Title
High Pressure Xenon Detectors for Rare Physics Searches

Permalink
https://escholarship.org/uc/item/88j6318f

Author
Renner, Joshua Edward

Publication Date
2014

Peer reviewed|Thesis/dissertation
High Pressure Xenon Detectors for Rare Physics Searches

By

Joshua Edward Renner

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 James Siegrist, Chair
Professor Yury Kolomensky
Professor Kai Vetter

Spring 2014
High pressure xenon gas detectors have the potential to significantly contribute to searches for rare physics processes such as neutrinoless double beta ($0\nu\beta\beta$) decay and interactions of dark matter with ordinary matter. We summarize the physics and implications of these two phenomena and discuss experimental developments conducted with an electroluminescent high pressure xenon time projection chamber (TPC).

The detector was constructed as a prototype for the NEXT experiment - Neutrino Experiment with a Xenon TPC - and was intended to demonstrate an ability to detect electron recoils with an energy resolution sufficient for a strong neutrinoless double beta decay search. Using the low-gain ionization signal amplification process of electroluminescence, we are able to obtain an energy resolution of approximately 1% FWHM in the detection of 662 keV gamma rays. In addition, we demonstrate basic particle tracking capabilities using a silicon photomultiplier-based tracking plane, a technique also relevant to the NEXT approach in searching for $0\nu\beta\beta$ decay.

To investigate the possibility of a gaseous xenon dark matter detector, we also characterize the ionization and scintillation signals produced by nuclear recoils in the TPC and compare them to the corresponding signals produced by electron recoils. The nuclear recoils are produced using radioisotope neutron sources. Our measurements demonstrate the ability to discriminate between electron and nuclear recoils using the ratio of ionization to scintillation produced on an event-by-event basis and provide an estimate of the ionization and scintillation yields of nuclear recoils in gaseous xenon.

We end with a discussion of the physical principles behind another approach to obtaining nearly-intrinsic energy resolution in a noble gas detector. This approach, called negative ion drift, consists of drifting negative ions produced by intentionally introducing impurities into a gaseous medium that capture ionization electrons. These
negative ions are then counted individually by detaching and amplifying the electron they carry. We develop a basic formalism in an attempt to understand the ion drift and detachment processes that are critical to the implementation of this technique.
This thesis is dedicated to the memory of my Grandfather, Allan Dale Schilling.
Acknowledgments

I must first thank Azriel Goldschmidt, who in practice advised me in the research leading to this thesis. Azriel has taught me many things about physics and life in general. His guidance was present in some way in all of my work as a graduate student, and none of it would have been possible without his instruction and encouragement.

And I thank Carlos Oliveira, who as both a friend and co-worker has made my time as a graduate student much more meaningful. Thanks for everything Carlos - for all of our adventures from Berkeley to Cacia and beyond.

I am also grateful for instruction and support from David Nygren. It was an honor to be lead by a scientist whose work and ideas have made such a significant impact in physics and science as a whole.

Thanks also to my advisor, Jim Siegrist, for introducing me to the lab and high pressure xenon research and keeping me on the right track for five years. Jim made time for me even amidst an extremely busy schedule and many responsibilities at the DOE.

The TPC at LBNL was a team effort, and this thesis would not have been possible without the contributions from all the members of the LBNL xenon group. Tom Miller and Derek Shuman provided excellent technical and engineering support, and designed/fabricated any piece of hardware we needed. Yasuhiro Nakajima made key contributions to the acquisition and analysis of the neutron data at times when it was becoming most difficult. Thanks to them, and to all others in the LBNL Xenon group for their contributions throughout the years, especially - Dennis Chan, Maxim Egorov, Victor Gehman, Daniel Hogan, Kevin Hsieh, Daniel Kiang, Megan Long, Howard Matis, and Tom Weber.

Our LBNL xenon group was fortunate to have become a part of the NEXT collaboration, and I am grateful to have been given the opportunity to visit the NEXT-IFIC group in Valencia, Spain several times in the past few years, and I thank my group and Juan Jose Gómez-Cadenas and the others at IFIC for allowing me this opportunity. I thank especially Francesc Monrabal, who introduced me to Spain, and in addition came to Berkeley to help assemble the TPC here. I look forward to working with all of you!

During most of my time as a graduate student I was supported by the DOE NNSA Stewardship Science Graduate Fellowship (SSGF), and so I thank the SSGF program, the other SSGF fellows, and the people from the Krell Institute who administered the fellowship, especially Lucille Kilmer. They took care of me for 4 years and brought me to interesting places to meet other researchers and discuss my research. They made my time as a graduate student much more enjoyable.

As an SSGF fellow I was also given the opportunity to work at Lawrence Livermore National Lab during the summer of 2010, and am grateful for the guidance and collaboration of Adam Bernstein, Mike Heffner, Peter Sorensen, and Melinda Sweany that arose from this project, and for the friendship and support of Bradley White.
and Broc Miramontes.

As an undergraduate at Georgia Tech, I was instructed in undergraduate research by John Wood, who also taught my first physics course there. We had many inspiring conversations and the things he taught me have echoed in my mind throughout undergraduate and graduate school and will continue to guide me as I continue on. I also worked many days alongside Scott Bustabad, and we had many discussions, often physics-related and often not! I am indebted to both Scott and John for their guidance during this time. I am also indebted to Karen Adams for her advice in making the transition from undergraduate to graduate school and to my roommates Ryan Curtin and Chad Kersey for their friendship and support.

During most of my time as a graduate student, I lived at the International House in Berkeley, CA. There I met many different people from around the world, and I thank all the friends who I have met there, especially Yi He, Dario Malnati, Markus Pelger, and Francesco Rossi, for all they have taught me.

And finally, I thank my parents and family for their constant support over the years, and for always welcoming me home no matter how far away I may go!
## Contents

1 Rare Physics ......................... 1
   1.1 Neutrinos and 0νββ ..................................... 1
      1.1.1 The Neutrino .................................... 1
      1.1.2 Fermions and Helicity ............................. 2
      1.1.3 Majorana Neutrinos ................................ 6
      1.1.4 The Neutrino Mass .................................. 8
      1.1.5 Neutrino Mixing and Oscillations ..................... 12
      1.1.6 Neutrinoless Double-Beta Decay ......................... 17
      1.1.7 Neutrinoless Double-Beta Decay: Experimental Considerations 21
      1.1.8 Neutrinoless Double-Beta Decay: Experiments .................. 27
      1.1.9 The NEXT-100 Experiment .............................. 29
   1.2 Dark Matter .......................... 32
      1.2.1 Evidence for Dark Matter ................................ 33
      1.2.2 ΛCDM: A Cosmological Model .......................... 37
      1.2.3 Dark Matter Constituents ............................... 39
      1.2.4 Direct Detection of WIMP Dark Matter ..................... 40
      1.2.5 WIMP Direct Detection Experiments ....................... 46

2 High Pressure Xenon Gas Detectors ..... 48
   2.1 Xenon Gas ............................. 48
   2.2 Energy Loss in Xenon Gas ..................... 49
      2.2.1 Ionization and Scintillation from Electron Recoils ............... 51
      2.2.2 Ionization and Scintillation from Nuclear Recoils ............... 53
      2.2.3 Electroluminescence .................................. 56
   2.3 Energy Resolution .......................... 56
      2.3.1 General Formalism .................................... 57
      2.3.2 Energy Resolution in Negative Ion TPCs ...................... 60
      2.3.3 Energy Resolution in Electroluminescent Xenon TPCs .............. 61
      2.3.4 Electroluminescent Gain, Charge Attachment and Multiplication 62
   2.4 Summary ..................................... 65
List of Figures

1.1 The possible configurations of neutrino helicity eigenstates for Dirac and Majorana neutrinos ........................................... 7
1.2 Illustration of the see-saw mechanism ........................................ 12
1.3 Probability of an electron neutrino oscillating into a muon neutrino . 16
1.4 The squared sine of twice the effective mixing angle in matter for various electron densities ............................................. 18
1.5 Feynman diagram for 0νββ .......................................................... 19
1.6 Monte Carlo simulation of the possible values of the effective neutrino mass for a given minimum neutrino mass ......................... 21
1.7 Energy schematic for most nuclei that are candidates for double-beta decay ................................................................. 22
1.8 Separation of 2νββ and 0νββ events with varying detector energy resolution and relative decay lifetimes ............................ 23
1.9 Cross section of the proposed NEXT-100 detector ......................... 30
1.10 Simulated two-electron track produced by a ββ event ....................... 31
1.11 The predicted sensitivity of several 0νββ experiments to mββ vs. total detector exposure time .................................................... 32
1.12 The bullet cluster ..................................................................... 35
1.13 The equilibrium density of a WIMP plotted as the temperature of the universe decreases relative to the WIMP mass, along with the freeze-out densities for different values of ⟨σAv⟩ ........................................... 41
1.14 The form factor for a Gaussian distribution of nuclear matter .......... 44
1.15 Fit of CMB temperature power spectrum to a ΛCDM model built on data from the Planck mission and other CMB experiments including the WMAP survey .............................................. 37
1.16 Calculation of nuclear recoil spectrum in a hypothetical detector ........ 45
1.17 The different detection strategies used in direct dark matter detection experiments and the associated experiments ..................... 47
1.18 The current status of dark matter claims and limits ........................ 47
2.1 Pressure vs. density for xenon gas at several different temperatures 50
2.2 Cross sections for electron interactions in pure xenon 51
2.3 The standard process responsible for scintillation in noble gases 52
2.4 The Lindhard factor $f_n$ as a function of nuclear recoil energy for xenon 54
2.5 The ratio of the observed ionization (S2) to primary scintillation (S1) signals in liquid xenon observed with a dual-phase xenon TPC plotted against recoil energy 55
2.6 Energy resolution vs. density for xenon gas 57
2.7 The fractional variance of single electron EL gain in the presence of electron attachment 64

3.1 Summary of the key components of NEXT-DBDM 67
3.2 General detection scheme in NEXT-DBDM 68
3.3 A typical event in NEXT-DBDM 69
3.4 The NEXT-DBDM TPC and gas circulation/purification system during operation 70
3.5 The gas system for NEXT-DBDM 71
3.6 Computer drawing of the pressure vessel and internal detector components 72
3.7 The 19-PMT array shown through two wire meshes through which electroluminescent photons must pass to be detected 73
3.8 The SiPM tracking plane placed behind the anode mesh in the TPC 74
3.9 The SiPM tracking plane electronics before (left) and after (right) installation in the TPC 74
3.10 Schematic of the strategy used to record SiPM data in NEXT-DBDM 75
3.11 A single EL electron emitting light instantaneously at a location $(x, y, z)$ in the EL gap 81
3.12 Initial samples of peaks in PMT-10 for reasonable cuts on charge 84
3.13 SPE charge histograms with final peak time at least 1 sample after the initial peak time and with final peak time at least 2 samples after the initial peak time 84
3.14 SPE charge histograms for PMTs 1-8 85
3.15 SPE charge histograms for PMTs 9-16 86
3.16 SPE charge histograms for PMTs 17-19 87
3.17 Demonstration of afterpulsing in NEXT-DBDM PMT signals 91
3.18 Relative total afterpulsing to signal ratio for PMT 1 92
3.19 Comparison of PMT signal values determined from the two integration methods 92
3.20 Several waveforms produced by gammas of energy 662 keV from $^{137}$Cs 96
3.21 S2 charge vs. event number for events near the 662 keV peak 97
3.22 Energy vs. drift time for events near the 662 keV peak 98
3.23 Spectra of S1 and S2 charges before any corrections due to geometrical dependencies ........................................... 98
3.24 Mathematical setup for calculation of photon flux through a disc of radius R when emitted at a distance z above the center of the disc . . 99
3.25 S1 charge vs. drift time ........................................... 100
3.26 Position dependence of S2 ........................................... 101
3.27 FWHM energy resolution of gamma peaks in NEXT-DBDM .......... 102
3.28 Expected and observed energy resolution in NEXT-DBDM under various configurations ........................................... 103
3.29 Schematic of the muon tracking test setup for NEXT-DBDM .......... 104
3.30 Waveforms showing signal measured by sum of PMT plane and sum of SiPM plane for a muon track ........................................... 105
3.31 Reconstructed muon track ........................................... 105
3.32 (x-y) projection of muon track showing relative SiPM intensities .... 106
3.33 PMT summed waveform for a gamma deposition with an x-ray .... 106
3.34 Reconstructed track and projections for a gamma event accompanied by an x-ray ........................................... 107
3.35 SiPM waveforms and 2D projection for the energy slice corresponding to the energy deposition of an x-ray accompanying a gamma event . . 108
4.1 Total cross section for $^9\text{Be}(\alpha,n)^{12}\text{C}$ reactions ................. 111
4.2 Center-of-mass frame angular distributions of emitted neutrons for several incident alpha energies with resulting $^{12}\text{C}$ nucleus in the ground, first excited, and second excited state .................... 112
4.3 Stopping power of alpha particles in Be in energy range of interest to the $^9\text{Be}(\alpha,n)^{12}\text{C}$ reaction ........................................... 114
4.4 The neutron spectra produced by alphas emitted in the $^9\text{Be}(\alpha,n)^{12}\text{C}$ reactions ........................................... 116
4.5 The total elastic scattering cross section for neutrons incident on several xenon isotopes ........................................... 117
4.6 Angular distributions for elastically scattered neutrons on xenon nuclei for several xenon isotopes ................................. 118
4.7 Spectra of nuclear recoils produced by neutrons emitted from a $^{238}\text{Pu}/\text{Be}$ neutron source ........................................... 120
4.8 NEXT-DBDM in its modified, shorter configuration with no buffer regions ........................................... 122
4.9 The S2 peak for 662 keV gamma rays, corrected with a central radial cut and an exponential attachment correction .................... 123
4.10 The S1 peak for 662 keV gamma rays, corrected for z-dependence as determined by a linear fit to the S1 data vs. drift time .......... 124
4.11 The S1 signals from several events in NEXT-DBDM in which the full-energy deposition of a 662 keV gamma ray was measured......... 125
4.12 The fraction of the total S1 signal present in the first 20 samples for full-energy S1 signals .......................................................... 125
4.13 The experimental setup for the acquisition of data with a $^{238}$Pu/Be source .......................................................... 126
4.14 Event rate vs. live time for the nuclear recoil characterization run with a $^{238}$Pu/Be source .......................................................... 127
4.15 Pressure and temperature vs. operation time for nuclear recoil data obtained with a $^{238}$Pu/Be source .......................................................... 128
4.16 Summary of trigger logic for neutron runs .......................................................... 129
4.17 Two candidate nuclear recoil events in NEXT-DBDM .......................................................... 130
4.18 Start times of NaI and S1 pulses .......................................................... 131
4.19 A candidate nuclear recoil event and its corresponding NaI signal .......................................................... 132
4.20 The $\sigma$ of a Gaussian fit to S2 pulses vs. mean drift time for events in the experimental data meeting the tagging and single-pulse cuts and in the corresponding Monte Carlo simulation .......................................................... 133
4.21 Distribution of reconstructed average event locations for single-pulse, tagged events passing diffusion cuts .......................................................... 134
4.22 The x and y projections of the reconstructed average event locations in data and Monte Carlo .......................................................... 135
4.23 The fiducial cut in data after position calibration and Monte Carlo .......................................................... 136
4.24 S2 vs. S1 for events at various stages of the analysis cuts .......................................................... 137
4.25 S1 and S2 experimental spectra with superimposed best-fit 1$\sigma$ uncertainties from Monte Carlo .......................................................... 139
4.26 Statistical analysis of the data/Monte Carlo fit procedure .......................................................... 140
4.27 The absolute S1 and S2 yields for nuclear recoils determined in this study assuming Lindhard energy dependence, compared to the yields in liquid xenon at various fields .......................................................... 141
4.28 A schematic of the experimental setup for data taken with a $^{241}$Am/Be neutron source .......................................................... 143
4.29 A photograph of the experimental setup for data taken with a $^{241}$Am/Be neutron source .......................................................... 144
4.30 Region of interest of the calculated nuclear recoil spectrum produced by neutrons emitted from a $^{241}$Am/Be source .......................................................... 144
4.31 Time-of-flight spectra describing cuts on neutron-induced events .......................................................... 145
4.32 S2 vs. S1 for low-energy events passing tagging, time-of-flight, diffusion, and radial cuts .......................................................... 146
4.33 S2 spectrum of nuclear recoil events for several different series of cuts .......................................................... 147
4.34 The logarithm of the ratio of S2/S1 plotted against nuclear recoil equivalent energy for events meeting all cuts, after calibration .......................................................... 149
4.35 The logarithm of the ratio of S2/S1 plotted against S1 in data and Monte Carlo for events meeting all cuts .......................................................... 150
4.36 The S2 plotted against S1 signals in data and Monte Carlo for events meeting all cuts ........................................... 150
4.37 S2 vs. S1 for low-energy events at varying drift field .................... 152
4.38 The ratio of 40 keV inelastic events to 30 keV x-ray events and the ratio of nuclear recoil events to 30 keV x-ray events plotted against drift electric field .................................................. 153

5.1 Diagram of a cross section of the negative ion TPC developed at Lawrence Livermore National Laboratory .................................... 155
5.2 Reduced mobility for several ion/neutral systems .......................... 159
5.3 Field-induced detachment lifetime vs. electric field for several ions calculated within the zero-range potential model ......................... 161
5.4 Detachment probability distributions for a constant field over a distance $d = 0.4$ mm at several different field strengths ..................... 168
5.5 Detachment probability distributions for a wire field configuration for wire radius $a = 50 \mu$m .................................................. 169
5.6 Detachment probability distributions for a LEM with pitch $a = 0.7$ mm, hole radius $r = 0.15$ mm, and thickness $d = 0.4$ mm .............. 170
5.7 Total detachment probability for an ion drifting through the entire field configuration of three different detector geometries ............ 171

A.1 Snapshot of the geometry setup in the GEANT4 Monte Carlo ........ 186
A.2 Illustration of energy resolution correction procedure for S1 ........ 188
A.3 Illustration of energy resolution correction procedure for S2 ........ 189

B.1 Electron mobility vs. density in xenon measured along the vapor-liquid coexistence curve ................................................. 191
B.2 W-value vs. density in xenon gas ........................................... 192
B.3 A schematic description of exciton energy levels .......................... 193
B.4 The number of atoms in a cluster of 10 or more vs. density in xenon obtained from a molecular dynamics simulation .................... 194
B.5 Pair distribution functions for atoms in xenon gas at several densities 195
B.6 Simulated $\delta$-ray distribution for 662 keV gamma events ............. 196
B.7 Energy resolution in liquid xenon with and without light and charge signal summation ....................................................... 199
B.8 Correlation between the gaseous xenon slow scintillation time and deduced probability of a xenon excitation to decay by means of a self-trapped exciton ....................................................... 201

C.1 The relative minimum and maximum electric field for electrode voltage of $V = 400$ V vs. source activity .................................. 207
C.2 The total space charge for electrode voltage of $V = 400$ V vs. source strength assuming an activity of 10 mCi produces $N = 150000$ ions/s in the active volume.
List of Tables

1.1 Key neutrino parameters .......................................................... 14
1.2 Results of the calculation of the number of $2\nu\beta\beta$ decays that will occur in the $0\nu\beta\beta$ peak for a 20-ton $^{136}\text{Xe}$ double-beta search experiment with varying energy resolution ...................................................... 26
1.3 Selected candidate $0\nu\beta\beta$ isotopes, their natural abundance, Q-values for double beta decay, and phase space factors ................. 26

2.1 Relative abundances (%) of isotopes in natural xenon ..................... 49

3.1 Mean SPE values for the 19 PMTs determined from the PMT calibration study ........................................................................... 87
3.2 Results for afterpulsing calculations using fixed-window and pulse-based analyses ................................................................. 94

4.1 $^{238}\text{Pu}$ $\alpha$-decay energies ....................................................... 110
4.2 Rates of neutron emission for a 11.6 mCi $^{238}\text{Pu/Be}$ source .......... 117
4.3 Error analysis for calculation of quenching factors and yields using the $^{241}\text{Am/Be}$ data ................................................................. 148
4.4 Comparison of nuclear recoil yields in the liquid and gas phases ... 148

5.1 Comparison of reduced mobilities calculated in the polarization limit to those measured in negative ion TPCs ................................. 158
5.2 Comparison of detachment factors across several ion/Ar and ion/Xe systems ................................................................. 171
Chapter 1

Rare Physics

The search for rare physics is an exercise in probability and statistics. If a number were randomly selected between one and one million, a single guess of that number is very unlikely to be correct. In fact, even repeating this select-and-guess process hundreds of times is unlikely to produce a correct guess in any single attempt. However, given one million attempts, the chance that one guess is correct would be significant. Rare physics can be defined as any physical process that occurs with an extremely low probability, and therefore may require many “attempts” to observe. Both neutrinoless double-beta decay and the scattering of a particle of dark matter off of a nucleus are examples of rare physics processes that could potentially be observed using a detector made of xenon gas. In both cases, the more xenon atoms present in a particle detector, the more attempts one has to observe the process. Here we outline the physics and scientific implications of both of these processes.

1.1 Neutrinos and $0\nu\beta\beta$

1.1.1 The Neutrino

Physicists first studying nuclear $\beta$-decay were confronted with an apparent violation of the conservation of energy. In $\beta$-decay, a neutron becomes a proton with the accompanied emission of an electron and, as we now know, a nearly massless neutral particle called the neutrino. Unable to detect the neutrino which could carry away a significant amount of the emitted energy in the decay, physicists were not able to explain the observed energy of emitted electrons. Neglecting the kinetic energy of recoil of the nucleus, the electron, if emitted alone, was expected to have a definite energy equal to the total energy released (the Q-value) of the decay. As it was found that the energy of the emitted electron was actually continuous, it was unclear where the rest of the energy was going until in 1930 Wolfgang Pauli proposed the neutrino as the means by which some amount of “disappearing” energy could be carried off undetected [1].
Since then our knowledge of the neutrino has improved significantly. Enrico Fermi was able to describe $\beta$-decay theoretically as a 4-fermion interaction with a coupling strength $G$, such that the rate of the process is $\propto G^2$. In 1962, it was discovered that certain neutrinos, now called muon neutrinos $\nu_\mu$, only produce muons in the products of their interactions [2], confirming the existence of different types (or “flavors”) of neutrinos. We now know of three neutrino flavors; the electron neutrino ($\nu_e$), the muon neutrino ($\nu_\mu$), and the tau neutrino ($\nu_\tau$). They are organized into groups, each with an associated lepton number ($n_e, n_\mu, n_\tau$),

$$\begin{pmatrix} e^- \\ \nu_e \end{pmatrix}, \quad \begin{pmatrix} \mu^- \\ \nu_\mu \end{pmatrix}, \quad \begin{pmatrix} \tau^- \\ \nu_\tau \end{pmatrix}$$

(1.1)

where in each case the negatively charged lepton and neutrino each carry lepton number $+1$ and the positively charged lepton and the antineutrino carry lepton number $-1$. Each individual lepton number $n_l$ is a flavor lepton number, while the total lepton number is defined as

$$L = \sum_{i=e,\mu,\tau} n_i.$$  

(1.2)

It was believed that both flavor and total lepton number are conserved quantities in all physical interactions until neutrinos were found to have mass. For massive neutrinos, each neutrino flavor is actually a superposition of neutrino mass states, and therefore neutrinos propagating through space can “oscillate” between flavors - a process in which flavor lepton number is not conserved. Despite these recent findings, questions remain concerning the physics that gives rise to the neutrino mass, and still total lepton number has not been shown to be violated in any physical process. Both of these unknowns are addressed in the search for $0\nu\beta\beta$ decay.

1.1.2 Fermions and Helicity

The neutrino is a fermion, that is its quantum mechanical spin is half-integer. More specifically it is a spin-1/2 fermion, and such particles are described mathematically as 4-component objects $\psi$ that in the absence of interactions obey the Dirac equation (see for example [3])

$$(i\gamma^\mu \partial_\mu - m)\psi = 0,$$

(1.3)

where $\partial_\mu = (\partial/\partial t, \nabla)$ and we are using the Einstein summation notation with metric $g_{00} = +1, g_{ii} = -1$, and $g_{\mu\nu,\mu\neq\nu} = 0$. We will use the chiral representation for $\gamma^\mu$, 

$$2$$
given by

\[
\gamma^0 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}, \quad \gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} -I & 0 \\ 0 & I \end{pmatrix}
\]

(1.4)

where \( I \) is the 2x2 identity and \( \sigma_i \) is the \( i \)th Pauli spin matrix. The 4-component object \( \psi \) is actually composed of two spinors of two components each, and we call these the “left-chirality” \( \psi_L \) and “right-chirality” \( \psi_R \) spinors,

\[
\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}.
\]

(1.5)

Labeling a spinor with L or R will therefore means that when written in 4-component form it contains only an upper (L) or lower (R) component, that is, we define \( \psi_L \equiv \frac{1}{2}(1 - \gamma^5)\psi \) and \( \psi_R \equiv \frac{1}{2}(1 + \gamma^5)\psi \).

Chirality is related to the orientation of a fermion’s spin along the axis of its linear momentum. A fermion is said to be “left-handed” if its spin points anti-parallel to its direction of travel and “right-handed” if its spin points parallel to its direction of travel. For example, in our chiral representation, we have two possible solutions to the Dirac equation for states with momentum along the +z axis with spin up (\( \uparrow \)) and down (\( \downarrow \)) [3]

\[
\psi_{\uparrow}(x, p) = \frac{1}{\sqrt{2E}} \begin{pmatrix} \sqrt{E - p} & 1 \\ \sqrt{E + p} & 0 \end{pmatrix} e^{-p \cdot x},
\]

\[
\psi_{\downarrow}(x, p) = \frac{1}{\sqrt{2E}} \begin{pmatrix} \sqrt{E + p} & 1 \\ \sqrt{E - p} & 0 \end{pmatrix} e^{-p \cdot x},
\]

(1.6)

where \( E = \sqrt{|p|^2 + m^2} \) is the fermion energy, \( p^\mu = (E, \mathbf{p}) \), \( x^\mu = (t, \mathbf{x}) \), and \( p \cdot x = p_\mu x^\mu \).

We are interested in the projection of the fermion’s spin along its direction of propagation, a quantity called helicity, defined [3] as the eigenvalue of the operator \( \hat{h} = \hat{p} \cdot \hat{\Sigma} \), where \( \hat{\Sigma} = \sigma I \) and \( I \) is the 4x4 identity matrix and \( \sigma = \sigma_x \hat{x} + \sigma_y \hat{y} + \sigma_z \hat{z} \). Both of the states in equation 1.6 are eigenstates of \( \hat{h} \), but by a Lorentz transformation, one can boost into a frame in which these fermions have different helicities, for example a frame moving in the direction of the fermion but faster than the fermion itself. This is not possible if the fermion is massless and moving at the speed of light, and in fact setting \( E = p \) in equation 1.6, one can see that the states \( \psi_{\uparrow} \) and \( \psi_{\downarrow} \) reduce to a single component. The right-handed (positive helicity) state \( \psi_{\uparrow} \) contains
only a right-chirality component $\psi_R$, and the left-handed (negative helicity) state $\psi_L$ contains only a left-chirality component $\psi_L$.\(^1\) In the high energy limit, chirality and helicity coincide, and therefore the terms chirality and helicity are often used interchangeably (for example $\psi_L$ may be called a left-handed spinor rather than a left-chirality spinor). It is important to note however that a massive fermion can be left-handed (have negative helicity) but still have some nonzero right-chirality component.

Chirality is important in describing fermion interactions. In quantum field theory, the interactions of fermions can be described by products of currents with vector gauge bosons such as the photon ($A^\mu$). One example of such a product can be found in quantum electrodynamics,

\[
\text{(coupling)} \times \text{(current)} \times \text{(gauge boson)} = e \times \bar{\psi} \gamma^\mu \psi \times A_\mu = e \bar{\psi} \gamma^\mu \psi A_\mu. \tag{1.7}
\]

where $\bar{\psi} = \psi^\dagger \gamma^0$. The form of the current dictates which particles participate in the interaction. Neutrinos are electrically neutral fermions and the gravitational interactions between neutrinos are negligible for our purposes here. Therefore we will be most concerned with the weak interaction, which mediates all processes of interest involving neutrinos, including $\beta$-decay. In the accepted model of weak interactions (the Glashow-Salam-Weinberg or GSW model), particles interact by exchanging massive gauge bosons called $W$-bosons $W^\pm$ and $Z$-bosons $Z^0$. The currents that couple to the $W$-bosons are called charged currents and those that couple to the $Z$-bosons are called neutral currents. In charged weak interactions, the current $J^\mu_W$ takes the form

\[
J^\mu_W \propto \bar{\psi}_a \gamma^\mu \frac{1}{2} (1 - \gamma^5) \psi_b, \tag{1.8}
\]

where $\psi_a$ and $\psi_b$ are fermions whose absolute difference in electric charge is 1. And in fact using $\frac{1}{2} (1 - \gamma^5)^2 = \frac{1}{2} (1 - \gamma^5)$ and $\gamma^\mu \frac{1}{2} (1 - \gamma^5) = \frac{1}{2} (1 + \gamma^5) \gamma^\mu$, we can write this as

\[
J^\mu_W \propto \bar{\psi}_{a,L} \gamma^\mu \psi_{b,L}, \tag{1.9}
\]

where we have used $\bar{\psi}_L = \frac{1}{2} (1 - \gamma^5) \bar{\psi} = \bar{\psi}_L \frac{1}{2} (1 + \gamma^5)$. This form of the current demonstrates several important aspects of weak interactions.

First, unlike in electromagnetic interactions, weak interactions do not conserve parity, that is they are not left unaltered by spatial inversion. This is because $\bar{\psi}_a \gamma^\mu \psi_b$ (called the “vector current”) is unaltered by a parity transformation while $\bar{\psi}_a \gamma^\mu \gamma^5 \psi_b$ (called the “axial current”) is negated \(^2\). In 1956, C. S. Wu et. al. \(^5\) confirmed

\(^1\)Note that for antifermions, the solutions to the Dirac equation are different and this relationship is reversed; a left-handed massless antifermion would contain only a right-chirality component.

\(^2\)Because of the factor $(a - b \gamma^5)$ with $a = b = 1$, both $\bar{\psi}_a \gamma^\mu \psi_b$ and $\bar{\psi}_a \gamma^\mu \gamma^5 \psi_b$ are present in the charged weak current with equal weighting, and it is for this reason that charged weak interactions are referred to as “maximally parity violating.”
experimentally that nuclear $\beta$-decay is parity violating by showing that the emission rate of the electron depended on its angle with the nuclear spin vector of the parent nucleus. Such a $\sigma \cdot \mathbf{p}$ interaction term violates parity, as the spin does not change sign under a parity transformation, while the vector $\mathbf{p}$ does.

Second, weak interactions conserve chirality. This can be seen by forming the product, using equation 1.4,

$$
\bar{\psi} \gamma^\mu \psi = \psi^\dagger \gamma^0 \gamma^\mu \psi
= (\psi_L^\dagger \psi_R^\dagger) \left( \begin{array}{cc} 0 & I \\ I & 0 \end{array} \right) \left( \begin{array}{c} 0 \\ \sigma^\mu \\ 0 \end{array} \right) \left( \begin{array}{c} \psi_L \\ \psi_R \end{array} \right)
= \psi_L^\dagger \sigma^\mu \psi_L + \psi_R^\dagger \sigma^\mu \psi_R,
$$

where here we have used the notation $[3] \sigma^\mu = (1, \sigma)$, $\sigma = (1, -\sigma)$. Here only left chirality and right chirality states are "coupled," and so interactions can only occur between same-chirality states. One can make a similar argument for the term $\bar{\psi} \gamma^\mu \gamma^5 \psi$.

This means that as $m \to 0$ where states of definite chirality become states of definite helicity, weak interactions conserve helicity in the massless limit.

Finally, charged weak interactions only involve left-chirality fermion (right-chirality antifermion) components. Neutral weak interactions, on the other hand, take the form

$$
J_Z^\mu \propto \bar{\psi} \gamma^\mu \frac{1}{2}(c_V^f - c_A^f \gamma^5)\psi,
$$

where $c_V^f$ and $c_A^f$ are called the vector and axial couplings for fermion $f$. It is at this point of the story that neutrinos begin to stand out if they have not done so already. In the standard model, $c_V^\nu = c_A^\nu = 1$ for neutrinos, so that even neutral weak currents are maximally parity-violating. This reflects important experimental observations that began in the 1950s [1] when Frauenfelder et al. [6] found that the electrons emitted in $\beta$-decay were predominantly left-handed, meaning that the neutrino had a single handedness, though which still could still have been left or right. Soon afterward Goldhaber et al. [7] determined that this handedness is in fact left, and this conclusion has held since then - right-handed neutrinos have never been detected.

If neutrinos were massless, one could conclude that only left-handed (and therefore left-chirality) neutrinos exist. But, again, for a massive particle, we can always Lorentz boost into a frame moving faster than it and observe the particle with its momentum flipped but its spin in the original direction, and thus with opposite helicity. As neutrinos are now known to have mass, right-handed neutrinos must, in principle, exist. Since this mass is very small, however, the left-chirality component in such a right-handed state would be very small, and therefore few right-handed neutrinos would be produced in a weak interaction that involves only left-chirality states.

Helicity may have significance beyond the above discussion of chirality. It may be that the neutrino and the antineutrino are actually the same particle, but our
tendency to distinguish them in experimental observations is a result of the strong connection between helicity and chirality in the massless limit [8]. For example, in nuclear beta ($\beta^-$) decay, a charged weak process,

$$n \rightarrow p + e^- + \nu_e,$$

we must have a left-chirality electron and a right-chirality antineutrino due to equation 1.8. If the neutrino is equivalent to the antineutrino, that is both participate in all interactions in the same way [8], the right-chirality state we call the antineutrino may be the right-chirality state of the neutrino itself. In $\beta^+$ decay

$$p \rightarrow n + e^+ + \nu_e.$$  

Due to lepton number conservation ($L = 0$ before and after the decay), we required the production of the antineutrino ($n_e = -1$) with the electron ($n_e = 1$), and the neutrino ($n_e = 1$) with the positron ($n_e = -1$), but the only way we have thus far distinguished between neutrinos and antineutrinos experimentally is through their helicity and stating that only left-chirality neutrinos and right-chirality antineutrinos interact. There is nothing preventing the neutrino from being equivalent to its antiparticle and having two interacting chirality states (see figure 1.1). If the neutrino is equivalent to its antiparticle, it is said to be a “Majorana” fermion, as opposed to a “Dirac” fermion with distinct particle and antiparticle states. From this point of view, there is no conserved lepton number - the concept of lepton number could be a reflection of helicity near-conservation in the nearly massless limit rather than a physically conserved quantity.

**1.1.3 Majorana Neutrinos**

The Italian physicist Ettore Majorana proposed the theory of Majorana fermions in [9], published in 1937. Shortly thereafter, in March of 1938, he disappeared [10].

We have said that a Majorana neutrino is a particle that interacts in exactly the same way as its antiparticle, and therefore the two are physically the same particle. Formally, a Majorana fermion transforms into itself under the operation of charge conjugation. This means that if the neutrino field is represented by $\nu(x)$, then the charge conjugation operation

$$[\nu(x)]^c = C \bar{\nu}(x)^T,$$  

does not physically alter the field, that is $[\nu(x)]^c = e^{i\phi} \nu(x)$ where $\phi$ is some real phase. Here $C$ is the charge conjugation matrix which in the chiral representation (eqn. 1.4) is $C = i\gamma^0\gamma^2$. The alternative to a Majorana fermion is called a Dirac fermion, one that is distinct from its antiparticle.$^3$

$^3$The Dirac neutrino field, under the same charge conjugation operation, is altered by more than
Figure 1.1: The possible configurations of neutrino helicity eigenstates for Dirac and Majorana neutrinos. In the massless case, eigenstates of helicity are purely left or right chirality, while neutrinos with small but nonzero mass contain some contribution from each chirality. If the neutrino is Dirac, the crossed-out states have not yet been observed. If the neutrino is Majorana, the observed right-handed neutrino state, which we have been calling the anti-neutrino, is really just the right-handed Majorana neutrino state.

Despite this important difference between Dirac and Majorana fermions, it has not yet been discovered which description applies to the neutrino. It is evident that charged particles such as quarks and the charged leptons are Dirac fermions, as their antiparticles have the opposite electric charge and therefore interact differently and are distinguishable. However, because the neutrino is electrically neutral, one cannot immediately make a similar conclusion. In fact, as described in section 1.1.2, in charged weak interactions one cannot distinguish between the existence of two helicity states of a Majorana neutrino, or one left-chirality Dirac neutrino and its right-chirality antineutrino.\(^4\) Perhaps the most significant candidate for an experiment that could answer this important question is the detection of neutrinoless double-beta decay.

\(^4\)It can be shown that in electromagnetic interactions [11] (despite being neutral, a neutrino can still couple to the photon [8]) and neutral weak interactions [12], Dirac and Majorana neutrinos are indistinguishable in the limit of zero neutrino mass. This concept, called the “practical Majorana-Dirac confusion theorem” [11] makes most ordinary experimentally observable processes involving neutrinos incapable of revealing their Dirac or Majorana nature.
1.1.4 The Neutrino Mass

We have discussed several implications of a nonzero neutrino mass. Massive neutrinos may be identified as Majorana particles, must exist with both positive and negative helicities, and as described further in section 1.1.5, oscillate between flavors. In fact, such neutrino oscillations have been observed, and so neutrinos must be massive. However, measurements of the oscillations alone can determine only differences in neutrino masses, and experiments aiming to measure absolute neutrino masses have so far only managed to set upper limits. Such attempts include measurements of the spectrum of energies of electrons emitted in $\beta$-decay, which will differ slightly near its endpoint between the massive and massless neutrino cases, and cosmological measurements (see for example [13]).

In quantum field theory, the mass of a fermion is expressed as a mass term in a Lagrangian. For example, the Dirac free field Lagrangian is given by

$$L = \bar{\nu}(i\gamma^\mu - m)\nu.$$  \hspace{1cm} (1.14)

As in section 1.1.2, the objects $\nu$ and $\bar{\nu} = \nu^\dagger\gamma^0$ are field operators consisting of linear superpositions of solutions to equation 1.3. The “mass term” for the field $\nu$ is $-\bar{\nu}\nu m$.

We now decompose $\nu$ into its two components $\nu_L$ and $\nu_R$, as we did in section 1.1.2

$$\nu_L = \frac{1}{2}(1 - \gamma^5)\nu, \quad \text{and} \quad \gamma^5\nu_L = -\nu_L,$$

$$\nu_R = \frac{1}{2}(1 + \gamma^5)\nu, \quad \text{and} \quad \gamma^5\nu_R = \nu_R.$$  \hspace{1cm} (1.15)

We will annotate two-component spinors as $\psi$ and four-component fields as $\nu$. We can express a neutrino field as $\nu = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}$, so that $\nu_L = \begin{pmatrix} \psi_L \\ 0 \end{pmatrix}$ and $\nu_R = \begin{pmatrix} 0 \\ \psi_R \end{pmatrix}$.

Then the mass term can be written as

$$-\bar{\nu}m\nu = -\nu^\dagger\gamma^0m\nu = -\begin{pmatrix} \psi_L^\dagger & \psi_R^\dagger \end{pmatrix} \begin{pmatrix} 0 & m \\ m & 0 \end{pmatrix} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}$$

$$= -[\psi_R^\dagger m\psi_L + \psi_L^\dagger m\psi_R].$$  \hspace{1cm} (1.16)

We see that a fermion mass term contains a combination of right and left-handed fields.\(^5\)

\(^5\) Note that in the GSW model, such mass terms do not come from free field components of the Lagrangian as shown in equation 1.14 but appear via the Higgs mechanism (see for example [4]); due to the extreme smallness of the neutrino mass, it is speculated that the Higgs mechanism is unlikely to also be the process behind generation of the neutrino mass [10].
We will still consider the possibility of neutrino mass terms similar to those of equation 1.16, but we will also attempt to construct additional mass terms. Since the neutrino may be a Majorana particle, it would be interesting to create mass terms in which the antineutrino and neutrino are both present. From [3], the charge conjugation operation acts on a fermion to give its antifermion without changing its spin projection. For a fermion field $\nu$, we have from equation 1.13 that the operation of charge conjugation $\nu \rightarrow \nu^c$ gives

$$ (\nu_{L,R})^c = C \nu_{L,R}^T. $$

where $C$ is the charge conjugation operator which in the chiral representation (eqn. 1.4) is $C = i\gamma^0 \gamma^2$. Note that

$$ (\nu_L)^c = C \nu_L^T = (i\gamma^0 \gamma^2) \begin{pmatrix} 0 & \psi_L^1 \\ -i\sigma_2 \psi_L^* \end{pmatrix}. $$

That is, the conjugated left-handed component of $\nu$ contains only a right-handed component, so it is effectively “right-handed,” and a similar argument holds for $(\nu_R)^c$. We will therefore attempt to construct mass terms coupling $(\nu_L)^c$ with $\nu_L$, and $(\nu_R)^c$ with $\nu_R$. That is, we will write following [10]

$$ L_{mass} = -\frac{1}{2} [\overline{\nu}_R m\nu_L + \overline{\nu}_L m\nu_R] - \frac{1}{2} [\overline{\nu}_L m_L (\nu_L)^c + (\nu_L)^c m_L \nu_L] $$

$$ - \frac{1}{2} [\overline{\nu}_R m_R (\nu_R)^c + (\nu_R)^c m_R \nu_R]. $$

(1.19)

Note that we have used $\overline{\nu}_R m\nu_L = (\overline{\nu}_L)^c m(\nu_R)^c$ [14] to combine the first and third terms. The first group of terms with mass $m$ makes up the original mass term, which will be called the Dirac mass term and we will denote $m = m_D$, and the other two groups with masses $m_L$ and $m_R$ will be called Majorana mass terms.

At this point, we will comment on some technical aspects of the notation. We may also write the individual field projections, for example $\nu_L$, as components in a larger column vector to rewrite $L_{mass}$ in a more compact form. In this case, one can still consider these projections as two-spinor objects, for example $\nu_L = \begin{pmatrix} \psi_L \\ 0 \end{pmatrix}$. The additional matrix notation is used for organizational purposes to rewrite the mass term in terms of fields with definite mass. It lies outside the usual fermion matrix manipulations, so for example the $\gamma$ matrices do not act on these larger matrices.
Following [10], we define an object
\[
v = \begin{pmatrix} \nu_L \\ (\nu_R)^c \end{pmatrix} = \begin{pmatrix} \psi_L \\ 0 \\ -i\sigma^2\psi_R^* \\ 0 \end{pmatrix}
\]
(1.20)
\[
\bar{v} = \begin{pmatrix} \nu_L^* \\ (\nu_R)^c \end{pmatrix} = \begin{pmatrix} 0 \\ \psi_L^T \\ 0 \\ i\sigma^2\psi_R^T \end{pmatrix}.
\]

Now the mass terms can be written as [10]
\[
\mathcal{L}_{\text{mass}} = -\frac{1}{2} \bar{v} \begin{pmatrix} m_L & m_D \\ m_D & m_R \end{pmatrix} v^c - \frac{1}{2} v^c \begin{pmatrix} m_L & m_D \\ m_D & m_R \end{pmatrix} v.
\]
(1.21)

We wish to diagonalize the mass matrix in equation 1.21, so that we can write the mass terms in \( \mathcal{L} \) using states of definite mass, i.e., in the form \(-m\nu\nu'\). The eigenvalues of this matrix are
\[
m_\pm = \frac{m_L + m_R}{2} \pm \sqrt{\left(\frac{m_L - m_R}{2}\right)^2 + m_D^2}.
\]
(1.22)

Note that if the term under the square root is larger than \((m_L + m_R)/2\), \(m_-\) will be negative. We assume this is the case and diagonalize the matrix such that the mass eigenvalue in \( \mathcal{L} \) will be \(-m_-\) [10],
\[
\begin{pmatrix} m_L & m_D \\ m_D & m_R \end{pmatrix} = U M U^T
\]
(1.23)
with
\[
\tan 2\theta = \frac{2m_D}{m_R - m_L}, \quad \cos 2\theta = \frac{m_L - m_R}{\sqrt{(m_L - m_R)^2 + 4m_D^2}}.
\]
(1.24)

Now we can write the entire \( \mathcal{L}_{\text{mass}} \) as
\[
\mathcal{L}_{\text{mass}} = -\frac{1}{2} (\bar{v}' M (v')^c + (v')^c M v')
\]
(1.25)
where
\[
v' = U^\dagger v = \begin{pmatrix} \nu_L \cos \theta - (\nu_R)^c \sin \theta \\ -i(\nu_L \sin \theta + (\nu_R)^c \cos \theta) \end{pmatrix}.
\]
(1.26)
Finally, we can add $v'$ and its charge conjugate $(v')^c$ to give a vector of mass eigenstate fields $v_m = v' + (v')^c = \begin{pmatrix} v_{m+} \\ v_{m-} \end{pmatrix}$, where

$$v_{m+} = (\nu_L + (\nu_L)^c) \cos \theta - (\nu_R + (\nu_R)^c) \sin \theta = \begin{pmatrix} \psi_L \\ i\sigma^2\psi_L^* \end{pmatrix} \cos \theta - \begin{pmatrix} -i\sigma^2\psi_R^* \\ \psi_R \end{pmatrix} \sin \theta,$$

$$v_{m-} = - (\nu_L + (\nu_L)^c) i \sin \theta - (\nu_R + (\nu_R)^c) i \cos \theta = - \begin{pmatrix} \psi_L \\ i\sigma^2\psi_L^* \end{pmatrix} i \sin \theta - \begin{pmatrix} -i\sigma^2\psi_R^* \\ \psi_R \end{pmatrix} i \cos \theta. \tag{1.27}$$

One can write the mass terms in terms of the mass eigenstates as

$$\mathcal{L}_{mass} = -\frac{1}{2} \overline{v_m} M v_m = -\frac{1}{2} (\overline{v_{m+}} v_{m+} m_+ + \overline{v_{m-}} v_{m-} (-m_-)), \tag{1.28}$$

noting that terms $\overline{v'} v'$ and $(v')^c (v')^c$ will be zero just as for example $\overline{\nu_L} \nu_L = 0$. We are left with 2 fields and can now abandon our extra matrix notation; $\nu_{m+}$ and $\nu_{m-}$ consist of two spinor components, and we note that we have constructed them such that $\nu_{m_i} = (\nu_{m_i})^c$, that is, these definite mass states are Majorana particles. This formalism can be extended to the case where there exists more than one neutrino flavor (see for example [10]).

To avoid potential confusion, note that if $m_L = m_R = 0$, this formalism still gives two mass eigenvalues, but they are degenerate and equal to the Dirac mass; $m_+ = -m_- = m_D$. In this case the left- and right-chirality states form the two spinors of a 4-component Dirac fermion. If only one of the two chirality states actually exists (for example, if $\psi_R = 0$), one cannot have a Dirac mass term, but one can still have a Majorana mass term (mass $m_L$), and in this case the corresponding neutrino is not a Dirac fermion - it is fundamentally different. One may argue that by construction we have written $v_{m+}$ and $v_{m-}$ to obey the Majorana condition $v_{m+} = (v_{m+})^c$, so even if we have only a Dirac mass we can still write the mass term in terms of Majorana fields, and this is correct. However, we are saying that in the special case for which the two masses are degenerate, we can also write the same mass term as $\overline{\nu} m_D \nu$, where $\nu = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}$ is a Dirac fermion which does not have to obey the Majorana condition.

In this sense the Dirac fermion is a special case of the Majorana formalism developed. If one or more Majorana masses is nonzero, we cannot assign the two neutrino states as the left- and right-chirality states of a single particle with one single mass; the masses are no longer degenerate, and we have two states with two different physically observable masses.

In this scheme, a Majorana mass offers an explanation for why the neutrino mass is so small via the “see-saw mechanism”\(^6\) for neutrino mass generation (see for example

\(^6\)More specifically, here we describe the “type I” see-saw mechanism.
[10, [14]), in which for each flavor of neutrino, two Majorana neutrino states exist, one with a very small mass, and one with a very large mass. Assume \( m_L = 0 \), \( m_D \) is of order the mass of one of the three charged leptons (\( e^\pm, \mu^\pm, \text{or } \tau^\pm \)), and \( m_R \gg m_D \). Then we have two Majorana neutrinos, with masses, from equation 1.22,

\[
m_+ \approx \frac{m_R}{2} \quad \text{and} \quad -m_- = -\frac{m_R}{2} \left( 1 - \sqrt{1 + 2 \frac{m^2_D}{m^2_R}} \right) \approx m_D \left( \frac{m_D}{m_R} \right).
\]  

(1.29)

One neutrino has the very large mass \( m_R \), and the other a mass much smaller than the ordinary charged lepton masses (see figure 1.2). If the neutrino is a Majorana particle, the see-saw mechanism could explain why the observed neutrinos have such small masses. This is another reason why the investigation of the Majorana nature of the neutrino is important.

