upper limit of $g_{a\bar{e}e}$ could be determined. The exclusion calculated from this result is presented in Figure 5.3 along with a comparison to other results, including
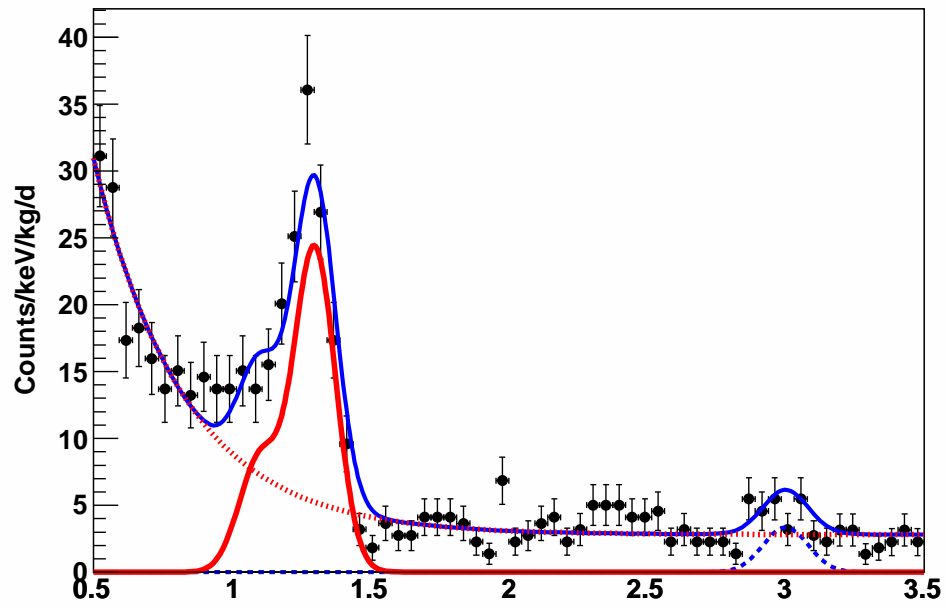
Figure 5.2: Example of an excluded non-relativistic axioelectric signal at $m_a = 3$ keV at 90% CL. In this fit performed in the high-gain channel, the excluded value N_{axion} is 36.1 counts. The components of the fits are split: red solid, L-Lines; red dotted, flat plus exponential background; blue dashed, excluded axioelectric signal.

previous results of the CoGeNT collaboration [11], the CDMS collaboration [89], and an acceptance region from the DAMA collaboration [87]. As noted in both references [90, 28], the limit calculation performed in [87] did not correctly treat the leading term in the Hamiltonian, producing instead a reduced rate around 3 orders of magnitude lower for a given $g_{a\bar{e}e}$. An estimation of the corrected result from DAMA as outlined in [90] appears in Figure 5.3. A comparison to limits derived from both solar neutrinos [23] and globular clusters [24] is included as well.

The strongest limits obtained from astronomical observation arise from determining how much 'hidden energy' may be carried off by axions from the solar core (in the case of solar neutrinos) and similarly from the cooling observed in the evolution of red-giant stars in globular clusters. All direct measurements provide stronger limits than those from solar neutrinos, but the mass-independent limits from globular-cluster stars still surpass all others. Supernovae generally provide the best constraints on axions and other exotic particles, but at their high temperatures ($O(10\,\mathrm{MeV})$) axion-electron interactions are additionally suppressed according to m_e^2/T^2 [28]. Other limits on pseudoscalars from cosmological observations are possible, including searches for decays to photons and estimates of axion abundance from the big bang. These limits are discussed with respect to the sensitivity of the MAJORANA DEMONSTRATOR in Section 5.2.5. Even though the presented direct-detection limits are already well within the space disallowed by astronomical constraints, it is still important to explore these parameter regions since limits from both cosmological observation and experiment can depend strongly on choices of models and their parameters.

Pospelov et al. noted the axioelectric process should be to first order invariant to the velocity of the incoming dark matter particle and so would not generate an appreciable annual modulation. This would limit any interpretation of the DAMA annual modulation as due to an axioelectric interaction and would therefore disallow the DAMA acceptance region in Figure 5.3. Collar et al. point out the possibility of recovering the annual modulation effect by considering the axioelectric interaction rate as not arising from changes in incident dark matter velocity but rather from the annual variation of the number density of the particles [90].
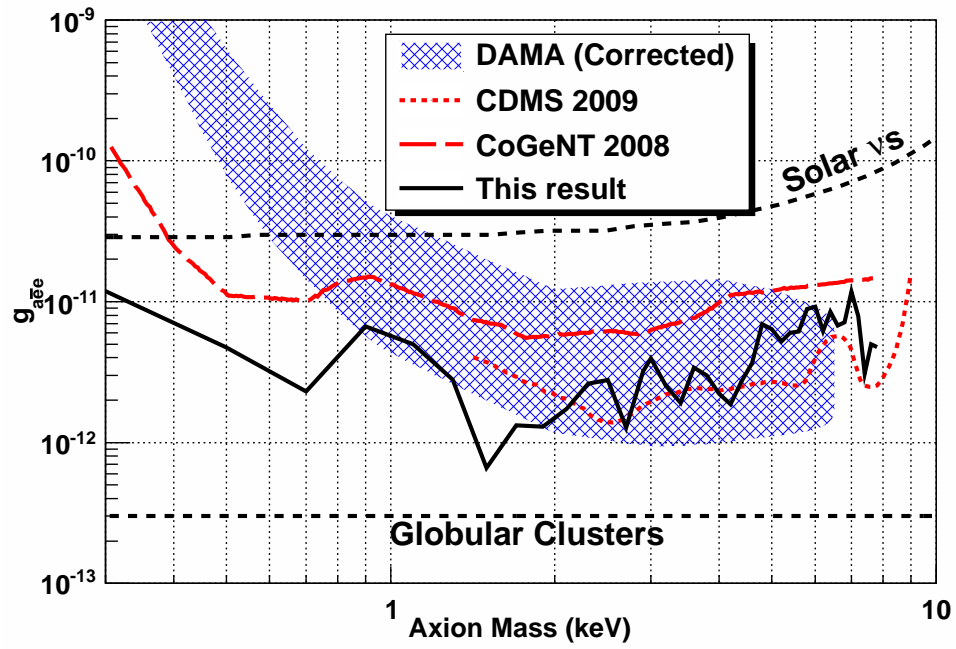
Figure 5.3: Limits on the axioelectric coupling constant $g_{a\bar{e}e}$. Results from this work appear in comparison to previous results from CoGeNT [11], CDMS [89], and DAMA [87]. The DAMA results have been corrected per reference [90]. Limits derived from both solar neutrinos [23] and globular clusters [24] are included as noted, see text for details.

5.1.3   Conclusions and Discussion

The results from this section underscore and emphasize the physics reach of P-PC detectors. At higher energies the modified-BEGe data yielded results comparable to those from CDMS, but was additionally able to generate exclusions at lower axion mass values due to the reduced threshold. Though limits derived from astronomical observations still exceed those from direct-detection experiments, it remains important to verify these limits and, if possible, better them with larger exposure times and smaller backgrounds. The need to explore dark matter candidates beyond those offered by WIMP theories remains evident and the only method to ensure experimental sensitivity to rare events with an *a priori* unknown signature is to reduce known backgrounds through appropriate shielding and detector component radio-purity. The MAJORANA experiment seeks to do this to search for $0\nu\beta\beta$, but the same efforts to reduce backgrounds in the double-beta decay signal region (~ 2 MeV) should benefit searches for dark matter in low-energy regions. The following sections discuss and calculate the sensitivity of the MAJORANA DEMONSTRATOR for such searches.

5.2   Sensitivity of the MAJORANA DEMONSTRATOR to Dark Matter Signals

5.2.1   Introduction

Since a framework was established to determine exclusion limits for low-mass WIMPs and for the axioelectric coupling constant, it is simple to apply this same framework to determine the sensitivities of the MAJORANA DEMONSTRATOR to these two dark matter signals. The ultra-clean composition of the DEMONSTRATOR should provide an excellent detector for searching for dark matter without relying upon significant background reduction cuts (e.g. discrimination between nuclear and electron recoils). Calculating the sensitivity of the experiment involves making some assumptions of the background at low energies and of the makeup of the experiment. A discussion about the estimation of the background follows in Section 5.2.2. The general prescription for calculating the sensitivity is outlined:

- Generate a background model, including an expected rate of background

- Simulate a spectrum according to the background model

- Fit to the simulated spectrum the background model plus a signal model (e.g. WIMP or axioelectric spectrum)

- Calculate the upper limit on the amplitude of the signal at 90% CL.

- Repeat a large number of times ($O(1000)$) to generate an ensemble of limits.

The generated ensemble of limits for a particular signal would create a distribution of limits. The 90% CL limit was chosen as the limit which was above 99% of the entries in this distribution. Details of the MAJORANA DEMONSTRATOR are given in Section 1.2, but for these purposes we have conservatively assumed that the MAJORANA DEMONSTRATOR will be composed of 20 kg of material and that it will accumulate between 1 and 5 years of live-time.

5.2.2 Low-energy Background Model

The estimation of background generally comes from verified simulations and from extrapolations from previous experiments. In this work, the background is estimated by assuming it arises from two main sources: (1) a continuum from higher energy processes, and (2) counts from the beta decay of cosmogenically-produced tritium in the detector. A simulation to estimate (1) is wrought with challenges since a large number of contributions can affect the result. However, it is possible to use previous results from low-background germanium-based experiments to produce an estimate of this background. The IGEX experiment measured a flat background rate of ~ 0.1 counts/keV/kg/day in this low energy region ($4 \rightarrow 10$ keV) [91]. The MAJORANA DEMONSTRATOR plans to reduce background above 200 keV by a factor of 100 [92] and so it is reasonable to expect the flat background at low energies will follow this same reduction to be roughly 0.001 counts/keV/kg/day. This background was assumed stable in time as well, so that it was flat in both time and energy. Estimating the flat background amplitude is not as critical assuming that the low-energy background is dominated by tritium.

The estimate of a background to tritium involves understanding the activation rate of tritium for germanium at the surface of the earth. This activation rate has been estimated to

be $\lesssim 200$ ^{3}H-atoms-per-kilogram-of-Ge-per-day for natural germanium at the surface of the earth, and almost a factor of 2 less for germanium with enriched ^{76}Ge content [42]. Other references have suggested that this rate is roughly an order of magnitude too high [93, 44]; we conservatively use the enhanced rate, but assuming a smaller activation would correspondingly lengthen an allowed exposure time. The exposure is determined by the time of manufacture beginning with the pulling of the germanium crystal and ending when the crystal is brought underground. For example, if the entire process from crystal pulling to detector development and then final deployment (or storage) underground takes 15 days, then this integrated time is the total tritium activation period for the detector. For a kilogram natural-germanium detector created over such a time scale, we would expect there to be $15 \times 200 \sim 3000$ atoms of ^{3}H within the detector. Given the slow time of decay of ^{3}H (12.32 year half-life), it is critical to minimize the time above ground and certainly necessary to avoid any time at high altitudes, e.g. storage or transport via airplane. In this simple background model, two optimistic exposure times are chosen – 15 and 30 days – and it is assumed that the detectors begin taking data immediately after arriving underground. In practice, the assumption of immediate detector commissioning is justified due to the long decay time of tritium: the detector will not significantly cool down during a period of time underground much shorter than the tritium lifetime. The average rates due to these exposures once underground are then roughly 0.03 and 0.06 counts/keV/kg/day for 15 and 30 days, respectively, more than an order of magnitude greater than the assumed flat-background contribution.

The pdf of the tritium decay function was constructed in both time and energy using the corrected kinematic equation:

$$f_{^3\text{H}}(E, t) = g_{^3\text{H}}(E) \times h_{^3\text{H}}(t) \tag{5.2}$$

with

$$
\begin{aligned}
g_{^3\mathrm{H}}(E) &= \sqrt{(E + M_e)^2 - M_e^2}\,(E + M_e)\,(Q_{^3\mathrm{H}} - E)^2\,F(Z, E) \\
h_{^3\mathrm{H}}(t) &= e^{-t/\tau_{^3\mathrm{H}}} \\
F(Z, E) &= y(1 - e^{-y})^{-1}(1.002037 - 0.001427(v/c)) \\
y &= \frac{2\pi Z \alpha}{v/c}
\end{aligned}
$$

where $F(Z, E)$ is an approximation of the screened and relativistic Fermi function with $Z = 2$ for the daughter nucleus (as determined by J. J. Simpson in [94]), $Q_{^3\mathrm{H}} = 18.6$ keV, $M_e = 511$ keV, and $\tau_{^3\mathrm{H}} = 12.32/\log(2)$ years.

The background from neutrons has been estimated previously for the IGEX experiment [95] located at the Canfranc underground laboratory. This reference considered both neutrons from cosmic-ray muons interacting in the rock as well as neutrons arising from spontaneous fission and (α, n) reactions. Estimates from this – see Figure 7 and Tables 2,3 in [95], Figure 7 reproduced here in Figure 5.4 – suggested that contributions to IGEX from both sources should be below 0.01 counts/keV/kg/day down to 0 keV. Additionally, it was demonstrated that neutrons from rock radioactivity could be effectively eliminated with additional shielding. Though the MAJORANA DEMONSTRATOR will have a different geometry to that of the IGEX experiment, it is expected that these numbers are a conservative upper limit due to the fact that the DEMONSTRATOR will be in a deeper location (Sanford Underground Lab ($\lesssim 4500$ m.w.e.) vs. Canfranc (2500 m.w.e.)). From these conclusions and because the background from $^3\mathrm{H}$ should at least initially dominate the DEMONSTRATOR, any background contribution from neutrons was omitted from the background model.

In general, the detector response must be convolved with the spectrum to get a realistic shape. However, since point-contact detectors have such excellent energy resolution ($\sigma \sim 70$ eV), the response has limited effect on the spectra. Therefore, no correction for finite resolution was taken into account.
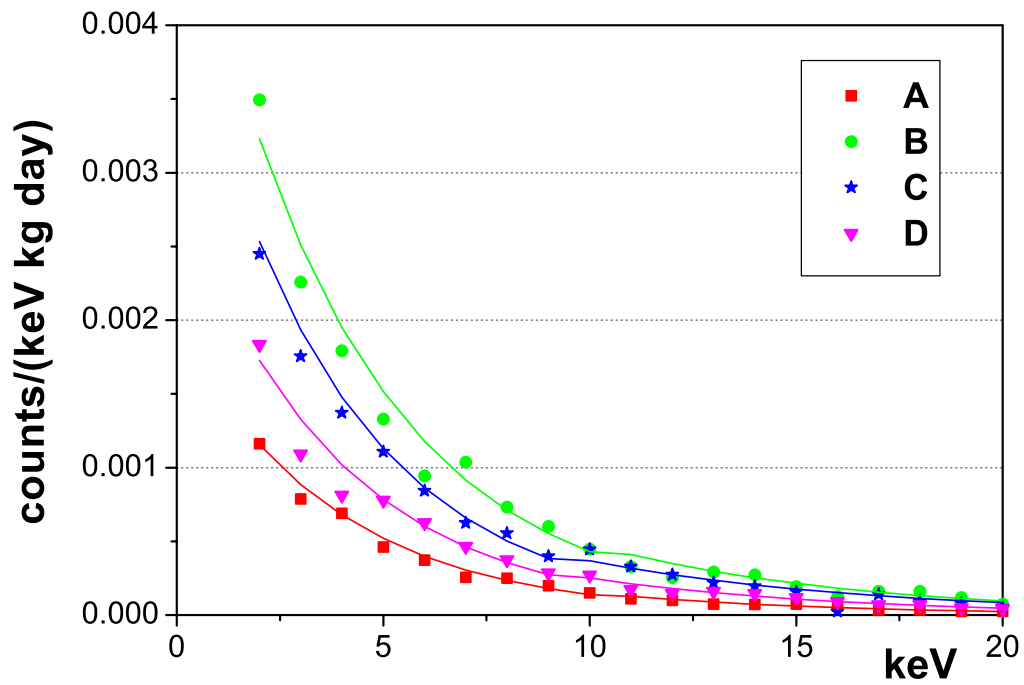
Figure 5.4: Simulated muon-induced-neutron spectrum for IGEX, reproduced from Figure 7 in reference [95]. A, B, C, and D are different neutron moderator configurations for the IGEX experiment, and the lines are exponential fits to guide the eye. A conservative extrapolation suggests that the smallest of these spectra (A) should be well below 0.01 counts/keV/kg/day at 0.5 keV and that all the spectra are much less than 0.001 counts/keV/kg/day above 10 keV.

Table 5.1: Variations on background and fitting for Majorana Demonstrator sensitivity calculations

Variable	Values
^{3}H exposure time	15, 30 days
Threshold	0.3, 0.5 keV
Demonstrator exposure	1, 5 years (20, 100 kg-years)

5.2.3 Sensitivity Fitting

The fitting procedure used binned maximum likelihood instead of an unbinned fit to reduce the time required in estimating an upper limit for each toy data model. Three parameters affecting the fits were varied: tritium exposure time, threshold, and exposure time of the Demonstrator. Each of these parameters could take two different values (see Table 5.1) so that for each signal eight different sets of fits were done. The data sets and fitting PDFs were both fully two-dimensional in energy and time and were fit over a range from threshold to 20 keV. The binning was chosen as follows: 256 bins for energy and 16 bins for time. For each of the eight set of fits for a signal, an ensemble of upper limits with a population $\gtrsim 1500^1$ was generated on the Athena cluster at the University of Washington [80]. The final sensitivity calculations outlined in the following sections took roughly 1500 hours of real-time or 460 cpu-days.

5.2.4 Sensitivity to WIMPs

The sensitivity calculations for WIMPs employed the 2-dimensional (time, energy) signal described in the previous chapter; more details can be found in Section 4.1. The values for all constants in the WIMP signal were used as previously noted. The fits proceeded as outlined in the previous section (Section 5.2.3). The shape of the WIMP signal was parameterized

---

[1] An exact population number was not specified for the ensembles to maximize the efficiency of the calculation. It was more efficient to specify the *time* for the calculation to run than to define a specific number of iterations.

solely by the mass of the WIMP, M_W, and so fits were performed with different values of this parameter on a variable grid (in GeV) of spacing given in ΔM_W, $M_W = 2.9$; 3; 3.5; $4 \rightarrow 10, \Delta M_W = 1$; $12 \rightarrow 24, \Delta M_W = 2$; $30 \rightarrow 100, \Delta M_W = 10$; $100 \rightarrow 1000, \Delta M_W = 100$. An example sensitivity fit to a WIMP signal at $M_W = 10$ GeV is given in Figure 5.5. In this particular example, the threshold was 0.3 keV, ^{3}H exposure time was 15 days, and the total exposure time was 1 year. It is clear that the amplitude of the excluded signal is very small and perhaps smaller than one might expect given the average count rate dominated by the ^{3}H decay. However, the shape of a WIMP signal is significantly different from that of a tritium beta spectrum making a stronger exclusion possible because the shape of the spectra are taken into account. A simple integration analysis looking solely at counts in a defined energy window should not produce the same limits.

The final results of the sensitivity calculation are presented in Figure 5.6. This figure includes results from the eight different sets of fits with the modification of the exposure times and the threshold. The factor of 2 increase in exposure time doesn't have a significant impact other than softening the exclusion limits as the WIMP mass increases. The threshold has a fairly significant impact below 10.5 GeV yielding better exclusions for WIMP masses and pushing the exclusion region diagonally left and down in the plot. This is certainly due to the sharp reduction of the tritium spectrum at low energies and to the sharp turn on of the exponential-like WIMP spectrum. As expected, the reduced threshold of 0.3 keV enables limits to be made down to $M_W = 3.5$ GeV.

