UNIVERSITY OF CALIFORNIA,
IRVINE

Measurement of the Radiative Muon Decay Branching Fraction in the MEG Experiment

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Physics

by

Feng Xiao

Dissertation Committee:
Professor William Molzon, Chair
Professor Henry Sobel
Professor David Kirkby

2010

DEDICATION

To my parents who have always supported me with their love and patience.

献给我善良慈爱的父母

梦还家

昨梦回故里，萱堂鬓催灰。

年来操持重，岁去所报微。

双慈尽思盼，游子迟不归。

养儿能防老？未曾奉一杯！

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

ACKNOWLEDGMENTS

Working on a PhD degree in experimental particle physics typically means spending a significant part of one's life on a project and the length of time and success of the project are often unpredictable. The work presented here would have not been possible without the help and support of many people to whom I owe many thanks.

I would like to thank my committee chair Prof. William Molzon, who encouraged optimal performance under self-supervision. With him I was trained to work independently and to handle unexpected, often stressful, situations.

I am very grateful to Prof. Henry Sobel, who despite being constantly busy himself spent time to improve part of my thesis.

I would like to express my gratitude to Ben Golden. We had many useful and thoughtful discussions, encouraged and supported each other and shared sympathies during difficult times, and created a lot of fun together to keep each other's sanity and to kill boring time which would have otherwise killed us.

I spent over half of my PhD research time working on-site at Paul Scherrer Institute, Switzerland. I have many thanks to various people there. Dr. Stefan Ritt, who was my on-site advisor, gave me insightful advices on the hardware project that I once worked on, and helped me a lot on administrative issues; Malte Hildebrandt was nice and soft, and had no hesitance in offering help whenever needed; Oleg Kiselev was my lunch time favorite and listened to my problems with a warm heart; Peter Raymond-Kettle, being the most senior person at PSI, had a special charm to assure a smooth experiment and steady progress and gave confidence to young physicists, like me; and Reinhard Schmidt helped me solder complicated wire connections of a gas control system.

Being a part of a multi-national collaboration, I had a great opportunity and honor to work with enthusiastic people from across the world. I would like to thank Ryu Sawada and Yusuke Uchiyama, the two of whom I bothered most with many, sometimes trivial, technical problems. They offered a lot of help on debugging and fixing my coding problems. I was also inspired by them from the physics and analysis discussions we had. I am very thankful for their patience and lenience and their insightful comments. Satoshi Mihara was always very understanding to junior scientists. He left his car to me and made my life in Switzerland more joyful and pleasant. Wataru Ootani became a UCI shift specialist who happened to have many shifts with me and made my overnight shifts smooth and easy. Toshiyuki Iwamoto had a special sense of humor. His remarks and comments will always be cherished. Shuei Yamada helped me a lot when I first started and struggled with the Unix operating system.

I would like to thank my Italian collaborator Giovanni Signorelli who gave me several small lectures on various topics of which I had little knowledge. He and Donato Nicolo once let me stay in their apartment for two weeks and saved me from being homeless. I have my gratitude to Gianluigi Boca, whose enjoyable humor made the tedious research work more colorful.

I appreciate my fellow students Ben Ziemer, Ben Golden, Kirk Bays, Felix Yu and Peter Porazik who proof read my thesis. They excused my unpolished English and poor grammar. Without their generous help, this work would not be in its current shape.

I am grateful to my funding agencies, the US Department of Energy (grant DE-FG03-91ER40679) and Paul Scherrer Institut, Switzerland for supporting my study, research and travel. I also owe thanks to Julie Aird and Michelle Arons, who assisted me with necessary documents to allow a foreign student to travel aboard to a third country.

最后，我要感谢我的父母。是他们在我身后的默默支持，才能使我完成这七年漫长的学业。这些年，虽然我不在他们的身边，但是他们给我的爱和信念，让我不言放弃、不断进取，帮我渡过各种意想不到的困难。我将我的博士论文，题献给他们。

CURRICULUM VITAE

Feng Xiao

EDUCATION

Doctor of Philosophy in Physics **2010**
University of California, Irvine *Irvine, California*

Master of Science in Physics **2005**
University of California, Irvine *Irvine, California*

Bachelor of Science in Physics **2003**
Zhejiang University *Hangzhou, China*

RESEARCH EXPERIENCE

Graduate Research Assistant **2005–2010**
University of California, Irvine *Irvine, California*

PhD Researcher **2007–2009**
Paul Scherrer Institute *Villigen, Switzerland*

TEACHING EXPERIENCE

Graduate Teaching Assistant **2003–2005**
University of California, Irvine *Irvine, California*

SELECTED HONORS AND AWARDS

UCI Regent Fellowship **2003–2004**
University of California, Irvine *Irvine, California*

Excellence Academic Scholarship **2000–2003**
Zhejiang University *Hangzhou, China*

PUBLICATIONS

A limit for the $\mu \rightarrow e\gamma$ decay from the MEG experiment **2010**
J. Adam et. al. [MEG Collaboration] Nucl. Phys. B, **834**, 1

ABSTRACT OF THE DISSERTATION

Measurement of the Radiative Muon Decay Branching Fraction
in the MEG Experiment

By

Feng Xiao

Doctor of Philosophy in Physics

University of California, Irvine, 2010

Professor William Molzon, Chair

A measurement of the radiative muon decay $\mu^+ \rightarrow e^+\nu_e\overline{\nu}_\mu\gamma$ branching fraction was done on the basis of the radiative decay data of MEG experiment. The data were taken periodically one day per week with the beam intensity of 1.2×10^6 μ^+/s from September to December, 2008. The positron was measured by a spectrometer with gradient magnetic field. The photon was detected by an innovative ∼900 liter liquid xenon scintillation detector.

The measurement was carried out with a cut-and-count approach. We report the radiative muon decay branching fraction to be

$$B(\mu \rightarrow e\nu\bar{\nu}\gamma) = (2.84 \pm 0.20(stat) \pm 0.05(syst)) \times 10^{-7},$$

with $E_e > 46$ MeV, $E_\gamma > 30$ MeV, and the opening angle between the positron and photon $\theta_{e\gamma}$ in its full kinematic range.

Distributions of three observables, E_e, E_γ, and $\theta_{e\gamma}$, as well as their kinematic correlations were verified to be consistent with the radiative decays.

The measurement is in excellent agreement with the prediction of the Standard Model.

Chapter 1

Introduction

Ever since the discovery of the muon, physicists have been puzzled by the existence of more than one generation of fundamental fermions and the interactions between the generations. The standard model (SM) of particle physics is so far the most comprehensive and precise theory describing the interactions of the elementary particles.

In the quark sector, the quark mixing in the flavor-changing weak decays are described by the Cabibbo-Kobayashi-Maskawa matrix. Recent discoveries of neutrino oscillations (e.g. by SNO, Super-Kamiokande, K2K, KamLAND, etc.) provide a direct evidence of the lepton flavor violation among the neutral leptons. Therefore, understanding the mixings among the charged leptons will bestow a new insight into the generation mechanism of the fundamental particles and their interactions.

In the SM, lepton flavor is conserved by the assumption of zero neutrino masses. Observation of neutrino oscillations corresponds to the fact that neutrinos have finite masses. Introduction of small neutrino masses by the seesaw mechanism through a Majorana mass term as an extension of the SM predicts the fraction of charged lepton flavor violation (cLFV) at the level that is tens of orders of magnitude lower than the capability of the experimental technique of the present day. Therefore, a search for cLFV is a sensitive probe to new physics beyond the SM.

The cLFV signal has been experimentally investigated in many processes, e.g. $\mu \to e\gamma$, $\tau \to \mu\gamma$, $\tau \to e\gamma$, $\mu \to eee$, $\mu N \to eN$, etc., among which $\mu \to e\gamma$ is the simplest and most famous one, whose current best limit on the branching fraction, $B(\mu \to e\gamma) = 1.2 \times 10^{-11}$, is set by the MEGA experiment[1, 2]. The $\mu \to e\gamma$ decay may likely occur with a branching fraction just below this limit, according to many of the promising theories for physics beyond the Standard Model, in particular, the Supersymmetric theories of Grand Unification or Supersymmetric Standard Model.

An endeavor that searches for $\mu \to e\gamma$ decay aiming at a sensitivity two orders of magnitude below the MEGA limit has been undertaken at the Paul Scherrer Institut (PSI), Switzerland since 1999 by the MEG (Muon to Electron and Gamma decay) collaboration, composed of some 60 physicists from Italian, Japanese, Russian, Swiss and American institutions. The MEG experiment uses the world's most intense DC muon beam to reach high muon luminosity, employs a novel positron spectrometer with a gradient magnetic field to work efficiently in a high rate environment, and exploits an innovative $\sim$900 liter liquid xenon photon detector to simultaneously measure photon energy, position, and time with fine resolutions. MEG started physics data taking in 2008 for three months.

This thesis focuses on the analysis of the radiative muon decay, $\mu \to e\bar{\nu}\nu\gamma$, which is the physics background of the MEG decay signal. With a high muon luminosity and the designed detector acceptance for high energy photons and positrons, we are able to measure the radiative decay branching fraction in an uncharted region.

In Chapter 2, we briefly overview the theoretical framework of muon decays. Chapter 3 describes the detectors and parts of the MEG experimental apparatus in details. A general description of the MEG software is given in Chapter 4. In Chapter 5, we summarize the run 2008 condition and various data types taken. Event reconstruction algorithms and detector calibrations and performance are presented from Chapter 6 to 8. Selection of events is addressed in Chapter 9 followed by the physical analysis of the radiative muon decay in Chapter 10. Finally, the thesis is concluded in Chapter 11.

Chapter 2

Overview

The study of muon physics has yielded great impact on the building blocks of the SM, in particular, the vector minus axial vector (V-A) framework of weak interaction and the validity of quantum electrodynamics. In this chapter, we briefly review the phenomenology of radiative muon decays.

The dominant decay mode of the muon, $\mu \to e\overline{\nu}\nu$, is often called a 'Michel decay'[3]. It has a branching fraction of nearly 100%. The muon also decays radiatively, $\mu \to e\overline{\nu}\nu\gamma$, or with an associated e^+e^- pair, $\mu \to e\overline{\nu}_e\nu_\mu e^+e^-$. The branching fractions for these decay modes and the upper limits of some lepton flavor violating decay modes at 90% confidence level are summarized in Table 2.1.

The $\mu \to e\gamma$ decay has been searched for for decades as a sensitive probe of physics beyond the SM. The MEG experiment conducted at PSI aims to yield a limit on $B(\mu \to e\gamma)$ of the order of $\sim 10^{-13}$. Experimental features as well as the history of the $\mu \to e\gamma$ search are summarized in Appendix A.

Decay mode	Branching ratio	Reference
$\mu^- \to \mathrm{e}^-\nu_\mu\overline{\nu}_\mathrm{e}$	~ 100 %	
$\mu^- \to \mathrm{e}^-\nu_\mu\overline{\nu}_\mathrm{e}\gamma$	(1.4 ± 0.4) %[1]	[4]
$\mu^- \to \mathrm{e}^-\nu_\mu\overline{\nu}_\mathrm{e}\mathrm{e}^+\mathrm{e}^-$	$(3.4 \pm 0.4) \times 10^{-5}$	[5]
$\mu^- \to \mathrm{e}^-\nu_\mathrm{e}\overline{\nu}_\mu$	< 1.2 %	[6]
$\mu^- \to \mathrm{e}^-\gamma$	$< 1.2 \times 10^{-11}$	[2]
$\mu^- \to \mathrm{e}^-\mathrm{e}^+\mathrm{e}^-$	$< 1.0 \times 10^{-12}$	[7]
$\mu^- \to \mathrm{e}^-\gamma\gamma$	$< 7.2 \times 10^{-11}$	[8]

Table 2.1: Decay modes and branching fractions of the muon (Listed in Particle Data Group table [9]).

---

[1] This only includes events with $E_\gamma > 10$ MeV. The radiative mode cannot be clearly separated from the Michel decay in the soft-photon limit. It can be regarded as a subset of the Michel decay mode.

2.1 Radiative Decay

At tree level in the SM Lagrangian, the muon interacts with the gauge bosons (photon, and $W^\pm$ and Z^0 bosons) and with the Higgs boson. The Lagrangian for those interactions are given,

$$\begin{aligned}
\mathcal{L} &= \mathrm{e}\bar{\mu}\gamma^\mu \mu A_\mu \\
&\quad -\frac{g}{\sqrt{2}}(\bar{\nu}_{\mu L}\gamma^\mu \mu_L W_L^+ + \bar{\mu}_L\gamma^\mu \nu_{\mu L} W_\mu^-) \\
&\quad -\sqrt{g^2+g'^2}\left[\bar{\mu}_L\gamma^\mu\left(-\frac{1}{2}+\sin^2\theta_W\right)\mu_L + \bar{\mu}_R\gamma^\mu \sin^2\theta_W \mu_R\right]Z_\mu^0 \\
&\quad -\frac{m_\mu}{v}\bar{\mu}\mu H, \qquad (2.1)
\end{aligned}$$

where g and g' are gauge coupling constants, and the Weinberg angle θ_W is defined by $\sin\theta_W \equiv g'/\sqrt{g^2+g'^2}$, and $e = g\sin\theta_W$ at tree level. Muon decay is described by a charged weak-current interaction mediated by $W^\pm$ gauge bosons, as seen in the second line in Equation 2.1.

In the SM, the muon decay is described by the V-A interaction. From late 1960s to 1980s, many experiments were carried out to measure the Michel decays precisely and all measurements, such as the electron energy spectrum, decay asymmetry and polarization vector, are found to be consistent with the V-A framework of weak interaction in the SM.

Theoretically, radiative decay is considered as an inner bremsstrahlung process in the Michel decay. Feynman diagrams for radiative decay are shown in Figure 2.1.

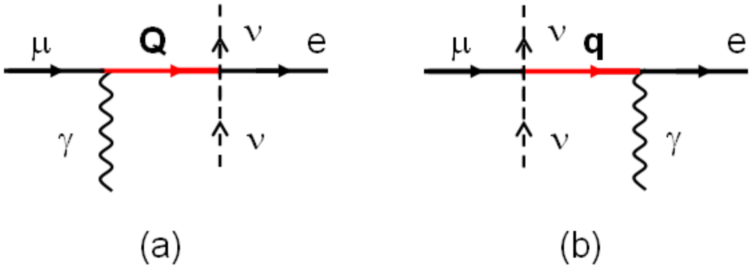


Figure 2.1: Feynman diagrams for radiative decay

The branching fraction as well as the electron and photon energy spectra of the radiative muon decay have been calculated by several authors[10, 11, 12, 13]. Within the framework of the V-A interaction, the differential branching fraction of the radiative muon decay in the muon rest frame is expressed by[14]

$$\begin{aligned} &dB(\mu^{\pm} \rightarrow e^{\pm}\nu\bar{\nu}\gamma) = \\ &\frac{\alpha}{64\pi^3}\beta dx \frac{dy}{y} d\Omega_e d\Omega_\gamma \cdot \left[F(x,y,d) \mp \beta \vec{P}_\mu \cdot \hat{p}_e G(x,y,d) \mp \vec{P}_\mu \cdot \hat{p}_\gamma H(x,y,d)\right], \end{aligned} \tag{2.2}$$

where the final positron (electron) and photon are emitted at energy intervals of dx and dy with solid angles of $d\Omega_e$ and $d\Omega_\gamma$, respectively, while x and y being normalized positron and photon energies, $x = 2E_e/m_\mu$ and $y = 2E_\gamma/m_\mu$; β is defined as

$\beta \equiv |\vec{p}_e|/E_e$; d is given by $d \equiv 1 - \beta\hat{p}_e \cdot \hat{p}_\gamma$; $\vec{P}_\mu$ is the muon polarization vector; $\vec{p}_e$ and $\vec{p}_\gamma$ are the momenta of the positron and photon in the muon rest frame, and $\hat{p}_e$ and $\hat{p}_\gamma$ are their unit vectors, respectively. $F(x, y, d)$, $G(x, y, d)$, and $H(x, y, d)$ in the SM are given in Appendix B. Figure 2.2 shows the differential branching fraction as a function of photon energy, $y = 2E_\gamma/m_\mu$.

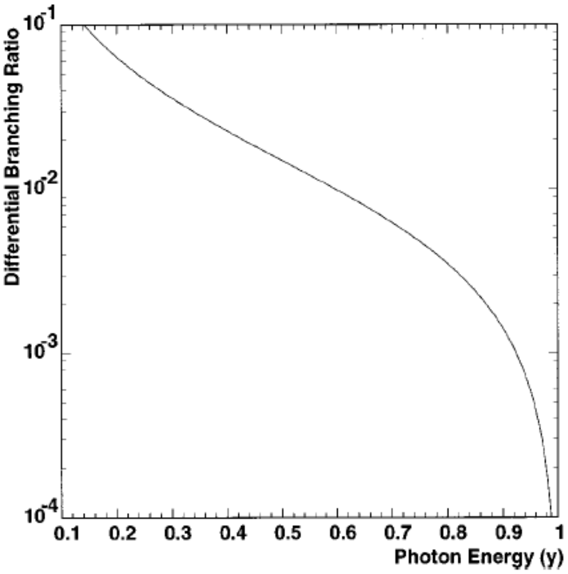


Figure 2.2: Differential branching fraction as a function of photon energy, $y = 2E_\gamma/m_\mu$. It is obtained by integrating over the e energy and the angle between the e and photon, $\theta_{e\gamma}$.

The measurement of the radiative decay branching fraction, (1.4 ± 0.4) % (for $E_\gamma >$ 10 MeV), listed in the Particle Data Group (PDG) table (Table 2.1) was done by using a 5-in freon bubble chamber in 1961. Recently a more precise measurements reported by the PIBETA experiment[15, 16] in 2006 gave

$$B(\mu \rightarrow e\nu\bar{\nu}\gamma) = [4.40 \pm 0.02(stat.) \pm 0.09(syst.)] \times 10^{-3}, \tag{2.3}$$

under the conditions that $E_\gamma > 10$ MeV and $\theta_{e\gamma} > 30°$. Both agree with the SM calculation; however, both measurements were limited by statistics for photons at high energy, which leaves $B(\mu \to e\nu\bar{\nu}\gamma)$ with relatively high positron and photon energies an uncharted region. But $B(\mu \to e\nu\bar{\nu}\gamma)$ with high positron and photon energies is technically difficult to measure and is very tiny, e.g. $B(\mu \to e\nu\bar{\nu}\gamma)|_{x>0.8,y>0.5} \sim 10^{-7}$, and thus requires high muon luminosity. Measuring $B(\mu \to e\nu\bar{\nu}\gamma)$ at relatively high energies with good precision gives another test of the V-A interaction.

2.2 Summary

The radiative muon decay, $\mu \to e\bar{\nu}\nu\gamma$, as a major background to the MEG experiment is worthy of being studied. The MEG experiment provides an opportunity to observe radiative muon decay with high positron and photon energies, e.g. $x > 0.8$ and $y > 0.5$, and measure its branching fraction. This is a measured test of the V-A theory of the weak interaction in a region that has not been previously charted. It, due to the smallness in value, requires high experimental sensitivity. It also needs high muon luminosity. The MEG experiment has both features. The design of the MEG, however, which is optimized for two-body decay of back-to-back products with monochromatic energies, introduces additional difficulties in the analysis of radiative decay and the precise measurement of its branching fraction.

Chapter 3

Experimental Apparatus

The MEG experiment is located at the Paul Scherrer Institute (PSI), Switzerland. It is designed to search for $\mu^+ \rightarrow e^+ + \gamma$ decay, whose signature is back-to-back positron and photon pair, coincident in time, both with energy of 52.8 MeV (half of the muon rest energy). In order to eliminate background events, precise measurements of the time, energy, and emission angle of the two particles are crucial to the success of the experiment. An overview of the MEG apparatus is shown in Figure 3.1. Before going into the details of each component, the global coordinate system used throughout the thesis will be defined. The origin (0, 0, 0) is placed at the center of the hollow COBRA[1] magnet, and is also ideally the center of the muon stopping target which lies inside of the COBRA magnet. The positive z direction is along the muon beam direction; the y-axis is the vertical axis pointing upward; and the x-axis is the third axis that, together with y- and z-axes, forms a right-handed coordinate system. The liquid xenon detector is located at negative x. In terms of a spherical description, θ is the polar angle from the positive z direction and ϕ is the azimuthal angle from the positive x direction.

---

[1] Acronym is defined in Section 3.2.1

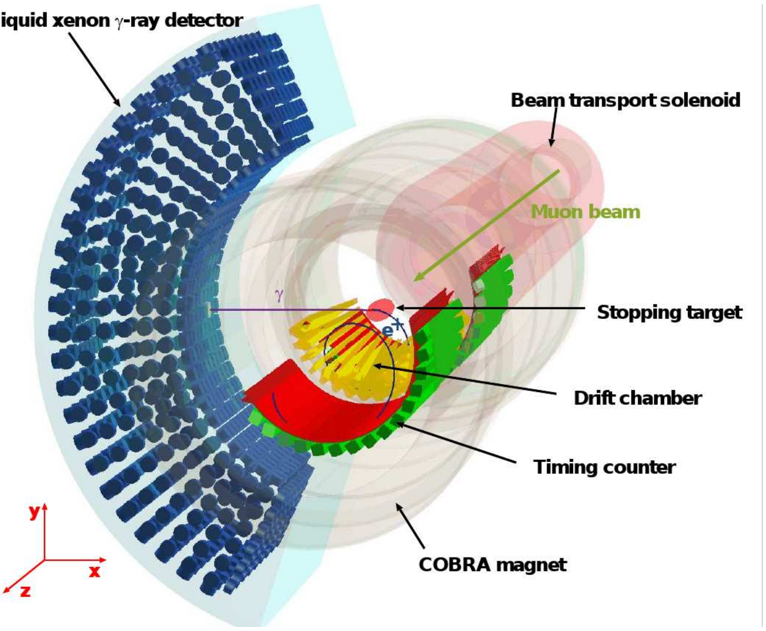


Figure 3.1: An overview of the MEG apparatus

The positive muon beam is transported to the apparatus by the beam transport solenoid (BTS). The muons are stopped on a thin target, then promptly decay. Any emitted photons enter the liquid xenon calorimeter (XEC), where their energy, position and time are all measured. Emitted positrons can be detected and measured by the positron spectrometer. Positron tracks are confined by the the superconducting solenoidal COBRA magnet, and tracked by the drift chamber (DCH) system. The DCH measures positron position and determines the positron trajectory, from which positron momentum and direction are further deduced. Positron timing is determined by an array of plastic scintillators serving as timing counters (TIC).

Detailed descriptions of the concept and design of each part and sub-detector in the MEG apparatus are covered in this chapter.

3.1 Beam and Target

The MEG experiment is conducted at the 590 MeV proton ring cyclotron facility at PSI, Switzerland, which is currently the world's most intense DC muon beam.

To achieve sufficient $\mu \rightarrow e\gamma$ sensitivity, a large number of muon decays is necessary. Therefore, a very intense muon beam is required. The accidental background rate increases with the instant beam intensity, so a DC beam is better than a pulsed beam for an optimal signal to noise ratio.

To avoid nuclear capture of muons by the stopping target, a positive muon beam is preferred than a negative one.

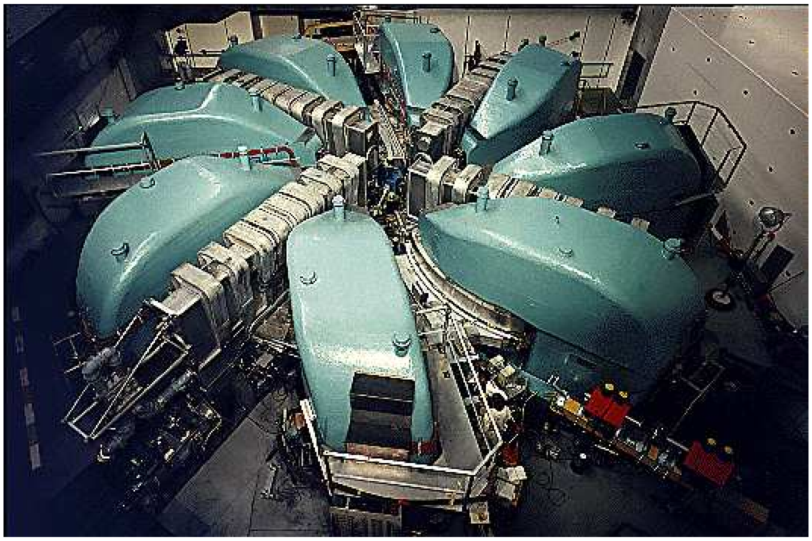


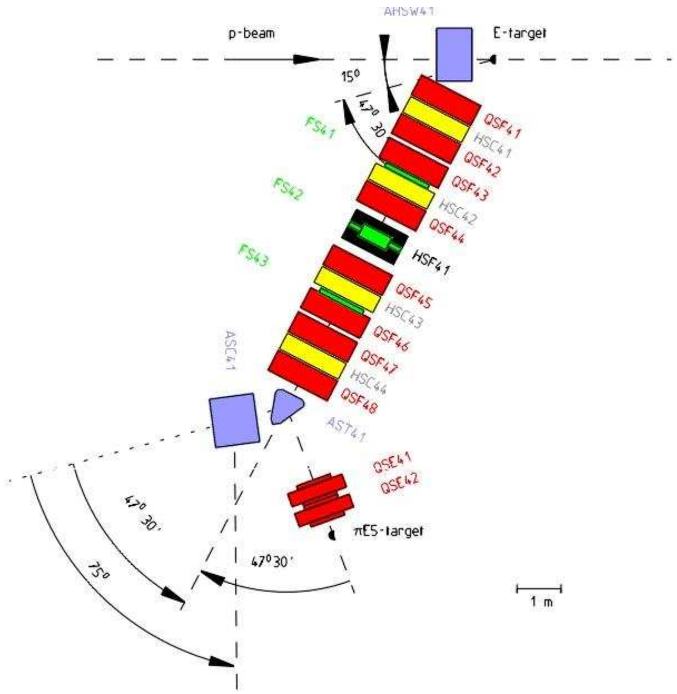
Figure 3.2: The PSI 590 MeV Ring Cyclotron

3.1.1 $\pi E5$ beamline

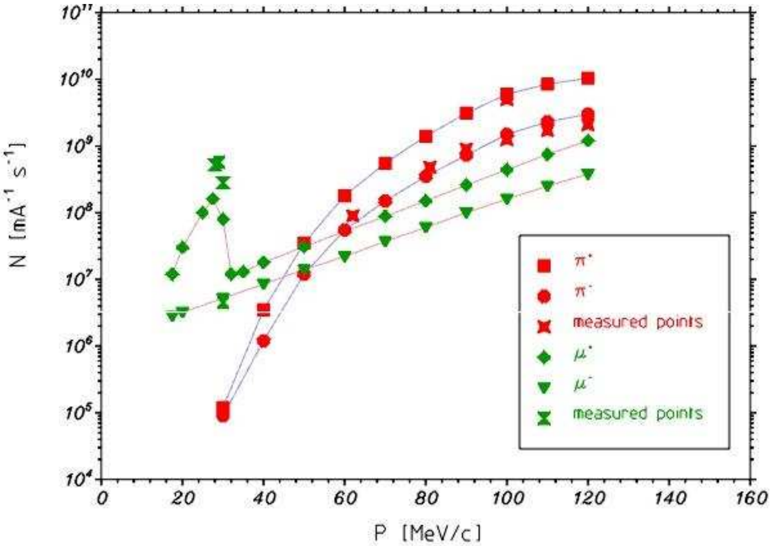
PSI operates a 590 MeV proton cyclotron, which delivers an up to 2 mA proton beam with 1.2 MW beam power (see Figure 3.2). Its design is based on criteria that allow operation at very high beam intensities. A production target is placed on the primary proton beamline and the produced pions and muons are extracted to the $\pi E5$ channel. The $\pi E5$ is a low energy (10-120 MeV/c) pion and muon beamline. Figure 3.3(a) shows the layout of $\pi E5$ beamline. A surface muon [17][18] beam is produced by selecting muons originating from pions that stop near the surface of the production target and decay at rest. The surface muons have a momentum of $\sim$29.8 MeV/c. Muons lose some energy after going through the thin layer of the target surface. But the momentum spread is still low. By tuning the beamline to accept $\sim$28 MeV/c, we can collect high intensity positive muons. The measured muon and pion rates in the $\pi E5$ channel are shown in Figure 3.3(b). A peak is seen at around 28 MeV/c.

3.1.2 Beam Transport System

Muons are transported from the production target to the muon stopping target through a beam transport system consisting of, in order of beam flow, a quadrupole triplet (Triplet I), a crossed-field separator (Wien filter), a second quadrupole triplet (Triplet II) and a beam transport solenoid (BTS). A schematic layout of the beam transport system is shown in Figure 3.4. The triplets are used to focus the beam. The Wien filter, applied with a horizontal magnetic field (133 Gauss) and a vertical electric field (195 kV), vertically separates the positive muons and positrons by 8.1σ, where σ is RMS width of the muon and positron beam profile. A momentum degrader



(a) $\pi E5$ beamline layout



(b) Muon and pion flux in $\pi E5$ channel

Figure 3.3: $\pi E5$ beamline layout and properties

made of 300 μm Mylar is installed in the BTS to reduce the muon momentum. The BTS adjusts the oscillation of the beam profile to minimize the spot size on the stopping target. At beam rate of $3.0 \times 10^7 \mu/s$, the beam can be focused in an elliptical spot with measured (σ_x, σ_y) of (9.5 mm, 10.2 mm) on the muon stopping target.

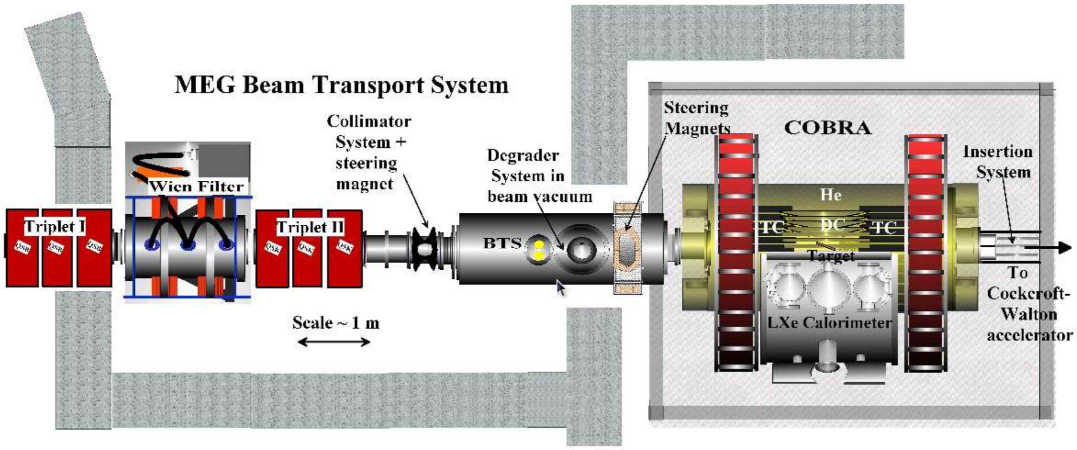


Figure 3.4: MEG beam transport system layout. Muon beam enters from the left.

3.1.3 Muon Stopping Target

The muon stopping target is installed in the center of the MEG detector. It should be thin enough to reduce the multiple scattering and the annihilation of the resulting positrons. The target is made of an elliptically shaped polyethylene foil with thickness of 205 μm (18 mg/cm^2). The foil is supported by a Rohacell[19] frame. The minor and major axes of the foil ellipse have diameters of 79.8 and 200.5 mm, respectively, and the frame has a thickness of 2×5.25 mm and a width of 6.5 mm. Six holes (10 mm diameter) are made on the foil to study vertex reconstruction performance and to align the target position using data. The MEG detector has its full acceptance around $\theta = 90^\circ$, so a smaller slant angle (from the z-axis) of the target is favored to decrease multiple scattering of the out going positrons and to reduce photon conversions; and, on the other hand, a larger slant angle is needed to enhance the muon stopping power. The slant angle is optimized to 20.5° for RUN 2008. Pictures of the target, both uninstalled and installed, are shown in Figure 3.5.

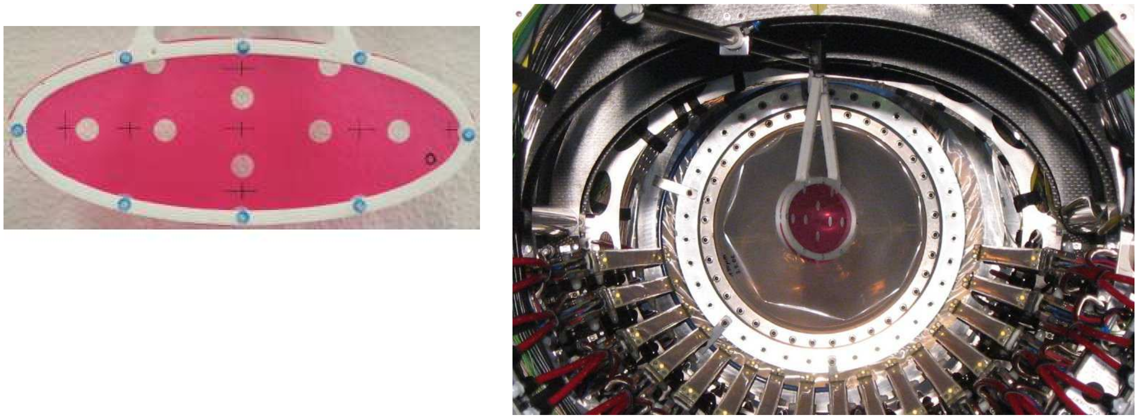

(a) (b)

Figure 3.5: The MEG target. (a), a picture of the uninstalled target; (b), a picture of target installed in position, viewed from downstream.

3.2 Detector

A schematic layout of the MEG detector is shown in Figure 3.6. Positrons emitted from muon decays are analyzed by the magnetic spectrometer, which consists of a thin-walled superconducting solenoid magnet called the COBRA (COnstant Bending RAdius) magnet, a system of 16 low-mass drift chambers, and fast scintillating timing counters located both upstream and downstream of the drift chambers. The positron spectrometer is also called the COBRA spectrometer. A liquid xenon scintillation detector is used to measure the photon kinematics, i.e. the energy, time and position of photons at the first conversion point in the detector. In this chapter, we describe the concept and design of each sub-detector.

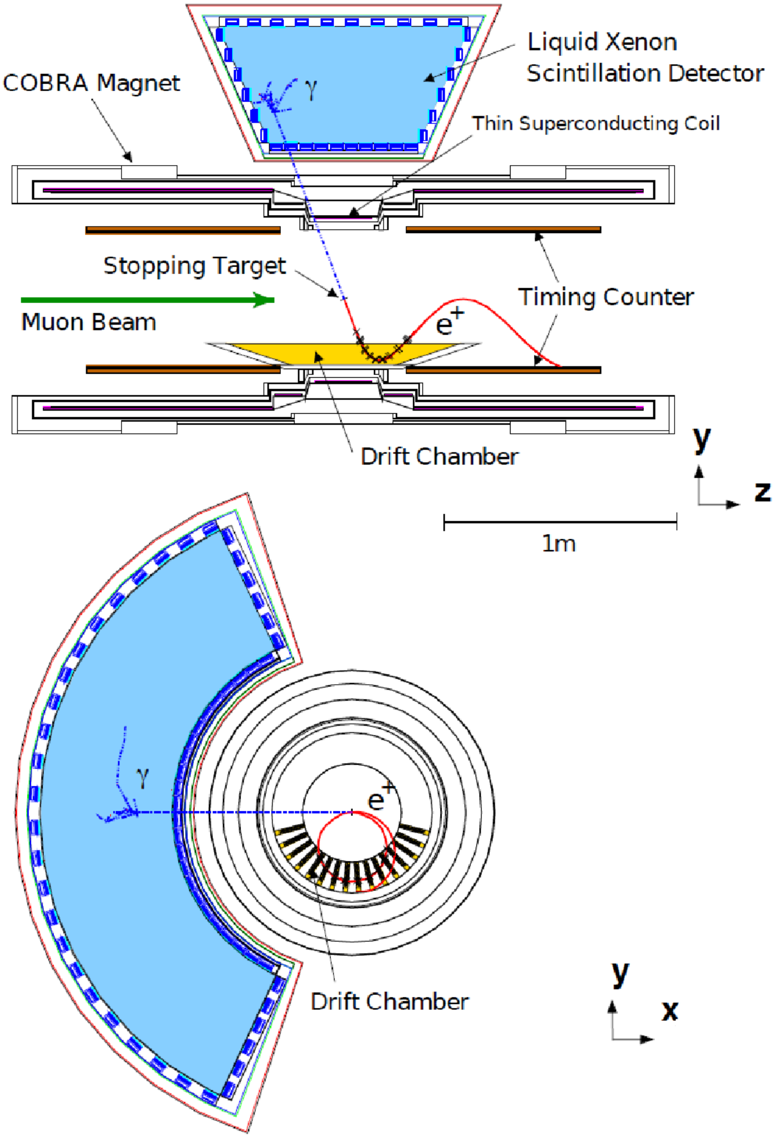


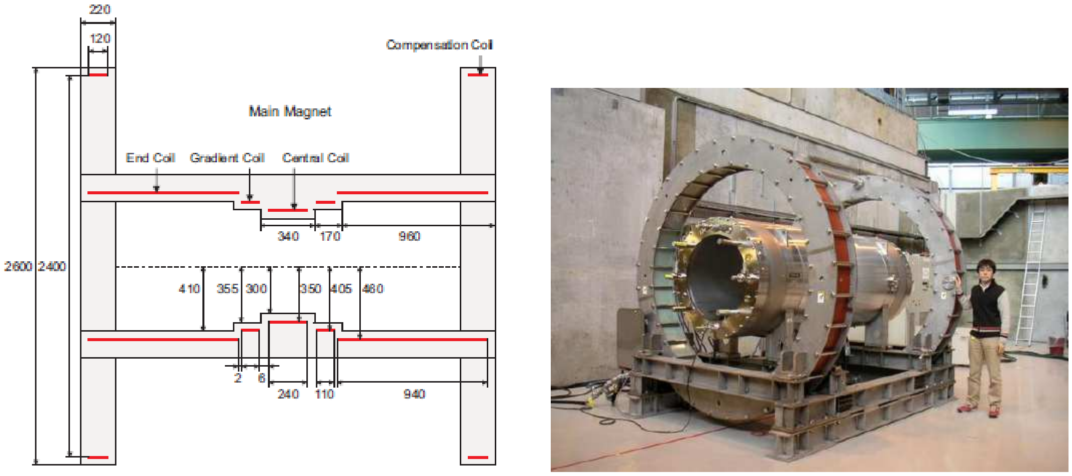
Figure 3.6: A schematic layout of the MEG detector.

3.2.1 The Positron Spectrometer

The nature of the MEG experiment requires its positron spectrometer to work in a high rate environment ($\sim 3 \times 10^7 s^{-1}$) with good momentum, timing and direction resolutions. The design of the spectrometer is based on these requirements.

The COBRA Magnet

We used a solenoidal magnet, called the COBRA magnet, with a gradient field [20] as opposed to a more conventional uniform field. As shown in Figure 3.7, the COBRA magnet system consists of a main superconducting magnet and two resistive compensation coils.



(a) A schematic layout of the COBRA magnet (b) A picture of the COBRA magnet

Figure 3.7: The COBRA magnet

The magnet is designed to increase the radius of the positron helix motion as it travels longitudinally away from the center of the magnet, as shown in Figure 3.9(a). This is realized by arranging the coils in the main superconducting magnet so that the magnetic field decreases as $|z|$ increases, from 1.27 T at z=0 to 0.49 T at $|z|$=1.25 m as shown in Figure 3.8. In a uniform magnetic field, positrons emitted nearly perpendicular to the beam axis undergo many turns in the spectrometer, whereas a gradient field sweeps out such positrons more efficiently (Figure 3.9).

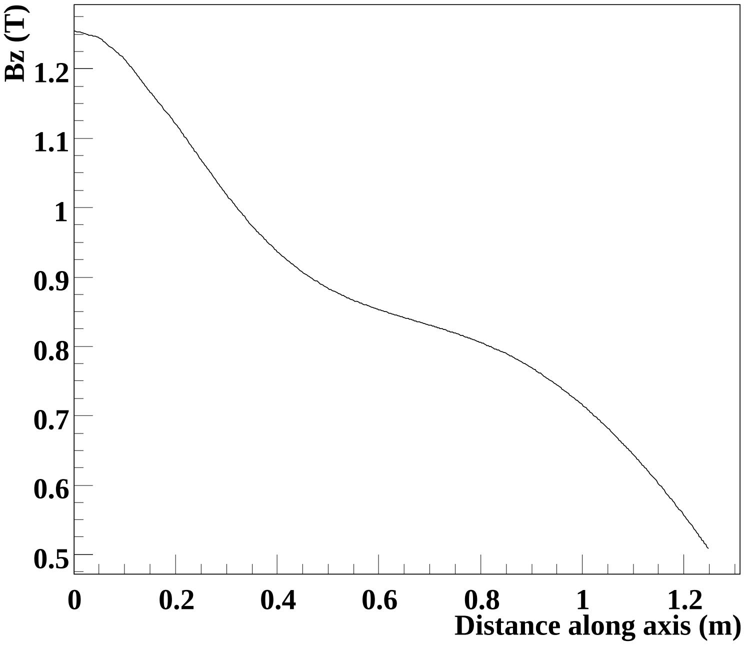
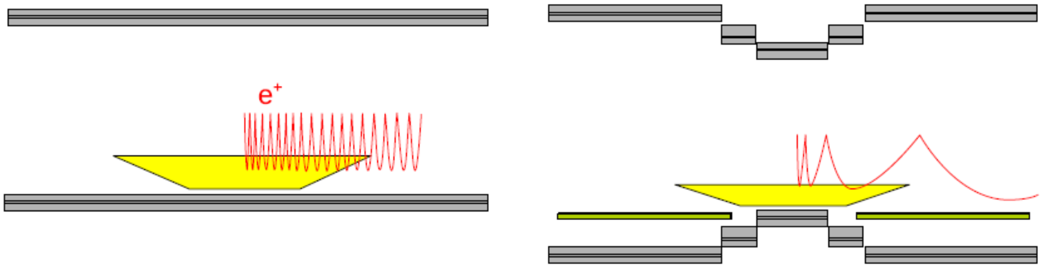


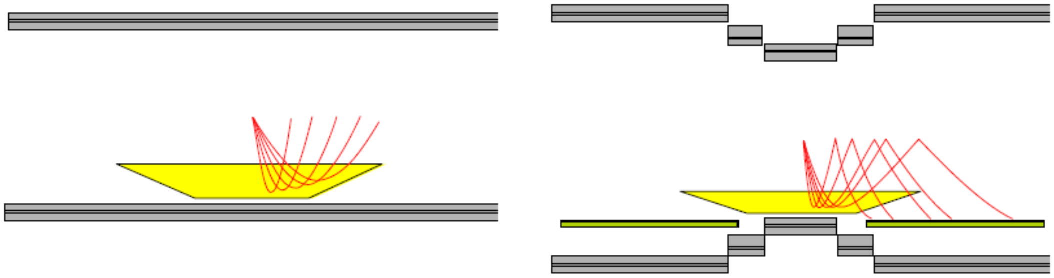
Figure 3.8: Profile of the magnet field along the z axis

In addition, the gradient field is designed so that positrons with the same momentum follow trajectories with a COnstant Bending RAdius independent of their emission angles, as shown in Figure 3.10(b). This is also how the name COBRA is coined. However, in a uniform field, the positron bending radius depends on its emission angle, as shown in Figure 3.10(a). Therefore, it allows us to discriminate high momentum signal positrons from the low momentum background positrons, by designing the position of drift chambers so that they cover a proper range of radius. It avoids accepting low energy positrons which come with very high intensity and could easily saturate the drift chambers. Compared to a uniform field, the COBRA field reduces the hit rate in the region of the drift chamber tracking sub-detector (Figure 3.11).



(a) A uniform magnetic field (b) The COBRA field

Figure 3.9: The COBRA magnetic field compared with a uniform field. (a) positrons emitted close to 90° undergo many turns in the tracking volume in a uniform magnetic field; whereas (b) they get swept out much more quickly in a gradient field.



(a) A uniform magnetic field (b) The COBRA field

Figure 3.10: (a) in a uniform field, the bending radii of positrons with the same momentum depend on their emission angles; whereas (b) in the COBRA field, they are independent of the emission angles.

The outputs of the PMTs used in the liquid xenon photon detector are very sensitive to the magnetic field as shown in Figure 3.12. Therefore, the residual magnetic field produced by the COBRA magnet near the liquid xenon detector can deteriorate the PMT performance. The tolerance is estimated to be less than 5×10^{-3} T. To reduce the residual field, compensation coils were installed at each end of the main magnet. A contour plot of the residual magnetic field near the liquid xenon region is shown in Figure 3.13.

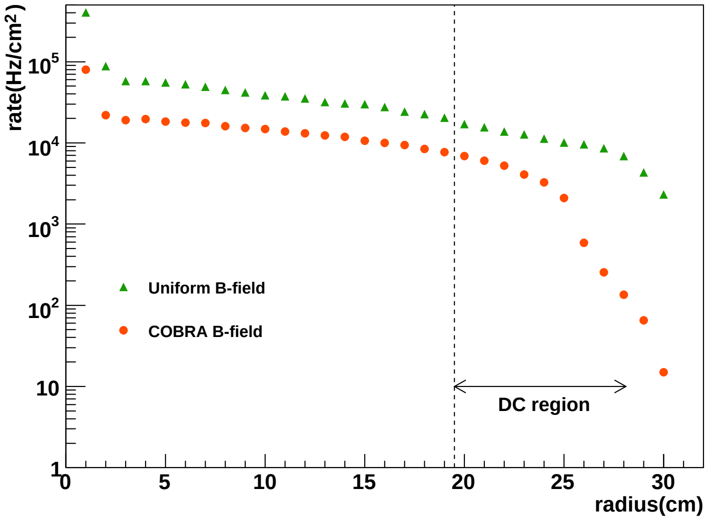


Figure 3.11: Rate of Michel positrons in the unit of $e^+/(cm^2 \cdot s)$ as a function of radius assuming muon decay rate of $3 \times 10^7 \mu/s$.

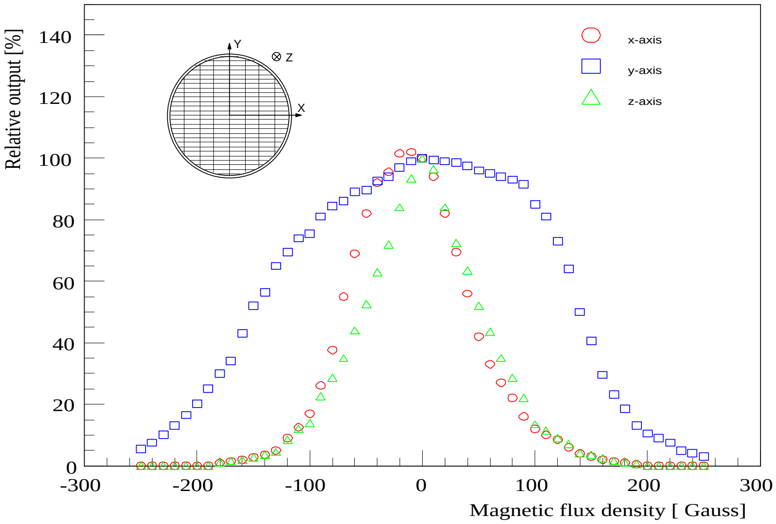


Figure 3.12: PMT response in magnetic field

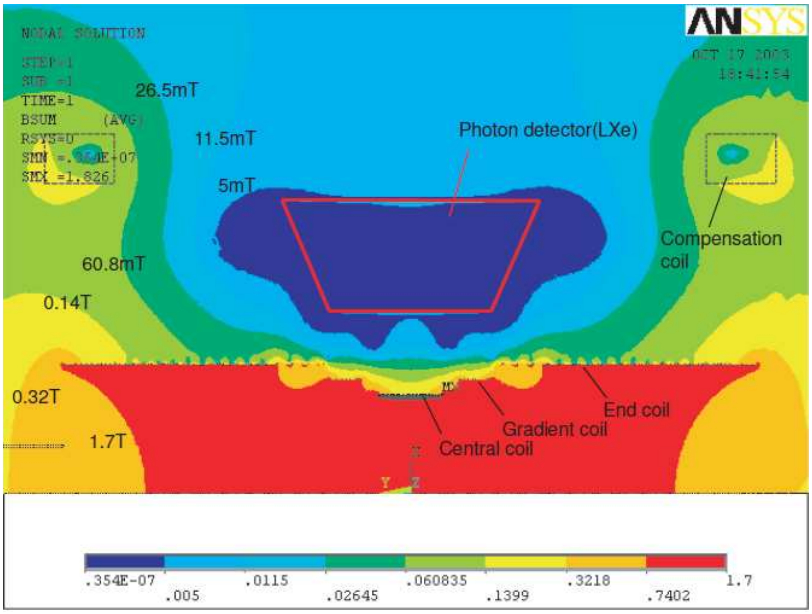


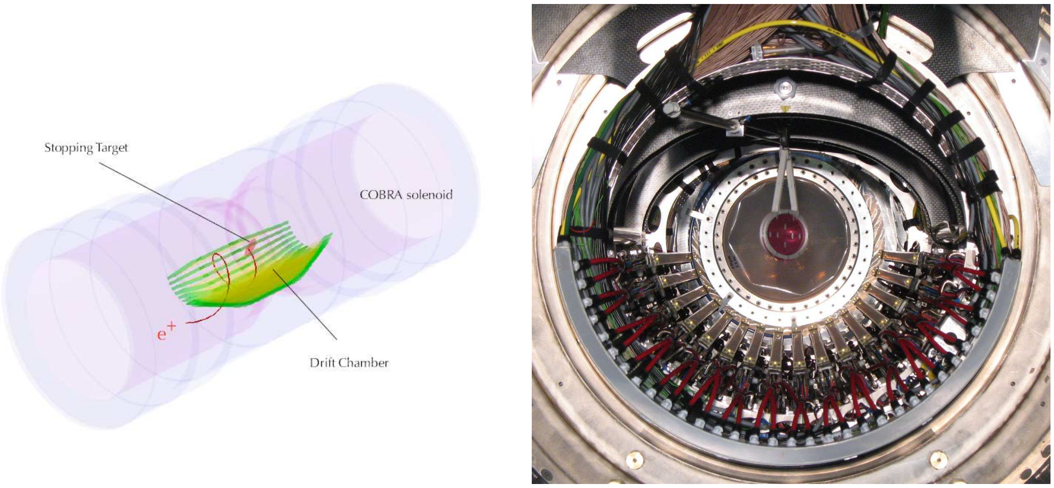
Figure 3.13: A contour plot of the residual magnetic field

The wall of the COBRA magnet is designed to be very thin, since photons from muon decays have to travel through the magnet wall to enter the liquid xenon detector as shown in Figure 3.6. This is to reduce the photon interaction on the wall. The total thickness of the material used in the coil and the cryostat of the magnet in front of the liquid xenon detector acceptance region corresponds to 0.197 radiation lengths of a MEG signal photon (52.8 MeV).

Drift Chambers

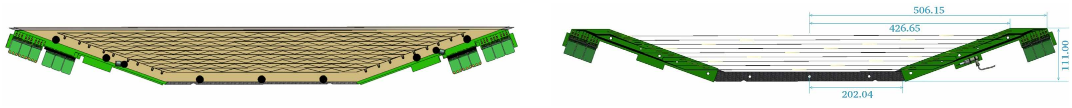
Positron tracks are measured by the drift chamber (DCH) detector located inside of the COBRA magnet. It reconstructs the positron trajectory to deduce the positron momentum, emission angle, positron time-of-flight and the muon decay vertex. The design of DCH is to optimize its performance in a high rate environment.

The DCH detector is a system of 16 drift chambers radially aligned with 10.5° intervals in azimuthal ϕ angle forming nearly a half circle from 11.25° to 168.75° around the muon stopping target. A schematic view of the DCH system and a picture of the installed DCH system are shown in Figure 3.14. The DCH module radially extends from 19.3 to 27.9 cm, which is designed to only have acceptance of high momentum positrons (>40 MeV) as a result of the COBRA gradient field. In the longitudinal direction, the active region extends up to $|z| = 43$ cm at the innermost radius and $|z| = 20$ cm at the outermost. The dimensions of a DCH module is shown in Figure 3.15. MEG signal positrons emitted from the target with $|cos\theta_e| < 0.35$ and $|\phi_e| < 60°$ are covered by the DCH geometry.



(a) (b)

Figure 3.14: (a) a schematic view of the DCH system; (b) a picture of the DCH system installed in position inside of the COBRA magnet.



(a) A DCH module (b) Dimensions of a DCH module

Figure 3.15: A DCH module with its dimensions (unit in mm)

A schematic layout of a single DCH module with geometrical information is shown in Figure 3.16. Each DCH module contains two layers of anode wires aligned along the beam axis, which we call planes. In each plane, there are 9 anode wires separated by 9 mm intervals. In the middle of two adjacent anode wires, there is a potential wire. Anode wires are made of Ni/Cr with a diameter of 25 μm and potential wires are of Be/Cu with a diameter of 50 μm. Anode wires have a resistivity of $2200\Omega/m$. The volume between two adjacent potential wires is called a drift cell. Two planes are shifted by one-half cell. Cathodes are made of 12.5 μm thick ultra-thin polyimide foil deposited with 250 nm aluminum. The cathodes are designed to be very thin to reduce the probability of positron multiple scattering. For the purpose of distinguishing the inner and outer cathodes, the outer ones are sometimes referred as the hoods. Anode wires are supplied with high voltages; 1850 V is the nominal working voltage. Potential wires and cathodes are grounded to maintained a potential difference in a cell.

