

NORTHWESTERN UNIVERSITY

Ionization yield measurements and low-energy background investigation using a
SuperCDMS-HVeV detector

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Physics

By

Runze Ren

EVANSTON, ILLINOIS

December 2022

© Copyright by Runze Ren 2022

All Rights Reserved

ABSTRACT

Ionization yield measurements and low-energy background investigation using a
SuperCDMS-HVeV detector

Runze Ren

Dark matter (DM) is one of the most outstanding problems in physics and is a promising hint for physics beyond the Standard Model. Many dark matter detection experiments have been built, with weakly-interacting massive particles (WIMP) as a popular DM candidate. There are also growing interest in light (keV-GeV) dark matter (LDM). The Super Cryogenic Dark Matter Search (SupreCDMS) experiment, which is based on transition-edge sensor (TES), is designed for low-mass WIMPs detection. In the meanwhile, SuperCDMS is using the same technology to build more sensitive detectors to probe some LDM models.

In this thesis, I will introduce one of the SuperCDMS R&D programs called HVeV which develops high-voltage detectors with eV-scale resolution. The outstanding performance of HVeV detector enabled two topics that are important for DM research and related fields. First, ionization yield is an essential parameter to calibrate the detector response for low-mass WIMP but is not characterized in the target energy region. I used the HVeV detector to measure the ionization yield in silicon down to 100 eV. Second, I studied the backgrounds in HVeV detectors and identified one that dominates. The sensitivity of HVeV detectors can be increased by two

orders of magnitude if this background source can be eliminated. I also show a DM exclusion limit with a half day of measurement from the HVeV detector.

Acknowledgements

This work is only possible with the support of multiple people and the SuperCDMS collaboration. It is difficult to give proper thanks to everyone who contributed to my growth as a scientist.

I am deeply indebted to Ziqing Hong, who is like a big brother that leads my way in both research and daily life. I miss the long, stimulating discussions on various topics on our way to Fermilab or to the Dim-sum place. Most of my initial insights about the fridge and detector were obtained during these commuting trips. Ziqing also spent a lot of valuable time helping me through my darkest days when finishing graduate school seemed impossibly far away.

I'm extremely grateful to my advisor, Enectalí Figueroa-Feliciano, a.k.a. Tali, who provided us with an inclusive and productive research environment. I still remember how nervous and intimidated I was the first day I came to the lab, but soon I learned that it is a safe place where I can make mistakes and learn. In Tali's group, I was able to keep a good work-life balance despite the rumor that PhD students need to give up a lot of personal time to succeed. Tali and Ziqing are excellent models of scientists with high standards.

To the collaborators that I worked closely with: Yen-Yung Chang, Miriam Diamond, Noah Kurinsky, Belina von Krosigk, Emile Michaud, Brian Nebolsky, Valentina Novati, Doug Pinckney, Tyler Reynold, Benjamin Schmidt, Tarek Saab, Matt Wilson, Yunru Zheng, Alexandra Zaytsev. I would also like to thank Dan Bauer, Adam Mayer, Elham Azadbakht, Wolfgang Rau, Emanuele Michielin, Amy Roberts, Caleb Fink, Matt Pyle, Anthony Villano, and Ruslan Podviianiuk for reviewing the data analyses and papers that I participated in.

To my thesis committee, Tali, Michael Schmitt, Tim Kovachy, and Eric Dahl for helping me get to the end. To TUNL staff, especially Long Li and Phillips Barbeau, for providing the neutron beam that is crucial to my research. And to my friends, Yubo, Bingbin, Gong, Yunpeng, and William who made good companies in a foreign country.

Finally, to my girlfriend, Yishan, and my parents. This dissertation stands as a testament to your love and encouragement.

Table of Contents

ABSTRACT	3
Acknowledgements	5
Table of Contents	7
List of Tables	9
List of Figures	10
Chapter 1. Dark matter and direct detection	13
1.1. Evidence for Dark Matter	13
1.2. Dark matter candidates	15
1.3. Searching for Dark Matter: SuperCDMS experiment	19
1.4. Motivation and outline of this work	28
Chapter 2. SuperCDMS HVeV detector design and characterization	31
2.1. Detector optimization	31
2.2. Experimental setup and Event reconstruction	34
2.3. QET characterization	42
2.4. Detector performance: small signal response	48
2.5. Detector performance: extending the dynamic range	52
Chapter 3. Measuring Ionization Yield of Silicon	58
3.1. Introduction	58

3.2. Experimental set-up	60
3.3. Simulations	67
3.4. Neutron tagging detector analysis	71
3.5. Coincidence analysis	75
3.6. Signal model and fit algorithm	82
3.7. Systematic uncertainty	84
3.8. Result	87
Chapter 4. Investigation of the background in HVeV detector	91
4.1. Introduction	91
4.2. Experimental setup and data collection	92
4.3. Event reconstruction	93
4.4. Dark matter constraints with 0V spectrum	97
4.5. Pulse shape anomalies	103
4.6. Comparison between 0V and HV	109
4.7. Discussion	116
Chapter 5. Conclusion and outlook	121
5.1. Further study of the low-energy background	122
5.2. Understanding the calibration	125
5.3. Dark matter experiments outlook	127
References	128
Appendix A. QET dynamic range and efficiency	137
A.1. QET Dynamic Range	137
A.2. QET Efficiency Modeling	140

List of Tables

2.1	Detector parameters of QP.4 and NFC detector	43
2.2	TES bias circuit parameters	48
3.1	Products created by protons on a LiF	69
3.2	Measured mean recoil energy for the LW detectors.	81
3.3	Measured ionization yield in silicon as well as the measured Fano factor and signal normalization	90

List of Figures

1.1	Dark matter candidates	17
1.2	Overview of QET energy transport (top) and design geometry (bottom).	21
2.1	Optimizing the detector in terms of Al fin and TES lengths	32
2.2	Picture of the detector inside the ADR	35
2.3	Block diagram of the TES readout electronics.	37
2.4	Illustration of the triggering algorithm.	40
2.5	Illustration of the matched-filter-based energy estimator	41
2.6	TES I-V curve and energy efficiency	44
2.7	Noise of NF-C detector, measured and modeled	47
2.8	Example laser energy spectrum	49
2.9	Trigger efficiency	50
2.10	Matched-filter-based energy estimator and gain-baseline variation	52
2.11	^{55}Fe energy spectra	54
2.12	Calibration curve	55
2.13	Energy resolution versus energy	57
3.1	Existing ionization yield measurements and Lindhard model	60
3.2	A photo of the IMPACT@TUNL experimental	61

		11
3.3	Neutron to silicon cross-section	62
3.4	Neutron energy vs emission angle	63
3.5	A sketch of the IMPACT@TUNL experimental setup.	65
3.6	The overall geometry of the IMPACT simulation.	67
3.7	Simulation spectrum before coincidence	69
3.8	TRIM calculated neutron energy distribution	70
3.9	Simulated recoil energy distribution	71
3.10	Fitting the simulation to data before coincidence	72
3.11	Example of events that have a long chain of peaks.	74
3.12	PMT cut demonstration	75
3.13	dt selection on the ring detectors	78
3.14	TOF selection on the LW1 detector	79
3.15	Measured coincidence energy spectra	80
3.16	Fit for the recoil energy	82
3.17	Signal model for IMPACT	83
3.18	Fit to the coincidence spectra	84
3.19	Demonstration of PE generation	86
3.20	Fit result on PEs for systematic uncertainty	87
3.21	Summary of the systematic uncertainties	88
3.22	Measured ionization yields	89
3.23	Measured Fano factors	90
4.1	Calibration curve for 0VeV analysis	96

4.2	Phase cut demonstration	98
4.3	0 V dark matter search energy spectrum	99
4.4	0 V cut efficiency	99
4.5	DM exclusion limit	101
4.6	Example of a “burst” event	104
4.7	Event selection of 0V data	105
4.8	Individual-pulse inter-arrival time (dt) distribution	107
4.9	Pulse arrival time distribution	108
4.10	Energy of primary and secondary pulses of burst events	109
4.11	Selected events for pulse shape comparison	111
4.12	0V vs. HV pulse-shape comparison	111
4.13	Number of secondary pulses N_s vs the primary pulse energy	113
4.14	0 V data selection efficiencies	114
4.15	Response matrices	115
4.16	Comparison of the converted HV spectra with the 0V spectrum.	116
4.17	Example “slow events	118
4.18	Picture of the new copper detector holder	119
4.19	Improvement of background with the new holder	120

CHAPTER 1

Dark matter and direct detection

One of the current most outstanding problems in physics is that ordinary baryonic matter could not explain a lot of observational phenomena. A new form of matter called “dark matter (DM)” is hypothesised to account for approximately 85% of the matter in the universe. There are many other alternative hypotheses, but they could explain only part of the observational evidence since the evidence comes from many independent approaches. Dark matter remains the simplest way to explain multiple pieces of evidence. Thus, the existence of dark matter is generally accepted, even though no one has directly observed it.

No known particles can account for a significant part of the DM. This makes DM one of the most evident hints for new physics beyond the Standard Model (SM). In this section, I summarise some of the main evidence for DM at different scales and introduce one of the DM detection experiments — the SupreCDMS experiment. I describe the motivations of this work at the end of this section. There have been in-depth reviews of DM [1, 2, 3], so I will keep the first two sections in this chapter minimal.

1.1. Evidence for Dark Matter

This section provides a non-exhaustive list of the most intriguing evidence for dark matter. For more information please refer to Ref. [1, 2].

1.1.1. Galaxies

Galaxy rotation curves One of the earliest and probably most known evidences is the rotation curve of spiral galaxies. The velocity of stars v at the distance R from the galactic centre is given as

$$(1.1) \quad v(R) = \sqrt{\frac{GM(R)}{R}}$$

in which $M(R)$ is the total mass within a radius R . As the distance increases, the total mass reaches constant and we expect the velocity to decrease as $1/\sqrt{R}$. However, in most of spiral galaxies the observed velocity far from the centre is approximately constant (see for example Ref. [4]), which is called the flat rotation curve. A flat rotation curve leads to a $M(R)$ proportional to R . This distribution of $M(R)$ extends beyond the visible disc, forming a invisible halo which is considered to be composed of dark matter.

Velocity dispersion Stars in bound systems like elliptical galaxies or globular galaxy clusters must obey the virial theorem. Together with the measured velocity distribution, one can calculate the mass distribution of the system or vice versa. The observed velocity distribution usually does not match the calculated velocity dispersion from the observed mass distribution.

1.1.2. Galaxy clusters

A Galaxy cluster is a structure that consists of many galaxies that are bound together by gravity, which contains large quantities of gases in the intergalactic medium. Galaxy clusters are very important for DM since their mass can be measured in multiple independent ways: the velocity dispersion mentioned above, x-ray emission, and the gravitational lensing. The x-ray emission measures only the visible (baryonic) mass, while gravitational lensing is based on gravitational effect only and does not depend on dynamics or electromagnetic radiation.

Gravitational lensing The presence of mass between a distant source and an observer acts as a lens to bend the light from this source, which is called gravitational lensing. The mass along the path can be calculated by observing the lensing effect and is therefore used to weigh the galaxy clusters. Many studies are consistent with 10-20% of visible mass.

Gravitational lensing, together with x-ray emission, can be a paradigmatic evidence of DM. The Bullet Cluster is one of the best known examples.

1.1.3. Cosmological scale

Dark matter also leaves imprints at the cosmological scale. Dark matter affects the cosmic microwave background (CMB) by its effects on the density and velocity of baryonic matter, which turns into small temperature anisotropies of CMB. Dark matter also provides a solution to problems in structure formation, without which the density perturbations are unable to grow into the current galaxies and clusters. Baryon acoustic oscillations (BAO), fluctuation in the density of the visible baryonic matter, is another evidence of dark matter. BAO can be seen in CMB, but can also be measured independently from galaxy redshift surveys, which in combination with CMB provide a precise estimate of the baryonic matter and dark matter density in the Universe.

1.2. Dark matter candidates

Dark matter can have a mass of any value. Many candidates have been proposed based on a list of properties from the multiple evidences:

- (1) Electrically neutral. Otherwise they would scatter light and thus not be dark.
- (2) Non-Baryonic. Result of CMB together with BBN shows that the majority of dark matter must be non-baryonic.

- (3) Nonrelativistic. Study of structure formation shows that dark matter needs to be non-relativistic to agree with observations.
- (4) Long-lived, with lifetime on the order of the age of the Universe.
- (5) Weak self-interaction. Dark matter’s lack of deceleration in the bullet cluster limits the upper bound of its self-interaction cross-section.

Most of the candidates assume that dark matter interacts with standard model particles not purely through gravity. This includes popular models like classes of axions, Weakly Interacting Massive Particle (WIMP), sterile neutrino, etc. WIMP is considered to be an important model with theoretical and experimental motivations, especially for the first two decades of searches for dark matter. Light dark matter (LDM) has been gaining popularity in recent years. There are also many other dark matter models, such as ultra-light dark matter, black holes, etc. In this section, I will introduce WIMP and LDM, which are related to the current work. See Ref [1, 2, 3] for in-depth review of dark matter.

1.2.1. WIMP dark matter

WIMP is a widely discussed and tested category of dark matter [5]. “Weakly” means that their interactions are only through the weak nuclear force and gravity or other possible interactions no higher than the weak scale; “massive” means that they are massive enough (GeV scale) compared to standard particles.

WIMP was first proposed to solve a list of astrophysical and cosmological evidence related to dark matter. The model is based on relic particles from the early Universe when all particles were in a state of thermal equilibrium. A new unknown particle is assumed to be in equilibrium with standard model particles in the early Universe. With this assumption, the velocity averaged annihilation cross-section of dark matter $\langle\sigma_{AV}\rangle$ can be predicted. The predicted cross-sections are around weak-scale interactions ($\sigma \sim 10^{43}\text{cm}^2$, $m_\chi \sim 20 - 400\text{GeV}/c^2$). On the other

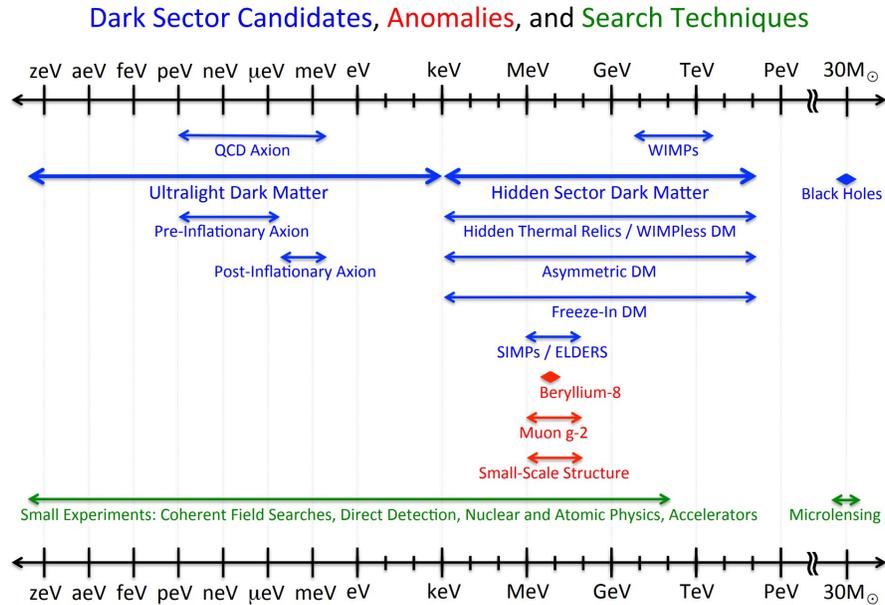


Figure 1.1. (Figure from Ref. [1]) Mass ranges for a subset of dark matter candidates. The ranges are only representative.

hand, in particle physics, WIMP-like particle interactions on weak-scale are predicted by several extensions of Standard Model theories such as supersymmetry (SUSY), extra dimension (UED) and little Higgs. This coincidence is referred to as the “WIMP miracle”.

A large part of the existing effort in DM detection is motivated by the WIMP hypothesis. The large-scale Xenon-based experiments have great sensitivity in the high-mass (> 10 GeV) region, and the low-threshold Si/Ge experiments cover the low-mass region. SuperCDMS is one of the low-threshold experiments designed for sub-10 GeV DM search. After years of searching, the parameter space of WIMP models has shrunk but is still viable. WIMP searches remain of great interest — both the large-scale experiments and low-threshold experiments keep improving.

There is a theoretical constraint on WIMP mass (the Lee-Weinberg bound) which limits the energy to be roughly above 2 GeV [6]. Below this limit, there will be an overabundance of DM that does not match observations. This number is estimated with a large uncertainty —

a 1-4 GeV range is quoted in Ref. [6]. Thus, the exact 2 GeV should not be taken as a hard constraint.

1.2.2. Light dark matter

Light dark matter shares a similar motivation as the WIMP: the starting point is the early history of the Universe, and the coupling to ordinary matter generates the observed DM abundance. The difference is that LDM is assumed to not interact under Standard Model forces to avoid the Lee-Weinberg bound. Instead, there needs to be a new mediator, which is sometimes called a “dark sector” or “hidden sector.” The most interesting subset of the models lies in the keV to GeV range. See Ref. [7, 8] for more details about LDM.

The low mass of the LDM makes them deposit very small energy in elastic scattering, based on the simple two-body scattering kinematics:

$$(1.2) \quad \Delta E \leq \frac{2\mu^2 v_\chi^2}{m_T} \sim \frac{1}{2} m_\chi v_\chi^2 \cdot \frac{4m_\chi}{m_T} \text{ (when } m_\chi \ll m_T\text{),}$$

in which m_T is the mass of target nucleus, m_χ is the mass of dark matter particle and $\mu = m_\chi m_T / (m_\chi + m_T)$ is the reduced mass. Fortunately, inelastic scattering is not limited by this kinematics. The maximum energy transfer is then

$$(1.3) \quad \Delta E \leq \frac{1}{2} \mu v_\chi^2 \sim \frac{1}{2} m_\chi v_\chi^2 \text{ (when } m_\chi \ll m_T\text{).}$$

which is no longer suppressed by a factor of $\frac{4m_\chi}{m_T}$. For a dark matter particle with the speed of $v_\chi^2 \sim 600\text{km/s}$, and a mass of 5 MeV, the energy transfer is around 10 eV, which is enough to create ionization in semiconductors with eV scale bandgap. This is the reason why the DM-SM cross section is given in the electron-scattering cross-section for LDM, as opposed to the nucleon

cross section for WIMPs. The single-charge sensitive detectors developed by SuperCDMS, which will be introduced later in Chapter 2, is one near-term tool to detect LDM.

1.3. Searching for Dark Matter: SuperCDMS experiment

There are three main experiment types used to search for DM: direct detection, indirect detection, and collider searches. At particle colliders, in particular the LHC, one can look for the creation of DM, which have final states with missing transverse energy. On the contrary, indirect detection looks for the Standard Model product through DM annihilation or decay. Finally, direct detection is hoping to directly measure the interaction of DM on ordinary matter, which is the focus of the SuperCDMS SNOLAB experiment[9]. SuperCDMS SNOLAB mainly searches for sub-10 GeV/c² WIMPs-like DM via spin-independent DM-nucleus elastic scattering.

SuperCDMS detectors are based on cryogenic phonon-mediated calorimeters, which employ large target volumes coupled to smaller volume superconductors to channel energy into a small heat capacity that can be read out at high signal-to-noise ratio. The detector consists of two main parts: 1) a macroscopic substrate as the particle-sensing target (a Si or Ge crystal for SuperCDMS), and 2) phonon sensors instrumented on the substrate. There is another family of SuperCDMS detector with dedicated charge readout called iZip. In this thesis, I will focus on the phonon-only detector.

The phonon-mediated detectors have two major benefits. First, large masses can be achieved while maintaining good energy resolution. Second, the phonon signal measures the full energy of the interaction regardless of the recoil type, which makes it possible to achieve a lower threshold on nuclear recoil than a traditional ionization-based detector. The SuperCDMS detectors also have the additional feature of being able to measure the ionization signal, making them also sensitive to LDM models.

This section gives an introduction to the detector technology used by the SuperCDMS experiment. First, I introduce the mechanism of Quasiparticle-trap-assisted Electrothermal-feedback TESs (QETs) [10] phonon sensors. Then I describe the physical process of nuclear and electron recoil in the substrate to clarify the total energy measured by the detector. Next, I discuss how we measure the charge signal. In the end, I briefly introduce the charge trapping and impact ionization processes that affect the charge measurement. Detailed description of the SuperCDMS SNOLAB experiment can be found in Ref. [9].

1.3.1. QET phonon sensor

The phonon sensor in our design is a parallel array of QETs for each readout channel. As shown in Fig. 1.2, a QET is made up of a superconducting thin film as a phonon collector (the Al fins) and a Transition-Edge Sensor (TES) [11] as a phonon detector.

On a microscopic level, phonon energy in the target from particle interactions is converted to superconducting quasiparticles in the Al fin phonon collectors. The Al fins employ their superconducting gap energy ($\sim 350 \mu\text{eV}$ for Al) to separate athermal phonons from the residual thermal phonons at low temperature ($\sim 1 \mu\text{eV}$ at 10 mK), thus providing a relatively fast sensor response. The Al/W overlap region has a lower gap than the Al bulk, forming a quasiparticle trap which funnels quasiparticles into the much smaller TES volume as they shed energy via the emission of phonons.

The energy transport in QET is shown in 1.2. Athermal phonons generated by events in the substrate propagate with high efficiency to the Al/substrate interface, where they are either transmitted or reflected. The transmitted phonons break Cooper pairs in the Al, creating free, athermal quasiparticles (QPs), which diffuse through the fin from the initial event. When these QPs encounter the lower-gap energy region of the Al/W quasiparticle trap, they convert most of the initial gap energy to phonons, heating the TES.

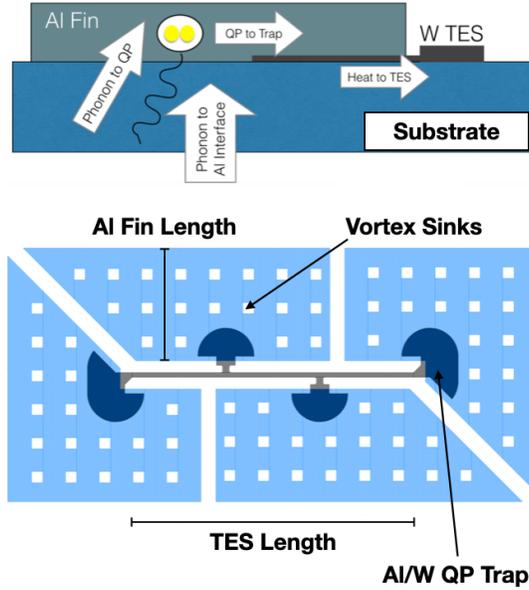


Figure 1.2. Overview of QET energy transport (top) and design geometry (bottom).

The TESs connected to these traps are operated in their superconducting transition with a voltage bias, producing an electro-thermal feedback effect [11]. They convert the phonon energy into a current change which can be sensed using cryogenic amplifiers. The parallel array of QET cells is spread out over the crystal surface, with the number of cells, coverage pattern, and individual QET design all affecting the performance of the device.

As detailed in Refs. [11, 12, 13], the intrinsic energy resolution of a TES calorimeter can be written in terms of the detector bandwidth (expressed as the time constant τ_{BW}), the efficiency of phonon energy collection ϵ , the thermal conductance G between the TES and the crystal (and associated power-law constant n), and the calorimeter operating temperature T_0 as

$$(1.4) \quad \sigma_{ph} = \frac{1}{\epsilon} \sqrt{2Gk_b T_0^2 \tau_{\text{BW}}},$$

where k_b is the Boltzmann constant. For a TES with a narrow transition width, T_0 can be reasonably approximated by the critical temperature T_c of the TES.

For these devices, the G is set by the volume of the TES and its electron-phonon coupling. The thermal power between the TES and the crystal is described by the equation

$$(1.5) \quad P_0 = \Sigma \frac{v_{\text{TES}}}{\zeta_{\text{TES}}} T_c^n \left(1 - \left[\frac{T_b}{T_c} \right]^n \right)$$

such that, when linearized around T_c , the thermal conductance is

$$(1.6) \quad G \approx n \Sigma \frac{v_{\text{TES}}}{\zeta_{\text{TES}}} T_c^{n-1} \approx \frac{n P_0}{T_c} \Big|_{T_b \ll T_c},$$

where Σ is the electron-phonon coupling constant for a W TES, ζ_{TES} is the fraction of the W volume contained in the TES length, v_{TES} is the total TES volume and n is the thermal conductance power-law exponent for the power equation, nominally taken to be $n \sim 5$ for electron-phonon coupling [14]. T_b is the base temperature of the cryostat, also known as bath temperature. Because there is a strong thermal conductance between the crystal and the bath in our setup, we refer to the crystal temperature also as T_b when the system is at equilibrium. This allows us to substitute G in the resolution scaling, giving

$$(1.7) \quad \sigma_{ph} \approx \frac{T_c^3}{\epsilon} \sqrt{2n \Sigma \frac{v_{\text{TES}}}{\zeta_{\text{TES}}} k_b (\tau_{ph} + \tau_-)},$$

where the bandwidth has been broken into phonon collection time τ_{ph} and effective TES response time τ_- (see e.g. Ref [15]). This result for athermal phonon detectors shows that the energy resolution scales as T_c^3 when phonon dynamics limit the integration time and the TES is limited by its own thermal fluctuations.

An additional consideration in detector design which becomes relevant for more general purpose TES detectors is dynamic range, and the related quantity, saturation energy. The

resolution model described above applies strictly in the small-signal limit; away from this limit, the TES response becomes non-linear, and for large enough events, enough energy is supplied to drive the TES into the normal state, which is referred to as the saturation energy. For transition width ΔT_c and specific heat c_W we find a saturation energy E_{sat} of

$$(1.8) \quad E_{sat} \approx \frac{1}{\epsilon} C(\Delta T_c) = \frac{1}{\epsilon} c_W \frac{v_{TES}}{\zeta_{TES}} T_c(\Delta T_c).$$

We thus see that many of the design drivers that minimize resolution (e.g. reducing TES volume and bias power) also reduce saturation energy. The total pulse integral is still a singular function of event energy above this point, but the saturation energy sets a rough scale where the TES goes from the linear to non-linear regime, and the resolution becomes energy dependent. The linear dynamic range is thus roughly the ratio of saturation energy to resolution, which scales as roughly

$$(1.9) \quad DR \sim \frac{E_{sat}}{\sigma_{ph}} \propto \frac{\sqrt{v_{TES}}}{T_c^2 \sqrt{\tau_{BW}}} (\Delta T_c)$$

and we see that, for fixed T_c , smaller TES volume decreases overall dynamic range. The subject of this paper is largely how to balance the typical TES resolution model, summarized above, with the dynamic range model we present for the first time in this paper. We also compare the model predictions with the measured detector response. We make the model in this paragraph more precise by including the TES response model; a reader interested in that modeling can jump to Appendix [A.1](#) before proceeding to the next section for more detail.

1.3.2. Detector response to nuclear and electron recoil

The detector substrate senses a particle by nuclear or electron recoil between the particle and the substrate, which leaves recoil energy E_r in the substrate. Although all of E_r will end up as

phonon energy eventually, the actually measured phonon energy E_{ph} can be different from E_r for various reasons. It is necessary to understand some physics of the nuclear and electron recoil in the substrate to discuss the energy scale of the detector.

We first need to take a big step to simplify the physics process of collective many-body scattering into a cascade of two-body scattering problems. A particle interacts with an electron or nucleus in the substrate crystal, resulting in a primary recoiling electron or nucleus. The recoiling electron or nucleus continues to scatter by ionization and by phonon emission, producing a cascade of electron-hole pairs (e^-/h^+) and phonons. At the end of the cascade that comes almost immediately after the primary recoil, the recoil energy distributes between phonons and electron-hole pairs:

$$(1.10) \quad E_r = E_{ph0} + E_g n_{eh}$$

where E_{ph0} is the phonon energy immediately after the cascade that we call “prompt phonon”, E_g is the bandgap energy and n_{eh} is the number of e^-/h^+ produced by the event.

Eventually, all of e^-/h^+ will recombine and release the bandgap energy by emitting phonons. If this happens within the normal time scale of an event, the total phonon energy E_{ph} equals the recoil energy

$$(1.11) \quad E_{ph} = E_r = E_{ph0} + E_g n_{eh}.$$

However, there is the possibility that the recombination happens at a much longer time scale or at traps with a depth less than the bandgap energy. In these cases, the total phonon energy is less than the recoil energy, with a lower bound of

$$(1.12) \quad E_{ph} = E_r - E_g n_{eh} = E_{ph0}.$$

when none of the bandgap energy of generated e^-/h^+ is released. I assume this is an extreme case and will be using Equation 1.11 in this thesis. In Chapter 5 I will discuss the effect on detector calibration if charges do not recombine timely.

The number of e^-/h^+ produced is a function of the recoil energy and is very different for electron and nuclear recoil.

For electron recoil, the number of e^-/h^+ in an ideal crystal is described as

$$(1.13) \quad n_{\text{eh}} = \frac{E_r}{\varepsilon_\gamma(E_r)},$$

where $\varepsilon_\gamma(E_r)$ is the average energy per e^-/h^+ . At low energy below the bandgap E_g , no ionization is possible. For energies between E_g and $2E_g$, only one electron-hole pair is allowed by energy conservation. At the high energy limit $E_r \gg E_g$, the statistical nature of this ionization process leads to asymptotic value for $\varepsilon_{\gamma,\infty}$ which shows good agreement between model and measurement for multiple materials. The region between $2E_g$ and where the asymptotic is a good approximation is still not well studied. To summarize the behaviour of $\varepsilon_\gamma(E_r)$:

$$(1.14) \quad \varepsilon_\gamma(E_r) = \begin{cases} \infty & E_r < E_g \\ E_r & E_g < E_r < 2E_g \\ \varepsilon_{\text{imp}}(E_r) & 2E_g < E_r \\ \varepsilon_{\gamma,\infty} & E_r \rightarrow \infty \end{cases}$$

where $\varepsilon_{\text{imp}}(E_r)$ is the unknown function. This relation will not always hold for real crystal with impurities and defects. For example, we know that infrared photons with sub-bandgap energy can also generate e^-/h^+ .

Recent measurements of $\varepsilon_{\text{imp}}(E_r)$ and $\varepsilon_{\gamma,\infty}$ in Si is summarized in [16, 17]. Reference [17] also reproduced a model of $\varepsilon_\gamma(E_r)$ first introduced by Ref. [16]. The exact number of $\varepsilon_{\text{imp}}(E_r)$

depends on temperature because of the temperature dependence of bandgap, which makes the choice of this value different among experiments. In this dissertation, we use $\varepsilon_{\gamma,\infty} = 3.8$ eV, which is a choice of SuperCDMS collaboration.

For nuclear recoil, the number of e^-/h^+ generated is usually less than an electron recoil of the same energy. Since the ionization of electron recoil is more well studied, people usually describe the n_{eh} of nuclear recoil relatively to the electron recoil and call the ratio Y “quenching factor” or “ionization yield”:

$$(1.15) \quad n_{eh} = Y(E_r) \cdot \frac{E_r}{\varepsilon_{\gamma,\infty}}.$$

Y is a function of recoil energy and is smaller than one for nuclear recoil in general. Equation 1.13 can be treated as a special case where the yield $Y = 1$. Moreover, the value of mean energy-per-pair $\varepsilon_{\gamma}(E_r)$ is taken to be its high energy asymptotic value for $\varepsilon_{\gamma,\infty}$ in this definition.

1.3.3. Measuring charge with phonons

The SuperCDMS detectors are also able to measure ionization signal through the Neganov-Trofimov-Luke (NTL) effect [18, 19] by applying a voltage across the substrate. Under the electrical field, the initial electron-hole pairs are accelerated and drift across the crystal, resulting in an amplified phonon signal. The total phonon energy, E_{ph} , is now the initial energy deposition E_r plus what is produced by the NTL effect, which is the potential energy of all e^-/h^+ :

$$(1.16) \quad E_{ph} = E_r + n_{eh} \cdot e \cdot V_{\text{NTL}}$$

where e is the elementary charge, n_{eh} is the number of electron-hole pairs produced, V_{NTL} is the bias applied across the detector. If we take the previously defined n_{eh} from Eqn. 1.15, we can

organize the equation into:

$$(1.17) \quad E_{ph} = E_r \left(1 + \frac{e \cdot V_{\text{NTL}}}{\varepsilon_{\gamma, \infty}} \cdot Y(E_r) \right) = E_r \cdot G_{\text{NTL}},$$

The signal is effectively amplified by a factor of G_{NTL} . When the voltage is high enough, the signal can be amplified to the point that the detector is sensitive to a single electron-hole pair.

We can invert Eq. 1.17 to obtain the charge resolution (σ_q) as a function of the phonon energy resolution (σ_{ph}) when the NTL amplification is significantly larger than the initial energy deposition:

$$(1.18) \quad \sigma_q \approx \frac{\sigma_{ph}}{e \cdot V_{\text{NTL}}}.$$

In this thesis, we talk a lot about energy depositions. From now on, we follow the convention that we'll use $\text{keV}_{\text{nr}}/\text{eV}_{\text{nr}}$ as the unit for energy transfer in nuclear recoil in silicon and liquid scintillator, $\text{keV}_{\text{ee}}/\text{eV}_{\text{ee}}$ as the unit for electron-equivalent energy, which is the recoil energy times the ionization yield $E_r \cdot Y(E_r)$ with the convention that $Y=1$ for electron recoil, and $\text{keV}_{\text{t}}/\text{eV}_{\text{t}}$ to denote total phonon energy in a silicon detector, which is the sum of the initial recoil energy (nuclear recoil or electron recoil) plus the phonon energy induced by NTL effect. Also, we refer to the mode without the voltage on the substrate as the “0V mode”, and the mode with voltage as the “HV mode”.

1.3.4. Charge trapping and impact ionization

In HV mode, the detector response is dominated by the number of e^-/h^+ produced. We've already discussed the number of e^-/h^+ generated (n_{eh}) for electron and nuclear recoil in Sec. 1.3.3 for ideal crystal. Each e^-/h^+ needs to travel across the crystal to the surface to have full NTL gain. However, in a real crystal with defects and impurities, the charge carrier may be trapped

or generate new charge carriers before they reach the surface, which is named as charge trapping (CT) and impact ionization (II) process. These two process affects the expected total phonon energy.

