Reconstructing Nuclear Recoil Tracks in the Dark Matter Time Projection Chamber

by

William Bradley Fedus

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Bachelor of Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2010

© Massachusetts Institute of Technology 2010. All rights reserved.
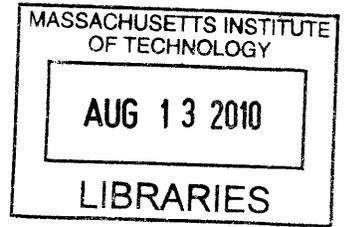
Author . . . .
Department of Physics
May 7, 2010

Certified by . . . .
Professor Gabriella Sciolla
Associate Professor of Physics
Thesis Supervisor

Certified by . . . .
Dr. James Battat
Pappalardo Fellow, Department of Physics
Thesis Supervisor

Accepted by . . . .
Professor David E. Pritchard
Senior Thesis Coordinator, Department of Physics

Reconstructing Nuclear Recoil Tracks in the Dark Matter Time Projection Chamber

by

William Bradley Fedus

Abstract

Astrophysical evidence indicates that 23% of our universe's energy density is in the form of nonluminous, nonbaryonic matter referred to as dark matter. One theoretically appealing dark matter candidate is the Weakly Interacting Massive Particle (WIMP). Because of astrophysical dynamics, the detectable signal from the expected WIMP dark matter halo should exhibit a unique daily directional modulation for which experiments can search .

The Dark Matter Time Projection Chamber (DMTPC) group aims to provide an unequivocal detection of WIMP particles through the anisotropy in the angular recoil spectrum. DMTPC uses a low-pressure time projection chamber filled with CF_4 gas to search for WIMPs via elastic collisions. Crucial to this experiment is the fidelity of nuclear recoil track reconstruction. By extracting parameters such as the angle and vector direction of nuclear recoils, DMTPC has sensitivity to the anisotropic WIMP signal.

This thesis develops a new track reconstruction algorithm motivated by the physics of nuclear energy loss in a diffuse gas medium. The algorithm is applied to simulated nuclear recoils and is compared to the existing track reconstruction algorithm. The new fitting algorithm outperforms the old algorithm in determining vector direction of nuclear recoils for recoil energies between 20 and 300 keV. The algorithm shows little sensitivity to CCD read noise. The length reconstruction of the new algorithm, however, fails to outperform the old algorithm below 100 keV.

Thesis Supervisor: Professor Gabriella Sciolla
Title: Associate Professor of Physics

Thesis Supervisor: Dr. James Battat
Title: Pappalardo Fellow, Department of Physics

Acknowledgments

All work presented here was completed on behalf of the DMTPC group whose members include C. Deaconu, D. Dujmic, J. Battat, T. Caldwell, P. Fisher, S. Henderson, R. Lanza, A. Lee, J. Lopez, A. Kaboth, G. Kohse, J. Monroe, R. Vanderspek, T. Sahin, G. Sciolla, I. Wolf, R. Yamamoto, H. Yegorian, S. Ahlen, A. Inglis, K. Otis, H. Tomita and H. Wellenstein. I would like to thank everyone for his/her contribution to my work. A few members have had a particularly significant influence on this work and I recognize them here.

First, I could not have wished for a more capable and more dedicated mentor and thesis adviser than James. Only through his countless hours of work and late nights of last-minute revisions was this thesis made possible. Furthermore, his enormous contributions and guidance to my work continued even while he welcomed his new baby girl to his family - a tribute to his work ethic and commitment to his students.

Of course, I need to thoroughly thank all of my labmates - Shawn, Jeremy, Cosmin, and Asher. Never perturbed by my continuous stream of coding and algorithmic issues, each one offered his full attention and support. Each of them has been an incredible resource in physics and other unrelated fields and I am lucky to have been able to work with them over this last year. Furthermore, I thank Michael Betancourt, my pseudo-labmate, for his extremely insightful inputs into the algorithm development and for his constant prodding in between.

And finally, I cannot fully express my gratitude towards my adviser, Gabriella, for inviting me to join DMTPC several years ago. Working with the group has been intellectually challenging and it has helped me develop a robust approach to solving complex problems. As both my thesis and academic adviser, Gabriella has been there as an expert guide for physics and general life questions and I've always admired her commitment.

4

Contents

Chapter 1

Introduction

To date, all observational evidence of dark matter is gravitational. The first indication that astronomers were missing a significant portion of mass in the universe emerged in 1933 when Fritz Zwicky developed a new method to determine the masses of galaxy clusters. Previously, a cluster's mass was inferred from the cluster galaxy's luminosities and internal rotational velocities. However, when Zwicky used the rotational velocities of the galaxies near the edges of the cluster with the virial theorem, he found the implied mass of the cluster to be many times larger than what other methods would suggest [19].

Further investigations into the nature of this hidden mass have led to the conclusion that dark matter is non-baryonic. This has been led by the analysis of the Cosmic Microwave Background (CMB) Radiation within our universe. The CMB is an electromagnetic radiation that permeates the universe and has small anisotropies in intensity. The power spectrum of the CMB anisotropies has several peaks which describe the density of various types of matter within the universe. Importantly, it was determined that the density of dark matter does not match that of the density of the baryonic matter [16], implying that dark matter is not only non-luminous but also non-baryonic.

Finally, and most important for direct detection experiments, dark matter is believed to exist within our Milky Way galaxy. This has been deduced from the rotation curves within our Galaxy [10].

Dark matter is expected to be non-relativistic and stable since it appears to have sustained minute density fluctuations in the universe and its gravitional effects are still observed. Because dark matter appears to interact primarily through gravitational interactions, valid theoretical particle candidates must have extremely small interaction cross sections with other dark matter particles and with baryonic matter. However, the properties of dark matter are poorly constrained and a broad landscape of candidate particles has been proposed. One of these candidate particles, the WIMP, can naturally provide the observed cosmological dark matter density observed today, and also arises separately as a stable particle in supersymmetry theories. Like many other direct detection experiments, DMTPC is sensitive to WIMP dark matter particles.

1.1 Nuclear Recoil Physics

WIMPs can interact elastically with nuclei via the weak interaction. In this section, we derive the properties of WIMP-induced nuclear recoils. The spiral arms of the Milky Way galaxy are embedded in a much larger halo of dark matter particles. In the simplest models of the halo, the dark matter particles have a Maxwellian velocity distribution with a mean velocity ≈ 200 km s^{-1} [1]. The orbital velocity of the Earth through the dark matter halo produces a headwind of WIMPs with a typical velocity, v_0 of 230 km s^{-1} [15].

The CF$_4$ molecules within the DMTPC detector have a Maxwellian velocity distribution. We consider the limit where the target gas particle is at rest in the lab frame. An assumption we now justify. Typical thermal velocities v of gas particles of mass m and temperature T are:

$$\langle v \rangle \approx \sqrt{\frac{3kT}{m}}. \tag{1.1}$$

In the DMTPC detector, $T \approx 295$ K, and the mass of the CF$_4$ molecule is 1.46×10^{-25} kg and the mean gas molecule velocity is $\langle v \rangle \sim 236$ m s^{-1}. The ratio of this approximate CF$_4$ velocity, v, to the expected WIMP velocity, v_D, is,

$$\frac{\langle v \rangle}{v_D} \approx \frac{236 \text{ m s}^{-1}}{230000 \text{ m s}^{-1}} \approx .001 \tag{1.2}$$

and so therefore to simplify the interaction, the CF_4 gas may be treated as stationary ($v = 0$). Within the CF_4 nucleus, the interaction usually occurs with the F nuclei since it has a larger cross section - therefore, we model the kinematics of an interaction of a WIMP traveling along the $+\hat{x}$ direction and a stationary F nuclei. Furthermore, this collision within the time projection chamber may be accurately treated to be perfectly elastic.

m_D	Mass of WIMP
m_T	Mass of F

Where the energies are defined as follows,

E_0	Total initial kinetic energy in Lab Frame
E_0'	Total initial kinetic energy in CM Frame
E_D	Final kinetic energy of WIMP in Lab Frame

And the relevant angles of the problem are given by,

ψ_D	Angle through which WIMP is scattered in Lab Frame
ψ_T	Angle through which F is scattered in Lab Frame

This is drawn with labels in Figure 1.1.

To estimate the energy of WIMP-induced nuclear recoils, we first begin with the conservation of energy

$$E_0 = E_D + E_T \tag{1.3}$$

and the conservation of momentum equations in x-direction ($\hat{x}$) and the y-direction ($\hat{y}$)

$$\hat{x} : \sqrt{E_0} = \sqrt{E_D} \cos\psi_D + \sqrt{ME_T} \cos\psi_T \tag{1.4}$$

9
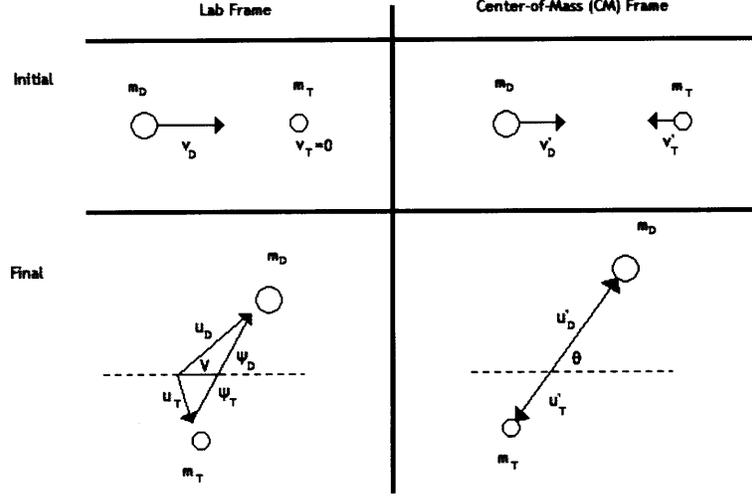
Figure 1-1: Diagram of the elastic collision between a WIMP and a Fluorine nucleus.

$$\hat{y}: E_D \sin^2 \psi_D = M E_T \sin^2 \psi_T \tag{1.5}$$

where

$$M \equiv \frac{m_T}{m_D}. \tag{1.6}$$

and the momentum of a particle i is written as

$$p_i = \sqrt{2 m_i E_i}. \tag{1.7}$$

To get E_T as a function of E_0 and one angle, first rearrange Equation 1.4 to isolate the $\cos \psi_D$ term and then square the result

$$E_D \cos^2 \psi_D = E_0 + M E_T \cos^2 \psi_T - 2\sqrt{2 E_0 E_T M} \cos \psi_T \tag{1.8}$$

then add Equation 1.8 to Equation 1.5 to get

$$E_D = E_0 + M E_T - 2\sqrt{E_0 E_T M} \cos \psi_T. \tag{1.9}$$

10

Now, use Equation 1.3 to eliminate E_D from Equation 1.9

$$
\begin{aligned}
E_0 - E_T &= E_0 + ME_T - 2\sqrt{E_0 E_T M}\cos\psi_T \\
2\sqrt{E_0 E_T M}\cos\psi_T &= E_T(1 + M) \\
4E_0 E_T M \cos^2\psi_T &= E_T{}^2 (1 + M)^2
\end{aligned}
$$

$$
E_T = E_0 \frac{4M}{(1+M)^2}\cos^2\psi_T \tag{1.10}
$$

Which gives the recoil energy of a nucleus struck by a WIMP particle and scattered at an angle θ. Using this equation, we can gain some sense of the nuclear recoil energies we will expect.

Using Equation 1.10, we may gain a sense of what range of energies we expect our nuclear recoils to be within. First, we choose $\psi_T=0$ to maximize E_T. Next, the other parameters are expected to have the following approximate numerical values

$$
\begin{aligned}
m_D &= 100 \text{ GeV/c}^2 \\
m_T &= 19 \text{ GeV/c}^2 \\
v_D &= 300 \text{ km s}^{-1} \rightarrow \frac{v_D}{c} = 10^{-3}
\end{aligned}
$$

and so $M = 0.19$ and $\frac{4M}{(1+M)^2} = 0.54$

$$
\begin{aligned}
E_0 &= \frac{1}{2}m_D v_D{}^2 \\
E_0 &= \frac{1}{2}\left(m_D c^2\right)\left(\frac{v_D}{c}\right)^2
\end{aligned}
$$

$$E_0 \approx \frac{1}{2} \cdot (100 \text{ GeV}) \cdot \left(10^{-3}\right)^2 \tag{1.11}$$

$$E_0 = 50 \times 10^{-6} \text{ GeV} = 50 \text{ keV}$$

and plugging in these values into Equation 1.10,

$$E_T \approx (50 \text{ keV}) (0.54) = 27 \text{ keV} \tag{1.12}$$

Therefore Equation 1.10 shows us that a dark matter detector must be sensitive to recoils with energies of tens of keV in order to detect WIMP particles.

With the basic kinematics of the interaction in place, we now investigate the expected nuclear recoil spectrum. First it is assumed that the WIMP halo is isotropic and obeys a Maxwellian dark matter velocity distribution,

$$f(v, v_E) \propto \exp{-\frac{(v + v_E)^2}{v_0{}^2}} \tag{1.13}$$

where v is the velocity onto the target, v_E is the Earth velocity relative to the dark matter distribution and v_0 is the average WIMP velocity. By assuming this and that the scattering is isotropic (uniform in $\cos\theta$ or equivalently $0 \leq E_T \leq Er$) we arrive at the following unmodified nuclear recoil spectrum [15]

$$\frac{dR(0, \infty)}{dE_T} = \frac{R_0}{E_0 r} \exp{\frac{-E_T}{E_0 r}}, \tag{1.14}$$

where R is the recoil rate, E_T is the energy of the target recoil, E_0 is the most probable incident kinetic energy of the WIMP particle, r is the kinematic factor $\frac{4m_D m_T}{(m_D + m_T)^2}$ for a target nucleus of mass m_T, a dark matter mass of m_D, R is the event rate per unit mass and R_0 is the total event rate. Due to the exponentially falling nature of the nuclear recoil rate, there is a heavy premium for being able to detect and reconstruct low energy events.

1.2 Directional Detection

As mentioned above, a simple model of our Milky Way galaxy suggests that the familiar spiral arms are embedded within a much larger dark matter halo. The spiral arms rotate through this larger halo, resulting in a flux of WIMPs through the Earth. This flux of WIMPs, often referred to as a "WIMP wind" has two primary modulations that may be used to detect the presence of dark matter.

First, there exists an annual modulation in the signal because of the Earth's orbit around the Sun. This is a result of a partial alignment and misalignment of the Earth's and the Sun's orbital angular velocity vectors. When the Earth's orbital velocity vector aligns with that of the Sun in the northern hemisphere's summer, the WIMP speed distribution is boosted higher and thus the energies of interacting nuclear recoils are higher. Conversely, we expect the opposite effect to occur in the northern hemisphere's winter when the two orbital velocity vectors are misaligned. Equation 1.15 is an approximation of this annual modulation for a smooth dark matter halo [9]

$$\frac{dR}{dE}(E,t) \approx S_0(E) + S_m(E)\cos\omega(t - t_c) \qquad (1.15)$$

where t_c is the time at which $v_{obs}(t)$ (the Earth's relative speed to the dark matter Halo) is at its maximum. $S_0(E)$ is the average differential recoil rate over a year and $S_m(E)$ is the modulation amplitude. However, this modulation is only of order a few percent and thus requires a large exposure to detect. We do not consider the annual modulation in this work.

Secondly, there exists a daily directional modulation of the signal due to the Earth's rotation about its axis. The Earth's axis of rotation is oriented at approximately 48 degrees from the anticipated direction of the WIMP wind, so the expected signal will be modulated in angle. After 12 hours the direction of the incident WIMP wind will change by approximately 90 degrees. Figure 1.2 shows the expected time-averaged recoil flux distribution in galactic coordinates. Coincidentally, this daily angular variation of the WIMP wind corresponds closely to the direction pointing

13
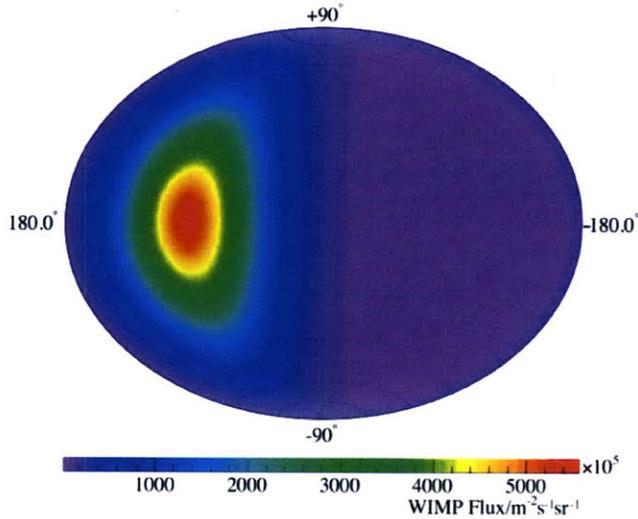
Figure 1-2: Hammer-Aitoff projection of the anticipated WIMP wind in galactic coordinates. The peak of this distribution coincides with the position of the Cygnus constellation. A WIMP mass of 100 GeV c^{-2} is assumed. Figure taken from [17].

towards the constellation Cygnus.

The measurement of this modulating signal, which cannot be mimicked by any known background source, would thus provide unequivocal evidence for the detection of dark matter.

1.3 WIMP Detection

Currently, many experiments are attempting to directly detect WIMPs. Most direct detection experiments attempt to identify these nuclear recoils in gaseous, liquid, or solid detectors. Measurement of the recoil is typically done by charge readouts and/or scintillation light created from the interaction. As mentioned above, our experiment falls under the gaseous category, along with DRIFT and NEWAGE. Liquid detectors include, XENON, and COUPP, and solid state detectors include CDMS, DAMA, CRESST, KIMS.
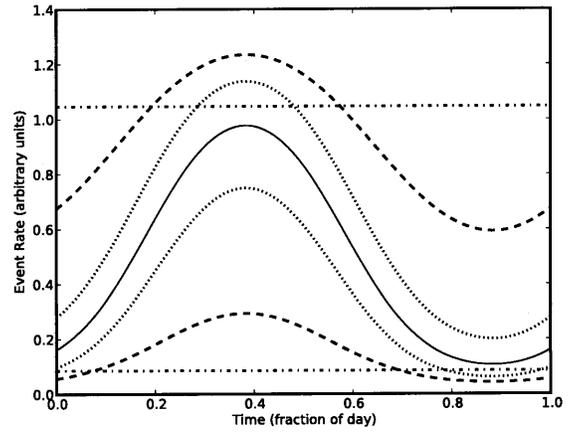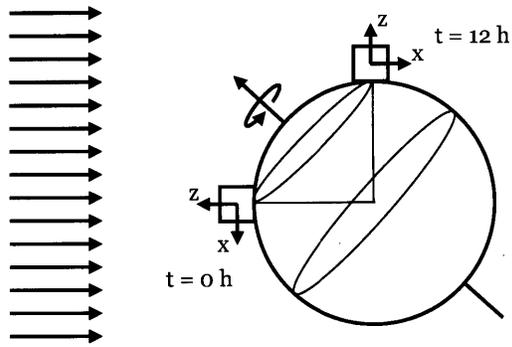
Figure 1-3: (Left) A schematic of the WIMP wind incident upon the Earth. The daily rotation introduces an angular modulation in the direction of arrival of the WIMP wind. Figure taken from [18]. (Right) The expected magnitude of this variation for seven directions in the Earth frame. The directions are 0° (solid line), ±18° (dotted), ±54° (dashed), and ±90° (dash-dot) respectively. Note that positions at ±90° point along the Earth's axis of rotation and thus have no modulation in signal intensity. Here a WIMP mass of 100 GeV c^{-2} is assumed and CS_2 is the target gas. Figure taken from [18].