![Figure 1.2: Illustration of the see-saw mechanism. As the Majorana mass \( m_R \) becomes larger relative to the Dirac mass \( m_D \), the separation between the value of the two physical masses \( m_+ \) and \( m_- \) becomes larger; \( m_+ \) becomes larger relative to the Dirac mass and \( m_- \) becomes relatively smaller.](image)

1.1.5 Neutrino Mixing and Oscillations

In the 1960s, the first solar neutrino measurement [15] was made by capturing solar neutrinos (\( \nu_e \)) on \( ^{37}\text{Cl} \) nuclei and extracting and counting the radioactive \( ^{37}\text{Ar} \) atoms produced in the resulting reaction. When the measured flux of neutrinos was found
to be much smaller than that predicted by the standard solar model (SSM), it was
unclear whether the apparent lack of neutrinos was due to experimental error, serious
errors in the SSM, or some physical reason such as neutrino oscillations. Over the
years, similar observations of $\nu_e$ solar neutrino fluxes falling well below SSM predic-
tions were observed by other experiments (see [16] for example). It was not until the
SNO experiment [17], which was able to measure interactions of all three neutrino
flavors, that the total measured neutrino flux was found to be in agreement with the
SSM, but that roughly 65-70% of the neutrinos arrived as $\nu_\mu$ or $\nu_\tau$. Some of these
solar neutrinos, originally all produced with $\nu_e$ flavor, had oscillated to a different
flavor.

In section 1.1.4, we considered a general neutrino mass term constructed from
right-handed and left-handed neutrinos along with their antiparticles and found two
eigenstates of definite mass. Now we restrict the mass term to only left-handed
neutrinos, yielding only one mass eigenstate per neutrino, but we will assume there
are three flavors of neutrinos,

$$\nu_l = \begin{pmatrix} \nu_{e,L} \\ \nu_{\mu,L} \\ \nu_{\tau,L} \end{pmatrix},$$  \hspace{1cm} (1.30)

where $e$, $\mu$, and $\tau$ represent the three flavors of electron, muon, and tau. These three
flavor states are not assumed to be also mass eigenstates, rather the mass eigenstates
are linear superpositions of the flavor eigenstates,

$$\nu_l = \sum_i U_{li} (\nu_m)_i,$$  \hspace{1cm} (1.31)

where $U$ is the 3x3 neutrino mixing matrix [16],

$$U = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13}e^{i\alpha_{21}/2} & s_{13}e^{i(\alpha_{31}/2-\delta)} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & [c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta}]e^{i\alpha_{21}/2} & s_{23}c_{13}e^{i(\alpha_{31}/2-\delta)} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & [-c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta}]e^{i\alpha_{21}/2} & c_{23}c_{13}e^{i(\alpha_{31}/2-\delta)} \end{pmatrix}. $$  \hspace{1cm} (1.32)

If the states $(\nu_m)_i$ are Dirac, the phases $\alpha_1 = 0$ and $\alpha_2 = 0$. Note that as opposed
to the mixing in section 1.1.4, we are now mixing several left-chirality flavor neutrino
states rather than several left- and right-chirality states of the same flavor and their
charge conjugates. The neutrino in this case could still be Majorana or Dirac, but we
are not considering the mixing between charge conjugate states (or our matrix would
be 6x6, see [10]).

Because neutrinos participate in weak interactions as flavor eigenstates but travel
through space as mass eigenstates, neutrinos can originate as one flavor and “oscillate”
into another flavor during the course of their travel. This is possible because [8]:

1. Neutrinos have mass.
Table 1.1: Key neutrino parameters, from [16].

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sin^2 \theta_{12}$</td>
<td>$0.857 \pm 0.024$</td>
</tr>
<tr>
<td>$\sin^2 \theta_{23}$</td>
<td>$&gt; 0.95$</td>
</tr>
<tr>
<td>$\sin^2 \theta_{13}$</td>
<td>$0.095 \pm 0.010$</td>
</tr>
<tr>
<td>$\Delta m_{21}^2$</td>
<td>$(7.50 \pm 0.20) \times 10^{-5}$ eV$^2$</td>
</tr>
<tr>
<td>$\Delta m_{32}^2$</td>
<td>$0.00232^{+0.00012}_{-0.00008}$ eV$^2$</td>
</tr>
</tbody>
</table>

2. The mixing matrix $U$ has nonzero off-diagonal elements.

We follow [8] for this discussion and consider a neutrino emitted at $t = 0$ and propagating through space in a state of definite energy $E_\nu$. The neutrino propagates through a distance $x$, and we apply a translation $\exp(i \hat{p}_x x)$

$$\nu_l(x) = e^{i \hat{p}_x x} \nu_l = \sum_i U_{li} e^{i p_i x} (\nu_m)_i$$  \hspace{1cm} (1.33)

where we have assumed that since the different mass eigenstates all must have the same energy, they must have different momenta $p_i$. Now obtaining the time dependence with $\exp(-i \hat{H} t) \nu_l = \exp(-i E_\nu t) \nu_l$ for a state of definite energy, we can write

$$\nu_l(x, t) = \sum_i U_{li} e^{i (p_i x - E_\nu t)} (\nu_m)_i.$$  \hspace{1cm} (1.34)

Evaluating the exponent at $t = x$, that is assuming the neutrino moves at the speed of light, the exponent becomes

$$i(p_m - E_\nu)x = i(\sqrt{E_\nu^2 - m_i^2} - E_\nu)x \approx i(E_\nu - m_i^2/2E_\nu^2 - E_\nu)x = -im_i^2 x/2E_\nu^2.$$  \hspace{1cm} (1.35)

Using the unitarity of $U$, that is $\sum_i U_{li} U_{l'i}^* = \delta_{li}$, we can write $(\nu_m)_i = \sum_{l'} U_{l'i}^* \nu_{l'}$ and rewrite equation 1.34 as

$$\nu_l(x, t) = \sum_{l'} \left( \sum_i U_{li} U_{l'i}^* e^{-im_i^2 x/E_\nu^2} \right) \nu_{l'}.$$  \hspace{1cm} (1.36)

Therefore the probability of finding the neutrino in a flavor state $\nu_{l'}$ after traveling a distance $x$ is

$$P(\nu_l \rightarrow \nu_{l'}) = \left| \sum_i U_{li} U_{l'i}^* e^{-im_i^2 x/E_\nu^2} \right|^2.$$  \hspace{1cm} (1.37)
It can be shown that if we assume that CP is conserved in interactions among leptons, all components of $U$ can be taken to be real [14], and in this case

$$P(\nu_l \rightarrow \nu_{l'}) = \sum_i U_{li}^2 + \sum_{i \neq j} U_{li} U_{lj} U_{l'i} U_{l'j} \cos(2\pi x / L_{ij}),$$  \hspace{1cm} (1.38)$$

where $L_{ij}$ is called the oscillation length,

$$L_{ij} = L_{ji} = \frac{2\pi}{\Delta m^2_{ij}}$$  \hspace{1cm} (1.39)$$

and $\Delta m^2_{ij} = |m^2_i - m^2_j|$ is the difference in the squares of the masses of eigenstates $\nu^m_i$ and $\nu^m_j$. Equation 1.38 shows that a neutrino originating as a specific flavor $l$ and propagating through space will have an oscillating probability of being observed as a neutrino of a different flavor $l'$. The oscillations will be observable in an experiment observing neutrinos traveling distances of order the values of the $L_{ij}$.

Considering the simplest case of mixing between only two neutrinos $\nu_a$ and $\nu_b$ requires only 1 mixing angle $\theta$ so that

$$U = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}.$$  \hspace{1cm} (1.40)$$

Then the probability of finding a neutrino $a$ to have oscillated to flavor $b$ through the propagation of mass eigenstates 1 and 2 through a distance $x$ is

$$P(a \rightarrow b) = \sin^2 \theta \sin^2(\pi x / L_{12}).$$  \hspace{1cm} (1.41)$$

The above discussion applies to neutrino oscillations in vacuum, while for neutrinos moving through matter, neutrino-electron scattering processes involving exchange of $W^\pm$ and $Z^0$ bosons can affect the oscillation probabilities through the Mikheyev-Smirnov-Wolfenstein (MSW) effect [18]. For example, all neutrinos produced within sun are subject to scattering by processes involving $Z^0$-exchange as they move outward through the solar matter, but only $\nu_e$ neutrinos can interact via $W^-$ exchange because $e^-$ are readily present in the sun while the corresponding charged leptons for other neutrinos are not present. This can be described by subjecting the $\nu_e$ component of each propagating mass eigenstate to an additional potential energy $V_e = \sqrt{2}G_F N_e$, where $G_F \approx 1 \times 10^{-5}$GeV$^{-2}$ is the Fermi constant and $N_e$ is the number density of electrons [19, 20]. To see how this affects the oscillation probability, we first write down the Hamiltonian for a propagating neutrino in the flavor basis [16]. In the basis of mass eigenstates, we have

$$H_0 = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix},$$  \hspace{1cm} (1.42)$$
Figure 1.3: Probability of an electron neutrino oscillating into a muon neutrino. This calculation was done in a two-neutrino mixing formalism (using equation 1.41) with $\Delta m^2 = \Delta m^2_{21} = 7.5 \times 10^{-5}\text{eV}^2$ and $\sin^2 2\theta = \sin^2 2\theta_{12} = 0.857$ (see table 1.1).

where $E_i = \sqrt{m_i^2 + p^2}$ and we have assumed that the neutrino has a fixed momentum $p$ (see [8]). We can rewrite this in the basis of flavor eigenstates using the transformation matrix $U$ from equation 1.40,

$$H'_{ll'} = \sum_{i,j} H_{ij} U_{li} U^*_{l'j} = \sum_i E_i U_{li} U^*_{l'i}.$$ \hspace{1cm} (1.43)

Therefore, for example, we will have

$$H_{0,ee} = E_1 (\cos \theta)^2 + E_2 (\sin \theta)^2$$
$$= (E_1 - E_2) \cos^2 \theta + E_2$$
$$= \frac{(E_1 - E_2)}{2} (1 + \cos 2\theta) + E_2$$ \hspace{1cm} (1.44)

where we have used $E_i \approx p(1 + \frac{1}{2}(m_i/p)^2)$ and $\Delta m^2 \equiv m_2^2 - m_1^2$. Calculating all of the other elements in a similar way, one finds

$$H_0 = \frac{\Delta m^2}{4E} \begin{pmatrix} -\cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix} + \begin{pmatrix} \frac{E_1 + E_2}{2} & 0 \\ 0 & \frac{E_1 + E_2}{2} \end{pmatrix}. \hspace{1cm} (1.45)$$
We can redefine the zero of energy and therefore eliminate the diagonal component. Now that we are in the flavor basis, we can add the additional contribution $V_e$ discussed above, which contributes only to $H_{ee}$, and if we subtract $V_e/2$ multiplied by the identity [20], we can again redefine our zero of energy and write

$$H = \frac{\Delta m^2}{4E} \left( \begin{array}{cc} -(\cos 2\theta - 2EV_e/\Delta m^2) & \sin 2\theta \\ \sin 2\theta & (\cos 2\theta - 2EV_e/\Delta m^2) \end{array} \right). \quad (1.46)$$

This can be rewritten in the form of the first term in equation 1.45 with a new squared mass splitting $\Delta m^2_M$ and mixing angle $\theta_M$ [20],

$$\sin^2 2\theta_M = \frac{\sin^2 2\theta}{\sin^2 2\theta + (\cos 2\theta - 2EV_e/\Delta m^2)^2}, \quad (1.47)$$

$$\Delta m_M = \Delta m \sqrt{\sin^2 2\theta + (\cos 2\theta - 2EV_e/\Delta m^2)^2}.$$ 

Therefore, $V_e$, determined by the electron density, can influence the amount of neutrino mixing present and even produce the condition of maximal mixing ($\theta = \pi/4$) when $V_e = \Delta m^2 \cos 2\theta/2E$. Note that because the ability of the effect to enhance mixing depends on the relative signs of $V_e$ (which is known, though different in the case of neutrinos and antineutrinos [16]) and $\Delta m^2 \cos 2\theta$, the effect is sensitive to the sign of this product involving the cosine of the mixing angle and the squared mass splitting. Current experimental data from observation of solar neutrinos is consistent with $\Delta m^2_{21} \cos 2\theta_{12} > 0$ and $\cos 2\theta_{12} > 0$ [16] given that $m_2 > m_1$, but the sign of $\Delta m^2_{31}$ is still unknown. Neutrinoless double-beta could play a key role in determining this sign and thus the neutrino mass hierarchy.

### 1.1.6 Neutrinoless Double-Beta Decay

We are now ready to discuss the main process of interest, $0\nu\beta\beta$. While there has not yet been a confirmed experimental observation of this process, its existence has many significant implications, and a large amount of thought and effort that has been devoted to it. More information in addition to what we describe here can be found in the reviews [10, 21, 22]. We will first make some reasonable assumptions in determining under what conditions the process can occur. Since new physics beyond the Standard Model of Particle Physics is required to enable $0\nu\beta\beta$ to occur, we cannot

---

7 In principle we can call either the heavier or the lighter neutrino “neutrino 1,” but since $\theta_{12}$ is a mixing angle, it determines how much of mass eigenstates “1” and “2” make up each flavor eigenstate and vice versa. Essentially this sign restriction tells us that the lighter of the two neutrinos whose masses are separated according to $\Delta m^2_{21}$ contains more electron flavor. Since a similar sign restriction has not been determined experimentally for $\Delta m^2_{31}$, a similar statement cannot yet be made as to whether the lightest of the three neutrinos contains the greatest or least amount of electron flavor.
be sure that we should even be discussing the process within our general framework of weak interactions. However, it is reasonable to start by making minimal extensions of our current thinking and impose some conditions on the neutrino required in order for $0\nu\beta\beta$ to exist. First we will assume that, like all weak interactions, we must conserve chirality (conservation of helicity as $m \to 0$) as discussed in section 1.1.2. The $0\nu\beta\beta$ process can be broken up into two stages (two $\beta$ decays) [19]

\[ d \to u + e^- + \bar{\nu}_e \quad (1) \]
\[ \bar{\nu}_e + d \to u + e^- \quad (2). \]  

(1.48)

Process 2 violates the concept of flavor lepton number conservation introduced in section 1.1.1, but as we know from neutrino oscillations, flavor lepton number is not conserved in all physical processes. Because lepton number in general could be nothing more than a manifestation of helicity, let us assume all fermions are massless at this point and examine the helicities of the particles in both processes. Both of these interactions are mediated by a charged $W$-boson, so only left-handed fermion components are involved. All particles are therefore in negative helicity states, and all anti-particles are in positive helicity states. In process 1 helicity is conserved, but it is not in process 2. Therefore for massless fermions process 2 would not be allowed, and so unless neutrinos are massive, $0\nu\beta\beta$ cannot occur.
It is also reasonable to assume that the particle theory describing weak interactions is a local gauge theory, meaning that the Lagrangian describing the theory is unchanged by certain “gauge” transformations that depend on the spatial coordinates (see [4] for a basic introduction to gauge theories). Under these conditions it was shown in [23] that the neutrino must be Majorana (the Schechter-Valle theorem), that is, unless neutrinos are Majorana $\nu\beta\beta$ cannot occur.

Under the above conditions of massive Majorana neutrinos, process 2 may be allowed but still involves a left-handed component, the smallness of which can be seen from the coefficient of the left-handed component of $\psi_1$ in equation 1.6, $\propto \sqrt{E - p} \approx \sqrt{m_\nu^2/2E}$. This leads to a suppression of the rate of such a physical process by a factor of $m_\nu^2/2E$. Therefore anti-neutrinos could simply be neutrinos with opposite helicity, and no processes such as (2) in equation 1.48 have been observed because the tiny neutrino mass $m_\nu$ is acting to suppress them.

If neutrinoless double-beta decay does occur, it is most anticipated to be described by the Feynman diagram shown in figure 1.5. This diagram describes the decay in terms of the participating initial and final particles and those exchanged between them. Note that the neutrino is not an initial or final particle in the decay and therefore is not observed physically - it is virtual. The existence of this virtual neutrino is not allowed unless the neutrino is Majorana [14], so in the Standard Model $0\nu\beta\beta$ is forbidden. The process involving the emission of two neutrinos ($2\nu\beta\beta$), however, is allowed in the Standard Model.

![Feynman diagram for $0\nu\beta\beta$. For the decay to proceed in this manner, the neutrino exchanged between the two electrons must be Majorana.](image)

Assuming Majorana neutrino exchange and therefore a nonzero decay rate, the calculated rate of neutrinoless double beta is given by [14]

$$\Gamma_{0\nu\beta\beta} = |m_{\beta\beta}|^2 |M^{0\nu}|^2 G(Q, Z),$$  

(1.49)
where \( M^{0\nu} \) is the amplitude for the decay represented by the Feynman diagram, \( G \) is a phase space factor that depends on the total energy released in the reaction, or Q-value, and the atomic number of the isotope in which the decay occurs, and \( m_{\beta\beta} \) is the the “effective” neutrino mass \([14]\),

\[
m_{\beta\beta} = \sum_{i=e,\nu,\tau} U_{ei}^2 m_i. \tag{1.50}
\]

Here only the “\( \nu_e \)” row of the mixing matrix \( U \) contributes to the effective neutrino mass. An exact rate cannot be calculated for any isotope, as the value of \( M \) depends on several nuclear matrix elements which cannot be calculated exactly, and the value of \( m_{\beta\beta} \) is unknown, though experiments continue to constrain its value.

As of this writing, the absolute values of the neutrino masses are unknown. While a clear observation of \( 0\nu\beta\beta \) cannot determine exactly the mass of any individual neutrino, it can determine the value of \( m_{\beta\beta} \) to the accuracy permitted by the uncertainty in the calculation of \( M^{0\nu} \). In terms of the mixing angles and phases in a Majorana mixing matrix (see equation 1.32), the value of the effective neutrino mass can be written

\[
m_{\beta\beta} = \cos^2 \theta_{13} \cos^2 \theta_{12} m_1 + \cos^2 \theta_{13} \sin^2 \theta_{12} e^{i\alpha_2} m_2 + \sin^2 \theta_{13} e^{i(\alpha_{31} - 2\delta)} m_3. \tag{1.51}
\]

Having measured experimentally the mass separations \( \Delta m^2_{21} \) and \( |\Delta m^2_{32}| \), and the mixing angles \( \theta_{12}, \theta_{23}, \) and \( \theta_{13} \), one can constrain the value of \( m_{\beta\beta} \) in several possible scenarios concerning the neutrino mass.

Figure 1.6 shows the results of a Monte Carlo simulation in which the experimentally determined parameters in equation 1.51 were varied according to their reported values and errors, and \( m_{\beta\beta} \) was determined for a random value of the minimum neutrino mass \( m_{\text{min}} \). All values with a reported mean and \( \sigma \) were drawn from a Gaussian distribution, \( \sin^2 \theta_{23} \) was chosen from a uniform distribution with values between its lower bound and 1, and the unknown phases \( \alpha_2 \) and \( \alpha_{31} - 2\delta \) were chosen from a uniform distribution between 0 and \( 2\pi \). The mass hierarchy or ordering of the neutrino masses depends on the unknown sign of \( \Delta m^2_{23} \) and was determined randomly to be “normal” \( (m_1 < m_2 \ll m_3) \) or “inverted” \( (m_3 \ll m_1 < m_2) \) with equal probability, and a random minimum neutrino mass was chosen between \( 10^{-5} \) eV and 1 eV. For each set of chosen values \( |m_{\beta\beta}| \) was computed, and 1000000 points were generated in total. For values of the minimum neutrino mass greater than about 0.1 eV, regardless of the sign of \( \Delta m^2_{23} \) the mass hierarchy is referred to as “degenerate,” as in this case the absolute values of all masses are significantly greater than either of the mass differences \( (m_1 \approx m_2 \approx m_3) \).

Figure 1.6 gives a concise picture of the physics that can be extracted from a \( 0\nu\beta\beta \) experiment even when the decay is not observed. Several limits are shown. The upper limit on \( |m_{\beta\beta}| \) from the combined results of the KamLAND-Zen and EXO experiments is shown as a horizontal line, and was determined in \([25]\) using the lower
bounds of the $0\nu\beta\beta$ half life measured in the two experiments at 90% confidence level. The value at which the line is drawn corresponds to the highest upper bound of 250 meV in the quoted range (120-250 meV) of values allowable by differences in calculated nuclear matrix elements. The vertical lines give the upper limit on the minimum neutrino mass given in [24],

$$\sum_{\nu=e,\mu,\tau} m_\nu < 0.66 \text{ eV}. \quad (1.52)$$

As $0\nu\beta\beta$ experiments and others that measure neutrino mixing parameters and/or place direct limits on neutrino masses continue to improve, more of the possible parameter space shown is excluded even if a discovery of an exact combination of values is not made.

### 1.1.7 Neutrinoless Double-Beta Decay: Experimental Considerations

The strategy behind a $0\nu\beta\beta$ experiment is to gather as much of a candidate isotope as possible, place it in a large detector capable of measuring the energy deposition
of two energetic (∼ several MeV) electrons, and wait long enough that a statistically significant number of decays occurs. This experimental challenge involves several decisions including which and how much isotope to use, how to design and fabricate a detector that can contain the isotope and observe the decay, and where to place the final detector. Few isotopes are a candidate for double beta decay since the single beta decay mode, a first-order weak process with a much higher rate, in most cases would provide excessive background in a search for the double beta mode. However, in some nuclei, nuclear pairing forces create the energy schematic shown in figure 1.7. Here the single beta decay mode cannot occur spontaneously due to energy conservation, but the double beta mode can.

![Energy schematic](image)

Figure 1.7: Energy schematic for most nuclei that are candidates for double-beta decay. Nuclear pairing forces produce a configuration in which single-β decay is disallowed by energy conservation, but double-β decay is allowed (figure after [22]).

The goal of a neutrinoless double-beta experiment is to measure a signal that confirms with sufficient statistical significance that the $0\nu\beta\beta$ decay occurs, and most commonly the energy of the two emitted electrons is measured. Other aspects of the reaction, such as the tracks formed by the electrons as they lose energy in the detector and the charged ion left over after the nucleus has decayed, may serve as additional confirmation that the energy signal measured corresponds to a double beta decay. Three guiding principles of any modern $0\nu\beta\beta$ experiment are good energy resolution, low background, and the use of a large amount of isotope mass.

1 - Good energy resolution

Because nearly all of the energy in $0\nu\beta\beta$ is released in the form of two electrons, an infinitely precise measurement of the energy of the two electrons will yield a single value equal to the Q-value of the reaction, $Q_{\beta\beta}$.

$$Z^A_N \rightarrow Z^{Z+2}_{N-2} + A + Q_{\beta\beta}.$$  \hspace{1cm} (1.53)

The value $Q_{\beta\beta}$ is different for each candidate nucleus, and several values are listed in table 1.3. However, in the two-neutrino mode ($2\nu\beta\beta$), a variable amount of energy
is carried off by the two neutrinos, and with a finite resolution one may not be able to distinguish $2\nu\beta\beta$ events for which very little energy was carried off by neutrinos from $0\nu\beta\beta$ events for which no energy was carried off by neutrinos. Figure 1.8 shows a normalized Gaussian peak with varying resolution added to a normalized $2\nu\beta\beta$ spectrum calculated using equation 18b of [26],

$$
\frac{d\omega}{dT} \sim \tilde{E} (Q_{\beta\beta} - \tilde{E})^5 [\tilde{E}^4 + 10\tilde{E}^3 + 40\tilde{E}^2 + 60\tilde{E} + 30],
$$

where $\tilde{E} = E/m_e$ is the ratio of the energy to the electron mass. In the figure the ratio of the total $2\nu\beta\beta$ and $0\nu\beta\beta$ rates was also varied. Better resolution allows one to more easily distinguish the $0\nu\beta\beta$ peak, but also shows that if the rate of $0\nu\beta\beta$ is very low this becomes increasingly difficult.

Figure 1.8: Separation of $2\nu\beta\beta$ and $0\nu\beta\beta$ events with varying detector energy resolution and relative decay lifetimes. A normalized Gaussian $0\nu\beta\beta$ peak at $Q = 2458$ keV (isotope $^{136}$Xe) with varying FWHM energy resolution was added to a normalized $2\nu\beta\beta$ spectrum with varying total rate ratios. The FWHM resolution was applied to the $2\nu\beta\beta$ spectrum via a convolution. The insets in the plots with the lowest $\Gamma_{0\nu}/\Gamma_{2\nu}$ show a close-up of the $0\nu\beta\beta$ peak region, and in the upper inset the dotted line shows the $0\nu\beta\beta$ contribution to the spectrum, while the solid line in all cases shows the sum of both $2\nu\beta\beta$ and $0\nu\beta\beta$ contributions. The approach taken here is similar to that of figure 1 of [27].

For the majority of detectors, energy resolution improves as $\sim E^{-1/2}$, so a higher Q-value is an advantage assuming events with increasing energy do not become gradually more difficult to detect due to the design of the detector. In general a higher
Q-value will also lead to a higher phase space factor $G(Q, Z)$, thus increasing the rate (see equation 1.49), and increased immunity to natural background, which becomes greatly reduced above the $^{208}$Tl gamma line at $E = 2614.5$ keV [28, 29].

2 - Low background

Because the process of interest is so rare, very little or no background is necessary to make a claim of observation. Primary backgrounds include gammas from naturally present radioactive isotopes, neutrons resulting from the interactions of muons in rock, and radiation emitted from the components of the detector itself [30]. Note that even if all other forms of background are eliminated, there still remains the irreducible background of the $2
\nu\beta\beta$ mode.

To emphasize the importance of low background, we demonstrate its influence on the sensitivity of a given experiment to the effective neutrino mass $m_{\beta\beta}$ (see section 1.1.6), a property of only the neutrinos and their mixing parameters and independent of the specific isotope chosen for the experiment. In this calculation we follow [21] and use the same notation. If we have a radioactive decay process with half-life $T_{1/2}$, we must have that the number of decays $N$ occurring in a time $t \ll T_{1/2}$ is

$$N = N_0(1 - e^{-t\ln2/T_{1/2}}) \approx N_0 t \ln 2/T_{1/2},$$

(1.55)

where $N_0$ is the initial number of atoms of the decaying species $N_0 = N_A \cdot M_{\beta\beta}/W_{\beta\beta}$. Here $N_A$ is Avogadro’s number, $M_{\beta\beta}$ is the amount of $\beta\beta$ candidate isotope mass present, and $W_{\beta\beta}$ is its molar mass. The observed number of decays is $N\epsilon$ where $\epsilon$ is the total detection efficiency of the experiment, so from equations 1.49 and 1.55 we have

$$|m_{\beta\beta}|^2 = \frac{1}{G^{0\nu}|M^{0\nu}|^2 T_{1/2}} = \frac{1}{G^{0\nu}|M^{0\nu}|^2} \left( \frac{N}{N_A(M_{\beta\beta}/W_{\beta\beta})t\epsilon \ln 2} \right).$$

(1.56)

Since $N_A$, $G^{0\nu}$, $|M^{0\nu}|^2$, and $W_{\beta\beta}$ are all constant for a given isotope, we can combine them into a single constant. To set a limit on the decay, we wish to know the lowest value of $|m_{\beta\beta}|^2$ consistent with the experiment but that is not equal to 0 (in which case one would never observe any events). This is the value for which $N = 1$, so our limit is

$$|m_{\beta\beta}| > K_1 \sqrt{\frac{1}{M_{\beta\beta} \cdot t \cdot \epsilon}}.$$  

(1.57)

Now if we consider a nonzero background, equation 1.57 is modified. Our sensitivity or confidence in a number of counted $\beta\beta$ events can be expressed as follows. If
$N$ events are observed with a confidence level of $n\sigma$ (assuming a Poisson distribution for $N$), then the following relation will hold \[31\],

$$N = n\sqrt{B + \overbar{N}},$$

(1.58)

where $B$ is the number of background events. $B$ can be expressed as

$$B = bt\cdot M_{\beta\beta}\cdot \Delta E,$$

(1.59)

where $b$ is the background rate in units of counts/(time\cdot mass\cdot energy), and $\Delta E$ is the energy region from which events are drawn and over which $b$ is assumed to be constant. Equation 1.58 tells us that if we assume a given signal to noise ratio $N = rB$ and a given confidence level of a result, the number of observed $\beta\beta$ events will scale as $N \sim \sqrt{B}$. Using this relation in equation 1.56 and using equation 1.59, we find

$$|m_{\beta\beta}| > K_2 \frac{1}{\sqrt{\epsilon}} \left( \frac{b \cdot \Delta E}{M_{\beta\beta} \cdot t} \right)^{1/4}.$$

(1.60)

Because of background, the limit that can be set on $m_{\beta\beta}$ decreases as the inverse fourth root of the running time rather than the inverse square root. Energy resolution becomes again important as the larger the value of $\Delta E$, the greater the observed background.

In fact, no $0\nu\beta\beta$ experiment can be completely background free without perfect energy resolution. Even if one managed to eliminate all other background processes, the $2\nu\beta\beta$ mode still produces background events that cannot be distinguished from $0\nu\beta\beta$ events except in the precise value of the total energy; the angular distribution of the emitted electrons is even the same \[28\]. In \[27\], a formula is given for the fraction of $2\nu\beta\beta$ decays that occur with energy that would fall in the $0\nu\beta\beta$ peak for a given FWHM energy resolution $\delta$,

$$F = \frac{F_0Q\delta^6}{m_e},$$

(1.61)

where $F_0$ is some factor that depends on $\delta$. Taking $^{136}$Xe to be the isotope of interest and assuming a mass of 20 tons ($N_0 = 20000 \text{ kg}/136 \text{ amu} = 8.856 \times 10^{28}$ atoms), and using the value for the $2\nu\beta\beta$ half life $T_{1/2}^{2\nu} = 2.11 \times 10^{21}$ years from the EXO collaboration \[32, 33\] and $Q = 2458$ keV \[34\], we have a number $N$ of $2\nu\beta\beta$ decays/year,

$$N = N_0 - N_0 e^{-\frac{T\ln 2}{T_{1/2}^{2\nu}}} \approx N_0 \left( \frac{T\ln 2}{T_{1/2}^{2\nu}} \right) = 2.909 \times 10^7 \text{ decays},$$

(1.62)

The values of $\delta$ and the corresponding factors from \[27\] are shown with the calculation of $F$ and $F \cdot N$ (the total number of decays in the peak for a 20-ton detector) in table 1.2.\footnote{This calculation was suggested and first carried out by Mike Heffner (LLNL).}
Table 1.2: Results of the calculation of the number of $2\nu\beta\beta$ decays that will occur in the $0\nu\beta\beta$ peak for a 20-ton $^{136}$Xe double-beta search experiment with varying energy resolution.

<table>
<thead>
<tr>
<th>FWHM</th>
<th>$F_0$</th>
<th>$F$</th>
<th>$2\nu\beta\beta$ Decays in $0\nu\beta\beta$ Peak</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>8.5</td>
<td>$4.09 \times 10^{-11}$</td>
<td>0.001</td>
</tr>
<tr>
<td>5%</td>
<td>7</td>
<td>$5.26 \times 10^{-7}$</td>
<td>15</td>
</tr>
<tr>
<td>10%</td>
<td>5</td>
<td>$2.40 \times 10^{-5}$</td>
<td>700</td>
</tr>
</tbody>
</table>

Table 1.3: Selected candidate $0\nu\beta\beta$ isotopes, their natural abundance, Q-values for double beta decay, and phase space factors $G(Q, Z)$. Natural abundances are taken from [35], and Q-values and phase space factors are taken from table III of [28] (see references therein).

<table>
<thead>
<tr>
<th>Isotope</th>
<th>Abundance (%)</th>
<th>$Q_{\beta\beta}$ (MeV)</th>
<th>$G(Q, Z)$ ($\times 10^{-14}$ yr$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^{48}$Ca</td>
<td>0.187</td>
<td>4.272</td>
<td>6.4</td>
</tr>
<tr>
<td>$^{76}$Ge</td>
<td>7.44</td>
<td>2.039</td>
<td>0.6</td>
</tr>
<tr>
<td>$^{82}$Se</td>
<td>8.73</td>
<td>2.995</td>
<td>2.7</td>
</tr>
<tr>
<td>$^{96}$Zr</td>
<td>2.80</td>
<td>3.350</td>
<td>5.7</td>
</tr>
<tr>
<td>$^{100}$Mo</td>
<td>9.63</td>
<td>3.034</td>
<td>4.6</td>
</tr>
<tr>
<td>$^{116}$Cd</td>
<td>7.49</td>
<td>2.805</td>
<td>4.9</td>
</tr>
<tr>
<td>$^{130}$Te</td>
<td>33.8</td>
<td>2.529</td>
<td>4.1</td>
</tr>
<tr>
<td>$^{136}$Xe</td>
<td>8.9</td>
<td>2.468</td>
<td>4.4</td>
</tr>
<tr>
<td>$^{150}$Nd</td>
<td>5.64</td>
<td>3.367</td>
<td>19</td>
</tr>
</tbody>
</table>

3 - Large mass

In order to have a reasonable chance of observing a process with a half life as long as that of $0\nu\beta\beta$, one must have a detector containing a significant amount of mass of the candidate isotope of interest. Modern $0\nu\beta\beta$ experiments as of this writing consist of detectors with isotope masses of 10’s - 100’s of kilograms, and as no convincing evidence has yet been observed for the decay, future experiments are considering masses on the scale of tons. Equations 1.57 and 1.60 show that the sensitivity of an experiment scales with isotope mass as it scales with running time.

One important consideration in striving for a high candidate isotope mass is the natural abundance of that isotope. Though many experiments use material enriched to the isotope of interest, the enrichment process can be expensive, so a greater natural abundance is normally favorable for the purposes of $0\nu\beta\beta$ experiments. Table 1.3 lists the natural abundances of several isotopes of interest along with their calculated phase space factors and Q-values.
1.1.8 Neutrinoless Double-Beta Decay: Experiments

Here we briefly summarize the experimental efforts in the search for neutrinoless double-beta decay and note the current techniques adopted for present and future experiments. General information for this section was found in [21], [22] [36], and [37].

Experiments using Semiconductors

One method of detecting double beta decay is by measuring the ionization produced in a semiconducting detector by the two high-energy electrons emitted in the decay. A bias voltage is applied across the semiconducting medium so that a strong electric field is present inside. When an energetic particle deposits energy in the medium, thus launching some electrons from the valence band into the conduction band and producing electron-hole pairs, the electrons and holes are drifted in opposite directions by the field and the resulting induced current on the electrodes constitutes the observed signal. In general, only the ionization signal is available, and the track left by the particle is pointlike and cannot be imaged (though one may be able to distinguish particles that interacted in more than one distinct location or “multi-site” events).

Several recent and present day $0\nu\beta\beta$ experiments consist of Ge detectors containing the candidate isotope $^{76}$Ge, including the Heidelberg-Moscow experiment [38], GERDA [39], and MAJORANA [40]. COBRA [41] makes use of CdZnTe (CZT) semiconductor detectors to search for $0\nu\beta\beta$ in $^{116}$Cd. A controversial claim of discovery of $0\nu\beta\beta$ has been made in $^{76}$Ge by a subset of the Heidelberg-Moscow experiment [42], which has since been ruled out by GERDA [43].

Experiments using Bolometry

Experiments based on bolometry determine the energy deposited in a crystal by measuring a resulting temperature change. The crystals are maintained at a low temperature, and as their heat capacity is proportional to the cubed temperature, $C = dQ/dT \propto T^3$, a very low temperature will result in a very low $C$, meaning a relatively large change in temperature for even a small energy deposition. The crystals are in thermal equilibrium with a heat reservoir via a path of low thermal conductivity and are monitored by a high-precision thermistor, so that when an energy $\Delta E$ is deposited, it is readily measured by the thermistor [44],

$$\Delta T = \frac{\Delta E}{C} e^{-t/\tau},$$

where the temperature returns to that of the heat reservoir with time constant $\tau$. The candidate double-beta isotope forms at least part of the composition of the material from which the crystals are constructed. This approach in general offers very good
energy resolution. Several experiments that will search for $0\nu\beta\beta$ using bolometry are CUORE [45, 44] (isotope $^{130}$Te with TeO$_2$ crystals), AMoRE [46] (isotope $^{100}$Mo with CaMoO$_4$ crystals), and LUCIFER [47] (isotope $^{82}$Se with ZnSe crystals). In AMoRE and LUCIFER, scintillating crystals are used, providing an additional signal (scintillation light) produced by the energetic electrons that can be measured simultaneously with $\Delta T$. Due to a difference in the ratio of energy deposited in the heat and light channels depending on the nature of the incident radiation, it is possible to use this additional signal to identify events as energy depositions by alpha particles or by beta/gamma rays, providing additional background rejection capabilities.

**Experiments using Noble Element TPCs**

Time projection chambers are detectors that drift electrons ionized by incident radiation and by design measure the deposited energy and the location at which an event occurs along an axis parallel to the direction of ionization drift. In double-beta experiments, the only noble element currently in use is xenon, as $^{136}$Xe is a candidate isotope for $0\nu\beta\beta$. Energetic electrons produce both ionization and scintillation in xenon, and therefore one can in principle measure both signals. While the energy resolution is not as good as that of bolometry or semiconductor-based experiments, xenon TPCs offer several advantages including the ability to purify the xenon continually during operation and the ability to use the mass on the edges of the detector to shield the inner region of the detector. The event localization capability of the TPC provides the means of making such a fiducial cut. Xenon in the gas phase offers in addition room temperature operation conditions, improved energy resolution over the liquid phase, and the ability to image the tracks formed by the incident radiation.

The EXO experiment [48] is a dual-phase (liquid/gas) TPC in which the energy is deposited in the liquid phase, and the gas phase provides the gain mechanism via electroluminescence (see chapter 2). The NEXT experiment (see section 1.1.9) makes use of a gaseous xenon TPC and will be described in more detail in section 1.1.9 and chapter 3.

**Experiments with Independent Source and Detection Medium**

All of the experiments we have discussed so far in this section have had one common characteristic; the detector medium itself with which the electrons produced in double-beta decay would interact is composed in full or in part by the candidate isotope that would undergo the decay. This is not the case in all double-beta experiments.

One approach that has been taken by the KamLAND-Zen [49] (isotope $^{136}$Xe) and SNO+ [50] (isotope $^{130}$Te) experiments is the dissolution of the candidate isotope in a liquid scintillator. The scintillator itself yields the photons that will be detected to determine the energy of the double-beta electrons, but the dissolved isotope is still present inside the detector. The SuperNEMO [51] and MOON [52, 53] experiments have adopted an approach in which thin films containing the candidate isotope are
placed in a setup in which they are surrounded by tracking detectors and calorimeters. These approaches are not straightforward to scale to higher masses but allow versatility in isotope selection.

1.1.9 The NEXT-100 Experiment

NEXT-100 is a modern neutrinoless double-beta decay search experiment that will use a high-pressure xenon TPC with 100 kg of xenon enriched 91% to the $\beta\beta$-decay candidate $^{136}$Xe. The work described in chapter 3 was performed as research and development for this experiment and demonstrates proof-of-concept for its two primary operating principles: energy measurement with electroluminescent gain and 3D particle tracking using electron drift and a 2D array of silicon photomultipliers. The electroluminescent gain serves as a low-noise amplification process providing good energy resolution, and the particle tracking allows for the identification of the two energetic electrons emitted in the $\beta\beta$ reaction, allowing for the rejection of events that do not produce this signature and thereby greatly reducing the rate of background. The results in section 3.6 show that with electroluminescence a resolution should be obtainable in high pressure xenon of near 0.5% FWHM at the energy $Q_{\beta\beta} = 2.468$ MeV (see table 1.3), and a two-electron event can be identified by a long track with regions of relatively denser energy deposition at the two ends. With these principles and the large amount of xenon present in the detector, NEXT-100 addresses all three important experimental considerations discussed in section 1.1.7.

Detector

A cross section of the proposed NEXT-100 detector [54] is shown in figure 1.9. The detector will be placed in the Laboratorio Subterráneo de Canfranc in Canfranc, Spain. Its functionality is largely similar to that of the NEXT prototypes, one of which will be described in detail in chapter 3. According to the plan at the time of this writing [54], its main components will be enclosed in a stainless steel pressure vessel of inner diameter 1.36 m and length of 2.28 m. The interior of the vessel will be lined with a 12 cm thick layer of low-activity copper which will provide additional shielding for the interior sensitive region, particularly shielding from radioactivity emitted from the pressure vessel itself. The field cage will lie inside the copper shielding and will consist of copper strips on high density polyethylene connected through a resistor network to establish the electric field in a 130 cm long drift region. Three wire meshes divide the TPC into the drift region and a 5 mm long EL region. The tracking plane will lie behind the EL region and will consist of several thousand silicon photomultipliers (SiPMs) spaced by approximately 1 cm coated with the wavelength shifter tetraphenyl butadiene to wavelength shift the UV xenon light into a regime in which the SiPMs are sensitive. The energy plane will lie at the opposite end of the TPC and will consist of 60, 3-inch diameter photomultiplier tubes. Electrons will drift horizontally
in the experiment and the total mass of the pressure vessel, copper shielding, field cage, and tracking and energy planes is estimated to be 12,500 kg. The entire TPC will be housed inside a 20 cm thick wall constructed from lead bricks.

Figure 1.9: (Figure from [54]) Cross section of the proposed NEXT-100 detector.

**Background Rejection**

A radiopurity campaign has been ongoing to measure the radioactivity of all of the components used in constructing the detector using Glow-Discharge Mass Spectrometry (GDMS) and gamma ray spectroscopy with high-purity Ge detectors [55]. The campaign measures the rate of gamma emission due to several different isotopes of interest. The primary backgrounds of concern consist of gammas from the decays of isotopes $^{214}$Bi (2448 keV) and $^{208}$Tl (2615 keV) [29] with energies near the value of $Q_{\beta\beta}$ for $^{136}$Xe. These isotopes are produced in the decay chains of $^{238}$U and $^{232}$Th respectively, and the amount of gamma activity resulting from each of these chains is reported amongst others in the radiopurity studies.

Background rejection [56] in NEXT-100 is anticipated to be realized by application of the following series of cuts to each event (consisting of energy and tracking information) acquired in the detector$^9$: First, a broad cut in energy near $Q_{\beta\beta}$ (a window of several hundred keV) is applied. From these events, a fiducial cut is made selecting only those whose entire energy was deposited some distance (several cm)

---

$^9$This procedure was developed by Paola Ferrario, Justo Martín-Albo, Javier Muñoz, and J.J. Gómez Cadenas at the Instituto de Física Corpuscular, and recent work done by them is summarized here.
from the walls of the active region. Events consisting of multiple tracks are also
discarded. The reconstructed track of the event is then analyzed to determine whether
one or two regions of high ionization density are present, as tracks produced by the
two high-energy electrons emitted in $\beta\beta$ decay should contain two of such regions
while background gamma-induced tracks only one (see figure 1.10). Note that this
topological signature is present for both $2\nu\beta\beta$ and $0\nu\beta\beta$ events, so a final energy
cut near $Q_{\beta\beta}$ is made to isolate $0\nu\beta\beta$ events. Preliminary studies have shown that
tracking and energy cuts are expected to reject background events due to $^{214}$Bi and
$^{208}$Tl by a factor of order $10^{-7}$, and initial estimates give an overall background rate
of order $5 \times 10^{-4}$ counts/keV-kg-yr for NEXT-100 [56]. Though the track identifica-
tion capabilities give a significant advantage in background reduction, they are not
100% efficient and therefore reduce the overall detection efficiency; the efficiency of
NEXT-100 is projected to be around 30% [57].

![Figure 1.10: (Figure from [58]) Simulated two-electron track produced by a $\beta\beta$ event.](image)
The two regions of higher ionization density mark the ends of the ionization tracks
produced by each of the two electrons. The presence of two such regions can be used
to distinguish double beta decay events from single-electron background events which
would contain only one of such regions.

**Projected Sensitivity**

Though the NEXT-100 experiment is feasible from the standpoint of cost and avail-
ability of facilities and materials, one must examine its physics discovery potential to
evaluate its overall merit. This is done for a $0\nu\beta\beta$ experiment by calculating its sen-
sitivity to $m_{\beta\beta}$; effectively the largest value of $m_{\beta\beta}$ that could be the true value and
still yield a null result in the experiment to some degree of confidence. The sensitivity
to $m_{\beta\beta}$ for several $0\nu\beta\beta$ experiments was evaluated in [57] using a Feldman-Cousins
unified approach for determining, given a measured number of events and an esti-
mated background rate, upper limits on the mean of the Poisson distribution that
would describe the statistics of the detected events. The sensitivity was computed using a selected value of the nuclear matrix element $M^{0\nu}$ (see equation 1.49) for each isotope based on computations using several different nuclear models, assuming a $0\nu\beta\beta$ detection efficiency and an energy resolution and background rate at $Q_{\beta\beta}$. The results for 7 different $0\nu\beta\beta$ experiments are summarized in figure 1.11, taken from [57] for “reference” and “optimistic” values of background rate and isotope mass.

![Graph](image)

Figure 1.11: (Figure from [57]) The predicted sensitivity of several $0\nu\beta\beta$ experiments to $m_{\beta\beta}$ vs. total detector exposure time. Plot (a) shows the results for a “reference” case of assumed background rate and $\beta\beta$ isotope mass, and plot (b) shows the same results for “optimistic” estimates (see table 2 of [57] for more details). The solid dots give the sensitivity value for each experiment assuming the detector runs for 10 years.

From figure 1.11, one can conclude that NEXT-100 is comparable to other current-generation $0\nu\beta\beta$ experiments in $m_{\beta\beta}$ sensitivity. However, as noted in [57], xenon-based experiments appear to be the most straightforward to scale to higher masses and therefore higher sensitivities for future generation experiments. This along with the advantages offered by xenon gas make NEXT-100 a valuable experiment in both the current and future studies of $0\nu\beta\beta$ decay.

### 1.2 Dark Matter

Dark matter is a mysterious type of matter that has never been directly observed on earth but is believed to exist in the universe due to the outstanding amount of cosmological evidence in its favor (see for example [59], summarized here). Beginning in the 1930’s, observations of the motion of mass in galaxies seemed to require the presence of some amount of mass in addition to the light-emitting mass observed in order to make sense. By the end of the 1970’s, more detailed observations including
measurements of rotational velocities of matter in galaxies, limits on temperature fluctuations in the cosmic microwave background (CMB), studies of x-rays emitted by baryonic gas between clusters of galaxies, and mapping of the large-scale structure of galaxies and clusters, made it clear that the massive components dictating the formation and development of the universe could not be known baryonic matter. This lead to the development of the currently accepted model of cosmology, the \( \Lambda \)CDM model, in the 1980's. Since then, precise measurements of CMB temperature fluctuations, baryon acoustic oscillations, and mass distributions of galaxies using x-ray emission and gravitational lensing have lead to the refinement and development of this model to arrive at our current understanding of the universe; a collection of spacetime that began expanding at a moment we call the “Big Bang” and whose expansion has been driven by the energy contained within - 70% an unknown “dark energy,” and the remaining 30% mass energy, 26% of which is dark matter, and only 4% the baryonic matter that makes up our familiar world.

Here we briefly review the major evidence for dark matter and the standard \( \Lambda \)CDM model of cosmology, discuss the possible candidates for dark matter, focusing specifically on WIMP dark matter, and describe the modern attempts and experimental strategies behind the direct detection of dark matter.