It is instructive to compare the estimated sensitivity of the DEMONSTRATOR to that of other WIMP dark matter experiments. The MAJORANA experiment is distinct from most dark matter experiments since, for one, it is not primarily seeking to detect dark matter, but also because it will not rely upon any significant background reduction technique to enhance its sensitivity to WIMP-induced nuclear recoils. That is, the MAJORANA experiment will not have the ability to distinguish between nuclear and electron recoils. Despite this, it can still be competitive because it can access parameter space unreachable by other experiments with higher thresholds. A comparison of the sensitivity of the MAJORANA DEMONSTRA-TOR (assuming 0.3 keV threshold, 15 day tritium exposure, and 100 kg-yr exposure) to two characteristic dark matter experiments is given in Figure 5.7. These two experiments
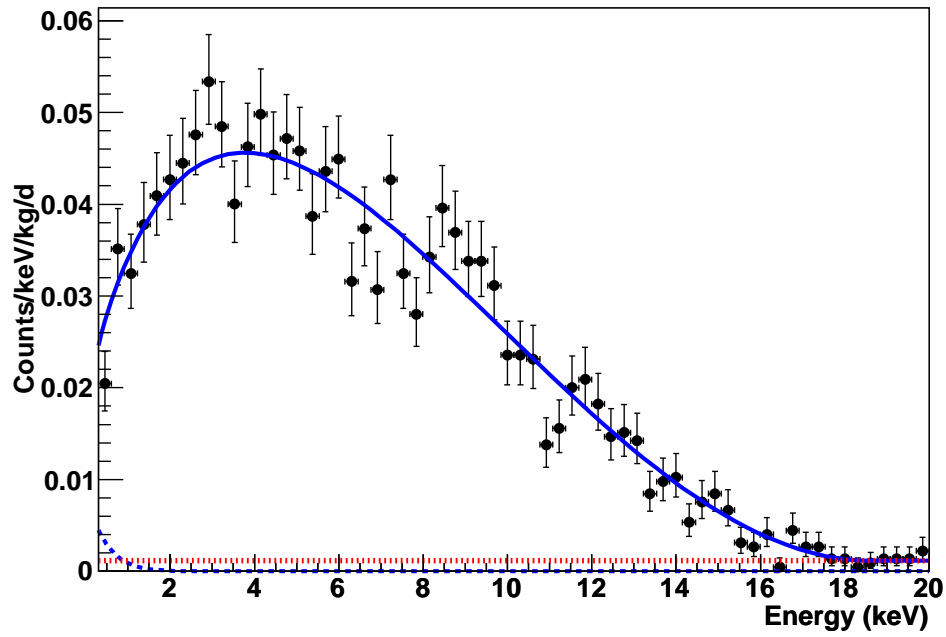
Figure 5.5: A sensitivity fit example with WIMP signal at $M_W = 10$ GeV, with σ_{nucl} excluded at 90% CL with value 8.6×10^{-9} pb. Components of the fit are broken out including WIMP signal (blue dashed) and flat background (red dotted). The beta spectrum from ^{3}H dominates the fit. Multiple fits like this were performed to estimate sensitivity (see Section 5.2.1).

168
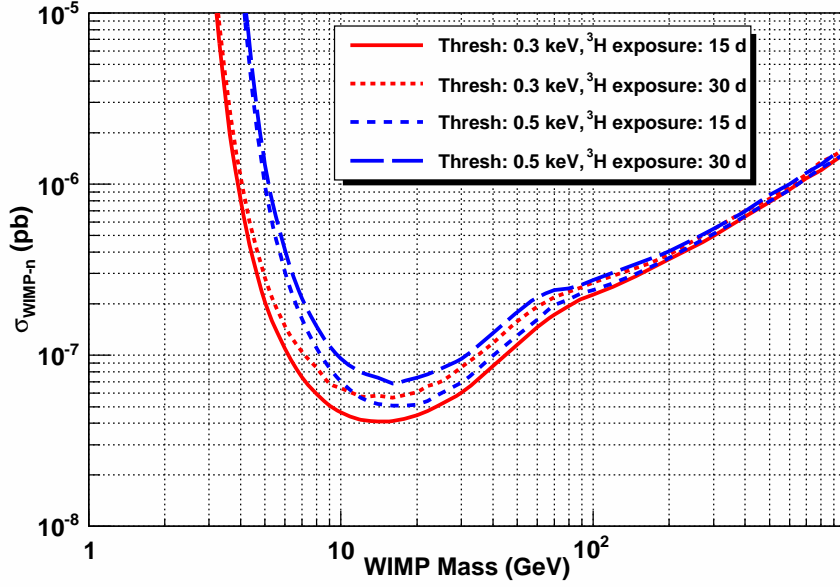
(a) 1 year (20 kg-yr) exposure time
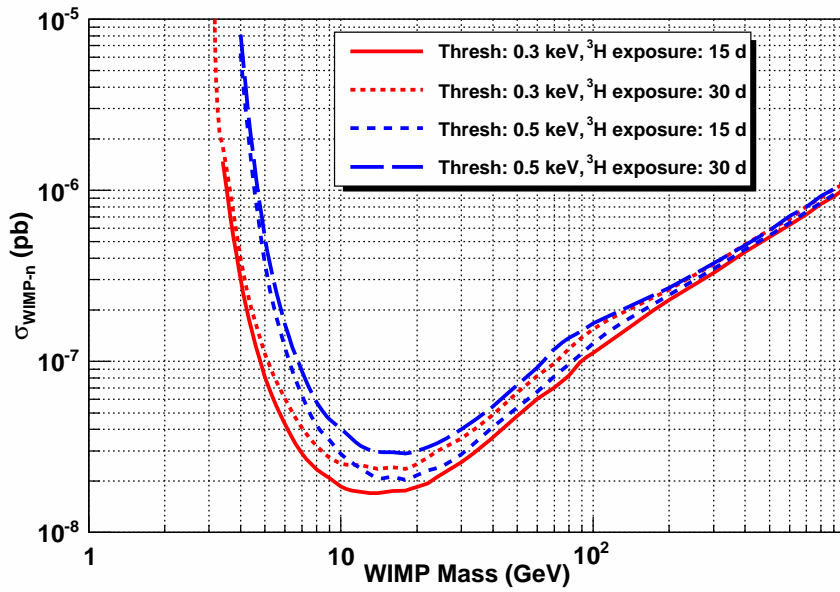
(b) 5 year (100 kg-yr) exposure time

Figure 5.6: MAJORANA DEMONSTRATOR sensitivity. Lines are 90% confidence level exclusions.

include Phase A of SuperCDMS [96] and the 300 kg LUX experiment [97]. The former is a proposed Ge-based dual-mode bolometer and ionization detector, the latter a proposed liquid-Xe-based detector that would read out both scintillation light and ionization. The dual-mode nature of these detectors gives them the ability to tag nuclear recoils and reject the electron recoils which dominate the background spectrum. Whereas it is obvious that the reach of these experiments in the higher-WIMP-mass range (> 10.5 GeV) far exceeds that of the MAJORANA DEMONSTRATOR, the low-energy threshold of the DEMONSTRATOR enables it to push well into the lower-WIMP-mass region.

5.2.5 Sensitivity to axioelectric effect

The sensitivity calculations for the axioelectric effect employed the signal model outlined earlier in Section 5.1.1. For simplicity and because it is expected that the rate of interaction does not vary significantly in time [28], this model was only 1-dimensional in energy. The mass of the axion, m_a, was chosen on a variable grid (in keV) with spacing given in Δm_a: $0.1 \to 1, \Delta m_a = 0.1$; $1 \to 9, \Delta m_a = 0.2$. An example sensitivity fit to an axioelectric signal centered at $m_a = 5$ keV is shown in Figure 5.8. In this example fit, a threshold of 0.3 keV and a tritium exposure time of 15 days is assumed. As before, components of the fit appear separately for comparison.

The sensitivity for the MAJORANA DEMONSTRATOR to the axioelectric effect is presented in Figure 5.9. All of the exclusion plots have similar characteristics, including a sharp drop at $m_a = 1.3$ keV and a rise at low m_a. The former arises from the corresponding feature in the rate (Figure 5.1) due to the L-capture line edge. The latter comes as the signal moves below threshold and the 'leakage' of signal from lower energies pushes above threshold due to the finite energy resolution of the detector. Larger tritium exposure time softens the sensitivity, but not significantly because $g_{a\bar{e}e}$ appears squared in the rate equation (Equation 5.1). Also, as expected at higher m_a the sensitivity has no dependence on the threshold.

To help compare with results presented before in Figure 5.3, Figure 5.10 is shown including exclusion limits from globular clusters [24]. Additional limit estimates from [28] are included, in particular those from searching for decays of exotic particles to photons and
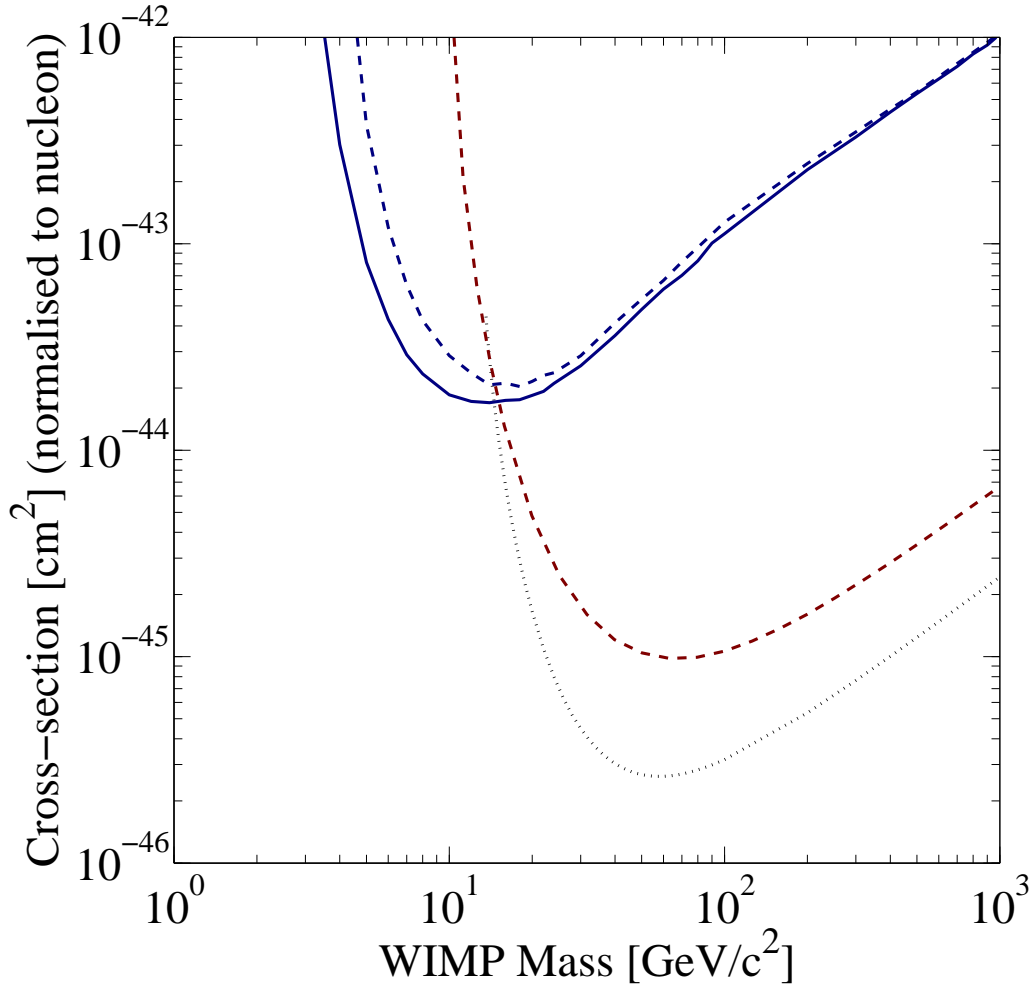
Figure 5.7: MAJORANA DEMONSTRATOR sensitivity to a WIMP signal (blue solid, 0.3 keV threshold; blue dashed, 0.5 keV threshold), comparing to SuperCDMS Phase A [96](red dashed) and LUX 300 [97] (black dotted). Plot generated with DMTools [86], lines are 90% CL exclusions.
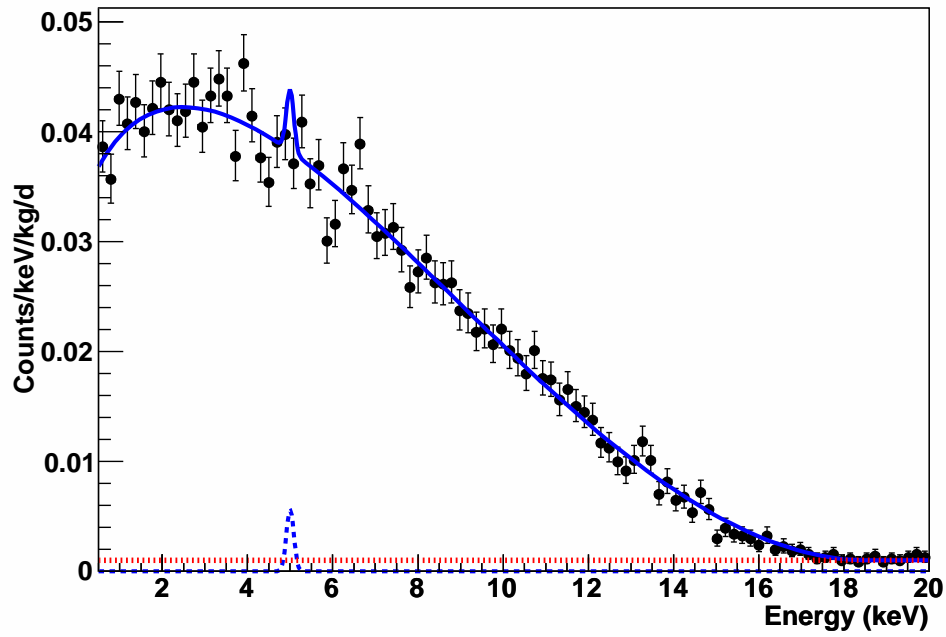
Figure 5.8: A sensitivity fit example with axioelectric signal at $m_a = 5$ keV, with $g_{a\bar{e}e}$ excluded at 90% CL with value 4.4×10^{-11} pb. Components of the fit are broken out including axioelectric signal (blue dashed) and flat background (red dotted). Multiple fits like this were performed to estimate sensitivity (see Section 5.2.1).
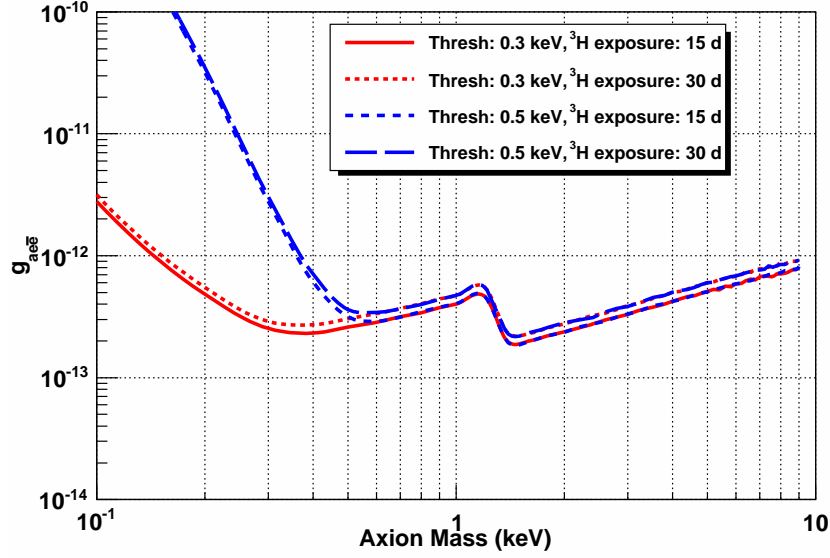
(a) 1 year (20 kg-yr) exposure time
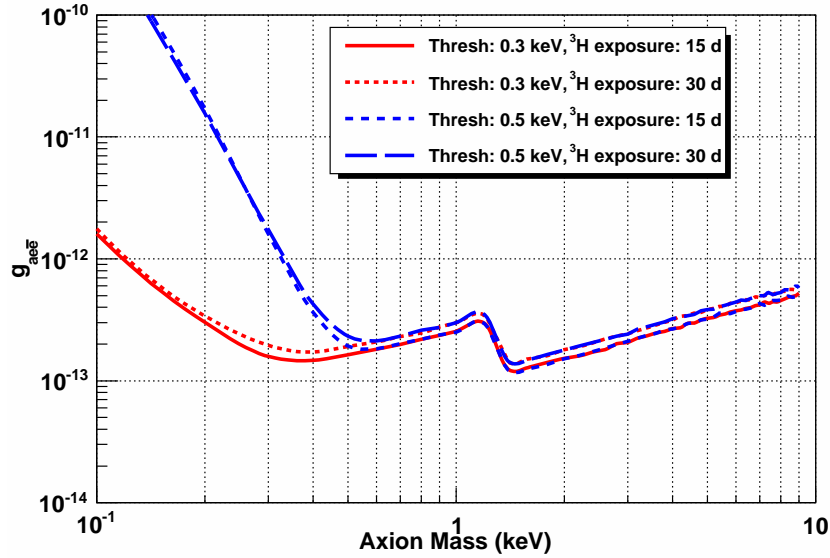
(b) 5 year (100 kg-yr) exposure time

Figure 5.9:   MAJORANA DEMONSTRATOR sensitivity at 90% CL to an axioelectric signal. Exposure time for ^{3}H and threshold were varied and results displayed as noted in the figure legends.

for big bang dark matter abundance considerations. For the first limit, the lifetime of the decay of the axion to 2 photons is given by (see [28]):

$$\Gamma_{a \to 2\gamma} = \frac{C_\gamma}{4\pi f_a^2} m_a^3 \tag{5.3}$$

with

$$C_\gamma \sim \frac{\alpha}{\pi} \frac{m_a^2}{m_e^2} \tag{5.4}$$

Following [28] and references therein, evidence from astronomical searches for monochromatic photons puts a limit in the 10 keV region of $\Gamma_{a \to 2\gamma} < 10^{-27} s^{-1}$ which yields an expected exclusion of

$$g_{a\bar{e}e} < 10^{-11} \left(\frac{\text{keV}}{m_a}\right)^{7/2} \tag{5.5}$$

As well, Pospelov et al. considered abundance limits on dark matter axions generated from b-quarks after the big bang, estimating that

$$g_{a\bar{e}e} < 10^{-13} \sqrt{\left(2\frac{\Omega_a}{\Omega_{\text{baryon}}}\right)\left(\frac{\text{keV}}{m_a}\right)} \tag{5.6}$$

where Ω_a is estimated as Ω_{DM} so that $\Omega_a/\Omega_{\text{baryon}} \sim 10$ and therefore:

$$g_{a\bar{e}e} < 4.5 \times 10^{-13} \sqrt{\frac{\text{keV}}{m_a}} \tag{5.7}$$

It is interesting to note that the MAJORANA experiment will demonstrate a better sensitivity to $g_{a\bar{e}e}$ than limits derived from astronomical observation in the limited m_a mass range $0.2 \to 2.5$ keV. Additionally, the sensitivity of the experiment will provide an important verification of the limits from globular clusters as well as of the constraints from the primordial abundance of axions and from the galactic γ-background.

5.3   Conclusions and Discussion

The results of this chapter indicate the usefulness of P-PC detectors for putting limits on dark matter, both for standard WIMP interactions as well as for more exotic interactions such as those from pseudoscalars. The MAJORANA DEMONSTRATOR will play an important role in the dark matter field especially since it will provide an experimental platform for searching for dark matter particles other than WIMPs and should be competitive with
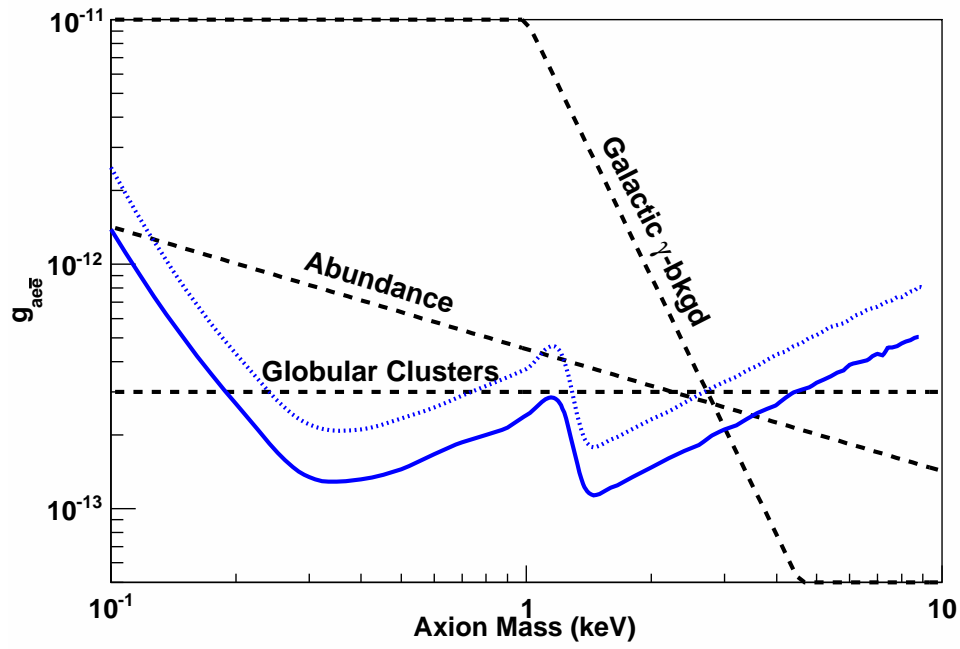
Figure 5.10: MAJORANA DEMONSTRATOR sensitivity at 90% CL (blue curved lines, dotted 20 kg-yr exposure, solid 100 kg-yr exposure) to an axioelectric signal, comparing to limits derived from astronomical observations including globular clusters [24], axion abundance after the big bang, and constraints from observed galactic γ backgrounds [28].

limits from astronomical observation. Of course, this chapter was unable to fully treat the rich landscape of available dark matter parameter space and some obvious candidates of study remain. A spin-dependent WIMP interaction could be studied using the DEMONSTRATOR since there exists one species of Ge with a non-zero nuclear spin:(^{73}Ge, $J = 9/2^+$). Since the relative content of this isotope in the detectors will be reduced by ^{76}Ge enrichment, a comparison with detectors with natural ^{73}Ge abundance (7.7%) could provide an opportunity for systematic studies. The size of the detector array should also make it an excellent candidate for searching for a generic annual modulation signature of interactions with velocity dependence (e.g. WIMP nuclear recoils). As well, other more exotic hypotheses might already exist for which MAJORANA could generate constraints, but those theories and analogous analyses of MAJORANA's sensitivity are left for the next thesis.

