

UNIVERSITY OF CALIFORNIA,
IRVINE

In Pursuit of the Rare Decay $\mu^+ \rightarrow e^+\gamma$:
A Window to the UV Completion of the Standard Model

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Physics

by

Benjamin Alexander Golden

Dissertation Committee:
Professor William Molzon, Chair
Professor Mu-Chun Chen
Professor Jonas Schultz

2012

TABLE OF CONTENTS

	Page
LIST OF FIGURES	vi
LIST OF TABLES	xii
ACKNOWLEDGMENTS	xiii
CURRICULUM VITAE	xv
ABSTRACT OF THE DISSERTATION	xvii
1 Introduction	1
2 Theoretical Background	3
2.1 Standard Model Particle Physics	3
2.1.1 Electroweak Theory	4
2.1.2 The Higgs Emerges	8
2.1.3 Limits of the Standard Model	13
2.2 The Case for Supersymmetry	14
2.3 Lepton Flavor Violation	22
2.3.1 LFV in the Standard Model	22
2.3.2 LFV Beyond the SM	25
3 Muon Physics	34
3.1 History	34
3.2 $\mu \rightarrow e\gamma$ detection	37
3.2.1 Signal Event Signature	38
3.2.2 Radiative Decay Background	38
3.2.3 Accidental Background	39
4 MEG Hardware	43
4.1 Beam	44
4.1.1 Proton Accelerator Complex	44
4.1.2 $\pi E5$	45
4.1.3 Beam Transport System	48
4.2 Coordinate System	48
4.3 Stopping Target	49

4.4	Magnetic Spectrometer	50
4.4.1	COBRA Magnet	51
4.4.2	Drift Chamber	57
4.4.3	Timing Counter	61
4.5	Liquid Xenon Calorimeter	65
4.6	Trigger and Data Acquisition	71
4.6.1	Data Acquisition with MIDAS	72
4.6.2	Waveform Digitization with DRS	73
4.6.3	Trigger System	74
4.7	Calibration Hardware	76
4.7.1	Charge Exchange Apparatus	77
4.7.2	Cockcroft-Walton Accelerator	79
5	Event Reconstruction	82
5.1	XEC	82
5.1.1	Waveform Analysis	83
5.1.2	Photon Reconstruction	85
5.2	DCH	92
5.2.1	Hit Reconstruction	93
5.2.2	Cluster Reconstruction	101
5.2.3	Track Reconstruction	103
5.2.4	Track Fitting	107
5.3	TIC	110
5.3.1	Waveform Analysis	110
5.3.2	Hit Reconstruction	111
5.4	DCH-TIC Association	111
5.5	Formation of MEG Candidates	112
6	Calibration	114
6.1	DRS Calibration	114
6.1.1	Voltage Calibration	114
6.1.2	Time Calibration	115
6.2	DCH Calibrations	115
6.2.1	Waveform Time Calibration	115
6.2.2	Z Calibration	120
6.2.3	Chamber Alignment	132
6.3	XEC	140
6.3.1	PMT Gain	140
6.3.2	Quantum Efficiency	142
6.3.3	Timing Calibration	142
6.4	TIC	143
6.4.1	Z-coordinate calibration	143
6.4.2	Bar-to-bar time offsets	143
6.4.3	Time Walk	144

7	Event Selection	146
7.1	Photon Cuts	146
7.1.1	Fiducial Volume	146
7.1.2	Cosmic Ray Veto	147
7.1.3	Pileup Rejection	148
7.2	Positron Cuts	148
7.2.1	Fiducial Volume	148
7.2.2	Drift Chamber Hit Pattern	149
7.2.3	Quality of Track Fit	150
7.2.4	Stopped Muon Consistency	150
7.2.5	Projection to Timing Counter	151
7.2.6	Ghost Selection	151
8	Hardware and Software Performance	157
8.1	Detector Resolutions	157
8.1.1	XEC	157
8.1.2	DCH	165
8.1.3	TIC	173
8.2	Resolutions of Multi-detector Measurements	175
8.2.1	Relative Time Resolution	175
8.2.2	Relative Angle Resolutions	176
8.3	Detector Efficiencies	181
8.3.1	Photon Detection Efficiency	181
8.3.2	Positron Detection Efficiency	183
9	Maximum Likelihood Analysis	186
9.1	Blinding and Analysis Windows	186
9.2	Likelihood Analysis Framework	187
9.2.1	The Likelihood Function	188
9.2.2	Accidental Background PDFs	190
9.2.3	Signal Event PDFs	196
9.2.4	Radiative Decay Background PDFs	205
9.3	Background Estimates for the Signal Window	212
9.3.1	Accidental Background	212
9.3.2	Radiative Decay Background	212
9.4	Setting a Confidence Interval	213
9.5	Normalizing the Experiment	215
9.6	Inclusion of Systematic Uncertainties	218
9.6.1	Signal PDFs	218
9.6.2	Radiative Decay PDFs	220
9.6.3	Accidental Background PDFs	221
9.7	Blind Sensitivity Estimation	222
9.7.1	Fitting the Sidebands	222
9.7.2	Simulating Experiments	223
9.8	Results of the Fit to the Signal Window	224

10 Conclusion	229
Bibliography	231
Appendices	236
A Radiative Muon Decay Branching Ratio	237

LIST OF FIGURES

	Page
2.1 Active neutrino loop contribution to $\mu \rightarrow e\gamma$. The internal fermion line is a neutrino mass eigenstate. By diagonalizing the neutrino propagator in flavor space, off-diagonal flavor couplings, $U_{\mu i}$ and U_{ei} , are induced.	24
2.2 Flavor violation in the left-handed lepton sector brought about through quantum effects involving right-handed neutrinos.	28
2.3 $\mu \rightarrow e\gamma$ mediated by slepton mixing.	28
2.4 Predictions for $BR(\mu^+ \rightarrow e_R^+\gamma)$ at various values of $\tan\beta$ as a function of the left-handed selectron mass for $M_R = 10^{13}$ GeV, no trilinear coupling, and a 200 GeV wino mass.	29
2.5 Predictions for $BR(\mu \rightarrow e\gamma)$ are shown as a function of the universal gaugino mass for two cases of $\tan\beta$, scanning an LHC relevant space in the parameters describing the Planck scale masses. Both the PMNS case (green) and the CKM case (red) are explored.	31
2.6 Example diagram mediating $\mu N \rightarrow eN$ in supersymmetric models.	33
3.1 A plot of the progress made in improving the upper limit on $BR(\mu \rightarrow e\gamma)$ in the 20th century.	36
3.2 An illustration of the event signatures for signal and background. Detected particles are indicated in red, while other particles are indicated in blue. $\Theta_{e\gamma}$ refers to the opening angle between the positron and photon, E_e refers to the positron energy, E_γ refers to the photon energy, t_e refers to the positron emission time, and t_γ refers to the photon emission time.	37
3.3 The accidental background energy spectra. Each differential branching ratio is normalized to be unity at its maximum value.	41
4.1 An aerial view of the 590 MeV ring cyclotron at PSI.	45
4.2 A diagram of the $\pi E5$ beamline components.	46
4.3 Measured flux of beam content.	47
4.4 An illustration of the beam transport system.	48
4.5 A sketch of the experimental layout for MEG.	49
4.6 A photograph of the muon stopping target labeled with dimensions.	50
4.7 A view of the mounted target, looking upstream.	50
4.8 A comparison of uniform magnetic field attributes with those of a graded magnetic field.	52

4.9	The rate of Michel positrons at a given radius for two magnetic field configurations.	53
4.10	A photograph of the magnet.	54
4.11	A schematic of the magnet design.	55
4.12	The magnetic field strength within the spectrometer as a function of $ z $. . .	56
4.13	PMT gain versus magnetic field strength for different orientations.	57
4.14	Magnetic field contours throughout the detector.	58
4.15	A photograph of the 16 drift chamber modules mounted on the support structure.	59
4.16	A single drift chamber module.	62
4.17	A cross-section of a drift chamber module, looking down the wires. The length measurements are in units of mm.	62
4.18	A picture of the three-sided frame.	62
4.19	An illustration of the technique for measuring Z from pad signals.	63
4.20	An schematic of the pad orientations.	63
4.21	A map of drift lines (red) and isochrones (green) within a cell for a fixed magnetic field strength of $B_z=0.9$ T.	64
4.22	A photograph of a timing counter assembly before installation.	66
4.23	A photograph of the calorimeter before installation.	68
4.24	A diagram of the calorimeter labeling the six faces.	69
4.25	A photograph of the mounted LED and alpha sources. The orange circle indicates a LED and the yellow arrow indicates an alpha source wire.	70
4.26	A close up of an ^{241}Am source on a wire.	71
4.27	A visual representation of the data acquisition system.	73
4.28	A diagram representing the technique for digitizing waveforms.	74
4.29	A pictorial layout of the trigger system.	75
4.30	Kinematics of the $\pi^0 \rightarrow \gamma\gamma$ decay.	77
4.31	A photograph of the liquid hydrogen target used with the π^- beam.	78
4.32	The setup for the external NaI calorimeter.	79
4.33	A photograph of the Cockcroft-Walton accelerator.	80
4.34	Components of the target setup for calibration with the Cockcroft-Walton. . .	81
5.1	Various representations of a PMT waveform from data.	84
5.2	An example of a saturated PMT waveform. The time over threshold is pointed out by the black arrow.	85
5.3	The relationship between expected charge and time over threshold based on the template waveform.	85
5.4	The 5.5 MeV data point is from the in-situ α sources; the 4.4 MeV, 11.7 MeV, and 17.6 MeV data points come from $p + \text{Li}$ and $p + \text{B}$ data; the 55 MeV and 83 MeV data points come from charge exchange data, and the 129 MeV line is from the radiative π^- capture reaction. The red line is a fit to a line with a floating y-intercept. The blue line is a fit to a line with the y-intercept fixed to zero.	89
5.5	Distribution of $\hat{\chi}_{time}^2$ on data (dotted line) and signal MC without pileup (solid line).	91

5.6	The six waveforms associated with a hit. On the left are waveforms from upstream readouts, and on the right are waveforms from downstream readouts. The top two are from the anode, the middle two are from the hood pads, and the bottom two are from the cathode pads. The horizontal red lines are the calculated baselines, the vertical red lines are the leading edge time, and the region enclosed by the two dashed blue lines is the allowed window to search for hits.	94
5.7	Distribution of $t_{end0} - t_{end1}$ for hits with both wire ends having a charge of at least .3 mV*ns. The region $ t_{end0} - t_{end1} < 6$ ns is fit to a Gaussian.	95
5.8	The blue data points are direct measurements of the Z resolution while the red data points are the calculated contributions of electric noise to the resolution at different integration times. Since the calculations for the latter involve measuring pedestal fluctuations for over a thousand channels, they are performed for a limited set of integration times.	98
5.9	Distribution of the difference in anode and pad determined Z coordinates of hits. The region $ Z_{anode} - Z_{pad} < 1$ cm is fit to a Gaussian.	100
5.10	An example of the waveforms used in the timing counter reconstruction. The black line is a DRS output, which is fit to a template (red). The blue line is the NIM pulse output, which is fit to another template (green). The time delay of the NIM pulse with respect to the DRS waveform is due to an electronic delay in the double threshold discriminator.	110
6.1	Some plots illustrating features of the time calibration on a single DRS chip.	116
6.2	The drift time distribution of a particular anode channel fit to a 5th order polynomial.	118
6.3	A pictorial summary of improvement in anode time alignment by calibration.	118
6.4	A fit to the distribution of the difference in peak amplitude times at the two ends of the cathode pad for a particular cell.	119
6.5	A fit to the distribution of the difference in peak amplitude times as measured from the downstream anode and cathode for a particular cell.	120
6.6	A pictorial summary of improvement in pad time alignment by calibration.	121
6.7	A fit of Eq. 6.1 to the pad asymmetry versus anode Z for the hood of a particular wire.	122
6.8	A pictorial summary of improvement in anode parameters by calibration.	124
6.9	The same as figure 6.7 except that fitted track Z coordinates are used in place of anode Z coordinates. The advantage is that the fits are typically more stable.	126
6.10	A pictorial summary of improvement in relative pad gains by calibration.	127
6.11	Scatter plots of the difference in h using anode determined Z coordinates and fitted track Z coordinates.	128
6.12	A profile histogram of the charge ratio versus $ Z $. This is fit to a cubic polynomial (solid black), which is extrapolated to the end of the wire (dashed red) to correct for the relative cathode-to-hood gain.	130

6.13	A fit to the asymmetry ratio dependence on the charge ratio, using data from all functioning wires. Enforcing the constraint that $\frac{A^{cathode}}{A^{hood}} = 1$ when $\frac{Q^{cathode}}{Q^{hood}} = 1$ leaves one free parameter, the slope of a line passing through that point. The slope is ~ 0.24 units of change in the asymmetry ratio per unit of change in the charge ratio.	131
6.14	The same as figure 6.13 but after rescaling the hood asymmetry to correct for bowing. The constraint that $\frac{A^{cathode}}{A^{hood}} = 1$ when $\frac{Q^{cathode}}{Q^{hood}} = 1$ is again enforced. The slope is not statistically significant.	132
6.15	A diagram of the target marker locations.	133
6.16	Figures showing the location of chamber survey markers.	134
6.17	Some plots illustrating the features of the radial chamber alignment.	135
6.18	Radial corrections to be applied to each chamber after alignment.	136
6.19	Some plots illustrating the features of the Z chamber alignment.	137
6.20	The track angle is the angle in the X-Y plane made with respect to the normal to the central chamber plane as marked by θ in this diagram. It is positive when the track is headed toward positive R, zero when the track is normal to the central chamber plane, and negative otherwise.	138
6.21	Some plots illustrating the features of the rotational and transverse chamber alignment.	139
6.22	Plots illustrating the PMT gain calibration method.	141
6.23	Scatter plot of time offsets versus bar number.	144
7.1	The accepted regions of the calorimeter are shown in red in the u-v plane (left) and the u-w plane (right).	147
7.2	Distributions of the positron coordinates when backwards projected into the calorimeter. The regions between the red lines are retained by the cuts.	149
7.3	Distributions of various positron variables related to the hit pattern. The red lines mark the cuts.	154
7.4	Distributions of various positron variables related to track fit quality. The red lines mark the cuts.	155
7.5	Distributions of vertex locations normalized to elliptical cuts. The cuts reject everything above the red lines.	156
7.6	Distributions of variables related to DCH-TIC matching. The regions between the red lines are retained by the cuts.	156
8.1	Lead brick apparatus for measuring XEC position resolution.	159
8.2	Measured distributions of u, v positions from the lead brick data.	159
8.3	A representation of the PMT grouping scheme for intrinsic timing evaluation. Blue circles correspond to odd group PMTs while orange circles correspond to even group PMTs.	161
8.4	Distribution of the NaI time and XEC time difference for coincident photon pairs.	162
8.5	Distribution of reconstructed photon energies from a 55 MeV source. Shallow events, $w < 2$ cm, are excluded.	164
8.6	A diagram of the technique for measuring the intrinsic radial position resolution.	165

8.7	A Gaussian fit to the distribution of $R_{plane\ 0} - R_{plane\ 1}$ out to ± 1 mm.	166
8.8	A diagram of the technique for measuring the intrinsic Z position resolution.	167
8.9	A Gaussian fit to the distribution of $Z_{late} - Z_{early}$ out to ± 4 mm.	169
8.10	Fits for the average positron angle resolutions using double turn events. The fit parameters shown are for the double Gaussian that is convolved with itself and fit to the distribution.	170
8.11	Two methods for finding the average positron energy resolution.	172
8.12	Fits for the average positron vertex resolutions using double turn events. The fit parameters shown are for the double (or triple) Gaussian that is convolved with itself and fit to the distribution.	173
8.13	Timing counter resolution as a function of bar number.	174
8.14	Muon radiative decay timing peaks.	176
8.15	The correlations between angle error and vertex error.	178
8.16	The correlations affecting the $\phi_{e\gamma}$ PDF mean.	180
8.17	The average relative angle resolutions.	182
8.18	The red curve is the standard error function model for the acceptance, while the blue curve is an alternate Gaussian model. Both are normalized to unit area in the range of positron energy on the plot.	184
8.19	The conditional timing counter efficiency given a positron passing drift chamber-related selection cuts versus positron momentum for various selection criteria on the match with the timing counter.	185
9.1	Some maximum likelihood fits to the accidental $t_{e\gamma}$ distribution. The fit PDFs are normalized to the data sample, constraining their form.	192
9.2	Accidental angle distributions used to build the PDFs.	193
9.3	An example showing the technique for interpolating the accidental $\theta_{e\gamma}$ PDF between bins of u	194
9.4	The accidental E_e distribution for different bins of δE_e	195
9.5	The accidental E_γ distribution for different regions of conversion depth. A PDF is formed from the distribution in the first bin and used to fit the other bins as shown in blue.	197
9.6	Fitted values of the parameters for the double Gaussian E_e resolution for several bins in δE_e . Each data point is plotted at the average value of δE_e within the bin.	199
9.7	The variation of σ_{up} with u and w for fixed v : $v \in [0, 6.19)$ cm. The plot on the left shows the discrete number of bins used to measure σ_{up} . The surface on the right is constructed by a linear interpolation.	200
9.7	Fits to the difference in θ_e on double turn events in several bins of δE_e . The fit parameters shown are for the double Gaussian that is convolved with itself and fit to the distribution.	204
9.8	The distribution of $\delta E_e^{1st\ turn} - \delta E_e^{2nd\ turn}$. The region between the red lines is used to construct positron angle PDFs.	205
9.8	Fits to the difference in ϕ_e on double turn events in several bins of ϕ_e for fixed $\delta E_e \in [0, 0.258)$ MeV. The fit parameters shown are for the double Gaussian that is convolved with itself and fit to the distribution.	207

9.9	The variation of depth resolution with $ PMT U $ and w for fixed $ PMT V $: $ PMT V \in [1.033, 2.067)$ cm. The plot on the left shows the discrete number of bins used to measure σ_W . The surface on the right is constructed by a second order interpolation.	208
9.10	Simulation of $\phi_{e\gamma}$ vs. $\theta_{e\gamma}$ for the resolution functions specific to a certain event.	208
9.11	The one-dimensional true radiative decay distributions. Each plot is made by integrating out the other three variables. They are normalized to unity at their maximum value.	210
9.12	The one-dimensional radiative decay PDF distributions for a certain event. Each plot is made by integrating out the other three variables. They are normalized to unity at their maximum value. Note that the projected angle PDFs do not reach zero within the analysis window.	211
9.13	Simulated distribution of upper limits.	224
9.14	Results of the likelihood fit projected onto each observable by integrating out the others. The green curve is the signal PDF, the red is the radiative decay PDF, the purple is the accidental PDF, and the blue is the sum of those three. The area under each curve is normalized to its best fit result.	226
9.15	Plots of the best fit value of N_{Sig} as each energy or time PDF parameter is changed. The spacing between data points on the horizontal axis is $\frac{\sigma}{2}$, where σ is the quoted uncertainty in the parameter.	227
9.16	Plots of the best fit value of N_{Sig} as each angle PDF parameter is changed. The spacing between data points on the horizontal axis is $\frac{\sigma}{2}$, where σ is the quoted uncertainty in the parameter.	228

LIST OF TABLES

	Page
3.1 Current upper limits on exotic muon decay and interaction modes.	37
7.1 A table of inefficiencies for each cut. The individual positron cut inefficiencies are tabulated as the fraction of events rejected by making only the cut with respect to making no cuts. The combined losses are tabulated similarly for each group of cuts made simultaneously with respect to making no cuts. In tabulating the photon acceptance cut inefficiencies, the positron cuts, pileup cut, and cosmic ray cut are already applied.	153
9.1 A summary of the systematic uncertainties in the signal PDFs.	221
9.2 Results of the fictitious signal searches in the sidebands. The asymmetric errors on the best fit parameters are the 1.645σ uncertainties calculated by MINOS.	223
9.3 Median lower limits at various confidence levels obtained for different amounts of simulated signal.	225

ACKNOWLEDGMENTS

The process of earning a PhD in experimental particle physics is a lengthy, trying experience that is at times exciting and at times disheartening. It would be remiss of me not to point out the contributions of various collaborators to that experience. Some graduate students join a mature experiment and essentially work directly toward a thesis-worthy physics result beginning from day one without incurring the educational benefit of helping to bring a young experiment to fruition. I was lucky in gaining a great deal of experience developing reconstruction code and performing a host of calibrations before ever beginning to do a thesis analysis. Working on MEG taught me to remain productive in any atmosphere, regardless of the ups and downs that come my way.

From my home institute, I must pay homage to several important people. I would like to express my gratitude to my thesis advisor, William Molzon, for unhesitatingly taking me as a student. He provided me with a highly unique education in both physics and communicating with collaborators. I learned at least as much from his epigrams as I did from any physics textbook, not the least of which is that, despite my preconceptions about particle physics, “Look, it’s not rocket science.”

Vladimir Tumakov laid the groundwork for the drift chamber software structure and algorithms to which I was later heir.

I owe Feng Xiao a particularly sincere acknowledgment. We worked together in developing drift chamber reconstruction algorithms and calibrations. More importantly, we shared the vicissitudes thrust upon us, like two lone soldiers hunted by a thousand adversaries.

Gordon Lim joined MEG just in time to help me focus on this thesis by absorbing many of my responsibilities. He was a refreshingly pleasant addition to the UCI group.

I gratefully acknowledge support from the U.S. Department of Energy under grant DEFG03-91ER40679.

It was a pleasure to work with my collaborators from the University of Tokyo. I shared many delightful conversations with Toshiyuki Iwamoto, who was always willing to answer my questions about the calorimeter and liberated me from several midnight journeys through the dark forest that separated PSI and my living quarters by giving me rides home at the end of our night shifts.

Wataru Ootani convened physics analysis meetings with a certain nonpartisan calm, which helped mitigate the rifts that arose during them. I greatly appreciate his help on physics analysis and the many rides home he gave me as well.

Without the help of Ryu Sawada and Yusuke Uchiyama, much of my work would have proceeded far less smoothly than it did. Their immense help with my questions on computing, calorimeter analysis, and physics analysis were vital to my work and this thesis.

I am also indebted to the team on site, at PSI, for their welcoming spirit. Malte Hildebrandt was indispensable in helping me better understand the drift chamber hardware, and his genial character vivified the rainiest of days.

Life would not have been the same without Oleg Kiselev, whose sense of humor provided a source of much needed joy.

One could always depend on Peter-Raymond Kettle to be the voice of reason amidst mayhem. He handled the difficult task of assigning shifts with great diplomacy.

Stefan Ritt aided me in comprehending the DRS and was responsive as host to the needs of foreign collaborators.

From the collaborating institutions in Italy, I worked with a number of notable individuals. My interactions with Angela Papa were highly affable, and she was happy to answer my questions about the calorimeter.

Giovanni Signorelli was consistently constructive to the experiment in his work and in his comments on other's work, and he generously offered me rides home during several shifts.

My encounters with Gianluigi Boca, though few, were memorable because of his merry and hilarious nature.

CURRICULUM VITAE

Benjamin Alexander Golden

EDUCATION

Doctor of Philosophy in Physics **2012**
University of California, Irvine *Irvine, California*

Bachelor of Science in Physics **2004**
University of California, Los Angeles *Los Angeles, California*

RESEARCH EXPERIENCE

Graduate Research Assistant **2007–2011**
University of California, Irvine *Irvine, California*

Research Assistant **2002–2004**
Institute of Geophysics and Planetary Physics at UCLA *Los Angeles, California*

TEACHING EXPERIENCE

Graduate Teaching Assistant **2005–2007**
University of California, Irvine *Irvine, California*

SELECTED HONORS AND AWARDS

GAANN Fellowship **2004–2007**
United States Department of Education

REFEREED JOURNAL PUBLICATIONS

- A limit for the $\mu \rightarrow e\gamma$ decay from the MEG experiment** 2010
Nucl. Phys. **B834** (2010) 1-12
- Calibration and monitoring of the MEG experiment by a proton beam from a Cockcroft-Walton accelerator** 2011
Nucl. Instrum. Meth. **A641** (2011) 19-32
- New Limit on the lepton-flavour-Violating Decay $\mu^+ \rightarrow e^+\gamma$** 2011
Phys. Rev. Lett. **107** (2011) 171801

REFEREED CONFERENCE PUBLICATIONS

- Results from the 2009 data of the MEG experiment in the search for $\mu^+ \rightarrow e^+\gamma$** 2011
Nucl. Phys. **B218** (2011) 62-67

CONFERENCES AND SEMINARS

- New Results from the MEG Experiment** Sep. 2010
The 11th International Workshop on Tau Lepton Physics *Manchester, UK*
- New Results from the MEG Experiment** May 2011
High Energy Experiment Seminar, Boston University *Boston, Massachusetts*

ABSTRACT OF THE DISSERTATION

In Pursuit of the Rare Decay $\mu^+ \rightarrow e^+\gamma$:
A Window to the UV Completion of the Standard Model

By

Benjamin Alexander Golden

Doctor of Philosophy in Physics

University of California, Irvine, 2012

Professor William Molzon, Chair

The MEG experiment is the latest in a series of progressively more precise rare decay searches designed to exploit an observation of $\mu^+ \rightarrow e^+\gamma$ or lack thereof as a sensitive low energy probe of new physics. A search for $\mu^+ \rightarrow e^+\gamma$ is performed using the 43 days worth of data acquired by the MEG experiment during 2009.

To best estimate the number of signal events in the data, a maximum likelihood fit is performed. A 90% confidence interval is then constructed using the Feldman-Cousins technique. The analysis is done blind by hiding signal-like events until the selection cuts and likelihood function are settled upon.

The result is consistent with the null hypothesis and an upper limit is set: $BR(\mu^+ \rightarrow e^+\gamma) < 7.9 \times 10^{-12}$ (90% CL).

Chapter 1

Introduction

Although the Standard Model of particle physics has been strongly supported by experimental evidence, it is thought to be a low energy effective field theory of a more fundamental microscopic theory for very compelling reasons. This implies that new physics comes in at some mass scale that has remained elusive to past experiments (i.e., the existence of heavy particles that have not yet been detected). Experimentally, this can be explored in at least two ways: doing high energy experiments to observe heavy particles directly or doing high precision experiments to measure small effects in the low energy effective regime, suppressed by inverse powers of a heavy mass scale.

Pursuing the latter route, the separate conservation of lepton number for each generation comes into question in theories beyond the Standard Model. While this symmetry is known to be violated among the neutral leptons by observations of neutrino oscillations, this implies lepton flavor violating processes among the charged leptons at one-loop order with rates far too small to detect at present, and indeed every experiment so far has yielded results consistent with this being a good symmetry among the charged leptons. In the context of models for new physics, such as supersymmetry, the smallness of lepton flavor violating

amplitudes is unnatural and extra assumptions are made to address the issue. Even with additional constraints in place, some models predict rates for lepton flavor violating processes near the current experimental limits. The decay mode $\mu \rightarrow e\gamma$, in particular, has been the target of numerous search experiments.

The MEG experiment is a $\mu \rightarrow e\gamma$ search aimed at either a discovery or improving the limit set by the MEGA experiment of $BR(\mu^+ \rightarrow e^+\gamma) < 1.2 \times 10^{-11}$ (90% CL)[1] down to 2×10^{-13} . The MEG experiment achieves fast data-taking at comparably low accidental background rates using the world's most intense continuous muon beam at PSI, superior photon timing with a liquid xenon calorimeter, and low-rate positron tracking with a drift chamber immersed in a graded magnetic field. The most recent result from the MEG experiment, combining data from its 2009 and 2010 physics runs, gives the world's best current limit: $BR(\mu^+ \rightarrow e^+\gamma) < 2.4 \times 10^{-12}$ (90% CL)[2].

This thesis presents an independent $\mu^+ \rightarrow e^+\gamma$ search using data from MEG's second physics run in 2009. Chapter 2 reviews the Standard Model and the reasons why it is thought to be incomplete, and discusses the presence of lepton flavor violation in theoretical scenarios for new physics. Chapter 3 briefly summarizes the history of past $\mu \rightarrow e\gamma$ searches and event signatures for signal and background. Chapters 4-6 outline the MEG detector setup, the reconstruction algorithms, and the calibration procedures. Chapters 7-8 examine the current level of detector performance. Lastly, chapter 9 presents the results of a maximum likelihood analysis to set a confidence interval on $BR(\mu^+ \rightarrow e^+\gamma)$.

Chapter 2

Theoretical Background

The Standard Model of particle physics is well-tested but regarded as incomplete. A great deal of models have been proposed to solve its problems. These models predict relationships between their free parameters and $BR(\mu \rightarrow e\gamma)$, which allow for a rate near the current experimental limit on $\mu \rightarrow e\gamma$ in some cases. This chapter provides a compendium on the Standard Model, its shortcomings, and various scenarios for new physics along with their predictions for $BR(\mu \rightarrow e\gamma)$.

2.1 Standard Model Particle Physics

As reviewed below, the Standard Model of particle physics is a quantum field theory that comes about uniquely given its field content, gauge groups, and the requirement that it be renormalizable.

2.1.1 Electroweak Theory

An ordinary spin- $\frac{1}{2}$ fermion may be coupled to the electromagnetic field by subsuming the inherent global U(1) symmetry

$$\psi(x) \rightarrow e^{-i\theta} \psi(x) \tag{2.1}$$

of its free field Lagrangian

$$\mathcal{L}(x) = \bar{\psi}(x)(i\cancel{D} - m)\psi(x) \tag{2.2}$$

into a more general local U(1) symmetry group

$$\psi(x) \rightarrow e^{-i\theta(x)} \psi(x). \tag{2.3}$$

Eq. 2.3 leaves the action invariant if a simultaneously transforming massless field

$$A_\mu \rightarrow A_\mu - \frac{1}{e} \partial_\mu \theta(x) \tag{2.4}$$

is introduced to form the most general (aside from one conveniently fixed constant) U(1) gauge invariant Lagrangian with, at most, second derivatives

$$\mathcal{L} = \bar{\psi}(x)(i\cancel{D} - m)\psi(x) - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \tag{2.5}$$

where

$$D_\mu \equiv \partial_\mu - ieA_\mu(x) \tag{2.6}$$

and

$$F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu. \quad (2.7)$$

The Euler-Lagrange equation for ψ gives the Dirac equation in an electromagnetic field if A_μ is interpreted as the vector potential and e is the electric charge of the ψ field, while the Euler-Lagrange equations for A_ν give Maxwell's equations if $e j^\nu$ with

$$j^\nu \equiv \bar{\psi} \gamma^\nu \psi \quad (2.8)$$

is identified as the electromagnetic current density for the ψ field. In fact, j^ν is precisely the Noether current associated with the global symmetry Eq. 2.1 of the Lagrangian.

If the weak force is to follow from a comparable symmetry, its form must conform to some highly constraining experimental results that were historically instrumental in suggesting the eventual Glashow-Weinberg-Salam model. Apart from their charge and mass difference, the electron e and neutrino ν_e look the same from the point of view of weak interactions. Although it becomes more approximate with heavier lepton generations, the symmetry is still there, and interactions appear to take place with a universal coupling strength. Experimental establishment of maximal parity violation then seals the chiral nature of the assertion that the Lagrangian obeys an $SU(2)_L$ gauge symmetry of the lepton doublets

$$L_1 = \begin{bmatrix} \nu_e \\ e \end{bmatrix}_L \quad L_2 = \begin{bmatrix} \nu_\mu \\ \mu \end{bmatrix}_L \quad L_3 = \begin{bmatrix} \nu_\tau \\ \tau \end{bmatrix}_L. \quad (2.9)$$

Their lack of participation in weak interactions places right-handed leptons in $SU(2)_L$ singlets:

$$R_1 = e_R \quad R_2 = \mu_R \quad R_3 = \tau_R, \quad (2.10)$$

while the complete failure to even observe the neutrino partners $(\nu_{e,R}, \nu_{\mu,R}, \nu_{\tau,R})$ prompts their exclusion from the theory entirely. This choice of symmetry group finds validation in the Noether currents,

$$J_\mu^i = \sum_{l=1}^3 \bar{L}_l \gamma_\mu \frac{\tau^i}{2} L_l \quad i = (1, 2, 3), \quad (2.11)$$

which reproduce the Fermi currents of the successful low energy V-A theory as well as an additional neutral current upon redefinition:

$$\begin{aligned} J_\mu^+ &= J_\mu^1 + iJ_\mu^2 = \frac{1}{2} \sum_{l=1}^3 \bar{\nu}_{lL} \gamma_\mu l_L \\ J_\mu^- &= J_\mu^1 - iJ_\mu^2 = \frac{1}{2} \sum_{l=1}^3 \bar{l}_L \gamma_\mu \nu_{lL} \\ J_\mu^3 &= \frac{1}{2} \sum_{l=1}^3 \bar{\nu}_{lL} \gamma_\mu \nu_{lL} - \bar{l}_L \gamma_\mu l_L. \end{aligned} \quad (2.12)$$

As always, the Noether charges,

$$T^i = \int J_0^i(x) d^3x \quad (2.13)$$

mimic the algebra of the generators:

$$[T^i, T^j] = i\epsilon^{ijk} T^k. \quad (2.14)$$

Similar arguments applied to the quark sector lead to the classification:

$$\begin{aligned}
Q_{L1} &= \begin{bmatrix} u \\ d \end{bmatrix}_L & Q_{L2} &= \begin{bmatrix} c \\ s \end{bmatrix}_L & Q_{L3} &= \begin{bmatrix} t \\ b \end{bmatrix}_L \\
U_{R1} &= u_R & U_{R2} &= c_R & U_{R3} &= t_R \\
D_{R1} &= d_R & D_{R2} &= s_R & D_{R3} &= b_R
\end{aligned} \tag{2.15}$$

with the caveat that generation changing currents, albeit suppressed, do occur. With so much inequity between the left and right-handed fermions, the theory will never recover electromagnetism without a parity respecting $U(1)_{em}$ symmetry. Electric charge, however, cannot possibly generate a simultaneous $U(1)$ symmetry at this point because the $SU(2)_L$ doublets house fields of dissimilar charge. Rather, one must look to the hypercharge,

$$\frac{Y}{2} = Q - T^3, \tag{2.16}$$

as the only possible linear combination of the neutral current charges Q and T^3 capable of generating a $U(1)_Y$ symmetry. Thus, the electroweak Lagrangian depicts a Yang-Mills gauge theory that is invariant under the direct product group $SU(2)_L \otimes U(1)_Y$ of local transformations:

$$\begin{aligned}
SU(2)_L &: e^{-i\alpha^i(x)\frac{\tau^i}{2}} \\
U(1)_Y &: e^{-i\beta(x)\frac{Y}{2}},
\end{aligned} \tag{2.17}$$

where τ^i are the Pauli matrices when acting on left-handed doublets and zero when acting on right-handed singlets, and of course Y is the field's hypercharge.

Implementing the symmetry outlined in Eq. 2.17 results in the generation ex nihilo of interaction-mediating gauge bosons in the $SU(2)_L \otimes U(1)_Y$ invariant Lagrangian:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^i F^{i\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} + \sum_{l=1}^3 (\bar{L}_l i \not{D} L_l + \bar{R}_l i \not{D} R_l + \bar{Q}_{Ll} i \not{D} Q_{Ll} + \bar{U}_{Rl} i \not{D} U_{Rl} + \bar{D}_{Rl} i \not{D} D_{Rl}). \quad (2.18)$$

Left-handed doublets with hypercharge eigenvalue Y experience the covariant derivative

$$D_\mu = \partial_\mu - ig \frac{\vec{\tau}}{2} \cdot \vec{A}_\mu - ig' \frac{Y}{2} B_\mu \quad (2.19)$$

while for right-handed singlets of hypercharge Y

$$D_\mu = \partial_\mu - ig' \frac{Y}{2} B_\mu \quad (2.20)$$

with gauge boson fields A_μ^i ($i=1,2,3$) needed for local $SU(2)_L$ invariance and B_μ for local $U(1)_Y$ invariance as well as their respective gauge coupling constants g and g' . The corresponding gauge invariant field strength tensors are

$$F_{\mu\nu}^i = \partial_\mu A_\nu^i - \partial_\nu A_\mu^i + g\epsilon_{ijk} A_\mu^j A_\nu^k \quad (2.21)$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu. \quad (2.22)$$

2.1.2 The Higgs Emerges

Inhibited by the very $SU(2)_L \otimes U(1)_Y$ invariance that defines the theory, not one field can have a mass term. Furthermore, the Lagrangian, as it stands, fails to represent the $U(1)_{em}$

symmetry that is empirically observed. For it to be realistic, this theory must undergo a symmetry breaking of $SU(2)_L \otimes U(1)_Y$ down to $U(1)_{em}$ by some mechanism. In the absence of any force with a much larger energy scale than QCD, the symmetry breaking must take place at tree level to give large enough masses to the gauge bosons. This approach leads to the Higgs mechanism adopted by the Standard Model. By acquiring a non-zero VEV, an additional field could break the $SU(2)_L \otimes U(1)_Y$ symmetry and generate masses for the fermions and gauge bosons. It cannot be coupled to the fermions in a gauge-invariant and renormalizable way unless it comes in a $Y = 1$ $SU(2)_L$ doublet and has mass dimension 1, and its VEV will not retain Lorentz invariance unless it is a scalar field. This scalar doublet,

$$\phi = \begin{bmatrix} \varphi^+ \\ \varphi^0 \end{bmatrix}, \quad (2.23)$$

can then enter the Lagrangian in several new gauge-invariant terms:

$$\begin{aligned} \mathcal{L}_\phi = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi^\dagger \phi) - \sum_{l,m=1}^3 G_{lm} \bar{L}_l \phi R_m + \Gamma_{lm}^D \bar{Q}_{Ll} \phi D_{Rm} \\ + \Gamma_{lm}^U \bar{Q}_{Ll} (i\tau_2 \phi^*) U_{Rm} + h.c. \end{aligned} \quad (2.24)$$

Renormalizability demands the early termination of the power series expansion of $V(\phi^\dagger \phi)$:

$$V(\phi^\dagger \phi) = m^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2. \quad (2.25)$$

When $\lambda > 0$ and $m^2 < 0$, the scalar potential assumes a festive Mexican hat shape, which is minimized by any one of a continuous set of degenerate vacua satisfying

$$\phi^\dagger \phi = \frac{v^2}{2}, \quad \text{with } v = \sqrt{\frac{|m^2|}{\lambda}}. \quad (2.26)$$

Not just any vacuum is acceptable, but the choice

$$\langle \phi \rangle = \begin{bmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{bmatrix} \quad (2.27)$$

will indeed leave the electric charge generator Q unbroken:

$$Q \langle \phi \rangle = 0. \quad (2.28)$$

When the Lagrangian is then re-expressed in terms of the field shifts from the classical vacuum, two separate mass matrices arise from the Higgs coupling to the gauge bosons through the covariant derivative $(D_\mu \phi)^\dagger (D^\mu \phi)$. One gives rise to the charged vector bosons as its mass eigenstates

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(A_\mu^1 \mp iA_\mu^2) \quad (2.29)$$

with mass eigenvalue

$$M_W = \frac{1}{2}gv. \quad (2.30)$$

The other contains a zero eigenvalue corresponding to the massless photon

$$A_\mu = \frac{1}{\sqrt{g^2 + g'^2}}(g'A_\mu^3 + gB_\mu) \quad (2.31)$$

with the coupling constant (electron charge)

$$e = \frac{gg'}{\sqrt{g^2 + g'^2}}, \quad (2.32)$$

as well as a massive neutral boson

$$Z_\mu = \frac{1}{\sqrt{g^2 + g'^2}}(gA_\mu^3 - g'B_\mu) \quad (2.33)$$

with mass eigenvalue

$$M_Z = \frac{1}{2}v\sqrt{g^2 + g'^2}. \quad (2.34)$$

From Eq. 2.24, no symmetry constrains the nine independent parameters of the tree-level lepton mass matrix:

$$\mathcal{L}_{lepton}^{(mass)} = - \begin{bmatrix} \bar{e}_L & \bar{\mu}_L & \bar{\tau}_L \end{bmatrix} \begin{bmatrix} m_{ee} & m_{e\mu} & m_{e\tau} \\ m_{\mu e} & m_{\mu\mu} & m_{\mu\tau} \\ m_{\tau e} & m_{\tau\mu} & m_{\tau\tau} \end{bmatrix} \begin{bmatrix} e_R \\ \mu_R \\ \tau_R \end{bmatrix} + h.c. \quad (2.35)$$

This is taken into the mass eigenstate basis by suitable separate unitary transformations on the chiral components:

$$\begin{bmatrix} e'_L \\ \mu'_L \\ \tau'_L \end{bmatrix} = U^\dagger \begin{bmatrix} e_L \\ \mu_L \\ \tau_L \end{bmatrix} \quad \begin{bmatrix} e'_R \\ \mu'_R \\ \tau'_R \end{bmatrix} = V^\dagger \begin{bmatrix} e_R \\ \mu_R \\ \tau_R \end{bmatrix} \quad (2.36)$$

thereby diagonalizing the mass matrix

$$U^\dagger M V = \begin{bmatrix} m_e & 0 & 0 \\ 0 & m_\mu & 0 \\ 0 & 0 & m_\tau \end{bmatrix}, \quad (2.37)$$

which is real once all phases are absorbed into field redefinitions. In terms of mass eigenstates,

the lepton fields,

$$L_l = \begin{bmatrix} \nu_{lL} \\ U_{l1}e'_L + U_{l2}\mu'_L + U_{l3}\tau'_L \\ R_l = V_{l1}e'_R + V_{l2}\mu'_R + V_{l3}\tau'_R \end{bmatrix} \quad (2.38)$$

appear to mix generations; however, we are free to affect the field redefinition:

$$L'_l = \sum_{j=1}^3 U_{lj}^\dagger L_j = \begin{bmatrix} U_{l1}^\dagger \nu_{eL} + U_{l2}^\dagger \nu_{\mu L} + U_{l3}^\dagger \nu_{\tau L} \\ l'_L \end{bmatrix} = \begin{bmatrix} \nu'_{lL} \\ l'_L \end{bmatrix} \quad (2.39)$$

$$R'_l = \sum_{j=1}^3 V_{lj}^\dagger R_j = l'_R.$$

As long as neutrinos are treated as massless, then any linear combination of them is a viable physical basis of mass eigenstates. In the basis of leptons l' and ν'_l , no flavor mixing takes place in the Lagrangian.

With no massless components, the quark sector splits into two independent 3×3 mass mixing matrices by virtue of the two different values of electric charge present in each of the three quark doublets:

$$\mathcal{L}_{quark}^{(mass)} = - \sum_{i,j=1}^3 \bar{D}_{Li} M_{ij}^D D_{Rj} - \sum_{i,j=1}^3 \bar{U}_{Li} M_{ij}^U U_{Rj} + h.c. \quad (2.40)$$

By convention, off-diagonal elements in M^D give rise to generation changing quark transitions, and the very scheme of 3 generations is itself the minimal possible model involving a relative lack of quark fields available for absorbing all the phases of the CKM matrix hence providing the means to reproduce CP violation as observed in K_L decays.

Electroweak “unification” is realized by the vacuum and its distance from the field space origin, which effectuates mass mixing of gauge particles from separate gauge groups, and, because it leaves one combination of generators unbroken rather than none, allows one particular quantum number to emerge unscathed from an otherwise ruinous agglomeration of shattered symmetries. In a stronger sense, however, this is not the same kind of single coupling constant unification sought after by GUTs, for despite its accurate portrayal of two different forces of nature, their alleged union is merely the peculiar mingling of two independently postulated gauge groups with two different coupling constants.

2.1.3 Limits of the Standard Model

The Standard Model is based on 19 free parameters (9 fermion masses, 4 fermion mixing parameters, the QCD vacuum angle, 3 coupling constants, and 2 parameters in the Higgs potential) whose values cannot be predicted from any underlying principles. Various features at tree-level have been experimentally verified, for example, all of its particles except the Higgs have been observed. An extraordinary amount of precision tests of its predictions at the quantum level have also been carried out with good agreement.[3] The state of blissful repose found in that harmony between theory and experiment, however, is disquieted by a number of fundamental shortcomings and aesthetic problems that plague the Standard Model.

The Standard Model does not attempt to include general relativity. Since these effects become important near the Planck scale, $M_{Pl} \sim 10^{18}$ GeV, the Standard Model can work well at everyday energies as a low-energy effective field theory. The Standard Model also lacks a dark matter candidate. While one can extend the Standard Model to include neutrino masses, that part of the Lagrangian is not well constrained (see section 2.3.2), and the number of free parameters increases.

In addition, there are cases of required unnatural fine-tuning for the theory to be consistent. The three forces require three coupling constants that do not quite intersect when run to high energies. Charge quantization is left unexplained. The existence of three generations, and the sheer number of input parameters in general is unappealing. Perhaps the most outstanding fine tuning problem is that of the Higgs self-energy, which is discussed at length in section 2.2.

2.2 The Case for Supersymmetry

With its construction based partially on the deliberate omission of operators with dimension greater than four from the Lagrangian, the Standard Model technically indulges the physicist with the power to circumvent infinite loop corrections to a particle mass and renormalize them against its bare mass into an appropriate phenomenological value regardless of the extent to which fine-tuning is required. As one descends the ladder of spin, various symmetries play a curious role in the degree of “naturalness” in this procedure. Spin-1 gauge bosons, which arise in the Standard Model as constructs to enforce its three gauged symmetries, are shielded from acquiring mass by the same symmetries that their existence maintains. Although gauge invariance is vacuously true at tree level, anomalies can potentially invalidate gauge invariance at one loop. As a result of the vanishing of

$$tr [\{T_\alpha, T_\beta\} T_\gamma], \tag{2.41}$$

however, gauge invariance remains intact in the Standard Model and its gauge bosons naturally remain massless. Considering the 173 GeV-top quark, spin- $\frac{1}{2}$ fermions apparently did not procure such effective sanctuary. More importantly, one finds that the fermion self-energy

diverges logarithmically; the bare electron mass, in particular, is shifted by

$$\delta_{m_e} \approx 2 \frac{\alpha_{em}}{\pi} m_e \log \frac{\Lambda}{m_e} \quad (\Lambda \rightarrow \infty). \quad (2.42)$$

From the low-energy effective field theory stance, however, the Standard Model is not valid beyond some large energy scale $\Lambda = M_{Pl}$, thus, neglecting effects beyond this scale, one finds a more placid correction: $\delta_{m_e} \approx .24m_e$. The dependence on m_e is a vestige of the breaking of an otherwise exact chiral symmetry by a non-zero tree level electron mass. In the limit $m_e \rightarrow 0$, an exact chiral symmetry keeps the electron massless to all orders in perturbation theory.

Ironically, the simplest representation of the Lorentz group demands the most byzantine fine tuning procedure of all. The Standard Model spin-0 neutral higgs scalar receives quadratically divergent corrections to its propagator due to fermion loops. Even with a cutoff $\Lambda = M_{Pl}$, the correction tends to drive the Higgs mass toward the Planck scale. From Eq. 2.24, any transformation $\phi' = A\phi + B$ that is a symmetry of the Lagrangian has $B = 0$ due to the Yukawa couplings, and a unitary operator A . Whether $m_{\phi^0} = 0$ or not, the symmetry group is exactly the same, which means that no symmetry protects the Higgs mass and the corrections are therefore independent of m_{ϕ^0} . Now all naturalness is lost, for one must balance large corrections with a similarly large bare mass just delicately enough to leave an observable mass of order 1 TeV or less. This gauge hierarchy problem makes for one convincing argument in favor of physics beyond the Standard Model. One stoic gaze at the one-loop correction to m_H^2 due to a fermion coupling $-\lambda_f H \bar{f} f$:

$$\Delta m_H^2 = -\frac{|\lambda_f|^2}{8\pi^2} \Lambda_{UV}^2 + \text{non-quadratically divergent terms} \quad (2.43)$$

reveals at once to the trained eye that its Λ_{UV}^2 term would be precisely canceled by the one-loop corrections to m_H^2 due to the coupling of 2 complex scalars $-\lambda_f |H|^2 |S_L|^2$ and

$-\lambda_f|^2 |H|^2 |S_R|^2$. One expects that a hypothetical symmetry relating fermions and bosons is a space-time symmetry because the gap between a spin-0 scalar and a spin- $\frac{1}{2}$ fermion can only be bridged by an operator that shifts (spin) angular momentum. This is seemingly at odds with the no-go theorem of Coleman-Mandula[4] that constrains the Lie algebra of the S-matrix symmetry generators of any consistent quantum field theory to be a direct product of the Poincaré group and an internal symmetry group. By assuaging the harshness of that theorem's assumptions through the permission of generators that form a graded Lie algebra, the Haag-Lopuszanski-Sohnius theorem[5] was precisely the necromancer that resurrected the possibility of non-trivial extension of the Poincaré group. In that case, the most general algebra that generates symmetries of the S-matrix consistent with relativistic quantum field theory is (with some reasonable assumptions and the omission of central charges for simplicity):

$$\{Q_\alpha^A, \bar{Q}_{\beta B}\} = 2\sigma_{\alpha\beta}^m P_m \delta_B^A \quad (2.44)$$

$$\{Q_\alpha^A, Q_\beta^B\} = \{\bar{Q}_{\alpha A}, \bar{Q}_{\beta B}\} = 0 \quad (2.45)$$

$$[P_m, Q_\alpha^A] = [P_m, \bar{Q}_{\alpha A}] = 0 \quad (2.46)$$

$$[P_m, P_n] = 0 \quad (2.47)$$

where Greek letters are Weyl spinor indices running from 1 to 2; Latin letters are Lorentz indices from 1 to 4; and finally, capital letters span an internal space from 1 to N. Further discussions will perpetuate the pigeonholing of central charges and consider only the N=1 case.

Clearly no ordinary field transformation suffices to capture a SUSY transformation, since the generators do not satisfy a Lie algebra. For this, one introduces anti-commuting Grassmann

spinors:

$$\{\theta^\alpha, \theta^\beta\} = \{\theta^\alpha, \bar{\theta}_{\dot{\beta}}\} = \{\bar{\theta}_{\dot{\alpha}}, \bar{\theta}_{\dot{\beta}}\} = 0, \quad (2.48)$$

which allows one to construct combinations of generators that *do* satisfy a Lie algebra:

$$[\theta Q, \bar{\theta} \bar{Q}] = 2\theta\sigma^m\bar{\theta}P_m \quad (2.49)$$

$$[\theta Q, \theta Q] = [\bar{\theta} \bar{Q}, \bar{\theta} \bar{Q}] = 0. \quad (2.50)$$

A general SUSY transformation

$$e^{i(\theta Q + \bar{Q}\bar{\theta} - x_\mu P^\mu)} \quad (2.51)$$

must then operate upon objects that contain bosons and fermions and have functional dependence on x , θ , and $\bar{\theta}$. These “superfields” have two important manifestations that form irreducible representations of the SUSY algebra. In the left and right chiral representations of the SUSY generators:

$$P_\mu = i\partial_\mu, \quad Q_\alpha^L = \frac{\partial}{\partial\theta^\alpha}, \quad \bar{Q}_{\dot{\beta}}^L = -\frac{\partial}{\partial\bar{\theta}^{\dot{\beta}}} + 2i\theta^\alpha\sigma_{\alpha\dot{\beta}}^\mu\partial_\mu \quad (2.52)$$

$$P_\mu = i\partial_\mu, \quad Q_\alpha^R = \frac{\partial}{\partial\theta^\alpha} - 2i\sigma_{\alpha\dot{\beta}}^\mu\bar{\theta}^{\dot{\beta}}\partial_\mu, \quad \bar{Q}_{\dot{\beta}}^R = -\frac{\partial}{\partial\bar{\theta}^{\dot{\beta}}} \quad (2.53)$$

it is possible to forge SUSY covariant derivatives

$$\{D_\alpha, Q_\alpha\} = \{\bar{D}_{\dot{\alpha}}, Q_\alpha\} = \{D_\alpha, \bar{Q}_{\dot{\alpha}}\} = \{\bar{D}_{\dot{\alpha}}, \bar{Q}_{\dot{\alpha}}\} = 0 \quad (2.54)$$

whose asymmetric yet complimentary structure in the fermionic coordinates:

$$D_\alpha^L = \frac{\partial}{\partial\theta_\alpha} + 2i\sigma_{\alpha\dot{\beta}}^\mu \bar{\theta}^{\dot{\beta}} \partial_\mu, \quad \bar{D}_{\dot{\beta}}^L = -\frac{\partial}{\partial\bar{\theta}^{\dot{\beta}}} \quad (2.55)$$

$$D_\alpha^R = \frac{\partial}{\partial\theta_\alpha}, \quad \bar{D}_{\dot{\beta}}^R = -\frac{\partial}{\partial\bar{\theta}^{\dot{\beta}}} - 2i\theta^\alpha \sigma_{\alpha\dot{\beta}}^\mu \partial_\mu \quad (2.56)$$

leads one to define left and right chiral superfields by imposing the blatantly chiral SUSY covariant constraints:

$$\bar{D}\Phi_L = 0 \quad (2.57)$$

$$D\Phi_R = 0. \quad (2.58)$$

Then the most general left chiral superfield, for example, takes a simple form in the left chiral representation due to the nilpotency of Grassmann numbers (i.e. $\theta^1\theta^1 = \theta^2\theta^2 = 0$):

$$\Phi_L(x, \theta) = \phi(x) + \sqrt{2}\theta^\alpha \psi_\alpha(x) + \theta^\alpha \theta^\beta \epsilon_{\alpha\beta} F(x). \quad (2.59)$$

The SUSY covariance property Eq. 2.54 parents the closure of the SUSY algebra in this representation; no transformation of the form Eq. 2.51 can turn a left (or right) chiral superfield into one that is not. With two physical fermionic degrees of freedom (on-shell), the second term in Eq. 2.59 represents a left or right-handed SM fermion. Since θ has mass dimension $-1/2$, which follows from Eq.'s 2.52 and 2.53, $\psi(x)$ and $\phi(x)$ may be assigned the customary mass dimensions of $3/2$ and 1 for SM fermions and their superpartners (sfermions) respectively. Off-shell, the two real propagating degrees of freedom in $\phi(x)$ are augmented by the two real degrees of freedom in the residual field, F , to match the four real degrees of freedom in $\psi(x)$.

To capture the rest of the known spectrum, one needs superfields for the spin-1 gauge bosons and their superpartners (gauginos). Borrowing the characteristic gauge field property of self-conjugacy, one defines vector superfields by:

$$V(x, \theta, \bar{\theta}) \equiv V^\dagger(x, \theta, \bar{\theta}). \quad (2.60)$$

The most general vector superfield satisfying Eq. 2.60 is a heinous divagation from simplicity that retains several unphysical degrees of freedom. Simplification requires the generalized gauge transformation law for a chiral (matter) superfield Φ charged under a gauge group with generators T_a and the vector (gauge) superfield(s) V_a to which it couples:

$$\Phi \rightarrow e^{\Lambda_a T_a} \Phi \quad (2.61)$$

$$e^{V_a T_a} \rightarrow e^{-(\Lambda_a T_a)^\dagger} e^{V_a T_a} e^{\Lambda_a T_a} \quad (2.62)$$

where Λ_a is a set of chiral superfields. Capitalizing on the freedom in Eq. 2.62, a vector superfield may be expressed compactly in the Wess-Zumino gauge:

$$V(x, \theta, \bar{\theta}) = \bar{\theta} \bar{\sigma}^\mu \theta A_\mu(x) + i\theta\theta\bar{\theta}\bar{\lambda}(x) - i\bar{\theta}\bar{\theta}\theta\lambda(x) + \frac{1}{2}\theta\theta\bar{\theta}\bar{\theta}D(x). \quad (2.63)$$

Under a SUSY transformation, Eq. 2.51, the spatial and fermionic coordinates undergo a translation, which cannot introduce non-real terms to a real superfield, Eq. 2.63. Thus, the constraint Eq. 2.60 is once again SUSY covariant. As the only candidate with two real degrees of freedom (on-shell), the real field A_μ can be identified as a spin-1 gauge field, duly requiring the complex spinor λ to assume a mass dimension of $\frac{3}{2}$, a felicitous attribute for a spin- $\frac{1}{2}$ gauge field superpartner or gaugino. Off-shell, the three real propagating degrees of freedom in $A_\mu(x)$ are augmented by one real degree of freedom in the residual field, D , to match the four real degrees of freedom in $\lambda(x)$.

To assemble a supersymmetric action from chiral and vector superfields, it is sufficient that the Lagrangian transforms as a total derivative under a supersymmetry transformation, for vanishing boundary conditions, by the divergence theorem. In fact, the ‘‘F-term’’ of a chiral superfield (Eq. 2.59) and the ‘‘D-term’’ (Eq. 2.63) of a vector superfield transform in precisely this manner.

Starting with the latter, one can build vector superfields out of a set of chiral superfields, Φ_i , to produce the supersymmetric, renormalizable, and gauge invariant Lagrangian:

$$\mathcal{L}_{Kinetic} = \int d^2\theta d^2\bar{\theta} [K^{ij} \Phi_i e^{2gQ_j V_a T_a} \Phi_j^\dagger + (H^{ij} \Phi_i e^{2gQ_j V_a T_a} \Phi_j + h.c.)] \quad (2.64)$$

where K^{ij} is real and symmetric under the interchange of i and j . [6] The integrand is called the Kähler potential and is responsible for kinetic terms and gauge interactions.

As for the former, one can build chiral fields out of any product of chiral fields to produce the supersymmetric and renormalizable Lagrangian:

$$\mathcal{L}_{Superpotential} = \int d^2\theta (k^i \Phi_i + \frac{1}{2} m^{ij} \Phi_i \Phi_j + \frac{1}{3} \lambda^{ijk} \Phi_i \Phi_j \Phi_k). \quad (2.65)$$

Here the integrand is called the superpotential, a completely holomorphic function of the chiral superfields, which is responsible for mass and interaction terms. Gauge invariance typically forces certain components of k , m , and λ to vanish.

To construct a supersymmetric version of the kinetic term for the gauge fields, one first defines the left chiral superfield:

$$W_\alpha = (\bar{D}_{\dot{\alpha}} \bar{D}_{\dot{\beta}} \epsilon^{\dot{\alpha}\dot{\beta}}) e^{-gV} D_\alpha e^{gV}. \quad (2.66)$$

The product $W_\alpha W^\alpha$, also a left chiral superfield, can be shown to be gauge invariant, and

the supersymmetric Lagrangian:

$$\mathcal{L}_{Pure\ Gauge} = \frac{1}{32g^2} \int d^2\theta W_\alpha W^\alpha \quad (2.67)$$

gives the usual kinetic term for gauge fields as well as a corresponding term for the gauginos and a coupling between the two. W_α can be thought of as the supersymmetric field strength tensor.

With all members of a super multiplet forced into mass degeneracy by the form of the superpotential (Eq. 2.65), any supersymmetric model built from Eq.s 2.64-2.65 stands in vexing incongruity with experiment. No charged scalar with the same mass as the electron, for example, has ever been detected. For this reason, supersymmetry must be broken to give heavier masses to the, as of yet, undetected superpartners of Standard Model particles. This can be incorporated into a model by adding soft breaking terms that retain the cancellation of quadratic divergences. Scalar mass terms ($-m_{\phi_i}^2|\phi|^2$), trilinear scalar interactions ($-A_{ijk}\phi_i\phi_j\phi_k + h.c.$), gaugino mass terms ($-\frac{1}{2}m_l\bar{\lambda}_l\lambda_l$), bilinear terms ($-B_{ij}\phi_i\phi_j + h.c.$), and linear terms ($-C_i\phi_i$) all meet this requirement[7].

Since any realistic supersymmetric theory must break down to the Standard Model at low energies, a bottom-up approach can be used to construct a minimal supersymmetric model that contains the necessary gauge groups, fields, and interactions needed to reproduce Standard Model physics. This leads to the Minimal Supersymmetric Standard Model (MSSM). It is defined by its field content, gauge groups, superpotential, and soft breaking terms. Two Higgs doublets are required because one would not be enough to give masses to both up and down types of quarks given that the superpotential must be holomorphic. Then the MSSM consists of the Standard Model gauge group, a superfield for each Standard Model particle

(and its superpartner), two Higgs superfields (H and \bar{H}) instead of one, a superpotential:

$$W_{MSSM} = \sum_{i,j=1}^3 [(\lambda_E)_{ij} H L_i E_j^c + (\lambda_D)_{ij} H Q_i D_j^c + (\lambda_U)_{ij} \bar{H} Q_i U_j^c] + \mu H \bar{H}, \quad (2.68)$$

and all possible soft breaking terms. While it is tempting to label this as a low energy effective supersymmetric field theory, it contains greater than 100 free parameters due to the large number of possible soft breaking terms and is not very predictive. Phenomenologically, supersymmetry breaking cannot originate from renormalizable tree level interactions involving MSSM fields; it must occur in a hidden sector that shares interactions with but does not couple directly to the visible sector of MSSM fields, giving rise to the soft breaking terms through some messenger field(s). [8] Consequently, the scale of supersymmetry breaking and dynamics of the theory can be very different depending on what this mechanism is. In particular, it is impossible to make model-independent predictions about lepton flavor violation and the result is completely at the mercy of the UV completion of the MSSM.

2.3 Lepton Flavor Violation

This section discusses the status of lepton flavor violation in the Standard Model and beyond.

2.3.1 LFV in the Standard Model

In the standard model with massless neutrinos, it is possible to find a basis of lepton doublets (L'_l) and singlets (R'_l), given explicitly in Eq. 2.38, that classically conserves lepton flavor for each generation ($l = e, \mu, \tau$). The Lagrangian is invariant under three global U(1)

transformations:

$$\begin{aligned} L'_l &\rightarrow e^{-iq_l} L'_l \\ R'_l &\rightarrow e^{-iq_l} R'_l, \end{aligned} \tag{2.69}$$

giving rise to three conserved Noether charges

$$Q_l = \int d^3x (l'^{\dagger}_L l'_L + l'^{\dagger}_R l'_R + \nu'^{\dagger}_{lL} \nu'_{lL}) = N_l - \bar{N}_l, \tag{2.70}$$

which correspond to the familiar definition of lepton number as the number of particles minus the number of antiparticles of that flavor. Although Eq. 2.69 is a symmetry of the action, it is not a symmetry of the measure of the path integral, and separate lepton flavor conservation is spoiled by instantons at the quantum level. Nonetheless, certain non-anomalous global symmetries involving baryon number and lepton number are retained: $B - L$, $L_\mu - L_e$, and $L_\tau - L_e$. Together they prevent the occurrence of processes such as $\mu \rightarrow e\gamma$ and the generation of any possible neutrino mass term at even the non-perturbative level.[9] In light of definitive observations of neutrino oscillations, the Standard Model Lagrangian must be augmented by neutrino mass terms, whose precise form remains unknown, and these symmetries no longer hold in general. The sector describing neutrino masses takes the general form:

$$-2\mathcal{L} = \bar{\nu}^c M_\nu \bar{\nu} + h.c. = \begin{bmatrix} \bar{\nu}_L & \bar{\nu}_R^c \end{bmatrix} \begin{bmatrix} m_L & m_D^T \\ m_D & M_R \end{bmatrix} \begin{bmatrix} \bar{\nu}_L^c \\ \bar{\nu}_R \end{bmatrix} + h.c. \tag{2.71}$$

The mass matrix, M_ν , in Eq. 2.71 must somehow reproduce a set of light neutrino masses as measured by several experiments. Regardless of its true form, $\mu \rightarrow e\gamma$ does receive contributions at one loop (shown in Figure 2.1) from neutrino mixing; however, this implies

$BR(\mu \rightarrow e\gamma) \lesssim 10^{-54}$, which is far from measurable [10, 11]. This result is obtained from

$$BR(\mu \rightarrow e\gamma) = \frac{3\alpha}{32\pi} \left| \sum_i (V_{PMNS})_{\mu i}^* (V_{PMNS})_{ei} \frac{m_{\nu_i}^2}{M_W^2} \right|^2 \quad (2.72)$$

using the current measurements of the differences in the squared neutrino masses:

$$|\Delta m_{32}^2| \approx 2.5 \times 10^{-3} eV^2 \quad (2.73)$$

$$\Delta m_{21}^2 \approx 7 \times 10^{-5} eV^2. \quad (2.74)$$

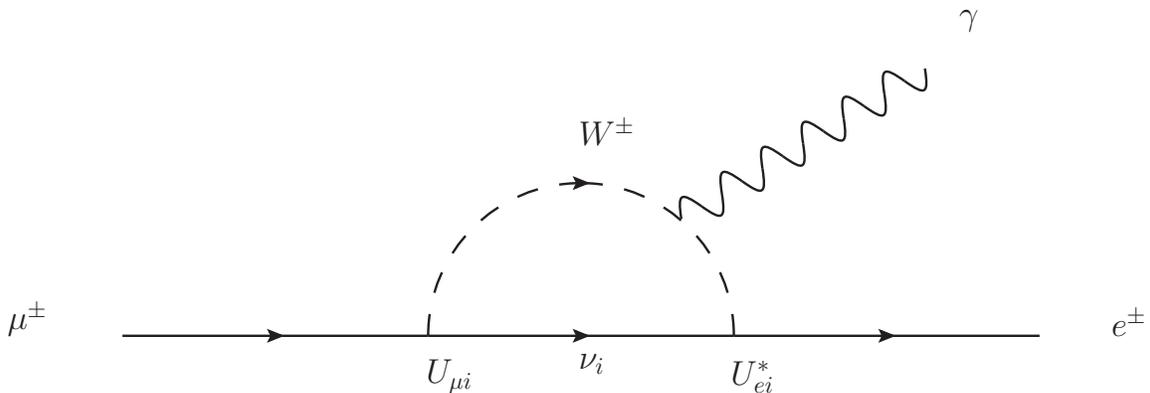


Figure 2.1: Active neutrino loop contribution to $\mu \rightarrow e\gamma$. The internal fermion line is a neutrino mass eigenstate. By diagonalizing the neutrino propagator in flavor space, off-diagonal flavor couplings, $U_{\mu i}$ and U_{ei} , are induced.

An observation of $\mu \rightarrow e\gamma$ would thus demonstrate the existence of new physics beyond the most minimal extension of the Standard Model to include massive neutrinos. In other words, additional diagrams with new matter content would be needed to enhance the branching fraction given in Eq. 2.72 to something detectable. The following section gives several examples of such a scenario.

2.3.2 LFV Beyond the SM

The details behind the mechanism responsible for generating neutrino masses have important consequences for lepton flavor violation. It's possible to generate neutrino masses without introducing any new light particles. The gauge-invariant, dimension-5, irrelevant operator:

$$\frac{\lambda_{ij}}{M}(L_i^T \sigma_2 \phi) C^{-1} (\phi^T \sigma_2 L_j) \quad (2.75)$$

could enter the effective Lagrangian after integrating out some heavy degree of freedom, M . This produces neutrino masses:

$$m_{ij} \sim \frac{\lambda_{ij} \langle \phi \rangle^2}{M}. \quad (2.76)$$

If no new physics occurs below the Planck scale, then by setting $\lambda_{ij} = 1$ and $M = M_{Pl}$, Eq. 2.76 predicts neutrino masses $m_{ij} \sim 10^{-5}$ eV, which are too small. With right-handed neutrinos, a Dirac mass term is possible:

$$Y_{ij}^\nu \bar{L}_i \phi \nu_{Rj} + h.c., \quad (2.77)$$

yielding neutrino masses:

$$m_{ij}^D \sim Y_{ij}^\nu \langle \phi \rangle. \quad (2.78)$$

If this is the sole source of neutrino mass, unnaturally small values of $Y_{ij}^\nu \leq 10^{-12}$, compared to the Yukawa couplings for charged leptons of $\sim 10^{-6} - 10^{-3}$, are required to reproduce experimental results. Right-handed neutrinos can also occur in bare mass terms, which respect gauge invariance because they are singlets under Standard Model gauge groups:

$$M_R \nu_R^T C^{-1} \nu_R + h.c. \quad (2.79)$$

One popular way to achieve small masses for the light neutrinos in a natural way exploits the fact that M_R in Eq. 2.71 is not constrained by any naturalness principle. If one takes $m_L=0$, then diagonalizing the mass matrix in Eq. 2.71 gives $\mathcal{M}_\nu = -m_D^T M_R^{-1} m_D$ for the light neutrinos. m_D is expected to be of the same order as the lepton masses, since they both arise through Yukawa terms. This requires right-handed neutrino masses in the range of $10^6 - 10^{15}$ GeV to get the right scale for the light neutrinos.[12] This is referred to as the type 1 seesaw mechanism, and is in fact equivalent to Eq. 2.76 with M being the right-handed neutrino mass scale.

SUSY In the context of the MSSM, with no inherent guiding principles to limit the form of the soft breaking terms, it's unclear why $\mu \rightarrow e\gamma$ does not take place at large rates. Off-diagonal slepton mass terms ($m_{\tilde{L}_{ij}}^2 \sim m_S^2 \delta_{LL}^{ij}$) contribute:

$$BR(l_i \rightarrow l_j \gamma) \sim \frac{\alpha^3}{G_F^2} \frac{|m_{\tilde{L}_{ij}}^2|^2}{m_S^8} \tan^2 \beta \lesssim (\delta_{LL}^{ij})^2 \tan^2 \beta \times 10^{-5}. \quad (2.80)$$

Here the slepton mass matrix elements are normalized to a typical SUSY mass scale of $m_S \gtrsim 100$ GeV. To predict acceptable rates for Eq. 2.80, the lepton and slepton mass matrices must be very well aligned; $BR(\mu \rightarrow e\gamma) \lesssim 10^{-11}$ requires $\delta_{LL}^{12} \lesssim 10^{-3}$. Such an alignment is quite puzzling because the lepton masses come from Yukawa interactions and electroweak symmetry breaking, while the slepton masses arise mostly from SUSY breaking. This is known as the SUSY flavor problem.[13] A number of solutions have been proposed. In one class of models, a high degree of degeneracy among slepton masses of different generations results from dynamics without specific assumptions about flavor. In gauge mediation scenarios, for example, supersymmetry breaking is delivered to the visible sector by messenger fields charged under Standard Model gauge groups. In these models, the degree of flavor violation is tied to the scale of supersymmetry breaking; if the scale of supersymmetry breaking is low, then the right handed neutrinos decouple before they can contribute to off-diagonal

slepton mass elements and there may be no observable lepton flavor violation. Other proposals involve postulating various flavor symmetries or kinematic suppression through heavy superpartner masses.[14] The following models focus on the gravity mediation scenario, in which the slepton mass matrix is taken to be diagonal and proportional to the unit matrix at the Planck scale, $M_{Pl} \sim 10^{18}$ GeV.