### 1.2.1 Evidence for Dark Matter

**Galactic Rotation Curves**

The existence of dark matter is widely accepted despite our inability so far to produce or observe any in the laboratory for a number of reasons. One of the earliest pieces of evidence came from observing the rotational velocity of luminous (photon-emitting) matter in galaxies. For an object orbiting a galactic center under Newtonian gravitational forces (in a circular orbit), one would expect a velocity

\[
v(r) = \sqrt{\frac{GM(r)}{r}},
\]

where \( G \) is Newton’s gravitational constant and \( M(r) \) is the total mass enclosed within radial distance \( r \). Therefore, unless \( M(r) \) is increasing significantly with \( r \), a galactic rotation curve should show decreasing \( v \) with increasing \( r \). However, as shown by the example in figure 1.12, measurements of rotational velocities in spiral galaxies have consistently shown a constant velocity at large radial distances, and a resulting \( M \sim r \) dependence. The presence of a “halo” of non-luminous, massive dark matter provides a valid explanation for these observations. Note that this behavior can also be explained [60] using a modification to the laws of Newtonian gravity [61].
Gravitational Lensing and X-ray Emission

Significant evidence in favor of dark matter has been put forth based on mapping the location and quantity of galactic mass using data from x-ray emission and gravitational lensing. X-rays emitted from hot gas in galaxies and galaxy clusters have been measured and used to draw conclusions on the total mass present and its distribution given the density and temperature profiles for the gas. Since most of the baryonic matter in the clusters is found in this hot gas, the fraction of the total mass that is baryonic can be determined. Such observations show that only of order 10% of the mass in typical clusters and galaxies is baryonic [59].

Gravitational lensing identifies mass based on its ability to distort light emitted from objects behind it. As the light passes nearby a massive object, its path is diverted in the strong gravitational field produced by the object according to Einstein’s theory of general relativity. The result is a stretched or otherwise altered object and/or multiple images of the same object, and an analysis of such effects yields information about the mass of the object producing them.

Both of these approaches have been taken in the observation of the “bullet cluster” [66] in figure 1.13. Due to the distinct behaviors of two apparent populations in this observation, one cannot explain this based on the assumption of a modification of the laws of gravitation for one single type of matter.

Large Scale Structure and Simulation

It is believed that the structure of the universe was formed from a nearly uniform distribution of matter, in which small fluctuations in density grew via gravitational interactions into the stars and galaxies observable today. Dark matter was the driving
force behind this structure formation, and one can imagine that it paved the way for baryonic matter, setting out the path along which stars and galaxies settled into a “lacework” of clusters and voids. Detailed simulations such as the Millennium [67] and Bolshoi [68] simulations that operate based on a cosmological model assuming the existence of cold dark matter (see section 1.2.2) are capable of reproducing galactic structure that agrees well with observations. Figure 1.14 shows a portion of the sky for which galaxies were identified and mapped by the Sloan Digital Sky Survey, and the corresponding map from the Bolshoi simulation, which should be compared not based on the exact locations of galaxies and clusters but based on statistics [62] in the distribution of galaxies.

The existence of large scale structure provides another key piece of evidence [59] for dark matter. At some point in the early universe, baryonic matter existed in the form of gas so hot that electrons could not remain bound to their nuclei, but as the temperature of the gas fell electrons recombined with nuclei to form neutral atoms. Only at this point, called “recombination,” was gravity able to begin its work of structure formation, and so the fluctuations in density at this time were the starting point from which structure could form. This is also the point at which photons were able to propagate freely amidst the baryons, and so the fluctuations in density can be deduced by observation of the Cosmic Microwave Background (CMB). It was found that these fluctuations were not nearly great enough in magnitude to produce the structure seen in the universe today. Therefore some other type of matter must exist that interacts gravitationally but very little if at all with photons and was distributed with density fluctuations greater than those deduced from the CMB. That is, baryonic
matter could not have provided the “seeds” necessary from which the universe could grow, and dark matter once again is needed to guide the baryonic matter in structure formation.

The Cosmic Microwave Background

When the universe reached the point in its evolution at which its expansion rate carried photons apart faster than their rate of creation and annihilation could maintain an equilibrium density, these photons began to propagate freely and became what is now known as the Cosmic Microwave Background (CMB) [70]. They exist now throughout the universe, and on average the energies of these photons are consistent with the spectrum of blackbody radiation at temperature $T = 2.726$ K. However, small deviations from this temperature of order $\sim 10^{-5}$ K have been observed when surveying different parts of the sky and can be expressed in terms of spherical harmonics $Y_{lm}$ as [70]

$$\frac{\delta T}{T}(\theta, \phi) = \sum_{m=-l}^{l} \sum_{l=2}^{\infty} a_{lm} Y_{lm}(\theta, \phi).$$

A plot of the variance of the coefficients $C_l \propto \sum_m |a_{lm}|^2/(2l + 1)$ against $l$ is called a CMB power spectrum and contains several peaks that can be used to constrain the parameters of cosmological models such as the ΛCDM model discussed in section 1.2.2. Figure 1.15 shows such a power spectrum produced by the Planck collaboration [13] and fit to a cosmological model.
Figure 1.15: (Figure from [13]) Fit of CMB temperature power spectrum to a ΛCDM model built on data from the Planck mission and other CMB experiments including the WMAP survey. Error bars and shaded regions include variations due to motion of the objects under observation (cosmic variance). The quantity plotted is $l(l+1)C_l/2\pi$ (see equation 1.65).

The fit shown in figure 1.15 produced the parameters, amongst others [24],

$$\Omega_b h^2 = 0.02207 \pm 0.00027, \quad \Omega_c h^2 = 0.1198 \pm 0.0026,$$

(1.66)

where $\Omega_b$ is the baryon density in the universe and $\Omega_c$ is the density of cold dark matter in the universe, and the range given above comes from 68% confidence limits of the fit. The “density” values $\Omega_i$ are the fraction of the total energy in the universe belonging to the $i$th component (see section 1.2.2), and in this case $h = 0.673$. Therefore approximately 26.4% of the energy in the universe is due to dark matter, while only about 4.9% is due to baryonic matter. The strong agreement between the model and the data in this case provides evidence for the existence of an abundance of dark matter in the universe.

1.2.2 ΛCDM: A Cosmological Model

Based on Einstein’s theory of general relativity, one can construct a cosmological model that describes the evolution of the universe. Here we define the main components of the ΛCDM model (following [71]), currently accepted as the “standard model” of cosmology, which has been shown to agree well with current experimental data. Building a model involves choosing a metric which describes the geometry of
space and time, and choosing a form in which to express the energy and momentum that will live in this space and time. Once these are chosen, Einstein’s equation tells us how spacetime is affected by the presence of the energy and momentum, and how the energy and momentum propagates through spacetime. This equation is

\[ R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 8\pi G T_{\mu\nu}. \]  

Here \( g_{\mu\nu} \) is the metric, and it is a tensor object, as is \( R_{\mu\nu} \), called the Ricci tensor, which contains derivatives of the metric and is related to the curvature of the spacetime it specifies. \( G \) is Newton’s gravitational constant, and \( T_{\mu\nu} \) is the energy-momentum tensor whose elements are the energy and momentum densities \((T^0^0\) and \( T^0^i\)) and the momentum fluxes \((T^{ij})\) of the matter it describes. This tensor is chosen to describe a “perfect fluid,”

\[ T_{\mu\nu} = (\rho + p) U_\mu U_\nu + p g_{\mu\nu}, \]

where \( U_\mu \) is the four-velocity of the fluid, \( \rho \) is its energy density, and \( p \) is its pressure. The metric chosen in the \( \Lambda \)CDM model is usually written as a relativistically invariant interval \( ds^2 \) which is a product of the metric and differentials of the coordinates in terms of which the metric is written,

\[ ds^2 = g_{\mu\nu} dx^\mu dx^\nu = -dt^2 + a^2(t) \left[ \frac{dr^2}{1-kr^2} + r^2 d\Omega^2 \right], \]

where here we are using 4D polar coordinates \((t, r, \theta, \phi)\). The variable \( k \) is the curvature, and if \( k > 0 \), the space is called “closed,” if \( k = 0 \) the space is “flat,” and if \( k < 0 \) the space is “open.” The variable \( a \) is known as the “scale factor” which determines the relative size of the spatial and temporal parts of the invariant interval.

The mathematical objects defined above merit much further explanation than that given here, and we will simply state the relationships between them implied by general relativity. In short we have just written down the mathematical description of a universe that is capable of expanding in space (through the factor \( a \)) and doing so isotropically (all of space is expanding by the same factor). In addition, this space may be curved (through \( k \)), and in it lies matter and energy that has a well-defined energy density \( \rho \) and isotropic pressure \( p \) in its rest frame. We will now impose the additional condition that \( \rho \) and \( p \) are related by the equation \( p = w \rho \). The application of energy-momentum conservation to \( T^{\mu\nu} \) requires

\[ \rho = \rho_0 a^{-3(1+w)}, \]

and using \( T^{\mu\nu} \) and our metric from equation 1.69 in Einstein’s equation,

\[ H^2 \equiv \left( \frac{\dot{a}}{a} \right)^2 = \frac{8\pi G}{3} \rho - \frac{k}{a^2}. \]
Equation 1.71 is known as the Friedmann equation (it is actually one of two Friedmann
equations obtainable from Einstein’s equation), and $H$ is called the Hubble expansion
rate (the value of which at the current time is called the Hubble constant and denoted
as $H_0$). We rewrite this equation by noting that the total energy density $\rho = \sum_i \rho_i$, where the $\rho_i$ are the energy densities of the several components that constitute the
total energy. If we define the quantities called density parameters $\Omega_i = (8\pi G/3H^2)\rho_i$,
we can rewrite equation 1.71 as

$$\sum_i \Omega_i = 1 + \frac{\kappa^2}{H^2 a^2}. \quad (1.72)$$

The $\Lambda$CDM model contains a density parameter for matter, $\Omega_m$ (split into $\Omega_b$
for baryonic matter and $\Omega_c$ for cold dark matter) and a density parameter $\Omega_\Lambda$, which
describes the vacuum energy. Through detailed observations of the CMB (producing
data such as that shown in figure 1.15), it is possible to determine the value of these
parameters and of $H_0$. From the Planck mission [24]

$$\Omega_m = 0.314 \pm 0.020, \quad \Omega_\Lambda = 0.686 \pm 0.020, \quad H_0 = 67.4 \pm 1.4, \quad (1.73)$$

and the values for $\Omega_c$ and $\Omega_b$ are given in equation 1.66. Note that $\Omega_m + \Omega_\Lambda \approx 1$, and
so Planck data gives $\kappa \approx 0$, or a spatially flat universe.

### 1.2.3 Dark Matter Constituents

The astronomical and cosmological evidence for dark matter discussed in sections
1.2.1 and 1.2.2 does not give a detailed indication of the properties of dark matter. It
tells only that dark matter must be massive, interact gravitationally, move non-
relativistically, and have total mass density of $\Omega_m$. A variety of candidates have
been proposed as the individual constituents of dark matter. These include black
holes that were created before light elements formed, massive neutrinos that do not
interact according to standard model weak interactions (“sterile” neutrinos), axions,
which are particles postulated to eliminate charge-parity (CP) violation in strong
interactions, and weakly interacting massive particles (WIMPs) [16]. Our main focus
will be on the WIMP, a general term which refers to a category of dark matter
candidates rather than any single particle in specific until the exact nature of the
WIMP, if the dark matter is composed of WIMPs, is discovered. Other candidates
(in addition to WIMPs) are discussed in more detail in [70].

The idea behind WIMPs is that they exist today as a “thermal relic” (see for
example [72], which we follow here). At some point the WIMPs were in thermal
equilibrium with the rest of the universe, and in this situation WIMP-antiWIMP pairs
were produced by interactions at some rate, and WIMPs and anti-WIMPs annihilated
into lighter particles at some rate. If thermal equilibrium had been maintained the
resulting equilibrium density of WIMPs today would be determined by Bose-Einstein
or Fermi-Dirac statistics of a particles at a temperature $T$, and the equilibrium number density of WIMPs would be $n_{\chi}^e \propto T^3$ at high temperatures and $n_{\chi}^e \propto (m_{\chi}T)^{3/2}e^{-m_{\chi}/T}$ at low temperatures. However, $n_{\chi}$ is changed also by the expansion of the universe (the increase of the scale factor $a$, see equations 1.69 and 1.71) and therefore is not necessarily equal to $n_{\chi}^e$. The Boltzmann equation describes the time rate of change of $n_{\chi}$,

$$\frac{dn_{\chi}}{dt} = -3Hn_{\chi} - \langle \sigma_A v \rangle [(n_{\chi})^2 - (n_{\chi}^e)^2],$$  

(1.74)

where $\sigma_A$ is the WIMP annihilation cross section and $v$ is the relative WIMP velocity, and $\langle \sigma_A v \rangle$ is the thermal average of their product. In this equation, the first term on the right describes the change in $n_{\chi}$ due to expansion, and the second and third (in square brackets) due to WIMP annihilation and creation respectively. Because $H \propto T^2$, at high temperatures, the term $-3Hn_{\chi} \propto T^5$ is much less in magnitude than the creation and annihilation terms which scale as $T^6$, and so expansion does not interfere significantly with thermal equilibrium. However, at low temperatures when $n_{\chi}$ begins to decrease exponentially, the time rate of change of the WIMP number density is completely dictated by expansion. The point at which expansion begins to dominate is called “freeze-out” and occurs when $Hn_{\chi} \sim \langle \sigma_A v \rangle n_{\chi}^2$ or

$$H \sim \langle \sigma_A v \rangle n_{\chi}.$$  

(1.75)

Figure 1.16 shows the development of $n_{\chi}$ (the “comoving number density”) as time increases (and temperature decreases). When the “freeze-out” condition is met, the $n_{\chi}$ deviates from its thermal equilibrium value and remains relatively constant. It can be shown that the density parameter $\Omega_c = (8\pi G/3H^2)\rho_{\chi}$, where $\rho_{\chi} = m_{\chi}n_{\chi}$ of WIMP cold dark matter can be written, independent of the value of $m_{\chi}$, as

$$\Omega_c h^2 = \frac{3 \times 10^{-27}}{\langle \sigma_A v \rangle} \text{cm}^3 \text{s}^{-1},$$  

(1.76)

where $h \approx 0.673$ is related to the current value of the Hubble constant $H$ as $h = H/(100 \text{ km s}^{-1} \text{ Mpc}^{-1})$. For weakly interacting particles, it turns out that the expected value of $\langle \sigma_A v \rangle \approx 10^{-25} \text{ cm}^3 \text{ s}^{-1}$, giving $\Omega_c h^2 \approx 0.03$, which is not far from the observed value (see equation 1.66). This consistency is known as the “WIMP miracle” and provides strong evidence that dark matter may interact weakly with ordinary matter and therefore that WIMPs give rise to the observed density $\Omega_c$.

### 1.2.4 Direct Detection of WIMP Dark Matter

If our galaxy is embedded in a dark matter “halo,” as one may infer from the observations discussed in section 1.2.1, particles of dark matter may be constantly moving through our world. If these particles are WIMPs and therefore interact with baryonic matter via the electroweak force, we may be able to observe such interactions and
Figure 1.16: (Figure from [72], originally from [73]) The equilibrium density of a WIMP plotted as the temperature of the universe decreases relative to the WIMP mass, along with the freeze-out densities for different values of $\langle \sigma_A v \rangle$. If the universe were not expanding, the WIMP density would follow the curve $N_{EQ}$, but due to freeze-out, the density instead remains relatively constant in time as the WIMP decouples from thermal equilibrium.

thereby “directly” detect dark matter on Earth. This approach involves building a detector with a massive sensitive region, shielding it from known interactions that would produce background events, and waiting to observe interactions that could only be attributed to dark matter. Other methods of detecting dark matter include “indirect” detection of some product of dark matter interactions (such as an excess of neutrinos coming from the center of the sun due to dark matter annihilations), and dark matter production as products of high-energy collisions of particles in accelerators such as the Large Hadron Collider (LHC). In studying high pressure xenon gas-based detectors, we are most interested in the direct detection of dark matter, and so we will now discuss this approach in more detail, mainly following source [74].

The signature of a WIMP produced in a dark matter detector is the recoil of a nucleus in the sensitive medium that occurs due to an elastic scatter of the dark matter particle. The kinetic energy of the recoiling nucleus is then deposited in the detector and measured. Even for a single type of dark matter particle moving with a fixed kinetic energy, a spectrum of nuclear recoil energies is produced. In actuality, the dark matter halo may contain many types of dark matter particles moving with a distribution of kinetic energies, thus producing an even more complex spectrum.
From [74], the observed spectrum of recoils in a direct detection experiment takes the form

\[
\frac{dR}{dE_R} = R_0 S(E_R) F^2(E_R) I, \tag{1.77}
\]

where \(R_0\) is the rate of recoils per unit target mass assuming a stationary earth and no upper velocity limit for dark matter particles contained in the halo, \(S(E_R)\) is the energy dependence incurred due to detector-specific effects and movement of the earth and the sun in the dark matter halo, \(F^2(E_R)\) is the nuclear form factor which gives the energy dependence due to the non-pointlike nature of the nucleus, and \(I\) is the interaction function which describes the effects of the WIMP interacting coherently with the entire nucleus which may or may not depend on spin. In the simplest case one would expect an exponentially decreasing spectrum

\[
\frac{dR}{dE_R} \sim e^{-E_R/E_0}, \tag{1.78}
\]

where \(E_0\) is some characteristic energy. Here we will consider each of the contributing factors individually with an idealized example of a dark matter particle of mass \(m_\chi = 39.6\) GeV/c\(^2\) in a halo with a Maxwellian velocity distribution with variance \(v_0\) scattering off of a Xe nucleus of mass \(m_{xe} = 127\) GeV/c\(^2\) with cross section \(\sigma_0\) in the limit of zero momentum transfer.

First, the base rate \(R_0\) is given in [74] as

\[
R_0 = \frac{503}{m_\chi m_{xe}} \left( \frac{\sigma_0}{1\ \text{pb}} \right) \left( \frac{\rho_D}{0.4\ \text{GeV/cm}^3} \right) \left( \frac{v_0}{230\ \text{km/s}} \right) \ d^{-1} \text{kg}^{-1}, \tag{1.79}
\]

where \(\rho_D\) is the local dark matter density and \(m_\chi\) and \(m_{xe}\) are expressed in GeV/c\(^2\). Note we have chosen \(m_\chi\) so that \(m_\chi m_{xe} \approx 5030\), and if we choose \(v_0 = 230\) km/s, \(\rho_D = 0.4\) GeV/cm\(^3\), and \(\sigma_0 = 10^{-9}\) pb, we have \(R_0 = 10^{-10}\) interactions/(kg\cdot day).

In the factor \(S(E_R)\) we will include the motion of the dark matter particles in a halo with infinite escape velocity and the detection efficiency. Assuming the elastic scattering is isotropic in the center-of-mass frame of the collision, the spectrum of recoil energies \(E_R\) integrated over the velocity distribution is given in [74] as

\[
\left. \frac{dR}{dE_R} \right|_{v_E, \infty} = R_0 \cdot \frac{\sqrt{\pi}}{16E_0} \frac{(m_\chi + m_{xe})^2 v_0}{m_\chi m_{xe}} v_E \left[ \text{erf} \left( \frac{v_{\min} + v_E}{v_0} \right) - \text{erf} \left( \frac{v_{\min} - v_E}{v_0} \right) \right], \tag{1.80}
\]

where \(v_E\), the velocity of the earth relative to the dark matter halo, appears here because it entered into the velocity distribution. \(E_0 = \frac{1}{2} m_\chi v_0^2\), \(v_{\min} = v_0 \sqrt{E_R (m_\chi + m_{xe})^2 / (4E_0 m_\chi m_{xe})}\) is the velocity corresponding to the minimum energy dark matter particle capable of producing a recoil with energy \(E_R\), and \(\text{erf}(x) \equiv \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} \, dt\). To apply the effects of a detector with a finite energy threshold, assume
that the signal used to trigger the detector consists of scintillation photons produced by the WIMP-induced nuclear recoils, and a photon detection system with efficiency $\epsilon = 0.1$ is in place. We can approximate the detection probability, given an energy deposition producing $N$ detectable photons, by a Gaussian distribution with mean $\mu = N\epsilon$ and variance $\sigma^2 = N\epsilon(1 - \epsilon)$. Given a calibration factor $\gamma$ equal to the number of photons produced per unit recoil energy, so that $E_R = N/\gamma$, and a detector threshold of $x_t$ photons corresponding to an average threshold recoil energy of $E_t = x_t/\gamma\epsilon$, the probability of detecting an event with recoil energy $E_R$ is equal to the probability of detecting $x > x_t$ photons. This is the same as $1 - (the\ probability\ to\ detect \leq x_t\ photons)$, giving a detection efficiency $\eta(E_R; E_t)$ of

$$
\eta(E_R; E_t) = 1 - \frac{1}{\sqrt{2\pi}\gamma E_R\epsilon(1 - \epsilon)} \int_{-\infty}^{\gamma E_t} \exp \left( \frac{-(x - \gamma E_R\epsilon)^2}{2\gamma E_R\epsilon(1 - \epsilon)} \right) dx.
$$

(1.81)

In addition, a finite detector resolution would affect the recoil spectrum at all energies, but we will not include this effect here. Our final factor of $S(E_R)$ will then be the quantities multiplying $R_0$ in equation 1.80 multiplied by the detector efficiency,

$$
S(E_R) = \frac{1}{R_0} \frac{dR}{dE_R} \bigg|_{v_E, \infty} \times \eta(E_R; E_t).
$$

(1.82)

The factor $F^2(E_R)$ is the nuclear form factor which accounts for the fact that the nucleus is not a single pointlike object but has physical size and composition. It accounts for the energy dependence of the cross section [74]

$$
\sigma(E_R) = \sigma_0 F^2(E_R),
$$

(1.83)

which always decreases with increasing transferred momentum to the nucleus as greater momentum transfer corresponds to a shorter deBroglie wavelength that essentially "sees" a smaller part of the nucleus for the purpose of interaction. For a static distribution of charge the form factor is a Fourier transform of the density $\rho(r)$ of the distribution [4], and here we use the approximation given in [74, 75] for a Gaussian $\rho(r)$

$$
F^2(qr_n) = e^{-q^2r_n^2/3},
$$

(1.84)

where $r_n = 1.2A^{1/3}$ fm is the approximate nuclear radius, $q = (2m_x eE_R)^{1/2}$ is the momentum transfer and we have now written $F^2$ in terms of the dimensionless (assuming $\hbar = 1$) quantity $qr_n$. The form factor given in equation 1.84 is shown in figure 1.17.

The final factor $I$ varies based on the type of interaction between the dark matter and the nucleus and whether or not it depends on their spins. If the interaction does not depend on the nuclear spin and is equivalent for each of the $Z + N = A$ protons
and neutrons in the nucleus, we will have $A$ of the same scattering amplitude, which when squared to obtain the cross section gives [74]

$$I = A^2,$$

(1.85)

assuming our cross section describes the interaction with only one of the $A$ nucleons and that the contributions add coherently (which will be the case for low momentum transfer $qr_n \ll 1$). Note that in this sense it is favorable to use a nucleus with high mass $A$ to search for WIMPs. The factor $I$ can be more complex, especially for spin-dependent interactions which have canceling contributions from paired nucleons with opposite spins and therefore only exist for nuclei with an odd number of protons or neutrons [74].

The recoil spectrum calculated using equations 1.77, 1.79, 1.80, 1.81, 1.82, and 1.84 is shown in figure 1.18 along with the detector efficiency used in the calculation. To illustrate the effect of Earth moving through the dark matter halo, also shown is the difference between the recoil spectra calculated using the minimum and maximum velocity of the Earth relative to the halo. We have used the approximation for this velocity [74]

$$v_E \approx 244 + 15 \sin(2\pi y) \text{ km/s},$$

(1.86)

where $y$ is the fraction of the year elapsed since March 2.

In a direct detection experiment, one would determine the rate $dR/dE_R$ in a given range of recoil energies, and with a thorough understanding of the above calculation...
Figure 1.18: (Middle) The nuclear recoil spectrum calculated from equation 1.77 via the methods discussed in this section. Note that $\sigma_0 = 10^{-9}$ pb was chosen to be the single-nucleon cross section, and the velocity of Earth was assumed to be the average velocity $v_E = 244$ km/s from equation 1.86. The detector efficiency calculated from equation 1.81 is shown (top) assuming $\gamma = 10$, $\epsilon = 0.1$, and $x_t = 5$. Also shown is the difference in recoil spectra calculated using the maximum and minimum velocities of Earth from equation 1.86 ($v_{\text{min}} = 229$ km/s, $v_{\text{max}} = 259$ km/s).
including the form factor, motion through the dark matter halo, and detector efficiency and resolution, one could calculate a minimum observable \( R_0 \) for a given combination of \( m_\chi \) and \( \sigma_0 \). These \((m_\chi, \sigma_0)\) values specify a limit curve above which all \((m_\chi, \sigma_0)\) combinations have been excluded by the experiment (see for example the curves shown in figure 1.20). The region of greatest sensitivity is typically found where \( m_\chi \) is approximately equal to the target mass, and the sensitivity decreases from that point at lower masses due to finite detector thresholds and at higher masses because assuming a constant \( \rho_D \), the flux of WIMPs falls as \( 1/m_\chi \) (since \( \phi_\chi = \langle v_\chi \rangle N_\chi = \langle v_\chi \rangle \rho_\chi/m_\chi \) [16].

1.2.5 WIMP Direct Detection Experiments

As is the case of neutrinoless double-beta decay searches, there are a variety of experiments attempting to directly detect dark matter using several different types of detectors. In all cases, the background must be low and so the experiment must be performed underground and away from significant radioactivity. It is also an advantage to be able to distinguish energy depositions due to nuclear recoils from those due to interactions of gammas and electrons to further reduce the number of background events. The way this is done is detector-dependent (if it is possible at all for a given detector), and in xenon it usually involves comparing the ratio of ionization to primary scintillation for a given energy deposition (see chapter 4).

Direct detection experiments have been performed and are in development using a number of strategies (see [16] and figure 1.19). Measuring energy spectra in a low background environment using detectors made of semiconductors such as Ge and Si or noble liquids such as xenon and argon have been popular experimental strategies, though materials such as NaI, CsI, and CaWO\(_4\) have also been used. Other strategies include the use of low-pressure gas to be able to image the track of a nuclear recoil and thereby potentially detect a directional dependence in WIMP interactions, and the use of superheated or supersaturated media that respond to very small energy depositions with the production of droplets or bubbles.

Though no widely accepted claim of direct dark matter detection has been established, there are several disputed claims and many limits have been set. The status of the field has been summarized in figure 1.20. Several experiments have claimed observation of a signal that could be attributed to WIMPs in the form of an excess of events in a given energy regime or a modulation of the observed rate due to the motion of the earth through the halo, though often these claims have been ruled out by one or more ultra-low background experiments that aim to identify WIMP interactions on an event-by-event basis. At some point dark matter searches will become so sensitive that they will encounter nuclear recoil events produced due to coherent neutrino interactions, at which point additional understanding of the dark matter signal (such as measurement of the direction of the nuclear recoil) will be necessary to prove a positive signal, as such a neutrino background cannot be shielded [77].
Figure 1.19: (Figure from [76]) The different detection strategies used in direct dark matter detection experiments and the associated experiments. Some experiments measure two different energy channels, and this often provides one way of discriminating electron and nuclear recoils.

Figure 1.20: (Figure from [77]) The current status of dark matter claims and limits in (mass, cross-section) parameter space, and the region of space at which coherent neutrino scattering will produce an unshieldable background, labeled as “WIMP discovery limit.”
Chapter 2

High Pressure Xenon Gas Detectors

Xenon gas has been studied as a detector medium for many years, and measurements over the past two decades characterizing high pressure xenon gas for the detection of gamma rays [78, 79, 80] have drawn attention to its potential for good energy resolution in rare physics searches. It is inert, relatively easy to purify, and provides both scintillation and ionization signals at room temperature. In the interest of rare physics searches, one isotope, \(^{136}\)Xe, is a candidate for \(0\nu\beta\beta\), and xenon nuclei are relatively high-mass and therefore benefit from the \(\sim A^2\) dependence of the spin-independent dark matter scattering cross section (see section 1.2.4). Here we discuss the general properties of xenon as a detector medium, focusing specifically on its energy resolution and scintillation and ionization signals.

2.1 Xenon Gas

Xenon is a heavy noble gas with \(Z = 54\) protons. Natural xenon is composed of the isotopes shown in table 2.1 with the stated abundances.

Though xenon is a noble gas, it is polarizable, and so the presence of one xenon atom can induce a dipole moment in another and thus create an electromagnetic interaction between the two. This produces an attractive force (known as the “van der Waals” force) between the two atoms which can be described along with the repulsive force of overlapping electrons due to the Pauli Exclusion Principle as the gradient of a Lennard-Jones potential

\[
U(r) = 4\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^{6} \right],
\]

(2.1)

where \(\sigma = 4.055\) Å and \(\epsilon = 229\) K (here \(\epsilon\) expressed in Kelvin should be multiplied by Boltzmann’s constant \(k\) to give units of energy) are used to describe xenon in [81]. Such an interaction gives rise to non-ideal behavior at high pressures, as shown in
Table 2.1: Relative abundances (%) of isotopes in natural xenon from [35].

<table>
<thead>
<tr>
<th>Isotope</th>
<th>Abundance (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>124</td>
<td>0.10</td>
</tr>
<tr>
<td>126</td>
<td>0.09</td>
</tr>
<tr>
<td>128</td>
<td>1.91</td>
</tr>
<tr>
<td>129</td>
<td>26.4</td>
</tr>
<tr>
<td>130</td>
<td>4.1</td>
</tr>
<tr>
<td>131</td>
<td>21.2</td>
</tr>
<tr>
<td>132</td>
<td>26.9</td>
</tr>
<tr>
<td>134</td>
<td>10.4</td>
</tr>
<tr>
<td>136</td>
<td>8.9</td>
</tr>
</tbody>
</table>

Figure 2.1. The nonideal behavior can be described by corrections to the ideal gas law $P = \rho kT$ in a virial expansion (see for example [82])

$$P = \rho kT(1 + B(T)\rho + C(T)\rho^2 + \ldots),$$

where $B(T)$ and $C(T)$ can be determined by fitting experimental data or by calculation/simulation assuming an interaction such as that of equation 2.1. As discussed later, properties of the xenon medium such as the electron mobility, W-value (energy required to produce an electron-ion pair), and energy resolution all change significantly at higher densities. The pressure range of greatest interest to this work is 10-20 bar, which corresponds to a density range of $\rho \approx 0.05 - 0.11$ g/cm$^3$ at $T = 293.15$ K (20° C), where xenon still largely behaves as an ideal gas and such changes in properties have not yet taken effect.

Xenon gas is relatively easy to purify, and commercial purifiers are available that can be used to remove impurities such as H$_2$O and O$_2$ to give electron drift lifetimes on the order of milliseconds.

2.2 Energy Loss in Xenon Gas

From the point of view of xenon as a detector medium, we are most interested in what occurs when a particle deposits its energy in xenon gas and what observable signals are produced as a result. Here we will focus on the detection of high-energy photons (such as gamma rays and x-rays), nuclear recoils, and high-energy electrons. In most cases we will discuss properties of xenon gas in the density regime for which it behaves like an ideal gas, which we call the “ideal regime.” We discuss the higher pressure regime in which the gas is effectively transitioning to the liquid phase in appendix B.
Gamma rays and x-rays interact in xenon in several ways [85], all resulting in the production of one or more high-energy electrons (and possibly a high-energy positron). A photon could transfer all of its energy to a single electron in a xenon atom via the process of photoelectric absorption or some its energy via Compton scattering, and that electron then continues through the xenon medium depositing its energy via electromagnetic interactions. If this ejected electron originated from a tightly-bound inner atomic shell, the xenon atom from which it came will be left in an excited state and will decay to emit a characteristic x-ray [37] which may also interact via the photoelectric effect or Compton scattering. In xenon, such an x-ray will have an energy near 30 keV (“K-shell” x-rays) or 5 keV (“L-shell” x-rays). A very high energy gamma ray may, in the electromagnetic field of the nucleus, produce an energetic electron along with its antiparticle the positron. The energy that went into producing the mass of the electron and positron will eventually reappear if the positron annihilates with another electron in the medium in the form of two 511-keV gamma rays.

Electrons lose energy in xenon primarily by ionizing and exciting xenon atoms as they traverse the medium, and some small fraction of energy is lost to elastic collisions with these atoms. Once at an energy below the lowest atomic excitation level, an electron loses the remainder of its energy via elastic collisions. The density of ionization is such that an electron of several hundred keV in kinetic energy can leave a track of ionization with total length on the order of centimeters at 10 bar pressure. These energetic electrons are known as “electron recoils.”

Nuclear recoils, unlike electron recoils, lose a large fraction of their energy to the
motion of other atoms. Because the xenon nucleus has the same mass as the other nuclei in the medium, transfer of energy in elastic collisions is much more efficient in a xenon nuclear recoil. To produce a nuclear recoil, some amount of kinetic energy is transferred to the nucleus of a xenon atom via interaction with, for example, a dark matter particle or a fast neutron. In this case the nucleus is the energetic particle that moves through the medium (dragging most of its electrons with it), ionizing and exciting xenon atoms. A much higher ionization density leaves a pointlike track for the purposes of imaging. Due to the extra energy loss in the form of heat, both the amount of ionization and scintillation produced per unit kinetic energy of the nuclear recoil are lower than the analogous quantities for electron recoils. In addition, the relative amounts of ionization and scintillation produced are different, and this particular difference can be used to differentiate between the two (see section 2.2.2).

### 2.2.1 Ionization and Scintillation from Electron Recoils

The physics of electron energy depositions is well understood in the density regime in which xenon behaves like an ideal gas. Figure 2.2 shows the cross sections for the various processes that can occur as an energetic electron traverses the xenon medium.

![Cross sections for electron interactions in pure xenon](image)

Figure 2.2: (Figure from [86]) Cross sections for electron interactions in pure xenon, including elastic collisions (el), ionizing collisions (ion), excitations (exc), and the total interaction cross section (T). The different point styles correspond to different sources from which the data was extracted (see [86] figure 2 for details).

An energetic electron leaves a “track” of ionized electrons in its path as it traverses the xenon medium. These electrons are normally collected, using an externally applied electric field to drift them through the xenon gas, and amplified into a macroscopic
signal proportional to the energy of the original electron. Ionization can occur only for electrons with kinetic energy above the ionization potential, which is 12.13 eV [85] for xenon gas in the ideal regime. Since an energetic electron and the electrons it ionizes can lose energy in several other channels including excitation and heat (elastic scattering), the amount of energy required to produce a single electron-ion pair, called the W-value, will be greater than the ionization potential. If the W-value is known, the number of electron-ion pairs \( N \) produced by an electron of energy \( E \) that deposits all of its energy in the xenon medium is \( N = \frac{E}{W} \). For electron recoils, the W-value was shown to vary with density and converge to a constant value at low densities [78], and the low-density value was assumed to be \( W = 21.9 \text{ eV} \) from [87].

Scintillation in gaseous and liquid xenon arises from excitation of the atoms but not from the direct decay of those excitations. For xenon gas densities greater than approx. 0.002 g/cm\(^3\) (\( \sim 0.4 \) bar pressure), it takes of order picoseconds after excitation for a triple collision between an excited atom and two other atoms to occur, resulting in the formation of an excited dimer [85]. The dimer then decays emitting a UV photon with a wavelength of approximately 170 nm. The entire process can be described as [85]

\[
\begin{align*}
(1) \quad e^- + A & \rightarrow A^*, \\
(2) \quad A^* + 2A & \rightarrow A^*_2 + A, \\
(3) \quad A^*_2 & \rightarrow 2A + h\nu.
\end{align*}
\]

The process is depicted in figure 2.3. Because the 170 nm photon did not originate from the xenon excited state directly, it will not be absorbed by individual xenon atoms and can escape the xenon medium. This process is what allows xenon and other noble gases to be effective scintillators. This is also the process responsible for the scintillation generated as a result of recombination.

![Diagram of scintillation process](image)

Figure 2.3: The standard process responsible for scintillation in noble gases. The decaying excited dimer emits a photon that cannot be recaptured by an individual xenon atom.

Recombination is the phenomenon by which an ionized electron encounters an ion in the xenon medium and is captured by it. In fact, by the time the electron
encounters the ion, the original ionized xenon atom may have already joined with another neutral xenon to create a charged dimer Xe$_2^+$. This original ion may have been its “parent” ion that was left behind when it was originally ionized, or the ion produced by a different ionization. Recombination proceeds, following ionization and thermalization of the resulting electron, as [88, 89]

\[
\begin{align*}
(1) & \quad e^- + A \rightarrow A^+ + 2e^- , \\
(2) & \quad A^+ + 2A \rightarrow A + A_2^+ , \\
(3) & \quad A_2^+ + e^- \rightarrow A^{**} + A , \\
(4) & \quad A^{**} + 2A \rightarrow A_2^{**} + A , \\
(5) & \quad A_2^{**} + A \rightarrow A^* + 2A , \\
(6) & \quad A^* + 2A \rightarrow A_2^* + A , \\
(7) & \quad A_2^* \rightarrow 2A + h\nu.
\end{align*}
\]

Note that in this process some excess excitation energy is converted to atomic motion (heat) in steps 4-5.

The amount of scintillation produced per unit deposited energy of an energetic electron can also be quantified by a scintillation $W$-value, $W_{sc}$. In principle this quantity is reported as if no external field is applied in the region of gas where the energy is deposited, and it is assumed that all ionization produced recombines. Measurements of this value have been inconsistent up to this point and include $W_{sc} = 76 \pm 12$ eV [90] (actually performed with a 90%Xe, 10%He mixture, and the applied field in the region of interaction is unclear) obtained using 60 keV x-rays from a $^{241}$Am source, and $W_{sc} = 111 \pm 16$ eV [89] obtained using 5.9 keV x-rays from a $^{55}$Fe source at 800 torr pressure (the field in the region of interaction was determined to affect the scintillation yield only slightly). $W_{sc} \sim 28$ eV is also given in [85] but was measured with relativistic heavy ions rather than an electron or gamma source. In chapter 4, we find a value of $W_{sc} = 52.0 \pm 7.8$ eV by observing primary scintillation produced by 662 keV gamma rays in $\sim$ 14 bar xenon gas in which a field of about 370 V/cm was present in the interaction region.

### 2.2.2 Ionization and Scintillation from Nuclear Recoils

The details of ionization and scintillation production due to nuclear recoils in gaseous xenon have not been thoroughly studied in the past.\(^1\) Cross sections of ionization, elastic, and excitation processes for energetic xenon ions incident on neutron xenon atoms are not known as precisely as those of energetic electrons incident on neutral xenon, though some experimental data [91] and theoretical estimates exist. In [92], Lindhard uses approximations for electronic stopping cross sections and differential

\(^1\)From this point forward we use the term “nuclear recoil” to refer specifically to xenon nuclei recoiling in a xenon medium.
nuclear stopping cross sections\(^2\) in an integral equation formalism to determine the amount of energy transferred to electrons and nucleons (atomic motion). A good approximation to Lindhard’s results can be given by a semi-empirical formula which states that the fraction \(f_n\) of the total recoil energy \(E\) transferred to electrons by a nuclear recoil of the same element and atomic (\(Z\)) and mass (\(A\)) numbers as those of the medium is given by [74, 92]

\[
\frac{f_n}{k g(\epsilon)} = 1 + k g(\epsilon),
\]

(2.3)

where \(k = 0.1332^{2/3} A^{-1/2}\), \(\epsilon = 11.5EZ^{-7/3}\) (with \(E\) in keV), and \(g(\epsilon)\) is a function that can be described approximately by 

\[
g(\epsilon) = 3\epsilon^{0.15} + 0.7\epsilon^{0.6} + \epsilon.
\]

Figure 2.4: The Lindhard factor \(f_n\) as a function of nuclear recoil energy for xenon (\(Z = 54\)) assuming the average atomic weight \(A = 131.293\) [16].

Lindhard’s theory gives one explanation for the reduced amount of detectable electronic energy deposition (“quenching”) produced by a nuclear recoil relative to that produced by an electron recoil of the same energy. However, it does not describe how much of that energy appears as excitation (and eventually scintillation) and ionization. For electron recoils in gaseous xenon, the number of excitations vs. ionizations can be determined via Monte Carlo methods using the cross sections shown in figure 2.2, but in the presence of recombination this does not map directly to the amount of

\(^2\)The stopping cross section is the quantity which when multiplied by the number of atoms per unit volume gives the energy lost by the recoiling nucleus per unit distance traveled in the medium.
scintillation and ionization produced. In low pressure gas and in the presence of a high electric field, recombination is minimal for electron recoils, but as pressure increases and the external field decreases, recombination will convert an increasing amount of ionization into scintillation (see for example figure 1 of [78]). Recombination should be even stronger for nuclear recoils which have a relatively shorter range for a given energy and therefore produce more closely spaced ionizations. In either case, the ratio of ionization to scintillation produced by nuclear recoils is found experimentally to be lower than that for electron recoils for a given recoil energy in both liquid (see figure 2.5) and gaseous (see chapter 4) xenon.

Figure 2.5: (Figure from [93]) The ratio of the observed ionization to primary scintillation signals in liquid xenon observed with a dual-phase xenon TPC plotted against recoil energy. As discussed in section 2.2.3, the ionization electrons can be amplified by a process called electroluminescence to produce a proportional number of photons, a signal called S2, while the detected primary scintillation photons are denoted as S1. Neutrons from an AmBe neutron source with energies of several MeV were used to produce the nuclear recoils, and the scintillation signal was calibrated assuming the ratio of observed nuclear recoil energy to electron recoil energy is $L_{\text{eff}} = 0.0984E_r^{0.169}$ and correcting for effects of the applied electric field on scintillation yield. Note the clear presence of a band of events at low energies with a significantly lower S2/S1 ratio.
2.2.3 Electroluminescence

The scintillation properties of xenon can also be used as a means of signal amplification in a process called electroluminescence. Ionization produced by incident radiation can pass through a region of high field, still within the xenon medium, that accelerates the electrons enough to excite but not ionize the xenon atoms. That is, electrons are capable of gaining energy beyond the 8.3 eV first excited level of xenon but lose this energy in an atomic excitation before they are capable of being accelerated beyond the $\sim 12$ eV ionization potential. The result is a number of xenon excitations (and therefore UV photons) produced per electron proportional to the length of high field region the electron traversed. This number of UV photons $g$ has been determined to follow \[94\]

$$ g = 140(E/p - 0.83)px, $$

(2.4)

where $E$ is the applied electric field in kV/cm, $p$ is the gas pressure in bars, and $x$ is the length in cm of the high field region. This “proportional scintillation” is not subject to the large fluctuations in the number of produced quanta that are present in electron avalanches, and therefore it serves as a gain mechanism with energy resolution that is superior to methods based on electron multiplication. Energy resolution in detectors using electroluminescent gain is discussed in more detail in section 2.3.3.

2.3 Energy Resolution

The precision with which a detector can determine the energy of an interaction is a key factor in determining its suitability for a specific application. In xenon detectors based on ionization and scintillation signals, it is determined by several factors including the collection efficiency of light and charge, the signal readout process including any amplification, and the physics of the ionization and scintillation processes. In gaseous xenon in the ideal regime, one is most likely to derive energy information from the ionization signal due to the ability to engineer near 100% charge collection efficiency and relative difficulty to engineer good light collection efficiency (often in the several percent regime). Several studies, for example \[78, 79, 95, 96, 97, 98\], indicate that the properties of xenon as a detection medium change in a way that is not fully understood when the density of the xenon medium is increased beyond that of a dilute gas. Amongst the most significant of such changes is a degradation in energy resolution, as is demonstrated clearly in figure 2.6 from reference \[78\]. This issue is discussed in more detail in appendix B. Because the operating densities in the experimental results discussed in chapters 3 and 4 are much lower than those at the onset of this effect, here we will discuss energy resolution at a more general level and develop formulas that can be applied directly to these results.

Fano \[99\] determined that in a gaseous detector, the number of electron-ion pairs
Figure 2.6: (Figure from [78]) Energy resolution vs. density for xenon gas. The energy resolution at constant temperature begins to degrade with increasing density beginning at $\rho \approx 0.55 \text{ g/cm}^3$.

produced by incident radiation with a fixed energy fluctuates about some mean $N$ as

$$\sigma_N^2 = FN,$$  

(2.5)

where $F$ is known as the Fano factor. Though $F = 1$ corresponds to a Poisson variance, typical values of $F$ are much smaller. For xenon gas, $F = 0.14$ [78]. We now discuss how this and other factors together contribute to the energy resolution of a gaseous detector.

2.3.1 General Formalism

We will now develop a general formalism for describing the energy resolution in particle detectors in the presence of several mechanisms that together are ultimately responsible for the final signal read out from the detector. We assume that the detection process consists of several steps, each containing an input signal $x_i$ and an output signal $x_{i+1}$, and each characterized by a distribution function $G_i(x_i, x_{i+1})$, all except the first step which contains only an output $x_0$, and is described by a distribution function $G_0(x_0)$. Each distribution function gives a mean and sigma of the output
given a specific value of the input,

\[
\mu_{i+1}(x_i) \equiv \langle x_{i+1} \rangle_{G_i} = \int x_{i+1} G_i(x_i, x_{i+1}) dx_{i+1}
\]

and for each input, \( G_i(x_i, x_{i+1}) \) is a normalized distribution,

\[
\int G_i(x_i, x_{i+1}) dx_{i+1} = 1.
\]

For example, in a detector system often ionization produced by an energetic particle is collected, and the collection process is described by some statistical distribution which yields a number of collected electrons (or some signal proportional to this quantity) \( y = x_2 \). This process is described by a distribution that depends on the number of ionized electrons produced initially \( n = x_1 \). For each \( n \), this distribution is normalized. However, since \( n \) can fluctuate, and \( y \) can fluctuate for a given \( n \), we need to determine a distribution \( G'(y) \) of the collected electrons taking into account all possible values of \( n \). The final distribution is one in which each \( G'_i(n, y) \) is weighted by \( G_0(n) \), \( G'(y) = \int G_0(n) G_1(n, y) dn \), or in the more general case,

\[
G'_i(x_{i+1}) = \int G'_{i-1}(x_i) G_i(x_i, x_{i+1}) dx_i \quad (i > 0).
\]

This recursive formula must terminate at \( i = 1 \), so the initial input distribution \( G'_0(x_1) \) is postulated and not calculated from a previous one. One can see by integrating \( G'_i \) over all \( x_{i+1} \) and using equation 2.7 that \( G'_i \) is normalized to 1 given that \( G'_{i-1} \) is normalized to 1, and therefore as long as \( G_0 \) is normalized all distributions \( G'_i \) are normalized.

The quantities of interest are the mean \( \mu'_{i+1} \) and standard deviation \( \sigma'^2_{i+1} \) under the distribution \( G'_i(x_{i+1}) \). We compute these by first computing that \( \langle x_{i+1} \rangle_{G'_i} \) and \( \langle x^2_{i+1} \rangle_{G'_i} \),

\[
\langle x_{i+1} \rangle_{G'_i} = \int x_{i+1} G'_i(x_{i+1}) dx_{i+1}
\]

\[
= \int G'_{i-1}(x_i) \left( \int x_{i+1} G_i(x_i, x_{i+1}) dx_{i+1} \right) dx_i
\]

\[
= \int G'_{i-1}(x_i) \mu_{i+1}(x_i) dx_i
\]

\[
= \langle \mu_{i+1} \rangle_{G'_{i-1}}
\]
and using, in terms of the quantities defined in equation 2.6, that \( \langle x_{i+1}^2 \rangle G_i = \sigma_i^2(x_i) + \mu_{i+1}^2(x_i) \),

\[
\langle x_{i+1}^2 \rangle G_i = \int x_{i+1}^2 G_i(x_{i+1}) dx_{i+1} \\
= \int G_i'(x_i) \left( \int x_{i+1}^2 G_i(x_i, x_{i+1}) dx_{i+1} \right) dx_i \\
= \int G_i'(x_i) [\sigma^2_{i+1}(x_i) + \mu_{i+1}^2(x_i)] dx_i \\
= \langle \sigma^2_{i+1}(x_i) + \mu_{i+1}^2(x_i) \rangle G_i'_{i-1},
\]

we can write

\[
\mu_{i+1}' = \langle \mu_{i+1} \rangle G_{i-1}', \\
\sigma^2_{i+1}' = \langle (x_{i+1} - \langle x_{i+1} \rangle G_{i+1}^2) \rangle G_i' \\
= \langle x_{i+1}^2 \rangle G_i' - \langle x_{i+1} \rangle G_i'^2 \\
= \langle \sigma^2_{i+1} + \mu_{i+1}^2 \rangle G_{i-1}' - \langle \mu_{i+1} \rangle G_{i-1}'^2 \\
= \langle \sigma^2_{i+1}(x_i) \rangle G_{i-1}' + \langle (\mu_{i+1}(x_i) - \langle \mu_{i+1}(x_i) \rangle G_{i-1}')^2 \rangle G_{i-1}'.
\]

Equation 2.11 tells us how to obtain the variance on an output signal \( x_{i+1} \) under a process \( G_i(x_i, x_{i+1}) \) given an input signal \( x_i \) distributed according to some previous process \( G_{i-1}(x_i) \). We will see in the following examples that if we can write expressions such as \( \langle \sigma^2_{i+1}(x_i) \rangle G_{i-1}' \) in terms of \( \mu_i' \), then we can compute each \( \mu_i' \) by knowing the previous \( \mu_i' \), and thus recursively propagate error through multiple processes.