BIBLIOGRAPHY

[1] L. Camilleri, E. Lisi, and J. F. Wilkerson, "Neutrino Masses and Mixings: Status and Prospects," *Ann. Rev. Nucl. Part. Sci.* **58** (2008) 343–369.

[2] L. Baudis *et al.*, "Limits on the Majorana neutrino mass in the 0.1 eV range," *Phys. Rev. Lett.* **83** (1999) 41–44, `arXiv:hep-ex/9902014`.

[3] S. R. Elliott and P. Vogel, "Double beta decay," *Ann. Rev. Nucl. Part. Sci.* **52** (2002) 115–151, `arXiv:hep-ph/0202264`.

[4] A. Barabash, "Review of modern double-beta-decay experiments," *Phys. Atom. Nucl.* **70** (July 2007) 1191–1202(12).

[5] MAJORANA Collaboration. `http://majorana.npl.washington.edu`, July, 2010.

[6] R. Cooper, D. Radford, K. Lagergren, J. F. Colaresi, L. Darken, R. Henning, M. Marino, and K. Yocum, "A Pulse Shape Analysis Technique for the MAJORANA Experiment," *submitted to Nucl. Inst. & Meth. A* (2010) .

[7] P. Luke, F. Goulding, N. Madden, and R. Pehl, "Low capacitance large volume shaped-field germanium detector," *IEEE Trans. Nucl. Sci.* **36** no. 1, (Feb 1989) 926–930.

[8] P. S. Barbeau, J. I. Collar, and O. Tench, "Large-mass ultra-low noise germanium detectors: performance and applications in neutrino and astroparticle physics," *J. Cosm. Astro. Phys.* **0709** (2007) 009.

[9] E. Hull, R. Pehl, J. Lathrop, G. Martin, R. Mashburn, H. Miley, C. Aalseth, and T. Hossbach, "Segmentation of the outer contact on p-type coaxial germanium detectors," *Proceedings of the 27th Seismic Research Review: Ground-Based Nuclear Explosion Monitoring Technologies* (Sept., 2007) 764–769. `http://handle.dtic.mil/100.2/ADA519832`.

[10] E. Hull, R. Pehl, J. Lathrop, P. Mann, and R. Mashburn, "P-Type Point Contact Germanium Detectors for Low-Level Counting," *Proceedings of 30th Monitoring Research Review: Ground-Based Nuclear Explosion Monitoring Technologies* (Sept., 2008) 768–771. `http://handle.dtic.mil/100.2/ADA517246`.

[11] **CoGeNT** Collaboration, C. E. Aalseth *et al.*, "Experimental Constraints on a Dark Matter Origin for the DAMA Annual Modulation Effect," *Phys. Rev. Lett.* **101** no. 25, (Dec., 2008) 251301, `arXiv:0807.0879`. Erratum: *ibid.* 102 (2009) 109903.

[12] C. E. Aalseth, *Germanium spectrometer pulse-shape discrimination for germanium-76 double-beta decay.* PhD thesis, University of South Carolina, 2000.

[13] D. Budjáš, M. Barnabé Heider, O. Chkvorets, N. Khanbekov, and S. Schönert, "Pulse shape discrimination studies with a Broad-Energy Germanium detector for signal identification and background suppression in the GERDA double beta decay experiment," *J. Inst.* **4** (2009) P10007, `arXiv:0909.4044 [nucl-ex]`.

[14] K. G. Begeman, A. H. Broeils, and R. H. Sanders, "Extended rotation curves of spiral galaxies: Dark haloes and modified dynamics," *Mon. Not. Roy. Astron. Soc.* **249** (1991) 523.
`http://articles.adsabs.harvard.edu/full/1991MNRAS.249..523B`.

[15] C. Amsler *et al.*, "Review of Particle Physics," *Phys. Lett. B* **667** no. 1-5, (2008) 1 – 6.

[16] D. Clowe, M. Bradač, A. H. Gonzalez, M. Markevitch, S. W. Randall, C. Jones, and D. Zaritsky, "A Direct Empirical Proof of the Existence of Dark Matter," *Astrophys. J. Lett.* **648** no. 2, (2006) L109.

[17] G. Jungman, M. Kamionkowski, and K. Griest, "Supersymmetric dark matter," *Phys. Rev.* **267** (1996) 195–373, `arXiv:hep-ph/9506380`.

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

[19] J. L. Feng, "Supersymmetry and cosmology," *Ann. Phys.* **315** no. 1, (2005) 2 – 51. Special Issue.

[20] J. Kopp, V. Niro, T. Schwetz, and J. Zupan, "DAMA/LIBRA data and leptonically interacting dark matter," *Phys. Rev. D* **80** no. 8, (Oct, 2009) 083502.

[21] R. D. Peccei and H. R. Quinn, "CP Conservation in the Presence of Pseudoparticles," *Phys. Rev. Lett.* **38** no. 25, (Jun, 1977) 1440–1443.

[22] M. Kuster, G. Raffelt, & B. Beltrán, ed., *Axions*, vol. 741 of *Lecture Notes in Physics, Berlin Springer Verlag.* 2008.

[23] P. Gondolo and G. G. Raffelt, "Solar neutrino limit on axions and keV-mass bosons," *Phys. Rev. D* **79** no. 10, (May, 2009) 107301.

[24] G. Raffelt and A. Weiss, "Red giant bound on the axion-electron coupling reexamined," *Phys. Rev. D* **51** no. 4, (Feb, 1995) 1495–1498.

[25] E. Arik *et al.*, "Probing eV-scale axions with CAST," *J. Cosm. Astro. Phys.* **2009** no. 02, (2009) 008.

[26] S. J. Asztalos, G. Carosi, C. Hagmann, D. Kinion, K. van Bibber, M. Hotz, L. J. Rosenberg, G. Rybka, J. Hoskins, J. Hwang, P. Sikivie, D. B. Tanner, R. Bradley, and J. Clarke, "SQUID-Based Microwave Cavity Search for Dark-Matter Axions," *Phys. Rev. Lett.* **104** no. 4, (Jan, 2010) 041301.

[27] R. J. Gaitskell, "Direct Detection of Dark Matter," *Ann. Rev. Nucl. Part. Sci.* **54** no. 1, (2004) 315–359.

[28] M. Pospelov, A. Ritz, and M. B. Voloshin, "Bosonic super-WIMPs as keV-scale dark matter," *Phys. Rev. D* **78** (2008) 115012, `arXiv:0807.3279 [hep-ph]`.

[29] R. Bernabei, P. Belli, A. Bussolotti, F. Cappella, R. Cerulli, C. Dai, A. d'Angelo, H. He, A. Incicchitti, H. Kuang, J. Ma, A. Mattei, F. Montecchia, F. Nozzoli, D. Prosperi, X. Sheng, and Z. Ye, "The DAMA/LIBRA apparatus," *Nucl. Inst. & Meth. A* **592** no. 3, (2008) 297 – 315.

[30] J. Orrell, "Note on the Purchase and Initial Operation of the PNNL Point Contact Detector," Tech. Rep. M-TECHDOCDET-2008-018, MAJORANA internal document, 2007.

[31] J. Orrell, "Pulse Shape Analysis of a p-Type Point Contact Germanium Detector for Neutrinoless Double-beta Decay and Dark Matter Searches," Tech. Rep. M-TECHDOCPHYS-2008-008, MAJORANA internal document, 2008.

[32] J. Anderson, R. Brito, D. Doering, T. Hayden, B. Holmes, J. Joseph, H. Yaver, and S. Zimmermann, "Data Acquisition and Trigger System of the Gamma Ray Energy Tracking In-Beam Nuclear Array (GRETINA)," *IEEE Trans. Nucl. Sci.* **56** (Feb, 2009) 258.

[33] T. Howe, M.A.and Bergmann, A. Kopmann, F. McGirt, M. Marino, K. Rielage, J. Wilkerson, and J. Wouters, "ORCA: Object-oriented Real-time Control and Acquisition." `http://orca.physics.unc.edu/`.

[34] The Apache Software Foundation, "The CouchDB Project." `http://couchdb.apache.org/`, April, 2010.

[35] R. Brun and F. Rademakers, "ROOT: An object oriented data analysis framework," *Nucl. Inst. & Meth. A* **389** (1997) 81–86.

[36] V. T. Jordanov and G. F. Knoll, "Digital synthesis of pulse shapes in real time for high resolution radiation spectroscopy," *Nucl. Inst. & Meth. A* **345** (June, 1994) 337–345.

[37] S. Rab, "Nuclear Data Sheets Update for A = 133," *Nuclear Data Sheets* **75** no. 3, (1995) 491 – 666.

[38] J. Morales, E. Garcia, A. O. de Solorzano, A. Morales, R. N. nez Lagos, J. Puimedon, C. Saenz, and J. A. Villar, "Filtering microphonics in dark matter germanium experiments," *Nucl. Inst. & Meth. A* **321** no. 1-2, (1992) 410 – 414.

[39] J. Orrell, "Pulse Shape Analysis of a p-Type Point Contact Germanium Detector for Neutrinoless Double-beta Decay and Dark Matter Searches," Tech. Rep. M-TSPCONFPROC-2008-027, Majorana internal document, 2008.

[40] A. G. Schubert, "Soudan PPC-II high-energy data," Tech. Rep. M-TECHDOCDET-2010-094, Majorana internal document, 2009.

[41] S. Cebrián, J. Amaré, B. Beltrán, J. M. Carmona, E. García, H. Gómez, I. G. Irastorza, G. Luzón, M. Martínez, J. Morales, A. O. de Solórzano, C. Pobes, J. Puimedón, A. Rodríguez, J. Ruz, M. L. Sarsa, L. Torres, and J. A. Villar, "Cosmogenic activation in germanium double beta decay experiments," *J. Phys. Conf. Ser.* **39** no. 1, (2006) 344.

[42] F. T. Avignone *et al.*, "Theoretical and experimental investigation of cosmogenic radioisotope production in germanium," *Nucl. Phys. B (Proc. Suppl.)* **28** no. 1, (1992) 280–285.

[43] S. R. Elliott, V. E. Guiseppe, B. H. LaRoque, R. A. Johnson, and S. G. Mashnik, "Fast-Neutron Activation of Long-Lived Isotopes in Enriched Ge," *submitted to Phys. Rev. C.* (2009) , arXiv:0912.3748 [nucl-ex].

[44] D.-M. Mei, Z.-B. Yin, and S. Elliott, "Cosmogenic production as a background in searching for rare physics processes," *Astroparticle Phys.* **31** no. 6, (2009) 417 – 420.

[45] E. Schönfeld, U. Schötzig, E. Günther, and H. Schrader, "Standardization and decay data of 68Ge/68Ga," *Applied Radiation and Isotopes* **45** no. 9, (1994) 955 – 961.

[46] J. A. Bearden and A. F. Burr, "Reevaluation of X-Ray Atomic Energy Levels," *Rev. Mod. Phys.* **39** no. 1, (Jan, 1967) 125–142.

[47] H. V. Klapdor-Kleingrothaus, L. Baudis, A. Dietz, G. Heusser, I. Krivosheina, B. Majorovits, and H. Strecker, "GENIUS-TF: a test facility for the GENIUS project," *Nucl. Inst. & Meth. A* **481** no. 1-3, (2002) 149 – 159.

[48] I. Barabanov, S. Belogurov, L. Bezrukov, A. Denisov, V. Kornoukhov, and N. Sobolevsky, "Cosmogenic activation of germanium and its reduction for low background experiments," *Nucl. Inst. & Meth. B* **251** no. 1, (2006) 115 – 120.

[49] J. Back and Y. Ramachers, "ACTIVIA: Calculation of isotope production cross-sections and yields," *Nucl. Inst. & Meth. A* **586** no. 2, (2008) 286 – 294.

[50] **CoGeNT** Collaboration, C. E. Aalseth *et al.*, "Results from a Search for Light-Mass Dark Matter with a P- type Point Contact Germanium Detector," `arXiv:1002.4703 [astro-ph.CO]`.

[51] W. Verkerke and D. Kirkby, "The RooFit toolkit for data modeling," `physics/0306116`.

[52] P. S. Barbeau, *Neutrino and Astroparticle Physics with P-Type Point Contact High Purity Germanium Detectors*. PhD thesis, University of Chicago, 2009.

[53] D. L. Donoho, I. M. Johnstone, G. Kerkyacharian, and D. Picard, "Wavelet Shrinkage: Asymptopia?," *J. R. Stat. Soc. B* **57** no. 2, (1995) 301–369. `http://www.jstor.org/stable/2345967`.

[54] D. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inf. Th.* **41** no. 3, (May, 1995) 613 –627.

[55] R. R. Coifman and D. Donoho, "Translation-Invariant De-Noising," in *Wavelets and statistics*, A. Antoniadis and G. Oppenheim, eds., vol. 103 of *Lecture Notes in Statistics*, pp. 125–150. Springer-Verlag, 1995. `http://www-stat.stanford.edu/~donoho/Reports/1995/TIDeNoise.pdf`.

[56] G. Nason and B. Silverman, "The Stationary Wavelet Transform and some Statistical Applications," in *Wavelets and statistics*, A. Antoniadis and G. Oppenheim, eds., vol. 103 of *Lecture Notes in Statistics*, pp. 125–150. Springer-Verlag, 1995.

[57] F. Wasilewski, "The PyWavelets python Wavelet package." `http://wavelets.scipy.org/moin/`, April, 2010.

[58] D. Donoho and I. M. Johnstone, "Adapting to Unknown Smoothness via Wavelet Shrinkage," *J. Amer. Stat. Assoc.* **90** (1995) 1200–1224. `http://www.jstor.org/stable/2291512`.

[59] A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Analytical Chemistry* **36** (1964) 1627–1639.

[60] H. S. W. C. Tseung, *Simulation of the Sudbury Neutrino Observatory Neutral Current Detectors.* PhD thesis, Wadham College, University of Oxford, 2008.

[61] F. Gatti *et al.*, "Study of Sensitivity Improvement for MARE-1 in Genoa," *J. Low Temp. Phys.* **151** (2008) 603–606.

[62] M. H. Chen, B. Crasemann, and H. Mark, "Relativistic K-shell Auger rates, level widths, and fluorescence yields," *Phys. Rev. A* **21** (Feb., 1980) 436–441.

[63] M. H. Chen, B. Crasemann, and H. Mark, "Widths and fluorescence yields of atomic L-shell vacancy states," *Phys. Rev. A* **24** (July, 1981) 177–182.

[64] E. J. McGuire, "Atomic M-Shell Coster-Kronig, Auger, and Radiative Rates, and Fluorescence Yields for Ca-Th," *Phys. Rev. A* **5** (Mar., 1972) 1043–1047.

[65] P. Morrison and L. I. Schiff, "Radiative K Capture," *Phys. Rev.* **58** (July, 1940) 24–26.

[66] A. D. Rújula, "A new way to measure neutrino masses," *Nuclear Physics B* **188** no. 3, (1981) 414 – 458.

[67] J. N. Bahcall, "Exchange and Overlap Effects in Electron Capture and in Related Phenomena," *Phys. Rev.* **132** no. 1, (Oct, 1963) 362–367.

[68] M. Strauss and R. Larsen, "Pulse height defect due to electron interaction in the dead layers of Ge(li) [gamma]-Ray detectors," *Nucl. Inst. & Meth.* **56** no. 1, (1967) 80 – 92.

[69] E. Sakai, "Slow Pulses from Germanium Detectors," *IEEE Trans. Nucl. Sci.* **18** no. 1, (Feb., 1971) 208 –218.

[70] E. L. Hull, R. H. Pehl, N. W. Madden, P. N. Luke, C. P. Cork, D. L. Malone, J. S. Xing, K. Komisarcik, J. D. Vanderwerp, and D. L. Friesel, "Temperature sensitivity of surface channel effects on high-purity germanium detectors," *Nucl. Inst. & Meth. A* **364** no. 3, (1995) 488 – 495.

[71] R. H. Helm, "Inelastic and Elastic Scattering of 187-Mev Electrons from Selected Even-Even Nuclei," *Phys. Rev.* **104** no. 5, (Dec, 1956) 1466–1475.

[72] G. Alner *et al.*, "First limits on nuclear recoil events from the ZEPLIN I galactic dark matter detector," *Astroparticle Phys.* **23** no. 5, (2005) 444 – 462.

[73] J. Lindhard and M. Scharff, "Energy Dissipation by Ions in the kev Region," *Phys. Rev.* **124** no. 1, (Oct, 1961) 128–130.

[74] D. J. Venzon and S. H. Moolgavkar, "A Method for Computing Profile-Likelihood-Based Confidence Intervals," *J. R. Stat. Soc. C* **37** no. 1, (1988) 87–94. `http://www.jstor.org/stable/2347496`.

[75] S. Yellin, "Finding an upper limit in the presence of an unknown background," *Phys. Rev. D* **66** no. 3, (Aug, 2002) 032005.

[76] W. A. Rolke and A. M. López, "Confidence intervals and upper bounds for small signals in the presence of background noise," *Nucl. Inst. & Meth. A* **458** no. 3, (2001) 745 – 758.

[77] G. Angloher *et al.*, "Limits on WIMP dark matter using sapphire cryogenic detectors," *Astroparticle Phys.* **18** no. 1, (2002) 43 – 55.

[78] W. A. Rolke, A. M. López, and J. Conrad, "Limits and confidence intervals in the presence of nuisance parameters," *Nucl. Inst. & Meth. A* **551** (Oct., 2005) 493–503, `physics/0403059`.

[79] F. James and M. Roos, "Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations," *Comput. Phys. Commun.* **10** (1975) 343–367.

[80] Physics and A. C. S. at the Unversity of Washington, "The Athena Cluster." `http://librarian.phys.washington.edu/athena/index.php/Main_Page`.

[81] A. G. S. Ocampo and D. C. Conway, "*LK*-Capture Ratio of *Zn*65," *Phys. Rev.* **128** no. 1, (Oct, 1962) 258–261.

[82] H. Brysk and M. E. Rose, "Theoretical Results on Orbital Capture," *Rev. Mod. Phys.* **30** no. 4, (Oct, 1958) 1169–1177.

[83] A. H. Wapstra, *Nuclear spectroscopy tables, by A.H. Wapstra, G.J. Nijgh [and] R. Van Lieshout.* North-Holland Pub. Co.; Interscience Publishers, Amsterdam, New York,, 1959.

[84] **The CDMS-II** Collaboration, Z. Ahmed *et al.*, "Results from the Final Exposure of the CDMS II Experiment," `arXiv:0912.3592 [astro-ph]`.

[85] C. Savage, G. Gelmini, P. Gondolo, and K. Freese, "Compatibility of DAMA/LIBRA dark matter detection with other searches," *J. Cosm. Astro. Phys.* **0904** (2009) 010, `arXiv:0808.3607 [astro-ph]`.

[86] R. Gaitskell and V. Mandic, "Dark Matter Results Plotter."
http://dmtools.brown.edu, July, 2010.

[87] R. Bernabei *et al.*, "Investigating pseudoscalar and scalar dark matter," *Int. J. Mod. Phys. A* **21** (2006) 1445–1470, arXiv:astro-ph/0511262.

[88] C. T. Chantler, "Detailed Tabulation of Atomic Form Factors, Photoelectric Absorption and Scattering Cross Section, and Mass Attenuation Coefficients in the Vicinity of Absorption Edges in the Soft X-Ray (Z=30–36, Z=60–89, E=0.1 keV–10 keV), Addressing Convergence Issues of Earlier Work." Online, 22 june, 2010. http://physics.nist.gov/ffast. *J. Phys. Chem. Ref. Data*, 4 (2000) 597.

[89] **CDMS** Collaboration, Z. Ahmed *et al.*, "Search for Axions with the CDMS Experiment," *Phys. Rev. Lett.* **103** no. 14, (Oct, 2009) 141802.

[90] J. I. Collar and M. G. Marino, "Comments on arXiv:0902.4693v1 'Search for Axions with the CDMS Experiment'," arXiv:0903.5068 [hep-ex].