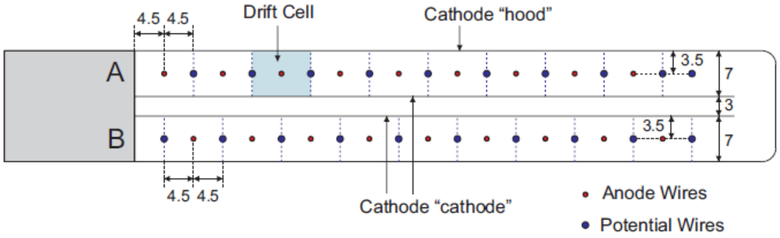


Figure 3.16: A cross-sectional view of a DCH module with geometrical information (unit in mm)

Anode wires are read out on both upstream and downstream sides of a chamber. The z coordinate of a hit is first roughly estimated by the charge ratio of two ends of a wire. A precise estimation is done by using the so-called Vernier pattern[21][22][23][24] on the cathode foil, as shown in Figure 3.17. The Vernier pad is formed by shaping the aluminum deposition on the polyimide foil in a zig-zag fashion with a period of 5 cm into two strips. The two strips are read out on opposite ends of each Vernier pad. The amount of induced charge on each pad is related to the z coordinate due to this zig-zag shape. Phases of the Vernier pattern are shifted by 1/4 of a period for each strip so that a precise z position can be measured from the ratio of charges on the pads.

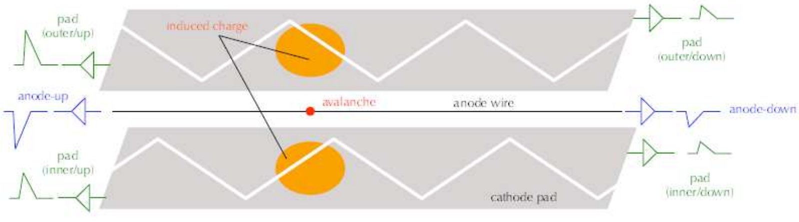


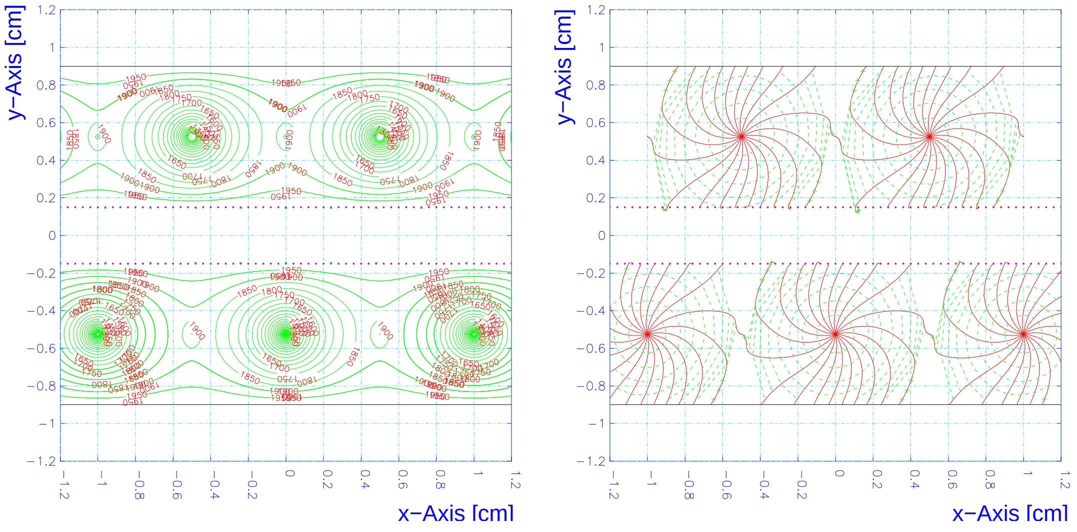
Figure 3.17: Schematic view and conceptual illustration of the Vernier pattern on the cathode pads

The r coordinate of a hit can be first roughly determined from the hit time. It is further measured at the stage of track reconstruction when several hits associated with the same positron passing through are connected and the local left-right ambiguities are then resolved.

The DCH system together with its support structure are installed in the center of the COBRA magnet. The COBRA magnet volume is filled with helium gas to reduce the amount of material in the spectrometer. The active gas inside the DCH modules is a mixture of 50 % helium and 50 % ethane (C_2H_6). This is, on one hand, to have enough ionization and, on the other hand, to minimize scattering[25][26]. The field map and the drift lines calculated by the GARFIELD[27] simulation are shown in Figure 3.18.

Details of the design and construction of the DCH detector, as well as DCH performance during engineering Run 2007, is described in a thesis [28].

The wires and cathodes of a DCH module are supported by a thin carbon-fiber frame.



(a)

(b)

Figure 3.18: (a) contour plot of the potential, and (b) drift lines with isochrone map calculated by the GARFIELD simulation.

It is designed to be a three-edged open frame, such that there is no support on the target side (see Figure 3.15(b)). This is to reduce the amount of material where many positrons pass through. The shape of the drift cell is therefore only formed by the 12.5 μm thick ultra-thin chamber wall. To maintain a stable field during operation, the differential pressure between the inside and outside of the DCH modules has to be precisely controlled.

A dedicated gas flow and pressure control system is developed to provide He in the COBRA magnet volume, to mix and provide He and C_2H_6 gas mixture to the DCH system, and most importantly to precisely control the differential pressure between the DCH volume and the COBRA magnet volume. We require that $|dP| = |P_{dch} - P_{cobra} - P_{setpoint}| < 1$ Pa, where $P_{setpoint}$ is the pressure difference we want to maintain between the inside and outside of the DCH system. We also want to regulate the deviation from $P_{setpoint}$ to less than 1 Pa to keep the cathode shape.

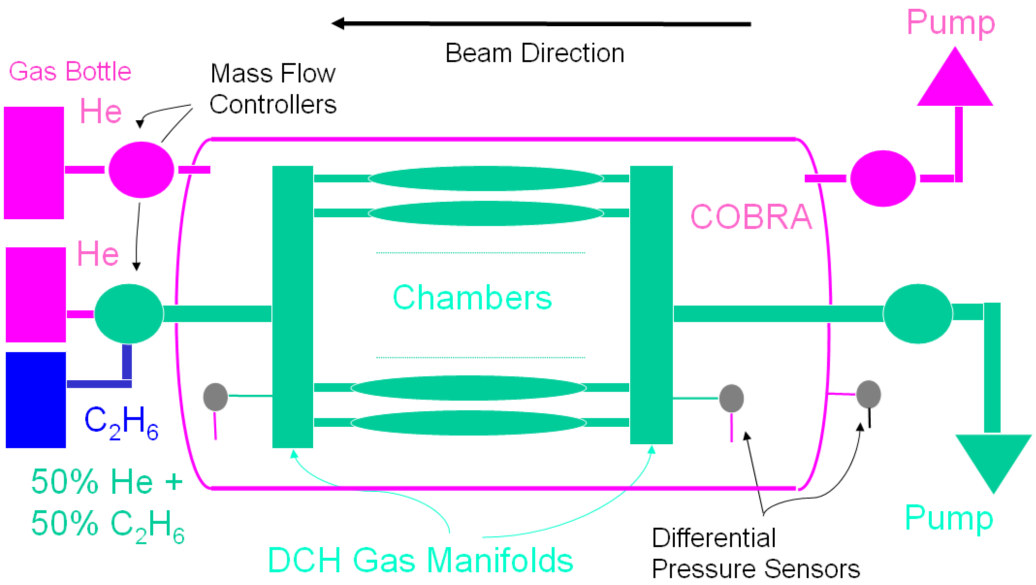


Figure 3.19: A layout of the gas system

A layout of the gas system is shown in Figure 3.19. 16 DCH modules are connected by gas tubes together with two manifolds where differential pressure sensors are installed to measure the pressure difference between the manifolds and the ambient COBRA magnet volume. Since there is a gas flow from the DCH inlet to the outlet manifold, there exists a pressure difference between them. The differential pressure between the DCH modules and the ambient COBRA magnet volume is taken as the average of the two pressure sensor readings. Mass flow controllers in the gas inlet lines maintain a constant gas flow during stable operation, and the ones in the gas outlet lines regulate according to the readings of the differential pressure sensors to adjust the pressure difference. This feedback control regulation is implemented by a slow control system. We operate the DCH at a little over pressure by setting $P_{setpoint} = 1.2$ Pa. This is to avoid He in COBRA magnet volume from leaking into the DCH volume and contaminating the chamber gas mixture. An example performance of the pressure control regulation is shown in Figure 3.20. It shows a stable performance of the pressure control regulation during two days, with a maximum fluctuation amplitude of ∼0.02 Pa, and an RMS fluctuation of less than 0.005 Pa. An additional gas line provides the COBRA gas and maintains the pressure difference between the COBRA magnet volume and the ambient atmosphere to be close to zero.

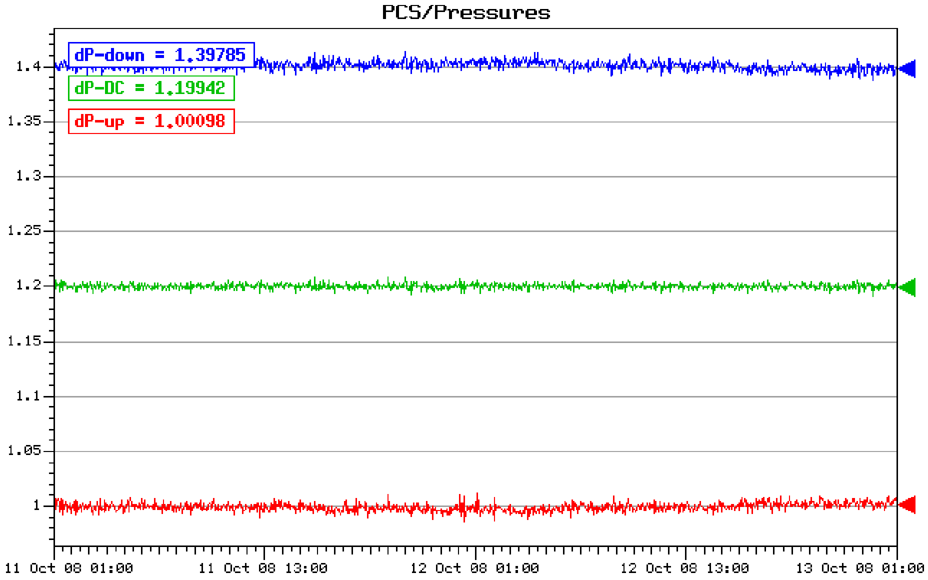


Figure 3.20: $P_{dch} - P_{cobra}$ [Pa] vs. time. The differential pressure measured at the inlet manifold (blue) is higher than that measured at the outlet manifold (red); while the differential pressure at the DCH modules (green) is taken as the average of the two. Stable performance of the pressure control regulation is shown over two days. The maximum fluctuation amplitude is ∼0.02 Pa and the RMS fluctuation is less than 0.005 Pa.

A safe mode exists in case of an emergency, i.e. over pressure, hardware failure, etc. A separate electronic safety board is installed serving as a redundant safety in case of a control unit failure. The gas system is proven to be the most stable and reliable sub-system in the experiment.

Timing Counters

The positron timing is measured by the timing counter (TIC) detector. Two groups of timing counters are installed in the COBRA magnet, on both the upstream and downstream sides of the DCH system. After passing through the DCH, positrons hit the TIC which measures the timing and impact position. The TIC consists of two orthogonally placed arrays of plastic scintillators: the outer layer counter and the inner layer one. The inner layer is stacked on top of the outer one. Each layer is designed to have specific tasks.

The outer layer counter (Figure 3.21), also called the ϕ-counter or the TIC bars, consists of 15 plastic scintillation bars at each end. They are made of BICRON BC-404[29] with a dimension of $4 \times 4 \times 8\ cm^3$. A 2-inch fine-mesh HAMAMATSU R5924[30] PMT is attached at both ends of each bar. They are placed at a radius of 32 cm from the beam axis. Each bar has a radial coverage of 10.5°. Therefore, the entire ϕ-counter covers 160° in total, from $-150°$ to 10°. The bars extend over the region $29 < |z| < 109$ cm. The primary task for TIC bars is to precisely measure the positron impact time and ϕ angle, though the z coordinate of a TIC hit can be deduced by comparing the PMT timings at two ends of a bar.

The inner layer counter (Figure 3.22), also called the z-counter or the TIC fibers, consists of 128 plastic curved scintillation fibers at each end placed on top of the ϕ-counter at a radius of 29 cm from the beam axis. The TIC fiber is made of SAINT-GOBAIN BCF-20[29] with a dimension of $6 \times 6\ mm^2$. Each fiber is read out by HAMAMATSU S8664-55[30] silicon avalanche photo-diodes (APD). The primary task of the z-counter is to obtain a precise and fast measurement of the z position of the impact point, which is designed to use in the trigger. However, in RUN 2008 (see Chapter 5), we

(a) (b)

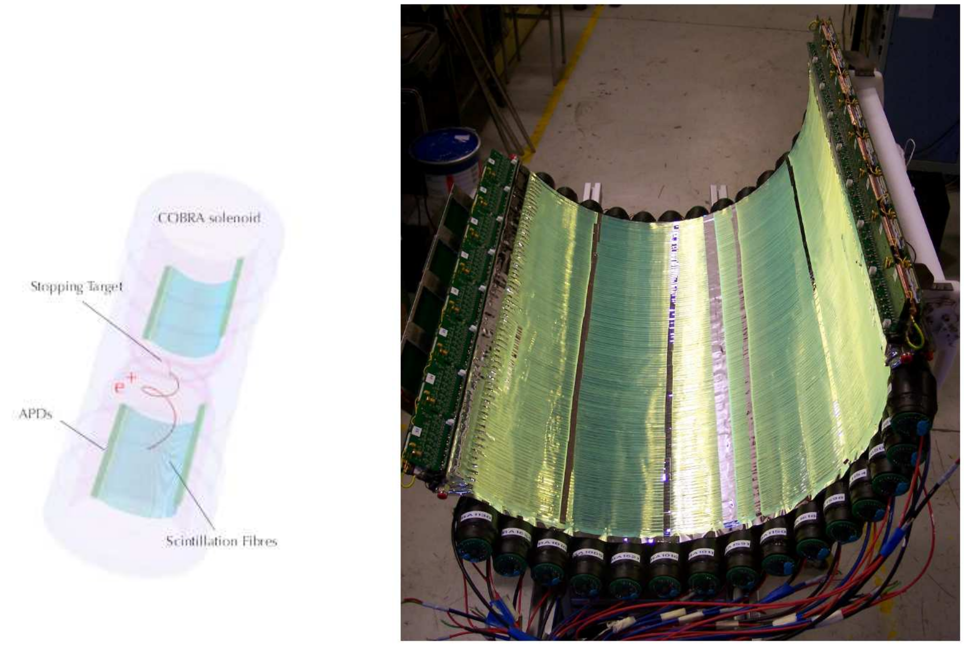
Figure 3.21: (a) a schematic view of the ϕ-counter, or TIC bars; (b) a picture of the ϕ-counter before installation into the COBRA magnet volume.

encountered problems when operating the z-counter. Therefore, the fibers were not used in triggers; ϕ-counter determined z information was used instead. Throughout this thesis, the z-counter is not used, either.

Details of the design, construction, and beam test performance of the TIC detector are described in theses [31] and [32].

3.2.2 Liquid Xenon Scintillation Detector

Photons from muon decays, unlike positrons which are bent by the magnetic field, pass straight through the thin wall of the superconducting coil and enter the photon detector, a liquid xenon calorimeter (XEC).



(a) (b)

Figure 3.22: (a) a schematic view of the z-counter, or TIC fibers; (b) a picture of the z-counter placed on top of the ϕ-counter before installation into the COBRA magnet volume.

The Liquid Xenon and the Detector

A layout of the XEC is shown in Figure 3.23 and a schematic view shown in Figure 3.24. The XEC is a scintillation detector filled with a 900 liter homogeneous volume of liquid xenon. The photons interact with the liquid xenon and generate scintillation light which is collected by 846 PMTs surrounding the XEC active volume. The detector measures the total energy released by the incident photon as well as the position and time of its first interaction point simultaneously.

Liquid xenon (LXe) is widely used in experimental particle physics and astrophysics[33] as a detector material. Details of the LXe scintillation mechanism are described in [34], [35] and [36]. Physical and scintillation properties of LXe are summerized in Table 3.1. The liquid phase makes it possible to build a single homogeneous detector

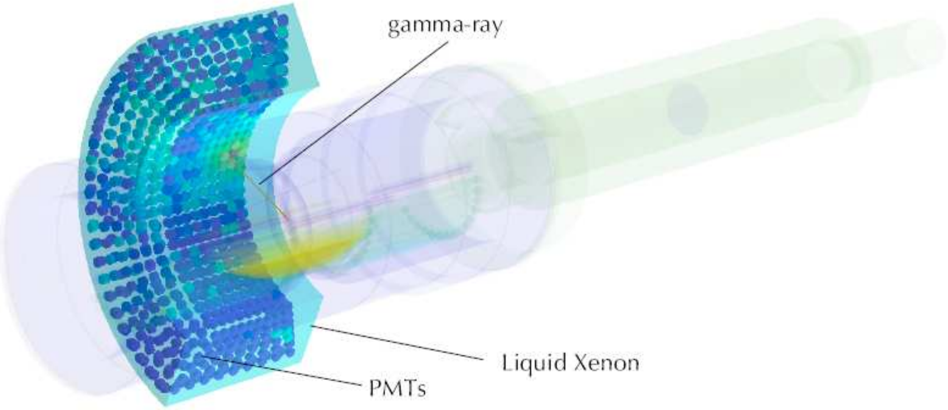


Figure 3.23: A layout of the liquid xenon detector

with arbitrary size and shape. LXe has a fast response (45 ns), which enables high timing resolution. It also has a high light yield (80% that of NaI, see Table 3.2), which is crucial for precise energy measurement. The absorption length for the scintillation photon (> 100 cm) is much bigger than the depth of the active LXe volume (38.5 cm, see Figure 3.25), which enables high energy resolution with a large-sized detector, as we have. Comparisons of scintillation properties between LXe and various other scintillation materials are listed in Table 3.2.

The PMTs are immersed in the liquid xenon to observe scintillation photons directly. They are installed on all six faces of the detector. The definition of the six faces, i.e. inner, outer, upstream, downstream, top and bottom, is illustrated in Figure 3.25. As shown in the figure, the PMT arrangement and density are different depending on the face. A local coordinate (u, v, w) is also defined and illustrated. The u-axis is identical to the global z-axis. The v-axis is defined as a direction of negative ϕ along the surface of the inner face of the XEC. Its center is at global $y = 0$. The w-axis is defined as the depth in the XEC. The direction is the same as the radial direction but the origin of the axis is at the surface of the inner face or, in other words, at global r_0

Properties	Value & Unit	Ref.
Atomic Number	54	
Mass Number	131.293	[9]
Density at 161.4 K	2.978 g/cm^3	[37]
Boiling point	165.1 K	[9]
Melting point	161.4 K	[9]
Radiation length	2.77 cm	[9]
Mollier radius	4.2 cm	[38]
Scintillation wavelength (peak±FWHM)	(178 ± 14) nm	[39]
Refractive index at 175 nm	1.57 to 1.72	[40]
Decay time (recombination)[1]	45 ns	[41]
Decay time (fast components)[1]	4.2 ns	[41]
Decay time (slow components)[1]	22 ns	[41]
Absorption length	> 100 cm	[42]
Scattering length	29 cm to 50 cm	[43]

---

[1] Due to different scintillation processes, details seen [34], [35] and [36]

Table 3.1: Properties of LXe.

	LXe	LAr	NaI	CsI	BGO	LSO	$PbWO_4$
Density (g/cm^3)	2.98	1.40	3.67	4.51	7.13	7.40	8.3
Radiation length (cm)	2.77	14	2.59	1.86	1.12	1.14	0.89
Mollier radius (cm)	4.2	7.2	4.13	3.57	2.23	2.07	2.00
Decay time (ns)	45	1620	230	1300	300	40	30/10[1]
Emission peak (nm)	178	127	410	560	480	420	425/420[1]
Relative light yield (%)	80	90	100	165	21	83	0.083/0.29[1]

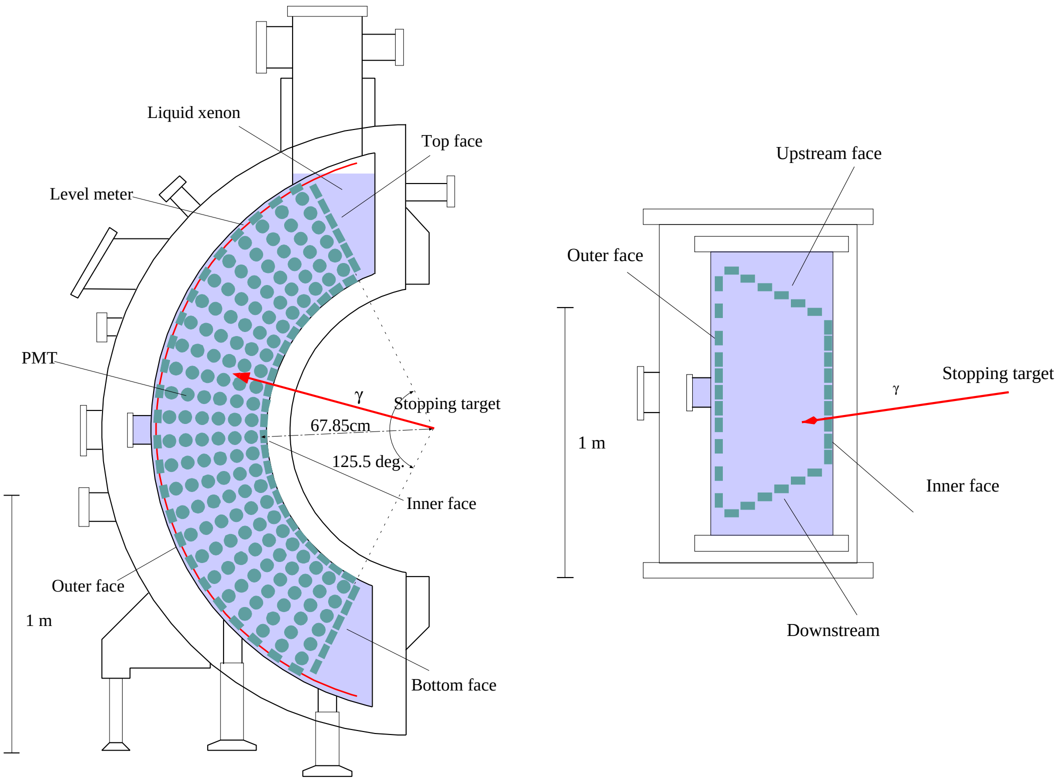
---

[1] slow/fast component

Table 3.2: Comparisons of scintillation properties between LXe and various other scintillation materials.

= 67.85 cm (dimension information seen in Figure 3.24). Conversion from the global coordinate to the XEC local one can be written as,

$$u = z, v = r_0 \cdot tan^{-1}(-y/x), w = \sqrt{x^2 + y^2} - r_0. \tag{3.1}$$



(a) Side view.

(b) Top view.

Figure 3.24: A schematic view of the XEC.

Sometimes multiple photons enter the XEC and are measured in the same time window due to operating in a high rate environment and the fact that the detector has a large active volume without any segments. We identify these kind of pileup events from the light distribution among the PMTs. The timing distribution and the PMT waveform shapes are also used in identifying pileup events.

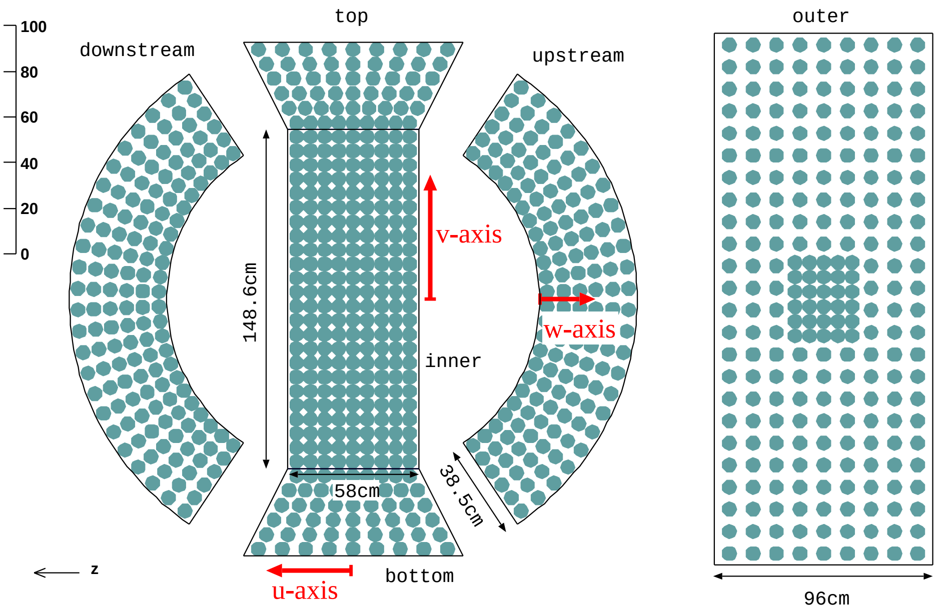


Figure 3.25: Illustration of the XEC faces, PMT arrangement, and the XEC local coordinate.

Though the scintillation photons do not get re-absorbed by the LXe itself, they can be easily absorbed by contaminants, like water or oxygen. Therefore, the purity of the LXe is crucial to the detector performance. Two purification systems, in both liquid and gaseous phases were developed. The gaseous purification system circulates at a slow speed ($< 100\ cm^3$ liquid/hour), and is operated in parallel with normal data acquisition. The liquid-phase purification, though having a high circulation speed (up to ∼35 l/hour), is only used before data taking and during beam maintenance breaks. Liquid-phase purification is not used during data taking due to the high noise level generated by the pump. The scintillation light of liquid xenon is vacuum ultraviolet(VUV) light with a peak at 178 nm. Normal PMT windows made of glass are not transparent to VUV light. The PMT model, R9288[30], developed by Hamamatsu Photonics for MEG, is equipped with a quartz window which is 80% transparent to the scintillation light. From Table 3.1, we can see that the liquid phase of xenon has a narrow temperature range around 4 K; therefore, a dedicated cryostat system exists to maintain the xenon temperature. Technical details of the development, installation, testing and performance of the XEC in engineering Run 2007 are described in a thesis [44].

LED and Alpha Sources

In the MEG experiment, the PMT quantum efficiency (QE) and gain are measured separately. The PMT gain is used to convert the PMT voltage output to the number of photoelectrons (N_{pe}), and the QE to convert N_{pe} to the number of photons (N_{ph}). The QE should be stable for long times, therefore, it is measured only a few times. On the other hand, it takes time to stabilize the gain when the PMT working environment, such as temperature or high voltage, changes.

Blue LED sources are installed in the XEC, 6 each on both the upstream and downstream faces, as constant light sources for the purpose of PMT gain calibrations. The LEDs are covered with Teflon[45] sheets with some small pinholes to attenuate emitted light so that the statistical fluctuation of LED light intensity is small. Typically 10 LEDs are flashed at the same time in normal calibration to illuminate all the PMTs uniformly. A picture of some LED light sources is shown in Figure 3.26.

In order to measure the QE, we use the same wavelength of light as LXe scintillation light since the QE highly depends on the wavelength. Therefore, the LED is not suitable for this purpose. Instead, we obtain the scintillation light by using radioactive sources. A calibration method with a radioactive point-source lattice was developed solely for this purpose[46]. Alpha rays from ^{241}Am have an almost monochromatic energy spectrum, with 5.485 MeV (84.5 %) and 5.443 MeV (13.0 %) components, whose range in LXe is as short as $40\mu m$, and thus can be treated as a point-like light

Figure 3.26: A picture of LED (red circle) and alpha (yellow arrow) sources installed in the XEC

source from a known position. In addition, it is very stable even at low temperatures in LXe. The half life of ^{241}Am is long enough that the source intensity can be regarded as constant. Alpha sources of 1mm in length are put on tungsten wires, with a diameter of 100 μm. A picture of an alpha source on a wire is shown in Figure 3.27. There are in total 5 wires with 5 alpha sources attached to each wire at constant intervals of 12.4 cm. A lattice of 25 alpha-source spots is formed in the XEC active volume. Example of placements of alpha sources is also shown in Figure 3.26.

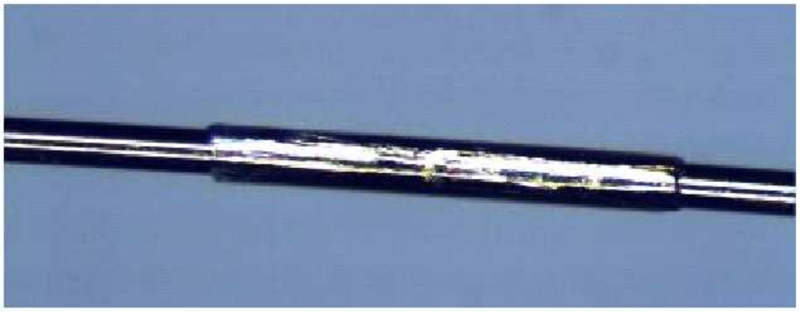


Figure 3.27: A magnified picture of an alpha source on a wire

3.3 MEG Calibration Devices

The XEC does not give an absolute photon energy scale. It has to be calibrated against known energy sources. The MEG detector as a whole also has to calibrate its timing coincidence signal position due to different cable lengths, etc. In addition to calibration, we have to monitor the running stability of all detectors under a high beam intensity, which may have variations with running time. Therefore, several calibration and monitoring methods are developed, which are covered in this section.

3.3.1 Pion Charge Exchange

Principles

To calibrate the energy scale of the XEC at around and above the MEG signal region (52.8 MeV), high energy photons from neutral pion decays ($\pi^0 \rightarrow \gamma\gamma$) are used. The neutral pion is produced by a charge exchange (CEX) process of a negative pion and a proton ($\pi^-\mathrm{p} \rightarrow \pi^0\mathrm{n}$). The produced neutral pion has a momentum of 28 MeV/c in the laboratory frame and immediately decays into two photons ($\pi^0 \rightarrow \gamma\gamma$) with a branching fraction of 98.80%. The decay kinematics are shown in Figure 3.28. Due to the decay kinematics and the residual pion momentum, the photon spectrum contineously ranges from 54.9 to 82.9 MeV. But if we select back-to-back photons, we are able to get monochromatic photons at energies of 54.9 and 82.9 MeV. In addition to this decay mode, the so called Dalitz decay of neutral pions ($\pi^0 \rightarrow \gamma\mathrm{e}^+\mathrm{e}^-$, with a branching fraction of 1.20%), which has a similar decay topology as the MEG signal ($\mu^+ \rightarrow e^+\gamma$), and radiative muon decays ($\mu^+ \rightarrow e^+\nu_e\overline{\nu}_\mu\gamma$), is used to calibrate the timing coincidence offset between the XEC and the TIC, as well as to study the resolutions of the detectors.

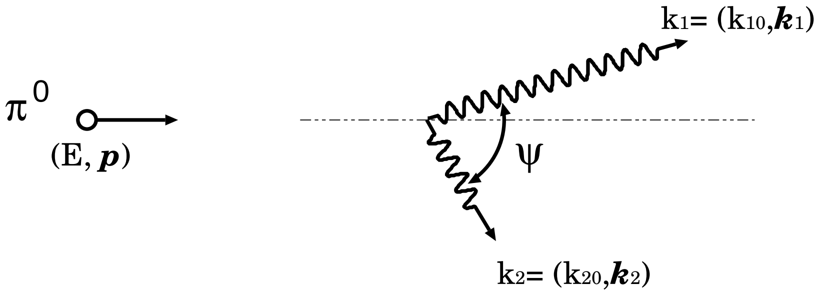


Figure 3.28: Neutral pion decay kinematics

Beam and Target

The charged pion beam is provided by the $\pi E5$ beamline. The momentum of the pion beam is set to 70.5 MeV/c, optimized to separate pions from muons and electrons. The pion beam intensity is calculated to be $\sim 1.5\times 10^6$ pions/s, and the beam profile is measured to be an elliptical spot with (σ_x, σ_y) of (8.5 mm, 7.5 mm). Several settings with reduced beam intensities are also prepared.

Liquid hydrogen (LH_2) is used as the charge exchange reaction target. It is liquefied by a flow of liquid helium and confined in a cylindrical cell of 50 mm in diameter and 75 mm in length. The cylindrical cell has a thin 135 μm Mylar film as the beam entrance window. A picture of the LH_2 target cell is shown in Figure 3.29.

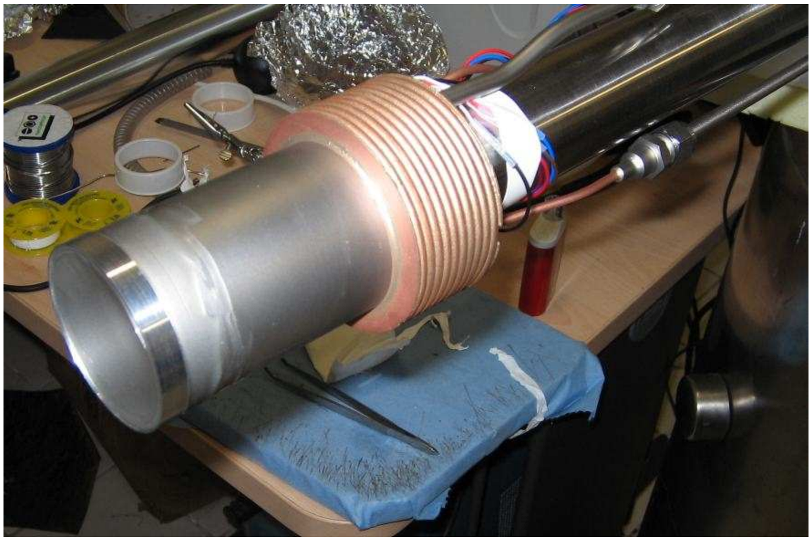


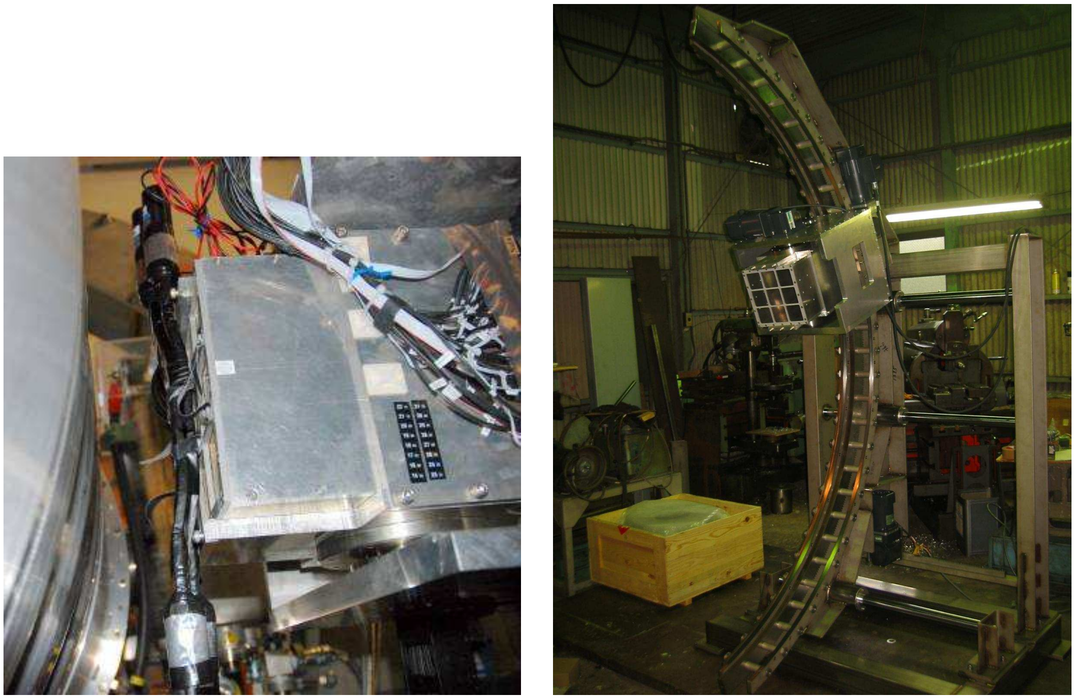
Figure 3.29: A picture of the LH_2 target cell for the pion charge exchange reaction before its installation

NaI Detector

In order to capture back-to-back photons from π^0 decay, a tagging NaI detector is placed at the opposite side of the XEC. The NaI detector consists of nine NaI crystals ($62.5 \times 62.5 \times 300.5$ mm^3), two plastic counters ($60 \times 60 \times 7$ mm^3) and a lead plate ($70 \times 70 \times 5$ mm^3). A picture of the NaI detector is shown in Figure 3.30(a). The lead plate is used to convert photons into electron/positron pairs. The plastic counters are used to precisely measure the incident time of the resulting charged particles. Two fine-mesh PMTs (Hamamatsu H6152-70[30]) are mounted, with one at each end of a counter. The NaI crystals act as a scintillation medium which allows a precise energy measurement. Each crystal is connected to an APD (Hamamatsu S8664-1010[30]). The tagging detector is mounted on a movable stage (see Figure 3.30(b)). It allows the detector to move in both the r and ϕ directions, and to rotate so that the detector can be put at an opposite direction to any selected triggering region. 2 PMT × 3 PMT regions are used until the entire XEC inner face has been scanned. The lead converter can be removed when better energy resolution is needed.

3.3.2 Cockcroft-Walton Proton Beamline

The Cockcroft-Walton (CW) proton accelerator[47] is used to generate the reaction $\mathrm{p} + {}^{7}_{3}\mathrm{Li} \rightarrow {}^{8}_{4}\mathrm{Be} + \gamma$. This reaction is resonant at proton energy $E_p = 440$ keV and gives a 17.6 MeV photon[48]. This is used to calibrate the XEC energy scale at a lower photon energy. This reaction was used in a previous $\mu^+ \rightarrow e^+\gamma$ search experiment[8][49]. Protons from the CW beamline are also used to cause the reaction $\mathrm{p} + {}^{11}_{5}\mathrm{B} \rightarrow {}^{12}_{6}\mathrm{C} + \gamma$. This reaction is resonant at $E_p = 163$ keV and gives photons at 16.1 MeV, 11.7 MeV and 4.4 MeV, depending on the nucleus transition. If ${}^{12}_{6}\mathrm{C}$ directly reaches its ground



(a) (b)

Figure 3.30: (a) a picture of the NaI detector; (b) a picture of the movable stage on which the NaI is mounted.

state, the emitted photon energy is 16.1 MeV. But if $^{12}_{6}$C reaches its first excited state at 4.4 MeV, then a 11.7 MeV photon is emitted in addition to the 4.4 MeV photon. The 11.7 MeV and 4.4 MeV photons are emitted in timing coincidence, which provides a unique opportunity to calibrate the relative timing between the XEC and the TIC, as well as the TIC bar-to-bar offset using XEC time as a reference.

A picture of the CW proton accelerator is shown in Figure 3.31(a). It is mounted on the downstream side of the MEG apparatus. A layout and a picture of the CW beamline is shown in Figures 3.31(c) and 3.31(d). We use a lithium tetraborate ($Li_2B_4O_7$) target (see Figure 3.31(b)) with the purpose of getting photons from both Li and B reactions simultaneously. The target is inserted by a bellows system (see Figure 3.31(e)) into the center of the COBRA magnet during the CW calibration, while the MEG target is removed to its parking position. When the CW calibration is done, the CW target gets extracted out and the MEG target is restored to its nominal working position for normal MEG data acquisition. The CW target insertion/extraction causes big volume changes, hence, high pressure changes, in the COBRA magnet volume. Therefore, the DCH gas control system has a dedicated mode to flux He with a high rate (up to 2 l/min) to compensate for the pressure change in the COBRA magnet volume and to maintain the RMS deviation of differential pressure between the DCH and the COBRA magnet volume less than 0.005 Pa. The target exchange enables us to switch from the normal muon beam to the CW beam in $\sim$20 minutes. This allows us to use the CW beam to monitor the stability of the XEC performance on a 3-day basis without losing much muon beam time.

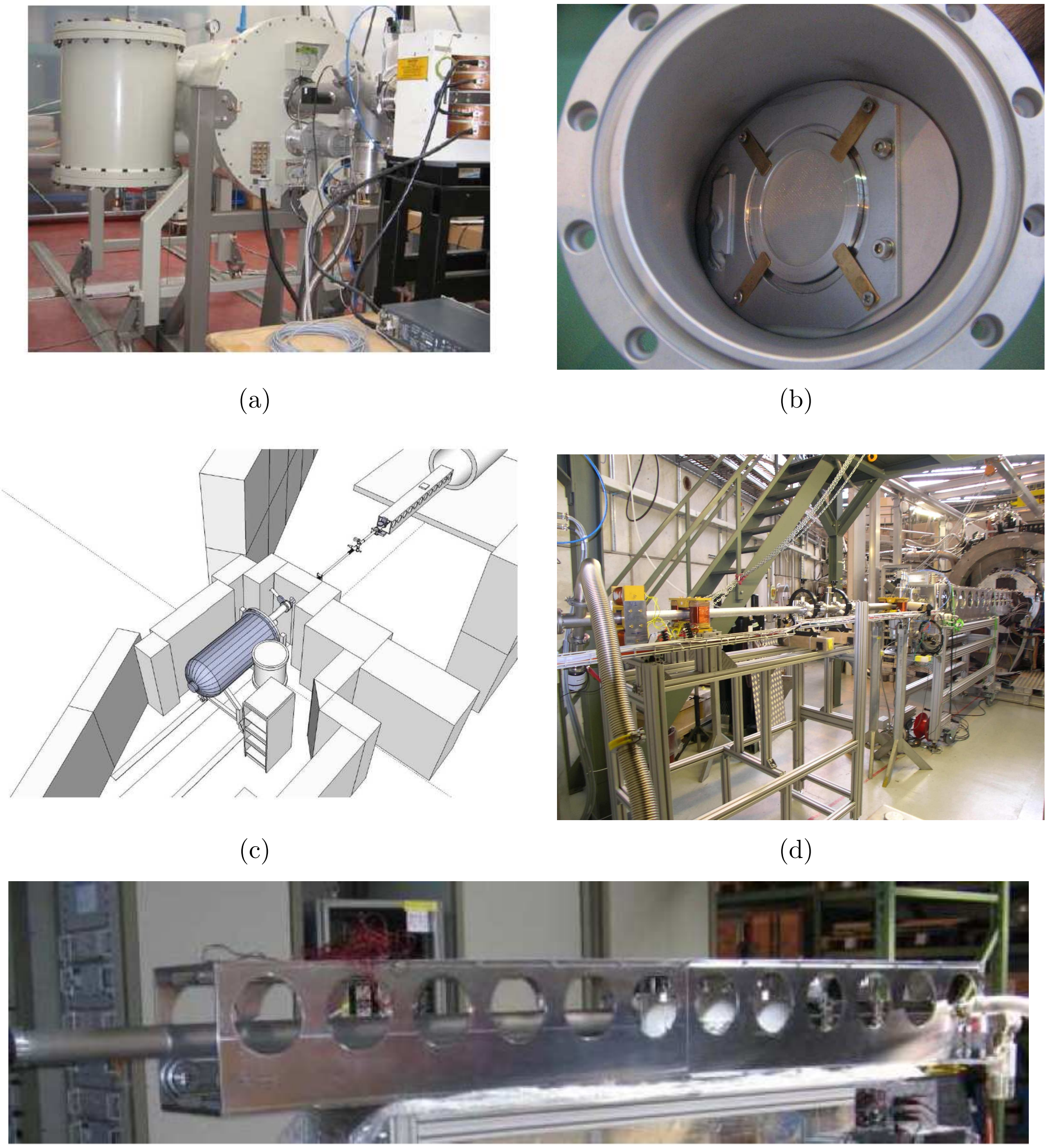
3.4 Electronics, Trigger and Data Acquisition

In this section we summarize the trigger electronics and the data acquisition (DAQ) system.

3.4.1 Data Flow

Analog signals from each detector pass through various devices before finally being digitized by waveform digitizers. A schematic view of the data flow from the detectors to the digitizers is presented in Figure 3.32.

PMT outputs from the XEC are sent to active splitters through coaxial cables. Each splitter board has 16 input channels and three types of outputs. The first output sig-



(a)

(b)

(c)

(d)

(e)

Figure 3.31: (a) a picture of the front face of the CW accelerator; (b) a picture of the $Li_2B_4O_7$ target; (c) a layout of the CW accelerator, the beamline and the πE5 area; (d) a picture of the CW proton beamline; (e) a closeup view of the CW bellows system to insert and extract the $Li_2B_4O_7$ target into and out of the center of the COBRA magnet.

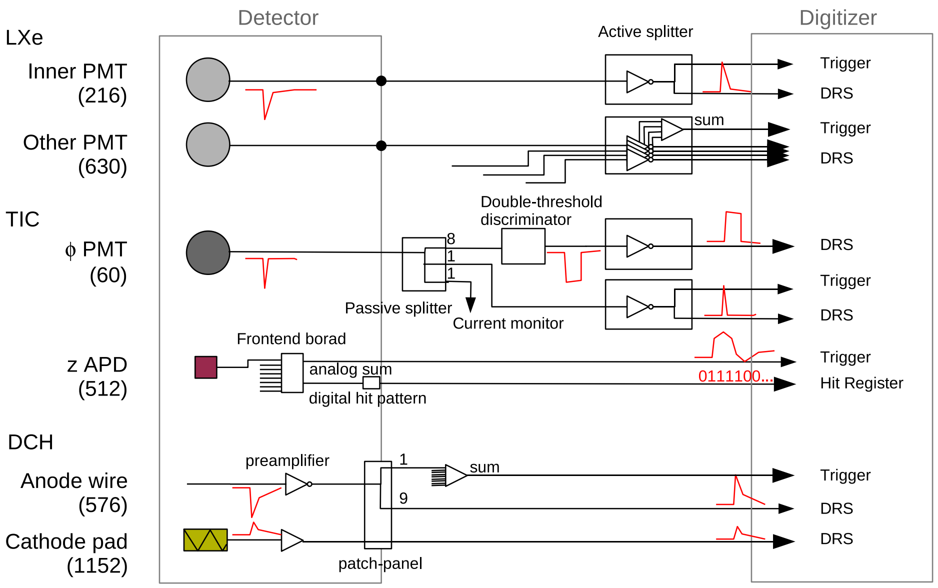


Figure 3.32: A schematic view of the data flow from detectors to digitizers. Different analog signals pass through different devices.

nal is from wide band (1900 MHz) fully differential amplifiers (THS4509[50]), which is sent to the primary waveform digitizer, the Domino Ring Sampler (DRS). The second and third types of outputs are used for triggering. The second output is from a low distortion differential ADC driver (AD8138[51]) with a low bandwidth (320 MHz). The third output is an analog sum of four of the input channels. The active splitters also invert the negative PMT pulses into positive signals, which fit in the DRS dynamic range.

PMT signals from TIC bars are split into three parts with a fraction of 8:1:1 by a passive splitter. The largest signal portion is sent to a double threshold discriminator to separate positron hits from noise. It finally goes to the DRS after an active splitter to invert the pulse signs. One of the smaller signal portions is sent to the DRS and the trigger after splitting. The other smaller signal portion is used in current monitoring.

Anode signals of the DCH are divided into two groups with a fraction of 9:1 by a resistive splitter. The larger signal portion is sent to the DRS. The smaller one is amplified to recover the pulse size and summed up over several wires according to trigger logic and eventually sent to the trigger system. The cathode signals, which are only to refine the resolution and hence not used in triggers, are only sent to the DRS.

3.4.2 Trigger System

The trigger system is equipped with flash analog to digital converters (FADC[51]) and field programmable gate arrays (FPGA[52]). Analog signals are digitized by FADCs, the output of which is analyzed by the FPGAs. The system is arranged in a tree structure. It consists of three layers of boards. A schematic layout of the trigger system is shown in Figure 3.33.

In the first layer, boards receive the analog signals and digitize them using 100 Hz FADCs. They also implement a basic reconstruction algorithm using FPGAs. The analyzed information is sent to the second layer of boards, which determine the trigger conditions of each sub-system. The last layer of boards makes the final trigger decision. When a trigger condition occurs, a stop signal is sent to DRS to read out the waveforms. The synchronization between the trigger and DRS is ensured by an ancillary system. It distributes a squared reference clock signal of 19.44 MHz generated by a clock oscillator (SaRonix SEL3935[53]) to all trigger and DRS boards. The jitter of this reference clock is measured to be less than 30 ps.

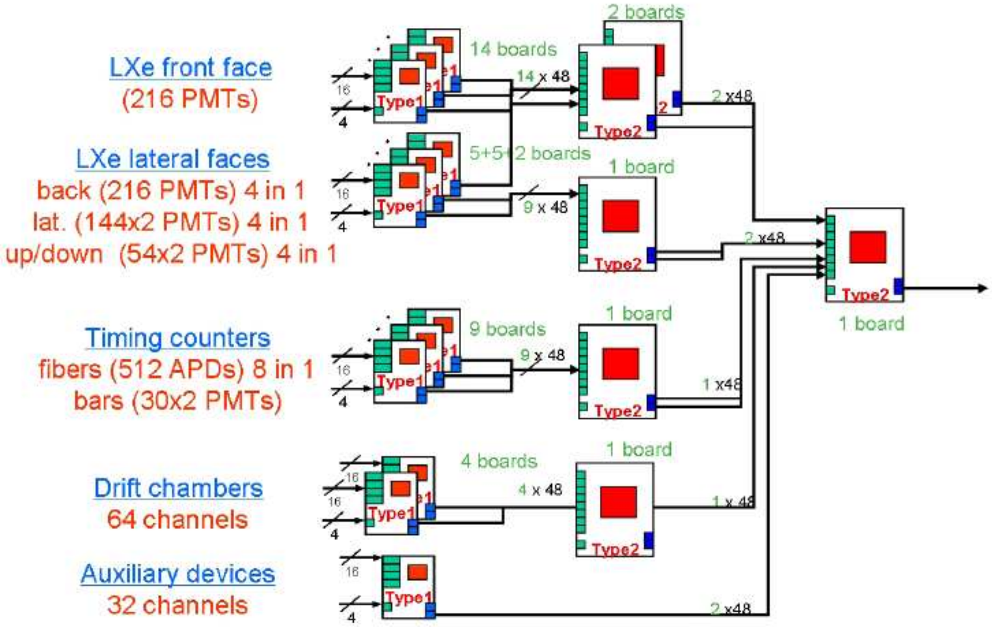


Figure 3.33: A schematic layout of the trigger system

The MEG trigger is set by three conditions: the photon energy, the positron-photon timing coincidence, and positron-photon direction matching. These conditions are analyzed by the fast reconstruction algorithms implemented with the FPGAs. Photon energy is estimated using the sum of all the PMT pulse heights. The threshold is set to be ~ 40 MeV to guarantee high efficiency near the MEG signal energy. The online resolution is measured to be 13.8% at the FWHM. A time window of 20 ns is set on the time difference between the XEC PMT time and the TIC bar hit time to ensure the positron-photon timing coincidence. The time window is set to be big enough to compensate for the spread of the positron time-of-flight and also the poor online timing resolution, which has a σ measured to be 3.4 ns. The position of the inner face XEC PMT which has the maximum pulse is an estimator of the photon direction. At any given positron energy, there is a correlation between the z position of the positron hit on the TIC and the emission angle of the positron from the target.

A lookup table of the XEC inner face PMT number vs. the TIC hit z position mapping the positron-photon direction matching is built by the Monte Carlo simulation. In RUN 2008, due to the unsolved APD problem of the TIC fibers, the TIC hit z position was estimated by the charge ratio of the two PMTs on the TIC bar. The MEG trigger rate is $\sim$ 6.5 Hz at the muon beam rate of $3\times10^7 s^{-1}$.

The radiative decay (RD) trigger only requires sufficient photon energy (to reduce the trigger rate) and positron-photon timing coincidence. They are set the same as their counterparts in the MEG trigger. At the muon beam rate of $3\times10^7 s^{-1}$, the RD trigger saturates. Therefore, the RD trigger is only taken with a reduced beam rate of $1.2\times10^6 s^{-1}$.

Various other kinds of trigger settings are prepared for calibration, normalization, and efficiency calculation purposes, such as self trigger of each sub-system, random trigger, etc. Multiple trigger settings can be used together simultaneously with a prescaling factor on each trigger. Trigger live time as well as total time is measured by the trigger system itself.

3.4.3 Domino Ring Sampler

The DRS[54] is a PSI developed high frequency high resolution analog signal digitizer. Each DRS chip has 10 channels and each channel has 1024 capacitive sampling cells. 8 out of the 10 channels are used as signal inputs, 1 to record the trigger stop signal, and the last to read the common clock signal. The sampling frequency ranges from 0.5 to 4.5 GHz. A simplified schematic of the DRS chip in shown in Figure 3.34 to illustrate its working principle. The input analog signal is stored in a ring of 1024 switched-capacitor arrays (SCA) with the sampling frequency generated by the inverter chain. Once an external trigger is received, the sampled signal is frozen in the SCA and read out in a shift register at a lower frequency (33 MHz) and digitized with high resolution by an external FADC.

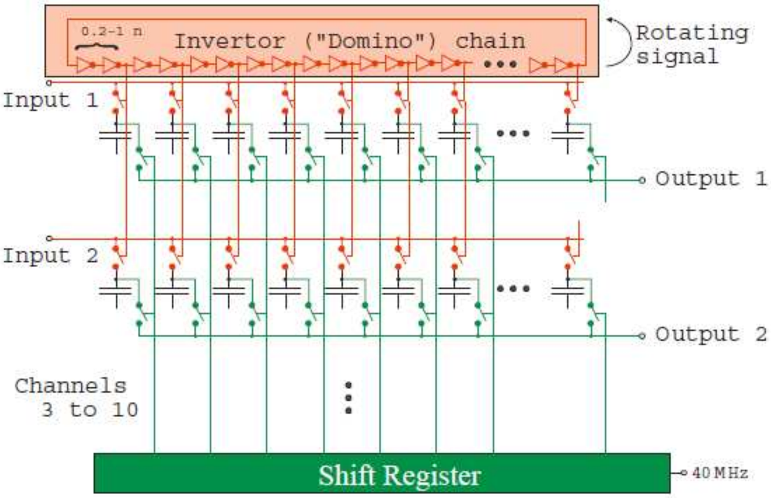


Figure 3.34: A schematic of DRS working principle

A sampling rate of 1.6 GHz is used for the XEC and TIC outputs. The 640 ns window size is large enough to buffer the trigger latency and cable delays. The DCH outputs are sampled with a frequency of 0.5 GHz. This is because high timing resolution is not needed and the DCH waveforms usually extend over several hundred nanoseconds.