Although gas detectors have masses that are typically much smaller than liquid or solid detectors (due to the low density of gas), the primary advantage of such devices is that nuclear recoils will travel a longer distance through the detector, measuring not just energy of interaction, but also the range, angle and vector direction of the nuclear recoil. This directional sensitivity allows gaseous detectors to factor out and reject isotropic backgrounds.

These following three paragraphs regarding the relevant cross section components are referenced from [2]. The weak interaction of the target material with the WIMP particle will in general have a cross section that will have both spin-independent and spin-dependent couplings

$$\sigma = \sigma_{SI} + \sigma_{SD}. \tag{1.16}$$

With a nucleus with Z protons, A-Z neutrons, and nuclear form factor F(E), the SI contribution to the WIMP-nucleus cross section is

$$\sigma_{SI} = \frac{4\mu^2}{\pi}[Zf_p + (A-Z)f_n]^2|F(E)|^2, \tag{1.17}$$

where f_p and f_n are WIMP couplings to the proton and neutron respectively. Next, the other component to equation (2) is the SD coupling which is given below for a nucleus with spin J and momentum exchange dependent structure functions S_{pp}(q), S_{pn}(q), S_{nn}(q)

$$\sigma_{SD} = \frac{32\mu^2 G_F^2}{2J+1}[a_p^2 S_{pp}(q) + a_p a_n S_{pn}(q) + a_n^2 S_{nn}(q)], \tag{1.18}$$

where G_F is the Fermi coupling constant, a_p and a_n are the axial four-momentum WIMP-nucleon couplings in units of $2\sqrt{2}G_F$. The DMTPC utilizes the S_{pp} coupling of $^{19}F(J = 1/2)$ in the form of CF_4 gas to make the SD measurements.

A. Green and B. Morgan [12] consider how each aspect of a directional detector's track reconstruction performance impacts the corresponding number of events which must be observed in order to reject isotropy with some confidence. As the base configuration for a directional detector, Green and Morgan assumed a theoretically

16

perfect, albeit currently experimentally intractable, CS_2 gaseous directional detector at 40 Torr with a 3-D vector (i.e. meaning recoil sense or "head-tail" of the recoil is measurable) read-out, an energy threshold of $E_T = 20$ keV (i.e. energy above which track direction may be reconstructed), and no background (Signal/Noise $\equiv$ S/N $= \infty$). Figure 1-4 lists specifications which are modified and the number of events needed to reject isotropy.

Figure 1-4 shows the results of this simulation and motivates this research. Specifically, the ability to reconstruct the head-tail of the recoil, rather than just determining the axial direction of the track, allows a detector to reject isotropy with a factor ~ 15 fewer reconstructed events (comparing "2-d axial read-out in optimal plane, reduced angles" vs. "2-d vector read-out in optimal plane, reduced angles"). Therefore, reconstruction of head-tail is of utmost importance in order to maximize the sensitivity of our experiment and to allow us to reject isotropy with shorter exposures per unit mass. Furthermore, Figure 1-4 details the importance of adding in a 3-D measurement into the experiment. By determining the 3-D orientation of the nuclear recoil, we may reject isotropy with an exposure that is $\sim$2x smaller.

The number of events necessary to reject isotropy may be re-framed in terms of expected exposure, E, needed to accumulate the statistically required number of events. The exposure required to accumulate the necessary number of events for a 95% confidence detection in 95% of experiments as a function of the WIMP-nucleon elastic scattering cross section, σ_0, was calculated using

$$E = \frac{N}{\sigma_0 R_T} \tag{1.19}$$

R_T is the total event rate above the threshold energy/unit cross section and N is the number of events required to be statistically significant. Below is a comparison over 2-D read-outs (vector and axial and also raw angles and reduced angles)

Figure 1-5 demonstrates the differences in exposure required to reject isotropy across detectors of various readout capabilities. It is shown that the requisite exposure is decreased by a factor of $\sim$15 by reconstructing the head-tail of the nuclear recoil.

difference from baseline configuration	N_{90}	N_{95}
none	7	11
$E_T = 0$ keV	13	21
no recoil reconstruction uncertainty	5	9
$E_T = 50$ keV	5	7
$E_T = 100$ keV	3	5
$S/N = 10$	8	14
$S/N = 1$	17	27
$S/N = 0.1$	99	170
3-d axial read-out	81	130
2-d vector read-out in optimal plane, raw angles	18	26
2-d axial read-out in optimal plane, raw angles	1100	1600
2-d vector read-out in optimal plane, reduced angles	12	18
2-d axial read-out in optimal plane, reduced angles	190	270

Figure 1-4: Dependence of the number of events above an energy threshold, E_T, required to reject isotropy of the signal. The baseline configuration has 3-D vector read-out, energy threshold $E_T = 20$ keV, and no background, $S/N = \infty$. For non-zero background configurations, the numbers displayed are the number of signal events required. Two confidence intervals were used, a 90% confidence interval and a 95% confidence interval. N_{90} (N_{95}) is the number of events needed to reject isotropy of signal to the 90% (95%) confidence interval. Taken from [12].
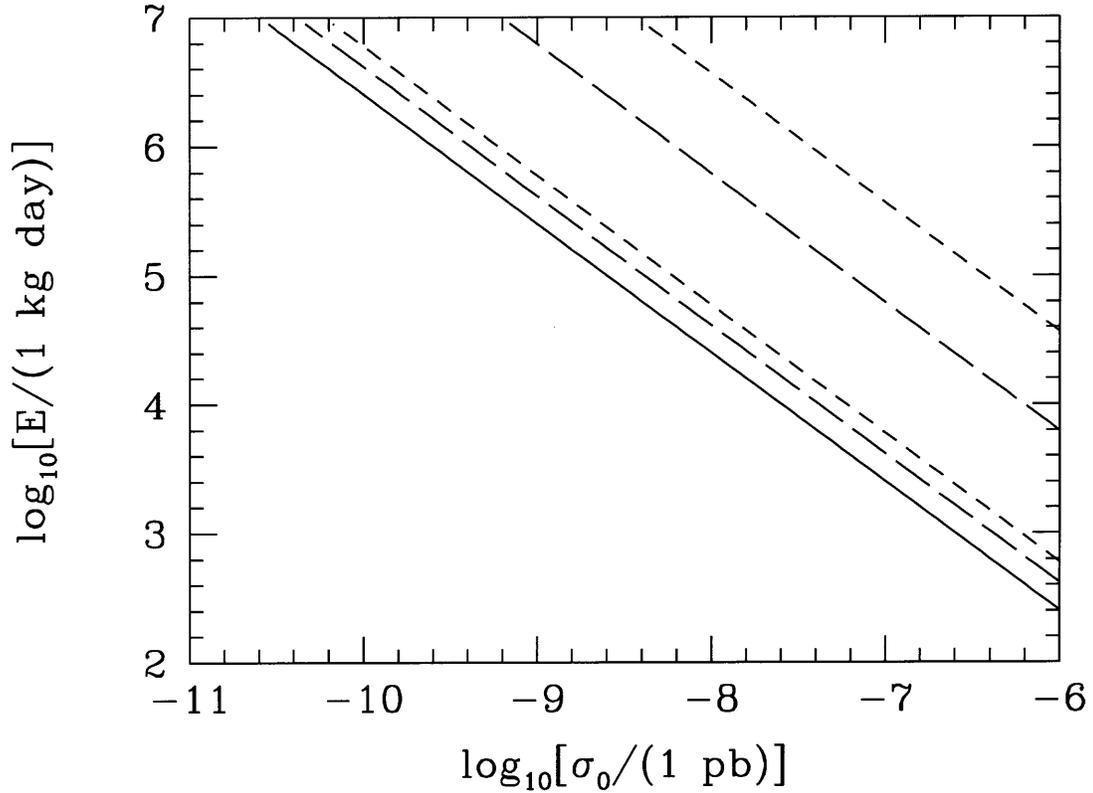
Figure 1-5: Plot of the required exposure to reject isotropy with 90% confidence. The solid black line is the benchmark detector: a CS_2 gaseous directional detector at 40 Torr with a 3-D vector read-out, an energy threshold of $E_T = 20$ keV and no background. The short dashed lines use raw angle and the long dashed lines use reduced angles. The two lines in the top right are for axial read-outs only where the two dashed lines nearer the benchmark line are for vector read-outs. The ability to reconstruct vector direction vs. only axial direction decreases necessary exposure by a factor of ~15. Further, the ability to reconstruct 3-D angle vs. only 2-D angle decreases necessary exposure by a factor of ~2. Figure taken from [12]

Furthermore, this necessary exposure may be further reduced by a factor of ~2 by determining the 3-D angle of the nuclear recoil. Therefore, the combined decrease in necessary exposure for being able to reconstruct both head-tail and the 3-D angle allows a factor of ~30 fewer events. This again highlights the importance of a robust reconstruction algorithm that accurately determines recoil sense or head-tail.

Chapter 2

Experimental Apparatus and Data Acquisition

2.1 DMTPC Detector Setup

The DMTPC is a low-pressure time projection chamber which uses CF_4 gas at 75 Torr as a target material. The chamber has two readouts: a CCD camera that images scintillation light produced in the amplification region and a charge amplifier connected to the anode that measures ionization generated by a particle moving through the detector. The CCD camera is mounted on the outside of the vacuum vessel and the charge readout electronics are located inside the vessel. Currently, the structure of the mesh-based amplification allows for two-dimensional imaging of charged particles, however, research is currently underway to image the three-dimensional track. CF_4 gas has an average ionization energy of 34 eV and therefore nuclear recoils produce on order of 10^3 primary electrons which are then drifted via electric fields toward the amplification region. As the fields intensify near the amplification region, the electrons accelerate and cause the ionization and excitation of other CF_4 molecules, creating an avalanche effect. The result of this avalanche multiplication is that when excited CF_4 molecules fall to the ground state, scintillation photons are produced (see Figure 2.1). Tracks may then be reconstructed to ascertain 2D-angle, head-tail, range and energy of the recoiling nucleus [5].
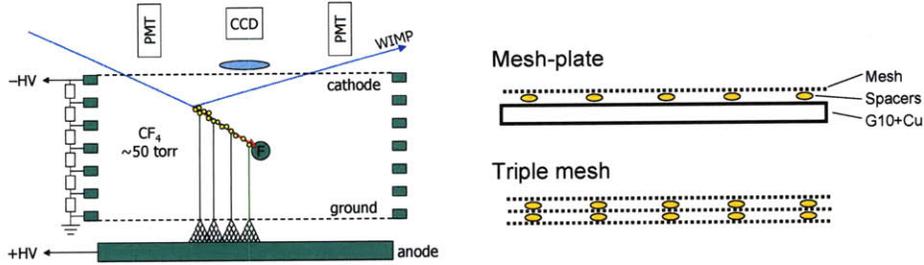
Figure 2-1: (Left) A schematic of an incident WIMP elastically scattering off of a nucleus and the resulting drift of electrons. (Right) Two designs for the amplification region: "mesh-plate" (top) and "triple-mesh" (bottom). Figure taken from [7].

Figure 2.1, is a schematic of the Dark Matter Time Projection Chamber detector. A dual-TPC is housed within a stainless steel vacuum vessel. In this work, we consider only the track reconstruction in images produced by the CCD cameras. Our drift region has been created using a woven mesh cathode, typically of a potential -5 kV which is separated from a wire mesh (28 μm wire, 256 μm pitch) ground grid 20 cm away. The chamber's vertical drift field is kept uniform within a tolerance of 1% by stainless steel field-shaping rings which are spaced 1 cm apart (see Figure 2.1).

An amplification region is formed between the ground grid and a copper-coated, G10 anode plane (at 720 V) which are separated from each other by 500 μm using resistive spacers. Our CCD camera is an Apogee Alta U6 with a thermoelectric cooler that maintains the temperature within the range [-20, -18]C. The read noise is 12 electrons/pixel. The photographic lens has an aperture ratio of f/1.2 and a focal length of 55 mm [4].

CF$_4$ at a pressure of 75 Torr allows for gas gains of approximately 10^5 without significant sparking. Using the current reconstruction algorithm our energy resolution for charge readout is 10% at 5.9 keV (measured using a ^{55}Fe source) and 15% at 50 keV for the CCD readout (measured using an alpha source). CF$_4$ was chosen since it has many advantageous properties as a target gas for a dark matter TPC. First, as
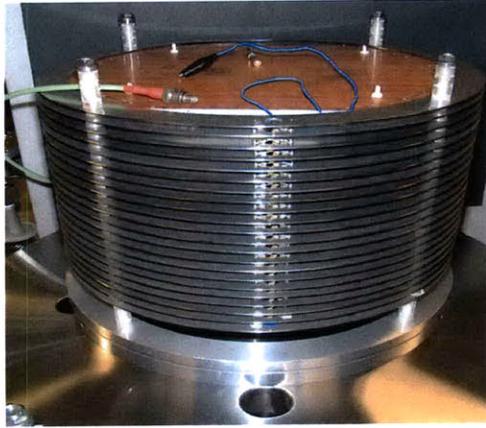
Figure 2-2: Image of the 10-liter DMTPC prototype. Figure taken from [6].

mentioned above, it has strong sensitivity to spin-dependent interactions due to the unpaired proton. The scintillation spectrum of CF_4 is broad and peaked near 625 nm (closely matched to the peak quantum efficiency of CCDs) [3]. Further, CF_4 gas is non-toxic, non-flammable and perfect for use underground (a necessary requirement since our experiment is performed in a mine). Finally, transverse diffusion of electrons in CF_4 is highly suppressed and there is minimal electron attachment over a 20 cm drift length [8].

2.2 Charge-Coupled Device (CCD) Readouts

2.2.1 CCDs in DMTPC

A good understanding of the operation and properties of Charge-Coupled Devices (CCDs) is important to DMTPC because the detector uses CCDs to optically image nuclear recoil events. However, CCD readouts do have many different types of noise that must be understood and isolated. The sources of noise in our readout are the

Figure 2-3: (Left) Photograph of the 10-liter DMTPC detector with an overlaid image of the dual TPC. Inside the chamber, stainless steel field shaping rings modulate the drift field. The CCD cameras are located outside the chamber on the top and bottom where they each image an amplification region. Image taken from [11].

Poisson noise associated with signal photons, read noise and dark noise (or dark current) both in and off the CCD chip; each will be elaborated upon the Efficiency and Noise subsection.

2.2.2   Operation of CCDs

Silicon CCDs operate on the principle of the photoelectric effect. The photoelectric effect is the process in which an incoming photon with an energy exceeding that of the binding energy of an electron in the material, ejects an electron from its current state. In a CCD, the incoming photons strike the CCD pixel and are absorbed if they possess the correct wavelength. CCDs are made of silicon, which is a semiconductor and thus has the property that its electrons are confined to remain within certain bands of energy within the material. Specifically, two major energy bands exist within semiconductors and are referred to as the valence and conduction band. The valence band is the highest range of energies that electrons are present in a material at absolute zero. Within the valence band, electrons are constrained to individual atoms. The conduction band is a higher energy band, separated from the valence band by what is referred to as the band gap. Electrons within the conduction band are free to move throughout the atomic lattice and are not bound to any individual atoms. Silicon has a band gap of 1.14 eV and so it easily absorbs light within the energy range 1.1 to 4eV (corresponding to wavelengths of 11,000 to 3,000 Å). The photon absorption within this energy range ejects an electron from the valence band and moves it into the conduction band.

Once freed to the conduction band, however, these electrons must be held in place in order to be collected from the readout system. This is done by applying voltages onto subpixel sized electrodes called gates. Gate structures hold the electrons in place until they are collected from the readout system and are amplified into a signal. The collection from the readout system includes a complex process of cycling voltages in order to transfer the charge into an output register, however, the details of the operation are not pertinent to this paper. This output voltage for each given pixel must then be converted into a digital number and from then on is measured in counts

or ADUs (Analog-to-Digital Units). The amount of voltage needed to correspond to a single ADU is referred to as the gain of the the CCD.

2.2.3 Efficiency and Noise

With the basic mechanics of the CCD in place, it is necessary to understand the efficiency of conversion of signal and of the types of noise inherent to this measurement process. We measure the CCD's sensitivity to light in terms of its quantum efficiency. Quantum efficiency of a CCD is defined as the ratio of photons detected by the device to the number of incoming photons. For our detector, the quantum efficiency of the CCD is 65% at 600 nm, well matched to the scintillation spectrum of CF_4. However, as mentioned above, many types of noise are manifest within this measurement process and we elaborate upon them here.

Read noise is typically quoted in terms of electrons/pixel. Read noise has two primary sources. The first, referred to as on-chip read noise, arises in the conversion of an analog signal to a digital number. On-chip amplifiers produce some statistical distribution around the initial signal; this process is not perfectly repeatable in practice so multiple read-outs with the same initial charge on the pixel will give different digital values. Second, read noise manifests itself "off the chip" in the electronics of the device. Specifically, noise is added to the signal as it passes through the camera's circuitry to the Analog to Digital Converter (ADC) to be converted into a digital pixel value. The sum of these two independent effects is referred to as the read noise of the device. For DMTPC, our total read noise is 13 electrons/pixel.

Dark noise is the result of thermally generated electrons that build up in the pixels of the CCD. Any material held at a temperature above absolute zero will be subject to this thermal noise. When the thermal agitation is large enough, electrons are freed from the valence band of the silicon chip and are collected within the potential well of a given pixel. Therefore, to reduce this effect, cooling is necessary for low noise imaging. However, given our current imaging setup, the noise primarily arises from read noise since dark noise does not become significant until longer exposure times.

