

Ionization Collection in Detectors of the Cryogenic Dark Matter Search

by

Arran Thomas James Phipps

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Bernard Sadoulet, Chair

Professor Robert G. Jacobsen

Professor Kai Vetter

Spring 2016

Ionization Collection in Detectors of the Cryogenic Dark Matter Search

Copyright 2016
by
Arran Thomas James Phipps

Abstract

Ionization Collection in Detectors of the Cryogenic Dark Matter Search

by

Arran Thomas James Phipps

Doctor of Philosophy in Physics

University of California, Berkeley

Professor Bernard Sadoulet, Chair

Determining the composition of dark matter is at the forefront of modern scientific research. There is compelling evidence for the existence of vast quantities of dark matter throughout the universe, however it has so-far eluded all direct detection efforts and its identity remains a mystery. Weakly interacting massive particles (WIMPs) are a favored dark matter candidate and have been the primary focus of direct detection for several decades. The Cryogenic Dark Matter Search (CDMS) has developed the Z-dependent Ionization and Phonon (ZIP) detector to search for such particles. Typically made from germanium, these detectors are capable of distinguishing between electromagnetic background and a putative WIMP signal through the simultaneous measurement of ionization and phonons produced by scattering events. CDMS has operated several arrays of these detectors at the Soudan Underground Laboratory (Soudan, MN, USA) resulting in many competitive (often world-leading) WIMP exclusion limits.

This dissertation focuses on ionization collection in these detectors under the sub-Kelvin, low electric field, and high crystal purity conditions unique to CDMS. The design and performance of a fully cryogenic HEMT-based amplifier capable of achieving the SuperCDMS SNOLAB ionization energy resolution goal of 100 eVee is presented. The experimental apparatus which has been used to record electron and hole properties under CDMS conditions is described. Measurements of charge transport, trapping, and impact ionization as a function of electric field in two CDMS detectors are shown, and the ionization collection efficiency is determined. The data is used to predict the error in the nuclear recoil energy scale under both CDMSlite and iZIP operating modes. A two species, two state model is developed to describe how ionization collection and space charge generation in CDMS detectors are controlled by the presence of “overcharged” D^- donor and A^+ acceptor impurity states. The thermal stability of these states is exclusive to sub-Kelvin operation, explaining why ionization collection in CDMS detectors differs from similar semiconductor detectors operating at higher temperature. This work represents a solid foundation for the understanding ionization collection in CDMS detectors.

To Emily, Finnegan, Margaret, Annaliese, Lucian, and Kayleigh.

Contents

Contents	ii
List of Figures	vii
List of Tables	xii
1 Dark Matter	1
1.1 A Brief History of Dark Matter	1
1.2 Evidence for Dark Matter	2
1.2.1 Galactic Rotation Curves	2
1.2.2 Gravitational Lensing	3
1.2.3 Large Scale Structure	5
1.2.4 Big Bang Nucleosynthesis	5
1.2.5 Cosmic Microwave Background	6
1.3 Candidate Particles and Theories	7
1.3.1 Modified Newtonian Dynamics (MOND)	7
1.3.2 Massive Compact Halo Objects (MACHOs)	8
1.3.3 Neutrinos	8
1.3.4 Axions	9
1.3.5 Weakly Interacting Massive Particles (WIMPs)	10
1.4 Detection Techniques	12
1.4.1 Collider Production	13
1.4.2 Indirect Detection	13
1.4.3 Direct Detection	14
1.5 Principles of Direct Detection	16
1.5.1 Interaction Rate in a Terrestrial Detector	16
1.5.2 Annual Modulation	20
1.5.3 Halo Model	21
1.5.4 Sensitivity Limits and Backgrounds	22
2 The Cryogenic Dark Matter Search	23
2.1 Overview	23

2.2	ZIP Detectors	24
2.2.1	Phonon Measurement	25
2.2.2	Ionization Measurement	27
2.2.3	Discrimination by Ionization Yield	28
2.2.4	Luke-Neganov Amplification	30
2.2.5	Detector Types	32
2.3	Problematic Phenomena and Events	34
2.3.1	Space Charge Buildup	34
2.3.2	Surface Events	36
2.3.3	Zero Charge Events	37
2.4	The Soudan Underground Laboratory	38
2.5	WIMP Search Results	39
2.5.1	CDMS I	39
2.5.2	CDMS II	41
2.5.3	SuperCDMS Soudan	42
2.5.4	CDMSlite	43
2.6	The Future - SuperCDMS SNOLAB	44
3	Charge Amplification of Cryogenic Semiconductor Detectors	46
3.1	Introduction	46
3.2	Shockley-Ramo Theorem	46
3.2.1	Qualitative Understanding	46
3.2.2	Quantitative Calculation	47
3.2.3	Planar Geometry	50
3.2.4	oZIP and mZIP Geometry	51
3.2.5	iZIP Geometry	51
3.3	The Basic Charge Amplifier	51
3.3.1	Input Signal	52
3.3.2	Bias Resistor	53
3.3.3	Open Loop Amplifier	54
3.3.4	Coupling Capacitor	54
3.3.5	Feedback	55
3.3.6	Output Signal	58
3.4	Noise Sources	59
3.4.1	Bias Resistor Johnson Noise	59
3.4.2	Feedback Resistor Johnson Noise	59
3.4.3	Open Loop Amplifier Voltage Noise	59
3.4.4	Open Loop Amplifier Current Noise	60
3.4.5	Shot Noise Due to Leakage Currents	61
3.4.6	Microphonics	62
3.4.7	Other Sources of Noise	62
3.4.8	Total Input Equivalent Voltage Noise	62

3.5	Optimal Filtering and Energy Resolution	62
3.6	CDMS JFET Amplifier	65
3.6.1	Open Loop Voltage Amplifier	65
3.6.2	DC Biasing	66
3.6.3	Closed Loop Charge Amplifier with Detector	68
3.6.4	Performance	70
4	Ionization Readout with HEMTs	71
4.1	Introduction	71
4.2	CNRS/LPN HEMTs	72
4.3	Modified CDMS Amplifier	74
4.4	Fully Cryogenic HEMT Amplifier	78
4.4.1	Open Loop Voltage Amplifier	79
4.4.2	Output Follower	81
4.4.3	Reset Switch and Leakage Current	82
4.4.4	Capacitive Feedback	82
4.4.5	Predicted Resolution	82
4.5	Prototype Amplifier Performance	83
4.6	Comparing Topologies	86
5	The Physics of Electrons and Holes in sub-Kelvin Germanium	90
5.1	Introduction	90
5.2	Band Structure and Effective Mass	90
5.2.1	The Schrödinger Equation	90
5.2.2	Dispersion Relations, Effective Mass, and the Crystal Momentum	91
5.2.3	Electron Effective Mass and Oblique Propagation	94
5.2.4	Hole Effective Mass	96
5.2.5	Conductivity Mass and Density of States Mass	98
5.3	Bulk Transport	99
5.3.1	Scattering	99
5.3.2	Fundamental Distributions and Averages	105
5.3.3	Monte Carlo Method and the Results of Sundqvist	107
5.3.4	Effective Carrier Temperature	112
5.3.5	Signal Shape	113
5.4	Capture (Trapping)	113
5.4.1	Capture Rates, Trapping Times, and Trapping Lengths	113
5.4.2	Signal Shape	116
5.4.3	Cascade Coulomb Capture	116
5.4.4	Neutral Capture Across the Band Gap	119
5.4.5	Neutral Capture Producing Overcharged D^-/A^+ States	119
5.4.6	Other Processes	122
5.5	Impact Ionization	123

5.5.1	Generation Rates, Times, and Lengths	123
5.5.2	Signal Shape	123
5.5.3	Cross Sections	126
6	Experimental Apparatus for the Measurement of Electron and Hole Properties	128
6.1	Introduction	128
6.2	Fiber Optic Probe	128
6.2.1	Vacuum Interface Fiber Flange	130
6.2.2	300K to 4K Can Fibers	131
6.2.3	From single mode to multi-mode	132
6.2.4	4K Can to 4K Tower Fibers	133
6.2.5	Tower Fibers and Detector Lid	133
6.2.6	Laser Diodes	134
6.3	Detector Geometry and Biasing	134
6.4	High Bandwidth Charge Amplifier	135
6.4.1	Open Loop JFET	135
6.4.2	Coaxial Cables	137
6.4.3	Akash Terminator	137
6.4.4	Amptek A250	137
6.4.5	AD8000 Follower	138
6.4.6	Gain, Output Pulse Shape, and the Drift Current Signal	139
6.5	Data Acquisition System	140
6.6	Data Processing	143
6.7	Drift Velocity Measurement	144
6.8	Trapping Length Measurement	146
6.9	Impact Ionization Measurement	147
7	Analysis of Measured Electron and Hole Properties	149
7.1	Introduction	149
7.2	Drift Velocities	150
7.3	Qualitative Dependence on Electric Field	151
7.4	Trapping Lengths	151
7.4.1	G22Q (n-type)	153
7.4.2	G13H (p-type)	156
7.4.3	EDELWEISS data	158
7.5	Impact Ionization	161
7.6	Temperature Dependence	162
7.7	Ionization Collection Efficiency	164
7.7.1	Relative Calibration	164
7.7.2	Absolute Calibration	166
7.7.3	Including Impact Ionization	168

7.7.4	Position Dependence	172
7.8	A Simple Model for iZIP Collection Efficiency	174
7.8.1	Bias Voltage, Electric Field, and the Ramo Potential	174
7.8.2	Approximated Ionization Collection Efficiency	176
7.8.3	Approximated Charge z-Partition	176
8	Implications for CDMS Detectors	179
8.1	Space Charge and the Two Species, Two State Model	179
8.1.1	Grounded, Steady State Condition	180
8.1.2	Generation of Space Charge via Capture	183
8.1.3	Detector Reset	184
8.2	Ionization Collection Efficiency and the Total Phonon Energy	185
8.3	Nuclear Recoil Energy Scale	187
8.3.1	CDMSlite Mode	188
8.3.2	iZIP Mode	189
8.4	Silicon Detectors	191
8.5	Moving to Larger Crystals	193
9	Conclusions	197
	Bibliography	199
A	Optimal Filtering	211
A.1	The Fourier Transform	211
A.2	Noise and the Power Spectral Density	211
A.3	Discrete Sampling	213
A.4	Optimal Amplitude Estimator	213
A.5	Optimal Start Time Estimator	215
A.6	Interpretation as a Filter	216
B	HEMT Calculations	218
B.1	Noise Contribution from the HEMT Cascode	218
B.2	Modified Effective Load Resistor Topology for Improved Gain	220
C	Electrical Shorts at Soudan	222
C.1	CDMS Cold Hardware	222
C.2	Characteristics	225
C.3	Possible Causes	226
C.3.1	Loss of NbTi Wire Tension	228
C.3.2	Metal Debris	228

List of Figures

1.1	<i>Left:</i> Vera Rubin. <i>Right:</i> An image of the Andromeda galaxy (M31) from the Palomar Sky Survey, compared to measured rotational velocities.	2
1.2	The universal energy distribution of dark energy, dark matter, and atoms. . . .	3
1.3	Galaxy rotation curves of several spiral galaxies.	4
1.4	<i>Left:</i> Color image of the Bullet cluster from Magellan. <i>Right:</i> X-ray emissions from Chandra.	4
1.5	The distribution of 67,676 galaxies within 5° of the Equatorial plane, color coded by their absolute magnitudes.	5
1.6	The all-sky CMB temperature fluctuations from the 9-year WMAP data.	6
1.7	Nine-year WMAP data shown with 6-parameter Λ CDM fits.	7
1.8	Current and proposed limits in the axion parameter space.	9
1.9	The comoving number density of WIMPs of mass M as a function of temperature.	11
1.10	The three basic dark matter detection techniques.	12
1.11	WIMP-nucleon cross sections limits from ATLAS (left) and CMS (right) at the LHC.	13
1.12	Visual representation of the various recoil energy deposition channels and their use by direct detection experiments.	15
1.13	Current direct detection limits.	17
1.14	The Helm form factor for various target nuclei.	19
1.15	Differential scattering rate versus recoil energy for various WIMP masses.	20
1.16	Experimental residual rate of single-hit scintillation events measured by DAMA/LIBRA in the 2-6 keV energy range.	21
2.1	A modern CDMS ZIP detector.	24
2.2	CDMS detector response to a typical scattering event.	25
2.3	Schematic of athermal phonon collection and quasiparticle diffusion.	26
2.4	Ionization energy and yield as a function of energy, for a detector exposed to a neutron source.	28
2.5	Comparison between the Lindhard model and measured ionization yield for nuclear recoils in Ge and Si.	31
2.6	Ionization and phonon sensor layout for the oZIP, mZIP, and iZIP designs. . . .	32
2.7	Phonon and ionization sensor layout for iZIP detectors.	33

2.8	Loss of charge collection efficiency due to space charge buildup.	35
2.9	Illustration demonstrating the geometric cause of the reduced ionization collection efficiency observed in surface events.	36
2.10	Ionization yield versus phonon timing parameter.	37
2.11	Zero charge events.	38
2.12	Dependence of muon flux with depth, shown with various underground facilities.	39
2.13	Layout of the CDMS Soudan installation.	40
2.14	SuperCDMS Soudan detector payload, consisting of 15 iZIP detectors.	41
2.15	WIMP-nucleon versus WIMP mass upper limits from the CDMS II experiment.	42
2.16	SuperCDMS Soudan low threshold analysis WIMP limits.	43
2.17	CDMSlite WIMP search limits.	44
2.18	Proposed SuperCDMS SNOLAB limits.	45
3.1	The position dependence in electric field line termination for a charge carrier placed between two planar electrodes.	47
3.2	Illustration showing how the flow of induced charge between electrodes can be sensed by an amplifier connected across the electrodes.	48
3.3	Setup for calculating the weighting potentials for two infinitely long, planer electrodes separated by a distance d	50
3.4	The basic charge amplifier topology.	52
3.5	Amplifier coupling efficiency as a function of the coupling capacitance.	55
3.6	Closed loop signal flow diagram.	56
3.7	Equivalent circuits for calculating the current-to-voltage and feedback factors.	57
3.8	Open loop amplifier noise sources.	60
3.9	Charge noise PSD comparison with cryocooler on and off.	63
3.10	The open loop CDMS JFET amplifier in the small signal limit.	66
3.11	The basic DC biasing of the CDMS JFET amplifier.	67
3.12	The CDMS closed loop JFET-based charge amplifier.	68
3.13	The closed loop current-to-voltage gain of the CDMS JFET amplifier.	69
3.14	The closed loop voltage-to-voltage gain of the CDMS JFET amplifier.	70
4.1	Basic physical structure of a HEMT.	71
4.2	The drain-source current vs drain-source voltage, as a function of the gate-source voltage, of a typical CNRS/LPN HEMT.	73
4.3	Comparison of the CNRS/LPN HEMT and Si JFET noise.	74
4.4	CDMS JFET amplifier modified to replace the JFET with a HEMT.	75
4.5	Experimental setup to measure the input voltage noise of the modified CDMS amplifier.	76
4.6	Measured voltage noise of the modified CDMS amplifier, referred to the input.	76
4.7	Measured output voltage noise of the modified CDMS amplifier in a high resistance feedback configuration.	77
4.8	Electron collection spectrum of a CDMS II detector exposed to a ^{241}Am source.	78

4.9	The fully cryogenic HEMT amplifier, which consists of an open loop HEMT voltage amplifier, HEMT follower, and HEMT switch.	79
4.10	Open loop HEMT voltage amplifier.	80
4.11	The HEMT follower, along with its small signal model.	82
4.12	Noise performance of the HEMT amplifier.	84
4.13	A typical 356 keV pulse from the fully cryogenic HEMT amplifier.	85
4.14	PSD of the fully cryogenic HEMT amplifier coupled to a CDMS oZIP detector (G11).	85
4.15	Comparison of the various amplifier topologies using active and passive reset schemes.	87
4.16	Ionization resolution as a function of low frequency cutoff for the various amplifier topologies discussed in the text.	88
5.1	The band structure of germanium.	92
5.2	Average thermal occupation as a function of energy for $T=45$ mK.	93
5.3	A three dimensional depiction of the surfaces of constant energy in conduction band of Ge.	95
5.4	Distribution of energy between inner and outer charge channels for the hole-collecting (side 1) and electron-collecting (side 2) surfaces of a CDMS iZIP detector.	97
5.5	The band structure of holes in germanium.	98
5.6	The various categories of scattering processes which can influence the bulk transport of charge carriers in sub-Kelvin germanium.	100
5.7	The phonon dispersion relation for germanium, idealized for a longitudinal branch and a single transverse branch for both acoustic and optical phonons.	102
5.8	The total phonon scattering rates for electrons, at $T=40$ mK, under isotropic assumptions.	104
5.9	Total phonon scattering rates for holes, at $T=40$ mK, under isotropic assumptions.	104
5.10	The energy and momentum relaxation times as a function of electric field for electrons and holes.	108
5.11	Drift velocities as a function of electric field for electrons and holes.	109
5.12	The total velocity as a function of electric field for electrons and holes.	110
5.13	The carrier energy distribution for electrons and holes at an electric field of 1 V/cm.	111
5.14	The effective carrier temperature T_C as a function of electric field for electrons and holes.	112
5.15	Examples of a single carrier event and mixed carrier event.	114
5.16	The cascade capture process.	117
5.17	Various capture and impact ionization cross sections for electrons in sub-Kelvin germanium.	120
5.18	Various capture and impact ionization cross sections for holes in sub-Kelvin germanium.	121
6.1	Experimental setup for the measurement of electron and hole properties.	129

6.2	The vacuum interface fiber flange.	130
6.3	A single 4K can to 4K tower fiber.	132
6.4	Tower fiber mounting fixture.	133
6.5	The high bandwidth charge amplifier.	136
6.6	The Amptek A250 charge sensitive preamplifier.	138
6.7	By taking the time derivative of the output voltage, the low frequency integration pole is cancelled and the drift current signal recovered.	141
6.8	A typical average output pulse and its derivative.	143
6.9	Determination of the drift time for a typical pulse.	145
6.10	Exponential fit to determine the trapping time.	146
6.11	Drift current pulse showing exponential growth in the primary pulse and the expected stop time of the secondary current due to impact ionization.	148
7.1	Measured electron and hole drift velocities as a function of electric field for the n-type and p-type detector.	150
7.2	Qualitative change in drift current pulse shape for electrons and holes as the electric field is increased in detector G22Q (n-type).	152
7.3	Qualitative change in drift current pulse shape for electrons and holes as the electric field is increased in detector G13H (p-type).	152
7.4	Electron and hole trapping lengths for G22Q.	153
7.5	Inverse trapping lengths with fits for G22Q.	155
7.6	Trapping lengths for G13H.	158
7.7	Inverse trapping lengths with fits for G13H.	159
7.8	Inverse trapping lengths with fit for the p-type EDELWEISS detector.	160
7.9	Temperature dependence of the hole and electron trapping lengths in G13H (p-type).	163
7.10	The relative calibration (seagull plot) of the ionization collection efficiency for G22Q.	165
7.11	Absolute calibration of the ionization collection efficiency in G22Q.	167
7.12	Calculated ionization collection efficiency for the n-type (G22Q) and p-type (G13H) detectors using the effective trapping length.	169
7.13	Ionization collection efficiency as calculated using the integrated method, which includes the effects of impact ionization.	171
7.14	Ionization collection efficiency as a function of event position for various electric fields in G22Q, as determined by the primary trapping lengths.	173
7.15	Ionization collection efficiency as a function of event position for various electric fields in G13H, as determined by the primary trapping lengths.	173
7.16	Ionization collection efficiency as a function of event position for various electric fields in G22Q, approximated using the integrated, effective lengths.	175
7.17	Ionization collection efficiency as a function of event position for various electric fields in G13H, approximated using the integrated, effective lengths.	175
7.18	The ionization collection efficiency of G22Q using the simplified iZIP model. . .	177

7.19	The predicted charge z-partition using the G22Q iZIP model collection efficiency.	178
8.1	Steady state ionized impurity fraction versus approximate bulk neutrality conditions for various ratios of donors to acceptors.	182
8.2	Biased (1 V/cm) steady state ionized impurity fractions.	184
8.3	Modified Luke gain in comparison to the standard Luke gain for G22Q.	187
8.4	Error in the CDMSlite nuclear recoil energy scale using the ionization collection efficiency parameters for G22Q.	190
8.5	Error in the iZIP nuclear recoil energy scale using the ionization collection efficiency parameters for G22Q.	192
8.6	Comparison of the ionization collection efficiencies for a 33 mm, 1 inch, and 1 cm detector.	193
8.7	33 mm charge collection efficiency as a function of z-position and electric field, using the trapping lengths derived from the integrated efficiency of G22Q. . . .	194
8.8	33 mm iZIP charge collection efficiency, using the trapping lengths derived from the integrated efficiency of G22Q.	195
8.9	Error in iZIP nuclear recoil scale as a function of the symmetric bias voltage, for varying detector thicknesses.	196
9.1	The author, Arran Phipps, and his children Emily, Margaret, and Finnegan. . .	197
B.1	The equivalent circuit and its small-signal model for a signal appearing on the gate of Q_2	219
B.2	The equivalent circuit and its small-signal model to boost the effective load resistance formed by Q_3 and Q_4	221
C.1	Cross sectional view of the CDMS tower and associated cold hardware components.	223
C.2	The CDMS SQUET assembly.	224
C.3	Side coax assembly board.	225
C.4	Physical distribution of the shorts amongst the towers, and their status before warming up at the end of R134.	227
C.5	Various cold hardware subsystems, separated by temperature.	227
C.6	Rendered views of an assembled CDMS tower.	229

List of Tables

4.1	Typical 100 pF CNRS/LPN HEMT parameters.	72
4.2	Measured energy resolution of the modified CDMS amplifier (with HEMT) in comparison to the standard JFET amplifier.	77
4.3	Predicted ionization resolution for the various amplifier topologies described in the text.	88
5.1	Electron band parameters.	95
5.2	Band structure and effective mass parameters for holes in Ge.	99
5.3	Phonon parameters.	103
6.1	Total number of traces acquired for different electric field ranges.	142
7.1	Properties of the two detectors studied using the optical probe and high bandwidth charge amplifier.	149
7.2	G22Q charge trapping processes for electrons and holes for the the trapping model described in the text.	156
7.3	G13H charge trapping processes for electrons and holes for the the trapping model described in the text.	157
7.4	EDELWEISS p-type trapping center concentrations from the inverse trapping length fits.	161
8.1	Approximate capture cross sections used to determine the steady state ionized impurity fractions.	181
C.1	Number of observed shorts during Soudan R133-134, organized by the type of signal affected.	226

Acknowledgments

I have made so many genuine friendships over the course of this dissertation and there is no hope of properly thanking everyone. But seriously — thank you — to everyone I have had the pleasure of collaborating with over the years. While there is always pride in performing good science, it is the human element of this work which I will forever look back on with fond memories.

My sincerest gratitude to my dissertation advisor, Bernard Sadoulet. It amazing just how broad *and* deep Bernard’s knowledge of physics is, and it is an honor to have been trained by such a brilliant mind. Bernard was kind enough to give me a chance to pursue my passion for physics, and I’m not sure I would have made it this far without his support.

Kyle Sundqvist deserves special thanks for being my day-to-day mentor while I was just learning the ropes. Most of the results of this dissertation are built upon his initial work, and he always motivated me to keep chugging along even when things looked bleak. I don’t think anybody else can truly appreciate the horrors of installing fibers in the fridge, but it was totally worth it. Thanks for teaching me how to properly analyze circuits, and showing me that most of the good books are actually in the engineering library. I’ll never be able to look at anything again without thinking of the equivalent circuit, and that’s a good thing.

Special thanks also to my Berkeley colleagues. Thank you to Dennis Seitz for his always-practical advice, often resulting in many “that’s so stupid it just might work” moments. Thank you to Rachel Winheld for your hard work (the group would be lost without you), and for always just being there to listen to whatever troubles, personal or otherwise, I was facing. Thanks to Bruno Serfass for always going above and beyond what was required of him, just to help out. Thanks to Nader Mirabolfathi for many stimulating discussions. Thanks to Miguel Daal for our shared struggles in running non-standard detector configurations. Danielle Speller and Todd Doughty, thanks for the companionship in navigating graduate school. Thanks to Albert Lam, Nate Earnest, Akash Dixit, Sergiy Nesterenko, Geoff Gaswint, and Alvaro Villalpando for being my undergraduates. Thank you to Colin Song for all the laughs, and for beating Bruno at tennis. Thank you to Alex Juillard for bringing an outside perspective and helping out with the HEMT amplifier. Thanks to Tal Peretz, Nick Zobrist, Xavier Defay and Bill Page for putting up with my jaded outlook as I approached graduation — it’s a normal part of the life cycle of the graduate student.

Finally, thank you to my family. It is only with your love and support that any of this is possible. Emily, Finnegan, and Margaret, thank you for being the best children a dad could ever ask for. Kayleigh, where do I even begin? Ever since we met my life has only gotten better. You love me for who I am, what more could I ask? Thank you for forever being my better half.

Hard work truly does pay off in the end.

Chapter 1

Dark Matter

1.1 A Brief History of Dark Matter

In 1933, an application of the virial theorem to observations of the Coma cluster by noted astronomer Fritz Zwicky led him to the conclusion that the total mass of the cluster was at least ten times greater than what was predicted by its luminous matter alone.[1, 2] This first observation of the effects of what Zwicky called “*dunkle materie*” — dark matter — went largely ignored for nearly 50 years. It wasn’t until the 1970s, when Vera Rubin studied the rotational velocities of stars in Andromeda and other high luminosity galaxies, that scientists began to realize there was something wrong with our understanding of the universe at a fundamental level. Her results clearly showed the velocities to remain relatively constant at high radius, in disagreement with the models based on only luminous matter which predicted a sharp decrease.[3, 4] The data supported the existence of a large, unknown dark matter background in which these galaxies were embedded, extending far beyond the visible galactic edges. The dark matter was responsible for most of the mass of the galaxies. This same behavior was observed in many other galaxies and by 1980 most astronomers were convinced that dark matter exists around galaxies and clusters.[5] The immediate question was obvious — *what exactly is this dark matter?* To this day, we still do not know the answer.

The theory of cold dark matter was proposed and developed throughout the 1980s.[6–10] Cold dark matter, produced in large abundance in the early universe, is a new type of matter which does not strongly interact with ordinary matter or radiation through electromagnetism. Astonishingly, over 80% of the matter in the universe is believed to be dark matter. Its gravitational effects are responsible for the structure of the universe. Individual structures form in a “bottom-up” fashion in the cold dark matter paradigm. The initial localized over-densities of dark matter first collapse under their own gravity before merging with similiary-sized structures in a continuous hierarchy to form ever larger objects. In some sense, dark matter can be thought of as the scaffolding over which the visible universe has been constructed.

The tiny initial variations in dark matter density were predicted to cause correlated

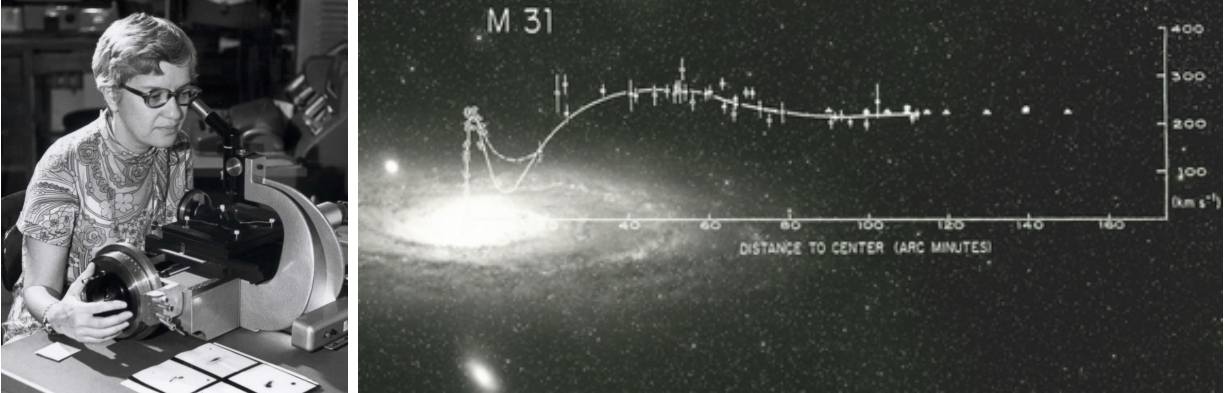


Figure 1.1: *Left:* Vera Rubin. *Right:* An image of the Andromeda galaxy (M31) from the Palomar Sky Survey, compared to measured rotational velocities. Note that the velocities remain high far beyond the visible disk. (Image by Vera Rubin and Janice Dunlap)

variations in the temperature of the cosmic microwave background (CMB). In 1992, the Nobel Prize winning COBE experiment measured these fluctuations and found them to be consistent with cold dark matter.[11] Combined with the discovery of dark energy, this led to the creation of today's standard cosmological model, Λ CDM.[12–14] The Λ represents the dark energy, also known as the cosmological constant, which makes up 71.4% of the energy of the universe. CDM — short for cold dark matter — makes up 24%. Ordinary matter makes up only a tiny fraction, a meager 4.6%. While Λ CDM has certainly earned its usage as the standard cosmology, successfully predicting an impressive number and variety of phenomena, there is still much to understand. Specifically in regards to this dissertation, we still do not know what the dark matter is made of and we have never directly detected a dark matter particle. To be the first to do so and shed light on this question is the fundamental motivation behind the Cryogenic Dark Matter Search.

1.2 Evidence for Dark Matter

1.2.1 Galactic Rotation Curves

It has been well-established that Newtonian dynamics describes the orbital motion of astrophysical objects bound together by a gravitational potential. The virial theorem relates the time-averaged kinetic energy of a stable, bound system of particles to the total potential energy. For objects within a galaxy, we expect the orbital velocities to be given by

$$v(r) \approx \sqrt{\frac{GM(r)}{r}}, \quad (1.1)$$

where $M(r)$ is the total mass enclosed at a radius r .

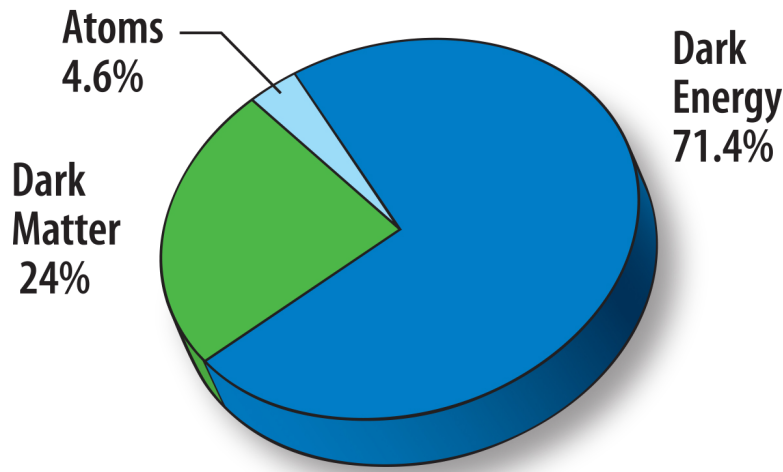


Figure 1.2: The universal energy distribution of dark energy, dark matter, and atoms (ordinary matter). The figure is based on the WMAP 9 year mission results.[15]

Assuming the total mass of the galaxy is determined by its visible stars, we can take $M(r) \approx \text{constant}$ beyond the visible edge of the galaxy. The rotational velocities of objects far away from the center of the galaxy should therefore fall as $r^{-1/2}$. Rotational velocity curves (also known as galaxy rotation curves) may be constructed by measuring the Doppler shift of several spectral lines emitted from objects located at various distances from the center of the galaxy. An excellent review of the technique can be found in [16]. As can be seen in Fig. 1.3, the rotational velocity curves of spiral galaxies are clearly flat at high radius. This implies the mass density outside the visible disc is not zero, but varies as r^{-2} — spiral galaxies appear to be embedded in gigantic clouds (halos) of dark matter which dominate the total mass of the galaxy. Models based on this assumption accurately reproduce the observed galactic rotation curves.[17]

1.2.2 Gravitational Lensing

The path of light is bent by gravitational potentials due to the effects of general relativity. A distant background image behind a massive object will become strongly lensed to the observer, duplicating and distorting the image in a predictable fashion. Strong lensing measurements can be used to determine the mass distribution of clusters. The maps show cluster masses are dominated by widespread, smooth distributions of dark matter, with visible matter concentrated at their centers.[18]

Weak lensing focuses on finding smaller distortions that appear in background images, and can also be used to make mass maps. Weak lensing surveys have also identified massive

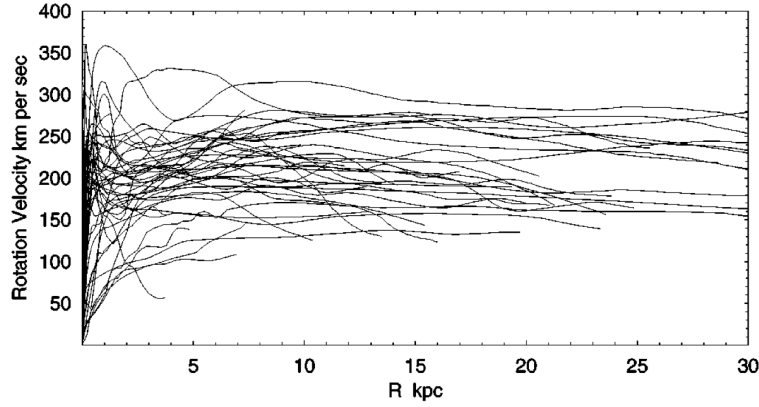
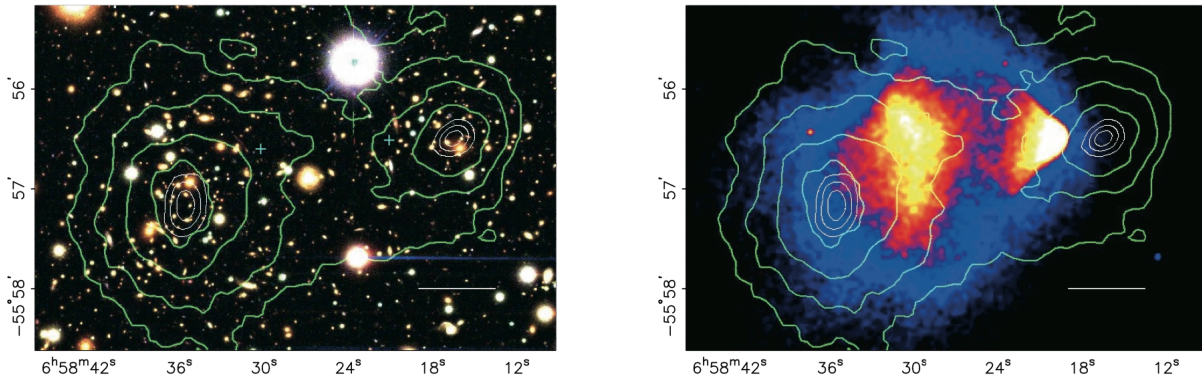


Figure 1.3: Galaxy rotation curves of several spiral galaxies.[16]

Figure 1.4: *Left*: Color image of the Bullet cluster from Magellan. *Right*: X-ray emissions from Chandra. Shown in both images are the mass map contours determined from weak lensing.[19]

clumps of dark matter, with a distribution consistent with the predictions of cold dark matter. Weak lensing has also been used to study the collision of galaxy clusters. The most well-known result comes from the merging of the Bullet cluster, in which weak lensing was used to map the mass distribution which was then compared to x-ray emissions from hot gas.[19] The x-ray emissions from hot gas indicate the ordinary matter has slowed down due to friction and is concentrated towards the center of the merger. On the other hand, the mass map clearly shows two large distributions of dark matter which have passed through one another with little interaction. The dark matter is the dominant source of mass.

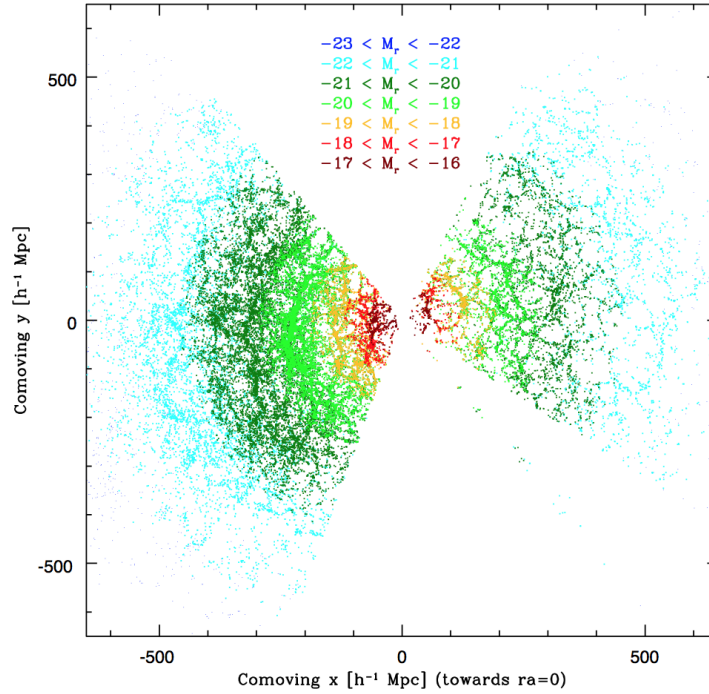


Figure 1.5: The distribution of 67,676 galaxies within 5° of the Equatorial plane, color coded by their absolute magnitudes. From the Sloan Digital Sky Survey.[21]

1.2.3 Large Scale Structure

The large scale structure of the universe is consistent with Λ CDM cosmology.[20] As can be seen in Fig. 1.5, the distribution of galaxies is not uniform but clumpy. Galaxies tend to cluster along string-like filaments and this behavior is accounted for by cold dark matter.[21] Simulations of large scale structure formation based on Λ CDM are only capable of reproducing the observed distribution of baryonic matter if long-range gravitational forces are dominated by cold dark matter.[22]

1.2.4 Big Bang Nucleosynthesis

The primordial abundances of light elements such as deuterium, helium, and lithium also provide strong evidence for the existence of a new type of matter. These light elements are produced within the first few minutes after the big bang, while atoms heavier than helium were produced in stars much later in the history of the universe. The physics of the nucleosynthesis of the light elements is well understood. The primordial abundances depend critically on the baryon density of the universe at the time of creation. The observed abundances require a relatively low baryon density, however we know a large amount of

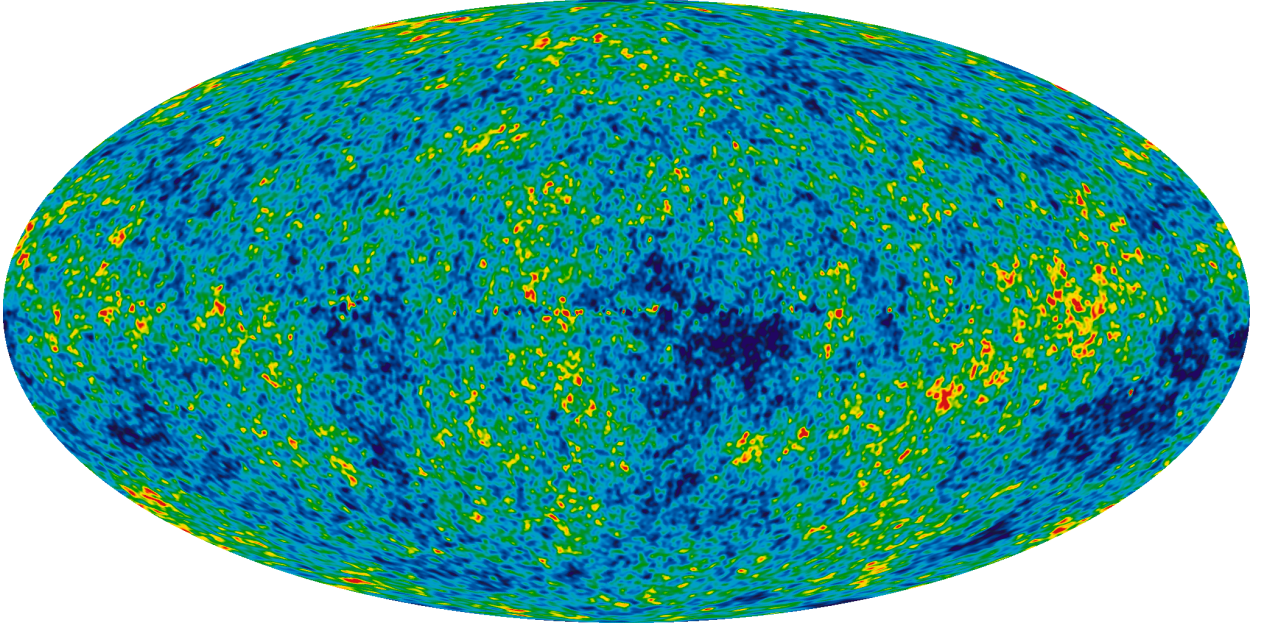


Figure 1.6: The all-sky CMB temperature fluctuations from the 9-year WMAP data. The temperature range shown is $\pm 200 \mu\text{K}$. Image from NASA.

dark matter must exist in the universe to properly account for many other phenomena. The majority of this dark matter therefore cannot be baryonic, indicating new physics is required to explain its nature. A recent review of big bang nucleosynthesis can be found in [23].

1.2.5 Cosmic Microwave Background

Initially after the big bang, the universe was too hot for neutral hydrogen to form. Photons were frequently Compton scattered by free charges and could not retain any information about the distributions of matter from which they came. The structure of the universe was masked in a thick fog. Approximately 378,000 years after the big bang, cooling due to the expansion of the universe finally brought the temperature low enough that electrons and protons could bind. When this occurred, the mean free path of photons became essentially larger than the size of the universe and the fog was lifted. The photons created at this time have continued to propagate through space, albeit continuously stretched to lower frequencies due to the universal expansion. We can detect these photons — the cosmic microwave background (CMB) — which take the form of a perfect blackbody distribution at $T \sim 2.73\text{K}$ which fills the universe and can be detected in all directions. The cosmic microwave background is the so-called afterglow of the big bang.

The photons of the CMB have not been scattered since their creation, and thus their fluctuations encode information about the distribution of matter within the universe at the

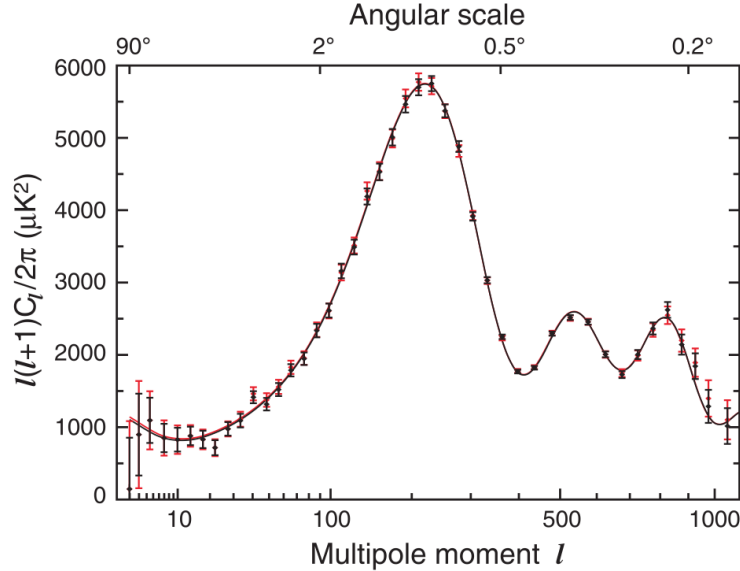


Figure 1.7: Nine-year WMAP data shown with 6-parameter Λ CDM fits. The agreement is excellent. From [15].

time of recombination. The Λ CDM model provides concrete predictions for the angular distribution of CMB temperature fluctuations, which strongly depends on the initial densities of dark matter, dark energy, and ordinary matter. The temperature anisotropies of the cosmic microwave background have been measured with increasing sensitivity by the COBE[24], WMAP[25], and Planck[26] satellites. The expansion of the CMB fluctuation map into the basis of spherical harmonics produces a number of peaks and valleys at specific multipole moments which can be compared to the predictions of the Λ CDM model. As can be seen in Fig. 1.7, the fit is excellent and strongly supports the presumed abundance of cold dark matter.

1.3 Candidate Particles and Theories

There are a number of candidate particles and theories which have been proposed to explain the dark matter. We now briefly describe several of these possibilities, although this list is by no means exhaustive.

1.3.1 Modified Newtonian Dynamics (MOND)

One possibility that is not a new particle — but instead a change in our laws of gravity — is modified Newtonian dynamics or MOND.[27] MOND proposes that Newton’s laws are

modified at the very low accelerations typically experienced by galaxies, and that this instead explains the variety of phenomena attributed to dark matter. While this is still an active area of research, MOND faces serious challenges in successfully matching observation with a single theory at all length scales. In addition, it has been argued that even if MOND is correct, it cannot account for all the dark matter needed to explain the distribution of mass in galaxy clusters.[28] It is also difficult to explain the observations of the Bullet cluster using MOND. Particle dark matter is the simplest explanation for the observational evidence and is supported over MOND by an overwhelming majority of active dark matter researchers.

1.3.2 Massive Compact Halo Objects (MACHOs)

Massive compact halo objects (MACHOs) are clumps of ordinary baryonic matter which do not emit (or emit very little) light. Possible MACHOs include black holes, neutron stars, and brown dwarfs. The presence of a large number of such objects is a possible source of dark matter. Such objects can be searched for by their microlensing signatures when passing in front of bright stars, as has been performed by the EROS and MACHO collaborations.[29–31] While some evidence for MACHOs is seen, the results show that MACHOs cannot account for all the dark matter.[32]

1.3.3 Neutrinos

As it became clear that most of the dark matter must be non-baryonic, attention quickly turned to neutrinos as a likely candidate. Developed in the early 1980s by Zel’dovich and others, the Hot Dark Matter (HDM) scenario proposed that light neutrinos were responsible for the invisible matter observed throughout the universe.[33–35] These models, however, were unable to properly predict structure formation in the universe. The relativistic motion of neutrinos dampen density fluctuations on small scales (“free-streaming” effects) until the universe cools to a temperature below the mass of the neutrino. This causes structure formation to occur in a “top-down” fashion, with superclusters forming first and galaxies much later. The universe predicted by such a model leads to a much more inhomogeneous distribution of matter than is observed, effectively ruling out standard neutrinos as an explanation for the dark matter.[36]

While the neutrinos of the standard model have been excluded, so-called *sterile neutrinos* which only interact gravitationally still remain as a possible dark matter candidate.[37] We have already discussed how such neutrinos would need to be relatively massive in order to account for structure formation, but either a heavy sterile neutrino or exotic production mechanisms have been shown to be capable of preserving the correct structure formation hierarchy. One difficulty of sterile neutrinos is that they should decay in ordinary standard model neutrinos and emit x-rays. Large-scale searches for these x-ray emissions have failed to discover anything of significance, making sterile neutrinos an unlikely explanation of the dark matter.[38, 39] There have been, however, some observations of weak x-ray signals

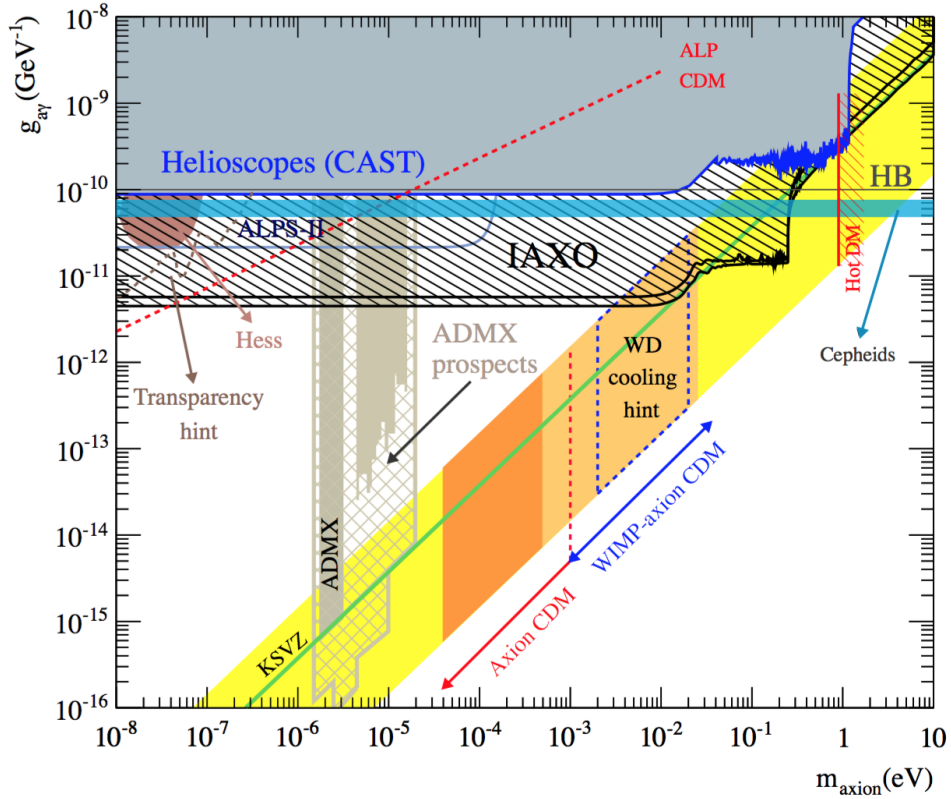


Figure 1.8: Current and proposed limits in the axion parameter space.[46]

originated from galaxies which can be interpreted through sterile neutrino decay and this is still an active area of research.[40]

1.3.4 Axions

There is an unresolved problem of the standard model involving the strong force. Strong force interactions allow for charge parity (CP) violation, however no CP violation has ever been observed in strong force processes. A solution to the strong CP problem was proposed by Peccei and Quinn.[41] CP violation can be suppressed by introducing a new $U(1)$ symmetry which is spontaneously broken. The axion is the Nambu-Goldstone boson which results from the breaking of the symmetry and would be present throughout the universe in large quantities if this mechanism were to occur. Axions are very light particles, with typical masses in the μeV range. Due to their non-thermal production mechanism, light axions are still a valid cold dark matter candidate. [42–45]

The ADMX collaboration has searched for axions through their conversion to microwave

photons in a resonant cavity immersed in a static magnetic field.[47–49] Current limits from ADMX and other experiments are shown with the proposed limits from future axion experiments in Fig. 1.8. A detailed introduction to axions and associated detection methods can be found in [50].

1.3.5 Weakly Interacting Massive Particles (WIMPs)

Weakly interacting massive particles — WIMPs — are easily the most popular proposed dark matter candidate. WIMPs are a generic class of new particles (symbolically labeled by χ) with coupling strengths characteristic of weak-scale interactions. WIMPs are generally assumed to have a mass $1 \text{ GeV} \lesssim m_\chi \lesssim 1 \text{ TeV}$. The popularity of the WIMP hypothesis comes from their emergence through general lines of reasoning in both cosmology and particle physics.

Consider a massive particle χ which is created by the big bang. Such a particle will remain in thermal and chemical equilibrium with the other constituents of the early universe until the expansion rate of the universe is roughly equal to the particle’s annihilation rate. After this condition is reached, the number density of χ particles becomes low enough that annihilation becomes rare, “freezing out” the comoving number density as seen in Fig. 1.9. This relic density of WIMPs would therefore still be present in today’s universe.

It can be easily shown that in order for the relic density to explain the dark matter density inferred through cosmological observations, we must have

$$\Omega_{DM} h^2 = \frac{3 \cdot 10^{-27} \text{ cm}^3 / \text{s}}{\langle \sigma_A v \rangle} \approx 0.1, \quad (1.2)$$

where Ω_{DM} is the ratio of the dark matter density to the critical density ($\Omega_{DM} \approx 0.22$), and h is the Hubble parameter. This leads to

$$\langle \sigma_A v \rangle \approx 3 \cdot 10^{-26} \text{ cm}^3 / \text{s} \rightarrow \sigma_A^2 \approx \frac{\alpha^2}{M_{EW}^2}, \quad (1.3)$$

where M_{EW}^2 is the electroweak scale ($\sim 100 \text{ GeV}$). This result — the so-called “WIMP miracle” — is quite remarkable. The cosmological data points to a link between the dark matter and the weak force. Any new stable particle with a mass of about 100 GeV which interacts through the weak force is a natural dark matter candidate.[52] Such particles are naturally non-relativistic at the time of freeze out making them cold dark matter, and simulations have shown they will correctly reproduce the large scale structure of the universe as we have previously discussed.

Completely independent from cosmological observations, particle physics suggests the need for new physics at the electroweak scale. The masses of the W and Z boson have been measured to be about 100 GeV . This requires some form of symmetry breaking as they are described by an $SU(2)$ gauge theory, however gauge theory bosons are massless. The symmetry breaking is assumed to be performed by the Higgs mechanism[53], through which

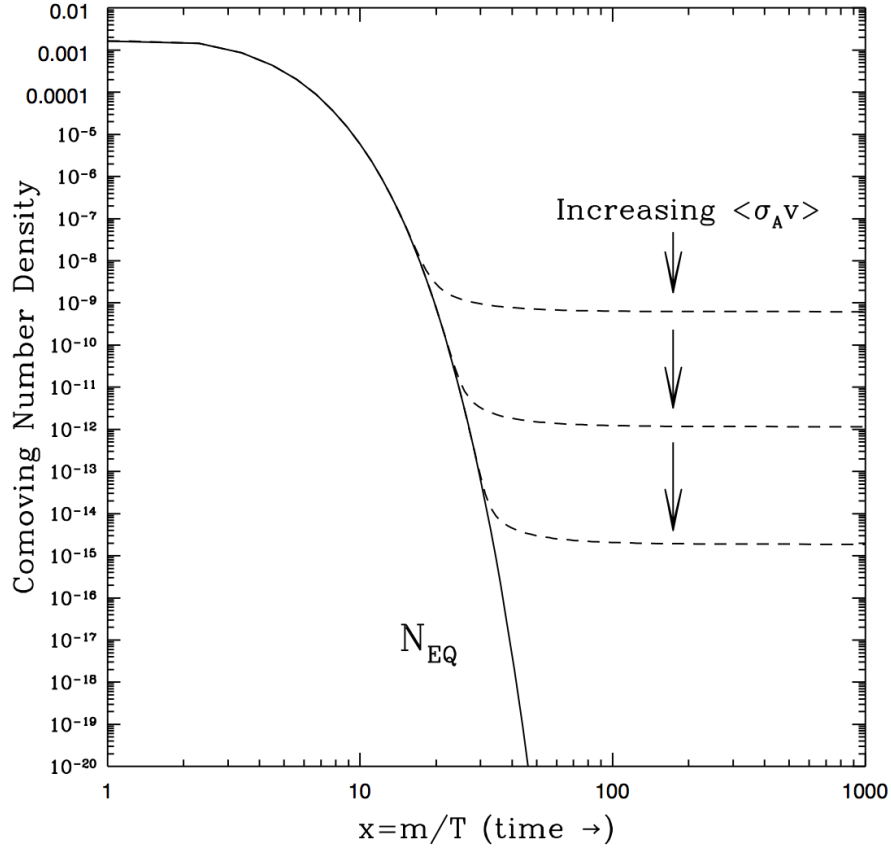


Figure 1.9: The comoving number density of WIMPs of mass M as a function of temperature. Note the clear decoupling of the WIMPs as the temperature decreases. Higher annihilation cross sections result in a lower relic density of dark matter. Figure from [51].

the elementary gauge bosons acquire their mass. Both ATLAS and CMS have detected the presumed Higgs boson with a mass of ~ 125 GeV via collider experiments at the LHC.[54, 55] Quantum (loop) corrections to the calculation of the Higgs mass, however, would predict it to be on the order of 10^{18} GeV without new physics.[56] This is known as the hierarchy problem.

A popular solution to the hierarchy problem is through the introduction of *supersymmetry* (SUSY) to the standard model. In this class of models, introduction of additional symmetries cause every boson to have a new supersymmetric fermion partner and vice-versa. A detailed discussion of SUSY models is beyond the scope of this thesis and more information can be found several in-depth reviews such as [56, 57]. The important point is that a large class of these models predict the lightest supersymmetric particle to be the neutralino — a stable particle which only interacts through gravity and the weak force, making it an ideal WIMP

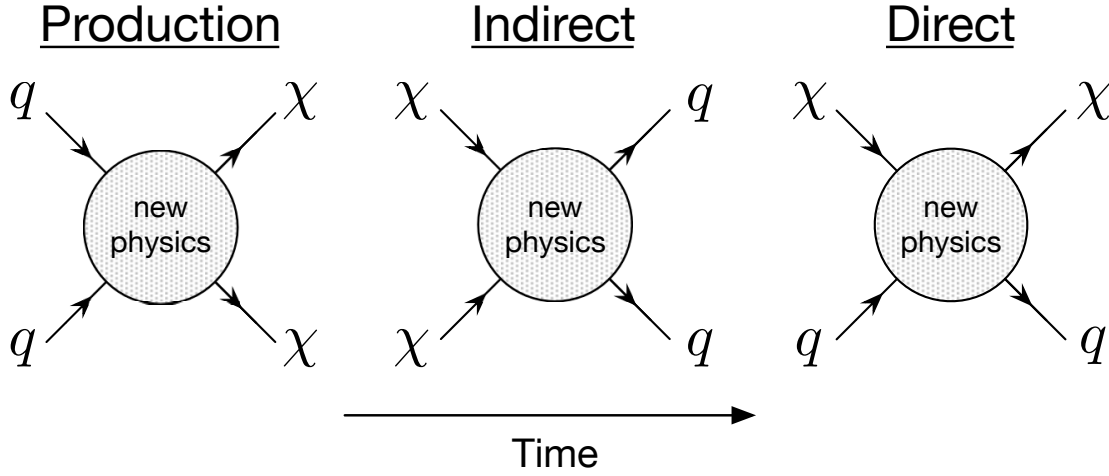


Figure 1.10: The three basic dark matter detection techniques. WIMPs (χ) are coupled to standard model particles (q) through some type of new physics.

candidate.[58, 59]

It is easy to understand how the independent prediction of WIMPs through both cosmology and particle physics, along with the success of models based on these ideas, have made WIMPs the most popular candidate. It is this form of dark matter which the Cryogenic Dark Matter Search aims to detect. To date, however, no direct evidence of WIMPs has been observed and there has been no evidence of supersymmetry at the LHC. There still, however, remains large amounts of unexplored parameter space. The remainder of this dissertation therefore focuses on WIMPs and their detection and we will often use the term dark matter with an implicit assumption that it is in the form of WIMPs.

1.4 Detection Techniques

There are three separate yet complimentary basic techniques which may be used to search for dark matter — collider production, indirect detection, and direct detection. As shown in Fig. 1.10, WIMPs may couple to standard model particles resulting in production, annihilation, and scattering processes which we can detect. Results are usually presented in the plane of WIMP mass versus WIMP-nucleon interaction cross section. Lines denote exclusion limits above which the parameter space has been ruled out at some level of statistical confidence, typically $\sim 90\%$. Closed contours indicate regions of parameter space which are consistent with possible WIMP detections, again within statistical confidence limits.

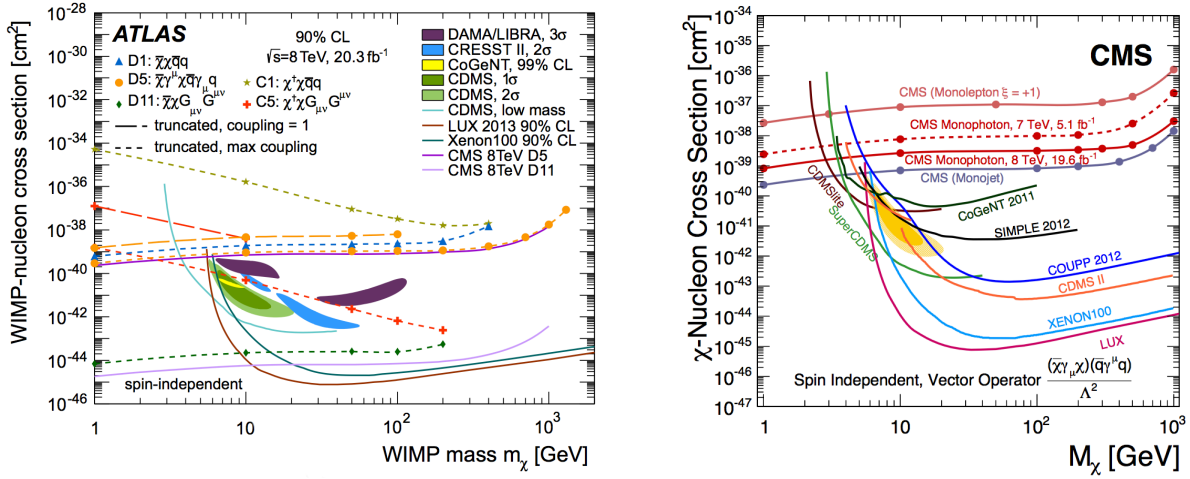


Figure 1.11: WIMP-nucleon cross sections limits from ATLAS (left) and CMS (right) at the LHC.[60, 61]

1.4.1 Collider Production

Dark matter may be produced by highly energetic collisions of standard model particles, as performed in particle accelerators such as the Large Hadron Collider. Since dark matter is assumed to be both stable and weakly interacting, any produced WIMPs would escape undetected. Accelerator experiments therefore search for events which show signs of missing energy and momentum. Limits can be set by comparison to the calculated background of known standard model processes. Collider experiments are not particularly competitive at high mass due to the fixed collision energy, however they are capable of probing new parameter space at lower WIMP masses, where kinematics limits the capabilities of direct detection experiments. Limits from both ATLAS and CMS are shown in Fig. 1.11.

1.4.2 Indirect Detection

WIMPs may be their own antiparticle, resulting in self-annihilation. Dark matter annihilations which produce high energy photons or rare antiparticles may be evident in regions of high expected dark matter density. Observation of these decay products would serve as an indirect detection of the dark matter. The lack of anomalous signal can also be used to set limits on the WIMP parameter space, however there is a decent amount of model-dependence in the production processes which complicates the analysis.

Several space-based experiments have performed searches near the center of the Milky Way, however backgrounds introduced by normal astrophysical sources make this a difficult analysis. The announcement of an observed positron excess by the PAMELA collaboration[62] generated a lot of initial excitement, and more recent measurements by AMS-02[63]

show this excess persists to energies of 300 GeV. This points towards a massive species of annihilating particle, however it has been argued that the expected signal from this dark matter if it were to be the relic abundance is much too small to account for the measurements.[63] The Fermi Gamma Ray Telescope has also seen some evidence of a gamma ray excess near the center of the Milky Way which may be attributed to annihilation of a low mass WIMP.[64–66]

Dwarf spheroidal galaxies are also expected to contain a large amount of dark matter, with a much lower abundance of astrophysical background sources as compared to the Milky Way center. Several searches for signals from these objects have been conducted, however no convincing evidence of an excess attributable to dark matter has been reported.[67, 68] This has placed some constraints below WIMP masses of ~ 20 GeV.

1.4.3 Direct Detection

As our own galaxy is embedded in a dark matter halo, billions of WIMPs pass through our bodies every second. Essentially every single one of these particles passes through the Earth without interacting, however there is a finite possibility for WIMP-nucleon scattering to occur. A properly-designed terrestrial detector may therefore search for evidence of WIMPs through deposition of recoil energy by a scattering event. The measured spectrum of recoil events depends on properties of both the WIMPs and the target material. If the events are attributable to WIMPs, it can be used to infer information about the WIMP-nucleon scattering rate. The lack of an observed recoil spectrum places strict limits on the dark matter parameter space. This technique is known as direct detection and is the method used by the Cryogenic Dark Matter Search.

Direct detection experiments are essentially just classical particle detectors searching for extremely rare processes. Current limits predict ton-scale experiments would only see a few dark matter events over exposures of a year or even longer. Background rejection and control is exceedingly important in this discipline. Most direct detection experiments are operated in remote, deep underground locations under strict cleanroom conditions. Care must be taken to avoid contaminating the detectors with radioactive background, forcing selection of individual components of the detector which have a high degree of radiopurity. While originally a fringe field, direct detection has grown into the realm of big science — collaborations including hundreds of scientists are now common and over 50 million dollars has been allocated for the current USA DOE/NSF G2 Dark Matter program. The Cryogenic Dark Matter Search has been a pioneering experiment in this field for some time.

Broadly speaking, there are two major types of technology currently used in direct detection experiments — cryogenic semiconductor detectors and liquid noble detectors. Cryogenic semiconductor detectors are crystalline substrates instrumented with a variety of sensors to measure the recoil energy deposited by a scattering particle. The recoil energy is generally distributed via some combination of ionization, scintillation, and phonons (heat). Some detectors are also able to provide active background rejection based on the simultaneous measurement of two or more of these deposition channels. A wide variety of crystal sub-

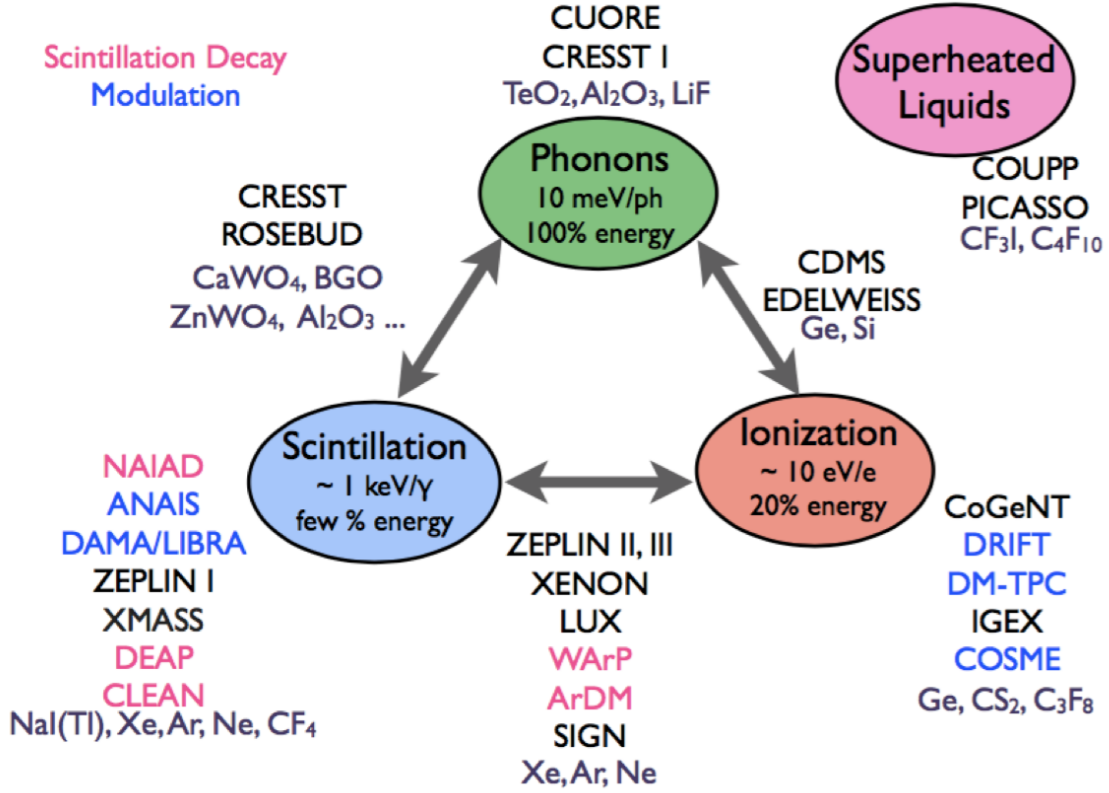


Figure 1.12: Visual representation of the various recoil energy deposition channels and their use by direct detection experiments. Experiments which appear between two channels measure both. Highlighted in blue are detectors specifically designed for annual modulation searches. Figure from [69].

strates, geometries, and sensors can be used to optimize for detection of specific WIMP mass ranges. Payloads of several detector modules are operated simultaneously to increase the total target mass. Notable cryogenic semiconductor detector experiments include the Cryogenic Dark Matter Search (CDMS)[70–73], EDELWEISS[74–76], CoGeNT[77], CRESST[78], CDEX[79] and DAMA/LIBRA[80, 81].

Liquid noble experiments instead use giant tanks of liquified noble gases, such as xenon and argon, as target materials. These detectors measure the scintillation light produced by scattering particles. Liquid noble detectors can also operate as dual-phase time projection chambers capable of performing background rejection through the simultaneous measurement of both scintillation and ionization, or by pulse shape rejection. The liquids themselves typically contain significant radioactive backgrounds, requiring increasing levels of purification as the total volume of liquid is increased. Most external backgrounds are absorbed

near the surfaces of the liquid, creating a very low background active detection region in the center. This self-shielding effect is enhanced in larger detectors. Notable liquid noble experiments include XENON[82], LUX/LZ[83, 84], ZEPLIN[85, 86], DarkSide[87], and ArDM[88].

A final type of detector not as widely used are the so-called bubble chambers. Superheated liquids are prepared such that interactions with heavy particles produce nucleation points, creating visible bubbles within the chamber which can be imaged. These threshold detectors do not provide direct information about the deposited recoil energy but have essentially no sensitivity to light electromagnetic interactions. Some discrimination between the remaining backgrounds such as alpha particles is performed through acoustic measurements of the bubble formation process. In particular, these experiments provide excellent sensitivity to spin-dependent dark matter due to the composition of the various liquid targets, however current detectors are dominated by a poorly understood background believed to originate at the liquid-gas-glass interfaces of the detector. Notable experiments include COUPP[89], PICASSO[90], and PICO[91].

This list is by no means exhaustive, and we have neglected to discuss several directional detection experiments for brevity. Needless to say, the field of direct detection is an active and rapidly growing. Several hints of signals have been reported in numerous experiments, although most are in contention with non-observation by others. A visual representation of the direct detection measurement channels and their use by direct detection experiments can be seen in Fig. 1.12. A good introductory review to direct detection techniques can be found in [69]. The current state of direct detection at the time of writing, not including projected limits from future experiments, is shown in Fig. 1.13.

1.5 Principles of Direct Detection

1.5.1 Interaction Rate in a Terrestrial Detector

The expected rate of WIMP collisions with a terrestrial detector depends upon the local density of dark matter and the composition of the target material. In the vicinity of the Earth, the average density of dark matter is 0.3 GeV/cm^3 . Note that heavier WIMP masses imply a lower *number* density, as it is the *mass* density that is assumed.

Generically, the WIMP velocity distribution is taken to be an isotropic Maxwellian ($P(v) \propto \exp(-v^2/v_0^2)$) with characteristic velocity $v_0 = 220 \text{ km/sec}$ in the vicinity of the Sun. The velocity is taken with respect to the rest frame of the galactic halo. The distribution is truncated at the galactic escape velocity of $v_{esc} \approx 540 \text{ km/sec}$. For a terrestrial detector — moving with the Earth's velocity — we make the replacement $\vec{v} \rightarrow \vec{v} + \vec{v}_E$, where $\vec{v}$ is now the velocity of the WIMP relative to the target and $\vec{v}_E$ the velocity of the Earth relative to the galactic halo. We typically take $v_E = 232 \text{ km/sec}$.

Consider a target nucleus of mass m_N . Generic WIMP scattering is an elastic process.

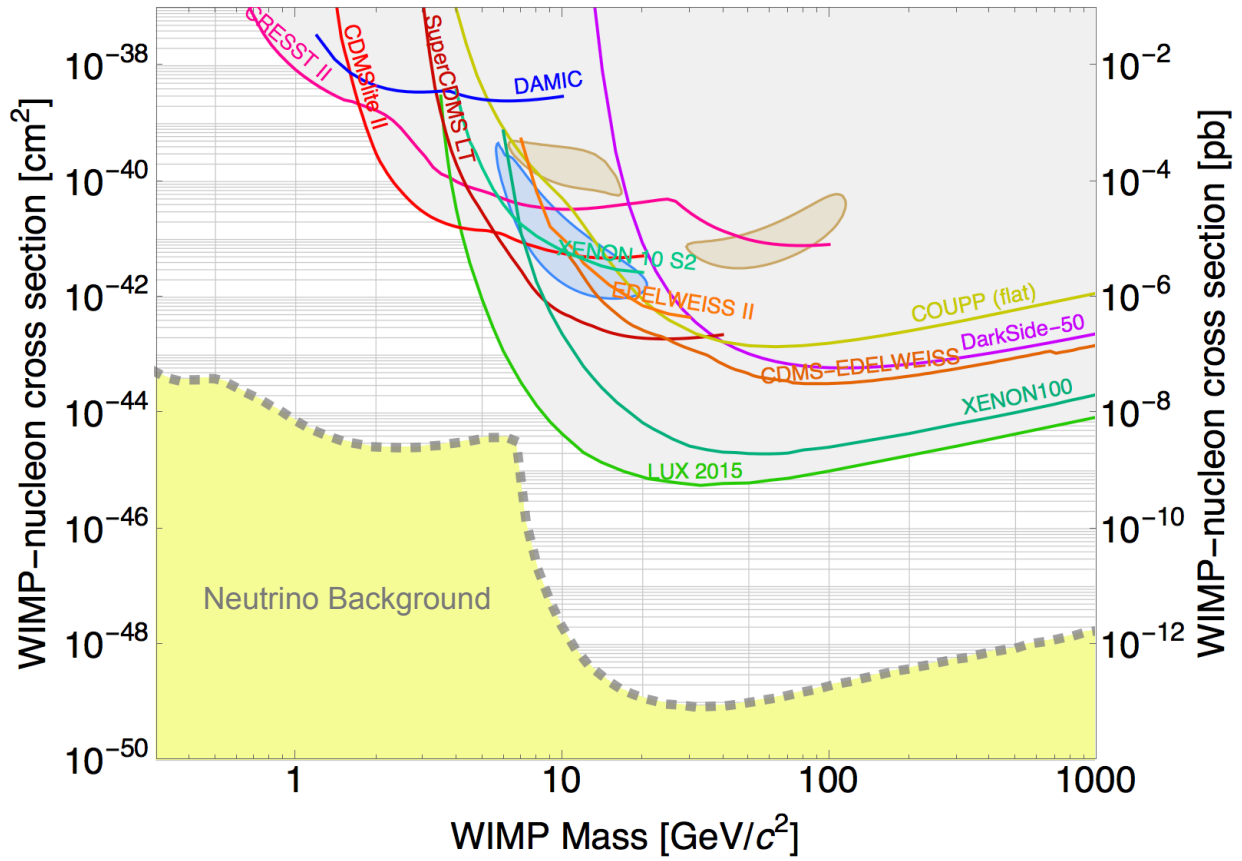


Figure 1.13: Current direct detection limits. Lines indicate exclusion limits, while contours indicate regions of interest (possible detections). Figure generated using the CDMS internal limit plotter.

The recoil energy is given by

$$E_r = \frac{\mu_N^2 v^2 (1 - \cos \theta^*)}{m_N}, \quad (1.4)$$

where θ^* is the scattering angle in the center-of-mass frame and $\mu_N = m_\chi m_N / (m_\chi + m_N)$ is the reduced mass. Backscattering ($\theta^* = \pi$) deposits the maximum recoil energy, given by

$$E_{r,max} = \frac{2\mu_N^2 v^2}{m_N}. \quad (1.5)$$

For fixed recoil energy, this corresponds to a required minimum WIMP velocity of

$$v_{min} = \sqrt{\frac{E_r m_N}{2\mu_N^2}}. \quad (1.6)$$

A 100 GeV WIMP has a kinetic energy of roughly 50 GeV. The mass of a germanium nucleus is 67 GeV. The average ratio of kinetic energy transferred to a target nucleus is given by

$$\frac{E_r}{E_\chi} = \frac{2 \left(\frac{m_N}{m_\chi} \right)}{\left(1 + \frac{m_N}{m_\chi} \right)^2}. \quad (1.7)$$

A typical collision will therefore result in about 25 keV of deposited recoil energy, easily measured by low threshold detectors. If we instead consider a target electron ($m_N = m_e$), we find an energy transfer of less than 1 eV — extremely difficult to detect.[92] It is therefore only nuclear scattering which needs to be considered.

The differential scattering rate per kg of target material may be written as

$$\frac{dR}{dE_r} = \frac{1}{m_N} \frac{d(N_\chi \langle \sigma v \rangle)}{dE_r} = \frac{\rho_0}{m_N m_\chi} \int_{v_{min}}^{v_{max}} \left(\frac{d\sigma}{dE_r} \right) v f(\vec{v}) d^3v, \quad (1.8)$$

where N_χ is the number density of WIMPs, ρ_0 the average WIMP mass density, $f(\vec{v})$ the normalized velocity distribution function, and $d\sigma/dE_r$ the differential cross section. The differential cross section is typically modeled as a point-contact interaction with cross section σ_0 coupled to a nuclear form factor F and is given by

$$\frac{d\sigma}{dE_r} = \frac{\sigma_0}{E_{r,max}} F^2(q), \quad (1.9)$$

where $q = \sqrt{2m_N E_r}$ is the momentum transfer.

The deBroglie wavelength of a 100 GeV WIMP is about the diameter of a large nucleus. This causes the WIMP to interact coherently with the entire nucleus, rather than just with individual nucleons. The standard (spin-independent) point-contact interaction cross section is taken to be

$$\sigma_0 = \frac{4\mu_N^2}{\pi} (Zf_p + (A - Z)f_n)^2, \quad (1.10)$$

where Z is the total number of protons, A the total number of nucleons, and f_p and f_n are the coupling strengths to protons and neutrons, respectively. Most supersymmetric models assume $f_p \approx f_n = f$, simplifying this to

$$\sigma_0 = \frac{4\mu_N^2}{\pi} A^2 f^2. \quad (1.11)$$

The A^2 dependence causes WIMP scattering to become significantly enhanced by heavy nuclei ($A^2 \sim 5000$), making germanium and xenon favored targets.

The nuclear form factor $F(q)$ arises from the fact that as the momentum transfer increases, the structure of the atom becomes more apparent and constructive coherence is lost.

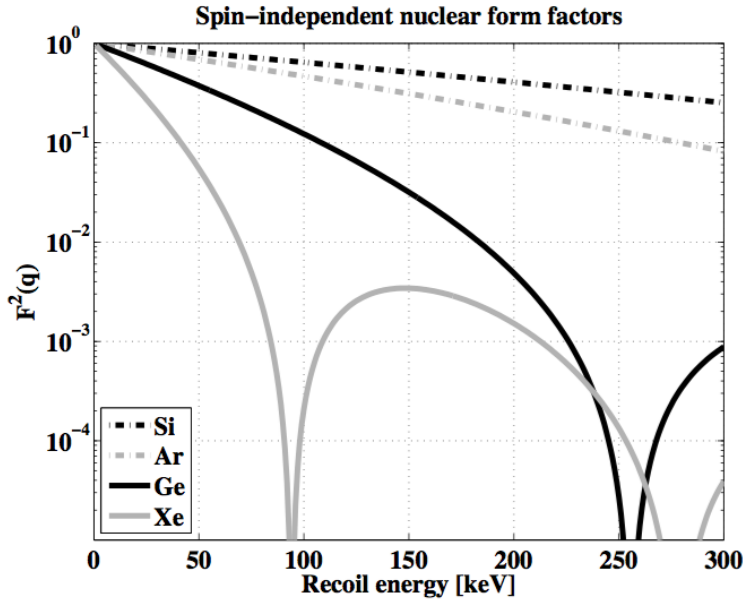


Figure 1.14: The Helm form factor for various target nuclei. Note the troughs at high recoil energy due to the loss of constructive coherence. Figure from [92].

The standard form factor used is the Helm form factor.[93–95] The Helm form factor for several target nuclei are shown in Fig. 1.14.

Eq. 1.11 gives the WIMP interaction cross section per target *nucleus*. As different experiments use different target materials, it is more useful to express the cross section per target *nucleon*. The WIMP-nucleon cross section is normalized to the value for a single proton and is given by

$$\sigma_p = \frac{\mu_p^2}{A^2 \mu_N^2} \sigma_0, \quad (1.12)$$

where μ_p is the WIMP-proton reduced mass. This can be inserted into Eq. 1.8 and simplified to obtain (per kg of target material)

$$\frac{dR}{dE_r} = \frac{1}{2} \frac{\rho_0 A^2 \sigma_p}{\mu_p^2 m_\chi} \int_{v_{min}}^{v_{max}} \frac{F^2(q)}{v} d^3v \quad (1.13)$$

The spin-independent differential scattering rates for several WIMP masses and target materials are shown in Fig. 1.15, in the community-standard units of WIMP scattering events per keV (of recoil energy) per kg (of target material) per day (of exposure). Note that for lower WIMP masses, lighter target nuclei have a greater rate at higher recoil energy due to the favorable kinematics. Spin-dependent cross sections can be constructed in a similar fashion.[97, 98]

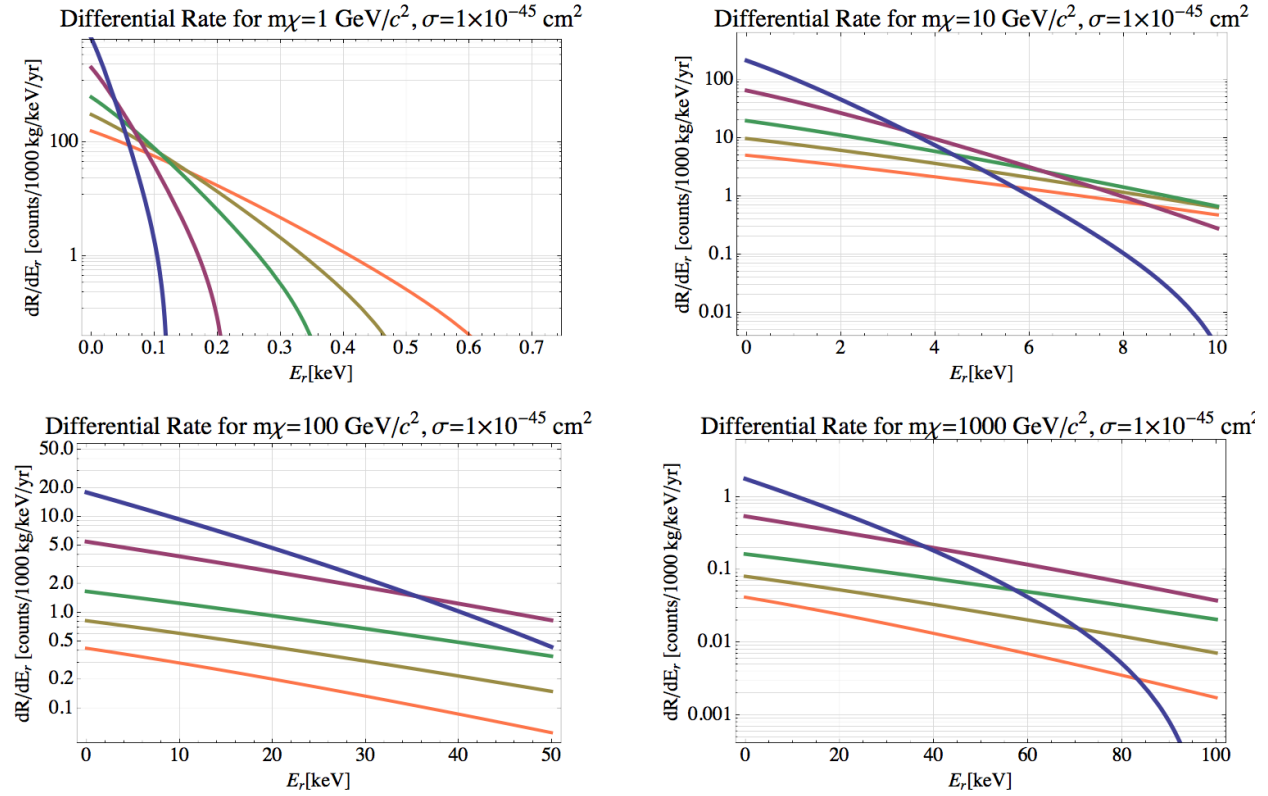


Figure 1.15: Differential scattering rate versus recoil energy for 1 GeV (*upper left*), 10 GeV (*upper right*), 100 GeV (*lower left*) and 1 TeV (*lower right*) WIMPs, shown for xenon (blue), germanium (purple), argon (green), silicon (brown) and neon (orange) target nuclei. Standard halo parameters and a spin-independent cross section of 10^{-45} cm^2 is assumed. Figure from [96].

1.5.2 Annual Modulation

As the Earth orbits the Sun, its velocity with respect to the WIMP halo varies. This causes a periodic variation — an annual modulation — in the WIMP-detector event rate, which is expected to reach a maximum in June and a minimum in December.[99] The change in event rate predicted by most WIMP models is expected to be on the order of a few percent. Observation of a modulation signal matching these expectations may be interpreted as evidence of WIMPs. Most notably, the DAMA/LIBRA collaboration[80] has observed a strong modulation signal (9.2σ significance) over 7 annual cycles, as seen in Fig. 1.16. The NaI crystals used by DAMA/LIBRA, however, have no discrimination capability between electromagnetic background and nuclear recoils, and several backgrounds are also expected to undergo a similar annual modulation. A signal of this magnitude, if attributable to nuclear recoils, would be easily observable in other direct detection experiments (such as CDMS).