We incorporate the model for charge trapping and impact ionization process described in Ref. [20]. The model describes charge trapping and impact ionization with two parameters: f_{CT} and f_{II} , which are the probability a charge carrier to undergo CT and II, respectively. This model treats the electrons and holes with the same f_{CT} and f_{II} , since measurements have found similar values for electrons and holes. We assume that CT and II happen uniformly across the crystal, which means if a charge is trapped, on average, it still creates phonon energy equal to $0.5 \cdot e \cdot h$. The average total phonon energy can then be calculated as

$$(1.19) \quad E_{ph} = E_r \left(1 + \frac{e \cdot V_{NTL}}{\varepsilon_{\gamma, \infty}} \cdot Y(E_r) \cdot (1 - 0.5f_{CT} + 0.5f_{II}) \right).$$

A more complicated model of IICT can be found in M.Wilson's thesis [21].

1.4. Motivation and outline of this work

This dissertation is arranged as the following. Chapter 1 (the current chapter) lays out the background of dark matter search and the detector physics of the SuperCDMS experiment. Chapter 2 is about the design, characterization and performance of the SuperCDMS-HVeV detector. Chapter 3 summarizes the experiment of nuclear-recoil ionization yield measurement and chapter 4 is the study of background events in HVeV detector.

1.4.1. The need of nuclear-recoil calibration

Experiments searching for WIMP dark matter and coherent neutrino-nucleus scattering ($CE\nu NS$) often rely on detecting nuclear recoil events in the target material. Many detectors operate based on the ionization signal of the nuclear recoil. It is vital to understand the material's ionization

response to nuclear recoil, which is characterized as nuclear recoil ionization yield function $Y(E_r)$.

One can calibrate the nuclear recoil response of their detector with a neutron source in principle. However, a neutron source is far less convenient than a photon source, making it harder to do a direct neutron calibration for each individual detector. Since $Y(E_r)$ is believed to be a material intrinsic property, many experiments choose to calibrate the detector with a photon source and apply a previously measured $Y(E_r)$ to get the nuclear recoil response.

Silicon is a commonly used target material. The $Y(E_r)$ in silicon has been measured by multiple experiments in the energy range above 1 keV_{nr}. With the development of detectors, the energy threshold has been lowering, and there are growing interest in detecting sub-1 GeV WIMP and measuring CE ν NS. A sub-1 GeV WIMP produces nuclear recoil less than 1 keV_{nr}. The only measurement of $Y(E_r)$ in this region deviates from a model which has been quite successful in the high-energy region. Moreover, there are new models which predict an ionization threshold at O(100 eV), which adds the uncertainty of the ionization yield and changes the expected reach of the DM-search experiments.

In order to better understand the response of nuclear recoils, SuperCDMS initiated an ionization yield measurement program called Ionization Measurement with Phonons At Cryogenic Temperatures (IMPACT). Multiple measurements are foreseen, including different material targets (silicon and germanium) and different neutron energies (including but not limited to neutrons produced with a DD neutron generator at Fermilab or a proton beam at Triangle Universities Nuclear Laboratory (TUNL)). In Chapter 3 I report the result of the first campaign of IMPACT, which investigated the ionization yield for a silicon detector by using a 56-keV neutron beam produced at TUNL.

1.4.2. Unexplained excess of events

Low-mass (sub-GeV/ c^2) dark matter searches have benefited from the detector developments with low energy threshold and low readout noise. Despite this progress, their reach has been challenged by unexpected, excess event rates. These include reports from experiments using cryogenic calorimeters instrumented for readout of phonon signals, such as EDELWEISS [22, 23], CRESST [24], NuCLEUS [25, 26], and SuperCDMS-CPD [27]. Unexpected events are also present in detectors instrumented for charge readout, such as the CCD-based experiments SENSEI [28] and DAMIC [29], as well as the phonon-based measurement of ionization signals [30, 31].

The excess of events is a great obstacle to the DM-search and the CE ν NS community. We also saw a similar excess in the SuperCDMS-HVeV detector. In Chapter 4 I studied and identified one dominating origin of the excessive events in the HVeV detector. Although other experiments may not be seeing the exact same mechanism, the result of Chapter 4 can still be helpful in guiding a good design for low-background detectors.

CHAPTER 2

SuperCDMS HVeV detector design and characterization

This chapter has been published on Physics Review D as Ref. [32]. Printed with permission of the co-authors and American Physical Society (APS). Copyright © 2020 by American Physical Society.

SuperCDMS high-voltage eV-scale (HVeV) detectors are a family of detector that is designed based on the theory of the previous chapter. The development of HVeV aims for good resolution in both phonon detection and phonon-assisted charge detection. These detectors have demonstrated eV-scale resolution and single- e^-/h^+ sensitivity in recent years. The good performance of HVeV detectors enabled multiple research projects such as ionization yield measurement, dark matter search, and the investigation of background events.

The HVeV detectors have iterated over three generations. The first generation proved the concept of single-charge resolution. The second generation is optimized based on the measurement of the first. It is used in this work and is also the focus of this chapter. Section 2.1 outlines the design and optimization of the second-generation HVeV detectors. Section 2.2 describes the experimental setup and data processing algorithm. Section 2.3 compares the tested sensor performance to the model. Section 2.4 and 2.5 discusses the performance for events near threshold and at high energy.

2.1. Detector optimization

The detector described in this paper (referred to as NF-C, one of the Gen-2 HVeV) was designed for ionization yield measurements in a neutron beam at the Triangle Universities Nuclear

Laboratory (TUNL) [33]. This application required a device that could measure large energy depositions (~ 100 keV) while maintaining excellent baseline resolution. NF-C is a re-optimization of the detector mask from Ref. [34] (referred to as QP.4, a Gen-1 HVeV), which attained the desired energy resolution (3 eV), but not the dynamic range. We apply the modeling framework described in Ref. [13] to map out the response of detectors as we varied design parameters.

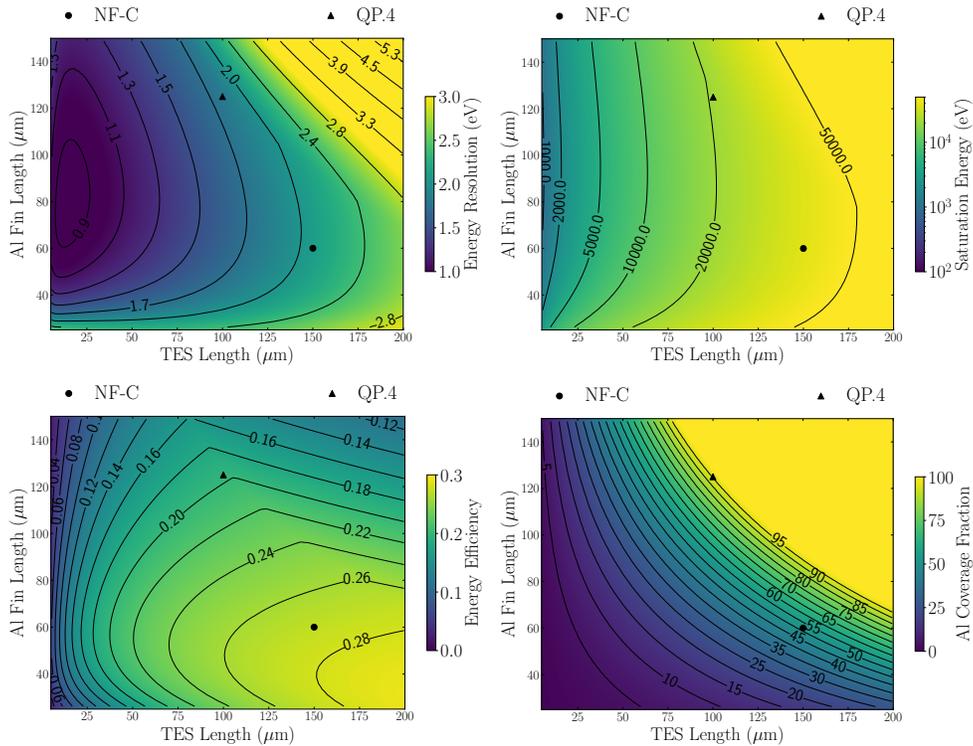


Figure 2.1. Detector energy resolution (top left), saturation energy (top right), energy efficiency (bottom left) and Al coverage fraction (bottom right) as a function of the Al fin and TES lengths. The model predictions for detectors QP.4 (triangle) and NF-C (circle) are also shown.

This detector response scan can be done independently of readout considerations by fixing the QET channel's overall normal resistance R_n . We also fix all TES properties (including T_c , width and thickness of the W) except for TES length (l_{TES}) and Al fin length (l_{fin}) to those measured from QP.4 [34]. The number of QETs (N_{QET}) in a channel is set to be a function of

TES length such that R_n is kept constant, with

$$(2.1) \quad R_n = \frac{R_{\text{TES,QET}}}{N_{\text{QET}}} = \frac{\rho_{\text{TES}} \cdot l_{\text{TES}}}{w_{\text{TES}} \cdot t_{\text{TES}} \cdot N_{\text{QET}}},$$

where ρ_{TES} is the W resistivity (which is T_c dependent), w_{TES} and t_{TES} are the width and thickness of the TES, respectively, and $R_{\text{TES,QET}}$ is the normal resistance of each QET cell. In this limit, the volume of TES per channel ($v_{\text{TES,ch}}$) scales as

$$(2.2) \quad v_{\text{TES,ch}} = N_{\text{QET}} v_{\text{TES,QET}} = \frac{\rho_{\text{TES}}}{R_n} l_{\text{TES}}^2,$$

where $v_{\text{TES,QET}}$ is the volume of the TES per QET cell.

We parametrize the geometry of a QET cell in the 2-dimensional space of $(l_{\text{TES}}, l_{\text{fin}})$. Because l_{TES} determines the number of QETs in a channel, the overall Al coverage fraction (a key parameter in the efficiency ϵ) is also set by these two parameters. With these design rules we can parametrize the detector energy resolution, saturation energy, energy efficiency, and Al coverage fraction in the 2-dimensional space of $(l_{\text{TES}}, l_{\text{fin}})$. The results of this modeling are shown in Fig. 2.1, along with the design points for QP.4 and NF-C. The efficiency model from Ref. [13] is qualitatively described in Appendix A.2.

As stated above, the NF-C design goal was to retain the QP.4 energy resolution while increasing the dynamic range, which is a function of the saturation energy. The dynamic range can be extended by increasing the volume of the TES (see Appendix A.1). At the same time, from Eq. 1.7 we see that we can avoid degrading the energy resolution by simultaneously increasing the collection efficiency such that we keep the ratio $\sqrt{v_{\text{TES}}}/\epsilon$ approximately constant. The chosen parameters for NF-C increase the efficiency projection from around 20% to 27% as the TES length increases from 100 to 150 μm , maintaining a relatively constant ratio of TES length

to energy efficiency; as a result, the overall energy resolution is largely constant. The model predicts the dynamic range is increased by 50% relative to the QP.4 detector.

2.2. Experimental setup and Event reconstruction

The fundamental need for operating the detector is a cryogenic refrigerator to cool the detector around the critical temperature. In this case, we use a VeriCold Adiabatic Demagnetization Refrigerator (ADR), which is cooled to ~ 50 mK and can maintain the temperature for ~ 10 hours.

The detector is clamped between two printed circuit boards (PCBs) for thermal sinking and to facilitate electrical connections. The PCBs are later found to be scintillating, which is a dominating source of the background in HVeV detector, see chapter 4. The QETs on the top side of the detector and the aluminum grid on the back side are connected via wirebonds to traces on the PCB's top surface. A light-tight copper box surrounds the detector and the PCB clamps. Traces on PCB are connected to the SQUID readout system and outside of the fridge. The SQUID readout system is crucial to convert the small current change in the detector into an easy-to-digitize voltage signal. The details of the SQUID readout can be found in Sec. 2.2.1.

The output of the SQUID readout system is logged down with a continuous streaming data acquisition (DAQ) system based on a NI PCIe-6374 card with a sampling frequency of 1.51 MHz. Offline software phonon triggering and further data processing are described in Sec. 2.2.2.

In order to calibrate the detector, we employ two photon feedthrough systems for optical photons and soft X-rays. First, a plastic optical fiber with a core diameter of 1 mm was fed through the detector box, with the gap between the fiber and the box filled with Eccosorb epoxy [35]. The plastic optical fiber was coupled to a single-mode optical fiber [36] through two pieces of KG-3 glass at 1.4 K. The single-mode fiber and the KG-3 glass filter were chosen to attenuate infrared photons from ambient and black body radiation from higher temperature

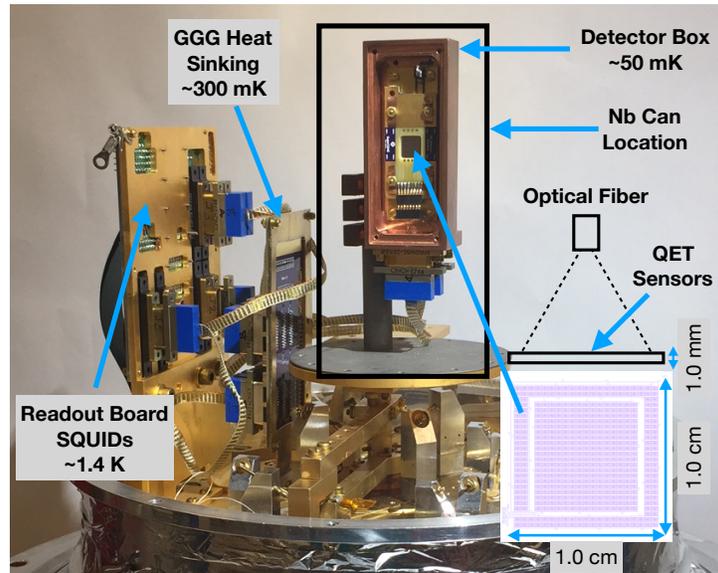


Figure 2.2. A side view of the detector box mounted inside the ADR with the outer shielding removed. The inset picture shows the schematics of the detector used, together with the optical fiber and its field of illumination. The cartoon shows the detector and laser from the side; the detailed diagram of the two phonon channels is a top-down view.

stages. The other end of the single mode fiber was connected to a vacuum feed-through at room temperature, then to a laser diode with a wavelength of 635 nm (corresponding to 1.95 eV per photon)[37]. For the second feedthrough system for soft X-rays, a 1 cm² square opening was cut on the copper box lid and re-sealed with a piece of aluminum foil 0.17 mm thick. The opening aligned with a Beryllium window installed on the ADR, serving as an X-ray input port. Multiple layers of Aluminized mylar sheets were placed between the opening and the Beryllium window at different thermal stages to block black body radiation from higher temperature stages while presenting minimal X-ray attenuation.

2.2.1. SQUID readout system

The signal of a HVeV detector is the current change in the TES. We use a SQUID [38] as a transimpedance amplifier to convert the tiny current from the TES into a voltage signal that

can be easily read out with a data acquisition card. The SQUID current amplifier consists of a SQUID array coupled to the input and feedback coil. Any current in the coils generates magnetic field, which is then turned into a voltage signal via the SQUID array. See Ref. [38] for more details. The SQUID amplifier is installed at 1 K temperature in the ADR.

The entire SQUID readout system also contains room-temperature circuits. The wires from room temperature are connected to the SQUID array to provide a bias current, measure the voltage across the SQUID, and provide a feedback flux to the SQUID to linearize its output [39]. Figure 2.3 shows the scheme of the SQUID readout system. The dashed box indicates the components present on the SQUID chip. The parasitic resistance on the TES branch is represented by R_p , and parasitic resistance on the shunt branch is included in the total shunt resistance R_{sh} . L_{in} and L_{fb} are the self-inductance of the input coil and the feedback coil in the SQUID array, respectively. The SQUID array is run in a closed feedback loop, with the warm feedback electronics represented by a single amplifier block in Fig. 2.3 and with the feedback voltage converted to a current via the cold feedback resistor R_{fb} . This resistor is run at 1 K to ensure that it does not contribute additional Johnson current noise to the SQUID feedback coil. Finally, the current bias for the TES is provided by a warm voltage source across a cold bias resistor R_b , with similar noise considerations taken as for R_{fb} to ensure that excess voltage and current noise do not couple into the TES circuit via the bias lines. While not shown in this figure, there are additional filtering stages on the warm electronics to minimize ground loops and coupling of EMI back into the cryostat. The output voltage of the room-temperature circuit is proportional to the current in the input coil:

$$(2.3) \quad V_{output} = I_{input} \cdot \frac{R_{fb}L_{in}}{L_{fb}}$$

The main reason for not using a traditional transistor-based transimpedance amplifier is the requirement of low noise, low input impedance, \sim MHz bandwidth, and low power consumption. It is still possible to design a transistor-based transimpedance amplifier with good enough specifications, but using SQUID is a much more straightforward solution.

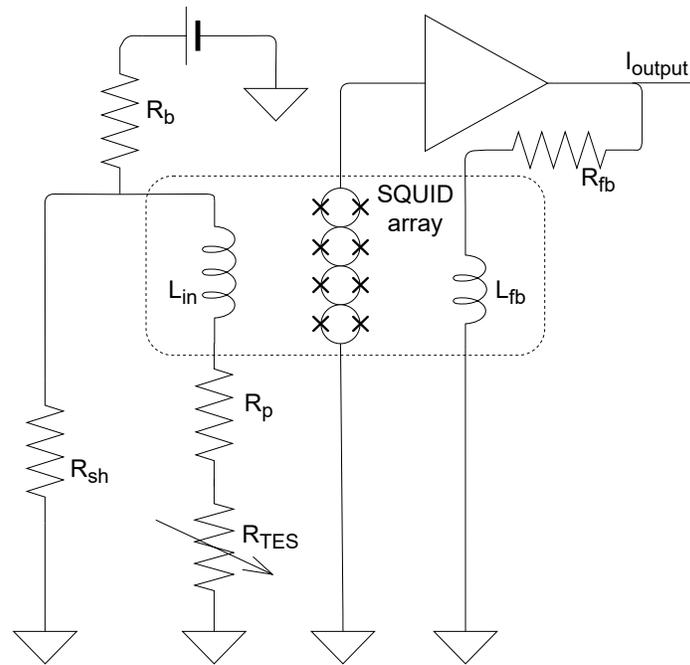


Figure 2.3. Scheme of the readout circuit for one TES channel represented by R_{TES} . The dashed box indicates the components present on the SQUID chip. The parasitic resistance on the TES branch is represented by R_p , and parasitic resistance on the shunt branch is included in the total shunt resistance R_{sh} . L_{in} and L_{fb} are the self-inductance of the input coil and the feedback coil in the SQUID array respectively. The SQUID array is run in a closed feedback loop, with the warm feedback electronics represented by a single amplifier block in the figure above, and with the feedback voltage converted to a current via the cold feedback resistor R_{fb} . This resistor is run at 1 K to ensure that it does not contribute additional Johnson current noise to the SQUID feedback coil. Finally, the current bias for the TES is provided by a warm voltage source across a cold bias resistor R_b , with similar noise considerations taken as for R_{fb} to ensure that excess voltage and current noise do not couple into the TES circuit via the bias lines. While not shown in this figure, there are additional filtering stages on the warm electronics to minimize ground loops and coupling of EMI back into the cryostat.

2.2.2. Trigger and filter

There is a lot of freedom in event triggering thanks to the continuously sampled timestream of the current flowing through each QET detector channel. One can tune the trigger algorithm for different purposes.

We apply a threshold trigger to the summation of all channels from one detector after filtering the summation with a finite impulse response (FIR) filter. The filtering process is to calculate the convolution between the timestream with the filter kernel. We mainly use two types of filter kernels: optimum filter (OF) for a lower threshold or Gaussian derivative filter (GF) for less artifact.

The OF is a minimum variance estimator of the amplitude of a pulse, with a known shape, in the presence of stationary noise, as described in Refs. [40, 41]. We use the OF in order to trigger with the lowest achievable threshold similar to Ref. [42].

The OF kernel is calculated with a pulse shape template and a noise power spectrum density (PSD). Several laser data sets—approximately equally distributed in time over operations—were used to construct the pulse template. Pulses coincident with the laser signal were collected by triggering on a digital signal from the laser driver. These events were averaged to produce the pulse template. The length of the template was optimized empirically to get the best energy resolution on the laser data and set to 16384 samples (10.8 ms). The noise PSD was evaluated by collecting noise traces of the same length, using a random trigger, and applying a pulse rejection algorithm to select pulse-free traces. This algorithm makes an iterative rejection of outlier events based on the mean, range, slope, and skewness of the traces. The outlier rejection procedure iteratively removes events furthest from the median of the distribution until the skewness of the distribution is less than 0.05.

After constructing both the noise PSD and the pulse template, a data stream was filtered, and a threshold trigger was applied to the filtered trace. A peak search window was defined, spanning 8192 samples after the crossing point for each threshold crossing occurrence. The trigger point was then adjusted to the point where the filtered trace reached its maximum value within the peak search window. A snippet of the raw trace within ± 8192 samples around the adjusted trigger point defines a triggered event which then undergoes further processing, where various event parameters are being evaluated, such as the template fit chi-square, the integral of the trace, the mean value of the pre-pulse region. An example of the trigger algorithm applied to a pulse can be seen in Fig. 2.4. A raw trace is shown in blue, while the optimum filtered trace is shown in orange. The dashed line is the trigger threshold. The green vertical bands show the trace regions selected by the triggering algorithms as events. The inset of Fig 2.4 shows a zoomed-in piece of the trace around a small pulse. The filtered trace reaches its maximum value at the pulse onset.

Although the OF has the best signal-to-noise (SNR) ratio — which means the lowest trigger threshold — under the assumption of known shape and stationary noise, it can have bad spurious responses in cases where the noise PSD has peaked. The spurious responses will create false triggers when an event isn't actually happening. This kind of trigger artifact can be removed during analysis but with the downside of making the analysis unnecessarily complicated. For analysis where the lowest threshold is not needed, we switch to the GF to avoid the artifacts.

The GF is a kind of filter using a kernel of the derivative of a Gaussian function. The GF is derived based on a different set of criteria and assumptions than OF, see Canny 1986 [43]. It does not assume a known pulse shape and requires minimum spurious responses that OF lacks. The σ of the Gaussian derivative kernel is tuned empirically to balance the SNR and the separation between events.

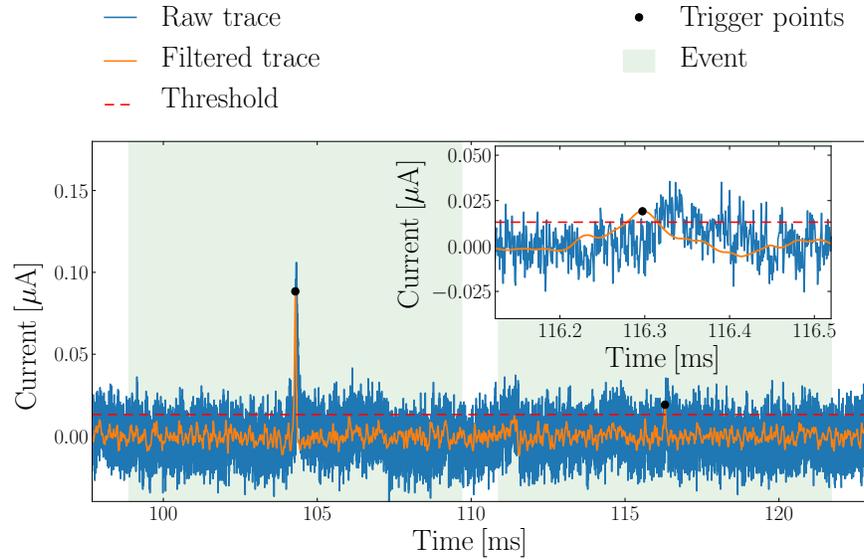


Figure 2.4. Illustration of the triggering algorithm. A raw trace is shown in blue, while the optimum filtered trace is shown in orange. The dashed line is the trigger threshold. The green vertical bands show the trace regions selected by the triggering algorithms as events. Inset: a zoomed in piece of the trace around a small pulse. The filtered trace reaches its maximum value at the pulse onset.

2.2.3. Integral-based energy estimator: “Matched Filter (MF) integral”

At energies below ~ 1 keV, the amplitude provided by the OF is used as an energy estimator to get the best possible resolution. However, TES saturation effects at higher energies distort the pulse shape, producing a large non-linear response and eventually saturating the OF estimator itself.

A hybrid of a pulse integral and a template fit was used to increase the dynamic range for high-energy analyses. The goal was to get the best estimation of the area of the pulse with a direct integral for the part where the pulse amplitude is high but distorted by saturation effects while using a fit to a pulse template to estimate the area where the detector behaves linearly but the signal to noise is low. We integrated the region where the pulse is above $2 \mu\text{A}$. The rest of the pulse was fit to a pulse template and then integrated from the $2 \mu\text{A}$ crossing to the end of

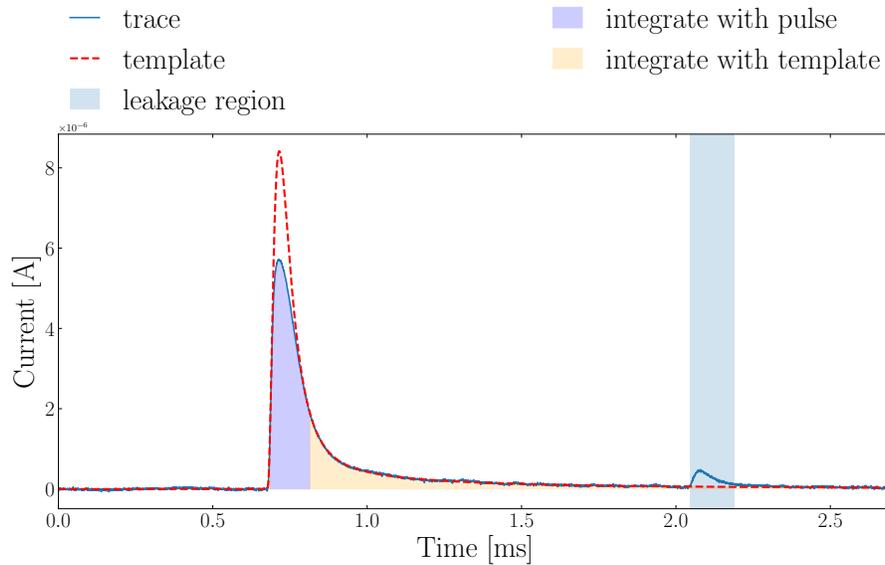


Figure 2.5. Illustration of the matched-filter-based energy estimator, MF integral, used to enlarge the energy range of the detector. The primary event is shown around 0.7 ms. The area highlighted in violet corresponds to the part integrated using the pulse itself. The area highlighted in yellow is integrated as the area below the red template, which is fit to the pulse tail in the yellow range. The pileup of the “leakage” pulse is identified through a threshold trigger and excluded from the tail fit to minimize its impact on the energy estimate of the primary pulse.

the pulse window. A 2.7-ms-length window was used, where the pre-pulse corresponds to 0.7 ms. The choice of a shorter trace with respect to the OF was dictated by a looser requirement for the energy resolution, which was in any case limited by the integral-based energy estimator. The $2 \mu\text{A}$ threshold was chosen as the level where the signal level is much higher than the noise level before the onset of heavy saturation.

The resulting estimator is functionally a hybrid of pure integration and a matched filter (MF), integrating the high signal to the noise region of the pulse directly and using the MF to estimate the contribution of the tail to the total pulse energy to reduce integrated noise. In addition, if there is a pileup pulse present between the $2 \mu\text{A}$ crossing and the end of the

trace, the pileup-pulse region is excluded from the fit. This exclusion region is defined as $10 \mu\text{s}$ before the pileup-pulse trigger to $130 \mu\text{s}$ after it, which is effective for preventing the dominant source of pileup pulses (single-electron-hole-pair leakage) from significantly affecting the fit. This energy estimator is referred to as the *MF integral* in the rest of the paper. Figure 2.5 illustrates the described procedure. The primary event is shown around 0.7 ms. The area highlighted in violet corresponds to the part integrated using the pulse itself. The area highlighted in yellow is integrated as the area below the red template, which is fit to the pulse tail in the yellow range. The pileup of the “leakage” pulse is identified through a threshold trigger and excluded from the tail fit to minimize its impact on the energy estimate of the primary pulse.

2.3. QET characterization

To validate the detector model, we measured various QET array properties such as bias power, energy efficiency and power noise. These measurements are key to understanding any differences between estimated and measured energy resolution. The measurements presented in this section are in good agreement with those predicted by our detector model, as shown in Table 2.1.

2.3.1. Resistance and Bias Power

Basic TES parameters can be evaluated by scanning through values of the TES bias voltage, V_b , and measuring the DC current response from the QET channel. In the first row of Fig. 2.6 (left) we show the variation of signal current I_s with bias voltage for both channels of NF-C. In the second row, we have calculated the inferred channel resistance [11], R , and in the final row the Joule power produced, $P = I_s^2 R$. We note that these scans were performed on the two channels simultaneously.

Parameter	Description	Unit	QP.4 [34]	NF-C			
				~50		~10	
T_b	Bath Temperature	mK	50				
	Design/Cryostat		ADR	Design	ADR	Design	NEXUS
A_{det}	Detector Area	cm^2		1			
T_c	TES Critical Temperature	mK	~65	60 – 65			
η	Detector Thickness	mm	1	4			
m_{det}	Detector Mass	g	0.24	0.96			
N_{QET}	QETs per Channel	-	170/300	504/536			
l_{TES}	TES length	μm	100	150			
v_{TES}	TES Volume (Per Channel)	μm^3	1360/2400	7.39×10^3			
ζ_{TES}	Fraction of W in TES	-	0.5	0.5			
v_{eff}	Effective W Volume	μm^3	2720/4800	1.48×10^4			
l_{fin}	Al Fin Length	μm	125	60			
ρ/η_{TES}	Resistivity/Thickness Ratio	Ω	2.88	2.88	3.0 ± 0.3	2.88	3.0 ± 0.3
	Normal Resistance (Inner)	$\text{m}\Omega$	400	350	332 – 396	350	332 – 396
R_n	Normal Resistance (Outer)	$\text{m}\Omega$	700	350	311 – 371	350	311 – 371
	Bias Power (Channel)	pW	1.2/2	4.6 – 8.4	4.0 ± 0.6	7.5 – 11.5	6 – 7.5
P_{TES}	Bias Power (Per TES)	fW	~7	8.8 – 16.0	7.6 ± 1	14.3 – 21.9	13 ± 2
G_{Chan}	Thermal Conductance (Channel)	pW/K	120/200	640 – 880	350 – 650	640 – 880	460 – 625
G_{TES}	Thermal Conductance (Per TES)	fW/K	225/375	1220 – 1680	660 – 1250	1220 – 1680	880 – 1190
Σ_{eph}	TES Electron-Phonon Coupling Constant	$\text{GW}/(\text{K}^5 \cdot \text{m}^3)$	0.47	0.47	$0.27 - 0.67$	0.47	$0.35 - 0.65$
τ_{BW}	Pulse Fall time	μs	~ 100	70 – 160	~ 80	55 – 100	~30
ϵ	Energy Efficiency	-	$\gtrsim 22\%$	27%	$\gtrsim 29\%$	27%	-
σ	Resolution	eV	3.0 ± 0.5	2.3 – 2.4	2.65 ± 0.02	1.8 – 2.1	~2.9 eV
S_p	NEP (Channel)	$\text{aW}/\sqrt{\text{Hz}}$	5.3	11 – 14	-	11 – 14	10
S_p	NEP (Per TES)	$\text{zW}/\sqrt{\text{Hz}}$	0.23	0.5 – 0.6	-	0.5 – 0.6	0.4

Table 2.1. Detector design parameters for the QP.4 prototype device (Ref. [34]) and the NF-C device described in this paper. Numbers for QP.4 are measured values from the previous reference, while numbers for NF-C are model predictions based on changes in the detector mask design. Both detectors have two channels, an inner grid of QETs surrounded by an outer frame of QETs. The outer channel in the QP.4 device had fewer QETs than the inner channel. Where multiple numbers are presented, the first/second number in the column is for the outer/inner channel on that device. The NEXUS analysis has not yet been extended to measure energy efficiency, as we are trying to improve the precision of the measurement, but the resolution implies it will be comparable to the efficiency found in the ADR. Similarly, we were unable to measure complex impedance in the ADR, so a precise estimation of NEP was not possible.

The bias power is affected by other heating effects like operating both channels simultaneously or changing the bath temperature. The reciprocal over-heating of the two channels lowers the required joule heating to stay in transition. Also, a higher bath temperature lowers the bias power needed to stay in transition. We note that the bias power is lower than was predicted by the NF-C model in the ADR measurement. For this reason, we repeated the measurement at NEXUS operating only one channel and at a lower bath temperature in Fig. 2.6 (left). The contribution of reciprocal heating was observed to be 1 to 2 pW at NEXUS. The bias power measured at NEXUS is then used in Sec. 2.3.3 for the noise modeling.

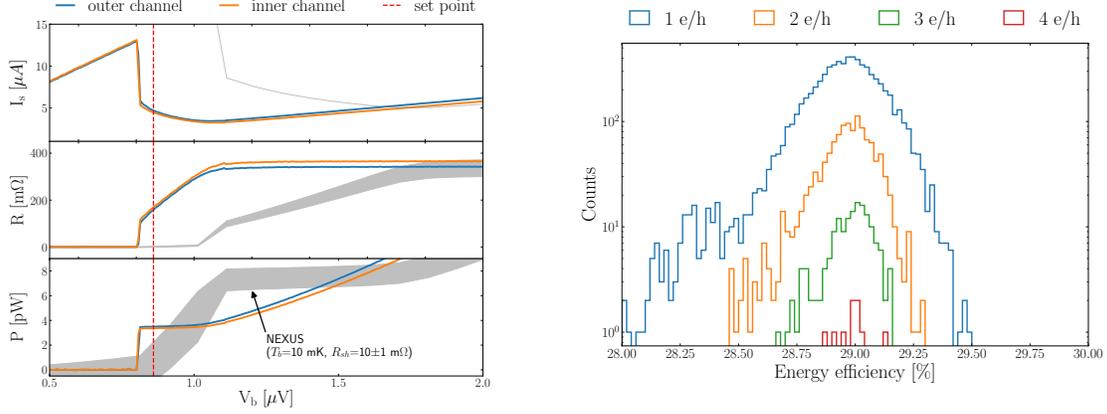


Figure 2.6. Left panel: TES signal current, resistance and power curves as a function of bias voltage $V_b = I_b R_{sh}$ for the two channels of NF-C operated at 50 mK in the Vericold ADR. In addition, the same variables are measured with a single channel operated at 10 mK in the NEXUS cryostat. The thickness of all the curves represents the associated systematic uncertainty band. The band is larger in the NEXUS data because of its lower value of shunt resistor ($R_{sh} \approx 10$ m Ω) and associated larger systematics uncertainty compared to the 50 m Ω shunt resistor used in the ADR. The detector was operated at 45% of the normal resistance value when operated in the ADR, and at 30% of the normal resistance during the measurements at NEXUS. Right panel: Reconstructed energy efficiency for different numbers of electron-hole pairs, see the text for details on the circuit parameter used.