And finally, the last type of noise intrinsic to our read out is the Poisson noise

associated with the signal photons. Poisson noise is a result of the quantum nature of light - individual photons, which are distributed according to Poisson statistics, constitute the signal. Because this is an intrinsic property of light, improving electronics or other techniques cannot remove this source of noise. The effects of this noise may be summarized in the equation of the signal-to-noise ratio unofficially termed the "CCD Equation" [14]

$$\frac{S}{N} = \frac{N_*}{\sqrt{N_* + n_{pix}\left(N_S + N_D + N_R{}^2\right)}} \qquad (2.1)$$

where S/N is the signal-to-noise ratio, N_* is the total number of photons collected from the event, n_{pix} is the number of pixels, N_S is the total number of photons per pixel from the background, N_D is the total number of dark current electrons per pixel, and $N_R{}^2$ is the total number of electrons per pixel from the read noise. This equation has the expected properties, for instance, if the total noise is dominated by the signal, N_*, then the S/N equation reduces to

$$S/N \approx \frac{N_*}{\sqrt{N_*}} = \sqrt{N_*} \qquad (2.2)$$

which is the expected Poisson characteristic in the signal dominated regime. In the situation where the read noise dominates, the Equation 2.1 becomes

$$S/N \approx \frac{N_*}{\sqrt{N_R{}^2}} = \frac{N_*}{N_R} \qquad (2.3)$$

and intuitively the signal-to-noise ratio becomes just the ratio of the signal to the read noise added into the system. Finally, we may consider the regime in which dark noise dominates, $N_D >> N_*, N_R$

$$S/N \approx \frac{N_*}{\sqrt{n_{pix}N_D}} = \frac{N_*}{\sqrt{n_{pix}N_D}}. \qquad (2.4)$$

Here we see that noise scales differently than the read noise component, instead the noise scales as $\propto \sqrt{N_D}$. For DMTPC, we operate within the regime where Poisson

noise from the signal photons dominates both read noise and dark noise (Equation 2.2).

Chapter 3

A Physical Model for Nuclear Recoils

3.1 Reconstructing Nuclear Recoils

Our current software package of extracting track parameters from nuclear recoils is named AnalysisFramework and although it uses sophisticated and well-tested algorithms, it is not based on the physics of the interaction. The purpose of this paper is to develop and evaluate a physically motivated fitting functions that can be used in the reconstruction of nuclear recoils and to optimize the extraction of track parameters. Because this new algorithm utilizes convolutions in its implementations, it is termed ConvolutionFramework.

3.1.1 Current AnalysisFramework

AnalysisFramework and ConvolutionFramework are operated in ROOT, a program built upon the C++ language at CERN which is designed to efficiently handle and analyze large amounts of data. The procedure for data analysis is as follows: first, AnalysisFramework takes in raw CCD images (either actual events or Monte Carlo generated events) and runs a series of processing functions. These processing functions include creating background images from empty images to find the overall bias of the

CCD. The bias of the CCD is directly attributable to the DC voltage maintained by the camera electronics in order to keep the signal detected from the ADC from ever being negative. Therefore, this bias has no physical meaning in the measurement of nuclear recoils and is subtracted off in order to do analysis on unbiased image (background noise levels average to 0 over the image).

Additionally, AnalysisFramework removes hot pixels, that is, pixels which are consistently above a high threshold from event to event. Events containing sparks, electrical discharges inside the DMTPC chamber that occur at high voltages, are also removed from the data set. With this complete, a pattern recognition algorithm is run over individual images to identify nuclear recoil candidates. Tracks are defined as five or more contiguous pixels above some threshold, currently set to be a group of contiguous pixels above 3.7σ, where σ is the standard deviation in the intensity of the pixels on the image.

Once individual tracks are found, parameters of each track are then extracted. The primary parameters of interest are the energy, length, in-plane angle (ϕ) and head-tail. Energy is determined by integrating the value of all bins within the identified track. Length is found by extracting all pixels above threshold and then the distance of the two farthest pixels is the set as the length. Angle is determined by constructing the moment of inertia tensor and then finding the eigenvalues. The moment of inertia tensor is defined as

$$\mathbf{I} = \begin{pmatrix} \langle x^2 \rangle & -\langle xy \rangle \\ -\langle xy \rangle & \langle y^2 \rangle \end{pmatrix} \tag{3.1}$$

where

$$\langle x^2 \rangle = \sum_i p_i \left(x_i - x_0 \right)^2$$
$$\langle xy \rangle = \sum_i p_i \left(x_i - x_0 \right) \left(y_i - y_0 \right)$$
$$\langle y^2 \rangle = \sum_i p_i \left(y_i - y_0 \right)^2$$

30
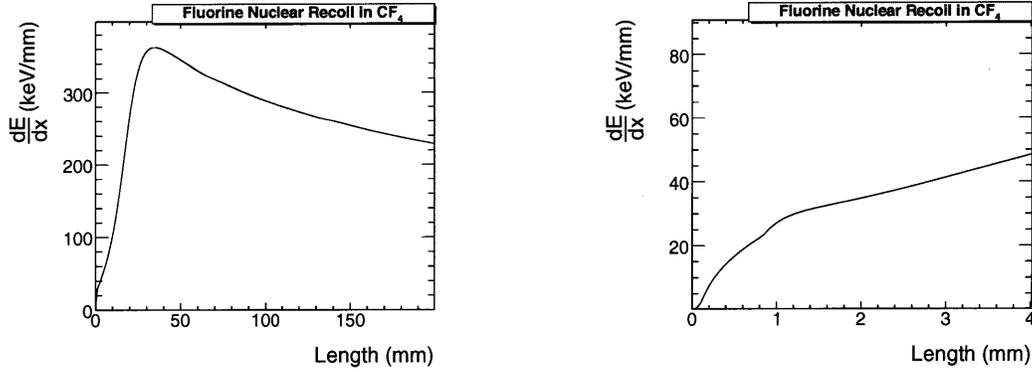
Figure 3-1: (Left) Ionization energy deposition of Fluorine nuclear recoil (also known as the Bragg curve). The point on this curve of highest energy loss is referred to as the Bragg peak. (Right) Zoomed image of left chart over the range of interest ($\approx$ few mm). The stopping power, $\frac{dE}{dx}$ is approximately linear for these low energy recoils and thus justifies our fitting algorithm (described later).

and p_i is the intensity of pixel i, x_i is the x-position of pixel i, x_0 is the x-mean position of the track, y_i is the y-position of pixel i and y_0 is the y-mean position of the track. It should be noted that this approach is equivalent to maximizing the moment of inertia of the two dimensional track with respect to an angle ψ through the center of the track.

Length is determined by finding the two most distant points in the contiguous cluster and calculating the distance between these. Head-tail is determined by splitting the detected cluster into two bins. Then, the bin with the higher integrated flux is assigned as the head of the track and the lower peak is assigned as the tail of the track.

3.1.2   Development of ConvolutionFramework

In order to better extract parameters from these tracks, this research tested whether the a priori knowledge of the physical interaction could result in more accurate track reconstruction. Specifically, the information inputted is the shape of the energy deposition of a recoiling F nuclei as a function of distance (dE/dx vs. x) through a gas of particles. Figure 3-1 shows the SRIM calculation for dE/dx vs. x for a F nucleus in 75 Torr CF_4.

The nuclear recoils of interest for our group lie ($E_{recoil} < 200$ keV) well below the Bragg peak. In this regime, dE/dx vs. x is linear (Figure 3-1 right). As described in Section 2.1, our time projection chamber is subject to transverse diffusion due to electron drifts. Transverse diffusion is the process where electrons produced in an ionization process are perturbed by random thermal fluctuations and they approach a thermal energy distribution corresponding to that of the surrounding gas. Their average energy, ϵ, becomes

$$\langle \epsilon \rangle = \frac{3}{2}kT \tag{3.2}$$

where k is the Boltzmann constant and T is the temperature of the gas in Kelvin. The electrons will be distributed over energy according to a Maxwell-Boltzmann distribution

$$F(\epsilon) = \text{const} \cdot \sqrt{\epsilon} \exp\left(\frac{-\epsilon}{kT}\right) \tag{3.3}$$

Locally produced ionization diffuses through the gas, resulting in a Gaussian uncertainty in initial position

$$\frac{dN}{N} = \frac{1}{\sqrt{4\pi Dt}} \exp\left(\frac{-x^2 - y^2}{4Dt}\right) dx dy \tag{3.4}$$

where $\frac{dN}{N}$ is the fraction of charge found in a length element $\sqrt{dx^2 + dy^2}$ at a distance $\sqrt{x^2 + y^2}$ at time t and D is the diffusion coefficient [13]. However, in our detector, a voltage is applied across the height and so consequently the electrons drift towards the anode with velocity v

$$v = \mu E \left(\frac{p_0}{p}\right) \tag{3.5}$$

where μ is the electron mobility, a constant of proportionality between the drift velocity and the electric field, E is the electric field in the chamber, p_0 is a normalized pressure and p is the pressure of the gas. Using this velocity, v, we may then plug this in for $t = z/v$ in Equation 3.4 and arrive at

$$\frac{dN}{N} = \frac{\mu E p_0}{4\pi D z p} \exp\left(\frac{-x^2 - y^2}{4Dz} \frac{p_0}{p}\right) dx dy, \tag{3.6}$$

where z is the vertical drift of the electrons along the field lines. Therefore, the width of this diffusion is

$$\sigma_{tot}^2 = \int \left(x^2 + y^2\right) \cdot A \exp\left(-\left(x^2 + y^2\right) \frac{\mu E}{4Dz} \frac{p_0}{p}\right) dx dy = 2 \left(\frac{2Dz}{\mu E} \frac{p_0}{p}\right), \tag{3.7}$$

where A is the normalization factor

$$A = \sqrt{\frac{\mu E}{4\pi D z} \frac{p_0}{p}}. \tag{3.8}$$

Therefore, the functional form of the energy deposited perpendicularly to the track direction may be written as a Gaussian

$$\frac{dE}{dy} = \text{const} \cdot \exp - \left(\frac{(y - y_0)^2}{2\sigma^2}\right), \tag{3.9}$$

where y is the distance perpendicularly from the track, y_0 is the position of the track and $\sigma = \frac{2Dzp_0}{\mu E p}$.

Furthermore, the approximated functional form of energy deposited along the length of the track should be that of a convolution of a Gaussian with a line. For a convolution of two general functions, $f(\cdot)$ and $g(\cdot)$, it is defined as follows

$$f * g = g * f \equiv \int_{\tau=-\infty}^{\infty} f(\tau) g(t - \tau) \, d\tau. \tag{3.10}$$

For our example the convolution then becomes a convolution of the two functions, $f(x)$ and $g(x)$ defined as

$$\begin{aligned} f(x) &= mx + b, \\ g(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/2\sigma^2}. \end{aligned}$$

Explicitly writing out the convolution equation we have

$$f(x) * g(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\tau=-\infty}^{\infty} [m(x-\tau) + b] e^{-\tau^2/2\sigma^2} d\tau . \tag{3.11}$$

$$\frac{dE}{dx} = y_0 + \frac{1}{\sqrt{2\pi\sigma^2}} \int_{t=x_0-x}^{x_0+l-x} [m(x-t) + b] \cdot \exp\left(-t^2/(2\sigma^2)\right) dt. \tag{3.12}$$

which when solved explicitly may be written as

$$\frac{dE}{dx} = y_0 + \frac{\sigma}{2\sqrt{\pi}} \left[e^{-(\frac{x-x_0}{\sigma})^2} - e^{-(\frac{x-x_0-l}{\sigma})^2} \right] + \frac{mx+b}{2} \left[\text{Erf}(\frac{x-x_0}{\sigma}) - \text{Erf}(\frac{x-x_0-l}{\sigma}) \right],$$
$$\tag{3.13}$$

where

$$\text{Erf(x)} \equiv \frac{2}{\sqrt{\pi}} \int_0^x \exp\left(-t^2\right) dt, \tag{3.14}$$

and y_0 is the offset bias of the total image (the underlying noise level in the CCD with a 0-exposure time), σ is the standard deviation in position as a result of diffusion processes, m is the slope of the convolved line, b is the y-intercept of the convolved line, x_0 is the starting position of the convolved line and l is the length of the convolved line along the x-axis. However, this functional form is only an approximation for the projection of the line parallel to the track through space. Orthogonal to the particle's trajectory, diffusion processes perturb the electron and the position is described by a Gaussian distribution. This information can be extracted relatively easily by breaking the fit into a series of two projections. Further, by breaking the fit of the projection of the track into a series of two one-dimensional fits, much of the complexity of two-dimensional fits may be avoided. Information regarding the amount of diffusion can be extracted from the orthogonal gaussian distribution; specifically, the standard deviation of this orthogonal projection may be fixed as a parameter in the gaussian convolved with the line to arrive at more accurate fits.

3.1.3 ConvolutionFramework Algorithm

This code is developed to be used with the currently implemented AnalysisFramework tool set developed by the DMTPC group. To begin, a Monte Carlo code developed by members of the DMTPC group is used to create thousands of instances of individual nuclear recoil events. Once this has been generated, AnalysisFramework is run upon the Monte Carlo result. The output is a processed file of all recognized tracks within the Monte Carlo data set. Image processing includes several procedures specific to this experiment. First, AnalysisFramework runs a track detection algorithm as defined in Section 3.1.1. Next, the algorithm searches for cosmic events termed by the DMTPC group to be worms, that is events that occur when cosmic particles pass through the silicon CCD and leaving one pixel of value signficantly above all others. Once the removal of worms is complete, AnalysisFramework subtracts off the bias of the image. And finally, a blurring method is applied to the image where a Gaussian noise is multiplied across contiguous bins to create higher continuity between pixels.

Once all tracks have been identified and processed by AnalysisFramework, the results are fed into ConvolutionFramework. First, the image of each identified track histogram is fit with a best fit line using standard ROOT fitting functonality. ConvolutionFramework performs two projections, so the projection area must first be defined. The projection region is constructed to be wide enough in order to get an accurate measure of the baseline while not being too large so as to introduce unnecessary noise in the projection histograms. To construct this, a box is placed around the track and then each point within the box is sampled and checked whether it is within five bins of the identified track. The projection boundary is then defined as that region which is within the initial box and exactly five bins distance from the nearest point on the track. The two-dimensional track histogram is then projected orthogonally to the best fit line.

The result is a gaussian distribution orthogonal to the projected path of the particle and centered on the axis of the line. The fit parameter σ, the standard deviation of the gaussian, is then saved to be later used as a fixed parameter for the fit parallel
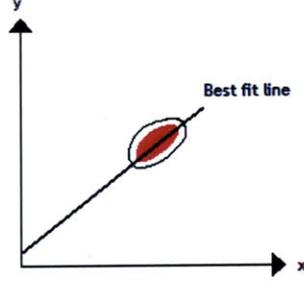
Figure 3-2: Schematic of the fitting box and the best fit line through the nuclear recoil track (red).

to the track's axis. Next, the two-dimensional track histogram is projected parallel to the initial best fit line. The functional form of this second fit is a line convolved with a gaussian, where the gaussian has a fixed parameter, σ, which is set from the orthogonal fit. This fit has five free parameters, y_0, x_0, l, m, and b as defined above.

In order for ROOT to complete both the orthogonal and parallel fits, it is necessary to feed it guess parameters. The guesses for the orthogonal gaussian,

$$\frac{dE}{dy} = y_0 + \frac{A}{\sqrt{2\pi\sigma_\perp}} \cdot \exp\left(\frac{(y - y_0)^2}{2\sigma_\perp{}^2}\right) \tag{3.15}$$

were determined by using the following methods

$$A = \max(H_\perp(x)) \tag{3.16}$$

$$x_0 = \operatorname{argmax}(H_\perp) \tag{3.17}$$

$$\sigma_\perp = \min\left(\sigma_{\text{imageRMS}}, \frac{1}{4}(x_{\text{maxbin}} - x_{\text{minbin}})\right) \tag{3.18}$$

$$y_0 = \operatorname{mean}(\text{image pixel values}) \tag{3.19}$$

36

where,

$$\text{argmax}\left(f(x)\right) \equiv x \mid \forall y : f(y) \leq f(x) \tag{3.20}$$

and $H_\perp(\cdot)$ is the perpendicular histogram value, σ_{imageRMS} is the image RMS, x_{minbin} is the minimum bin on the projection and x_{maxbin} is the maximum bin on the projection. The guess parameters for the parallel projection, Equation 3.13, were determined using the following

$$y_0 = \text{mean(image pixel values)} \tag{3.21}$$

$$m = 0 \tag{3.22}$$

$$x_0 = x \ni H_\parallel(x) > 5\sigma_{\text{RMS}} \tag{3.23}$$

$$l = x_1 - x_0, x_1 \ni H_\parallel(x) < 5\sigma_{\text{RMS}} \tag{3.24}$$

$$b = \max(H_\parallel) \tag{3.25}$$

where σ in 3.13 is a fixed parameters and is set to $\sigma_\perp$, $H_\parallel(\cdot)$ is the parallel histogram value, σ_{RMS} is the RMS value of the histogram. In addition to the fitting guesses, we restrict the parameter space to be only physically possible values. Therefore, we implement the following parameter constraints,

$$A \in [0, \infty) \tag{3.26}$$

$$x_0 \in [x_{minbin}, x_{maxbin}] \tag{3.27}$$

$$y_0 \in [y_{min}, y_{max}] \tag{3.28}$$

$$l \in [0, \infty) \tag{3.29}$$

$$\sigma \in [0, \infty) \tag{3.30}$$

and the variables for slope, m, and y-intercept, b are not constrained.

Chapter 4

Analysis

Several factors intrinsic to the experiment complicate the analysis and reconstruction of these tracks. In particular, these features include various types of readout noise, out-of-plane angle (θ), and drfit length. The angles are defined in the coordinate system in Figure 4.

An understanding of these effects on the two parameter extraction algorithms, AnalysisFramework and ConvolutionFramework, is crucial.

4.1 Complicating Factors

The accurate extraction of true parameters of nuclear recoils contains many difficulties, especially at the important low energy scale (100 keV and less). First, our experiment is subjected to read noise, Poisson noise, and dark noise as described by Equation 2.1. Currently, our setup has a total read noise of 7.3 ADU/pixel, i.e. each pixel has an additional randomized value added to the initial signal sampled from a Gaussian distribution with standard deviation 7.3 ADU. Second, our experiment is being redesigned in order to reconstruct the out-of-plane angle of the track through the chamber by using charge readout. However, as it stands currently, DMTPC is only able to image a two-dimensional projection of the nuclear recoil through the chamber. Therefore at steep angles, i.e. $\theta \in [0, 30] \cap [150, 180]$, important information encoding the head-tail and length of the track are lost. Thirdly, each pixel in
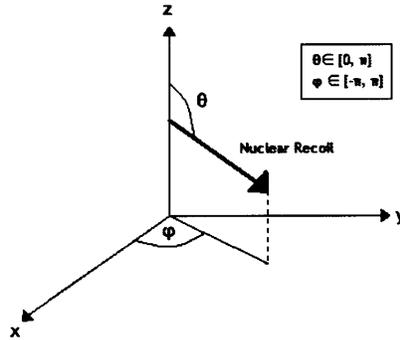
Figure 4-1: Schematic of our coordinate system for nuclear recoil reconstruction. Here, the thick dark line represents a recoiling nucleus with the arrow displaying direction of motion.

our camera images 0.143 x 0.143 mm^2 of the amplification region. On readout, the camera groups pixels 4x4 into a single bin. At low energies, tracks may only extend over 2-3 bins, presenting a serious issue for a fitting function. Figure 4-2 is an example of the typical length versus energy relation.