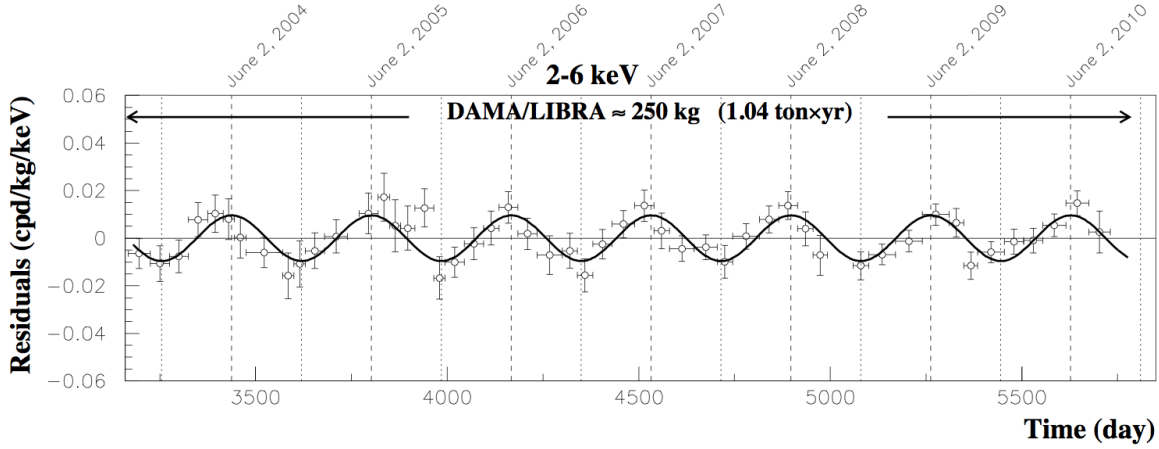


Figure 1.16: Experimental residual rate of single-hit scintillation events measured by DAMA/LIBRA in the 2-6 keV energy range. The dashed vertical lines show the expected maxima from the standard WIMP signal. From [80].

The standard WIMP parameter space this corresponds to has been ruled out by several experiments. This requires an obscure dark matter model which preferentially scatters off NaI over several other target nuclei. While the DAMA/LIBRA collaboration stand behind their result, it is difficult to see how it could arise from interactions with WIMPs. Several additional NaI experiments are planned to settle this conflict and it will be interesting to see the end result.

The CoGeNT collaboration also reported observation of an annual modulation signal in a p-type point contact germanium detector.[100] WIMP models disagree on whether this signal is consistent with DAMA/LIBRA. CDMS, however, has also used germanium detectors in the same mine—just down the hallway—and does not observe an annual modulation signal.[101] Combined with other null results, this signal has been essentially ruled out as WIMPs.

1.5.3 Halo Model

So far, we have assumed a Maxwell-Boltzmann velocity distribution for the dark matter halo. Several other functional forms have been proposed, such as the Navarro-Frank-White and Einasto distributions, which would modify the expected interaction rate.[102, 103] Recent work by Bozorgnia, however, using high-resolution hydrodynamic simulations of Milky Way analogues has shown that event rates in direct detection experiments obtained from the best fit Maxwellian distribution (with peak speed of 223-289 km/sec) are similar to those obtained directly from the simulations.[104] The standard halo model is robust and a good choice for use in interpreting experimental results.

1.5.4 Sensitivity Limits and Backgrounds

The experimentally-observed interaction rate, which is to be compared to Eq. 1.13, is given by the ratio of observed events to the total exposure (total target mass $\times$ total search time $= MT$). The minimum measurable sensitivity will generally decrease as the total exposure increases. As discussed in [92, 105], the increase in sensitivity with exposure depends on the level of background contamination in the signal region. There are three important cases:

- **Background free:** The optimal situation is obviously if there are no background events in the signal region. In this case, the 90% Poisson upper limit on the interaction rate (and hence inverse cross section) is given by $r \approx 2.3/(MT)$. Thus sensitivity scales linearly with exposure.
- **Background subtraction:** If the signal region is polluted by background events *which come from a well-understood distribution*, background subtraction can be performed. Assuming the number of background events also scales linearly with the exposure, the statistical uncertainty in the number of background events to subtract will be proportional to $\sqrt{MT}$. Since this also translates into the uncertainty in true WIMP events, the sensitivity scaling with exposure reduces to $\sqrt{MT}$.
- **Background limited:** While the relative statistical error incurred in the presence of a well-understood background can be reduced by increasing the exposure, at some point systematic error or poorly-understood backgrounds dominate. In this case, the *error* in the background distribution also tends to scale linearly with exposure. Increasing exposure will not improve sensitivity limits as signal and background can no longer be distinguished — the additional error in the number of backgrounds introduced by the longer exposure is equal to the expected increase in WIMP events. This is the worst situation. The experiment is background limited and no improvement in sensitivity is possible.

Considerable amount of effort is spent in the data analysis of a direct detection experiment as a number of event cuts and background models must be taken into consideration to define the expected WIMP signal region. Generically, the end result is some number (often zero) of possible candidate WIMP events and associated recoil energies for a given amount of exposure. Statistical exclusion limits (Fig. 1.13) are then set using a variety of methods, the most common being the Yellin optimal interval method.[106] Finally we note that it is much easier to increase exposure by increasing the total target mass rather than WIMP search time — most would prefer to take the time to build a 10x bigger detector and perform a 1-year search rather than use the same detector for 10 years.

Chapter 2

The Cryogenic Dark Matter Search

2.1 Overview

In the mid-1980s, there was great interest within the scientific community to develop a cryogenic dark matter detector. As summarized by Sadoulet[107], such a detector must have a low energy threshold, high sensitivity to WIMP interactions, and excellent background rejection. It was recognized that semiconductor crystals, cooled to sub-Kelvin temperatures and instrumented with phonon sensors, would provide all but the necessary background rejection for a large-scale dark matter search.

In 1990, Shutt and Sadoulet at Berkeley made a breakthrough when they observed the collection of ionized charge carriers in addition to the thermal phonon signal in a germanium NTD thermistor-based detector.[108] While it was naively expected that an ionization signal would not be observable with such weak ($\sim V/cm$) electric fields and low temperatures, the simultaneous measurement of ionization and phonons produced within the detector by incident particles would allow for the discrimination between electron recoils produced by electromagnetic background and nuclear recoils produced by scattering dark matter. This led to the development of the Berkeley Large Ionization and Phonon (BLIP) detector, capable of event-by-event discrimination between electron and nuclear recoils.[109] These detectors, however, misidentified events occurring near the detector surfaces resulting in a large population of background events.

During the same period of time, the Cabrera group at Stanford was developing thin-film superconductors for the purpose of building a solid-state neutrino detector.[110, 111] By depositing thin films onto a semiconductor substrate and biasing them along their superconducting transition, these transition edge sensors (TESs) are able to measure the higher-energy *athermal* phonons produced by particles interacting with the atoms of the semiconductor. The athermal phonons encode information about the location of the event within the semiconductor and could identify the surface events which plagued the BLIP detectors, replacing the NTD technology.

A collaboration was formed between Berkeley, Stanford, and Santa Barbara to construct a

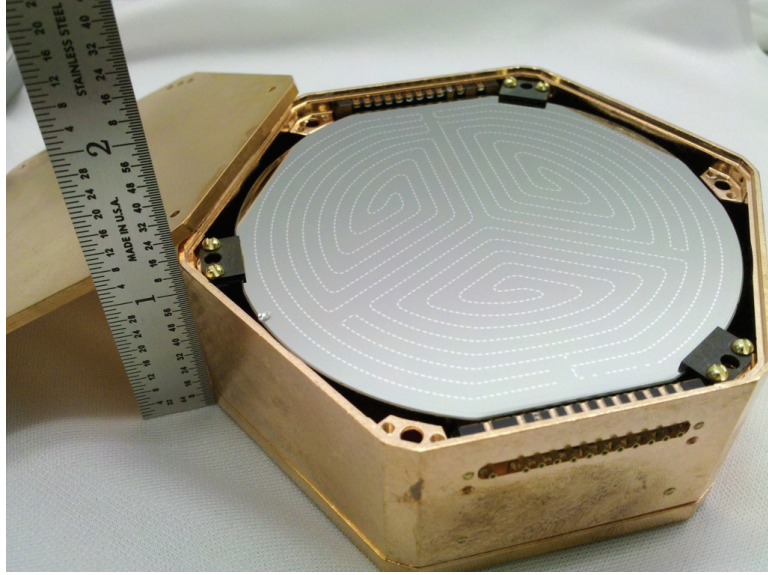


Figure 2.1: A modern CDMS ZIP detector.

direct detection dark matter experiment using cryogenic semiconductor detectors. A shielded cryostat was built and installed at the shallow-site Stanford Underground Facility (SUF), which allowed for rapid development of a new detector incorporating the ideas of Sadoulet and Cabrera.[112] The result of this collaboration, the Z-dependent Ionization and Phonon (ZIP) detector, was a great success and the first WIMP limits using these devices were published in 2000.[113] This initial collaboration, named the Cryogenic Dark Matter Search (CDMS), blossomed into a much larger project, culminating in ~ 13 years of operation half a mile underground at the Soudan Underground Laboratory which has led to several world-leading WIMP detection limits. The collaboration now consists of 22 separate institutions who are presently working on the next stage of the experiment, SuperCDMS SNOLAB. This chapter aims to summarize the CDMS experiment, discussing the basic detector design, operation, major WIMP search results, and the future at SNOLAB.

2.2 ZIP Detectors

The Z-dependent Ionization and Phonon (ZIP) detector is a cylindrical puck of high purity germanium or silicon equipped with ionization collecting electrodes and athermal phonon transition edge sensors. The detectors are cooled to about 40 mK, allowing the TESs to be biased on their superconducting transition and eliminating bulk thermal phonons and charge carriers within the crystal — the detectors are said to be "frozen out." A weak electric field (~ 1 V/cm) is applied between the ionization electrodes and phonon sensors, through the bulk of the crystal. Incident particles scatter off atoms in the crystal and produce free charge carriers and athermal phonons, generating an electrical response (an event, see Fig. 2.2) in

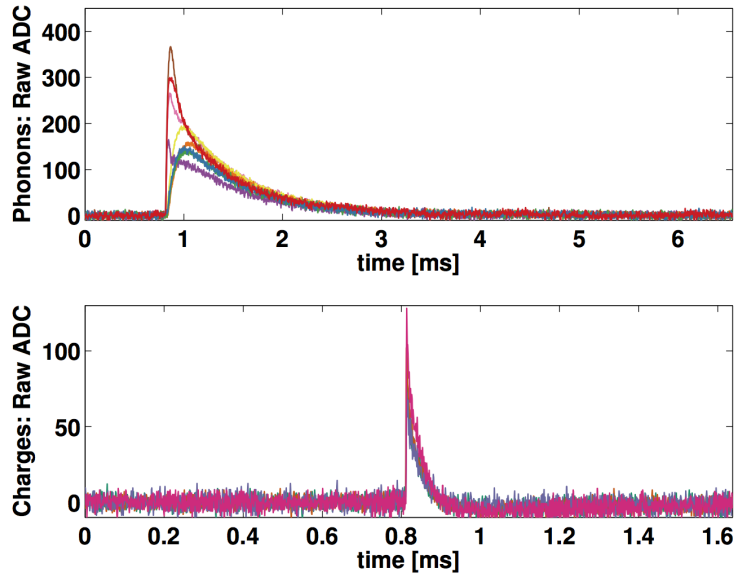


Figure 2.2: CDMS detector response to a typical scattering event. Phonon traces are shown in the top panel, while ionization traces are shown in the bottom panel.[114]

the sensors which is amplified and recorded. Information about the recoil energy, location, and recoil type is encoded in the size, shape, and distribution of the signal amongst the various sensors. This allows for event-by-event discrimination between a possible WIMP signal and uninteresting background events.

2.2.1 Phonon Measurement

An interacting particle will create a local deposition of energy confined to a small volume of the detector. Some of this energy goes into the production of quantized vibrational modes of the crystal lattice — out-of-equilibrium *athermal* phonons — which then propagate through the crystal. The initial high energy phonons move away from the interaction point in a quasi-diffusive fashion, quickly decaying via anharmonic processes into multiple lower energy phonons. When the mean free path of the phonons becomes comparable to the dimensions of the detector, they travel ballistically towards the surfaces where they are either absorbed by the sensors or reflected back into the crystal. If the phonons spend enough time in the crystal before being absorbed (due to multiple reflections, for example) they are said to be thermalized, randomizing their momentum, and no longer provide information about the location of the initial particle interaction. The power of CDMS detectors comes from their ability to measure the low energy athermal phonons before the onset of thermalization.

The phonons are measured using Quasiparticle-trap-assisted Electrothermal-feedback Transition edge sensors (QETs). A QET consists of overlapping tungsten (W) and aluminum (Al) thin film superconductors deposited on the top and bottom surfaces of the

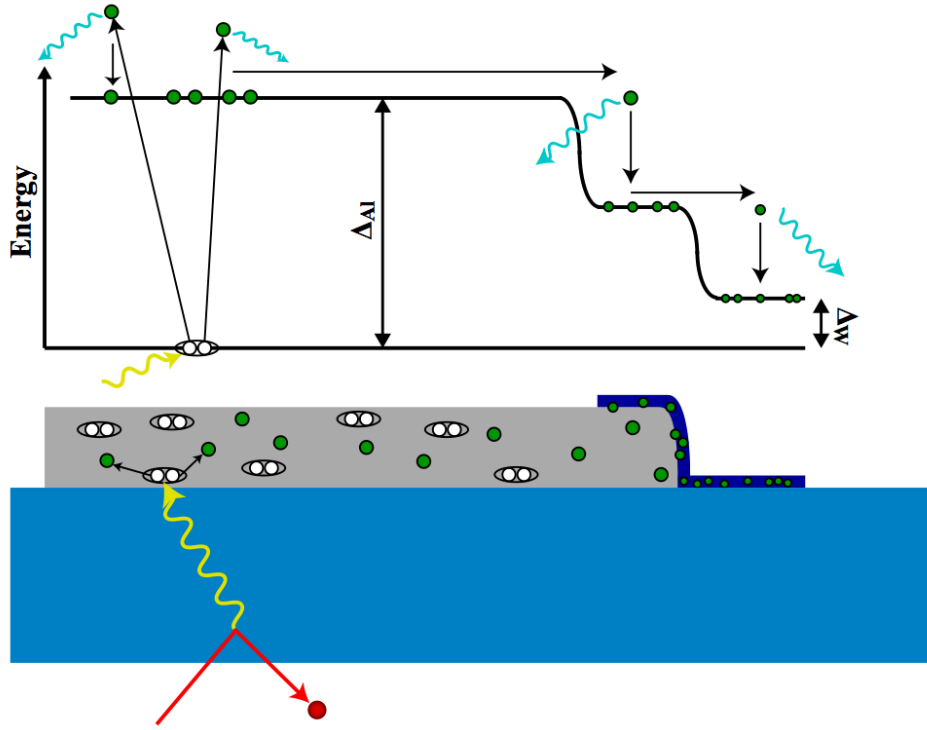


Figure 2.3: Schematic of athermal phonon collection and quasiparticle diffusion. High energy phonons created by an interaction break Cooper pairs in the Al absorbing film. The created quasiparticles diffuse into the W TES, where they deposit their energy and become trapped.[115]

germanium or silicon crystal. The aluminum fins serve as phonon collectors which couple to the tungsten TESs.

At the operating temperature of the detector, the aluminum is fully superconducting. Phonons from the germanium/silicon are absorbed by the aluminum and break Cooper pairs, creating quasiparticles which diffuse towards the tungsten. Due to the lower superconducting gap energy, quasiparticles which enter the tungsten are effectively trapped and deposit their energy into this thin film.

The tungsten is electrically biased to lie on the transition between its superconducting and normal resistive states. This is accomplished by Joule heating using an external voltage source. This voltage biasing scheme, originally proposed by Irwin[116], places the tungsten in an electrothermal-feedback configuration. The resistance of the tungsten increases with temperature, while the power dissipation due to the voltage bias is $\propto V^2/R$ and thus decreases with increasing temperature. The negative feedback creates a stable bias point which holds the TES at a fixed temperature on long time scales and results in a pulse-like change in the current through the sensor when phonons are absorbed. By placing the input coil of

a Superconducting Quantum Interference Device (SQUID) in series with the TES, the small change in current due to phonon absorption can be sensed and amplified.

Modern CDMS detectors are equipped with 8 phonon channels, 4 per side, each with many QETs placed in parallel to allow sensitivity over a large crystal surface area. The channels are arranged such that an outer ring surrounds three evenly-distributed inner channels, as can be seen in Fig. 2.2. As the QETs are sensitive to athermal phonons, comparison of the relative pulse height and shape of each sensor allows an event-by-event determination of the approximate position within the detector where the recoil occurred. This information can be used to exclude problematic events which occur near the edges and surfaces of the detector from the WIMP search data. The traditional neutron transmutation doped (NTD) Ge sensors used, for example, by the EDELWEISS experiment are mostly sensitive to thermal phonons and only provide a measure of the energy of the recoil event.

2.2.2 Ionization Measurement

Germanium, being a semiconductor, has a Fermi level which lies in the band gap separating the conduction and valence bands. At milliKelvin temperatures there is not enough thermal energy to excite an electron across the band gap into the conduction band. This means at equilibrium no free charge carriers are present in the crystal and it does not conduct — it is said to be frozen out.

The initial localized energy deposition from a recoiling particle will excite a number of electrons across the band gap into the conduction band. Each now-vacant state in the valence band will also act as a positively-charged particle, known as a hole. A cloud of free electron-hole pairs is generating in vicinity to the interaction site.

In the absence of an external field the electron-hole pairs will quickly recombine. If an electric field is present, however, the charge cloud is pulled apart. The electrons and holes drift in opposite directions preventing recombination and results in a current — the *drift current* — through the detector.

Typically the external field is established by applying a voltage across metallic ionization electrodes deposited on the flat surfaces of the crystal. By Ramo's theorem[117], the drift current through the crystal will also produce a current in the metallic electrodes. This current can essentially be thought of as the moving image charges generated in a metallic conductor due to the movement of free charge carriers within the crystal. This image current can be sensed by attaching a charge amplifier (discussed in detail in later chapters) to the ionization electrode, effectively providing a measurement of the ionization produced by the recoiling particle. Once the free charge carriers reach the electrodes, they quickly relax to the Fermi level of the metal. This be visualized as the free carriers in the germanium recombining with the image current in the electrode at the interface.

The ionization electrodes of CDMS detectors are formed by first placing a layer of amorphous silicon (a-Si) onto the bare germanium which is then covered with boron-doped aluminum. The a-Si forms a blocking region which partially prevents back-diffusion of carriers in charge clouds produced near the electrodes before the external field is able to separate

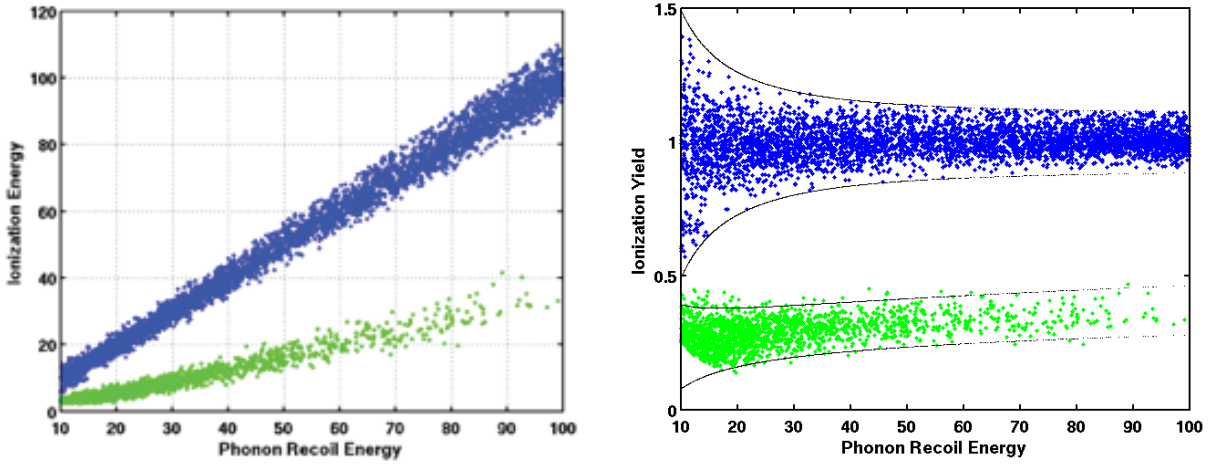


Figure 2.4: *Left:* Ionization energy versus phonon recoil energy for a detector which has been exposed to a neutron source. *Right:* Ionization yield as a function of energy. The discrimination between electron recoil background (blue) and nuclear recoil signal (green) regions is clearly observed.

them.[118] This back-diffusion causes an undesirable reduction in the ionization signal for events near the surface of the crystal. These surface events are a major background for the WIMP search experiment and are discussed later.

A modern CDMS detector consists of four ionization electrodes, two on each surface. In standard operation, non-zero bias voltages are applied to these electrodes while the phonon sensors serve as the electrical ground. On a given side of the detector, the two ionization electrodes differ in shape — an outer ring surrounds a much larger disc. Comparison of the relative amplitude of the signals produced in the inner and outer electrodes provides additional information on the radial position of the event. Events with a large outer electrode signal, indicating the event occurred close to the problematic edges of the detector, can be removed from the WIMP search.

2.2.3 Discrimination by Ionization Yield

The energy deposited by a recoiling particle creates both phonons and electron-hole pairs. The distribution of the recoil energy between these two types of excitations depends on the identity of the interacting particle. By measuring the *ionization yield*, the ratio of energy used to create electron-hole pairs to the energy used to create phonons, CDMS detectors are able to distinguish between uninteresting background events and possible WIMP candidates. The following provides a brief description of the origin of this difference in ionization yield

between electron recoils and nuclear recoils. A detailed discussion of the energy loss by a recoiling particle can be found in [119].

First consider the case of an interacting electron. Due to kinematics (an electron has a much smaller mass than a nucleus) this electron cannot transfer a significant energy to any nuclei — the energy initially goes into the creation of highly energetic electron-hole pairs. These electrons and holes shed their energy through the emission of phonons and creation of additional electron-hole pairs. This process continues until the electrons and holes no longer have enough energy to produce new charge carriers and have relaxed to their respective band minima via phonon emission. The end result is a population of free electrons and holes and a population of phonons.

Due to energy losses from phonon production, on average it takes roughly three times the band gap energy to create an electron-hole pair. For a given recoil energy, the number of electron-hole pairs created is thus given by

$$n_{eh} = E_{recoil}/E_{create} \quad (2.1)$$

where E_{create} is 2.96 eV for Ge and 3.81 eV for Si. At the end of the cascade process, the electron-hole pairs have relaxed to the band gap energy E_{gap} (0.785 eV for Ge, 1.21 eV for Si). The amount of energy stored in the produced phonons is therefore

$$E_{phonon} = E_{recoil} - E_{charge} = E_{recoil} - n_{eh}E_{gap} = E_{recoil} \left(1 - \frac{E_{gap}}{E_{create}} \right) \quad (2.2)$$

where E_{charge} is the energy stored in electron-hole pairs.

When the free electrons and holes reach their respective electrodes, they relax to the Fermi level of the metal through more phonon emission. The energy $n_{eh}E_{gap}$ lost in the creation of the electron-hole pairs is recovered in the form of these recombination phonons. In the absense of charge trapping, discussed later, *all of the initial recoil energy has been converted to phonons.*

Now consider the case of an interacting neutron. In addition to the creation of electron-hole pairs, which follows the same cascade process outlined above, some energy can now be transferred to the nuclei of the crystal. These nuclei lose energy through the excitation of other nuclei, the emission of phonons, and the creation of additional electron-hole pairs. As the energy of the nuclei decreases, it becomes much more probable to lose energy through excitation of other nuclei/phonon emission rather than the creation of additional charge carriers. Since there is no E_{gap} for the nuclei, their energy is fully liberated as phonons. The net result is that this additional energy loss mechanism through interaction with the nuclei causes less electron-hole pairs to be produced. Again, at the end of the event all of the recoil energy is converted to phonons however **the intermediate ionization signal has been reduced.**

The theory of the distribution of recoil energy between phonon and electron-hole pair production was studied in detail by Lindhard *et al.*[120], resulting in integral equations which must be numerically solved. The *ionization yield*, defined as the ratio of recoil charge

energy to recoil phonon energy, normalized such that an electron recoil has a yield of one, describes the variation of this partitioning as a function of recoil energy. For nuclear recoils, the yield can be approximated as

$$Y_{NR}(\epsilon) = \frac{k \cdot g(\epsilon)}{1 + k \cdot g(\epsilon)} \quad (2.3)$$

where

$$\begin{aligned} k &= 0.133 Z^{2/3} A^{-1/2} \\ \epsilon &= 11.5 Z^{-7/3} E_{recoil} \text{ (in keV)} \\ g(\epsilon) &\approx 3\epsilon^{0.15} + 0.7\epsilon^{0.6} + \epsilon \end{aligned} \quad (2.4)$$

A comparison of ionization yield measurements to the Lindhard theory predictions are shown in Fig. 2.2.3.

Due to their lack of electromagnetic interactions and relatively heavy mass, WIMPs are expected to primarily interact with the nuclei of CDMS detectors. These interactions should have an ionization yield described by Eq. 2.3. The nuclear recoil response of CDMS detectors is determined by exposure to radioactive ^{252}Cf , a strong neutron emitter. As seen in Fig. 2.2.3, electron recoils are clearly distinguished from the nuclear recoils induced by the source. Any candidate WIMP interactions must lie within this nuclear recoil band. As it is relatively easy to shield against neutrons, discrimination by ionization yield is a powerful method to eliminate essentially all of the electromagnetic background events.

2.2.4 Luke-Neganov Amplification

Up to this point, we have only discussed phonons which are generated by the primary particle interaction. Drifting charge carriers produce additional phonons, known as Luke-Neganov (or simply Luke, for brevity) phonons[121], which are also sensed by the QETs. The applied electric field causes free charge carriers to rapidly accelerate and gain energy. Although there are essentially no ambient phonons present in the crystal, the charge carriers interact with the zero-point motion of the atoms of the lattice. This interaction causes the charge carriers to spontaneously emit phonons — Luke phonons — which cause the charge carriers to lose energy and randomizes their momentum. These processes happen very rapidly, at frequencies of several GHz or higher, and results in the charge carriers moving with a constant average drift velocity which depends on the magnitude of the electric field. All of the work performed on the charge carriers by the electric field is dissipated as Luke phonons.

Consider a detector with potential difference ΔV between the top and bottom surfaces. Remembering that all the initial recoil energy is eventually converted into phonons, the *total* energy measured by the phonon sensors due to an event is

$$E_{meas} = E_{recoil} + E_{luke} = E_{recoil} + q \Delta V n_{eh} = E_{recoil} \left(1 + \frac{q \Delta V}{E_{create}} \right). \quad (2.5)$$

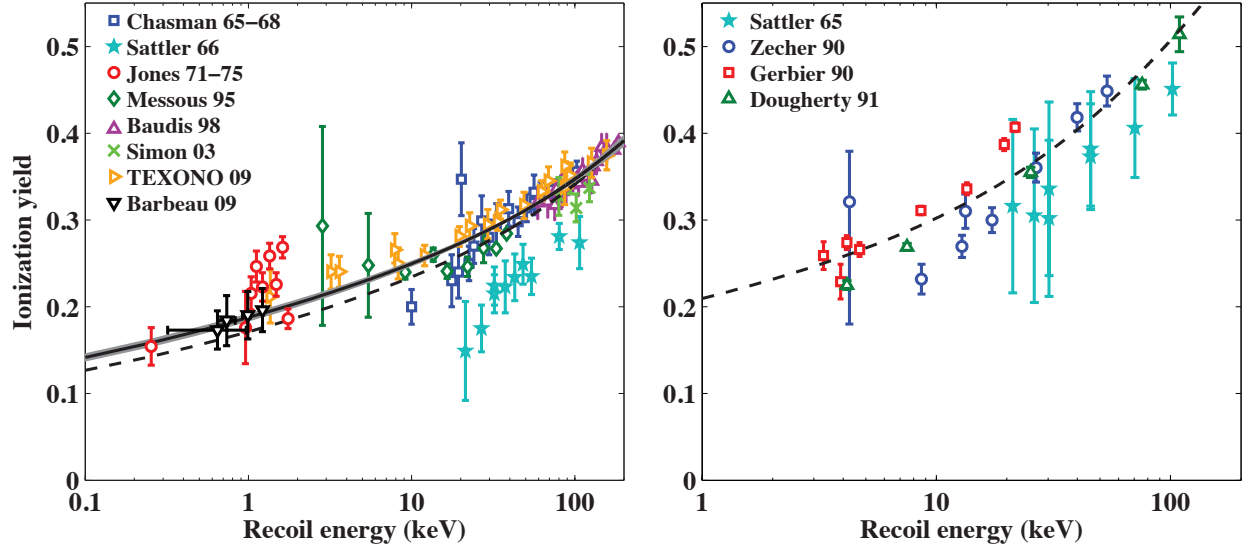


Figure 2.5: Comparison between the Lindhard model and measured ionization yield for nuclear recoils in Ge (left) and Si (right). The dashed black lines indicate the standard Lindhard predictions (Eq. 2.3) for each material. The solid black line through the Ge data is the best-fit line from allowing k to vary and including a scaling multiplicative prefactor. The data is from a number of references and the plots from a CDMS paper (in preparation) on the ionization yield.

This is the reason CDMS detectors typically operate with such weak (~ 1 V/cm) bulk electric fields. The potential difference is kept small in order to prevent Luke phonons from masking the primary recoil phonon signal. The measured charge energy from the ionization readout is used to calculate the expected Luke phonon contribution (assuming 100% charge collection), which is then subtracted from the total measured energy to recover the recoil energy. Note that electron and nuclear recoils will produce different amounts of Luke phonons — in the case of nuclear recoils E_{recoil} should be replaced with E_{recoil}/Y_{NR} .

Clearly, the number of Luke phonons produced per electron-hole pair can be increased by simply raising the bias voltage across the detector. At the same time, the electronic noise of the phonon readout amplifier should remain unchanged. Thus the signal-to-noise ratio of the *total* phonon signal increases with detector bias. Measurement of the Luke phonons can therefore result in an indirect, extremely low threshold measurement of the ionization produced by interacting particles. This technique, known as *Luke amplification*, allows CDMS detectors to probe the low-mass WIMP regime (see section on CDMSlite) at the expense of event-by-event background rejection. In principle, Luke amplification can be used to detect the production of a single electron-hole pair.

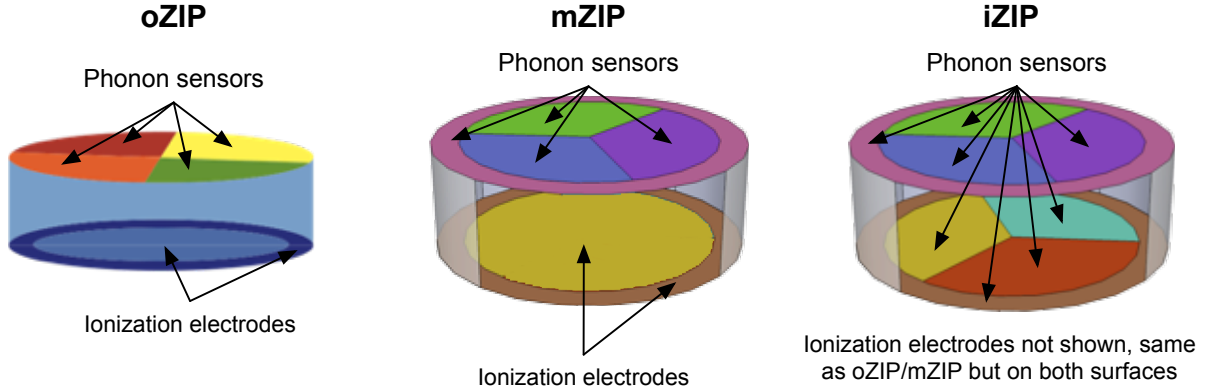


Figure 2.6: Ionization and phonon sensor layout for the oZIP, mZIP, and iZIP designs.

2.2.5 Detector Types

oZIP

The oZIP — shorthand for “original” ZIP — detector was the workhorse of the CDMS II experiment. Each oZIP detector is a 7.6 cm diameter, 1 cm thick cylindrical puck of high purity germanium or silicon with a mass of 250g (Ge)/100g (Si). One surface of the crystal is instrumented with 4 phonon sensors, segmented into quadrants. The other face consists of a large inner disk electrode and thin outer ring electrode, both connected to independent charge amplifiers. The detectors are typically operated with a 3V bias on both inner and outer charge electrodes. Surface event rejection is performed by comparing the timing and partitioning of the phonon signals, as described in [92]. These detectors are rarely used nowadays due to advances in CDMS detector technology.

mZIP

The mZIP, or “mercedes” ZIP detector, was intended to be the successor to the oZIP detector. Two major differences were introduced in the mZIP design. The first was to increase the thickness of the detector from 1 cm to 1 inch, increasing the germanium mass to 635g (250g for Si). Obviously, these detectors have better sensitivity to WIMP interactions due to the increase in the number of atoms in the crystal. Another advantage is an increase in the ratio of volume to surface area which decreases fraction of surface events.

The second major design change involved the phonon sensors. While 4 phonon sensors are still used, they are partitioned into 3 inner channels and one outer guard ring (shown in Fig. 2.6) as opposed to quadrants. This change was made to resolve degeneracies in the position reconstruction which made it difficult to correctly identify the location of certain events in the oZIP detectors. Due to the similarity in appearance of this sensor partitioning to a certain automotive company logo, these detectors were referred to as “mercedes” ZIPs.

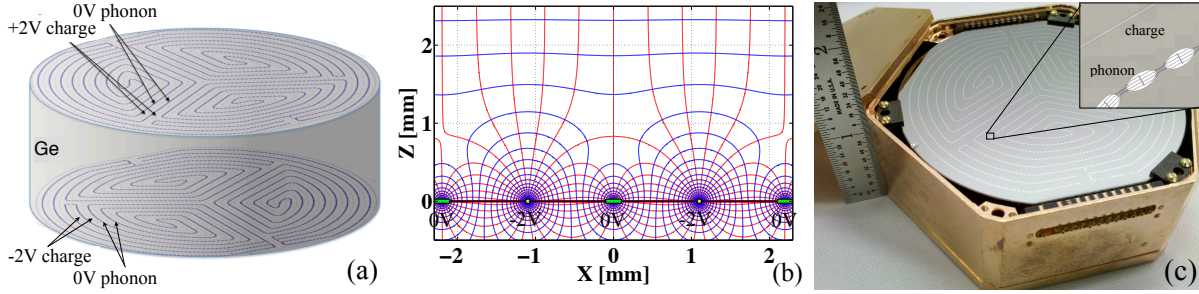


Figure 2.7: (a) Phonon and ionization sensor layout for iZIP detectors. Both faces are instrumented with ionization sensors (one face with +2 V and the other with -2 V) that are interleaved with phonon sensors (0 V) on a ~ 1 mm pitch. (b) Magnified cross section view of electric field lines (red) and equipotential contours (blue) near the bottom face of a SuperCDMS iZIP detector. The yellow ionization electrodes are thinner than the green phonon sensors. (c) Packaged iZIP detector, with a magnified image of the ionization and phonon rails.[123]

Additional improvements in phonon collection efficiency were made by changing the geometry of the individual QETs.[119, 122]

During the development and testing of the mZIP detectors, we discovered that a significant number of surface events which occurred near the charge side of the crystal would leak into the nuclear recoil band. Due to the increased thickness of the crystal, the phonon signals from these events no longer retained the identifying timing signatures of a surface event. This background leakage would strongly limit the WIMP search capabilities of the mZIP detectors and the design was abandoned. While not particularly useful as dark matter detectors, the mZIPs are incredibly valuable for the purpose of charge transport research. The large crystal thickness and uniform electric field produced by the electrode layout serve as a perfect sample for studying bulk charge transport properties under CDMS operating conditions. Essentially all the charge transport results of this thesis were obtained by studying mZIP detectors.

iZIP

The iZIP detector was developed to address the need for complete surface event rejection on both faces of the crystal and was a radical departure from the typical CDMS detector design. Instead of placing the charge and phonon sensors on opposite surfaces of the crystal, both types of sensors are placed on each surface in such a way as to approximate concentric rings of alternating charge and phonon sensors (see Fig. 2.7). This *interleaving* of the charge and phonon sensors led to adoption of the name "iZIP" for this detector design. While the physical size of the detector is the same as the mZIP, the total number of phonon sensors has been increased to 8 and the total number of charge sensors increased to 4. Each side

of the detector possesses 4 phonon sensors partitioned as in the mZIP design, and 2 charge electrodes partitioned into a main inner electrode and outer guard ring. Note that the inner electrode is no longer a disc, but a collection of rings of increasing radius wired in parallel.

The top and bottom charge electrodes are respectively biased to $\pm 2V$, while the phonon sensors are held at 0V. This bias configuration produces a bulk axial electric field, similar to the oZIP/mZIP, but near the surfaces the electric field becomes tangential. The electric field lines near the surfaces terminate on the phonon sensors and do not extend to the opposite detector face, while the bulk field lines do terminate on both faces. Bulk events will therefore produce an ionization signal on both sides, while surface events will only produce a signal on one side — one polarity of charge carriers has been absorbed by the grounded phonon sensors. The primary surface rejection cut is now based on the symmetry of the ionization signals rather than phonon timing parameters and does not depend on which surface the event occurred.

A lengthy research and development program focusing on iZIPs was undertaken, producing numerous revisions to the basic iZIP design. Detailed descriptions of the detector design and evolution can be found in [119, 122]. The iZIP detector is now the standard CDMS design and a payload of 15 iZIPs were used in the Soudan Underground Laboratory from late 2011 to the final warmup in 2015.

2.3 Problematic Phenomena and Events

A perfect detector would measure the recoil energy deposited as ionization and phonons with 100% collection efficiency, with no degradation in performance over time. Reality, however, is far from perfect and there are a number of problematic phenomena and events observed in ZIP detectors. I will now summarize the most important of these and discuss the strategies used to lessen their impact.

2.3.1 Space Charge Buildup

The bulk electric field in a CDMS detector is established by applying a voltage difference between the top and bottom surfaces of the crystal. The strength of this electric field determines the overall bulk charge collection efficiency due to its influence on the fundamental transport and capture properties of the produced charge carriers. As shown in Fig. 2.8, the charge collection efficiency is observed to degrade over time while the detectors are biased. The degradation is known to speed up with increasing total event rate. The traditional explanation for this effect was the generation of *space charge* within the crystal via charge capture processes which produce ionized impurities. These localized charged impurities produce their own electric field which competes with the external bias, resulting in a reduced net electric field — increasing the capture rate and thus reducing the ionization collection efficiency. Poor collection efficiency will result in reduced ionization yield which weakens

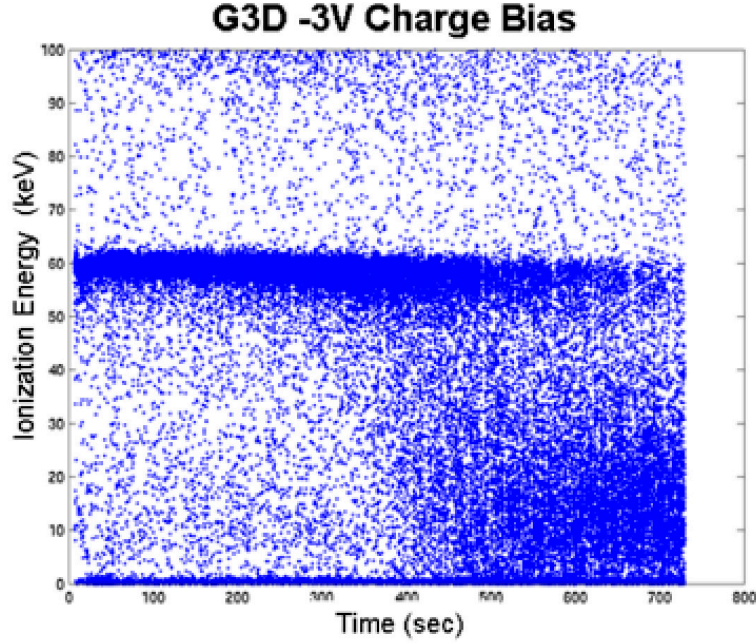


Figure 2.8: Loss of charge collection efficiency due to space charge buildup. The blue dots represent the recorded ionization energy versus time for a CDMS detector exposed to a ^{241}Am radioactive source. The collection efficiency of the 60 keV photons (thick blue band) from the source is clearly seen to degrade over time.

the electromagnetic background rejection capabilities, causing leakage into the nuclear recoil band.

Empirically, it was found that the charge collection efficiency can be reset by removing the external voltage bias (grounding the detector) and generating a large number of electron-hole pairs within the crystal. This is generally performed by flashing LEDs placed in close proximity to the crystal surface, however strong radioactive sources can also be used. This process is known within CDMS as “flashing” or “baking” the detector. Upon initial cooldown, the detectors are baked until no improvement in charge collection is observed — typically overnight. After the initial bake, periodic flashing of 5 to 10 seconds every few hours (while shielded deep underground, about every 10 minutes on the surface) is enough to maintain good charge collection. The free charge carriers created during the flash seek out and balance the number positive, negative, and neutral impurities until bulk neutrality is achieved. Understanding the detailed physics behind charge trapping and the reset procedure has been a major focus of this dissertation and is discussed in detail in later chapters.

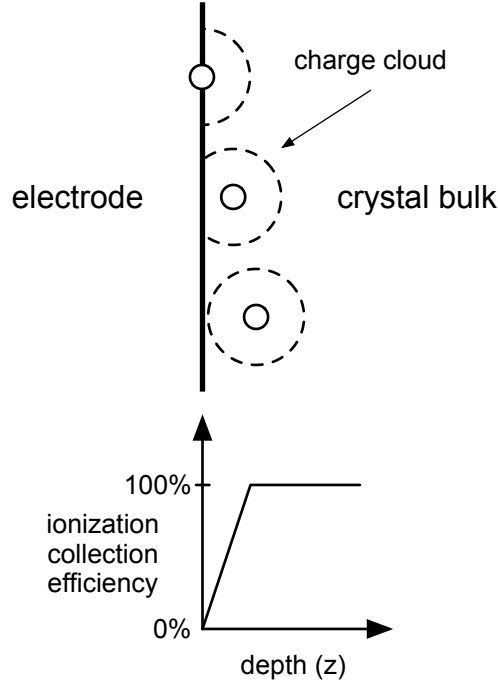


Figure 2.9: Illustration demonstrating the geometric cause of the reduced ionization collection efficiency observed in surface events.

2.3.2 Surface Events

Events which occur within $\sim 10\mu\text{m}$ of the detector surfaces suffer from reduced ionization collection and can be misidentified as nuclear recoils. Immediately after the initial energy deposition, the local electron-hole pair density is very high and the Coulomb interaction between the carriers screens the external field. The charge carriers diffusively expand in a spherical fashion until the screening is reduced and the external field separates the electrons from the holes. As shown in Fig. 2.9, this diffusive sphere is not fully contained within the bulk germanium for surface events and carriers of the wrong polarity diffuse into the wrong contact resulting in a suppressed ionization signal.

Surface events are the primary CDMS background and much of the technological innovation associated with these detectors has been for the purpose of identifying these events. For the oZIP and mZIP detectors, surface events are identified by timing parameters associated with the athermal phonon signal. The first parameter used is the 10%-40% risetime of the primary phonon pulse. The second is the delay between the ionization signal, which is collected first, and 20% of the maximum amplitude of the primary phonon signal. These two parameters are summed to form a single “phonon timing parameter” capable of rejecting electron recoils at a level $>10^6$ when combined with ionization yield, shown in Fig. 2.10.[124]

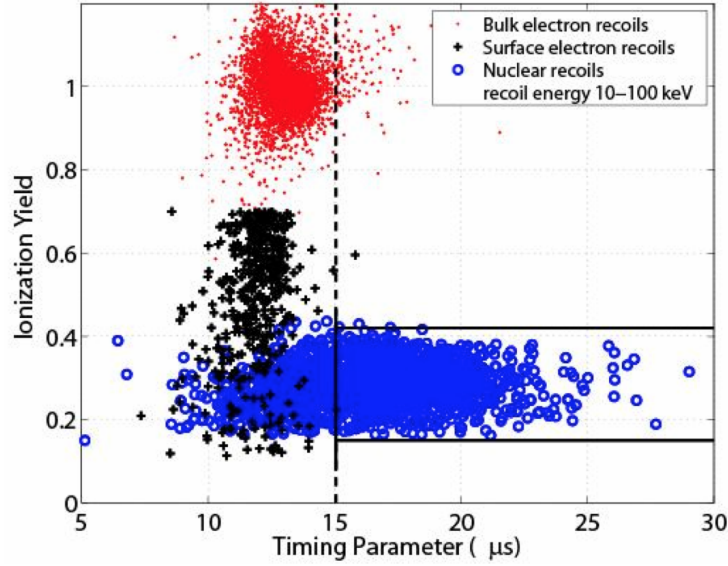


Figure 2.10: Ionization yield versus phonon timing parameter. The box in the bottom right of the plot indicates events which pass both the ionization yield and timing cut, indicating good nuclear recoils. The electron recoil surface events are clearly rejected.

While this discrimination method can still be used with iZIP detectors, a much more powerful discriminator is the ionization z-partition, which measures the difference in amplitude between charge collected on the top and bottom surfaces. Using only this z-partition, iZIP detectors are able to distinguish surface events at a level of $>10^5$ and can reject electron recoils better than 1 in 10^6 when combined with ionization yield.[123]

2.3.3 Zero Charge Events

As seen in Fig. 2.11, there is a significant population of events with ionization yield consistent with zero. These so-called zero charge events are the dominant background when extending CDMS analysis to recoil energies below ~ 10 keVnr. The phonon energy partitioning of these events indicates that they are occurring at high radius, near the edges of the detector. It is believed that these events originate from interactions near the sidewalls of the detector. All of the produced charge ends up trapped in surface states of the bare germanium, producing essentially no ionization signal. There has been some success in identifying these events through a phonon radial partition cut. A better solution would be to place ionization electrodes on the sidewalls, as has been done in the EDELWEISS experiment. For now, we accept some background due to these events when performing low energy analysis.

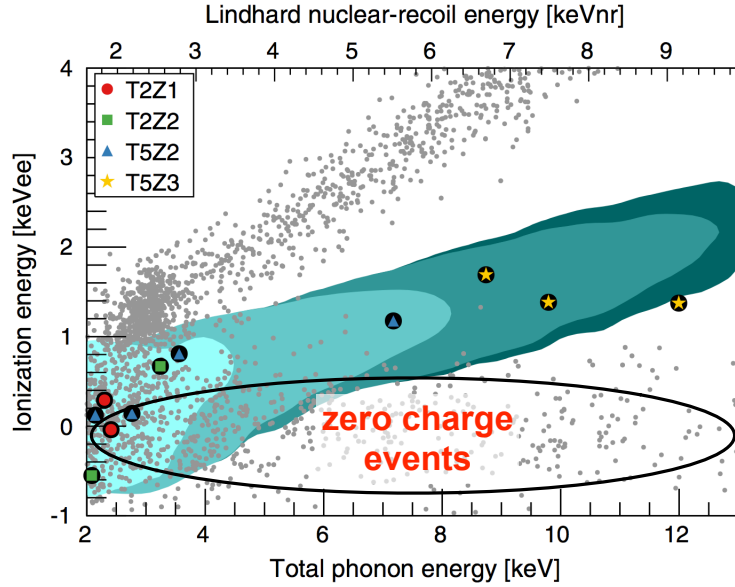


Figure 2.11: Zero charge events.

2.4 The Soudan Underground Laboratory

For over a decade, CDMS has been searching for WIMPs at the Soudan Underground Laboratory (SUL), located in the former Soudan iron mine in northern Minnesota. The nearly half-mile depth of SUL reduces the muon flux by a factor $\sim 5 \times 10^4$ compared to the surface and serves as the perfect environment for a WIMP search (see Fig. 2.12). While muons and muon-induced electromagnetic secondaries can be easily rejected by an active muon veto, energetic muons far from the veto can interact with the rock and produce neutrons which are indistinguishable from WIMPs. Reducing this neutron background is the reason CDMS and other direct detection experiments must operate deep underground. The 2090 meters water equivalent (m.w.e) overburden provided by SUL, when combined with external passive shielding which surrounds the detector payload, is sufficient to reduce the neutron background to 0.1 event for the SuperCDMS Soudan experiment.

The layout of the CDMS experiment is shown in Fig. 2.13. The detector payload is enclosed in a number of copper cans, known as the icebox, which connect to the various temperature stages of a ^3He - ^4He dilution refrigerator through a set of “tails” known as the c-stem. The icebox is surrounded by alternating layers of lead, to shield against gammas, and polyethylene, a neutron moderator. The innermost layer is formed from “ancient” lead, recovered from the ballast of a sunken ship near Nantes, France. When initially created, lead contains a large amount of radioactive ^{210}Pb , with a half life of ~ 22 years, and the decay chain leads to a significant release of electromagnetic radiation. By using ancient lead, a large

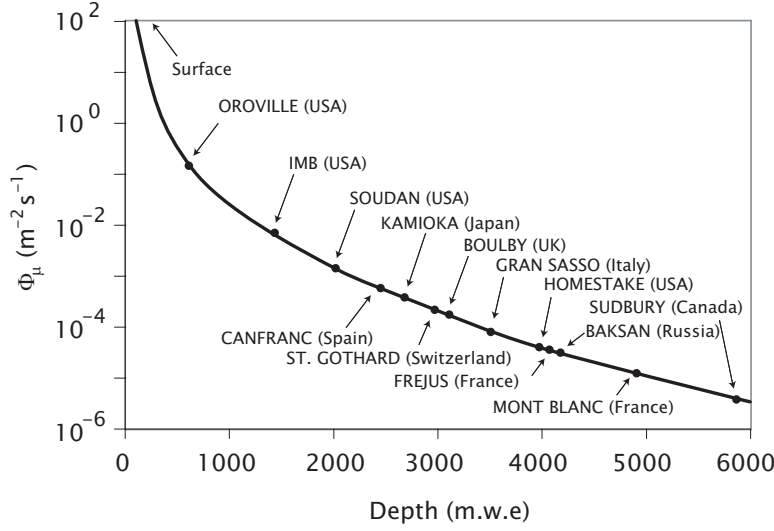


Figure 2.12: Dependence of muon flux with depth, shown with various underground facilities.[125]

fraction of the ^{210}Pb has decayed away and lowers the electromagnetic background incident upon the detectors. The layers of passive shielding are then surrounded by an active muon veto made from plastic scintillator panels, with 99.98% efficiency of detecting through-going muons. This shielding is 99% hermetic, with the only penetrations being from the c-stem for cryogenics and e-stem for electronics cabling. A Gifford-McMann cryocooler is mounted to the e-stem to mitigate the heat load introduced by the cabling. Vacuum feedthroughs at the end of the e-stem allow the detector cabling to be connected to the associated room temperature electronics. This system is capable of maintaining the detectors at a continuous base temperature of ~ 40 mK, for periods of over a year. The final incarnation of the experiment, SuperCDMS Soudan, consisted of 15 iZIP detectors arranged in 5 towers of 3 detectors, shown in Fig. 2.14. After ~ 13 years of successful operation, the CDMS Soudan experiment began its final warmup on November 18, 2015.

2.5 WIMP Search Results

2.5.1 CDMS I

The original CDMS experiment was performed at a shallow site — the Stanford Underground Facility (SUF) — from 1996-2002. The Soudan experiment was based on this design which included a shielded icebox connected to a dilution refrigerator through “tails”. The initial stages of the experiment used the Berkeley BLIP detectors[126], in which the phonon measurement was performed using NTDs. This facility served as the testbed for the first ZIP

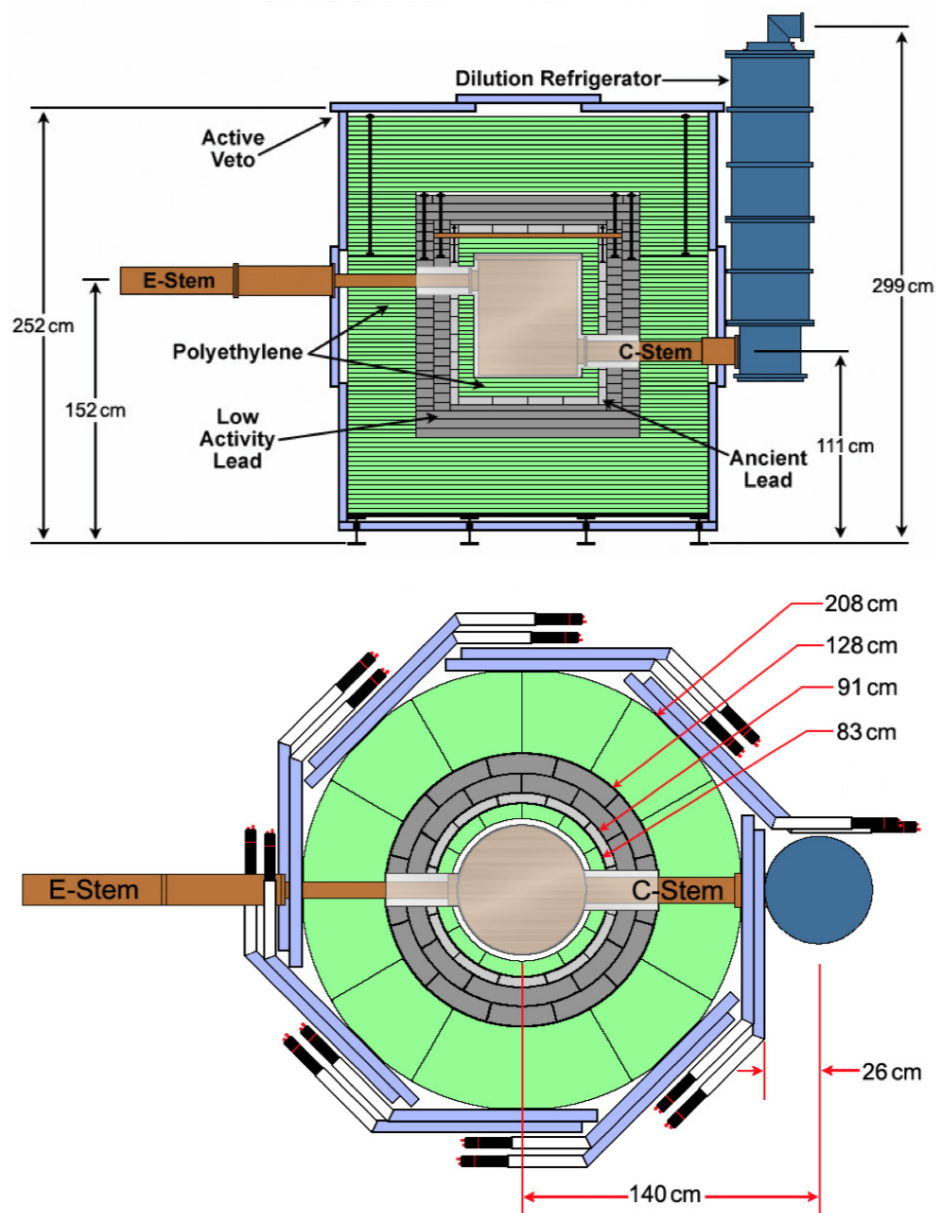


Figure 2.13: Layout of the CDMS Soudan installation. A cross sectional view is shown in the top panel, while a top-down view is shown in the bottom panel. Figure courtesy Joel Sanders.

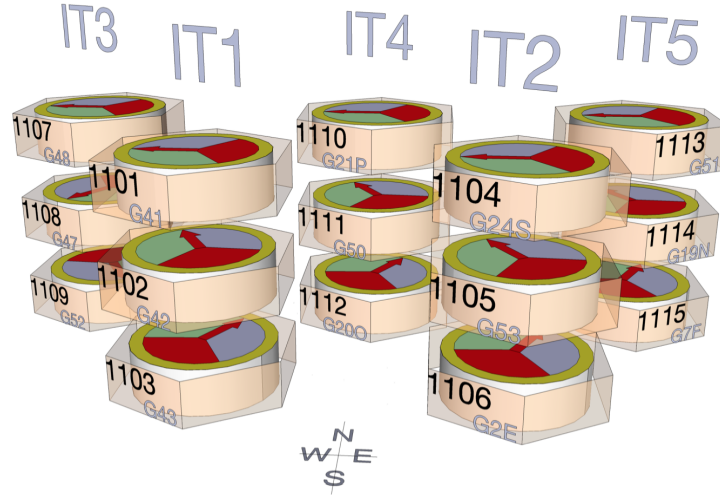


Figure 2.14: SuperCDMS Soudan detector payload, consisting of 15 iZIP detectors.

detectors, with a final payload of 4 Ge and 2 Si oZIPs and achieved 90% C.L. WIMP-nucleon elastic cross-section exclusion limits of $\sim 4 \times 10^{-42}$ for a WIMP mass of $100 \text{ GeV}/c^2$. [113] With only 16 m.w.e of overburden, the sensitivity of the experiment was ultimately limited by the neutron background however it was considered a huge success, proving the effectiveness of the CDMS detector technology.

2.5.2 CDMS II

Based on the accomplishments of CDMS I, the construction of a deep underground WIMP search facility at SUL began in 2001. The cryogenics system was fully assembled and operated in December 2002 with the first WIMP search limits published at the end of 2004. [127] The CDMS II payload consisted of 30 CDMS oZIP detectors (19 Ge/11 Si) arranged in 5 towers of 6 detectors. The final raw exposure of 612 kg-days consisted of 2 candidate events, resulting in a WIMP-nucleon spin-independent cross-section upper limit of $7.0 \times 10^{-44} \text{ cm}^2$ for a WIMP mass of $70 \text{ GeV}/c^2$. Combined with all previous CDMS II data, this upper limit is reduced to $3.8 \times 10^{-44} \text{ cm}^2$. [71] These results are shown in Fig. 2.15 along with limits set on the spin-dependent WIMP-nucleon cross-section. At the time, CDMS II was the world-leader in WIMP direct detection.

A number of other WIMP analyses have been performed using the CDMS II data, including combined limits with the EDELWEISS experiment [128], a low energy analysis which reduces the energy threshold from 10 keVnr to 2 keVnr [72], a search for inelastic dark matter [129], and an annual modulation study. [101] In addition, these datasets have been used to conduct searches for non-WIMP particles such as axions [130] and lightly ionizing particles

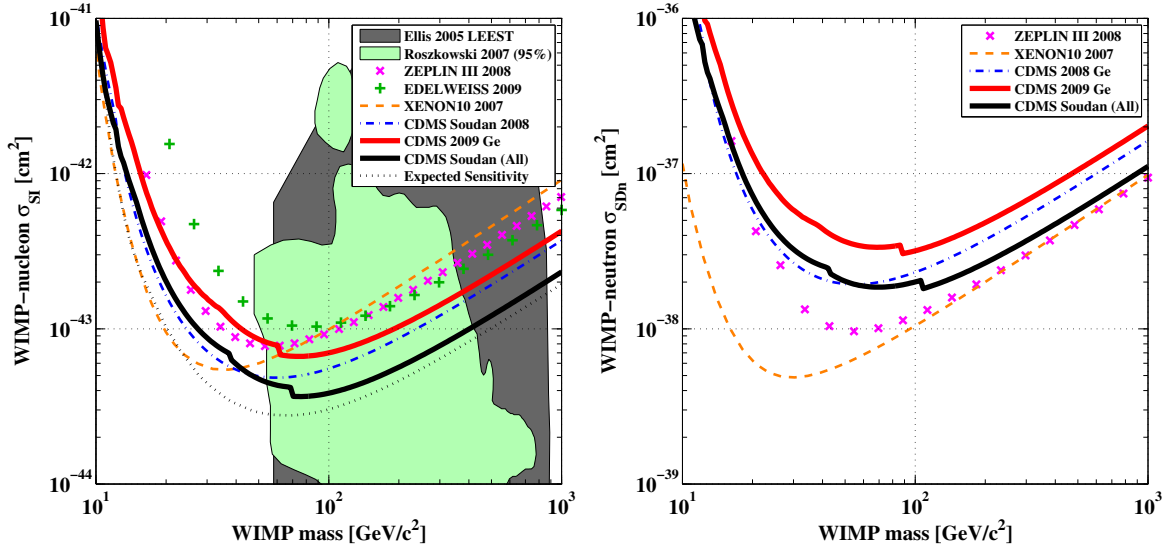


Figure 2.15: *Left:* Spin-independent WIMP-nucleon versus WIMP mass upper limits from the CDMS II experiment. *Right:* Spin-dependent limits from the same experiment.

(LIPs).[131]

In 2013, an analysis of the data from the CDMS II silicon detectors revealed 3 candidate events with an expected background of less than 1 event for a raw exposure of 140.2 kg-days.[132] A profile likelihood ratio test of the three events that included the measured recoil energies gave a 0.19% probability for the known-background-only hypothesis when tested against the alternative WIMP+background hypothesis. The highest likelihood occurred for a WIMP mass of $8.6 \text{ GeV}/c^2$ and WIMP-nucleon cross-section of $1.9 \times 10^{-41} \text{ cm}^2$. There was significant interest in this result, as many believed it could be the beginning of a WIMP signal, however subsequent experiments have essentially ruled out this interpretation.

2.5.3 SuperCDMS Soudan

After a lengthy research, development, and production phase — partially slowed by the science funding climate in the United States at the time — an upgraded payload of 15 0.6 kg Ge iZIP detectors was installed at Soudan and began operations in March 2012. The immediate goals were to demonstrate the background rejection capabilities of the iZIP design and probe the preferred WIMP region suggested by the CDMS II silicon results. Two detectors were installed with ^{210}Pb sources and were used to show surface event rejection at the level needed for the future SuperCDMS SNOLAB experiment.[123] An initial low threshold analysis, which included background events, was undertaken using 7 detectors with the lowest energy resolution and published in June, 2014.[73] With a total exposure of 577 kg-days, eleven events were observed after unblinding which set an upper limit (see

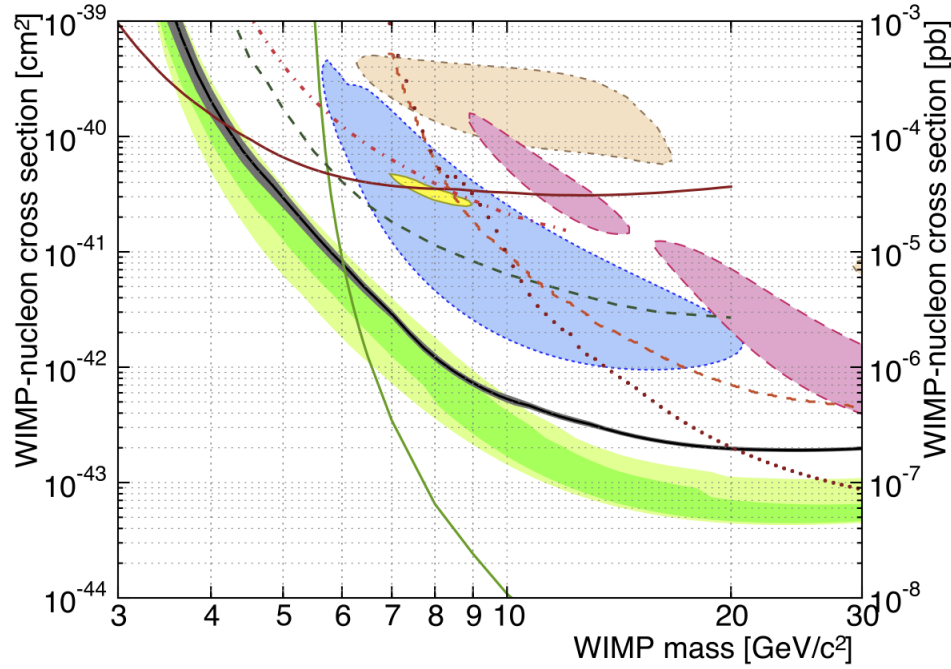


Figure 2.16: SuperCDMS Soudan low threshold analysis WIMP limits. The solid black line shows the 90% confidence upper limit based on all observed events. The expected sensitivity in the absence of a signal is shown by the dark green (68% C.L.) and light green (95% C.L.) bands. Closed contours are the preferred parameter space from the results of CDMS II Si (blue), CoGeNT (yellow), CRESST-II (pink), and DAMA/LIBRA. Solid lines are exclusion limits from CDMS II Ge (dotted dark red), CDMS II Ge low threshold (dashed-dotted red), CDMSlite (solid dark red), LUX (solid green), XENON10 S2 only (dashed dark green), and EDELWEISS low threshold (dashed orange) analyses.[73]

Fig. 2.16) on the spin-independent WIMP-nucleon cross section of $1.2 \times 10^{-42} \text{ cm}^2$ at $8 \text{ GeV}/c^2$. The result excluded the WIMP interpretation of the CDMS II silicon results, along with several other hints of WIMP detection from other competing experiments. A high threshold zero background analysis is currently underway.

2.5.4 CDMSlite

As stated before, Luke-Neganov amplification can be used to significantly lower the ionization energy threshold at the expense of losing event-by-event background rejection. In 2013, a CDMS iZIP detector was configured with a 69V planar bias and used for a low mass WIMP search with a small exposure of only 6.3kg-days, known as the CDMSlite experiment.

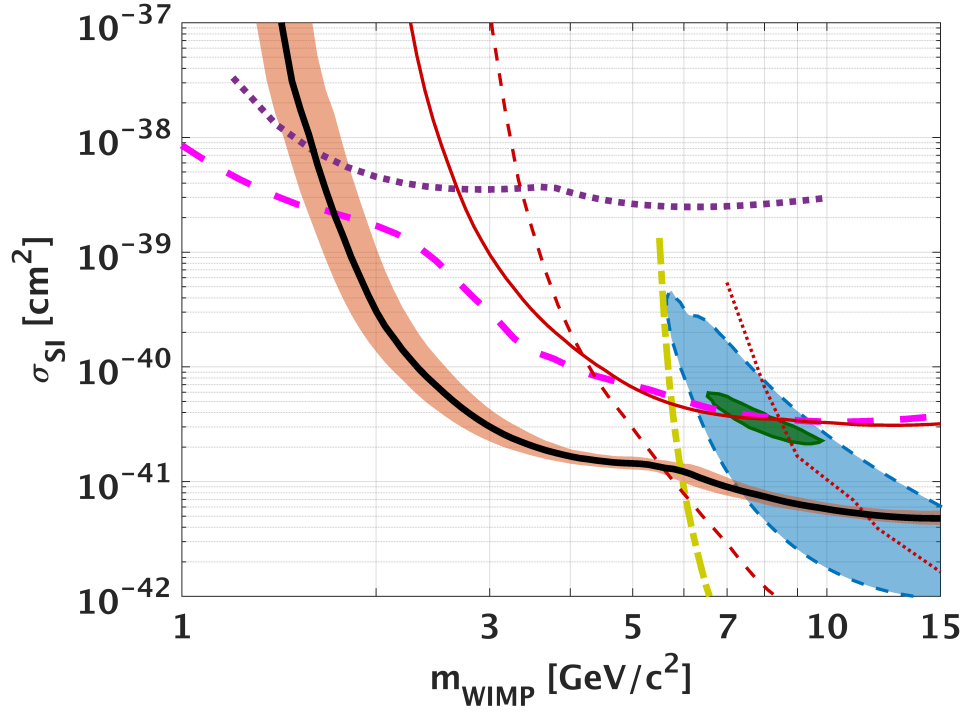


Figure 2.17: CDMSlite WIMP search limits. Median (90% C.L.) and 95% interval from the analysis are shown as the black curve surrounded by the salmon-shaded band. Other 90% C.L. exclusion limits shown are from CDMSlite Run 1 (thin, solid red line), SuperCDMS low threshold (thin, dashed red line), EDELWEISS-II (thin, red dotted line), LUX (thick, dark-yellow dot-dashed line), CRESST (thick, magenta dashed line), and DAMIC (thick, purple dotted line). Closed contours are from CDMS II Si (blue) and CoGeNT (dark green).[133]

The achieved ionization energy resolution of 170 eVee allowed for new exclusion limits below WIMP masses of 6 GeV/ c^2 and demonstrated the use of Luke amplification for dark matter detection.[134] A second CDMSlite run with improved readout electronics and a total exposure of 70 kg-days was performed in 2014 and set new limits between 1.6 and 5.5 GeV/ c^2 , shown in Fig. 2.17.[133]

2.6 The Future - SuperCDMS SNOLAB

In order to probe even smaller cross sections, it is necessary for CDMS to move to a deeper location than Soudan to limit the neutron background. The SNOLAB facility in Sudbury, Canada — famous for the Nobel Prize winning SNO neutrino detection experiment[135] — is about 1.5 miles below the surface and provides 6010 m.w.e of overburden. The next iteration of CDMS, the SuperCDMS SNOLAB experiment, will be located in the SNOLAB

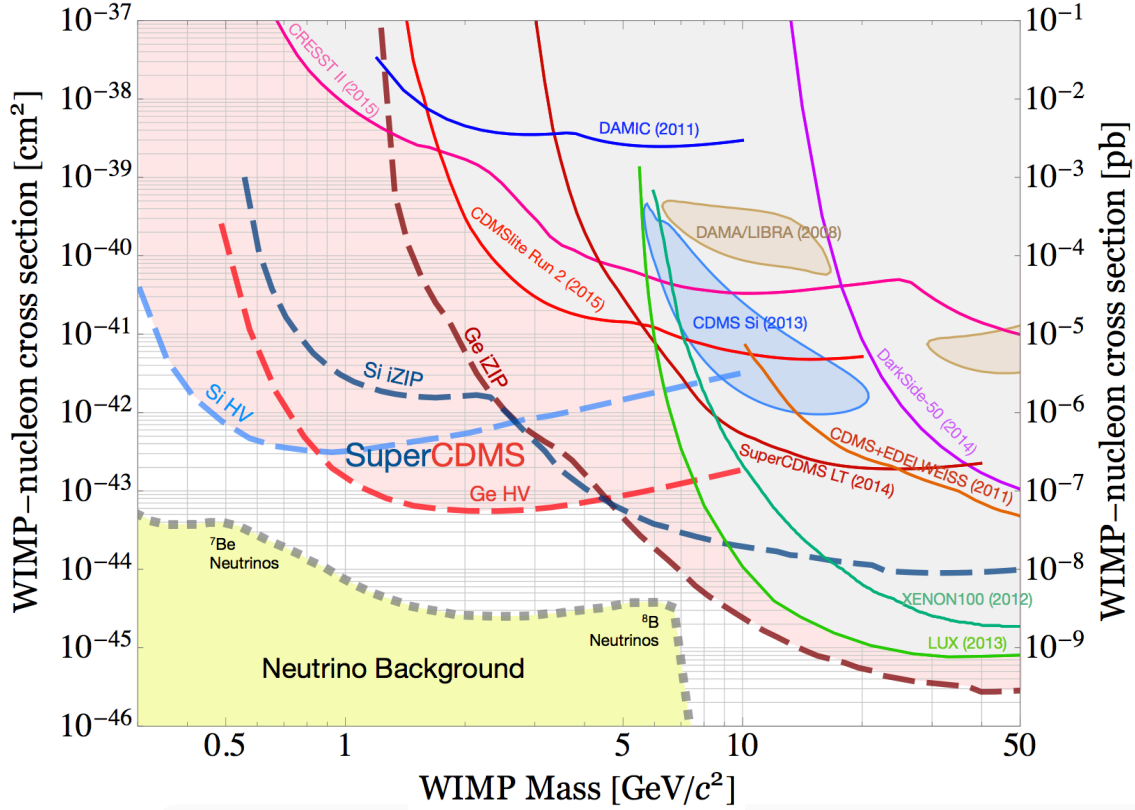


Figure 2.18: Proposed SuperCDMS SNOLAB limits. The light red shaded regions indicate new parameter space which will be explored by the project.

facility and is presently under design and construction. Due to the success of liquid noble experiments such as LUX at high mass, SuperCDMS SNOLAB will focus on probing the low-mass regime of WIMP masses below $10 \text{ GeV}/c^2$, as shown in Fig. 2.18.

SuperCDMS SNOLAB detectors will still follow the basic iZIP design, however the dimensions of the detector have been increased to a diameter of 100 mm and a thickness of 33.3 mm. Each Ge detector has a mass of 1.39 kg, and each Si detector has a mass of 0.61 kg. The number of phonon sensors is increased to 12, with six per side. Readout of the ionization signal will be performed with a HEMT-based front end, discussed in Chapter 4. Six detectors will be integrated into a detector tower. The SuperCDMS SNOLAB baseline payload consists of three towers of Ge iZIP detectors, one tower of Si iZIP detectors, and one tower of high voltage (HV) detectors (four Ge and two Si) optimized for the CDMSlite mode. The intent is for 5 years of running with 80% livetime, giving raw exposures of 22 kg-yr (Ge HV), 5 kg-yr (Si HV), 100 kg-yr (Ge iZIP), and 15 kg-yr (Si iZIP).

Chapter 3

Charge Amplification of Cryogenic Semiconductor Detectors

3.1 Introduction

An experimentalist needs to have a strong understanding of the process through which the signal produced by a detector is converted to the form in which it is recorded. For ionization collection in cryogenic semiconductor detectors such as those used by CDMS, this conversion is performed using charge amplifiers. In this chapter, we discuss the formation and amplification of the *electrical* ionization signal produced by CDMS detectors. This understanding is important as the charge amplifier noise performance is directly related to the ionization energy resolution of the experiment.

3.2 Shockley-Ramo Theorem

Events produce free charge carriers which drift towards their respective charge-collecting electrodes. What is the electric current induced in the electrodes due to these moving charges? The answer to this question takes the form of the Shockley-Ramo theorem (also known as simply Ramo's theorem) which first appeared in a 1938 paper by Shockley[136], however a more general (and more elegant) derivation appeared a year later in a paper by Ramo.[117] A common misconception is that the electrode current only appears once the free carriers have reached the electrodes. The Shockley-Ramo theorem shows that the induced current is in fact instantaneous¹, which we will now derive.

3.2.1 Qualitative Understanding

First, consider a free charge q in the presence of an infinitely large electrode. All of the electric field lines emanating from the charge terminate on the electrode. Gauss's law states

¹Up to speed-of-light delays, obviously!

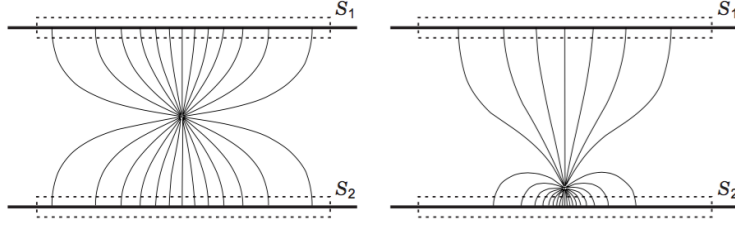


Figure 3.1: The position dependence in electric field line termination for a charge carrier placed between two planar electrodes.[137]

that the electric flux passing through a surface is proportional to the charge enclosed by that surface. If we surround our charge q by such a surface, we have

$$\oint_S \vec{E} \cdot d\vec{a} = \frac{q}{4\pi\epsilon\epsilon_0}. \quad (3.1)$$

We can find the induced charge on the electrode by simply switching the sign of the differential element $d\vec{a}$, equivalent to surrounding the electrode by a surface which does not enclose our free charge. The induced charge on the electrode is clearly $-q$.

Now consider the case of a free charge in the presence of two infinitely large electrodes, as shown in Fig. 3.1. When the free charge is equidistant from both electrodes, half of the field lines terminate on the top electrode and half of the field lines terminate on the bottom. Each electrode has an induced charge of $-q/2$. If the free charge is closer to one of the electrodes, it will terminate more of the field lines emanating from the free charge, resulting in a greater amount of induced charge on that electrode. As less field lines terminate on the opposite electrode, and since the total induced charge must be equal to $-q$, a smaller amount of induced charge is present on the electrode further away. Thus we see that if a free charge begins near one electrode and moves towards the other, the induced charge flows from one electrode to the other. This current can be measured by connecting the input of an amplifier across the two electrodes, as shown in Fig. 3.2.

3.2.2 Quantitative Calculation

Now consider any number of electrodes in any geometrical configuration. Ground all the electrodes so that the electric potential on their surface is 0. As before, place a charge q in the vicinity of the electrodes and surround it by a small equipotential sphere such that Eq. 3.1 holds true. Call the electric potential of the sphere V_s and the potential between the sphere and the electrodes V . As there are no free charges in this region, we have

$$\nabla^2 V = 0. \quad (3.2)$$

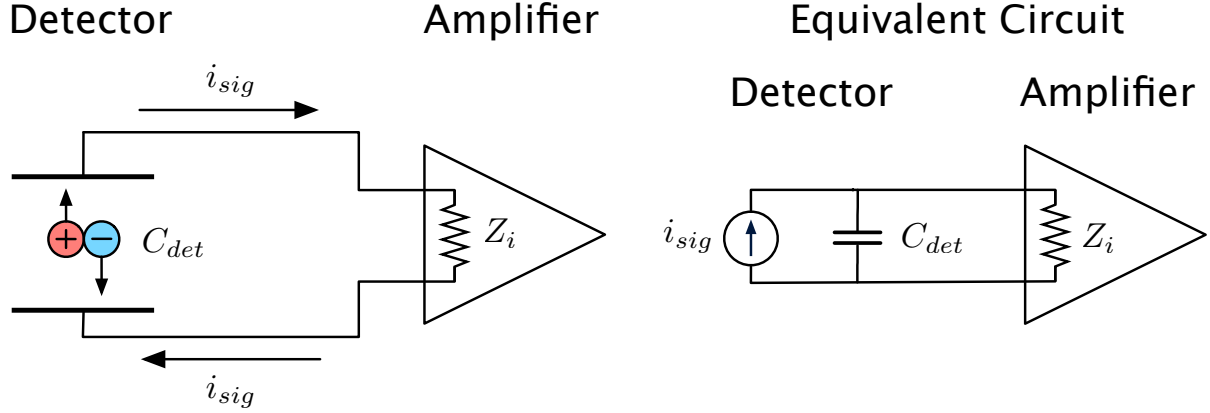


Figure 3.2: The flow of induced charge between electrodes can be sensed by an amplifier connected across the electrodes. The equivalent circuit model is a current source in parallel to the capacitance between the electrodes.

Remove the charge and raise the potential of one of the electrodes, which we label electrode A , to a potential of 1 Volt. Label the potential of this configuration U , so again we have

$$\nabla^2 U = 0 \quad (3.3)$$

in the region between the electrodes, including the point where the former charge q was located. Let U_q be the new potential on the sphere which surrounded this point.

We can now apply Green's second identity to the volume between the sphere and the electrodes, labeled W , using these potential definitions. We denote the boundaries of this volume as ∂W . We have

$$\int_W (U \nabla^2 V - V \nabla^2 U) d^3x = - \oint_{\partial W} \left(U \frac{\partial V}{\partial n} - V \frac{\partial U}{\partial n} \right) ds, \quad (3.4)$$

where n is the direction normal to the boundary surface and ds is the differential surface element. The left hand side is clearly 0. The right hand side can be split into three integrals:

1. Over the surfaces of all electrodes other than A . On the surface of these electrodes, $V = U = 0$ so the surface integral is 0.
2. Over the surface of electrode A . On the surface of A , $V = 0$ and $U = 1$ so we have

$$- \oint_A \frac{\partial V}{\partial n} ds. \quad (3.5)$$

3. Over the surface of the sphere. We have

$$-U_q \oint_{\text{sphere}} \frac{\partial V}{\partial n} ds + V_s \oint_{\text{sphere}} \frac{\partial U}{\partial n} ds. \quad (3.6)$$

Combining these three contributions, we have

$$\oint_A \frac{\partial V}{\partial n} ds = -U_q \oint_{\text{sphere}} \frac{\partial V}{\partial n} ds + V_s \oint_{\text{sphere}} \frac{\partial U}{\partial n} ds. \quad (3.7)$$

Note that all of these surface integrals are equivalent to Gauss's law ($\vec{E} \cdot d\vec{a} = -(\partial V/\partial n)ds$) and are therefore proportional to the charge enclosed by the surface. The configuration with electric potential U contains no free charge, so the second term on the right hand side of Eq. 3.7 is 0 and we are left with

$$q_A = U_q q \quad (3.8)$$

after canceling the common term $4\pi\epsilon\epsilon_0$ from in the denominators from application of Gauss's law. q_A is the induced charge on electrode A. If we take the limit as the volume of the sphere approaches 0, U_q represents the value of the potential U at the point where our free charge was situated.

We now let our charge q move with velocity $\vec{v}$. The induced current on electrode A is

$$i_A = \frac{dq_a}{dt} = q \frac{dU_q}{dt} = q (\vec{v} \cdot \nabla U_q) \quad (3.9)$$

as

$$\frac{dU_q}{dt} = \frac{\partial x}{\partial t} \frac{\partial U_q}{\partial x} + \frac{\partial y}{\partial t} \frac{\partial U_q}{\partial y} + \frac{\partial z}{\partial t} \frac{\partial U_q}{\partial z} = v_x \frac{\partial U_q}{\partial x} + v_y \frac{\partial U_q}{\partial y} + v_z \frac{\partial U_q}{\partial z} = \vec{v} \cdot \nabla U_q \quad (3.10)$$

where we identify $\partial x/\partial t$ as the x-component of the charge's velocity, and so forth. The induced current is equal to the scalar product of the drift current $q\vec{v}$ and the gradient of the potential U at the location of the moving charge.

We therefore redefine U as the *weighting potential* Φ , also known as the Ramo potential, and derive a weighting field $\vec{E}_W \equiv -\nabla\Phi$. Eq. 3.9 becomes

$$i_A = q\vec{v} \cdot \nabla\Phi = -q\vec{v} \cdot \vec{E}_W, \quad (3.11)$$

where the weighting potential and field for a particular electrode in any geometrical configuration are found by raising that electrode's potential to unity while grounding all others. A charge moving from position $\vec{r}_1$ at time t_1 to position $\vec{r}_2$ at time t_2 will therefore induce a net charge

$$\Delta q_A = \int_{t_1}^{t_2} i_A dt = \frac{1}{|\vec{v}|} \int_{\vec{r}_1}^{\vec{r}_2} i_A dr = \frac{q}{|\vec{v}|} \int_{\vec{r}_1}^{\vec{r}_2} (\vec{v} \cdot \nabla\Phi) dr = q (\Phi(\vec{r}_2) - \Phi(\vec{r}_1)). \quad (3.12)$$

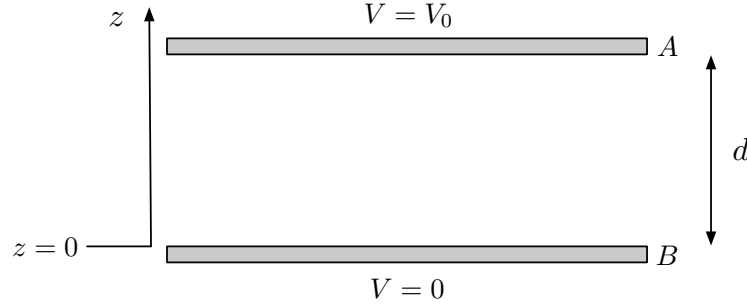


Figure 3.3: Setup for calculating the weighting potentials for two infinitely long, planer electrodes separated by a distance d .

The induced charge is independent of the path taken — it only depends on the initial and final position.

It is important to note that **the electric field and the weighting field are distinctly different**. While the weighting field determines the induced charge on a given electrode, it is the electric field which dictates the motion of the charge. In general, the two fields are different and care must be taken to understand the time-dependence of the induced charge due to a mobile carrier. Only in the case of a two-electrode system are the electric field and weighting field proportional.

3.2.3 Planar Geometry

The case of two infinitely large, parallel electrodes separated by a distance d is simple to calculate. As shown in Fig. 3.3, let the z -axis be perpendicular to both electrodes, the electrode at $z = d$ be held at potential V_0 , and the electrode at $z = 0$ be held at electrical ground. The electric and Ramo potentials are

$$\begin{aligned} V &= V_0 z/d, \\ \Phi_A &= z/d, \\ \Phi_B &= 1 - z/d. \end{aligned} \tag{3.13}$$

The electric field and Ramo fields are

$$\begin{aligned} \vec{E} &= -V_0 \hat{\mathbf{z}}/d, \\ \vec{E}_{WA} &= -\hat{\mathbf{z}}/d, \\ \vec{E}_{WB} &= \hat{\mathbf{z}}/d. \end{aligned} \tag{3.14}$$

Note the symmetry of the two Ramo fields and their direct proportionality to the electric field. The Ramo fields are constant, implying that the induced current in each electrode perfectly mirrors the drift current within the detector.