2.3.2. Energy Efficiency

For a TES in strong feedback [11], the phonon energy absorbed by a TES can be inferred from the change in signal current and circuit parameters as

$$(2.4) \quad E_{abs} \approx \left(1 - 2 \frac{R_\ell}{R_\ell + R_0}\right) I_b R_{sh} \int \delta I_s(t) dt + R_\ell \int \delta I_s^2(t) dt$$

where $R_\ell = R_{sh} + R_p$ is the total resistance (shunt and parasitic) in the TES bias loop [11] apart from the TES, R_0 is the TES operating resistance, and $I_b = V_b/R_{sh}$ is the TES bias current. Here we have defined $\delta I_s(t) = I_0 - I_s(t) > 0$ as the change in signal current during a

phonon pulse relative to the quiescent value, I_0 . This absorbed energy can be compared to the calibrated total phonon energy to define the detector’s energy efficiency, $\epsilon = E_{abs}/E_{ph}$.

The efficiency was evaluated using a laser calibration dataset with a mean number of photons per pulse $\lambda \sim 0.3$, the detector operated at $V_{N\text{TL}} = 100$ V, and the cryostat temperature stable at 50.00 ± 0.01 mK. Data selection criteria were applied to select pulses which were coincident with the laser trigger signal, had energy above the noise threshold, and had a stable baseline signal before the pulse.

Figure 2.6 (right) shows the energy collection efficiency that was calculated for individual phonon pulses using a particular set of circuit parameters. For this figure, we selected the most conservative set of assumptions to obtain a lower estimate of the energy collection efficiency of $\epsilon \gtrsim 29\%$. As reported in Table 2.1 and detailed in Appendix A.2, this is compatible with design expectations. The current measurement is dominated by the systematic uncertainties in TES circuit parameters (e.g. R_p and R_0); future measurements will include more precise characterization of these components to place tighter constraints on this value.

2.3.3. Noise Modeling

The resolution model for a QET described in Section 2.1 relies on the assumption that the QET noise is dominated only by thermal fluctuations across the thermal conductance G between the TES and the crystal. In reality, the bias circuit has its own intrinsic noise from both passive components and the SQUID current amplifier. Optimization of the detector normal resistance takes these expected contributions into account to ensure that the TES is dominated by its own quantum noise. Modeling the current noise, and converting to Noise Equivalent Power (NEP), allows us to compare the intrinsic power noise of the QET to that expected by the resolution model. The NEP for a generic thermal detector with thermal conductance G at temperature

$T = T_c$ is [11]

$$(2.5) \quad NEP = \sqrt{4k_b T_c^2 G}$$

and thus we can compare the noise power inferred from the current noise to the expectation from the measured bias power and transition temperature, which predicts the magnitude of the thermal fluctuation noise and is expected to be flat in NEP. The NEP expected for these detectors is summarized in Table 2.1.

In order to validate this noise model and demonstrate that this detector achieved near quantum-limited noise, we employed the TES bias circuit noise model described in past work (see e.g. Refs. [15, 13, 12, 11]). Due to the less constrained input inductance and parasitic resistance parameters on the ADR electronics circuit, we carried out a dedicated measurement at NEXUS. We characterized the noise inherent to the SQUID bias circuit using a SQUID with the TES coil disconnected. We then fit the contribution of passive noise to the total transition noise by adjusting the effective noise temperature of the fit to jointly match the noise in the normal and superconducting states. We also measured the complex impedance of the TES both with a square wave impulse and swept sine wave measurements to characterize the TES thermal poles, with results summarized in Table 2.2. The superconducting noise combined with complex impedance measurements constrained the inductance in the loop. In addition, we were able to extract estimates of TES response characteristics in Table 2.2 (similar to the method used in Ref. [15]) to constrain the TES power to current response. The measurements of bias power in the lower temperature environment in NEXUS allows us to bound thermal conductance and better constrain the parameters in Table 2.1.

The measured current noise for a single QET channel at the operating bias point of $R_0/R_n = 0.43$ is shown in Fig. 2.7 (left), along with the model incorporating systematic uncertainties, demonstrating that the TES response is dominated by the quantum (thermal fluctuation) noise.

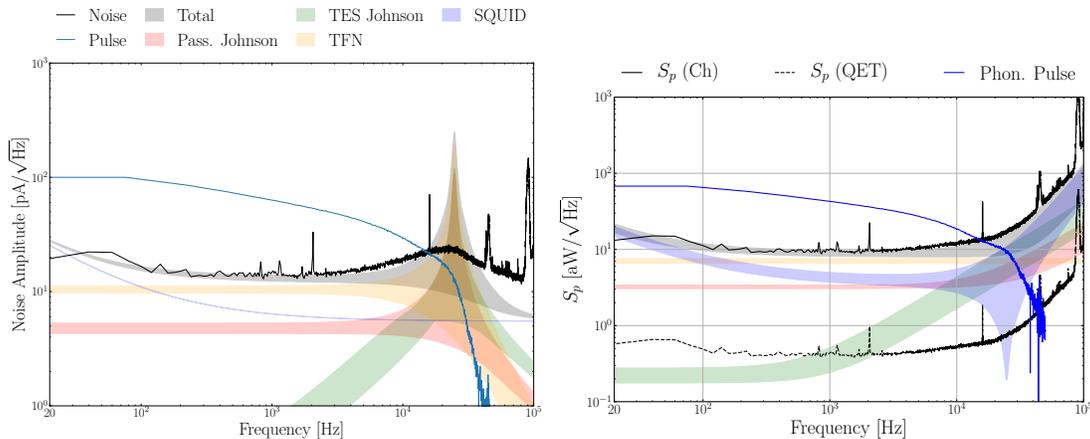


Figure 2.7. Left panel: Current noise for NF-C run in NEXUS (black) compared to the best-fit model informed by complex impedance measurements taken in the same facility, highlighting the dominance of thermal fluctuation noise (TFN) assumed for the detector modeling. The pulse shape found by averaging pulses near threshold is also shown. The pulse shape is scaled arbitrarily relative to noise to better visualize atop the noise. Right panel: Power noise inferred from noise modeling, computed by dividing the current noise by the power to current transfer function derived from the complex impedance measurement [15]. The total QET channel power noise, as well as the noise per individual QET cell are shown in black, compared to the pulse shape (blue) in power space. In both cases, the closed-loop SQUID gain begins to drop around 50 kHz, where the phonon pulse is cutoff. This also artificially broadens the electrothermal oscillation peak at ~ 25 kHz.

At high frequency, the signal to noise was degraded by electrothermal oscillation due to the high inductance of the readout system (~ 800 nH), which impacts both the QET pulse and the noise. Dividing out the electrical response of the TES bias loop using complex impedance measurements gives the estimates for noise equivalent power in Fig. 2.7 (right). With around 525 QETs/channel, we obtain a total power noise of ~ 10 aW/ \sqrt{Hz} , which is equivalent to 500 zW/ \sqrt{Hz} per individual QET cell. This is consistent with the NEP used to estimate TES resolution in Eq. 1.4.

The large error bands in the noise model come from the same source of systematic uncertainty as for energy efficiency, namely the uncertainty in overall resistance scale. This becomes

Table 2.2. TES bias circuit parameters measured at NEXUS, which were extracted from complex impedance measurements used to fit TES noise in Fig. 2.7 (parameters refer to the definitions employed in Ref. [15]). Fall time and feedback gain in electrothermal feedback (ETF) parameterize the effect of the voltage bias feedback on the TES response. The reported uncertainties are dominated by the systematic uncertainty on the shunt resistor and affected by the limited bandwidth of the readout circuit. Distortions in the driving signal above 10 kHz required a correction for finite bandwidth in the bias circuit. Higher precision measurements will better constrain these parameters as a function of bias point and base temperature in future work.

Parameter	Description	Value
L	Inductance	850 ± 50 nH
R_{sh}	Shunt Resistance	8 ± 1 m Ω
R_p	Parasitic Resistance	19 ± 2 m Ω
R_0/R_n	Bias Point	0.43 ± 0.01
R_0	TES Resistance	125 ± 25 m Ω
$ \tau_{ETF} $	ETF Fall time	$7-8$ μ s
τ_0	Thermal Fall time	200 ± 50 μ s
\mathcal{L}	ETF Gain	30 ± 5
β	Current Response	0.2-0.3

a systematic uncertainty on bias power, leading to a large range in the measurement for G , but is also degenerate with measurements of inductance. In addition, some uncertainty comes from the limited bandwidth of the measurement technique used for the data taken in this run. Future measurements will further constrain QET properties by carefully calibrating out these uncertainties and by improving the precision of the complex impedance characterization. In particular, four-wire measurements of the detector R_n will reduce the large systematic uncertainties on the resistance scale, which dominates the uncertainty of all measurements described in this paper.

2.4. Detector performance: small signal response

The results of the Section 2.3 suggest that the parameters that feed into the energy resolution estimate match expectation, and thus we should find the energy resolution to be close to the

design expectation. In this section, we report a measured baseline resolution¹ comparable to the design value—2.65(2) eV compared to 2.3 – 2.4 eV expectation—and explore the small signal response.

Calibration of the low-energy region (below ~ 1 keV) is performed with laser data sets as described previously. The single-charge resolution leads to discrete peaks in the spectrum corresponding to quantized charge excitation. This produces a set of well defined lines of known energy that can be used for calibration.

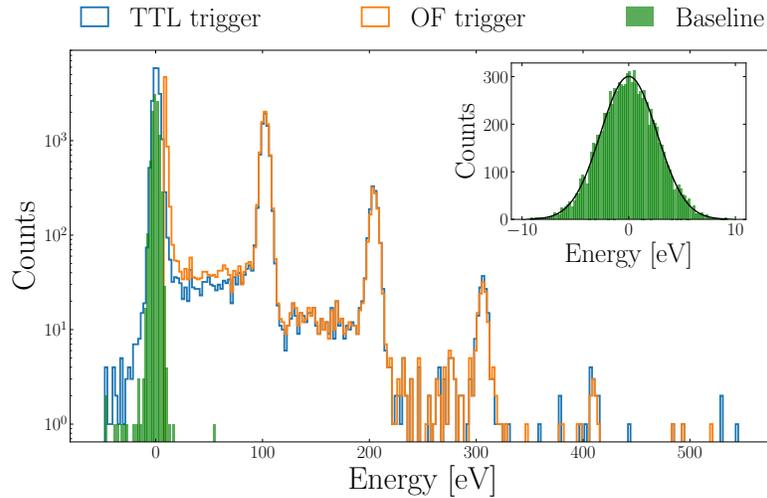


Figure 2.8. Laser distribution acquired with 100 V NTL bias. The data are triggered with the digital laser signal (TTL signal) and with the OF trigger. An energy resolution $\sigma = 3.25(4)$ eV was measured at the first electron-hole pair peak. Inset: Zoom of the green histogram, which represents the random triggers used to estimate the baseline resolution. A baseline resolution $\sigma = 2.65(2)$ eV was measured from random triggers. The Gaussian fit is shown with a black line. The discrepancy between the baseline and peak resolution is due to additional variance from absorption of photons in the QETs [34].

Figure 2.8 shows the energy distribution of a laser calibration dataset in which the average number of photons λ absorbed in the Si substrate is of order 1 ($\lambda \sim 1$). The statistics of the

¹We refer to baseline resolution as the detector energy resolution when no pulses are recorded.

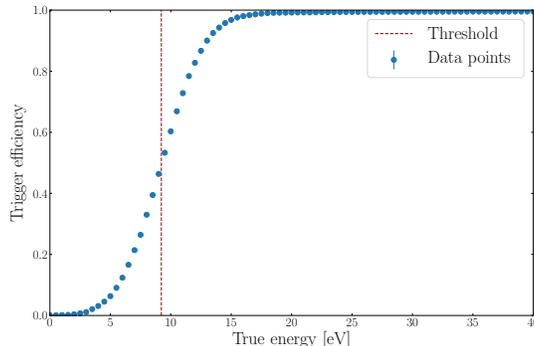


Figure 2.9. Trigger efficiency, measured by injecting the pulse template into randomly triggered noise traces. The dashed line shows the 9.2 eV threshold set on the trigger energy estimator. Uncertainties on the data points are included in the figure, but are not visible at this scale due to the high statistics of the laser data used to determine the trigger efficiency.

dataset shown is large enough to extend the calibration to the fourth electron-hole pair peak, corresponding to a maximum energy of ~ 400 eV.

The fill-in between the laser peaks can be explained via both charge trapping and impact ionization as charges propagate across the crystal [20]. In the former case, a charge is trapped in the crystal lattice, reducing the amount of phonon energy produced by shortening the drift length through the crystal. In the latter case, a charge kicks off a second loosely-bound unpaired charge increasing the total amount of energy collected. The charge trapping and impact ionization probabilities were evaluated for this detector by fitting the laser data with the model described in Ref. [20]. From these fits, we obtain a charge trapping of 12.7% and an impact ionization of 0.6% with the data acquired at TUNL.

We employed the OF estimator described in Section 2.2.2 to evaluate the detector performance in the linear regime of the detector. We obtained a phonon energy resolution at the first electron-hole pair peak (corresponding to 101.95 eV for a NTL bias of 100 V) of 3.25(4) eV, which corresponds to a charge resolution at the level of $\sim 3\%$ at 100 V bias. The measured

phonon energy resolution was observed to be independent with respect to the applied NTL voltage below the point at which charge leakage begins to increase exponentially, as discussed in Refs. [44, 45].

The baseline energy resolution was evaluated from a set of pseudo-random triggers on 0 V data. The amplitude was evaluated at the random trigger position with an optimum filter-based estimator without allowing the algorithm to search for the maximum. A Gaussian fit results in a reconstructed energy resolution of $\sigma = 2.65(2)$ eV, which is very close to the value predicted in Sec. 2.1. We see a discrepancy between the baseline resolution and the resolution at the first electron-hole-pair peak, implying an additional source of energy smearing in the latter. This is likely due to surface absorption in the QETs [34].

The QET direct absorption is a known effect, and both the offset and variance of the laser peaks have been shown to correlate with the laser intensity [34]. For 100 V laser data with $\lambda \sim 1$, the expected energy shift is of the order of 0.9 eV, which corresponds to less than a 1% effect on the position of the first electron-hole-pair peak. This effect is taken into account during the calibration using laser data and thus will not impact reconstruction of events caused by a single bulk energy deposition.

The trigger efficiency was studied by injecting pulses into randomly triggered noise traces. The OF pulse template, which is the averaged laser pulse, was used as the shape of the injected pulses. A trigger time cut around the expected position of the injected pulse ($\pm 3\sigma$ of the timing resolution which is equal to $\sigma = 440$ ns for 15 eV events) ensures the correspondence between the injected pulse and the triggered one. The efficiency was calculated as the fraction of injected pulses that were triggered by the OF trigger algorithm. We achieved a threshold of 9.2 eV, which corresponds to 3.5σ of the baseline resolution, while maintaining the trigger rate as low as 20 Hz. The resulting efficiency curve is shown in Fig. 2.9.

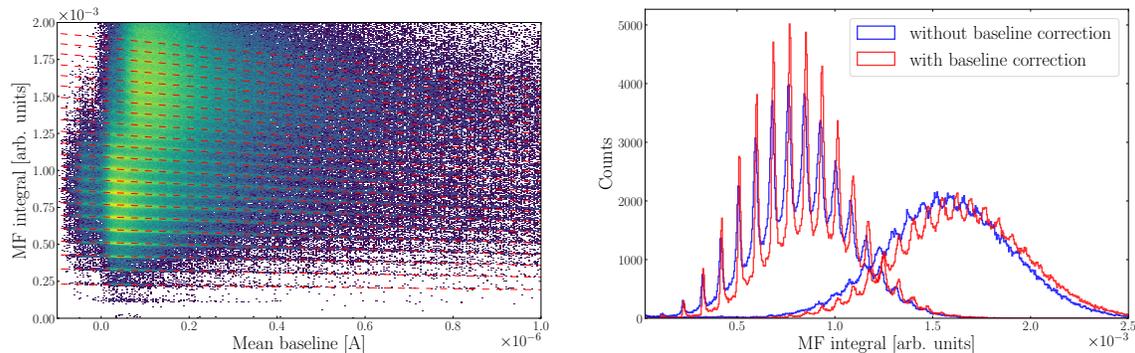


Figure 2.10. Left panel: 2D histogram of the matched-filter-based energy estimator as a function of the mean baseline for the laser data. Two data series with a different average number of photons were used, which is visible by the two clusters at low mean-baseline values. The red lines highlight the trend of the detector energy as a function of the mean baseline. Right panel: Laser spectrum before and after the mean-baseline gain correction for the two laser data series used in the correction.

2.5. Detector performance: extending the dynamic range

As the energy deposition increases, the TES gradually saturates until a point the pulse amplitude no longer increases. Instead, the pulse duration becomes longer. In this highly-saturated region, the OF, which is an estimation of the amplitude, no longer have good sensitivity. By using the MF integral estimator, we can extend the energy scale to 120 keV, corresponding to an effective dynamic range of 4 orders of magnitude.

Calibration of energies above the nominal linear region of ~ 1 keV was accomplished by combining: (1) high-intensity laser data, up to 6 keV at HV bias; (2) an ^{55}Fe source, which extends the calibration up to 120 keV by applying a voltage bias of 70 V; (3) data taken with a ^{57}Co source without NTL bias. Past work demonstrated that laser data can be used to calibrate energies below 700 eV [44, 30, 45], as shown in the previous section, and other groups have used the ~ 6 keV double peak from ^{55}Fe , and associated 1.5 keV Al fluorescence, to calibrate the detector energy scale above 1 keV [46]. Here we demonstrate, for the first time, a combined

approach to linearize the energy scale across four orders of magnitude in energy, combining the low-amplitude linear response region with the high-amplitude saturation region of the QET channels.

The first step was to model the response of the detector to the laser calibration signal at higher average photon number. The number of photons emitted by the laser was Poisson-distributed and was controlled by increasing the laser excitation current. The number of events populating the peaks (which are then used for the calibration) were reduced, due to the charge trapping and the impact ionization effects mentioned in the previous section. A longer acquisition time (~ 5 hours) and a high laser rate (~ 101 Hz) were used to collect sufficient statistics for this first calibration step.

The high pulse rate, combined with a non-shielded cryostat operated in an above-ground facility, greatly increased the probability of pileup pulses. This caused the working point of the detector to shift, leading to a reduction in pulse height for a given energy deposit. The mean pre-pulse baseline, defined as the average value of 900-samples in the pre-pulse trace, directly measured the detector bias current, and was used to correct for this gain variation [47].

Figure 2.10 (left) shows the reconstructed pulse amplitude as a function of mean baseline for two data sets of laser data, demonstrating that the MF integral of each peak decreases as the mean baseline value increases. The correlation between laser peak positions in the mean baseline and amplitude plane has been approximated with a linear function and is shown in red for each electron-hole-pair peak in Fig. 2.10. The mean-baseline correction was achieved by rotating the red lines around the zero-point on the mean-baseline axis, corresponding to the nominal detector baseline level.

We rejected events above $1 \mu\text{A}$ in mean-baseline, limiting ourselves to the linear regime of this dependence and neglecting small nonlinear effects that are appreciable only over a larger mean-baseline range beyond $1 \mu\text{A}$. Figure 2.10 (bottom) shows the laser spectrum before and

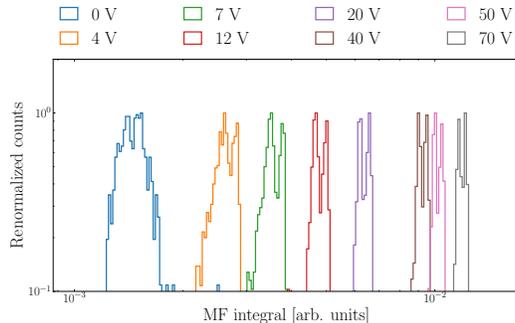


Figure 2.11. ^{55}Fe distribution for different NTL bias; the distributions are renormalized by the maximum of the distribution. The two-peak structure corresponds to the ^{55}Mn K_α and K_β X-rays at 5.9 and 6.5 keV, respectively

after this correction; the improvement in the energy resolution and peak definition is evident. These laser data, acquired with NTL bias of 100 V and 250 V, provided a calibration up to 6 keV by using the first 24 peaks.

The calibration at high energy used an external ^{55}Fe source, which emits two ^{55}Mn X-rays at 5.9 and 6.5 keV. The data were acquired at eight different NTL biases in order to uniformly cover the energy region between 6 and 120 keV. Figure 2.11 shows the measured ^{55}Fe energy distributions used for this calibration. The use of a source outside of the cryostat produced an unusual event ratio between the K_α and K_β lines caused by a decrease in attenuation of the X-rays with increasing energy.

We also incorporated the trapping and impact ionization effect to model the expected energy distribution of these peaks at high voltage. In the many charge limit, charge trapping and impact ionization effects can be included in the energy calibration using the relation:

$$(2.6) \quad E_{ph} = E_r \cdot G_{\text{NTL}}(1 - 0.5 \cdot P_{\text{CT}} + 0.5 \cdot P_{\text{II}}),$$

where P_{II} and P_{CT} are the impact-ionization and charge-trapping probabilities, $G_{\text{NTL}} = 1 + e \cdot V_{\text{NTL}}/\epsilon_\gamma$ is the NTL gain and E_{ph} and E_r are the final phonon energy and the initial recoil

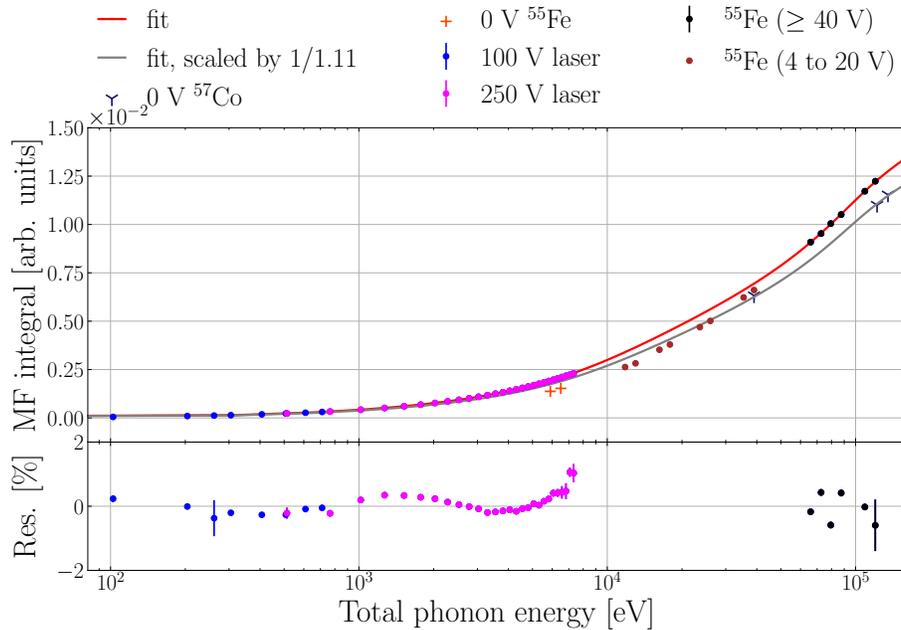


Figure 2.12. Combined calibration, including laser and ^{55}Fe data. Uncertainty on each point is included in this figure, but is small enough that it is not visible in the upper panel; error bars for some points become appreciable in the residual calculation, as seen in the lower panel.

energy. The factor 0.5 assumes that the charge trapping and impact ionization occur evenly across the detector. We expected a decrease in the energy scale of the order of 4.5% by using the probabilities measured by the fit, as discussed in the previous section. This factor was included in the final energy calibration in Fig. 2.12. It was relevant for both calibration and background data at high voltage. This correction to the energy scale assumes that the trapping and impact ionization at the detector surface are the same as in the volume.

One finding from these data post-calibration was a mismatch between the calibration obtained with the laser source at high voltage and the calibration obtained with the ^{55}Fe source at low voltage. The most likely mechanisms which could account for this discrepancy are: (1) the NTL phonons have a different response with respect to the phonons generated by charge recombination; (2) the penetration length of X-rays in Si ($\sim 30 \mu\text{m}$) is not sufficient to reach the

bulk and there is some signal degradation due to surface effects; (3) the deposition of a single X-ray could generate local saturation in the sensor, because the ^{55}Fe source was directed at the QET-instrumented face. The temperature distribution of the individual QETs can be strongly nonuniform due to a near-surface energy deposition causing only those QETs in the local vicinity of the deposition to saturate.

In the high-electric-field regime, (1) the charges are quickly drifted to the detector bulk, and (2) the phonon signal is dominated by NTL amplification such that the original energy deposition is negligible in comparison. This second point ensures that the phonon distributions with laser and ^{55}Fe events are produced by the same mechanism. For these reasons, we only included the ^{55}Fe calibration for nonzero voltage bias when the NTL effect accounts for more than 90% of the expected phonon energy, corresponding to voltage biases in the range 40–70 V. In this limit, a smooth energy reconstruction was possible for data acquired in the presence of strong NTL amplification.

We used an external ^{57}Co source with its two gamma rays at 122 keV and 136 keV and the 39-keV Compton edge in order to calibrate the data acquired without the NTL amplification. The 0 V calibration data were fit using the same curve shape as for the HV data—a sixth order polynomial—multiplied by a scaling factor. The scaling factor was extracted from the fit and corresponds to 1/1.11. This curve is represented in Fig. 2.12 in gray. The ^{55}Fe data acquired at 0 V were not compatible with this curve, we suspect that this is due to the aforementioned local saturation and surface effects.

Figure 2.13 shows the variation of the energy resolution as a function of the energy of the event. The OF energy estimator demonstrated an energy resolution of 3.25(4) eV at 101.95 eV, as discussed in the previous section. The MF integral trades energy resolution for dynamic range, allowing us to probe much higher energies while maintaining an energy resolution lower than 5%.

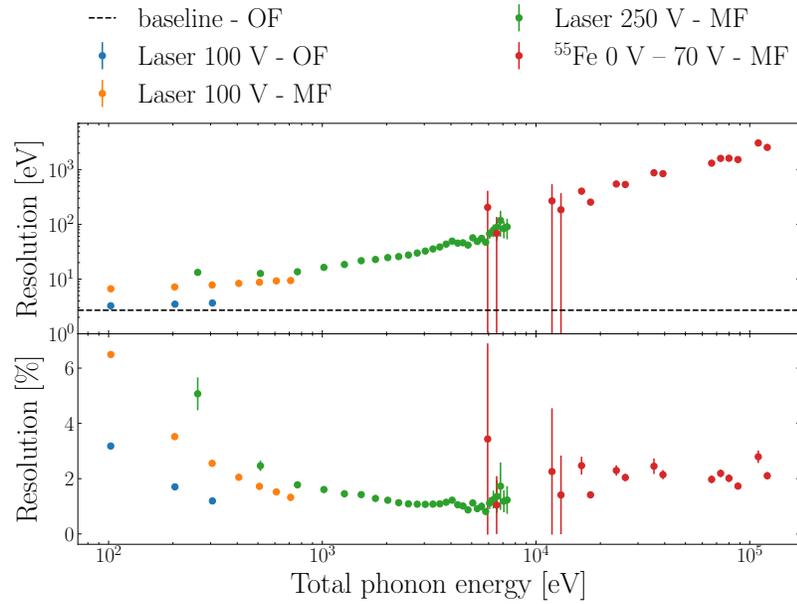


Figure 2.13. Energy resolution expressed as a function of the energy both for the OF and the MF energy estimators. The OF processing achieved the best energy resolution of $\sigma = 3.25(4)$ eV for low energies. The MF estimator allowed us to obtain a fractional energy resolution less than 5% for energies above 1 keV.

CHAPTER 3

Measuring Ionization Yield of Silicon

This chapter is being prepared for publication. Printed with permission of SuperCDMS collaboration and APS.

3.1. Introduction

The identity of dark matter and determination of neutrino properties are problems at the forefront of physics beyond the Standard Model. Rare event searches focused on dark matter detection [48, 49, 50, 51, 52, 53] or coherent neutrino nucleus scattering (CE ν NS) [54, 55, 56, 57, 58, 59, 60, 61] often detect the products generated by a particle interacting in a target material, making a strong understanding of that material's response to energy depositions vital.

Silicon is a commonly used target material. Particle interactions with the silicon nuclei or electrons generate ionization, with nuclear recoils generating less ionization than electron recoils of the same energy. The relative ratio of the ionization produced, called the ionization yield (also known as quenching factor) Y defined in Eqn 1.17, is crucial to understanding the response of such detectors and is believed to be an intrinsic material property.

Nuclear scattering can generate damage effects in the matter, such as the dislocation of ion in the lattice (Frenkel defect ??) or the subsequent ionization. Nuclear scattering, which on the atomic level is a screened Coulomb potential scattering, is not easy to solve even for a single recoil. In high-energy limit, Lindhard established a model [62] for the average damage effect of nuclear scattering, including Y . The model has been actively used for over 50 years. The

original book from Lindhard [62] may be hard to read, Ref. [63] contains a good summary of the Lindhard model.

Experimental measurements of Y in silicon for nuclear recoils above 4 keV (denoted keV_{nr}) have been consistent with the Lindhard model [64, 65, 66]. However, in the sub-keV_{nr} range, which is typical in low mass dark matter or CE ν NS searches, recent measurements of Y indicate a significant deviation from the Lindhard model [67, 68], see Fig. 3.1. Furthermore, measurements of the low energy yield in another commonly used semiconductor, germanium, have been inconsistent with each other [69, 70], hinting at the possibility of an electric field or temperature dependence in the yield.

On the other hand, recent modeling [71, 63] has focused on understanding the origin of these deviations. These models use Lindhard model as a starting point and remove some approximations in the Lindhard model (mainly the binding energy). The models lead to lower-than-Lindhard yield at the sub-keV region, and predict an ionization threshold of $\mathcal{O}(100)$ eV. If the models are correct, it is bad news for low-threshold experiments using ionization signals from Si/Ge, because the sensitivity to low-energy events will be zero below the ionization threshold.

These observations and the recent models cement the need for a range of ionization yield measurements. SuperCDMS initiated an ionization yield measurement program called Ionization Measurement with Phonons At Cryogenic Temperatures (IMPACT). Multiple measurements are foreseen, including different material targets (silicon and germanium) and different neutron energies (including but not limited to neutrons produced with a DD neutron generator at Fermilab or a proton beam at Triangle Universities Nuclear Laboratory (TUNL)). This work reports the first campaign of IMPACT, which investigated the ionization yield for a silicon detector by using a 56-keV neutron beam produced at TUNL.

The main goal of this measurement is to measure the ionization yield of silicon from 100 eV_{nr} to 4 keV_{nr}. The lower limit roughly corresponds to the lower end of the SuperCDMS SNOLab

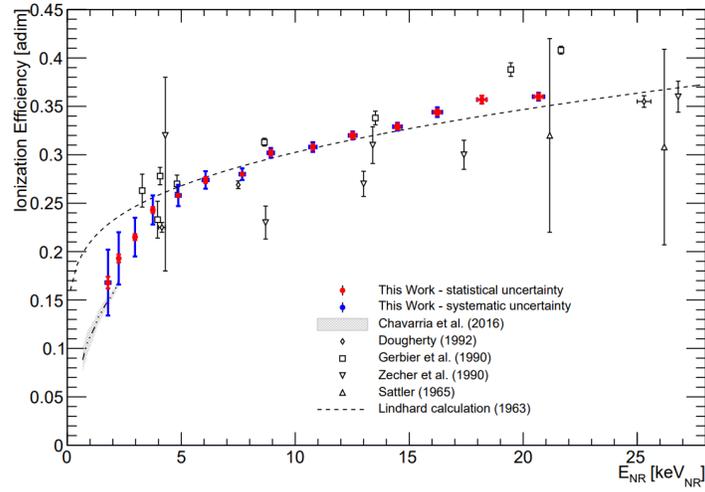


Figure 3.1. Existing ionization yield measurements and Lindhard model. This plot was taken from F. Izraelevitch et al. [68].

low-mass projection limit, while the higher end extends to regions where the existing data agree with the Lindhard model. The energy interval from 100 eV to 800 eV is a new parameter space for the ionization yield in silicon.

3.2. Experimental set-up

IMPACT measures the ionization yield based on a neutron scattering setup. When a neutron scatters on the nucleus, the recoil energy can be calculated from the neutron scattering angle according to the following equation:

$$(3.1) \quad \Delta E = 2E_n \frac{M_n^2}{(M_n + M_T)^2} \left(\frac{M_n}{M_T} + \sin^2\theta - \cos\theta \sqrt{\left(\frac{M_n}{M_T}\right)^2 - \sin^2\theta} \right),$$

where E_n is the energy of the incident neutron (56 keV in our case), M_n and M_T are the masses of the neutron and the target, and θ is the scattering angle. With a mono-energetic neutron beam, the recoil energy solely depends on the scattering angle. Thus, we can select events of particular recoil energy by using neutron tagging detectors at fixed angles. The ionization

yield can then be calculated with events that are happening in both the HVeV and the neutron tagging detectors by using Eqn. 1.17

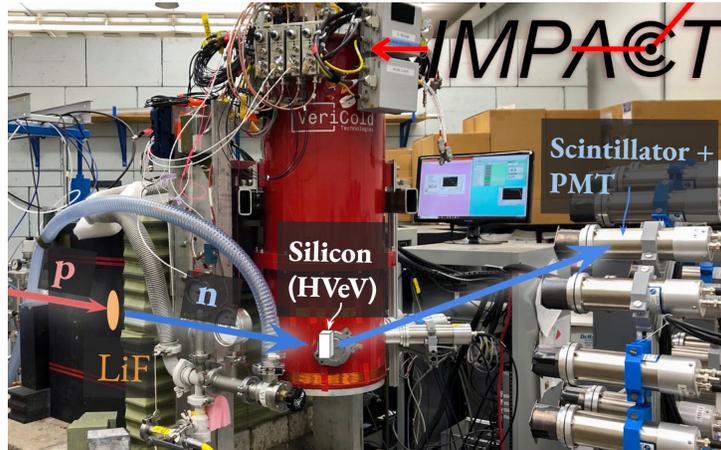


Figure 3.2. A photo of the IMPACT@TUNL experimental. The silicon HVeV detector is depicted with a grey cube. The position of the PMT detector and neutron beam is pointed out.