[91] I. G. Irastorza *et al.*, "Present status of IGEX dark matter search at Canfranc underground laboratory," *Nucl. Phys. B (Proc. Suppl.)* **110** (2002) 55–57, arXiv:hep-ex/0111073.

[92] MAJORANA Collaboration, R. Gaitskell *et al.*, "White paper on the Majorana zero-neutrino double-beta decay experiment," arXiv:nucl-ex/0311013.

[93] J. I. Collar. PhD thesis, University of South Carolina, 1992.

[94] J. J. Simpson, "Measurement of the β-energy spectrum of ^{3}H to determine the antineutrino mass," *Phys. Rev. D* **23** no. 3, (Feb, 1981) 649–662.

[95] J. M. Carmona, S. Cebrián, E. García, I. G. Irastorza, G. Luzún, A. Morales, J. Morales, A. O. de Solórzano, J. Puimedón, M. L. Sarsa, and J. A. Villar, "Neutron background at the Canfranc underground laboratory and its contribution to the IGEX-DM dark matter experiment," *Astroparticle Phys.* **21** no. 5, (2004) 523 – 533.

[96] **CDMS-II** Collaboration, D. Akerib *et al.*, "The SuperCDMS proposal for dark matter detection," *Nucl. Inst. & Meth. A* **559** no. 2, (2006) 411 – 413. Proceedings of the 11th International Workshop on Low Temperature Detectors - LTD-11.

[97] **LUX** Collaboration, "Projected Dark Matter Limits."
http://lux.brown.edu/experiment_sens.html, June, 2010.

[98] M. Howe, G. Cox, P. Harvey, F. McGirt, K. Rielage, J. Wilkerson, and J. Wouters, "Sudbury neutrino observatory neutral current detector acquisition software overview," *IEEE Trans. Nucl. Sci.* **51** no. 3, (June, 2004) 878 – 883.

[99] M. A. Howe, M. G. Marino, and J. F. Wilkerson, "Integration of embedded single board computers into an object-oriented software bus DAQ application," in *Nuclear Science Symposium Conference Record, 2008. NSS '08. IEEE*, pp. 3562 –3567. 19-25, 2008.

[100] I. Abt, A. Caldwell, K. Kröninger, J. Liu, X. Liu, and B. Majorovits, "Identification of photons in double beta-decay experiments using segmented germanium detectors–Studies with a GERDA phase II prototype detector," *Nucl. Inst. & Meth. A* **583** no. 2-3, (2007) 332 – 340.

[101] J. L. Orrell, C. E. Aalseth, M. W. Cooper, J. D. Kephart, and C. E. Seifert, "Radial position of single-site gamma-ray interactions from a parametric pulse shape analysis of germanium detector signals," `arXiv:nucl-ex/0703022`.

[102] **GERDA** Collaboration. `http://www.mpi-hd.mpg.de/gerda`, July, 2010.

[103] M. Frigo and S. G. Johnson, "The Design and Implementation of FFTW3," *Proceedings of the IEEE* **93** no. 2, (2005) 216–231. Software available: `http://www.fftw.org/`.

[104] E. Gatti, P. F. Manfredi, M. Sampietro, and V. Speziali, "Suboptimal filtering of $1/f$-noise in detector charge measurements," *Nucl. Inst. & Meth. A* **297** no. 3, (1990) 467 – 478.

[105] **KATRIN** Collaboration. `http://www-ik.fzk.de/~katrin`, July, 2010.

[106] **SNO+** Collaboration. `http://snoplus.phy.queensu.ca`, July, 2010.

Appendix A

DEVELOPMENT OF A DIGITAL DATA ACQUISITION SYSTEM FOR P-TYPE POINT-CONTACT DETECTORS

A.1 Introduction

The detectors in the MAJORANA DEMONSTRATOR will be readout using a digital data acquisition system using fast ADCs to digitize raw preamp traces to be used in later pulse-shape analysis. The requirements of a digital DAQ system for the DEMONSTRATOR were investigated, in particular focusing on refining the hardware specifications needed to achieve the physics goals of the project. These specifications must be sufficient to achieve two general goals: (1) background reduction through pulse-shape analysis, especially in the $0\nu\beta\beta$ region-of-interest, and (2) low-energy performance enabling both background reduction at $Q_{0\nu\beta\beta}$ and sensitivity to low-energy physics (dark matter).

This appendix focuses on investigations of several issues of the DAQ:

- Sampling rate

- Operational stability

- Low-energy performance

  - Triggering

  - Low-energy resolution

but does not detail the development of all the hardware and software to run most of these tests. For more information regarding some of that work, please see Appendix C.

Because P-type Point Contact detectors will be used in the MAJORANA DEMONSTRATOR, it was necessary to use such a detector for some of these tests. A P-PC detector (henceforth referred to as P-PC2) was procured by PNNL in January 2008 to be used in studies

aimed at more clearly understanding the detector technology [30]. Initial tests took place in the laboratory, focusing on basic operation and pulse-shape analysis techniques relevant for background reduction in the $0\nu\beta\beta$ region-of-interest and for Dark Matter searches [31]. Some of these results are presented here (see Section A.3). Following these tests, in November 2008 the detector was deployed underground at Soudan National Laboratory at a depth of 2100 m.w.e. to investigate both the low-energy performance and operational stability.

A.2 Baseline DAQ System

The baseline system for development used the Gretina Mark IV digitizer designed by the GRETA collaboration [32]. This VME64x-based ADC digitizes 10 independent channels at 100 Ms/s with a resolution of 14 bits. The card was read out using the ORCA DAQ software developed at the University of Washington and at the University of North Carolina [98, 99, 33].

A.3 Digitizer Comparison Tests

Since the Gretina card is being considered as a candidate card for the MAJORANA DEMONSTRATOR, head-to-head tests between it and digitizers with different characteristics were necessary. The purpose of these measurements was to investigate how the specifications of each digitizer affected its ability to perform pulse-shape analysis for background reduction through multi- and single-site event selection. The decided method was to take data from P-PC2 concurrently with different digitizers using a ^{232}Th source. ^{208}Tl, a daughter in the thorium chain, produces a double-escape peak (DEP) at 1592.5 keV which is predominantly composed of single-site events and is commonly used (e.g. [100, 101]) to test pulse-shape analysis algorithms. After taking concurrent data, the waveform data were processed using the same algorithm to determine how each digitizer performed.

A.3.1 Measurement

The measurement with P-PC2 took place at Pacific Northwest National Laboratory running a round-robin test with 3 different digitizers, two cards at a time. The digitizers used were the Gretina Mark IV, XIA DGF4-c, and the XIA Pixie-4, details about which are given

Table A.1: Comparison of digitizer characteristics used in this test.

Name	Sampling Rate	Bit Resolution	Form Factor
XIA DGF4-c	40 Ms/s	14 bits	CAMAC
XIA Pixie-4	75 Ms/s	14 bits	PCI
Gretina Mark IV	100 Ms/s	14 bits	VME64x

in Table A.1. The P-PC2 preamp had two identical signal outputs, and so each of these was AC-coupled to one of the two digitizers being used for each test. The inhibit output, which generated a logic pulse when the reset circuitry of the preamp was active, was split and input in to each digitizer. The time signature of the inhibit pulse was later used to synchronize the data sets (see Section A.3.2).

The DAQ software used for both XIA cards was an IGOR Pro-based package designed and shipped by XIA. The setup for the Gretina card was similar but simpler to that described in Section 2.2, again using ORCA to readout the card. The data from each DAQ system were converted to a common format and stored in ROOT files using MGDO data objects. More details on MGDO can be found in Appendix B. Each XIA card was run in parallel with the Gretina card and around 2 hours live-time of source data were taken for each pair, though the wall clock time was roughly 2 (4) times this during runs with the DGF4-c (Pixie-4) due to induced the dead-time of the XIA cards. (The Gretina card is designed to be dead-time-less.)

A.3.2   Analysis

The data were converted to the ROOT-based MGDO format using the OrcaROOT processing libraries. The amplitudes of each Gretina pulse were calculated using an offline trapezoidal filter. The value of the on-board trapezoidal calculation was used for each XIA card. Once converted, the runs with each pair of cards were synchronized together using the timing of the inhibit pulse. The basic algorithm searched for common time differences between reset pulse events, accommodating the fact that the XIA cards tended to miss
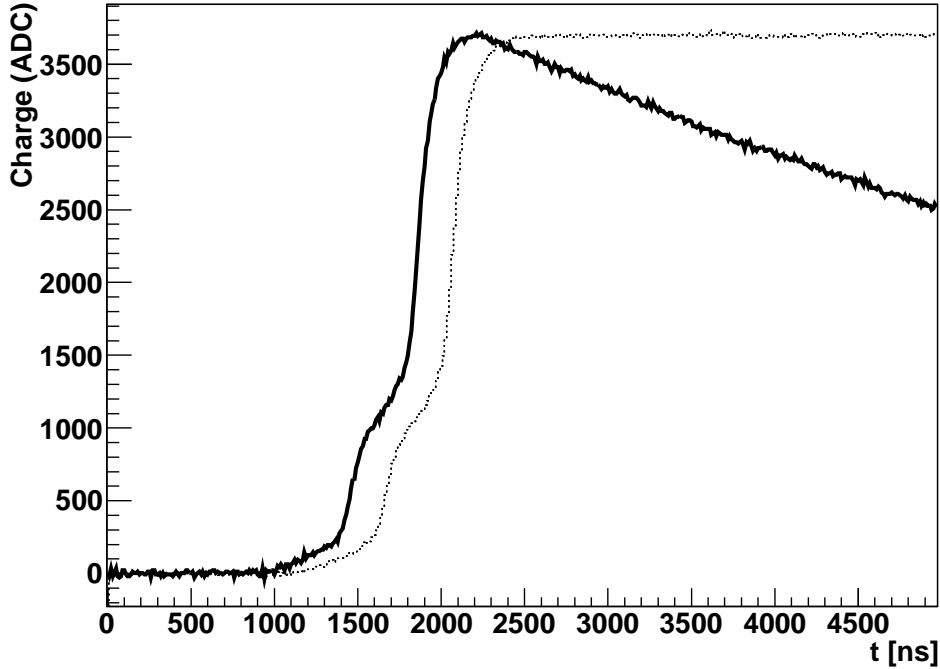
Figure A.1: Visualization of two synchronized pulses from the Gretina (bold) and Pixie digitizers. The different decay constant are due to the different input impedances of each card.

numbers of these events given their dead-time and that their 32-bit on-board clocks rolled over more often than the 48-bit clock on the Gretina card. Once a catalog of synchronized events was determined, the data were compiled into a single set of ROOT files for later pulse-shape analysis. An example of a synchronized event is shown in Figure A.1. The difference in decay time constants is due to the different input impedances of each card. (Same-valued capacitors were used to AC couple each card.) The synchronization between the two cards ensured that each event processed was generated by the same physics event.

After synchronization, a simple pulse-shape analysis algorithm [13] developed by the GERDA collaboration was applied to the data. In this algorithm, the ratio A/E is calculated, where A is the maximum of the current pulse (derivative of the charge pulse) and E

is the amplitude (energy) of the charge pulse. This ratio can then be used to discriminate between single- and multi-site events due to the fact that events with multiple interaction sites tend to have wider current pulses for a given energy, E. A version of this algorithm was implemented using waveform processing code in MGDO.

Before processing, the data sets were reduced to include only events in a ~50 keV window around the ^{208}Tl DEP at 1592.5 keV. For each event of this subset, the current waveform was generated by using a Savitzky-Golay derivative filter [59] of size 5 and degree 4 and the current maximum was saved. Cuts were then applied to the data for a range of values for A/E and the data were fit using binned maximum likelihood using the RooFit toolkit [51] over the range 1570 → 1615 keV. This fit was used to estimate the survival probability of the ^{208}Tl DEP and background reduction in the continuum and nearby peaks, including the 1588.2 and 1580.5 keV gamma peaks of ^{228}Ac. An example of such a fit is shown in Figure A.2. A direct comparison between the pairs of cards is possible by looking at the background reduction and survival probability versus the cut parameter, A/E, as calculate for each digitizer. These results are shown for DGF4-c and Gretina cards in Figure A.3, and for the Pixie-4 and Gretina cards in Figure A.4. It is clear from these plots that the 3 cards behave comparably with this particular algorithm. The only discrepancy exists in the DGF4-c vs. Gretina results: the DGF4-c has a slightly steeper acceptance curve around an A/E value of 62.

A.3.3   Conclusions

For the simple pulse-shape algorithm presented here the 3 digitizer cards perform similarly, suggesting that the A/E calculation is not significantly sensitive to the sampling frequency of the fast ADC. More advanced PSA, such as those based upon comparing pulses to a library of single-site events [6], are more likely to depend on the sampling characteristics of the digitizer. The following sections explore other characteristics of the Gretina digitizer to determine how well its hardware and firmware configuration enable it to perform at low energies.
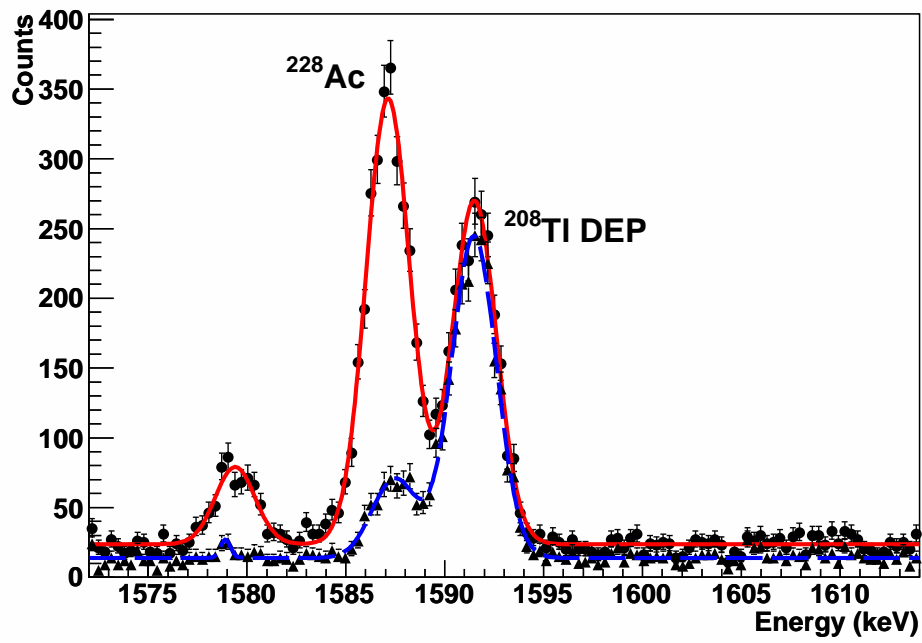
Figure A.2: An example set of fits using the Gretina card. The solid line (circles) is without cuts, the dashed (triangles) with cuts, yielding a signal acceptance of $94.8 \pm 1.77\%$ in the ^{208}Tl DEP and a $89.1 \pm 0.99\%$ reduction in the adjacent 1588.2 keV ^{228}Ac peak.
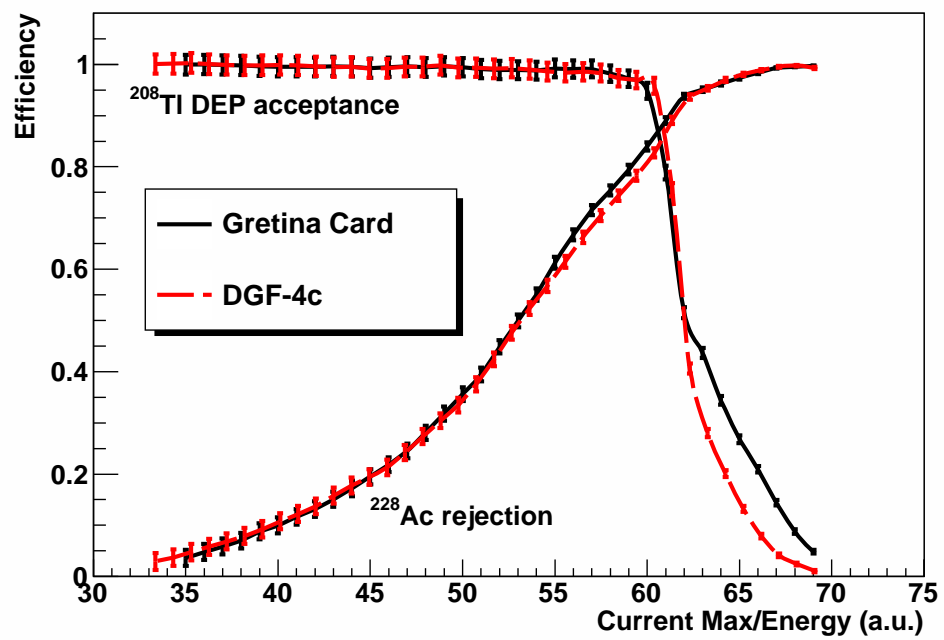
Figure A.3: Cuts for Gretina vs. DGF-4c.
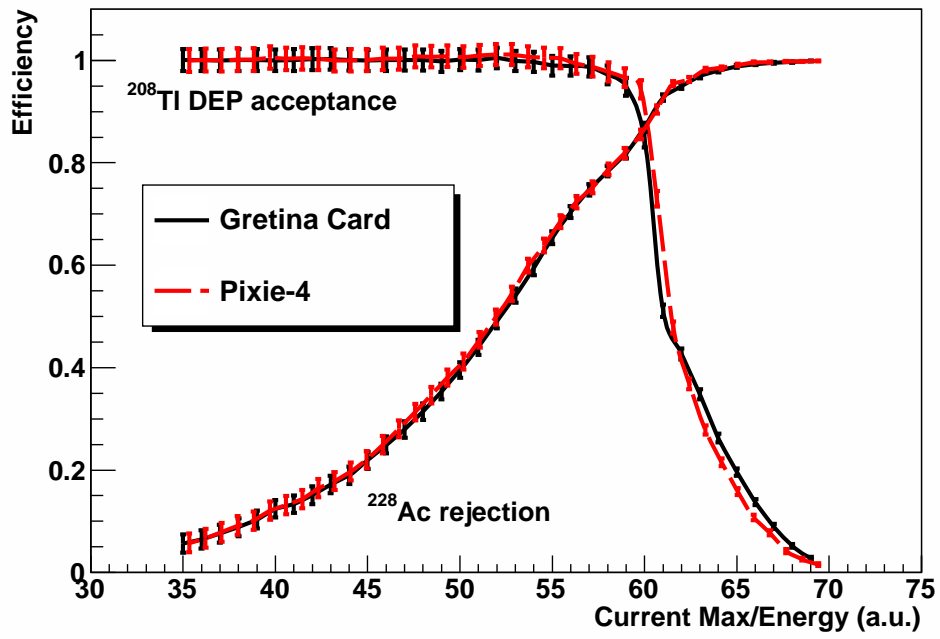
Figure A.4: Cuts for Gretina vs. Pixie-4c.

A.4 Development and Testing of the Gretina Mark IV Digitizer

A.4.1 Trigger Design and Tests

Before deployment of the DAQ system underground in Soudan, triggering tests were performed to determine the optimum conditions for minimizing the energy threshold of the electronics. The goal of these initial measurements was also to develop an automated set of tests to perform regularly on P-PC2 *in situ*. Triggering efficiency refers to the probability of inducing a trigger at a particular signal amplitude. It should be 1 well above threshold and decrease to 0 as the amplitude of the signal reduces below threshold. The basic technique to measure the trigger efficiency of a detector system is to inject a pulse of known amplitude into the test port of the electronics and complete a series of measurements, adjusting the amplitude of the pulse to sample around the threshold. One can determine the probability of detecting a pulse by either knowing the rate of the injected pulse and performing the test for a known period of time, or by having an independent measure of the timing of the injected pulse (i.e. a synchronization pulse) and performing a coincidence measurement. For these tests, the latter method was chosen as it was deemed a cleaner technique to extract both the trigger efficiency given a certain pulse amplitude as well as the false trigger rate at a particular threshold setting.

The test setup included the pulser and computer-controlled attenuators described in Section 2.2 and was run by the ORCA DAQ software. To vary the amplitude of the injected pulse, it was decided to keep the output pulse from the waveform generator constant and change the attenuator settings. Scripts were designed in ORCA to perform these variations automatically. The attenuated pulse was amplified using a Phillips 777 before being injected directly into a Gretina card to ensure that the noise of the input signal dominated the intrinsic noise of the digitizer. The amplitude of the measured pulse was estimated using an offline trapezoidal filter as was done in later analysis (see Section 2.3). Because a similar detector system to P-PC2 was unavailable above ground at the time of the tests, it was impossible to simulate the exact noise environment of P-PC2's electronics. To circumvent this limitation, all the measurements were determined in terms of signal-to-noise ratios to be able to compare directly to the detector system. For example, a measured signal-to-noise

ratio could be multiplied by the separately measured magnitude of the detector noise to provide a rough calibration of the results. Manufacturer specifications quoted this value as 180 eV FWHM[1] [30].

