

Techniques and challenges in low-mass dark matter searches

using CDMS style detectors

by

Richard Germond

under the supervision of

Wolfgang Rau

A thesis submitted to the Graduate Program in Physics, Engineering Physics, &
Astronomy in conformity with the requirements for the Degree of Doctor of Philosophy

Queen's University

Kingston, Ontario, Canada

February, 2023

Copyright © Richard Germond, 2023

Page intentionally left blank.

To my parents, Anne and Colin.

Abstract

Astronomical and cosmological observations suggest that roughly 85% of matter in the universe is in the form of non-luminous dark matter that interacts predominantly via gravity. The preferred explanation is that dark matter is comprised of one or more new particles, which may solve open questions in particle physics. Direct detection experiments look for interactions of galactic dark matter in sensitive detectors using a variety of technologies, including: cryogenic solid state detectors, time projection chambers with liquid noble elements, and bubble chambers with superheated fluids. The cryogenic operating temperatures of solid state detectors afford these devices excellent energy resolutions and low energy thresholds, providing sensitivity to low-mass dark matter particles. The Super Cryogenic Dark Matter Search (SuperCDMS) is a direct detection experiment that uses cryogenic semiconductor detectors instrumented with superconducting transition edge sensors. The first phase of SuperCDMS took place at the Soudan mine in Northern Minnesota, and preparations are being made for the next phase at SNOLAB in Sudbury, ON. Prior to the completion of SuperCDMS SNOLAB, the detectors will be tested in the neighboring experimental bay at SNOLAB in the Cryogenic Underground TEst (CUTE) facility. This thesis describes contributions to the design and construction of CUTE and testing data from SuperCDMS detectors in this facility. As well, results of a dark matter search using data from SuperCDMS Soudan, and results from a prototype detector operated at the University of Massachusetts Amherst will be presented. These data sets provide constraints on axion-like particles and dark photons with masses as low as $40 \text{ eV}/c^2$, and thermal relics with masses down to $40 \text{ MeV}/c^2$. When probing such low dark matter masses with cryogenic detectors, several experimental and analytical considerations must be taken into account, which will be detailed in this thesis.

Acknowledgements

As I approach the completion of my PhD, I reflect on the countless people who made it possible, while naming them all and discussing their impact could surely fill up many pages, but I list here some of the most influential people. First and foremost, this could not have been done without my parents, Anne and Colin, who supported me in all my endeavours, and continue to do so to this day. In addition to my parents, friends and family such as my partner, Mina Papic, my sister, Caitlin, and close friends like Ken Bonder motivated me to keep pushing forward, even when the path was unclear. The friends/colleagues in my local group, especially: Eleanor, Ryan, Emanuele, Jonathan, Aditi, Adam, and Muad made the work highly enjoyable, especially during the late nights and early mornings in the lab.

Beyond the support of my family and friends, the mentorship and guidance provided by my supervisor, Wolfgang Rau, was invaluable, in terms of both the technical material and in navigating a life in academia. Additionally, I benefited from the mentorship and guidance of many other people, including: Philippe Camus, Matt Pyle, Bruno Serfass, Ziqing Hong, Andy Kubik, Silvia Scorza, Serge Nagorny, Scott Hertel, and Scott Oser. On top of the support I received from members of SuperCDMS, many professors like Jordan Morelli, Aaron Vincent, and Joseph Bramante, influenced me greatly during my time at Queen's through the great deal of passion they brought to teaching.

The collective wisdom I received from the Queen's machinists, Chuck Hearn and Pat Given, was invaluable in completing the mechanical design work in this thesis. Finally, this work could not have been completed without the help of the Queen's, TRIUMF, and SNOLAB staff, whose contributions were of the utmost importance.

Statement of Originality

While the work described in this thesis is largely my own, it could not have been completed without contributions from a number of people. The focus of my experimental work was largely the CUTE facility in the context of the SuperCDMS experiment; my primary contributions were the development of the cold hardware necessary to mount a SuperCDMS SNOLAB detector tower and the slow-control system used to monitor the facility performance. The development of the CUTE facility's monitoring webpage was largely done by a team of students from BCIT, where my role was to guide the design and integrate the front-end webpage with the back-end servers. I also contributed to the CUTE run planning, and the data taking (and subsequent analysis) of the detectors operated in CUTE. Chapter 6 describes the analysis of a detector operated at the University of Massachusetts Amherst, with the purpose of setting new constraints on low-mass thermal relics. Several analyzers worked on this analysis of which my main contribution was the development of the analysis cuts, along with the calibration of the detector's energy scale. I also contributed to the efficiency calculation by providing ideas and suggestions for their implementation. Chapter 7 describes constraints on axion-like particle and dark photon dark matter, set using electron recoil data from SuperCDMS Soudan. This analysis was a group effort where my primary contribution was the CDM-Slite portion, specifically developing the limit setting technique and the determination of the efficiency in the below threshold regime.

Contents

Abstract	i
Acknowledgements	ii
Statement of Originality	iii
Contents	viii
List of Figures	ix
List of Tables	xii
List of Terms	xvi
1 Introduction	1
2 The case for particle dark matter	4
2.1 Observational evidence	4
2.1.1 Galactic rotation curves	5
2.1.2 Velocity dispersion of galaxy clusters	5
2.1.3 Gravitational lensing	6
2.1.4 Cluster dynamics	7
2.1.5 Cosmic microwave background radiation	8
2.2 Dark matter candidates	10
2.2.1 Astrophysical objects	11
2.2.1.1 MACHOs	11
2.2.1.2 Primordial black holes	12

CONTENTS

2.2.2	Particle candidates	12
2.2.2.1	Thermal relics	13
2.2.2.2	Axions	16
2.2.2.3	Dark photons	18
2.2.3	Modified gravity	19
3	The detection of dark matter	21
3.1	Production at accelerators	22
3.2	Indirect detection	25
3.3	Direct detection	26
3.3.1	Annual modulation searches	28
3.3.2	Liquid noble element detectors	30
3.3.3	Superheated bubble chambers	31
3.3.4	Cryogenic solid state detectors	32
3.3.5	Results	34
3.3.5.1	Low-energy excesses	35
4	The SuperCDMS experiment	37
4.1	A brief history of SuperCDMS	37
4.2	Detection principle	39
4.2.1	Semiconductor physics	40
4.2.2	Phonon physics	41
4.2.3	Athermal calorimeters	42
4.2.4	Transition edge sensors	43
4.2.5	Charge readout	47
4.3	SuperCDMS Soudan	48
4.3.1	Detectors	48
4.3.2	Shielding	51
4.3.3	Backgrounds	51
4.3.4	Calibration	53
4.4	SuperCDMS SNOLAB	54

CONTENTS

4.4.1	Detectors	55
4.4.2	Shielding	56
4.4.3	Backgrounds	57
4.4.4	Calibration	58
4.5	Other devices	59
4.5.1	High Voltage detector with eV scale resolution (HVeV)	59
4.5.2	Cryogenic PhotoDetector (CPD)	61
5	The cryogenic underground test facility	62
5.1	Motivation	62
5.2	Facility description	63
5.2.1	Cryostat	64
5.2.2	Suspension system	66
5.2.3	Shielding	68
5.2.4	Calibration systems	70
5.3	Detector implementation	71
5.3.1	Heat sinking	72
5.3.2	Infrared radiation shielding	76
5.3.3	Detector installation procedure	79
5.4	Facility monitoring	80
5.4.1	Facility control interface	83
5.4.2	Suspension system performance	85
5.4.3	Purge gas vibrations	90
5.5	Detector testing	92
5.5.1	Noise studies	92
5.5.2	Detector calibration results	94
5.5.2.1	Barium calibration	95
5.5.2.2	Sensor response	96
5.5.3	Background measurements	98
5.5.4	Low-energy excess	98

6	A search for low-mass thermal relics using a wide area photon detector	102
6.1	Description of the data	103
6.2	Data quality selection criteria	106
6.2.1	Livetime cuts	107
6.2.2	Baseline cuts	108
6.2.3	χ^2 cuts	110
6.2.4	Noise cuts	112
6.3	Calibration	116
6.4	Efficiency determination	119
6.4.1	Expected probability distribution	120
6.4.2	A comment on a new analysis	124
6.5	Results	127
6.6	Discussion	129
6.6.1	Low energy excess	130
7	Constraints on sub-keV dark matter from electron recoils in Super- CDMS Soudan	133
7.1	Description of the data	134
7.2	Energy resolution	135
7.3	Efficiency model	139
7.3.1	CDMSlite	139
7.3.2	iZIP	141
7.4	Rate calculation	144
7.4.1	Photoelectric cross section	145
7.5	Limit setting	147
7.5.1	Combination of data sets	148
7.5.2	Uncertainty band estimate	150
7.6	Results	150
7.7	Discussion	154
7.7.1	ALP constraints	154
7.7.2	Dark photon constraints	155

CONTENTS

8 Conclusion	157
Bibliography	160
A The optimal filter	177
A.1 The 1-template optimal filter	177
A.2 N-signal M-background optimal filter	179

List of Figures

2.1	Example galactic rotation curve	6
2.2	Composite image of the bullet cluster	8
2.3	Cosmic microwave background power spectrum	9
2.4	Thermal relic abundance of WIMPs through freeze out	14
3.1	Diagram representing the different strategies for detecting dark matter .	22
3.2	Diagram to explain the annual modulation of dark matter	29
3.3	Exclusion plot for spin-independent dark matter set by direct detection experiments	35
4.1	Circuit diagram of the SuperCDMS phonon circuit	46
4.2	Circuit diagram of the SuperCDMS charge circuit	48
4.3	Schematic of the SuperCDMS Soudan experiment	49
4.4	SuperCDMS Soudan iZIP detector and its sensor layout	50
4.5	Schematic of the SuperCDMS SNOLAB experiment	54
4.6	SuperCDMS SNOLAB HV and iZIP detectors, with their sensor layouts	55
4.7	Energy spectrum of a CDMS HVEV device	60
5.1	Schematic of the CUTE facility	64
5.2	Schematic of the CUTE gamma calibration system	71
5.3	Photograph of the 50 K heat sink	73
5.4	Photograph of the custom bottom of the 4 K can	73
5.5	Photograph of the 100 mK tower stage heat sink	74
5.6	Photograph of the copper membrane at the still stage	75
5.7	Photograph of the 4 K tower stage heat sink	76

LIST OF FIGURES

5.8	Photograph of the custom mixing chamber can	78
5.9	Screenshot of the CUTE facility monitoring webpage	84
5.10	Plot showing the detector's response to changes in air pressure.	86
5.11	Plot showing the detector's response to changes in air pressure.	87
5.12	Noise spectra in two detectors for a vibration study	89
5.13	Accelerometer measurements under different purge gas flow rates	91
5.14	Noise spectra under different purge gas flow rates	92
5.15	Noise spectra highlighting the effect of the Peltier cooler controller	93
5.16	Noise spectra highlighting the effect of Wi-Fi	94
5.17	Noise spectra highlighting the effect of Bluetooth	95
5.18	Energy spectra with the barium source at different positions	97
5.19	Energy spectrum of the background at CUTE	99
5.20	Low-energy rate of PD2 at CUTE over various cryogenic runs	100
6.1	Photograph of the second version of the CPD	103
6.2	Signal template used by the optimal filter	106
6.3	Periodogram of the noise in CPDv2 over time	107
6.4	Prepulse baseline current over time for CPDv2	109
6.5	Slope cut definition	110
6.6	Average of several traces with sine wave feature	112
6.7	Sine wave cut definition	113
6.8	Noise spectra through successive cut applications	115
6.9	Measurement of the optimal filter resolution	115
6.10	Energy spectrum from the integral energy estimator	118
6.11	Calibration of the optimal filter energy scale	119
6.12	Combined trigger and cut efficiency for CPDv2	121
6.13	Efficiency map determined from salted data	121
6.14	Analytical form of the efficiency map without the anti-coincidence trigger	125
6.15	Analytical form of the efficiency map with the anti-coincidence trigger . .	125
6.16	Recoil energy spectra for different dark matter masses	128
6.17	Exclusion plot for spin-independent, low-mass dark matter	129

LIST OF FIGURES

6.18	Low energy rate in CPDv2 at UMass in two different cryogenic runs . . .	131
7.1	Energy resolution model for the three CDMSlite data sets used in the dark photon/ALP analysis	137
7.2	Energy resolution model for an iZIP detector used in the dark photon/ALP analysis	138
7.3	Efficiency curves for the three CDMSlite data sets used in the dark pho- ton/ALP analysis	141
7.4	Efficiency curves for an iZIP detector used in the dark photon/ALP analysis	144
7.5	Photoelectric cross section of germanium	146
7.6	Upper limits on the rate determined from the dark photon/ALP analysis	151
7.7	Exclusion plot for the axioelectric coupling	152
7.8	Exclusion plot for the dark photon's kinetic mixing parameter	153

List of Tables

7.1	Exposures of the three CDMSlite data sets used in the dark photon/ALP analysis	135
7.2	Exposures of the four iZIP detectors used in the dark photon/ALP analysis	135
7.3	Fitted parameters for the energy resolution model used by the three CDM-Slite data sets	136
7.4	Fitted parameters for the energy resolution model used by the four iZIP detectors	137

List of Terms

ADMX Axion Dark Matter experiment. 17

ALP Axion-Like Particle. 17

ATLAS A Toroidal LHC Apparatus. 23

BBN Big Bang Nucleosynthesis. 12

BSM Beyond Standard Model. 21

CCD Charge Coupled Device. 34

CDMS Cryogenic Dark Matter Search. 37

CDMSlite CDMS low-ionization threshold experiment. 50

CE ν NS Coherent Elastic Neutrino-Nucleus Scattering. 32

CMB Cosmic Microwave Background. 8

CMS Compact Muon Solenoid. 23

COBE Cosmic Background Explorer. 9

CP Charge-Parity. 16

CPD Cryogenic PhotoDetector. vi, 60

CRESST Cryogenic Rare Event Search with Superconducting Thermometers. 33

CUTE Cryogenic Underground TEst facility. 2

- DAQ** Data Acquisition system. 47
- DCRC** Detector Control and Readout Card. 55
- DEAP** Dark Matter Experiment using Argon Pulseshape discrimination. 30
- DFSZ** Dine-Fischler-Srednicki-Zhitnitsky model. 153
- DUNE** Deep Underground Neutrino Experiment. 25
- ER** Electron Recoil. 39
- FET** Field-Effect Transistor. 47
- GERDA** Germanium Detector Array experiment. 152
- GHS** Gas Handling System. 63
- HDPE** High Density Polyethylene. 56
- HEMT** High Electron Mobility Transistor. 47
- HFC** Horizontal-Flexible Cable. 55
- HV** High Voltage. 39
- HVeV** High Voltage detector with eV scale resolution. 59
- IR** Infrared Radiation. 76
- iZIP** interleaved Z-sensitive Ionization and Phonon. 38
- KDE** Kernel Density Estimator. 117
- LHC** Large Hadron Collider. 23
- MACHO** Massive Astrophysical Compact Halo Object. 11
- MOND** Modified Newtonian Dynamics. 19

List of Terms

NR Nuclear Recoil. 39

NTD Neutron Transmutation Doped. 33

NTL Neganov-Trofimov-Luke. 33

OI Optimum Interval. 125

OVC Outer Vacuum Chamber. 65

PBH Primordial Black Hole. 12

PCB Printer Circuit Board. 72

PMT Photomultiplier Tube. 30

PSD Power Spectral Density. 88

PT Pulse Tube cooler. 65

PWM Pulse Width Modulation. 92

QCD Quantum Chromodynamics. 16

QET Quasiparticle Trap-Assisted Electrothermal Feedback TES. 44

RQ Reduced Quantity. 106

SiPM Silicon Photomultiplier. 30

SLAC Stanford Linear Accelerator Center. 79

SQUID Superconducting Quantum Interference Device. 45

SUF Stanford Underground Facility. 37

SUL Soudan Underground Laboratory. 38

SuperCDMS Super Cryogenic Dark Matter Search. 2

SUSY Supersymmetry. 15

List of Terms

- TES** Transition Edge Sensor. 2
- TeVS** Tensor Vector Scalar gravity. 20
- TPC** Time Projection Chamber. 30
- UED** Universal Extra Dimensions. 15
- VFC** Vertical-Flexible Cable. 55
- VIB** Vacuum Interface Board. 55
- WIMP** Weakly Interacting Massive Particle. 13
- WISPDMMX** WISP Dark Matter experiment. 19
- WMAP** Wilkinson Microwave Anisotropy Probe. 9

Chapter 1

Introduction

Many open questions exist in the fields of particle physics and cosmology. Observations on a number of different scales suggest that a large fraction of the matter in the universe is in the form of non-luminous, feebly-interacting, dark matter particles, beyond our current understanding of particle physics [1]. Moreover, there exist several problems within the standard model of particle physics, for instance: the non-zero mass of neutrinos [2], the anomalous magnetic moment of muons [3], and the yet unobserved electric dipole moment of the neutron [4]. Each of these issues may require new, beyond standard model (BSM) physics, generally in the form of new fields and particles. Some of the proposed solutions may also be viable dark matter candidates [4], spurring the physics community to develop search techniques that could detect signatures of BSM physics [5].

There are three general strategies used in the search for dark matter particle candidates [5]: production in accelerators, signatures from astrophysics, and directly measuring galactic dark matter in a terrestrial detector. The latter of these three approaches, known as direct detection, is the only type that is directly sensitive to galactic dark matter recoiling with a detector. Direct searches involve operating a highly-sensitive, well-shielded detector for extended periods of time to potentially measure a signal from a recoiling dark matter particle. A variety of different technologies are used by direct detection experiments, including: liquid noble element detectors [6, 7, 8], gaseous time projection chambers [9], superheated bubble chambers [10], and cryogenic solid-state detectors [11, 12], which each have their own unique advantages and challenges.

When compared to other direct detection technologies, cryogenic detectors can have excellent sensitivity to low-mass dark matter particles, due to their exceptional energy resolutions and low energy thresholds. The Super Cryogenic Dark Matter Search (SuperCDMS) is a direct detection experiment that operates cryogenic semiconductor detectors, instrumented with superconducting transition edge sensors (TES), at milliKelvin temperatures. The first phase of the SuperCDMS experiment took place in the Soudan mine in Northern Minnesota, and preparations are currently underway for the next phase of the experiment, which will be in SNOLAB near Sudbury, ON. Prior to the construction of SuperCDMS SNOLAB, the detectors will be tested in the Cryogenic Underground TESt facility (CUTE), located in SNOLAB adjacent to the site of SuperCDMS SNOLAB. Testing of the detectors in CUTE before the commissioning of SuperCDMS SNOLAB is complete offers a number of advantages, e.g. testing of the detectors and readout electronics can be performed in a low background environment while minimizing the exposure of the detectors to cosmogenic radiation. Furthermore, an early dark matter search may be performed with the detectors in CUTE well before the main experiment is online, allowing for processing and analysis techniques to be refined while also improving the current constraints on particle dark matter.

Despite the advantages of cryogenic detectors, a number of technical challenges on all aspects of the experiment must be addressed in order to perform a world-leading dark matter search. For one, in order to achieve the milliKelvin scale temperatures necessary to operate cryogenic detectors, adequate heat-sinking and infrared shielding is required; leakage of infrared photons through the shielding can produce low-energy events that resemble noise. Additionally, vibrations produced by the cryogenic equipment can introduce substantial noise into the detectors which must be mitigated so as to not limit the sensitivity of the experiment. On the detector side, stress induced microfractures in the crystal of the detector can produce low-energy backgrounds [13], which increase the low-energy rate of events. Finally, after the data are collected, special care must be taken in their analysis to set constraints on the interaction of dark matter particles that deposit energy near or below the trigger threshold.

I will begin by motivating the case for particle dark matter in Chapter 2, and in

Chapter 3 will discuss some of the detection techniques that are used. In Chapter 4 I will describe the detection principle utilized by SuperCDMS and the construction of SuperCDMS Soudan, and SuperCDMS SNOLAB. Chapter 5 will focus on the setup of the CUTE facility, showing how the configuration of the CUTE cryostat impacts detector operations and describing some methods used to mitigate these negative effects. Once the work related to the physical setups have been detailed, I will present two separate searches for low-mass dark matter, which study data from both nuclear and electron recoils, respectively in Chapters 6 and 7. These two searches are each sensitive to below threshold recoil energies, due to the finite, non-zero energy resolution of the detectors. After describing the methods to prepare the data and obtain the constraints on the relevant dark matter models, I will discuss the results in the context of comparable direct detection experiments and astrophysical constraints.

This thesis demonstrates that cryogenic solid state detectors are well suited for the discovery of low-mass dark matter particles, and describes future developments that could improve the technology. Advancements in this field will not only improve the viability of the direct detection of low-mass dark matter, but may also bring about improvements in other fields, such as quantum computing based on superconducting qubits [14, 15].

Chapter 2

The case for particle dark matter

Over the past century, an abundance of observational evidence on vastly different scales has pointed to a large fraction of the matter in the universe consisting of non-luminous “dark matter”. The current consensus is that the dark matter is a new particle, outside the current standard model of particle physics. This chapter will detail the observational evidence supporting the existence of dark matter, give examples of popular dark matter particle candidates, and describe some alternatives to particle dark matter, along with their successes and shortcomings.

2.1 Observational evidence

So far, all of the observational evidence for dark matter has come from its gravitational effects on different scales. These observations range from galactic scales, where deviations from the expected Keplerian motion are seen, to clusters of galaxies, and even to scales as large as the observed universe. While the nature of the effect that dark matter has varies depending on the scale, each observation shares a commonality, namely: there is more gravitational potential in the universe than can be ascribed to just the luminous matter. A potential common explanation for these observations is the existence of some form of non-luminous, non-baryonic matter, termed dark matter. In fact, the various observations generally agree that approximately 85% of the matter in the universe is in the form of dark matter [16], making it about five times as abundant as regular matter. This section will discuss the various pieces of evidence for dark matter, going from smaller,

astrophysical scales to larger, cosmological scales.

2.1.1 Galactic rotation curves

In the 1970s Vera Rubin studied the galactic rotation curves of several galaxies, through measurements of the 21 cm transition line in hydrogen and other emission lines, and was able to measure the velocity dispersion far beyond the galactic nucleus [17]. The velocity profile of these galaxies was found to remain flat well outside the central bulge of the galaxy, which suggests that the integrated mass of the galaxy increases radially outward beyond the center of the galaxy, where most of the luminous matter is located. An explanation for the observed velocity curves is the presence of a massive, dark halo, that surrounds the galaxy and extends well beyond the visible edge of the galaxy. Improved results were later obtained, confirming the initial observations of Rubin [18]. An example of one of these rotation curves along with the respective fits to models of the different components (stars, gas, dark halo) is shown in Figure 2.1. While the observations made by Vera Rubin sparked great interest within the astronomy community about the possibility of the presence of dark matter, this was not the first hint at its existence.

2.1.2 Velocity dispersion of galaxy clusters

The first evidence for dark matter came in the 1930s from a Swiss astronomer named Fritz Zwicky. Using redshift measurements of the Coma cluster performed by Edwin Hubble, Zwicky noticed that several nebulae exhibited large velocity dispersions, on the order of several thousand kilometers per second [19]. By applying the Virial theorem to these data, which relates the potential and kinetic energies of a system, Zwicky concluded that the actual mass of the cluster must be roughly 400 times more than what is expected from just the luminous matter. Zwicky proposed that the anomalously high velocity dispersions observed could be explained by a form of non-luminous, dark matter that exists in a much greater density than the luminous matter. While Zwicky's initial calculation remains in good agreement with modern day measurements of the Coma cluster and other similar clusters, it was not until several decades later that his proposed explanation of dark matter became well accepted. It should be noted at the time of Zwicky's work, he was unable to observe the hydrogen gas in the cluster, which greatly outweighs the stars.

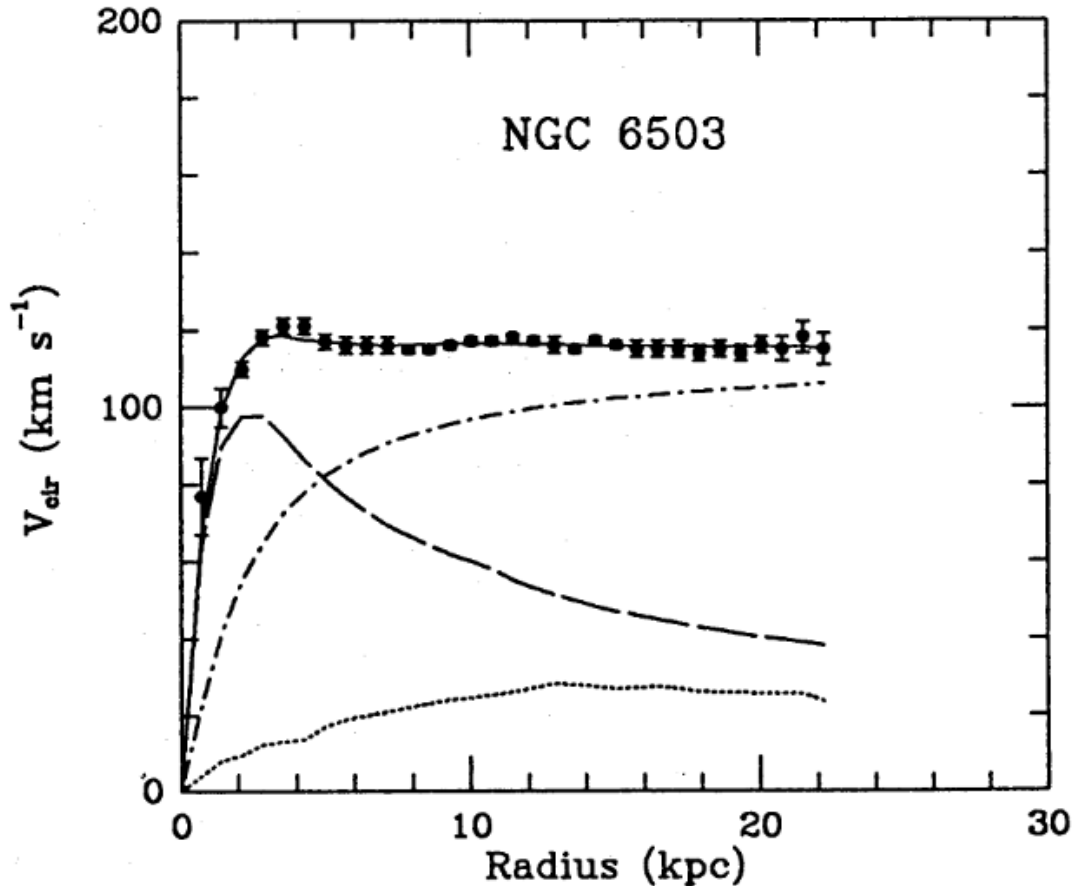


Figure 2.1: Measured rotation curve of the galaxy NGC 6503 (points with error bars), with the best fit value of the dark-halo fit shown in the solid black line. Also shown are the rotation curves of the individual components: the dashed curve for the visible components, the dotted curve for the gas, and the dotted-dashed curve is for the dark halo. The dark halo component is required in order to explain the flatness of the rotation curve far from the center of the galaxy. Image taken from Reference [18].

While this hydrogen gas is unable to explain the discrepancy, it does reduce the required amount of dark matter.

2.1.3 Gravitational lensing

Other measurements of the total mass of galaxies or galaxy clusters comes from gravitational lensing observations. The principle of gravitational lensing is that light from a distant background source (such as a galaxy) will be distorted along multiple paths by a large mass in the foreground (such as a galaxy cluster or a dark matter halo), and converge back to the observer. In this sense the gravitational lense can act as a means

to focus light from distant sources, sometimes producing multiple distorted images. The strength of the gravitational lense depends on the mass of the foreground object; as such, gravitational lensing effects vary in intensity, the most drastic example being strong lensing, while weak lensing or microlensing correspond to less distortion of the light. The precise amount of distortion is predicted by Einstein's theory of general relativity; in fact, a gravitational lensing event observed in 1919 provided the first experimental verification of Einstein's theory of general relativity [20]. However, it was Zwicky who proposed that gravitational lensing could be used to estimate the mass of galaxy clusters [21]. Thanks to the advancement of astronomical techniques, many lensing surveys have now been completed that show the distribution of dark matter on many different scales, which is in good agreement with simulations of large scale structure in the universe [22].

2.1.4 Cluster dynamics

Mergers of galaxy clusters provide another environment to test the dark matter hypothesis. The dominant component of baryonic matter in a galaxy cluster is hot gas (such as hydrogen), and the fraction of matter in the gas is much larger than the visible matter (such as stars). During a merger of galaxy clusters, the baryonic matter can be observed from X-ray imaging [23], while gravitational lensing can be used to independently track the gravitational potential of the two clusters. Since dark matter interacts very weakly, its motion will only be affected by gravitation, whereas baryonic matter will be slowed down through electromagnetic interactions. If the cluster contains a significant amount of dark matter, then the intergalactic gas in the two clusters will interact during a merger, while the dark matter will remain (relatively) collision free and continue travelling in its original direction. In this case the gravitational potential will be spatially separated from the visible matter, since the expectation is that there is much more dark matter than ordinary matter in a cluster. Conversely, if galaxy clusters are devoid of dark matter then the gravitational potential during a merger would spatially coincide with the baryonic matter. Observations of merging galaxy clusters [23], most famously the bullet cluster, show that the luminous matter is not spatially coincident with the observed gravitational potential, again supporting the hypothesis that most of the matter in the universe is com-

2.1. OBSERVATIONAL EVIDENCE

posed of dark matter. Figure 2.2 shows an optical image of the bullet cluster, overlaid by a depiction of the gravitational potential and the baryonic matter density, showing the disparity of the distributions.



Figure 2.2: Composite image of an optical observation of two neighbouring galaxy clusters, the gravitational potential determined by weak lensing (in blue), and the X-ray emission of hot gas (in pink). This famous observation [23], referred to as the bullet cluster due to the observed shock wave in the gas distribution, has provided compelling evidence for the existence of dark matter, since the hot gas and gravitational potentials do not spatially coincide. Image credit: NASA [24].

2.1.5 Cosmic microwave background radiation

Perhaps the most compelling piece of evidence supporting the existence of dark matter comes from observations of the cosmic microwave background (CMB) [16]. In the plasma of the early universe, photons were strongly coupled to charged particles (electrons and protons); at this time the mean free path of photons was small, so photons interacted before propagating any appreciable distance. As the universe expanded and cooled, neutral atoms eventually formed which allowed photons to freely propagate; these photons

form the CMB, the first light in the universe that can be observed today. Anisotropies in the angular power spectrum of the CMB have been measured by a number of satellite based experiments, including COBE [25], WMAP [26], and Planck [1]; these temperature fluctuations are believed to arise from baryonic oscillations in the plasma. In the framework of the standard cosmological model, Λ CDM, gravitational wells cause an attraction between baryonic matter and dark matter. The infalling baryons are eventually repelled out of the wells due to electromagnetic repulsion, producing an oscillatory behaviour. The modes of the baryonic oscillations are imprinted on the CMB anisotropies, which can be fit to a cosmological model like Λ CDM to provide information about the early universe, such as the energy density of dark and baryonic matter. The CMB anisotropy as function of angular scale, along with the best fit from Λ CDM and residuals is shown in Figure 2.3.

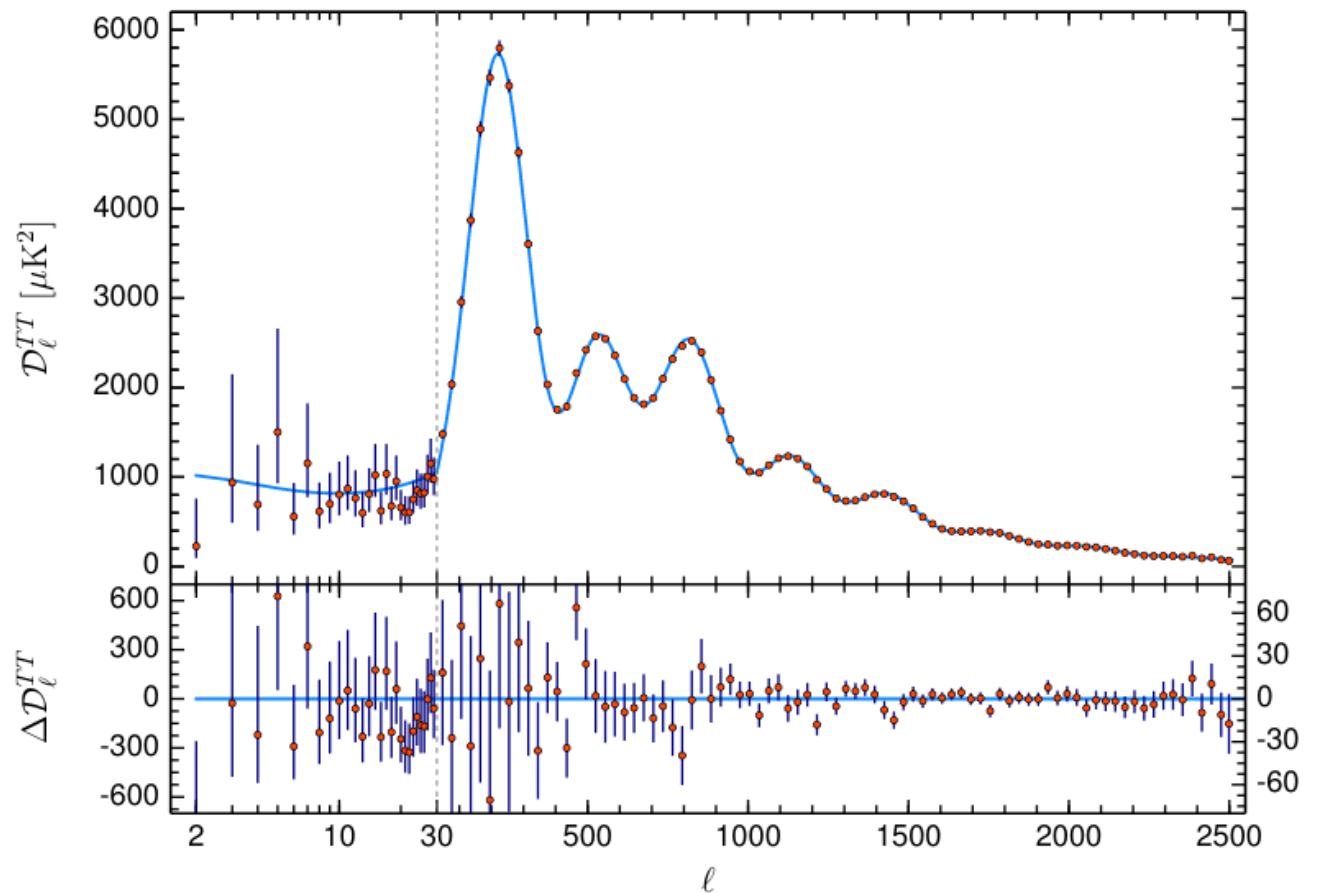


Figure 2.3: Power spectrum of the cosmic microwave background anisotropies measured by the Planck satellite, along with the best fit to the Λ CDM model and residuals. Image taken from Reference [1].

2.2. DARK MATTER CANDIDATES

The position of the first peak in the CMB power spectrum is sensitive to the curvature of the universe; in fact, the position of this peak indicates that the universe is very close to being spatially flat. The existence of the second peak confirms the existence of the oscillatory behaviour, and the ratio of the first peak to the second peak provides information on the baryon content of the universe. The overall height of each peak is affected by the energy density of the dark matter, however at least three peaks are required to properly constrain the amount of dark matter. The height of the third peak being comparable to the height of the second peak indicates that dark matter dominated the matter density of the plasma before recombination. If the dark matter density were lower, then the amplitudes of the oscillations would be suppressed, since the depth of the gravitational potentials would be shallower. In Λ CDM, the energy density of different components is parameterized as Ω_i , where i could be b , c , γ , or ν , corresponding respectively to the energy density of baryonic matter, cold dark matter, photons, and neutrinos. These parameters are generally quoted as $\Omega_i h^2$, which reflects the energy density of the given component per comoving volume (since the universe is expanding). Measurements from the Planck collaboration show that the energy density of baryons is $\Omega_b h^2 = 0.02233$ [1], which is in agreement with the prediction from Big Bang nucleosynthesis (BBN) [27], a theory that predicts the abundance of light elements in the early universe. The energy density of dark matter was measured by Planck to be $\Omega_c h^2 = 0.1198$ [1], which suggest that approximately 85% of the matter in the early universe is dark matter. So, the dark matter makes up roughly 25% of the total energy density of the universe while the baryonic matter makes up about 5%; the remaining 70% is attributed to dark energy, an unknown form of energy responsible for the accelerating expansion of the universe.

2.2 Dark matter candidates

In Section 2.1 some of the observational evidence for dark matter was discussed. Over the years many different hypotheses have been proposed to explain these observations. The proposed solutions to the dark matter problem span a broad range of mass scales, from new particles outside of the standard model to macroscopic astrophysical objects that

emit little to no light (such as brown dwarf stars or primordial black holes). Some theories attempt to explain the observations through alternative means, such as modifications to the behaviour of gravity on large scales. The following sections will describe and discuss some of these different hypotheses in the context of recent observations.

2.2.1 Astrophysical objects

Perhaps the simplest explanation of dark matter is that it is composed of massive, astrophysical objects, such as black holes or brown dwarf stars, i.e. massive objects that emit no light, and could account for many of the observations described above. Such explanations are simple because they require no modifications to the standard model of particle physics, and only minor modifications to astrophysical processes. These theories can be tested by gravitational micro-lensing surveys, where the massive objects would appear as shears in the gravitational potential of distance objects in the line of sight. In general, astrophysical objects that formed after the time of recombination are not able to form the entirety of the dark matter in the universe, since the shape of the CMB power spectrum would be drastically altered. It is however possible that a small fraction of the galactic dark matter is comprised of astrophysical objects.

2.2.1.1 MACHOs

One class of astrophysical objects that could make up some of the dark matter are massive astrophysical compact halo objects (MACHOs). Examples of MACHOs include: neutron stars, black holes, rogue planets, brown dwarfs and other dim stars. Constraints on the fraction of the dark matter in the Milky Way halo composed of MACHOs in the mass range $6 \cdot 10^{-8} M_{\odot}$ - $15 M_{\odot}$, where $M_{\odot}$ is the mass of the Sun, come from microlensing surveys performed by the EROS-2 collaboration [28]. Using a survey of bright stars in the Milky Way galaxy, the EROS-2 collaboration [28] determined that less than 8% of the halo mass of the Milky Way is composed of MACHOs. Constraints on MACHOs in the mass range of $10 M_{\odot}$ to several thousands of $M_{\odot}$ are set by requiring the timescale of growth for star clusters in ultra-faint dwarf galaxy be less than the estimated age of the dwarf galaxy [29]. The combination of these constraints, along with several other independent constraints [28, 29], rules out all of the dark matter being composed of

2.2. DARK MATTER CANDIDATES

MACHOs for masses above $10^{-7}M_{\odot}$.

2.2.1.2 Primordial black holes

One massive astrophysical object that could reasonably make up a large fraction of the dark matter in the universe are primordial black holes (PBHs), which are black holes that formed in the early universe before the time of Big Bang nucleosynthesis (BBN). A number of formation mechanisms for PBHs have been proposed, although these mechanisms are still widely debated. In general, a necessary condition for the formation of PBHs is the high density state of the early universe; however additional conditions are required [30]. A typical requirement is some sort of inhomogeneity, generally arising from quantum fluctuations; alternatively, a phase transition in the early universe could generate or enhance any inhomogeneities, thereby driving the production of PBHs [30]. The mass of PBHs depends on how early in the universe they formed; for instance PBHs that formed at the Planck time (10^{-43} s) would have masses on the order of the Planck mass (10^{-8} kg), whereas PBHs that formed at around 1 s would have masses on the order of roughly $10^5 M_{\odot}$ [30]. PBHs can span a wide range of masses, from 10^{-8} kg to several thousands of solar masses ($\approx 2 \cdot 10^{33}$ kg). However, PBHs with masses less than roughly 10^{11} kg are excluded from being the dark matter, as they would have decayed via Hawking radiation before the present day. While they could not compose all of the present day dark matter, PBHs with masses less than 10^{12} kg could have been present in the early universe. However, their evaporation would have affected the anisotropies in the CMB and the abundance of light elements determined from BBN [30], and so the mass range for PBHs from 10^7 - 10^{12} kg is also ruled out. PBHs in the mass range of 10^{17} - 10^{37} kg are ruled out as the dominant component of dark matter by lensing constraints, gravitational waves, cosmic structure formation, and dynamical constraints [30].

2.2.2 Particle candidates

Another natural explanation for dark matter is that it consists of a new, yet-unobserved particle. Despite the success of the standard model of particle physics in its predictive power, there still exist a number of open questions that remain to be addressed. Examples include the apparent fine-tuning of model parameters and the hierarchical nature of

the model. In most cases, extensions to the standard model have been proposed by introducing new symmetries or mechanisms that result in new particles, some of which could be the dark matter in the universe. These particles make excellent dark matter particle candidates as they are motivated independently of the dark matter problem, while their properties make them viable dark matter candidates in light of astronomical and cosmological observations. In this section, several popular dark matter particle candidates will be discussed.

2.2.2.1 Thermal relics

One of the most popular dark matter particle candidates are Weakly Interacting Massive Particles (WIMPs). It is assumed that dark matter particles self-annihilate into standard model fermions (such as quarks or leptons) in the process; standard model particles can also annihilate into dark matter particles, providing there is sufficient energy for the interaction. If the dark matter and standard model particles in the early universe are in thermal equilibrium, the rates of these interactions are balanced, and the number of dark matter particles remains constant. As the universe expands and cools, the available energy to support the production of dark matter eventually becomes too low, and only the process corresponding to the annihilation of WIMPs into standard model fermions remains. At some point the expansion rate of the universe exceeds the self-annihilation rate of dark matter and the abundance of dark matter remains constant, in a process called freeze-out [31]. As discussed in Reference [31], the dark matter abundance throughout the freeze-out process follows the differential equation:

$$\frac{dY}{dx} = -\frac{\lambda}{x^2} (Y^2 - Y_{EQ}^2), \quad (2.1)$$

where $Y \equiv n_\chi/T^3$ is a reparameterization of the dark matter's number density n_χ in terms of the temperature of the universe T , $x \equiv m_\chi/T$ is a term that is analogous to the dimensionless time the universe has had to evolve, $\lambda \equiv \frac{m_\chi^3 \langle \sigma v \rangle}{H(m_\chi)}$ is a reparameterization of the annihilation rate, and Y_{EQ} is the equilibrium density of the dark matter. Solving this differential equation yields the relic density of the dark matter as a function of time (more precisely inverse temperature which scales linearly with time). The result of this calculation is shown in Figure 2.4 for several values of λ , along with the equilibrium

2.2. DARK MATTER CANDIDATES

number density. It can be seen that as the annihilation rate (λ) increases, the point at which freeze-out occurs happens at lower temperatures (later times); before freeze-out, the dark matter is in equilibrium with the thermal bath and the number density tracks the equilibrium number density.

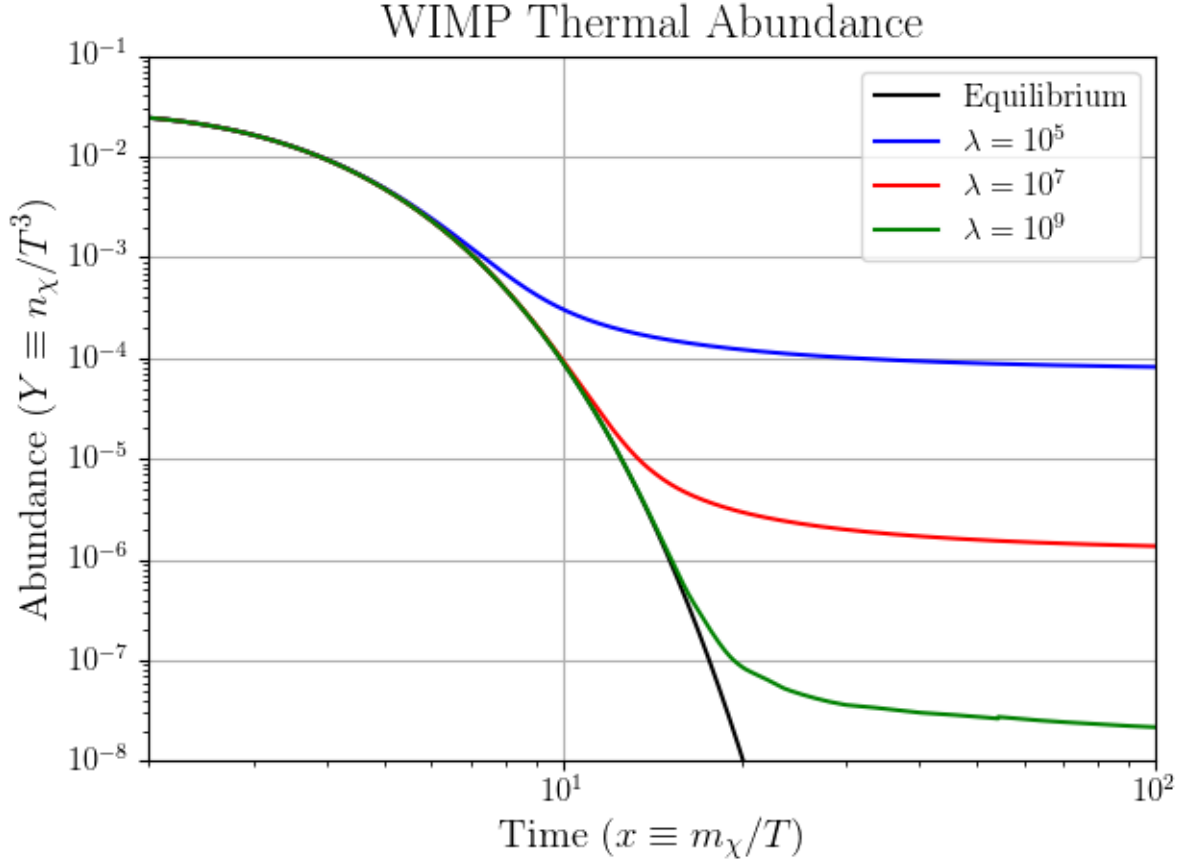


Figure 2.4: Schematic showing the thermal abundance Y of WIMPs produced through the freeze out mechanism, calculated using Equation 2.1 for different values of λ .

Particles produced through the freeze-out mechanism are generally classified as WIMPs. The present day abundance of WIMPs depends on their mass and annihilation cross section, if the cross section is high, the self-annihilation of WIMPs is likely and the present day abundance becomes low (since most WIMPs annihilate before freezing out). Conversely, if the cross section is small then annihilations are rare and freeze out occurs earlier, producing an over abundance of dark matter. The fact that for particles with masses in the range of $10 \text{ GeV}/c^2$ to a few $100 \text{ GeV}/c^2$ and a cross section on the same order as the weak force interaction yields the correct present day dark matter abundance

is referred to as the WIMP miracle.

Many extensions to the standard model include particles that align with this definition of WIMPs. In theories of supersymmetry (SUSY) the standard model fermions and bosons have supersymmetric partners [32]; the lightest, neutral, supersymmetric particle is a potential WIMP. In most SUSY models this is the neutralino [33], which is a mixture of the supersymmetric partners of the Higgs and the neutral electroweak gauge bosons. In theories that introduce universal extra dimensions (UED) the standard model fields may propagate in one or more compactified dimensions [34]. The introduction of extra dimensions allows for the existence of additional particles, the Kaluza-Klein partners of the standard model, which are the excitations of standard model particles in the extra dimension. The lightest Kaluza-Klein particles are stable, making them viable WIMP candidates [34].

Inspired by the WIMP miracle and the number of potential WIMP candidates in supersymmetric theories, a number of experiments were designed with the intent to discover WIMPs; these experiments include direct detection, indirect detection, and collider based production experiments, all of which will be discussed in more detail in Chapter 3. While WIMPs are certainly a promising dark matter candidate, so far no experiment has detected any definitive signatures from WIMP dark matter, or any hints of SUSY. Experiments are still testing viable parameter space for WIMPs in the mass range from 10-1000 GeV/ c^2 , but WIMP dark matter has largely fallen out of favor as the leading dark matter candidate as the WIMP miracle no longer applies to the yet untested parameter spaces. New particle models and production mechanisms have been proposed that would allow a lighter (sub-GeV mass) thermal relic to be produced copiously in the early universe. Such production mechanisms typically introduce a new mediator (for example a dark photon) in order to produce the correct relic abundance, but the overall production mechanism is largely the same, i.e. a particle that was previously in thermal equilibrium with the plasma of the early universe freezes out. These light thermal relics would generally interact in much the same way as the canonical WIMP, although they may have a small coupling to electrons. For the remainder of this thesis, any thermal relics produced in the early universe will be referred to as WIMPs.

2.2. DARK MATTER CANDIDATES

2.2.2.2 Axions

Quantum chromodynamics (QCD), the theory describing the strong force interactions between quarks and gluons, includes a charge-parity (CP) violating term θ that parameterizes the strength of the violation; experimental measurements of the neutron's electric dipole moment constrain θ to be less than 10^{-10} [35]. The strong CP problem is the apparent fine tuning of the θ parameter to be zero, when it could take on any value between 0 and 2π . One solution to the strong CP problem is to introduce a new global $U(1)$ symmetry that is spontaneously broken; the spontaneous symmetry breaking introduces a new massive pseudo Nambu-Goldstone boson called the axion [4]. Axions may be produced in the early universe through thermal and non-thermal mechanisms [36], and could account for the present day dark matter abundance. The expected mass range of axions is in the $10^{-6} - 10^{-4}$ eV/ c^2 range [27].

The CP violating term in the QCD Lagrangian is roughly given by:

$$\mathcal{L}_\theta \sim \theta G \tilde{G}, \quad (2.2)$$

where G is the gluon field strength tensor (and $\tilde{G}$ is its complement); this term is CP violating if θ takes on any non-zero value. As θ can take on any value from 0 to 2π , it would be (naively) expected to be of order unity. However, since observations constrain it to be very close to zero, a dynamical solution explaining the supposed zero value of θ is required. The Peccei-Quinn solution [4] to the strong CP problem is to essentially promote θ to a field, such that the relevant term in the Lagrangian becomes

$$\mathcal{L}_\theta = \left(\frac{a}{f_a} - \theta \right) \frac{\alpha_s}{8\pi} G^{\mu\nu\alpha} \tilde{G}_{\mu\nu\alpha}. \quad (2.3)$$

This minimal extension to the standard model introduces a new global $U(1)$ symmetry that is spontaneously broken at an energy scale f_a , and explicitly broken at the quantum scale by a QCD anomaly. The result of this mechanism is the introduction of a pseudo-Nambu-Goldstone boson, the QCD axion a , which dynamically drives the new effective angle $\theta_{eff} \equiv \left(\frac{\langle a \rangle}{f_a} - \theta \right)$ to zero, thereby solving the strong CP problem. The mass of the axion m_a can be determined through chiral perturbation theory [27], and is approximately equal to

$$m_a^2 \approx \frac{f_\pi^2 m_\pi^2}{f_a^2} \frac{m_u m_d}{(m_u + m_d)^2}, \quad (2.4)$$

where m_π is the pion mass, f_π is the pion decay constant, and m_u and m_d are the up and down quark masses, respectively.

The modified Lagrangian that includes the axion field generates a coupling between the axion and two gluons, however, this coupling is difficult to probe so experiments instead rely on the axion's coupling to two photons $g_{a\gamma\gamma}$, which arises in order to cancel the color anomaly. The relevant Lagrangian term for this process is:

$$\mathcal{L}_{a\gamma\gamma} = -\frac{g_{a\gamma\gamma}}{4}aF_{\mu\nu}\tilde{F}^{\mu\nu} = g_{a\gamma\gamma}a\vec{E} \cdot \vec{B}, \quad (2.5)$$

where $F_{\mu\nu}$ is the electromagnetic field strength tensor, and $\vec{E}$ and $\vec{B}$ are the electric and magnetic fields, respectively. Dedicated experiments known as axion haloscopes search for dark matter axions by probing this interaction. Axion haloscopes work by using a microwave cavity to resonantly amplify the electric fields generated by axions in a large, static magnetic field. The axion mass that the haloscope is sensitive to depends on the geometry of the microwave cavity and the amplitude of the magnetic field. Axion haloscopes, such as ADMX [37], are sensitive to axions with masses in the range of neV/c^2 to $\mu\text{eV}/c^2$. However, this range of axion masses is too low to be observed in typical direct detection experiments through a scattering type interaction, and so the rest of this thesis will not discuss the detection of dark matter (QCD) axions.

While axions are introduced as the solution to the strong CP problem [4], it is possible that other spontaneously broken global symmetries exist (unrelated to the strong CP problem) that introduce axion-like particles (ALPs). Unlike the QCD axion, ALPs are not required to solve the strong CP problem, and so the range of masses that ALPs can take on is much larger, spanning from roughly $10^{-20} \text{ eV}/c^2$ up to a few eV/c^2 [38]. ALPs arise from the spontaneous symmetry breaking of a global $U(1)$ symmetry, and often appear in higher energy theories such as string theory [27]. Somewhat recently, an ALPogenesis scenario has been proposed in which ALPs simultaneously explain the observed abundance of dark matter and the cosmological baryon asymmetry [39]. A generic feature of many ALP theories is a tree-level coupling to standard model leptons, which can be probed through a mechanism known as the axioelectric effect [40], an analogue to the photoelectric effect where an ALP is absorbed and its total energy is transferred to an electron. Direct detection experiments are generally sensitive to this process, and it will

2.2. DARK MATTER CANDIDATES

be discussed further in Chapter 7.