An event occurring at position $z = z_0$ will produce some number of electron-hole pairs N_{eh} . If V_0 is positive, the holes will drift towards electrode B and the electrons will drift towards electrode A . Although the electrons and holes are drifting in opposite directions, their charges are also opposite. Their respective drift currents are the same sign and add together. By Eq. 3.12 the total induced charge on electrode A is

$$\begin{aligned} q_A &= N_{eh}q (\Phi_A(0) - \Phi_A(z_0)) - N_{eh}q (\Phi_A(d) - \Phi_A(z_0)) \\ &= N_{eh}q (0 - z_0/d) - N_{eh}q (1 - z_0/d) \\ &= -N_{eh}q \end{aligned} \tag{3.15}$$

and is independent of the position of the event. Similarly, the induced charge on electrode B is $N_{eh}q$ and the sum of the two electrodes is zero, as expected. An amplifier whose input is connected between the electrodes thus produces a pulse whose amplitude is proportional to the number of electron-hole pairs produced by the event.

3.2.4 oZIP and mZIP Geometry

The CDMS oZIP and mZIP detectors have Ramo potentials which are well-described by the planar geometry result at small radius. Near the edges of the detector, however, the Ramo potential deviates from this geometry and results in two populations of events described as the “funnel”, which have a reduced total charge signal, and the “ear”, which have a negative outer charge channel signal.[138] These events are effectively removed from the WIMP search analysis by ensuring the signal in the outer charge electrode be consistent with noise.

3.2.5 iZIP Geometry

Due to the interleaved nature of the iZIP geometry, the Ramo potential map is complicated. The presence of phonon sensors on the same face as the ionization collection electrode causes the Ramo potential to only strongly vary nearby. As opposed to the mZIP and oZIP geometries, in which the charge amplifiers respond uniformly throughout the crystal, the iZIP geometry causes charge to only be measured once it is near the collecting surface. Thus, for a symmetric bias, one side of the detector measures the holes while the other measures the electrons. A simple model for the iZIP is to assume only the charge which reaches the electrode is collected.

3.3 The Basic Charge Amplifier

Dark matter detectors are generally interested in events producing recoil energies of a few tens of keV. The total amount of produced charge is quite small — a 10 keV event only generates about 3,000 electron-hole pairs. It is therefore necessary to amplify the signal. Since the recoil energy is proportional to the number of produced electron-hole pairs, the

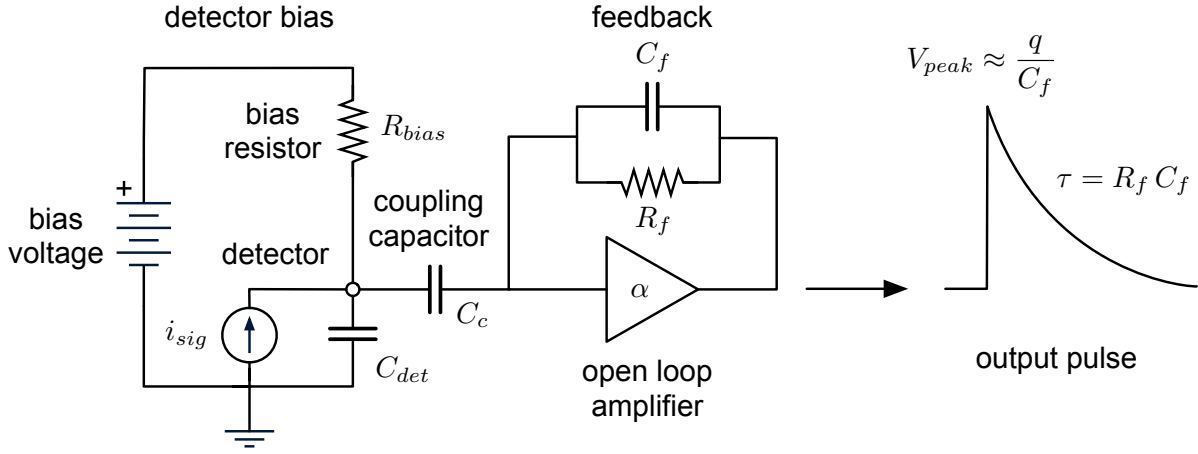


Figure 3.4: The basic charge amplifier topology.

use of *charge amplifiers* which integrate the Ramo current are commonplace. These produce an output pulse whose amplitude is proportional to the total number of electron-hole pairs, with a falltime determined by the components of the amplifier. We will now develop a basic small-signal model of a dark matter detector coupled to a charge amplifier. This model is valid for any generic semiconductor radiation detector and the basic topology is shown in Fig. 3.4.

3.3.1 Input Signal

First we need to develop a suitable electronic model of the detector response. As the crystal is cold enough to be non-conducting, the resistance of the detector is assumed to be infinite. The geometrical configuration of the detector electrodes forms a capacitance network between themselves and the electrical 0V reference (ground) of the system. While more complicated models involving the inter-electrode capacitance can be developed to simulate effects such as crosstalk, for our purposes we will focus on a single electrode which we represent by its capacitance to ground C_{det} . The Ramo current induced by moving charges within the detector represents a current source in parallel to this capacitance. As discussed in the previous section, the shape of this current pulse depends on the trajectory of the charges (as determined by the electric field) and the Ramo field for that particular electrode.

Note that the typical detector electrode capacitance is on the order of 100 pF, and is biased to ~ 1 V. The ~ 5 μ V change in electrode potential due to a 10 keV event has essentially no effect on the electric field and we can take it to be constant over the course of the pulse. Drift times for electrons and holes at ~ 1 V/cm fields are roughly 1 μ s for a 1 inch thick detector. If we take this to be the 10% – 90% risetime, an amplifier bandwidth of 2

MHz would be required to track the drift current. Due to the lengthy runs of cabling from the detector to the readout electronics, charge amplifier bandwidths are usually limited to several hundred kHz. Thus we see that these charge amplifiers do not respond fast enough to follow the drifting of the charge carriers and the shape of the signal can be represented as a delta-function excitation of current whose amplitude is equal to the total change in charge of the electrode due to the pulse. We have

$$i_{sig}(t) = q\delta(t). \quad (3.16)$$

3.3.2 Bias Resistor

The electrode bias voltage is provided by a room-temperature power supply. It is important to place a bias resistor R_b between the supply and the electrode so that the Ramo current is not immediately shorted out by the supply. In order to limit the Johnson noise, discussed later in this chapter, this resistor should have as large a value as is feasible, typically several tens of M Ω s, and be located at the lowest temperature stage of the cryostat, generally the temperature of the detectors/mixing chamber. The limitations on the value of the bias resistor come from instabilities in large-valued resistances at low temperature and the need to limit the electrode biasing time constant, determined by the product of the bias resistor and total capacitance present on the input, to a maximum of several seconds. The resistor should be placed in close proximity to the detector to limit stray capacitance due to cabling from appearing in parallel to the detector capacitance, which will both increase the detector biasing time constant and degrade the noise performance.

In the small signal limit, the Ramo current now sees the detector capacitance in parallel to the bias resistor (voltage sources appear as AC grounds). This forms an integrator with a time constant equal to $R_b C_{det}$. If this time constant is much greater than the drift time of the charge carriers, the voltage at the electrode will peak with an amplitude $\Delta V = Q/C_{det}$ before decaying exponentially. The voltage amplitude is inversely proportional to the detector capacitance, which should be as small as possible in order to maximize sensitivity to charge. This is the reason why point contact detectors, with an electrode capacitance ~ 1 pF, can have much better energy resolution than CDMS detectors. CDMS does not use point contacts because they require a high electric field for good charge collection which would produce too many Luke phonons. Any additional stray capacitance present on the input should be added in parallel to C_{det} .

Noting that the Fourier transform of $i_{sig}(t)$ is simply q for all frequencies, we have

$$V(f) = q \left(\frac{R_b}{1 + j2\pi f R_b C_{det}} \right) \quad (3.17)$$

in the frequency domain.

3.3.3 Open Loop Amplifier

By connecting the electrode to a high input impedance voltage amplifier, represented by the triangle in Fig. 3.4, this signal can be amplified to a level suitable for digitization. Note that this input impedance also appears in parallel to the detector so a high input impedance is required so that the signal is not shorted out and a measurable input voltage is produced.

While modern operational amplifiers can have a very high input impedance, they only operate at room temperature. The capacitance of the cabling required to bring cryogenic signals to room temperature, which would also appear in parallel to the input and degrade charge sensitivity, prohibits their use as a first-stage amplifier. The traditional solution has been to use a single silicon field effect transistor (FET) placed within the cryostat, heated to an operating temperature of ~ 150 K, as a first-stage buffer. The FET provides the necessary high input impedance with a capacitance that can be adapted to the electrode to maximize noise performance, discussed later. While the input resistance of the FET can generally be ignored, the input capacitance C_{FET} must be taken into consideration — in Eq. 3.17 the total input capacitance $C_{total} = C_{det} + C_{stray} + C_{FET}$ should be used.

The FET is connected to a room temperature amplification stage to provide gain without the cable capacitance appearing directly on the input. The combination of the cryogenic FET and room temperature amplification stage form the *open loop amplifier*, named as such due to the lack of feedback, which amplifies the electrode voltage with a frequency-dependent gain $\alpha(f)$. At the output of the open loop voltage amplifier, in the absence of feedback, we have

$$V_{out} = \alpha(f) \cdot q \left(\frac{R_b}{1 + j2\pi f R_b C_{total}} \right). \quad (3.18)$$

3.3.4 Coupling Capacitor

In order to isolate the DC detector bias from the open loop amplifier input, which typically operates at a much different DC level, a *coupling capacitor* is used. The capacitor should be much larger in value than the detector and open loop amplifier input capacitance to minimize the loss of charge to the input of the open loop amplifier.

It is easiest to understand the effects of the coupling capacitor using complex impedances. Let the impedance of the coupling capacitor be Z_c , the impedance on the detector side be Z_{det} , and the impedance on the amplifier side be Z_{in} . A current divider is formed between Z_{det} and $Z_c + Z_{in}$ and the voltage at the amplifier input is reduced to

$$V_{in} = \left(\frac{Z_{det}}{Z_{det} + Z_c + Z_{in}} \right) Z_{in}. \quad (3.19)$$

The coupling efficiency assuming assuming 100 pF of detector capacitance, 100 pF of amplifier input capacitance, and an infinitely large bias resistor as a function of the coupling capacitance is shown in Fig. 3.5. Note that CDMS II and SuperCDMS Soudan used a 300 pF

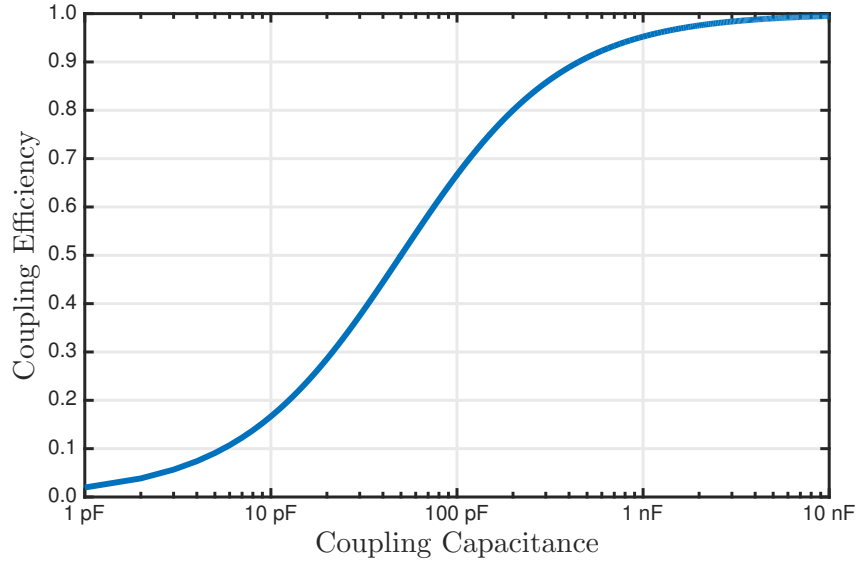


Figure 3.5: Amplifier coupling efficiency as a function of the coupling capacitance. The calculation is made assuming 100 pF of detector capacitance, 100 pF of amplifier input capacitance, and an infinitely large bias resistor.

coupling capacitor and $\sim 15\%$ of the charge signal was lost. SuperCDMS SNOLAB plans to use 10 nF coupling capacitors which have near-perfect coupling.

3.3.5 Feedback

Searching for WIMPs is a time-consuming process and datasets can span over the course of a year or longer. The overall gain of the charge amplifier system must remain stable throughout this time period. While open loop amplifiers provide huge voltage gain, often exceeding one million, they are very sensitive to variations in the individual components which make up the circuit. Even changes in operating temperature can cause the gain to drift by several percent. By introducing *feedback*, in which a portion of the amplified signal is returned to the input, the total gain can be made relatively insensitive to variations in the open loop gain.

Consider the signal flow diagram shown in Fig. 3.6. We first work in an abstract space where the quantities can represent either voltages or currents. A source produces a signal x_s which is connected to the closed loop amplifier. By definition, the output x_o is related to the open loop amplifier input x_i by

$$x_o = A \cdot x_i. \quad (3.20)$$

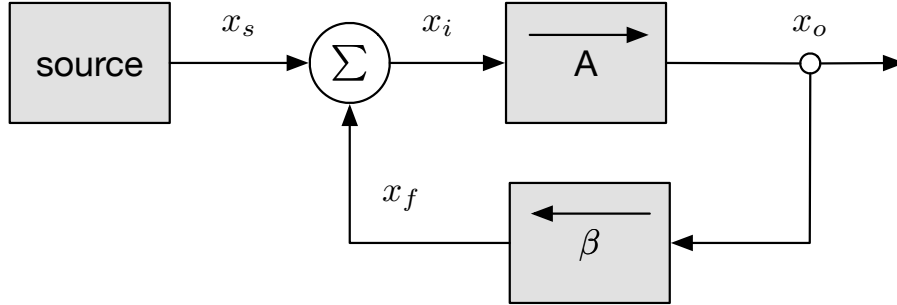


Figure 3.6: Closed loop signal flow diagram.

A portion of that output is fed back to the input through the feedback factor β by

$$x_f = \beta \cdot x_o. \quad (3.21)$$

This feedback signal is summed with the source signal to produce the amplifier input

$$x_i = x_s + x_f. \quad (3.22)$$

These three equations can be combined to solve for the closed loop gain, given by

$$\frac{x_o}{x_s} = \frac{A}{1 - A\beta}. \quad (3.23)$$

If the open loop amplifier is inverting with magnitude α , we have

$$A = -\alpha \quad (3.24)$$

and

$$\frac{x_o}{x_s} = -\frac{\alpha}{1 + \alpha\beta}. \quad (3.25)$$

We see that the closed loop amplifier is also inverting, however the gain has been reduced by $1 + \alpha\beta$. In the limit that $\alpha\beta \gg 1$, the closed loop gain reduces to

$$\frac{x_o}{x_s} \approx -\frac{1}{\beta}. \quad (3.26)$$

This shows the advantage of using feedback — the closed loop gain is determined only by feedback components, which can be chosen to have high precision and stability, with little sensitivity to variations in the open loop gain.

We now need to take this formalism and apply it to our basic charge amplifier. It is easiest to consider the open loop amplifier and feedback factors to be voltage-to-voltage converters.

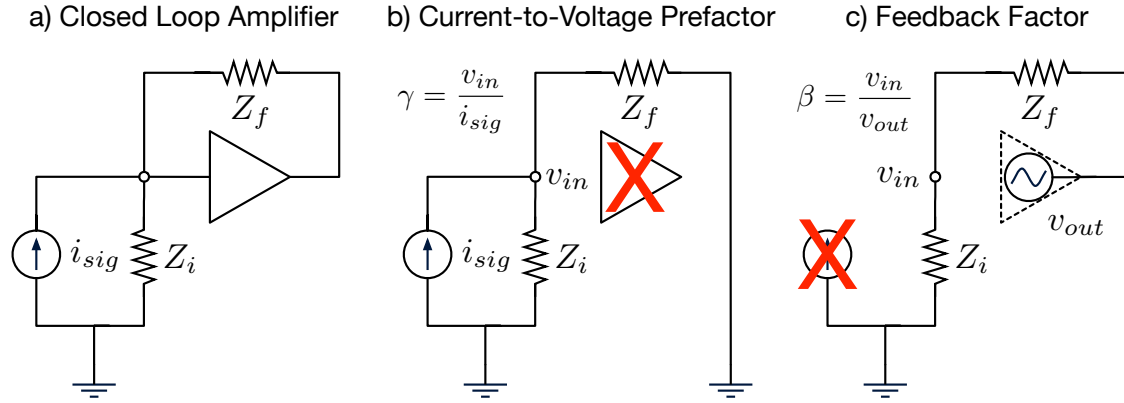


Figure 3.7: a) The closed loop amplifier. b) Equivalent circuit for calculating the current-to-voltage prefactor. c) Equivalent circuit for calculating the feedback factor.

Consider the closed loop configuration shown in Fig. 3.7a. As our input signal model is a current source, we must first define the *current-to-voltage prefactor* γ . This factor determines the relationship between the source voltage x_s and the signal current. This is equivalent to the voltage at the input of the amplifier with the output shorted, which effectively removes the amplifier from the circuit as shown in Fig. 3.7b. The feedback impedance appears in parallel to the input impedance and we have

$$\gamma = \left. \frac{v_{in}}{i_{sig}} \right|_{v_{out}=0} = \frac{Z_i Z_f}{Z_i + Z_f}. \quad (3.27)$$

The feedback factor β is relationship between the voltage which appears at the input due to a voltage at the output, with the input signal shorted. This configuration is shown in Fig. 3.7c, and we see a voltage divider is formed between the input impedance and the feedback impedance. We therefore have

$$\beta = \left. \frac{v_{in}}{v_{out}} \right|_{i_{sig}=0} = \frac{Z_i}{Z_i + Z_f}. \quad (3.28)$$

Assuming we are in the limit that $\alpha\beta \gg 1$, we find the combined closed loop gain of the charge amplifier to be

$$G_{cl} = \frac{v_{out}}{i_{sig}} = -\frac{\gamma}{\beta} = -Z_f, \quad (3.29)$$

as would be expected from a basic operational amplifier analysis of the circuit. Based on this result, it may appear that the input impedance plays no role in determining the output of a charge amplifier, however it is important to note that Eq. 3.29 is the closed loop gain

for a *current signal*. For a *voltage signal* the closed loop gain is

$$H_v = \frac{v_{out}}{v_{in}} = -\frac{1}{\beta} = -\frac{Z_i + Z_f}{Z_i}, \quad (3.30)$$

which does depend on the input impedance. If we now consider Z_i to be the detector capacitance and Z_f to be a feedback capacitance, we get

$$|H_v| = \frac{1/j2\pi f C_{det} + 1/2\pi f C_f}{1/j2\pi f C_{det}} = \frac{C_{det} + C_f}{C_f} \approx \frac{C_{det}}{C_f} \quad (3.31)$$

for $C_d \gg C_f$. If C_d is too much larger than C_f , the closed loop voltage gain approaches the open loop gain and the approximation $\alpha\beta \gg 1$ by which we derived our results no longer holds. The large detector capacitance would make the feedback ineffective, taking $\alpha\beta \rightarrow 0$, and the output voltage once again becomes Eq. 3.18. The full advantages of feedback are only present if the closed loop gain is significantly less than the open loop gain, otherwise these second order effects begin to dominate.

3.3.6 Output Signal

We will now discuss the output signal for the most typical feedback configuration of a parallel RC combination. Let the feedback resistance be R_f and the feedback capacitance be C_f . Eq. 3.29 shows that the output pulse is entirely determined by the impedance of this combination. Recalling that we are modeling our input current signal as a simple delta function of magnitude q , we can Fourier transform back from the frequency domain to the time domain and find

$$\begin{aligned} v_{out}(t) &= -\frac{q}{C_f} \int_{-\infty}^{\infty} df \left(\frac{R_f C_f}{1 + j2\pi f R_f C_f} \right) \exp(j2\pi f t) \\ &= -\frac{q}{C_f} \exp\left(-\frac{t}{R_f C_f}\right). \end{aligned} \quad (3.32)$$

The output pulse peaks with an amplitude q/C_f and decays by the $R_f C_f$ time constant, determined by the feedback. One way to interpret this result is to think of large open loop gain of amplifier as forcing all the charge to accrue on the feedback capacitor, resulting in a voltage change q/C_f , which is then discharged by the feedback resistance.

We can treat the case of no feedback resistance by looking at the limit of Eq. 3.32 as $R_f \rightarrow \infty$. We see the output signal shape becomes a step function of amplitude q/C_f and never returns to the baseline level before the pulse. In this configuration, the feedback capacitor must be periodically reset once the DC output level has drifted beyond the operating range of the amplifier.

3.4 Noise Sources

The minimum amount of charge an amplifier can measure is determined by the amount of noise present in the system. The voltage peak at the output of the amplifier can only be measured if it is discernible above the random voltage fluctuations also present at the output. From the basic charge amplifier topology of Fig. 3.4 we can identify a number of potential noise sources which we now describe. All of the sources in question typically have equivalent input equivalent voltage levels on the order of 0.1-10 nV/ $\sqrt{\text{Hz}}$.

3.4.1 Bias Resistor Johnson Noise

All resistive elements exhibit an intrinsic, temperature-dependent white noise known as Johnson noise. While often modeled as a voltage noise in series with the resistance, it is easiest for us to consider the Johnson noise as a parallel current noise source with a single-sided power spectral density of

$$i_{Rb} = \sqrt{\left(\frac{4k_B T_{Rb}}{R_b}\right)}, \quad (3.33)$$

where k_B is Boltzmann's constant, T_{Rb} the temperature of the bias resistor, and the units are A/ $\sqrt{\text{Hz}}$. This current noise source appears in the exact same way as the signal current, and thus is also amplified by Eq. 3.29. It is clear that this noise source is minimized by making R_b as large as possible while keeping T_{Rb} small. Note that use of a coupling capacitor also influences the coupling of this noise, as previously discussed for the signal current.

3.4.2 Feedback Resistor Johnson Noise

The presence of a feedback resistor will add an additional Johnson noise source. Similar to the bias resistor, the feedback resistor noise is best treated as a parallel current source. As expected, the single-sided power spectral density is given by

$$i_{Rf} = \sqrt{\left(\frac{4k_B T_{Rf}}{R_f}\right)}, \quad (3.34)$$

where T_{Rf} is the temperature of the feedback resistor. Once again the noise is reduced by maximizing R_f while minimizing T_{Rf} , and care should be taken to properly include the effects of a coupling capacitor.

3.4.3 Open Loop Amplifier Voltage Noise

The components which make up the open loop amplifier will add noise to the system and must be taken into consideration. One part of this noise is modeled as a voltage noise source in series with the input to the amplifier, known as the *open loop amplifier voltage noise*. If

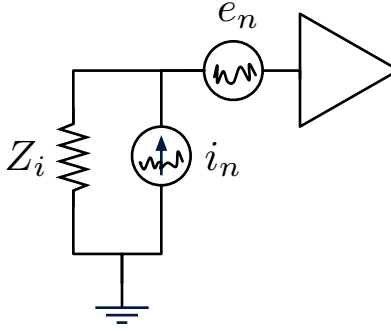


Figure 3.8: Open loop amplifier noise sources. The voltage noise e_n is always present, while the voltage induced by the current noise depend on the total impedance at the input Z_i .

the amplifier is properly designed, this noise will be completely dominated by the component forming the input to the amplifier. This has traditionally been a silicon JFET. Even in the case that other sections of the amplifier contribute meaningfully to the noise, the series voltage source model holds true.

The voltage noise can be found by shorting the input of the open loop amplifier, measuring the output noise, and dividing by the open loop gain. The noise e_n is generally modeled as the uncorrelated combination of a frequency-independent white noise and frequency-dependent flicker noise. We have

$$e_n^2(f) = \left(\frac{A}{f^\alpha} \right)^2 + e_w^2, \quad (3.35)$$

where e_w determines the magnitude of the white noise and A determines the magnitude of the flicker noise. The parameter α determines the frequency dependence of the flicker noise and is usually between 0.1 and 1. The flicker noise tends to dominate at low frequency and is often referred to as $1/f$ noise even when $\alpha \neq 1$. As this noise is a voltage directly at the input of the amplifier, it appears at the output having been multiplied by the closed loop voltage gain H_v .

3.4.4 Open Loop Amplifier Current Noise

The remainder of the amplifier noise is modeled as a current noise source i_n in parallel to the amplifier input, which we call the *open loop amplifier current noise*. The equivalent input voltage noise generated by this source depends on the total input impedance and is given by

$$e_{ni} = \gamma \cdot i_n = \left(\frac{Z_i Z_f}{Z_i + Z_f} \right) i_n, \quad (3.36)$$

making it distinctly different from the amplifier input voltage noise.

The current noise is found by first measuring the input voltage noise. A small capacitance is then connected to the input and the new total input equivalent voltage noise is measured. The impedance of the capacitance, which goes as $1/C$, causes the current noise to induce a large input voltage noise. The intrinsic input voltage noise is then subtracted in quadrature and the result divided by the input impedance to recover i_n . This process can be repeated for several input capacitance values to confirm the dependence on the total input impedance. The intrinsic input voltage noise should be recovered for a large enough input capacitance, which effectively shunts the current noise. The measured data is fit to a simple model of the form

$$i_n^2(f) = a + bf + cf^2, \quad (3.37)$$

which usually produces good results.

If the input impedance is small, the current noise does not contribute. For our charge amplifier, however, the input impedance is at least several tens of $M\Omega$ s in parallel with several hundred pF of capacitance and cannot be ignored. In fact, since this source is a current noise on the input, it responds in the same way as the signal current and appears on the output having been amplified by Eq. 3.29. Any attempts to shunt the current noise will also shunt the signal.

3.4.5 Shot Noise Due to Leakage Currents

Although we usually assume the detector to be perfectly non-conducting, the electric fields due to the potential differences between the electrodes can cause small DC currents to flow. This has been particularly problematic in the iZIP detector design due to the close proximity of the interleaved electrodes to one another. Even with biases of only $\pm 2V$, the surface fields can be as strong as 2 kV/cm. Above a detector-dependent bias voltage which is still not well understood, a DC current begins to flow between the charge and phonon sensors. The phonons emitted by these charge carriers cause an increase in noise and temperature of the phonon sensors preventing normal operation. Other possible sources of leakage current are the readout cabling and bulkhead connectors.

Regardless of the source, these leakage currents produce noise due to the finite nature of electrical charge. Along the conduction path, the charges are assumed to arrive in a Poisson-like manner. The fluctuations in their arrival times leads to *shot noise* and we have

$$i_{shot} = \sqrt{2eI_{leak}}, \quad (3.38)$$

where I_{leak} is the total DC leakage current and e is the elementary charge. This shot noise appears as current noise source in parallel to the detector, and thus the induced input equivalent voltage noise is

$$e_{nshot} = \gamma \cdot i_{shot} = \left(\frac{Z_i Z_f}{Z_i + Z_f} \right) i_{shot}. \quad (3.39)$$

As in the case of the amplifier current noise, the shot noise appears on the output having been multiplied Eq. 3.29.

3.4.6 Microphonics

The physical wires which carry the signal have a geometrical, parasitic capacitance between themselves and the electrical ground of the system. Changes in this capacitance due to vibrations can cause an injection of charge via the term

$$\Delta Q = V \Delta C. \quad (3.40)$$

As can be seen in Fig. 3.9, microphonic noise is assumed to be peak-like as various physical vibrational modes of the system are excited. In general, microphonic-induced noise is difficult to predict and is not modeled when predicting noise contributions. If long enough traces are used and the microphonic peaks not too broad nor numerous, the use of an optimal filter significantly reduces the effects of microphonic peaks.

3.4.7 Other Sources of Noise

While the previous sources of noise represent the most commonly occurring contributions, the list is by no means exhaustive. In particular, high valued resistors at low temperature often exhibit extra noise in addition to their Johnson noise contribution. Many other sources of noise may be present in the system and a comparison between predicted and measured noise spectra is crucial to ensure a full understanding of the situation. Studying changes in output noise for various input and feedback impedance configurations can help to determine the best way to model an unknown noise noise.

3.4.8 Total Input Equivalent Voltage Noise

We are now ready to combine all the various noise contributions. Although the output noise is what is actually measured, it is most useful to perform the combination at the input to obtain the total input equivalent voltage noise. By using the input, we can compare different amplifiers without having to worry about differences in their input-to-output gain. The various noise sources are assumed to be uncorrelated so we add them in quadrature, obtaining

$$\begin{aligned} e_{n,total}^2 &= e_n^2 + \gamma^2 (i_n^2 + i_{Rb}^2 + i_{Rf}^2 + i_{shot}^2) \\ &= \left(\frac{A}{f^\alpha} \right)^2 + e_w^2 + \gamma^2 (i_n^2 + i_{Rb}^2 + i_{Rf}^2 + i_{shot}^2). \end{aligned} \quad (3.41)$$

Multiplying Eq. 3.41 by the closed loop voltage gain H_v (Eq. 3.30) gives the total noise at the output.

3.5 Optimal Filtering and Energy Resolution

With our knowledge of both the signal shape and noise contributions, we are now ready to predict the amplifier's ability to measure charge. As the amount of charge produced by an

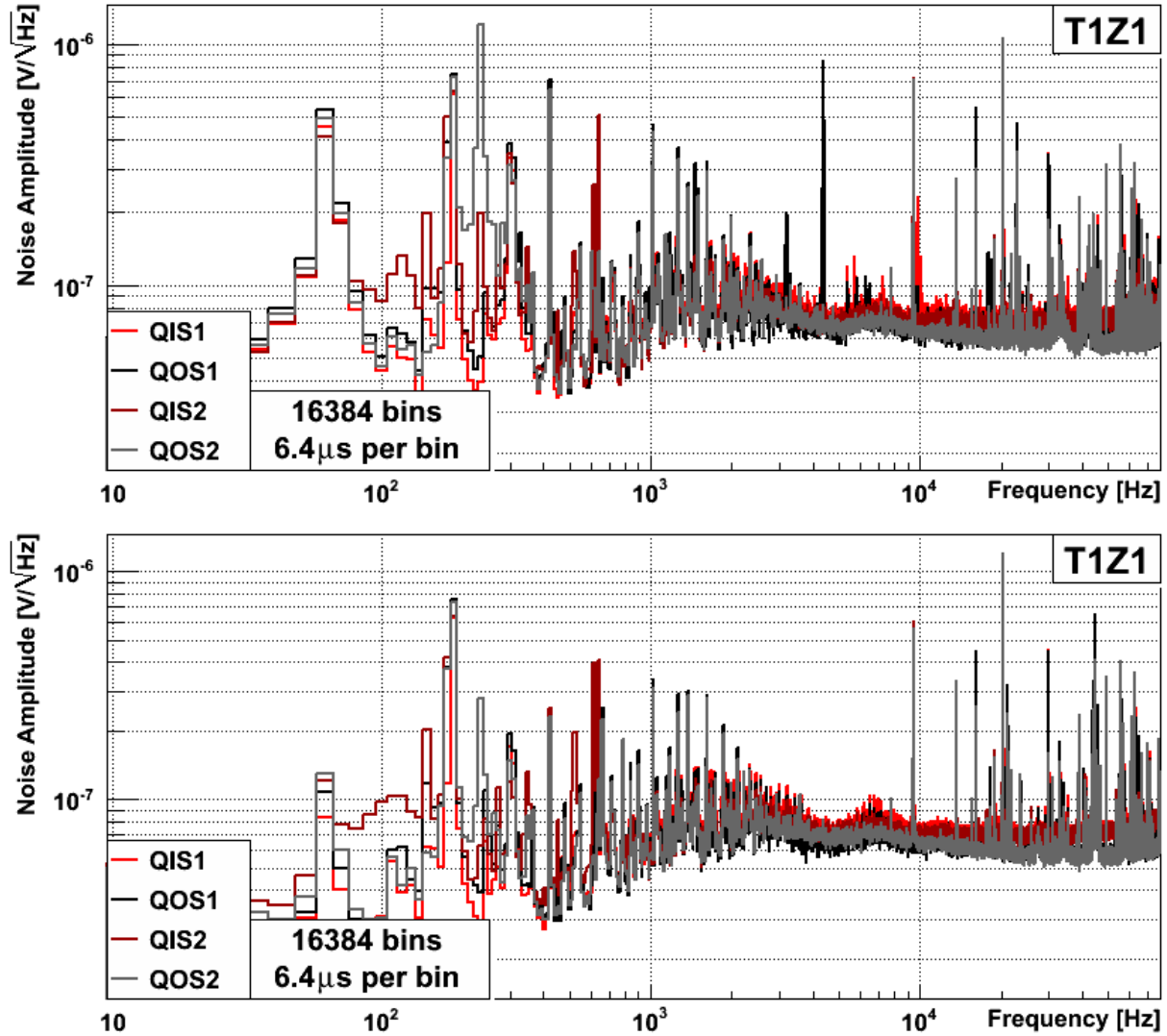


Figure 3.9: Charge noise PSD comparison with vibration-inducing cryocooler on (top) and off (bottom). Although difficult to see, some peaks —assumed to be due to microphonics— are not present or significantly reduced with the cryocooler off

event is proportional to the recoil energy, we usually speak of the *energy resolution* of an amplifier. As a reminder, we assume it takes 3 eV of recoil energy to produce 1 electron-hole pair, which the amplifier sees as 1 Coulomb of charge.

Due to the linearity of the closed loop amplifier, all events will have the same basic pulse shape which is scaled by the number of electron-hole pairs produced. Let $s(t)$ be the pulse formed by creation of a single electron-hole pair, by which all other pulses can be normalized. A measured signal will take the form

$$V_{meas}(t) = As(t) + n(t), \quad (3.42)$$

where A is the true amplitude of the pulse and $n(t)$ the noise. An amplitude estimator $\hat{A}$ is extracted from the noisy pulse through the use of an *optimal filter*, described in detail in Appendix A. The optimal filter performs a frequency-domain noise weighted fit of $s(t)$ to the measured pulse, taking into account the power spectral density of the noise. From Eq. A.11, the R.M.S fluctuation of the amplitude estimator $\sigma_{\hat{A}}$ is

$$\sigma_{\hat{A}}^2 = \left(4 \int_{f_{min}}^{f_{max}} \frac{|s(f)|^2}{J(f)} df \right)^{-1} \rightarrow \left(4 \sum_{f_{min}}^{f_{max}} \frac{|s(f)|^2}{J(f)} \Delta f \right)^{-1}, \quad (3.43)$$

where $s(f)$ is the Fourier transform of $s(t)$, $J(f)$ is the single-sided noise power spectral density (in V^2/Hz), and $f_{max} - f_{min}$ is the measurement bandwidth. In case of discrete sampling, the right hand side is used, where Δf is the spacing between measured frequencies.

The input and output voltages are related by the closed loop voltage gain H_v . We have the following relations:

$$s_{out} = H_v \cdot s_{in}, \quad (3.44)$$

$$J_{out} = e_{n,out}^2 = (H_v \cdot e_{n,total})^2 = H_v^2 \cdot J_{in}, \quad (3.45)$$

$$s_{in} = \gamma \cdot i_{sig}, \quad (3.46)$$

$$J_{in} = e_{n,total}^2. \quad (3.47)$$

Note that $i_{sig} = 1$ in the frequency domain under the assumption of an impulse of a single charge. The ratio between $s(f)$ and $J(f)$ of Eq. 3.43 reduces to

$$\frac{|s_{out}(f)|^2}{J_{out}(f)} = \frac{|s_{in}(f)|^2}{J_{in}(f)} = \frac{|\gamma|^2}{e_{n,total}^2}, \quad (3.48)$$

independent of the voltage gain. We see that the closed loop gain of the amplifier has no effect on the resolution — it only needs to be high enough to overcome the input noise of the output digitization system. Given that this is true, the resolution is determined solely by the total input equivalent voltage noise and the current-to-voltage prefactor γ , which we recall from Eq. 3.27 is the parallel combination of the input and feedback impedance. Thus, for a given amplifier topology, we can calculate $\sigma_{\hat{A}}$ by Eq. 3.43 using Eq. 3.48 for the integrand.

The final step is to convert $\sigma_{\hat{A}}$, which is in units of Coulombs, to units of recoil energy. A subtlety arises because this conversion factor depends on the type of recoil event — nuclear recoils produce less charge per unit energy than electron recoils. The accepted standard is to quote energy resolutions in terms of electron recoils, which are assumed to require 3 eV of recoil energy per produced electron-hole pair. The energy resolution is

$$\sigma_E = \frac{3 \cdot \sigma_{\hat{A}}}{e} = \frac{3}{e} \left(4 \int_{f_{min}}^{f_{max}} \frac{|\gamma|^2}{e_{n,total}^2} df \right)^{-1/2} \quad (\text{in eVee}), \quad (3.49)$$

where “ee” stands for electron equivalent, indicating we are speaking of electron recoils.

When working with data from an actual detector, what we measure is the voltage at the output of the amplifier. It is difficult to know the absolute signal gain because of individual variations in the amplifier components. Instead, an empirical relationship between the output amplitude and recoil energy is made by exposing the detector to a radioactive source with a known emission spectrum and measuring the output voltage amplitude of the resulting photopeaks. CDMS detectors, for example, are calibrated to the 356 keV line from ^{133}Ba . A large sample of these pulses of known energy can be used to generate $s(t)$. $J(f)$ can be found by averaging the individual PSDs of a large sample of random traces, with care taken to ensure the sample is not contaminated by real pulses. The achieved resolution is given by the width of the amplitude estimator distribution formed by applying the optimal filter to a separate large sample of random triggers.

One complication of this calibration scheme is that the charge collection efficiency is assumed to be 100%. While the photopeaks represent recoil events of a known energy, charge trapping and impact ionization influence the number of electron-hole pairs measured. This changes the relationship between the number of produced electron-hole pairs and the recoil energy from 3 eV per pair and enters into the calculation of Luke phonon production. These effects are discussed in later chapters.

3.6 CDMS JFET Amplifier

3.6.1 Open Loop Voltage Amplifier

Ionization readout for the CDMS II and SuperCDMS Soudan experiment was performed using a JFET-based charge amplifier. A basic small signal model of the open loop amplifier is shown in Fig. 3.10. The input is the high impedance gate of JFET Q_1 , which is physically located inside the cryostat. The JFET is placed on a Kapton membrane heatsunk to the T=4K stage and heated to an operating temperature of $\sim 150\text{K}$. The JFET converts the input voltage v_{in} to a current through its transconductance g_m . The InterFET IF4501 JFETs used by CDMS typically have $g_m \approx 25\text{mS}$.

This signal current passes through the 4K-to-300K stripline, after which it sees the parallel combination of the JFET drain biasing resistor R_D and the impedance looking into the

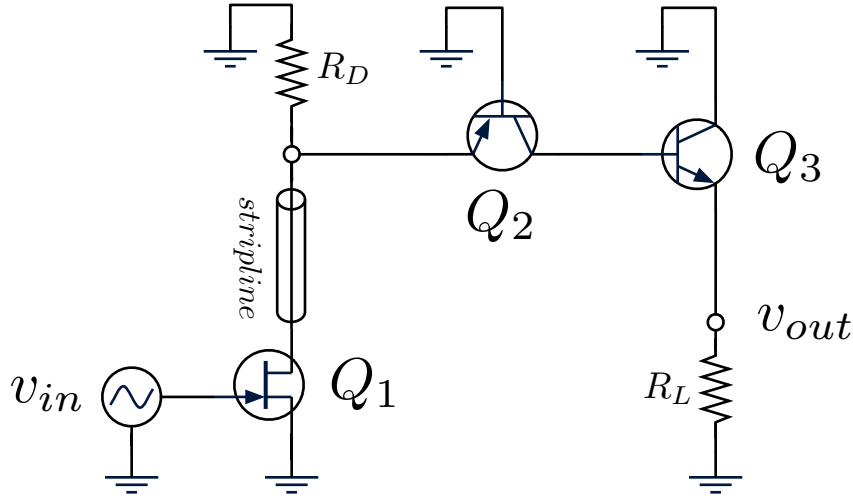


Figure 3.10: The open loop CDMS JFET amplifier in the small signal limit. The DC biasing components are not shown.

emitter of p-type BJT Q_2 . As Q_2 is configured as a common base amplifier, its input impedance is much smaller than R_D ($\sim 40\Omega$ vs $3.24k\Omega$). Essentially all the signal current flows into the emitter of Q_2 and R_D can be ignored. This combination of Q_1 and Q_2 forms what is known as a *cascode* — Q_2 serves as a buffer and allows the signal current, which now flows out of the Q_2 collector, to be further amplified without appreciably changing the drain voltage of Q_1 . This suppresses the Miller effect, which would degrade the signal amplification.[139]

The collector of Q_2 is connected to the base of an n-type BJT Q_3 , configured as an emitter-follower. All of the signal current from the collector of Q_2 flows through the base of Q_3 . The transistor action amplifies the current by a factor $(1 + \beta)$, where β is a BJT-dependent parameter. The amplified current flows out of the emitter and through R_L , producing an amplified voltage at the output. Q_3 should be a high- β transistor to maximize the gain. The total voltage gain is

$$H_v = g_m \cdot (1 + \beta) \cdot R_L \approx 48,000 \quad (3.50)$$

for $g_m = 25\text{mS}$, $\beta = 1000$, and $R_L = 1.91k\Omega$ which are typical values for the CDMS JFET amplifier.

3.6.2 DC Biasing

The basic DC biasing scheme for the amplifier is shown in Fig. 3.11. The DC base voltage of the cascode BJT is set by the voltage divider formed by R_{B1} and R_{B2} . Large filtering

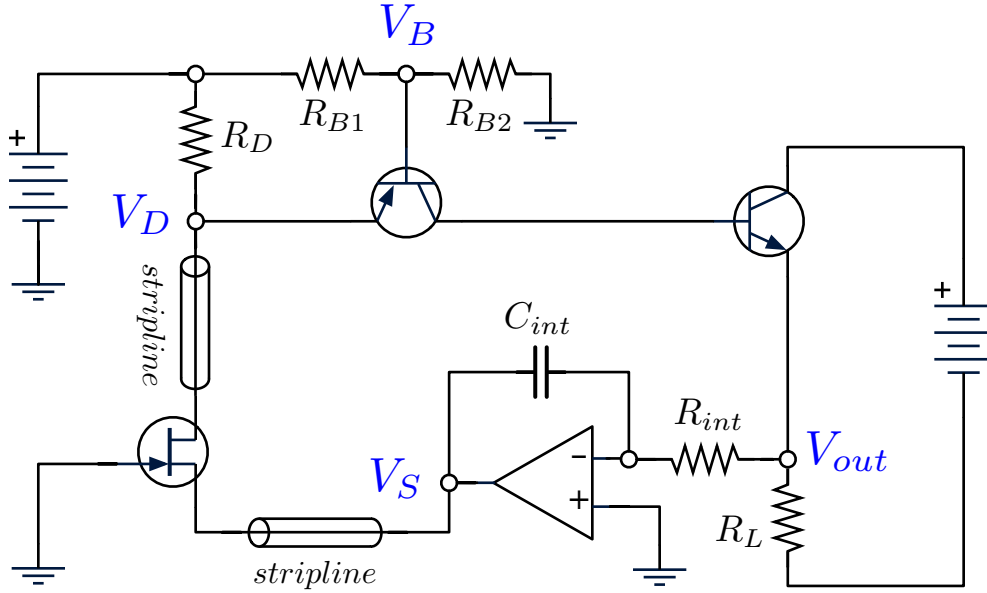


Figure 3.11: The basic DC biasing of the CDMS JFET amplifier. Not all components are shown.

capacitors (not shown) are placed in parallel to these resistors to ensure the base still appears as an AC ground. Due to the pn-junction formed between the emitter and the base, the emitter voltage (and hence the JFET drain voltage) is

$$V_D \approx V_B + 0.6V. \quad (3.51)$$

The total current flowing from the power supply into the node connecting the JFET drain and BJT emitter is therefore

$$I_{DE} = \frac{V_{supply} - V_D}{R_D} = I_{FET} + I_{BJT}, \quad (3.52)$$

as the current is divided between the JFET and the BJT.

The DC output voltage of the amplifier is set by the DC current flowing into the cascode BJT, which is amplified by the emitter-follower. Note that the emitter-follower is connected to a bipolar supply, allowing the output to swing between positive and negative voltages. To minimize the effects of microphonics, we desire the DC output voltage of the amplifier to be 0V because this will also become the JFET gate voltage with the addition of feedback. To accomplish this, we adjust the current division between the JFET and the cascode BJT through the use of an integrating amplifier. It can easily be shown through basic operational amplifier analysis that

$$V_S = - \int \frac{V_{out}}{R_{int}C_{int}} dt. \quad (3.53)$$

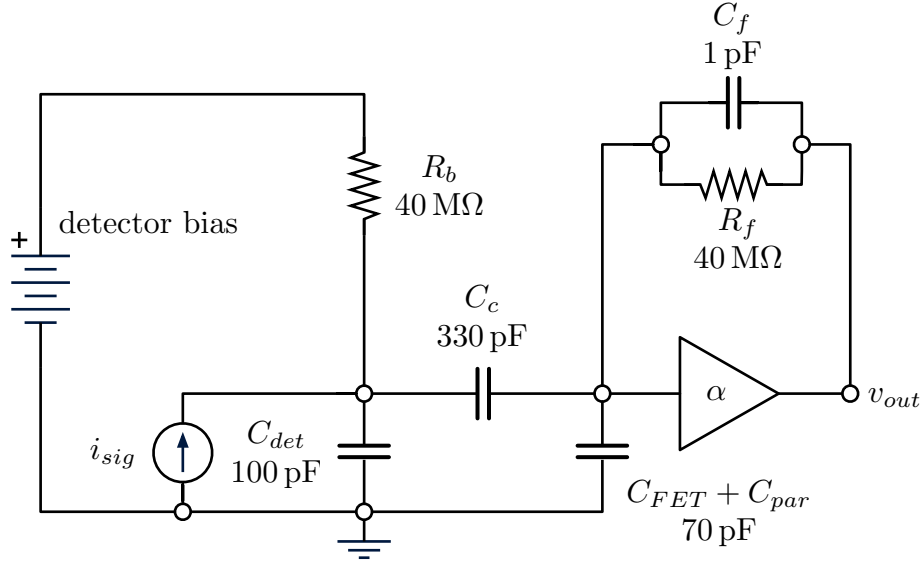


Figure 3.12: The CDMS closed loop JFET-based charge amplifier. The amplifier output is a pulse with amplitude proportional to the total charge with a $40\mu\text{s}$ falltime.

The source voltage is automatically adjusted through integration of the output voltage and changes until $V_{out} = 0$. CDMS uses a long time constant of $R_{int}C_{int} = 4.84$ seconds and low-pass filters the integrating amplifier output (not shown) such that it can only respond on time scales much longer than that of a typical signal pulse. The integrating amplifier can be ignored in the small signal limit.

3.6.3 Closed Loop Charge Amplifier with Detector

Fig. 3.12 depicts the closed loop CDMS JFET-based charge amplifier, connected to a detector electrode. The detector is biased using a room temperature power supply in series with a $40\text{ M}\Omega$ bias resistor, which is located at the base temperature ($\sim 50\text{ mK}$) stage of the fridge. In reality, the bias resistor is four individual $10\text{ M}\Omega$ resistors placed in series.

The detector is connected to the amplifier through a 330 pF coupling capacitor, also located at the base temperature stage of the fridge. As stated before, the relatively small value of this capacitor caused a 15% loss in charge signal. CDMS originally intended to use 1 nF coupling capacitors, however they were found to be easily damaged by electrostatic discharge. This was particularly problematic during tower installation at Soudan — the low humidity cleanroom environment made everything susceptible electrostatic buildup and shorted coupling capacitors would often be discovered during continuity checks. The compromise was to use 330 pF capacitors, which were less susceptible to electrostatic-induced damage.

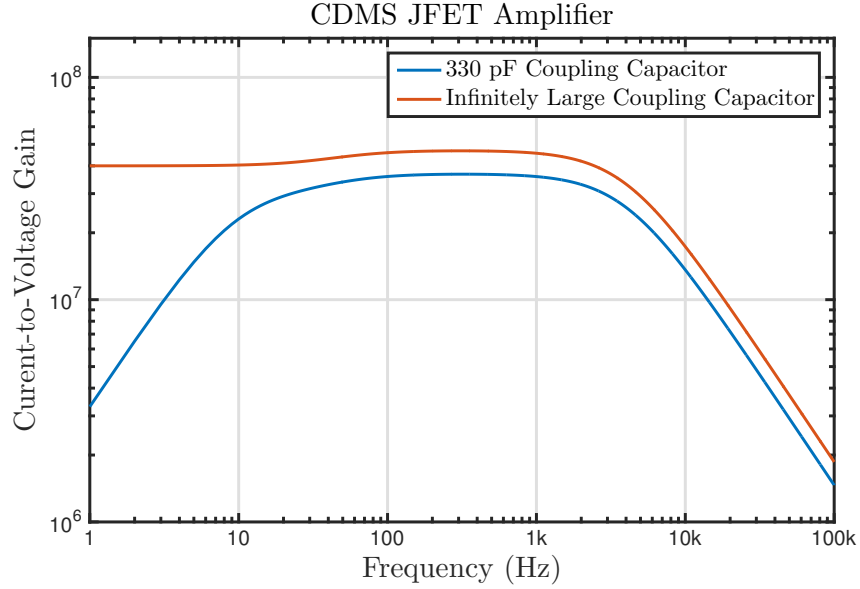


Figure 3.13: The closed loop current-to-voltage gain of the CDMS JFET amplifier. The blue curve shows the effects of the 330 pF coupling capacitor.

The feedback is formed by a $40\text{ M}\Omega$ resistor in parallel to 1 pF of capacitance. As with the bias resistor, the feedback resistor is four individual $10\text{ M}\Omega$ resistors placed in series, located at the base temperature stage. The feedback capacitance is the parasitic capacitance across these resistors, which has consistently been found to be $\sim 1\text{ pF}$. The $R_f C_f$ time constant is $40\text{ }\mu\text{s}$, which corresponds to a frequency of 4 kHz. Note that due to the coupling capacitor, the DC gate voltage will be equal to the DC output voltage. This is how the autozero circuit, which forces the output to 0 V by adjusting the JFET source voltage, also sets the gate to 0 V.

The closed loop current-to-voltage gain is shown in Fig. 3.13. We note that the current gain is maximized between $\sim 50\text{ Hz}$ and 4 kHz, and is relatively flat with frequency which indicates a direct conversion between current and voltage. Above the 4 kHz $R_f C_f$ pole, the the gain falls like $1/f$ indicating the voltage is proportional to the integral of the current, making this a charge amplifier. In the time domain, an impulse of charge from an event will produce a pulse on the output which decays with the $40\text{ }\mu\text{s}$ $R_f C_f$ time constant. In the frequency domain, the pulse will have the same shape as Fig. 3.13, scaled by the amplitude of the event. Compared to the closed loop voltage-to-voltage gain, shown in Fig. 3.14, we see that the current gain dominates up to the feedback integration pole. After this pole, the current gain decreases while the voltage gain remains constant.

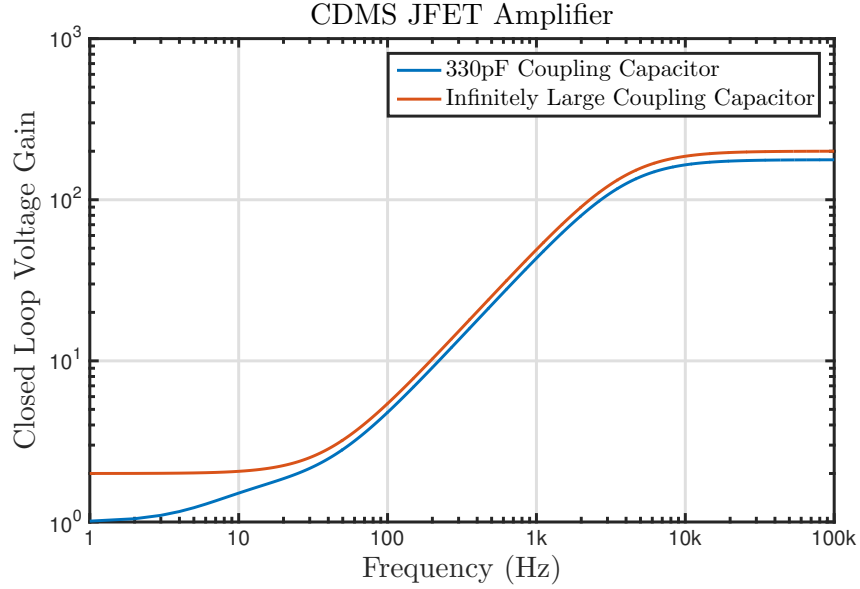


Figure 3.14: The closed loop voltage-to-voltage gain of the CDMS JFET amplifier. The blue curve shows the effects of the 330 pF coupling capacitor.

3.6.4 Performance

Putting all the pieces together, the predicted charge resolution of the CDMS JFET amplifier is roughly 250 eVee. As seen in Fig. 3.9, there is a forest of peaks throughout the signal bandwidth which significantly degrades the resolution. This behavior appears on essentially all Soudan detectors, although some are worse than others. The source of this peaks is believed to be related to the power and grounding of the CDMS Soudan installation. The typical charge resolution for Soudan detectors is about 460 eVee.[73] The noise situation is significantly better at test facilities, and a resolution of 260 eVee has been achieved at the UC Berkeley test facility using this amplifier.

Chapter 4

Ionization Readout with HEMTs

4.1 Introduction

So far, our discussion has focused on the use of a JFET as the input stage of the charge amplifier. High electron mobility transistors (HEMTs) are a different class of transistor-like 3-terminal device in which the current flowing between two terminals is controlled by the voltage on a third. This control terminal also exhibits high input impedance, similar to a JFET. While the underlying physics is different, we can model the HEMT in the same way as the JFET — a voltage-controlled current source with transconductance g_m and output resistance r_o .

HEMTs are formed from the heterojunction created at the interface of a highly-doped wide bandgap n-type material (for example, AlGaAs) and a non-doped narrow bandgap

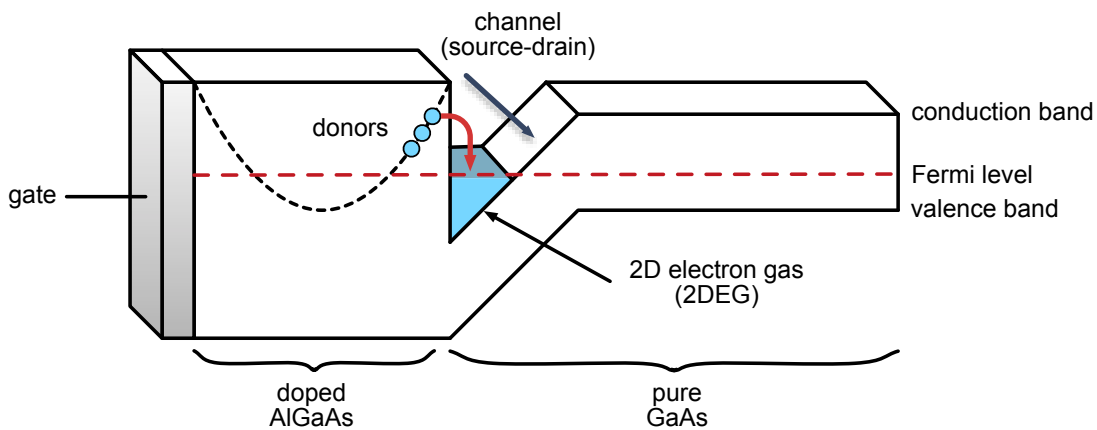


Figure 4.1: Basic physical structure of a HEMT. A voltage applied to the gate changes the location of the Fermi level, varying the number of electrons in the conducting channel.

Parameter	Value
Device Gate Geometry	32 μm x 2 mm
Drain-Source Voltage	100 mV
Drain-Source Current	1 mA
Power Dissipation	100 μW
Transconductance	36 mS
Output Resistance	1250 Ω
Input Capacitance	100 pF
Noise at 50 kHz	0.2 nV/ $\sqrt{\text{Hz}}$
Noise at 10 Hz	2.5 nV/ $\sqrt{\text{Hz}}$

Table 4.1: Typical 100 pF CNRS/LPN HEMT parameters.

channel (for example, GaAs). The basic structure is shown in Fig. 4.1. Free electrons from the donors in the n-type material fall into the deep conduction band quantum well on the GaAs side of the structure, forming a 2D electron gas (2DEG). The high purity of the undoped GaAs allows the electrons to be transported without frequent impurity collisions typical of doped semiconductors, giving the electrons a very high mobility.[140] This heterostructure also allows HEMTs to operate at cryogenic temperatures, unlike silicon-based JFETs which freeze out. Traditionally, HEMTs have not been used for readout of CDMS detectors due to high $1/f$ noise and have not matched the performance of Si JFETs from DC-100 kHz. High $1/f$ noise also prevents the use of MOSFETs which are somewhat capable of operating under cryogenic conditions.[141]

4.2 CNRS/LPN HEMTs

In 2011, while attending the 14th International Workshop on Low Temperature Detectors, Bernard Sadoulet and myself found a poster by Dr. Yong Jin of the CNRS Laboratoire de Photonique et de Nanostructures on the topic of HEMTs for readout of cryogenic semiconductor detectors.[142, 143] Dr. Jin reported producing GaAs HEMTs with better noise performance than the JFETs used by CDMS, specifically designed to have a capacitance which is well-matched to CDMS-style detectors. Intrigued, Bernard arranged to receive some samples for me to experiment with. The initial testing program has grown into the full-fledged adoption of the CNRS/LPN HEMT technology by SuperCDMS SNOLAB, inclusion of Dr. Jin as a SuperCDMS collaboration member, and an international collaboration between CDMS, EDELWEISS, and CNRS/LPN to optimize HEMT circuit architecture for the readout of cryogenic semiconductor detectors.

The characteristics of a typical CNRS/LPN HEMT are given in Table 4.1. A typical IV curve is shown in Fig. 4.2. The HEMTs have a number of properties which make them

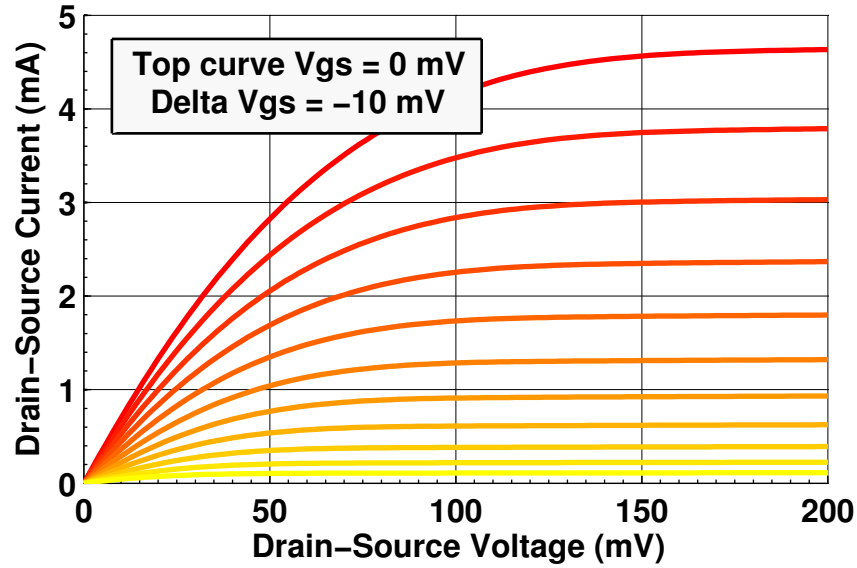


Figure 4.2: The drain-source current vs drain-source voltage, as a function of the gate-source voltage, of a typical CNRS/LPN HEMT. The bright red curve at the top of the figure corresponds to a gate-source voltage of 0 V. Each successive curve represents a -10 mV change in the gate-source voltage.

attractive for ionization readout of CDMS detectors:

- As shown in Fig. 4.3, the gate voltage noise of the HEMT is lower than the JFET at all frequencies. These HEMTs make it possible for SuperCDMS SNOLAB to reach its goal of 100 eVee ionization energy resolution.
- The power dissipation per device is much lower — only 100 μ W in comparison to 5 mW for a single JFET. Originally, SuperCDMS SNOLAB was intending to deploy at least 100 detectors (each with 4 charge channels) giving a total power dissipation of 2 W at the 4K stage of the cryostat due to only the JFETs. This is well above the cooling capacity of the cryostat. Using HEMTs, the power dissipation for the full SNOLAB payload would be only 40 mW — a factor 7.5 less than that of the much smaller SuperCDMS Soudan experiment!
- The HEMTs are designed to operate at a temperature of 4K and do not require a thermal standoff to operate. The Kapton membrane used to isolate the JFETs has vibrational modes which can be induced by the cryogenics system, creating microphonic noise. This is eliminated with the use of the HEMTs. In addition, the blackbody radiation emitted by the hot JFETs can be absorbed by detector, generating detri-

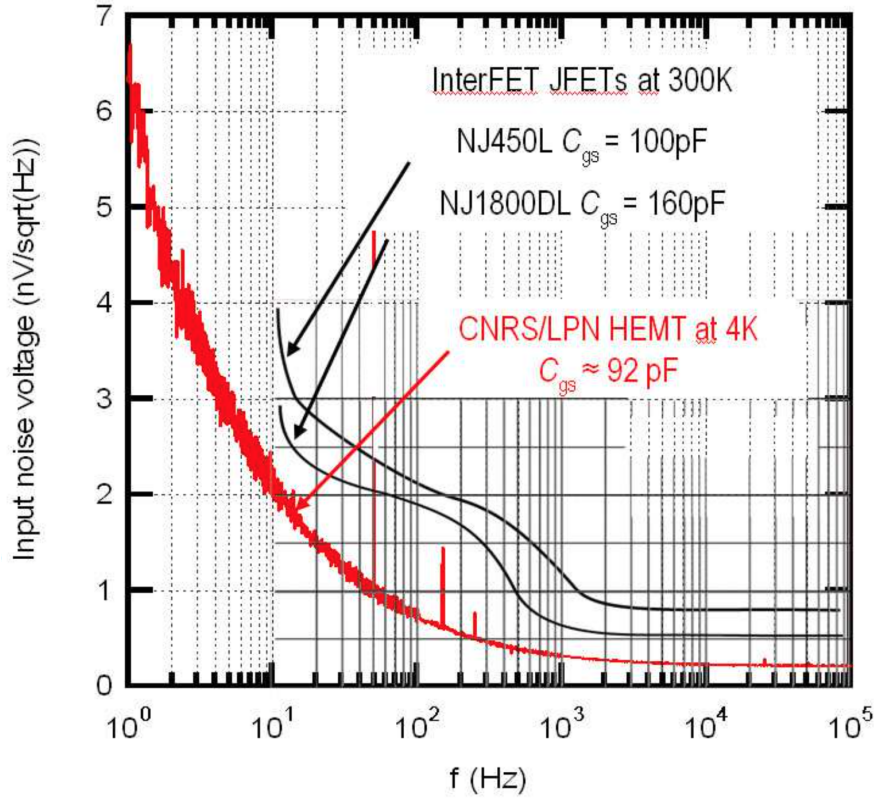


Figure 4.3: Comparison of the CNRS/LPN HEMT and Si JFET noise. The InterFET IF4501 JFETs used by CDMS are made by the InterFET NJ450L process, which has worse noise performance than the CNRS/LPN HEMT at all frequencies. Figure courtesy of Yong Jin.

mental space charge. The lower operating temperature of the HEMTs suppresses the blackbody radiation.

4.3 Modified CDMS Amplifier

In principal, nothing about the design of the CDMS JFET amplifier prevents the replacement of the JFET by a HEMT. In practice, this proves to be much more difficult. The drain-source, gate-source, and gate-drain voltages at the operating point of the HEMT are about 10 times smaller than those of the JFET. This prevents proper biasing of the Q_2 and Q_3 transistors, as the drain-gate voltage drop of ~ 150 mV of the HEMT must also be the voltage drop between the Q_2 and Q_3 emitters, as seen in Fig. 3.11, if the gate is to be held at a DC potential of 0V. The solution is to introduce a p-type Darlington pair (Q_4 and Q_5 in Fig. 4.4), which follows the original output voltage before connecting to the feedback

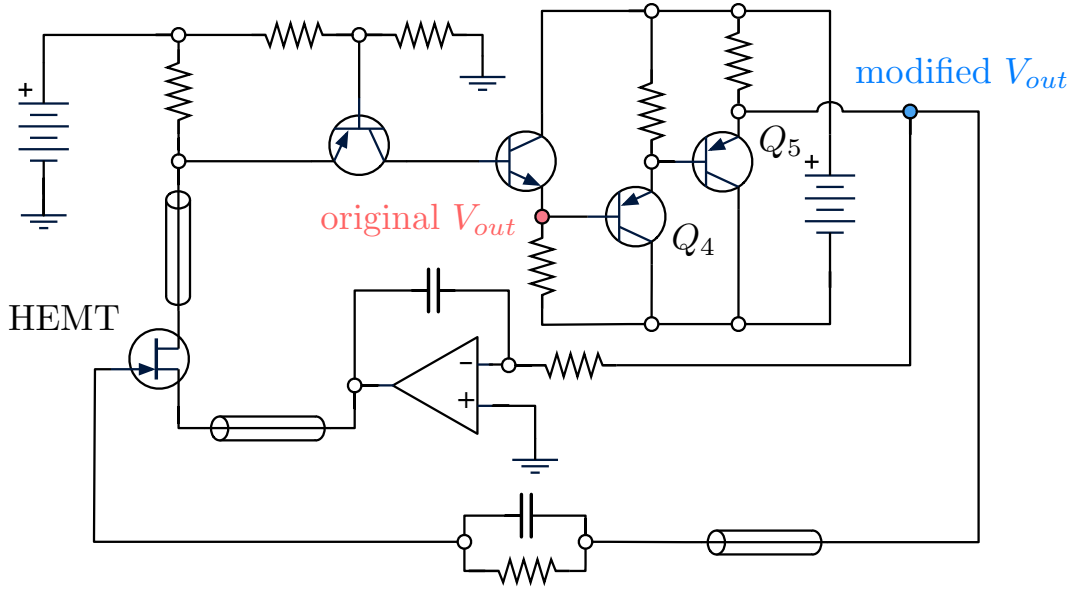


Figure 4.4: CDMS JFET amplifier modified to replace the JFET with a HEMT. Note the p-type Darlington pair on the output which allows the original output to run at a negative voltage, while still maintaining 0V on the new output.

and autozero circuit. The original output is DC biased at a negative value to allow for proper biasing of the HEMT, Q_2 , and Q_3 . The DC voltage is then stepped back up by the Darlington pair. Since the signal has already been converted to a voltage at the original output, no additional signal gain is added through this modification. The noise introduced by Darlington pair is negligible.

The input voltage noise of the modified CDMS amplifier was measured using the setup shown in Fig. 4.5. A low resistance feedback network is used to set the voltage gain to 101, while any current noise is shunted by the $100\ \Omega$ resistor and does not generate an appreciable voltage. The measured noise referred to the input is shown Fig. 4.6. Using the HEMT, the noise is clearly lower above $\sim 1\ \text{kHz}$. The high frequency white noise of the HEMT is $0.3\ \text{nV}/\sqrt{\text{Hz}}$, in comparison to $0.5\ \text{nV}/\sqrt{\text{Hz}}$. The gate voltage noise of a single HEMT, as measured by CNRS/LPN, is also shown in Fig. 4.6. The discrepancy between the CNRS/LPN measurements and the modified CDMS amplifier are due to additional noise sources in the amplifier circuit and should not be attributed to the HEMT.

In order to test whether the *current noise* of the HEMTs would be an issue, the noise was also measured using the standard CDMS $40\ \text{M}\Omega$ high impedance feedback network. The measured output noise in this configuration is shown in Fig. 4.7. Again, we see that use of the HEMT reduces the high frequency noise, but below the $4\ \text{kHz}$ integration pole the noise

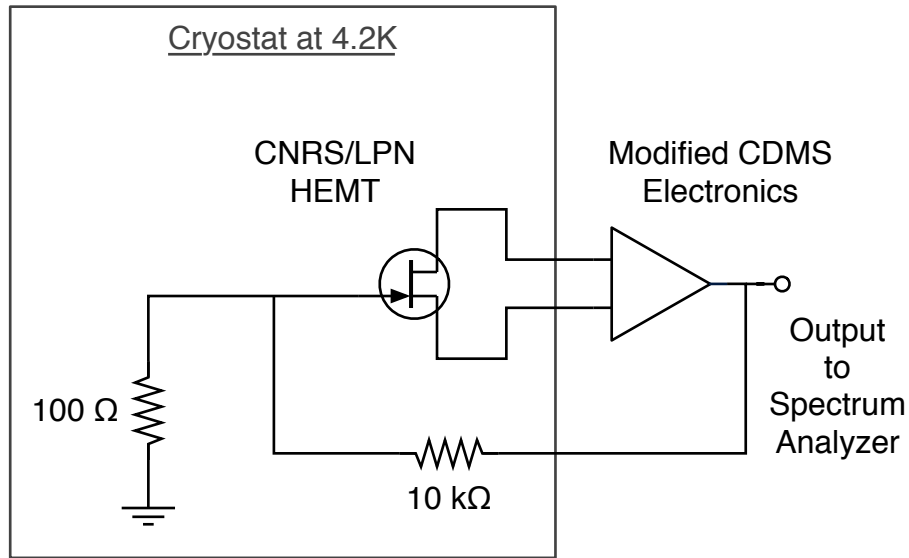


Figure 4.5: Experimental setup to measure the input voltage noise of the modified CDMS amplifier.

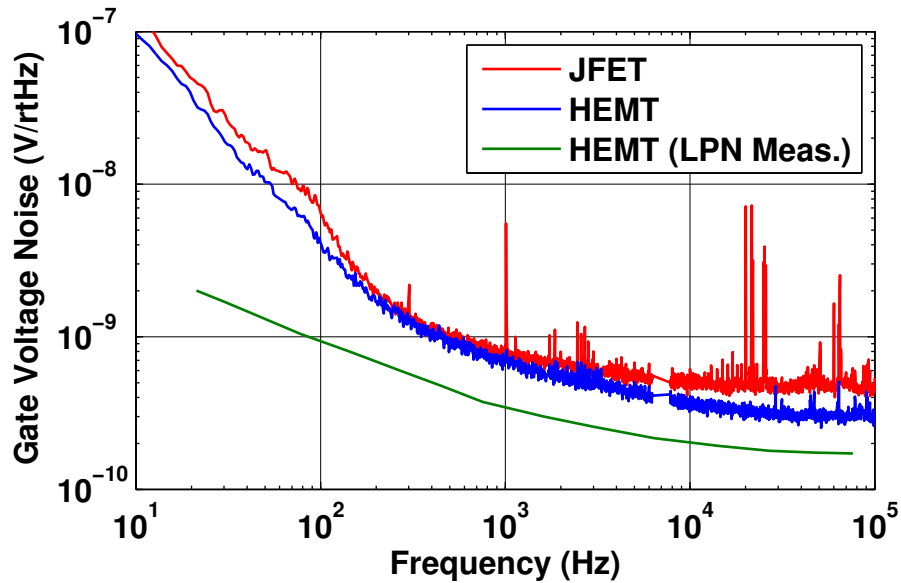


Figure 4.6: Measured voltage noise of the modified CDMS amplifier, referred to the input (*blue*). Also shown is the voltage noise of amplifier when used with a JFET (*red*), and the voltage noise of the HEMT as measured by CNRS/LPN (*green*).[144]

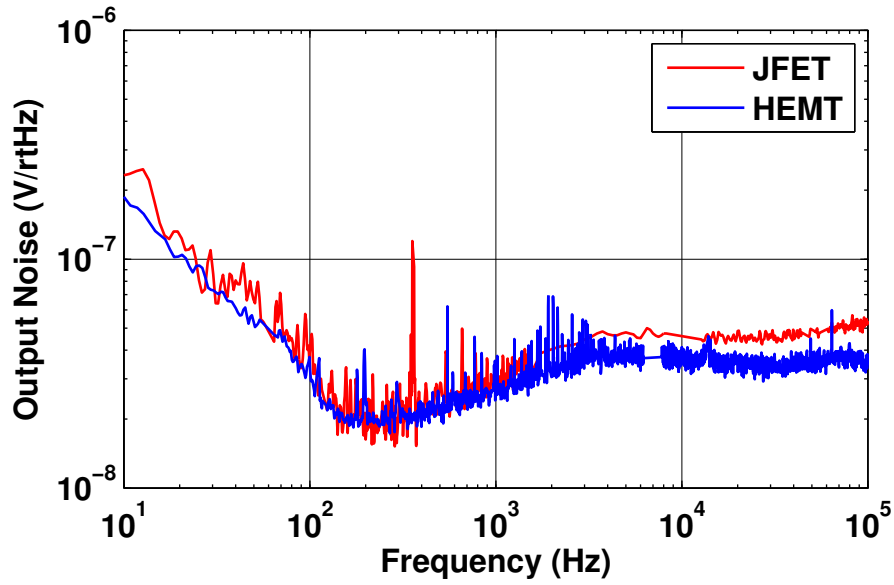


Figure 4.7: Measured output voltage noise of the modified CDMS amplifier in a high resistance feedback configuration (*blue*), in comparison with using the standard JFET (*red*).[144]

Quantity (FWHM)	HEMT	JFET
Holes 60 keV	1.040 ± 0.013 keV	1.059 ± 0.012 keV
Electrons 60 keV	0.998 ± 0.012 keV	0.970 ± 0.012 keV
Baseline Noise	0.607 ± 0.016 keV	0.605 ± 0.016 keV

Table 4.2: Measured energy resolution of the modified CDMS amplifier (with HEMT) in comparison to the standard JFET amplifier.

is essentially unchanged. The current noise of the HEMT is unimportant compared to the noise of the other components of the amplifier.

A CDMS II detector exposed to a collimated ^{241}Am source was used to assess the energy resolution of the modified amplifier. Energy calibration was performed using the 60 keV photopeak. The spectrum obtained while collecting electrons can be seen in Fig. 4.8. The resulting FWHM of the 60 keV photopeak when collecting electrons, when collecting holes, and the baseline noise is shown in Table 4.2. For all cases we find the resolution using HEMTs to be equivalent to that using JFETs, indicating that the amplifier resolution is mostly dominated by other noise sources in the amplifier. The difference between 60 keV peaks and the baseline resolution is due to detector effects smearing the energy peak resolution and is unrelated to the amplifier performance.

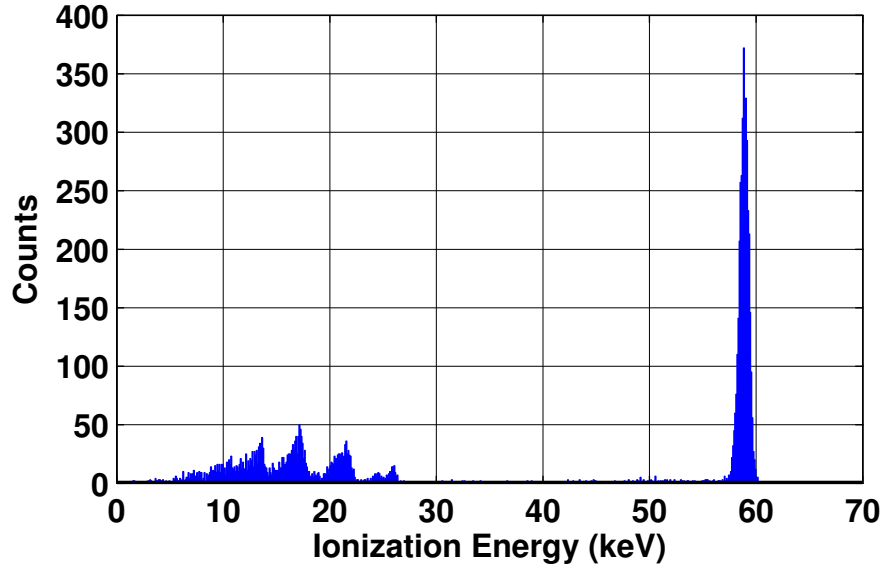


Figure 4.8: Electron collection spectrum of a CDMS II detector exposed to a ^{241}Am source. Ionization readout was performed using the modified CDMS amplifier.[144]

This work is discussed in more detail in [144]. The results demonstrated the feasibility of using HEMTs for CDMS ionization readout, and show that at least an equivalent energy resolution to the JFET amplifier can be achieved. This led to the selection of HEMTs as the input stage component for SuperCDMS SNOLAB, and the inclusion of Yong Jin as a SuperCDMS collaboration member.

4.4 Fully Cryogenic HEMT Amplifier

Instead of using the HEMT as a cryogenic input stage, with amplification occurring on the room temperature electronics, multiple HEMTs can be used to create a fully cryogenic charge amplifier operating at $T=4\text{K}$. While cryogenic amplification has always been desirable, the challenges of operating multiple silicon-based transistors in a single circuit within the cryostat has historically made this impossible. The low power dissipation of the HEMTs removes this restriction. Not only does the fully cryogenic amplifier allow us the full benefits of the HEMT's low noise, by providing amplification within the cryostat the signal will be less susceptible to environmental noise introduced by the long stretch of cable running from 4K to room temperature.

The closed loop configuration of the fully cryogenic HEMT amplifier is shown in Fig. 4.9. The triangle represents a HEMT-based open loop voltage amplifier. A feedback capacitor closes the loop. A HEMT switch is added to reset the DC input voltage/feedback capacitor

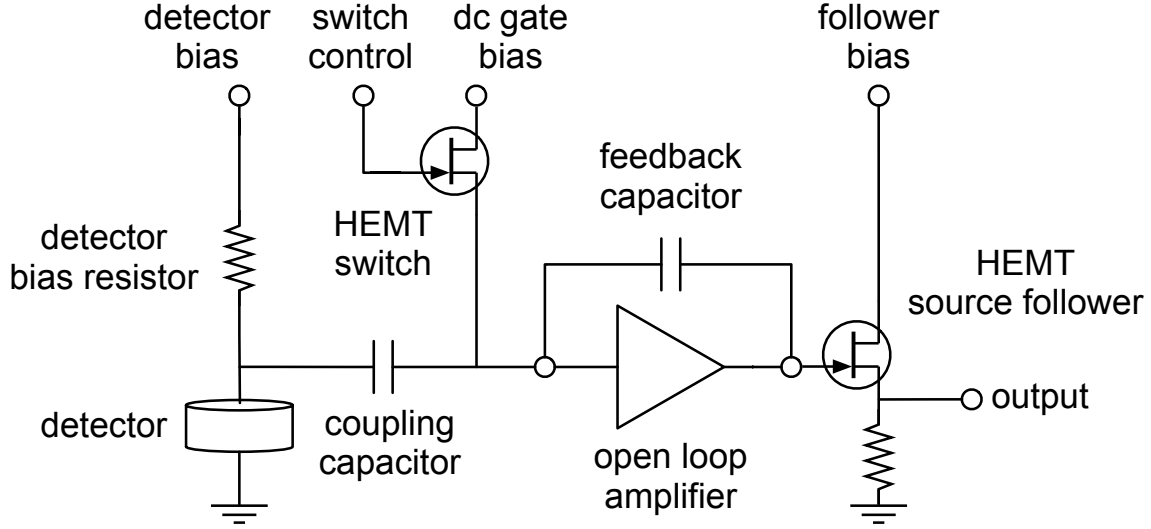


Figure 4.9: The fully cryogenic HEMT amplifier, which consists of an open loop HEMT voltage amplifier, HEMT follower, and HEMT switch.[145]

and is open during normal operation. A HEMT source follower is added to isolate the amplifier output from high capacitance cabling. The detector is AC-coupled to the input. We now discuss each of these in more detail.

4.4.1 Open Loop Voltage Amplifier

The open loop amplifier is shown in Fig. 4.10 and requires four HEMTs, labeled Q_1 - Q_4 . Two types of HEMT gate geometries are used to maximize performance. Due to its low noise, the 100 pF HEMT is still used as the input component. Noisier HEMTs with smaller input capacitance (30 pF) but higher transconductance are used in the remainder of the circuit, where higher noise can be tolerated. The amplifier is powered by two room-temperature voltage sources. The gate of Q_1 should be held at a DC value of approximately -25 mV and can be adjusted to maximize gain. All components, other than the external power sources, are within the cryostat at the T=4K layer.

Q_1 and Q_2 form a HEMT cascode which converts the input voltage into a signal current via the transconductance of Q_1 . As with the JFET/BJT cascode of the CDMS JFET amplifier, the addition of Q_2 boosts the effective output impedance of Q_1 so that it is able to drive a larger load. The signal current is converted back into an inverted, amplified voltage on the drain of Q_2 by passing through the effective load resistance formed by Q_3 and Q_4 . Both are individual self-biasing current sources, placed in series to increase the load resistance as seen by the output of the cascode and enhancing the overall voltage gain.[139]

The equivalent small signal model of the amplifier is also shown in Fig. 4.10. The cascode

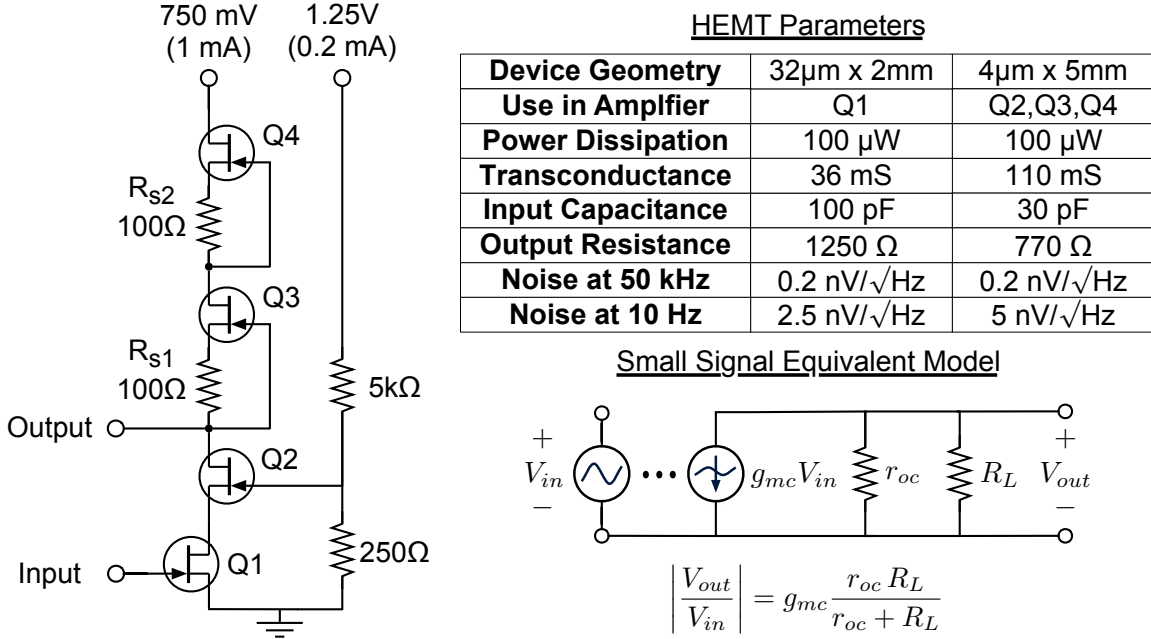


Figure 4.10: *Left:* The open loop voltage amplifier schematic. The HEMTs are labeled Q_1 - Q_4 . *Top Right:* Parameters for two types of HEMTs used in the amplifier. *Bottom Right:* The equivalent small signal model of the open loop amplifier.

is represented by a voltage-controlled current source with transconductance g_{mc} in parallel to its output resistance r_{oc} . The effective load resistance R_L of Q_3 and Q_4 appears in parallel to the cascode. The voltage gain is given by

$$G = \left| \frac{V_{out}}{V_{in}} \right| = g_{mc} \frac{r_{oc} R_L}{r_{oc} + R_L} \approx \text{several hundred} \quad (4.1)$$

and strongly depends on the parameters of the specific HEMTs used, typical of open loop amplifiers. The gain is limited by effective load resistance. The effective transconductance of the cascode is

$$g_{mc} = \frac{g_{m1} r_{o1} (1 + g_{m2} r_{o2})}{r_{o1} + r_{o2} + g_{m2} r_{o1} r_{o2}} \approx g_{m1}, \quad (4.2)$$

where lowercase letters refer to the small signal parameters of the numbered HEMTs. The transconductance of the cascode is essentially the same as the transconductance of the input HEMT.