3.4.4 Data Acquisition and Management

The DAQ system, consisting of the trigger and DRS, is implemented with a PSI designed Maximum Integration Data Acquisition System (MIDAS[55]) which is controlled and interfaced by Linux PCs. MIDAS is a multi-functional system which builds events, controls front end processes, inputs online running parameters, fires and resets alarms, regulates the slow control systems, monitors and logs the experiment history and provides web interfaces.

The raw data size is typically ~ 5 MB/event, if the waveforms of all channels in all detectors are recorded without any data reduction or compression. Such an event size is very inconvenient. Therefore, at the online data taking stage, some reductions are applied. In case of the DCH and TIC, only a small number of channels have signal hits while most of the channels only contain noise. Therefore, only channels with large enough pulse height and their related channels (for example, a hit cell and its adjacent cells in the DCH) are recorded, while the rest are zero-suppressed. Furthermore, even in the recorded channels, the pulse width is relatively narrower than the full readout time window. Hence, a region of interest is defined around the pulse appearance position of each trigger type (pulse appearance position depends on the trigger type). For the signal from the DCH pads, only information in the region of interest is fully recorded, while outside of this region, waveforms are re-binned. In case of the XEC channels, PMT waveforms are re-binned with wider bin width except around the peak.

At the offline stage, data files are further compressed by a factor of 2 using ZIP compression. The typical final data size is about 1.5 MB/event. In RUN 2008, we took 31 TB data in total.

Chapter 4

Simulation and Analysis Softwares

The MEG experiment uses a known kinematics and signature to search for a rare decay mode. The simulation of the experiment is important in order to validate and understand details of the detector response. In this chapter, we summarize the MEG software structure. The Monte Carlo simulation of the experiment is also covered.

4.1 Software Structure

The MEG software consists of two parts, the online MIDAS experiment data acquisition and the offline simulation and analysis software. The offline part is composed of three packages: megmc, megbartender and meganalyzer. Megmc and megbartender generate simulation data and meganalyzer analyzes both experiment and simulation data. The software structure is shown in Figure 4.1.

Megmc is a Monte Carlo (MC) simulation software based on the GEANT3[56] simulation tool. It processes physics event (a muon decay of a certain type, a elec-

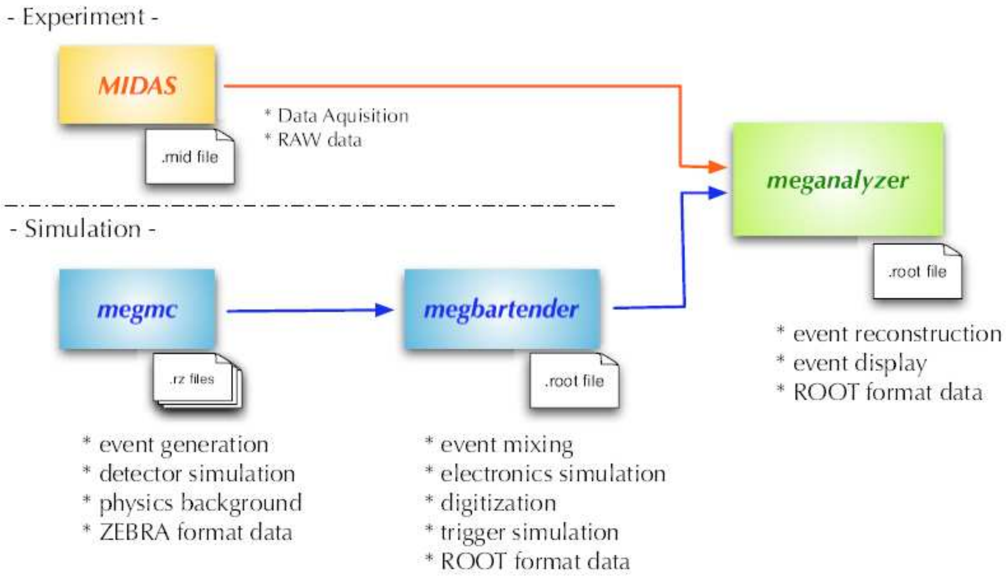


Figure 4.1: Structure of the MEG software

tron/positron pair production, an annihilation in flight, etc.) generation and detector simulation. Table 4.1 shows a list of examples (not all) of physics event types in the MC simulation. The simulated physics event output of megmc is in the ZEBRA[57] format. However, since this is an experiment conducted in a high rate environment, many physics events overlap in the same time window. Therefore, a ROOT[58] based simulation package, megbartender, is developed to serve the purpose of mixing multiple physics events generated by megmc to simulate accidental overlap. Megbartender also handles the electronics simulation, such as waveforms, etc.. It outputs ROOT files.

The event reconstruction and analysis are processed by the ROOT based meganalyzer software package, which takes both the experiment data (MIDAS format) and the simulation data (ROOT format). It, as well as megbartender, is organized by the ROME[59] framework toolkit. Analysis processes and relevant information, such as data, parameters and analysis results, are treated as objects, tasks and folders in

$\mu^+ \rightarrow e^+\gamma$ signal event
Radiative decay (RD) event
Michel decay event
μ^+ beam
e^+ at 52.8 MeV
e^+ from RD decay
e^+ from Michel decay
γ at 52.8 MeV
γ from RD
γ from e^+ annihilation-in-flight
γ from flat spectrum
Cosmic-ray μ
LED light
α from ^{241}Am source on wire
$\pi^0 \rightarrow \gamma\gamma$ decay
photon from $\mathrm{p} + {}^{11}_{5}\mathrm{B} \rightarrow {}^{12}_{6}\mathrm{C} + \gamma$ reaction

Table 4.1: List of examples of physics events implemented in the MC.

ROME. ROME, as an extension of the MIDAS, has features to connect the DAQ system, to access the database, to support the graphical interface, etc.. Therefore, meganalyzer provides an event display as a graphical user interface tool. It gives final reconstructed variables, such as positron momentum, photon energy, positron and photon times, etc. in a ROOT tree.

During data taking, meganalyzer is running online and its graphical event display serves as a monitor to demonstrate the experiment behavior. After each run, an offline meganalyzer process is run automatically to provide detailed information.

4.2 Simulation

4.2.1 Detector Simulation

All detector components, including detailed geometries and materials, are described in megmc in the standard GEANT3 manner. The GEANT particle-tracking system processes the simulation of the particle propagations, which correctly handles particle interactions with relevant materials such as electron/positron pair production, bremsstrahlung, Compton scattering, photo-electric effect, ionization losses, multiple scattering and positron annihilation in flight, etc.. It also traces the secondary particles down to 10 keV. During the process, hit information, such as space-time and energy deposition, is recorded. A brief summary of the simulation of each subdetector is covered in this section.

XEC Scintillation Photons

At each interaction point in the LXe active volume, scintillation photons are generated isotropically. The energy deposition and the scintillation efficiency depends on the particle type during the propagation. Optical processes, such as Rayleigh scattering, scintillation photon absorption, reflection on the walls and PMT windows, transmittance and refraction through the PMT window, and photo-electric effect on the photo-cathode, are simulated. Arrival times of photons at each PMT are recorded.

DCH drift cells

The ionization process inside the drift chamber active-gas region is carried out by the GEANT particle-tracking. The GARFIELD[27] program calculates the isochrone maps of drift cells for the drift electrons. It also provides the time-to-distance relationship and simulates the primary ionization and the electron diffusion effect. The HEED[60] program handles the energy losses of fast charged particles in a gaseous volume, whereas the electron propagation parameters in the He/C_2H_6 gas mixture are taken from the MAGBOLTZ[61] program.

TIC Scintillation Photons

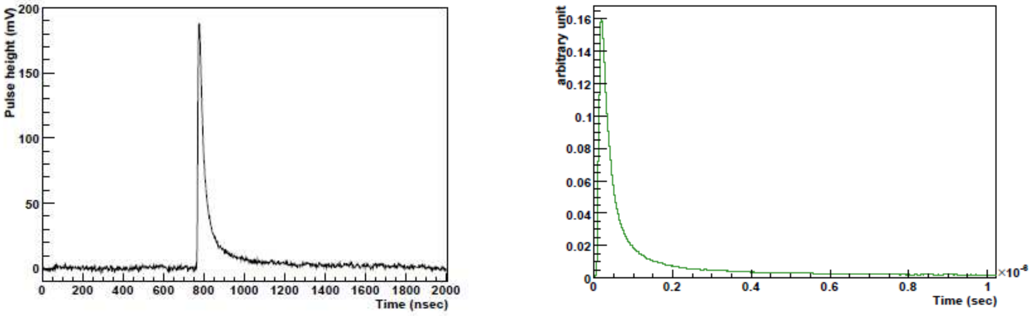
The propagation of scintillation photons in TIC bar is computed analytically taking into account the attenuation in the bar and the reflection on the surface. Poisson statistics is applied when simulating the photo-electric effect of the photo-cathode.

4.2.2 Electronics Simulation

The megbartender program prepares a cocktail of physics events in accordance with the realistic high rate situation and processes the electronics simulation. Electric signals from the detectors pass through various devices and finally reach the DRS. Therefore, the DRS waveforms include all electrical effects from the wires, pre-amplifiers, cables, and digitizers, etc., during the signal propagation. We adopt the impulse-response[62] method to simulate the electronics of each sub-detector.

In the real data, the XEC PMT waveform is a convolution of the photon arrival time in the PMT with an impulse response function. Therefore, the impulse response function used in simulation can be obtained by a deconvolution of the average waveform of the real signal from the MC simulated PMT hit time. The deconvolution was done by a fast Fourier transform technique.

For the DCH impulse response function, we use the 5.4 keV soft X-ray from a ^{54}Mn radioactive source. The 5.4 keV soft X-ray usually excites a single electron avalanche. Therefore, the DRS output of the single electron avalanche characterizes the impulse response function of the DCH system. An example of the anode-waveform output of a 5.4 keV X-ray event and the impulse response function obtained is shown in Figure 4.2.



(a) (b)

Figure 4.2: (a) an example anode waveform output of the 5.4 keV soft X-ray ; (b) the DCH impulse response function obtained by averaging one thousand 5.4 keV soft X-ray anode waveforms.

The TIC bar impulse response function is obtained in a similar way as in case of the XEC. The TIC bar PMT waveform is much simpler than that of the XEC PMT. So instead of deconvolution, we analytically apply a parameterized response function as well as a series of digital filters to represent transit time spread, reflection, etc., on top of the simulated PMT hit time. The parameters of the impulse response function are adjusted such that the output waveform reproduces the real data well.

Baseline noise (electronic noise) is superimposed on the constructed waveforms of each sub-detector. The profile of the baseline noise is extracted from the data. Finally, the waveforms are digitized to simulate the DRS output.

Chapter 5

RUN 2008

The physics run was carried out in 2008 after an engineering run in 2007. It was started in June with a calibration Michel run. A π^0 calibration run was conducted from late July to early September. We started the physics run on September 12 after a short break for trigger setup and background check and took 13 weeks worth of data till December 17. A brief π^0 run was conducted again at the end to close RUN 2008.

In this chapter, we summarize the details of data types and run conditions.

5.1 Runs

Table 5.1 lists all triggers types and requirements used in RUN 2008. Some of those triggers were taken with their own prescaling factors respectively in parallel with the MEG physics data acquisition, while the others were taken alone.

MEG Run The MEG run data taking was mixed with 10 other triggers for calibration, $\mu \rightarrow e\gamma$ decay normalization, detector efficiency calculation purposes. During the MEG data taking, CW calibration runs were taken three times per week, dedicated radiative decay (RD) runs 24 hours per week. Muon beam rate was set to 3.0×10^7 μ/s for the MEG run.

Dedicated RD Run The dedicated RD runs were taken periodically 1 day per week with low beam intensity at 1.2×10^6 μ/s. Besides being a background source for MEG events, RD decay serves as a tool to understand the detector's timing resolution as well as to demonstrate the timing coincidence between positron and photon. It also gives an opportunity to calculate the radiative decay branching fraction in an uncharted high energy region. Triggers 4 and 5 were taken in parallel with triggers 14, 27 and 31 all prescaled by 10^7. Several hundred runs of trigger 2 were also taken at the beginning, which were not used, due to small data size and huge uncertainty in direction-match evaluation.

π^0 Calibration Run Both decay modes of π^0 were used in calibrations. $\pi^0 \rightarrow \gamma\gamma$ is used to determine the energy scale, to patch-scan the XEC in order to check its uniformity, and to understand the XEC time and energy resolutions. The XEC inner face was divided into 12 patches with the dimension of 3 PMTs×3 PMTs. We aligned the NaI to the opposite direction of a certain patch, depending on its position, and took triggers 6 and 7 to collect time-coincident back-to-back events. The Dalitz decay ($\pi^0 \rightarrow \gamma e^+ e^-$), which has a similar decay topology to RD decays, was taken with the RD trigger (4) for the timing resolution and timing coincidence studies.

Periodic Calibration Runs A number of XEC related calibration runs were

#	Name	Prescale Factor	Beam	Description	Settings and Requirements
0	MEG	1	μ^+	$\mu \rightarrow e\gamma$	$(Q_{XEC} > Q_{high}$[1]$) \wedge (\lvert T_{XEC} - T_{TIC}\rvert < 20$ ns $\wedge$ direction-match
1	MEG-LowQ	150	μ^+	$\mu \rightarrow e\gamma$	same as # 0 but with low Q_{XEC} threshold[1]
2	MEG-WideD	500	μ^+	$\mu \rightarrow e\gamma$	same as # 0 but with wide direction-match window
3	MEG-WideT	100	μ^+	$\mu \rightarrow e\gamma$	same as # 0 but with wide time-coincidence window (40 ns)
4	RD	1000	μ^+/π^0	RD/Dalitz	$(Q_{XEC} > Q_{high}) \wedge (\lvert T_{XEC} - T_{TIC}\rvert <)20ns)$
5	RD-WideT	-	μ^+/CW	RD/CW-Boron	same as # 4 but with wide time-coincidence window
6	Pi0	-	π^0	$\pi^0 \rightarrow \gamma\gamma$	$Q_{XEC} \wedge Q_{NaI} \wedge$ time-coincidence b/w LXe and NaI
7	Pi0	-	π^0	$\pi^0 \rightarrow \gamma\gamma$	same as # 6 but with wide time-coincidence window
9	XECHighQ	8000	μ^+/π^0	XEC self	$(Q_{XEC} > Q_{high})$
10	XECLowQ	-	π^0/CW/-	XEC self/CW-Li/alpha	$(Q_{XEC} > Q_{low}$[1]$)$
14	LED	10	-	LED	trigger signal from LED driver
16	Michel	-	μ^+	Michel (DCH+TIC)	3 out of 4 consecutive chambers having outer hits $\wedge$ TIC hits
18	DCH	10^7	μ^+	Michel (DCH self)	3 out of 4 consecutive chambers having outer hits
22	TIC	10^7	μ^+/-	Michel (TIC self)/Cosmic ray	TIC hits
27	XEC-CR	600	-	Cosmic-ray in XEC	$Q_{XEC} > Q_{CR}$[1]
31	Pedestal	20000	μ^+/-	random trigger	-

---

[1] Q_{high} corresponds to $\sim$ 40 MeV in energy threshold, Q_{low} to $\sim$ 30 MeV and Q_{CR} to $\sim$ 65 MeV

Table 5.1: Trigger types and requirements used in RUN 2008

taken periodically three times per week for different calibration and monitoring purposes. It took about three hours per set of runs. It consists of 5 trigger types, in which the LED run is used to calibrate the PMT gains, the alpha run is for PMT QE calibration, the cosmic ray run is to monitor the XEC detector, the CW-Li run is to calibrate the energy scale at the low energy level, and the CW-Boron run is used to understand the timing coincidence between the XEC and TIC.

XEC Self Runs Two types of XEC self runs (9 and 10) were taken. Trigger 9 has the same charge threshold as in MEG and RD triggers, while trigger 10 has a lower threshold. Therefore, trigger 10 data is used to measure the XEC trigger efficiency in MEG and RD triggers (i.e. efficiency of trigger 9). In addition, special trigger 9 data was taken with a π^0 beam and a lead slit mounted in front of the XEC entrance window. By reconstructing the shape of the lead slit, the XEC positron resolution can be measured.

Michel Runs A dedicated Michel run with trigger 16 was taken at the beginning of RUN 2008 to diagnose DCH channels and to calibrate the DCH detector. Two kinds of other Michel triggers, i.e DCH self (18) and TIC self (22) were taken in parallel with the MEG trigger. Trigger 18 is used to calculate the TIC acceptance efficiency and the TIC trigger efficiency.

Pedestal Run Pedestal data of random trigger was taken in parallel with the MEG data. It is used to check the electronics and to evaluate the baseline background.

5.2 Run Conditions

Many hardware problems occurred, especially in the positron detector, during the 5-month time span of RUN 2008. Run conditions of different sub-detectors are described in this section.

DCH Tripping Problem

A severe DCH operational problem was observed in RUN 2008. The nominal working high voltage (HV) for the DCH planes was designed to be 1850 V, which was applied independently on each DCH plane. However, many DCH planes tripped frequently due to discharging under this voltage. Several planes were dead completely and several were set to work under lower voltages. It was understood after a careful analysis following the 2008 operation that the problem was caused by helium penetration into the enclosed volumes between the HV and ground lines. Figure 5.1 shows a cross sectional view of a fraction of the printed circuit board (PCB) of a DCH module and illustrates the DCH tripping problem. DCH HV and ground lines were molded with glues, between which small volumes filled with air were left and kept enclosed. Helium gas slowly permeated into them during the course of time and caused discharging. Since it took time for He to penetrate the glue, the DCH tripping problem also worsened over the course of time. Figure 5.2 shows the history of HV applied on the DCH planes. At the end of the run, only 12 planes out of 32 were still working under the voltage of 1800 V, with 6 additional planes under slightly lower voltages which could have valid hits to be considered operational.

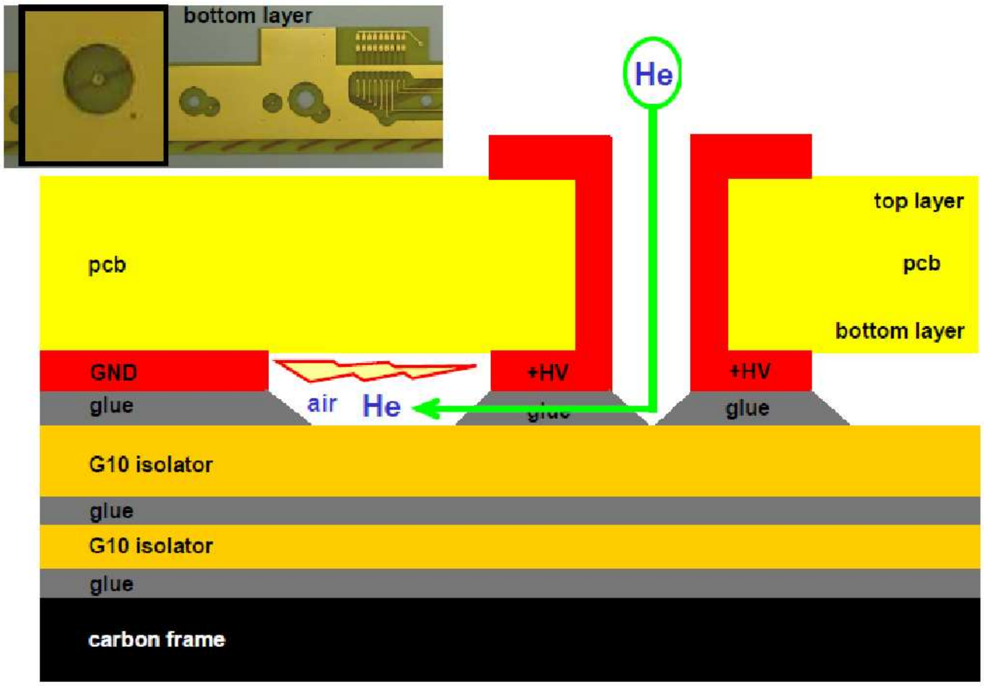


Figure 5.1: DCH tripping problem illustration

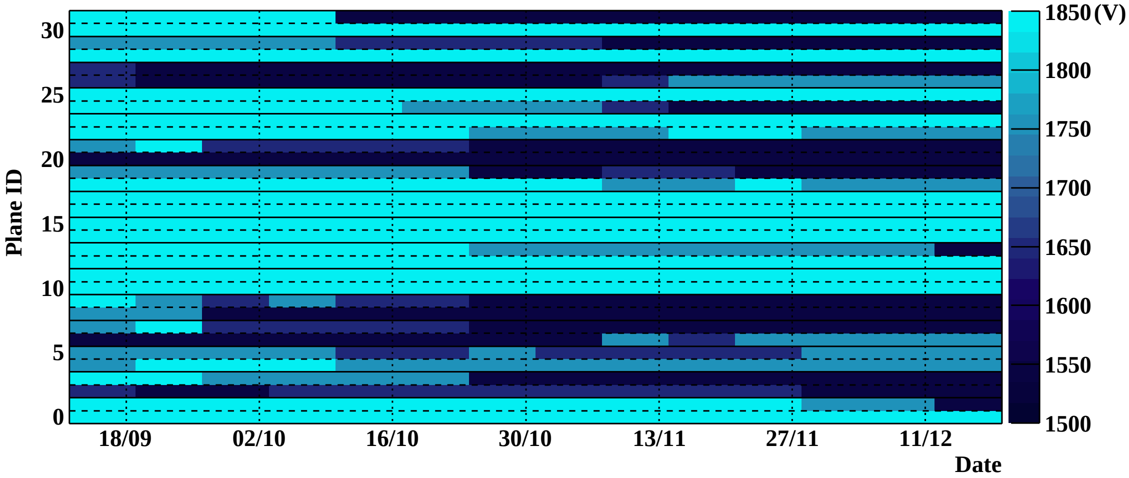


Figure 5.2: High voltage history of DCH planes. DCH tripping problem worsened over the course of time. Nominal working voltage is 1850 V. Voltages above 1700 V could be considered as operational.

COBRA He Concentration

A small amount of air was introduced into the COBRA volume attempting to solve the DCH tripping problem. The COBRA He concentration was kept around 94% on average. However, the beam tuning was done before the introduction of air, and was not corrected to account for the effect of the additional air.

TIC Fibers APD Problem

The TIC fibers (or the z-counter) were designed to have a fast z determination of the positron hit to be used in the trigger decision. However, we encountered an APD output problem that the analog output had a noise level that was too high to be useful. Consequently, the z-counter was not used in the run. As for the trigger decision, the z information was estimated by the charge ratio of the PMT readouts of the two ends of the TIC bars.

Bad Channels of Sub-detectors

Besides the dead planes, there were bad channels discovered on the operational DCH planes. Out of 1728 channels (288 wires × 6 channels/wire), 17 dead channels were found. They were mainly due to bad connections inside the closed COBRA volume where chambers were placed and, therefore, could not be recovered during the run. A list of dead channels is shown in Table 5.2 labeled in the universal wire number (see Figure 6.3). A dead anode connection loses the hit information entirely, whereas a dead cathode pad connection still has a good r determination and a rough z estimation from the charge division on the two ends of the wire. There were more channels with subtle problems such as low charges, high noise levels, strange pulse shapes, etc., which have yet to be understood.

Description	wire #
Dead Anode	60, 260, 278
Dead Cathode Pad	17, 22, 44, 45, 48, 67, 162, 163, 226, 242, 252, 254, 272, 274

Table 5.2: Dead DCH channel list in RUN 2008

One bad channel was found in the TIC bars, which was fixed at the end of October and worked properly afterward.

Seven out of 846 PMT channels were found to be bad in the XEC, out of which three were dead, one had strange pulse shape and three showed unstable PMT gains. Nonetheless, all front face PMTs were working properly. Therefore, no efficiency loss was suffered.

Chapter 6

Event Reconstruction

6.1 Overview

The event reconstruction algorithms are developed and optimized in accordance with the electronics characteristics and the detector geometries, performances, and constraints. The reconstruction starts with the waveform analysis, in which basic physical quantities, such as arrival times and output charges, are obtained. Detector dependent algorithms further analyze the time and charge to deduce particle kinematic variables, such as incident time, momentum, energy, position, emission angle, etc.. At the event reconstruction stage, positrons and photons are treated separately.

The positron kinematic variables are obtained by analyzing data from two subdetectors, the DCH and TIC. The positron analysis structure is illustrated in Figure 6.1. The DCH analysis reconstructs hits from the waveforms, groups relevant adjacent hits into clusters, further associates clusters to form possible track candidates, and finally fits the tracks of positron trajectories. Thus, it calculates positron momentum. By projecting the track to the muon stopping target, we obtain the decay vertex point and the positron emission angle. For the TIC, we analyze the PMT waveform to reconstruct hits, and further form clusters from the hits. The TIC information provides a precise positron hit time. By matching the DCH track with the TIC time, we can calculate the positron time-of-flight and deduce the muon decay time at the vertex.

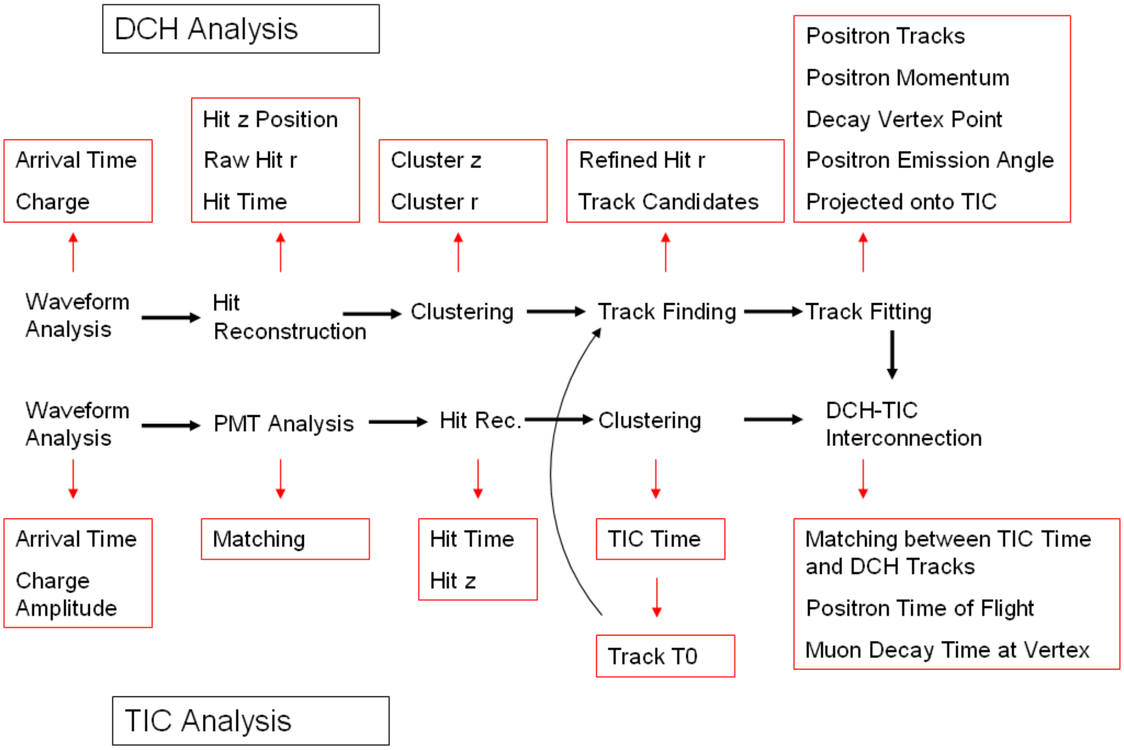


Figure 6.1: Positron analysis structure

As for the photon, kinematic variables such as energy deposition in the XEC, the photon interaction point and time are obtained from the XEC analysis. Figure 6.2 shows the flow of the XEC analysis chain.

In this chapter, we describe the details of the reconstruction algorithms. In real-

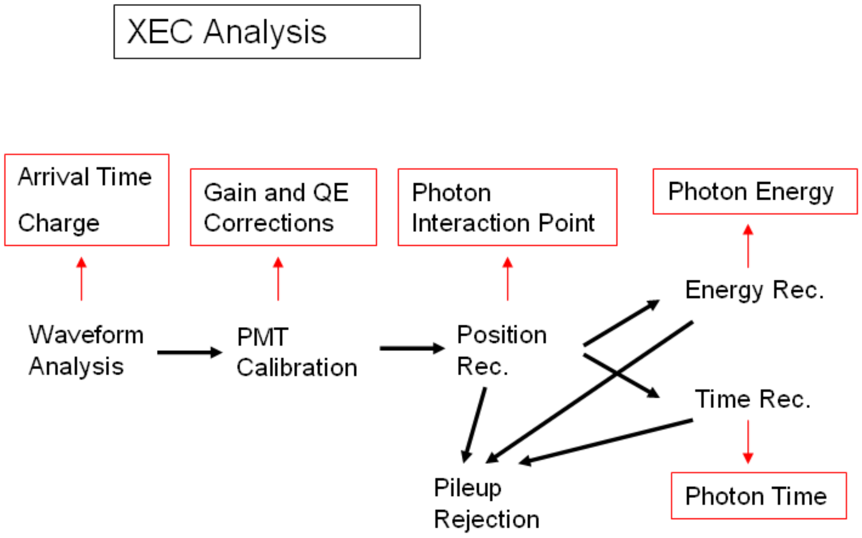


Figure 6.2: Photon analysis structure

ity, reconstruction algorithms and calibration techniques are co-developed and hence inseparable. Positron and photon information are further combined to serve the advanced physical analysis. Both calibration techniques and physical analysis are covered in later chapters.

6.2 DCH Analysis

The DCH analysis has been divided into a number of tasks that together allow the full analysis of this detector system. Interface between tasks is provided through ROOT folders[58], which allows any of the tasks to be replaced with a different code so long as the folder variables are consistent.

We define the DCH numbering scheme here, as shown in Figure 6.3. We adopt a cross sectional view against the beam direction. 16 DCH modules (or simply chambers) are labeled from 0 to 15 from right most to the left most, which is also in the direction of positron flight. For each chamber, the left plane is numbered as plane 0 and the right one as plane 1. Note that, due to historical reasons, plane number is in fact the opposite of the overall DCH numbering convention. 288 wires or cells are labeled from 0 to 287, which we call the universal wire or cell numbers. The wire numbering starts from the innermost cell of the right most plane, first counts radially outward on the same plane and then move to the adjacent left plane and keeps counting in the same fashion.

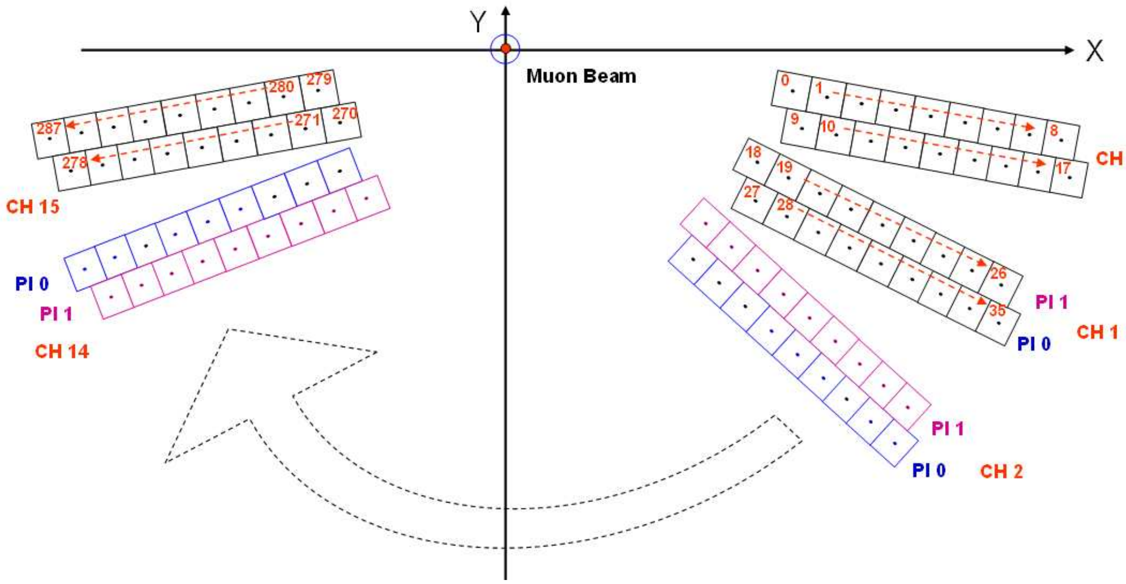


Figure 6.3: The DCH numbering scheme (not in exact scale and for illustration purpose only). The muon beam is coming out of the plane towards the direction of positive z.

6.2.1 Waveform Analysis

We have six waveforms for each DCH cell: two ends of the anode wire and four cathode pads (the inner two are simply called cathodes, and the outer two are called hoods), as shown in Figure 3.17. They are recorded on DRS digitizers with a nominal bin size of 2 ns. The DRS bin-to-bin voltage gain is already corrected at the online stage. We rotate the waveforms to a common start time so that waveforms from a particle that causes the trigger will appear in the center of the DRS window. The waveform provides vectors of charge and time for each DRS bin. An example of the six waveforms associated with one cell is shown in Figure 6.4.

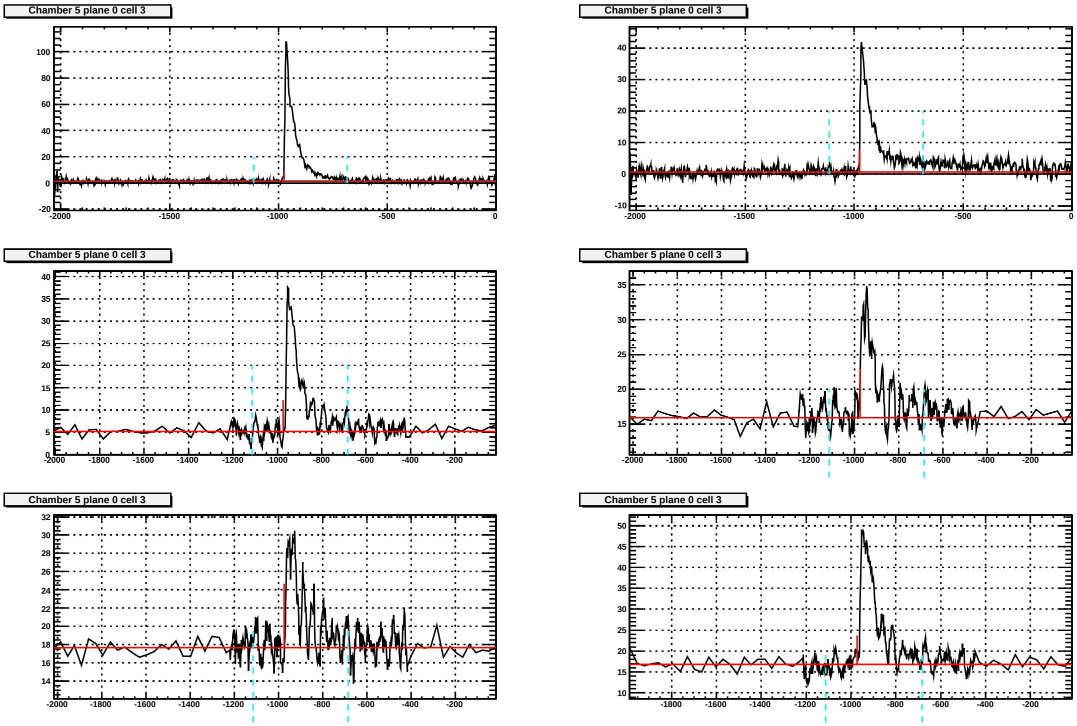


Figure 6.4: An example of the six waveforms associated with one cell. Two ends of the anode wire are displayed on the first row; cathodes are on the second row and hoods on the third. Vertical axises have the unit of mV and horizontal axises are in the unit of ns. The horizontal red lines indicate the averaged baseline levels. The vertical red lines mark the pulse leading edge times. The region between the dashed cyan lines is the expected region where signals appear with respect to the trigger. The pad waveforms are re-binned outside of the region of interest, which is set bigger than the expected signal region.

6.2.2 Hit Reconstruction

The hit reconstruction algorithm finds hits, defined as the information on a single cell associated with the passage of a single charged particle through that cell. It takes the DRS signal with time and charge as the input, from which the hit finding can deduce the primary information, such as the total charge and leading edge time of the pulse in the two ends of the anode, the best time associated with the hit, the inferred Z coordinate from the anode signals, and the inferred Z from the pad signals.

The algorithm proceeds by first expanding the pad waveforms into nearly constant width bins, since some bins of the waveforms away from the pulses have been rebinned at the online stage to reduce the data size. The algorithm then calculates the baseline for each of the six signals on each cell. The baseline is assumed to be constant. The algorithm histograms the voltage in all time bins before the nominal signal region (set to be the first 400 bins) for a given waveform, finds the voltage corresponding to the peak value of this histogram, then averages all voltages within a ±5 mV range about the voltage of the peak bin.

We measure the precision of the predicted baseline in the signal region by selecting events without a signal and comparing the measured baseline from the first 400 bins with the average value of the waveform in the signal region. Figure 6.5 shows the distribution in the difference of the predicted baseline and the actual mean baseline in the signal region (in this particular example, the signal region has a width of 40 ns). This shows that the typical systematic shift in the predicted baseline is -0.17 mV, with an RMS error of 0.7 mV. The negative mean value means that the baseline in the signal region is typically slightly less than the prediction. This is because the DRS is a rotating digitizer, thus earlier hit pulses not having fully decayed give a value higher than the actual baseline. This observed effect is not corrected in the analysis.

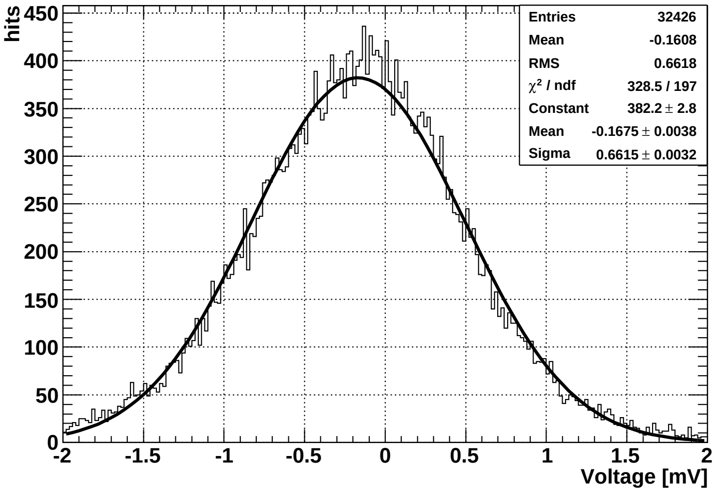
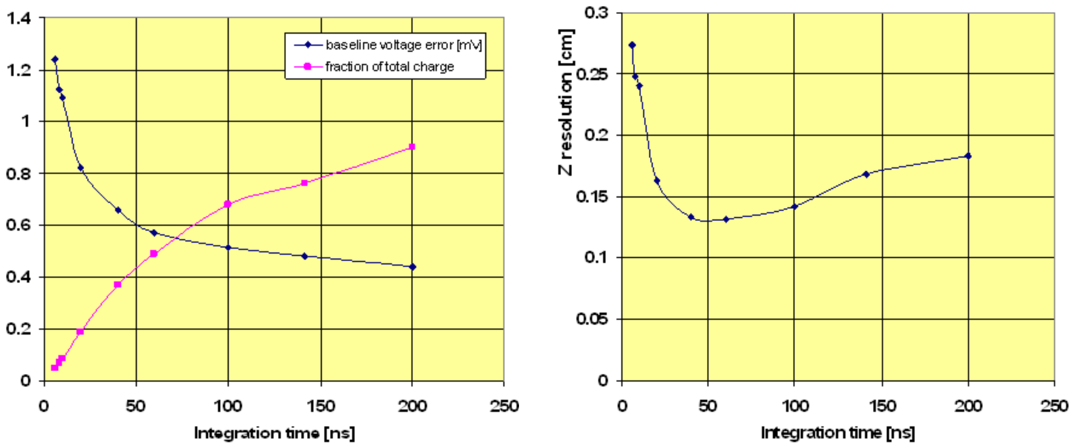


Figure 6.5: Distribution in the difference of the predicted baseline and the actual baseline in the signal region with a width of 40 ns.

The error in the hit charge due to a baseline error is proportional to the baseline voltage error times the signal window width. The optimal signal window width is the one that minimizes the total error in Z, accounting for the error in the baseline and also for fluctuations in the fraction of the charge that is in the signal window width (see Figure 6.6(a)). This has been studied by varying the signal region width and plotting the charge error due to baseline shift error divided by the mean fraction of the signal in the chosen width. Figure 6.6(b) shows this ratio as a function of the signal window width. A broad minimum around 50 ns is apparent, which is, therefore, set as the optimal integration time.

Having found the best estimate of the baseline voltage, it is subtracted from the waveform, and the baseline corrected waveform is then used to find hits. In order to



(a) (b)

Figure 6.6: Plots of finding the optimal signal window width. (a) baseline voltage and fraction of integration charge vs. integration time window; (b) contribution of noise in the anode charge to the Z uncertainty (averaging over all positions) vs. the integration time. Z information is obtained by the charge ratio of two anode wire ends.

reduce the high-frequency noise in the waveform, the hit finding algorithm smooths the waveform by constructing a running average of the bin voltages. The algorithm then finds the bin in the smoothed waveform with the largest voltage above 5 mV. Having found the peak, the algorithm finds the time and charge of the hit on each end of the wire and finds the charges on the pads in that cell. The charge integration limits are set by the times that the signal waveform of the larger peak of the two anodes is above two times its pedestal RMS sigma. There is also a minimum integration time window of 5 ns before and 10 ns after the peak.

After integrating charges on both ends of the anode wire and the pads, the algorithm then first estimates the Z position by charge division of the two ends of the anode and further refines Z by the charge asymmetry on the Vernier patterned pads. For the anode wire, we define the charge division or charge asymmetry as

$$\epsilon_a = \frac{Q_d - Q_u}{Q_d + Q_u} \tag{6.1}$$

where $Q_{u,d}$ are the measured charge at the upstream and downstream ends of the wire, respectively. The cathode and hood charge asymmetries are also defined in the same fashion. The anode-determined Z can be expressed as

$$z_{anode} = \left(\frac{L}{2} + \frac{R}{\rho}\right)\epsilon_a \tag{6.2}$$

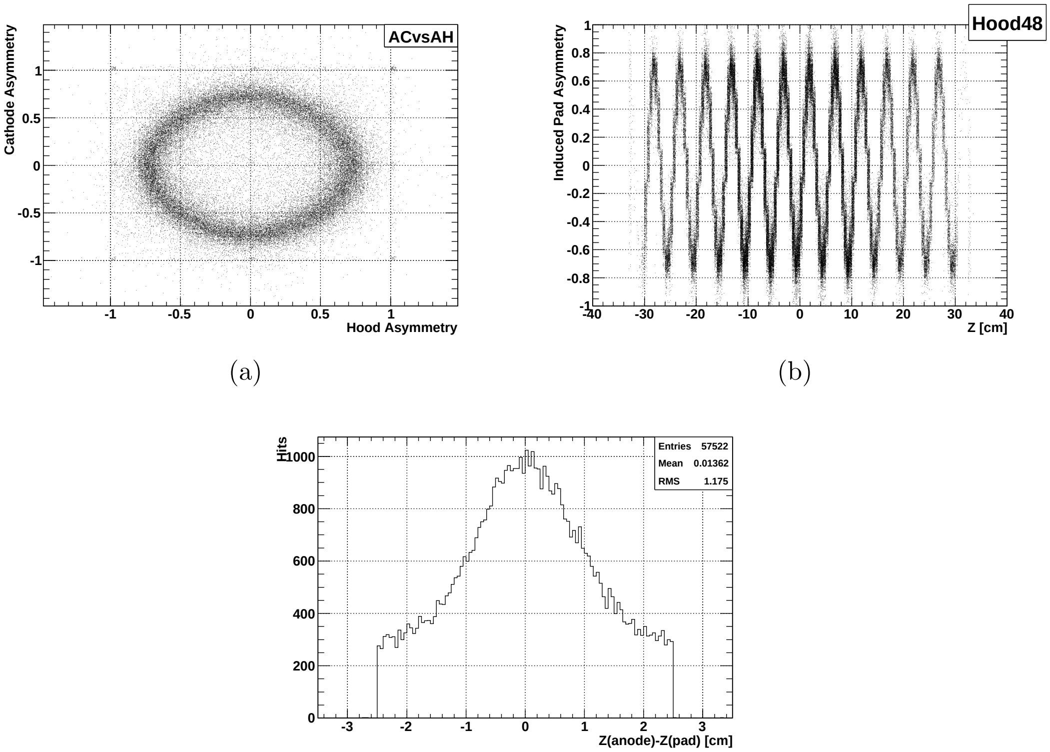
where L is the resistive wire length, R is the preamp input impedance, assuming the same for both upstream and downstream ends, and ρ is the wire resistance per unit length. The factor L_{eff}, defined as $L_{eff} \equiv \left(\frac{L}{2} + \frac{R}{\rho}\right)$, is also referred to cias the effective wire length.

We then refine Z more precisely using the Vernier pattern. The charge asymmetry for the vernier pads are similarly defined for the cathode and the hood. Then the distribution of cathode charge asymmetry vs. hood charge asymmetry forms a circular pattern due to the 5 cm Vernier periodicity. We call each period a pad cycle. Figures 6.7(a) and 6.7(b) show the Vernier circle and the pad charge asymmetry periodicity of 5 cm in anode Z. From these figures, we can write the pad refined z_{pad} as,

$$z_{pad} = \arctan\left(\frac{\epsilon_{hood}}{\epsilon_{cathode}}\right) \cdot \frac{5cm}{2\pi} + \delta_{pad} \tag{6.3}$$

where δ_{pad} is an offset term that only depends on which pad cycle z_{anode} is in.

The difference of the anode determined Z and the pad refined Z is shown in Figure 6.7(c).



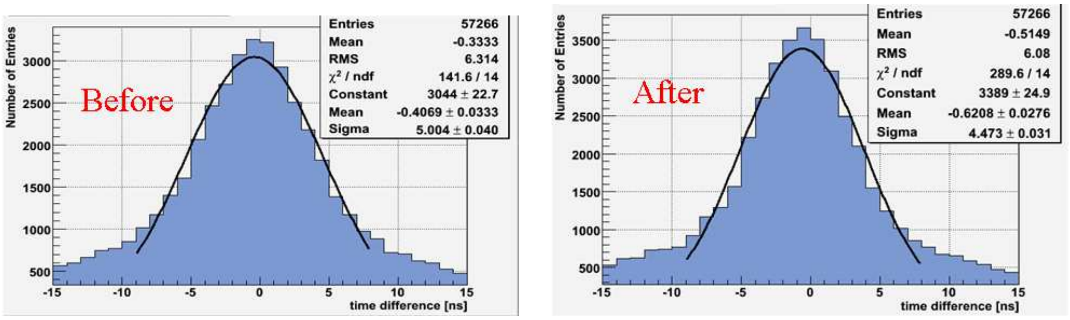
(a)

(b)

(c)

Figure 6.7: The DCH Z position determination. (a) circle pattern formed by cathode charge asymmetry vs. hood charge asymmetry due to the Vernier periodicity; (b) an example of the pad charge asymmetry vs. anode Z, which shows the pad charge asymmetry periodicity of 5 cm; (c) difference of the anode determined Z and the pad refined Z

The leading edge time at each end is then found by moving to earlier time from the time of the voltage peak, and taking the leading edge time to be that of the earliest bin that has a voltage at least 3 times the RMS noise level on that event and wire end. Figure 6.8 shows the distributions in the difference of the leading edge time for the two wire ends, before and after the correction for signal propagation delay, which is the travel time of the signal to the preamplifiers located at the end of the wires. Propagation delay is corrected by subtracting the travel time from the leading edge time based on the Z location of the hit and a travel speed of 30 cm/ns. The width of the distribution after the signal propagation delay correction has contributions from uncorrelated noise at the two ends and effects of the algorithm used to find the leading edge.



(a) (b)

Figure 6.8: Correction of signal propagation delay. (a) difference of the leading edge time for the two wire ends, dt, before the signal propagation delay correction; (b) dt after the signal propagation delay correction. We fit a Gaussian in the core part of each distribution. The RMS of dt improves from 5.0 ns to 4.5 ns after the signal propagation delay correction.

The drift time determination of a hit requires knowledge of the track time. Therefore, it can only be obtained in the tracking stage. At this level, the R position of a hit is estimated by the wire position. It will be further refined in the tracking algorithm upon knowing the incident angle and the drift time to resolve the left-right ambiguity[1].

---

[1] For a given drift time and the angle that a particle passes through a cell, there is still a two-fold ambiguity of the hit position, which is illustrated in Figure 6.11

6.2.3 Cluster Finding

The cluster finding algorithm operates on hits, obtained in hit finding, and creates groups of hits called clusters. A cluster is a collection of hits (typically a minimum of two, although single hit clusters can be used) in a single chamber with a pattern that is consistent with the hits being made by a single charged particle traversing the chambers.

The algorithm starts clustering hits by R adjacency, allowing a gap of no more than 3 radially adjacent cells. The next step is to split clusters if the hits are at very different Z. The average Z of the hits is calculated. If a hit in the cluster deviates by more than 3.75 cm from the average, it is removed from the cluster and set aside for future processing. The value 3.75 cm is set by the requirement that a Z coordinate that is mis-measured due to finding the wrong pad cycle will not be removed (this is explained further in the description of cluster cloning). The average is recalculated and this process iterated until no hit fails this requirement. At this point, the cluster is again checked for the requirement of a maximum radial gap between successive hits. If an individual hit fails the check in R, it is removed from the current cluster and set aside for further processing. The hits removed from the original cluster may qualify to be reassembled into a cluster of their own. This check for sub-clusters continues until all removed hits are either part of a cluster or have been deemed inadequate for any cluster. This removal and reassembling process may proceed for many iterations before all individual hits find a home as part of a cluster of at least 2 hits or become an isolated hit.

All remaining hits that do not belong to any cluster are defined as single hit clusters.

An example of clustering in a real event is shown in Figure 6.9. Hits of the same color in each chamber belong to the same cluster.

Due to many chambers operating at low voltage, the anode determined Z has a poor resolution and may end up in a wrong pad cycle inconsistent with its actual position. Therefore, a 'cluster cloning' technique is developed to correct for hits that have Z coordinates wrong by 5 cm because of an incorrect determination of the cathode pad cycle. If a cluster spreads by more than 2.5 cm in Z, the algorithm attempts to extract the correct pad cycle by creating artificial clusters in anticipation that the 'true' cluster will later be picked up at the tracking stage. In case of a 2-hit cluster, 2 clones are created by increasing (decreasing) Z of the smaller (bigger) Z hit. For 3-hit clusters, the one hit that is not in the same pad cycle as the remaining 2 hits is simply shifted by 5.0 cm accordingly. For clusters having more than 3 hits, the correct pad cycle is determined by the 'majority rule'; and in case of an equal weight scenario, for example, 2+2 in 4-hit clusters, two clones are created. Figure 6.10(a) demonstrates the effect of the cluster cloning and Figure 6.10(b) shows the number of hits in clusters.

The R and Z coordinates of a cluster are simply the average of all hits in the cluster.

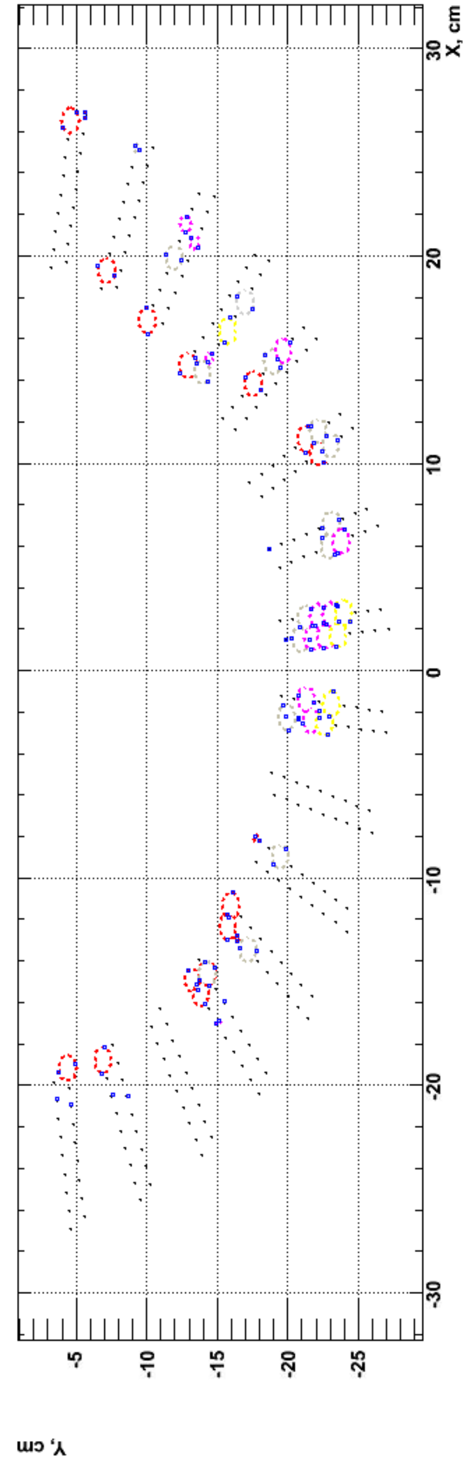
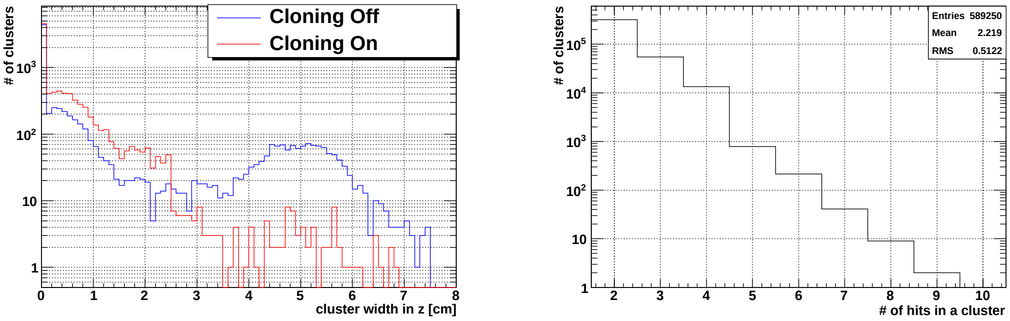


Figure 6.9: An example of clustering in a real event.