MSSM with Type 1 Seesaw Right-handed neutrinos can be incorporated into the MSSM through a superpotential for the leptons:

$$W = \sum_{i,j=1}^3 \left[(\lambda_E)_{ij} H L_i E_j^c + (\lambda_\nu)_{ij} \bar{H} L_i N_j^c + \frac{(M_R)_{ij}}{2} N_i^c N_j^c \right]. \quad (2.81)$$

Even if a universal scalar mass is assumed at the gravitational scale, slepton mixing is brought about through the running of the renormalization group equations from the Planck scale down to the electroweak scale, as illustrated by the diagrams in Figure 2.2. Since this only affects the left handed sleptons, it predicts $\mu^+ \rightarrow e_R^+ \gamma$ via the class of diagrams shown in Figure 2.3. If neutrino mixing is assumed to come entirely from the Yukawa couplings in the second term of Eq. 2.81 and not from the third term, then solar and atmospheric neutrino observations may be used to estimate typical values for $BR(\mu^+ \rightarrow e_R^+ \gamma)$, displayed in Figure 2.4.[15]

SUSY GUTS Supersymmetric grand unified theories possess the attractive feature of Standard Model gauge coupling unification at an energy scale ($\sim 10^{16}$ GeV) that is very close to the estimated seesaw scale ($\sim 10^{15}$ GeV) and thus may explain the origin of right-handed neutrinos. The well studied $SU(5)$ unification group cannot achieve this naturally because the matter representations, $\bar{\mathbf{5}} \oplus \mathbf{10}$, do not contain right-handed neutrinos and they must be added manually.[12] Moreover, minimal $SU(5)$ is ruled out by strong limits on proton decay from Super Kamiokande.[16]

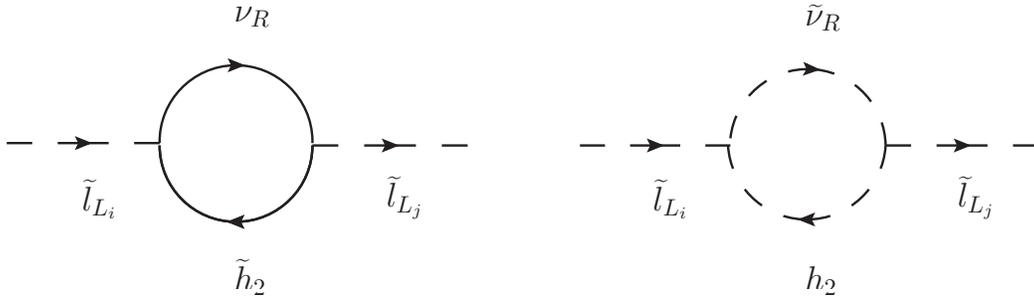


Figure 2.2: Flavor violation in the left-handed lepton sector brought about through quantum effects involving right-handed neutrinos.

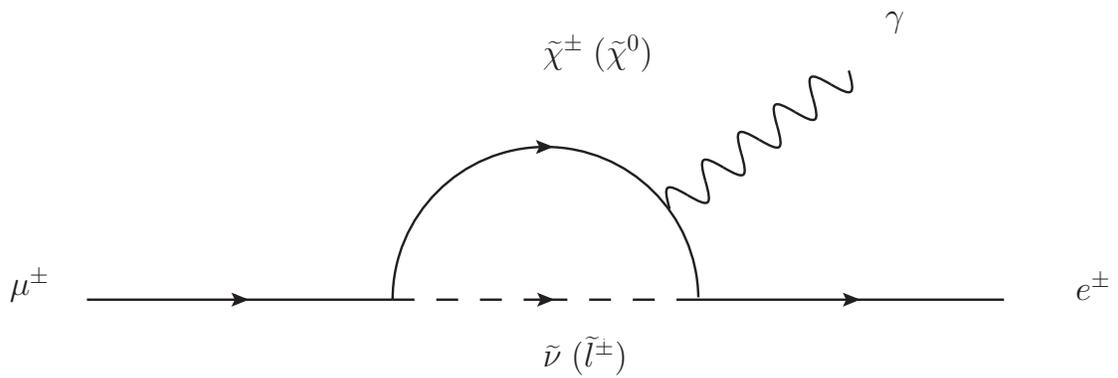


Figure 2.3: $\mu \rightarrow e\gamma$ mediated by slepton mixing.

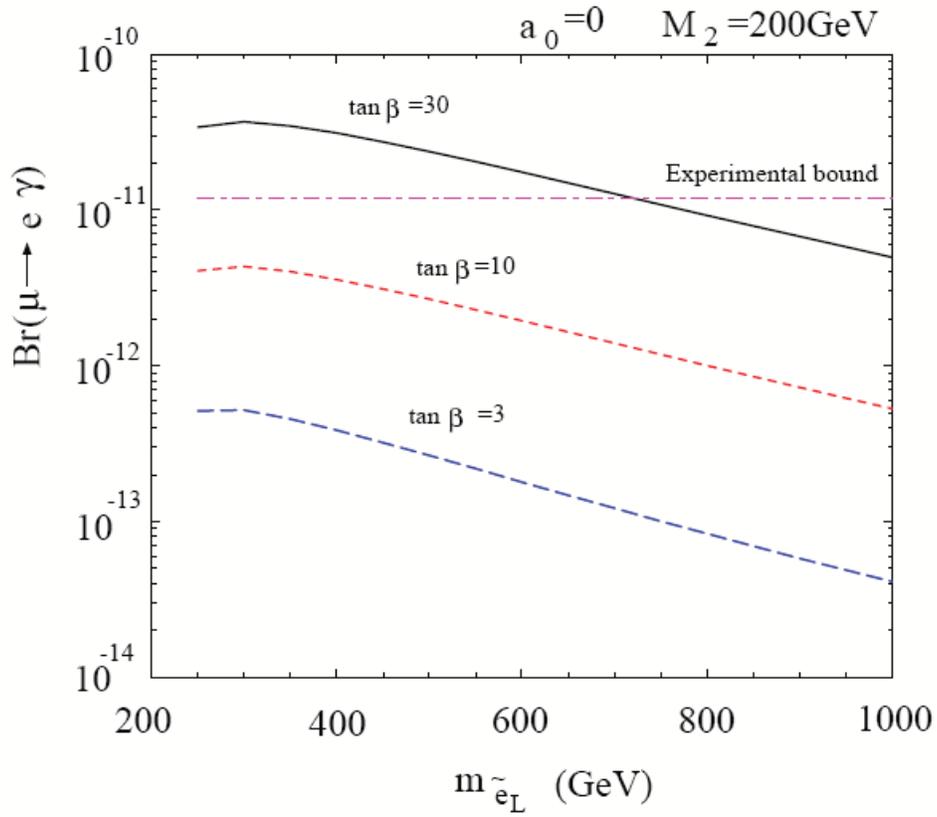


Figure 2.4: Predictions for $BR(\mu^+ \rightarrow e_R^+ \gamma)$ at various values of $\tan \beta$ as a function of the left-handed selectron mass for $M_R = 10^{13}$ GeV, no trilinear coupling, and a 200 GeV wino mass.

In contrast, an $SO(10)$ gauge group has a 16-dimensional representation, which can naturally accommodate all 15 fermions of a single generation (two quarks and two anti-quarks per family with three possible colors, a charged lepton and anti-lepton, and a left-handed Majorana neutrino that is its own anti-particle) as well as a corresponding right-handed neutrino into a single multiplet. The minimal $SO(10)$ model has the superpotential:

$$W_{SO(10)} = \frac{1}{2}(y_u)_{ij}\Psi_i\Phi_u\Psi_j + \frac{1}{2}(y_d)_{ij}\Psi_i\Phi_d\Psi_j \quad (2.82)$$

containing the 16-dimensional representation Ψ_i and two ten-dimensional Higgs fields Φ_u and Φ_d . It can be extended in various ways to include a seesaw mechanism. $B - L$ is a gaugeable subgroup of $SO(10)$ because anomalies cancel, and right-handed neutrino masses emerge naturally through the spontaneous symmetry breaking of this subgroup.[12] Most models work in either the CKM (Cabibbo-Kobayashi-Maskawa) case with small mixing in the neutrino Yukawa couplings or the PMNS (Pontecorvo-Maki-Nakagawa-Sakata) case with large mixing in the neutrino Yukawa couplings. In the context of minimal supergravity, the universal scalar mass m_0 , trilinear couplings A_0 , and universal gaugino mass $M_{\frac{1}{2}}$ are free parameters that tie down the theory at the Planck scale. These are accompanied by the parameters in the superpotential, an undetermined sign in the Higgs potential parameters, and of course $\tan\beta$. For fixed $\tan\beta$, a range in the remaining parameters can be scanned to make predictions for $BR(\mu \rightarrow e\gamma)$. This is shown in Figure 2.5 for a range of parameter space that allows a squark mass below 2.5 TeV, which is LHC accessible.[17]

Other Approaches to the Gauge Hierarchy Problem The presence of measurable charged lepton flavor violation is not unique to the supersymmetric path beyond the Standard Model. It can appear in other solutions to the gauge hierarchy problem for various reasons.

In technicolor models, electroweak symmetry is broken dynamically by a fermion condensate formed by new particles that are charged under a new gauge group. Realistic models require

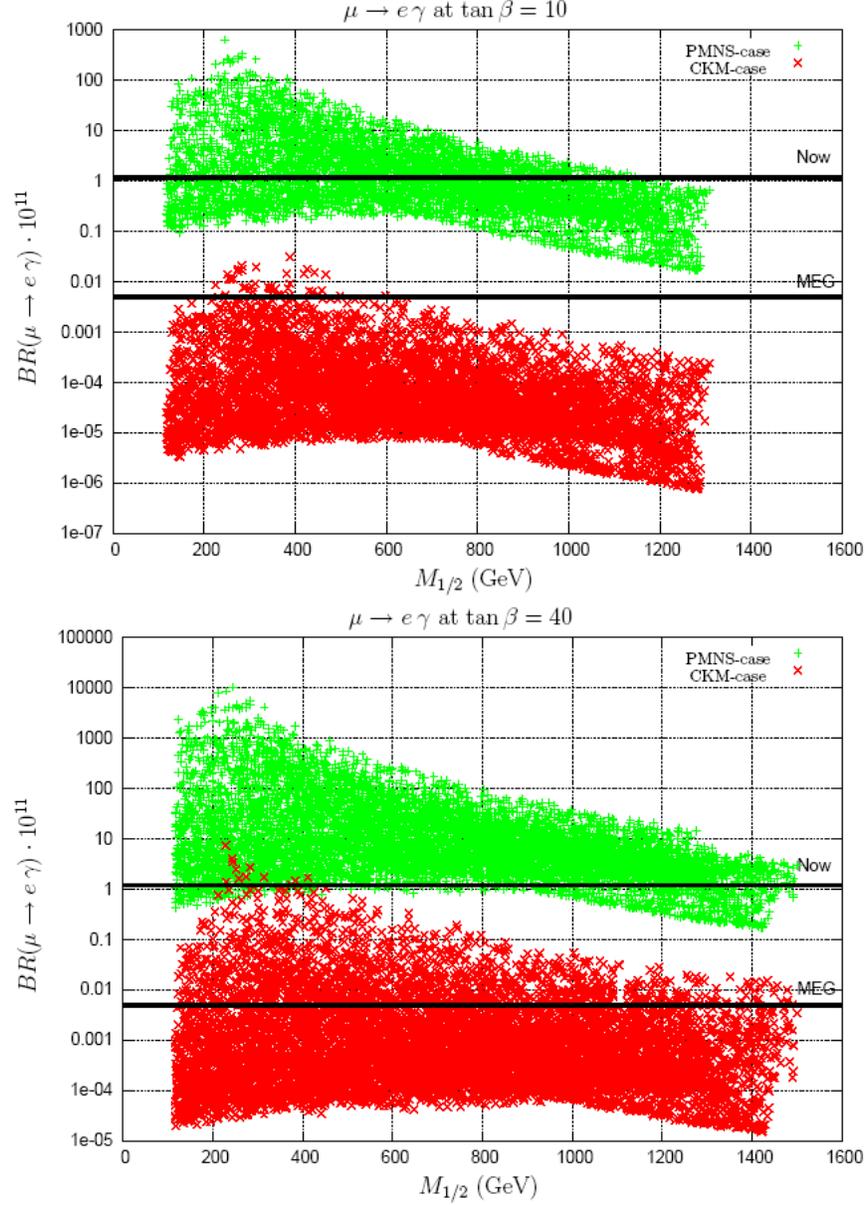


Figure 2.5: Predictions for $BR(\mu \rightarrow e\gamma)$ are shown as a function of the universal gaugino mass for two cases of $\tan\beta$, scanning an LHC relevant space in the parameters describing the Planck scale masses. Both the PMNS case (green) and the CKM case (red) are explored.

non-universal gauge groups that induce lepton flavor violating couplings of gauge bosons to lepton mass eigenstates producing $\mu \rightarrow e\gamma$ at loop level.[18]

In little Higgs models, a nonlinear symmetry is invoked to get cancellation in the one loop corrections to the Higgs mass by particles of the same spin. Charged lepton flavor violation can arise through gauge bosons and exotic scalar multiplets.[18]

Extra dimension models limit the size of the loop corrections to the Higgs mass by reducing the size of the Planck mass; gravity is only observed to be weak due to loss of flux to extra dimensions making the gravitational coupling constant larger and the Planck mass smaller ($M_{Pl} \propto \frac{1}{\sqrt{G}}$). In models explaining neutrino masses in the context of extra dimensions, right-handed neutrinos can generate $\mu \rightarrow e\gamma$ in the bulk with Kaluza-Klein states playing a role similar to that of sparticles in SUSY scenarios.[18]

Model Discriminating Power Clearly a number of different models could be substantiated by an observation of $\mu \rightarrow e\gamma$, but while the predicted rate is highly model dependent, a positive measurement or more stringent limit would provide a great deal of guidance in narrowing down the allowed parameter space in each class of models by constraining their phenomenology. Lastly, a linear correlation between $BR(\mu \rightarrow e\gamma)$ and $BR(\mu N \rightarrow eN)$ would strongly disfavor scenarios where muon conversion is not dominated by on-shell photon exchange.[18] Examples of models that would be supported by such an observation are supersymmetric scenarios (see figure 2.6) and doubly-charged Higgs models.[19]

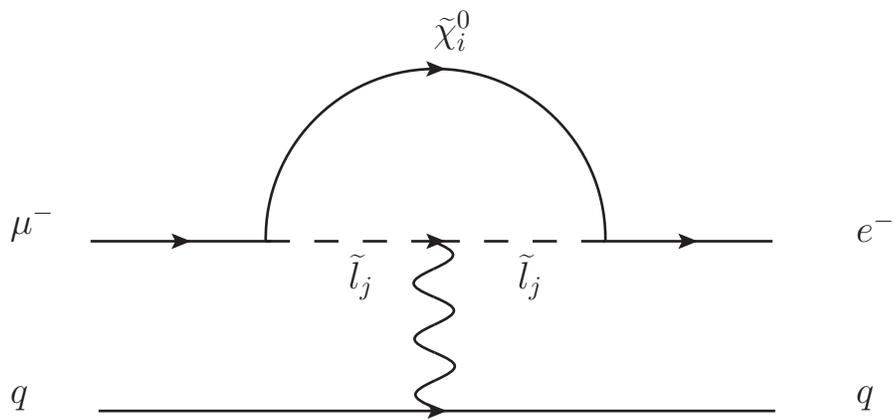


Figure 2.6: Example diagram mediating $\mu N \rightarrow e N$ in supersymmetric models.

Chapter 3

Muon Physics

This chapter gives an overview of past $\mu \rightarrow e\gamma$ searches and explains the event signatures for signal and background.

3.1 History

The muon was initially discovered by Anderson and Nedermeyer[20] in an experiment that measured the fractional energy loss of cosmic ray particles penetrating a 1 cm plate of platinum inside a cloud chamber when a class of penetrating particles was observed to have a mass consistent with being between that of an electron and that of a proton. At first, it was identified as the mediator of the force between protons and neutrons as predicted by Yukawa[21] on the basis of its mass. When experiments later showed that they interacted very weakly with atomic nuclei[22] and that another type of particle with similar mass, the pion[23], was found in cosmic rays, that classification was reassigned to the pion. The decay modes were not known, and the process $\mu \rightarrow e\gamma$, which conserves charge and total lepton number, seemed like a natural candidate mode.

The first search for $\mu \rightarrow e\gamma$ was performed by stopping cosmic ray muons in a graphite absorber and detecting outgoing particles in Geiger-Müller counters giving an upper limit on the branching ratio of less than 10% [24]. In the first measurement to exploit a man-made source of muons, the limit was lowered several orders of magnitude to 2×10^{-5} . [25] Improvements in experimental searches for $\mu \rightarrow e\gamma$ continued. [26, 27, 28, 29, 30, 31]

In the late 1950's, it was pointed out that if the form of the 4-fermion interaction was, as a number of experiments supported, indeed a universal V-A one in which currents interacted with themselves through the exchange of a heavy charged boson, then the predicted branching ratio for $\mu \rightarrow e\gamma$ was about 10^{-4} , at odds with the current experimental limit, although the result was somewhat ambiguous because it depended on the unknown anomalous magnetic moment of the intermediate boson and the regularization of potentially divergent integrals. [32] It was suggested that the apparent absence of $\mu \rightarrow e\gamma$ could be explained by associating separate lepton numbers for muons and electrons, and requiring them to both be conserved. In order to explain the abundantly observed $\mu \rightarrow e\nu\bar{\nu}$, the outgoing neutrinos had to also carry muon (ν_μ) and electron number (ν_e). Both the two-neutrino hypothesis and the separate conservation of lepton number were first confirmed in 1962 at Brookhaven in an experiment that observed $\bar{\nu}_\mu + p^+ \rightarrow \mu^+ + n$ but not $\bar{\nu}_\mu + p^+ \rightarrow e^+ + n$. [33]

For over a decade after that result, $\mu \rightarrow e\gamma$ experiments ceased, until a new era of searches began in 1977 with an experiment making use of the intense muon beam at the Swiss Institute for Nuclear Research (SIN) [34, 35], which is the site of present day PSI. More improvements in the limit followed. [36, 37] The best pre-MEG upper limit of $BR(\mu^+ \rightarrow e^+\gamma) < 1.2 \times 10^{-11}$ (90% *CL*) comes from the MEGA experiment at the Los Alamos National Laboratory Meson Physics Facility (LAMPF). [1] While the process is ruled out exactly in the Standard Model with massless neutrinos, $\mu \rightarrow e\gamma$ appears generically in extensions of the Standard Model, as reviewed in Section 2.3, possibly at rates near the best current limit of $BR(\mu^+ \rightarrow e^+\gamma) < 2.4 \times 10^{-12}$ (90% *CL*), set by MEG. [2] In an era where beyond the

Standard Model physics is pervaded by numerous free parameters weakly constrained by experiments, the prospect for discovery or even a more stringent limit provides an alluring motivation to look for $\mu \rightarrow e\gamma$. At the time of writing, MEG is the latest such experiment in a historically rich ongoing hunt. A summary of the measurements of $BR(\mu^+ \rightarrow e^+\gamma)$ leading up to MEG is provided in Figure 3.1.

Other lepton flavor violating muon interactions are predicted by physics beyond the Standard Model; Table 3.1 summarizes their current limits.

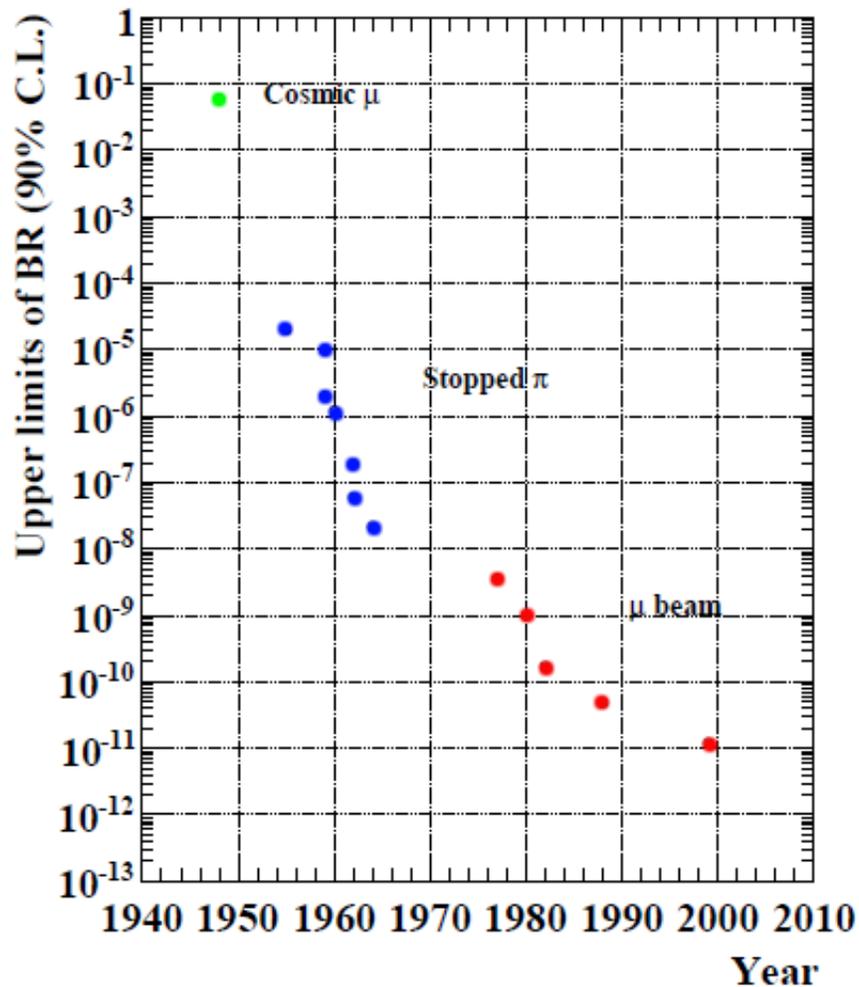


Figure 3.1: A plot of the progress made in improving the upper limit on $BR(\mu \rightarrow e\gamma)$ in the 20th century.

Table 3.1: Current upper limits on exotic muon decay and interaction modes.

Mode	UL (90% CL)	Reference
$\text{BR}(\mu^+ \rightarrow e^+\gamma)$	2.4×10^{-12}	[2]
$\text{BR}(\mu^+ \rightarrow e^+\gamma\gamma)$	7.2×10^{-11}	[37]
$\text{BR}(\mu^+ \rightarrow e^+e^+e^-)$	1.0×10^{-12}	[38]
$\text{BR}(\mu^+ \rightarrow e^+\nu_\mu\bar{\nu}_e)$	1.2×10^{-2}	[39]
$\frac{\Gamma(\mu^- + N \rightarrow e^- + N)}{\Gamma(\text{ordinary muon capture})}$	4.3×10^{-12}	[40]

3.2 $\mu \rightarrow e\gamma$ detection

The use of μ^- 's is made untenable by the processes of muon capture and momentum transfer to the nucleus, hence $\mu \rightarrow e\gamma$ searches rely on μ^+ 's. Signal events can be imitated by two primary sources of background: radiative muon decays and accidental matches of an unrelated positron and photon. These cases are shown in Figure 3.2.

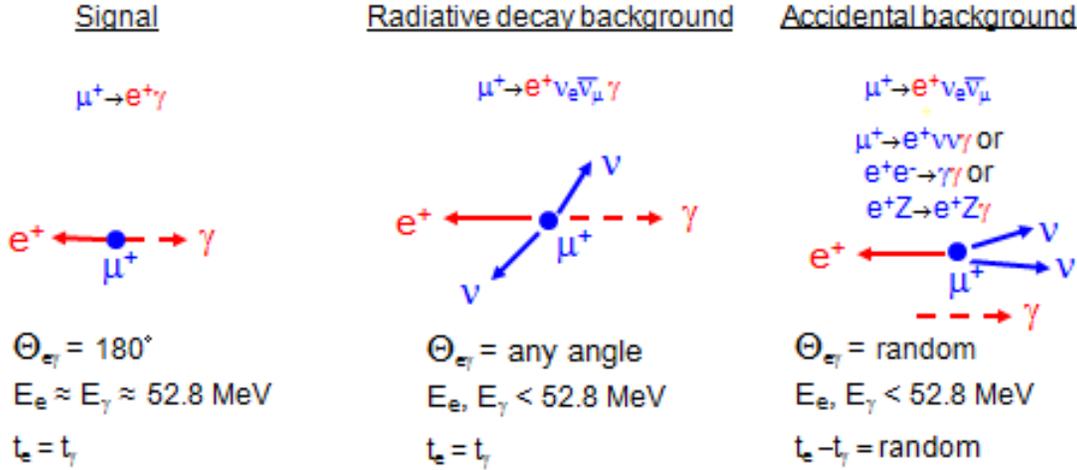


Figure 3.2: An illustration of the event signatures for signal and background. Detected particles are indicated in red, while other particles are indicated in blue. $\Theta_{e\gamma}$ refers to the opening angle between the positron and photon, E_e refers to the positron energy, E_γ refers to the photon energy, t_e refers to the positron emission time, and t_γ refers to the photon emission time.

3.2.1 Signal Event Signature

In the muon rest frame, a $\mu^+ \rightarrow e^+\gamma$ event is identified by a back-to-back positron and photon consistent with simultaneous emission from a common vertex, and with energies approximately equal to half the muon mass. More precisely, conservation of the energy-momentum 4-vector requires $E_{e^+} = 52.830$ MeV and $E_\gamma = 52.828$ MeV. Signal events are discerned from background based on the measured positron energy (E_e), photon energy (E_γ), difference in emanation times ($t_{e\gamma}$), and opening angle ($\Theta_{e\gamma}$).

3.2.2 Radiative Decay Background

One source of background comes from radiative muon decays ($\mu^+ \rightarrow e^+\nu_e\bar{\nu}_\mu\gamma$) with a nearly back to back positron and photon saturating their kinematic energy limits; however, the branching ratio for this process is suppressed in that region of phase space. For $E_e > 46$ MeV, $E_\gamma > 30$ MeV, and no constraint on the opening angle, the radiative decay branching fraction is of order 10^{-7} , as predicted by theory and as measured by none other than MEG itself.[41]

The expected amount of background can be calculated by integrating the tree level differential decay width over a finite signal box. Defining $x \equiv \frac{2E_e}{m_\mu}$, $y \equiv \frac{2E_\gamma}{m_\mu}$, and $z \equiv \pi - \Theta_{e\gamma}$, and δx , δy , δz as the respective half-widths of a signal box centered at $x = y = 1$ and $z = 0$, the expected branching ratio for unpolarized muons is [19]:

$$BR_{signal\ box}(\mu^+ \rightarrow e^+\nu_e\bar{\nu}_\mu\gamma) = \frac{\alpha}{16\pi}[J_1 + J_2]. \quad (3.1)$$

In the relevant case where $\delta z \leq 2\sqrt{\delta x \delta y}$:

$$J_1 = \frac{8}{3}(\delta x)^3(\delta y)\left(\frac{\delta z}{2}\right)^2 - 2(\delta x)^2\left(\frac{\delta z}{2}\right)^4 + \frac{1}{3}\left(\frac{1}{(\delta y)^2}\right)\left(\frac{\delta z}{2}\right)^8 \quad (3.2)$$

$$J_2 = 8(\delta x)^2(\delta y)^2\left(\frac{\delta z}{2}\right)^2 - 8(\delta x)(\delta y)\left(\frac{\delta z}{2}\right)^4 + \frac{8}{3}\left(\frac{\delta z}{2}\right)^6. \quad (3.3)$$

Thus, for energy resolutions (δx and δy) of order 1%, a box analysis with high signal efficiency is subject to radiative decay background at the level of merely 10^{-15} compared to the signal sensitivity goal of order 10^{-13} .

3.2.3 Accidental Background

Another source of background stems from accidental occurrences of a positron coming from Michel decay ($\mu^+ \rightarrow e^+ \nu_e \bar{\nu}_\mu$) and a photon coming either from radiative muon decay, annihilation in flight of a positron in the stopping target, or bremsstrahlung. The effective accidental background branching ratio is proportional to the fraction of the accidental spectrum in each variable contained in the signal region. Random coincidences of uncorrelated positrons and photons naturally populate values of $t_{e\gamma}$, and $\cos \Theta_{e\gamma}$ (opening angle) uniformly and the same is approximately true of E_e because the Michel spectrum is roughly flat near the signal energy as shown in Figure 3.3c. A signal window of width $2\delta t_{e\gamma}$ contributes a factor of $(2\delta t_{e\gamma} R_\mu)$ to the effective accidental branching ratio where R_μ is the instantaneous beam rate. If a cut on the opening angle of $|\Theta_{e\gamma} - \pi| < \delta \Theta_{e\gamma}$ is made, it will accept a portion of the full solid angle that is given by the solid angle subtended by a cone with apex angle $2\delta \Theta_{e\gamma}$: $2\pi(1 - \cos \delta \Theta_{e\gamma}) \approx \pi(\delta \Theta_{e\gamma})^2$. The accepted fraction of the solid angle is then $\frac{\delta \Theta_{e\gamma}^2}{4}$. The fraction of the Michel spectrum falling in a signal window centered at $x = 1$ of width $2\delta x$ is given by integrating the differential unpolarized Michel decay branching ratio over $1 - \delta x \leq x < 1$,

yielding $f_e \approx 2\delta x$. The fraction of accidental photons falling in a signal window centered at $y = 1$ of width $2\delta y$ is well approximated, in the limit that radiative decay photons are the dominant source of accidental photons, by integrating the differential branching ratio for radiative muon decay over $1 - \delta y \leq y < 1$, giving $f_\gamma = \frac{\alpha}{2\pi}(\delta y)^2[\ln(\delta y) + 7.33]$, which is roughly proportional to $(\delta y)^2$ because the radiative decay photon spectrum descends in an approximately linear way near $y = 1$ as shown in Figure 3.3d. The total effective branching ratio for accidental background events in the signal window can then be written as:

$$BR_{signal\ box}(acc.) = R_\mu(2\delta x)\left(\frac{\alpha}{2\pi}(\delta y)^2[\ln(\delta y) + 7.33]\right)(2\delta t_{e\gamma})\left(\frac{\delta\Theta_{e\gamma}^2}{4}\right). \quad (3.4)$$

For fixed $\mu^+ \rightarrow e^+\gamma$ acceptance, the signal window in each variable is proportional to its resolution. As an example, taking 90% efficient windows in each variable and assuming Gaussian resolution functions, the effective branching ratio for accidentals at the resolutions and beam rate quoted by the MEG proposal [42] can be estimated by setting:

$$R_\mu = 3 \times 10^7 \frac{\mu}{s} \quad (3.5)$$

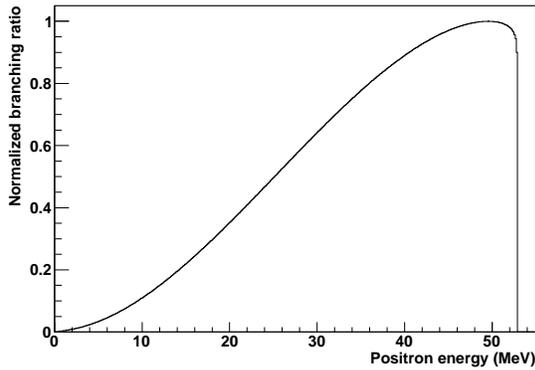
$$\delta t_{e\gamma} = 1.64 \times 64 \text{ ps} \quad (3.6)$$

$$\delta x = 1.64 \times 0.004 \quad (3.7)$$

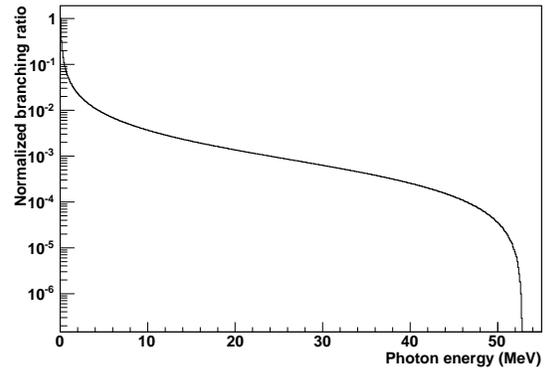
$$\delta y = 1.64 \times 0.017 \quad (3.8)$$

$$\delta\Theta_{e\gamma} = 1.64 \times 8.7 \text{ mrad}. \quad (3.9)$$

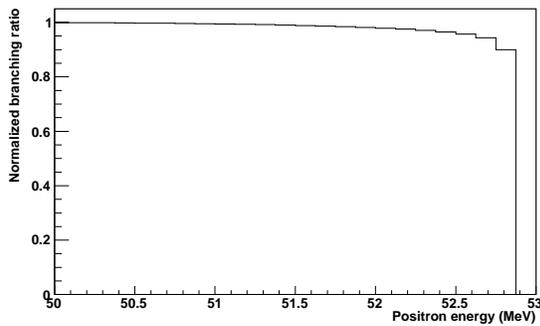
With these values, Eq. 3.4 predicts an effective accidental background branching ratio for MEG at the proposal resolutions of 1.42×10^{-14} . Thus, to suppress the accidental background rate to the level of about one tenth of the signal sensitivity goal of $\approx 10^{-13}$, it is necessary to achieve these resolutions. Since the radiative decay background is about an order of



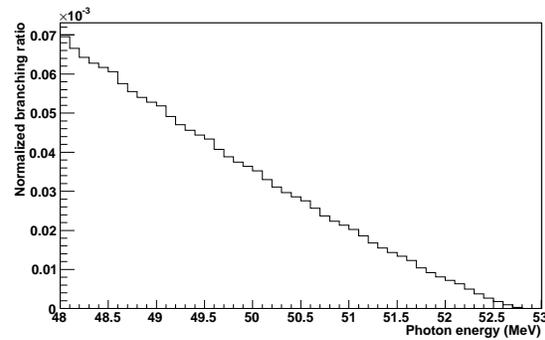
(a) The Michel positron energy spectrum calculated from theory to one-loop for unpolarized muons.



(b) The muon radiative decay photon energy spectrum calculated from theory at tree-level for unpolarized muons, integrating over the positron energy and opening angle. The vertical axis is shown in logarithmic scale to illustrate the steep descent in the high energy region.



(c) A close-up of the high energy region in figure 3.3a.



(d) A close-up of the high energy region in figure 3.3b. The vertical axis is shown in linear scale to illustrate the approximately linear shape in this energy region.

Figure 3.3: The accidental background energy spectra. Each differential branching ratio is normalized to be unity at its maximum value.

magnitude smaller at comparable resolutions, the accidentals constitute the dominant source of background.

Chapter 4

MEG Hardware

This chapter describes the hardware components of the MEG experiment.

Antimuons are brought to rest in a thin stopping target and detectors are positioned to measure photons and positrons from $\mu^+ \rightarrow e^+\gamma$. MEG uses a continuous muon beam with a stopping intensity of ~ 30 MHz. In the previous $\mu^+ \rightarrow e^+\gamma$ search, the MEGA experiment, a pulsed beam of instantaneous stopping intensity ~ 250 MHz and 6.7% duty cycle was used. MEG is able to take data 1.8 times faster than MEGA and faces 12% of the accidental background rate of MEGA for the same region in $E_e, E_\gamma, t_{e\gamma}$ and $\Theta_{e\gamma}$.

Positrons are tracked in a magnetic spectrometer with a graded magnetic field designed to limit the rate without sacrificing acceptance. The momentum is measured by a set of low-mass drift chambers and the time is measured by an array of scintillation counters. In the MEGA experiment, the magnetic field was uniform, thus, as discussed in section 4.4.1, the bending radius depended on the emission angle and the number of turns made in the spectrometer was not limited as in the case of a graded field.

Photons are detected in a fully-absorbing liquid xenon calorimeter, which can achieve good

timing resolution (~ 40 ps) and high detection efficiency ($\sim 60\%$). In MEGA, a thin layer of lead was used to induce photons to pair produce, and the resulting positrons and electrons were detected in drift chambers and plastic scintillators. Because the lead was thin, resulting in good energy resolution, the detection efficiency for photons was low ($\sim 5\%$). The timing resolution was ~ 600 ps.[43]

4.1 Beam

To make a discovery or lower the limit on $BR(\mu^+ \rightarrow e^+\gamma)$ down to the $\sim 10^{-13}$ level, MEG requires an extremely intense source of μ^+ 's. Since a continuous beam operates at a lower instantaneous rate than a pulsed beam for the same amount of muons, it is the optimum choice from the perspective of reducing accidental background. For these reasons, the MEG experiment uses the 590 MeV proton ring cyclotron facility at the Paul Scherrer Institut (PSI) in Villigen, Switzerland, the world's most intense continuous muon source.

4.1.1 Proton Accelerator Complex

A Cockcroft-Walton accelerator first injects 870 keV protons into a ring cyclotron called Injector 2. They are then injected with an initial energy of 72 MeV into the center of another ring cyclotron and reach 590 MeV energies to form a beam current of 2.2 mA. The 590 MeV ring cyclotron is shown in Figure 4.1. They are then directed to a series of meson production targets, one of which, dubbed target E, is a 4 cm thick graphite target.

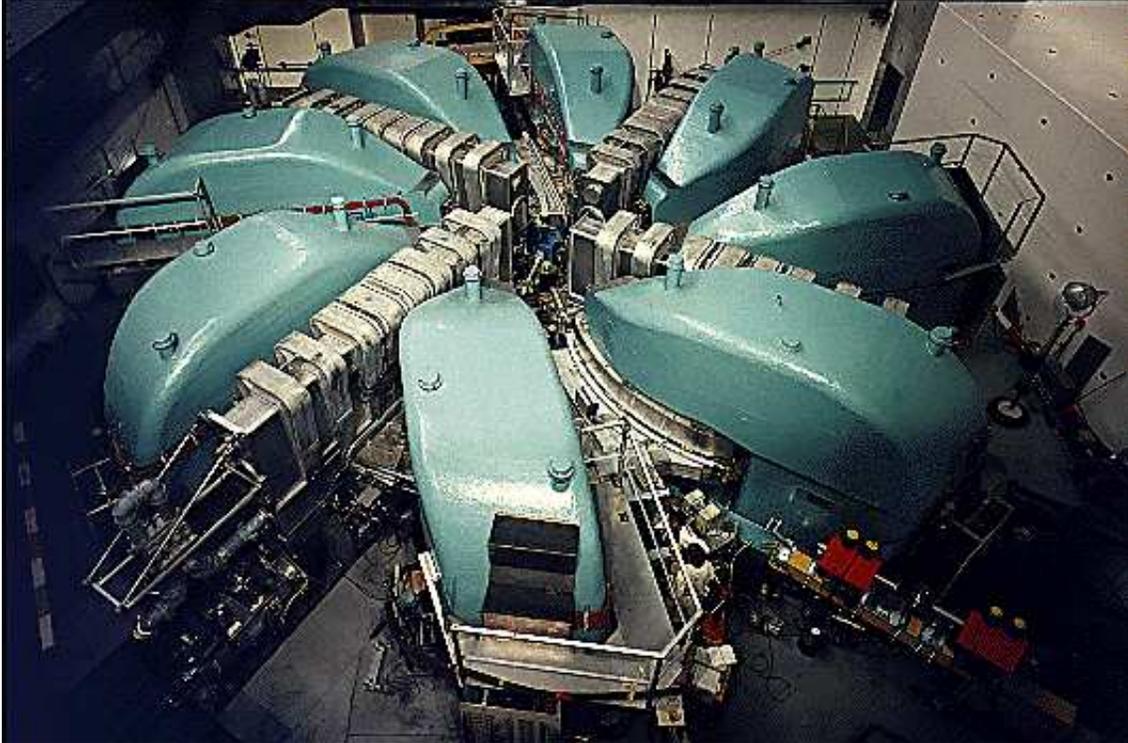


Figure 4.1: An aerial view of the 590 MeV ring cyclotron at PSI.

4.1.2 π E5

π E5, one of several secondary beam lines sharing target E as their source, supplies low energy pions and muons. This beam line is tuned to accept “surface” muons originating from pions decaying at rest near the surface of target E. Such muons have 29.8 MeV/c of momentum by virtue of being produced by pions at rest and travel through a sufficiently small remaining portion of target E that they exit with a small momentum spread ($\sigma \sim 0.85\%$). Figure 4.2 lays out the π E5 beam line components. The measured flux of pions and muons in π E5 at the end of the last bending magnet (AST) is presented in Figure 4.3.

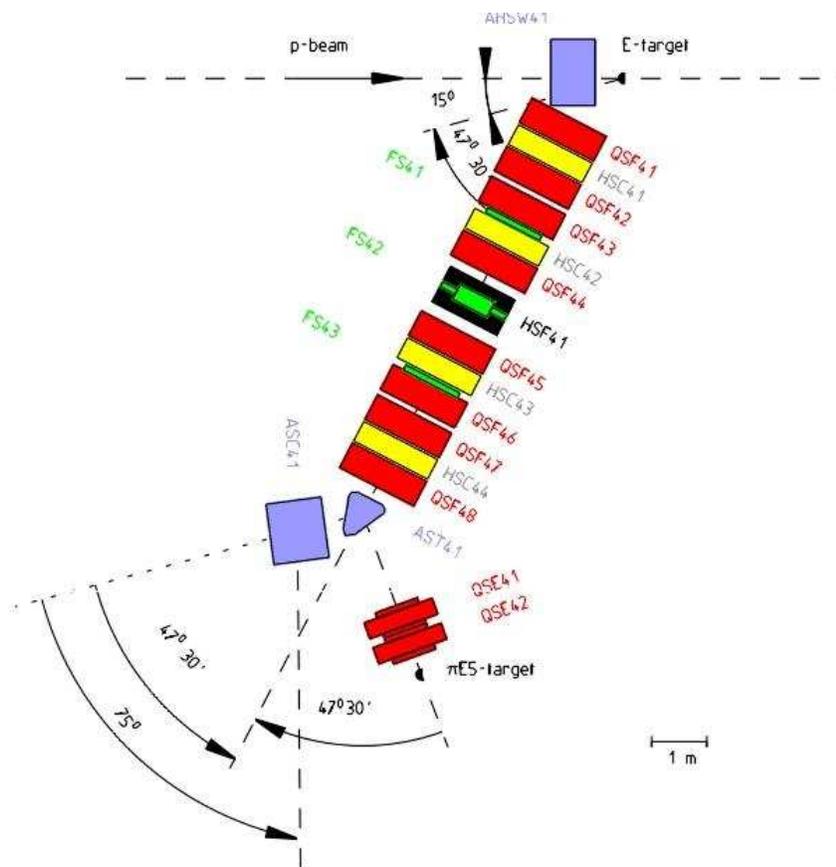


Figure 4.2: A diagram of the π E5 beamline components.

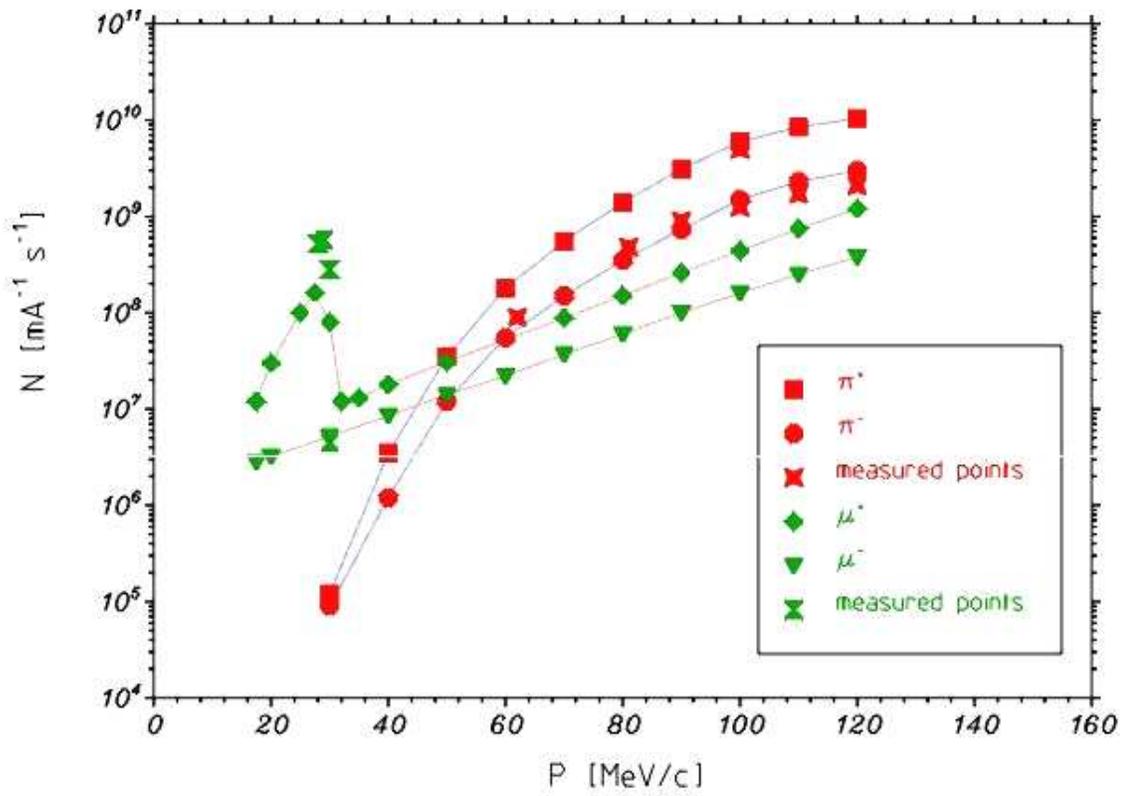


Figure 4.3: Measured flux of beam content.

4.1.3 Beam Transport System

Figure 4.4 illustrates the remaining beam path. The outgoing $\pi E5$ muon beam is directed through two sets of quadrupole triplet magnets that sandwich a Wien filter. The Wien filter applies perpendicular electric and magnetic fields optimized to allow muon passage while deflecting the paths of contaminant positrons. A spatial separation between the positrons and muons of 7.5 times the RMS dispersion of the beam profile is achieved by directing the positrons into a collimator system that functions as a beam-dump. Next, the muon beam is focused by a superconducting beam transport solenoid magnet (BTS) where it encounters a 200-300 μm thick Mylar degrader to slow it down before delivering it to the experiment. The final spot size on the stopping target is $\sigma_x \approx \sigma_y \approx 1 \text{ cm}$.

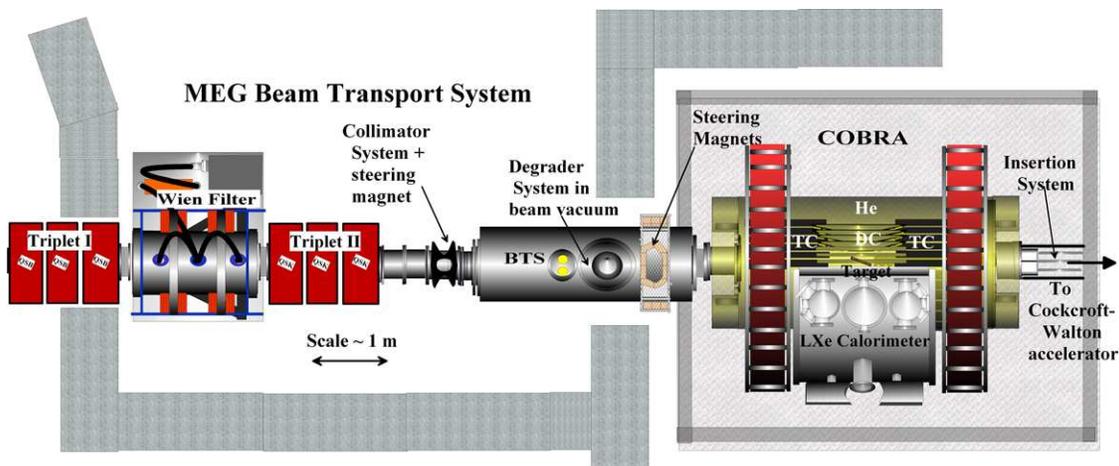


Figure 4.4: An illustration of the beam transport system.

4.2 Coordinate System

The MEG detector setup consists principally of a magnetic spectrometer and an electromagnetic calorimeter located near a stopping target. This setup is outlined in Figure 4.5. The positive z -axis is defined to lie parallel to the beam direction, pointing downstream. Inside the magnet bore, the stopping target is placed in the beam. Nominally, the origin of the

global coordinate system coincides with the center of the target and of the magnet. The positive y -axis points upward and the x -axis is oriented in the manner of a right-handed coordinate system. The calorimeter is confined to the half-volume $x < 0$. When we refer to particle emission angles, θ is the zenith angle made with respect to the positive z -axis ($0 \leq \theta < \pi$), and ϕ is the azimuthal angle made in the x - y plane with respect to the positive x -axis ($-\pi \leq \phi < \pi$).

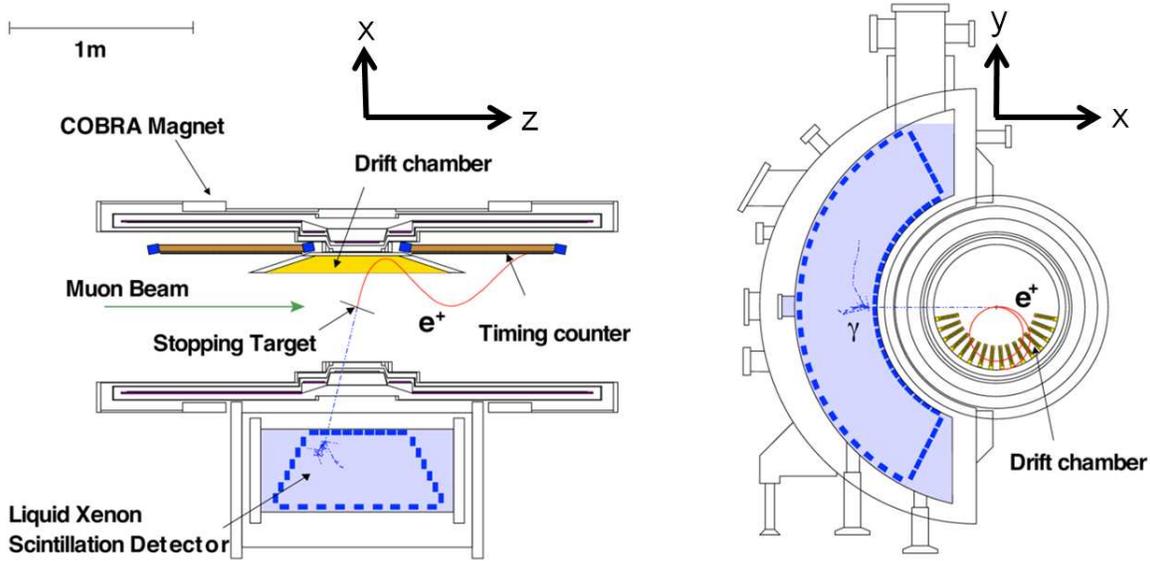


Figure 4.5: A sketch of the experimental layout for MEG.

4.3 Stopping Target

A 205 μm thick, elliptical, polyethylene target, of density 0.922 g/cm^3 , is placed inside the magnetic spectrometer. A Rohacell [44] frame, of density 0.0513 g/cm^3 , supports the polyethylene film. Figure 4.6 shows the standalone target and frame structure, while figure 4.7 represents the target in its fully mounted state. The target material and thickness are optimized to minimize scattering of positrons in the target and annihilation in flight. Since the muons are fully polarized after pion decay, it is also chosen for its depolarizing characteristics. Several holes, with 5 mm radii, are placed in the target to check its positioning using

reconstructed positron tracks projected back to their decay vertex. The target plane is not perpendicular to the z-axis, but instead makes an angle of 20° with it to increase muon stopping power without significantly increasing the amount of target material typically traversed by exiting positrons. Knowledge of the position and orientation of the target is crucial for determining the positron angles. A displacement of the target in the direction perpendicular to its plane of $500 \mu\text{m}$, for example, causes an error in the inferred ϕ_e of about 3.6 mrad.



Figure 4.6: A photograph of the muon stopping target labeled with dimensions.

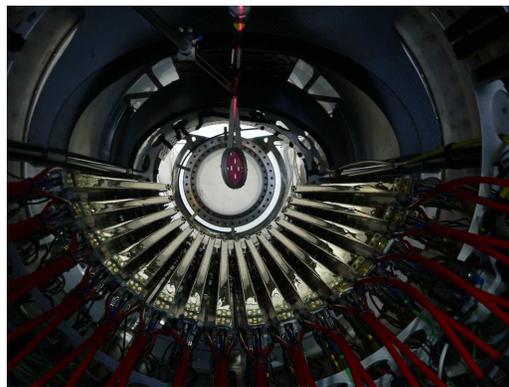


Figure 4.7: A view of the mounted target, looking upstream.

4.4 Magnetic Spectrometer

Positrons are detected and measured in a spectrometer consisting of several components fully immersed in a solenoid magnetic field graded along the beam axis. The trajectories are traced by a discrete set of position measurements in a drift chamber, which is used to infer the particle momentum, the decay vertex, and the path length to the timing counter. Two sets of plastic scintillating bars, the timing counters, are placed at the upstream and downstream ends within the magnetic field to measure the positron time of impact after exiting the drift chamber. Together, these measurements fully characterize the positron kinematics.

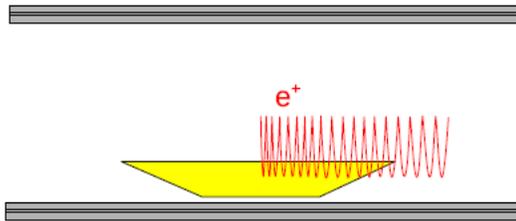
4.4.1 COBRA Magnet

The magnetic field is graded along the beam axis to fall in magnitude as $|z|$ increases. The grading is tuned to compensate the dependence of transverse momentum on the zenith angle and keep the bending radius approximately independent of it. From this property of the field, the name of the magnet (COnstant Bending RADIUS or COBRA) originates. This enables the placement of the drift chambers at large radius, where positrons near the signal momentum travel, and where the rate is low. It also limits the number of turns made in the drift chamber to typically one or two, regardless of the emission angle. Figure 4.8 summarizes these highly advantageous features pictorially. Figure 4.9 compares the positron rate as a function of radius for a uniform and graded field.

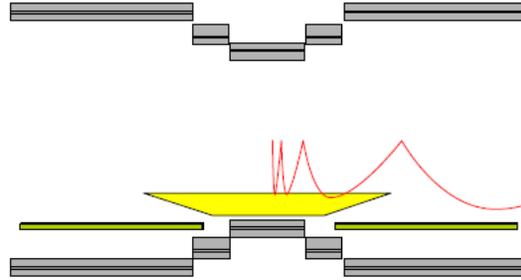
The graded axial magnetic field is set up by a series of superconducting coils of current (360 A), pictured in Figure 4.10. The central coil, centered at $z = 0$, is placed between two gradient coils of larger radius, which themselves are located between two end coils of even larger radius. A pair of normal conducting compensation coils reduces the field near the calorimeter. This structure is mapped out in Figure 4.11. The resulting magnetic field ranges from 1.27 T at $z = 0$ to 0.49 T at $|z| = 1.25$ m, as shown in Figure 4.12.

Since the target is placed within the bore of the magnet, photons must travel through its structure before reaching the calorimeter. For this reason, the wall of the magnet is made very thin, and the total thickness of the superconducting coil, support structure, and cryostat amounts to $0.197X_0$.

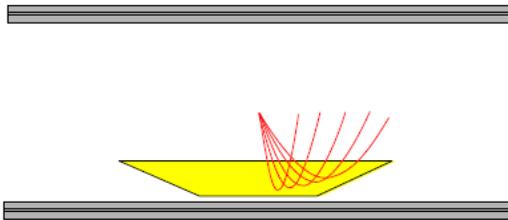
Although the calorimeter is placed outside the magnetic solenoid, the fringe field still presents concerns regarding the performance of PMTs in a magnetic field. The magnetic force on the electrons in a PMT moving from dynode to dynode is most severe when the PMT orientation sets the path of internally moving electrons perpendicular to the direction of the magnetic field, as seen in Figure 4.13. To counteract this, compensation coils are installed in the



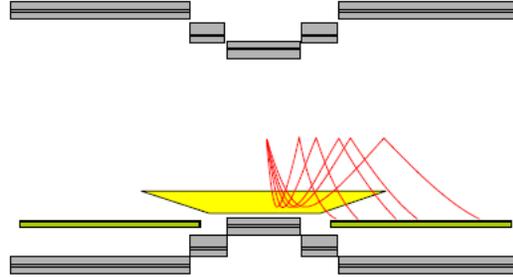
(a) In a uniform axial magnetic field, positrons emitted nearly perpendicular to the beam axis make numerous turns in the drift chamber complicating the pattern recognition and increasing the rate in the drift chambers.



(b) In a graded magnetic field, the bending radius of the positron grows with $|z|$, which quickly sweeps it out of the chambers before making many turns.



(c) In a uniform axial magnetic field, the positron bending radius is a linear function of its transverse momentum.



(d) In a graded magnetic field, the bending radius depends only on the magnitude of its momentum. This enables a precise radial selection window for high momentum tracks.

Figure 4.8: A comparison of uniform magnetic field attributes with those of a graded magnetic field.

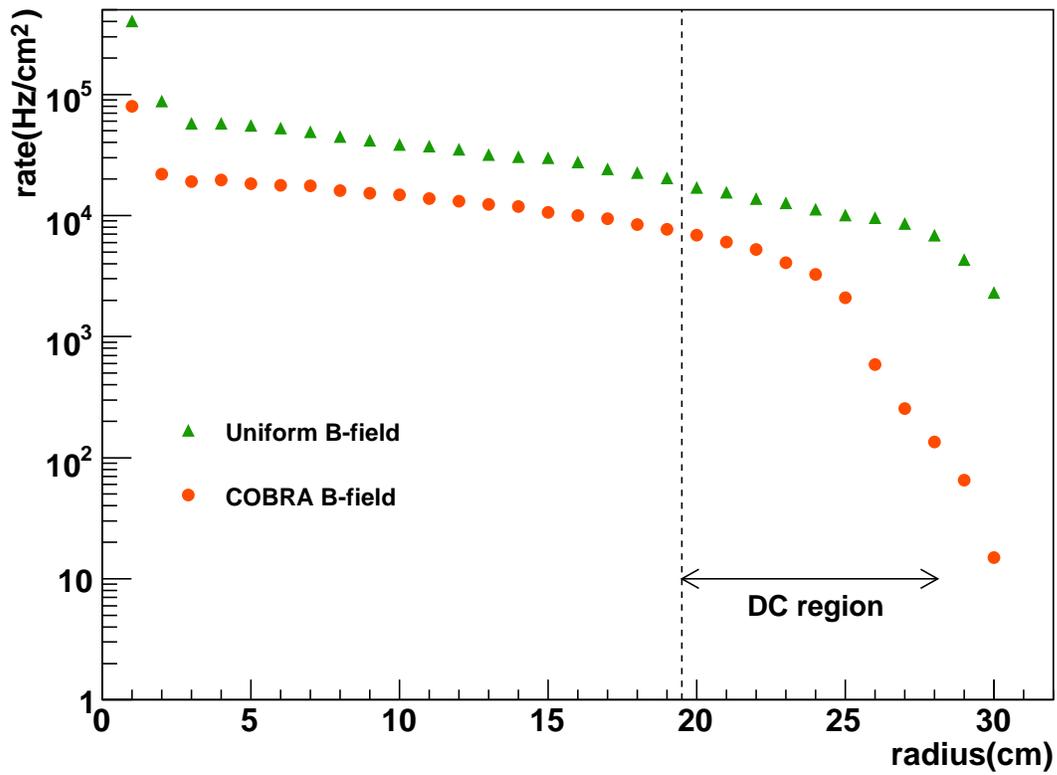


Figure 4.9: The rate of Michel positrons at a given radius for two magnetic field configurations.



Figure 4.10: A photograph of the magnet.

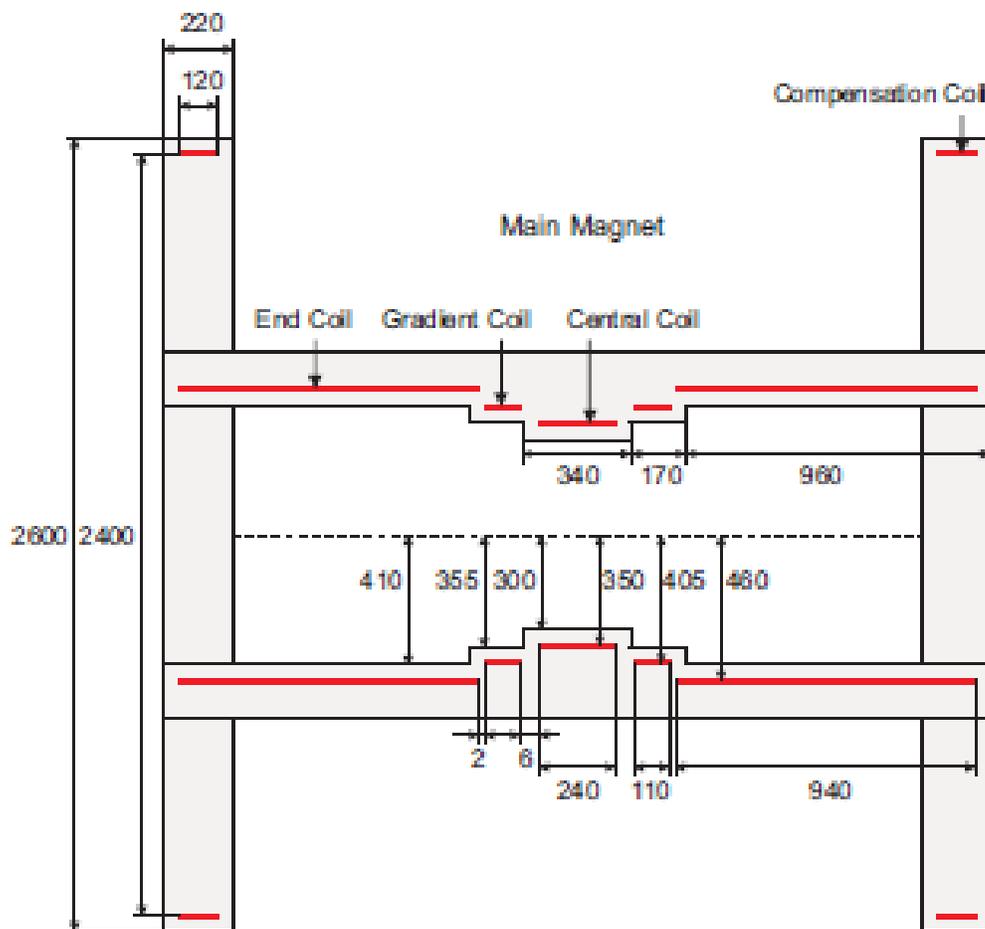


Figure 4.11: A schematic of the magnet design.

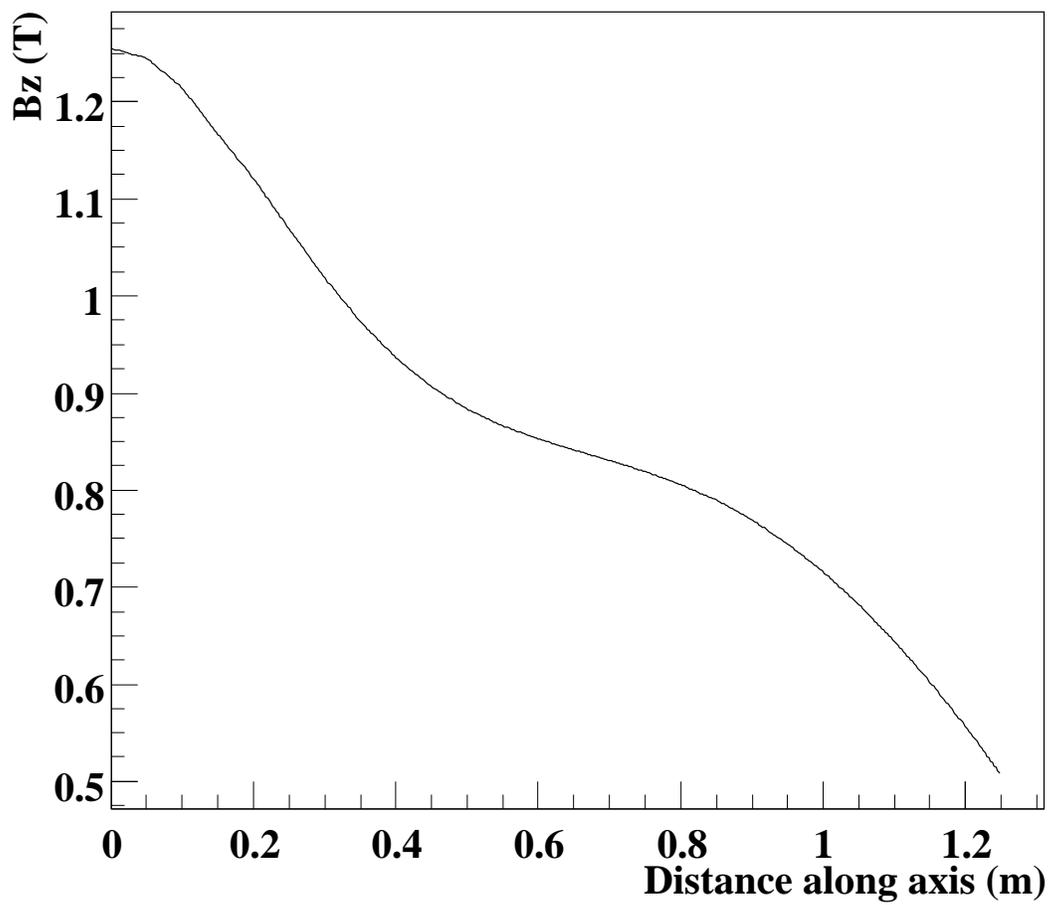


Figure 4.12: The magnetic field strength within the spectrometer as a function of $|z|$.

magnet to cancel the magnetic field in the calorimeter region down to the level of 50 Gauss. Figure 4.14 charts the magnetic field contours achieved throughout the detector as a result of the compensation coils.

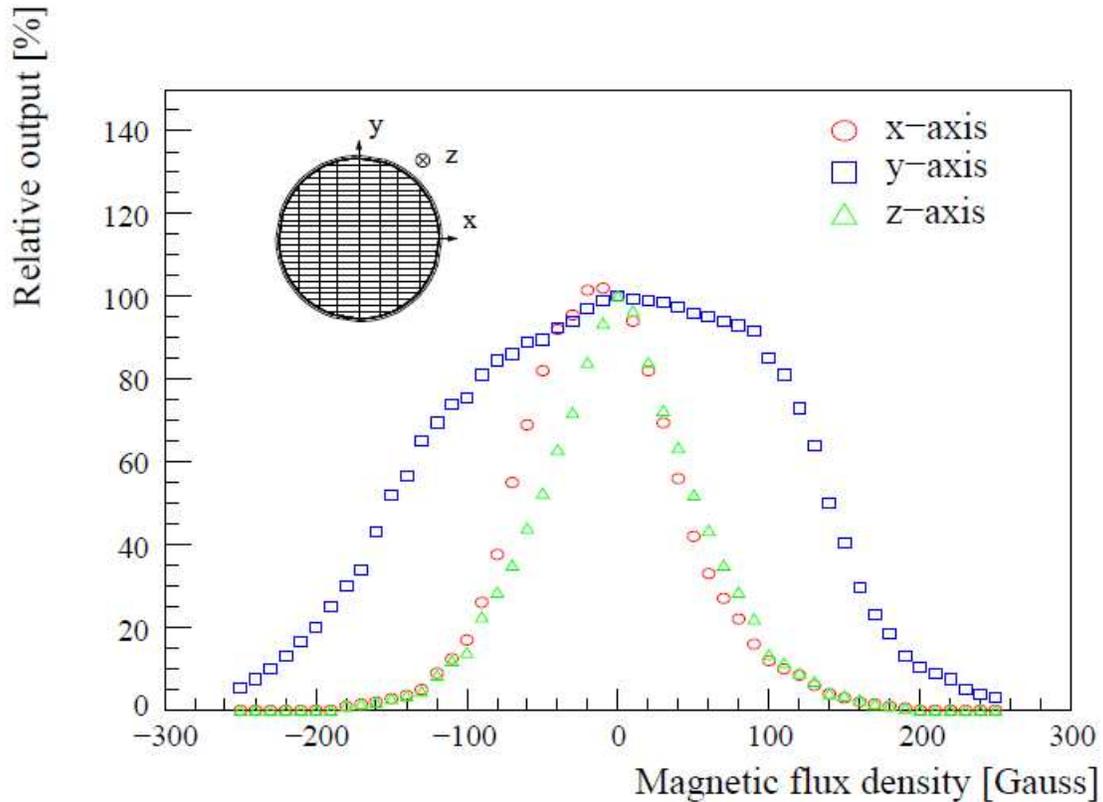


Figure 4.13: PMT gain versus magnetic field strength for different orientations.

4.4.2 Drift Chamber

The drift chamber system is made up of 16 modules, nominally arranged to lie at constant ϕ and spaced from one another by 10.5° , populating the half-volume defined by $y < 0$. Figure 4.15 presents a photograph of the fully installed drift chamber system.

The drift chamber acceptance is designed to extend in zenith emission angle for $|\cos \theta_e| < 0.35$ and in azimuth to lie in the region $|\phi_e| < 60^\circ$. Positrons emitted nearly perpendicular to the beam axis ($|\cos \theta_e| < 0.08$), however, make many turns in the spectrometer and the

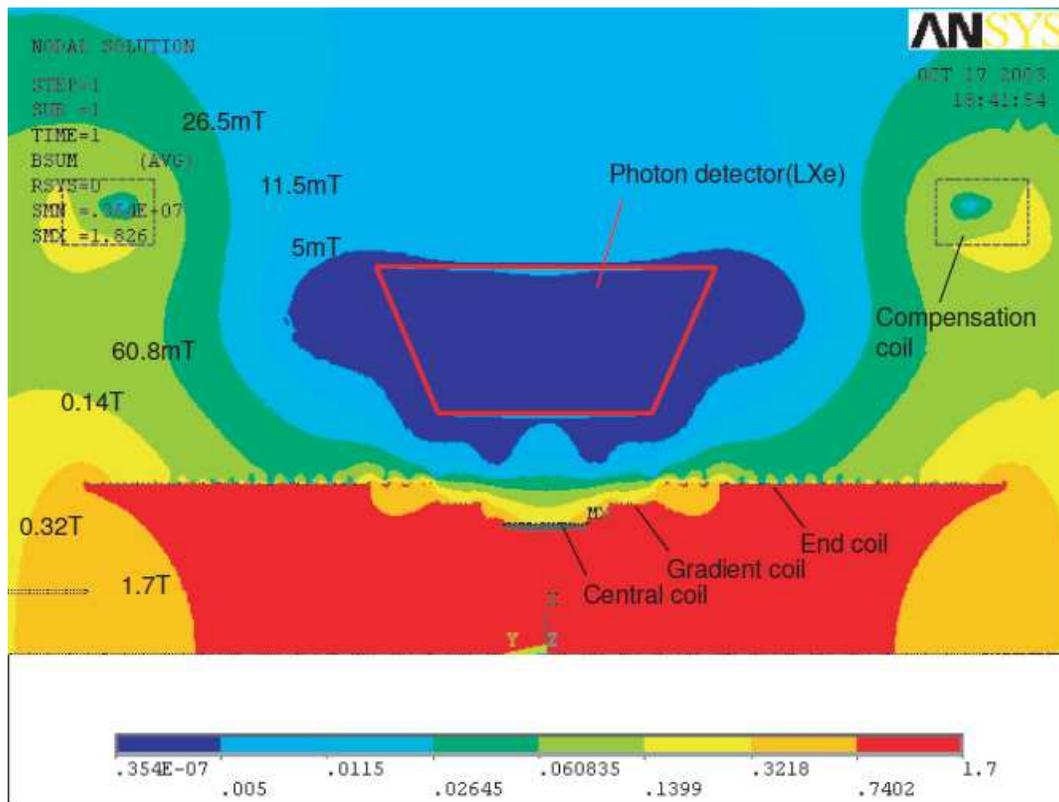


Figure 4.14: Magnetic field contours throughout the detector.