The effective output resistance of the cascode is given by

$$r_{oc} = r_{o1} + r_{o2} + g_{m2} r_{o1} r_{o2} \approx g_{m2} r_{o1} r_{o2} \approx 100 \text{ k}\Omega. \quad (4.3)$$

The power of the cascode comes from the product of the two individual output resistances. This allows the cascode to drive a much bigger load, increasing the voltage gain of the

amplifier. This is also why it is better to use a 30 pF HEMT rather than a 100 pF HEMT for Q_2 in the cascode as $g_{m2} r_{o2} > g_{m1} r_{o1}$.

The effective load resistance of the Q_3 - Q_4 combination is given by

$$R_L = r_{o1} + R_{s1} + g_{m1} r_{o1} R_{s1} + R_{s2} + g_{m2} r_{o2} R_{s2} \approx 10 \text{ k}\Omega. \quad (4.4)$$

This resistance depends strongly on the DC biasing points of Q_3 and Q_4 , which determine their small signal parameters. As the gate-source voltage of each HEMT is set by a fixed bias resistor ($R_{s1,2}$), the HEMTs must be well-matched to ensure both are in saturation. In the case that the effective load resistance is much greater than the output resistance of the cascode ($R_L \gg r_{oc}$), $G \approx g_{mc} R_L$ and the maximum achievable gain is ≈ 3600 . We believe this may be achieved by eliminating R_{s2} and connecting the gate of Q_4 to the source of Q_3 , which would boost R_L to several M Ω , however this topology has not yet been tested.

The -3 dB bandwidth of the open loop amplifier is approximately determined by

$$f_{-3dB} = \frac{1}{2\pi R_{out} C_{out}}, \quad (4.5)$$

where R_{out} is the parallel combination of r_{oc} and R_L and C_{out} the capacitance to ground present on the output. If the output capacitance can be minimized, the amplifier can be fast.

4.4.2 Output Follower

Due to the length and low thermal conductivity required of the 4K-to-300K cabling, a direct connection between the amplifier output and room temperature is undesirable. The cable represents a load of about 300 to 400 pF, limiting the open loop amplifier bandwidth to ~ 50 kHz. To minimize these effects, we buffer the output using a 5 pF input capacitance HEMT configured as a source follower. The source resistor can either be placed inside the cryostat and connected to a local ground, or placed at room temperature and connected to a warm ground to minimize power dissipation. The DC bias current and load resistor must result in the DC voltage of the follower source being high enough above the open loop amplifier direct output so that the follower is put into saturation. This can be determined by increasing the voltage on the follower drain until the source voltage (the follower output) stops changing. Using a 1 k Ω source resistor, we found a drain voltage of 750 mV would result in a drain current of 0.3 mA, with a DC output voltage about 100 mV above the direct output.

The follower is not perfect. Using the small signal model shown in Fig. 4.11, the voltage gain of the amplifier is

$$H_{fol} = \frac{g_m Z_{total}}{1 + g_m Z_{total}}, \quad (4.6)$$

where Z_{total} is the total parallel impedance to ground on the source of the follower. About 8% of the voltage gain is lost, however since the both the output signal and noise are multiplied by this same factor the signal-to-noise is preserved.

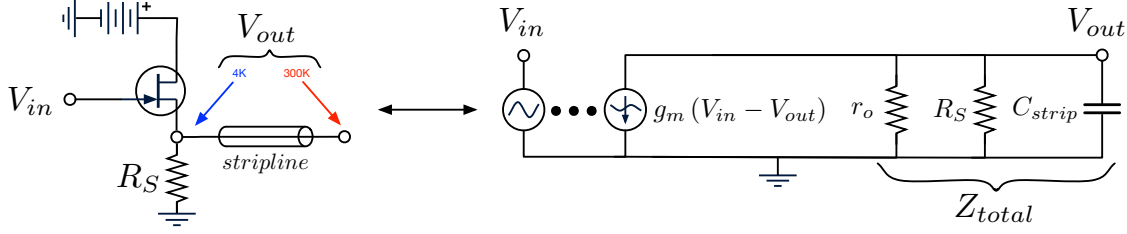


Figure 4.11: The HEMT follower, along with its small signal model.

4.4.3 Reset Switch and Leakage Current

The input of the open loop amplifier requires a DC voltage of about -25 mV in order to put Q_1 into saturation. If this were to be provided by a direct connection to a room temperature power supply, the current signal from the detector would be shunted to ground. Instead, a HEMT switch is used to provide the DC bias voltage to the amplifier input. The switch can be opened and closed by varying the switch gate voltage. When closed, the input capacitance is charged to the appropriate DC level however any current from the detector is shunted through the switch. When the switch is open, the DC level persists and the detector signal can be integrated until the voltage on the input capacitance drifts out of the operating range of the amplifier. The switch can then again be closed and the process repeated. A HEMT with small gate-source capacitance (≈ 5 pF) was used to limit charge injection into the input capacitance during switching and noise injection via capacitance coupling when open.

4.4.4 Capacitive Feedback

Feedback is provided by connecting a capacitor between the input and the output, stabilizing the gain. As discussed earlier, at frequencies above the integration pole ($1/2\pi R_{bias} C_{in}$) the closed loop voltage gain is

$$G_{cl} = \frac{\alpha}{1 + \alpha\beta} = \frac{\alpha}{1 + \alpha \left(\frac{C_f}{C_{in} + C_f} \right)} \approx \frac{C_{in}}{C_f} \quad (4.7)$$

where α is the open loop gain, β is the feedback factor, and C_{in} is the total input capacitance. The approximation holds as long as $\alpha\beta \gg 1$ and $C_{in} \gg C_f$.

4.4.5 Predicted Resolution

The amplifier has been designed such that the noise on the output is completely dominated by the noise of the input (Q_1) HEMT. The reported 100 pF HEMT voltage and current

noise[142, 143] can be used to perform an optimal filter calculation and predict the charge resolution of the amplifier. Using a total input capacitance of 300 pF (200 pF from the detector+parasitics, 100 pF from the HEMT), we calculate a predicted resolution of 83 eVee (electron equivalent) in the absence of bias resistor Johnson noise. Note that achieving this resolution does not require a fully-cryogenic amplifier — as long as the output noise is completely dominated by the noise of the input HEMT, the same resolution can be achieved.

4.5 Prototype Amplifier Performance

A prototype amplifier was constructed based on this design and is described in detail in [145]. The total power consumption of the complete amplifier (with follower) was ~ 1 mW. The open loop voltage gain was determined to be 340. The closed loop voltage gain with a total input capacitance of 280 pF and a feedback capacitance of 1.6 pF was determined to be 120. Note that the full expression of Eq. 4.7 should be used as the open loop gain is not much larger than the closed loop gain. The 100 pF HEMT used for the Q_1 transistor came from a fabrication series with slightly worse noise performance than the best CNRS/LPN HEMTs produced, however the performance ($\sigma_q = 150$ eVee) is the best ever observed from a CDMS charge amplifier.

The measured open and closed loop noise performance was used to develop an empirical noise model. The open loop voltage noise can be measured by leaving the reset switch closed, and dividing the recorded output noise by the open loop gain. In our model, the open loop voltage noise is given by

$$e_n^2 = \left(\frac{10 \times 10^{-9}}{f} \right)^2 + (0.23 \times 10^{-9})^2 \quad (\text{in V}^2/\text{Hz}). \quad (4.8)$$

The current noise is modeled as

$$i_n = 1.8 \times 10^{-17} \sqrt{f} \quad (\text{in A}/\sqrt{\text{Hz}}), \quad (4.9)$$

based on the measurements of Yong Jin.[142, 143] The model can be verified by measured the output noise of the amplifier with the reset switch open (closed loop mode) and the input connected to a capacitance of known value. The input equivalent voltage noise is the combined contributions of the open loop voltage noise and voltage-equivalent current noise. A comparison between the model and measured noise can be seen in Fig. 4.12. The noise of the complete amplifier is consistent with being fully determined by the noise of the input HEMT.

The HEMT switch transitioned from fully closed to fully open in $\approx 10 \mu\text{s}$. With the switch open, the input leakage current was determined to be $\leq 3 \cdot 10^{-17} \text{A}$. Shot noise from this current is clearly inconsequential and the limit on the leakage current is very conservative. While there was some slight drift in the DC output level, we were able to run the amplifier in closed loop mode without reset for periods as long as one hour. This improves the limit

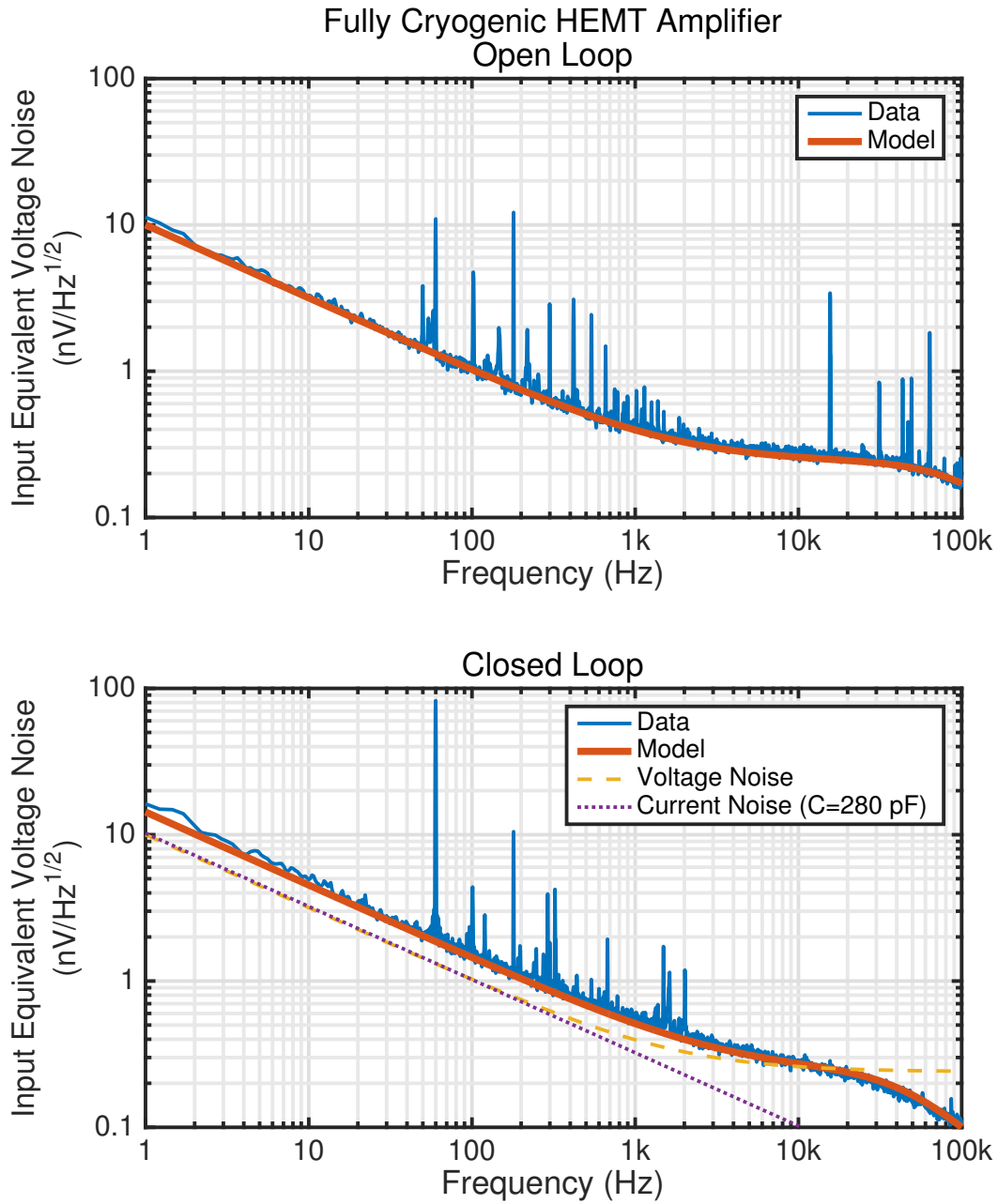


Figure 4.12: *Top:* Open loop noise performance of the HEMT amplifier. *Bottom:* Closed loop noise performance of the HEMT amplifier, for a total input capacitance of 280 pF. The difference in high frequency roll off is due to an increase in output cable capacitance when taking the closed loop measurements, which were performed in the dilution refrigerator.

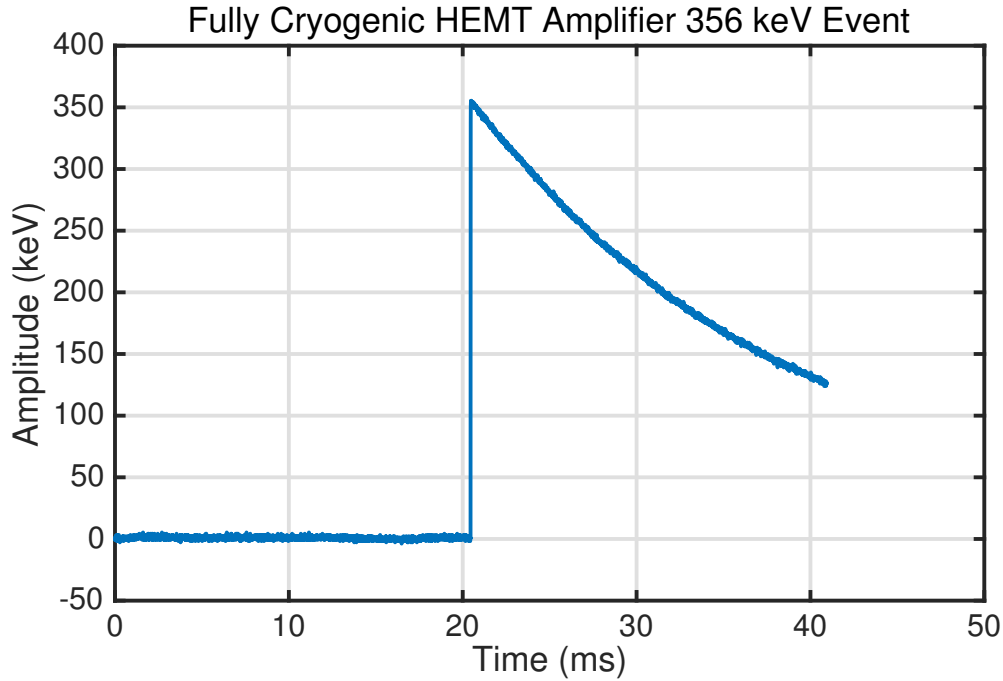


Figure 4.13: A typical 356 keV pulse from the fully cryogenic HEMT amplifier.

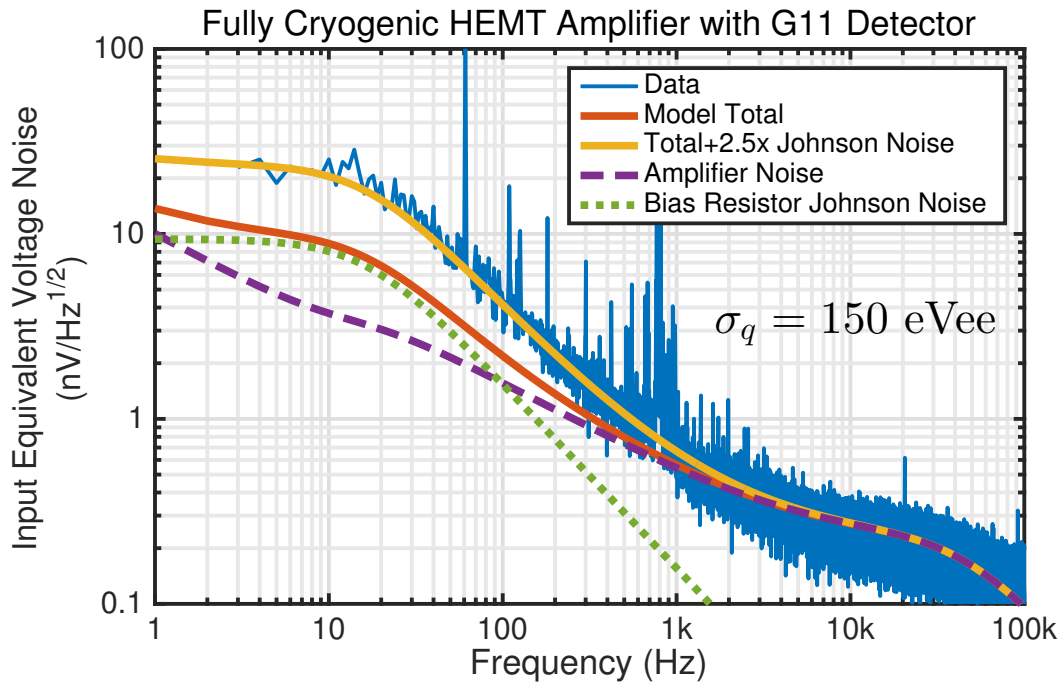


Figure 4.14: PSD of the fully cryogenic HEMT amplifier coupled to a CDMS oZIP detector (G11). The 40 M Ω bias resistor has an additional noise component — the measured resistor noise is roughly 2.5x the intrinsic Johnson noise.

by a factor of 30, but the maximum recordable trace length used to determine the leakage was limited to 100 seconds.

The amplifier performance when coupled to a CDMS II oZIP detector (G11) was also measured. The energy scale was calibrated to the 356 keV peak from ^{133}Ba . A typical 356 keV pulse is shown in Fig. 4.13. Note the ~ 25 ms faltime even though we are using purely capacitive feedback. This is due to the relatively low open loop gain in comparison to the closed loop gain. An amplifier with much higher open loop gain would instead have a step-function response as the falltime increases with open loop gain.

In this setup, the standard CDMS II 40 M Ω bias resistor was used. We observed the bias resistor to possess extra noise equivalent to roughly 2.5 times the standard Johnson noise, confirmed in a separate measurement performed without the detector. The measured and modeled PSDs are shown in Fig. 4.14. Even with this additional noise, the optimal filter **ionization resolution is 150 eVee** — *the lowest ever recorded for a CDMS-style detector*. The resolution is limited due to the bias resistor. By using an active detector reset, or by increasing the value of the bias resistor, $\sigma_q \approx 100$ eVee should be achievable.

4.6 Comparing Topologies

There has been considerable discussion within the CDMS collaboration with regards to what specific charge amplifier topology should be used for the SuperCDMS SNOLAB baseline design. The fundamental question is whether to use an active reset, in which only a feedback capacitor is used, versus a passive reset, in which a feedback resistor is also used. When using an active reset, there is no DC connection to the input and a switch must be used to occasionally reset the DC biasing as has been done in the fully cryogenic HEMT amplifier. With a passive reset, the feedback resistor provides a path for DC biasing but adds temperature-dependent Johnson noise to the system. Additionally, one may consider biasing the *detector* through a switch to eliminate the Johnson noise of the bias resistor.

In this section, we compare the expected resolution for 4 separate HEMT amplifier topologies:

- No bias or feedback resistor. Both detector and amplifier would be periodically reset using HEMT switches.
- 2 G Ω bias resistor at T=50 mK. The detector has a dedicated DC bias, while the amplifier uses an active reset.
- 2 G Ω bias resistor at T=50 mK, and 2 G Ω feedback resistor at T=50 mK. The amplifier uses a passive reset, however additional wires from 4K to base temperature are required to hold the feedback resistor at 50 mK.
- 2 G Ω bias resistor at T=50 mK, and 2 G Ω feedback resistor at T=4 K. The feedback resistor has been placed at the T=4 K stage of the cryostat, reducing the number of wires required to reach base temperature.

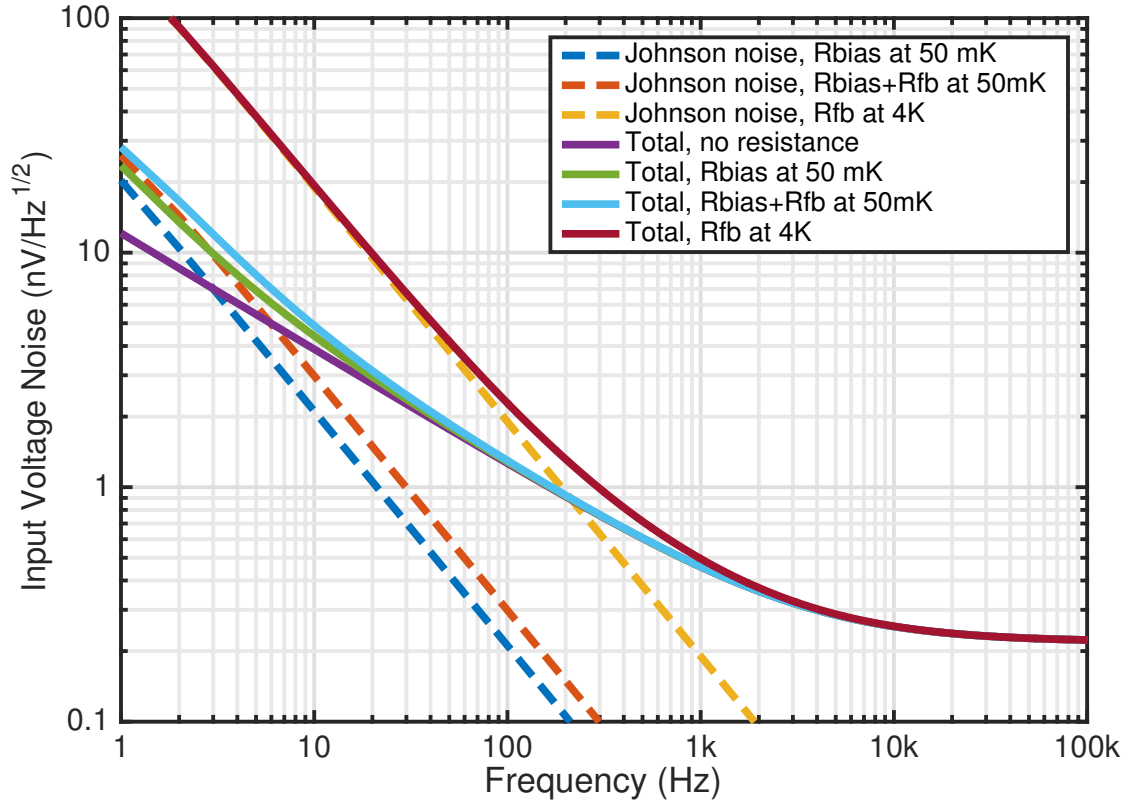


Figure 4.15: Comparison of the various amplifier topologies using active and passive reset schemes. Note the bias resistor is always assumed to be at $T=50$ mK. The degradation in noise performance due to placing the feedback resistor at $T=4$ K is clear.

PSDs are calculated assumed a total input capacitance of 300 pF, with 200 pF from the detector/parasitics and 100 pF from the HEMT. The PSDs, along with the various contributions from the resistors, are shown in Fig. 4.15. The optimal filter calculation is performed over a bandwidth of 5 Hz-100 kHz in 1 Hz steps. The signal shape is total input impedance, which varies between topologies. Results are given in Table 4.3.

We see that while there is not much degradation if the passive components are kept at $T=50$ mK, placing the feedback resistor at $T=4$ K significantly degrades the performance. The goal of 100 eVee cannot be achieved unless an active amplifier reset is used, or the feedback resistor kept at base temperature. Note that these results are somewhat optimistic and assume that the only noise contribution from the resistors is due to their Johnson noise. As we saw with the 40 M Ω bias resistor used for G11, this style of resistor has additional noise contributions. It is not known whether the 2 G Ω resistors intended for SuperCDMS SNOLAB will also have this additional noise.

Bias Resistor	Temperature	Feedback Resistor	Temperature	Resolution (eVee)
No	-	No	-	83
Yes	50 mK	No	-	86
Yes	50 mK	Yes	50 mK	88
Yes	50 mK	Yes	4 K	126

Table 4.3: Predicted ionization resolution for the various amplifier topologies described in the text.

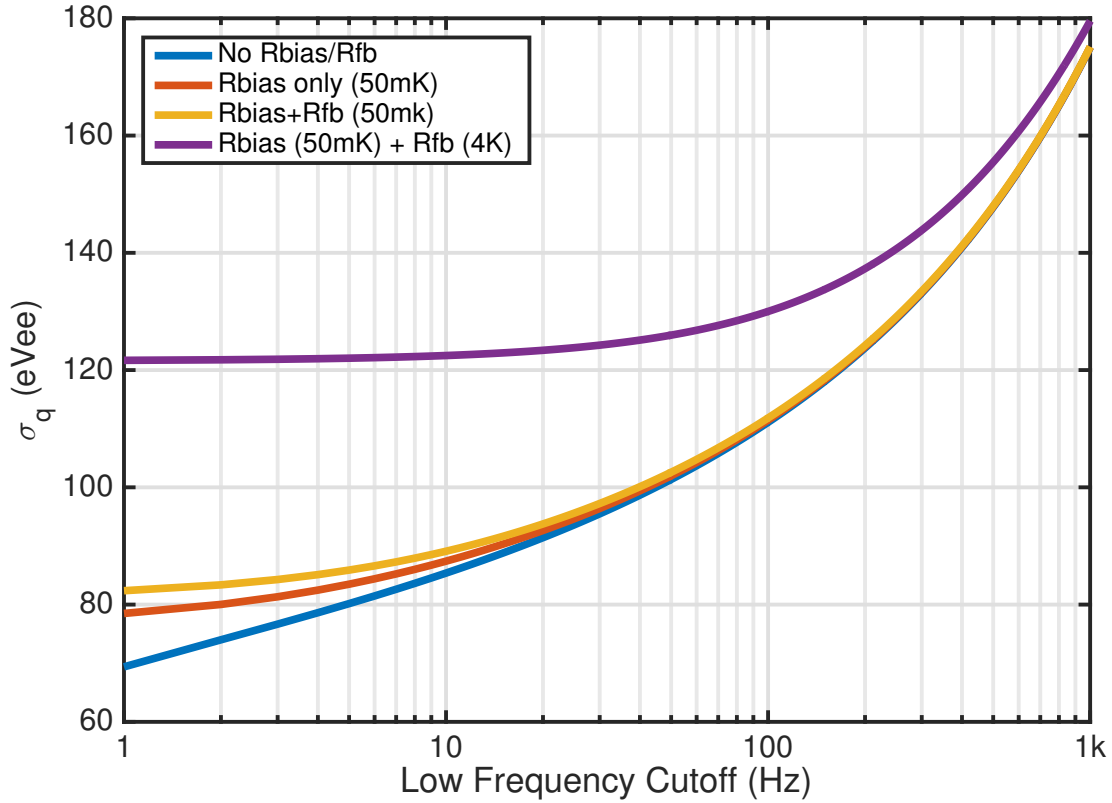


Figure 4.16: Ionization resolution as a function of low frequency cutoff for the various amplifier topologies discussed in the text.

The variation in ionization resolution as a function of the optimal filter low frequency cutoff is shown in Fig. 4.16. The benefits of going to lower frequency (longer trace length) are reduced as the Johnson noise increases above the HEMT $1/f$ noise. In order to reach 100 eVee resolution, a low frequency cutoff of about 50 Hz or lower is required.

Chapter 5

The Physics of Electrons and Holes in sub-Kelvin Germanium

5.1 Introduction

The charge signal measured by CDMS detectors is determined by the microscopic physics of electrons and holes in sub-Kelvin germanium. While semiconductors (in particular germanium) have been heavily studied for many years, there is a very limited amount of literature on the unique conditions of high purity, sub-Kelvin temperature, and low electric field — most of the work has been performed by the CDMS and EDELWEISS collaborations themselves. An understanding of the detailed physics behind charge carrier transport is becoming increasingly important as these experiments moves towards larger detectors and lower energy thresholds. The only comprehensive reference on charge transport in CDMS-style detectors (to my knowledge) is the dissertation of Kyle Sundqvist[146], who was also a graduate student in the Sadoulet lab and my mentor. Following his work, this chapter is an attempt to introduce the important microphysics which dictate ionization collection in CDMS detectors, discuss how we can determine the fundamental carrier distributions through simulation, present capture cross sections for the dominant processes, and connect the microphysics to its macroscopic effects on the drift current generated by an event.

5.2 Band Structure and Effective Mass

5.2.1 The Schrödinger Equation

The repeating, geometric arrangement of germanium atoms which make up the crystal create a periodic potential which matches the symmetries of the lattice. Solving the Schrödinger equation with this potential produces solutions which take the form

$$\psi_{k,n}(\mathbf{r}) = u_k(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}} e^{i\mathcal{E}_n(\mathbf{k})t/\hbar}, \quad (5.1)$$

where $\mathbf{k}$ is known as the wavevector and the u_k are Bloch functions which obey the periodicity of the lattice. Each value of k is associated with an infinite set of discrete energy eigenvalues, which we label by n , such that

$$\hat{H}\psi_{k,n} = \mathcal{E}_n(\mathbf{k})\psi_{k,n}, \quad (5.2)$$

where $\hat{H}$ is the Hamiltonian. The variation of $\mathcal{E}_n$ with $\mathbf{k}$ traces out an *energy band*, which is labelled by n . In general, there exists an energy gap between adjacent bands. The location of the chemical potential μ , also known as the Fermi level, and the size of the energy gap determines whether the material acts as a metal (μ lies within a band), an insulator (μ lies within an energy gap, and the gap is large), and a semiconductor (μ lies within an energy gap, and the gap is small).

Germanium and silicon are both semiconductors and hence the Fermi level lies between two bands. The band below the Fermi level is known as the valence band, while the band above the Fermi level is known as the conduction band. The band structure of germanium is shown in Fig. 5.1. The minimum energy between bands — the band gap energy E_{gap} — is 0.66 eV at room temperature. The band gap has temperature dependence [147] and at milliKelvin temperatures is 0.785 eV.

Electrons are fermions and hence in thermal equilibrium, the average thermal occupancy N of a particular energy level will be given by the Fermi-Dirac distribution:

$$N_{k,n} = \frac{1}{e^{(\mathcal{E}_{k,n}-\mu)/k_B T} + 1}. \quad (5.3)$$

While the value of the Fermi level in CDMS detectors is not known precisely, its exact location is not particularly important under our operating conditions of $T=45$ mK. As shown in Fig. 5.2, in which we take $\mu = 0.4$ eV, the width of the transition between occupied and unoccupied states is extremely sharp, very much smaller than E_{gap} . Thus we see that the detector is truly “frozen out”, the valence band is completely full while the conduction band is completely empty — there are no thermally generated free carriers. Since the valence band is full, no current flows even under application of a weak electric field. The addition of the field will not change the occupancy of the states in either band. The only way in which current can flow is through the generation of free carriers through interactions of external particles with the crystal, which excite some electrons from the valence band into the conduction band.

5.2.2 Dispersion Relations, Effective Mass, and the Crystal Momentum

The mathematical relationship between energy and wavevector is known as the *dispersion relation*, and is also sometimes referred to as the band structure. Due to various energy relaxation processes, which we will soon discuss, free carriers quickly occupy states near the band energy minima. Starting first in one dimension, we can Taylor expand around these

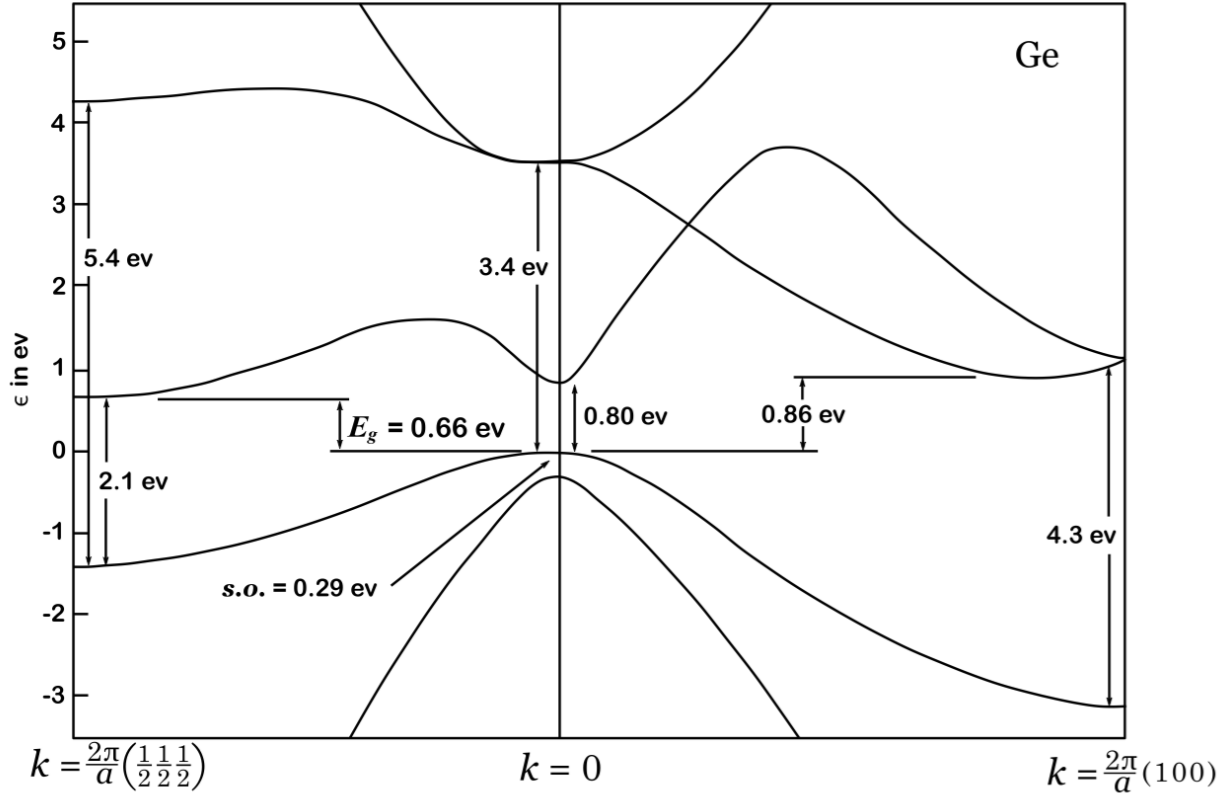


Figure 5.1: The band structure of germanium, taken from [148]. Note that the valence band maximum and conduction band minimum are not aligned in momentum space — germanium is an indirect-gap semiconductor.

minima to obtain

$$\mathcal{E}(k) = \mathcal{E}(k_0) + \frac{1}{2} \frac{\partial^2 \mathcal{E}(k)}{\partial k^2} (k - k_0)^2, \quad (5.4)$$

where the first order derivative has vanished. We see the bands take an approximately parabolic form near the band minima. This equation can be modified to include non-parabolic terms to better approximate the band structure, but are generally unimportant for CDMS applications.[149]

As has been derived in several standard texts, such as [149, 150], an electron in a state specified by Eq. 5.1 has a nonzero average velocity, given by

$$v_g(\mathbf{k}) = \frac{1}{\hbar} \nabla_{\mathbf{k}} \mathcal{E}(\mathbf{k}). \quad (5.5)$$

Under influence of an external force F (such as an electric field) the wavevector evolves

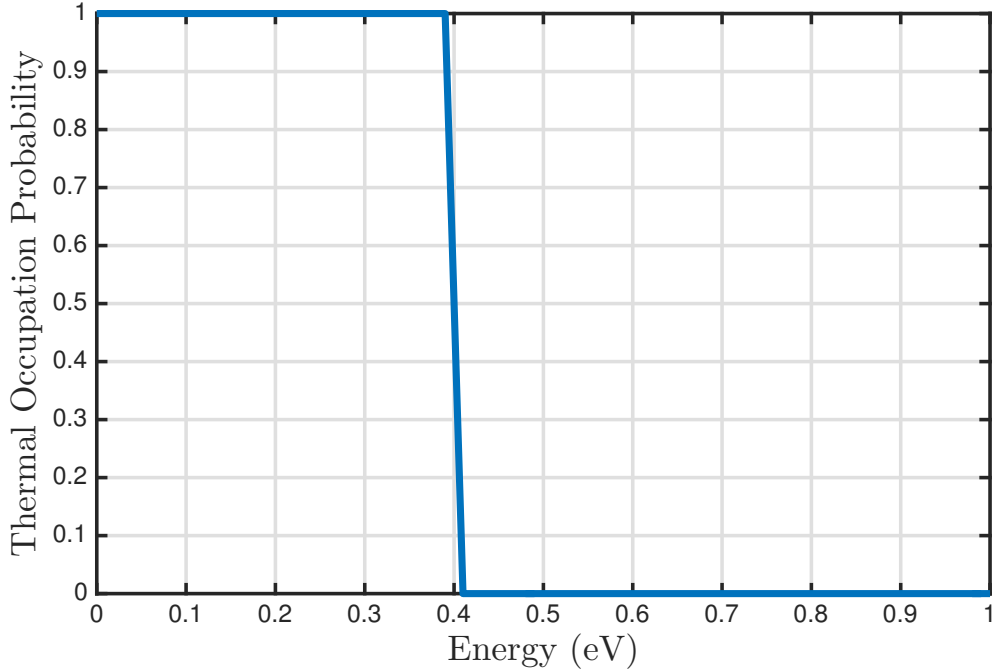


Figure 5.2: Average thermal occupation as a function of energy for $T=45$ mK. We take the chemical potential to be 0.4 eV. The sharpness of the distribution shows that the crystal is frozen out — there are no thermally-excited carriers in the conduction band, whose minimum is at 0.785 eV.

according to

$$\mathbf{F} = \hbar \dot{\mathbf{k}}. \quad (5.6)$$

In one dimension, we therefore have

$$v_g = \frac{1}{\hbar} \frac{d}{dt} \left(\frac{d\mathcal{E}}{dk} \right) = \frac{1}{\hbar} \frac{d^2 \mathcal{E}}{dk^2} \frac{dk}{dt} = \frac{1}{\hbar^2} \frac{d^2 \mathcal{E}}{dk^2} F. \quad (5.7)$$

To preserve the standard classical relationship between force, mass, and acceleration we define the *scalar effective mass* as

$$m^* = \left(\frac{1}{\hbar^2} \frac{d^2 \mathcal{E}}{dk^2} \right)^{-1}. \quad (5.8)$$

Note that the $\psi_{k,n}$ are not spatially localized and simply serve as a complete set of basis vectors for the Hilbert space. Even after an event has occurred, the density of free carriers within the crystal is extremely low. Each carrier can therefore be approximated by an independent Gaussian wavepacket, centered around some $\mathbf{k}$, such that the wavepacket's

group velocity obeys Eq. 5.5. Thus we see through Eq. 5.7 and Eq. 5.8 that the free carriers follow the classical equations of motion, but with a mass m^* rather than the free electron mass. This is known as the semiclassical approximation, valid under conditions that carrier density is low and the dimensions of the detector are much larger than the spatial extent of the wavepackets. It describes the motion of free carriers in the absence of scattering and only depends on knowledge of the dispersion relation.

These equations lead one to define the momentum of the carrier as

$$\mathbf{p} = \hbar \mathbf{k}. \quad (5.9)$$

It is important to recognize that this is **not** the true momentum of the carrier, which varies with position as seen in Eq. 5.1. Eq. 5.9 defines what is known as the *crystal momentum*. The rate of change of the carrier's true momentum is given by the total force, which includes the effects of the crystal potential. The crystal momentum is only changed by external forces and is only related to the crystal potential through the dispersion relation. The crystal momentum is a useful concept as it is generally conserved between scattering events and responds to external forces as would be expected of a classical particle with mass m^* .

5.2.3 Electron Effective Mass and Oblique Propagation

As can be seen in Fig. 5.1, the minimum of the conduction band is along the $\langle 111 \rangle$ direction of the crystal lattice, at the L -point. Due to the symmetries of the crystal, there are 8 equivalent L -points. The surfaces of constant energy are ellipses centered around each point, depicted in Fig. 5.3. Due to the non-spherical nature of the constant energy surfaces, the curvature of $\mathcal{E}(\mathbf{k})$ varies with the direction of $\mathbf{k}$. The simple scalar mass of Eq. 5.8 is generalized to an *inverse mass tensor*, given by

$$[\mathbf{m}^{*-1}]_{ij} = \frac{1}{\hbar^2} \frac{\partial^2 \mathcal{E}}{\partial k_i \partial k_j}. \quad (5.10)$$

The rate of change of the group velocity vector becomes

$$\dot{\mathbf{v}}_{\mathbf{g}} = [\mathbf{m}^{*-1}] \cdot \mathbf{F}. \quad (5.11)$$

For a particular L -valley, we can choose a coordinate frame such that the principal axis of the ellipse lies along $\hat{\mathbf{z}}$. In this frame, the dispersion relation becomes

$$\mathcal{E}(\mathbf{k}) = C_1(k_x^2 + k_y^2) + C_2 k_z^2, \quad (5.12)$$

where $\mathbf{k}$ is measured from the band minimum at $\mathbf{k}_0$. This leads to a diagonal inverse mass tensor with entries

$$[\mathbf{m}^{*-1}] = \begin{pmatrix} 1/m_t & 0 & 0 \\ 0 & 1/m_t & 0 \\ 0 & 0 & 1/m_l \end{pmatrix}. \quad (5.13)$$

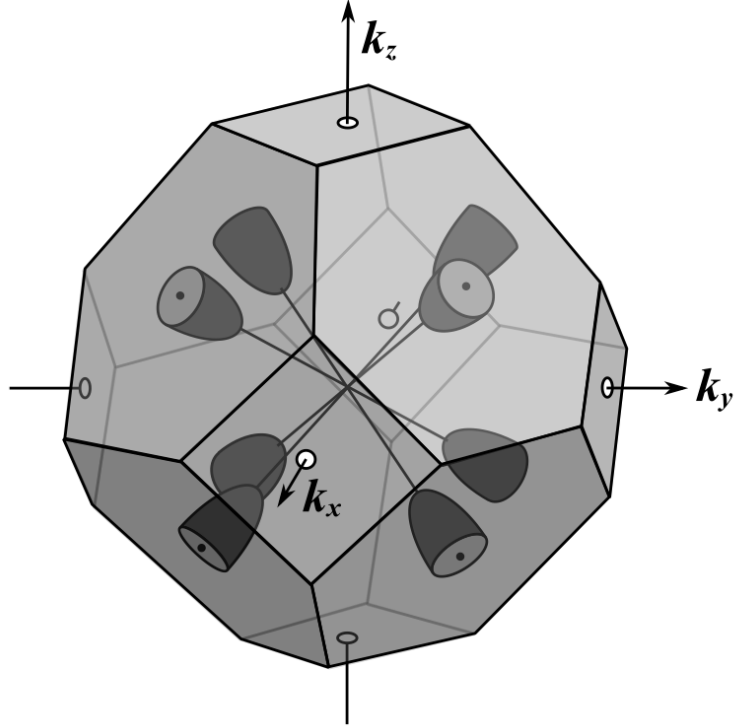


Figure 5.3: A three dimensional depiction of the surfaces of constant energy in conduction band of Ge. Note that the ellipses of constant energy are centered around the conduction band minima at the L -points of the lattice, along the $\langle 111 \rangle$ directions.[148]

Electron Parameter	Symbol	Value
Longitudinal mass	m_l	$1.570 m_e$
Transverse mass	m_t	$0.081 m_e$
Effective mass	m^*	$0.577 m_e$
Density of states mass	m_D	$0.217 m_e$

Table 5.1: Electron band parameters. m_e is the free electron mass (9.11×10^{-31} kg).

where $1/m_t = 2C_1/\hbar^2$ is the transverse mass and $1/m_l = 2C_2/\hbar^2$ is the longitudinal mass (see Table 5.1).

Recall that all standard CDMS and EDELWEISS detectors are aligned such that the longitudinal axis ($\hat{\mathbf{z}}$) of the *detector* is aligned with the $\langle 001 \rangle$ axis of the crystal lattice. We must apply a rotation to Eq. 5.13 in order to use the standard detector coordinate system. Upon performing the rotation, which will be unique for each valley, the inverse mass tensor will have off-diagonal elements. The external bias applied to the detector electrodes creates an electric field and the electrons feel a force $\mathbf{F} = -eE\hat{\mathbf{z}}$. We see from Eq. 5.11 that the acceleration is **not parallel** to the electric field due to the off-diagonal elements of the inverse mass tensor.

Electrons in the conduction band act very much like a bad grocery store shopping cart — they move at an angle relative to the direction in which you push them. This phenomena, called *oblique propagation*, is unique to the low temperature and low electric field operating conditions of CDMS detectors. While ellipsoidal anisotropies also exist in silicon, the conduction band minima occur in the X -valleys, whose $\hat{\mathbf{z}}$ -axis is aligned with the standard detector coordinate system, and the longitudinal and transverse masses are similar in magnitude. The high ratio of ~ 16 between the longitudinal and transverse mass makes this anisotropy particularly important in germanium. Oblique propagation of electrons has been imaged directly by Robert Moffatt[148, 151], and causes an increased sharing of the ionization energy between inner and outer channels for the electron-collecting side of the detector, as can be seen in Fig. 5.4. At higher fields, the effects of oblique propagation are greatly reduced due to large amounts of intervalley scattering, which causes components of the drift velocity which are not parallel to the electric field to cancel.

5.2.4 Hole Effective Mass

Throughout this dissertation, we have spoken of holes as if they were independent particles from the electrons without justification. Intuitively, we understand that a hole is an unoccupied state within the valence band that has been created by exciting an electron into the conduction band and that this vacancy can be treated as a positively-charged particle. To understand why this is a valid representation, first consider an unfilled band in which we can label the states as either being occupied or unoccupied. The instantaneous current density is given by

$$\mathbf{j} = (-e) \int_{\text{occupied}} \frac{d\mathbf{k}}{4\pi^3} \mathbf{v}_{\mathbf{g}}(\mathbf{k}). \quad (5.14)$$

We know that if the band were filled, the current density would be zero. In this case, Eq. 5.14 becomes

$$0 = -e \left(\int_{\text{occupied}} \frac{d\mathbf{k}}{4\pi^3} \mathbf{v}_{\mathbf{g}}(\mathbf{k}) + \int_{\text{unoccupied}} \frac{d\mathbf{k}}{4\pi^3} \mathbf{v}_{\mathbf{g}}(\mathbf{k}) \right), \quad (5.15)$$

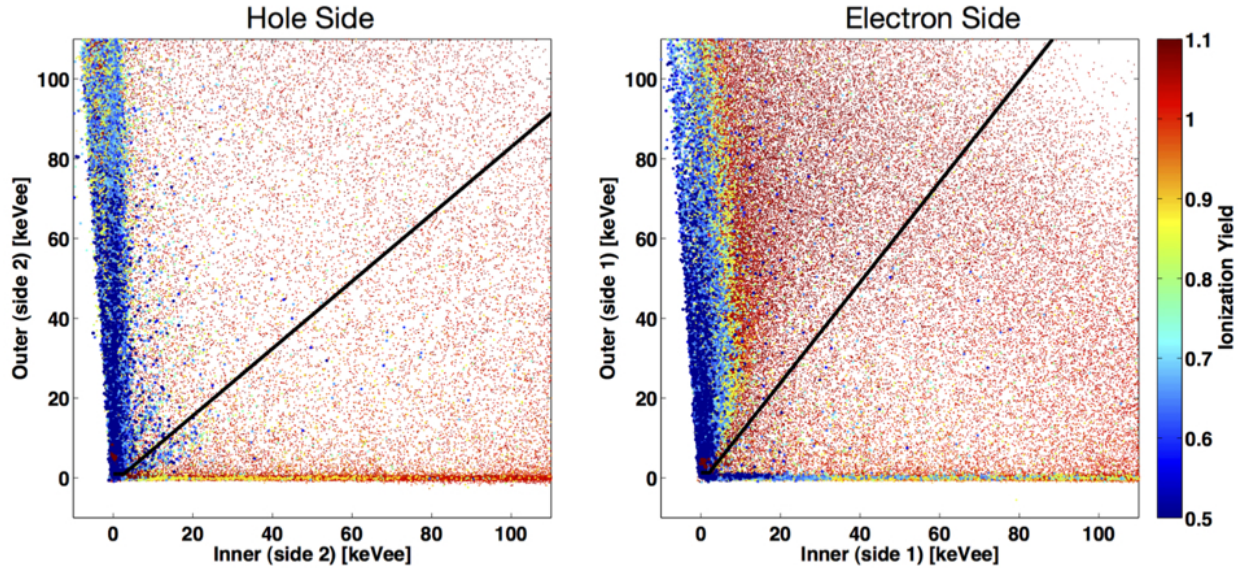


Figure 5.4: Distribution of energy between inner and outer charge channels for the hole-collecting (side 1) and electron-collecting (side 2) surfaces of a CDMS iZIP detector. The greater fraction of shared events on the electron-collecting surface is due to oblique propagation.

where we have still labelled the states by whether they were occupied or unoccupied in the unfilled case. Thus, for the unfilled band, we have

$$\mathbf{j} = (-e) \int_{\text{occupied}} \frac{d\mathbf{k}}{4\pi^3} \mathbf{v}_{\mathbf{g}}(\mathbf{k}) = (+e) \int_{\text{unoccupied}} \frac{d\mathbf{k}}{4\pi^3} \mathbf{v}_{\mathbf{g}}(\mathbf{k}). \quad (5.16)$$

In an unfilled band, the instantaneous current density produced by the negatively charged electrons in occupied states is equivalent to the current density produced by *positively* charged particles in *unoccupied* states — what we call holes. As derived in [150], holes will move in the opposite direction of electrons under influence of an electric field. After a recoil event, since the number of unoccupied states in the valence band will be small, it makes more sense to speak of hole transport. In general, since the dispersion relation in the valence band is different than that of the conduction band, holes and electrons will have different effective masses.

There are 3 separate band structure minima for holes in the valence band, as seen in Fig. 5.5. The heavy and light bands are degenerate at $k = 0$. Due to their proximity to one another, the dispersion relations within these bands are quite complicated. They are given

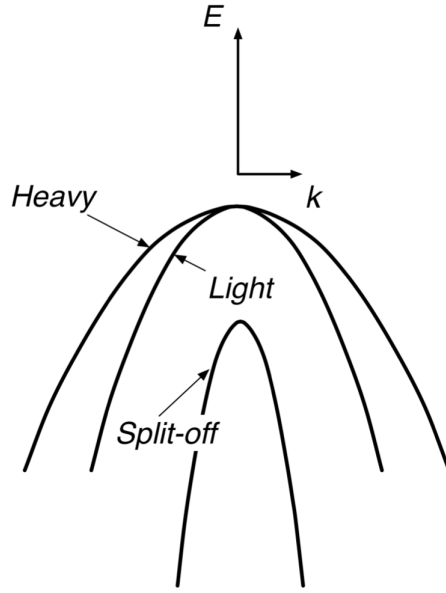


Figure 5.5: The band structure of holes in germanium.

by[152]

$$\begin{aligned}\mathcal{E}_h(\mathbf{k}) &= \frac{-\hbar^2}{2m_e} \left[Ak^2 - \sqrt{B^2k^4 + C^2(k_x^2k_y^2 + k_y^2k_z^2 + k_z^2k_x^2)} \right], \\ \mathcal{E}_l(\mathbf{k}) &= \frac{-\hbar^2}{2m_e} \left[Ak^2 + \sqrt{B^2k^4 + C^2(k_x^2k_y^2 + k_y^2k_z^2 + k_z^2k_x^2)} \right].\end{aligned}\tag{5.17}$$

The split-off band also has a minimum at $\mathbf{k} = 0$, however it is shifted in energy from the light and heavy bands. The split-off band is typically ignored, as the holes spend most of their time in the light and heavy bands. Due to the complicated nature of the dispersion relations, all elements of the inverse mass tensor for light band and heavy band holes depend on mixed $\mathbf{k}$ -vector components. To avoid this complication, the effective mass for each band is typically approximated as a scalar.[146, 153, 154] Typical holes parameters for germanium can be found in Table 5.2.

5.2.5 Conductivity Mass and Density of States Mass

When performing calculations for processes which occur within an anisotropic valley or band, there are two different definitions of mass which are used. The *conductivity mass* is the mass which enters into kinematic calculations, and represents the typical value of the mass tensor during the free-flight period. We refer to this mass as m^* and it is generally taken as the

Hole Parameter	Symbol	Value
Valence band parameter	A	13.38
Valence band parameter	B	8.48
Valence band parameter	C	13.14
Heavy band scalar mass	m_h	$0.73m_e$
Light band scalar mass	m_l	$0.25m_e$

Table 5.2: Band structure and effective mass parameters for holes in Ge.[153]

harmonic mean of the different masses. As an example, for electrons in an L -valley we have

$$m^* = \frac{1}{3}(m_l + 2m_t). \quad (5.18)$$

The *density of states mass* is the mass which must be used when determining the number of states available to an isotropic carrier. In this case, we take the geometric mean. For electrons in an L -valley, for example, we have

$$m_D = (m_l m_t^2)^{1/3}. \quad (5.19)$$

When performing calculations, care must be taken to ensure the correct expression for mass is being used for the various portions of the final expression.

5.3 Bulk Transport

5.3.1 Scattering

There are a variety of possible scattering processes which cause the trajectory of the charge carriers to deviate from Eq. 5.11. The major categories and subcategories of processes which can influence bulk transport are shown in Fig. 5.6. Scattering occurs due to perturbations in the lattice potential which cause the charge carrier to undergo quantum-mechanical transitions, taking it from state $\mathbf{k}$ to state $\mathbf{k}'$. When we speak of scattering, the final state remains in the conduction/valence band and the carrier is still free to propagate — carrier capture/trapping is considered separately and discussed in a later section.

For a scattering potential U_S , Fermi's golden rule (actually developed by Dirac[155]) states that the scattering rate from $\mathbf{k}$ to $\mathbf{k}'$ is given by

$$\Gamma(\mathbf{k}, \mathbf{k}') = \frac{2\pi}{\hbar} \left| \left(\int \frac{d\mathbf{r}}{\Omega} \psi_{k'} U_S \psi_k \right) \right|^2 \delta(\mathcal{E}(\mathbf{k}') - \mathcal{E}(\mathbf{k})), \quad (5.20)$$

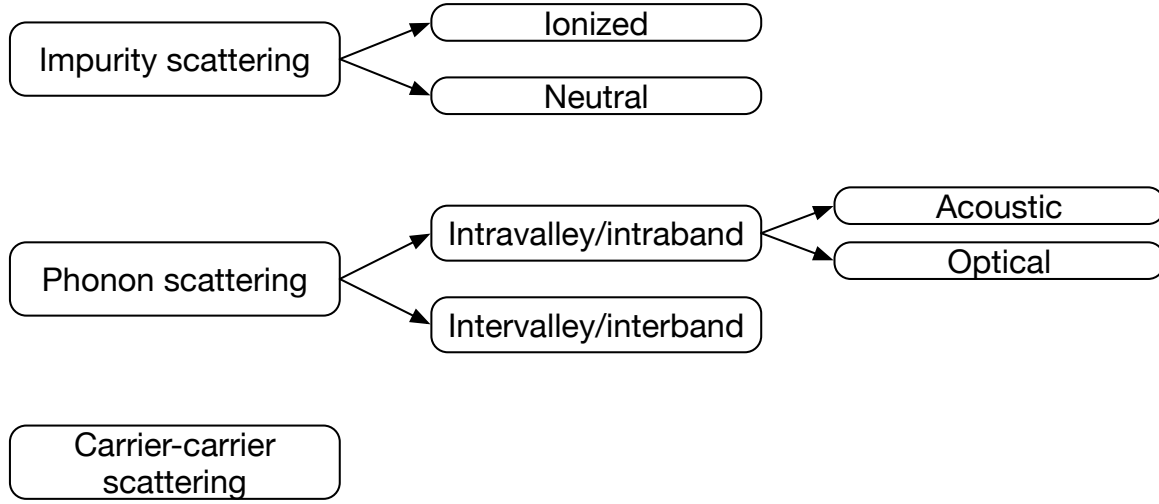


Figure 5.6: The various categories of scattering processes which can influence the bulk transport of charge carriers in sub-Kelvin germanium.

where Ω is the normalization volume. In general, these *differential* rates will not be isotropic and vary depending on the direction of $\mathbf{k}'$ relative to $\mathbf{k}$. The differential rates can be used to create an anisotropic Monte Carlo simulation to predict the transport properties of the charge carriers, as has been done by Sundqvist[146].

From the differential rate, we can define a *total* scattering rate for initial state $\mathbf{k}$, which is more useful for judging the relative importance of different processes. The total rate is given by

$$\Gamma(\mathbf{k}) = \frac{\Omega}{2\pi^3} \int \Gamma(\mathbf{k}, \mathbf{k}') d\mathbf{k}'. \quad (5.21)$$

Each scattering processes will have a particular form for U_S and the energy-conserving δ -function. We will now briefly discuss each process in more detail — a deeper discussion can be found in [146].

Carrier-carrier scattering

Recoil events produce numerous electron-hole pairs which can interact with one another through the Coulomb force. As our detectors are frozen out, we do not have a thermal population of free carriers and the free carrier density generated by a recoil event is relatively low. Under these conditions, we are safe in neglecting carrier-carrier scattering as simulation has indicated it only influences the carrier trajectories for the first ~ 5 ns or so of transport, much less than the typical $\sim 1 \mu\text{s}$ drift time.

Impurity Scattering

CDMS detectors have a net shallow impurity concentration of $\sim 10^{10}$ per cm^3 . These impurities are localized scattering centers for free carriers. As the impurities are much more massive than the free carriers, these processes are generally assumed to be elastic. While the momentum of the carriers will be randomized by the scattering process, their energy is conserved.

Ionized impurities will produce unscreened Coulomb potentials — again due to the lack of a thermal population of carriers — which are felt by the free charge carriers produced by a recoil event. Ionized impurity scattering under these conditions is given by the Conwell-Weisskopf formalism.[156] This leads to a total scattering rate of

$$\Gamma_{ion}(\mathbf{k}) = \frac{\pi N_I}{\sqrt{2m^* \mathcal{E}(\mathbf{k})}^3} \left(\frac{Ze^2}{4\pi\kappa\epsilon_0} \right)^2 \log \left[1 + \left(\frac{Ze^2}{4\pi\kappa\epsilon_0} \frac{N_I^{1/3}}{\mathcal{E}(\mathbf{k})} \right)^{-2} \right], \quad (5.22)$$

where N_I is the number density of the ionized impurities, Z the atomic number, and m^* the appropriate scalar effective mass. Due to the low impurity concentration of CDMS detectors, ionized impurity scattering only plays a role at fields well below typical CDMS bias conditions where the average carrier energy is very small.

Neutral impurities also represent possible scattering centers and are typically represented as a square well potential. Sclar developed an advanced formalism for treating neutral impurity scattering which depends on a *threshold energy* $\mathcal{E}_T$, typically assumed to be ~ 0.5 meV.[146, 157] The total scattering rate is given by

$$\Gamma_{neut}(\mathbf{k}) = N_I \left(\sqrt{\frac{2\mathcal{E}(\mathbf{k})}{m^*}} \right) \left(\frac{4\pi\hbar^2}{2m^* \sqrt{\mathcal{E}(\mathbf{k}) + \mathcal{E}_T}} \right), \quad (5.23)$$

where N_I is the number density of neutral impurities. Assuming $\sim 10^{10}$ impurities per cm^3 , one finds a mean free path of over a meter. Neutral impurity scattering is unimportant in high purity germanium.

Phonon Scattering

The final and most important type of scattering process for sub-Kelvin germanium we will discuss is phonon scattering. Phonons are the quantized vibrational modes of the crystal lattice. If one imagines the atoms of the crystal lattice as being connected to each other through springs which tend to hold them in their equilibrium position, it is easy to see how the displacement of an atom can propagate throughout the crystal. Phonons are treated in essentially all standard solid state physics and semiconductor texts, such as [149, 150, 158]. As wavefunctions of the charge carriers depend upon the periodic potential of the lattice ions, it is easy to predict there will be a strong coupling between electrons and phonons.

Phonons come in two types — *optical* and *acoustic*, and are labelled by their wavevector $\mathbf{q}$, analogous to the crystal momentum $\mathbf{k}$ of the charge carriers. Each type of phonon has

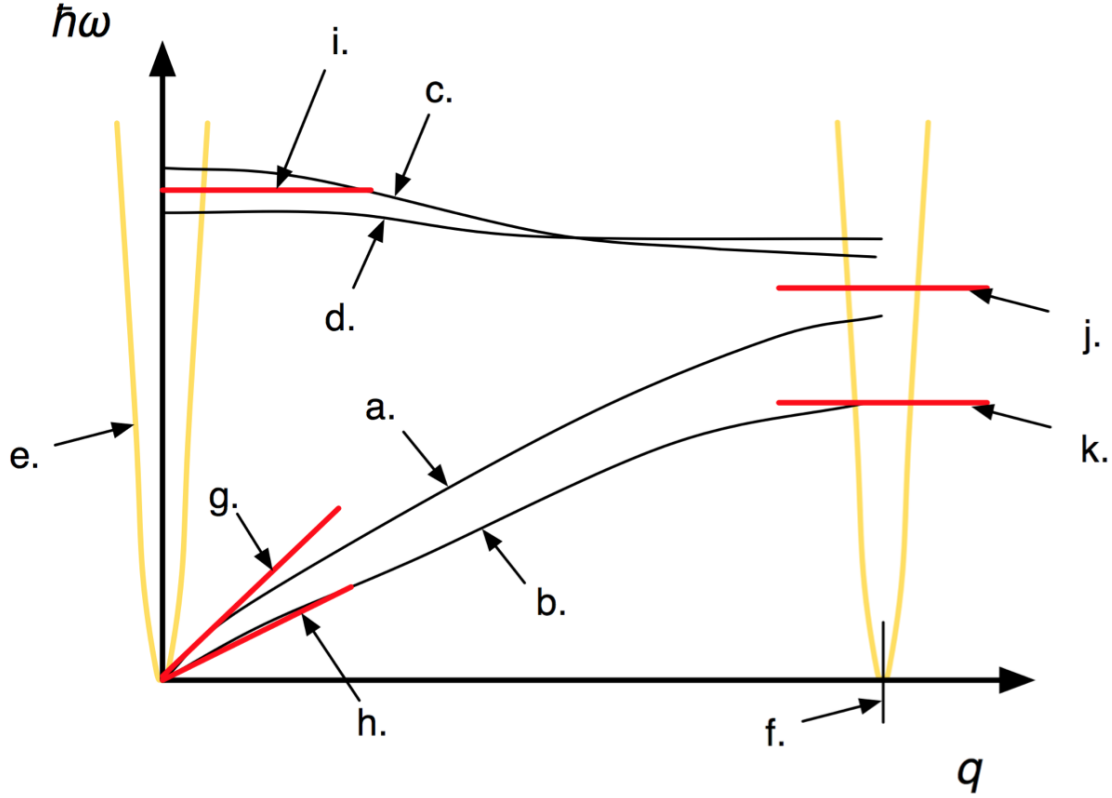


Figure 5.7: The phonon dispersion relation for germanium, idealized for a longitudinal branch and a single transverse branch for both acoustic and optical phonons. **a.** longitudinal acoustic branch (LA) **b.** transverse acoustic branch (TA) **c.** longitudinal optical branch (LO) **d.** transverse optical branch (TO) **e.** a superimposed dispersion relation for parabolic electrons (plotting the corresponding $\mathbf{k}$ -vector for electrons — not to scale) **f.** the Brillouin Zone boundary **g.** slope for the longitudinal speed of sound **h.** slope for the transverse speed of sound **i.** the optical phonon energy, ω_O **j.** the intervalley phonon energy, ω_i **k.** the transverse acoustic intervalley phonon energy, ω_{it} . Taken from [146]. See Table 5.3 for values.

three different *branches* related to their polarization — two transverse branches and one longitudinal branch. The dispersion relations for the various types of phonons are shown in Fig. 5.7. For clarity, only one transverse branch is shown for each type of phonon.

Acoustic phonons are identified as having zero energy ($\hbar\omega = 0$) at zero wavevector ($\mathbf{q} = 0$). The dispersion relation is linear near this point, and its slope is the *speed of sound*, v_s . The acoustic dispersion relation is thus approximated as

$$\mathcal{E}_{ac}(\mathbf{q}) = v_s \cdot |\mathbf{q}|. \quad (5.24)$$

Phonon Parameter	Symbol	Value
Density of germanium	ρ	5320 kg/m ³
Longitudinal speed of sound	v_{sl}	5.4×10^5 cm/s
Transverse speed of sound	v_{st}	3.2×10^5 cm/s
Acoustic deformation potential	Ξ_A	11 eV
Optical deformation potential	Ξ_O	5.5×10^8 eV/cm
Optical phonon energy	$\hbar\omega_O$	37 meV
Intervalley phonon energy	$\hbar\omega_{it}$	10 meV

Table 5.3: Phonon parameters.[146, 153]

Each acoustic phonon branch has a different speed of sound, with the longitudinal polarization being the fastest. For simplicity, the transverse branches are often combined into a single effective branch, and often even the transverse and longitudinal branches will be combined into a single effective acoustic branch, using an average speed of sound. The speed of sound contains slight anisotropies, however they are generally ignored.

Optical phonon have a much different dispersion relation, which remains roughly constant near $\mathbf{q} = 0$. While there are three branches of optical phonons, their dispersion relations are all very similar and they are combined into a single effect branch, with a *constant* dispersion relationship given by

$$\mathcal{E}_{op}(\mathbf{q}) = \hbar\omega_O, \quad (5.25)$$

where ω_O is the constant average optical phonon frequency. Note that optical phonons require a minimum energy unlike acoustic phonons, however there is no dependence on the momentum of the optical phonon.

Phonon scattering processes can cause charge carriers to transition between states within a given valley/band, known as *intravalley scattering*. This process can occur for either acoustic or optical phonons as long as energy and momentum are conserved. At low fields, acoustic phonons dominate the total intravalley scattering rate.

Phonon scattering can also cause a charge carrier to transition from one band or valley into another, which we call *intervalley scattering*. Particularly for the L-valleys of the germanium conduction band, these processes generally result in large changes in momentum. Both acoustic and optical processes can contribute to intervalley transitions and it is unclear as to which plays the dominant role. Intervalley scattering is usually approximated as an optical phonon process, with some average intervalley scattering energy $\hbar\omega_i$. At large wavevectors, the linear approximation to the acoustic phonon dispersion relations no longer hold true. The slowest transverse acoustic branch has a substantially smaller energy at large wavevector, however these phonons can also contribute to intervalley processes and is sometimes modeled as a separate intervalley process with energy $\hbar\omega_{it}$.

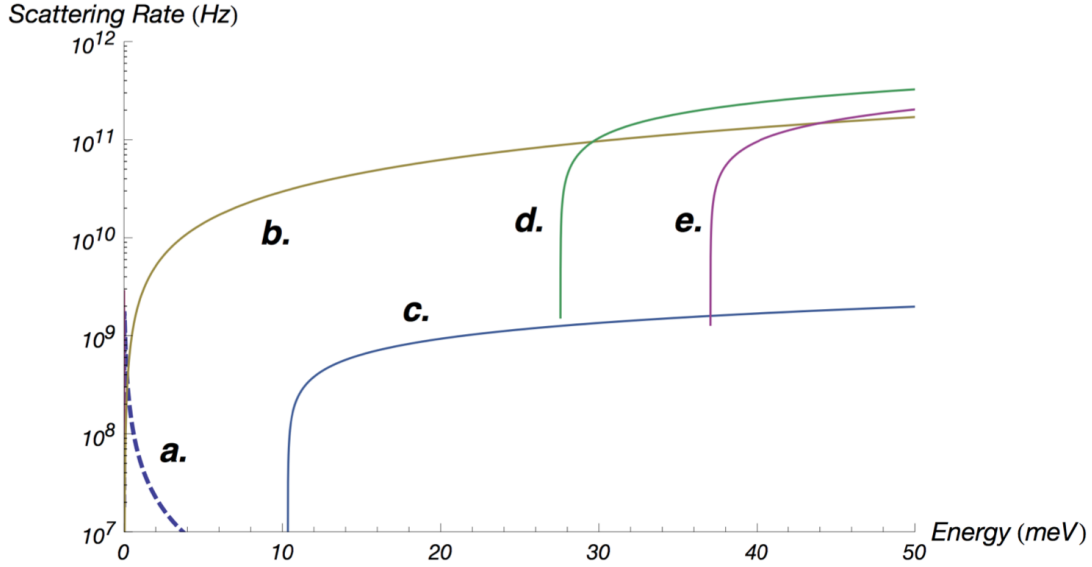


Figure 5.8: The total phonon scattering rates for electrons, at $T=40\text{mK}$, under isotropic assumptions.[146] **a.** Conwell-Weisskopf ionized impurity scattering rate at $N_I = 10^{10}\text{cm}^{-3}$ **b.** acoustic phonon emission **c.** slow-transverse intervalley phonon emission **d.** intervalley phonon emission **e.** optical phonon emission

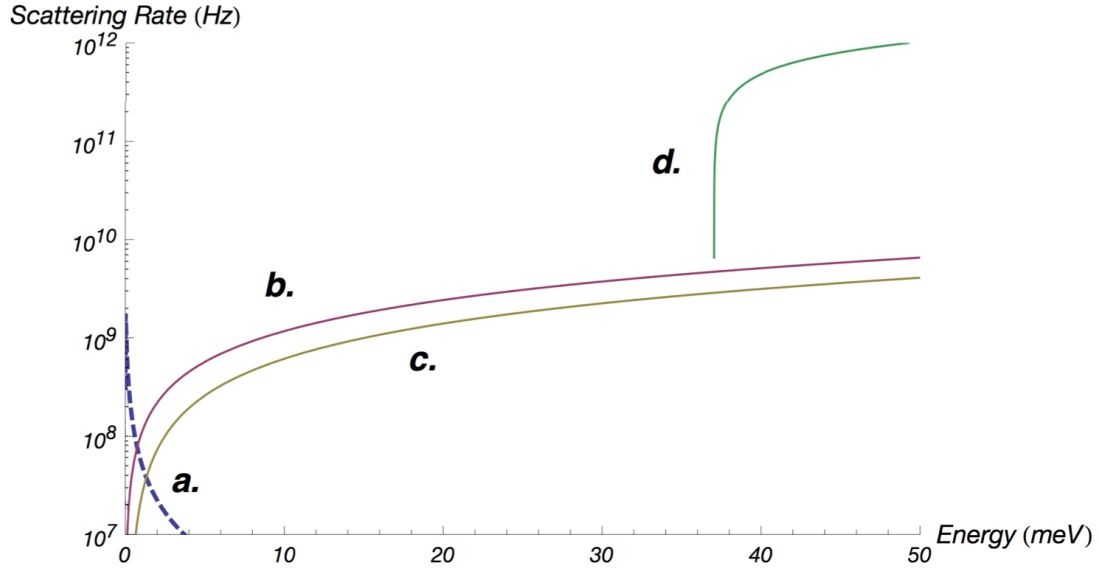


Figure 5.9: Total phonon scattering rates for holes, at $T=40\text{mK}$, under isotropic assumptions.[146] **a.** Conwell-Weisskopf ionized impurity scattering rate at $N_I = 10^{10}\text{cm}^{-3}$ **b.** acoustic phonon emission for heavy-to-heavy band transitions **c.** acoustic phonon emission for heavy-to-light band transitions **d.** optical phonon emission for heavy-to-heavy band transitions

Charge carrier scattering can result from either the emission or absorption of phonons. Charge carriers in germanium typically reach energies of a few meV under CDMS operating conditions. At $T=40\text{mK}$, the thermal occupancy of phonon modes with enough energy and momentum to significantly interact with the carriers is essentially zero. Phonon *absorption* does not play an important role in these detectors. Note, however, that the quantum *zero-point motion* of the lattice ions — which we recognize as coming from the $\frac{1}{2}\hbar\omega$ energy minimum of the quantum harmonic oscillator — causes a substantial amount of phonon emission, similar to the spontaneous emission of radiation. A thorough discussion of the kinematics of both phonon emission and absorption by charge carriers can be found in [146].

The electron-phonon interaction is parameterized by a *deformation potential* which is related to the physical displacement of the lattice ions.[159] Assuming the initial state and final state effective mass are the same, the total zero-point scattering rate for due to acoustic phonon emission is given by

$$\Gamma_A(\mathbf{k}) = \frac{m^*\Xi_A^2}{3\pi^2\rho v_s\hbar^5k}(\hbar k - m^*v_s)^3, \quad (5.26)$$

where Ξ_A is the acoustic deformation potential and v_s is the average speed of sound for the three branches.[160]. v_s is given by

$$v_s = \frac{v_{sl} + 2v_{st}}{3}. \quad (5.27)$$

For optical (and intervalley) phonons, the total isotropic zero-point scattering rate is given by

$$\Gamma_O(\mathbf{k}) = \frac{\Xi_O^2}{\rho\omega_O} \frac{2m^*\sqrt{\mathcal{E}(k) - \hbar\omega_O}}{4\pi\hbar^3}, \quad (5.28)$$

where Ξ_O is the optical deformation potential. Isotropic deformation potentials for electrons and holes can be found in Table 5.3. The interested reader is directed to [161–164] for more information regarding phonon scattering and deformation potentials.

The total scattering rates for several processes are shown in Fig. 5.8 for electrons and Fig. 5.9 for holes. With typical charge carrier energies of a few to tens of meV, we see impurity scattering plays a very small role and the effective total rate is completely dominated by phonon scattering processes. Charge transport in CDMS detectors is dictated by the interactions with the zero-point motions of the lattice ions, as expected for such high purity crystals operating at low operating temperature. Note the threshold energies for intervalley and optical phonon emission, which become increasingly important as the electric field is increased. At high field, the effects of anisotropy are reduced due to the high rate of intervalley scattering. We recognize that these **spontaneously emitted phonons are the Luke phonons** generated during transport.

5.3.2 Fundamental Distributions and Averages

The average scattering rate is clearly very high in comparison to the typical drift time of $1\ \mu\text{s}$ in CDMS detectors. In our semiclassical picture, electrons and holes gain energy through

numerous brief acceleration periods described by Eq. 5.6 and Eq. 5.7. Each acceleration period is terminated due to a random scattering process which instantaneously changes the energy and momentum, but not the position, of the charge carrier. As the random initial momentum after a scattering event is not parallel with the electric field, the trajectory of the charge carrier during the acceleration (or free-flight) period will only have a small preference to move towards the appropriate ionization-collecting electrode. Over a large number of scatters, this small preference results in the average movement of charge carriers towards the electrodes.

The increase in the total phonon scattering rate as the carrier energy increases causes an equilibrium to be reached between the energy gained through free-flight and the energy lost through phonon emission. The properties of the charge carriers produced by a recoil event can therefore be described by equilibrium distributions whose width comes from the variations due to random scattering processes. These distributions will have well-defined averages which can be compared to experimental measurements, as is done in later chapters. Note that the equilibrium condition here is strictly **not thermal equilibrium**, but an electric field-dependent equilibrium which is strongly influenced by the charge carriers properties and phonon scattering rates.

There are two basic distributions which are used to describe the properties of the electrons and holes — the energy distribution and the velocity distribution. Obviously, the average carrier energy $\mathcal{E}$ can be found from the energy distribution. Since the average energy is usually much greater than the temperature of the crystal lattice, transport under these conditions is sometimes referred to as *hot carrier transport*.

There are three different velocities derived from the velocity distribution:

Drift velocity The drift velocity $\mathbf{v}_d$ is simply the average of the velocity distribution. It represents the average velocity at which the charge carriers move. Note that for electrons, each of the 4 L-valleys has equal probability of being populated and the velocity components perpendicular to the electric field average away, so we are really speaking of the average drift velocity in the $\langle 100 \rangle$ direction.

Random/thermal velocity The random velocity v_{th} , also called the thermal velocity, is given by the standard deviation of the velocity distribution. The fluctuations in carrier velocity are generally much greater than the drift velocity. In a later section we will see that the random/thermal velocity can be used to define an effective carrier temperature, hence the name.

Total velocity The velocity distribution is often approximated by a displaced Maxwellian, such that we have $\mathbf{v} = \mathbf{v}_d + \mathbf{r} \cdot v_{th}$, where r is a randomly sampled direction from a Gaussian distribution with a standard deviation of unity. The total velocity is defined as $v_{tot} = \langle \sqrt{\mathbf{v} \cdot \mathbf{v}} \rangle$. Since the velocity fluctuations are much bigger than the drift velocity, the total velocity essentially *is* the thermal velocity, and the approximation $v_{tot} = v_{th}$ is often used.

There are two important timescales associated with these equilibrium distributions:

Energy relaxation time In the zero-point regime, we define the energy relaxation time as

$$\tau_{\mathcal{E}} = \frac{\mathcal{E}}{eEv_d}. \quad (5.29)$$

This can be thought of as a time constant determining how long it takes a carrier's energy to come to equilibrium with the electric field.[149]

Momentum relaxation time We define the momentum relaxation time as

$$\tau_p = \frac{m^*v_d}{eE}. \quad (5.30)$$

This time constant determines how long it takes a carrier's momentum to become randomized, reaching equilibrium with the electric field.[149]

5.3.3 Monte Carlo Method and the Results of Sundqvist

Analytical determination of the basic distributions and averages is essentially impossible due to the numerous bands and scattering processes which play a role in charge transport. These distributions are instead found numerically through Monte Carlo simulations, which use random number generators to simulate the probabilistic semiclassical transport and scattering processes for a large number of carriers. A good introduction to Monte Carlo simulations can be found in [149]. With knowledge of the scattering rates, the basic Monte Carlo method proceeds as follows:

1. A specific electric field is chosen for study.
2. After initialization of a carrier, a random free-flight timestep is determined using the total combined scattering rate of all processes, which we label Γ_0 . Since the total true rate depends on carrier energy, a fictitious *self-scattering* process is introduced such that the total rate always remains constant. A self-scattering process does not change the carrier energy or momentum. A uniformly-sampled random number between 0 and 1 is chosen, and the timestep determined by

$$t_c = \frac{\ln(r)}{\Gamma_0}. \quad (5.31)$$

Free-flight is assumed to occur for a time t_c and the position, momentum, and energy of the charge carrier are updated.

3. Free-flight is interrupted by a random scattering event. The charge carrier's energy and momentum are used to determine the instantaneous rate of all scattering processes under consideration. A random number is chosen to determine which scattering process occurs, based on the ratios of the individual rates to the total rate.

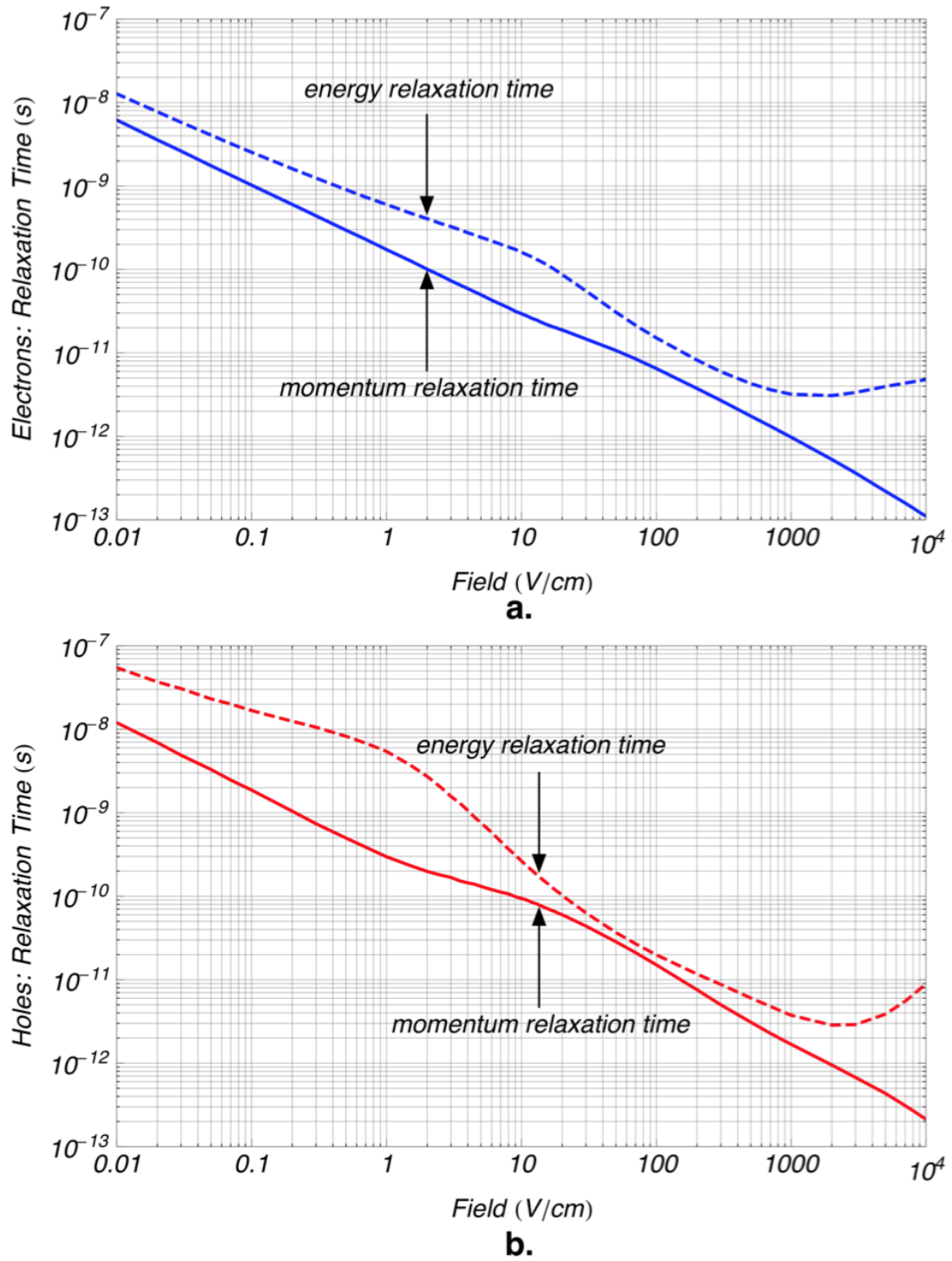


Figure 5.10: The energy and momentum relaxation times as a function of electric field for (a) electrons and (b) holes.[146]

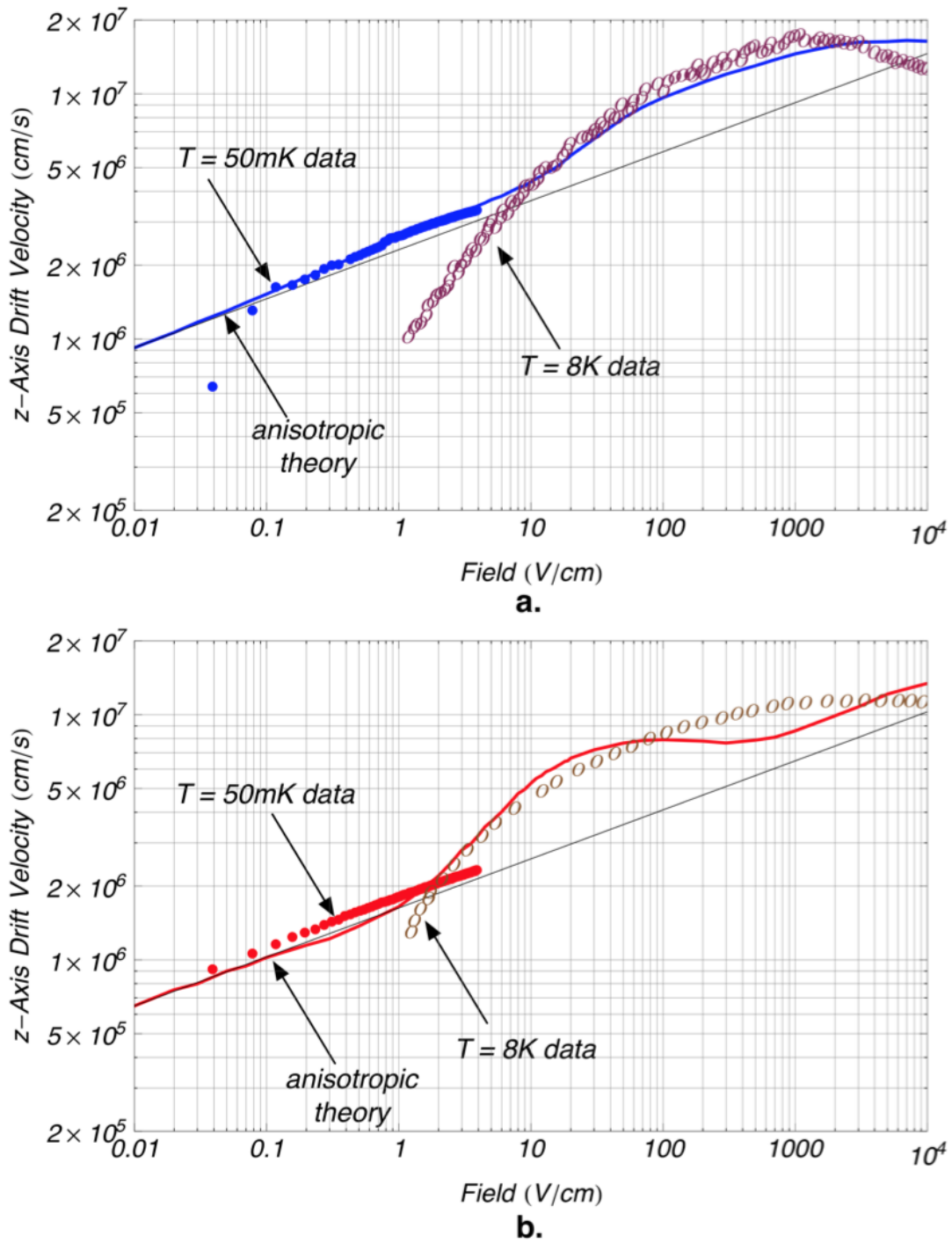


Figure 5.11: Drift velocities as a function of electric field for (a) electrons and (b) holes.[146]

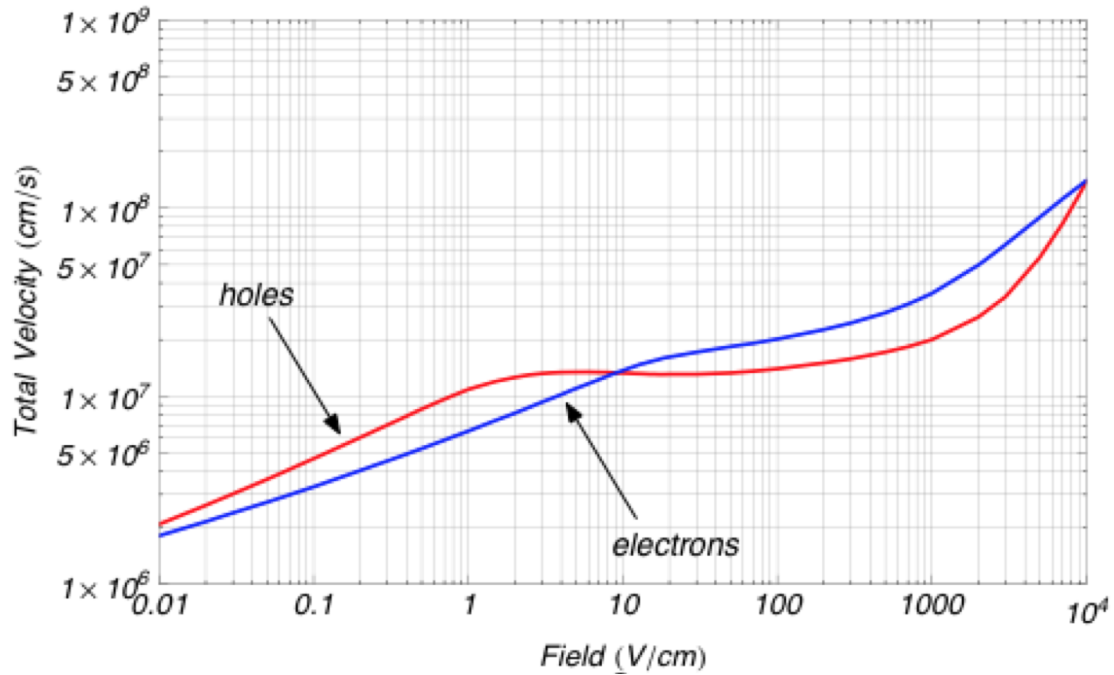


Figure 5.12: The total velocity as a function of electric field for electrons and holes.[146]

4. The post-scatter energy and momentum of the charge carrier are determined by selection of at least two more random numbers. Appropriate energy and momentum conservation is performed for the chosen type of scatter event. Self-scatter events make no changes to the charge carrier state and are simply a computational penalty for simplifying the timestep determination. If an intervalley scatter occurs, the effective mass and dispersion relation for the carrier is updated accordingly.
5. All desired information about the charge carrier and scattering event properties are recorded.
6. A new timestep is determined and the free-flight and scattering routines repeat until a satisfactory amount of scattering events have occurred.
7. The recorded data is used to determine the basic distributions for the electric field under study.
8. The process is repeated for all desired electric fields.