2.2.2.3 Dark photons

A useful test of the standard model electroweak interaction is the magnetic moment of the muon; somewhat recently, precision measurements of the magnetic moment of anti-muons were found to be above the standard model prediction [41]. The anomalous additional component of the muon's magnetic moment could be explained by the existence of a light vector boson -called the dark photon- with masses on the order of a few GeV/c^2 [42]. While the dark photon solution to the anomalous magnetic dipole moment of the muon was ruled out by the BaBar experiment [43], some argue that the dark photon can still simultaneously explain the anomalous magnetic moment of the muon while being the dominant component of the dark matter in the universe [42].

Dark photons are a natural extension to the standard model and could serve as a potential mediator between standard model and dark matter particles by coupling to the standard model hypercharge through a kinetic mixing with the standard model photon and Z boson [44]. The introduction of the dark photon charges dark matter under a new Abelian interaction [45]. The dark photon could mediate interactions between dark matter and the standard model, thereby changing the constraints on the mass and cross section of the dark matter required to thermally produce dark matter in the early universe. This means that sub-GeV dark matter could have been copiously produced in the early universe, something that would not be possible in the canonical WIMP freeze-out scenario. Depending on the mass of the dark photon it is possible that the new gauge boson itself is the dominant component of dark matter. In this case, dark photon dark matter could be produced through athermal mechanisms such as the misalignment mechanism [46]. In order for dark photons to be a viable dark matter candidate their mass must be less than twice the electron mass; otherwise the relic abundance would be depleted through decays into electron-positron pairs [46].

Experiments searching for dark matter dark photons are often axion haloscopes that have reinterpreted the limits in terms of the dark photon's kinetic mixing parameter [46]. There are however a few key differences between the detection of dark photons and axions that are important to consider. The first is that dark photon's have a specific

polarisation, and so unlike the axion case where the induced electric field is aligned with the applied magnetic field, the electric field induced by the dark photon can be in any direction. The second important consideration is that the coupling of the dark photon in the microwave cavity is not zero when there is no applied magnetic field, as is the case for axions. Several dedicated dark photon dark matter experiments have begun that are similar to axion haloscope experiments, but account for these important differences in terms of the shielding, noise vetoing, and polarisation detection. Some examples of these experiments are: WISPDMMX [47], and Dark E-field [48]. Similar to the case of dark matter axion detection, the mass range that these experiments are sensitive to is well below what can be detected by direct detection experiments. Instead, direct detection experiments look for the absorption of a dark photon in a process that ejects an electron with a well-defined energy [49], which will be discussed further in Chapter 7.

2.2.3 Modified gravity

While the existence of dark matter is widely accepted, alternate explanations have been proposed. Since all of the evidence for dark matter is based on its gravitational influence, a natural hypothesis is that something is wrong with the current laws of gravity. While it is well known that the theory of gravity is incomplete, in particular it lacks a proper theory for quantum gravity, Newton’s law of universal gravitation and Einstein’s theory of general relativity have each been highly successful in explaining observations on astronomical scales. Modifications to these laws have been proposed in order to account for the observed effects that are otherwise attributed to dark matter.

Modified Newtonian dynamics (MOND), originally proposed by Milgrom in the 1980s [50], provides a modification to Newton’s equation describing the force of gravity by changing its behaviour for low acceleration values. The modification to Newton’s force of gravity F_G is:

$$F_G = \frac{GMm}{\mu(a/a_0)r^2}. \quad (2.6)$$

Here $\mu(a/a_0)$ is a factor that modifies Newton’s law of gravitation depending on the acceleration experienced by the object, and $a_0 = 1.2 \cdot 10^{-10} \text{ m/s}^2$ is a new constant of nature, whose value was determined empirically [50]. If $a > a_0$ then $\mu \approx 1$, and Newton’s

2.2. DARK MATTER CANDIDATES

original formulation is returned. However if $a < a_0$ then $\mu \approx a/a_0$, which means that for very small accelerations the force of gravity experienced by an object is enhanced. The end result is that the dynamics of objects experiencing high accelerations (such as the planets in our solar system) behave normally, whereas objects in the low acceleration regime (in regions far from the galactic center) experience an increased velocity relative to standard Newtonian dynamics. This formulation of MOND ensures that the observed rotation curves of galaxies is explained, by design.

MOND is generally quite successful at explaining effects on galactic scales, such as rotation curves and planar structures orbiting the Milky Way and Andromeda galaxies [51]. In fact, Λ CDM has been shown to fail to describe certain galactic phenomena, such as the brightness of satellites in the Milky Way, whereas MOND has succeeded [52]. However, MOND also faces a number of challenges in reproducing observables on cosmological scales, such as the anisotropies in the CMB and the shape of the matter power spectrum [53], and gravitational lensing phenomena. Work in this field has progressed, and a relativistic extension to MOND, known as TeVeS, was able to explain the observed effects of gravitational lensing without the inclusion of dark matter [54]. Even more recently a group was able to reproduce the observed CMB anisotropies and matter power spectrum by proposing another relativistic extension to MOND [55].

The relative successes and failures of dark matter and modified gravity have spurred research into hybrid solutions. Proposals of superfluid dark matter have been made that effectively combine the successful aspects of dark matter and modified gravity [51]. Essentially, a new scalar field is added that behaves like a dark matter fluid on large scales, and modifies the standard behaviour of gravity on non-linear scales. The advantage of this description is that the standard cosmological observables, such as the CMB anisotropies and matter power spectrum, and smaller scale features that the standard Λ CDM model fails to accurately predict are both explained. Whether such a theory should be seen as the addition of superfluid dark matter or a modification of gravity is a debate left to the scholars, but it is clear that further thought should be given to such models, specifically to determine unique, observable features that can be tested.

Chapter 3

The detection of dark matter

The abundance of observational evidence that suggests the existence of dark matter, and the open questions in particle physics that require beyond standard model (BSM) physics, makes it only natural to devise experimental strategies to test these theories. Over the years, many different experiments to detect signatures of BSM physics have been performed, and certainly more will be developed in the coming years. This section will describe the three main classes of experiments used to detect particle dark matter, namely: direct detection, indirect detection, and accelerator searches. These three processes are represented diagrammatically in Figure 3.1, where each process corresponds to a different direction of the same fundamental interaction. In direct detection experiments, a dark matter particle interacts with a standard model target to produce a measurable signal, usually in the form of heat, ionization, scintillation light or some combination of these. Indirect detection experiments look for spectral features of standard model particles produced by the annihilation (or decay) of dark matter particles in places of expected enhanced dark matter density (core of the Sun, center of our galaxy, and other galaxies). Finally, accelerator searches accelerate standard model particles (usually protons or electrons) to highly relativistic speeds, collide them with other standard model particles, and look for signatures of missing energy or transverse momentum. Each of these techniques have relative advantages and disadvantages, and provide sensitivity to different ranges of dark matter masses, providing complementarity between the methods. Using all three of these methods allows for a broad range of dark matter models to be tested, which would

3.1. PRODUCTION AT ACCELERATORS

not be possible using only one of these techniques. Within a given search technique, such as direct detection, complementarity exists between different experiments, meaning some experiments have better sensitivity to certain types of dark matter particles than other experiments. After broadly discussing these three detection strategies, more information on direct detection experiments will be given, as this is the class of experiment relevant to this thesis.

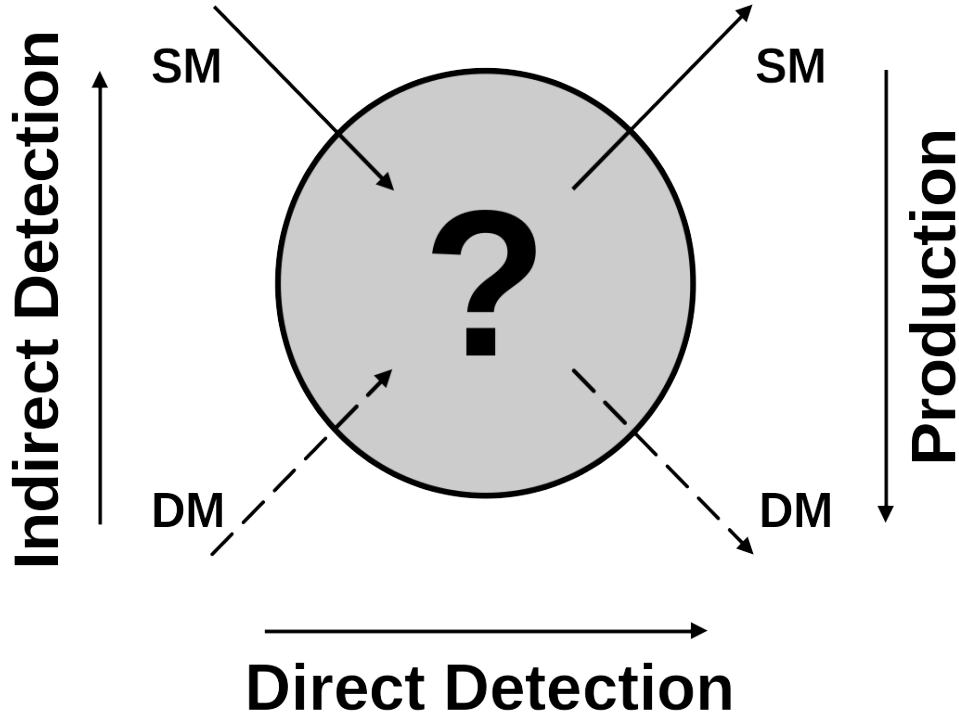


Figure 3.1: Representation of the general search strategies for particle dark matter. Direct detection experiments look for signatures of dark matter scattering with a standard model particle, corresponding to the left-to-right process in the diagram. Indirect detection experiments look for signals produced by dark matter annihilating into standard model products, corresponding to the process that goes from bottom-to-top. The process that goes from top-to-bottom represents searches that attempt to produce dark matter particles at high energy particle accelerators.

3.1 Production at accelerators

A popular technique in the search for BSM physics is to produce new particles at high-energy accelerators. In these searches, standard model particles, such as protons or

electrons, are accelerated to relativistic speeds and collided with other standard model particles. Such a collision could produce BSM particles, signatures of which can be searched for in detectors near the interaction site. Searches for BSM particles typically take place at large particle accelerators like the Large Hadron Collider (LHC) [56, 57], although lower energy accelerators have performed searches in beam dump style experiments [58]. Strictly speaking, these search techniques cannot determine if the BSM interaction by-products are stable, and therefore good dark matter candidates. However, the results of these searches are still valuable as they do not depend on the local density of dark matter. The main limiting factors of such experiments are the maximum center-of-mass energies that are achievable, and the integrated luminosity.

At the LHC, two rings of highly-relativistic protons are collided at fixed interaction points in the middle of large detectors at several sites around the facility. By colliding hadrons in this fashion the center-of-mass energy is twice the beam energy, allowing for heavy BSM particles to be produced. The basic idea is to study the collision by-products around the interaction site using several different detectors that have excellent tracking and calorimetry for charged particles. If a BSM particle, that could potentially be a component of dark matter, were produced in the interaction it would not interact in the subsequent detectors; instead the signature of interest would be missing transverse momentum in the final reaction by-products [59]. As the net momentum in the plane perpendicular to the beam of particles must be zero, any imbalance in the observed transverse momentum of the interaction by-products would mean that BSM particles were produced (and did not interact in the surrounding detectors). The two largest experiments at the LHC, ATLAS [56] and CMS [57], have both published a number of searches for WIMPs and dark photons, and while identifying such a particle would not necessarily be the dark matter of the universe, it would provide valuable information about potential dark sectors and inform other experiments where to search.

Searches for new physics have also been performed using collisions of leptons. The BaBar experiment collided electrons and positrons with a 10.6-11.2 GeV center-of-mass energy to produce B and anti- B pairs [60], and perform searches for new physics like dark photons [44]. The process of interest in this experiment is the annihilation of the

3.1. PRODUCTION AT ACCELERATORS

electron-positron (e^-/e^+) pair into a single photon (γ) and a dark photon (A'), which subsequently decays into invisible end products: $e^- + e^+ \rightarrow A' + \gamma$. The electron-positron pair that usually annihilates into two photons, instead produces only one photon as the other kinetically mixes into a dark photon. No such signature was seen by BaBar, and instead strong constraints were set on dark photons with masses between $200 \text{ MeV}/c^2$ and $10 \text{ GeV}/c^2$; these constraints exclude the values of the A' coupling required by the dark photon interpretation of the $(g-2)_\mu$ anomaly [44], although alternative decay topologies have reopened this possibility [42].

Another form of searches for BSM physics through accelerator production are beam dump experiments [58]. In a beam dump experiment, a high energy beam of standard model particles (either protons or electrons) impinge a fixed target (such as aluminum). In the collision, a number of high energy standard model particles are produced, mostly in the form of neutrinos, muons, or pions. In addition to these standard model particles, dark photons may also be produced, which could subsequently decay into other long lived dark-sector particles. These dark sector particles would be sufficiently long lived to travel from the target through additional shielding, up to a detector that would have some sensitivity to recoiling dark matter particles [58]. The advantage of beam dump experiments is the significant reduction in background, afforded by placing the detector far away from the interaction site. Most of the beam related backgrounds (pions and muons) are blocked from the actual detector, making the main backgrounds neutrino-induced events. Other, non-beam related, backgrounds can be excluded by requiring events in the downstream detector to be coincident with the beam timing. The disadvantage of beam dump experiments is that they require a large number of electrons or protons on the target, two interactions are required, but due to the low interaction cross section, the expected rate per beam particle is small [58]. While the downstream detector is sensitive to interactions of dark matter particles in the detector, beam dump experiments are distinct from direct detection experiments as they do not depend on the local density of dark matter, and are unable to show that the lifetime of the accelerator made BSM particle is sufficient to be the dark matter of the universe.

3.2 Indirect detection

Indirect searches look for the standard model products of annihilating (or decaying) dark matter particles in space. The annihilation rate of dark matter depends on its density, so indirect detection experiments observe regions where the density is elevated, such as at the center of galaxies or the Sun [61]. Indirect detection experiments use a number of different detector technologies, and look at many different channels, including: neutrinos, gamma rays, and positrons [61]. Indirect searches offer a number of advantages over accelerator production and direct detection searches as they provide sensitivity to higher energies and longer decay lengths than terrestrial dark matter searches [61]. There are however disadvantages to indirect searches, in particular backgrounds that are not well understood, and systematic uncertainties that can be difficult to quantify. Since the background contamination from galaxies can often drown out potential signals, dwarf galaxies are often used by indirect searches as a source that is abundant in dark matter but low in potential backgrounds [62].

Searches for signatures of dark matter annihilating into neutrinos come mostly from observations of the galactic center, where the local density of dark matter is greater. However, the limits on the annihilation cross section are generally much weaker than those set from the gamma or cosmic ray channels, due to the difficulty in detecting neutrinos. Currently, a number of Cherenkov detectors, such as IceCube [63], Super-K [64], and ANTARES [65], have set strong constraints on the annihilation of dark matter into neutrinos for a wide range of masses, although these constraints are still well above the expected cross section required for the thermal relic abundance [61]. These constraints are expected to be greatly improved in the next decade, with new experiments like DUNE [66] and Hyper-K [67] (the successor to Super-K) coming online soon.

In addition to neutrinos, X-rays and gamma rays are both used to set constraints on the decay and annihilation cross sections of dark matter. Gamma rays produced from the annihilation of dark matter in the galactic center, the halo of the Milky Way, and other extragalactic environments are studied [61]. Similar to neutrinos, X-rays and gamma rays will propagate directly from their source, without being scattered by astrophysical

3.3. DIRECT DETECTION

magnetic fields, unless they are first absorbed. The Fermi Gamma-Ray Telescope [68] has been collecting data for over 10 years, and provides leading sensitivity to GeV gamma ray sources thanks to its excellent angular and energy resolution. One interesting signature that may have been caused by decaying dark matter is a 3.5 keV X-ray line [61]. Such a line may be caused by sterile neutrino dark matter decaying to an active neutrino and a photon through the process: $\chi \rightarrow \nu + \gamma$, where χ is a 7 keV sterile neutrino. The validity of this signal is widely debated, and it has been attributed by some to systematic uncertainties in the analysis [61], although further investigation is required. Further constraints on this anomalous line will soon be obtained by a number of new X-ray telescopes, such as Micro-X [69, 70].

3.3 Direct detection

Direct detection experiments look for a signal produced by a dark matter particle interacting with a standard model target. A wide variety of target materials are used for this purpose, including cryogenic solid state detectors [71], liquid noble elements [6, 7], superheated fluids [10], and ancient minerals [72]. Depending on the target material, different signal channels such as heat, ionization, or scintillation light are used. In some cases, multiple signal channels are combined to provide discrimination between different particles; for example, the ratio of heat to ionization signals can be used to discriminate between interactions with neutrons and electrons [71]. The main challenges in direct detection experiments are reducing unwanted backgrounds [11]. To reduce the background from cosmic rays, most direct detection experiments are located deep underground. Additional shielding is generally required to further reduce the environmental backgrounds, caused by radioactive decays in the materials surrounding the experiment. Low activity materials are selected for the shielding and construction materials, and clean fabrication procedures are developed, but backgrounds can still be introduced from dust and radon in the air. Cosmogenic backgrounds, in the detector and surrounding materials, are mitigated by limiting the exposure to cosmic rays.

The general search approach of direct detection experiments is to count particle interactions in a detector, and use this measured rate to claim evidence of a particular dark

matter particle model (or exclude that model in the absence of a detection). In practice this is done by assuming that dark matter will interact in the detector through some mechanism, with a certain event rate that can be calculated a priori [73]. The number of events N seen in a direct detection experiment is related to the number density of dark matter, the dark matter particle's velocity v , the number of target particles' of a certain mass m_T (typically nuclei or electrons), the duration of the experiment, and the interaction cross section σ . Many direct detection experiments are sensitive to the energy deposited by the interacting particle, and can make use of this information by comparing the measured spectrum to the expected shape of the dark matter's recoil spectrum. Defining the exposure of the experiment ϵ as the product of target mass and livetime, and knowing the local density of dark matter to be $\rho_\chi = (0.3 - 0.4) \text{ GeV}/(c^2 \text{ cm}^3)$ [74], the differential number of expected dark matter interactions $\frac{dN}{dE_R}$ is [73]:

$$\frac{dN}{dE_R} = \epsilon \frac{\rho_\chi}{m_\chi m_T} \int v f(v) \frac{d\sigma}{dE_R} dv. \quad (3.1)$$

Here $f(v)$ is the velocity distribution of the dark matter, which is typically assumed to follow a Maxwellian distribution [73, 75].

Equation 3.1 can be used to compare the measured energy spectrum of an experiment to different dark matter models, and constrain the interaction cross section. Combining the knowledge of the expected background and signal spectra, the presence or absence of a signal can be tested for in a spectrum.

There are a number of interaction mechanisms that can be considered [75], depending on the underlying mechanism, spin-dependent and spin-independent effects may have to be considered separately. For a purely spin-independent interaction the differential cross section is given as [73]:

$$\frac{d\sigma}{dE_R} = \frac{m_N}{2\mu^2 v^2} \sigma_0 F^2, \quad (3.2)$$

where σ_0 is the spin-independent scattering cross section with each nucleon in the target, and $\mu \equiv \frac{m_\chi m_N}{m_\chi + m_N}$ is the reduced mass of the system. For typical momentum transfers in searches involving scattering with nuclei, the dark matter particle scatters coherently with the entire nucleus. As such, the scattering cross section receives a correction to account for the shape and size of the nucleus [73]; these corrections are accounted for by the form

3.3. DIRECT DETECTION

factor F . The study of form factors is an active field of research that depends heavily on complex nuclear physics. One example of a form factor used for spin-independent interactions is the Helm form factor [73], which is given by:

$$F(q) = \left(\frac{3j_1(qR_1)}{qR_1} \right)^2 e^{-q^2 s^2/2}, \quad (3.3)$$

where j_1 is a spherical Bessel function, q is the momentum transferred to the nucleus, s is the nuclear skin-thickness (roughly 1 fm), and $R_1 \approx 1.25 \text{ fm} \times A^{1/3}$ is the effective nuclear radius, where A is the atomic mass number.

As this thesis focuses on a direct detection experiment that uses cryogenic semiconductor detectors, further details will be given on the main competing experiments in this field. While these experiments use vastly different detector technologies, the general principle is the same: a well-shielded detector counts the number of particle interactions (ideally with additional information), and uses the measured rate to constrain the interaction strength of dark matter. Below, a description of some technologies and techniques used in this field will be given, although it is by no means a complete list.

3.3.1 Annual modulation searches

As the Earth orbits the Sun, the expected flux of dark matter seen by a direct detection experiment modulates annually due to the relative motion of the Earth with respect to the galactic dark matter halo [76]. When the Earth travels in the direction of the incoming dark matter, the flux is maximized; this maximum is expected to occur in June, while the flux is at a minimum in December [76]. Figure 3.2 shows a schematic representation of this effect; during June the Earth travels in the direction of the dark matter wind, so the dark matter flux seen by a terrestrial experiment is enhanced. The observation of such an annual modulation can be a useful signal if the rate of events is measured over the course of a year (or more), as most backgrounds are not expected to fluctuate in this way. Over the past 10 years the DAMA/LIBRA experiment [77], which uses 250 kg of NaI(Tl) scintillator, has shown with high statistical significance the presence of such a modulation. The peak of the phase informs about the mass of the dark matter particle, and the DAMA result [77] implies an 80 GeV (10 GeV) WIMP elastically scattering with predominantly iodine (sodium). Several explanations for this result have been proposed,

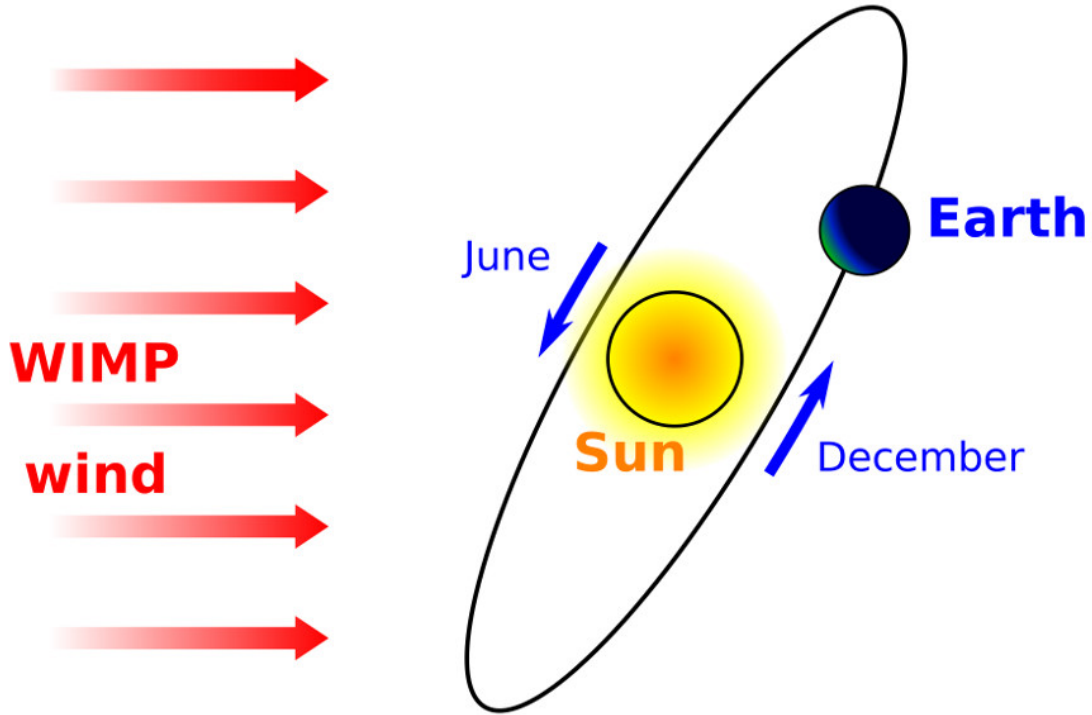


Figure 3.2: Schematic representation of the effect that causes an annual increase/decrease in the flux of dark matter seen by terrestrial dark matter experiments. In June, the flux is maximized as the Earth is travelling in the direction of the dark matter wind, while in December the flux is minimized. Image taken from Reference [76].

such as a modulation in the rate of cosmogenic muons caused by the seasonal expansion of the troposphere [78], or neutrons induced by muons and neutrinos [79]. The DAMA result is highly contested as several other experiments have not seen the modulation signal, and the parameter space of the implied dark matter particle explanation has been excluded by many experiments. As such, dedicated efforts to repeat the DAMA result have been made; the COSINE-100 [80] experiment has reported that with 3 years of data they observe a signal compatible with both the modulation and no modulation cases, spurring the need for more data. It has recently been proposed [81] that the way in which DAMA calculates the residual rate in the presence of a decreasing background, from ^{210}Pb or tritium, may induce a measurable modulation effect. This effect has been demonstrated by COSINE-100 [81], suggesting that the positive signal of an annual modulation may in fact be an artifact from the analysis.

3.3.2 Liquid noble element detectors

In the search for medium to high mass WIMPs, detectors based on liquid noble elements are currently the most competitive available, due to their large sizes, low backgrounds, and excellent energy resolutions. In such a detector, a noble element (typically Ar or Xe) is liquefied into a cryostat instrumented with the necessary sensors to read out the signals. When a particle interaction occurs in xenon or argon, an amount of scintillation light proportional to the interaction energy is produced and collected by photomultiplier tubes (PMTs) or silicon photomultipliers (SiPMs). Xenon based experiments, such as LZ [6] or XENON [7, 82, 83], generally have lower backgrounds than argon based experiments, such as DEAP-3600 [84, 8] and DarkSide [85, 86], but are also more costly, and limited in size by the amount of xenon available.

Xenon based experiments [6, 7] often make use of a dual-phase technology, where a gaseous phase of xenon sits atop a liquid phase. A particle interaction in xenon will produce prompt 178 nm scintillation light, and ionization (free electrons). By applying an electric field across the xenon, electrons liberated by the primary interaction are drifted through the liquid into the gaseous phase (without recombining), such a device is known as a time projection chamber (TPC). An avalanche process occurs when the electrons reach the liquid surface, and scintillation light proportional to the number of electrons and the strength of the electric field is produced. This technique provides two observable signals: the prompt scintillation light, referred to as the S1 signal, and the secondary scintillation produced by the electrons, referred to as the S2 signal. The ratio between the S1 and S2 signals provide discrimination between electron and nuclear recoils, since an electron recoil will produce a greater number of electrons than a nuclear recoil. By using the S2 signal [83], events corresponding to individual electrons have been observed, thereby greatly reducing the energy threshold. The interaction site can be reconstructed by using the timing information of the S1 and S2 signals, since the drift velocity of electron's in xenon is constant. Improved limits can be set by only considering particle interactions that occurred in the central region of the detector, which is shielded by the surrounding xenon.

Liquid argon based experiments use PMTs or SiPMs to measure the scintillation light

produced by recoiling particles [84, 85]. A powerful feature of liquid argon detectors is their ability to use pulse-shape discrimination to distinguish electron and nuclear recoils. The triplet (singlet) state of liquid argon is preferentially excited by electron (nuclear) recoils; the different life times of the triplet and singlet states can be used to distinguish between the initial recoil type. Liquid argon experiments are an excellent technology for dark matter detection due to the large amount of high purity argon that can be obtained, along with its high scintillation yield. Experiments such as Darkside [85, 86] use a dual phase technology, which allows for the rejection of surface backgrounds while maintaining the other advantages of a single phase technology.

3.3.3 Superheated bubble chambers

Another excellent technology for the detection of dark matter are superheated bubble chambers [10]. In a bubble chamber, a fluid is put in the superheated state through careful regulation of the temperature and pressure. When a particle deposits energy in the fluid, it changes the phase from liquid to gas, resulting in the nucleation of a bubble. To form a bubble, enough energy must be deposited within some critical volume. Multiple scatters occur during an electron recoil event, so the energy deposition is not well localized. The threshold of a bubble chamber can be controlled by changing the pressure, and can be set so electron recoil events (a background for WIMP searches) will not trigger the detector. The nuclear recoil background is dominant above about 1 keV, while the electron recoil dominates below that.

Bubble chambers are typically instrumented with pressure sensors, cameras, and piezoelectric sensors [87]. To remove events caused by multiply-scattering neutrons, events showing multiple bubbles on the cameras are excluded from the WIMP search data. Discrimination between neutrons and alpha particles is achieved by using the acoustic measurements from the piezoelectric sensors. A useful feature of bubble chambers is the ability to change out the target material; the PICO experiment [87] uses C_3F_8 as a target, while its predecessor COUPP [88] has used CF_3I . Atoms such as fluorine and iodine provide excellent sensitivity to spin-dependent dark matter models, and the ability to constrain both spin-independent and spin-dependent cross sections with various

3.3. DIRECT DETECTION

targets can greatly constrain supersymmetric models of dark matter [89].

A new development in bubble chamber technology is the scintillating bubble chamber [90], which uses superheated liquid xenon as a target material. The advantage of this approach is that it combines the excellent background discrimination of a bubble chamber, with the scintillation based energy reconstruction of noble liquid detectors. The operation of scintillating bubble chambers is similar to a standard bubble chamber, except the temperature required to superheat the xenon is lower than the typical target materials of C_3F_8 and CF_3I . The other main difference is the use of PMTs to measure the scintillation light. A primary advantage of the scintillating bubble chamber is the lower threshold for bubble formation, which is possible through the dissipation of energy into scintillation light. While scintillating bubble chambers are currently under development, they could prove a promising and scalable technology for future dark matter and $\text{CE}\nu\text{NS}$ experiments [90].

3.3.4 Cryogenic solid state detectors

There are a variety of dark matter detectors based on cryogenic solid state materials [71, 91, 12]. In general, when a particle interacts in a solid it produces electrons, phonons (quanta of lattice vibrations), and scintillation light -depending on the target. One or more of these channels can be read-out by instrumenting the detector appropriately; multiple channels provide event-by-event discrimination of electron and nuclear recoils [71, 91, 12]. Thermal fluctuations are suppressed by the cryogenic temperatures, allowing the detectors to have excellent energy resolutions and low thresholds. Furthermore, the average energy to produce an ionization event in a semiconductor is on the order of a few eV, compared to the typical tens of eV in liquid noble or gas based detectors. In general, this makes cryogenic detectors better suited for low threshold measurements, i.e. better sensitivity to lower dark matter masses.

When a particle interaction occurs in a semiconductor (such as silicon or germanium), ionization (electron-hole) pairs and phonons are produced, however if there is no electric field in the detector, the electron-hole pairs will recombine, releasing phonons in the process. While the total phonon energy would provide information about the energy

deposited in the crystal, it provides no discrimination as to whether the interaction was a nuclear or electron recoil. Experiments like SuperCDMS [71] and EDELWEISS [91] distinguish between nuclear and electron recoils by applying an electric potential across the detector. Instead of recombining, the electrons and holes produced by the interaction are drifted to the positive and negative electrodes, respectively. The phonon signal can be measured using cryogenic sensors such as transition edge sensors (TES) or neutron-transmutation-doped (NTD) thermal sensors, while the charge signal in the electrodes can be amplified and read-out using a suitable charge amplifier. The ratio between the charge and phonon signals provides discrimination between electron and nuclear recoils [71, 91]. As the electron-hole pairs are drifted across the detector, they produce additional phonons proportional to the potential difference; this process is known as Neganov-Trofimov-Luke (NTL) amplification [92], and will be discussed in further detail in the Chapter 4. If a large enough voltage is applied across the detector the initial recoil phonons will be hidden in the NTL phonons, losing the discrimination between electron and nuclear recoils. In this case a small ionization signal can be amplified into a large phonon signal, this is the basis for several SuperCDMS detectors [92, 11, 93], and the ability to detect single electron-hole pairs has been demonstrated [93]. Further details will be given in Chapter 4. While the list of cryogenic solid state detector technologies described below is not comprehensive, it gives a picture of the general techniques used in the field.

Instead of a semiconductor target, the CRESST experiment uses CaWO_4 [12], a scintillating material. When a particle interaction occurs in the scintillating target, it generates phonons and ionization, along with scintillation light. Thermal sensors, such as TESs, measure the energy deposited in the crystal through the phonon signal; note that the ionization pairs will recombine, depositing their energy into the phonon system. A light detector above the scintillator, composed of a TES instrumented on a thin silicon-on-sapphire wafer, measures the scintillation light [12]. Discrimination between electron and nuclear recoils is achieved using the light yield from the scintillation signal [12]. The separation of nuclear and electron recoils using the scintillation light is excellent for events with energy greater than 2 keV, but falls off below this. At high energies, the light yield can even distinguish between nuclear recoils off of the oxygen and tungsten atoms. The

3.3. DIRECT DETECTION

low mass of the oxygen atoms provides excellent sensitivity to low-mass dark matter [12]. One of the main challenges faced by CRESST are low energy excesses [94, 95], and radiogenic backgrounds intrinsic to the crystal [96]. In order to mitigate these backgrounds, special care must be taken in the production of the crystal.

Charge-coupled devices (CCDs) have also been used as detectors for a dark matter search. The DAMIC experiment [97] uses silicon CCDs consisting of millions of pixels only a few microns in size. Particle interactions in the silicon generate electron-hole pairs; the ionization pairs are collected at the surfaces of the CCD and read out by moving the charges pixel-to-pixel to the read-out amplifier. Noise introduced by the actual read out is greater than noise generated by thermal fluctuations, so the CCD is generally read out once every 8 hours. A new development in CCD technology is the skipper-CCD, which is similar in function to the normal CCD except it repeats the read out of each pixel many times in a non-destructive fashion to minimize the read-out noise. The SENSEI experiment [98] has performed dark matter searches using skipper-CCDs, and has set very strong constraints on several low-mass dark matter models [98], thanks to its ability to distinguish single electron-hole pair events.

3.3.5 Results

The technologies discussed above constrain a variety of different dark matter models, by using data from both electron and nuclear recoils. A number of assumptions go into the actual calculation, such as the dark matter’s local density and velocity distribution [73], but the sensitivities of different experiments may still be compared. Shown in Figure 3.3 is the exclusion plot for spin-independent WIMPs, set by a variety of direct detection experiments [86, 8, 99, 82, 83, 87, 80, 92, 12, 100, 11]; the light grey region in Figure 3.3 shows the WIMP masses and cross sections that have been excluded by direct detection experiments. In general, experiments that use liquid noble elements (xenon or argon) as a target have better sensitivity (can exclude smaller cross sections) to larger WIMP masses, whereas experiments that use lighter targets are more sensitive to lighter WIMP masses. This is due to the kinematic matching between the dark matter and target nucleus, as lighter dark matter transfers more momentum to a lighter target than a heavier target.

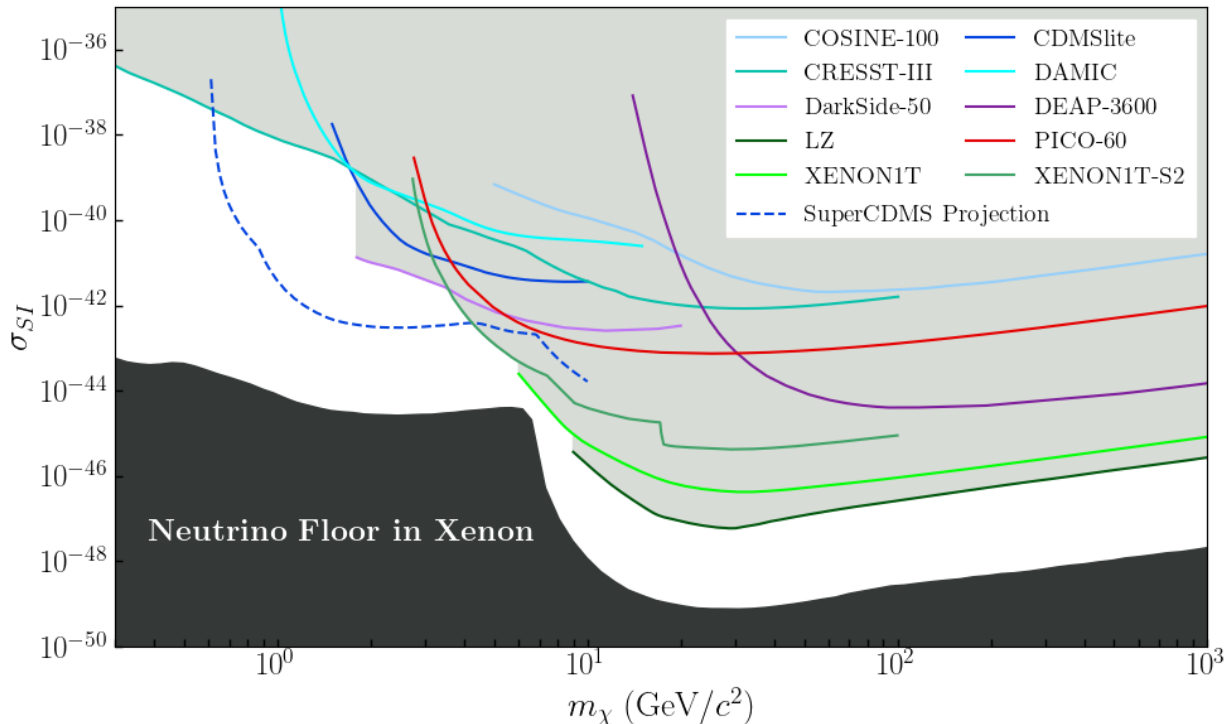


Figure 3.3: Exclusion plot for the spin-independent WIMP-nucleon cross section showing results from the latest experiments. The light grey shaded region shows the envelope of the parameter space excluded by current experiments (solid lines), while the dashed lines show projections for future experiments. The dark grey shaded region corresponds to the neutrino floor for xenon, below which solar neutrinos become an irreducible background. Limits from a variety of different detector technologies are shown in this plot, including liquid noble element detectors [86, 8, 99, 82, 83], cryogenic solid state detectors [80, 92, 12, 100, 11], and superheated bubble chambers [87].

Also shown in Figure 3.3 is the projected sensitivity for SuperCDMS SNOLAB [11]. From this projection it can be seen that SuperCDMS SNOLAB expects to provide new sensitivity to low mass dark matter. The dark grey shaded region at the bottom of Figure 3.3 is known as the neutrino floor for xenon, and corresponds to the region of parameter space where direct detection experiments (using xenon) will become limited by an irreducible background of solar neutrinos [101].

3.3.5.1 Low-energy excesses

Radioactive backgrounds, both internal and external to the target, are important for all direct detection experiments. However, one of the primary challenges currently faced by

3.3. DIRECT DETECTION

direct detection experiments are low-energy (below 100 eV) excesses not associated with any known radioactive backgrounds [95]. As this thesis focuses on the direct detection of dark matter using cryogenic detectors, some more details will be given on the low-energy excesses in solid state devices.

Each of the cryogenic solid state detection technologies described above have seen evidence of low-energy (below about 100 eV) excesses [95]. The proposed explanations for such excesses fall under one of two categories: particle interactions in the detector or surrounding materials, and excesses generated by structural features of the detector. The first category relates to low-energy particle backgrounds such as Cherenkov interactions or scintillation light produced in materials near the detector. Such events could be produced by particle interactions in the detector clamps that produce scintillation light, or impurities on the surfaces of the detector. The second category of low-energy excesses are structural issues with the detector. Stress induced microfractures in the crystal [95, 14] can deposit excess heat in the thermal sensors. These could originate from stress introduced by the detector clamping scheme, or even stress that is intrinsic to the crystal [14]. Microfractures in the crystal could produce low-energy phonons as the stress is released, which would be seen as a background that decays over time and may or may not reappear after the crystal has been warmed up [14, 102]. It is unlikely that the low-energy excesses in each experiment have the same cause, due to the shapes and energies of the low-energy excesses [95], so to better understand these effects dedicated studies are required [14].

Chapter 4

The SuperCDMS experiment

The Super Cryogenic Dark Matter Search (SuperCDMS) experiment is a direct detection experiment that uses cryogenic semiconductor detectors instrumented with superconducting transition edge sensors (TES) to search for dark matter particle interactions [71, 11]. SuperCDMS is the advancement of the Cryogenic Dark Matter Search (CDMS) experiment [103], which used the same detection principle with a slightly different implementation. The first phase of the SuperCDMS experiment took place at the Soudan Underground Laboratory [71], and the collaboration is currently making preparations for the second phase of the experiment, which will take place at SNOLAB [11]. This chapter gives a brief history of the SuperCDMS experiment, details the detection principle and detectors used, and describes the experimental setups of SuperCDMS Soudan and SuperCDMS SNOLAB.

4.1 A brief history of SuperCDMS

The first occurrence of the CDMS experiment, CDMS-I, began collecting data in the Stanford Underground Facility (SUF) at Stanford university in 1998 [104]. The SUF is a shallow underground site with approximately 16 m water equivalent (mwe) of overburden, which shields the detectors from the hadronic component of cosmic radiation, while also reducing the muon flux by approximately a factor of 5 compared to the surface [104]. The experiment used six detectors, made of germanium or silicon, that were 1 cm thick and 7 cm in diameter [104].

4.1. A BRIEF HISTORY OF SUPERCDMS

Several years after the first results from CDMS-I were published, additional funding was secured to begin the next iteration of the experiment: CDMS-II [103]. For the second phase of CDMS, the experiment was moved from SUF to a laboratory hosted in an iron mine in Soudan, Minnesota, called the Soudan Underground Laboratory (SUL). The SUL is a deep underground facility, providing roughly 2090 mwe overburden [105]. In addition to the improved underground facility, new detectors were acquired for CDMS-II: 19 germanium and 11 silicon detectors [103]. CDMS-II used a well-shielded cryostat, with lead shielding to block gamma rays and polyethylene shielding for neutrons [105]. The experiment also featured an active muon veto made of plastic scintillator panels, which helped remove events that were coincident with a muon passing through the experiment. Analysis of the CDMS-II silicon detectors yielded a detection of three WIMP candidate events, which is above the expected background [103]. These events, when interpreted as a WIMP signal, would be consistent with a roughly $8.6 \text{ GeV}/c^2$ WIMP with a WIMP-nucleon cross section of $1.9 \cdot 10^{-41} \text{ cm}^2$ [103], although this interpretation was later excluded by several experiments [71]. However, in order to fully exclude this interpretation, more data must be taken using silicon detectors, as is planned for SuperCDMS SNOLAB [11].

SuperCDMS is the next generation of the CDMS experiment, of which the first phase (SuperCDMS Soudan) used the same experimental setup as CDMS-II, in the SUL [71]. SuperCDMS Soudan featured an improved suite of 15 interleaved Z-sensitive Ionization and phonon (iZIP) germanium detectors, which were cylindrically shaped with a thickness of 2.5 cm and a diameter of 7.6 cm [71]. During SuperCDMS Soudan, some of the detectors were operated in a new mode called the CDMS low ionization threshold experiment (CDMSlite) [92]. A relatively high voltage (about 70 V) was applied between the faces of the CDMSlite detectors in order to amplify a small charge signal into a large phonon signal; the details of CDMSlite operation will be discussed later. The success of the CDMSlite detectors inspired further development of this technique, to be used in SuperCDMS SNOLAB.

The second and current phase of the SuperCDMS experiment is located near Sudbury, Ontario, in SNOLAB [11]. SNOLAB boasts the world's deepest operational clean

room facility, thanks to a class 2000 (or better) clean room at a depth of roughly 2 km [106]. SuperCDMS SNOLAB will use 24 germanium and silicon cylindrical detectors that are 30 mm thick with a diameter of 100 mm [11]. In contrast to SuperCDMS Soudan, SuperCDMS SNOLAB will feature two types of detectors in their search [11]: the next generation of iZIP detectors, as well as the new high voltage (HV) detectors, which were developed to build on the success of CDMSlite. In addition to the improved detector technology, SuperCDMS SNOLAB will also feature improved shielding, making SuperCDMS SNOLAB the most sensitive iteration of the CDMS experiment to date.

4.2 Detection principle

The SuperCDMS detection principle is based on the measurement of heat and ionization signals in a semiconductor crystal. The heat signal corresponds to the presence of phonons (quanta of lattice vibrations) in the crystal, measured by superconducting transition edge sensors, while the ionization signal corresponds to electrons (and holes) that are excited from the valence to the conduction band and collected by charge electrodes. This section will describe the relevant physics surrounding the generation of ionization and phonon signals in the crystal, and describe the transduction of the phonon and ionization signals into measurable voltage signals.

A number of different interaction mechanisms are possible in a detector depending on the particle type and energy. The particle can be fully absorbed, scatter elastically, or scatter inelastically; these processes can happen on an electron, corresponding to an electron recoil (ER), or on an atomic nucleus, corresponding to a nuclear recoil (NR). Particles that couple to electromagnetism, such as electrons or photons, can produce electron recoils, while nuclear recoils are mostly caused by the scattering of neutral particles, such as neutrons. Traditionally, direct detection experiments have looked for signals produced by nuclear recoils, as the energy transfer to an electron is expected to be on the order of a few eV, whereas the energy transferred to nuclei could be up to a few 100 keV. However, in recent years there has been an increase in alternative dark matter particle candidates (to WIMPs) that feature a (small) coupling to electromagnetism, which would produce an electron recoil signal in the detector [107]. As such, there is great value in

4.2. DETECTION PRINCIPLE

performing dark matter searches using both electron and nuclear recoil signals. Developments in the phenomenology of dark matter interactions introduce new mechanisms for experimentalists to test. For instance, the Migdal effect is a mechanism where a nuclear recoil can produce a signal in a detector that resembles an electron recoil [108]. This is one example of a mechanism that falls outside of the typical description of nuclear and electron recoils, and almost certainly new mechanisms will be discovered.

4.2.1 Semiconductor physics

When sufficient energy is deposited in a semiconductor, electrons are promoted from the valence band to the conduction band, leaving holes in the otherwise fully populated valence band. The electrons in the conduction band can move freely and the holes in the valence band can be filled with electrons from neighbouring atoms, leaving them with a hole. In effect, the holes can be considered as a freely moving positively charged particle. In the presence of an electric field across the detector, the electrons and holes drift towards oppositely charged electrodes on the surface of the detector; electrons will move to the positive electrode with velocities greater than 20 km/s in germanium [109], and holes will drift towards the negative electrode. In reality, the motion of the electrons and holes is not entirely straightforward, as holes will travel in a straight line in the direction of the electric field lines, while electrons interact with the crystal at oblique angles due to the shape of the conduction band. At low voltages the electrons generally follow the valleys of the conduction band, while at high voltages inter-valley scattering dominates and they appear to go in straight lines [109, 110]. The large drift velocity of the charge carriers means that in a large SuperCDMS detector, the charges traverse the crystal and induce a signal in under 2 μ s.

When a particle interaction occurs in a semiconductor, the recoil energy (E_R) will be deposited in the crystal, producing electron-hole (e/h) pairs and phonons. The amount of ionization pairs produced depends on the interaction type (ER or NR); for a given recoil energy nuclear recoils will produce fewer electron-hole pairs than electron recoils. The ionization yield $Y(E_R)$ gives the ratio of the interaction energy (E_Q), if a standard electron recoil is assumed, to the recoil energy:

$$Y(E_R) = \frac{E_Q}{E_R}. \quad (4.1)$$

For an ER the ionization yield is defined as 1, while the ionization yield for a NR is less than 1 since less energy goes into the charge carriers for a nuclear recoil. The expected number of electron-hole pairs $\langle N \rangle$ produced by an interaction is related to the recoil energy, and the average energy required to produce an electron-hole pair ϵ_γ :

$$\langle N \rangle = \frac{E_R}{\epsilon_\gamma}. \quad (4.2)$$

The average energy required to produce an electron-hole pair in germanium and silicon is 3.0 eV and 3.8 eV, respectively, while the minimum energy required to produce an electron-hole pair is given by the band gap energy, which is 1.12 eV in germanium and 0.66 eV in silicon, at low temperatures [111, 112]. The variance in the expected number of e/h pairs σ_N^2 is defined as the product of the expected number of e/h pairs and the Fano factor F , which quantifies the deviation from Poisson statistics and is a fundamental property of the detector material; the Fano factor for germanium and silicon is 0.11 and 0.12, respectively [113].

As the ionization yield is greater for an electron recoil than a nuclear recoil, the type of recoil can be determined on an event-by-event basis by comparing the ratio of the ionization signal to the phonon signal. At low energies this discrimination breaks down since the signals are so small [114], but this effect has been used with great success in the iZIP detectors, which will be described in more detail later in this chapter.

4.2.2 Phonon physics

Phonons are quanta of lattice vibrations in a crystal. Germanium (silicon) forms crystals with a face-centered cubic structure with two (eight) germanium (silicon) atoms per cell, and a lattice spacing of 5.66 (5.43) Å [115]. Phonons in a crystal can be modelled as a system of bosons, with a thermal number density given by a Planck distribution [115]:

$$\langle n(\omega) \rangle = \frac{1}{e^{\hbar\omega/k_B T} - 1}. \quad (4.3)$$

As the temperature of a typical SuperCDMS detector is on the order of tens of millikelvin, the thermal phonons are negligible compared to the athermal phonons produced

4.2. DETECTION PRINCIPLE

in a particle interaction. There are two main types of phonons: acoustic and optical, which correspond to the lattice vibrating in or out of phase, respectively. The main phonons that will be detected in a SuperCDMS detector are acoustic phonons, as the optical phonons will decay to acoustic phonons on a timescale smaller than the time they take to traverse the detector.

After a particle interaction occurs in the crystal, the initial energy in the phonon system is given by:

$$E_{initial} = E_R - nE_{gap}, \quad (4.4)$$

where n is the number of e/h pairs produced in the interaction, and E_{gap} is the band gap energy. The energy in the phonon system is reduced by the energy that went into producing the electron-hole pairs, however, in the absence of a bias voltage across the detector, the e/h pairs will recombine almost immediately, releasing recombination phonons with energy equal to nE_{gap} . If there is a detector bias voltage, then the electrons and holes will first drift to their respective charge electrodes before eventually recombining and depositing nE_{gap} into the phonon system.

If a bias voltage is applied across the detector, an additional source of phonons is introduced into the system as the charge carriers are drifted across a voltage potential, this is called the Neganov-Trofimov-Luke (NTL) effect [116, 117]. As electrons and holes are drifted through the crystal they quickly reach a terminal velocity, so the electric potential energy is deposited into the phonon system rather than the kinetic energy of charge carriers. The energy of the NTL phonons is related to the detector bias voltage V_{bias} and the number of charge carriers n by:

$$E_{NTL} = neV_{bias}, \quad (4.5)$$

where e is the electric charge. Combining these different contributions to the phonon system, the total energy in the phonons E_T from a particle interaction can be written as:

$$E_T = E_{initial} + E_{recombination} + E_{NTL} = E_R + neV_{bias}. \quad (4.6)$$

4.2.3 Athermal calorimeters

When a particle interaction occurs in a detector, the target material absorbs some of the energy and produces athermal excitations in the form of phonons, ionization, or photons

(in the case of a scintillator). These excitations are then thermalized through inelastic scattering processes and the detector cools back down to its equilibrium temperature. In a calorimeter, the target material (absorber) has some heat capacity C_{abs} , and is connected to a thermal bath with a thermal conductance G_{ab} [118]. The absorber can be instrumented with a thermal sensor in order to accurately measure the thermal excitations. When a particle interaction occurs in the absorber the temperature measured by the sensor rises, reaches a peak, and then falls; this behaviour is called a pulse. The rise and fall times of the pulse are determined by two (or more in practical calorimeters) characteristic times: the thermalization time and $\tau = C_{abs}/G_{ab}$, the faster time is the rise time and the slower time is the fall time [118, 119]. In this basic implementation of a calorimeter, the theoretical energy resolution σ_E is limited by thermal fluctuations across the thermal link between the absorber and the bath, such that $\sigma_E \approx k_B C_{abs} T^2$ [118], which suggests that lower temperatures will improve the sensitivity of the detector.

The thermal conductance between the absorber and the sensor is caused by electrons in the thin metal film of the sensor being driven out of thermal equilibrium [120], which scales as T^4 , while the thermal conductance between the bath and the absorber is driven by the mismatch of different semiconducting substrates [121], which scales as T^3 . As the temperature of the absorber is lowered it becomes difficult to thermally couple the sensor and absorber. To alleviate this issue a sensor with a fast response time, such as a TES, can be used to measure the athermal excitations before thermalization; this configuration is known as an athermal calorimeter [118, 119]. For the case of TESs, the relevant heat capacity is instead the heat capacity of the sensor (which scales with the volume of the sensor), and the time constants are determined by the collection time of the athermal excitations and the thermal time constant of the sensor. Athermal calorimeters based on transition edge sensors can achieve excellent energy resolutions [122], which is a desirable feature for low-mass dark matter searches.

4.2.4 Transition edge sensors

Transition edge sensors are a superconducting sensor commonly used in astronomy and rare event searches [71, 12, 123]. Transition edge sensors are generally operated in the

4.2. DETECTION PRINCIPLE

transition regime between the superconducting and normal state of a superconductor by applying a bias current to the sensor [124]. A small amount of energy absorbed in the TES can break Cooper pairs in the TES, causing a steep, measurable change in the resistance [124]. As the effective band gap (of Cooper pairs) for a superconductor in transition is zero, the TES is sensitive to very small depositions of energy. SuperCDMS uses an array of TESs in parallel that are instrumented on the surface of the substrate by a sputtering technique and patterned into TESs by lithography [11]. The transition edge sensors used by SuperCDMS are made of tungsten, with aluminum fins to increase the collection efficiency [71].