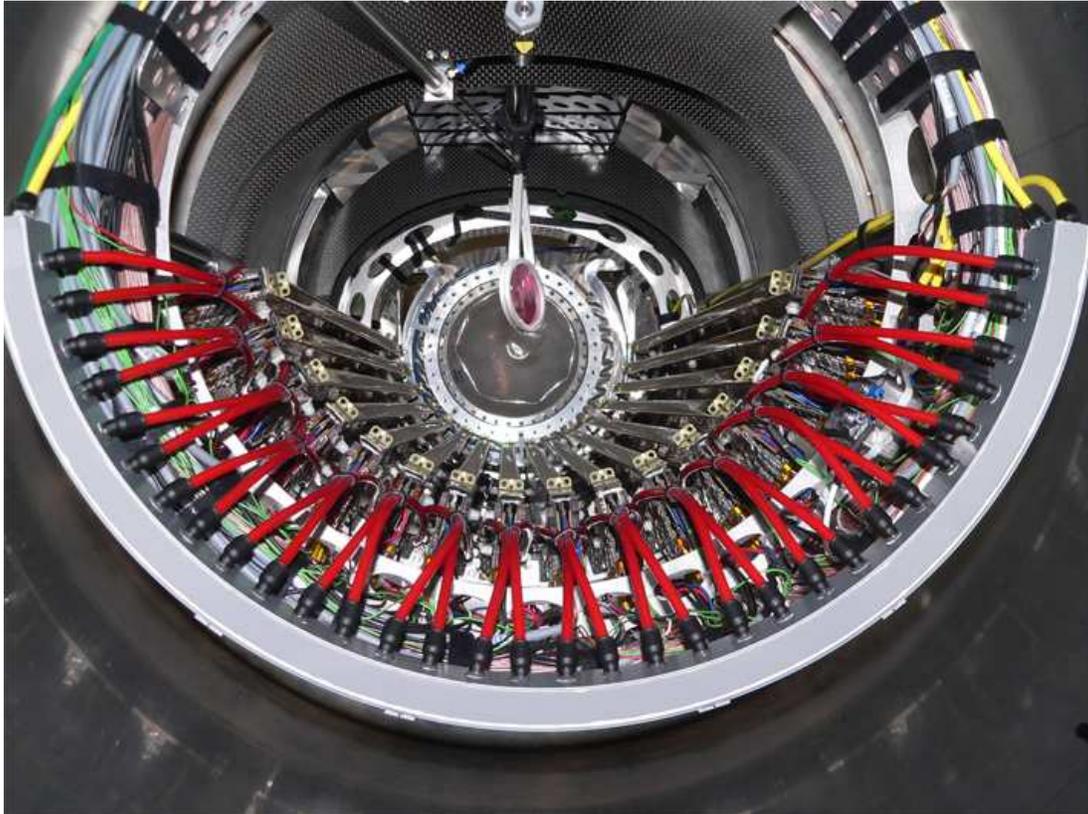


Figure 4.15: A photograph of the 16 drift chamber modules mounted on the support structure.

reconstruction efficiency degrades. The longitudinal extent ranges from 40 cm in length for the shortest wire at the largest radius to 86 cm in length for the longest wire at the smallest radius. Radially, the acceptance of a drift chamber module covers 8.55 cm from approximately $R=20$ cm to $R=28.55$ cm. Maximum efficiency is expected to occur near the signal momentum, 52.8 MeV/c, but non-zero reconstruction probability reaches down to roughly 40 MeV/c.

As viewed from the side (see Figure 4.16), the drift chamber modules are trapezoids of stretched wires increasing in length as one nears the innermost radius. This shape is mandated by the desire to place the timing counters as close to the chambers as possible. Figure 4.17 presents a diagram of a module as viewed from a line of sight parallel to the wires. As shown in the figure, a single module is made up of two planes, each containing nine drift cells. To help resolve the left/right ambiguity, the two planes are staggered from one another by half a cell. The physical volume of a plane is defined by two parallel foils, composed of 12.5 μm thick polyimide with a 250 nm thick aluminum deposition. The outermost planes are called hoods and the innermost planes are called cathodes; they are sometimes referred to as pads. This design brings the average material crossed per module to $2.5 \times 10^{-4} X_0$. At the center of each drift cell, a NiCr anode (or sense) wire, 25 μm in diameter, is held at a nominal operating potential of 1850 V. Field shaping potential wires are placed on the side(s) of a drift cell. A carbon frame (see Figure 4.18), equipped with electronics (PCBs made of G-10), surrounds the module on three of the remaining sides and is left open on the side closest to the target.

Pulses are read out at both the upstream and downstream ends of the wire. A Vernier pattern is etched into the cathode and hood foils, surrounding the drift cell, as illustrated in Figure 4.19. The charge deposited on the anode wire in turn induces charge on the pads. Each of the zig-zag patterns on a pad, which are separated by a gap, is read out electronically. This results in four pad signals per cell: two from the cathode and two from

the hood. The pattern is periodic every 5 cm, and the phases of the patterns are shifted from one another as shown in Figure 4.20. A description of how the anode and pad signals are used for reconstruction is given in section 5.2.1.

Each drift chamber module is filled with a mixture of equal parts helium (He_2) and ethane (C_2H_6) by volume, which provides a substantial amount of ionization (about 22.6 sites/cm for a minimum ionizing particle[45]) for a minimal amount of additional scattering. To equalize the pressure outside the modules with that inside the modules, the surrounding volume within the magnetic spectrometer is filled with pure helium. This choice minimizes scattering and energy loss for the particles coming off the target. This is controlled by a gas system consisting of several pumps, sensors, and controllers hooked up to gas reserve tanks. Since it is less damaging to leak gas from the modules into the magnet volume than vice-versa, the pressure inside the drift chamber modules is held at 1.2 Pa above the ambient pressure inside the magnet volume in case of leaks. To avoid fluctuations in the shape of the outer hood foils, this pressure difference must be maintained to within 1 Pa; in reality, it is regulated to better than 0.02 Pa.

Given the module geometry, its electromagnetic boundary conditions, and the helium-ethane gas pervading it, a map of the field lines within a cell and the resulting arrival times are simulated by the GARFIELD [46] program. Figure 4.21 displays an example Garfield simulation. Garfield is also used to estimate the magnitude of gravitational wire sag and electromagnetic deflection, which are both found to be negligible compared to the wire thickness.

4.4.3 Timing Counter

The timing counters are placed at the upstream and downstream ends on both sides of the drift chamber, within the magnetic field, as shown in Figure 4.5. Each timing counter comprises two layers of scintillating material, approximately 80 cm long, populating the

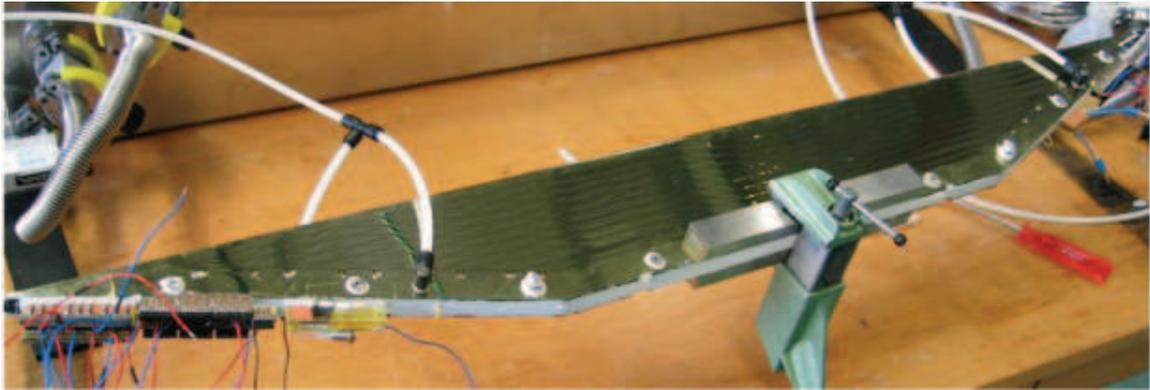


Figure 4.16: A single drift chamber module.

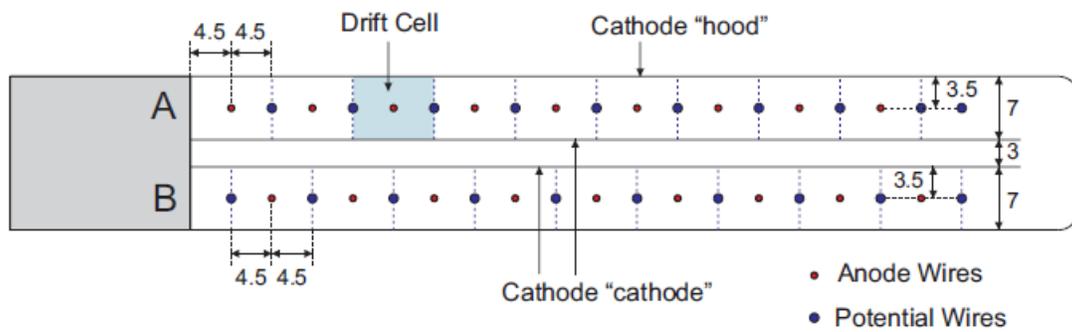


Figure 4.17: A cross-section of a drift chamber module, looking down the wires. The length measurements are in units of mm.



Figure 4.18: A picture of the three-sided frame.

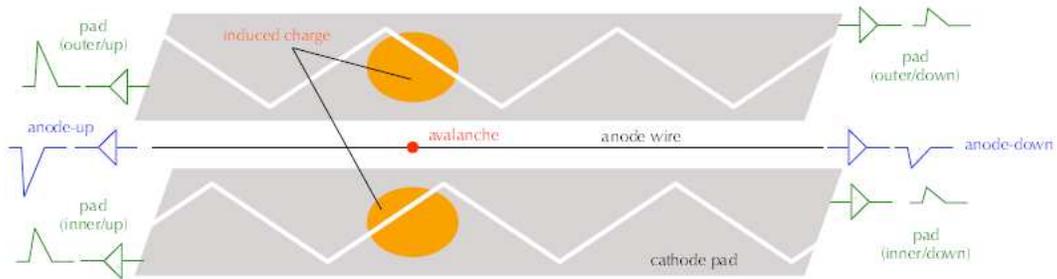


Figure 4.19: An illustration of the technique for measuring Z from pad signals.

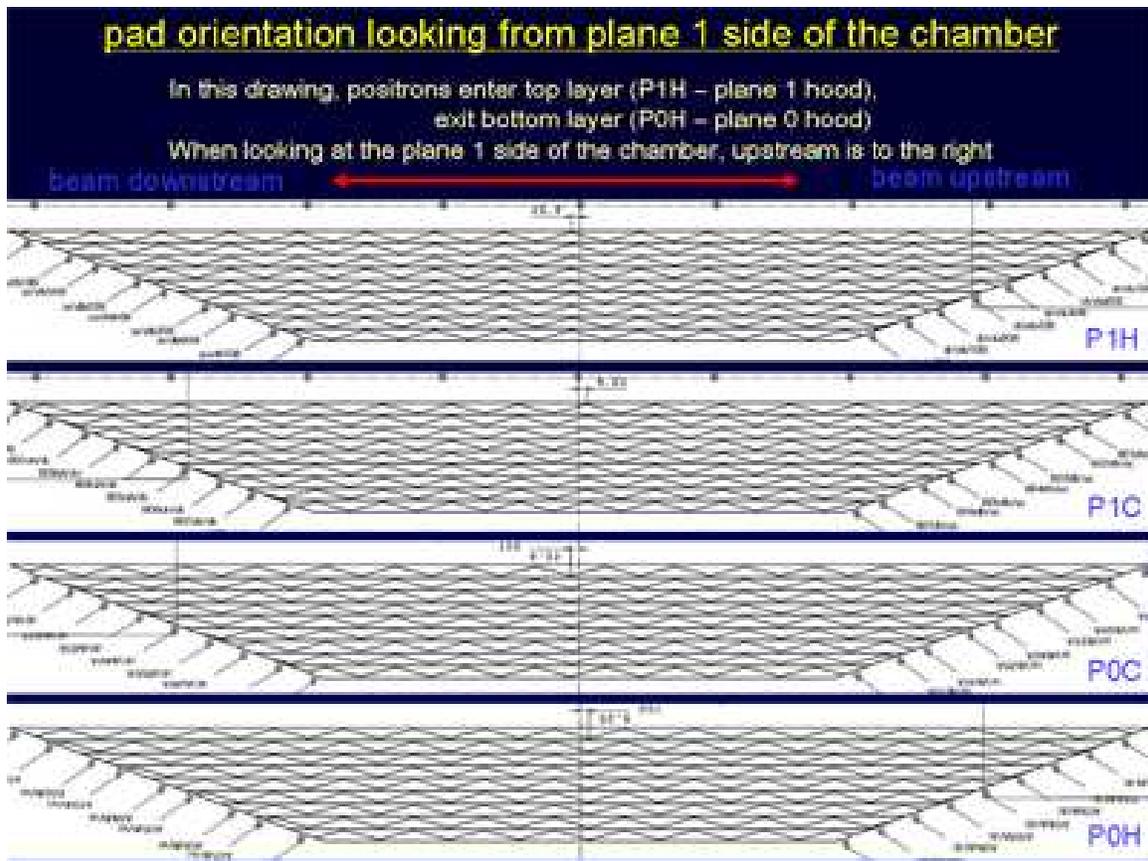


Figure 4.20: An schematic of the pad orientations.

regions $29 \text{ cm} < |z| < 109 \text{ cm}$ and covering about 160° in ϕ .

Nearest to the origin, at $R=29 \text{ cm}$, are 128 scintillating plastic fibers, made of $6 \times 6 \text{ mm}^2$ SAINT-GOBAIN BCF-20 [47], arranged at constant z . These are visible as the top layer in Figure 4.22. Both ends of a fiber are coupled to a $5 \times 5 \text{ mm}^2$ silicon avalanche photo-diode (APD), specifically a HAMAMATSU R5924[48]. APDs have the advantage of being small and easy to align with the tightly spaced fibers, as well as being insensitive to magnetic fields, unlike conventional PMTs. The fiber system is intended to provide a precise measurement of the positron impact location in Z for use in the trigger as well as the analysis. Due to technical problems with the APDs, the fibers were not used for either triggering or analysis in the 2009 data.

Behind the fibers, at $R=32 \text{ cm}$, 15 scintillating plastic bars, made of $4 \times 4 \times 80 \text{ cm}^3$ BICRON BC-404 [47], are arranged at constant ϕ in 10.5° intervals. These are visible beneath the fibers in Figure 4.22. Each end of a bar is read out by a 2-inch fine-mesh HAMAMATSU PMT S8664-55 [48], slanted with respect to the z -axis by 10° to reduce the effects of the magnetic field on their performance. In order to avoid damaging helium penetration into the PMTs, each of the two timing counter assemblies is contained in a nitrogen filled bag. The bars provide the impact location in ϕ and are also used to infer Z from the ratio of the light collected at the two ends. They are, in fact, the only source of information on impact time and location in 2009 data. The use of the timing counter bars for event reconstruction is discussed in section 5.3.

4.5 Liquid Xenon Calorimeter

The calorimeter, shown in Figure 4.23, is a portion of a cylindrical shell, within the region $x < 0$, whose inner surface is placed flush with the outer surface of the cylindrical magnet

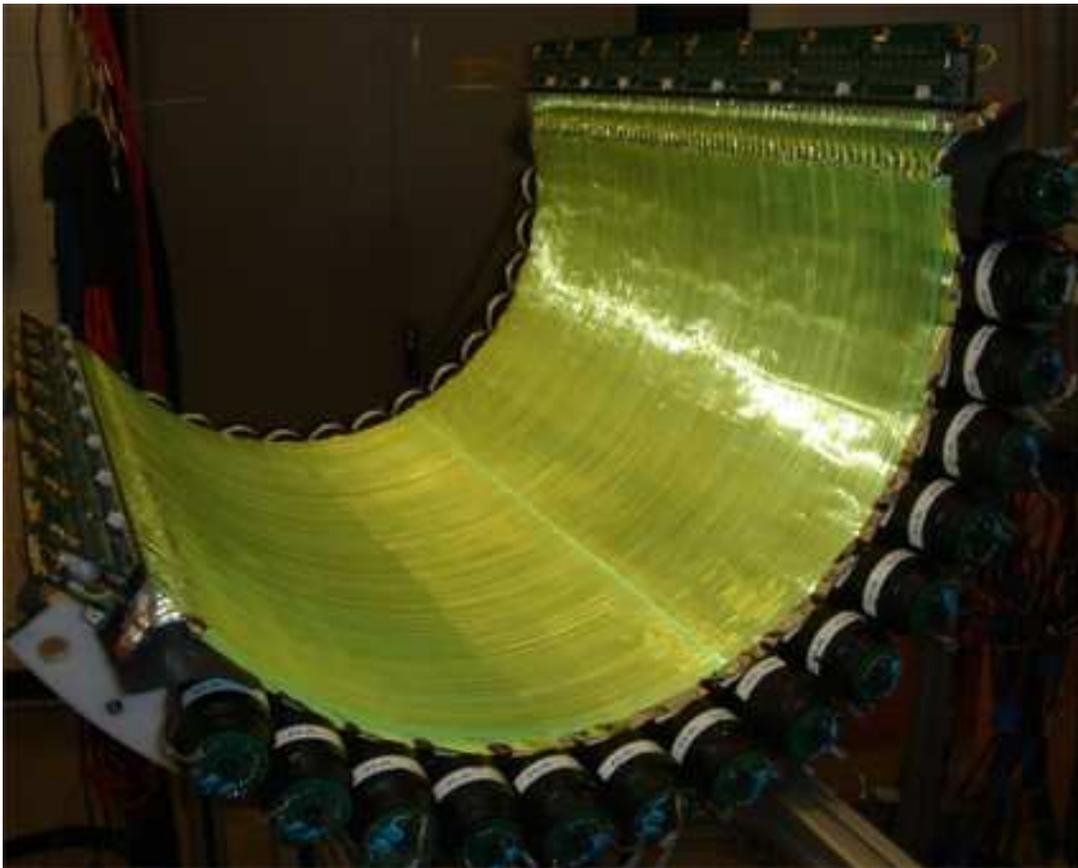


Figure 4.22: A photograph of a timing counter assembly before installation.

wall. Its volume is enclosed by six faces, labeled in Figure 4.24. Aluminum honeycomb panels, covered with carbon fiber plates, form a $0.075X_0$ -thick photon entrance window at the inner face. 900 liters of liquid xenon fill the interior of the detector as the scintillation medium, and a total of 846 PMTs are distributed on the six faces to collect the scintillation light; the ionization is not measured. PMT placement is densest on the inner face with about 35% photocathode coverage because it is the closest face to the bulk of the energy deposition, the least affected by Rayleigh scattering and absorption, and therefore the most important for locating the shower. The acceptance is matched to that of the drift chamber for back-to-back pairs of positrons and photons assuming the decay occurs at the origin, hence the solid angle coverage is a modest 11%, while the sensitive depth extends to 38.5 cm ($14X_0$), preventing energy from leaking out of the outer face.

A variety of processes influence the development of an electromagnetic shower: pair production, bremsstrahlung, photo-electric absorption, scattering, and pair annihilation. For photon energies near the signal energy (52.8 MeV), the dominant interaction is pair production. Due to the short radiation length of liquid xenon (2.77 cm), the first interaction occurs 3.56 cm from the inner face on average. As estimated from Monte Carlo simulation, the mean longitudinal extent of the shower profile is approximately 5 cm and the mean transverse spread is about 2 cm.

Two types of scintillation processes are possible. In one case, the reaction begins with an *excited* xenon atom and gives rise to a scintillation photon with a decay time of either 4.2 ns or 22 ns depending on how it proceeds [49]. In the other case, the reaction begins with an *ionized* xenon atom and produces scintillation light with a decay time of 45 ns [49]. Either way, the achievable time resolution is very good, and the emitted photon cannot be reabsorbed by xenon. One disadvantage stems from the scintillation light falling in the vacuum ultraviolet regime (around 178 nm) [50], which can be easily absorbed by water or oxygen. For this reason, the xenon is purified before data acquisition by pumping the liquid



Figure 4.23: A photograph of the calorimeter before installation.

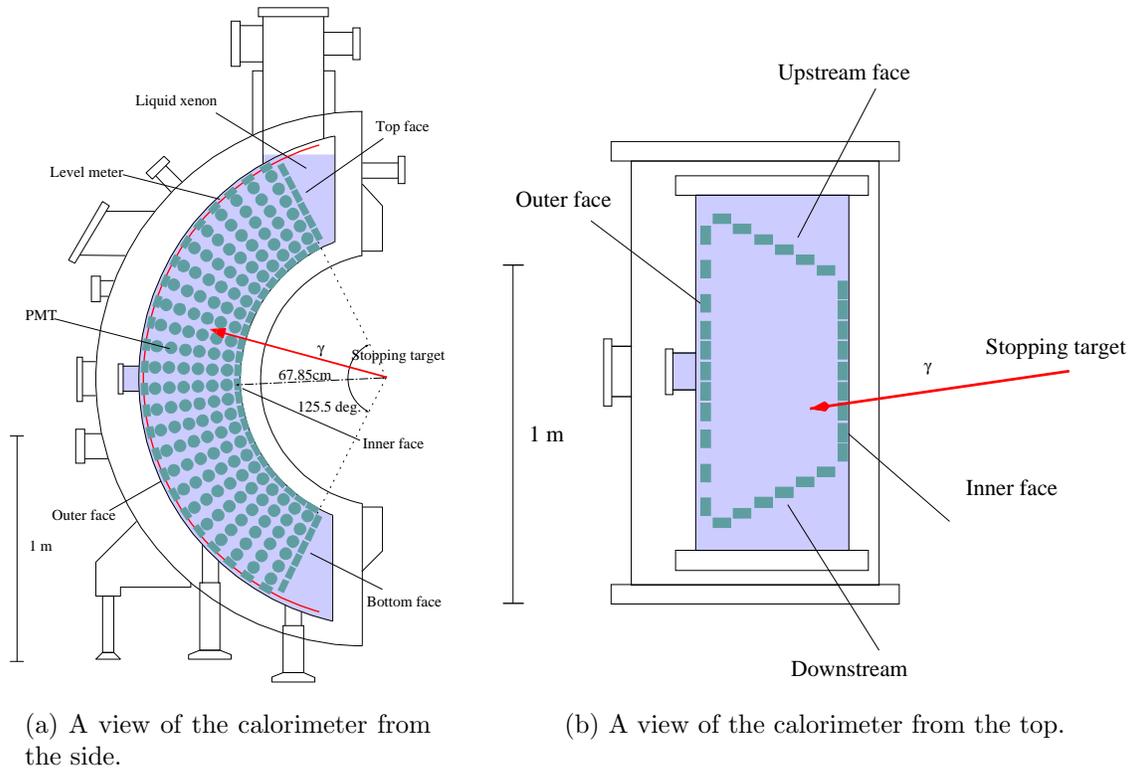


Figure 4.24: A diagram of the calorimeter labeling the six faces.

through filters, and can be purified during data acquisition by filtering evaporated xenon. A special PMT (R9869), sensitive to this light and able to function at cold temperatures, was developed by HAMAMATSU PHOTONICS specifically for MEG. Another difficulty is the additional machinery required to keep the xenon liquid (at around 165 K). The xenon is cooled by a 200 W pulse tube refrigerator mounted on top of the detector.

To help calibrate PMT gains, a dedicated system of remote-controlled, blue light-emitting diodes (LED) are installed within the active volume at six positions on both the upstream and downstream faces. Each LED is attenuated by a filter made of an aluminum sheet with holes so that they can be operated at a high voltage optimal for stable performance. During calibration, uniform illumination over each photocathode is achieved by flashing 10 LEDs simultaneously. The setup is pictured in Figure 4.25.

For calibrating the PMT quantum efficiencies, radioactive sources are placed at known po-



Figure 4.25: A photograph of the mounted LED and alpha sources. The orange circle indicates a LED and the yellow arrow indicates an alpha source wire.

sitions throughout the calorimeter and the resulting scintillation light is used. Five 100 μm -diameter, tungsten wires with five ^{241}Am sources each are mounted as shown macroscopically in Figure 4.25 and microscopically in Figure 4.26. These sources emit α particles at 5.485 MeV (84.5%) and 5.443 MeV (13%), which have a sufficiently short range in liquid xenon (about 40 μm) that they can be regarded as point-like interactions.

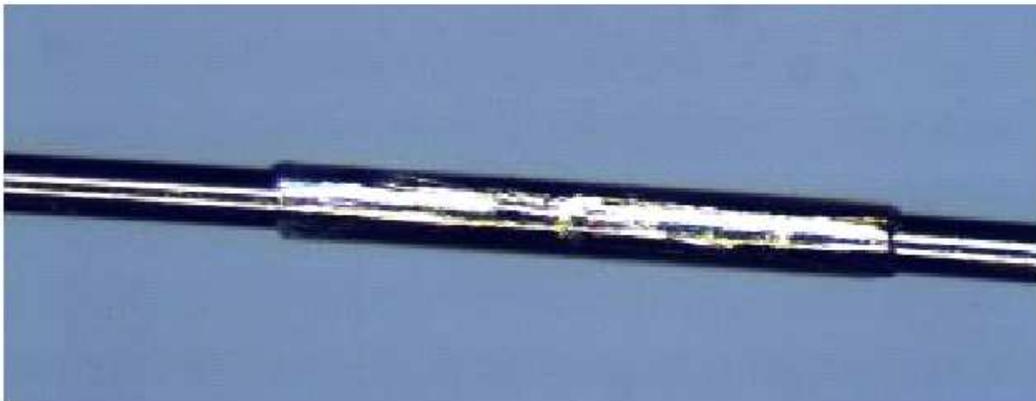


Figure 4.26: A close up of an ^{241}Am source on a wire.

The photon reconstruction is described in section 5.1.

4.6 Trigger and Data Acquisition

When the experiment is live, thousands of signals from the detectors are routed to waveform digitizers to store the waveforms, sent to a trigger system that decides if the event is of interest or not, and permanently logged in software file format if the trigger conditions are satisfied. The essential components of this system are shown in Figure 4.27.

PMT signals from the calorimeter are sent to active splitters through coaxial cables. There they are split into three outputs, all of which are inverted into a positive pulse to lie in the dynamic range of the waveform digitizers. One is a wide-band (1.9 GHz) fully differential output that is delivered to a waveform digitizer by a 2 m-long high-density twisted-pair

cable. The other outputs are a 320 MHz-bandwidth differential output, which is used for PMTs on the inner face, and a four-to-one sum output, which is used for other PMTs; both are fed to the trigger by a 2 m-long lower-density cable.

Signals from the PMTs associated with the timing counter bars are sent to a passive splitter where they are divided into three outputs in an 8:1:1 ratio. The largest output is passed through a double threshold discriminator (DTD) to distinguish real positron hits from spurious ones, yielding 50 ns-wide standard NIM pulses that go to waveform digitizers. One of the smaller outputs exiting the passive splitter is sent to an active splitter where one output is coupled to the waveform digitizer and the other is coupled to the trigger system. The remaining passive splitter output is used for online current monitoring.

The anode wire signals from the drift chamber are split in the ratio 9:1 by a passive splitter. The largest output goes to the waveform digitizer while the other is amplified, summed over several wires, and then used in the trigger. All signals induced on the pads are sent to the waveform digitizer.

4.6.1 Data Acquisition with MIDAS

Data acquisition is controlled by the MIDAS [51] framework. The full output of the experiment is divided into nine crates, each controlled by a Linux PC. Four of them are connected to the trigger system, and five are connected to the waveform digitizers. MIDAS controls all of the crates, an online database of parameters for data-taking and triggering, an alarm system, a live monitoring system, and a system for logging events to disk.

Several techniques are applied to reduce the data size. Only drift chamber waveforms consistent with a pulse above the noise on an anode are recorded, and the waveforms are then re-binned outside the region of interest for coarser granularity. The timing counter

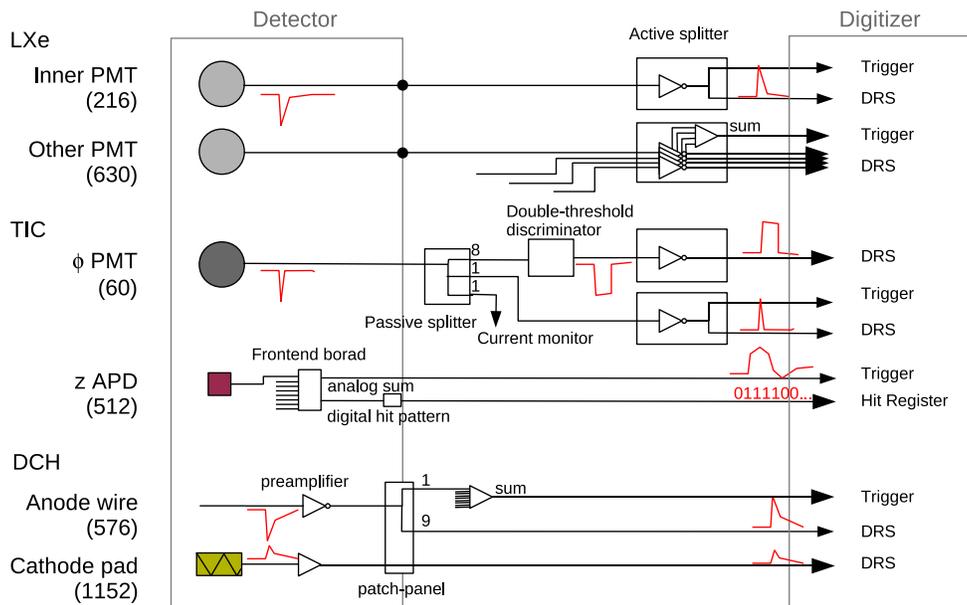


Figure 4.27: A visual representation of the data acquisition system.

waveforms are similarly zero-suppressed and re-binned. The calorimeter waveforms are not zero-suppressed, but they are also re-binned outside the region of interest. This results in a total reduction in data size by a factor of greater than 3. In total, 93 TB of data was collected during 2009.

4.6.2 Waveform Digitization with DRS

All waveform digitization in the MEG experiment is performed by a PSI-designed switched capacitor array named the DRS (Domino Ring Sampler). A single DRS chip contains eight sampling channels, a channel to accept trigger signals, and a channel to accept a clock signal. Each of the eight sampling channels can store waveforms in a ring of 1024 capacitors. During operation, the DRS is continuously storing information in the 1024 sampling cells at a fixed frequency until a trigger signal causes it to stop and the sampled waveform is read out by a shift register and digitized by a commercial FADC (fast analog-to-digital converter). This

process is summarized in Figure 4.28.

Throughout the 2009 run, all detectors used version DRS4 [52] except the timing counter, which used version DRS3 [53]. The calorimeter and timing counter waveforms were sampled at 1.6 GHz for fine granularity in discerning pileup, while the drift chamber waveforms were sampled at 0.8 GHz because the drift processes happen on a slower time scale with typical rise times of ~ 10 ns.

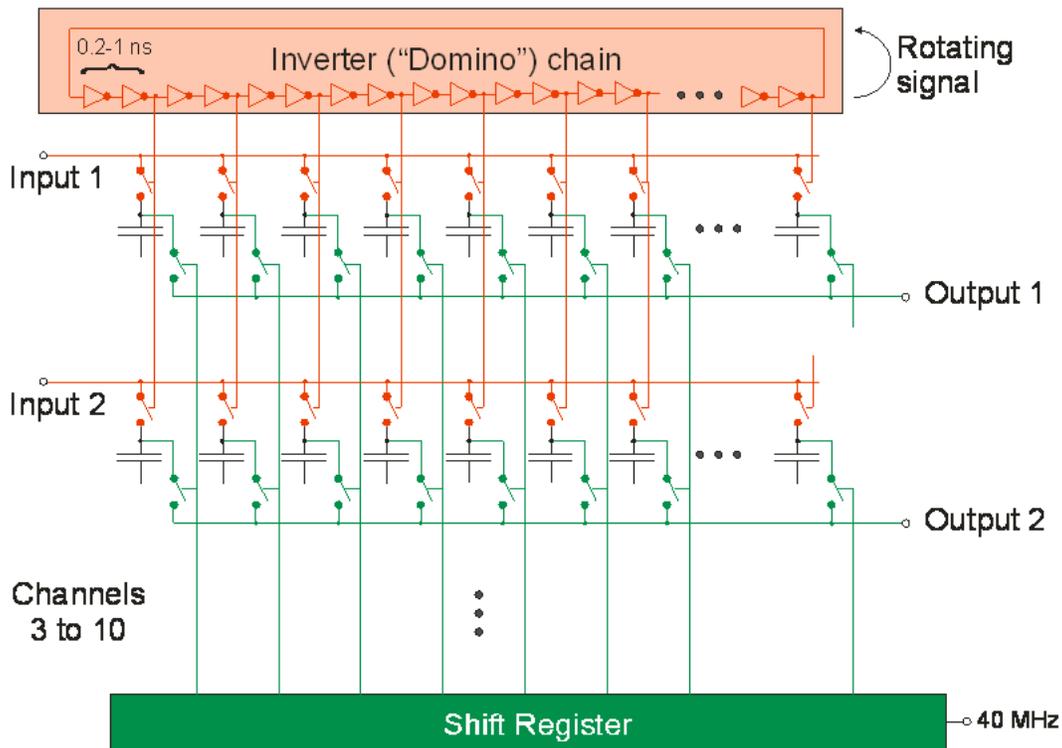


Figure 4.28: A diagram representing the technique for digitizing waveforms.

4.6.3 Trigger System

The trigger system consists of flash analog to digital converters (FADC)[54], which digitize the waveforms received from the detectors, and field programmable gate arrays (FPGA)[55], which analyze the digitized waveforms. The chain of command is displayed in Figure 4.29.

At the first stage of triggering, Type 1 boards digitize analog input signals with FADCs

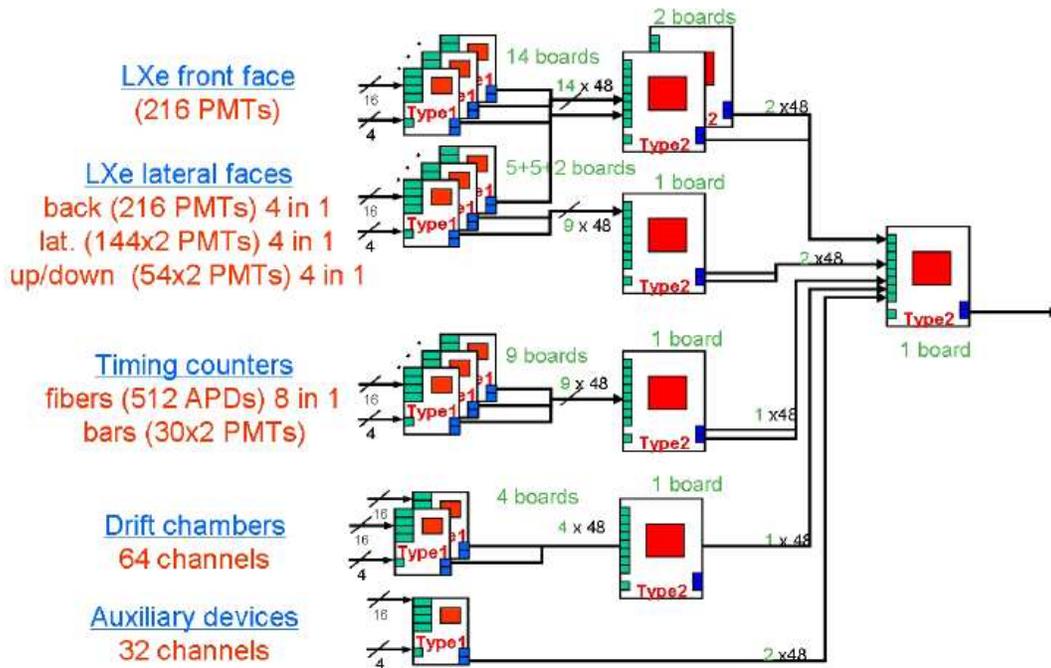


Figure 4.29: A pictorial layout of the trigger system.

(AD9218 [54]) at 100 MHz, carry out some fast reconstruction algorithms on FPGAs (Xilinx Virtex-II Pro [55]), and pass the results onward. The remaining triggering stages make use of Type 2 boards. The second triggering stage figures out the state of each detector, and, at the third and final stage, a decision is made. If an event satisfies the trigger, then all of the waveform digitizers are stopped and read out.

A master board contains a 19.44 MHz reference square pulse generator (SaRonix SEL3935 [56]) and accepts decision commands from the third stage of the trigger system. The clock signals are distributed to all boards of the trigger and digitizer systems to guarantee that they are temporally synchronized.

The MEG event trigger requires a photon energy above threshold, and a space-time match between the positron and photon. The online algorithm estimates the photon energy by a global sum of calorimeter PMT waveforms, which achieves an 8% FWHM resolution, and the threshold is set to 44 MeV. Both the positron and photon emission times are taken from

a linear interpolation of PMT pulse leading edges; neither is corrected for flight time from the target. The resolution of the positron photon time difference is 3.3 ns (σ), and the MEG trigger requires this to be within 10 ns of zero. To check for a direction match, the photon conversion position is given by the most illuminated inner face PMT and the positron z impact location is taken from the charge ratio of PMTs in the hit bar. The positron path from the target to the timing counter is complicated due to the non-trivial magnetic field, thus, for a fast online determination of the positron emission angle, a lookup-table is created from a Monte Carlo simulation of the correlation between target emission angle and timing counter impact location for signal positrons. The trigger rate for events satisfying these conditions during the 2009 data taking period was 9 Hz.

A variety of different triggers are taken with high prescale factors along with the MEG data for purposes such as normalization and detector efficiency measurements.

4.7 Calibration Hardware

Additional apparatus is used to calibrate the photon energy scale; measure the photon energy, time, and position resolutions; and calibrate the calorimeter-timing counter relative timing. Back-to-back photon pairs from π^0 decay, as described in section 4.7.1, are used to calibrate the photon energy scale and measure the photon energy, time, and position resolutions. The photon energy scale is monitored during data taking using monochromatic photons from proton reactions with Lithium and Boron, as described in section 4.7.2, and the calorimeter-timing counter relative timing is calibrated with back-to-back photons from one of those reactions.

4.7.1 Charge Exchange Apparatus

The reconstruction algorithm eventually calculates the number of scintillation photons observed in the calorimeter for a given event. This light level is converted into a photon energy using a conversion factor determined from calibration. A monochromatic source of photons near the signal energy is used to calibrate the absolute energy scale. The monochromatic source is obtained by using the reaction $\pi^- p \rightarrow \pi^0 n$ at rest followed by $\pi^0 \rightarrow \gamma\gamma$ decay and selecting events where the two photons are in anti-parallel directions. As illustrated in Figure 4.30, the π^0 momentum is fixed at 28 MeV/c. When the photons are anti-parallel along the π^0 flight direction, they have energies of 54.9 and 82.9 MeV to conserve energy and momentum. The 55 MeV photons are close to the signal energy of 52.8 MeV. Dalitz decays, $\pi^0 \rightarrow e^+ e^- \gamma$, are also recorded for other uses. Additionally, the radiative capture reaction, $\pi^- p \rightarrow \gamma n$, can occur giving an 8.9 MeV neutron and a 129.4 MeV photon. Since it takes about a week to prepare the setup for this calibration, described below, it was only performed once, just before the 2009 data taking.

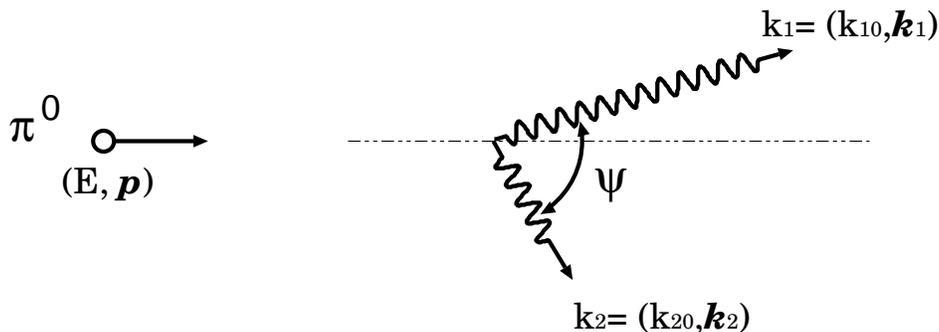


Figure 4.30: Kinematics of the $\pi^0 \rightarrow \gamma\gamma$ decay.

Beam and Target Setup The $\pi E5$ channel provides a 70.5 MeV/c π^- beam, whose momentum is selected as that most effective for filtering beam contaminants. During charge exchange data taking, the traditional muon target is replaced with a cylindrical target filled with about 150 cc of liquid hydrogen, pictured in Figure 4.31. It is 50 mm in diameter, 75

mm in length, and has a $135\ \mu\text{m}$ thick Mylar film at the beam entrance. The charged pions come to rest in the target and the outgoing neutrons and neutral pions from the charge exchange reaction are emitted back-to-back, each with $28\ \text{MeV}/c$ of momentum.



Figure 4.31: A photograph of the liquid hydrogen target used with the π^- beam.

NaI Detector To select events with back-to-back photons, a miniature calorimeter, pictured in Figure 4.32a, is placed opposite the liquid xenon calorimeter. For the energy measurement, nine NaI crystals ($62.5 \times 62.5 \times 300.5\ \text{mm}^3$) comprise the scintillation medium and each is coupled to an APD ($10 \times 10\ \text{mm}^2$, HAMAMATSU S8664-1010 [48]). To measure the time, a 5 mm-thick lead plate and two plastic scintillators are placed in front of the central NaI crystal. The e^+e^- pairs produced in the lead plate are collected in the two plastic scintillating counters ($70 \times 70 \times 7\ \text{mm}^3$), each coupled to two fine-mesh PMTs (HAMAMATSU, H6152-70 [48]). The entire assembly is mounted on a movable platform, pictured in Figure 4.32b, which allows for full coverage of the liquid xenon calorimeter acceptance when selecting back-to-back pairs.

Details on the use of the charge exchange data for measuring photon energy, time, and

position resolutions are presented in section 8.1.1.



(a) The NaI calorimeter, consisting of a lead plate, plastic scintillating bars, and NaI crystals.



(b) The movable platform for positioning the NaI calorimeter.

Figure 4.32: The setup for the external NaI calorimeter.

4.7.2 Cockcroft-Walton Accelerator

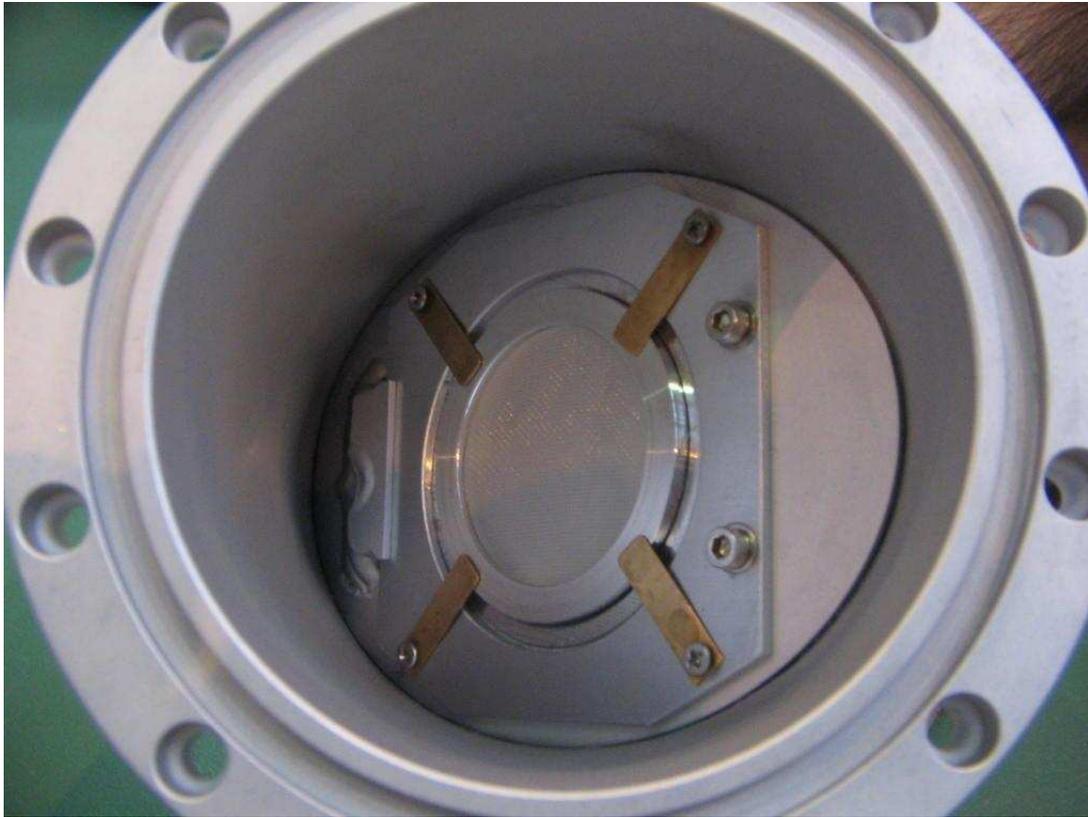
A Cockcroft-Walton proton accelerator[57], pictured in Figure 4.33, is used to provide photons at several different fixed energies to check the calorimeter linearity. Photon pairs are also produced and used to calibrate the relative time offset between the calorimeter and the timing counter. One reaction, $p + {}^7_3\text{Li} \rightarrow {}^8_4\text{Be} + \gamma$, produces a 17.6 MeV photon and has a high cross-section at low proton energies (resonant at 440 keV)[58]. Another reaction, $p + {}^{11}_5\text{B} \rightarrow {}^{12}_6\text{C} + \gamma$, which is resonant at 163 keV, produces a 16.1 MeV photon when the ${}^{12}_6\text{C}$ is in its ground state, or an 11.7 MeV photon when the ${}^{12}_6\text{C}$ is in its first excited state, as well as a simultaneous 4.44 MeV photon from the transition to the ground state.

During calibration runs, the Cockcroft-Walton supplies a 10^{12} Hz proton beam to a lithium tetraborate target ($\text{Li}_2\text{B}_4\text{O}_7$) target, pictured in Figure 4.34a, catalyzing both of the above

reactions with a single target. A quick exchange between the muon stopping target and the lithium tetraborate target, requiring only ~ 20 minutes, is performed remotely with a bellows system, pictured in Figure 4.34b. Daily Cockcroft-Walton calibration runs are used to monitor the calorimeter light yield throughout the MEG data taking.



Figure 4.33: A photograph of the Cockcroft-Walton accelerator.



(a) A photograph of the $Li_2B_4O_7$ target for use with the Cockcroft-Walton proton accelerator.



(b) The bellows system for insertion and extraction of the $Li_2B_4O_7$ target.

Figure 4.34: Components of the target setup for calibration with the Cockcroft-Walton.

Chapter 5

Event Reconstruction

An event recorded by the experiment consists of raw detector waveforms that are used to reconstruct the kinematics of the detected particles. Photons are detected in the liquid xenon calorimeter (XEC), and positrons are detected in the drift chamber (DCH) and timing counter (TIC). This chapter describes the algorithms used to reconstruct the energy, time, and angle of the photon and positron.

5.1 XEC

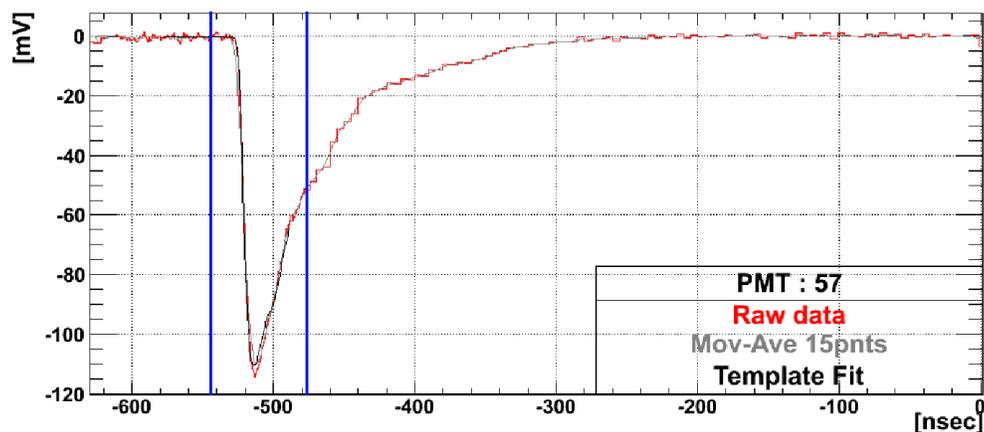
Photons entering the calorimeter deposit their energy inside its volume, giving off characteristic patterns of scintillation photons. The ensuing series of waveforms measured by the PMTs are the exclusive source of information on the photon energy, time, and position. A set of software algorithms first reconstructs the number of detected scintillation photons and their arrival time in each PMT, and from these PMT measurements infers the energy, time, and first conversion position of the detected photon.

5.1.1 Waveform Analysis

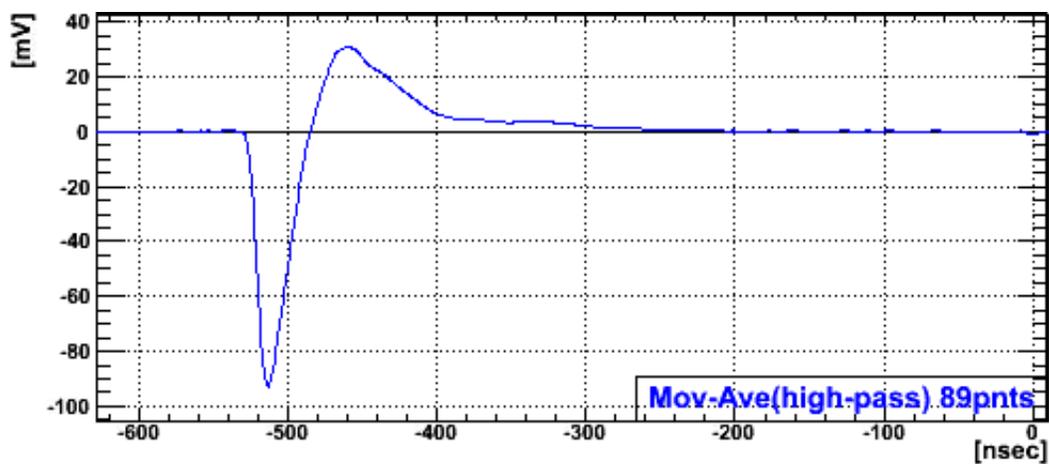
The analysis begins by finding the leading edge time and charge of each PMT waveform. The leading edge times of the raw PMT waveforms are estimated by a constant fraction discriminator, namely, the time at which the waveform reaches 30% of its full pulse height. To reduce the effects of waveform noise and stochastic pulse shape fluctuations, the pulse height is measured indirectly from the charge by assuming a linear scaling relationship between charge and pulse height, which is inferred from a template waveform shape that is constructed from the accumulation of many events.

The charge integration proceeds by first averaging the voltages in a region before the pulse on an event-by-event, and PMT-by-PMT basis to form a constant estimator of the baseline during the pulse. A high-pass filter is constructed by subtracting a moving average of the waveform (itself a low-pass filtered waveform) from the original waveform. The number of bins in the moving average, 89, is tuned to achieve a cut-off frequency of approximately 11 MHz. Figure 5.1 shows an example of a raw and high-pass filtered waveform. A common charge integration region is applied to all PMTs in an event. The end of the integration region is determined by the zero-crossing time of the sum of all PMT waveforms in the event. The length of the integration range is 48 ns, which acts as a low-pass filter with a 21 MHz cut-off frequency.

The energy estimator is an appropriately weighted sum of all PMT charges in the event. The performance of this estimator is limited by PMTs receiving signals large enough to saturate the electronics. With about 15% of photon interactions occurring within 1 cm of the front face, the probability for an inner face PMT to saturate is significant. Figure 5.2 displays a saturated waveform. An improved estimator of the charge in these situations is actualized through a *time over threshold* technique. The time span during which the waveform is above some threshold, 150 mV, is related to the expected charge based on the assumption of the



(a) The red line shows a raw PMT waveform. The grey line is a low-pass filtered version of the waveform, constructed from a moving average. The black line is the template waveform with its amplitude scaled by the measured charge and its leading edge aligned with the measured one.



(b) A high-pass filtered version of the same waveform.

Figure 5.1: Various representations of a PMT waveform from data.

fixed template waveform shape discussed above. Figure 5.3 shows this relation. In cases of PMT saturation, the charge is calculated directly from the time over threshold using this relation.

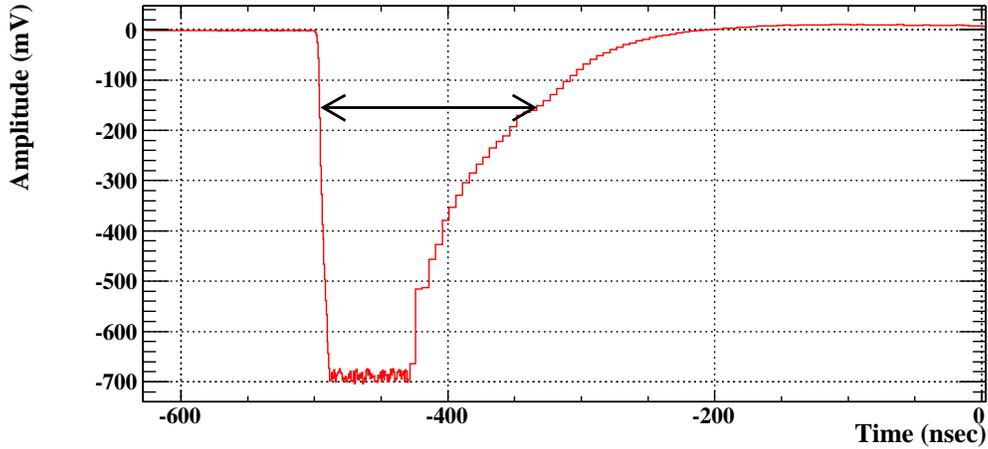


Figure 5.2: An example of a saturated PMT waveform. The time over threshold is pointed out by the black arrow.

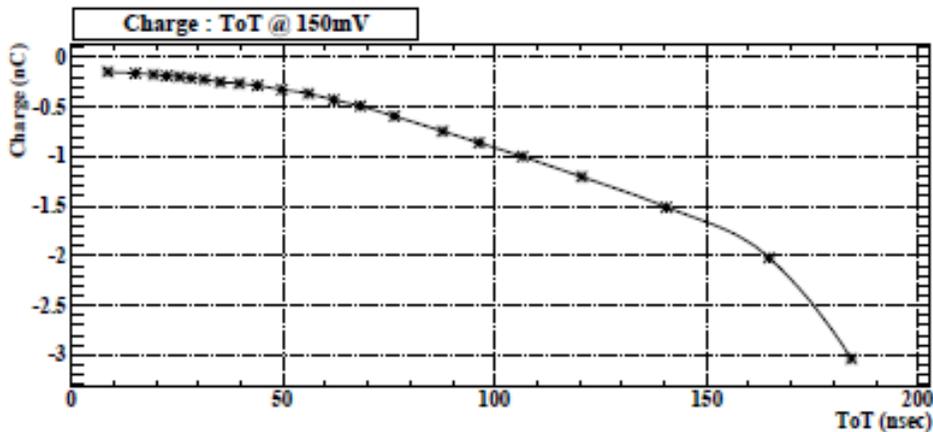


Figure 5.3: The relationship between expected charge and time over threshold based on the template waveform.

5.1.2 Photon Reconstruction

The number of photoelectrons ($N_{pe,i}$) and scintillation photons ($N_{pho,i}$) detected by the i 'th PMT are inferred from its measured charge (Q_i), its electronic gain (G_i), the quantum

efficiency of its photocathode (QE_i), and the elementary charge, e :

$$N_{pe,i} = \frac{Q_i}{(e \cdot G_i)} \quad (5.1)$$

$$N_{pho,i} = \frac{N_{pe,i}}{QE_i}. \quad (5.2)$$

Sections 6.3.1 and 6.3.2 discuss the determination of the PMT gains (G_i) and quantum efficiencies (QE_i), respectively. With this information, the analysis proceeds to reconstruct the shower position, time, and energy.

Position To reconstruct its position, the shower is first treated as a point-like interaction that emits scintillation photons isotropically. At each one of a series of test locations $(x_\gamma, y_\gamma, z_\gamma)$, the solid angle subtended by the i 'th PMT ($\Omega_i(x_\gamma, y_\gamma, z_\gamma)$) is computed numerically. Both the photon position $(x_\gamma, y_\gamma, z_\gamma)$ and a factor (c) are varied to minimize:

$$\chi_{pos}^2 = \sum_i \frac{N_{pho,i} - c \times \Omega_i(x_\gamma, y_\gamma, z_\gamma)}{\sigma_{pho,i}(N_{pho,i})}. \quad (5.3)$$

The measurement uncertainty is that due to Poisson fluctuations in the number of observed photoelectrons:

$$\sigma_{pho,i}(N_{pho,i}) = \frac{\sigma_{pe,i}(N_{pe,i})}{QE_i} = \frac{\sqrt{N_{pe,i}}}{QE_i} = \sqrt{\frac{N_{pho,i}}{QE_i}}. \quad (5.4)$$

The position fit is actually performed twice to smooth the effects of non-point like shower development. In each case, only inner face PMTs are used since they are the closest to the bulk of the energy deposition and thus the least sensitive to Rayleigh scattering and absorption. The starting point is a charge-weighted mean using PMTs around the PMT with maximum signal. In the first fit, all PMTs whose center lies within a 3.5-PMT radius

of the starting point are used, typically 45. In the final fit, all PMTs whose center lies within a 2-PMT radius around the position from the first fit are used, typically 15. The final fitted position can be taken into the (u,v,w) coordinate system by the transformation equations:

$$u = z \tag{5.5}$$

$$v = r_0 \arctan\left(\frac{-y}{x}\right) \tag{5.6}$$

$$w = \sqrt{x^2 + y^2} - r_0, \tag{5.7}$$

where $r_0 = 67.85$ cm, the radius of the inner face.

u is just the usual coordinate along the beam axis, v is a coordinate along the circumference of the inner face, and w measures the radial depth inside the calorimeter volume.

Based on MC studies, no bias is observed in v , while $|u|$ and w tend to preferentially reconstruct toward larger values than the true position. Consequently, MC deduced corrections are applied to u and w . Additional corrections for fluctuations in the shape of the shower are administered, predicated on the information contained in the difference of the two position fit results.

A discussion of the photon position resolution is given in section 8.1.1.

Energy The algorithm for photon energy reconstruction is based on the principle that the photocathode coverage is uniform in a given face and approximately independent of shower position. In that case, the deposited photon energy is proportional to the weighted sum of all detected scintillation photons:

$$N_{sum} = \sum_i w_i \times N_{pho,i}. \tag{5.8}$$

w_i is the inverse photocathode coverage of the i 'th PMT. This approach works well when the first interaction occurs at a significant depth, $w > 2 - 3$ cm. For shallower conversions, the solid angle subtended by PMTs on the inner face varies strongly with shower position and the assumption of uniform coverage on the inner face is violated. Monochromatic 55 MeV photons from π^0 decay are used to study the dependence of N_{sum} on the solid angle subtended by the inner face PMT with maximum signal, $\Omega_{in,max}$, for shallow depth ($w < 3$ cm). The results of this study are used to correct the shower position dependence of N_{sum} for shallow events. After the correction for solid angle dependence, a residual position dependence of energy is observed due to finite absorption length and differences in the effective coverage of PMTs with conversion position. A position-dependent correction factor, $F(u, v, w)$, is estimated from the 17.6 MeV photon energy peaks in the $p + Li$ data. Finally, the photon energy is given by:

$$E_\gamma = \alpha F(u, v, w) N_{sum} \quad (5.9)$$

The proportionality constant α is obtained from 55 MeV calibration photons. The degree to which α is constant with energy is verified with several monochromatic sources of photons as shown in figure 5.4.

A discussion of the photon energy resolution is given in section 8.1.1.

Time Each PMT gives an independent measurement of the first interaction time:

$$t_{interaction,i} = t_{PMT,i} - \Delta t_{delay,i} - \Delta t_{offset,i}. \quad (5.10)$$

The first term, $t_{PMT,i}$, is the leading edge time of the i 'th PMT waveform. Event dependent time delays are encoded in the second term, and constant electronic delays are represented by the last term.

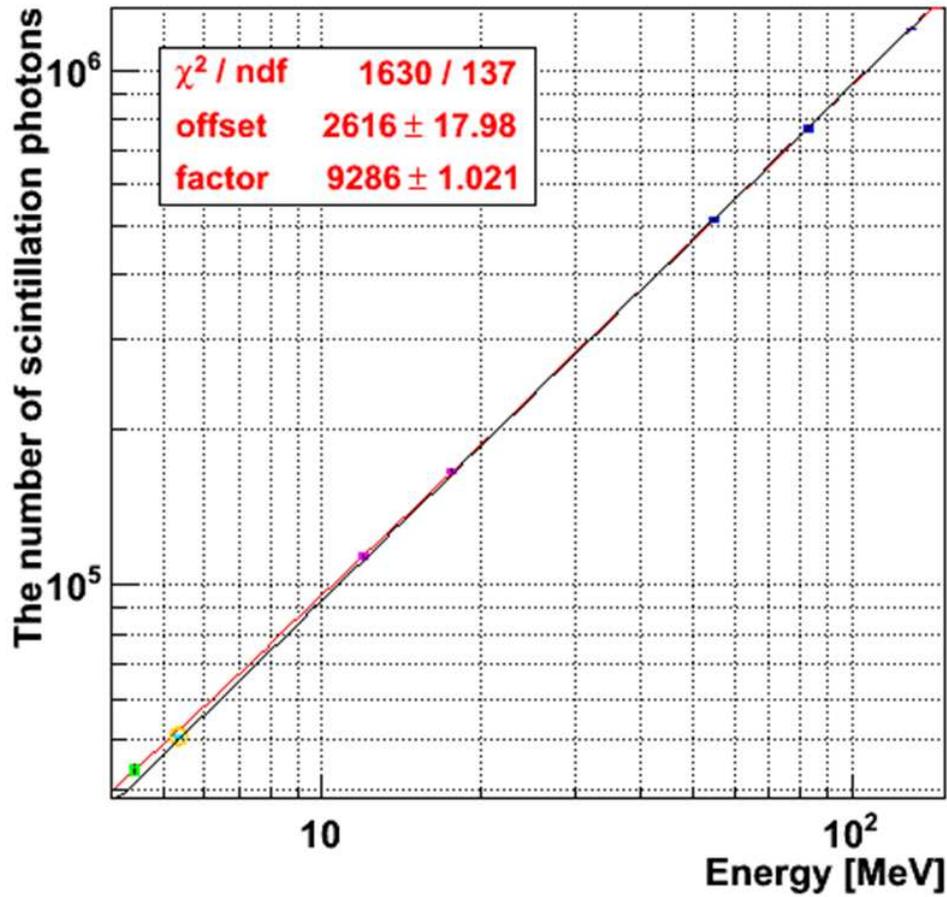


Figure 5.4: The 5.5 MeV data point is from the in-situ α sources; the 4.4 MeV, 11.7 MeV, and 17.6 MeV data points come from $p + Li$ and $p + B$ data; the 55 MeV and 83 MeV data points come from charge exchange data, and the 129 MeV line is from the radiative π^- capture reaction. The red line is a fit to a line with a floating y-intercept. The blue line is a fit to a line with the y-intercept fixed to zero.

The event dependent piece is composed of three parts:

$$\Delta t_{delay} = \Delta t_{direct}(d, v_{eff}) + \Delta t_{scatter}(\eta) + \Delta t_{walk}(N_{pe}). \quad (5.11)$$

Here, the first term encapsulates the contribution from scintillation photons propagating directly from the source to the PMT for a finite amount of time. It depends on the distance between source and detector, d , and the speed of scintillation photons in liquid xenon, v_{eff} . Photons reaching a PMT after reflection off walls or scattering in the xenon comprise the second term. This delay depends on the incident angle of incoming photons with respect to the normal to the photocathode (η); this part becomes severe when the angle is far from zero. The last term describes time walk effects in getting the leading edge of the PMT waveform. This is a function of the pulse height, or equivalently, the number of observed photoelectrons, N_{pe} . These corrections are determined phenomenologically as discussed in section 6.3.3.

The best estimate of the first interaction time, t_{XEC} , is then the one that minimizes:

$$\chi_{time}^2 = \sum_i \frac{(t_{interaction,i} - t_{XEC})^2}{\sigma_{t,i}(N_{pe,i})^2}. \quad (5.12)$$

The measurement uncertainty, $\sigma_{t,i}(N_{pe,i})$, is the measured i 'th PMT time resolution as a function of N_{pe} . The sum is over PMTs that satisfy $N_{pe} > 50$, typically about 150 of them.

A discussion of the photon timing resolution is given in section 8.1.1.

Pileup Identification At the high muon rates necessary to achieve the desired sensitivity, the calorimeter reconstruction is especially susceptible to events containing multiple photons. Such pileup events are identified spatially by the light distribution in both the inner face and outer face PMTs, and temporally by the distribution in the PMT leading edge times.

One characteristic signature of pileup is an unusually large value of χ_{time}^2 divided by the number of degrees of freedom, $\hat{\chi}_{time}^2$, which occurs when different groups of PMTs observe distinct interaction times of different photons. The threshold for pileup recognition is $\hat{\chi}_{time}^2 > 3$, as deduced from the distribution in that variable from data and MC (figure 5.5).

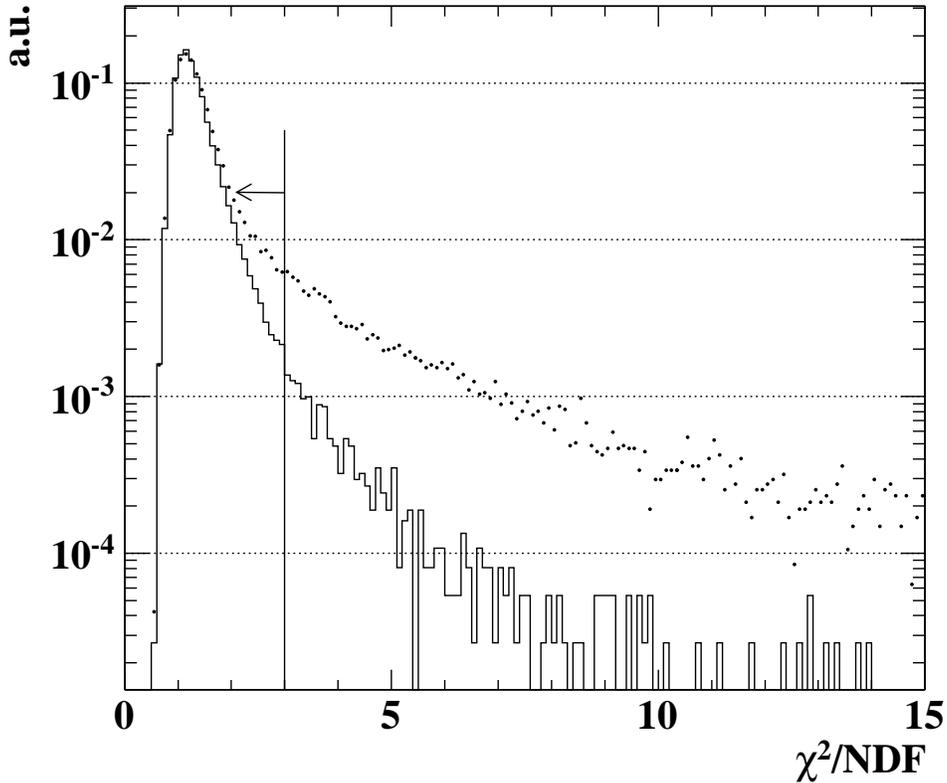


Figure 5.5: Distribution of $\hat{\chi}_{time}^2$ on data (dotted line) and signal MC without pileup (solid line).

Another conspicuous indication of pileup is the presence of more than one peak in the PMTs on either the inner or outer face. Secondary peaks are distinguished by locating PMTs with $N_{pho} > 200$ on either face at positions isolated from the primary peak.

Pileup events are recovered by first removing PMTs near the location of one photon when reconstructing another photon in the same event. This introduces the complication that a precise energy determination still requires a better estimator than $N_{pho} = 0$ in the affected

PMTs. For this purpose, a table of average PMT N_{pho} yields for all possible shower positions using a $(1.55 \times 1.55 \times 1.55)$ cm grid is prepared using the 17.6 MeV photons from $p + Li$ data. The photon energy can then be estimated by replacing the outputs of those PMTs with expected yields from the tables. Pileup events that are identified by $\hat{\chi}_{time}^2$ but not by spatially separated peaks are not recovered.

5.2 DCH

A charged particle traversing the drift chamber produces primary ionization along its path in the sensitive volume of, typically, one or two cell(s) per intercepted plane. The waveform information associated with the resulting avalanches is used to calculate the particle's position at each plane and ultimately the particle momentum, a decay vertex and set of emission angles from the projection to the target, a projected impact location at the timing counter, and the overall path length of the trajectory. A chain of software algorithms first measure the arrival time and charge on each wire and pad associated with the passage of a charged particle through a cell, which are converted into spatial coordinates. The information from each cell, referred to as a hit, is cross-checked with other hits on the same chamber, and groups of hits consistent with coming from the same particle are then collected into clusters. Patterns of clusters consistent with coming from the same particle are then collected into tracks. Finally, a fit to the global track trajectory is made to estimate the track momentum, the interception at the timing counter and target plane, and the total path length from the target to the timing counter.

5.2.1 Hit Reconstruction

This process begins by identifying hits. A hit refers to the information from a cell associated with the passage of a single charged particle through its sensitive volume.

At every event and for each waveform, the algorithm forms a binned distribution of voltage measurements (with 1 mV bin widths) from the region preceding the nominal location of minimum drift time hits associated with triggering particles. A voltage baseline is computed by averaging all entries in the binned voltage distribution within ± 5 mV of its mode, additionally excluding 140 bins after any 12 consecutive bins exceeding the mode voltage by 5 mV or more. This procedure reduces contamination from early hits and electronic noise fluctuations in determining the baseline. The baseline is subtracted from the waveform for the remainder of the analysis.

Each anode waveform is then searched for hits in a region near the nominal triggered event time. Such a window should be large enough to include maximum drift time hits and allow for some jitter in the event time, yet small enough to avoid including excess hits with no pertinence to the trigger (from 80 ns before to 350 ns after the nominal event time). To reduce high-frequency noise sensitivity, a low-pass filter is implemented by means of a moving average of voltages with a period of 5 bins. This is effective at removing noise above the cutoff frequency of approximately 160 MHz. A candidate hit is generated upon detecting a voltage maximum above 5 mV in both smoothed anode waveforms of the same wire. Figure 5.6 displays waveforms associated with a hit.