We will now discuss the special case in which \( \mu_{i+1}(x_i) \) and \( \sigma^2_{i+1} \) are linear in \( x_i \), so that \( \mu_{i+1} = kx_i \) and \( \sigma^2_{i+1} = \lambda x_i \) where \( k \) and \( \lambda \) are constants. This will occur if a process \( G_i \) takes some integer number of inputs and acts on each one in the same way regardless of the total number. Then the first term in equation 2.11 is

\[
\langle \sigma^2_{i+1}(x_i) \rangle G_{i-1}' = \lambda \langle x_i \rangle G_{i-1}' = \lambda \mu_i(x_{i-1}),
\]

and the second term is

\[
\langle (\mu_{i+1}(x_i) - \langle \mu_{i+1}(x_i) \rangle G_{i-1}')^2 \rangle G_{i-1}' = k^2 \langle (x_i - \langle x_i \rangle G_{i-1}')^2 \rangle G_{i-1}' = k^2 \sigma_i^2.
\]

Noting that \( \mu_{i+1}' = k \langle x_i \rangle G_{i-1}' = k \mu_i' \), we can write

\[
\left( \frac{\sigma_{i+1}'}{\mu_{i+1}'} \right)^2 = \left( \frac{\sigma_{i+1}(x_i)}{\mu_{i+1}(x_i)} \right)^2 + \left( \frac{\sigma_i'}{\mu_i'} \right)^2.
\]
where \( \bar{x}_i = \langle x_i \rangle_{G_i} \) is the average value of the input signal. Equation 2.14 states that the contribution to the energy resolution by a process that acts on an input signal \( x_i \) and has a variance \( \sigma(x_i) \) and mean \( \mu(x_i) \) is given by adding \( \sigma/\mu \) of that process alone, evaluated at the mean value of the input parameter, in fractional quadrature with that of the previous processes\(^3\). If all statistical processes in a detector system through which signals must pass during the detection process are such that the process acts on each input signal in the same way regardless of the total number of inputs, the final energy resolution of the detector is the quadrature sum of \( \sigma/\mu \) of each process evaluated at the mean value of their input signals. In this case, the process and its preceding input process are statistically independent.

### 2.3.2 Energy Resolution in Negative Ion TPCs

We will now apply the formalism developed in section 2.3.1 to a detection scheme simplified to two statistical processes. First, \( x_1 \) electrons are ionized according to a distribution \( G'_0 \) that yields Fano statistics, and next \( x_2 \) electrons are recovered from a binomial loss process \( G_1 \) with loss probability \( \eta \). We write down the relevant quantities from equation 2.6 for a binomial distribution

\[
\mu_2(x_1) = x_1(1 - \eta) \\
\sigma_2^2(x_1) = x_1\eta(1 - \eta)
\]

Now using equation 2.11, and defining \( \bar{x}_i = \langle x_i \rangle_{G_i} \),

\[
\sigma_2^2 = \bar{x}_1\eta(1 - \eta) + (1 - \eta)^2 \langle (x_1 - \bar{x}_1)^2 \rangle_{G'_0} \\
= \bar{x}_1\eta(1 - \eta) + (1 - \eta)^2 F\bar{x}_1,
\]

where we have used Fano statistics [99] in the second step. Therefore, calling \( \bar{n} = \bar{x}_1 \) the number of electron-ion pairs produced, the energy resolution is

\[
\frac{\Delta E}{E} = \sqrt{\frac{\sigma^2_2}{\mu^2_2}} = \sqrt{\frac{1}{\bar{n}} \left( (F - 1) + \frac{1}{1 - \eta} \right)}.
\]

Note that we could have also added a binomial process with mean \( \mu_2 = \bar{x}_1(1 - \eta) \) in fractional quadrature with the Fano process with mean \( \mu_1 = \bar{x}_1 \) and obtained the same answer, since both \( \mu_2 \) and \( \sigma_2 \) from equation 2.15 are linear in \( x_1 \).

\(^3\)This simplification was first noted by Peter Sorensen at LLNL in the case of the resolution for the Negative Ion TPC discussed in section 2.3.2.
2.3.3 Energy Resolution in Electroluminescent Xenon TPCs

We will first describe the several processes that will contribute to the energy resolution of an EL TPC. Each will be statistically independent, so the final energy resolution will be quantified by the quadrature sum of each value of $\sigma/\mu$.

Production of Ionization: As in section 2.3.2 this process is governed by Fano statistics, so given a mean number of electron/ion pairs produced $\bar{n}$, we have $\sigma_n^2 = F\bar{n}$. Therefore the contribution to the energy resolution will be

$$\left( \frac{\sigma_n}{\bar{n}} \right)_{\text{Fano}}^2 = \frac{F}{\bar{n}}.$$ (2.18)

Production of EL Gain: Similar to the Fano distribution, for a single electron traversing the EL gap, the fluctuation in the number of EL photons $g$ produced follows $\sigma_g^2 = J_{CP}\bar{g}$, where $J_{CP}$ is the Conde-Policarpo factor [100]. Thus for $\bar{n}$ electrons, the fluctuation is $J_{CP}\bar{g}\bar{n}$. The contribution to the energy resolution is then

$$\left( \frac{\sigma_{EL}}{\bar{g}} \right)_{\text{EL}}^2 = \frac{J_{CP} \bar{n}}{\bar{g}}.$$ (2.19)

Photon Detection Efficiency: Given $x$ photons produced, we model the detection process as a random event with probability $\varepsilon$ of success. For a binomial process $\mu = \varepsilon x$ and $\sigma^2 = \varepsilon(1 - \varepsilon)x$. Therefore we have, for $x = \bar{n}\bar{g}$, the contribution to the energy resolution

$$\left( \frac{\sigma_{\text{det}}}{\bar{n}\bar{g}} \right)_{\text{eff}}^2 = \frac{\varepsilon(1 - \varepsilon)\bar{g}}{(\varepsilon\bar{n}\bar{g})^2} = \frac{1}{\bar{n}\bar{g}} - \frac{1}{\varepsilon\bar{n}\bar{g}}.$$ (2.20)

PMT Gain: For the detection of a photon that has produced a single photoelectron (SPE) in a PMT, the integrated response yields an SPE charge distributed about a mean $\bar{q}$ with variance $\sigma_q^2$. For $x$ detected photons, we have mean $x\bar{q}$ and variance $x\sigma_q^2$. Therefore for $x = \varepsilon\bar{n}\bar{g}$, the contribution to the energy resolution is

$$\left( \frac{\sigma_{\text{PMT}}}{\varepsilon\bar{n}\bar{g}} \right)_{\text{PMT}}^2 = \frac{\varepsilon\bar{n}\bar{g}\sigma_q^2}{(\varepsilon\bar{n}\bar{g}\bar{q})^2} = \frac{1}{\varepsilon\bar{n}\bar{g}} \left( \frac{\sigma_q}{\bar{q}} \right)^2.$$ (2.21)

Now adding all of the above contributions and defining the average number of electrons produced in the initial ionization as $n_{\text{ion}} = \bar{n}$, the number of photons produced in the EL process as $n_{EL} = \bar{n}\bar{g}$, and the number of photons detected as $n_{\text{det}} = \varepsilon\bar{n}\bar{g}$, we have the final expression for the energy resolution of an EL TPC

$$\left( \frac{\Delta E}{E} \right) = \sqrt{\frac{F}{n_{\text{ion}}} + \frac{J_{CP} - 1}{n_{EL}} + \frac{(\sigma_q/\bar{q})^2 + 1}{n_{\text{det}}}}.$$ (2.22)
Note that this does not include effects due to attachment in the drift region and attachment and multiplication in the EL gap. We will consider these effects in section 2.3.4.

2.3.4 Electroluminescent Gain, Charge Attachment and Multiplication

Here we will consider the EL gain process in more detail in the presence of attachment or multiplication. EL gain of a single electron traveling a distance \( x \) in xenon at a pressure \( p \) (bar) in a uniform electric field \( E \) (kV/cm) follows equation 2.4 which we rewrite here as

\[
g(x) = 140(E/p - 0.83)px \equiv kx,
\]

where we have assumed a constant \( E \) and \( p \) and defined \( k \equiv 140(E/p - 0.83)p \). With no attachment or multiplication, every electron produces a gain of exactly \( g(d) = kd \), where \( d \) is the length of the EL gap. We will now consider the energy resolution when the gain is influenced by attachment in the EL gap, or by electron avalanche multiplication.

Attachment

The attachment process can be characterized by an attachment coefficient \( \alpha_a \), which we will define as follows. For \( n_0 \) electrons entering the EL gap, the number remaining which have not attached after drifting a distance \( x \) is \( n(x) = n_0e^{-\alpha_a x} \). Thus the number of electrons that attached in an interval \( dx \) after drifting a distance \( x \), \(-dn/dx)dx\), divided by the initial number of electrons, is a differential probability of attachment.

\[
p(x)dx = -\frac{1}{n_0} \frac{dn}{dx} dx = \alpha_a e^{-\alpha_a x} dx.
\]

Note that since for \( x > d \), the electron has escaped attachment completely, we can calculate the probability of escape \( p_e \) as

\[
p_e = \int_d^\infty \alpha_a e^{-\alpha_a x} dx = e^{-\alpha_a d}.
\]

Since the EL gain of each electron is directly proportional to its drifted distance, we can turn \( p(x) \) into a probability distribution of gains by changing variables with \( x = g/k \),

\[
p(g)dg = \frac{\alpha_a}{k} e^{-(\alpha_a/k)g} dg.
\]
We now consider a distribution consisting of \( p(g) \), the probability that \( g \) lies in an interval \([g, g + dg]\), and \( e^{-\alpha_d} \), the probability of gain \( g = kd \). The mean of \( g \) and \( g^2 \) are given by

\[
\bar{g} = \int_0^{kd} g \frac{k}{\alpha_a} e^{-(\alpha_a/k)g} dg + kde^{-\alpha_a d} = \frac{\alpha_a}{k} (1 - e^{-\alpha_a d})
\]

and

\[
\bar{g^2} = \int_0^{kd} g^2 \frac{\alpha_a}{k} e^{-(\alpha_a/k)g} dg + (kd)^2 e^{-\alpha_a d} = 2 \left( \frac{k}{\alpha_a} \right)^2 (1 - (1 + \alpha_a d)e^{-\alpha_a d}),
\]

and so the variance on \( g \) is

\[
\sigma^2_g = \bar{g^2} - \bar{g}^2 = \left( \frac{k}{\alpha_a} \right)^2 (1 - 2\alpha_a de^{-\alpha_a d} - e^{-2\alpha_a d}) = 2 \left( \frac{k}{\alpha_a} \right)^2 e^{-\alpha_a d} [\sinh \alpha_a d - \alpha_a d].
\]

Noting that \( \bar{g}^2 = (k/\alpha_a)^2 (1 - e^{-\alpha_a d})^2 = 2(k/\alpha_a)^2 [\cosh \alpha_a d - 1] \), we can write the fractional variance as

\[
\left( \frac{\sigma_g}{\bar{g}} \right)^2 = \frac{\sinh \alpha_a d - \alpha_a d}{\cosh \alpha_a d - 1}.
\]

In the case of zero attachment \( \alpha_a d \to 0 \), we have 0 fluctuation in the EL gain process due to attachment. However, in the case of extreme attachment \( \alpha_a d \to \infty \), \( \sigma_g/\bar{g} \to 1 \); despite any amount of attachment, the fractional variance per electron can only be at worst equal to 1. This is a result of the fact that even if an electron has attached, it still performs EL up to the point of attachment. If this were not the case, and we only record gain from an electron that traversed the entire gap, we would have \( \bar{g} = kde^{-\alpha_a d} \) and \( \bar{g^2} = (kd)^2 e^{-\alpha_a d} \). Therefore

\[
\left( \frac{\sigma_g}{\bar{g}} \right)^2_{\text{EL gap only}} = \frac{\bar{g^2} - \bar{g}^2}{\bar{g^2}} = e^{\alpha_a d} - 1.
\]

As one can see from figure 2.7, the ability of an electron to produce EL gain up to the point of attachment is an advantage to the energy resolution of an EL-based detector.

**Charge Multiplication**

Now assume that a single electron may ionize electrons while traversing the EL gap, such that in a distance \( dx \), the probability of ionization is \( \alpha_m dx \). We will consider the multiplication and EL processes separately, in two steps:

1. \( M \) electrons are created by an electron avalanche in the EL gap.

2. For each of these \( M \) electrons, an EL gain \( g = k(d - x) \) is produced, where \( x \) is the location at which the electron was created in the avalanche.
First we will analyze step 2. Given a particular avalanche gain $M$, we have $M$ electrons at $x = d$, and defining a new coordinate $y = (d - x)$, so that $dy = -dx$, the number of electrons present in the avalanche at a location $y$ decreases as $dm = -m(y)\alpha_m dy$. Therefore we have the distribution of electrons

$$m(y) = Me^{-\alpha_my}.$$ (2.31)

It is this number of electrons $-dm = M\alpha_me^{-\alpha_my}dy$ that will perform EL over the distance $y$. Therefore, the probability that a single electron performs EL over a distance $y$ is $-dm/M = \alpha_me^{-\alpha_my}dy$. This is the same problem as the one we solved in the attachment case, now with coefficient $\alpha_m$. The variance in EL gain for a single electron in the avalanche will therefore be given by equation 2.29.

For the avalanche process, one must assign a statistical distribution with mean $\bar{m}$ and variance $\sigma_m^2$. For example, for an exponential distribution on $m$, $p(m) = 

\footnote{Note that this will only hold in the case that $m \gg 1$, as for small $m$ we cannot accurately make $m$ a continuous variable when describing a discrete number of electrons.}
We have $\sigma_m^2 = \bar{m}^2$. In [101], it is shown that $(\sigma_m/\bar{m})^2 = f(\bar{m})$, that is the fluctuations in the avalanche process are not necessarily as strong as those in the case of the exponential distribution. If we assume the avalanche and gain processes are statistically independent, we have $m$ electrons produced according to the avalanche statistics that each undergo an EL process with (apparent) attachment. Adding the relative variances from both processes, the final relative variance on the mean number of EL photons $\mu = \bar{m}(k/\alpha_m)(1 - e^{-\alpha_md})$ is

$$
\left(\frac{\sigma}{\mu}\right)^2 = \frac{1}{\bar{m}} \left[ \frac{\sinh \alpha_md - \alpha_md}{\cosh \alpha_md - 1} \right] + f(\bar{m}). \quad (2.32)
$$

Note that an increase in gain produces a larger number of EL photons, which should improve the energy resolution according to equation. However, once the avalanche fluctuations become large (i.e., $f(\bar{m})$ begins to dominate), the resolution ultimately degrades.

### 2.4 Summary

We have briefly described some of the basic properties of gaseous xenon as a detector medium. In general, the use of xenon as a detector medium relies on the measurement of the light and charge produced when incident radiation creates electron and nuclear recoils in the gas. In chapter 3, we present a study demonstrating the energy measurement and tracking capabilities of xenon gas for detecting particles that produce electron recoils. In chapter 4 we examine the response of xenon gas to nuclear recoils.
Chapter 3

NEXT-DBDM High Pressure Xenon TPC

3.1 Introduction

The NEXT-DBDM (Double Beta, Dark Matter) TPC was constructed as a prototype for the NEXT-100 experiment primarily to study energy resolution in high pressure xenon. Later studies in silicon photomultiplier (SiPM) tracking and performance with gas additives were carried out, as well as attempts to characterize neutron elastic scattering (see chapter 4). All in all the TPC is an excellent demonstration of the capabilities of xenon gas-based detectors. Primary scintillation and ionization amplified by electroluminescent gain are observed with 19 photomultiplier tubes. The pressure vessel typically holds of order 1 kg of xenon gas and operates at $\sim 10$-15 bar pressure and at room temperature, and the active volume is large enough to contain tracks of several centimeters in length. Here we describe the operational principles of the detector, details of the hardware and data acquisition, and results obtained from observations of electron recoils. Note that some content in the SiPM tracking sections in this chapter appears in the NEXT report to the Laboratorio Subterráneo de Canfranc (LSC), Nov. 2012.

3.2 Basic Principles of Operation

A high-pressure electroluminescent TPC allows for both precise energy measurement and accurate determination of the location of the interaction along the z-axis (parallel to the drift field). With the addition of a tracking plane, position information in the x-y plane can also be obtained. The layout of the active area in the TPC is shown in figure 3.1. The detector itself is sealed in a pressure vessel and filled with xenon at some operating pressure (10-15 bar in the studies presented here). An array of 19 PMTs observes the active area which is divided into several smaller regions by wire
Figure 3.1: Summary of the key components of NEXT-DBDM. The TPC is divided by wire meshes into several regions including the drift region R2 where the detectable interactions occur, the EL (gain) region R3, and two gas buffer regions. High voltages placed on the meshes establish the electric fields in the different regions. PMTs observe the drift and EL regions, an SiPM plane may be placed behind the EL region in R4, and teflon reflectors surround the drift region. A teflon reflector may also be stationed behind buffer region R4 and is only relevant in the absence of the SiPM plane.

The basic flow of operation for a typical event is shown in figure 3.2. Particles interact in the drift region and deposit their energy, producing scintillation and ionization. Some fraction (several percent at best) of the total scintillation yield produced is detected immediately by the PMTs, and the resulting signal is denoted as “S1.” The ionization is transported by the applied electric field to the EL region in which each electron creates a number of photons proportional to the distance through which it drifted and that is dependent on the magnitude of the applied field (see section 2.2.3). Again some fraction of these EL photons is detected by the PMTs, and the
resulting signal is denoted as “S2.”

Figure 3.2: General detection scheme in NEXT-DBDM. (1) Incident radiation (for example, a gamma ray) interacts in the sensitive region and creates both a track of ionization and primary scintillation. The primary scintillation is immediately detected by the PMTs and constitutes the S1 signal. (2) The ionization is drifted in the applied electric field to the EL plane. (3) Electroluminescence is produced by each electron traversing the EL gap and detected by the PMTs and constitutes the S2 signal.

The digitized waveform constructed from the sum of the 19 PMT signals for a typical electron recoil event is shown in figure 3.3. The S1 signal is a small pulse
of light denoting the initial interaction, and the S2 signal a larger pulse exhibiting features dependent on the geometry of the track. In this detection scheme, the energy is manifest in the total amount of S2 light produced (the integral of the S2 pulse), and therefore a high-energy electron recoil of a fixed energy can give rise to a wide variety of S2 signals, but still the same detected energy.

Figure 3.3: (Top) A typical event in NEXT-DBDM. (Bottom) The same waveform enlarged in the region of S1.
3.3 TPC Hardware and Design

Figure 3.4: The NEXT-DBDM TPC and gas circulation/purification system during operation.

3.3.1 Pressure Vessel and Gas System

During operation the detector and xenon gas are contained within a stainless steel pressure vessel with a volume of approximately 10 L rated to a maximum pressure of 17 atm. The vessel has several sideports used for gas input/output and for connection of instruments to monitor temperature and pressure. The detector hardware is attached via PEEK rods to the lid of the pressure vessel, and the entire assembly (lid + detector) can be attached via two metal pegs to sliding metal supports that provide a straightforward means of transporting the detector into and out of the vessel. The lid is sealed to the vessel with a copper gasket and 24 bolts tightened to a torque of 500 ft-lb. A stainless steel octagon containing 8 small flanges capable of adaptation to various inputs and outputs (for example, for gas flow and signal connections) is attached to the lid at the end of a long steel tube, and all signals and PMT high voltage cables are fed down the tube and connected via ports on the octagon. A narrow opening of diameter 1.7 cm extends through the octagon and down the center of the steel tube and leads to a thin steel window to the inside of the pressure vessel so that a collimated source can be directed down the hole to emit gammas axially into the sensitive region of the detector. The gas system is shown in figure 3.5.
Figure 3.5: (Figure from [102], created by Tom Miller, LBNL) The gas system for NEXT-DBDM.

3.3.2 Detector

The detector itself consists of a PTFE (teflon) cover, the 19-PMT array and its electronics, a field cage and the wire meshes dividing it into different functional regions, an SiPM plane, and a teflon reflector placed behind the electroluminescent region. In different operating configurations, the SiPM plane and/or teflon reflector may be removed, and the field cage may be modified. The field cage is made up of 18 rectangular teflon panels on which copper strips are mounted and connected by resistors to grade the field between the wire meshes. The wire meshes are made from crossed wires spaced with a pitch of 0.5 mm\(^1\), and each is supported by stainless steel frames consisting of two steel hexagons that clamp down on the mesh. The tension on the mesh can be controlled via screws on the inside perimeter of the frames. Electrical connection between the panels and the wire meshes is provided by fastening metal contacts on the panels to the mesh frames with screws. The detector components are

\(^1\)Thanks to Carlos Faham (LBNL) for making this measurement.
supported by PEEK rods which pass through holes in the mesh frames separated by PEEK spacers.

Initially the 5 mm electroluminescent gap was formed by separating two mesh frames with plastic panels, and the PEEK support rods passed through both mesh frames in the hexagonal perimeter. In an attempt to reduce high voltage breakdown between the two meshes, the “c-clamp” support approach was designed, in which the detector is constructed in two pieces, each ending with an EL mesh on one side, and the two EL meshes connected only by 3 c-shaped clamps which hook onto bolts fastened to the mesh frames by screws and supported by nuts. This strategy proved to be effective in maintaining a constant EL gap while minimizing electrical breakdown. The main components of the detector are highlighted in figure 3.6.

Figure 3.6: (Figure from [102]) Computer drawing of the pressure vessel and internal detector components. In this diagram the EL gap is maintained by smooth panels which were later replaced by “c-clamps” to reduce electrical breakdown at high voltages.

---

2In earlier configurations in which the electroluminescent gap was maintained by smooth plastic panels that extended straight across the gap, electrical breakdown lead to the production of dark tracks on the plastic which facilitated further breakdown.
PMT Array

The PMT array consists of 19 Hamamatsu R7378A, 1-in. diameter PMTs arranged forming a regular hexagon. The PMTs are plugged into an electronic board containing the resistor network necessary for distributing the total applied voltage across the dynodes. A grounded mesh is placed directly in front of the array to ensure the PMTs are not exposed to excessive fields that would interfere with their operation. The PMTs were calibrated with a pulsed LED placed inside of the TPC as described in section 3.5.

![Figure 3.7: The 19-PMT array shown through two wire meshes through which electroluminescent photons must pass to be detected.](image)

SiPM Tracking Plane

The tracking plane installed in the NEXT-DBDM prototype\(^3\) is an 8×8 array of Hamamatsu S10362-11-025P silicon photomultipliers (SiPMs), each with a 40×40 array of Geiger cells covering an area of 1 mm\(^2\) and yielding a gain of 2.75×10\(^5\). The SiPMs are coated with a thin film of tetraphenyl butadiene (TPB) and placed in the center of the hexagonal plane just in front of the electroluminescent region of the TPC. The SiPMs produce an electronic pulse for each detected photon, and the TPB serves to shift the incoming ∼ 175 nm light from xenon to visible light detectable by the SiPMs. The plane of SiPMs is shown in figure 3.8, and the electronics used in powering and reading out the SiPMs are shown in figure 3.9. Note that the electronics are stored inside the pressure vessel along with the tracking plane to minimize the number of cables that must be fed out of the pressure vessel.

\(^3\)The SiPM tracking plane design and much of the hardware came from our collaborators at the Instituto de Física Corpuscular in Valencia, Spain. The SiPM electronics were designed, installed, and debugged at LBNL by Azriel Goldschmidt, John Joseph, and Maxim Egorov.
3.3.3 Signals and Trigger

All high voltage signals and the grounding connection to the mesh in front of the PMT array are passed inside the pressure vessel via high voltage feedthroughs and attached to the appropriate mesh frames via teflon-coated cable. These high voltages, which range between -15 kV to +20 kV, are produced by three Glassman high voltage power supplies. The PMT high voltage is provided by a single supply and divided into 19 channels by a voltage divider module. The module allows for fine adjustment of the voltage delivered to each channel to vary the individual PMT gains which differ significantly at a fixed voltage. The PMT high voltage and signal cables are connected with coaxial cables via ports on the octagon and passed down the steel source tube into the pressure vessel, along the outside of the field cage, and to the PMT array. All PMT and SiPM signals from NEXT-DBDM are sent through coaxial cables from
feedthrough ports on the octagon and sampled by three 8-channel Struck SIS3302
digitizer cards which are read out via USB into a Linux machine through a SIS3150
controller card. The PMT signals are passed through a Phillips 777 amplifier with a
gain of x40 and then through an RC filter (50 MHz bandwidth) to extend their pulses
to lengths visible to the digitizer, which samples at 100 MHz. The SiPMs output fast
pulses that pass through several stages of processing in the electronics before final
readout. First, the pulses encounter two amplification stages, each with a gain of
×21. The resulting amplified pulses are then extended with an RC-shaper of time
constant $\tau = 500$ ns and enter a final amplification stage with a gain of ×11. These
extended, amplified pulses are then sampled using a compression scheme in which the
64 SiPM signals are read out in 4 waveforms with samples 10 ns in length and from
each of which 16 smaller waveforms are extracted.

The outputs of the SiPMs are multiplexed such that each SiPM is sampled $k = 4$
times before moving to the next, and after all have been sampled, 4 samples of
“buffer signal” are recorded (see figure 3.10). The buffer signal lies at a voltage
that is otherwise unreachable during normal operation, and the set of 4 buffer values
marks the end of a sampling cycle. Later, the 4 waveforms are disentangled into 64
waveforms by locating the end of the first sampling cycle (which determines some
initial offset $t_0$) and, knowing the order in which the SiPMs were sampled, extracting
the values for each SiPM thereafter. Note that in this multiplexing scheme, a single
SiPM is only sampled during its turn in the sampling cycle, and therefore the RC-
shaping is essential to ensure that the originally fast pulses are not lost while the
other SiPMs are being sampled.

![Figure 3.10: Schematic of the strategy used to record SiPM data in NEXT-DBDM.](image)

At a rate of 100 MHz, 4 samples are taken from each of 16 SiPMs, followed by 4
samples of an unreachable “buffer value,” marking the end of the sampling cycle. 4
of these waveforms are created, and are later disentangled to produce 64 individual
waveforms that can be used in track reconstruction.

The digitizer is controlled by a PC through a user interface written in TCL which
directs the digitizer to record data in blocks of 512 events and transfers the data from
the digitizer to a local drive in binary format. The binary data is then processed
block by block using an automated data management system written in Perl. The
raw data are converted to waveforms stored event by event in ROOT trees, and the
waveforms are then analyzed, producing key analysis quantities such as the energy
of each event and coordinates of points forming its reconstructed track (see section
3.4 Analysis

Here we describe how the signals from the detector are digitized and the resulting waveforms processed into meaningful quantities for physics analysis. Signals include those output by the 19 PMTs, the 64 SiPMs, and at times an external signal of interest to the particular dataset, for example a signal from a PMT coupled to an NaI scintillator. The analysis is performed on an event-by-event basis using the FMWK [103] analysis package. The samples for each digitizer channel originally stored as raw binary data are first converted to integer form and stored in ROOT trees. Each event is then processed through several steps.

1. Pedestal subtraction: For the 19 PMT waveforms, the mean and RMS are calculated using the first 1/16 of the total samples in the waveform, and the mean is subtracted from all samples. These values are calculated via an iterative algorithm\(^4\) in which the mean and RMS are calculated using only samples within

\(^4\)Note this algorithm was taken from analysis code originating from the Instituto de Física Corpuscular, Valencia, Spain.
some multiple of the previously calculated RMS from the previously calculated baseline. This minimizes the influence of actual signal on the calculation of the noise. The calculation is repeated until the RMS begins to increase, and the minimum RMS is selected.

2. Waveform sum: All of the PMT waveforms are summed into a single waveform, weighted by the single-photon charge (the average number of integrated ADC counts corresponding to a single detected photon). The RMS of the summed waveform is taken to be the quadrature sum of the RMS values of the individual PMT waveforms, again weighted by their single-photon charges. Here the 4 standard waveforms containing SiPM data are broken up into 64 individual SiPM waveforms and their mean and RMS values are calculated using the iterative algorithm described in step 1, except using the entire waveform rather than the first 1/16 of the total number of samples.\(^5\)

3. Peak finding: All peaks, that is subsets of the waveform that have dropped below some threshold and the samples immediately before and after according to predetermined parameters, are identified and classified as S1 or S2 based on their width and rise time. An S1 peak and one or more S2 peaks are selected for the event based on a preference of selecting the peaks with the most charge first.

4. Track reconstruction: Each peak identified in the summed waveform and selected as S1 or S2 is divided into “slices” of approximately equal charge. For each slice, information from individual PMTs and SiPMs is used to assign an x-y location to the energy deposited in the slice. The x-y coordinates are calculated from the individual PMT charges with a weighted average and from the SiPMs using both a weighted average and a maximum likelihood fit.

5. Analysis tree: Key variables such as total energy in S1 and S2, reconstructed track coordinates, etc. are stored in a ROOT tree for access on an event-by-event basis.

Additional details of several of these steps are given below.

### 3.4.1 Peak finding

The peak finding algorithm consists of two major steps: peak identification and peak classification. The summed waveform, constructed by adding the values for each sample in all 19 PMT waveforms (after determination of baseline and noise RMS in each), is analyzed in search of peaks, and the peaks are then classified as either S1 or S2 signals. To find peaks the following procedure is implemented in the analysis.

\(^5\)Attempting to avoid application of noise finding in later parts of the waveforms in which larger signals were likely to be present was not found to be necessary in the case of the SiPMs.
In peak identification, the summed waveform is scanned sample by sample beginning from \( i = 0 \). When the signal value at sample \( i \) has dropped below some threshold value, set as some multiple (usually between 3 and 5) of the calculated RMS of the sum waveform, peak integration begins. The peak is integrated sample by sample, beginning some number \( n_b \) samples before the sample \( i_0 \) that dropped below threshold and ending some number of samples \( n_f \) after the sample \( i_f \) whose value returns above the threshold. Typical values are \( n_b = 2 \) and \( n_f = 20 \). If a sample should drop below threshold again after \( i_f \), the integration is allowed to continue as normal until the next sample that returns above threshold. To avoid combining two legitimate peaks into one single peak, “peak split” samples are recorded during the integration, starting with sample \( i_0 \) and including any sample which returns above threshold between the previously determined \( i_f \) and \( i_f + n_f \) samples after which some sample drops below a second threshold higher than the first. At the end of integration, the integrated peak is split into subpeaks if 1 or more peak split sample has been recorded in addition to the initial sample. \(^6\) All samples encountered during the integration are stored in a peak data structure so that they can be accessed later to determine properties of the peak (such as its charge, or total integral) or to perform some operation such as fitting the peak to a Gaussian.

In peak classification, the list of peaks found in the peak identification step is examined, and a peak is classified to be “S1-like” or “S2-like” based on its width and total charge. \(^7\) An S1-like peak has total width \(^8\) between some minimum \( w_{1,\text{min}} \) and maximum \( w_{1,\text{max}} \) widths. To be considered S2-like, the peak had to meet similar criteria of minimum width \( w_{2,\text{min}} \) and minimum charge \( q_{2,\text{min}} \). Typical values are \( w_{1,\text{min}} = 0 \), \( w_{1,\text{max}} = 180 \), \( w_{2,\text{min}} = 50 \), and \( q_{2,\text{min}} = 200 \) (photons). If the conditions on the widths or charges of S1 or S2-like peaks overlapped, the S2 condition was only considered if the S1 condition was found to be false. Once all peaks have been classified, the two with the greatest charge values are selected. The peak with the greatest charge must be S2-like to continue, otherwise only S1 is set for the event and peak classification ends. If the peak with the greatest charge is S2-like and the other is S1-like, the two peaks are chosen as S2 and S1 for the event. If both peaks are S2-like, both are recorded as S2 for the event and the peak with the third-greatest charge is found, and this process repeats until an S1-like peak is found that comes earlier in time (samples) than all other recorded S2 peaks or until no peaks remain. Only one S1 peak can be recorded per event, and all recorded S1 and S2 peaks are passed to the next phase of the analysis and have their values saved in the final analysis tree.

\(^6\)Note that only S2 peaks are able to be split.

\(^7\)Often additional criteria, such as a check on the rise time (time between the first sample of the peak and the sample with the maximum signal), were implemented as additional classification factors.

\(^8\)Note that the width was calculated as the number of samples between the initial and final threshold crossings or the number of samples between the first and second determined threshold crossing, depending on the analysis.
3.4.2 Track reconstruction

The information recorded in the PMTs and SiPMs together can be used to reconstruct the track of a particle that interacts in the TPC. The weighted sum of the PMT signals constitute a waveform containing S1 and S2 light pulses that effectively represents the energy deposited by the particle per unit time. By dividing individual pulses of the waveform into smaller intervals of time (“slices”) and observing the integrated light detected by the SiPMs during each interval, one can reconstruct one or more points indicative of the x-y location(s) of the track at that time. From the time values of the interval itself one can determine a z-coordinate to accompany these x-y positions, and the collection of all reconstructed points constitutes the 3D track for that particle.

The algorithm used in track reconstruction consists of three primary steps:

1. The event is divided into slices, and for each slice the integrated SiPM charges are computed, yielding a 2D projection of that slice in x-y.

2. For each 2D projection, individual SiPMs are selected that are expected to be closest to a point on the track at which a single electron performed electroluminescence in the EL gap. These SiPMs are called “cores.”

3. For each core, one reconstructed point will be computed using the core itself and its neighboring SiPMs in a likelihood fit to an SiPM pattern that would be produced by electroluminescence from a single electron at a single point in the x-y plane.

Slicing the Event

At this point we assume that we have a waveform containing samples consisting of a number of photons (proportional to energy) per time bin, and that we have identified the initial and final sample values of each (S1 or S2) pulse in the waveform. Each pulse is then divided into slices by integrating from the initial sample value until the integrated charge $q$ reaches some specified amount $q_{s,0}$. At this sample the slice is ended, and integration begins for a new slice. Once the end of the pulse is reached, one or more slices have been recorded with a given start time $t_i$, end time $t_f$, and charge $q_s$. Note that if a pulse has been divided into more than one slice, and the final slice does not accumulate charge $q_{s,0}$, its charge and time interval are added on to the previous slice.

Once slices have been determined for all pulses, the SiPM charges collected during the slice time intervals must be determined. Since the SiPMs are sampled at a lower effective rate than the PMTs, the samples defining the slice interval must be converted to a corresponding interval in SiPM samples. For a given SiPM multiplexed from a mixed waveform with block length $k$ and initial block offset $t_0$ (see section 3.3.3), the initial and final SiPM samples, $t_{s,i}$ and $t_{s,f}$ containing signal observed in the slice time interval from $t_s$ to $t_e$ are

79
\[ t_{s,i} = (\text{int}) \left( \frac{t_s - t_0}{17k} \right) \]  
\[ t_{s,f} = (\text{int}) \left( \frac{t_f - t_0}{17k} \right), \]  
(3.1)  
(3.2)

where \((\text{int})\) signifies that the decimal portion of the result should be dropped. Since one SiPM sample spans a longer time interval than one PMT sample, overlap may occur between two consecutive slices when the end sample of one maps to the same SiPM sample as the start sample of the next. In order to avoid overlap, the SiPM charge in sample \(t_{s,i}\) should be multiplied by the fraction \(1 - f_1\), where \(f_1\) is the fraction dropped in the calculation of \(t_{s,i}\) (that is, the floating point value of \((t_s - t_0)/(17k)\) minus the integer part from equation 3.1). Likewise the SiPM charge in sample \(t_{s,f}\) should be multiplied by \(f_2\), where \(f_2\) is the fraction dropped in calculating \(t_{s,f}\). Once SiPM charges have been calculated for all SiPMs in a slice, the analysis of the resulting 2D projection can be carried out.

Core Selection

The 2D projection consisting of all SiPM charges for a given slice is analyzed to give one or more reconstructed points by first identifying core SiPMs. These are essentially those SiPMs with which a single point-like energy deposition in the x-y plane can be associated. If energy was deposited in more than 1 x-y location for a given z coordinate, then one should expect more than 1 core SiPM in the reconstruction.

In determining core SiPMs, the SiPM with the maximum charge in the projection is identified. To be considered a core, an SiPM must have a charge greater than some fraction \(f_{\text{max}}\) (typically \(\sim 0.4\)) of the maximum SiPM charge, and this value must be greater than some preset noise value \(q_n\). Therefore the SiPM with the greatest charge is automatically considered the first core if its charge value lies above \(q_n\). Once an SiPM has been identified as a core (we will now call its charge \(q_c\)), all of its neighbors are ruled out as potential cores (“disabled”) if they have charge less than or equal to \(q_c f_{\text{nb}}\), where \(f_{\text{nb}}\) is some preset parameter (typically \(\sim 0.85\)). This is done to ensure that two adjacent SiPMs with similar charges can both be considered cores. A “neighbor” SiPM is defined as one within some preset radius \(r_{\text{nb}}\) of the core SiPM.

Once a core has been identified, the SiPM with the greatest charge (excluding those of all previously identified cores and those neighbors of previously identified cores set to “disabled”) is identified as a core in a similar way. When no other SiPM meeting the criteria for a core remains, the core identification process ends, and each core and its associated neighbors are together assigned a single reconstructed point using a likelihood fit.

Likelihood Fit

The SiPM charges in each core and its associated neighbors are used to perform a likelihood fit to choose a single point best indicative of the light pattern they display.
In this fit, the relative SiPM charges are compared to a theoretical model for the light pattern that should be produced when light is emitted at a single location \((x, y)\) uniformly across the EL gap. We first develop this model by considering the uniform emission of light from a point \((x, y, z)\) in the EL gap and asking what fraction of the emitted light is incident on an SiPM of area \(A = 1 \text{ mm}^2\) at location \((x_i, y_i, 0)\) (see figure 3.11).

![Figure 3.11: A single EL electron emitting light instantaneously at a location \((x, y, z)\) in the EL gap. A fraction of this light strikes the face of the SiPM located at \((x_i, y_i)\) on the SiPM plane.](image)

Assuming \(z^2 \gg A\) and the flux density a distance \(r\) from the emission point is \(I = (I_0/4\pi r^2)\hat{r}\), we have a flux at the SiPM face of \(I \cdot A\), where \(A = A\hat{z}\). We consider the emission of a single photon and interpret the quantity \(I \cdot A/I_0\) as the probability of this photon striking the face of SiPM \(i\). Thus the probability that a single photon emitted at \((x, y, z)\) will strike the face of the SiPM at \((x_i, y_i)\) is

\[
p(x, y, z; x_i, y_i) = \frac{A}{4\pi} \left( \frac{z}{((x - x_i)^2 + (y - y_i)^2 + z^2)^{3/2}} \right), \tag{3.3}
\]

where we have used \(\hat{r} \cdot \hat{z} = z/\sqrt{(x - x_i)^2 + (y - y_i)^2 + z^2}\). For a given number of photons emitted uniformly over the EL gap, we average this probability over the interval \(z \in [z_e, z_e + d]\)

\[
\tilde{p}(x, y; x_i, y_i) = \frac{1}{d} \int_{z_e}^{z_e + d} p(x, y, z; x_i, y_i)dz
\]

\[
= \frac{A}{4\pi d \sqrt{(x - x_i)^2 + (y - y_i)^2 + z_e^2}} \left[ 1 - \frac{(x - x_i)^2 + (y - y_i)^2 + z_e^2}{(x - x_i)^2 + (y - y_i)^2 + (z_e + d)^2} \right]. \tag{3.4}
\]

The above function should give the probability of a photon hitting an SiPM, though it does not take into account the quantum efficiency of the SiPM. In fact, rather
than attempting to determine the exact number of photons emitted, one can let the normalization float as a free parameter in the fit. In this case, one is fitting to the shape of the light distribution on the SiPMs, and thus we normalize the probability to its value at \((x = x_i, y = y_i)\) and multiply by a normalization factor \(N\) to obtain

\[
k_i(x, y; x_i, y_i) = N \cdot \frac{z_e}{\sqrt{(x - x_i)^2 + (y - y_i)^2 + z_e^2}} \left\{ 1 - \sqrt{\frac{(x - x_i)^2 + (y - y_i)^2 + z_e^2}{1 - \frac{z_e}{z_e + d}}} \right\}.
\]  

(3.5)

The likelihood fit consists of choosing a point \((x, y)\) and a normalization \(N\), calculating \(k_i\) for each SiPM used in the fit at locations \((x_i, y_i)\), and comparing the calculated values to the actual charge values observed in the SiPMs for the particular slice. The comparison generates a value of the likelihood function \(\mathcal{L}\), and the best fit point \((x, y)\) is the one that maximizes this function (or in the actual implementation, minimizes the negative logarithm of this function, \(-\ln \mathcal{L}\)). The likelihood function is constructed assuming a Poisson distribution of detected photons on each SiPM with a mean \(\mu_i\) equal to the observed SiPM charge value,

\[
P_i(k_i; \mu_i) = \frac{\mu_i^{k_i} e^{-\mu_i}}{k_i!},
\]  

(3.6)

where each \(k_i\) is the theoretical normalized charge value calculated using equation 3.5. The function \(\mathcal{L}\) is then the product of the Poisson probabilities of the observed charge values \(\mu_i\) for each SiPM included in the fit

\[
\mathcal{L} = \prod_i P_i(k_i; \mu_i).
\]  

(3.7)

As the chosen location \((x, y)\) varies, the theoretical values \(k_i\) will vary for each SiPM. The point that best represents the energy deposition for a given core is the one that produces the \(k_i\) values such that the likelihood is maximized. In practice, we use the MINUIT package in ROOT [104] to vary the parameters \((x, y)\) and \(N\) until the negative logarithm of \(\mathcal{L}\) is minimized. Removing any constant factors, we seek to minimize the value \(M\)

\[
M = \sum_i \left[ (k_i + 1/2) \ln(k_i + 1) + \ln \left( 1 + \frac{0.0833}{k_i + 1} \right) - k_i (\ln n_i + 1) \right].
\]  

(3.8)

Note that \(M\) consists of the \(k_i\)-dependent terms of \(-\ln \mathcal{L} = -\sum_i \ln P_i\), and the other terms not dependent on \(k_i\) need not be included in the minimization, as they will be constant for all possible reconstructed points. Once the minimum value of \(M\) is found, the parameters \((x, y)\) that generated the \(k_i\) corresponding to the minimum \(M\) are the coordinates of the reconstructed point assigned to the core. This process is repeated for all cores to give the reconstructed track.
3.5 PMT Characterization

3.5.1 Determination of Single Photoelectron (SPE) Charge

In order to determine how many photons are present in the S1 and S2 signals produced in the TPC, one must know how much integrated signal corresponds to a single photon in each PMT. A calibration study that yields this information was performed with an LED connected through one of the signal feedthroughs in an octagon port and driven by a pulser module. The data acquisition was triggered on the driving of the pulser, and the magnitude of the voltage pulse sent to the LED was such that only a small fraction of the pulses produced one or more detected photons in a given PMT. This kept the probability of the production of multiple SPEs low. Waveforms of size 1024 samples were acquired and the baseline and noise RMS values were determined for each waveform by calculating the mean and variance of samples in the window [100,200). All peaks were located and integrated (similar to the integration procedure described in 3.4.1) in each waveform of each PMT using a single threshold to determine the initial and final samples of each peak, and each peak was integrated 1 sample before the initial threshold crossing and 3 samples beyond the final threshold crossing. The threshold was set to a value equal to twice the determined noise RMS below the baseline.

The histogram of the integrals of all SPE peaks should give the distribution of single photon charge values for each PMT. Ideally one expects a single maximum in this distribution and a width corresponding to the variance in SPE charge, but in practice there will be some contamination from noise pulses and possibly multiple SPE events. In this case, we have an exponentially rising background at low charge values due to noise superimposed on the SPE charge distribution. To isolate the single-photoelectron candidate events and reduce the number of noise pulses, several cuts were applied on the initial and final times (threshold crossings) determined for each peak. Figure 3.12 shows the initial times for all peaks found for PMT 10.9 Note the peak near sample 270 which represents the time at which pulser-induced signals must have occurred.

In addition to a narrow range of start times, SPEs produced by the pulser have some appreciable width, so the number of noise peaks selected can be further reduced with a cut requiring the final sample in the pulse to be greater than some distance from the initial. The histogram of all peak charges meeting the initial and final peak timing cuts is shown in figure 3.13. The histogram with a somewhat relaxed final peak time cut is also shown to demonstrate how the more restrictive cut eliminates noise in the spectrum. The final histograms for the remaining PMTs are shown in figures 3.14, 3.15, and 3.16. Timing cuts varied with PMTs to produce the most reasonable SPE spectrum.

9As the knowledge of when the trigger occurred was known, peaks were only located between samples [256,512).
Figure 3.12: Initial samples of peaks in PMT-10 for reasonable cuts on charge $q \in (100, 1000)$. Note the peak corresponding to the pulser-induced signals. The remaining structure in the distribution is likely to be due to periodic noise present in the waveforms.

Figure 3.13: SPE charge histograms (left) with final peak time at least 1 sample after the initial peak time and (right) with final peak time at least 2 samples after the initial peak time. Note that the restriction on the peak width significantly eliminates noise pulses. The Gaussian fit to the region of the spectrum corresponding to the SPE charge is shown in the plot on the right.

Gaussian fits to the histograms are superimposed in figures 3.13 and 3.14, 3.15, and 3.16. From their means we estimate the mean SPE charge for each PMT (rounding to the nearest whole number). These values are shown in table 3.1.
Figure 3.14: SPE charge histograms for PMTs 1-8.
Figure 3.15: SPE charge histograms for PMTs 9-16.
Figure 3.16: SPE charge histograms for PMTs 17-19.

Table 3.1: Mean SPE values for the 19 PMTs determined from the PMT calibration study.

<table>
<thead>
<tr>
<th>PMT</th>
<th>SPE Mean</th>
<th>PMT</th>
<th>SPE Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>354</td>
<td>11</td>
<td>266</td>
</tr>
<tr>
<td>2</td>
<td>310</td>
<td>12</td>
<td>378</td>
</tr>
<tr>
<td>3</td>
<td>353</td>
<td>13</td>
<td>240</td>
</tr>
<tr>
<td>4</td>
<td>293</td>
<td>14</td>
<td>304</td>
</tr>
<tr>
<td>5</td>
<td>323</td>
<td>15</td>
<td>320</td>
</tr>
<tr>
<td>6</td>
<td>414</td>
<td>16</td>
<td>279</td>
</tr>
<tr>
<td>7</td>
<td>443</td>
<td>17</td>
<td>173</td>
</tr>
<tr>
<td>8</td>
<td>394</td>
<td>18</td>
<td>314</td>
</tr>
<tr>
<td>9</td>
<td>284</td>
<td>19</td>
<td>297</td>
</tr>
<tr>
<td>10</td>
<td>334</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.5.2 Afterpulsing Analysis

One potential concern when working with PMTs is the presence of afterpulsing. In a PMT, a single photon can be detected when it liberates an electron from a photocathode which is then accelerated and strikes another electrode (the first dynode), causing the ejection of more electrons which then strike the next dynode, and this process is
repeated until a number of electrons sufficient to produce a macroscopic current is produced and they reach the final electrode (the anode) from which a current is read out (see for example [105]). Though the inside of a PMT is kept at a good vacuum, a small amount of gas may still be present.\textsuperscript{10} If a gas atom or molecule is ionized, it becomes an ion and travels in the fields established in the electrode chain in the opposite direction. If the ion gains enough energy to liberate one or more electrons and strikes the cathode or a dynode at an early stage in the gain process, it may be able to make enough electrons to produce an SPE-like signal (or stronger). This afterpulsing occurs typically hundreds of nanoseconds after the true pulse and results in an artificially high number of photoelectrons. In NEXT-DBDM where S2 signals typically last microseconds, there is no way to disentangle the afterpulsing and the true signal, and afterpulsing may have a negative effect on the energy resolution which we attempt to quantify here.