A thorough treatment of Monte Carlo simulation in low temperature germanium was performed by Jacoboni and Reggiani in 1983[153], however only down to temperatures of 8K. Other notable simulations are those of Sundqvist[165], Cabrera[166], Aubry-Fortuna[167],

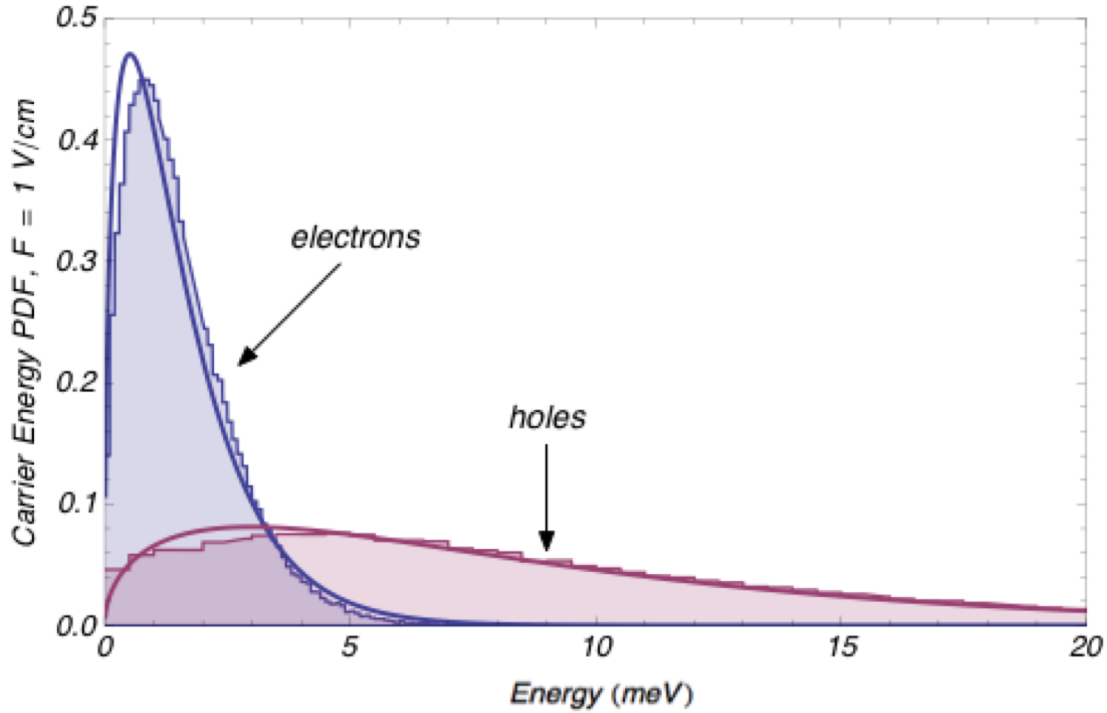


Figure 5.13: The carrier energy distribution for electrons and holes at an electric field of 1 V/cm. The distributions are well-described by displaced Maxwellians.[146]

and Brandt[168], however all of these simulations use integrated, isotropic rates. Given the importance of anisotropy in low temperature germanium, particularly for the electrons, Sundqvist developed a new approach to the Monte Carlo scattering routine using the differential anisotropic rates, as described in [146]. We present his major results here, pertinent to our previous discussion.

The energy and momentum relaxation rates are shown in Fig. 5.10. We see the relaxation times are much faster than the typical drift time of 1 μ s, indicating the carriers quickly reach equilibrium with the electric field. The drift velocities are shown in Fig. 5.11 and are in decent agreement with the experimental data. We will take a closer look at the comparison between measured and predicted drift velocities in a later chapter. The total velocities are shown in Fig. 5.12. We see that the total velocity is about 10 times greater than the drift velocity, showing how the fluctuations in instantaneous carrier velocity are much greater than the average velocity. Finally, in Fig. 5.13, we plot the carrier energy distribution for electrons and holes at a field of 1 V/cm. A log-log sweep of energy distribution as a function of field can be found in [146].

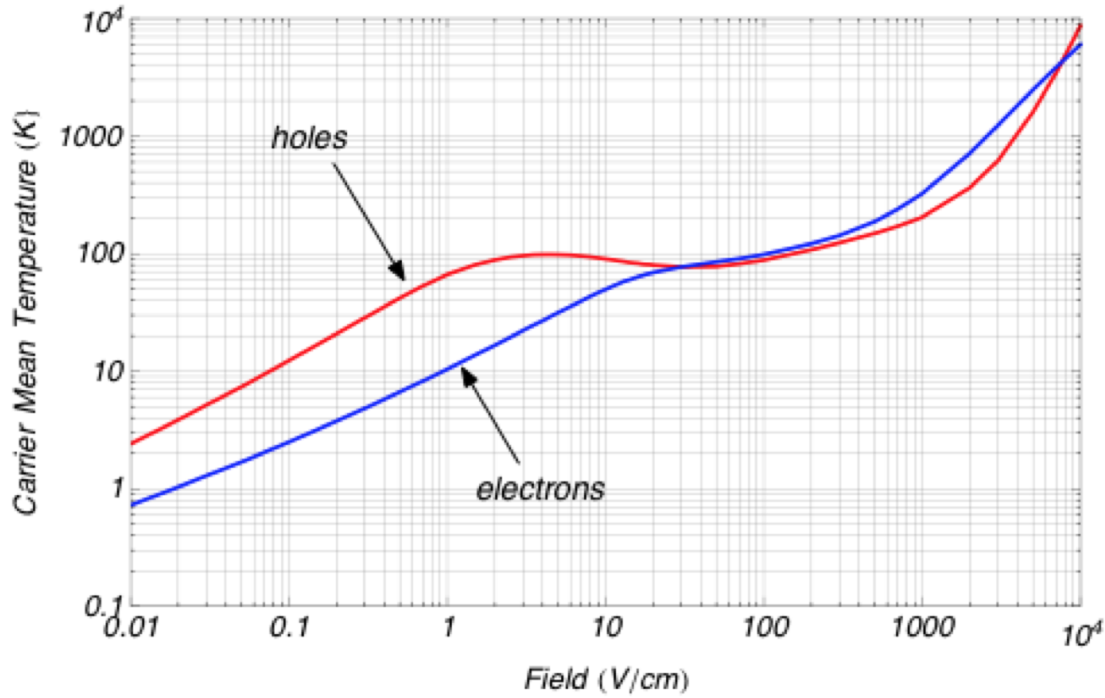


Figure 5.14: The effective carrier temperature T_C as a function of electric field for electrons and holes.[146]

5.3.4 Effective Carrier Temperature

Since the charge carriers reach an equilibrium with the applied field, it makes sense to define an effective carrier temperature. While the carrier distributions may not be perfectly Maxwellian in nature, an approximate temperature can still be formed by calculating the variance of the various degrees of freedom. Sundqvist defines the *thermal tensor* as

$$\frac{1}{2}k_b T_{ij} = \frac{\hbar}{2} \text{cov}(\delta k_i, \delta v_j). \quad (5.32)$$

The scalar effective temperature T_C is proportional to the trace of the thermal tensor. We have

$$\frac{3}{2}k_b T_C = \text{trace}\left(\frac{1}{2}k_b T_{ij}\right). \quad (5.33)$$

The effective carrier temperature as a function of electric field is shown in Fig. 5.14. We see that the carrier temperature is much greater than the lattice temperature, indicating we are in the regime of hot carrier transport. As we will see, the carrier temperature is a useful for the calculation of capture cross sections.

5.3.5 Signal Shape

A recoil event occurring at some depth z_0 within the detector will produce electrons and holes which drift in opposite directions due to the applied field. As we have seen, the carriers quickly reach equilibrium with the field and drift in the $\pm z$ -direction with drift velocity v_{de} for electrons and v_{dh} for holes. If the electric field is constant, as in the case of the planar geometry, the carrier drift velocities remain constant throughout the crystal. Ignoring the effects of trapping or impact ionization, the drift current generated by each polarity of carrier is constant and may be described by a square pulse with an amplitude equal to $n_{eh}ev_d$, where the appropriate drift velocity is chosen for electrons or holes and n_{eh} is the number of electron-hole pairs created by the event. If the detector thickness is L and we assume the holes drift towards $z = L$, the width of the hole drift current signal is $(L - z_0)/v_{dh}$ while the width of the electron signal is z_0/v_{de} . In general, these two times will not be equal due to the differences in both drift length and drift velocity.

As the typical charge amplifiers used with these detectors *integrate* the drift current, we expect a ramp-like waveform with a kink corresponding to the collection of the carrier with the shorter drift time, as shown in Fig. 5.15. If the event occurs near $z = 0$ or $z = L$, one polarity of carrier will reach its collection electrode nearly instantaneously and will not contribute to the measured drift current. In this case, the integrated waveform will be a pure ramp with no kink. This difference in pulse shape can be used to discriminate between bulk and surface events, as we did in [169] to perform the first drift velocity measurements of electrons and holes in sub-Kelvin germanium.

5.4 Capture (Trapping)

As the charge carriers drift through the germanium, they encounter impurities. While we have already discussed how impurities can serve as scattering centers, another set of processes exist in which a charge carrier becomes bound to a spatially-localized impurity state and no longer contribute to the drift current signal. The charge carriers are said to be captured, or trapped, by the impurity atoms. It was always assumed that capture to impurities was responsible for the presumed build-up of space charge observed in CDMS detectors, however the detailed physics behind this process were poorly understood. A major portion of this dissertation is dedicated to advancing the understanding of charge capture physics and their effect on ionization collection in sub-Kelvin germanium, building upon the work of Sundqvist.[146]

5.4.1 Capture Rates, Trapping Times, and Trapping Lengths

Consider a uniform distribution of a single type of impurity with number density N . Assume there is a process by which one species of charge carrier can become locally captured onto one of these impurities. We say the impurities serve as capture *centers*. Let the energy-dependent cross section for this process at a fixed electric field strength by $\sigma(E)$. For the

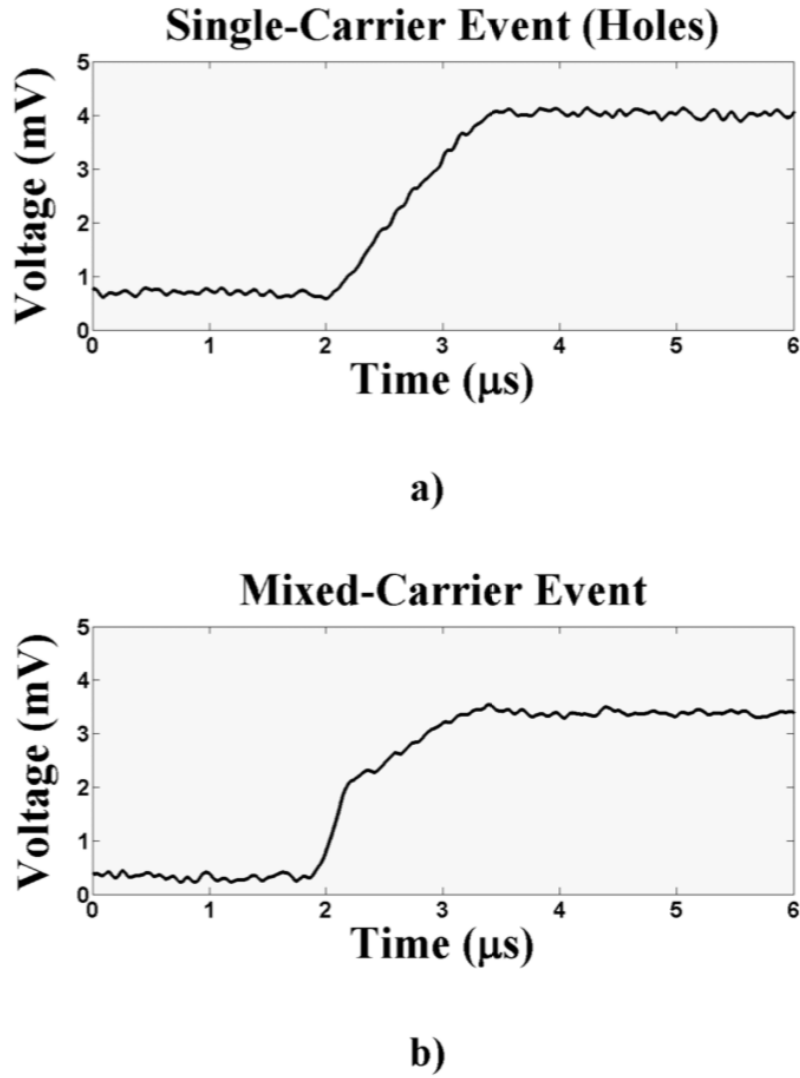


Figure 5.15: (a) A single carrier (holes) event. A single slope is observed in the integrated charge signal. (b) An event occurring within the bulk of the detector. Transport is shared between electrons and holes. The clear kink in the integrated charge during the pulse indicates the carrier with the shortest drift time has been collected.

charge carrier being captured, let $v_{tot}(E)$ be the total velocity distribution function, and let $g(E)$ be the energy distribution function. The *capture rate per unity density* represents the rate at which carriers are captured onto this center when $N = 1$ and is given by

$$C = \langle \sigma v_{tot} \rangle = \frac{\int \sigma(E) v_{tot}(E) g(E) dE}{\int g(E) dE}. \quad (5.34)$$

The *total capture cross section* is defined such that

$$C = \langle \langle \sigma \rangle \rangle \langle v_{tot} \rangle, \quad (5.35)$$

where $\langle v_{tot} \rangle$ is the expectation value of v_{tot} over the carrier energy distribution. Equating Eq. 5.34 and Eq. 5.35, we see

$$\langle \langle \sigma \rangle \rangle = \frac{\langle \sigma v_{tot} \rangle}{\langle v_{tot} \rangle}. \quad (5.36)$$

Typically, it is the total cross section which is measured by experiment and reported in the literature as there is no way to distinguish the contributions from individual portions of the underlying energy distribution. When making theoretical predictions, the carrier energy and velocity distributions from Monte Carlo simulation can be used to compute the total cross section.

The total capture rate due to this process is related to the trapping time by

$$\frac{1}{\tau_{trap}} = NC = N \langle \langle \sigma \rangle \rangle \langle v_{tot} \rangle. \quad (5.37)$$

To convert this time to a *trapping length*, which is useful to compare to the dimensions of the detector, we make use of the drift velocity. We have

$$\frac{1}{\lambda} = \frac{1}{\langle v_z \rangle \tau_{trap}} = \frac{NC}{\langle v_z \rangle} = N \langle \langle \sigma \rangle \rangle \frac{\langle v_{tot} \rangle}{\langle v_z \rangle}. \quad (5.38)$$

Here, we take $\langle v_z \rangle$ to be the average drift velocity along the $\langle 100 \rangle$ crystal axis, in accordance with the orientation of CDMS detectors.

Note the important difference between the use of the total velocity versus the drift velocity in Eq. 5.38. When calculating the capture *rate*, we multiply the cross section by the instantaneous total velocity as this determines the flux of carriers incident upon the center. Scattering processes tend to randomize the total velocity, however, and it is not directly related to a physical distance within the crystal. As the carriers rapidly come to equilibrium with the electric field, their position is well-described by the drift velocity ($z - z_0 = v_d t$) allowing a conversion to units of length to be made.

If multiple processes contribute to the capture of a single species of carrier, their inverse trapping lengths are summed to give a total effective trapping length of

$$\frac{1}{\lambda} = \sum_i \frac{1}{\lambda_i}, \quad (5.39)$$

where the λ_i are the trapping lengths of each process. The trapping times may be combined in the same manner.

5.4.2 Signal Shape

Now consider a packet of charge carriers of a single polarity moving through a slab of germanium with uniform electric field and impurity concentration. For trapping processes, the rate of change of charge carriers of number density n is given by

$$\dot{n} = -\frac{n}{\tau_{trap}}. \quad (5.40)$$

Temporarily ignoring the drifting of the charge carriers, this is easily solved to give

$$n(t) = n_0 \exp\left(-\frac{t}{\tau_{trap}}\right). \quad (5.41)$$

The charge carriers decays in an exponential fashion with a time constant equal to the trapping time.

The drift current is related to the number density of charge carriers via

$$I_d = n q v_d. \quad (5.42)$$

Since we assume uniformity of the crystal and electric field, there is no position dependence in the trapping process and we can directly substitute in Eq. 5.41. We thus obtain

$$I_d = I_0 \exp\left(-\frac{t}{\tau_{trap}}\right) = n_0 q v_d \exp\left(-\frac{t}{\tau_{trap}}\right), \quad (5.43)$$

and we see the drift current also follows the same exponential decay.

5.4.3 Cascade Coulomb Capture

Previously, we discussed how ionized impurities create Coulomb potentials which cause free charge carriers to scatter in their vicinity. While ionized impurities play an insignificant role in scattering, they strongly influence carrier *capture*. Intuitively, we recognize that attractive Coulomb potentials possess a series of localized, bound states. Charge carriers in the vicinity of a Coulomb potential can transition into the bound states, becoming trapped and neutralizing the impurity. Shallow acceptors and donors in germanium are generally hydrogen-like, with binding energies of ~ 10 meV or more.[170–173]

Due to the indirect band gap in germanium, capture due to the emission of *photons* is rare and inconsequential.[174] Photons have a difficult time conserving energy and momentum in indirect gap semiconductors due to the photon dispersion relation $E = pc$ — large amounts of energy are required for even small changes in momentum. It is again the *phonons*, whose energy increase significantly slower with momentum, which play a dominant role in carrier capture. Historically, the large capture cross sections ($\sim 10^{-12}$ cm²) in low temperature germanium observed by Koenig[175, 176] were difficult to explain by the emission of a single phonon which would cause the charge carrier to fall into the ground state of the impurity.

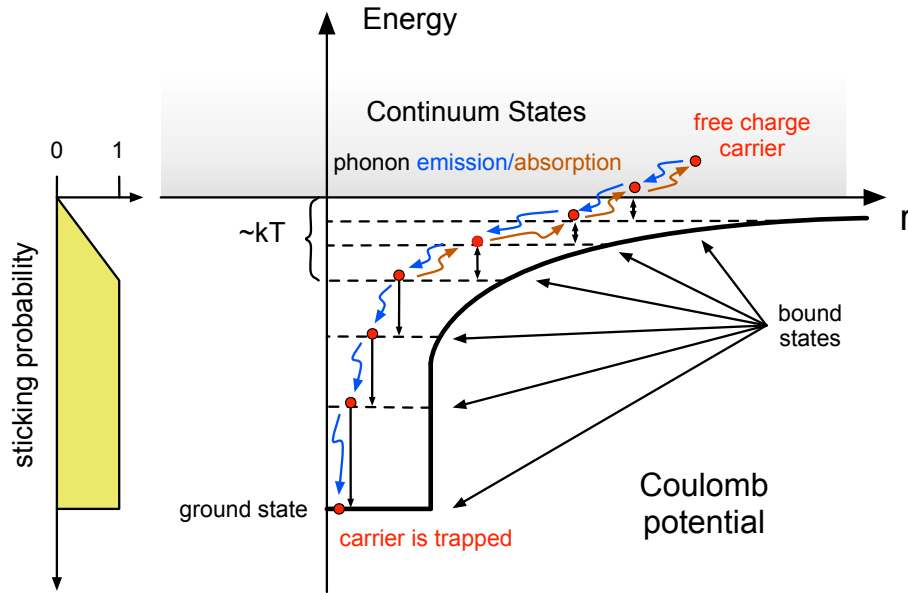


Figure 5.16: The cascade capture process.

The typical emitted phonon energy is a few times $\frac{1}{2}m^*v_s^2$ — much smaller than the impurity binding energy!

Melvin Lax realized that carrier capture in germanium (and other semiconductors) was not caused by the emission of a single phonon, but by a *cascade* of many phonons emitted in succession.[177] As shown in Fig. 5.16, a charge carrier in the vicinity of an attractive potential can emit a low energy phonon and transition into a highly excited (and thus high radius) state. This carrier can then fall down into the ground state by emitting a number of additional low energy phonons as it transitions through intermediate, less energetic excited states. At higher temperatures in which a *thermal* phonon population is present, there is some probability for phonon absorption to knock the carrier up the energy ladder and back into the continuum states, preventing the carrier from becoming trapped. As the carrier falls deeper into the potential well, it becomes more and more likely for the carrier to become fully captured. This led Lax to introduce an energy-dependent *sticking probability* which gives the likelihood for the carrier to escape a given bound state. This line of reasoning led to cross section predictions which were able to explain the so-called “giant traps” observed in the data. Note that the process of cascade capture does not only apply to the Coulomb potential — any attractive potential with ladder of bound states will have an enhanced cross section under the right conditions.

The original formalism of Lax has been extended and applied to a variety of potentials, resulting in many useful analytic formula for calculating capture cross sections. Most of

these approaches, however, assume thermal equilibrium between the charge carriers and the crystal lattice — a situation which is clearly not true in CDMS detectors. Sundqvist[146] has performed the first calculations of cascade capture under CDMS operating conditions, building upon the work of Abakumov.[178] Separating the carrier temperature T_C from the lattice temperature T_L , Sundqvist finds the Coulomb capture cross section *with zero applied field* to be

$$\sigma_{eq} = \frac{4\pi}{15l_0} \left[\frac{2}{3} \left(\frac{e^2}{(4\pi\epsilon\epsilon_0)kT_C} \right) \left(\frac{e^2}{(4\pi\epsilon\epsilon_0)m^*v_s^2} \right)^2 + 2 \left(\frac{kT_C}{m^*v_s^2} \right) \left(\frac{e^2}{(4\pi\epsilon\epsilon_0)m^*v_s^2} \right)^3 + \left(\frac{e^2}{(4\pi\epsilon\epsilon_0)m^*v_s^2} \right)^3 \right], \quad (5.44)$$

where we define

$$l_0 = \frac{\pi\hbar^4\rho}{2m_D^3\Xi^2I^2} \quad (5.45)$$

as per Abakumov, m_D is the density of states effective mass, Ξ is the acoustic deformation potential, and I is the Bloch wavefunction overlap factor (often approximated to be 1).[178] With an applied field, he finds

$$\sigma_c = \left(\frac{T_L}{T_C} \right)^{3/2} \left(1 + 0.98 \frac{\delta U_0}{kT_L} \right)^{-1} \sigma_{eq}. \quad (5.46)$$

The effects of barrier lowering by the applied field are encoded in δU_0 , given by

$$\delta U_0 = 2\sqrt{\frac{e^3 E}{4\pi\epsilon\epsilon_0}}. \quad (5.47)$$

Note that this capture cross section only takes into account capture due to acoustic phonon emission.

Kogan and Lifshits have shown that although the energy levels near the ground states of various chemical impurity species are somewhat different in energy, the highly excited states all appear to have very similar energies.[179] Since cascade Coulomb capture is initiated by transitions into these states, we expect the capture cross sections to be essentially the same for various species of ionized impurities. In particular, we note that overcharged D^- and A^+ states, discussed soon, can serve as cascade Coulomb capture centers for holes and electrons, respectively. The fundamental cascade Coulomb capture processes are thus



Calculated cross sections from Sundqvist are shown in Fig. 5.17 and Fig. 5.18 for electrons and holes. The labelling of the capture centers as overcharged states has been chosen based on the field dependence of the measured trapping lengths, discussed in a later chapter.

5.4.4 Neutral Capture Across the Band Gap

Neutral donors and acceptors may become slightly polarized due to the presence of an applied field or due to the field produced by a mobile carrier in the vicinity of the impurity. The induced dipole is described by the quadratic Stark effect. This slight dipole potential can also serve as an attractive center which may cause carriers to become trapped. Trapping through this process results in the formation of so-called “standard” D^+/A^- states — the particular processes are



This type of neutral capture is referred to as taking place “across the band gap”, as the shallow impurity states for these processes lie near the conduction band for hole capture and near the valence band for electron capture. The free carrier must therefore fall through the band gap in order to become captured.

Although there is some disagreement upon whether trapping to these states can be properly described by cascade capture, the problem has been treated using Lax’s formalism by Abakumov[178] and Sundqvist[146]. Abakumov derives the cascade capture cross section for a dipole to be

$$\sigma_d = 2\pi \frac{d}{l_0} \left(\frac{e^2}{(4\pi\epsilon\epsilon_0)k_B T_L} \right)^2, \quad (5.50)$$

where d is the dipole length, reported by Sundqvist to be

$$d = (4\pi\epsilon\epsilon_0) \frac{E}{e} a_B^3, \quad (5.51)$$

where a_B is the effective Bohr radius of the impurity atom.

Calculations performed using this method are a good match in magnitude and field dependence to the experimental measurements of Sanada.[180] The experimental data indicates neutral capture cross sections of $\sim 10^{-16} - 10^{-15} \text{ cm}^2$, which are several orders of magnitude smaller than both the cascade Coulomb capture and neutral capture forming overcharged states cross section. **Neutral capture across the band gap does not play a dominant role in sub-Kelvin germanium.** The implications of this result are discussed in a later chapter.

5.4.5 Neutral Capture Producing Overcharged D^-/A^+ States

Although not often discussed, the hydrogen atom possesses a second bound state in which *two* electrons orbit the proton, forming an negatively charged H^- atom. First predicted

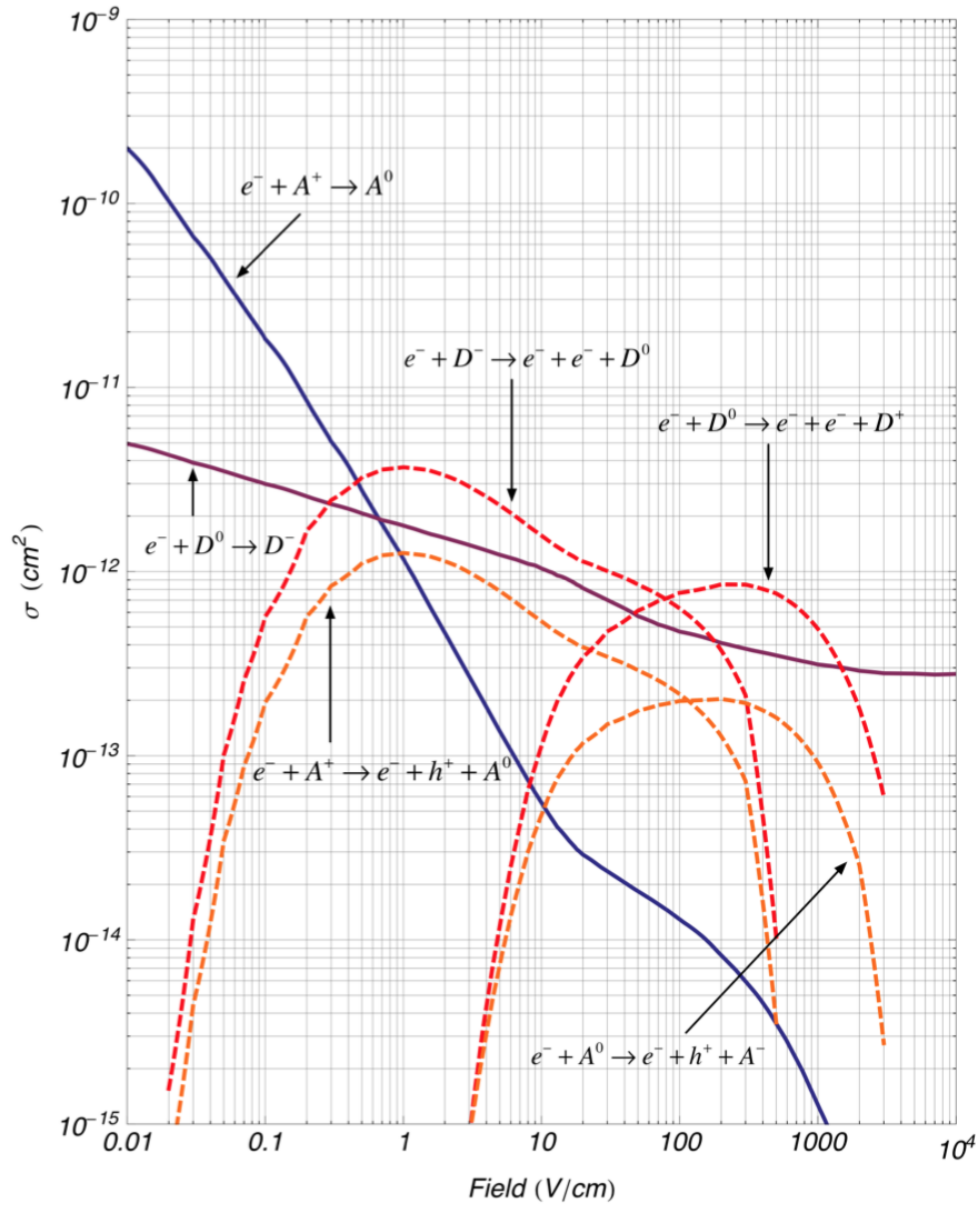


Figure 5.17: Various capture and impact ionization cross sections for electrons in sub-Kelvin germanium, as calculated by Sundqvist.[146]

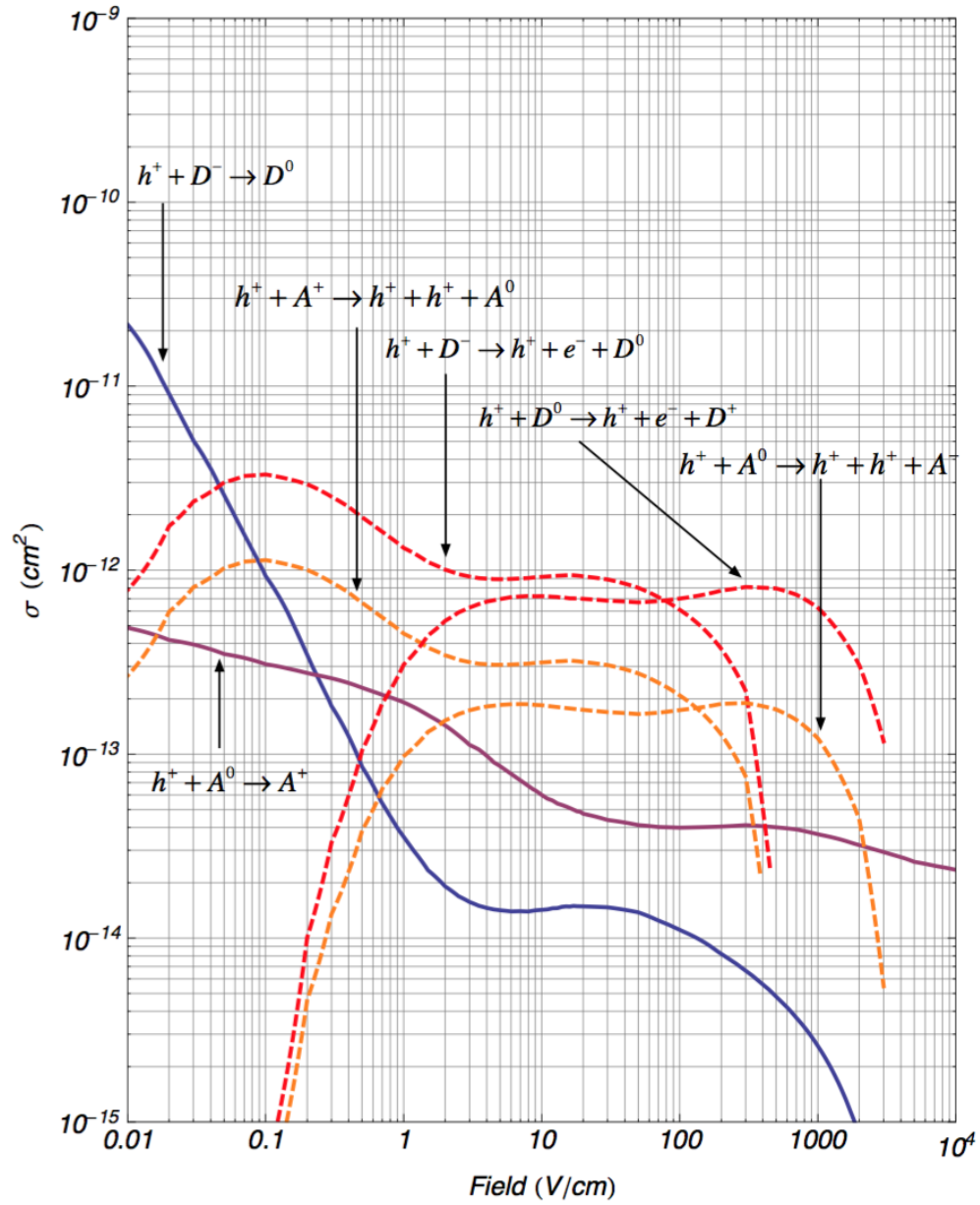


Figure 5.18: Various capture and impact ionization cross sections for holes in sub-Kelvin germanium, as calculated by Sundqvist.[146]

by Bethe[181], this state is referred to the negative hydrogen ion, the hydrogen anion, or overcharged hydrogen. The equivalent semiconductor impurity states have been shown to exist in germanium.[182–186] These overcharged D^- and A^+ states have very shallow binding energies, typically ~ 0.7 meV. An electron in the vicinity of a neutral donor (or a hole in the vicinity of a neutral acceptor) may emit a single low energy phonon and become captured via



Note that as we have electrons trapping onto donors, and holes onto acceptors, these are *same-side* capture processes as opposed to capture across band gap discussed earlier.

Abakumov[178] derives the capture rate for these processes based on the zero-radius-potential model. He finds

$$C_{ovc} = \frac{\pi^3 B^2}{16l_0} \sqrt{\frac{2\mathcal{E}_{ovc}}{m^*}} \left(\frac{\hbar}{m_D v_s} \right)^3 \psi(\mathcal{E}_K), \quad (5.53)$$

where $B = 1.1$ is a normalization constant associated with the zero-radius-potential model, l_0 is the same as for cascade capture, $\mathcal{E}_{ovc}$ is the binding energy of the overcharged state, and $\psi(\mathcal{E}_K)$ is given by

$$\begin{aligned} \psi(\mathcal{E}_K) &= \frac{4}{\pi^2} \left(\arctan \frac{1}{x} - \frac{x}{x^2 + 1} \right)^2, \\ x &= \frac{\sqrt{2m^* v_s^2 \mathcal{E}_{ovc}}}{\mathcal{E}_{ovc} + \mathcal{E}_K}, \end{aligned} \quad (5.54)$$

where $\mathcal{E}_K$ is the kinetic energy of the carrier.

Assuming equal binding energies of 0.7 meV for both D^- and A^+ states, Sundqvist[146] uses the expression for the capture rate in combination with his predicted total velocities to find the neural capture cross sections shown in Fig. 5.17 and Fig. 5.18. Note that these cross sections are quite large, $\sim 10^{-12}$ cm², much larger than neutral capture across the band gap. Typically, semiconductor radiation detectors are operated at much higher temperatures ($T=77$ K is commonplace) which renders these extremely shallow states thermally unstable. Under these conditions, neutral capture producing overcharged states does not play a role in charge collection.

At the sub-Kelvin operating temperatures of CDMS, these states *are* thermally stable and the large capture cross section means that **neutral capture producing overcharged states plays a dominant role in sub-Kelvin germanium**. As we will see in later chapters, the thermal stability of the D^- and A^+ states is what causes CDMS detectors to behave much differently than similar detectors operating at higher temperature.

5.4.6 Other Processes

There are several other capture processes which are readily reported in the literature, such as capture by deep centers or by static dipoles. There also exist impurity states which can

be doubly- or triply-charged which would represent enhanced Coulomb capture centers. For these processes, the number densities or cross sections are typically too small to have much of an effect on CDMS detectors due to the low operating temperature and high purity. As we will see when we compare to our experimental measurements, essentially all capture processes in CDMS detectors are well-explained by Coulomb capture onto overcharged states, or neutral capture producing overcharged states. We therefore decide to forego discussion of these other processes, which can instead be found in both Abakumov[178] and Sundqvist[146].

5.5 Impact Ionization

Impact ionization is the general name given to semiconductor processes in which free charge carriers gain enough energy to *produce* additional free carriers through interactions with either impurity atoms or the crystal lattice itself. At the field strengths used with CDMS detectors (in either iZIP for CDMSlite mode), free electrons and holes do not have enough kinetic energy to excite an electron from the valence band into the conduction band. Any evidence of impact ionization must therefore involve impurity states, namely overcharged D^-/A^+ and neutral D^0/A^0 states. [146, 174] Studies of impact ionization of impurities in germanium at liquid helium temperatures performed by Sclar and Burstein used specimens with impurity concentrations over 1,000 times greater than those of CDMS detectors.[187] At a temperature of 4.2 K, they reported electrical breakdown in germanium due to impact ionization for electric fields between ~ 5 -10 V/cm, which is in agreement with our observations of impact ionization[188] above roughly 5 V/cm at 50 mK which we discuss in a later chapter.

5.5.1 Generation Rates, Times, and Lengths

In some sense, there is nothing fundamentally different about process of capture versus impact ionization other than the fact that capture removes carriers from the conduction/valence bands, while impact ionization contributes carriers. Impact ionization processes will have associated cross sections, by which we can define a generation rate per unit density, generation time, and generation length as in Eqs. 5.35-5.38. These quantities determine the number of excess carriers created by the impact ionization process, rather than destroyed. As we will see in the next section, when the effects of both trapping and impact ionization are to be included in the same equations we should define the generation rates, times, and lengths as *negative* quantities, which can be combined with positive trapping quantities as per Eq. 5.39.

5.5.2 Signal Shape

In the case that impact ionization produces new charge carriers of the same type as those incident upon the impurity center, the resulting signal shape is straightforward to calculate. The produced carriers come to equilibrium with the electric field essentially instantaneously

compared to the drift time. Following our derivation of the signal shape in the case of trapping, we start with

$$\dot{n} = \frac{n}{\tau_{imp}}, \quad (5.55)$$

where τ_{imp} is the generation time. Note we have made the time derivative a *positive* quantity to indicate carriers are being generated through impact ionization. This is easily solved to give us

$$n(t) = n_0 \exp\left(\frac{t}{\tau_{imp}}\right). \quad (5.56)$$

This is the same as Eq. 5.43 but with exponential growth instead of decay and is equivalent to impact ionization representing a *negative* “trapping time”.

When both impact ionization and trapping of the same carrier type occur, we can combine the two effects. The total rate of change is given by

$$\dot{n} = \frac{n}{\tau_{imp}} - \frac{n}{\tau_{trap}} = -n \left(\frac{1}{\tau_{trap}} - \frac{1}{\tau_{imp}} \right) = -\frac{n}{\tau_{eff}}, \quad (5.57)$$

demonstrating the use Eq. 5.39 with impact ionization represented as a negative trapping time to form an effective total trapping time τ_{eff} . Note that τ_{eff} (and the associated length λ_{eff}) can be positive or negative. Whether the drift current depicts exponential growth or decay depends on which process is dominant.

In the case that the impact ionization process generates carriers of the opposite polarity — drifting holes generating free electrons or electrons generating holes — the signal shape becomes much more complicated. In our transport data, discussed in Chapters 6 and 7, we see clear evidence of an “afterglow” current which persists past the stop time of the primary drifting carriers before decaying back to the baseline. An example is shown in Fig. 6.11. If the liberated carriers are of the opposite sign than the primary drift carriers, they drift back towards the opposite detector contact. Carriers liberated near the end of the primary drift time will still be drifting even after primary carriers have reached the contact. Thus we see opposite carrier impact ionization predicts an “afterglow” current whose signal shape we can calculate.

We define the boxcar function as

$$\Pi(t_{start}, t_{stop}, t) = \begin{cases} 1 & : t_{start} \leq t \leq t_{stop} \\ 0 & : \text{otherwise} \end{cases} \quad (5.58)$$

The primary drift current pulse is given by

$$I_{prim}(t) = A \exp\left(-\frac{t}{\tau_1}\right) \times \Pi(t_{start}, t_{stop}, t), \quad (5.59)$$

where A is the initial amplitude of the current and τ_1 is the trapping time for the primary carrier.

Let v_{d1} be the drift velocity of the primary carriers, v_{d2} be the drift velocity of the impact ionized carriers, τ_2 the trapping time of the impact ionized carriers, and assume an impact ionization event happened at time t_{imp} . The primary carriers will have traveled a distance $v_{d1}(t_{imp} - t_{start})$ into the crystal. The impact ionized carriers will then need to drift for $v_{d1}(t_{imp} - t_{start})/v_{d2}$ additional seconds to traverse this same distance back to injection contact. The resulting current at time t from a unit-amplitude impact ionization pulse occurring at t_{imp} will be

$$I_{imp}(t) = \exp\left(-\frac{t - t_{imp}}{\tau_2}\right) \times \Pi\left(t_{imp}, t_{imp} + \frac{v_{d1}}{v_{d2}}(t_{imp} - t_{start}), t\right). \quad (5.60)$$

Carriers will be generated via impact ionization throughout the entirety of the primary drift signal. The total contribution to the current at time t due to impact ionization will be the sum of all impact-ionized carriers still drifting at that time. The probability to impact ionize at a specific location depends on the number of primary carriers, so we should multiply by the primary current signal and integrate over t_{imp} from t_{start} to t . We obtain

$$I_{Timp}(t) = \int_{t_{start}}^t \left[A \exp\left(-\frac{s}{\tau_1}\right) \times \Pi(t_{start}, t_{stop}, s) \right] \times \left[B \exp\left(-\frac{t - s}{\tau_2}\right) \times \Pi\left(s, s + \frac{v_{d1}}{v_{d2}}(s - t_{start}), t\right) \right] ds, \quad (5.61)$$

where B is a constant related to the impact ionization cross section, which we discuss later.

Although somewhat daunting, Eq. 5.61 can be integrated by hand and results in a piecewise function which is split at the end of the primary pulse. Letting $t_{start} = 0$ for simplicity, the result can be written as

$$I_{Timp}(t) = \begin{cases} AB \kappa \exp(-t/\tau_2) (\exp(t/\kappa) - \exp(t/\eta\kappa)) & \text{for } 0 < t \leq t_{stop}, \\ AB \kappa \exp(-t/\tau_2) (\exp(t_{stop}/\kappa) - \exp(t/\eta\kappa)) & \text{for } t_{stop} < t \leq t_{final}. \end{cases} \quad (5.62)$$

The new variables are defined as

$$\begin{aligned} \kappa &= \frac{\tau_1 \cdot \tau_2}{\tau_1 - \tau_2} \\ \eta &= 1 + \frac{v_{d1}}{v_{d2}} \\ t_{final} &= \eta \cdot t_{stop} \end{aligned} \quad (5.63)$$

Note that Eq. 5.62 only describes the current due to impact ionization of the opposite carrier. The total current is the sum of the primary and secondary current and is given by

$$I(t) = I_{prim}(t) + I_{Timp}(t). \quad (5.64)$$

We now need to connect the parameter B to the impact ionization cross section. Consider the case in which a carrier of number density n generates the opposite carrier p through σ_{imp} , and that the impact ionization center has number density N . We start by relating the total impact ionization rate to the change in p through

$$\dot{p} = \frac{dp}{dt} = n N \langle \langle \sigma_{imp} \rangle \rangle \langle v_{tot1} \rangle, \quad (5.65)$$

where v_{tot1} is total velocity of the primary carrier.

Using Eq. 5.42 we can relate n the primary drift current I_{prim} of Eq. 5.59, and we have

$$dp = \left(\frac{I_{prim}}{q v_{d1}} \right) N \langle \langle \sigma_{imp} \rangle \rangle \langle v_{th} \rangle dt. \quad (5.66)$$

The small number of produced carriers dp rapidly come to equilibrium with the electric field, and move in the opposite direction with drift velocity v_{d2} . This small portion of the impact ionization current thus takes the form

$$dI_{imp} = dp q v_{d2} \exp \left(-\frac{t}{\tau_2} \right). \quad (5.67)$$

Comparing Eq. 5.67 to the integrand of Eq. 5.61, with proper redefinition of t and the inclusion of the boxcar functions, we see

$$I_{prim} = A \exp \left(-\frac{s}{\tau_1} \right) \times \Pi(t_{start}, t_{stop}, t), \quad (5.68)$$

$$\frac{dI_{imp}}{I_{prim}} = B \exp \left(-\frac{t-s}{\tau_2} \right) \times \Pi \left(s, s + \frac{v_{d1}}{v_{d2}} (s - t_{start}), t \right) ds. \quad (5.69)$$

$$(5.70)$$

Solving for B gives

$$B = v_{d2} \frac{N \langle \langle \sigma_{imp} \rangle \rangle \langle v_{tot1} \rangle}{v_{d1}} = \frac{v_{d2}}{\lambda_{imp}}, \quad (5.71)$$

where λ_{imp} (in this equation, a strictly positive quantity) is the generation length, or the magnitude of the “negative” trapping length.

5.5.3 Cross Sections

The kinematics and calculation of the impact ionization cross section for shallow impurity states has been performed by Landsberg.[174] A notable difference between shallow state impact ionization and impact ionization across the band gap is that momentum ($\mathbf{k}$) does not need to be conserved when dealing with impurities. This is because $\mathbf{k}$ is not a good quantum number for the localized, bound state Hamiltonian. Landsberg’s full expression for

the impact ionization cross section is complicated. An excellent piece-wise approximation to the cross section was developed by Palmier[189, 190], and is given by

$$\sigma_{imp}(\mathcal{E}) = \begin{cases} 0 & \text{for } \mathcal{E} < \mathcal{E}_i, \\ \sigma_i \left(\frac{\mathcal{E} - \mathcal{E}_i}{\mathcal{E}_i} \right) & \text{for } \mathcal{E}_i \leq \mathcal{E} < 2\mathcal{E}_i, \\ 2\sigma_i \left(\frac{\mathcal{E}_i}{\mathcal{E}} \right) & \text{for } \mathcal{E} \geq 2\mathcal{E}_i, \end{cases} \quad (5.72)$$

where $\mathcal{E}_i$ is the ground state energy of the localized impurity state and σ_i is given by Sundqvist[146] to be

$$\sigma_i = \frac{\pi}{2} \left(\frac{\hbar^2}{2m^* \mathcal{E}_i} \right). \quad (5.73)$$

The energy dependence of the cross section can be understood as follows. Obviously, the incident carrier must possess enough energy to knock the bound carrier out of its localized ground state and into the conduction band — this explains the first case. As the incident carrier energy exceeds the binding energy, additional final states become energetically allowed, increasing the cross section — this explains the second case. At energies above twice the binding energy, the velocity of the incident carrier becomes so high that it does not spend an appreciable time in the vicinity of the impurity and the cross section decreases — this explains the third case. Modifications to the impact ionization cross section due to barrier lowering have also been calculated by Sundqvist[146], but we do not repeat them here.

Sundqvist's predicted cross sections for impact ionization of both overcharged and neutral donors and acceptors are shown in Fig. 5.17 and Fig. 5.18. We see that there are 8 possible processes which we expect to be evident in sub-Kelvin germanium:

$$\begin{aligned} e^- + D^- &\rightarrow 2e^- + D^0 \\ e^- + A^+ &\rightarrow e^- + h^+ + A^0 \\ e^- + D^0 &\rightarrow 2e^- + D^+ \\ e^- + A^0 &\rightarrow e^- + h^+ + A^- \\ \\ h^+ + A^+ &\rightarrow 2h^+ + A^0 \\ h^+ + D^- &\rightarrow h^+ + e^- + D^0 \\ h^+ + A^0 &\rightarrow 2h^+ + A^- \\ h^+ + D^0 &\rightarrow h^+ + e^- + D^+ \end{aligned} \quad (5.74)$$

We note that the impact ionization cross sections are on the order of 10^{-12} cm^2 , and that both same and opposite carrier processes contribute at about equal levels. Impact ionization should therefore be present in CDMS detectors when operated at high field. As we will see when discussing our experimental results, our observations of the drift current at high field clearly indicates that these processes do indeed occur. Impact ionization will be important for future high voltage detectors.

Chapter 6

Experimental Apparatus for the Measurement of Electron and Hole Properties

6.1 Introduction

In order to measure the fundamental charge transport physics responsible for CDMS detector performance, a dedicated experimental apparatus was constructed by myself and Kyle Sundqvist, shown in Fig. 6.1. While radioactive sources and standard charge amplifiers can be used to gain some insight into charge collection, the results lack precision and can often be confusing. We instead use a fiber optic probe coupled to a room temperature laser diode to create small packets of charge which drift across the crystal. A custom high bandwidth charge amplifier is used to track the drift *current*, rather than just the integrated charge. This tracking capability allows the time dependence of the drift current to be observed as the charges move through the bulk germanium, encoding the signatures of the fundamental processes at work. The setup is essentially a cryogenic version of the Haynes-Shockley experiment.[191] The apparatus is now described in detail.

6.2 Fiber Optic Probe

The fiber optic probe couples light from a room temperature laser diode to the surface of the detector, where it creates the free electron-hole pairs which are tracked by the charge amplifier. The use of a long optical fiber connected to a room-temperature light source provides a fast, reliable, and repeatable method of creating controlled concentrations of electron-hole pairs at the surface of the detector. In contrast, cold LEDs mounted near the detector (such as those used for flashing/baking) are slow to turn on, being initially frozen out, and it is not clear when they actually begin to emit light making precise time-of-flight measurements impossible. Moreover, they tend to significantly warm up the experimental

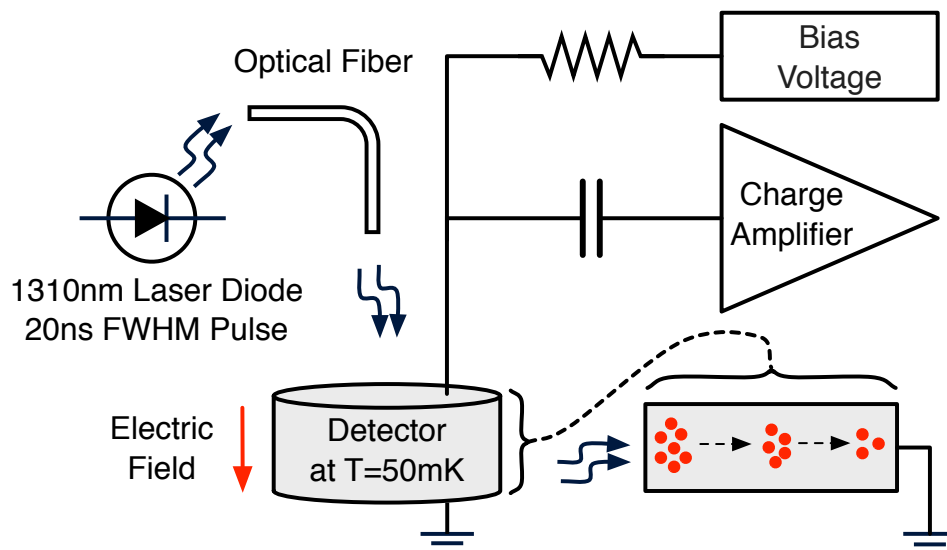


Figure 6.1: The experimental setup. Light from the laser diode traverses the optical fiber and is absorbed by the detector, producing electron-hole pairs near one face of the crystal. The positive bias voltage produces an electric field which causes the holes to drift to the opposite face of the detector. The charge amplifier records the resulting hole drift current. The electron drift current can be studied by changing the sign of the bias voltage, reversing the direction of the electric field.

setup. There are several advantages to using an optical probe rather than a radioactive source for carrier generation.

- Electrons and holes can be studied separately. By choosing a wavelength of light such that the electron-hole pairs are made close to the surface of the crystal, only one species of carrier will be injected into the bulk germanium. The direction of the electric field, determined by the electrode biasing, determines whether it will be electrons or holes. Due to the proximity to the surface, the opposite carrier will be immediately absorbed and will not contribute to the measured ionization signal. Although 60 keV photons from ^{241}Am can be used to approximate these effects, there is still some error associated with their penetration depth of ~ 1 mm in Ge.
- The light pulses can be controlled. The amplitude, duration, and rate of light pulses emitted by the laser can be varied to optimize signal quality and study the effects on detector response.
- The light pulses represent repeatable events. Although radioactive sources can be collimated, there is still substantial event-to-event variation. By using the optical

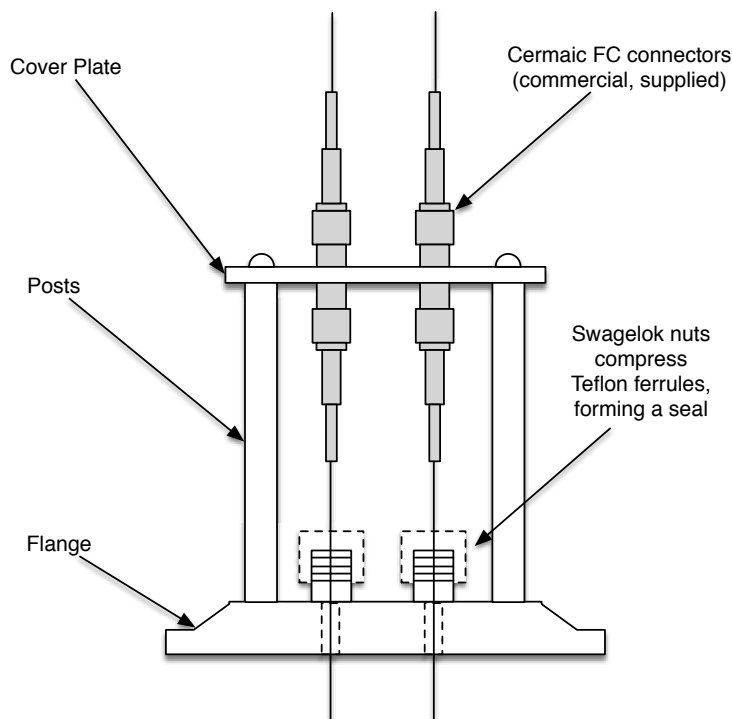


Figure 6.2: The vacuum interface fiber flange. A seal is formed by tightening Swagelok nuts to compress a Teflon ferrule which surrounds the fiber. A connector is added to the exposed end of the fiber, allowed light to be coupling into the dilution refrigerator.

probe, the same light pulse is incident upon the same location, minimizing position dependence.

- The data acquisition system can be triggered by the firing of the laser. The random emissions from radioactive sources are impossible to predict. Since we control the laser, we know exactly when an event will occur. Since the events are also repeatable, this means that they can be averaged to improve the signal-to-noise. This is a huge advantage when using the high bandwidth tracking amplifier, as good signal-to-noise is sacrificed for speed. Care must be taken to ensure that the accrual of space charge has not corrupted later pulses so as not to contaminate the average.

6.2.1 Vacuum Interface Fiber Flange

A custom vacuum interface is required to couple the room temperature light source to the fiber. A diagram of the interface is shown in Fig. 6.2 and is based on the design of Abraham and Cornell.[192] Swagelok pipe fittings are welded into holes drilled through the flange. The

room temperature end of the fiber passes from the vacuum space through the pipe fitting and out to room temperature. A Teflon ferrule with a small center bore is slipped around the fiber and is pushed into the pipe fitting, filling the remaining space. A Swagelok nut (with an open center face to allow the fiber to pass through) is screwed onto the threaded end of the pipe fitting. As the nut is screwed down, it compresses the Teflon around the fiber and forms a vacuum seal, good down to pressures of at least 2×10^{-10} Torr.

A ceramic FC connector is attached to the room temperature end of the fiber and screwed into an FC-FC barrel, allowing it to be coupling to any other room temperature fiber. The barrel is held in a cover plate, rigidly secured to the flange via posts. This prevents the weight of the connector from stressing the fiber — dangling fibers with heavy connectors will break almost immediately.

6.2.2 300K to 4K Can Fibers

The vacuum side of the fiber flange opens into a relatively large drum which connects to a long experimental access tube extending from room temperature to just inside the 4K can. Each room temperature fiber extends from the room temperature fiber flange, down the experimental access tube and along the wall of the 4K can to a separate set of coupling barrels mounted to the bottom lip of the can. To prevent a substantial amount of room temperature blackbody radiation from being coupled into the fridge, small diameter 8.2 μm single mode fiber (Corning SMF-28e+) was used.

Due to fears of outgassing, we were forced to use unjacketed fiber consisting of only the core and the cladding. While the fiber is quite strong along its primary axis, without the jacketing it is extremely susceptible to breakage due to a combination of shearing forces and sharp copper edges within the cryostat. To add some protection, along with increasing the likelihood that at least some small fraction of the fibers would survive the perilous act of being pulled through the experimental access tube, up 8 fibers would be inserted into Teflon tubing as a makeshift form of jacketing. Small (~ 1 inch) slits were periodically cut into the tubing to prevent air from becoming trapped during pump out. Long sections (~ 20 feet) of fiber would be unrolled and taped to butcher paper. Using markers, each fiber was color coded along its length in a repeatable pattern so that it could be identified after insertion. The warm ends would be gathered together and the fiber bundle literally pushed through the Teflon tubing, a process I hope to never have to repeat.

The Teflon bundles would then be securely taped to a long string with a heavy nut tied on the other end. The nut would be dropped through the top of the experimental access tube and jiggled about until it finally emerged out the bottom, after which the fiber bundles would be dragged through by pulling on the string. Long sections of slack were kept at both ends, either by taping to the top of the drum or wrapping around the inner lip of the 4K can, so that the whole process need not be repeated in case the connectors broke — as they usually did. Two fibers from the bundle were screwed into barrels as part of a fixture attached to the 4K can.

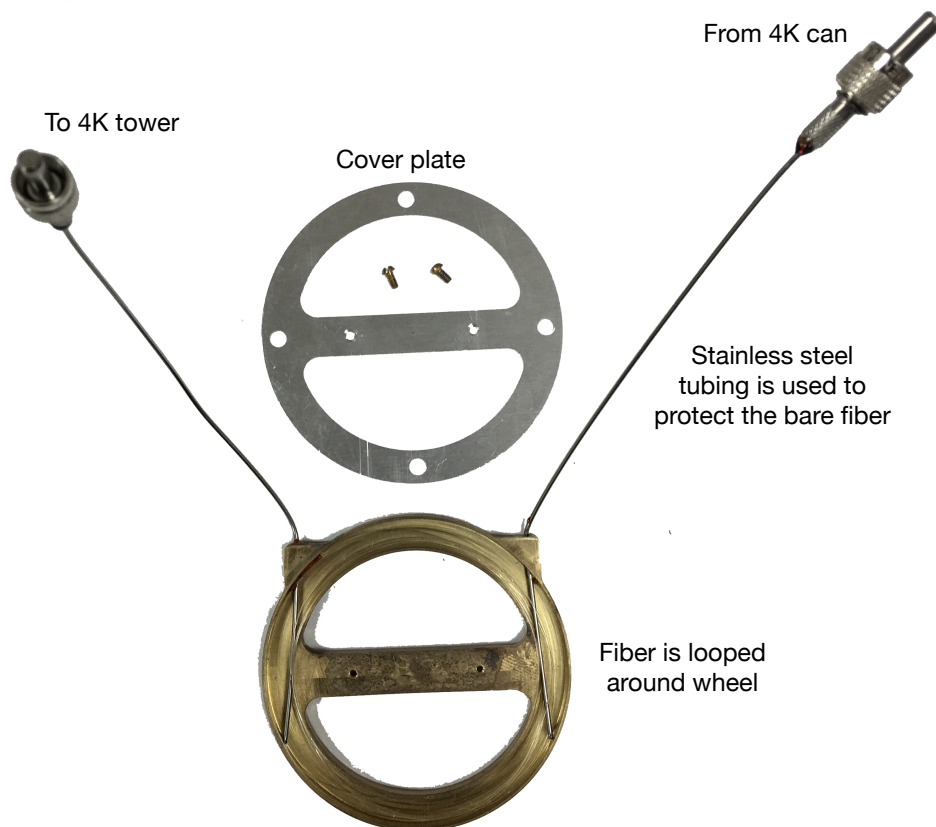


Figure 6.3: A single 4K can to 4K tower fiber.

6.2.3 From single mode to multi-mode

It was discovered that the typical ceramic connectors were prone to cracking after repeated thermal cycling. Instead, SMA-standard stainless steel connectors were used within the cryostat. Likely due to our inexperience in putting connectors on bare fiber, coupling two single mode fibers using SMA connectors and barrels resulted in enormous power losses at the interface. We instead used multi-mode fiber with a $250\text{ }\mu\text{m}$ diameter core within the cryostat. Thinking of the fibers as water pipes, the much larger diameter of the multi-mode fiber allows for much better coupling to the 300K to 4K single mode fiber when using the SMA connectors. Coupling two multi-mode fibers still introduced some losses, however they were small enough to be deemed acceptable.

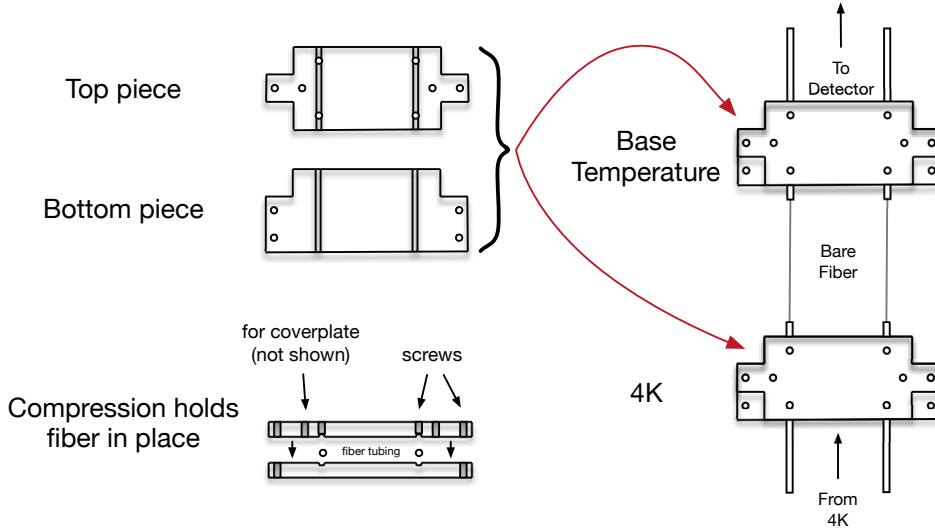


Figure 6.4: Tower fiber mounting fixture. The bottom piece is screwed into the tower wire cover plate of the respective temperature stage. The top piece screws into the bottom, holding the stainless steel tubing surrounding the fiber in place. The section of bare fiber prevents a thermal short between the temperature stages of the tower. A cover plate (not shown) can be screwed between the two temperature stages to protect the bare fiber during transport and allow the two fixtures to remain rigidly separated when removed from the tower.

6.2.4 4K Can to 4K Tower Fibers

Although we can now send photons from room temperature to 4K, we still need to get them to the detector. We must first bridge the gap between the edge of the 4K can and the 4K stage of the CDMS detector tower at the center of the cryostat. A dummy SQUET (see Appendix C) which screwed into the tower was made with a fixture holding two SMA barrels. A long fiber was looped into the metal housing shown in Fig. 6.3, with the connectors extending out of stainless steel tubing for protection. The fiber was looped 30 times within the housing in an attempt to help both thermalization and allow for attenuation of the room temperature blackbody radiation. One end would be screwed into the 4K can barrel, while the other into the tower, thus linking the room temperature end of the fiber to the tower.

6.2.5 Tower Fibers and Detector Lid

The photons now need to be routed up the tower to the surface of the detector. The detector is mounted at the end of the stack, which at test facilities is inserted up inside the mixing chamber can, opposite of the icebox at Soudan. Two SMA barrels were inserted into a detector lid. Only the barrel located above the center of the detector face was used. The

inside of the lid was blackened with Lampblack to absorb the photons which reflect off the crystal surface. A fiber was inserted into two sections of stainless steel tubing and connectors added to each end. A small patch of exposed fiber was left unprotected so as not to thermally link the 4K and base temperature stages of the tower. A two-piece fixture, shown in Fig. 6.4, was screwed over the 4K and base temperature tower wire covers. Screws inserted into the fixture would clamp down onto the two separate pieces of stainless steel tubing, leaving the exposed fiber between the two temperature stages and preventing a thermal short. One end of the fiber was screwed into the dummy SQUET barrel, while the other was taped along the side of the tower and the detector stack. The tubing would be bent as to make the 180 degree rotation necessary to be screwed into the tower lid. The optical connection between room temperature and the detector surface has been formed.

By placing the detector lid directly on the housing of the detector under study, the tip of the fiber was placed only 5 mm from the detector surface. The high intensity of light on such a small area caused it to locally charge up very rapidly making measurements difficult. By placing a blank 1 inch housing at the end of the stack before adding the lid, the light could be spread out over a much larger area preventing rapid local buildup of space charge. Using the housing, the spot size was increased to 1.4 cm^2 .

6.2.6 Laser Diodes

Photons were injected into the room temperature end of the fiber using pigtailed laser diodes. Fast ($\sim 20 \text{ ns}$ FWHM) pulses of light from the diode were created by a commercial pulse generator. While in principal any wavelength compatible with the fiber can be used, by choosing a wavelength which allows excitation of the direct band gap in germanium (0.8 eV, $\lambda \lesssim 1500 \text{ nm}$) the photons will be absorbed very close to the crystal surface. Initial measurements were performed using a 5 mW 785 nm laser, but we have since settled on an 80 mW 1310 nm laser. The injected number of carriers was equivalent to a roughly 1 MeV recoil event, but can be tuned as needed.

6.3 Detector Geometry and Biasing

The properties of electrons and holes depend upon the electric field and impurity concentration of the bulk germanium. The most useful sample is therefore a thick slab of germanium with a constant field and known impurity concentration. As the iZIPs were specifically designed to have a *non*-uniform field, they can be difficult to work with. We instead prefer the older mZIP detectors, which serve as a fine approximation to a 1 inch germanium slab as the crystal diameter is about three times the thickness. To enhance field uniformity, the inner and outer charge electrodes are connected together on one face, as are the four phonon channels on the other. The electric field points from one face to the other, along the axial

direction of the crystal, and is simply given by

$$E = \frac{\Delta V}{L}, \quad (6.1)$$

where L is the thickness of the detector. The Ramo field is constant, and is directly proportional to the electric field since we have only two electrodes.

To allow measurements at high field, the detector was biased from the phonon side of the crystal so as not to blow out the coupling capacitor between the amplifier and the charge electrode. All QETs and their returns were connected to a common biasing line extending from 4K to 300K. Carriers were injected at the charge electrode due to the CDMS convention of having the charge side of the mZIP detectors facing the end cap. A positive bias voltage on the phonon electrode therefore results in electron transport, while a negative bias corresponds to hole transport.

6.4 High Bandwidth Charge Amplifier

The mobility of both electrons and holes at low temperature is extremely high in the crystals used by CDMS, resulting in drift times on the order of $1 \mu\text{s}$. A high bandwidth ($\sim 20 \text{ MHz}$) charge amplifier is necessary to adequately track the charge carriers as they move through the bulk of the crystal, however the standard CDMS charge amplifier is limited to only a few hundred kHz of bandwidth. We developed a custom high bandwidth charge amplifier, shown in Fig. 6.5, which consists of an open loop JFET coupled to a fast current-to-voltage amplifier. We now describe the components of the amplifier.

6.4.1 Open Loop JFET

The input to the amplifier is the gate of an IF4501 Si JFET, the same as used in the CDMS JFET amplifier. To run the JFET in a closed loop configuration, as is done in the CDMS JFET amplifier, the JFET signal must first be sent outside of the cryostat, amplified by warm electronics, and fed back into the cryostat. Just due to the length of the stripline extending from 4K to 300K, the minimum round trip distance is 6 m. Assuming the electric signals travel at roughly $2/3$ the speed of light, this introduces a delay of at least $\sim 30 \text{ ns}$. This delay shows up in the feedback signal as an additional phase shift, determined by

$$\phi_{\text{delay}} = 2\pi f t_{\text{delay}} \approx 100^\circ \quad (6.2)$$

for $f = 10 \text{ MHz}$. Instead of the feedback being the opposite polarity to the input signal at the JFET gate, this additional phase shift will cause the feedback signal to become in-phase with the input, creating undesirable oscillations which render the amplifier useless. To avoid these complications, we operate the JFET in an open loop, common source configuration.

The detector is AC-coupled to the gate through a 10 nF coupling capacitor. The source of the JFET is grounded to the cryostat chassis. DC biasing of the gate is performed through

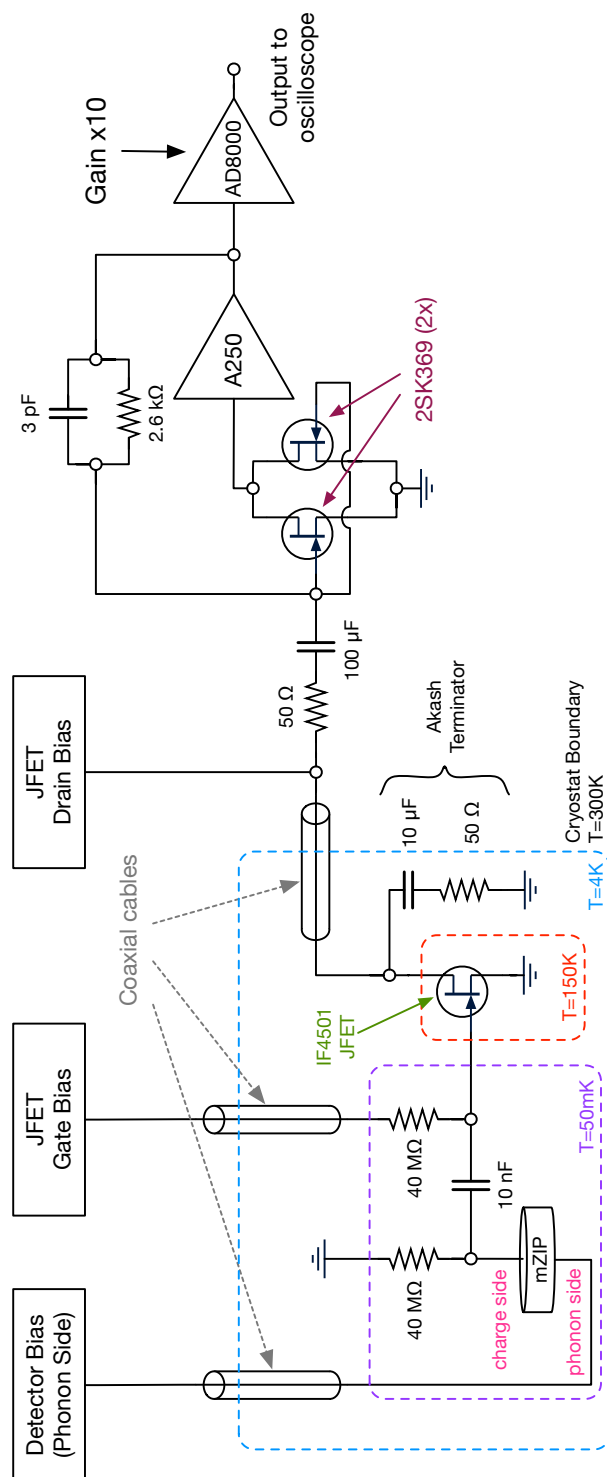


Figure 6.5: The high bandwidth charge amplifier. The standard CDMS JFET is operated in an open loop configuration. The current signal produced by the JFET is amplified by a closed loop, room temperature amplifier consisting of two parallel JFETs coupling to an Amptek A250. The output is buffered by an AD8000 and sent to the oscilloscope for digitization.

a room temperature power supply, in series with a $40\text{ M}\Omega$ resistor at $T=50\text{ mK}$ (this is the feedback resistor of the standard CDMS JFET amplifier). The change in JFET gate voltage due to an event in the detector results in a change in drain-source current, which is accessed through the drain and amplified.

6.4.2 Coaxial Cables

The CDMS stripline was designed to transmit relatively low frequencies, less than about 500 kHz. The intrinsic capacitance and inductance of the stripline results in a broad resonance around 2 MHz, which is obviously problematic for high speed signals. Luckily, the Berkeley cryostat is equipped with 4 long 4K-to-300K coaxial cables with a characteristic impedance of $50\ \Omega$, originally installed by graduate student Miguel Daal for testing of kinetic inductance devices.[193] We modified a CDMS SQUET so that the drain, gate, and phonon side of the detector were instead connected to these coaxial cables rather than the CDMS stripline.

6.4.3 Akash Terminator

The coaxial cables only act as good transmission lines if they are properly terminated at *both* ends by their characteristic impedance of $50\ \Omega$. While this is relatively unimportant for the detector and JFET gate biasing, it is important for the JFET drain coax which carries the signal. The small signal output resistance of the JFET is $\sim 40\text{ k}\Omega$ and clearly does not represent proper termination at the cold end. When running without any additional termination, we observed noticeable ringing in our output pulses near the $\sim 12\text{ MHz}$ characteristic resonant frequency of the coax.

A direct $50\ \Omega$ connection to ground at the cold end of the JFET drain coax cannot be made as it would shunt the JFET bias current. We mimic a direct termination by placing a large ($10\ \mu\text{F}$) capacitor in series with a $50\ \Omega$ connected to ground on a small printed circuit board placed within cryostat, which we refer to as the *Akash terminator*. The board is named after Akash Dixit who originally suggested this termination trick while working as an undergraduate researcher in the Sadoulet lab. The capacitor blocks the DC JFET bias current from being shunted through the resistor, while allowing the node to appear properly terminated within the AC bandwidth of our signal, eliminating the 12 MHz ringing. The downside to this termination scheme is that half the signal amplitude is lost through the terminator, but it is a necessary price to pay to maintain signal fidelity.

6.4.4 Amptek A250

The current signal from the JFET drain must be amplified and converted to a voltage before it can be digitized. This was performed using a commercial Amptek A250 preamplifier (Fig. 6.6). The A250 is essentially a high bandwidth current amplifier — to permit optimization for a wide range of detectors, a user-selectable external JFET must be connected to the

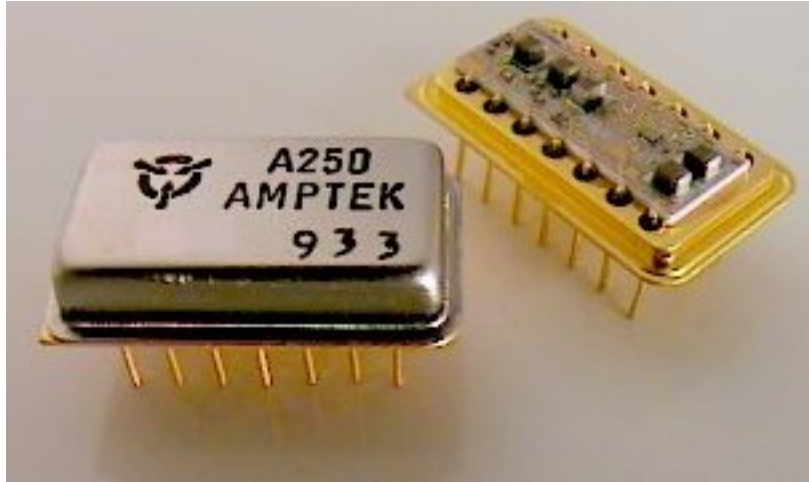


Figure 6.6: The Amptek A250 charge sensitive preamplifier.

A250 input. A feedback connection between the A250 output and JFET gate is required to set the gain and keep the amplifier stable.

Typically, the A250 is used in the same way as the warm electronics portion of the CDMS JFET amplifier — the signal from a cryogenic JFET connected to the detector is amplified by the A250 and fed back to JFET gate, hence the A250 is advertised as a charge sensitive preamplifier. For our purposes, since we do not want to send feedback to the cryogenic JFET, we use the A250 as a high speed current-to-voltage converter. Two room temperature Toshiba 2SK369 are connected in parallel to the A250 input to boost the effective open loop gain. A room temperature feedback loop connecting the warm FET gates to the A250 output is made using a $2.6\text{ k}\Omega$ in parallel to $\sim 3\text{ pF}$ of capacitance. The resistor sets the current-to-voltage gain, while the amplifier bandwidth of 20 MHz is determined by the feedback RC time constant.

6.4.5 AD8000 Follower

Due to radiated noise from the display of the oscilloscope, a long BNC cable was required between it and the amplifier box. Loading the output of the A250 with such a long BNC cable led to unstable oscillations. An ultrahigh speed Analog Devices AD8000 operational amplifier with a voltage gain of 10 was used as a buffer to isolate the A250 output from the oscilloscope cabling. A high bandwidth of several hundred MHz was required to ensure that the pole of AD8000 did not interact with the 20 MHz pole of the A250, which would lower the effective bandwidth.

6.4.6 Gain, Output Pulse Shape, and the Drift Current Signal

Since there is no direct feedback to the gate of the cold JFET, the drift current signal from the detector passes through the parallel impedance of the two 40 M Ω resistors and the total input capacitance (~ 200 pF). The coupling capacitor is large enough to be ignored. As discussed previously, this impedance represents the current-to-voltage transfer function

$$\gamma = \frac{R_{eff}}{1 + j 2\pi f R_{eff} C_{in}}, \quad (6.3)$$

where $R_{eff} = 20$ M Ω . We recognize this as an integrator — the voltage at the gate of the cold JFET will be equal to the integral of the drift current. Note that the integration pole is at a very low frequency of ~ 40 Hz so we are guaranteed to integrate the full drift current pulse. While some might assume that this long integration window means that the high speed timing information is lost, the basic RC integrator is perfect in the sense that the voltage is the true time integral of the current. The high speed components of the signal are still tracked in the rising edge of the signal, as the integration proceeds through the finite width of the pulse.

The transconductance of the JFET converts the integrated gate voltage into a change in drain-source current. The bandwidth of the JFET is high enough as not to distort the signal during this process. The amplifier action of the A250 makes the gates of the room temperature JFETs appear as a virtual ground. Referring to Fig. 6.5, the drain-source current signal sees the impedance of the Akash terminator in parallel to a room temperature series RC of 50 Ω and 100 μ F. The impedance looking into the JFET drain bias is large enough to be ignored. The 100 μ F capacitor is for AC-coupling purposes, while the 50 Ω resistor acts as a room temperature termination for frequencies at which the virtual ground approximation holds. While this effectively terminates the long coaxial cable at both ends, as required to maintain signal fidelity, half of the current is lost through the Akash terminator. The output drain current transmitted to room temperature, referred to the detector drift current is

$$\frac{i_{drain}}{i_{drift}} = \frac{\gamma \cdot g_m}{2}. \quad (6.4)$$

Due to the high input impedance of the two room temperature JFETs, the drain current flowing into the virtual ground must pass through the feedback network between their gates and the output of the A250. The amplifier action of the JFETs+A250 allows the output voltage to swing negative in order to sink the current, and we have

$$\frac{v_{out}}{i_{drain}} = \frac{R_f}{1 + j 2\pi f R_f C_f}, \quad (6.5)$$

where $R_f = 2.6$ k Ω and $C_f = 3$ pF. We again have an integrator, however the integration pole is at a much higher frequency of ~ 20 MHz. Below the pole, the the drain current is converted to an amplified voltage by R_f . In the time domain, this translates into the output voltage responding to changes in the drain current with a 10%-90% risetime of ~ 17 ns, which

is short enough compared to the drift time of $1 \mu s$ for adequate timing resolution. As the drift velocity is the fundamental property of the charge carriers, thinner detectors will have shorter drift times and require higher bandwidth to properly track the signal.

Let f_{lo} be the input impedance integration pole at the gate of the cold JFET. Let f_{hi} be the high frequency integration pole formed by R_f and C_f . Ignore the effects of AC-coupling, which is at low enough frequency to be unimportant. The transfer function from drift current to output voltage can be written as

$$\frac{v_{out}}{i_{drift}} = \frac{1}{2} \frac{R_{eff} \cdot g_m \cdot R_f}{\left(1 + j \frac{f}{f_{lo}}\right) \left(1 + j \frac{f}{f_{hi}}\right)}. \quad (6.6)$$

The output is an amplified copy of the integral of the drift current signal from f_{lo} to f_{hi} . Frequencies above f_{hi} have already been integrated once (by virtue of the second pole at f_{hi}) and cannot be directly tracked.

Taking the time derivative of the output pulse is equivalent to multiplying by $j 2\pi f$ in the frequency domain. The transfer function becomes

$$\frac{\left(\frac{dv_{out}}{dt}\right)}{i_{drift}} = \frac{j 2\pi f}{2} \frac{R_{eff} \cdot g_m \cdot R_f}{\left(1 + j \frac{f}{f_{lo}}\right) \left(1 + j \frac{f}{f_{hi}}\right)} \approx \frac{1}{2} \cdot \frac{1}{C_{in}} \cdot \frac{g_m \cdot R_f}{\left(1 + j \frac{f}{f_{hi}}\right)}, \quad (6.7)$$

where the approximation holds for $f \gg f_{lo}$ and we use $f_{lo} = 1/2\pi R_{eff} C_{in}$. We have undone the low frequency integration and recovered a high bandwidth, amplified copy of the drift current signal as illustrated in Fig. 6.7.