Proceeding backwards in time from the maximum voltage, the earliest bin time whose voltage still exceeds three times the bin-to-bin pedestal RMS ($\sigma_{pedestal}$) defines the leading edge of the pulse. The leading edge of the end with largest maximum amplitude is assigned as the absolute hit arrival time in order to reduce errors due to pulse-height slewing and noise. Once the Z coordinate of the hit is known, the hit time is corrected for propagation time

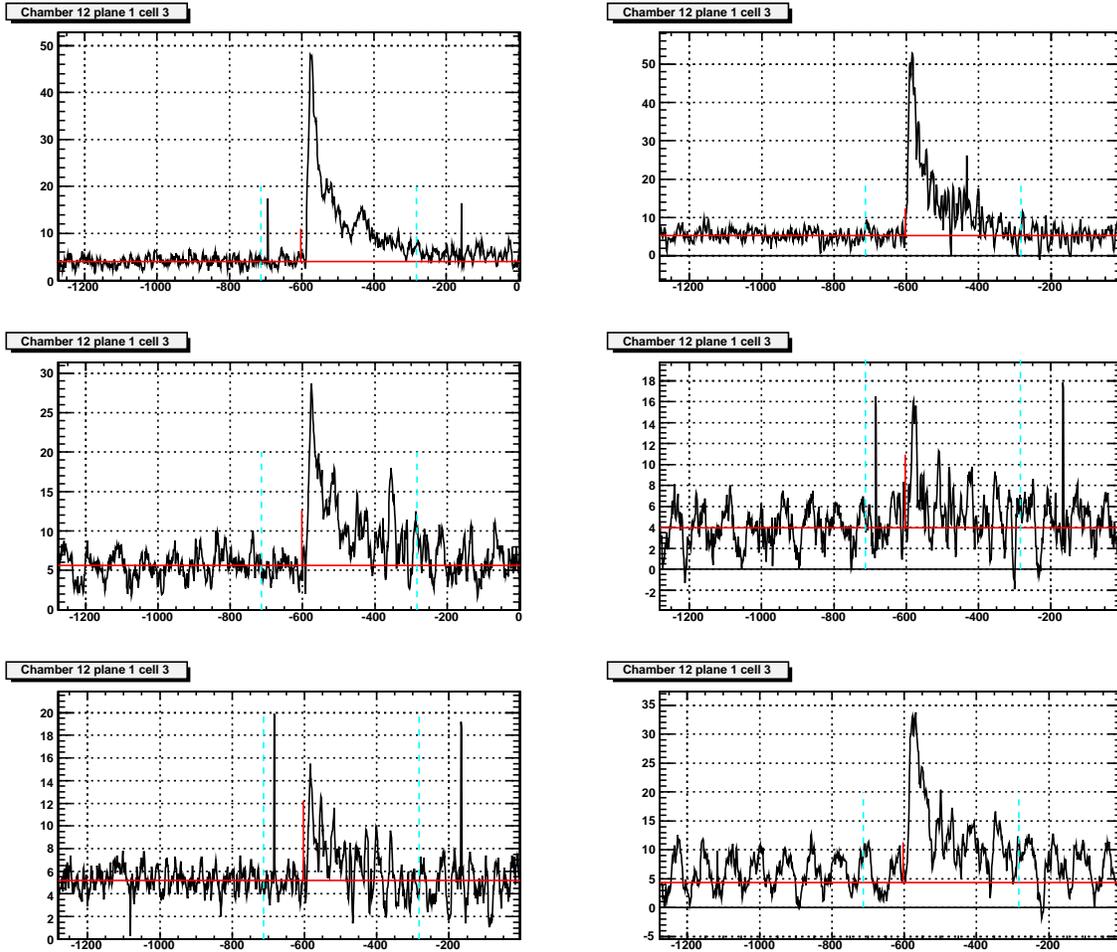


Figure 5.6: The six waveforms associated with a hit. On the left are waveforms from upstream readouts, and on the right are waveforms from downstream readouts. The top two are from the anode, the middle two are from the hood pads, and the bottom two are from the cathode pads. The horizontal red lines are the calculated baselines, the vertical red lines are the leading edge time, and the region enclosed by the two dashed blue lines is the allowed window to search for hits.

along the wire.

Two independent measures of the hit time are supplied by the two anode waveforms. Figure 5.7 presents the distribution in the difference of these measurements. Neglecting possible correlations end-to-end, one may infer a resolution in the hit time of $\sigma_t = \frac{2.4}{\sqrt{2}} = 1.7$ ns in the central Gaussian along with collateral non-Gaussian tails. While uncorrelated noise and effects of the algorithm are present in this resolution estimate, it is not sensitive to electron drift diffusion, random motion due to thermal energy of the drift electrons.

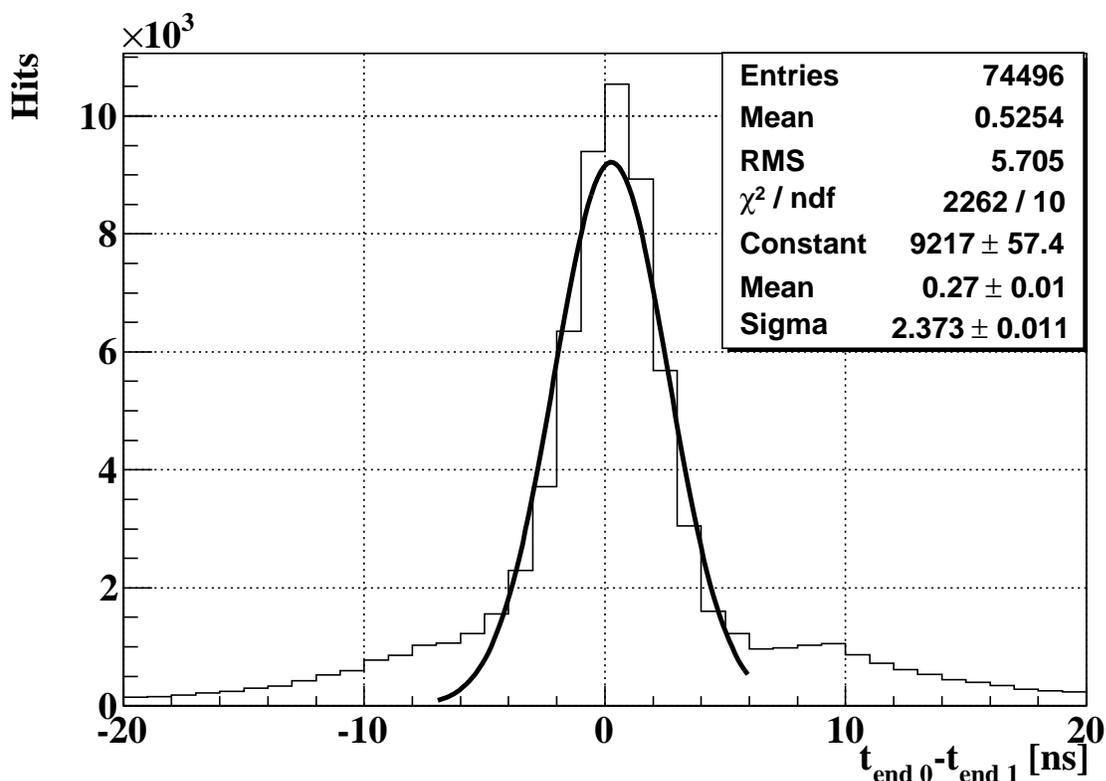


Figure 5.7: Distribution of $t_{\text{end}0} - t_{\text{end}1}$ for hits with both wire ends having a charge of at least .3 mV*ns. The region $|t_{\text{end}0} - t_{\text{end}1}| < 6$ ns is fit to a Gaussian.

For the purpose of charge integration, the pulse duration on each anode is defined by the nearest $2\sigma_{\text{pedestal}}$ crossing on either side of the peak. The integration region is bounded from below by a minimum of 15 ns (5 ns before the peak, 10 ns after it) and from above by a maximum of 80 ns (24 ns before the peak, 56 ns after it). If either waveform contains fewer

than 8 bins above 5 mV, it is no longer considered a hit candidate. The charge integration region for all six waveforms is taken to be that of the anode signal with largest maximum amplitude. A rectangular summation is employed to measure the charge contained within the integration range:

$$Q = \sum_{i \in \text{integration range}} V_i \cdot \Delta t_i. \quad (5.13)$$

V_i and Δt_i are respectively the voltage and time width of the i 'th bin. When the start or end time of the integration region falls inside a bin, only the fraction of the bin inside the integration limit weights its voltage. Note that the charge is measured in units of voltage * time.

There are several grounds upon which to base the optimization of the integration time. At short integration times, timing errors such as systematically misaligned waveforms and uncertainties in the DRS bin widths tend to produce large errors in Z . On the other hand, long integration times enlarge the path length of the track that contributes to the measurement, thus accepting ionization sites over a broader extent in Z . N contributing ionization sites spanning a distance dZ will impart a component of $\frac{dZ}{\sqrt{12N}}$ to the Z resolution. Minimizing the integration time also reduces problems from multiple hits not well separated in time.

Characteristics of the electronic noise spectrum present additional concerns. For a sharp component in the frequency domain, integrating over a multiple of its period is preferred. Random noise bestows an integration time dependent constituent to the Z resolution as well and, in general, possesses bin-to-bin correlations. By rewriting the measured charge in Eq.

5.13 in terms of the average voltage, \bar{V} , and the total integration time, ΔT :

$$Q = \frac{\sum_i V_i \cdot \Delta t_i}{\sum_i \Delta t_i} \sum_i \Delta t_i = \bar{V} \Delta T, \quad (5.14)$$

the problem of determining noise-related contributions to the resolution is reduced to measuring noise related contributions to $\sigma_{\bar{V}}$:

$$\sigma_Q = \sigma_{\bar{V}} \Delta T. \quad (5.15)$$

Comparing the predicted baseline, V_P , to the average baseline in the signal region, \bar{V}_M , in events without a signal allows one to isolate the ingredient of $\sigma_{\bar{V}}$ due to pedestal fluctuations as $\sigma_{(V_P - \bar{V}_M)}$. This is found to vary significantly over the numerous readout channels of the detector. Consequently, the integration time dependence of the uncertainty in Z originating from fluctuations in the baseline can be quantified through the following prescription:

- Measure $\sigma_{\bar{V}}$ from pedestal fluctuations for all pad channels in the drift chamber.
- Run the reconstruction on data, calculating σ_Q from Eq. 5.15 and the resulting error in Z from Eq. 5.18, assuming no correlations from end to end, on an event-by-event basis, for many hits.
- Repeat the above procedure for different values of the integration time limit, ΔT .
- For each value of the integration time, plot a single data point representing the resulting error in Z averaged over all hits.

Figure 5.8 shows the dependence of the measured noise contribution to the Z resolution, as well as the measured overall Z resolution itself, on the integration window. Both curves acquire a minimum at 80 ns, which also matches the period of a known harmonic in the noise spectrum at 12.5 MHz, hence the 80 ns limit on the integration window. The validity of

the baseline time-independence assumption is similarly studied by comparing the precision with which constant, linear, and quadratic time-dependence fits predict the pedestal in the signal region in events with no signal. The conclusion is that a constant fit displays the best performance.

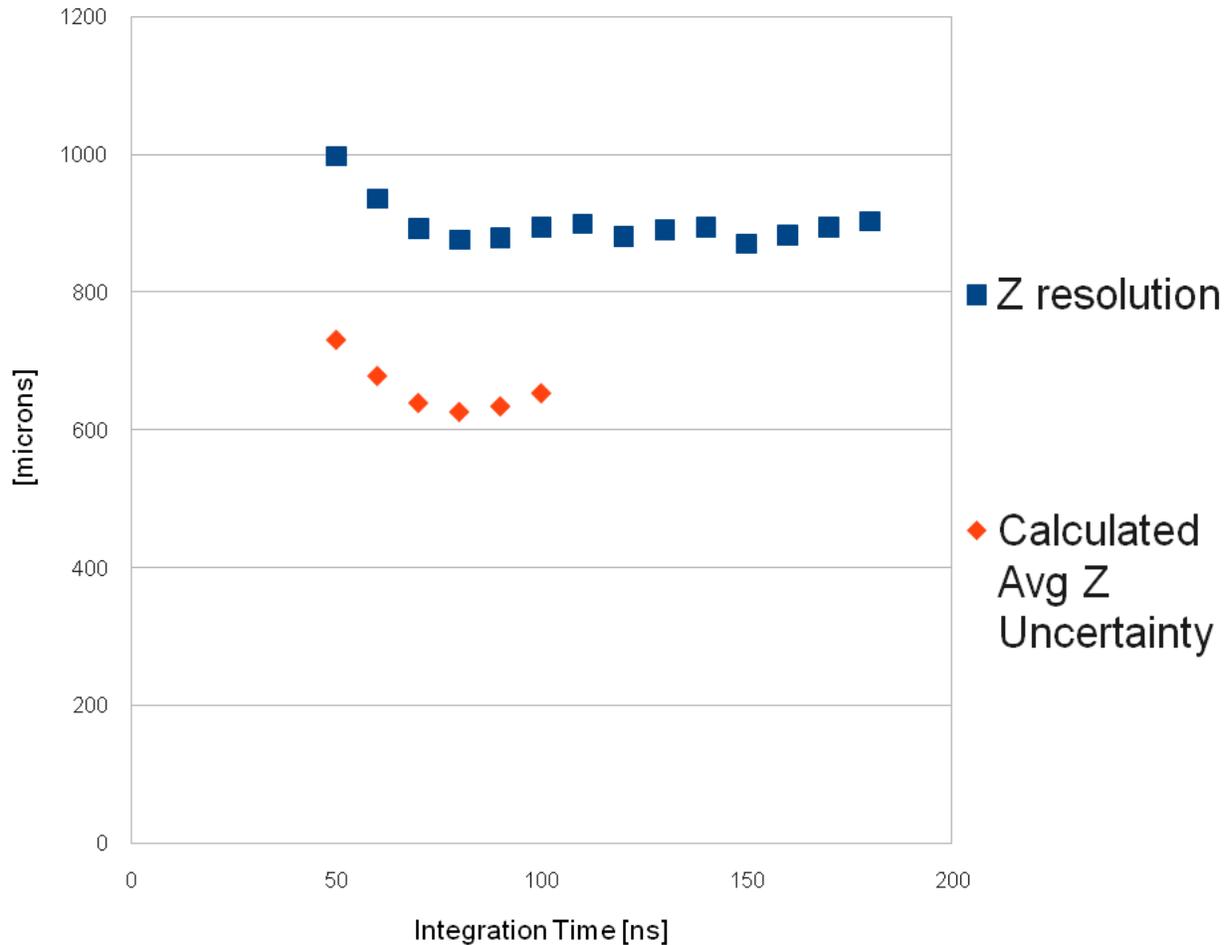


Figure 5.8: The blue data points are direct measurements of the Z resolution while the red data points are the calculated contributions of electric noise to the resolution at different integration times. Since the calculations for the latter involve measuring pedestal fluctuations for over a thousand channels, they are performed for a limited set of integration times.

Through anode charge division, one obtains a first estimate of the Z coordinate. Assuming equal preamplifier input impedance, R , at each end of the anode, and a uniform resistance per unit length, ρ , the Z coordinate along a wire of length L centered at the origin may be

calculated from the anode charges as:

$$Z_{anode} = \left(\frac{L}{2} + \frac{R}{\rho}\right) \frac{Q_d^{anode} - Q_u^{anode}}{Q_d^{anode} + Q_u^{anode}}. \quad (5.16)$$

$Q_{u(d)}^{anode}$ is the measured anode charge at the upstream(downstream) end. The factor

$$A^{anode} \equiv \frac{Q_d^{anode} - Q_u^{anode}}{Q_d^{anode} + Q_u^{anode}} \quad (5.17)$$

is termed the anode charge asymmetry. From the measured pad charge asymmetries one attains further precision on the Z coordinate:

$$Z_{pad} = \frac{\lambda_{pad}}{2\pi} \arctan\left(\frac{A^{hood}}{A^{cathode}}\right) + \delta_{pad\ cycle}(Z_{anode}). \quad (5.18)$$

$\lambda_{pad} = 5$ cm is the length of a pad cycle and $\delta_{pad\ cycle}(Z_{anode})$ describes the occupied pad cycle; it is based solely on Z_{anode} .

If one or more of the four pad channels associated with a cell is not functioning, the anode-determined Z coordinate is used.

Figure 5.9 compares the two calculated Z coordinates: Z_{anode} and Z_{pad} . Because the uncertainty of Z_{pad} is an order of magnitude smaller than Z_{anode} , the width of this distribution yields a measure of the resolution on Z_{anode} , about 8 mm. If the error on Z_{anode} exceeds 2.5 cm, the hit will be placed on the wrong pad cycle. The shape of figure 5.9 suggests that this will occasionally happen; however, this situation is partially corrected during clustering and tracking.

After recording a hit, the algorithm zeros out the bins in both the smoothed and raw waveforms within the hit integration limits, and iterates the process to find additional hits until either no peaks above threshold are encountered or no bins remain in the waveforms.

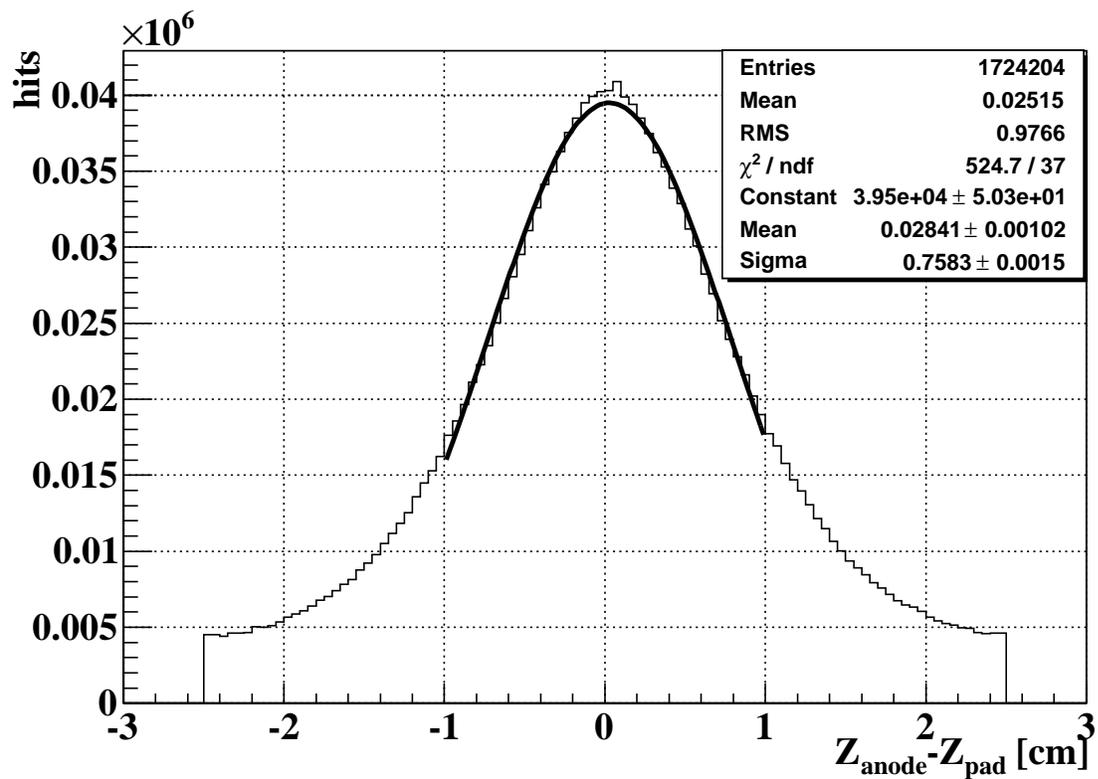


Figure 5.9: Distribution of the difference in anode and pad determined Z coordinates of hits. The region $|Z_{anode} - Z_{pad}| < 1$ cm is fit to a Gaussian.

5.2.2 Cluster Reconstruction

Given that a chamber may contain many hits, it is a useful intermediate step toward tracking to group collections of hits on a single chamber into clusters, if they are consistent with coming from the same charged particle.

Because no trajectory information is available at this stage, the cluster formation requirements are intentionally kept very loose. The hits on a chamber are first assembled into non-overlapping sets with no gap larger than 2 cells between hits on the same plane and no more than 10 hits per set. Next, the clusters are trimmed of hits until no hit deviates from the average Z coordinate by more than 3.75 mm. This threshold is established to retain hits in a cluster whose pad cycle may have been measured incorrectly. Consider the benchmark case of a two-hit cluster from a charged particle path in which one hit is reconstructed on the pad cycle adjacent to its true location; the maximum possible difference in the hit Z coordinates is 7.5 cm and each deviates from the average Z by 3.75 cm. The clusters are further split into sub-clusters requiring no gap larger than 1 cell between hits. All hits not associated with some cluster are checked for consistency with other clusters on the basis of the above criteria iteratively until each chamber contains only two kinds of objects: clusters of spatially correlated hits and disjunct hits that cannot be matched to any cluster. Both objects are treated as independent clusters.

As a result of the loose requirement on the Z consistency of hits within a cluster, some clusters may span a total distance along Z (>2.5 cm) beyond that expected for a set of hits coming from primary ionization sites distributed only along a track's comparatively short path in the chamber. The algorithm assumes that this is due to an incorrect pad cycle assignment of one or more hits in the cluster and attempts to correct the situation. Clusters composed of either a single hit or more than four are not treated for this; the remaining cases are handled as described below.

A two-hit cluster with a Z span greater than 2.5 cm is compatible with either of its hits being wrong by one pad cycle. Any such cluster is then replaced by two alternate clusters, each created by shifting one of the two hits by one pad cycle, or 5 cm, toward the other. This technique is based on the philosophy that, although the true manifestation of this cluster is not known at this stage, the tracking algorithm will discern between the two possibilities by choosing the one that best fits a track, if any.

A three-hit cluster with a Z span between 2.5 cm and 7.5 cm is subject to a similar procedure. In this instance, the most outlying Z coordinate is translated by one pad cycle toward the Z coordinates of the other two hits.

Four-hit clusters are presumed to fall into one of three categories. In the simplest case, all four hits are consistent with correct pad cycle placement and nothing is done. In another scenario, three hits are within one pad cycle of each other and one hit deviates by more than 2.5 cm from at least one of these three. Then there are two ways in which one hit could have been mis-measured to produce such a scenario. Either the hit at smallest Z originated one pad cycle toward positive Z away from its measured position, or the hit at largest Z originated one pad cycle toward negative Z away from its measured position. The original cluster is then replaced by both possibilities. Lastly, there may be two pairs of hits in which a hit in a pair is consistent with the other hit in the pair, but one pair has measured Z coordinates that are inconsistent with those of the other pair. The two ways in which this could have come about by mis-measuring the pad cycle of both hits in a pair are recreated in place of the original cluster. Once again, these methods rely on the tracking to distinguish among two possible forms of the true cluster.

Note that when a hit is mis-measured by more than one pad cycle, no attempt is made to remedy this circumstance.

A parent cluster is assigned a global R and Z coordinate by individually averaging the wire

R coordinates and inferred Z coordinates of its daughter hits in each plane of the cluster and then averaging those two measures. This is the best estimate of the particle's coordinates at the chamber center.

5.2.3 Track Reconstruction

The arrangement of clusters in an event is then examined for patterns, called track candidates, that resemble one turn of a positron trajectory in the DCH.

Signal positrons tend to reach large radii by virtue of COBRA's design, where hit rates are also lowest. This is advantageous for pattern recognition because it facilitates searching for tracks in a region less prone to accidental hits. This is to be balanced with the desire to abstain from introducing strongly momentum-dependent acceptance effects aside from those already inherent in the design of the drift chamber. Exploiting these attributes, the algorithm begins by finding combinations of three adjacent clusters, called seeds, whose center cluster resides at the maximum R coordinate of the potential track, with the mild requirement that it exceed 20 cm to conserve acceptance in a substantial region of positron momentum. A seed must have spatial agreement between adjacent clusters, no more than one skipped chamber in its span, and at least one multiple-hit cluster among the three. The difference in the radius of any two consecutive clusters is restricted to 2 cm, or 4 cm when a skipped chamber exists between them. The difference in Z coordinates between any two sequential clusters must lie within a range spanning 4 cm, or 8 cm when separated by a skipped chamber. The center of this range is a linear function of the Z coordinate, instead of zero, to account for the fact that clusters at large $|Z|$ could only have come from positrons with large $|p_z|$ and are thus naturally expected to travel a large distance in Z between chambers. The parameters of this function are chosen to give high efficiency by studying data.

Bereft of any drift time information, the algorithm has not yet pinned down the seed's radial

coordinates with much precision. Toward that goal, one can make use of the independently reconstructed impact times at the TIC; however, it is not possible at this stage to determine which TIC cluster, if any, is associated with the seed, whose path is only roughly known. This impediment is overcome by constructing as many duplicate seeds as there are TIC clusters and attempting to reconstruct each possibility, thus $N_{TIC\ clusters} * N_{seeds}$ independent seeds will now exist. Since a positron takes of order 1 ns to make one turn in the DCH, the primary ionization times along the track are treated as simultaneous to good approximation. The drift time of the i 'th hit in the seed is therefore expressed as

$$t_{drift}^i = t_{hit}^i - T_0, \tag{5.19}$$

where t_{hit}^i is the i 'th hit's absolute arrival time as defined during the hit reconstruction stage. T_0 , also known as the track time, is calculated by correcting the TIC time by the average time of flight between the DCH and the TIC. The uncertainty in T_0 , which is affected by the TIC time resolution (~ 90 ps) and the dispersion in the time of flight (~ 1 ns), makes a negligible contribution to the radial precision (~ 40 μm), except when two turns are made in the DCH. In this instance, one of the turns (the first one) will be reconstructed with a T_0 that is systematically late by typically ~ 3 ns, delivering a ~ 120 μm component to its radial resolution. From Eq. 5.19, the drift time refers to the time elapsed between a primary ionization event and the arrival of the earliest part of the subsequent avalanche. By treating the track as a straight line at a known angle through the cell, with a continuum of ionization sites, the measured drift time can be used to infer the distance of closest approach to the wire whose ionization site of minimum drift time is the closest match. The angle is inferred by computing the circular path in the XY plane defined by the triplet of X,Y coordinates in the seed. A massive, finely-binned lookup table derived from the GARFIELD software (see section 4.4.2) acts as a Rosetta Stone for converting the track angle, drift time, and local electromagnetic field, to the (x,y) location of the original ionization site. By

symmetry under reflection about the axis of the wire, the situation is still obscured by a two-fold degeneracy as to the side of the wire upon which the track impinges. This “left/right ambiguity” (l/r) is resolved by testing all $2^{N_{hits}}$ possible configurations in each cluster and choosing the configuration that minimizes the figure of merit:

$$\sum_{i=1}^{N_{hits}} (R_i - \bar{R})^2, \quad (5.20)$$

where R_i is the R coordinate of the i'th hit projected onto the central chamber plane, and \bar{R} is the average of this quantity over all hits in the cluster. The radial cluster coordinate is then replaced with \bar{R} to better estimate the local trajectory. Resolving l/r for single-hit clusters is postponed until the full extent of the track is known.

Next, the algorithm attempts to extend each seed in both directions, adding clusters to the track until either no more clusters are able to pass the selection criteria or the last chamber is reached. In projecting the track to the next chamber, the instantaneous track circle and momentum vector are estimated using only the nearest three clusters. The radius of curvature is only approximately constant in the non-uniform magnetic field; however, one can take advantage of the adiabatic invariants for slowly varying axial magnetic fields:

$$\frac{p_T^2}{B_z} = k_1 \quad (5.21)$$

$$R_{track} B_z = k_2, \quad (5.22)$$

where p_T and R_{track} are respectively the instantaneous transverse momentum and radius of curvature, and B_z is the z-component of the magnetic field vector. Based on the conservation law of Eq. 5.22, an estimate of the expected radial point of intersection of the track with the next chamber is made. The expected radial intersection is then coupled with two constants of the motion, $|\vec{p}|$ and Eq. 5.21, to predict the intersection in Z. A cluster falling within 4 cm

of the predicted radial location and 2.5 cm of the predicted Z location is added to the track. If more than one cluster on the same chamber satisfies the track projection criteria, the track is split into multiple possibilities, and each is continued along the chamber separately. The skipping of one additional chamber per direction is allowed in tracking beyond the skipped chambers from seeding.

Equipped with trajectory information and a track time, the algorithm checks for badly deviant hits to remove from their liberally formed clusters. On every track, each cluster is stripped of irregular hits until either all remaining hits satisfy all criteria or only one hit remains on the chamber. Hit removal is carried out with highest priority on hits with unphysical drift times (either negative or beyond the maximum possible drift time), then on hits whose radial or Z projection onto a common plane fails to agree well with the other hit projections. The projection methods are explained in greater detail in section 8.1.2. Given a set of $N_{hits} > 2$ projected R or Z coordinates ordered from smallest to largest, some parameters are defined to consider for the removal of the most outlying hit. The set of hits in the cluster except the one being considered for removal is called the sub-cluster. $\delta Z(R)$ is the largest spacing between two adjacent hits in the sub-cluster in R(Z), and $\Delta R(Z)$ is the distance in R(Z) between the candidate for removal and the closest hit coordinate. If either $\frac{\Delta R}{\delta R} > 3$ and $\delta R > 1$ mm or $\frac{\Delta Z}{\delta Z} > 3$ and $\delta Z > 2$ mm, the hit is expelled from the track.

As stated previously, $1/r$ resolution for single-hit clusters cannot be based on the figure of merit Eq. 5.20. In fact, that method is also problematic for two-hit clusters when the track direction is nearly parallel to a line connecting the two wires, for then Eq. 5.20 varies only slightly among the two possible configurations consistent with that track angle. A supplementary approach is invoked in these cases. Since 3 clusters are sufficient to estimate the track momentum, there are $N_{clusters} - 2$ measurements available, and one can resolve

these lingering 1/r ambiguities by minimizing a surrogate figure of merit:

$$\sum_{i=1}^{N_{clusters}-2} (|\vec{p}_i| - |\vec{p}|)^2. \quad (5.23)$$

A more sophisticated fitting routine is then supplied with a list of hits associated with each track. It is provided with the measured Z coordinate, the distance of closest approach to the wire, the 1/r resolution, and a set of measurement uncertainties for each hit. The uncertainty in Z is calculated from the pedestal fluctuation contribution discussed in section 5.2.1 together with an estimated contribution from charge calibration discussed in section 6.2.2. A simple model, linear in the distance of closest approach, estimates the uncertainty in the radial impact parameter.

5.2.4 Track Fitting

All track candidates with at least 4 clusters are sent to a Kalman filter [59] to fit a trajectory to the set of hits. At any instant, the charged particle is described by a six-dimensional state vector:

$$\vec{x} = \left\{ R, Z, \arctan\left(\frac{p_y}{p_x}\right) - \phi, \frac{p_z}{p_T}, \frac{1}{|\vec{p}|}, \phi \right\}. \quad (5.24)$$

From a set of sequential measurements of the state vector, subject to random errors assumed to be normally distributed, evolving with time according to a known model, the Kalman filter recursively estimates the true state vectors. Based on the information of k-1 hits, the state vector at the k'th hit can be predicted:

$$x_k^{k-1} = F_k^{k-1} x_{k-1}. \quad (5.25)$$

Here the best estimate of the state vector at the (k-1)'th hit, x_{k-1} , is propagated to the location of the k'th hit, x_k^{k-1} , by the operator F_k^{k-1} . Both the effect of the magnetic field and the average energy loss are encoded in F_k^{k-1} . In the extrapolation from k-1 to k, the errors on the state vector components will also change:

$$C_k^{k-1} = F_k^{k-1} C_{k-1} (F_k^{k-1})^T + Q_k. \quad (5.26)$$

The predicted covariance matrix at k, C_k^{k-1} , receives contributions from F_k^{k-1} as well as stochastic processes in the propagation described by Q_k , namely multiple scattering. The actual measurements of Z and distance of closest approach, m_k , are mapped onto the true state, x_k , by a projection matrix, H_k , and the uncertainties of the measurements, ϵ_k , according to:

$$m_k = H_k x_k + \epsilon_k. \quad (5.27)$$

The predicted state at k, x_k^{k-1} , is then corrected by the information in the measured state at k, m_k , to produce a filtered estimate of the true state vector at k:

$$x_k^k = x_k^{k-1} + K_k (m_k - H_k x_k^{k-1}). \quad (5.28)$$

K_k is the Kalman gain matrix, defined as:

$$K_k = C_k^{k-1} H_k^T (V_k + H_k C_k^{k-1} H_k^T)^{-1}, \quad (5.29)$$

where V_k is the covariance matrix of measurement errors, ϵ_k . Concurrently, the full covariance matrix at state k, C_k^k , is updated:

$$C_k^k = [I - K_k H_k] C_k^{k-1}. \quad (5.30)$$

Once all hits are processed, the filter updates, or smooths, all of the filtered state vectors and covariance matrices with the full information from all hits:

$$x_k^n = x_k^k + A_k(x_{k+1}^n - x_{k+1}^k) \quad (5.31)$$

$$C_k^m = C_k^k + A_k(C_{k+1}^m - C_{k+1}^k)A_k^T, \quad (5.32)$$

where

$$A_k = C_k^k F_k^T (C_{k+1}^k)^{-1}. \quad (5.33)$$

A χ^2 that tests the agreement between measured and predicted state vectors is used to gauge badly measured hits for removal and incorrect l/r solutions for switching. A similar χ^2 , calculated by comparing state vectors at the closest point near the beam line, serves as a means by which to check whether or not two fitted tracks should be merged as two turns of one trajectory. Only tracks that use the same timing counter cluster in defining their track time are allowed to merge.

The estimated track momentum, before energy loss in going through the chambers, is corrected for energy loss in the target. The decay vertex and positron emission angles are obtained by propagating the state vector back to the target plane, requiring knowledge of the target location and orientation, and the TIC impact location is predicted by propagating the state vector forward to the TIC. An event-by-event indicator of the resolution in each state vector component is provided by the diagonal elements of the covariance matrix; this is later exploited in the likelihood analysis.

A discussion of the resolutions in the drift chamber measurements is given in section 8.1.2

5.3 TIC

After passing through the drift chamber, positrons that reach the timing counter deposit energy in one or more scintillating bars. The waveforms read out by the PMTs at either end of each bar provide a measurement of the impact time and an estimate of the impact position that together serve as the only discriminating tools for matching a TIC hit with a DCH track. A group of software algorithms first measure the impact time and location in each hit bar from the PMT waveforms; this information is called a hit. Groups of hits consistent with coming from the same particle are then identified as clusters.

5.3.1 Waveform Analysis

Recall from section 4.6 that 80% of a PMT waveform is sent to a double threshold discriminator and then digitized. These NIM pulses form the basis of the waveform analysis. Figure 5.10 displays an example PMT waveform.

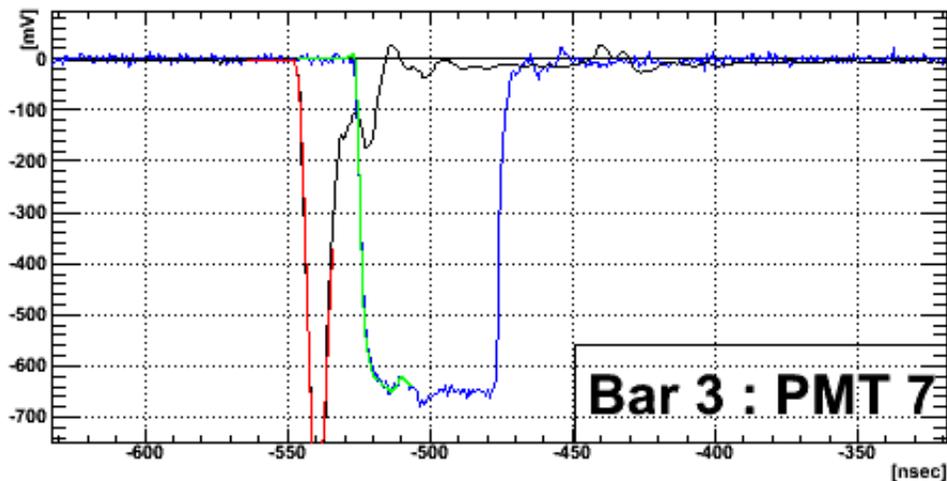


Figure 5.10: An example of the waveforms used in the timing counter reconstruction. The black line is a DRS output, which is fit to a template (red). The blue line is the NIM pulse output, which is fit to another template (green). The time delay of the NIM pulse with respect to the DRS waveform is due to an electronic delay in the double threshold discriminator.

A template waveform shape for each PMT is constructed from the accumulation of many events. This template shape is then fit to the NIM pulse (see section 4.6) in each hit for a flat baseline and a leading edge time.

In addition to the leading edge time, the charge and amplitude are also measured using the DRS waveforms (one of the 10% outputs of the passive splitter described in section 4.6) primarily for calibration purposes. The charge is measured by integrating the waveform over a 30 ns window, and the amplitude is found as the difference between the peak voltage and the fitted baseline.

5.3.2 Hit Reconstruction

A hit in a bar is defined as the information from the upstream and downstream PMTs associated with an energy deposit from a charged particle passing through the bar. The leading edge time from each PMT is corrected for time walk effects, and the difference between the two times gives a Z coordinate. The hit time is the average of the two leading edge times corrected for photon propagation time in the bar based on the measured Z coordinate. Hits are then clustered by closeness in time and Z location. Multi-hit clusters take the time of the first hit bar as the cluster time.

A discussion of the timing counter time resolution is given in section 8.1.3.

5.4 DCH-TIC Association

At first, the tracking and timing counter reconstruction proceed independently; the tracks are built using the times of the timing counter hits, but no judgment on whether or not they are spatially correlated has taken place. The next step is to select pairs of DCH tracks and

TIC clusters that are consistent with belonging to a single charged particle trajectory.

Each track is propagated to the ϕ location of the first bar of the TIC cluster whose time was used to define the track time. If the projected impact point on the bar deviates by more than 20 cm from the TIC hit Z coordinate ($\Delta Z_{DCH-TIC}$) or by more than 10 cm from the radial coordinate of the bar center ($\Delta R_{DCH-TIC}$), the pair is rejected. Lastly, the spatial matching $\chi_{TIC-DCH}^2$ is required to be less than ten to accept the pair.

The time of flight of the trajectory from the muon decay vertex to the impact location in the TIC is computed by dividing the total path length by the speed of light. The projection from the last chamber to the TIC is complicated by its length and materials in its path such as cable ducts and preamplifiers. The accuracy is thus limited by scattering and energy loss. An additional correction to the time of flight is made according to the value of $\Delta Z_{DCH-TIC}$, whose form is determined from Dalitz data (see section 4.7.1).

In events with two (or more) hit bars, the projected path length of the trajectory between bars is coupled with the two hit time measurements to get two measurements of the impact time at the first bar. The TIC time estimator is the average of the two measurements in this case.

The best estimate of the hit time at the first bar is subsequently corrected by the best estimate of the time of flight to obtain the time of positron emission from the decay vertex, T_e .

5.5 Formation of MEG Candidates

The initial event formation strategy is to fuse the reconstructed positrons and photons in a single event to construct all possible combinations. In each case, the photon vertex is

assumed to coincide with the positron vertex, $\vec{\mathbf{x}}_\mu$. A straight line path from the vertex to the position of the first interaction, $\vec{\mathbf{x}}_\gamma$, is used to calculate the photon direction vector:

$$\hat{\mathbf{p}}_\gamma = \frac{\vec{\mathbf{x}}_\gamma - \vec{\mathbf{x}}_\mu}{|\vec{\mathbf{x}}_\gamma - \vec{\mathbf{x}}_\mu|} = (\sin \theta_\gamma \cos \phi_\gamma, \sin \theta_\gamma \sin \phi_\gamma, \cos \theta_\gamma), \quad (5.34)$$

and to correct the photon interaction time back to the photon emission time by the time of flight, $t_{\gamma\mu}$:

$$T_\gamma = t_{XEC} - t_{\gamma\mu}. \quad (5.35)$$

With this information, one can engineer three quantities that will reconstruct at zero with maximum probability in a signal event:

$$t_{e\gamma} = T_\gamma - T_e \quad (5.36)$$

$$\phi_{e\gamma} = (\phi_e + \pi) - \phi_\gamma \quad (5.37)$$

$$\theta_{e\gamma} = (\pi - \theta_e) - \theta_\gamma. \quad (5.38)$$

The set of five kinematic observables $\{E_e, E_\gamma, t_{e\gamma}, \phi_{e\gamma}, \theta_{e\gamma}\}$ forms the basis of the likelihood analysis.

Chapter 6

Calibration

This chapter details the calibration methods used in the experiment. This primarily consists of calibrating out the effects of electronics properties (e.g., time delays, gain) on the timing, position, and energy measurements in the detectors, and measuring the orientations of the detector elements.

6.1 DRS Calibration

The DRS chips, described in section 4.6.2, are calibrated for both voltage-dependent response and fluctuations in the sampling intervals among each of the 1,024 cells in a chip.

6.1.1 Voltage Calibration

Every cell of every DRS chip has a different non-linear output voltage response to an input voltage. An internal DC voltage generator on the DRS board is used to prepare detailed voltage response maps for each cell in each chip. The calibration is done online, so that the

waveforms stored in the raw data files are already calibrated for voltage response.

6.1.2 Time Calibration

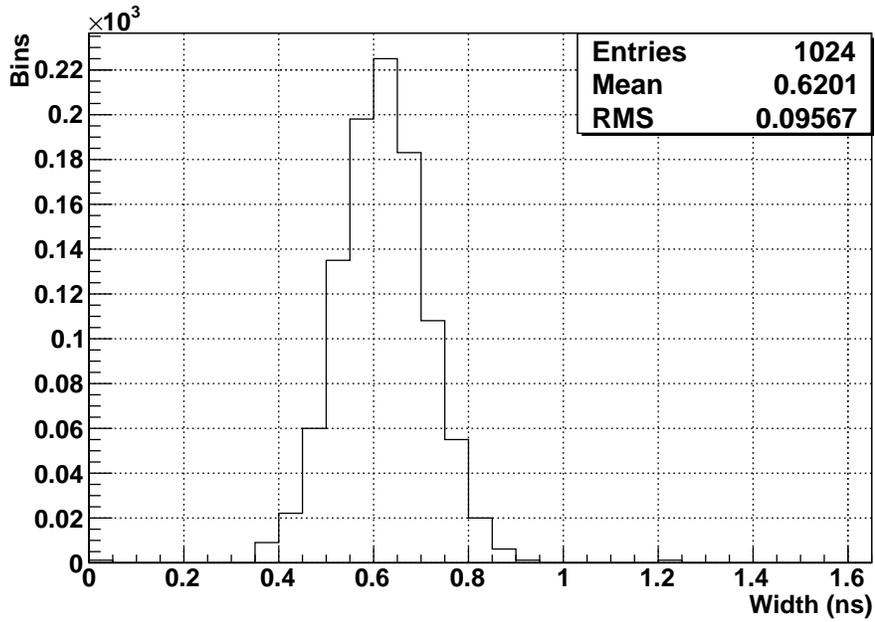
The sampling interval of each cell in a DRS chip can vary due to non-uniformities in the manufacturing process. These are calibrated by presenting sine waves of known frequency and random phase from a function generator to each chip. The time between zero crossings is measured, and all bins widths within the period are uniformly rescaled to match the true period with the measured period. This is repeated many times until the plots shown in figure 6.1 stabilize. The sampling interval of each cell of every chip is stored in a database and applied offline, during event reconstruction.

6.2 DCH Calibrations

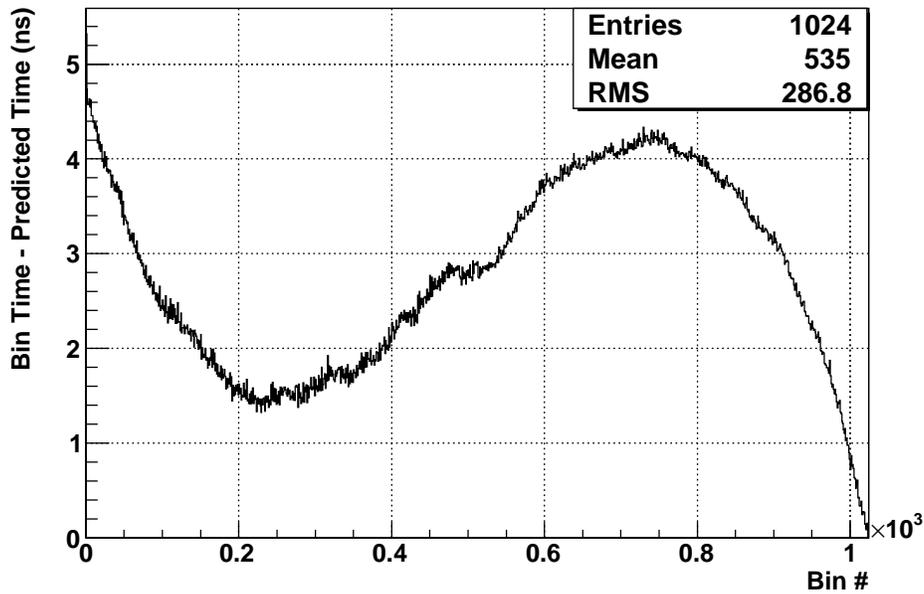
6.2.1 Waveform Time Calibration

Upon triggering, the 1,728 DCH channels are read out by the DAQ, and the contents of each of their 1,024 bins are stored. During the reconstruction, the 1,728 waveforms must be temporally aligned with one another. A 19.44 MHz internal trigger clock signal is distributed with precisely cut cables to all of the DRS chips in the experiment and synchronized with a specific bin, referred to as cell 0, in every DRS. The recorded location of cell 0 in each waveform serves as a reference for aligning the waveforms in time. Square pulses are simultaneously input to a few DRS boards to test the full set of effects from hardware and the software waveform alignment, and the resolution in the pulse time difference between different DRS channels is measured to be about 260 ps.

On top of this, there are relative electronics delays between channels due, for example, to



(a) Distribution of the 1,024 bin widths of a particular DRS chip as found by the time calibration.



(b) The accumulated error made by assuming nominal bin widths (i.e. before calibration) as a function of cell number with both schemes aligned at the last cell for a single DRS chip. The error is as much as 5 ns.

Figure 6.1: Some plots illustrating features of the time calibration on a single DRS chip.

varying cable lengths connecting the DCH outputs to the DRS boards or varying electronics delays. The calibration procedure for compensating these systematic delays takes different approaches for handling the anodes and the pads in view of the different roles they play in the reconstruction.

Hit times are measured exclusively from anode signals; therefore, time offsets in the anode waveforms of one wire with respect to those of other wires will deteriorate the drift time measurement and ultimately the radial position resolution. Time delays between the two ends of the same wire are also important because they result in misaligned charge integration windows on the two ends, which in turn damages the anode Z measurement. To determine the comparative time delays among anode ends, the distribution of measured drift times on each channel (defined by Eq. 5.19 but with t_{hit}^i representing the hit time measured from the correct end) is fit in a restricted range to a quintic polynomial. Figure 6.2 displays a typical fit. The leading edge of the distribution is calculated as the time when the fitted polynomial reaches 15% of its peak. An iterative calibration is performed by applying time shifts to each wire end to move the drift time leading edge toward the mean leading edge position for that end. Figure 6.3 illustrates the level of convergence achieved on the dispersion of leading edge positions. The systematic error in the technique is estimated to be around 2 ns by comparing the leading edge deduced from alternative fitting functions (e.g., third order polynomial, fourth order polynomial). Persisting offsets in the difference of leading edge positions among two ends of the same wire are then corrected by distributing the necessary correction symmetrically across the two ends.

The extent to which the same charge fraction is integrated on both ends of a pad figures crucially into the Z calculation, while their alignment with waveforms from other cells is irrelevant. Since the integration region is set by the largest amplitude anode waveform, synchronized pad and anode waveforms on a cell promotes optimal signal-to-noise ratios in the charge integration windows. First, systematic time delays between the upstream

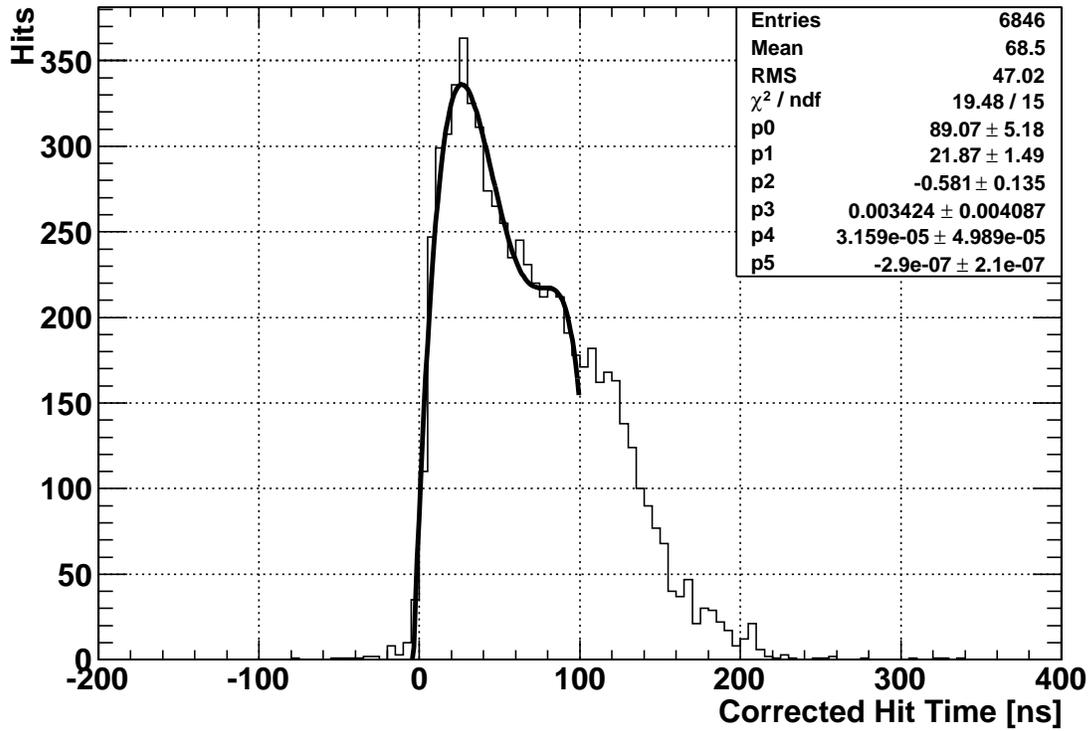
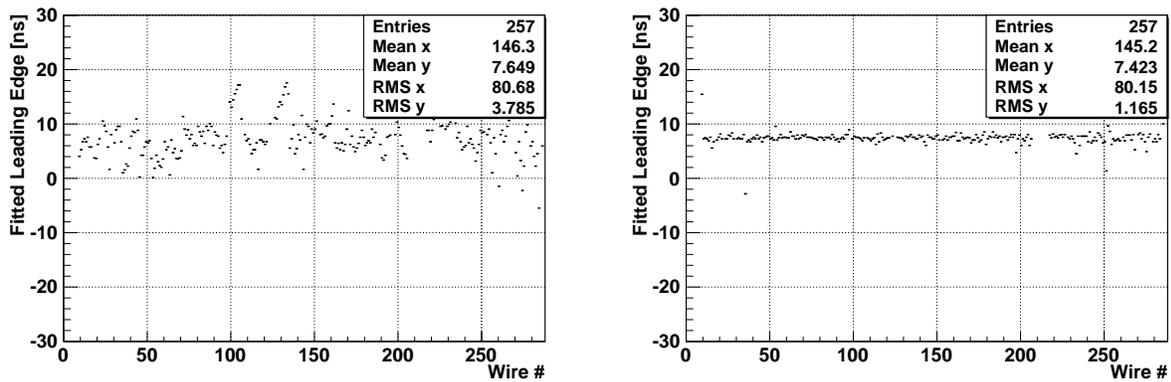


Figure 6.2: The drift time distribution of a particular anode channel fit to a 5th order polynomial.



(a) A scatter plot of pre-calibration leading edge time offsets on the upstream anode versus wire.

(b) A scatter plot of post-calibration leading edge time offsets on the upstream anode versus wire.

Figure 6.3: A pictorial summary of improvement in anode time alignment by calibration.

and downstream ends of a pad are found by plotting the difference in the peak amplitude times of the two ends of a pad and fitting for the mean. Figure 6.4 contains an example fit. Constraining the upstream/downstream time difference to vanish after calibration still leaves the freedom to shift the times of both ends by any constant amount. This allows the pad times to be aligned with the anode times by then plotting the difference in the peak amplitude times as measured by an upstream or downstream pad and anode, fitting for the means, and applying a constant shift to both ends of the pad based on the relative pad-anode time offset averaged over the upstream and downstream ends. See figure 6.5 for a typical fit to such a distribution.

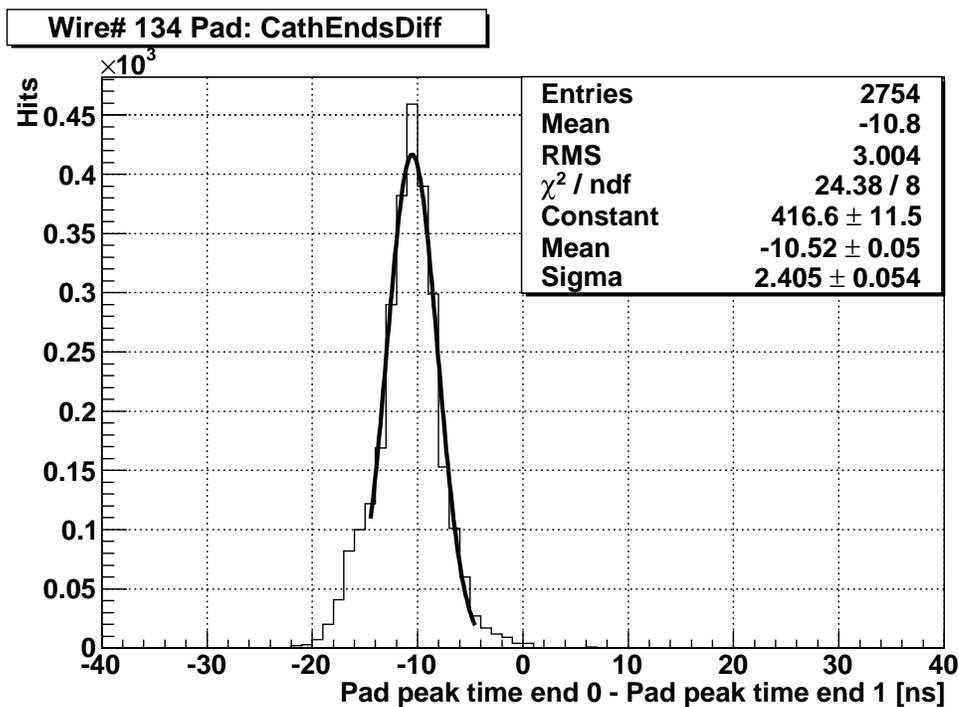


Figure 6.4: A fit to the distribution of the difference in peak amplitude times at the two ends of the cathode pad for a particular cell.

Figure 6.6 shows the level of convergence achieved on reducing these systematic offsets.

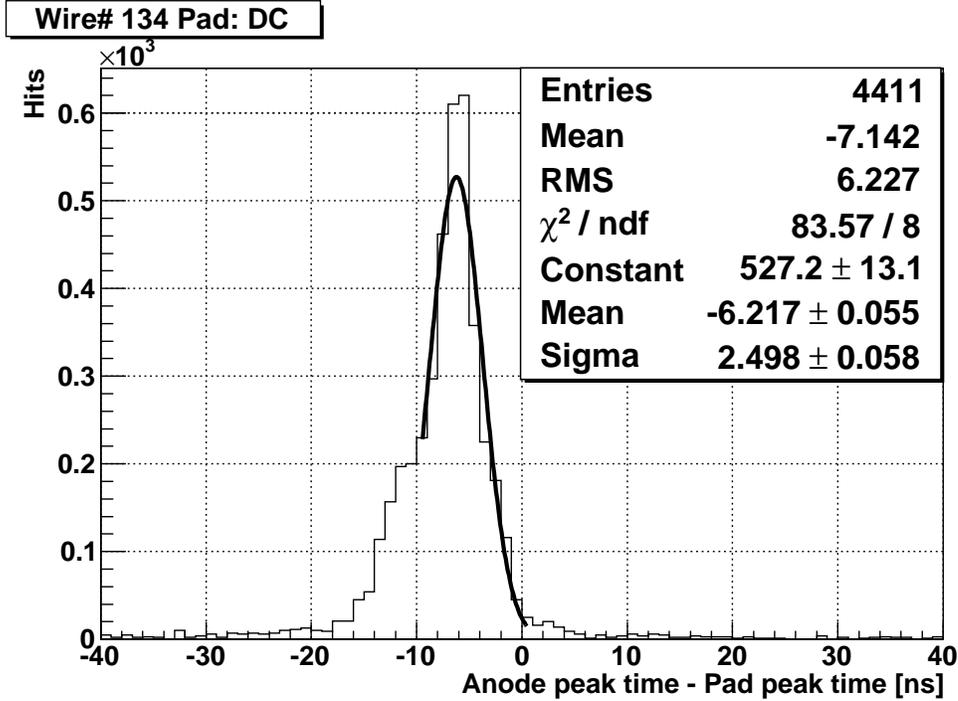
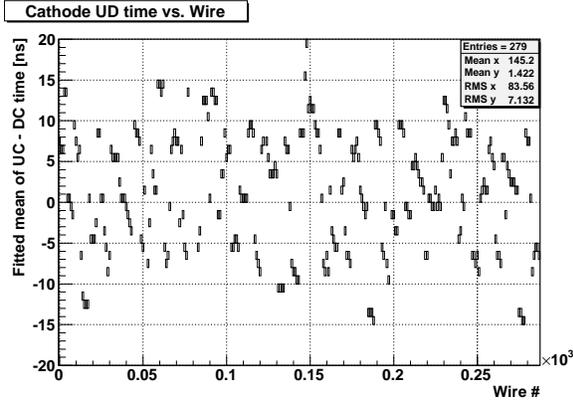


Figure 6.5: A fit to the distribution of the difference in peak amplitude times as measured from the downstream anode and cathode for a particular cell.

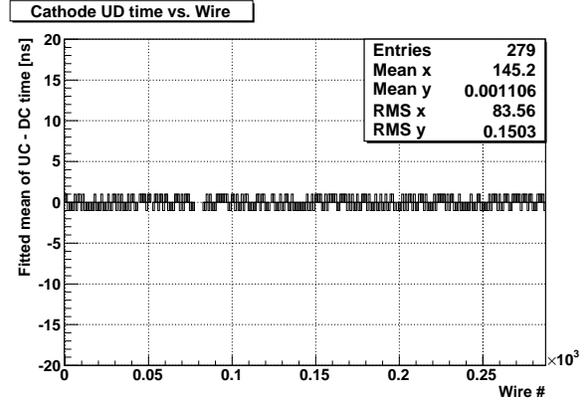
6.2.2 Z Calibration

A variety of electronic hardware properties influence the accuracy with which the anode and pad determined Z coordinates are measured. Errors in the knowledge of electronics parameters result in a different class of coherent effects than errors in the assumed wire positions in real space or in the relative positioning of the chambers. A set of Z calibrations is executed using only measurements from within the same cell and is thus only sensitive to the former class of effects.

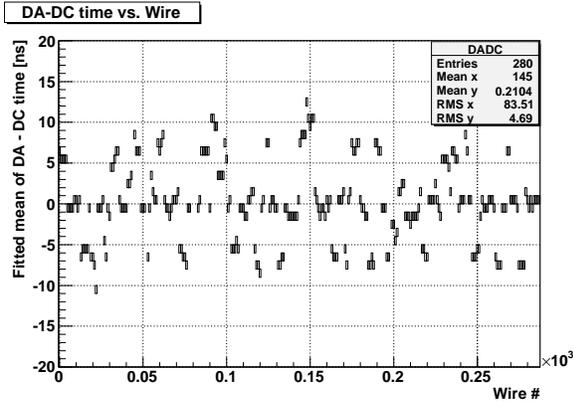
One group of calibrations involves adjusting the parameters entering into the anode Z calculation. These calibrations improve the probability of obtaining the correct pad cycle from the initial, anode charge division determination of Z. Recall that, under simplified assumptions, the anode Z coordinate is given by Eq. 5.16. The pad geometry, etched onto the foils with a 5 cm period known to a precision of 100 μm , is easily sufficient to serve as a calibration



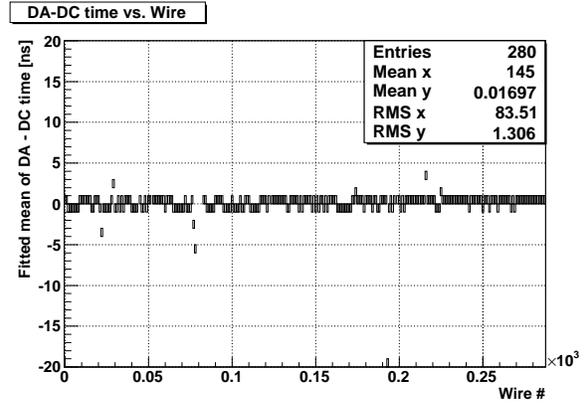
(a) A scatter plot of pre-calibration cathode time offsets between the two pad ends versus wire.



(b) A scatter plot of post-calibration cathode time offsets between the two pad ends versus wire.



(c) A scatter plot of pre-calibration cathode time offsets between downstream anode and cathode versus wire.



(d) A scatter plot of post-calibration cathode time offsets between downstream anode and cathode versus wire.

Figure 6.6: A pictorial summary of improvement in pad time alignment by calibration.

reference. Figure 6.7 demonstrates the fundamental calibration tool: profile histograms of the measured pad charge asymmetry as a function of the measured anode Z coordinate for each cell. Experimentally, these are well fit to a five-parameter function:

$$A^{pad}(Z_{anode}) = A_0(1 + m|z|) \sin\left(\frac{2\pi}{\lambda}(Z_{anode} - \delta)\right) + h. \quad (6.1)$$

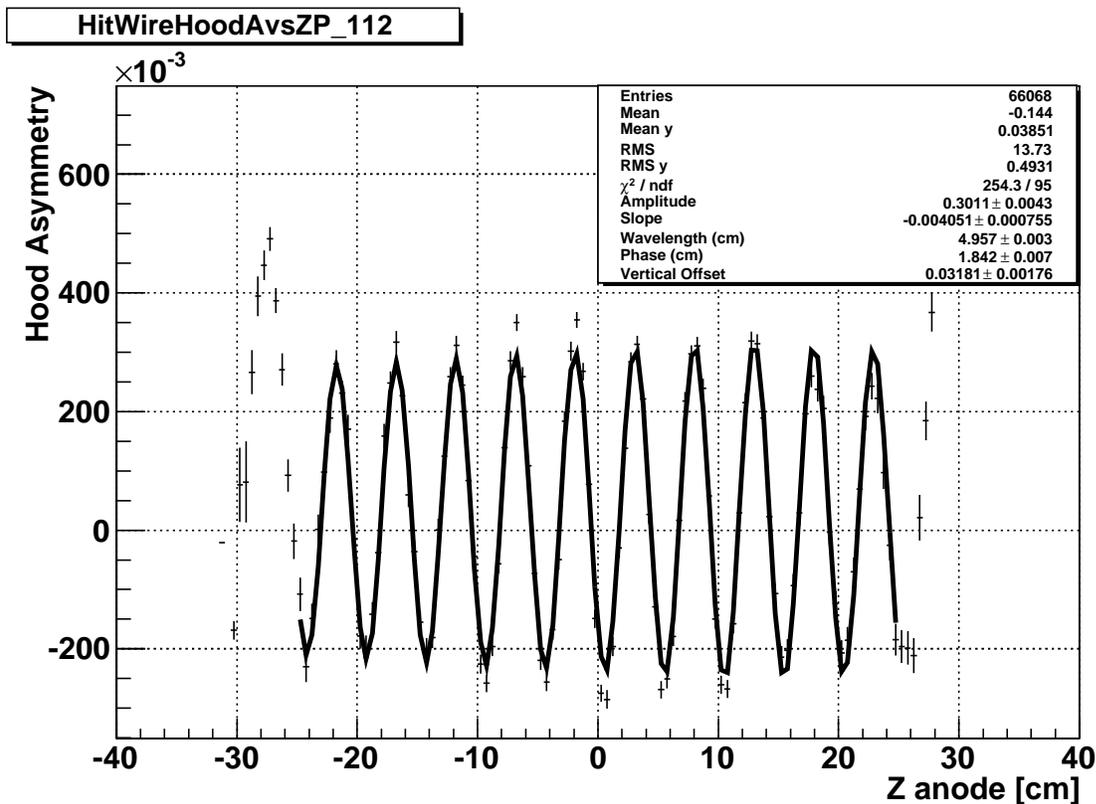


Figure 6.7: A fit of Eq. 6.1 to the pad asymmetry versus anode Z for the hood of a particular wire.

Calculation of the induced charge asymmetry agrees well with the sinusoidal dependence. While the function is not technically periodic, the roots are still equidistant and λ represents the effective length of a pad cycle, and δ is an effective phase (the smallest positive Z at which the pad asymmetry crosses zero while increasing with Z). h is a vertical offset that can arise from a relative gain between pad charges, A_0 is the amplitude at $Z = 0$, and m functions as a slope for the amplitude, which is observed to decrease linearly as one moves away from the

origin. The anode calibration then consists of measuring the wavelength, λ_{meas} , and phase, δ_{meas} , on each wire and applying corrections to the measured Z coordinate, Z_{meas} , to get a calibrated Z coordinate, Z_{true} , possessing the correct wavelength, $\lambda_{true} = 5 \text{ cm}$, and phase, δ_{true} (known from the pad geometry) when fit by Eq. 6.1:

$$Z_{true} = \frac{\lambda_{true}}{\lambda_{meas}} Z_{meas} + \delta_{true} - \frac{\lambda_{true}}{\lambda_{meas}} \delta_{meas}. \quad (6.2)$$

The scaling factor, $\frac{\lambda_{true}}{\lambda_{meas}}$, in the first term of Eq. 6.2 will differ from unity if the value used for $(\frac{L}{2} + \frac{R}{\rho})$ in Eq. 5.16 disagrees with the correct one. This part is calibrated out by making the substitution:

$$\left(\frac{L}{2} + \frac{R}{\rho}\right) \rightarrow \frac{\lambda_{true}}{\lambda_{meas}} \left(\frac{L}{2} + \frac{R}{\rho}\right). \quad (6.3)$$

The remaining terms in Eq. 6.2 produce a translation in the Z coordinate, which can emerge from a relative gain between upstream and downstream anode preamplifiers or DRSs. Both the absolute gain on any charge and the relative gain between the cathode and hood pads are inconsequential to the calculation of Z. Under a relative downstream/upstream gain of $1 + \epsilon$, Eq. 5.16 becomes:

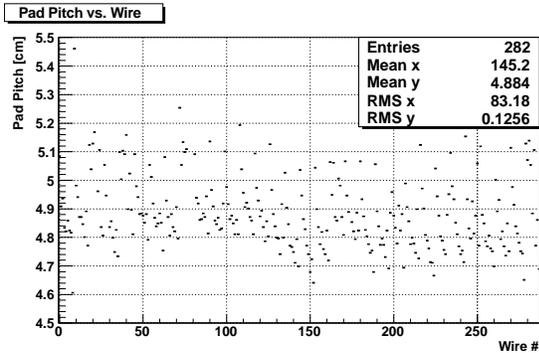
$$\left(\frac{L}{2} + \frac{R}{\rho}\right) \frac{1 - \frac{Q_d}{Q_u}(1 + \epsilon)}{1 + \frac{Q_d}{Q_u}(1 + \epsilon)} = \left(\frac{L}{2} + \frac{R}{\rho}\right) \frac{1 - \frac{Q_d}{Q_u}}{1 + \frac{Q_d}{Q_u}} + d = z + d. \quad (6.4)$$

That is, the true Z coordinate is shifted by an amount d, such that:

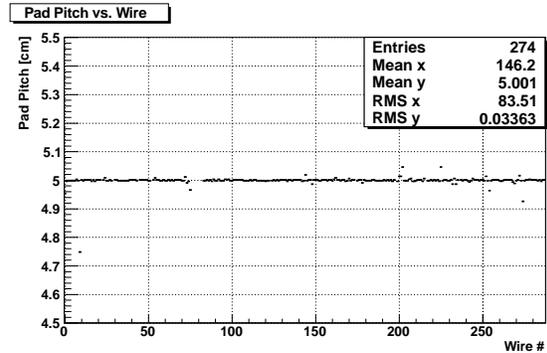
$$\epsilon = -\frac{(1 + \frac{Q_d}{Q_u})^2 d}{\frac{Q_d}{Q_u} (2(\frac{L}{2} + \frac{R}{\rho}) + d + d\frac{Q_d}{Q_u})}. \quad (6.5)$$

Equivalently, one must translate the rescaled Z coordinate by an amount $d = \delta_{true} - \frac{\lambda_{true}}{\lambda_{meas}} \delta_{meas}$ according to Eq. 6.2 to turn it into Z_{true} . For this value of d, one can find the necessary gain correction from Eq. 6.5. Since this expression depends on $\frac{Q_d}{Q_u}$ and hence Z_{meas} , it should

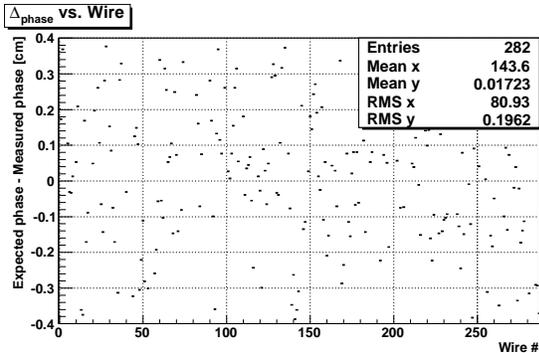
be averaged over the values of Z_{meas} populating the fit range. This is roughly estimated by setting $\frac{Q_d}{Q_u} = 1$, which corresponds to $Z_{meas} = 0$, the center of the fit range. From the two sets of pads per cell, two independent measures of the scaling factor on Z and the gain correction are obtained and averaged. Figure 6.8 encapsulates the distributions in λ and $\delta_{true} - \delta_{meas}$ before and after calibration.



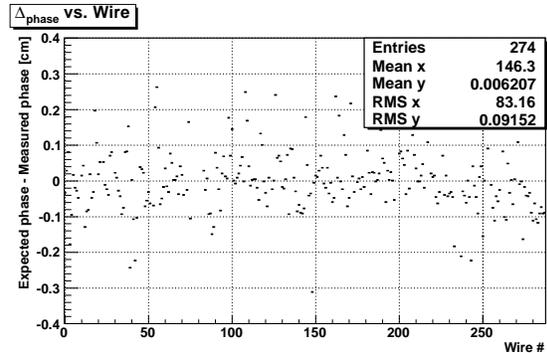
(a) A scatter plot of pre-calibration fitted values of λ versus wire.



(b) A scatter plot of post-calibration fitted values of λ versus wire.



(c) A scatter plot of pre-calibration deviations of anode phase from expected value versus wire.



(d) A scatter plot of post-calibration deviations of anode phase from expected value versus wire.

Figure 6.8: A pictorial summary of improvement in anode parameters by calibration.