Finally, once all these issues are addressed and the algorithm detects and extracts parameters, the results of the fit may not be statistically significant and must be abandoned. As mentioned before, one primary feature we are concerned with is the head-tail of the nuclear recoil. The DMTPC detector has been shown to be capable of determining head-tail to low energies, This information is encoded within the slope of the Equation 3.13. In the simulations performed, a negative slope, $m < 0$, in the convolution function meant a proper reconstruction.
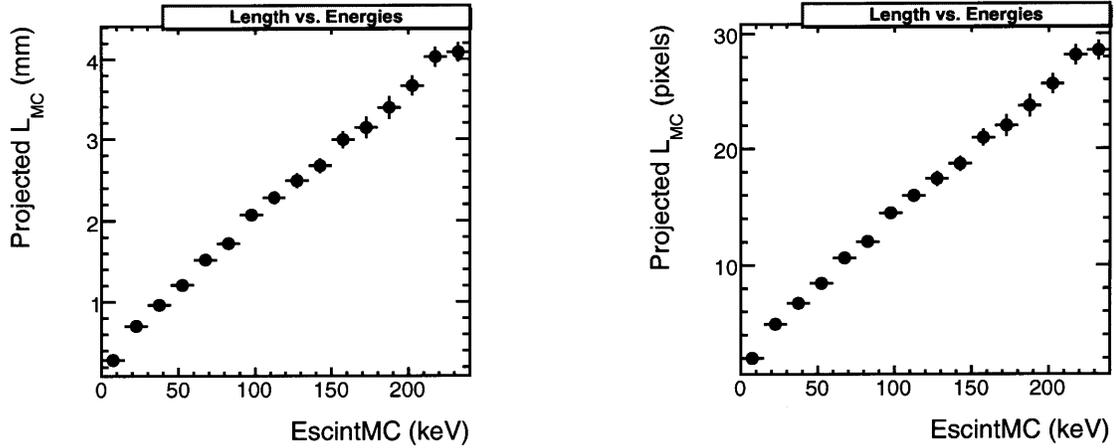
Figure 4-2: 12,000 Event Monte Carlo which demonstrates the length versus energy characteristic of the nuclear recoils. EscintMC is defined as the amount of energy left in the scintillation process. Projected $L_{MC} = L_{MC} \cdot \sin\theta$, i.e. the projection of the true Monte Carlo generated length of the track in the plane of the detector. Note events below 100 keV only span approximately 4 superpixels (excluding effects of drift) and tracks of energy 50 keV only span 2 superpixels, complicating fitting algorithms.

4.2 Fitting Results

Figure	Read Noise Level	Energy Level
4-3	High	High
4-4	High	Low
4-5	Low	High
4-6	Low	Low

In each of the primary fitting screens, there are four subplots. In the top left corner is the raw image of one Monte Carlo event before processing by AnalysisFramework. The red box is centered on the x_0 and y_0 position of the head of the track. The top right corner is this image duplicated with the best fit line drawn through the track. Projections are then carried on using this line. In the bottom left, the histogram of the projection perpendicular to the best fit line is displayed. Overlaid on the histogram is the fitted Gaussian result which is used to derive the diffusion coefficient. Finally, in the bottom right is the histogram of the projection parallel to the best fit line.

Overlaid on this histogram is the fitted convolution function, equation (3.13), in black. The red vertical lines represent the limits of the bin summation in order to ascertain the energy.

To test the sensitivity of the algorithms as a function of the read noise, we investigated the dynamics of each as the noise ramped up from 0.5 ADU/pixel read noise up to the full 7.3 ADU/pixel read noise. These Monte Carlo simulations are generated with 12,000 events each. The parameters inputted are 20cm drift region, isotropic in the in-plane angle (ϕ) and isotropic in the z-coordinate angle (θ). The read-noise is run on noises of 0.5 ADU/pixel, 2.0 ADU/pixel, 3.65 ADU/pixel and finally the full read noise of 7.3 ADU/pixel.
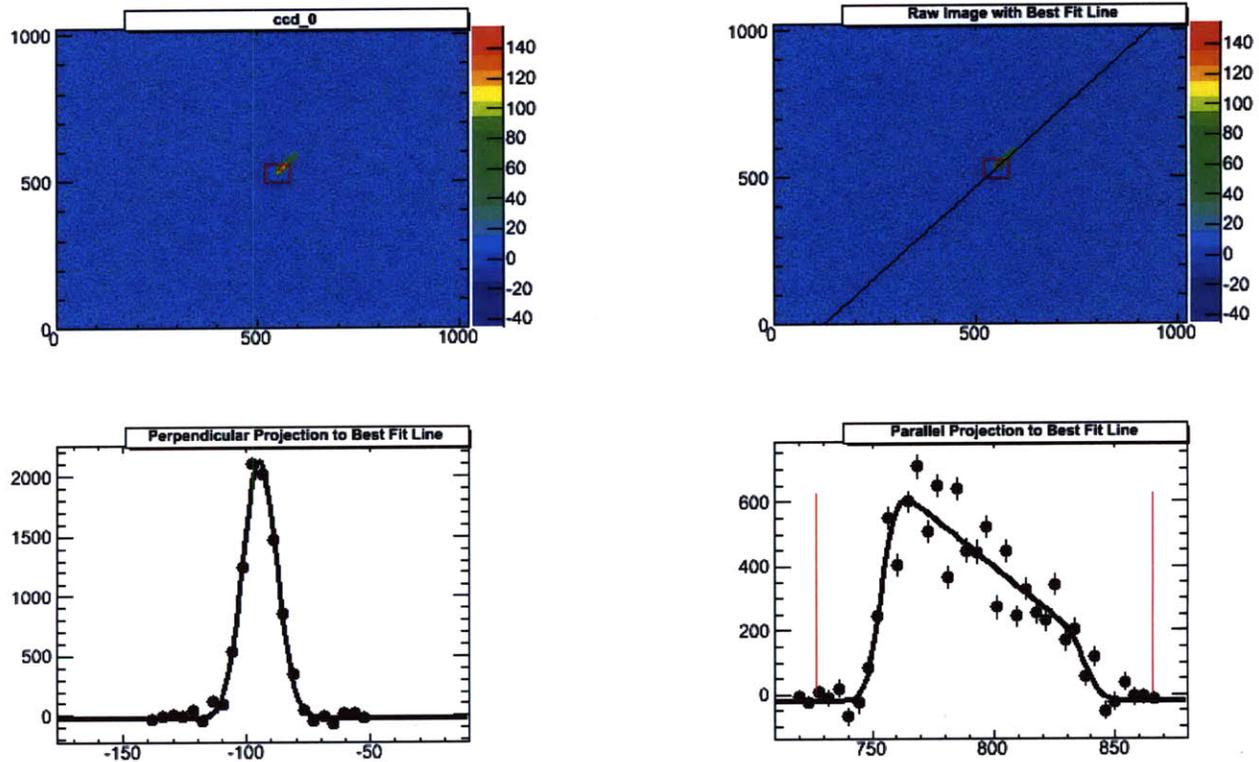
Figure 4-3: An example of a high energy nuclear recoil track, 983 keV, with high read noise of 7.3 ADU/pixel. At these high energies, the signal of the nuclear recoil is easily selected and head-tail is readily reconstructed.
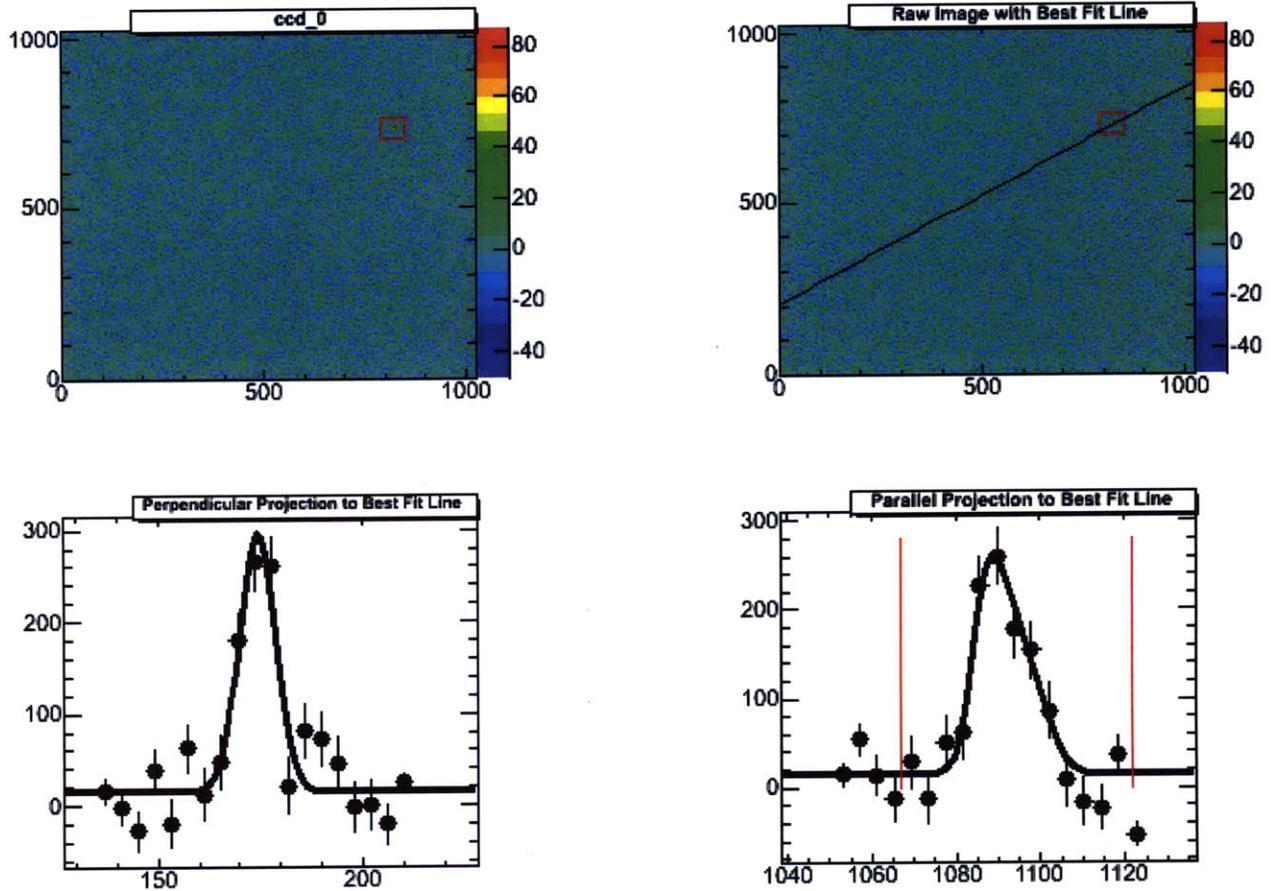
Figure 4-4: An example of a low energy nuclear recoil track, 50 keV, with high read noise of 7.3 ADU/pixel. In the low energy regime, 50 keV and lower, we can still, in certain circumstances detect and reconstruct head-tail.
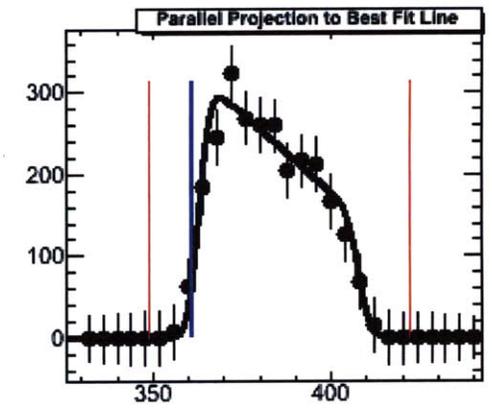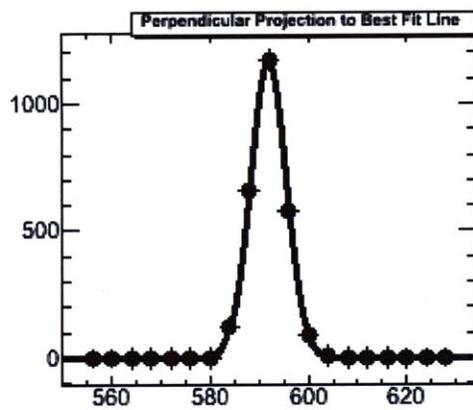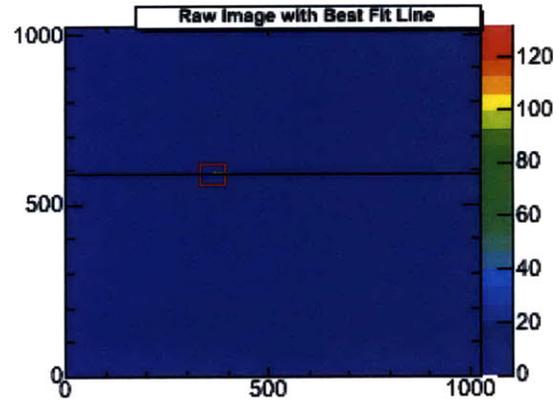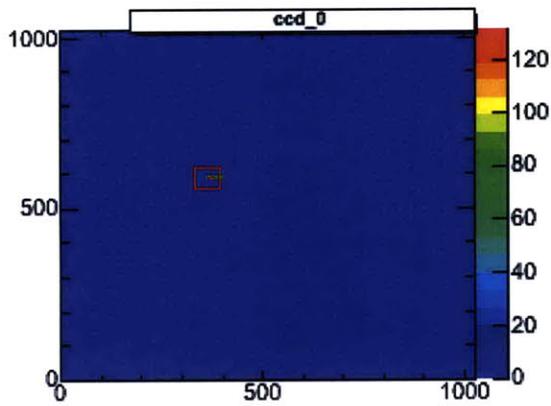
Figure 4-5: An example of a high energy nuclear recoil track, 270 keV, with low read noise of 0.0001 ADU/pixel. Now with the read noise nearly eliminated the head-tail is easily reconstructed for high energy events.
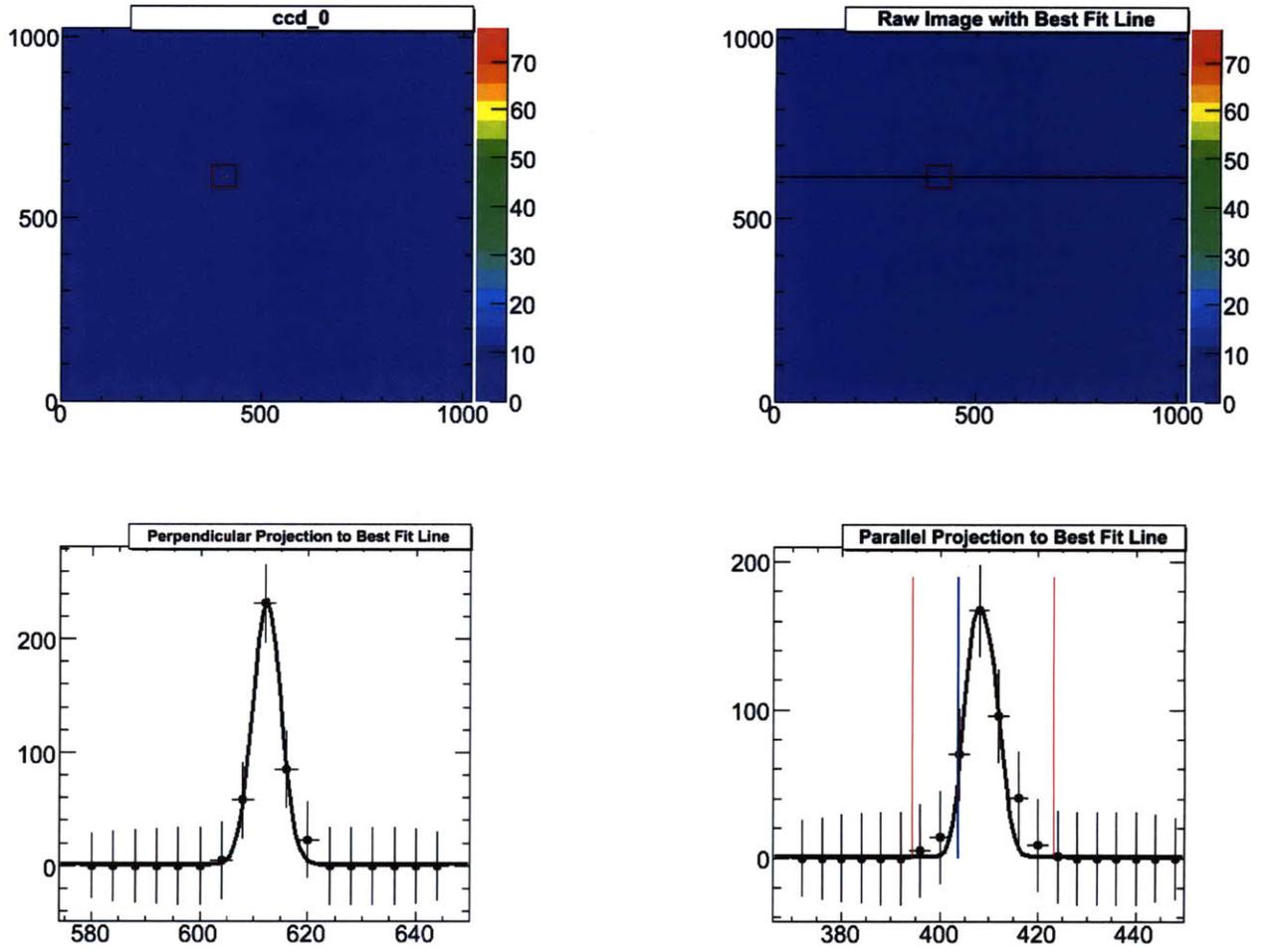
Figure 4-6: An example of a low energy nuclear recoil track, 40 keV, with low read noise of 0.0001 ADU/pixel.