The energy resolution of a TES is related to its heat capacity, which is proportional to its volume, this means that physically smaller TESs will be more sensitive. However, the TES must also be able to collect phonons with a high efficiency, which requires a larger area. To do this efficiently, the small TES ($10^2 \mu\text{m}^2$) is expanded with large ($10^4 \mu\text{m}^2$) aluminum collecting fins [124]. This combination of a TES with collecting fins is called a quasiparticle trap-assisted transition edge sensor (QET) [124]. The aluminum collecting fin covers a large area on the surface of the substrate, which allows for a large number of phonons to be collected within several reflections off of the surface. The TES without the collecting fins requires a trade-off between the collection efficiency and heat capacity; the inclusion of the fins greatly improves the collection efficiency allowing for a high collection efficiency and a low heat capacity.

When a phonon interacts in the superconducting aluminum, it breaks Cooper pairs into Bogoliubov quasiparticles (superpositions of electrons and holes) if more than twice the band gap of the Cooper pairs is deposited in the film [124]. The binding energy $E_b = 2\Delta \approx 3.5k_B T_c$ is twice the superconducting band gap Δ for temperatures well below the critical temperature T_c ($T_c \approx 1$ K for aluminum [124]), which is roughly 0.36 meV for aluminum [115]. After the Cooper pairs are broken, the quasiparticles will diffuse through the aluminum until they recombine or become trapped in the tungsten. Since the T_c of tungsten is lower than aluminum's (and equivalently so is the binding energy), the change in energy levels makes it energetically favorable for quasiparticles to move into the overlap region of the tungsten and aluminum without being able to move back,

effectively becoming trapped. These quasiparticles interact with the local electrons in the tungsten TES, depositing heat and causing the resistance of the TES to increase [124].

At temperatures well below the critical temperature, the TES behaves as a superconductor and dissipates virtually no power; at temperatures well above T_c , the TES behaves as a normal metal with a (roughly) constant normal resistance R_N . When being used as a sensor, TESs are operated in the narrow transition between the superconducting and normal regimes, which typically has a width of only a few mK. This transition is driven by the structure of the thin film, which is not purely crystalline like a bulk superconductor, but has small patches with different transition temperatures that pass through the transition regime at different temperatures. Transition edge sensors can be operated in a voltage-biased mode, where they self-regulate through negative electrothermal feedback [124]. In this case, when heat is introduced to the sensor and the resistance increases, the Joule power being dissipated in the sensor decreases since $P_J = V_{TES}^2/R_{TES}$, which in turn causes the temperature of the sensor to decrease and return to its steady state [124]. The alternative mode of operation would be to current-bias the sensor, which would result in positive electrothermal feedback as a small change in resistance would increase the Joule heating (since $P_J = I_{TES}^2 R_{TES}$), causing the temperature of the sensor to continually increase [124]. For this reason, TESs are generally operated in the voltage-biased regime through use of a circuit similar to the one in Figure 4.1. In this diagram the TES is shown as a variable resistor R_{TES} in series with an inductor L_{in} (the input coil), also shown in this branch is a small, parasitic resistance R_p that arises from finite resistances in the connections and line. The TES branch of the circuit is in parallel with a small shunt resistor R_{sh} (typically on the order of a few m Ω), this entire circuit is supplied by a current source I_b . Once the TES is in the transition with $R_{TES} \gg R_{sh}$, $I_{sh} \approx I_b$ so that the voltage drop across the TES $V \approx I_b R_{sh}$, which is independent of R_{TES} .

The current in the TES branch of the circuit decreases with increasing TES resistance, this change in current is measured by inductively coupling the in-series inductor to a superconducting quantum interference device (SQUID). A SQUID is a sensitive magnetometer that is used to amplify the magnetic field produced by the inductor into a measurable voltage; the changing current through the inductor changes the magnetic

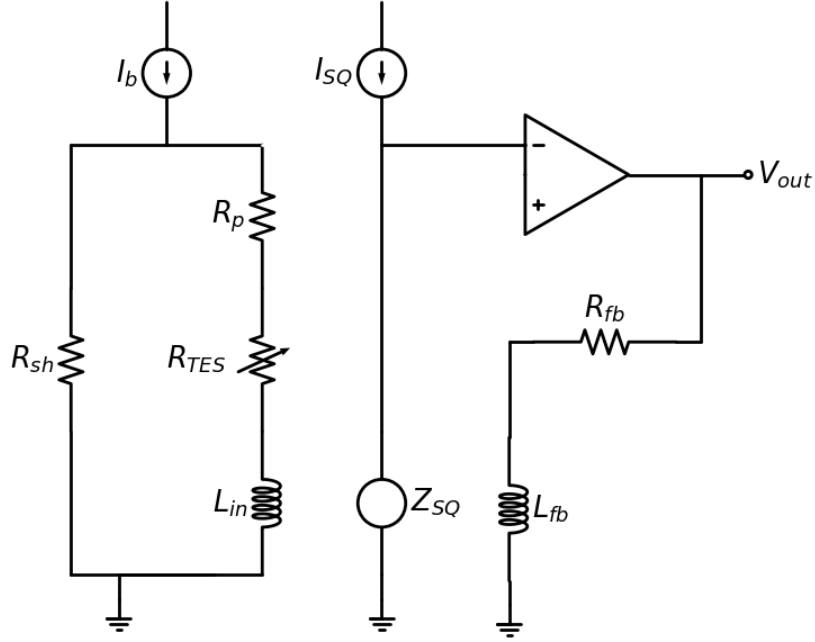


Figure 4.1: Simplified TES readout circuit diagram, the respective symbols are defined in text. The TES is depicted as a variable resistor and is voltage-biased by placing it in parallel with a small shunt resistor. The current through the TES can be measured by inductively coupling the input coil to a SQUID. To linearize the SQUID response a portion of the amplifier output is fed back to the SQUID through a feedback resistor and inductor.

field through the SQUID, resulting in a change in potential across the SQUID [124]. The SQUID is connected to a room temperature amplifier, the output of which is connected to a feedback resistor R_{fb} in series with a cold feedback coil L_{fb} . The feedback coil is inductively coupled to the SQUID to compensate the flux through the SQUID, thereby linearizing the SQUID response. After setting the bias current through the SQUID I_{SQ} and connecting the SQUID to the feedback circuit, the output voltage of the SQUID is:

$$V_{out} = I_{TES} R_{fb} \frac{N_i}{N_{fb}}, \quad (4.7)$$

where $\frac{N_i}{N_{fb}}$ is the ratio of the turns between the input coil and feedback coil. In the SuperCDMS Soudan electronics, the value of the feedback resistor was approximately 1 k Ω and the loop ratio was about 10; in the SNOLAB electronics, the value of the feedback resistor is approximately 5 k Ω and the loop ratio is about 2.4 [102].

The output of this preamplifier circuit is conditioned by a main amplifier then connected to an analog-to-digital converter (ADC), which digitizes the analog output of the amplifiers so it can be stored by the data acquisition system (DAQ). The quantity stored by the DAQ is the measured number of ADC bins n_{ADC} , which can be converted to the current through the TES as:

$$I_{TES} = \frac{n_{ADC} V_{ADC}}{N_{ADC} R_{fb}} \frac{N_{fb}}{N_i} \frac{1}{G}, \quad (4.8)$$

where V_{ADC} is the total voltage range of the ADC (8 V in the SNOLAB electronics), N_{ADC} is the total number of ADC bins (2^{16} in the SNOLAB electronics), and G is the adjustable gain of the amplifiers. It should be noted that this quantity is sensitive to the offset of the ADC, but changes in current, which is what is mainly of interest, are still accurate.

4.2.5 Charge readout

A simplified diagram of the SuperCDMS charge read-out circuit is shown in Figure 4.2. A bias voltage V_{bias} is applied to the detector through a bias resistor R_{bias} . The drifting charges induce a mirror charge in the electrodes, which is coupled by a coupling capacitor C_c , to a sensitive charge amplifier with a feedback capacitor C_{fb} . Stray capacitance in the circuit is represented as C_s . The charge amplifiers used by SuperCDMS Soudan were based on FETs [105], while SuperCDMS SNOLAB will use charge amplifiers based on HEMTs [125] for the iZIPs. The output voltage of the amplifier is proportional to the induced charge Q_{ind} by $V_{out} = Q_{ind}/C_{det}$, where C_{det} is the capacitance of the detector. The signal then decays away as the capacitor discharges with a fall time of $\tau_Q = R_{fb}C_{fb}$.

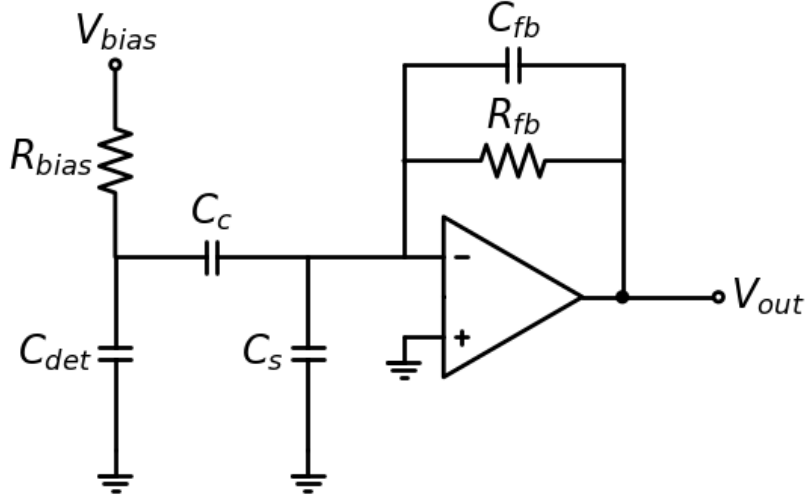


Figure 4.2: Simplified diagram of the circuit used to read out the charge signal from the detector, see text for details.

4.3 SuperCDMS Soudan

The first phase of the SuperCDMS experiment took place in one of the caverns in the Soudan Underground Laboratory. SuperCDMS Soudan used the same experimental apparatus as CDMS-II, but with improvements to the detectors [71]. A dilution refrigerator was used to cool the detectors to temperatures below 50 mK. The cryostat was surrounded by shielding, and connected to the dilution refrigerator by the C-stem. A schematic of the experimental setup can be seen in Figure 4.3. This section will describe the detectors used in SuperCDMS Soudan, the shielding, calibration methods, and experimental backgrounds.

4.3.1 Detectors

SuperCDMS Soudan used 15 high-purity germanium detectors, mounted in “tower” configurations with three detectors per tower. The towers contained all of the cold electronics necessary to operate detectors, and were connected to room temperature read-out electronics by long read-out cables. The detectors used in SuperCDMS Soudan were cylindrical in shape, 25 mm thick with a diameter of 76 mm, and weighed roughly 0.6 kg, an image of a detector is shown in Figure 4.4a. Each detector was instrumented on

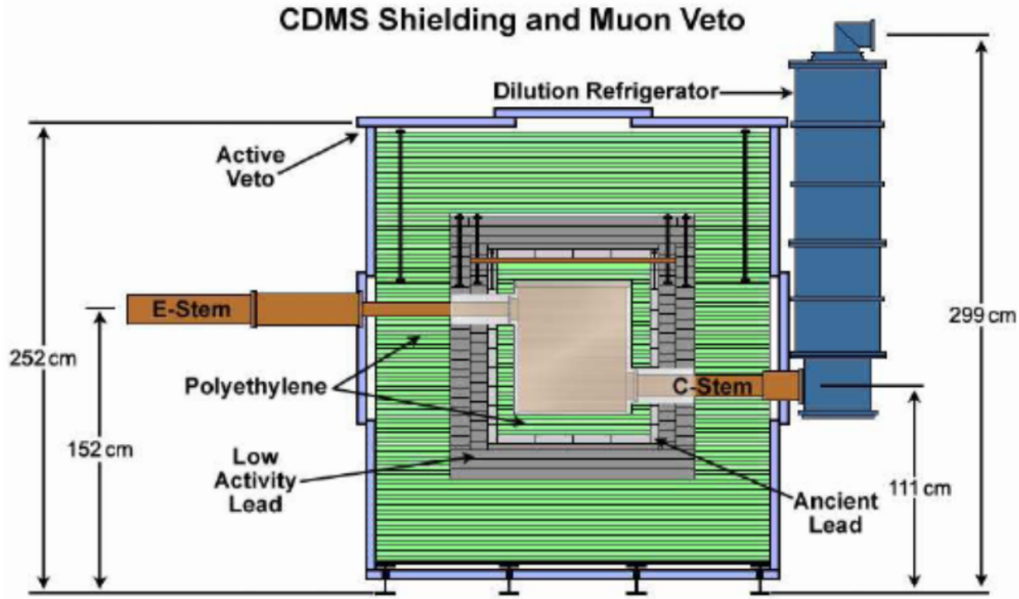
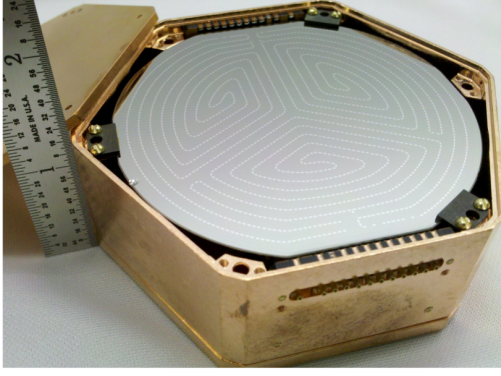


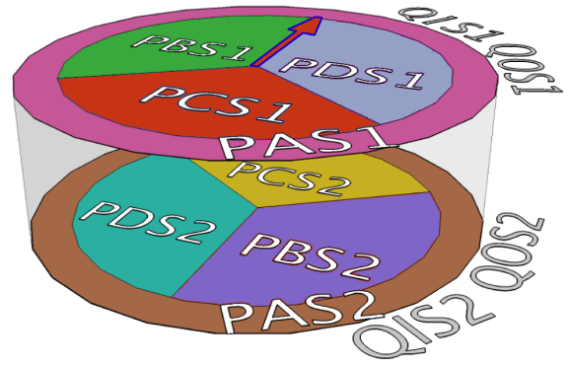
Figure 4.3: Experimental setup of SuperCDMS Soudan. The detectors are mounted inside the cryostat at the center of the image. The cryostat is surrounded by layers of passive polyethylene and lead shielding, which are in turn surrounded by an active muon veto made of scintillator panels. The dilution refrigerator on the right is connected to the cryostat through the C-stem, and the cabling is passed out of the cryostat through the E-stem. Image taken from Reference [126]

both of the flat faces with an array of QETs interleaved with charge electrodes, hence the name interleaved Z-sensitive Ionization- and Phonon-mediated (iZIP) detectors [71]. Each detector had 4 phonon channels and 2 charge channels per side, the layout is shown in Figure 4.4b; there were three phonon channels in the center and another surrounding these, the charge channels were on the inner and outer parts of the detector respectively.

In the normal (iZIP) operating mode of the detectors, the TESs were put into transition and a small bias voltage of ± 2 V was applied to the electrodes on opposite faces of the detector. In this mode both the phonon and charge signals were measured by their respective sensors; the ratio between the charge and phonon signal provided discrimination between ER and NR events [71]. When a particle interaction occurs near the flat surfaces, the electric field produced between the interleaved phonon and charge sensors causes the electrons and holes to be collected entirely on that side, in contrast to the symmetric charge collection for interactions in the bulk. This interleaved layout enables



(a) A SuperCDMS Soudan germanium iZIP detector in its copper housing. Image from Reference [127]



(b) Layout of the four charge (labelled with Q) and eight phonon channels (labelled with P) on a SuperCDMS Soudan iZIP detector. Image from Reference [128]

Figure 4.4: A SuperCDMS Soudan iZIP detector and a schematic of its sensor layout.

discrimination between events that occurred near the flat surfaces and in the bulk by comparing the charge collection [129]. Additionally, events that occur near the round detector sidewalls can be excluded by comparing the ratio of the signal in the inner channels to outer channels, this feature is useful when considering background events introduced during the detector fabrication, which will be discussed later.

During SuperCDMS Soudan, a new operating mode was tested to take advantage of the NTL amplification and gain sensitivity to lower energy recoils. In this mode, known as the CDMS Low Ionization Threshold Experiment (CDMSlite) [92], a relatively large bias voltage (70 V) was applied across the faces of the detector, by doing this a small ionization signal can be amplified into a large phonon signal. The original Soudan electronics were not designed to apply detector bias voltages greater than 10 V, so custom electronics were implemented to hold one side of the detector's phonon sensors and charge electrodes at the elevated voltage. As the electronics were not designed to read-out channels at an elevated voltage, only the grounded (non-biased) side of the detector was read-out, resulting in a loss of phonon collection efficiency by half. The charge signal that was read out was only used to tag periods of bad noise.

4.3.2 Shielding

The depth of the Soudan Underground Laboratory provides an overburden of 2090 mwe, reducing the surface muon flux by a factor of $5 \cdot 10^4$ [105]. The dilution refrigerator, cryostat, and shielding were contained within an RF tight room to reduce noise from external electromagnetic sources. Within the RF room, passive and active shielding layers were used to reduce backgrounds. The outermost layer of shielding was an active veto, which consisted of 40, 5 cm thick scintillator panels read out by PMTs [105]. The scintillator panels were arranged so that adjacent panels had some overlap, ensuring no direct line of sight to the detectors. The veto was able to distinguish between muons and ambient photons (from the cavern), and despite the depth the rate of muons is roughly 1 muon per minute [105]; photons from the cavern dominate the total veto rate, which is roughly 600 Hz. The first layer of passive shielding within the muon veto was a 40 cm thick cylindrical polyethylene shield that moderated low-energy neutrons from radioactive decays. Inside the outer layer of polyethylene was a 22.5 cm thick cylindrical lead shield, of which the innermost 4.5 cm was composed of ancient lead. Ancient lead is low in ^{210}Pb , which decays via beta-decay and produces a broad range of photons through Bremsstrahlung, so the ancient lead served to moderate the ^{210}Pb decays from the outer part of the lead shield, while not contributing to this background. Within the lead shielding was an additional 10 cm thick cylindrical polyethylene shield that further moderated the neutrons that passed through the outer layer. The copper cans of the cryostat provided an additional shielding of roughly 3 cm of copper. The SuperCDMS Soudan shielding was roughly 99% hermetic, with the only gaps being along the stems that connected the cryostat to the electronics and dilution refrigerator.

4.3.3 Backgrounds

The backgrounds in the experiment are comprised of four types of particles: alphas, betas, gammas, and neutrons. As neutrons scatter via nuclear recoils, they cannot be rejected from a nuclear recoil based dark matter search on an event-by-event basis, unless they scatter in multiple detectors. Some gammas and electrons can also be misidentified as nuclear recoils if they scatter near the detector's surface where the poor charge collection

4.3. SUPERCDMS SOUDAN

leads to a low ionization yield. These surface events make up the dominant background component of the experiment, but can be rejected thanks to the sensor layout of the iZIP detectors.

Surface backgrounds arise as a result of contamination on the detector's surface or adjacent materials; the main sources of the surface beta background are: ^{210}Pb , ^{40}K , and ^{14}C [71]. In addition to betas, surface contamination also contributes to the alpha background. Alpha particles produce large recoil energies and low ionization energies, which makes them relatively easy to identify. However, the recoiling nucleus from the alpha decay may deposit energy in the range of 20-100 keV, which could be mistaken for a WIMP event if the alpha particle is not also identified. The detection of alphas in the detector is not only useful for rejecting the signal from the recoiling nucleus, but also for estimating the level of contamination with ^{210}Pb , the decay chain of which has ^{210}Po , an alpha emitting daughters.

The neutron background in the experiment arises from natural radioactivity, for example decays in uranium and thorium along with their decay products, and muon interactions in the surrounding materials. The rate of neutrons in the rock surrounding the experiment is high, but their energies are only on the order of a few MeV, meaning that the polyethylene shield moderates the flux by a factor of roughly 10^6 [11], making this contribution to the neutron background negligible. Cosmogenic neutrons produced by muons passing through the shielding will be vetoed by the scintillator panels, so this contribution is also small. The main contribution to the neutron background is from muon-induced neutrons in the surrounding rock that have energies higher than about 50 MeV, which can penetrate through the polyethylene shield and produce additional secondary neutrons [11] (for instance by scattering off of lead). The number of neutrons produced from decays in the rock is negligible as their energies are too low to make it through the shielding [11].

The gamma background comes mostly from contamination of the copper cryostat cans and the shielding materials. Radioactive isotopes such as uranium, thorium, and potassium in the copper cans, polyethylene, and lead shields can contribute to the gamma background at energies around a few MeV. ^{210}Pb in the lead shielding beta-decays, and

the emitted electrons produce Bremsstrahlung X-rays that may interact with the detector. A final source that contributes to the gamma ray background is radon inside the shielding volume. To mitigate this background source the volume is purged with nitrogen that was stored for long enough to allow the radon to decay ($T_{1/2} = 3.8$ days) [11]. The dominant electron recoil background is from cosmogenic activation of the germanium detectors. Isotopes like ^{55}Fe and ^{65}Zn introduce monoenergetic peaks in the spectrum, produced from the decay via electron-capture [11]. One of the dominant backgrounds is tritium (^3H), produced from nuclear spallation during interactions of high-energy cosmic rays, which beta-decays with an endpoint of roughly 20 keV [11].

4.3.4 Calibration

To calibrate the detectors' energy scale and response to nuclear recoils, several external sources were used. Two ^{133}Ba sources were fed into the cryostat along the electronics stem and cryogenics stems. ^{133}Ba has several high energy lines at 276 keV, 303 keV, 356 keV, and 384 keV, all of which are sufficiently energetic to penetrate through the copper cans. However, only the ionization (charge) channels have a linear response at these energies, as the phonon channels become highly non-linear due to saturation of the TES above about 200 keV [11]. As a result, the barium source is used to calibrate the ionization channels, and the phonon channels are calibrated against this for electron recoils.

A ^{252}Cf source is used for calibrating the detector's response to nuclear recoils. This also validates the calibration of the energy scale for the phonon channels obtained from the barium source, which remains in good agreement even for nuclear recoils [11]. The ^{252}Cf also has the added benefit of activating the germanium detectors, by producing ^{71}Ge through neutron capture on ^{70}Ge . ^{71}Ge decays via electron-capture, emitting a low-energy line at either 10.37 keV, 1.30 keV, or 0.16 keV, corresponding respectively to the K-, L-, and M-shell energies of ^{71}Ga . These peaks are useful in calibrating the CDMSlite detectors, which saturate at much lower energies (due to the high NTL gain) and cannot benefit from the cross calibration of the ionization channels.

4.4 SuperCDMS SNOLAB

The second phase of the SuperCDMS experiment will take place at SNOLAB; SuperCDMS SNOLAB will feature reduced backgrounds, thanks to improved shielding and material screening, as well as more sensitive detectors. A schematic of the SuperCDMS SNOLAB experiment can be seen in Figure 4.5. While SuperCDMS Soudan used most of the same infrastructure as CDMS-II [105, 71], SuperCDMS SNOLAB will feature an entirely new cryostat, shielding, detectors and electronics [11]. While the main design of the SuperCDMS SNOLAB facility will be similar to that of SuperCDMS Soudan, the primary difference is that SuperCDMS SNOLAB will have dedicated high-voltage detectors, with electronics designed accordingly, whereas SuperCDMS Soudan used the same detectors and electronics used for the iZIP detectors.

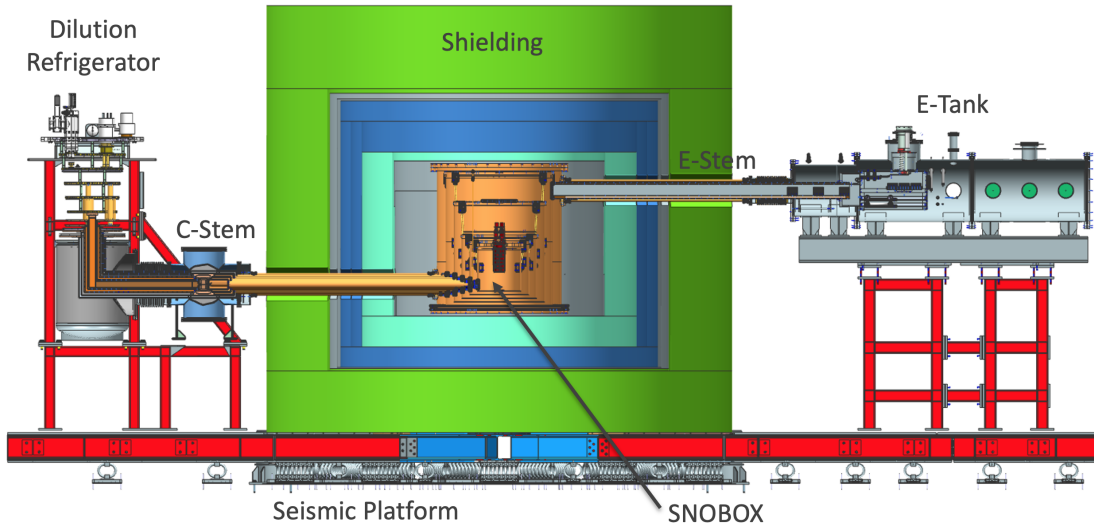
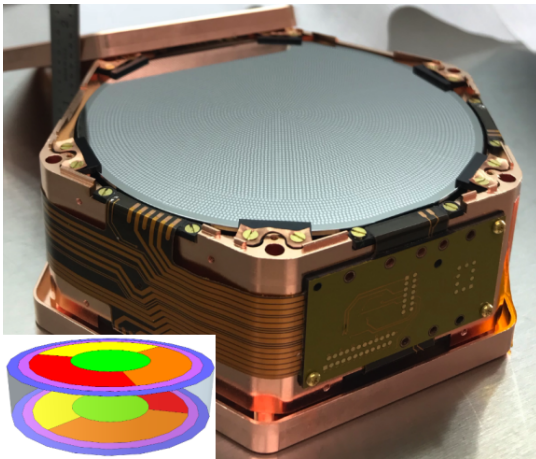


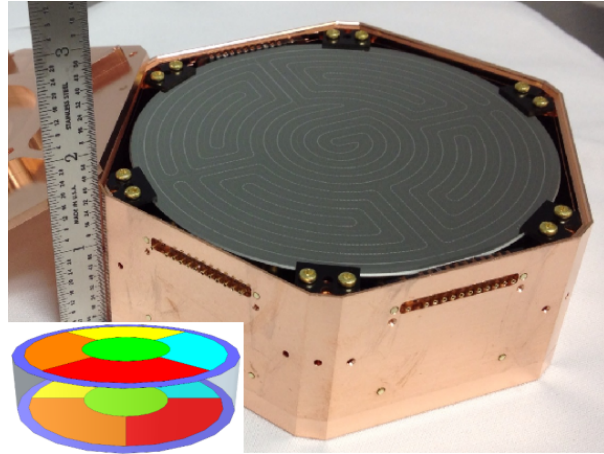
Figure 4.5: Schematic of the SuperCDMS SNOLAB experiment, highlighting the shielding and cryostat. The cryostat, called the SNOBOX, is mounted within layers of neutron and gamma shielding. The SNOBOX is connected to the dilution refrigerator by the C-Stem, and the electrical signals are passed to the E-Tank through the E-Stem. The entire setup sits on a seismic isolation platform. Image taken from Reference [11].

4.4.1 Detectors

SuperCDMS SNOLAB will use cylindrical detectors, with a thickness of 33 mm and a diameter of 100 mm, made of either silicon or germanium. In total there will be four different types of detectors, iZIP and high-voltage (HV), made of either silicon or germanium. The iZIP detectors will have 12 phonon channels (6 per side), and two interleaved charge electrodes on the inner and outer parts of each detector face (making four charge channels in total). The HV detectors will not have charge electrodes, but will have 12 phonon channels in a slightly different layout than the layout on the iZIP detectors. An image of a SuperCDMS SNOLAB HV detector is shown in Figure 4.6a, while a SuperCDMS SNOLAB iZIP detector is shown in Figure 4.6b.



(a) SuperCDMS SNOLAB HV detector, with the layout of the phonon sensors shown in the bottom left.



(b) A SuperCDMS SNOLAB iZIP detector in a prototype copper housing, the layout of the phonon sensors is shown in the bottom left.

Figure 4.6: The HV and iZIP detectors to be used for SuperCDMS SNOLAB, along with their respective sensor layouts.

The detectors will be mounted in tower configurations, with six detectors per tower and four towers in total. Two of the towers will be composed of only iZIP detectors while the other two will include only HV detectors; one iZIP tower will include only germanium detectors, while the other towers will each have two silicon detectors and four germanium detectors. The channels of each detector are wirebonded to a horizontal-flexible cable

(HFC) that wraps around the detector housing, and connects to a vertical-flexible cable (VFC) that supports the cold electronics and goes along the length of the tower. The flexible cables have superconducting traces backed by Kapton with the electronic components and connectors mounted on Cirlex stiffener boards. The bottom end of the VFC is connected to a short superconducting cable, which connects to a long read-out cable at the 4 K stage that brings the signals up to room temperature. A vacuum-interface board (VIB) connected hermetically to a vacuum port on the cryostat passes the signals to outside the cryostat. The VIB is connected to the room temperature read-out electronics, known as the detector control and read-out cards (DCRCs). The grounds of the SuperCDMS SNOLAB DCRCs float with respect to the chassis of the experiment, allowing for the phonon channels of the HV detectors to be read-out while biased, something that was not possible with the SuperCDMS Soudan electronics. While the electronics used by SuperCDMS SNOLAB are different than those used in SuperCDMS Soudan, the general operation principle remains the same.

4.4.2 Shielding

A schematic of the SuperCDMS SNOLAB experiment is shown in Figure 4.5. A dilution refrigerator is used to cool the detectors to temperatures below 30 mK. The cold region of the experiment, called the SNOBOX, includes several layers of copper cans, the innermost of which is where the detector towers are mounted; the SNOBOX is thermally connected to the dilution refrigerator by the C-stem. The SNOBOX, dilution refrigerator, shielding, and room temperature electronics are all mounted on a seismic platform to isolate from seismic activity, such as rockbursts (from mining activities) or earthquakes. The SNOBOX is enclosed in several layers of movable shielding, in which the only gaps are for connecting the SNOBOX to the dilution refrigerator via the C-stem, and the electronics tank (E-tank) via the E-stem. High-density polyethylene (HDPE) plates that are 60 cm thick are used at the bottom of the experiment to block neutrons, the rest of the shielding sits atop these HDPE plates. The outermost layer of shielding consists of several, 60 cm thick, water tanks that moderate the flux of neutrons; the area within the water tanks is purged with low-radon air [11]. The next layer of shielding within the water tanks

are pieces of 23 cm thick low-activity lead, which provide a nearly full 4π coverage from gamma rays. Within the lead shield is an additional 40 cm of HDPE to further moderate the flux of neutrons. The final layer of shielding are the copper cans of the SNOBOX.

4.4.3 Backgrounds

The expected backgrounds of SuperCDMS SNOLAB come from impurities in the detector, contaminations in the surrounding materials, and radioactive decays from the rock in the cavern [11]. These sources can produce both ER and NR type events in the detectors.

Radioactive impurities in the crystals are expected to be the dominant background for the HV detectors. A major contaminant of the detectors is tritium (^3H), which is produced as a spallation product from high-energy cosmic rays interacting with the crystal (while the crystal is on surface). Tritium has a long half-life ($T_{1/2} = 12.3$ years) and decays via beta-decay with an endpoint energy of 18.6 keV, which covers the entire region of interest of the HV detectors [11]. For the silicon detectors, naturally occurring ^{32}Si will be present, which has a very long half-life ($T_{1/2} \approx 153$ years) [11]. Both of these isotopes are modelled using a generic beta spectrum, although their exact quantities are not easily predicted (especially for ^{32}Si). For the germanium detectors there are several contaminants are produced by cosmogenic activation, the most important of which are: ^{68}Ge , ^{68}Ga , ^{65}Zn , ^{73}As , ^{57}Co , ^{55}Fe , ^{54}Mn , and ^{49}V [11]. Each of these can decay via electron-capture from the K-, L-, or M-shell, giving rise to a total of 24 monoenergetic peaks.

Radioactive impurities such as ^{238}U , ^{232}Th , ^{60}Co , and ^{40}K can be introduced during the manufacturing process or from cosmic ray secondaries [11]. Cosmogenic activation of the materials surrounding the detectors, in particular the copper cans and tower assembly can have several radioactive cobalt isotopes, and introduce additional gamma and beta backgrounds. These contaminants are generally present at low levels in most materials, such as the detector clamps and the VFCs, requiring the careful screening of materials prior to their inclusion in the experiment.

The Norite rock surrounding the cavern at SNOLAB introduces several gamma and neutron backgrounds to the experiment. Gamma backgrounds from the decay of ^{238}U ,

^{232}Th , and ^{40}K could penetrate through the shielding, although this is not expected to be a dominant contribution to the gamma background. Neutrons produced from the spontaneous fission of ^{238}U , along with (α, n) reactions in the rock from uranium and thorium decay chains are expected to contribute to the NR background [11]. Finally, radon in the air could introduce several high energy gammas via the ^{214}Pb and ^{214}Bi daughters, although this background should be mitigated by the radon-free purge gas.

Surface backgrounds are expected to be a dominant source of backgrounds, in particular from radon daughters plating onto surfaces near or on the detectors (such as the copper detector housing) during the manufacturing process. The three main isotopes of interest are: ^{210}Pb , ^{210}Bi , and ^{210}Po . ^{210}Pb and ^{210}Bi will beta-decay, producing a near-surface and bulk ER background, along with several X-ray lines. ^{210}Po decays via alpha decay, and if the recoiling nucleus is on the surface of the detector it will deposit roughly 100 keV of energy in the detector. If the alpha decay occurs on a nearby surface (such as the detector housing), then the energy of the recoiling ^{206}Pb nucleus will be degraded, resulting in a continuum of nuclear recoils up to about 100 keV [11].

Another source of single-scatter NR events is coherent neutrino scattering from solar neutrinos, produced from the decay of ^8B formed at the end of the pp-III solar fusion reaction chain [11]. While not a dominant background now, in the future the irreducible background of coherent neutrino scattering from solar neutrinos will limit the sensitivity of SuperCDMS (and other direct detection experiments) [11].

4.4.4 Calibration

Calibration of the SuperCDMS SNOLAB detectors will use a scheme similar to what was used at SuperCDMS Soudan. An automated ^{133}Ba source will be used to calibrate the ER energy scale, and a manual ^{252}Cf source will be used to calibrate the detector response to NR. The energy scale of the HV detectors will be calibrated using the ^{71}Ga electron capture lines at 10.37 keV, 1.30 keV, and 0.16 keV for the K-, L-, and M-shell energies respectively [11]. One challenge is the calibration of the silicon HV detectors, which do not have a similar feature; the expected dominant backgrounds in the silicon detectors are tritium and ^{32}Si which are both beta spectrums without easily identifiable features

[11]. The current plan is to introduce a high rate of gammas using the ^{133}Ba source, the distribution of the photons' energy depositions produces well defined Compton steps as they eject electrons from the different orbitals; the expected energies of these steps are: 1.84 keV, 150 eV, and 99 eV [130]. At energies that exceed these binding energies, a step like feature is produced as the probability of scattering off an electron in an inner orbital increases.

4.5 Other devices

The detectors discussed above were developed specifically for SuperCDMS Soudan and SuperCDMS SNOLAB, but the SuperCDMS collaboration has also been working with research and development (R&D) devices to better understand intrinsic detector effects and push to lower thresholds. These devices are much smaller than the detectors used in the main experiments, with masses on the order of grams; the small detector sizes result in good energy resolutions and low thresholds. This section will describe the different R&D devices that have been used, and discuss their relative advantages and disadvantages.

4.5.1 High Voltage detector with eV scale resolution (HVeV)

The CDMS HVeV devices [93] are gram-scale implementations of the SuperCDMS HV detectors, which make use of the NTL effect to measure individual electron-hole pairs. The devices are made of silicon with dimensions of roughly $1 \times 1 \times 0.4 \text{ cm}^3$; one side of the detector is instrumented with QETs while the other side has an electrode used to apply high voltage. A high voltage of approximately 140 V, up to 160 V, is applied to the side of the device with the electrode [93], while the QETs on the other side are read-out in the normal fashion. The large NTL gain allows for individual peaks to be resolved, with a resolution of better than 0.1 electron-hole pairs [93], which corresponds to one to a few electron-hole pairs being drifted across the device.

The spectrum obtained from an HVeV device is shown in Figure 4.7. After the application of the data quality cuts five peaks can be resolved, corresponding to up to five charge carriers being drifted across the detector bias voltage. As charge carriers drift across the crystal they can become trapped by defects or impurities in the crystal lattice, thereby reducing the number of NTL phonons produced. The drifting charge carriers

4.5. OTHER DEVICES

can also liberate a trapped charge in the crystal, in which case an additional charge carrier is drifted across a fraction of the electric potential and the NTL phonon energy is enhanced. These situations, known respectively as charge trapping and impact ionization, are particularly important for low-energy interactions with high bias voltages, where even one additional or fewer drifting charge can result in a significant change in the measured phonon signal. The regions between the peaks in Figure 4.7 are caused by charge-trapping and impact-ionization effects, where a charge carrier is stopped before drifting across the entire potential, or an additional charge carrier is freed and drifted across part of the potential difference. These events between the peaks have been attributed to effects like sub-gap infrared radiation that can excite a trapped charge [93]. Such backgrounds can be mitigated by improved infrared shielding, highlighting the need for a light-tight experimental setup. The ability to resolve individual electron-hole pairs has enabled these devices to set world-leading limits on dark matter [93].

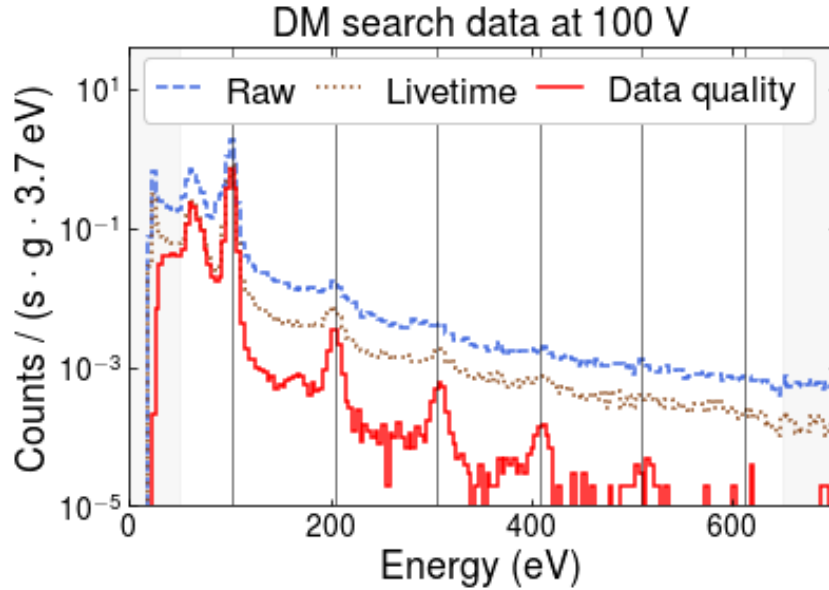


Figure 4.7: Spectrum from a CDMS HVeV device operated at Northwestern University [93]; raw data is shown in dashed blue, dotted brown shows the spectrum after livetime cuts, and solid red shows the cuts after the application of all data quality cuts. A detector bias voltage of 100 V was applied across the detector, and individual peaks can be seen at multiples of 100 eV, up to 500 eV. These peaks correspond to one to five electron-hole pairs being drifted across the detector. The regions between the peaks are caused by charge trapping and impact ionization effects. Image taken from Reference [93].

4.5.2 Cryogenic PhotoDetector (CPD)

CPDs are small scale devices with excellent energy resolution [122]. They are made out of 1 mm thick silicon wafers with a diameter of roughly 76 mm, corresponding to a total weight of about 10.6 g. One side of the detector is instrumented with QETs to measure the athermal phonons, with no applied voltage. The wide coverage of the single QET channel provides fast phonon collection, which minimizes position dependent effects that can degrade the energy resolution. The disadvantage of these devices is their small dynamic range, which is reduced by saturation effects. The baseline energy resolution of the first generation of CPDs was roughly 3-4 eV, although improvements to the second generation's QET mask improves this to almost 2 eV (see Chapter 6). The good energy resolutions and low thresholds of the devices provide excellent sensitivity to low-energy nuclear recoils, which has been used to set world-leading limits on the spin independent WIMP-nucleon cross section at the lowest WIMP masses [122]. This result was achieved on surface with an exposure of less than 10 g·days, which corresponds to a total livetime of less than a day. The main factors that affect the sensitivity of CPDs are the noise environment of the experiment, which can worsen the baseline resolution, and an excess of low-energy events believed to be caused by stress induced microfractures [102, 119]. These low-energy events are generally attributed to how the wafer is supported, so improvements to the clamping/detector holding scheme can mitigate these excesses [14].

Chapter 5

The cryogenic underground test facility

The Cryogenic Underground TEst facility (CUTE) is a detector testing facility, located approximately 2 km underground in SNOLAB, built to operate and test cryogenic detectors for low background particle physics experiments. This chapter describes the motivation and design of the facility, the design of the cold hardware used for the implementation of a SuperCDMS detector tower, the slow-control system used to monitor the environmental conditions at the facility, and some detector testing results from CUTE.

5.1 Motivation

The primary motivation of the CUTE facility is to perform early detector characterization studies with SuperCDMS SNOLAB detectors, although the facility is well suited for other small scale cryogenic experiments that require low levels of background radiation. Testing SuperCDMS SNOLAB detectors at CUTE can confirm that the detectors are in working order after their delivery underground. As the SuperCDMS SNOLAB cryostat will take about a month to reach base temperature, it is impractical to cool down for this long only to realize that something in the setup doesn't work. The preliminary data from SuperCDMS detectors operated in the CUTE facility can also be used to develop new analysis techniques, for instance related to the reconstruction of the energy and position of particle interactions in the detector. New calibration schemes can also be

developed, which is particularly important for the silicon HV detectors. Besides from the tests described above, many other detector characterization studies can be performed, which will greatly improve the ability for results of good scientific quality to be produced once SuperCDMS SNOLAB becomes operational. In addition to testing the SuperCDMS detectors, CUTE also provides an opportunity to test the entire system, which includes the readout electronics and data acquisition software. The CUTE facility can be especially useful for identifying issues in the readout chain that may have otherwise gone unnoticed. For example, the CUTE facility uses the same power-over-ethernet switch as the main experiment. It was noticed that when only a single ethernet port was powered, all of the connected boards remained powered. The cause was identified as a grounding issue on the DCRC boards, and because the problem was identified during tests at CUTE, a solution could be implemented before the main experiment.

5.2 Facility description

Shown in Figure 5.1 is a schematic of the CUTE facility. The main structure of the facility is a cylindrical water tank, roughly 3.6 m in height and 3.6 m in diameter, used to shield the experiment. The cryostat, which houses the detectors, sits on an active suspension system within the well-shielded drywell of the water tank, which is securely mounted to the stainless steel deck. While the cryostat is in the drywell it can be accessed by personnel via the deck in order to connect the cryogenic systems and perform continuity tests of the detectors. The detectors themselves are installed in the cryostat while it sits in the adjacent cleanroom, which is provided with low-radon air from the SNOLAB compressed air system or the SuperCDMS radon filtration system. Once the detectors are installed and the cryostat is closed, it is transferred from the cleanroom to the drywell by an overhead mono-rail crane. On the ground next to the water tank is the gas handling system (GHS), which contains the pumps, pneumatic valves, and the helium tank necessary to operate the dilution refrigerator that cools the CUTE cryostat. Also on the ground next to the water tank are several computers used to control the dilution refrigerator and the detector's readout.

5.2. FACILITY DESCRIPTION

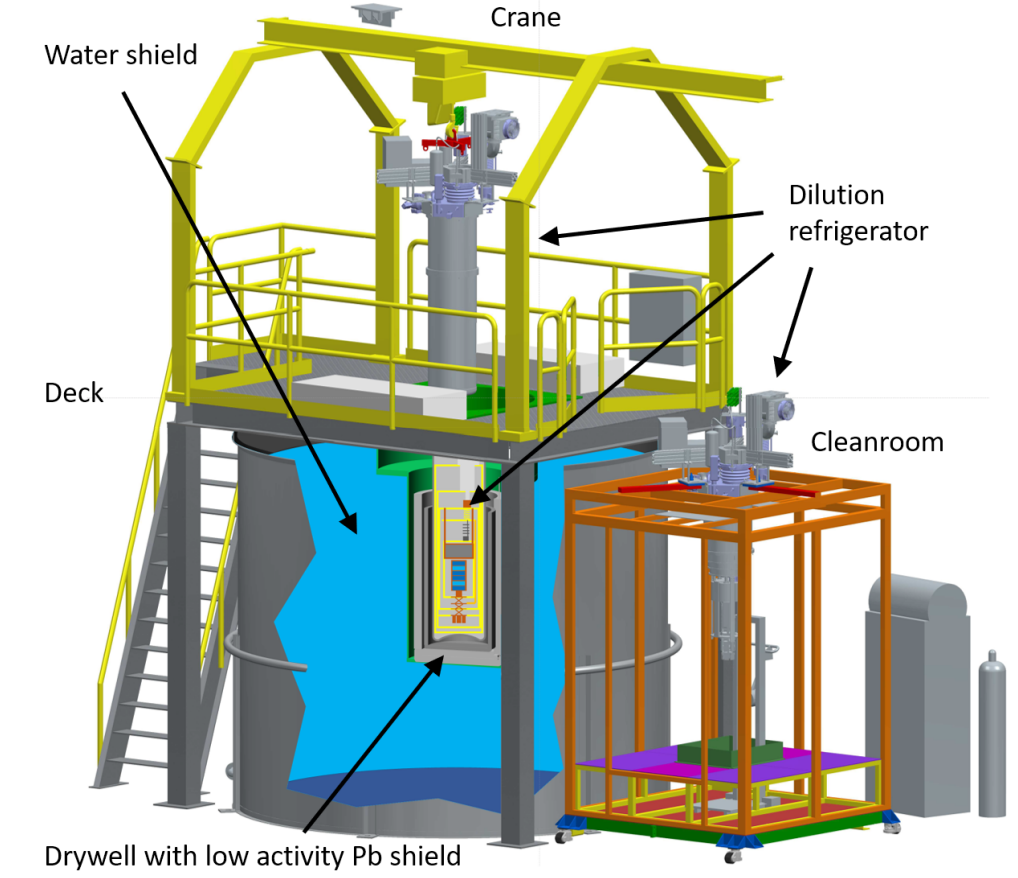


Figure 5.1: Schematic of the CUTE facility. The primary feature of the facility is a large water tank. The experimental payload is mounted to the cryostat in the cleanroom, and the cryostat is moved into the shielded drywell by a monorail crane. The upper deck of the facility can be accessed by personnel using stairs. The gas-handling system and computers are on the ground next to the water tank and cleanroom.

5.2.1 Cryostat

The cryostat is the system that supports the milliKelvin temperatures required to operate the SuperCDMS detectors. This includes the vacuum chamber and the various thermal stages necessary to achieve the base temperature of approximately 13 mK. The CUTE cryostat consists of several thermally isolated stages: the mixing chamber (base temperature) stage, a heat exchanger stage (cold plate), the still (nominally at 0.7 K), and two precooling stages at approximately 4 K and 50 K. These thermal stages are thermally separated by stainless steel or G-10 standoffs, which provide strong mechanical support and low thermal conductivities. To minimize the radiative heat load from the

warmer to colder stages, the three warmest stages are separated with copper cans. In some ways, the cryostat resembles a nesting doll, with the colder stages being nested within the warmer stages. A large stainless steel Dewar surrounds the cold stages of the cryostat, and supports the low pressure of approximately 10^{-7} mbar necessary to reach the base temperature. The vacuum is achieved by a turbo pump backed by a scroll pump, these systems can reach a pressure inside the outer vacuum can (OVC) of approximately $10^{-5} - 10^{-4}$ mbar; the final pressure of 10^{-7} mbar is achieved by cryopumping.

The cryostat is cooled by a dilution refrigerator that uses two systems to reach Kelvin and milliKelvin temperatures, respectively. Traditionally, cooling to the Kelvin range was achieved using liquid cryogens (nitrogen followed by helium), but more modern dilution refrigerators like the one at CUTE use two-stage pulse tube (PT) coolers, which repeatedly compress and expand ^4He into a fixed volume. As no liquid cryogens are required, this system is referred to as a dry dilution refrigerator; these have a number of advantages over their wet counterpart, namely that no additional liquid helium is required to cool down, so there is no dependence on the availability of helium. Pulse tube cryocoolers are commonly used in dry dilution refrigerators over other cryocoolers, such as Gifford-McMahon cryocoolers, as they have no moving parts at the cold end which minimizes vibrations and the need for maintenance. Still, pulse tube coolers introduce a high level of vibrations into the cryostat. To reduce these vibrations the cryostat is mechanically decoupled from the pulse tube, the specific implementation at the CUTE facility is described later in this chapter.

In order to reach milliKelvin temperatures, dilution refrigerators make use of the small enthalpy of mixing between ^3He and ^4He . When a mixture of ^3He and ^4He is cooled below about 0.8 K it separates into two phases: a concentrated phase of almost pure ^3He , and a dilute mixture of a few percent ^3He in superfluid ^4He [131]. Since ^3He is lighter than ^4He , the concentrated phase sits atop the dilute phase. The binding energy of ^3He is larger in ^4He than in ^3He , so ^3He atoms from the pure phase will mix into the dilute phase [131]. This mixing process takes place in the mixing chamber and absorbs energy from the surroundings since the enthalpy of ^3He in the dilute phase is greater than in the concentrated phase [131]. In practice the mixing will not continue

5.2. FACILITY DESCRIPTION

indefinitely, as eventually the concentration of ^3He in the dilute phase will exceed 6.6% [131] and the binding energy of ^3He in the mixture will be smaller than in pure ^3He [131]. To counteract this, a tube is connected from the bottom of the mixing chamber (in the dilute phase) to the still stage, at approximately 0.7-1 K. There the ^3He evaporates out of the ^4He (which remains in its superfluid state) and is recirculated back through the dilution refrigerator. An osmotic pressure gradient is formed by pumping the ^3He out of the still, which draws ^3He out of the dilute phase in the mixing chamber and allows the mixing process to proceed [131]. The circulation in the CUTE dilution refrigerator is driven by a turbomolecular pump backed by a scroll pump; during the precool phase of the cool-down a higher circulation rate is achieved by a helium compressor.

To increase the available space for the experimental payload (i.e. the SuperCDMS detector towers), the experimental stage of the CUTE cryostat is physically separated from the actual mixing chamber stage. The experimental stage is attached to the actual mixing chamber plate by three large copper bars, such that the detectors are mounted roughly 50 cm below the actual mixing chamber stage. As the vertical spacing between the mixing chamber and the still stages of the dilution refrigerator dictate the required quantity of helium mixture for proper cooling, it would be impractical to space these stages to hold the full SuperCDMS detector tower. Instead, by having the experimental stage separate from the mixing chamber stage, the actual dilution refrigerator is designed independently from the detector tower, and only the length of the copper bars that attach the experimental stage need to be properly sized for the detector tower. This separate experimental stage is also beneficial when considering that the CUTE facility could run non-SuperCDMS detectors, in which case the extra space provides flexibility for different types of detectors to be mounted in the cryostat with little concern of space restrictions. There are a few challenges with the separated experimental stage, such as bringing in the detector cabling, but these design considerations will be discussed later on in Section 5.3

5.2.2 Suspension system

Vibrations from the environment or the cryogenic systems (in particular the pulse tube cooler) can be detrimental to the usefulness of the facility. Large levels of vibrations

can produce low-frequency noise that appears in the signal band of the detector, and in extreme cases can impact the performance of the dilution refrigerator. To minimize the transmission of vibrations from outside the cryostat a number of special design features are implemented. The systems in place at CUTE to minimize the level of vibrations are collectively referred to as the suspension system.

The first part of the suspension system that minimizes the transmission of vibrations is the coupling of the pulse tube cooler to the two warmest stages of the cryostat. In a typical dry dilution refrigerator, the cold head of the pulse tube is connected directly to the corresponding stages of the cryostat, generally through a soft copper braid. However, in this case excess vibrations can be transferred from the cold head of the pulse tube down to the experimental stage. In the CUTE cryostat there is no mechanical coupling between the cold head of the pulse tube and the 4 K stage; instead, the cold head is convectively coupled to the cryostat using a portion of the helium mixture as an exchange gas. The helium mixture is circulated through a closed volume that contains the cold head of the pulse tube, thereby introducing a convective coupling to the corresponding thermal stages.

As the pulse tube is thermally coupled to the cryostat by a gaseous coupling, instead of a solid one, the level of vibrations transferred to the cryostat is greatly reduced. The top part of the dilution refrigerator, which supports the pulse tube's rotary valve and the circulation turbo pump, is rigidly connected to the drywell. The cryostat sits on a steel frame supported by three soft, independent dampers on movable stages. The only connection between the upper and lower parts of the cryostat is through soft bellows, which greatly diminish the transmission of vibrations. While the soft dampers that support the lower part of the cryostat do permit some transmission of vibrations through the drywell, this is greatly dampened by the performance of the dampers.