Statistics of Afterpulsing

We begin with a discussion of the statistics of PMT afterpulsing and develop a formalism for understanding afterpulsing within the context of a simple model. First consider the detection of a single photon in a PMT. We assume that the resulting signal consists of two components: a single photoelectron charge $q_s$ drawn from a distribution with mean $\overline{q}_s$ and RMS $\sigma_{q_s}$, and an afterpulsing charge $q_a$, drawn from a distribution with mean $\overline{q}_a$ and RMS $\sigma_{q_a}$. We assume that the $q_s$ component is present for every detected photon, but the $q_a$ component on average is present in only a fraction $p_a$ of detected photons. That is, we set $q_a = 0$ with probability $1 - p_a$, and with probability $p_a$ we draw $q_a = q'_a$ from a distribution with mean $\overline{q}'_a$ and RMS $\sigma_{q'_a}$. Our single photon charge will be $q = q_s + q_a$.

We first set out to understand how to express $\sigma_{q_a}$ in terms of the afterpulsing distribution RMS $\sigma_{q'_a}$ and the afterpulsing probability $p_a$. We will consider average values of the binomial process in which the occurrence of afterpulsing is determined by a probability $p_a$ and use these values in formulas for the estimators of $\overline{q}_a$ and $\sigma_{q_a}$. This should be accurate in the case of a large number of afterpulses. For example, from a set of $N$ afterpulsing values we have mean

$$\overline{q}_a = \frac{1}{N} \sum_{i=1}^{N} q_{a,i} = \frac{1}{N} [N(1 - p_a)(0) + Np_a \sum_{i=1}^{N} q'_{a,i}] = \frac{1}{N} p_a N \overline{q}'_a = p_a \overline{q}'_a. \tag{3.9}$$

That is, the mean on the observed afterpulsing distribution is equal to the mean of the distribution from which the afterpulsing values are drawn with probability $p_a$ multiplied by the probability of having an afterpulse. A similar calculation for $\sigma_{q_a}$ gives

\textsuperscript{10}In particular helium is known to be able to pass inside of PMTs and remain there, causing afterpulsing.
\[ \sigma_{q_a}^2 = \frac{1}{N} \sum_{i=1}^{N} (q_{a,i} - q_a)^2 = \frac{1}{N} \sum_{i=1}^{N} (q_{a,i} - p_a \overline{q}_a)^2 \]
\[ = \frac{1}{N} [N(1-p_a)(p_a \overline{q}_a)^2 + \sum_{i=1}^{N_{pa}} (q'_{a,i} - p_a \overline{q}_a)^2] \]
\[ = \overline{q}_a^2 p_a^2 (1-p_a) + \frac{p_a}{N_{pa}} \sum_{i=1}^{N_{pa}} (q'_{a,i} - p_a \overline{q}_a)^2 \]
\[ = \overline{q}_a^2 p_a^2 (1-p_a) + p_a[\overline{q}_a^2 - 2q^2_a p_a + q^2_a \overline{p}_a] \]
\[ = \overline{q}_a^2 p_a^2 (1-p_a) + p_a \sigma_{q_a}^2 + \overline{q}_a^2 - 2q^2_a p_a + q^2_a \overline{p}_a \]
\[ = p_a \sigma_{q_a}^2 + \overline{q}_a^2 - p_a \overline{q}_a \]

where we have assumed that \( N_{pa} \) is large enough to provide good estimators in going from lines 3 to 4 of the above equation, and in line 5 we have used \( \sigma_{q_a}^2 = \overline{q}_a^2 - q_a^2 \).

Note that if we measure \( \sigma_{q_a} \) and \( \overline{q}_a \) directly, we do not need to know about the specifics of the distribution of \( q_a \) or the probability \( p_a \), as our final variance of interest will ultimately be on \( q = q_s + q_a \). That is we are interested in finding

\[ \frac{\sigma_q}{\overline{q}} = \sqrt{\frac{\sigma_{q_s}^2 + \sigma_{q_a}^2}{q_s + q_a}}. \quad (3.10) \]

To do this, we will investigate quantities containing sums of single photons, separable into SPE and afterpulsing components. We will measure quantities \( Q = Q_s + Q_a \), where \( Q_s = \sum_{i=1}^{N_{p}} q_{s,i} \), and \( Q_a = \sum_{i=1}^{N_{p}} q_{a,i} \), in which \( N_p \) is the number of photons observed in a single measurement. In these measurements,

\[ \overline{Q}_s = \frac{1}{N} \sum_{i=1}^{N} \left( \sum_{j=1}^{N_{p,i}} q_{s,j} \right) = \frac{1}{N} \sum_{i=1}^{N} (N_{p,i} \overline{q}_{s,i}), \quad (3.11) \]

with a similar equation for \( \overline{Q}_a \). Note that if \( N_{p,i} = \overline{N}_p \) is a constant across all values of \( Q_s \), the above expression is equal to \( \overline{N}_p \) times the mean of the mean values \( \overline{q}_{s,i} \). In our measurements this will not be the case, and we will incur an additional error term in \( \sigma^2_{Q_s} \) due to \( \sigma_{N_p} \) (for example, equal to \( q^2_s \sigma^2_{N_p} \) if we assume \( Q_s = \overline{N}_p \overline{q}_s \)). Therefore we will be interested in a quantity that does not rely on a constant number of photons per measurement. We examine the quantity \( f = Q_a/Q_s \), so for a given measurement \( f_i \)

\[ f_i = \left( \frac{\sum_{j=1}^{N_{p,i}} q_{a,j}}{\sum_{k=1}^{N_{p,i}} q_{s,k}} \right) = \left( \frac{N_{p,i} \overline{q}_{a,i}}{N_{p,i} \overline{q}_{s,i}} \right) = \left( \frac{\overline{q}_{a,i}}{\overline{q}_{s,i}} \right) \overline{q}_a. \quad (3.12) \]

Now we assume that the error on the means \( \overline{q}_s \) and \( \overline{q}_a \) follow the equation, for example, \( \sigma_{\overline{q}_s} = \sigma_{q_s}/\sqrt{\overline{N}_p} \), where \( \overline{N}_p \) is the mean number of photons observed per
measurement. Note that if we have detected enough photons, the contribution to the error of a varying number of photons should not affect this result. In particular, noting that $\sigma_{q,s,j} = \sigma_{q,s}$ for all $j$,

$$
\sigma_{q,s}^2 = \left[ \frac{\partial}{\partial N_p} \left( \frac{1}{N_p} \sum_{i=1}^{N_p} q_{s,i} \right) \right]^2 \sigma_{N_p}^2 + \sum_{j=1}^{N_p} \left[ \frac{\partial}{\partial q_{s,j}} \left( \frac{1}{N_p} \sum_{i=1}^{N_p} q_{s,i} \right) \right]^2 \sigma_{q,s}^2.
$$

(3.13)

The first term evaluates to 0 since $(\partial/\partial N_p) \sum_{i=1}^{N_p} q_{s,i} = q_{s}^{11}$, and the second term evaluates to $\sigma_{q,s}^2/N_p$. Using this result and standard error propagation formalism to equation 3.12, we find that

$$
\frac{\sigma_f}{\bar{f}} = \frac{1}{\sqrt{N_p}} \sqrt{\left( \frac{\sigma_{q,s}}{q_{s}} \right)^2 + \left( \frac{\sigma_{q,a}}{q_{a}} \right)^2},
$$

(3.14)

where note that we have used that the mean value $\bar{f} = \bar{q}_a/\bar{q}_s$, where $\bar{q}_a$ and $\bar{q}_s$ are the means from the single-photon distribution.

We conclude that given an SPE calibration that yields known values for $\bar{q}_s$ and $\sigma_{q,s}$ and measurements that yield known values of $\bar{f}$, $N_p$, and $\sigma_f$, we can determine the quantity of interest $\sigma_{q}/\bar{q}$ (equation 3.10) by finding $\bar{q}_a$ and $\sigma_{q,a}$,

$$
\bar{q}_a = \bar{q}_s \bar{f},
$$

(3.15)

and

$$
\sigma_{q,a} = \bar{q}_a \sqrt{N_p \left( \frac{\sigma_f}{\bar{f}} \right)^2 - \left( \frac{\sigma_{q,a}}{\bar{q}_s} \right)^2}.
$$

(3.16)

Note that in our calculations we will take $N_p = \bar{Q}_s/\bar{q}_s$. As long as $N_p$ is sufficiently large, we can assume $\bar{q}_{s,i}$ in equation 3.11 to be well-determined and set it equal to a constant $\bar{q}_s$.

**PMT Characterization**

The statistical procedure outlined above was applied to measurements taken using an LED driven to create a $\sim 100$ ns pulse producing approximately 100-200 detected photons in each of 19 PMTs. 2048 waveforms in total were acquired, each containing the prompt signal from the PMTs corresponding to the arrival of the photons followed by the afterpulsing signal. For channel 1, the first waveform and the average of the first 128 waveforms are shown in figure 3.17.

---

11This result is sensible but we will not show it to hold rigorously.
Figure 3.17: Demonstration of afterpulsing in NEXT-DBDM PMT signals. Channel 1: Event 1 (left); Average of first 128 events (right) zoomed in on the relevant region to show afterpulsing. The initial large pulse corresponds to a signal produced by approximately 100-200 photoelectrons.

For each waveform, a mean baseline and noise sigma $\sigma_n$ value were determined using samples [50, 250). Then integration of the relevant parts of the waveform was performed in two ways:

1. Fixed-window integration: Each waveform was integrated over samples [275, 300] to determine the prompt charge $Q_s$, and then over samples [301,1300] to determine the afterpulsing charge $Q_a$.

2. Pulse-based integration: The region of the waveform over samples [275, 1300] was searched for pulses. A pulse begins when a sample in the waveform drops below a threshold value of $5\sigma_n$ below the determined baseline and lasts until the signal again returns above this threshold (several samples before these beginning and end points are also included in the integration). The integral of the first pulse found was considered to be the prompt signal if its end sample fell in some region of sensible sample values [291,311]. If this was not the case, the event was not considered in the pulse-based analysis (note that this means we have $\leq 2048$ events in the pulse analysis). The sum of the integrals of the remaining pulses was considered to be the afterpulsing signal.

The histogram of values of $f$ is shown for channel 1 in figure 3.18 for both analysis methods. The mean $\bar{f}$ and RMS $\sigma_f$ were determined from this histogram, and the number of photons was determined from the mean of the histogram of $Q_s$ (see for example figure 3.19), divided by $\bar{q}_s$. The values for $\bar{q}_s$ determined in the SPE calibration were used, and $\sigma_{q_s} = 0.7q_s$ was chosen by examining the fits in figures 3.14, 3.15, and 3.16, and choosing an approximate relation that is consistent across most PMTs. Equations 3.15 and 3.16 were then applied, and the resulting quantities were used in equation 3.10 to determine the quantity of interest $\sigma_q/\bar{q}$.

Table 3.2 shows the various key quantities obtained using the fixed-integration and pulse-based methods. Note that the significant discrepancy between the results
Figure 3.18: Relative total afterpulsing to signal ratio for PMT 1. A histogram of \( f = Q_a/Q_s \) for PMT 1 is shown using the fixed-window integration method (left) and pulse integration method (right).

Figure 3.19: Comparison of PMT signal values determined from the two integration methods. A histogram of \( Q_s \) values for PMT 1 is shown using the fixed-window integration method (left) and pulse integration method (right).
for the two methods in PMT 17 is a result of a large $\sigma_{q_a}$ due to heavy influence from noise in the fixed-window method. Because $q_a$ was low for this PMT, the results for fixed-window integration were largely influenced by noise, and therefore some events were found to have negative $q_a$ ($q_a$ was then set to 0 when this occurred). The pulse-based method, which only begins integration when the signal drops below a threshold, provides a more reliable result in this case.

Overall, the PMTs yielded a wide distribution of integrated single photoelectron charges with a width greater than half of the mean value in nearly all cases. The presence of afterpulsing introduced further uncertainty in the value of the total single photoelectron charge in cases where the afterpulses must be integrated along with the prompt signal. This is the case for S2 pulses in NEXT-DBDM, which have widths $\gtrsim 2\mu s$, and therefore photoelectrons produced as afterpulses, usually arriving 200-300 ns after the initial true photoelectron, will be produced at the same time as other true S2 photons in the pulse. The end effect is a degradation in energy resolution via a larger value of $\sigma_{q}/\bar{q}$ (see section 3.6.5). Including the effects of afterpulsing, most PMTs were found to have a value of $\sigma_{q}/\bar{q}$ between 0.9 and 0.95.
Table 3.2: Results for afterpulsing calculations using fixed-window integration, with pulse-based analysis results in parentheses. $\bar{q}_s$ was taken from SPE calibration data. It was assumed that $\sigma_{qs} = 0.7\bar{q}_s$, which seemed to match most PMTs based on Gaussian fits to SPE spectra, though note that some fits did not provide a clear measurement of $\sigma_{qs}$ and others gave a somewhat larger value of $\sigma_{qs}$.

<table>
<thead>
<tr>
<th>PMT</th>
<th>$N_0$</th>
<th>$\bar{q}_s$</th>
<th>$\sigma_{qs}$</th>
<th>$\bar{q}_a$</th>
<th>$\sigma_{qa}$</th>
<th>$\bar{q}$</th>
<th>$\sigma_q$</th>
<th>$\sigma_{qs}/\bar{q}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>103 (103)</td>
<td>354.0</td>
<td>247.8</td>
<td>16.99 (21.51)</td>
<td>230.3 (185.2)</td>
<td>371.0 (375.5)</td>
<td>338.3 (309.4)</td>
<td>0.9119 (0.8239)</td>
</tr>
<tr>
<td>2</td>
<td>97 (97)</td>
<td>310.0</td>
<td>217.0</td>
<td>16.04 (17.51)</td>
<td>204.9 (148.5)</td>
<td>326.0 (327.5)</td>
<td>298.5 (263.0)</td>
<td>0.9154 (0.8029)</td>
</tr>
<tr>
<td>3</td>
<td>112 (112)</td>
<td>353.0</td>
<td>247.1</td>
<td>15.74 (18.59)</td>
<td>214.3 (185.6)</td>
<td>368.7 (371.6)</td>
<td>327.1 (309.1)</td>
<td>0.8871 (0.8318)</td>
</tr>
<tr>
<td>4</td>
<td>123 (124)</td>
<td>293.0</td>
<td>205.1</td>
<td>16.02 (17.11)</td>
<td>209.2 (164.3)</td>
<td>309.0 (310.1)</td>
<td>293.0 (262.8)</td>
<td>0.9481 (0.8474)</td>
</tr>
<tr>
<td>5</td>
<td>113 (113)</td>
<td>323.0</td>
<td>226.1</td>
<td>18.13 (20.87)</td>
<td>219.4 (186.8)</td>
<td>341.1 (343.9)</td>
<td>315.1 (293.3)</td>
<td>0.9235 (0.8530)</td>
</tr>
<tr>
<td>6</td>
<td>104 (104)</td>
<td>414.0</td>
<td>289.8</td>
<td>21.61 (27.19)</td>
<td>265.2 (217.7)</td>
<td>435.6 (441.2)</td>
<td>392.9 (362.4)</td>
<td>0.9019 (0.8215)</td>
</tr>
<tr>
<td>7</td>
<td>104 (104)</td>
<td>443.0</td>
<td>310.1</td>
<td>27.69 (30.15)</td>
<td>304.5 (268.8)</td>
<td>470.7 (473.1)</td>
<td>434.6 (410.4)</td>
<td>0.9234 (0.8673)</td>
</tr>
<tr>
<td>8</td>
<td>120 (120)</td>
<td>394.0</td>
<td>275.8</td>
<td>21.56 (24.33)</td>
<td>267.9 (223.1)</td>
<td>415.6 (418.3)</td>
<td>384.5 (354.7)</td>
<td>0.9253 (0.8479)</td>
</tr>
<tr>
<td>9</td>
<td>153 (153)</td>
<td>284.0</td>
<td>198.8</td>
<td>11.69 (15.87)</td>
<td>185.1 (169.1)</td>
<td>295.7 (299.9)</td>
<td>271.6 (261.0)</td>
<td>0.9186 (0.8703)</td>
</tr>
<tr>
<td>10</td>
<td>107 (108)</td>
<td>334.0</td>
<td>233.8</td>
<td>17.75 (21.26)</td>
<td>231.2 (196.7)</td>
<td>351.7 (355.3)</td>
<td>328.8 (305.6)</td>
<td>0.9348 (0.8601)</td>
</tr>
<tr>
<td>11</td>
<td>144 (145)</td>
<td>266.0</td>
<td>186.2</td>
<td>10.78 (14.48)</td>
<td>172.7 (150.3)</td>
<td>276.8 (280.5)</td>
<td>254.0 (239.3)</td>
<td>0.9176 (0.8531)</td>
</tr>
<tr>
<td>12</td>
<td>115 (115)</td>
<td>378.0</td>
<td>264.6</td>
<td>19.22 (26.02)</td>
<td>261.6 (232.5)</td>
<td>397.2 (404.0)</td>
<td>372.1 (352.2)</td>
<td>0.9367 (0.8718)</td>
</tr>
<tr>
<td>13</td>
<td>136 (137)</td>
<td>240.0</td>
<td>168.0</td>
<td>9.717 (12.25)</td>
<td>155.0 (120.9)</td>
<td>249.7 (252.2)</td>
<td>228.6 (207.0)</td>
<td>0.9153 (0.8206)</td>
</tr>
<tr>
<td>14</td>
<td>127 (127)</td>
<td>304.0</td>
<td>212.8</td>
<td>15.46 (21.06)</td>
<td>211.8 (178.2)</td>
<td>319.5 (325.1)</td>
<td>300.3 (277.5)</td>
<td>0.9399 (0.8538)</td>
</tr>
<tr>
<td>15</td>
<td>137 (138)</td>
<td>320.0</td>
<td>224.0</td>
<td>14.89 (19.73)</td>
<td>206.5 (179.2)</td>
<td>334.9 (339.7)</td>
<td>304.7 (286.8)</td>
<td>0.9097 (0.8444)</td>
</tr>
<tr>
<td>16</td>
<td>135 (135)</td>
<td>279.0</td>
<td>195.3</td>
<td>18.21 (20.36)</td>
<td>229.2 (176.9)</td>
<td>297.2 (299.4)</td>
<td>301.1 (263.5)</td>
<td>1.013 (0.8803)</td>
</tr>
<tr>
<td>17</td>
<td>128 (128)</td>
<td>173.0</td>
<td>121.1</td>
<td>9.123 (10.32)</td>
<td>143.8 (95.51)</td>
<td>182.1 (183.3)</td>
<td>188.0 (154.2)</td>
<td>1.032 (0.8413)</td>
</tr>
<tr>
<td>18</td>
<td>108 (108)</td>
<td>314.0</td>
<td>219.8</td>
<td>14.90 (19.67)</td>
<td>191.9 (158.1)</td>
<td>328.9 (333.7)</td>
<td>291.8 (270.8)</td>
<td>0.8872 (0.8114)</td>
</tr>
<tr>
<td>19</td>
<td>157 (157)</td>
<td>297.0</td>
<td>207.9</td>
<td>14.92 (18.33)</td>
<td>196.3 (168.6)</td>
<td>311.9 (315.3)</td>
<td>285.9 (267.7)</td>
<td>0.9166 (0.8489)</td>
</tr>
</tbody>
</table>
3.6 Energy Resolution Studies of Electron Recoils in Xenon

A primary goal of NEXT-DBDM was to attain near-intrinsic energy resolution in high-pressure xenon gas for electron recoils with the low-noise amplification process of electroluminescence. To address this goal, electron recoils produced by gamma rays from a $^{137}\text{Cs}$ were studied intently. Simultaneous development of the data analysis and improvements to the hardware over the course of nearly a year lead to increasingly better resolution and eventually to arrival at the target resolution of an energy peak corresponding to the 662 keV gamma ray emitted by the source with a full width at half maximum equal to 1% of the mean (1% FWHM). All data in sections 3.6.1 through 3.6.5, unless otherwise noted, comes from a single dataset using a $^{137}\text{Cs}$ gamma ray source taken at a pressure of 10.1 atm. The TPC was configured as shown in figure 3.1. The voltages on the wire meshes were set such that the electric field in the drift region was approximately 530 V/cm and in the EL region was approximately 26.4 kV/cm, giving \( E/p = 2.6 \text{ kV/(cm.bar)} \) and by equation 2.4, an EL gain of \( g = 1250 \).

3.6.1 Source and General Event Characteristics

A 10 mCi $^{137}\text{Cs}$ source was placed in a lead enclosure with a small opening at one end and into which a small cylindrical lead collimator of diameter 3.5 mm was fitted. The resulting collimated beam of gamma rays was directed down the source tube and incident axially on the active region of the TPC. $^{137}\text{Cs}$ decays via $\beta^-$ decay, emitting a 661.657 keV gamma ray with 85.1% probability [29], and therefore the uncollimated source emits about $3.15 \times 10^8$ gammas/second. Figure 3.20 shows waveforms for several events produced by a 662 keV gamma ray in the TPC and for which the full energy was observed. Note the significant differences in the geometry of the S2 pulses despite nearly the same amount of deposited energy.

The mesh voltages and therefore the fields in the drift and EL regions may be varied for different datasets, and environmental conditions such as temperature may fluctuate slightly between runs. Therefore the number of S1 and S2 photons produced for a given energy electron recoil and the distribution of drift times for the events will differ between datasets, but nevertheless it is useful to give a general idea of some of the typical values characterizing the data. In general, for a 662 keV deposition, several hundred S1 photons and several hundred thousand S2 photons are detected. Electron drift velocity in the drift region is about 1 mm/$\mu$s, and typical widths of S2 pulses normally fall between 10-20 microseconds and depend on the orientation of the track. Electrons produced in tracks oriented parallel to the EL (x-y) plane will reach the plane after nearly the same drift time and produce a relatively narrower and more intense S2 pulse, whereas tracks oriented perpendicular to the x-y plane will produce a relatively longer and less intense S2 pulse. Pointlike depositions (such as 30 keV
x-rays) give an S2 pulse width of about 5 microseconds. Due to the presence of the teflon reflector panels and the buffer region, and since the EL plane is located on the side of the TPC opposing the PMT array, both the S1 and S2 light patterns cast on the PMTs are roughly uniform.

Even within the same dataset some variation in measured quantities with time is possible and even likely. A strong contributor to such variations is the change in temperature with the state of the PMTs. If the PMTs have been powered off for a significant amount of time and then powered on, the temperature change due to the heat produced in powering the PMTs is significant enough to affect PMT gains. The S2 charge for events near the 662 keV peak is plotted against event number (effectively time) in figure 3.21. Instead of attempting to correct for these effects we choose an event regime in which the variation is relatively small.

### 3.6.2 Gas Purity and Electron Attachment

The purity of the xenon gas significantly affects the performance of the TPC. After reclaiming the xenon from the pressure vessel, servicing the TPC, and then refilling it again, the gas must be recirculated through the hot getter for at least several days before stable performance can be expected. The purity can be monitored by measuring the electron attachment lifetime in a plot of the S2 charge against the drift
Figure 3.21: S2 charge vs. event number (effectively time) for events near the 662 keV peak. Though all variations shown are relatively small (note the zero-suppressed axes), for subsequent analyses of this data, we use only events 50000-75000 at which the S2 yield appears to be relatively stable.

time for each event. The drift time is measured as the average time between the start of the S1 signal and the S2 pulse. That is, the S2 pulse is assumed to be concentrated at one sample determined by its charge-weighted average. If the initial sample of the S1 peak is denoted $t_i$ and the S2 pulse has initial sample $t_1$ and final sample $t_2$, then we define the $n$th moment as

$$M_n = \int_{t_1}^{t_2} (t - t_i)^n q(t) dt,$$

where $q(t)$ is the number of S2 photons arriving at sample $t$ and $Q = \int_{t_1}^{t_2} q(t) dt$. The drift time is then given by $t_d = M_1 / M_0$. Figure 3.22 shows a scatter plot of the S2 charge vs. drift time for events near the 662 keV peak and the corresponding fit to an exponential. The high purity of the xenon gas achieved by the hot getter is evident in the fitted attachment lifetime constant of over 30 ms.

### 3.6.3 S1 and S2

The energies of events in the S2 peak with no corrections except for the event selection criteria established in section 3.6.1 are shown in figure 3.23. The resolution of the peak is currently about 1.5% FWHM, though with further analysis cuts this can be further improved (see sections 3.6.4 and 3.6.5). Before continuing with these improvements we consider the S1 signals produced by these events.

Since S1 is produced at a range of distances 0 to 8 cm from the PMT array, there will be some z-dependence on the amount of light detected. We consider a simple geometry to estimate the form of this dependence. Consider light emitted uniformly
Figure 3.22: Energy vs. drift time for events near the 662 keV peak (left) and its profile (right) fit to an exponential $\sim e^{-t/\tau}$. The error bars in the profile plot show the error on the mean values in each y-bin. As electrons spend a maximum of approximately 80 microseconds in the drift region, the gas in this run was very pure and few electrons attached.

Figure 3.23: Spectra of S1 (left) and S2 (right) charges before any corrections due to geometrical dependencies.

from a point placed a distance $z$ above the center of a circular plane as shown in figure 3.24.

If the total amount of light emitted is $Q_0$ photons, the intensity of the light incident on the circular area at a distance $r$ from the center of the circle is $I = Q_0/[4\pi(r^2+z^2)]$. The flux $Q$ of light through the surface is then $dQ = \hat{I} \cdot d\mathbf{A}$ where $d\mathbf{A}$ is the differential area element $d\mathbf{A} = rdrd\phi \hat{z}$ and $\hat{I}$ is the unit vector in the direction of the emitted light. We can determine the number of photons striking the area by integrating the total flux over all $\phi$ and $r$ from 0 to $R$. Noting that the cosine of the
angle between $\mathbf{\hat{i}}$ and $\mathbf{\hat{z}}$ is equal to $z/\sqrt{z^2 + r^2}$,

$$Q = \int_0^{2\pi} d\phi \int_0^R \frac{Q_0}{4\pi \sqrt{r^2 + z^2}} \cdot \frac{z}{\sqrt{r^2 + z^2}} \cdot r \, dr = \frac{Q_0}{2} \left( 1 - \frac{z}{\sqrt{z^2 + R^2}} \right). \quad (3.18)$$

Equation 3.18 predicts that half of the photons emitted will be observed at $z = 0$. However, the reflectivity of the chamber is such that even some of the light emitted facing away from the PMT plane will be detected. We assume that the reflected light will contribute a constant $Q_r$ to the detected number of photons and rewrite equation 3.18 as

$$q(z) = Q_r + \frac{Q_0}{2} \left( 1 - \frac{z}{\sqrt{z^2 + R^2}} \right), \quad (3.19)$$

where we now denote the number of detected photons as $q(z)$. Note that if $z \ll R$ we can make the approximation

$$q(z) \approx Q_r + Q_0 (1 - z/R)/2. \quad (3.20)$$

This linear approximation will turn out to describe the data well for our drift region in which the maximum value of $z$ is of order $R = 7.16$ cm, where $R$ is the center to vertex distance of the hexagon formed by the PMT plane. In assigning this value to $R$ we have approximated the hexagonal area of the PMT plane as its inscribed circle and assumed the entire area inside this circle is sensitive to photons. In practice fitting the S1 position dependence data is facilitated by the use of the drift time $\Delta t$ in place of distance from the EL plane $z$, as the drift time is the quantity reported by the analysis, and converting between $\Delta t$ and $z$ requires knowledge of the exact drift velocity. This can be determined from the data but requires the identification of the drift time corresponding to the end of the drift region closest to the PMT array, and this value can only be determined to some uncertainty. Thus for the purpose of
correcting the observed S1 charge for dependence on \( z \), it is most straightforward to use the drift time directly. Figure 3.25 shows the S1 values from the events shown in figure 3.23 plotted against drift time along with a linear fit of the form

\[
q(\Delta t) \approx q_0 + k\Delta t.
\]  

(3.21)

Figure 3.25: S1 charge vs. drift time for events shown in figure 3.23 (left) and the corresponding profile histogram (right). The fit gives \( q_0 = 218.5 \) photons and \( k = 2.736 \) photons/\( \mu s \).

The results from the fit can be used to correct the S1 charge \( q_{S1} \)

\[
q_{S1,c} = \frac{q_{S1}}{q(\Delta t)}q_0 \approx \frac{q_{S1}}{1 + (k/q_0)z},
\]

(3.22)

where \( q_{S1,c} \) is the corrected S1 charge. Note that the charge is corrected towards the lowest value \( q_0 \) produced on average furthest from the PMTs rather than the highest value produced on average nearest to the PMTs. This is again because the maximum drift time must be identified (introducing some uncertainty) in order to determine the greatest average S1 charge. Given maximum drift time \( \Delta t_{\text{max}} \) and total drift distance \( d \), one can substitute \( z(\Delta t) = d(1 - \Delta t/\Delta t_{\text{max}}) \) into equation 3.20 to find \( Q_0 = 2Rk\Delta t_{\text{max}}/d \) and \( Q_r = q_0 - k\Delta t_{\text{max}}(R/d - 1) \). For the fit shown in figure 3.25, if one takes the maximum drift time to be \( \Delta t_{\text{max}} = 72\mu s \), and for \( d = 8 \) cm, \( R = 7.16 \) cm, one finds that \( Q_0 = 352.6 \) photons and \( Q_r = 239.2 \) photons.

### 3.6.4 Radial Dependence of Measured Energy

In section 3.6.2, we discussed the \( z \)-dependence of the measured energy (S2) due to attachment of electrons during drift, and now we address the x-y dependence of S2 due to geometric effects (varying light collection efficiency). Although the light pattern produced on the PMT array during the production of S2 light is roughly uniform, some information still remains about where in x-y the event occurred. Coordinates
$\bar{x}$ and $\bar{y}$ can be assigned to each event based on a weighted average over the total amount of light collected in each PMT over the entire event.

$$\bar{x} = \frac{1}{Q} \sum_{i=1}^{19} x_i q_i \quad \bar{y} = \frac{1}{Q} \sum_{i=1}^{19} y_i q_i,$$

(3.23)

where $Q = \sum_{i=1}^{19} q_i$ is the total charge in the event. Figure 3.26 shows the x-y distribution of events over a wide energy range (not restricted to the 662 keV peak or the event range cuts determined in section 3.6.1), and the energy of the events (in the peak region and with event range cuts) plotted against radial distance from the center of the x-y distribution. Note that due to the nearly uniform illumination of the PMT array, the values $\bar{x}$ and $\bar{y}$ from the weighted averaging process are not in units of actual distance, and therefore all distances have been scaled by a factor of 23.28 to match the physical dimensions of the PMT array.

![Figure 3.26: Position dependence of S2. (Left) Distribution of events in x-y (not subject to S2 or event range cuts) for $\bar{x}$ and $\bar{y}$ calculated by equation 3.23. (Right) Energy (S2) of events near the peak (subject to event range cuts) vs. radial distance from a central point $(x_0, y_0) = (2.56, -2.63)$.](image)

It is clear that as events are produced further from some central point of maximum light collection efficiency the collected S2 signal steadily decreases. Rather than correcting for this effect we make a radial cut, eliminating events beyond some radius $r^*$, minimizing the geometric effects of light collection efficiency.

### 3.6.5 Energy Resolution

The best energy resolution attainable in the TPC was determined by choosing a set of events least affected by the systematic effects in the analyses discussed above. The events chosen were subject to several analysis cuts: only events 50000-75000 were considered (see figure 3.21), the S2 charge was within 30000 photons of the mean of the peak in the plot on the right in figure 3.23, the S1 was
chosen to be between 100 and 600 photons (note that some events had a misidentified S1, and only those with a sensible S1 are therefore kept), the radial distance from the center of the x-y event distribution (see figure 3.26) was less than the distance equivalent to a physical distance of 0.58 cm. Due to the long attachment lifetime no correction for electron attachment was applied. The resulting peak is shown in figure 3.27, showing a FWHM energy resolution of 1.07 ± 0.08%. Also shown is a peak for which the energy resolution was determined at approximately 30 keV using xenon x-rays and a similar strategy. The x-rays can be found as individual S2 peaks isolated from the remainder of the S2 in an event (see figure 3.20). Figure 3.27 shows a 5% energy resolution for x-rays from a dataset taken at 10 bar pressure.

Figure 3.27: (Above) FWHM energy resolution of the $^{137}\text{Cs}$ gamma peak after radial cut. Note that the axis has been calibrated assuming 662 keV $\rightarrow$ 547500 photons. (Below, figure from [102]) FWHM energy resolution of xenon x-ray peak near 30 keV.

Note that from [78] we know that the Fano factor in pure gaseous xenon is $F = 0.14$ (see section 2.3) and we will assume $W_i = 24.8$ eV as in [102]. We have EL gain $g = 1250$ and we will assume $J_{CP} \ll 1$. For $n_{\text{ion}} = E/W = 662000/24.8 = 26694$, we have $n_{\text{EL}} = gn_{\text{ion}} = 3.34 \times 10^7$, and $n_{\text{det}} = 547529$ from figure 3.23. We can then use equation 2.22 to calculate the best case energy resolution. We will assume $\sigma_q/\bar{q} = 1$ (see table 3.2) and note that due to the very large value of $n_{\text{EL}}$, the second term under the square root in 2.22 will be negligible. Using the above values, for $E = 662$ keV one would expect an energy resolution of

$$\left(\frac{\Delta E}{E}\right)_{\text{int}} \text{ (}% \text{ FWHM) } \approx 235 \cdot \sqrt{\frac{F}{E/W} + \left(\frac{\sigma_q/\bar{q}}{n_{\text{det}}}ight)^2 + \frac{1}{n_{\text{det}}}} = 0.54\%.$$ (3.24)

We have not reached this resolution most likely due to the inability to completely correct for discrepancies in geometry and systematic effects. Figure 3.28 gives a more detailed comparison of the obtained resolution with NEXT-DBDM compared to the expected theoretical resolution for a given number of detected photons. For gamma
Figure 3.28: (Figure from [102]) Expected and observed energy resolution in NEXT-DBDM under various configurations. The resolution is shown for different numbers of detected photons for 30 keV x-rays (triangles) at 10 bar xenon, 662 keV gamma rays (squares) at 15 bar xenon, and a pulsed LED (circles). The expected theoretical resolutions are also shown as the dot-dot-dashed line for 30 keV x-rays, the dot-dot-dot-dashed line for 662 keV gammas, and the dashed line for the LED. The resolution assuming only the Fano factor contribution is shown for the x-rays and gamma rays as dot-dashed lines, and the statistical limit $\sigma = 1/\sqrt{N}$ on the number of photons detected as the dotted line.

3.7 Preliminary Tracking Studies of Electron Recoils in Xenon

The data in section 3.6 was acquired in a configuration in which the SiPM plane was not installed in the TPC. Because the SiPM plane is coated with a layer of TPB to wavelength shift the UV xenon light to a wavelength detectable by the SiPMs, and because this coating may not be uniform and thus introduce an additional position dependence on the amount of detected light, energy resolution comparable to that
obtained in section 3.6 may not be obtainable with the SiPM plane installed. However, improved tracking using the algorithm described in section 3.4.2 was realized. Note: In the following examples, a preliminary version of the track reconstruction algorithm described above was used. Specifically, the overlap correction procedure described below equation 3.1 was not used (simply the integer values were taken), and prior to the likelihood fit, the all observed SiPM charges $\mu_i$ were normalized such that the value of the core charge was equal to 1. In addition, the distance between the SiPM plane and the EL region used to compute the theoretical light response functions in the analysis detailed here was not the same as the distance in the actual TPC, though a later revision of this analysis showed that this did not make a significant difference.

### 3.7.1 Example - Muon Track

The NEXT-DBDM TPC was set up to detect muons as shown in figure 3.29. The trigger was set on the output of an NaI scintillator placed below the TPC. A characteristic muon event was selected, and the reconstruction analysis was performed.

![Schematic of the muon tracking test setup for NEXT-DBDM.](image)

Figure 3.29: Schematic of the muon tracking test setup for NEXT-DBDM.

Figure 3.30 shows the waveform acquired summed over the PMT plane and the corresponding waveform showing the sum of all the SiPMs. The extent of each waveform is 16.384 $\mu$s, though the binning differs due to an effective SiPM sampling rate of $100$ MHz / $(17 \times 4) = 1.47$ MHz resulting from the multiplexing described in section 3.3.3. Though the PMTs record more precisely the features of the ionization track, the light observed by the SiPMs reflects a similar pattern. Note that certain features near the beginning and end of the S2 signal produced on the PMT plane are not shown in the SiPM signal. These are those produced at times in which the x-y location of the track was not covered by the 8×8 SiPM plane.

Figure 3.31 shows the reconstructed muon track in several projections (xy, xz, yz) and in three dimensions. The z coordinate extends into the drift region, where $z = 0$ corresponds to the beginning of the EL region.
Figure 3.30: Waveforms showing signal measured by sum of PMT plane (above) and sum of SiPM plane (below) for a muon track.

Figure 3.31: Reconstructed muon track.

In figure 3.32, 64 SiPMs are shown as circles shaded relatively according to the total amount of light observed by each (with some minimum value). The x-y coordinates at each point in the reconstructed track are given by red squares. By fitting a line to this projection and summing the squares of the difference between the fit and actual x-coordinates of each point, the RMS error in position determination is estimated to be approximately 1.2 mm.\textsuperscript{12}

\textsuperscript{12}Note that in a revised analysis this error was found to be about 2.1 mm.
3.7.2 Example - Gamma Track with X-ray

Here we present a gamma ray event in which a xenon x-ray was produced and interacted some significant distance from the main track. The summed PMT waveform for this event is shown in figure 3.33.

![PMT summed waveform for a gamma deposition with an x-ray.](image)

The reconstructed track is shown with several projections in figure 3.34. Note that the point corresponding to the x-ray is significantly separated from the rest of the track. The 2D projection used to reconstruct the x-ray location is shown in figure 3.35, along with waveforms for 4 SiPMs relevant to the reconstruction. Note that two SiPMs observe clearly light from the EL produced by the x-ray while two adjacent SiPMs see very little.

3.8 Summary

We have demonstrated that the good energy resolution attributed to gaseous xenon in chapter 2 is realizable in practice. We have also shown the potential for detailed mea-

---

Figure 3.32: (x-y) projection of muon track showing relative SiPM intensities.
Figure 3.34: Reconstructed track and projections for the event shown in figure 3.33.

measurement of tracks produced by high-energy electron recoils in gaseous xenon. Both of these achievements are highly relevant to a xenon-based $0\nu\beta\beta$ search experiment as both aid in background rejection in a context where low-background is essential.
Figure 3.35: SiPM waveforms and 2D projection for the energy slice corresponding to the x-ray deposition. In the 2D projection, the SiPMs are shaded relatively according to their intensities and the reconstructed point is shown as a red square.
Chapter 4

High Pressure Xenon for Dark Matter Searches

4.1 Introduction

The most straightforward way to calibrate a detector for the detection of low-energy nuclear recoils is to subject it to incident neutrons of MeV energies and observe the energy deposited in the elastic scattering interactions between the neutrons and the nuclei. This calibration measurement has been performed successfully in several detectors, for example using liquid xenon [93, 106] and liquid argon [107]. In addition to energy calibration purposes, such an experiment demonstrates the ability of a detector to distinguish between nuclear recoil events that would be produced by a heavy dark matter particle such as a WIMP, and electron recoil events resulting from gamma ray interactions. The first detection of nuclear recoils in gaseous xenon and demonstration of this rejection ability in the gas phase was presented by Prof. James White of Texas A&M University (see [108]). Here we study in detail the response of nuclear recoils in the NEXT-DBDM TPC. Note that due to instability, one of the 19 PMTs was turned off for the duration of all datasets described in this chapter.

4.2 Characterization of a $^{238}$Pu/Be Neutron Source

The present study was performed with an $\alpha$-Be source using $^{238}$Pu as the alpha emitter. The alphas emitted by the radioactive decay of this isotope and their intensities are shown in table 4.1. On March 13, 1996, the activity of the source was recorded to be 13.29 mCi, and the yield was stated to be $2.09 \times 10^4$ neutrons/s. We are interested in determining the neutron spectrum of this source at the time of use and the resulting spectrum of xenon nuclear recoils produced by elastic neutron scatters on xenon nuclei. To determine this we perform an analytical calculation using methods similar to those employed in [109] and [110, 111]. All ($\alpha, n$) and ($n, n$) cross sections

\begin{table}[h]
\centering
\begin{tabular}{|c|c|}
\hline
\textbf{Isotope} & \textbf{Activity (mCi)} \\
\hline
$^{238}$Pu & 13.29 \\
\hline
\end{tabular}
\caption{Activity of the $^{238}$Pu/Be source.}
\end{table}
used in calculations in this section were processed from Evaluated Nuclear Data File (ENDF) format [112] to Generalized Nuclear Data (GND) format using the Fudge [113] Python interface.

Table 4.1: $^{238}$Pu $\alpha$-decay energies ($Q = 5593.20$ keV), from [29].

<table>
<thead>
<tr>
<th>Alpha energy (MeV)</th>
<th>Intensity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5357.7</td>
<td>0.105</td>
</tr>
<tr>
<td>5456.3</td>
<td>28.98</td>
</tr>
<tr>
<td>5499.03</td>
<td>70.91</td>
</tr>
</tbody>
</table>

4.2.1 Spectrum of Emitted Neutrons

To determine the energies of the neutrons emitted, we adopt the following description of the $\alpha$-Be reaction. The $\alpha$ of mass $m_\alpha$ is emitted isotropically in the $^{238}$Pu decay and begins to travel through the surrounding $^9$Be (mass $m_b$). We assume that the amount of Be is significantly larger than the amount of Pu in the source and that it is uniformly distributed throughout the active region of the source, so that each emitted $\alpha$ particle travels through pure Be of constant density$^1$. We define the differential cross section for an incident $\alpha$ particle to produce a neutron at an angle $\theta$ with respect to the direction of incidence in the center of mass frame of an $\alpha$-Be collision as $d\sigma(\theta)/d\Omega$, noting that the cross section depends only on the angle $\theta$ in the plane of collision and not on the angle $\phi$ that would specify the orientation of the center of mass frame. Given the total cross section for the reaction $\sigma_T(E_{\alpha})$ (shown in figure 4.1) at alpha energy $E_{\alpha}$, we write the probability, given that the reaction occurs, for neutron emission at center of mass angle in the range $[\theta, \theta + d\theta]$ as

$$P(\theta; E_{\alpha})d\theta = \frac{1}{\sigma_T(E_{\alpha})} \int \frac{d\sigma(\theta; E_{\alpha})}{d\Omega} d\Omega = \frac{1}{\sigma_T(E_{\alpha})} \frac{d\sigma(\theta; E_{\alpha})}{d\Omega} 2\pi \sin \theta d\theta,$$  

(4.1)

The function $2\pi(d\sigma(\theta; E_{\alpha})/d\Omega)/\sigma_T(E_{\alpha})$, called the angular distribution of emitted neutrons, is shown in figure 4.2 for several alpha energies.

Our end goal is to compute a spectrum consisting of the number of neutrons at each neutron energy $N_i(E_n)$ produced per $\alpha$ particle emitted from Pu decay at energy $E_{\alpha,i}$. Therefore we are interested in the probability of neutron emission at energy $E_n$, given the reaction occurs with incident alpha energy $E_{\alpha}$. This quantity $G(E_n; E_{\alpha})$

$^1$We believe this is a valid assumption though the details of the interior of the source used to make the experimental measurements discussed in this study are unknown.
can be determined by noting that each angle $\theta$ corresponds to a given emitted neutron energy, and (see for example [115])

$$G(E_n; E_\alpha) dE_n \equiv P(\theta; E_\alpha) d\theta = P(E_n; E_\alpha) \frac{d\theta}{dE_n} dE_n,$$

(4.2)

where $E_n$ can be written as a function of $\theta$. To derive this function, we consider the kinematics of the ($\alpha$, n) reaction. For each reaction in the lab frame, an alpha particle with momentum $p_\alpha$ and energy $E_\alpha = p_\alpha^2/2m_\alpha$ impinges on a stationary Beryllium atom. The center of mass frame is the frame with origin at the location

$$\mathbf{R} = \frac{m_\alpha \mathbf{r}_\alpha + m_b \mathbf{r}_b}{m_\alpha + m_b},$$

(4.3)

where $\mathbf{r}_\alpha$ and $\mathbf{r}_b$ are the position vectors of the $\alpha$ particle and Be atom in the lab frame. We will define $M_{ab} \equiv m_\alpha + m_b$. Note that the velocity of the center of mass origin is $\dot{\mathbf{R}} = (m_\alpha/M_{ab})\dot{\mathbf{r}}_\alpha = \mathbf{p}_\alpha/M_{ab}$. To convert any position vector $\mathbf{r}$ to the center of mass frame one must subtract $\mathbf{R}$, so $\mathbf{r}_{CM} = \mathbf{r} - \mathbf{R}$, and $\mathbf{p}_{x,CM} = \mathbf{p}_x - (m_x/M_{ab})\mathbf{p}_\alpha$. Therefore in the center of mass frame, we have alpha and neutron momenta

$$\mathbf{p}_{\alpha,CM} = \mathbf{p}_\alpha - \frac{m_\alpha}{M_{ab}}\mathbf{p}_\alpha = \frac{m_b}{M_{ab}}\mathbf{p}_\alpha,$$

$$\mathbf{p}_{n,CM} = \mathbf{p}_n - \frac{m_n}{M_{ab}}\mathbf{p}_\alpha.$$

(4.4)
From conservation of momentum in the center of mass frame in which the total momentum is equal to 0 before the collision, we must have that the total final momentum is equal to 0. The magnitudes of the momenta of both initial particles are equal to each other and equal to $|p_{\alpha,\text{CM}}|$, and those of both final particles must also be equal to each other and equal to $|p_{n,\text{CM}}|$. Therefore we have from conservation of energy in the center of mass frame

$$\frac{p_{\alpha,\text{CM}}^2}{2m_{\alpha}} + \frac{p_{\alpha,\text{CM}}^2}{2m_{b}} = \frac{p_{n,\text{CM}}^2}{2m_{n}} + \frac{p_{n,\text{CM}}^2}{2m_{c}} - Q,$$

and defining $\kappa \equiv (m_n m_c / (m_{\alpha} m_b))(M_{ab} / M_{nc})$, where $M_{nc} \equiv m_n + m_c$, one can rewrite
this equation as

\[ p_{n,CM}^2 = \kappa (p_{a,CM}^2 + 2QM_{ab}). \]  

(4.6)

Using equations 4.4 and 4.6 one can write the energy of the neutron in the lab frame

\[ E_n = \frac{p_n^2}{2m_n} = \frac{p_{n,CM}^2}{2m_n} + \frac{p_a \cdot p_{n,CM}}{M_{ab}} + \left( \frac{m_n}{M_{ab}} \right)^2 \frac{p_a^2}{2m_n} \]

\[ = \frac{\kappa p_{a,CM}^2}{2m_n} + \frac{\kappa Qm_α m_{b}}{m_n M_{ab}} + \frac{p_a p_{n,CM} \cos \theta}{M_{ab}} + \left( \frac{m_n}{M_{ab}} \right)^2 \frac{p_a^2}{2m_n} \]

(4.7)

\[ = E_α \left( \frac{m_b m_c}{M_{ab} M_{nc}} \right) + \frac{Q m_c}{M_{nc}} + \sqrt{2E_α m_α p_{n,CM} \cos \theta} + \frac{m_n m_α E_α}{M_{ab}^2}, \]

where \( \theta \) is the center of mass scattering angle of the outgoing neutron. Defining \( x \equiv M_{nc}/M_{ab} \) and solving equation 4.7 for \( \cos \theta \) gives

\[ \cos \theta = \frac{E_n M_{nc}^2 - E_α x (m_b m_c + m_n m_α) - Q m_c M_{nc}}{2[E_α m_α m_n x^2 (E_α m_b m_c x + Q m_c M_{nc})]^{1/2}}. \]

(4.8)

Working in atomic mass units, we have \( m_n = 1, m_α = 4, m_c = 12, \) and \( m_b = 9. \)

Though these masses are not exact, we will have approximately \( M_{nc}/M_{ab} \approx (1 + 12)/(9 + 4) = 1, \) and so the final equation relating the center of mass scattering angle \( \theta \) to the lab energies \( E_α \) and \( E_n \) is [110]

\[ \cos \theta = \frac{E_n (m_n + m_c)^2 - E_α (m_b m_c + m_n m_α) - Q m_c (m_n + m_c)}{2[E_α m_α m_n (E_α m_b m_c x + Q m_c (m_n + m_c))]^{1/2}}. \]

(4.9)

We now have an expression \( E_n(\theta) = a + b \cos \theta, \) where therefore \( b = (E_n(0) - E_n(\pi))/2 \) (see [109]), so we can write \( dE_n = -b \sin \theta d\theta \) or

\[ G(E_n; E_α)dE_n = \frac{1}{\sigma_T(E_α)} \frac{d\sigma(E_n; E_α)}{d\Omega} \cdot \frac{4\pi}{E_n(\theta = 0) - E_n(\theta = \pi)} dE_n. \]

(4.10)

(We have dropped the negative sign as it reflects only that \( E_n \) decreases with increasing \( \theta \).) To determine \( N(E_n) \), we will have to integrate over all alpha energies, weighting the integration by the number of interactions that occur at each energy. The total cross section for the reaction (see figure 4.1) depends on the energy \( E_α \), so this weighting factor will vary over the track of the alpha as it moves through the Be. The spectrum can be found by integrating along the path of the alpha with the following logic,

\[ N(E_n)dE_n = dE_n \int dx \left( \text{differential number of interactions at } x \right) \times \]

\[ \left( \text{differential probability that emitted neutron has energy } E_n \right) \]

(4.11)
We have already determined the second term in the integral; it is $G(E_n; E_\alpha)$. To determine the first term, we note that the probability of interaction in a small segment $dx$ of the path of the emitted alpha is $dx/\lambda_{\alpha,n}$, where $\lambda_{\alpha,n} = [\rho \sigma_T(E_\alpha)]^{-1}$ is the mean interaction length of the reaction and $\rho \approx 1.23 \times 10^{23}$ atoms/cm$^3$ is the density of Be [16]. Note that for this to be valid, we must have $dx \ll \lambda_{\alpha,n}$. Even a total cross section greater than any $\sigma_T(E_\alpha)$ relevant in this problem (see figure 4.1), for example 0.5 barns, gives an interaction length of

$$\lambda_{\alpha,n} = (\rho \sigma_T)^{-1} \approx \frac{1}{1.23 \times 10^{23} \text{cm}^{-3} \cdot 0.5 \times 10^{-24} \text{cm}^2} \approx 16.2 \text{ cm}, \quad (4.12)$$

However, the track length $l_\alpha$ in Be of a 5.5 MeV alpha particle can be determined by the stopping power of alphas in Be [116] (see figure 4.3)

$$l_\alpha = \int_0^{E_{\alpha,i}} \frac{dx}{dE_\alpha} dE_\alpha \approx 28 \mu\text{m}. \quad (4.13)$$

Figure 4.3: Stopping power of alpha particles in Be in energy range of interest to the $^9\text{Be}(\alpha, n)^{12}\text{C}$ reaction for Beryllium (density $\rho = 1.848 \text{g/cm}^3$). The data from [116] was linearly interpolated to obtain this plot.