The Monte Carlo Simulation generates recoils in the expected amount generated according to a spherically symmetric nuclear recoil spectrum to Equation 1.14. Additionally, the track recognition efficiency approximately follows a sigmoidal curve and so for small E_T few events are detected but eventually the efficiency reaches 100% as $E_T \rightarrow \infty$. However, we choose to conduct analysis over a roughly even distribution of number of events vs. recoil energy so that certain regimes are not statistics limited. Therefore, we arrive at the following energy distributions, Figures 4-7, 4-8, 4-9 and 4-10.
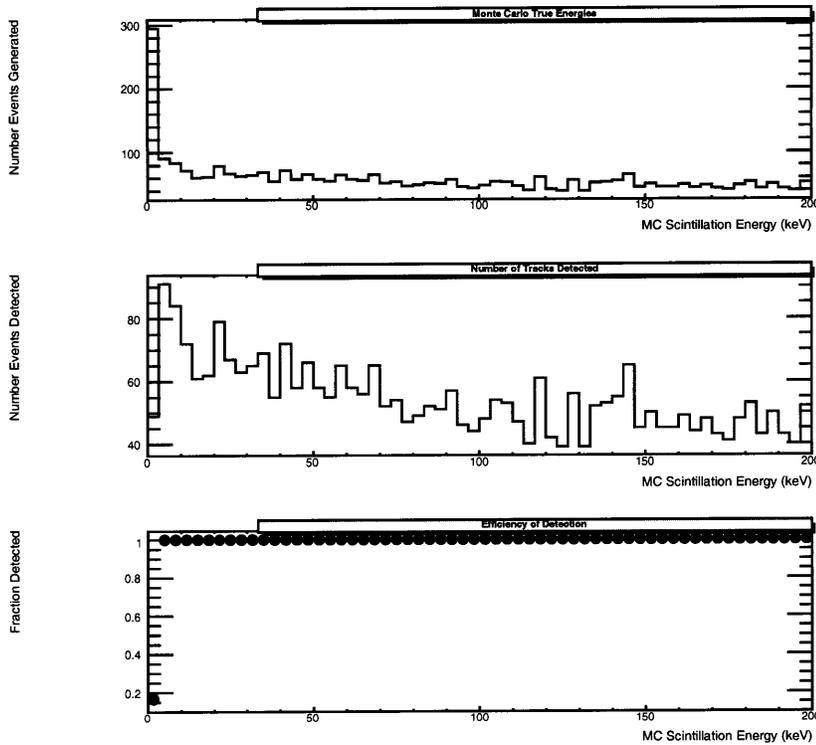
Figure 4-7: Efficiency of detection for 0.5 ADU/pixel read noise Monte Carlo Simulation.

4.3 Read Noise Dependence

We now address the effect of CCD read noise on the nuclear recoil reconstruction ability of both AnalysisFramework and ConvolutionFramework. To begin, we summarize

how the values of the nuclear recoil are extracted using ConvolutionFramework. First, head-tail was determined to be correct in a relatively simple fashion. In the simulations generated for this thesis, all sources were incident from the left side of the chamber or $+\hat{x}$ direction in Figure 4. This implies that a correct head-tail reconstruction is one in which the line of the Equation 3.13 has a negative slope. This is directly derived from the characteristic Bragg curve for a neutron induced nuclear recoil at these energies, Figure 3-1.

L_{recon} was found directly by taking the Δx of the Convolution line after the fit on the image. And finally, E_{recon} was determined by integrating over the histogram, $H_\parallel$, on the length,

$$l_{integration} = [x_0 - 2\sigma, x_1 + 2\sigma] \qquad (4.1)$$

where x_0 is the x position of the beginning of the Convolution line, x_1 is the x
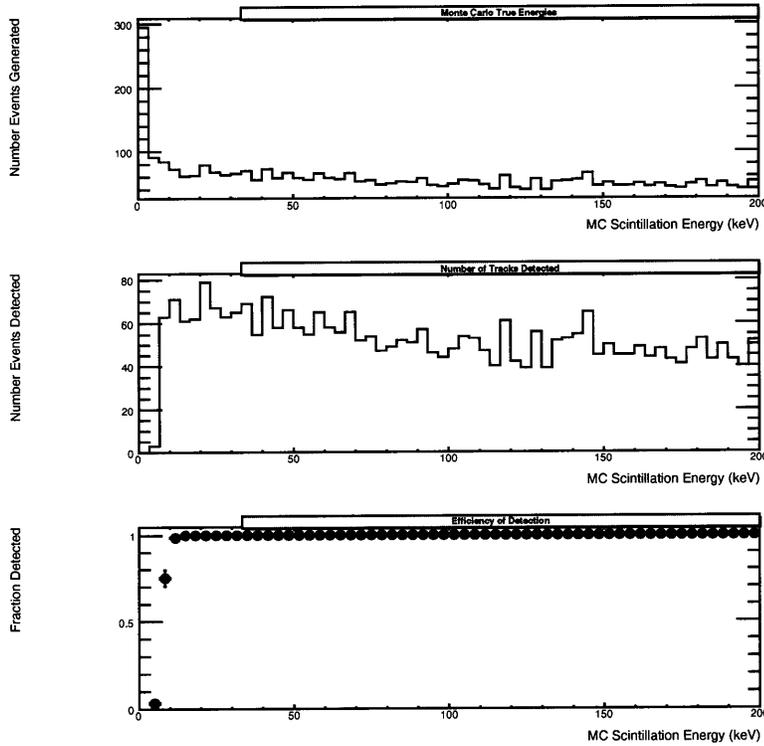
Figure 4-8: Same as Figure 4-7 but with 2.0 ADU/pixel read noise.

position of the ending of the Convolution line, and σ is the standard deviation found from the perpendicular fit, Equation 3.9.

4.3.1 ConvolutionFramework vs. AnalysisFramework

Figure 4-11 is a direct comparison of the performance of the two frameworks at full read noise, isotropic angle distributions and a 20 cm drift length. This figure and many of the plots to follow in this thesis, shows the performance of three quantities: head-tail, energy and length reconstruction effectiveness.

The first plot on head-tail reconstruction shows the fraction of events reconstructed correctly, i.e. $N_{correct}/N_{total}$, where $N_{correct}$ is the number of events in that energy bin reconstructed correctly and N_{total} is the total number of events in that energy bin. This is plotted against the scintillation energy, EscintMC. The second plot measures energy reconstruction ability, specifically it is a plot of $\frac{E_{recon}-E_{MC}}{E_{MC}}$, or
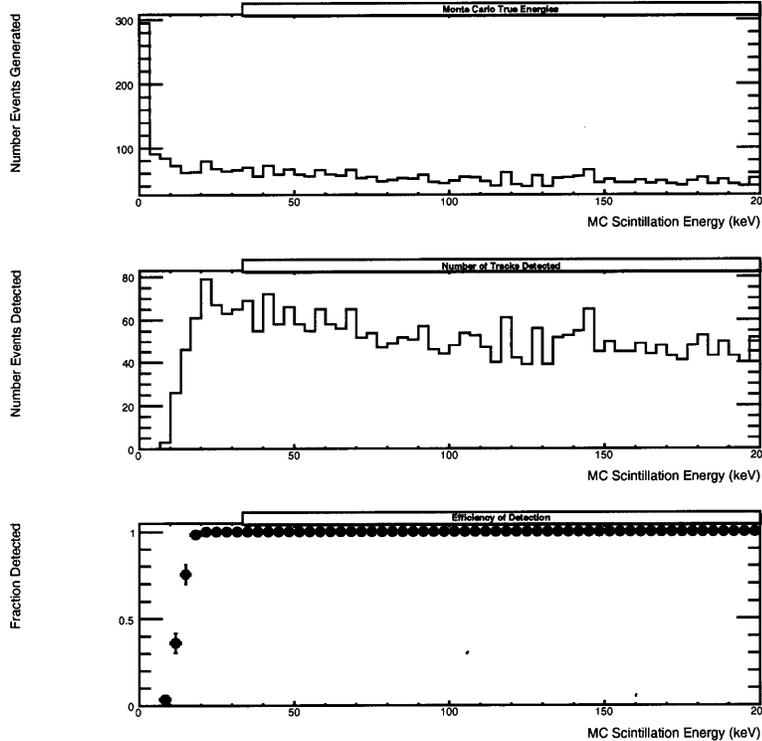
Figure 4-9: Same as Figure 4-7 but with 3.65 ADU/pixel read noise.

energy reconstruction bias versus EscintMC. Finally, the third plot measures the projected length reconstruction ability, specifically it is a plot of $\frac{L_{recon}-L_{MC}}{L_{MC}}$, or projected length reconstruction bias versus EscintMC.

For the given full read noise situation (7.3 ADU/pixel), we find that the reconstruction ability of head-tail for ConvolutionFramework surpasses that of Analysis-Framework for all energies studied here. Energy reconstruction is closely comparable because both algorithms follow a similar process of histogram integration. Reconstructed length for both algorithms is roughly similar until about 80 keV and then ConvolutionFramework begins to diverge and dramatically overestimate projected length of the track. This length bias of the ConvolutionFramework will be analyzed throughout this paper in order to understand its basis and the implications for the DMTPC.
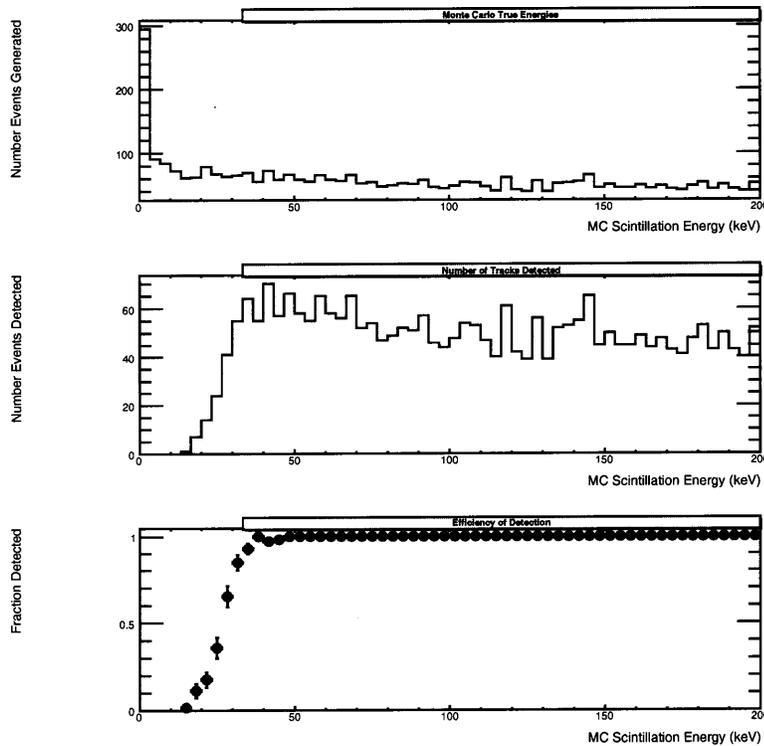
Figure 4-10: Same as Figure 4-7 but with 7.3 ADU/pixel read noise.

4.3.2 ConvolutionFramework Over Various Read Noises

Now, we probe into the dynamics and efficiency of ConvolutionFramework over different read noises. Figure 4-12 demonstrates this varying ability to reconstruct over various read noises.

Interestingly, we find that read noise has no statistically significant bearing on
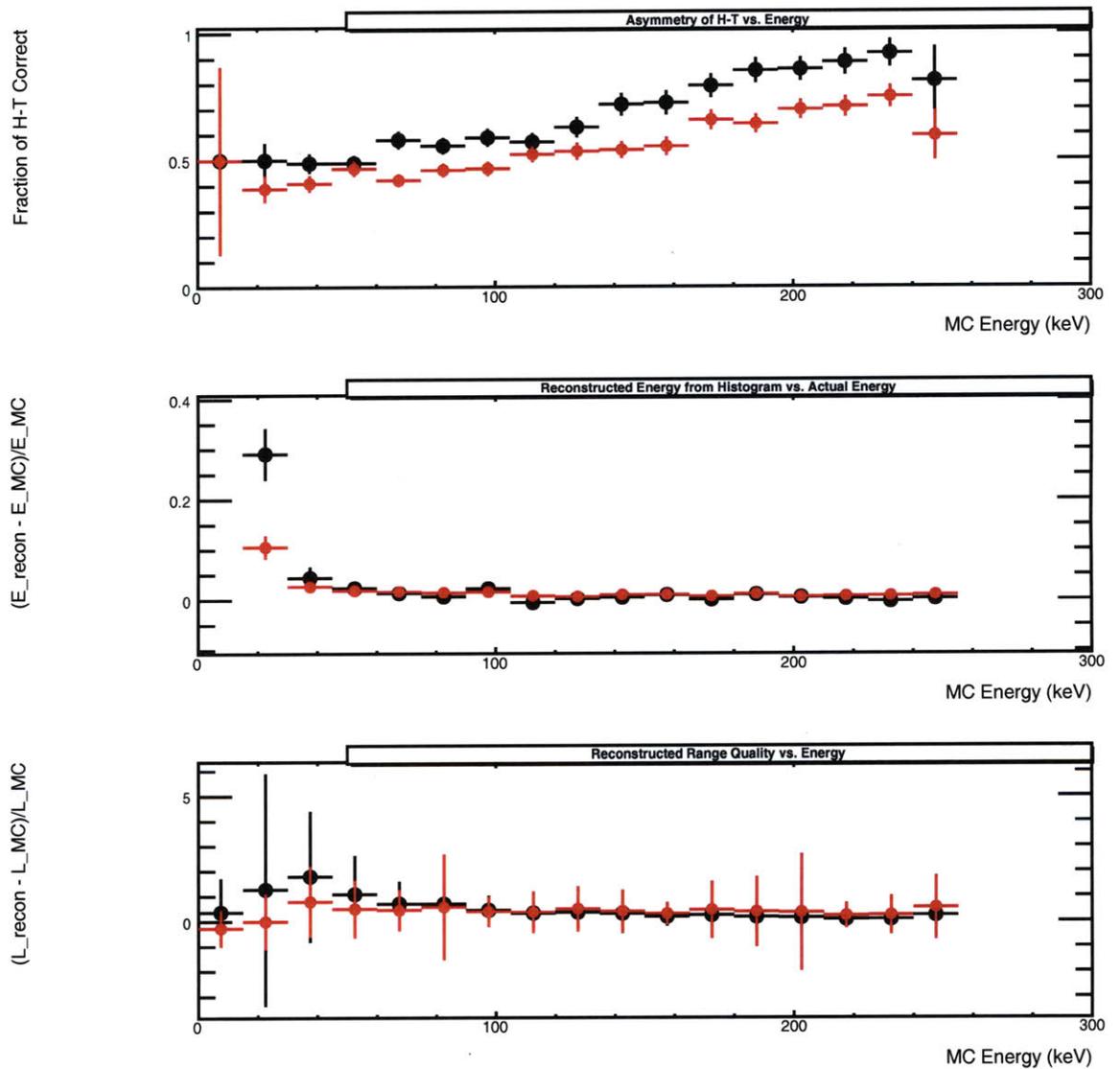
Figure 4-11: Comparison of ConvolutionFramework (Black) and AnalysisFramework (Red) for 12,000 event Monte Carlo at 7.3 ADU/pixel read noise, isotropic theta distribution, and a 20 cm drift length.

head-tail reconstruction ability over the energy range [20, 250 keV]. All points within a bin are found to be approximately within 1.5σ, where σ is the standard deviation of a uniform distribution. This result implies that the read noise is not the dominating factor in Equation 2.1. This implies that we are in the high signal limit, where $N_R{}^2 << N_*$ and Equation 2.1 then becomes the high signal regime, Equation 2.2. In
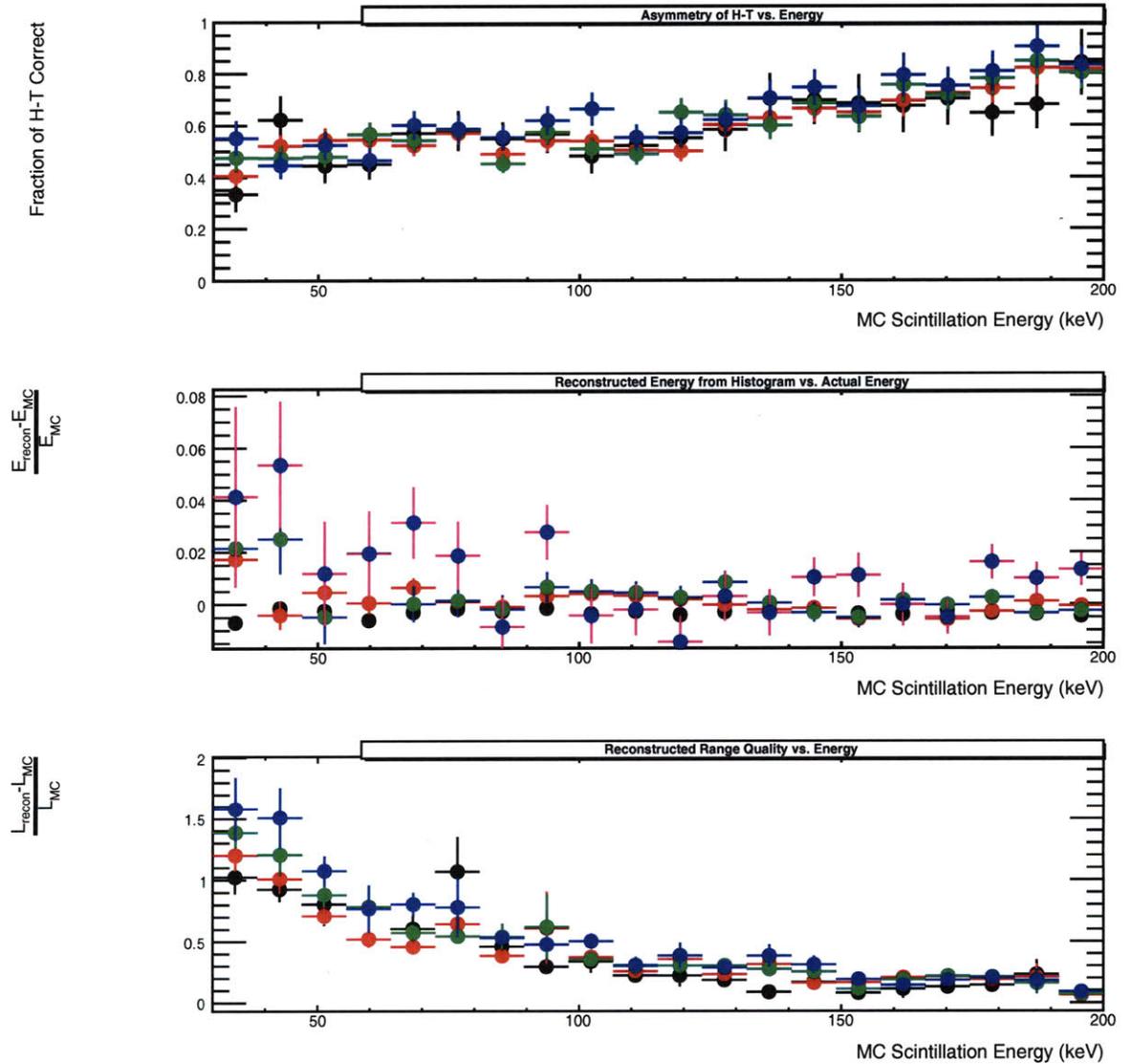
Figure 4-12: Comparison of ConvolutionFramework over different read noises. The black time series is the 0.5 ADU/pixel read noise simulation, the red time series is the 2.0 ADU/pixel read noise simulation, the green time series is the 3.65 ADU/pixel read noise simulation, and the blue time series is the 7.3 ADU/pixel read noise simulation.