One disadvantage to rigidly connecting the top of the cryostat to the drywell while resting the lower part on the dampers is that the position of the cryostat is sensitive to fluctuations of the atmospheric pressure in SNOLAB. The pressure differential between the interior of the bellows and the air pressure in the lab results in a large lifting force on the cryostat. As the air pressure in SNOLAB regularly varies over the course of the

5.2. FACILITY DESCRIPTION

day, due to the opening/closing of ventilation shafts in the mine, this can result in a time varying force that causes the cryostat to drift upwards slowly. If the cryostat moves too much it is possible that a contact between the cold head of the pulse tube and the cryostat is produced, which would drastically increase the level of vibrations. To combat this effect, an active control system is utilized on each damper of the suspension frame. Position sensors independently monitor the position of the dampers, and regulate it to within some fixed tolerance by driving the movable stages of the dampers with stepper motors, which are controlled by an AVR microcontroller [132].

5.2.3 Shielding

The CUTE facility requires specialized shielding to minimize radioactive backgrounds. Cosmic rays are shielded by the roughly 2 km of overburden (6000 mwe) [106], which blocks almost all particles except for neutrinos. The hadronic component of cosmic rays (such as pions) is entirely blocked by the first several meters of rock [133], however some small fraction of the muonic component can make it through to SNOLAB [106]. The primary backgrounds that are external to the detectors are neutrons and other radioactive decay products in the cavern walls at SNOLAB. The 3.6 m diameter water tank is the primary shielding for detectors at CUTE, and does an excellent job at blocking neutrons and some gamma rays. However, additional lead shielding inside the drywell is required to block high energy photons that can still penetrate through to the detectors. Lead is naturally radioactive; in particular the beta-decay of ^{210}Pb to ^{210}Bi , which eventually beta-decays again to ^{210}Po , produces Bremsstrahlung radiation in the lead shield that would increase the background of the experiment. To mitigate this effect, the lead shielding in the drywell consists of two parts: a larger, outer shield produced with low-activity lead, and a smaller, inner shield made from very-low-activity lead. In this sense the outer lead shield blocks most gamma rays from outside the water tank, while the inner lead shield blocks any additional photons produced in the outer shield. While it would be possible to build just one shield out of the very-low-activity lead, this would be very expensive, and so it is more cost effective to design the shielding in two parts such as this.

The shielding described so far provides coverage on all sides of the experimental payload except the top. A movable polyethylene shield on top of the deck is used to shield neutrons along the vertical axis of the cryostat. Polyethylene is an excellent neutron blocker, and the roughly 20 cm thick slabs of polyethylene enable a near 4π coverage of the detectors from neutrons. The polyethylene shield comes in two equal sized halves, and is attached to wheels that roll in a fixed track on the deck. When the cryostat is being moved from the cleanroom into the drywell the shield can be opened, and once all the hoses and electronics are connected, the shield can be closed again.

To block the detectors from high energy photons along the cryostat's vertical axis and shield from the intrinsically dirty parts of the dilution refrigerator, a 120 kg, 20 cm thick, cylindrical lead plug encased in copper is mounted inside the cryostat. The lead plug is thermally attached to the still stage of the cryostat by three copper rods, and sits directly above the experimental stage (below the mixing chamber stage). Since lead is a superconductor with a transition temperature of roughly 7 K, there are some challenges in cooling it down to 1 K (the nominal temperature of the still stage). When a metal enters the superconducting regime its electrons pair up to form Cooper pairs, which no longer contribute to thermal conduction in the metal, and so all of the thermal conductivity is driven by the phonon system [131, 115]. The reduced thermal conductivity of lead in its superconducting state makes cooling such a massive block of lead quite time intensive; cooling the internal lead to 1 K dominates the cool-down time of the CUTE cryostat.

In addition to the neutron and gamma shielding in the drywell, a mu-metal magnetic shield is used to shield the SQUIDS and detectors from external magnetic fields, such as the Earth's magnetic field. The magnetic shield is cylindrical with an end-cap on the bottom, and is situated between the two external lead shields in the drywell. Such a magnetic shield, with a high magnetic permeability, is required by almost all cryogenic experiments that utilize superconductors. If a large magnetic field were present while the detectors are being cooled down, it could lead to trapped magnetic flux in the superconductors which can be detrimental to the performance of the experiment. In the QETs the trapped flux leads to a decreased collection efficiency, while in the SQUIDS the trapped flux can make the SQUIDS unusable. In the best case the presence of trapped flux can

5.2. FACILITY DESCRIPTION

lead to increased noise, while in the worst case it can prevent the normal operation of the detectors. To release the trapped flux from the superconductor, the temperature of the superconductor must be raised above the superconducting transition temperature; note that the magnetic shield must be around the cryostat during the subsequent cool-down, otherwise there will still be trapped flux after cooling back down.

5.2.4 Calibration systems

Two calibration systems will be available at the CUTE facility in order to perform detector calibrations: a gamma and a neutron calibration system [134]. The gamma calibration system uses a ^{133}Ba source, while the neutron calibration system uses a ^{252}Cf source. Currently, only the gamma calibration system is available, although the neutron system is expected to be installed and commissioned soon. As other experiments at SNOLAB must be given at least a week's notice when neutron sources are moved through the lab, it is inconvenient to store the sources away from the CUTE facility since it is not always well known when calibration data sets should be collected. Instead, both of the radioactive sources are stored in housings made of appropriate shielding materials, and the systems are designed so that the sources can be remotely deployed without any on-site personnel.

The ^{133}Ba source is stored in a lead housing mounted on the drywell, and can be deployed into the shielding by stepper motor system, a schematic of this system can be seen in Figure 5.2. The barium source is contained in a copper encapsulation that is attached to a beaded chain which winds around a sprocket connected to the stepper motor. When the stepper motor is turned it can move the source out of the lead housing along a copper tube into a groove in the lead shielding. When the source is lowered into the inner most lead shield, it moves up or down freely to a user defined height. A remote user can drive the stepper motor by accessing a webpage that sends commands over a websocket to an AVR microcontroller [132].

The principle of the neutron calibration system is comparable to the gamma calibration system, although the implementation is different. In order to reduce the flux of neutrons at neighboring experiments (SuperCDMS and PICO) the source is stored in a polyethylene housing within the water tank. Inside of the polyethylene housing is a tung-

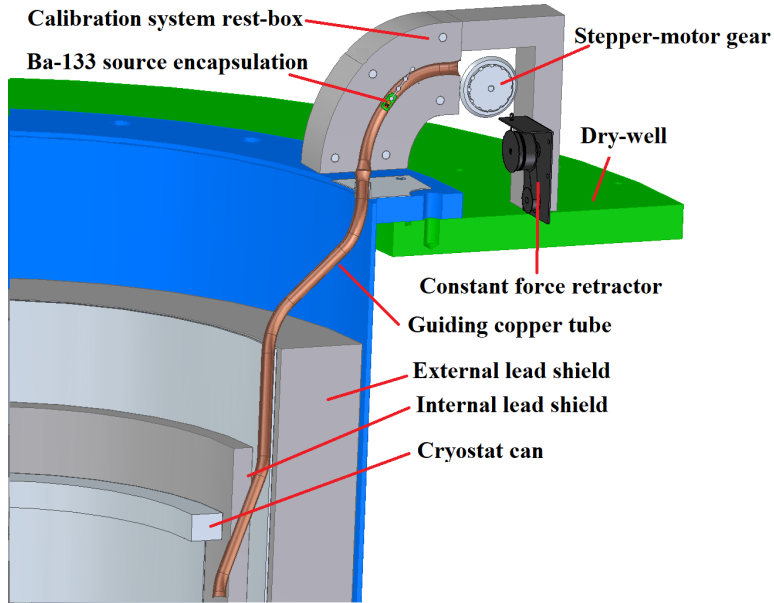


Figure 5.2: Schematic of the gamma calibration system, which lowers an encapsulated ^{133}Ba source into the shielding by a stepper motor system.

sten based alloy called HeavyMet [135] that provides additional shielding to block the flux of gammas produced by the californium source. In a similar fashion to the gamma source, the doubly-encapsulated neutron source is connected to a stepper motor on top of the deck by a beaded chain, and can be driven to a desired position remotely by sending commands to the AVR [132] via the websocket server.

5.3 Detector implementation

Additional hardware was designed and fabricated to mount the SuperCDMS detector tower and connect the readout cables in a way that they can be properly operated. One challenge in cryogenic experiments is minimizing the heat load along the cable connecting the cold detector to the room temperature electronics; this is done by adequately heat-sinking the cable to each stage that it passes through. Improper heat sinking can lead to an excessive heat load on one or more thermal stages, which can negatively impact the fridge performance (prevent the cooling process). The detectors must also be shielded from infrared radiation from the warmer stages of the cryostat, which can produce excess low-energy events in the detector. Gaps in the thermal shielding between different thermal stages of the cryostat can produce an excessive heat load that negatively affects fridge

5.3. DETECTOR IMPLEMENTATION

performance, so any such gaps must be minimized. This section will describe the heat sinking and infrared shielding that was designed and implemented in order for CUTE to properly operate a SuperCDMS detector tower.

5.3.1 Heat sinking

The readout cable for a SuperCDMS detector consists of 100 twisted wire pairs. The cables are passed through a steel elbow into the cryostat via a hermetic vacuum interface board (VIB), designed by the SuperCDMS collaboration, which mates with a flange on the elbow. The warm end of the cable is soldered to a PCB that slots into a connector on the VIB, and the signals are passed through copper traces on the VIB that go to a connector for the DCRC on the opposite (non-vacuum) side of the VIB. The 100 wires per detector represents a significant heat load, so each cable must be adequately heat-sunk to the 50 K and 4 K cryostat stages.

The first heat sink is at the 50 K stage of the cryostat, where the cable is attached to an L-shaped copper bracket that sits on a circular opening in the 50 K plate, which allows the cabling to pass through. To maximize heat transfer from the cable to the plate, the cables are pressed between the bracket and a copper plate in a way that increases the contact area between the cable and the bracket. To minimize gaps in the 50 K layer, the rest of the hole that the cable passes through is covered by a copper plate and any small gaps that remain are covered by copper tape. An image of the 50 K heat sink is shown in Figure 5.3. The readout cable must then be heat-sunk to the 4 K stage of the cryostat before it connects to the detector tower. This is done by affixing it to several points along the outside of the 4 K can with copper shims, an example of one of these shims is shown in Figure 5.4. For heatsinking as well as to avoid damage to the cable, the cable was eventually cast in an epoxy block that fits the cut-out in the flange at the bottom of the 4 K can.

The SuperCDMS detector tower consists of four different stages that are thermally isolated by titanium and graphite standoffs. A low thermal conductivity between stages is required to adequately cool the detectors; otherwise heat would flow from the warmest to the coldest stage and prevent the detectors from reaching the base temperature. If

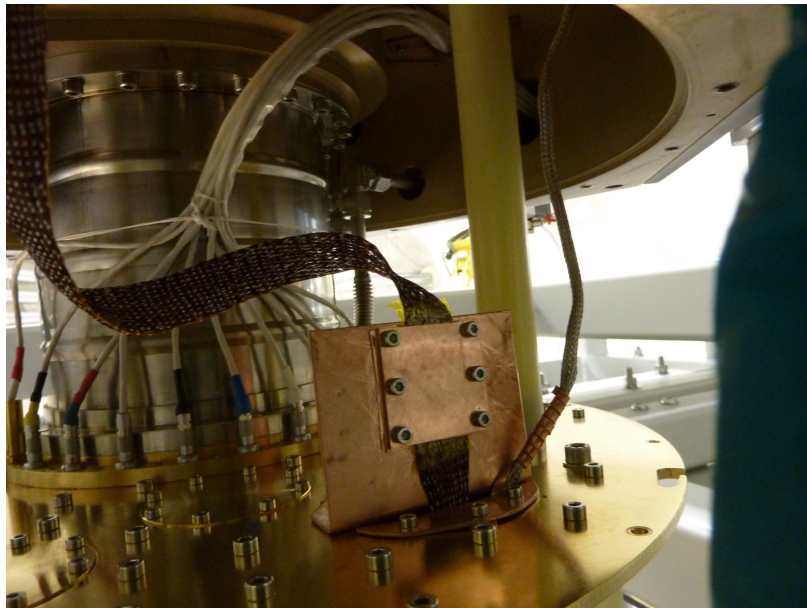


Figure 5.3: Photograph of the 50 K heat sink for the readout cable. The flat readout cable is sandwiched between a copper plate and bracket and the bracket is screwed onto the 50 K plate. Another copper plate is used to close off the large gap that the cable passes through, while any smaller gaps are closed off on the underside by copper tape (not shown). The smaller, circular, silver cable is for additional thermometry.



Figure 5.4: The specialized bottom of the 4 K can that was designed to pass the readout cabling into the 4 K can in a light tight manner. Above the feedthrough in the can is one of the copper shims used to affix and heat sink the cable to the 4 K can.

5.3. DETECTOR IMPLEMENTATION

the various stages of the detector tower are not adequately heat-sunk to the cryostat, there would be a much higher heat load on the detector housing from the cables and the detector would not go superconducting.

The cold plate stage of the tower (nominally at 100 mK), where the shunt resistors are mounted, must be thermally attached to the cold plate stage of the cryostat. Since the tower hangs below the actual dilution refrigerator, a long copper rod is attached to the cold plate stage of the cryostat and reaches down to below the experimental stage. A heat sink made of a highly conductive, flexible copper braid attaches this rod to the respective tower stage, shown in Figure 5.5.

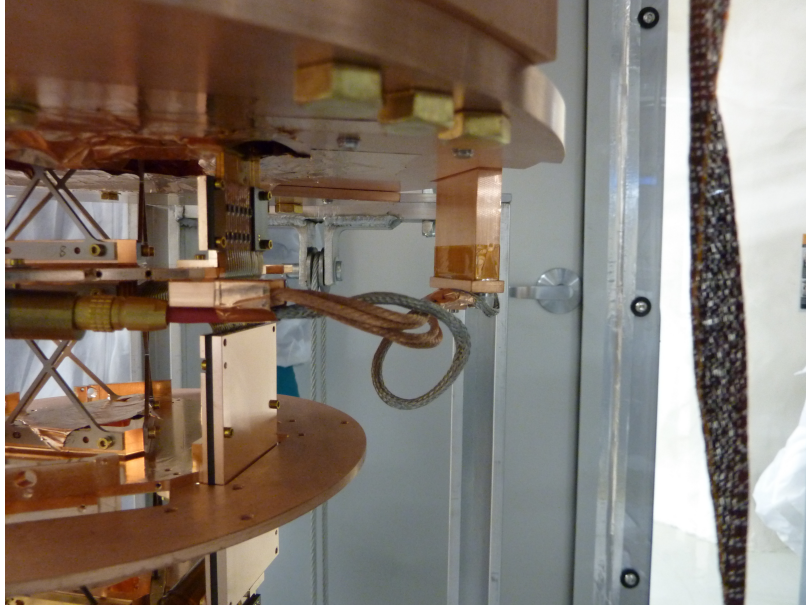


Figure 5.5: The copper heat sink that attaches the 100 mK stage of the tower to the cold plate stage of the cryostat. The braided silver cable is additional thermometry that was installed for a few runs to measure the actual temperature of this stage of the tower.

The still stage of the tower (nominally at 1 K), where the SQUIDs are mounted, is heat-sunk by a thin copper foil membrane. A copper adapter plate is attached to the still stage of the tower, and the membrane is screwed in to this adapter plate by being sandwiched between the adapter and a copper ring. The outer edge of the membrane is screwed to the bottom of the still can (which is open on both ends) in a similar fashion, by being sandwiched between a copper ring and the edge of the still can. A small circular undulation at about the midpoint of the foil accommodates the vertical

differential thermal contractions between the still can and the inner assembly with the tower, which occur during the cool-down and may put undue forces on the tower. The membrane design is multi-functional as it provides heat sinking for the still stage of the tower, and also blocks infrared radiation from the 4 K can. Figure 5.6 shows the membrane, adapter plate, and the copper rings installed in the CUTE cryostat.



Figure 5.6: Photograph of the copper membrane that provides infrared shielding and heat sinking for the still stage of the tower. The membrane attaches to the tower stage by a copper adapter plate and ring, and attaches to the bottom edge of the still can with a larger diameter ring.

The final stage of the tower that must be heat-sunk is the 4 K stage, where the readout cable connector is attached. This heat sinking is accomplished by a thick, braided copper strap, shown in Figure 5.7. In early detector runs at CUTE, the vertical flex cable (VFC) had traces made of NbTi, which has a superconducting transition temperature around 9 K, and the 4 K stage of the tower was connected to the 4 K stage of the cryostat by a rim on the inside of the 4 K can. However, in the final design, the VFC traces consist of Ti with a transition temperature somewhere below 4 K. As a result, the 4 K stage of the tower is now attached to the still stage of the cryostat instead of the 4 K stage, to allow the VFC to go fully superconducting. In this case, a short superconducting readout cable is attached to the VFC and connects to the long readout cable at a bracket attached to

5.3. DETECTOR IMPLEMENTATION

the 4 K can. Figure 5.7 shows an earlier run in which the copper strap connected the 4 K stage of the tower to the 4 K can of the cryostat.



Figure 5.7: Photograph of the copper strap that connects the 4 K stage of the tower to a rim on the inside of the 4 K can. With the new Ti VFC, this strap now connects to the bottom of the still can.

5.3.2 Infrared radiation shielding

The colder stages of the cryostat must be shielded from infrared radiation (IR) to allow for proper cooling and detector operations to occur. The various cans of the cryostat effectively block IR from warmer stages, but additional measures must still be taken. As the lead plug (which is thermally attached to the still) sits below the mixing chamber, the innermost can is on the still stage since a can on the mixing chamber would provide no benefit due to the 1 K heat source. This section details the additional measures taken to block IR from the warmer stages to the detectors.

As the detector readout cables are brought from room temperature to the detector, they pass through a number of heat sinks and gaps. The heat sink at the 50 K stage was described in the previous section, and the only measures taken to block IR here are the copper plate and some copper tape to close off any small gaps. The cable then passes along the outside of the 4 K can where it is connected to the bottom of the detector tower. To achieve this, a special extension for the 4 K can was designed and fabricated,

shown in Figure 5.4. This can has six slots (one for each cable) where the cables can be passed from the outside of the 4 K can to the inside. Each cable passes through a slot in the can, and a three-part copper cover plate goes around the cable and screws into the sidewall of the can, ensuring that any gaps in the slot are effectively covered up; any remaining small gaps can be covered with copper tape.

In the previous section, the foil membrane that provides heat sinking to the still stage of the tower was described (see Figure 5.6). This membrane thermally connects the still stage of the tower to the cryostat, while also closing the can to IR radiation from the 4 K stage. The membrane closes off the main gap at the bottom of the still can, but small gaps along the center of the tower still exist. These gaps could be closed by additional IR shields that form along the cable, but the simpler solution implemented for testing was just to close these gaps with copper tape. Note that copper tape can introduce some additional background (due to the adhesive), but for early tests at CUTE this additional background was considered negligible for the purposes of the tests being performed. When the testing of the actual SuperCDMS SNOLAB towers begins, copper tape can not be used as the adhesive could introduce additional backgrounds, so the extra IR shields will be required.

The last required piece of additional IR shielding is for the experimental stage where the detectors are mounted. As the lead plug sits directly above this stage, having cans on the mixing chamber or cold plate stages would not be effective. Instead, additional shielding is placed directly around the detector tower at the experimental stage to block IR radiation from the still stage. The primary IR shield for the detector towers is a custom made copper can, dubbed the “top hat”, that surrounds the tower and sits directly on the experimental stage, see Figure 5.8. One of the main design constraints was to make this can virtually light tight without using any welding, since welding is an additive process that can introduce contamination which would increase the background seen by the detector. A secondary design constraint along the same lines as the first is to minimize the number of screws used, again in an effort to minimize backgrounds; the final design used only four screws, two at the top and two at the bottom. There are several components in the can: a circular disk at the top, an O-shaped ring at the

5.3. DETECTOR IMPLEMENTATION

bottom, two rods, and two curved sidewalls. The disk, ring, and rods feature grooves that the sidewalls sit in. These grooves ensure that the can is light tight, since light would need to make two 90° turns to get in the can, and also provide additional stability. After putting the sidewalls in the grooves of the rods (to form an open cylinder), they are placed in the appropriate grooves within the top and bottom plates. The rods have threaded holes in their ends to accommodate screws from the top and the bottom. To block IR from gaps at the bottom of the tower, copper tape has been used on the experimental stage. However, this is not suitable for the SuperCDMS SNOLAB towers, so additional IR shields were developed.

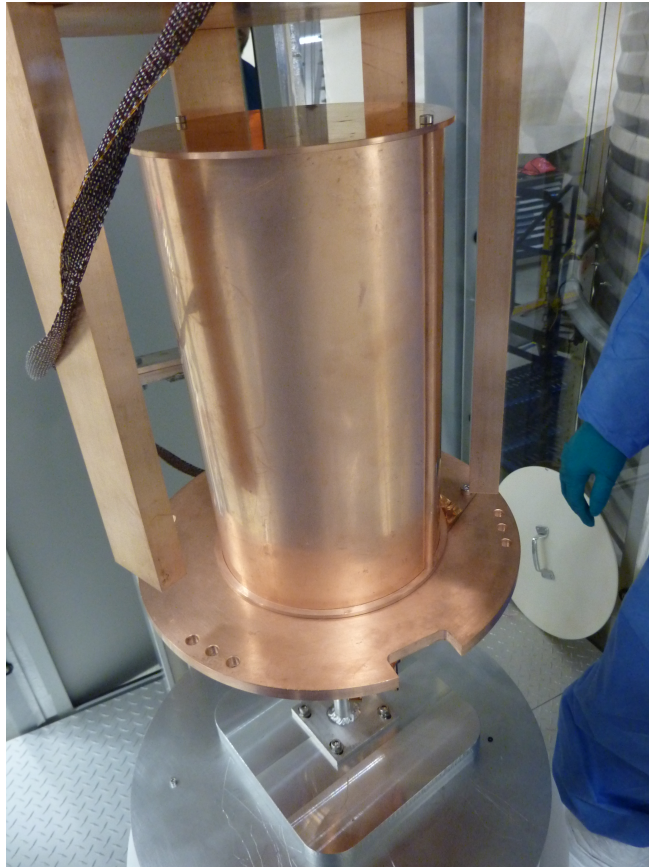


Figure 5.8: The copper can that provides IR shielding for the detector tower, also known as the top hat. This photo is taken during the installation of the detectors into the cryostat, before the experimental stage is attached to the copper bars that connect to the mixing chamber plate.

5.3.3 Detector installation procedure

The previous sections described the various components necessary to mount a Super-CDMS detector tower in the CUTE cryostat; this section will describe the full installation procedure in order of the steps performed from when the cryostat is fully opened in the cleanroom, to the point where it is closed up and connected in the drywell.

Before the procedure can be started, the readout cabling must be installed. This only needs to be performed once (or at least once per cable) as after it is installed it essentially remains in place for the duration of the experiment. The readout cables are first connected to the appropriate slot in the VIB and passed into the cryostat through the vacuum port. The readout cable is then pulled through the elbow of the vacuum port into the main part of the cryostat, and attached to the respective heat sinks and IR shields described above.

With the cabling in place, the experimental payload is prepared on a tower installation stand designed by SLAC. The detectors are mounted to the tower and the vertical flex cables are attached to the corresponding detectors. At this point any other desired components for the experiment, such as radioactive sources or thermometry can also be installed. After the payload is prepared, another stand with a vertical threaded rod is screwed into the bottom of the tower, which can be seen in Figure 5.8. The tower is connected to the experimental stage plate and the assembled top hat is mounted over the tower. The tower and experimental stage is then moved below the cryostat. After the experimental stage is properly aligned, the plate is screwed to the copper bars that connect to the mixing chamber plate, and the copper heat sink from the cold plate is attached to the cold plate stage of the tower. Any gaps along the tower are closed with either copper tape or custom IR shields.

After mounting the experimental stage in the cryostat, the still membrane is connected between the tower and the still can. The readout cable is then passed through the 4 K can extension (the one shown in Figure 5.4) and connected to the tower. After attaching the extension of the 4 K can to the middle part of the 4 K can, excess length of cable is tied up in the volume and the removable lid is screwed onto the extension. The cable feedthrough cover plates are then attached to the can, and the heat sinking shims along

5.4. FACILITY MONITORING

the can are installed. The cable is then taped with copper tape at several points along the can to prevent it from inadvertently touching the 50 K can, which could cause a thermal short.

Once the readout cable is attached, the remaining cans of the cryostat are mounted, and the top of the cleanroom is opened. The monorail crane is attached to a lifting device on top of the cryostat, and the cryostat is slowly lifted from the cleanroom to the drywell. The cryostat is lowered onto the frame of the suspension system, and the top part of the cryostat is bolted to the drywell by long, aluminum beams. During the handling of the cryostat, the top and bottom frames are rigidly connected by steel rods. These rods are essential for moving the cryostat, because otherwise the lower part of the cryostat would not lift with the top part, and the soft bellows would be damaged. When the cryostat is securely mounted in the drywell, these rods are removed to decouple the cryostat from the upper frame and transfer its weight to the suspension system. The cryostat is then connected to the various cryogenic lines, which include: the high and low pressure lines for the pulse tube cooler, the line connecting to the vacuum chamber, two inlet and two outlet lines for the helium mixture, and several smaller tubes that enable the opening and closing of the pneumatic valves on top of the cryostat. Various electrical connections must be made in order to read out the thermometers and pressure gauges, and control the Peltier cooler that cools the circulation turbo pump. Once all of the necessary lines are connected, the cool-down procedure can begin.

5.4 Facility monitoring

The CUTE facility requires a number of independent systems to properly operate a detector. Information from these different systems can provide insight as to how the facility and the detector is performing, and any anomalous behaviours can be identified and corrected. This section will detail the various systems in place for monitoring the performance of the facility, and show how monitoring facility data can improve the performance of the detector.

The time period between where the cryostat is cooled down and warmed up is termed a “Run”; generally each run has unique scientific goals that may require different payloads.

Each run has a unique number that acts as an identifier. Sometimes, an issue with the system makes it so the cryostat's base temperature is never reached, so the cryostat is warmed back up to room temperature, and an attempt is made to repair the cause of the failure. In the case where the cryostat fails to cool down the run number is still incremented.

The primary feedback for controlling the dilution refrigerator comes from the thermometers inside the cryostat, as well as pressure gauges and a flow meter in the helium mixture circuit. Inside the cryostat are six thermometers that work at low temperatures; two on the mixing chamber stage and one on each of the other four thermal stages. The three thermometers on the warmest stages (50 K, 4 K, and still) and one on the mixing chamber function well at all temperatures, while the thermometer on the cold plate and the other one on the mixing chamber stage only give accurate readings for temperatures below about 20 K. The full-scale thermometers provide important information as to how the cool down is proceeding. At low temperatures the still temperature is an important indicator of fridge performance. Later on in this thesis, evidence for how the still temperature can affect the detector readout will be shown.

The helium mixture in the tank is moved into the dilution refrigerator by a scroll pump and a compressor, which sit on the ground in the gas handling system (GHS). The mixture enters the fridge through one of two inlets, the fast- or normal-injection line. The normal-injection line passes through Joule-Thomson impedances in the fridge that liquefy the helium, while the fast-injection line bypasses these impedances to maximize helium flow through the fridge. The fast-injection line is used only during the cool down, while the normal-injection line is used during low temperature operations. During the cool down, helium is circulated out of the dilution refrigerator through the fast-pumping line using a compressor, while at low temperatures the helium is circulated out of the still line by a turbo pump. The turbo pump sits directly on top of the cryostat and is connected to the still line, through which the exchange gas for the pulse tube cooler flows; the fast-pumping line is a separate line altogether.

The pressure gauges in the GHS and on top of the cryostat provide invaluable information to the operation of the dilution refrigerator. A dilution refrigerator can be cooled

5.4. FACILITY MONITORING

down without thermometers, but the same cannot be said for pressure gauges. There are six pressure gauges in the helium circuit line: P1- at the inlet of the turbo pump (outlet of the still), K3- at the outlet of the scroll pump, K4- at the outlet of the compressor (inlet to the cold trap), and K5- at the inlet to the fridge (outlet of the cold trap). The vacuum line pumps the OVC out through a turbo pump backed by a scroll pump; the auxiliary turbo pump is on top of the cryostat, while the auxiliary pump is in the GHS. There are two pressure gauges in the vacuum line before and after the turbo pump: P2- at the outlet, and P3- at the inlet (inside the OVC).

The thermometers and pressure gauges are logged to a MySQL database at a fixed interval by the fridge control software, along with the status of the pumps and valves, and the room temperature. On a separate interval, the pulse tube cooler's helium compressor inlet and outlet pressures, and cooling water temperatures are logged to a MySQL database by a Python program. Another Python program logs the temperature and output percentage of the Peltier cooler to a MySQL database. A few other variables related to the fridge performance are logged, but these are less important.

Data from the suspension system are also logged. On each Labjack damper there are two sensors: an optical sensor that measures the distance between the suspension system and the drywell, and a Hall effect sensor that tracks the position of the movable stage. When the optical sensor reads a value of 0 mm, this corresponds to the suspension system resting on the drywell, while a value of 2 mm corresponds to the maximum value (due to a physical stop on the Labjacks). The output of the sensors are calibrated against in-situ measurements of the relevant distances. Automatic control software reads the values from the optical sensors, and ensures that it falls within a user defined range, typically 1 ± 0.04 mm, if the value is low (high) the stepper motors will drive up (down). The Hall effect sensors are useful for ensuring that the motors do not attempt to drive past the maximum or minimum of the Labjack stage. Data from an atmospheric pressure sensor logs the air pressure in the lab. Later on we will see how the lab air pressure affects the position of the suspension system. The tank water level is also recorded, although this typically remains constant at 3 m, so the logging interval is very sparse. The main function of the water level sensor is to ensure that there is no leak in the tank.

In addition to the other sensors there are also two accelerometers (one single axis and one triaxial) mounted on top of the cryostat. Output from these sensors are not logged continually, but they have been used to check the level of vibrations on top of the fridge.

The fridge/facility database and the respective servers are based on a Mac mini computer, which is mounted in the electronics rack on top of the deck. By running these relatively resource intensive processes on a separate computer from the fridge control computer there is little risk of slowing down the fridge control, which is essential for the operation of the dilution refrigerator.

5.4.1 Facility control interface

A webpage, see Figure 5.9, was designed and developed to interface with the various facility subsystems, including the suspension system, gamma calibration system, pulse tube cryocompressor, and the Peltier cooler. The webpage also displays the temperatures throughout the cryostat, the pressures measured throughout the dilution refrigerator, and the status of the various pumps and compressors used by the dilution refrigerator. The code for the webpage is written using the React framework, and interfaces with the respective servers using websockets and PHP calls. Web frameworks, such as React, are beneficial to front end development as they allow for components (such as buttons) to be easily programmed without needing to program the necessary HTML, CSS, and Javascript. Moreover, web frameworks provide a highly modular code base, allowing components to be easily added or removed.

The ribbon at the top of the page displays the current value of several relevant facility parameters, such as the lab air pressure, the weight of liquid nitrogen in the cold trap Dewar, the temperature of the cooling water provided to the pulse tube’s cryocompressor, along with several other parameters. The panel labelled “Fridge Diagram” updates the status of the dilution refrigerator every minute, which includes the temperatures of the thermometers, the pressures of the gauges in the system, and the status of the various pumps and compressors. The panel labelled “Suspension System” allows for the suspension system’s active control functionality to be turned on and off, and allows for the position of the active control to be changed. It also includes functionality to perform

5.4. FACILITY MONITORING

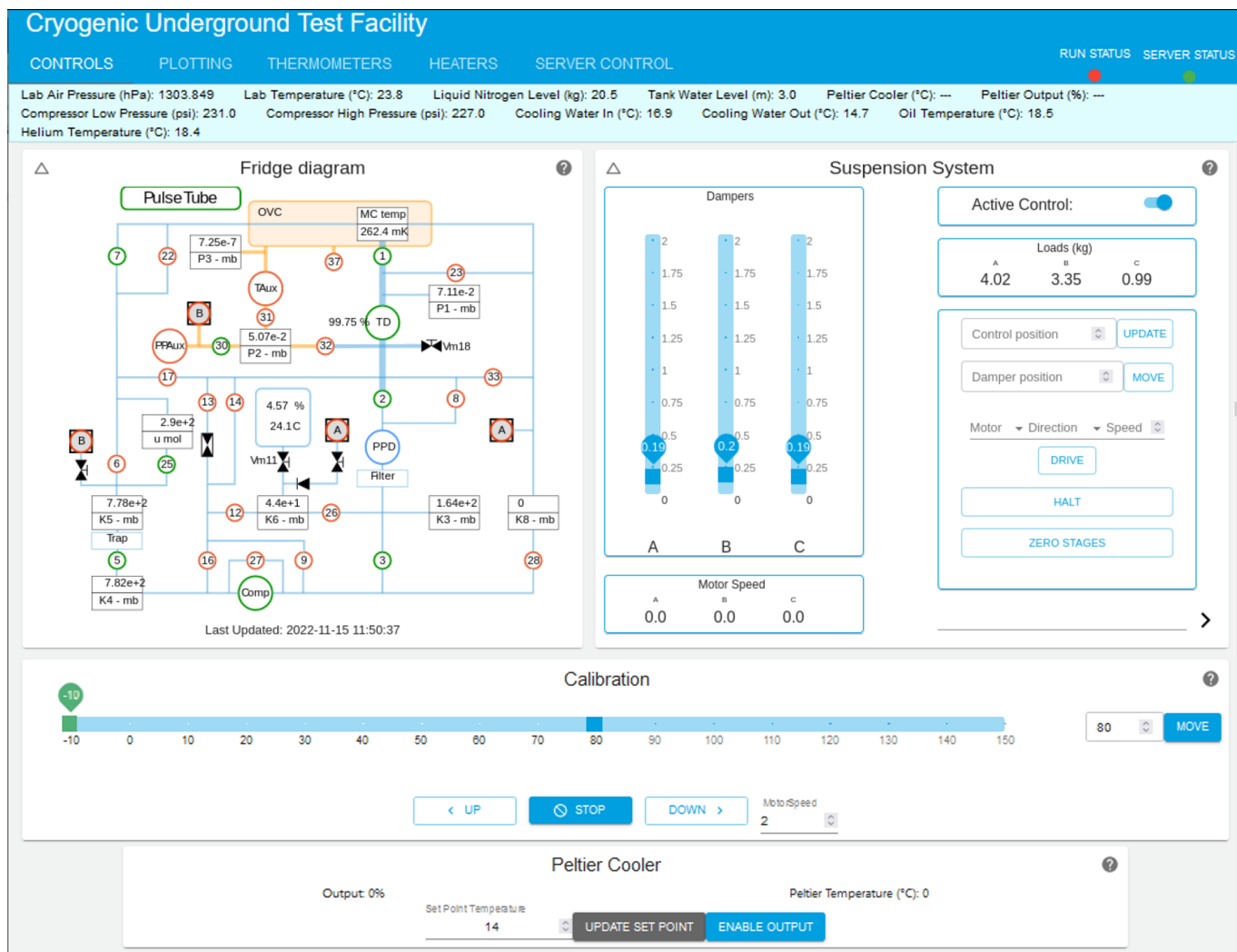


Figure 5.9: Webpage used to monitor and control the CUTE facility’s various slow control systems. This page displays the status of the dilution refrigerator, and provides the ability to control the suspension system, gamma calibration system, and the Peltier cooler. The upper ribbon gives real time information about important facility metrics, such as the weight of the liquid nitrogen in the cold trap and the cooling water temperature. The other tabs on the page provide functionality to plot the relevant slow control information, control the thermometry and heater settings, and start and stop the relevant servers.

fine control operations on the suspension system, such as driving each motor at different speeds and stopping the motors entirely. A command line interface at the bottom of the panel provides even more control over the system, allowing for commands to be sent directly to the AVR microcontroller, which can be useful for troubleshooting. The “Calibration” panel displays and controls the position of the gamma calibration source. When

the neutron source is installed, a similar panel will be developed to control the neutron source [134]. The “Peltier Cooler” panel displays the temperature of the Peltier cooler and the output percentage of the controller (which corresponds to the cooling power of the Peltier cooler). Additionally, the set point temperature of the Peltier cooler can be changed through this panel.

There are other tabs at the top of the page that provide additional functionality. The “Plotting” tab provides the ability to plot and download data from the facility database. The “Thermometers” tab allows for the settings (currents, excitation voltages, etc.) of the dilution refrigerator’s thermometers to be changed; the “Heaters” tab allows for the settings (set point, maximum power, etc.) of the dilution refrigerator’s heaters to be changed. The “Server Control” tab allows for the various servers to be started and stopped through the webpage, which allows for people with no knowledge of the underlying server codes to restart them in case the computer crashes.

5.4.2 Suspension system performance

When the pressure in the still line drops below about 1 mbar (and the top and bottom of the cryostat are decoupled), the suspension system will begin to float due to the atmospheric pressure in the lab. The lifting force experienced on the cryostat is so great in fact, that additional weights must be added to the suspension frame to compensate. This section will demonstrate that the suspension system functions as expected, and effectively reduces the level of vibrations from the pulse tube cooler.

During an overnight data taking series in Run 16, the detector trigger rate drastically increased. Over this same period, it was noticed that the still temperature increased, as did the lab air pressure. By examining the positions of the Labjack dampers, it was confirmed that the suspension system had moved up during the same time, in correlation with the lab air pressure. As the suspension frame moves upwards with increasing lab air pressure, the volume in the still pumping line decreases, thereby reducing the still pumping efficiency and leading to an increase in still temperature. Figure 5.10 shows the change in detector trigger rate in the top panel, the change in the lab air pressure in the middle panel, and the deviation of the three dampers from their nominal control point.

5.4. FACILITY MONITORING

As seen in this plot, the trigger rate was correlated with the lab air pressure. The bottom panel shows the damper position increasing with the lab air pressure; the solid black lines represent the control limit for the suspension system, which was only engaged once. After this series the active control limit was decreased from ± 0.1 mm to ± 0.04 mm, and data was taken again. Figure 5.11 shows the same information as in Figure 5.10 for a series taken immediately after adjusting the control limit, using identical detector and trigger settings as in the previous series. In this series, a similar change in lab air pressure as the previous series is seen, although now as the damper position increases it exceeds the control limit and is adjusted much more frequently. The net effect of this is a much more stable detector trigger rate. These plots show that the suspension system functions in an expected way, while also demonstrating the importance of the automatic control.

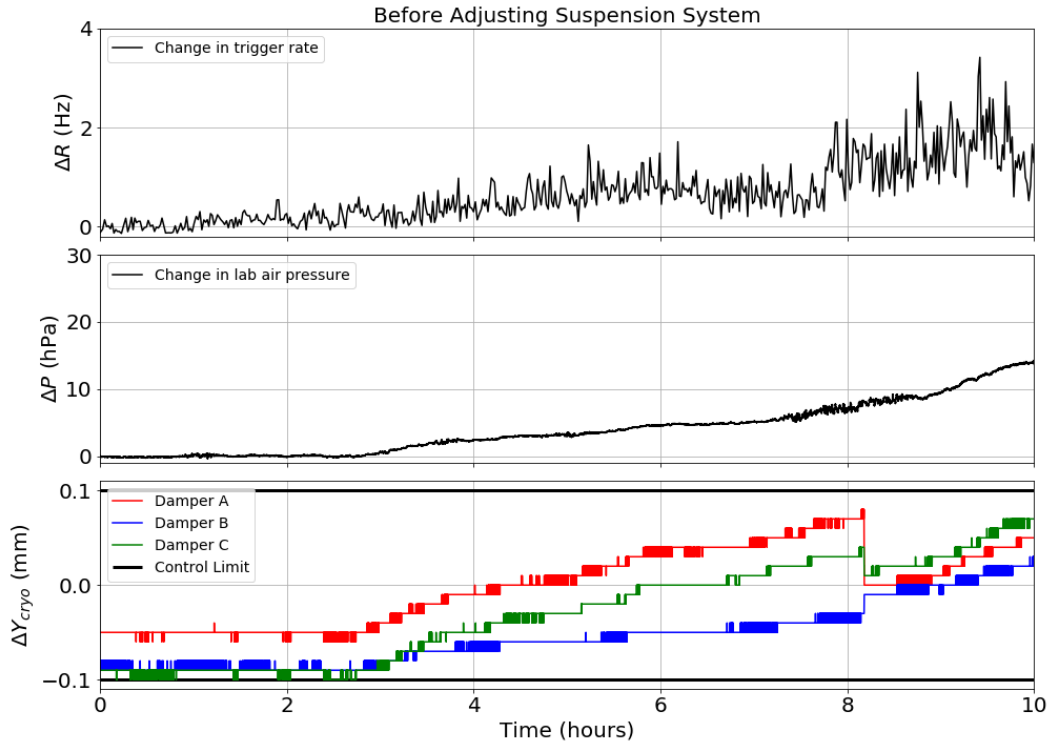


Figure 5.10: Plot showing data taken from a series in Run 16. The top panel shows the change in detector trigger rate, the middle panel shows the change in the lab air pressure, and the bottom panel shows the deviation of the dampers from their nominal position. The horizontal black lines in the bottom panel indicate the upper and lower limits of the active control utilized by the suspension system. The dampers move upwards with the increasing lab air pressure, of which the net effect is an increase in the detector trigger rate.

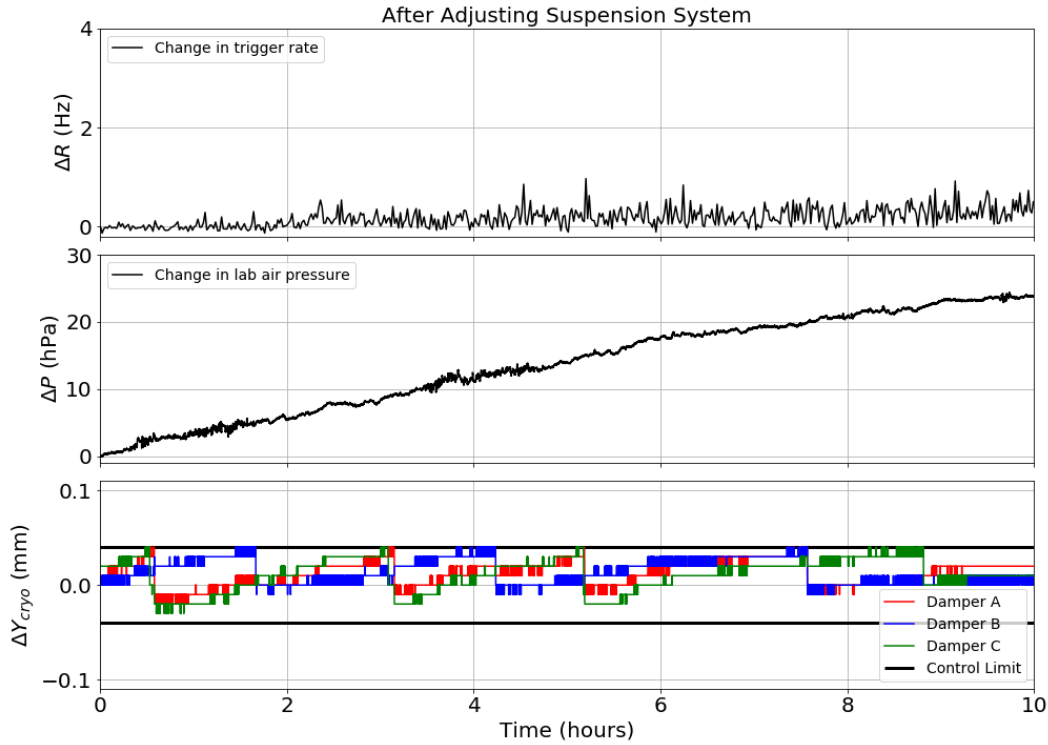


Figure 5.11: Plot showing the same information as Figure 5.10 for the series taken after the suspension system’s control limit was adjusted. In this series, the damper positions still increase with the increasing lab air pressure, although now they engage the upper control limit more frequently, resulting in a more stable detector trigger rate.

While it is reassuring to see that the suspension system functions as expected, it is also important to check that it minimizes vibrations from the pulse tube cooler (otherwise it is an unnecessary complication). Excess low-frequency noise (relative to previous results measured at SLAC) was observed in the detectors during an earlier CUTE run (Run 9), the cause for which was later identified as a contact between the cold head of the pulse tube cooler and the cryostat. After the cold head of the pulse tube was realigned between runs, the excess low-frequency noise disappeared. In a later fridge run, some tests were performed to attempt to reproduce this effect and validate the vibration isolation of the suspension system.

During Run 18 the detector payload included several different detectors, including a 1.4 kg germanium, prototype SuperCDMS HV detector (called G124), and a TES sensor on thin silicon substrate (hereafter referred to as the chip). Due to its large mass, G124 is more sensitive to vibrations, whereas the TES chip is less sensitive to vibrations but

5.4. FACILITY MONITORING

more sensitive to electronic noise from cross-talk or electromagnetic radiation. Figure 5.12 shows the low-frequency noise spectrum for both of these detectors collected during several different series. Two states of the suspension system were defined for the purpose of these tests: a balanced state that corresponds to the normal operation of the suspension system, and a coupled state, where the suspension system was intentionally tilted to introduce a contact between the cold head and the cryostat. The contact was verified by measuring the resistance between the pulse tube cold head and the cryostat; with no contact the resistance should measure as open. When the pulse tube was stopped for a particular series, it was restarted after the data collection finished and the fridge was allowed to cool down for about 30 minutes before proceeding with the next test. For each test, the detectors were triggered randomly in order to obtain detector noise, and any traces that included pulses (which are present due to the background) were removed from this analysis. The power spectral density (PSD) of each noise trace was computed, and all PSDs (for traces with no pulses) were averaged to produce an averaged PSD, which corresponds to the average noise in the detector for that particular series.

The noise during the normal operating condition of the detectors, with the suspension system balanced and the pulse tube cooler running serves as a baseline noise to compare to. The data taken with the suspension system balanced and the pulse tube cooler switched off show the noise level drop a bit, which can be attributed to acoustic coupling of the rotary valve to the lower part of the cryostat, but overall the noise is very similar to the baseline situation. These two measurements provide good evidence that the suspension system minimizes vibrations coupled to the cryostat from the pulse tube cooler. The data taken while the pulse tube is running and the pulse tube cooler's cold head is directly coupled to the cryostat shows a large increase in the noise level (similar to what was observed in the earlier runs when the pulse tube was misaligned). After switching the pulse tube cooler off (with the suspension system still tilted), the noise returns to a similar level as in the balanced configuration. This supports the idea that most of the noise in the coupled configuration with the pulse tube on is related to the pulse tube cooler being on, however it is not clear whether this noise is related to the excess vibrations, or from some electromagnetic interference introduced by the coupling of the cold head.

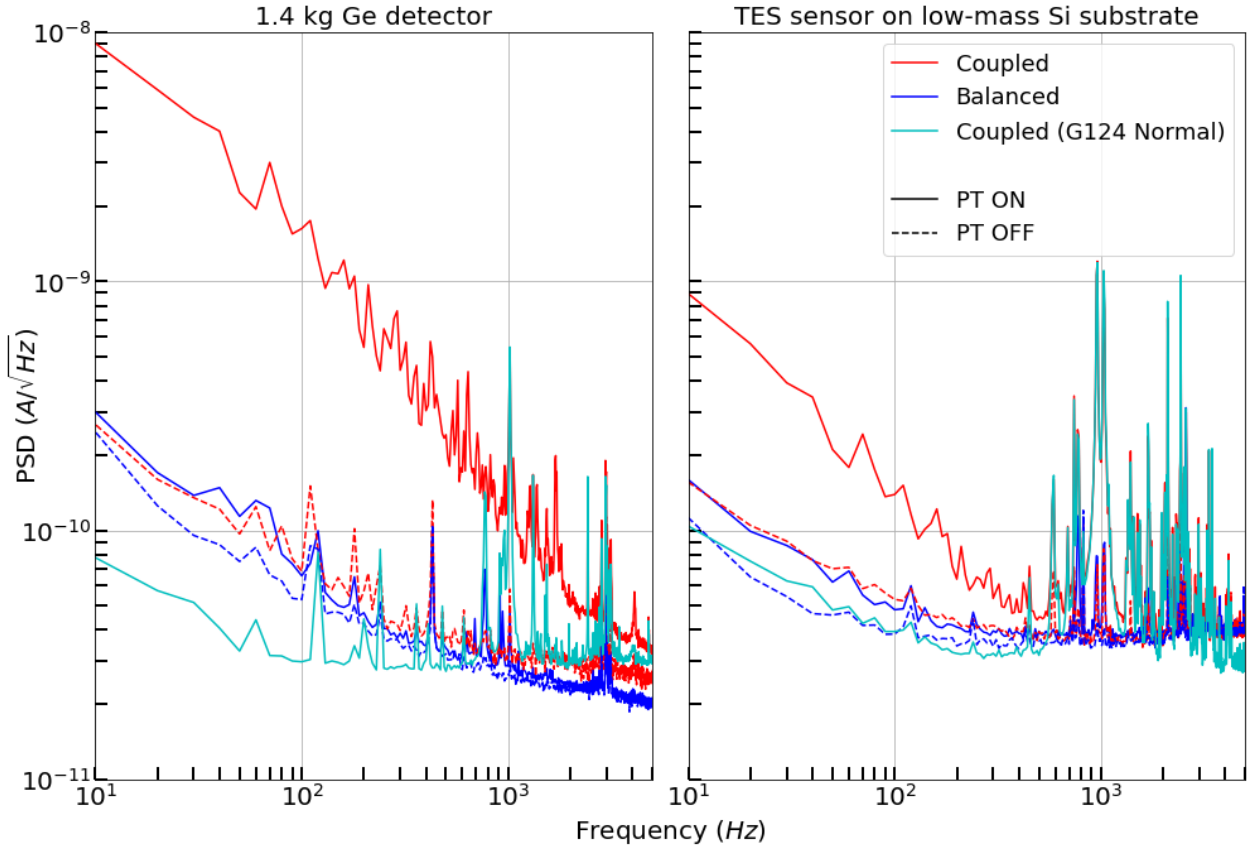


Figure 5.12: Noise measurements collected from one of the G124 channels and the TES chip. This plot demonstrates a large increase in the level of vibrations when the suspension system is misaligned, indicating that excess vibrational noise couples in to the detector from the pulse tube cooler. Additional details can be found in the text.

To determine whether the excess noise in the coupled state with the pulse tube running is related to vibrations or electromagnetic interference, the TES chip was operated in the superconducting transition regime (as it was for all other series), while G124 was operated in the normal conducting regime (instead of the superconducting transition). By putting G124 normal, the TES resistance is drastically increased, meaning most of the current goes through the shunt resistor instead of the sensor (recall the circuit diagram in Figure 4.1). This has the effect of removing any TES power noise in G124, such that the only remaining noise is related to the down-stream electronics. The overall noise level in G124 dropped well below the baseline noise, which is expected as the power noise in the TES is generally larger than the electronic noise. The more interesting result of this test is the

5.4. FACILITY MONITORING

noise in the TES chip drops below the baseline noise level for frequencies under 500 Hz. As the only thing that changed is G124 being normal, the excess noise seen before in the chip was cross-talk from G124. For frequencies in the range of (0.5-3) kHz, when the pulse tube is running and coupled to the cryostat the noise levels in the chip are almost identical when G124 is normal and in transition. Additionally, the noise level in the chip is noticeably higher than in G124 when G124 is normal, so this cannot be cross-talk. This means that the higher frequency noise is related to electromagnetic interference introduced by the coupling of the cold head and the cryostat, while the lower frequency noise is related to the physical vibrations. The low-frequency noise seen in the large detector is therefore related to vibrations, while the higher frequency noise seen in the small device is related to electromagnetic interference; since the big detector is more sensitive to vibrations, and the small chip is more sensitive to electromagnetic noise. The excess low-frequency noise below 500 Hz seen by the chip while G124 was in transition is a result of cross-talk, whereas when G124 was in the normal conducting regime this noise disappeared. This supports the conclusion that a coupling of the pulse tube cooler can introduce vibration noise at frequencies below 500 Hz, while also introducing some higher frequency noise related to electromagnetic interference.

5.4.3 Purge gas vibrations

Radon in the volume between the cryostat's vacuum can and the inner lead shield can introduce an additional background of gamma rays that is not blocked by the lead shields. To mitigate this background, compressed air from the surface (which is low in radon) is forced into the drywell with a plastic hose; this compressed air is known as the purge gas. The optimization of this flow rate is non-trivial, as from a backgrounds perspective the flow rate should be set maximally (higher flow rate means less radon in the volume), while from a noise perspective the flow should be as low as possible (higher flow rate means higher vibration induced noise). To determine the optimal flow rate, measurements of the radon concentration in this volume of air were performed using a RAD7 device [134, 136], while the level of vibrations were measured with accelerometers. Figure 5.13 shows measurements from one of the channels (aligned with the vertical axis of the

cryostat) of the triaxial accelerometer on top of the cryostat, for different flow rates of the purge gas. Between 0 and 30 cfh (cubic feet per hour), the increase in the level of vibrations is modest, while at 90 cfh the increase is quite pronounced. The RAD7 measurements indicated that a flow rate of 15 cfh kept the radon concentration in the drywell sufficiently low [134], so this setting was used going forward. Measurements of transition noise in T5Z2 demonstrates that the increase in low-frequency noise with the flow rate set to 15 cfh is minimal, see Figure 5.14.