Another calibration procedure is carried out on the pads to achieve optimal resolution in the final estimate of the Z coordinate. If there were no systematic, relative downstream/upstream pad gain, then the pad charge asymmetry distribution would be centered at the origin. Consequently, the parameter h in Eq. 6.1 can be translated into a relative pad gain. The

relation is constrained by the orientation of the Vernier pattern. The upstream charge, u , and downstream charge, d , along the length of one pad cycle can be parameterized as:

$$u = 1 + a_0 \sin \theta \quad (6.6)$$

$$d = 1 - a_0 \sin \theta, \quad (6.7)$$

where a_0 is the asymmetry amplitude and θ ranges from 0 to 2π . If the upstream and downstream charges are respectively subjected to gains G_u and G_d , the pad asymmetry becomes:

$$A = \frac{u - d}{u + d} = \frac{G_u - G_d + a_0(G_u + G_d) \sin \theta}{G_u + G_d + a_0(G_u - G_d) \sin \theta} = \frac{x + a_0 \sin \theta}{1 + x a_0 \sin \theta} \quad (6.8)$$

where $x = \frac{1 - \frac{G_d}{G_u}}{1 + \frac{G_d}{G_u}}$. The Maclaurin series expansion about x is:

$$A = a_0 \sin \theta + (1 - a_0^2 \sin^2 \theta)x + (a_0^3 \sin^3 \theta - a_0 \sin \theta)x^2 + \mathcal{O}(x^3). \quad (6.9)$$

Keeping terms up to order x , the average pad charge asymmetry over a single pad cycle is:

$$\langle A \rangle \approx \left(1 - \frac{a_0^2}{2}\right)x, \quad (6.10)$$

since $\langle \sin \theta \rangle = 0$ and $\langle \sin^2 \theta \rangle = \frac{1}{2}$ over a pad cycle. Note that Eq. 6.10 vanishes in the limit $x \rightarrow 0$ to any order as expected. From direct data measurements, $a_0 \approx .68$. Then since the h parameter is a direct measure of $\langle A \rangle$, Eq. 6.10 can be used to find corrective factors on the pad charges to cancel the observed gains. In the case of the anode calibration, it is clearly necessary to use Z_{anode} in Eq. 6.1 because the stretching and displacement along the Z_{anode} axis are of particular interest. When only the displacement along the A axis is of interest, however, it is possible to use the more precisely determined coordinate from fitted

track trajectories, Z_{track} . Figure 6.9 shows an example of such a fit. Figure 6.10 exhibits some example fitted values of h (those for the hood pads) before and after calibration. The initial dispersion of $\sigma_{\langle A \rangle} = .04$ or $\sigma_Z = 300 \mu m$ is reduced to $\sigma_{\langle A \rangle} = .006$ or $\sigma_Z = 40 \mu m$. In fact, the pad calibration is performed for both Z_{anode} and Z_{track} . Figure 6.11 compares the fitted values of h from both measurements summed over all wires. The width of that distribution, $\sigma_{\Delta h} = .009$, provides an estimate of the systematic error in the calibration procedure, $\sigma_Z = 70 \mu m$.

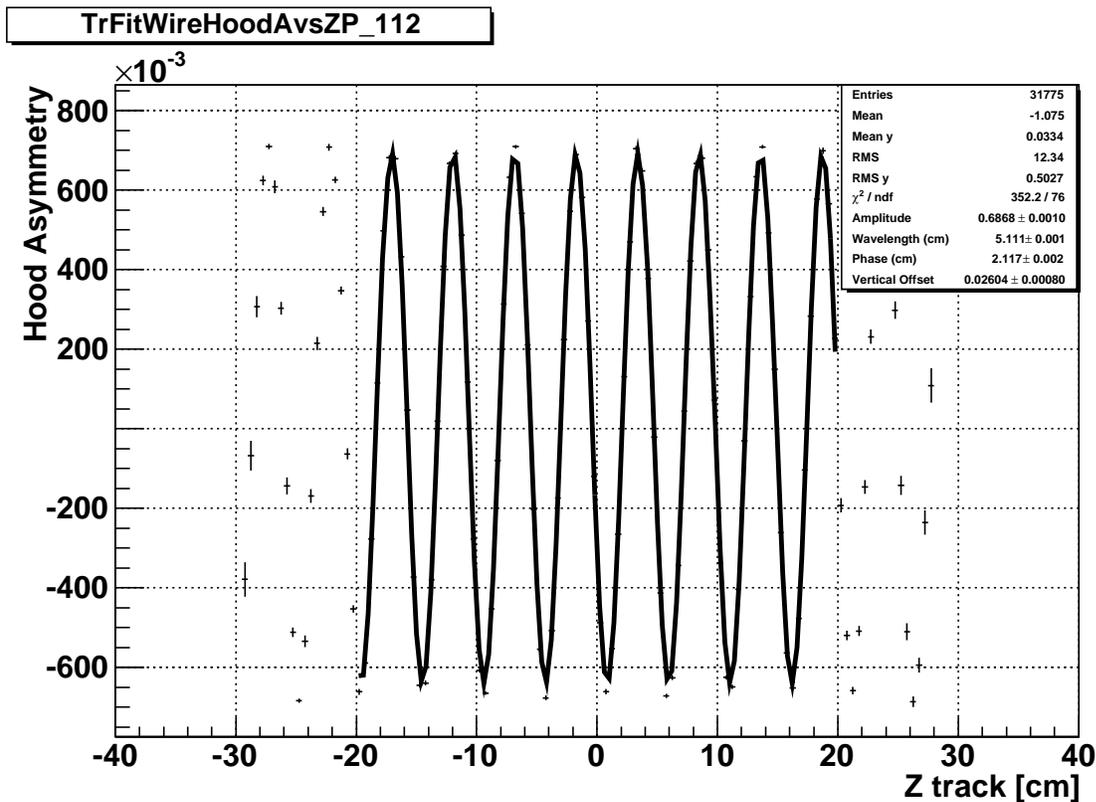
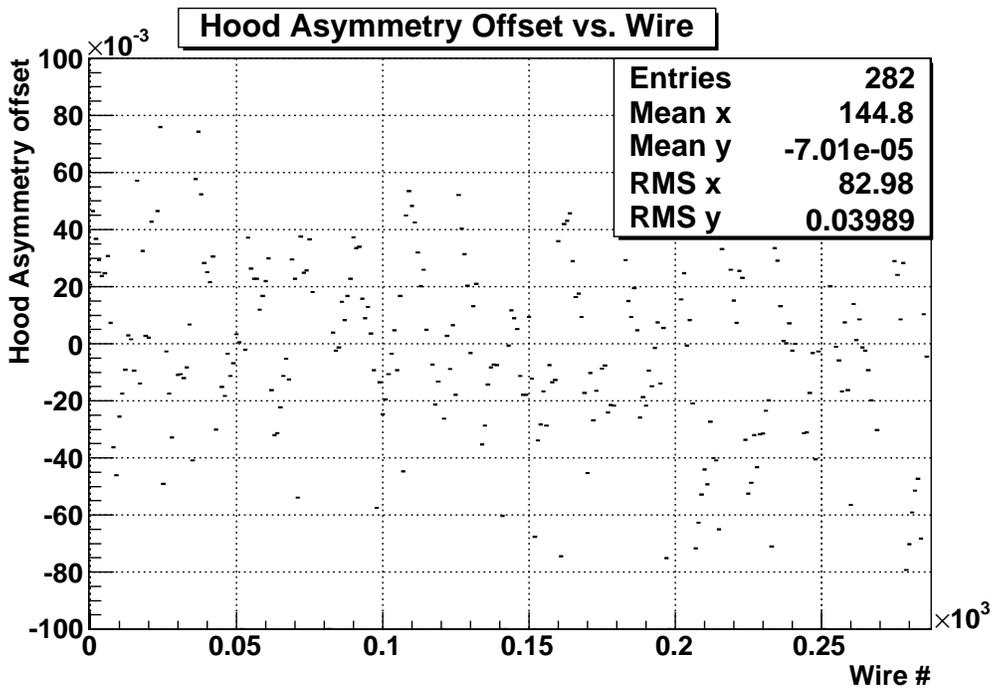
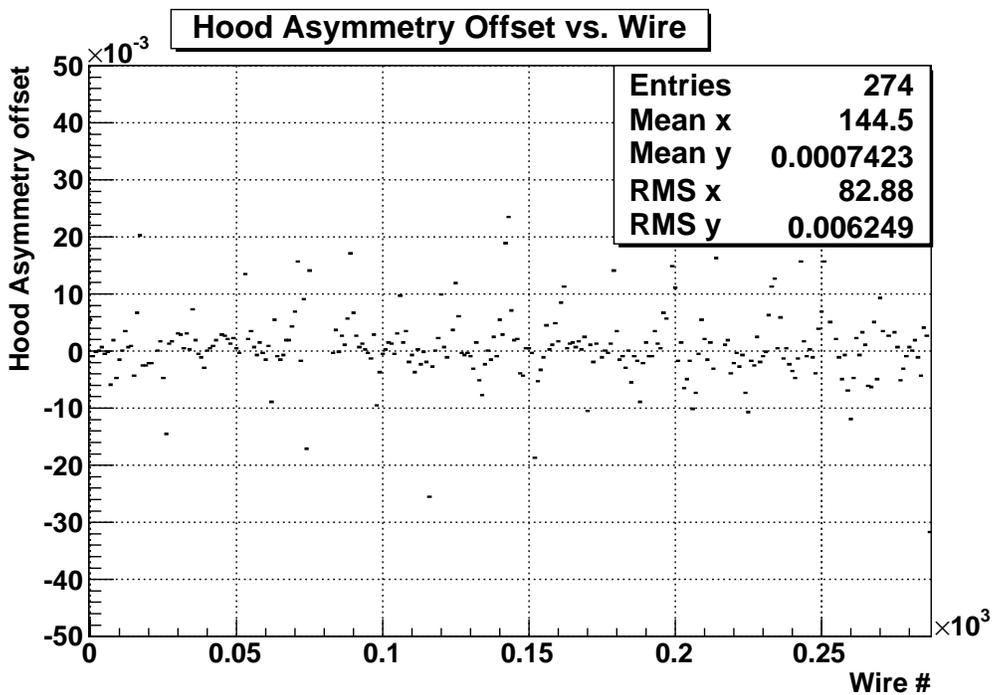


Figure 6.9: The same as figure 6.7 except that fitted track Z coordinates are used in place of anode Z coordinates. The advantage is that the fits are typically more stable.

The method of measuring the Z coordinate using the Vernier pads assumes that the charge asymmetry on both (cathode and hood) pads varies sinusoidally with equal amplitudes, that is a_0 in Eq.s 6.6 and 6.7 must be the same for both pads of a cell. That assumption is known to be violated by at least one effect. The chambers, being operated at 1.2 Pa above the pressure in the surrounding magnet volume, experience bowing of the outer hood foils.

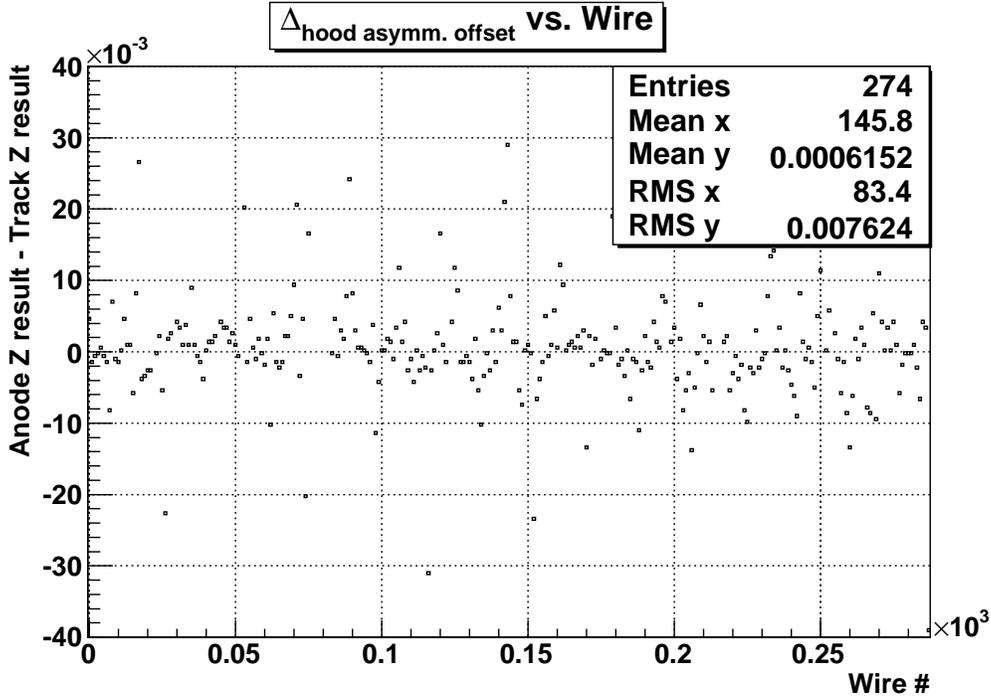


(a) A scatter plot of pre-calibration fitted values of h versus wire.

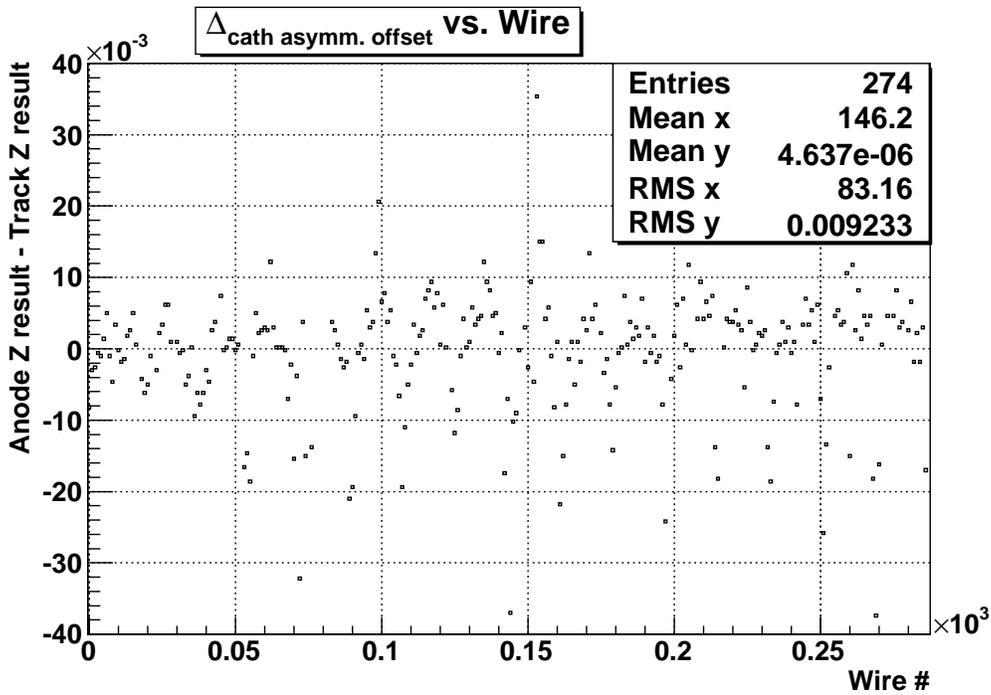


(b) A scatter plot of post-calibration fitted values of h versus wire.

Figure 6.10: A pictorial summary of improvement in relative pad gains by calibration.



(a) The difference in the values of h found by two different methods for the hood pads versus wire.



(b) The difference in the values of h found by two different methods for the cathode pads versus wire.

Figure 6.11: Scatter plots of the difference in h using anode determined Z coordinates and fitted track Z coordinates.

The bulge of the hood is largest near $Z = 0$ and at middle cell numbers because the foil is only attached at the edges of the chamber. This means that the two pad foils are not equidistant to the anode wire, but vary instead as a function of Z even within the same cell. The induced charge as well as the magnitude of the charge asymmetry also differ between the two pads depending on Z . The situation is further complicated by additional effects from changes in pressure with elevation and gravitational pressure on the pads. As a result, each wire is treated separately for calibration purposes. To correct this phenomenon, one needs the ratio of the pad asymmetry magnitudes, $\frac{A^{cathode}}{A^{hood}}$, as a function of Z on each cell. From the distribution of charge asymmetry versus Z coordinate, one can extract the asymmetry magnitude every 5 cm by making use of the relation between the amplitude and RMS over a full period of a pure sine wave:

$$amplitude = RMS \times \sqrt{2}. \quad (6.11)$$

This information is, in principle, sufficient; however, trends in $\frac{A^{cathode}}{A^{hood}}$ along Z are, in practice, hidden by large statistical fluctuations and this approach is not feasible. On the other hand, $\frac{Q^{cathode}}{Q^{hood}} = \frac{Q_u^{cathode} + Q_d^{cathode}}{Q_u^{hood} + Q_d^{hood}}$, which is related to $\frac{A^{cathode}}{A^{hood}}$ for geometric reasons, does exhibit a clear, expected dependence on Z . As in figure 6.12, such plots are fitted to a cubic polynomial in $|Z|$ of the form:

$$\frac{Q^{cathode}}{Q^{hood}}(Z) = a + b|Z|^2 + c|Z|^3. \quad (6.12)$$

By symmetry, the charge ratio must be invariant under $Z \rightarrow -Z$. Then, in order for the derivative to exist at the origin: $\left. \frac{d}{dz} \frac{Q^{cathode}}{Q^{hood}} \right|_{z=0} = 0$. This is guaranteed by the omission of a linear term in Eq. 6.12. In this instance, the relative gain between hood and cathode charges is relevant and is obtained from the constraint that the charge induced on either pad should be equal at the end of the wire where they are equidistant to the anode wire by virtue of

ChargeRatio Wire= 129

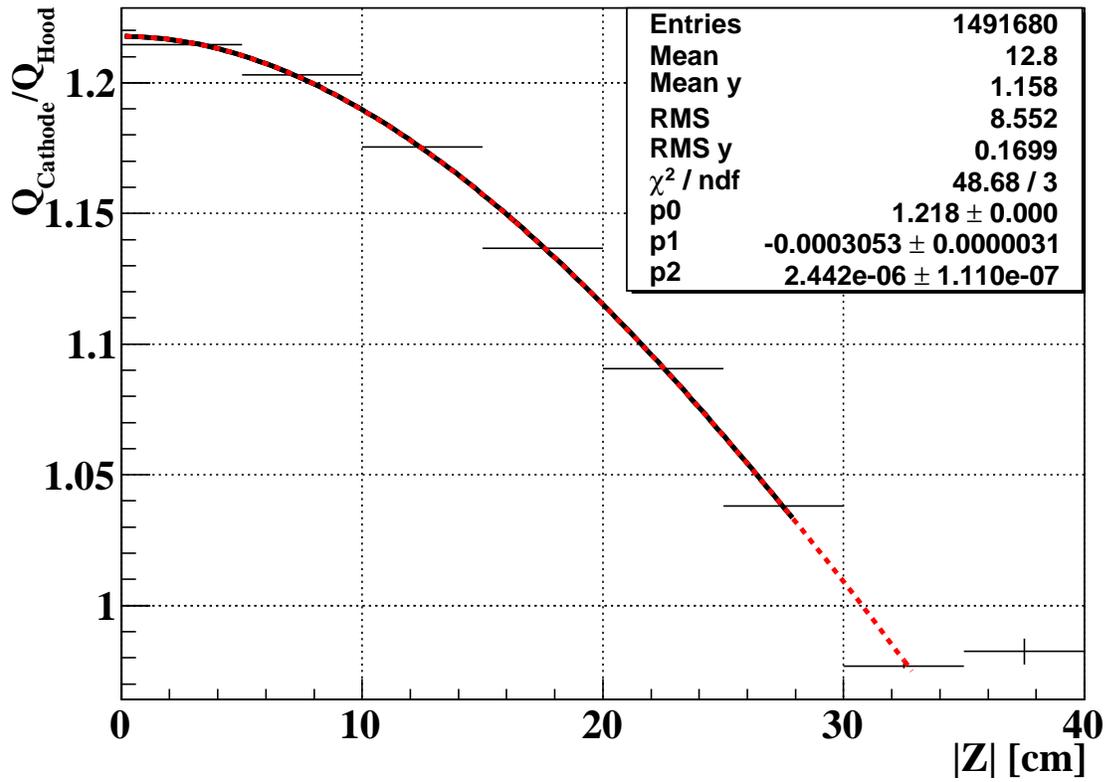


Figure 6.12: A profile histogram of the charge ratio versus $|Z|$. This is fit to a cubic polynomial (solid black), which is extrapolated to the end of the wire (dashed red) to correct for the relative cathode-to-hood gain.

the chamber construction. Thus for a wire of length L , the gain corrected charge ratio must obey: $\frac{Q^{cathode}}{Q^{hood}}(\frac{L}{2}) = 1$. For these measurements to be of use, the charge ratio must be related to the ratio of charge asymmetry magnitudes of the two pads. This relationship is extracted directly from the data from a plot of the charge ratio dependence of the charge asymmetry magnitude ratio, calculated from Eq. 6.11, summed over all wires and values of Z . As shown in figure 6.13, this is fit to a line forced to obey the physical constraint that $\frac{A^{cathode}}{A^{hood}} = 1$ when $\frac{Q^{cathode}}{Q^{hood}} = 1$.

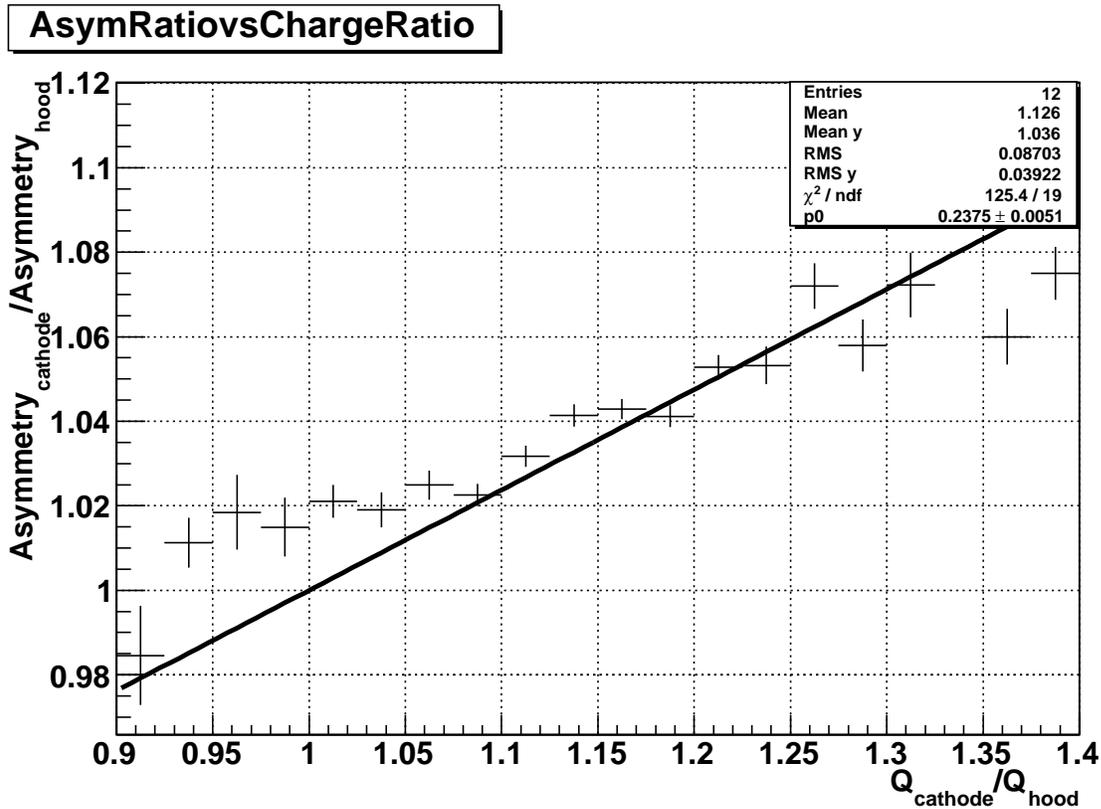


Figure 6.13: A fit to the asymmetry ratio dependence on the charge ratio, using data from all functioning wires. Enforcing the constraint that $\frac{A^{cathode}}{A^{hood}} = 1$ when $\frac{Q^{cathode}}{Q^{hood}} = 1$ leaves one free parameter, the slope of a line passing through that point. The slope is ~ 0.24 units of change in the asymmetry ratio per unit of change in the charge ratio.

Finally, a calibration is done to remove the effect of the bowing on the Z coordinate calculation during the reconstruction by finding the charge ratio for each hit based on an initial evaluation of Z and applying Eq. 6.12, then rescaling the ratio of asymmetries as they appear in Eq. 5.18 for a final evaluation of Z . Figure 6.14 exhibits the level of success in reducing

the correlation between $\frac{A^{cathode}}{A^{hood}}$ and $\frac{Q^{cathode}}{Q^{hood}}$ after calibration.

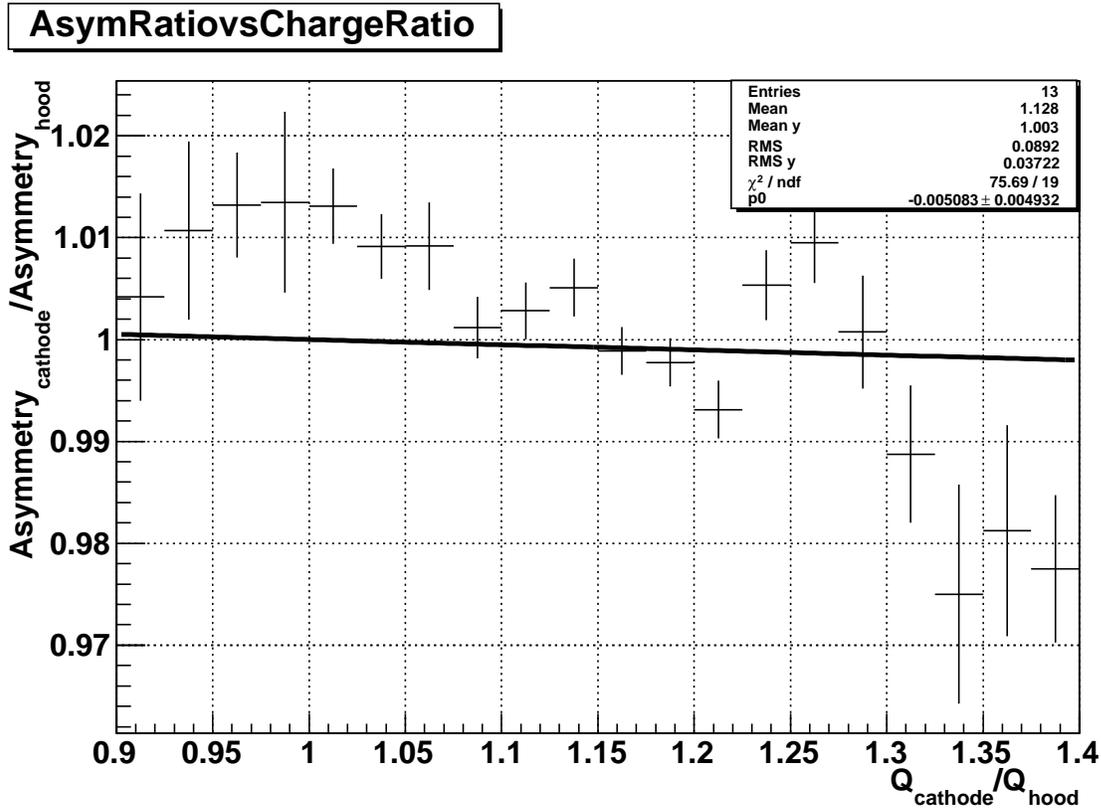


Figure 6.14: The same as figure 6.13 but after rescaling the hood asymmetry to correct for bowing. The constraint that $\frac{A^{cathode}}{A^{hood}} = 1$ when $\frac{Q^{cathode}}{Q^{hood}} = 1$ is again enforced. The slope is not statistically significant.

6.2.3 Chamber Alignment

The muon stopping target, drift chamber modules, and the support structure are outfitted with reference marks that are measured by an optical survey technique to determine their absolute positions. Figure 6.15 shows the locations of the target markers. A plane is fit to the measured target cross positions, which is then used to infer the orientation and location of the target plane. Figure 6.16 illustrates the placement of reference marks on the chamber modules and support structure. On top of each chamber module, a thin plate with a cross mark is glued at both the upstream and downstream ends. The chamber modules

are fastened between carbon blocks on the support structure. The blocks each contain two sets of pins. The line connecting the midpoint of the pins and the cross position gives the full orientation of the module in the XY plane for fixed Z. From the measured positions of the upstream and downstream crosses, the three-dimensional orientation is then deduced by treating the module as a rigid body. These measurements are ultimately translated into a set of wire positions and directions that are used in the reconstruction. The survey measurement precision is thought to be about $200 \mu\text{m}$ in Z and negligible in X and Y (about $20 \mu\text{m}$).

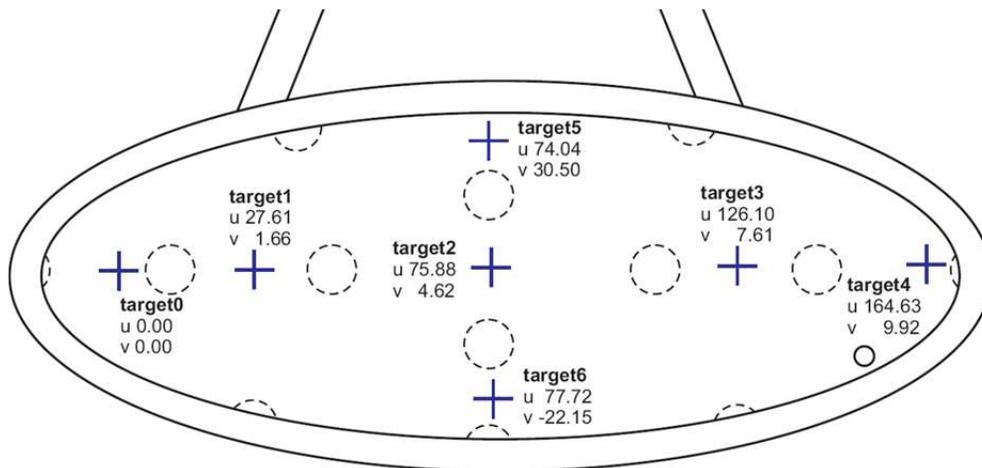
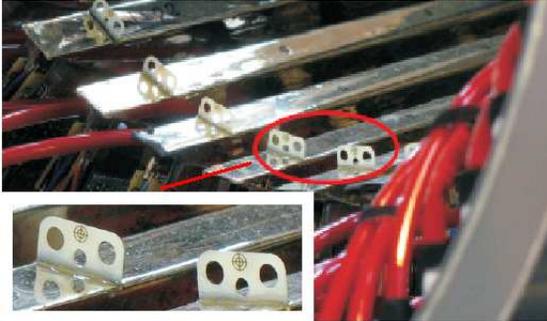


Figure 6.15: A diagram of the target marker locations.

The survey-determined chamber positioning is checked and refined by a technique based on the Millipede algorithm[60] using cosmic rays.

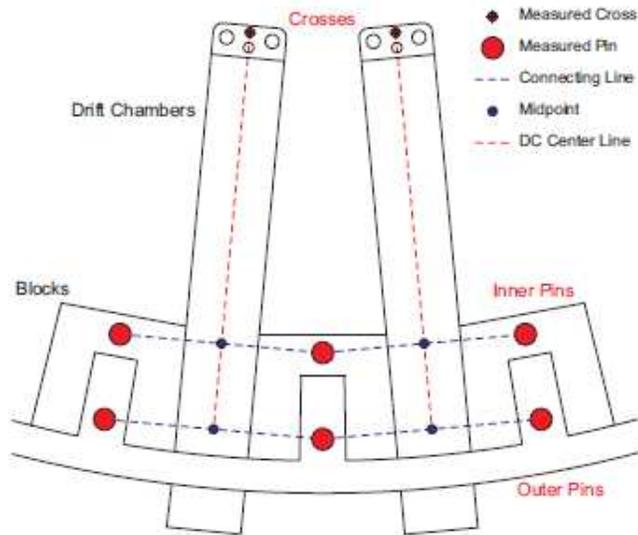
The Millipede alignment is cross-checked by a relative chamber alignment technique using Michel positrons. The radial chamber alignment is analyzed by taking three non-consecutive hit chambers on a track and using the circle defined by their $R - \phi$ coordinates to project, as described in section 5.2.3, to the single chamber they enclose and plotting the difference in the measured and projected R coordinate. Figure 6.17 shows this pull distribution for a single chamber as well as radial shifts to be applied as a function of chamber number. Because this diagnostic is insensitive to absolute chamber positions, the overall location of the support structure in real space is constrained by fixing the positions of the edge chambers



(a) A photograph of survey cross marks.

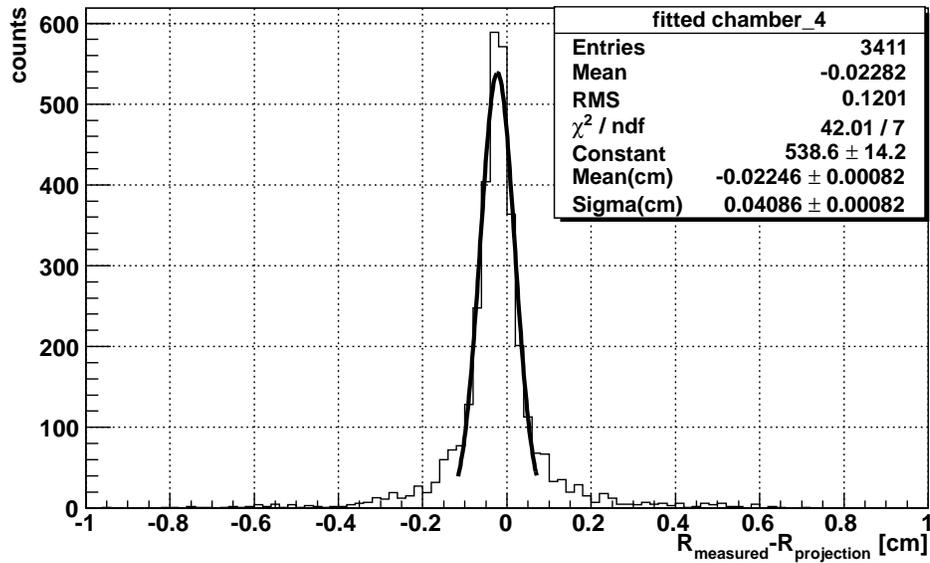


(b) A photograph of survey pins.

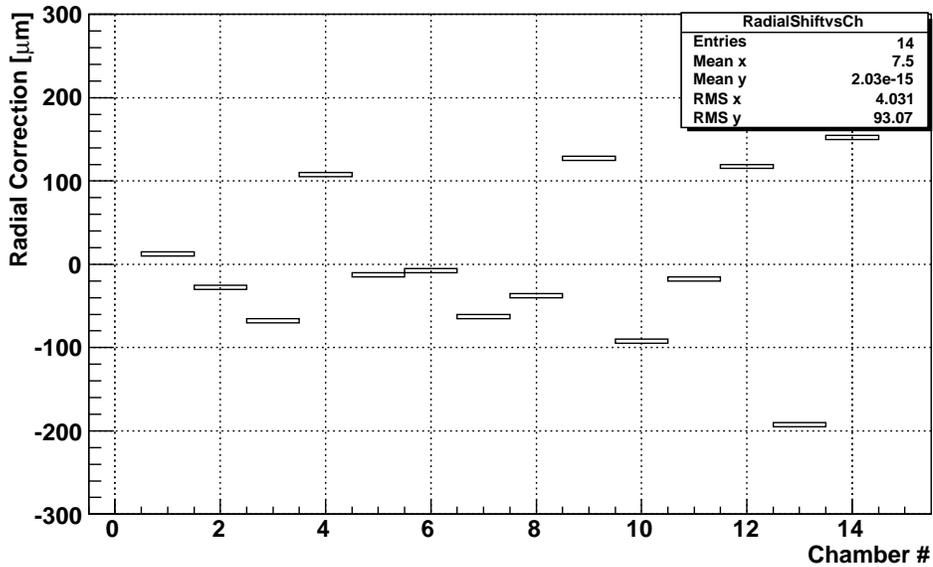


(c) A schematic of survey markers on the chamber modules.

Figure 6.16: Figures showing the location of chamber survey markers.



(a) A Gaussian fit to the radial pull distribution for a chamber.



(b) Radial corrections to be applied to each chamber before alignment.

Figure 6.17: Some plots illustrating the features of the radial chamber alignment.

(chamber 0 and chamber 15) to their optical survey values. The other 14 radial chamber positions are then shifted to bring the means of the pull distributions closer to zero. This procedure is iterated until the desired level of convergence is achieved. Figure 6.18 presents the level of convergence after the final iteration; the residual corrections are far below the single-hit position resolution.

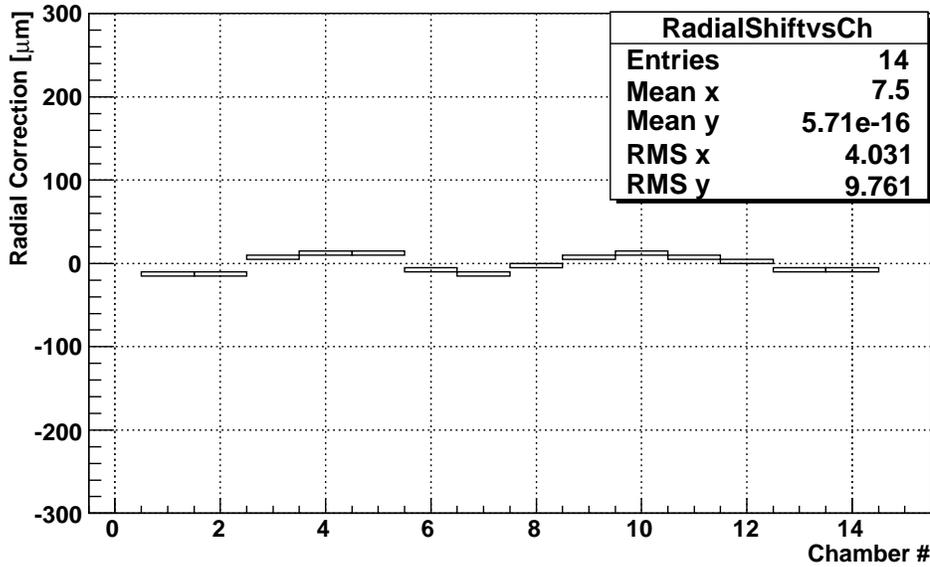
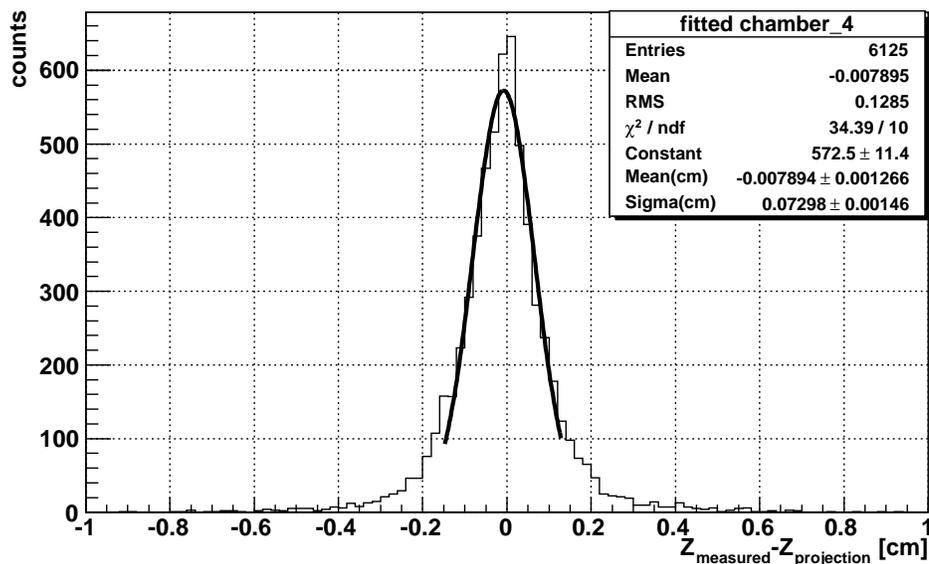


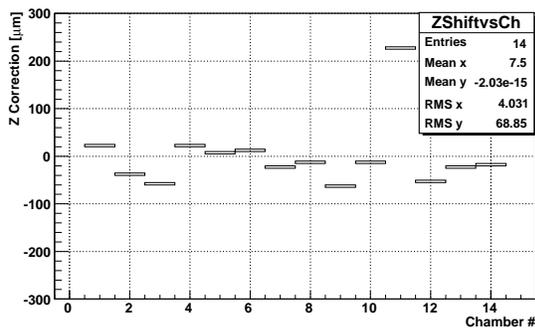
Figure 6.18: Radial corrections to be applied to each chamber after alignment.

Exactly the same process is applied to correct the relative chamber positions along Z, except that the pull distributions are formed with fitted tracks by plotting the difference between the z component of the state vector (see section 5.2.4) and the measured Z coordinate of the corresponding hit. Figure 6.19 illustrates various features of this alignment.

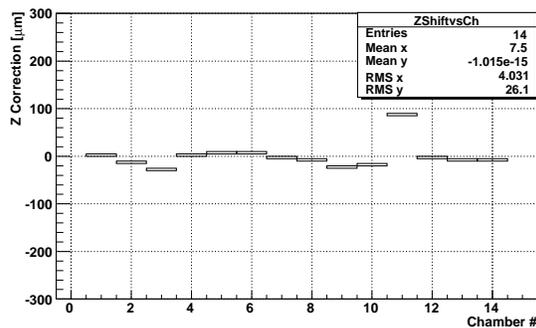
A similar set of diagnostics are utilized to search for tilting of the chamber modules beyond what is captured in the optical survey. The radial pulls are plotted as a function of the measured Z coordinate for each chamber. A uniform tilt in the R-Z plane (a rotation about the ϕ -axis) is signaled by a line with a slope significantly different than zero. Lastly, the radial pulls are plotted as a function of track angle for each wire. Figure 6.20 defines the



(a) A Gaussian fit to the Z pull distribution for a chamber.



(b) Z corrections to be applied to each chamber before alignment.



(c) Z corrections to be applied to each chamber after alignment.

Figure 6.19: Some plots illustrating the features of the Z chamber alignment.

track angle. This diagnostic is sensitive to displacements of the cell in the X-Y plane (and equivalently chamber tilts in this plane): (1) Track angles near zero show little effect, (2) the radial pull distribution mean tends toward larger absolute values as the track angle moves away from zero, (3) and track angles with opposite sign have an average pull with opposite sign. Figure 6.21 presents some representative examples of these plots. They are all consistent with no such chamber tilts or transverse displacements, so no corrections are made.

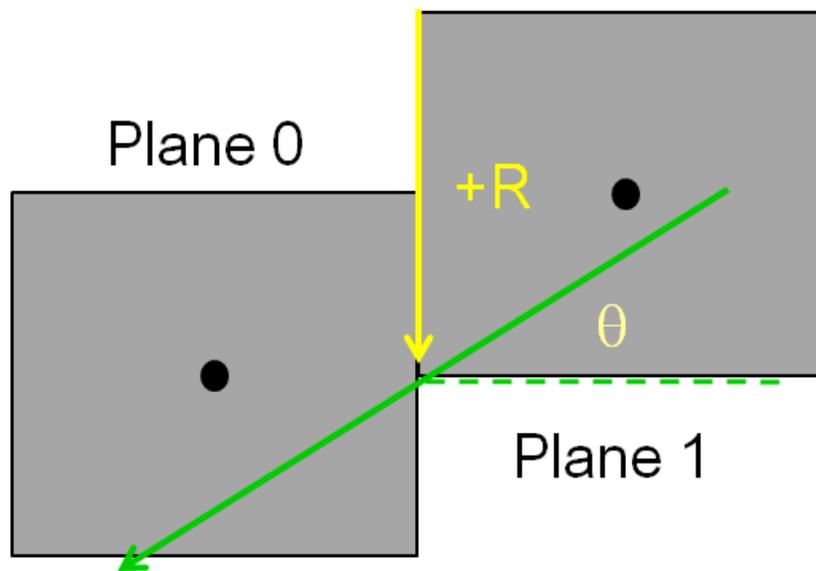
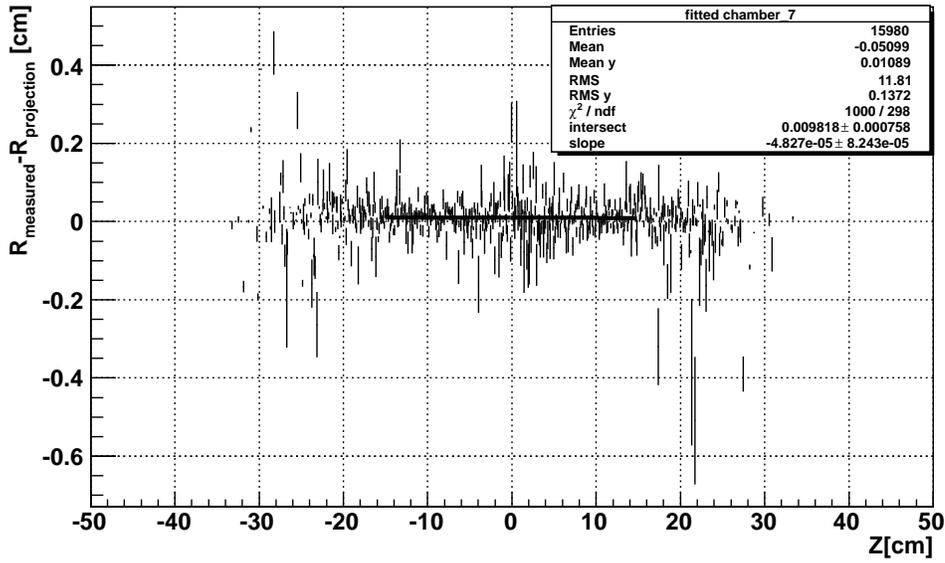
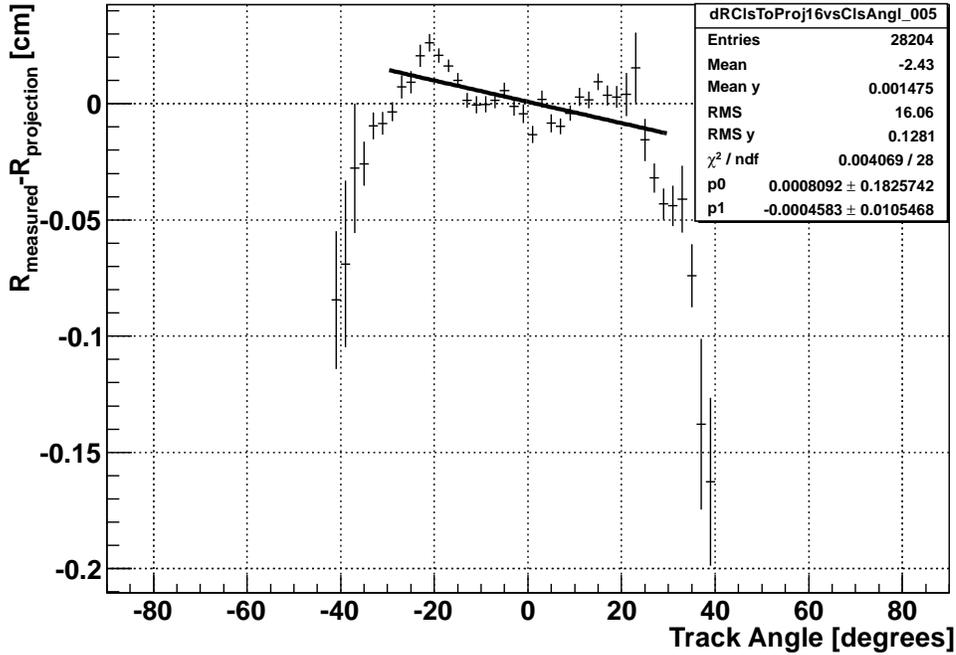


Figure 6.20: The track angle is the angle in the X-Y plane made with respect to the normal to the central chamber plane as marked by θ in this diagram. It is positive when the track is headed toward positive R, zero when the track is normal to the central chamber plane, and negative otherwise.

Both the Millipede alignment using cosmic rays and the Michel alignment have some shortcomings. An alignment using tracks at fixed track angle with respect to a chamber plane is insensitive to a translation of the chamber in the direction of the tracks, hence any alignment scheme requires a range of track angles at each chamber to constrain the position perpendicular to the chamber plane. Rotations of a chamber about the z-axis, which change the ϕ coordinates of different wires in the chamber by different amounts, require a range of track



(a) A line fit to the radial pull versus Z for a chamber.



(b) A line fit to the radial pull versus track angle for a chamber. There are wiggles that are not consistent with chamber offsets and possibly due to the time to distance relationship with track angle. The points are given equal weight in the fit to account for these effects.

Figure 6.21: Some plots illustrating the features of the rotational and transverse chamber alignment.

angles for each wire in the chamber to measure. One disadvantage of the cosmic ray data is that the range of track angles and the number of wires with high hit rates is limited. Because the cosmic ray data is collected with the magnetic field turned off, it also cannot detect a misalignment of the assumed magnetic field map with respect to the assumed location of the drift chamber system. An advantage of the Millipede algorithm is that the relative chamber alignment is less sensitive to the assumed starting positions of the chambers.

6.3 XEC

6.3.1 PMT Gain

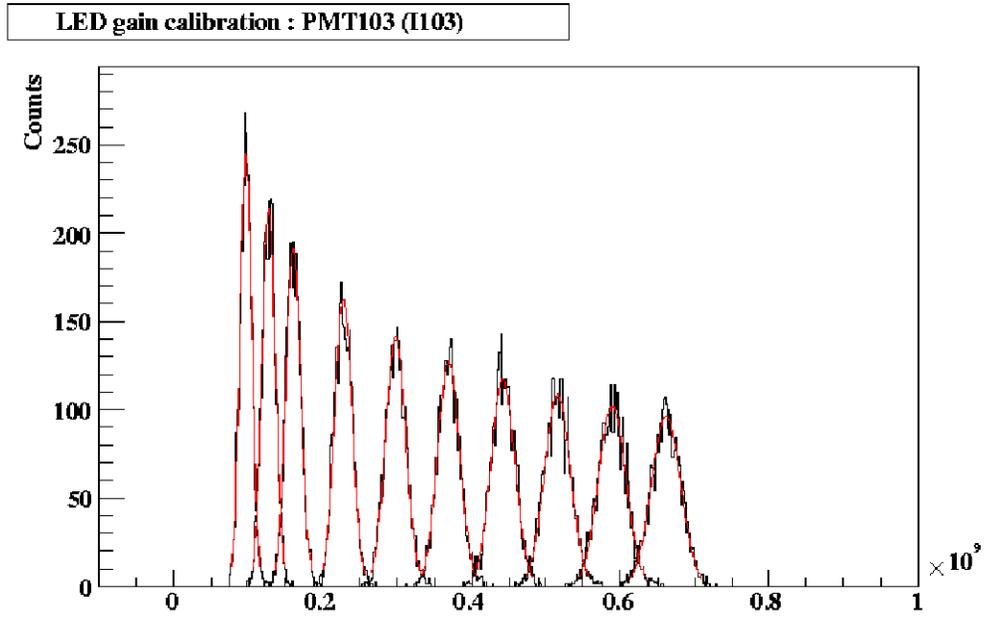
The PMT gain calibration presumes that the LEDs function as a fixed source of photons, and that the resulting number of photoelectrons produced in the photocathode of a recipient PMT is Poisson-distributed. By these assumptions, the mean number of photoelectrons is $\bar{N}_{pe} = N_{pho} \times QE$, the variance is $\sigma_{N_{pe}} = \sqrt{\bar{N}_{pe}}$, and the measured charge is $Q = Gain \times N_{pe}$, where N_{pho} is the number of photons in an event and N_{pe} is the number of photoelectrons from some PMT. The uncertainty in the measured charge can be modeled as the union of two uncorrelated parts: statistical fluctuations in N_{pe} and electronic noise effects. In this model,

$$\bar{Q} = Gain \times \bar{N}_{pe} \tag{6.13}$$

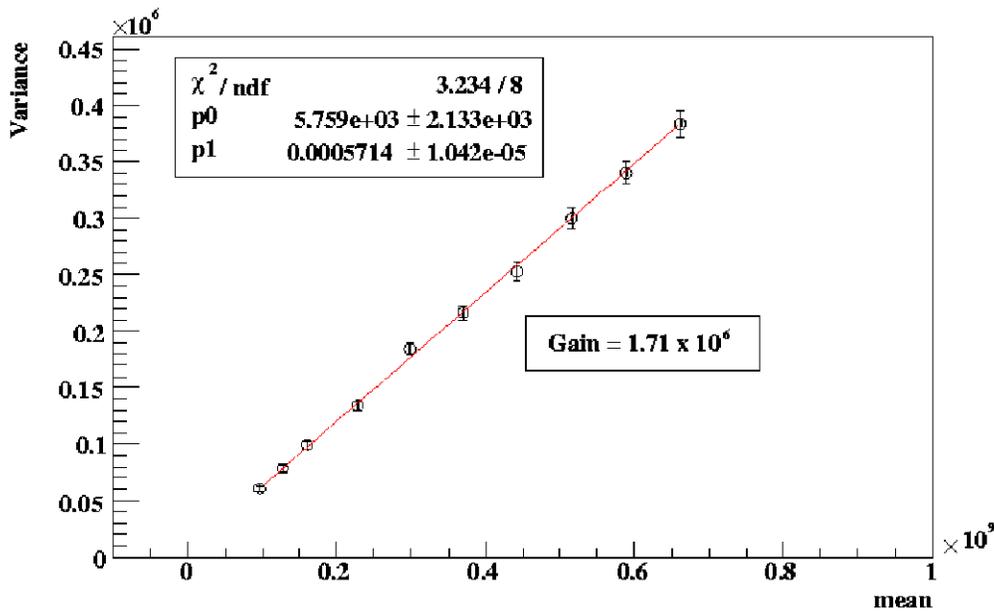
$$\sigma_Q^2 = Gain \times \bar{Q} + \sigma_{noise}^2. \tag{6.14}$$

A large sample of events is acquired for each of ten different, fixed LED intensities. Each of the ten distributions of N_{pe} is fit for the mean and variance to supply a data point for

a graph of σ_Q^2 versus \bar{Q} . Figure 6.22 shows these distributions. The slope of this plot is a measure of the PMT gain.



(a) An example of superimposed charge distributions for ten different LED intensities for a single PMT.



(b) A scatter plot of variance versus mean of the charge distributions for a single PMT. Each data point corresponds to a unique LED intensity.

Figure 6.22: Plots illustrating the PMT gain calibration method.

6.3.2 Quantum Efficiency

The dedicated α radiation sources mounted in the calorimeter serve as the reference for calibrating PMT quantum efficiencies. The expected amount of observed light from an α source at a fixed position with respect to a PMT is simulated by Monte Carlo. Discrepancies of the measurements are calibrated out by adjusting the quantum efficiencies of the PMTs.

The precision of the product $Gain \times QE$ is investigated by selecting groups of four PMTs arranged in a rectangular pattern on a given face and checking N_{pho} in each of the tubes, averaged over many events that are selected to be symmetric with respect to the 4 tubes. If the overall photon distribution is isotropic and homogeneous, then the illumination at the four tubes should be equal. Deviations from unity of the ratio of the average number of photons in one of the PMTs to that of all four provide a measure of $\sigma_{Gain \times QE}$ for that PMT. The uncertainty, $\sigma_{Gain \times QE}$, inferred from this procedure is generally below 5%, which is estimated to produce a contribution to the photon energy resolution of about 0.6%.

6.3.3 Timing Calibration

The PMT time delays are taken from charge exchange data ($\pi^0 \rightarrow \gamma\gamma$) with the NaI detector serving as a reference timer. To determine the various parameters involved in Eq. 5.11, v_{eff} is initially set to 10 cm/ns, the expected group velocity for the spectrum of scintillation photons in liquid xenon, and the Δt_{direct} term is accordingly corrected out. The $\Delta t_{scatter}$ piece is then removed by modeling the dependence of Δt_{delay} on the angle η and calibrating it away. Any residual systematic correlation between Δt_{delay} and N_{pe} for this PMT is treated as a time-walk effect. The consistency of the resulting calibration is checked by looking for any dependence of t_{XEC} on the total number of photoelectrons in an event, N_{pe}^{total} . A linear correlation is observed and then eradicated by tuning $v_{eff} = 8$ cm/ns.

After t_{delay} is fully corrected, $t_{offset,i}$ in Eq. 5.10 is obtained for each PMT by aligning $t_{interaction,i}$ among them.

6.4 TIC

6.4.1 Z-coordinate calibration

Relative time offsets between the two PMTs coupled to a single bar are calibrated by adjusting them to center the $\Delta Z_{DCH-TIC}$ distributions on each bar. The effective velocity of scintillation light in the bars is measured from data by again using the extrapolated Z position of the track at the bar from the DCH measurements as a reference and measuring the slope of it versus the difference in PMT times at the two ends of the bar.

6.4.2 Bar-to-bar time offsets

The time offsets of one bar relative to another are calibrated using Michel events in which a positron crosses two bars. The time measured at one bar can be compared to the other by correcting for the time of flight between bars, which is estimated event-by-event, typically 200 ps. The time offsets are adjusted so that all downstream bars are aligned with bar 0, and all upstream bars are aligned with bar 15. The upstream and downstream timing counter bars are aligned with each other and the XEC using Dalitz data. The effectiveness of this technique is confirmed on Boron data ($p + B \rightarrow \gamma(4.4 \text{ MeV}) + \gamma(11.7 \text{ MeV}) + C$). Events where the 4.4 MeV photon is detected in the calorimeter and the 11.7 MeV photon collides with a TIC bar are selected, and the mean time difference between the two simultaneous photons is plotted as a function of bar number (see figure 6.23). The residual dispersion in this time difference is at the level of 28 ps among upstream bars and 49 ps among downstream

bars. A residual timing offset between the upstream bars (15-29) and downstream bars (0-14) is apparent in the plot. It is also seen as an offset in the muon radiative decay timing peaks for upstream and downstream events as discussed in section 8.2.1, and it is accounted for in the likelihood analysis as discussed in section 9.2.3.

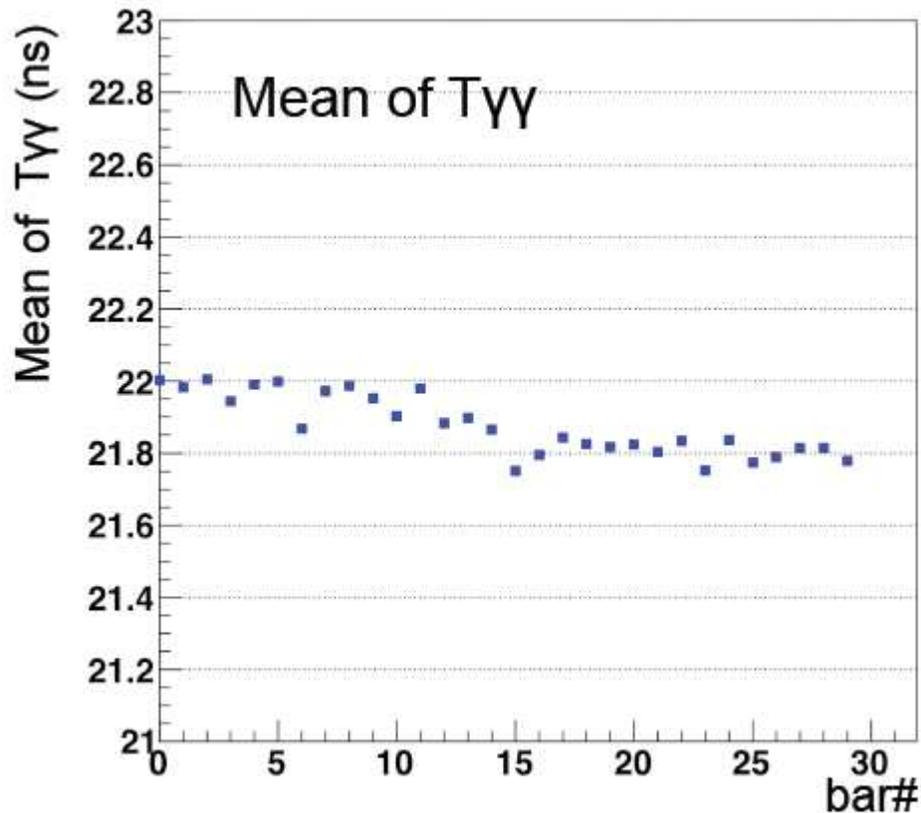


Figure 6.23: Scatter plot of time offsets versus bar number.

6.4.3 Time Walk

The severity with which pulse-height slewing affects the leading edge extraction is actually visible in the template waveform shape, as described in section 5.3.1, used in the fit. The time dependence on $x \equiv \frac{\text{low threshold}}{\text{pulse height}}$ is fit to a function with three floating coefficients:

$$T(x) = A + B\sqrt{x} + C \log x. \tag{6.15}$$

The result is then used on an event-by-event, and PMT dependent basis to correct for the fact that as the pulse height increases, the fitted leading edge time would otherwise systematically decrease. Typical values of the coefficients are $A=2.2$, $B=4.8$, and $C=0.2$, with $T(x)$ varying by ~ 2 ns over the range of interest of $x < 0.04$.

Chapter 7

Event Selection

This chapter reports the selection criteria placed on both the positron and the photon to reduce poorly reconstructed events. The quality of the resolutions vary with respect to various event properties, but this is largely incorporated into the likelihood analysis with appropriate event-by-event PDFs, whose details are discussed in chapter 9. Thus, the strategy for choosing selection criteria is based primarily on eliminating those types of events upon which claiming a discovery would be dubious, and not on making cuts with small acceptance losses to improve the resolutions.

7.1 Photon Cuts

7.1.1 Fiducial Volume

When a photon interacts near the edges of the calorimeter, the quality of reconstruction deteriorates due to increased energy leakage and PMT saturation. These events are removed by requiring the first conversion point to satisfy: $|u| < 25$ cm and $|v| < 71$ cm, which

excludes half the sensitive area of the outermost PMTs in u and v . The entire sensitive depth is retained: $0 < w < 38.5$ cm. These cuts are illustrated in figure 7.1. Table 7.1 gives the acceptance losses for these cuts.

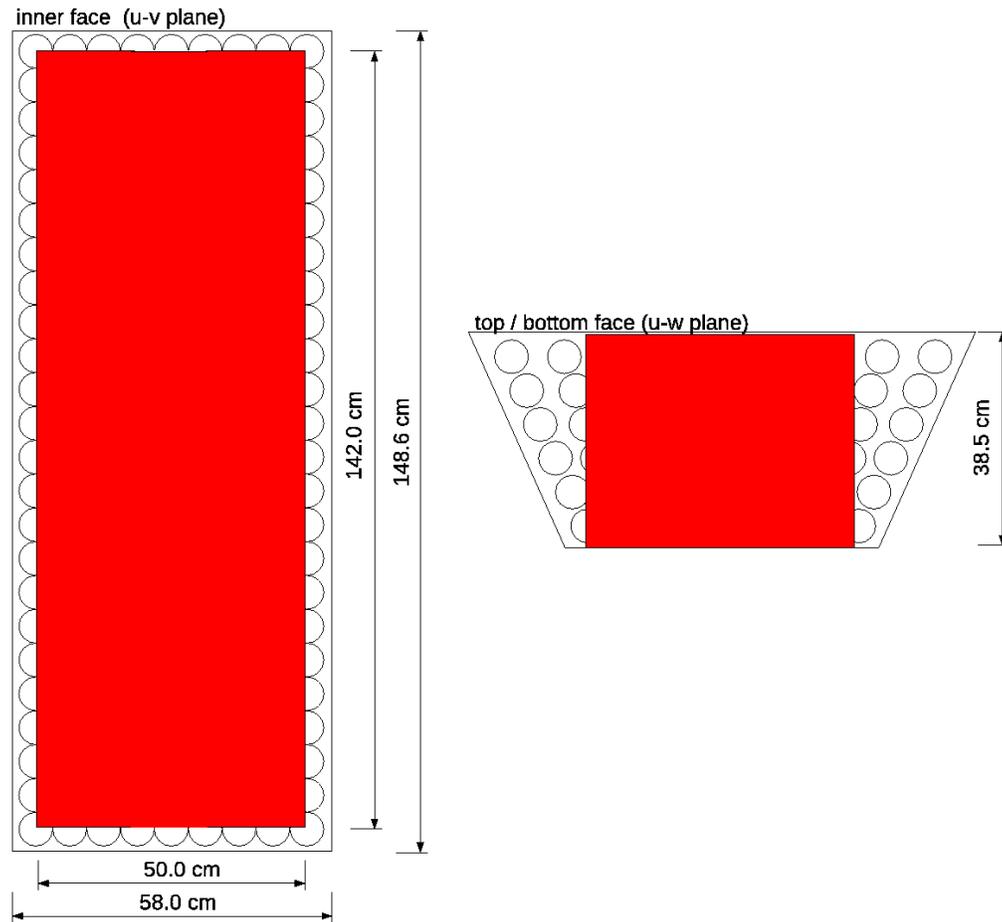


Figure 7.1: The accepted regions of the calorimeter are shown in red in the u - v plane (left) and the u - w plane (right).

7.1.2 Cosmic Ray Veto

Photons originating from the stopping target tend to interact close to the inner face and strongly illuminate the PMTs there. Deep events, and those which strongly illuminate the outer face PMTs, are characteristic of cosmic rays entering the calorimeter through the outer

shell. To reject these events, a cut is made on a combination of the ratio of inner face to outer face scintillation photons, $\frac{N_{inner}}{N_{outer}}$, and conversion depth: $-420\frac{N_{inner}}{N_{outer}} + 120 < w < 15\frac{N_{inner}}{N_{outer}} + 15$. This effectively places a minimum requirement on $\frac{N_{inner}}{N_{outer}} > 0.194$. From Monte Carlo, this cut is estimated to be 99% efficient and to reject 56% of cosmic rays.

7.1.3 Pileup Rejection

Section 5.1.2 describes how pileup events are handled when two peaks in the inner and outer face light distribution are detected. An additional requirement is made on these events that: $0 < \frac{E_{\gamma}^{normal} - E_{\gamma}^{pileup\ corrected}}{E_{\gamma}^{normal}} < 0.1$. Here E_{γ}^{normal} is the total deposited energy and $E_{\gamma}^{pileup\ corrected}$ is the energy after removing the secondary shower centroid. This cuts events in which the secondary photon contributes an unusually large number of scintillation photons. In the event that pileup is recognized by the $\hat{\chi}_{time}^2$ criteria of section 5.1.2 but not by the presence of multiple peaks in the light distribution, the event is considered irrecoverable and rejected. From Monte Carlo, the efficiency of the pileup rejection is estimated to be 95.5%.

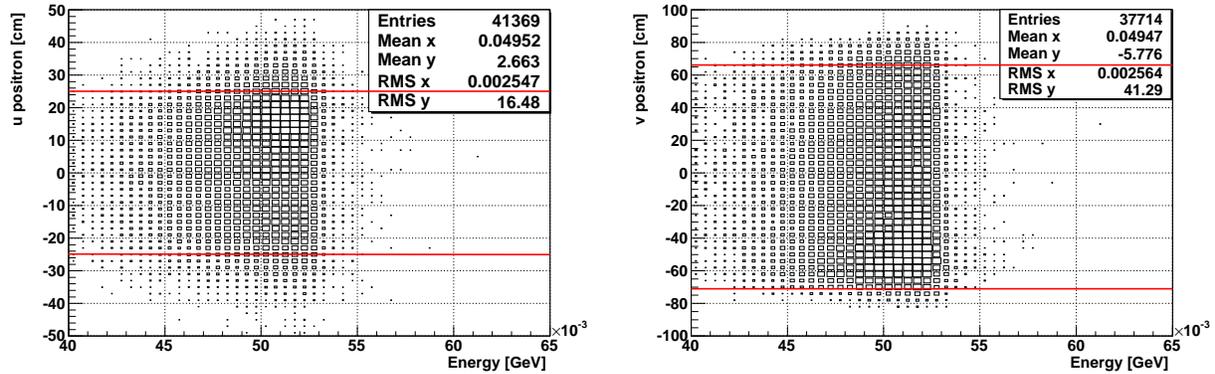
7.2 Positron Cuts

In this section, each plot is made by imposing all cuts on the positron except the one being examined.

7.2.1 Fiducial Volume

The direction of the positron momentum vector must be oriented such that a back-to-back photon enters the calorimeter acceptance defined in section 7.1.1. Since this cut must be made on positrons in the normalization sample, it must be made on MEG candidates as

well. Figure 7.2 shows these cuts overlaid on data distributions of those variables. Table 7.1 gives the acceptance losses for these cuts.



(a) Distribution of the positron u coordinate when backwards projected into the calorimeter versus positron energy.

(b) Distribution of the positron v coordinate when backwards projected into the calorimeter versus positron energy.

Figure 7.2: Distributions of the positron coordinates when backwards projected into the calorimeter. The regions between the red lines are retained by the cuts.

7.2.2 Drift Chamber Hit Pattern

A number of selection criteria are imposed based on the pattern of hits in the drift chamber. For the following hit and chamber counts, only reliable hits are counted; those which derive their Z coordinate from the anode wires because the pads do not function are excluded (see 5.2.1).

At least seven hits are required. This is a minimal cut, considering that in the simplified scenario of a charged particle moving in a uniform magnetic field, already five parameters are required to specify the helix trajectory: (x_0, y_0) coordinates at a reference plane $z = z_0$, the magnitude of the momentum, and two angles to specify the direction. At least four chambers must contain hits, the span of the track must cover at least five chambers, and at least two chambers must contain more than just a single hit. Tracks failing these criteria have very few

constraints, and tracks producing only a single hit in a certain chamber provide very little information for resolving its left/right ambiguity. Tracks making more than two turns in the drift chamber are further eliminated. Lastly, if the projected maximum radial coordinate of the fitted track exceeds that of the drift chamber modules, the event is rejected. This is because positrons that go through the carbon frame will have increased scattering and poor resolution. Figure 7.3 illustrates these cuts overlaid on data distributions of those variables. Table 7.1 gives the acceptance losses for these cuts.

7.2.3 Quality of Track Fit

The Kalman filter discussed in section 5.2.4 provides several indicators of the quality of the track fit. One is the normalized $\hat{\chi}^2$ that compares predicted and measured state vectors along the track. A requirement is made that $\hat{\chi}^2 < 12$. Estimates of the energy and angle uncertainties are also produced by the Kalman filter. A loose cut on the energy uncertainty is made, $\delta E_e < 1.1$ MeV, because the dependence of the likelihood PDFs on δE_e are incorporated into the analysis. The positron angle uncertainties are required to pass $\delta\phi_e < 1.5^\circ$ and $\delta\theta_e < 0.6^\circ$. The selections generally retain most of the distributions in those variables. Figure 7.4 displays these cuts overlaid on data distributions of those variables. Table 7.1 gives the acceptance losses for these cuts.

7.2.4 Stopped Muon Consistency

Additional cuts are made to ensure that the positron is consistent with coming from the stopping target. The reconstructed vertex is required to be within an ellipse that excludes the target frame by 5 mm. This is to avoid accepting events that originate in the Rohacell target frame. An even tighter cut is made to remove events falling outside an ellipse whose center and axes are chosen to retain $\sim 2\sigma$ of the beam stopping distribution, since events

failing this have increased probability of originating from a muon stopping in the helium gas. Two-turn events in which the first turn is missed during the reconstruction also have an increased chance of failing this cut. Figure 7.5 gives these cuts overlaid on data distributions of those variables. Table 7.1 gives the acceptance losses for these cuts.