Because the interaction length is much larger than the track length, $l_\alpha/\lambda_{\alpha,n} \approx 1.7 \times 10^{-4} \ll 1$, any small length of the track $dx$ meets the required condition $dx \ll \lambda_{\alpha,n}$.  

114
Now we have
\[ N_i(E_n) dE_n = \left( \int \frac{dx}{(\sigma_T \rho(x) - 1)} G(E_n; E_\alpha(x)) \right) dE_n. \]  (4.14)

We can rewrite the integral for \( N(E_n) \) as an integral over alpha energies using
\[ dx = \left( \frac{dx}{dE_\alpha} \right) dE_\alpha, \]
and obtain
\[ N_i(E_n) dE_n = \left( \int_{E_\alpha,i}^{E_\alpha,i+\delta} \frac{4\pi [d\sigma(E_n; E_\alpha)/d\Omega]}{dE_\alpha/(\rho dx)[E_n(\theta = 0) - E_n(\theta = \pi)]} dE_\alpha \right) dE_n \]  (4.15)
where \( E_{\alpha,i} \) is the energy of the emitted alpha. Computing the spectrum of emitted neutrons \( N(E_n) \) for a given emitted alpha energy amounts to computing the integral in equation 4.15. The final spectrum of emitted neutrons will be the weighted average of the spectrum for each emitted alpha energy,
\[ N(E_n) = \sum_i x_i N_i(E_n), \]  (4.16)
where \( x_i \) is the alpha emission intensity for alpha energy \( E_{\alpha,i} \) from table 4.1. The final spectrum obtained from this calculation is shown in figure 4.4.

This spectrum can be integrated to determine the number of neutrons produced per alpha particle, and knowing the amount of radioactive isotope (in this case \(^{238}\text{Pu}\)) in the source along with its half-life and the branching ratio for alpha emission, one can determine the source rate of neutron emission. As 1 mCi = \( 3.7 \times 10^7 \) decays/s, the number of neutrons/s \( n \) emitted by the source is
\[ n = 37 NR, \]  (4.17)
where \( N = \int N(E_n) dE_n \) is the integrated spectrum and \( R \) is the source rate in mCi. The rates for neutrons emitted with the resulting \(^{12}\text{C}\) nucleus in different states are listed in table 4.2 for a 11.6 mCi source.

### 4.2.2 Spectrum of Nuclear Recoils

The spectrum of emitted neutron energies does not give enough information to predict the signals produced by neutron interactions in the TPC. The neutron-induced events of interest will be nuclear recoils, and even monoenergetic neutrons will not produce a recoil of fixed energy but lead to a spectrum of recoils. We first describe the calculation of the spectrum of a monoenergetic neutron, and we then integrate such spectra weighted by the distribution of emitted neutrons from the source (figure 4.4) to obtain the final spectrum of recoil energies expected in the TPC.
Figure 4.4: The neutron spectra produced by alphas emitted in the $^9$Be($\alpha$,n)$^{12}$C reaction in which the resulting $^{12}$C nucleus was left in the ground state, 1st excited state, and 2nd excited state. The sum of all three spectra is shown in a dark solid line. Compare to the spectra in figure 3 of [109], noting that in this case we did not consider the low-energy “break-up” neutrons discussed in this reference.

The problem of neutron elastic scattering on a xenon nucleus can be approached in the same way as the problem addressed in section 4.2.1 of an alpha incident on a beryllium atom. Now we have the process

$$n + Xe \rightarrow n' + Xe',$$  \hspace{1cm} (4.18)

and can write an equation analogous to equation 4.2 describing the probability $G_r(E_{x'}; E_n)$ per unit recoil energy of obtaining a recoil of energy $E_{x'}$ given incident neutron energy $E_n$,

$$G_r(E_{x'}; E_n)dE_{x'} \equiv P_r(\theta; E_n)d\theta = P_r(E_{x'}; E_n)\frac{d\theta}{dE_{x'}}dE_{x'}.$$ \hspace{1cm} (4.19)

The total cross section for neutron elastic scattering on several isotopes of xenon is shown in figure 4.5 for the range of neutron energies relevant to this discussion, and figure 4.6 shows several angular distributions of scattered neutrons for several neutron energies and xenon isotopes (see the discussion of equation 4.1).
Table 4.2: Rates of neutron emission for a 11.6 mCi $^{238}$Pu/Be source. Note that these values include only neutrons produced by the $(\alpha,n)$ reaction. From these results, approximately 65% of neutron emissions accompanied by a 4.4 MeV gamma.

<table>
<thead>
<tr>
<th>$^{12}$C state</th>
<th>$\gamma$ (MeV) [117]</th>
<th>Rate (neutrons/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ground</td>
<td>—</td>
<td>8756</td>
</tr>
<tr>
<td>$^{12}$C$^*$</td>
<td>4.439</td>
<td>21824</td>
</tr>
<tr>
<td>$^{12}$C$^{**}$</td>
<td>7.654</td>
<td>2963</td>
</tr>
<tr>
<td>Total</td>
<td>—</td>
<td>33543</td>
</tr>
</tbody>
</table>

Figure 4.5: The total elastic scattering cross section for neutrons incident on several xenon isotopes; data from ENDF [112].

Analogous to equation 4.10, one can express the differential recoil probability as

$$G(E_{x'}; E_n) dE_{x'} = \frac{1}{\sigma_T(E_n)} \frac{d\sigma(E_{x'}; E_n)}{d\Omega} \cdot \frac{4\pi}{E_{x'}(\theta = 0) - E_{x'}(\theta = \pi)} dE_{x'},$$  \hspace{1cm} (4.20)

where now we must express the recoil energy $E_{x'}$ in terms of the center of mass scattering angle $\theta$. This is the same problem as the one addressed in section 4.2.1.
Figure 4.6: Angular distributions for elastically scattered neutrons on xenon nuclei for several xenon isotopes; data from ENDF [112].

and for which equation 4.9 is the solution, with the replacements,

\[ \begin{align*}
    m_\alpha & \rightarrow m_n, \quad E_\alpha \rightarrow E_n \\
    m_b & \rightarrow m_x, \quad E_b \rightarrow E_x = 0 \\
    m_n & \rightarrow m_{n'}, \quad E_n \rightarrow E_{n'} \\
    m_c & \rightarrow m_{x'}, \quad E_c \rightarrow E_{x'}
\end{align*} \]

(4.21)

where \( m_x = m_{x'} \) is the mass of the xenon atom and \( E_x = 0 \) is the energy of the stationary xenon nucleus. Making the above replacements in equation 4.9 and noting that for elastic scattering, \( Q = 0 \), we have

\[ E_{n'} = E_n \left[ 1 - \frac{2m_nm_x}{(m_n + m_x)^2} (1 - \cos \theta) \right], \]

(4.22)
and by conservation of energy in the lab frame (where $E_x = 0$), $E_n = E_n' + E_{x'}$. Therefore (c.f. [37])

$$E_{x'} = \frac{2E_n m_n m_x}{(m_n + m_x)^2} (1 - \cos \theta), \quad (4.23)$$

and $E_{x'}(\theta = 0) = 0, E_{x'}(\theta = \pi) = 4E_n m_n m_x/(m_n + m_x)^2$. Given that the reaction occurs, then, our spectrum of recoils will be given by

$$G(E_{x'}; E_n) = \frac{(m_n + m_x)^2}{m_n m_x} \cdot \frac{\pi}{E_n} \left( \frac{1}{\sigma_T(E_n)} \frac{d\sigma(E_{x'}; E_n)}{d\Omega} \right). \quad (4.24)$$

where again we have ignored the negative sign as it reflects only the direction of change of $E_{x'}$ relative to $\theta$. Note that unlike the incident alpha particle in section 4.2.1, the neutron will not be stopped in the xenon and is unlikely to interact and lose energy before producing the recoil of interest, and so the number of recoils per incident neutron will depend on the density $\rho$ and thickness of xenon traversed $\Delta x$ and therefore the gas pressure and the direction of incidence of the neutron. The probability of elastic scattering will be equal to $\rho \sigma_T(E_n) \Delta x \ll (\rho \sigma_T)^{-1}$. For the purposes of a single incident neutron, this number is an overall normalization factor, but as we will combine spectra of many incident neutron energies, we will keep this factor in defining the number of interactions producing recoil energy $E_{x'}$ for a neutron of energy $E_n$. Note that the neutron scattering cross sections will differ with xenon isotope, and so we will have a different differential number of interactions for each target xenon isotope $a$,

$$N_{r,a}(E_{x'}; E_n) = \frac{(m_n + m_{x,a})^2}{m_n m_{x,a}} \cdot \frac{\pi}{E_n} \left[ \frac{d\sigma(E_{x'}; E_n)}{d\Omega} \right]_{a} \rho \Delta x. \quad (4.25)$$

Our final neutron spectrum will be a sum of such spectra weighted over the spectrum of incident neutron energies, equation 4.16, and weighted over the abundance of each isotope of xenon $f_a$ (the “representative isotopic composition” from [35]),

$$N(E_{x'}) = \sum_a f_a \int_0^\infty N(E_n) N_{r,a}(E_{x'}; E_n) dE_n. \quad (4.26)$$

Note that for our present calculation we will exclude isotopes $^{124}\text{Xe}$ and $^{126}\text{Xe}$ due to their very low (< 0.1%) natural abundance. Figure 4.7 shows the calculated and normalized spectrum of nuclear recoils along with normalized spectra of nuclear recoils obtained from the Monte Carlo simulation, in one (“ideal”) case considering only events that contained a single xenon recoil and a single neutron (no additional neutrons were generated from interactions, for example with the lead), and in another case considering the sum of all xenon recoils produced in a single event regardless of the number of neutrons produced. It is interesting to note that about 3.9% of events
containing at least 1 xenon nuclear recoil consisted of multiple xenon recoils, in which case the recoil energy observed would be the sum of their energies. The presence of the lead seems to have a significant influence on the recoil spectrum produced, and in particular, the “bumps” in the ideal spectrum around 40 keV and 80 keV, which would make S2 calibration much more straightforward, are not present in the final spectrum of all recoils.

Figure 4.7: Spectra of nuclear recoils produced by neutrons emitted from a $^{238}$Pu/Be neutron source. The influence of the lead block on the spectrum of generated recoils can be seen by comparing the calculated spectrum (dark solid line) with that of single-recoil, single-neutron events from Monte Carlo (solid line) and the spectrum of the sum of all recoil energies in a single Monte Carlo event (dashed line). The spectra agree well except at very low energies. Note that several events with total summed recoil energy > 200 keV were produced in the Monte Carlo and are not shown as part of this plot.

4.3 Experimental Configuration of NEXT-DBDM

The NEXT-DBDM detector was modified from the original configuration discussed in chapter 3 to characterize nuclear recoils. We will discuss the changes made to the TPC hardware including the addition of the wavelength shifting molecule tetraphenylbutadiene (TPB) which served to increase the light yield by shifting much of the $\sim 175$ nm xenon scintillation to the visible $\sim 430$ nm range at which the quantum efficiency of PMTs is generally significantly higher [118]. We will also discuss effects
on the time dependence of the observed scintillation light that apparently resulted from the presence of the TPB or some other agent introduced in the TPB coating process. As we are interested in how the behavior of nuclear recoils compares to that of electron recoils in xenon gas, we will also clarify our assumptions about electron recoils in NEXT-DBDM. That is, we will make specific choices of the amount of energy required to produce an ionization electron \( W_i \) and to produce a primary scintillation photon \( W_{sc} \).

### 4.3.1 Hardware Modifications

Two potential issues with the hardware configuration used for energy resolution and tracking studies in chapter 3 include low light collection efficiency and the presence of gas buffer regions. In these buffer regions, primary scintillation can be produced that is visible to the PMTs while the corresponding ionization produced cannot be collected. In an event consisting of some ionization in the sensitive drift region but one or more gamma rays that interact in a buffer region, the primary scintillation produced in the buffer contributes to \( S_1 \) while the ionization produced there does not contribute to \( S_2 \), thus lowering the observed \( S_2/S_1 \) ratio and leading to possible erroneous identification of the event as a nuclear recoil. To minimize the areas in which \( S_2 \) could not be collected, the two 5 cm buffer regions were eliminated by removing the mesh located 5 cm from the PMT array, leaving only the mesh placed directly in front of the PMTs, and placing a piece of teflon directly behind the mesh marking the end of the EL region. The teflon was not coated with TPB because in a previous configuration in which a TPB-coated 3M film was placed behind the EL region, the energy resolution was found to be considerably worse. This was likely due to nonuniformities in the TPB coating. The mesh directly in front of the PMTs was grounded, and positive voltages were placed on the two meshes defining the EL region. This modified configuration is shown in figure 4.8.

In addition, the teflon reflectors surrounding the drift region were coated with TPB. The presence of the TPB surrounding the sensitive region produced an overall increase in light collection by about a factor of 2. The TPB was dissolved in a toluene solution and deposited directly on the teflon panels using an airbrush. Though the application process was performed under fluorescent lighting, the panels were stored in a light-tight box and only exposed to incandescent lighting during installation in the TPC. This precaution was taken due to concerns of TPB degradation due to UV light.

### 4.3.2 \( S_1 \) and \( S_2 \) Yields for Electron Recoils

Before looking in detail at the response of NEXT-DBDM to nuclear recoils we first measure some characteristics of electron recoils. With the hardware modifications described in section 4.3.1, we examine the response due to electron recoils from inter-
actions of gamma rays emitted from a collimated $^{137}$Cs source directed axially along the TPC (similar to the configuration used to produce the data analyzed in section 3.6). The fields were configured such that the drift field was $E_d = 369 \pm 9$ V/cm and the EL gain was $g = 720 \pm 24$ photons/e$^-$.\footnote{The uncertainties in $g$ and $E_d$ were calculated by varying the voltages on the two anode grids by $\pm$ 100 V, the minimum increment on their digital displays.} Using the number of detected primary scintillation photons $S_1$ and electroluminescent photons $S_2$ for full-energy depositions of 662 keV gamma rays, we are interested in determining the energy required to produce an S1 photon $W_{sc}$. This can be calculated as

$$W_{sc} = \left( \frac{E}{S_1} \right) \epsilon,$$

where $E = 661657 \pm 3$ eV \cite{29} is the gamma energy and $\epsilon$ is the light collection efficiency at the EL plane. By knowing the EL gain and the energy required to produce an ionization electron $W_i$, the $\epsilon$ can be calculated as

$$\epsilon = \frac{S_2 \cdot W_i}{g \cdot E}.$$

Figure 4.8: NEXT-DBDM in its modified, shorter configuration with no buffer regions. The PMTs observe an 8 cm drift region that is terminated immediately at the EL plane.
We will assume a value of $W_i = 24.7 \pm 1.1$ eV for a moderate drift field (see the discussion in section 3.2 of [119]) and apply standard error propagation so that

$$\sigma^2 = \left( \frac{\sigma_{S_2}}{S_2} \right)^2 + \left( \frac{\sigma_{W_i}}{W_i} \right)^2 + \left( \frac{\sigma_g}{g} \right)^2 + \left( \frac{\sigma_E}{E} \right)^2. \quad (4.29)$$

From this we use the mean value $S_2$ determined from figure 4.9 and calculate

$$\epsilon = 0.0291 \pm 0.0016. \quad (4.30)$$

Figure 4.9: The $S_2$ peak for 662 keV gamma rays, corrected with a central radial cut and an exponential attachment correction with electron lifetime $\tau \approx 8.325$ ms. From this we determine $S_2 = 561617 \pm 92$ photons.

The light collection efficiency is therefore about 3% at the EL plane in this configuration. In determining $W_{sc}$ we will need to determine $S_1$ in a similar way, correcting for the $z$-dependence of the light collection. This is shown in figure 4.10, and in particular we will use the factor

$$k/q_0 = 8.8 \times 10^{-3} \ (\mu s \cdot photon)^{-1} = 8.8 \times 10^{-5} \ (\text{samples} \cdot \text{photon})^{-1} \quad (4.31)$$

to correct $S_1$ for $z$-dependence (see section 3.6.3) throughout the nuclear recoil studies.

Finally, eliminating $\epsilon$ from equation 4.27 using 4.28, we find

$$W_{sc} = \frac{S_2 \cdot W_i}{S_1 \cdot g} = 52.0 \pm 7.8 \ eV, \quad (4.32)$$

where we have propagated the uncertainties in the same way that lead to equation 4.29. Before continuing, we will make an additional note on the effects of the TPB
Figure 4.10: The S1 peak for 662 keV gamma rays (left), corrected for z-dependence as determined by a linear fit to the S1 data vs. drift time (right). From this we determine the mean $S_1 = 371 \pm 1$ photons and the quantities $q_0 = 370.6$ and $k = 3.272$ analogous to those in figure 3.25.

coatings on light collection. The presence of the TPB coating on the drift panels seems to have lead, by some mechanism not currently understood, to the presence of long tails in both the S1 and S2 signals. These tails are much longer than those expected from pure xenon scintillation. Due to the natural width of the S2 signals of several microseconds, the time characteristics of the tails are better extracted by examining the S1 pulses. Figure 4.11 shows dual-exponential fits to several S1 signals produced by full-energy 662 keV depositions. The signals appear to be well-described by a short decay component $\tau_s$ and a long decay component of $\tau_l$, where

$$\tau_s \approx 100 \text{ ns}, \quad \tau_l \approx 1400 \text{ ns.} \quad (4.33)$$

If we assume the total S1 signal is composed of two exponential signals $s_s(t)$ and $s_l(t)$ with time constants $\tau_s$ and $\tau_l$ and total area $I$, then we can write $s_s(t) = (f_s I / \tau_s) e^{-t/\tau_s}$ and $s_l(t) = (f_l I / \tau_l) e^{-t/\tau_l}$.

Figure 4.12 shows the integral the first 20 samples ($200 \text{ ns} = 2\tau_s$) of the full-energy S1 signals divided by the total S1 integral. The fraction of the total area in the initial 20 samples appears to be on average equal to 0.32. One can perform the integral of the two signals out to $2\tau_s$, which must equal $0.32 I$, and obtain

$$\int_0^{2\tau_s} [s_s(t) + s_l(t)]dt = I \left[f_s(1 - e^{-2}) + f_l(1 - e^{-2\tau_s/\tau_l})\right] = 0.32 I. \quad (4.34)$$

And this can be solved in conjunction with $f_s + f_l = 1$ to obtain

$$f_s = \frac{1 - 0.68 e^{1/7}}{1 - e^{(-2+1/7)}} \approx 0.26, \quad f_l = 1 - f_s = 0.74. \quad (4.35)$$

Therefore for the purposes of characterizing the nuclear recoil data, we will assume all scintillation is produced in the form of two exponentially decaying signals, one
Figure 4.11: The S1 signals from several events in NEXT-DBDM in which the full-energy deposition of a 662 keV gamma ray was measured. The signals are fit to a function of the form $p_0 + p_1 e^{-(t-p_2)/p_3} + p_4 e^{-(t-p_2)/p_5}$.

Figure 4.12: The fraction of the total S1 signal present in the first 20 samples for full-energy S1 signals. The average value of this quantity is about 0.32, meaning only about 30% of the S1 light produced is present in the prompt component.
with time constant $\tau_s = 100$ ns and containing 26% of the total light, and one with time constant $\tau_l = 1400$ ns and containing 76% of the total light.

4.4 Characterization of Nuclear Recoils with NEXT-DBDM

The response of Xe nuclear recoils in high pressure xenon gas (at approximately 14 bar pressure) was studied by collecting events with NEXT-DBDM produced by neutrons from the $^{238}$Pu/Be neutron source characterized in section 4.2 and from a $^{241}$Am/Be neutron source, which yields a similar spectrum of neutrons. The majority of data discussed in this section were acquired with the $^{238}$Pu/Be source, and most of the discussion will focus on this dataset except for that of section 4.4.3b, in which data acquired with the $^{241}$Am/Be source is discussed. The experimental setup for data taken with the $^{238}$Pu/Be source is shown in figure 4.13. The idea was to detect 4.4 MeV gamma rays emitted from the source in coincidence with an S1 signal in the TPC and thereby “tag” events induced by neutrons from the source. The neutron source was placed between a 2 inch thick (6.75 inch width x 4 inch height) lead brick and a NaI scintillator detector consisting of a NaI crystal coupled to a PMT. Though the brick attenuates the flux of neutrons on the TPC and changes the spectrum of the incident neutrons, it greatly reduces the probability that gammas originating from the source interact in the TPC. A detailed Monte Carlo study of the experimental setup (briefly described in appendix A) allows us to quantify these effects and will be used to guide our understanding throughout the study.

Figure 4.13: The experimental setup for the acquisition of data with a $^{238}$Pu/Be source.
4.4.1 Data Acquisition and Systematics

The $^{238}\text{Pu}/\text{Be}$ data were taken over a period of about 8 days. Events were acquired at a rate of approximately 1.4 Hz until the digitizer buffer was filled at 512 events, at which point the events were read out (requiring 1-2 minutes in which no new events were taken). This process was repeated, and figure 4.14 shows the acquisition rate calculated for each block. The total live time for the run was $T = 50422$ seconds or about 140 hours, though because 41581 events out of 715264 were eliminated to remove possible discrepancies due to disruption of the PMTs by internal electrical breakdown, we will consider an effective live-time of $T_{\text{eff}} = (1 - 41581/715264)T = 474911$ s.

![Figure 4.14: Event rate vs. live time for the nuclear recoil characterization run with a $^{238}\text{Pu}/\text{Be}$ source. Each pair of (time, rate) values was calculated per block. The rate was determined as $512/t_b$, where $t_b$ is the time difference in seconds between the recorded timestamp of the first and last events acquired in the block, and the live time on the x-axis is the sum of all times $t_b$ up to and including that block.](image)

The pressure and temperature were also monitored during the run and are shown in figure 4.15. The pressure was stable throughout the run to within 0.1 bar, and the temperature for a given location to about 1.5 °C. A serious issue was the presence of dielectric breakdown (sparking) in regions of the TPC at which high potential difference was maintained across short distances. The most prominent location of sparking was found to be localized at the high-voltage feedthrough delivering the highest voltage to the interior of the TPC (14 kV). In the past sparking has also been found to occur across the 5 mm electroluminescent gap. The large amount of light emitted in a spark is normally sufficient to cause the PMTs to draw significant current, and the HV supplies were set to automatically shut down the PMTs in the event of a spark and slowly (over the duration of about 1 minute) restore the applied voltage to the nominal operating values. Because this was likely to register events...
when the PMTs were operating at lower gain, all events occurring within 5 minutes after a spark were discarded.

![Graph showing pressure and temperature vs. run time](image)

Figure 4.15: Pressure (above) and temperature (below) vs. operation time for nuclear recoil data obtained with a $^{238}$Pu/Be source. The three different temperature curves correspond to three different measurement locations: inside the TPC at the top of the pressure vessel (green), inside the TPC on the side of the pressure vessel (blue), and outside the TPC (red). The vertical gray lines mark instances in which the PMTs were shut off, most likely due to the presence of sparking. Note that operation time spans the entire duration of the run, including the live time, all time required to perform readout, and any inactive time due to PMT tripping.

The trigger was designed using NIM electronics modules with two main components; one to identify an S1 signal in coincidence with an NaI signal corresponding to a 4.4 MeV gamma ray, and another to identify a single S2 pulse with integrated charge between some lower and upper limit$^3$. A successful S1/NaI component pulse is delayed beyond the possible arrival time of the corresponding S2, and a successful S2 identification pulse opens a gate during which time the delayed S1/NaI pulse is capable of triggering a single event. The trigger logic is shown in schematic form in figure 4.16. This combined S1/S2/NaI trigger gives a rate of $\sim 1.4$ Hz (see figure 4.14).

$^3$The S2 trigger was designed and implemented by Yasuhiro Nakajima (LBNL).

128
Figure 4.16: Summary of trigger logic for neutron runs. NaI pulses that fall between two thresholds are selected and delayed so as to enter a coincidence module with a signal produced from a discriminator when the sum of 16 of the PMTs passes some low threshold intended to identify an S1 pulse. The result of this coincidence is delayed 100 microseconds to be placed in coincidence with the S2 component of the trigger. This S2 component is formed by an S2 pulse detected in the 16-PMT sum signal by a discriminator that outputs a pulse continuously while the S2 signal remains above threshold. This discriminator pulse must end in coincidence with a 10 microsecond gate delayed 2 microseconds from the beginning of the S2 pulse. Effectively, S2 pulses with width between 2 and 10 microseconds are selected. An overall S2 upper threshold is imposed using a veto constructed from a single PMT. Additional trigger logic ensures that the S1 signal appears at the same location in the waveform by delaying the S1/NaI coincidence a fixed amount and using the resulting signal to form the final coincidence.
4.4.2 Selection of Neutron Events

In selecting valid events that could be analyzed as potential nuclear recoils, a series of cuts is applied. These include a tagging cut, a single-pulse selection cut, a diffusion cut, a fiducial cut, and an selection cut in S1-S2 space. A nuclear recoil in xenon is expected to produce a pointlike deposition of ionization that is drifted in the TPC to the EL gap and generates a Gaussian-like S2 pulse with a width dependent on the drift time (see figure 4.17). To eliminate backgrounds interacting near the walls of the TPC, events within a given radius of the center of the TPC are selected. Each of these cuts is described in more detail below.

Figure 4.17: Two candidate nuclear recoil events in NEXT-DBDM. Note the slight increase in width and decrease in amplitude with increased drift time (S1 occurred at about 20 µs in both waveforms). The S2 pulse in the upper waveform contains about 5465 photons, and in the lower waveform contains about 5346 photons.

NaI Tagging and Single-Pulse Selection Cuts

To improve timing resolution, both the S1 (19-PMT sum) and NaI pulses were fit to the functional form\(^4\)

\[
f(x) = p_0 + \frac{p_1 e^{-(x-p_2)/p_3}}{1 + e^{-(x-p_2)/p_4}},
\]

\(^4\)The use of this functional form was proposed by Yasuhiro Nakajima (LBNL).
and the start time of the pulse was set to \( p_2 \) as determined by the fit. Note that this procedure was only accurate for well-formed S1 pulses, and signals on the order of one or several photons may not have had enough ADC counts for a proper fit. In addition, the fit is meant to be executed on the prompt component of the S1 light and therefore would not return the correct start time for an event for which only S1 in the tail was observed, though in this case the correct start time would not be identifiable even by other means. The resulting S1 and NaI start time distributions are shown in figure 4.18.

Figure 4.18: Start times of NaI (left) and S1 (right) pulses as determined by a fit to equation 4.36. Note that the fit extended from the initial time as determined by the pulse-finding algorithm (see section 3.4.1) to 15 samples past the sample at which the pulse reached its peak amplitude. For the NaI pulses, the start time of the pulse with the greatest charge in an event was placed in the histogram shown here. However during the actual application of the tagging cut, the NaI pulse whose start time was determined to lie in the range described by the above histogram and closest to that of the S1 signal was chosen. Note that these histograms were generated before events were removed due to sparking as described in section 4.4.1.

The tagging cut was applied as follows: the event must contain an S1 pulse with start time \( t_0 \in (1885, 1935) \) and a NaI pulse with start time \( t_0 \in (1880, 1945) \). If multiple NaI pulses were present in the valid range, the one with initial time closest to that of the S1 signal was chosen. In addition, a single-pulse selection cut was applied, which demanded that the total S2 value be between 95% and 105% of the S2 in a single pulse. This ensured that the S2 integration was not distributed over several pulses in the waveform. An example of a properly-tagged neutron-like event is shown in figure 4.19.

**Diffusion Cut**

By plotting the variance of the Gaussian fit to each S2 pulse passing the tagging and selection cuts described previously against the mean drift time \( \bar{t} \) determined
Figure 4.19: A candidate nuclear recoil event (above) and its corresponding NaI signal (below).

from the analysis, one can observe a clear correlation between the drift time and the pulse width. The events for which S1 was properly tagged and identified lie in this band shown in figure 4.20. Similar to the approach taken in [120], this band can be described by a line,

$$\sigma^2 = \sigma_0^2 + \alpha \bar{t}. \quad (4.37)$$

Here $\sigma_0^2$ is related to the length of time required for a single electron to traverse the EL gap, and $\alpha$ is related to the longitudinal diffusion of electrons drifting in xenon gas under the specific conditions of pressure, temperature, and drift field. To convert $\alpha$ (in $\mu s^{-1}$) to an amount of spatial diffusion in mm/$\sqrt{\text{cm}}$, one can ask how much drift distance corresponds to a given pulse width time. The answer is given by the drift velocity $v_d$ which we will express in units of mm/$\mu s$. If $\sigma$ is expressed in $\mu s$ then the corresponding width in distance $z$ is $\sigma_z = v_d \sigma$. If $\alpha \bar{t}$ gives a result in $\mu s^2$ for $\bar{t}$ in $\mu s$, then $\alpha$ also has dimensions of $\mu s$. One can rewrite $\bar{t} = z/v_d$ where $z$ is the drift distance in mm, and

$$\sigma_z^2 = (v_d \sigma)^2 = (v_d \sigma_0)^2 + v_d^2 \alpha (z/v_d) = (v_d \sigma_0)^2 + v_d \alpha z.$$ 

When S2 ionization drifts to the EL region, it must then experience diffusion in $z$ with a sigma of $\sqrt{v_d \alpha z}$ mm, or if we express the drift length $z$ in cm, $\sqrt{(v_d \alpha) 10z}$, so
that dividing out \( z \), we have

\[
D = \sqrt{10v_d \alpha} \text{ mm}/\sqrt{\text{cm}},
\]

(4.38)

where again \( \alpha \) is the quantity from equation 4.37 expressed in \( \mu \text{m} \) and \( v_d \) is in \( \text{mm}/\mu\text{s} \).

Figure 4.20 shows the variance \( \sigma_p \) vs. drift time from a Gaussian fit to the S2 pulses in both data and Monte Carlo. The centroid was computed using a method similar to the one described in section of [121], in which the range of drift times was divided into several bins and the values of the variance \( \sigma_p^2 \) were histogrammed and fit to a Gaussian. The set of points in (drift time, \( \sigma_p^2 \)) space from these Gaussian fits were then fit to a second order polynomial, and the value of this polynomial at the center of each drift time bin was subtracted from all \( \sigma_p^2 \) values with drift times in that bin. This process was repeated iteratively until the difference between the computed centroids between iterations converged. Fits of second order polynomials to the set of points lying \( \pm 2 \) standard deviations from the centroids are shown as red dotted lines, and all events within these lines were considered to have passed the diffusion cuts.

Figure 4.20: The \( \sigma \) of a Gaussian fit to S2 pulses vs. mean drift time for events in the experimental data meeting the tagging and single-pulse cuts and in the corresponding Monte Carlo simulation. The solid red line defines the centroid of the distribution, and the dotted red lines encompass events that lie \( \pm 2 \) standard deviations from the centroid. Note that the experimental diffusion cuts differ slightly from those in Monte Carlo, most likely due to the presence of long tails observed in the S2 pulses that were not modeled in Monte Carlo. In this case, the lack of such tails in Monte Carlo required a longer EL light emission time to match the zero-drift \( \sigma_p^2 \) value observed in experiment, causing the S2 pulse widths to be dominated by this parameter and remain more constant at earlier drift times.

The diffusion centroid of the experimental data can be described by

\[
\sigma_p^2 = 0.83 + 0.017 \bar{t} + 0.000020 \bar{t}^2.
\]

(4.39)
If we consider only the linear and constant terms (a good approximation, as even at $t = 100 \mu s$ these two terms will give the greatest contribution), we have $\alpha = 0.017$. Given a maximum drift time of about $95 \mu s$ for an $8$ cm drift length, we have a drift velocity of $v_d \approx 80/95 = 0.84 \text{ mm/\mu s}$, and thus we find $D \approx 0.38 \text{ mm/\sqrt{cm}}$.

**Fiducial Cut**

Though the TPC was not configured for detailed tracking for the purpose of detecting neutrons, it was possible to use the distribution of light on the PMTs to perform some fiducialization. Note that the absence of signal in the disabled PMT was taken into account by using as the observed charge value the average of the charge observed in the surrounding PMTs. The distribution in reconstructed $\overline{x}$ and $\overline{y}$ of the events passing tagging, single-pulse, and diffusion cuts is shown in figure 4.21. The $(\overline{x},\overline{y})$ values were calculated using a light-weighted position average as in section 3.6.4, and due to the uniformity of the light pattern on the PMT plane, do not correspond to physical lengths.

![Figure 4.21: The distribution of reconstructed $(\overline{x},\overline{y})$ values for single-pulse, tagged events passing the diffusion cuts. The distribution is somewhat biased due to uncertainties in PMT calibration, particularly in quantum efficiency, as only the single-photon charges were determined in the calibration. The resulting biases include a shift in the center and a scaling along both the x and y axes and will be corrected using Monte Carlo (see figure 4.22).](image)

Due to differences in PMT quantum efficiency, possible inaccuracies in PMT calibration, and potential geometric effects introduced by the presence of TPB, the reconstructed $(\overline{x},\overline{y})$ distribution has a central value different from $(0,0)$ and is elongated in the x and y directions by different scale factors. To correct for these effects the projections of figure 4.21 along the x and y axes were plotted and compared to
those from the Monte Carlo run (this was done before calibration of nuclear recoil yields, but this should not make a significant difference as the S2 charge value of the event did not seem to have a significant effect on the (x,y) pattern produced). These projections are shown in figure 4.22. Gaussian fits were made to the peaks at both ends of each projection, and differences in the resulting means calculated. The centers of the experimental distributions were set such the point halfway between the two means was set to zero, and values in each distribution were scaled so that the difference between the two means matched that of the Monte Carlo distributions.

Figure 4.22: The x (top) and y (bottom) projections of the reconstructed (x, y) locations of events in data (left) and Monte Carlo (right). The center of the experimental distributions was shifted by (-0.24, 0.27), the x-values scaled by 0.58, and the y-values scaled by 0.80 to match the general characteristics of the Monte Carlo distribution.

A single fiducial cut, $\sqrt{x^2 + y^2} < 0.3$ was then applied in both data and Monte Carlo. The resulting cut is shown in figure 4.23 and, assuming the distance from vertex to vertex across the $y = 0$ center of the plane is about 14 cm, corresponds to a radial cut of about $r < 5.5$ cm.
Energy Selection

After all of the above described cuts, one can identify clearly the nuclear recoil candidate events in S1-S2 space. As in liquid xenon, nuclear recoils are likely to be quenched relative to electron recoils, and therefore we should look for a low, correlated band of events at low S2 and S1 values. Figure 4.24 shows S1 vs. S2 for the tagged, single-pulse events at different stages of the analysis. A band of low-energy events is identified to be the nuclear recoil events, and the region of events selected for the fit is shown.

4.4.3 Calibration of Nuclear Recoil Yields

Here we attempt the difficult task of obtaining information about the absolute yields of nuclear recoils in xenon gas from the limited fraction of the nuclear recoil spectrum observed in this study. This is done in two different ways. The present data do not allow for the measurement of primary scintillation and ionization yields for nuclear recoils of distinct energies, as the energies and directions of the incoming and scattered neutron are not known precisely, but by comparing the spectra of S1 and S2 signals in experiment and Monte Carlo, one should find consistency if the number of photons produced and detected for a given nuclear recoil energy are correctly modeled. Using additional data with a $^{241}$Am/Be source, we also examine an apparent feature of the recoil spectrum that should be identifiable in the absence of lead shielding between the source and TPC. Though not statistically significant enough for a strong claim, evidence of this sudden change in behavior of the recoil spectrum is present in the S2 spectrum, and yields are extracted under this assumption.
Figure 4.24: S2 vs. S1 for events passing the tagging and single-pulse cuts (top left), with added diffusion cuts (top right), and added diffusion and fiducial cuts (bottom). The S1 has been corrected for each event to its corresponding value at the EL plane using the relation shown in figure 4.10. The space enclosed by the dotted red lines is the region in S2 vs. S1 selected to perform the fit and encloses the nuclear recoil candidate events. One can observe three additional lines at higher S2, two (≈ 25000 photons and ≈ 28000 photons) corresponding to xenon x-rays of energies near 30 keV, and one (≈ 35000 photons) corresponding to gammas near 40 keV resulting from inelastic scatters of neutrons off of $^{136}$Xe nuclei.
a) Monte Carlo Spectral Fits

The Monte Carlo (described in further detail in appendix A) was run assuming electron recoil yields consistent with those discussed in section 4.3.2, and some constant quenching factors for S1 and S2 production in nuclear recoils relative to electron recoils. More specifically, the Monte Carlo assumed $W_1 = 25$ eV, $W_{sc} = 50$ eV and constant quenching factors of 1.63 for primary scintillation and 5.33 for ionization\(^5\). The S1 and S2 of selected events of the Monte Carlo were further corrected to match those of experiment by means of the fit procedure described below.

The fit attempts to identify the nuclear recoil yields that, when applied to the nuclear recoils produced in the Monte Carlo, most closely give the S1 and S2 spectra in the energy range of interest determined in section 4.4.2. As the Monte Carlo generates waveforms that are processed in the same way as the experimental data, we are able to apply nearly the same cuts as those described in 4.4.2 to Monte Carlo events, though in practice we have to account for the possibility that events containing nuclear recoils may have their energies modified during the fit procedure and may pass all cuts under certain yield assumptions and may not under others. In addition, no tagging is done in Monte Carlo, rather an event is only recorded if a neutron elastic or inelastic scatter occurred during the simulation of the event. There are three parameters in the fit, including a rate scale factor $\beta$, an S1 yield factor $\eta_1$, and an S2 yield factor $\eta_2$. If we denote the yield (the number of photons or electrons produced per unit of total deposited energy) as $Y_1$ for primary scintillation and $Y_2$ for ionization, for a given $\eta_1$ and $\eta_2$, the S1 and S2 yields are assumed to follow

$$ Y_1 = \frac{\eta_1 f_n(E)}{W_{sc}}, \quad Y_2 = \frac{\eta_2 f_n(E)}{W_1}, \quad (4.40) $$

where $f_n$ is the Lindhard factor. That is, though we are allowing for a constant multiplicative factor, we are assuming Lindhard energy dependence of the scintillation and ionization yields for nuclear recoils. Because no published data exists as a starting point, we assume this as a reasonable initial assumption. Note that we are also assuming constant yields for electron recoils and using W-values determined in a calibration to an electron recoil energy of 662 keV. This appears to be a valid assumption based on the detected S1 and S2 for $\sim 30$ keV x-rays and 662 keV gamma rays. The Monte Carlo fit is applied as follows: first, all Monte Carlo events that contain at least 1 xenon nuclear recoil or meet the single-pulse selection cuts are selected. For these events, the S1 and S2 charges are subject to a smearing procedure to better match the observed detector resolution (see section A.4). Then for a given parameter set $(\beta, \eta_1, \eta_2)$, the S1 and S2 values of each event are corrected assuming that all nuclear recoils have yields as given by equation 4.40.\(^6\) After the yield correction is performed,

---

\(^5\)By “constant quenching factor” we mean, for example, that the nuclear recoil scintillation yield at all energies was equal to that of an electron recoil of the same energy divided by 1.63.

\(^6\)The number of S1 photons and ionized electrons produced by each nuclear recoil in each event can be identified after the Monte Carlo has been run, and therefore the fraction of total energy that
the diffusion, fiducial, and energy selection cuts are applied. The events passing all cuts are used to generate S1 and S2 spectra that are fit to the experimental spectra using a binned $\chi^2$ method, in which

$$
\chi^2 = \sum_i \left( \frac{p_{1,i} - \beta q_{1,i}}{p_{1,i}} \right)^2 + \sum_j \left( \frac{p_{2,j} - \beta q_{2,j}}{p_{2,j}} \right)^2,
$$

(4.41)

where the $i$ sum over the bins in the S1 histograms and the $j$ sum over the bins in the S2 histograms, and the values $p_{1,i}$ ($q_{1,i}$) and $p_{2,i}$ ($q_{2,i}$) are the number of counts in the $i$th bin of the experimental (Monte Carlo) S1 and S2 spectra.

The fit procedure is performed on a three-dimensional grid of values of the parameters ($\beta, \eta_1, \eta_2$), and the calculated values of $\chi^2$ are recorded for each parameter set. Figure 4.25 shows the best-fit S1 and S2 spectra, and 4.26 shows the values of the parameters that gave the minimum values of $\chi^2$ and gives an idea of the statistical uncertainties involved in the fit. The S1 and S2 yields for nuclear recoils resulting from the fit can be calculated using equation 4.40 and are shown in figure 4.27.

![Figure 4.25: The S1 (left) and S2 (right) experimental spectra (points) with superimposed best-fit 1σ uncertainties from Monte Carlo (right). All bins with $> 50$ counts were used in the fit.](image)

The characterization of the S1 and S2 yields via comparison to Monte Carlo relies on the integration of multiple calculations and assumptions that could introduce significant systematic errors in the final result. In fact, these systematics should prevail strongly over the statistical errors described in figure 4.26. Here we will not consider systematics due to changing operating conditions (pressure, temperature, and HV sparking and stability, for example, which are discussed in section 4.4.1). Other potential sources of systematic error include the many aspects of the Monte Carlo simulation, including inaccuracies in the calculation of the neutron spectrum (see section 4.4.3b) and in modeling the TPB and geometry for the purposes of light was produced by each recoil can be determined. The total S1 and S2 can then be adjusted based on these fractions and the change in the yield (equation 4.40) for each recoil relative to the initial constant nuclear recoil yields with which the Monte Carlo was generated.
Figure 4.26: Statistical analysis of the data/Monte Carlo fit procedure: the difference in $\chi^2$ (minimized with respect to $(\eta_1, \eta_2)$) from its absolute minimum as a function of the scale factor $\beta$ (left), and the 1σ, 2σ, and 3σ contours in $(\eta_1, \eta_2)$ space for the value of $\beta$ which minimizes $\chi^2$ overall (right). The parameter set at the absolute minimum is $(\beta = 2.76, \eta_1 = 3.525, \eta_2 = 0.7375)$.

collection. These errors are believed to be less significant than two main sources of systematic error which we will investigate here. These include the determination of the detector efficiency and the amount of gamma-induced background present in the fit region in (S1,S2) space.

The total efficiency of the detector is modeled in Monte Carlo through the construction of data-like waveforms that are then passed through the same analysis procedure as the experimental data\textsuperscript{7}. Small discrepancies between these constructed waveforms and the data and/or erroneous confidence in the trigger’s ability to pass events with S1’s visible to the reconstruction algorithm may significantly alter the low-energy portion of the S1 spectrum and therefore also the S2 spectrum. This may then be accounted in the fit for by an unphysical change in the fit parameters. In order to quantify this effect, a modified fit was performed alongside the usual fit computation. In this modified fit, if the Monte Carlo rate scale factor $\beta$ were adjusted such that the Monte Carlo S1 spectrum rose above the experimental spectrum, an additional efficiency factor would be applied to all bins from the lowest energy up to the bin at which the Monte Carlo spectrum dropped below the experimental spectrum. That is, the Monte Carlo S1 spectrum would be “efficiency corrected” such that it matched the experimental spectrum up until the bin at which the spectra matched without any correction. In the application of an efficiency factor $f$ to a bin, events were excluded from both the S1 and S2 spectra with a probability $1 - f$, and the minimization was performed using only the resulting S2 spectrum. This gave an idea of what the best $\eta_1$ and $\eta_2$ parameters would have been if the efficiency were

\textsuperscript{7}In this analysis we did not account for a threshold on the S2 signal that was later determined to have a potentially significant effect on the trigger efficiency. This should not affect, however, the alternative analysis given in section 4.4.3b
Figure 4.27: The absolute S1 (above) and S2 (below) yields for nuclear recoils in \( \sim 14 \) bar xenon gas and at a drift field of about 370 V/cm (left) determined in this study assuming Lindhard energy dependence, compared to the yields in liquid xenon at various fields from NEST [122] (right). The shaded regions show the \( \pm 1\sigma \) uncertainties for the gaseous xenon measurements, considering only the systematic errors discussed in the text, which are by far the dominant contributions. The dotted vertical lines give the recoil energies corresponding to the S1 at the lower bound of the region in (S1,S2) space selected for the fit (see figure 4.24) assuming the S1 yield is given by the mean (solid line) in the S1 yield plot, and thus the yields at all energies below these lines should be considered extrapolated, as the corresponding recoils were not actually observed in the data.

such that the S1 spectra matched for the given \( \beta \). This procedure produced results in which \( \eta_1 \) and \( \eta_2 \) were lower than the best-fit values found in section 4.4.3 by \( \sim 2 \) and 0.35 respectively, assuming rates (\( \beta \)) about a factor of 2 higher than the best-fit value.

If gamma background is causing one or more fit parameters to become biased, one
could in principle reduce the amount of gamma background at the expense of statistics by performing a tighter fiducial cut. The fit procedure was run with several different fiducial cuts, including the nominal $r < 0.3$ (see figure 4.23), $r < 1$, and $r < 0.15$. $\eta_1$ was found to vary by about amounts on the order of 1-1.5 and $\eta_2$ by amounts on the order of $< 0.1$. We anticipate in this case that gamma background may be artificially increasing the value of $\eta_1$ in particular. From the above observations, we assign a systematic uncertainty of $\sigma_{\eta_1} = 1.5$ and $\sigma_{\eta_2} = 0.20$.