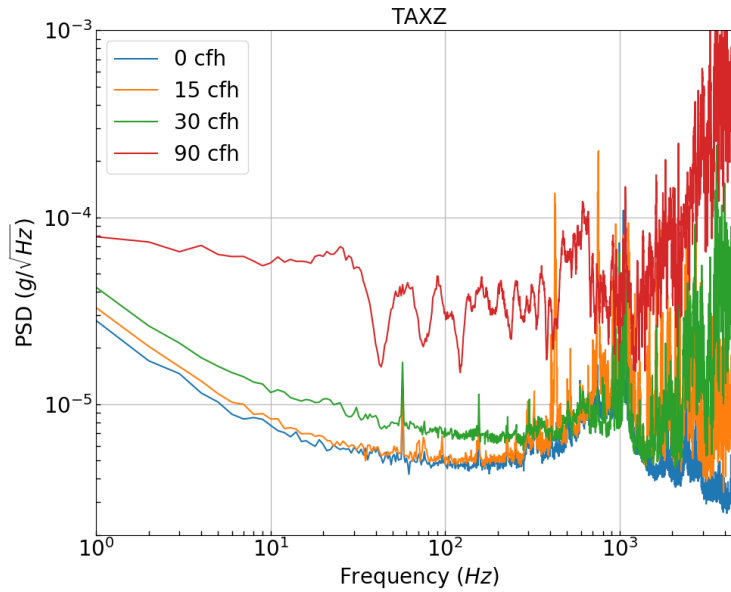


Figure 5.13: Measurements from the vertical axis of the triaxial accelerometer collected during CUTE Run 14 while increasing the flow rate of the purge gas. A large increase in the level of vibrations is seen when the flow rate is set to 90 cfh, but for lower flow rates the increase in the level of vibrations is minimal. Below a flow rate of 15 cfh, the intrinsic noise of the accelerometer and charge amplifier limits the sensitivity of the measurement.

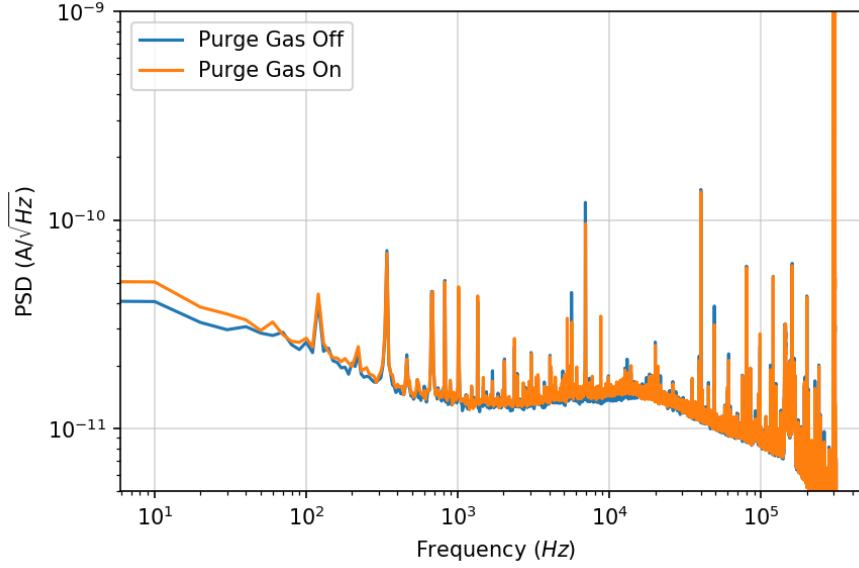


Figure 5.14: Transition noise data in T5Z2 collected during CUTE Run 11 with the drywell purge gas off and the drywell purge gas set to 15 cfh. A minimal increase in the low-frequency noise is observed with the purge gas turned on.

5.5 Detector testing

Since the CUTE facility has been operational, several different SuperCDMS detectors have been tested including: a SuperCDMS Soudan iZIP detector (T5Z2), a SuperCDMS SNOLAB HV Ge prototype (G124), a SuperCDMS SNOLAB HV Si prototype (S101), two wide area athermal silicon detectors (PD2 and CPD), and several TES chips. This section will discuss some of the key results obtained with these devices.

5.5.1 Noise studies

The low-frequency noise was greatly improved after realignment of the pulse tube; however, there was still excess noise that impacted detector performance. Dedicated tests were performed to identify these noise sources by systematically turning on and off different devices in the facility, to determine the effect on the detectors.

The controller for the Peltier cooler was found to produce peaks in the noise spectrum at 330 Hz and the corresponding harmonics [134], see Figure 5.15. This noise arises from the controller’s pulse-width modulation (PWM), which regulates the cooling power of the Peltier cooler. Attempts were made to reduce this noise by shielding the wires connecting

the controller to the Peltier cooler with aluminum foil, and electrically isolating the Peltier cooler from the turbo pump with kapton tape. In addition to these modifications, the controller was moved outside of the drywell to the slow-control rack on top of the deck. Unfortunately, none of these changes were sufficient to completely remove the 330 Hz noise, so a new switching power supply was obtained with a switching frequency much higher than the signal band of interest. The analog output of the new power supply was connected directly to the Peltier cooler, and the output voltage was set by the controller; this provided a direct signal to the Peltier cooler and bypassed the 330 Hz PWM. By making these changes the Peltier cooler could still be operated as normal (which is required for sustained operation of the turbo pump and by extension the dilution refrigerator), but the 330 Hz noise was removed.

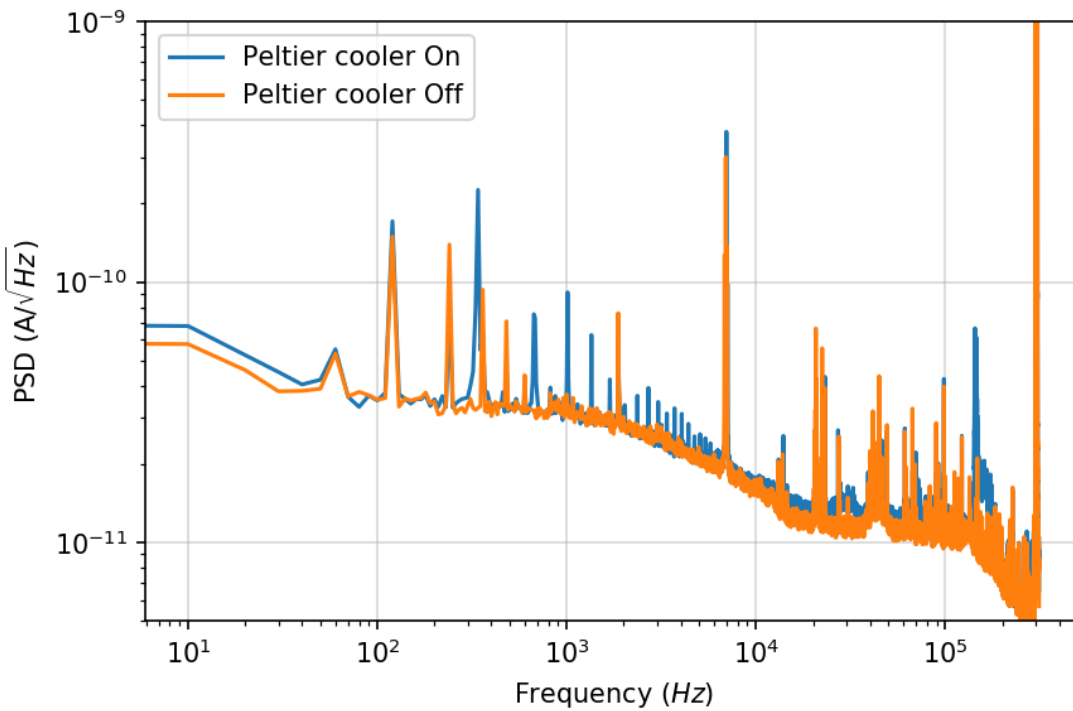


Figure 5.15: Noise in a TES chip that demonstrates the effect of turning the original Peltier cooler controller on and off. When the controller is running 330 Hz and harmonics peaks are present in the noise spectrum. This noise was eventually mitigated by changing the power supply that runs the Peltier cooler.

Additional studies were performed to determine the effects on the detectors of radio frequency waves produced by transmitting devices. A laptop was placed on the deck near

5.5. DETECTOR TESTING

the DCRCs to observe the impact of Wi-Fi and Bluetooth on the detectors, these effects can be seen in Figures 5.16 and 5.17 respectively. In each of these configurations there is excess noise below 4 kHz, with some broad harmonic structures. For the case of a laptop connected to Wi-Fi near the DCRCs, there is evidence of a peak at about 1 kHz with a first harmonic at about 2 kHz. For a laptop emitting a Bluetooth signal there are several peaks, with the lowest peak at just below about 300 Hz. As a result of this excess noise, no devices connected to Bluetooth or Wi-Fi are used on the deck while data taking is happening, and preferably none are used near the facility.

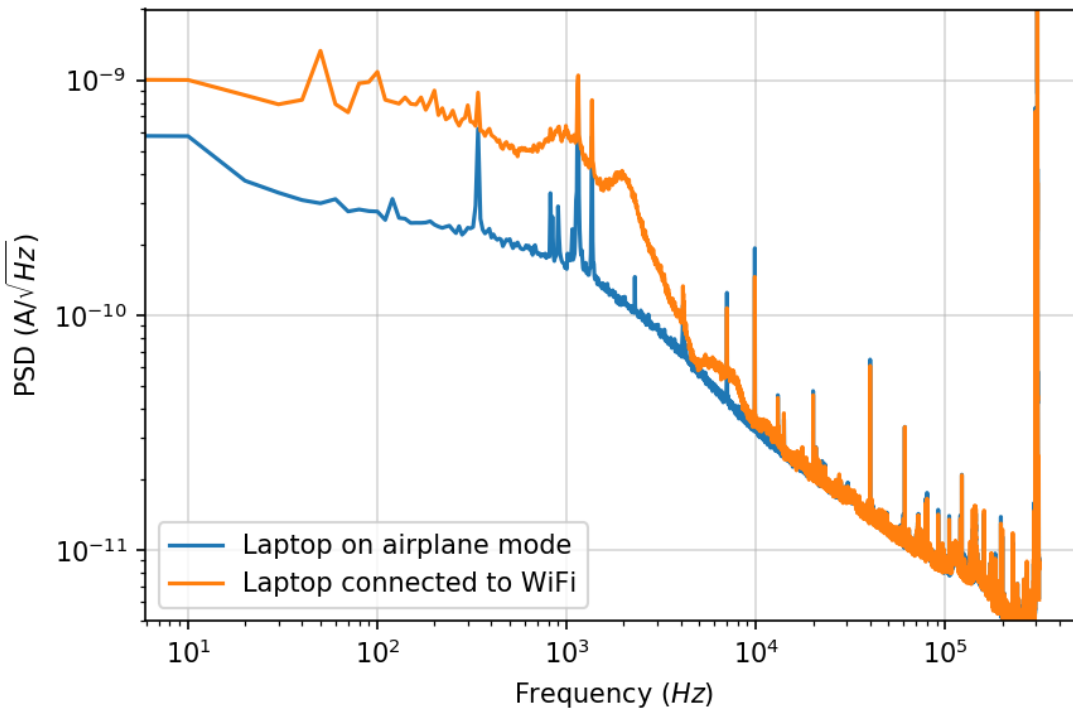


Figure 5.16: Comparison of the noise in PD2 when a device on the deck of CUTE is connected to Wi-Fi and when the same device is in airplane mode.

5.5.2 Detector calibration results

To interpret the output of a detector in a meaningful way, the energy scale must be calibrated. Output traces from the detector that issue a trigger are stored on disk and processed afterwards to extract the energy of the event. The triggered events are processed using an optimal filter algorithm, which uses a pulse template and the detector noise to return an optimal filter amplitude, an estimate of the deposited energy. More

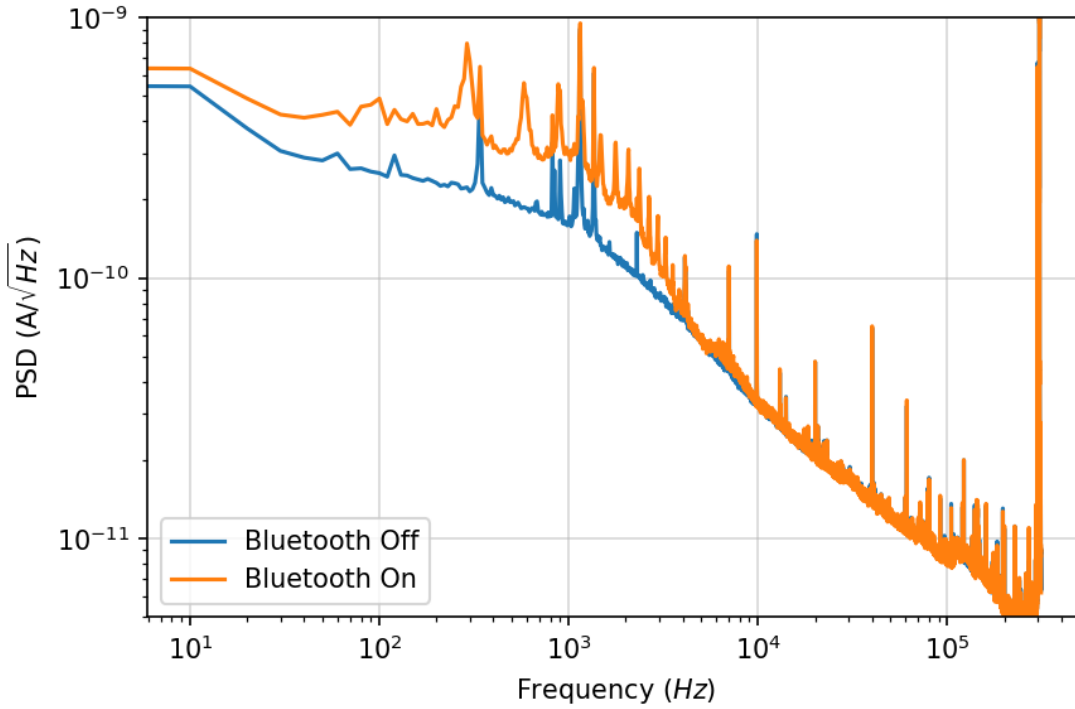


Figure 5.17: Comparison of the noise in PD2 when a device on the deck of CUTE has its Bluetooth capabilities enabled and when the same device is in airplane mode.

details on the optimal filter can be found in Appendix A. The optimal must be converted to units of energy (typically electron-Volts) in order to perform physics analyses (such as a dark matter search). This section discusses the calibration of T5Z2 using the barium source, and the calibration of PD2 using the sensor response.

5.5.2.1 Barium calibration

During CUTE Run 11, the barium source was deployed to multiple positions to calibrate T5Z2 and determine the optimal source position for calibrations, these spectra can be seen in Figure 5.18. Comparing the spectra of the series where the source is deployed to series where the source was not deployed (labelled as Control in Figure 5.18), there is a clear difference in the rate measured by the detector, indicating that the source is in fact present. In several of the series there is a clear peak featured in the spectra, which is the 356 keV gamma from the barium source; it is this peak that the optimal filter amplitude of T5Z2 is calibrated to in Figure 5.18. The calibration is done by fitting the position of the optimal filter amplitude of this peak with a Gaussian, and aligning to the known

5.5. DETECTOR TESTING

energy of 356 keV. A secondary peak can be seen at 303 keV. In some series where the barium source was deployed this feature is not present, in particular when the source position was at 67 cm and 140 cm. The absence of the peak in this case is attributed to the fact that the barium source was further away from the detector, and the probability of Compton scattering increased, distorting the spectrum. For the series at 67 cm, the source is above the internal lead shield, whereas for the series taken with the source at 140 cm the source is far below the detector and the flux seen at the detector is reduced. For the series taken with the source between 75 cm and 95 cm, the peak is displayed prominently.

When the source is very close to the detector tower (as in the series taken with the source at 95 cm), the rate of events seen at the detector is high. The high rate means that pileup in the detector (multiple pulses in a single trace) is very likely, which impacts the ability of the optimal filter to accurately reconstruct the energy. To account for this effect the trace length was shortened, which means that the probability of pileup is reduced. For these reasons it was determined that the optimal source position for barium calibrations was 83 cm, so the source is not blocked by the internal lead shield but still far enough away that the rate of pileup is low.

5.5.2.2 Sensor response

Smaller devices, such as the CPD detectors, cannot be calibrated with the barium source since the energy of the gammas is too high and well above where saturation effects set in (resulting in the detector response being non-linear). In order to calibrate these small devices an ^{55}Fe source near the detector can be used, this will introduce a constant background to the detector that limits the sensitivity of a dark matter search. For runs where no ^{55}Fe source is present it is still possible to calibrate the detector based on the current measured in the sensor, although this requires knowledge of the sensor's collection efficiency. Measuring the collection efficiency of CPD detectors will be discussed in more detail in Chapter 6, but for the purpose of this chapter the collection efficiency will be assumed to be the same value measured at SLAC before PD2 was shipped to SNOLAB, which was 13.3% [102, 119].

The rate of change of power in the sensor with respect to current as a function of

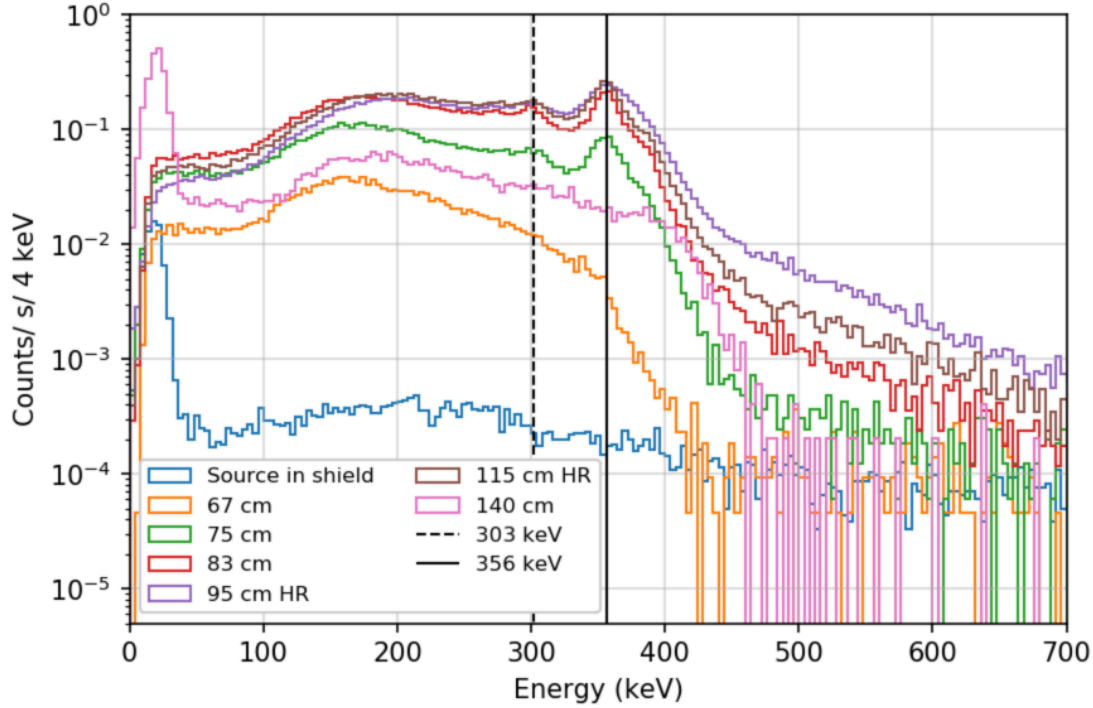


Figure 5.18: Spectra measured by T5Z2 with the barium source deployed at various positions. The spectra have been calibrated to the position of the 356 keV peak from the series where the source was deployed to 83 cm. When the source was very close to the detector, the high rate (HR) from the barium source required the trace length to be shortened to reduce the number of traces with multiple pulses; the optimal filter used for these series was different, but the overall calibration factor remained the same.

frequency $\frac{\partial P}{\partial I}(\omega)$ can be measured, and it is this quantity that is important for determining the calibration factor. The energy deposited in the phonon system E is related to the energy absorbed in the sensor by the collection efficiency ϵ as:

$$E = \frac{\partial P}{\partial I}(0) \frac{A}{\epsilon}, \quad (5.1)$$

where A is the integral of the pulse with dimensions of current times time. Equation 5.1 has dimensions of energy and can be converted between units of Joules or electron-Volts for convenience.

The change in power with current is related to the change of voltage with current $\frac{\partial V}{\partial I}(\omega)$, which represents the sensor's response to an input current. In practice $\frac{\partial V}{\partial I}(\omega)$ is measured by sending a small square wave signal through the TES while measuring the response. The Fourier transform of the measured traces is computed and divided by

5.5. DETECTOR TESTING

the Fourier transform of the square wave signal, which yields the complex admittance of the sensor as a function of frequency $\frac{\partial I}{\partial V}(\omega)$. This quantity can then be inverted and multiplied by the current through the sensor to yield the parameter of interest $\frac{\partial P}{\partial I}(\omega)$, which is the relevant quantity in determining the calibration factor. More details on the electrical response of a TES can be found in Reference [124].

With this method devices such as PD2 can be calibrated without an actual source, so long as the collection efficiency is known. The collection efficiency depends primarily on the detector geometry, sensor layout, and thermal connection to the bath, none of which are expected to change after shipment to SNOLAB, so this method provides a reasonable estimate of the energy calibration. It should be noted that between Runs 15 and 18, the tightness of the clamps holding PD2 were adjusted, which did affect the collection efficiency. However, the assumption of a collection efficiency of 13.3% as measured at SLAC should be reasonable for runs performed before the clamps were adjusted.

5.5.3 Background measurements

Measurements of the facility background were performed in CUTE Run 15 using a SuperCDMS Soudan iZIP detector (T5Z2); T5Z2 is well suited for this measurement due to its large size (relative to CPD devices) and high dynamic range. During the operation of T5Z2, there was no way to read-out the charge channels and only a limited number of phonon channels were functional. Even without using the charge channels, energy depositions up to 1 MeV can be measured (although no particle discrimination is possible). Figure 5.19 shows the spectrum measured by T5Z2 in the absence of a radioactive source. The upper limit on the facility background rate predicted by simulations in the energy range of 1-1000 keV is 5.6 ± 0.6 events/(keV·kg·day) [137], while the rate measured in Run 15 over the same energy range was 5.81 ± 0.04 events/(keV·kg·day). The broad peak structure around 200 keV in the measured spectrum is a result of ^{210}Pb contamination in the shielding, and materials surrounding the detector [137].

5.5.4 Low-energy excess

While testing PD2 in Run 11, it was noticed that the low-energy rate was well above the facility background level. Specifically, the expected facility background rate is on

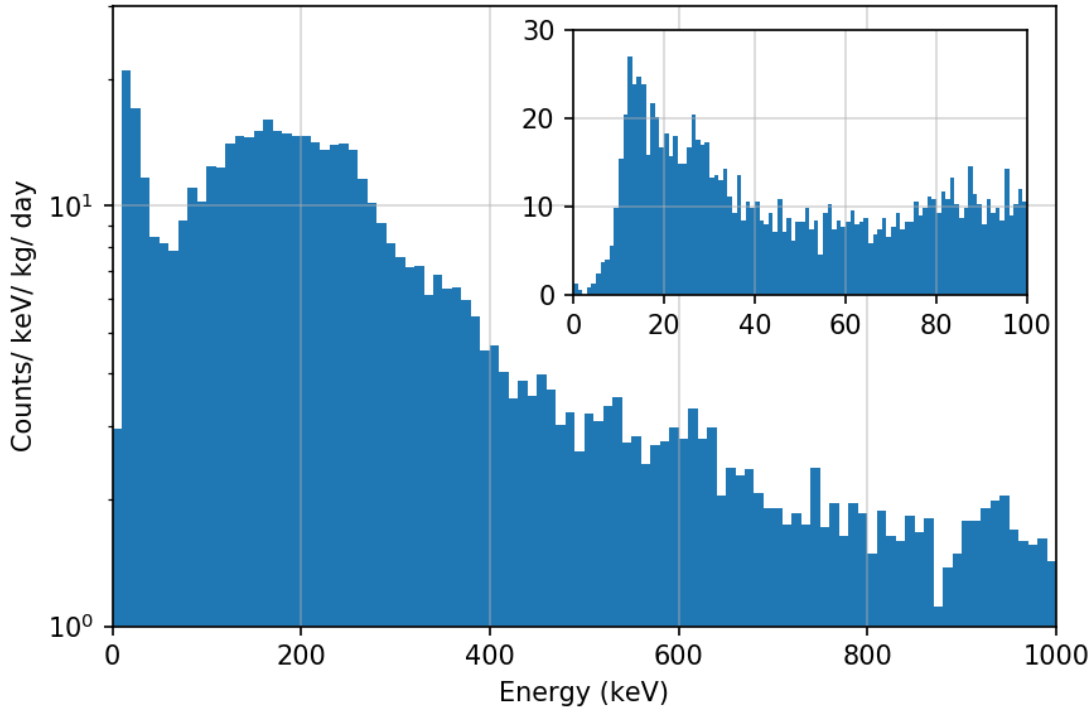


Figure 5.19: CUTE facility background as measured by a Soudan iZIP detector (T5Z2) during Run 15. Note that only three phonon channels were operational for this measurement. The inset of the plot shows the low-energy portion of the same spectrum with finer binning.

the order of 10 events/(keV·kg·day), while the rate observed in PD2 in the energy range of 1-7 keV was on the order of 10^4 events/(keV·kg·day) (as seen in Figure 5.20). The explanation of this excess low-energy rate is still not fully understood, but it has been attributed to stress in the detector and defects in the crystal producing low-energy events [102, 119]. The excess rate is even more pronounced in the lower energy bins, where the excess rate is on the order of 10^6 events/(keV·kg·day) in the energy range of 50-100 eV, as seen in Figure 5.20. This excess low-energy rate decreases over time, falling more rapidly while the detector is cold. It should be noted that in the last few series of Run 15, the rate in the lowest energy bin (50-100 eV) is seen to increase. This effect is understood to be caused by a change in the noise conditions and trigger settings, which allowed more triggers caused by noise to contaminate the data stream [102].

Over the course of Run 11 (and Runs 14/15) the low-energy excess was seen to decay with a half life on the order of 1-2 weeks (depending on the energy bin used in the fit), see Figure 5.20. While such a background (that decays exponentially) would normally be

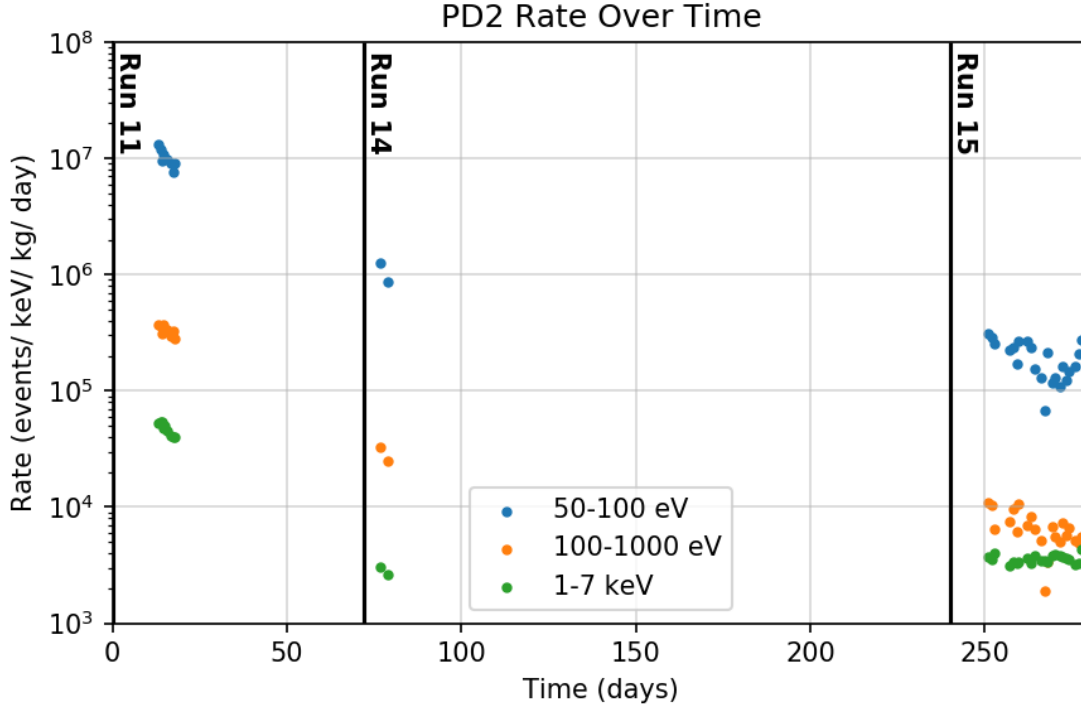


Figure 5.20: Measurements of the excess low-energy rate in PD2 taken over multiple series throughout CUTE Runs 11, 14, and 15. The rate in each of the energy bins (indicated by the legend) are several orders of magnitude above the expected facility rate, indicating some effect that is intrinsic to PD2. The cooldown of Runs 12 and 13 both failed, so no data could be collected. The large gap between Runs 14 and 15 is due to broken thermometry that was sent away to be fixed.

attributed to some radioactive background, this decay rate is not well explained by any known isotopes. One isotope with a half-life in the right range is ^{32}P , which has a half-life of roughly 14 days and an atomic weight close to that of silicon. However, only trace amounts of ^{32}P are expected to occur in natural materials, and the ^{32}P should be in secular equilibrium with ^{32}Si , so the decay rate should be identical to that of ^{32}Si . If the excess rate were in fact due to some radiogenic background, then the rate would be expected to drop between runs (as the isotope decays). However, the excess rate appears to only decay exponentially while the substrate is cold. This evidence suggests that it could be related to thermal stress in the crystal being released at low (Kelvin scale) temperatures, and the stress stays at that level while at room temperature [102]. Measurements of the resistance and bias power of the sensor throughout the same run indicated that the sensor

was staying at the same point in the transition, which would affect the calibration of the energy scale [102]. This excludes the theory that the decreasing low-energy rate can be explained by changes in the operating point of the TES.

Dedicated experiments have shown that additional stress from the mounting configuration can introduce a low-energy excess much like the one seen at CUTE [14]. An understanding of this stress-related background could be beneficial to fields other than the direct detection of dark matter with cryogenic detectors. It has been shown that phonon-mediated breaking of Cooper pairs in superconducting qubits can result in reduced coherence times [138], an essential operating feature of superconducting qubits [15].

Despite the excess low-energy background in the CPD operated at CUTE, relatively competitive constraints on spin-independent dark matter were set using the data collected during Run 15 [102]. While the constraints would be quite strong, there was a large degree of uncertainty on them due to the lack of a proper calibration. The most competitive constraints came from the data collected during Run 15, as the excess low-energy rate was much lower than in Runs 11 and 14 (see Figure 5.20). The ^{55}Fe source that could be used for calibration was only installed in Run 18, and unfortunately between Run 15 and Run 16 the clamps that held the silicon wafer in the detector housing were loosened in an attempt to study the low-energy excess. When the clamps were loosened, the phonon collection efficiency (as determined by the ^{55}Fe calibration) appeared to worsen relative to the measurements at SLAC [122, 102]. The lack of a proper determination of the collection efficiency introduced a large uncertainty into the constraints. Assuming that the collection efficiency remained the same as measured at SLAC, the CPD device operated during CUTE Run 15 could set constraints on dark matter masses as low as $100\text{-}150\text{ GeV}/c^2$, which would have been world leading for the cross-sections excluded [102].

Chapter 6

A search for low-mass thermal relics using a wide area photon detector

Building on the success of the original Cryogenic PhotoDetector (CPD) [122], a second version of the detector (CPDv2) was produced. The overall geometry of the new device remained the same as the original, a 1 mm thick silicon wafer with a diameter of roughly 76 mm, but an improved QET mask was implemented. The new mask was designed to increase the phonon collection efficiency relative to the original CPD. As a consequence, the new device has an improved energy resolution (lower energy threshold) but is more sensitive to saturation effects in the QET. The wafer is clamped to a copper housing with sapphire balls between phosphor-bronze clamps and the silicon substrate; an image of the detector can be seen in Figure 6.1. To calibrate the detector, an ^{55}Fe source irradiates several layers of aluminum foil, which produces several monoenergetic peaks. This chapter will discuss the analysis of the data collected from CPDv2 while operated at the University of Massachusetts Amherst (UMass), in a collaboration between SuperCDMS and SPICE/HeRALD [139]. These data are used to set new constraints on the interaction cross section of low-mass thermal relics.

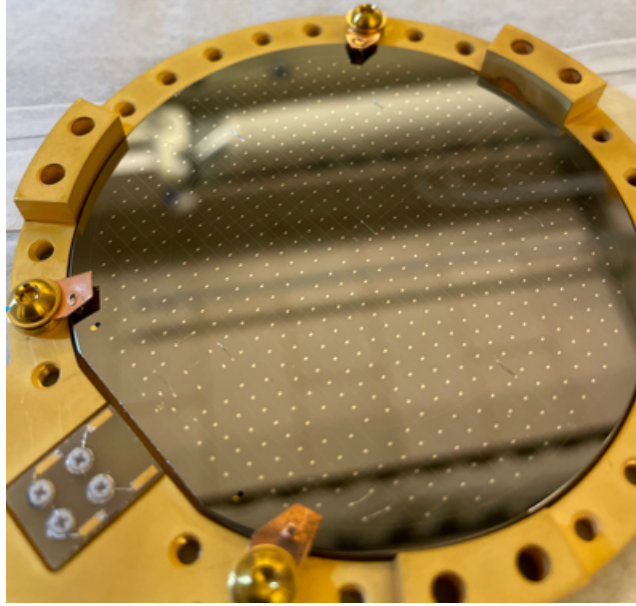


Figure 6.1: Photograph of the second version of the CPD, which was operated at the University of Massachusetts Amherst. An array of QETs is instrumented on a silicon wafer, which is clamped to a copper housing.

6.1 Description of the data

A dilution refrigerator cooled CPDv2 in two separate fridge runs, Runs 26 and 28. Excess noise in Run 26 worsened the energy resolution of the device, so only data from Run 28 were used in the analysis. During Run 28, data were collected during two periods, Run 28a and Run 28b, separated by one week. The readout current of the single QET channel was amplified using a Magnicon SQUID and digitized by a National Instruments DAQ. The output from the detector was logged continually, in contrast to the typical data acquisition used by SuperCDMS where only triggered events are recorded due to the large amount of data. By storing the output from the detector continuously, new data analysis techniques can be implemented that are not available if only individual triggered events are stored.

To avoid introducing an unintentional bias to the analysis, a “blinded” analysis is performed; the analysis is developed based on a fraction of the data set, while the rest is not looked at (blinded) until the analysis is fully developed. The final results are then extracted from applying the analysis to the blinded data. The continuous raw data

6.1. DESCRIPTION OF THE DATA

stream is stored in files that each contain five minutes of data. The analysis is developed using only the odd files (the unblinded data), and the final result is determined with the even files (the blinded data).

Events are extracted from the continuous data stream using a trigger algorithm based on an optimal filter [140, 102]. The optimal filter takes a signal and a noise template to extract the best fit amplitude of a signal in the presence of stationary noise by deweighting frequency components in the raw trace caused by noise; more details on the optimal filter can be found in Appendix A. One limitation of the optimal filter is its sensitivity to variations in the pulse shape, which can result in non-linearities in the reconstructed energy due to saturation effects in the detector. The trigger algorithm applies the optimal filter to the continuous data and registers events that have an optimal filter amplitude (OFamp) above a certain trigger threshold (in this analysis 4.5 times the resolution of the optimal filter). A coincidence window is used to determine the optimum trigger point by selecting the event time with the best fit (lowest χ^2) as the event time. In addition to these physics triggers, triggers are issued randomly throughout the data series in order to understand the noise in the detector.

The signal (pulse) template is generated by first collecting triggered events using a best-guess of the pulse template taken from another cryogenic run. A first data-based template is then created by averaging all triggered events with a reconstructed energy between 40 and 100 eV, while removing outlier events that had a poor fit quality. To improve the alignment of the pulse with the trigger point, the triggering algorithm is repeated using the rough template and the new set of low-energy events are averaged again. Note that the alignment of the trigger has a minimal impact on the estimate of the energy absorbed in the QET, the main purpose of re-triggering is to improve the alignment of the pulses. After producing a template from the average of low-energy traces of the second trigger run (that should be reasonably well aligned), a fit to the averaged template is done with an analytical form of the pulse template. The functional form of a pulse in the detector depends on the transfer function of the TES and the dynamics of phonons in the substrate. The analytical form of the pulse used here is obtained by

CHAPTER 6. A SEARCH FOR LOW-MASS THERMAL RELICS USING A WIDE AREA PHOTON DETECTOR

convolving a two-pole TES response with a three-pole phonon response, leading to:

$$f_{pulse}(t) = (1 - e^{-t/\tau_{TES1}})e^{-t/\tau_{TES2}} \times (1 - e^{-t/\tau_{ph1}})(ae^{-t/\tau_{ph2}} + (1 - a)e^{-t/\tau_{ph3}}). \quad (6.1)$$

The best-fit values of the TES response times are given by $\tau_{TES1} = 19.6 \pm 0.1\mu s$, and $\tau_{TES2} = 19.8 \pm 0.4\mu s$. The phonon response times are given by $\tau_{ph1} = 19.8 \pm 0.4\mu s$, $\tau_{ph2} = 90.9 \pm 0.4\mu s$, and $\tau_{ph3} = 452.3 \pm 1.8\mu s$. The parameter $a = 0.2520 \pm 0.0002$ is dimensionless. The units of this template are not important as the optimal filter requires a template normalized such that the maximum value is equal to 1; the units of the optimal filter are inherited from the noise PSD. Figure 6.2 compares the analytical pulse template to the averaged template, showing good agreement. Using the analytical form of the pulse template improves the energy resolution of the optimal filter.

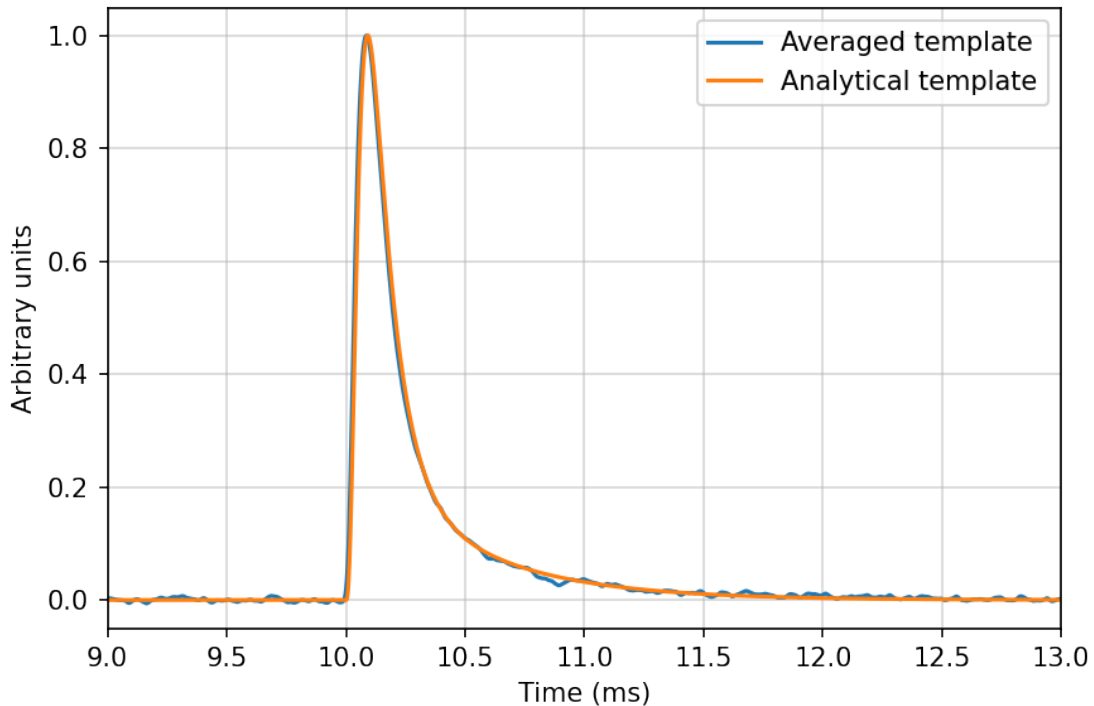


Figure 6.2: Comparison of the analytical form of the signal (pulse) template used by the optimal filter to the empirical pulse template (determined by averaging low-energy events). The template is determined by averaging many low-energy pulses and then fitting an analytical form of the pulse to the averaged template. Shown here is the template produced from the averaging along with the analytical pulse template determined from the fit. Note that the entire pulse template is 20 ms long, but a zoomed-in portion around the trigger point (10 ms) is shown here.

6.2 Data quality selection criteria

After a trigger is determined, reduced quantities (RQs) are extracted from the raw trace to provide information about the event that can be used in the analysis; the processing of the traces is performed once and the RQs are stored. The RQs include the optimal filter amplitude and χ^2 (goodness-of-fit), along with several other quantities discussed in this section. Triggered events that pass selection criteria (cuts) based on one or more RQs can be selected for use in the analysis. To determine the noise template (PSD) used by the optimal filter, additional cuts are defined to select traces from the random triggers. This section will describe the selection of the livetime used in this analysis, and the definitions

of cuts to select good physics events and noise traces.

6.2.1 Livetime cuts

For analyses with thresholds just above the noise level, non-stationary (time varying) noise can introduce unwanted triggers. In this analysis, an increased trigger rate was observed during time periods with elevated noise. A heatmap of the noise over time for Run 28a compared to the average trigger rate is shown in Figure 6.3, confirming that periods with elevated noise generally have an increased trigger rate.

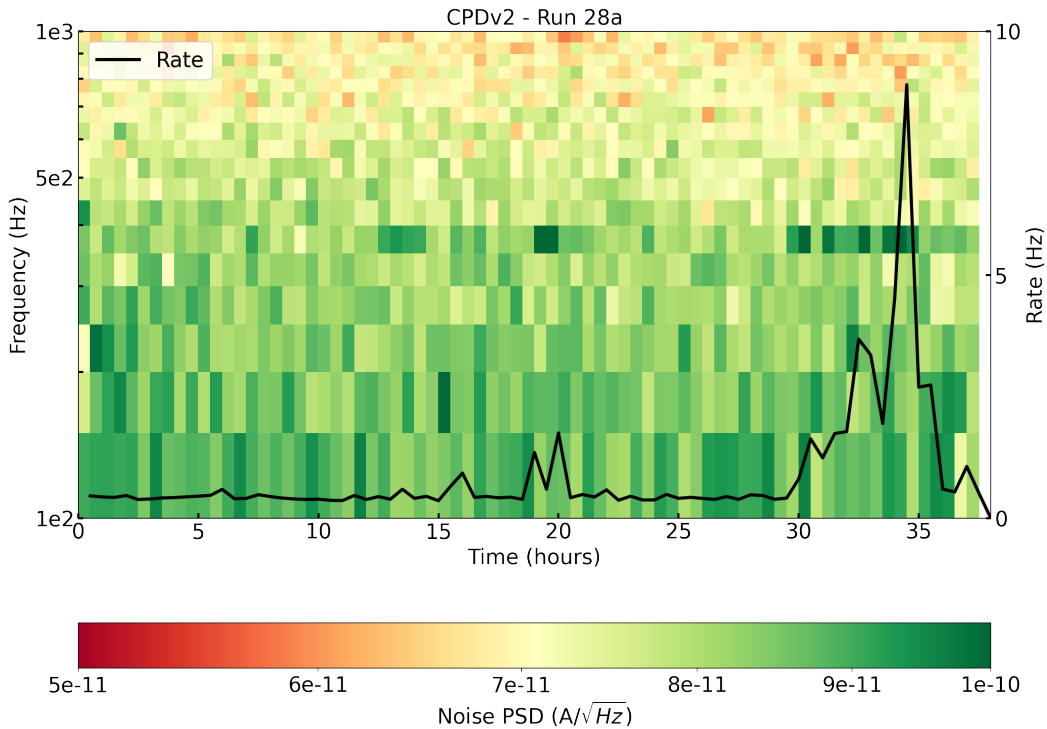


Figure 6.3: Periodogram of the noise below 1 kHz over time for Run 28a, frequency bins are shown on the left hand axis while the amplitude is given by the color bar; the low energy trigger rate is shown in black on the right hand axis. The trigger rate can be seen to increase during periods with excess noise in the 350 Hz range.

Since this analysis is background limited (the livetime does not drastically impact the final limit), a good-time cut was developed to discard noisy time periods. The good-time cut consists of two parts: a broad and a fine time cut. The broad time cut was set by looking at the average noise PSD over time, and excluding periods that had elevated noise; in Run 28a, hours 12-25 were selected, while for Run 28b only the first 9 hours

6.2. DATA QUALITY SELECTION CRITERIA

were used in the analysis. The fine time cut was developed to remove time periods with an excess trigger rate within the broad periods. As alternating files in the continuous data set were blinded, the fine time cut was developed to remove time periods from the blinded portion based on information from the unblinded portion. Using the unblinded portion of the data set, a cut line on the low energy rate is defined by considering events in the energy range between 4.5 to 40 times the optimal filter energy resolution. The procedure for this is as follows:

1. Calculate the event rate in the energy range of interest for each unblinded data file
2. Take the minimum event rate as the baseline
3. Calculate the standard deviation of the difference in the event rate between nearest neighbour unblinded files
4. Define the rate cut line as the minimum event rate plus five times the standard deviation of the differences

After determining the rate cut line, the time periods to remove are determined in a way that allows for the blinded periods to be removed as well. The midpoints (in time) of all the unblinded data files are determined and the low energy event rate is calculated for the halves of the unblinded data files that shoulder a blinded data file. In other words, the event rate of the later half of the unblinded file before a blinded file, and the first half of the unblinded file after a blinded file is calculated. If this event rate exceeds the rate cut line then the entire time period between the two midpoints of the unblinded files is excluded. This method removes blinded (and unblinded) time periods that could have an elevated low energy event rate, without biasing the analysis by opening the blinded data files.

6.2.2 Baseline cuts

As the detector was operated on surface, the rate of cosmic ray interactions is quite high. Cosmic ray interactions in the detector can heat the crystal, leading to subsequent particle interactions sitting on the falling tail of a high energy cosmic ray interaction. This can

CHAPTER 6. A SEARCH FOR LOW-MASS THERMAL RELICS USING A WIDE AREA PHOTON DETECTOR

result in improper energy reconstruction as the gain may be different, so an important set of cuts used in the analysis are the prepulse baseline cuts. For this purpose two data quality cuts are used: one to remove events with a high prepulse baseline average, and another to remove events that have a non-negligible slope.

The prepulse baseline average RQ is calculated by taking the average of the bins in the first 9.5 ms in the trace (recall the trigger point is at 10 ms). The prepulse baseline average RQs are grouped into 30 minute intervals over the entire livetime, and a Gaussian function is fit to each of these groupings. Events in each time period that have a prepulse baseline average RQ within a window (defined as the mean of the Gaussian plus or minus three times the width of the Gaussian) pass the cut, while events with a prepulse baseline average outside this window fail. Figure 6.4 shows the prepulse baseline average RQ over time for Run 28a and Run 28b, with events passing the cut shown in orange.

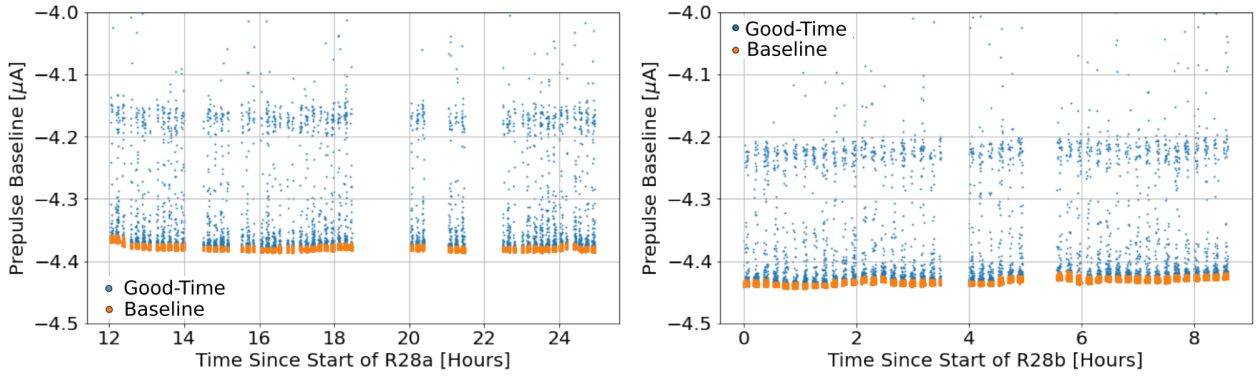


Figure 6.4: Baseline RQ over time for CPDv2 in UMass Run 28a and Run 28b. All events passed the good-time cuts, while the events in orange also passed the prepulse baseline cut.

To remove events that display evidence of having a slope in the baseline current (for example from muonic heating), a slope RQ ($\hat{m}$) is calculated as:

$$\hat{m} = \frac{y_{post} - y_{pre}}{t_{post} - t_{pre}}. \quad (6.2)$$

Here y_{pre} is the average of the baseline current in the first 9.5 ms of the trace, y_{post} is the average of the baseline current in the last 5 ms of the trace, t_{pre} is the mean time of the first 9.5 ms (4.25 ms), and t_{post} is the mean time of the last 5 ms (17.5 ms). The cut plane used to remove events with a large slope is shown in Figure 6.5. For energies below

6.2. DATA QUALITY SELECTION CRITERIA

1.5 keV, the cut is defined by fitting a Gaussian to the slope RQ of random triggers and passing events that fall within the mean of the Gaussian plus or minus three times the width of the Gaussian. As saturation effects begin to set in at higher energies, the slope RQ tends to increase; to avoid removing real events that are saturated, above 1.5 keV the definition of the slope cut changes. A Gaussian is fit to the slope RQ of ^{55}Fe events, the mean of this Gaussian plus three times its width is used as the upper limit at the position of the ^{55}Fe peak (2.65 keV in calibrated optimal filter amplitude). A straight line is fit between the upper limit of the aluminum fluorescence peak energy and the ^{55}Fe peak energy, this straight line is used as the upper limit of the slope cut for energies above 1.5 keV.

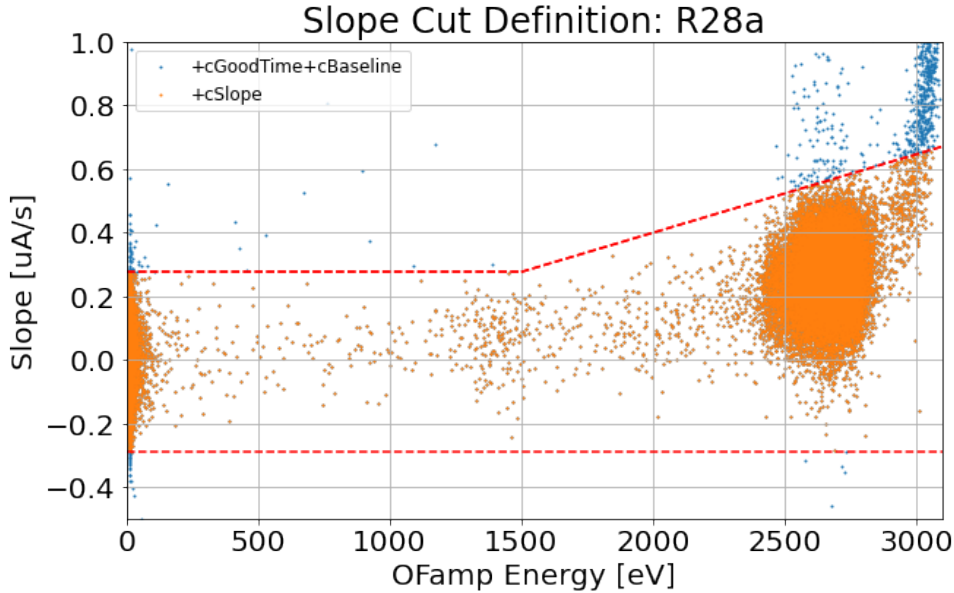


Figure 6.5: Scatter plot of the slope RQ ($\hat{m}$) against energy, the upper and lower cut bounds are shown with dashed red lines. All events passed the good time and prepulse baseline cuts, while events in orange also passed the slope cut.

6.2.3 χ^2 cuts

Another important metric in determining if an event should be used in the analysis is the shape of the pulse in a triggered event. At low energies, noise in the trace can alter the shape of the pulse, while at higher energies the pulse shape changes due to saturation effects. Each of these effects can impact the reconstructed energy determined by the optimal filter. The χ^2 of the optimal filter fit parameterizes the goodness-of-fit; the χ^2

CHAPTER 6. A SEARCH FOR LOW-MASS THERMAL RELICS USING A WIDE AREA PHOTON DETECTOR

of a good fit should be close to the number of degrees of freedom used in the fit. In practice, the χ^2 is generally a distribution centered on the number of degrees of freedom used in the fit, with some width. Events that have a distorted pulse shape (relative to the template) can be excluded from the analysis by removing events with a high χ^2 .