(a) (b)

Figure 6.10: Plots of cluster cloning. (a) effect of cluster cloning; (b) number of hits in clusters.

6.2.4 Track Finding

The track finding algorithm exploits the property that particles of interest will have clusters at large radii and the property that the hit rates are significantly lower at large radii. It starts the track-finding at large radius (greater than 24 cm) by finding a track seed with clusters in three chambers, then extending the seed to find track candidates. Having found clusters consistent with a particle track, it then resolves hit left-right ambiguities (explained and illustrated in Figure 6.11), refines the measured track positions and calculates the track time.

Seed Finding

A seed is defined as the outer most three clusters of a track. The seeding algorithm first looks for a cluster beyond a minimum radius of 24 cm on a non-edge chamber (not 0 or 15), which is then considered as a candidate for a center cluster of a seed and hence the cluster having the largest R coordinate in the track to be reconstructed. On each side of such a cluster, the algorithm requires a nearby cluster satisfying some criteria on the difference in R and Z values and with at most 1 skipped chamber

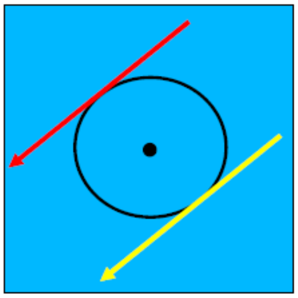


Figure 6.11: Two-fold left-right ambiguity of the hit position given a track incident angle. At the hit reconstruction stage, only a leading edge time is obtained. The drift time is obtained by the leading edge time subtracted by the track time. Providing the drift time, track incident angle and local B-field, the TXY relationship returns a hit position, which still has a two-fold ambiguity to be resolved.

between them. The R coordinate of the side cluster must be smaller than that of the center one, and the difference must be no more than 2 cm. The difference in the Z coordinate of an added cluster from the center one must be in a range that has a full width depending on whether the new cluster is in an adjacent chamber or the next adjacent one. The full width is 4 cm for adjacent chambers, 8 cm if one is skipped. A seed must contain at least one multi-hit cluster. A triplet of clusters of this kind qualifies as a seed.

Once seeds are found, an initial estimate of the track time, which we call the track T_0, is computed. A technique is developed to find the track T_0, to resolve the hit left-right ambiguity, and to refine the cluster position. A rough guess of the track T_0 is simply the earliest hit time of all the hits in the seed clusters. Starting from this initial guess a surrounding 120 ns range of time is checked. We first calculate a circle that goes through the original three clusters. The track incident angle at each cluster is estimated by this circle. For this incident angle and a given trial time, each hit on a cluster has two possible left-right solutions. Each combination of the left-right solutions of hits in one cluster gives an average R and an RMS deviation of hits from this average R. The R-error-squared is defined as the square of the total R RMS of hits in the cluster. The position that gives the minimum RMS deviation out of all combinations is the refined cluster position, and the left-right ambiguity is therefore resolved. To ensure the correct assignment of left-right solutions, we only resolve the left-right ambiguity if the hit is sufficiently far away from the wire ($> 500\mu m$) and leave the hits close to the wires to the track fit algorithm. The trial time that minimizes the sum of the R-error-squared contributions of all clusters is the track time. Figure 6.12 shows an illustration of this technique. Single hit clusters are excluded in using this technique.

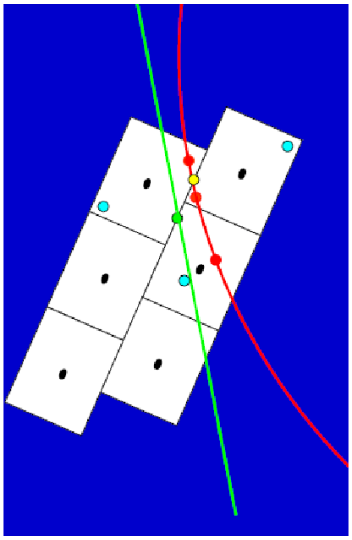
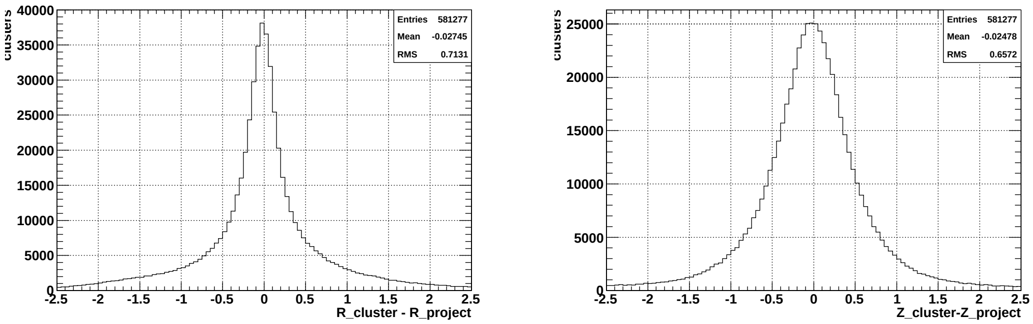


Figure 6.12: Colored coded illustration of the technique that does track time finding, hit left-right ambiguity resolving, and cluster position refining all together. As a first approximation, the track goes through the green point, which is the first estimate of the cluster position by averaging R positions of all hits (in this case, all 3 hits) in the cluster. The track incident angle is estimated by the three consecutive clusters that form the track circle. Red and cyan points are the left-right solutions of the hits for this incident angle and a given trial track time. Each combination of 3 hit positions gives an average R and an RMS deviation of these three hits from this average R. The position that gives the minimum deviation out of all combinations is the refined cluster position (yellow point) and that combination of the left-right solutions resolves the ambiguity (right points). The trial time that minimizes the sum of the R-error-squared contributions of all clusters is the track time.

Track Extension

After finding the seeds, the algorithm attempts to extend each seed on both sides, adding clusters on other chambers that are consistent with a track. In each direction, it adds clusters to the track until either no more clusters are able to pass the selection criteria or the last chamber is reached. Starting from a cluster on the edge of a seed, a projection in R and Z is made to the adjacent chamber. The seed gives a measurement of the instantaneous track circle, which in turn allows for a measurement of transverse momentum. As the B-field changes from one chamber to the next, so does the radius of the track circle. Knowledge of the transverse momentum as well as the $(X,\, Y,\, Z)$ displacement of the track through the seed allows for an estimate of the total track momentum, which should be constant up to multiple scattering effects. Coupling this knowledge of the invariant total track momentum with the adiabatic invariant P_T^2/B_Z for slowly varying axial magnetic fields, an estimate of the Z intersection is made. A cluster to be added must be located within the range of ±1.5 cm in R and ±2.5 cm in Z from the projected R and Z coordinates. Distributions of the the errors in R and Z projection are shown in Figure 6.13.



(a) (b)

Figure 6.13: (a) distribution of the error in R projection; (b) distribution of the error in Z projection.

Once a cluster passes these checks, it is then used to calculate an updated track circle and track angle. For the cluster to be added to the track the updated track circle must curve in the correct direction (center towards the beam axis) and the track angle at this chamber must be in the allowable range. The newly added cluster's position is then refined using the same technique as mentioned previously, the track time is re-calculated and the left-right ambiguities of the hits on this cluster are resolved.

If more than one cluster on the same chamber satisfies the track projection criteria, duplicate track candidates are made and each is continued along the chamber separately. The skipping of one total chamber per direction is allowed in tracking in addition to any skipped chambers from seeding. Lastly, there is also the possibility that the projection falls outside the range of the DCH entirely, which halts the tracking in one direction only if the deviation goes beyond an allowable error of 2 cm.

Figure 6.14 shows an example of track finding in a real event. Clusters belonging to the same track candidate are connected by the line segments.

At the end of the tracking stage, all track candidates have a DCH self determined track T_0. Figure 6.15(a) shows the distribution of the DCH self determined track T_0. All hits but the ones belonging to single-hit clusters have their left-right ambiguities resolved and all cluster positions refined by the track T_0. The momentum of each track candidate is also calculated. It is done by averaging the momenta calculated for each three consecutive clusters on the track, using the refined cluster positions due to the track time and final left-right solutions. The distribution of the track candidate momentum from Michel run is shown in Figure 6.15(b).

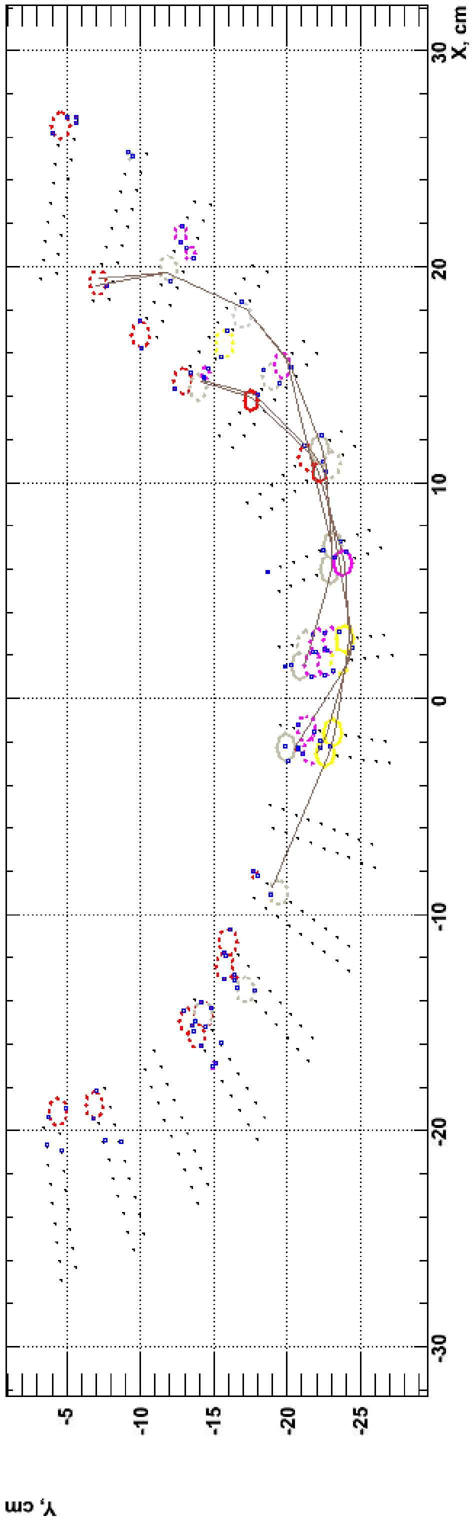
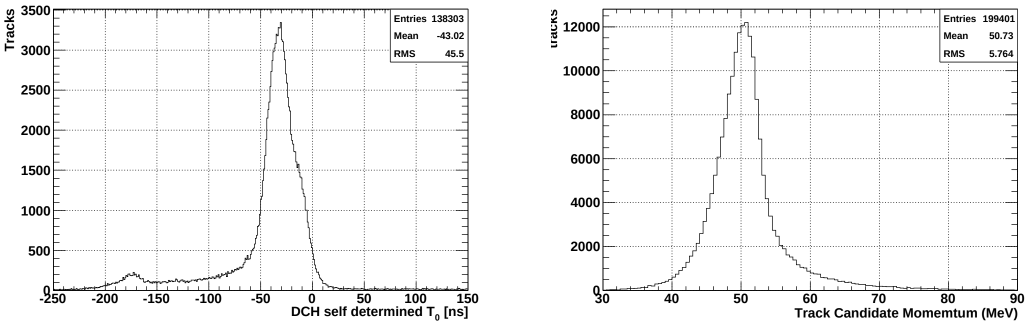


Figure 6.14: An example of tracking in a real event. Tracks are split if more than one cluster on the same chamber satisfies the track projection criteria during the track extension.



(a) (b)

Figure 6.15: (a) distribution of the DCH self determined T_0; (b) distribution of the track candidate momentum.

6.2.5 TIC Determined Track Time, T_0^e

The track finding algorithm is a self-sufficient one that, without any external information, determines a track time, T_0, from the hits, resolves the left-right ambiguities of the hits, and improves cluster positions.

The TIC hit provides a much better time measurement, if matched with a track, which will eventually be used for positron timing. The DCH track and TIC hit matching, however, can only be done after extrapolating the fitted track to the TIC surface. Nonetheless, we attempt to use the better timing information in the track finding by trying to assign every TIC time to every track seed. So after finding a seed, we make as many copies of the seed as the number of TIC clusters, $N_{TIC\ Cluster}$, in the event. For each one of the $N_{TIC\ Cluster}$ copies, we re-run the tracking algorithm providing the TIC determined track time of each TIC cluster, T_e^0, obtained by the TIC time subtracted by the overall DCH-TIC time offset, until all TIC clusters are used, anticipating that the track candidates developed with invalid TIC time assignments will either fail in track fitting or be rejected by the DCH-TIC interconnection criteria. The DCH-TIC time offset is extracted by finding the peak time of the distribution of the time difference between the TIC hit time and DCH self determined T_0 (see

Figure 10.9[2]). From the figure, we know that, using the TIC time as a reference, the DCH self determined T_0 has a precision of 6.4 ns. The DCH self determined T_0 is not great because there are not many hits in tracks. This is the purpose of using TIC determined track time T_0^e.

Using this assigned TIC determined track time, T_0^e, the left-right ambiguities of the hits can still be resolved and the cluster positions improved by the same technique.

At the end, for each unique track candidate, there are $N_{TIC\ Cluster}+1$ copies, one having the DCH self determined T_0, and the rest are TIC determined by the different TIC times. All track candidates are treated equally and sent to the track fitting for a further filtering process.

6.2.6 Track Fitting

After finding track candidates, the last step of reconstruction is to fit the tracks. The track fitting precisely analyzes the positron trajectory and finds the positron state vector at any given point along the track. The state vector, consisting of five parameters, i.e. two for the position, two for the direction and one for the momentum, fully and uniquely describes the track of a charged particle in a magnetic field at a given point. The Kalman filter[63] technique is used in the track fitting. A Kalman filter is a method of estimating the state of dynamic systems. It produces estimates of the true values of measurements by predicting a value, estimating the uncertainty of the predicted value, and computing a weighted average of the predicted value and the measured value. The smaller the uncertainty of a predicted value is, the more

---

[2] This plot is used to correct an error in the normalization sample for the radiative decay branching fraction measurement. Therefore, it is shown there.

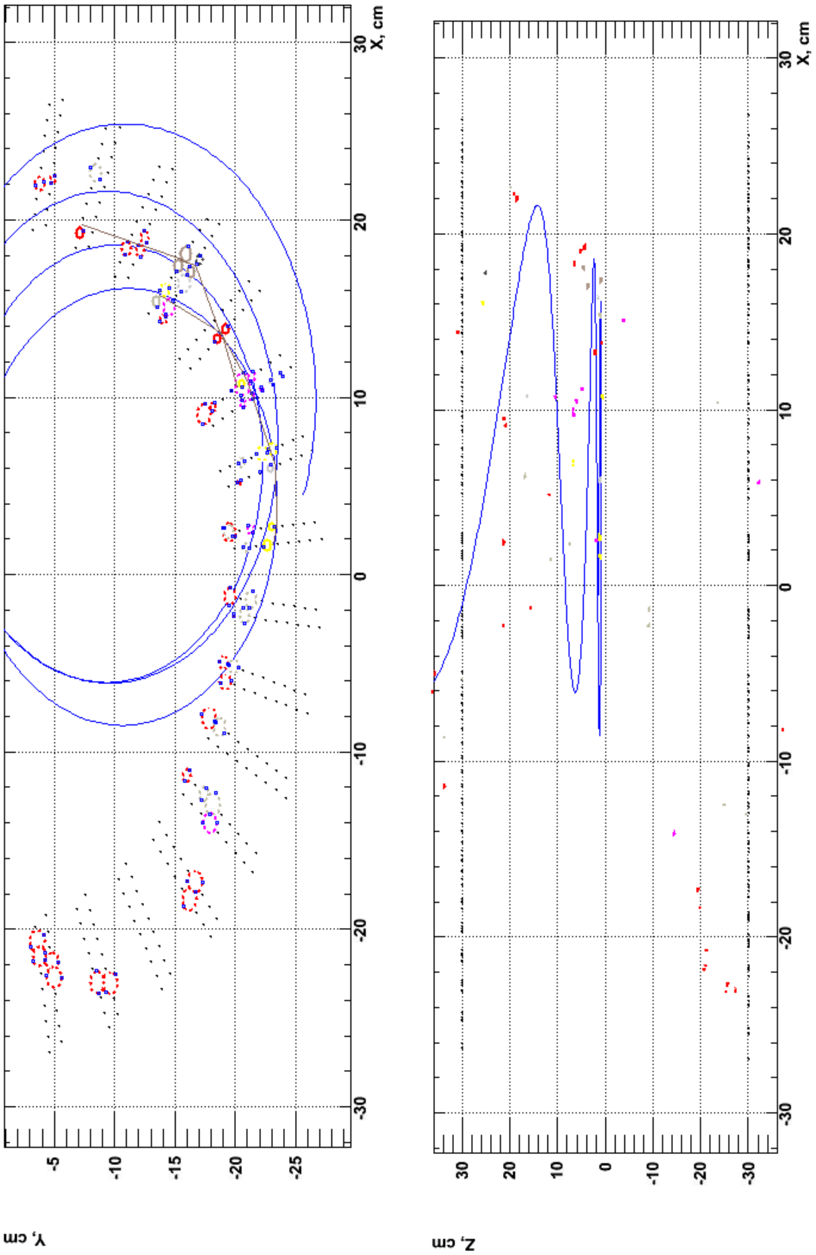
weight the method puts into it. Therefore, the estimates produced by the filter tend to be closer to the true values than the original measurements because the weighted average has a better estimated uncertainty than either of the values that went into the weighted average. In our application, the Kalman filter correctly handles the multiple scattering effects and energy loss of the charged particle and provides control of error propagation.

The state of a charged particle moving in the magnetic field can be described by the track parameters, such as position and momentum. The Kalman filter fits the tracks in the following steps:

1. **Prediction**: $\hat{\mathbf{x}}_{k|k-1}$ ($\mathbf{P}_{k|k-1}$) is the predicted value of the state vector $\hat{\mathbf{x}}$ (covariant error matrix $\mathbf{P}$) at the 'present' k-th point using the information up to and including the previous $(k-1)$-th point.
2. **Filtering**: $\hat{\mathbf{x}}_{k|k}$ ($\mathbf{P}_{k|k}$) is the estimate of the state vector (covariant error matrix) at the 'present' k-th point by taking a weighted average of the predicted value and the measured value of this point.
3. **Smoothing**: The previous two steps iterate until all measurements are filtered. After the last measurement (n-th) is filtered, the algorithm runs backward in time updating all filtered state vectors ($\hat{\mathbf{x}}_{k|n}$) using information from all n points.

Details of the Kalman filter track fitting algorithm can be found in a thesis [28]. Figure 6.16 shows an example of track fitting in a real event.

In the filtering process, a local χ^2_k is calculated based on the predicted and measured values. It is used to identify outlying hits, which, once identified, are removed from



(a) X-Y view (b) X-Z view

Figure 6.16: An example of track fitting. The fitted track is shown in a blue curve. Note that many track fitting algorithm features are demonstrated. It starts with two distinctive track candidates (X-Y view), which have very close state vectors at a certain point and, therefore, are merged as one track of two turns (X-Z view). A single-hit cluster on chamber 1 is identified as an outlier and is hence eliminated from the fitted track.

the fitted track (also see Figure 6.16). It is also used to correct the hit left-right ambiguity if assigned incorrectly in the tracking stage. A global χ^2 is also calculated for a fitted track to evaluate the track quality.

Sometimes multiple tracks may have very close state vectors at certain points to be considered as different turns of one track in the spectrometer, and are hence merged by the fitting algorithm (again see Figure 6.16).

The fitted track is extrapolated in both directions, one into the muon stopping target and the other into the TIC. The state vector at the stopping target gives the muon decay vertex, the emission angles, and the momentum with the covariant error matrix. The total track length from the target to the TIC is calculated and used to deduce the positron time-of-flight, t^e_{ToF}.

The performance of the tracking algorithm is presented in Section 8.1.

6.3 TIC Analysis

The TIC analysis provides the TIC hit time and position extracted from the waveforms.

6.3.1 Waveform Analysis

As mentioned in Section 3.4.1, the PMT signals from TIC bars are split into three parts with a fraction of 8:1:1 by a passive splitter. The largest signal portion goes to a double threshold discriminator (DTD), which outputs a NIM[3] logic pulse. One of the smaller portions is digitized by the DRS. An example of a PMT waveform and the generated NIM logic pulse is shown in Figure 6.17.

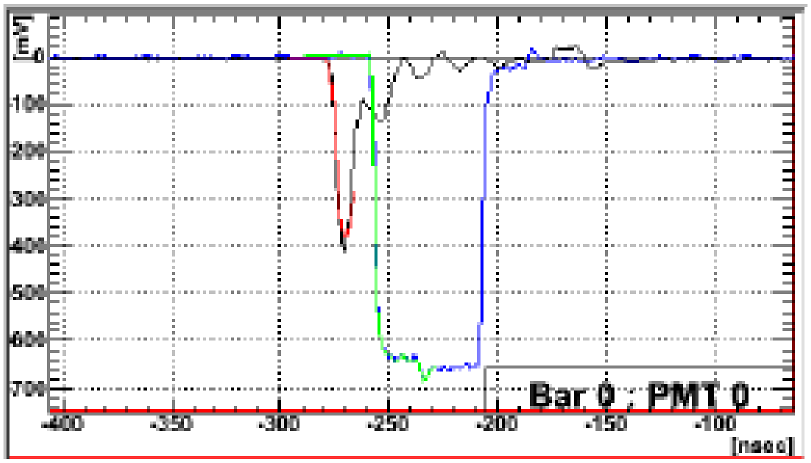


Figure 6.17: An example of the TIC waveforms. The black line shows the DRS digitized PMT pulse, which is fitted to a template indicated by the red line. The NIM logic pulse of the same signal outputted from the DTD is shown in blue line. Its template is fitted and overlaid on top of it in a green line. The NIM time is lagged by ∼20 ns due to the electronics delay in the DTD.

The charge and amplitude are obtained from the DRS waveform. The baseline of the DRS waveform is assumed to be constant. The charge is calculated by integrating the pulse over a 30 ns window around the pulse peak. With a large enough sample of cosmic events, TIC hits distribute uniformly along the bars. Therefore, by equalizing the peaks of the PMT charge distributions at the two ends of a TIC bar, we can correct the product of PMT gain and QE. Hence, the total number of input optical photons to a PMT can be derived. The pulse amplitude is simply taken as the difference between the estimated baseline and the pulse peak.

The pulse time is extracted from the corresponding NIM logic pulse of the same

---

[3] The fast-negative logic pulse of Nuclear Instrument Module (NIM). The voltage of 0 V for logic 0 and that of 0.8 V for logic 1, where logic 1 corresponds to the presence of a hit.

channel by a template waveform fitting which is obtained by averaging a large number of pulses from all channels. The fitting has the baseline level and the leading edge time as two free parameters. The NIM logic signal has a sharp leading edge close to a step function; whereas the leading edge of the actual PMT pulse (the DRS pulse), defined as a constant fraction of the pulse height, is pulse amplitude dependent. Therefore, the NIM pulse time has to correct a time walk effect due to the fact that different rising slopes of the original pulses pass the DTD threshold at different fractions of the pulse peak. We subtract a term that is proportional to $1/\sqrt{A}$, where A is the PMT pulse amplitude.

6.3.2 Hit Reconstruction and Clustering

A TIC hit is identified by the two signals on each end of a TIC bar. The hit time is first estimated as the weighted average of the two pulse times after time-walk corrections. The weighting factor is the number of input optical photons to each PMT. The TIC hit Z coordinate is determined by the time difference between the two ends times the effective speed of light in the TIC bar. After knowing the hit Z, the hit time is further corrected by the propagation time from the Z location to the PMTs.

Sometimes positrons pass through not only one timing bar, but through multiple bars as shown in Figure 6.18. In these cases, TIC hits that are consistent with a single charged particle passing through are clustered by their time closeness and Z adjacency. The TIC hit time for such a cluster is given by the time of the hit on the first bar.

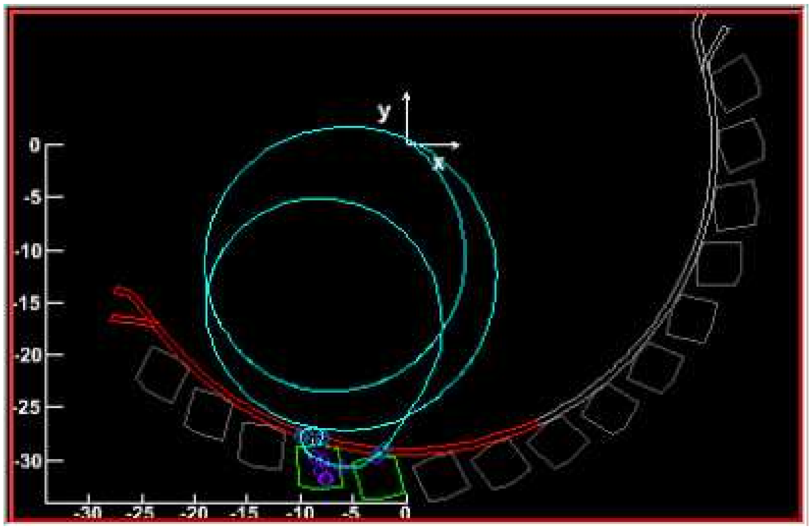
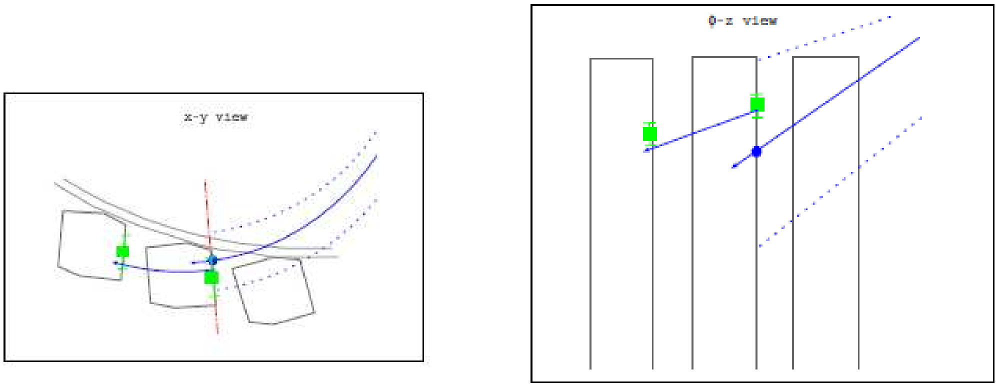


Figure 6.18: An example of a positron hitting multiple bars. The cyan curve indicates the positron trajectory, and green boxes are the bars hit.

6.4 DCH-TIC Interconnection

The last step of positron reconstruction is to associate the DCH tracks with the TIC hits, which gives a well measured time to a well reconstructed track. This is done by extrapolating the track onto the surface of the TIC bars and checking the Z and R adjacencies between the track incident point and the TIC hits. A TIC hit to be matched with the track must be located within the range of ±10 cm in R and ±20 cm in Z from the projected track incident point after correcting the offsets. Figure 6.19 illustrates the track extrapolation to the incident point on the TIC.

In addition to the spacial requirement of R and Z adjacencies, the DCH-TIC interconnection algorithm also calculates a $\chi^2_{DCH-TIC}$, which is required to be less than 10 to guarantee a good match. The matching criteria on R, Z and $\chi^2_{DCH-TIC}$ are not very strict because the projected track from the last measurement point in the DCH to the TIC surface has large uncertainties due to the scattering and energy loss in materials between the DCH and TIC such as chamber support structure, cables,



(a) (b)

Figure 6.19: An illustration of the track extrapolation at the end point on the TIC. The DCH-TIC connection requires spacial matching.

pre-amplifiers, etc..

TIC Hit Time Refining

After associating the DCH tracks with TIC hits, we are able to improve the TIC hit time measurement for the multi-hit clusters by the knowledge of the matched DCH track. In a multi-hit cluster, the hit time on the second bar is delayed from the first one by the track propagation time between the two, which can be calculated from the track extrapolation and subtracted off. Therefore, the TIC hit time, t_{TIC}, can be taken as the average of the first hit time and the corrected second hit time. Assume both time measurements have the same precision, then the TIC time is hence improved by a factor of $\sqrt{2}$.

Although we sometimes get TIC clusters having more than two hits, we only use the first two hits on consecutive bars. This is because after passing through a few bars, positrons often undergo hard scattering, thus their extrapolations are hard to estimate.

Positron Emission Time

At one end, the track is extrapolated to the TIC surface; at the other end, the track is back-projected onto the plane of the muon stopping target, where the intersecting point is regarded as the muon decay vertex. If a track is matched with a TIC time, t_{TIC}, subtracting the positron time-of-flight, t^{e}_{ToF}, the positron emission time at the decay vertex, t_e, can be derived as the following,

$$t_e = t_{TIC} - t^{e}_{ToF} \tag{6.4}$$

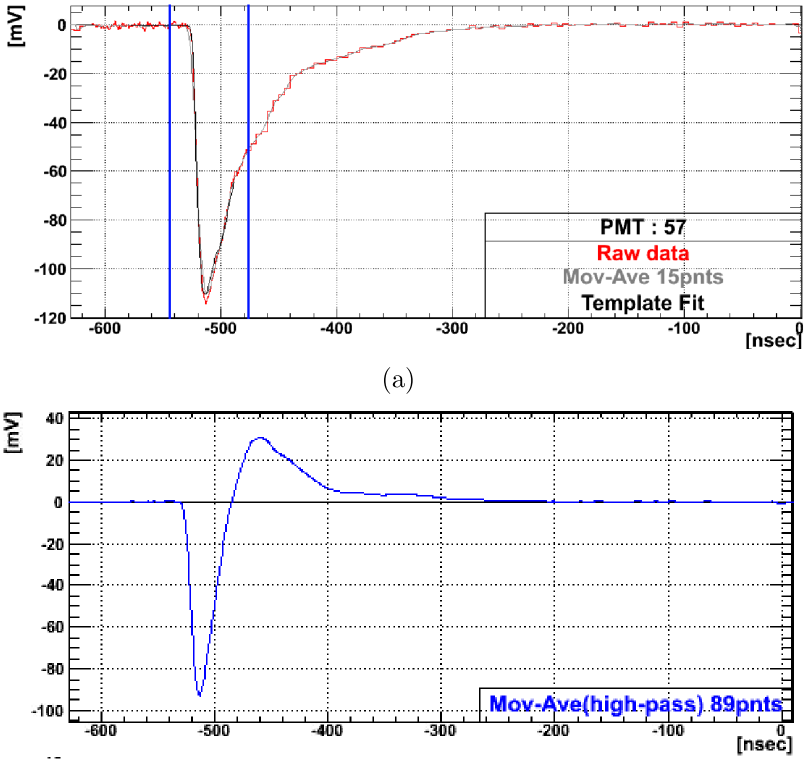
where t^{e}_{ToF} can be calculated as the positron track length divided by the positron speed that the high energy particle travels at the speed of light.

6.5 XEC Analysis

Photon information is solely from the XEC analysis, which reconstructs photon energy, position and time simultaneously. However, the XEC detector is not sensitive to the photon direction, hence it is assumed that the photon is emitted at the muon decay vertex determined as the intersecting point of the positron track and the muon stopping target plane.

6.5.1 Waveform Analysis

We have 846 PMT waveforms for every event. An example of a typical XEC PMT waveform is shown in Figure 6.20(a). We extract the PMT time and charge from the waveforms.



(a)

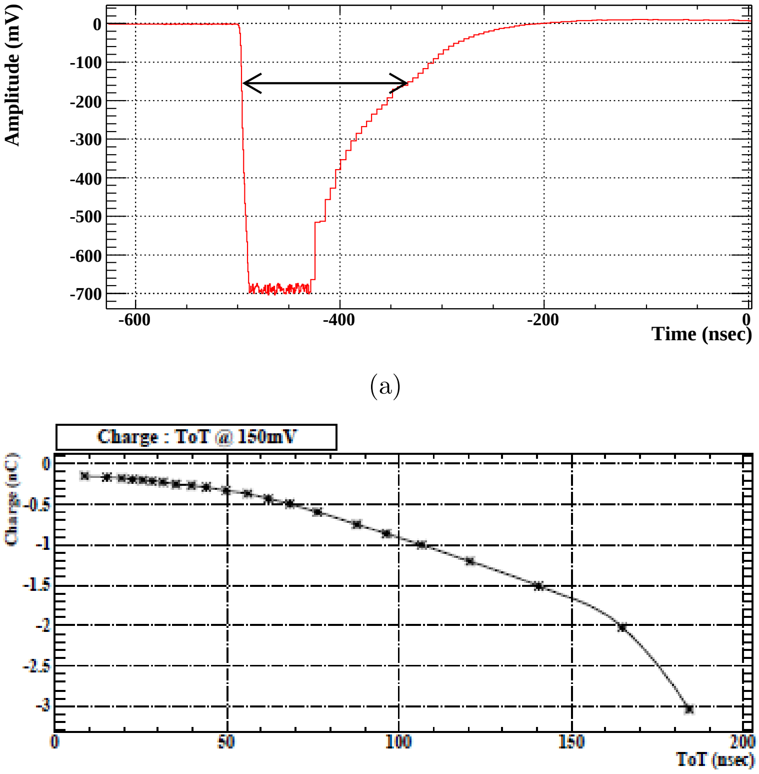
(b)

Figure 6.20: Plots of the XEC waveforms. (a) an example of a typical XEC PMT waveform. The red curve shows the raw waveform; the grey curve is the smoothed waveform obtained by a running average for every 15 bins of the raw waveform; and the black curve is the template fitted to this waveform. The region between two blue lines is the so called region of interest, outside of which the raw waveform is rebinned at the DAQ stage to reduce data size. (b) the filtered waveform after a high pass digital filter. The region between red bars in (a) is the integration range, which is from the zero crossing time in (b) to 48 ns before it.

The leading edge time of a pulse is determined by the constant fraction method, which sets the leading edge time as the time of the voltage that reaches 30% of the full pulse height. This method gives a time that is independent of the pulse height. The pulse height is obtained by finding the peak of the fitted template to the waveform, instead of looking for the pulse amplitude directly from the raw waveform, to avoid noise and fluctuation of photo-statistics. The template is taken as the average pulse shape of a large number of events.

The total number of photoelectrons of a PMT is derived from the charge integration of the waveform. The baseline voltage is assumed to be constant and evaluated for each PMT by averaging the waveform before the pulse at every event. However, a low frequency noise $\sim$1 MHz is observed whose interval ($\sim$1000 ns) is of the same order as the DRS window (625 ns). Therefore, the baseline subtraction cannot remove this low frequency noise completely. A high-pass filter is applied for the purpose of eliminating slow-component noise, and the high frequency noise is filtered by the integration itself. The filter is built by subtracting a running average of the raw waveform (which is in fact a low-pass filter) from the raw waveform itself. The number bins of the running average is optimized as 89, which corresponds to a cut-off frequency at 10.7 MHz. Figure 6.20(b) shows the filtered waveform of the raw waveform in Figure 6.20(a). The common charge integration window for all channels is determined by the sum of all filtered PMT waveforms, which is set from the zero crossing time and 48 ns before it.

In reality, photons sometimes (15%) interact with the liquid xenon right after they enter the detecter (<1 cm), which results in large signals being observed by the close-by inner face PMTs. These large pulses are often saturated due to the limited DRS dynamic range. An example of a saturated signal is shown in Figure 6.21(a). A time-over-threshold (ToT) technique is developed to recover the satuated signals. ToT is the time span of the saturated pulse over a relatively low threshold of 150 mV. We can thereby estimate the charge from the template waveform having the same ToT as the saturated signal. The ToT to charge relation is shown in Figure 6.21(b).



(a)

(b)

Figure 6.21: Illustration of the ToT method. (a) a saturated PMT signal. The black arrow indicates the time span of the pulse over the threshold of 150 mV, the ToT. (b) the ToT to charge relation.

6.5.2 Photon Reconstruction

In the photon reconstruction, the algorithm uses thef number of scintillation photons (N_{pho}) hitting a PMT as the physical observable to determine the photon kinematics, instead of the PMT charge (Q). We first convert the PMT charge to the number of photoelectrons (N_{phe}) detected by the PMT,

$$N_{phe,i} = Q_i/(e \cdot G_i), \tag{6.5}$$

where the subscript i is the PMT number, G_i is the gain of the PMT, and e is the elementary electric charge. We further convert N_{phe} to N_{pho},

$$N_{pho,i} = N_{phe,i}/QE_i, \tag{6.6}$$

where QE is the quantum efficiency of the PMT.

Photon Position

In reality, photons pair produce and cause 'showers' (though having only a few steps) with certain profiles in the XEC. The algorithm, however, attempts to find the photon interaction position assuming the scintillation light all comes from a point-like source. Therefore, N_{pho} should be proportional to the solid angle of the photo-cathode viewed from the photon interaction point. The algorithm finds the photon position by fitting the distribution of N_{pho} of inner face PMTs in a limited region to minimize a χ^2,

$$\chi^2_{\mathbf{x}_\gamma} = \sum_i \frac{N_{pho,i} - \alpha \cdot \Omega_i(x, y, z)}{\sigma_{pho,i}(N_{pho,i})}, \tag{6.7}$$

where $\Omega_i(x, y, z)$ is the solid angle of the photo-cathode of the i-th PMT viewed from point (x, y, z), α is a constant coefficient and a free fitting parameter, and $\sigma_{pho,i}(N_{pho,i})$ is the statistical uncertainty of the PMT output, which is based on the statistical fluctuation in the observed number of photoelectrons,

$$
\begin{aligned}
\sigma_{pho,i}(N_{pho,i}) &= \frac{1}{QE_i} \times \sigma_{phe,i}(N_{phe,i}) && (6.8) \\
&= \frac{1}{QE_i} \times \sqrt{N_{phe,i}} && (6.9) \\
&= \sqrt{\frac{N_{pho,i}}{QE_i}}. && (6.10)
\end{aligned}
$$

The limited region on the inner face is around an initial estimated point (u_0, v_0) in the XEC local coordinate (see Figure 3.25), where (u_0, v_0) is the center position of the PMT having the maximum charge in the event. The fit is done twice to minimize the effect of shower fluctuation. The first fit uses inner face PMTs within a 3.5-PMT distance around (u_0, v_0) and improves the position with better estimation (u_1, v_1), and the second fit uses PMTs within a 2-PMT distance around the improved position, and finally gives the photon interaction position (u_γ, v_γ).

The fit is in fact a three-dimensional one; therefore, not only the 2-d (u_γ, v_γ) coordinates, which are projected to the inner face, are reconstructed, but also the depth, w_γ, is estimated.

Nevertheless, the reconstructed position by the fitting still has residual bias in u- and w-coordinates, while there is essentially no bias in v-coordinate. These biases are studied with MC simulation, and corrections are applied.

Photon Energy

The basic concept of energy reconstruction is based on the fact that high energy photons typically deposit all their energy in the XEC and almost all deposited energy is converted to scintillation light.

The algorithm finds the photon energy by summing weighted N_{pho} of all PMTs,

$$N_{sum} = \sum_i \frac{1}{c_i} \cdot N_{pho,i}, \tag{6.11}$$

where c_i is the photo-cathode coverage of the i-th PMT, which takes the different coverage of photo-cathodes due to the different PMT location into account. Figure 6.22 shows the different PMT photo-cathode coverage on different XEC faces.

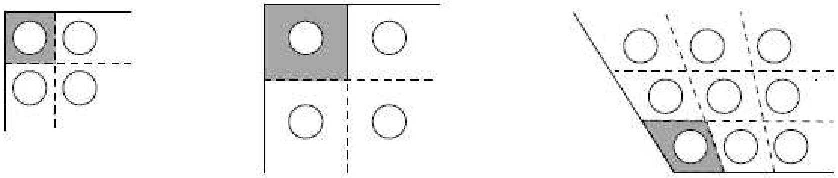


Figure 6.22: The PMT photo-cathode coverage (ratio of the area of the circle indicating a PMT cathode to the area confined by the solid and dashed lines) on different XEC faces. From left to right: inner, outer, top faces.

The XEC detector, however, does not provide an absolute photon energy scale, which has to be determined by CW and CEX calibrations. The linearity of N_{sum} to photon energy E_γ is obtained and shown in Figure 6.23.

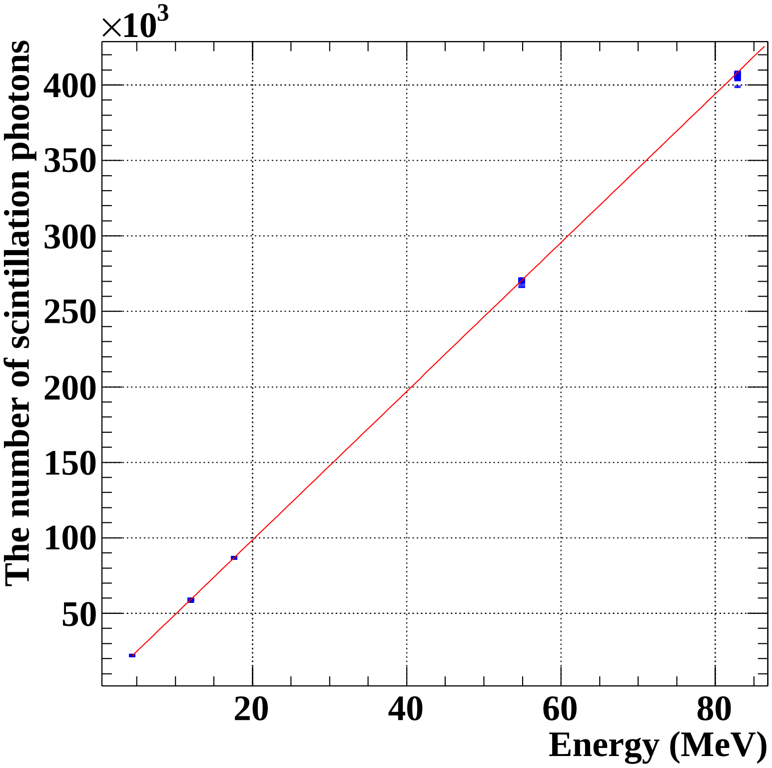


Figure 6.23: The linearity of N_{sum} vs. E_γ, where 4.4 and 11.7 MeV are from CW-B calibration, 17.6 MeV from CW-Li calibration, and 54.9 and 82.9 MeV from π^0 runs. The red line is the best-fit linear function requiring a fixed intercept at zero.

Photon Conversion Time

The photon deposits its energy at the first interaction point and the scintillation light propagates to and is measured by the PMTs. Therefore, the photon conversion time is reflected by each and every PMT time. For the i-th PMT, the photon conversion time is measured and expressed as,

$$t_{XEC\ hit,i} = t_{PMT,i} - t_{propagate,i}(d, v_{eff}) - t_{offset,i}, \tag{6.12}$$

where $t_{PMT,i}$ is the PMT leading edge time measured by the constant fraction method described in Section 6.5.1; $t_{propagate,i}(d, v_{eff})$ is the propagation time of the scintillation light from the conversion point to the PMT, which is modeled by the distance d between the light source and PMT and the effective speed of light in the liquid xenon v_{eff}; and $t_{offset,i}$ is the time delay of each channel due to cable length and electronics. The factor v_{eff} is determined from π^0 data, where $t_{propagate,i}$ is extracted by comparing the time of a reference counter. The value of v_{eff} is tuned to 8 cm/ns.

The time reconstruction algorithm finds the best estimated time by minimizing a χ_t^2,

$$\chi_t^2 = \sum_i \frac{(t_{XEC\ hit,i} - t_{XEC\ hit})^2}{\sigma_{t,i}(N_{phe})^2}, \tag{6.13}$$

where $\sigma_{t,i}(N_{phe})$ is the time resolution of each PMT as a function of the number of photoelectrons and $t_{XEC\ hit}$ is the best estimated photon conversion time. Only

PMTs with more than 50 photoelectrons are used in the minimazation process.

6.5.3 Pileup Identification

Operating in a high rate environment, multiple photons sometimes enter the XEC detector separated only by a time interval short enough to be recorded in the same trigger window and hence are considered as one event, which is called a pileup event. From the Monte Carlo study, ∼8.5% of events have pileup photons, tyically two, at the beam rate of $3{\times}10^7 s^{-1}$. The pileup events are identified by i), light distribution of the inner surface PMTs, and ii), time distribution of all PMTs.

A threshold of ∼8 mV is set on the PMT pulse height in the peak search process on the light distribution of inner face PMTs. This finds the position of each photon in the calorimeter and separates multiple γ rays spatially (see Figure 6.25(a)).

Two photons enter the XEC having different interaction time so that the PMTs close to each photon's conversion point observe different leading edge times. Therefore, the normalized χ^2_t value of the time fit, $\hat{\chi}^2_t$, can be used to separate photon overlap temporally. A comparison of the $\hat{\chi}^2_t$ distribution of MC and data is shown in Figure 6.24. A threshold of $\hat{\chi}^2_t$ is set to 3 to identify pileup events.

These two methods are complementary to each other.

After finding the pileup photons, the algorithm recovers the pileup events by eliminating the overlapping one. A table is prepared using CW data, in which average

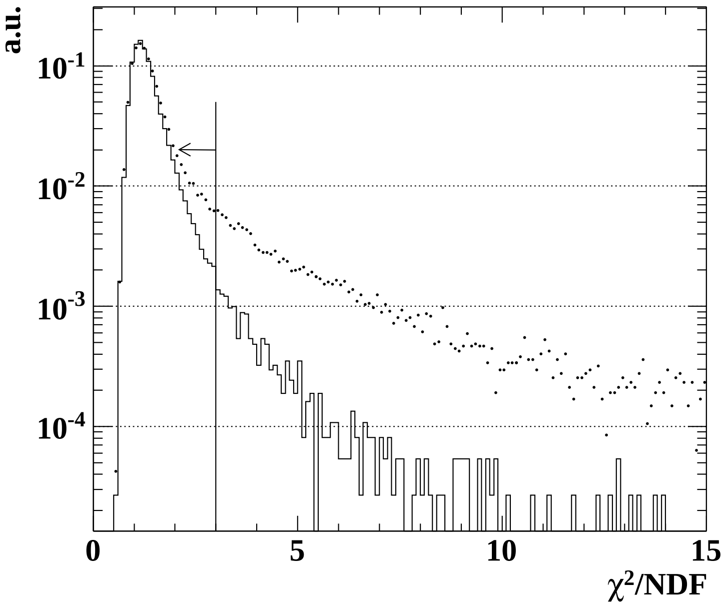
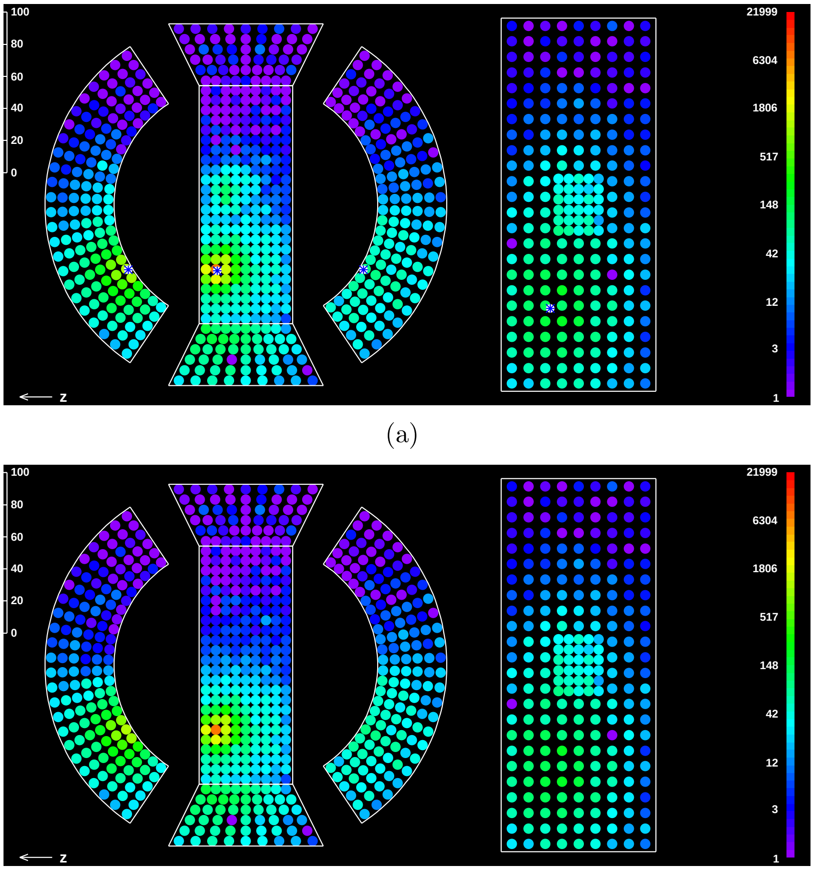


Figure 6.24: Comparison of $\hat{\chi}_t^2$ distribution from MC simulation (dotted) and data

outputs of each PMT for every photon conversion position mesh ($1.55 \times 1.55 \times 1.55$ cm^3) are written. Rough position and energy estimates are first done by the PMTs excluding the ones around the overlapping photon, whose expected outputs (assuming only one photon in the XEC without the overlapping one) are calculated from the table. These PMTs' outputs are then replaced by the expected values. The energy is therefore re-calculated using the new PMT charges. Figure 6.25 shows an example of pileup elimination.

The performance of the photon reconstruction is covered in Section 8.3.



(a)

(b)

Figure 6.25: An example of pileup elimination: before (a) and after (b). The PMT output is color coded. (a) is also an example of spatial identification of pileup events.

6.6 Positron-Photon Combined Analysis

The XEC detector only finds the photon conversion point, time and energy deposit. It does not provide the direction of photon flight. The photon is assumed to be emitted at the muon decay vertex, which is obtained as the intersecting point between the positron track and the muon stopping target plane. The photon travels in a straight line and its travel distance is expressed as $|\mathbf{x}_\gamma - \mathbf{x}_\mu|$. Therefore, the photon time-of-flight, t^γ_{ToF}, can be derived. Hence, the photon time at the decay vertex, t_γ, is obtained as the XEC hit time, $t_{XEC\ hit}$ subtracted by the photon time-of-flight,

$$t_\gamma = t_{XEC\ hit} - t^\gamma_{ToF}. \tag{6.14}$$

And the relative time difference between positron and photon is then,

$$t_{e\gamma} = t_\gamma - t_e. \tag{6.15}$$

The photon travel direction is found as

$$\hat{\mathbf{p}}_\gamma = \frac{\mathbf{x}_\gamma - \mathbf{x}_\mu}{|\mathbf{x}_\gamma - \mathbf{x}_\mu|}, \tag{6.16}$$

and the positron direction, $\hat{\mathbf{p}}_e$ is obtained in the DCH analysis. Finally, the opening angle between the photon and positron is given as

$$\cos\theta_{e\gamma} = \hat{\mathbf{p}}_{\gamma} \cdot \hat{\mathbf{p}}_{e}. \tag{6.17}$$

Chapter 7

Calibrations

In order to have a good detector performance, it is essential to make various calibrations to all sub-detectors. In this chapter, calibration techniques and procedures are described in detail.

7.1 DCH Calibrations

7.1.1 Channel to Channel t_0 Calibration

A channel to channel t_0 calibration is done by the method of finding the leading edge of distribution of hit time corrected by track T_0 of each individual anode channel.

We use the dedicated calibration runs for this purpose. The data were taken with trigger 16, which required hits in at least four out of five consecutive chambers and a timing counter hit in time.

In the analysis, for each wire, both ends are corrected by the propagation delays from the hit point to the anode ends. We make histograms of hit time subtracted by track T_0 of each channel after requiring that tracks satisfy minimum quality criteria. We select tracks that pass sufficient number of chambers and thus have better determined track time for timing calibration purpose. It also requires that the tracks used should be in a relative narrow time window with respect to the trigger time and have good reconstructed momentum and momentum error. The list of criteria is summarized in Table 7.1. The histogram represent distributions of corrected hit times in a cell, which have sharp leading edges as the characteristic shape (Figure 7.1).

Quantities	Requirements
number of found tracks in an event	1
track time	$\|T_0\| < 40$ ns
reconstructed momentum	43 MeV $< p_e <$ 57 MeV
momentum rms of momenta measured by every 3 consecutive clusters in a track	$\sigma^{rms}_{p_e} < 10$ MeV
number of hits	$N_{hits} \geq 10$
number of clusters	$N_{cluters} \geq 5$

Table 7.1: Track quality cuts used in calibration

The distribution, from the leading edge to 100 ns, is fitted to a fifth order polynomial and relative offsets of different wire-ends are obtained by the constant fraction method. The time offset is defined as the time when the fit reaches 15% of its peak. The wire-end time offsets are further adjusted by the constraint that the average correction is zero to maintain the same average event time.