Initial measurements found that the Gretina on-board trigger, a leading-edge discrimination (LD) differential algorithm with fixed shaping, achieved ∼90% efficiency at an S/N of ∼6, suggesting that a similar efficiency would be found at ∼1 keV on the detector system. This was at least a factor of 2 worse performance than demonstrated in analog readout systems with a detector of similar noise characteristics [8]. The degraded trigger performance was due to the limited shaping associated with the LD trigger which had not originally been designed to trigger on very-low-amplitude preamp signals. To solve this issue, a hybrid digital/analog system was designed: the signal was split after the 777, one line running directly to the digitizer and the other into a spectroscopy amplifier with 1 μs shaping time. The output from the spectroscopy amplifier was input into the Gretina card and this channel was used to trigger the unshaped input channel. Longer shaping times were tested, but were found to trigger poorly since the differential algorithm was insensitive to the leading edge of a slower-rising pulse. Tests with this hybrid system indicated an improvement of triggering efficiency. Results comparing the two methods are presented in Figure A.5.

A.4.2   Measured Electronic Noise

The electronic noise was measured by injecting a pulse from the waveform generator and calculating the FWHM of the width of the peak. Since this calculation was performed offline, the parameters of the trapezoidal filter (i.e. integration time and collection time) could be varied over the same data set to determine the values which would yield the best resolution. The trace length of the waveform was limited to 10 μs and, since the rising edge of the waveform was positioned in the middle of the digitization window, the offline filter integration length was limited to less than 5 μs. In practice, the limitation on the integration time was closer to 3.5 μs to account for variations in the position of the rising edge of the

---

[1]This value was measured using an analog shaping amplifier and is dependent upon the shaping times of the amplifier. In practice, this value is different than one calculated using digital shaping (i.e. with a trapezoidal filter, as was done in this analysis) and so was interpreted as an estimate when compared directly to digital measurements.
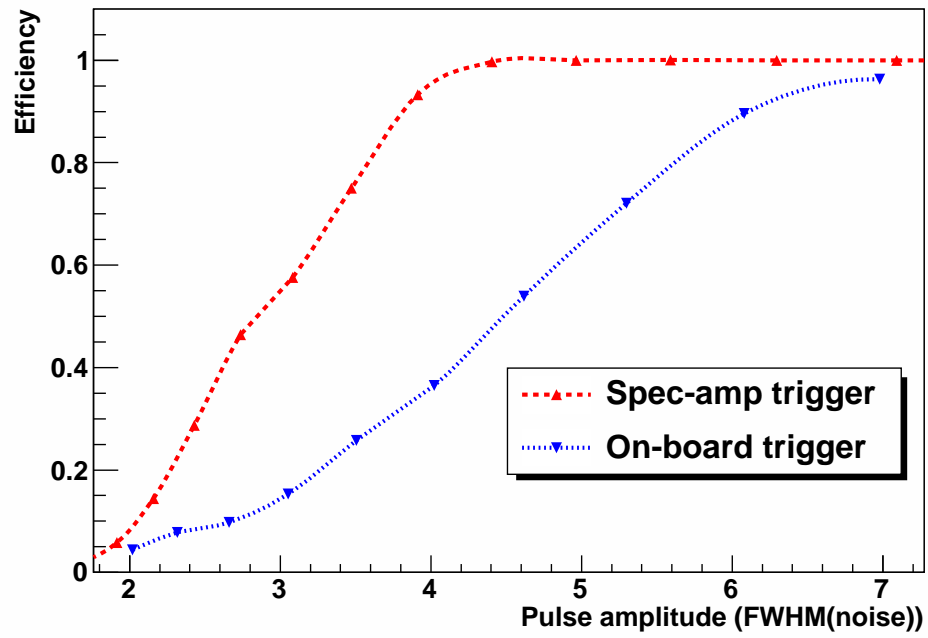
Figure A.5: Triggering efficiency test results comparing the on-board LD trigger to the hybrid system. The hybrid system was found to have a factor of ∼2 improvement.
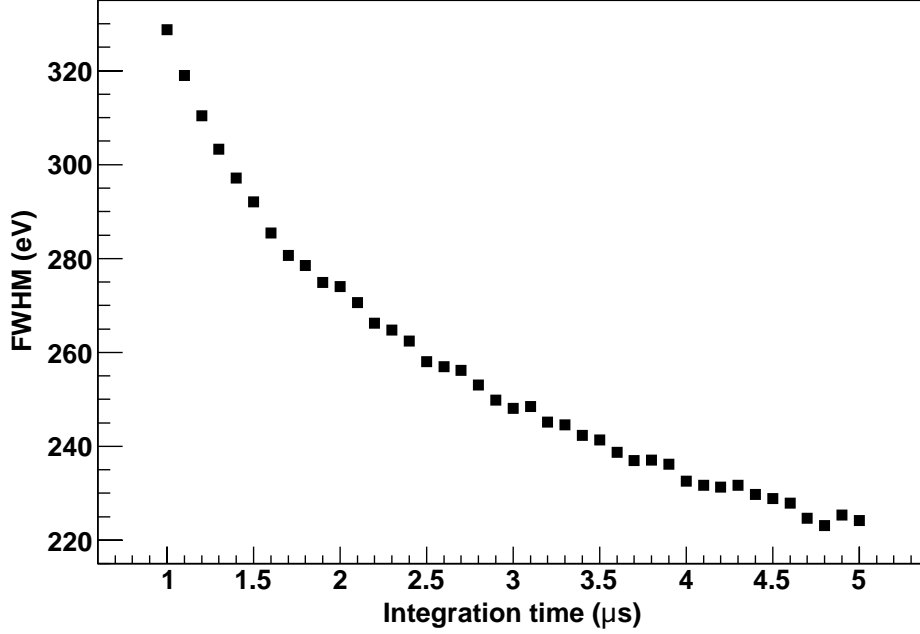
Figure A.6: Noise versus integration time of the trapezoidal filter.

pulse with different pulse amplitudes. Results, shown in Figure A.6, indicate that the best resolution comes at the longest shaping time and also suggest that the minimum resolution could not be achieved with this setup. The minimum value measured in this test, 225 eV, was larger than the value measured with an analog electronics system, 180 eV, but it is expected that this difference would shrink if longer integration times were available.

A.4.3 Conclusions

The Gretina card provides sufficient characteristics for physics near the $0\nu\beta\beta$ Q-value (2 MeV), but does not clearly achieve good enough threshold or resolution at low energies in comparison to results from analog systems. Because of this, additional work beyond the scope of this thesis began, focusing on refining the triggering algorithms and onboard hardware to enable better resolution and triggering capabilities. This work is focusing on developing these capabilities in two cards, including the Struck 16-bit, 100 MS/s 3302 and the

Gretina digitizer. Despite the low-energy limitations of the tested firmware on the Gretina card, this DAQ system was deployed to readout the P-PC2 detector placed underground at Soudan Underground Laboratory. Results of this deployment are presented in Chapter 2.

Appendix B

ANALYSIS TOOLS

This appendix outlines several technical aspects related to the software work in this dissertation and can serve as a full or partial reference for the following:

1. MAJORANA GERDA Data Objects (MGDO) – software framework for encapsulation and processing of waveforms

2. Analysis, Run database and processing framework

3. pyWIMP, a software framework for generating limits on a WIMP signal

B.1 MAJORANA *GERDA Data Objects (MGDO)*

The MAJORANA GERDA Data Objects (MGDO) C++-based software package has been jointly developed by the MAJORANA [5] and GERDA [102] collaborations as a framework for encapsulating and processing waveform data. This section will reference some of the functionality and layout of this software package, focusing on the overall structure as well as particular aspects specifically relevant to this dissertation. This is meant to supplement the reference materials already available, including a basic user and installation guide that ships with the distribution. MGDO is available on specific request to the MAJORANA or GERDA software groups at the following subversion repository `svn://pclg-soft.mppmu.mpg.de/MGDO`. This section will outline the basic structure and functionality of MGDO and then proceed to describe in more detail the waveform transformations that ship with the software and have been developed related to this dissertation work, emphasizing aspects important for users rather than programmers of the package.

B.1.1 Structure

The MGDO software is composed in the following package structure:

1. Base – Base classes

2. Root – Root wrappers for base classes, allowing serialization of encapsulation classes and usage of calculation classes from ROOTCINT or pyROOT.

3. Transforms – Transform classes for objects performing a calculation or modification on a waveform

4. Majorana – MAJORANA-specific classes intended for encapsulation of data in the MAJORANA data format

5. Gerda – GERDA-specific classes intended for encapsulation of data in the GERDA data format.

Further description will focus only on the first three sub-packages and will leave the reference of objects particular to the MAJORANA or GERDA collaborations to their respective owners or developers.

B.1.2 Base

This package includes base classes for all the objects within MGDO, including encapsulation classes and others providing virtual interfaces for derived. The main encapsulation classes are as follows:

`class MGWaveform` Provides a base encapsulation class for waveforms, storing e.g. raw trace, sampling information, type (charge, current, etc.) and providing access functions and overloaded operators for simple calculations.

`class MGWaveformFT` Provides a base encapsulation class for frequency waveforms, and is specifically tailored to hold data from the fast Fourier transform of real-data. Provides similar access and storage functionality to `MGWaveform`.

The main virtual base classes are:

class `MGVWFFastFourierTransform` Provides a virtual interface for fast Fourier transform (FFT) classes. Derived from this class are particular implementations, including `MGWFFastFourierTransformDefault` and `MGWFFastFourierTransformFFTW` which provide a basic FFT and a wrapped version of FFTW3 [103], respectively.

class `MGVWaveformTransformer` Provides a virtual interface for generating waveform transformations or calculations, see Section B.1.4. class `MGMultiWaveformTransformer` provides a base class for chaining multiple waveform transformations together.

class `MGVFreqDomainTransformer` Provides a virtual interface for generating transformations or calculations on a waveform in frequency space, see Section B.1.4. This base class handles the necessary transformations to Fourier space. class `MGMultiFreqDomainTransformer` provides a base class for chaining multiple frequency-domain transformations together.

B.1.3 Root

The Root sub-package includes objects that may depend on the ROOT analysis package. In particular, this directory provides wrapped code for Base class objects, enabling them to be serialized or loaded during a ROOT interactive session in ROOTCINT or pyROOT. It also introduces a class: class `MGTDataObject : public TNamed` which provides an inheritance layer between ROOT objects and MGDO objects. In general, every MGDO class requiring ROOT dependence should derive from this class. Some base classes are merely wrapped:

```
1  ...
2  class MGTVWaveformTransformer : public MGTDataObject, public
       MGVWaveformTransformer
3  {
4    public:
5      MGTVWaveformTransformer(const std::string& aTransformationName);
6      virtual ~MGTVWaveformTransformer();
7      virtual void Transform(MGWaveform* anInput, MGWaveform* anOutput = NULL)
           = 0;
8
```

```
 9    protected:
10    ClassDef(MGTVWaveformTransformer, 0)
11  };
12  ...
```

whereas others add some functionality, e.g. `class MGTWaveform` :

```
 1  ...
 2  class MGTWaveform : public MGTDataObject, public MGWaveform
 3  {
 4      ...
 5      // Returns a histogram named MGTWaveformHist_[label/ID]. If the user
            does not
 6      // supply an ID, the waveform ID is used. The function first searches
 7      // gROOT for MGTWaveformHist_[label/ID]. If it is found, that hist is
            used.
 8      // If it is not found, the function returns a newly allocated histogram
 9      // of the waveform. Either way, it is the user's responsibility to
10      // delete the histogram.
11      virtual TH1D* GimmeHist(const char* label="");
12
13      // Deprecated versions
14      virtual TH1D* GetNewHist();
15      virtual TH1D* GetNewHist(int id);
16
17      // When you don't care what the lable is and just want a brand-new,
18      // unique histogram, call the following function
19      virtual TH1D* GimmeUniqueHist();
20
21      // Loads the waveform into a user-supplied hist. This function is safer
22      // because it is obvious that the user owns (and therefore must later
23      // delete) this histogram.
24      virtual void LoadIntoHist(TH1D* hist);
25
26      // Returns a TF1 object that is based on the
27      // loaded waveform.  There are three parameters for this TF1:
28      //    parameters[0] - normalization (a)
29      //    parameters[1] - constant factor (b)
```

```
30        //     parameters[2] - offset (c)
31        //
32        //    MGWaveform is a function of x == wf(x)
33        // This function returns a TF1 with the underlying structure:
34        //   a * wf( b*x + c )
35        //
36        // Linear interpolation is applied when values are requested between
37        // data indices.
38        //
39        // iCopy specifies a different number of the function to return.
40        // The name of the function takes the form:
41        //
42        // MGTWaveformFunction_[ID]_[Copy]
43        //
44        // where ID is the id of the histogram and Copy is the passed in iCopy.
45        // If this function already exists, it is deleted and
46        virtual TF1* GetFunction( int iCopy = 0 );
47
48   protected:
49        enum EWFFunctionPars { kNumberOfParameters = 3 };
50        TF1* fWaveformFunction;
51        double InterpolatingWaveformFunction( const double* x,
52                                              const double* parameters );
53
54   ClassDef(MGTWaveform, 2) // Waveform class, holds waveforms
55 };
56
57 ...
```

which adds the capabilities of interpolation and export to other ROOT class formats (e.g. `TF1`, `TH1D`).

B.1.4 Transforms

This package includes basic waveform transforms which are derived from `class MGVWaveformTransformer` and `class MGVFreqDomainTransformer`, which define the basic interfaces

for time-domain and frequency-domain transformations, respectively. The latter derives
from the former. The use of the transformations involves calling the class function `virtual`
`void MGVWaveformTransformer::Transform(MGWaveform* input, MGWaveform* output = NULL)` to gen-
erate the transformation. If the function is called with `output=NULL` (the default) then the
result of the transformation is placed in the `input` variable. The particulars of the calcula-
tion depend on whether or not the transformation can be done in-place or out-of-place. For
example, a transformation done in-place can directly modify the input waveform, whereas
a out-of-place transformation would require an additional auxiliary waveform to perform
the transformation. The following sections describe the general transformations available
in the Transforms package. If algorithms are described, $g[n]$ is used to denote the output
waveform and $f[n]$ the input waveform.

class MGWFAddNoise

Adds Gaussian (white) noise to the input waveform. The amplitude of the noise can be set
by calling `virtual void MGWFAddNoise::SetNoiseAmplitude(double aVal)` defined as the σ of the
gaussian noise.

class MGWFAddNoiseFromFT

Adds noise to a waveform given an input noise spectrum in frequency space. The input spec-
trum is defined using the class function `virtual void MGWFAddNoiseFromFT::SetNoiseWaveform(`
`MGWaveformFT* aWF)` and is assumed to be an average noise spectrum with added together
Fourier transforms of noise (i.e. the Fourier transforms should be added using the class
function `MGWaveformFT::AddNormsSquared(const MGWaveformFT& other)`). This information is used
to add noise to the waveform since each bin of the average power spectrum is related to the
sigma of the real and imaginary parts of the noise (see [60] and Section 3.4.3).

class MGWFBandpassFilter

Performs a bandpass filter on a waveform, using upper and lower bandpass frequen-
cies defined using the class functions `MGWFBandpassFilter::SetLowerBandpass(double lower)` and

`MGWFBandpassFilter::SetUpperBandpass(double upper)`. The input frequencies use CLHEP units, for example `SetLowerBandpass(1*CLHEP::GHz)`. The algorithm provides a hard cutoff at the defined frequencies with no roll-off. In other words, all frequency components of the waveform not within the lower and upper bounds are set to zero.

`class MGWFBaselineRemover`

Removes the baseline from a waveform. The baseline is estimated using a simple average beginning at a start time or start sample number of the waveform defined by one of the class functions `MGWFBaselineRemover::SetStartSample( size_t iSample )` or `MGWFBaselineRemover::SetStartTime( double aTime )`. The length of the baseline averaging is defined by either `MGWFBaselineRemover::SetBaselineSamples( size_t nSamples )` or `MGWFBaselineRemover::SetBaselineTime( double aTime )`. A utility function is included in this class to just estimate the baseline of a waveform: `virtual double MGWFBaselineRemover::GetBaseline(const MGWaveform& waveform)`.

`class MGWFCalculateChiSquare`

Calculates the chi-square difference (RMS) between the input and output waveforms. This transformation leaves both waveforms unchanged and the total chi-square can be retrieved using the the class function `double MGWFCalculateChiSquare::GetChiSquareValue()`.

`class MGWFCountZeroCrossings`

Counts the number of zero crossings in a region of the waveform defined using either class function `SetSearchRegion( const MGWaveformRegion& aVal )` or `SetSearchTimeRegion(double tlower, double tupper)`. This can be used in conjunction with, for example, `MGWFDerivative` to determine the number of peaks in a waveform. The class function `SetRequiredDifference( double aVal )` can be used to set the lower threshold for a zero crossing. This can be set to ignore noise oscillations around zero. Results from this transformation can be obtained using the following class functions: `size_t GetNumberOfZeroCrossings()`, which returns the number of zero crossings seen, and `const std::vector<double>& GetZeroCrossingTimeVector`

(), which returns the time of the zero crossings in a vector. This class was used extensively when performing the pulse-shape analysis outlined in Section A.3.2.

class MGWFDerivative

Performs a simple derivative on the waveform, implementing the following function:

$$g[n] = (f[n+1] - f[n-1])/sp$$

where sp is the sampling period of the waveform.

class MGWFDerivativeFourthOrder

Performs a fourth-order derivative on the waveform, estimating the derivative using additional points:

$$g[n] = (-f[n+2] + 8f[n+1] - 8f[n-1] + f[n-2])/(12sp)$$

where sp is the sampling period of the waveform.

class MGWFDigitizer

Takes a continuous waveform and 'digitizes' it, returning a waveform that has been resampled (see MGWFResampler) 'digtized' according to some range and bit depth. The range may be defined using the functions SetMaxSignal(double aVal) and SetMinSignal(double aVal), and the bit depth can be specified by SetNumberOfBits(size_t nBits). The input value from the waveform is assigned an integer value based upon the nearest 'digital' assignable value. The output sampling frequency is defined using the function SetSamplingFrequency(double freq).

class MGWFExtremumFinder

Does a simple search to find the maximum or minimum of a waveform. The behavior of the class is defined by the class function SetFindMaximum(bool findMax) which will enable either finding the maximum or minimum. Results are obtained from the class functions virtual const size_t& GetTheExtremumPoint(), which returns the point in the waveform, and virtual const double& GetTheExtremumValue(), which returns the value at that point.

`class MGWFIntegral`

Performs a simple integral of the waveform according to the following algorithm:

$$g[n] = g[n-1] + f[n]$$

`class MGWFMovingAverage`

Smooths a waveform by performing a moving average of a defined width. The width is defined using the class function `SetSmoothSize( size_t aVal )` and region of smoothing (the part of the waveform that is smoothed) can be defined using the function `SetSmoothRegion( const MGWaveformRegion& aVal )`.

`class MGWFMovingWindow`

Performs a moving window transformation on a waveform. This algorithm can be used together with `MGWFPoleZeroCorrection` to determine the amplitude of a preamplifier waveform. However, it is generally more desirable to use `MGWFStaticWindow` defined later. The algorithm used:

$$g[n] = g[n-1] + f[n] - f[n-rt] - f[n - ft - rt] + f[n - ft - 2rt]$$

where ft and rt are the flat and ramp times, respectively, defined by the class functions `SetFlatTime( double aVal )` and `SetRampTime( double aVal )`.