In this analysis the low-frequency χ^2 (χ_{LF}^2) is used to increase the sensitivity. The χ_{LF}^2 is defined as the χ^2 of the optimal filter fit using only the frequency bins below a certain cut-off frequency, which is 10 kHz in this analysis. The main benefit in using the low-frequency χ^2 is that it considers the portion of the fit to the frequency bins where the signal template is larger, as the noise is more dominant at higher frequencies. With a sampling frequency of 1.25 MHz and 20 ms traces, a cut-off frequency of 10 kHz corresponds to 400 bins used in the fit. A cut based on the χ_{LF}^2 RQ is used to remove events with a distorted pulse shape; the cut is defined such that events with a χ_{LF}^2 below 500 pass the cut. In order to keep events that have a high χ_{LF}^2 due to saturation effects (such as the ^{55}Fe events), this cut is only applied for energies below about 300 eV, allowing all events with higher energies to pass.

It was noted during the analysis that some triggered events appeared to have a feature similar to a 360 Hz sine wave; the average of several traces that display this feature is shown in Figure 6.6. These events were typically reconstructed in the energy range of 12-18 eV and generally had an elevated prepulse baseline average and χ_{LF}^2 , meaning they were mostly removed by other cuts. However, to ensure no events with this feature were accepted into the final spectrum, a dedicated cut was developed. A type of optimal filter known as the N-signal-M-background ($N_s M_b$) optimal filter was used to find traces with evidence of the sine wave feature. The $N_s M_b$ optimal filter fits N signal templates and M background templates to a trace, in the presence of stationary noise; more details on the $N_s M_b$ optimal filter can be found in Appendix A. For this analysis, the $N_s M_b$ optimal filter is used to fit a pulse in the presence of a sine wave with a fixed frequency. Due to the implementation of the $N_s M_b$ code, a sine and cosine with a fixed frequency (360 Hz) are used as the background templates to account for the phase through the relative amplitudes, which is equivalent to fitting the amplitude and phase of a single sine wave. The $N_s M_b$ optimal filter can also be used as a background only (M_b) optimal filter by

6.2. DATA QUALITY SELECTION CRITERIA

removing the fit to the signal templates. For this analysis, the difference in χ^2 ($\Delta\chi^2$) between the χ^2 of the normal (1-signal) optimal filter and the χ^2 of the background only (sine wave) optimal filter was calculated. The $\Delta\chi^2$ against optimal filter energy plane can be seen in Figure 6.7. To remove events with a large $\Delta\chi^2$, a cut line (shown in Figure 6.7) was defined as:

$$y = 9 - x^2, \quad (6.3)$$

where x is the optimal filter amplitude divided by the resolution of the optimal filter. Events above the line had a better fit to the sine wave template than to the pulse template, and are removed from the final spectrum.

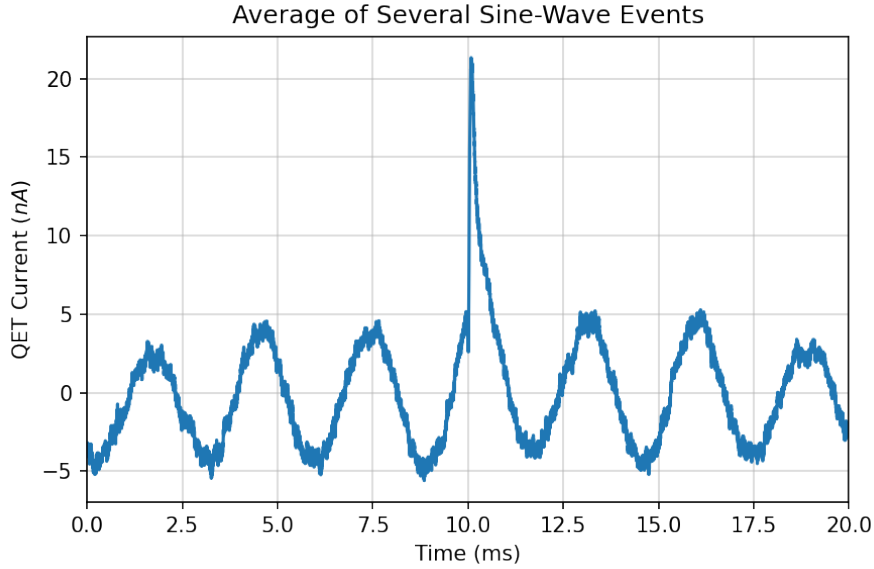


Figure 6.6: Average of several raw traces that display evidence of the 360 Hz sine wave feature. These traces had energies in the range 12-18 eV.

6.2.4 Noise cuts

Additional cuts were developed to determine the final noise PSD for the optimal filter, using traces from the random triggers. The initial processing is done using a noise PSD determined by the autocuts algorithm from the QETpy module¹, which is efficient at removing traces with pulses in them. However, the autocuts algorithm tends to allow

¹<https://github.com/spice-herald/QETpy>

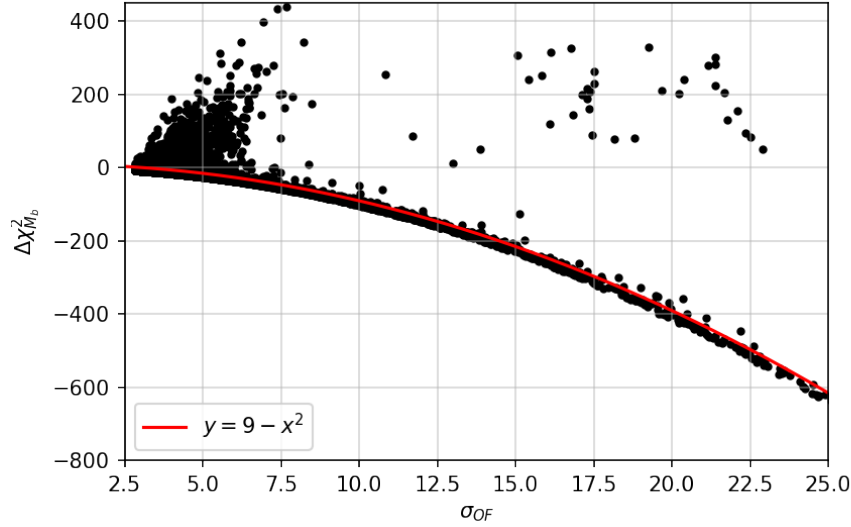


Figure 6.7: Scatter plot of the $\Delta\chi^2$ RQ against the optimal filter amplitude divided by the resolution of the optimal filter. The cut line is shown in red, events below the red line pass the cut while events above fail.

more low frequency noise into the spectrum, which results in a worse energy resolution and an accordingly elevated trigger threshold.

The first cut used to determine the final noise PSD is one that removes traces with evidence of a pulse. The χ^2 of the optimal filter where the signal template is allowed to shift in time is calculated, since a randomly triggered trace could have a pulse at any position. Next, the χ^2 is calculated by assuming that there is no pulse in the trace, this is called the “no-pulse” optimal filter. The difference between these two χ^2 values is calculated as $\Delta\chi^2_{noise} = \chi^2_{pulse} - \chi^2_{no-pulse}$, and events in the bottom 40th-percentile of the $\Delta\chi^2_{noise}$ distribution of all the randoms are removed.

The second cut used in determining the final noise PSD is one that removes traces that display evidence of the 360 Hz sine wave. The difference between the χ^2 of the M_b optimal filter (described above) and the no-pulse optimal filter is taken, and events in the bottom 1th-percentile are removed. Only the bottom 1% of events in this distribution are removed because this cut is very sensitive to the particular frequency of the sine wave, and can remove events that do not actually display the large sine wave feature but just have excess noise in that frequency bin. For example see the inset of Figure 6.8, if the cut

6.2. DATA QUALITY SELECTION CRITERIA

were placed at a higher percentile, the amplitude of the 350 Hz noise bin would decrease below the rest of the other noise bins.

Using these two cuts, in combination with the other livetime and data quality cuts described above, this procedure (of reapplying all cuts as described above) is repeated iteratively through several rounds of processing, until the noise PSD eventually converges. Once the noise PSD has converged, only that noise PSD is used for any subsequent rounds of processing. Figure 6.8 shows the evolution of the noise PSD after subsequent quality cuts are applied. The cut that has the largest effect in reducing the calculated noise power at low frequencies is the No-Pulse cut, which removes traces with evidence of contamination by pulses that contribute to the excess power level in the low frequency range. The No-Sine cut has the effect of reducing the 350 Hz bin (which is the closest bin to 360 Hz), as evidenced by the inset of Figure 6.8. The noise PSD used for the final optimal filter is the PSD calculated after the application of all the cuts described above.

The random traces that pass all of the good noise cuts can be used to determine the resolution of the optimal filter. By applying the no delay optimal filter (with the updated noise PSD) to these traces, a (roughly) Gaussian distribution of optimal filter amplitudes is obtained. It is important that the no delay optimal filter is used, as the introduction of a delay will bias the distribution towards higher energies (where the fit is better). The width of this distribution, which is determined by fitting a Gaussian, represents the resolution of the optimal filter. The Gaussian function represents the probability of reconstructing an event with a true energy of (essentially) zero. The distribution of the optimal filter amplitudes of these events is shown in Figure 6.9, along with the Gaussian fit. The measured width of the distribution is $3.95 \cdot 10^{-9}$ OFAmp, which corresponds to an energy resolution of roughly 2.3 eV based on the energy scale calibration described in the following section.

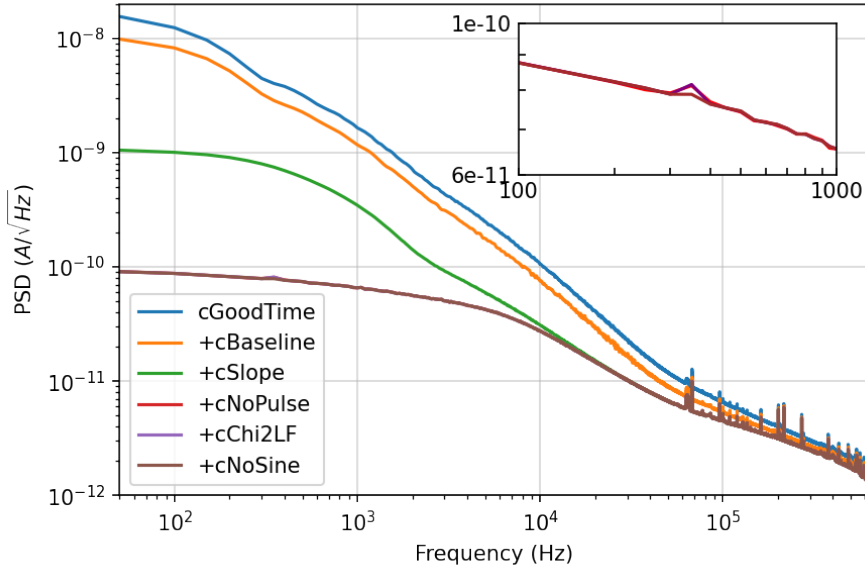


Figure 6.8: Evolution of the noise PSD after successive cuts are applied. The No-Pulse cut has the largest impact as it removes traces that contain pulses, which increase the overall low frequency noise level. The inset zooms into the low frequency portion of the lowest noise PSDs, highlighting the effect of the No-Sine cut which primarily reduces the level of the 350 Hz bin. The noise PSD used for the final optimal filter is labelled as +cNoSine.

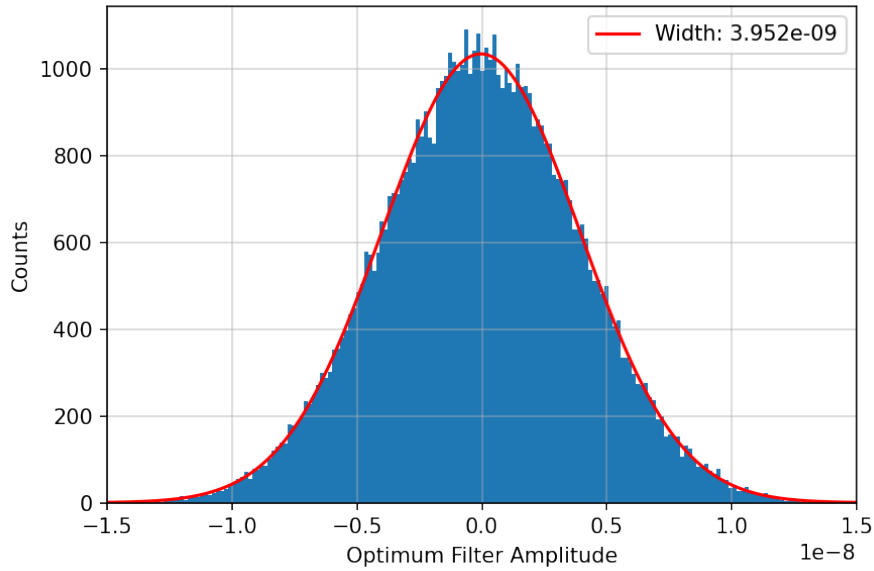


Figure 6.9: Distribution of the optimal filter amplitudes for the random trigger events that passed all of the good noise quality cuts. A Gaussian function is fit to the distribution to determine the energy resolution of the optimal filter, which corresponds to roughly 2.3 eV based on the energy scale calibration.

6.3 Calibration

To perform a meaningful analysis, the energy scale of the optimal filter amplitude must be converted to units of energy. In this experimental setup, monoenergetic peaks are present in the spectrum due to the ^{55}Fe source, which shines through several layers of aluminum foil. The electron capture decay of ^{55}Fe to ^{55}Mn produces monoenergetic photons with energies of approximately 5.9 keV and 6.5 keV, corresponding respectively to the K-alpha and K-beta transitions. When these photons interact with the aluminum, there is some chance of producing fluorescence photons, with an energy of roughly 1.49 keV. Here the aluminum fluorescence peak is used to calibrate the energy scale of the detector, as the ^{55}Fe peaks are highly saturated and would lead to an underestimation of the absorbed energy in the detector.

The first step in calibrating the detector's absolute energy scale is to determine the QET's collection efficiency using the energy absorbed in the QET. Recall the QET's read out circuit diagram, shown in Figure 4.1. The power-to-current transfer function ($\partial P/\partial I$) for a TES is given by [124]:

$$\frac{\partial P}{\partial I}(\omega) = \left(I_0 \left(1 - \frac{1}{\mathcal{L}} \right) \left(1 + \frac{j\omega\tau_0}{1 - \mathcal{L}} \right) \right) \left(R_L + R_0(1 + \beta) + j\omega L + \frac{R_0 \mathcal{L}(2 + \beta)}{(1 - \mathcal{L})(1 + \frac{j\omega\tau_0}{1 - \mathcal{L}})} \right). \quad (6.4)$$

Here, I_0 is the quiescent current through the TES, and R_L is the load resistor. The term β represents the current sensitivity of the TES, a typical value for SuperCDMS detectors is $\beta < 3$ [118]; $\tau_0 = C/G$ is the TES's natural time constant in the absence of electrothermal feedback [124]. In the DC limit ($\omega = 0$) and assuming the low-frequency loop gain $\mathcal{L}$ (related to the TES's current sensitivity to temperature changes) to be infinite [124], the current-to-power transfer function becomes:

$$\frac{\partial I}{\partial P}(0) = -\frac{1}{I_0(R_0 - R_L)}. \quad (6.5)$$

Rearranging and substituting in $R_0 = V_b/I_0 - R_L$, where $V_b \approx I_b R_{sh}$ is the TES bias voltage, gives:

$$dP_0 = -I_0(V_b - 2I_0 R_L)dI_0. \quad (6.6)$$

Through energy conservation, the change in the power absorbed in the TES from a

CHAPTER 6. A SEARCH FOR LOW-MASS THERMAL RELICS USING A WIDE AREA PHOTON DETECTOR

perturbation is $\Delta P_{abs} = -\Delta P_0$, and can be found by carrying out the integral:

$$\Delta P_{abs} = \int (V_b - 2I_0 R_L) dI_0 = (V_b - 2I_0 R_L) \Delta I - \Delta I^2 R_L, \quad (6.7)$$

where ΔI is the deviation of current through the QET from its baseline value. The energy absorbed E_{abs} in the QET is found by integrating this expression over time:

$$E_{abs} = \int (V_b - 2I_0 R_L) \Delta I - \Delta I^2 R_L dt, \quad (6.8)$$

noting that ΔI is the measured quantity that changes over time, for instance during a particle interaction. The remaining quantities in Equation 6.8 can be measured, and for Run 28 these values are: $V_b = 0.13\mu\text{V}$, $I_0 = 4.1\mu\text{A}$, and $R_L = 4.1\text{m}\Omega$. The integral is performed from the trigger time (10 ms) to 2.5 ms after the trigger time.

The spectrum of the absorbed energy for triggered events (that passed the quality cuts described above) is shown in Figure 6.10. Using these events, the position of the aluminum fluorescence peak in terms of the absorbed energy is determined using a Gaussian kernel density estimator (KDE). KDEs are a method of estimating a univariate distribution based on a series of independent observations [141] (here the absorbed energy in the QET). The density distribution of the variable is estimated by taking the sum of Gaussian kernel functions (with a fixed width) at the position of each observation. The position of the peak was determined to be 379 eV, which corresponds to a collection efficiency of 25.4 % (379/1490) given the true energy of the aluminum fluorescence peak of 1490 eV.

When a high-energy particle interacts in the substrate, the sensor response saturates due to the large number of phonons being absorbed in the QET. As the sensor's resistance increases, it approaches its normal resistance and the response is no longer linear. The net effect of this non-linearity is that the shape of the pulse near the top flattens out, thereby reducing the area under the curve and, correspondingly, the absorbed energy parameter. This saturation effect explains why the ^{55}Fe peaks in Figure 6.10 appear at a lower energy than would normally be expected (in the absence of saturation effects).

While the spectrum of total deposited energy could be used to set constraints on the interaction strength of dark matter, the optimal filter amplitude is (in general) a better energy estimator. In the absence of saturation effects that alter the pulse shape, the optimal filter amplitude should have a better energy resolution than the integral

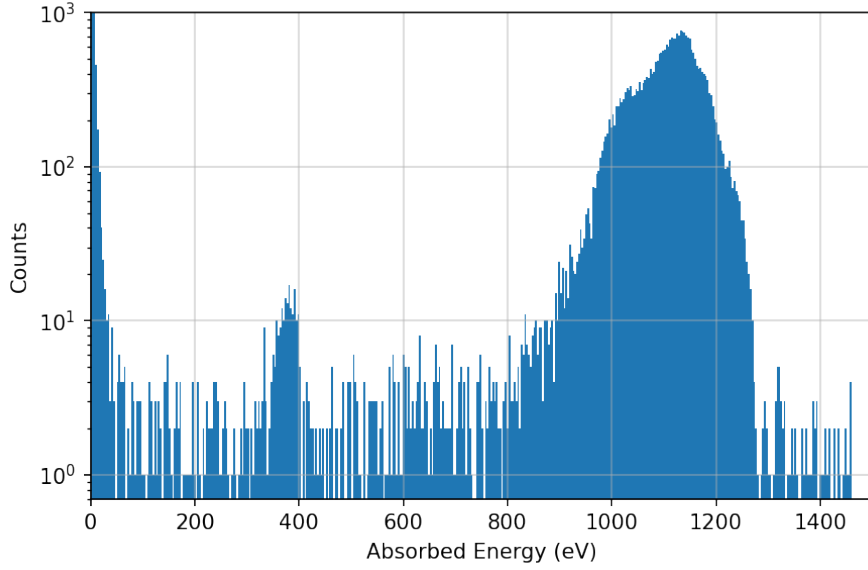


Figure 6.10: Spectrum of the absorbed energy RQ generated using events that passed all quality cuts. The aluminum peak (centered at 379 eV) is used to determine the collection efficiency of the detector. The wide blob centered around 1100 eV is from ^{55}Fe events and cosmic ray interactions; the resolution of these events is degraded and the true energy is underestimated, due to saturation effects.

(absorbed energy) estimator; the integral estimator is less sensitive to saturation effects, so at higher energies the resolution of the absorbed energy is expected to be better than the optimal filter amplitude. However, for this analysis the energy range of interest is below 100 eV, which is where the pulse template for the optimal filter was determined, so the optimal filter amplitude should be a better energy estimator.

The absolute energy scale of the optimal filter amplitude is determined by performing a linear fit between the total deposited energy and the optimal filter amplitude. The fit is fixed to pass through the origin, and the slope gives the calibration factor of the optimal filter to absolute energy (see Figure 6.11). As saturation effects begin to distort the optimal filter amplitudes at high energies and noise effects worsen the integral estimator at low energies, only a small energy range of events is used to calibrate the optimal filter amplitudes. The result of this linear fit gives the calibration factor of the optimal filter as: $5.892 \cdot 10^5$ keV/OFamp.

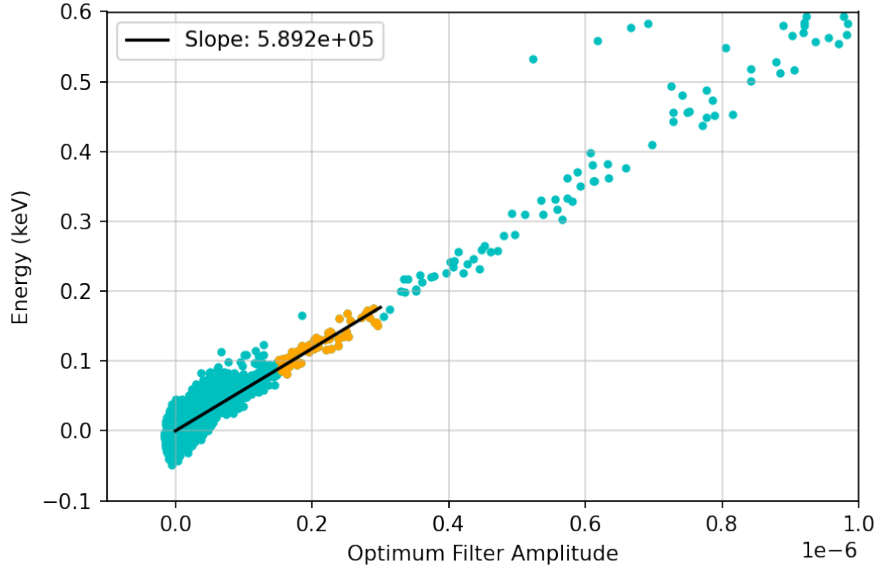


Figure 6.11: Calibration of the absolute energy scale of the optimal filter amplitude by fitting to the deposited energy (integral estimator) in the detector. Shown are triggered events that pass all livetime and data quality cuts; only the orange points are used in the fit for the calibration factor to avoid bias by noise (at low energy) and saturation effects (at high energy).

6.4 Efficiency determination

To determine the passage fraction of the data quality cuts and the trigger efficiency, a procedure known as “salting” is performed, where simulated pulses are randomly injected into the data stream. By randomly injecting pulses with a known energy, events with an energy below the trigger threshold can be forced to trigger, which allows the trigger efficiency as function of true energy to be measured. The fake pulses are identical to the pulse template used by the optimal filter template, and scaled to absorbed energies between 0.1 eV and 50 eV (between 0.1 eV and 10 eV in steps of 0.1 eV; above that in steps of 2 eV). The continuous data stream is salted at a rate of roughly 5 Hz. The salting procedure is described below:

1. The trigger algorithm is performed on the raw (unsalted) data, real triggers are recorded.
2. Salted pulses are injected into the continuous data stream with random energy and time. A $100 \mu\text{s}$ time separation is required between adjacent salt pulses.

6.4. EFFICIENCY DETERMINATION

3. The trigger algorithm is performed on the salted trace, searching for the optimum trigger around the true salt time with a $\pm 50 \mu\text{s}$ coincident window.
4. Information about each injected salt trace is recorded.
5. The above steps (except for 1) are repeated about 10 times, until sufficient statistics are obtained.

After processing the salted data, an efficiency map can be generated that relates the probability of triggering (and passing the quality cuts) to the the true (salt) energy and the reconstructed (optimal filter) energy. The passage fraction of triggered salt events as a function of true energy is directly measured, since the true salt energy is known. For energy depositions below threshold it is possible that the noise in the detector causes the event energy to be reconstructed above threshold. This effect, referred to as noise boosting, can lead to below threshold events being reconstructed above threshold, due to the finite energy resolution of the detector. To reduce the sensitivity to events with a true energy of zero (which would be unphysical), an anti-coincidence trigger is implemented where events are rejected if a trigger was issued on both the salted and unsalted trace. This will be described in further detail in the following section, but the main point of the anti-coincidence is to minimize the impact of erroneously claiming sensitivity to a below threshold event being reconstructed above threshold. In fact, the noise boosting effect can be quite beneficial as it provides additional sensitivity to below threshold events, however this effect must be appropriately quantified.

The combined trigger and cut efficiency as a function of the true recoil energy determined from the salted data is shown in Figure 6.12. The non-zero efficiency of events with a true recoil energy below the threshold is possible due to the detector's non-zero energy resolution, since the optimal filter can reconstruct events above the trigger threshold due to noise fluctuations. Above threshold the efficiency is roughly constant at about 80%, due to the quality cuts.

6.4.1 Expected probability distribution

The efficiency map (that is the efficiency as a function of true and reconstructed energy) determined using the anti-coincidence trigger scheme on the salted data is shown in Figure

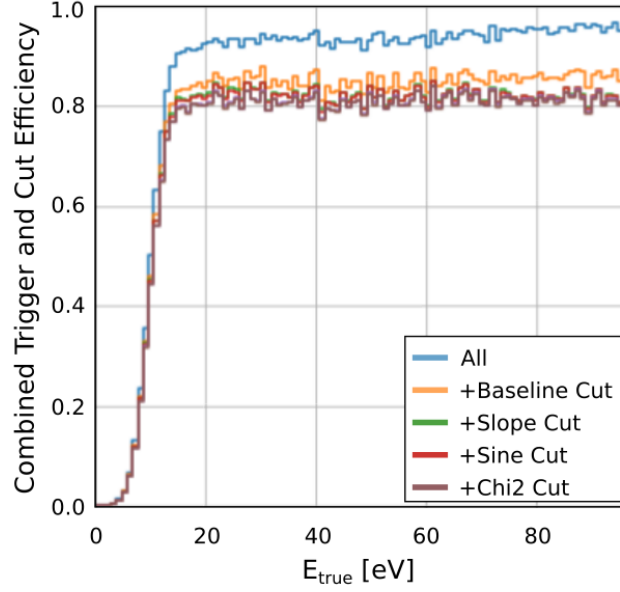


Figure 6.12: Combined trigger and cut efficiency as a function of the true simulation energy. The efficiency is shown after successive quality cuts are applied, and the final efficiency above threshold is roughly 80%.

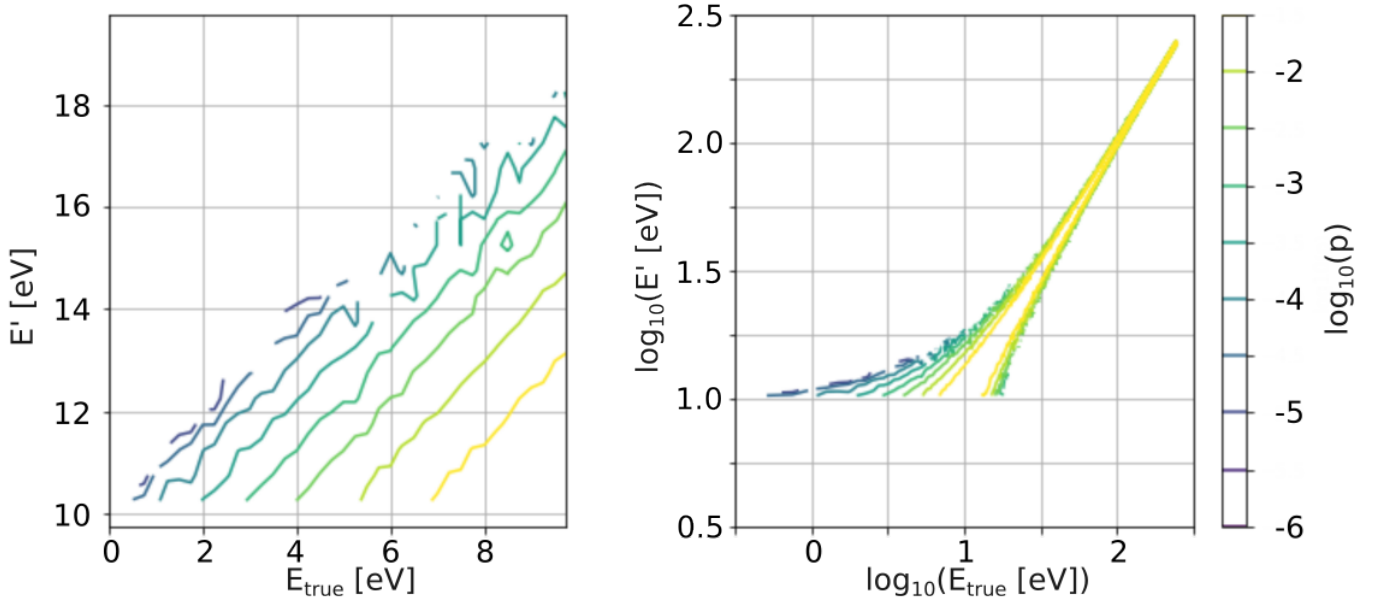


Figure 6.13: Efficiency map ($p_a(E'|E_{true})$) with the anti-coincidence trigger scheme determined from the salted data. The left hand plot shows a zoom into the low-energy portion on a linear scale, while the right hand plot shows a wider energy range on a logarithmic scale.

6.13. The difference between this map and the expected map (shown in Figure 6.15) can be accounted for by limited statistics in the salted data set. To determine if the efficiency map obtained from the salting method is reasonable, an analytical distribution of the

mapping function is derived.

One important quantity to derive is $p(E'|E_{true})$, which is the probability of an event with true energy E_{true} being triggered and reconstructed as E' . The distribution of the reconstructed energy due to noise fluctuations $P_0(E'|E_{true})$ is typically assumed to be a Gaussian, although it may have non-Gaussian tails. Independent of the shape of this distribution, it is assumed that this distribution is linear in energy (and its width is energy-independent), such that for a shift in energy ΔE :

$$P_0(E' + \Delta E|E_{true} + \Delta E) = P_0(E'|E_{true}). \quad (6.9)$$

The trigger efficiency $\eta(E')$ can be described as a Heaviside step function $H(E)$, and depends on the reconstructed energy and the trigger threshold E_{th} such that:

$$\eta(E') = H(E' - E_{th}). \quad (6.10)$$

At the time a salt pulse with energy E_{true} is injected, the probability of finding a trigger with energy E' is given by:

$$p_0(E'|E_{true}) = \eta(E')P_0(E'|E_{true}). \quad (6.11)$$

As the optimal filter is linear with the simulated pulse energy, the probability of finding a trigger at the same position in the unsalted trace is:

$$p_0(E' - E_{true}|0) = \eta(E' - E_{true})P_0(E' - E_{true}|0). \quad (6.12)$$

So, injecting a simulation pulse with energy E_{true} can be viewed as reducing the trigger threshold for the unsalted trace by E_{true} . As previously mentioned, to remove sensitivity to events with a true energy of zero, events are rejected if a trigger was issued on both the salted and unsalted trace. This scheme, known as the anti-coincidence trigger, is equivalent to requiring that the maximum fluctuation of E' due to noise is the same as the trigger threshold. The probability of finding a trigger that passes the anti-coincidence is given by:

$$\begin{aligned} p_{a0}(E'|E_{true}) &= p(E'|E_{true}) - p(E' - E_{true}|0) \\ &= \eta(E')P_0(E'|E_{true}) - \eta(E' - E_{true})P_0(E' - E_{true}|0) \\ &= (\eta(E') - \eta(E' - E_{true}))P_0(E'|E_{true}) \\ &\equiv \eta(E')(1 - \eta(E' - E_{true}))P_0(E'|E_{true}). \end{aligned} \quad (6.13)$$

CHAPTER 6. A SEARCH FOR LOW-MASS THERMAL RELICS USING A WIDE AREA PHOTON DETECTOR

From this it can be seen that $p_{a0}(E'|0) = 0$, irrespective of the form of $P_0(E'|E_{true})$; so, the anti-coincidence trigger scheme reduces the apparent sensitivity to zero energy events.

The previous results did not account for the fact that the optimum trigger time may not be exactly at the time of the salt. The probability of reconstructing an event with true energy E_{true} as E' within some coincident window around the true simulation time has been derived in Reference [142] to be:

$$P(E'|E_{true}) = P_0(E'|E_{true})F_0(E'|0)^{N_s-1} + (N_s - 1)P_0(E'|0)F_0(E'|E_{true})F_0(E'|0)^{N_s-2}, \quad (6.14)$$

where $F_0(E'|E_{true})$ is the cumulative distribution function (CDF) of $P_0(E'|E_{true})$, and N_s is the number of independent trace samples within the coincidence window. Note that in Reference [142] there was no constraint on the polarity of the pulse, whereas here the search is restricted to positive pulses, so the error functions in Equation A3 of Reference [142] have been changed to CDFs. The number of independent samples is given by the sampling frequency times the width of the coincident window, but due to correlations in the noise, the effective number is reduced. With this, the probability of finding a trigger becomes:

$$p(E'|E_{true}) = \eta(E')P(E'|E_{true}). \quad (6.15)$$

The second term in Equation 6.14 remains the same with time shifting because the terms remain the same for each independent sample. The first term in Equation 6.14 can be described by one of two scenarios:

1. The amplitude measured at the time of the salt is above threshold ($E' - E_{true} > E_{th}$).
2. The amplitude measured at the time of the salt is below threshold, but one or more independent samples are above threshold.

The first case contributes to the probability of finding a trigger:

$$p_1 = \eta(E' - E_{true})P_0(E' - E_{true}|0)F_0(E'|0)^{N_s-1}, \quad (6.16)$$

while the second case contributes:

$$p_2 = \eta(E')(1 - \eta(E' - E_{true}))P_0(E' - E_{true}|0)(F_0(E'|0)^{N_s-1} - F_0(E_{th}|0)^{N_s-1}). \quad (6.17)$$

6.4. EFFICIENCY DETERMINATION

With this, the probability of finding a trigger within the coincident window can be written as:

$$\begin{aligned}
p_a(E'|E_{true}) &= p(E'|E_{true}) - p_1 - p_2 \\
&= \eta(E')P_0(E'|E_{true})F_0(E'|0)^{N_s-1} - p_1 - p_2 \\
&= \eta(E')(1 - \eta(E' - E_{true}))P_0(E'|E_{true})F_0(E'|0)^{N_s-1} - p_2 \\
&= \eta(E')(1 - \eta(E' - E_{true}))P_0(E'|E_{true})F_0(E_{th}|0)^{N_s-1}.
\end{aligned} \tag{6.18}$$

This means that the boosting factor (chance of reconstructing above threshold) due to allowing the time shift is reduced by the anti-coincidence trigger scheme. Additionally, as $F_0(E_{th}|0) \leq 1$, $N_s > 1$ introduces a penalty to the trigger efficiency. This can be understood by considering a situation where the salted time results in: $E' > E_{th}$ and $E' - E_{true} < E_{th}$, which will generate a trigger without the time shift. If the time shift is allowed, any independent samples that fluctuate above E_{th} will generate the anti-coincidence; the more samples allowed, the higher the chance of anti-coincidence becomes. In practice the coincidence window must be wide enough to allow at least one independent sample to account for the trigger time fluctuation [142].

With Equation 6.18, the estimation of $p_a(E'|E_{true})$ determined from the salting method can be compared to the analytical form for two cases:

1. P_0 is a Gaussian with width 2.3 eV, and $N_s = 1$ (no time shifting allowed).
2. P_0 is a Gaussian with width 2.3 eV, and $N_s = 3$ (time shifting is allowed).

For both of these cases the trigger threshold is set to 4.5 times the energy resolution. Figure 6.14 shows the analytical form of the efficiency map for both of these cases with no anti-coincidence triggering implemented. The key takeaway is that without the anti-coincidence trigger scheme there is some sensitivity (non-zero efficiency) to zero energy events. Figure 6.15 shows the analytical form of the efficiency map for the two cases with the implementation of the anti-coincidence trigger, which is seen to remove the sensitivity to arbitrarily low-energies.

6.4.2 A comment on a new analysis

At the time of writing this thesis, an updated analysis on the UMass Run 28 data set was under development. The new analysis utilizes the same cut definitions and energy scale

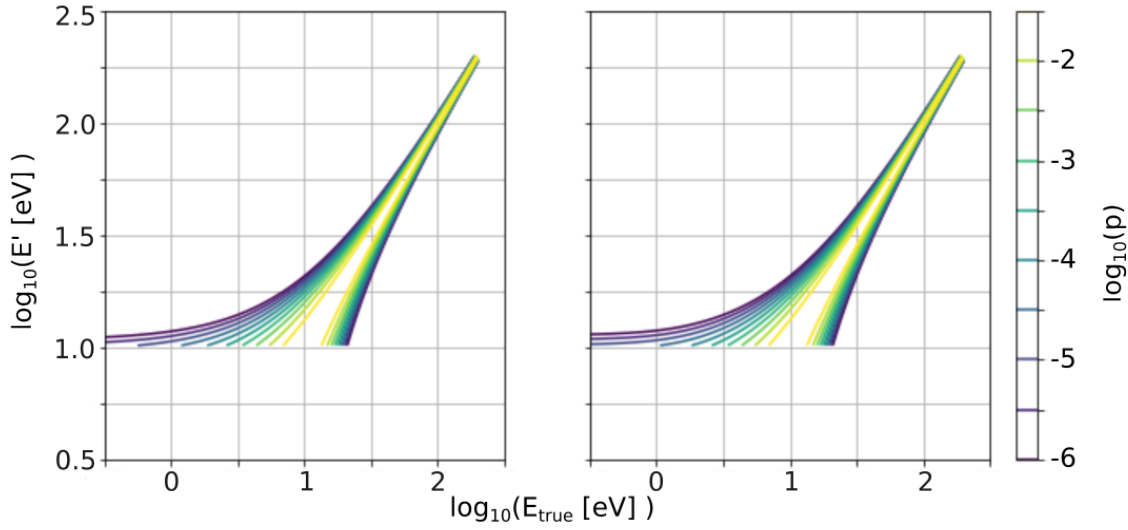


Figure 6.14: Analytical form of the efficiency map ($p(E'|E_{true})$) with no anti-coincidence trigger scheme for a detector with 2.3 eV energy resolution and a trigger threshold of 4.5 times the resolution. The left hand plot shows the case with no time shifting allowed, while the right hand plot shows the case with time shifting allowed. In both cases there is still sensitivity to zero energy events.

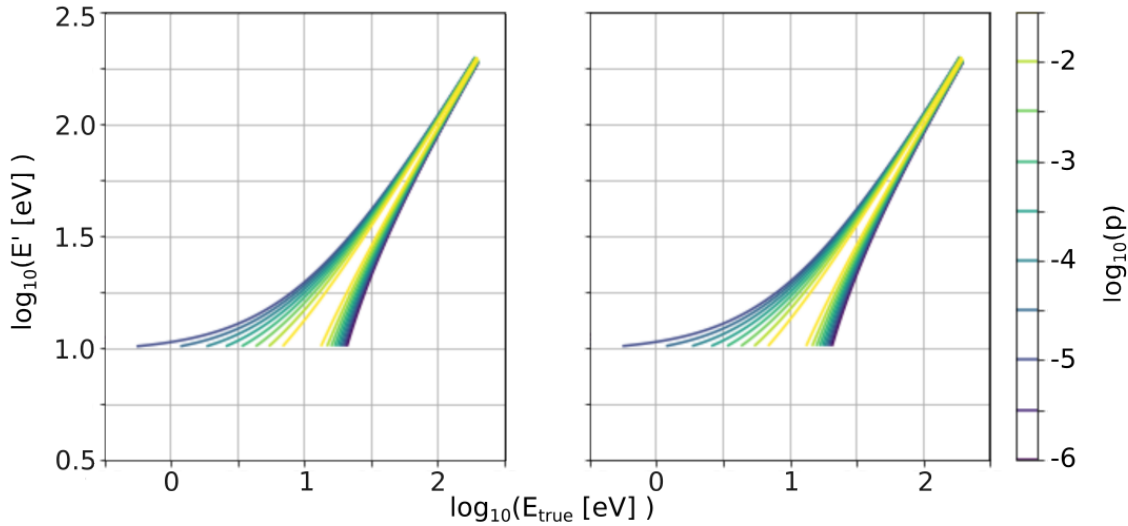


Figure 6.15: Analytical form of the efficiency map ($p_a(E'|E_{true})$) with the anti-coincidence trigger scheme for a detector with 2.3 eV energy resolution and a trigger threshold of 4.5 times the resolution. The left hand plot shows the case with no time shifting allowed, while the right hand plot shows the case with time shifting allowed. In both cases the sensitivity to zero energy events has been removed.

6.4. EFFICIENCY DETERMINATION

calibration as described above, but the method of determining the efficiency differs. In the updated version of the analysis the efficiency is determined by salting the data with a fixed rate of mono-energetic pulses, running the trigger algorithm, and measuring the reconstructed energy spectrum given a known rate of mono-energetic pulses are in the spectrum. The difference of the salted and unsalted spectra (given a known rate of “dark matter” events) is related to the efficiency to quantify the noise boosting effect [143].

The differential rate (in reconstructed energy) as a function of a known interaction rate R_{DM} of dark matter with mass M_{DM} is given by [143]:

$$\frac{dR}{dE'}(E'|R_{DM}, M_{DM}) = \frac{dR}{dE'}(E'|0) + R_{DM}(f(E'|E_{DM}) - f(E'|0)), \quad (6.19)$$

where $f(E'|E)$ is the parameter to be derived. The main benefit to this formalism is that in the limit where $M_{DM} \rightarrow 0$:

$$\begin{aligned} \frac{dR}{dE'}(E'|R_{DM}, M_{DM} \rightarrow 0) &= \frac{dR}{dE'}(E'|0) + R_{DM}(f(E'|E_{DM} \rightarrow 0) - f(E'|0)) \\ &= \frac{dR}{dE'}(E'|0), \end{aligned} \quad (6.20)$$

while in the original CPD analysis at SLAC [122] the subtraction of $f(E'|0)$ was not included such that:

$$\begin{aligned} \frac{dR}{dE'}(E'|R_{DM}, M_{DM} \rightarrow 0) &= \frac{dR}{dE'}(E'|0) + R_{DM}(f(E'|E_{DM} \rightarrow 0)) \\ &= \frac{dR}{dE'}(E'|0)(1 + R_{DM}). \end{aligned} \quad (6.21)$$

Note that this includes some dependence on the interaction rate of dark matter, implying that interactions of an ultra-light axion (10^{-19} eV) can be increased above a 10 eV threshold by noise fluctuations, which is clearly unphysical. The analysis presented here and the new (yet unreleased) analysis seek to reconcile this issue through two different methods. In this analysis, a scheme known as the anti-coincidence trigger is used to reduce the sensitivity to noise related triggers, while in the updated version this is done by subtracting the unsalted and from the salted spectra. The updated scheme should be more correct than the limit presented here, which may be overly conservative, and the overall reach to low masses should remain similar.

6.5 Results

CPDs are well suited for measuring low-energy nuclear recoils, making low-mass thermal relics an ideal search candidate [122, 102, 119]. The spin-independent WIMP-nucleus scattering rate ($\frac{dR}{dE_{true}}$) depends on the motion of the Earth in the galaxy, the nuclear form factor of the target, and the WIMP's phase space [144, 145]; a Python implementation of the rate calculation is provided by the RQpy package². Figure 6.16 shows the expected rate of dark matter events for various dark matter masses (and an interaction cross section fixed to 10^{-41} cm²), before and after accounting for the detector's energy resolution and the trigger efficiency. The energy resolution and trigger efficiency are accounted for by convolving the dark matter spectra with the efficiency map (determined from the salting). Figure 6.16 shows that the non-zero resolution provides sensitivity to recoil energies that would otherwise be sub-threshold, since there is some expected number of events above threshold for dark matter masses that should otherwise deposit all of their energy below threshold. This effect allows for the noise in the detector to increase the reconstructed energy of an event above threshold.

To set a limit on the interaction cross section of a dark matter particle (with a fixed mass), the optimum interval (OI) method is used [146, 147]. This method tests the compatibility (to a given confidence level) of the expected number of dark matter events with the observed number of events in different energy intervals. The OI method selects an interval with a particularly small number of observed events, compared to what would be expected for a true signal, and determines the interaction cross section that would be incompatible with the observed number of events at a given confidence interval [146, 147]. For instance, the 90% confidence interval would result in 90% of random experiments having more events than the number observed in the given experiment.

The 90% OI limit for the analysis discussed here is shown in Figure 6.17 (labeled as CPDv2 @ UMass) along with several other comparable limits from direct detection experiments. Two limits are calculated, one with the efficiency map determined with, and another without allowing for time shifting (Figure 6.15 left and right respectively). For

²<https://github.com/spice-herald/RQpy>

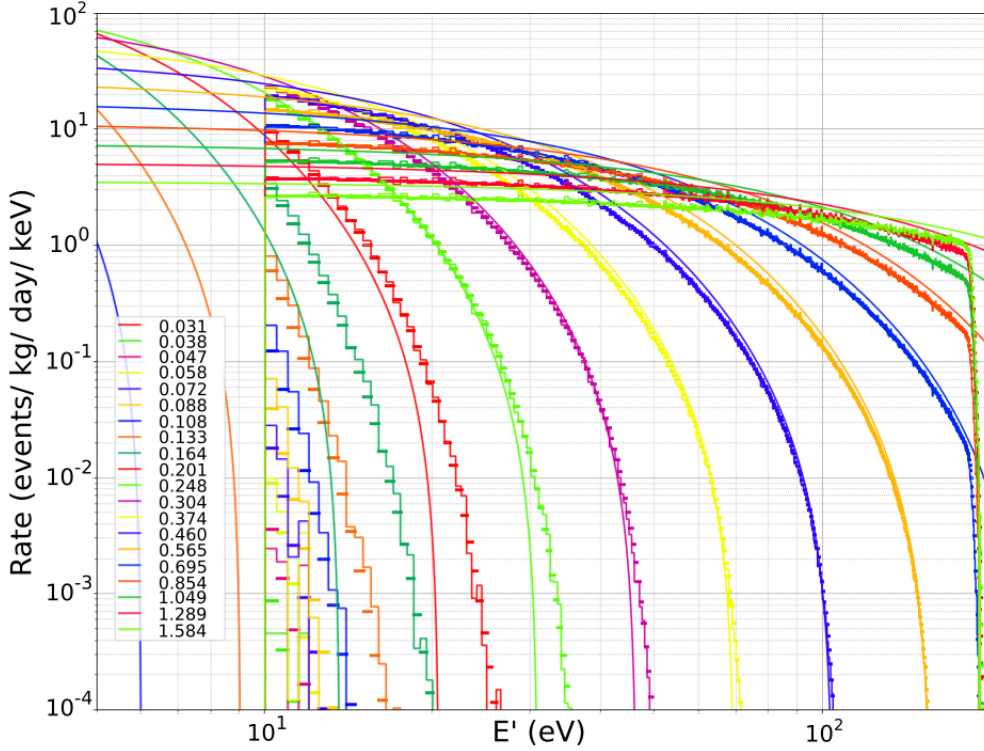


Figure 6.16: Dark matter recoil energy spectra as a function of reconstructed energy, for a fixed cross section of 10^{-41} cm^2 , shown by solid lines for different dark matter masses (given in the legend in units of GeV/c^2). The histograms correspond to the convolution of these spectra with the efficiency map. A toy model of this effect, shown by the horizontal bars, is generated by convolving the recoil spectra with a Gaussian with width 2.3 eV and applying a flat 80% efficiency above threshold.

dark matter masses that have the recoil energies mostly above threshold ($200 \text{ MeV}/c^2$), the two limits are comparable. However, for dark matter masses where a significant portion of the recoil energy is below threshold, the two curves begin to diverge, and the limit calculated with the efficiency map that allows the time shift is weaker, and thus more conservative. This is because of the penalty factor that is introduced by allowing a time shift, which increases the probability of an anti-coincidence and correspondingly reduces the efficiency.

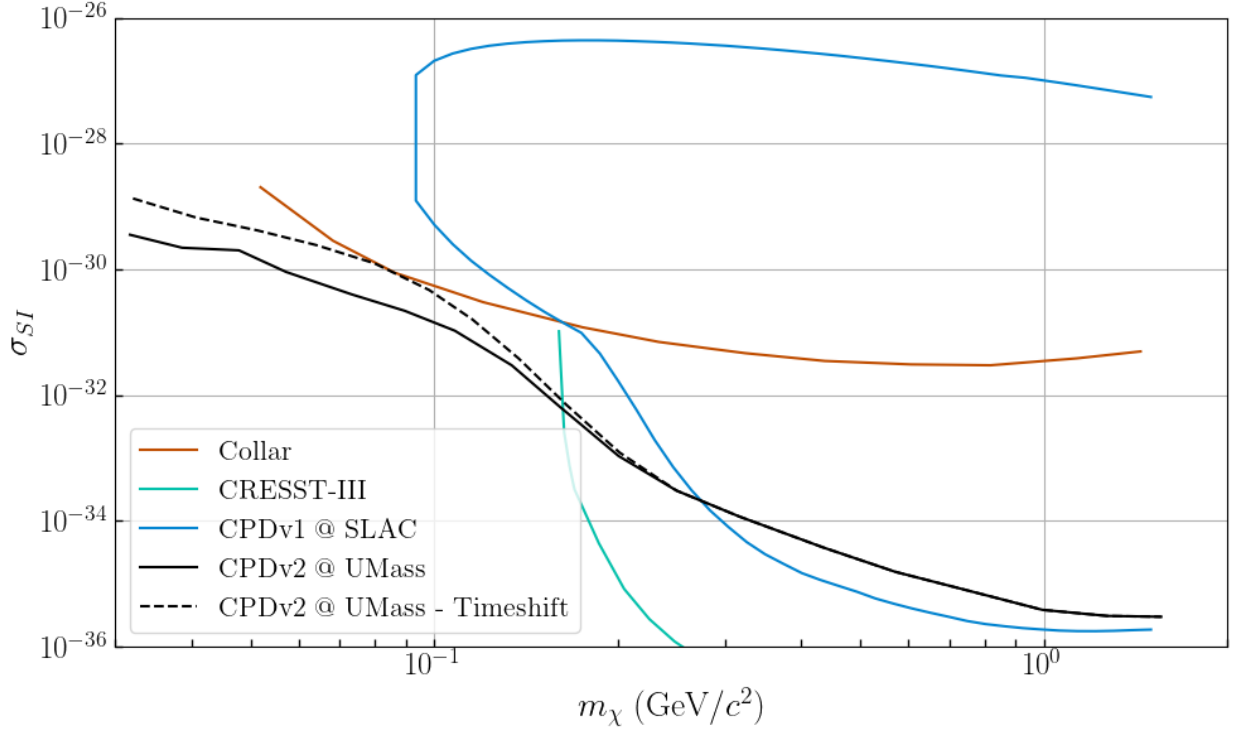


Figure 6.17: Constraints on the spin-independent coupling of WIMPs to nucleons, set by the analysis of the data from the second version of the CPD operated at the University of Massachusetts Amherst. Two limits are calculated with the same data, one where the calculated efficiency map allows for time shifting, and another where time shifting is not allowed; more details are given in the text. Shown alongside this result are comparable limits from other direct detection experiments [122, 148, 149].

6.6 Discussion

The improved energy resolution of CPDv2 provides sensitivity to lower mass dark matter particles, with respect to the original CPD [122]. The result from Collar [148] achieves a low threshold due to the hydrogenated scintillator used for the target, although the low energy background rate is elevated with respect to the UMass data set. The result from CPDv2 at UMass extends to dark matter masses of about 30 MeV/ c^2 , while the Collar result [148] has reach down to about 50 MeV/ c^2 . The limit from CPDv2 at UMass surpasses the Collar limit whether or not the time shift was allowed, although in the case with the time shift this margin is smaller.

The hydrogenated scintillator used by Reference [148] has better kinematic matching

6.6. DISCUSSION

to low mass dark matter ($< 1 \text{ GeV}/c^2$) than the (silicon) CPD, due to the high content of hydrogen. However, the excellent energy resolution of the CPD still provides sensitivity to lower mass dark matter particles. This is because the hydrogenated scintillator requires energy depositions on the order of 10 eV to produce scintillation light [148], while lower energy depositions in the silicon substrate can still produce detectable phonons. For dark matter masses larger than $160 \text{ MeV}/c^2$, the CPDv2 limit is surpassed by CRESST-III [149], which was operated underground and had a considerably lower background.

The original CPD analysis at SLAC [122] includes an upper limit that corresponds to the cross section in which the atmosphere shields the detector from the dark matter particles. For cross sections above this upper bound no limit can be set since the dark matter would not make it through the atmosphere to interact in the detector. No such calculation was performed here, but the cross section that such an effect would become relevant is around 10^{-28} cm^2 .

6.6.1 Low energy excess

The averaged differential rates for the second version of CPD operated during UMass Runs 26 and 28 are shown in Figure 6.18. A high rate of low-energy events is observed, which is similar to the observation from the original CPD operated at CUTE. During both Runs 26 and 28, this low energy rate appears to decay over a relatively short time. However, since CPDv2 was not operated at UMass as long as the original CPD was operated at CUTE, it is difficult to get an accurate measurement of the drop in rate.