Some wire-ends have low statistics due to low chamber voltage, whose time offsets can not be obtained by finding the leading edge with a fit. For these wire-ends, the time offsets are taken as the average offset of all wires of the same end on the same chamber.

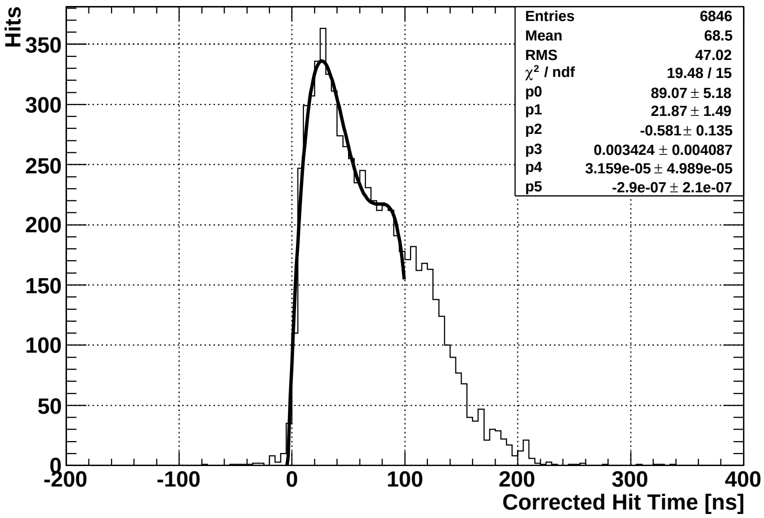


Figure 7.1: The corrected hit time distribution with a fifth order polynomial fit. The characteristic shape is a sharp leading edge.

Several iterations of the channel to channel t_0 calibration are made. At each iteration, new offsets are calculated by adding corrections obtained from current iteration following the above procedure. Deviations of corrections on time offsets at every end as well as the difference of both ends at each iteration should converge. The convergence is shown in Figure 7.2.

An example of distribution of hit time differences between two ends on a single wire, wire 163, after time calibration is shown in Figure 7.3. The core of the distribution is fitted with a Gaussian function. Non-Gaussian tails might be due to noise or pileup hits on the same wire in the DRS window. Two ends of the same wire should have the same intrinsic time resolution, σ_t. Hence, the RMS of the distribution of hit time difference between two ends is thus, $\sigma_{dt} = \sqrt{2}\sigma_t$. In this instance, it shows that $\sigma_t \sim 2.2$ ns.

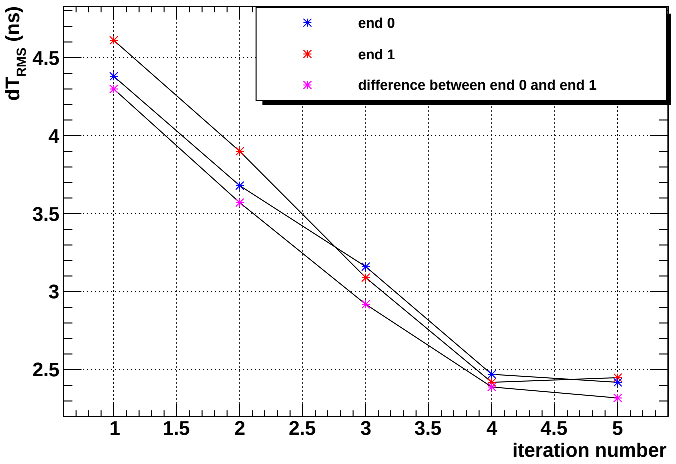


Figure 7.2: The RMS of corrections on all time offsets at each iteration. A converging process demonstrates that the calibration method works.

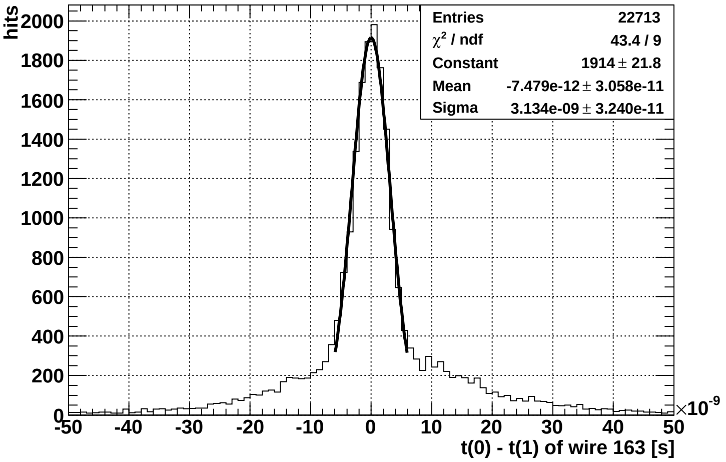
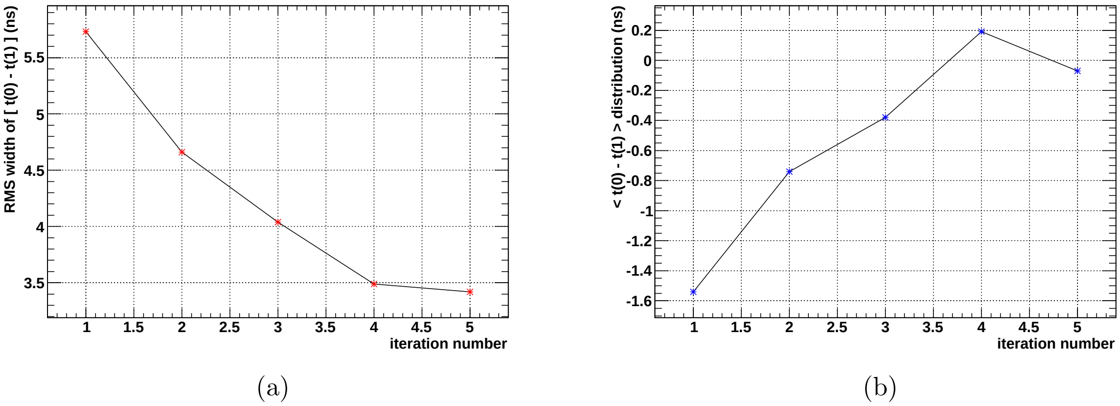


Figure 7.3: Distribution of hit time differences between two end on a single wire after time calibration. The core part of the distribution is fitted with a Gaussian function, $\sigma_{dt} = 3.13$ ns.

The channel to channel t_0 calibration should minimize the width of the distribution of hit time difference between two ends to the level of the intrinsic drift chambers time resolution. It should also center the mean of the difference in hit time between two ends at zero. So if one fits the core of the distribution of hit time difference between two end with a Gaussian function, the RMS width of that distribution should be minimized and the mean value should approach zero. Distributions of these two variables at each iteration are shown in Figure 7.4.



(a)

(b)

Figure 7.4: Certificate check of the iterative approach of time calibration. (a), RMS sigma of distribution of hit time differences between two ends of all the wires at each iteration. It is minimized during the iterative process. (b), mean of the distribution of hit time differences between two ends of all the wires at each iteration. It approaches zero.

All wire-end time offsets are stored in the database and used in the analysis. And the final set of calibration offsets is set to be used for the standard analysis for RUN 2008.

7.1.2 Optical Survey and Position Alignment

An optical survey of the drift chamber system was done to provide drift chamber anode wire positions in the MEG coordinate system to a good precision. The opti-

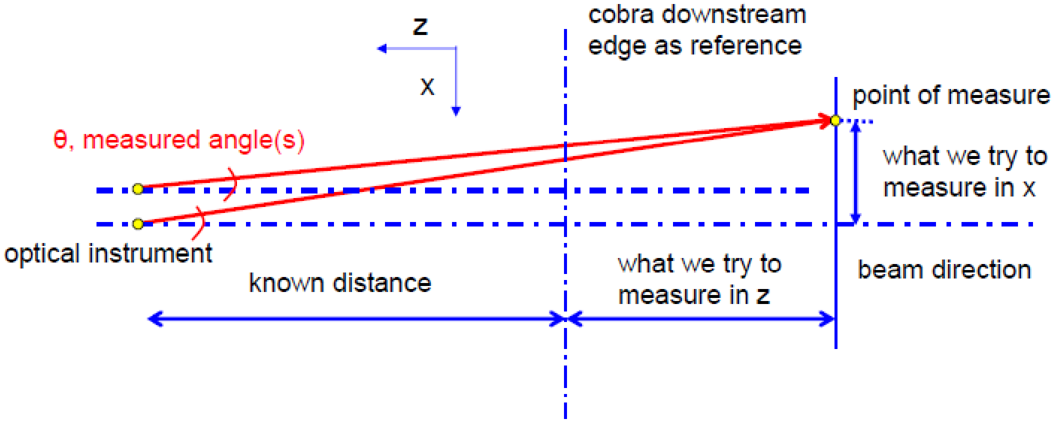


Figure 7.5: Survey methodology

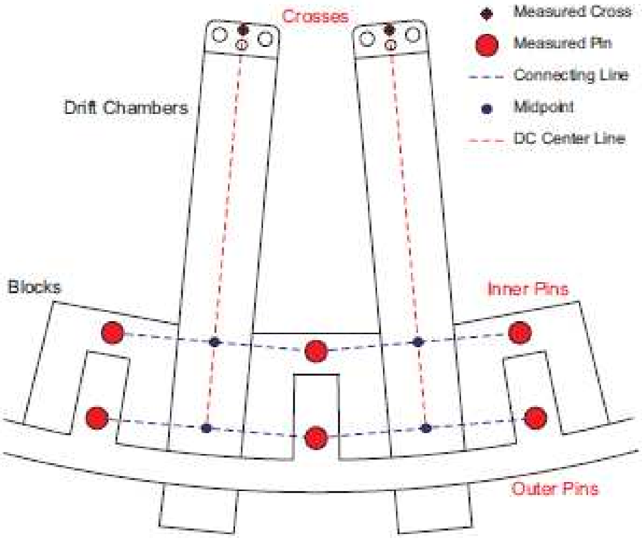
cal survey is based on simple trigonometry, whose methodology is shown in Figure 7.5.

The survey is based on laser measurements of two sets of reference marks on individual chambers and on their support structure. Two small plates with cross marks are glued on top of each chamber module, both upstream and downstream, to be measured to infer the chamber upper edge position. The chambers are mounted on the support structure by clamping them between carbon blocks. Thus, two radially aligned pins are inserted in each block; one serves as the other's auxiliary, in case one is obscured from visible access during the survey. Figure 7.6 shows the schematic of survey marks and their pictures. It is assumed that the midpoint of two adjacent pins is on the chamber central plane, thus a chamber slant angle is derived.

Each chamber has a position reference pin at the bottom of the frame. Distances from the survey crosses and anode wires to the reference pin were measured with a precision of less than 10 μm in a mechanical workshop. This enables us to transform survey measurements to anode wire positions. The distance between upstream and downstream crosses in every chamber are both measured in the mechanical workshop

(a) Picture of survey crosses

(b) Picture of survey pins



(c) Schematic of survey marks

Figure 7.6: Schematic and pictures of survey marks. An assumption is made that the midpoint of two adjacent pins is on the chamber central plane.

with high precision (less than 10 microns) and by the survey. By comparing these two results, one can deduce the survey precision in Z direction. Table 7.2 shows that the sigma of difference between survey and mechanical workshop distance measurements between two cross marks on each chamber is 290 microns. In the survey, the distance of two cross marks on each chamber is obtained from two individual measurements, i.e. measurement of upstream cross mark and that of downstream. Thus if the survey measurement precision in Z is δ_Z, $\sigma_{diff} = \sqrt{2}\delta_Z$ ignoring the error of mechanical workshop measurement. Therefore, $\delta_Z \sim 200\mu m$. The error in X related to that in Z in the following way, $\delta X \sim \tan(\theta) \times \delta Z$, where θ is the measured angle shown in Figure 7.5. The setup of the survey instrument gave $\tan(\theta) \sim 0.1$. This indicates that $\delta_X \sim 20\mu m$. Similarly, one can get the error in Y, which is even smaller than that in X.

Figure 7.7 shows the optical survey result. Both ideal wire positions and survey wire position with displacement from ideal magnified by 10 are plotted.

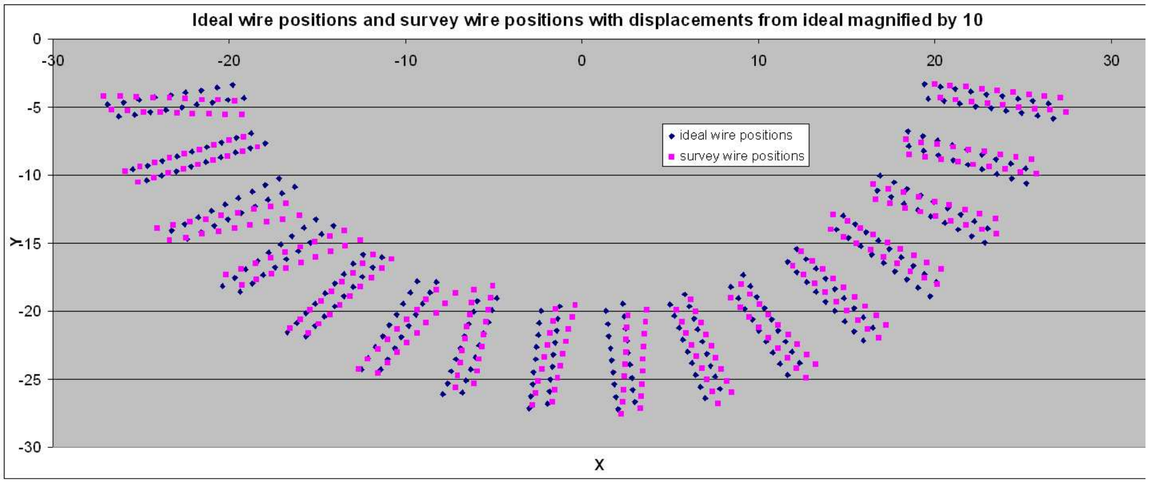


Figure 7.7: Survey results: chamber displacements

A software alignment technique was developed and implemented as a cross check of survey result as well as a possible further improvement. The technique is described

chamber number	survey (mm)	mechanical workshop (mm)	difference (mm)
0	958.99	959.26	0.26
1	958.46	959.13	0.67
2	959.78	960.46	0.68
3	961.30	961.35	0.05
4	959.46	959.75	0.29
5	960.15	959.91	- 0.25
6	960.00	959.92	- 0.08
7	960.41	960.59	0.18
8	961.07	960.98	- 0.08
9	961.56	961.41	- 0.15
10	961.64	961.45	- 0.19
11	958.74	958.72	- 0.02
12	959.19	958.98	- 0.21
13	958.89	958.92	0.03
14	961.09	961.34	0.25
15	959.69	959.89	0.20
		sigma (mm)	0.29

Table 7.2: Comparison between survey measurements and mechanical workshop measurements of distance between two cross marks on each chamber

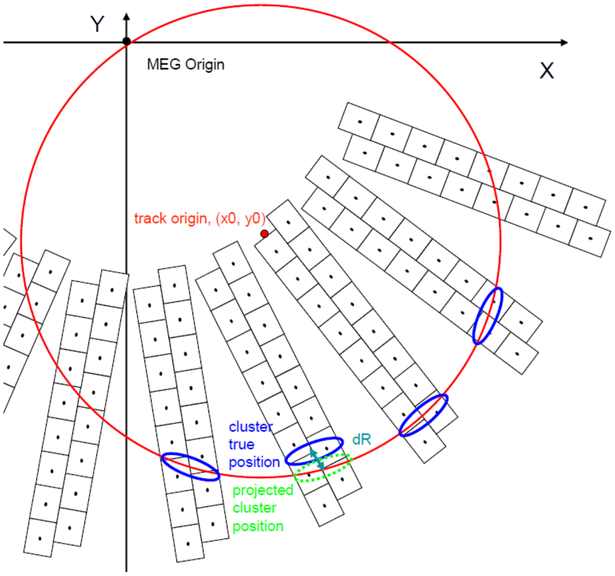


Figure 7.8: Alignment technique illustration

in Figure 7.8. We run the analysis software using survey wire positions. We first select well reconstructed positron tracks satisfying criteria listed in Table 7.1. For each cluster on a chamber, we use the surrounding three clusters (two on one side and one on the other side) to calculate a circle and plotted the difference between the circle and measured radial positions (dR). We fit the core of the distribution to a Gaussian function (Figure 7.9), whose mean gives the offset of one chamber to its adjacent ones. This technique may introduce an overall displacement systematically to all the chambers because the radial alignment of one chamber is based on the knowledge of its adjacent ones. We further corrected all the radial offsets by the overall average. This method fixes the two edge chambers, whose positions are taken from the survey. Any significant non-zero offset after the correction indicates a radial misalignment. The result is shown in Table 7.3. Typical radial offsets are below 50 μm, which is consistent with the survey precision.

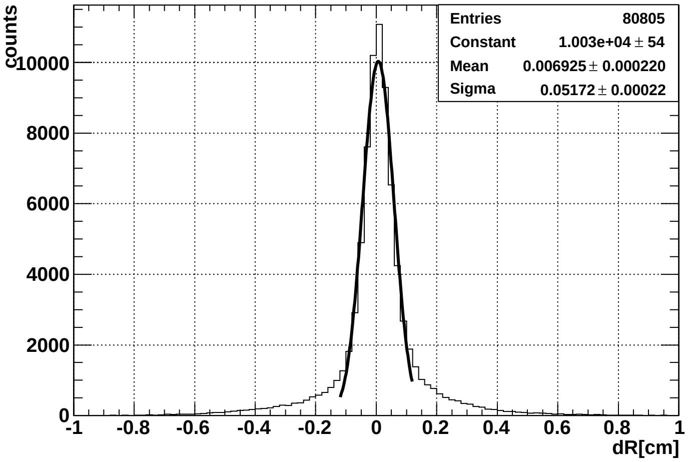


Figure 7.9: Distribution of the difference between the projected and measured radial positions (dR) in chamber 4, as an example, used in alignment radial offset study. Note the fitted mean is not corrected by the overall average yet. The corrected offset of chamber 4 is listed in Table 7.3.

Optical survey numbers were put into a database and used in the standard analysis

chamber number	radial offsets (μm)
0	0
1	- 54.4
2	35.4
3	- 38.0
4	46.1
5	- 77.6
6	33.3
7	50.8
8	- 13.8
9	44.7
10	49.6
11	- 32.2
12	- 48.5
13	32.7
14	- 28.1
15	0
sigma	45.9

Table 7.3: Software radial alignment result, on top of survey wire positions

for RUN 2008.

7.1.3 Z Calibration

The Z coordinate is first measured by anode charge division and further refined by charge asymmetry on the vernier patterned pads. Thus it is important to align the anode determined Z coordinate with that determined from the pads. Since this is only a relative alignment, it cannot hope to account for a shift in the position of the wires or pads from their assumed position in real space, which is done by optical survey instead.

Recall Equation 6.2 that, to the first order, the anode determined Z position is estimated by the charge division, ϵ_a (Equation 6.1).

Since the geometry of the pads is periodic every 5 cm, one expects a plot of pad asymmetry (ϵ_{pad} of the cathode or the hood) vs. anode determined Z coordinate to yield a periodic distribution with a 5 cm wavelength and a phase determined by the alignment of the pad relative to the origin. A typical example of ϵ_{pad} vs. z_{anode} is shown in Figure 7.10. The data is fit to a periodic function

$$\epsilon_{pad}(z_{anode}) = A_0(1 + m\,|z_{anode}|)\sin(\frac{2\pi}{\lambda}(z_{anode} - \delta)) + h. \tag{7.1}$$

It is clear that any systematic error in the effective wire length, $L_{eff} \equiv \left(\frac{L}{2} + \frac{R}{\rho}\right)$ in Eq. 6.2, will cause only a scale shift in the Z coordinate, which results in a wavelength change in Eq. 7.1. This is equivalent to a renormalization of the factor L_{eff}. A relative gain error, on the other hand, affects factor $\frac{Q_u}{Q_d}$ and hence produces a translation of the Z coordinate, which corresponds to the phase term in Eq. 7.1. A relative pad gain error will cause an offset in the pad charge asymmetry ϵ_{pad}, which is reflected by the vertical offset h in Eq 7.1. A parameter m is introduced as a slope for the amplitude to best fit the data.

Therefore, three calibration factors are extracted therein: the wire dependent z-scaling factor (related to the wavelength), the relative anode gain between upstream and downstream (related to the phase), and the relative pad gain between upstream and downstream (related to the vertical offset). All calibrations are stored in the database for the standard analysis. Results of those are shown in Figure 7.11.

A certificate check on the anode calibration is done by comparing the difference of the

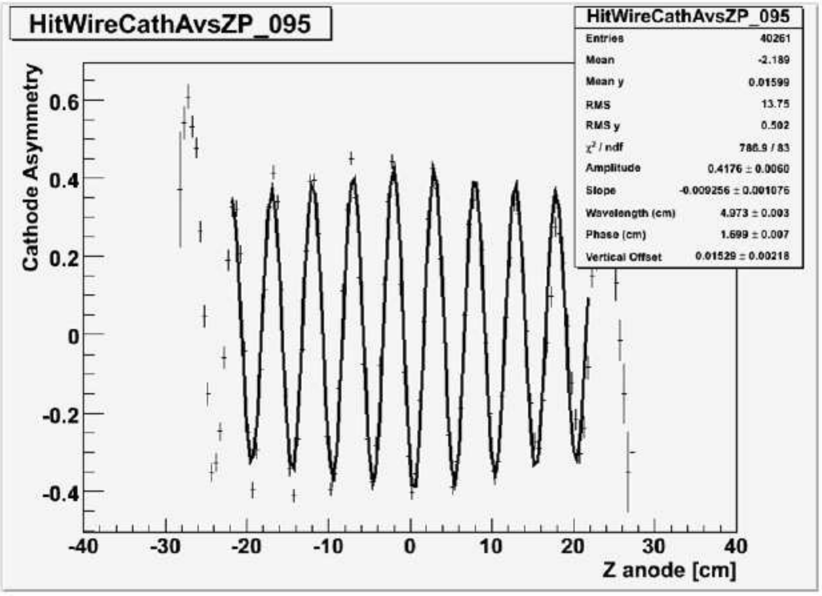
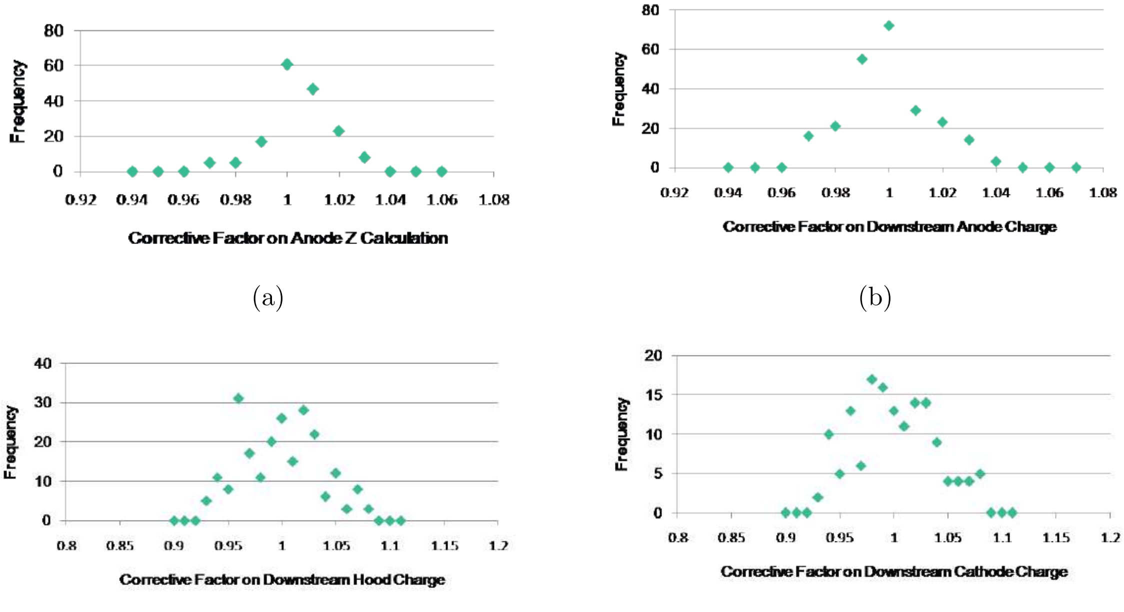


Figure 7.10: An example of fit to the cathode asymmetry vs. anode determined Z plot of wire



(a)

(b)

(c)

(d)

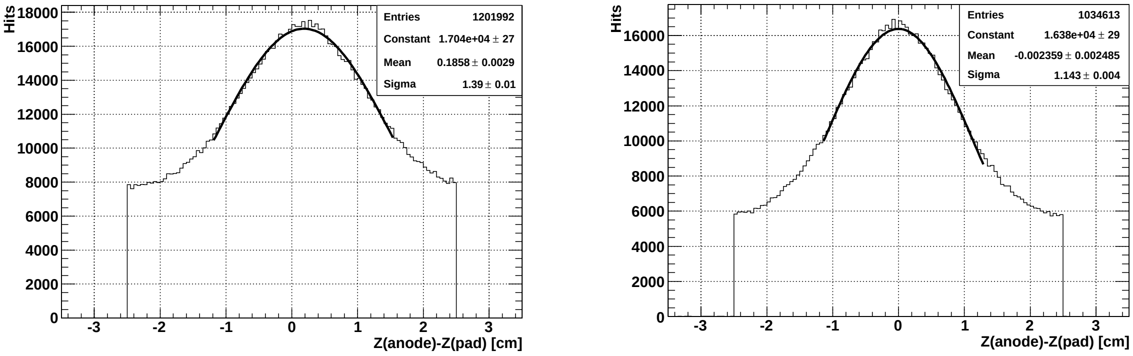
Figure 7.11: Distribution of Z calibration factors. (a), distribution of wire Z scaling factor; (b), distribution of relative downstream anode gain, assuming upstream gain = 1; (c), distribution of relative downstream hood gain, assuming upstream gain = 1; (d) distribution of relative downstream cathode gain, assuming upstream gain = 1

anode determined and the pad determined Z coordinate ($dz = z_{anode} - z_{pad}$) pre- and post-Z calibration (Figure 7.12). It was done with the early calibration Michel run data, when the chambers were in good condition. One can see that the anode-pad Z determination agreement had been improved, as was expected. It is to be noted that Z determination is very sensitive to the chamber gain, which is strongly related to the chamber high voltage; as the high voltage increases, so does the gain. During our normal data taking of RUN 2008, the chambers were not in a good condition, with many planes at low voltage and several dead ones as well. The same dz plots are made with runs during the normal data taking. One can see worsening of data quality in Figure 7.13, while the Z calibration still improves the anode-pad Z determination agreement comparing to non-calibrated data.

7.1.4 Time-to-Distance Relation Calibration

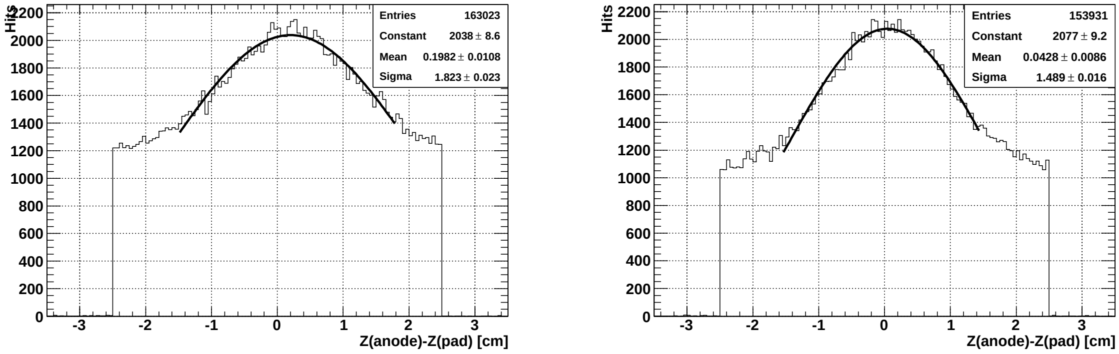
The hit position in a cell is determined by a time-to-distance relation function, which takes the corrected hit time ($t_{hit} - T_0$), the local B field and the track incident angle as input arguments and outputs the position of the ionization site that has the shortest drift time to the wire as the hit position. Since the function returns the cell local x, y coordinates as the hit position, it is termed the TXY function. However, the TXY function has a two-fold left-right ambiguity with any track incident angle. The track could pass through either side of the wire as shown in Figure 6.11, which is resolved at the tracking stage.

The TXY function was initially calculated by GARFIELD program, which simulates the field lines and electron drift lines and isochrones inside the ionization chamber. Should the calculated TXY relation have any deviation from reality, it can be seen



(a) before Z calibration (b) after Z calibration

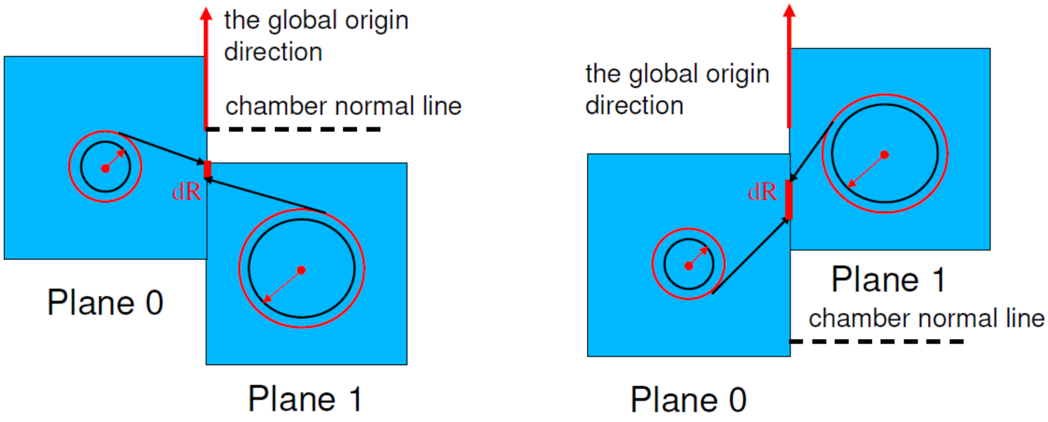
Figure 7.12: Certificate of Z calibration with early calibration Michel runs. Difference in the anode and pad determined Z coordinates before (a) and after (b) Z calibration. Anode-pad agreement improved as expected.



(a) before Z calibration (b) after Z calibration

Figure 7.13: Certificate of Z calibration with normal MEG runs. Difference in the anode and pad determined Z coordinates before (a) and after (b) Z calibration. Comparing to Figure 7.12, one can see the worsening of the data quality due to low chamber voltages. However, the Z calibration still improves the anode-pad Z determination agreement.

and then calibrated by data. After track finding, if the hits in two-hit clusters are on adjacent cells of different planes, we project both hits to the chamber central plane by the track incident angle. The local coordinate in the wire frame for plane 0(1) is denoted as $r_{hit,0(1)}$, and the projected position on the chamber central plane in global coordinate is $R_{cent,0(1)}$. We also define the difference of the two global radial coordinates as $dR = R_{cent,0} - R_{cent,1}$. To first order, there are two configurations of L/R assignment of any two-hit cluster, i.e. the track passes through same side of both wires (Figure 7.14(a)) or opposite sides of two wires (Figure 7.14(b)).



(a) Same side configuration

(b) Opposite side configuration

Figure 7.14: Illustration of hit projection onto chamber central plane in two L/R assignment configurations. The track incident angle is defined to be positive if the track enters the chamber above the chamber normal line and leaves the chamber blow it, as shown in (b). This is also a conceptual illustration of R resolution measurement.

If we assume the average drift velocity obtained from the TXY function is bigger than in reality, for a given drift time r_{hit} should be bigger than reality for both hits. This scenario is illustrated in Figure 7.14; the calculated drift circle is color coded in red and the real drift circle in black. It is easy to realize that the impact of the error in average drift velocity differs in two L/R configurations regardless of track incident angle. So for the same side configuration, the errors of both hits are on the same direction and thus only has an insignificant impact on dR; whereas for the opposite side configuration, the errors of two hits are on the opposite directions and thus has a significant impact on dR. Therefore, we used opposite side configuration as our calibration tool. Furthermore, if we don't separate positive and negative track incident angles, some effects would possibly cancel out. So we only look at opposite side configuration with positive track angle(Figure 7.14(b)). When we examined the distribution of dR vs. $r_{hit,0}$ of this group of two-hit clusters, we found that there was a tilt of the mean of dR vs. $r_{hit,0}$(Figure 7.1.4). This could be interpreted as a systematic average drift velocity error. It is more understandable if we project the 2D histogram to different slices in $r_{hit,0}$. Some examples of the projected histograms are shown in Figure 7.16. It is obvious from Figure 7.16(e) that the mean of dR has a strong linear correlation with $r_{hit,0}$, which increases 270 μm with 3 mm change of $r_{hit,0}$. This indicates a roughly 10% error in the TXY function, which could be interpreted as the average drift velocity being 10% faster than in reality. Note that this technique integrates all $r_{hit,1}$. Thus it is only a first order calibration. To the next order, similarly, we projected dR distribution by slicing through $r_{hit,1}$ with a fixed range of $r_{hit,0}$. We iterate this process until mean of dR vs. $r_{hit,0(1)}$ is flat. Finally the drift velocity is corrected accordingly.

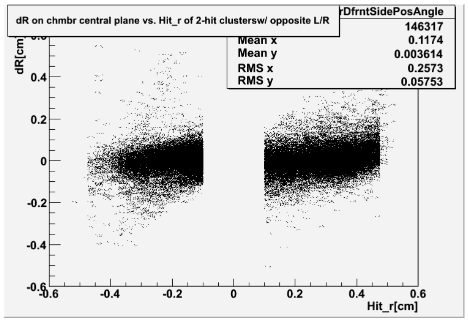
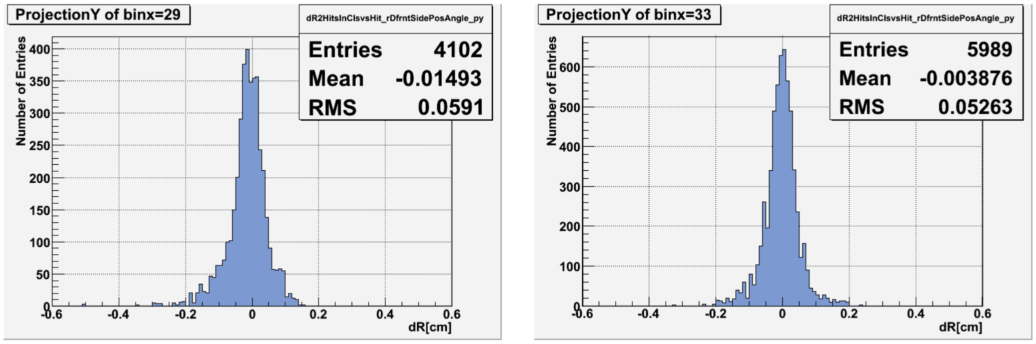
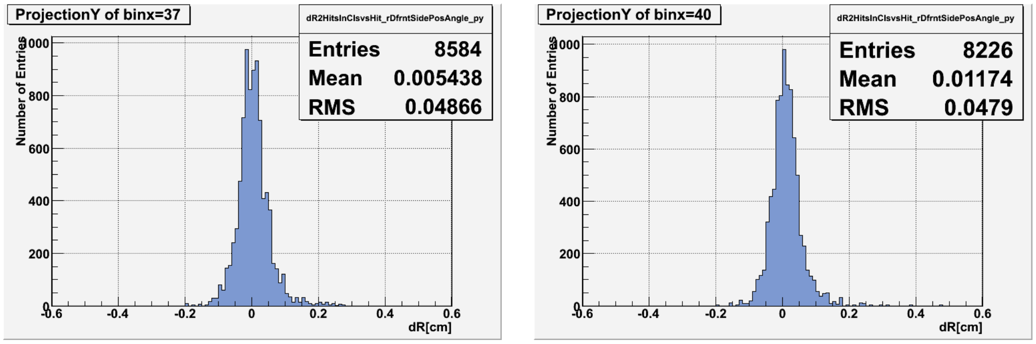


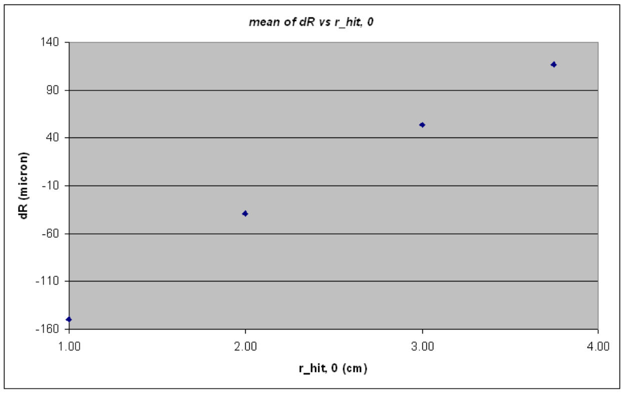
Figure 7.15: dR vs. $r_{hit,0}$ before TXY function calibration



(a) dR at $r_{hit,0}$ between 1 to 1.25 mm (b) dR at $r_{hit,0}$ between 2 to 2.25 mm



(c) dR at $r_{hit,0}$ between 3 to 3.25 mm (d) dR at $r_{hit,0}$ between 3.75 to 4 mm



(e) mean of dR vs $r_{hit,0}$

Figure 7.16: dR distribution slicing through different $r_{hit,0}$

7.2 TIC Calibrations

7.2.1 Same Bar PMT Time Offset Calibration

The TIC hit Z position is reconstructed by the time difference between the two PMTs on the same bar. The time offset between the two is calibrated by a large number of arbitrary cosmic events whose distribution of the impact point on a bar should be flat. So the mean of the time difference between the two ends, $\langle t_0 - t_1 \rangle$, gives the PMT offset directly. Figure 7.17 shows the $\langle t_0 - t_1 \rangle$ of each bar, derived by a large sample of cosmic events.

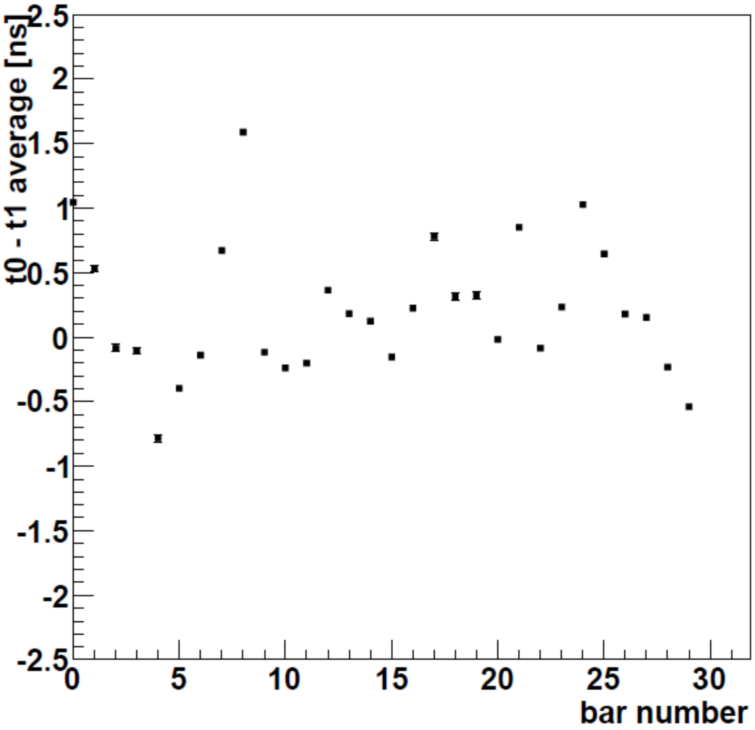


Figure 7.17: PMT time offset of each TIC bar

PMT time offsets are stored in the database and used in the standard analysis for RUN 2008.

7.2.2 Bar-to-Bar Time Offsets Calibration

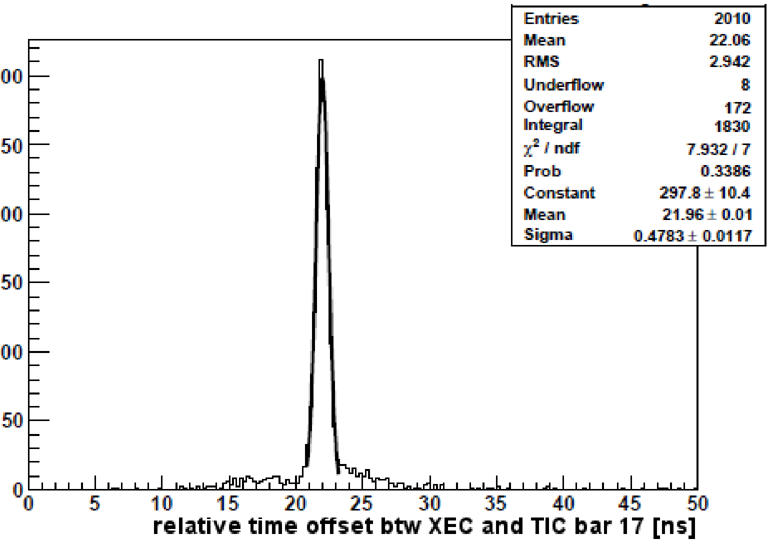
As explained in Section 3.3.2, nuclear reaction $\mathrm{p} + {}^{11}_{5}\mathrm{B} \rightarrow {}^{12}_{6}\mathrm{C} + \gamma$ from CW beamline sometimes produces simultaneous photons of 11.7 MeV and 4.4 MeV without any angular correlation. Therefore, we select such two-photon CW-B events with 4.4 MeV photon detected by the XEC and 11.7 MeV photon measured by a single TIC bar. Using the XEC time as a reference, the time difference of the two photons, $T_{\gamma\gamma\ i}$, after correcting the time-of-flight of each, gives the relative timing between the XEC and the i-th TIC bar. The mean of the $T_{\gamma\gamma\ i}$ distribution for each bar gives the measurements of TIC bar-to-bar time offsets. Results of the XEC-TIC relative timing and the TIC bar-to-bar offsets are shown in Figure 7.18.

Results of bar-to-bar time offsets are stored in the database and used in the standard analysis for RUN 2008.

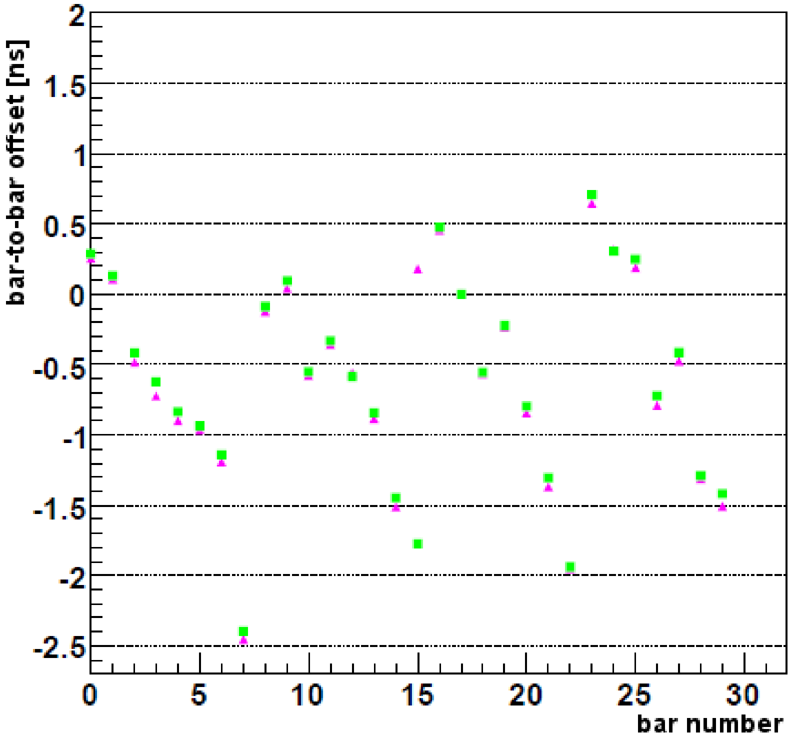
7.3 XEC Calibrations

7.3.1 PMT Gain Calibration

PMT gains are calibrated by the LEDs. With the assumption that LED outputs are constant and statistics of the number of photoelectrons, N_{phe}, obeys the Poisson distribution. The variance of the observed charge of the PMTs, σ_Q^2, has two uncorrelated contributions, i.e. that of the photo-statistics, σ_{phe}, and of the pedestal fluctuation,



(a) $T_{\gamma\gamma\ 17}$



(b) bar to bar offsets, referenced by bar 17

Figure 7.18: Plots of the PMT time offset calibration. (a) $T_{\gamma\gamma\ i}$ distribution of bar 17. It gives the relative timing between the XEC and TIC bar 17. (b) bar to bar offsets relative to bar 17. Different marks show calibrations done with two different data samples taken at different run periods, September (green) and November (purple). Discrepancy of bar 15 is due to a recovery of a bad channel.

σ_{ped}. Therefore, σ_Q^2 can be written as,

$$\sigma_Q^2 = (G \cdot e \cdot \sigma_{phe})^2 + \sigma_{ped}^2, \tag{7.2}$$

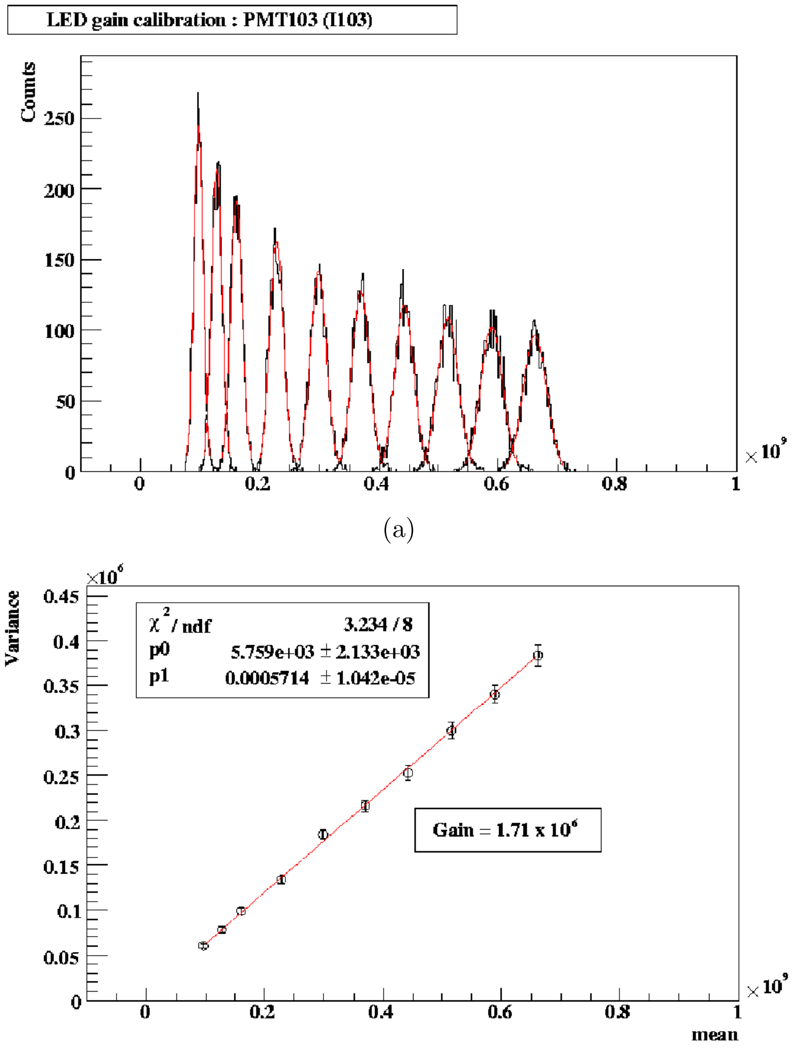
where G is the PMT gain and e is the elementary electric charge. Recall that $\sigma_{phe} = \sqrt{N_{phe}}$ and that the observed charge, Q, is related to number of the photoelectrons, N_{phe}, as $Q = N_{phe} \cdot G \cdot e$. We hence have,

$$\sigma_Q^2 = G \cdot e \cdot Q + \sigma_{ped}^2, \tag{7.3}$$

This gives an approach of extracting the PMT gain as the inverse of the slope of the variance-mean plot by taking LED data with different intensities. The benefit of this technique is that it is insensitive to the pedestal fluctuation, σ_{ped}^2. For each set of LED calibration runs, 10 different intensities are taken with N_{phe} at each step sufficiently large to be regarded as a Gaussian distribution. Figure 7.19(a) shows the PMT charge distribution of each intensity and Figure 7.19(b) shows the variance-mean plot from which the PMT gain is extracted.

7.3.2 PMT Quantum Efficiency Calibration

PMT quantum efficiency is calibrated by the α particles emitted from point-like light sources as described in Section 3.2.2 and shown in Figures 3.26 and 3.27. The use of α rays instead of LEDs is because that the QE highly depends on the wavelength of



(a)

(b)

Figure 7.19: PMT gain calibration. (a), PMT charge distributions with different LED each intensity. (b), variance-mean plot whose inverse of the slope is the PMT gain.

scintillation light and LED light wavelength is different from that of the scintillation light. We estimate the amount scintillation light observed of a given PMT by the knowledge of the positions of α sources and α spectrum and with the help of Monte Carlo simulation of a fixed QE (16%) as a reference. QE is therefore measured by comparing the PMT output and that of the simulation.

Details of XEC calibration methods and results can be found in a PhD thesis [64].

Chapter 8

Detector Performance

Detector performance is covered in this chapter. The detector resolutions are evaluated using data directly with a few exceptions which are first roughly estimated by data and further confirmed and validated by MC simulation. Data used in the performance evaluation are chosen with the selection criteria described in Chapter 9, unless otherwise mentioned. These selection cuts are also applied in the final physics analysis.

Detector efficiencies, another crucial aspect of performance, are described in details in Chapter 10, because they are closely related to radiative decay branching fraction calculation.

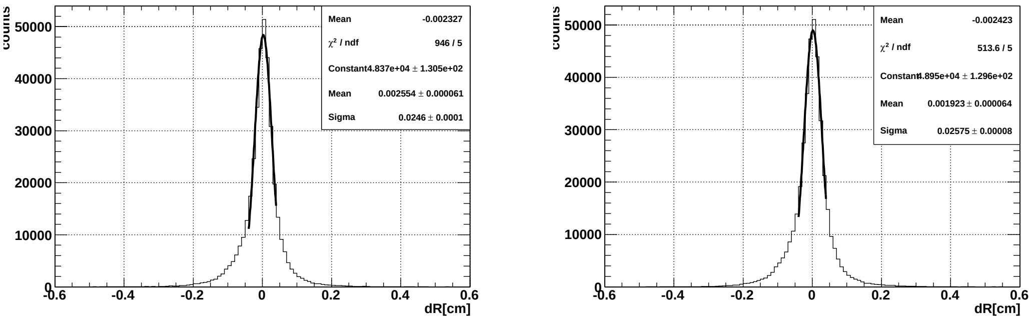
8.1 DCH Performance

The detector intrinsic resolutions of various observables are measured after the DCH calibrations completed. Different techniques of the same concept are developed to obtain R, Z, θ, ϕ and momentum resolutions.

8.1.1 Intrinsic R Resolution

The channel to channel t_0 calibration contributes to improve R resolution. We select two-hit clusters with adjacent hits on different planes of a chamber and find the difference in the positions of the two hits to infer the intrinsic R resolution. This position difference is calculated by using clusters in three consecutive chambers on a track to calculate a helix trajectory. The angle of the trajectory at the center chamber is used to project the hits on the two planes of the center chamber to the chamber midpoint. A conceptual illustration is shown in Figure 7.14.

The width of the distribution in the difference in these two projected positions is a measure of the single hit spatial resolution. This technique avoids the positron multiple scattering effects happened between two chambers, because it only compares the position difference of two hits on the same chamber. Figure 8.1 shows the distributions in this difference for hits on the same and opposite sides of the wires in the two planes. The core part of the distributions are each fitted to a Gaussian. For hits on the same (opposite) side, the fitted Gaussian has a σ_{dR} of 248(258) μm, corresponding to a R resolution on each wire of $\sigma_{dR}/\sqrt{2}$ =175(182) μm, assuming both wires having the same R resolution.



(a) (b)

Figure 8.1: Distributions in the R position difference at the central chamber plane for two hits on the same side (a) and opposite sides (b) of the wires.

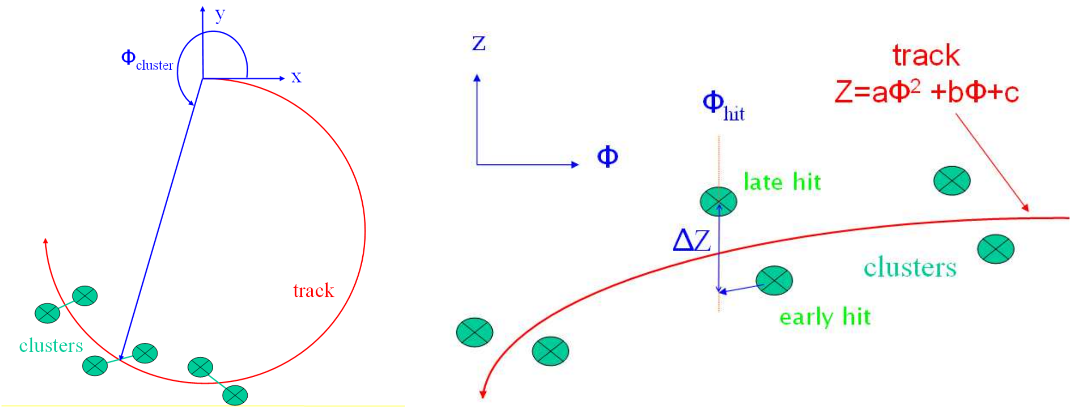
The difference in the measurements of the R resolution obtained from two hits on the same side and opposite sides is negligible. This fact, on the other hand, proves that both the channel to channel t_0 calibration and the time-to-distance relation calibration are done correctly and in good precisions.