7.2.5 Projection to Timing Counter

Cuts are also made on the closeness of the spatial projection of the track in the drift chamber to the measured timing counter impact location. Some requirements are made on the variables defined in section 5.4: $|\Delta R_{DCH-TIC} - 1.8| < 5$ cm and $|\Delta Z_{DCH-TIC}| < 12$ cm. The radial cut is off-centered to account for the non-zero mean observed in that distribution. Events well outside these cuts are likely to have undergone large scattering in going from the drift chamber to timing counter. Figure 7.6 presents these cuts overlaid on data distributions of those variables. Table 7.1 gives the acceptance losses for these cuts.

7.2.6 Ghost Selection

Because the tracking algorithm attempts to reconstruct all possible tracks in a high rate environment, more than one manifestation of the same track may be identified in a single event. For example, if two clusters of hits are found on a chamber, it may be possible to place either of them on and successfully fit an otherwise identical track. In this case, two nearly identical tracks may survive selection criteria. Accordingly, a procedure is developed to classify all manifestations or “ghosts” of the same track and to select among them. All of the fitted tracks in an event are compared with one another and any two are considered to be ghosts of the same underlying track if the following conditions are met:

1. They both derive their track time from the same timing counter cluster.

2. The number of chambers with at least one hit in common is more than half the total number of hit chambers on the shorter of the two tracks.

Each track in a set of ghost tracks is then assigned a ranking to indicate its quality of reconstruction. This rank, R , is assigned based on the $\widehat{\chi}^2$ of the track fit and the chamber span, S , of the track:

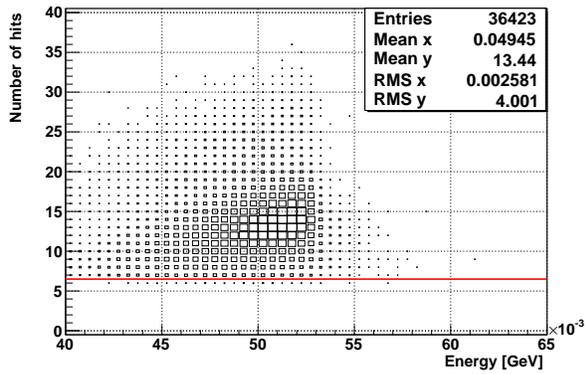
$$R = \frac{1}{S} + \alpha \widehat{\chi}^2. \tag{7.1}$$

$\alpha = .025$ is chosen so that the first term dominates the ranking when tracks have very similar $\widehat{\chi}^2$ and vice-versa.

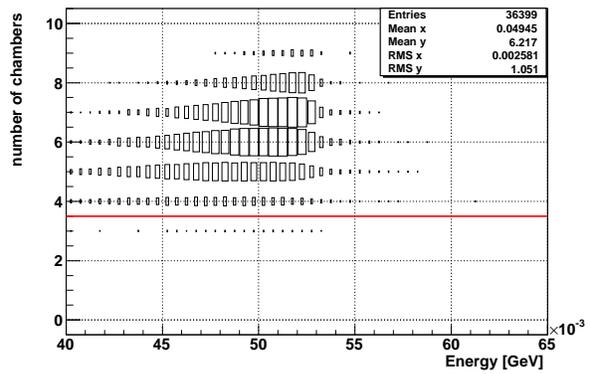
The selection procedure begins by eliminating all tracks in an event that do not pass the above selection criteria. In the case that more than one ghost of the same underlying track passes cuts, only the one with the smallest ranking, R , is selected. Thus, it is possible for more than one track to pass selection cuts in a single event only if they are not found to be ghosts of the same track.

Table 7.1: A table of inefficiencies for each cut. The individual positron cut inefficiencies are tabulated as the fraction of events rejected by making only the cut with respect to making no cuts. The combined losses are tabulated similarly for each group of cuts made simultaneously with respect to making no cuts. In tabulating the photon acceptance cut inefficiencies, the positron cuts, pileup cut, and cosmic ray cut are already applied.

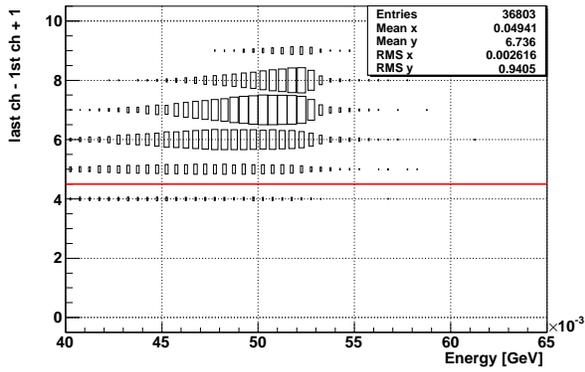
Class of Cuts (Combined Loss (%))	Cut Definition	Individual Loss (%)
γ Fiducial Volume (3.0)	$ u < 25$ cm	2.5
	$ v < 71$ cm	0.6
e^+ Fiducial Volume (26.1)	$ u < 25$ cm	14.9
	$ v < 71$ cm	12.7
DCH Pattern (8.6)	$N_{hit} \geq 7$	2.2
	$N_{chamber} \geq 4$	1.5
	Chamber Span ≥ 5	6.2
	$N_{multihit} \geq 2$	1.1
	$N_{turns} \leq 2$	0.1
Track Fit Quality (15.9)	R_{max} within DC module	0.5
	$\hat{\chi}^2 < 12$	1.6
	$\delta E_e < 1.1$ MeV	3.0
	$\delta\phi_e < 1.5^\circ$	14.1
	$\delta\theta_e < 0.6^\circ$	0.1
Vertex Cuts (16.3)	Target Frame Cut	12.0
	Beam Spot Cut	14.7
DCH-TIC Matching (9.2)	$ \Delta R_{DCH-TIC} - 1.8 < 5$ cm	2.3
	$ \Delta Z_{DCH-TIC} < 12$ cm	8.1



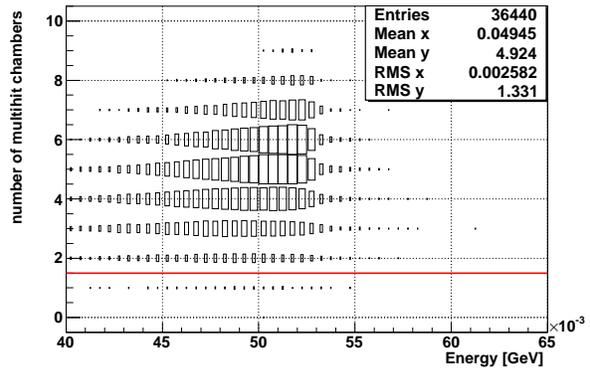
(a) Distribution of the number of hits versus positron energy.



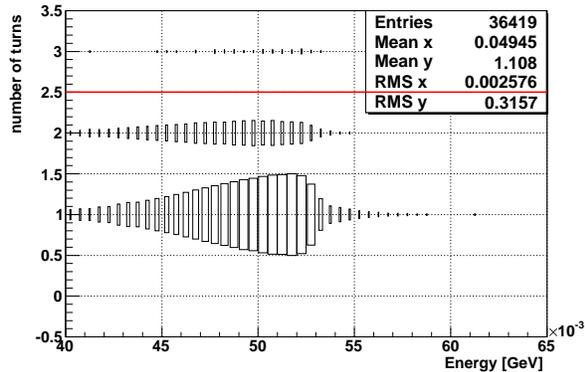
(b) Distribution of the number of hit chambers versus positron energy.



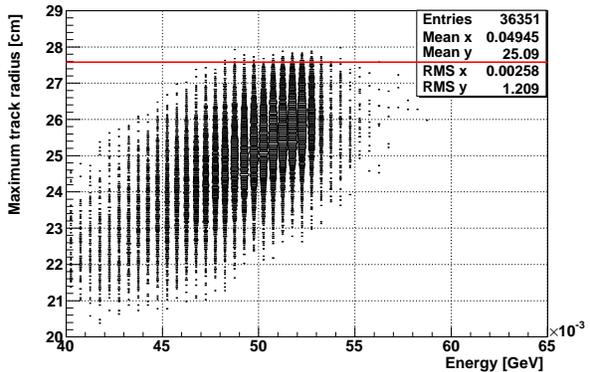
(c) Distribution of the chamber span versus positron energy.



(d) Distribution of the number of multiple hit chambers versus positron energy.

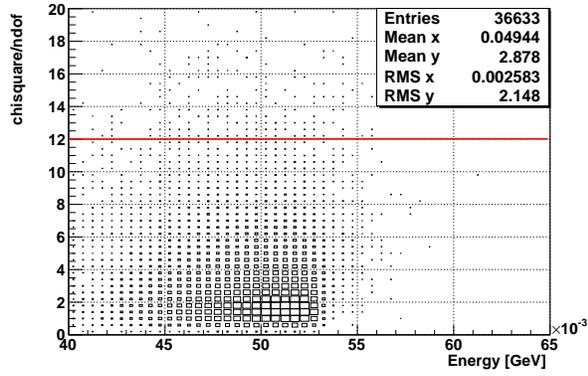


(e) Distribution of the number of turns versus positron energy.

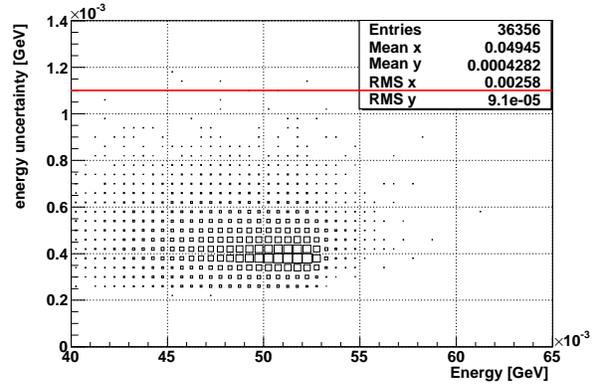


(f) Distribution of the projected maximum track radius versus positron energy.

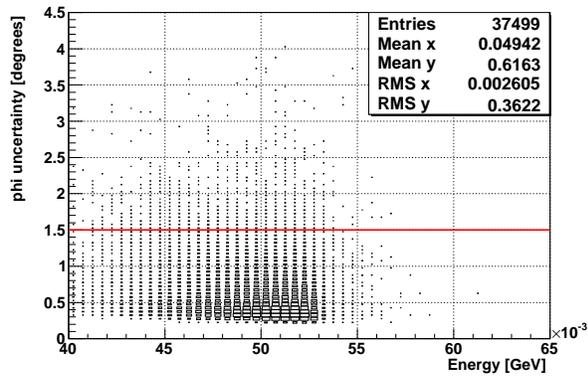
Figure 7.3: Distributions of various positron variables related to the hit pattern. The red lines mark the cuts.



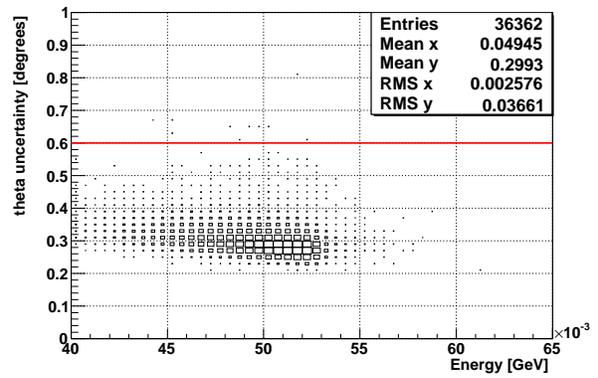
(a) Distribution of $\hat{\chi}^2$ versus positron energy.



(b) Distribution of δE_e versus positron energy.

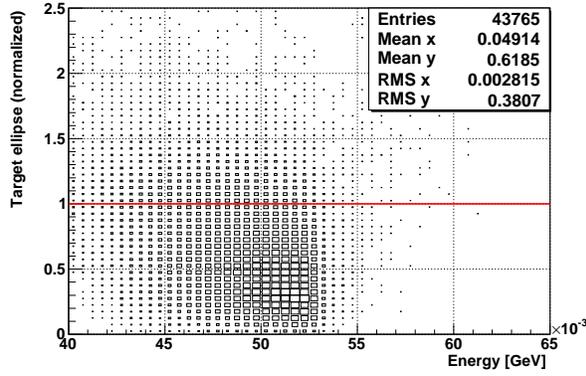


(c) Distribution of $\delta\phi_e$ versus positron energy.

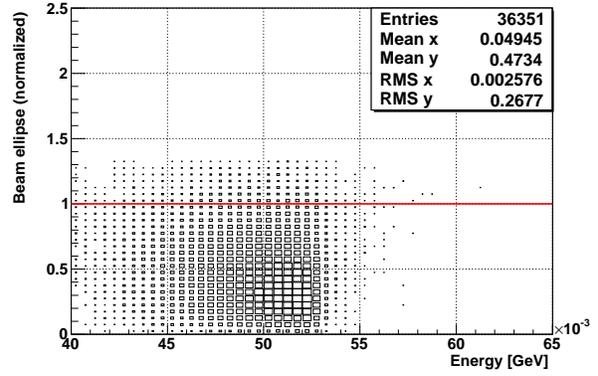


(d) Distribution of $\delta\theta_e$ versus positron energy.

Figure 7.4: Distributions of various positron variables related to track fit quality. The red lines mark the cuts.

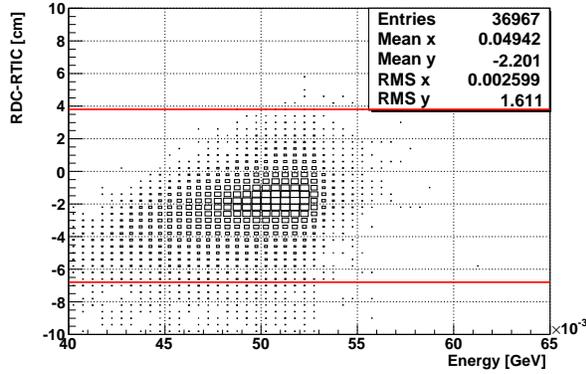


(a) Distribution of vertex locations normalized to the ellipse of the target frame cut versus positron energy.

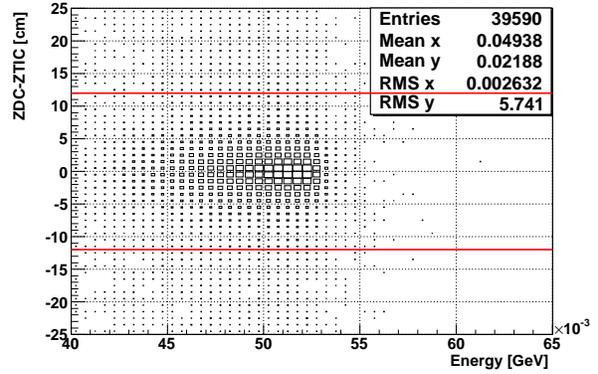


(b) Distribution of vertex locations normalized to the ellipse of the beam cut versus positron energy.

Figure 7.5: Distributions of vertex locations normalized to elliptical cuts. The cuts reject everything above the red lines.



(a) Distribution of $\Delta R_{DCH-TIC}$ versus positron energy.



(b) Distribution of $\Delta Z_{DCH-TIC}$ versus positron energy.

Figure 7.6: Distributions of variables related to DCH-TIC matching. The regions between the red lines are retained by the cuts.

Chapter 8

Hardware and Software Performance

A maximum likelihood analysis of the data requires a good knowledge of the detector resolutions and acceptances, which are combinations of many effects at both the hardware and software reconstruction level. This chapter summarizes the overall performances achieved, attempts to isolate contributing factors to the various resolutions and efficiencies, and explains the methods by which they are measured. The selection criteria of chapter 7 are applied in all cases unless an exception is explicitly stated.

8.1 Detector Resolutions

This section outlines the individual detector performances.

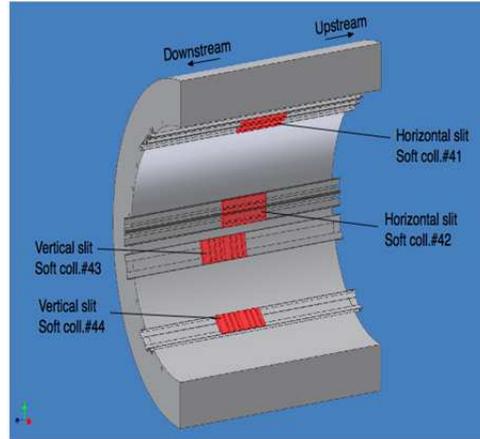
8.1.1 XEC

Position Resolution The position of the first conversion in the calorimeter is used to infer the photon emission angle. Thus, one needs the resolution in this measurement in

order to obtain the relative angle response functions. A set of charge exchange runs taken with lead bricks positioned directly in front of the calorimeter entrance window, functions as a basis for extracting the position resolutions along u and v from data. This setup is outlined in Figure 8.1. Figure 8.2 gives the reconstructed position distributions for the lead brick data. The projection of the v coordinate shows three peaks corresponding to the three slits. This distribution is modeled as three Gaussian resolution functions, two error functions representing the effects at the edges of the brick, and a flat component representing a set of events penetrating the brick that are uniformly distributed in v . Since the Gaussian parts receive contributions from the width of the slit and the spread of the pion decay vertex, they do not directly represent the position resolution. A Monte Carlo simulation of the lead brick setup is performed to infer the true position resolution. It predicts a Gaussian width that is systematically smaller than that obtained from the data, and the difference is assumed to be due to a resolution component from PMT quantum efficiency errors of about 1.8 mm. The position dependence of the v resolution is then taken from Monte Carlo simulation and a 1.8 mm resolution contribution is added in quadrature by assuming that it affects the v resolution in the same uncorrelated way at all positions. A similar procedure is applied to get the u resolution, which can also be checked with data, and to get the w resolution, which cannot be checked with data. Deep events have worse resolution since the light distribution on the inner face gets broader, and very shallow events suffer from PMT saturation. The best resolution also tends to occur when the light is distributed to more PMTs, which is controlled by the u and v location of the shower. The u and v resolutions are well fit to double Gaussians, both having an effective resolution of $\sigma_{u,v} \sim 5$ mm (70% of events in core with resolution of ~ 4 mm, tail resolution of ~ 8 mm), averaging over the position dependence. The w resolution is treated as a single Gaussian, which is $\sigma_w \sim 6$ mm when averaged over its position dependence. These are to be compared with goal resolutions of ~ 3.5 mm.

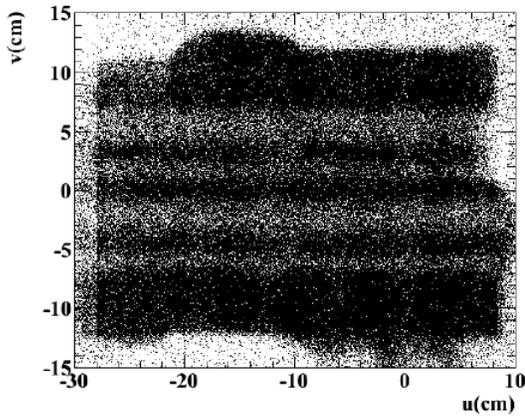


(a) Lead bricks with 1 cm wide slits. A PMT is shown for reference on top of the bricks on the right.

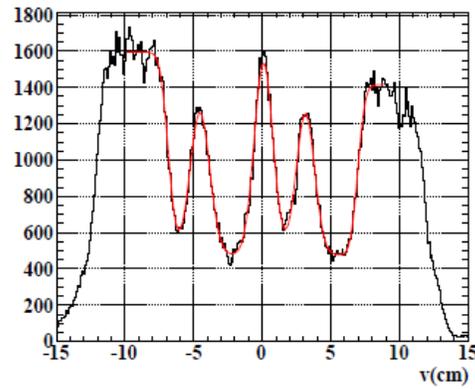


(b) Configuration of bricks at the calorimeter entrance window.

Figure 8.1: Lead brick apparatus for measuring XEC position resolution.



(a) Scatter plot of u, v positions.



(b) Projected distribution of v positions.

Figure 8.2: Measured distributions of u, v positions from the lead brick data.

Timing Resolution The time of the first conversion in the calorimeter directly enters the calculation of the relative positron-photon emission time, thus the resolution of this measurement constitutes a contribution to the resolution of the relative time as used in the likelihood analysis.

Intrinsic timing resolution An estimate of the intrinsic resolution of the time reconstruction is performed by dividing the PMTs into two groups and comparing the reconstructed conversion time from each PMT group event-by-event. There is a set of odd numbered PMTs and a set of even numbered PMTs as illustrated in Figure 8.3. One conversion time measurement, t_{odd} , is reconstructed based on only the measurements of odd PMTs, and another measurement, t_{even} , is made from even PMTs. The dispersion in the quantity $\frac{t_{odd}-t_{even}}{2}$ is mathematically equivalent to the uncertainty in $\frac{t_{odd}+t_{even}}{2}$, which is practically the same as t_{XEC} from section 5.1.2. This technique tests the precision of the PMT waveform time extraction and the dependence on photo-statistics and location within the calorimeter. Effects from the position reconstruction and the event-by-event shower spread are largely canceled out, however, because both PMT groups tend to experience a similar effect. When averaged over position, this gives Gaussian resolutions of 37.5 ps at 55 MeV and 30.5 ps at 83 MeV. Interactions occurring very close to the lateral faces (in u and v) or at very shallow depth (in w) have deteriorated time resolutions due to shower leakage and thus fewer photoelectrons.

Effective timing resolution To evaluate the full timing resolution of the calorimeter, simultaneous photons from $\pi^0 \rightarrow \gamma\gamma$ events are exploited by using the reconstructed time from the scintillators of the NaI detector of one photon as a reference with which to compare the XEC reconstructed time of the other photon. The full calorimeter volume is scanned by moving the NaI detector to different positions.

Figure 8.4 displays the difference between t_{XEC} and the NaI reference time over many events.

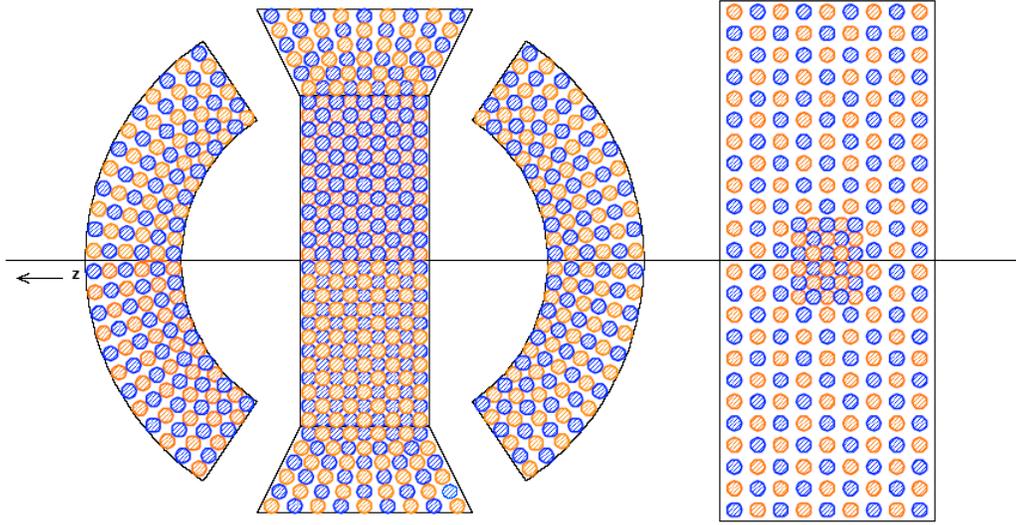


Figure 8.3: A representation of the PMT grouping scheme for intrinsic timing evaluation. Blue circles correspond to odd group PMTs while orange circles correspond to even group PMTs.

The Gaussian width of this distribution is 171 ps at 55 MeV and 163 ps at 83 MeV. These values contain a contribution of about 72 ps from the NaI time resolution and a component of about 58 ps from the spread of the pion decay vertex. The inferred calorimeter time resolutions are therefore 144 ps at 55 MeV and 134 ps at 83 MeV. An additional component is present in these estimates from the synchronization of the NaI and XEC digitizers due to DRS timing problems when they were first installed that only affected the charge exchange run. Its size is estimated to be roughly 130 ps. In a previous MEG run (during 2008), when this problem was not present, the intrinsic timing resolution was comparable at 44 ps, and the effective timing resolution was found to be 78 ps. This is to be compared with a goal resolution of 43 ps.

Photon Energy Resolution The photon energy response function is taken from the reconstructed energy distribution of monochromatic 55 MeV photons from π^0 decay (see

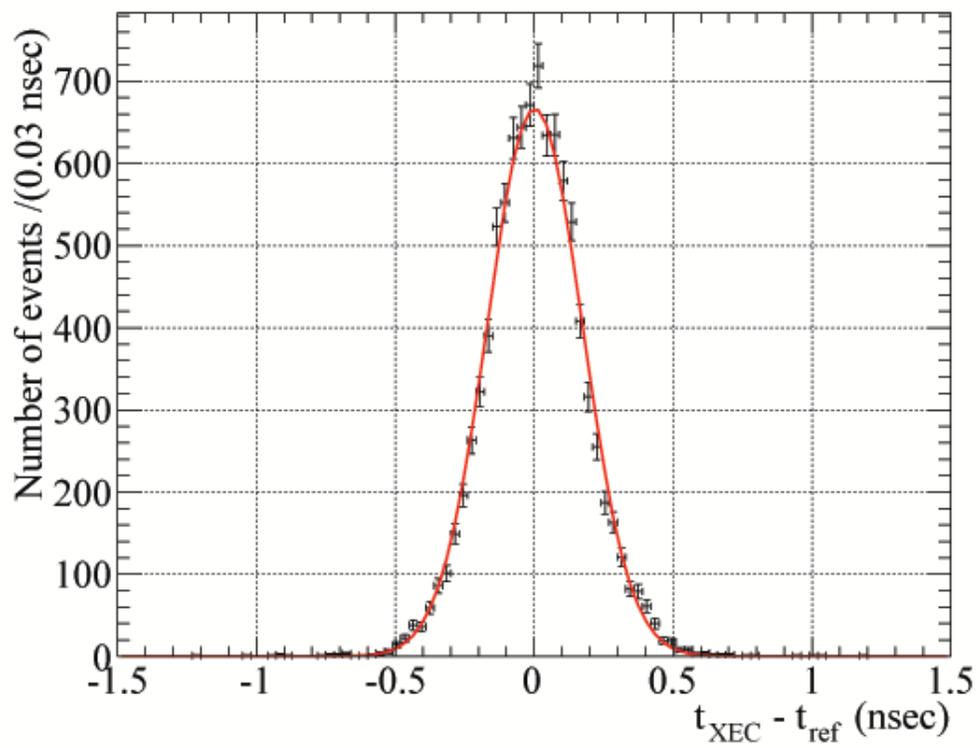


Figure 8.4: Distribution of the NaI time and XEC time difference for coincident photon pairs.

Figure 8.5), which are very close to the signal energy. The 55 MeV photon sample is isolated by requiring an 83 MeV photon in the NaI detector and an opening angle greater than 175° . The low energy tail is made up of photons interacting in front of the calorimeter and shallow events with energy leakage through the front face. This distribution is sensitive to not only the intrinsic resolution introduced by the energy reconstruction algorithm and the detector itself, but also to pileup effects that are unique to the π^- beam environment. The latter component looks different in the normal μ^+ beam environment. The intrinsic, beam-independent photon energy resolution is modeled by a piecewise function:

$$f_0(x) = \begin{cases} A \exp\left(\frac{t}{\sigma_{up}^2} \left[\frac{t}{2} - (x - x_0)\right]\right) & x \leq x_0 + t \\ A \exp\left(\frac{(x-x_0)^2}{-2\sigma_{up}^2}\right) & x > x_0 + t. \end{cases} \quad (8.1)$$

A is constrained by normalization in a likelihood fit, x_0 is the location of the peak, t indicates the distance from the peak position at which the piecewise definition changes, and σ_{up} is the resolution of the high energy part. The π^- beam specific part is captured in a histogram of the energy distribution in noise-triggered events, $h_\pi(x)$. The convolution, $f_0(x) \otimes h_\pi(x)$, is then fit to the 55 MeV photon data, as in Figure 8.5.

A μ^+ beam specific distribution, $h_\mu(x)$, is obtained similarly, and the convolution, $f_0(x) \otimes h_\mu(x)$, represents the signal photon energy response function. The effective σ_{up} of that convolution varies strongly with depth. For $w > 2$ cm (about 77% of the event distribution), that resolution is 2.1% (1.1 MeV). As the interaction occurs closer to the front face, the resolution worsens due, for example, to PMT saturation and faster variation of the solid angle subtended by nearby PMTs. The average resolution is 2.8% (1.5 MeV) for $1 \leq w < 2$ cm and 3.3% (1.7 MeV) for $w < 1$ cm. These are to be compared with a goal resolution of 1.7%. There are also smaller variations with u and v ; the resolution tends to worsen near the edges of the detector for reasons similar to those explaining the depth dependence.

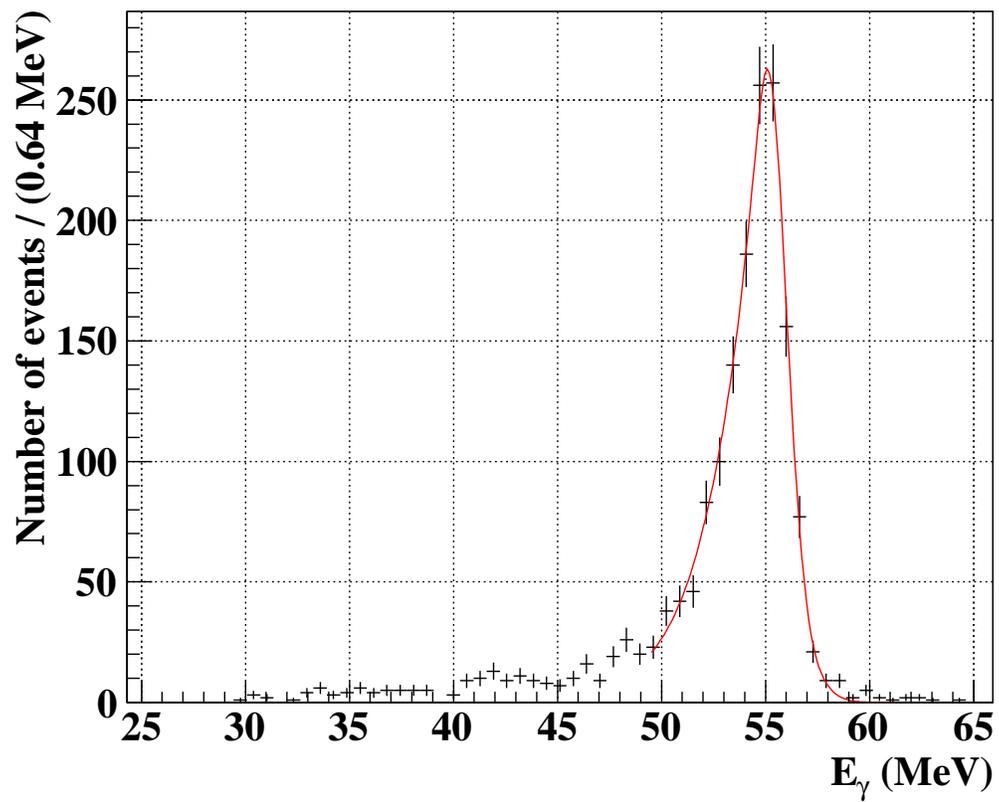


Figure 8.5: Distribution of reconstructed photon energies from a 55 MeV source. Shallow events, $w < 2$ cm, are excluded.

xy plane is calculated from the position of this cluster and the two neighboring clusters. Cases in which there is a skipped chamber within the track circle or a neighboring cluster consists of only one hit are further eliminated. Each hit is then propagated to the central chamber plane using this trajectory and the difference in the two radial positions at the central plane is interpreted as a measure of the single hit position resolution. Since the trajectory is determined from local information and only used to propagate the hit positions over a small distance, this technique is not very sensitive to multiple scattering or errors in the chamber alignment. Figure 8.7 shows the distribution of $R_{plane\ 0} - R_{plane\ 1}$. Assuming uncorrelated hit errors, the width of this distribution is to be divided by $\sqrt{2}$, yielding a single hit radial position resolution of $\sim 250\ \mu\text{m}$ with additional non-Gaussian tails. This is to be compared with a goal resolution of $\sigma_R = 200\ \mu\text{m}$. This resolution gets contributions from drift diffusion, the precision of the hit time, the precision of the track time, the accuracy with which the magnetic field is known, and the accuracy of the time-to-distance relationship.

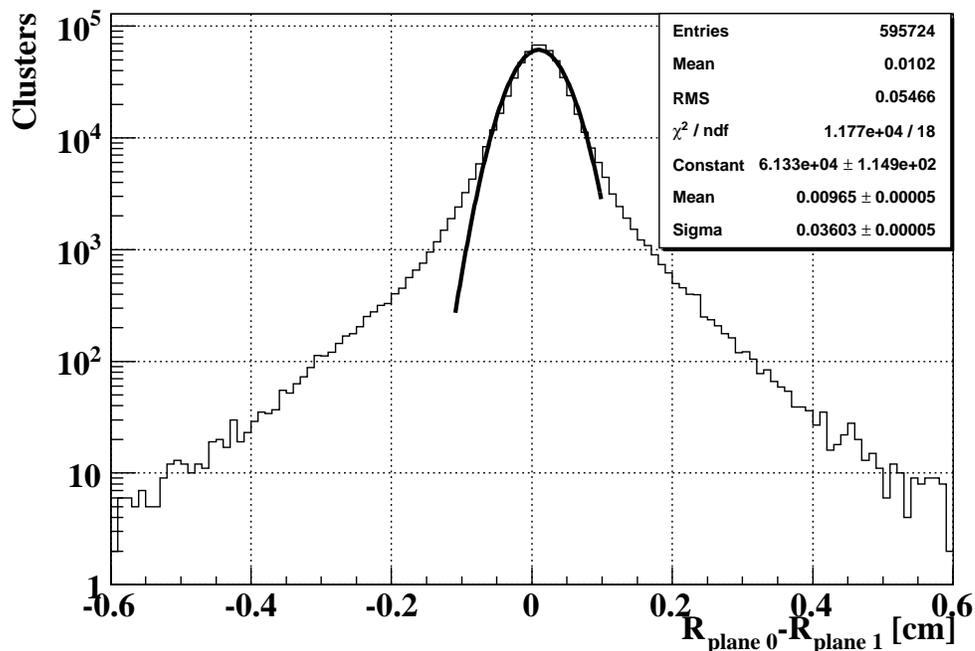


Figure 8.7: A Gaussian fit to the distribution of $R_{plane\ 0} - R_{plane\ 1}$ out to $\pm 1\ \text{mm}$.

Intrinsic Z Resolution A technique for measuring the intrinsic Z position resolution is illustrated in Figure 8.8. Two-hit clusters belonging to a track segment are selected in the same way described for the radial position resolution measurement to compare the two measured Z positions. The calculated track circle in the xy plane is used to define a local coordinate system, with the center of the circle defining the origin. Within this coordinate system, a polar angle (Φ) is calculated for each overall cluster position and for each hit within the cluster, as shown in Figure 8.8a. From the three cluster coordinates, a quadratic trajectory for Z as a function of the local Φ is computed. This trajectory is used to project the measured Z coordinates of the two hits in question to a common Φ and the resulting distribution in the difference of the projected Z coordinates, $Z_{late} - Z_{early}$, is used to infer the intrinsic Z resolution, as pictured in Figure 8.8b. This procedure is also

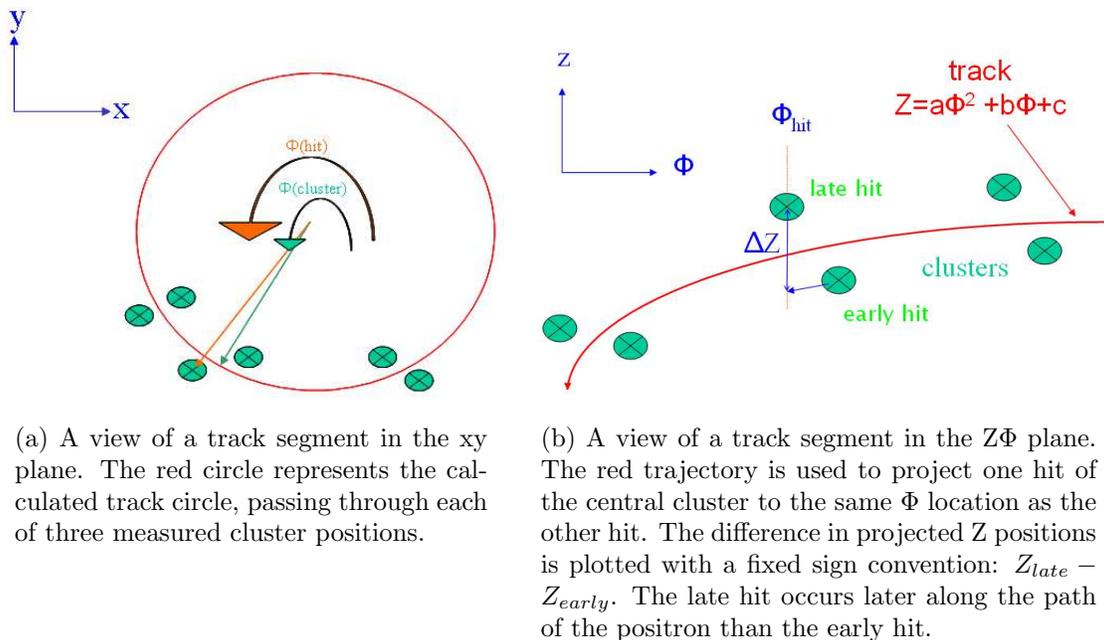


Figure 8.8: A diagram of the technique for measuring the intrinsic Z position resolution.

largely insensitive to multiple scattering and errors in chamber alignment. Figure 8.9 shows the distribution of $Z_{late} - Z_{early}$. This implies a single hit longitudinal position resolution of $\sim 900 \mu\text{m}$ and additional non-Gaussian tails. This is to be compared with a goal resolution of $\sigma_Z = 300 \mu\text{m}$. The largest known contribution to this number comes from the precision

with which the electronic voltage baseline is known in the presence of noise fluctuations, as discussed in section 5.2.1. This gives a contribution to the Z resolution whose magnitude depends on the pattern of charge induced on the pads and on the location of the wire within the detector. Averaging over many events distributed throughout the detector gives a contribution of $\sigma_{Z,noise} = 550 \mu\text{m}$ to the Z resolution from the precision of the baseline prediction. Relative gains between the upstream and downstream readout channels of a given pad, coming either from different preamplifiers or DRS chips, give another contribution to the Z resolution. In section 6.2.2, the residual uncertainty in relative gains after calibration is estimated to produce a small error in Z of $\sigma_{Z,gain} = 70 \mu\text{m}$. An additional uncertainty arises from potential mismatches in the integration times for measuring the charge at both ends of a pad. A stochastic error of 1 ns RMS on the difference in integration times is found to cause an error in Z of about $\sigma_{Z,timing} = 350 \mu\text{m}$. Another source of error is that from fluctuations in the Z position of the ionization sites contributing to the charge measurement. Averaging over different drift distances and angles, the RMS error in the mean Z coordinate is estimated to be about $\sigma_{Z,ionization} = 90 \mu\text{m}$. If the amplitude of the charge asymmetry is systematically different on the hood than on the cathode, this constitutes another effect on the Z resolution. This can result from chamber bowing due to gas pressure; however, such effects are calibrated out as described in section 6.2.2. Another source of this effect that is not calibrated is asymmetric shower development, in which the first arriving drift electrons come preferentially from the hood (or cathode) side of the cell. It is difficult to estimate the magnitude of such effects since the expected magnitude of shower offsets is unknown. More details on the preceding effects are available in a technical note.[61] Lastly, there is a contribution from extra charge induced on a pad from a hit in an adjacent cell. If in addition to the charge induced on the pad foils of a cell by a hit on its wire, there is also extra charge induced from a hit in a nearby cell at significantly different Z, it will produce an error in the measured Z coordinate. The size of this effect is not precisely known. Summing only the known or estimated contributions gives an expected Z resolution of $660 \mu\text{m}$.

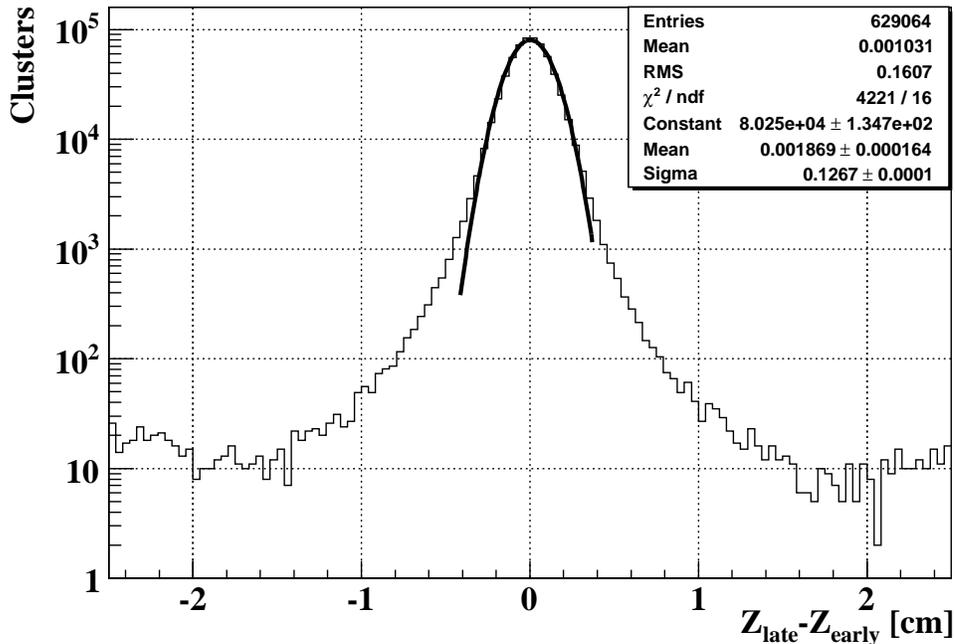
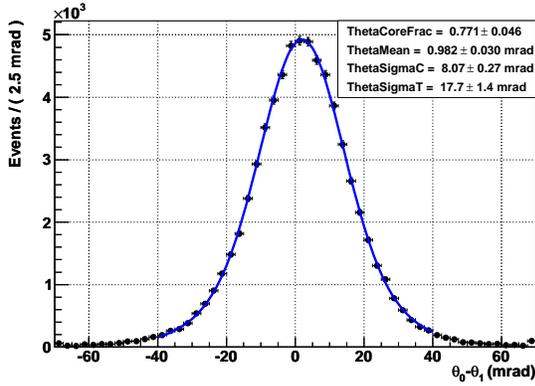


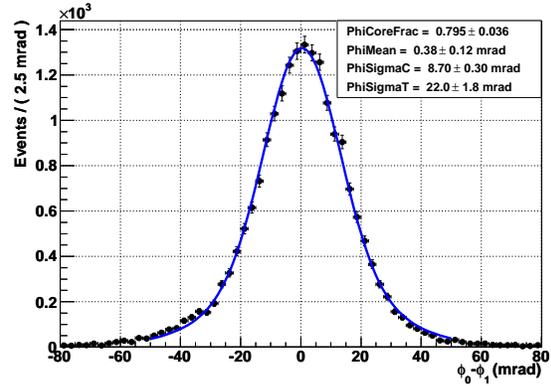
Figure 8.9: A Gaussian fit to the distribution of $Z_{\text{late}} - Z_{\text{early}}$ out to ± 4 mm.

Positron Angle Resolutions The resolution in the measurement of the positron angle at the target is another necessary ingredient in determining the relative angle response functions. The resolutions in both θ_e and ϕ_e are measured from data by exploiting events where the positron makes two turns in the drift chamber. Each turn is treated as an independent track, fitted, and propagated back to the target plane where the emission angles are compared. Since the target plane is the same in both cases, this method is not sensitive to a displacement of the target from its assumed position, which would produce a systematic error in the positron angles. The distribution in the difference of the two angle measurements in double-turn events is shown in figure 8.10 for both θ_e and ϕ_e . Both the correlation between the positron energy error and the ϕ_e angle, and the correlation between the θ_e angle error and the ϕ_e angle, discussed in section 8.2.2, are naturally present in the distribution of the difference in ϕ_e measurements from the two turns. If the fitted energies or the extrapolated θ_e angles of the two turns differ, the mean of the distribution in $\Delta\phi_e \equiv \phi_e^{\text{1st turn}} - \phi_e^{\text{2nd turn}}$ will be offset from zero. Averaging over different values of $\Delta E_e \equiv E_e^{\text{1st turn}} - E_e^{\text{2nd turn}}$ and

$\Delta\theta_e \equiv \theta_e^{1st\ turn} - \theta_e^{2nd\ turn}$ would tend to inflate the inferred ϕ_e resolution. To minimize this effect, the ϕ_e resolution is measured on a sample of two turn events with cuts imposed: $|\Delta E_e| < 500$ keV and $|\Delta\theta_e| < 10$ mrad. The resolution in each turn is assumed to be the same and these distributions are fit to the convolution of a sum of two Gaussians with itself. This sum of two Gaussians then represents the response function of the positron angle. The average θ_e resolution is 8.07 mrad in the core (77.1% of events) and 17.7 mrad in the tail, and the average ϕ_e resolution is 8.70 mrad in the core (79.5% of events) and 22.0 mrad in the tail. The goal resolution is $\sigma_{\theta_e, \phi_e} \sim 5$ mrad.



(a) A fit to the distribution of $\Delta\theta_e$ on double turn events.



(b) A fit to the distribution of $\Delta\phi_e$ on double turn events.

Figure 8.10: Fits for the average positron angle resolutions using double turn events. The fit parameters shown are for the double Gaussian that is convolved with itself and fit to the distribution.

Positron Energy Resolution One method for measuring the positron energy resolution consists of fitting the accidental positron energy spectrum in the sidebands to a model for the true energy spectrum convolved with a response function:

$$Probability\ density(E_e^{measured}) = (Michel * Acceptance)(E_e^{true}) \otimes Resolution. \quad (8.2)$$

Here the true energy distribution is represented by the unpolarized Michel positron energy spectrum multiplied by an acceptance function. One must assume a functional form for both the acceptance and the resolution function. Based on the guidance provided by Monte Carlo, the acceptance function is taken to be of the form:

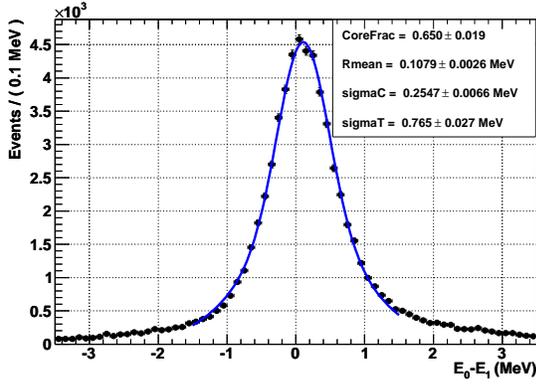
$$\frac{1 + \operatorname{erf}\left(\frac{E_e - \mu_{acc}}{\sqrt{2}\sigma_{acc}}\right)}{2}, \quad (8.3)$$

and the resolution is taken to be a sum of two Gaussians:

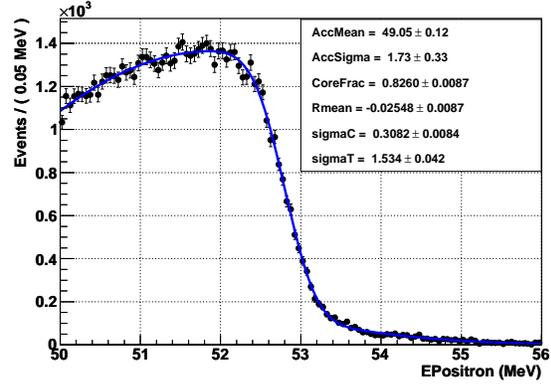
$$f_{core} \frac{1}{\sqrt{2\pi}\sigma_{core}} e^{(E_e - \mu_e)/2\sigma_{core}^2} + (1 - f_{core}) \frac{1}{\sqrt{2\pi}\sigma_{tail}} e^{(E_e - \mu_e)/2\sigma_{tail}^2}. \quad (8.4)$$

A total of six parameters are floated in the fit: μ_{acc} , σ_{acc} , σ_{core} , σ_{tail} , f_{core} , μ_e . Figure 8.11 presents such a fit to all positrons passing selection criteria. This gives an average resolution of 308 keV in the core (82.6% of events) and 1.534 MeV in the tail. There is also a small 25 keV systematic underestimate of the energy. This is to be compared with the goal resolution of $\sigma_{E_e} = 180$ keV (0.8%). This technique suffers from the drawback that it is affected by correlations among the acceptance and resolution function since both are floated.

A complimentary approach to determining the positron energy resolution is possible by using two-turn events as in the positron angular resolution measurements. Figure 8.11a shows the distribution of the energy difference between the two turns. This is fit to the convolution of a double Gaussian with itself, the same shape assumed in the Michel edge fit. A disadvantage of this technique is its inability to detect a global shift in the positron energy scale. This technique gives an average resolution of 255 keV in the core (65% of events) and 765 keV in the tail. A systematic offset of 108 keV between the energies of the two turns also appears; the energy of the first turn is systematically larger than the energy of the second turn. This can originate from errors in the assumed energy loss in the track fitting or errors in the magnetic field.



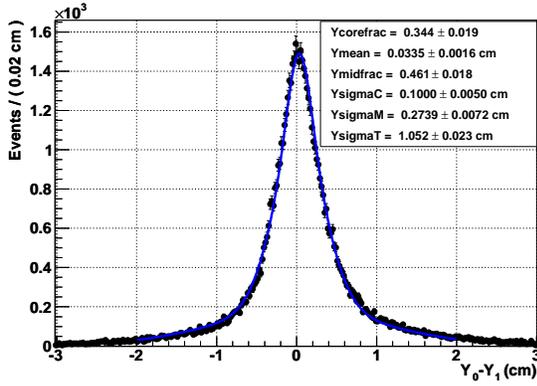
(a) A fit to the distribution of $\Delta E_e \equiv E_e^{1st\ turn} - E_e^{2nd\ turn}$ on double turn events.



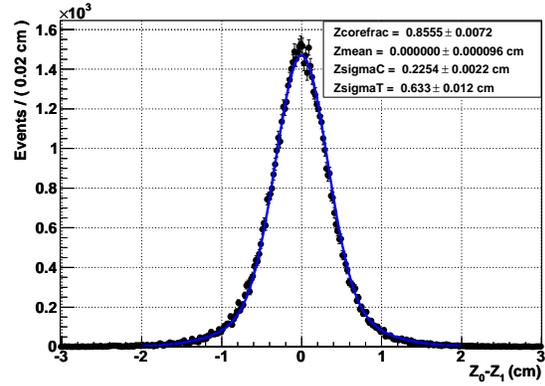
(b) A fit to the accidental (Michel) positron energy spectrum.

Figure 8.11: Two methods for finding the average positron energy resolution.

Muon Vertex Position Resolution The resolution in the vertex position at the stopping target is dominated by the positron angular resolution with a small contribution from the position resolution in the drift chamber. For a proper evaluation of the relative angle response function, a precise knowledge of the correlations between positron angle error and vertex position error is required. To simply find the average vertex position resolutions, however, it is sufficient to measure it directly by comparing the projected point of interception at the target plane on double-turn events. Figure 8.12 presents the distribution of the difference in Y and Z vertex coordinates from the two turns. ΔZ_e is fit to the convolution of a double Gaussian with itself, while ΔY_e is fit to the convolution of the sum of three Gaussians with itself. The resolutions are found to be 1 mm in the core (34.4% of events), 2.739 mm in the middle Gaussian (46.1% of events) and 1.052 cm in the tail for Y; and 2.254 mm in the core (85.55% of events) and 6.33 mm in the tail for Z. The goal resolution is $\sigma_{Y,Z} \sim 1$ mm.



(a) A fit to the distribution of ΔY_e on double turn events.



(b) A fit to the distribution of ΔZ_e on double turn events.

Figure 8.12: Fits for the average positron vertex resolutions using double turn events. The fit parameters shown are for the double (or triple) Gaussian that is convolved with itself and fit to the distribution.

8.1.3 TIC

The resolution of the positron impact time at the timing counter contributes to the relative time resolution. This is measured on events where a positron passes through two adjacent bars. The distribution in the difference of the two measured impact times gives a measure of the impact time resolution. This is shown in Figure 8.13 as a function of bar number where the resolution is the σ of a fitted Gaussian divided by $\sqrt{2}$. When two bars are crossed that are connected to different DRS chips, the measured resolution tends to be worse than when the bars are connected to the same chip due to the clock synchronization. The average resolution is 75 ps, compared to the goal resolution of 40 ps. Note that this technique does not account for possible systematic errors in the impact time that vary with Z .

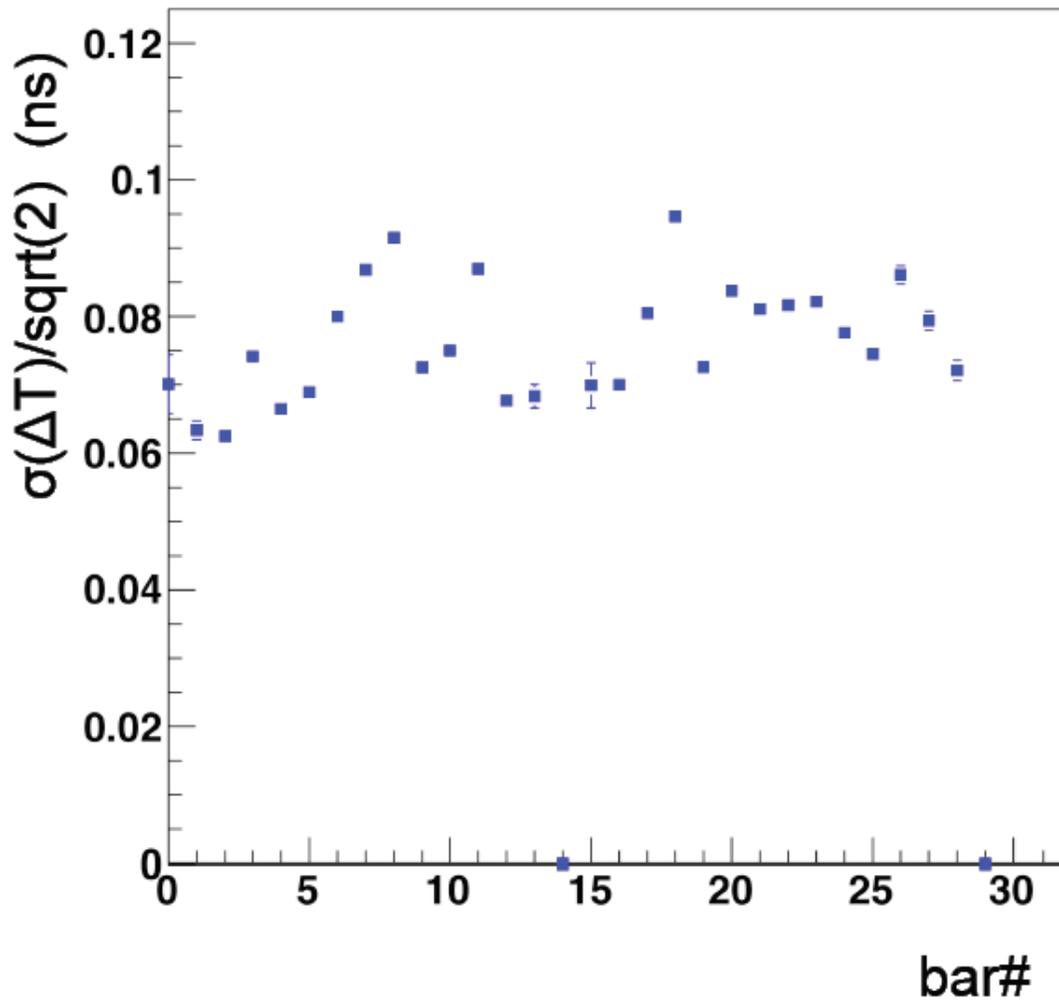


Figure 8.13: Timing counter resolution as a function of bar number.

8.2 Resolutions of Multi-detector Measurements

This section outlines the total resolutions achieved in measurements that involve multiple detectors.

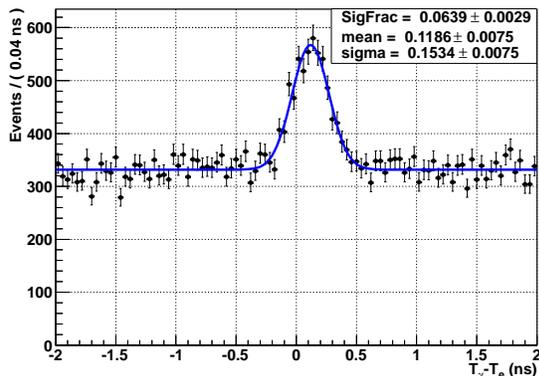
8.2.1 Relative Time Resolution

Radiative muon decays provide a means to measure the relative timing resolution in $t_{e\gamma}$ under experimental conditions that are identical to those under which a signal would be taken. Figure 8.14 presents the radiative decay timing peak, whose mean is found to differ significantly in upstream and downstream events. These distributions are fit to the sum of a flat line describing the accidental time spectrum and a single Gaussian describing the timing resolution for radiative decays. The energy ranges are fixed ($40 < E_\gamma < 45$ MeV, $40 < E_e < 56$ MeV) and relative angle cuts ($|\theta_{e\gamma}, \phi_{e\gamma}| < 400$ mrad) are made to avoid edge effects of the detector acceptances. An additional cut is made to make the kinematics consistent with radiative decays. For fixed positron and photon momenta, the energy carried off by the two neutrinos can be calculated (neglecting neutrino masses) as: $E_{\nu_e} + E_{\bar{\nu}_\mu} = |\vec{p}_{\nu_e}| + |\vec{p}_{\bar{\nu}_\mu}| \geq |\vec{p}_{\nu_e} + \vec{p}_{\bar{\nu}_\mu}|$. This triangle inequality saturates when the two neutrinos are emitted with zero opening angle, thus, one may calculate the minimum energy carried off by the neutrinos, $E_{\nu\nu}$, by balancing the positron and photon momenta with two parallel neutrinos. Then the data is required to satisfy:

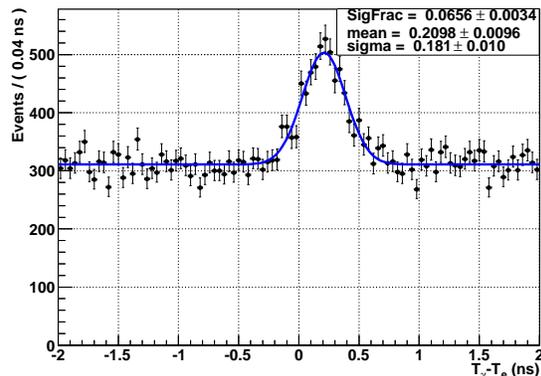
$$E_e + E_\gamma + E_{\nu\nu} \leq m_\mu \tag{8.5}$$

for consistency with a radiative muon decay at rest. The relative timing resolution on this sample is $\sigma_{t_{e\gamma}} \approx 180$ ps. Since photo-statistics increases with photon energy, the timing resolution at the signal energy is expected to be slightly better than on the sample with

$40 < E_\gamma < 45$ MeV. The photon timing resolution dependence on energy is modeled as $\sigma_{t_\gamma} \propto \frac{A}{\sqrt{E_\gamma}}$. By comparing the measured photon timing resolution at 55 MeV and 83 MeV, the proportionality constant is estimated to be $A = 0.659 \text{ ns}\sqrt{\text{MeV}}$. The average photon energy in the radiative decay sample is 42 MeV. The timing resolution can then be naively extrapolated into the signal region to obtain $\sigma_{t_{e\gamma}} = \sqrt{180^2 - \frac{0.659^2}{42} + \frac{0.659^2}{52.8}} = 174$ ps. This is to be compared with a goal resolution of 64 ps.



(a) A fit to the radiative decay timing peak for upstream positrons.



(b) A fit to the radiative decay timing peak for downstream positrons.

Figure 8.14: Muon radiative decay timing peaks.

The $t_{e\gamma}$ resolution receives three comparably sized contributions: the positron impact time resolution (~ 75 ps from 8.1.3), the positron time of flight resolution (~ 75 ps estimated from Monte Carlo studies), and the photon timing resolution (~ 78 ps from 8.1.1).

8.2.2 Relative Angle Resolutions

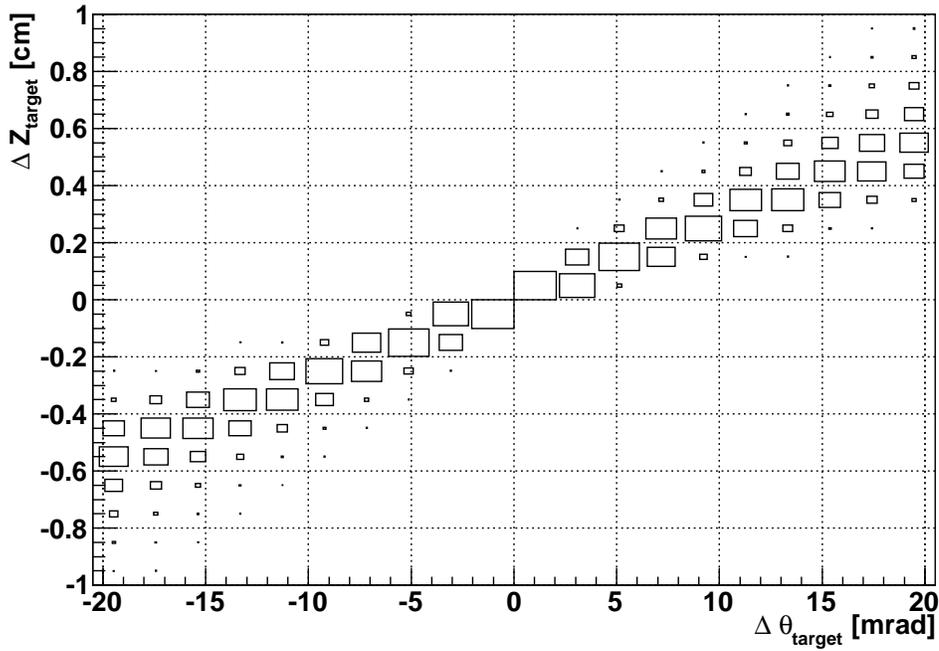
With no available source of particles at fixed angle, the resolutions in $\theta_{e\gamma}$ and $\phi_{e\gamma}$ are not directly measurable. The angular resolutions are, however, composed of things that can be measured. The relative angle PDFs are complicated functions of the positron angular resolutions, the photon position resolutions, and the vertex position resolutions, and are

further complicated by strong correlations of the vertex position resolutions with the positron angle resolutions. Such a complex scenario is most amenable to a Monte Carlo simulation according to the measured component PDFs.

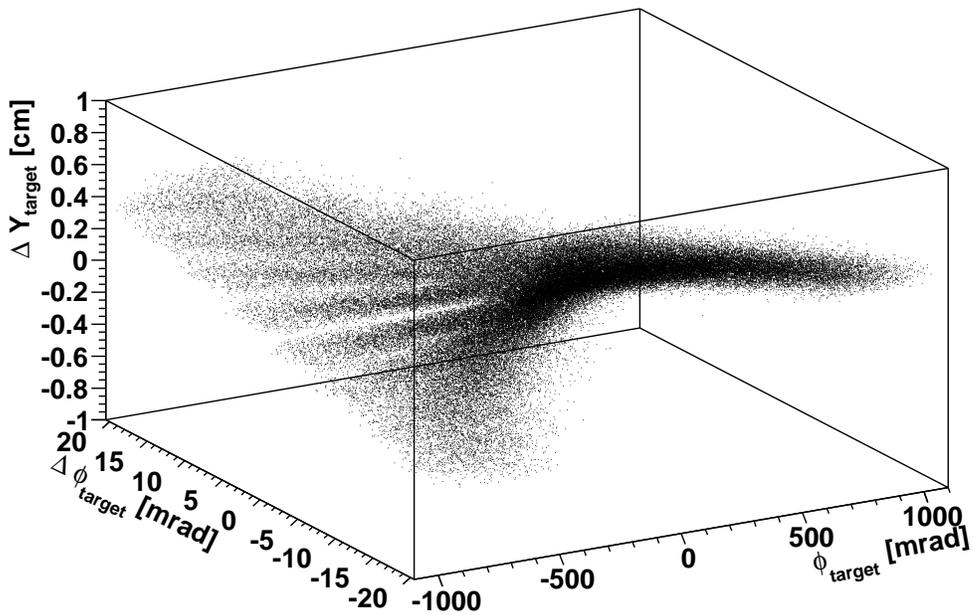
For this simulation to be accurate, one needs the form of the correlations between positron angle error and vertex position error. These correlations are governed by the geometry of the drift chamber and target, the magnetic field, and the algorithm for propagating drift chamber tracks back to the target plane.

A technique for determining these correlations directly on data using the same algorithm for track propagation as in the analysis is developed. Tracks passing the standard selection criteria are chosen, and the track angles at the first hit chamber are incrementally changed, one at a time. Each time the direction of the momentum vector at the first chamber is modified, the track is projected to the target plane and the change in emission angle at the target is plotted as a function of the change in vertex position at the target. Figure 8.15a shows the results of the error in Z at the target as a function of the error in θ_e at the target. A fit to a line obeying the constraint that zero error in θ_e should produce zero error in Z gives the result that a 1 mrad error in θ_e causes a 288 μm error in Z . Figure 8.15b shows a similar plot for the error in Y at the target versus the error in ϕ_e at the target, with an additional dependence on ϕ_e itself for geometrical reasons. Depending on ϕ_e , similarly constrained line fits to ΔY versus $\Delta\phi_e$, in 10 mrad bins of ϕ_e , give slopes with magnitudes ranging from 160 $\frac{\mu\text{m}}{\text{mrad}}$ to 506 $\frac{\mu\text{m}}{\text{mrad}}$ depending on ϕ_e .

While they do not affect the width of the angular resolutions if they are known precisely, there are two additional correlations that do affect the mean of the $\phi_{e\gamma}$ PDF. An error in E_e alters the path length and changes the value of ϕ_e inferred by the projection back to the target plane. This correlation is measured as a function of ϕ_e with techniques identical to the those mentioned above. The magnitude of this correlation varies from 0 to 70 $\frac{\text{mrad}}{\text{MeV}}$ depending on ϕ_e . Finally, an error in θ_e produces an error in Z , which produces an error in



(a) Correlation between error in θ_e and error in Z as measured from data.



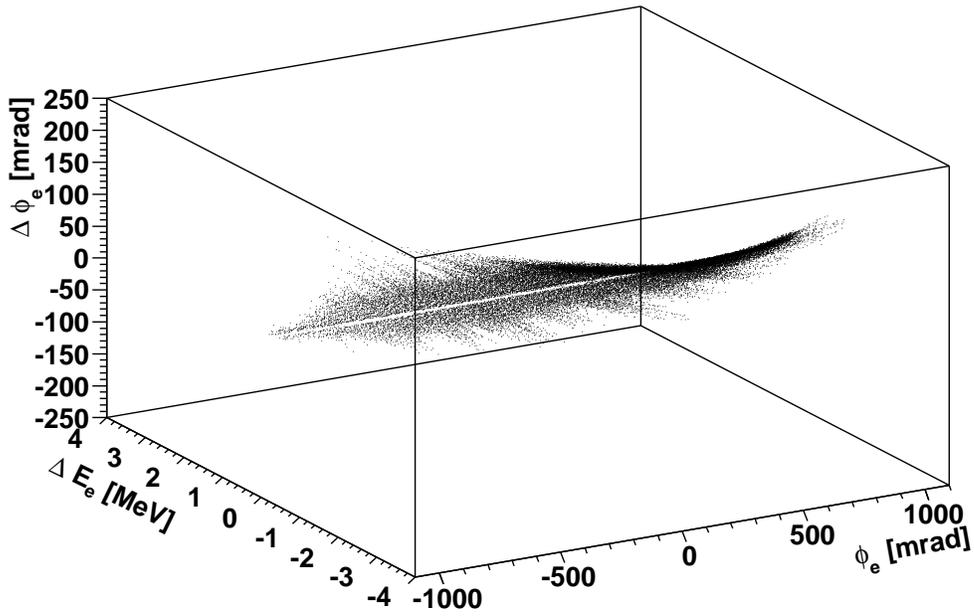
(b) Correlation between error in ϕ_e and error in Y as measured from data.

Figure 8.15: The correlations between angle error and vertex error.

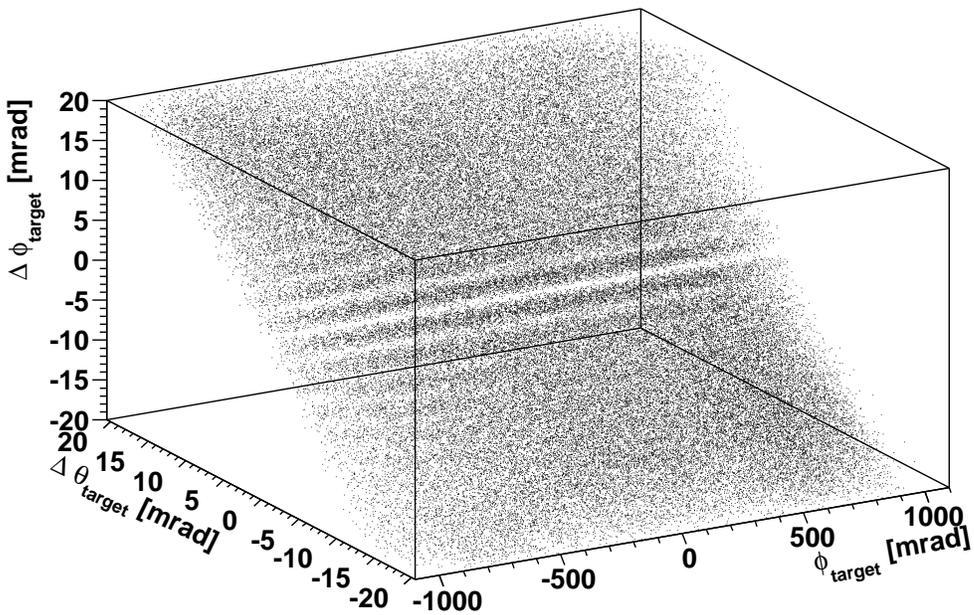
X because they are related by the slant of the target plane. Such an error in X results in an error in ϕ_e . The overall effect is a correlation between θ_e and ϕ_e , which is measured in a way identical to the above and results in slopes ranging from 0.8 to $2 \frac{\Delta\phi_e}{\Delta\theta_e}$ depending on ϕ_e . Figure 8.16 maps these correlations as a function of ϕ_e .