It can also be determined, given the experimental live time and nearly identical cuts in the Monte Carlo and experimental datasets, what source and tagging rate would give a number of nuclear recoil events that is consistent between the two. Let $N_E(s)$ and $N_{MC}(s)$ be the total number of events in experiment and Monte Carlo passing all cuts and appearing in the spectra in figure 4.25 for a given S1 signal $s$ (similar arguments could be made for S2). Define $\eta_{NR}(s)$ to be the probability that a neutron emitted from the source produces a nuclear recoil event in the TPC with detected S1 signal $s$. We do not have a tagging or trigger efficiency in Monte Carlo, but we construct waveforms similar to those of the actual data (see section A.3), and we will assume that if enough light was placed in such a waveform for the S1-finding algorithm to identify the S1 correctly, the S1 trigger would have passed this event. That is, we assume the probability that an event is triggered if its S1 is properly found is equal to unity and thus absorb all trigger effects into our definition of $\eta_{NR}(s)$. For a tagging efficiency $\gamma$, experimental live time $T$, and NaI singles trigger rate $\gamma R = 1000$ Hz, we have

$$N_E(s) = \gamma RT \cdot \eta_{NR}(s). \quad (4.42)$$

In Monte Carlo, we also know from the beginning the total number of generated events $N_{tot,MC}$, so we can calculate

$$N_{MC}(s) = N_{tot,MC} \cdot \eta_{NR}(s). \quad (4.43)$$

The fit parameter $\alpha$ is the scaling factor by which the value in each bin of the Monte Carlo spectra must be multiplied to give the value in the experimental spectrum. This is the ratio $N_E/N_{MC}$, so

$$\beta = \frac{N_E}{N_{MC}} = \frac{\gamma RT \cdot \eta_{NR}(s)}{N_{tot,MC} \cdot \eta_{NR}(s)} = \frac{\gamma RT}{N_{tot,MC}}. \quad (4.44)$$

We can then compute the expected measured rate of tagged neutrons, using the effective live time $T_{eff}$ from section 4.4.1 in place of $T$:

$$\gamma R = \frac{\beta N_{tot,MC}}{T_{eff}} = 581 \text{ Hz}, \quad (4.45)$$

for $N_{tot,MC} = 10^8$ neutrons, $\beta = 2.76$, and $T_{eff} = 474911$ s. This is about a factor of 2 lower than expected assuming the entire NaI singles rate comes from detecting 4.4
MeV gammas. However, often the neutrons themselves or the 7.7 MeV gamma rays that accompany the lower-energy neutrons of the source spectrum produce a signal in the NaI scintillator that falls within the 4.4 MeV spectrum. Therefore our target value of $\gamma R$ is actually expected to be significantly lower than the measured singles rate of approximately 1 kHz.

b) S2-based calibration

Here the above calibration obtained from the Monte Carlo fit is compared to that obtained with an alternative methodology, a different assumption on the energy dependence of the nuclear recoil yields (constant rather than Lindhard), and a different experimental setup. Figures 4.28 and 4.29 show the experimental configuration in which neutrons were emitted from a 56 mCi Am/Be source and again the 4.4 MeV gamma ray was tagged. This is also a radioisotope neutron source for which the emitted neutron and corresponding nuclear recoil spectra can be computed with the methods discussed in section 4.2. In fact, because the $\alpha$ particles emitted by $^{241}\text{Am}$ are similar in energy to those emitted by $^{238}\text{Pu}$, the calculated spectra are nearly identical.

Figure 4.28: A schematic of the experimental setup for data taken with a $^{241}\text{Am}/\text{Be}$ neutron source. The source was surrounded by lead and polyethylene shielding, and effectively collimated along the source entrance tube leading to the TPC. The NaI scintillator was placed within the shielding to tag 4.4 MeV gamma rays.

The results obtained from this setup are likely to be different from those obtained in the previous setup with the $^{238}\text{Pu}/\text{Be}$ source in two main ways. First, there is a direct path from the neutron source to the TPC, and so the nuclear recoil spectrum should not be altered as it was in the presence of a 2-inch thick lead block (see figure 4.7). Second, the source is placed far enough away from the TPC that neutron and gamma-induced events should be distinguishable by time-of-flight. As evident
from figure 4.30, a shoulder-like feature is expected to be present in the S2 spectrum observed in this configuration.

A similar analysis was performed for this dataset, which included 3682304 events taken over a live-time of 346 hours, 3.8% of which were removed as they occurred.
within 5 minutes of a spark in the TPC. Events tagged by a valid NaI pulse were included, and only events corresponding to a time of flight of 6-40 ns were selected (see figure 4.31). The time of flight was computed as the difference in start times \( p_2 \) in equation 4.36 of the S1 and NaI pulses. A diffusion cut similar to that shown in figure 4.20 was applied, along with a fiducial cut corresponding to a cylindrical volume with \( r < 3 \) cm. The resulting distribution of S1 vs. S2 is shown in figure 4.32. The nuclear recoils, considered to be the events lying within the dotted lines in the figure, are present at low S1 and S2, and the xenon x-rays (approx. 30 keV) and gammas produced by inelastic scattering of neutrons on \(^{129}\text{Xe}\) are also shown.

Figure 4.31: (Left) Time-of-flight spectra for high-energy events (S2 > 175000 photons, 30 < S1 < 500 photons) showing two separate populations corresponding to gamma-induced and neutron-induced events. The times are shown in samples as computed in the analysis, and an offset is present due to the timing of the electronics. A fit to a sum of two Gaussians gives \( \mu_1 = 6.24, \sigma_1 = 0.429 \) (left peak) and \( \mu_2 = 4.66, \sigma_2 = 0.647 \) (right peak). (Right) Time-of-flight spectra for low-energy events (S2 < 40000 photons, S1 < 50 photons). The low-energy events are mostly neutron-induced, and no gamma peak is therefore observed. An offset of 6.4 samples was removed and the time-of-flight values were converted to ns. The range of time-of-flight values corresponding to the neutron energies emitted by the source were determined in a Monte Carlo calculation, and this region was selected along with an additional selection on each end of roughly 2\( \sigma_1 \) to account for the finite timing resolution.

The S2 spectrum plotted for nuclear recoil events with varying radial and S1-S2 cuts is shown in figure 4.33. Though not statistically significant enough to make a strong claim, if we consider the ranges denoted by dotted lines in the figure, assuming the mean \( S_{2,n} \) of the shoulder corresponds to the central value of the range and 1/2 of the width is considered to be systematic uncertainty, we can obtain a value for the S2 quenching of ionization in nuclear recoils relative to electron recoils using

\[
\alpha_2 = \frac{S_{2,n}}{(E_n/W_i) \cdot g \cdot \epsilon} = \frac{(S_{2,n}/E_n)}{(S_{2,\gamma}/E_{\gamma})},
\]

(4.46)
Figure 4.32: S2 vs. S1 for low-energy events passing tagging, time-of-flight, diffusion, and radial cuts. The plot on the right shows the same data as that on the left in contour form. The solid black line shows the centroid of the nuclear recoil band and the dotted lines denote a particular choice of cuts used to isolate nuclear recoil events. The centroid was determined by binning the S2 data in 5 bins along the S1-axis, fitting a Gaussian to the resulting peak in low S2-values, fitting the means of these Gaussian fits to a polynomial, and then repeating the procedure with the subtracted means determined by the fit polynomial, until the means $\mu$ and widths $\sigma$ of the Gaussian fits converged. The dotted lines were determined by linear fits to the points $1.5\sigma$ away from each fit value $\mu$. Nuclear recoil events are considered to be those that lie within the dotted lines.

where $E_n = 80.46$ keV is the mean energy obtained from the fit in figure 4.30, $S_{2,\gamma} = 561617$ photons is the S2 peak in gamma ray calibration spectrum (see figure 4.9), and $E_{\gamma} = 661657$ eV is the energy of the gamma to which the calibration was made. The quantities $W_i$, $g$, and $\epsilon$ are the W-value for ionization, EL gain, and light collection efficiency as discussed in section 4.3.2. By using the centroid determined in figure 4.32, one can calculate the value $S_{1,n}$ corresponding to $E_n$ as $S_{1,n} = (S_{2,n} - S_{2,0})/\lambda$, assuming the centroid line is described by $S_2 = S_{2,0} + \lambda S_1$, and the S1 quenching factor as

$$\alpha_1 = \frac{S_1^*}{(E_n/W_{sc}) \cdot \epsilon} = \frac{S_{2,n} - S_{2,0}}{\lambda S_{1,\gamma}} \cdot \left( \frac{E_{\gamma}}{E_n} \right)$$

(4.47)

The errors incurred in the above procedure are described in table 4.3. To determine the error on $E_n$, the coefficients in the Legendre expansion of the angular distributions of the $(\alpha,n)$ reactions used in the calculation were varied by 5% those of the xenon $(n,n)$ reactions were varied by 10%. These variations were determined by examining the experimental data on $(\alpha,n)$ cross sections and tabulated data on the $(n,n)$ cross sections to get an idea of their errors. Equations 4.46 and 4.47 ultimately give quenching factors of $\alpha_1 = 0.530 \pm 0.103$ (stat) $\pm 0.126$ (sys) and $\alpha_2 = 0.170 \pm 0.001$ (stat) $\pm 0.030$ (sys). The corresponding yields can be computed
Figure 4.33: S2 spectrum of nuclear recoil events for several different series of cuts. The S1-S2 cuts made in plot (1) correspond to those shown in figure 4.32. (1) \( r < 3 \text{ cm}, -1513 + 278S_1 < S_2 < 9437 + 364S_1 \) (2) \( r < 4 \text{ cm}, -1513 + 278S_1 < S_2 < 9437 + 364S_1 \) (3) \( r < 3 \text{ cm}, -3513 + 278S_1 < S_2 < 11437 + 364S_1 \) (4) \( r < 6 \text{ cm}, -3513 + 278S_1 < S_2 < 11437 + 364S_1 \). Note that the shoulder-like feature is more visible under some cuts than others. The dotted gray lines denote the region selected as the shoulder, the mean of which is assumed to correspond to the approximately 80 keV mean determined in figure 4.30.

by dividing by the appropriate W-value as \( Y_1 = \alpha_1/W_{sc} = 10.2 \pm 2.0 \) (stat) \( \pm 2.5 \) (sys) ph/keV and \( Y_2 = \alpha_2/W_i = 6.9 \pm 1.3 \) (sys) e\(^-\)/keV (the statistical error on \( Y_2 \) is \( < 0.1 \) e\(^-\)/keV). The results from both this calculation and the Monte-Carlo fit are consistent; the quenching factors computed in this section at 80.46 keVr would correspond to \( \eta_1 = 2.6 \) (compare to 3.5) and \( \eta_2 = 0.84 \) (compare to 0.74) of section 4.4.3a, both within the determined systematic errors \( \sigma_{\eta_1} = 1.5 \) and \( \sigma_{\eta_2} = 0.20 \).

The calculated nuclear recoil yields at \( E_n = 80.46 \text{ keVr} \) are compared to those of liquid xenon from NEST [122] in table 4.4. The primary scintillation yields are similar, while the ionization yield in gas is significantly higher, though due to the nontrivial energy dependence of both electron and nuclear recoil yields in liquid xenon and the lack of detailed knowledge about the energy dependence of nuclear recoils in gas, we will not use this single point to make a general statement about any advantages or disadvantages of the gas phase over the liquid phase in detecting nuclear recoils. These results encourage further study and more precise measurements in the gas phase.
Table 4.3: Error analysis for calculation of quenching factors and yields using the $^{241}$Am/Be data.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Error (stat)</th>
<th>Error (sys)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_{2,n} = 11625$ ph</td>
<td>—</td>
<td>16%</td>
<td>Width of selected range</td>
</tr>
<tr>
<td>$S_{2,0} = 3962$ ph</td>
<td>14%</td>
<td>13%</td>
<td>Fit error (stat); Est. variation (sys)</td>
</tr>
<tr>
<td>$\lambda = 321$</td>
<td>13%</td>
<td>9.3%</td>
<td>Fit error (stat); Est. variation (sys)</td>
</tr>
<tr>
<td>$S_{2,\gamma} = 561617$ ph</td>
<td>0.016%</td>
<td>—</td>
<td>Error on mean of S2 peak fit</td>
</tr>
<tr>
<td>$S_{1,\gamma} = 370.6$ ph</td>
<td>0.35%</td>
<td>—</td>
<td>Error on mean of S1 peak fit</td>
</tr>
<tr>
<td>$g = 720$ ph</td>
<td>—</td>
<td>3.3%</td>
<td>Inaccuracy in setting HV</td>
</tr>
<tr>
<td>$E_n = 80.46$ keV</td>
<td>0.61%</td>
<td>7.6%</td>
<td>Fit err.(stat); angular dist. (sys)</td>
</tr>
<tr>
<td>$W_i = 24.7$ eV</td>
<td>—</td>
<td>4.5%</td>
<td>Error on reported value</td>
</tr>
</tbody>
</table>

Table 4.4: Comparison of yields in the liquid phase (from [122]) to those determined in the gas phase for $E = 80.46$ keVr and a drift field of 370 V/cm.

<table>
<thead>
<tr>
<th>Phase</th>
<th>S1 Yield (ph/keVr)</th>
<th>S2 Yield (e$^-$/keVr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gas</td>
<td>10.2</td>
<td>6.9</td>
</tr>
<tr>
<td>Liquid</td>
<td>12.4</td>
<td>2.5</td>
</tr>
</tbody>
</table>
4.4.4 Electronic/Nuclear Recoil Discrimination

As has been shown to be the case in liquid xenon, the ability to distinguish between electronic and nuclear recoils in xenon gas lies in the ratio of $S_2/S_1$. Nuclear recoils produce less ionization relative to primary scintillation than electron recoils, and thus for nuclear recoil events this ratio is significantly lower. Figure 4.34 shows the logarithm of the ratio $S_2/S_1$ plotted against the nuclear recoil energy $\text{keV}_r$ calculated by using the $S_1$ yield determined in section 4.4.3b and assuming a constant energy dependence to convert the $S_1$ signal to a nuclear recoil energy. Two bands are apparent; the band with the higher ratio is produced by gamma events and the band with the lower ratio is made up of low energy nuclear recoil events.

![Figure 4.34: The logarithm of the ratio of $S_2/S_1$ plotted against nuclear recoil equivalent energy for events meeting all cuts, after calibration of $S_1$ axis to energy in keV$_r$. The band of events with a lower $S_2/S_1$ ratio is composed of nuclear recoils events, while the upper band is composed of events due to gamma rays, mainly due to xenon x-rays of energies $\sim 30$ keV and $40$ keV gamma rays produced in inelastic scatters of neutrons on $^{129}$Xe nuclei.](image)

While it is clear that high pressure xenon offers $S_2/S_1$ discrimination capabilities, due to low photon statistics and poor energy resolution in $S_1$ (see section A.4), we will not attempt to quantify its ability and efficiency of identifying nuclear recoils and rejecting electron recoils. Figures 4.35 and 4.36 give an idea of the consistency of the data and Monte Carlo, showing for both the logarithm of $S_2/S_1$ plotted against $S_1$ (corrected for $z$-dependence in the TPC) and $S_2$ plotted against corrected $S_1$. Note that the Monte Carlo values of $S_2$ and $S_1$ were corrected slightly from those originally generated to reflect the nuclear recoil yields determined in section 4.4.3b.
Figure 4.35: The logarithm of the ratio of S2/S1 plotted against S1 for events meeting all cuts, after calibration of Monte Carlo data using the quenching factors obtained in section 4.4.3b.

Figure 4.36: The S2 plotted against S1 signals for events meeting all cuts, after calibration of Monte Carlo data using the quenching factors obtained in section 4.4.3b.

Note the significant lack of events in the gamma bands relative to the nuclear recoil band. Relatively more nuclear recoil events are present in Monte Carlo, as events in the simulation were only recorded if they originated from some neutron interaction (elastic or inelastic scattering), and therefore the data contains more x-ray events caused by gamma background. In addition, some fraction of the neutrons at the lower energies in the source spectrum (see figure 4.4) that are emitted with a $\sim 7.7$ MeV (see table 4.2) are also being tagged in the experimental data when the 7.7 MeV gammas deposit part of their total energy in the NaI scintillator, leaving a pulse that falls within the threshold set up to tag the 4.4 MeV gammas. These lower-energy neutrons are not modeled in Monte Carlo, but have a significantly higher
cross section for inelastic scattering on $^{129}$Xe and producing a 40 keV gamma ray. Therefore we see a strong band corresponding to 40 keV gammas in the data, but not in the Monte Carlo. The nuclear recoils produced by elastic scatters of these low-energy neutrons are thus also absent from the Monte Carlo, though the majority of these recoils are lower in energy and unlikely to influence the fraction of the spectrum visible in experiment.

4.4.5 Preliminary Analysis of Drift Field Dependence in Nuclear Recoils

We briefly describe an apparent dependence of nuclear recoil yields in high pressure xenon gas on the magnitude of the drift electric field present in the region in which the nuclear recoil occurs. Data was taken using the $^{238}$Pu/Be neutron source in the configuration described in figure 4.13 under the same gain and trigger conditions as described this section, but at several values of drift electric field. Note that in this data events that occurred near PMT current trips were not removed, though this should not significantly influence the conclusions presented here.

To each dataset, tagging, diffusion, and radial cuts were applied as described in section 4.4.2 (the radial cut used here was $r < 0.6$ on the scale of figure 4.23). From the remaining events, the number of events present in 3 different regions in S1-S2 space was counted for each dataset; one region corresponding to the 40 keV gammas produced by inelastic neutron scattering on $^{129}$Xe, one corresponding to the 30 keV xenon x-rays, and one corresponding to nuclear recoil events. The three regions are shown for each additional dataset in figure 4.37. The same procedure was also applied to the main ($^{238}$Pu/Be) dataset.

The ratio $r_i$ of the number of 40 keV gammas to 30 keV x-rays and the ratio $r_n$ of nuclear recoil events to 30 keV x-rays are plotted vs. drift field in figure 4.38. There is a clear trend in the case of $r_n$ in which a relatively larger number of nuclear recoil events are present as the field increases. This could be interpreted as the decreasing of the S2 yield of nuclear recoils with decreasing electric field due to electron-ion recombination. As the field is lowered, fewer nuclear recoil events pass the S2 threshold present in the trigger. If this were the case, the ratio $r_i$ consisting of only gamma-induced events should remain approximately constant with changing drift field. Aside from the lowest-field point, this appears to be true. Though further investigation is required, the behavior shown is strong evidence for recombination effects in nuclear recoils that are sensitive to experimentally reasonable drift fields, a result encouraging to recent ideas on using recombination to determine nuclear recoil directionality [123].
Figure 4.37: S2 vs. S1 for low-energy events at varying drift field. Events corresponding to 40 keV gammas produced by neutron inelastic scattering are counted in the red dot-dashed box, events corresponding to 30 keV gammas are counted in the solid red box, and events corresponding to nuclear recoils are counted in the dashed red box.
Figure 4.38: The ratio of 40 keV inelastic events to 30 keV x-ray events, and the ratio of nuclear recoil events to 30 keV x-ray events plotted against drift electric field.

4.5 Summary

We have described in detail a simultaneous observation of the primary scintillation and ionization produced by recoils of xenon nuclei in pure xenon gas, presented a rough determination of the their scintillation and ionization yields, and confirmed that gaseous xenon is capable of discrimination of electron and nuclear recoils via the ratio S2/S1. These results establish the potential of xenon in the gaseous phase as a medium for dark matter searches and encourage more detailed measurements. A detector with improved light collection and tracking capabilities should be able to further study potential advantages xenon may have in electron/nuclear recoil discrimination and dark matter searches.
Chapter 5

Negative Ion TPCs

5.1 Introduction

We now consider a different method used to measure the ionization produced in a gaseous medium by an energetic particle - negative ion drift [124], which intends to incorporate electron attachment to negative ions followed by ion drift to a collection plane to provide an ionization measurement with optimal resolution in both location and energy. Built upon the workings of the time projection chamber (TPC) technology, this method uses the negative ion to transport the ionization that would be otherwise drifted freely, greatly reducing the amount of diffusion in the resulting signal and providing the possibility of improved energy resolution. Such improvements would provide advantages in background rejection in searches for neutrinoless double-beta decay and WIMPs.

The suitability of the technique was first demonstrated in several studies in which electrons were produced by illumination of a photocathode with a UV light source and attached to CS$_2$ molecules to form negative ions through drift in a low electric field. Later these ions were detected by removing the attached electron and in a high electric field and multiplying it in an electron avalanche. Using such a process, the drift velocity and transverse diffusion of CS$_2$ were measured in various gas mixtures [125], in which it was found that r.m.s. diffusion widths of negative CS$_2$ ions in mixtures of Ar/CH$_4$/CS$_2$ and Xe/CH$_4$ could be less than 0.15 mm over a 15 cm drift length, that is less than a factor of 10 times lower than values expected for electron drift in a similar setup. Later both transverse and longitudinal diffusion of CS$_2$ ions were confirmed to follow basic principles of thermal drift in mixtures of CS$_2$ with different amounts of Ar [126]. Similar measurements were made for longitudinal diffusion of CS$_2$ in various mixtures of He gas [127].

Measurements of energy depositions by ionizing radiation have also been performed with negative ion drift. A $^{55}$Fe pulse-height spectrum was obtained in the observation of drifted negative CS$_2$ ions [128] using the gas electron multiplier (GEM) micropattern detector [129]. This was done by detaching and multiplying the electrons
from the drifted negative ions in the high electric field of the GEM and measuring the resulting amount of charge produced. Theoretically, the intrinsic resolution dictated by the Fano factor [99] of the gaseous medium could be reached by extending this method to count the avalanche produced by each ion individually, thereby counting each individual electron produced in an ionization deposition [127, 130]. Indeed, the foundations of such an operation have been demonstrated using a large electron multiplier (LEM) or thick gem (THGEM) [131] and negative ion drift of \( \text{O}_2^- \) ions in a 70:30 Ar/\( \text{CO}_2 \) gas mixture [132] (see figure 5.1).

![Diagram of a cross section of the negative ion TPC developed at Lawrence Livermore National Laboratory. Ionization produced in an Ar/\( \text{CO}_2 \) gas mixture by 5.9 keV x-rays from a \( ^{55}\text{Fe} \) source was attached to \( \text{O}_2 \) molecules and drifted to a readout plane where the electron was detached and multiplied in a large electron multiplier (LEM). A simulation of the LEM avalanche process performed in Garfield [133] is shown at the right.](image)

Here we seek to present a concise and semi-quantitative description of the negative ion detection process, beginning after the attachment of the electrons to the chosen electronegative impurity, meant to guide the selection of the ion/neutral system and field configuration used to remove and multiply the electron from the negative ion. Methods of estimating ion mobility and diffusion are discussed along with a model of the ion detachment process, and these concepts are then applied to describe various aspects of negative ion detectors.

### 5.2 Physics Concepts

Here we rely on kinetic theory and statistics to concisely describe the physics of negative ion drift without invoking a more complex and precise mathematical formalism, such as the solution to the Boltzmann transport equation. Our goal is, for a given
ion drifting in a given neutral gas, to be able to calculate to a reasonable approxi-
mation the mobility and diffusion coefficients, and to be able to understand how and
where the ion will detach in different detector geometries. Momentum transfer theory
(MTT) uses arguments of conservation of momentum and energy to deduce equations
for mobility and diffusion coefficients under certain simplifying approximations, and
most of the discussion and notation used here will be based on the treatment of this
theory given in [134].

5.2.1 Transport Properties of Ions

We will begin from the picture of a negatively charged ion drifting in some electric
field $E(x)$ through a neutral gas of particles (atoms or molecules) moving in thermal
equilibrium at temperature $T$. We assume that $E(x)$ is constant between collisions.
During its journey through the gas, the ion may undergo a collision with a single
neutral particle resulting in one of two possible outcomes; the ion and neutral particle
will scatter elastically and isotropically, and continue on with their trajectories, or
the excess electron will be detached from the negative ion, and the ion will cease to
drift in the electric field. Other outcomes such as resonant charge exchange [134], in
which an electron is passed from the ion to a neutral particle, and other mechanisms
of inelastic scattering, are not considered here. Only collisions involving exactly one
ion and exactly one neutral particle will be considered, and it is thus assumed that the
local concentration of ions is low enough that no two ions interact appreciably. We
consider first the elastic collisions, ignoring the possibility of detachment until section
5.3, and show how the information obtained from elastic scattering cross sections can
produce estimates of mobility and diffusion in a given configuration.

In the following expressions for the collision cross section, mobility, and diffusion
coefficient in an ion/neutral system, the assumption of low fields is implied. One can
take this to mean that once the field is high enough that the average energy of the
drifting ion becomes greater than the thermal energy $(3/2)kT$ of the medium, these
expressions lose accuracy. In negative ion TPCs, this is only likely to occur at the
fields reached during the ion detachment process.

Collision Cross Section

The atomic physics involved in accurately describing the ion/neutral collisions is
formidable, and a full quantum-mechanical treatment of the subject will not be dis-
cussed here. The details of the elastic ion collisions will be approached from the
perspective of the cross section as determined by approximate classical methods. We
consider the motion of a negative ion in a gas of neutral particles and define the total
elastic collision cross section $\sigma(\varepsilon)$, at a given mean ion-neutral center of mass energy
$\varepsilon$, by writing the frequency of elastic collisions $\nu(\varepsilon)$ [134]

$$\nu(\varepsilon) = N\sigma(\varepsilon)\bar{v}_r.$$  \hspace{1cm} (5.1)
Here $N$ is the number of neutral atoms per unit volume and $\bar{v}_r$ is the magnitude of the relative velocity of the ion-neutral system, that is if we denote the ion mass as $m$ and velocity as $v$ and the neutral mass as $M$ and velocity as $V$, then $\bar{v}_r = |v - V|$ and $\bar{\varepsilon} = \frac{1}{2} \mu \bar{v}_r^2$, where $\mu = mM/(m + M)$ is the reduced mass of the ion-neutral system.

We will now choose a particular approximation discussed in [135] in which to calculate $\sigma$ that is valid at near-thermal energies $\bar{\varepsilon}$ and expresses $\sigma$ in terms of information that is straightforward to obtain for many ion-neutral systems. We ignore all interactions between the ion and neutral atom except those arising from the polarization of the neutral by the ion. Such an interaction is described by the radial potential

$$V(r) = -\frac{e^2 \alpha_d}{2r^4},$$

(5.2)

where $e$ is the electron charge and $\alpha_d$ is the dipole polarizability of the neutral atom. $\alpha_d$ can be expressed in units of cm$^3$, in which case $e$ must be expressed in the cgs unit of esu. The resulting cross section is given in as $\pi b_0^2(\varepsilon)$, where $b_0(\varepsilon)$ is the impact parameter at which the orbits of the polarization potential become unstable at a given energy of approach $\varepsilon$, and can be written [135]

$$\sigma_{\text{pol}}(\varepsilon) = \frac{1}{2} \pi \sqrt{\frac{2e^2 \alpha_d}{\varepsilon}}.$$

(5.3)

Note that $\sqrt{\bar{\varepsilon}}\sigma_{\text{pol}}$ only depends on the polarization of the neutral particle and not on any specific property of the negative ion. For our purposes we will evaluate $\sigma_{\text{pol}}$ at an average ion energy $\bar{\varepsilon}$ rather than an instantaneous energy $\varepsilon$, noting that the results of this approximation will depend on the underlying distribution of $\varepsilon$ that gives rise to $\bar{\varepsilon}$; a distribution which may change with the magnitude of the applied field in which the ion is drifting.

In equation 5.3, we claim to have packaged all of the necessary information to describe approximately the elastic collisions in an ion-neutral system. We will now use it to show how the mobility and diffusion of an ion/neutral system can be quantified approximately knowing only the masses of the ion and neutral and the polarizability $\alpha_d$ of the neutral, quantities available for a number of atoms and molecules (see for example [134]).

**Mobility and Diffusion**

MTT produces an expression for the mobility $K(\bar{\varepsilon})$ of negative ions in a neutral medium in terms of the average ion energy and $\sigma$ [134], and using $\sigma \rightarrow \sigma_{\text{pol}}$ from equation 5.3,

$$K = \frac{e}{N} \left( \frac{1}{2\mu \bar{\varepsilon}} \right)^{\frac{1}{2}} \frac{1}{\sigma(\bar{\varepsilon})} = \frac{1}{2\pi N} \left( \frac{1}{\mu \alpha_d} \right)^{\frac{1}{2}}.$$

(5.4)
Table 5.1: Comparison of reduced mobilities calculated in the polarization limit to those measured in negative ion TPCs. The polarizabilities in Å³ used in the calculations are as follows: $\alpha_{d,\text{Ar}} = 1.642$, $\alpha_{d,\text{CH}_4} = 2.59$, $\alpha_{d,\text{Xe}} = 4.044$, $\alpha_{d,\text{CF}_4} = 3.86$, $\alpha_{d,\text{He}} = 0.2050$, and $\alpha_{d,\text{Ne}} = 0.3946$, all from [134], and $\alpha_{d,\text{CS}_2} = 8.21$ from [138]. Note that for $\text{CS}_2$, which has differing polarizabilities along different axes, the mean polarizability is used.

<table>
<thead>
<tr>
<th>System</th>
<th>$\mu N,_{\text{calc}}$ (cm²-torr/V·s)</th>
<th>$\mu N,_{\text{expt}}$ (cm²-torr/V·s)</th>
<th>% Difference</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{Ar}/\text{CH}_4/\text{CS}_2$ (9:1:14.5)</td>
<td>953</td>
<td>915.0 ± 38</td>
<td>4.2</td>
<td>[125]</td>
</tr>
<tr>
<td>$\text{Xe}/\text{CS}_2$ (10:14.5)</td>
<td>773</td>
<td>456.0 ± 38</td>
<td>69.6</td>
<td>[125]</td>
</tr>
<tr>
<td>$\text{CS}_2$</td>
<td>707</td>
<td>410.4 ± 20</td>
<td>72.3</td>
<td>[139]</td>
</tr>
<tr>
<td>$\text{CS}_2/\text{CF}_4$ (3:1)</td>
<td>762</td>
<td>456.0 ± 20</td>
<td>67.2</td>
<td>[139]</td>
</tr>
<tr>
<td>$\text{CS}_2/\text{CF}_4$ (1:1)</td>
<td>827</td>
<td>524.4 ± 20</td>
<td>57.7</td>
<td>[139]</td>
</tr>
<tr>
<td>$\text{CS}_2/\text{CF}_4$ (1:3)</td>
<td>904</td>
<td>615.6 ± 20</td>
<td>46.8</td>
<td>[139]</td>
</tr>
<tr>
<td>$\text{CS}_2/\text{Ar}$ (7:1)</td>
<td>734</td>
<td>448.4 ± 20</td>
<td>63.6</td>
<td>[139]</td>
</tr>
<tr>
<td>$\text{CS}_2/\text{He}$ (7:1)</td>
<td>802</td>
<td>456.0 ± 20</td>
<td>76.0</td>
<td>[139]</td>
</tr>
<tr>
<td>$\text{CS}_2/\text{Ne}$ (7:1)</td>
<td>792</td>
<td>463.6 ± 20</td>
<td>70.8</td>
<td>[139]</td>
</tr>
</tbody>
</table>

This expression now depends on the masses of both the ion and neutral particle in addition to the neutral polarizability and is independent of the collision energy $\bar{\varepsilon}$. Employing cgs units, the value given by equation 5.4 is in units of cm²(sV)⁻¹s⁻¹, where (sV) is the statvolt. To convert the value yielded by equation 5.4 to the SI units m²V⁻¹s⁻¹, one must multiply it by 100/c, where c is the speed of light in m/s.

Figure 5.2 shows how the experimentally measured mobilities of several systems compare to that calculated by equation 5.4. In general, there is agreement at low fields, though as $E/N \sim 10$ Td, the mobility may increase or decrease depending on the system. Note that at a pressure of 1 bar and a temperature of 293 K, fields of interest to ion drift ($E \lesssim 5$ kV/cm) correspond to $E/N \lesssim 20$ Td.

One may often be interested in using neutral gas mixtures as opposed to single-component neutral gases in a negative ion TPC. In this case it is useful to have an approximate expression for the mobility of an ion in the mixture given equation 5.4 for a single component. The equation we will use is Blanc’s law [134],

$$\frac{1}{K_{\text{mix}}} = \frac{1}{N} \sum_j N_j K_j = 2\pi N \sum_j N_j \sqrt{\mu_j \alpha_{d,j}}.$$  \hfill (5.5)

where $N_j$ and $K_j$ are the number density and mobility of the $j$th component. To give an idea of the results one may expect table 5.1 compares several reduced mobilities measured in negative ion TPCs to the values obtained using equation 5.5.

Once the mobility of an ion-neutral system is known, its transverse and longitudinal diffusion coefficients, $D_T$ and $D_L$, can be found in the approximation of low field using the Nernst-Einstein-Townsend relation (see [134] for example). They will be
Figure 5.2: Reduced mobility $NK$ for several ion/neutral systems, where $N$ is the gas density. The horizontal dot-dash line shows the calculated value using equation 5.4, and the points are experimental data, from [136] ($\text{O}_2^-/\text{O}_2$, $\text{C}_2\text{H}_2^-/\text{He}$, and $\text{CO}_3^-/\text{Ar}$), and from [137] ($\text{F}^-/\text{Ar}$). The polarizabilities used in the calculations are, in Å$^3$, $\alpha_{d,\text{Ar}} = 1.642$, $\alpha_{d,\text{O}_2} = 1.57$, and $\alpha_{d,\text{He}} = 0.2050$, all from [134].
equal, with $D_T = D_L \equiv D$ and

$$D = \frac{kT}{eK} = \frac{kT}{2\pi eN} \left( \frac{1}{\mu \alpha_d} \right)^{\frac{1}{2}},$$  \hspace{0.5cm} (5.6)

where we have substituted equation 5.4 for $K$. In this approximation, the longitudinal and transverse coefficients are both equal to $D$.

### 5.2.2 Detachment of Negative Ions

Once the processes by which the negative ion arrives at the readout location are understood, the detachment of its excess electron becomes the next subject of concern. Detachment could occur by one of several processes which involve the presence of a high electric field, collisions between the ion and neutral gas particles, or both. In many cases it is useful to model the electron-ion system as a single electron bound in a zero-range potential or “ZRP” (see [140] and references therein), that is a potential of the form $V(r) \propto \delta(r)$. The constant of proportionality is chosen such that the single bound state of the system has an energy $E_0 = E_A$, where $E_A$ is the electron affinity of the system. Especially for ions with a low value of $E_A$ in which the excess electron is weakly bound, this “ZRP” model which effectively assumes that the influence of the nucleus is concentrated at $r = 0$, is likely to be a good approximation. We will first discuss, within this model, electron detachment as a result of solely the application of an external electric field, ignoring the presence of the neutral atoms. Though the ZRP model can also be used in solving the problem of electron detachment due to energetic collisions between the negative ion and neutral gas particles the knowledge of the potential of interaction between the ion and neutral must be known at all values of $r$ relevant to the collision (see [141] for example).

#### Field-Induced Detachment

The introduction of an external electric field [140] to the system adds an imaginary component to the bound state energy $E_0 \rightarrow E' = i\frac{\Gamma}{2}$. For a negative ion in a bound state denoted by $|\phi_0\rangle$, from elementary quantum mechanics the state after a time $t$ will be $|\phi(t)\rangle = e^{-iEt/\hbar} |\phi_0\rangle$. The probability of detachment, that is finding this time-evolved state in some state other than the original bound state $|\phi_0\rangle$ is

$$P(t) = 1 - |\langle \phi_0 | \phi(t) \rangle|^2 = 1 - e^{-\Gamma t/\hbar}.$$  \hspace{0.5cm} (5.7)

For the negative ion bound in a ZRP, we quote the result for $\Gamma$ from [140]

$$\Gamma = \frac{e\varepsilon}{4 \cdot (2m_eE_0)^{1/2}} \exp \left( -\frac{2^{5/2}m_e^{1/2}E_0^{3/2}}{3e\varepsilon} \right),$$  \hspace{0.5cm} (5.8)
where $m_e$ is the electron mass, and the equation is valid under the condition $\gamma \equiv (\hbar^2e^2/2m_e)^{1/3}(\delta^{2/3}/E_0) \ll 1$. We denote the lifetime for field-induced detachment by writing the exponential in equation 5.7 as $e^{-t/\tau_F}$, so that $\tau_F = \hbar/\Gamma$. To cast this into a more practical form, we have for $\delta'$ given in kV/cm and $E_0$ in eV

$$\tau_F = 0.1349 \cdot \frac{\sqrt{E_0}}{\delta'} \exp \left( 6.831 \times 10^4 \cdot \frac{E_0^{3/2}}{\delta'} \right) \text{ ns,}$$

valid for $\gamma = 7.249 \times 10^{-4}(\delta^{2/3}/E_0) \ll 1$. Figure 5.3 shows several curves of field-induced detachment lifetimes $\tau_F$ produced using the electron affinities of several ions. It is clear from these calculations that large fields ($\sim$ MV/cm) are required to have any reasonable chance of detaching an electron due to the field alone. Other mechanisms such as collisional detachment are more likely to be relevant to negative ion based detectors.

Collisional Detachment

When an energetic negative ion collides with a neutral atom or molecule, the excess electron of the negative ion may be forced out of its bound state and into the continuum of free states. This process of collisional detachment can be modeled quantum
mechanically, though the analysis is more complex than in the case of field detachment. Multiple atoms are now involved which contain multiple electrons whose interactions dictate the dynamics of the collision and the instantaneous binding energy of the excess electron in the negative ion.

Here we present an approach that captures the basic physics of collisional detachment without resorting to quantum mechanical calculations. The ion-neutral interaction is replaced by a single parameter $\varepsilon_D$ that can be adjusted to match experimental findings. The parameter $\varepsilon_D$ is the threshold energy for electron detachment, such that if an ion-neutral collision occurs with center of mass energy $\varepsilon > \varepsilon_D$, the excess electron is detached from the ion. This is similar to the approximation of Dion et. al. [142] which assumes that collisional detachment of a negative ion occurs when the center of mass energy $E$ of a single ion-neutral pair becomes greater than or equal to the electron affinity $E_A$ of the ion. Dion et. al. argue that the field required to meet this condition must be able to accelerate the ion to $E = E_A$ in between two collisions, as the ions will remain at thermal energies during drift. By examining experimental data on collisional ion detachment (see [143] for a compilation of detachment cross sections), it can be seen that the first of the above statements is incorrect. The electron affinity alone does not dictate the threshold for detachment, but it is set by the interactions of the electrons in the ion and the neutral atom during the collision. We will continue the analysis from the perspective of the second statement, that the ion moves between collisions starting effectively from rest, but will use a threshold energy of detachment we call $\varepsilon_D$. This threshold would not be strict in a quantum mechanical sense but could be thought of as the energy at which a collision can occur such that the potential energy binding the ion to the neutral molecule becomes 0 [141]. One should be able to calculate this value from first principles, though here we will leave it as a free parameter to be fit to a specific experimental setup.

We begin by revisiting the picture of a negative ion moving in a gas of neutral molecules as described in section 5.2.1. According to standard collision theory (see e.g. [115]) we define a constant collision rate $\nu = \tau^{-1}$ such that the probability that a particle undergoes a collision in time $dt$ is $\tau^{-1}dt$. This rate in general is given by equation 5.1 and depends on the mean energy of a particle, but as the instantaneous energy $\varepsilon$ of a particle varies frequently during the course of its travel through the gas, we will assume that for a given mean energy $\bar{\varepsilon}$, we have a constant $\tau^{-1}$. This mean energy corresponds to a particular value of the drift velocity $v_d$ of the ion in the neutral gas. Using this collision rate one can express the probability $P_c(t)$ that a particle undergoes a collision after time $t$ in an interval $[t, t+dt]$ as [115]

$$P_c(t)dt = e^{-t/\tau} \frac{dt}{\tau}. \quad (5.10)$$

In a time $t$ the ion will be accelerated by the field to an energy of

$$E = \frac{1}{2} mv^2 = \frac{1}{2m} e^2 \bar{\varepsilon}^2 t^2 \quad (5.11)$$
where we have used \( v = at = eE t / m \). We will now assume that the neutral atom is at rest, a valid assumption if \( kT \ll \varepsilon_D \). In this case, the relative velocity of the ion-neutral system will be the velocity of the ion, \( \bar{v}_r = v \). In the center of mass frame the total energy will be

\[
\varepsilon = \frac{1}{2} \mu \bar{v}_r^2 = \frac{\mu}{m} E
\]

Thus the detachment condition becomes \( E = \frac{m}{\mu} \varepsilon_D \) and

\[
t_d = \frac{m}{eE} \sqrt{\frac{2 \varepsilon_D}{\mu}}
\]

is the amount of time the ion must accelerate without elastic collision to gain enough energy to undergo detachment in its next collision. Note that in [142] it is assumed that \( \mu = m/2 \) so the detachment condition is \( \varepsilon = 2 \varepsilon_D \). Therefore the probability \( P_d(t) \) that the negative ion, having just undergone a collision at \( t = 0 \), will undergo detachment during its next collision is

\[
P_d(t) = \int_{t_d}^{\infty} \frac{e^{-t/\tau}}{\tau} dt = e^{-t_d/\tau}.
\]

From this we obtain a new collision rate for detachment equal to the original collision rate multiplied by the probability for detachment, yielding the lifetime for detachment,

\[
\tau_d = \tau e^{t_d/\tau}.
\]

In the limit \( t_d \gg \tau \) the mean detachment lifetime becomes quickly very long, while when \( t_d \approx 0 \), the mean lifetime for detachment is equal to the mean lifetime for collision, and every collision produces detachment. Note that \( \tau \) is not the lifetime for elastic collisions, \( \tau_{el} \), but the total mean lifetime that includes both elastic and detachment collisions. Therefore \( \tau \) contains information about \( \tau_d \), and we must further manipulate equation 5.15 to isolate \( \tau_d \) from the corresponding elastic lifetime \( \tau_{el} = \nu(\bar{v})^{-1} \). In the simplest approximation, we can consider only the regime in which detachment collisions are rare and all collisions contributing to \( \tau \) are elastic. In this case we will have

\[
\tau_d = \tau_{el} e^{t_d/\tau_{el}},
\]

where using \( \bar{v}_r = \sqrt{2\varepsilon / \mu} \), the mean lifetime for elastic collisions is

\[
\tau_{el} = \frac{1}{N\sigma(\bar{v})} \sqrt{\frac{\mu}{2\varepsilon}}.
\]

In the more general case, if we take \( \tau \) to include detachment collisions in equation 5.15, then we will have a cross section for elastic collisions and a cross section for
detachment collisions $\sigma_d$. These effective areas add to give the total cross section so that $\tau^{-1} = \bar{v}_r N (\sigma_d + \sigma_{el})$, or $\tau^{-1} = \tau_{el}^{-1} + \tau_d^{-1}$. Defining $y \equiv 1 + \tau_d^{-1}/\tau_{el}^{-1}$, inverting and adding 1 to both sides of equation 5.15, and solving for $y$ yields

$$y \left(1 - e^{-t_d y/\tau_{el}}\right) = 1. \quad (5.18)$$

This equation must be solved numerically, but the solutions can be approximated as

$$y \approx \frac{1}{1 - e^{-t_d/\tau_{el}}}, \quad t_d/\tau_{el} \gg 1 \quad (5.19)$$

$$y \approx \frac{1}{\sqrt{t_d/\tau_{el}}}, \quad t_d/\tau_{el} \ll 1. \quad (5.20)$$

Note that in the limit $t_d/\tau_{el} \gg 1$, equation 5.19 gives equation 5.16. In using this approximation, a < 20% error on $\tau_d$ is produced for all values $t_d/\tau_{el} \gtrsim 0.5$. Once $y$ is known, the mean detachment lifetime can be written

$$\tau_d = \frac{\tau_{el}}{y - 1}. \quad (5.21)$$

Invoking the same arguments as those that lead to equation 5.7, we can write a probability of ion detachment after traveling for time $t$. By changing variables from time $t$ to position $z$ in a field configuration, one can determine a detachment probability distribution for a given detector geometry. However, since detachment collisions do not occur as frequently as elastic collisions, the applied field may vary significantly between them. We therefore define a position-dependent mean detachment length,

$$\chi(z) \equiv v_{dr}(z) \tau_d(z), \quad (5.22)$$

where $v_{dr} = KE$, and follow the arguments set forth in [115] that lead to equation 5.7. Let $P(z; z_0)$ be the probability that the ion drifts without detachment to a location $z$ from an initial location $z_0$. The probability of detaching in an infinitesimal interval $[z, z + dz]$ is $dz/\chi(z)$, and so the probability that an ion survives until $z$ and then detaches immediately in this small interval is $P(z; z_0) dz/\chi(z)$. The ion will have a lower survival probability $P(z + dz; z_0)$ after passing through the infinitesimal interval by exactly this amount, and thus

$$dP(z; z_0) = -\frac{P(z; z_0) dz}{\chi(z)}. \quad (5.23)$$

Integrating this equation, we find that the probability of drifting without detaching from $z_0$ to $z$ is

$$P(z; z_0) = \exp\left(-\int_{z_0}^{z} \frac{dz'}{\chi(z')}\right). \quad (5.24)$$
Thus for some field configuration $E$, the probability of detaching after drifting a distance $z$ beginning at $z_0$ is equal to $1 - P(z; z_0)$. To obtain the probability of detachment in an interval $[z, z + dz]$, which we will call $p(z) \equiv P(z; z_0)/\chi(z)$, we substitute equation 5.24 into equation 5.23 to obtain

$$p(z) = \frac{1}{\chi(z)} \exp \left( - \int_{z_0}^{z} \frac{dz'}{\chi(z)} \right).$$ (5.25)

At this point we have a model that predicts detachment probabilities in a one-dimensional field configuration for an ideal system consisting of elastic collisions and, when the ion is capable of gathering an amount of energy between two collisions that surpasses some threshold, detachment collisions. Due to the many assumptions, it is unlikely that this model will yield accurate quantitative results for real systems. However, by examining the results of the model for different field configurations, we can gain some insight into what experimental conditions produce optimal detachment for a collision-based detachment process. To simplify the results further, we write equation 5.24 in the polarization limit discussed in section 5.2.1. The polarization cross section is not valid at the higher energies that produce detachment, though the comparison of the results across different geometries should yield some insight. Using equations 5.1, 5.4, 5.16, 5.17, and 5.22, the approximate detachment length is

$$\chi(z) = \frac{\mathcal{E}(z)/N}{4\pi^2 eN\alpha_d} \exp \left( \frac{2\pi m \sqrt{2\varepsilon_D \alpha_d}}{\mu [\mathcal{E}(z)/N]} \right).$$ (5.26)

Already from this equation we note two important points that should hold in general for negative ion TPCs. The first is that $\chi(z) \propto (\mathcal{E}/N)/N$, while only $\mathcal{E}/N$ appears in the exponent. By examining equation 5.24, we see that increasing $N$ decreases the probability of survival if $\mathcal{E}/N$ is held fixed. Therefore increasing the pressure in a negative ion TPC should increase the probability of ion detachment, provided that the high electric field used for detachment can be increased to keep $\mathcal{E}/N$ constant - often a difficult task. If $\mathcal{E}$ is allowed to remain constant, the influence of $N$ in the exponential is dominant over the increase in the prefactor, and overall the probability of survival increases, thus producing less detachment.

The second point involves the factor $\mu/m\sqrt{\varepsilon_D}$ in the exponential. This factor remains in general even outside of the use of the polarization cross section to describe the elastic collisions. The greater this factor is made, the greater the probability of detachment, and so it may be taken as an indicator of the potential for an ion/neutral system to be used in a negative ion TPC.

### 5.3 Detector Concepts

The concepts of section 5.2 will be applied here to model various aspects of detectors that rely on negative ion drift.
5.3.1 Drift and Diffusion

In a negative ion TPC, the detected radiation interacts with the neutral gas and produces ionization over a length scale that we will assume to be much smaller than the drift length $d$ over which the ionization will then be drifted once attached to a chosen electronegative agent. The attachment occurs over some characteristic length $\alpha_a$, which we will also assume to be much less than $d$. The details of the ionization and attachment processes will be unique to the medium, though both can be understood and quantified for a given medium using the Monte Carlo techniques supported by the Garfield [144] and Magboltz [145] computer codes.