Figure 4-13 we verify that indeed noise is primarily due to Poisson statistics of signal photon count.
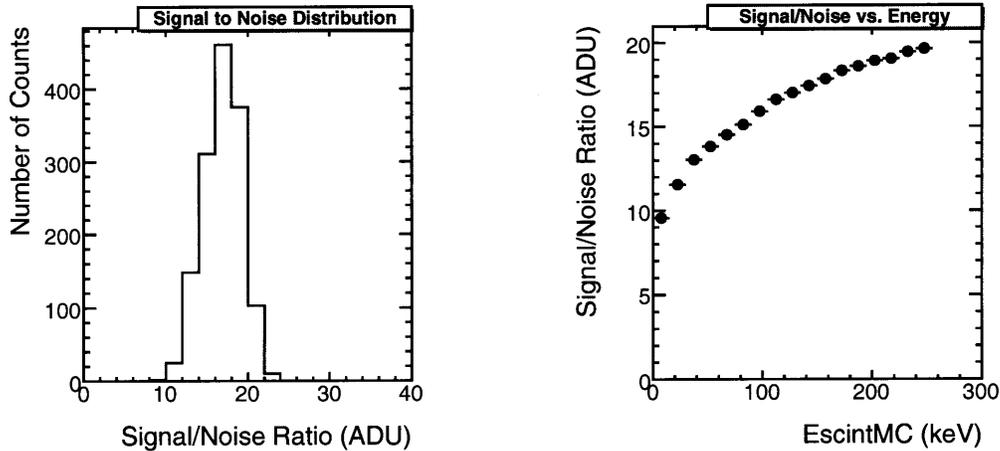
Figure 4-13: (Left) Demonstration of the Signal/Noise (S/N) per pixel distribution with peak near 16 ADU/pixel. (Right) Scaling of the Signal/Noise per pixel ratio with EscintMC. Note that only on the smallest energy scales, EscintMC $\in [0, 20$ keV$]$ is the full read noise of 7.3 ADU/pixel $\approx S/N$. For all higher energy ranges, $S/N >$ 7.3 ADU/pixel.

Both the energy and length reconstruction follow a similar trend and do not show any significant differences over various read noises. Again, this supports the claim that we are operating within the high signal regime where read noise is not the most pertinent source of noise. However, one important feature is that the length reconstruction appears to be heavily biased at low energies. This consistent over estimate of length may not be explained by read noise since this feature is common across all read noises.

4.4 Theta Dependence

Now because our detector currently has no three-dimensional angle resolution, we test the sensitivity of ConvolutionFramework over various angles. Each of the following subsections tests a certain set of nuclear recoil θ angles and the reconstruction ability. θ is defined off of the z-axis, therefore, $\theta = 0$ corresponds to a track moving vertically

upward, a $\theta = 90$ corresponds a nuclear track in the plane of the detector window and $\theta = 180$ corresponds to a nuclear track moving vertically downward. Because the DMTPC may only measure the projections of these nuclear recoils in the plane of the detector window, θ angles nearer to 90 degrees display the most information.

4.4.1 Theta $\in [0, 30] \cap [150, 180]$

First we test the steepest angle regime where $\theta \in [0, 30] \cap [150, 180]$ which is plotted in Figure 4-14.
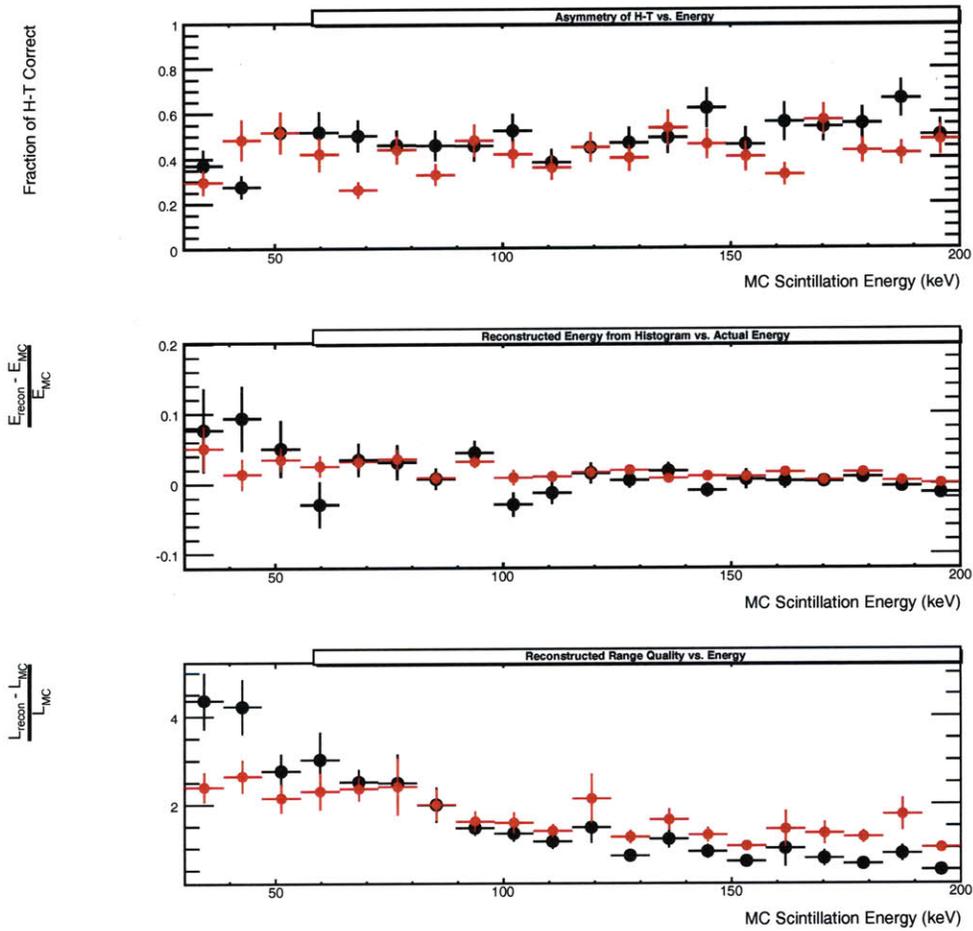
Figure 4-14: 19,000 Events run on $\theta \in [0,30] \cap [150, 180]$. ConvolutionFramework is plotted in black and AnalysisFramework is plotted in red.

Figure 4-14 shows that as expected, reconstruction of head-tail and length are

particularly difficult at these extremely steep angles. Similarly, energy as expected is reconstructed as well as before since the integration of the projection onto the plane of the detector window at any arbitrary angle does not effect the total signal measured. Therefore, for the remainder of these subsections, the analysis will be focused primarily on head-tail and length reconstruction. In this steep angle regime, the head-tail reconstruction is scarcely better than 0.5, i.e.random guesses for the signs of the slopes. However, ConvolutionFramework appears promising and actually consistently outperforms AnalysisFramework on EscintMC $\in$ [140, 250 keV]. One particularly troubling aspect is the large length biases introduced. In the low energies of [40, 70 keV], ConvolutionFramework dramatically underperforms AnalysisFramework
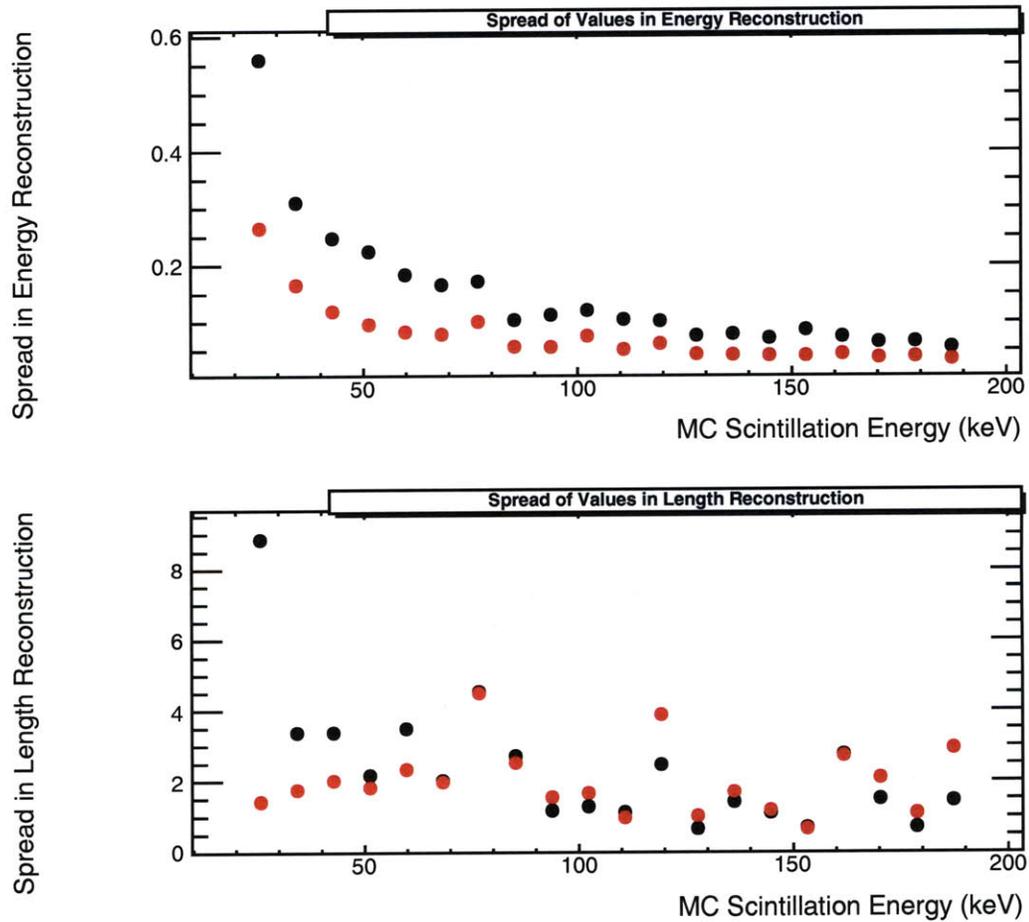
Figure 4-15: 19,000 Events run on $\theta \in$ [0,30] $\cap$ [150, 180]. ConvolutionFramework is plotted in black and AnalysisFramework is plotted in red.

54

and overestimates lengths by as much as 500%.

Figure 4-15 displays the spread of reconstructed values for energy bias and length bias for both algorithms. These points represent the variance in reconstruction performance around the mean reconstructed value. For each angle regime, this plot is displayed.

4.4.2 Theta $\in [30, 45] \cap [135, 150]$

Now, we explore shallower angles at the same read noise.
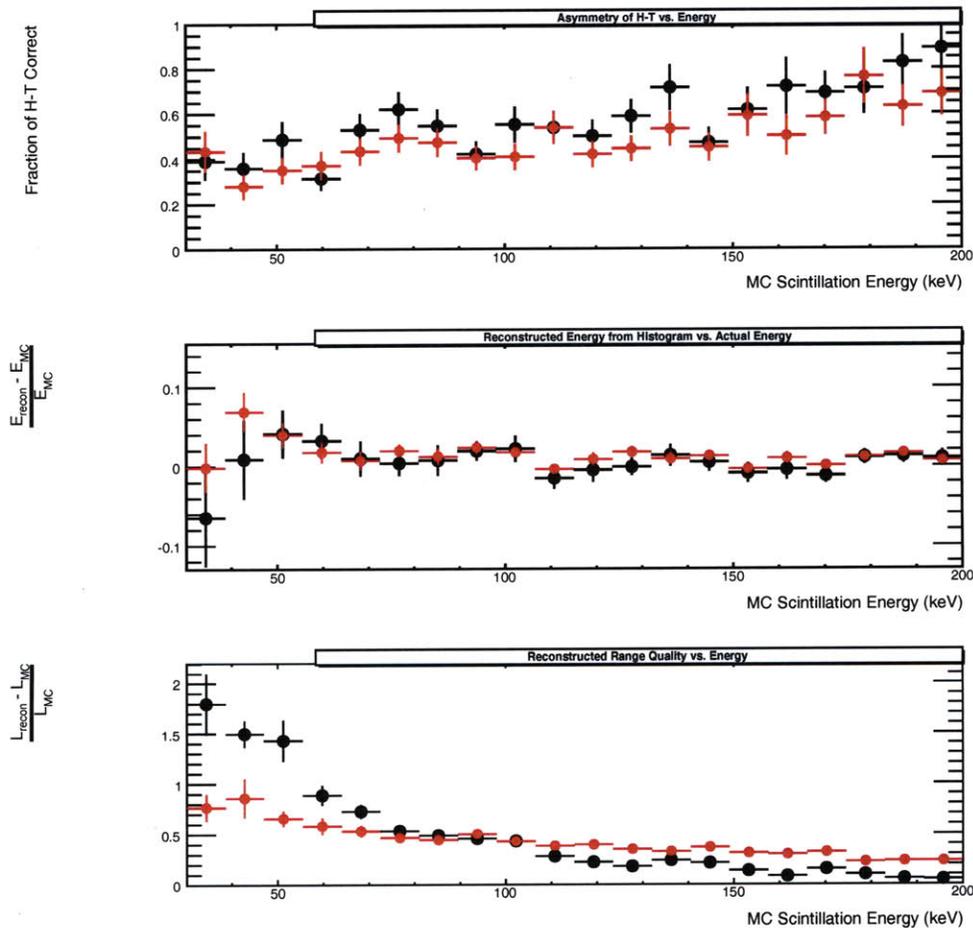
Figure 4-16: 19,000 Events run on $\theta \in [30,45] \cap [135, 150]$. ConvolutionFramework is plotted in black and AnalysisFramework is plotted in red.

We see in Figure 4-16 that by reducing the steepness of the angle by 15°, we have

55

dramatically improved head-tail reconstruction for both frameworks. Now, ConvolutionFramework reconstructs head-tail more effectively than AnalysisFramework on the same range energy as before, Escint $\in$ [140, 250 keV]. However, greatly improved is the large length bias in ConvolutionFramework. Now on the energy range, [40, 70 keV], the bias on length is decreased to near 190%. AnalysisFramework improves by roughly the same percent and continues to outperform ConvolutionFramework on this low energy range. Further, Figure 4-17 shows that the spread in the energy reconstruction values for ConvolutionFramework remains a factor of $\sim$ 2 higher than that for AnalysisFramework. Spread between the two algorithms is approximately equal in length reconstruction.
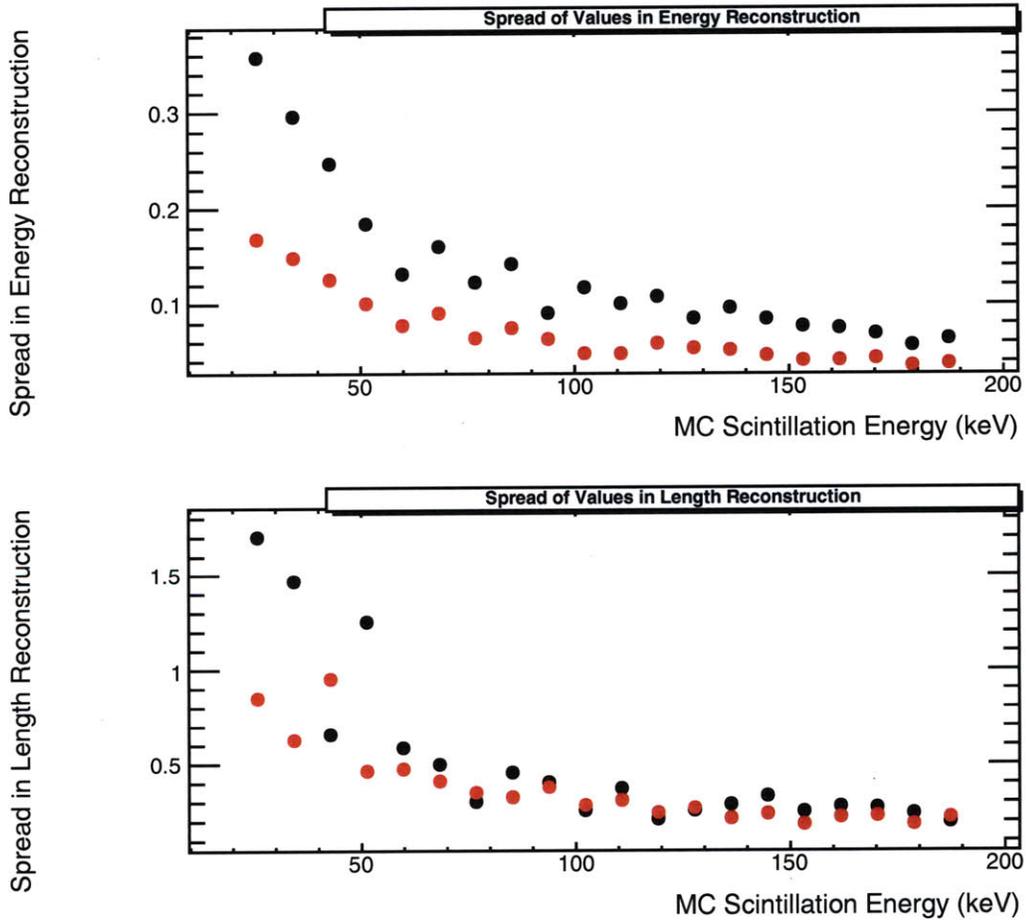
Figure 4-17: 19,000 Events run on $\theta \in$ [30,45] $\cap$ [135, 150]. ConvolutionFramework is plotted in black and AnalysisFramework is plotted in red.