Figure 3.2 shows a photo of this measurement. Multiple different systems needed to be operated in this measurement to obtain this result. I will skip the HVeV detector that is already thoroughly described in Chapter 2. Sec. 3.2.1 describes the neutron beam; Sec. 3.2.2 introduces the neutron tagging detectors. Sec. 3.2.3 describes the alignment between the beam, the silicon detector, and the liquid scintillator detectors.

3.2.1. The neutron beam

The neutron beam at TUNL is generated with a pulsed proton beam, which is directed onto a 100 nm-thick lithium fluoride (LiF) target. The subsequent ${}^7\text{Li}(p,n){}^7\text{Be}$ reaction produces a nearly mono-energetic neutron beam with a controllable energy after collimation [72, 73]. The neutron beam is pulsed at 400 ns intervals. A beam pickup monitor (BPM) installed upstream from the target provides a nanosecond-resolution time measurement of the beam. Together with

the neutron tagging detectors that is introduced in Sec. 3.2.2, the neutron time of flight (TOF) can be measured.

The neutron energy is tuned by the proton energy, which is set to 1.889 MeV to generate 55.7 keV neutrons for this measurement. The particular neutron energy is chosen to take advantage of the narrow resonance of the silicon-to-neutron cross-section, which consequently maximizes the number of events detected by the silicon detector. Fig 3.3 shows the neutron-to-silicon cross-section as a function of the neutron energy, showing a resonance at 55.7 keV. This resonance keeps the energy spread narrow and increases the event rate.

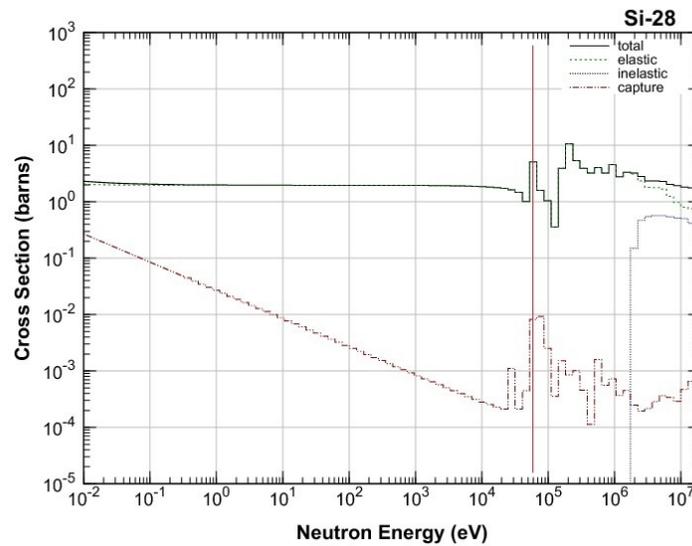


Figure 3.3. Neutron to silicon cross-section [74]. The vertical red line highlights the 56 keV neutron energy.

The spread of the neutron energy also increases with respect to the thickness of the LiF target. Therefore, we would like to have a thin target. A thinner LiF target also lowers the gamma background. On the other hand, the event rate will decrease when the target is too thin. We have reached a compromise between various effects and chose a thickness of 100 nm. An

initial measurement (performed before the TUNL campaign) and calculations indicate a neutron production rate of $23 \text{ /cm}^2\text{/s}$ at 70 cm from the target with an average beam current of 400 nA.

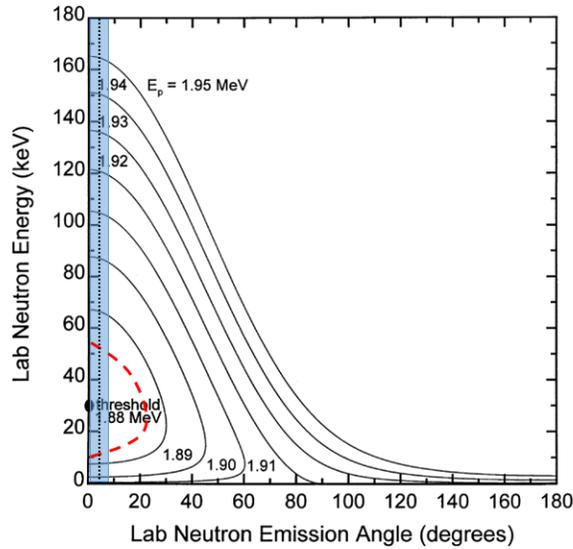


Fig. 1. Proton energy contours for a thick lithium target.

Figure 3.4. This plot was taken from Ref. [75]. Neutron energy as a function of the neutron emission angle for a thick LiF target. The contours represent curves with a constant proton energy. The light-blue area shows the angular region in which we are interested that is limited by the collimator. We added an additional contour in red to highlight the two energies that we are expecting given the proton energy in use. The vertical black dotted line intersect the contour curves and is meant to highlight the two different energy that we expect to see in the spectrum.

Another feature of the neutron is that it is not mono-energetic, even if the beam is perfectly stable. This effect is explained in Fig. 3.4 taken from Ref. [75] where neutron energy is expressed as a function of the neutron emission angle. For a fixed given angle, we expect two neutron energies when the proton energy is above the reaction threshold, and the neutron emission angle is below 60 degrees. In this measurement, the neutrons' angular distribution is constrained by the collimator to within 4 degrees. The light-blue area highlights the region of interest considering the effect of the collimator. The red dashed contour is added by hand (not precise)

to mimic the contour of 1.88 MeV incoming proton energy. The intersections between the red contour and the vertical black dotted line show the two solutions that we expect at a given fixed angle, corresponding to about 56 keV and 11 keV. A more detailed simulation of the neutron beam energy distribution that goes through the detector for our proton beam energy is reported in Sec. 3.3.1.2.

3.2.2. The neutron tagging detectors (PMT)

Neutron tagging detectors are used to measure the scattering angle of events. They are fixed at six different angles that we planned to measure. If an event happens in both the neutron tagging detector and the silicon HVeV, we can know its recoil energy from the angle.

Neutron tagging detectors we use are 2-inch liquid scintillator units coupled to 2-inch photomultiplier tubes (PMTs). The liquid scintillators are EJ309 and EJ301 produced by Eljen, and are designed for neutron measurement. For simplicity, in this dissertation, I will refer to the liquid scintillator neutron tagging detector as the "PMT" detector.

The PMT detector and the aforementioned BPM share the same data acquisition system (SIS 3316) running at 250 Msps or 4 ns per sample. The PMT DAQ is hardware triggered, and only the traces triggered were recorded for offline processing.

The PMTs were divided in two main groups:

- **the lone-wolf (LW) detectors:** 3 PMTs were fixed on a mechanical arm at the side of the ADR. These detectors will be referred as LW1, LW2 and LW3, and are tagging around 750 eV, 2000 eV, and 3870 eV of recoil angles for 56 keV neutrons.
- **the backing array:** 26 detectors were arranged in two concentric rings on a movable structure. We will refer the two rings as the inner and outer ring. The inner ring included 8 PMTs while the outer ring had 18 PMTs. The structure was positioned at two distances (85.6 cm and 131.4 cm) from the ADR in order to access a total of

4 possible scattering angles. The two distances were chosen to have a common angle in the two configurations (the closer-position inner-ring angle is equal to the farther-position outer-ring angle) corresponding to 220 eV elastic nuclear recoils from 56 keV neutrons. The other two recoil angles correspond to 100 eV and 440 eV for 56 keV neutrons.

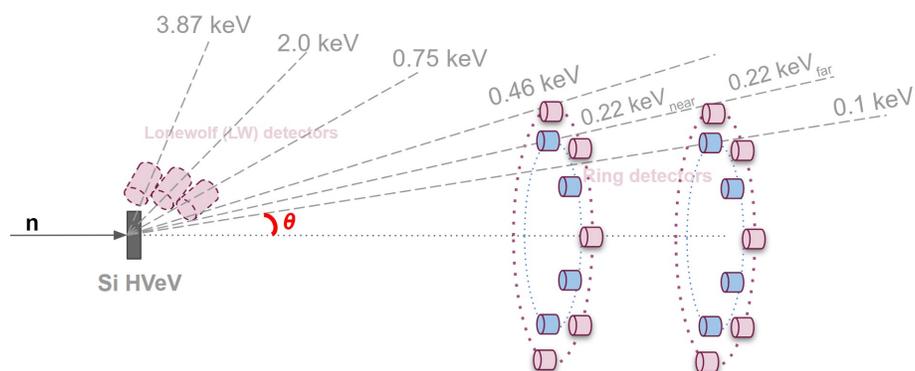


Figure 3.5. A sketch of the IMPACT@TUNL experimental setup.

3.2.3. Alignment

It is crucial to align the HVeV detector and the PMT detector correctly with respect to the beam because the experiment relies on the scattering angle to determine the recoil energy. The technique to align the system is not difficult but can be complicated. I document the general procedure in case it is useful for future experiments. The equipment needed is a self-leveling laser that projects a horizontal and vertical red line.

The alignment between the beam and the HVeV detector was done in three steps.

- (1) Finding the position of the detector on the fridge in the “beam direction”. I put the beam direction quoted because this step is actually done before installing the fridge in the beam. The beam direction means the 45° to the fridge’s symmetrical axis, which

is perpendicular to the top side of the HVeV detector. While the fridge is upside-down and open, we perform the following steps:

- Point the laser 45 degrees to the fridge and roughly pointing to the detector center.
- Move the laser until it is pointing to the detector center while maintaining 45 degrees to the fridge.
- Put the ADR shells on without moving the laser or the mobile cart, mark the laser spot on the outer vacuum chamber (OVC).

(2) Aligning the whole ADR to the beam. This step is achieved thanks to an adjustable (up/down, left/right) stand fabricated by the University of Montreal. The method is very similar to the first step:

- Mount the ADR to the stand and make sure the stand is perpendicular to the beam
- Align the self-leveled laser to the beam pipe, pointing at the center of the LiF target
- Adjust the fridge left/right and up/down on the stand until laser spot hits the marker on OVC

The alignment of PMT backing array is easier with the provided marks on the array. First, we set the distance between the array and the fridge with a laser range finder. Second, we make sure the array is centered to the beam. We set up the self-leveled laser aligned to the beam pipe and moved the array so that the vertical marks and horizontal marks coincided with the laser red lines. Third, we check that the array is perpendicular to the beam. In the end, we measure the distance between the array and the fridge again.

3.3. Simulations

In this section we describe the effort of simulating the neutron beam, and the interaction between the product of the beam with the collimator, the ADR, and various detectors. The main purpose of the simulation is to understand the expected recoil energy distributions for each secondary neutron detector. The simulation model built with GEANT4 [76] includes the neutron beam collimator, the ADR, the silicon detector, and the scintillator cells of the secondary neutron detectors. These simulations are also used to evaluate the stability of the neutron beam. More details of this simulation can be found in Reynolds' dissertation [77].

3.3.1. Simulation overview

The simulation consists of several elements: the Northwestern ADR, a neutron beam collimator, three lone wolf PMTs, and a set of 28 PMTs mounted on an aluminum ring. The full geometry is shown in Fig. 3.6. Throughout this chapter, the colors in images of the simulation follow the same conventions as Fig. 3.6



Figure 3.6. The overall geometry of the IMPACT simulation.

3.3.1.1. Simulation geometry. The ADR is modeled with its aluminum thermal shields, the niobium magnetic shield, as well as the two of its internal electrical circuit PCBs: the SQUID board and the GGG heat sinking board. The fridge also has its cryogenic reservoirs modeled,

filled with liquid nitrogen and liquid helium. Internal to the niobium can is the full detector housing, consisting of a copper box containing an HVeV sandwiched between two PCBs. The HVeV is a Si chip without the Al electrodes modeled. It serves as the origin of the simulation and is placed at its expected location based on the CAD drawings.

A collimator is placed near the fridge such that the nominal position of the LiF target is 92 cm from the detector, in line with the measured distance. The collimator consists of high-density polyethylene, borated polyethylene, and lead, with dimensions meeting the specifications provided by TUNL.

The active elements of the PMTs are modeled with polyvinyltoluene scintillator plastic. The number density of hydrogen in this material closely matches that of the EJ-301 and EJ-309 used in the real PMTs, so there should be no significant difference in the number of interactions in the PMTs. Three of these are lone-wolf detectors placed around the fridge based on the expected location of the HVeV. The remaining 28 are placed on an Al backing ring. The backing array PMTs have a 1 mm thick piece of lead in front of each of them to mimic the shielding used in the real experiment. Other than the lead plates and active scintillator elements, no parts of the PMT bodies or responses are modeled.

3.3.1.2. Simulation input. The interactions of protons on LiF are not simulated since the physics for this is not in Geant4. Instead, the products are created using a thin cylindrical source at the nominal target location. This cylinder has a radius of 1 mm and fires the products of the interactions towards the HVeV isotropically in a cone of 4 degrees half angle. These products include neutrons as well as four gamma lines: 110, 197, 478, and 6048 keV. A summary of the products, including the reaction and their production cross-section, is given in Table 3.1. The contribution of the gamma lines can be seen in Fig. 3.7

All the gamma lines have fixed energy. However, the energy of produced neutrons are dependent on the proton energy in the target. Given a specific proton beam energy, the proton

Product	Reaction	cross-section (mb/sr)
Neutrons	${}^7\text{Li}(p,n){}^7\text{Be}$	57.26
110 keV Gammas	${}^{19}\text{F}(p,p\gamma_1)$	2.31
197 keV Gammas	${}^{19}\text{F}(p,p\gamma_2)$	3.66
478 keV Gammas	${}^7\text{Li}(p,p'\gamma)$	6.1
6048 keV Gammas	${}^{19}\text{F}(p,\alpha\gamma)$	16.3

Table 3.1. The products created by protons incident on a LiF target, with natural Li composition, and their production cross-sections. The cross-sections are from [78, 79, 80]

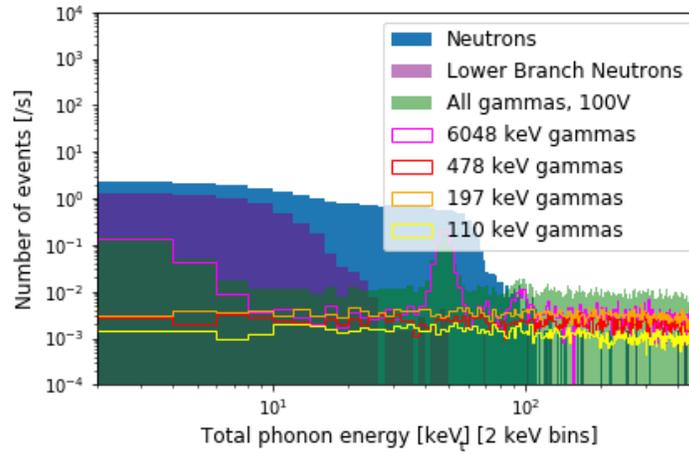


Figure 3.7. Neutron and gamma simulation detected by the silicon detector with a NTL bias of 100 V. This spectrum is before selecting particular recoil angle and does not include the TOF cut.

energy in the target can go down as the proton goes deeper into the target. The proton energy loss of a proton traveling in the LiF target is determined through a calculation using the TRIM program, part of the SRIM [81] package that handles the transport of ions in matter. Within TRIM, a proton of energy 1881 keV is taken to be incident with an angle of 0 degrees on a LiF crystal target. Using the "Monolayer Collision Steps/Surface Sputtering" method, the proton energy as it leaves the target is determined. The target thickness is varied from 5 nm to 100 nm in 5 nm intervals, representing different (p,n) interaction points. The energy loss in LiF, as determined by TRIM, can then be converted into a neutron energy following the kinematics

outlined in [73]. We sampled neutron energy distributions in this manner for targeted nominal beam energies ranging from 46 to 60 keV in 1 keV steps. Figure 3.8 shows the TRIM calculated neutron energy distribution used for the simulation input.

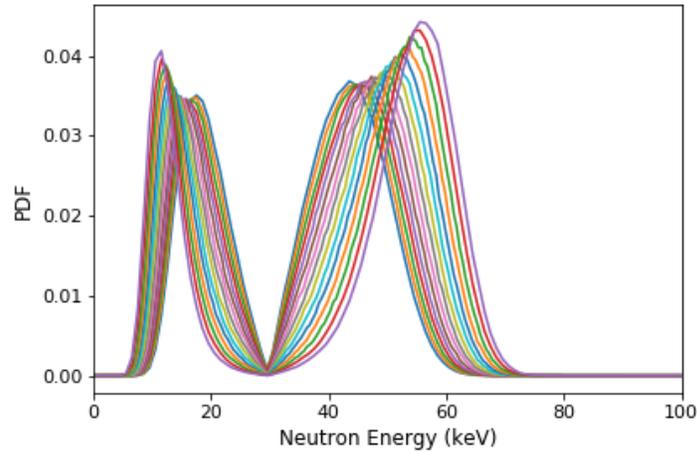


Figure 3.8. The distribution of neutron energies as determined by the TRIM calculation, for a 2 keV Gaussian spread in the proton beam energy.

3.3.1.3. Simulation result. The simulation generates the recoil energies that occur in the HVeV detector when the products of protons on LiF interact with it. This include all scatters in the HVeV detector, without any selection of the recoil angles (before coincidence). Outside of the simulation, we did a post-processing to select events that hit both the HVeV and PMT detectors (after coincidence). The coincidence spectra of 56 keV neutron beam is shown in Fig. 3.9, which is the expected energy spectra in HVeV detector at 0V. These spectra include multiple scatters in the silicon detector, which make up no more than 4% of the simulated events.

3.3.2. Beam energy estimation

We use these simulations to evaluate the stability of the neutron beam energy hourly based on the best fit of the spectrum of all neutron energy depositions in the silicon detector between 4 and 8.2keV_{nr} . The lower bound is selected so that this fit is performed outside our region

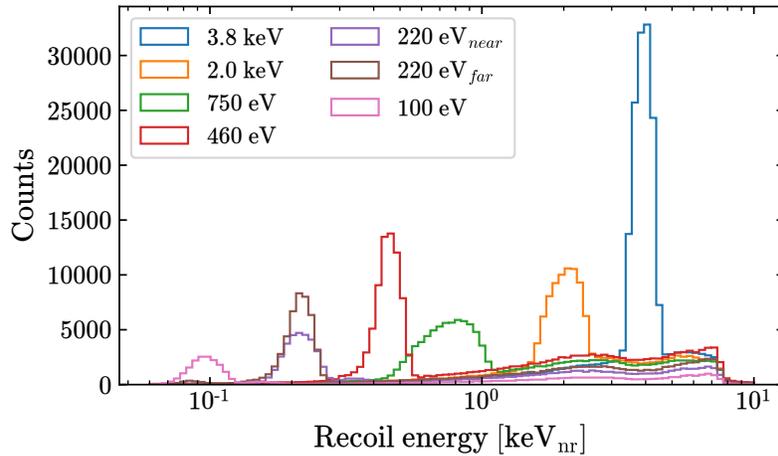


Figure 3.9. The simulated recoil energy spectra for neutrons that hit both the silicon and neutron detectors associated with the given neutron detector position. The nominal beam energy is 56 keV.

of interest, while the upper bound limits the comparison to an energy region where neutron interactions dominate. Figure 3.10 shows an example of the fit on one hour of data. Since the discrepancy between the Lindhard model and existing measurements is of order 20% at most for this energy range, we use the Lindhard model to convert the simulated recoil energies to total phonon energies via Eq. 1.17. The neutron beam simulation that matches the data the best is chosen to represent the beam behavior for that hour of data. This method yields a spread in the best-fit beam energy of about 3 keV, which we take to be the systematic uncertainty on the beam energy.

3.4. Neutron tagging detector analysis

In this section, we discuss the neutron tagging detector (PMT) analysis which includes the energy calibration of the PMT signals and the various selections investigated for the selection of the neutron events.

The PMT DAQ uses an online trigger after a trapezoid filter on the raw trace, with the threshold studied in Sec. 3.4.1. We record the raw trace of all the PMT channels, the raw trace

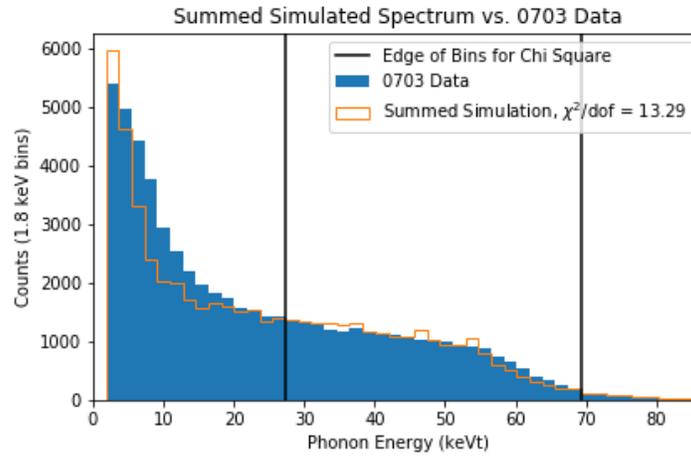


Figure 3.10. Comparison of a summed spectrum of the best-fit simulations on an hourly basis for a given day against that day’s data.

of the BPM channel, and the trigger time, when an online trigger fires. In each PMT trace, the expected trigger bin is at sample 120, with a total trace length of 400 samples. The raw traces are then processed offline to generate RQs for analysis.

Below is a list of reduced quantities we use in this study.

- **Trigger time:** we calculate a more precise trigger time reconstructed with a constant fraction discriminator (CFD). We use CFD to find the trigger time of the first peak it encounters starting from sample 100, denoted as `bd0` and from the beginning of the event, denoted as `bd02`.
- **Pre/post-pulse baseline:** Several different quantities are calculate to evaluate the baseline of each event.
 - `baseline`: them mean of the first 50 samples of an event
 - `baseline2`: them mean of the last 50 samples of an event
 - `baselinemid`: them mean of the sample 60 to 110
 - `stdbaselinemid`: them standard deviation of the sample 60 to 110
- **Pulse amplitude:** The maximum of the trace, denoted as `PeakHighValue`

- **Pulse integral:**

- SG: short integral, sum of the 5 samples before `bd0` and 7 samples after
- LG: long integral, sum of the 5 samples before `bd0` and 170 samples after

3.4.1. Calibration

The neutron tagging detector is only used to select events at certain scattering angles, thus the absolute energy scale is not necessary. Nevertheless, we still calibrate the PMT detector to monitor the threshold, as well as rejecting high energy γ events. We need a stable low threshold since a 56 keV neutron may deposit close-to-threshold energy in the liquid scintillator. With 10% quenching factor [82], we expect at most 5keV_{ee} energy deposition in the liquid scintillator.

The PMT calibration was performed with a strong ^{137}Cs source every day, where a 2-minutes calibration dataset was taken before we turn the beam on. We chose ^{137}Cs as it was the only strong radioactive source available, albeit the energy being very high thus not ideal for our purpose. For the calibration dataset of each day, we fit for the Compton edge (477 keV for ^{137}Cs) of each PMTs using the convolution of an ideal Compton edge into a Gaussian as our fitted function:

$$(3.2) \quad R(E) = \frac{1}{2} [a \cdot (E^2 + \sigma^2) + b \cdot E + c] \cdot \text{erfc} \left(\frac{E - E_c}{\sqrt{2}\sigma} \right) + \left[\frac{-\sigma}{\sqrt{2\pi}} \cdot a \cdot (E + E_c) + b \right] \cdot \exp \left(-\frac{(E - E_c)^2}{2\sigma^2} \right)$$

where E_c is the Compton edge energy or the maximum energy that can be transferred by the photon via a Compton scattering. We use E_c for a linear energy scale at $477\text{keV}_{\text{ee}}$.

We estimate the threshold of each PMT on every day. The measured thresholds are mostly around 2.5-3.5 keV_{ee} . To determine what an acceptable threshold would be for the PMT analysis, the PMT response was simulated using SuperSim. One million 55.7 keV neutrons were incident on a cylinder of polyvinyl toluene plastic scintillator. The recoil energies from hydrogen atoms

in the scintillator were taken from the simulation and had a 10% quenching factor [82] applied to them to account for the PMT response. The simulation showed a $\sim 80\%$ efficiency with 3 keV_{ee} threshold, which confirms that the PMT detectors are in good conditions.

3.4.2. Event selection

Data quality selections are used to select the events whose traces match idealized neutron events. We use three types of selections on PMT detector.

First, the two reconstructed trigger time bd0 and bd02 are required to be both positive, and consistent with each other within 10 ns. The number 10 ns is set empirically.

Secondly, the reconstructed energy from integral is limited to be 1.5-10 keV_{ee} . Based on the simulation, most neutrons deposit energy less than 10 keV_{ee} . We also reject events whose amplitude reaches the maximum range of the ADC.

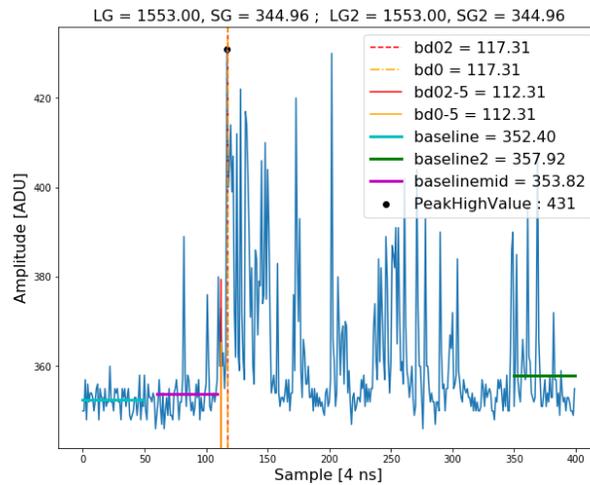


Figure 3.11. Example of events that have a long chain of peaks.

Third, two kinds of noise events are rejected. One kind of events have a long chain of peaks in the waveform, see Fig. 3.11. The other kind have a pulse shape that is narrower than normal

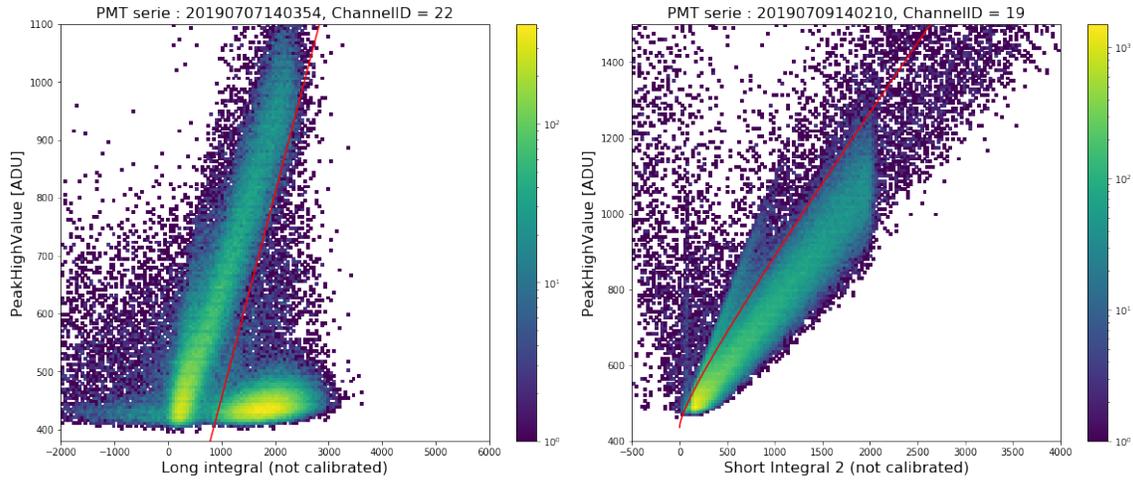


Figure 3.12. Amplitude versus long integral (left) and short integral (right). The region below the red line in the left plot and above the red line in the right plot is rejected.

events. Both are found to be not consistent with neutron events. These events are rejected by checking the pulse amplitude versus long/short integral, see Fig. 3.12.

3.5. Coincidence analysis

For the ionization yield measurement, we are interested in events that are 1) from the beam, 2) scatters in the HVeV, and 3) scatters at certain angles. This requires a triple coincidence between the beam pickup monitor (BPM), the silicon detector (HVeV), and the neutron tagging detector (PMT). In this section, I describe the analysis of the coincidence events. The algorithm to sync time between two systems is outlined in Sec.3.5.1. The event selection criteria are described in Sec. 3.5.2. The accidental coincidence background is estimated in Sec. 3.5.3. Section 3.5.4 includes the correction of recoil energy.

3.5.1. Time synchronization

The foundation of finding the coincidence events is establishing an accurate and synchronized time base among the three detectors. As described in the introduction, BPM and PMT signals

are recorded with one DAQ (the PMT DAQ), while the HVeV signal is on another DAQ (the HVeV DAQ) which does not share the same clock. To align the time base of the two DAQs, we use a time-syncing signal with the help of offline processing.

The time-syncing signal is a one-per-minute pulse that is recorded by both DAQs. During the first round of offline processing, each event in the PMT DAQ or the HVeV DAQ is assigned a timestamp based on the clock of the DAQ it comes from. Then we pick out the events on the channel for the time-syncing signal on both DAQs. We then use the timestamps of these events from two DAQs to construct a piece-wise linear function that maps the timestamp from one DAQ to another.

Before applying to the real data, the time syncing algorithm is tested first on a mock-up dataset. The mock-up dataset is generated by feeding a signal generated by a function generator into the two DAQs, with a shaping amplifier inline in the TES signal to slow down the signal. The time syncing signals are injected in an identical way as the real data. A $\sigma \sim 0.3\mu\text{s}$ resolution is observed. A shifted mean of $\sim 9\mu\text{s}$ is due to the shaping amplifier applied to the TES signal fed into the TES DAQ.

In the future, there is a simple way to improve the resolution of the time synchronization, which will then improve the background rejection. The method is to generate the time-syncing signal with the HVeV DAQ, because the HVeV DAQ is the one with a slower time base and cannot measure time as precisely as the PMT DAQ. By generating the signal with HVeV DAQ, we no longer need to measure time on the HVeV side. Thus, the time resolution of HVeV DAQ is eliminated from the time alignment.

3.5.2. Event selection

Three types of selections were considered in this analysis:

- **HVeV event quality selections** remove events whose energy reconstruction would not be accurate;
- **PMT event quality selections** select neutron events occurring in the liquid scintillator present in front of the PMTs;
- **coincidence selections** are used to identify the event that occurred in the BPM, HVeV and the PMTs.

The main challenge of setting the selection in this analysis is to optimize the selection in order to maximize the signal/background ratio (in other word to preserve as much neutron as possible while rejecting all the backgrounds) when the expected number of signal events is as low as 10-20 events in some configurations. The PMT event selection has already been discussed in Sec. 3.4.2. Details of the HVeV and coincidence selections are described as following.

3.5.2.1. HVeV selection. Two selections are applied to HVeV events. Both selections aim to remove events that have inaccurately reconstructed energy.

1) Baseline selection removes events that have an unstable pre-pulse baseline. This could be the event that occur on the tails of large energy depositions or a time period where the fridge temperature is unstable.

2) Pile-up selection removes events that have secondary pulses after the main trigger, which may result in an inaccurate estimate of the energy. We allow up to 2 pile-ups per event because the MF estimator for determining an event's energy has the ability to automatically reject several pile-ups without noticeably biasing the result.

3.5.2.2. Coincidence selection. There are three types of events with time information: BPM, HVeV, and PMT. We call the time difference between the HVeV and PMT event "dt", and the time difference between the BPM and PMT event "time of flight (TOF)". For beam-neutron events that scatter in HVeV and get tagged by PMT, the dt is the neutron time of flight from

the HVeV to the PMT, and the TOF is the neutron time of flight from the LiF target to the PMT.

dt has a resolution of ~ 500 ns, which is limited by the time resolution of the HVeV detector. TOF has a resolution of ~ 1 ns. The different time resolution of dt and TOF makes their spectrum looks very different. The speed of 55.7 keV neutron is 3.2×10^6 m/s. It takes 200 ns for a neutron to travel from HVeV to PMT backing array. Due to the large resolution of dt, the neutron events will show a wide peak around 0 in the dt histogram. Meanwhile, the TOF histogram shows the detailed shape of the distribution.

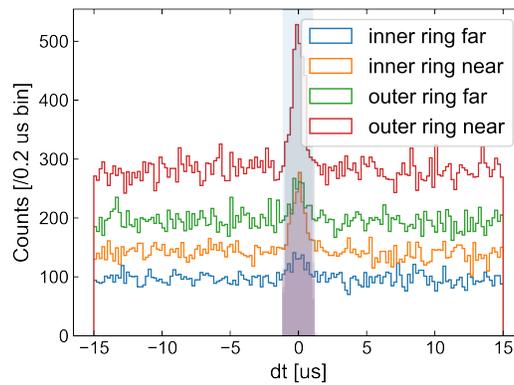


Figure 3.13. dt selection on the ring detectors. The shaded vertical region shows the dt cut window.

We apply a " ± 3 -sigma" cut on dt, see Fig. 3.13. Note that we use the same sigma value from a fit to the dt of all data instead of using a different width for different recoil locations. We made this choice for the simplicity of exposure estimation when we modeled the background.

We use the simulation to set the TOF cut since the statistics of data are too low. The first step is to correct the discrepancy between the simulation and the data. As shown in Fig. 3.14 left, the simulation is different from the data in two aspects: the peak locations and the flat background. The peak location is due to the possible small geometry difference from the real setup. A 1 cm shift along the recoil angle will result in a TOF difference of 3.1 ns. The higher

flat background rate in the data is not reflected in the simulation because of unmodeled random coincidence events. With this understanding, we fit the simulation TOF distribution to data with three parameters: the time offset, the amplitude, and a flat background rate. The fit result is shown in Fig. 3.14 right panel. The difference between the fit and data may come from the modeling of the beam energy profile. The beam energy is drifting with an unknown pattern, while in a single simulation we use a Gaussian spread of beam energy.