The selection criteria applied in R resolution study are the following:

1. Only fitted, single-turn, TIC-matched tracks with DCH self determined track time satisfying, $|T_0| < 40$ ns.
2. Only two-hit clusters with two adjacent hits on two different planes.
3. No skipped chambers.
4. No clusters having too small drift distance (< 500 μm) hits, which typically means the TXY function fails to evaluate the hits.

8.1.2 Intrinsic Z Resolution

Similar to R resolution, the intrinsic Z resolution is inferred by the position difference of the two hits in two-hit clusters. The method is illustrated in Figure 8.2. In Z-Φ plane, the positron track is well described by a quadratic trajectory defined by three sequential clusters on the track. The 'early hit' is projected along the track direction to the Φ_{hit} position of the 'late hit' and the difference in Z, dZ, at this constant Φ_{hit}, is used as a measure of the intrinsic Z resolution. This technique is again free from the positron multiple scattering effects between the chambers.



(a) X-Y view (b) Z-Φ view

Figure 8.2: Illustration of Z resolution measurement method. The position difference ΔZ is used to infer the intrinsic Z resolution.

Distribution of dZ is shown in Figure 8.3, the core part of which is fitted with a Gaussian. The Gaussian has a σ_{dZ} of 0.182 cm, which corresponds to a single hit Z resolution of $\sigma_{dZ}/\sqrt{2} = 0.129$ cm, assuming both wires having the same Z resolution.

The selection criteria applied in Z resolution study are the same as those in R solution study listed in Section 8.1.1.

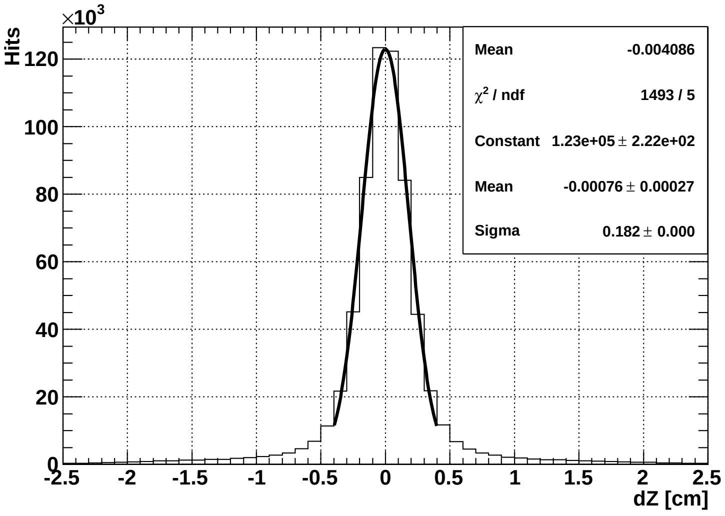
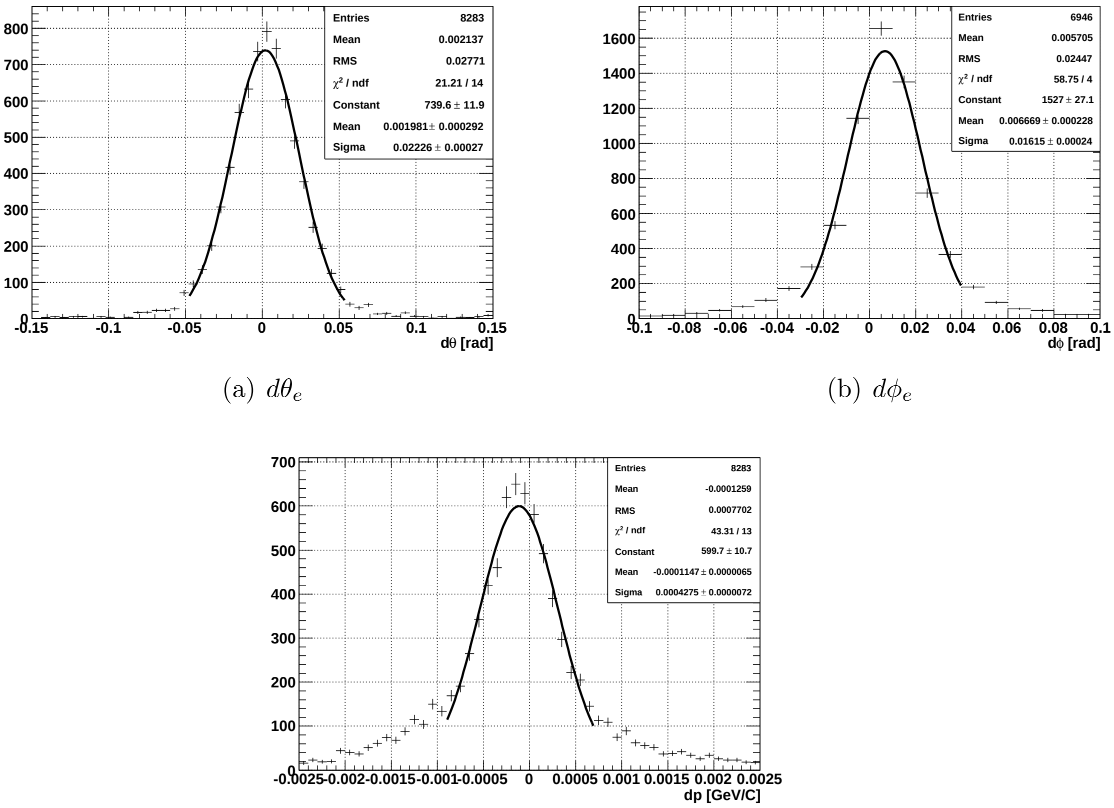


Figure 8.3: Distribution of position difference of the two hits in two-hit clusters projected on a constant Φ.

8.1.3 θ_e, ϕ_e, and Momentum Resolutions

Unlike R and Z positions, which are local measurements, θ_e, ϕ_e, and positron momentum, p_e, are measured globally by the tracks. θ_e and ϕ_e on the decay vertex are deduced by the track extrapolation onto the target plane, and the positron momentum is determined by the track trajectory. A technique using two-turn tracks to infer the θ_e, ϕ_e, and positron momentum resolutions is developed. It first selects events with well fitted two-turn tracks and then re-runs the track fitting algorithm with the track merging option off to have two separate tracks. The algorithm extrapolates the first turn track forward from the last chamber measurement and the second turn track backward from the first chamber measurement to a common point. Comparing the difference of the state vectors at that common point, $d\theta_e$, $d\phi_e$, and, dp_e are used as measures of the θ_e, ϕ_e, and momentum resolutions.

Distributions of $d\theta_e$, $d\phi_e$, and, dp_e are shown in Figure 8.4. The core parts of these three distributions are each fitted with a Gaussian. The fitted Gaussian to $d\theta_e$ dis-



(a) $d\theta_e$

(b) $d\phi_e$

(c) dp_e

Figure 8.4: Distributions of $d\theta_e$, $d\phi_e$, and, dp_e of two turns of a fitted track separately extrapolated to a common point

tribution has a $\sigma_{d\theta_e}$ of 0.022 rad corresponding to a single track θ_e resolution of $\sigma_{d\theta_e}/\sqrt{2} = 0.016$ rad; to $d\phi_e$ distribution has a $\sigma_{d\phi_e}$ of 0.016 rad corresponding to a single track ϕ_e resolution of $\sigma_{d\phi_e}/\sqrt{2} = 0.011$ rad; to dp_e distribution has a σ_{dp_e} of 428 keV/c corresponding to a single track momentum resolution of $\sigma_{dp_e}/\sqrt{2} = 303$ keV/c; assuming that both turns experience the same amount of multiple scattering and therefore have the same θ_e, ϕ_e, and momentum resolutions.

An alternative approach estimating the momentum resolution is to fit the end point of the observed Michel spectrum to the theoretical Michel spectrum convovled with a resolution function characterized as the sum of two Gaussians (core and tail components). Figure 8.5 shows the fitting result and the obtained momentum resolution function. The fit gives a core part of 81% with a resolution of $\sigma_{p_e} = 370$ keV. This result is consistent with the one obtained by the two-turn track method.

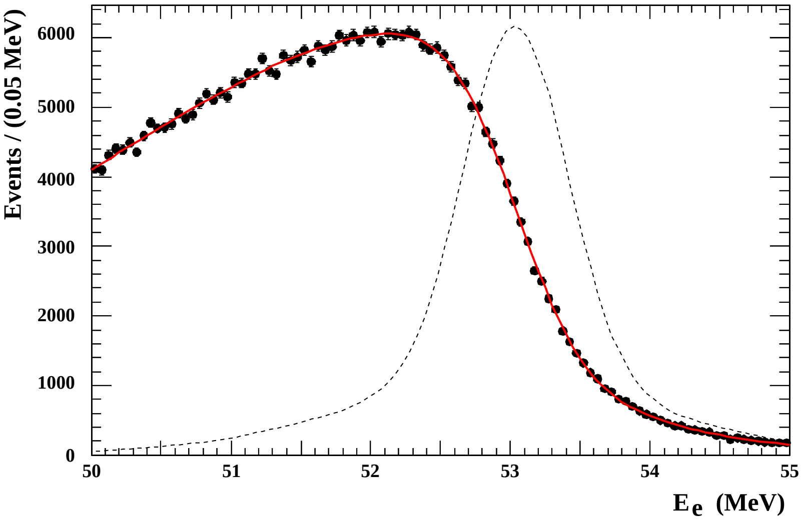
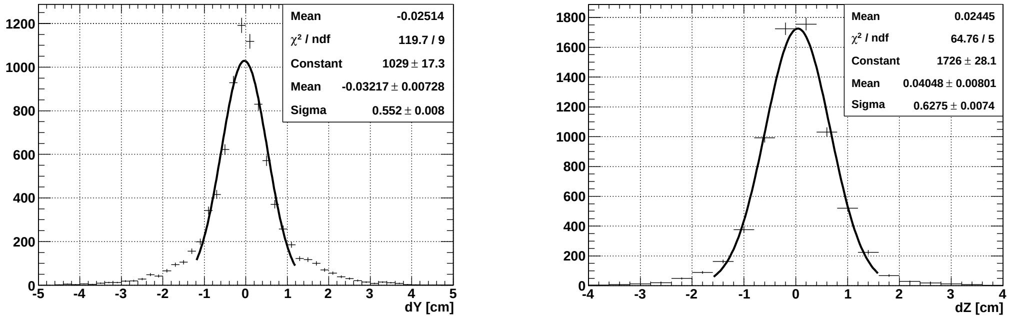


Figure 8.5: Fit of the theoretical Michel spectrum convovled with a resolution function to the observed spectrum. The fitting range is 51 to 55 MeV and the resolution function (dotted) is characterized by the sum of two Gaussians, representing the core and tail components, respectively.

8.1.4 Muon Decay Vertex Position Resolution

The muon decay vertex is assumed to be the intersecting point of the positron track extrapolated onto the muon stopping target. The target is mounted with a small slant angle (20.5°) from the z-axis. Therefore, a set (Y, Z) can best describe a given point on the target plane.



(a) dY (b) dZ

Figure 8.6: Distributions of dY and dZ using the two-turn track technique

Tracks typically have one and half turns from the muon target to the TIC. Suffered from multiple scattering effects along the positron trajectory, position resolution of the decay vertex indicates the quality of Y and Z determination. The technique of evaluating position resolution uses two-turn tracks, separately extrapolated to a common point. Position differences of the two turns, dY and dZ, give a good evaluation of such resolutions. Distributions of dY and dZ are shown in Figure 8.6. The core parts of these two distributions are each fitted with a Gaussian. The fitted Gaussian to dY distribution has a σ_{dY} of 5.5 mm corresponding to decay vertex Y resolution of $\sigma_{dY}/\sqrt{2} = 3.9$ mm; to dZ distribution has a σ_{dZ} of 6.3 mm corresponding to decay vertex Z resolution of $\sigma_{dZ}/\sqrt{2} = 4.5$ mm. This is based on the assumption that both turns have the same resolution.

8.2 TIC Performance

The intrinsic TIC time resolution is inferred by the time difference measured by two consecutive bars caused by the same positron passing through. Plot of intrinsic time resolution (estimated by $1/\sqrt{2}$ times the RMS σ of the time difference) vs. bar number is shown in Figure 8.7. The average resolution for single bar is estimated to be 67 ps.

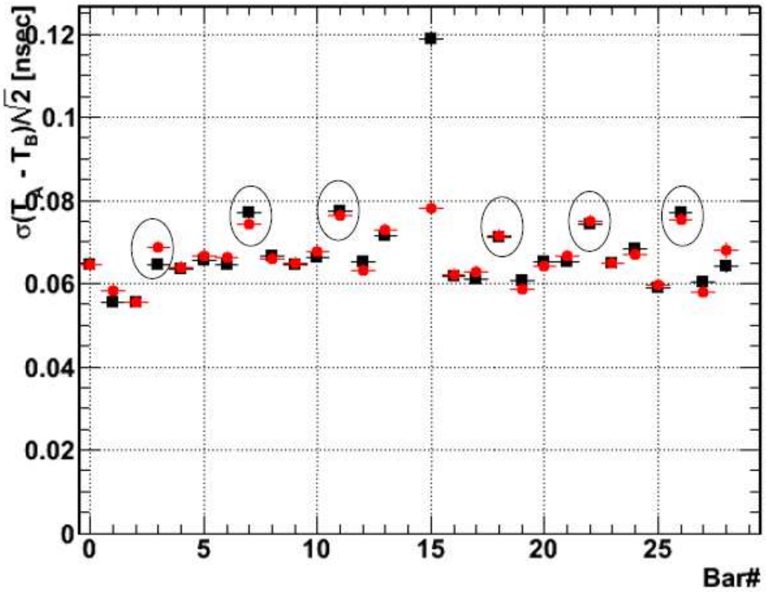


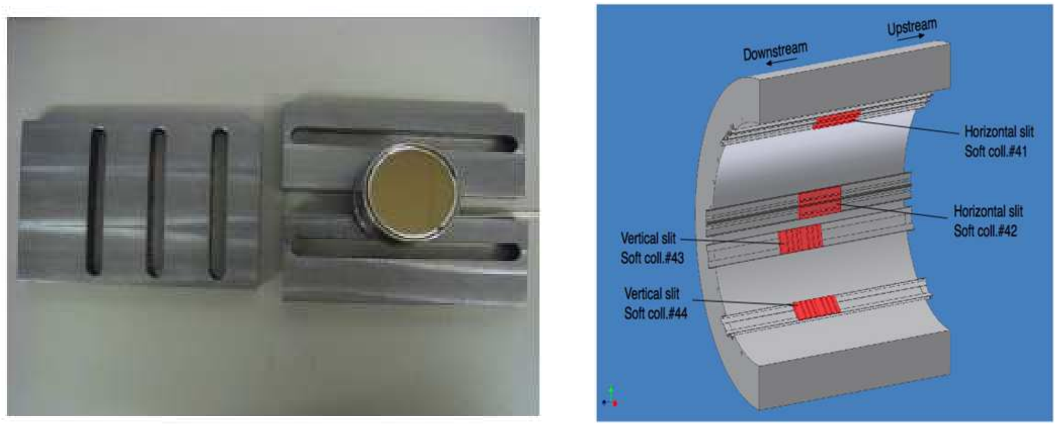
Figure 8.7: Intrinsic time resolution vs. bar number. Different marks show evaluations done with two different data samples taken at different run periods, September (black) and November (red). Discrepancy of bar 15 is due to a recovery of a bad channel. Channels marked with circles had DRS board changed between the two run periods.

8.3 XEC Performance

8.3.1 Position Resolution

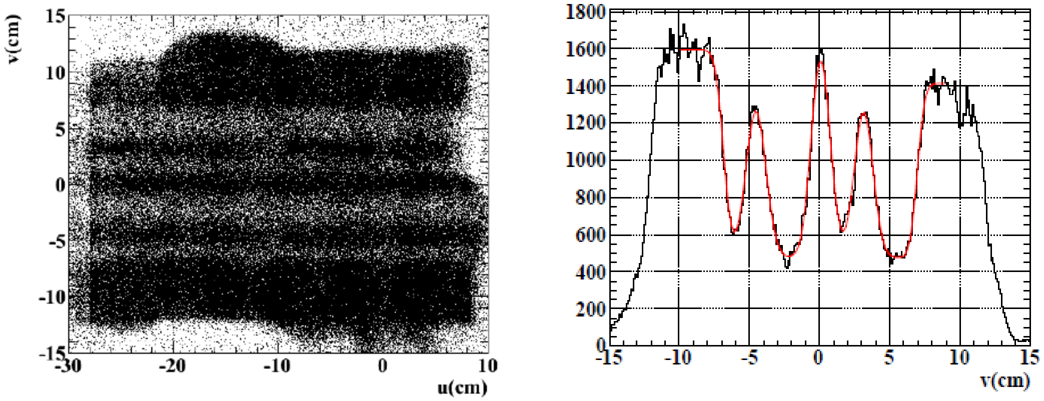
Photon position resolution is evaluated by measuring the shadow of lead bricks with several 1 cm wide slits mounted in front of the XEC entrance window in the π^0

runs. A picture of the lead brick and the configuration of its installation on the XEC entrance wall are shown in Figure 8.8.



(a) (b)

Figure 8.8: Lead bricks for photon position resolution study. (a), a picture of lead bricks compared with a PMT. (b), installation of lead bricks on the XEC entrance wall.



(a) (b)

Figure 8.9: Photon position resolution study. (a), 2D distribution of reconstructed photon position in local uv frame. (b), projection onto $v-$axis.

Some bricks are installed with slits parallel to u direction, while others with slits parallel to v direction. Figure 8.9(a) shows an example of the distribution of the reconstructed position with in the local uv plane, which has the image of a brick. Its projection onto, in this case, $v-$axis (Figure 8.9(b)) has three peaks corresponding to the three slits on the brick. Each slit peak is fitted with a Gaussian and each edge plateau with an error function. The average of the sigmas is $\langle\sigma\rangle = 6.9$ mm. However, the width of the fitted Gaussian contains effects from the width of slit, photon direction, depth of the conversion point as well as the contribution from the spread of π^0 decay point. A dedicated MC simulation with the same configuration is performed to deconvolve these effects and to investigate the response of position measurement. We find the effective resolution of ∼5 mm in both u and v directions on average and ∼6 mm in the depth direction (w).

Incorporating the photon position resolution and the uncertainty of the muon decay vertex, the angular resolutions of the photon direction are evaluated to be σ_{θ_γ} =9.9 mrad and σ_{ϕ_γ} =9.2 mrad, on average.

8.3.2 Intrinsic XEC Time Resolution

The intrinsic XEC time resolution is estimated by the difference of times measured by two groups of PMTs. 846 PMTs are numbered from 1 to 846 and are divided equally and evenly into two groups: odd numbered and even numbered PMTs. The photon time is reconstructed independently by these two groups of PMTs, termed as t_{odd} and t_{even}, respectively. Both odd and even groups are uncorrelated to each other, therefore, each group's timing resolution is $\sqrt{2}$ times of the intrinsic XEC time resolution. Hence, the uncertainty of half of the time difference between the two, $(t_{odd} - t_{even})/2$ is statistically equivalent to that of the XEC time.

Distribution of $(t_{odd} - t_{even})/2$ for 55 MeV photons from the π^0 runs is shown in Figure 8.10. It is fitted to a Guassian function with a $\sigma_{dt/2}$ of 45 ps, which is an estimate of the intrinsic XEC time resolution at 55 MeV.

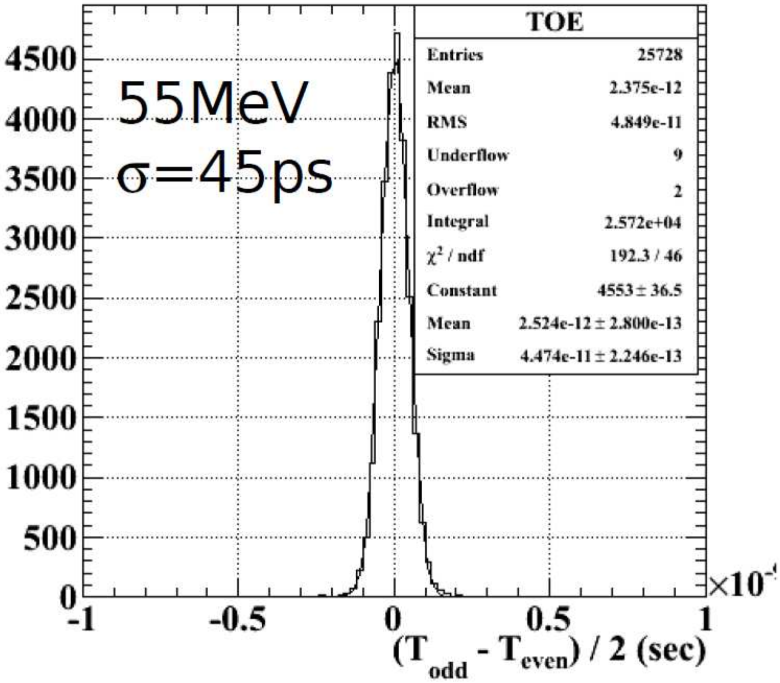
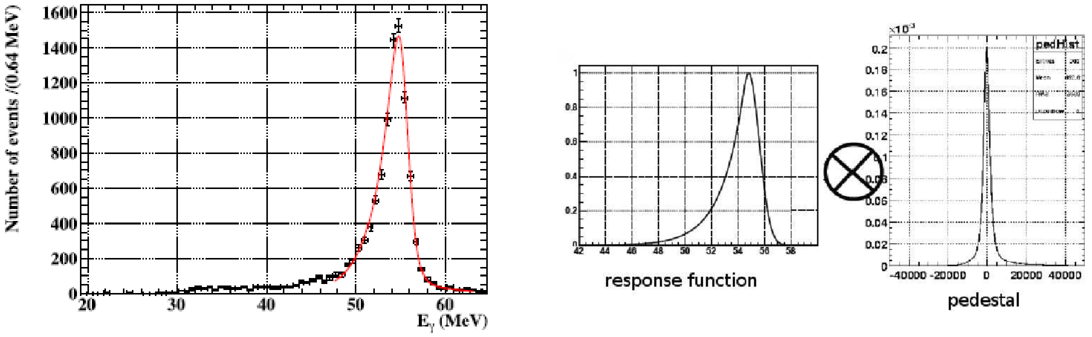


Figure 8.10: Distribution of $(t_{odd} - t_{even})/2$ for 55 MeV photons

The intrinsic XEC time resolution is estimated to be 36 ps at 83 MeV (the other back-to-back photon from π^0 run). This is because that the number of photoelectrons increases as the scintillation light increases. The intrinsic time resolution is completely dominated by the photo-statistics.

8.3.3 Photon Energy Resolution

Photon energy resolution is evaluated by using 55 MeV photons from π^0 runs. The distribution of the reconstructed 55 MeV photon spectrum is shown in Figure 8.11(a). It has a long lower tail mainly due to photon interacting with the XEC wall and losing some energy before entering the XEC active volume. It also has a higher tail component because of the higher energy photons in the π^0 decay. The reconstructed spectrum is fitted with a response function convovled with a pedestal function (Figure 8.11(b)). The response function is characterized by a Gaussian above the peak energy, and an exponential decay component below the peak energy. The fit is shown in Figure 8.11(a). It has a σ_{E_γ} of 850 keV as a measure of the photon energy resolution at 55 MeV, which corresponds to a resolution of 1.54%.



(a) (b)

Figure 8.11: Plots of photon energy resolution. (a) reconstructed photon energy spectrum for 55 MeV photons in π^0 decay. After unfolding the pedestal, the width of the distribution is a good measure of the energy resolution at 55 MeV. (b) fitting function built by a convolution of a response function and a pedestal distribution.

8.4 Combined Resolutions

8.4.1 Positron-Photon Relative Time Resolution

The experiment provides a direct measurement of the positron-photon relative time. The positron-photon timing coincidence is the signature of radiative decay. The resolution of positron-photon relative time is therefore measured by the dedicated radiative decay runs; it has a width of 202.2±9.7 ps. Details of the analysis technique and description is found in Section 10.1. The position of the relative time is confirmed by the Dalitz decay ($\pi^0 \rightarrow \gamma e^+ e^-$) positron-photon timing coincidence.

8.4.2 Positron-Photon Angular Resolution

The experiment does not provide a direct measurement of the photon emission angle. Assuming the photon travels from the muon decay vertex to the photon energy deposit point in the XEC, the angular resolution of the photon direction is evaluated to be around ~ 9 to ~ 10 mrad for both θ and ϕ direction. The evaluation also includes the uncertainty of the vertex position. The resolution of the positron-photon opening angle $\theta_{e\gamma}$ is then obtained by combining the angular resolutions of the two particles. It is evaluated to be $\sim$30 mrad.

8.5 Summary of Detector Resolutions

A summary of the detector resolutions is listed in Table 8.1

	Resolution (σ)
Positron energy, E_e	300 - 370 keV
Positron emission angle, (θ_e, ϕ_e)	(16, 11) mrad
Muon decay vertex, (y, z)	(3.9, 4.5) mm
TIC single bar time	67 ps
XEC interaction position (u, v and w)	(5, 5, 6) mm
XEC intrinsic timing	45 ps @ 55 MeV
Photon energy, E_γ	850 keV @ 55 MeV
Positron-photon relative time, $t_{e\gamma}$	202 ps
Positron-photon opening angle, $\theta_{e\gamma}$	$\sim$ 30 mrad

Table 8.1: Summary of detector resolutions in sigma.

Chapter 9

Event Selection

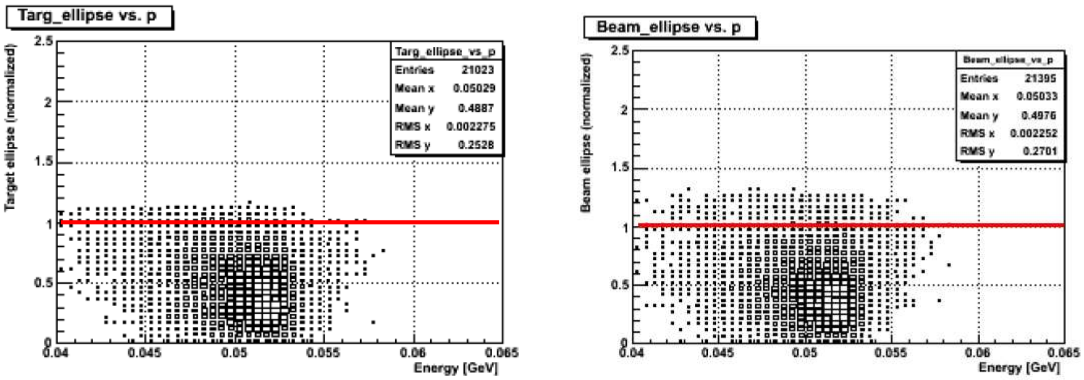
In order to reduce background and to ensure a good observation of radiative decay timing coincidence signal, a set of event selection criteria has to be identified. The selection criteria apply on both positron and photon analysis, which depend crucially on detector performance and the operating environment. They are applied in the physics analysis at all stages, serving as the common ground for all conditional probability calculations.

9.1 Positron Selection

In the study of identifying selection criteria for positron analysis, we use a sample of data that spans uniformly the full run period. We examine the distribution of certain selection criterion vs. positron momentum. We choose selection cuts to eliminate as many unphysical events (>55 MeV) as possible and, at the same time, to maintain a minimal loss of events in the physical region of interest(∼45-52.8 MeV). We also eliminate poorly measured positrons, which would contaminate the quality of selected data sample. The positron selection criteria further divide into several categories: decay vertex cuts, fiducial volume cuts, basic track cuts, track fitting quality cuts and the DCH-TIC matching cuts.

9.1.1 Decay Vertex Cuts

To ensure the muons decay on the stopping target, we require that the projected positron track end points intersecting the target plane are within the target ellipse (whose dimensions are set by the target geometry). Because of the air doping in COBRA, the beam distribution on the target plane is slightly off-centered. We further require that positron track end points on the target should also be within an ellipse that is featured by the muon beam profile. The major and minor axes of the beam ellipse are obtained from the beam distribution in Y and Z directions, and are set to roughly 2σ of the beam spot size. Distributions of normalized target and beam ellipses vs. positron energy are shown in Figure 9.1.



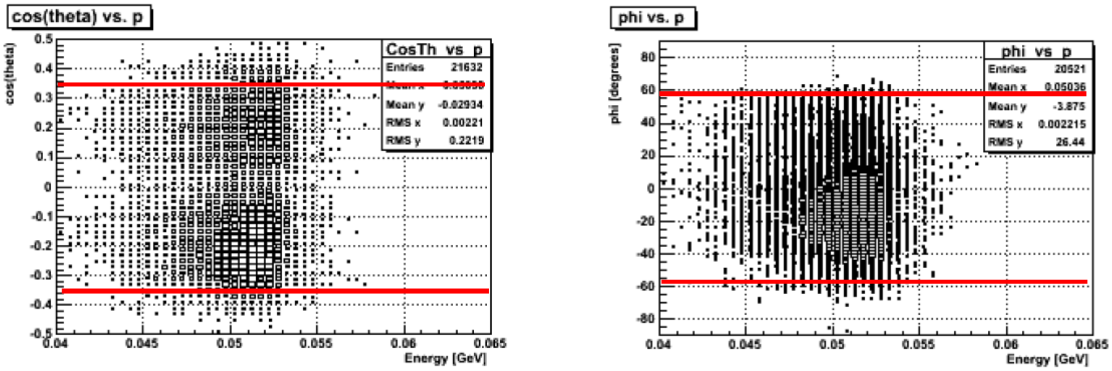
(a) Normalized target ellipse vs. E_e

(b) Normalized beam ellipse vs. E_e

Figure 9.1: Decay vertex cuts

9.1.2 Fiducial Volume Cuts

The spectrometer covers roughly 9% of the full solid angle. The configuration of the spectrometer fiducial volume is defined by the detector geometry. This requires the positrons coming off the target with $|\cos\theta_e| < 0.35$ and $|\phi_e| < 57.3°$. Distributions of reconstructed positron $\cos\theta_e$ and ϕ_e at the decay vertex vs. positron energy are shown in Figure 9.2.



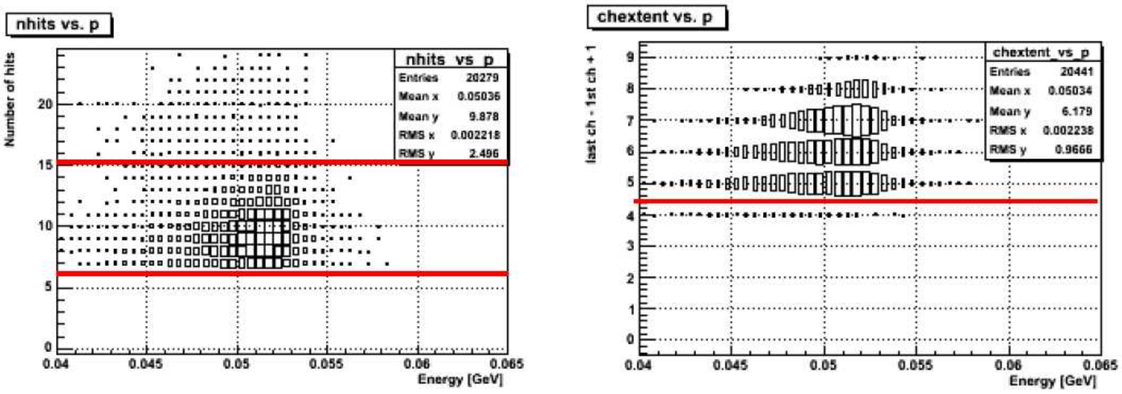
(a) $\cos\theta_e$ on target vs. E_e

(b) ϕ_e on target vs. E_e

Figure 9.2: Fidicial volume cuts

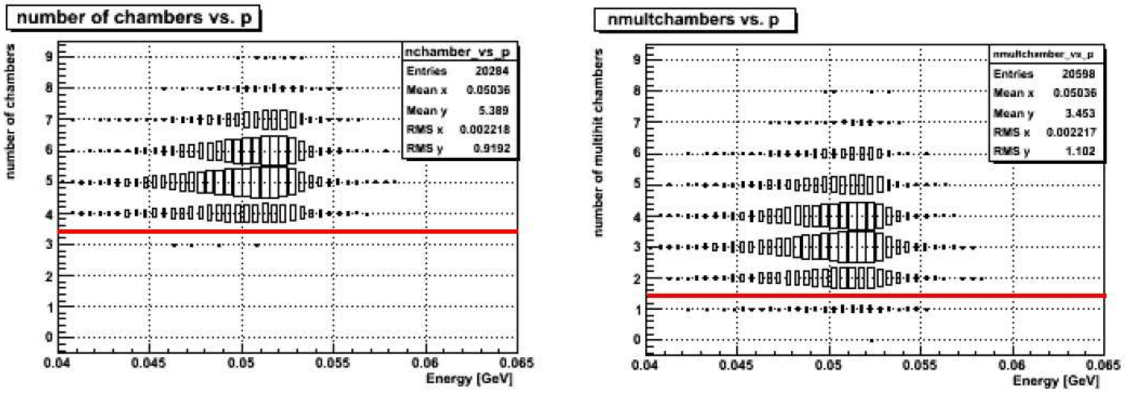
9.1.3 Basic Track Cuts

For each track, we require the following: total number of hits to be no less than 7 ($n_{hit} \geq 7$); number of chambers that a track passes through to be between 5 and 9 ($5 \leq n_{chext} \leq 9$); number of chambers that have hits to be no less than 4 ($n_{ch} \geq 4$); and number of chambers that have multiple hits to be no less than 2 ($n_{multch} \geq 2$). Distributions of these quantities vs. positron energy are shown in Figure 9.3. The first two cuts nearly have full efficiency. Tracks that failed the last two cuts do not have enough information to deduce the DCH self-determined T_0.



(a) n_{hit} vs. E_e

(b) n_{chext} vs. E_e



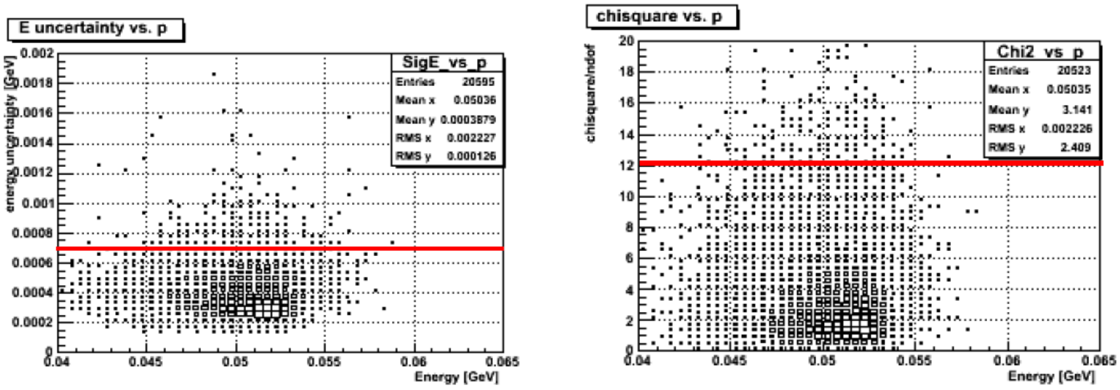
(c) n_{ch} vs. E_e

(d) n_{multch} vs. E_e

Figure 9.3: Basic track cuts

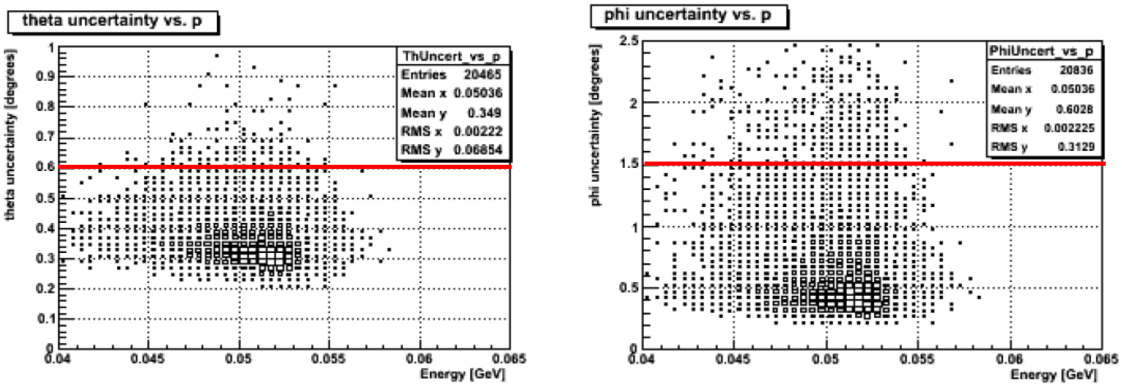
9.1.4 Track Fitting Quality Cuts

In the fitting procedure, the Kalman filter evaluates the quality of the fit by providing uncertainties of physical observables as well as the normalized $\hat{\chi}^2$ per degree of freedom of the fit. We require the following criteria on the fitted track to select good fitting quality: positron energy uncertainty, $\delta E_e < 0.7$ MeV; track normalized $\hat{\chi}^2 < 12$; uncertainty of positron θ_e at decay vertex, $\delta\theta_e < 0.6°$; and uncertainty of positron ϕ_e at decay vertex, $\delta\phi_e < 1.5°$. Distributions of these quantities vs. positron energy are shown in Figure 9.4.



(a) δE_e vs. E_e

(b) normalized $\hat{\chi}^2$ vs. E_e



(c) $\delta\theta_e$ vs. E_e

(d) $\delta\phi_e$ vs. E_e

Figure 9.4: Track fitting quality cuts

The Kalman filter fitting algorithm also merges tracks into multiple-turn ones, i.e.

tracks having more than one turn in the spectrometer, if possible. The COBRA magnet was designed to sweep positrons out of the spectrometer with typically only one to two turns. Tracks that have more than two turns are either by a failure in the merging algorithm or by accident that two or even more tracks are close in space and time. We thus eliminate tracks having three or more turns. The distribution of the number of positron turns vs. positron energy is shown in Figure 9.5.

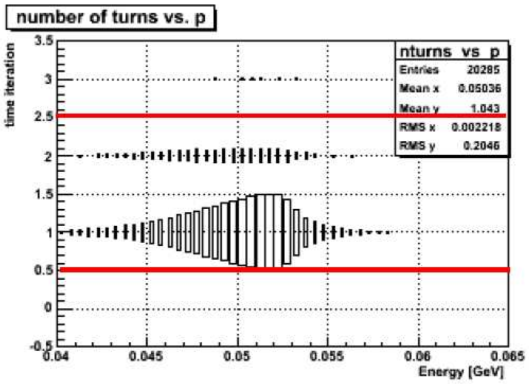
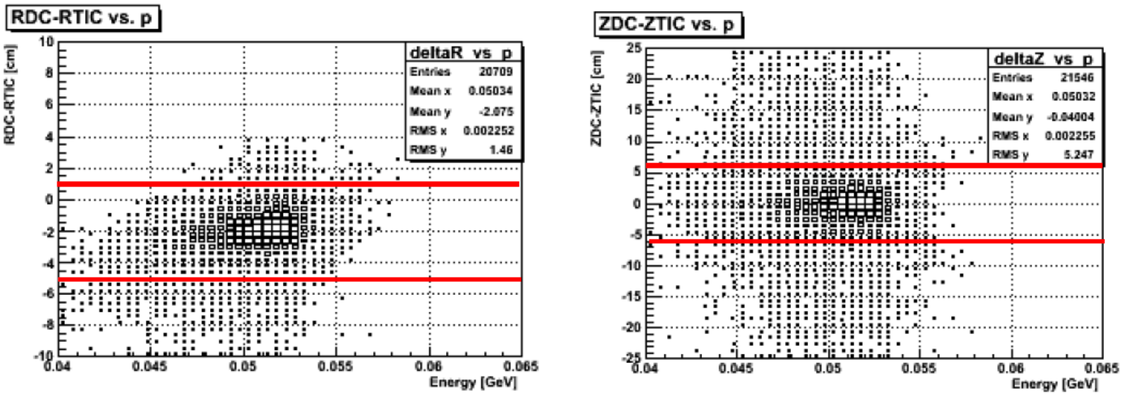


Figure 9.5: n_{turns} vs. E_e

9.1.5 DCH-TIC matching cuts

The DCH-TIC connection is crucial to our analysis. It connects a DCH track with its correlated TIC hit, which serves as the basis of the positron timing. The DCH-TIC connection algorithm extrapolates the fitted track to the ϕ of every TIC hit, then it calculates $dZ_{DCH-TIC} = Z_{DCH} - Z_{TIC}$ and $dR_{DCH-TIC} = R_{DCH} - R_{TIC}$, between the track projected end point on the TIC and the TIC hit. Therefore, we apply matching cuts on $dZ_{DCH-TIC}$ and $dR_{DCH-TIC}$ to select good DCH-TIC connections. The matching is affected by scattering and energy loss in materials between the DCH and TIC such as chamber support structure, cables, pre-amplifiers, etc. Thus, $dZ_{DCH-TIC}$ and $dR_{DCH-TIC}$ distributions are slightly off-centered. We require that

$|R_{DCH} - R_{TIC} - \delta_R| < 3cm$ and $|Z_{DCH} - Z_{TIC} - \delta_Z| < 6cm$ (Figure 9.6) as criteria of good DCH-TIC matching, where δ_Z and δ_R are the $dZ_{DCH-TIC}$ and $dR_{DCH-TIC}$ distribution offsets from zero. We call these two cuts spacial matching criteria.



(a) dR_{DC-TIC} vs. E_e (b) dZ_{DC-TIC} vs. E_e

Figure 9.6: DCH-TIC spacial matching cuts

Besides spacial matching between the DCH and TIC, we also require that the DCH track and the TIC hit have a time match. Besides the DCH self-determined T_0, each track also has a track time that is determined by the associated TIC hit, which is the TIC-determined track time, T_0^e. We center the distribution of T_0^e by offsetting the global DHC-TIC time difference. Since T_0^e is determined by the TIC hit that is external to DCH. The TIC has a much better timing. T_0^e distribution typically has a narrow peak. However, it is possible that the TIC hit used to reference the track T_0^e is in fact irrelevant to the track even though they may be spatially adjacent; in which case the T_0^e will be an outlier in the distribution. Therefore, we require $|T_0^e| < 12$ ns (Figure 9.7) as a criterion of a good DCH-TIC timing match.

Distributions of all selection cut variables vs. positron momentum are shown from Figure 9.1 to Figure 9.7.

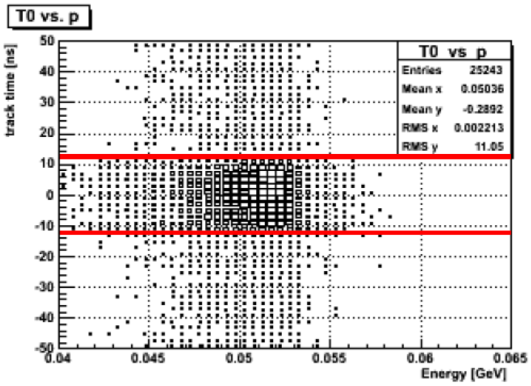


Figure 9.7: Positron T_0^e vs. E_e

9.1.6 Ghost Track Selection

In the tracking stage, some tracks may have essentially the same hits, differed by only a few. These tracks after fitting, extrapolation and DCH-TIC connection processes may still exist, and furthermore, may even use the same TIC hit to extract timing information. At this point, these tracks share enough information so that they should not be considered as independent tracks. A set of tracks that share sufficient information are grouped as a unique track. This group of tracks is referred as a set of ghost tracks.

An effort has been put to identify and resolve ghost tracks. The issue though is two-folded. If certain tracks share some hits in an event, one has to first identify whether they are ghost tracks to each other or they are truly separate tracks; and among the ghost tracks, one has to select the one that could best represent that unique track to use in the analysis. Two conditions must be satisfied to define any two tracks as ghosts to each other:

1. have the same track time, i.e. they both have to have good DCH-TIC connection with the same TIC hit and use this time as their time reference

2. number of chambers which have at least one hit in common must be more than half of the number of hit chambers on shorter one of the two

We then rank the ghost tracks to select one that best represents the unique track. The normalized $\hat{\chi}^2$ and the chamber span (S) of a track are the two most straightforward quality measurements of a track. We construct the following rank order to represent the track quality,

$$Rank = \frac{1}{S} + \alpha\hat{\chi}^2 \tag{9.1}$$

The smaller the value is, the higher a ghost track ranks. The goal is to choose α so that the 2nd term dominates the ranking when two ghost tracks have very different $\hat{\chi}^2$ and the 1st dominates when two ghost tracks have similar $\hat{\chi}^2$. The coefficient α is tuned to be 0.0025.

After applying all selection cuts, if there is still more than one ghost of a unique track survived in an event, we choose the highest ranking ghost as the positron track used in the analysis.

9.2 Photon Selection

The active volume of the XEC detector defines the photon fiducial volume (Figure 9.8), which is written, in the local coordinate system, as,

$$|u| < 25 \text{ cm}, \quad |v| < 71 \text{ cm}, \quad 0 < w < 38.5 \text{ cm}. \tag{9.2}$$

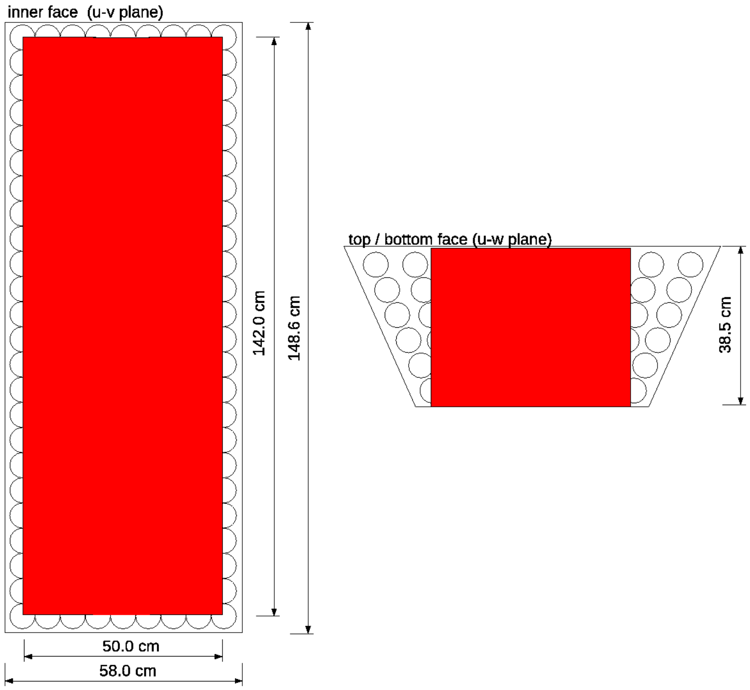


Figure 9.8: Photon fiducial volume indicated as the red area

For a radiative decay event, the photon has a certain probability entering its fiducial volume given the positron entering the spectrometer fiducial volume. This conditional probability is termed as conditional photon geometrical acceptance. Therefore, a fiducial volume cut is not applied in the photon selection, but implied by the conditional photon geometrical acceptance in the radiative decay branching fraction calculation (see 10.5.1).

To ensure the quality of the event and to reject cosmic background, pileup rejection and cosmic elimination are applied in the photon selection, which are covered in the following sections.

9.2.1 Pileup Rejection

In the analysis, we have algorithms to identify pileup events and to recover them by eliminating the overlapping photons as described in Section 6.5.3. Pileup events are identified by two methods: light distribution ($N^{light\ peak} > 1$) and time distribution ($\hat{\chi}_t^2 > 3$). The elimination algorithm is not applied if an event is only identified as pileup by time distribution but not by light distribution. Therefore, such event are rejected as pileup events by the criterion $\hat{\chi}_t^2 < 3 \vee N^{light\ peak} > 1$.

We also reject pileup events if the elimination fails. We require $0.9 < E_\gamma^1/E_\gamma^0 < 1$ as a criterion of proper working elimination, where E_γ^0 is the reconstructed energy before eliminating the second photon and E_γ^1 is that after. Figure 9.9 shows the effect of pileup rejection.

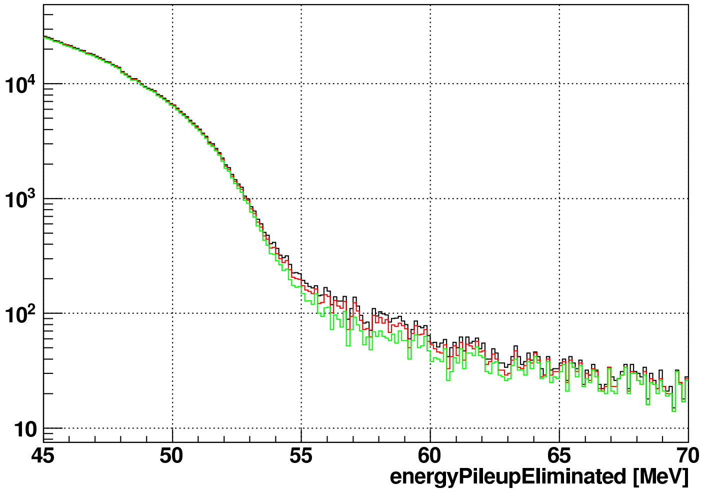


Figure 9.9: Effect of pileup cut. Black line shows the original reconstructed photon spectrum, red line is applied with $E_\gamma^1/E_\gamma^0 < 1$, and green line is applied additionally with $0.9 < E_\gamma^1/E_\gamma^0 < 1$

9.2.2 Cosmic Elimination

Sometimes cosmic rays having hits in both XEC and TIC are able to cause a radiative decay trigger. Some of these events are well reconstructed and recognized as radiative decay events due to the presence of a positron track in both timing and spatial coincidence.

There are two possible configurations for cosmic rays coming from the top to cause a radiative decay trigger, as shown in Figure 9.10. Due to the hardware geometry, most cosmic rays hit the XEC first and enter the XEC from its outer face. They have the first interaction point closer to the outer face, and deposit a large energy near it. Therefore, the ratio of the light observed by the inner face PMTs to that by the outer face ones is different from the ratio for photons generated from the target (i.e. entering the XEC from its inner face). We require $N_{pho}^{inner}/N_{pho}^{outer} > 0.3$ to distinguish a normal photon entering the XEC from the inner face from a cosmic ray entering the XEC from the outer face. Application of this cut successfully reduces the rate of cosmic-ray background below that of photons from muon decays.

However, cosmic events with cosmic rays entering the XEC from its inner face are similar to photons from muon decay and hence are not eliminated by this cut. They are rejected by the constraint of radiative decay kinematics, which is described in Section 10.1.2.

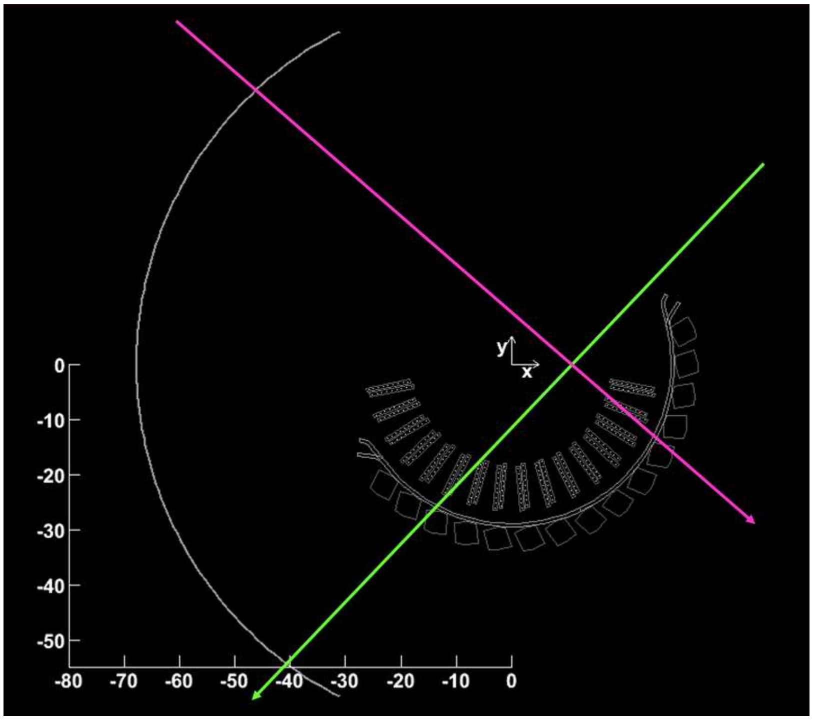


Figure 9.10: Cosmic rays coming from the top have two possible configurations hitting both XEC and TIC in order to cause a radiative decay trigger. Purple ray hits the XEC first and then the TIC, which enters the XEC from its outer face and hence has the first interaction point to deposit energy close to the outer face. Green ray hits the TIC first and then the XEC. It enters the XEC from the inner face like radiative decay photons generating from the target.

Chapter 10

Radiative Muon Decay Analysis

In this chapter, a detailed analysis on measuring the branching fraction of the radiative muon decay is presented. We first show the timing coincidence signal, which is the event signature of the radiative decay. We further describe the normalization scheme developed in the circumstance that a normalization trigger was not taken during the data taking. We then calculate detector efficiencies that are dependent on the detected particle energies.

To measure the branching fraction of $\mu \rightarrow e\bar{\nu}\nu\gamma$ decay, we adopt a cut-and-count approach. We count the number of radiative decay events in the signal region after subtracting background and measure the total number of muon decay by counting Michel decays in the data sample. Comparisons between the experiment result and the theoretical prediction is discussed.