The average resolutions in $\theta_{e\gamma}$ and $\phi_{e\gamma}$ are simulated as follows:

- Back-to-back pairs of positrons and photons (i.e. $\theta_{e\gamma}^{true}, \phi_{e\gamma}^{true} = 0$) are generated uniformly in θ_e and ϕ_e over the detector acceptance, with the vertex fixed at the origin (i.e. $X^{true} = 0, Y^{true} = 0, Z^{true} = 0$). These conditions uniquely determine the first conversion position of the photon except for its depth, w , which is sampled from a probability distribution based directly on a histogram of conversion depths in data. This gives the true coordinates of the first conversion: $(u^{true}, v^{true}, w^{true})$.
- The resolution effect on the positron angles is introduced by generating measurement errors in ϕ_e and θ_e according to the average resolutions in those variables presented in section 8.1.2. The measured positron angles are related to the true angles by the simulated measurement errors: $\theta_e^{meas} = \theta_e^{true} + \Delta\theta_e$, $\phi_e^{meas} = \phi_e^{true} + \Delta\phi_e$.
- Based on the generated error in θ_e , $\Delta\theta_e$, a corresponding error in the vertex Z position is calculated: $Z^{meas} = \Delta Z = \frac{dZ}{d\theta_e} \Delta\theta_e$, where $\frac{dZ}{d\theta_e}$ is the measured slope of the correlation discussed above. A similar error in the vertex Y position is calculated from the generated error in ϕ_e , $\Delta\phi_e$, and the measured angle itself, ϕ_e^{meas} : $Y^{meas} = \Delta Y = \frac{dY}{d\phi_e}(\phi_e^{meas}) \Delta\phi_e$, where $\frac{dY}{d\phi_e}(\phi_e^{meas})$ is the ϕ_e^{meas} dependent slope of that correlation discussed above. The associated error in X follows immediately from the angle at which the target is slanted with respect to the beam axis: $X^{meas} = \Delta X = -\Delta Z \tan 20.05^\circ$.
- The resolution effect on the measured conversion position is included by drawing errors in the coordinates of the first conversion $(\Delta u, \Delta v, \Delta w)$ from their respective average resolutions outlined in section 8.1.1. This gives a simulated set of measured coordinates



(a) Correlation between error in E_e and error in ϕ_e as measured from data.



(b) Correlation between error in θ_e and error in ϕ_e as measured from data.

Figure 8.16: The correlations affecting the $\phi_{e\gamma}$ PDF mean.

for the first conversion: $u^{meas} = u^{true} + \Delta u, v^{meas} = v^{true} + \Delta v, w^{meas} = w^{true} + \Delta w$.

- The measured photon angles, θ_γ^{meas} and ϕ_γ^{meas} , are calculated from the measured vertex position and the measured coordinates of the first conversion.
- Finally, the simulated measured relative angles, $\theta_{e\gamma}^{meas}$ and $\phi_{e\gamma}^{meas}$, are calculated from $\theta_e^{meas}, \phi_e^{meas}$ and $\theta_\gamma^{meas}, \phi_\gamma^{meas}$. The distributions in $\theta_{e\gamma}^{meas} - \theta_{e\gamma}^{true}$ and $\phi_{e\gamma}^{meas} - \phi_{e\gamma}^{true}$ are plotted over a large number of simulated events, which represent the shape of the average relative angle PDFs.

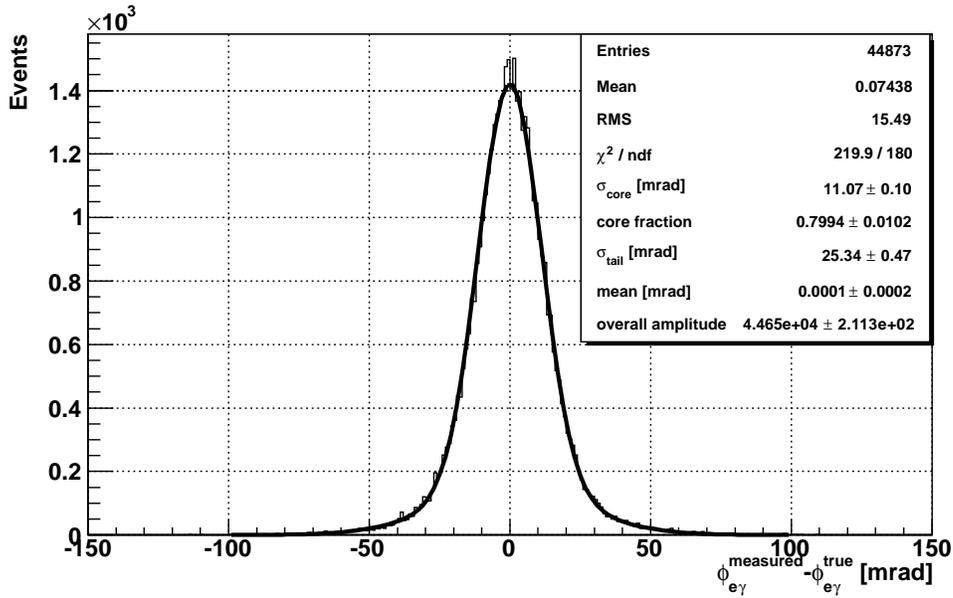
These distributions are presented in figure 8.17. The average resolution in $\theta_{e\gamma}$ is 13.1 mrad in the core (78% of events) and 25.2 mrad in the tail, and the average resolution in $\phi_{e\gamma}$ is 11.1 mrad in the core (79.9% of events) and 25.3 mrad in the tail. These are to be compared with a goal resolution in the opening angle of 8.7 mrad.

8.3 Detector Efficiencies

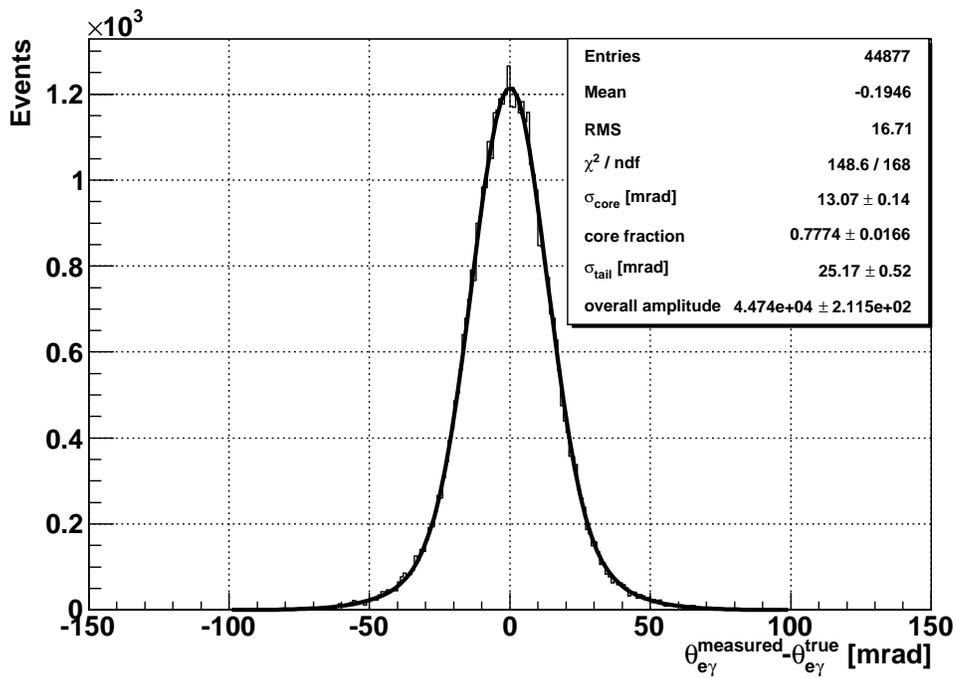
This section outlines the individual detector efficiencies.

8.3.1 Photon Detection Efficiency

One crucial aspect of the calorimeter performance is the photon detection efficiency for signal photons headed toward its fiducial volume. This number is, in fact, required for normalizing the experiment. Several sources of efficiency loss contribute: photons may fail to penetrate the magnet wall and cryostat, photons may be rejected by pile-up cuts, and photons may be rejected by energy cuts especially when they interact before reaching the active volume. The photon detection efficiency is measured primarily from signal Monte Carlo. It is defined as the conditional probability to reconstruct a signal photon that passes pile-up and energy



(a) A fit to the simulated average $\phi_{e\gamma}$ resolution.



(b) A fit to the simulated average $\theta_{e\gamma}$ resolution.

Figure 8.17: The average relative angle resolutions.

cuts, given a reconstructed signal positron. The uncertainty in the result is estimated by comparing it with the probability to detect 54.9 MeV photons from charge exchange data given a tagged 82.9 MeV photon in the NaI detector. The result is $\epsilon_{\mu \rightarrow e\gamma}^{\gamma} = 0.58 \pm 0.03$.

8.3.2 Positron Detection Efficiency

For the purposes of normalizing the experiment, the absolute positron detection efficiency is not needed. Only the ratio of signal positron to Michel positron detection efficiency is required. This information is contained in the acceptance part of the fit to the Michel spectrum of section 8.1.2. The ratio of detection efficiency for signal to Michel positrons in the interval $50 < E_e < 52.8$ MeV is calculated by evaluating the fitted acceptance function at the signal energy and dividing by the theoretical Michel spectrum-weighted average of the acceptance function in the interval $50 < E_e < 52.8$ MeV. The uncertainty in this number is estimated by refitting the measured Michel spectrum with an alternate model of the acceptance function, a Gaussian, and recalculating the efficiency ratio. The result is $\frac{\epsilon_{\mu \rightarrow e\gamma}^e}{\epsilon_{\mu \rightarrow e\nu\nu}^e} = 1.14 \pm 0.06$. Figure 8.18 shows both acceptance curves superimposed.

The positron detection efficiency is actually the product of the drift chamber detection efficiency and the conditional timing counter detection efficiency given a reconstructed positron. The latter can be measured directly. By using a drift chamber-only trigger, and selecting events with a positron in the drift chamber passing the drift chamber-related selection cuts, the probability to also have a detected timing counter hit passing timing counter-related selection cuts can be measured. One impediment to this is the fact that in the default reconstruction algorithm, tracks in the drift chamber already require the presence of a timing counter hit because it must be used to calculate the track time (see section 5.2.3). To do this study, an alternate version of the algorithm is used that calculates the track time by testing a range of track times, and finding the one that minimizes the sum over all clusters on the

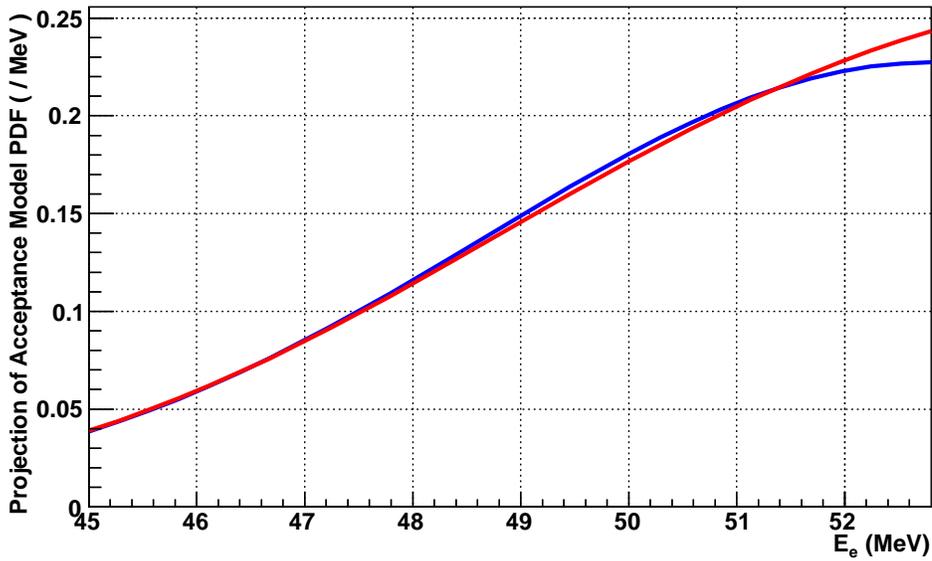
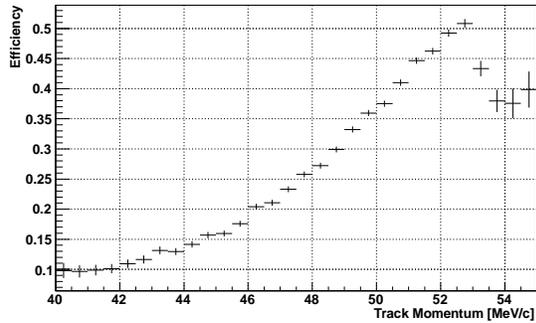


Figure 8.18: The red curve is the standard error function model for the acceptance, while the blue curve is an alternate Gaussian model. Both are normalized to unit area in the range of positron energy on the plot.

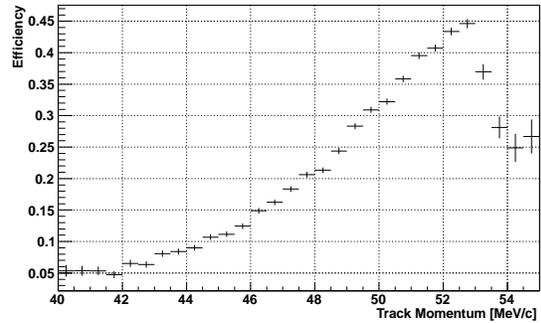
track of the figure of merit in Eq. 5.20. There is some ambiguity as to what constitutes a good match between a drift chamber track and a timing counter cluster. Figure 8.19 shows the conditional timing counter detection efficiency as a function of positron momentum for a few different choices of selection criteria on the DCH-TIC matching. The efficiency peaks near the signal energy at 45-51% depending on the timing counter-related selection cuts. This efficiency is seen to fall for unphysical values of the momentum.

Total efficiency for a loose spatial match



(a) The conditional timing counter efficiency requiring $|\Delta Z_{DCH-TIC}| < 20$ cm and $|\Delta R_{DCH-TIC}| < 10$ cm, which is just the minimal selection in the DCH-TIC matching algorithm.

Total efficiency for a time match & full spatial match



(b) The conditional timing counter efficiency requiring the same spatial matching as in section 7.2.5 as well as requiring the reconstructed track time and the calculated track time based on the timing counter cluster to be within 50 ns.

Figure 8.19: The conditional timing counter efficiency given a positron passing drift chamber-related selection cuts versus positron momentum for various selection criteria on the match with the timing counter.

Chapter 9

Maximum Likelihood Analysis

This chapter describes the methods and results of a maximum likelihood analysis to place a confidence interval on $BR(\mu^+ \rightarrow e^+\gamma)$. A blind analysis is performed by initially hiding signal-like events. A likelihood function is constructed from probability density functions (PDFs) of the distributions in $E_e, E_\gamma, t_{e\gamma}, \theta_{e\gamma}$ and $\phi_{e\gamma}$ for signal, muon radiative decay, and accidental background events. Signal-like events are then unblinded, and the likelihood function is used to obtain a best fit value of $BR(\mu^+ \rightarrow e^+\gamma)$. Finally, a 90% confidence interval is calculated for $BR(\mu^+ \rightarrow e^+\gamma)$.

9.1 Blinding and Analysis Windows

A rare decay search is particularly sensitive to the selection criteria. If the experimenter knows which events satisfy them during the analysis, a biased result can be obtained either by preferentially choosing cuts to throw away certain events giving a better upper limit than deserved or by choosing cuts that retain certain events producing an artificially significant signal.[62] A blind analysis is performed by initially hiding all events that satisfy: $|t_{e\gamma}| < 0.7$

ns and $E_\gamma > 48$ MeV, so that no signal events are expected in the unhidden sample. The blinding is held in place until the selection cuts, PDFs, and likelihood analysis procedure are settled upon. Three windows in the unhidden sample are used for testing and preparing the final likelihood fit procedure. Each window is defined by a unique set of selections in addition to the minimal criteria of chapter 7:

1. Left sideband: $-2.1 < t_{e\gamma} < -0.7$ ns, $48 < E_\gamma < 58$ MeV, $50 < E_e < 56$ MeV, $|\theta_{e\gamma}, \phi_{e\gamma}| < 50$ mrad
2. Right sideband: $0.7 < t_{e\gamma} < 2.1$ ns, $48 < E_\gamma < 58$ MeV, $50 < E_e < 56$ MeV, $|\theta_{e\gamma}, \phi_{e\gamma}| < 50$ mrad
3. Bottom sideband: $|t_{e\gamma}| < 0.7$ ns, $44 < E_\gamma < 48$ MeV, $50 < E_e < 56$ MeV, $|\theta_{e\gamma}, \phi_{e\gamma}| < 50$ mrad.

The analysis to search for $\mu^+ \rightarrow e^+\gamma$ is performed in the signal window, which consists of the left and right sidebands together with most of the hidden events: $|t_{e\gamma}| < 2.1$ ns, $48 < E_\gamma < 58$ MeV, $50 < E_e < 56$ MeV, $|\theta_{e\gamma}, \phi_{e\gamma}| < 50$ mrad. The purpose of the relatively large time window ($\sim \pm 11\sigma$) is to use events well outside the region where signal is expected in order to get high precision on the number of accidental background events in the fit. This is equivalent to fitting in a smaller time window and adding a constraint on the number of accidental events in the likelihood function based on extrapolation from the sidebands if the accidental background is uniformly distributed in $t_{e\gamma}$.

9.2 Likelihood Analysis Framework

This section defines the likelihood function and describes how it is measured.

9.2.1 The Likelihood Function

The data to be analyzed consists of a set of N_O independent observations (or events) of a detected photon and positron. The measurements associated with the i 'th event are collected into an observation vector: \vec{x}_i , and the union of the N_O observation vectors form the entire set of observations from the experiment: $\vec{X} = \{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_{N_O}\}$. By assumption, the space of possible detected events is spanned by signal events ($\mu^+ \rightarrow e^+\gamma$), accidental background events (a positron and photon from different sources), and radiative muon decay events ($\mu^+ \rightarrow e^+\gamma\bar{\nu}_\mu\nu_e$). Given N_{Sig} signal events, N_{Acc} accidental background events, and N_{RD} radiative muon decay events, the conditional probability to observe an arbitrary data set \vec{X} is the product of the conditional probabilities, $p(\vec{x}_i|N_{Sig}, N_{Acc}, N_{RD})$, to observe each of the N_O events:

$$P(\vec{X}|N_{Sig}, N_{Acc}, N_{RD}) = \prod_{i=1}^{N_O} p(\vec{x}_i|N_{Sig}, N_{Acc}, N_{RD}). \quad (9.1)$$

Since it is only the data set \vec{X} that is known, it is more useful to consider the “likelihood” of a given set of parameters (N_{Sig}, N_{Acc}, N_{RD}), given the observed data. If the experimental data is substituted for \vec{X} in Eq. 9.1, and the expression is considered to be only a function of the parameters to be estimated, one obtains the likelihood function:

$$\mathcal{L}(N_{Sig}, N_{Acc}, N_{RD}|\vec{X}) = \prod_{i=1}^{N_O} p(\vec{x}_i|N_{Sig}, N_{Acc}, N_{RD}). \quad (9.2)$$

The best estimate of the parameters (N_{Sig}, N_{Acc}, N_{RD}) is simply the one that maximizes the likelihood function.

If an ensemble of identical experiments were performed, the observed number of events would fluctuate according to Poisson statistics because it is a counting experiment. For this reason, the overall normalization is floated in the likelihood fit by introducing an additional factor

into the likelihood function, making it an extended likelihood function:

$$\mathcal{L}_{ext}(N_{Sig}, N_{Acc}, N_{RD}|\vec{X}) = \frac{N^{N_O} e^{-N}}{N_O!} \prod_{i=1}^{N_O} p(\vec{x}_i|N_{Sig}, N_{Acc}, N_{RD}). \quad (9.3)$$

Here $N \equiv N_{Sig} + N_{Acc} + N_{RD}$ is the expected number of events, which is not necessarily the same as the observed number of events, N_O . This technique has the advantage that error estimates on the numbers of event types automatically include the statistical error on the total number of observed events.[63]

Following the strategy of imposing minimal selection cuts and incorporating the dependence of detector responses on various properties into event-by-event PDFs, the observation vector, \vec{x}_i , contains two distinct types of variables. One class of variables, $\vec{m} \equiv \{E_e, E_\gamma, t_{e\gamma}, \theta_{e\gamma}, \phi_{e\gamma}\}$, consists of the kinematic event properties that follow different probability distributions depending on the event type, and another class of variables, \vec{y} , consisting of event properties that determine the shape of the former probability distributions, but do not themselves have varying probability distributions among the three event types. The observation vector now takes a more specific form:

$$\vec{x}_i = \{\vec{m}_i, \vec{y}_i\}. \quad (9.4)$$

The conditional probability to obtain \vec{x}_i given a certain set of values for the numbers of event types can be written as:

$$p(\vec{x}_i|N_{Sig}, N_{Acc}, N_{RD}) = \frac{N_{Sig}}{N} S(\vec{m}_i|\vec{y}_i)p(\vec{y}_i) + \frac{N_{Acc}}{N} A(\vec{m}_i|\vec{y}_i)p(\vec{y}_i) + \frac{N_{RD}}{N} R(\vec{m}_i|\vec{y}_i)p(\vec{y}_i). \quad (9.5)$$

$S(\vec{m}_i|\vec{y}_i)$, $A(\vec{m}_i|\vec{y}_i)$ and $R(\vec{m}_i|\vec{y}_i)$ are respectively the probabilities for a signal, accidental background, and radiative decay event to result in the set of observables \vec{m}_i given a set of event properties \vec{y}_i . Since the probability to obtain \vec{y}_i is common to all event types, it factors out and becomes irrelevant to the maximum likelihood fit (since it always cancels in

a likelihood ratio test); thus it can be dropped from the likelihood function:

$$\mathcal{L}_{ext}(N_{Sig}, N_{Acc}, N_{RD}|\vec{X}) = \frac{N^{N_O} e^{-N}}{N_O!} \prod_{i=1}^{N_O} \left(\frac{N_{Sig}}{N} S(\vec{m}_i|\vec{y}_i) + \frac{N_{Acc}}{N} A(\vec{m}_i|\vec{y}_i) + \frac{N_{RD}}{N} R(\vec{m}_i|\vec{y}_i) \right). \quad (9.6)$$

In the following sections, we discuss the measurement of $A(\vec{m}_i|\vec{y}_i)$, $S(\vec{m}_i|\vec{y}_i)$, and $R(\vec{m}_i|\vec{y}_i)$ using data.

9.2.2 Accidental Background PDFs

The accidental background PDF is written as the product of five statistically independent PDFs because the positron and photon are by definition uncorrelated with each other:

$$A(\vec{m}_i|\vec{y}_i) = P_A(E_e|\vec{y}_i)P_A(E_\gamma|\vec{y}_i)P_A(t_{e\gamma}|\vec{y}_i)P_A(\theta_{e\gamma}|\vec{y}_i)P_A(\phi_{e\gamma}|\vec{y}_i). \quad (9.7)$$

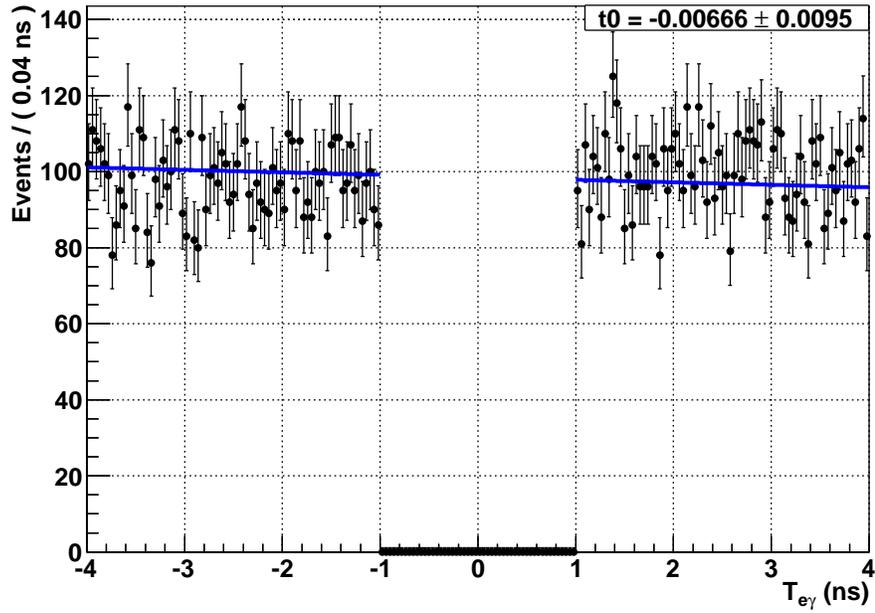
While this is obviously true of E_e and E_γ , there could be a correlation between E_e and $\phi_{e\gamma}$, for example, due to the reasons discussed in section 8.2.2. For accidental events, however, the true positron energy is unknown (as with all the other kinematic variables), and the flatness of the Michel positron energy spectrum within the analysis window effectively washes out the effects of any such correlation.

The accidental $t_{e\gamma}$ spectrum is expected to be flat, but could be modified by trigger effects. This is checked directly on data using events well outside the signal region and within the MEG trigger threshold: $1 < |t_{e\gamma}| < 4$ ns. Figure 9.1 gives the results of both a linear and quadratic fit to the accidental $t_{e\gamma}$ distribution. Since the implied deviations from a uniform (flat) PDF in $t_{e\gamma}$ are not statistically significant, a uniform PDF is used. Furthermore,

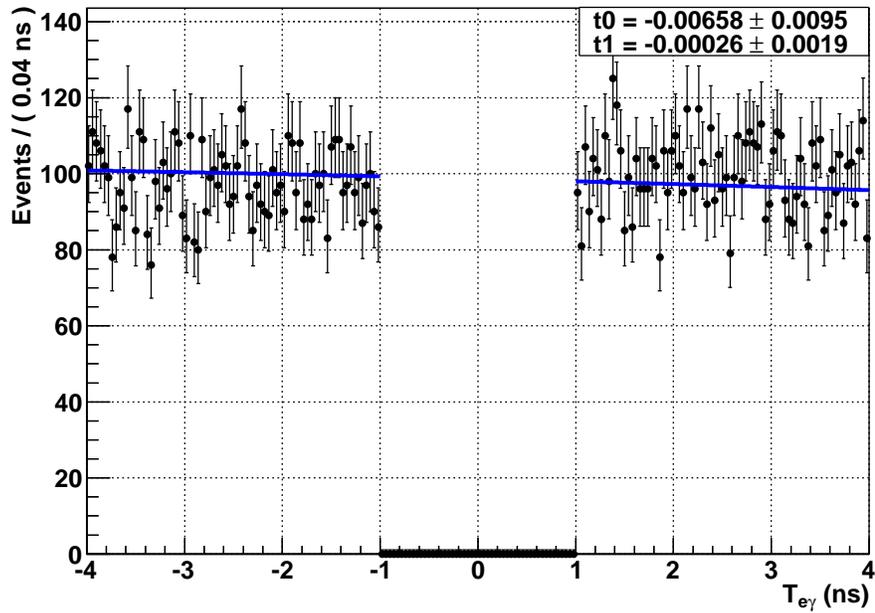
$P_A(t_{e\gamma}|\vec{y}_i) = P_A(t_{e\gamma})$ is assumed because the PDF shape does not change in conjunction with any specific event property.

The accidental relative angle PDFs are measured directly on data. One effect that may be present is a possible bias in the direction match criteria used in the trigger. Aside from that, these distributions should change near the edge of the detector acceptance. Fixing one of either the photon conversion position or the positron emission angles almost fully constrains the other for a given value of $\theta_{e\gamma}$ and $\phi_{e\gamma}$ except for a spread in decay vertex positions at the target. To deal with acceptance and trigger effects, the accidental $\phi_{e\gamma}$ and $\theta_{e\gamma}$ distributions are measured on data, for $1 < |t_{e\gamma}| < 4$ ns, in bins of v and u of the photon conversion position respectively. Figure 9.2 displays these distributions. Each bin in v is fit to a cubic polynomial in $\phi_{e\gamma}$ and the same is done for each bin of u in $\theta_{e\gamma}$. The fit to a given bin describes the relative angle PDF ($\phi_{e\gamma}$ or $\theta_{e\gamma}$) at the average value of v or u in that bin. This ties down the shape of those PDFs at a discrete set of positions, one for each bin. The PDF shapes must then be interpolated between bins. Between two bins, this is done by continuously deforming the shape of the PDF in $\phi_{e\gamma}$ and $\theta_{e\gamma}$ at one bin into the shape of the PDF at the adjacent bin along v and u respectively through a technique described in [64]. Figure 9.3 shows an example of how the shape of the PDFs change between bins. At the edge bins, where no PDF is available to tie down the shape beyond the outermost bins, the fitted PDF shape in that bin is assumed to be constant from the average value of v or u in that bin to the edge of the acceptance cut in v or u , so there is no interpolation at the edges. The complete set of accidental $\phi_{e\gamma}$ distributions for all values of v gives the PDF: $P_A(\phi_{e\gamma}|v)$, and the complete set of accidental $\theta_{e\gamma}$ distributions for all values of u gives the PDF: $P_A(\theta_{e\gamma}|u)$.

The accidental positron energy spectrum is that of the Michel positron energy spectrum. The measured positron energy spectrum includes the effects of the detector acceptance and resolution. As mentioned in section 5.2.4, the Kalman filter used in the positron track recon-

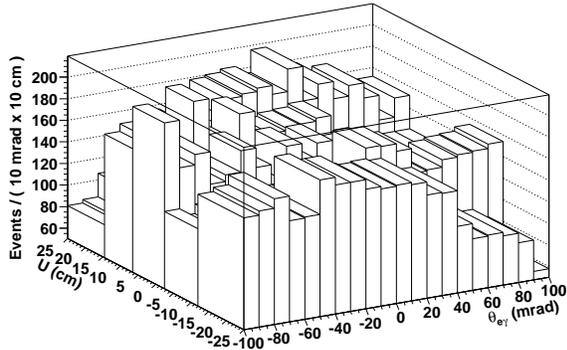


(a) A fit to $1 + t_0 * t_{e\gamma}$.

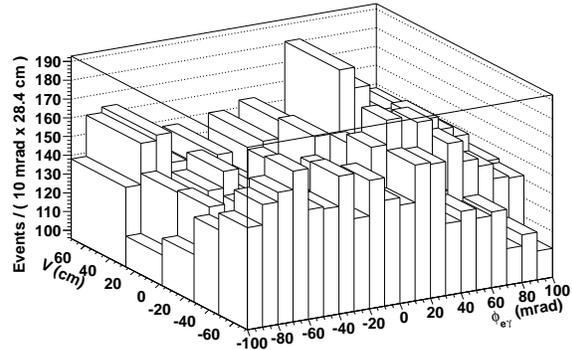


(b) A fit to $1 + t_0 * t_{e\gamma} + t_1 * t_{e\gamma}^2$.

Figure 9.1: Some maximum likelihood fits to the accidental $t_{e\gamma}$ distribution. The fit PDFs are normalized to the data sample, constraining their form.



(a) Accidental distribution of $\theta_{e\gamma}$ in different bins of u .

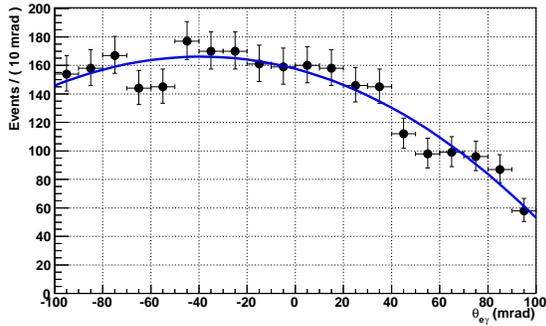


(b) Accidental distribution of $\phi_{e\gamma}$ in different bins of v .

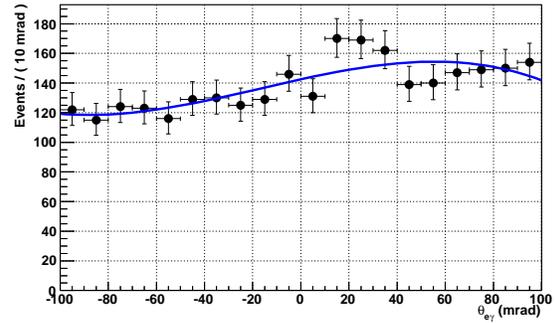
Figure 9.2: Accidental angle distributions used to build the PDFs.

struction provides an event-by-event estimator of the uncertainty in the measured positron energy, δE_e . The energy resolution is expected to worsen as δE_e does, and the acceptance may change as well; both result in the measured accidental positron energy distribution changing its shape according to δE_e . To quantify this, all positrons passing selection criteria are sorted into several bins of δE_e , and each is fit to the model described in section 8.1.2 with a floating acceptance and resolution function. Figure 9.4 illustrates how the spectrum changes with δE_e . This ties down the shape of the spectrum in each of the δE_e bins, and an interpolation between them is done in the same way described for the accidental angle PDFs. The complete set of accidental positron energy distributions for all values of δE_e gives the PDF: $P_A(E_e|\delta E_e)$.

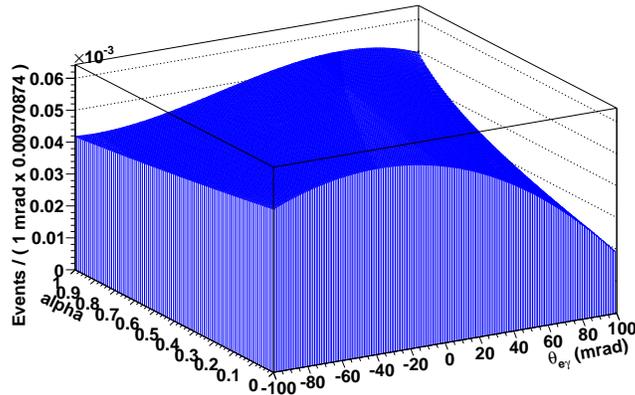
The measured accidental photon energy spectrum is a mixture of photons coming from muon radiative decay, annihilation in flight, and bremsstrahlung, combined with the effects of the detector acceptance, resolution, and pileup, making it difficult to reliably model. For reasons discussed in section 8.1.1, the photon energy resolution is sensitive to the position of the first conversion point in the calorimeter, thus the accidental photon energy spectrum is expected to depend on the conversion point. To ascertain which variations are significant, the photon



(a) A cubic fit to the accidental $\theta_{e\gamma}$ distribution for $u \in [-25, -12)$ cm, the downstream most bin in u .

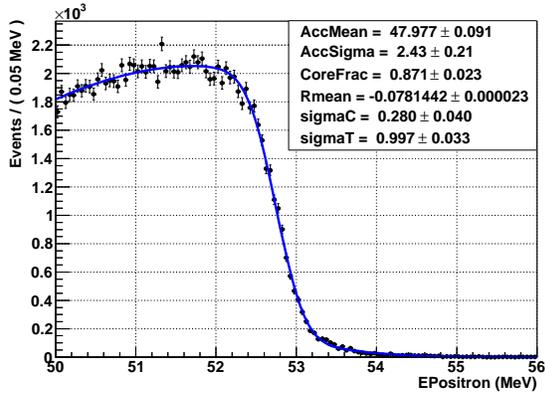


(b) A cubic fit to the accidental $\theta_{e\gamma}$ distribution for $u \in [-12, -2.5)$ cm, the bin adjacent to that of Figure 9.3a.

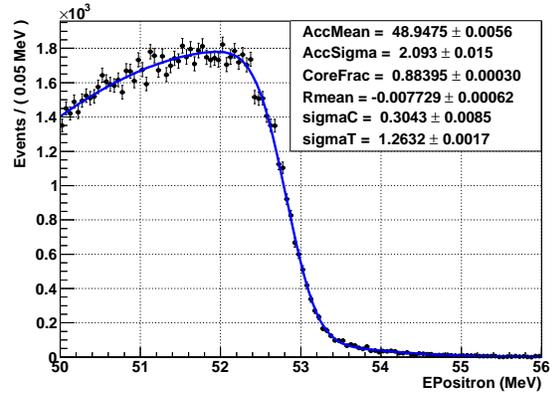


(c) The deformation of the accidental $\theta_{e\gamma}$ PDF from its shape in one bin to its shape in the next bin. Here, $\alpha=0$ represents the PDF as measured in the downstream most bin of u (Figure 9.3a) and $\alpha=1$ is the shape as measured in the adjacent bin of u (Figure 9.3b).

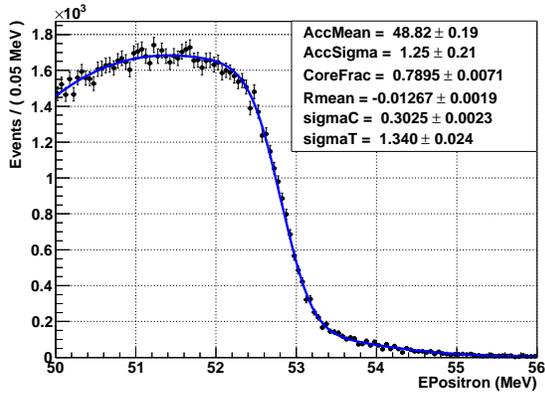
Figure 9.3: An example showing the technique for interpolating the accidental $\theta_{e\gamma}$ PDF between bins of u .



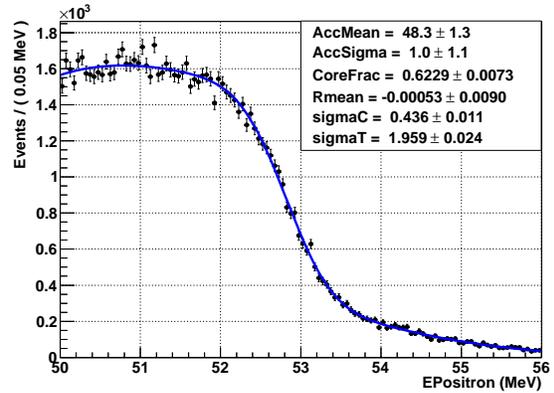
(a) Accidental E_e distribution for $\delta E_e \in [0, 0.345)$ MeV.



(b) Accidental E_e distribution for $\delta E_e \in [0.38, 0.395)$ MeV.



(c) Accidental E_e distribution for $\delta E_e \in [0.4115, 0.433)$ MeV.



(d) Accidental E_e distribution for $\delta E_e \in [0.515, 1.1)$ MeV.

Figure 9.4: The accidental E_e distribution for different bins of δE_e .

energy distribution in the sidebands ($1 < |t_{e\gamma}| < 10$ ns) is independently split into several bins of u, v , and w . A PDF is formed directly from an E_γ histogram of one slice in u (or v or w) and then fit to the E_γ distribution of other slices in u (or v or w), and the reduced χ^2 is assessed. It is found that fitting one slice to other slices gives statistically plausible χ^2 values except in the case of w , indicating that the accidental photon energy spectrum changes most significantly with conversion depth. The accidental photon energy is therefore plotted in different bins of w , and an interpolation is done in the same way described in the above paragraphs. No fit to the spectrum is made; the histogram itself ties down the shape in each bin. Figure 9.5 shows the accidental photon energy PDF in different regions of conversion depth. The complete set of accidental photon energy distributions for all values of w gives the PDF: $P_A(E_\gamma|w)$.

The full accidental PDF can now be written in a specific form:

$$A(\vec{m}_i|\vec{y}_i) = P_A(E_e|\delta E_e)P_A(E_\gamma|w)P_A(t_{e\gamma})P_A(\theta_{e\gamma}|u)P_A(\phi_{e\gamma}|v). \quad (9.8)$$

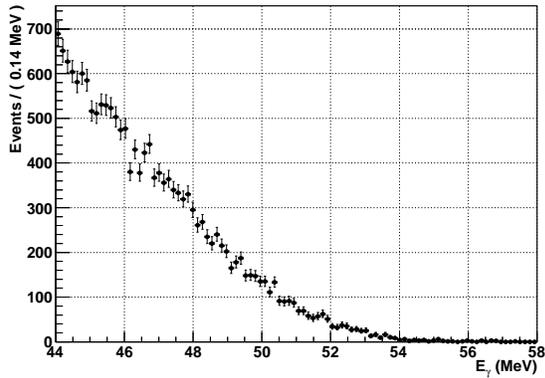
The absence of some or all of the elements in \vec{y}_i in each of the factors on the right hand side is an implicit assumption that the spectrum in each of those kinematic variables depends significantly on only some of the elements in \vec{y}_i , i.e., the ones that appear in the expression.

9.2.3 Signal Event PDFs

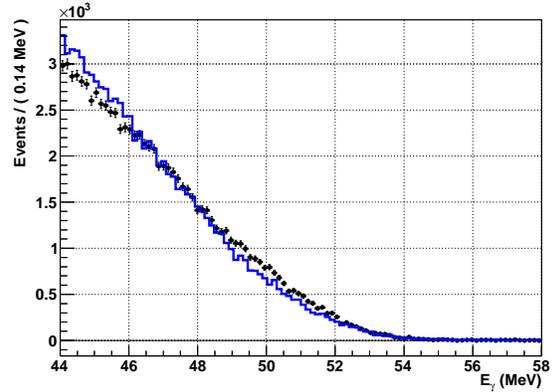
Because of the correlations discussed in section 8.2.2, the signal PDF is written as:

$$S(\vec{m}_i|\vec{y}_i) = P_S(\phi_{e\gamma}|\vec{y}_i, E_e, \theta_{e\gamma})P_S(\theta_{e\gamma}|\vec{y}_i)P_S(E_e|\vec{y}_i)P_S(E_\gamma|\vec{y}_i)P_S(t_{e\gamma}|\vec{y}_i). \quad (9.9)$$

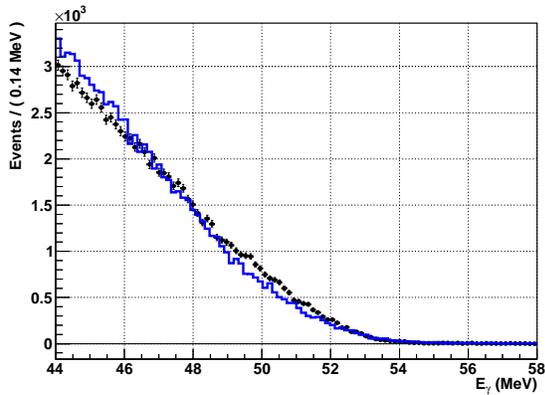
The signal $t_{e\gamma}$ PDF is measured in the way described in section 8.2.1, with the extrapolation to the signal photon energy applied. The relative time resolution is checked for



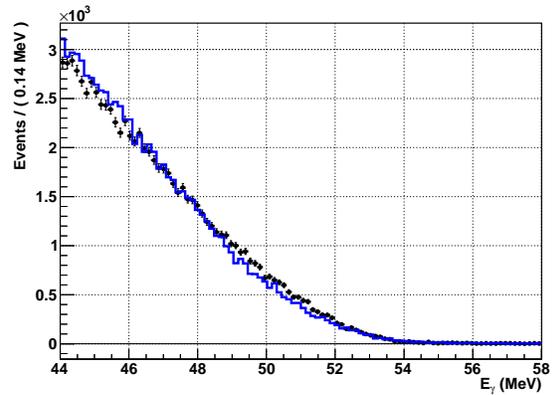
(a) Accidental E_γ distribution for $w \in [0, 1.5)$ cm.



(b) Accidental E_γ distribution for $w \in [1.5, 3.7)$ cm.



(c) Accidental E_γ distribution for $w \in [3.7, 7.2)$ cm.



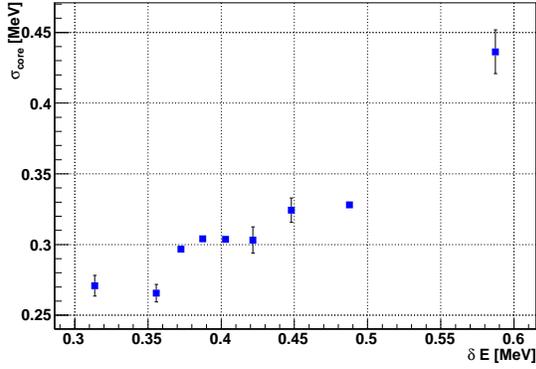
(d) Accidental E_γ distribution for $w \in [7.2, 40)$ cm.

Figure 9.5: The accidental E_γ distribution for different regions of conversion depth. A PDF is formed from the distribution in the first bin and used to fit the other bins as shown in blue.

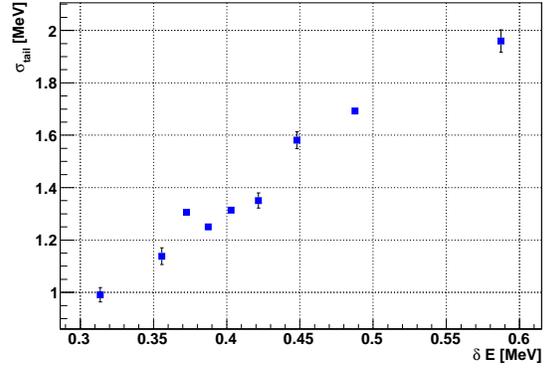
dependencies on a number of different event properties with no statistically significant effects found: $\delta E_e, u, v, w, \Delta R_{DCH-TIC}, \Delta Z_{DCH-TIC}$ and positron track length. Hence, the only event property used for this PDF is θ_e , since the mean of the resolution function is significantly different for upstream and downstream events as discussed in section 8.2.1. The upstream and downstream $t_{e\gamma}$ resolutions give the PDF: $P_S(t_{e\gamma}|\theta_e)$.

Two options are available for measuring the signal E_e PDF, as outlined in section 8.1.2. In order to make the best use of the event-by-event information provided by the Kalman filter, the positron energy resolution must be measured as a function of δE_e . This is difficult to do using two turn events because the two turns have different values of δE_e ; the number of two turn events available in the data set is also statistically inferior to the number of Michel positron events. For these reasons, the fit to the Michel edge is used to obtain the signal E_e PDF. Positrons passing selection cuts in $|t_{e\gamma}| > 1$ ns are binned into several slices of δE_e , and the fit is performed on each. The double Gaussian resolution component of the fit represents the signal E_e PDF at the average value of δE_e within the bin, and an interpolation between them is done in the same way described for the accidental angle PDFs in section 9.2.2. Figure 9.6 plots the change in the components of the resolution with δE_e . The complete set of positron energy resolutions for all values of δE_e gives the PDF: $P_S(E_e|\delta E_e)$.

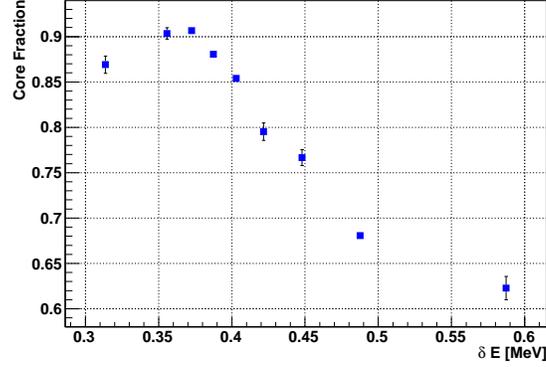
The signal E_γ PDF is measured in several divisions of conversion point location (u,v,w) by the method outlined in section 8.1.1. The usual method of PDF interpolation does not work in this case because the binning is in three dimensions rather than one. Instead, the parameters describing the PDF (σ_{up}, t , and x_0 from Eq. 8.1) are themselves interpolated independently. Because the dependence on v is particularly weak, the measured values of each of three PDF parameters are plotted as a function of u and w for each bin in v. Next, a linear interpolation is done in 2 dimensions (u,w) for each of these plots, resulting in surfaces like the one shown in figure 9.7. For an event with first conversion coordinates (u,v,w), the signal E_γ PDF parameters are first obtained from the u-w surface of the correct bin in v,



(a) Variation of σ_{core} with δE_e .



(b) Variation of σ_{tail} with δE_e .



(c) Variation of the fraction of events in the core with δE_e .

Figure 9.6: Fitted values of the parameters for the double Gaussian E_e resolution for several bins in δE_e . Each data point is plotted at the average value of δE_e within the bin.

as well as from the two neighboring bins in v at the same (u,w) location. From these three values of the PDF parameters at different v locations (but identical u and w locations), a quadratic interpolation in v is performed. Finally, the result represents the signal E_γ PDF parameters for this event. The complete set of photon energy resolutions for all values of u,v , and w gives the PDF: $P_S(E_\gamma|u, v, w)$.

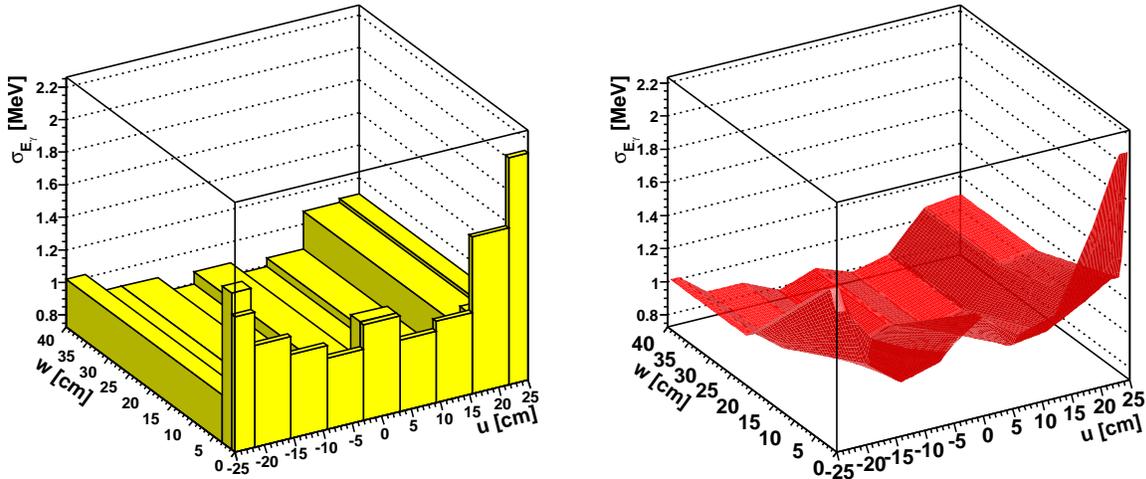


Figure 9.7: The variation of σ_{up} with u and w for fixed v : $v \in [0, 6.19)$ cm. The plot on the left shows the discrete number of bins used to measure σ_{up} . The surface on the right is constructed by a linear interpolation.

The shapes of the signal $\theta_{e\gamma}$ and $\phi_{e\gamma}$ PDFs are constructed from their components as documented in section 8.2.2. The resolution in θ_e is measured from the two-turn method in several bins of δE_e as shown in figure 9.7. Here the problem of ambiguity associated with the two turns having different values of δE_e is unavoidable. It is dealt with by using events satisfying $|\delta E_e^{1st\ turn} - \delta E_e^{2nd\ turn}| < 200$ keV (chosen to balance precision with statistics) and assigning the average of δE_e from the two turns to the event. Figure 9.8 gives the distribution in $\delta E_e^{1st\ turn} - \delta E_e^{2nd\ turn}$ on two-turn events for reference. The double Gaussian resolution fit is taken to represent the θ_e PDF at the average value of δE_e within the bin, and an interpolation between them is done in the usual way described for the accidental angle PDFs in section 9.2.2. The ϕ_e resolution is measured as a function of both δE_e and ϕ_e ; the

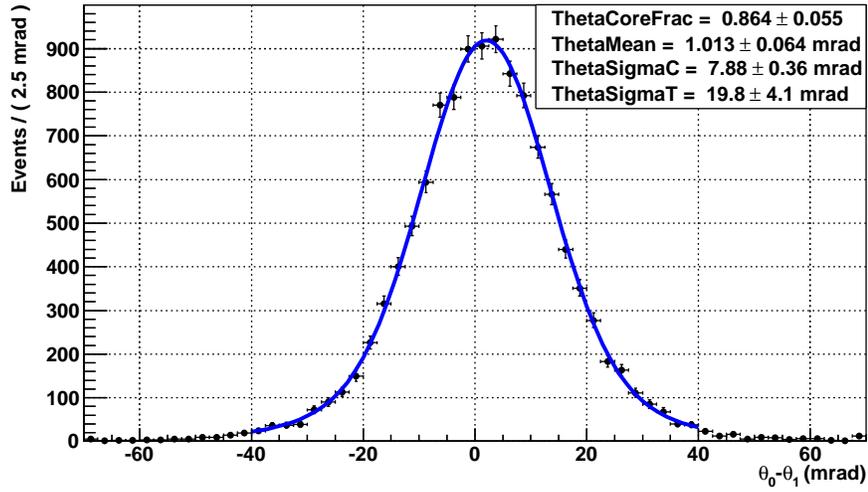
dependence on the latter is a geometric effect: for larger values of $|\phi_e|$, the projection back to the target is more sensitive to errors in the drift chamber position measurements. Since the variation with δE_e is less severe, the ϕ_e resolution is measured in 2 bins of δE_e with the two-turn method. For each slice in δE_e , the ϕ_e resolution is measured in several bins of ϕ_e , and an interpolation in ϕ_e is done in the usual way, with no interpolation in δE_e . The ϕ_e emission angle associated with a two turn event is obtained from the global track fit to both turns and then assigned to both turns. Some example fits are given in figure 9.8.

The calorimeter position resolutions are measured as described in 8.1.1 in bins of position. In lieu of u and v , they are measured as functions of PMT U and PMT V, which are defined as the distance from the center of the nearest PMT in u and v respectively. The $u(v)$ resolutions are three-parameter double Gaussians (σ_{core} , σ_{tail} , and core fraction) measured in bins of PMT U and w (PMT V and w). The three parameters are interpolated independently as in the E_γ PDF by constructing interpolation surfaces in two dimensions. The w resolution is a single parameter Gaussian (σ) measured in bins of PMT U, PMT V, and w . The three dimensions are interpolated in exactly the same way as the E_γ PDF, with PMT U- w surfaces constructed for each bin of PMT V. Figure 9.9 shows an example of this.

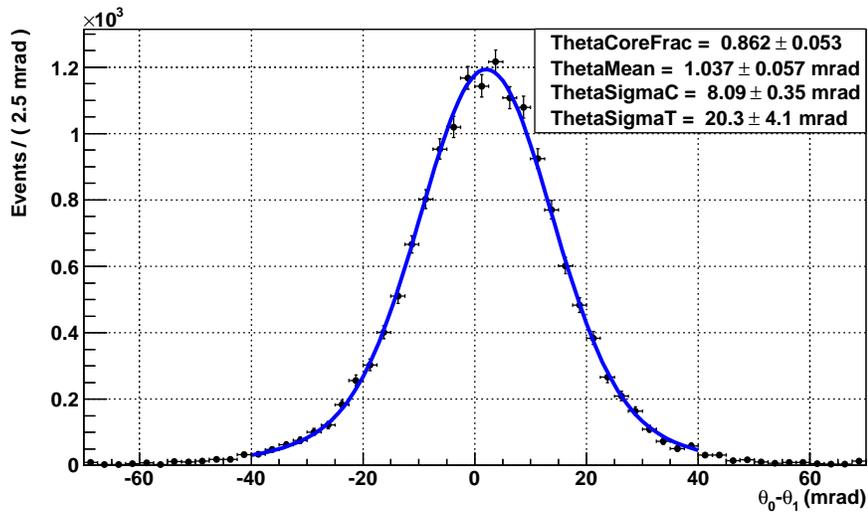
Lastly, the correlations between θ_e and Z , and between ϕ_e and Y as a function of ϕ_e are taken from the measurement described in section 8.2.2. The shapes of the relative angle PDFs are simulated for each event by fixing the shapes of the component resolutions and correlations depending on the values of u, v, w, ϕ_e , and δE_e . The complete set of $\theta_{e\gamma}$ resolutions for all values of u, v, w , and δE_e gives the PDF: $P_S(\theta_{e\gamma}|u, v, w, \delta E_e)$.

The $\phi_{e\gamma}$ PDF requires knowledge of E_e and $\theta_{e\gamma}$ because its mean depends on them. A shift of E_e away from the signal energy shifts the mean of the ϕ_e and hence the $\phi_{e\gamma}$ PDF in a calculable way using the measured correlation between E_e and ϕ_e from section 8.2.2. Such a shift in the mean of the ϕ_e PDF produces a shift in the Y_e PDF because of that correlation (discussed in section 8.2.2), which produces a shift in ϕ_γ since the photon vertex is assumed

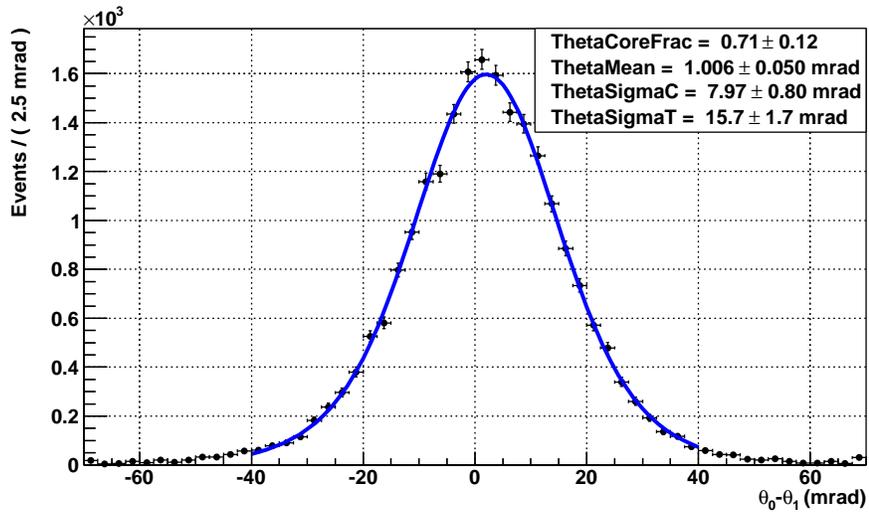
to coincide with the positron. This is taken into account as an additional shift in the mean of the $\phi_{e\gamma}$ PDF. A non-zero value of $\theta_{e\gamma}$ indicates an additional shift of the signal $\phi_{e\gamma}$ PDF mean; however, it is not known how much of this is due to an error in θ_e and how much is due to an error in θ_γ , and it is only the former that causes the signal $\phi_{e\gamma}$ PDF mean to change. Based on the shape of the component resolutions, it is possible to predict an average error in $\phi_{e\gamma}$ due to an error in $\theta_{e\gamma}$. If, for example, the resolution in θ_e is wide compared to the resolution in θ_γ , then a signal event with a non-zero $\theta_{e\gamma}$ is likely to be due mostly to an error in θ_e , and the slope $\frac{\Delta\phi_{e\gamma}}{\Delta\theta_{e\gamma}}$ is expected to be large. For each event, the angle and position resolutions are known as described above, so they are used to simulate the error in $\phi_{e\gamma}$ due to an error in $\theta_{e\gamma}$ for the event. Figure 9.10 gives an example of such a distribution. These plots are fit to a line to obtain the correlation between $\phi_{e\gamma}$ and $\theta_{e\gamma}$ for use in shifting the mean of the signal $\phi_{e\gamma}$ PDF. The slopes in $\frac{\Delta\phi_{e\gamma}}{\Delta\theta_{e\gamma}}$ range from -0.16 to -1.1. The shifts due to errors in E_e and $\theta_{e\gamma}$ are uncorrelated and simply added to get the total shift of the signal $\phi_{e\gamma}$ PDF mean from zero. The complete set of $\phi_{e\gamma}$ shapes and means for all values of $v, w, \phi_e, \delta E_e, E_e$ and $\theta_{e\gamma}$ gives the PDF: $P_S(\phi_{e\gamma}|v, w, \delta E_e, E_e, \theta_{e\gamma})$.



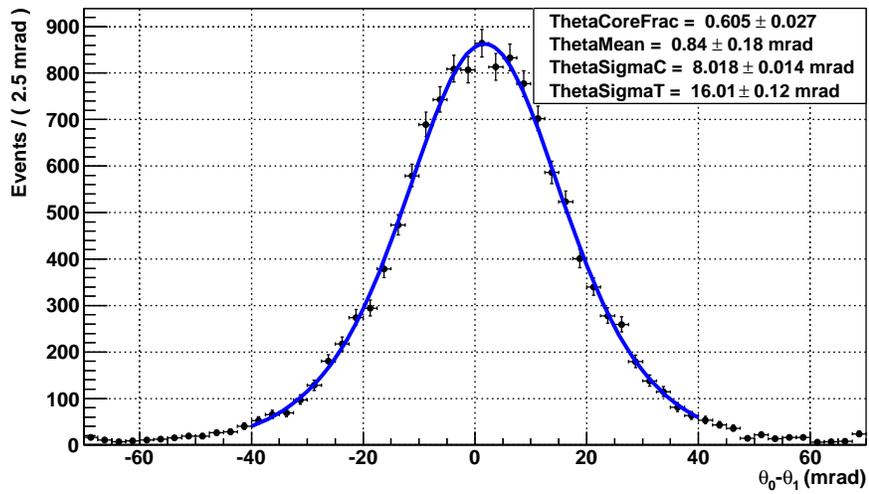
(a) A fit to the θ_e resolution for $\delta E_e \in [0, 0.242)$ MeV.



(b) A fit to the θ_e resolution for $\delta E_e \in [0.242, 0.266)$ MeV.



(c) A fit to the θ_e resolution for $\delta E_e \in [0.266, 0.32)$ MeV.



(d) A fit to the θ_e resolution for $\delta E_e \in [0.32, 1.1)$ MeV.

Figure 9.7: Fits to the difference in θ_e on double turn events in several bins of δE_e . The fit parameters shown are for the double Gaussian that is convolved with itself and fit to the distribution.

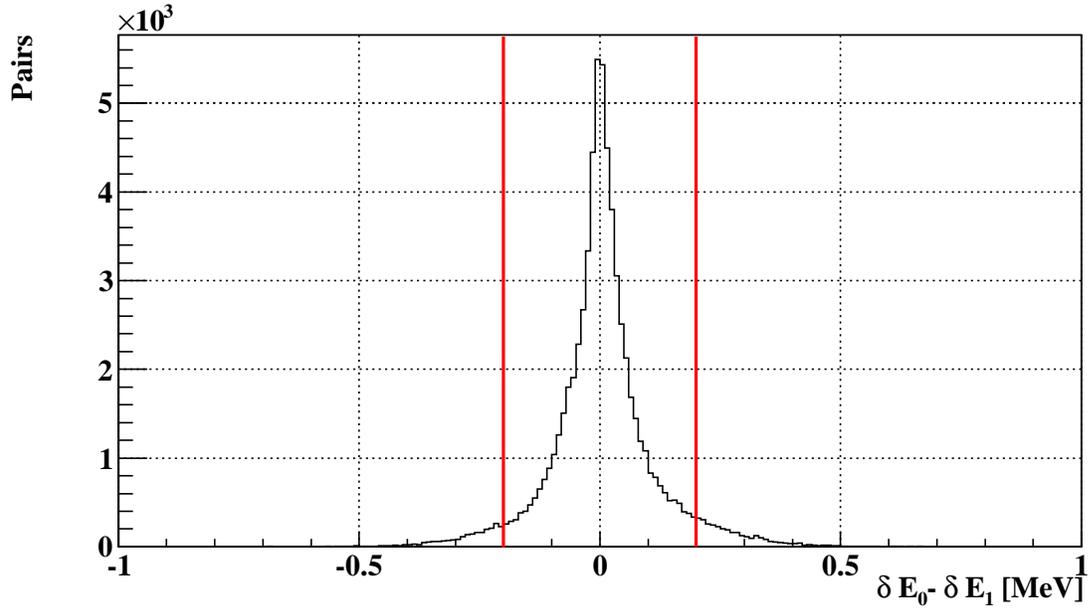


Figure 9.8: The distribution of $\delta E_e^{1st\ turn} - \delta E_e^{2nd\ turn}$. The region between the red lines is used to construct positron angle PDFs.

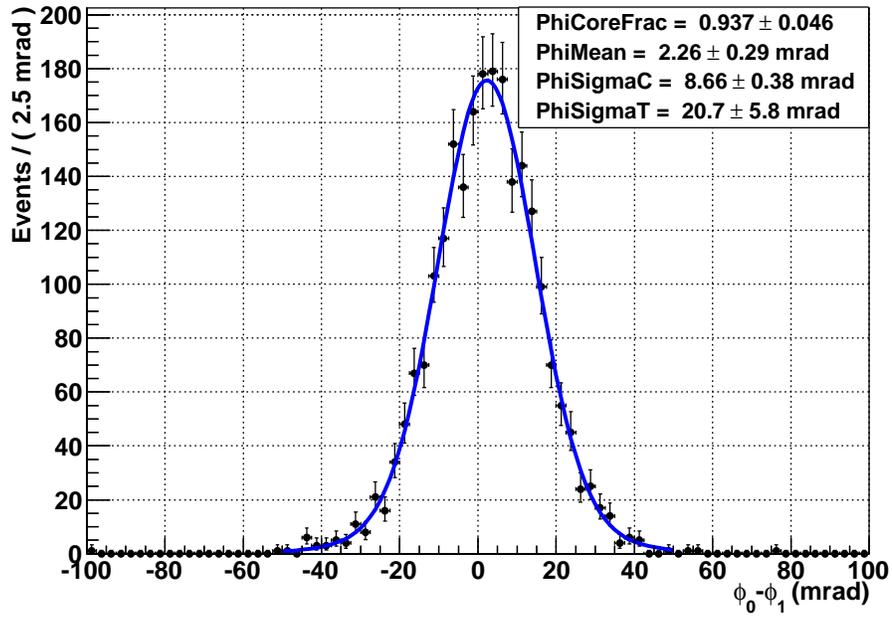
The full signal PDF can now be written in a specific form:

$$\begin{aligned}
 S(\vec{m}_i|\vec{y}_i) &= P_S(\phi_{e\gamma}|v, w, \delta E_e, E_e, \theta_{e\gamma})P_S(\theta_{e\gamma}|u, v, w, \delta E_e) \\
 &\quad \times P_S(E_e|\delta E_e)P_S(E_\gamma|u, v, w)P_S(t_{e\gamma}|\theta_e).
 \end{aligned}
 \tag{9.10}$$

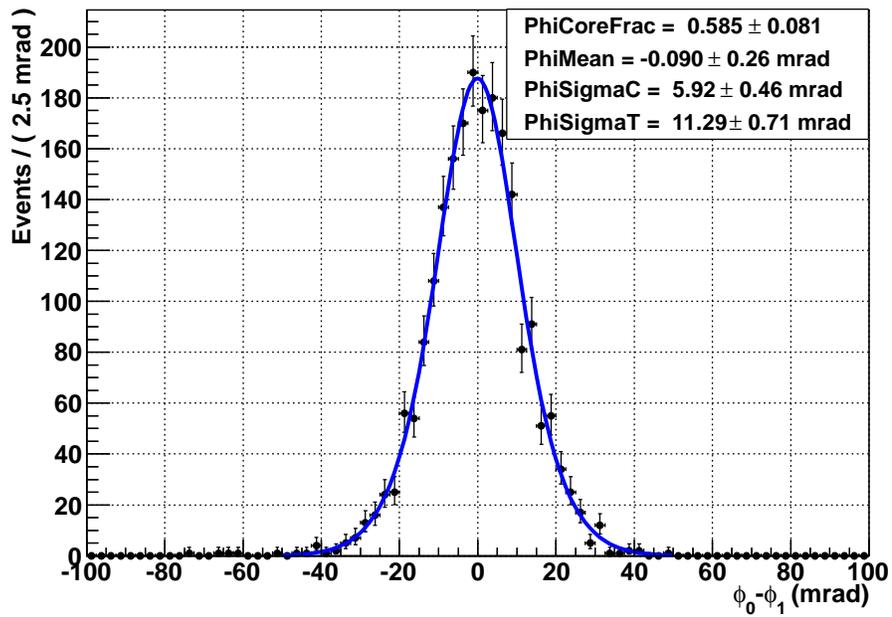
9.2.4 Radiative Decay Background PDFs

The radiative decay branching ratio from theory is a function of $E_e, E_\gamma, \theta_{e\gamma}$ and $\phi_{e\gamma}$, making these automatically correlated, but not $t_{e\gamma}$. It follows that the radiative decay background PDF can be written as:

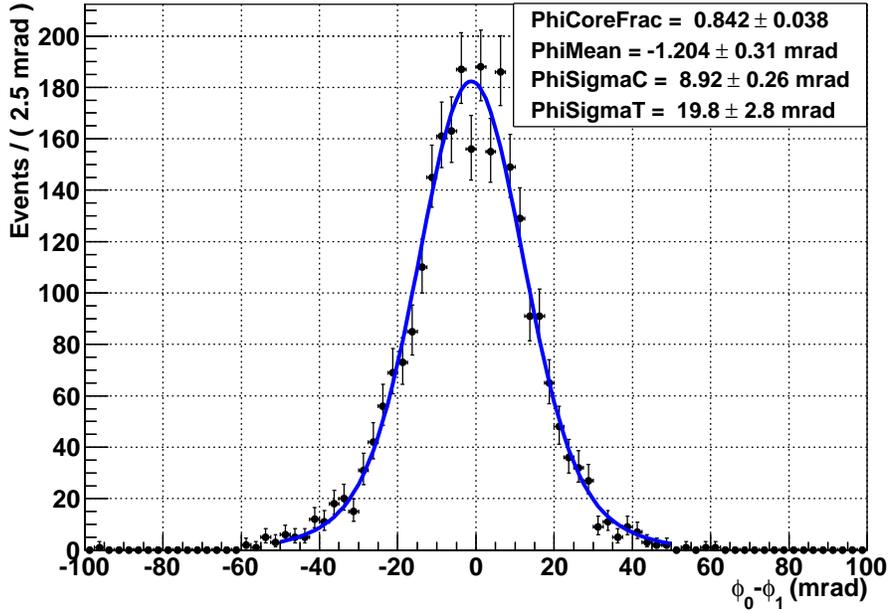
$$R(\vec{m}_i|\vec{y}_i) = P_R(E_e, E_\gamma, \theta_{e\gamma}, \phi_{e\gamma}|\vec{y}_i)P_R(t_{e\gamma}|\vec{y}_i).
 \tag{9.11}$$



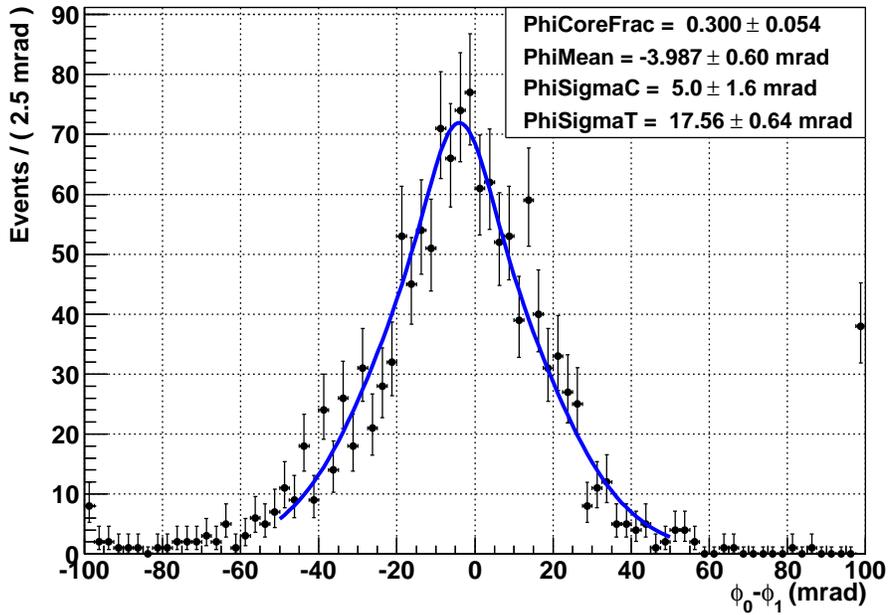
(a) A fit to the ϕ_e resolution for $\phi_e \in [-950, -400]$ mrad.