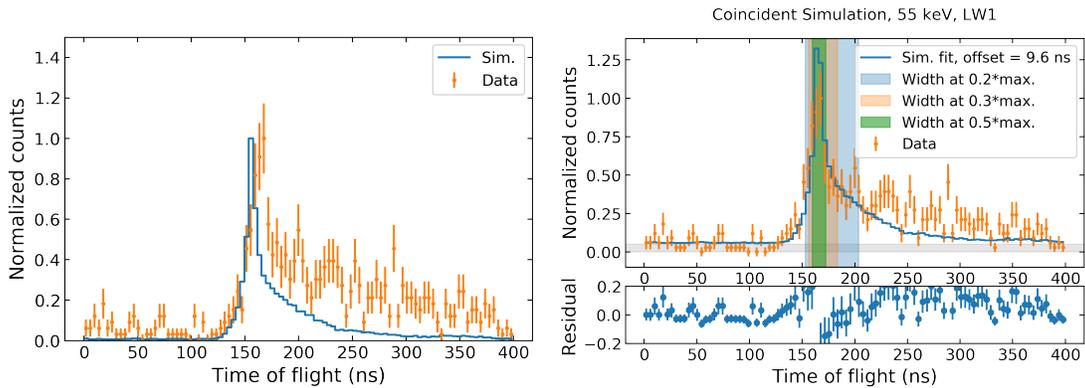


Figure 3.14. TOF cut for LW1. Left: before correcting the offset in simulation. Right: after correct for the time and amplitude offset in simulation. The three vertical bands corresponds to the cut window at 0.2, 0.3 and 0.5 of the maximum value.

We then set three cut regions by finding the 0.2, 0.3, and 0.5 of the maximum value in the simulation. We treat the 0.3-max one as the nominal value of this cut and use the other two for the systematic uncertainty study. The three TOF intervals roughly correspond to neutron energy spread of ± 1 keV, ± 2 keV, and ± 3 keV, which are reasonable choices based on the estimated standard deviation of neutron energy of 3 keV.

3.5.3. Background estimation

We use events outside the main dt peak to estimate the random coincidence background, including events caused by environmental gamma and neutrons, as well as leakage events in the detector. We use a background dt window between $\pm 50 \mu\text{s}$ and $\pm 20 \mu\text{s}$.

The number and the threshold of PMTs in each recoil position are different, which makes the coincidence efficiency different. The background spectrum then needs to be estimated for each recoil angle individually, though we do not expect them to be significantly different except for a normalization factor.

For the yield measurement, we don't need to subtract the background since we build the background into the signal model in Sec. 3.6. However, here we still show the background-subtracted data just for demonstration purposes.

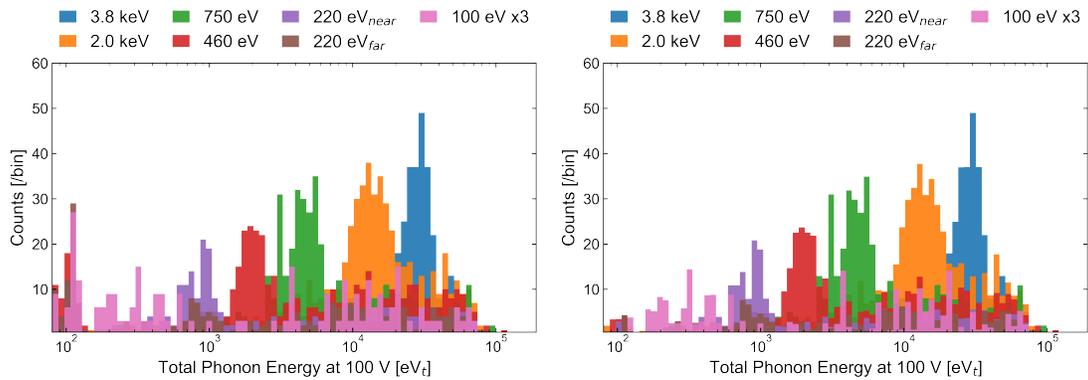


Figure 3.15. Total phonon energy spectrum of all recoil angles, before and after subtracting the background. The 100 eV histogram bin counts are multiplied by 3 to make it visible compared to other histograms.

3.5.4. Recoil energy correction

The recoil energy depends on the scattering angle, which relates to the position of HVeV and PMT detectors. The HVeV detector position is measured indirectly to an accuracy of 1 cm. We

do not have a more accurate measurement because the position can change when the fridge cools down as different materials contract at different rates. As a result of this change in position, the scattering angles are altered slightly, which in turn is tied to the recoil energy. Therefore, shifts in the detector position represent a systematic uncertainty in our recoil energy determination and, ultimately, the yield.

The measured position of the HVeV detector offset from CAD position by $dx = -1.6$ cm, $dy = 1.5$ cm and $dz = 0.9$ cm (with +y along the beamline). This offset can create a significant change in the recoil energy in the lone-wolf detectors while leaving the ring detectors unaffected. This is because the ring detectors are far away from the HVeV detector.

Due to this large effect on LW, we have decided not to use both the CAD position and the simulation position for LW detectors. We would like to constrain further the impact of detector position for LW detectors only since the effect on the ring detector is negligible. Although we could not measure the detector position more accurately than the previous numbers, we have a direct measurement of the recoil energy from the 0 V data. We bypass the step of finding where the detector actually is and scale the simulation spectrum with CAD geometry to fit the 0 V data.

We scale the simulated spectra with a scale factor and calculate the Poissonian negative log-likelihood (NLL) between the simulation and the data. We scan over the scale factor and find the best fit, which is where NLL reaches its minimum. The NLL as function of scale factor is shown in Fig. 3.16, and the best fit energy is summarized in Table 3.2.

Nominal recoil energy	Measured recoil energy	Scale factor
0.75 keV	0.89 keV	$1.18^{+0.03}_{-0.15}$
2.00 keV	2.33 keV	$1.16^{+0.05}_{-0.14}$
3.87 keV	3.91 keV	$1.01^{+0.01}_{-0.11}$

Table 3.2. Measured mean recoil energy for the LW detectors.

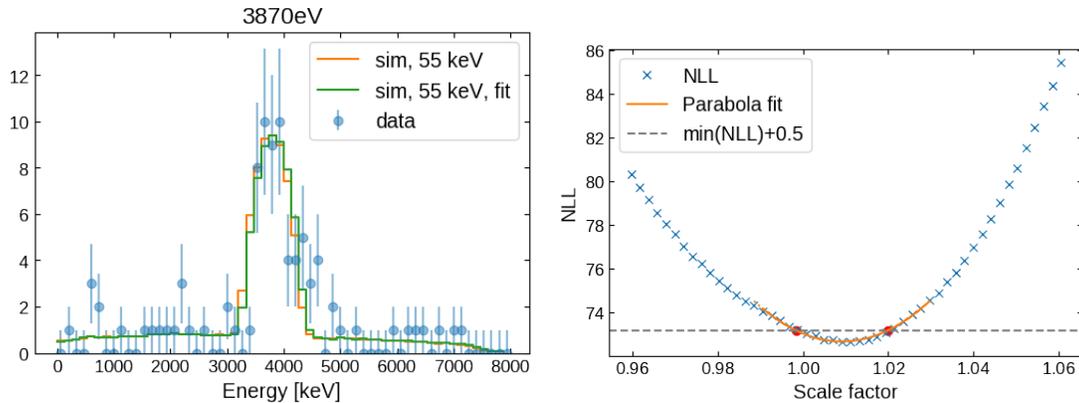


Figure 3.16. Left: The recoil energy spectrum of data (coincidence spectrum of all 0V data), simulation and the fit result. Right: NLL as function of scale factor. The yellow line is the parabola fit, and the solid part shows the fit region. The red dots are where the NLL increases by 0.5, which correspond to 1 sigma uncertainty under Gaussian assumption.

Now we will use the fit result as the nominal recoil energy for LW detectors. The simulation of LW detectors with CAD geometry is scaled by the scale factor and becomes the nominal simulation dataset. The upper bound of the scale factor is obtained from the fit. The lower uncertainty of this scaling is dominated by our knowledge of the 0V calibration. According to Sec. 2.5, the 0V calibration is obtained from the scaling of HV calibration with a factor of 1/1.1. Since we don't have data to bind the uncertainty of this scaling, we will treat the difference between the 0V calibration and HV calibration as a one-sided systematic uncertainty of the 0V calibration. This uncertainty then translates to the uncertainty of the scale factors: we perform the fit again on 0V data but with the HV calibration function. We created another two sets of simulations with the lower and upper bound of the scale factor.

3.6. Signal model and fit algorithm

We build a detector response model that transforms the simulated nuclear recoil energy spectra, shown in Fig. 3.7, into the total phonon energy spectra, which is sensed in the detector

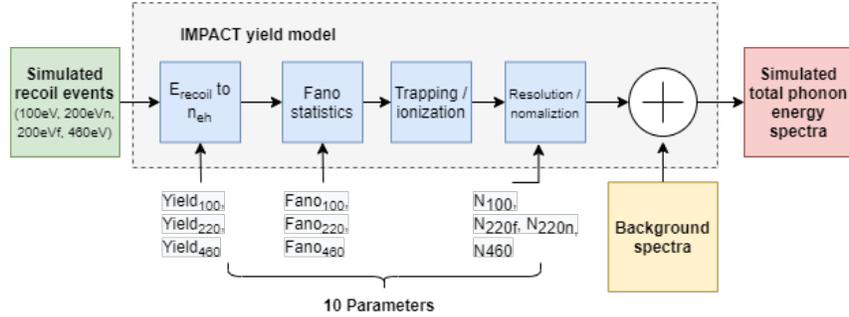


Figure 3.17. An overview of the signal model used in the MCMC for the backing array detectors. Simulated recoil energies are converted into phonon energy distributions via a set of 10 parameters, then combined with a data-estimated background spectrum to make a final simulated total phonon energy spectrum.

and used to extract the ionization yield from the data. We parameterize $Y(E_r)$, the energy-dependent ionization yield, as a piecewise linear function characterized by the ionization yield values at the nominal nuclear recoil energies that we are measuring. The function is fixed at low and high energies to be $Y(E_r = 0 \text{ keV}) = 0$, and $Y(E_r = 10 \text{ keV}) = 0.3$ from the Lindhard model. For each simulated event with E_r , $\langle n_{\text{eh}} \rangle$ is calculated with Eq. 1.17. The n_{eh} for this event is sampled from a distribution with this mean and a variance ($\sigma^2(n_{\text{eh}})$) characterized by a Fano factor $F = \sigma^2(n_{\text{eh}}) / \langle n_{\text{eh}} \rangle$. Poisson, binomial, and negative binomial distributions are used to model the scenarios of $F=1$, $F>1$, and $F<1$, respectively. n_{eh} is further smeared to account for charge trapping (12.7% probability) and impact ionization (0.6% probability) in the detector [32]. A Gaussian distributed detector from Fig. 2.13 is applied to the total phonon energy E_t after conversion from n_{eh} with Eq. 1.17. This procedure is visualized in Fig. 3.17.

We compare the simulated E_t spectra to data after applying their normalization factors, denoted as n . Each simulated E_t spectrum has three free parameters, Y , F , and n . For the 220 eV recoil energy point that is measured twice, the yield and Fano factor are constrained to be the same. We sample these parameters with a Markov-Chain Monte Carlo (MCMC) method using a binned likelihood loss function implemented in BAT [83].

To improve the speed of convergence, we perform the fit in an iterative way. The first fit is for the ring detector parameters while keeping the yield at the LW detector energies fixed to the previous measured ionization yield from Ref. [68]. We then fit for the LW detector parameters while keeping the ring parameters fixed to the previous fit result. Finally, the fit is rerun on the ring parameters with the LW parameters fixed to the previous fit result. The fit model with the best-fit parameters is shown in Fig. 3.18.

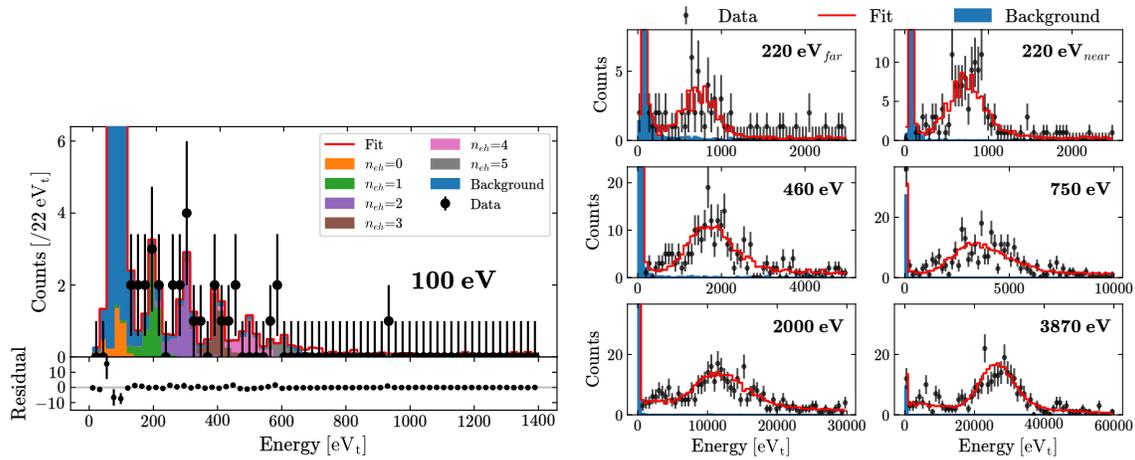


Figure 3.18. Comparison of the fit model using the best-fit parameters to the data for all the recoil energies. For the 100 eV recoil energy, the integer number of electron-hole (eh) pair contributions to the model are shown colored in by pair number up to $n_{eh} = 5$. Individual eh pairs are not shown for other recoil energies because they are indistinguishable based on our energy resolution and the number of events.

3.7. Systematic uncertainty

We consider the following systematic uncertainties:

- (1) Recoil energy uncertainty: The recoil energy uncertainty has two contributions. There is a 1.3% uncertainty arising from the total phonon energy scale calibration of the 0 V mode data. Additionally, there are uncertainties on the scale factors used for the LW detectors given in Table 3.2.

- (2) Neutron beam energy uncertainty: The central energy of the neutron beam was evaluated to have a spread of ± 3 keV. We vary the beam energy in the simulation and use the resulting spectra in the fit model.
- (3) Charge trapping and impact ionization probability uncertainties: The values of these probabilities are varied by the uncertainty from Ref. [32]. They are varied together such that when one probability is increased, the other is decreased.
- (4) Time of flight cut uncertainty: The neutron time of flight is correlated with their energy. The effect of the TOF selection is evaluated by choosing a narrower (wider) window selecting events within $\pm 50\%$ ($\pm 20\%$) of the simulated TOF distribution maximum.
- (5) Uncertainty in modeling the Fano factor: To evaluate the effect of mismodeling the Fano factor, we perform a fit with the Fano factor fixed to one.

Although we are using MCMC, we choose not to model the nuisance parameters in the MCMC for the evaluation of systematic uncertainty. The model we used in the MCMC is already complicated and is facing convergence issues whenever we add in more parameters due to the shallow minimum of the NLL function. Another difficulty is that some of our nuisance parameters directly affect the input data or change the fit model of MCMC, which makes these parameters very computationally expensive to be included in the likelihood function in MCMC. Instead of building an MCMC model, we use pseudo-experiments (PE) to evaluate systematic uncertainty. PEs are generated by sampling events from the simulation and converting the recoil energy into total photon energy with our signal model. Following are the detailed steps in generating PEs:

- PE at each recoil energy is characterized by a three parameters: yield Y , Fano factor F and signal normalization N . The recoil energy of simulated scatters is converted to total phonon energy via the signal model with Y and F .

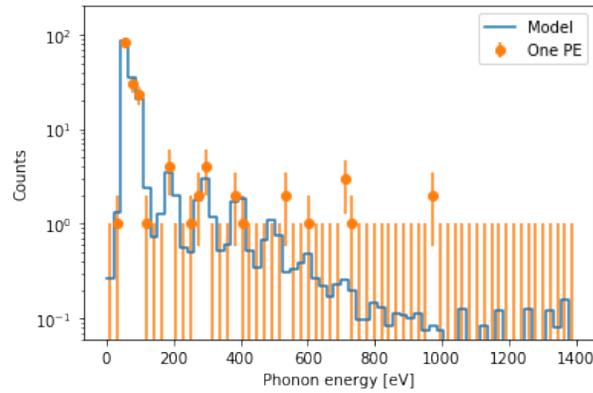


Figure 3.19. The expected total phonon energy distribution from the signal model and one PE sampled from this distribution.

- The total phonon energy is binned into histogram with then same binning as data and normalized to the specified signal normalization N .
- The measured background spectra is added to the total phonon energy spectra. We treat this final spectra as the expected distribution of total phonon energy with a given (Y, F, N) .
- Finally, we sample from this distribution to get a PE histogram. For each bin in this distribution, we generate a random number from a Poisson distribution with mean value equal to the bin content.

Figure 3.19 shows an example of the PE generation.

For each of the items that we would like to evaluate, we generate 100 PEs. We then run the fit algorithm on the PEs and calculate the deviation from the nominal setting. The statistical uncertainty is reduced by calculating the average deviation of all the PEs. Figure 3.20 shows the averaged deviation from nominal values. For the first four items of systematic uncertainty, they are varied to both directions and the two markers with the same color correspond the upper and lower bound of each item. For the Fano factor, it is only varied to one side and only one marker is shown.

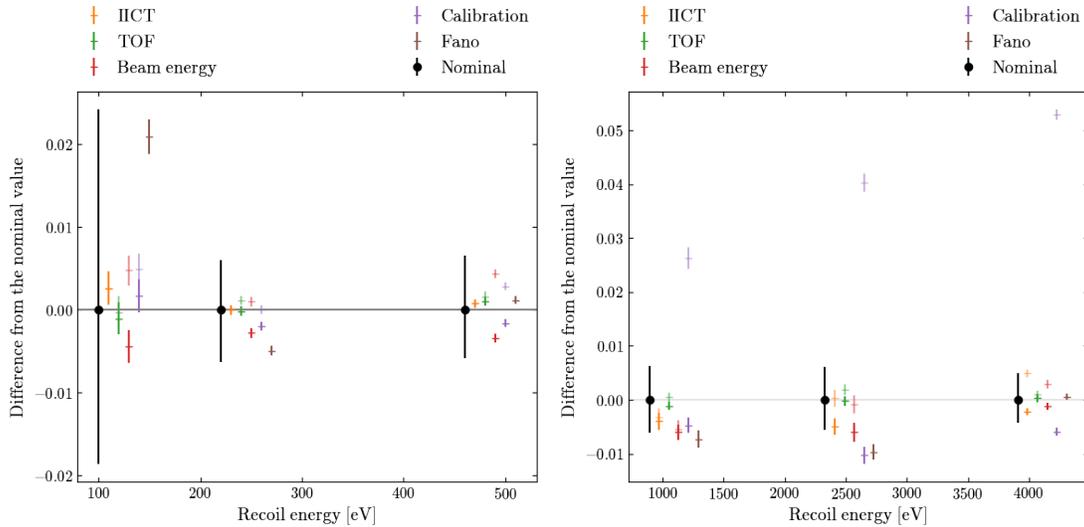


Figure 3.20. Deviation of the ionization yield from the nominal values when varying each parameter, evaluated with pseudo-experiments. Left panel shows the ring positions, right plot shows the LW positions. The horizontal offset in the points for a given recoil energy is for clarity.

We then combined the systematic uncertainty items. For the first four parameters (IICT, TOF, beam energy, calibration), if the deviations are on different sides, then we treat each side as a separate uncertainty. If the deviations are on the same side, then we use the largest deviation as a two-sided uncertainty. The Fano factor is treated as a two-sided uncertainty. The absolute deviation is added on both sides.

This procedure is reflected in Fig. 3.21. For example, the IICT of 100eV deviates only to the positive side, so the total uncertainty spans symmetrically with the same number; the beam energy uncertainty of 100 eV deviates on both sides, so the total uncertainty is just as the MCMC result.

3.8. Result

We have measured the ionization yield of nuclear recoils in silicon at six different energies down to 100 eV, the first such measurement to do so. The previously noted deviation from the

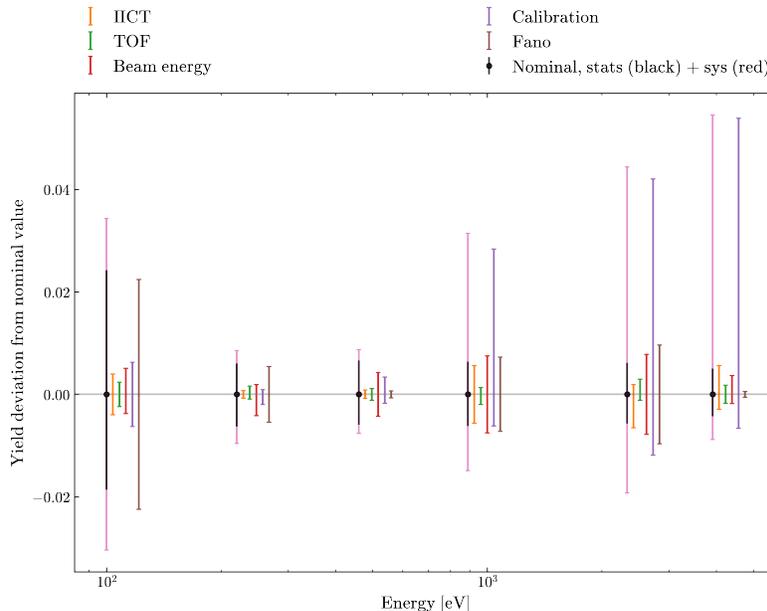


Figure 3.21. Summary of the systematic uncertainties. The end-cap of the error bars show the range of each systematic uncertainty. The horizontal offset in the points for a given recoil energy is simply to provide a visual aid in distinguishing the points.

Lindhard model extends down to 100 eV with no indication of the existence of an ionization production threshold. This latter fact should be of great interest to rare event search experiments. Furthermore, our results are consistent with the previous experiments using a photo-neutron source [67] and a similar neutron-scattering setup [68] above 2.3 keV.

The measured yield and the uncertainties are provided in Table 3.3 and summarized in Fig. 3.22. The correlations between the fit parameters were found to be small. We provide a least-square fit to our results with an empirically chosen power-law function that is constrained to go through the Lindhard model at 10 keV_{nr}. The function form is $Y(E_r) = Y_{10keV}/(10000^B) \cdot E_r^B$, in which $Y_{10keV} = 0.302$. The resulting best fit is $B = 0.261$ with confidence interval of [0.250,0.278]. Note that this is only an empirical line to guide the eye without physical motivations.

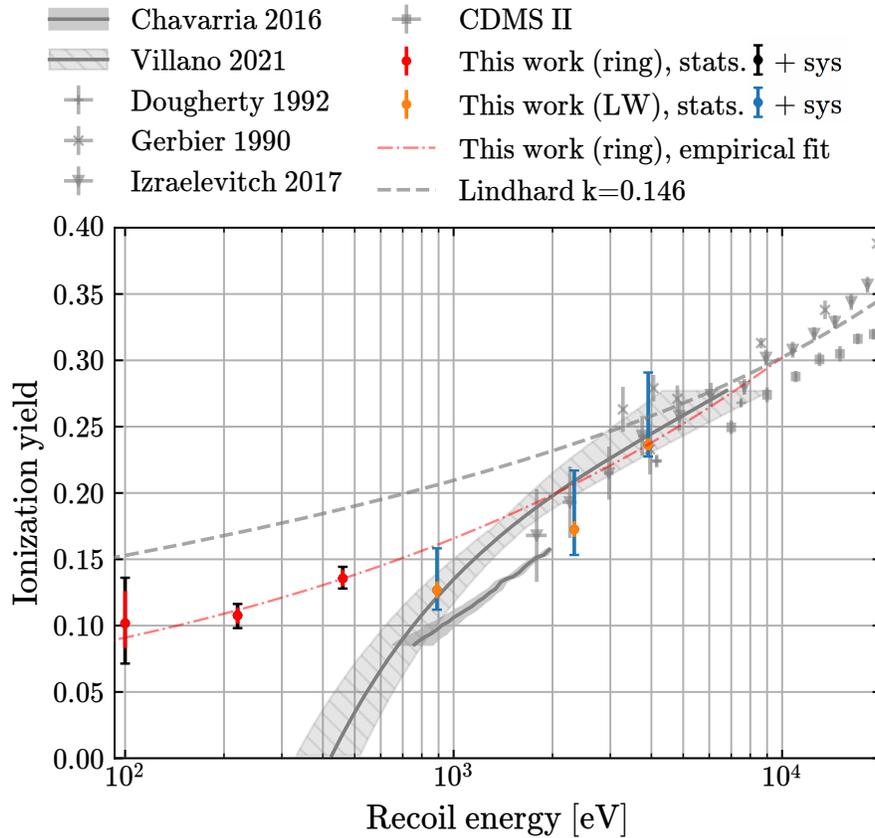


Figure 3.22. The measured ionization yields, along with their statistical, systematic uncertainties and a fit with power-law function. Also shown are data points from previous measurements [64, 65, 66, 67, 68, 84]. The dashed line shows the Lindhard model with the k value expected for silicon [85].

On the theoretical side, the most recent work of Sakris et al. [86] further improved the calculation of Lindhard model by using more realistic electronic stopping power (S_e) and atomic binding energy models. Compared to their previous work using a simple constant atomic binding energy, the modeled ionization threshold is lowered from ~ 300 eV to ~ 50 eV, which can potentially be consistent with the IMPACT measurement. Future ionization yield measurement to lower energy is needed to test the newly predicted ionization threshold.

E_r [keV _{nr}]	ionization yield Y	Fano fac- tor	normalization	Stat.	Recoil energy	Neutron beam energy	CT/II	TOF	Fano factor mismodeling
0.1	$0.102^{+0.034}_{-0.030}$	$0.9^{+0.7}_{-0.4}$	28^{+6}_{-5}	$+0.024$ -0.019	$+0.006$ -0.006	$+0.005$ -0.004	± 0.004	± 0.002	± 0.022
0.22	$0.108^{+0.009}_{-0.010}$	$0.5^{+0.2}_{-0.1}$	48^{+7}_{-7} (118 ⁺⁹)	$+0.006$ -0.006	$+0.001$ -0.002	$+0.002$ -0.004	± 0.001	-0.001	± 0.005
0.46	$0.136^{+0.009}_{-0.008}$	$1.8^{+0.6}_{-0.5}$	230^{+14}_{-13}	$+0.007$ -0.006	$+0.003$ -0.002	± 0.004	± 0.001	± 0.001	± 0.001
0.89	$0.127^{+0.031}_{-0.015}$	$3.7^{+0.8}_{-0.9}$	288^{+11}_{-12}	$+0.006$ -0.006	$+0.028$ 0.006	± 0.008	± 0.006	$+0.001$ -0.002	± 0.007
2.33	$0.173^{+0.044}_{-0.019}$	$7.7^{+3.2}_{-2.2}$	377^{+16}_{-14}	$+0.006$ -0.006	$+0.042$ -0.012	± 0.008	$+0.002$ -0.007	$+0.003$ -0.001	± 0.010
3.91	$0.236^{+0.055}_{-0.009}$	$8.4^{+2.4}_{-1.9}$	318^{+12}_{-15}	$+0.005$ -0.004	$+0.054$ -0.007	$+0.004$ -0.002	$+0.006$ -0.003	± 0.002	± 0.001

Table 3.3. Measured ionization yield in silicon as well as the measured Fano factor and signal normalization, provided with their total uncertainties. The 220 eV normalization given without (with) parentheses is for the far (near) position of the ring detectors. The remaining columns provide the statistical (stat.) uncertainty as well as the systematic uncertainties associated with the uncertainty on the recoil energy, the neutron beam energy, the charge trapping and impact ionization probabilities (CT/II), the time of flight cut, and mismodeling the Fano factor.

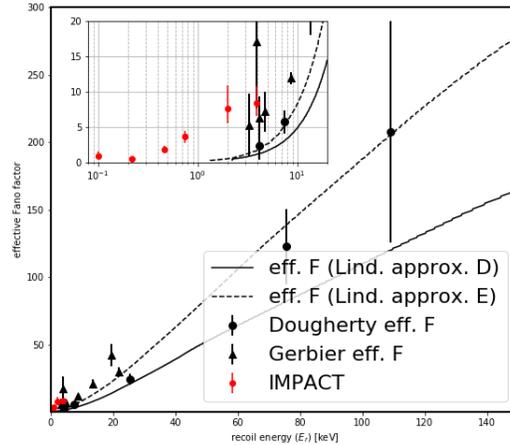


Figure 3.23. Comparison of the measured Fano factor with other experiments and Lindhard model. Code take from Ref. [87].

The measured Fano factor is presented in Fig. 3.23 and Table 3.3. The numbers are higher than the Lindhard model prediction. Measurement of the Fano factor is not the main purpose of this measurement, but the result could be useful to some research. Reference [87] has a more detailed discussion of the effective Fano factor of nuclear recoil.

CHAPTER 4

Investigation of the background in HVeV detector

This chapter has been published on Physics Review D as Ref. [88]. Printed with permission of the SuperCDMS collaboration and American Physical Society (APS). Copyright © 2022 by American Physical Society.

4.1. Introduction

Many searches of low-mass (sub-GeV/ c^2) dark matter have recently been challenged by unexpected excess event rates, along with the success of developing lower-threshold detectors. These include reports from experiments using cryogenic calorimeters instrumented for readout of phonon signals, such as EDELWEISS [22, 23], CRESST [24], NuCLEUS [25, 26], and SuperCDMS-CPD [27]. Unexpected events are also present in detectors instrumented for charge readout, such as the CCD-based experiments SENSEI [28] and DAMIC [29], as well as the phonon-based measurement of ionization signals [30, 31].

These latter measurements were made possible by the development by the SuperCDMS Collaboration of silicon-based gram-scale detectors: the high-voltage eV-resolution (HVeV) detectors [44, 32]. These detectors can be operated in high voltage (HV) mode in which an applied electric field amplifies the signal from electron-hole pairs (e^-/h^+) via the Neganov-Trofimov-Luke (NTL) effect [18, 19]. If the voltage is sufficiently high, the signal represents the number of e^-/h^+ , and a trigger threshold of well below a single e^-/h^+ was reached. However, these devices can also be operated in zero-voltage (0V) mode. In this case the measured signal represents the actual interaction energy.

We undertook an above-ground search for dark matter with a second-generation Si HVeV detector. An analysis of the data taken in the HV mode (100 V) was described in Ref. [31] and measured an unexplained excess of events similar to those observed with a previous version of the detector [30]. In order to better understand this excess event rate, in this manuscript we analyze the data taken in the 0V mode alongside the data taken at two different high-voltage settings: 60 V and 100 V. We infer information about the origin of the observed events by comparing how the spectrum scales with the applied voltages.

This chapter is arranged as following: We review the experimental setup in Sec. 4.2 and present the event reconstruction algorithms in Sec. 4.3. We present a dark matter analysis of the data taken in the 0V mode in Sec. 4.4. The investigation of the low-energy events starts in Sec. 4.5, where we discuss a class of events with anomalous pulse shape found in the dark matter search data, and in Sec. 4.6 we compare the pulse shapes and energy spectra from data taken at different voltages. In Sec. 4.7, we discuss a plausible explanation for the low-energy events with the anomalous pulse shape.

4.2. Experimental setup and data collection

The experimental setup and data collection conditions used in this analysis are identical to those described in Ref. [31], and are very similar to what is used in Chapter 3. We use the identical NF-C HVeV detector in the Animal ADR and the same data acquisition system.

We collected data during April 29–May 16 of 2019, including calibration data and dark matter (DM) search data at 0 V, 60 V and 100 V. Each day during the data-taking campaign, the ADR was re-cooled down from above 4 K. The ADR base temperature was stabilized at 50 mK from April 29th to May 7th and at 52 mK from May 8th to May 16th. Both channels of QETs were operated at 45% of their normal-state resistance. We calibrated the detector energy response daily using a 635 nm laser that was fiber-coupled from room temperature to

the detector housing. We also took calibration data on May 14 with an ^{55}Fe source at crystal biases up to 60 V to extend the detector calibration to ~ 100 keV.

4.3. Event reconstruction

In this section we describe the triggering and energy-reconstruction algorithms, and the energy calibration procedure. In-depth discussions of the algorithms and calibration procedure can be found in Refs. [31, 32].

4.3.1. Triggering and energy reconstruction

We read out a continuously sampled timestream of the current flowing through each QET detector channel. The sum of the traces from the two channels is filtered with an optimum filter (OF) [40, 13] before applying the threshold trigger as part of the offline data processing. In this analysis we use an OF time window of 10.8 ms, with equal pre- and post-trigger regions of 5.4 ms. We build a pulse template for the OF using events with a total phonon energy of ~ 100 eV from laser-calibration data, an energy deposition where the detector is far from its saturation regime and thus its response is expected to be linear. We also calculate the noise power spectral density (PSD) on a daily basis from randomly-selected sections of the data that lacks pulses. We set a 9.2 eV trigger threshold for the dark matter constraints discussed in Sec. 4.4, which results in a 20 Hz trigger rate. For the comparison of 0V and HV data discussed in Sec. 4.5 and onward, we use a higher threshold of 15 eV to reduce the contribution from triggers caused by noise.

We use the amplitudes calculated by the OF algorithm as the energy estimator for low energy events, and use an integral-based energy estimator for high-energy events. At higher energies the TESs approach their normal-state resistances, resulting in “flat-topped” pulses. These saturated pulses have shapes that deviate significantly from the pulse template, resulting in degradation of the energy sensitivity of the OF amplitude. The integral-based energy estimator integrates over

the raw trace when the detector is saturated and the signal-to-noise ratio is high, and integrates the area below a fit to the tail of the pulse using the average pulse template where the signal-to-noise ratio is low. We refer to this estimator as the “Matched Filter (MF) integral” [32]. The detector energy reconstruction is based on the OF amplitude below 600 eV and MF integral above 800 eV, with a linear transition in between.

4.3.2. Energy calibration

In this section, we discuss the calibration procedure using HV data, the application of the daily gain corrections, and their combination into a single calibrated energy estimator. We also discuss how this calibration is applied to 0V data in the end.

We calibrate the detector from the threshold to ~ 100 keV. The calibration is divided into two parts: (1) low-energy calibration using a laser, up to 700 eV at 100 V bias; (2) high-energy calibration using a combination of laser data up to 700 eV and ^{55}Fe source data up to 104 keV with bias voltages up to 60 V. We collected laser data every day for robust low-energy calibrations that accounts for the daily gain change due to thermal cycling of the ADR. In contrast, it was not practical to conduct daily high-energy calibration, because of the extended source exposure required to acquire sufficient event statistics. We, therefore, took the high-energy calibration data only once during the data-taking campaign on a dedicated day (“Fe-day”).