The penalty for using this method to recover the drift current is degradation of the signal-to-noise. The intrinsic voltage noise of the cryogenic FET, which tends to be white and dominant at high frequencies, is not integrated by the input impedance and also appears white on the output. When taking the derivative, this white noise is multiplied by f as in Eq. 6.7 and thus increases with frequency — there is no low frequency pole to cancel. While still present, the effects of this noise can be greatly reduced by both filtering and averaging.

6.5 Data Acquisition System

A custom LabVIEW-based data acquisition system was created to control and synchronize the various pieces of equipment. Traces were digitized using a Tektronix TDS3014 digital phosphor oscilloscope. The detector was biased using a Keithley 2612A SourceMeter to allow for voltages beyond the ± 10 V range of the standard CDMS electronics. A square pulse from a signal generator was used as a global trigger. The trigger initiated both the firing of the laser diode and the oscilloscope acquisition. By controlling the creation of electron-hole pairs and the acquisition of the resulting drift current with the same trigger, the start time of the pulse always occurs at the same position of the oscilloscope trace. This allows for multiple pulses to be averaged together, improving the signal-to-noise.

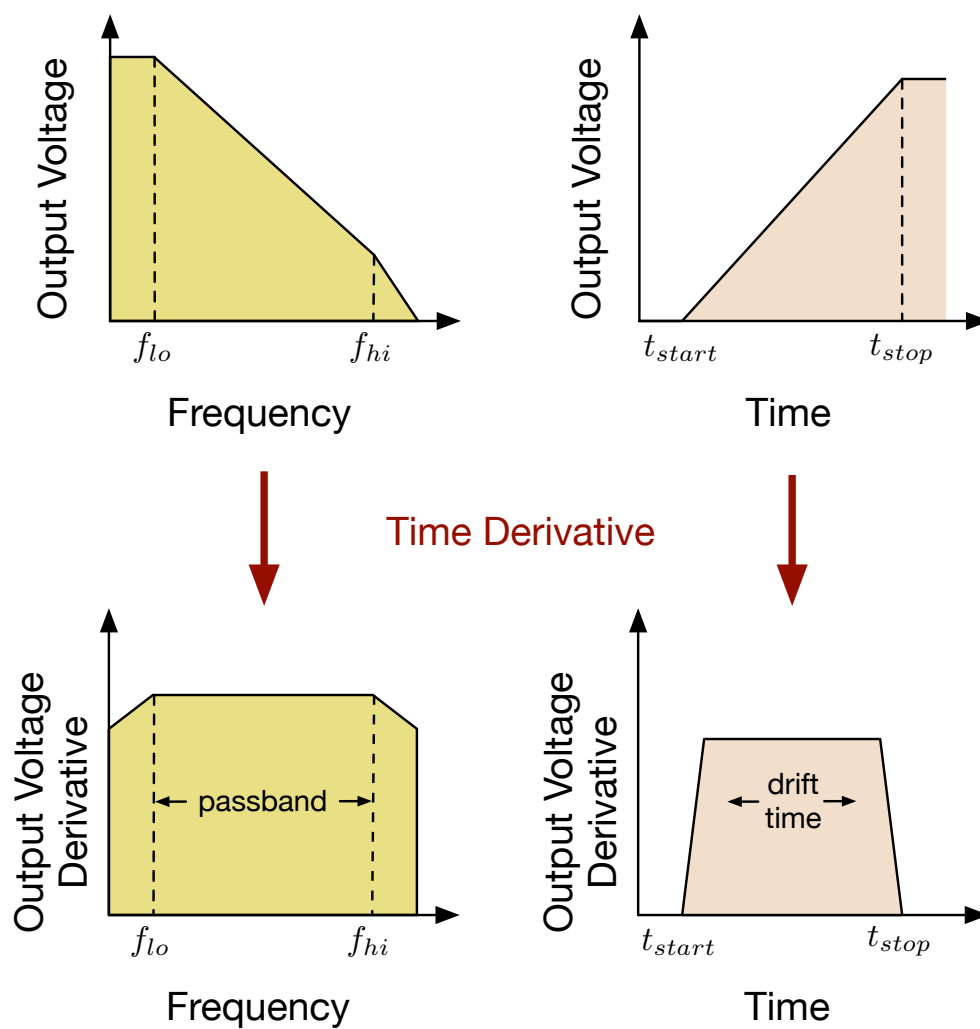


Figure 6.7: By taking the time derivative of the output voltage, the low frequency integration pole is cancelled and the drift current signal recovered.

Electric Field (V/cm)	Oscilloscope Averaging	Oscilloscope Traces	Total Traces
0.0—0.2	1	24	24
0.2—0.8	4	8	32
0.8—2.0	8	8	64
>2.0	32	16	512

Table 6.1: Total number of traces acquired for different electric field ranges. The total number of traces is given by the oscilloscope averaging times the number of oscilloscope traces.

Under ideal circumstances, the output of the amplifier would have been recorded using a dedicated high speed digitizer. Unfortunately, all the on-hand digitizers in the Sadoulet lab had sampling rates far too slow to record the drift current, and digitizers with the required ~ 10 ns sampling rate were prohibitively expensive. We were stuck using the oscilloscope to record the traces. The limited data transfer rate of the GPIB interface between the oscilloscope and the computer caused it to take about a second to record a single trace, making acquisition of a large number of individual traces for a given bias difficult. To overcome this limitation, we made use of the on-board oscilloscope averaging to combine several traces before transferring to the computer, reducing the dead time. The total number of acquired traces per bias was dependent upon the electric field strength. At the lowest fields, the trapped charge from a single pulse would be sufficient to significantly alter the shape of subsequent pulses and the detector reset procedure would need to be performed after every pulse. As the field is increased, the amount of trapped charge necessary to substantially alter the applied field increases, while at the same time the amount of trapping decreases, allowing for many more pulses to be acquired before needing to reset the detector. The approximate number of traces acquired for a given electric field range is given in Table 6.1.

After initial cooldown, the detector undergoes a lengthy overnight bake (repeat firing of cold LEDs in proximity to the grounded detector) to put the detector into its grounded, sub-Kelvin steady state. An energy calibration is performed using the 1.17 MeV and 1.33 MeV photopeaks from a ^{60}Co source. Data taking proceeds as follows:

1. The detector is biased to the desired voltage.
2. The global trigger is sent.
3. The laser diode is pulsed the same number of times as the oscilloscope averaging.
4. The averaged oscilloscope trace is stored to the computer, and the process repeated until the total number of traces has been reached.

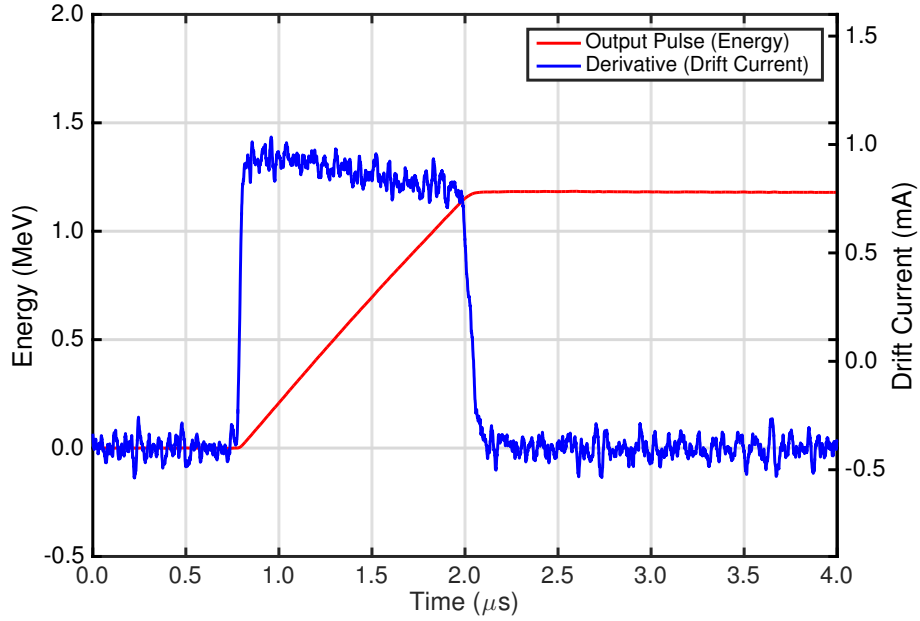


Figure 6.8: A typical averaged output pulse, whose amplitude represents the total charge energy, and its derivative, which represents the drift current.

5. The detector is grounded and the cold LEDs are briefly flashed to return the detector to the grounded, steady state. This process heats the detector to a few hundred mK.
6. The experiment is paused while the detector cools back to base temperature. This typically takes 5-10 minutes.
7. The detector is biased to the next desired voltage and the process repeated until all bias points have been recorded.

The total number of traces was usually not recorded in a single sweep of the biasing, particularly at the lower fields where only a small number of pulses can be fired before the detector must be reset. Multiple sweeps of the bias points would be performed until the desired number of traces were acquired.

6.6 Data Processing

Data processing was done using a custom analysis package developed in MATLAB. There are 10,000 points in each trace, with different time windows being used depending upon the bias voltage/drift time of the carrier under study. After passing initial data quality checks

to ensure good individual pulse fidelity, traces were separated by bias and subjected to a number of processing steps:

1. Any baseline offset was removed by subtracting the average of the first 1,000 points of each trace.
2. A zero-phase shift 2nd order lowpass Butterworth filter was applied to all pulses using MATLAB's `filtfilt` command.[194] For bias fields below 0.4 V/cm, a 5 MHz low frequency cutoff was used. For bias fields above 0.4 V/cm, a 35 MHz low frequency cutoff was used.
3. All traces for a given bias were averaged together.
4. Recalling that the raw output of the charge amplifier is the integral of the drift current, the total charge collected is proportional to the amplitude of the pulse. Since the faltime of the amplifier is much longer than measurement window, the pulse amplitude was determined by averaging the last 1,000 points of the trace. The voltage amplitudes were converted to electron-equivalent recoil energy using ^{60}Co calibration data.
5. The numerical derivative of each output pulse was calculated, recovering the drift current. A typical averaged output pulse and the corresponding drift current pulse are shown in Fig. 6.8.

At the end of processing, we are left with a set of well-averaged, calibrated electron and hole drift current pulses which are ready for more detailed analysis.

6.7 Drift Velocity Measurement

Due to our choice of laser wavelength and detector geometry, charge carriers of either type which are injected into the bulk must traverse the full thickness of the detector before they are absorbed at the opposite face. The *drift time* is the field- and carrier-dependent time it takes travel this distance, and is equal to the width of the measured drift current pulse. By measuring the drift time τ_d , we have a direct measurement of the drift velocity of the charge carriers as

$$v_d = \frac{L}{\tau_d}. \quad (6.8)$$

We are in essence performing a time-of-flight measurement.[195]

A subtlety arises when discussing the drift velocity of electrons because of the germanium band structure and electrode configuration of the detectors under study. Recall that CDMS detectors are aligned such that the axial (or z) direction is aligned with the $\langle 100 \rangle$ crystal lattice axis. At lower electric fields, where the mean free path of intervalley scattering events is on the same order as the thickness of the detector, a given electron will be confined to one of the 4 symmetric L-valleys along the $\langle 111 \rangle$ crystal axis directions. The total electron

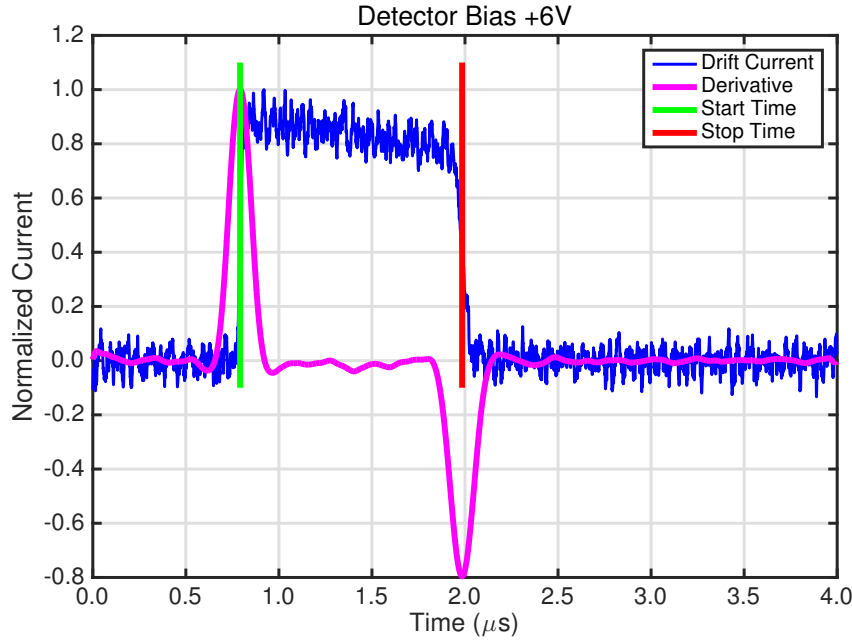


Figure 6.9: Determination of the drift time for a typical pulse. The drift current is shown in blue, and its smoothed derivative shown in magenta. Note the clear peaks indicating the beginning and end of the current pulse. Start and stop times for the drift velocity measurement are shown as the green and red vertical lines.

drift velocity is **not parallel** to the $\langle 100 \rangle$ axis and what we measure is the projection of the electron's true drift velocity along $\langle 100 \rangle$ axis. In any case, since what we actually measure is the collective drift signal from a large number of electrons, and since each L-valley is equally likely to be populated, the perpendicular components of the collective drift velocity cancel and we are again left with only the $\langle 100 \rangle$ component remaining.

The start and stop times of each averaged drift current pulse are determined by examining its smoothed derivative. As shown in Fig. 6.9, a large positive spike is present at the beginning of the drift current pulse, while a large negative spike exists at the end. We fit a Gaussian to each peak and take the resulting means to be the start and stop times. Their difference is the drift time, with the error taken to be

$$\Delta\tau_d^2 = \Delta t_{start}^2 + \Delta t_{stop}^2 + \Delta t_{sys}^2, \quad (6.9)$$

where $\Delta t_{start}/\Delta t_{stop}$ are the errors on the Gaussian fit means which determine the start/stop time, and Δt_{sys} is a constant 10 ns error added to account for the finite risetime of amplifier response. The drift velocity is found using Eq. 6.8 with $L=1$ inch, with error

$$\Delta v_d = \left(\frac{v_d}{\tau_d^2} \right) \Delta\tau_d. \quad (6.10)$$

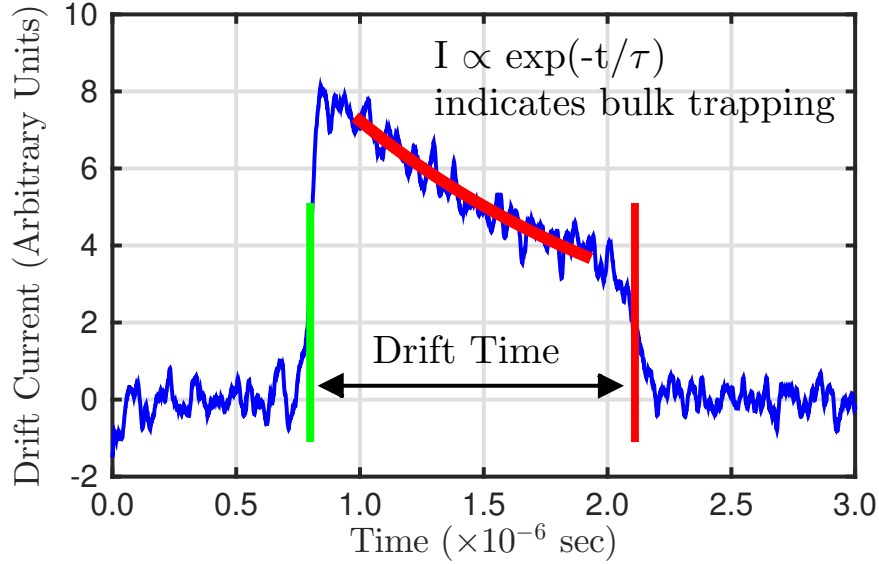


Figure 6.10: Exponential fit to determine the trapping time, and hence trapping length when combined with knowledge of the drift velocity. The exponential decay of the current throughout the entirety of the pulse is a clear indication of bulk trapping within the germanium. The red curve is the exponential fit to the drift current to obtain the trapping time. Start and stop times for the drift velocity measurement are shown as the green and red vertical lines.

6.8 Trapping Length Measurement

The exponential decay of the drift current is a clear indication of bulk trapping within the germanium. As shown in Fig. 6.10, a *trapping time* τ_{trap} is determined for each pulse through a simple exponential fit. To avoid the effects of the finite amplifier risetime, the exponential fit is performed in a window starting 100 ns after the previously-determined start time, and ends 100 ns before the previously-determined stop time. The *trapping length* is determined by the product of the trapping time and drift velocity,

$$\lambda = v_d \cdot \tau_{trap}. \quad (6.11)$$

The error on the trapping length is given by

$$\left(\frac{\Delta\lambda}{\lambda}\right)^2 = \left(\frac{\Delta v_d}{v_d}\right)^2 + \left(\frac{\Delta\tau_{trap}}{\tau_{trap}}\right)^2. \quad (6.12)$$

The trapping length is the preferred quantity to quote when discussing detector physics as it is easily compared to the dimensions of the detector — a much longer trapping length

than the detector thickness would indicate the effects of trapping may be safely ignored, and become more significant as the trapping length approaches the detector thickness. Note that determination of the trapping length does not directly depend on the amplitude of the pulse.

6.9 Impact Ionization Measurement

As we have discussed in the previous chapter, there are 8 impact ionization processes we expect to observe as the bias across the crystal is increased. In the case of same carrier impact ionization which dominates over trapping, the signal shape will exhibit exponential growth instead of exponential decay. If opposite carrier impact ionization also occurs the pulse shape becomes much more complicated, however we know the expected stop time of the secondary current should simply be the sum of the drift times for electrons and holes. The multiple exponentials of Eq. 5.62 combined with multiple processes within the same pulse (see next chapter) make accurate determination of impact ionization parameters exceedingly difficult.

Faced with these difficulties, we simply fit the primary portion of the drift current signal to a simple exponential (which can either grow or decay) to find an effective trapping length, which may be positive or negative. We also calculate the expected stop time of the secondary pulse based on the primary pulse stop time and the drift velocity of the opposite carrier, as shown in Fig. 6.11.

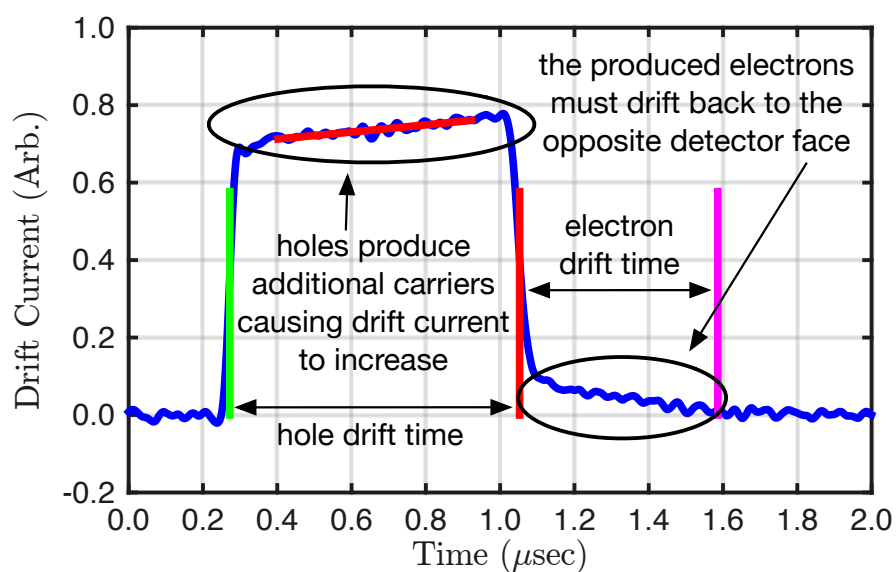


Figure 6.11: Drift current pulse showing exponential growth in the primary pulse due to impact ionization, the associated fit to determine an effective trapping length, and the expected stop time of the secondary current under the assumption of opposite carrier impact ionization. In this case, the primary carriers are holes while the secondary carriers are electrons.[188]

Chapter 7

Analysis of Measured Electron and Hole Properties

7.1 Introduction

To understand the underlying charge transport physics which influence CDMS detector ionization collection, it is necessary to study the electric field dependence of parameters such as the drift velocities and trapping lengths. The previously described experimental apparatus was used to perform detailed charge transport studies of two CDMS mZIP detectors, G22Q and G13H, whose properties are summarized in Table 7.1.

The germanium boules from which the detectors were formed were both purchased from ORTEC. The net shallow impurity concentration $|N_D - N_A|$ for both detectors was reported to be 1.3×10^{10} per cubic centimeter. In G22Q, the donors outnumber the acceptors making the bulk germanium n-type. In G13H, the acceptors dominate making it p-type. G22Q was chosen for study because it represents the typical germanium used by CDMS to make detectors — n-type, with a net impurity concentration of $\sim 10^{10}$ per cm^{-3} . Historically, detectors made from n-type germanium have performed better than those made from p-type germanium. Nevertheless, a handful of p-type CDMS detectors exist. In truth, G13H was chosen for study due to its availability for testing, however it is an excellent counterpart to G22Q as the only essential difference is that it is p-type rather than n-type.

Detector	Type	Impurity Conc. (cm^{-3})	Etch Pit Density (cm^{-2})	Soudan ID
G22Q	n	1.3×10^{10}	4370	MT2Z6
G13H	p	1.3×10^{10}	5290	MT1Z2

Table 7.1: Properties of the two detectors studied using the optical probe and high bandwidth charge amplifier.

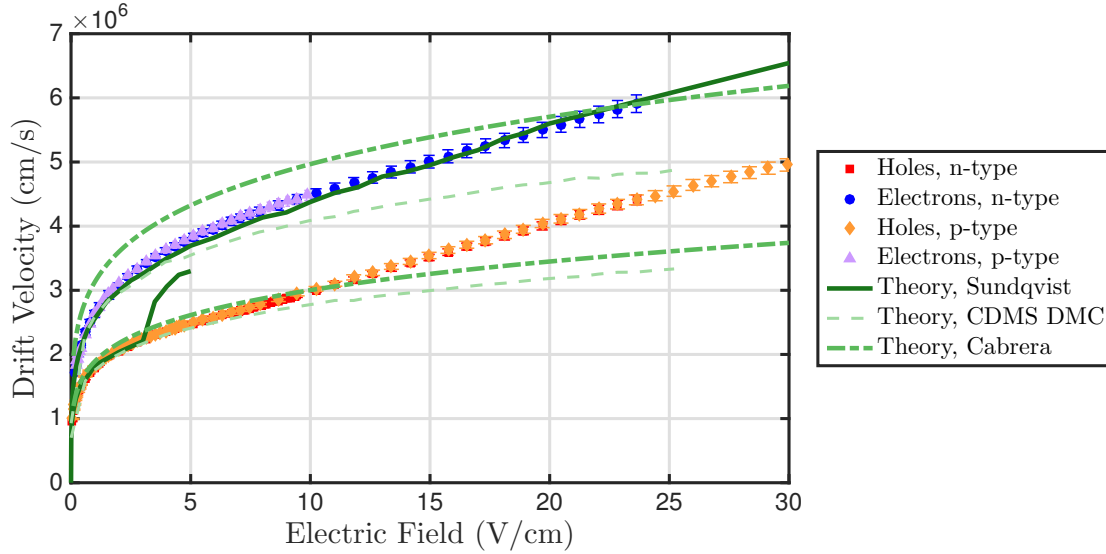


Figure 7.1: Measured electron and hole drift velocities as a function of electric field for the n-type and p-type detector. Theory from Sundqvist[146], CDMS DMC[168], and Cabrera[166].

7.2 Drift Velocities

The measured electron and hole drift velocities as a function of electric are shown in Fig. 7.1. We immediately see that there is no discernible difference between the n-type and p-type detectors, indicating that impurities play a minimal role in determination the underlying energy and momentum distributions of the charge carriers. As expected, transport properties are determined by the electron-phonon interaction in these high purity crystals.

Also shown in Fig. 7.1 the drift velocity predictions of a detailed charge transport Monte Carlo by Sundqvist[146], the CDMS Detector Monte Carlo (DMC)[168], and Cabrera[166]. In particular for the electrons, Sundqvist's predictions are in very good agreement with the measured data. The kink in the predicted holes velocity curve at ~ 4 V/cm is due to the expected onset of optical phonon emission, which is not observed in the data. Rather than a threshold effect, it appears that optical emission turns on in a much more gradual fashion. Based on these results, Sundqvist chooses to focus on acoustic phonon rates when calculating capture cross sections. The excellent agreement between Sundqvist's predictions at low field justifies the use of his effective carrier temperatures for cascade Coulomb capture, which is only significant below a few V/cm.

7.3 Qualitative Dependence on Electric Field

The qualitative change in drift current pulse shape as the electric field is increased is shown in Fig. 7.2 for G22Q (n-type) and in Fig. 7.3 for G13H (p-type). All traces are plotted with the same vertical and horizontal limits so they may be compared directly. For both detectors, we see that the initial drift current increases with the applied field, however the drift time decreases. Since we are discussing the drift *current*, this is expected. Higher drift velocities will result in a larger current, as $I_d \sim qv_d$, which must drift for a shorter period of time through the constant detector thickness. In the absence of trapping or impact ionization, the total charge will be the same for all pulses and the area under the curve conserved between the various fields. The field range studied for the p-type electrons is limited due to breakdown of the crystal at higher field for the electron transport bias polarity.

Three effects are observed which influence the charge collection. First, for both electrons and holes, the drift current decreases with time throughout the entirety of the pulse at lower electric fields. This is a clear indication of bulk charge trapping within the germanium. With the exception of the p-type electrons, the trapping is observed to decrease in magnitude as the field is increased. The p-type electrons initially follow this trend, although the trapping again increases before the onset of breakdown.

Second, again for both carriers, a small but significant secondary drift current is observed to persist after the primary pulse has reached the opposite face of the detector for fields above ~ 5 V/cm. The amplitude of this secondary current increases with applied field. We interpret this as impact ionization generating carriers of the opposite charge, as discussed earlier — electrons produce holes and holes produce electrons. Note that this effect appears to be suppressed for the p-type holes.

Finally, when holes are the primary charge carrier, we observe the onset of a positive slope to the primary component of the drift current as the applied field is increased. We believe this is at least partially due to additional same sign impact ionization overcoming the trapping — holes produce additional holes. In particular, we see a positive slope in the p-type holes data without a large secondary current signal.

This qualitative change in pulse shape demonstrates the rich underlying physics which influence ionization collection in CDMS detectors. While approximating the ionization signal as a charge impulse is a decent first order approximation, a deep understanding of CDMS detector behavior requires more detail.

7.4 Trapping Lengths

The dependence of the trapping length on the electric field encodes information about the underlying processes responsible for carrier capture. The effective trapping length at each measured field point is determined through the process described in the previous chapter. The results are plotted in Fig. 7.4 and Fig. 7.6 for the n-type and p-type detectors. We im-

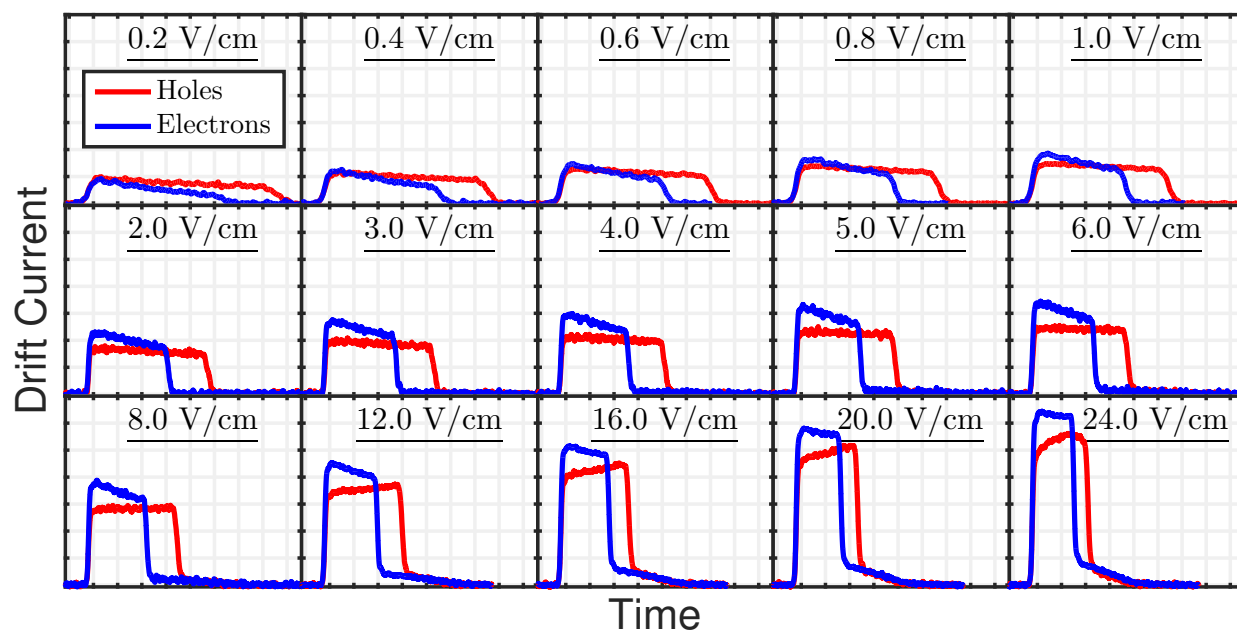


Figure 7.2: Qualitative change in drift current pulse shape for electrons and holes as the electric field is increased in detector G22Q (n-type).

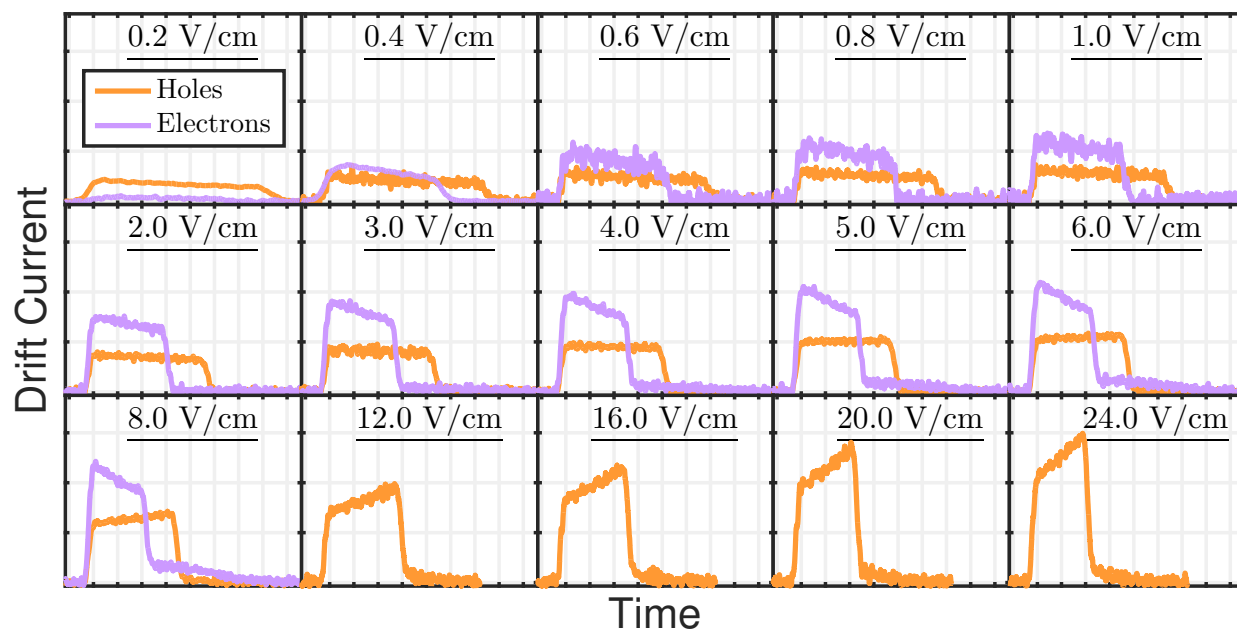


Figure 7.3: Qualitative change in drift current pulse shape for electrons and holes as the electric field is increased in detector G13H (p-type). Detector breakdown prevented acquisition of high field electron data.

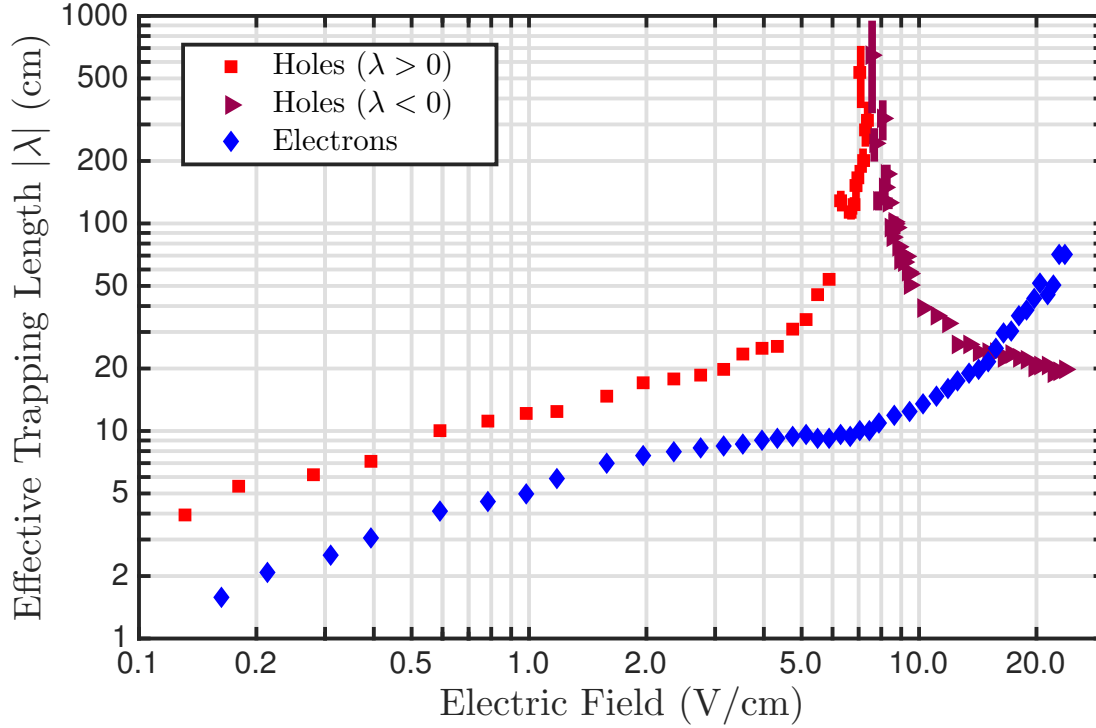


Figure 7.4: Electron and hole trapping lengths for G22Q.

mediately note the non-trivial field dependence and substantial differences between electrons and holes, and between detectors.

7.4.1 G22Q (n-type)

Focusing first on the n-type data, we can split the trapping lengths into three approximate regions.

1. **Low field (0-1 V/cm):** Trapping lengths less than $\sim 5x$ the detector thickness. The trapping length starts out short (even shorter than the detector thickness for electrons) and grows with electric field. The high amount of trapping and strong field dependence is indicative of cascade Coulomb capture.
2. **Mid field (1-5 V/cm):** Trapping lengths between ~ 5 - $10x$ the detector thickness. The rate of growth with electric field decreases. As we will show, this region is dominated by neutral capture producing overcharged states.
3. **High field (>5 V/cm):** The trapping length begins to grow rapidly with increasing field, reaching several tens to hundreds of times the detector thickness. The energy

threshold for impact ionization processes to occur has been surpassed. For holes, the trapping length diverges and changes sign, indicating impact ionization has overcome the trapping during the primary drift current pulse. It appears the electrons will follow this trend if the field strength is further increased.

The behavior of the trapping length at high field, along with the qualitative change in drift current shape discussed earlier, is clear evidence of impact ionization. A non-negligible fraction of opposite carrier impact ionization is responsible for the tails which are observed after the end of the primary pulse. As we have derived in Eq. 5.62, the primary portion of the pulse is no longer described by a simple exponential. We therefore limit our quantitative interpretation of the trapping lengths in terms of capture cross sections to the low and mid field region.

In order to compare our data to the cross sections described in Chapter 5, it is most useful to look at the inverse trapping length. Based on the field dependence, magnitude of the measured lengths, and order-of-magnitude knowledge of the total shallow impurity concentration, we fit our inverse trapping length data to the combination of neutral capture forming overcharged states and cascade Coulomb capture. The functional form of the model is given by

$$\frac{1}{\lambda} = N_q \sigma_q \frac{\langle v_{tot} \rangle}{v_d} + N_0 \sigma_0 \frac{\langle v_{tot} \rangle}{v_d}, \quad (7.1)$$

where σ_q is the total Coulomb capture cross section, N_q is the number density of Coulomb capture centers, σ_0 is the neutral capture cross section, and N_0 the number density of neutral capture centers. The cross sections, total velocities, and drift velocities are taken from Sundqvist.[146] The fit is performed by allowing the number density of the trapping centers to vary for both processes. The inverse trapping lengths for electrons and holes, along with the fits, are shown in Fig. 7.5.

We see the model is a good qualitative fit to the data over most of the field range in question. The only discrepancy in shape is in the high field holes data, where we begin to be dominated by impact ionization. Although plotted in the figure, the holes inverse trapping lengths above 5 V/cm are not used for the fit as the model is not valid in this region. The resulting impurity concentrations and the assumed capture processes, summarized in Table 7.2, are in general agreement with the impurity concentration of the detector however they are in contention with the empirical bulk neutrality of the crystal. A net difference of only 10^8 per cm^3 charged impurities would generate an observable space charge field of ~ 1 V/cm within the detector which was not seen.¹ We also find a higher concentration of acceptors even though the detector was made from n-type germanium. We believe there are several reasons for these discrepancies and do not take them to indicate any fundamental error in the theory.

¹As noted by Abakumov[178], the physical locations of oppositely-charged impurities in low temperature germanium are correlated. A high degree of compensation would be expected even though $\sim 10^{10}$ per cm^3 charged impurities of both polarities are present.

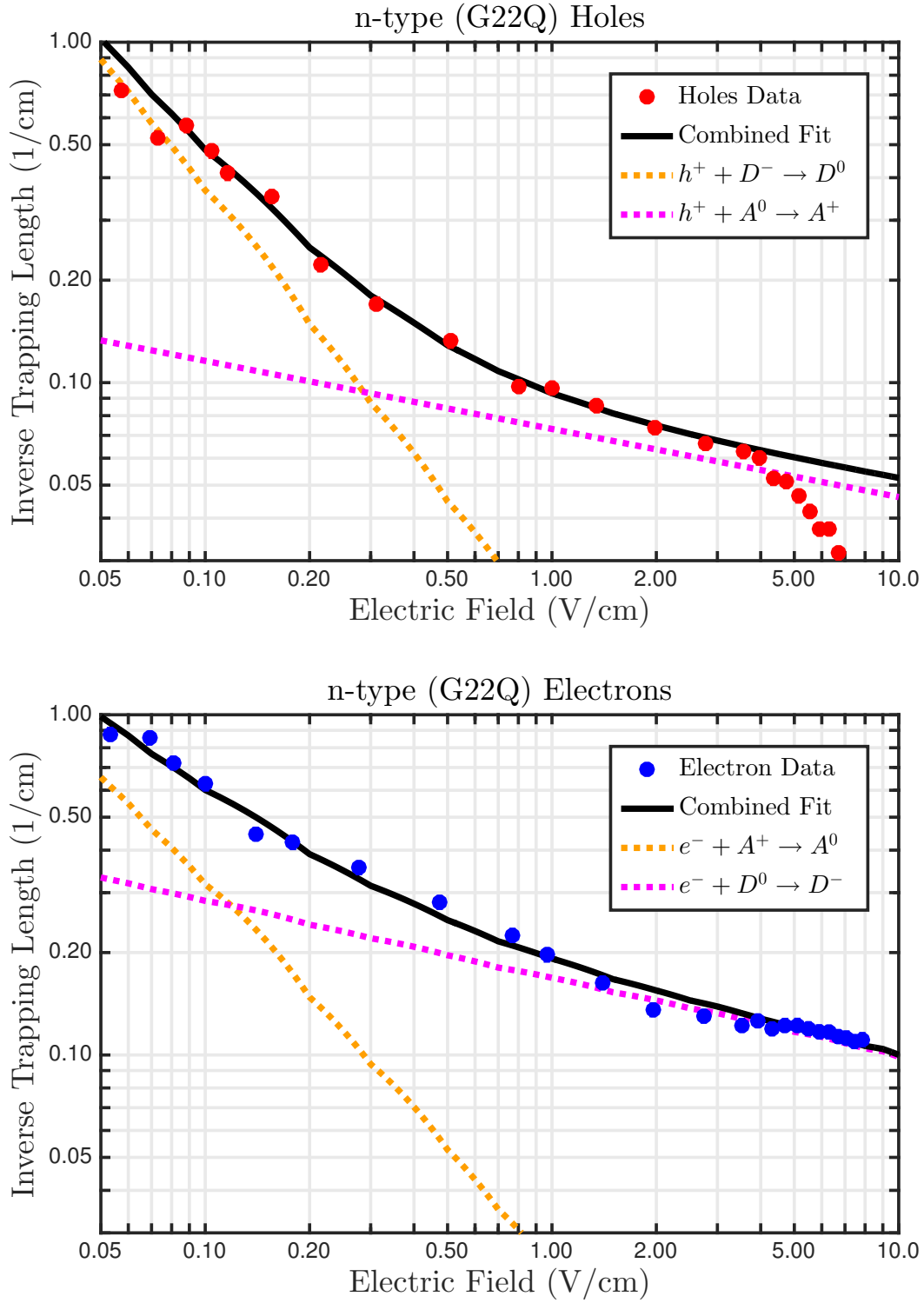


Figure 7.5: Inverse trapping lengths with fits for G22Q.

G22Q (n-type)			Number Density
Process	Type	Trapping Center	from Fit (cm^{-3})
<u>Hole Trapping</u>			
$A^0 + h^+ \rightarrow A^+$	Neutral	A^0	3.6×10^{11}
$D^- + h^+ \rightarrow D^0$	Charged	D^-	8.3×10^{10}
<u>Electron Trapping</u>			
$D^0 + e^- \rightarrow D^-$	Neutral	D^0	8.6×10^{10}
$A^+ + e^- \rightarrow A^0$	Charged	A^+	1.4×10^{10}

Table 7.2: G22Q charge trapping processes for electrons and holes for the the trapping model described in the text. The number densities are the trapping center concentrations from the fits in Fig. 7.5.

First, the capture model is isotropic although charge carriers in milliKelvin germanium experience a high degree of anisotropy. Second, it is unclear from the literature whether the conductivity mass or density-of-states mass should be used for portions of the cascade Coulomb capture calculation. Finally, as we discuss later, the temperature dependence of the trapping indicates that the A^+ and D^- states have different binding energies while our model assumes they are the same. These effects will not significantly change the order of magnitude or field dependence of the capture rates but can introduce errors in the fitted impurity concentrations at the observed level.

7.4.2 G13H (p-type)

In a broad sense, the trapping lengths from the p-type detector are qualitatively similar to those of the n-type. There are, however, some notable differences. We again split the field range studied in 3 regions, although the electron high field range is somewhat limited due to the breakdown of the detector.

1. **Low field (0-1 V/cm):** Again we find short trapping lengths at low field, which rapidly grow as the bias is increased, indicative of cascade Coulomb capture. This is very similar to the n-type detector.
2. **Mid field (1-5 V/cm):** For the holes, it appears the rate of increase in the trapping length begins to slightly slow, however we see the onset of impact ionization around 3-4 V/cm, nearly a factor of 2 lower in field than the n-type detector. For the electrons, we clearly entire a regime dominated by neutral capture, however the change in trapping length reverses direction and actually becomes shorter from 2.5-5V/cm.

G13H (p-type)			Number Density
Process	Type	Trapping Center	from Fit (cm^{-3})
<u>Hole Trapping</u>			
$A^0 + h^+ \rightarrow A^+$	Neutral	A^0	3.6×10^{11}
$D^- + h^+ \rightarrow D^0$	Charged	D^-	1.7×10^{10}
<u>Electron Trapping</u>			
$D^0 + e^- \rightarrow D^-$	Neutral	D^0	5.4×10^{10}
$A^+ + e^- \rightarrow A^0$	Charged	A^+	5.8×10^9

Table 7.3: G13H charge trapping processes for electrons and holes for the the trapping model described in the text. The number densities are the trapping center concentrations from the fits in Fig. 7.7.

3. **High field ($>5 \text{ V/cm}$):** Both carriers show the onset of impact ionization above 5 V/cm , although as we have stated the electron high field range is limited. For the holes, impact ionization overcomes capture and the trapping length changes sign.

Due to early onset of impact ionization and bump in the mid-field electron data, we fit the same model used for the n-type detector, however only up to a fields of 2 V/cm . The fits are shown in Fig. 7.7, with resulting impurity concentrations given in Table 7.3. While the quality of the fits are not as good as with the n-type detector, we still find a decent match with the data up to 2 V/cm . As before, the magnitude of the resulting impurity concentrations are in general agreement with that reported for the boule. We find the total number of acceptors exceeds the donors, as expected for a p-type crystal. The same considerations regarding errors in the model when discussing the n-type data also apply here.

The bump near 5 V/cm in the electron inverse trapping length data is not expected from capture processes dominated by acoustic phonon emission. This shape indicates an energy threshold is crossed and an additional capture process suddenly contributes to the total trapping rate. The field range points towards neutral capture forming overcharged D^- states, enhanced through optical phonon emission. Sundqvist considers such optical phonon capture processes in his “semi-ballistic” approach² and even predicts such a bump in the neutral capture cross section, however there are other difficulties associated with the model at low field. This bump has not been observed in any other detector and appears to be a peculiarity of G13H — perhaps a specific impurity species within this detector has a binding energy which is a near-perfect match to the energy of the average emitted optical phonon.

²Appendix A of [146], a separate model from what we have been discussing so far.

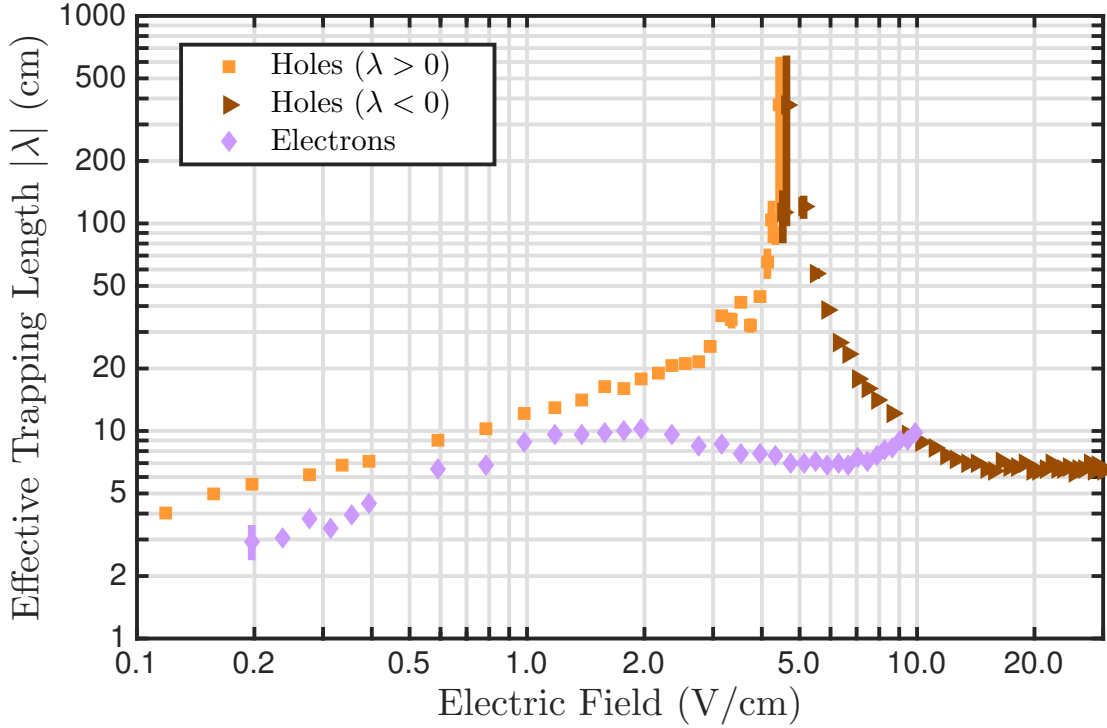


Figure 7.6: Trapping lengths for G13H.

7.4.3 EDELWEISS data

The EDELWEISS collaboration has also studied charge trapping in sub-Kelvin germanium, performing relative trapping length measurements using ^{241}Am for several detectors.[196] Unfortunately, only data from 1 detector spans a large enough electric field range to capture all of the underlying processes. The detector in question is made from p-type germanium, with a net shallow impurity concentration of $\sim 10^{11}$ per cm^{-3} . In addition, the data has been normalized to hole collection at 18 V/cm, in a region where the effects of impact ionization appear to be dominant, which will skew the reported trapping lengths. Our attempt to fit their data using our model is shown in Fig. 7.8, with the resulting impurity concentrations given in Table 7.4. We still find a decent fit to the theory, and note that the total number of acceptors is greater than the donors, matching the p-type nature of the germanium. The onset of impact ionization appears to be observed at higher fields, which we have stated will skew the normalization.

For the other EDELWEISS detectors studied, the range of electric fields is within the region we expect to be dominated by neutral capture forming overcharged states. They observe a linear relationship between drift velocity and trapping length, which is only expected for this neutral capture process, as discussed in Chapter 5. These trends in the EDELWEISS

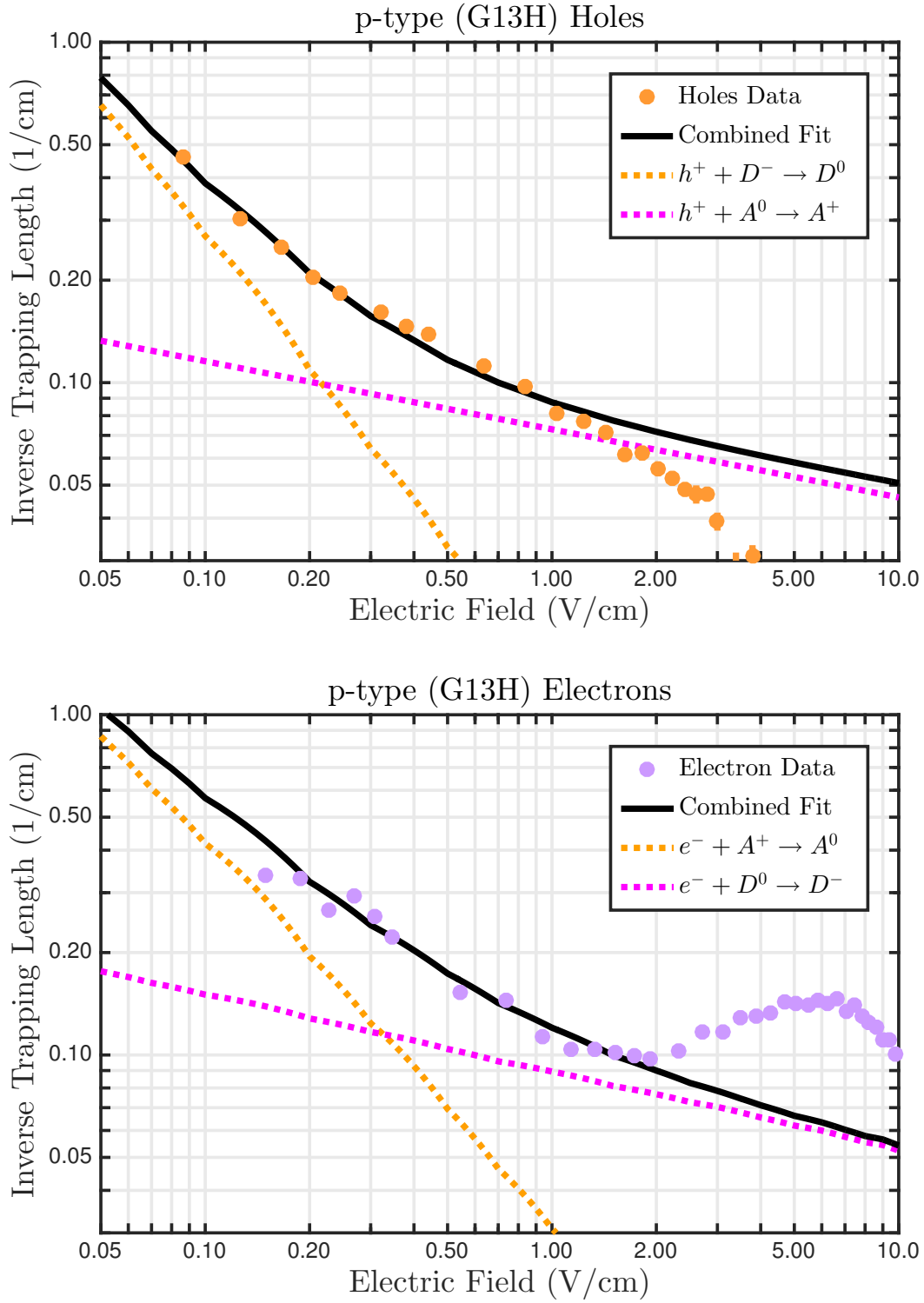


Figure 7.7: Inverse trapping lengths with fits for G13H.

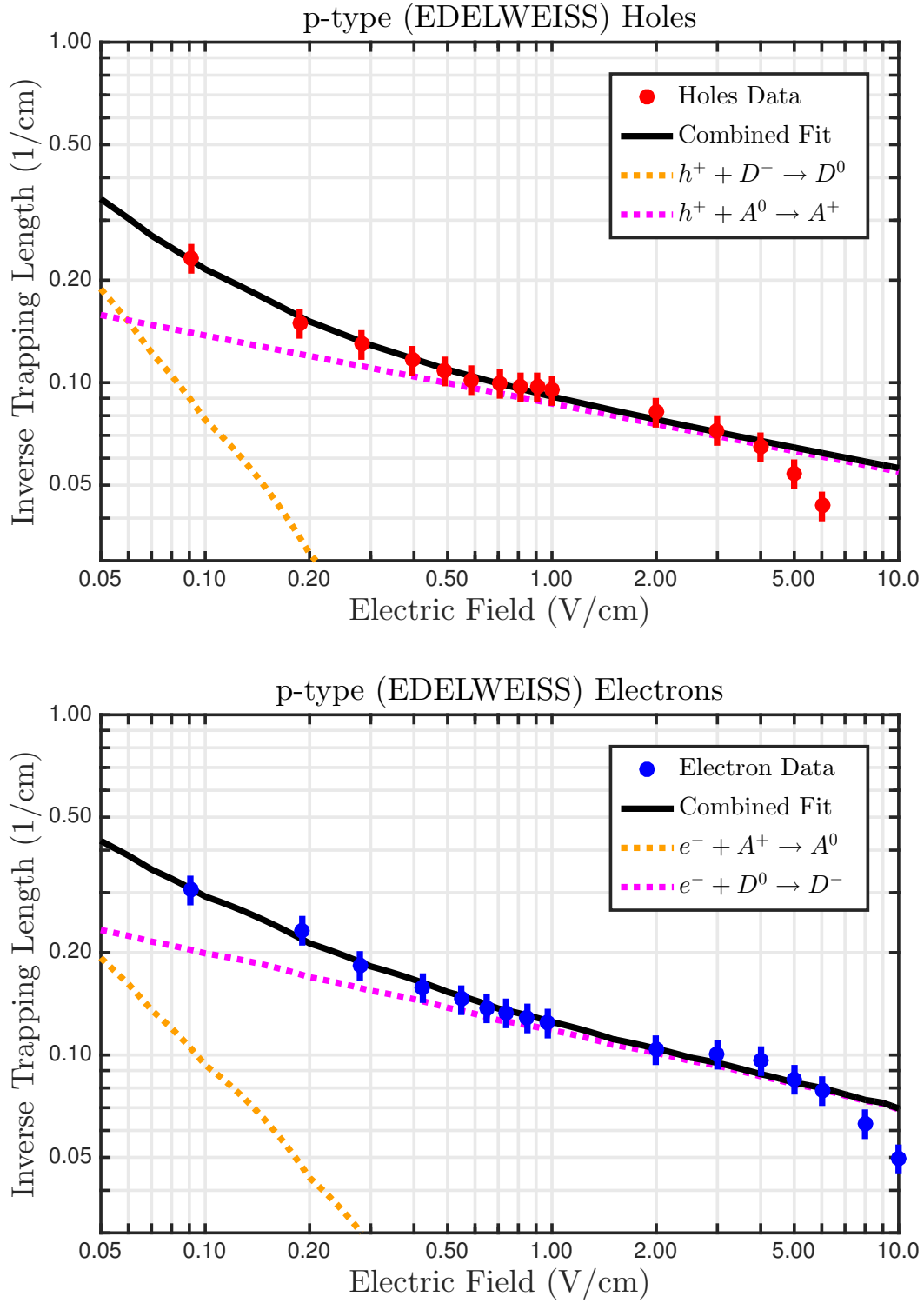


Figure 7.8: Inverse trapping lengths with fit for the p-type EDELWEISS detector.[196]

EDELWEISS detector (p-type)			Number Density
Process	Type	Trapping Center	from Fit (cm^{-3})
<u>Hole Trapping</u>			
$A^0 + h^+ \rightarrow A^+$	Neutral	A^0	4.3×10^{11}
$D^- + h^+ \rightarrow D^0$	Charged	D^-	1.6×10^{10}
<u>Electron Trapping</u>			
$D^0 + e^- \rightarrow D^-$	Neutral	D^0	6.7×10^{10}
$A^+ + e^- \rightarrow A^0$	Charged	A^+	2.2×10^9

Table 7.4: EDELWEISS p-type trapping center concentrations from the fits in Fig. 7.8.

data support our trapping model.

7.5 Impact Ionization

We see clear evidence of both same carrier and opposite carrier impact ionization in our measured drift current pulses. When a secondary current is present, the secondary stop time agrees well with the expected drift time of the secondary carrier. This is most easily seen in Fig. 7.2 — for high field, when opposite carrier impact ionization is present for both electron and hole primary carrier pulses, the *total* pulse width is the same. The width of the primary pulse is the primary carrier drift time, while the width of secondary current is the opposite carrier drift time. These times are independent of the polarity of the electric field. For a constant electric field magnitude, it does not matter if the primary carrier are holes and secondary carrier electrons or the other way around — the sum of their drift times is the same, hence the total pulse widths are the same. This is only expected for opposite carrier impact ionization when the primary carriers have been generated near a detector surface. Events which occur in the bulk will have a more complicated shape.

As we have discussed earlier, the complicated impact ionization pulse shape makes detailed parameter determination unfeasible. We perform an order-of-magnitude estimate of the impact ionization cross section from our data by using the relationship between the impact ionization generation length and cross section.[188] For the n-type holes at high field, the generation length reaches a nearly-constant value of $\lambda = 20$ cm. We assume the number density of impact ionization centers to be $\sim 10^{10} \text{ cm}^{-3}$. The total velocity is ~ 10 x the average drift velocity. We therefore have

$$\langle \sigma \rangle = v_d / (v_{total} \cdot N \cdot \lambda) \approx 1 / (10 \cdot 10^{10} \text{ cm}^{-3} \cdot 20 \text{ cm}) = 5 \times 10^{-13} \text{ cm}^2, \quad (7.2)$$

which is the same order-of-magnitude predicted by Sundqvist. For the p-type holes at high field, the generation length stabilizes around 6 cm, giving $\langle \sigma \rangle = 1.6 \times 10^{-12} \text{ cm}^2$, which is

in the same range. It is difficult to perform a more detailed quantitative analysis of specific processes with this data, however we will see in the next sections that we can still include the effects of impact ionization on the ionization collection efficiency.

7.6 Temperature Dependence

The temperature dependence of the trapping was studied in G13H, the p-type crystal. Electron and hole trapping lengths were measured up to a temperature of 1.5 K. At each temperature, the detector was reset in the standard fashion before applying the external bias. Data taking at higher temperature was challenging, especially at low electric field, due to rapid loss of bulk neutrality, possibly caused by the blackbody radiation present at these temperatures. This prevented us from taking very low bias data as the temperature was increased. Data was limited to a maximum temperature of 1.5 K due to instabilities in the dilution refrigerator above this value.

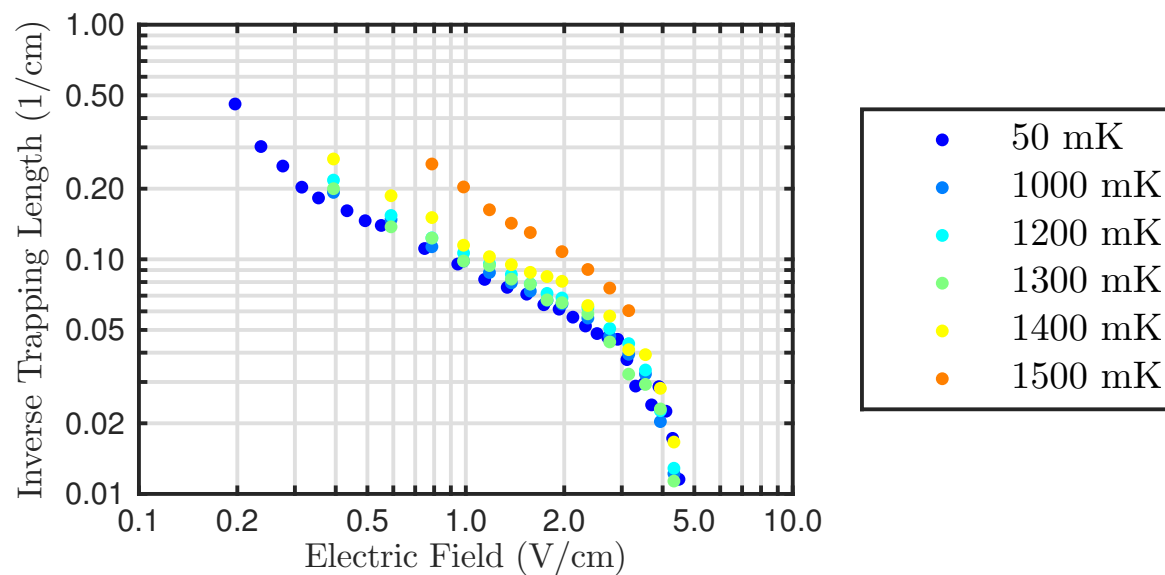
The inverse trapping lengths as a function of electric field and temperature are shown in Fig. 7.9. We see no difference in the trapping lengths up to a temperature of about 1 K. As the temperature is increased, less trapping is observed in the low field electron data while the high field electron data remains unchanged. For the holes, we see an increase in the trapping lengths before the onset of impact ionization.

This behavior can be explained by the very shallow binding energies (~ 0.7 meV) of the overcharged states. At 1.5 K, $k_B T \sim 0.1$ meV. As the temperature is increased, phonons in the high energy tail of the thermal distribution become energetic enough to neutralize overcharged states. The low field reduction in trapping of electrons indicates it is the A^+ states which are neutralized, as they serve as Coulomb capture centers for the electrons. The high field electron trapping is unchanged because this is dominated by neutral D^0 states, whose number density has not been modified by the increased temperature. The thermal neutralization of A^+ states causes an increase in hole trapping due to the higher density of A^0 states, which serve as neutral trapping centers for holes, compared to 50 mK conditions. Unfortunately, we were unable to go low enough in field to prove that the density of D^- states is also unchanged, however this is indirectly supported by the apparent lack of change in the density of D^0 states.

The keen observer will note that the apparent increase in the number of neutral A^0 states would indicate a large fraction of A^+ states are being thermally neutralized. This indicates that greater than 10% of the total acceptors are actually in the A^+ state under sub-Kelvin, grounded, steady state conditions. In the next chapter, we will see this might even be expected with a small modification to Sundqvist's capture rates. For now, it is an interesting observation which warrants further investigation.

The thermal neutralization of ionized impurities by such low energy phonons strongly supports the presence of overcharged states with very shallow binding energies, as there are still no phonons at $T=1.5$ K energetic enough to interact with the "standard" D^+/A^- states. We also know from experience that partial warmups (from base temperature to $T=3-4$ K)

G13H (p-type) Holes



G13H (p-type) Electrons

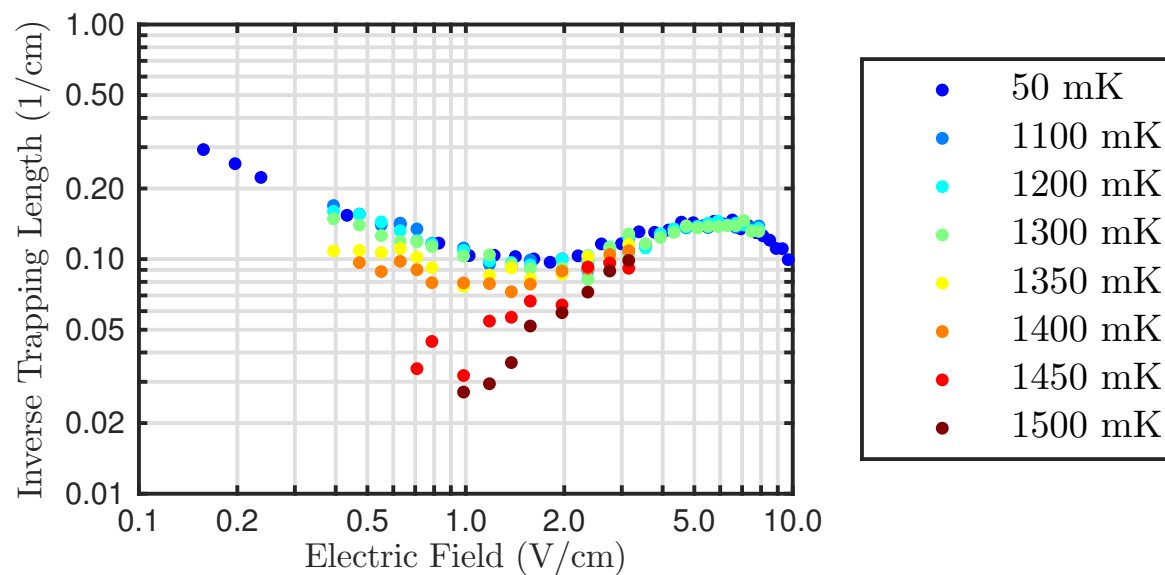


Figure 7.9: Temperature dependence of the hole (top) and electron (bottom) trapping lengths in G13H (p-type).

of the Soudan payload requires all detectors be re-neutralized after returning to sub-Kelvin conditions, as would be expected by this model. The thermal neutralization of overcharged states disrupts the compensation of A^+ and D^- states, leading to a loss of bulk neutrality.

The fact that we appear to see changes in the number of A^+ states, but not the number of D^- states, gives reason to believe that the D^- binding energy must be larger than that of the A^+ state. We expect to see a similar neutralization of D^- states occur at even higher temperatures, possibly around $T=2$ K. As we have stated earlier, these differences in binding energies will change the capture cross sections and resulting impurity concentrations determined by fits to the inverse trapping lengths. Not accounting for these differences may be the reason why the fits to the n-type detector gave a greater concentration of acceptors over donors. A larger D^- binding energy would increase in the number of neutral donors predicted by the fit. Unfortunately, the same temperature measurements were unable to be performed on the n-type detector due to scheduling conflicts.

7.7 Ionization Collection Efficiency

As discussed in the next chapter, the conversion from event pulse amplitudes to recoil energies depends upon the total charge measured by the amplifier. The important quantity is known as the *ionization collection efficiency*, often shortened to charge collection efficiency, and represents the ratio of the measured charge to the true charge produced by the recoil event. Due to the electric field dependence of the trapping and impact ionization processes, this quantity can vary substantially with field. A value less than 1 indicates a net loss of charge, while a value greater than 1 indicates a net excess of charge above that which would be expected for a given recoil event. Obviously it is important for CDMS to have some measure of this parameter.

7.7.1 Relative Calibration

The traditional setup used by CDMS and others to estimate the ionization collection efficiency utilizes a collimated ^{241}Am source placed in close proximity to one face of the detector. Ionization data is taken for various applied bias voltages. The output voltage amplitude of the resulting ~ 60 keV photopeak is measured for each bias voltage. Since the initial recoil energy deposited by the photons does not depend on the electric field, the variations in photopeak amplitude must be due to variations in the ionization collection efficiency.

^{241}Am is a very useful radioactive source to use for ionization collection measurements. The attenuation length of a 60 keV photon in germanium is about 1 mm, indicating that the recoil events are occurring near the surface of the detector. Electron and hole collection can therefore be studied separately by simply switching the polarity of the applied electric field, and the drift distance is known to be essentially the thickness of the detector. A number of other low energy x-ray peaks are also present in the spectrum of ^{241}Am , allowing for cross checks at lower energies if needed.

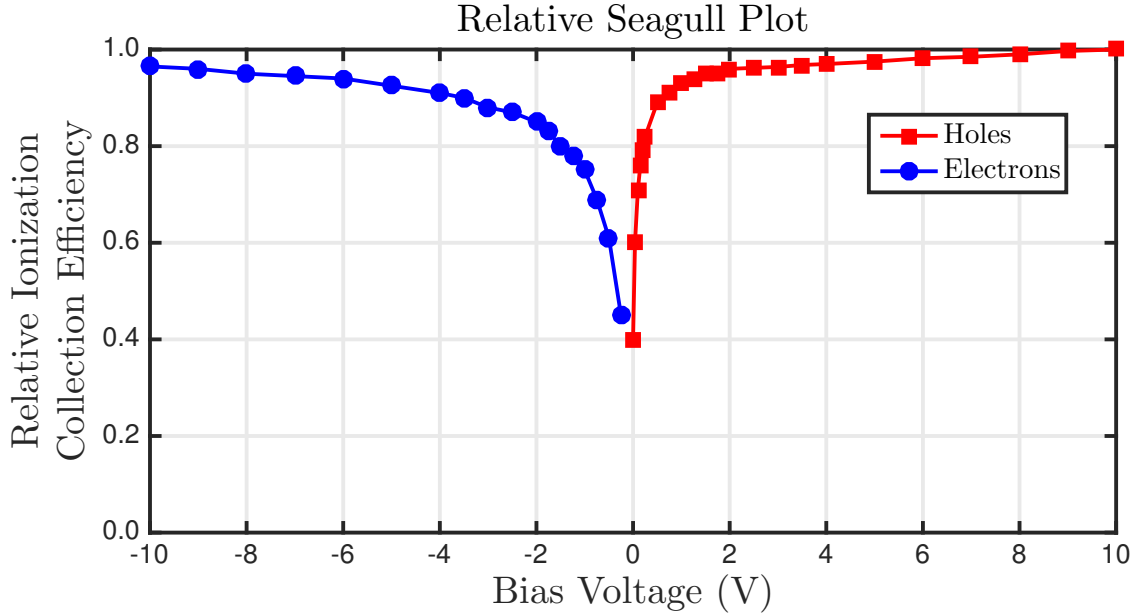


Figure 7.10: The relative calibration (seagull plot) of the ionization collection efficiency for G22Q.

Due to the uncertainties in the component values and parasitic impedances in the charge amplifier circuit, an absolute conversion from number of collected charge carriers to peak output voltage is imprecise. Partially due to a lack of other options, and partially due to human prejudices when observing trends in data, the field-dependent 60 keV photopeak amplitudes are normalized to the field point with the highest measured amplitude. This is almost always the hole collection amplitude at the largest applied voltage. These results, erroneously assumed to represent the true charge collection efficiency, are only a **relative** measurement to this normalization amplitude — the collection as the normalization field is assumed to collect charge with 100% efficiency.

The relative calibration of G22Q using this method is shown in Fig. 7.10. Due to the shape formed by the data, this plot is referred to as the seagull plot within the CDMS collaboration. The relative collection starts out very poor for low bias voltages. As the bias voltage is increased, the efficiency rapidly improves until roughly ± 2 V, after which the rate of growth slows considerably. We see that hole collection is better than electron collection. While the specific collection values differ from crystal to crystal, these general trends are observed in all CDMS and EDELWEISS detectors.

While it obviously makes sense to normalize to the highest measured amplitude if an absolute calibration cannot be performed, it is crucial that it be recognized as only a relative calibration. Historically, it was assumed that since rate of growth in collection efficiency slows significantly at the higher bias voltages, the collection efficiency for holes at a 10 V bias must

be very close to 100%. Comparing our measured trapping lengths to the dimensions of a CDMS detector show that this not the case. This normalization scheme has also led to the assumption that high voltage is equivalent to full charge collection. Our observations of impact ionization are a clear indication that the effective charge collection efficiency can be in excess of 100%, so how has this believe remained?

For many years the CDMS detector bias electronics were limited to supplying ± 10 V, slightly before the onset of impact ionization in a 1 inch crystal. With their thinner geometry, CDMS II oZIPs have reduced sensitivity to both trapping and impact ionization, making their seagull plots less likely to become skewed by improper normalization. Finally, we note that it makes more sense to plot the data with respect to the electric field rather than the bias voltage, as it is the electric field which is fundamental in determining the cross sections of the various underlying transport and capture processes.

7.7.2 Absolute Calibration

With knowledge of the trapping lengths and detector thickness, we can perform an absolute calibration of the ionization collection efficiency. Assume we have a single polarity of carrier drifting from one detector face to the other in a constant electric field, as we have in our optical probe experimental setup. The drift current as a function of time is

$$I(t) = I_0 \exp\left(-\frac{t}{\tau_{trap}}\right) = Q_0 \cdot v_d \cdot \exp\left(-\frac{t}{\tau_{trap}}\right), \quad (7.3)$$

where we have used the fact that the initial current is simply the total charge multiplied by the drift velocity. The charge measured by the amplifier is found by integrating the Ramo-weighted drift current over the drift time. For the planar geometry, the Ramo field is simply $1/L$ throughout the entire crystal, giving

$$\begin{aligned} Q &= \int_{t=0}^{t=\tau_d} \frac{I(t)}{L} dt = \frac{Q_0 \cdot v_d}{L} \int_{t=0}^{t=\tau_d} \exp\left(-\frac{t}{\tau_{trap}}\right) dt \\ &= \frac{Q_0}{L} \int_{z=0}^{z=L} \exp\left(-\frac{z}{\lambda}\right) dz \\ &= Q_0 \cdot \frac{\lambda}{L} \left(1 - \exp\left(-\frac{L}{\lambda}\right)\right), \end{aligned} \quad (7.4)$$

where we have changed from t to z using $z = v_d t$ and $\lambda = v_d \tau_{trap}$. The ionization collection efficiency is thus

$$\frac{Q}{Q_0} = \frac{\lambda}{L} \left(1 - \exp\left(-\frac{L}{\lambda}\right)\right) \quad (7.5)$$

and is independent of the amplitude of the pulse and does not depend on the amplifier gain. By measuring the fundamental parameter determining the collection efficiency —

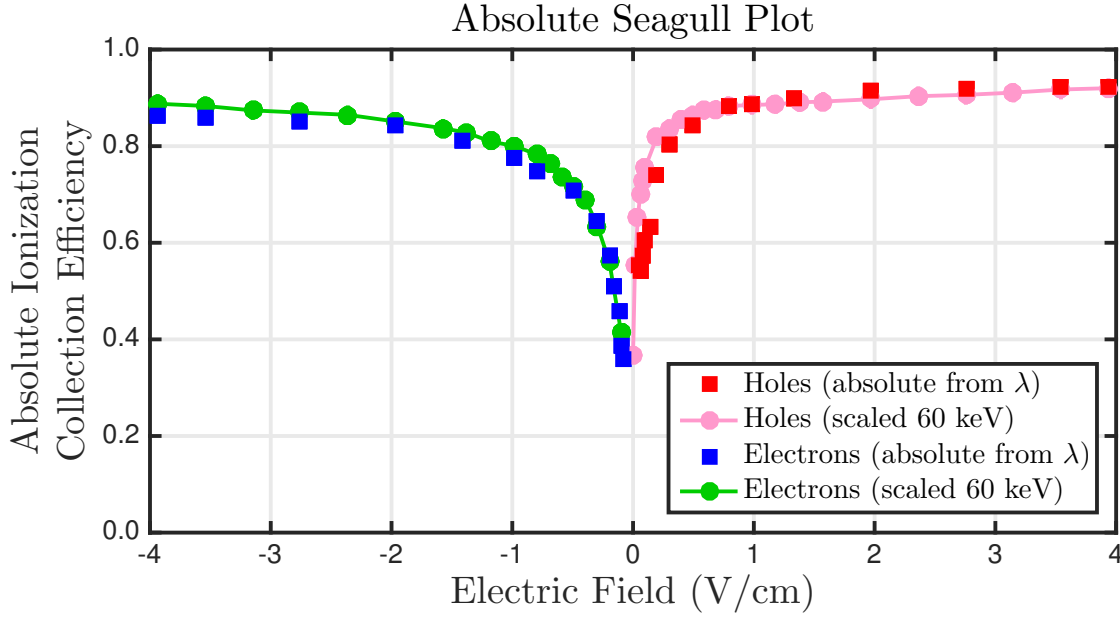


Figure 7.11: Absolute calibration of the ionization collection efficiency based on λ , compared to the scaled ^{241}Am 60 keV recoil data for the n-type detector (G22Q).

the trapping length — we can calculate the true collection efficiency without the errors associated with only measuring the integrated charge. Eq. 7.5 is used to calculate the collection efficiency for each point, creating a properly-calibrated seagull plot.

One might wonder whether the ionization collection of carriers created by the optical probe is somehow fundamentally different from standard recoil events. Using the probe, a relatively large number of electron-hole pairs are created in comparison to typical recoil energies of tens of keV, and they are spread over a large initial area. Recoil events create a smaller number of electron-hole pairs, but with a much higher initial density. Using G22Q, we compare the relative measurements using the ^{241}Am source to the absolute calibration by re-normalizing the data to the absolute hole collection at the same field as the normalization point, shown in Fig. 7.11.

We find good agreement between the absolute calibration and the scaled 60 keV data by using a normalization factor of 0.92 for the 10 V (~ 4 V/cm) ^{241}Am holes amplitude. The results show that the ionization collection efficiency under the standard CDMS biasing conditions of ~ 1 V/cm is between 80%-90%. Even though the physical distribution of charge carriers produced by the optical probe is different in nature from the ^{241}Am photons, they clearly obey the same transport physics responsible for determining the ionization collection efficiency. One can also conclude that charge carrier loss through recombination due to the self-shielding effects of the initial charge ball do not play a dominant role.

While there are some small discrepancies between the two datasets, we believe they

are likely due to other effects. Although their absorption length in Ge is small, the 60 keV photons do slightly penetrate the bulk germanium and are not true surface events. In addition, the proximity of the 60 keV events to the surface of the crystal may result in some loss of charge carriers due to back diffusion, particularly at lower fields. As the standard CDMS electronics only measure the total charge produced, these effects would be masked in the ^{241}Am data. Since the ionization collection efficiency of the optical probe data depends only on the pulse shape and not the amplitude, it is not influenced by these effects. In fact, the un-normalized total charge collected by the high speed amplifier shows a much different field dependence than an analysis using only the trapping lengths. This difference is solely due to back diffusion effects and can be ignored. These differences, particularly if compared at the same field but using a variety of photon sources with varying penetration depths, could allow CDMS to better probe the size of the dead layer near the crystal surfaces.

Due to the limited bias range of ± 10 V in the standard CDMS electronics, ^{241}Am data was not taken above ~ 4 V/cm. Based on the agreement between the absolute calibration and scaled relative calibration, it was deemed unnecessary to check the validity of the ionization collection efficiency measurements at higher fields, nor was the data taken for the p-type detector due to time constraints. In addition, the use of a single trapping length to predict the ionization collection efficiency at higher fields is invalid because of the effects of impact ionization. Nevertheless, Fig. 7.12 shows the ionization collection efficiency as determined by the effect trapping lengths for both n-type and p-type detectors over the full field range studied.

7.7.3 Including Impact Ionization

As we have previously discussed, impact ionization of shallow impurity states is expected to influence charge collection for electric fields above ~ 4 V/cm. We have directly observed changes in the shape of the drift current signal at higher bias consistent with impact ionization processes. If impact ionization only produced additional charges of the same polarity as the primary carriers, the use of a single “trapping” length (which may be positive or negative) to predict the ionization collection efficiency would be warranted. The creation of additional carriers of the opposite polarity complicates the situation, making Eq. 7.5 invalid. Not only does the primary pulse no longer follow a simple exponential shape, an additional secondary current signal persists after the primary pulse as the opposite polarity carriers drift back to the illuminated detector face. These additional charge carriers need to be included in the calculation of the ionization collection efficiency.

The raw peak output voltage of the high speed charge amplifier (before taking the time derivative) is in fact proportional to the total integrated charge. Using this as a measure of the ionization collection efficiency, however, has the same normalization problems as with radioactive sources and can only be a relative measurement. How, then, should we attempt to perform an absolute calibration which includes impact ionization? Recall that in the case of perfect charge collection, the drift current pulse of a single polarity of carrier moving from one detector face to the other will be a square pulse with a width equal to the drift time. The

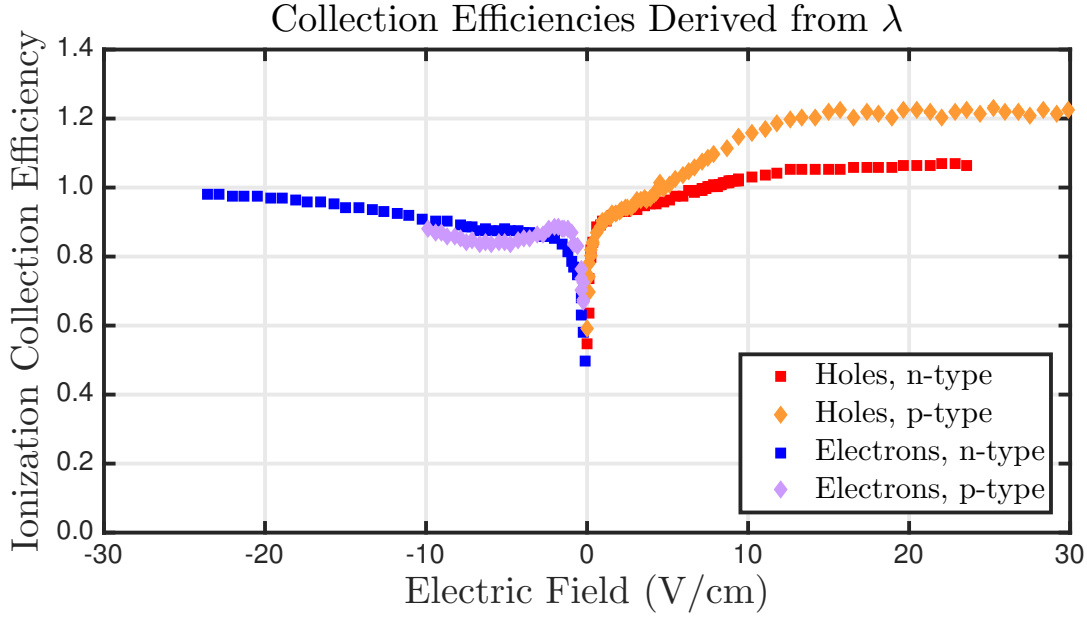


Figure 7.12: The calculated ionization collection efficiency for the n-type (G22Q) and p-type (G13H) detectors using the effective trapping length λ , over the entire field range studied for both detectors. Note that these efficiencies are misleading as they do not include the full effects of impact ionization.

area under the amplified, output drift current pulse is proportional to the total produced charge, as the collection efficiency is 100%. If we can predict this area for each bias point it can be used to individually normalize the integral of the *measured* drift current signal, giving us the charge collection efficiency including impact ionization.