`class MGWFPoleZeroCorrection`

Removes a pole from a waveform defined by a time constant. Before calling this transformation it is essential to remove the baseline from a waveform. The following algorithm is used:

$$g[n] = g[n-1] + f[n] + \tau sf(f[n] - f[n-1])$$

where sf is the sampling frequency and τ is the time constant of the pole defined using the class function `SetDecayConstant(double aVal)`. This transformation can be used to remove a pole induced by a preamplifier in order to estimate the amplitude of the waveform.

```
class MGWFPulseFinder
```

Finds pulses (or regions of pulses) in a waveform. The baseline must be removed from a waveform before passing it through this transformation. The algorithm searches the waveform and determines when it goes either above a positive threshold or below a negative threshold and stores the results. These saved regions can then be retrieved using the class function `const std::vector< MGWaveformRegion >& GetThePulseRegions()`. The threshold is set using `SetThreshold( double aThresh )`.

```
class MGWFRCDifferentiation
```

Handles an RC differentiation of a waveform simulating the passing of a waveform through a simple preamp. The following algorithm is applied:

$$g[n] = g[n-1] + f[n] - f[n-1] - g[n-1]/(\tau s f)$$

where sf is the sampling frequency and τ is the RC time constant defined using the class function `SetTimeConstant( double aVal )`.

```
class MGWFRCIntegration
```

Handles an RC integration of a waveform simulating the passing of a waveform through a simple preamp integrator. The following algorithm is applied:

$$g[n] = f[n-1] + g[n-1](1 - 1/(tausf))$$

where sf is the sampling frequency and τ is the RC time constant defined using the class function `SetTimeConstant( double aVal )`. The output waveform has an amplitude multiplied by $\tau s f$.

```
class MGWFResampler
```

Compresses or resamples a waveform. The parameters to resample the input waveform are defined by the parameters of the output waveform. For example, to downsample a waveform sampled at 1 GHz to 100 MHz, the sampling frequency of the output waveform must be

defined to be 100 MHz. Various methods are used to resample the waveform and these different methods are selected using `SetMode(EMode mode)` where `EMode` is an enum taking the following values:

- `kInterpolate` (default): take a strict interpolation from the input waveform at the requested output sample time

- `kIntegrate`: integrate over the input waveform for one sample of the output waveform. The position of the integration region relative to the output sample time is chosen via `SetSampleReference()`. Each sample is multiplied by the sampling period (like a proper integral), and hence the waveform type of the output is different than the input (e.g. current becomes charge)

- `kRebin`: like `kIntegrate`, but doesn't multiply by the sampling period or change the waveform type (as if the waveform is a histogram whose contents merely need to be "rebinned")

- `kAverage`: like `kIntegrate`, but divides by the number of samples (including fractional samples) involved in the integral over the bin width. Doesn't change the waveform type.

`class MGWFRisetimeCalculation`

Finds the rise-time of a waveform. Before this transform is called the baseline of the pulse should be removed and the maximum value of the waveform should be set using the class function `SetPulsePeakHeight(double maximum)`. This is to allow the user to use different methods to find the maximum and maximize usefulness of this class. This algorithm always assumes the pulse is rising/falling left to right. This class contains a number of access and behavior-modification functions, which are most easily expressed in the following:

- `virtual void SetScanFrom( size_t scanfrom )` – Sets the position to begin scanning forward on the waveform for the mid point of the rising edge

- `virtual void SetInitialThresholdPercentage( double threshold )` – Sets the initial threshold percentage (should be between 0, 1)

- `virtual void SetFinalThresholdPercentage( double threshold )` – Sets the final threshold percentage (should be between 0, 1)

- `virtual void SetInitialScanToPercentage( double threshold )` – Sets the initial scanning percentage to find the middle of the pulse (should be between 0, 1)

- `virtual double GetRiseTime()` – Returns rise-time with CLHEP units of time

- `virtual size_t GetInitialThresholdCrossing()` – Returns point of the initial threshold crossing in the waveform.

- `virtual size_t GetFinalThresholdCrossing()` – Returns point of the final threshold crossing in the waveform.

- `virtual double GetInitialThresholdCrossingEstimate()` – Returns the initial threshold crossing estimate, based upon linear interpolation

- `virtual double GetFinalThresholdCrossingEstimate()` – Returns the final threshold crossing estimate, based upon linear interpolation

class MGWFSavitzkyGolaySmoother

Performs Savitzky-Golay smoothing on the waveform [59]. Behavior of the transformation can be changed by calling the class function `ResetSmootherAttributes( size_t smoothSize, size_t derivativeOrder, size_t polynomialDegree )` to set the size of the smoothing, the derivative order (0 for no derivative), and the polynomial degree.

class MGWFSmoother

Performs triangular smoothing on a waveform which is essentially a convolution of the waveform with another triangular waveform of defined size. The size of the smoothing can

be changed by calling `SetSmoothSize( size_t aVal )` and region of smoothing (the part of the waveform that is smoothed) can be defined using the function `SetSmoothRegion( const MGWaveformRegion& aVal )`.

`class MGWFStaticWindow`

Performs an estimation of the amplitude of the pulse, it should be used after baseline removal and pole-zero correction. This algorithm is the preferred method for performing an offline estimation of a preamplifier pulse height. Instead of performing an offline convolution, the transformation begins after a delay, averages for the first ramp time, skips ahead for the flat time, averages for the second ramp time and then subtracts the first ramp time from the average of the second ramp time. The integration (ramp) times are independent so that you can essentially have an asymmetric trapezoidal filter. The behavior of this class is therefore defined by the following functions: `SetDelayTime( double aVal )`, `SetFirstRampTime( double aVal )`, `SetSecondRampTime( double aVal )`, `SetFlatTime( double aVal )`. The peak height can be read out using the function `double GetPeakHeight()`.

`class MGWFStaticWindowCusp`

This class performs essentially the same function as the `MGWFStaticWindow`, but instead of performing a straight average, it applies a truncated cusp-like weighting function to the average:

$$\sum_n W[n]f[n]$$

This is equivalent to using a truncated cusp filter and can give improved performance in certain noise environments, see e.g. [104]. The behavior is equivalent to the `MGWFStaticWindow` class, with the addition of functions `SetFirstRampTruncation( double truncate )` and `SetSecondRampTruncation( double truncate )` used to define the truncation of the cusp. The weighting function used for a particular ramp time, r, and a truncation, t, is given by:

$$W[n] = \exp(-nt/r)$$

where $0 \leq n \leq r$ denotes a sample in the waveform. It is clear that as $t \to 0$, this becomes equivalent to a straight average.

B.1.5 Usage Examples

Some examples of possible usage for these transformations are given with particular emphasis on types of analyses performed in this dissertation.

Listing B.1: Calculating amplitude of pulse in C++

```cpp
1  MGWFBaselineRemover baseline;
2  MGWFStaticWindow staticWindow;
3  MGWFPoleZeroCorrection pz;
4  double tauUsed = 50e3;
5  ...
6  // Setting baseline average time
7  baseline.SetBaselineTime( baselineAverageTime );
8  // Taking advantage of the fact that the baseline has been removed
9  // Don't have to do it with the staticWindow *and* the baseline tool.
10 // Setup staticWindow
11 staticWindow.SetFirstRampTime( 0 );
12 staticWindow.SetDelayTime( baselineAverageTime );
13 staticWindow.SetSecondRampTime( secondRiseTime );
14 staticWindow.SetFlatTime( flatTime );
15 // Setup up pole-zero correction
16 pz.SetDecayConstant(tauUsed);
17 // Perform transforms, all in-line (results saved back in wf)
18 baseline.Transform( wf );
19 pz.Transform( wf );
20 staticWindow.Transform( wf );
21 // Grab energy (amplitude)
22 double energy = staticWindow.GetPeakHeight()/(tauUsed*wf->
        GetSamplingFrequency());
```

Listing B.2: Calculating general waveform characteristics in Python

```python
1  # Baseline Transformer
2  baseline = ROOT.MGWFBaselineRemover()
3  init_baseline_time = 280e3
4  # Extremum Transformer
5  extremum = ROOT.MGWFExtremumFinder()
6  # Bandpass transformer, this is used to smooth
```

```python
7  # the shaped waveforms before energy estimation
8  first_bandpass = ROOT.MGWFBandpassFilter()
9  shaped_bandpass = 0.0001 # 100 kHz low-bandpass
10 first_bandpass.SetUpperBandpass(shaped_bandpass)
11 ...
12 for wf in all_waveforms:
13     # All channels
14     for chan_num in (0,1,2,4,5):
15         baseline.SetBaselineTime(init_baseline_time) # 250 mus
16         wf = event.GetWaveform(chan_num)
17         extremum.SetFindMaximum(True)
18         extremum.Transform(wf)
19
20         # Find parameter of waveform, max, min, etc.
21         max_value = extremum.GetTheExtremumValue()
22         avg_value = max_value
23         extremum.SetFindMaximum(False)
24         extremum.Transform(wf)
25         min_value = extremum.GetTheExtremumValue()
26         baseline_factor = 1
27
28         # only process the shaped waveform with a bandpass filter
29         if chan_num in (0,1,2):
30             first_bandpass.Transform(wf)
31             extremum.SetFindMaximum(True)
32             extremum.Transform(wf)
33             avg_value = extremum.GetTheExtremumValue()/wf.GetLength()
34             baseline_factor = wf.GetLength()
35         baseline_value = baseline.GetBaseline(wf)/baseline_factor
```

These limited examples only touch on a very small set of the functionality of MGDO Transforms, further examples are available at git repositories located at http://github.com/mgmarino.

B.1.6   Application of MGDO Package to the Modified-BEGe Analysis

This section outlines how the analysis of the tier1 (rootified raw data) files was performed using the MGDO package, demonstrating how tier2 files were generated from the tier1 data. In particular, this is a detailed description of how the analysis of the waveforms recorded in the DAQ was performed, focusing on *how* characteristics of the waveforms (e.g. baseline, rise-time, etc.) were extracted. In addition, several analysis 'container classes' were created to hold the data generated by these calculations and these are described in the first section.

Modified-BEGe Analysis Container Classes

Several classes were created to encapsulate data into ROOT files for further processing. In addition to wrapping the data to simplify its storage, several utility functions were provided by the classes to return information. All the classes described here can be found in a git repository at `http://github.com/mgmarino/BeGeAnalysisClasses`.

`class MGMBeGeChannelInfo`

Provides classes to encapsulate basic information about the waveform.

```
1   class MGMBeGeOneChannelInfo
2   {
3     /* Class to encapsulate a single channel of the data. */
4     public:
5       MGMBeGeOneChannelInfo() : baseline(0), maximum(0), minimum(0),
             averagepeak(0) {}
6       MGMBeGeOneChannelInfo(Double_t aBase, Double_t aMax, Double_t aMin,
             Double_t aPeak) :
7        baseline(aBase), maximum(aMax), minimum(aMin), averagepeak(aPeak) {}
8       Double_t baseline;    // Waveform baseline
9       Double_t maximum;   // Waveform max
10      Double_t minimum;     // Waveform min
11      Double_t averagepeak;   // Waveform average at its peak (relevant for
             shaped pulses)
12    // For ROOT dictionary generation
13    ClassDef(MGMBeGeOneChannelInfo,2)
```

```
14 };
15
16 class MGMBeGeChannelInfo: public TObject
17 {
18   /* Class to encapsulate multiple channels, with utility functions for
19       access during TTree::Draw commands.  */
20   public:
21     std::vector<MGMBeGeOneChannelInfo> channels;
22     MGMBeGeOneChannelInfo& GetChannel(size_t i) { return channels[i]; }
23     size_t GetNumChannels() { return channels.size(); }
24   // For ROOT dictionary generation
25   ClassDef(MGMBeGeChannelInfo,1)
26 };
```

class MGMRisetimeInfo

Provides classes to encapsulate rise-time information about the waveform.

```
1 class MGMRisetimeOneChannelInfo
2 {
3   // Class for encapsulation of rise-time information for a single channel
4   public:
5     MGMRisetimeOneChannelInfo() : start(0), stop(0),
6                                   risetime(0),
7                                   maximum(0), minimum(0),
8                                   max_point(0), min_point(0) {}
9     MGMRisetimeOneChannelInfo(Double_t aStart, Double_t aStop,
10                               Double_t aRT,
11                               Double_t amax, Double_t amin,
12                               UInt_t max_pt, UInt_t min_pt) :
13       start(aStart), stop(aStop),
14       risetime(aRT),
15       maximum(amax), minimum(amin),
16       max_point(max_pt), min_point(min_pt) {}
17
18   public:
19     Double_t start;   // Start of rt, in units of time
```

```
20      Double_t stop;       // End of rt, in units of time
21      Double_t risetime;     // Total risetime
22      Double_t maximum; // maximum calculated in de-noised pulse
23      Double_t minimum; // minimum, as above
24      UInt_t max_point;    // position of maximum
25      UInt_t min_point;    // position of minimum
26    // ROOT Dictionary generation
27    ClassDef(MGMRisetimeOneChannelInfo,3)
28  };
29
30  class MGMRisetimeInfo: public TObject
31  {
32    // Encapsulates all the channels
33    public:
34      std::vector<MGMRisetimeOneChannelInfo> channels;
35      MGMRisetimeOneChannelInfo& GetChannel(size_t i) { return channels[i]; }
36      size_t GetNumChannels() { return channels.size(); }
37    // ROOT Dictionary generation
38    ClassDef(MGMRisetimeInfo,1)
39  };
```

```
class MGMMuonVeto
```

Provides classes to encapsulate muon-veto information about an event. This is useful, because the muon veto logic pulses are much shorter (10 μs) than the total length of the waveform (400 μs) and can fire several times in parts of the waveform (see Figure B.1). In particular, it exports several convenience functions for determining whether or not a muon veto has fired at a particular point in the waveform. For example, `IsInVetoRegion(size_t pos)` will return true or false depending upon if `pos` falls during a veto fire.

```
1  class MGMMuonVeto: public TObject
2  {
3    public:
4      std::vector<MGWaveformRegion> regions;
5      bool IsInVetoRegion(size_t);
6      bool RangeIsInVetoRegion(size_t beginning, size_t end);
```
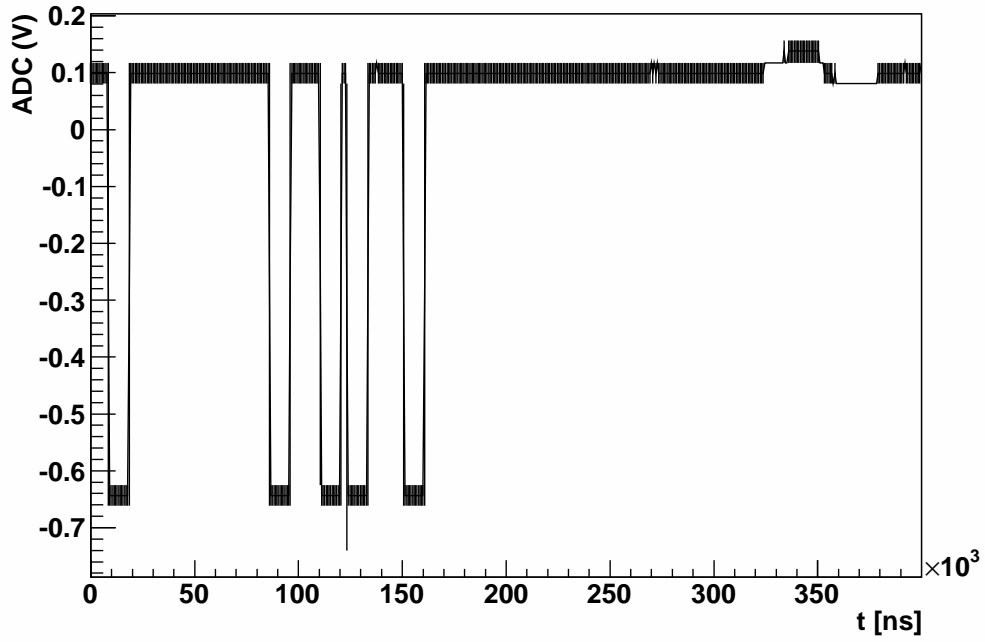
Figure B.1: Example of the digitized muon veto channel, the low signal indicates a veto fire.

```
 7    public:
 8      size_t GetNumberOfRegions() { return regions.size(); }
 9
10    ClassDef(MGMMuonVeto,1)
11  };
```

Modified-BEGe Tier0→Tier1 Analysis

The analysis script for the modified-BEGe is unfortunately too long to reproduce here, but is freely available online at: `http://github.com/mgmarino/BeGeAnalysisClasses/blob/master/BEGeAnalyzeWaveforms/analyze_waveforms.py`. This script made extensive use of MGDO classes as well as the container classes described above and is provided as a reference for using MGDO in production.

B.2   Analysis Database and Processing Framework

The analysis database was set up to facilitate the movement and analysis of data from beginning of processing (data acquisition) to the end (exclusion plots, physics analyses, etc.). The database ran using a CouchDB backend, a document-based database with several desirable features, including scalability using the 'view' Map-Reduce functionality and simple backup and mirror propagation through built-in replication. Replication allows quick and easy distribution of data to several servers, very desirable for collaborations wishing to share or mirror data in multiple locations. As well, the CouchDB server employs a RESTful interface, accepting basic http commands to access and manipulate data and distributing and storing data in a JSON (JavaScript Object Notation) format. This interface simplifies communication with the database server and has fostered the development of CouchDB 'wrappings' in several languages including Ruby and Python. More information on this database can be found online at `http://couchdb.apache.org/`. This section will describe the basic functionality of the run/analysis database as well as describe particular details of its implementation.

B.2.1   Overview

The central CouchDB server for the run/analysis database was located on a machine at the University of Washington and backed up to several other instances including personal laptops and online at CloudAnt (`https://cloudant.com/`). The code written to encapsulate and read out/run the database was written in Python using an older version of couchdb-python (available `http://code.google.com/p/couchdb-python/`). Other software with CouchDB Python 'bindings' is available including couchdbkit (`http://couchdbkit.org/`), but the former was initially chosen for arbitrary reasons. The functionality of both frameworks is similar and so a translation from one to the other should not entail a significant amount of effort. Three databases were run with this code, one for each of three different detectors: P-PC1, P-PC2, and the low-background modified-BEGe detector. P-PC2 and the modified-BEGe are described previously in this dissertation. The code is freely available in a browsable git repository online at: `http://github.com/mgmarino/SoudanDB`.

> `management` – sub-module for management objects
>
> `utilities` – sub-module for utility functions
>
> `update` – sub-module housing objects for updating the database
>
> `views` – sub-module housing objects and base classes for database views
>
> `databases` – sub-module housing database objects
>
> > `bege_gretina` – Database for Gretina DAQ readout of the modified BEGe
> >
> > `bege_jc` – Database for NI DAQ readout of the modified BEGe (see Section 3.1)
> >
> > `ppctwo` – Database for test stand readout of P-PC1 (not described in this dissertation

The `database` directories each had the following sub-structure with relevant code for the particular database:

> `db.py` – module providing particular code for the database
>
> `views` – module providing separate Python files for each view
>
> `update` – module providing separate Python files for each update function

Instead of describing the code in depth, the following sections will focus on key functional aspects of the database as they pertain to the creation of future databases.

B.2.2 Functional Overview

The purpose of the run/analysis database was to track run information and data and manage the analysis process. A schematic of this process is shown in Figure B.2. Metadata (e.g. names of run files, time information, parameter settings, etc.) generated by the DAQ system or slow controls is automatically inserted into the database as a record when a new run file is generated or run settings are changed. Once this information appears in the database, a process management class can then determine what types of processing needs to happen to the metadata record and then execute the necessary tasks to complete these. In the schematic in Figure B.2, the process management class queries the run database to determine which records have associated data that still requires back up or analysis processing and then calls the associated program(s) to complete this. This methodology can scale well to larger numbers of data records since processing (e.g. cpu usage) only occurs when needed.

Another advantage of using a CouchDB database is that it is possible to insert records with different internal structure or different metadata. For example, a single database can include records of associated run metadata as well as other records with slow control information. Additionally, metadata need not be consistent across records of a certain type. This has the advantage that it is not required to fully define the data structure within a database during initial deployment, therefore allowing simple adaptability to any changes that will inevitably happen to the data-taking process such as adding a new channel or additional sets of information.

B.2.3 Functional Implementation

Implementing the functionality described in the previous section involved taking advantage of several features built into the CouchDB database. The database update could occur via several methods, for example: a daemon which runs regularly to insert the metadata when
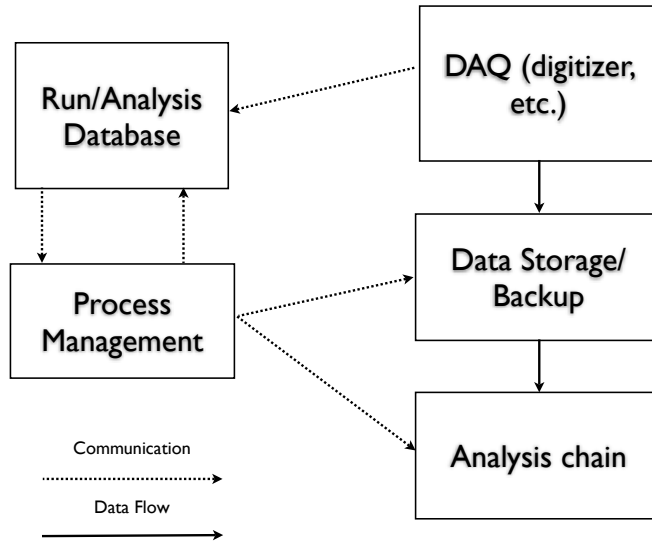
Figure B.2: Schematic of the run/analysis database functionality.

new files or information from the DAQ are generated; a system interrupt process calling an insertion function or program when a new file appears in a certain directory; active insertion (direct communication) by the DAQ software into the database. This software employed the first method, regularly running a daemon via a cron job to determine when a new file was written into a certain directory by the DAQ software and inserting the necessary metadata as a record into the database.