The low-energy calibration follows a similar method as described in Ref. [31]. We use laser data to calculate calibration functions $E_{\text{OF},i}$ to convert OF amplitudes (A_{OF}) to energies up to 700 eV. The subscript i denotes the i^{th} day of data-taking. The function is a second order polynomial

$$(4.1) \quad E_{\text{OF},i} = a_i \cdot A_{\text{OF}} + b_i \cdot A_{\text{OF}}^2,$$

where a_i and b_i are the two calibration coefficients for the i^{th} day. An example of the OF calibration curve is shown in Fig. 4.1.

We derive a second calibration function based on the MF integral up to 98 keV with the laser data and the ^{55}Fe data at 40 V and 50 V as well as 60 V with data at the additional voltages used to map out the non-linearity in the high-energy range. The ^{55}Fe source emits X-rays with two characteristic energies of 5.9 keV and 6.5 keV [89]. The total phonon energy of the two characteristic lines at the applied voltages are calculated according to Ref. [32]. We use a 4th-order polynomial to model the MF integral as a function of the total phonon energy, as shown in Fig. 4.1. This parameterization is used to accommodate the high-energy data points which suffer from saturation effects. These effects cannot be described by a 2nd-order polynomial as they are intrinsically of higher order, driven by the response of a TES to large energy depositions. The resulting calibration function is denoted as $E_{\text{MF,Fe}}$, where the subscript ‘‘Fe’’ specifies that it is derived from data acquired on the dedicated high-energy calibration day.

To account for daily variation of the detector working point relative to Fe-day, we calculate a day-by-day correction factor as the ratio of the low-energy calibration’s linear-term coefficient between the i^{th} and Fe-day: $k_i = \frac{a_i}{a_{\text{Fe}}}$. We then scale the calibration function based on the MF integral by this multiplicative factor. The corrected high-energy calibration function for the i^{th} day is $E_{\text{MF},i} = k_i \cdot E_{\text{MF,Fe}}$.

We combine the low-energy and high-energy estimators with the smooth transition shown in Eq. (4.2),

$$(4.2) \quad E_{\text{ph}} = \begin{cases} E_{\text{OF}}, & E_{\text{OF}} < 600 \text{ eV} \\ (1 - c)E_{\text{OF}} + c \cdot E_{\text{MF}}, & 600 \text{ eV} \leq E_{\text{OF}} \leq 800 \text{ eV} \\ E_{\text{MF}}, & 800 \text{ eV} < E_{\text{OF}} \end{cases}$$

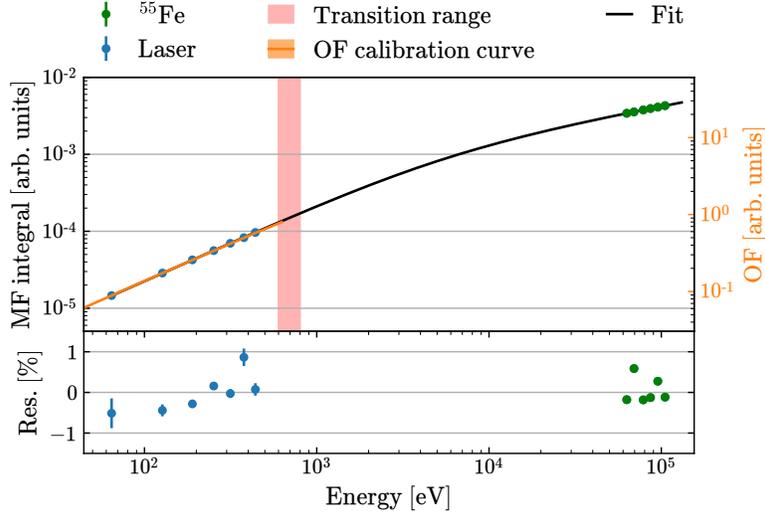


Figure 4.1. Top: Application of the energy calibration curve. The MF integral calibration curve (black) and a representative example of the OF calibration curve from Fe-day (orange, corresponds to Eq. 4.1). The OF calibration curve includes an $\sim 11\%$ systematic uncertainty band, and corresponds to the y-axis on the right side. The vertical red shaded region marks the 600–800 eV transition range. Bottom: residuals between the data points and the calibration curve expressed as a percentage.

in which $c = \frac{E_{\text{OF}} - 600 \text{ eV}}{200 \text{ eV}}$. E_{ph} is the calibrated total phonon energy of an event, and E_{OF} and E_{MF} are the energy of an event calculated by the low-energy and high-energy calibrations, respectively.

We note that the above calibration is derived with data collected in the HV mode. Ref. [32] shows that the calibration of the same detector for the 0V mode can be different from that for the HV mode by $\sim 11\%$. For this study, we use the above described calibration for both HV-mode and 0V-mode data, and use the difference as the systematic uncertainty of the calibration (as shown in Fig. 4.1). As shown in Sec. 4.6, this systematic uncertainty is negligible in the comparison of the 0V and HV mode data.

4.4. Dark matter constraints with 0V spectrum

In this section we consider the energy spectrum with zero bias across the crystal to constrain the spin-independent DM–nucleon elastic scattering cross section. We describe the live-time and event-based selection criteria and present the dark matter exclusion limit.

4.4.1. Live-time estimate

We apply the following live-time selection criteria to ensure a stable data-taking environment: (1) a fridge temperature selection discards time intervals with unstable ADR temperatures; (2) an average pre-pulse baseline selection removes time intervals that lie on the tail of high energy particle hits; and (3) a 120 Hz selection removes time intervals affected by the power-line noise. The selection criteria (1) and (2) are similar to the ones used in the electron recoil dark matter search in Ref. [31] with the only difference being that we use a time bin of 0.1 s instead of 1 s to preserve more live-time. The necessity of the 120 Hz selection (3) arises from the use of a much lower trigger threshold for the 0V data compared to the HV data in Ref. [31]. We observe that the trigger rate fluctuates with a 120 Hz frequency. We relate this feature to the power-line-induced noise and identify its phase by clustering triggered events in the phase vs. time plane as shown in Fig. 4.2, where the phase is defined as time modulo 1/120 s. The average phase of the event clusters varies in time due to the varying AC power phase relative to the stable data acquisition clock cycle. We fit the time-dependent phase trend of the increased trigger rate with a 5th-order polynomial and remove a 50% live-time band around the fit, as shown in Fig. 4.2. After applying all three live-time selection criteria to the 0 V data, the remaining science exposure is 0.185 gram·days.

4.4.2. Event-based selections

We perform pulse-shape-based selections to remove pulses not consistent with particle energy depositions in the region of interest (ROI) between 9.2 and 250 eV. The reduced- χ^2 , in both the time and frequency domains, between the pulses and the pulse template serves as the metric. We refer to the reduced- χ^2 as χ^2 in this paper for simplicity. We reject events for which the χ^2 quantity deviates from the corresponding mean of the laser calibration data by over three standard deviations, which rejects anomalous triggers such as those caused by electromagnetic interference (EMI) pickup. Figure 4.3 shows the energy spectrum of the dark matter search data before and after applying the χ^2 selections. The combined efficiency of the two selections is calculated as the passage fraction of the laser data with an energy-independent fit and is shown in Fig. 4.4. We tested how the selection-efficiency uncertainty affects the dark matter limit and found that even a large uncertainty of up to 20% is subdominant to the other uncertainties, as discussed in the following subsection. Therefore, the χ^2 selection-efficiency uncertainty is not included in the estimate of the systematic uncertainty shown in Fig. 5.

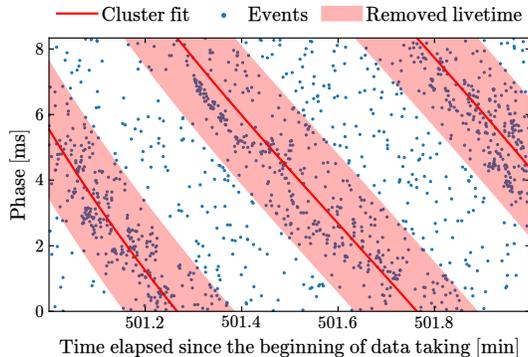


Figure 4.2. Triggered events in the phase vs. time plane (blue dots) with the 5th-order polynomial fit of the event clusters (red line) and the live-time removed by the 120 Hz selection (red area). The selection is used to identify and remove periods of high trigger rate associated with power-line-induced noise. The time interval shown in this figure corresponds to approximately 0.3% of the entire data sample analyzed in this report.

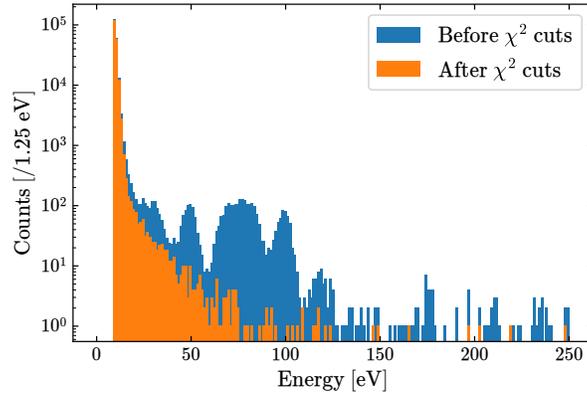


Figure 4.3. 0 V dark matter search energy spectrum before and after applying the χ^2 selections. The live-time selection criteria are applied to both spectra.

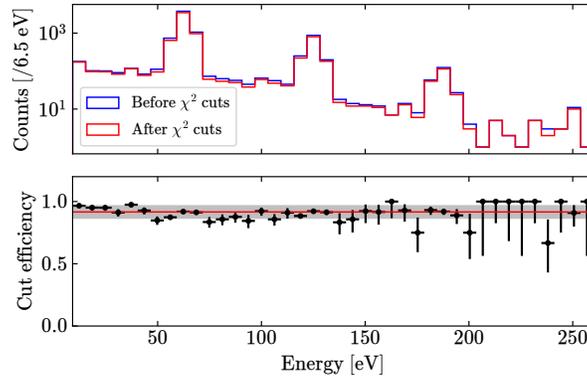


Figure 4.4. Evaluation of the cut efficiency. Top: 60 V laser calibration spectrum before (blue) and after (red) applying the χ^2 selections. Bottom: selection efficiency versus total phonon energy (black data points) fitted by an energy-independent efficiency model (red line) and 1σ statistical uncertainty (gray band).

4.4.3. Dark matter limit

We obtain an exclusion limit on the spin-independent DM-nucleon scattering cross section using a signal-only hypothesis and the data described in the previous subsection. The calculation uses the standard signal model in Ref. [90] with the following parameters: an asymptotic value of the Maxwellian velocity distribution $v_0 = 220$ km/s, a galactic escape velocity $v_{\text{esc}} = 544$ km/s,

a local DM mass density $\rho_0 = 0.3 \text{ GeV}/(c^2 \cdot \text{cm}^3)$ and a mean orbital velocity of the Earth $v_{\text{lab}} = 232 \text{ km/s}$ [91, 92, 93].

To account for the effect of detector resolution on the energy reconstruction, we perform a detector response simulation. We scale the pulse template to energies between 0.5 and 260 eV in 0.5 eV steps, and inject these scaled template pulses into randomly triggered noise traces collected throughout the data-taking period. We use the same triggering and energy-reconstruction algorithms that are used for the experimental data to reconstruct the energy of an injected pulse, thus obtaining detector response probability distributions $P(E'|E_0)$, where E_0 is the true energy of the injected pulses and E' is the reconstructed energy. We use a trigger-time selection to ensure that the triggered events correspond to the injected pulses. The dark matter signal model as a function of true energy is then convolved with the detector response probability distributions to construct the signal model as a function of reconstructed energy:

$$(4.3) \quad \frac{\partial R}{\partial E'}(E'|M_{\text{DM}}) = \Theta(E' - \delta)\varepsilon \times \int_{E_0=0\text{eV}}^{260\text{eV}} \left[\Theta(E' - E_0 + 3\sigma(E_0)) \times \Theta(E_0 + 3\sigma(E_0) - E') \times P(E'|E_0) \frac{\partial R}{\partial E_0}(E_0, M_{\text{DM}}) \right] dE_0.$$

Here $\frac{\partial R}{\partial E_0}$ is the differential DM-nucleon scattering rate, M_{DM} is the dark matter candidate mass, δ is the trigger threshold, and ε is the selection efficiency (assumed energy-independent in this analysis, see Fig. 4.4). The trigger efficiency is included in the detector response probability distributions $P(E'|E_0)$. The two Heaviside functions Θ inside the integral perform a 3σ cutoff of the detector response function, where $\sigma(E_0)$ is the width of the Gaussian fits to each $P(E'|E_0)$ distribution. This cutoff simplifies the numerical calculation by restricting the convolution of

the detector response with the signal model to a range of $\sim 1.7\text{eV}$ to 258.7eV and avoids an undefined recoil rate at zero energy.

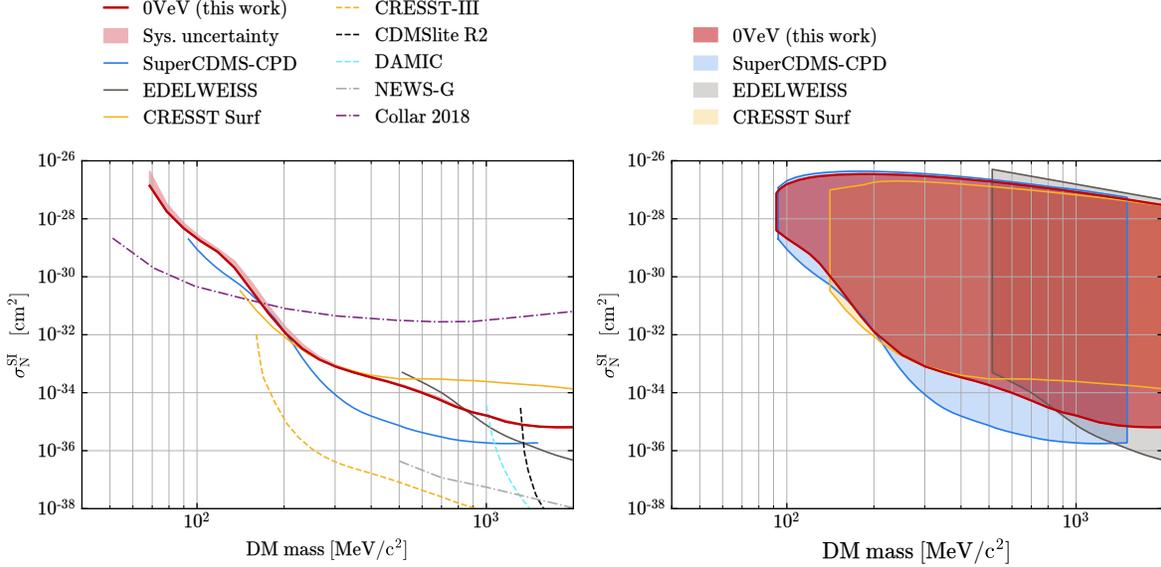


Figure 4.5. Left: 90% confidence-level exclusion limit on the spin-independent DM-nucleon scattering cross section. The result of this work is depicted in solid red with an estimate of the systematic uncertainties in light red. The results of other surface experiments using solid-state detectors are depicted as solid lines: blue for SuperCDMS-CPD [27], dark gray for EDELWEISS [22], and gold for CRESST-surface [25]. Underground searches using solid-state detectors are depicted as dashed lines: gold for CRESST-III [24], dark gray for CDMSlite [94], and cyan for DAMIC [95]. Other experimental constraints are shown as dash-dotted lines: light gray for NEWS-G [96] and purple for Collar [97]. Right: the same results with upper- and low-mass boundaries on the exclusion areas derived from the atmosphere and Earth shielding effect [27, 22, 98]. The upper boundary limits the low-mass reach of the current experiment to $92\text{MeV}/c^2$.

We utilize the Optimum Interval (OI) method [99, 100] to set a 90% confidence level exclusion limit on the DM-nucleon scattering cross section, using the experimental spectrum and the signal model described above. Figure 4.5 (left) compares our result to other experimental results in the same mass region [27, 22, 25, 24, 94, 95, 96, 97]. The systematic uncertainty propagated from the energy calibration uncertainty, discussed in Sec. III B, is shown as the

filled area. We estimate the systematic uncertainty by rescaling the energy calibration by 11% (see Fig. 4.1) and recalculating the limit. The resulting limit differs from the main result by up to $6\times$ at the lowest mass (up to $2\times$ at masses above $100 \text{ MeV}/c^2$). The other systematic uncertainties are not included in Fig. 5 as they were found to be subdominant: up to 20% from the uncertainties in the detector response simulation and less than 20% from the cut-efficiency uncertainty.

A very-low-energy threshold allows us to reach dark matter masses below $100 \text{ MeV}/c^2$, but the relatively high cross-section values in this mass range require us to consider the shielding by the atmosphere and Earth. At high values of the cross section, a presumed dark matter particle would not reach the detector with its original kinematic energy due to its interactions with the atmosphere and the Earth, therefore such cross-section values cannot be probed by our experiment. To calculate the upper bound on the cross-section exclusion region (Fig. 4.5, right), we use the `verne` package [101], which takes into account the mean direction of the DM flux at the location and the time of the experiment and estimates the impact of shielding on the standard halo model velocity distribution, assuming straight-line particle trajectories and continuous energy loss in the shielding (atmosphere and Earth). While these assumptions are in general only valid for high-mass particles ($> 10^5 \text{ GeV}/c^2$), a comparison with a more complete Monte Carlo approach demonstrates that the simplified approach used in the `verne` package leads to similar results [102]. Accounting for shielding removes the sensitivity of this analysis to dark matter masses below $92 \text{ MeV}/c^2$. To make a comparison to other experimental results in the same parameter space [27, 22, 98], we do not correct the lower bound of the exclusion region for shielding. However, this correction should be done in general at cross sections $\gtrsim 10^{-33} \text{ cm}^2$, especially for experiments probing new parameter spaces. Further efforts are required to consider shielding in the OI method, as it introduces a dependency of the DM spectrum shape on the value of the cross-section. In the current analysis, if the entire energy ROI is used instead of the

OI method, considering DM shielding would increase the lower bound of the exclusion region by a factor of ~ 2.1 at $100 \text{ MeV}/c^2$.

4.5. Pulse shape anomalies

We observe populations of events with pulse shapes different from the calibration data in the data-set even after the χ^2 cut. Anomalously shaped events exist in both the 0V and HV DM exposures with different characteristics. In the 0V data, we observe events that have a significantly longer pulse decay time than the laser-pulse shape. In HV data, we notice a large population of events with more than one pulse closely packed in time, which we refer to as “burst” events in this manuscript. Figure 4.6 shows one example of a burst event. To study these anomalous events, we do not use the event-based selections described in Sec. 4.4 because they tend to remove these events. We instead establish looser selections described in this section and use them to investigate the pulse shape anomalies in the 0V and HV data. We then discuss the pulse shape anomalies in 0V and HV data in the rest of this section.

4.5.1. Data selection

To study the pulse shape anomalies and facilitate the comparison of the 0V and HV datasets, we apply the same live-time selections (1) and (2) described in Sec. 4.4 to both datasets. We increase the analysis threshold for this investigation to 25 eV to avoid near-threshold noise effects such as the the 120 Hz power-line-induced noise events, which allows us to preserve more exposure because live-time selection (3) is not needed. The resulting exposures are 0.4 gram·days at 0 V, 0.7 gram·days at 60 V, and 1.7 gram·days at 100 V.

We use a loose χ^2 selection to remove trigger artifacts caused by the OF. We also use a pulse-width selection to reject EMI noise, for which the average pulse width is wider ($> 160 \mu\text{s}$)

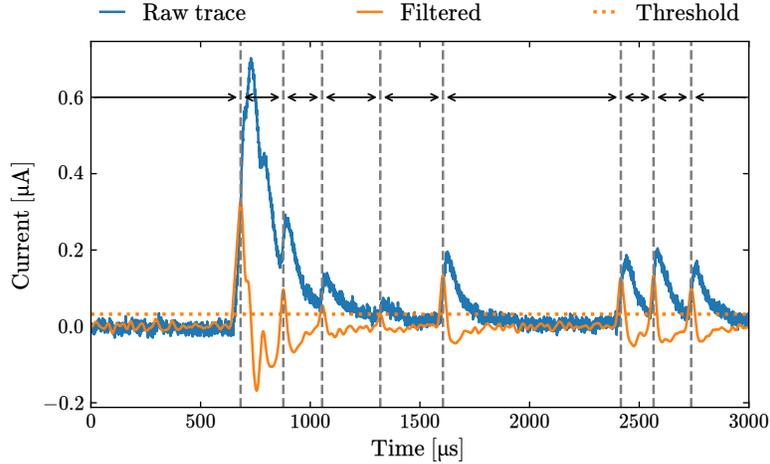


Figure 4.6. Example of a “burst” event at 60 V. The blue trace is the raw trace, whereas the orange trace results after applying a Gaussian derivative filter (described in Sec. 4.5.3), which peaks at the rising edges in the raw trace. The dotted orange line is the threshold for peak-finding. Each peak above the threshold in the filtered trace corresponds to a pulse in the raw trace. Note that the filter has limited time resolution, which results in the second pulse being below the threshold and not identified. The vertical dashed guide lines show the rising edge of the events identified above the threshold. The inter-arrival time of two events is defined as the time distance between their rising edge.

than for particle-interaction events ($< 100 \mu\text{s}$). The two selections are applied to both HV and 0V data. The selection efficiencies are evaluated in Sec. 4.5.

For the pulse-shape study reported in this section, we also remove a population of “slow events” from the 0V data. These events have pulse-decay times two orders of magnitude slower than the decay time for laser-calibration events. Such a slow time constant indicates that these events are the result of a different type of energy deposition in the detector. We discuss this class of events further in Sec. 4.7 B.

4.5.2. 0V mode: long-tail events

The χ^2 metric is sensitive to differences in pulse shape relative to the pulse template, and different event populations are apparent in the χ^2 versus reconstructed-energy plane (Fig. 4.7

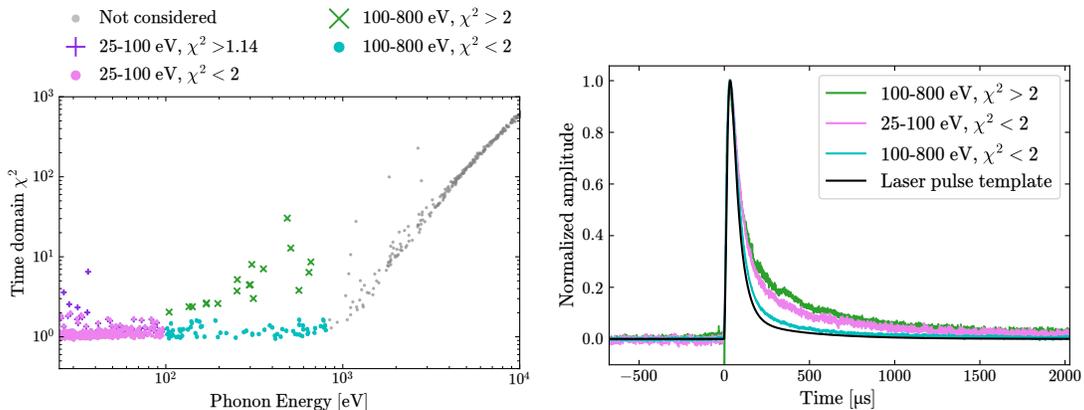


Figure 4.7. Event selection (left) of 0V data and averaged pulses of selected 0V events (right). Four groups of events are selected with two energy ranges (25–100 eV and 100–800 eV; events >800 eV, shown in gray, are not considered due to detector saturation) and two χ^2 ranges ($\chi^2 > 2$ and $\chi^2 < 2$). The 25 – 100 eV events with $\chi^2 > 2$ (marked with purple +) have square pulse shape which is consistent with Radio Frequency (RF) induced noise, and are not plotted in the right panel. The legend in the right panel is ordered from top to right with decreasing fall time of the averaged pulse shape.

left) for the 0V data. Using event selections in this plane, we create average pulses for each group (Fig. 4.7 right). We split the data into a low-energy region (up to 100 eV) where the signal-to-noise ratio is modest and a high-energy region from 100–800 eV where pulse-shape differences are more easily distinguishable by χ^2 . Events above 800 eV are subject to strong detector saturation effects and have hence been excluded in this pulse-shape study. For each energy region, we select events with a template-like shape with an empirical selection of $\chi^2 < 2$ and an anomalous shape with $\chi^2 > 2$. We compare these to the aforementioned template made with laser pulses. To rule out pulse-shape differences associated with different interaction types, we verified that this pulse template is also consistent with the pulse shape of nuclear recoil events both at 0 V and 100 V, using data taken at a neutron beam [103].

The average pulse of the anomalous $\chi^2 > 2$ events between 100 eV and 800 eV, shown in green in Fig. 4.7 (bottom), exhibits a pronounced slower decay time, or “long tail”, compared to

the pulse template. The average pulse of events in this energy range with $\chi^2 < 2$ is very similar in shape to the pulse template, see the cyan pulse in Fig. 4.7 (bottom). The small deviation of the 100-800 eV average pulse (cyan) from the template is a result of including some events with slight saturation and some of the long-tail events. As is visible in Fig. 4.7 (top) the discrepancy in χ^2 diminishes with decreasing energy and is close to our selection boundary at ~ 100 eV. Hence, we do not expect a full event-by-event separation of these long-tail events for the low-energy selection of $\chi^2 < 2$ events (in pink). Curiously, we observe an average pulse from this population that is much closer to the pulse shape of the anomalous events in the 100–800 eV range than that of the laser-pulse template. This suggests that the low-energy data are dominated by long-tail events.

4.5.3. HV mode: burst events

When the detector is operated in HV mode, we classify all events with more than one pulse in the 5.4 ms post-trigger time window as a burst event, as exemplified by the event shown in Fig. 4.6. We divide the pulses in a burst event into two categories: the primary pulse occurring at the trigger time of the event, and the secondary pulses occurring after the primary pulse. Pulses from both categories are treated as a single event.

To study the time distribution of the individual pulses, we identify the individual pulses inside a burst event with an edge detection algorithm. This algorithm searches for peaks after filtering the raw event with a first-order Gaussian derivative kernel. The inter-arrival time (dt) is defined as the time distance between sequential rising edges as shown in Fig. 4.6. The dt distribution of all pulses is shown in Fig. 4.8. If all the pulses were from a random Poissonian process with uncorrelated pile-up probabilities, the dt distribution would follow a single exponential function. We note that the distribution roughly follows such an exponential function in the region of $0.5 \text{ s} < dt < 1.5 \text{ s}$, while deviating from it at smaller and larger time scales. The

deviation at larger time scales suggests there may be long-time correlation between events, though this is not investigated in this report. Meanwhile, at smaller time scales the non-Poissonian component dominates. For example, within the post-trigger trace length of 5.4 ms, the Poissonian component contributes only 2% of all pulses. This suggests that the majority of the individual pulses in burst events are correlated in time and likely have a common origin.

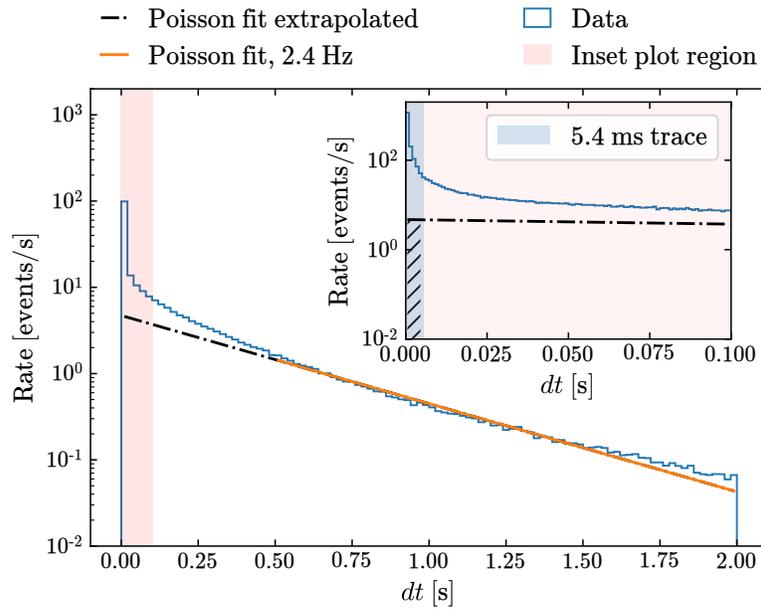


Figure 4.8. Individual-pulse inter-arrival time (dt) distribution of 100 V data. The bin width of the main plot is 0.02s. The inset panel is a zoom-in of the highlighted pink region. The blue area in the inset plot indicates the post-trigger duration used in our standard event-reconstruction algorithm.

We further characterize the burst events via the distribution of secondary-pulse arrival times relative to the primary pulse, Fig. 4.9. This time distribution is used later in Sec. 4.6 B to simulate burst events. The rate of secondary pulses decreases non-exponentially, which suggest there are multiple time scales.

The high rate of secondary pulses within a short time requires a special methodology to reconstruct their individual energies. First, we use a much shorter trace length of $\sim 150 \mu\text{s}$ as

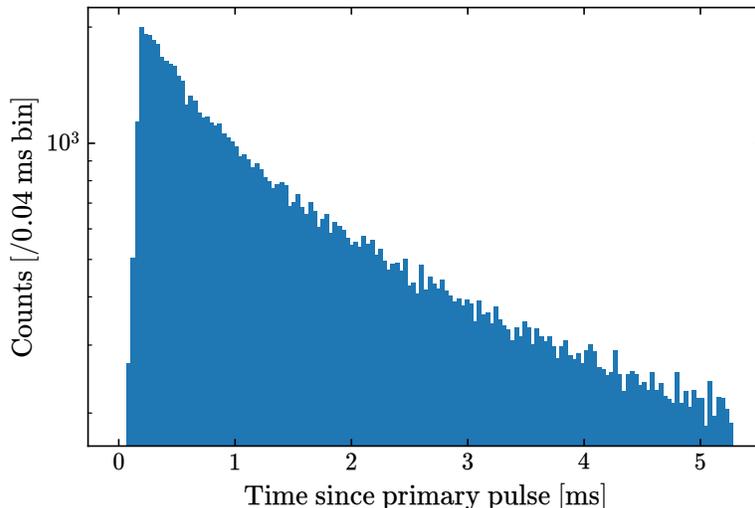


Figure 4.9. Time distribution of secondary pulses with respect to the primary (triggering) pulse in burst events from HV-mode data with the detector operated at 100 V. Note that the first five time bins (starting at 0 ms) have few counts due to the limited time resolution of the peak finding algorithm.

opposed to the 10.8 ms used in our standard event reconstruction. We then fit the pre-pulse baseline with an exponential function and subtract this function from the trace to minimize the impact of the preceding pulse on the reconstructed energy. Finally, we correct for the baseline-dependent gain variations as defined in Ref. [32] and use the best-fit OF amplitude to estimate the energy.

The energy spectra of the primary pulses and the secondary pulses are shown in Fig. 4.10. We note that the primary pulse energy goes up to several keV, while the secondary pulse energy peaks around the energy of a single e^-/h^+ . The energy of single e^-/h^+ events is given by the initial recoil energy E_r and the NTL phonon energy, $e \cdot V_{\text{NTL}}$. The distribution of secondary pulses peaks at ~ 2 eV above $e \cdot V_{\text{NTL}}$; this excess is interpreted as the recoil energy, where the systematic uncertainty of the energy calibration is estimated to $\mathcal{O}(1)$ eV.

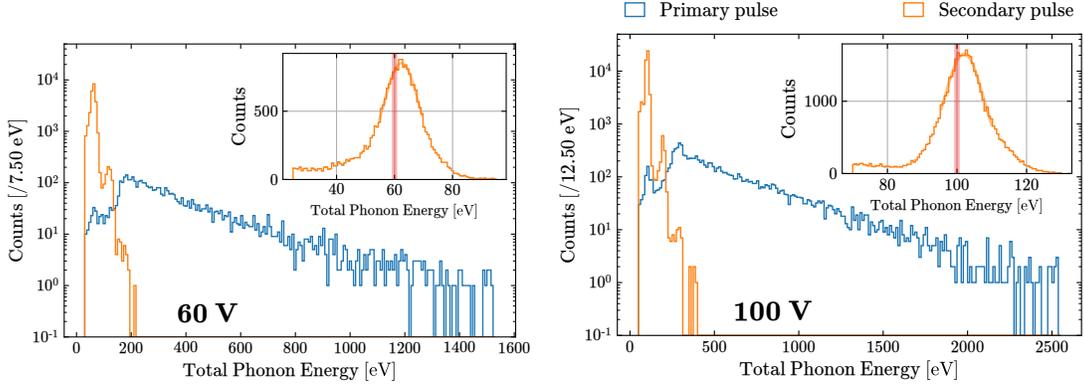


Figure 4.10. Energy of primary and secondary pulses of burst events in 60 V and 100 V data. The first peak in the primary pulse spectrum is consistent with 60 eV and 100 eV, respectively. The inset plot shows the zoom-in of around the energy of $1 e^-/h^+$ for secondary pulses. The red vertical line in the inset plot indicates the energy of the NTL contribution ($e \cdot V_{\text{NTL}}$) for a single e^-/h^+ , which is 60 eV and 100 eV, respectively.

4.6. Comparison between 0V and HV

The only difference between the 0V and HV datasets is the crystal voltage bias; so, we consider the possibility that the anomalous pulse shapes in the 0V data have the same origin as the burst events in the HV data. Under this assumption, we compare the 0V and HV pulse shapes based on ensemble averages which will be done in Sec. 4.6 A. In order to also make a spectral comparison and take into account potential effects of the event selections and detector response, we develop a burst event simulation to estimate the detector response for burst events with and without NTL amplification. The simulation is described in Sec. 4.6 B. Note that while we expect a nonzero voltage bias to introduce charge-leakage events in the HV data that will not be present in the 0V data, these events are below the energy region of interest for the comparison discussed in this section. We also note that we cannot rule out the alternative hypothesis that the crystal voltage bias can induce time correlated events. We will elaborate on this point in Sec. 4.7.

4.6.1. Pulse shape comparison

At 0 V, we cannot distinguish events with an energy that would typically produce a single e^-/h^+ from random noise fluctuations, making it difficult to identify potential burst events at 0 V. Thus, we focus on the averaged pulse shape when comparing between the 0V and HV data. We select 0V data in the energy range between 25 eV and 100 eV (pink events in Fig. 4.7). The 60 V events shown as orange dots in the χ^2 vs. energy plane in Fig. 4.11 are chosen to match this energy range with an NTL gain of 16.8, assuming $\epsilon_{\text{eff}} = 3.8$ eV [104]. Additionally, a subset of 60 V events that are not burst events (blue crosses in Fig. 4.11) are also selected at the higher end of this energy range, which have no more than one pulse identified within the 5.4 ms post-trigger window and thus are less likely to be burst events. We use this group of “non-burst” events from the HV data to produce an average pulse shape for events that have some saturation. The resulting averaged pulse shapes are shown in Fig. 4.12.