We proceed as follows:

1. The trapping time for each pulse has been found through an exponential fit to the drift current signal within a window filled by some portion of the primary drift signal. The determination of the trapping time is independent of the amplitude of the pulse, hence the fitting window was chosen to begin at a time $t' = 0$. The amplitude of the exponential from the fit is time-shifted from the true amplitude at the start of the pulse. Let Δt_{win} be the time difference between the true start time, which we take to occur at $t = 0$, and the start time of the exponential fit window. We have

$$\begin{aligned}
 I_{meas} &= A \exp\left(-\frac{t'}{\tau_{trap}}\right) = A \exp\left(-\frac{t - \Delta t_{win}}{\tau_{trap}}\right) \\
 &= A \exp\left(\frac{\Delta t_{win}}{\tau_{trap}}\right) \exp\left(-\frac{t}{\tau_{trap}}\right) = B \exp\left(-\frac{t}{\tau_{trap}}\right).
 \end{aligned} \tag{7.6}$$

We see the true initial current as determined by the fit is given by

$$B = A \exp\left(\frac{\Delta t_{win}}{\tau_{trap}}\right), \quad (7.7)$$

where Δt_{win} is fixed in software as a constant time offset from the predicted start time of the pulse obtained during the drift velocity measurement. We prefer to use this method rather than attempting to find the initial amplitude directly from the measured drift current pulse as it is smeared by the risetime of the amplifier. We take the error on B to be

$$\left(\frac{\Delta B}{B}\right)^2 = \left(\frac{\Delta A}{A}\right)^2 + \left(\frac{\Delta t_{win} \cdot \Delta \tau_{trap}}{\tau_{trap}^2}\right)^2. \quad (7.8)$$

2. We know the drift time of the primary carrier from the drift velocity measurement. As B represents the amplified initial current value, which does not change over the course of the pulse in the case of 100% collection efficiency, the amplified area representing the true collected charge is

$$Q_0 = B \cdot v_d, \quad (7.9)$$

with error

$$\left(\frac{\Delta Q_0}{Q_0}\right)^2 = \left(\frac{\Delta B}{B}\right)^2 + \left(\frac{\Delta v_d}{v_d}\right)^2, \quad (7.10)$$

giving an independent normalization for each bias point.

3. We now sum the entire drift current pulse to obtain the measured, amplified total charge

$$Q = \sum_{\text{trace}} I_{meas}(t) \cdot \Delta t, \quad (7.11)$$

where Δt is the sampling time between measured drift current points.

4. Finally, we take the ratio Q/Q_0 of the total measured charge to the predicted normalization. The error is given by

$$\Delta\left(\frac{Q}{Q_0}\right) = \frac{Q}{Q_0^2} \cdot \Delta Q_0. \quad (7.12)$$

The amplifier gain is common for both the measured and predicted charge and cancels when taking the ratio. We have approximated the absolute charge collection efficiency including the effects of impact ionization. The integrated calibrations of the n-type and p-type detector are shown in Fig. 7.13. We see good agreement between the absolute calibration using only the trapping lengths and the integrated calibration at low field, before the onset of impact ionization. This implies that our integrated method is well-calibrated. At higher fields, particularly when opposite carrier impact ionization begins to appear, the integrated calibration depicts a higher ionization collection efficiency as expected.

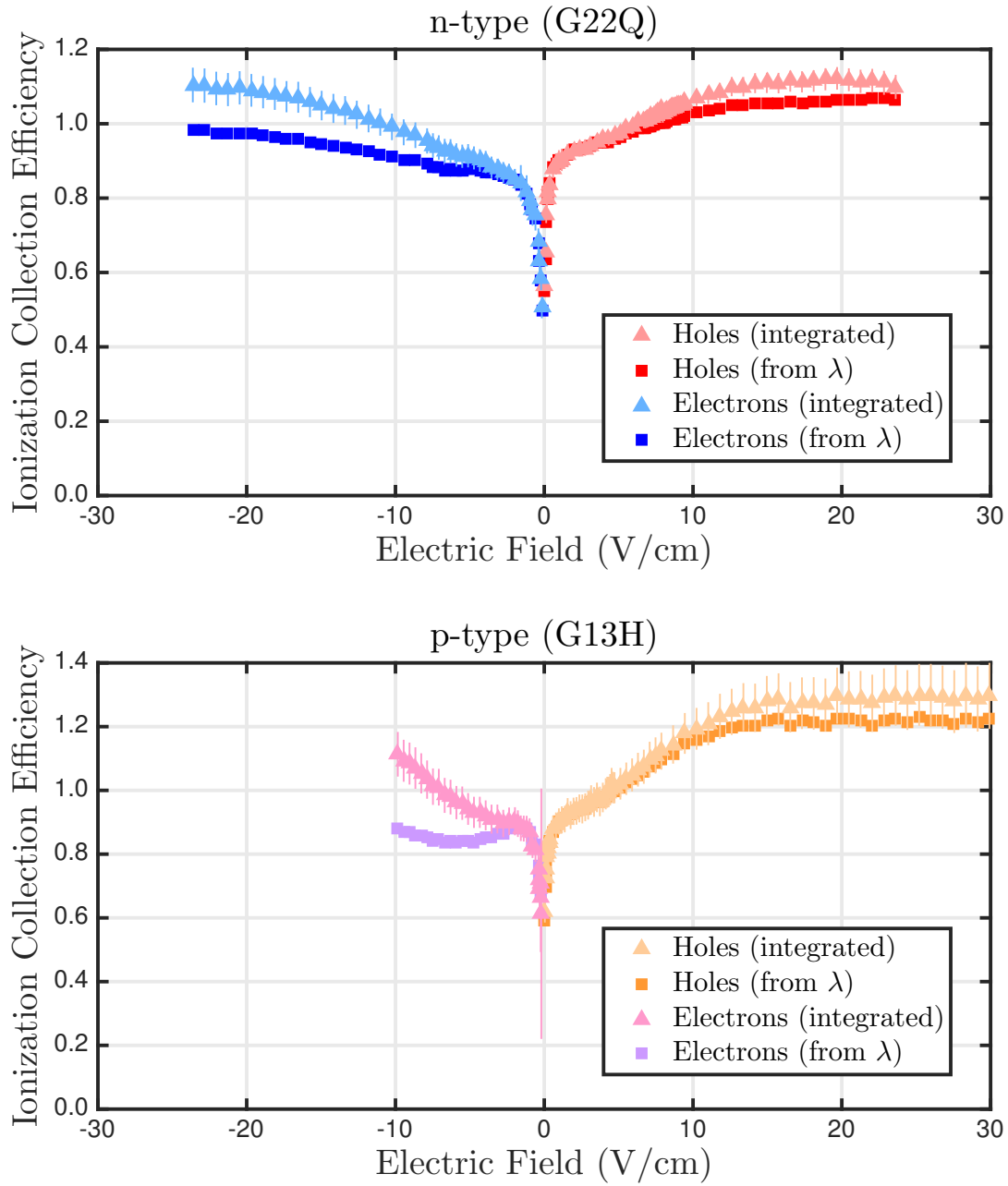


Figure 7.13: The ionization collection efficiency as calculated using the integrated method, which includes the effects of impact ionization, in comparison to the collection efficiency calculated using only the effective trapping length of the primary pulse. The two methods begin to differ at higher fields due to the onset of impact ionization.

7.7.4 Position Dependence

So far, the ionization collection efficiencies shown have been for a single primary carrier type drifting through the full one inch thickness of the germanium. Actual recoil events will occur throughout the volume of the detector, and the total charge measured by the amplifier will be a combination of the electron and hole signals. As the effects of trapping and impact ionization depend on the drift left of each carrier type, there will be a z-position dependence to the charge collection.

A recoil event at a depth of z_0 produces some number of electron-hole pairs, which drift in opposite directions due to the applied field. The electrons drift towards the electrode at $z = L$, while the holes drift towards the electrode at $z = 0$. In the case that the evolution of each polarity of carrier may be described its own effective trapping length (which may be positive or negative), we have the following relations defining the drift distances, drift times and position:

$$\begin{aligned}
 L_h &= z_0, \\
 \tau_{dh} &= \frac{L_h}{v_{dh}} = \frac{z_0}{v_{dh}}, \\
 z_h &= z_0 - v_{dh} \cdot t, \\
 L_e &= L - z_0, \\
 \tau_{de} &= \frac{L_e}{v_{de}} = \frac{L - z_0}{v_{de}}, \\
 z_e &= z_0 + v_{de} \cdot t,
 \end{aligned} \tag{7.13}$$

where the h and e subscripts refer to the relevant hole and electron properties.

Since the Ramo field of the planar geometry is constant with z-position, both carrier polarities contribute equally to the measured drift current signal. The ionization collection efficiency is thus

$$\begin{aligned}
 \frac{Q}{Q_0} &= \frac{1}{L} \left(v_{dh} \int_{t=0}^{t=\tau_{dh}} \exp\left(-\frac{t}{\tau_{trap,h}}\right) dt + v_{de} \int_{t=0}^{t=\tau_{de}} \exp\left(-\frac{t}{\tau_{trap,e}}\right) dt \right) \\
 &= \frac{1}{L} \left(\int_{z_h=z_0}^{z_h=0} \exp\left(-\frac{z_0 - z_h}{\lambda_h}\right) dz_h + \int_{z_e=z_0}^{z_e=L} \exp\left(-\frac{z_e - z_0}{\lambda_e}\right) dz_e \right) \\
 &= \frac{\lambda_h}{L} \left(1 - \exp\left(-\frac{z_0}{\lambda_h}\right) \right) + \frac{\lambda_e}{L} \left(1 - \exp\left(-\frac{L - z_0}{\lambda_e}\right) \right).
 \end{aligned} \tag{7.14}$$

Note that for $z_0 = 0$ or $z_0 = L$, we recover Eq. 7.5 for the appropriate carrier.

The ionization collection efficiency as a function of z-position, for various electric field strengths, is shown in Fig. 7.14 for G22Q and Fig. 7.15 for G13H. This efficiency has been

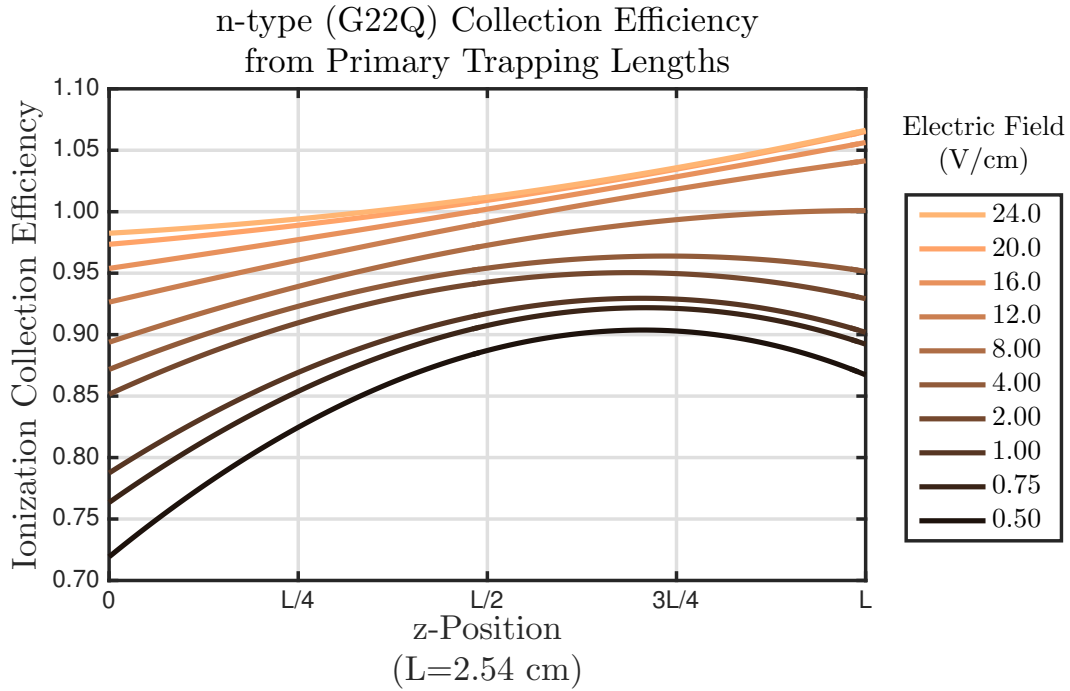


Figure 7.14: Ionization collection efficiency as a function of event position for various electric fields in G22Q, as determined by the primary trapping lengths.

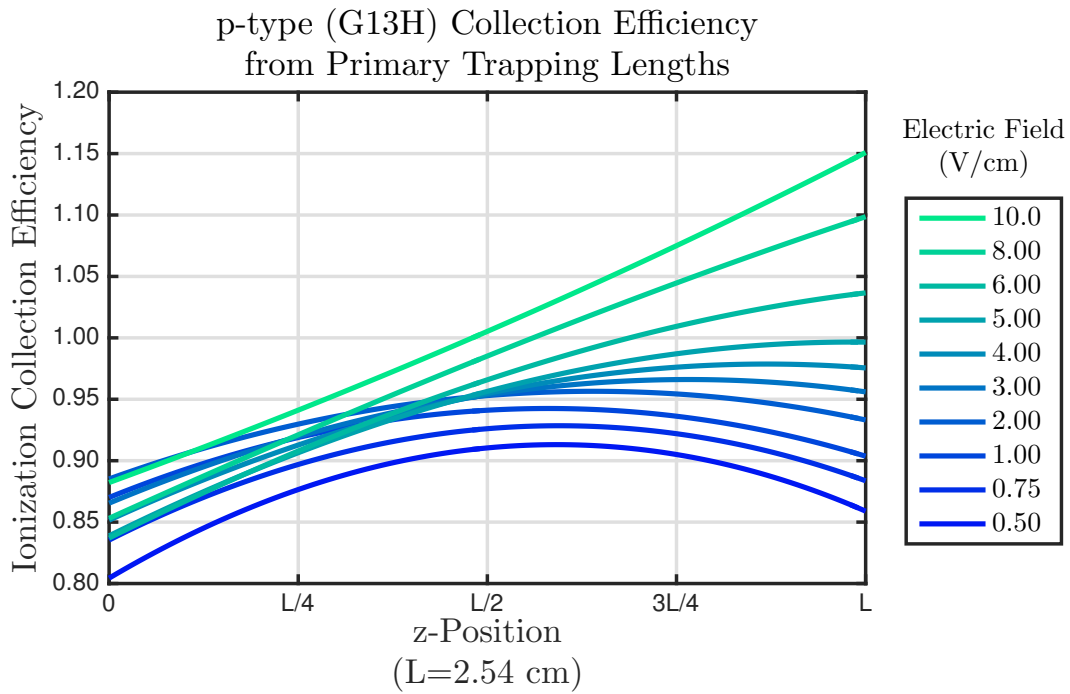


Figure 7.15: Ionization collection efficiency as a function of event position for various electric fields in G13H, as determined by the primary trapping lengths.

calculated using the primary trapping length, obtained from the exponential fit to the primary carrier pulse. To include the effects of impact ionization, we use the effective trapping lengths derived from the integrated ionization collection efficiency in Fig. 7.16 and Fig 7.17. Using the effective trapping lengths, we find differences in charge collection on the order of $\sim 5\%$ (for fixed field) which will broaden the charge energy resolution for uniformly distributed events of a fixed energy. In the case that impact ionization is not dominant (both λ_h and λ_e are positive), it is easy to show that the collection efficiency is maximized at position

$$z_{max} = \frac{L}{1 + \frac{\lambda_e}{\lambda_h}}. \quad (7.15)$$

We see that if the hole trapping length is much greater than the electron trapping length, the point of maximum collection approaches $z = L$ and holes become the primary carriers. The opposite holds true (maximum collection point approaches $z = 0$) if the electron trapping length is much greater than that of the holes. In the case that one of the trapping lengths is negative, there may be no local maximum (or minimum) within the detector, other than at the contacts.

The strange behavior of the p-type detector at its higher field values demonstrates the difficulties in using the effective trapping length, which ignores opposite carrier impact ionization. We make a rough attempt at approximating these effects by determining what single-carrier trapping lengths would be needed to give the integrated ionization collection efficiencies of Fig. 7.13. As there is no closed-form solution when attempting to invert Eq. 7.5, this is performed through numerical methods. The results are plotted in Fig. 7.16 and Fig. 7.17.

7.8 A Simple Model for iZIP Collection Efficiency

The iZIP detectors have a complicated electrode geometry which causes large deviations from the constant field approximations we have been using. While the bulk field is relatively uniform, the interleaved ionization electrodes and grounded phonon sensors produce an electric field which varies with both depth and radial position within the first ~ 1 mm near the contacts. The Ramo potential is also nontrivial and is not directly proportional to the electric field. About half of the ionization signal comes from bulk transport, while the other half is only from the charge carriers which reach the surface. We will now develop a simple model by which to approximate ionization collection efficiency in an iZIP detector using the measured properties of G22Q and G13H.

7.8.1 Bias Voltage, Electric Field, and the Ramo Potential

iZIP detectors are typically biased in a ± 2 V configuration, while the phonon sensors serve as electrical ground. While the total potential across the crystal is 4 V, the electric field in the bulk is not simply V/L . Due to the proximity of the ionization and phonon rails, each

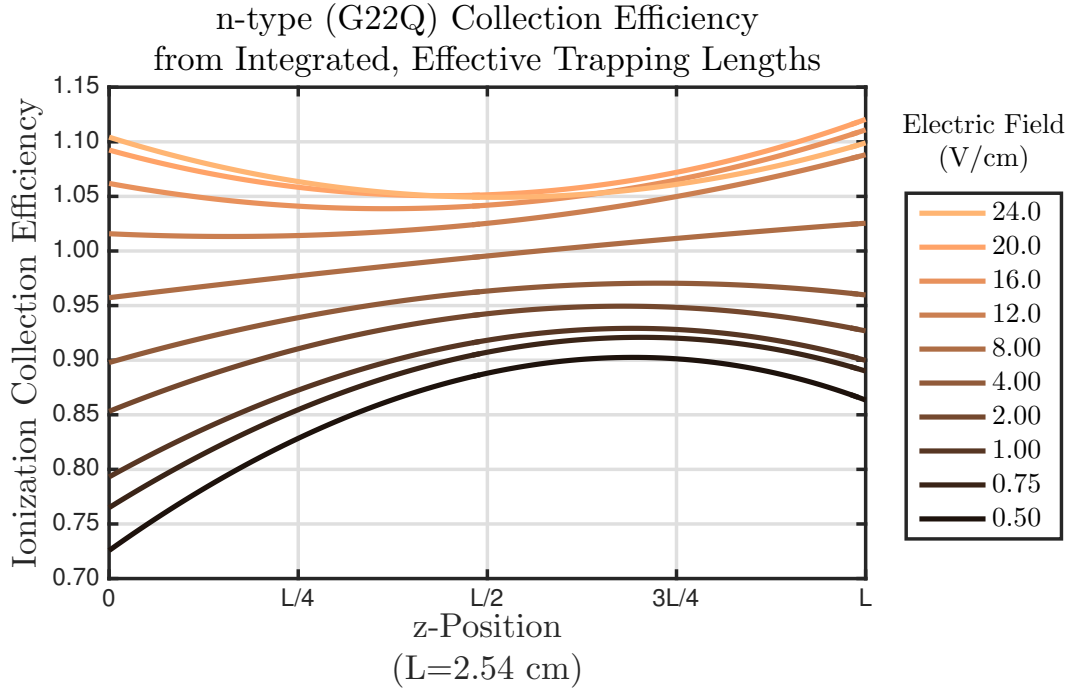


Figure 7.16: Ionization collection efficiency as a function of event position for various electric fields in G22Q, approximated using the integrated, effective lengths.

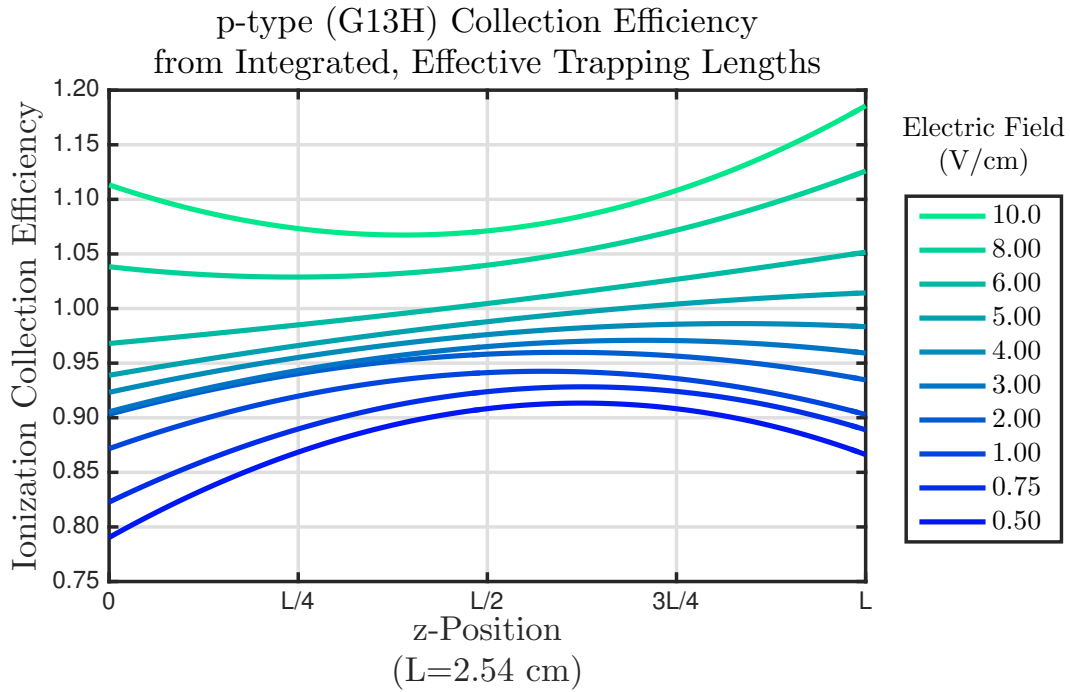


Figure 7.17: Ionization collection efficiency as a function of event position for various electric fields in G13H, approximated using the integrated, effective lengths.

face has an approximate potential of half the applied value, hence the bulk field is half of what would be typically expected. We plot the results in terms of the symmetric bias voltage rather than electric field, so as to avoid confusion as to how the detector is biased.

We model the Ramo potentials for the side 1 ($z = 0$) and side 2 ($z = L$) in the following way, where we will focus on the side 1 potential to understand the shape. Recall that Ramo's theorem dictates we hold the voltage of the side one electric at unit potential, while grounding all other conductors to calculate the weighting field. Far away from side 1, the effective potential appears to be only 0.5 instead of 1. We approximate the Ramo potential as increasing linearly away from $z = L$ towards $z = 0$ as in the planar case, however only reaching a value of 0.5 at $z = L_1 = L - L_2$. At this point we are close enough to the side 1 electrode to distinguish it from the grounded phonon rail, and the Ramo potential rapidly increases from 0.5 to 1.

7.8.2 Approximated Ionization Collection Efficiency

In the iZIP design, we have $L_1 \approx 1$ mm. In this region, both the Ramo potential and electric field rapidly increase in value. As this is only $\sim 4\%$ of the detector thickness, we assume the contribution to the signal is the charge which remains at $Z = L_1$, without any modification due impact ionization or trapping, weighted by a factor of 0.5. With this simple model, Eq.7.14 becomes

$$\begin{aligned} \frac{Q_1}{Q_0} = & \frac{1}{2} \left[\frac{\lambda_h}{L_2} \left(1 - \exp \left(-\frac{z_0}{\lambda_h} \right) \right) + \frac{\lambda_e}{L_2} \left(1 - \exp \left(-\frac{L - z_0}{\lambda_e} \right) \right) \right] \\ & + \frac{1}{2} \exp \left(-\frac{z_0}{\lambda_h} \right) \end{aligned} \quad (7.16)$$

for side 1, where we have assumed hole collection. The side 2 signal will be equivalently given by

$$\begin{aligned} \frac{Q_2}{Q_0} = & \frac{1}{2} \left[\frac{\lambda_h}{L_2} \left(1 - \exp \left(-\frac{z_0}{\lambda_h} \right) \right) + \frac{\lambda_e}{L_2} \left(1 - \exp \left(-\frac{L - z_0}{\lambda_e} \right) \right) \right] \\ & + \frac{1}{2} \exp \left(-\frac{L - z_0}{\lambda_e} \right) \end{aligned} \quad (7.17)$$

The results of this model, using trapping lengths derived from the integrated ionization collection efficiency of G22Q, are shown in Fig. 7.18. Note that under the standard iZIP operating mode of ± 2 V, the collection efficiency is only $\sim 90\%$.

7.8.3 Approximated Charge z-Partition

The iZIP's ability to measure a charge signal on both surfaces of the crystal serves as a powerful discrimination tool against problematic surface events.[123] The primary discriminator

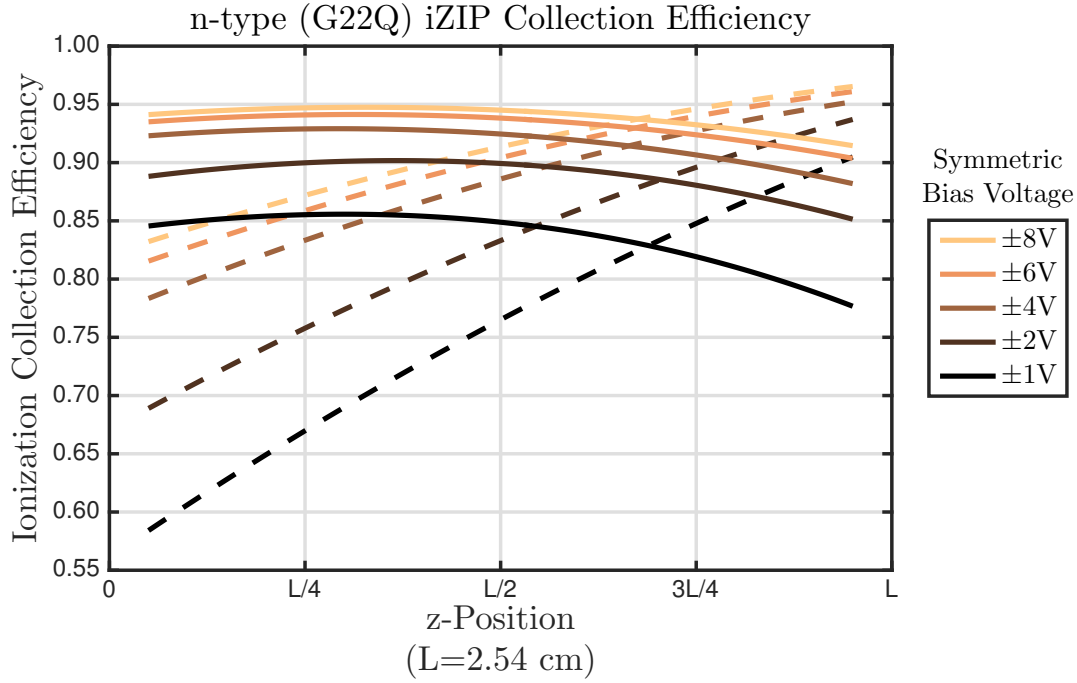


Figure 7.18: The ionization collection efficiency of G22Q using the simplified iZIP model described in the text. Side 1 ($z=0$, hole) collection is shown with solid lines, while side 2 ($z=L$, electron) collection is shown with dashed lines.

is the charge z -partition, or $qzpart$, defined as

$$qzpart = \frac{Q_{s1} - Q_{s2}}{Q_{s1} + Q_{s2}}, \quad (7.18)$$

where Q_{s1} and Q_{s2} are the total charge collected on side 1 and side 2, respectively. Events which occur within the bulk of the crystal will produce a signals of approximately equal magnitude on each side and will have $qzpart \approx 0$. Due to the tangential field in this region, surface events will result in a charge signal being recorded on only one side as one polarity of carrier is collected to the adjacent phonon rails rather than drifting through the bulk to the opposite face. Surface events will therefore have $qzpart \approx \pm 1$.

The z -dependence of the ionization collection efficiency seen in Fig. 7.18 will clearly influence appear of the z -partition. We plot our predicted $qzpart$ values as a function of z , using our iZIP ionization collection efficiencies for Q_{s1} and Q_{s2} , in Fig. 7.19. Note that we do not include the 1 mm of depth on either side of the detector faces, as in this region we assume the charge to be fully collected and only on on side ($qzpart = \pm 1$). We see under standard iZIP biasing that the spread in $qzpart$ is quite narrow even though the electron collection is relatively poor for events occurring near the hole collection face. These results

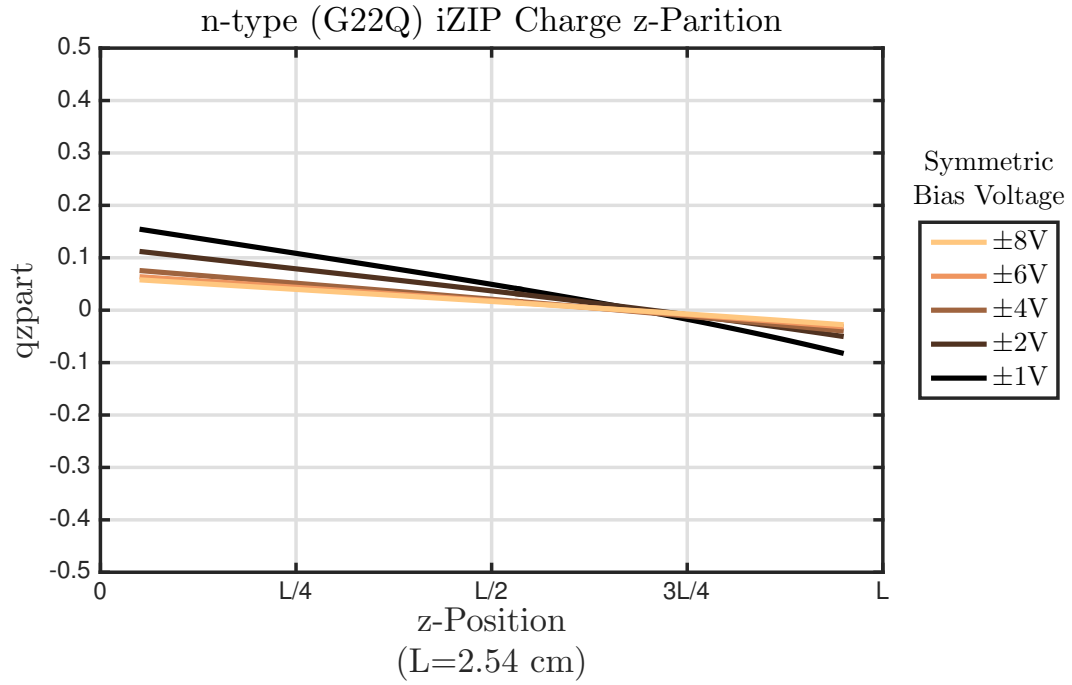


Figure 7.19: The predicted charge z-partition using the G22Q iZIP model collection efficiency. Note that the $qzpart$ axis is somewhat zoomed in, ranging from ± 0.5 rather than ± 1 . Events which occur within ~ 1 mm of either detector face will have $qzpart = \pm 1$ and are not shown.

demonstrate the power in using $qzpart$ to reject surface events and show the distribution is relatively unaffected by trapping to first order.

Chapter 8

Implications for CDMS Detectors

8.1 Space Charge and the Two Species, Two State Model

When a CDMS detector is cooled from room temperature to sub-Kelvin conditions, the bulk germanium transitions from being filled with thermal charge carriers and phonons to a state of freeze out. The absence of free charge carriers and thermal phonons of appreciable energy causes the occupation of impurity states to remain fixed without the application of an external stimulus, such as ionizing radiation. Under these conditions, it does not make sense to speak of thermal equilibrium with regards to the charge state of the detector. Thermal processes do not change the number densities of the impurity charge states.

Immediately after cooldown, the ionization collection efficiency of the detector is terrible. This is believed to be due to a large residual population of “standard” D^+ and A^- impurity states, which are present in abundance at higher temperature. The temperature of the crystal decreases rapidly enough for these charged states to remain initially frozen in, disrupting charge transport. It was found empirically that by leaving the detector electrodes grounded and generating large number of free carriers, acceptable ionization collection could be achieved. The free carriers are captured by the D^+/A^- impurities, converting them to neutral D^0/A^0 states. As there are no significant processes which lead to the *production* of D^+/A^- states under sub-Kelvin conditions, after the initial neutralization these states no longer contribute.

Traditionally, it was believed that after the initial LED “bake” all impurities are neutralized, leading to acceptable ionization collection conditions. The presence of a large neutral capture cross section which produces *overcharged* states, however, results in a sizable population of overcharged D^-/A^+ states, which compensate one another to result in approximate bulk neutrality of the crystal. Based on the initial work of Sundqvist[146], we now develop a two species (donor/acceptor), two state (neutral/overcharged) space charge model describing the behavior of CDMS detectors.

8.1.1 Grounded, Steady State Condition

In the absence of thermal processes, we need a reliable procedure for preparing detectors in a known state. By grounding the electrodes and creating a large number of free carriers, we eventually reach the *grounded, steady state condition*. Consider free electron-hole pairs are generated at a rate g . Electrons will be capture by neutral donors and overcharged acceptors, while holes will be captured by neutral acceptors and overcharged donors. Let n and p be the free electron and hole densities, as in previous chapters. We have

$$\begin{aligned}\dot{n} &= g - c_{D0} N_{D0} n - c_{A+} N_{A+} n, \\ \dot{p} &= g - c_{A0} N_{A0} p - c_{D-} N_{D-} p,\end{aligned}\tag{8.1}$$

where $c = \langle \sigma v \rangle$ are the capture rates per unit density for the capture center indicated by the subscript. In steady state conditions, we have $\dot{n} = \dot{p} = 0$. Solving for the carrier densities, we have

$$\begin{aligned}n &= \frac{g}{c_{D0} N_{D0} + c_{A+} N_{A+}}, \\ p &= \frac{g}{c_{A0} N_{A0} + c_{D-} N_{D-}}.\end{aligned}\tag{8.2}$$

We can also write the time derivatives of the impurity densities. For the overcharged states, we have

$$\begin{aligned}\dot{N}_{D-} &= c_{D0} N_{D0} n - c_{D-} N_{D-} p, \\ \dot{N}_{A+} &= c_{A0} N_{A0} p - c_{A+} N_{A+} n.\end{aligned}\tag{8.3}$$

Note that $\dot{N}_{D-} = -\dot{N}_{D0}$ and $\dot{N}_{A+} = -\dot{N}_{A0}$. By setting the time derivatives equal to zero and using Eq. 8.2 for n and p , we arrive at

$$\begin{aligned}\frac{c_{D0} N_{D0}}{c_{D0} N_{D0} + c_{A+} N_{A+}} &= \frac{c_{D-} N_{D-}}{c_{A0} N_{A0} + c_{D-} N_{D-}}, \\ \frac{c_{A0} N_{A0}}{c_{A0} N_{A0} + c_{D-} N_{D-}} &= \frac{c_{A+} N_{A+}}{c_{D0} N_{D0} + c_{A+} N_{A+}}.\end{aligned}\tag{8.4}$$

The total number of donors may be written as $N_D = N_{D0} + N_{D-}$ and the total number of acceptors as $N_A = N_{A0} + N_{A+}$. We only consider neutral and overcharged states as we have shown that the “standard” D^+/A^- states are not regenerated after their initial neutralization under sub-Kelvin conditions. We define the fraction of ionized impurities to be

$$\begin{aligned}\zeta_D &= \frac{N_{D-}}{N_D}, \\ \zeta_A &= \frac{N_{A+}}{N_A}.\end{aligned}\tag{8.5}$$

Process	Trapping Center	Rate	Cross Section (cm ²)
$D^- + h^+ \rightarrow D^0$	D^-	c_{D-}	5×10^{-12}
$D^0 + e^- \rightarrow D^-$	D^0	c_{D0}	4×10^{-12}
$A^+ + e^- \rightarrow A^0$	A^+	c_{A+}	6×10^{-11}
$A^0 + h^+ \rightarrow A^+$	A^0	c_{A0}	4×10^{-13}

Table 8.1: Approximate capture cross sections used in Fig. 8.1

Using these definitions in Eq. 8.4, we can solve for ζ_D in terms of ζ_A and vice-versa to find

$$\begin{aligned}\zeta_D &= \frac{c_{D0} c_{A0} (1 - \zeta_A)}{c_{D0} c_{A0} (1 - \zeta_A) + c_{A+} c_{D-} \zeta_A}, \\ \zeta_A &= \frac{c_{D0} c_{A0} (1 - \zeta_D)}{c_{D0} c_{A0} (1 - \zeta_D) + c_{A+} c_{D-} \zeta_D}.\end{aligned}\tag{8.6}$$

These equations have an inverse relationship. As $\zeta_D \rightarrow 1, \zeta_A \rightarrow 0$ and as $\zeta_D \rightarrow 0, \zeta_A \rightarrow 1$. **Complete neutrality of the impurities at equilibrium is not possible.** Neutralizing one species of impurity causes the other to become fully ionized. Approximate bulk neutralization must therefore arise from the compensation of overcharged donors and acceptors, realized mathematically through the relationship

$$\zeta_D \approx \left(\frac{N_A}{N_D} \right) \zeta_A.\tag{8.7}$$

While this result can be used with Eq. 8.6 to obtain an approximate expression for ζ_D only in terms of the capture rates, the form is not particularly enlightening. Adding an additional complication, the capture rates depend on the electric field, which will be determined by the residual ionized impurities as the detector electrodes are grounded. A self-consistent model which couples overcharged impurity concentrations and field-dependent capture rates to Poisson's equation is needed. An attempt at solving such a model using a finite element solver has been performed by Sundqvist[146], however the problem is computationally difficult and there is some confusion as to how to specify the boundary conditions at the electrodes. In any case, we expect a small residual electric field to be generated due to slight imbalances in the number of overcharged impurities.

Experimentally, we find that detectors placed in their grounded, steady state conditions retain a residual electric field of a few tens of mV/cm. At such weak fields, the total velocities of electrons and holes are approximately equal. We replace capture rates of Eq. 8.6 by the numerical capture cross sections for the various processes (Fig. 5.17/Fig. 5.18) under these field conditions, summarized in Table 8.1. In Fig 8.1, we plot the grounded, steady

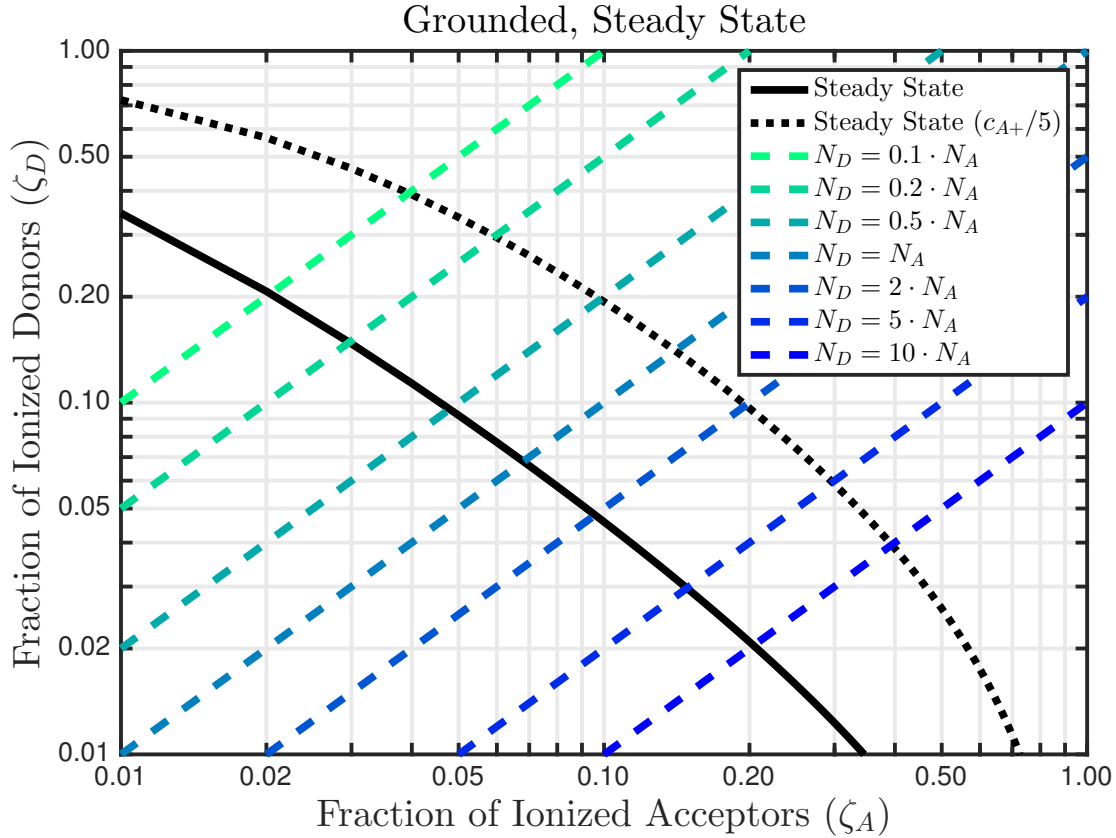


Figure 8.1: Steady state ionized impurity fraction (solid black) versus approximate bulk neutrality conditions for various ratios of donors to acceptors (dashed colored). The intersection between the solid and dashed lines predict the grounded, steady state impurity fractions. The dotted black line represents the steady state ionized impurity fraction by reducing the overcharged acceptor Coulomb capture rate (c_{A+}) by a factor of 5. See Table 8.1 for parameters.

state ionized impurity fractions (Eq. 8.6) in comparison to what is required for bulk neutrality (Eq. 8.7) using these parameters. The intersection between the steady state curve and approximate bulk neutrality lines indicate the expected grounded, steady state ionized impurity fractions.

Focusing on the intersections with the solid black line, we see that Sundqvist's rates predict an n-type detector ($N_D > N_A$) will have $\sim 10\%$ - 20% ionized acceptors and $\sim 2\%$ - 3% ionized donors. The opposite holds true for a p-type detector. A larger ionized fraction of the minority impurity species is needed to cancel the residual overcharged majority impurity states, as would be naively expected. While these percentages are somewhat smaller than what we appear to observe from our inverse trapping length fits from the previous chapter, the results show we should expect to find a substantial number of residual overcharged states in the grounded, steady state. These overcharged states serve as the cascade Coulomb capture centers responsible for the high amount of carrier trapping observed at low external bias. It is somewhat counterintuitive that we can achieve bulk neutrality while still retaining such large populations of overcharged states, however the presence of a neutral capture cross section comparable in size to Coulomb capture makes this unavoidable.

Our best experimental data comes from the n-type detector (G22Q), where we found a $\sim 5\times$ excess of D^- compared to A^+ states from the inverse trapping length fits. As stated earlier, Sundqvist's predicted capture cross sections do not take the electron transport anisotropy into account. If we assume the true Coulomb capture cross section onto A^+ centers is actually $5\times$ lower than the value predicted by Sundqvist, N_{A+} would increase $5\times$ and be in good agreement with N_{D-} — a requirement for approximate bulk neutrality. The grounded, steady state ionized impurity fractions under this hypothesis are shown via the dotted black line in Fig. 8.1. The reduction in A^+ -induced Coulomb capture results in 10% - 50% ionized impurity fractions for both species, which is also in better agreement with the inverse trapping length fits. A more detailed understanding obviously requires better precision in both theory and experiment, however we should assume that $\sim 10\%$ of the total impurities within a so-called “neutralized” CDMS detector are in fact overcharged.

8.1.2 Generation of Space Charge via Capture

When an external bias is applied, the field-dependent capture rates change and the detector is no longer in the steady state. Free carriers produced by events slowly cause the impurity densities to shift from their grounded values to a new distribution, also described by Eq. 8.6. The Coulomb capture rates decrease much more rapidly with electric field than the neutral capture processes. This increases the fraction of ionized impurities in the steady state as it becomes more likely for a free carrier to become captured by neutral impurity, generating overcharged states.

In Fig. 8.2, we plot the steady state ionized impurity fractions using Sundqvist's cross sections at 1 V/cm . A majority of the impurities of both species are ionized. Even if approximate bulk neutrality holds, the huge number of ionized impurities would result in a very high total Coulomb capture rate. This buildup of ionized impurities is responsible for

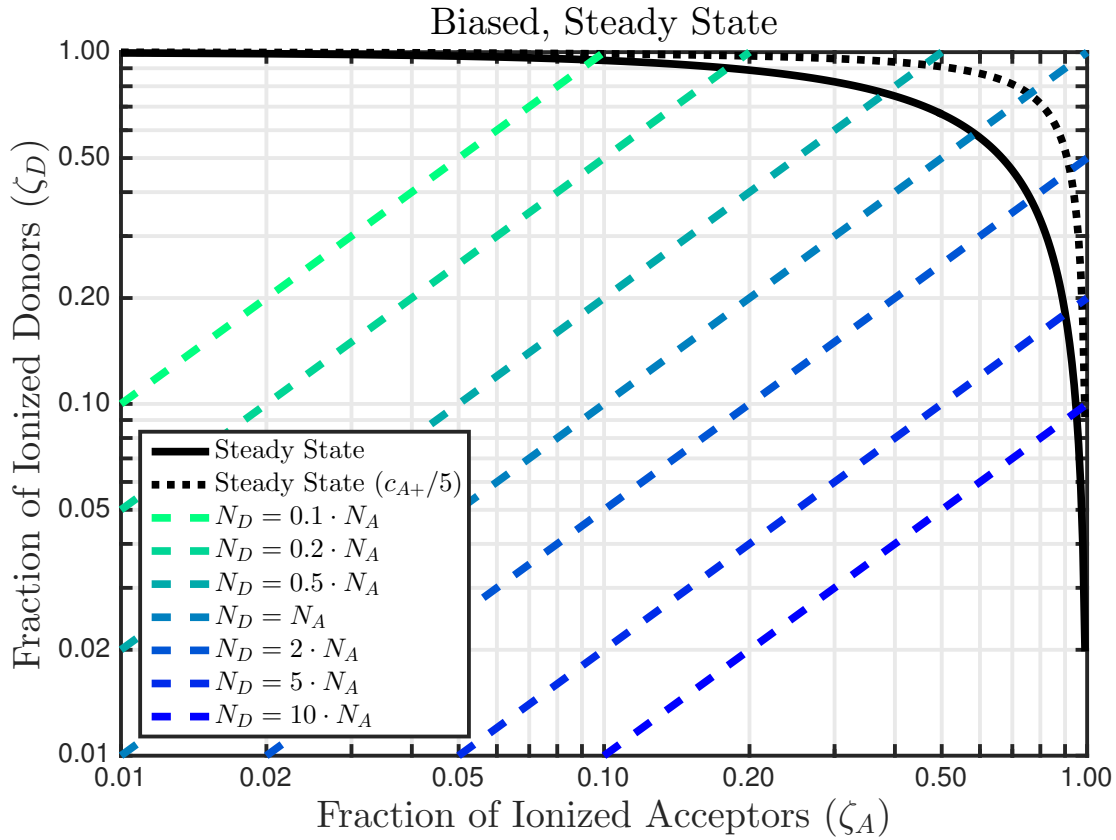


Figure 8.2: Biased (1 V/cm) steady state ionized impurity fractions. Dashed lines represent approximate bulk neutrality.

the degradation of ionization collection efficiency over time while the detectors are biased. Again, a full simulation including Poisson’s equation would be required to determine the specific steady state condition which arises, but these results show we expect large amounts of space charge to accrue within these detectors when biased.

8.1.3 Detector Reset

As the detector approaches the biased, steady state condition the reduction in ionization collection efficiency prevents it from effectively searching for WIMPs. The procedure for restoring the ionization collection efficiency is obvious — simply remove the external bias (ground the detector electrodes) and generate a large number of free carriers to bring the crystal back into the grounded, steady state. These periodic detector resets can be completed in much less time than the initial neutralization process after cool down since we only need to neutralize the excess overcharged states as opposed to a large population of “standard”

D^+/A^- states. In particular, if the reset is performed immediately upon the onset of a change in ionization collection efficiency (before the complete biased, steady state has been reached), good collection can be restored with LED flashes of only a few seconds rather than several hours.

We now have a good understanding of the evolution of space charge within CDMS detectors. Immediately after reaching base temperature, a lengthy neutralization procedure is required to place the detector in its grounded, steady state condition in which roughly 10% of the total impurities are in the form of overcharged states. The correlated formation of these overcharged states[178] results in the approximate bulk neutrality of the crystal. Upon biasing for WIMP search operation, the detector is immediately out of equilibrium from the biased, steady state. Over time, the free carriers created by events are preferentially captured by neutral impurities, generating the large amounts of space charge predicted under the biased, steady state and compromising the ionization collection efficiency. The external bias can then be removed, and a large number of free carriers generated to return to the grounded, steady state. The detector may once again be biased with good ionization collection efficiency temporarily restored.

8.2 Ionization Collection Efficiency and the Total Phonon Energy

Since both the recombination and Luke-Neganov phonons depend on the number of charge carriers, trapping and impact ionization influence the total phonon energy produced by a recoil event. Recall we have

$$E_{phonon} = E_{recoil} - n_{eh}E_{gap} \quad (8.8)$$

and

$$p_t = E_{phonon} + E_{recomb} + E_{luke}, \quad (8.9)$$

where p_t is the total measured phonon energy.

In the absence of trapping or impact ionization processes, all the initial recoil energy lost to the creation of the charge carriers is regained when the charge carriers relax to the Fermi level at the contacts. The total phonon energy is then simply the recoil energy plus the Luke phonon contribution. Since we believe carriers are trapped to shallow states, with binding energies far smaller than the bandgap, a fraction of the recombination phonons will not be released. In the case of impact ionization, the energy to excite the extra carriers to the conduction or valence bands comes from other charge carriers and not the initial recoil deposition, hence an excess of recombination phonons will be produced. Hence we have

$$E_{recomb} = n_{eh,f}E_{gap}, \quad (8.10)$$

where $n_{eh,f}$ is the number of electron-hole pairs which actually reach the contacts.

Eq. 8.10 is a simplification of the true situation. The Fermi level at the contact likely not equidistant in energy from the conduction and valence bands, and need not even be the same

at each contact. In addition, the differences in electron and hole trapping/impact ionization mean that we should not be counting by electron-hole pairs, as there will be an excess of one carrier over the other, but by individual carriers. The number of each carrier will also depend on the position of the event within the detector.

We simplify the situation by assuming the Fermi level lies in the middle of the bands at both electrodes. With a planar geometry, we therefore have

$$\begin{aligned} E_{recomb} &= \frac{1}{2} \left(\exp \left(-\frac{z}{\lambda_h} \right) + \exp \left(-\frac{L-z}{\lambda_e} \right) \right) n_{eh} E_{gap} \\ &= (1 - \beta) n_{eh} E_{gap}, \end{aligned} \quad (8.11)$$

which defines β .

Similarly, the assumed number of Luke phonons produced is given by

$$E_{luke} = n_{eh} eV. \quad (8.12)$$

The Luke phonon contribution from carriers traveling a distance $\vec{dz}$ is given by

$$dE_{luke} = -n_{eh} e \vec{E} \cdot \vec{dz} = -q_0 \exp \left(-\frac{z}{\lambda} \right) \vec{E} \cdot \vec{dz}. \quad (8.13)$$

In the case of the planar geometry, this becomes

$$dE_{luke} = q_0 \exp \left(-\frac{z}{\lambda} \right) \left(\frac{V}{L} \right) dz, \quad (8.14)$$

leading to

$$E_{luke} = \eta n_{eh} eV, \quad (8.15)$$

where $\eta = Q/Q_0$ is the charge collection efficiency. Note that this holds for either trapping or impact ionization, and that position dependence can be included by using Eq. 7.14 for η .

Substituting Eq. 8.11 and Eq. 8.15 into Eq. 8.9, we obtain

$$p_t = E_{recoil} - \beta n_{eh} E_{gap} + \eta n_{eh} eV. \quad (8.16)$$

The initial number of electron-hole pairs created is related to the recoil energy through the parameter ϵ via $n_{eh} = E_{recoil}/\epsilon$. Finally, we find

$$p_t = E_{recoil} \left(1 - \beta \frac{E_{gap}}{\epsilon} + \eta \frac{eV}{\epsilon} \right). \quad (8.17)$$

We recognize this as a modified form of the standard Luke gain equation, which is recovered by setting $\beta = 0$ and $\eta = 1$. We plot the modified Luke gain (p_t/E_{recoil}) as a function of the bias voltage using the integrated charge collection efficiency and derived trapping lengths for events occurring in the middle of the detector ($z = L/2$) in Fig. 8.3.

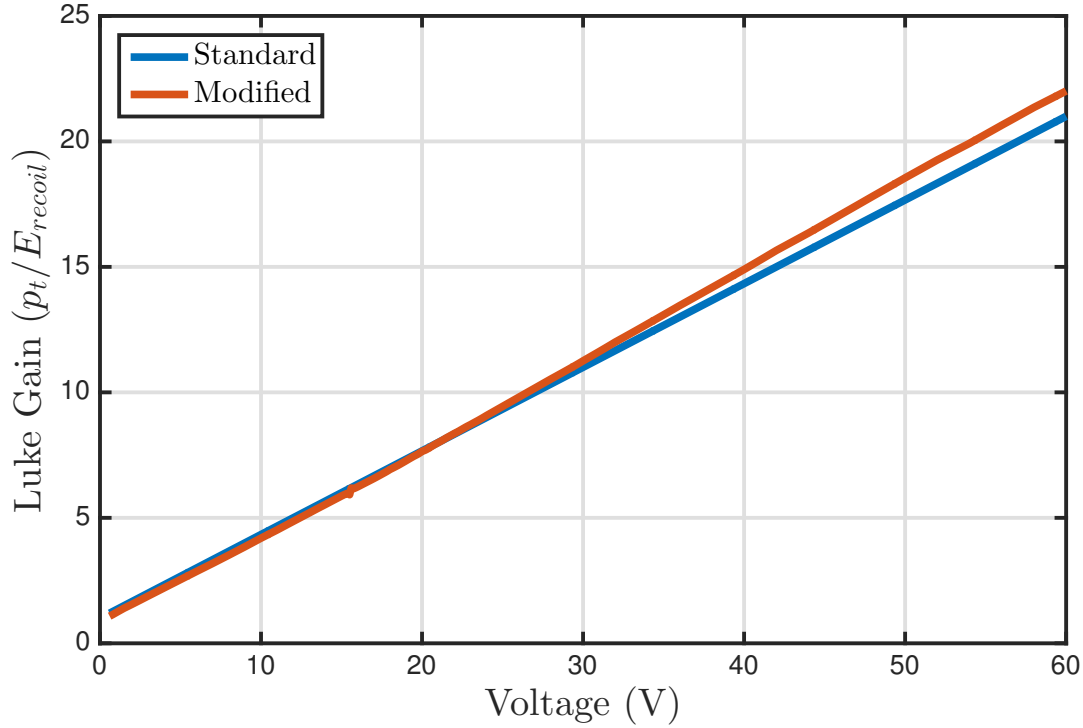


Figure 8.3: Modified Luke gain (Eq. 8.17) in comparison to the standard Luke gain for G22Q. We take $\epsilon = 3$ eV and $E_{gap} = 0.785$ eV.

8.3 Nuclear Recoil Energy Scale

The reduced ionization yield of nuclear recoils causes less Luke phonons to be produced than an equivalent-energy electron recoil event. Thus the conversion between measured total phonon energy and recoil energy differs between the two types of events. For CDMS and EDELWEISS detectors, it is the electron recoil scale which is determined through the calibration procedure as monoenergetic photons are easily obtainable through the use of radioactive sources. The output voltage of the corresponding photopeak represents a known electron recoil energy as the photons are fully absorbed by the crystal. WIMPs and neutrons will typically not deposit their full energy, as they are scattered and not absorbed. Thus even with a monoenergetic neutron source, there is no photopeak by which to calibrate to. Depending on the mode in which the detector is run, a conversion between measured total phonon energy and true recoil energy must be performed. We now determine how much error the modified Luke gain of Eq. 8.17 introduces to the CDMS nuclear recoil energy scale, which assumes perfect charge collection efficiency.

8.3.1 CDMSlite Mode

The high-field CDMSlite experiment, operating at a field of ~ 27 V/cm, is clearly in a regime where we would expect to see impact ionization. Unlike the standard iZIP mode, CDMSlite only measures the total phonons produced and takes advantage of Luke amplification to significantly reduce the ionization energy threshold. Since no event-by-event discrimination between electron and nuclear recoils is possible, the entire measured spectrum is interpreted as nuclear recoils.

Calibration of the electron recoil scale is performed by exposing the CDMSlite detector to a ^{252}Cf source. ^{71}Ge is created via neutron capture onto ^{70}Ge . The ^{71}Ge then decays via electron capture with a half-life of 11.43 days.[197] Depending on whether the electron is captured from the K-, L-, or M-shell, x-rays and auger electrons are emitted with a total energy of 10.37, 1.30, and 0.16 keV, respectively.[198] For the CDMSlite Run 1 and 2 results, the K-shell line was used.[133, 134]

We define the *measured* photopeak location as $p_{t,C}$. Since we know with certainty that this photopeak represents an electron recoil of 10.37 keV, we can convert between an arbitrary *measured* total phonon energy $p_{t,M}$ and the electron recoil energy scale using

$$E_{r,ee} = \frac{10.37 \text{ keV}}{p_{t,C}} \cdot p_{t,M}. \quad (8.18)$$

To convert the measured spectrum to the nuclear recoil energy scale, the total phonon energy is assumed to take the simplified form

$$E_t = E_r \left(1 + Y(E_r) \frac{eV}{\epsilon} \right). \quad (8.19)$$

For electron recoils, we have $E_r = E_{r,ee}$ and $Y(E_r) = 1$. For nuclear recoils, $E_r = E_{r,nr}$ and $Y(E_r)$ is assumed to follow the Lindhard yield.[120] Equating these two cases and rearranging, we have

$$E_{r,nr} = E_{r,ee} \left(\frac{1 + eV/\epsilon}{1 + Y(E_{r,nr})eV/\epsilon} \right). \quad (8.20)$$

Eq. 8.18 is substituted for $E_{r,ee}$, completing the conversion from measured total phonon energy to the nuclear recoil energy scale.

The problem arises in the use of Eq. 8.19. As we have seen in the previous section, including the effects of charge carrier trapping and impact ionization modify this equation, which takes the true form

$$E_t = E_r \left[1 + Y(E_r) \left(\eta \frac{eV}{\epsilon} - \beta \frac{E_{gap}}{\epsilon} \right) \right] \quad (8.21)$$

as in Eq. 8.17. Propagating this change through the derivation of the nuclear recoil energy scale, we arrive at

$$E_{r,nr} = E_{r,ee} \left[\frac{1 + \left(\eta \frac{eV}{\epsilon} - \beta \frac{E_{gap}}{\epsilon} \right)}{1 + Y(E_{r,nr}) \left(\eta \frac{eV}{\epsilon} - \beta \frac{E_{gap}}{\epsilon} \right)} \right]. \quad (8.22)$$

Note that Eq. 8.18 can still be used to convert from $E_{r,ee}$ to $p_{t,M}$ since variations in charge collection efficiency do not change the fact that $p_{t,C}$ still represents a known electron recoil energy, regardless of the total phonons collected.

The error introduced by taking the ratio of Eq. 8.22 to Eq. 8.20, which is the factor by which a nuclear recoil energy arrived at by the standard calibration would need to be multiplied by to recover the true recoil energy. We have

$$\delta E_{r,nr} = \left[\frac{1 + \left(\eta \frac{eV}{\epsilon} - \beta \frac{E_{gap}}{\epsilon} \right)}{1 + Y(E_{r,nr}) \left(\eta \frac{eV}{\epsilon} - \beta \frac{E_{gap}}{\epsilon} \right)} \right] \cdot \left(\frac{1 + Y(E_{r,nr}) eV/\epsilon}{1 + eV/\epsilon} \right), \quad (8.23)$$

which we plot as a function of bias voltage using our field-dependent values for η and β for G22Q from the previous section in Fig. 8.4, assuming a nuclear recoil yield of 0.2.

We see under the high field conditions of CDMSlite, impact ionization on the order of 10% will only result in a nuclear recoil energy scale error of $\sim 1\%$ or less. In fact, we see in the limit of very large bias voltage Eq. 8.23 becomes

$$\lim_{V \rightarrow \infty} \delta E_{r,nr} = \left(\frac{\eta \frac{eV}{\epsilon}}{Y(E_{r,nr}) \eta \frac{eV}{\epsilon}} \right) \cdot \left(\frac{Y(E_{r,nr}) \frac{eV}{\epsilon}}{\frac{eV}{\epsilon}} \right) = 1. \quad (8.24)$$

Impact ionization has no effect on the CDMSlite nuclear recoil energy scale at very high field. The error is completely determined by systematics on the ionization yield. Intuitively, this can be understood by noting that at very high field the entire signal is essentially all Luke phonons — both the recoil and recombination phonons make negligible contributions to the total number of measured phonons. As impact ionization does not depend on whether the initial energy deposition was an electron or nuclear recoil, the only difference between them is through the number of electron-hole pairs initially created. This is fully determined by the ionization yield, hence it is the only source of error.

8.3.2 iZIP Mode

When operating in iZIP mode, both the ionization and phonons produced by an event are measured. A specific model for the ionization yield is not needed due to the simultaneous measurement of both quantities. While this allows for the discrimination between electron and nuclear recoils, the ionization, p_t and nuclear recoil energy scales must still be determined.

The iZIP calibration is performed by exposure to a radioactive ^{133}Ba source, whose decay results in the production of a large number of 356 keV photons. We can define the ionization energy scale by fitting to this peak. As before, with q_M being an arbitrary measured charge signal and q_C being the location of the photopeak, we have

$$q = \frac{356 \text{ keV}}{q_C} \cdot q_M = E_{r,ee}. \quad (8.25)$$

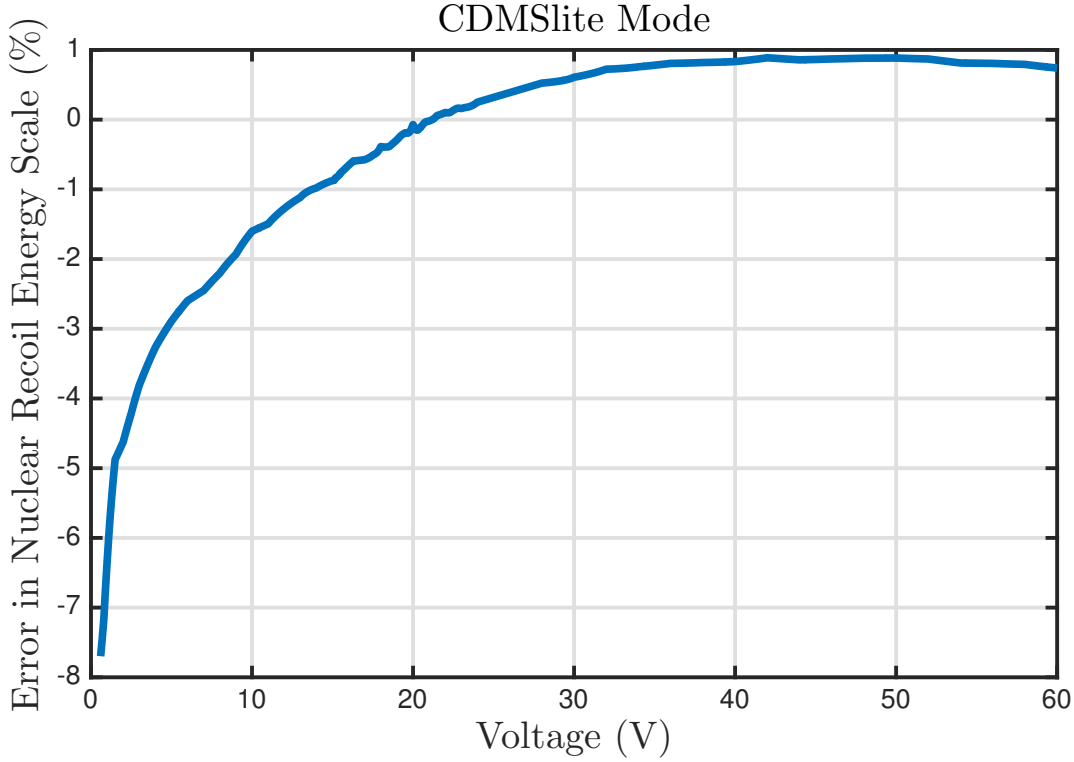


Figure 8.4: Error in the CDMSlite nuclear recoil energy scale using the ionization collection efficiency parameters for G22Q.

Again, since we know the true electron recoil energy, this calibration is unaffected by variations in ionization collection efficiency.

To calibrate p_t , the standard Luke gain formula is assumed with a yield of 1 to generate a predicted p_t value for each 356 keV electron recoil event based on the Eq. 8.25. The predicted p_t values are then used to define a conversion between the raw total phonon amplitude (in Volts) and the p_t energy scale, which we define as $p_{t,C}$, such that p_t follows the standard Luke gain. We have

$$p_t = q \left(1 + \frac{eV}{\epsilon} \right) \cdot \frac{p_{t,M}}{p_{t,C}} = q \left(1 + \frac{eV}{\epsilon} \right) \quad (8.26)$$

as $\frac{p_{t,M}}{p_{t,C}} \equiv 1$.

We are now faced with two options when trying to determine $E_{r,nr}$ for nuclear recoils. The first is to each event's own calibrated ionization signal. The calibration from Eq. 8.25 can be used to estimate the Luke phonon contribution — *even in the case of nuclear recoils*. The output voltage of the charge amplifier only depends on the number of electron-hole pairs

and does not care through which process they have been generated. We therefore have

$$E_{luke} = n_{eh}eV = \frac{q}{\epsilon}eV, \quad (8.27)$$

where n_{eh} is the number of electron-hole pairs created and we have assumed it takes $\epsilon = 3$ eV of recoil energy to produce an electron hole pair. In this equation, the use of 3 eV to convert from q to n_{eh} is correct even for nuclear recoils, as our energy calibration has been performed using electron recoils.

Under the assumption of complete charge collection, the nuclear recoil energy is simply

$$E_{r,nr} = p_t - E_{luke}. \quad (8.28)$$

The downside to this method is that the ionization measurement is used to directly estimate the Luke phonons. At low energies, the charge signals are much noisier than the phonon signals, degrading the resolution. To avoid this, we instead define a function $q_p = f(p_t)$ based on ^{252}Cf neutron calibration to replace q when calculating the Luke phonon contribution. [73]

Regardless of the method used, the departure from complete ionization collection due to trapping and impact ionizations alters the forms of Eq. 8.26 and Eq. 8.27. The true total phonon energy is

$$p_{t,true} = q \left(1 - \beta \frac{E_{gap}}{\epsilon} + \eta \frac{eV}{\epsilon} \right), \quad (8.29)$$

while the true Luke phonon energy is

$$E_{luke,true} = \eta \frac{q}{\epsilon} eV. \quad (8.30)$$

The ratio between the true recoil energy and the measured energy is

$$\frac{E_{r,nr,true}}{E_{r,nr}} = 1 - \beta \frac{E_{gap}}{\epsilon}. \quad (8.31)$$

We see that the errors in the Luke phonon predictions due to variations in charge collection are calibrated out, however the error in the number of predicted recoil phonons remains. We plot the error between the true energy and calibrated energy using our derived parameters for G22Q in Fig. 8.5.

8.4 Silicon Detectors

Charge collection in silicon CDMS detectors has not been studied in detail. The boules used to produce detectors typically have net shallow impurity concentrations 10x-100x greater than the germanium boules. As noted in Abakumov[178], similar impurity capture processes to those we have been discussing are observed in silicon, including neutral capture producing overcharged states.[199] Silicon is also frozen out at sub-Kelvin temperatures, so we expect

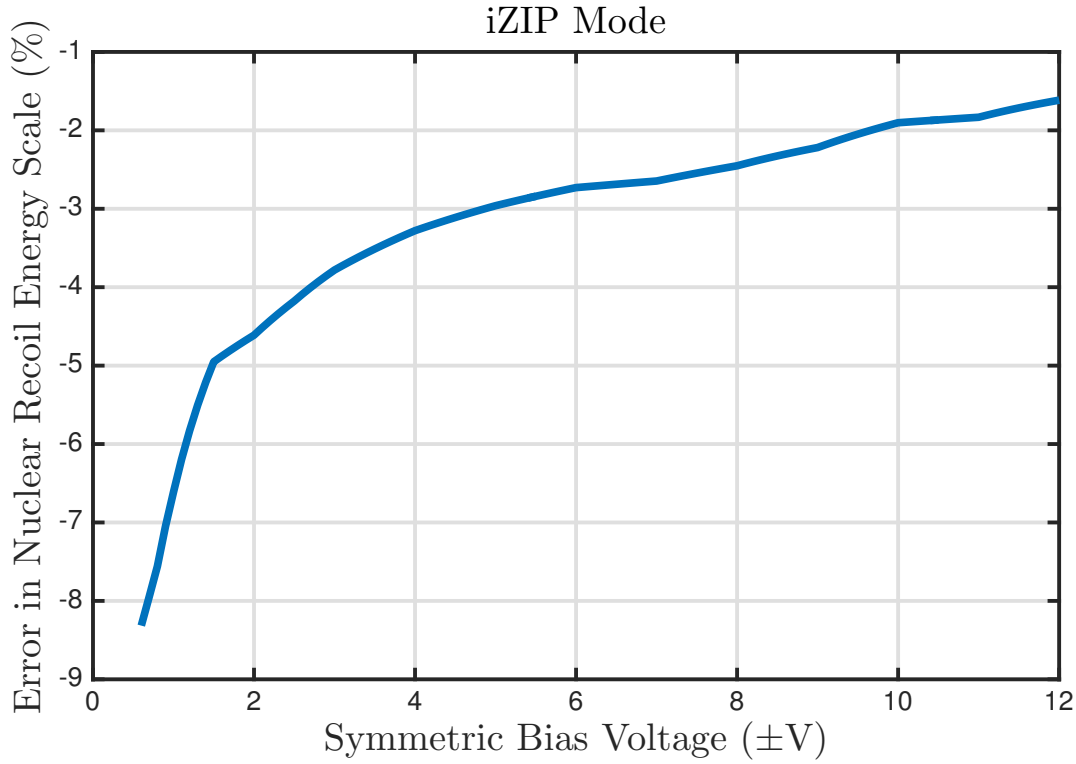


Figure 8.5: Error in the iZIP nuclear recoil energy scale using the ionization collection efficiency parameters for G22Q.

transport to be determined by spontaneous phonon emission and a two species, two state model to dictate carrier capture. Ionized impurity scattering may be more important at lower fields. As the total capture rate scales with the total impurity concentration, we expect trapping to be worse in silicon detectors. Similarly, impact ionization may be enhanced at high field.

A deeper understanding of ionization collection in sub-Kelvin silicon would require a dedicated investment of manpower and resources. A detailed Monte Carlo simulation such as Sundqvist's[146] would need to be written to obtain the fundamental carrier distributions, however the effects of anisotropy would be reduced. Cross sections for cascade Coulomb capture and neutral capture forming overcharged states would then need to be calculated. Finally, a few mZIP-style silicon detectors would need to be studied using the experimental apparatus described in this dissertation. As the collaboration intends to deploy silicon detectors for SuperCDMS SNOLAB, this work deserves some consideration.

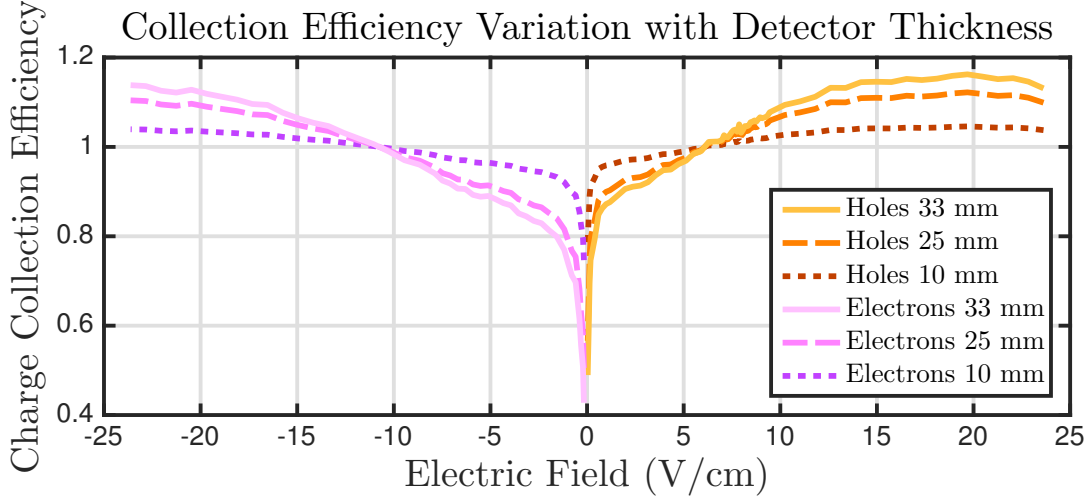


Figure 8.6: Comparison of the ionization collection efficiencies for a 33 mm, 1 inch, and 1 cm detector. Efficiencies are based on the integrated efficiencies of G22Q.

8.5 Moving to Larger Crystals

To maximize the exposure to possible WIMP interactions, it is desirable for the detectors to be made as large as possible. In moving from CDMS II to SuperCDMS Soudan, the detector thickness was increased from 1 cm to 1 inch. For SNOLAB, this thickness will be increased by an additional 30% to 33 mm. The increase in drift length due to the larger crystal dimensions means that the effects of charge trapping and impact ionization become enhanced because of the exponential dependence on the drift length. In iZIP mode, an additional complication arises because the number of Luke-Neganov phonons produced depends on the total *voltage* and not the electric field. For a fixed Luke gain, a thicker crystal will have a lower bulk electric field than a thinner detector. We have seen that the charge trapping greatly increases as the field is reduced. Increasing the bias voltage to reduce the trapping may cause too many Luke phonons to be produced, masking the primary recoil phonon signal.

Using our various measures of the trapping and generation lengths for the n-type G22Q detector, we now calculate the influence of charge transport effects for the SNOLAB detector geometry, using a detector thickness of $L = 33$ mm. In Fig. 8.6, we show the charge collection efficiency as a function of electric field, as derived from the integrated efficiencies. When trapping dominates, the extra detector thickness causes an additional reduction in the collection efficiency of $\sim 5\%$. When impact ionization dominates, the increased drift length causes a larger excess in charge carriers, as expected. This highlights the increasing dependence on charge transport properties as the thickness is increased.

Next, in Fig. 8.7, we plot the charge collection efficiency as a function of z-position

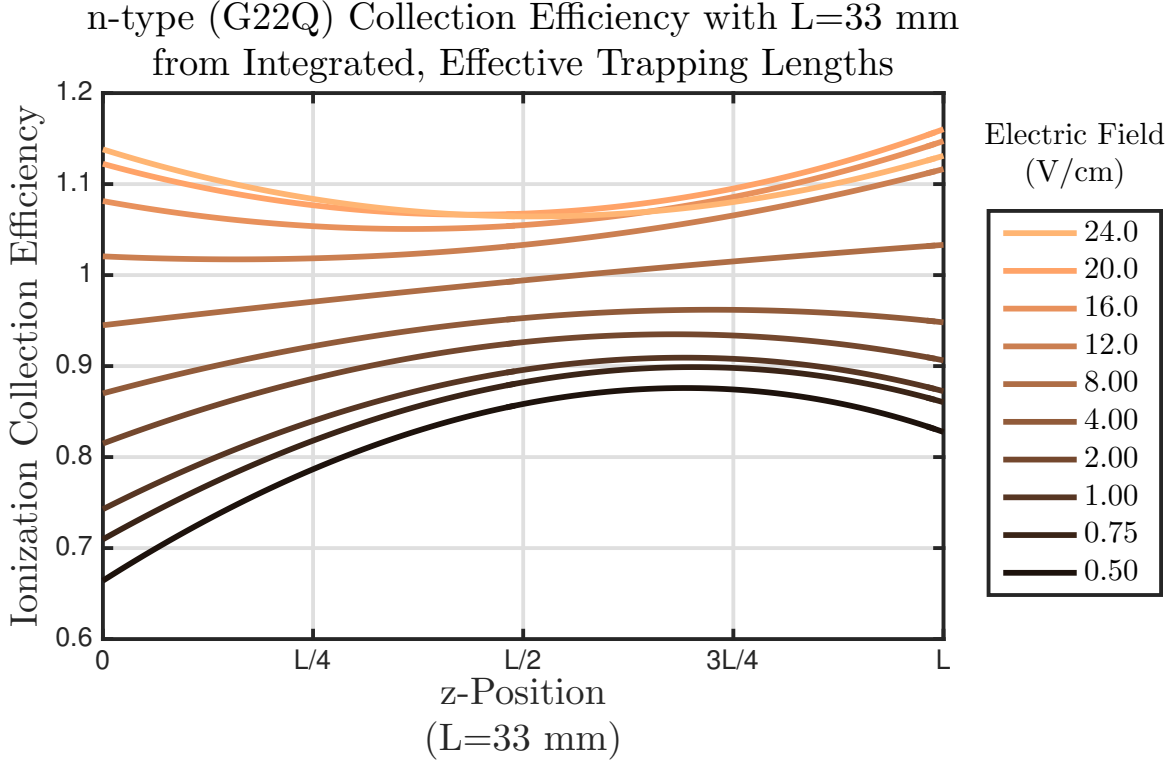


Figure 8.7: 33 mm charge collection efficiency as a function of z-position and electric field, using the trapping lengths derived from the integrated efficiency of G22Q.

for various electric field strengths. To approximate the effects of opposite carrier impact ionization, we use the trapping lengths derived from the integrated charge collection efficiency of G22Q. We see that the shape is qualitatively similar to the 1 inch detector (see Fig. 7.16), however the overall collection is decreased by ~ 5 -10% and the difference between minimum and maximum collection for a given field is increased. We should therefore expect some additional broadening of photopeaks due to this increased position dependence, although this will not influence the baseline energy resolution.¹ We plot the collection efficiency in iZIP mode in Fig. 8.8.

We have shown that the error in the CDMSlite energy scale does not depend on the charge collection efficiency as the voltage is increased. The SNOLAB high voltage detectors are expected to operate at fields above the range covered by our measurements. Increasing the detector thickness will not degrade the CDMSlite energy resolution as long as high bias voltage can be maintained and error in the nuclear recoil yield model kept low.

When operating in iZIP mode, the field must be kept relatively low in order to limit Luke phonon production. In Fig. 8.9, we plot the error in the iZIP nuclear recoil energy scale for

¹The baseline energy resolution is determined only by the electrical amplifier noise.

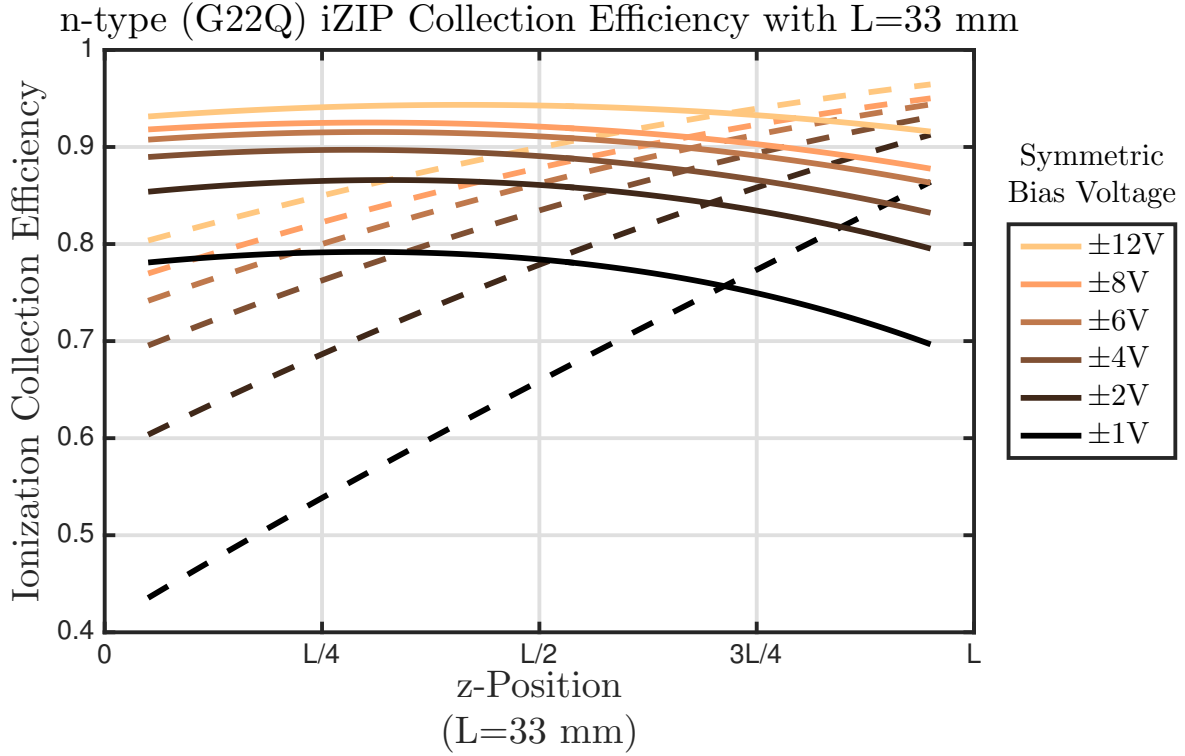


Figure 8.8: 33 mm iZIP charge collection efficiency, using the trapping lengths derived from the integrated efficiency of G22Q. Side 1 ($z=0$, hole) collection is shown with solid lines, while side 2 ($z=L$, electron) collection is shown with dashed lines.

various symmetric biases voltages and detector thicknesses. Clearly, the error increases with the detector thickness for fixed bias. SNOLAB iZIP detectors should be run with biases of $\sim \pm 4V$ in order to limit the error to 5% or less, although this will vary from detector to detector. If there is still a desire to continue with even thicker detectors, CDMS will need to increase the standard iZIP operating voltage further. Care must be taken to ensure good discrimination and event reconstruction are possible with the increased Luke phonon emission.

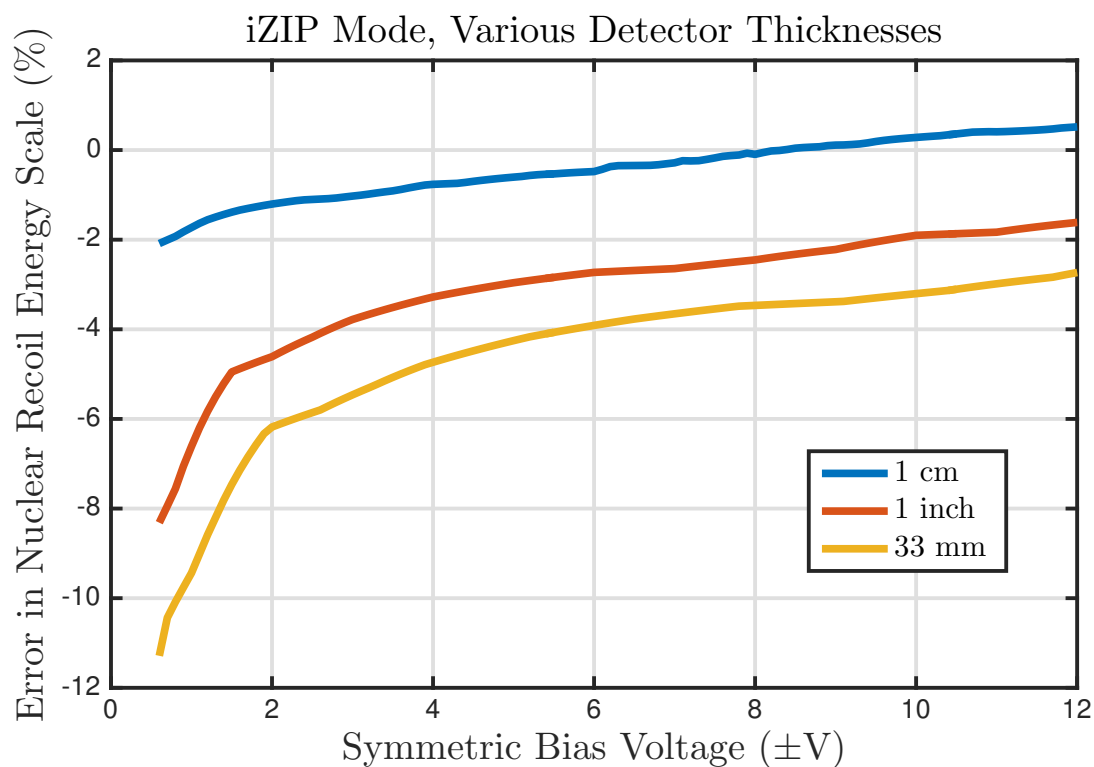


Figure 8.9: Error in iZIP nuclear recoil scale as a function of the symmetric bias voltage, for varying detector thicknesses. The results are derived from the integrated charge collection efficiency of G22Q.

Chapter 9

Conclusions



Figure 9.1: The author, Arran Phipps, and his children Emily, Margaret, and Finnegan.

The future of the Cryogenic Dark Matter Search lies within the regime of low mass WIMPs. Due to the sub-keV recoil energies deposited by low mass WIMPs, a deep understanding of the detector ionization collection physics and the development of extremely low-noise instrumentation for signal readout are a necessity for the success of the SNOLAB project. It is my hope that the work described in this dissertation will help the collaboration achieve its goals.

Reaching sub-100 eV ionization resolution clearly requires the use of CNRS/LPN HEMTs. We have shown that our fully cryogenic HEMT amplifier is fundamentally capable of reaching such low noise levels, and have achieved (to our knowledge) the best ever ionization resolution in a CDMS detector of 150 eVee. The fact that this resolution is strongly limited

by detector bias resistor Johnson noise shows the importance of careful component selection and consideration of more exotic amplifier topologies for the SNOLAB project. It may be in the project's best interest to remove all resistors entirely and perform DC biasing through the use of HEMT switches.