4.4.3 Theta $\in [45, 60] \cap [120, 135]$

We continue to move to shallower angles and continue the analysis on head-tail and length reconstruction. We see in Figure 4-18 that within this regime of angles, ConvolutionFramework begins dominate AnalysisFramework, especially at the higher energies. Furthermore, both ConvolutionFramework's length bias has decreased again, ConvolutionFramework to approximately 100% and AnalysisFramework to approximately 40% on [40, 70 keV]. Figure 4-19 sshows that the spread for ConvolutionFramework at these shallower angles is now higher for both energy and length reconstruction.
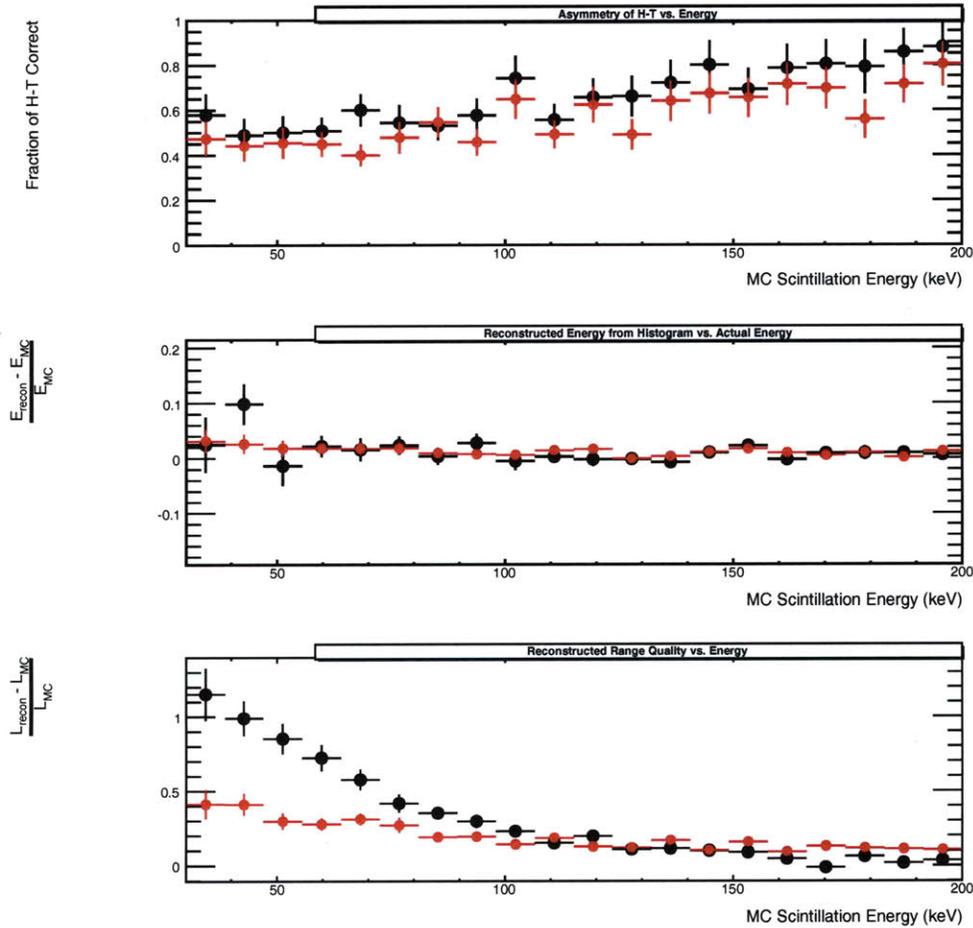
Figure 4-18: 19,000 Events run on $\theta \in [45,60] \cap [120, 135]$. ConvolutionFramework is plotted in black and AnalysisFramework is plotted in red.

4.4.4 Theta $\in [60, 75] \cap [105, 120]$

Continuing, now with all angles within 30° of the detector window plane we see an interesting stasis emerge.

Figure 4-20 shows that ConvolutionFramework does again improve head-tail reconstruction, soundly outdistancing AnalysisFramework, especially at higher energy events. However, length bias has improved by less than 5% while AnalysisFramework improved again significantly to a 20% length bias. This lack of change for ConvolutionFramework indicates that this length bias does not entirely emerge from steep angle tracks throwing the distribution off. Furthermore, we may infer that the bias
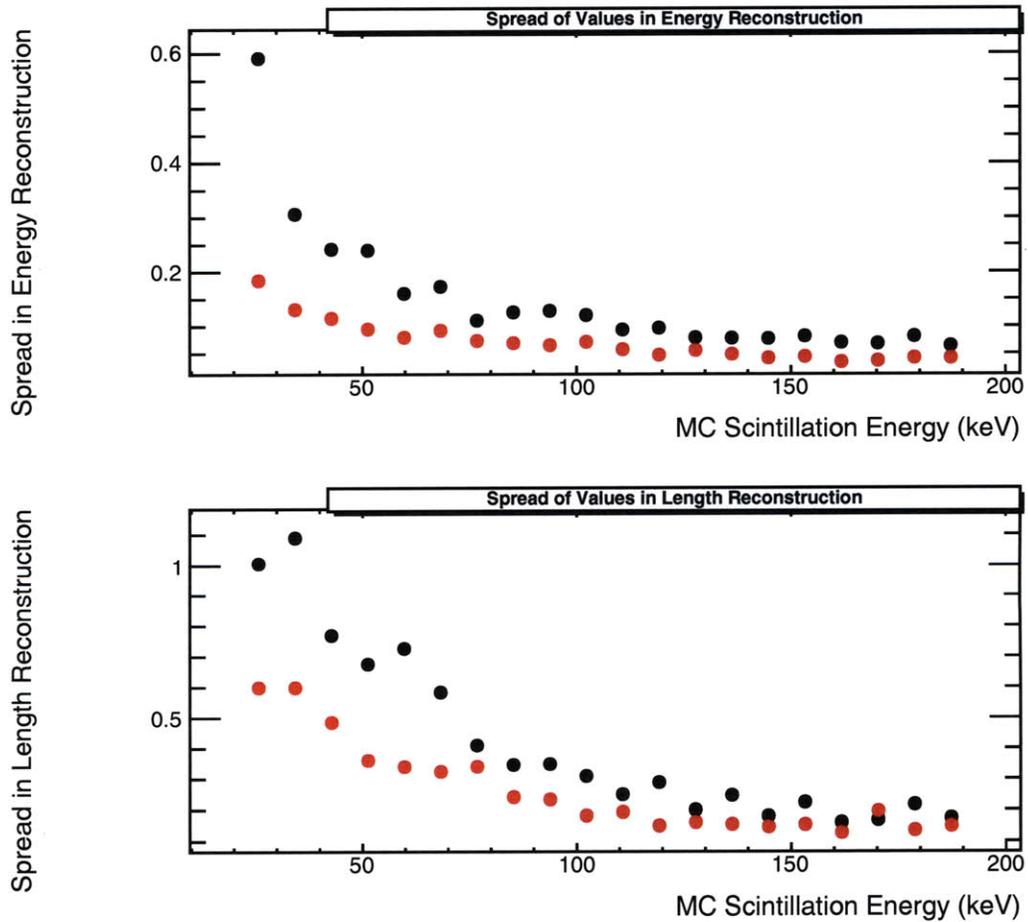
Figure 4-19: 19,000 Events run on $\theta \in [45,60] \cap [120, 135]$. ConvolutionFramework is plotted in black and AnalysisFramework is plotted in red.

that is effecting ConvolutionFramework is not the same one that is effecting Analy-sisFramework since AnalysisFramework's length reconstruction continues to improve over this new regime of shallower angles.

4.4.5   Theta $\in [75, 105]$

Finally, we constrain all θ values to be within 15° of the plane of the detector window.

In Figure 4-22 we see little improvement for either ConvolutionFramework or for AnalysisFramework on both head-tail reconstruction and on length reconstruction. Therefore, the θ values on [60, 120] are well handled by either algorithm, but once
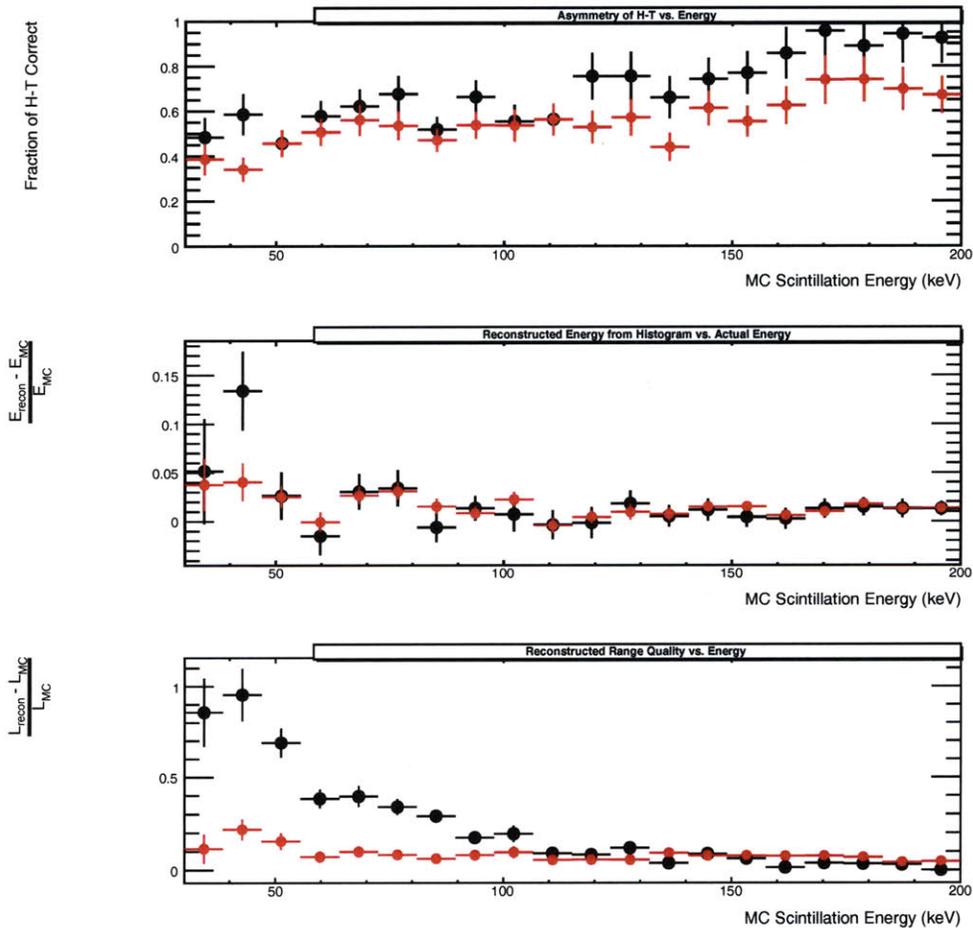
Figure 4-20: 19,000 Events run on $\theta \in [60,75] \cap [105, 120]$. ConvolutionFramework is plotted in black and AnalysisFramework is plotted in red.

θ extends beyond this window, reconstruction ability falls sharply. Also, in Figure 4-23 we note that the spread of reconstructed values from ConvolutionFramework is higher, again indicating a larger variance in reconstruction values.

4.4.6 Convolution Framework Over Various Angles

Now overlaying the plots of all the preceding angle regimes to better visualize how reconstruction ability of ConvolutionFramework varies,

we see more clearly see the rapid improvements in head-tail and length reconstruction by moving towards shallower angles. Furthermore, the remaining length
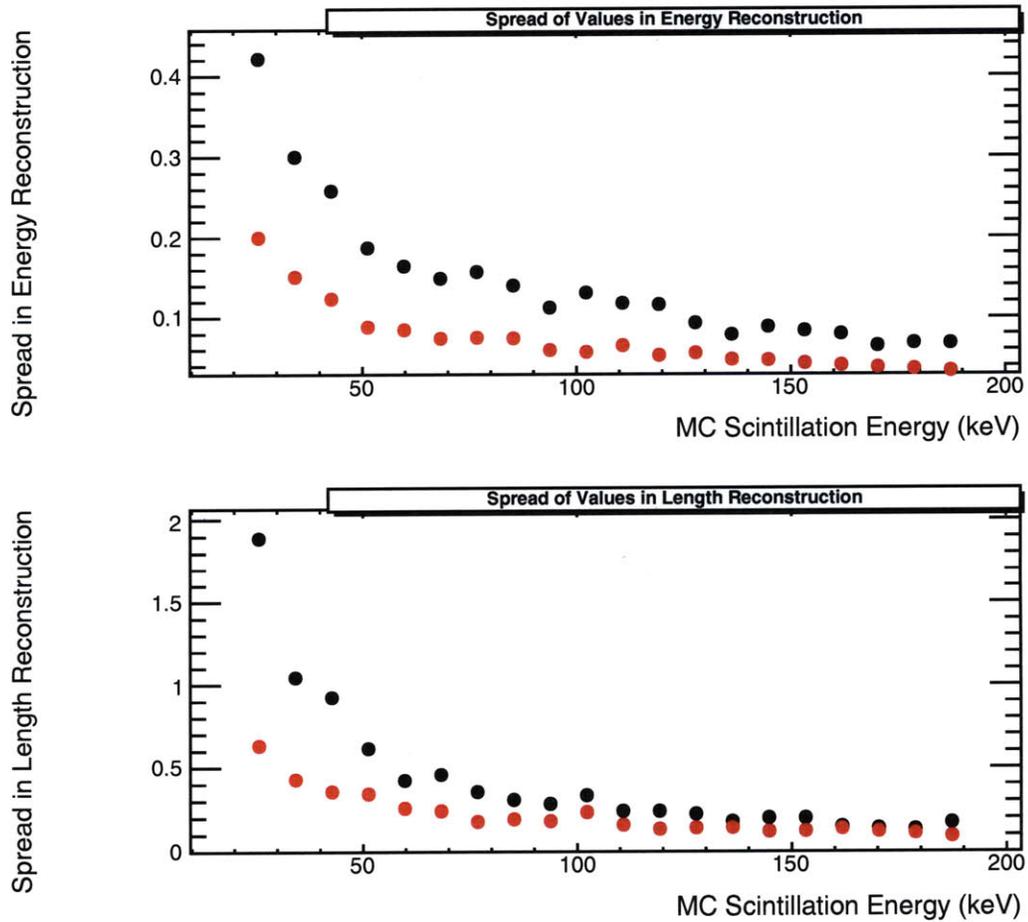
Figure 4-21: 19,000 Events run on $\theta \in [60,75] \cap [105, 120]$. ConvolutionFramework is plotted in black and AnalysisFramework is plotted in red.

bias over the shallowest angles remains particularly clear and necessitates further explanation. So therefore, although the steepest θ values do detract sharply from the ConvolutionFramework's reconstruction ability, there remains another bias latent within the algorithm. We therefore, analyze one of the last remaining complications in fitting, the resolution of the camera.

As mentioned above in the Complicating Factors section, the lowest energy tracks
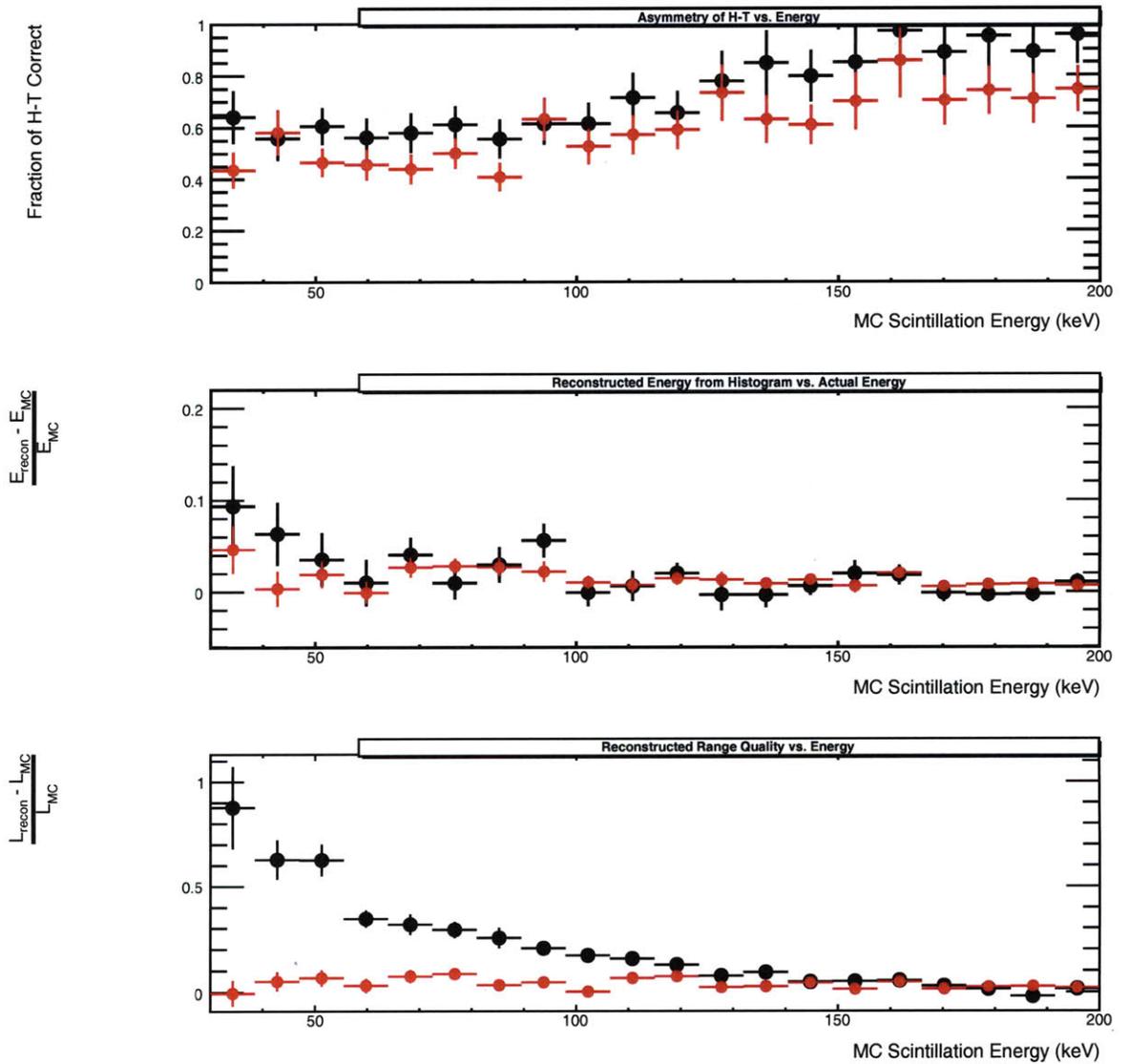
Figure 4-22: 19,000 Events run on $\theta \in [75, 105]$. ConvolutionFramework is plotted in black and AnalysisFramework is plotted in red.

will only extend across a few bins. This makes a fitting algorithm difficult to implement and essentially meaningless at the lowest energy limits (i.e. fitting the convolution of a line and a Gaussian to 1-2 pixel length tracks). We therefore consider the distributions of the true and reconstructed lengths.