As mentioned in Section 5.1, there are two types of data that were taken with a requirement on the positron-photon timing coincidence, i.e. the MEG trigger data and the dedicated RD trigger data. However, only the latter suits the purpose of measuring the RD branching fraction. The MEG trigger data does not serve this purpose because it had a requirement on the direction matching of positron and photon in addition to their timing coincidence. The direction matching condition was based on Monte Carlo simulation and was used during the DAQ. At the stage of offline analysis, it is impractical to correct the effect introduced by the direction matching[1]. Therefore, the MEG trigger data are not used.

10.1 Observation of Radiative Decays

10.1.1 Timing Coincidence

The dedicated RD runs were taken with a muon stopping rate of roughly $1.2 \times 10^6 \mu/s$. Figure 10.1 shows all trigger types taken.

The photon-positron time difference at the decay vertex ($t_{e\gamma} = t_\gamma - t_e$) is shown in Figure 10.2. The histogram has a bin size of 100 ps, where there is a clear radiative decay timing coincidence signal at 25.8 ns. It is notable that there is a bump at 29.5 ns in the distribution, which is identified by the geometry as the cosmic ray events of the second configuration described in Section 9.2.2.

---

[1] The direction matching favors back-to-back positron-photon pairs and thus introduces an angular trigger efficiency, which depends on the positron-photon opening angle. To measure the RD branching fraction, this has to be corrected. However, the experiment does not provide enough information to correct this.

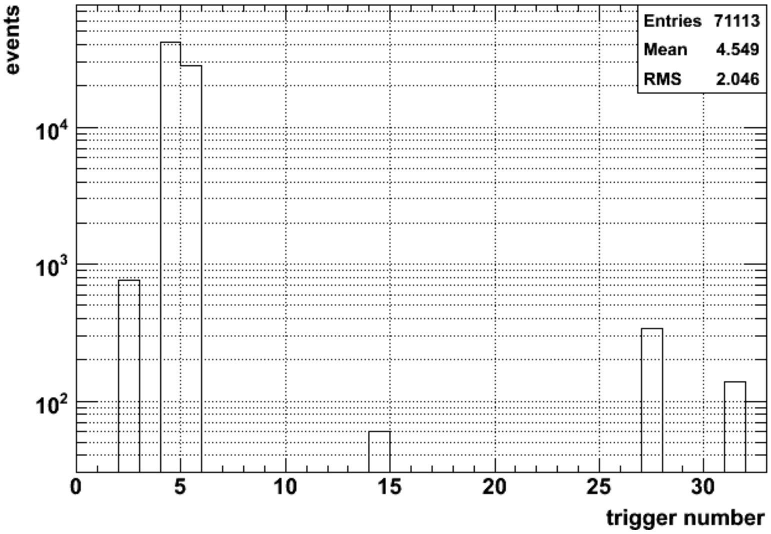


Figure 10.1: Triggers taken in the dataset used in the analysis

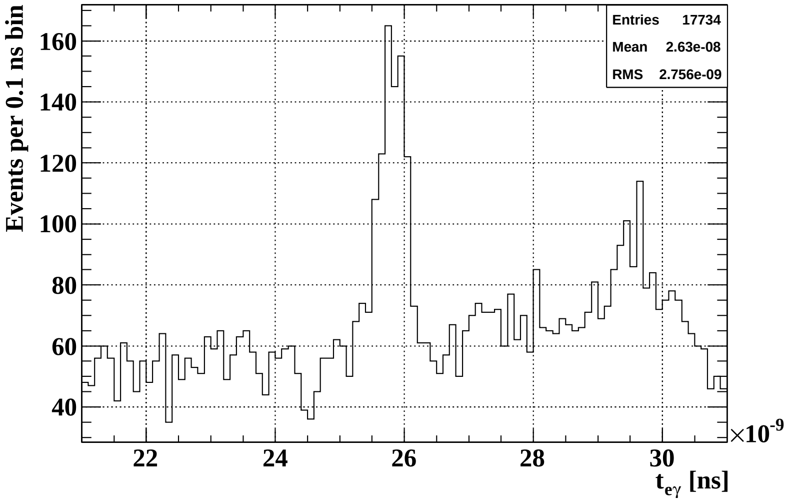


Figure 10.2: Distribution of $t_{e\gamma}$

10.1.2 Kinematic Constraint

Even after applying the selection criteria, most of the events in the remaining dataset are still inconsistent with originating from a muon decaying at rest, and are largely accidental events and some cosmic ray induced events. We eliminate these events by applying a kinematic cut. We calculate the energy of a single neutrino (or in other word, the sum of energies of two neutrinos coming off in parallel) that would be needed to balance the vector sum of the photon and positron momenta; this represents the minimum additional energy in the neutrinos necessary to ensure that the putative muon is at rest. Hence, the sum of the single neutrino, photon, and positron energies should give a lower limit on the rest mass of a particle that would decay at rest to produce the observed photon and positron. A diagram of the kinematic cut is illustrated in Figure 10.3. The distributions of single neutrino energy and of the putative muon mass are shown in Figure 10.4. Only events with putative muon mass below the muon rest energy ($M_\mu = 105.6$ MeV) could in fact have resulted from a single muon decay; the others are primarily accidental coincidences. Therefore, the kinematic constraint can be formulated as:

$$E_{tot} = E_e + E_\gamma + E_{\nu\nu} < M_\mu = 105.6 \text{ MeV}. \tag{10.1}$$

In practice, due to the reconstruction resolution effect, we choose $E_{tot} < 110$ MeV. Note that the application of the kinematic cut should not change the number of observed radiative decays, however, it will reduce the background more and thus improve the signal to background ratio.

The distribution of $t_{e\gamma}$ after applying the kinematic cut is shown in Figure 10.5. The histogram also has a bin size of 100 ps and is fitted with a normalized Gaussian func-

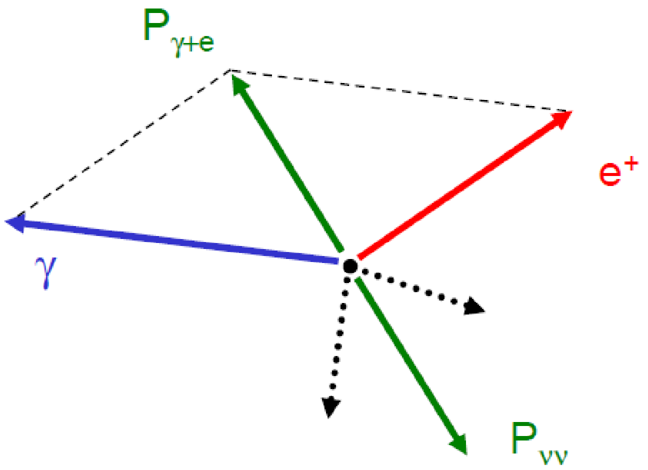
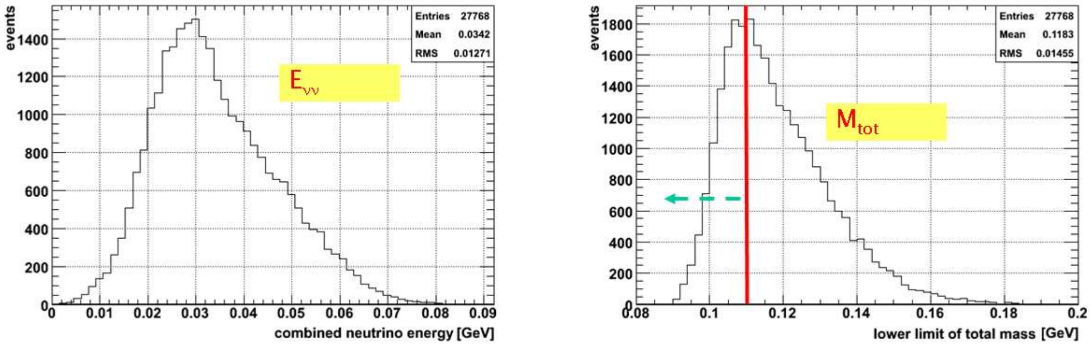


Figure 10.3: Illustration of the kinematic cut. Two dashed black arrows represent possible neutrino directions. If 2 neutrinos have zero opening angle; or effectively a single neutrino has momentum $P_{\nu\nu}$ balancing the vector sum of the photon and positron momenta, $P_{\gamma+e}$, this single neutrino (or the sum of the two neutirnos coming off in parallel) gives the minimum energy required. Otherwise additional momentum is needed to balance the transverse component.



(a) $P_{\nu\nu}$

(b) Distribution of the putative muon mass, which sets the lower limit of the total mass in the decay process. The kinematic cut is marked in the red line. Most of the events are not consistent with originating from a muon decaying at rest.

Figure 10.4: Distributions of the effective single neutrino energy ($P_{\nu\nu}$) and the putative muon mass

tion with a flat background. The radiative decay timing coincidence signal appears at $T_{e\gamma} = 25.8$ ns, the same as in Figure 10.2. It has a σ RMS width of 202.2 ± 9.7 ps, which is quoted as the detector resolution of positron-photon relative time. Comparing to Figure 10.2, the application of the kinematic constraint dramatically improved the signal to noise ratio. The bump of cosmic ray events that originally appeared at 29.5 ns is also removed by the kinematic constraint as well.

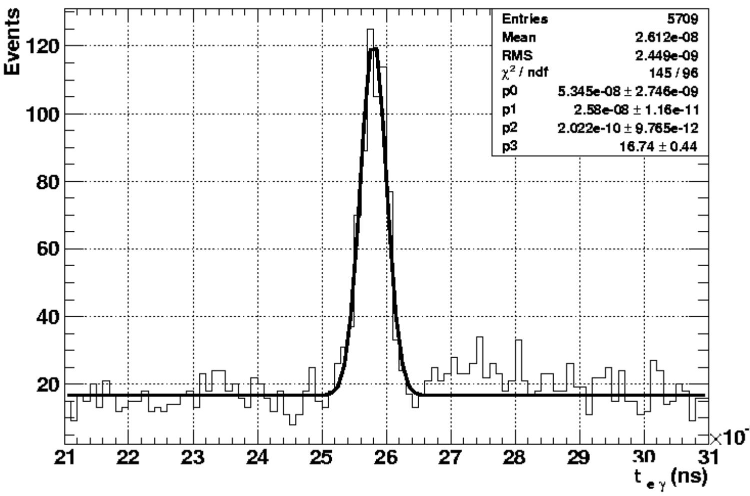
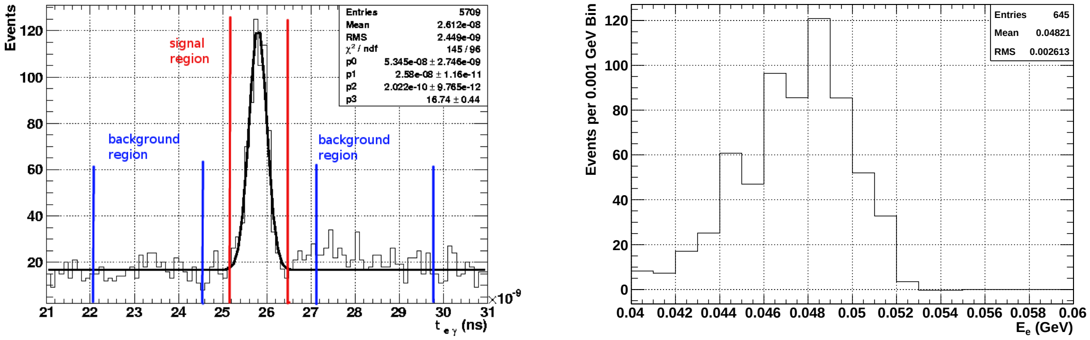


Figure 10.5: Radiative decay signal: distribution of $t_{e\gamma}$ after applying the kinematic cut

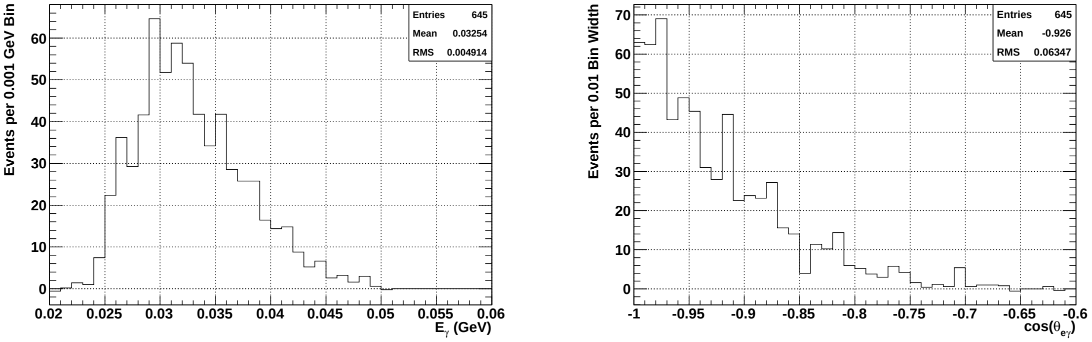
10.1.3 Kinematic Distributions of Signal Events

Figure 10.5 shows a clear timing coincidence signal over a flat background, which consists of accidental events. For the purpose of background subtraction, we divide the time domain around the timing coincidence signal peak time ($T_{e\gamma}$) into sections. We took ($T_{e\gamma} \pm 0.6$ ns) as the signal region, ($T_{e\gamma} + 1.2$ ns, $T_{e\gamma} + 4.2$ ns) and ($T_{e\gamma} - 4.2$ ns, $T_{e\gamma} - 1.2$ ns) as the background regions (Figure 10.6(a)). Note that the background region is 5 times as wide as the signal region. We apply a standard background subtraction for a signal peak over a flat background by subtracting the expected number of background events determined from the fit. The kinematic distributions of signal events are shown in Figure 10.6 after subtraction of background.



(a) background subtraction illustration

(b) signal events E_e



(c) signal events E_γ

(d) signal events $\cos\theta_{e\gamma}$

Figure 10.6: Distributions of kinematic variables of signal events

10.2 Normalization Sample

We summarized the theoretical calculation of unpolarized radiative muon decay in Section 2.1. In practice, recall that a branching fraction is defined as the ratio of the decay rate of a particular process, e.g. $\Gamma_{\mu\to e\nu\bar{\nu}\gamma}$ in case of radiative muon decay, to the total decay rate Γ_{tot}: $B(\mu \to e\nu\bar{\nu}\gamma) \equiv \Gamma_{\mu\to e\nu\bar{\nu}\gamma}/\Gamma_{tot}$. For muon decays, the total decay rate is nearly saturated by that for Michel decays, $\Gamma_{\mu\to e\nu\bar{\nu}}$, with a very small contribution from radiative decays ($B(\mu \to e\nu\bar{\nu}\gamma) \sim 1.4\pm0.4\%$ for $E_\gamma > 10$ MeV [4]).

However, in the MEG experiment, for Michel decays, only the final positron can be observed, so from the positron point of view, Michel decays and the radiative decays are indistinguishable. Thus, $\Gamma_{\mu\rightarrow e\nu\bar{\nu}(\gamma)} = \Gamma_{tot} \equiv 1$. As a convention, for normalization purposes, we use Michel decays to refer to both decay modes in the sense of resulting in a positron as one of the decay products.

Our normalization technique is developed using the radiative decay trigger data. In the positron reconstruction algorithm, at the tracking stage, each unique positron track has a DCH self-determined T_0. This T_0 calculation is not associated with any TIC hit, but measured only by the hits found on the track. Figure 10.7 shows the DCH self determined T_0 distribution of all positron tracks with E_e above 43 MeV in the entire RD trigger data. The peak in the T_0 distribution consists of tracks related to the triggered TIC hits. The flat tail, for example from 100 to 180 ns, consists of accidental positrons falling into the DCH DRS window, which are not associated with the TIC hits that caused the trigger. These tracks can be used for Michel analysis as the normalization sample.

The hardware trigger jitter is less than 40 ns, therefore, 100 ns away from the peak is sufficient to avoid bias from the trigger jitter. However, besides the accidental positron tracks, tracks in fact associated with the TIC hits that caused the trigger but having poorly determined DCH T_0 could also possibly enter this normalization sample window. This effect has a small contribution, and will be addressed. Figure 10.8 shows the DCH T_0 distribution in the normalization sample window from 100 to 180 ns, which is fitted with a first order polynomial. The error on the slope of the first order polynomial fit shows the fit is consistent with a constant. Therefore, these tracks are, to first order, accidental Michel positrons falling into this 80 ns window as we expected.

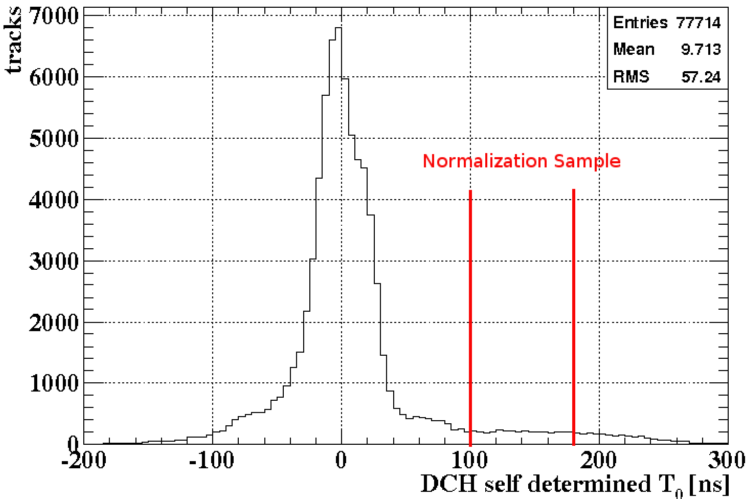


Figure 10.7: DCH self determined T_0 distribution of the RD trigger data. The left tail does not extend out as far as the right tail is due to a pre-set hit search window in the hit finding which extends further out in the positive time direction.

We may need to correct for the overflow of tracks in the normalization window due to poor determination of DCH T_0. Figure 10.9 shows the distribution of of the time difference between the TIC determined track time T_0^e (i.e. the TIC hit time) and DCH self determined T_0. The plot was done with the RD trigger data by requiring a spatial match between the DCH track and the TIC hit which caused the trigger. The core part is fitted which a Gaussian that has a σ of 6.4 ns, which means using the TIC time as a reference the DCH self determined T_0 has a precision of 6.4 ns. The DCH-TIC time offset is -513.9 ns (the fitted peak time). So if a DCH T_0 is in fact close to zero (i.e the track is in fact associated with the TIC hit that caused the trigger) but is calculated incorrectly by 100 ns later in time, it will enter the $T_0^e - T_0$ plot of Figure 10.9 100 ns before the peak. We call these tracks overflow tracks which have to be subtracted when counting the normalization positrons in the normalization window.

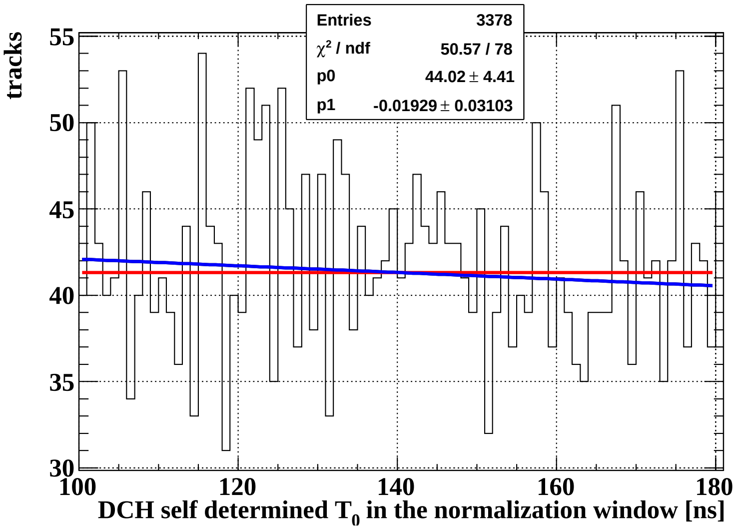


Figure 10.8: DCH T_0 distribution in the normalization sample window from 100 to 180 ns. It is fitted with a constant (red), and a first order polynomial (blue). The error on the slope of the first order polynomial fit shows the fit is consistent with a constant. Therefore, these tracks are accidental Michel positrons falling into this 80 ns window.

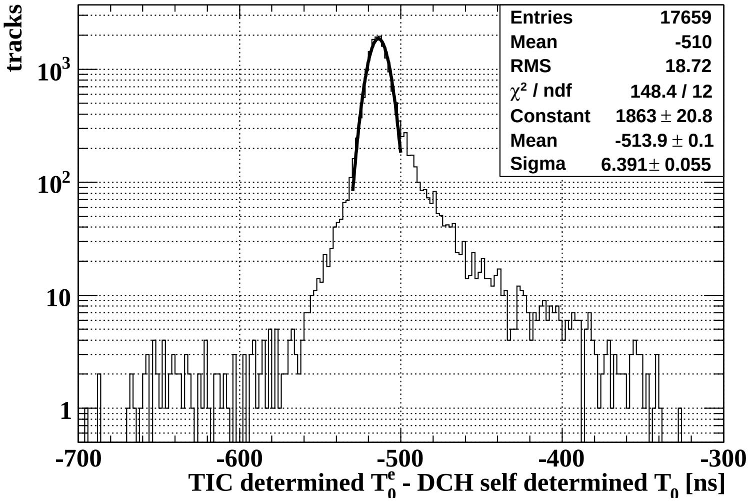


Figure 10.9: Distribution of the time difference between the TIC determined track time T_0^e and DCH self determined T_0

Let us define N_w as the number of tracks in the 80 ns window 100 ns prior to the peak time (i.e. from $\sim$ -690 to $\sim$ -610 ns) in Figure 10.9. N_w consists of two contributions, the overflow tracks ($N_{overflow}$) as well as a type of accidental tracks (N_{acc}). This type of accidental track does not relate to the the required TIC hit that caused the trigger but happens to satisfy the spatial matching criteria. Thus, we have $N_w = N_{overflow} + N_{acc}$. We define λ and ρ as ratios of N_w and $N_{overflow}$ to the number of entries in the peak in Figure 10.9. If η is the accidental matching probability of an irrelevant track that happens to spatially match with a particular TIC hit in a 80 ns window, we have $\lambda = \rho + \eta$. η is estimated to be 0.05%[2]. From Figure 10.9, we calculate that $\lambda = 0.32\%$. Finally we get $\rho = 0.27\%$ as a good estimator for the fraction of tracks that should be in the peak in Figure 10.7 but have poorly determined DCH T_0 and overflow into the 80 ns normalization sample window 100 ns later from the peak. For $E_e > 43$ MeV, the number of tracks in the peak region[3] in Figure 10.7 is 60311, and the number of tracks in the 80 ns normalization window is 3378. If we define κ as the fraction of overflow correction to be applied to the normalization sample, $\kappa = \rho \times 60311/3378 = 4.8\%$. Hence the normalization sample should be reduced by 4.8% due to poor DCH tme measurement.

Positron energy distribution of tracks having DCH self-determined T_0 between 100 to 180 ns and $E_e > 43$ MeV is shown in Figure 10.10.

The systematic uncertainty of measuring the normalization sample mostly comes from two parts, η and λ. Estimation of η relies on the assumption that the muon target

---

[2] Some information needed for this calculation are detailed in later sections. We summerize them here: beam intensity $1.2 \times 10^6 \mu/s$; positron geometrical acceptance of the spectrometer 9%; absolute DCH reconstruction efficieny $\sim$40% at this beam intensity for $E_e > 40$ MeV assuming the target stopping power being 100% (An estimation of the absolte DCH efficiency is done in Section 10.4.4 for a different E_e range); Z cut on DCH-TIC matching 6 cm; Z range of the TIC bar 80 cm. Therefore, $\eta = 1.2 \times 10^6 s^{-1} \times 9\% \times 40\% \times \frac{2 \times 6cm}{80cm} \times 80$ ns $= 0.05\%$.

[3] We take region $|T_0| < 40$ ns as the peak region.

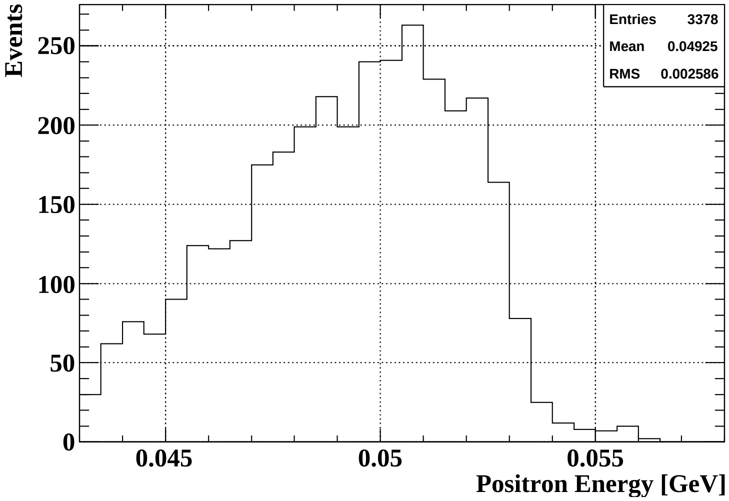


Figure 10.10: Positron energy distribution of tracks having DCH self determined T_0 between 100 to 180 ns

stopping power efficiency is 1. It is a non-measurable quantity in the experiment. If we assign a 50% uncertainty on η, it translates to a resulting systematic uncertainty of the normalization sample of 0.45%. Thus we have,

$$\delta_{syst}^{norm\ \eta} \sim 0.45\%. \tag{10.2}$$

On the other hand, uncertainty of λ can be from either statistics of number of tracks in the 80 ns window 100 ns prior to the peak time or from the systematics. The uncertainty from the statistics is $\sim$ 14%, which should be the dominant uncertainty. So on the conservative side, we assgin the total uncertainty of λ to be 20%. This translates to a resulting systematic uncertainty of the normalization sample of 1.1%. Thus we have,

$$\delta_{syst}^{norm\ \lambda} \sim 1.1\%. \tag{10.3}$$

These two parts are added in quadratures. Therefore, we find the total systematic uncertainty on the normalization sample to be

$$\delta_{syst}^{norm} \sim 1.2\%. \tag{10.4}$$

10.3 Normalization Scheme

One can write the number of detected Michel decays as a product of a series of measurable quantities,

$$N_M^{e,D} = R^{\mu} \cdot \tau_M^L \cdot \varepsilon(stop) \cdot B_M^e \cdot G_M^e \cdot \left\langle \varepsilon^{\mathrm{DCH}} \right\rangle \cdot \varepsilon_M^{frac} \tag{10.5}$$

where $N_M^{e,D}$ is the number of detected Michel positrons; R^{μ} is the muon beam rate; τ_M^L is the trigger live time in counting Michel events; $\varepsilon(stop)$ is the target stopping efficiency; B_M^e is the branching fraction of all decays resulting in a positron as one of the products, thus $B_M^e \equiv 1$; G_M^e is the spectrometer geometrical acceptance for Michel positrons subjected to the experiment setup; $\varepsilon^{\mathrm{DCH}}$ is the DCH acceptance and reconstruction efficiency (DCH efficiency); and ε_M^{frac} is the fraction of Michel positron momentum spectrum that is used to count Michel decays.

Similarly, for radiative decays, the number of detected decays can be written as

$$\begin{aligned} N_{e\nu\bar{\nu}\gamma}^{e\gamma,D} &= R^{\mu} \cdot \tau_{e\nu\bar{\nu}\gamma}^L \cdot \varepsilon(stop) \cdot B(\mu \rightarrow e\nu\bar{\nu}\gamma) \\ &\cdot G_{e\nu\bar{\nu}\gamma}^e \cdot \left\langle \varepsilon^{\mathrm{DCH}} \right\rangle \cdot \left\langle \varepsilon^{\mathrm{TIC}} \right\rangle \cdot \left\langle \varepsilon^{\mathrm{TIC\ Trig}} \right\rangle \\ &\cdot \left\langle G_{e\nu\bar{\nu}\gamma}^{\gamma(e)} \right\rangle \cdot \left\langle \varepsilon^{\mathrm{XEC}} \right\rangle \cdot \left\langle \varepsilon^{\mathrm{XEC\ Trig}} \right\rangle \end{aligned} \tag{10.6}$$

The factors are defined and labeled the same way as for Michel decays in Equation 10.5. Additional factors are defined in the following: ε^{TIC} is the TIC acceptance and efficiency factor (TIC acceptance) which is defined as a conditional probability of finding a correlated TIC hit given a well reconstructed DCH track; $\varepsilon^{\text{TIC Trig}}$ is the TIC trigger efficiency which is defined as a conditional probability of causing a timing counter trigger given a well reconstructed DCH track as well as a correlated timing counter hit; $G^{\gamma(e)}_{e\nu\bar{\nu}\gamma}$ is the conditional geometrical acceptance for the photon entering the XEC volume given the positron from the same radiative decay going into the spectrometer, which is subjected to both the experiment geometry as well as the decay kinematics; ε^{XEC} is the XEC detection and reconstruction efficiency; $\varepsilon^{\text{XEC Trig}}$ is the XEC trigger efficiency which is defined as a conditional probability of setting a XEC trigger given a well reconstructed photon in the XEC detector.

The factors ε^{TIC} and $\varepsilon^{\text{TIC Trig}}$ appear here in Equation 10.6 in the calculation of $N^{e\gamma,D}_{e\nu\bar{\nu}\gamma}$, but not in that of $N^{e,D}_{M}$ in Equation 10.5, because the Michel positrons we use in the analysis are from the tail events in the T_0 distribution, which are only required to pass DCH but not necessarily to hit the TIC or even further to cause a TIC trigger. ε^{DCH}, ε^{TIC} and $\varepsilon^{\text{TIC Trig}}$ form an efficiency chain of positron reconstruction. Among those, ε^{DCH} and ε^{TIC} are positron momentum dependent, i.e. $\varepsilon^{\text{DCH}}(E_e)$ and $\varepsilon^{\text{TIC}}(E_e)$; while $\varepsilon^{\text{TIC Trig}}$, to first order, only depends on the positron energy deposited in the TIC scintillation bars. The experiment geometry constrains positrons reaching the timing counter typically to have energy bigger than 40 MeV. Compared to the threshold of the energy deposit required to cause a trigger, at this energy level, the conditional timing counter trigger efficiency $\varepsilon^{\text{TIC Trig}}$ is independent of positron energy.

Similarly, ε^{XEC} and $\varepsilon^{\text{XEC Trig}}$ form an efficiency chain of photon reconstruction. Both are photon energy dependent, i.e. $\varepsilon^{\text{XEC}}(E_\gamma)$ and $\varepsilon^{\text{XEC Trig}}(E_\gamma)$.

The conditional geometrical acceptance $G^{\gamma(e)}_{e\nu\bar{\nu}\gamma}$ only depends on the opening angle between the positron and photon, i.e. $G^{\gamma(e)}_{e\nu\bar{\nu}\gamma}(\cos\theta_{e\gamma})$.

Equations 10.5 and 10.6 share common factors, R^{μ} and $\varepsilon(stop)$, which are independent from event-to-event variations. Because we deal with an ensemble of events in the cases of both Michel decays and radiative decay signals, efficiency factors in both equations are in fact weighted average efficiencies, integrated over the correct input spectra, respectively[4]. Combining Equations 10.5 and 10.6, aligning their common factors, leaving $\varepsilon^{\text{TIC Trig}}$ as a constant factor, and solving for $B(\mu \rightarrow e\nu\bar{\nu}\gamma)$, we get that

$$\begin{aligned} B(\mu \rightarrow e\nu\bar{\nu}\gamma) = \frac{N^{e\gamma,D}_{e\nu\bar{\nu}\gamma}}{N^{e,D}_{M}} \cdot B^{e}_{M} \cdot \varepsilon^{frac}_{M} \cdot \frac{\tau^{L}_{M}}{\tau^{L}_{e\nu\bar{\nu}\gamma}} \cdot \frac{G^{e}_{M}}{G^{e}_{e\nu\bar{\nu}\gamma}} \cdot \frac{\left\langle \varepsilon^{\text{DCH}}_{M} \right\rangle}{\left\langle \varepsilon^{\text{DCH}}_{e\nu\bar{\nu}\gamma} \right\rangle} \cdot \frac{1}{\varepsilon^{\text{TIC Trig}}_{e\nu\bar{\nu}\gamma}} \\ \cdot \frac{1}{\left\langle \varepsilon^{\text{TIC}}_{e\nu\bar{\nu}\gamma} \right\rangle} \cdot \frac{1}{\left\langle G^{\gamma(e)}_{e\nu\bar{\nu}\gamma} \right\rangle} \cdot \frac{1}{\left\langle \varepsilon^{\text{XEC}} \right\rangle} \cdot \frac{1}{\left\langle \varepsilon^{\text{XEC Trig}} \right\rangle} \end{aligned} \tag{10.8}$$

In Equation 10.8, we have put sub-scripts on positron related efficiency factors to

---

[4] This means, for example, $\langle \varepsilon^{\text{XEC}} \rangle$ and $\langle \varepsilon^{\text{XEC Trig}} \rangle$ are integrated over different spectra: $\langle \varepsilon^{\text{XEC}} \rangle$ is integrated over the input photon spectrum of the experiment, and $\langle \varepsilon^{\text{XEC Trig}} \rangle$, due to its definition as a conditional probability, is integrated over the spectrum of the well reconstructed photons in XEC. This gives $\langle \varepsilon^{\text{XEC}} \rangle \cdot \langle \varepsilon^{\text{XEC Trig}} \rangle = \langle \varepsilon^{\text{XEC}} \cdot \varepsilon^{\text{XEC Trig}} \rangle$. It can be proved as follows. Let us first ignore the resolution effect,

$$\begin{aligned} \langle \varepsilon^{\text{XEC}} \rangle \cdot \langle \varepsilon^{\text{XEC Trig}} \rangle &= \frac{\int I(E)\varepsilon^{\text{XEC}}(E)dE}{\int I(E)dE} \cdot \frac{\int I(E)\varepsilon^{\text{XEC}}(E)\varepsilon^{\text{XEC Trig}}(E)dE}{\int I(E)\varepsilon^{\text{XEC}}(E)dE} \\ &= \frac{\int I(E)\varepsilon^{\text{XEC}}(E)\varepsilon^{\text{XEC Trig}}(E)dE}{\int I(E)dE} \\ &= \langle \varepsilon^{\text{XEC}} \cdot \varepsilon^{\text{XEC Trig}} \rangle \end{aligned} \tag{10.7}$$

where $I(E)$ is the input photon spectrum and $I(E)\varepsilon^{\text{XEC}}(E)$ is the spectrum of well-reconstructed photons in XEC. To correctly handle the resolution effect, one has to either carefully define the efficiencies as functions of the observed energy or, in some cases, define the efficiencies as functions of the input energy with the condition that the observed spectrum to be used has to be de-convolved with the resolution function.

The same argument applies to positron efficiencies. And it is obvious that, for instance, $\langle \varepsilon^{\text{TIC}} \rangle \cdot \langle \varepsilon^{\text{XEC}} \rangle = \langle \varepsilon^{\text{TIC}} \cdot \varepsilon^{\text{XEC}} \rangle$, simply because they are integrated over different variables while averaging.

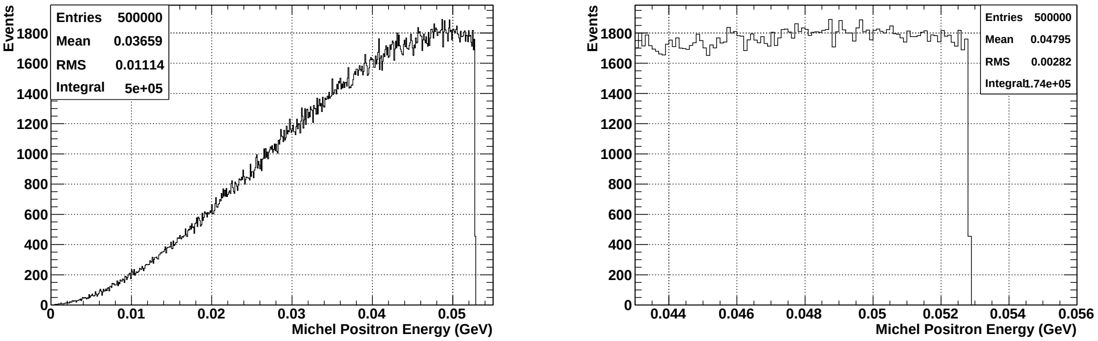
denote two decay modes. It is apparent that this method allows us to cancel certain common factors, which are typically hard to determine in data. This much simplifies the analysis. The first factor on the right hand side of the equation, B^e_M, is unity as discussed earlier. The factor $G^e_M/G^e_{e\nu\bar{\nu}\gamma}$, is unity as well, since the geometrical acceptance for positron, without any requirement on photon direction, only depends on the spectrometer coverage, which is about 9% of the solid angle, thus should be independent of the decay mode. As mentioned previously, timing counter trigger efficiency is independent of positron energies, which makes the factor $1/\varepsilon_{e\nu\bar{\nu}\gamma}^{\mathrm{TIC\ Trig}}$ a constant for all positron energy in the region of interest. This factor should be very close to unity, because the threshold of the energy deposit required for causing a TIC trigger is low.

The DCH efficiency $\varepsilon^{\mathrm{DCH}}(E_e)$ is energy dependent. $\left\langle\varepsilon_M^{\mathrm{DCH}}\right\rangle$ is the average DCH efficiency subjected to Michel spectrum, and $\left\langle\varepsilon_{e\nu\bar{\nu}\gamma}^{\mathrm{DCH}}\right\rangle$ to the radiative decay positron spectrum. They appear in pairs in the numerator and denominator in factor $\left\langle\varepsilon_M^{\mathrm{DCH}}\right\rangle / \left\langle\varepsilon_{e\nu\bar{\nu}\gamma}^{\mathrm{DCH}}\right\rangle$. Therefore, it is not necessary to calculate the absolute DCH efficiency, which is very difficult to measure in the data. Instead, a relative DCH efficiency is enough for the purpose of calculating the radiative decay branching fraction.

All the other factors in the equation are either experiment observables or directly measurable from data or Monte Carlo. The exact calculations of each factor are detailed in the following sections.

10.4 Positron Analysis

The positron energy distribution of tracks having DCH self determined T_0 between 100 to 180 ns and $E_e > 43$ MeV is shown in Figure 10.10. This is for positrons after applying all event selection criteria except the DCH-TIC matching cuts. The DCH has optimal acceptance at MEG signal positron ($E_e = 52.8$ MeV), since the experiment was designed to search for $\mu \rightarrow \mathrm{e}\gamma$ events. Therefore, E_e distribution falls at lower energy where the DCH has worse acceptance and efficiency. It also has a tail that extends out to 56 MeV which is due to resolution effect. For comparison purpose, Figure 10.11 shows the Michel spectrum obtained from the Monte Carlo simulation. Spectra for both the full kinematic range and for $E_e > 43$ MeV are presented.



(a) Full range Michel spectrum

(b) Michel spectrum at $E_e > 43$ MeV

Figure 10.11: Michel spectrum with radiative correction generated by the Monte Carlo simulation

10.4.1 Number of Detected Michel Positrons

The number of detected Michel positrons, $N_M^{e,D}$, is a direct observable of Michel decays. It can be easily counted in Figure 10.10. However, the statistics at lower energy is not great, thus will introduce big statistical error. Therefore, we want to select positrons with $E_e > 48$ MeV, where the DCH acceptance of these positrons is roughly flat and high. In Equation 10.5, in order to calculate the number of muon stops, N^{μ}_{stop}, one has to carry out all other factors in a way that is consistent in counting the number of detected Michel positrons, $N^{e,D}_{M}$. The calculations of the other factors in Equation 10.5 will be articulated in details in the following sections. It should be pointed out here that counting $N^{e,D}_{M}$ in different energy regions will have different fractions of the Michel spectrum used in counting the normalization Michel decays, ε^{frac}_{M}, and will also have different impacts on the DCH efficiency, $\varepsilon^{\mathrm{DCH}}$, and the TIC acceptance factor, $\varepsilon^{\mathrm{TIC}}$, both of which are positron energy dependent. Therefore, for the purposes of cross check and estimating the systematic uncertainty, we select $N^{e,D}_{M}$ in two different energy regions, i.e. $E_e > 48$ MeV and $E_e > 50$ MeV. Counting in the histogram of Figure 10.10 and applying the overflow correction κ, we get

$$N^{e,D}_{M}(E_e > 48 \text{ MeV}) = 2210 \tag{10.9}$$

and

$$N^{e,D}_{M}(E_e > 50 \text{ MeV}) = 1395 \tag{10.10}$$

10.4.2 Michel Counting Fraction

The factor ε^{frac}_{M} is the fraction of Michel positrons in the energy region used in counting the normalization Michels with respect to the entire kinematic range. It can be directly calculated from theoretical Michel spectrum formula. Alternatively, it can also derived from Michel spectrum generated from the Monte Carlo simulation with radiative decay correction (Figure 10.11(a)). Dividing number of entries in regions of

$E_e > 48$ MeV and $E_e > 50$ MeV by the total entries in the full range, we have

$$\varepsilon_M^{frac}(E_e > 48 \text{ MeV}) = 17.32\% \tag{10.11}$$

and

$$\varepsilon_M^{frac}(E_e > 50 \text{ MeV}) = 10.08\% \tag{10.12}$$

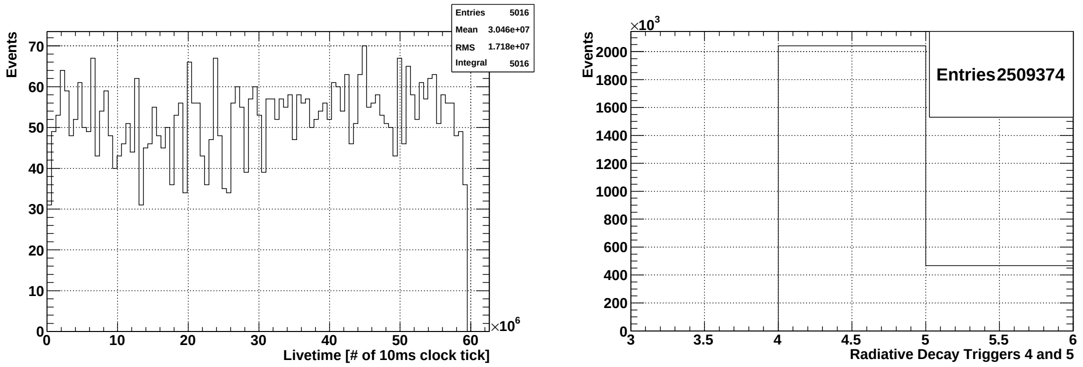
10.4.3 Live Time

There is a 10 ms clock tick built in the trigger system and the trigger live time is registered as the number of clock ticks accumulated event by event. An example plot of this cumulative registration of trigger live time is shown in Figure 10.12(a), where the example run is arbitrarily chosen. Note that since it is a cumulative registration, it is the entry of the last event, instead of the integral of the histogram, that gives the total live time of a run. For this particular example run, the total live time is 594.167 s.

We sum over the live time of all the radiative decay data to get the total live time $\tau^L_{e\nu\bar{\nu}\gamma} = 499677$ s. Since for every event, we only count Michel positron in an 80 ns window (100 ns $< T_0 <$ 180 ns), Michel counting live time should be 80 ns multiplied by the total number of the RD trigger events, $\tau^L_M = 80 \text{ ns} \times 2509374 = 0.201$ s, where 2509374 is the number of radiative decay run events used in the analysis, which is obtained in Figure 10.12(b). Therefore, we have that

$$\frac{\tau^L_M}{\tau^L_{e\nu\bar{\nu}\gamma}} = 4.018 \times 10^{-7} \tag{10.13}$$

The factor $\tau^L_M/\tau^L_{e\nu\bar{\nu}\gamma}$ is in fact equivalent to a prescale factor.



(a) (b)

Figure 10.12: Variables used in calculation of counting prescale factor. (a), accumulative registration of trigger live time, counted in number of 10 ms clock ticks. Note that it is the entry of the last event, instead of the integral of the histogram, gives the total live time of a run. (b), total number of the RD trigger events.

10.4.4 Relative DCH Efficiency

The DCH efficiency is a combination of efficiencies in both DCH acceptance and reconstruction, which are inseparable in the view of analysis. As discussed earlier, we must calculate the DCH efficiency as a function of positron energy, $\varepsilon^{\text{DCH}}(E_e)$, and then further apply it in the radiative decay branching fraction calculation (Equation 10.8) with the correct ensembles of events representing the two decay modes. As we have addressed earlier, a relative DCH efficiency is sufficient for the purpose of branching fraction calculation. Therefore, we have developed a method to derive the relative DCH efficiency as a function of positron energy, $\varepsilon^{\text{DCH}}(E_e)$.

MEG data were taken with a substantially higher beam rate at 3×10^7 μ/s. It is expected that the absolute DCH efficiency may get worse comparing to that of the low beam rate due to more overlaps and pileups in both space and time at higher rate and thus, resulting a higher background level. Additionally, RUN 2008's chamber operation was very sensitive to the beam condition. Higher beam intensity brought more positrons entering the spectrometer, which made more chambers tripped frequently and thus, worsened the chamber condition. Nonetheless, the relative DCH efficiency should not depend on the beam intensity, because the pileup, background and chamber condition affect the positron reconstruction in the same manner in the energy range of interest. We calculate $\varepsilon^{\mathrm{DCH}}(E_e)$ using LED trigger (trigger 14) events taken along with MEG data taking in high intensity beam. The LED trigger is used primarily as a XEC calibration trigger. Since it is not triggered on any positron information, it is a random trigger for the purpose of positron analysis. Thus, reconstructed positrons in these events should demonstrate the Michel spectrum subjected to the DCH efficiency with no bias.

We take the ratio of the reconstructed positron spectrum to the theoretical Michel spectrum numerically calculated by the Monte Carlo with appropriate energy range and binning. We thus obtain an energy bin by energy bin relative DCH efficiency, $\varepsilon^{\mathrm{DCH}}(E_e)$. The result of $\varepsilon^{\mathrm{DCH}}(E_e)$ using this method is shown in Figure 10.13. We have fitted it to a Gaussian function. Note that it is not normalized, and the vertical axis does not have a meaning; it only shows the magnitude of the ratio of the number of reconstructed positron to that of Monte Carlo generated. The mean and σ of the fitted Gaussian function contain the full parameterization of $\varepsilon^{\mathrm{DCH}}(E_e)$.

There is a subtlety in the technique of getting $\varepsilon^{\mathrm{DCH}}(E_e)$. Positrons above the kinematic limit, 52.8 MeV, are unphysical. But due the the resolution effect the data spectrum does not end at the Michel end point of 52.8 MeV. A deconvolution has to be done on the data spectrum before taking the ratio, by which we can adequately

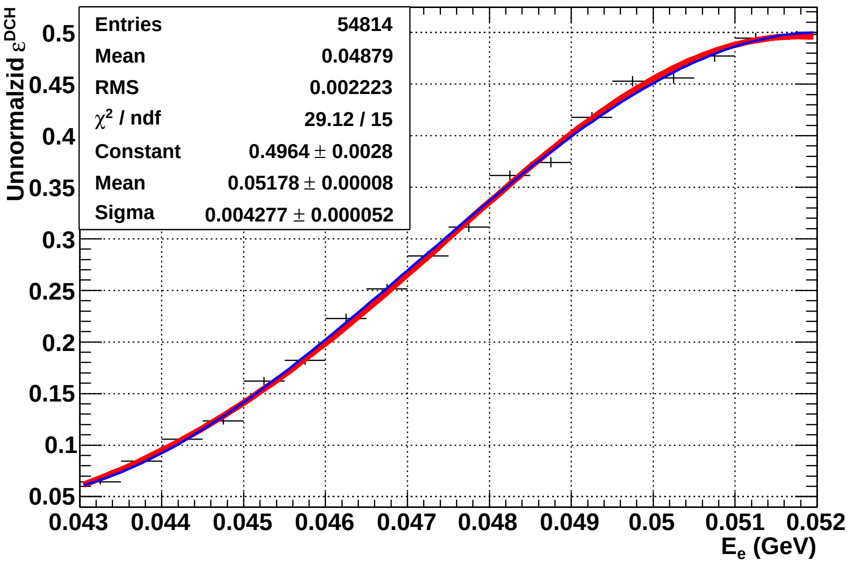


Figure 10.13: Unnormalized relative DCH efficiency as a function of positron energy, $\varepsilon^{\mathrm{DCH}}(E_e)$, obtained from Monte Carlo comparison method. It is fitted to a Gaussian function (red). Note that the vertical axis is arbitrary. The mean and σ of the fitted Gaussian function contain the full parameterization of $\varepsilon^{\mathrm{DCH}}(E_e)$. A third order polynomial fit (blue) is used to evaluate the systematic uncertainty.

compensate this resolution effect[5].

Therefore, we can construct $\varepsilon^{\mathrm{DCH}}(E_e)$ as

$$\varepsilon^{\mathrm{DCH}}(E_e) = \exp(-(E_e - E_{max})^2/2\sigma_E^2) \tag{10.14}$$

where E_{max}=0.0518 GeV and σ_E=0.00428 GeV. In this case, we have set the maximum $\varepsilon^{\mathrm{DCH}}(E_e)$ to 1; and the rest is normalized to it.

Errors in $\varepsilon^{\mathrm{DCH}}(E_e)$ do not directly enter the RD branching fraction calculation, but enters it in the ratio of $\left\langle \varepsilon_M^{\mathrm{DCH}} \right\rangle / \left\langle \varepsilon_{e\nu\bar{\nu}\gamma}^{\mathrm{DCH}} \right\rangle$. We discuss the uncertainty in Section 10.6.1.

---

[5] Therefore, $\varepsilon^{\mathrm{DCH}}(E_e)$ is in fact a function of the input positron energy. It has to be defined this way because the Michel spectrum to be used in factor $\left\langle \varepsilon_M^{\mathrm{DCH}} \right\rangle / \left\langle \varepsilon_{e\nu\bar{\nu}\gamma}^{\mathrm{DCH}} \right\rangle$ is taken from the theory.

Estimation of the Absolute DCH Efficiency

Though the absolute DCH efficiency is not essential to the radiative decay branching fraction calculation, it is still interesting to estimate its value. It can be calculated from Equation 10.5. We know that the beam intensity for radiative decay data are roughly $R_{beam} = 1.2 \times 10^6 \mu/s$. We also know that the counting live time of side band Michel tracks is $\tau_M^L = 0.251$ s (Section 10.4.3). We assume that all muons stop on the target, i.e. $\varepsilon(stop) = 1$. The spectrometer geometrical coverage is about 9% of the solid angle, i.e. $G_M^e = 9\%$; the number of detected positrons above 50 MeV is $N_M^{e,D}(E_e > 50 \text{ MeV}) = 1395$; the Michel branching fraction is $B_M^e = 1$; the Michel counting fraction above 50 MeV is $\varepsilon_M^{frac}(E_e > 50 \text{ MeV}) = 10.08\%$; the abosulte DCH efficiency $\varepsilon^{\mathrm{DCH}}$ is unknown, but is nearly flat in average above 50 MeV (see the relative $\varepsilon^{\mathrm{DCH}}(E_e)$ plot, Figure 10.13), which is the reason why we select positrons with energy above 50 MeV. Putting all the know factors into Equation 10.5 and solving for $\varepsilon^{\mathrm{DCH}}$, we estimate that the average absolute DCH efficiency at positron energy above 50 MeV is

$$\begin{aligned}
&\left\langle \varepsilon^{\mathrm{DCH}}(E_e > 50 \text{ MeV}) \right\rangle \\
&\quad = N_M^{e,D}(E_e > 50 \text{ MeV}) \cdot \frac{1}{R^\mu} \cdot \frac{1}{\tau_M^L} \cdot \frac{1}{\varepsilon(stop)} \cdot \frac{1}{B_M^e} \cdot \frac{1}{G_M^e} \cdot \frac{1}{\varepsilon_M^{frac}(E_e > 50 \text{ MeV})} \\
&\quad = 51.0\%
\end{aligned} \tag{10.15}$$

As argued previously, even though the relative DCH efficiency does not depend on beam intensity, the absolute DCH efficiency, nonetheless, does. At the beam rate of

$R_{beam} = 3 \times 10^7 \mu/s$, the absolute DCH efficiency is estimated using the same technique to be about 37%.

10.4.5 TIC Acceptance

The TIC acceptance factor is in fact a combination of both TIC geometrical acceptance and reconstruction efficiency of a well-measured TIC hit, which are, again, inseparable in the view of analysis. The MEG spectrometer was designed to optimize the acceptance of $\mu \rightarrow e\gamma$ decay positrons, which has a monochromatic energy at 52.8 MeV. Therefore, the acceptance is biased for the higher energy positrons, which reach TIC with higher probabilities. So the TIC acceptance is positron energy dependent. It is well defined as a conditional probability of finding a correlated TIC hit given a well reconstructed DCH positron track, $\varepsilon^{\mathrm{TIC}}$. This definition leads to a method to derive the TIC acceptance from data directly. The DC-TIC connection matching cuts require both spatial and timing match between the DCH track end point projected to the TIC and the associated TIC hit. We take the ratio of E_e distribution after applying the DC-TIC connection matching cuts to that before, we get the TIC acceptance as a function of positron energy, $\varepsilon^{\mathrm{TIC}}(E_e)$.