(b) A fit to the ϕ_e resolution for $\phi_e \in [-400, 100]$ mrad.



(c) A fit to the ϕ_e resolution for $\phi_e \in [100, 500]$ mrad.



(d) A fit to the ϕ_e resolution for $\phi_e \in [500, 950]$ mrad.

Figure 9.8: Fits to the difference in ϕ_e on double turn events in several bins of ϕ_e for fixed $\delta E_e \in [0, 0.258)$ MeV. The fit parameters shown are for the double Gaussian that is convolved with itself and fit to the distribution.

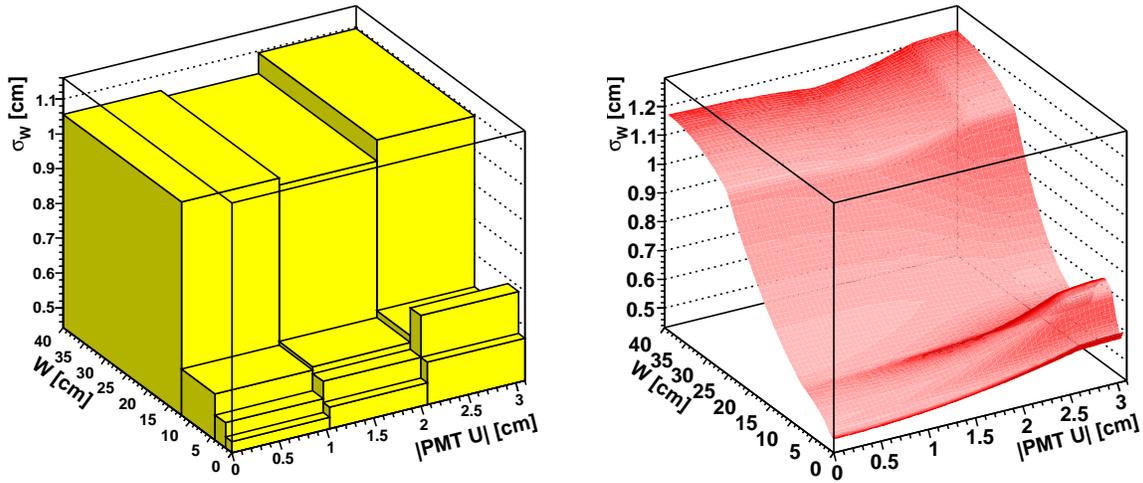


Figure 9.9: The variation of depth resolution with $|PMT U|$ and w for fixed $|PMT V|$: $|PMT V| \in [1.033, 2.067)$ cm. The plot on the left shows the discrete number of bins used to measure σ_W . The surface on the right is constructed by a second order interpolation.

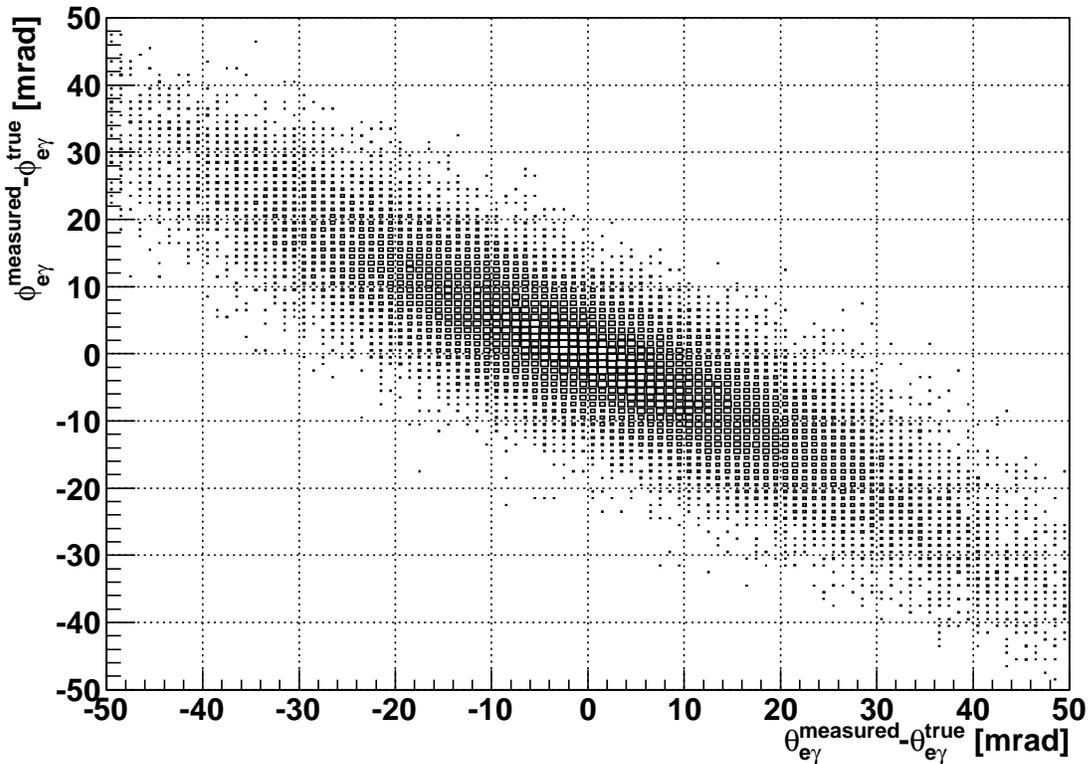
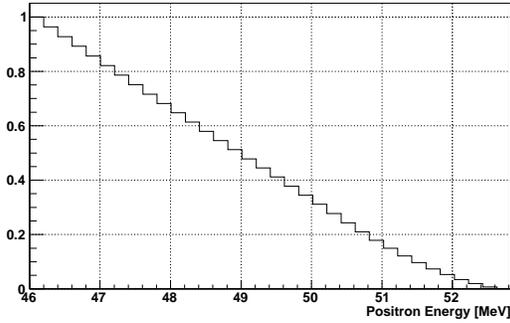


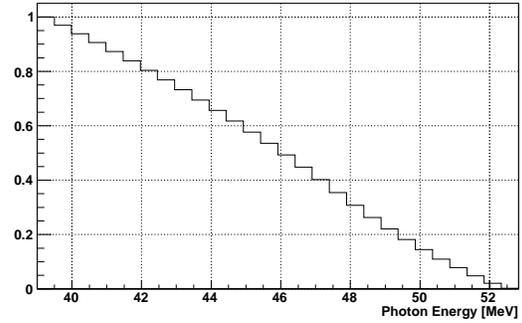
Figure 9.10: Simulation of $\phi_{e\gamma}$ vs. $\theta_{e\gamma}$ for the resolution functions specific to a certain event.

Appendix A gives the tree level differential branching ratio for $\mu^+ \rightarrow e^+ \gamma \bar{\nu}_\mu \nu_e$. This must first be converted into an expression in terms of $E_e, E_\gamma, \theta_{e\gamma}$ and $\phi_{e\gamma}$, the relevant observables. A binned probability density map in $E_e, E_\gamma, \theta_{e\gamma}$ and $\phi_{e\gamma}$ is prepared by numerically integrating the six input variables of this expression over the fiducial detector volumes, assuming zero muon polarization, and performing the change of variables. This represents the true radiative decay spectrum, and the measured radiative decay spectrum in the limit of perfect detector resolutions and the absence of acceptance effects. Figure 9.11 presents the projected one-dimensional distributions in each of these variables. In reality, the measured kinematic radiative decay PDF must be corrected for detector acceptance variations and response functions over each of these four variables as well as the correlations between errors in E_e and $\phi_{e\gamma}$, and $\theta_{e\gamma}$ and $\phi_{e\gamma}$. This involves multiplying by acceptance functions, and performing a four-dimensional convolution of the true spectrum with the resolution functions taking into account the correlations between errors. Since the acceptance effects, resolution functions, and strength of the correlations depend on the event properties (i.e. \vec{y}), a separate PDF is prepared for each individual event. Figure 9.12 shows an example of the resulting PDF for a particular event. The positron energy acceptance is taken from the fits to the Michel edge of section 9.2.3. The acceptance function from the fit to each bin of δE_e represents the acceptance at the average value of δE_e in that bin, and an interpolation between them is done in the same way described for the accidental angle PDFs in section 9.2.2. The relative angle acceptances are taken from the measured accidental angle distributions from section 9.2.2. As such, they are functions of u and v respectively. The photon energy acceptance, which comes entirely from the trigger efficiency, is estimated from the ratio of events that fire the MEG trigger to events that fire a trigger with a lower photon energy threshold. This is well modeled by an error function whose parameters are measured at different position bins in (u, v, w) and then interpolated in the same way the signal E_γ PDF of section 9.2.3 is interpolated. Since the acceptance is essentially flat above ~ 46 MeV, it is only important in the bottom sideband. The positron energy, photon energy, and

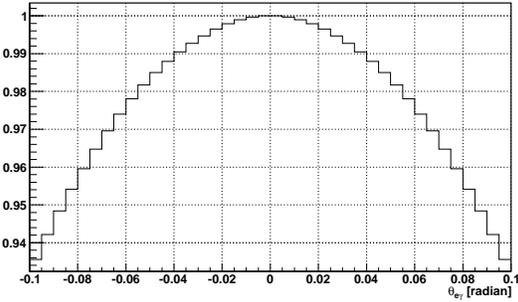
relative angle resolution functions are the same as those used in the signal PDFs. The correlations between errors in E_e and $\phi_{e\gamma}$ and between $\theta_{e\gamma}$ and $\phi_{e\gamma}$ are known from section 9.2.3 and included in performing the convolution. The resulting density map in the kinematic quantities for any given set of parameters $\{u, v, w, \delta E_e, \phi_e\}$ gives the kinematic part of the radiative decay PDF: $P_R(E_e, E_\gamma, \theta_{e\gamma}, \phi_{e\gamma}|u, v, w, \delta E_e, \phi_e)$.



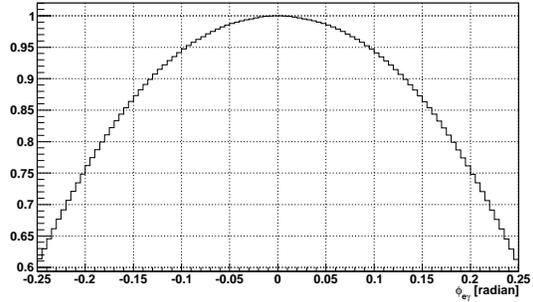
(a) Projection of true radiative decay spectrum in E_e .



(b) Projection of true radiative decay spectrum in E_γ .



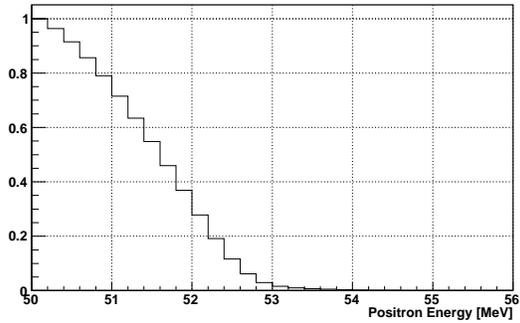
(c) Projection of true radiative decay spectrum in $\theta_{e\gamma}$.



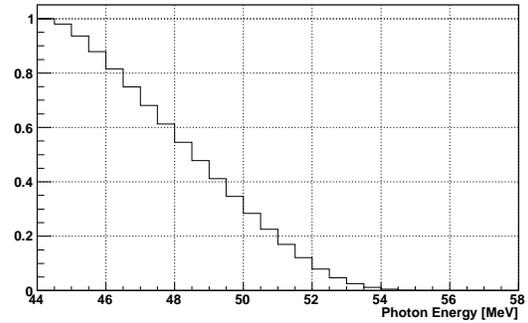
(d) Projection of true radiative decay spectrum in $\phi_{e\gamma}$.

Figure 9.11: The one-dimensional true radiative decay distributions. Each plot is made by integrating out the other three variables. They are normalized to unity at their maximum value.

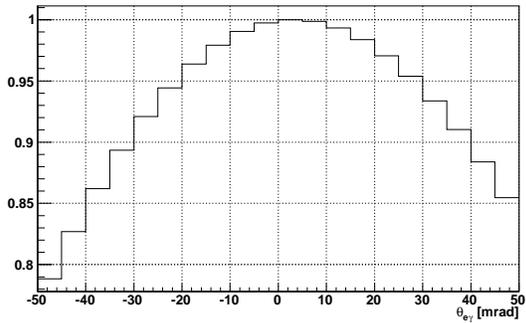
The radiative decay $t_{e\gamma}$ PDF is the same as for signal except that the extrapolation of the resolution to the signal energy is not applied. Because the true photon energy for radiative decays is unknown (unlike signal photons), the E_γ dependence of the timing resolution is not used. The upstream and downstream $t_{e\gamma}$ resolutions give the PDF: $P_R(t_{e\gamma}|\theta_e)$.



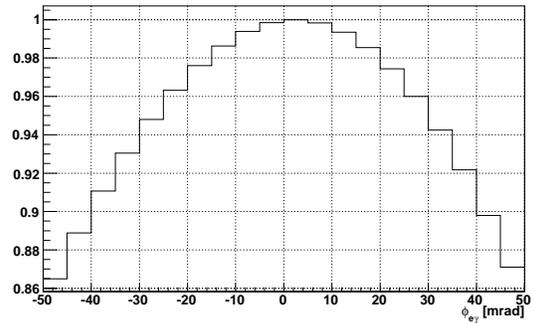
(a) Projection of radiative decay PDF in E_e .



(b) Projection of radiative decay PDF in E_γ .



(c) Projection of radiative decay PDF in $\theta_{e\gamma}$.



(d) Projection of radiative decay PDF in $\phi_{e\gamma}$.

Figure 9.12: The one-dimensional radiative decay PDF distributions for a certain event. Each plot is made by integrating out the other three variables. They are normalized to unity at their maximum value. Note that the projected angle PDFs do not reach zero within the analysis window.

The full radiative decay PDF can now be written in a specific form:

$$R(\vec{m}_i|\vec{y}_i) = P_R(E_e, E_\gamma, \theta_{e\gamma}, \phi_{e\gamma}|u, v, w, \delta E_e, \phi_e)P_R(t_{e\gamma}|\theta_e). \quad (9.12)$$

9.3 Background Estimates for the Signal Window

9.3.1 Accidental Background

Based on the assumption of a flat accidental background PDF for $t_{e\gamma}$, which is justified in section 9.2.2, the number of events in the left and right sidebands can be used to estimate the expected number of accidental background events in the signal window. Together, the two sidebands contain 500 events over a total of 2.8 ns in $t_{e\gamma}$, giving the result $\frac{500 \pm 22.36 \text{ accidentals}}{2.8 \text{ ns}}$. This predicts 250 ± 11.18 accidentals in the hidden time region ($|t_{e\gamma}| < 0.7$ ns), and a total of 750 ± 11.18 accidentals throughout the entire signal window ($|t_{e\gamma}| < 2.1$ ns).

9.3.2 Radiative Decay Background

The expected number of radiative muon decays in the signal window is calculated by performing a likelihood fit to the bottom sideband, in which the number of radiative decays is allowed to float, and scaling the result to the signal window. A fit to the bottom sideband with the constraint $N_{Sig} = 0$ gives $N_{RD} = 74.5^{+20.1}_{-21.0}$, where the errors are the 1σ uncertainties calculated by MINOS as part of the MINUIT package.[65] The impact of systematic uncertainties in the radiative decay PDF is estimated by reproducing the PDF with different acceptances, resolution functions, and correlations according to the uncertainties in each. Since recreating the kinematic radiative decay PDF for each event is extremely time consuming, the systematics are checked instead by recreating a single kinematic PDF using

average resolution functions and acceptances for each systematic uncertainty, and an alternate fit to the bottom sideband is performed in each case. The change in N_{RD} is a measure of the impact of each uncertainty. The largest effect comes from the uncertainty in the width of the positron energy resolution (4% change in N_{RD}). Even when all the changes in N_{RD} are added in quadrature, however, the total systematic uncertainty is negligible compared to the statistical uncertainty on N_{RD} of 28%. The result is extrapolated into the signal window by calculating the ratio of the integral of the radiative decay PDF over the signal window to the integral over the bottom sideband. This predicts $N_{RD} = 29.42 \pm 8.2$ in the signal window.

9.4 Setting a Confidence Interval

In addition to the best estimate of the number of signal events, one needs a measure of its precision. This is the purpose of constructing a confidence interval. In the frequentist interpretation, the true number of signal events is an exact but unknown number that does not follow a probability distribution, while the best estimate of it *does* follow a probability distribution. Thus, to the frequentist, a 90% confidence interval is one such that if an ensemble of similar MEG experiments were performed with measurement errors distributed according to the assumed PDFs, and with each doing a likelihood analysis and calculating a confidence interval in the same prescribed way, then 90% of the ensemble of confidence intervals would contain the true value of the number of signal events. Feldman and Cousins proposed a method[66] for computing confidence belts in the case of a likelihood function of a single variable. When nuisance parameters are present, N_{Acc} and N_{RD} in this case, one way of dealing with them is the profile likelihood method[67], which is adopted in this analysis. Since the radiative decay PDF density is sparse in the region where the signal PDF density is high when compared to the accidental PDF density in that region (see section 3.2),

errors in N_{RD} are not strongly correlated with errors in N_{Sig} . N_{RD} is therefore fixed in the likelihood fit based on the expectation to simplify the analysis. Additionally, the physical constraints $N_{Sig} \geq 0$ and $N_{Acc} \geq 0$ are enforced. The best fit values are denoted as \widehat{N}_{Sig} and \widehat{N}_{Acc} , and the fixed value used for N_{RD} is denoted as N_{RD}^{fix} .

Once a likelihood fit is performed on a data sample, a number of test values for N_{Sig} are scanned and the confidence level at each is evaluated. The union of all points with a confidence level of 90% or less forms the 90% confidence interval. When $N_{Sig} = 0$ falls within this region, only an upper limit is reported. The confidence level at each test point, N_{Sig}^i , is evaluated as follows:

1. The test statistic for data is calculated:

$$R_{data}^i = \frac{\mathcal{L}(N_{Sig}^i, N_{RD}^{fix}, \widehat{N}_{Acc}(N_{Sig}^i))}{\mathcal{L}(\widehat{N}_{Sig}, N_{RD}^{fix}, \widehat{N}_{Acc})}. \quad (9.13)$$

The denominator is simply the maximum value of the likelihood. The numerator is the likelihood evaluated at the test value N_{Sig}^i , the same fixed value of radiative decays N_{RD}^{fix} , and the value for the number of accidentals that maximizes the likelihood with the other parameters fixed as such $\widehat{N}_{Acc}(N_{Sig}^i)$.

2. An ensemble of similar experiments is simulated taking the true values of the event numbers to be N_{Sig}^i , N_{RD}^{fix} , and \widehat{N}_{Acc} . All three of these are fluctuated according to a Poisson distribution. For each simulated event in a simulated experiment, a set of event properties, \vec{y} , is drawn randomly from the actual events in the data sample. These are used to build the necessary PDFs and the kinematic observables, \vec{m} , are generated according to them.
3. A likelihood fit is performed on each simulated experiment to obtain the best fit parameters. For the j'th simulated experiment, the best fit parameters are denoted as: \widehat{N}_{Sig}^j , N_{RD}^{fix} (this is still fixed), and \widehat{N}_{Acc}^j . The test statistic for the j'th simulated experiment

is calculated:

$$R_{sim_j}^i = \frac{\mathcal{L}(N_{Sig}^i, N_{RD}^{fix}, \widehat{N}_{Acc}^j(N_{Sig}^i))}{\mathcal{L}(\widehat{N}_{Sig}^j, N_{RD}^{fix}, \widehat{N}_{Acc}^j)}. \quad (9.14)$$

Again, the denominator is the maximum value of the likelihood, and the numerator is the likelihood evaluated at the test value N_{Sig}^i , N_{RD}^{fix} , and the number of accidentals that maximizes the likelihood for this particular simulation with the other parameters fixed as such \widehat{N}_{Acc}^j .

4. The confidence level at the test point is the probability

$$P(R_{data}^i < R_{sim}^i) \quad (9.15)$$

calculated over the simulated experiments.

9.5 Normalizing the Experiment

In order to convert a number of signal events into a measurement of or upper limit for $BR(\mu^+ \rightarrow e^+\gamma)$, a normalization factor must be calculated for the experiment. The Michel decay ($\mu^+ \rightarrow e^+\nu_e\bar{\nu}_\mu$), which nearly saturates the total muon decay rate, is used as the normalization channel because many efficiency and acceptance factors affecting the total number of detected events are common to both Michel and signal modes and cancel. A special trigger type that makes the same requirement on the positron as the MEG trigger but does not require a photon to fire is used to collect a normalization data sample (of Michel events) and is taken at the same time as the MEG data but with a high prescale factor. The number of detected signal events, N_{Sig} , and the number of detected Michel events, N_{Michel} ,

within the signal window can both be expressed as a product of similar factors:

$$N_{Sig} = BR(\mu^+ \rightarrow e^+\gamma) \times N_\mu \times \tau \times P_S \times \epsilon_S^{eT} \times G_S^e \times \epsilon_S^e \times G_S^{\gamma(e)} \times \epsilon_S^{\gamma(e)} \quad (9.16)$$

$$N_{Michel} = BR(\mu^+ \rightarrow e^+\nu_e\bar{\nu}_\mu) \times N_\mu \times \tau \times P_M \times \epsilon_M^{eT} \times G_M^e \times \epsilon_M^e \times f_M. \quad (9.17)$$

The number of muons stopped during the data taking, N_μ , and the livetime, τ , are the same in both cases since the two samples are taken concurrently. P_S and P_M are the prescale factors for the MEG and Michel triggers respectively. ϵ_S^{eT} and ϵ_M^{eT} are respectively the probabilities to satisfy the MEG and Michel trigger conditions if an event satisfies all the analysis selection criteria. G_S^e and G_M^e are respectively the geometric acceptances for signal and Michel positrons; they are made equal by imposing the same acceptance cut on the positrons in both data samples. ϵ_S^e and ϵ_M^e are respectively the signal and Michel positron efficiencies for passing through event reconstruction and selection criteria. The remaining factors for signal account for the photon efficiency. $G_S^{\gamma(e)}$ is the conditional geometric acceptance for a signal photon given that the positron is in the acceptance, and $\epsilon_S^{\gamma(e)}$ is the signal photon efficiency for passing event reconstruction and selection criteria. Lastly, f_M is the fraction of the Michel positron energy spectrum that is used to count Michel events.

Solving Eq.s 9.16 and 9.17 for $BR(\mu^+ \rightarrow e^+\gamma)$ gives:

$$\frac{BR(\mu^+ \rightarrow e^+\gamma)}{BR(\mu^+ \rightarrow e^+\nu_e\bar{\nu}_\mu)} = \frac{N_{Sig}}{N_{Michel}} \times \frac{P_M}{P_S} \times \frac{\epsilon_M^{eT}}{\epsilon_S^{eT}} \times \frac{\epsilon_M^e}{\epsilon_S^e} \times f_M \times \frac{1}{G_S^{\gamma(e)} \epsilon_S^{\gamma(e)}}. \quad (9.18)$$

Michel positrons in the energy range $50 < E_e < 56$ MeV are used for the normalization sample, giving $N_{Michel} = 16,294$. The prescale factors are known: $P_S = 1$ and $P_M = 1.17 \times 10^7 \pm 1\%$. The conditional MEG trigger efficiency is measured using data taken with

a trigger type not containing a direction match requirement on the positron and photon. The probability to fire the MEG trigger is found as a function of the measured opening angle and this is extrapolated to the opening angle for signal. Comparing this result with an independent estimate from signal Monte Carlo gives the uncertainty. The result is $\epsilon_S^{eT} = 0.915 \pm 1\%$. The ratio of positron reconstruction/selection cut efficiencies can be decomposed: $\frac{\epsilon_M^e}{\epsilon_S^e} = \frac{\epsilon_M^{eR}}{\epsilon_S^{eR}} \times \frac{1}{\epsilon_S^{eEcut}} \cdot \frac{\epsilon_M^{eR}}{\epsilon_S^{eR}}$ is the ratio of efficiencies to reconstruct the positron and pass all but the energy cut, which is different than one because detector acceptance increases with energy; this measurement is explained in section 8.3.2. ϵ_S^{eEcut} is the efficiency for a signal positron to have a reconstructed energy inside the analysis window: $50 < E_e < 56$ MeV. This is measured from the fraction of the positron energy resolution function within this window by averaging over the event-by-event positron energy resolution shapes in the left and right sidebands. The uncertainty is estimated by comparing it with the fraction of the average E_e resolution function inside that window. The result is $\epsilon_S^{eEcut} = 0.993 \pm 0.2\%$. f_M is obtained by computing the fraction of the Michel momentum spectrum above 50 MeV, and the uncertainty is estimated by repeating this with a different lower bound consistent with the uncertainty in the absolute positron energy scale giving $f_M = 0.1004 \pm 1.8\%$. The final factor can be decomposed: $\frac{1}{G_S^{\gamma(e)} \epsilon_S^{\gamma(e)}} = \frac{1}{G_S^{\gamma(e)} \epsilon_S^{\gamma R(e)} \epsilon_S^{\theta cut} \epsilon_S^{\phi cut}} \cdot \epsilon_S^{\gamma R(e)}$ is the signal photon efficiency to reconstruct and pass all but the angle cuts. $\epsilon_S^{\theta cut}$ and $\epsilon_S^{\phi cut}$ are the efficiencies for a signal pair to have relative angles inside the analysis window of $|\theta_{e\gamma}, \phi_{e\gamma}| < 50$ mrad. The result for the product $G_S^{\gamma(e)} \epsilon_S^{\gamma R(e)}$ is discussed in section 8.3.1. The remaining factors are computed in the same way as ϵ_S^{eEcut} , resulting in $\epsilon_S^{\theta cut} = 0.987 \pm 0.2\%$ and $\epsilon_S^{\phi cut} = 0.917 \pm 5.7\%$.

The final result is

$$\frac{BR(\mu^+ \rightarrow e^+ \gamma)}{BR(\mu^+ \rightarrow e^+ \nu_e \bar{\nu}_\mu)} = N_{Sig} \times (9.7 \times 10^{-13} \pm 10\%). \quad (9.19)$$

9.6 Inclusion of Systematic Uncertainties

The systematic uncertainties in the shape of the PDFs and normalization are estimated and incorporated into the confidence level calculation as described below.

9.6.1 Signal PDFs

The systematic uncertainties on the mean and σ of the signal $t_{e\gamma}$ PDF are estimated from the statistical error on those parameters from the fit to the radiative decay timing peak. Both parameters are fluctuated from experiment to experiment in the fit PDF during the confidence level calculation by their uncertainties.

The absolute positron energy scale (or mean of the signal E_e PDF) is obtained by treating it as a floating parameter in the Michel fit to the accidental positrons. The uncertainty in determining this parameter by such a fit is estimated by changing each of the other parameters in the acceptance and resolution function by their statistical fit uncertainty and refitting to check the change in the fitted energy scale. Summing all of these effects in quadrature gives a measure of the uncertainty in the positron energy scale. The systematic uncertainty in the signal E_e PDF shape is encapsulated in a single parameter for the uncertainty of the full RMS within the signal window. This is done by measuring the signal E_e PDF as a function of δE_e on Monte Carlo in the same way described in section 9.2.3 and again using the same technique but measuring the PDF shapes from $E_e^{meas} - E_e^{true}$. This results in two sets of signal E_e PDFs for any value of δE_e : true and measured. The full RMS of each is calculated and compared over the distribution of δE_e values in the left and right sidebands. It is found that there is both a systematic overestimate of the full RMS of the measured compared to the true PDF on average as well as a stochastic fluctuation about that value. The systematic portion is handled by generating all experiments from an accordingly tighter PDF than that

used for fitting. The stochastic component is incorporated by fluctuating the full RMS of the fit PDF by its value from experiment to experiment.

The absolute photon energy scale (or peak of the signal E_γ PDF) is obtained from the 55 MeV peak in the charge exchange data. Several sources of uncertainty are added in quadrature: the deviation of the 82.9 MeV peak from expectation, the estimated uncertainty of the non-uniformity corrections of section 5.1.2, the estimated uncertainty in the PMT gains, and the statistical uncertainty of the fits to 55 MeV data. As for the signal E_γ PDF shape, both the uncertainties in σ_{up} and t as well as the correlation between them are taken from the statistical error on the fits to the 55 MeV peak. All of these parameters are fluctuated accordingly in the fit PDF from experiment to experiment during the confidence level calculation.

The relative alignment of the DCH and XEC detectors is checked using cosmic rays. A systematic shift in $\theta_{e\gamma}$ away from zero is found, and the data is corrected for this. The uncertainty in this measurement contributes to the uncertainty in the mean of both the $\phi_{e\gamma}$ and $\theta_{e\gamma}$ signal PDFs. θ_e has an additional uncertainty due to the precision of the DCH alignment and knowledge of the magnetic field, which is estimated using alternate alignments and magnetic fields, and inserted into the total $\theta_{e\gamma}$ signal PDF mean uncertainty. The uncertainty in the position of the target plane, estimated from the optical survey precision, results in an uncertainty in ϕ_e that is incorporated into the uncertainty of the $\phi_{e\gamma}$ signal PDF mean. Systematic uncertainties in the full RMS of both relative angle PDFs are estimated by calculating the change in the full RMS due to changes in each of the component PDFs used to build them and adding the results in quadrature. This receives contributions from uncertainties in the following:

- The photon first conversion position resolution. The uncertainty in the full RMS of the photon position PDFs are estimated to be $\sigma_{u_{RMS}} = 0.3$ mm, $\sigma_{v_{RMS}} = 0.3$ mm, and

$$\sigma_{w_{RMS}} = 0.7 \text{ mm.}$$

- The size of the positron angle/vertex position correlations. The slopes of $\frac{dZ}{d\theta_e}$ and $\frac{dY}{d\phi_e}$ measured in section 8.2.2 are compared with the results of signal Monte Carlo yielding the slope uncertainties: $\sigma_{dZ/d\theta_e} = 17\%$ and $\sigma_{dY/d\phi_e} = 61\%$.
- The width of the positron angle PDFs. The uncertainty in the full RMS of each is estimated by the same method as for the signal E_e PDF. They too give a coherent component (θ_e, ϕ_e PDF widths overestimated by 20% and 25% respectively) and a stochastic component ($\sigma_{\phi_e RMS} = 9.2\%$ and $\sigma_{\theta_e RMS} = 3.3\%$). The coherent parts are converted into a rescaling factor to generate using accordingly narrower relative angle PDFs than the fit PDFs.

The mean and full RMS of each relative angle fit PDF is fluctuated from experiment to experiment in the confidence level calculation. The uncertainties of the slopes in the remaining correlations are found to be $\sigma_{d\phi_{e\gamma}/dE_e} = 35\%$ by the same technique used for the angle/vertex correlations and $\sigma_{d\phi_{e\gamma}/d\theta_{e\gamma}} = 8.1\%$ by the same technique used in calculating the total uncertainty in the angle PDF widths from all contributions. These are also fluctuated in the fit PDFs.

These are all summarized in table 9.1.

Lastly, the 10% uncertainty in the normalization factor is included by fluctuating the number of generated signal events from experiment to experiment.

9.6.2 Radiative Decay PDFs

The radiative decay $t_{e\gamma}$ PDF is fluctuated when fitting simulated experiments in the same way as the signal $t_{e\gamma}$ PDF with the same uncertainties. This is not a feasible approach for

Table 9.1: A summary of the systematic uncertainties in the signal PDFs.

Observable	Uncertainty in fit PDF peak	Uncertainty in fit PDF width	Scale factor on generation PDF width
$t_{e\gamma}$	10 ps	7%	1
E_e	40 keV	10%	0.79
E_γ	0.31%	15% (σ_{up})	1
$\theta_{e\gamma}$	4 mrad	3.6%	0.87
$\phi_{e\gamma}$	4 mrad	9.5%	0.89

the kinematic radiative decay PDF because of the intense computation required to perform the convolution. Since the number of radiative decay events is fixed in the likelihood fit, uncertainties in the shape of the kinematic radiative decay PDF are equivalent to uncertainties in the number of radiative decays in the signal window. As discussed in section 9.3.2, the uncertainty in the number of radiative decays is dominated by statistics anyway. This uncertainty is integrated into the confidence level calculation by generating a true number of radiative decay events, in each simulated experiment, according to a Gaussian probability distribution centered at the expectation (the same number that N_{RD} is fixed to in the fit), and whose σ is the uncertainty.

9.6.3 Accidental Background PDFs

As discussed in section 9.2.2, the accidental $t_{e\gamma}$ PDF is consistent with being flat. The statistical error on the slope of a line from the fit mentioned in that section is used to represent the systematic uncertainty of the flatness. This is incorporated in the confidence

level calculation by introducing a non-zero slope in the accidental $t_{e\gamma}$ fit PDF that fluctuates from experiment to experiment according to its uncertainty.

The accidental $\phi_{e\gamma}, \theta_{e\gamma}, E_e$ and E_γ PDFs are each extracted by an interpolation in some parameter of distributions measured directly on data, and do not have an explicit functional form. To estimate the uncertainty in these PDFs, each is re-measured in the way described in section 9.2.2 on an alternate set of sideband events. In the confidence level calculation, the fit PDFs are a properly normalized superposition of the nominal and alternate distributions, where the coefficients are fluctuated from experiment to experiment so as to give equal weight to the nominal and alternate PDF on average.

9.7 Blind Sensitivity Estimation

An estimate of the upper limit at 90% CL that the analysis of the signal window is expected to obtain in the case that there are no true signal events is performed in a few different ways while still remaining blind.

9.7.1 Fitting the Sidebands

Since the left and right sidebands are located $> 2.7\sigma$ away from the peak of the signal and radiative decay $t_{e\gamma}$ PDF, they are presumed to be free of signal and radiative decay events. A fictitious $\mu^+ \rightarrow e^+\gamma$ search is carried out in the left and right sidebands by shifting the signal and radiative decay $t_{e\gamma}$ Gaussian PDF means away from zero by -1.4 ns and +1.4 ns respectively. The number of radiative decays is fixed to zero in each case. Table 9.2 summarizes the resulting likelihood fits and calculated upper limits.

Table 9.2: Results of the fictitious signal searches in the sidebands. The asymmetric errors on the best fit parameters are the 1.645σ uncertainties calculated by MINOS.

Fit Region	Number of Observed Events	Best Fit (N_{Sig}, N_{Acc})	N_{Sig} UL (90% CL)	$BR(\mu^+ \rightarrow e^+\gamma)$ UL (90% CL)
left sideband	262	$(0.00_{-0.00}^{+4.77}, 262_{-25.7}^{+27.5})$	5.3	5.1×10^{-12}
right sideband	238	$(0.00_{-0.00}^{+2.88}, 238_{-24.5}^{+26.3})$	3.0	2.9×10^{-12}

9.7.2 Simulating Experiments

Another estimate of the expected sensitivity is based on a pure simulation. An ensemble of experiments is simulated assuming no signal and setting the true number of accidentals and radiative decays to their signal window expectations from section 9.3. In each experiment, the event numbers are fluctuated according to Poisson uncertainties. For each simulated event in a simulated experiment, a set of event properties, \vec{y} , is drawn randomly from actual events in the left and right sidebands. These are used to build the necessary PDFs and the kinematic observables, \vec{m} , are generated according to them. A likelihood fit is done on each experiment, fixing N_{RD} to the expectation in section 9.3.2, and a 90% CL interval is computed. The resulting histogram of upper limits, shown in figure 9.13, can be interpreted as a probability distribution of upper limits to be expected from a null experiment due to the statistical nature of the PDFs. The median of that distribution is $N_{Sig} = 5.1$ (90% CL UL), which corresponds to $BR(\mu^+ \rightarrow e^+\gamma) < 4.9 \times 10^{-12}$ (90% CL). The distribution is also consistent with the upper limits obtained on the sidebands in section 9.7.1.

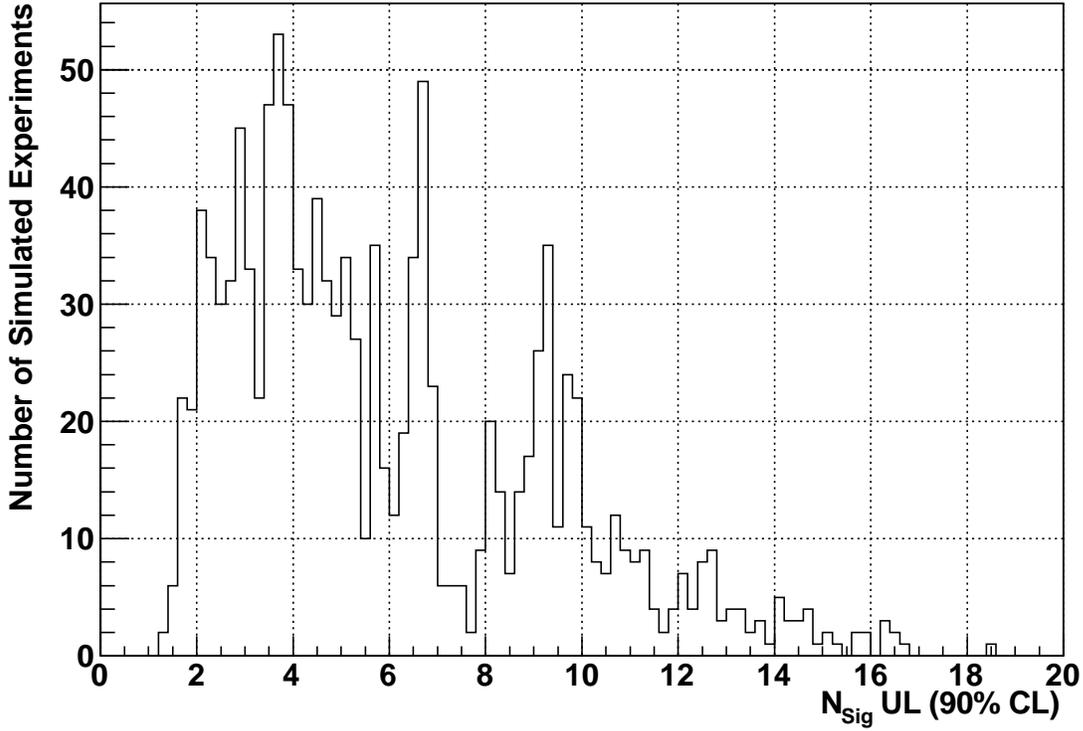


Figure 9.13: Simulated distribution of upper limits.

The same technique is used to investigate the discovery potential of the experiment. The simulation is performed for a few different non-zero numbers of true signal events. In each case, the median lower limit is calculated for various confidence levels. The results are reported in table 9.3.

9.8 Results of the Fit to the Signal Window

The likelihood fit to the signal window is done with the constraint that $N_{RD} = 29.42$, as discussed in section 9.3.2. In the signal window, 790 events are observed and the parameters that maximize the likelihood function, together with their 1.645σ uncertainties from MINOS,

Table 9.3: Median lower limits at various confidence levels obtained for different amounts of simulated signal.

True N_{Sig}	True $BR(\mu^+ \rightarrow e^+\gamma)$	$BR(\mu^+ \rightarrow e^+\gamma)$ LL (90% CL)	$BR(\mu^+ \rightarrow e^+\gamma)$ LL (95% CL)	$BR(\mu^+ \rightarrow e^+\gamma)$ LL (99% CL)
5	4.9×10^{-12}	1.1×10^{-12}	5.7×10^{-13}	2.3×10^{-14}
10	9.7×10^{-12}	3.6×10^{-12}	3.1×10^{-12}	1.3×10^{-12}
15	1.5×10^{-11}	7.1×10^{-12}	6.1×10^{-12}	4.2×10^{-12}

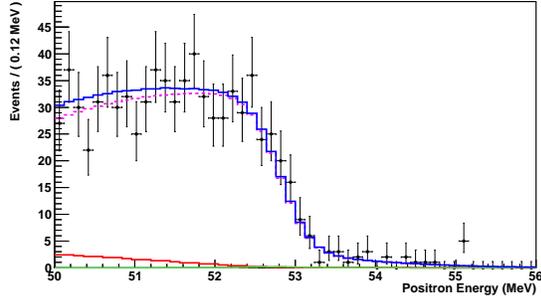
are:

$$(N_{Sig}, N_{Acc}) = (1.50^{+5.82}_{-2.61}, 764.5^{+47.1}_{-45.3}). \quad (9.20)$$

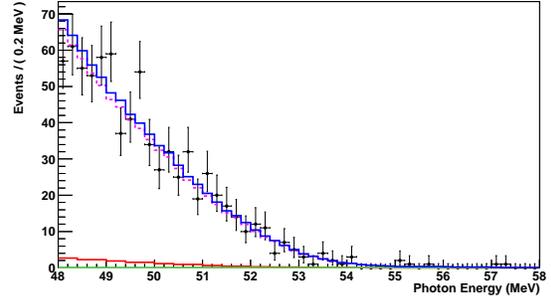
Figure 9.14 shows each of the PDFs projected onto each observable.

The confidence level at $N_{Sig} = 0$ is 78%, so only an upper limit is reported. The upper limit is $N_{Sig} = 8.1$ (90% CL), which corresponds to an upper limit on the branching fraction of $BR(\mu^+ \rightarrow e^+\gamma) < 7.9 \times 10^{-12}$ (90% CL). The probability of a null experiment to yield this upper limit or greater is estimated to be 28% from the simulated distribution of upper limits in figure 9.13.

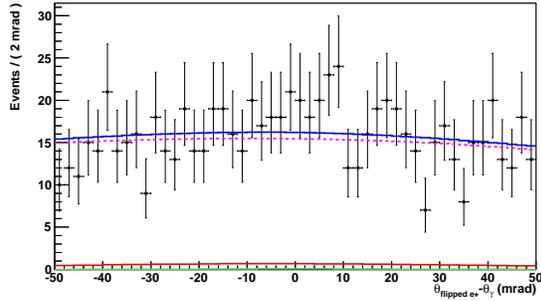
The effects of systematic uncertainties in the signal PDF shapes on the fit to the signal region are investigated by varying the parameters in increments of $\frac{\sigma}{2}$, refitting, and observing the change in the best fit value of N_{Sig} . Figures 9.15-9.16 show the results for the various components of the signal PDF shapes. The largest effects come from the uncertainties in the $\phi_{e\gamma}$ PDF center and the $t_{e\gamma}$ resolution, but the total uncertainty on N_{Sig} from all signal PDF parameter systematic uncertainties is 0.63, which is small compared to the statistical uncertainty of ~ 3 .



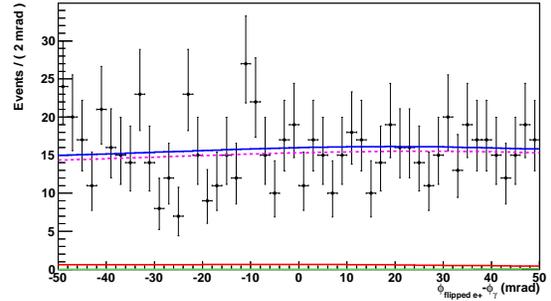
(a) Likelihood fit projected onto E_e .



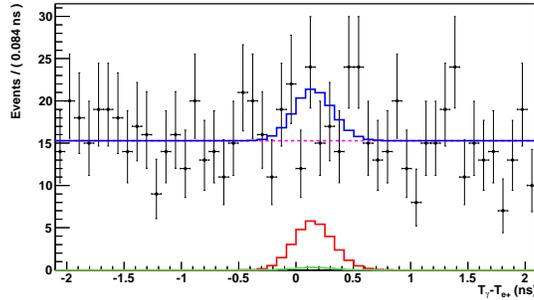
(b) Likelihood fit projected onto E_γ .



(c) Likelihood fit projected onto $\theta_{e\gamma}$.

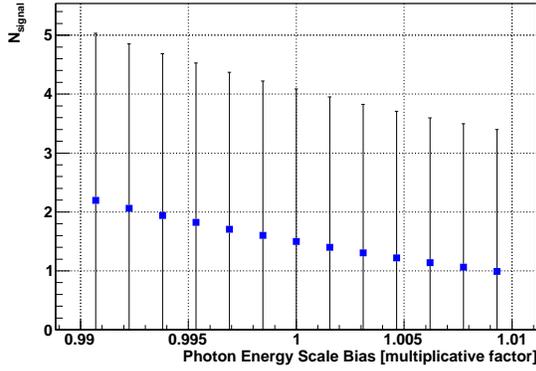


(d) Likelihood fit projected onto $\phi_{e\gamma}$.

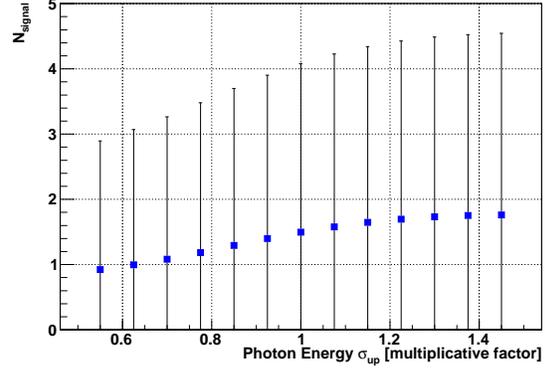


(e) Likelihood fit projected onto $t_{e\gamma}$.

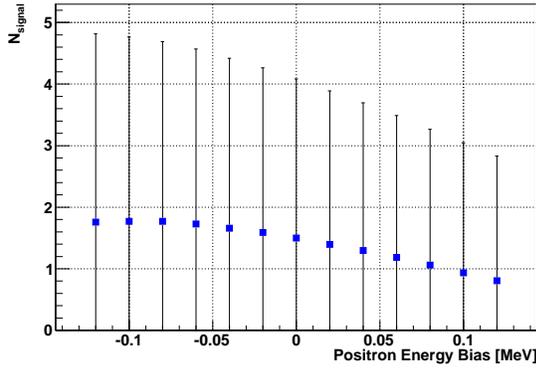
Figure 9.14: Results of the likelihood fit projected onto each observable by integrating out the others. The green curve is the signal PDF, the red is the radiative decay PDF, the purple is the accidental PDF, and the blue is the sum of those three. The area under each curve is normalized to its best fit result.



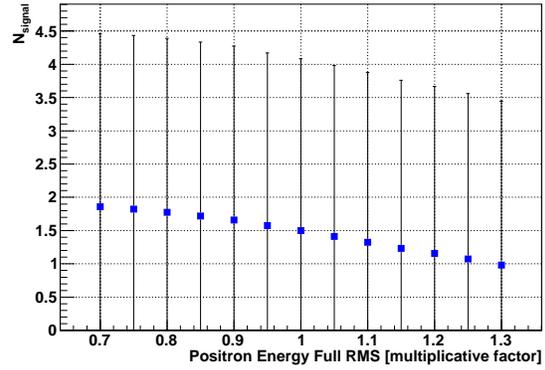
(a) Effects of varying the photon energy scale.



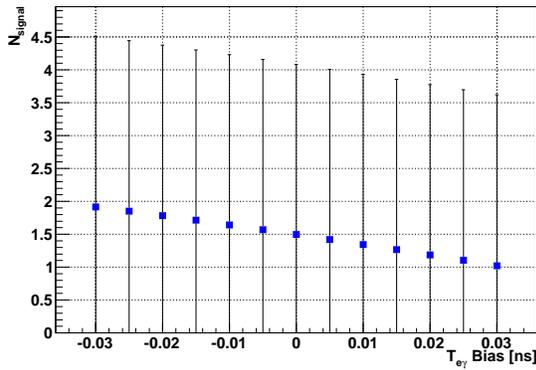
(b) Effects of varying the photon energy resolution (σ_{up}).



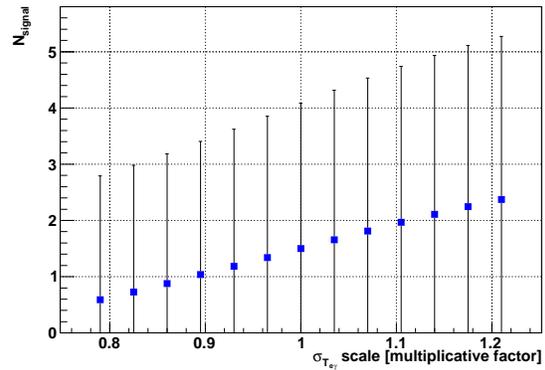
(c) Effects of varying the positron energy scale.



(d) Effects of varying the positron energy resolution (full RMS).

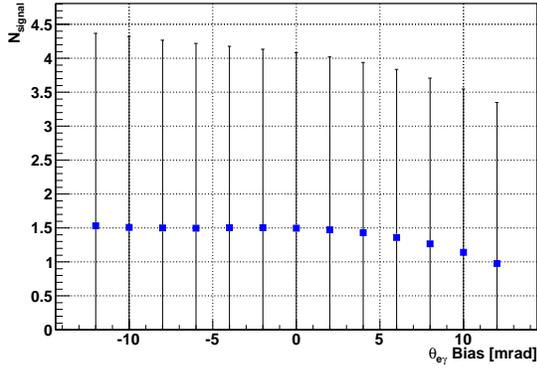


(e) Effects of varying the center of $t_{e\gamma}$.

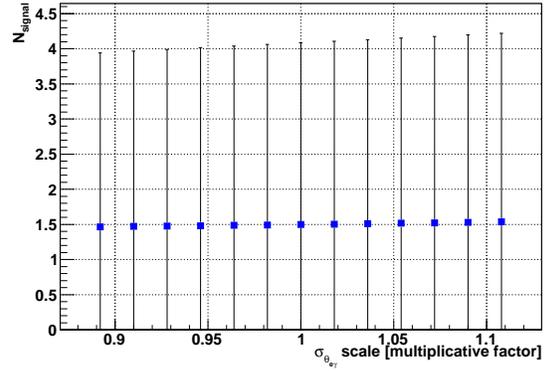


(f) Effects of varying the $t_{e\gamma}$ resolution.

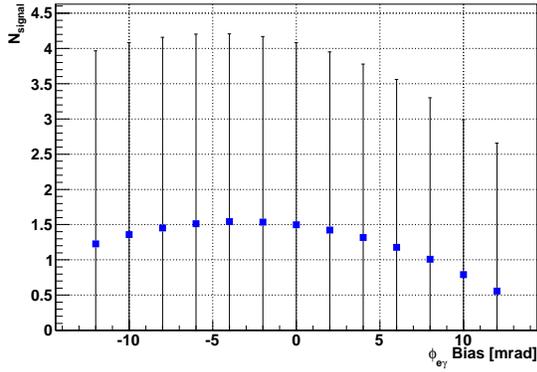
Figure 9.15: Plots of the best fit value of N_{Sig} as each energy or time PDF parameter is changed. The spacing between data points on the horizontal axis is $\frac{\sigma}{2}$, where σ is the quoted uncertainty in the parameter.



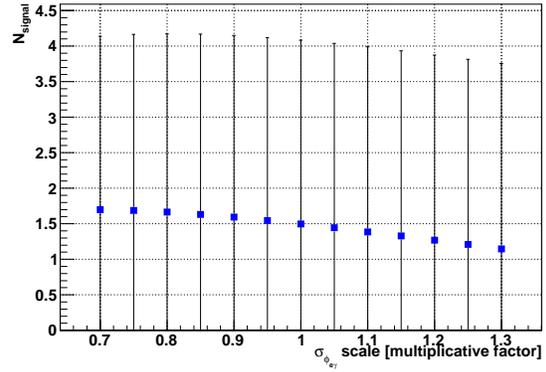
(a) Effects of varying the center of $\theta_{e\gamma}$.



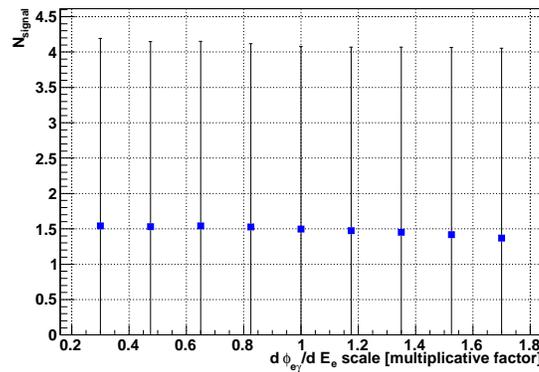
(b) Effects of varying the $\theta_{e\gamma}$ resolution (full RMS).



(c) Effects of varying the center of $\phi_{e\gamma}$.



(d) Effects of varying the $\phi_{e\gamma}$ resolution (full RMS).



(e) Effects of varying the slope of $\frac{d\phi_{e\gamma}}{dE_e}$.

Figure 9.16: Plots of the best fit value of N_{Sig} as each angle PDF parameter is changed. The spacing between data points on the horizontal axis is $\frac{\sigma}{2}$, where σ is the quoted uncertainty in the parameter.

Chapter 10

Conclusion

During 2009, the MEG experiment completed its second physics run to search for $\mu^+ \rightarrow e^+\gamma$, acquiring 43 days of data. Positrons were detected in a magnetic spectrometer with a graded magnetic field. Photons were detected in a 900 liter liquid xenon calorimeter.

We performed a blind analysis that extracted the number of signal events from a maximum likelihood fit to the data. A confidence interval was evaluated using the Feldman-Cousins prescription. The simulated sensitivity of the experiment, in case of no signal, was found to be

$$BR(\mu^+ \rightarrow e^+\gamma)_{expected} < 4.9 \times 10^{-12} \quad (90\% \text{ CL}). \quad (10.1)$$

The results of the $\mu^+ \rightarrow e^+\gamma$ search were a best fit number of signal events of 1.50, a 90% CL interval that included the null hypothesis, and an upper limit of:

$$BR(\mu^+ \rightarrow e^+\gamma) < 7.9 \times 10^{-12} \quad (90\% \text{ CL}). \quad (10.2)$$

The current best limit remains $BR(\mu^+ \rightarrow e^+\gamma) < 2.4 \times 10^{-12}$ (90% *CL*) from the combined 2009 and 2010 data sets; however, the result of this thesis lowers the previous limit obtained from the 2009 data set alone of $BR(\mu^+ \rightarrow e^+\gamma) < 9.6 \times 10^{-12}$ (90% *CL*)[2] by a factor of 1.2.

Bibliography

- [1] M. L. Brooks *et al.*, “New Limit for the Family-Number Non-conserving Decay $\mu^+ \rightarrow e^+\gamma$,” *Phys. Rev. Lett.* **83** (1999) 1521.
- [2] **MEG** Collaboration, J. Adam *et al.*, “New limit on the lepton-flavour violating decay $\mu^+ \rightarrow e^+\gamma$,” *Phys. Rev. Lett.* **107** (2011) 171801, [arXiv:1107.5547 \[hep-ex\]](#).
- [3] J. L. Hewett, “The Standard model and why we believe it,” [arXiv:hep-ph/9810316](#).
- [4] S. Coleman and J. Mandula, “All possible symmetries of the s matrix,” *Phys. Rev.* **159** no. 5, (Jul, 1967) 1251–1256.
- [5] R. Haag, J. T. Lopuszanski, and M. Sohnius, “All Possible Generators of Supersymmetries of the s Matrix,” *Nucl. Phys.* **B88** (1975) 257.
- [6] N. Polonsky, “Supersymmetry: Structure and phenomena. Extensions of the standard model,” *Lect. Notes Phys.* **M68** (2001) 1–169, [arXiv:hep-ph/0108236](#).
- [7] M. Drees, “An Introduction to supersymmetry,” [arXiv:hep-ph/9611409](#).
- [8] S. P. Martin, “A Supersymmetry Primer,” [arXiv:hep-ph/9709356](#).
- [9] M. C. Gonzalez-Garcia and Y. Nir, “Developments in neutrino physics,” *Rev. Mod. Phys.* **75** (2003) 345–402, [arXiv:hep-ph/0202058](#).
- [10] M. Maltoni and T. Schwetz, “Three-flavour neutrino oscillation update and comments on possible hints for a non-zero θ_{13} ,” *PoS IDM2008* (2008) 072, [arXiv:0812.3161 \[hep-ph\]](#).
- [11] T. P. Cheng and L. F. Li, “ $\mu \rightarrow e\gamma$ in theories with Dirac and Majorana neutrino mass terms,” *Phys. Rev. Lett.* **45** (1980) 1908.
- [12] R. N. Mohapatra and A. Y. Smirnov, “Neutrino Mass and New Physics,” *Ann. Rev. Nucl. Part. Sci.* **56** (2006) 569–628, [arXiv:hep-ph/0603118](#).
- [13] J. I. Illana and M. Masip, “Neutrino mixing and lepton flavor violation in SUSY-GUT models,” *Acta Phys. Polon.* **B34** (2003) 5413–5422, [arXiv:hep-ph/0310257](#).
- [14] M. Dine, Y. Grossman, and S. D. Thomas, “Slepton flavor physics at linear colliders,” *eConf C010630* (2001) P332, [arXiv:hep-ph/0111154](#).

- [15] K. Tobe, “Lepton flavor violation in SUSY models with and without R- parity,” [arXiv:hep-ph/0008075](https://arxiv.org/abs/hep-ph/0008075).
- [16] H. Murayama and A. Pierce, “Not even decoupling can save minimal supersymmetric SU(5),” *Phys. Rev.* **D65** (2002) 055009, [arXiv:hep-ph/0108104](https://arxiv.org/abs/hep-ph/0108104).
- [17] L. Calibbi, A. Faccia, A. Masiero, and S. K. Vempati, “Lepton flavour violation from SUSY-GUTs: Where do we stand for MEG, PRISM / PRIME and a super flavour factory,” *Phys. Rev.* **D74** (2006) 116002, [arXiv:hep-ph/0605139](https://arxiv.org/abs/hep-ph/0605139).
- [18] D. W. Maybury and B. Murakami, “Differentiating solutions to the gauge hierarchy problem through rare muon decays,” [arXiv:hep-ph/0401170](https://arxiv.org/abs/hep-ph/0401170).
- [19] Y. Kuno and Y. Okada, “Muon Decay and Physics Beyond the Standard Model,” *Rev. Mod. Phys.* **73** (2001) 151.
- [20] S. H. Neddermeyer and C. D. Anderson, “Note on the Nature of Cosmic Ray Particles,” *Phys. Rev.* **51** (1937) 884–886.
- [21] H. Yukawa, “On the Interaction of Elementary Particles,” *Proc. Phys. Math. Soc. Jap.* **17** (1935) 48–57.
- [22] M. Conversi, E. Pancini, and O. Piccioni, “On the Disintegration of Negative Mesons,” *Phys. Rev.* **71** (1947) 209–210.
- [23] C. M. G. Lattes *et al.* *Nature* **159** (1947) 453,486.
- [24] E. P. Hincks and B. Pontecorve, “Search for Gamma-Radiation in the 2.2-Microsecond Meson Decay Process,” *Phys. Rev.* **73** (1947) 257–258.
- [25] S. Lokanathan and J. Steinberger, “Search for Improbable Meson Decays,” *Phys. Rev.* **98** (1955) 240.
- [26] T. W. O’Keefe, M. Rigby, and J. R. Wormald, “Search for the process $\mu^+ \rightarrow e^+ + \gamma$,” *Proc. Phys. Soc.* **73** (1959) 951.
- [27] J. Ashkin, T. Fazzini, G. Fidecaro, N. H. Lipman, and A. W. Merrison, “Search for the Decay $\mu^+ \rightarrow e^+ + \gamma$ and Observation of the Decay $\mu^+ \rightarrow e^+ + \nu + \bar{\nu} + \gamma$,” *Nuovo Cim.* **14** (1959) 1266.
- [28] S. Frankel, V. Hagopian, J. Halpern, and A. L. Whetstone, “Further search for the decay $\mu^+ \rightarrow e^+ + \gamma$,” *Phys. Rev.* **118** (Apr, 1960) 589–590. <http://link.aps.org/doi/10.1103/PhysRev.118.589>.
- [29] D. Bartlett, S. Devons, and A. M. Sachs, “Search for the decay mode: $\mu \rightarrow e + \gamma$,” *Phys. Rev. Lett.* **8** (1962) 120–123.
- [30] S. Frankel *et al.*, “New Limit on the $e^+ + \gamma$ Decay Mode of the Muon,” *Phys. Rev. Lett.* **8** (1962) 123–125.

- [31] S. Parker, H. L. Anderson, and C. Rey, “Search for the decay $\mu^+ \rightarrow e^+ + \gamma$,” *Phys. Rev.* **133** (Feb, 1964) B768–B778.
<http://link.aps.org/doi/10.1103/PhysRev.133.B768>.
- [32] G. Feinberg, “Decays of the Mu Meson in the Intermediate-Meson Theory,” *Phys. Rev.* **110** (1958) 1482.
- [33] G. Danby, J. M. Gaillard, K. A. Goulianos, L. M. Lederman, N. B. Mistry, M. Schwartz, and J. Steinberger, “Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos,” *Phys. Rev. Lett.* **9** (1962) 36.
- [34] H. Povel *et al.*, “A New Upper Limit for the Decay $\mu^+ \rightarrow e^+ \gamma$,” *Phys. Lett. B* **72** (1977) 183.
- [35] A. van der Schaaf *et al.*, “A SEARCH FOR THE DECAY $\mu^+ \rightarrow e^+ \gamma$,” *Nucl. Phys.* **A340** (1980) 249.
- [36] W. W. Kinnison *et al.*, “A SEARCH FOR $\mu^+ \rightarrow e^+ \gamma$,” *Phys. Rev.* **D25** (1982) 2846.
- [37] R. D. Bolton *et al.*, “Search for Rare Muon Decays with the Crystal Box Detector,” *Phys. Rev.* **D38** (1988) 2077.
- [38] **SINDRUM** Collaboration, U. Bellgardt *et al.*, “Search for the Decay $\mu^+ \rightarrow e^+ e^-$,” *Nucl. Phys.* **B299** (1988) 1.
- [39] S. J. Freedman *et al.*, “Limits on neutrino oscillations from anti-electron- neutrino appearance,” *Phys. Rev.* **D47** (1993) 811–829.
- [40] **SINDRUM II**. Collaboration, C. Dohmen *et al.*, “Test of lepton flavor conservation in $\mu \rightarrow e$ conversion on titanium,” *Phys. Lett.* **B317** (1993) 631–636.
- [41] F. Xiao, *Measurement of the Radiative Muon Decay Branching Fraction in the MEG Experiment*. PhD thesis, University of California, Irvine, 2010.
- [42] T. Mori *et al.*, “Search for $\mu^+ \rightarrow e^+ \gamma$ Down to 10^{-14} Branching Ratio, Research Proposal to Paul Scherrer Institut, R-99-05,” 1999.
- [43] **MEGA** Collaboration, M. Ahmed *et al.*, “Search for the lepton-family-number nonconserving decay $\mu^+ \rightarrow e^+ \gamma$,” *Phys. Rev.* **D65** (2002) 112002,
[arXiv:hep-ex/0111030](https://arxiv.org/abs/hep-ex/0111030).
- [44] Evonik Industries GmbH, “ROHACELL®.”
<http://www.rohacell.com/product/rohacell/en/Pages/default.aspx>.
- [45] A. Sharma, “Properties of some gas mixtures used in tracking detectors,”
 SLAC-J-ICFA-16-3.
- [46] R. Veenhof, “Garfield - simulation of gaseous detectors.”
<http://garfield.web.cern.ch/garfield/>.

- [47] “Saint-Gobain Ceramics & Plastics.” <http://www.detectors.saint-gobain.com/>.
- [48] “Hamamatsu Photonics.” <http://www.hamamatsu.com/>.
- [49] A. Hitachi, T. Takahashi, N. Funayama, K. Masuda, J. Kikuchi, and T. Doke, “Effect of ionization density on the time dependence of luminescence from liquid argon and xenon,” *Phys. Rev. B* **27** no. 9, (May, 1983) 5279–5285.
- [50] J. Jortner, L. Mever, S. A. Rice, and E. G. Wilson, “Localized excitations in condensed Ne, Ar, Kr, and Xe,” *J. Chem. Phys.* **42** (1965) 4250–4253.
- [51] MIDAS, “Maximum Integration Data Acquisition System.” <https://midas.psi.ch/>.
- [52] S. Ritt, “Design and Performance of the 6 GHz Waveform Digitizing Chip DRS4,” in *IEEE Nuclear Science Symposium Conference Record*. 2008.
- [53] S. Ritt, “Design and Performance of the 5 GHz Waveform Digitizing Chip DRS3,” in *IEEE Nuclear Science Symposium Conference Record*. 2007.
- [54] A. Devices, “AD9218, 10 bit 3v Dual A/D converter.” <http://www.analog.com/en/index.html>.
- [55] Xilinx, “Virtex II Pro.” <http://www.xilinx.com/>.
- [56] SaRonix, “Crystal Clock Oscillator.” <http://www.pericom.com/saronix/>.
- [57] “High Voltage Engineering Europe.” <http://www.highvolteng.com/>.
- [58] NACRE, “European Compilation of Reaction Rates for Astrophysics.” <http://pntpm.ulb.ac.be/Nacre/nacre.htm>.
- [59] R. E. Kalman, “A new approach to linear filtering and prediction problems,” *Transactions of the ASME—Journal of Basic Engineering* **82** no. Series D, (1960) 35–45.
- [60] V. Blobel, C. Kleinwort, and F. Meier, “Fast alignment of a complex tracking detector using advanced track models,” *Comput. Phys. Commun.* **182** (2011) 1760–1763, [arXiv:1103.3909](https://arxiv.org/abs/1103.3909) [physics.ins-det].
- [61] B. Golden and W. Molzon, “Contributions to Z resolution from All Effects.” MEG Technical Note 59, Nov., 2009.
- [62] J. R. Klein and A. Roodman, “Blind analysis in nuclear and particle physics,” *Ann. Rev. Nucl. Part. Sci.* **55** (2005) 141–163.
- [63] R. Barlow, “Extended maximum likelihood,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **297** no. 3, (1990) 496 – 506. <http://www.sciencedirect.com/science/article/pii/0168900290913348>.

- [64] A. L. Read, “Linear interpolation of histograms,” *Nucl. Instrum. Meth.* **A425** (1999) 357–360.
- [65] “Minuit.” <http://root.cern.ch/root/html/TMinuit.html>.
- [66] G. J. Feldman and R. D. Cousins, “A Unified Approach to the Classical Statistical Analysis of Small Signals,” *Phys. Rev.* **D57** (1998) 3873–3889, [arXiv:physics/9711021](https://arxiv.org/abs/physics/9711021).
- [67] W. A. Rolke, A. M. Lopez, and J. Conrad, “Limits and confidence intervals in the presence of nuisance parameters,” *Nucl. Instrum. Meth.* **A551** (2005) 493–503, [arXiv:physics/0403059](https://arxiv.org/abs/physics/0403059) [physics].
- [68] C. Fronsdal and H. Uberall, “mu-Meson Decay with Inner Bremsstrahlung,” *Phys. Rev.* **113** (1959) 654–657.

Appendices

A Radiative Muon Decay Branching Ratio

The differential branching ratio, calculated at tree level [68, 19], is given by:

$$dB(\mu^\pm \rightarrow e^\pm \gamma \bar{\nu} \nu) = \frac{\alpha}{64\pi^3} \beta dx \frac{dy}{y} d\Omega_e d\Omega_\gamma \times \quad (\text{A.3})$$

$$\left[F(x, y, d) \mp \beta \vec{P}_\mu \cdot \hat{p}_e G(x, y, d) \mp \vec{P}_\mu \cdot \hat{p}_\gamma H(x, y, d) \right].$$

\vec{P}_μ is the muon polarization vector; \vec{p}_e and \vec{p}_γ are the positron and photon momenta in the muon rest frame; \hat{p}_e and \hat{p}_γ are the corresponding unit vectors; $\beta \equiv \frac{|\vec{p}_e|}{E_e}$; $d \equiv 1 - \beta \hat{p}_e \cdot \hat{p}_\gamma$; $x \equiv \frac{2E_e}{m_\mu}$; and $y \equiv \frac{2E_\gamma}{m_\mu}$.

Furthermore,

$$F = F^{(0)} + \left(\frac{m_e}{m_\mu}\right)^2 F^{(1)} + \left(\frac{m_e}{m_\mu}\right)^4 F^{(2)} \quad (\text{A.4})$$

$$G = G^{(0)} + \left(\frac{m_e}{m_\mu}\right)^2 G^{(1)} + \left(\frac{m_e}{m_\mu}\right)^4 G^{(2)} \quad (\text{A.5})$$

$$H = H^{(0)} + \left(\frac{m_e}{m_\mu}\right)^2 H^{(1)} + \left(\frac{m_e}{m_\mu}\right)^4 H^{(2)}, \quad (\text{A.6})$$

where

$$F^{(0)}(x, y, d) = \frac{8}{d} \{ y^2 (3 - 2y) + 6xy (1 - y) + 2x^2 (3 - 4y) - 4x^3 \} \quad (\text{A.7})$$

$$+ 8 \{ -xy (3 - y - y^2) - x^2 (3 - y - 4y^2) + 2x^3 (1 + 2y) \}$$

$$+ 2d \{ x^2 y (6 - 5y - 2y^2) - 2x^3 y (4 + 3y) \} + 2d^2 x^3 y^2 (2 + y)$$

$$\begin{aligned}
F^{(1)}(x, y, d) &= \frac{32}{d^2} \left\{ -\frac{y(3-2y)}{x} - (3-4y) + 2x \right\} \\
&\quad + \frac{8}{d} \{y(6-5y) - 2x(4+y) + 6x^2\} \\
&\quad + 8 \{x(4-3y+y^2) - 3x^2(1+y)\} + 6dx^2y(2+y)
\end{aligned} \tag{A.8}$$

$$F^{(2)}(x, y, d) = \frac{32}{d^2} \left\{ \frac{(4-3y)}{x} - 3 \right\} + \frac{48y}{d} \tag{A.9}$$

$$\begin{aligned}
G^{(0)}(x, y, d) &= \frac{8}{d} \{xy(1-2y) + 2x^2(1-3y) - 4x^3\} \\
&\quad + 4 \{-x^2(2-3y-4y^2) + 2x^3(2+3y)\} \\
&\quad \quad \quad - 4dx^3y(2+y)
\end{aligned} \tag{A.10}$$

$$G^{(1)}(x, y, d) = \frac{32}{d^2} (-1 + 2y + 2x) + \frac{8}{d} (6x^2 - xy) - 12x^2(2+y) \tag{A.11}$$

$$G^{(2)}(x, y, d) = -\frac{96}{d^2} \tag{A.12}$$

$$\begin{aligned}
H^{(0)}(x, y, d) &= \frac{8}{d} \{y^2 (1 - 2y) + xy (1 - 4y) - 2x^2 y\} \\
&\quad + 4 \{2xy^2 (1 + y) - x^2 y (1 - 4y) + 2x^3 y\} \\
&\quad + 2d \{x^2 y^2 (1 - 2y) - 4x^3 y^2\} + 2d^2 x^3 y^3
\end{aligned} \tag{A.13}$$

$$\begin{aligned}
H^{(1)}(x, y, d) &= \frac{32}{d^2} \left\{ -\frac{y(1-2y)}{x} + 2y \right\} + \frac{8}{d} \{y(2-5y) - xy\} \\
&\quad + 4xy(2y-3x) + 6dx^2 y^2
\end{aligned} \tag{A.14}$$

$$H^{(2)}(x, y, d) = -\frac{96y}{d^2 x} + \frac{48y}{d}. \tag{A.15}$$