The initial insertion of metadata placed a skeleton record in the database, populating the fields of the record with available data and leaving other fields that required information from further processing empty. The data resided in the database in JSON format, and an example of a portion of an inserted JSON record in the run/analysis database is shown in Listing B.3. In this example, one can see the left-empty fields denoted by `null`. It is simple to write a 'view' (a CouchDB query) to return records where an empty field might exist. This method can be used to select records for processing. For example, in the record snippet, the `root_data_file_tier_1` has daughter values set to `null` including `last_mod_time` (last modification time) and `md5hash` (calculated md5 hash) indicating that this particular

file in the analysis chain has yet to be created. The process management daemon could then run the proper program or function to generate this file and populate those fields. Other fields with `null` also require some further processing.

```
1   {
2       "_id": "20100713120006",
3       "_rev": "6-8ba079fbf9a7196c83b14aaf9976678b",
4       "modification_time": null,
5       "raw_data_file_tier_0": {
6           "last_mod_time": "2010-07-13T12:22:07Z",
7           "lfn": "tier0/100713120006",
8           "md5hash": "c669edfa8e5a671b0a7c4e84932decb1",
9           "pfn": "/data/Soudan/Data/BeGe/tier0/100713120006"
10      },
11      "number_of_entries_in_tier1_root_tree": null,
12      "quality_assurance": {
13          "qa_check_process_has_been_run": null,
14          "qa_accept_run": null
15      },
16      "livetime": {
17          "run_milliseconds": 9046984,
18          "run_milliseconds_error": 0
19      },
20      "local_time_of_start_of_run": "2010-07-13T12:00:06Z",
21      "root_data_file_tier_1": {
22          "last_mod_time": null,
23          "lfn": "tier1/100713120006_rootified.root",
24          "md5hash": null,
25          "pfn": "/data/Soudan/Data/BeGe/tier1/100713120006_rootified.root"
26      },
27      "run_settings": [
28          "8000",
29          "20000000",
30          "20"
31      ],
32      ...
```

```
33  }
```

Listing B.3: Truncated example of a JSON record in the run/analysis database.

In the database sub-modules, the framework associates 'views' with processing functions in the 'update' module. This allows updates to be performed on records returned by particular views. Each update module (e.g. `SoudanDB.databases.bege_jc.update.update_number_of_entries_in_root_file`) exports two functions: `get_view()` and `update_rundoc(run_doc)`. Therefore, a daemon management program would use the `get_view()` to return the records needing updates and then run `update_rundoc()` for each `run_doc`. The management process for doing this is located in `SoudanDB.update.update_calculations_on_database`. Newer versions of CouchDB should allow such processing to occur within the view server framework of the database itself. For example, the CouchDB server should eventually be able to handle both generating the views and running the update functions on those views. Therefore, future implementations should not use this daemon methodology and instead should migrate to using a Python view server.

B.3 WIMP PDFs Software Framework

This section outlines and describes the functionality of the pyWIMP software package for generating fits and limits to dark matter interactions such as a WIMP-nuclear recoil or the axioelectric effect. The software is freely available in a git repository online at `http://github.com/mgmarino/pyWIMP`. The software is written as a Python/C++ hybrid package, essentially using Python for process control and management and using compiled C++ code where speed is critical. As with most work performed for a thesis, some significant effort was placed in the initial design and development of this package, but as deadlines approached and the pressure to make it just *work* increased, some code not generally useful necessarily appeared. Because of this, the following sections will discuss the code that *is* generally useful in the hopes that its development will save the same effort from others.

B.3.1 Overview

The pyWIMP package makes extensive use of the RooFit fitting package [51] – developed for BaBar and now included in the ROOT analysis package [35]. The basic structure of pyWIMP is as follows:

`buildTools/` – directory for building utilities

`test_scripts/` – directory containing test scripts useful for checking for proper installation

`pyWIMP/` – base directory of Python package

> `Calculation` – sub-module containing Python calculation classes for deriving sensitivities and limits
>
> `DMModel` – sub-module containing Python objects for the construction of models used in fitting dark matter signals
>
> `WIMPPdfs` – sub-module containing compiled C++ classes used to construct dark matter signal models
>
> `utilities` – utilities sub-module

Since the most generically useful tools exist in the WIMPPdfs and DMModels sub-modules, the following sections will detail these.

B.3.2 `pyWIMP.WIMPPdfs`

WIMPPdfs contains C++-constructed objects of probability distribution functions and related classes. Almost all of the classes derive from `RooAbsPdf`, except for classes used for Form Factor calculation which derive from `MGMVWimpFormFactor`. All classes are described in detail here, including their functionality and primary constructor.

`class MGMVWimpFormFactor`

Base class for all classes describing a form factor. This class has a standard constructor, but should never need to be directly instantiated by a user:

```
MGMVWimpFormFactor(const char *name = "", const char *title = "")
```

`class MGMWimpHelmFFSquared`

Class describing a Helm nuclear form factor squared [71], the default constructor is:

```
MGMWimpHelmFFSquared(const char *name, const char *title,
  RooAbsReal& _q,
  RooAbsReal& _r_sub_n,
  RooAbsReal& _s)
```

where `_q` is the momentum transfer, `_r_sub_n` is the effective nuclear radius and `_s` is the skin thickness.

`class MGMBetaDecayFunction`

Class describing a generic beta decay. This class does not take into account any correction factors and instead gives the kinematic spectrum for the electron given the endpoint energy of the decay. The default constructor is

```
MGMBetaDecayFunction(const char *name, const char *title,
  RooAbsReal& _energy,
  RooAbsReal& _mass_of_electron,
  RooAbsReal& _qvalue)
```

where `_enery` is the energy of the electron, `_mass_of_electron` is the mass of the electron and `_qvalue` is the q-value of the decay, all in keV.

`class MGMErfcFunction`

Class implementing an error function pdf of the form: $\frac{1}{2}\operatorname{erf}((E-\mu)/(\sigma\sqrt{2})) + \rho$ with ρ the offset. The constructor:

```
MGMErfcFunction(const char *name, const char *title,
  RooAbsReal& _energy,
  RooAbsReal& _mean,
  RooAbsReal& _sigma,
  RooAbsReal& _offset)
```

```
class MGMWimpDiffRateBasicPdf
```

Class implementing the basic form of the WIMP interaction with no time dependence [18]:

$0.751 R_0/(E_0 r) \exp(-0.561 \frac{Q}{E_0 r})$

```
MGMWimpDiffRateBasicPdf(const char *name, const char *title,
  RooAbsReal& _R_sub_0,
  RooAbsReal& _E_sub_0,
  RooAbsReal& _Q,
  RooAbsReal& _r,
  MGMVWimpFormFactor& _form_factor = MGMVWimpFormFactor::DefaultFormFactor()
     )
```

```
class MGMWimpDiffRatePdf
```

Class implementing the time-dependent WIMP interaction assuming infinite escape velocity (see Section 4.1 and [17, 18]), where v_E is the velocity of the earth with possible time dependence. This class takes advantage of RooFit functionality which does not require parameters to have explicit dependencies. That is, this class does not need to know about any time dependence of v_E.

```
MGMWimpDiffRatePdf(const char *name, const char *title,
  RooAbsReal& _v_sub_0,
  RooAbsReal& _v_sub_min,
  RooAbsReal& _v_sub_E,
  RooAbsReal& _R_sub_0,
  RooAbsReal& _E_sub_0,
  RooAbsReal& _r,
  MGMVWimpFormFactor& _form_factor = MGMVWimpFormFactor::DefaultFormFactor()
     )
```

```
class MGMWimpDiffRateEscapeVelPdf
```

Class implementing the time-dependent WIMP interaction with the finite escape velocity of the dark matter halo included. This class is equivalent to the above class with the addition

of the escape velocity parameter, v_{esc}. See Section 4.1 for details on the functional form of this pdf.

```
MGMWimpDiffRateEscapeVelPdf(const char *name, const char *title,
  RooAbsReal& _v_sub_0,
  RooAbsReal& _v_sub_min,
  RooAbsReal& _v_sub_E,
  RooAbsReal& _R_sub_0,
  RooAbsReal& _E_sub_0,
  RooAbsReal& _r,
  RooAbsReal& _v_sub_esc,
  MGMVWimpFormFactor& _form_factor = MGMVWimpFormFactor::DefaultFormFactor()
      );
```

```
class MGMWimpTimeFunction
```

Class implementing a generic time pdf for a WIMP search with the functional form: $v_0 + v_1 \sin 2\pi t$. This is useful for deriving sensitivity or looking for limits on a generic oscillation signal.

```
MGMWimpTimeFunction(const char *name, const char *title,
  RooAbsReal& _velocity_0,
  RooAbsReal& _velocity_1,
  RooAbsReal& _time);
```

B.3.3 `pyWIMP.DMModels`

DMModels provides a sub-module of Python objects that return generically useful PDFs. The purpose of this module is to provide simple objects with interfaces to derive PDFs for sensitivity and limit calculations. Base classes are defined in `DMModels.base_model`, which are used by other Python objects in this sub-module. All of the model classes have a `get_model()` routine which returns a RooAbsPdf. Some model classes may return other PDFs; those are indicated where applicable.

`DMModels.base_model`

Defines `DMModels.base_model.BaseVariables` and `DMModels.base_model.BaseModel`. The first provides an encapsulation class for the observables time (in years) and energy (in keV) and the second provides a base class for other objects in this submodule. `DMModels.base_model.BaseVariables` can be instantiated by:

```
base_variabes =   DMModels.base_model.BaseVariables(
                  time_beginning,    # Define the beginning of time obs
                  time_in_years,     # Define the end of time observation
                  energy_threshold,  # Energy threshold (keV)
                  energy_max,        # Maximum energy range (keV)
                  use_tag = True     # Use a tag, useful when using
                                     # multiple instances of this class.

                  )
```

`DMModels.beta_decay_model`

Defines `DMModels.beta_decay_model.BetaDecalModel` which generates the kinematic electron spectrum of a beta decay. Also, if the lifetime of the decay is included, the returned PDF will be 2-D in energy and time. In can be instiated by

```
beta_decay =      DMModels.beta_decay_model.BetaDecalModel(
                  base_variabes,     # Base variables (BaseVariables class),
                  q_value,           # Q value of the beta decay (float)
                  lifetime = None    # Lifetime (in years, float), if None
                                     # then the returned PDF is flat in time

                  )
beta_decay_model = beta_decay.get_model() # Return the RooAbsPdf
```

`DMModels.tritium_decay_model`

Defines `DMModels.tritium_decay_model.TritiumDecayModel` which generates the kinematic decay of tritium. This class returns a 2-D tritium decay spectrum with an additional flat background. It returns an extended model with the relative amounts are set correctly (but allowed to float) This means that if you are doing toy model fitting with the returned model

before to generate the events (RooFit will automatically generate the correct number of events from an extended pdf) and SAVE all the variables.

```
tritium_decay =   DMModels.tritium_decay_model.TritiumDecayModel(
                  base_variables,
                  tritium_exposure,     # Tritium Exposure time (days)
                  tritium_activation,   # Tritium activation rate
                                        # in atoms/kg/day
                  mass_of_detector,     # Mass of the detector (kg)
                  flat_background_rate  # Flat background rate,
                                        # in counts/kg/keV/day
                  )
tritium_model = tritium_decay.get_model() # Return RooExtendPdf
```

DMModels.flat_model

Defines `DMModels.flat_decay_model.FlatModel` which generates a pdf flat in both time and energy.

```
flat =            DMModels.falt_decay_model.FlatModel(
                  base_variables)
flat_model = flat.get_model() # Return RooAbsPdf
```

DMModels.gamma_line_model

Defines `DMModels.gamma_line_model.GammaLineModel` and `DMModels.gamma_line_model.GammaLineFactory`. The first defines a basic object to handle a gamma (or x-ray line) at a certain energy and a defined sigma. The second is a convenience class to limit the generation of code when implementing a gamma line.

```
gamma =           DMModels.gamma_line_model.GammaLineModel(
                  base_variables,
                  mean,              # Mean, RooAbsReal
                  sigma,             # Sigma, RooAbsReal
                  lifetime = None    # Lifetime, float, if None,
                                     # return PDF is 1-D
                  )
```

```
gamma_model = gamma.get_model()
```

The factory includes two convenience functions to quickly return a gamma line:

```
gamma =           DMModels.gamma_line_model.GammaLineFactory.generate(
                  mean,              # Mean, float
                  lifetime,          # Lifetime, float
                  base_variables)
gamma_model = gamma.get_model()
# ... OR alternatively explicitly define errors ...
gamma =           DMModels.gamma_line_model.GammaLineFactory.generate(
                  mean,              # Mean, float
                  mean_error,        # Mean error, float
                  sigma,             # Sigma, float
                  sigma_error,       # Sigma error, float
                  lifetime,          # Lifetime, float
                  base_variables)
gamma_model = gamma.get_model()
```

```
DMModels.wimp_model
```

Defines `DMModels.wimp_model.WIMPModel` which is a class wrapping several different WIMP
PDFs. At the time of writing, the class was tuned to work with germanium detectors
as it included quenching and atomic information specific to a germanium detector. Modify-
ing the class to be applicable to other detectors should not be difficult. The models that this
class returns require a normalization factor, which is the *nuclear* cross section of the WIMP
interaction (i.e. not normalized per nucleon). Therefore, all classes should use a `RooExtendPdf`
to multiply the cross-section by the obtained `RooAbsPdf`. The class is instantiated:

```
wimp_class =      DMModels.wimp_model.WIMPModel(
                  base_variables,
                  mass_of_wimp,      # Mass of the WIMP (GeV)
                  kilograms,         # Mass of the detector (kg)
                  constant_quenching # If true, use constant, non-energy
                                     # dependent quenching
                                     # Else use energy-dependent quenching.
                  )
```

```python
# Get normalization: multiply cross-section by this to get
# per-nucleon xs (in pb ) (returns RooRealVar)
wimp_nucleon_norm = wimp_class.get_normalization()


# Get simple (only 1-D exponential) model (returns RooAbsPdf)
wimp_simple = wimp_class.get_simple_model()


# Get different form factors (returns RooAbsPdf)
helm_ff = wimp_class.get_helm_form_factor()
exponential_ff = wimp_class.get_exponential_form_factor()


# Get different 2-D WIMP models.   The following allow
# changing the value of the WIMP mass by passing
# in an (optional) float parameter

wimp_model_no_escape_velocity = \
    wimp_class.get_WIMP_model(20) # Change mass to 20
# WIMP Model, escape velocity, form factor = 1
wimp_model_esc_vel_no_ff = \
    wimp_class.get_WIMP_model_with_escape_vel_no_ff()


# WIMP Model, escape velocity, Helm form factor
wimp_model = wimp_class.get_WIMP_model_with_escape_vel()
# ... OR equivalently ...
wimp_model = wimp_class.get_model()
```

B.3.4 Calculating Sensitivities

The `pyWIMP` package includes a utility executable, `job_engine.py`, that can distribute many different calculations across several CPUs. (There are also several examples of using this software package with MPI using the Python mpi4py distribution. These can be found in e.g. in `test_Rolke.py`, `mpi_job_engine.py`, but the usage of this is not detailed here.) This executable may be called with several different options, which are presented here for reference:

```
> python /path/to/pyWIMP/job_engine.py WIMPModel --help
Usage: job_engine.py model [options]
```

```
Options:

  -h, --help              show this help message and exit
  --total_time=TOTAL_TIME
                          Total time (year)
  --print_out_plots       Print plot results of fit to eps format.   This will
                          cause the program to continue instead of stopping
                            once
                          a plot is made. Batch mode will be set so that no
                            plot
                          will be displayed during the program.
  --tritium_activation_rate=TRITIUM_ACTIVATION_RATE
                          Tritium activation rate (counts/kg/day)
  --constant_energy       Set energy as constant
  --do_axioelectric       Do axioelectric fitting
  --wimp_mass=WIMP_MASS
                          WIMP mass (GeV/c^-2)
  --num_time_bins=NUM_TIME_BINS
                          Number of time bins, 0 means unbinned
  --energy_max=ENERGY_MAX
                          Maximum energy (keV)
  --num_energy_bins=NUM_ENERGY_BINS
                          Number of energy bins, 0 means unbinned
  --debug                 Set debug flag, enables verbose output
  --axion_mass=AXION_MASS
                          Axion mass in keV [default 0]
  --constant_time         Set time as constant
  --variable_quenching  Set to use variable quenching
  --threshold=THRESHOLD
                          Threshold (keV)
  --tritium_exposure_time=TRITIUM_EXPOSURE_TIME
                          Tritium exposure time days
  --confidence_level=CONFIDENCE_LEVEL
                          Confidence level (0 -> 1)
  --show_plots            Show plot results of fit
```

```
--background_rate=BACKGROUND_RATE
                    Background rate (counts/keV/kg/day)
--mass_of_detector=MASS_OF_DETECTOR
                    Mass of detector (kg)
-o OUTPUT_FILE, --output_file=OUTPUT_FILE
                    Define the output file name (full path)
-n NUMPROCESSORS, --num_cpus=NUMPROCESSORS
                    Define the number of cpus used
-a MAX_TIME, --max_time=MAX_TIME
                    Set the max time [seconds] until this program shuts
                    down
-i NUM_ITER, --num_iter=NUM_ITER
                    Number of iterations per cpu
```

For example, to perform a sensitivity calculation on a WIMP mass of 20 GeV, with a MAJORANA-like detector of 20 kg over 5 years, one could execute:

```
/path/to/pyWIMP/job_engine.py WIMPModel   \
  -o output_root_file.root                \
  -n 8                                    \
  -i 1000                                 \
  --mass_of_detector=20                   \
  --wimp_mass=20                          \
  --tritium_exposure_time=30              \
  --tritium_activation_rate=200           \
  --total_time=5                          \
  --background_rate=0.001                 \
  --energy_max=20                         \
  --threshold=0.5                         \
  --confidence_level=0.9                  \
```

which would distribute the workload across 8 CPUs. Similarly, to calculate the sensitivity to an axioelectric effect at an axion mass of 3 keV, one might use the additional flags:

```
  --do_axioelectric                       \
  --axion_mass=3                          \
```

and drop the flag `--wimp_mass`.

Currently, limit calculations on data can also be performed using `'job_engine.py DataExclusion'`, but this is currently limited to a background model consistent with that from the modified-BEGe detector described in Section 4.4.1. Therefore, any application of this framework to other data sets would require the construction of a background model to suitably describe that dataset. It is the hope that this toolkit together with the tools already available in RooFit would provide everything necessary for building such a model.

Appendix C

DEVELOPMENT OF DAQ HARDWARE FOR Majorana

C.1 Tundra Universe IID PCI-VME Bridge Driver

This section describes a Linux driver for PCI-VME Tundra Universe II bridge chip. This
driver is used in Linux kernels newer than 2.6.20 and has been used extensively with the x86-
based lightweight CRUX distribution (`http://www.crux.nu`) as well as Fedora Core (`http://fedoraproject.org/`). The driver is available at a git repository `http://github.com/mgmarino/VMELinux` as well as in the ORCA subversion repository `svn://orca.physics.unc.edu/Drivers/SBCDrivers`. Installation instructions are given in the code package, so
those details are not covered here.

C.1.1 Char Driver

The Linux driver package installs a kernel-space driver and installs several devices in `/dev`,
including:

```
/dev/vme_m[0-7]
/dev/vme_ctl
/dev/vme_dma
```

for the 8 slave image windows of the Universe II chip, a control device, and a DMA transfer
device. The devices can be accessed directly using the Linux `ioctl`, `read`, `write` system
calls, though the simpler method for interaction is via the exported API, described later
(see Section C.1.2). The possible ioctl calls are:

```
UNIVERSE_IOCSET_CTL         // Set CTL of device, but device specific
                            // i.e. DMA sets CTL of DMA device, etc.
                            // See code for more details
UNIVERSE_IOCSET_BS          // Set base offset of image
UNIVERSE_IOCSET_BD          // Set bound (top) of image
UNIVERSE_IOCSET_VME         // Set VME address offset
```

```
UNIVERSE_IOCSET_IOREMAP      // Set to use ioremap (instead of mmap)
UNIVERSE_IOCCHECK_BUS_ERROR // Check for an error on the VME bus

// The following are for the control device ONLY:
UNIVERSE_IOCSET_HW_BYTESWAP // Set the HW byteswap (board specific
                            // to Concurrent Technologies boards)
UNIVERSE_IOCGET_MEM_SIZE    // Get size of PCI memory set aside by driver
UNIVERSE_IOCGET_BOARD_TYPE  // Return board type
UNIVERSE_IOCIO_PORT_READ    // Read the IO Port specified by input address
UNIVERSE_IOCIO_PORT_WRITE   // Write the IO Port at input address
```

The initialization procedures for each image should be followed:

```
For a normal image minor, initialize it by (in this order):
    UNIVERSE_IOCSET_CTL: Set control register (address space, data space,
        etc.)
    UNIVERSE_IOCSET_IOREMAP: set to ioremap pci mem to kernel memory.
    UNIVERSE_IOCSET_BS: Set base (offset) of window.   The driver will
        automatically determine where this is within its allowed PCI space.
    UNIVERSE_IOCSET_BD: Set bound (size) of window.
    UNIVERSE_IOCSET_VME: Use this to set the desired VME address the base
        of the window will point to.


For a DMA minor, initialize it by (in this order):
    UNIVERSE_IOCSET_CTL: Set control register (address space, data space,
        etc.)
    UNIVERSE_IOCSET_VME: Use this to set the desired VME address from which
        the
         DMA will start.
```

The CTL settings for each image device can be found in the Tundra Universe II manual. However, it is much simpler to use the API since this abstracts the behaviour of the driver.