Note here that the distribution of projected L_{MC} has a greater proportion of events nearer and beneath the one and two superpixel boundaries because the true values
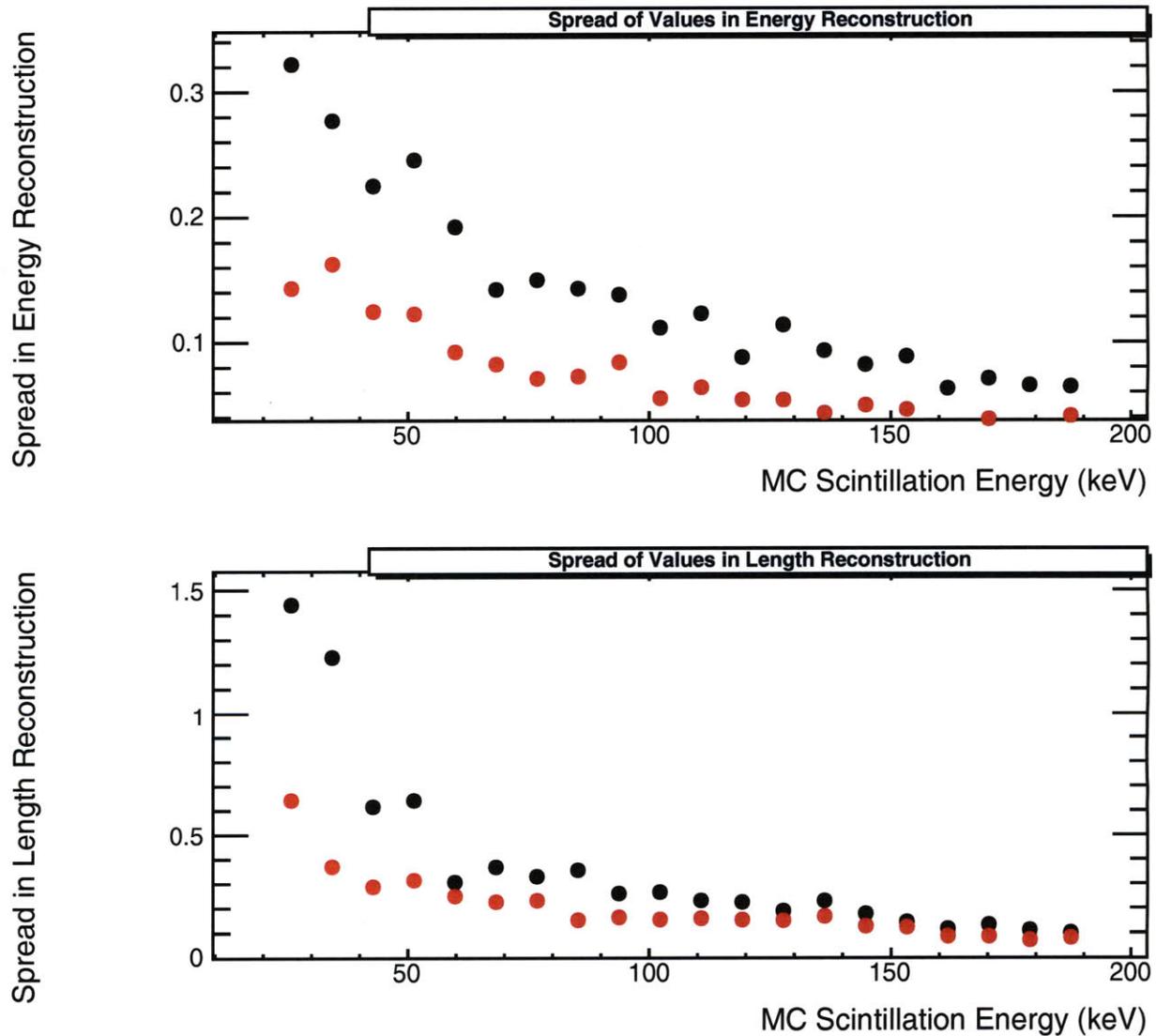
Figure 4-23: 19,000 Events run on $\theta \in [75, 105]$. ConvolutionFramework is plotted in black and AnalysisFramework is plotted in red.

have no imposed resolution limitations and may be made generated to be arbitrarily small. However, these boundaries represent the difficulty for reconstructing data points near the resolution threshold. Since a great deal of information of the functional form of $\frac{dE}{dx}$ is within one or two pixels, a functional fitting algorithm is ineffective in this regime. Resolution as a limiting factor partially explains the remaining length reconstruction bias of ConvolutionFramework.
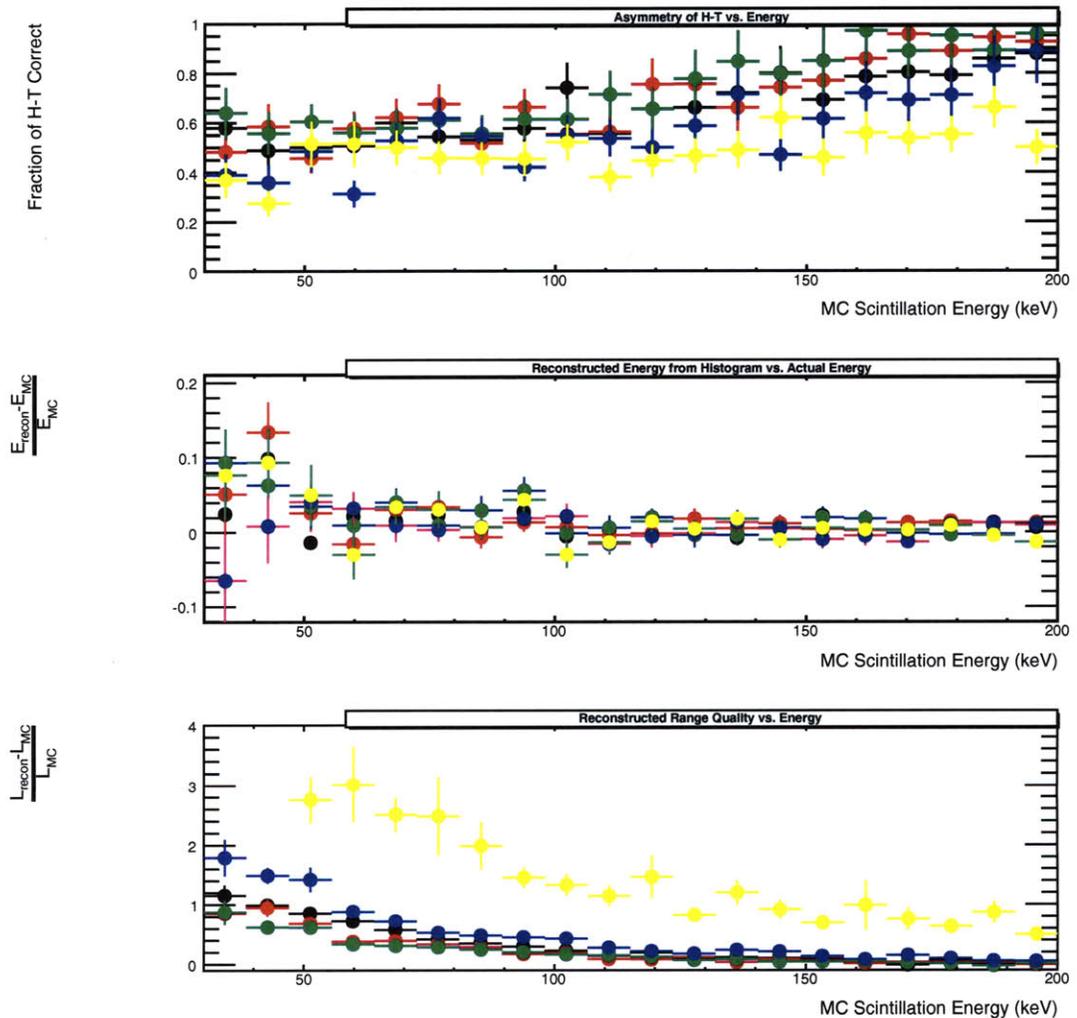
Figure 4-24: 19,000 Events on five different sets of angles. Yellow corresponds to θ angles on range [0, 30] ∩ [150, 180]. Blue corresponds to θ angles on range [30, 45] ∩ [135, 150]. Black corresponds to θ angles on range [45, 60] ∩ [120, 135]. Red corresponds to θ angles on range [60, 75] ∩ [105, 120]. And Green corresponds to θ angles on range [75, 105].

63
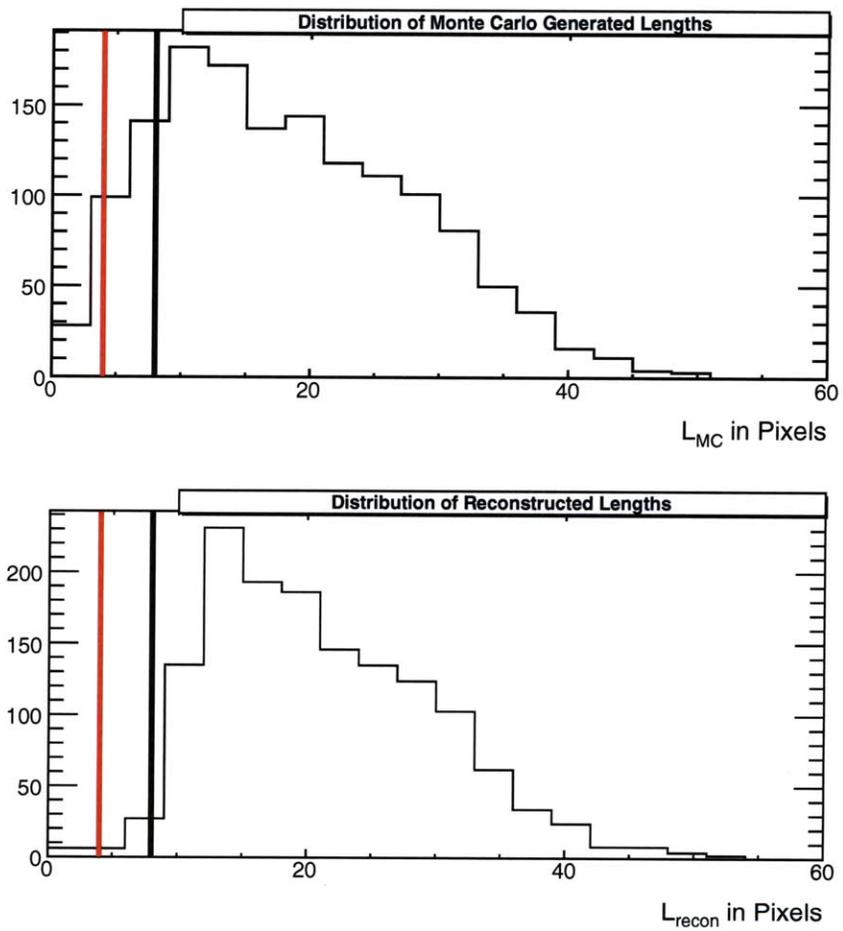
Figure 4-25: Distribution of projected L_{MC} and of L_{recon}. Red line represents the width of a single superpixel, the finest resolution the analysis is conducted upon. The black line represents the width of two superpixels.

Chapter 5

Conclusions

The purpose of this research was to develop and evaluate a physically motivated fitting function to be used in the reconstruction of nuclear recoils and to optimize the extraction of track parameters. The new algorithm, ConvolutionFramework, was shown to be a promising approach to reconstruction for the DMTPC group. Certain biases and uncertainties remain, however, and must be adressed before being incorporated into the standard analysis procedure.

One of the primary points demonstrated was that lowering the camera's read noise, below the 7.3 ADU/pixel current level, did not offer a significant improvement in reconstruction ability of head-tail, energy or length of the nuclear recoil when the camera is read out in 4x4 pixels (see Figure 4-12). This result was expected because our detector operates in the high signal regime as shown in Figure 4-13. In this regime, the Poisson noise from the photon signal represents a larger contribution to noise in the image than does the read noise.

Additionally, the newly introduced ConvolutionFramework algorithm outperforms the current AnalysisFramework algorithm in head-tail reconstruction over all energies tested. As expected, reconstruction of head-tail for both algorithms was found to be extremely poor at steep out-of-plane angles, θ. For θ within 30° of the x-y plane (i.e. the detector window plane), head-tail reconstruction was $\approx 60\%$ correct at $E = 80$ keV using ConvolutionFramework. However, increasing the steepness of θ by 30° completely destroyed head-tail reconstruction ability at $E = 80$ keV. This in-

dicates that 3-D angle reconstruction will be imperative for DMTPC moving forward. The current analysis of head-tail reconstruction is based on a 2-D projection of the nuclear recoil and is significantly negatively impacted for steep θ values because the skewness and profile of the ionization energy deposition, $\frac{dE}{dx}$ is lost in the projection of these steep angle nuclear recoils. By using full 3-D information of the nuclear recoil, DMTPC would be able to determine head-tail far more effectively in these difficult steep angle regimes.

Nuclear recoil energy reconstruction for ConvolutionFramework performed slightly worse than AnalysisFramework. Both are built upon the same underlying track recognition program and both rely on an integration of bins, however, ConvolutionFramework is more sensitive to poor fits and thus has a higher variance. Energy reconstruction ability for both algorithms is effective and at 50 keV, ConvolutionFramework estimates energy within 25% percent accuracy and AnalysisFramework is within 15% accurate. As expected, θ does not influence energy reconstruction ability since the projection of the total energy deposited onto the x-y plane makes the initial θ angle of the nuclear recoil irrelevant.

Nuclear recoil length reconstruction presented many interesting biases and features. First, the commonality between both algorithms was that at energies above 120 keV, both algorithms proved to be very effective in determining the length of the nuclear recoil track. At energies below 100 keV, though, both algorithms consistently overestimate the length of the track. This bias is partly a result of the difficulty in reconstructing length for steep θ values, however, the bias still exists even after constraining tracks to be ±30° within the x-y plane, indicating another source of bias which is believed to be a result of the limited camera resolution.

Both algorithms also display differences in nuclear recoil length reconstruction. ConvolutionFramework outperforms AnalysisFramework for energies above $\approx$ 120 keV, particularly at the steepest θ angles. This is a result of ConvolutionFramework explicitly identifying the contribution of transverse diffusion and factoring it out of the length determination. This ability to reconstruct lengths more accurately by factoring out transverse diffusion shows promise, however, this is not a robust technique

across all energies. As the nuclear recoil energy is decreased, ConvolutionFramework significantly underperforms AnalysisFramework and actually produces length estimates far larger than AnalysisFramework. Although this effect is partially a result of steep θ value nuclear recoils, we believe it also to be related to the limited camera resolution. ConvolutionFramework relies on a more sophisticated functional form to fit with and at the lowest energies (< 50 keV), nuclear recoils may only extend 1-3 pixels and therefore proper fitting is virtually impossible. This result indicates that resolution will be a significant limiting factor in low energy length reconstruction ability and should be a focus on improving moving forward.

5.1 Further Work

As with all research, this work presents many unexplored paths and also requires the improvement of certain procedures. One area of further investigation is the low energy length reconstruction bias of the ConvolutionFramework algorithm. As detailed above, a probable source of the bias is the limited resolution in our camera (0.143 x 0.143 mm^2 of the amplification region is imaged per camera pixel which are then binned 4 x 4). Tests of reconstruction ability as a function of camera pixel resolution could be generated using our current Monte Carlo Simulation and could provide a clear signal about whether this is in fact the primary limiting factor.

One procedure that necessitates reformulation is the determination of the track region around the 2-D projected nuclear recoil in the image. As described earlier, this is currently set to be the region within 5 pixels of the nearest point of the nuclear recoil track, where the nuclear recoil track is designated to be > 5 contiguous pixels above a certain threshold. However, this implicitly utilizes these 5 "border" pixels to determine the surrounding level of image bias or the pedestal. But these 5 border pixels incorporate the signal as well and thus is not an entirely accurate measure of the surrounding level pedestal. To reformulate this procedure, instead of taking the 5 pixels directly adjacent to the identified track, the algorithm will instead select 5 pixels which are a distance of 5 away from the track. This method allows for a more

careful sampling of the background since the signal from the nuclear recoil will be greatly diminished at these distances.

And finally, another potentially interesting topic to explore is the implemention of a Bayesian framework in determining head-tail for a given nuclear recoil track. Currently, the algorithm uses the ROOT numerical minimization program, MINUIT, to determine the best fit functions, Equations 3.9 and 3.13, for the perpendicular $(H_\perp)$ and parallel $(H_\parallel)$ projections respectively. Therefore, for the parallel fitting procedure, MINUIT selects a single set of parameters from all the physically possible parameter space and directly determines head-tail from the sign of the slope of the convolved line m. However, in determining head-tail, we are not concerned with the absolute value of m, only the sign of m. So an alternative approach is to instead create a physically motivated likelihood function of the nuclear recoil distribution and then partition the parameter space into two sets, one set where $m < 0$ and the other set where $m > 0$. With the parameter space partitioned, we may then perform a Bayesian model selection. This allows for higher sensitivity to the sign of the slope because this approach incorporates more information about the parameter space to determine the head-tail. Should this prove succesful, this Bayesian model selection could be extended to determination of length as well.

Bibliography

[1] K. Freese A. K. Drukier and D.N. Spergel. Detecting cold dark-matter candidates. *Physical Review D*, 33(3495), 1986.

[2] DMTPC and T. Caldwell. First limit from a surface run of a 10 liter dark matter time projection chamber, May 2009. Undergraduate Thesis.

[3] DMTPC and A. Kaboth et al. A measurement of photon production in electron avalanches in CF4. *Nuclear Instruments and Methods in Physics Research A*, 592(63), 2008.

[4] DMTPC and D. Dujmic et al. Charge amplification concepts for direction-sensitive dark matter detectors. *Astroparticle Physics*, 30(58), 2008.

[5] DMTPC and D. Dujmic et al. Observation of the 'head-tail' effect in nuclear recoils of low-energy neutrons. *Nuclear Instruments and Methods in Physics Research A*, 584(327), 2008.

[6] DMTPC and G. Sciolla et al. The DMTPC Detector. *arXiv*, 0811.2922, 2008.

[7] DMTPC and G. Sciolla et al. The DMTPC project. *Journal of Physics: Conference Series*, 179(012009), 2009.

[8] DMTPC and T. Caldwell et al. Transport properties of electrons in CF4. *arXiv*, 0905.2549, 2009.

[9] C. Savage et al. Compatibility of DAMA/LIBRA dark matter detection with other searches. *Journal of Cosmology and Astroparticle Physics*, 0904(010), 2009.

[10] G. Battaglia et al. The radial velocity dispersion profile of the galactic halo: Constraining the density profile of the dark halo of the milky way. *Royal Astronomical Society*, 364(433), 2005.

[11] S. Ahlen et al. The case for a directional dark matter detector and the status of current experimental efforts. *International Journal of Modern Physics A*, 25(1), 2010.

[12] A. Green and B. Morgan. Optimizing WIMP directional detectors. *Astroparticle Physics*, 27(142), 2006.

[13] C. Grupen and B. Shwartz. *Particle Detectors*. Cambridge University Press, 2008.

[14] S.B. Howell. *Handbook of CCD Astronomy*. Cambridge University Press, 2000.

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

[16] C. H. Lineweaver and D. Barbosa. Cosmic microwave background observations: implications for Hubble's constant and the spectral parameters n and q in cold dark matter critical density universes. *Astronomy and Astrophysics*, (9612146v2), 1997.

[17] B. Morgan, A. Green, and N.J.C. Spooner. Directional statistics for WIMP direct detection. *Physical Review D*, 71(103507), 2005.

[18] J.D. Vergados and A. Faessler. Direct WIMP detection in directional experiments. *Physical Review D*, 75(055007), 2007.

[19] F. Zwicky. On the masses of nebulae and of clusters of nebulae. *Astrophysical Journal*, 86(217), 1937.