The average pulse shapes for both the 0V sample and the HV data burst events show visibly longer decay times than the laser-pulse template, which suggests the potential for these 0V and HV events to have a common origin. Conversely, the average pulse shape of the non-burst HV sample is similar to the laser-pulse template, indicating that detector saturation effects are unlikely to be the cause of the longer decays times in the other samples.

4.6.2. Burst event simulation

The different energy estimators—OF amplitude in the low-energy region, and MF integral in the high-energy region—have different sensitivities to secondary pulses, which is expected to lead to a systematic bias when scaling the HV-mode spectra for comparison to the 0V spectrum. We correct for this bias by applying a response matrix evaluated with the burst event simulations described below. We also use the burst event simulation to evaluate the event selection efficiencies.

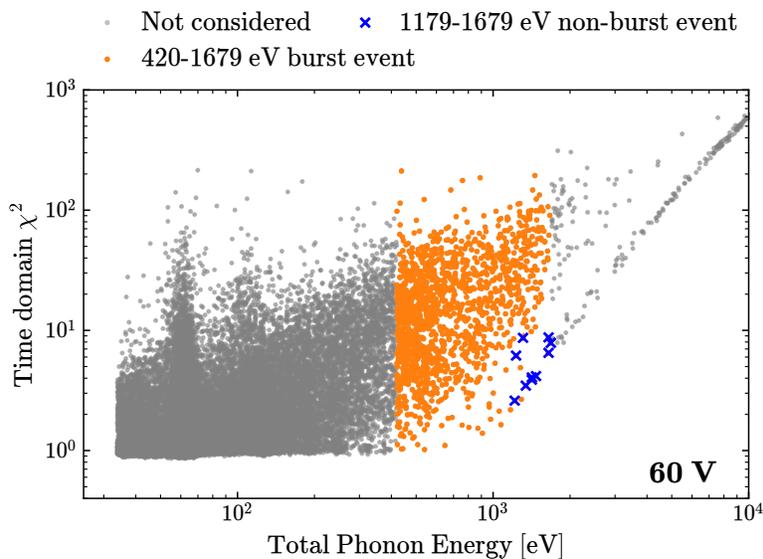


Figure 4.11. Selection of HV events for the comparison with 0V long-tail events. All events within 420-1679 eV and non-burst events at the higher end of that energy range are highlighted in orange dots and blue crosses, respectively.

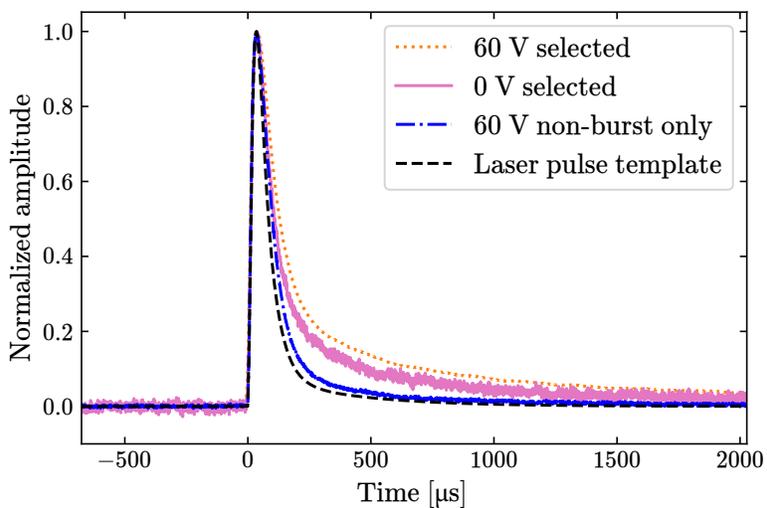


Figure 4.12. 0V vs. HV pulse-shape comparison. Pink and orange-dotted lines are the the average pulse shapes for the 0 V and 60 V events selected in Fig. 4.7 and 4.11. The black-dashed line is the laser-pulse template, which represents the non-saturated pulse shape. The blue line is the average pulse shape for the 60 V non-burst event sample and acts as a reference of the slightly saturated pulse shape.

We simulate the burst events with the time and energy distributions measured in the 60 V dataset. Burst events are characterized by the following parameters:

- Primary-pulse energy, E_p
- Number of secondary pulses, N_s
- Energy of the secondary pulses, E_s
- Time of each secondary pulse, $t_{s,i}$.

We modeled the distributions of E_p and $t_{s,i}$ with probability density functions extracted from the data, conforming to the distribution shown Fig. 4.9 and 4.10, respectively. E_s is set to 2 eV, which is consistent with single e^-/h^+ events. The distribution of N_s from data is shown in Fig. 4.13, and is modeled as a linear function of the energy of the primary pulse with a Gaussian distribution and standard deviation equal to its mean value, as a trial ansatz. The model with nominal parameters is shown as the center red line. The boundaries of the red shaded region, corresponding to double and half the number of secondary pulses compared to the red line, are chosen to bracket the mean number of secondary pulses we observed in data. We simulated three different scenarios corresponding to the red line and the upper and lower edges of the red shading region in Fig. 4.13.

We construct the trace of each event by summing a noise trace obtained from randomly triggered data, a primary pulse with the energy-dependent pulse shape empirically determined from calibration data, and N_s secondary pulses using the pulse template and onset times following the $t_{s,i}$ distribution. The simulated data sets are then reconstructed using the same algorithms as the detector data.

4.6.3. Energy spectra comparison

The energy spectra measured with a crystal voltage bias of 60 V and 100 V correspond to the total phonon energy with NTL gain, while the energy spectrum measured at 0 V represents

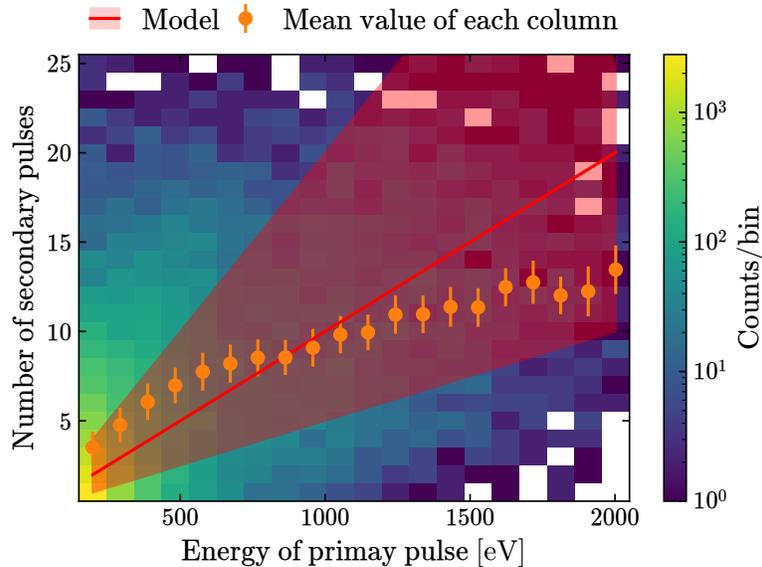


Figure 4.13. 2D histogram of the number of secondary pulses N_s as function of the primary pulse energy. Orange dots with error bar are the mean and standard deviation of each column of the histogram. The red line represents the relation between the number of secondary pulses and the energy of the primary pulse used in the burst event simulations. The boundaries of the red shaded area, corresponding to double and half the number of secondary pulses relative to the red line, are also simulated.

the recoil energy. The NTL gain depends on the averaged e^-/h^+ production energy, ϵ_{eff} . By comparing the spectra at different voltages we can estimate ϵ_{eff} of the anomalous events.

Before comparing the energy spectra, we correct the energy spectra for their event-selection efficiency. We evaluated the selection efficiency of the χ^2 and pulse width selections in the region of 25 – 150 eV of reconstructed recoil energy. We expect the 0V data to be a mix of both calibration-like events and the long-tailed events. The selection efficiency is thus evaluated on both the laser-calibration data and burst event simulation. We estimate the uncertainty for the latter from the three simulated secondary-pulse scenarios. We estimate the selection efficiency as the combination of the two efficiency curves and assign their total uncertainty as the systematic

uncertainty (see Fig. 4.14). We note that for the corresponding energy region in the 60 V and 100 V data, the selection efficiency evaluated with the burst event simulation is 100%.

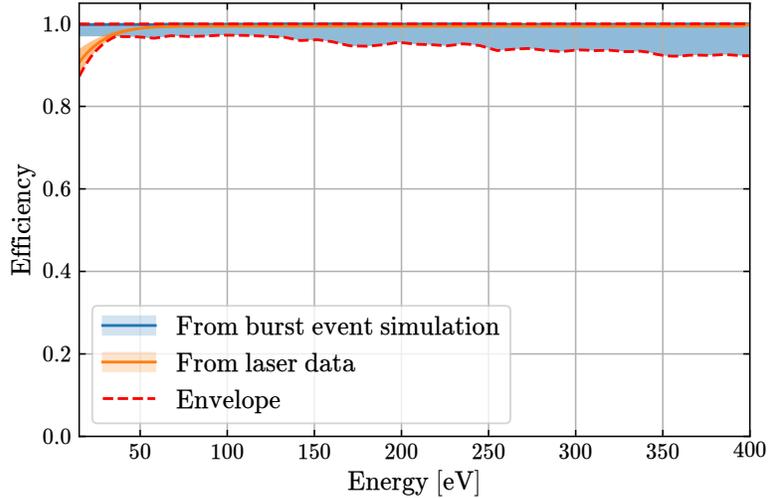


Figure 4.14. 0 V data selection efficiencies evaluated with laser calibration data (orange) and burst event simulation (blue), and associated uncertainties (shading). The dashed red lines are the envelope of the two uncertainty bands, which is used as the total uncertainty of the selection efficiency.

We then use response matrices to correct for the detector response difference between the HV mode and the 0V mode. The response matrices quantify the probability density function of an event being reconstructed in an energy bin with high voltage applied, provided that it is observed in a specific energy bin in the 0V data. The response matrices are evaluated with the burst event simulation. For each event in the simulation, traces at 0 V, 60 V, and 100 V are generated with ϵ_{eff} from 2-7 eV in steps of 0.5 eV. We processed the events at different voltages with the same algorithms as the detector data, and use the 2D histogram of the reconstructed energy of HV events versus 0V events to build response matrices. Examples of response matrices with the three different N_s models as described in Sec. 4.6 B are shown in Fig. 4.15, which also shows a fourth response matrix estimated from a simulation sample with no secondary pulses. We perform the correction by multiplying these matrices with the uncorrected recoil energy

spectra. For each HV-mode spectrum, we assign an envelope corresponding to the spread of the spectra calculated with the four matrices as the systematic uncertainty for the correction.

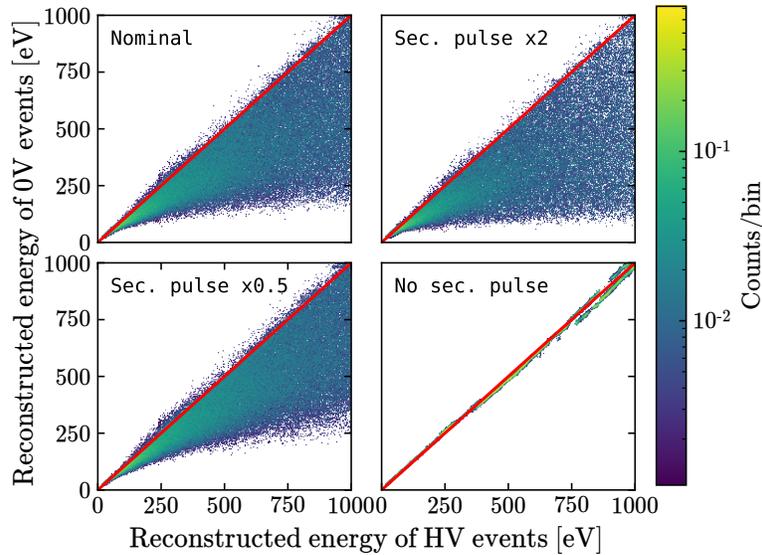


Figure 4.15. Response matrices that convert the 100 V spectrum to 0 V assuming $\epsilon_{\text{eff}} = 4 \text{ eV}$. The four panels correspond to different settings for the rate of secondary pulses in the burst event simulation.

Finally, we scan over ϵ_{eff} and compare the goodness of the fit (χ^2) between the converted HV spectra and the 0V spectrum in the recoil energy region of 25 eV to 150 eV. Figure 4.16 shows an example of the 0V spectrum along with the converted 100 V and 60 V spectra at $\epsilon_{\text{eff}} = 4 \text{ eV}$. We find that the converted HV spectra best match the 0V spectrum for an ϵ_{eff} of 4-5 eV, with a shallow minimum in χ^2 for these averaged e^-/h^+ production energies. We note that the χ^2 does not take into account the correlation and systematic uncertainties, thus we are not reporting the exact minimum and uncertainties of ϵ_{eff} . Figure 4.16 also shows the spectra before the conversion in the inset panel.

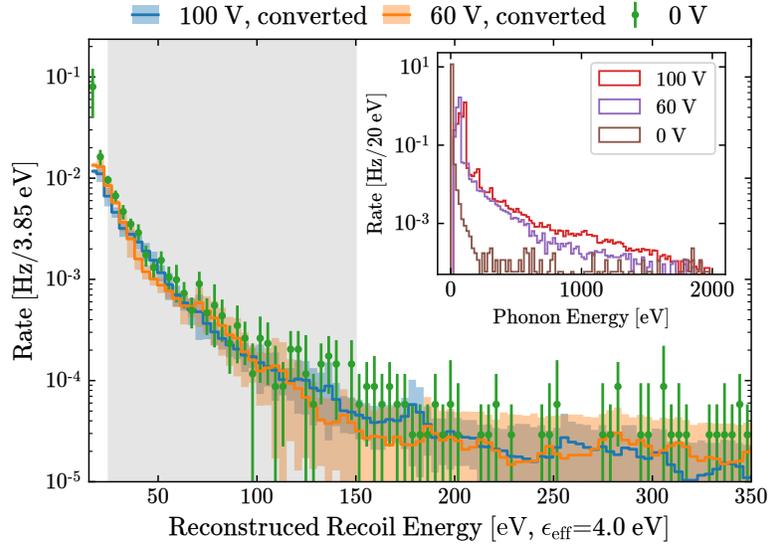


Figure 4.16. Comparison of the converted HV spectra with the 0V spectrum. The gray area shows the energy range (25-150 eV) where the χ^2 is calculated. The inset plot shows the phonon spectra before applying the response matrix conversion.

4.7. Discussion

The comparison of the pulse shapes and energy spectra in Sec. 4.6 suggests that the HV and 0V background may be dominated by events from the same origin. In this section, we discuss a model that is consistent with these observations drawing from the information in Sec. 4.6 and additional circumstantial evidences.

4.7.1. A possible explanation of the burst events

In Sec. 4.6 C we showed that the primary pulse of burst events has an ϵ_{eff} around 4-5 eV, with the assumption that the HV and 0V background events have the same predominant origin. There are at least two possible mechanisms that will result in an ϵ_{eff} close to 4-5 eV: 1) a single electron recoil event with an energy higher than 20 eV, which will have $\epsilon_{\text{eff}} = 3.8$ eV; 2) a group

of sub-10 eV electron recoil events that all occur within a couple of μs time scale (and thus look like a single higher-energy pulse) can have an ϵ_{eff} around 4-5 eV according to Ref. [105].

Furthermore, we found that the luminescence effect can explain what we have observed assuming that the primary pulse is a collection of 4-5 eV events. For example, SiO_2 , the primary component of the PCB that holds the detector, can create luminescence photons of 4.4 eV, 1.9 eV, and 2.7 eV with a decay time of 1.5 μs , 20 μs , and 7 ms, respectively [106, 107]. The energies and time scales of the 4.4 eV and 2.7 eV photons are consistent with the results of Sec. 4.5 and Sec. 4.6. The time constant of the 1.9 eV photons is close to the pulse fall time in our detector, and can be reconstructed as part of the primary pulse.

Besides luminescence, Cherenkov radiation and transition radiation have been suggested as possible sources of the low-energy excess seen in DM searches with an ER signal [108]. We do not evaluate these two mechanisms here because they will not produce a chain of events on the time scale of μs observed by our dominant source of background events, the burst events. They may become important once we can eliminate burst events.

4.7.2. Slow events

Interestingly, we also noticed a group of events in the 0 V dataset with a large slope in the pulse during the 5.4 ms post-trigger region. Upon further investigation, we found that all of these events have a long-timescale pulse with fall time > 10 ms following the initial pulse. Similar events also appear in 100 V data, as shown in Fig. 4.17. We refer to these events as “slow events”. We note that the first, fast pulses of the 0 V slow events have an average shape compatible with 0 V long-tail events within 0.5 ms, as shown in the inset plot in Fig. 4.17 top panel. We also note that about one third of the 100 V slow events are accompanied by a series of single e^-/h^+ size pulses, while the slow pulses are of similar sizes like those in the 0 V data.

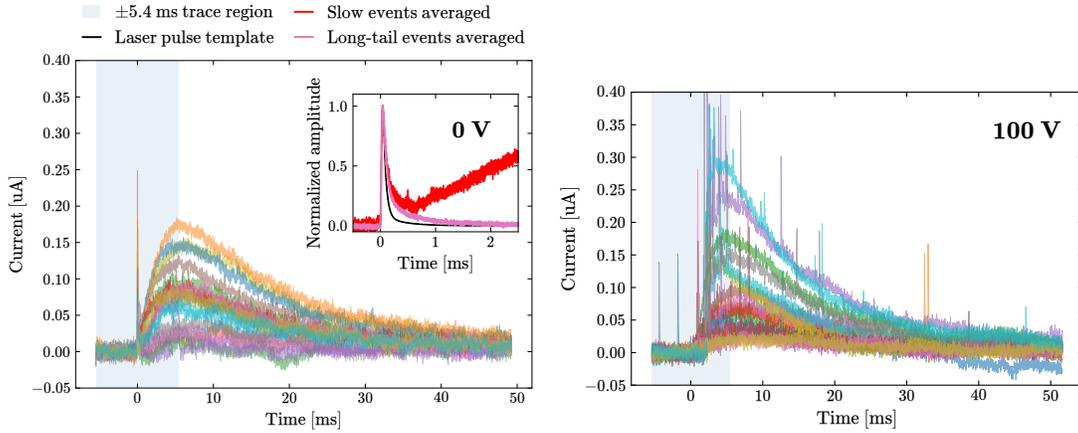


Figure 4.17. Example “slow events” that exhibit a second, slow pulse from 0 V (top) and 100 V (bottom) data. The shaded region shows the standard trace length that has been used elsewhere in this paper. The slow pulses extend far beyond the regular trace length. The inset plot of the top panel shows the zoom-in of the averaged pulse shape of the 0 V slow events in the main plot, compared with the averaged 0 V long-tail events (pink) as in Fig. 4.7 and the laser-pulse template.

The slow pulses could be from energy deposition of high energy particles in the detector holder PCBs of which we would expect a much longer time constant than of energy depositions in the detector directly. The energy deposition in the PCB may generate luminescence photons, some of which might then be absorbed in the HVeV detectors, causing slow pulses with single e^-/h^+ burst events as seen in the HV data. In 0V data these would show up as long-tail events combined with the slow pulses, consistent with our observation. The presence of these slow pulses with single e^-/h^+ burst events is then consistent with the luminescence explanation of the burst and long-tail events.

4.7.3. Ongoing development

Since we consider luminescence from the detector holder material to be a likely origin of these excess events, we have designed a new detector holder which minimizes insulator material inside the detector volume to reduce this potential background in our current science campaign.

Figure 4.18 shows the experimental setup of the new copper detector holder.

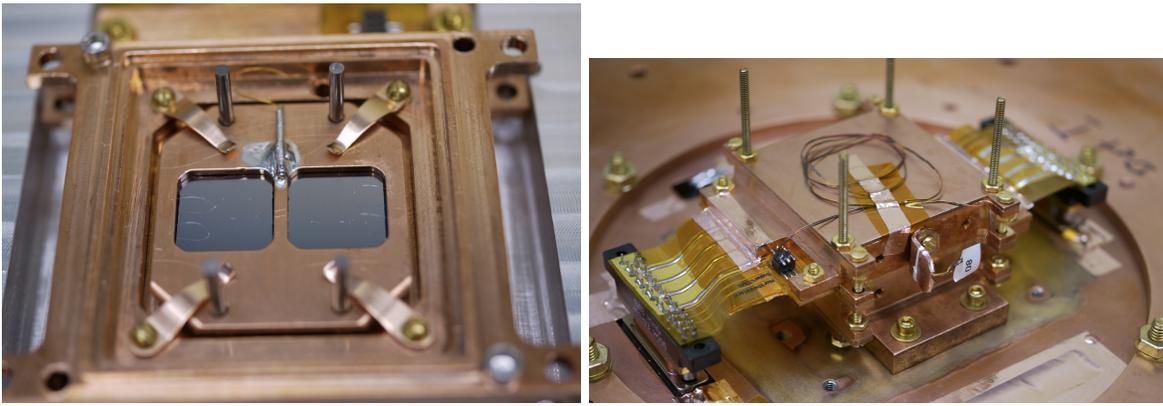


Figure 4.18. Picture of the new copper detector holder

The preliminary result shows a great reduction of background by about two orders of magnitude even without event selection. No more burst events are observed with the copper holder. However, there remains to be some excess of events at 0V mode. Figure 4.19 shows a comparison of energy spectrum with the new holder. The new setup shows a drastically decrease in the event rate, which proves the success of this analysis. Unfortunately, the scales are hidden and I cannot show a side by side comparison because SuperCDMS is currently using the data to perform a dark matter search.

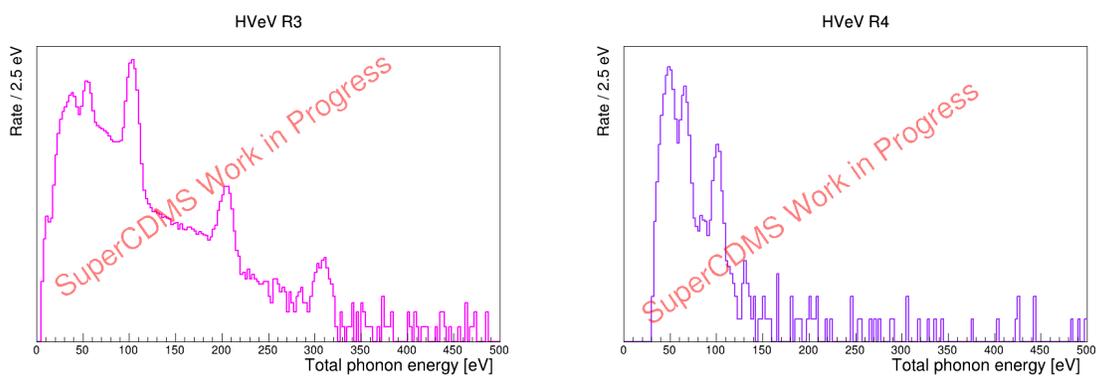


Figure 4.19. Comparison of the background energy spectrum with new (right) and old (left) detector holder. The y axis is hidden because the data in the same run is current in use for a dark matter search.

CHAPTER 5

Conclusion and outlook

In this thesis, I introduced the design and optimization of SuperCDMS-HVeV cryogenic calorimeters for dark matter searches. The HVeV detectors use Si/Ge as substrates and QETs as the phonon sensor with the capability of measuring charge signals when operated under high voltage. The low threshold of the HVeV detector makes it a powerful tool for exploring new parameter space of Dark Matter. HVeV can also be treated as a generic high-performance photon detector, which can be interesting to a broad range of low-background particle physics experiments.

I did two studies based on the HVeV detector. First, in Chapter 3, I calibrated the neutron ionization yield in Si to 100 eV, which is the lowest energy ever measured. This calibration helps to understand the material property and provides a trustworthy energy scale for WIMP and CE ν NS search. Second, in Chapter 4, I investigated the source of the low-energy background in the HVeV detector and identified a dominating origin. When this background is removed, we can better use the low threshold of the HVeV detector for better DM sensitivity.

In this chapter, I will discuss two interesting topics to be studied in the future. These topics are not only related to the HVeV detector but will benefit the main SuperCDMS SNOLAB experiment and the general DM-search community. Other important R&D perspectives, such as lowering the detector threshold and exploring new detector targets, are discussed in Noah Kurinsky's thesis [13].

5.1. Further study of the low-energy background

Although the previous dominating background source in the HVeV detector is found, there is a remaining unknown low-energy (LE) excess which limits the reach of the HVeV detector. The LE background is different in 0V and HV modes. In 0V mode, there is a roughly exponential increase in the event rate below $\mathcal{O}(100 \text{ eV})$. In HV mode, there is still a similar background as 0V but an additional peak at the energy of $1 \text{ e}^-/\text{h}^+$. This may look similar to what I discussed in Chapter 4. However, the picture is already much cleaner. After the burst events are eliminated, the 0V-mode excess is slightly reduced, and the HV-mode excess decreases to a peak at $1 \text{ e}^-/\text{h}^+$ instead of a broad range of energies. In this section, I will discuss the remaining background at 0V and HV, respectively.

5.1.1. 0V background

In 0V mode, the current situation (after getting rid of the burst events) of HVeV looks more similar to the commonly observed LE excess in different experiments, including SuperCDMS-CPD [109], EDLEWEISS [22, 23], and CRESST [24]. Finding out mitigation to this background has become a top priority since it is limiting the reach of the DM search. The community started a yearly workshop called EXCESS [110] to communicate the understanding of this background. A list of properties is found to be consistent across multiple experiments:

- **Non-ionizing:** no significant charge signal is generated during these events. There could still be a small amount of ionization
- **Time and temperature dependency:** the background rate decreases almost exponentially with respect to time. Heating up the detector to a few K will dramatically increase the rate, but the rate will start to decay again once the detector is cooled to $\sim\text{mK}$ operation temperature.

- **Stress dependency:** more stress on the detector is usually associated with more background events
- **Similar rate above/below ground**
- **Rate varies significantly between detectors**

We call this background “zero-charge low energy excess (0QLEE)” based on the feature that they generate almost no charge. The origin of these events can be narrowed down with the observed properties. It is unlikely to be electronic noise since the events have the same phonon pulse shape as calibration events. It is unlikely to be a high-energy photon/neutron since the rate does not change when moved underground with an improved shield. It is unlikely to be surface contamination since the rate decreases over time and surges when the temperature rises. It is unlikely to be some meta-stable electronic state because the energy scale of condensed matter is on the eV scale, which will need multiple particles rearranging at the same time. Moreover, there is no significant charge signal observed.

The exclusions above lead to one hypothesis that is consistent with all the evidence so far — stress-induced microfractures. Stress can come from three different levels: mounting structure, sensor films on the substrate, and localized stress in the substrate. Reference [111] shows that 0QLEE is still there after the stress in the mounting structure is minimized, so the next step would be studying the stress in sensor films. There is already ongoing effort measuring the stress in sensor film. In the meanwhile, we can use the same sensor design on a much larger substrate to test if the film stress is the issue. We expect the 0QLEE background rate to be similar when we use a larger substrate with the same sensor.

5.1.2. HV background

In HV mode, we have the aforementioned 0QLEE background plus an extra amount of “quantized” events — events with the energy around $1 e^-/h^+$. The $1 e^-/h^+$ peak is still almost

two orders of magnitude higher than the 0QLEE background and is the limiting factor for electron-recoil dark matter search.

We have gathered much information about the $1 e^-/h^+$ background. I will try to summarize the information we have and the upcoming tests.

1) The $1 e^-/h^+$ event rate is highly dependent on the operation history of the detector. We have found that the lowest stable rate is achieved by doing “pre-biasing”—keeping the detector at a voltage higher than the operating voltage for a time period before the operation. For example, if we would like to operate the detector at 100 V, then pre-biasing at 120 V for an hour before the operation will get us the lowest rate. If we directly apply 100 V, the rate starts high but keeps decreasing with a time constant on the order of days. We did not test if the rate without pre-biasing would eventually decrease to the lowest rate with pre-biasing.

2) There is an additional sub- $1 e^-/h^+$ peak that may show up after the pre-biasing; see the peak in the orange histogram in Fig. 4.19 right panel. The sub- $1 e^-/h^+$ peak varies significantly among detectors, and a higher pre-biasing voltage creates more sub- $1 e^-/h^+$ leakage. The rate of sub- $1 e^-/h^+$ events decays with a time constant of a few hours. The sub- $1 e^-/h^+$ leakage does not affect the ERDM search but can limit the usage of this detector as a general photon detector. It can also be helpful in understanding the $1 e^-/h^+$ leakage.

3) The $1 e^-/h^+$ rate depends on the polarity of the voltage. We see about a 10-100 times higher rates at negative voltages (E field from the TES to the backing grid) even with pre-biasing.

4) Some high-energy events will trigger a “breakdown”. We see that the rate of $1 e^-/h^+$ surges from 0.01 Hz to 1 Hz after some high-energy events, but not all of them. The rate will then decrease with a time constant of a few hours.

We do not find a single mechanism that can explain all the phenomena. Since changing the polarity has the most significant effect on the event rate, this suggests that the metal-semiconductor surface is playing some role in generating the leakage. We have made the third

generation of the HVeV detector by adding a SiO₂ insulation layer between the silicon substrate and the aluminum/tungsten deposition. This detector is currently being tested.

In the short term, there are still ways to improve our sensitivity for ERDM with the existing setup before the source of 0QLEE and 1 e⁻/h⁺ leakage is figured out. First, we can use the second e⁻/h⁺ peak to set an exclusion limit for ERDM because the leakage rate is much lower at 2 e⁻/h⁺. We can further boost the operating voltage to push the 2 e⁻/h⁺ peak to higher energy where the 0QLEE background is lower. Second, we can try to model the 0QLEE background at 0V and subtract it from the HV energy spectrum. This approach will need a long exposure to not be limited by statistical uncertainty.

5.2. Understanding the calibration

The discrepancy between the 0V and HV mode is one remaining major question about the calibration. We see in Fig. 2.12 that the ⁵⁵Fe and ⁵⁷Co γ peaks are on a different trend from the HV calibration curve. For example, in HV mode, a laser event with 24 e⁻/h⁺ at 250 V should generate a similar amount of total phonon energy as a 6 keV ⁵⁵Fe X-ray peak in 0V mode. However, the pulse integral of the laser event is 20% higher than the ⁵⁵Fe. This discrepancy implies two possibilities: either the estimation of total phonon energy is different, or the same total phonon energy gives a different response. Understanding this discrepancy is critical to get a more accurate energy scale of the detector.

If we assume that the same total phonon energy has a different response when voltage is applied, there is a hypothesis already discussed in Chapter 2: the NTL phonons have a different response with respect to the phonons generated by charge recombination; the penetration length of X-rays in Si ($\sim 30 \mu\text{m}$) is not sufficient to reach the bulk, and there is some signal degradation due to surface effects; the deposition of a single X-ray could generate local saturation in the sensor because the ⁵⁵Fe source was directed at the QET-instrumented face.

There is another possibility that the total phonon energies are actually different from the naive estimation. In Chapter 1, we discussed that the charge carriers might not recombine within the time window of a normal event, which will retain part of the recoil energy. Depending on how much recombination happened, in 0V mode, the phonon energy can be anywhere between the recoil energy E_r and the recoil energy minus the gap energy that all charge carries holds $E_r - n_{eh} \cdot E_{gap}$. The energy loss of $n_{eh} \cdot E_{gap}$ is a significant fraction of the expected phonon energy in 0V mode for gammas:

$$(5.1) \quad \frac{\Delta E}{E_{ph}} = \frac{n_{eh} \cdot E_{gap}}{E_r} = \frac{E_r/\epsilon_\gamma \cdot E_{gap}}{E_r} \sim 0.31$$

in which we use $E_{gap}=1.2$ eV. However, in HV mode, this effect will be relatively much smaller because the energy loss is the same, but the total energy is amplified by G_{NTL} :

$$(5.2) \quad \frac{\Delta E}{E_{ph}} = \frac{n_{eh} \cdot E_{gap}}{E_r \cdot G_{NTL}} \sim \frac{E_r/\epsilon_\gamma \cdot E_{gap}}{E_r \cdot V/\epsilon_\gamma} \sim 0.31/V.$$

where V is the voltage across the crystal. Thus, this hypothesis can explain with the observation that the ^{55}Fe at 0V gives a smaller phonon signal than laser events.

A similar argument can be made for nuclear recoils. Nuclear recoil dislocates atoms from their place in the lattice and creates vacancies (Frenkel defects). Reference [112] has a thorough description of the evolution of a collision cascade process. Minimum energy is needed for the dislocation to happen, which is called the binding energy. If the dislocated atom does not return to the vacancy and gives back this part of binding energy, the total phonon energy will be smaller. The modeling of the recombination of damage is still an open question.

Further tests are needed to study how much energy is retained in the form of gap energy and binding energy, and it is a necessary step when high-precision calibration of nuclear/electronic recoil is needed.

5.3. Dark matter experiments outlook

The traditional WIMP and the LDM are very promising dark matter models to detect. With the studies I summarized above, the dark matter search for WIMP and LDM using SuperCDMS technologies can be greatly improved to reach a lower mass and better sensitivity. It will benefit the community of dark matter search and CE ν NS if the origin of the 0QLEE is figured out following the discussion in this section.