C.1.2 API

The Universe driver API provides a C and C++ interface for programming simplicity. The C interface is simpler, providing a functional-based interface and will be described first. To use both, add

```
#include "universe_api.h"
```

to your code and make sure the API library is in your search path of your compiler, e.g.:

```
g++ myprog.c -o myprog -L/use/local/universe/lib -luniverse_api
```

C API

The C API is outlined in the following. If an error occurs during the function, the specified error is set in the `errno` global.

- ```
  extern TUVMEDevice* get_new_device(uint32_t vmeAddress, uint32_t
      addressModifier, uint32_t dataWidth, uint32_t sizeOfImage = 0);
  ```

  This function grabs a new device with the given specifications. It will return NULL if there is any error. An error can be caused if there are no more available devices, or if something is wrong with the input parameters. If `sizeOfImage` is 0, then the driver attempts to make the size of the returned device 1/8 of the total PCI memory space allocated to the Universe II chip. The size of the image specifies how many addresses from the base vmeAddress can be read/written. vmeAddress should be normally 64K aligned (0x10000, with the bottom 16 bits always 0), but there are two devices which have 4K resolution (0x100 aligned). The driver API will attempt to return one of these two devices when the `addressModifier` specifies A16, but it is not guaranteed. `dataWidth` specifies the width of the data in bytes.

- ```
  extern int32_t close_device(TUVMEDevice* device);
  ```

  Closes a device and releases it back into the available pool. This should be called if a device is no longer being used

- ```
  extern TUVMEDevice* get_dma_device(uint32_t vmeAddress, uint32_t
      addressModifier, uint32_t dataWidth);
  ```

  Grabs the DMA device and sets up the transfer. If NULL, this means that DMA device is busy. A transfer from the DMA is initiated with the `read_device` function.

- ```
  extern TUVMEDevice* get_ctl_device();
  ```

Grabs the control device. If NULL, this means that control device is busy.

- ```
  extern void set_dma_no_increment(bool noInc = true);
  ```

This specifies that the dma device should not increment a VME address. This is useful if a dma read of x bytes is required at one particular address. It should be called before `get_dma_device`.

- ```
  extern void set_hw_byte_swap(bool doSwap = true);
  ```

Sets byte swap in the hardware. This currently only works on the VX 40x/04x Concurrent technologies cpu boards and has undefined behavior for other boards.

- ```
  extern int32_t read_device(TUVMEDevice*, char* buffer, uint32_t
      numBytes, uint32_t offset = 0);
  ```

Reads `numBytes` bytes from a device into a buffer at an offset on the device and returns number of bytes read, or less than 0 if error occurs.

- ```
  extern int32_t write_device(TUVMEDevice*, char* buffer, uint32_t
      numBytes, uint32_t offset = 0);
  ```

Writes `numBytes` bytes into a device from a buffer at an offset on the device and returns number of bytes written, or less than 0 if error.

- ```
  extern uint32_t get_max_size_of_image(void);
  ```

Returns the maximum size of an image in bytes.

Some examples of using the functionality are in the following. Only the 'read' variants of the functions are shown, but the 'write' functions behave correspondingly.

1. Grab a device and read from it:

```
TUVMEDevice* device = get_new_device(0x0, 0x29, 2, 0x10000); //Map
    the entire A16 space beginning at 0x0
char buffer[4];
if (device == NULL) exit; // Error!
// read at address 0x3300
if (read_device(device, buffer, 4, 0x3300) < 0 ) {
  // Error!
  exit;
}
// Otherwise we have a successful read.


close_device(device);
/* This last call is an unnecessary call since the API library
    handles the closing automatically.  However, if the pool of 8
    devices is already empty, this will release a device to be opened/
    enabled anew via the get_new_device function.  Normally, readout
    code will set up devices at the beginning of the run so it doesn't
     need to be done multiple times. */
```

2. Perform a DMA transfer:

```
// Set this if this DMA does not auto-increment the address.
set_dma_no_increment(true);
// Set up a DMA transfer, A32, D32 BLT, beginning at address 0
    x2101000
TUVMEDevice* device = get_dma_device(0x2101000, 0xB, 4);
if (device == NULL) exit; // Error!
char buffer[4096];
if (read_device(device, buffer, 4096) != 4096) { // DMA transfer 4096
     bytes
  //Error!
}


/* If the setting on the dma device do not change, it is possible to
    not call get_dma_device multiple times, but rather hold on to the
    pointer to the device.  However, the programmer must take care not
     to change the device settings elsewhere in the code. */
```

3. Read/Write from/to the Universe registers

```
TUVMEDevice* device = get_ctl_device(); // Grab the ctl device.
char buffer[4];
// the ctl device can only read/write 4 bytes at a time!
if (read_device(device, buffer, 4, 0x4) < 0 ) {
  // Error
}
```

C++ API

The C API is actually a wrapper for the C++ framework which provides the underlying API control for the Universe chip. In the C++ framework, each slave image corresponds to an object, `TUVMEDevice`. As well there exists a singleton (global) which acts as a manager class for all the objects, `TUVMEDeviceManager`. It is suggested to use this singleton class to manage the open and closed image objects, but it is possible not to. The singleton class will not 'take control' of the images unless the static function `TUVMEDeviceManager* TUVMEDeviceManager ::GetDeviceManager()` is called.

`class TUVMEDevice`

A class encapsulating a PCI-VME image. The interface is defined as:

```
class TUVMEDevice {

  public:
    TUVMEDevice(uint32_t devNumber);
    virtual ~TUVMEDevice();

    /* Constants for the devices */
    enum ETUVMEDeviceEnum {kNumberOfDevices = 8};
    enum ETUVMEDeviceAddressSpace { kA16 = 0,
                                    kA24 = 1,
                                    kA32 = 2,
                                    kCRCSR = 3,
                                    kUser1 = 4,
```

```
                                       kUser2 = 5};
enum ETUVMEDeviceDataWidth { kD8 = 1,
                             kD16 = 2,
                             kD32 = 4,
                             kD64 = 8};
enum ETUVMEDeviceMode { kProgram = 0, kData };
enum ETUVMEDeviceType { kNonPrivileged = 0, kSuper };
/* End constants */

/* Set functions to define behaviour of image. */
inline void SetPCIOffset(uint32_t offset);
inline void SetSizeOfImage(uint32_t sizeOfImage);
inline void SetAddressSpace(ETUVMEDeviceAddressSpace addressSpace);
inline void SetDataWidth(ETUVMEDeviceDataWidth dataWidth)
inline void SetMode(ETUVMEDeviceMode mode);
inline void SetType(ETUVMEDeviceType type);
inline void SetUseBLTs(bool useBLTs);
inline void SetAllowPostedWrites(bool allowPostedWrites);
inline void SetUseIORemap(bool useIORemap);
void SetVMEAddress(uint32_t vmeAddress);
int32_t SetWithAddressModifier(uint32_t addressModifier);
/* End set functions. */

/* Introspection functions */
inline int32_t GetDevNumber() {return fDevNumber;}
inline uint32_t GetVMEAddress() {return fVMEAddress;}
inline uint32_t GetSizeOfImage() {return fSizeOfImage;}

/* Returns a pointer to the image base. */
inline volatile void* GetMappedAddress();


virtual std::string GetDeviceStringName();


/* Check if a bus error has occurred */
int32_t CheckBusError();
```

```
    /* Control functions, Open(), should be called,
       then the behaviour set and then Enable called.*/
       Enable will return 0 if successfull.
    int32_t Open();

    virtual int32_t Enable();

    void Close();


    /* Locking functions for thread safety. */
    virtual int32_t LockDevice() { return pthread_mutex_lock( &fLock ); }

    virtual int32_t UnlockDevice() { return pthread_mutex_unlock( &fLock );
       }


    /* Read/Write functions */
    int32_t Read(char* buffer, uint32_t numBytes, uint32_t offset = 0);

    int32_t Write(char* buffer, uint32_t numBytes, uint32_t offset = 0);


  protected:
    ...
  };
```

All PCI images are memory mapped, meaning that access to the devices occurs by using a volatile pointer. Direct access to this pointer is possible using `volatile void* TUVMEDevice::GetMappedAddress()`. This pointer can be used to access the VME bus, but it is up to the user to determine if an error occurred after access by calling `TUVMEDevice::CheckBusError()`.

`class TUVMEControlDevice`

Defines a control device for the Universe chip and simplifies the calls available to this device.

```
class TUVMEControlDevice: public TUVMEDevice {
  public:
    TUVMEControlDevice();
    virtual ~TUVMEControlDevice();


    /* Constants, cycle speed, board types */
    enum ECycleSpeeds { kNormal  = 0,
                        kFaster  = 1,
```

```
                            kFastest = 2 };
    enum EBoardType {    kUnknown = UNIVERSE_BOARD_TYPE_UNKNOWN ,
                         kCCT     = UNIVERSE_BOARD_TYPE_CCT };


    /* Access functions */
    int ReadIOPortMemory ( uint16_t address );
    int WriteIOPortMemory ( uint16_t address , uint8_t value );
    void SetHWByteSwap (bool doByteSwap = true);
    void SetDSNegationSpeed ( ECycleSpeeds speed = kNormal);
    void SetDSHighTimeBLTs ( ECycleSpeeds speed = kNormal);
    size_t GetPCIMemorySize ();
    EBoardType GetBoardType ();
    ...
  };
```

`class TUVMEDMADevice`

Defines the DMA device. The only additional interface functions are as follows:

```
class TUVMEDMADevice : public TUVMEDevice {
  public :
    TUVMEDMADevice ();
    virtual ~TUVMEDMADevice ();
    // Set device to not increment the address during DMA transfer
    void SetNoIncrement (bool noInc = true);
    ...
  };
```

`class TUVMEDeviceManager`

Defines a manager (singleton) class which is the preferred way to use this API. The interface
of this class is as follows:

```
class TUVMEDeviceManager {
  public :
    static TUVMEDeviceManager* GetDeviceManager ();
    TUVMEDevice* GetDevice ( uint32_t vmeAddress ,
```

```
                              uint32_t addressModifier,
                              uint32_t dataWidth,
                              uint32_t sizeOfImage = 0);
  TUVMEDevice* GetControlDevice();


  /* DMA read device. */
  TUVMEDevice* GetDMADevice(  uint32_t vmeAddress,
                              uint32_t addressModifier,
                              uint32_t dataWidth,
                              bool autoIncrement);
  /* Close the various devices */
  void ReleaseDMADevice();
  int32_t CloseDevice(TUVMEDevice* device);


  inline uint32_t GetSizePerImage();


  void SetUsePostedWrites(bool usePostedWrites = true);


  virtual int32_t LockDevice();
  virtual int32_t UnlockDevice();
  ...
};
```

Locking functions are included to enable usage of this object across threads. Using this functions is preferable to using a user-defined `pthread_lock_t` since these functions may be optimized for speed.

Usage of the singleton class is similar to usage of the C framework and examples are provided here:

1. Grab a device and read from it:

```
TUVMEDeviceManager* gMgr = TUVMEDeviceManager::GetDeviceManager()
TUVMEDevice* device = gMgr->GetDevice(0x0, 0x29, 2, 0x10000); //Map
    the entire A16 space beginning at 0x0
uint8_t buffer[4];
if (device == NULL) exit; // Error!
// read at adress 0x3300
```

```
if (device ->Read(buffer , 4, 0x3300) < 0 ) {
  // Error!
  exit;
}
// Otherwise we have a successfull read.

gMgr ->CloseDevice (device);
/* This last call is an unnecessary call since the API library
   handles the closing automatically.  However , if the pool of 8
   devices is already empty , this will release a device to be opened/
   enabled anew via the get_new_device function.  Normally , readout
   code will set up devices at the beginning of the run so it doesn't
    need to be done multiple times. */
```

2. Perform a DMA transfer:

```
TUVMEDeviceManager* gMgr = TUVMEDeviceManager :: GetDeviceManager ()
// Set up a DMA transfer , A32, D32 BLT, beginning at address 0
   x2101000
TUVMEDevice* device = gMgr ->GetDMADevice (0x2101000 , 0xB, 4, true);
if (device == NULL) exit; // Error!
char buffer [4096];
if (device ->Read(buffer , 4096) != 4096) { // DMA transfer 4096 bytes
  //Error!
}
gMgr ->ReleaseDMADevice ()
```

3. Read/Write from/to the Universe registers

```
TUVMEDevice* device = gMgr ->GetControlDevice (); // Grab the ctl
   device.
char buffer [4];
// the ctl device can only read/write 4 bytes at a time!
if (device ->Read(buffer , 4, 0x4) < 0 ) {
  // Error
}
```

C.2 Single-board Computers (SBCs)

This section briefly describes the framework of the development of Single-board computers (SBCs) for the MAJORANA experiment and related test stands. This software was built in conjunction with Mark Howe as part of the ORCA [99] software package and has been applied to SBCs with VME, cPCI, and IPE form factors and use with test stands for the MAJORANA experiment [5], the KATRIN experiment [105], and the SNO+ experiment [106]. The SBCs in VME crates used the driver described in the previous section, Section C.1. The software provides two essential run modes: (1) a slow-control mode to set register settings of cards and (2) a fast readout mode to pull data from cards resident on the bus. The general software design was to provide an interface to the ORCA DAQ program, which would allow the behavior of cards to be customized through register settings, etc. The fast readout would then be handled by the SBC to take advantage of the speed of being directly coupled to the bus. The SBC would read out the cards on the bus and fill a circular buffer with their data. This circular buffer would then be read out asynchronously by the ORCA program in large chunks over a TCP socket. This type of setup is most useful given the high bandwidth yet high latency of Gigabit ethernet. A schematic of the fast readout is shown in Figure C.1.

C.2.1 SBC Card Software Design

C++-based software was designed to handle the fast readout of cards on the bus. A virtual interface was designed so that a single management program could be created to handle cards of various form factors and designs. The virtual base class for this was `class ORVCard` described below:

```cpp
// Struct providing runtime information of the card.
// This is populated and sent from the ORCA DAQ program
struct SBC_card_info;

// Struct encapsulating Look-at-me (LAM) data
// to be returned to ORCA
```

246

Figure C.1: A schematic of the SBC data acquisition on a VME bus. An embedded SBC takes advantage of fast data transfers across the crate bus, filling a circular buffer which is then readout across a TCP socket. The SBC communicates on the VME bus via a PCI-VME 'bridge' chip. The DAQ program on the SBC is a threaded process which takes significant advantage of the dual-core processor onboard.

```
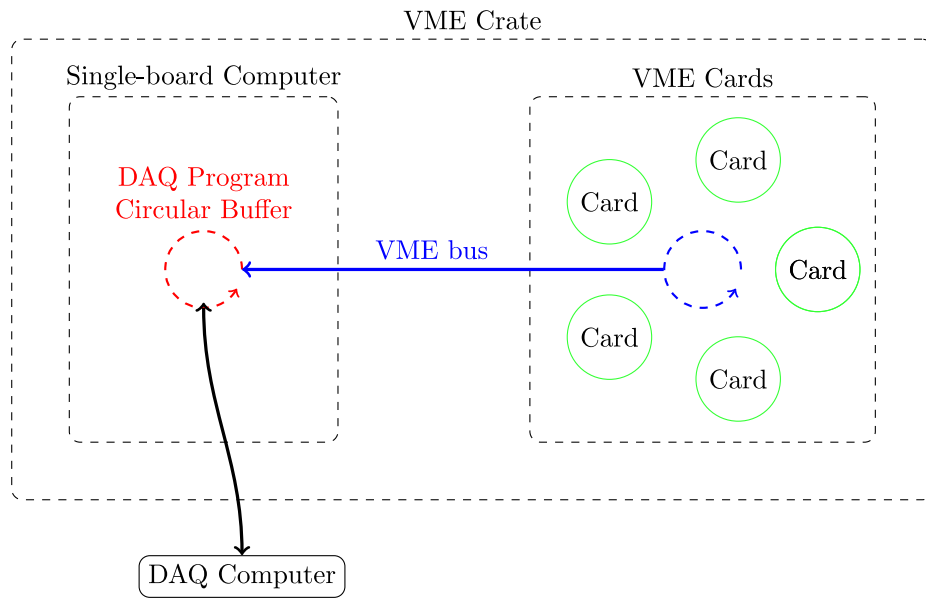struct SBC_LAM_Data;


class ORVCard
{
    public:
        ORVCard(SBC_card_info* ci);
        virtual ~ORVCard() {}


        // Overload the following three member functions
        // to enable readout functionality for a card
        virtual bool Start() { return true; }
        virtual bool Readout(SBC_LAM_Data*) = 0;
        virtual bool Stop() { return true; }

        // Readout function called by management routine
        virtual int32_t ReadoutAndGetNextIndex(SBC_LAM_Data*);

        inline uint32_t GetSlot();
        inline uint32_t GetCrate();
        inline uint32_t GetBaseAddress();
        inline uint32_t GetAddressModifier();
        inline uint32_t* GetHardwareMask();
        inline uint32_t* GetDeviceSpecificData();
        inline int32_t GetNextCardIndex();
        inline uint32_t* GetNextTriggerIndex();
        inline uint32_t GetHWTypeID();
        ...
};
```

This interface provides very basic functionality, essentially running the member functions Start and Stop at the start and stop of a run to initialize or shut down a card, and then calling Readout during a readout cycle. The interface could also easily be extended to particular form factors, for example the VME bus, which wraps calls to the Linux driver API (see Section C.1). The following is a truncated example of this:

```
class TUVMEDevice;
```

```cpp
class ORVVmeCard : public ORVCard
{
   public:
       ORVVmeCard(SBC_card_info* card_info);
       virtual ~ORVVmeCard();


    protected:
    /* All the following functions return the bytes written/read,
       or -1 if an error has occurred. */


       // Provide DMA access
       // A corresponding read exists as well
       static int32_t DMARead(uint32_t vme_address,
                              uint32_t address_modifier,
                              uint32_t data_width,
                              uint8_t* buffer,
                              uint32_t number_of_bytes,
                              bool auto_increment = true);
       ...
       // Provide normal read/write access
       static int32_t VMERead(uint32_t vme_address,
                              uint32_t address_modifier,
                              uint32_t data_width,
                              uint8_t* buffer,
                              uint32_t number_of_bytes);
       ...
       // Other convenience functions are provided
       // that specify the size of the buffer
       static int32_t VMERead(uint32_t vme_address,
                              uint32_t address_modifier,
                              uint32_t data_width,
                              uint32_t& buffer);
       ...
};
```

C.2.2   SBC Management, Fast Readout Design

By designing the code using abstracted functionality, the development of the readout code was simple and is described fully in the following pythonic pseudocode:

```python
...
allTheCards          # List of cards
allTheCardsMap = {} # Dictionary to hold the list
# This dictionary is to allow cards to tell the
# management program which card to trigger next
index = 0
for acard in allTheCards:
    allTheCardsMap[index] = acard

# First perform the initialization
for acard in allTheCards:
    acard.Start()

# Now do the readout
startIndex = 0
while doKeepRunning: # another thread will set this False
    acard = allTheCardsMap[startIndex]
    startIndex = acard.ReadoutAndGetNextIndex(lamData)
    ... # Handle LAM data

# First perform the shut down
for acard in allTheCards:
    acard.Stop()

...
```

This simple interface has already been used successfully with multiple different form factors and trigger types. As well, the base class has enough functionality to enable complex trigger development. For example, a readout of one card might depend on the triggering of another card. By calling the function `ReadoutAndGetNextIndex`, this enables the code in the card class to control the flow of the readout loop. This simple and robust design has simplified debugging, streamlining both the hardware and development process.

VITA

Michael Gregory Marino has lived many places, but probably not so many so as to make a nomad jealous. His life began in San Diego, California, but his family took to the road after he finished 2nd grade, residing in Spokane and Snohomish, Washington, before migrating to Fort Worth, Texas, for the last three years of his high school career. As his parents took to the next logical destination (Alaska), he moved north to attend the University of Notre Dame. His stay in Indiana was interrupted once for a year-long venture to Dublin, Ireland, to pursue studies at Trinity College, but he returned to finish his final year at Notre Dame. From there, he moved to Seattle, Washington to begin graduate studies in Physics at the University of Washington. After completion of these studies in 2010, he proceeded to the Cluster of Excellence for Fundamental Physics in Munich, Germany, for further physics research.