It can be seen in Figure 6.18 that the low energy rate is comparable between the Runs 26 and 28. If the decaying rate were due to some radiogenic background, it should decay at the same rate irrespective of the detector being cold or warm. Instead, the low-energy excess appears to decay predominantly while the detector is cold. The primary hypothesis for this (similar to CPD at CUTE) is that the excess rate is related to stress introduced to the substrate by the clamping scheme. A dedicated experiment at the University of California, Berkeley was performed to validate this hypothesis by producing two functionally identical silicon wafers with the same QET mask [14]. These devices were mounted in two different ways: a high stress configuration where the wafer was glued to

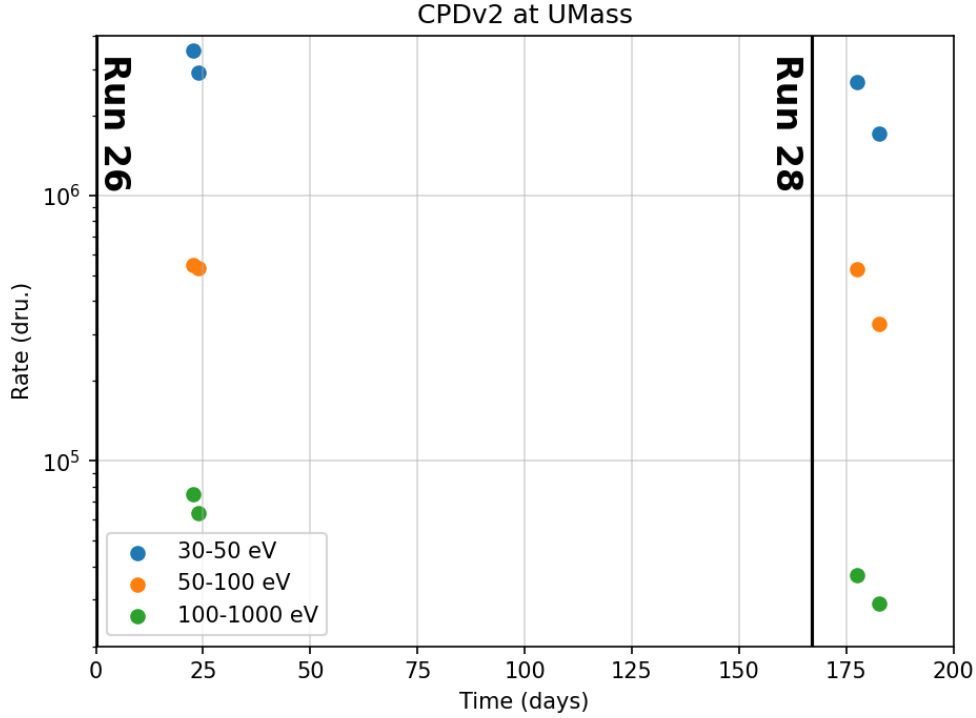


Figure 6.18: Rate (after quality cuts) in CPDv2 at UMass for the two cryogenic runs that it was operated in. The low-energy rate seems to decrease exponentially over time while the detector is cold. Radiogenic backgrounds are not a good explanation for this decay as decay rate seems to increase while the detector is cold. One possible explanation is stress induced microfractures in the silicon releasing while cold.

a copper mount, and a low stress configuration where the wafer was suspended with gold wire bonds [14]. It was found that the device mounted in the high stress configuration exhibited an excess low-energy rate (of more than two orders of magnitude) relative to the device mounted in the low stress configuration [14]. This dedicated experiment provides good evidence that stress can cause excess events at low energies, making the hypothesis that the low-energy excess in CPD at CUTE and UMass are also caused by stress plausible.

These low-energy events result in excess quasiparticles in the superconductor (of the QET), contributing to an effect that in a different context is often referred to as “quasi-particle poisoning” [14, 15]. Quasiparticle poisoning has been identified as a challenge in the development of superconducting qubits, as excess quasiparticles reduce the coherence time of the qubits [15]. A long coherence time is important for the development of

6.6. DISCUSSION

a practical quantum computer based on superconducting qubits. The source of quasiparticle poisoning in superconducting qubits is still poorly understood [15] and several mechanisms have been proposed, including interactions of environmental ionizing radiation [15] and microfracture events [14]. The development of dedicated measurements, such as the one at Berkeley, can be used to better understand mechanisms of quasiparticle poisoning that are relevant for particle detectors based on superconducting sensors that are sensitive to very low energy depositions (such as low-mass dark matter direct detection experiments), and quantum computing experiments that rely on superconducting circuits.

Chapter 7

Constraints on sub-keV dark matter from electron recoils in SuperCDMS Soudan

New particles beyond the standard model are often postulated to solve open questions in particle physics, generally related to the apparent fine tuning of model parameters or anomalous values outside the standard model expectation. In some cases, these particles may end up being good dark matter candidates. For instance, the axion was originally proposed as a solution to the strong CP problem but is also considered to be a good dark matter candidate [4, 36]. Unfortunately, the expected mass range of the axion that would solve the strong CP problem is too small to be detected by direct detection experiments that are sensitive to particle interactions [107], requiring dedicated axion search experiments to probe this parameter space [37]. However, there could exist other $U(1)$ symmetries in the universe (unrelated to the strong CP problem) that would give rise to axion-like particles (ALPs) [38]. ALPs correspond to the pseudo Nambu-Goldstone bosons of a broken $U(1)$ symmetry [38], and could have masses in a range that would be detectable by experiments like SuperCDMS [107]. In order to be a candidate for dark matter the mass of the ALP must be less than twice the electron mass, $m_a < 2m_e$, so that the abundance is not depleted through decays to electron-positron pairs. Another constraint on ALPs constituting the dark matter of the universe, is that their abundance

must not be depleted via decays to two photons, through a loop-level interaction [150]. A generic feature of ALPs is a tree-level coupling to electrons [40], so if the mass of ALPs is in the eV-keV range, they can potentially be detected by SuperCDMS through an absorption like process analogous to the photoelectric effect, often called the axioelectric effect [40]. A similar process exists for the absorption of dark photon dark matter [49], which has the same constraint on the mass being less than twice the electron mass, $m_V < 2m_e$, but are not subject to the constraint of decays to two photons [150].

The content of this chapter is largely similar to the published work found in Reference [107], and serves to provide additional context to certain aspects of the analysis. Using electron recoil data from four iZIP detectors and two CDMSlite detectors operated during SuperCDMS Soudan [107], constraints are set on the axioelectric coupling parameter g_{ae} of ALPs, and the kinetic mixing parameter ϵ of dark photons with standard model photons. The data from the CDMSlite detectors are more sensitive at recoil energies below about 10 keV, while the iZIP detectors provide a higher dynamic range, and sensitivity up to 500 keV. This chapter will describe the rate calculation used for ALPs and dark photons, the energy resolution models and detection efficiencies used for the iZIP and CDMSlite detectors, and finally the limit calculation procedure. Since data from multiple detectors were used in this analysis, a method of combining the data was implemented to ensure random fluctuations in the rate did not bias the limit downwards in regions where one detector measures a lower rate.

7.1 Description of the data

From 2012 to 2014, 15 iZIP detectors were operated in the SuperCDMS Soudan experiment [71]. Starting in 2012, individual detectors were operated in CDMSlite mode during SuperCDMS Soudan for three separate data series [151, 152, 153]. The first operation of a CDMSlite detector at Soudan, called Run 1, was a short run to demonstrate the working principle [151]. CDMSlite Run 2 [152] was a longer science run that used the same detector as Run 1, and the third CDMSlite run, Run 3 [153], used a different detector than Run 2 and had a slightly smaller exposure. This analysis uses data from CDMSlite Runs 2 and 3, and from four detectors operated in iZIP mode [107]. During CDMSlite

CHAPTER 7. CONSTRAINTS ON SUB-KEV DARK MATTER FROM ELECTRON RECOILS IN SUPERCDMS SOUDAN

Run 3, the phonon noise performance worsened due to a change in operating conditions, thereby degrading the energy resolution and requiring the separation of Run 3 into two separate periods [153]. Table 7.1 shows the exposures of the three CDMSlite data sets used in this analysis (Run 2 and the two Run 3 periods). The iZIP portion of the analysis uses data from the SuperCDMS Soudan low-mass WIMP search (see Reference [154]), which included data from 7 iZIP detectors. For this analysis, two of the detectors were excluded because they had shorts on one or more of the readout channels, and a third detector was not included since it showed evidence of incomplete charge collection on one side [107]. The exposures of the four iZIP detectors that were included in this analysis are shown in Table 7.2.

	Run 2	Run 3-1	Run 3-2
Exposure (kg·days)	70.1	31.5	29.4

Table 7.1: Exposures of the three CDMSlite data sets used in this analysis [107].

	T1Z1	T2Z1	T2Z2	T4Z3
Exposure (kg·days)	80.2	82.9	80.9	83.8

Table 7.2: Exposures of the four iZIP detectors used in this analysis [107].

7.2 Energy resolution

The energy resolution σ_T of the SuperCDMS detectors is modelled as [92, 107]:

$$\sigma_T(E) = \sqrt{\sigma_E^2 + \sigma_F^2(E) + \sigma_{PD}^2(E)}, \quad (7.1)$$

where E is the measured energy, σ_E is the resolution of the baseline noise, σ_F represents the uncertainty caused by the statistical fluctuations of the number of electron-hole pairs and accounts for the Fano factor. The term σ_{PD} is linear in energy and accounts for effects such as position dependence in the detector. The last two terms in the energy resolution model are energy dependent, and are parameterized as $\sigma_F = \sqrt{BE}$ and $\sigma_{PD} =$

7.2. ENERGY RESOLUTION

AE , where A and B are constants [92, 107]. The term σ_F^2 is linear in energy because the variation in the number of electron-hole pairs (N) is related to the Fano factor, so $\sigma_F^2 = FN\epsilon_\gamma^2 = F\epsilon_\gamma E$. Variations due to detector effects, such as position dependence, are expected to scale with energy, the σ_{PD} term also encapsulates any other effects that scale with energy [92, 107]. To determine the values of these model parameters, σ_E , A , and B , Gaussian functions are fit to the known peaks in the spectrum. The model given by Equation 7.1 is then fit to the measured widths of the relevant Gaussian peaks. For the CDMSlite detectors, the known peaks in the spectrum are the K-, L-, and M-shell electron capture peaks of ^{71}Ge , at energies of 10.37 keV, 1.30 keV, and 0.16 keV respectively [92, 107]. For the iZIP detectors, the known peaks used in the fit of the resolution model are: the ^{71}Ge K-shell at 10.37 keV, a 66.7 keV peak that appears from inelastic neutron scattering during the ^{252}Cf calibration, the 356 keV peak from ^{133}Ba calibrations, and a 511 keV peak that results from electron-positron annihilation. Table 7.3 shows the fitted model parameters (σ_E , A , and B) for the three CDMSlite data sets used in this analysis and Figure 7.1 shows the model with the points used to determine the fit. Table 7.4 shows the model parameters for the energy resolution of the four iZIP detectors used in this analysis, and Figure 7.2 shows the energy resolution model for T2Z2; the four iZIP detectors have very similar energy resolution models.

	σ_E (eV)	B (eV)	A ($\times 10^{-3}$)
Run 2	9.26 ± 0.11	0.64 ± 0.11	5.68 ± 0.94
Run 3-1	9.87 ± 0.04	0.87 ± 0.12	4.94 ± 1.27
Run 3-2	12.70 ± 0.04	0.80 ± 0.12	5.49 ± 1.13

Table 7.3: Fitted parameters for the energy resolution model for the three CDMSlite data sets used in this analysis [107].

	σ_E (eV)	B (eV)	A ($\times 10^{-3}$)
T1Z1	100.5 ± 3.4	2.7 ± 0.5	19.5 ± 0.1
T2Z1	69.9 ± 4.6	2.4 ± 1.1	15.4 ± 0.3
T2Z2	79.0 ± 2.9	1.2 ± 0.3	13.6 ± 0.1
T4Z3	80.2 ± 5.4	0.6 ± 1.2	19.0 ± 0.4

Table 7.4: Fitted parameters for the energy resolution model for the four iZIP detectors used in this analysis [107].

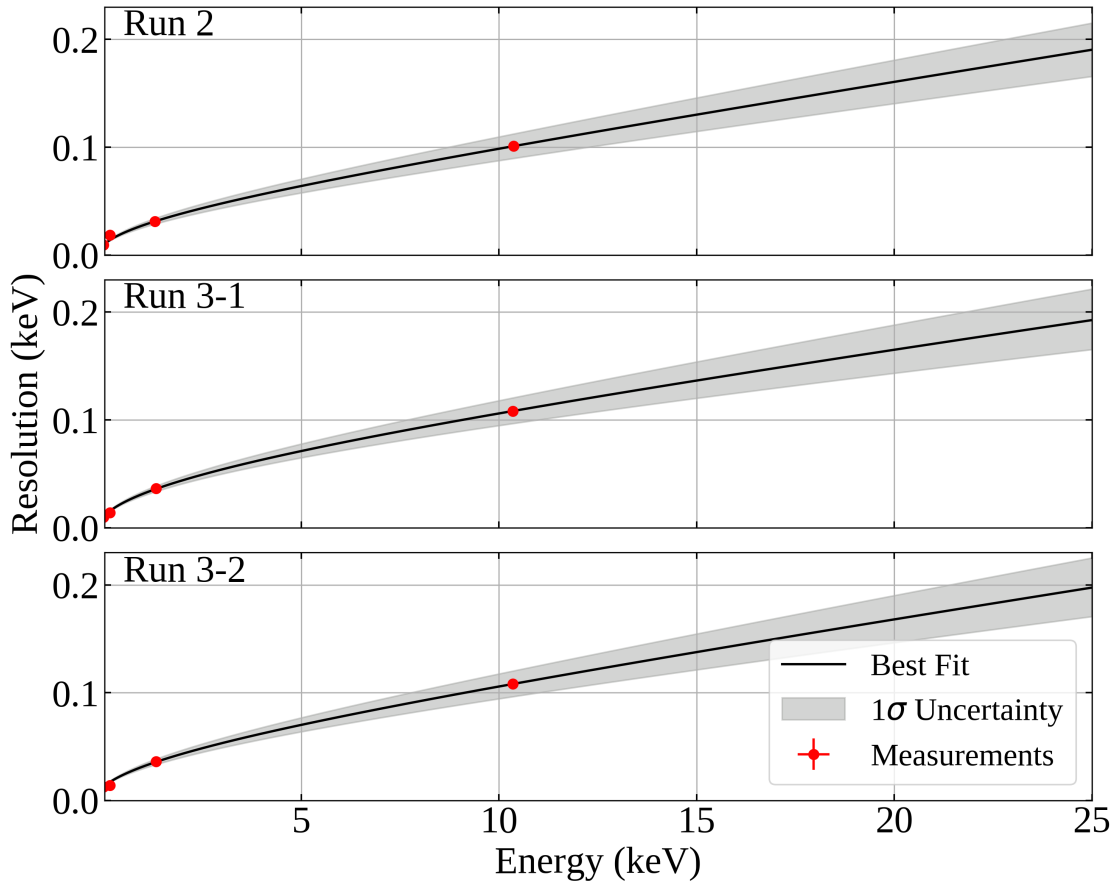


Figure 7.1: Energy resolution model for the three CDMSlite data sets used in this analysis. The fitted model is shown in black with a grey error band. The points that were used to determine the fit are shown in red. Image from Reference [107].

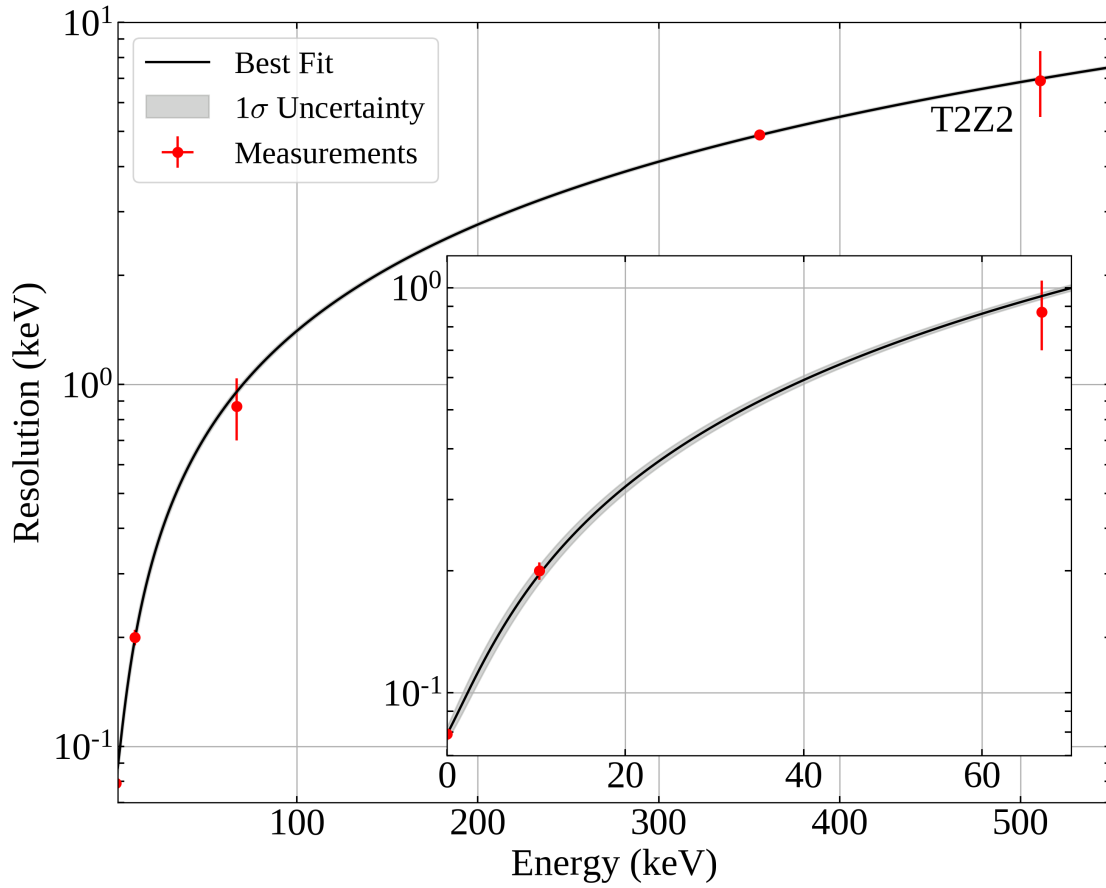


Figure 7.2: Energy resolution model for T2Z2, one of the iZIP detectors used in this analysis. Only one of the iZIP detectors is shown as all four iZIP detectors had qualitatively similar models. The fitted model is shown in black with a grey error band, while the red points were used to determine the fit. Image from Reference [107].

7.3 Efficiency model

When a signal in the detector produces a trigger, it is not necessarily because of a real, physical pulse, but could be related to a fluctuation in noise conditions. To remove unphysical triggers and poorly reconstructed events from the data sets, a number of selection criteria are applied to the raw traces. Events that pass these data quality selection criteria (cuts) are excluded from the analysis, at the cost of a loss in detection efficiency. The detection efficiency is the probability that a real particle interaction at a given energy is accepted into the final spectrum. The detection efficiency is zero below the trigger threshold, after which it typically rises to some constant value that is ideally just below unity but is generally much lower. In this section the detection efficiency curves for the CDMSlite and iZIP data sets will be described, without going into detail on the definitions of the underlying data selection criteria.

7.3.1 CDMSlite

The event selection criteria for CDMSlite Runs 2 and 3 are described in References [92, 153]. The data quality cuts used for the CDMSlite portion of the analysis were designed to remove triggers that were likely caused by noise, events with poor energy reconstruction, and time periods where the detector's behaviour was not as expected. Livetime cuts are any selection criteria that remove periods of time based on a certain condition. These include the removal of any periods where the high voltage power supply was improperly set, as well as time periods coincident with the NuMI neutrino beam that points at the MINOS experiment at the SUL [155], to remove events that were coincident in time with the neutrino beam. An additional livetime cut was produced by requiring that any trigger in the CDMSlite detector was not in coincidence with a trigger on another detector, since the expected rate of dark matter interactions is low enough that multiple triggers are not expected.

Data quality cuts remove events where the data may have some anomalous feature. One data quality cut removes events where the pre-trigger baseline current exceeds its time averaged value by more than four standard deviations, while another removes events that were likely triggered by excess electronic noise. In the triggered data, several classes

7.3. EFFICIENCY MODEL

of events with a peculiar pulse shape were identified. To remove these events, several glitch templates were generated and the events were fit to the regular pulse template as well as the glitch templates. The difference in fit quality (χ^2) was then used to determine if the event resembled a real pulse, or if it was more likely caused by noise. Finally, events that occurred close to the outer edges of the detector were removed using a fiducial volume cut. The removal of these events is important as they experience reduced NTL amplification, because of the inhomogeneous electric field in the detector, and are reconstructed at a lower energy. Additionally, the expected background around the edge of the detector is increased due to contamination from radioactive impurities.

The detection efficiency for CDMSlite Runs 2 and 3 is obtained using the livetime and data quality cuts described above, these efficiency curves are shown in Figure 7.3. The determination of the efficiency is non-trivial, and precise details can be found in References [92, 153, 102]. The low-energy portion of the efficiency curves are dominated by the trigger efficiency; the lowest observed energy in CDMSlite Run 2 was 56 eV, while in Run 3 it was about 70 eV. The efficiency of the fiducial volume cuts is determined by generating simulated pulses to understand the distribution of events, in terms of a parameter related to the radial distance from the center of the detector to the event site [102]. Noise triggers in Run 3 were reduced by the low-frequency noise cut (glitch cut) [153], however this had the effect of reducing the efficiency and effectively raising the trigger threshold. The efficiency curve used in the original CDMSlite Run 2 WIMP analysis was only calculated up to 2 keV, but was extended up to 30 keV for the purposes of a background study [156]. The extension of the efficiency curve was done by interpolating between the measured efficiencies at 2 keV and 10.37 keV (the Ge K-shell energy) [156, 107]. For energies above 10.37 keV, electron recoils from the ^{133}Ba calibration were compared to simulations, and it was found that the efficiency dropped above 17.3 keV [156], an effect that was attributed to saturation in the phonon sensors. The Run 3 efficiency curve was extended in the same fashion as the Run 2 curve [107].

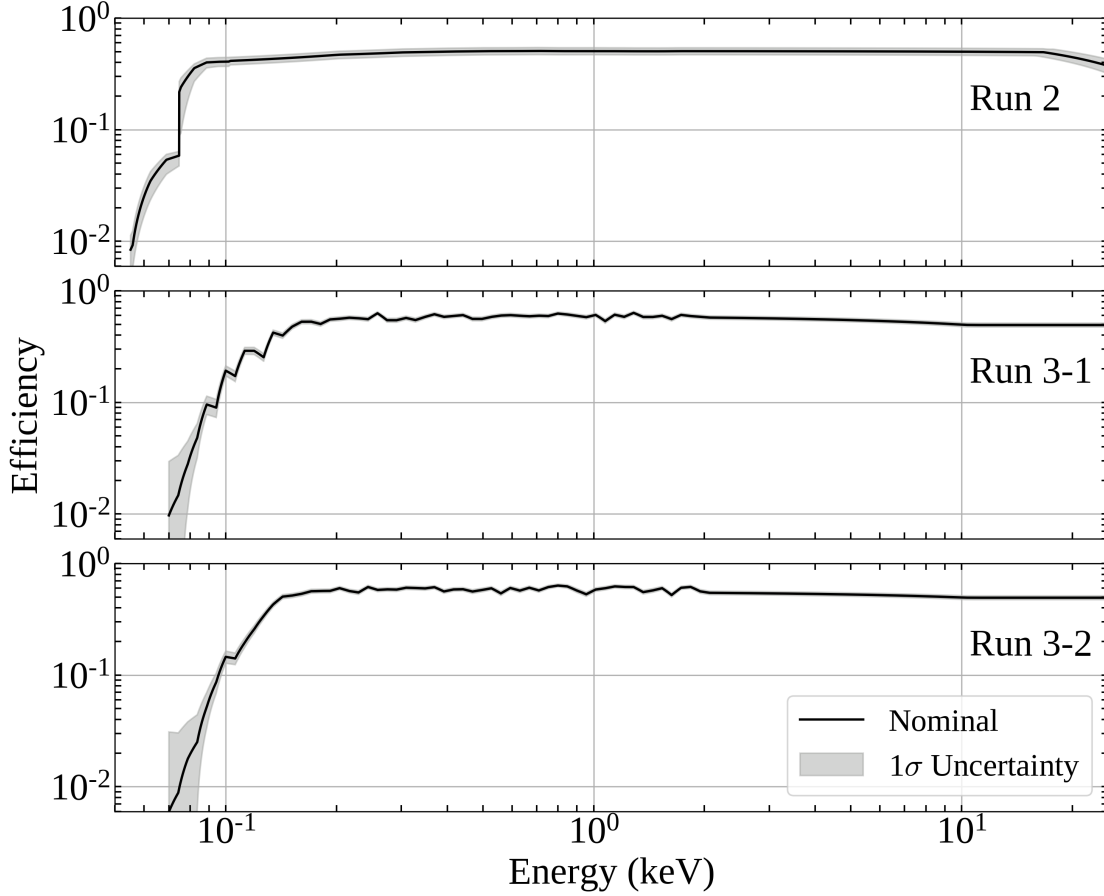


Figure 7.3: Efficiency curves (black) with uncertainty bands (grey shaded) for CDMSlite Run 2, and the two periods of CDMSlite Run 3. Image from Reference [107].

7.3.2 iZIP

The selection criteria for the iZIP portion of the analysis used the same definitions as in Reference [154] for the data quality cuts, the trigger cut, the single-scatter cut, and the muon veto cut. The data quality cuts remove events with elevated noise in the pre-trigger baseline that is inconsistent with normal periods, while also removing events with a pulse shape that is poorly fit to the template (based on the χ^2). The trigger cut requires that the event issued a trigger on the detector in question, while the single-scatter cut requires that it did not also issue a trigger on any other detector in the same time window. The muon veto cut removes any events that are coincident with a trigger in the muon veto detector, the same as was done in the CDMSlite portion of the analysis [107]. These four criteria have a combined efficiency of close to 95% for each detector, and are slightly energy dependent near the trigger threshold.

7.3. EFFICIENCY MODEL

While the four data selection criteria described above are the same as in Reference [154], there are a few differences between the analyses that must be accounted for. Since ALPs and dark photons are expected to interact with the electron orbitals, the electron recoil event band is used instead of the nuclear recoil band. As the original fiducial volume cut for the WIMP analysis was defined for nuclear recoils, and the efficiency was determined using events from the ^{252}Cf neutron calibration data set, the cut needed to be redefined for electron recoils. The main purpose of the fiducial volume cut is to remove the surface-event background caused by ^{210}Pb decays, and remove events that occur near the edge of the detector which have reduced NTL amplification because of the distorted electric field [92, 153, 102].

Since this analysis looks for point-like electron recoils, the fiducial volume cut should be redefined using a population of these events. One such population are the Ge K-shell events, which are expected to be distributed more or less homogeneously throughout the detector [107]. Since these events are mono-energetic (10.37 keV), they cannot be used to measure the energy dependence of the cut, so the cut is defined to be largely energy independent, and the efficiency is measured at the Ge K-shell energy.

The fiducial volume cut is defined using the charge signal in the detector, and consists of two separate parts: a charge symmetry cut and a radial charge cut [107]. Events that occur in the bulk of the detector are expected to have symmetric charge collection on both sides of the detector, while events that occur near one of the flat faces of the detector will have a larger fraction of the total charge collected on that side. The charge symmetry cut removes events that occur closer to the flat surfaces of the detector by requiring that the charge signal amplitude is similar on both sides. The radial charge cut removes events that occur near the outer edges of the detector that may experience reduced NTL amplification due to the inhomogeneous electric field in this part of the detector. By defining the fiducial volume cuts in this way, events that have improper energy reconstruction are rejected, as well as events that are likely caused by backgrounds.

The charge symmetry parameter is defined as the difference between the total charge collected on each side, divided by the total charge collected on both sides. The distribution of this parameter with energy is widest at low energies, and tapers off at higher

CHAPTER 7. CONSTRAINTS ON SUB-KEV DARK MATTER FROM ELECTRON RECOILS IN SUPERCDMS SOUDAN

energies. For energies below the Ge K-shell energy, the mean and standard deviation of the cut parameter is determined in 1 keV bins, and an exponential function is fit to the mean plus three standard deviations. For energies above the Ge K-shell energy, the cut line is set at the value of the cut parameter measured at the Ge K-shell energy [107]. Since the distribution tapers off at higher energies the efficiency will actually be slightly higher than the measured efficiency at 10.37 keV, but without a way to measure the efficiency at these energies, setting the efficiency to be constant above the Ge K-shell energy is conservative [107]. For energies below the Ge K-shell energy, the number of events that fall outside of the cut line is small, and so the assumption of constant efficiency is conservative.

The radial charge cut removes events where the charge collected in the outer channels exceeds a certain fraction of the total charge. Events are removed where the charge signal in the outer electrodes exceeds three times the baseline (charge) resolution of that sensor. This fraction is determined from Ge K-shell events that appear at the correct energy in the inner sensors [107]. This definition of the radial charge cut provides an energy independent efficiency above the Ge K-shell energy, since the distribution of the signals in inner and outer channels is independent of energy for point-like events occurring at a given position in the detector. Due to noise in the measurement, low-energy events with very little charge collected in the outer channel have a chance to be reconstructed above the cut limit, and therefore fail the cut. This effect is accounted for by modelling the signal distribution between inner and outer channels for Ge K-shell events, and scaling the model to lower energies and convolving it with the applicable charge resolution model [107].

The efficiency of the newly defined charge fiducial volume cuts is determined using the Ge K-shell events. The relatively short half-life (11.43 days) of ^{71}Ge is exploited by using a time period of data collected directly after the ^{252}Cf calibrations [107]. Events are selected where the recoil energy falls within $\pm 4\sigma_T$ of the expected Ge K-shell energy (10.37 keV), where σ_T is the phonon energy resolution at 10.37 keV. The charge fiducial volume cuts are applied, and the passage fraction is determined by comparing the fraction of events that passed over the total number of Ge K-shell events. The combined

7.4. RATE CALCULATION

efficiency of the two charge fiducial volume cuts falls between 30% and 36%, depending on the detector in question [107].

After combining the data quality, anti-coincidence, muon-veto, and charge fiducial volume cuts, the total efficiency above threshold is around 35%, varying somewhat between detectors [107]. Note that this efficiency is dominated by the efficiency of the charge fiducial volume cuts. An example of the efficiency curve for one of the iZIP detectors (T2Z2) is shown in Figure 7.4; the other iZIP detectors have slightly different efficiency curves, but the general shape is qualitatively similar.

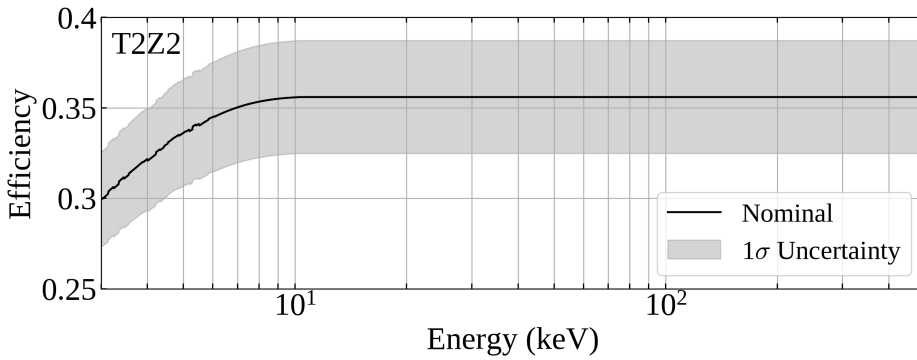


Figure 7.4: Efficiency curve (black) with uncertainty bands (grey shaded) for T2Z2, one of the iZIP detectors used in this analysis. The efficiency of roughly 30% is dominated by the efficiency of the charge fiducial volume cuts. Image from Reference [107].

7.4 Rate calculation

This analysis sets constraints on the axioelectric coupling of ALPs and the kinetic mixing parameter of dark photons by assuming that the respective particle constitutes the entirety of the dark matter in the universe. The underlying particle interaction that the SuperCDMS detectors are sensitive to is a process where the dark photon or ALP is wholly absorbed, transferring its total energy to an electron in the crystal [49, 40]. If the dark matter is made up of ALPs or dark photons with a definite mass, then the signal that will be produced in the detector is a Gaussian peak at the particle’s mass (since dark matter is non-relativistic and the kinetic energy is negligible) with a width equal to the detector’s energy resolution.

The interaction rate R in the detector is given as the product of the flux of dark

CHAPTER 7. CONSTRAINTS ON SUB-KEV DARK MATTER FROM ELECTRON RECOILS IN SUPERCDMS SOUDAN

matter Φ , by the interaction cross section σ :

$$R = \Phi\sigma. \quad (7.2)$$

The flux is given by:

$$\Phi = \frac{\rho_\chi v_\chi}{m_\chi}, \quad (7.3)$$

where ρ_χ is the local density of dark matter, v_χ is the dark matter velocity, and m_χ is the dark matter particle's mass. The interaction cross section for both ALPs and dark photons is proportional to the photoelectric cross section σ_{pe} of the target, in this case germanium. For ALPs, the interaction cross section σ_a is given by [40]:

$$\sigma_a(E_a) = \sigma_{pe}(E_a) \frac{g_{ae}^2}{\beta_a} \frac{3E_a^2}{16\pi\alpha m_e^2 c^4} \left(1 - \frac{\beta_a^{2/3}}{3}\right). \quad (7.4)$$

In this equation E_a is the total energy of the ALP, β_a is the ratio of the ALP's velocity to the speed of light c , and m_e is the electron mass. Given that the dark matter is non-relativistic ($E_a = m_a c^2$ and $\beta_a \ll 1$), the interaction rate for ALPs is given by [40]:

$$R_a(E_a) = \rho_\chi \sigma_{pe}(E_a) \frac{3g_{ae}^2 m_a c}{16\pi\alpha m_e^2}. \quad (7.5)$$

For dark photons the interaction cross section σ_V is given by [49]:

$$\sigma_V(E_V) = \sigma_{pe}(E_V) \frac{\epsilon^2}{\beta_V}, \quad (7.6)$$

where E_V is the energy of the dark photon and β_V is the dark photon's relativistic beta factor. Similarly as above, the interaction rate for dark photons is [49]:

$$R_V(E_V) = \frac{\rho_\chi}{m_V} \epsilon^2 \sigma_{pe}(E_V) c. \quad (7.7)$$

7.4.1 Photoelectric cross section

The germanium target's photoelectric cross section, which the interaction rate for dark photons and ALPs depends on, was obtained by performing a literature search in the relevant energy range (40 eV-500 keV) [107]. The photoelectric cross section for germanium is expected to be mostly independent of the temperature, so measurements performed at higher temperatures were used [157]. There existed some discrepancies in the reported cross section values in the literature, in particular for energies below about 1 keV, so to

7.4. RATE CALCULATION

account for this uncertainty in the analysis, two curves were produced: a nominal and conservative photoelectric cross section. The construction of the nominal photoelectric cross section followed the procedure used in Reference [157], while the conservative photoelectric cross section was produced by taking the smallest photoelectric cross section values at each energy found in the literature [107]. By constructing the conservative photoelectric cross section in this fashion, the result obtained for the upper limit of g_{ae} and ϵ using the conservative cross section will be higher (more conservative) than one obtained using the nominal cross section. Figure 7.5 shows the nominal and conservative photoelectric cross sections, as well as the difference between the two curves in percent.

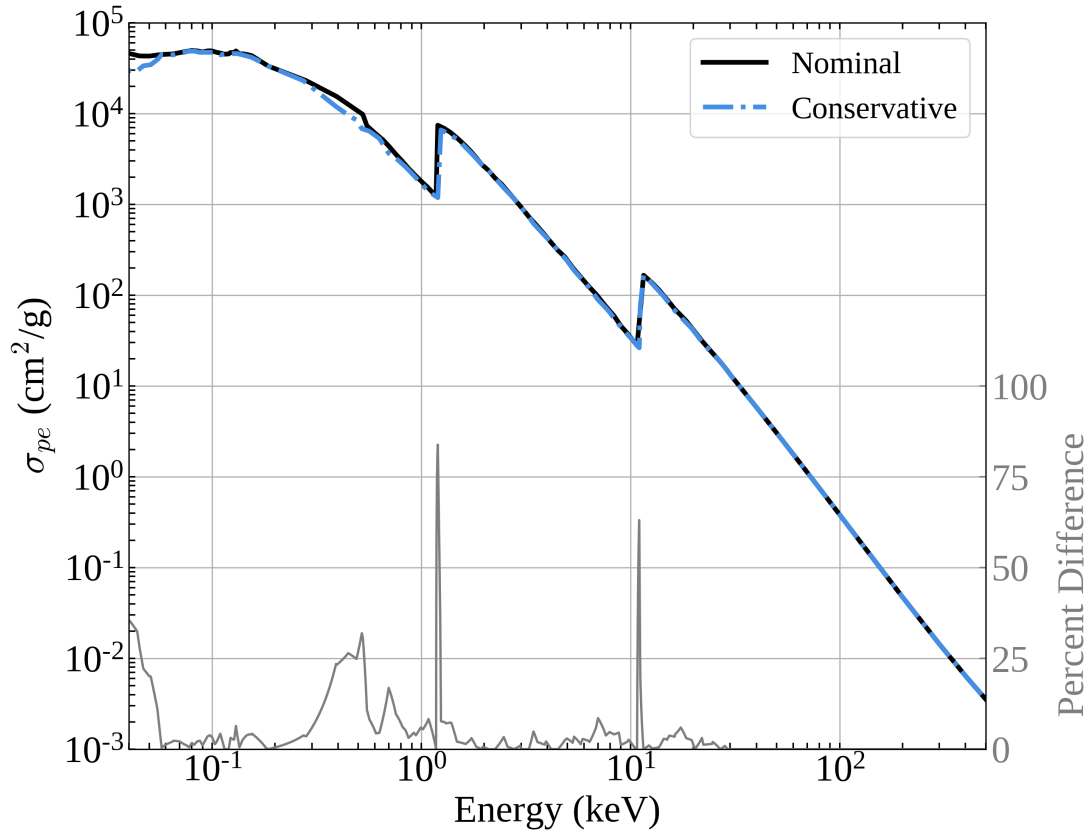


Figure 7.5: Photoelectric cross section of germanium. The nominal cross section produced following Reference [157], while the conservative curve is produced from the smallest values found in the literature search [107]. Image from Reference [107].

7.5 Limit setting

No background modelling was performed for this analysis, meaning that no detection of a dark matter signal can be claimed. Instead, an upper limit on the interaction strength is set by determining the signal rate (for a particular dark matter mass) that would exceed the measured rate significantly. An alternative way to consider this is to assume that all of the observed events might be caused by dark matter interactions, and test for the interaction strength of the dark matter for different dark matter masses that would produce the same measured spectrum in 90% of random experiments. Note that these two statements are equivalent and simply different ways to phrase the limit setting technique. It should be noted that limit setting techniques that do not include background modelling, such as the maximum gap or optimum interval methods [146, 147], will always set worse (higher) limits than techniques that do incorporate background modelling, such as profile likelihood methods.

The number N of dark matter interaction events in the detector is related to the interaction rate of dark matter R_χ , the exposure of the experiment, and the detection efficiency η :

$$N = MT \int \eta(E) R_\chi(E) dE. \quad (7.8)$$

Here the exposure is defined as the product of the mass of the detector M and the livetime T of the experiment. The dark matter rate R_χ in this case is either the rate of ALPs or dark photons, defined above in Equations 7.5 and 7.7 respectively. Since all of the terms in the dark matter interaction rate are known except for the coupling strength (g_{ae} or ϵ), the number of events seen in the detector can be used to constrain the parameters of interest.

Since the dark matter signal in question for this analysis is a Gaussian, a simple way to set a limit is by counting the number of events within a particular energy range, and calculating the 90% Poisson upper limit on the parameter of interest. The choice of the window size must be a good compromise between maximizing the signal (larger window), while minimizing the background (smaller window). The optimal window size depends on the event rate, so to avoid overtuning the window size based on the actual spectrum,

7.5. LIMIT SETTING

a window of $\pm 1\sigma$ of the dark matter mass being tested was used, where σ is the energy resolution of the detector in question at that particular energy.

At energies near the trigger threshold, the event selection efficiency drops quickly to zero, as was seen in Section 7.3. As such, the expected signal shape is no longer Gaussian, but instead skews towards higher energies where the efficiency is greater. For these cases, simply setting the upper edge of the window at $+1\sigma$ may result in a significant portion of the signal being excluded. To combat this effect for energies near the trigger threshold, the upper edge of the window is determined from the expected signal shape, instead of the Gaussian signal shape. At these energies, the upper edge of the window is set to the $+1\sigma$ -equivalent, defined as the point above which 15.9% of the signal would be found, the same that is found above the $+1\sigma$ position for a Gaussian [107]. This definition of the low-energy window is implemented for dark matter masses below 100 eV for the CDMSlite Run 2 data set, and below 200 eV for the two CDMSlite Run 3 data sets. For the iZIP portion of the analysis, an analysis threshold of 3 keV is imposed (which is well above the trigger threshold of about 1 keV), so the described effect is small and the standard window definition of $\pm 1\sigma$ is used throughout.

For dark matter masses corresponding to energies below the trigger threshold, the positive tail of the expected event distribution (in particular the portion of the Gaussian tail that overlaps with the non-zero detection efficiency) still provides some sensitivity. Since the efficiency estimate in these energy regions is strongly affected by potential non-Gaussian tails and uncertainties on the energy resolution, a lower analysis threshold on the CDMSlite data sets is chosen to be roughly 2σ below the trigger threshold.

For the CDMSlite data sets, an upper analysis threshold of 25 keV is set to avoid saturation effects caused by the large intrinsic NTL amplification. For the iZIP portion of the analysis, an upper analysis threshold of 500 keV is selected, which is just below the highest peak (511 keV) used in the calibration.

7.5.1 Combination of data sets

When setting a limit with multiple data sets, some consideration must be taken in combining the information from the different data sets. The simplest method would be to

CHAPTER 7. CONSTRAINTS ON SUB-KEV DARK MATTER FROM ELECTRON RECOILS IN SUPERCDMS SOUDAN

combine the measured rates from all detectors, however this will result in a worse limit as some detectors may have a truly higher rate from backgrounds. Another approach would be to use the best limit from a single detector for each dark matter mass being tested. This would set the best limit by definition, but downward statistical fluctuations in the rate will introduce a bias in the limit. To reduce the sensitivity to downward statistical fluctuations while also excluding detectors that see a truly higher (energy-dependent) rate, a method for combining rates from multiple data sets was developed [107].

First, limits for a set of discrete dark matter masses are calculated for each of the seven data sets (three CDMSlite, four iZIP). For each of these mass points, the following procedure is used:

1. Select the data set with the lowest measured rate.
2. Discard data sets in which the difference between the observed rate and the lowest measured rate exceeds three times the uncertainty on the difference (three standard deviations).
3. Include any data sets that did not fail the previous requirement.

Since the number of data sets is low, the requirement imposed to discard data sets will seldom occur if the detectors truly have the same rate. A Monte Carlo simulation was used to quantify this effect, and it was found that for detectors with identical rates, data sets are discarded less than one percent of the time due to statistical fluctuations [107].

After determining which data sets to include for each dark matter mass, the upper limit on the rate is set by counting the number of events in a window centered on the dark matter mass-energy being tested, using the events from all included data sets. The 90% Poisson upper limit on this number n_{90} is then calculated. If this number were the true value, then the measured count or less would be seen in 10% of random trials. The 90% Poisson upper limit for a given number n is the point on a Poisson cumulative distribution function (that has an average of n) that corresponds to 0.9.

7.5.2 Uncertainty band estimate

For each dark matter mass used in the calculation, the nominal limit is calculated using the nominal value of the resolution and efficiency model [107]. An upper (lower) bound on the uncertainty is generally determined by using the upper (lower) limit on the resolution and the lower (upper) limit on the efficiency. However, for the low-energy portion of the CDMSlite analysis (where the dark matter mass is below the energy threshold) a wider resolution may result in a better limit, since more of the signal would appear above threshold. At these masses, the lower and upper bounds of the uncertainty on the limit are determined by calculating the limit for all combinations of upper and lower limits on the resolution and efficiency, and assigning the highest (lowest) calculated limit as the upper (lower) bound of the limit curve.

7.6 Results

Using the seven data sets, three combined rate limits with uncertainty bands were produced: a combined iZIP limit from the four iZIP detectors, a combined CDMSlite limit from the three CDMSlite data sets, and a combined iZIP and CDMSlite limit in the energy region where the two data sets overlap (3-25 keV) [107]. These three limits on the rate are shown in Figure 7.6. In the overlap region between the CDMSlite and iZIP limits, the CDMSlite data sets generally have a better limit. This is to be expected since the energy resolution of the CDMSlite detectors is better than that of the iZIP detectors by roughly a factor of 2. In addition, the iZIP data sets were collected during the first measurement periods of SuperCDMS Soudan, whereas the CDMSlite data sets were acquired towards the end, roughly two years afterwards. As such, the intrinsic detector backgrounds from cosmogenic activation, such as ^{65}Zn at 9 keV and ^{55}Fe around 6 keV, were able to decay to lower levels [107]. However, the method of combining data sets (described above) still included the iZIP data sets for most of the mass points where the CDMSlite and iZIP data sets overlap. Around 7 keV there is a point in the combined limit that is lower than both the CDMSlite and iZIP limits, since the rate there is comparable and the gain in exposure results in a lower combined limit.

After producing the full combined limit on the rate for the entire energy range, the

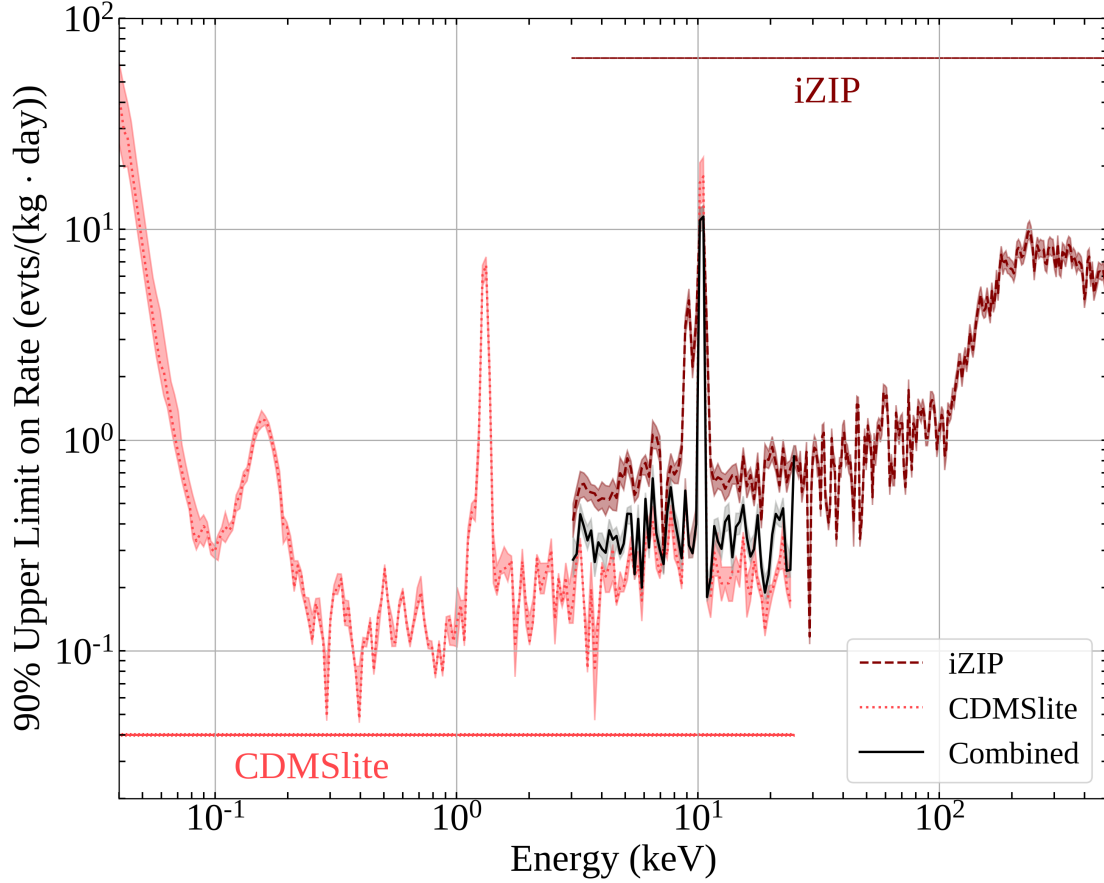


Figure 7.6: 90% upper limits on the rate with uncertainties on the combined CDMSlite data sets, the combined iZIP data sets, and the combined iZIP and CDMSlite data sets. Image from Reference [107].

limit on the rate can be converted to a limit on the parameters of interest g_{ae} and ϵ using Equations 7.5 and 7.7 respectively. Since the conservative photoelectric cross section will always result in a worse limit on the parameter of interest than the nominal photoelectric cross section, the upper bound of the limit curve is calculated using the nominal photoelectric cross section, and the nominal and lower bound of the limit curve is calculated using the nominal photoelectric cross section. At the lowest energies, the dominant source of uncertainty in the limit is the uncertainty on the photoelectric cross section. The limits on the parameters of interest, g_{ae} and ϵ , are shown in Figures 7.7 and 7.8 respectively, alongside constraints set by other direct detection experiments and astrophysical and cosmological observations.

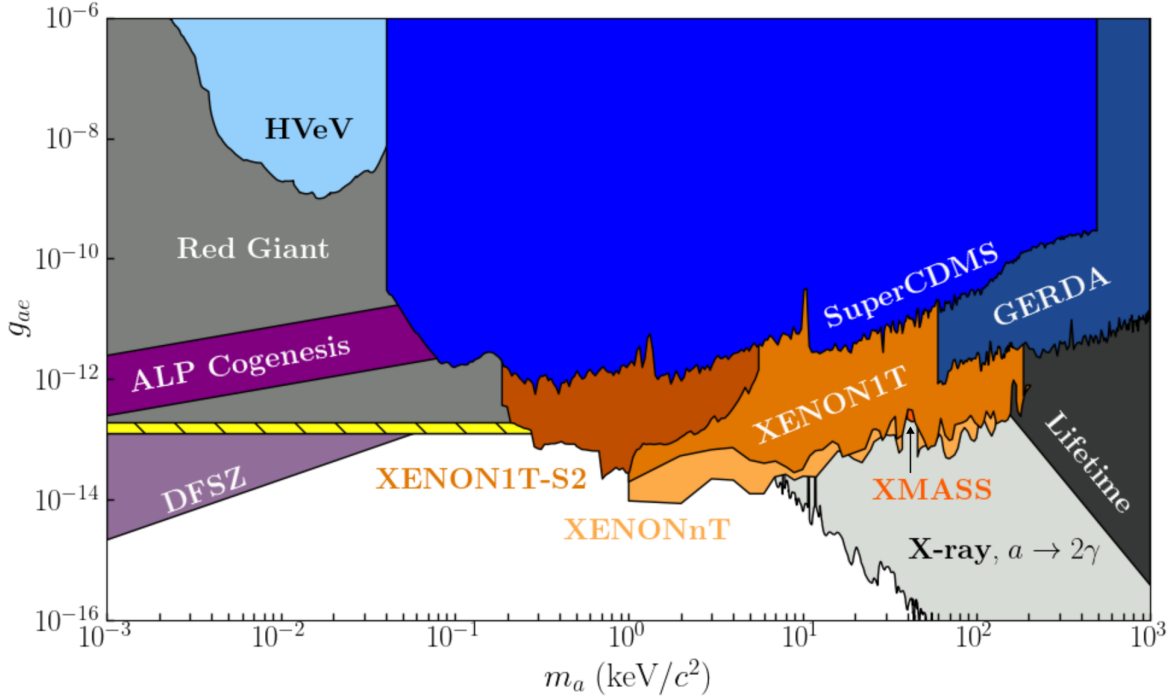


Figure 7.7: Astrophysical and direct constraints on the axioelectric coupling parameter g_{ae} . Astrophysical constraints [158, 150] are shown in grey, along with the parameter space corresponding to the hint of new physics determined by excess cooling of white dwarf stars [159] in yellow. Direct detection experiments that use semiconductor technologies [93, 107, 160] are shown in shades of blue, while xenon based technologies [83, 161, 162, 163] are shown in shades of orange. Shown in purple are two regions of parameter space that correspond to the correct relic abundance of axions/ALPs in specific cosmological models [39, 164].