We will not discuss the ionization and attachment processes in detail here, though we will assume that because both occur over distances much less than $d$, we can consider the initial distribution of ions to be that of a single point, starting at drift time $t = 0$. The mean drifting time $t_{dr}$ of the ions can then be found by equation 5.4 as $t_{dr} = d/K \varepsilon$,

$$t_{dr} = \frac{2\pi d}{\varepsilon/N} \sqrt{\mu \alpha_d}, \quad (5.27)$$

where we have assumed a single-component neutral gas, though $K_{mix}$ could be used instead of $K$ for a multi-component neutral gas. The drift time $t_{dr}$ will dictate how long the ions will diffuse before encountering the readout plane, thus determining in part the spread in arrival times and position resolution. For an initial pointlike distribution, the drifting ions will diffuse in a time $t$ such that the RMS displacement from the initial point in one dimension can be found as $\sqrt{x^2} = \sqrt{2Dt}$, where $D$ is the diffusion coefficient of the medium and $t$ is the drift time (see for example [134] or [37]). Using formula 5.6 and $t = t_{dr}$, we find that the spread in ion arrival time $\Delta t \equiv \sqrt{x^2/v_{dr}}$ can be written

$$\Delta t = t_{dr} \sqrt{\frac{2kT}{e\varepsilon d}}. \quad (5.28)$$

As long as the energy $e\varepsilon d$ imparted to the ion by the external field during drift is much greater than the thermal energy $kT$, the spread of arrival times due to diffusion will be a small fraction of the drift time.

5.3.2 Negative Ion Detection

Equation 5.26 will now be applied to several field configurations relevant to potential detector geometries. To simplify the notation, we will define for each configuration a field $\varepsilon'(z) = \varepsilon_0 f(z)$, where $\varepsilon_0$ is a constant characteristic of the field magnitude and $f(z)$ is 1-dimensional function yielding a dimensionless factor at each location $z$.
within the geometry. We also define

$$\chi_0 \equiv \frac{\mathcal{E}_0}{4\pi^2 e N \alpha_d}, \quad \beta \equiv \frac{\mathcal{E}_0}{2\pi \sqrt{2\varepsilon_D} \alpha_d} \left( \frac{\mu}{m} \right),$$  

(5.29)

so that equation 5.26 can be written as

$$\chi(z) = \chi_0 f(z) \exp \left[ (\beta f(z))^{-1} \right].$$  

(5.30)

In the following examples, we will use parameters corresponding to $O_2^-$ ions drifting in Ar; $m = 32.0$ amu, $M = 39.9$ amu from [146], $\alpha_d = 1.642 \text{ Å}^3$ from [134], and we select $\varepsilon_D = 3$ eV. The calculations will be done with $T = 293$ K and $N$ corresponding to that of an ideal gas at a pressure of $P = 190$ torr.

**Detection with a Constant Field**

Consider a constant field such as that produced by a parallel plate capacitor of thickness $d$ and applied voltage $V$. Such a configuration would be relevant for detectors such as the micromegas [147]. We take the characteristic field $\mathcal{E}_0 = V/d$, and $f(z) = 1$, yielding a constant $\mathcal{E}(z) = \mathcal{E}_0$. Using equations 5.24 and 5.29, we find that the total probability for detachment is

$$P_{\text{tot}} = 1 - \exp \left( -\frac{d}{\chi_0} e^{-\beta} \right).$$  

(5.31)

This result shows that the detachment probability will increase sharply near the threshold as expected, and that it is very sensitive to $\beta$. From equation 5.25, the detachment distributions will be exponentials, as shown in figure 5.4 for $d = 0.4$ mm. For any detachment that occurs, the instantaneous probability of detachment is greatest at $z = 0$.

**Detection with a Wire**

Another field configuration of interest is that of a single wire, $\mathcal{E}(r) \propto 1/r$, where $\mathcal{E}$ is entirely radial in direction. Assigning the characteristic field to be the highest obtainable field with a wire of radius $a$, $\mathcal{E}_0 = \mathcal{E}(a)$, the expression for the field becomes

$$\mathcal{E}(r) = a \mathcal{E}_0 / r,$$  

(5.32)

so that, now using $r$ in place of $z$ in equation 5.30 $f(r) = a/r$. In this case the solution to equation 5.24 is analytical, and is

$$P_{\text{tot}} = 1 - \exp \left( -\frac{a}{\chi_0} \beta(1 + \beta) e^{-\beta} \right).$$  

(5.33)
Figure 5.4: Detachment probability distributions for a constant field over a distance $d = 0.4 \text{ mm}$ at several different field strengths. The integrated total detachment probabilities for each field are approximately 97% (15 kV/cm), and 100% (25 kV/cm and 45 kV/cm).

The result is similar to the constant field case, except for the additional factor of $\beta(1 + \beta)$ in the exponential. Since in the regime leading up to detachment $\beta < 1$, and $\beta(1 + \beta) < 1$ for $\beta < 0.618$, the wire geometry has an additional disadvantage in detachment relative to the constant field, as is evident in the fact that the highest field in the wire configuration occurs only at the surface of the wire while it is maintained throughout a distance $d$ for the constant field case.

Detection with a Large Electron Multiplier

A large electron multiplier (LEM), a larger design of the gaseous electron multiplier (GEM) [129], is a micropattern detector used for amplification of single electrons via electron avalanche.

Consider a LEM with thickness $d$, pitch $a$, and hole radius $r$. The electric field along the axis through the center of a LEM hole will by symmetry arguments be directed along the axis and can be expressed approximately as
Figure 5.5: Detachment probability distributions for a wire field configuration for wire radius $a = 50 \mu m$. The integrated total detachment probabilities for each field are approximately 9% (15 kV/cm), 62% (25 kV/cm), and 99% (45 kV/cm).

$$\varepsilon_z(z) = \frac{\Delta V}{2d} \left( 1 + \pi \left( \frac{r}{a} \right)^2 \right) \times \left[ \frac{z - d/2}{\sqrt{r^2 + (z - d/2)^2}} - \frac{z + d/2}{\sqrt{r^2 + (z + d/2)^2}} \right]$$

(5.34)

where $\Delta V$ is the voltage across the LEM and $z = 0$ is the center of the LEM hole. Choosing $\varepsilon_0 = V/d$ as in the constant field case leaves a more complex expression for $f(z)$. Some examples of detachment probability distributions for the LEM are shown in figure 5.6. The total detachment probability will be solved numerically in this case, and since the fields of a LEM are similar to the fringing fields of a parallel plate capacitor, we would expect a somewhat lower detachment probability than in the constant field case, though for small enough hole size still better than the wire case. In addition, as the field within a LEM hole is smallest on its axis, we should expect a somewhat higher detachment probability than calculated.

### 5.3.3 Choice of Capture Agent and Geometry

Thus far it has been shown that several molecules may function as the electronegative agents in a negative ion detector, including $\text{CS}_2$ and $\text{O}_2$ (as previously discussed) and
Figure 5.6: Detachment probability distributions for a LEM with pitch $a = 0.7$ mm, hole radius $r = 0.15$ mm, and thickness $d = 0.4$ mm. The integrated total detachment probabilities for each field are approximately 72% (15 kV/cm), and 100% (25 kV/cm and 45 kV/cm).

nitromethane CH$_3$NO$_2$ [148]. The model described in section 5.2.2 does not include the information necessary to calculate from first principles the detachment lifetime of a given ion/neutral system in an arbitrary electric field. However, it does provide some guidance in what properties of a system are likely to facilitate detachment. The ratio $\mu/m\sqrt{\varepsilon_D}$ discussed in section 5.2.2, for example, is important in determining the total detachment lifetime. Though we do not know the underlying details of the interactions between the ion and neutral, if the detachment is dictated by a strict energy threshold as assumed in the model, this ratio should give the correct dependence on that threshold. By assuming $\varepsilon_D$ roughly follows the electron affinity, we can calculate this ratio and have some level of guidance in choosing an ion/neutral system, though a single system will not be recommended here based solely on this ratio. Examples for several systems are given in table 5.2.

For the examples presented in section 5.3.2, the total detachment probability is shown for different values of the characteristic field $\mathcal{E}_0$ in figure 5.7. From this it is clear that the constant field configuration provides optimal detachment between the three possible geometries with the selected dimensions. It is nevertheless difficult to determine which geometry would be optimal in the general case because the values of $\mathcal{E}_0$ obtainable may also vary between geometries. Note that $\beta < 0.6$ for all values of the field $\mathcal{E}_0$ shown in figure 5.7, so that equation 5.26 is a valid approximation.
Table 5.2: Comparison of detachment factors $\gamma = \mu/m\sqrt{\varepsilon_D}$ across several ion/Ar and ion/Xe systems. Atomic masses are taken from [146], and electron affinities from [149].

<table>
<thead>
<tr>
<th>Ion</th>
<th>Mass (amu)</th>
<th>EA (eV)</th>
<th>$\gamma_{\text{Ar}}$</th>
<th>$\gamma_{\text{Xe}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO$^-$</td>
<td>30.01</td>
<td>0.026</td>
<td>3.54</td>
<td>5.05</td>
</tr>
<tr>
<td>O$_2^-$</td>
<td>32.00</td>
<td>0.45</td>
<td>0.83</td>
<td>1.20</td>
</tr>
<tr>
<td>CS$_2^-$</td>
<td>76.14</td>
<td>0.80</td>
<td>0.38</td>
<td>0.71</td>
</tr>
<tr>
<td>CH$_3$NO$_2^-$</td>
<td>61.04</td>
<td>0.26</td>
<td>0.78</td>
<td>1.34</td>
</tr>
</tbody>
</table>

Figure 5.7: Total detachment probability for an ion drifting through the entire field configuration of three different detector geometries. The ion/neutral system and corresponding parameters are the same as those used throughout section 5.3.2, and the dimensions of the geometry are the same as those in figures 5.4, 5.5, and 5.6.
5.4 Summary

Basic methods for calculating the mobility and diffusion of ions in a neutral gas mixture have been presented and produce accurate results at low fields for some systems, though others show significant deviation, possibly due to inaccuracies of the polarization approximation. The ion detachment process depends on the complex interactions between the ion and corresponding neutral atom, and a detailed detachment model would require the knowledge of such interactions. However, with simplifying approximations the detachment process of interest can be identified as collisional in nature, and the resulting dependence on gas density, ion/neutral masses, and a proposed detachment energy threshold can be extracted. The ideal field configuration for electron detachment and amplification will depend on the field magnitude realizable in a given setup. Though not meant to provide precise quantitative predictions, the methods presented here can be used as a way to characterize an ion/neutral system initially and are intended to provide ways to think about the physics of negative ion detectors.
Chapter 6

Conclusions and Future Work

The present study has demonstrated the strong potential of high pressure xenon gas as a detector medium in searches for both dark matter and neutrinoless double beta decay ($0\nu\beta\beta$). 1% FWHM energy resolution was achieved for 662 keV electron recoils using electroluminescent readout at approximately 10 atm pressure, and this resolution extrapolates to 0.5% at the $Q_{\beta\beta}$ value for the neutrinoless double-beta decay of $^{136}$Xe. This is encouraging for the NEXT experiment, which intends to search for $0\nu\beta\beta$ decay with a high pressure xenon electroluminescent TPC. The S1 and S2 signals from nuclear recoils in xenon were also characterized and compared to those of electron recoils, confirming the ability of xenon gas to discriminate between electron and nuclear recoils using the ratio S2/S1. Finally, calculations were presented that seek to aid in the understanding of ion drift and detachment in negative ion TPCs. These results are intended to inform and guide potential future developments of high pressure xenon detectors.

The present results show that to benefit fully from electroluminescent readout, one must also have good tracking capabilities. In this study we relied on a tight fiducial cut to obtain the best possible resolution. In a similar TPC with a larger fiducial volume, one would need to use improved tracking to be able to account for the dependence of light collection on the x-y location at which the S2 light was produced. The tracking capabilities shown within the NEXT collaboration thus far have been encouraging, and studies in the NEXT-DEMO prototype [150] have shown the ability to correct for such effects and obtain good energy resolution in a cylindrical fiducial volume of several cm in radius. In principle, given a tracking resolution that is more precise than any variations in light production and collection, long tracks could be divided into slices and each slice corrected based on its reconstructed location.

While the nuclear recoil yields determined in this study are not conclusive and more detailed measurements are necessary to reveal quantitative information on details such as field and energy dependence of the yields, they can be used as a starting point for future measurements. For example, if one attempts to make a more precise measurement of nuclear recoil yields with a threshold < 10 keVr, one should...
aim for a detection efficiency of at least $\sim 10\%$, as of order 10 primary scintillations photons/keV can be expected based on this study. The primary scintillation yields in the gas and liquid phases appear to be similar, and as teflon is known to be more reflective in the liquid phase, light collection is more difficult in the gas phase. An alternate strategy, for example one in which wavelength shifting bars are used to capture and transport light to PMTs for readout (see [123]) may be necessary to achieve the light collection efficiency necessary to make gaseous xenon competitive in dark matter searches.

Significant further work is also required in understanding whether negative ion drift could be used in a rare physics search with a xenon TPC. Most importantly an electron capture agent that could be drifted in xenon as a negative ion and allow for operation at high pressure ($\sim 10$ bar) must be identified. One would need to drift the ion predictably to its readout location and efficiently detach and amplify the electron. The present study aims to provide ideas and guidance to future development in finding a gas additive and readout configuration capable of doing so.
Bibliography


Plots in this document not taken from external sources were produced using Matplotlib (http://matplotlib.org), GNUPLLOT (http://www.gnuplot.info), or ROOT (http://root.cern.ch/drupal).

This work was supported by the Department of Energy National Nuclear Security Administration Stewardship Science Graduate Fellowship, grant number DE-FC52-08NA28752. This work was also supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy, and used resources of the National Energy Research Scientific Computing Center (NERSC), supported by the Office of Science of the U.S. Department of Energy, both under Contract No. DE-AC02-05CH11231.
Appendix A

NEXT-DBDM Monte Carlo

The primary components of NEXT-DBDM involved in the particle detection process were simulated using a combination of several software packages including GEANT4 [151], BHEP [152], and FMWK [103]. The goal was to simulate the entire process of energy deposition (including the production of ionization and primary scintillation), drift, electroluminescence, and PMT signal production. Waveforms similar to those produced in the actual PMTs of the detector were generated and passed through the data analysis so that the final results could be compared with TPC data.

A.1 Physics and Energy Deposition

The physics component of the simulation was performed using GEANT4 [151] and included the steel pressure vessel, the PMTs, the field cage (teflon reflector panels and wire meshes), and the SiPMs. A snapshot of the geometry is shown in figure A.1. A sensitive region is defined for regions R2 and R3 of figure 3.1 so that the relevant ionization signals will only be produced for events with detectable S2.

A typical event in the GEANT4 simulation proceeds as follows. A neutron of random energy chosen according to the spectrum in figure 4.4 is launched in a random direction in coincidence with a gamma ray of energy 4.4 MeV, both originating from a position between the lead block and the NaI crystal similar to the position of the source in the experimental setup. The G4EmLivermorePhysics electromagnetic physics class was used to describe electromagnetic processes in the simulation, and the high-precision (HP) neutron models (G4NeutronHPElastic, G4NeutronHPInelastic, and G4NeutronHPCapture), which make use cross section data from evaluated nuclear data libraries, were used to describe neutron interactions. In the simulation, gammas and neutrons can interact anywhere in the TPC, though specifically if they interact

---

1 The base of the simulation and analysis code was written mainly by members of the NEXT Collaboration (Justo Martín-Albo, Michel Sorel, et. al. at the Instituto de Física Corpuscular, Valencia, España).
in the xenon primary scintillation will be produced and, if the interaction occurred in the sensitive region, some number of ionization clusters. Each primary scintillation photon is tracked individually by GEANT, and those that interact with the PMT window are registered as hits. Each ionization cluster contains some number of ionized electrons, and the amount of transverse and longitudinal diffusion are computed assuming drift to the anode and using the input diffusion constants and electron drift velocity in the drift region. At the end of each event, S1 hits (number of photons and timing) information is saved, as is position and diffusion information for each ionization cluster. The true locations and kinematic quantities of relevant particles (including S1 photons and ionization clusters, amongst others) are also stored for later reference.

The energies required to produce an S1 photon ($W_{sc}$) and an electron-ion pair ($W_{i}$) were input to the simulation for energy deposition by electrons and gammas as $W_{sc} = 50 \text{ eV}$ and $W_{i} = 25 \text{ eV}$. These values were then increased for nuclear recoils by some quenching factor $Q_{S1}$ or $Q_{S2}$ as $W_{N,sc} = Q_{S1}W_{sc}$ and $W_{N,i} = Q_{S2}W_{i}$.

A.2 Parameterized Simulation of Electroluminescence

After the initial GEANT step of the simulation, the stored quantities are further processed in an electroluminescence (EL) parameterization step. For each electron
in each ionization cluster, a final location and drift time are calculated considering transverse and longitudinal diffusion and the initial production location of the ionization cluster. For each electron, a specified number of photons is generated (equal to the EL gain). These photons are not tracked in GEANT as the S1 photons were, but for each PMT, a photon detection probability is determined from a table of values that was generated in a separate GEANT simulation step. In this EL table generation step, a large number of photons (typically 100000 or 1000000) were generated at each point in an x-y grid of points at the EL plane, and the number of photons detected by each PMT was recorded. For each PMT, the number of photons detected was divided by the total number of photons generated to give a detection probability, and a table was generated mapping an x-y location in the EL plane to that probability. The number of EL photons detected due to each electron traversing the EL gap at a point (x,y) can then be determined, and these photons are distributed in arrival time beginning with at drift time calculated for the original ionization cluster and extending for 3.1 microseconds. The number of photons and their arrival times are binned and recorded to be used in a waveform generation step.

A.3 Waveform Generation

In the waveform generation step, the S1 “hits” (number of photons + arrival times) generated in the GEANT step and the S2 (EL) hits generated in the parameterized step are transformed into data-like waveforms that can then be analyzed as if they were waveforms read from the TPC. These Monte Carlo waveforms are a sum of the detected photon signals on a Gaussian noise profile with a pre-specified sigma. Each photon is represented by an exponential pulse with a time constant of $\tau = 80$ ns and a total charge given by a random Gaussian distributed about the SPE mean $\mu$ (from table 3.2) with $\sigma = \mu$. Once the waveforms have been generated and analyzed according to the data analysis, both data-like and Monte Carlo truth analysis quantities are available on an event-by-event basis.

A.4 Energy Resolution

The full Monte Carlo takes into account effects causing finite energy resolution including the Fano factor (in ionization production) and the Poisson statistics of photon collection in both S1 and S2. However, due to the presence of TPB and PMT afterpulsing, the observed energy resolution in the TPC is significantly worse than that observed in the S1 and S2 spectra resulting from the Monte Carlo. In order to correct

---

2For the purpose of the table generation, the PMT quantum efficiency is assumed to be 100%, and later during the computation of the number of photons detected, the appropriate quantum efficiency is applied.
for these effects, which are likely to be mostly due to the radial dependence of light collection in the TPC, we multiply all S1 and S2 values generated in the Monte Carlo (before any attempts to fit for the nuclear recoil yields) by a Gaussian-distributed scale factor with a mean of 1. For S1, the sigma of this Gaussian distribution is $\sigma_1 = 0.6$, and for S2 we use $\sigma_2 = 0.036$. These values were chosen such that the approximate widths of S1 and S2 peaks in the spectra corresponding to x-ray energies matched. The results of this procedure are described in figures A.2 and A.3.

Figure A.2: Illustration of energy resolution correction procedure for S1. Restricting the values of S2 to those shown in the spectra of figure A.3, one finds the S1 experimental spectrum shown (above). The spectra in Monte Carlo before (left) and after (right) the application of the correction procedure are also shown.
Figure A.3: Illustration of energy resolution correction procedure for S2. Restricting the values of S1 to those shown in the spectra of figure A.2, one finds the S2 experimental spectrum shown (above). The spectra in Monte Carlo before (left) and after (right) the application of the correction procedure are also shown.
Appendix B

Xenon Gas: Energy Resolution and Density

The gaseous xenon medium undergoes significant changes as its density is increased and its properties approach those of the liquid phase. We seek to investigate these issues by summarizing relevant past experimental results to provide insight into how they may be understood. We also perform some calculations and simulations in an attempt to understand what happens at the microscopic level in the the xenon medium as its density is increased and how this may influence the ionization process. While a complete understanding of these issues would require a significant amount of further experimental and theoretical study, we hope to highlight here some of the results that may aid in understanding them.

B.1 Properties of Xenon Gas with Increasing Density

B.1.1 The Electron Mobility

The electron mobility in xenon (see figure B.1), measured along the vapor-liquid coexistence curve, decreases to a minimum at approximately $\rho \sim 0.6 \text{ g/cm}^3$, then rises and sharply peaks at $\sim 2.2 \text{ g/cm}^3$ and then begins to decrease again [153]. This effect has been attributed to the variation of the ground state energy of the free electron across the Ramsauer-Townsend minimum in its low-energy scattering cross section [154]. However, it was noted in path-integral Monte Carlo studies of free electrons in xenon that the interactions between electrons and the xenon medium change with density, creating a “push-pull” effect$^1$ that begins by pushing free electrons away from the bulk medium and pulling them into regions of highest density and eventually

$^1$This picture of electron motion in xenon was emphasized to the author by Prof. David Chandler of U.C. Berkeley.
balances to give rise to a mobility maximum [81]. In this same study it was observed that the kinetic energy of electrons remains relatively constant across densities and even decreases at lower temperatures.

![Figure B.1: (Figure from [153]) Electron mobility vs. density in xenon measured along the vapor-liquid coexistence curve.](image)

This is one example of the interactions between electrons and the xenon medium of changing density, though the decrease, increase, decrease trend does not seem to be correlated with the decreasing W-value and evolution of exciton states.

### B.1.2 The W-Value

As seen in figure B.2, the amount of energy required to produce an electron-hole pair, or the W-value, decreases linearly with increasing xenon density beginning at about \( \rho \sim 0.4 \text{ g/cm}^3 \) [78].

This behavior has been attributed to the rise and evolution of band structure in the xenon medium as it approaches a liquid state [78]. This claim could be tested by performing a calculation similar to that done in [155], [156], and [157] for the W-value in liquid Xe. These calculations make use of Platzman’s energy balance equation [87]

\[
W = \bar{E}_i + \bar{E}_{ex}(N_{ex}/N_i) + \bar{\epsilon}
\]

(B.1)

where \( \bar{E}_i \) is the average ionization energy, \( \bar{E}_{ex} \) is the average excitation energy, \( N_{ex}/N_i \) is the ratio of excitations to ionizations produced, and \( \bar{\epsilon} \) is the average energy of electrons once they can no longer excite the medium. The calculation may be performed
B.1.3 Exciton States

An exciton is a bound electron-hole pair, that is a hydrogen-like bond between an electron and a positively-charged ion within the medium (see for example [160] chapter 7 or [161]). In liquid xenon the exciton energies lie at hydrogen-like energy levels below the band gap energy (see [85] pg. 82 and [162])

\[ E_n = E_g - \frac{G}{n^2} \]  (B.2)

where \( G \approx 0.9 \text{ eV} \) and \( E_g \approx 8.9 \text{ eV} \) (see references within [162]). These energies are depicted in figure B.3.

In each of the above energy states \( E_n \), the electron is located on average a distance (see [162] and [163]) from the central positive core of

\[ a_n = 0.53 \frac{\epsilon n^2}{\mu} \approx 3.25n^2 \text{ Å} \]  (B.3)

if we take the dielectric constant of liquid xenon to be \( \epsilon \approx 1.9 \) (see [162] table 2) and...
Figure B.3: A schematic description of exciton energy levels. These energy levels are described by equation B.2.

\[ \mu_{Xe} \approx 0.31 \] (see [162] and [163]). As the minimum of the Xe Lennard-Jones potential is located at an atomic separation of \( \sigma \approx 0.4 \text{ nm} \) [81] (reference 23 therein), the radius of the \( n = 1 \) exciton spans less than an atomic separation.

To track the evolution of the properties of the xenon medium from gas-like to liquid-like, a standard has been developed in [164] for determining whether a given volume of xenon is in a “liquid-like state,” or one in which the first exciton state can be present. It was determined that the exciton exists when within a volume \( V = 1.5 \times 10^{-21} \text{ cm}^3 \) there exists at least 10 xenon atoms. Molecular dynamics simulations have been performed using the computer code accompanying [165] to investigate this claim, and the results are shown in figure B.4. Similar results can be obtained by analyzing the data of [166]. Note that in [167] it was found that the first exciton state (\( n = 1 \)) in xenon exists for \( \rho > 0.59 \text{ g/cm}^3 \), and the \( n = 2 \) state exists for \( \rho > 2.36 \text{ g/cm}^3 \).

The term “cluster” must not be misinterpreted as a tendency for the medium to spontaneously form dense bodies of atoms, as the gradual “clustering” shown figure B.4 is simply an artifact of the atoms moving closer together as the overall density of the medium is increased\(^2\). In fact, in [81] it was found that electron states in Xe remain delocalized at higher temperatures and densities (see figure B.5) and therefore do not tend to remain near higher-density “clusters.”

### B.2 Energy Resolution

It was found in several studies ([78] and references therein) that the energy resolution in xenon ionization chambers degrades with increasing density after a threshold value of about \( \rho \approx 0.55 \text{ g/cm}^3 \). The present explanations for this effect include fluctuating

\(^2\)Another insight gained through discussion with Professor D. Chandler
Figure B.4: The number of atoms in a cluster of 10 or more obtained from a molecular
dynamics simulation begins to increase around the density $\rho \approx 0.6 \text{ g/cm}^3$. Note that
a cluster was defined in the simulation code from [165] as a set of atoms in which
all were located within some radius $r_0$ of at least one other atom in the cluster (no
reference was made to the characteristic volume defined by Laporte). The radius
chosen was $r_0 = 1.24 \sigma$ as in [166], where $\sigma$ is the location of the Lennard-Jones
minimum for Xe [81]. This is coincides with the appearance of the first exciton state
[167].

electron-ion recombination [168] and the formation of “clusters” in the xenon medium
beginning at the threshold density [79] (see also [95]). The degrading resolution seems
to approach the resolution of the liquid xenon medium. Though Fano factors in liquid
xenon have been calculated to be 0.04-0.06 [157], the actual energy resolution obtained
in liquid xenon spectrometers has at best (see [169]) revealed a Fano factor of $F > 1$
[156]. Such observations of anomalous resolution have also been observed in the liquid
Ar detector medium [170]. Here we discuss several attempts that have been made to
approach this problem and we attempt to connect it with the phenomenon seen in
the high-pressure gas.

B.2.1 Electron-Ion Recombination

When electron-ion pairs are produced in a medium, most of these pairs recombine,
producing photons, unless an electric field is applied to separate them, in which case
the charge can be collected and recorded (see discussions in [79] and [162]). Two
general trends noted in [121] (pg. 58) show that greater electric fields tend to sup-
press recombination up to a point (see plots for example in [78] and [171]) and that
higher density of ionization leads to more recombination. If a fluctuating amount of
Figure B.5: (Figure and caption from [81]) Note that electrons do not seem to be located in regions of higher density, providing strong evidence that localization of free electron states does not occur.

Ionization is produced with high density, the energy partitioned between charge and light will fluctuate (see section B.2.3), an idea that is studied in [168] and references therein. This effect has been modeled in liquid xenon [172] from the point of view of a “δ-ray model” in which the number of low-energy, more densely-ionizing secondary (“delta”) electrons fluctuates, as could the absolute amount of recombination in such delta rays, to cause the collected charge signal to fluctuate. Attempts to test this model have been made by introducing photo-ionizing dopants into the liquid medium (see section B.2.2).

The fluctuation in the number of δ-rays produced by an electron passing through matter can be determined using “knock-on electron” distribution given in [146] pg.
\[
\frac{d^2 N}{dTdx} = \frac{1}{2} K z^2 Z A \beta^2 \frac{F(T)}{T^2}
\]

where \( K = 4\pi N Ar_e^2 m_e c^2 \) (\( r_e \) is the classical electron radius and \( m_e \) is the electron mass), \( z = 1 \) is the electron charge in units of \( e \), \( A \) is the atomic mass of xenon, \( \beta = v/c \) (\( v \) is the velocity of the incident electron), \( F(T) \) is an expression that depends on the spin of the particle, and \( T \) is the kinetic energy of the emitted \( \delta \)-ray. Note that \( dx \) is actually \( \rho dx \) so the quantity on the left hand side of the equation is in \( \text{cm}^2/\text{g} \). When a form of this distribution is used in a Monte Carlo simulation of the ionization distribution of a 662 keV electron, the number of \( \delta \)-rays produced in each event is shown in figure B.6 for 10000 events (note here we made a specific choice for \( F(T) \) that requires further explanation). Note the code to generate this distribution was written by Daniel Hogan (UC Berkeley).

Figure B.6: Simulated \( \delta \)-ray distribution for 662 keV gamma events. This histogram shows the simulated number of delta rays produced for 10000 events. Because the total number of delta rays produced per event is relatively small < 100, differences in the recombination physics of the delta rays and the main track may be more significant.

Because delta-rays are more densely ionizing and thus produce more recombination, if their number fluctuates, so does recombination. This would produce the charge-light fluctuations discussed in B.2.3.

The details of why a more densely ionizing medium supports greater recombination are still unclear. For example, one may ask whether the ions resulting from ionization are so densely packed that their corresponding electrons are quickly finding a neighboring ion (not their parent ion) with which to recombine, or if they are more
likely to remain bound with their parent ion. Note that the existence of exciton states in the high pressure gas and liquid would allow an electron to remain bound to its parent ion at greater distances and also become bound to a neighboring ion at greater distances than in the low pressure gas (see section B.1.3).

Another question addresses the issue of whether electrons are losing energy more efficiently with greater ionization density or the electrons that are ionized with lower energies are more likely to recombine. In a setting of greater ionization density, there may exist some process that causes the electrons to lose large amounts of energy quickly \(^3\). However it may be instructive to look at the energy distribution of ionized electrons, as those at the low end of this distribution are more likely to recombine. These energy distributions were measured for several gases (including xenon) and found to fit an approximate functional form in \([173]\)\(^4\). Note that this distribution refers to the gas phase, though it may still be instructive to examine, at least near the threshold of \(\rho \approx 0.55 \text{ g/cm}^3\).

### B.2.2 Photo-Ionizing Dopants

Attempts to test the model developed in \([172]\) and/or recover energy resolution by converting the recombination light back into a charge signal through the use of photo-ionizing dopants \([170, 97, 98, 174]\) have been conducted, and they conclude that the full intrinsic energy resolution cannot be recovered with this strategy. Further investigation of the use of photo-ionizing agents such as TMA may reveal that it is incapable of converting the ionization lost due to dense recombination back into a charge signal, \(^5\) and this would explain the failure of these agents to account for the fluctuations in recombination of densely ionizing \(\delta\)-rays.

### B.2.3 Observation of Both Light and Charge

We begin by noting that several studies have identified correlated fluctuations between scintillation and ionization signals in liquid noble gas detectors \([175, 176, 96, 177, 178, 121, 179]\). With such fluctuations, the observation of only one of the two signals would inhibit detector energy resolution. This is the key piece of evidence that the poor resolution in the liquid state is due to the partitioning of energy between light and charge through recombination. However, according to recent comments, one by Aprile noted in \([170]\) and another in \([177]\), attempts to observe both light and charge to form a recombination-independent signal have not been successful in obtaining intrinsic resolutions matching the low Fano factors calculated in \([157]\).

---

\(^3\)This idea is due to Dave Nygren (LBNL).

\(^4\)This was brought to our attention through its implementation in the Garfield \([133]\) detector simulation software.

\(^5\)This idea is also due to Dave Nygren (LBNL).
When combining the light and charge signals with the intention of regaining the charge lost to recombination, the light signal must be examined to determine not only how much signal constitutes a single photon but also to determine which photons were produced as a result of recombination and which were produced as a result of excitation. This is because the production of recombination photons should be governed by the processes of charge production while the excitation photons do not reflect a past ionization event. Only the recombination photons should be correlated with the charge signal. We now consider some of the approaches to the summation of charge and light taken in the studies of the correlated fluctuations.

It is stated in [177] that the ionization $Q$ and scintillation signals $S$ should be combined as $Q + aS$, where $a$ is computed by using the scintillation signal $S_0$ with no applied field, the energy deposited $\Delta E$, and the W-value for producing a single photon, in $S_0 = (1/a)\Delta E/W_{sci}$. However, this computation does not separate the photons produced by excitation and those produced by recombination but determines the signal corresponding to a single photon using the sum of both kinds of photons.

Conti et. al. [96] demonstrated correlated light-charge fluctuations using a liquid xenon ionization chamber with a single PMT. They state that their ADC used to read out the PMT signal was calibrated to a single photon, and they report a light signal in “photoelectrons” and compare it to the ionization signal. However, Doke [177] notes that the details of how the scintillation photons detected were used to determine the energy deposited, that is, how $S_0$ was calculated, was not discussed. The procedure for distinguishing scintillation photons due to excitation from those due to recombination was also not discussed. The best possible resolution calculated by considering both signals was determined by plotting the charge vs. light signals and measuring $\sigma$ of the “correlated combination” [96] (see the original source for full details on what this means). It was found that this resolution accounting for both light and charge was an improvement over that determined by measuring ionization only (see figure B.7).

Dahl [121] has presented the most detailed method for summing together light and charge signals. It is claimed in [121] that all losses in energy resolution can be completely accounted for solely by considering known statistical errors in light and charge production, instrumental error, and recombination fluctuations. This energy accounting is done by expressing the fluctuation in each signal in terms of a sum in quadrature of recombination fluctuations and independent single-signal fluctuations [121]:

\[
\sigma_{n_\gamma}^2 = \sigma_R^2 + \sigma_{S1}^2 \\
\sigma_{n_e}^2 = \sigma_R^2 + \sigma_{S2}^2 \\
\sigma_E^2 = W^2(\sigma_{S1}^2 + \sigma_{S2}^2)
\]

where $n_\gamma$ and $n_e$ are the number of photons (to within a multiplicative factor) and charges collected and $W$ is the amount of energy required to produce a single photon.
multiplied by the efficiency for producing a recombination photon. S1 is the actual signal measured for light and S2 is the signal measured for charge. These equations essentially say that the number of emitted photons and collected charges fluctuate even for a constant value of the electric field, and these fluctuations, which are due to recombination, can be separated from the individual “uncorrelated” fluctuations of each signal. The total energy deposited is the same regardless of how much recombination there is and how much it fluctuates. It is claimed in [121] that:

a.) The fluctuations in energy can be broken down into fluctuations in recombination and in the individual signals independent of recombination.

b.) Then the fluctuation in the individual signals can be broken down further into instrumentation error (goes linearly with the signal) and statistical fluctuations (go as the square root of the signal).

It is interesting that Dahl obtains a variance of the recombination fluctuations that is a factor of 100-150 greater than the expected result based on binomial statistics [121]. One significant aspect of this analysis is the ability to distinguish the recombination photons from the excitation photons, that is, to calculate the recombination

Figure B.7: (Figure from [96]) Energy resolution in liquid xenon with and without light and charge signal summation. Note that $\sigma_e$ is the resolution obtained by observing ionization only and $\sigma_{\text{min}}$ is that obtained by considering both ionization and light.
fraction $r$, with some approximation and using some information from theory, such that the number of recombination photons is equal to $rN_i$, where $N_i$ is the number of ionizations.

Overall, we argue that this light/charge fluctuation is a property of the liquid xenon medium, and that it is the same problem that is gradually becoming present in the high-pressure gas as density is increased. If this is the case, understanding the behavior in the energy resolution in the high pressure gas would lead to simultaneous understanding in the liquid phase. Bolotnikov and Ramsey have studied light and charge signals in the high pressure gas at densities as high as 0.74 g/cm$^3$ in [79] by observing the two components of the luminescence decay, with a “fast” decay time $\tau_1$ and a “slow” decay time $\tau_2$. Note that this study did not focus on the summation of light and charge signals to obtain better energy resolution but mostly on the dynamics of the light output. In this study [79], $\tau_1$ was found to decrease from 14 ns to 5 ns as the density of the medium was varied from 0.06 g/cm$^3$ to 0.74 g/cm$^3$. The argument is that “volume” recombination, that which results from electrons that have been freed from their parent ion but later return to recombine with it or others in the area, contributes at the lower densities to this decay time but not at higher densities. $\tau_2$ was found to decrease from 140 ns to 25 ns over the same density range, and here the argument is that the decay time becomes more like that of an exciton ($\tau = 21$ ns) and less like that of a Xe$_2^+$ dimer ($\tau = 103$ ns), such that the probability of decaying via a self-trapped exciton is very near to 1 for $\rho > 0.6$ g/cm$^3$ (see figure B.8). This supports the appearance of the $n = 1$ exciton state [167] at approximately this density.

### B.2.4 Comments

Here we make several notes that attempt to synthesize the details of this section into a reasonable explanation for the poor energy resolution in liquid and thus the degrading energy resolution in the gas.

First note that in the liquid medium all observed UV scintillation occurs at the same wavelength and was claimed to be a result of either “self-trapped” exciton decay (26%) or luminescence due to recombination that can be suppressed by a high electric field (74%) [175], see also [180]. It was also stated that the decay of a “localized” excited dimer causes the UV emission, and this localization occurs either before the electron becomes involved (the $R_2^+$ hole is “self-trapped”) or an exciton undergoes “dynamical self-trapping” and then decays [181]. Perhaps the observations of [175] confirm this claim in that the suppressing of the recombination emission occurs because the electron that would otherwise recombine with the “self-trapped” $R_2^+$ hole is driven away from it by the field. However, the field is not strong enough to remove the electron from an exciton, and so some luminescence remains. Note that since Bolotnikov found that the “volume” recombination does not contribute to the luminescence at higher densities [78], we may need to distinguish between two
different “volume”-like recombination processes - one in the high density regions of ionization in which electrons are attracted to self-trapped $R_2^+$ holes that were not formed by their parent ions and one in the low-density gas in which an electron may wander for a longer time before recombining with a non-parent ion.

Holes are localized into $R_2^+$ molecular ions “within a picosecond” (see references 14-16 and pg 3490 of [180]). As this process occurs quickly, the self-trapped exciton light seen in [175] was most likely produced by low-energy ionized electrons that recombined immediately or were never fully ionized. This is consistent with Bolotnikov’s [79] findings of increased “geminate” recombination at densities $\rho > 0.6 \text{ g/cm}^3$.

According to a comment by G. R. Freeman in the “Discussion” following [162], an estimated $< 10\%$ of excitons with $n = 6$ could be ionized by an electric field of 100 kV/cm. This comment indicates that once an exciton has formed an electron that has been nearly ionized can no longer be drawn away from its parent ion in nearly all cases. The typical UV emission lies around 170 nm in the gas ([85] pg. 76), while exciton energies are $E_G - G \leq E_{\text{exciton}} \leq E_G$ with $E_G = 9.28\text{ eV}$ and $G = 0.86\text{ eV}$ for solid Xe ([85] pgs. 82-83), though $E_G = 8.9\text{ eV}$ in the liquid (see ref. 15 of [162]). This would mean that $\sim 8.4\text{ eV}$ would be required to produce an exciton while only $\hbar c/\lambda \approx 1240\text{ eV} \cdot \text{nm} / 170\text{ nm} \approx 7.3\text{ eV}$ would be seen in the form of a photon.
and approximately 0.5-1 eV is lost, possibly due to non-radiative processes such as resonant energy transfer (see [181] and [182]), per exciton.

It was shown in [121] that higher densities of ionization give rise to greater recombination, which according to the above arguments must be fluctuation in the production of exciton states, despite whether the electron was ever completely freed or not. The fluctuation in the low number of $\delta$-rays shown in section B.2.1 would lead to fluctuation in ionization density. This evidence along with the increased appearance of exciton light found in [79] and [175] indicates that the exciton, a state that lies between excitation and ionization, plays a key role in the observed fluctuations between light and charge. By studying the energy resolution at the onset of the effect when the first $n = 1$ exciton appears in the gas phase, one should be able to confirm this reasoning.
Appendix C

Space-Charge in a Parallel Plate Region

C.1 Theory

We consider the case of a source producing ionization uniformly between two parallel plates of area $A$ and separation $d$ across which a voltage difference $V$ is present. The rate of ionization production is such that one may be concerned that enough charge remains in the region of interest to distort the field seen by the electrons, particularly charge from the ions which are relatively slow-moving compared to the electrons. Assuming the ions move at drift velocity $v$ in the field $E$ and are produced at a rate of $N$ ions/second, we attempt to find the ion charge per unit length $\eta(z)$ in the steady state created by the two opposing processes of ion production and ion removal by the drift field. We define the ion charge production per unit length per unit time as $\chi = eN/d$. A differential equation describing these processes can be formulated by considering the changes in the distribution in an infinitesimal spatial interval $dz$ during an infinitesimal time interval $dt$. The amount of ion charge in this interval $\eta(z,t)dz$ becomes $\eta(z,t+dt)dz$ due to ions drifting out of the volume $Adz$, ions drifting into the volume, and ions produced in the volume.

\[
\eta(z,t+dt)dz = \eta(z,t)dz + \chi dz dt + \eta(z - dz,t)v(z - dz,t)dt - \eta(z,t)v(z,t)dt. \quad (C.1)
\]

This equation reads: the ion charge in $dz$ at a time $t + dt$ is equal to the charge previously in $dz$, plus the charge in $dz$ produced by the source in $dt$, plus the charge entering $dz$ in $dt$, minus the charge leaving $dz$ in $dt$. Rewriting this using $f(x + dx) =
\[ f(x) + (\partial f/\partial x)dx, \text{ we have, denoting } \eta = \eta(z,t) \text{ and } v = v(z,t) \]

\[
\eta dz + \frac{\partial \eta}{\partial t}dzdt = \eta dz + \chi dzdt + [\eta - \frac{\partial \eta}{\partial z}] [v - \frac{\partial v}{\partial z}] dt - \eta vd t
\]

thus

\[
\frac{\partial \eta}{\partial t} dz dt = \chi dz dt - \eta \frac{\partial v}{\partial z} dz dt - v \frac{\partial \eta}{\partial z} dz dt + \frac{\partial \eta}{\partial z} \frac{\partial v}{\partial z} d z^2 dt \tag{C.2}
\]

and

\[
\frac{\partial \eta}{\partial t} + v \frac{\partial \eta}{\partial z} + \eta \frac{\partial v}{\partial z} = \chi.
\]

where in the last step we have taken the limit as the differentials \(dz\) and \(dt\) approach 0 so the last term vanishes. In the steady state all time derivatives are set equal to 0, and we can rewrite the left hand side of equation C.2 as

\[
\frac{d}{dz} (\eta v) = \chi. \tag{C.3}
\]

We have now reduced the problem to that of a single variable. Integrating we find that

\[
\eta(z)v(z) = \chi z + C(\chi), \tag{C.4}
\]

where \(C(\chi)\) is some constant that depends on \(\chi\). We note that \(\eta(z)v(z)\) is the charge passing through a point \(z\) per unit time, and so we choose \(C(\chi) = 0\) so that all ions that pass through a point \(z\) originated somewhere below it.\(^1\) Now we use Gauss’ law in differential form, and assuming a field in the \(z\) direction only,\(^2\)

\[
\frac{\partial E(z)}{\partial z} = \frac{\eta(z)}{A \epsilon_0}. \tag{C.5}
\]

And using \(v(z) = KE(z)\), where \(K\) is the ion mobility,\(^3\)

\[
\chi z = K \eta(z) E(z) = KA \epsilon_0 E \frac{\partial E(z)}{\partial z} = K \epsilon_0 A \frac{1}{2} \frac{d}{dz} (E^2). \tag{C.6}
\]

Integrating this equation and defining some constant \(k^2\), we have

\[
[E(z)]^2 = k^2 + \frac{\chi}{K \epsilon_0 A} z^2. \tag{C.7}
\]

\(^1\)If \(C(\chi) \neq 0\), then it would imply some constant ion flow through a point \(z_1\) without the same ions flowing through the point \(z_2 > z_1\).

\(^2\)This assumes that we are considering the region between the plates far from the edges of both the plates and the region described by \(A\).

\(^3\)We assume that over the range of \(E\) values studied the mobility remains a constant.
We can now use as boundary conditions the voltages on the plates,

\[-V = - \int_0^d E(z) dz = - \int_0^d \sqrt{k^2 + \frac{\chi}{K \epsilon_0 A} z^2} \, dz,\]

and defining \( u_0 \equiv (\chi d / K \epsilon_0 A) \), finding the unknown \( k \) amounts to solving the equation

\[ \frac{V}{u_0 d} = \frac{1}{2x} \left( \sqrt{1 + x^2} + \frac{1}{x} \ln(\sqrt{1 + x^2} + x) \right), \] (C.9)

for \( x = u_0 / k \). Once \( k \) has been found the electric field between the plates can be expressed as

\[ E(z) = \sqrt{k^2 + u_0^2 (z/d)^2}. \] (C.10)

And using equation C.5 we can compute the total space charge

\[ Q_{\text{space}} = \int_0^d \eta(z) dz = A \epsilon_0 k \left[ \sqrt{1 + (u_0/k)^2} - 1 \right]. \] (C.11)

Note that if \( u_0 \ll V/d \) we can solve equation C.9 in for small \( x \), obtaining

\[ k \approx \sqrt{(V/d)^2 - \frac{1}{3} u_0^2}, \] (C.12)

and therefore

\[ E(z) \approx \sqrt{(V/d)^2 + u_0^2 [(z/d)^2 - 1/3]}. \] (C.13)

From this we conclude that the effects of space charge are small if \( u_0 \ll V/d \), or

\[ \frac{\chi d^2}{K V \epsilon_0 A} \ll 1. \] (C.14)

### C.2 Example - Space charge in a parallel plate geometry

Given a parallel plate geometry with separation \( d = 0.5 \) cm illuminated by a source of \(^{241}\)Am x-rays, we calculate the effects of space charge. We assume that the number of x-rays deposited per second in the region between the plates is about 150000 s\(^{-1}\) for a source of activity 10 mCi and scales linearly with the activity. We also assume
that the area over which ions are deposited is uniform across the gap and equal to \( A = 1 \text{ cm}^2 \).

From chapter 5, the mobility of \( \text{Xe}^+ \) ions in xenon can be calculated in an approximation that assumes thermal ion energies and for which the interaction between the xenon ions and neutral xenon atoms can be described by a two-body potential produced by the polarization of the neutral by the ion, and the subsequent interaction between the ion and the polarized atom \[134\] (see equation 5.4),

\[
K = \frac{1}{2\pi \rho} \left( \frac{1}{\mu \alpha_d} \right)^{1/2},
\]

where \( \mu \) is the reduced mass of the system (in this case, \( 1/2 \) of the xenon atomic mass), \( \rho \) is the xenon gas density, and \( \alpha_d \) is the dipole polarizability of xenon. We use \( \alpha_d = 4.044 \times 10^{-24} \text{ cm}^3 \) (in cgs units) \[134\] and calculate \( N \) assuming xenon is an ideal gas at 1 atm. The calculated mobility is then

\[
K_{\text{Xe}} = 1.01 \text{ cm}^2/(\text{V} \cdot \text{s}).
\]

Note that for high fields \( (E/N = 40 \text{ Td}) \), the measured mobility of \( \text{Xe}^+ \) in \( \text{Xe} \), depending on the spin state of the ion, is between 0.558 and 0.591 \( \text{cm}^2/(\text{V} \cdot \text{s}) \) \[137\].

Using this mobility, one can calculate the electric field \( E(z) \) and total space charge \( Q_{\text{space}} \). The maximum electric field \( E_{\text{max}} = E(d) \) and minimum electric field \( E_{\text{min}} = E(0) \) are shown in figure C.1 relative to the pure parallel plate field with no space charge \( E_0 = V/d \), and the total space charge is shown in figure C.2 for EL gap voltage of \( V = 400 \text{ V} \) and different values of the source activity assuming \( N = 150000 \) ions/s corresponds to an activity of 10 mCi. Note that the effects of space charge will become more significant relative to the 0-space-charge field as the gap voltage \( V \) decreases. In the current configuration, space charge effects seem to be negligible until extremely high source activities are used.
Figure C.1: The relative minimum and maximum electric field for electrode voltage of $V = 400$ V vs. source activity assuming an activity of 10 mCi produces $N = 150000$ ions/s in the active volume. $E_0 \equiv V/d$ is the field in the absence of a source (space charge).

Figure C.2: The total space charge for electrode voltage of $V = 400$ V vs. source strength assuming an activity of 10 mCi produces $N = 150000$ ions/s in the active volume. The horizontal line denotes the charge that would lie on one parallel plate in an area of 1 cm$^2$ assuming no source (space charge) is present.