The ionization trace length will need to be substantially longer than has been typically used within the collaboration in order to reach the goal of 100 eV. In addition to creating problems with data acquisition, testing of detectors at the surface becomes increasingly difficult due to the large background event rate. This leads to a number of pileup events, complicating the assessment of both amplifier and detector performance. Radioactive source calibration of the detectors also causes concern as the event rate would need to be high enough to obtain adequate statistics, but low enough as not to induce too much dead time in the HEMT amplifier. The success of the prototype amplifier, however, would seem to indicate that these concerns can be addressed. The fully cryogenic HEMT amplifier is a project which the collaboration should continue to pursue, and the use of HEMTs for other specialized measurements (such as leakage currents) should be seriously considered.

Over the past several years, we have made substantial progress in understanding the physics of ionization collection in CDMS detectors. The predictions of Sundqvist appear to be in good agreement with our measured data. Transport properties are determined by the spontaneous emission of phonons due to free charge carrier interactions with the zero-point motion of the lattice ions. Carriers are trapped via Coulomb capture onto overcharged D^-/A^+ states, or via neutral capture producing these overcharged states. The thermal stability of overcharged states at sub-Kelvin temperature causes them to dictate the overall charge state of the detector. We have developed a two species, two state model which explains why good ionization collection can be recovered by placing detectors in their grounded, steady state condition, while also giving reason for the observed accrual of space charge when the detectors are biased. Bulk neutrality is achieved through the balance of positive and negative overcharged states, rather than a complete neutralization of all impurities. We have also shown direct evidence of high field impact ionization processes at levels consistent with the predictions of Sundqvist.

The effects of trapping and impact ionization will become more pronounced as the collaboration moves to larger detectors with lower energy thresholds. Impact ionization will be particularly important for future high voltage detectors which aim to achieve single electron-hole pair sensitivity. Bias voltages in iZIP mode will need to be kept high in order to limit error in the nuclear recoil energy scale. Variations in ionization collection efficiency will make it difficult to have much better than 5% accuracy in the nuclear recoil energy scale, however this does not strongly influence the final WIMP search results.

While there is still work which can be done to improve our understanding of the specific impurity species and associated cross sections present in these detectors, this lies closer to the discipline of condensed matter physics rather than particle astrophysics. Broadly speaking, we now have a solid foundation for understanding ionization collection in detectors of the Cryogenic Dark Matter Search.

Bibliography

1. Zwicky, F. Die rotverschiebung von extragalaktischen nebeln. German. *Helvetica Physica Acta* **6**, 110–127 (1933).
2. Zwicky, F. On the Masses of Nebulae and of Clusters of Nebulae. *Astrophys. J.* **86**, 217–246 (1937).
3. Rubin, V. C. & Ford Jr, W. K. Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. *Astrophys. J.* **159**, 379–403 (1970).
4. Rubin, V. C., Thonnard, N. & Ford Jr., W. K. Extended rotation curves of high-luminosity spiral galaxies. IV - Systematic dynamical properties, SA through SC. *Astrophys. J.* **225**, L107–L111 (1978).
5. Faber, S. M. & Gallagher, J. S. Masses and mass-to-light ratios of galaxies. *Ann. Rev. Astron. Astrophys.* **17**, 135–183 (1979).
6. Peebles, P. J. E. Large scale background temperature and mass fluctuations due to scale invariant primeval perturbations. *Astrophys. J.* **263**, L1–L5 (1982).
7. Blumenthal, G. R., Pagels, H. & Primack, J. R. Galaxy Formation by Dissipationless Particles Heavier than Neutrinos. *Nature* **299**, 37–38 (1982).
8. Bond, J. R., Szalay, A. S. & Turner, M. S. Formation of Galaxies in a Gravitino Dominated Universe. *Phys. Rev. Lett.* **48**, 1636 (1982).
9. Ellis, J. R., Hagelin, J. S., Nanopoulos, D. V., Olive, K. A. & Srednicki, M. Supersymmetric Relics from the Big Bang. *Nucl. Phys.* **B238**, 453–476 (1984).
10. Blumenthal, G. R., Primack, J. R., Rees, M. J. & Faber, S. M. Formation of galaxies and large scale structure with cold dark matter. *Nature* **311**, 517–525 (1984).
11. Wright, E. L. *et al.* Interpretation of the Cosmic Microwave Background radiation anisotropy detected by the COBE differential microwave radiometer. *Astrophys. J.* **396**, L13–L18 (1992).
12. Riess, A. G. *et al.* Observational evidence from supernovae for an accelerating universe and a cosmological constant. *Astron. J.* **116**, 1009 (1998).
13. Perlmutter, S. *et al.* Measurements of Ω and Λ from 42 High-Redshift Supernovae. *Astrophys. J.* **517**, 565 (1999).

14. Liddle, A. *An introduction to modern cosmology* (John Wiley & Sons, 2015).
15. Hinshaw, G. *et al.* Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results. *Astrophys. J. S.* **208**, 19 (2013).
16. Sofue, Y. & Rubin, V. Rotation Curves of Spiral Galaxies. *Annu. Rev. Astron. Astrophys* **39**, 137–74 (2001).
17. Klypin, A., Zhao, H. S. & Somerville, R. S. Λ CDM-based models for the Milky Way and M31. I. Dynamical models. *Astrophys. J.* **573**, 597 (2002).
18. Tyson, J. A., Kochanski, G. P. & Dell’Antonio, I. P. Detailed Mass Map of CL 0024+1654 from Strong Lensing. *Astrophys. J. Lett.* **498**, L107 (1998).
19. Clowe, D. *et al.* A Direct Empirical Proof of the Existence of Dark Matter. *Astrophys. J. Lett.* **648**, L109 (2006).
20. Springel, V., Frenk, C. S. & White, S. D. M. The large-scale structure of the Universe. *Nature* **440**, 1137–1144 (2006).
21. Tegmark, M. *et al.* The Three-Dimensional Power Spectrum of Galaxies from the Sloan Digital Sky Survey. *Astrophys. J.* **606**, 702 (2004).
22. Klypin, A. A., Trujillo-Gomez, S. & Primack, J. Dark matter halos in the standard cosmological model: Results from the bolshoi simulation. *Astrophys. J.* **740**, 102 (2011).
23. Fields, B. D., Molaro, P. & Sarkar, S. Big-Bang Nucleosynthesis. *Chin. Phys.* **C38**. arXiv: 1412.1408 [astro-ph.CO] (2014).
24. Smoot, G. F. *et al.* Structure in the COBE differential microwave radiometer first year maps. *Astrophys. J.* **396**, L1–L5 (1992).
25. Bennett, C. *et al.* Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: final maps and results. *Astrophys. J. S.* **208**, 20 (2013).
26. Planck Collaboration *et al.* Planck 2013 results. XXIII. Isotropy and statistics of the CMB. *A&A* **571**, A23 (2014).
27. Milgrom, M. A modification of the Newtonian dynamics as a possible alternative to the hidden mass hypothesis. *Astrophys. J.* **270**, 365–370 (1983).
28. McGaugh, S. S. A tale of two paradigms: the mutual incommensurability of Λ CDM and MOND. *Canadian Journal of Physics* **93**, 250–259 (2015).
29. Paczynski, B. Gravitational microlensing by the galactic halo. *Astrophys. J.* **304**, 1–5 (1986).
30. Renault, C. *et al.* Observational limits on MACHOs in the galactic halo. arXiv: astro-ph/9612102 [astro-ph] (1996).
31. Alcock, C. *et al.* The MACHO project: limits on planetary mass dark matter in the galactic halo from gravitational microlensing. *Astrophys. J.* **471**, 774 (1996).

- 32. Alcock, C. *et al.* EROS and MACHO Combined Limits on Planetary-Mass Dark Matter in the Galactic Halo. *Astrophys. J. Lett.* **499**, L9 (1998).
- 33. Doroshkevich, A., Khlopov, M. Y., Sunyaev, R., Szalay, A. & Zeldovich, Y. B. Cosmological impact of the neutrino rest mass. *Annals of the New York Academy of Sciences* **375**, 32–42 (1981).
- 34. Zeldovich, I. B., Einasto, J. & Shandarin, S. Giant voids in the universe. *Nature* **300**, 407–413 (1982).
- 35. Bond, J. R. & Szalay, A. S. The Collisionless Damping of Density Fluctuations in an Expanding Universe. *Astrophys. J.* **274**, 443–468 (1983).
- 36. White, S. D., Frenk, C. & Davis, M. Clustering in a neutrino-dominated universe. *Astrophys. J.* **274**, L1–L5 (1983).
- 37. Dodelson, S. & Widrow, L. M. Sterile neutrinos as dark matter. *Phys. Rev. Lett.* **72**, 17 (1994).
- 38. Horiuchi, S. *et al.* Sterile neutrino dark matter bounds from galaxies of the Local Group. *Phys. Rev. D* **89**, 025017 (2014).
- 39. Ng, K. C. Y., Horiuchi, S., Gaskins, J. M., Smith, M. & Preece, R. Improved limits on sterile neutrino dark matter using full-sky Fermi Gamma-ray Burst Monitor data. *Phys. Rev. D* **92**, 043503 (2015).
- 40. Abazajian, K. N. Resonantly Produced 7 keV Sterile Neutrino Dark Matter Models and the Properties of Milky Way Satellites. *Phys. Rev. Lett.* **112**, 161303 (2014).
- 41. Peccei, R. D. & Quinn, H. R. Constraints imposed by CP conservation in the presence of pseudoparticles. *Phys. Rev. D* **16**, 1791–1797 (1977).
- 42. Preskill, J., Wise, M. B. & Wilczek, F. Cosmology of the invisible axion. *Phys. Lett. B* **120**, 127–132 (1983).
- 43. Abbott, L. F. & Sikivie, P. A cosmological bound on the invisible axion. *Phys. Lett. B* **120**, 133–136 (1983).
- 44. Dine, M. & Fischler, W. The not-so-harmless axion. *Phys. Lett. B* **120**, 137–141 (1983).
- 45. Ipser, J. & Sikivie, P. Can galactic halos be made of axions? *Phys. Rev. Lett.* **50**, 925 (1983).
- 46. Carosi, G. *et al.* Probing the axion-photon coupling: phenomenological and experimental perspectives. A snowmass white paper. arXiv: 1309.7035 [hep-ph] (2013).
- 47. Sikivie, P. Axions, Domain Walls, and the Early Universe. *Phys. Rev. Lett.* **48**, 1156–1159 (1982).
- 48. Asztalos, S. J. *et al.* SQUID-Based Microwave Cavity Search for Dark-Matter Axions. *Phys. Rev. Lett.* **104**, 041301 (2010).
- 49. Asztalos, S. J. *et al.* Nuclear Instruments and Methods in Physics Research A. *Nucl. Instr. Meth. Phys. Res. A* **656**, 39–44 (2011).

50. Rosenberg, L. J. & van Bibber, K. A. Searches for invisible axions. *Phys. Rept.* **325**, 1–39 (2000).
51. Kolb, E. *The early universe* (Westview Press, New York, 1994).
52. Steigman, G., Dasgupta, B. & Beacom, J. F. Precise relic WIMP abundance and its impact on searches for dark matter annihilation. *Phys. Rev. D* **86**, 023506 (2012).
53. Higgs, P. W. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.* **13**, 508–509 (1964).
54. Aad, G. *et al.* Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B* **716**, 1–29 (2012).
55. Chatrchyan, S. *et al.* Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett. B* **716**, 30–61 (2012).
56. Murayama, H. Supersymmetry phenomenology. arXiv: hep-ph/0002232 [hep-ph] (2000).
57. Martin, S. P. A Supersymmetry primer. arXiv: hep-ph/9709356 [hep-ph] (1997).
58. Jungman, G., Kamionkowski, M. & Griest, K. Supersymmetric dark matter. *Phys. Rept.* **267**, 195–373 (1996).
59. Dawson, S. The MSSM and why it works. arXiv: hep-ph/9712464 [hep-ph] (1997).
60. Aad, G. *et al.* Search for new phenomena in final states with an energetic jet and large missing transverse momentum in pp collisions at $\sqrt{s}=8$ TeV with the ATLAS detector. *Eur. Phys. J. C* **75**, 299 (2015).
61. De Cosa, A. LHC results for dark matter from ATLAS and CMS. arXiv: 1510.01516 [hep-ex] (2015).
62. Adriani, O. *et al.* An anomalous positron abundance in cosmic rays with energies 1.5–100 GeV. *Nature* **458**, 607–609 (2009).
63. Aguilar, M. *et al.* First result from the Alpha Magnetic Spectrometer on the International Space Station: precision measurement of the positron fraction in primary cosmic rays of 0.5–350 GeV. *Phys. Rev. Lett.* **110**, 141102 (2013).
64. Goodenough, L. & Hooper, D. Possible Evidence For Dark Matter Annihilation In The Inner Milky Way From The Fermi Gamma Ray Space Telescope. arXiv: 0910.2998 [hep-ph] (2009).
65. Linden, T., Hooper, D. & Yusef-Zadeh, F. Dark matter and synchrotron emission from galactic center radio filaments. *Astrophys. J.* **741**, 95 (2011).
66. Abazajian, K. N., Canac, N., Horiuchi, S., Kaplinghat, M. & Kwa, A. Discovery of a new galactic center excess consistent with upscattered starlight. *Journal of Cosmology and Astroparticle Physics* **2015**, 013 (2015).
67. Drlica-Wagner, A. *et al.* Search for gamma-ray emission from DES dwarf spheroidal galaxy candidates with Fermi-LAT data. *Astrophys. J. Lett.* **809**, L4 (2015).

- 68. Ackermann, M. *et al.* Searching for Dark Matter Annihilation from Milky Way Dwarf Spheroidal Galaxies with Six Years of Fermi Large Area Telescope Data. *Phys. Rev. Lett.* **115**, 231301 (2015).
- 69. Saab, T. An Introduction to Dark Matter Direct Detection Searches & Techniques. arXiv: 1203.2566 [physics.ins-det] (2013).
- 70. Akerib, D. S. *et al.* New results from the Cryogenic Dark Matter Search experiment. *Phys. Rev. D* **68**, 082002 (2003).
- 71. Ahmed, Z. *et al.* Dark Matter Search Results from the CDMS II Experiment. *Science* **327**, 1619–1621 (2010).
- 72. Ahmed, Z. *et al.* Analysis of the low-energy electron-recoil spectrum of the CDMS experiment. *Phys. Rev. D* **81**, 042002 (2010).
- 73. Agnese, R. *et al.* Search for Low-Mass Weakly Interacting Massive Particles with SuperCDMS. *Phys. Rev. Lett.* **112**, 241302 (2014).
- 74. Sanglard, V. *et al.* Final results of the EDELWEISS-I dark matter search with cryogenic heat-and-ionization Ge detectors. *Phys. Rev. D* **71**, 122002 (2005).
- 75. Armengaud, E. *et al.* First results of the EDELWEISS-II WIMP search using Ge cryogenic detectors with interleaved electrodes. *Phys. Lett. B* **687**, 294–298 (2010).
- 76. Armengaud, E. *et al.* Final results of the EDELWEISS-II WIMP search using a 4-kg array of cryogenic germanium detectors with interleaved electrodes. *Phys. Lett. B* **702**, 329–335 (2011).
- 77. Aalseth, C. E. *et al.* Results from a Search for Light-Mass Dark Matter with a p-Type Point Contact Germanium Detector. *Phys. Rev. Lett.* **106**, 131301 (2011).
- 78. Angloher, G. *et al.* Results from 730 kg days of the CRESST-II Dark Matter search. *Eur. Phys. J. C* **72**, 1971 (2012).
- 79. Zhao, W. *et al.* First results on low-mass WIMPs from the CDEX-1 experiment at the China Jinping underground laboratory. *Phys. Rev. D* **88**, 052004 (2013).
- 80. Bernabei, R. *et al.* First results from DAMA/LIBRA and the combined results with DAMA/NaI. *Eur. Phys. J. C* **56**, 333–355 (2008).
- 81. Bernabei, R. *et al.* New results from DAMA/LIBRA. *Eur. Phys. J. C* **67**, 39–49 (2010).
- 82. Aprile, E. *et al.* Dark Matter Results from 225 Live Days of XENON100 Data. *Phys. Rev. Lett.* **109**, 181301 (2012).
- 83. Akerib, D. *et al.* First results from the LUX dark matter experiment at the Sanford Underground Research Facility. *Phys. Rev. Lett.* **112**, 091303 (2014).
- 84. Mallin, D. *et al.* After LUX: The LZ Program. arXiv: 1110.0103 [astro-ph.IM] (2011).

85. Alner, G. *et al.* First limits on nuclear recoil events from the ZEPLIN I galactic dark matter detector. *Astroparticle Physics* **23**, 444–462 (2005).
86. Alner, G. *et al.* First limits on WIMP nuclear recoil signals in ZEPLIN-II: A two-phase xenon detector for dark matter detection. *Astroparticle Physics* **28**, 287–302 (2007).
87. Agnes, P. *et al.* First Results from the DarkSide-50 Dark Matter Experiment at Laboratori Nazionali del Gran Sasso. *Phys. Lett.* **B743**, 456–466 (2015).
88. Rubbia, A. ArDM: A Ton-scale liquid Argon experiment for direct detection of dark matter in the universe. arXiv: hep-ph/0510320 [hep-ph] (2006).
89. Behnke, E. *et al.* First Dark Matter Search Results from a 4-kg CF₃I Bubble Chamber Operated in a Deep Underground Site. *Phys. Rev.* **D86**, 052001 (2012).
90. Barnabe-Heider, M. *et al.* Improved spin-dependent limits from the PICASSO dark matter search experiment. *Phys. Lett. B* **624**, 186–194 (2005).
91. Amole, C. *et al.* Dark Matter Search Results from the PICO-2L C₃F₈ Bubble Chamber. *Phys. Rev. Lett.* **114**, 231302 (2015).
92. Filippini, J. P. *A Search for WIMP Dark Matter Using the First Five-Tower Run of the Cryogenic Dark Matter Search* PhD thesis (University of California, Berkeley, 2008).
93. Helm, R. H. Inelastic and elastic scattering of 187-MeV electrons from selected even-even nuclei. *Phys. Rev.* **104**, 1466 (1956).
94. Lewin, J. & Smith, P. Review of mathematics, numerical factors, and corrections for dark matter experiments based on elastic nuclear recoil. *Astroparticle Physics* **6**, 87–112 (1996).
95. Dūda, G., Kemper, A. & Gondolo, P. Model-independent form factors for spin-independent neutralino–nucleon scattering from elastic electron scattering data. *Journal of Cosmology and Astroparticle Physics* **2007**, 012 (2007).
96. Cushman, P. *et al.* Snowmass CF1 Summary: WIMP Dark Matter Direct Detection. arXiv: 1310.8327 [hep-ex] (2013).
97. Engel, J. & Vogel, P. Spin-dependent cross sections of weakly interacting massive particles on nuclei. *Phys. Rev. D* **40**, 3132 (1989).
98. Tovey, D. R., Gaitskell, R., Gondolo, P., Ramachers, Y. & Roszkowski, L. A new model-independent method for extracting spin-dependent cross section limits from dark matter searches. *Phys. Lett. B* **488**, 17–26 (2000).
99. Ling, F.-S., Sikivie, P. & Wick, S. Diurnal and annual modulation of cold dark matter signals. *Phys. Rev. D* **70**, 123503 (2004).
100. Aalseth, C. E. *et al.* Search for an annual modulation in a p-type point contact germanium dark matter detector. *Phys. Rev. Lett.* **107**, 141301 (2011).

101. Ahmed, Z. *et al.* Search for annual modulation in low-energy CDMS-II data. arXiv: 1203.1309 [astro-ph.CO] (2012).
102. Navarro, J. F., Frenk, C. S. & White, S. D. A universal density profile from hierarchical clustering. *Astrophys. J.* **490**, 493 (1997).
103. Dutton, A. A. & Macciò, A. V. Cold dark matter haloes in the Planck era: evolution of structural parameters for Einasto and NFW profiles. *Monthly Notices of the Royal Astronomical Society* **441**, 3359–3374 (2014).
104. Bozorgnia, N. *et al.* Simulated Milky Way analogues: implications for dark matter direct searches. arXiv: 1601.04707 [astro-ph.CO] (2016).
105. Welliver, B. *Dedicated Searches for Low and High Mass WIMPs with the SuperCDMS Soudan iZIP Detectors* PhD thesis (University of Florida, 2016).
106. Yellin, S. Finding an upper limit in the presence of an unknown background. *Phys. Rev. D* **66**, 032005 (2002).
107. Sadoulet, B. *Prospects for detecting dark matter particles by elastic scattering* in *Proceedings of the 13th Texas Symposium on Relativistic Astrophysics* (1987), 14–19.
108. Shutt, T. *et al.* Development of a cryogenic dark matter detector. *IEEE Transactions on Nuclear Science* **37**, 547–552 (1990).
109. Shutt, T. *et al.* Measurement of ionization and phonon production by nuclear recoils in a 60 g crystal of germanium at 25 mK. *Phys. Rev. Lett.* **69**, 3425–3427 (1992).
110. Cabrera, B., Krauss, L. M. & Wilczek, F. Bolometric detection of neutrinos. *Phys. Rev. Lett.* **55**, 25–28 (1985).
111. Irwin, K. D., Nam, S. W., Cabrera, B., Chugg, B. & Young, B. A. A quasiparticle-trap-assisted transition-edge sensor for phonon-mediated particle detection. *Review of Scientific Instruments* **66**, 5322 (1995).
112. Hellmig, J. *et al.* The CDMS II Z-sensitive ionization and phonon germanium detector. *Nucl. Instr. Meth. Phys. Res. A* **444**, 308–311 (2000).
113. Abusaidi, R. *et al.* Exclusion Limits on the WIMP-Nucleon Cross Section from the Cryogenic Dark Matter Search. *Phys. Rev. Lett.* **84**, 5699–5703 (2000).
114. Thakur, R. *The Cryogenic Dark Matter Search Low Ionization-Threshold Experiment* PhD thesis (University of Illinois at Urbana-Champaign, 2015).
115. Moore, D. *A Search for Low-Mass Dark Matter with the Cryogenic Dark Matter Search and the Development of Highly Multiplexed Phonon-Mediated Particle Detectors* PhD thesis (California Institute of Technology, 2012).
116. Irwin, K. An application of electrothermal feedback for high resolution cryogenic particle detection. *Appl. Phys. Lett.* **66**, 1998–2000 (1995).
117. Ramo, S. Currents Induced by Electron Motion. *Proceedings of the IRE* **27**, 584–585 (1939).

118. Shutt, T. *et al.* A solution to the dead-layer problem in ionization and phonon-based dark matter detectors. *Nucl. Instr. Meth. Phys. Res. A* **444**, 340–344 (2000).
119. Hertel, S. *Advancing the Search for Dark Matter: from CDMS II to SuperCDMS* PhD thesis (Massachusetts Institute of Technology, 2012).
120. Lindhard, J., Nielsen, V., Scharff, M. & Thomsen, P. V. Integral equations governing radiation effects. *Danske Videnskab. Selskab. Mat. Fys. Medd.* **33** (1963).
121. Luke, P. Voltage-Assisted Calorimetric Ionization Detector. *Journal of Applied Physics* **64**, 6858–6860 (1988).
122. Pyle, M. *Optimizing the Design and Analysis of Cryogenic Semiconductor Dark Matter Detectors for Maximum Sensitivity* PhD thesis (Stanford University, 2012).
123. Agnese, R. *et al.* Demonstration of surface electron rejection with interleaved germanium detectors for dark matter searches. *Appl. Phys. Lett.* **103**, 164105 (2013).
124. Bailey, C. *The Cryogenic Dark Matter Search: First 5-Tower Data and Improved Understanding of Ionization Collection* PhD thesis (Case Western Reserve University, 2009).
125. Neutron background at the Canfranc underground laboratory and its contribution to the IGEX-DM dark matter experiment. *Astroparticle Physics* **21**, 523–533 (2004).
126. Golwala, S. *Exclusion Limits on the WIMP-Nucleon Cross Section from the Cryogenic Dark Matter Search* PhD thesis (University of California, Berkeley, 2000).
127. Akerib, D. S. *et al.* First Results from the Cryogenic Dark Matter Search in the Soudan Underground Laboratory. *Phys. Rev. Lett.* **93**, 211301 (2004).
128. Ahmed, Z. *et al.* Combined Limits on WIMPs from the CDMS and EDELWEISS Experiments. *Phys. Rev. D* **84**, 011102 (2011).
129. Ahmed, Z. *et al.* Search for inelastic dark matter with the CDMS II experiment. *Phys. Rev. D* **83**, 112002 (2011).
130. Ahmed, Z. *et al.* Search for Axions with the CDMS Experiment. *Phys. Rev. Lett.* **103**, 141802 (2009).
131. Agnese, R. *et al.* First Direct Limits on Lightly Ionizing Particles with Electric Charge Less than $e/6$. *Phys. Rev. Lett.* **114**, 111302 (2015).
132. Agnese, R. *et al.* Silicon detector dark matter results from the final exposure of CDMS II. *Phys. Rev. Lett.* **111**, 251301 (2013).
133. Agnese, R. *et al.* New Results from the Search for Low-Mass Weakly Interacting Massive Particles with the CDMS Low Ionization Threshold Experiment. *Phys. Rev. Lett.* **116**, 071301 (2016).
134. Agnese, R. *et al.* Search for Low-Mass Weakly Interacting Massive Particles Using Voltage-Assisted Calorimetric Ionization Detection in the SuperCDMS Experiment. *Phys. Rev. Lett.* **112**, 041302 (2014).

135. Ahmad, Q. *et al.* Direct evidence for neutrino flavor transformation from neutral-current interactions in the Sudbury Neutrino Observatory. *Phys. Rev. Lett.* **89**, 011301 (2002).
136. Shockley, W. Currents to Conductors Induced by a Moving Point Charge. *Journal of Applied Physics* **9**, 635 (1938).
137. Spieler, H. *Semiconductor detector systems* (Oxford University Press, 2005).
138. Daal, M., Mirabolfathi, N. & Sundqvist, K. M. Ionization collection in regions of distorted electric field in the CDMS ZIP detector. *Nucl. Instr. Meth. Phys. Res. A* **559**, 408–410 (2006).
139. Gray, P. R., Hurst, P., Meyer, R. G. & Lewis, S. *Analysis and design of analog integrated circuits* (Wiley, 2001).
140. Mimura, T. The early history of the high electron mobility transistor (HEMT). *IEEE Transactions on Microwave Theory Techniques* **50**, 780–782 (2002).
141. Van der Ziel, A. *Noise in solid state devices and circuits* (Wiley-Interscience, 1986).
142. Dong, Q., Liang, Y. X., Gennser, U., Cavanna, A. & Jin, Y. The Role of the Gate Geometry for Cryogenic HEMTs: Towards an Input Voltage Noise Below $0.5 \text{ nV}/\sqrt{\text{Hz}}$ at 1 kHz and 4.2 K. *J. Low Temp. Phys.* **167**, 626–631 (2012).
143. Dong, Q. *et al.* Ultra-low noise high electron mobility transistors for high-impedance and low-frequency deep cryogenic readout electronics. *Appl. Phys. Lett.* **105**, 013504 (2014).
144. Phipps, A., Jin, Y. & Sadoulet, B. Ionization Readout of CDMS Detectors with Low Power, Low Noise HEMTs. *J. Low Temp. Phys.* **176**, 470–475 (2014).
145. Phipps, A., Sadoulet, B., Juillard, A. & Jin, Y. An HEMT-Based Cryogenic Charge Amplifier for Sub-kelvin Semiconductor Radiation Detectors. *J. Low Temp. Phys.* 1–7 (2016).
146. Sundqvist, K. *Carrier Transport and Related Effects in Detectors of the Cryogenic Dark Matter Search* PhD thesis (University of California, Berkeley, 2012).
147. Varshni, Y. P. Temperature dependence of the energy gap in semiconductors. *Physica* **34**, 149–154 (1967).
148. Moffatt, R. *Two-Dimensional Spatial Imaging of Charge Transport in Germanium Crystals at Cryogenic Temperatures* PhD thesis (Stanford University, 2016).
149. Lundstrom, M. *Fundamentals of carrier transport* (Cambridge University Press, 2009).
150. Ashcroft, N. & Mermin, N. *Solid State Physics* (Saunders College, Philadelphia, 1976).
151. Moffatt, R. A. *et al.* Imaging the oblique propagation of electrons in germanium crystals at low temperature and low electric field. *Appl. Phys. Lett.* **108**, 022104 (2016).
152. Dresselhaus, G., Kip, A. F. & Kittel, C. Cyclotron Resonance of Electrons and Holes in Silicon and Germanium Crystals. *Phys. Rev.* **98**, 368–384 (1955).

153. Jacoboni, C. & Reggiani, L. The Monte Carlo method for the solution of charge transport in semiconductors with applications to covalent materials. *Rev. Mod. Phys.* **55**, 645–705 (1983).
154. Fischetti, M. V. Monte Carlo simulation of transport in technologically significant semiconductors of the diamond and zinc-blende structures. I. Homogeneous transport. *IEEE Transactions on Electron Devices* **38**, 634–649 (1991).
155. Dirac, P. A. M. The Quantum Theory of the Emission and Absorption of Radiation. *Proceedings of the Royal Society of London Series A* **114**, 243–265 (1927).
156. Conwell, E. & Weisskopf, V. F. Theory of Impurity Scattering in Semiconductors. *Phys. Rev.* **77**, 388–390 (1950).
157. Sclar, N. Neutral Impurity Scattering in Semiconductors. *Phys. Rev.* **104**, 1559–1561 (1956).
158. Kittel, C. & Kroemer, H. *Thermal physics* (Macmillan, 1980).
159. Herring, C. & Vogt, E. Transport and Deformation-Potential Theory for Many-Valley Semiconductors with Anisotropic Scattering. *Phys. Rev.* **101**, 944–961 (1956).
160. Ridley, B. K. *Quantum processes in semiconductors* (Oxford University Press, 2013).
161. Ehrenreich, H. & Overhauser, A. Scattering of holes by phonons in germanium. *Phys. Rev.* **104**, 331 (1956).
162. Wolfe, J. P. *Imaging phonons: acoustic wave propagation in solids* (Cambridge University Press, 2005).
163. Madarasz, F. L. & Szmulowicz, F. Transition rates for acoustic-phonon-hole scattering in p-type silicon with nonparabolic bands. *Phys. Rev. B* **24**, 4611 (1981).
164. Lawætz, P. Low-field mobility and galvanomagnetic properties of holes in germanium with phonon scattering. *Phys. Rev.* **174**, 867 (1968).
165. Sundqvist, K. M. & Sadoulet, B. Detectors of the Cryogenic Dark Matter Search: Charge Transport and Phonon Emission in Ge <100> at 40mK. *J. Low Temp. Phys.* **151**, 443–447 (2008).
166. Cabrera, B., Pyle, M., Moffatt, R., Sundqvist, K. & Sadoulet, B. Oblique propagation of electrons in crystals of germanium and silicon at sub-Kelvin temperature in low electric fields. arXiv: 1004.1233 [astro-ph.IM] (2010).
167. Aubry-Fortuna, V. & Dollfus, P. Electron transport properties in high-purity Ge down to cryogenic temperatures. *Journal of Applied Physics* **108**, 123706 (2010).
168. Brandt, D. *et al.* Semiconductor phonon and charge transport Monte Carlo simulation using Geant4. arXiv: 1403.4984 [physics.ins-det] (2014).
169. Sundqvist, K. M. *et al.* A Measurement of Electron and Hole Drift Velocities in a Germanium <100> CDMS Detector, at a Temperature of 31 milliKelvin. *AIP Conf. Proc.* **1185**, 128 (2009).

170. Balkanski, M. & Wallis, R. F. *Semiconductor physics and applications* (Oxford University Press, 2000).
171. Haller, E. E., Hansen, W. L. & Goulding, F. S. Physics of ultra-pure germanium. *Advances in Physics* **30**, 93–138 (1981).
172. Wilson, D. K. Electron Spin Resonance Experiments on Shallow Donors in Germanium. *Phys. Rev.* **134**, A265–A286 (1964).
173. Nisida, Y. & Horii, K. Shallow Donor Levels of Germanium in an Intermediate Magnetic Field. *Journal of the Physical Society of Japan* **31**, 776–782 (1971).
174. Landsberg, P. T. *Recombination in semiconductors* (Cambridge University Press, 2003).
175. Koenig, S. H. Recombination of Thermal Electrons in *n*-Type Germanium below 10°K. *Phys. Rev.* **110**, 988–990 (1958).
176. Koenig, S. H. Hot and warm electrons—A review. *Journal of Physics and Chemistry of Solids* **8**, 227–234 (1959).
177. Lax, M. Cascade Capture of Electrons in Solids. *Phys. Rev.* **119**, 1502–1523 (1960).
178. Abakumov, V., Perel, V. I. & Yassievich, I. *Nonradiative recombination in semiconductors* (Elsevier, 1991).
179. Kogan, S. M. & Lifshits, T. M. Photoelectric Spectroscopy - A New Method of Analysis of Impurities in Semiconductors. *physica status solidi (a)* **39**, 11–39 (1977).
180. Sanada, T., Matsushita, K., Ohyama, T. & Otsuka, E. Donor to Acceptor Electron Transfer in Germanium at Low Temperatures. *Journal of the Physical Society of Japan* **45**, 501–504 (1978).
181. Bethe, H. Berechnung der elektronenaffinität des wasserstoffs. German. *Zeitschrift für Physik* **57**, 815–821 (1929).
182. Haegel, N. M., Beeman, J. W., Luke, P. & Haller, E. E. *Phys. Rev. B* **39**, 3677 (1989).
183. Gershenzon, E. M., Mel'nikov, A. P., Rabinovich, R. I. & Serebryakova, N. A. H⁻-like impurity centers and molecular complexes created by them in semiconductors. *Physics-Uspekhi* **23**, 684 (1980).
184. Taniguchi, M., Hirano, M. & Narita, S. Very Shallow Trapping State in Doped Germanium. *Phys. Rev. Lett.* **35**, 1095–1098 (1975).
185. Nakata, H., Ichikawa, Y., Otsuka, E. & Kobayashi, M. Shallow Positive Acceptor in Germanium Doped with Deep Zinc Impurity. *Journal of the Physical Society of Japan* **55**, 2859–2866 (1986).
186. Domange, J., Olivieri, E., Fourches, N. & Broniatowski, A. Thermally-Stimulated Current Investigation of Dopant-Related D⁻ and A⁺ Trap Centers in Germanium for Cryogenic Detector Applications. *J. Low Temp. Phys.* **167**, 1131 (2012).

187. Sclar, N. & Burstein, E. Impact ionization of impurities in germanium. *Journal of Physics and Chemistry of Solids* **2**, 1–23 (1957).
188. Phipps, A., Sadoulet, B. & Sundqvist, K. M. Observation of Impact Ionization of Shallow States in Sub-Kelvin, High-Purity Germanium. *J. Low Temp. Phys.* <<http://dx.doi.org/10.1007/s10909-016-1472-5>> (2016).
189. Palmier, J. F. Impurity ionization in n-type germanium. *Phys. Rev. B* **6**, 4557 (1972).
190. Gutierrez-D, E. A., Deen, J. & Claeys, C. *Low temperature electronics: physics, devices, circuits, and applications* (Academic Press, 2000).
191. Haynes, J. R. & Shockley, W. Investigation of Hole Injection in Transistor Action. *Phys. Rev.* **75**, 691–691 (1949).
192. Abraham, E. R. & Cornell, E. A. Teflon feedthrough for coupling optical fibers into ultrahigh vacuum systems. *Applied optics* **37**, 1762–1763 (1998).
193. Daal, M. *Kinetic Inductance Detectors for Dark Matter Searches* PhD thesis (University of California, Berkeley, 2015).
194. Oppenheim, A. V., Schafer, R. W., Buck, J. R., *et al.* *Discrete-time signal processing* (Prentice-hall Englewood Cliffs, 1989).
195. Phipps, A., Sundqvist, K. M., Lam, A. & Sadoulet, B. Optically Induced Measurement of Electron and Hole Drift Velocities in a Germanium <100> CDMS Detector at 50 mK. *J. Low Temp. Phys.* **167**, 1112–1118 (2012).
196. Piro, M. C., Broniatowski, A., Marnieros, S., Dumoulin, L. & Olivieri, E. *J. Low Temp. Phys.* **176**, 796 (2014).
197. Hampel, W. & Remsberg, L. P. Half-life of Ge-71. *Phys. Rev.* **C31**, 666–667 (1985).
198. Bearden, J. A. & Burr, A. F. Reevaluation of X-Ray Atomic Energy Levels. *Rev. Mod. Phys.* **39**, 125–142 (1967).
199. Norton, P. Photoconductivity from shallow negative donor ions in silicon: A new far-infrared detector. *Journal of Applied Physics* **47**, 308–320 (1976).
200. Brigham, E. *The Fast Fourier Transform* (Prentice-Hall, Englewood Cliffs, N.J, 1974).

Appendix A

Optimal Filtering

A.1 The Fourier Transform

Functions of time may be converted to functions of frequency using the Fourier transform, which essentially expands the function $g(t)$ into the basis of complex exponentials of the form $e^{j\omega t}$. We define the Fourier transform pair as

$$\begin{aligned}\tilde{g}(f) &= \int_{-\infty}^{\infty} dt g(t) e^{-j\omega t}, \\ g(t) &= \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \tilde{g}(f) e^{j\omega t},\end{aligned}\tag{A.1}$$

where $\omega = 2\pi f$. It is important to note that $\tilde{g}(f)$ is defined for both positive *and* negative frequencies. If $g(t)$ is a real-valued function, as is the case for the measured voltage output of an amplifier, the positive and negative frequency components are correlated and we have $\tilde{g}(f) = \tilde{g}^*(-f)$. The necessity of negative frequency components is easily understood by considering the formula for a real-valued sine wave written in terms of complex exponentials — $e^{j\omega t}$ and $e^{-j\omega t}$ must be summed to cancel their imaginary parts. A component of the signal thus exists at both positive and negative frequencies in this basis.

A.2 Noise and the Power Spectral Density

We understand noise as the random fluctuations of a measured quantity around its true value due to a variety of different processes. In the time domain, the noise at a time $t + \tau$ is generally not independent from the noise at time t . For example, a voltage fluctuation at the input of a charge amplifier will result in a pulse which decays with the amplifier time constant, hence some information about the noise at a time t still remains at time

$t + \tau$. Mathematically, these correlations in the noise $v(t)$ are encoded in the autocorrelation function, defined as

$$R(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} dt v(t) v(t + \tau). \quad (\text{A.2})$$

These same correlations can be expressed in the frequency domain. The Fourier transform of the autocorrelation function defines the *power spectral density* $J(f)$, which has units of V^2/Hz for voltage fluctuations. Experimentally, there is a much simpler procedure for determining the power spectral density, as derived in [126] — a large number of noise traces are acquired and the squared magnitude of their Fourier components averaged together.

Considerable amounts of time have been lost due to confusion surrounding the definition of the power spectral density, its units, the frequency domain over which it is defined, and its relationship to the standard noise formula we have burned into our memories. In particular, the mysterious factor of 4 which appears in our final formula for the optimal filter amplitude resolution (Eq. A.11) depends crucially on proper definition of this term.

Recall that the Fourier transform of a time domain signal results in a function which is defined for both positive and negative frequencies. The Fourier transform of Eq. A.2 thus defines the *double-sided power spectral density*:

$$J_{ds}(f) = \int_{-\infty}^{\infty} dt R(t) e^{-j\omega t}, \quad (\text{A.3})$$

in which the noise power is split between positive and negative frequency components. Most discussions of optimal filtering, however, instead use the *single-sided power spectral density*, which combines the positive and negative frequency components and is only defined for $f > 0$. The single-sided power spectral density $J_{ss}(f)$ is defined as

$$J_{ss}(f) = 2J_{ds}(f), \quad (\text{A.4})$$

for $f > 0$. The use of the single-sided distribution is because what we measure in experiment are real-valued signals, therefore $J_{ds}(f) = J_{ds}(-f)$. The noise power at a specific *magnitude* of frequency is twice the value of the positive frequency component and thus no information is lost through the use of the single-sided distribution.

To add to the confusion, it is typically $\sqrt{J_{ss}(f)}$ which is discussed in text and displayed in noise plots, with units of $\text{V}/\sqrt{\text{Hz}}$. This distribution is *also* referred to as the power spectral density with no distinction other than through the physical units, including within this dissertation. In particular, we note that the standard $\sqrt{4k_B T R}$ definition of Johnson noise is referring to $\sqrt{J_{ss}(f)}$ — the double-sided power spectral density (in V^2/Hz) is in fact $2k_B T R$. It is easy to see how factors of 2 can be gained or lost when performing calculations if care is not taken to use the proper definitions.

A.3 Discrete Sampling

In the real world, signals are not measured continuously nor for an infinite amount of time. The rigorous conversion from the continuous to the discrete case is not particularly illuminating but can be found in [126, 200]. Discrete, digitized signals are described by their sampling rate f_s and total trace length T . The total trace length is related to the number of sampled points N via

$$T = \frac{N}{f_s} = N\Delta t. \quad (\text{A.5})$$

Instead of a continuous function, $g(t_k)$ is a set of N points representing the signal at times separated by $\Delta t = 1/f_s$. In order to preserve our physical units, we use a somewhat non-standard definition of the discrete Fourier transform. The Fourier transform pair becomes

$$\begin{aligned} \tilde{g}(f_n) &= \frac{1}{N} \sum_{k=1}^N \Delta t g(t_k) e^{-j\omega_n t_k}, \\ g(t_k) &= \sum_{n=1}^N \Delta f \tilde{g}(f_n) e^{j\omega_n t_k}, \end{aligned} \quad (\text{A.6})$$

where f_n ranges from $-f_s/2$ to $f_s/2$ in steps of $\Delta f = f_s/N$ and $\omega_n = 2\pi f_n$. Rigorous definitions omit the Δt and Δf terms, with their proper inclusion being introduced in a more detailed conversion from continuous to discrete equations. For our purposes, it suffices to work entirely with the continuous equations, making the substitution to the discrete case via the replacement

$$\int_{-\infty}^{\infty} df \rightarrow \sum_{n=1}^N \Delta f \quad (\text{A.7})$$

in our final results.

The effects of a finite sampling rate and trace length are two-fold. By the Nyquist-Shannon sampling theorem, all information at frequencies above half the sampling rate ($f_s/2$) is lost. Note that this is due to the finite sampling rate and not the trace length. The finite trace length instead limits the frequency resolution to f_s/N , the spacing between points in the frequency domain. Increasing the trace length allows information at lower frequencies to be recorded, while simultaneously increasing the resolution between frequencies. High resolution is necessary to resolve peaks occurring specific frequencies, especially if multiple peaks are close together.

A.4 Optimal Amplitude Estimator

With the formalism we have defined, we now develop a procedure to determine the optimal amplitude estimator $\hat{A}$ for a measured pulse. We assume the pulse takes the form

$$v(t) = As(t) + n(t), \quad (\text{A.8})$$

where A is the amplitude of the pulse, $s(t)$ is the normalized signal template, and $n(t)$ is a specific noise realization drawn from the noise distribution described by $J(f)$. Note the assumption of linearity — all pulses have the same shape $s(t)$ which is scaled by the scalar amplitude A .

It can be shown for stationary, Gaussian noise that parameters determined by minimizing the χ^2 value of the fit is equivalent to the maximum likelihood estimate, hence the use of the term “optimal” when described this procedure. While the χ^2 can be defined in the time domain, for linear systems such as a charge amplifier there is significant advantage to working in the frequency domain as the noise components at different frequencies are independent. This useful property comes from the fact that the response of a linear system at frequency f is independent of the response at other frequencies, which is simply the definition of a linear system. We therefore define χ^2 to be

$$\chi^2 = \int_{-\infty}^{\infty} df \frac{|\tilde{v}(f) - A\tilde{s}(f)|^2}{J_{ds}(f)}. \quad (\text{A.9})$$

In the time domain, we would need to include cross-terms describing the correlations of the output voltage at different times. While it is true that the positive and negative frequency components of a real signal are correlated, Golwala[126] argues that this definition of the χ^2 is properly normalized as the noise at each positive frequency component has two degrees of freedom — amplitude and phase. We thus expect the average χ^2 value to be equal to twice the number of positive frequency components which is the total number of degrees of freedom. As the phase information is lost by taking the squared magnitude in Eq. A.9, only integrating over the positive frequency components would give an average χ^2 of half the expected value — the other half is recovered by integrating over the negative frequencies. Sadoulet’s internal CDMS notes on noise explicitly deal with splitting the noise into its real and imaginary components and all final results are unchanged. We therefore stick with Golwala’s definition for simplicity.

The optimum amplitude is easily found by minimizing χ^2 with respect to A , giving

$$\hat{A} = \frac{\int_{-\infty}^{\infty} df \left(\frac{\tilde{s}^*(f) \cdot \tilde{v}(f)}{J_{ds}(f)} \right)}{\int_{-\infty}^{\infty} df \left(\frac{|\tilde{s}(f)|^2}{J_{ds}(f)} \right)}. \quad (\text{A.10})$$

The amplitude estimator is unbiased — while the individual noise fluctuations will cause a spread in the predicted values of $\hat{A}$ for the same physical amplitude, the average of $\hat{A}$ is equal to A . The proof of this is left as an exercise for the reader.

In terms of amplifier performance, it is much more important to find the fluctuations of $\hat{A}$, as these determine the energy resolution of the measurement. The standard deviation of

the amplitude estimator is calculated to be

$$\begin{aligned}
\sigma_A^2 &= \left(\frac{1}{2} \frac{\partial^2 \chi^2}{\partial A^2} \right)^{-1} \\
&= \left(\int_{-\infty}^{\infty} df \frac{|\tilde{s}(f)|^2}{J_{ds}(f)} \right)^{-1} = \left(2 \int_0^{\infty} df \frac{|\tilde{s}(f)|^2}{J_{ds}(f)} \right)^{-1} \\
&= \left(4 \int_0^{\infty} df \frac{|\tilde{s}(f)|^2}{J_{ss}(f)} \right)^{-1} \rightarrow \left(4 \sum_{f_{min}}^{f_{max}} \Delta f \frac{|\tilde{s}(f)|^2}{J_{ss}(f)} \right)^{-1},
\end{aligned} \tag{A.11}$$

where the summation used in the discrete case is only over the positive f_n . Note the factor of 4 in the final result comes from two independent substitutions. As our expression only contains the squared magnitude of the signal and noise Fourier components, which are assumed to be real-valued in the time domain, we can integrate only over the positive frequencies and multiply the integral by 2 since all terms are equal for positive and negative frequencies. The second factor of 2 comes from switching from the double-sided power spectral density J_{ds} to the single-sided power density J_{ss} , as $J_{ds} = J_{ss}/2$. As described in Chapter 3, this equation is very powerful as it allows for the calculation of the best-case resolution achievable for a given amplifier topology.

A.5 Optimal Start Time Estimator

We have so far implicitly assumed that we know the start time of the pulse. The position of the start time will vary from trace to trace when dealing with real data. Shifting the signal template by a time offset t_0 is equivalent to multiplying by a phase offset $e^{-j\omega t_0}$ in the frequency domain. We must therefore redefine the χ^2 value to be

$$\chi^2(A, t_0) = \int_{-\infty}^{\infty} df \frac{|\tilde{v}(f) - Ae^{-j\omega t_0} \tilde{s}(f)|^2}{J_{ds}(f)}. \tag{A.12}$$

For a given time offset t_0 , the optimum amplitude is still found using Eq. A.10 by making the substitution $\tilde{s}^*(f) \rightarrow e^{j\omega t_0} \tilde{s}^*(f)$ in the numerator. As can be seen from the form of Eq. A.11, the inclusion of an arbitrary time offset does not change the resolution of the amplitude estimator as the time offset phase factor cancels in every term.

In theory, we can minimize χ^2 with respect to t_0 to find the optimal start time estimator $\hat{t}_0$. As described in [92], the resulting equation has no analytic, closed-form solution for $\hat{t}_0$. It can be shown, however, that the value $\hat{t}_0$ which minimizes χ^2 also maximizes $\hat{A}$. An efficient method for determining the time offset is to calculate Eq. A.10 using the time offset phase factor for a range of possible t_0 values, and setting $\hat{t}_0$ to be the specific value which gives the

maximum amplitude $\hat{A}$. This procedure introduces a slight complication when attempting to experimentally determine the optimal filter resolution using noise traces in which no pulse is present, as the fit will preferentially select a start time (and hence amplitude) corresponding to the time of the highest noise fluctuation. This produces a double-peaked distribution which is not centered around $\hat{A} = 0$. The solution is to simply force $\hat{t}_0 = 0$ when applying the fitting routine to noise traces.

The resolution of the time offset estimator may be calculated in a similar fashion to the resolution of the amplitude estimator. We find

$$\sigma_{t_0}^2 = \left(\frac{1}{2} \frac{\partial^2 [\chi^2(A, t_0)]}{\partial t_0^2} \right)^{-1} = \left(A^2 \int_{-\infty}^{\infty} df \omega^2 \frac{|\tilde{s}(f)|^2}{J_{ds}(f)} \right)^{-1}. \quad (\text{A.13})$$

Unlike the resolution of the amplitude estimator, the time offset estimator is not independent of the amplitude of the pulse. The resolution is improved for higher energy events. The determination of the resolution is also strongly weighted towards the high frequency components of the signal. Assuming $J(f)$ is the same for both cases, an amplifier with lower bandwidth will have worse time offset resolution than an amplifier with higher bandwidth due to the decreasing amplitude of $\tilde{s}(f)$ at high frequency.

A.6 Interpretation as a Filter

We have so far given no reason as to why this pulse fitting procedure is called “optimal filtering”. The connection is made by defining the *optimal filter* $\tilde{\phi}(f)$, evaluated at optimal time offset $\hat{t}_0$, as

$$\tilde{\phi}(f) = \frac{e^{j\omega\hat{t}_0} \tilde{s}^*(f)}{J_{ds}(f)} \quad (\text{A.14})$$

with corresponding time domain representation $\phi(t)$. Returning to our expression for the amplitude estimator $\hat{A}$, we see that Eq. A.10 may be rewritten as

$$\hat{A} = \frac{\int_{-\infty}^{\infty} df \tilde{\phi}(f) \cdot \tilde{v}(f)}{\int_{-\infty}^{\infty} df \tilde{\phi}(f) \cdot \tilde{s}(f)}. \quad (\text{A.15})$$

The product of functions in the frequency domain is equivalent to their convolution in the time domain. We see that the numerator is simply the convolution of the measured signal $v(t)$ with the optimal filter. The denominator is the convolution of the normalized signal template with the optimal filter, and serves as a normalization factor to account for the fact that filtering the pulse changes its amplitude. The optimal amplitude is therefore proportional to the amplitude of the measured trace after being passed through the optimal

filter, hence the name. It is important to note that the optimal filter is acasual and can only be applied after the measurement of the trace is complete. Real-time implementation of an optimal filter through analog electronics is impossible, however the ever-increasing speed of digital electronics is beginning to allow near-instantaneous calculation of the optimal filter parameters directly on board the readout electronics as soon as digitization is complete.

Appendix B

HEMT Calculations

B.1 Noise Contribution from the HEMT Cascode

The fully cryogenic HEMT amplifier shown in Fig. 4.9 makes use of a cascode formed by the Q_1 and Q_2 HEMTs. There has been concern that noise present on the gate of Q_2 may appear on the output in comparable magnitude to the input noise of the Q_1 HEMT. In this section, we show that the voltage gain for a signal appearing on the gate of Q_2 is negligible in comparison to the gain of a signal appearing on the input (gate of Q_1).

The equivalent circuit for a signal v_{cg} appearing on the gate of Q_2 is shown in Fig. B.1. R_L is the effective load resistance formed by the Q_3 and Q_4 HEMTs of the fully cryogenic amplifier (not shown). Note that the gate of Q_1 , the input of the amplifier, is grounded as there is no input signal. This eliminates the transconductance of Q_1 in the small-signal limit as there is no change in the Q_1 gate-source voltage, however we still must consider its output resistance r_{o1} . The topology is equivalent to a source-loaded common source amplifier.

The transconductance of Q_2 causes v_{cg} to induce a signal current i_{ds} . Neglecting the output resistance of Q_2 , which does not significantly alter the final result, we have the following relations:

$$i_{ds} = g_{m2} (v_{cg} - v_s), \quad (\text{B.1})$$

$$= v_s / r_{o1}, \quad (\text{B.2})$$

$$= -v_{out} / R_L. \quad (\text{B.3})$$

Eq. B.2 and Eq. B.3 may be combined to solve for v_s , giving

$$v_s = i_{ds} r_{o1} = -v_{out} \frac{r_{o1}}{R_L}. \quad (\text{B.4})$$

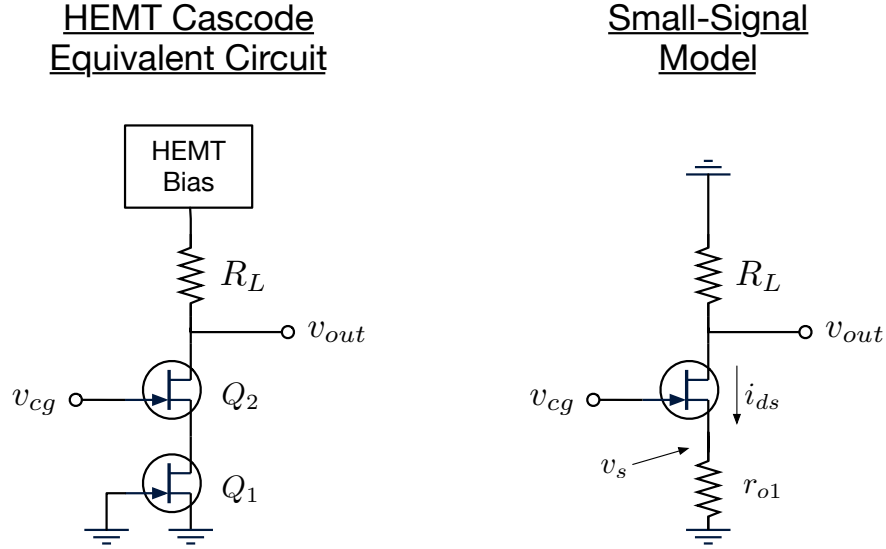


Figure B.1: The equivalent circuit and its small-signal model for a signal appearing on the gate of Q_2 .

We now have

$$\begin{aligned} \frac{-v_{out}}{R_L} &= g_{m2} (v_{cg} - v_s) \\ &= g_{m2} \left(v_{cg} + v_{out} \frac{r_{o1}}{R_L} \right). \end{aligned} \quad (\text{B.5})$$

This can be rearranging to solve for the voltage gain. We find

$$\frac{v_{out}}{v_{cg}} = \frac{-g_{m2} R_L}{1 + g_{m2} r_{o1}}. \quad (\text{B.6})$$

The voltage gain has been reduced by a factor $(1 + g_{m2} r_{o1})$. For the fully cryogenic HEMT amplifier, we use $r_{o1} = 1250 \, \Omega$, $g_{m2} = 110 \, \text{mS}$, $R_L = 10 \, \text{k}\Omega$ and find the gain to be reduced by a factor of about 140 to give $v_{out}/v_{cg} \approx -8$. Compared to the open loop voltage gain of a signal appearing at the amplifier input of 340, similar noise on the gate of Q_2 would represent an increase in the output noise level of

$$\frac{\sqrt{340^2 + 8^2}}{340} \approx 1.0003. \quad (\text{B.7})$$

Even relatively high noise levels on the gate of Q_2 are completely inconsequential in determining the total output noise. Similar arguments may be used to show that the noise

of Q_3 and Q_4 are even less significant. As we have seen in Chapter 4, the output noise of the fully cryogenic HEMT amplifier is completely determined by the input noise and is ultimately limited by the noise of the input (Q_1) HEMT.

B.2 Modified Effective Load Resistor Topology for Improved Gain

The effective load resistor formed by Q_3 and Q_4 in the fully cryogenic HEMT amplifier (Fig. 4.9) has several properties which can be improved upon. The effective load resistance is only $\sim 10 \text{ k}\Omega$, limiting the open loop voltage gain to just several hundred. Since we would like a closed loop voltage gain of about one hundred, this causes the closed loop amplifier response to have a significant dependence on the open loop gain. Increasing the open loop gain limits these effects and would result in a more predictable amplifier response which is less sensitive to the individual HEMT characteristics.

Both Q_3 and Q_4 need to be placed into saturation for proper small-signal operation. The saturated channel current depends strongly on the gate-source voltage, and varies little with drain-source bias. As the gate-source voltages of Q_3 and Q_4 are set independently by R_{s3} and R_{s4} in our original design, the HEMTs must be well-matched in order to properly function. This creates difficulties with DC biasing as several different values of R_{s3} and R_{s4} must be tested to maximize the effective load resistance. We had some trouble as our Q_3 and Q_4 HEMTs were not particularly well-matched, and often one HEMT would not enter saturation, degrading the open loop voltage gain. Each resistor also adds an additional $100 \text{ }\mu\text{W}$ of power dissipation, which we would like to avoid.

An improved topology which addresses these issues is shown in Fig. B.2. While the Q_3 HEMT topology remains unchanged, we remove R_{s4} and instead connected the gate of Q_4 to the source of Q_3 . This sets the magnitude of the gate-source voltage of Q_4 equal to the drain-source voltage of Q_3 , both on the order of 100 mV at typical biases for the 30 pF HEMTs. Note that it is because of this similarity that topology works — HEMTs with very different gate-source and drain-source values will not be properly DC biased. The advantage to this design is that the DC biasing of Q_4 now depends on the biasing of Q_3 , which is set by R_s and the bias supply. The bias supply can be increased until Q_3 reaches saturation, effectively fixing the source voltage of Q_3 . As the supply voltage is increased further, the voltage at the drain of Q_3 /source of Q_4 increases which can proceed until Q_4 is also in saturation. The elimination of R_{s3} also reduces the power consumption of the amplifier.

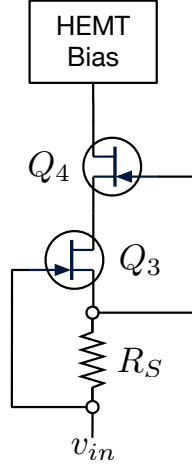
The coupling of Q_3 and Q_4 enhances the effective load resistance. Referring to the small-signal model of Fig. B.2, we have

$$i_{ds} = g_{m4}(v_{s3} - v_{s4}) - v_{s4}/r_{o4}, \quad (\text{B.8})$$

$$= g_{m3}(v_{in} - v_{s3}) + (v_{s4} - v_{s3})/r_{o3}, \quad (\text{B.9})$$

$$= (v_{s3} - v_{in})/R_s. \quad (\text{B.10})$$

Improved Effective Load Resistor Circuit



Small-Signal Model

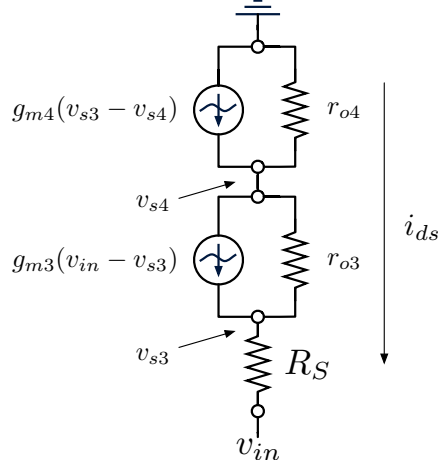


Figure B.2: The equivalent circuit and its small-signal model to boost the effective load resistance formed by Q_3 and Q_4 .

These equations can be combined to solve for the input impedance, which represents the effective load resistance seen by the output of the $Q_1 + Q_2$ HEMT cascode. We find

$$R_L = \left| \frac{v_{in}}{i_{ds}} \right| = g_{m3} g_{m4} r_{o3} r_{o4} R_s + g_{m3} r_{o3} R_s + g_{m4} r_{o4} r_{o3} + r_{o3} + r_{o4} + R_s \quad (\text{B.11})$$

$$\approx 770 \text{ k}\Omega$$

using $g_{m3} = g_{m4} = 110 \text{ mS}$, $r_{o3} = r_{o4} = 770 \Omega$, and $R_s = 97 \Omega$. This is significantly improved over the $\sim 100 \text{ k}\Omega$ effective load resistance of our original design, and boosts the open loop voltage gain to over 3000.

This new topology is clearly superior to our original design discussed in Chapter 4. These design changes have been incorporated into the next version of the fully cryogenic HEMT amplifier, which is still under development. If testing is succesful, the next step would be to fabricate the complete amplifier design on a single chip rather than using individual discrete HEMTs.

Appendix C

Electrical Shorts at Soudan

During CDMS II and SuperCDMS Soudan operations, a significant fraction of detector readout channels were found to be electrically shorted to the chassis of the cryostat. Shorted channels are essentially unusable due to non-standard pulse shapes, high levels of noise injection, and cryostat heating due to oscillating amplifiers. Shorted channels were left unbiased or completely disabled in order to prevent loss of live time caused by high trigger rates. While some effort has been made to analysis data taken from the detectors with shorted channels, there is considerable risk in including this data in a WIMP search. For example, the excess of events observed in the SuperCDMS Soudan low threshold analysis are almost all from a detector with a shorted outer charge electrode, and in hindsight this detector should have probably been excluded from the analysis. In this appendix, we discuss possible causes of these shorts.

C.1 CDMS Cold Hardware

CDMS II and SuperCDMS Soudan shared the same cold hardware, with the only major difference being the size of the detector. A cross sectional view of the various hardware components is shown in Fig. C.1. The detector stack is located near the bottom of the icebox and is physically connected to the base temperature stage of the six-sided tower via the connector tube. Aluminum wirebonds electrically connect the detector sensors to the detector interface board (DIB), located inside the detector housing. The neutralization LEDs are also located on the DIB. Pins on the detector end of the side coax assembly are inserted into the DIB sockets, bringing the electrical signals away from the detector. The bias resistors, feedback resistors, and coupling capacitors for the ionization readout channel are also located on the side coax assembly, near the detector end. The remainder of the side coax assembly consists of NbTi wires which carry the signals to the base temperature stage of the tower, through which the connection is made through another set of pins and sockets. The tower wiring is also made from NbTi and bring the signals from base temperature to the 4 K SQUET assembly. The tower wires are heatsunk to the various thermal stages of

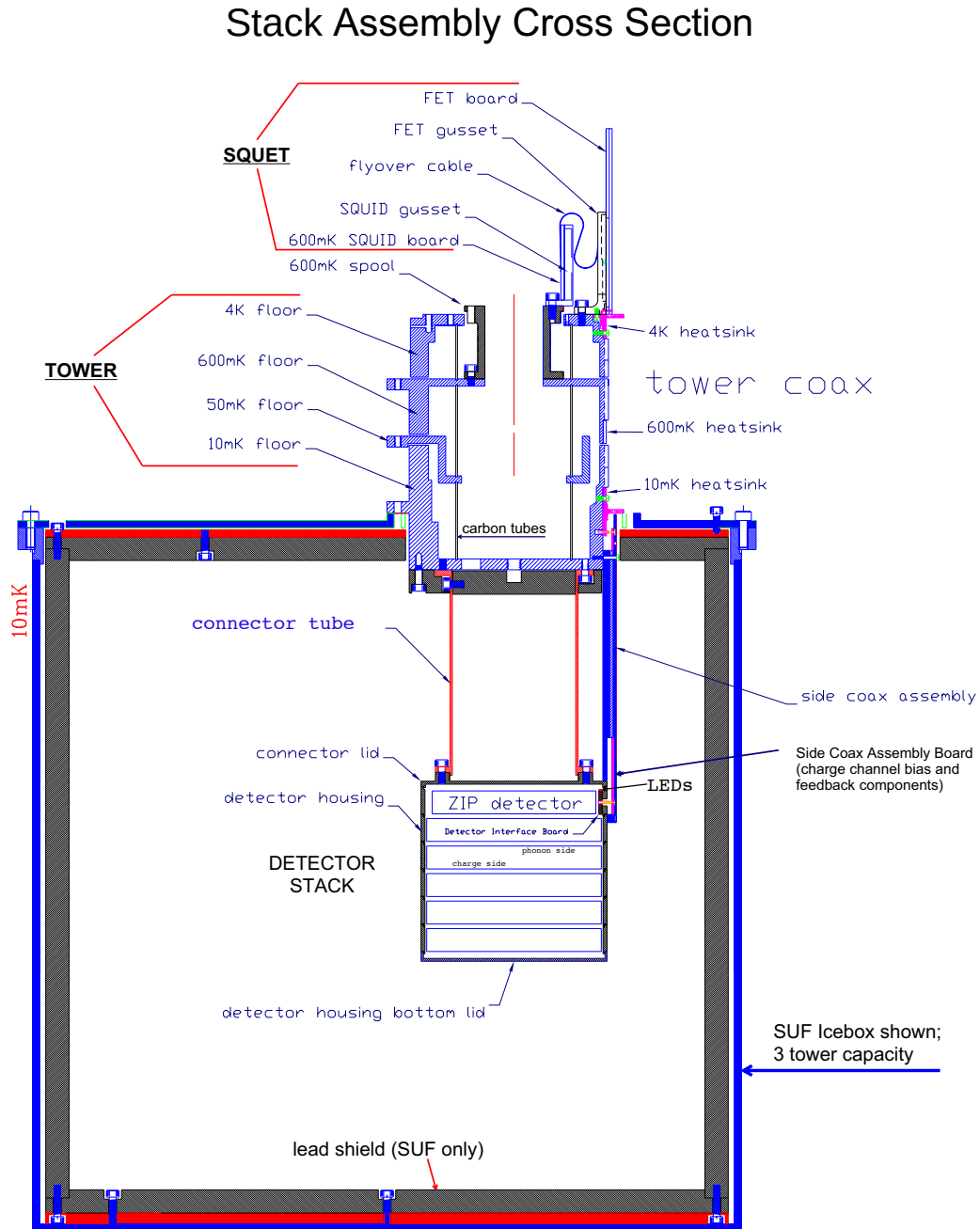


Figure C.1: Cross sectional view of the CDMS tower and associated cold hardware components. Figure courtesy Dennis Seitz.



The SQUET (SQUID+FET) assembly, shown in Fig. C.2 plugs into the tower and consists of two parts. The SQUID board holds the SQUID chips which are used to measure the phonon signals produced by the detector QET sensors. The SQUID board is heatsunk to the 600 mK temperature stage of the tower and is connected to the main SQUET card via a small flyover cable. The main SQUET card houses the JFETs used to sense the ionization signal. The JFETs are placed on a small Kapton membrane to allow them to self-heat to an operating temperature of 150 K. All signal lines are brought to a custom 50-pin connector at the top of the FET card. The FET card is then connected to a stripline, which brings the signals from 4 K to the 300 K vacuum bulkhead feedthrough. The room temperature side of the feedthrough is finally connected to the warm temperature electronics boards.

For the oZIPs and mZIPs, a single DIB-side coax-tower face-SQUET-stripline combination was enough to read out a single detector, allowing for six detectors to be placed in the stack. Due to the doubling in the number of detector sensors from CDMS II to SuperCDMS Soudan, two sets of hardware are required for a single iZIP detector. As two sets of tower face wiring are now required, only 3 iZIP detectors can be placed in the stack.

Returning to the side coax assembly, the side coax assembly board (SCAB) is shown in Fig. C.3. The SCAB is physically located adjacent to the detector housing, and the pins which plug into the DIB sockets are visible at the bottom of the figure. At the right of the

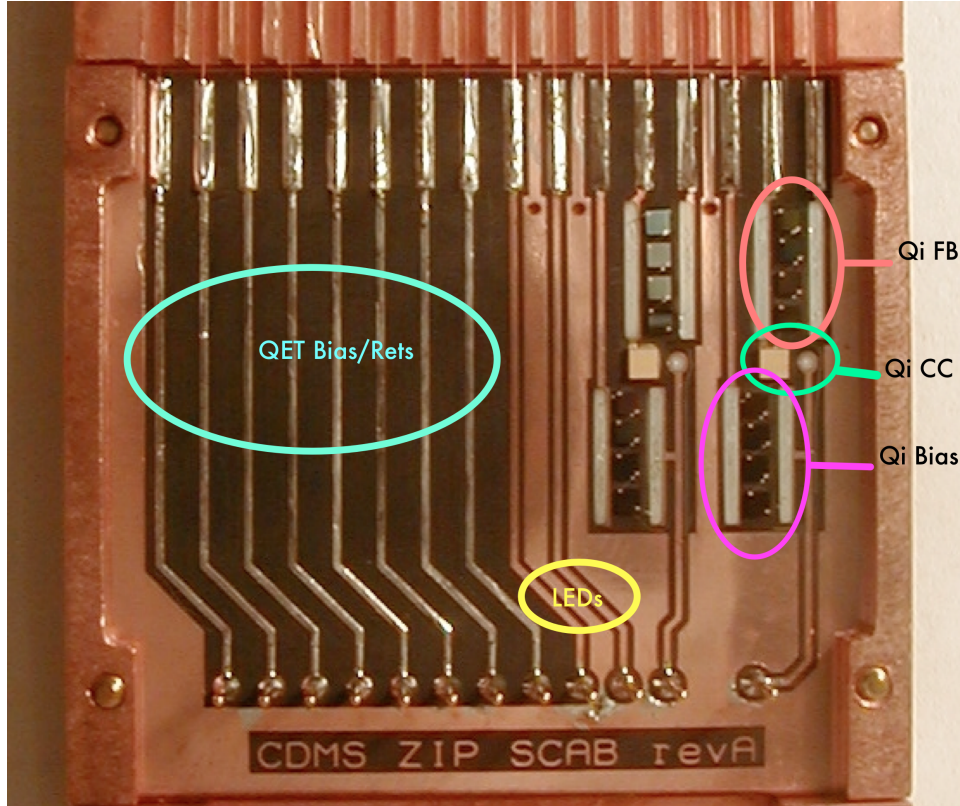


Figure C.3: Side coax assembly board.

figure, the individual components which make up the Q_{inner} bias resistor, feedback resistor, and coupling capacitor are circled. An uncircled, identical set of components can be seen for the Q_{outer} channel. Note that the Q_{inner} signal lines are closest to the *outer* edge of the SCAB, which sometimes can cause confusion. The signal lines for the QETs and LEDs are also circled. At the very top of the figure, the NbTi side coax wiring is visible. Each wire is centered to run through a square copper trench, approximating a coaxial cable. The lack of a physical insulator between the wire and trench has resulted in these lines (and sometimes the whole assembly) being called vacuum coaxes.

C.2 Characteristics

While chassis shorts have been a persistent issue at Soudan, their effect was particularly noticeable during the operation of the iZIPs. Half of the 15 detector payload experienced at least one channel which was shorted at some point during operations. There were a total of 18 chassis shorts reported, summarized in Table C.1, with 14 occurring on DIB1 and 4 on DIB2 (each iZIP has 2 DIBs). We also note that we never experience “opens”, which would

Type of Short	Number of Wires	Number Observed in R133-R134
Charge	6	8
Phonon	8	8
LED	2	2

Table C.1: Number of observed shorts during Soudan R133-134, organized by the type of signal affected. The number of wires refers to the number of tower/side coax wiring for a single tower face.

indicate an unexpected break in the signal line. It is solely unintended electrical connections to the cryostat chassis which are the problem.

The resistance of the shorts varies from several Ω to hundreds of $k\Omega$. They are not constant in time, and may suddenly disappear after remaining stable for months. A channel which has been shorted in the past is more likely to become shorted again, often with a similar resistance. The shorts also vary between thermal cycles of the cryostat, although it is unclear whether this is related to changes in temperature or somehow associated with the large amounts of physical labor in the vicinity of the icebox which typically occurs in parallel with a warm up/cool down.

All of the shorts occur on signal lines which extend from 4K to the detector, which we will refer to as the detector signals. These include the QET biases and returns, the LED lines, and the charge bias, JFET gate, and feedback lines. We never observe problems with lines which only extend from room temperature to 4K, such as the SQUID biases/returns and JFET drain/source. We refer to these lines as readout signals.

One final peculiarity of the Soudan shorts is that this seems to be the only location where they occur. CDMS test facilities, such as the one here at UC Berkeley, do not experience the same problems with chassis shorts. This is true even though test facilities are thermal cycled *significantly* more often than Soudan. There is one important difference between Soudan setup and the test facilities — at Soudan, the end of the detector stack points towards the floor, while at the test facilities, the tower is inverted and the end points towards the ceiling. Test facilities do not have the luxury (curse?) of a large icebox, and the detector stack/tower are inverted so that they may be placed directly under the dilution refrigerator mixing chamber. The cause of the shorts would hopefully explain this discrepancy.

C.3 Possible Causes

Looking at the block diagram of the individual cold hardware subsystems, shown in Fig. C.5, we are inclined to only consider those which are at base temperature. It is highly unlikely that we would only observe detector signal shorts if they are originating in a subsystem which

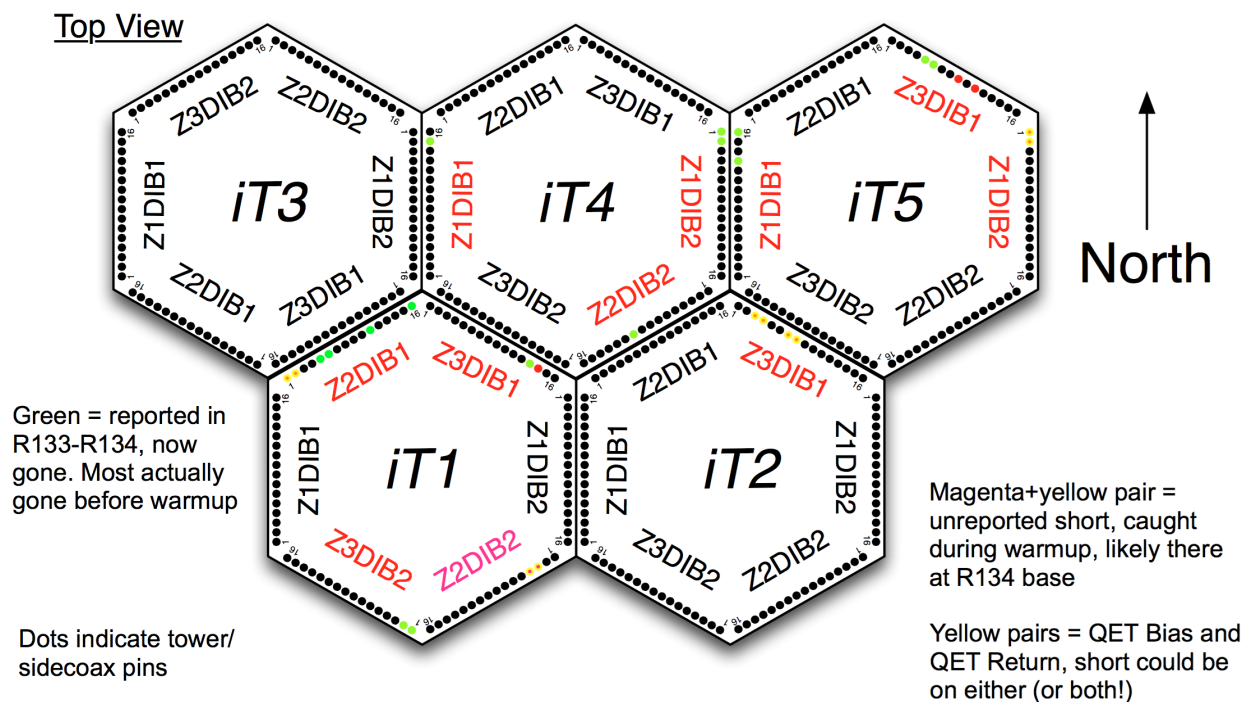


Figure C.4: Physical distribution of the shorts amongst the towers, and their status before warming up at the end of R134. The numbered pins on each tower face correspond to the side coax/tower wiring pin assignments.

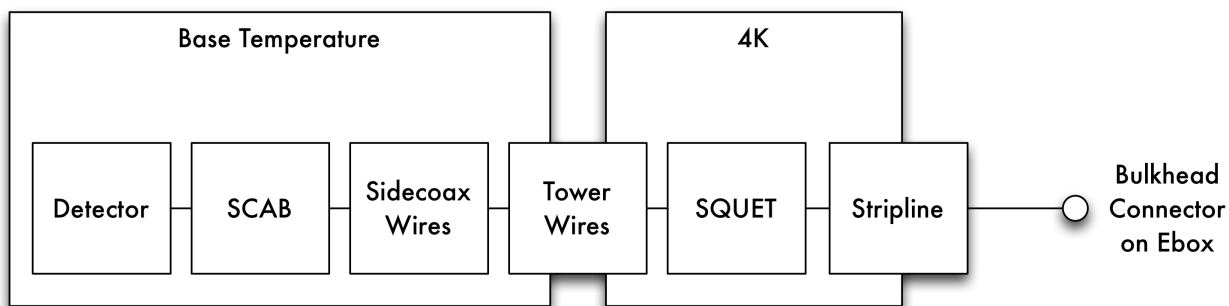


Figure C.5: Various cold hardware subsystems, separated by temperature.

also includes readout signals. There appears to be no apparent bias in the physical layout of the 4K and higher subsystems which would prevent the readout lines from also becoming shorted, so we do not consider them as possibilities for the location of the problem.

C.3.1 Loss of NbTi Wire Tension

The tower and side coax NbTi wires are soldered under tension to hold them in place. If these wires were to become slack, they may touch the chassis and cause a short. The most apparent reason for this to occur is due to differences in thermal contraction of the different materials which make up the tower and side coax. If the structural supports to which the wires are soldered contract more than the wires themselves, there will be a loss of tension.

For the tower wiring, the structural supports are carbon fiber rods to which the individual temperature stages of the tower are glued. The rods contract less than the wires, hence the tension increases as the experiment is cooled. We do not expect slack wires in the tower to be the cause of the chassis shorts.

The situation is more subtle for the side coax wiring. The underlying support is copper, which contracts more than the wires and we expect the tension to decrease. When soldered in place, however, the tension is set purposely high to account for losses due to thermal contraction. Complicating matters further, it seems that whether the pre-tensioning is high enough depends upon the work condition of the wire. If the wire is in a wrought condition, as is suspected, then there is no problem. If the wire is actually in an annealed condition, then it is possible enough tension is lost such that the wire may touch the chassis.

The evidence supporting this hypothesis is that the shorts seem to disappear upon warming up and extracting the towers. We note, however, there is significant physical movement around the icebox during warmup and extraction which may also cause the shorts to disappear. As we discuss in the next section, we also see shorts which occur after the 40 M Ω bias resistor, which cannot occur through slack side coax wiring. It is difficult for me to see how the intermittent nature of the shorts can occur while the fridge remains at a fixed temperature. Finally, we would expect to see the same problem at the test facilities.

C.3.2 Metal Debris

Numerous screws are used to connect all the various pieces of cold hardware to the cryostat. When inserting and removing screws, small flakes of metal debris are produced by the friction between the screw and the threads of the hole. This debris can land in the space between a signal trace and the chassis, especially in areas where the tolerances are small, and create a short. In particular, we note that proximity of the chassis plane of the SCAB board to the signal traces, especially for the charge signal lines.

There are several observations which support the debris hypothesis. Some of the observed chassis shorts have occurred between the 40 M Ω charge bias resistor and the detector electrode itself. This mode of failure can only occur if the short is directly on the SCAB, the DIB, or the detector. It is impossible for slack wiring to be the cause of these shorts.

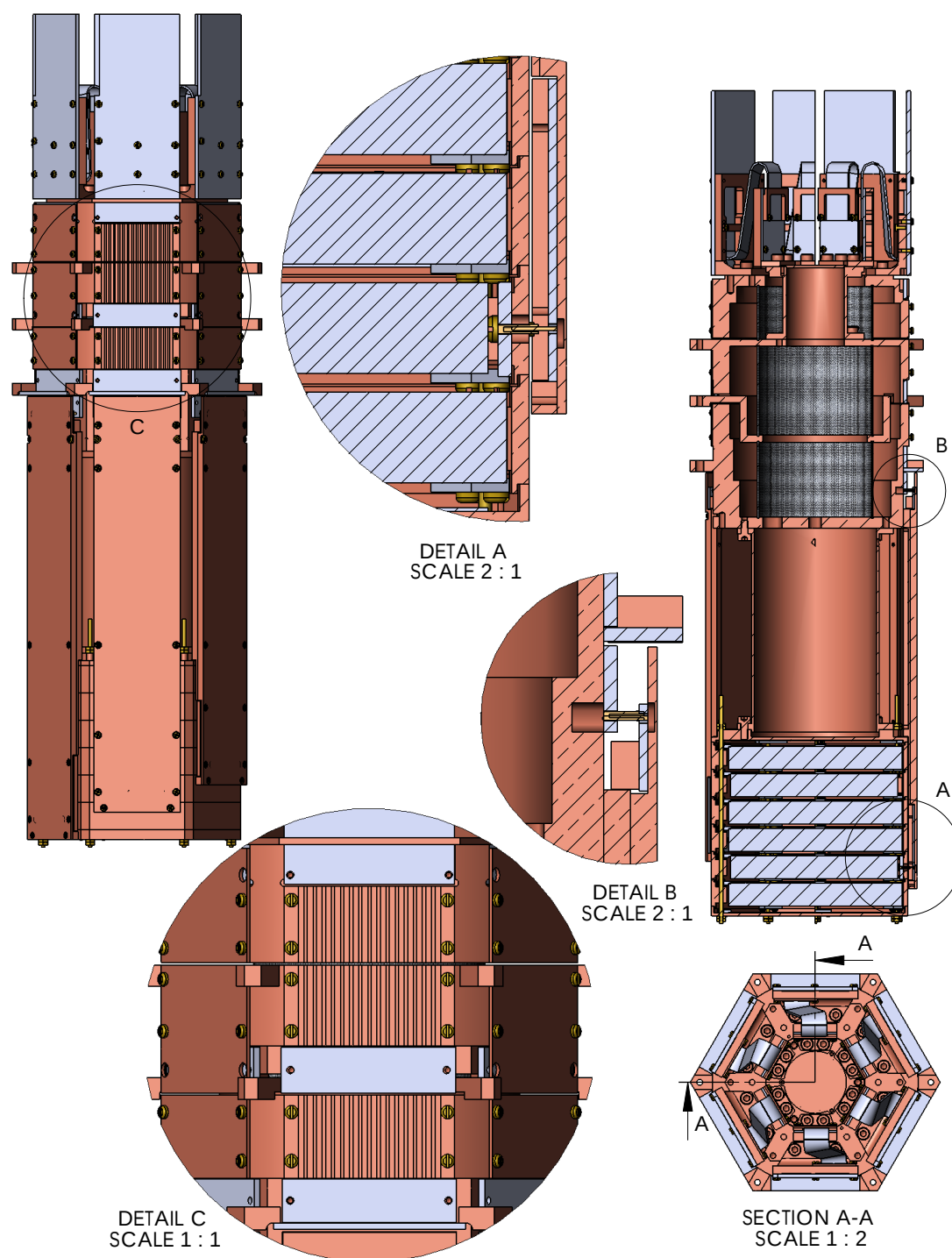


Figure C.6: Rendered views of an assembled CDMS tower. Figure courtesy of Miguel Daal.

If the debris is generated during the installation of the towers and lands upon the top of the detector stack, vibrations will cause the debris to migrate and fall off the edges. This falling debris could get caught in the gaps between the side coax SCAB and the detector housing facing, eventually inducing shorts on the SCAB. As we can see in Fig. C.6, the base temperature tower “ears” which screw into the cryostat cans are in direct line-of-sight with the top of the detector stack. They are also in an excellent position for the debris to fall into the gap between the SCAB board and the detector housing. This may explain why we tend to see shorts accumulate over time, and why they can be intermittent. Several shorts have been observed to appear and disappear by banging on the c-stem using a rubber mallet, which to me would indicate the movement of debris being the source of the problem.

The debris hypothesis also provides explanation for the differences between the Soudan installation and the test facilities. As the towers are mounted upside-down at the test facilities, debris generated by the screws falls away from the detector stack, towards the floor. Anyone experienced with CDMS tower installation at a test facility will attest to the fact that you can feel the debris falling on your face as you screw in the tower — presumably, this debris must be falling towards the detector stack at Soudan.

While we do see signs of debris upon extraction of the Soudan towers, all the shorts seem to disappear upon warm up. This would tend to support the slack NbTi wire hypothesis over the debris hypothesis, but all of the vibrations coupled with the changing temperature and pressure of the environment surrounding the towers make cause the debris to clear itself from the signal lines. In any case, now that operations at Soudan are over, forensic efforts are being made to identify the root cause of the Soudan shorts. This is not expected to be a problem for SuperCDMS SNOLAB due to changes in the cold hardware design.