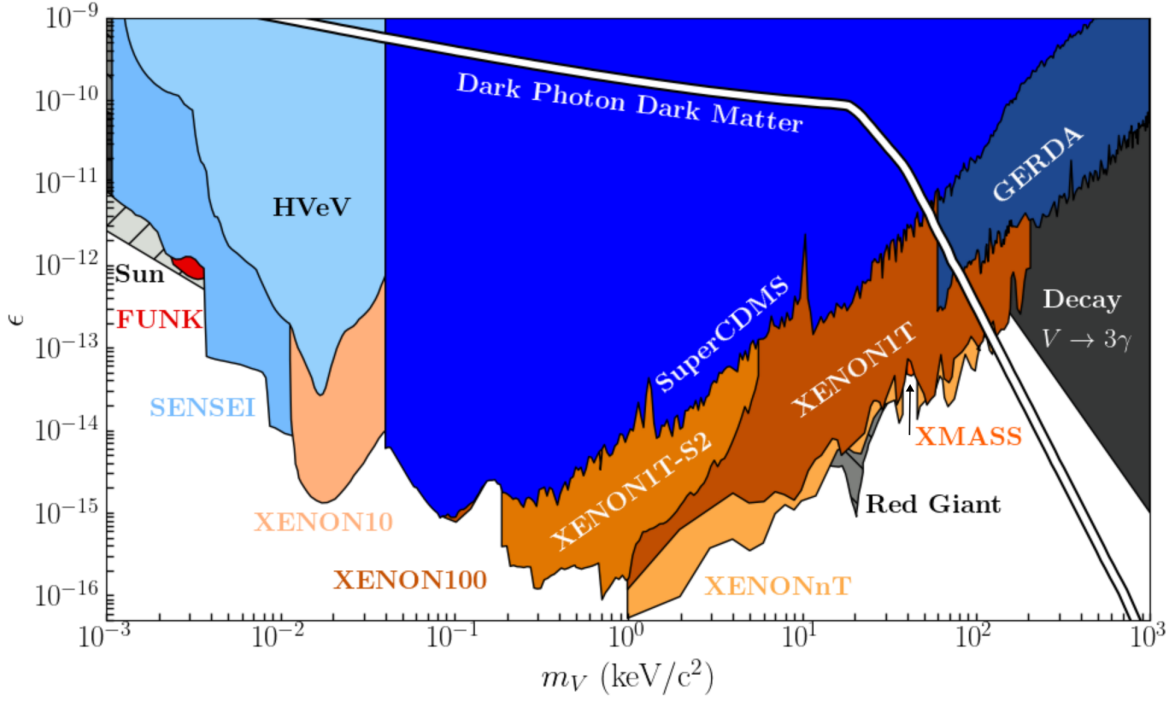


Figure 7.8: Constraints on the kinetic mixing parameter ϵ of dark photons, set by direct detection experiments and astrophysical constraints. Experiments that use semiconductor detectors [98, 93, 107, 160] are shown in shades of blue, while experiments using xenon based detectors [165, 83, 161, 162, 163] are shown in shades of orange. Astrophysical constraints are shown in shades of grey [165]. The region below the thick white line corresponds to the parameter space where dark photons could reasonably make up the entirety of the dark matter in the Universe. This limit was compiled by [46, 166] based on a number of different constraints [167, 168, 169, 170, 171, 172, 173].

7.7 Discussion

In this section, the results of this analysis are discussed in the context of astrophysical and cosmological constraints, as well as constraints obtained from other direct detection experiments. Many of the constraints discussed are obtained from a maintained repository¹ [166].

7.7.1 ALP constraints

Astrophysical and direct constraints on g_{ae} can be seen in Figure 7.7. The constraints on g_{ae} set using data from SuperCDMS Soudan are the best direct detection constraints for ALP masses between 40 eV and 200 eV. Below 40 eV down to 1.2 eV, the best direct detection constraints come from the SuperCDMS HVeV detector [93]. Due to the high NTL amplification of the HVeV devices, individual electron-hole pairs can be resolved, and the effective trigger threshold is roughly 0.4 eV, enabling the couplings of ALPs with order eV masses to be probed. While the low-threshold SuperCDMS technology is able to probe low mass ALPs, the parameter space probed by SuperCDMS below 1 keV is disfavored by constraints set from the brightness of red giant stars [158], indicated by the shaded grey region in Figure 7.7. Just below the red giant constraint in Figure 7.7, is a yellow patch that corresponds to excessive energy losses in white dwarfs that could be explained by an axion/ALP [159]. While there is no definite axion mass associated with this hint, it sets a clear target goal on the coupling of $g_{ae} \approx 10^{-13}$.

For ALP masses above 200 eV, constraints set by liquid xenon experiments, such as XENON1T [83, 161], XENONnT [163], and XMASS [162], exceed the SuperCDMS Soudan constraints [107] by roughly an order of magnitude due to their large exposures and low background rates. For masses above 200 keV, the best direct detection constraints are set by GERDA [160], an experiment searching for neutrinoless double beta decay using high purity germanium detectors. It should be noted that for ALP masses above 6 keV, the parameter space of g_{ae} constrained by all direct detection experiments is excluded by even more stringent constraints set by the decay of ALPs to two photons using X-ray and gamma-ray searches [150]; these constraints are labelled as “X-ray, $a \rightarrow 2\gamma$ ” in Figure

¹<https://github.com/cajohare/AxionLimits>

CHAPTER 7. CONSTRAINTS ON SUB-KEV DARK MATTER FROM ELECTRON RECOILS IN SUPERCDMS SOUDAN

7.7. The same decay mechanism (namely $a \rightarrow 2\gamma$) can also be used to set constraints on ALP dark matter by requiring the lifetime exceeds the age of the universe, although this constraint is weaker than those set by X-ray searches; this constraint is labelled as “Lifetime” in Figure 7.7.

In addition to the astrophysical and direct constraints shown in Figure 7.7, two shaded purple regions are shown which give the corresponding parameter space of the correct dark matter relic abundance, for specific cosmological models. The region labelled “ALP Cogenesis” corresponds to a range of parameters where ALPs would simultaneously solve the baryon asymmetry of the Universe while also making up the dominant component of dark matter [39]. For this patch it is assumed that the coupling of the ALP is generated by one-loop radiative corrections, arising from new heavy fermions that are charged under the new $U(1)$ symmetry [39]. While this is not the only region of parameter space for which ALPs could simultaneously resolve the baryon asymmetry and dark matter problems, there is value to probing this space with direct detection experiments, as there may be some kind of chameleon-mechanism that can circumvent the stellar bounds, but not the direct detection constraints [174]. The shaded purple region labelled as “DFSZ” corresponds to the relic density of the QCD axion produced through the misalignment mechanism. Specifically, the DFSZ model [164] introduces a new complex scalar field (the axion) that couples to an extended Higgs sector. While the mass of the QCD axion is required to be below $1.6 \cdot 10^{-2}$ eV by supernova 1987a [175, 176], the relic abundance patch is displayed for reference. Additionally, there is a lower bound of about 10^{-6} eV on the mass of the QCD axion set by requiring that the present day cosmological energy density of the axion does not exceed the critical density of the universe [177, 176]

7.7.2 Dark photon constraints

Figure 7.8 shows astrophysical and direct constraints on the dark photon’s kinetic mixing parameter. The constraints set by SuperCDMS Soudan are world-leading for dark photon masses between 40 eV and 80 eV [107]. Below 40 eV, the primary direct constraints on dark photon dark matter come from experiments with sensitivity to individual charge carriers, like SENSEI [98], SuperCDMS HVeV [93], and XENON10 [165]. In addition to

7.7. DISCUSSION

these canonical direct detection experiments, the FUNK experiment [178] sets constraints in the mass range of (2-3.5) eV on dark matter dark photons by using PMTs to search for the conversion of dark photons into standard model photons at the surface of a mirror. While FUNK [178] and SENSEI [98] are the most sensitive direct searches for dark photon masses below about 10 eV, the most stringent constraints on the dark photon’s kinetic mixing parameter below about 3.5 eV comes from modelling stellar energy loss in the Sun [165].

For dark photon masses above 200 eV, the best constraints come from liquid xenon experiments, such as XENON1T [83, 161], XENONnT [163], and XMASS [162], and the neutrinoless double beta decay experiment GERDA [160]. There is also a small range of masses between 15 keV and 30 keV where the best constraints come from the stellar cooling of red giants [165]. For dark photon masses above about 150 keV, all direct detection constraints are surpassed by a limit determined from the decay of dark photons to three photons, which would deplete the relic abundance of dark photon dark matter [165]; this patch is labelled as “Decay, $V \rightarrow 3\gamma$ ” in Figure 7.8. The thick white line (outlined in black) in Figure 7.8 is the upper limit for which dark photons could make up the entirety of the dark matter in the Universe. This limit results from several cosmological and astrophysical considerations of dark photon dark matter, such as that the dark photon condensate produced through the misalignment mechanism should be sufficiently long lived so as to still exist today [167]. Even if the condensate survives until the present epoch, it is possible that some evaporation of it in the early universe may occur, which could dump energy into the photon bath and lead to distortions of the CMB and small modifications to the effective number of relativistic neutrino species [167]. Finally, the decay of dark photons to three photons through an electron loop should be less than the diffuse X-ray background [167]. Additional constraints come from astrophysical considerations regarding the heating of the intergalactic medium [172] and the heating of gas-rich dwarf galaxies [173]. These and various other considerations [168, 169, 170, 171] were compiled by [46, 166] into this bound on dark photons making up the entirety of the dark matter in the universe.

Chapter 8

Conclusion

An abundance of evidence supports the idea that 85% of the matter in the universe is made up of non-baryonic matter, called dark matter [1]. While no direct observation of dark matter has been confirmed yet, the leading view is that dark matter is comprised of one or more new particles beyond the standard model. This view is supported by several open questions within the standard model of particle physics, such as the anomalous magnetic moment of muons [3], and the yet unobserved electric dipole moment of the neutron [4]. Proposed solutions to these open questions generally involve the introduction of new symmetries and mechanisms that result in new particles, some of which could be the dark matter of the universe [4, 42]. A variety of different experimental techniques have been proposed and performed to search for signatures of new physics. This thesis considers the detection potential of a direct detection experiment that use cryogenic solid-state detectors, specifically, the SuperCDMS experiment [11].

SuperCDMS is a direct detection experiment that uses cryogenic semiconductor (germanium and silicon) detectors instrumented with transition edge sensors. The first phase of the SuperCDMS experiment took place at the Soudan Underground Laboratory in Northern Minnesota [71]. The detectors operated at SuperCDMS Soudan yielded competitive constraints for several different dark matter particles, including WIMPs [71, 92]. Using the data from SuperCDMS Soudan, searches for non-WIMP dark matter candidates were performed, such models include dark photons and axion-like particles (ALPs) that would be wholly absorbed, ejecting an electron in the process [40, 49]. The excel-

lent resolution of SuperCDMS detectors, in particular the low-threshold achieved by the CDMSlite operating mode provided sensitivity to dark photon and ALP masses as low as $40 \text{ eV}/c^2$ [92, 107]. The constraints on sub-keV mass dark photons and ALPs were set using data obtained from both the iZIP and CDMSlite detector operating modes [107]. At the time of publication these results were world leading and they remain competitive to this day; more details can be found in Chapter 7. My contributions to this work included re-analyzing the data with the new signal model, and implementing a new method of calculating the efficiency for sub-threshold dark matter masses.

Preparations are currently underway for the next phase of SuperCDMS, which will be located at SNOLAB in Sudbury, ON. A variety of SuperCDMS detectors have been tested in the CUTE facility since 2019, before the construction of SuperCDMS SNOLAB is complete. Operating detectors in CUTE prior to the commissioning of SuperCDMS SNOLAB is beneficial as it provides an opportunity to test the read-out chain while minimizing the exposure of the detectors to cosmogenic activation. The devices tested so far in CUTE -which include a Soudan iZIP detector, a SNOLAB HV detector, and a CPD- have provided valuable information about the performance of the facility and detectors, as demonstrated in Chapter 5. Some of my main contributions to the work done at CUTE include: the design and construction of the cold hardware needed to mount a SuperCDMS detector tower, the commissioning of the dilution refrigerator and suspension system on surface and underground, the development of the slow-control system that monitors the environmental parameters of the facility, and aiding in the day-to-day operations required to collect and analyze the data from the various devices operated in the facility.

The tests at CUTE confirmed that the facility background matched the expected background rate, and the facility is now prepared to test the SuperCDMS SNOLAB towers. For the detectors that will be used in the main experiment, it is critical to minimize the cosmogenic activation in order to stay within the allocated background budget. These tests at CUTE will provide the opportunity to perform characterization studies, which would not otherwise be possible on surface. It is expected that a full SuperCDMS SNOLAB tower will be operated for several months sometime in the coming year (2023), and with these data a competitive dark matter search may be performed. The

tower testing at CUTE is also expected to reduce the commissioning time of SuperCDMS SNOLAB, thanks to the knowledge gained by operating the detectors and read-out chain at CUTE.

Another key finding from the tests at CUTE was the observation of an excess low-energy rate that decayed exponentially, but was not well matched to any known radioactive backgrounds [102]. A similar observation was seen when operating a comparable device at the University of Massachusetts (see Chapter 6), and the primary explanation is that the excess background is related to stress in the crystal [14]. Despite this excess background, competitive constraints on low-mass, spin-independent dark matter were set using the data from the CPD device operated at UMass. The continuous read out of the detector allowed for new analysis techniques to be developed, which helped in achieving sensitivity to dark matter masses below $50 \text{ MeV}/c^2$. For this work my primary contributions were the development of the livetime and quality cuts, and the calibration of the detector's energy scale.

The mitigation of the stress related background could be important for improving the results from direct detection experiments that use cryogenic solid-state detectors, as well as being an important step in increasing the coherence times of superconducting qubits [14]. CUTE is an ideal facility for performing dedicated studies with superconducting qubits to try and better understand the mechanisms that decrease the coherence time, thanks to its low-background environment. For instance, one mechanism that has been suggested to decrease the coherence time of superconducting qubits is through interactions of ionizing radiation (such as cosmic rays) [15], so performing dedicated tests with superconducting qubits in a low-background environment like CUTE could help better understand such proposals.

The results in this thesis have shown cryogenic detectors are well suited for probing the parameter space of low-mass dark matter models. While a variety of technical challenges exist, in both the underlying detector technology and its scaling, the potential of cryogenic detectors could allow for the testing of lighter dark matter masses. The upcoming tests of SuperCDMS SNOLAB detectors in the CUTE facility, followed by the data from SuperCDMS SNOLAB, should soon provide new competitive results.

Bibliography

- [1] N. Aghanim et al. Planck 2018 results. *Astronomy & Astrophysics*, 641:A6, 2020.
- [2] A B McDonald. Sudbury Neutrino Observatory results. *Physica Scripta*, T121:29–32, 2005.
- [3] T. Aoyama et al. The anomalous magnetic moment of the muon in the standard model. *Physics Reports*, 887:1–166, 2020.
- [4] R. D. Peccei. The strong CP problem and axions. *Lecture Notes in Physics*, 741:3–17, 2008.
- [5] A. Arbey and F. Mahmoudi. Dark matter and the early universe: A review. *Progress in Particle and Nuclear Physics*, page 103865, 2021.
- [6] D.S. Akerib et al. The LUX-ZEPLIN (LZ) experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 953:163047, 2020.
- [7] E. Aprile. First dark matter search results from the XENON1T experiment. *Physical Review Letters*, 119:181301, 2017.
- [8] P.-A. Amaudruz et al. First results from the DEAP-3600 dark matter search with argon at SNOLAB. *Physical Review Letters*, 121(7), 2018.
- [9] Q. Arnaud et al. First results from the NEWS-G direct dark matter search experiment at the LSM. *Astroparticle Physics*, 97:54–62, 2018.
- [10] C. Amole et al. Dark matter search results from the PICO-60 C₃F₈ bubble chamber. *Physical Review Letters*, 118(25):251301, 2017.

- [11] R. Agnese et al. Projected sensitivity of the SuperCDMS SNOLAB experiment. *Physical Review D*, 95(8):082002, 2017.
- [12] A.H. Abdelhameed et al. First results from the CRESST-III low-mass dark matter program. *Physical Review D*, 100(10), 2019.
- [13] J. Åström et al. Fracture processes observed with a cryogenic detector. *Physics Letters A*, 356(4-5):262–266, 2006.
- [14] Robin Anthony-Petersen et al. A stress induced source of phonon bursts and quasi-particle poisoning, 2022. arXiv: 2208.02790.
- [15] Antti P. Vepsäläinen et al. Impact of ionizing radiation on superconducting qubit coherence. *Nature*, 584(7822):551–556, 2020.
- [16] P. A. R. Ade et al. Planck 2015 results. XIII. Cosmological parameters. *Astronomy & Astrophysics*, 594:A13, 2016.
- [17] Vera C. Rubin and Jr. Ford, W. Kent. Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. *The Astrophysical Journal*, 159:379, 1970.
- [18] K. G. Begeman, A. H. Broeils, and R. H. Sanders. Extended rotation curves of spiral galaxies: dark haloes and modified dynamics. *Monthly Notices of the Royal Astronomical Society*, 249(3):523–537, 1991.
- [19] F. Zwicky. Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta*, 6:110–127, 1933.
- [20] Frank Watson Dyson, Arthur Stanley Eddington, and C. Davidson. A determination of the deflection of light by the sun’s gravitational field, from observations made at the total eclipse of May 29, 1919. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 220(571-581):291–333, 1920.
- [21] F. Zwicky. On the masses of nebulae and of clusters of nebulae. *The Astrophysical Journal*, 86:217, 1937.

- [22] Richard Massey, Thomas Kitching, and Johan Richard. The dark matter of gravitational lensing. *Reports on Progress in Physics*, 73(8):086901, 2010.
- [23] Douglas Clowe, Marusa Bradac, Anthony H. Gonzalez, Maxim Markevitch, Scott W. Randall, Christine Jones, and Dennis Zaritsky. A direct empirical proof of the existence of dark matter. *The Astrophysical Journal*, 648:L109–L113, 2006.
- [24] NASA. Astronomy Picture of the Day: The Matter of the Bullet Cluster. <https://apod.nasa.gov/apod/ap170115.html>. Online; accessed November 2022.
- [25] C. L. Bennett et al. Four-year COBE DMR Cosmic Microwave Background Observations: Maps and basic results. *The Astrophysical Journal*, 464(1):L1–L4, 1996.
- [26] G. Hinshaw et al. Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Cosmological parameter results. *The Astrophysical Journal Supplement Series*, 208(2):19, 2013.
- [27] C. Amsler et al. Review of Particle Physics. *Physics Letters B*, 667(1), 2008.
- [28] P. Tisserand et al. Limits on the MACHO content of the galactic Halo from the EROS-2 survey of the Magellanic Clouds. *Astronomy and Astrophysics*, 469(2):387–404, 2007.
- [29] Timothy D. Brandt. Constraints on MACHO dark matter from compact stellar systems in ultra-faint dwarf galaxies. *The Astrophysical Journal Letters*, 824(2):L31, 2016.
- [30] Bernard Carr, Kazunori Kohri, Yuuiti Sendouda, and Jun’ichi Yokoyama. Constraints on primordial black holes. *Reports on Progress in Physics*, 84(11):116902, 2021.
- [31] Scott Dodelson. *Modern cosmology*. Academic Press, San Diego, CA, 2003.
- [32] Leszek Roszkowski, Enrico Maria Sessolo, and Sebastian Trojanowski. WIMP dark matter candidates and searches—current status and future prospects. *Reports on Progress in Physics*, 81(6):066201, 2018.

- [33] Tao Han, Zhen Liu, and Shufang Su. Light neutralino dark matter: direct/indirect detection and collider searches. *Journal of High Energy Physics*, 08:093, 2014.
- [34] Hsin-Chia Cheng, Jonathan L. Feng, and Konstantin T. Matchev. Kaluza-Klein dark matter. *Physical Review Letters*, 89(21), 2002.
- [35] Jan Kuckei, Claudio Dib, Amand Faessler, Thomas Gutsche, Sergey Kovalenko, Valery E. Lyubovitskij, and Kem Pumsa-ard. Strong CP violation and the neutron electric dipole form-factor. *Physics of Atomic Nuclei*, 70:349–357, 2007.
- [36] Pierre Sikivie. Axion cosmology. *Lecture Notes in Physics*, 741:19–50, 2008.
- [37] C. Bartram et al. Axion Dark Matter eXperiment: Run 1B analysis details. *Physical Review D*, 103(3), 2021.
- [38] David J.E. Marsh. Axion cosmology. *Physics Reports*, 643:1–79, 2016.
- [39] Raymond T. Co, Lawrence J. Hall, and Keisuke Harigaya. Predictions for axion couplings from ALPogenesis. *Journal of High Energy Physics*, 2021(1), 2021.
- [40] K. Arisaka, P. Beltrame, C. Ghag, J. Kaidi, K. Lung, A. Lyashenko, R. D. Peccei, P. Smith, and K. Ye. Expected sensitivity to galactic/solar axions and bosonic Super-WIMPs based on the axio-electric effect in liquid xenon dark matter detectors. *Astroparticle Physics*, 44:59–67, 2013.
- [41] H. N. Brown et al. Precise measurement of the positive muon anomalous magnetic moment. *Physical Review Letters*, 86:2227–2231, 2001.
- [42] Gopolang Mohlabeng. Revisiting the dark photon explanation of the muon anomalous magnetic moment. *Physical Review D*, 99(11), 2019.
- [43] J.P. Lees et al. Search for invisible decays of a dark photon produced in e^+e^- collisions at BaBar. *Physical Review Letters*, 119(13), 2017.
- [44] J. P. Lees et al. Search for a dark photon in e^+e^- collisions at BaBar. *Physical Review Letters*, 113(20):201801, 2014.

- [45] Maxim Pospelov, Adam Ritz, and Mikhail B. Voloshin. Secluded WIMP dark matter. *Physics Letters B*, 662:53–61, 2008.
- [46] Andrea Caputo, Alexander J. Millar, Ciaran A. J. O’Hare, and Edoardo Vitagliano. Dark photon limits: A handbook. *Physical Review D*, 104(9), 2021.
- [47] Le Hoang Nguyen, Andrei Lobanov, and Dieter Horns. First results from the WISPDMM radio frequency cavity searches for hidden photon dark matter. *Journal of Cosmology and Astroparticle Physics*, 2019(10):014–014, 2019.
- [48] Benjamin Godfrey et al. Search for dark photon dark matter: Dark E-Field radio pilot experiment. *Physical Review D*, 104(1):012013, 2021.
- [49] Itay M. Bloch, Rouven Essig, Kohsaku Tobioka, Tomer Volansky, and Tien-Tien Yu. Searching for dark absorption with direct detection experiments. *Journal of High Energy Physics*, 06:087, 2017.
- [50] Mordehai Milgrom. MOND—theoretical aspects. *New Astronomy Reviews*, 46(12):741–753, 2002.
- [51] Niels C.M. Martens and Dennis Lehmkuhl. Dark matter = modified gravity? Scrutinising the spacetime–matter distinction through the modified gravity/ dark matter lens. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 72:237–250, 2020.
- [52] Michael Boylan-Kolchin, James S. Bullock, and Manoj Kaplinghat. The Milky Way’s bright satellites as an apparent failure of Λ CDM. *Monthly Notices of the Royal Astronomical Society*, 422(2):1203–1218, 2012.
- [53] Scott Dodelson. The real problem with MOND. *International Journal of Modern Physics D*, 20:2749–2753, 2011.
- [54] Jacob D. Bekenstein. Relativistic gravitation theory for the modified Newtonian dynamics paradigm. *Physical Review D*, 70:083509, 2004.
- [55] Constantinos Skordis and Tom Złóśnik. New relativistic theory for modified Newtonian dynamics. *Physical Review Letters*, 127(16), 2021.

- [56] G. Aad et al. Search for dark matter produced in association with a single top quark in $\sqrt{s}=13$ TeV pp collisions with the ATLAS detector. *The European Physics Journal*, C81(860), 2021.
- [57] A.M. Sirunyan et al. Search for dark matter particles produced in association with a Higgs boson in proton-proton collisions at $\sqrt{s}=13$ TeV. *Journal of High Energy Physics*, 2020(3), 2020.
- [58] Eder Izaguirre, Gordan Krnjaic, Philip Schuster, and Natalia Toro. New electron beam-dump experiments to search for MeV to few-GeV dark matter. *Physical Review D*, 88(11), 2013.
- [59] Björn Penning. The pursuit of dark matter at colliders—an overview. *Journal of Physics G: Nuclear and Particle Physics*, 45(6):063001, 2018.
- [60] B. Aubert et al. Measurement of the $e^+e^- \rightarrow b\bar{b}$ cross section between $\sqrt{s}=10.54$ and 11.20 GeV. *Physical Review Letters*, 102(1), 2009.
- [61] Jennifer M. Gaskins. A review of indirect searches for particle dark matter. *Contemporary Physics*, 57(4):496–525, 2016.
- [62] Benjamin Zitzer. A search for dark matter from dwarf galaxies using VERITAS. *Proceedings of Science*, ICRC2015:1225, 2016.
- [63] Marco Chianese, Damiano F.G. Fiorillo, Gennaro Miele, Stefano Morisi, and Ofelia Pisanti. Decaying dark matter at IceCube and its signature on high energy gamma experiments. *Journal of Cosmology and Astroparticle Physics*, 2019(11):046–046, 2019.
- [64] K. Abe et al. Indirect search for dark matter from the galactic center and halo with the Super-Kamiokande detector. *Physical Review D*, 102(7), 2020.
- [65] A. Albert et al. Search for dark matter towards the galactic centre with 11 years of ANTARES data. *Physics Letters B*, 805:135439, 2020.
- [66] B. Abi et al. Prospects for beyond the standard model physics searches at the Deep Underground Neutrino Experiment. *The European Physical Journal C*, 81(4), 2021.

- [67] Nicole F. Bell, Matthew J. Dolan, and Sandra Robles. Searching for dark matter in the Sun using Hyper-Kamiokande. *Journal of Cosmology and Astroparticle Physics*, 2021(11):004, 2021.
- [68] W. B. Atwood et al. The Large Area Telescope on the Fermi Gamma-ray Space Telescope Mission. *The Astrophysical Journal*, 697(2):1071–1102, 2009.
- [69] D. Iakubovskiy. Observation of the new emission line at ~ 3.5 keV in X-ray spectra of galaxies and galaxy clusters. *Advances in Astronomy and Space Physics*, 6(1):3–15, 2016.
- [70] E. Figueroa-Feliciano et al. Searching for keV sterile neutrino dark matter with X-ray microcalorimeter sounding rockets. *The Astrophysical Journal*, 814(1):82, 2015.
- [71] R. Agnese et al. Results from the Super Cryogenic Dark Matter Search experiment at Soudan. *Physical Review Letters*, 120(6):061802, 2018.
- [72] Sebastian Baum, Andrzej K. Drukier, Katherine Freese, Maciej Górski, and Patrick Stengel. Searching for dark matter with paleo-detectors. *Physics Letters B*, 803:135325, 2020.
- [73] Jodi Cooley. Dark matter direct detection of classical WIMPs. *SciPost Physics Lecture Notes*, 2022.
- [74] J. I. Read. The local dark matter density. *Journal of Physics G: Nuclear and Particle Physics*, 41(6):063101, 2014.
- [75] K. Schneck et al. Dark matter effective field theory scattering in direct detection experiments. *Physical Review D*, 91(9), 2015.
- [76] Katherine Freese, Mariangela Lisanti, and Christopher Savage. Annual modulation of dark matter: A review. *Reviews of Modern Physics*, 85(4):1561–1581, 2013.
- [77] R. Bernabei et al. The dark matter: DAMA/LIBRA and its perspectives, 2021. arXiv: 2110.04734.

- [78] R. Bernabei et al. No role for muons in the DAMA annual modulation results. *The European Physical Journal C*, 72(7), 2012.
- [79] Jonathan H. Davis. Fitting the annual modulation in DAMA with neutrons from muons and neutrinos. *Physical Review Letters*, 113(8), 2014.
- [80] Govinda Adhikari et al. Strong constraints from COSINE-100 on the DAMA dark matter results using the same sodium iodide target. *Science Advances*, 7(46):eabk2699, 2021.
- [81] G. Adhikari et al. An induced annual modulation signature in COSINE-100 data by DAMA/LIBRA’s analysis method, 2022. arXiv: 2208.05158.
- [82] E. Aprile et al. Dark matter search results from a one ton-year exposure of XENON1T. *Physical Review Letters*, 121(11), 2018.
- [83] E. Aprile et al. Light dark matter search with ionization signals in XENON1T. *Physical Review Letters*, 123(25), 2019.
- [84] P.-A. Amaudruz et al. Design and construction of the DEAP-3600 dark matter detector. *Astroparticle Physics*, 108:1–23, 2019.
- [85] P. Agnes et al. DarkSide-50 532-day dark matter search with low-radioactivity argon. *Physical Review D*, 98(10), 2018.
- [86] P. Agnes et al. Low-mass dark matter search with the DarkSide-50 experiment. *Physical Review Letters*, 121(8), 2018.
- [87] C. Amole et al. Dark matter search results from the complete exposure of the PICO-60 C₃F₈ bubble chamber. *Physical Review D*, 100(2), 2019.
- [88] E. Behnke et al. First dark matter search results from a 4-kg CF₃I bubble chamber operated in a deep underground site. *Physical Review D*, 86(5), 2012.
- [89] W.J. Bolte et al. Development of bubble chambers with enhanced stability and sensitivity to low-energy nuclear recoils. *Nuclear Instruments and Methods in Physics*

- Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 577(3):569–573, 2007.
- [90] D. Baxter et al. First demonstration of a scintillating xenon bubble chamber for detecting dark matter and coherent elastic neutrino-nucleus scattering. *Physical Review Letters*, 118(23), 2017.
- [91] Q. Arnaud et al. First germanium-based constraints on sub-MeV dark matter with the EDELWEISS experiment. *Physical Review Letters*, 125(14), 2020.
- [92] R. Agnese et al. Low-mass dark matter search with CDMSlite. *Physical Review D*, 97(2):022002, 2018.
- [93] D.W. Amaral et al. Constraints on low-mass, relic dark matter candidates from a surface-operated SuperCDMS single-charge sensitive detector. *Physical Review D*, 102(9), 2020.
- [94] J. Åström et al. Fracture processes studied in CRESST. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 559(2):754–756, 2006. Proceedings of the 11th International Workshop on Low Temperature Detectors.
- [95] Prakruth Adari et al. EXCESS workshop: Descriptions of rising low-energy spectra. *SciPost Physics Proceedings*, 9(001), 2022.
- [96] R.F. Lang et al. Electron and gamma background in CRESST detectors. *Astroparticle Physics*, 32(6):318–324, 2010.
- [97] A. Aguilar-Arevalo et al. Search for low-mass WIMPs in a 0.6 kg day exposure of the DAMIC experiment at SNOLAB. *Physical Review D*, 94(8), 2016.
- [98] Liron Barak et al. SENSEI: Direct-detection results on sub-GeV dark matter from a new skipper CCD. *Physical Review Letters*, 125(17), 2020.
- [99] J. Aalbers et al. First dark matter search results from the LUX-ZEPLIN (LZ) experiment, 2022. arXiv: 2207.03764.

- [100] A. Aguilar-Arevalo et al. Results on low-mass Weakly Interacting Massive Particles from an 11 kg-d target exposure of DAMIC at SNOLAB. *Physical Review Letters*, 125(24), 2020.
- [101] F. Ruppin, J. Billard, E. Figueroa-Feliciano, and L. Strigari. Complementarity of dark matter detectors in light of the neutrino background. *Physical Review D*, 90(8), 2014.
- [102] Ryan Underwood. *Dark matter searches with germanium and silicon detectors in SuperCDMS and CUTE*. PhD thesis, Queen’s University, 2022.
- [103] R. Agnese et al. Silicon detector dark matter results from the final exposure of CDMS II. *Physical Review Letters*, 111:251301, 2013.
- [104] R. Abusaidi et al. Exclusion limits on the WIMP-nucleon cross section from the Cryogenic Dark Matter Search. *Physical Review Letters*, 84:5699–5703, 2000.
- [105] D. S. Akerib et al. Exclusion limits on the WIMP-nucleon cross section from the first run of the Cryogenic Dark Matter Search in the Soudan Underground Laboratory. *Physical Review D*, 72(5), 2005.
- [106] F. Duncan, A.J. Noble, and D. Sinclair. The construction and anticipated science of SNOLAB. *Annual Review of Nuclear and Particle Science*, 60(1):163–180, 2010.
- [107] T. Aralis et al. Constraints on dark photons and axionlike particles from the SuperCDMS Soudan experiment. *Physical Review D*, 101(5), 2020.
- [108] Masahiro Ibe, Wakutaka Nakano, Yutaro Shoji, and Kazumine Suzuki. Migdal effect in dark matter direct detection experiments. *Journal of High Energy Physics*, 2018(3), 2018.
- [109] C. Jacoboni, F. Nava, C. Canali, and G. Ottaviani. Electron drift velocity and diffusivity in germanium. *Physical Review B*, 24:1014–1026, 1981.
- [110] B. Cabrera, M. Pyle, R. Moffatt, K. Sundqvist, and B. Sadoulet. Oblique propagation of electrons in crystals of germanium and silicon at sub-kelvin temperature in low electric fields. arXiv:1004.1233.

- [111] S.O.W. Antman, D.A. Landis, and R.H. Pehl. Measurements of the Fano factor and the energy per hole-electron pair in germanium. *Nuclear Instruments and Methods*, 40(2):272–276, 1966.
- [112] R.H. Pehl, F.S. Goulding, D.A. Landis, and M. Lenzlinger. Accurate determination of the ionization energy in semiconductor detectors. *Nuclear Instruments and Methods*, 59(1):45–55, 1968.
- [113] U. Fano. Ionization yield of radiations. II. The fluctuations of the number of ions. *Physical Review*, 72:26–29, Jul 1947.
- [114] R. Agnese et al. Improved WIMP-search reach of the CDMS II germanium data. *Physical Review D*, 92(7), 2015.
- [115] N. W. Ashcroft and N. D. Mermin. *Solid State Physics*. Holt-Saunders, 1976.
- [116] B. Neganov and V. Trofimov. 1985 USSR Patent No 1037771, *Otkrytiya, Izobret*, 146:215, 1985.
- [117] P. N. Luke. Voltage-assisted calorimetric ionization detector. *Journal of Applied Physics*, 64(12):6858–6860, 1988.
- [118] M.A. Lindemann. *Microcalorimetry and the transition-edge sensor*. PhD thesis, University of California, Davis, 2010.
- [119] Samuel Watkins. *Athermal phonon sensors in searches for light dark matter*. PhD thesis, University of California, Berkeley, 2022.
- [120] F. C. Wellstood, C. Urbina, and John Clarke. Hot-electron effects in metals. *Physical Review B*, 49:5942–5955, Mar 1994.
- [121] W. A. Little. The transport of heat between dissimilar solids at low temperatures. *Canadian Journal of Physics*, 37(3):334–349, 1959.
- [122] I. Alkhatib et al. Light dark matter search with a high-resolution athermal phonon detector operated above ground. *Physical Review Letters*, 127(6), 2021.

- [123] Jason R. Stevens et al. Characterization of transition edge sensors for the Simons Observatory. *Journal of Low Temperature Physics*, 199(3-4):672–680, 2020.
- [124] Kent Irwin and Gene Hilton. *Transition-Edge Sensors*. Springer-Verlag,, 2021-10-12 15:10:27 2021.
- [125] A. Phipps, A. Juillard, B. Sadoulet, B. Serfass, and Y. Jin. A HEMT-based cryogenic charge amplifier with sub-100 eVee ionization resolution for massive semiconductor dark matter detectors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 940:181–184, 2019.
- [126] Zeeshan Ahmed. *A dark matter search using the final CDMS-II dataset and a novel detector of surface radiocontamination*. PhD thesis, California Institute of Technology, 2012.
- [127] Catalin Vitelaru, Daniel Lundin, Nils Brenning, and Tiberiu Minea. Plasma reactivity in high-power impulse magnetron sputtering through oxygen kinetics. *Applied Physics Letters*, 103(10):104105, 2013.
- [128] William Page. *Searching for low-mass dark matter with SuperCDMS Soudan detectors*. PhD thesis, University of British Columbia, 2019.
- [129] R. Agnese et al. Demonstration of surface electron rejection with interleaved germanium detectors for dark matter searches. *Applied Physics Letters*, 103(16):164105, 2013.
- [130] D. Norcini et al. Precision measurement of Compton scattering in silicon with a skipper CCD for dark matter detection. *Physical Review D*, 106(9), 2022.
- [131] Frank Pobell. *Matter and Methods at Low Temperatures*. Springer, 1992.
- [132] Microchip Technology. *ATEVK1101*, 2007. Discontinued.
- [133] Thomas K. Gaisser. *Cosmic Rays and Particle Physics*. Cambridge University Press, 1990.

- [134] Jonathan Corbett. Neutron calibration system for the Cryogenic Underground Test Facility (CUTE). Master’s thesis, Queen’s University, 2021.
- [135] Mi-Tech Tungsten Metals. Tungsten alloys. <https://www.mttm.com/tungsten-alloys>. Online; accessed November 2022.
- [136] Durridge. RAD7 Manual. <https://durridge.com/documentation/RAD7%20Manual.pdf>. Online; accessed November 2022.
- [137] Silvia Scorza. Radioactive background characterization of the Cryogenic Underground TEst facility (CUTE). In *17th International Conference on Topics in Astroparticle and Underground Physics (TAUP 2021)*, 2021. <https://indico.ific.uv.es/event/6178/contributions/15932/>.
- [138] U. Patel, Ivan V. Pechenezhskiy, B. L. T. Plourde, M. G. Vavilov, and R. McDermott. Phonon-mediated quasiparticle poisoning of superconducting microwave resonators. *Physical Review B*, 96(22), 2017.
- [139] A. Biekert et al. Scintillation yield from electronic and nuclear recoils in superfluid ^4He . *Physical Review D*, 105(9), 2022.
- [140] Sunil Golwala. *Exclusion limits on the WIMP-nucleon elastic-scattering cross section from the Cryogenic Dark Matter Search*. PhD thesis, University of California at Berkeley, 2000.
- [141] Murray Rosenblatt. Remarks on Some Nonparametric Estimates of a Density Function. *The Annals of Mathematical Statistics*, 27(3):832 – 837, 1956.
- [142] E. Armengaud et al. Searching for low-mass dark matter particles with a massive Ge bolometer operated above ground. *Physical Review D*, 99(8), 2019.
- [143] Matt Pyle. Modeling the differential rate for dark matter interactions in coincidence with noise fluctuations, 2022. Internal note.
- [144] J D Lewin and P F Smith. Review of mathematics, numerical factors, and corrections for dark matter experiments based on elastic nuclear recoil. Technical report, RAL, Chilton, 1996.

- [145] Richard W. Schnee. Introduction to dark matter experiments. In *Physics of the Large and the Small*. World Scientific, 2011.
- [146] Steven Yellin. Finding an upper limit in the presence of an unknown background. *Physical Review D*, 66(3):032005, 2002.
- [147] Steven Yellin. Extending the optimum interval method, 2007. arXiv: 0709.2701.
- [148] J. I. Collar. Search for a nonrelativistic component in the spectrum of cosmic rays at Earth. *Physical Review D*, 98:023005, 2018.
- [149] G. Angloher et al. Results on MeV-scale dark matter from a gram-scale cryogenic calorimeter operated above ground. *The European Physical Journal C*, 77(9), 2017.
- [150] Ricardo Z. Ferreira, M. C. David Marsh, and Eike Müller. Do direct detection experiments constrain axionlike particles coupled to electrons? *Physical Review Letters*, 128(22), 2022.
- [151] R. Agnese et al. Search for low-mass Weakly Interacting Massive Particles using voltage-assisted calorimetric ionization detection in the SuperCDMS experiment. *Physical Review Letters*, 112(4), 2014.
- [152] R. Agnese et al. New results from the search for low-mass Weakly Interacting Massive Particles with the CDMS Low Ionization Threshold Experiment. *Physical Review Letters*, 116(7):071301, 2016.
- [153] R. Agnese et al. Search for low-mass dark matter with CDMSlite using a profile likelihood fit. *Physical Review D*, 99(6), 2019.
- [154] R. Agnese et al. Search for low-mass WIMPs with SuperCDMS. *Physical Review Letters*, 112(24), 2014.
- [155] D.G. Michael et al. The magnetized steel and scintillator calorimeters of the MINOS experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 596(2):190–228, 2008.

- [156] R. Agnese et al. Production rate measurement of tritium and other cosmogenic isotopes in germanium with CDMSlite. *Astroparticle Physics*, 104:1–12, 2019.
- [157] Yonit Hochberg, Tongyan Lin, and Kathryn M. Zurek. Absorption of light dark matter in semiconductors. *Physical Review D*, 95(2), 2017.
- [158] Francesco Capozzi and Georg Raffelt. Axion and neutrino bounds improved with new calibrations of the tip of the red-giant branch using geometric distance determinations. *Physical Review D*, 102(8), 2020.
- [159] Maurizio Giannotti, Igor G. Irastorza, Javier Redondo, Andreas Ringwald, and Ken'ichi Saikawa. Stellar recipes for axion hunters. *Journal of Cosmology and Astroparticle Physics*, 2017(10):010–010, 2017.
- [160] M. Agostini et al. First search for bosonic Superweakly Interacting Massive Particles with masses up to $1 \text{ MeV}/c^2$ with GERDA. *Physical Review Letters*, 125(1), 2020.
- [161] E. Aprile et al. Excess electronic recoil events in XENON1T. *Physical Review D*, 102(7), 2020.
- [162] K. Abe et al. Search for dark matter in the form of hidden photons and axion-like particles in the XMASS detector. *Physics Letters B*, 787:153–158, 2018.
- [163] E. Aprile et al. Search for new physics in electronic recoil data from XENONnT. *Physical Review Letters*, 129(16), 2022.
- [164] Michael Dine, Willy Fischler, and Mark Srednicki. A simple solution to the strong CP problem with a harmless axion. *Physics Letters B*, 104(3):199–202, 1981.
- [165] Haipeng An, Maxim Pospelov, Josef Pradler, and Adam Ritz. Direct detection constraints on dark photon dark matter. *Physics Letters B*, 747:331–338, 2015.
- [166] C.A.J. O'Hare. AxionLimits. <https://github.com/cajohare/AxionLimits>, 2020. GitHub repository; accessed November 2022.

- [167] Paola Arias, Davide Cadamuro, Mark Goodsell, Joerg Jaeckel, Javier Redondo, and Andreas Ringwald. WISPy Cold Dark Matter. *Journal of Cosmology and Astroparticle Physics*, 06:013, 2012.
- [168] Samuel D. McDermott and Samuel J. Witte. Cosmological evolution of light dark photon dark matter. *Physical Review D*, 101(6):063030, 2020.
- [169] Samuel J. Witte, Salvador Rosauero-Alcaraz, Samuel D. McDermott, and Vivian Poulin. Dark photon dark matter in the presence of inhomogeneous structure. *Journal of High Energy Physics*, 06:132, 2020.
- [170] Andrea Caputo, Hongwan Liu, Siddharth Mishra-Sharma, and Joshua T. Ruderman. Modeling dark photon oscillations in our inhomogeneous universe. *Physical Review D*, 102(10):103533, 2020.
- [171] Andrea Caputo, Hongwan Liu, Siddharth Mishra-Sharma, and Joshua T. Ruderman. Dark photon oscillations in our inhomogeneous universe. *Physical Review Letters*, 125(22):221303, 2020.
- [172] Sergei Dubovsky and Guzmán Hernández-Chifflet. Heating up the galaxy with hidden photons. *Journal of Cosmology and Astroparticle Physics*, 12:054, 2015.
- [173] Digvijay Wadekar and Glennys R. Farrar. Gas-rich dwarf galaxies as a new probe of dark matter interactions with ordinary matter. *Physical Review D*, 103(12):123028, 2021.
- [174] Sabyasachi Chakraborty, Tae Hyun Jung, Vazha Loladze, Takemichi Okui, and Kohsaku Tobioka. Solar origin of the XENON1T excess without stellar cooling problems. *Physical Review D*, 102(9), 2020.
- [175] Georg G. Raffelt. Astrophysical methods to constrain axions and other novel particle phenomena. *Physics Reports*, 198(1):1–113, 1990.
- [176] Guido Mueller, Pierre Sikivie, D. B. Tanner, and Karl van Bibber. Detailed design of a resonantly enhanced axion-photon regeneration experiment. *Physical Review D*, 80(7), 2009.

BIBLIOGRAPHY

- [177] L.F. Abbott and P. Sikivie. A cosmological bound on the invisible axion. *Physics Letters B*, 120(1):133–136, 1983.
- [178] Arnaud Andrianaivalomahefa et al. Limits from the FUNK experiment on the mixing strength of hidden-photon dark matter in the visible and near-ultraviolet wavelength range. *Physical Review D*, 102(4), 2020.

Appendix A

The optimal filter

Optimal filter algorithms are used to obtain the optimal fit of a signal template to a raw signal contaminated by noise. For a real particle event in the data, the pulse is expected to match the signal template (with some scaling of the amplitude) with some variations due to noise. The optimal filter method can be shown to provide the optimal fit of the template to the data in the case where there is no variation in the real pulse shape and the noise is stationary (does not vary with time) [140]. Optimal filter algorithms work by converting the signal and template to Fourier (frequency) space, and performing the fit in this domain. These methods have been used (for instance) by the CDMS [140] and SuperCDMS experiments [92, 102] to extract the energy (or more precisely a parameter related to the energy) from the raw traces.

A.1 The 1-template optimal filter

As stated above, the primary purpose of the optimal filter algorithm is to extract the amplitude of the expected signal (template) from a raw trace, in the presence of (stationary) noise. It is assumed that the raw trace $S(t)$ measured from the detector can be adequately described by the form:

$$S(t) = aA(t) + n(t), \tag{A.1}$$

where $A(t)$ is the signal template, $n(t)$ is the noise, and a is the signal amplitude (the energy estimator that is to be extracted). The signal template is generally estimated by taking the average of many raw traces, and either using this averaged value as the

A.1. THE 1-TEMPLATE OPTIMAL FILTER

template, or by performing a time domain fit to the averaged trace using a best-guess of the form of the pulse shape. Irrespective of how the signal template is generated, it is assumed that the signal template is a good description of the signal shape in the raw traces. The noise template is difficult to measure in the time domain, but can be easily estimated in the frequency domain by calculating the average power spectral density of traces that have no pulse in them.

For the case of Gaussian noise, a simple time-domain fitting routine that minimizes the χ^2 between the template and the raw trace can be used. However, time-domain fitting has drawbacks in the case of non-Gaussian but still stationary noise, namely, non-Gaussian noise introduces correlations in the noise at different times. To do an optimal fit in the presence of non-Gaussian noise, the fit must be performed in the frequency domain by converting the relevant signals and templates with a Fourier transform:

$$\tilde{\psi}(\nu) = \int_{-\infty}^{\infty} \psi(t) e^{i2\pi\nu t} dt. \quad (\text{A.2})$$

The different frequency components of stationary noise are uncorrelated, so the issue of correlations in time that was present for the time-domain fitting is no longer an issue. The variation of the noise with frequency is accounted for by de-weighting frequency components that are expected to have higher levels of noise. The optimal amplitude is determined by minimizing the χ^2 with respect to amplitude. The χ^2 is defined as [102, 140]:

$$\chi^2(a) \equiv \sum_{\nu} \frac{|\tilde{S}(\nu) - a\tilde{A}(\nu)|^2}{\tilde{J}(\nu)}, \quad (\text{A.3})$$

here $\tilde{J}(\nu)$ is the power spectral noise density $\langle \tilde{n}^2(\nu) \rangle$. Minimizing the χ^2 with respect to amplitude allows for the optimal amplitude to be determined as [102, 140]:

$$a = \sum_{\nu} \frac{\tilde{S}(\nu)\tilde{A}(\nu)/\tilde{J}(\nu)}{\tilde{A}(\nu)\tilde{A}^*(\nu)/\tilde{J}(\nu)}. \quad (\text{A.4})$$

This provides the optimal amplitude for the situation where the triggered pulse is aligned with the template, however, the quality of the fit may be improved by allowing a delay parameter to account for imperfect triggering. A small delay t_0 may be added such that the χ^2 becomes:

$$\chi^2(a, t_0) \equiv \sum_{\nu} \frac{|\tilde{S}(\nu) - ae^{-i2\pi\nu t_0}\tilde{A}(\nu)|^2}{\tilde{J}(\nu)}. \quad (\text{A.5})$$

The optimal amplitude can then be determined in a similar fashion as for no delay case, i.e. minimizing the χ^2 with respect to the amplitude and solving the resulting equation [140].

A.2 N-signal M-background optimal filter

A more generic form of optimal filter can be developed to fit multiple signal and background templates in the presence of stationary noise, this is known as the N-signal M-background ($N_s M_b$) optimal filter. The original purpose of the $N_s M_b$ optimal filter was to remove the effects of muons from data obtained in surface facilities, but other uses for the algorithm arose, such as fitting background pileup events in the presence of laser triggered events with known start times. The $N_s M_b$ optimal filter can also be used to fit glitch-like non-stationary noise in a trace, so that the glitches do not interfere with the fitting of the signal templates.

In a similar fashion to the 1-template optimal filter, the χ^2 for the $N_s M_b$ optimal filter (with delay) is given by:

$$\chi^2(\mathbf{a}, t_0) \equiv \sum_{\nu} \frac{|\tilde{S}(\nu) - \sum_{n=1}^N a_n e^{-2\pi i t_0 \nu} \tilde{A}_n(\nu) - \sum_{m=N+1}^{M+N} a_m \tilde{A}_m(\nu)|^2}{\tilde{J}(\nu)}. \quad (\text{A.6})$$

Here, $\tilde{A}_n(\nu)$ are the N-signal templates and $\tilde{A}_m(\nu)$ are the M-background templates, the vector $\mathbf{a}$ represents the amplitudes of the respective signal and background components. It should be noted that with this description of the $N_s M_b$ optimal filter only the N-signal templates are able shift in time, while the M-background templates are assumed to be fixed in time.

The derivation to the solution for the $N_s M_b$ optimum filter used in this thesis will be given below, specifically a 1-template 2-background optimal filter. The explicit form of the χ^2 for this optimal filter is given by:

$$\chi^2(\mathbf{a}, t_0) \equiv \sum_{\nu} \frac{|\tilde{S}(\nu) - a_1 e^{-2\pi i t_0 \nu} \tilde{A}_1 - a_2 \tilde{A}_2 - a_3 \tilde{A}_3|^2}{\tilde{J}(\nu)}. \quad (\text{A.7})$$

The first step in obtaining the amplitudes $\vec{a}$ is to solve $\frac{\partial \chi^2}{\partial a_1} = \frac{\partial \chi^2}{\partial a_2} = \frac{\partial \chi^2}{\partial a_3} = 0$, which yields the following system of equations:

$$\mathbf{a} = \mathbf{P}^{-1} \cdot \mathbf{q}. \quad (\text{A.8})$$

A.2. N-SIGNAL M-BACKGROUND OPTIMAL FILTER

Here, the $\mathbf{P}$ matrix is given by:

$$\mathbf{P} = \begin{pmatrix} \sum_{\nu} \frac{\tilde{A}_{1,\nu}^* \tilde{A}_{1,\nu}}{\tilde{J}_{\nu}} & \sum_{\nu} \frac{\tilde{A}_{1,\nu}^* \tilde{A}_{2,\nu} e^{2\pi i t_0 \nu}}{\tilde{J}_{\nu}} & \sum_{\nu} \frac{\tilde{A}_{1,\nu}^* \tilde{A}_{3,\nu} e^{2\pi i t_0 \nu}}{\tilde{J}_{\nu}} \\ \sum_{\nu} \frac{\tilde{A}_{1,\nu}^* \tilde{A}_{2,\nu} e^{2\pi i t_0 \nu}}{\tilde{J}_{\nu}} & \sum_{\nu} \frac{\tilde{A}_{2,\nu}^* \tilde{A}_{2,\nu}}{\tilde{J}_{\nu}} & \sum_{\nu} \frac{\tilde{A}_{2,\nu}^* \tilde{A}_{3,\nu}}{\tilde{J}_{\nu}} \\ \sum_{\nu} \frac{\tilde{A}_{1,\nu}^* \tilde{A}_{3,\nu} e^{2\pi i t_0 \nu}}{\tilde{J}_{\nu}} & \sum_{\nu} \frac{\tilde{A}_{2,\nu}^* \tilde{A}_{3,\nu}}{\tilde{J}_{\nu}} & \sum_{\nu} \frac{\tilde{A}_{3,\nu}^* \tilde{A}_{3,\nu}}{\tilde{J}_{\nu}} \end{pmatrix}, \quad (\text{A.9})$$

and the $\mathbf{q}$ vector is given by:

$$\mathbf{q} = \begin{pmatrix} \sum_{\nu} \frac{\tilde{S}_{\nu} \tilde{A}_{1,\nu}^* e^{2\pi i t_0 \nu}}{\tilde{J}_{\nu}} \\ \sum_{\nu} \frac{\tilde{S}_{\nu} \tilde{A}_{2,\nu}^*}{\tilde{J}_{\nu}} \\ \sum_{\nu} \frac{\tilde{S}_{\nu} \tilde{A}_{3,\nu}^*}{\tilde{J}_{\nu}} \end{pmatrix}. \quad (\text{A.10})$$

Both $\mathbf{P}$ and $\mathbf{q}$ (and hence $\mathbf{a}$) depend on the delay t_0 , so in order to determine the best fit (minimum χ^2) with delay the computation must be repeated for each value of t_0 (which is the same as the number of bins in the trace). While this may seem computationally intensive, it should be noted that $\mathbf{P}$ does not depend on the raw trace, so $\mathbf{P}$ can be precomputed (and inverted) for each t_0 .

Explicit Python implementations of the standard optimal filter and $N_s M_b$ optimal filter algorithms can be found in the QETpy repository ¹, along with further documentation on their usage.

¹<https://github.com/spice-herald/QETpy>