References

- [1] Marco Battaglieri, Alberto Belloni, Aaron Chou, Priscilla Cushman, Bertrand Echenard, Rouven Essig, Juan Estrada, Jonathan L Feng, Brenna Flaughner, Patrick J Fox, et al. Us cosmic visions: new ideas in dark matter 2017: community report. *arXiv preprint arXiv:1707.04591*, 2017.
- [2] A Arbey and F Mahmoudi. Dark matter and the early universe: a review. *Progress in Particle and Nuclear Physics*, 119:103865, 2021.
- [3] JD Lewin and PF Smith. Review of mathematics, numerical factors, and corrections for dark matter experiments based on elastic nuclear recoil. *Astroparticle Physics*, 6(1):87–112, 1996.
- [4] E. Corbelli and P. Salucci. The extended rotation curve and the dark matter halo of m33. *Monthly Notices of the Royal Astronomical Society*, 311(2):441–447, jan 2000.
- [5] Gerard Jungman, Marc Kamionkowski, and Kim Griest. Supersymmetric dark matter. *Physics Reports*, 267(5-6):195–373, 1996.
- [6] Benjamin W Lee and Steven Weinberg. Cosmological lower bound on heavy-neutrino masses. *Physical Review Letters*, 39(4):165, 1977.
- [7] Rouven Essig, Marivi Fernandez-Serra, Jeremy Mardon, Adrian Soto, Tomer Volansky, and Tien-Tien Yu. Direct detection of sub-gev dark matter with semiconductor targets. *Journal of High Energy Physics*, 2016(5):1–54, 2016.
- [8] Jim Alexander, Marco Battaglieri, Bertrand Echenard, Rouven Essig, Matthew Graham, Eder Izaguirre, John Jaros, Gordan Krnjaic, Jeremy Mardon, David Morrissey, et al. Dark sectors 2016 workshop: community report. *arXiv preprint arXiv:1608.08632*, 2016.
- [9] MF Albakry, I Alkhatib, DWP Amaral, T Aralis, T Aramaki, IJ Arnquist, I Atae Langroudy, E Azadbakht, S Banik, C Bathurst, et al. A strategy for low-mass dark matter searches with cryogenic detectors in the supercdms snolab facility. *arXiv preprint arXiv:2203.08463*, 2022.
- [10] K. D. Irwin, S. W. Nam, B. Cabrera, B. Chugg, and B. A. Young. A quasiparticle-trap-assisted transition-edge sensor for phonon-mediated particle detection. *Rev. Sci. Instrum.*, 66:5322, 1995.

- [11] K. D. Irwin and G. C. Hilton. *Transition-Edge Sensors*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2005.
- [12] M. C. Pyle. *Optimizing the design and analysis of cryogenic semiconductor dark matter detectors for maximum sensitivity*. PhD thesis, Stanford University, Phys. Dept., 2012.
- [13] N. A. Kurinsky. *The Low-Mass Limit: Dark Matter Detectors with eV-Scale Energy Resolution*. PhD thesis, Stanford U., Phys. Dept., 2018.
- [14] Francesco Giazotto, Tero T. Heikkilä, Arttu Luukanen, Alexander M. Savin, and Jukka P. Pekola. Opportunities for mesoscopics in thermometry and refrigeration: Physics and applications. *Rev. Mod. Phys.*, 78:217, 2006.
- [15] C. W. Fink, S. L. Watkins, T. Aramaki, P. L. Brink, S. Ganjam, B. A. Hines, M. E. Huber, N. A. Kurinsky, R. Mahapatra, N. Mirabolfathi, et al. Characterizing TES power noise for future single optical-phonon and infrared-photon detectors. *AIP Adv.*, 10:085221, 2020.
- [16] Stanislav Koc. The quantum efficiency of the photo-electric effect in germanium for the 0.3–2 μ wavelength region. *Czechosl. J. Phys.*, 7:91, 1957.
- [17] K. Ramanathan and N. Kurinsky. Ionization yield in silicon for eV-scale electron-recoil processes. *Phys. Rev. D*, 102:063026, 2020.
- [18] B. S. Neganov and V. N. Trofimov. Calorimetric method measuring ionizing radiation. *Otkryt. Izobret.*, 146:215, 1985.
- [19] P. N. Luke. Voltage-assisted calorimetric ionization detector. *J. Appl. Phys.*, 64:6858, 1988.
- [20] F. Ponce, C. Stanford, S. Yellin, W. Page, C. Fink, M. Pyle, B. Sadoulet, B. Serfass, S. L. Watkins, P. L. Brink, et al. Measuring the impact ionization and charge trapping probabilities in SuperCDMS HVeV phonon sensing detectors. *Phys. Rev. D*, 101:031101(R), 2020.
- [21] M Wilson. *A New Search for Low-mass Dark Matter and an Examination and Reduction of the Uncertainty due to the Photoelectric Absorption Cross Section using a Cryogenic Silicon Detector with Single-charge Sensitivity*. PhD thesis, University of Toronto, 2022.
- [22] E. Armengaud et al. Searching for low-mass dark matter particles with a massive Ge bolometer operated above ground. *Phys. Rev. D*, 99:082003, 2019.
- [23] Q. Arnaud et al. First Germanium-Based Constraints on Sub-MeV Dark Matter with the EDELWEISS Experiment. *Phys. Rev. Lett.*, 125:141301, 2020.
- [24] A. H. Abdelhameed et al. First results from the CRESST-III low-mass dark matter program. *Phys. Rev. D*, 100(10):102002, 2019.

- [25] G. Angloher et al. Results on MeV-scale dark matter from a gram-scale cryogenic calorimeter operated above ground. *Eur. Phys. J. C*, 77:637, 2017.
- [26] J. Rothe, G. Angloher, F. Ardellier-Desages, A. Bento, L. Canonica, A. Erhart, N. Ferreiro, M. Friedl, V.M. Ghete, D. Hauff, et al. NUCLEUS: Exploring Coherent Neutrino-Nucleus Scattering with Cryogenic Detectors. In *18th International Workshop on Low Temperature Detectors*, volume 199, page 433, Milano, Italy, 2019.
- [27] I. Alkhatib et al. Light dark matter search with a high-resolution athermal phonon detector operated above ground. *Phys. Rev. Lett.*, 127:061801, Aug 2021.
- [28] L. Barak et al. SENSEI: Direct-Detection Results on sub-GeV Dark Matter from a New Skipper CCD. *Phys. Rev. Lett.*, 125:171802, 2020.
- [29] A. Aguilar-Arevalo et al. Results on low-mass weakly interacting massive particles from a 11 kg-day target exposure of DAMIC at SNOLAB. *Phys. Rev. Lett.*, 125:241803, 2020.
- [30] R. Agnese et al. First Dark Matter Constraints from a SuperCDMS Single-Charge Sensitive Detector. *Phys. Rev. Lett.*, 121:051301, 2018.
- [31] D. W. Amaral et al. Constraints on low-mass, relic dark matter candidates from a surface-operated SuperCDMS single-charge sensitive detector. *Phys. Rev. D*, 102(9):091101, 2020.
- [32] R. Ren et al. Design and characterization of a phonon-mediated cryogenic particle detector with an ev-scale threshold and 100 keV-scale dynamic range. *Phys. Rev. D*, 104(3):032010, 2021.
- [33] Ziqing Hong. SuperCDMS IMPACT: an Ionization Yield Calibration Program. APS April Meeting, 2020.
- [34] Z. Hong, R. Ren, N. Kurinsky, E. Figueroa-Feliciano, L. Wills, S. Ganjam, R. Mahapatra, N. Mirabolfathi, B. Nebolsky, H. D. Pinckney, et al. Single electron-hole pair sensitive silicon detector with surface event discrimination. *Nucl. Instrum. Meth. A*, 963:163757, 2020.
- [35] For more information about Eccosorb, a microwave energy dissipating material, see <http://www.eccosorb.com/>.
- [36] For more information on the Thorlabs SM450 single mode fiber, see <https://www.thorlabs.com/thorproduct.cfm?partnumber=SM450>.
- [37] For more information on the Thorlabs LPS-635-FC laser diode employed, see <https://www.thorlabs.com/thorproduct.cfm?partnumber=LPS-635-FC>.
- [38] John Clarke and Alex I Braginski. *The SQUID handbook*, volume 1. Wiley Online Library, 2004.

- [39] Hans-Christian Stahl. *Cryogenic Particle Detection*. Springer Science & Business Media, 2005.
- [40] E. Gatti and P. F. Manfredi. Processing the signals from solid-state detectors in elementary-particle physics. *Riv. Nuovo Cim.*, 9:1, 1986.
- [41] J. W. Fowler, B. K. Alpert, W. B. Doriese, D. A. Fischer, C. Jaye, Y. I. Joe, G. C. O’Neil, D. S. Swetz, and J. N. Ullom. Microcalorimeter spectroscopy at high pulse rates: a multi-pulse fitting technique. *Astrophys. J. Suppl. Ser.*, 219:35, 2015.
- [42] S Di Domizio, F Orio, and M Vignati. Lowering the energy threshold of large-mass bolometric detectors. *J. Instrum.*, 6(02):P02007, feb 2011.
- [43] John Canny. A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, (6):679–698, 1986.
- [44] R. K. Romani, P. L. Brink, B. Cabrera, M. Cherry, T. Howarth, N. Kurinsky, R. A. Moffatt, R. Partridge, F. Ponce, M. Pyle, et al. Thermal detection of single e-h pairs in a biased silicon crystal detector. *Appl. Phys. Lett.*, 112:043501, 2018.
- [45] D. W. Amaral et al. Constraints on low-mass, relic dark matter candidates from a surface-operated SuperCDMS single-charge sensitive detector. *Phys. Rev. D*, 102:091101(R), 2020.
- [46] C. W. Fink et al. Performance of a large area photon detector for rare event search applications. *Appl. Phys. Lett.*, 118:022601, 2021.
- [47] A. Alessandrello, C. Brofferio, C. Bucci, O. Cremonesi, A. Giuliani, B. Margesin, A. Nucciotti, M. Pavan, G. Pessina, E. Previtali, et al. Methods for response stabilization in bolometers for rare decays. *Nucl. Instrum. Meth. A*, 412:454, 1998.
- [48] A. H Abdelhameed et al. First results from the cress-t-iii low-mass dark matter program. *Phys. Rev. D*, 100(10):102002, 2019.
- [49] E. Armengaud et al. Searches for electron interactions induced by new physics in the edelweiss-iii germanium bolometers. *Phys. Rev. D*, 98:082004, Oct 2018.
- [50] R Agnese et al. Projected sensitivity of the supercdms snolab experiment. *Phys. Rev. D*, 95(8):082002, 2017.
- [51] Hao Ma et al. Results of direct dark matter detection with CDEX experiment at CJPL. *J. Phys. Conf. Ser.*, 1468(1):012070, feb 2020.
- [52] Orr Abramoff et al. Sensei: Direct-detection constraints on sub-gev dark matter from a shallow underground run using a prototype skipper ccd. *Phys. Rev. Lett.*, 122:161801, Apr 2019.

- [53] A. Aguilar-Arevalo et al. Results on low-mass weakly interacting massive particles from an 11 kg d target exposure of damic at snolab. *Phys. Rev. Lett.*, 125:241803, Dec 2020.
- [54] D Akimov et al. Observation of coherent elastic neutrino-nucleus scattering. *Science*, 357(6356):1123–1126, 2017.
- [55] Belkis Cabrera-Palmer and David Reyna. Development of coherent germanium neutrino technology (cogent) for reactor safeguards. *Sandia Report SAND2012-8021*, 2012.
- [56] Henry T Wong. Low energy neutrino physics at the kuo-sheng reactor neutrino laboratory in taiwan. *Nucl. Phys. B-Proceedings Supplements*, 138:333–336, 2005.
- [57] V Belov et al. The ν gen experiment at the kalinin nuclear power plant. *J. Instrum.*, 10(12):12011, 2015.
- [58] C Buck et al. The conus experiment—coherent elastic neutrino nucleus scattering. In *15th Intern. Conf. on Topics in Astroparticle and Underground Physics, TAUP2017, Sudbury, Canada, July 24*, volume 28, 2017.
- [59] J Billard et al. Coherent neutrino scattering with low temperature bolometers at chooz reactor complex. *J. Phys. G*, 44(10):105101, 2017.
- [60] G Agnolet et al. Background studies for the miner coherent neutrino scattering reactor experiment. *Nucl. Instrum. Methods A*, 853:53–60, 2017.
- [61] A Aguilar-Arevalo et al. The connie experiment. *J. Phys. Conf. Ser.*, 761:012057, 2016.
- [62] Jens Lindhard, V Nielsen, M Scharff, and P. V Thomsen. Integral equations governing radiation effects. *Mat. Fys. Medd. Dan. Vid. Selsk*, 33(10):1–42, 1963.
- [63] Y Sarkis, Alexis Aguilar-Arevalo, and Juan Carlos D’Olivo. Study of the ionization efficiency for nuclear recoils in pure crystals. *Phys. Rev. D*, 101(10):102001, April 2020.
- [64] A. R Sattler. Ionization produced by energetic silicon atoms within a silicon lattice. *Phys. Rev.*, 138(6A):A1815, 1965.
- [65] G Gerbier et al. Measurement of the ionization of slow silicon nuclei in silicon for the calibration of a silicon dark-matter detector. *Phys. Rev. D*, 42(9):3211, 1990.
- [66] Brian L Dougherty. Measurements of ionization produced in silicon crystals by low-energy silicon atoms. *Phys. Rev. A*, 45(3):2104, 1992.
- [67] A. E Chavarria et al. Measurement of the ionization produced by sub-keV silicon nuclear recoils in a ccd dark matter detector. *Phys. Rev. D*, 94(8):082007, 2016.
- [68] F Izraelevitch et al. A measurement of the ionization efficiency of nuclear recoils in silicon. *J. Instrum.*, 12(06):06014, 2017.

- [69] J. I. Collar, A. R. L. Kavner, and C. M. Lewis. Germanium response to sub-keV nuclear recoils: A multipronged experimental characterization. *Phys. Rev. D*, 103:122003, Jun 2021.
- [70] M. F. Albakry et al. arXiv:2202.07043.
- [71] Peter Sorensen. Atomic limits in the search for galactic dark matter. *Phys. Rev. D*, 91(8):083509, April 2015.
- [72] C.L. Lee and X.-L. Zhou. Thick target neutron yields for the ${}^7\text{Li}(p,n){}^7\text{Be}$ reaction near threshold. *Nucl. Instrum. Methods B*, 152(1):1–11, 1999.
- [73] A. O. Hanson, R. F. Taschek, and J. H. Williams. Monoenergetic neutrons from charged particle reactions. *Rev. Mod. Phys.*, 21:635–650, Oct 1949.
- [74] Keiichi Shibata, Osamu Iwamoto, Tsuneo Nakagawa, Nobuyuki Iwamoto, Akira Ichihara, Satoshi Kunieda, Satoshi Chiba, Kazuyoshi Furutaka, Naohiko Otuka, Takaaki Ohsawa, et al. Jendl-4.0: a new library for nuclear science and engineering. *Journal of Nuclear Science and Technology*, 48(1):1–30, 2011.
- [75] CL Lee and X-L Zhou. Thick target neutron yields for the ${}^7\text{Li}(p,n){}^7\text{Be}$ reaction near threshold. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 152(1):1–11, 1999.
- [76] Rene Brun, L Urban, Federico Carminati, Simone Giani, M Maire, A McPherson, F Bruyant, and G Patrick. Geant: detector description and simulation tool. Technical report, CERN, 1993. We use version 4.10.05.p01, with the shielding physics list.
- [77] T. Reynolds. *A Measurement of the Ionization Yield at Sub-keV Recoil Energies in Silicon*, Ph.D. Thesis. PhD thesis, University of Florida, 2021.
- [78] CH Poppe, JD Anderson, JC Davis, SM Grimes, and C Wong. Cross sections for the ${}^7\text{Li}(p,n){}^7\text{Be}$ reaction between 4.2 and 26 meV. *Physical Review C*, 14(2):438, 1976.
- [79] P Cabanelas, J Cruz, Micaela Fonseca, A Henriques, F Lourenço, H Luís, Jorge Machado, J Pires Ribeiro, AM Sánchez-Benítez, P Teubig, et al. Cross sections for proton induced high energy γ -ray emission (pige) in reaction ${}^{19}\text{F}(p,\alpha\gamma){}^{16}\text{O}$ at incident proton energies between 1.5 and 4 meV. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 381:110–113, 2016.
- [80] D Bachiller Perea, P Corvisiero, D Jiménez Rey, V Joco, A Maira Vidal, A Muñoz Martín, and A Zucchiatti. Measurement of gamma-ray production cross sections in ${}^7\text{Li}$ and ${}^9\text{F}$ induced by protons from 810 to 3700 keV. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 406:161–166, 2017.

- [81] James F Ziegler, Matthias D Ziegler, and Jochen P Biersack. SRIM - The stopping and range of ions in matter (2010). *Nucl. Instrum. Methods B*, 268(11-12):1818–1823, June 2010.
- [82] C Awe, P. S Barbeau, J. I Collar, S Hedges, and L Li. Liquid scintillator response to proton recoils in the 10–100 keV range. *Phys. Rev. C*, 98(4):045802, 2018.
- [83] Allen Caldwell, Daniel Kollar, and Kevin Kröniger. Bat—the bayesian analysis toolkit. *Comput. Phys. Commun.*, 180(11):2197–2209, 2009.
- [84] A. N. Villano, M. Fritts, N. Mast, S. Brown, P. Cushman, K. Harris, and V. Mandic. First observation of isolated nuclear recoils following neutron capture for dark matter calibration. *Phys. Rev. D*, 105:083014, Apr 2022.
- [85] J. D. Lewin and P. F. Smith. Review of mathematics, numerical factors, and corrections for dark matter experiments based on elastic nuclear recoil. *Astropart. Phys.*, 6(1):87–112, 1996.
- [86] Y Sarkis, A Aguilar-Arevalo, and JC D’Olivo. Ionization efficiency for nuclear recoils in silicon from ~ 50 eV to 3 MeV. *arXiv preprint arXiv:2209.04503*, 2022.
- [87] M Matheny, A Roberts, A Srinivasan, and AN Villano. The effective fano factor of nuclear recoils for dark matter searches. *arXiv preprint arXiv:2206.13639*, 2022.
- [88] MF Albakry, I Alkhatib, DWP Amaral, T Aralis, T Aramaki, IJ Arnquist, I Atae Langroudy, E Azadbakht, S Banik, C Bathurst, et al. Investigating the sources of low-energy events in a supercdms-hvev detector. *Physical Review D*, 105(11):112006, 2022.
- [89] Gwyn P Williams. X-ray data booklet. *X-ray data booklet*, 2001.
- [90] Z Ahmed et al. Search for inelastic dark matter with the cdms ii experiment. *Phys. Rev. D*, 83(11):112002, 2011.
- [91] F. J. Kerr and Donald Lynden-Bell. Review of galactic constants. *Mon. Not. Roy. Astron. Soc.*, 221:1023, 1986.
- [92] M. C. Smith, G. R. Ruchti, A. Helmi, R. F. G. Wyse, J. P. Fulbright, K. C. Freeman, J. F. Navarro, G. M. Seabroke, M. Steinmetz, M. Williams, et al. The rAVE survey: constraining the local galactic escape speed. *Mon. Notices Royal Astron. Soc.*, 379(2):755, 2007.
- [93] Ralph Schönrich, James Binney, and Walter Dehnen. Local kinematics and the local standard of rest. *Mon. Notices Royal Astron. Soc.*, 403(4):1829, 2010.
- [94] R. Agnese et al. New results from the search for low-mass weakly interacting massive particles with the cdms low ionization threshold experiment. *Phys. Rev. Lett.*, 116:071301, Feb 2016.

- [95] A. Aguilar-Arevalo et al. Search for low-mass WIMPs in a 0.6 kg day exposure of the DAMIC experiment at SNOLAB. *Phys. Rev. D*, 94:082006, 2016.
- [96] Q. Arnaud et al. First results from the NEWS-G direct dark matter search experiment at the LSM. *Astropart. Phys.*, 97:54–62, 2018.
- [97] J. I. Collar. Search for a nonrelativistic component in the spectrum of cosmic rays at earth. *Phys. Rev. D*, 98:023005, Jul 2018.
- [98] J. H. Davis. Probing sub-gev mass strongly interacting dark matter with a low-threshold surface experiment. *Phys. Rev. Lett.*, 119:211302, 2017.
- [99] S. Yellin. Finding an upper limit in the presence of an unknown background. *Phys. Rev. D*, 66:032005, 2002.
- [100] S. Yellin. Extending the optimum interval method. *ArXiv:0709.2701*, 2007.
- [101] B. Kavanagh. Earth-scattering of super-heavy dark matter: updated constraints from detectors old and new. *Phys. Rev. D*, 97:123013, 12 2017.
- [102] Timon Emken and Chris Kouvaris. How blind are underground and surface detectors to strongly interacting dark matter? *Phys. Rev. D*, 97(11):115047, Jun 2018.
- [103] SuperCDMS Collaboration. Internal communication, 2021.
- [104] R. W. Ogburn. *A search for particle dark matter using cryogenic germanium and silicon detectors in the one-and two-tower runs of CDMS-II at Soudan*. PhD thesis, Stanford U., Phys. Dept., 2008.
- [105] F Scholze, H Henneken, P Kuschnerus, H Rabus, M Richter, and G Ulm. Determination of the electron–hole pair creation energy for semiconductors from the spectral responsivity of photodiodes. *Nucl. Instrum. Meth. A*, 439(2-3):208–215, 2000.
- [106] AN Trukhin, J Jansons, H-J Fitting, T Barfels, and B Schmidt. Cathodoluminescence decay kinetics in ge+, si+, o+ implanted sio2 layers. *J. Non Cryst. Solids*, 331(1-3):91–99, 2003.
- [107] AP Baraban, SN Samarin, VA Prokofiev, VA Dmitriev, AA Selivanov, and Y Petrov. Luminescence of sio2 layers on silicon at various types of excitation. *J. Lumin.*, 205:102–108, 2019.
- [108] Peizhi Du, Daniel Egana-Ugrinovic, Rouven Essig, and Mukul Sholapurkar. Sources of low-energy events in low-threshold dark-matter and neutrino detectors. *Phys. Rev. X*, 12:011009, Jan 2022.

- [109] I Alkhatib, DWP Amaral, T Aralis, T Aramaki, Isaac J Arnquist, I Atae Langroudy, Elham Azadbakht, Samir Banik, D Barker, Corey Bathurst, et al. Light dark matter search with a high-resolution athermal phonon detector operated above ground. *Physical review letters*, 127(6):061801, 2021.
- [110] A Fuss. Excess workshop: Descriptions of rising low-energy spectra. Technical report, Fermi National Accelerator Lab.(FNAL), Batavia, IL (United States); SLAC . . . , 2022.
- [111] Robin Anthony-Petersen, Andreas Biekert, Raymond Bunker, Clarence L Chang, Yen-Yung Chang, Luke Chaplinsky, Eleanor Fascione, Caleb W Fink, Maurice Garcia-Sciveres, Richard Germond, et al. A stress induced source of phonon bursts and quasiparticle poisoning. *arXiv preprint arXiv:2208.02790*, 2022.
- [112] Christopher Race. *The modelling of radiation damage in metals using Ehrenfest dynamics*. Springer Science & Business Media, 2011.
- [113] S. M. Griffin, Y Hochberg, K. Inzani, N Kurinsky, T Lin, and T. C. Yu. SiC Detectors for Sub-GeV Dark Matter, 2020.
- [114] S. B. Kaplan, C. C. Chi, D. N. Langenberg, J. J. Chang, S. Jafarey, and D. J. Scalapino. Quasiparticle and phonon lifetimes in superconductors. *Phys. Rev. B*, 14:4854, 1976.
- [115] T Guruswamy, D J Goldie, and S Withington. Nonequilibrium superconducting thin films with sub-gap and pair-breaking photon illumination. *Supercond. Sci. Technol.*, 28:054002, 2015.
- [116] J. J. Yen, B. Shank, B. A. Young, B. Cabrera, P. L. Brink, M. Cherry, J. M. Kreikebaum, R. Moffatt, P. Redl, A. Tomada, et al. Measurement of quasiparticle transport in aluminum films using tungsten transition-edge sensors. *Appl. Phys. Lett.*, 105(16):163504, 2014.
- [117] J. J. Yen, J. M. Kreikebaum, B. A. Young, B. Cabrera, R. Moffatt, P. Redl, B. Shank, P. L. Brink, M. Cherry, and A. Tomada. Quasiparticle Transport in Thick Aluminum Films Coupled to Tungsten Transition Edge Sensors. *J. Low Temp. Phys.*, 184:30, 2016.

APPENDIX A

QET dynamic range and efficiency**A.1. QET Dynamic Range**

To determine the dynamic range of a TES-based sensor, we want to calculate the ratio of saturation energy (energy required to drive the TES normal) to energy resolution. Given that the observable is TES current, we can calculate the ratio of the pulse height for an impulse of total energy equal to the energy resolution to the maximum current change from the bias point, which yields a dimensionless ratio useful for calculating current or energy quantities.

Consider an ideal voltage-biased TES, assuming the operating point is much greater than the shunt resistance in the TES bias loop [11]. We find that the saturation current scales as

$$(A.1) \quad I_{sat} \approx V_b \left[\frac{1}{R_0} - \frac{1}{R_n} \right] = \frac{V_b}{R_n} [a^{-1} - 1]$$

where the bias resistance $R_0 = aR_n$, R_n is the normal state resistance, and V_b is the TES bias voltage. The equilibrium bias condition tells us that Joule power and thermal conductance power from the TES to the crystal substrate will balance, which allows us to calculate bias voltage as

$$(A.2) \quad V_b = \sqrt{aR_n \Sigma \frac{v_{TES}}{\zeta_{TES}} T_c^n}$$

giving us an equation for saturation current:

$$(A.3) \quad I_{sat} \approx \sqrt{\frac{a}{R_n} \Sigma \frac{v_{TES}}{\zeta_{TES}} T_c^n} [a^{-1} - 1]$$

This clearly has similar scalings to the resolution.

In the small signal limit, we want to calculate the current amplitude for an injection pulse of energy equal to the energy resolution. We assume that the phonon pulse follows the simple exponential form

$$(A.4) \quad P(t) = \frac{\epsilon\sigma}{\tau_{ph}} e^{-t/\tau_{ph}}$$

and the Green's function response of the TES has the form [11]

$$(A.5) \quad \delta I(t) = \frac{\mathcal{L}}{1 + \beta} \frac{\Delta E}{V_b \tau} e^{-t/\tau_-}$$

where we have implicitly assumed that the rise time is much shorter than the fall-time of the TES (we assume we are operating in the limit of low inductance). If we write $\Delta E = P(t' - t)\Delta t'$, we can derive the QET response function by convolving the two pulses

$$(A.6) \quad \delta I(t) = \frac{\mathcal{L}}{1 + \beta} \frac{1}{V_b \tau} \epsilon\sigma \frac{1}{1 - \tau_{ph}/\tau_-} \left[e^{-t/\tau_-} - e^{-t/\tau_{ph}} \right]$$

This function has two limits: (1) when the TES response time is much larger than the phonon response time, the right-most term reduces to the TES Green's function, which has an amplitude given by the coefficient; (2) in the limit that the phonon response time is much larger than the TES response time, the amplitude is corrected by the fall time ratio. Solving precisely for maximum amplitude of the time-dependent part of this function, we find the formula for maximum amplitude

$$(A.7) \quad I_\sigma = \epsilon\sigma \frac{\mathcal{L}}{1 + \beta} \frac{1}{V_b \tau} \left(\frac{\tau_{ph}}{\tau_-} \right)^{\frac{-\tau_{ph}}{\tau_- - \tau_{ph}}}$$

where we can see that a long phonon fall-time reduces the maximum pulse height for the same TES response.

Finally, we can calculate dynamic range by taking the ratio of saturation current to pulse amplitude

$$(A.8) \quad DR = \frac{I_{sat}}{I_\sigma}$$

$$(A.9) \quad = \frac{1}{\epsilon\sigma} \frac{V_b^2}{R_n} [a^{-1} - 1] \frac{1 + \beta}{\mathcal{L}} \tau \left(\frac{\tau_{ph}}{\tau_-} \right)^{\frac{\tau_{ph}}{\tau_- - \tau_{ph}}}$$

$$(A.10) \quad = \frac{1}{\epsilon\sigma} \frac{\Sigma v_{TES} T_c^n}{\zeta_{TES}} [1 - a] \frac{1 + \beta}{\mathcal{L}} \tau \left(\frac{\tau_{ph}}{\tau_-} \right)^{\frac{\tau_{ph}}{\tau_- - \tau_{ph}}}$$

$$(A.11) \quad = \frac{1}{\epsilon\sigma} \frac{f_{sc} c_W v_{TES} T_c^2}{n \zeta_{TES}} [1 - a] \frac{1 + \beta}{\mathcal{L}} \left(\frac{\tau_{ph}}{\tau_-} \right)^{\frac{\tau_{ph}}{\tau_- - \tau_{ph}}}$$

where this last step follows from the T_c dependence of τ [**11**, **13**],

$$(A.12) \quad \tau = \frac{f_{sc} c_W}{n \Sigma} T_c^{2-n}.$$

Here, f_{sc} is the superconductivity enhancement to the specific heat and c_W is the normal state specific heat.

This last scaling, if we assume the TES transition shape is invariant with T_c and TES volume, shows us how to maximize dynamic range of a device without degrading resolution. If we fix device T_c and hold resolution constant by definition, we find that from Eq. 1.7, $\sigma \propto \frac{\sqrt{v_{TES}}}{\epsilon}$; so for a fixed resolution, $DR \propto \sqrt{v_{TES}}$. Scaling up total TES volume will only improve dynamic range, without degrading resolution, if we can also increase device efficiency as the square root of volume enhancement.

While Eq. A.11 is exact in the case that T_c and geometry dependence of the various device parameters are known, these scalings only hold in the specific limit that we can be reasonably

certain that TES response (\mathcal{L} and β) will not change with efficiency and volume scaling. In this paper, design changes were largely limited to the size and number of QETs and thus we could reasonably model efficiency and volume as independent of TES response, benchmarking TES constants to previous devices such as the QP.4 detector discussion in the text. We should note, however, that the fully general calculation should add back in considerations for TES rise time and more complex phonon response characteristics, and thus this scaling serves as more of a general design guide than a precise calculation.

A.2. QET Efficiency Modeling

A complete description of the energy efficiency model can be found in Section 3.4 of Ref. [13]. Here, we briefly summarize the key features of this model and discuss how further refining the measured efficiency of this device, and comparable designs, can inform this model.

The total energy efficiency for converting phonon energy into the TES (that is subsequently detected) can be split into four main components, as illustrated in the top panel of Fig. 2.1:

- (1) Phonon collection efficiency ϵ_{ph} , the probability that an initial phonon is absorbed by an Al fin;
- (2) Phonon to quasiparticle conversion efficiency ϵ_{qp} for a phonon absorbed in the fin;
- (3) QP collection efficiency ϵ_{coll} for QPs concentrated into the trapping regions;
- (4) Trapped QP to TES thermal energy conversion efficiency, ϵ_{trap} .

All efficiencies are applied on a per-phonon or per-quasiparticle basis and are assumed to be energy independent. The total efficiency used in the resolution calculation is thus $\epsilon = \epsilon_{ph}\epsilon_{qp}\epsilon_{coll}\epsilon_{trap}$. For a practical device, only the first and third efficiencies are readily tunable through design optimization; the conversion efficiencies are largely material-defined rather than geometric.

Overall phonon collection inefficiency can be further split into phonon losses in the bulk, at surfaces, and to non-instrumented absorption. For sufficiently pure crystals, bulk losses are negligible and surface effects dominate, as discussed in Ref. [113]. For the current device, phonon losses are further minimized with high surface coverage. This enhances the probability for phonon absorption before their energy drops below the Al band gap by down conversion on crystal surfaces. In addition, there is very little uninstrumented absorbing surface, as the bias rails of this device are integrated into the QET fins. Thus the dominant phonon losses are expected to be in the back-side grid and at the detector side-walls. These losses are further mitigated by (1) making the backside grid only 30 nm thick, compared to the 600 nm thick absorbing fins, allowing phonons to be reflected back into the substrate before breaking Cooper pairs, and (2) using a large aspect ratio device to minimize total side-wall area. Our model suggests that these design choices are consistent with $\epsilon_{ph} \gtrsim 95\%$, given that phonon losses are assumed to be fairly negligible in this limit (assuming that there are no bulk or sidewall phonon losses) and that the fraction of phonons absorbed in the backside grid scales linearly with thickness¹.

The largest fixed efficiency reduction comes from the limited efficiency of phonon to quasi-particle conversion in the Al fins, referred to as Kaplan down-conversion. Detailed studies of the energy dependence of this process can be found in e.g. Refs. [114, 115]. For typical phonon energies many times larger than the superconducting gap energy, this process is limited to an efficiency of around 50 – 60%. Close to the gap, the efficiency increases due to the reduced fraction of the energy which can be released as phonons. This means that a phonon sensor using a superconducting absorber is fundamentally limited by the mismatch between the phonon energy

¹We can model the backside absorption fraction, for equal surface coverage, as $f_{loss} = f_{front}/(f_{front} + f_{back}) \sim 5\%$ for the thicknesses used in this design. When the backside grid has a lower coverage than the QET pattern, as is the case with this detector, this should be an upper bound on total phonon loss.

distribution and superconducting gap. We take $\epsilon_{qp} \sim 50\%$ as an upper limit on the efficiency of our sensors.

A related efficiency is the down-conversion of QPs to phonons plus normal electrons in the TES. In principle, this efficiency can be as high as 100% if the phonon energy can be contained to the TES, but there are losses both in the transport regions between the trap and TES, and during the phonon emission process, to the TES. Experiments measuring efficiency difference between events absorbed in the fin, and directly by the TES, imply that this efficiency is roughly 62% for simple trap designs [13]. A dedicated study of the trap design used by more recent QETs has yet to be fully characterized, but is expected to be higher. If we take this efficiency as a bound, this implies that, for perfect QP and phonon collection, our devices can at best expect an efficiency of 30 – 40%, limited by the energy conversion efficiency of the phonon to QP to phonon process.

The final consideration, which can be highly optimized, is the QP collection. Past studies have shown that QPs in high-quality Al fins have diffusion lengths on the order of hundreds of microns, but that the collection fraction of QPs is a function of the Al fin length, fin thickness, and trap geometry [116, 117]. This can be understood as a quasi two-dimensional diffusion problem, in which the diffusion length is also a function of the film thickness (in the limit that the fins are much longer than this thickness), and the collection at the interface depends on the transmission probability through that interface as well as the probability of diffusing to the interface. A detailed discussion can be found in Ref. [13]; for fins shorter than 100 μm , all geometries can expect collection efficiency greater than 75%. As the TES gets longer, the Al area can be split into more QET fins with individual Al/W trap areas. The diffusion in the fins then becomes more 1-dimensional, which increases the effective diffusion length and leads to more efficient QP collection for a fixed fin length.

Bounding the QP collection efficiency thus implies that, for designs with quasi-1D QP diffusion and short fin lengths, we will expect efficiencies on the order of 20 – 30%. The quoted efficiency in this paper, a lower bound of 29%, suggests that these assumptions are realistic. A more precise measurement of the efficiency of multiple detectors with the same QET design, but different surface area scalings, will help better quantify the remaining uncertainty in the model.