During the MEG data taking, DCH self trigger (trigger 18) data were recorded; it requires that the DCH have sufficient number of consecutive chambers hit without any requirement on the TIC. Thus, these events can be used for the purpose of TIC acceptance study. Higher beam intensity will cause higher accidental matching between the DCH and TIC. It happens when a positron does not in fact reach the TIC (e.g. due to scatterings on hard materials), but there is an accidental TIC hit coincident in both space and time with the projected DCH track end point satisfying the DCH-TIC matching selection criteria. It is clear that this accidental matching probability, $\varepsilon_{acc}^{\mathrm{TIC}}$, increases as the beam intensity increases. We can write $\varepsilon_{acc}^{\mathrm{TIC}}(R_{beam})$ as a function of beam rate,

$$\varepsilon_{acc}^{\mathrm{TIC}}(R_{beam}) = R_{beam} \cdot G_M^e \cdot \left\langle \varepsilon^{\mathrm{DCH}}(R_{beam}) \right\rangle \cdot \frac{2\delta R}{D_{bar}} \cdot \frac{2\delta Z}{Z_{bar\ range}} \cdot 2\delta T_0^e \tag{10.16}$$

where the beam intensity $R_{beam} = 3 \times 10^7 \mu/s$; positron geometrical acceptance of the spectrometer $G_M^e = 9\%$; $\left\langle \varepsilon^{\mathrm{DCH}}(R_{beam} = 3 \times 10^7 \mu/s) \right\rangle$ is the average absolute DCH efficieny, which has been estimated about 37% for normal beam intensity (see Section 10.4.4); δR, δZ, and δT_0^e are limits of the spatial and timing cuts in the DCH-TIC matching, whose numbers are 2 cm, 6 cm, and 12 ns, seperately; D_{bar}, the depth of the TIC bar and $Z_{bar\ range}$, the Z range of the TIC bar are the spatial span, within which the DCH track projected end point and the TIC hit take place, whose numbers are from the detector geometry, $D_{bar} = 4\text{cm}$ and $Z_{bar\ range} = 80\text{cm}$. The reason why we put a factor of 2 in front of δR, δZ, and δT_0^e is because we cut on the absolute values of the spatial and time difference between DCH and TIC; thus, we should count both directions. Combining these numbers, we estimate that

$$\begin{aligned} &\varepsilon_{acc}^{\mathrm{TIC}}(R_{beam} = 3 \times 10^7 \mu/s) \\ &\quad = 3 \times 10^7 s^{-1} \times 9\% \times 37\% \times \frac{2 \times 2cm}{4cm} \times \frac{2 \times 6cm}{80cm} \times (2 \times 12 \times 10^{-9} s) \\ &\quad = 0.36\% \end{aligned} \tag{10.17}$$

At this level, the accidental matching probability is negligible. Therefore, we can use trigger 18 events to calculate $\varepsilon^{\mathrm{TIC}}(E_e)$. The result is shown in Figure 10.14, fitted with a fourth order polynomial function (red).

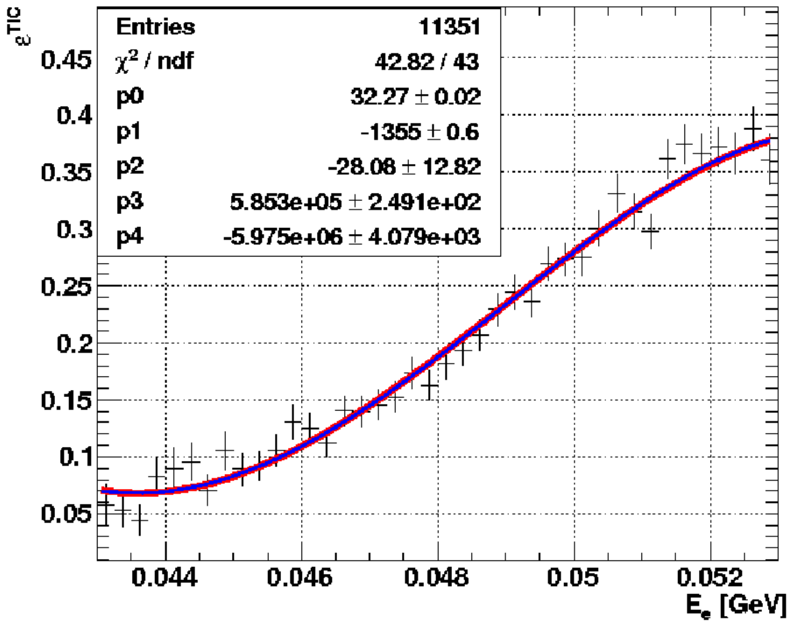


Figure 10.14: TIC acceptance as a function of positron energy, $\varepsilon^{\mathrm{TIC}}(E_e)$

We construct $\varepsilon^{\mathrm{TIC}}(E_e)$ function from a 4th order polynomial whose coefficients are the fitting parameters shown in Figure 10.14.

The average statistical uncertainty of $\varepsilon^{\mathrm{TIC}}$ is estimated to be 0.17 % using binomial error[6].

We also fitted $\varepsilon^{\mathrm{TIC}}(E_e)$ with a fifth order polynomial in Figure 10.14 (blue, barely seen) to estimate the systematic error, which is evaluated by a weighted average with the correct spectrum over the full kinematic range[7]. In average, it is evaluated to be 0.54 %.

---

[6] The calculation of $\varepsilon^{\mathrm{TIC}}(E_e)$ uses the same data sample with a condition that requires a DCH self triggered track has an associated TIC hit. This requirement can be considered as a binomial process with a probability of $\varepsilon^{\mathrm{TIC}}(E_e)$. In this case, 11351 positron tracks (number of entries in Figure 10.14) reach the TIC out of a total 54814 candidates (number of entries in Figure 10.13). Binomial errors are used as estimates for the average statistical uncertainties of all other efficiencies that are calculated in the same principle.

[7] The same technique is applied to estimate all other average systematic uncertainties of the efficiency factors

10.4.6 TIC Trigger Efficiency

The TIC trigger efficiency, $\varepsilon^{\text{TIC Trig}}$, is defined as the conditional probability of causing a TIC trigger given a well reconstructed DCH track as well as a correlated timing counter hit. As argued in Section 10.3, it is independent of the positron energy and is expected to be close to 1. Small possible inefficiencies may come from trigger algorithm failures, trigger-DAQ communication failures, and cable, board and other hardware inefficiencies, etc..

$\varepsilon^{\text{TIC Trig}}$ is used on the radiative decay part in the branching fraction calculation (Equation 10.8). The radiative decay trigger consists of three parts: a TIC hit with energy deposit above some threshold, a XEC hit with energy deposit above some threshold, and the time difference between the two is within 20 ns (trigger 4) or 40 ns (trigger 5). Threshold on the TIC in the radiative decay trigger (as well as the MEG trigger) is set the same as the one in trigger 22, the TIC alone trigger. So it is valid to use trigger 22 to evaluate the efficiency of the TIC trigger part of the radiative decay trigger. It is also the reason why we have referred the TIC trigger part of the radiative decay trigger (as well as the MEG trigger) as the TIC trigger.

We measure $\varepsilon^{\text{TIC Trig}}$ from the data directly. The trigger mechanism is designed to work in the way that besides the trigger type an event is triggered on the system also registers all other trigger types that the event satisfies. Figure 10.15 shows the trigger registration map of a trigger 18 event.

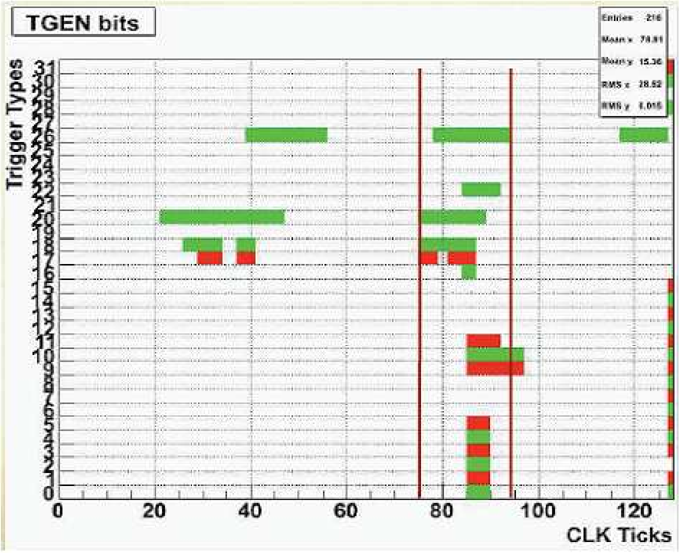


Figure 10.15: An example of the trigger mechanism. In a trigger 18 event, all triggers that satisfy their own firing requirements are marked in time, labeled in 10 ns clock ticks. Red and green marks alternate line by line for a clear viewing. The TIC trigger efficiency is the probability of registering a TIC trigger (22) in a well reconstructed DCH self trigger (18) event within a 200 ns window (marked between the two red bars) with respect to the latter's time.

Therefore we measure $\varepsilon^{\text{TIC Trig}}$ by selecting trigger 18 (DCH self trigger) events with well reconstructed track and associated TIC hit to the track, and then looking whether trigger 22 (TIC trigger) is also registered. Higher beam rate does introduce more positrons hit on the TIC in the same time interval. This could possibly bias the $\varepsilon^{\text{TIC Trig}}$ calculation, in the case that the DCH track correlated TIC hit fails to cause a trigger, but an accidental TIC hit causes one. To avoid this, we selected clean events with only one TIC hit. Positrons pass the DCH before they reach the TIC. The time of flight is in the order of a few nanoseconds. But the electron drift time in the DCH cells is in the order of 100 ns. So the DCH trigger comes later in time. We can find the global time difference between the DCH self determined T_0 and the consequential TIC time for trigger 18 events. Offsetting this time difference, we look for trigger 22 registrations in a window of 200 ns to compensate the drift time (also see Figure 10.15). Result of $\varepsilon^{\text{TIC Trig}}$ vs. E_e is shown in Figure 10.16. We have fitted

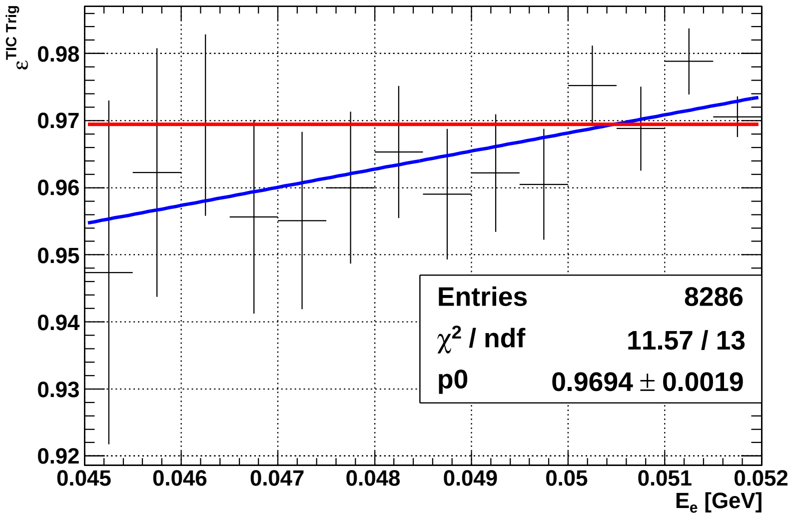


Figure 10.16: The TIC trigger efficiency

the histogram with a constant (red). We get that $\varepsilon^{\text{TIC Trig}} = 96.9\%$.

The statistical error of $\varepsilon^{\text{TIC Trig}}$ is estimated to be 0.19 % using binomial error.

We also fitted $\varepsilon^{\text{TIC Trig}}$ with a linear function to evaluate the systematic uncertainty, which is 0.67%, taking the difference between the two values as the systematic uncertainty.

The scale of the positron energy range applied in the branching fraction calculation is tens of times those of the positron resolution. Therefore, systematic uncertainty due to the positron detector resolution effects only have negligible contribution. This is also true in case of the photon analysis.

10.5 Photon Analysis

Unlike the positron analysis, which is coupled with the Michel normalization decays, photon analysis is intrinsic to the radiative decay. Therefore, we have to get absolute values for all photon related efficiency factors, the conditional photon geometrical acceptance ($G^{\gamma(e)}_{e\nu\bar{\nu}\gamma}$), the XEC efficiency ($\varepsilon^{\text{XEC}}$) and the XEC trigger efficiency ($\varepsilon^{\text{XEC Trig}}$). We will follow the same approach as we have done in the positron analysis to calculate a chain of well defined efficiencies in logical coherence and consistency.

10.5.1 Conditional Photon Geometrical Acceptance

The first photon related factor appearing in the branching fraction calculation (Equation 10.8) is the conditional photon geometrical acceptance, $G^{\gamma(e)}_{e\nu\bar{\nu}\gamma}$, defined as the conditional probability for the photon to enter the XEC volume given the positron from the same radiative decay process entering the DCH. It depends on the apparatus configuration and the decay kinematics. The photon and positron are bound to be back to back in the MEG decay, so the experiment apparatus was designed to have an optimal acceptance for MEG events: if the muon decays near the center of the target and if the positron enters the DCH, the photon from the MEG decay enters the XEC detector as well, i.e. $G^{\gamma(e)}_{MEG} \approx 1$. The reason that $G^{\gamma(e)}_{MEG}$ is not exactly 1 is because of the spread of the muon decay vertexes on the target.

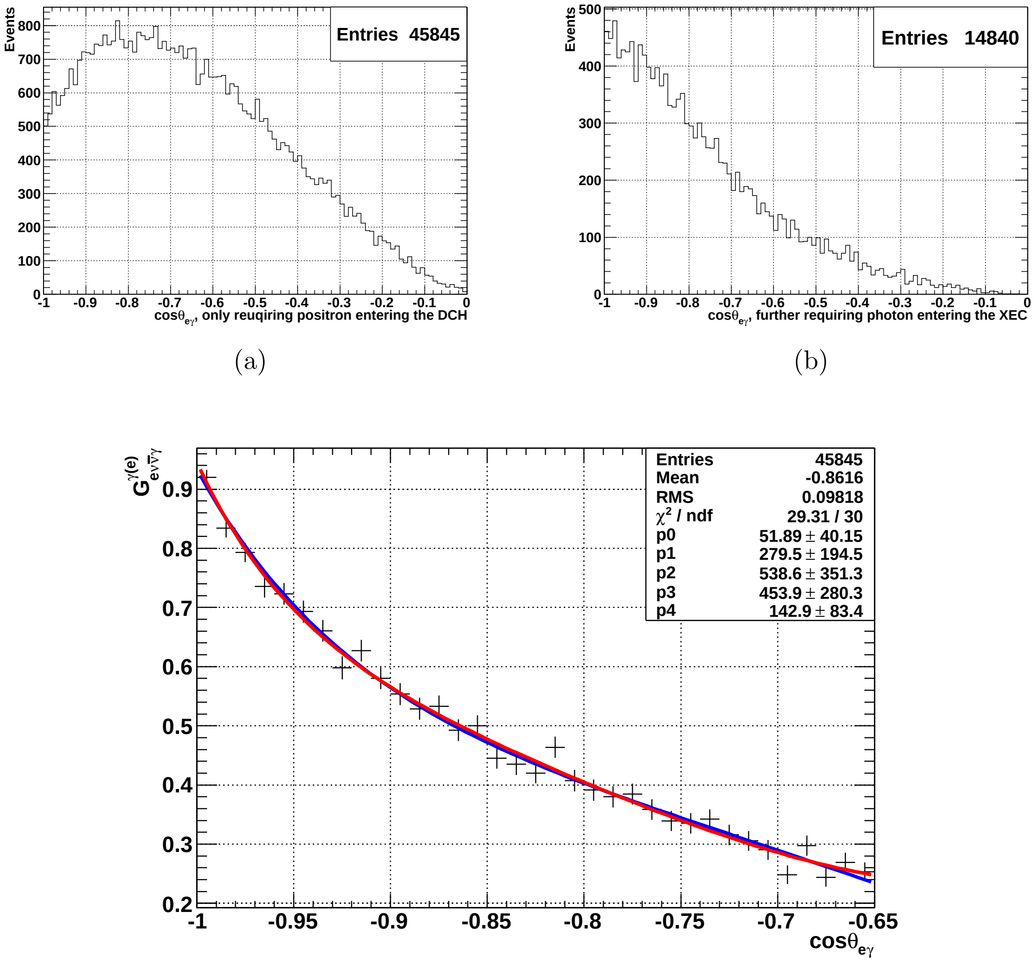
However, the radiative decay photon and positron, depending on their energies, can have a relatively wide range of opening angle between the photon and positron, $\theta_{e\gamma}$. For a given $\theta_{e\gamma} \neq 180°$, if the positron goes into the DCH, the photon does not necessarily enter the XEC volume. Thus, we have to calculate $G^{\gamma(e)}_{e\nu\bar{\nu}\gamma}$ to compensate this. $G^{\gamma(e)}_{e\nu\bar{\nu}\gamma}$ is a purely geometrical factor. Since the radiative decay differential branching fraction is proportional to $d\cos\theta_{e\gamma}$ (see Equation 10.29), instead of $d\theta_{e\gamma}$, in practice it is more intuitive to built $G^{\gamma(e)}_{e\nu\bar{\nu}\gamma}$ as a function of $\cos\theta_{e\gamma}$, i.e. $G^{\gamma(e)}_{e\nu\bar{\nu}\gamma}(\cos\theta_{e\gamma})$.

We generate a large sample of radiative decay Monte Carlo events in relatively broad energy ranges for both photon and positron to cover all possible $\cos\theta_{e\gamma}$ in the experimental limit. We generated both positron and photon into the full solid angle under the constraint of radiative decay kinematics. We first get the $\cos\theta_{e\gamma}$ distribution (Figure 10.17(a)) by only requiring the positron going into the DCH volume, which is only defined by ϕ_e and $\cos\theta_e$ as following, $|\phi_e| < 57.3^\circ$ and $|\cos\theta_e| < 0.35$; and further get the $\cos\theta_{e\gamma}$ distribution (Figure 10.17(b)) by requiring the photon entering the XEC volume (defined in Equation 9.2 and shown in Figure 9.8) in addition to the previous condition. Dividing the second $\cos\theta_{e\gamma}$ distribution by the first one, we have obtained the $G^{\gamma(e)}_{e\nu\bar{\nu}\gamma}(\cos\theta_{e\gamma})$ distribution, whose relevant range is shown in Figure 10.17(c). We have fitted the distribution with an exponential of a forth order polynomial (red) to describe the acceptance. Even though we use radiative decay simulation to calculate $G^{\gamma(e)}_{e\nu\bar{\nu}\gamma}(\cos\theta_{e\gamma})$, this technique does not rely on the theoretical radiative decay kinematics.

We can fit the $G^{\gamma(e)}_{e\nu\bar{\nu}\gamma}(\cos\theta_{e\gamma})$ with an exponential of a 4th order polynomial,

$$\begin{aligned}
&G^{\gamma(e)}_{e\nu\bar{\nu}\gamma}(\cos\theta_{e\gamma}) = \\
&\quad \exp(p0 + p1\cdot\cos\theta_{e\gamma} + p2\cdot\cos^2\theta_{e\gamma} + p3\cdot\cos^3\theta_{e\gamma} + p4\cdot\cos^4\theta_{e\gamma}) \\
&\quad (\text{for} \quad -1 \le \cos\theta_{e\gamma} \le -0.65)
\end{aligned}$$

where the coefficients are the fitting parameters shown in Figure 10.17(c).



(a)

(b)

(c)

Figure 10.17: Calculation of the conditional photon geometrical acceptance, $G^{\gamma(e)}_{e\nu\bar{\nu}\gamma}(\cos\theta_{e\gamma})$. (a) $\cos\theta_{e\gamma}$ distribution by only requiring the positron entering the spectrometer volume; (b) $\cos\theta_{e\gamma}$ distribution after further requiring the photon going into the XEC volume; (c) $G^{\gamma(e)}_{e\nu\bar{\nu}\gamma}(\cos\theta_{e\gamma})$ distribution fitted with an exponential of a forth order polynomial (red) and also with an exponential of a third order polynomial (blue).

The statistical error of $G^{\gamma(e)}_{e\nu\bar{\nu}\gamma}$ is estimated to be 0.22 % using binomial error[8]

We also fitted $G^{\gamma(e)}_{e\nu\bar{\nu}\gamma}(\cos\theta_{e\gamma})$ with an exponential of a third order polynomial (blue) to estimate the systematic uncertainty. It is evaluated to be 0.46 % taking the difference between the two fitting fuctions as the systematic uncertainty.

10.5.2 XEC Efficiency

After obtaining $G^{\gamma(e)}_{e\nu\bar{\nu}\gamma}(\cos\theta_{e\gamma})$, i.e. knowing a photon entering the XEC volume, the next thing to know is the XEC efficiency, $\varepsilon^{\mathrm{XEC}}$. A photon going in the direction of the XEC volume may convert or Compton scatter on the XEC inner face wall, which constitutes a certain loss on the photon detection rate. After it enters the XEC detector, the detector itself has an intrinsic efficiency to reconstruct the photon. Therefore, the XEC efficiency is a combination of XEC detection rate and its reconstruction efficiency, which are inseparable in the view of analysis.

The XEC efficiency is very likely photon energy dependent, i.e. $\varepsilon^{\mathrm{XEC}}(E_\gamma)$. Unfortunately, the experiment does not have a source of γ rays that has a known spectrum over the radiative decay photon energy range; we have to, therefore, evaluate $\varepsilon^{\mathrm{XEC}}(E_\gamma)$ with the help of Monte Carlo simulation.

We generated several samples of monochromatic γ rays entering the XEC detector with a uniform distribution over the geometrical range. This is meant to evaluate an efficiency of the average XEC performance. We define the probability density function

---

[8] 14840 photons enter the XEC volume out of 45845 events. These number are shown in Figures 10.17(a) and (b)

of the ratio of detected energy to the input energy, i.e. $R(E_{\gamma,detected}/E_{\gamma,input})$, as the response function of the XEC detector. Distributions of $R(E_{\gamma,detected}/E_{\gamma,input})$ with different input energies are shown in Figure 10.18. In fact, we find that the response function has similar distributions for all samples of the different monochromatic input photon energies in the relevant energy range, i.e. 0.7< $R < 1.2$. To simplify the calculation, an average response function is used.

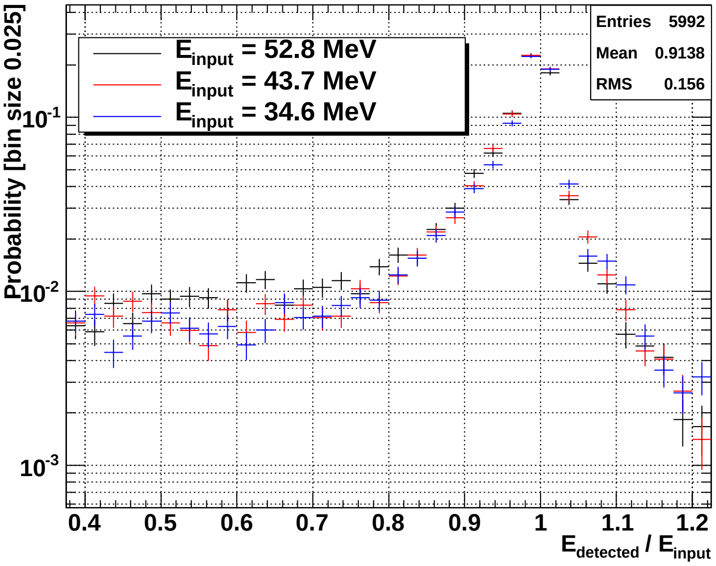


Figure 10.18: The XEC response functions with different input energies

The observed photon spectrum of the RD events in data(Figure 10.6(c)) and the XEC trigger efficiency (which is calculated from data directly [See Figure 10.20 in the next section]), assuming the XEC efficiency does not vary much in the relevant region, indicate that the original photon input distribution should decrease dramatically as E_γ increases. We use the theoretical radiative decay as an input spectrum (Figure 10.19(a), black) and convolve it with the response function, following Equation 10.18, to obtain the ‘observed’ spectrum (Figure 10.19(a), red).

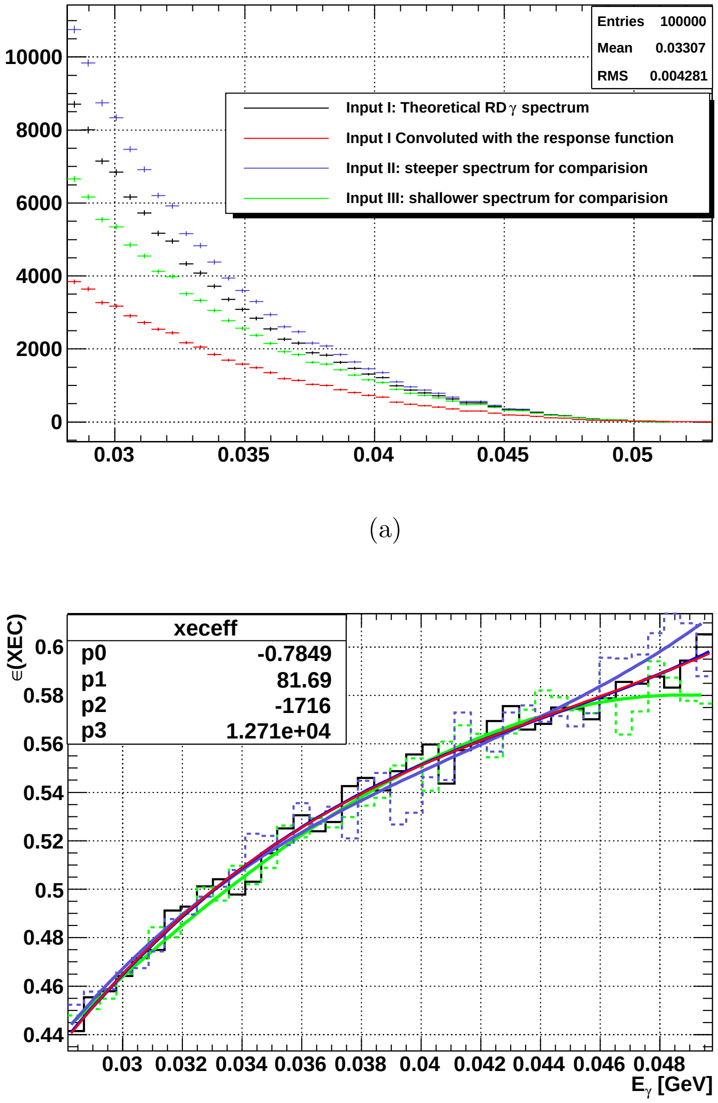
$$O(E) = \int I(E')R\left(\frac{E}{E'}\right)dE' \tag{10.18}$$

where $I(E)$ and $O(E)$ are the input and ‘observed’ spectra respectively. The XEC efficiency is calculated by taking the ratio of the ‘observed’ spectrum to the input one. $\varepsilon^{\mathrm{XEC}}(E_\gamma)$, obtained by means of this technique from the theoretical RD photon spectrum, is shown in Figure 10.19(b).

The reliance of the theoretical input into the $\varepsilon^{\mathrm{XEC}}(E_\gamma)$ is not totally satisfactory. Since the behaviors of different $\varepsilon^{\mathrm{XEC}}(E_\gamma)$ calculated from different input spectra may differ at the high energy end due to resolution effects. Two test spectra (Figure 10.19(a)) are used as inputs to calculate $\varepsilon^{\mathrm{XEC}}(E_\gamma)$ for the purpose of verification. These tests are shown in Figure 10.19 and coded by color. The results are consistent with $\varepsilon^{\mathrm{XEC}}(E_\gamma)$ obtained from the theoretical RD spectrum in the most relevant region. The divergence occurs when $E_\gamma > 46$ Mev, where there are only very few observed events (see Figure 10.6(c)).

The experiment does have one nearly monochromatic photon source at 55 MeV from π^0 decays by tagging back-to-back photons around 83 MeV in the NaI detector. The XEC efficiency is measured to be 63±4% at 55 MeV from data directly[64]. This is consistent with our estimation from the MC method.

The distribution for $\varepsilon^{\mathrm{XEC}}(E_\gamma)$ is fit to a third order polynomial whose coefficients are the fitting parameters shown in Figure 10.19(b).



(a)

(b)

Figure 10.19: Calculation of the XEC efficiency. (a) theoretical RD photon spectrum (black) and its convolution with the XEC response function (red). Two other test input spectra for verification purpose are also presented. (b) $\varepsilon^{\mathrm{XEC}}(E_\gamma)$ calculated from the theoretical RD photon spectrum and from other two test spectra, all fitted with third order polynomials.

The average statistical uncertainty of ε^{XEC} is calculated using the binomial error to be 0.05%, which has a negligible contribution.

The systematic uncertainty is estimated by evaluating different ε^{XEC} obtained from the different spectra, averaged over the full kinematic range. It is evaluated to be 0.79 %.

10.5.3 XEC Trigger Efficiency

The last factor in the line of photon analysis is the XEC trigger efficiency, i.e. $\varepsilon^{\text{XEC Trig}}$, which is defined as the conditional probability of setting a XEC trigger given a well reconstructed γ in the XEC detector. The XEC trigger fires if the sum of PMT charges is above some threshold, therefore, $\varepsilon^{\text{XEC Trig}}$ is a photon energy dependent factor, i.e. $\varepsilon^{\text{XEC Trig}}(E_\gamma)$.

$\varepsilon^{\text{XEC Trig}}(E_\gamma)$ is evaluated with trigger 18 (DCH trigger), 22 (Michel trigger) and 31 (Random Trigger) events, since they are all random triggers from the perspective of XEC detector. We require a well-reconstructed photon in the XEC and then ask for the probability that trigger 10 (XEC trigger) is also registered. $\varepsilon^{\text{XEC Trig}}(E_\gamma)$ is shown in Figure 10.20.

The $\varepsilon^{\text{XEC Trig}}(E_\gamma)$ distribution has the shape of an error function. Thus we fit it with an error function taking a first order polynomial of E_γ as its argument with floating boundaries (Equation 10.19). The result of the fit is shown as the red curve

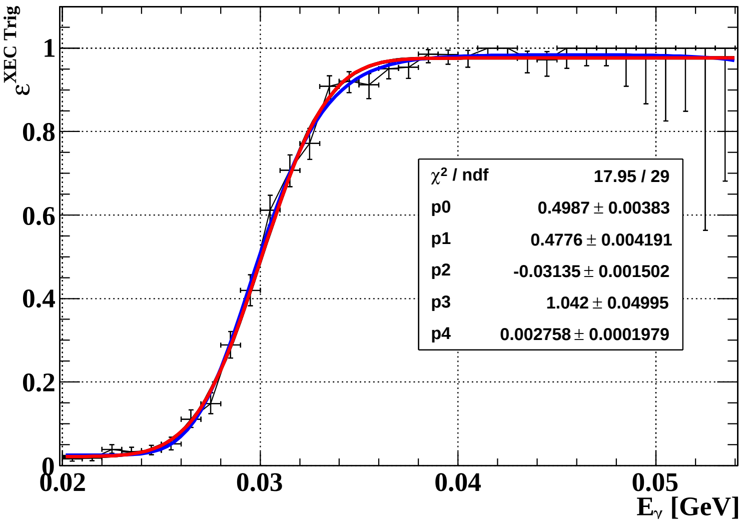


Figure 10.20: XEC trigger efficiency

in Figure 10.20. We can construct $\varepsilon^{\mathrm{XEC\ Trig}}(E_\gamma)$ from fitting parameters,

$$\varepsilon^{\mathrm{XEC\ Trig}}(E_\gamma) = p0 + p1 \cdot \mathrm{Erf}[(p2 + p3 \cdot E_\gamma)/(\sqrt{2} \cdot p4)] \tag{10.19}$$

where $p0 = 0.499$, $p1 = 0.478$, $p2 = -3.135 \times 10^{-2}$, $p3 = 1.04$ and $p4 = 2.76 \times 10^{-3}$.

The average statistical error of $\varepsilon^{\mathrm{XEC\ Trig}}$ is estimated to be 0.75 % using binomial error[9]

In order to evaluate the systematic uncertainty, we fit the data with an error function taking a second order polynomial of E_γ as its argument, i.e.

$$\varepsilon^{\mathrm{XEC\ Trig}}(E_\gamma) = p0 + p1 \cdot \mathrm{Erf}[(p2 + p3 \cdot E_\gamma + p4 \cdot (E_\gamma)^2)/(\sqrt{2} \cdot p5)].$$

The result of the second fit is shown as the blue curve in the same plot. We found

---

[9] Total number of well reconstructed photons in all events of triggers 18, 22, 31 is 4083, out of which 1477 events have trigger 10 registered. These numbers are not shown in any histogram.

that the average uncertainty over the full kinematic range is 0.83 %.

10.6 Branching Fraction Calculation

Having calculated all the factors used in Equation 10.8, we can calculate the radiative decay branching faction for certain energy ranges of the positron and photon. We do the calculation on an event-to-event basis[10].

10.6.1 Factor $\left\langle \varepsilon_M^{\mathbf{DCH}} \right\rangle / \left\langle \varepsilon_{e\nu\bar{\nu}\gamma}^{\mathbf{DCH}} \right\rangle$

One merit of our normalization scheme is to avoid calculating the absolute DCH efficiency. DCH efficiency appears in pairs both in the numerator and denominator in factor $\left\langle \varepsilon_M^{\mathrm{DCH}} \right\rangle / \left\langle \varepsilon_{e\nu\bar{\nu}\gamma}^{\mathrm{DCH}} \right\rangle$. Therefore a relative DCH efficiency is sufficient for this purpose. We have calculated the relative $\varepsilon^{\mathrm{DCH}}$ in Section 10.4.4 and parameterized it in Equation 10.14. We normalized the relative $\varepsilon^{\mathrm{DCH}}$ to its maximum, which occurs at the high energy end.

For a known input E_e spectrum, $P(E_e)$, the average DCH efficiency, $\left\langle \varepsilon^{\mathrm{DCH}} \right\rangle$, can be written as,

$$\left\langle \varepsilon^{\mathrm{DCH}} \right\rangle = \frac{\int P(E_e) \cdot \varepsilon^{\mathrm{DCH}}(E_e) dE_e}{\int P(E_e) dE_e}. \tag{10.20}$$

The high energy end of Michel E_e spectrum is shown in Figure 10.11(b). Together with Equation 10.14, we calculated that $\left\langle \varepsilon_M^{\mathrm{DCH}} \right\rangle$ for $E_e > 48$ MeV and $E_e > 50$ MeV

---

[10] This means to do weighted averages as discrete sum event by event instead of integration over the spectra.

are,

$$\left\langle \varepsilon_M^{\mathrm{DCH}}(E_e > 48\ \mathrm{MeV}) \right\rangle = 90.31\% \tag{10.21}$$

and

$$\left\langle \varepsilon_M^{\mathrm{DCH}}(E_e > 50\ \mathrm{MeV}) \right\rangle = 98.07\% \tag{10.22}$$

Figure 10.6(b) shows the detected radiative decay signal positron energy spectrum. It is the original radiative decay positron energy distribution with DCH efficiency $\varepsilon^{\mathrm{DCH}}$, TIC efficiency $\varepsilon^{\mathrm{TIC}}$, and TIC trigger efficiency $\varepsilon^{\mathrm{TIC\ Trig}}$ factors built in it. We correct these factors[11] and we get the reconstructed input radiative decay positron energy spectrum, which is shown in Figure 10.21.

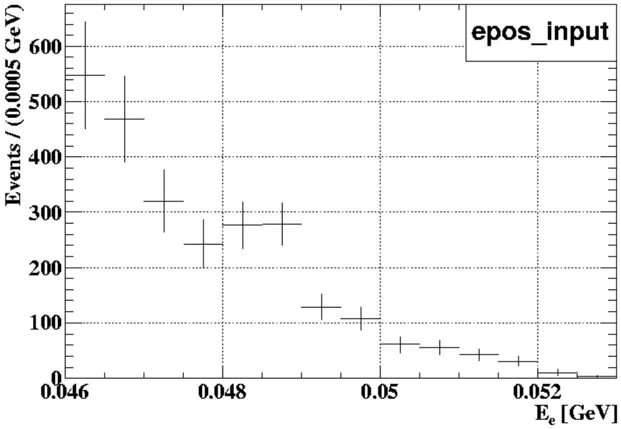


Figure 10.21: Reconstructed radiative decay positron spectrum, i.e. the input E_e spectrum

---

[11] We correct these factor by first dividing the observed positron spectrum by $\varepsilon^{\mathrm{TIC\ Trig}}$ and $\varepsilon^{\mathrm{TIC}}$, then de-convolving it with the resolution function, and finally dividing it by $\varepsilon^{\mathrm{DCH}}$. This is because that $\varepsilon^{\mathrm{TIC\ Trig}}$ and $\varepsilon^{\mathrm{TIC}}$ are functions of the observed energy, but $\varepsilon^{\mathrm{DCH}}$ is a function of the input energy. Calculation of factor Factor $\left\langle \varepsilon_M^{\mathrm{DCH}} \right\rangle / \left\langle \varepsilon_{e\nu\bar{\nu}\gamma}^{\mathrm{DCH}} \right\rangle$ needs a radiative decay positron spectrum of the input energy as well.

In the radiative decay branching fraction calculation, we are only interested in the E_e range where it has significant reconstruction efficiencies. It is because the efficiency factors are measured from data, low efficiency corresponds to small number of events in that region, thus bigger error. So we choose $E_e > 46$ MeV for this purpose. Thus using Figure 10.21, we calculate the average DCH efficiency for radiative decay positrons.

$$\left\langle \varepsilon^{\mathrm{DCH}}_{e\nu\bar{\nu}\gamma}(E_e > 46\ \mathrm{MeV}) \right\rangle = 62.12\% \tag{10.23}$$

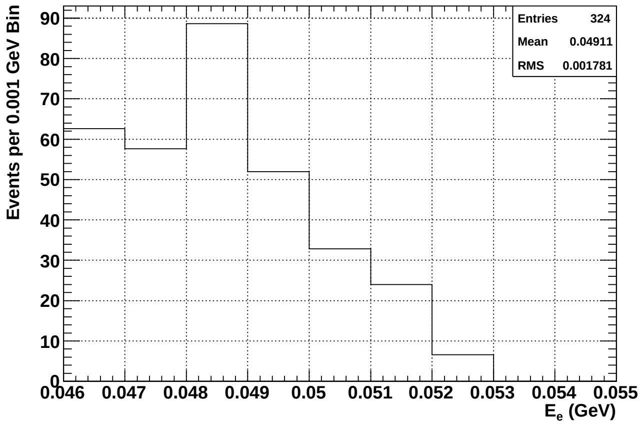
Therefore, we are able to calculate the ratio between the two relative DCH efficiencies,

$$\frac{\left\langle \varepsilon^{\mathrm{DCH}}_{M}(E_e > 48\ \mathrm{MeV}) \right\rangle}{\left\langle \varepsilon^{\mathrm{DCH}}_{e\nu\bar{\nu}\gamma}(E_e > 46\ \mathrm{MeV}) \right\rangle} = 1.45, \tag{10.24}$$

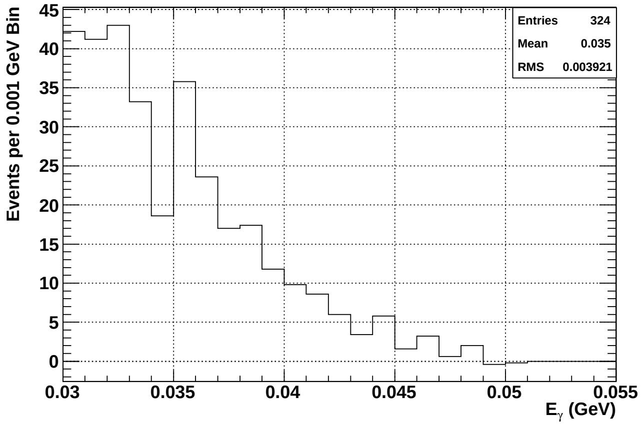
and

$$\frac{\left\langle \varepsilon^{\mathrm{DCH}}_{M}(E_e > 50\ \mathrm{MeV}) \right\rangle}{\left\langle \varepsilon^{\mathrm{DCH}}_{e\nu\bar{\nu}\gamma}(E_e > 46\ \mathrm{MeV}) \right\rangle} = 1.58, \tag{10.25}$$

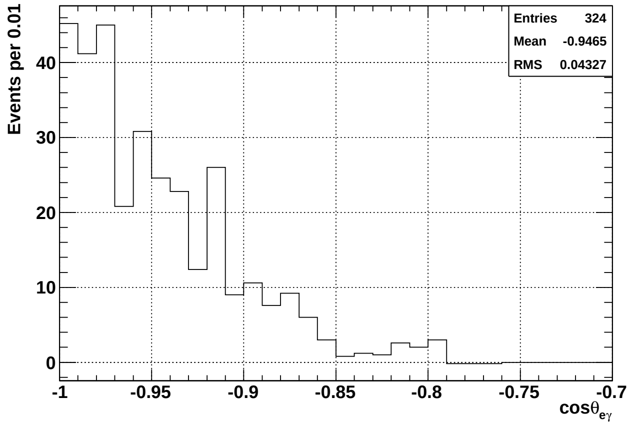
Estimations of both statistical and systematic uncertainties of $\left\langle \varepsilon^{\mathrm{DCH}}_{M} \right\rangle / \left\langle \varepsilon^{\mathrm{DCH}}_{e\nu\bar{\nu}\gamma} \right\rangle$ are done by taking the weighted averages integrated over the Michel spectrum and radiative decay spectrum. These results only have negligible contributions compared to the uncertainties of other factors.



(a) signal events E_e



(b) signal events E_γ



(c) signal events $\cos\theta_{e\gamma}$

Figure 10.22: Distributions of kinematic variables of signal events with condition $E_\gamma > 30$ MeV and $E_e > 46$ MeV

10.6.2 Radiative Decay Branching Fraction

We have now calculated all relevant factors in Equation 10.8. We have also estimated the systematic uncertainties of the efficiency factors. For the purpose of reducing systematic errors, we select radiative events only in the E_e and E_γ ranges where their efficiency factors are significant. We choose $E_\gamma > 30$ MeV and $E_e > 46$ MeV.

Physical Meaning	Symbol	Quantity
Number of detected radiative decay signals (background subtracted)	N_{sig}	324.2
Number of Michel events with $E_e > 50$ MeV in the normalization sample	$N_M^{e,D}(E_e > 50 \text{ MeV})$	1395
Fraction of $E_e > 50$ MeV Michel positron in the normalization sample	$\varepsilon_M^{frac}(E_e > 50 \text{ MeV})$	10.08%
Ratio of Michel live time to the radiative decay live time	$\tau_M^L / \tau_{e\nu\bar{\nu}\gamma}^L$	4.018×10^{-7}
Ratio of average relative DCH efficiencies	$\frac{\left\langle \varepsilon_M^{\text{DCH}}(E_e{>}50 \text{ MeV})\right\rangle}{\left\langle \varepsilon_{e\nu\bar{\nu}\gamma}^{\text{DCH}}(E_e{>}46 \text{ MeV})\right\rangle}$	1.58
Average TIC efficiency	$\left\langle \varepsilon^{\text{TIC}}(E_e > 46 \text{ MeV})\right\rangle^1$	18.2%
TIC trigger efficiency	$\varepsilon^{\text{TIC Trig}}$	96.9%
Average conditional photon geometrical acceptance	$\left\langle G_{e\nu\bar{\nu}\gamma}^{\gamma(e)}(\cos\theta_{e\gamma})\right\rangle^1$	72.3%
Average XEC efficiency	$\left\langle \varepsilon^{\text{XEC}}(E_\gamma > 30 \text{ MeV})\right\rangle^1$	50.5%
Average XEC trigger efficiency	$\left\langle \varepsilon^{\text{XEC Trig}}(E_\gamma > 30 \text{ MeV})\right\rangle^1$	81.3%
Radiative decay branching fraction in the range of $E_\gamma > 30$ MeV and $E_e > 46$ MeV	$B(\mu \to e\nu\bar{\nu}\gamma)\|_{E_\gamma > 30 \text{ MeV } E_e > 46 \text{ MeV}}$	2.84×10^{-7}

[1] The average efficiency factors are calculated in the same way as $\left\langle \varepsilon_{e\nu\bar{\nu}\gamma}^{\text{DCH}}(E_e > 46 \text{ MeV})\right\rangle$.

Table 10.1: Factors used in the radiative decay calculation

After background subtraction[12], the signal events kinematic distributions are shown

---

[12] As explained in Section 10.1.3, background subtraction is done using a background band that

in Figure 10.22. Note that these distributions are not merely the high energy/bigger opening angle end of distributions in Figure 10.6, but rather a subset of them, because all three kinematic variables are strongly correlated. The kinematic constraint also gives an upper limit on $\cos\theta_{e\gamma}$ at the chosen energy ranges, i.e. $E_\gamma > 30$ MeV and $E_e > 46$ MeV, which is calculated to be $\cos\theta_{e\gamma}$ <-0.78. The $\cos\theta_{e\gamma}$ distribution (Figure 10.22(c)) has falls smoothly to this upper limit, which confirms that, at these energy ranges, we indeed measure a branching fraction in the full $\cos\theta_{e\gamma}$ range. Using $E_e > 50$ MeV Michel events as the normalization sample and solving Equation 10.8, we have

$$B(\mu \to e\nu\bar{\nu}\gamma)|_{E_\gamma>30\text{ MeV }E_e>46\text{ MeV}} = 2.84 \times 10^{-7} \tag{10.26}$$

Table 10.1 lists values of the factors that go into the calculation.

10.6.3 Uncertainty Estimation

A summary of relative statistical and systematic uncertainties from the measurements of the efficiency factors is shown in Table 10.2. All efficiencies are uncorrelated, therefore, the total uncertainties of efficiencies, δ^{eff}_{stat} and δ^{eff}_{syst}, can be calculated from standard error propagation.

The statistical uncertainty in our calculation is dominated by the numerator, which has contributions from the total number of the observed events in the signal region

---

is 5 times as wide as the signal region. The number of events in the signal region is subtracted by the expected number of background events to get the detected number of signal events. Kinematic variables are handled separately. There are 456 events in the signal region and 659 in the background region. Thus, number of detected RD signal events is calculated as $456 - \frac{1}{5} \times 659 = 324.2$

relative uncertainties	statistical (%)	systematic (%)
$\delta(\varepsilon^{\mathrm{TIC}})$	0.17	0.54
$\delta(\varepsilon^{\mathrm{TIC\ Trig}})$	0.19	0.67
$\delta(G^{\gamma(e)}_{e\nu\bar{\nu}\gamma})$	0.22	0.46
$\delta(\varepsilon^{\mathrm{XEC}})$	-	0.79
$\delta(\varepsilon^{\mathrm{XEC\ Trig}})$	0.75	0.83
total	δ^{eff}_{stat}	δ^{eff}_{syst}
	0.8	1.5

Table 10.2: List of uncertainties from the measurements of the efficiencies

$N^{sig\ region} = 456$ and the number of background events in the background region $N^{bkgd\ region} = 659$. The background region time window is 5 times as wide as that of the signal region, therefore, the number of observed signal events (background subtracted) is $N^{sig} = N^{sig\ region} - \frac{1}{5}N^{bkgd\ region}$. Using Poisson error, the statistical uncertainty from the signal events is estimated to be

$$\delta^{sig}_{stat} = \frac{\sqrt{N^{sig\ region} + \frac{1}{25}N^{bkgd\ region}}}{N_{sig}} \sim 6.4\%. \tag{10.27}$$

The statistical uncertainty due to the normalization sample is found to be,

$$\delta^{norm}_{stat} = \frac{1}{\sqrt{N^{e,D}_{M}}} \sim 2.6\%. \tag{10.28}$$

A list of the uncertainties estimated for different aspects of the calculation is summarized in Table 10.3. The effective total uncertainty is calculated to be $\delta_{tot} \sim 6.9\%$.

After estimation of the uncertainties, we quote our experimental measurement of the radiative muon decay branching fraction at energy ranges of $E_\gamma > 30$ MeV and

	relative uncertainties (%)	
physical aspects	statistical	systematic
number of observed signals	δ_{stat}^{sig} 6.4	-
normalization	δ_{stat}^{norm} 2.6	δ_{syst}^{norm} 1.2
efficiencies	δ_{stat}^{eff} 0.8	δ_{syst}^{eff} 1.5
total	δ_{stat} 6.9	δ_{syst} 1.9

Table 10.3: Summary of uncertainties

$E_e > 46$ MeV with $\cos\theta_{e\gamma}$ in the fully allowed kinematic region is

$$B(\mu \rightarrow e\nu\bar{\nu}\gamma)|_{E_\gamma > 30 \text{ MeV } E_e > 46 \text{ MeV}} = (2.84 \pm 0.20(stat) \pm 0.05(syst)) \times 10^{-7}$$

Verification Checks of the Uncertainties

In Section 10.6.2, we use a normalization sample of Michel positron $E_e > 50 MeV$. The background subtraction applied to it is to choose the background window (ΔT_{bkgd}) 5 times as wide as the signal window (ΔT_{sig}). Using (i) a different normalization sample of Michel positron $E_e > 48$ MeV and (ii) a different background window size, $\Delta T_{bkgd}/\Delta T_{sig} = 8$ provides checks of the statistical and systematic uncertainties of certain aspects. Table 10.4 lists the values of these two calculations, which are compared with the previous one listed in the central column.

When the normalization sample and background window size are changed, there introduces only negligible changes in the branching fraction compared to the total statistical and systematic uncertainties.

	i		ii
window ratio	$\frac{\Delta T_{bkgd}}{\Delta T_{sig}} = 5$	$\frac{\Delta T_{bkgd}}{\Delta T_{sig}} = 5$	$\frac{\Delta T_{bkgd}}{\Delta T_{sig}} = 8$
normalization sample	$E_e > 48$ MeV	$E_e > 50$ MeV	$E_e > 50$ MeV
N_{sig}	-[1]	324.2	330.1
$N_M^{e,D}$	2210	1395	-
ε_M^{frac}	17.32%	10.08%	-
$\tau_M^L / \tau_{e\nu\bar{\nu}\gamma}^L$	-	4.018×10^{-7}	-
$\frac{\left\langle \varepsilon_M^{\mathrm{DCH}} \right\rangle}{\left\langle \varepsilon_{e\nu\bar{\nu}\gamma}^{\mathrm{DCH}} \right\rangle}$	1.45	1.58	1.57[2]
$\left\langle \varepsilon^{\mathrm{TIC}}(E_e > 46 \text{ MeV}) \right\rangle$	-	18.2%	18.2%
$\varepsilon^{\mathrm{TIC\ Trig}}$	-	96.9%	-
$\left\langle G_{e\nu\bar{\nu}\gamma}^{\gamma(e)}(\cos\theta_{e\gamma}) \right\rangle$	-	72.3%	72.2%
$\left\langle \varepsilon^{\mathrm{XEC}}(E_\gamma > 30 \text{ MeV}) \right\rangle$	-	50.5%	50.5%
$\left\langle \varepsilon^{\mathrm{XEC\ Trig}}(E_\gamma > 30 \text{ MeV}) \right\rangle$	-	81.3%	81.2%
$B(\mu \to e\nu\bar{\nu}\gamma)_{E_\gamma > 30 \text{ MeV}\ E_e > 46 \text{ MeV}}$	2.84×10^{-7}	2.84×10^{-7}	2.88×10^{-7}

---

[1] Value is unchanged from that in the central column.

[2] The change of background region slightly affects the observed radiative decay positron energy distribution and results the minor change of this quantity. $\left\langle \varepsilon_{e\nu\bar{\nu}\gamma}^{\mathrm{DCH}} \right\rangle$ is calculated to be 62.35%, comparing to the previous 62.12%. Therefore, $\left\langle \varepsilon_M^{\mathrm{DCH}}(E_e > 50 \text{ MeV}) \right\rangle / \left\langle \varepsilon_{e\nu\bar{\nu}\gamma}^{\mathrm{DCH}}(E_e > 46 \text{ MeV}) \right\rangle$ is 1.57 in this case. Other efficiencies are also slightly changed for the same reason.

Table 10.4: Verification checks of the uncertainties using (i) a different normalization sample and (ii) a different background window size

10.7 Comparison to the Theory

The differential branching fraction of the radiative decay has three kinematic variables, E_e, E_γ and $\theta_{e\gamma}$. In the V-A theory of weak interaction, it is given as Equation 2.2. The first term is independent of the muon polarization. There are also terms proportional to $\vec{P}_\mu \cdot \hat{p}_e$ and $\vec{P}_\mu \cdot \hat{p}_\gamma$. Hence, the prediction for the branching fraction may depend on the muon beam polarization.

In the MEG experiment, muons are totally polarized when they enter the experiment because muons are produced by pions decaying at rest. The stopping target is designed to totally depolarize muons; however, they may still have a small residual polarization $\vec{P'_\mu}$. However, in the detector acceptance region ($70° < \theta < 110°$), the integrand of the polarization terms containing factors $\hat{p_e} \cdot \vec{P'_\mu}$ and $\hat{p_\gamma} \cdot \vec{P'_\mu}$ are odd functions with respect to 90°. Thus the integrations of the polarization terms vanish. Therefore, considering only the unpolarized term, we calculate the theoretical branching fraction of radiative decay by integrating x from $46/52.8 = 0.8712$ to 1, and y from $30/52.8 = 0.5618$ to 1 and, $\cos\theta_{e\gamma}$ in the allowed kinematic region, we get

$$B(\mu \to e\nu\bar{\nu}\gamma) = \frac{\alpha}{8\pi} \int_{0.8712}^{1} dx \int_{0.5618}^{1} \frac{dy}{y} \int d\cos\theta_{e\gamma} \beta F^{(0)}(x, y, d), \tag{10.29}$$

where we have already integrated out extra freedoms, leaving the opening angle between photon and positron, $\theta_{e\gamma}$, as an integration variable. The integration yields,

$$B^{theory}(\mu \to e\nu\bar{\nu}\gamma)|_{E_\gamma > 30 \text{ MeV } E_e > 46 \text{ MeV}} = 2.99 \times 10^{-7}. \tag{10.30}$$

The experiment data agrees with the theory within one standard deviation of the uncertainty.

Another check of the agreement between theory and experiment is a comparision of the distributions in the three kinematic variables: E_e, E_γ, and $\cos\theta_{e\gamma}$. A likelihood fit method is used to compare the observed distributions of the kinematic variables with those of the theory.

Here, we briefly present our likelihood fitting procedure. For a kinematic variable x (E_e, E_γ, or $\cos\theta_{e\gamma}$), we first find the 'expected' signal distribution of x by integrating over other variables, convolving the theoretical distribution of x with the detected resolution function, and multiplying by the efficiency functions. We call this resulting distribution $S(x)$. We then directly measure the background distribution of x, $B(x)$ from background sidebands. Here, both $S(x)$ and $B(x)$ are functions of measured quantities. Thus, both signal and background probability density functions (PDFs) can be built by normalizing $S(x)$ and $B(x)$. These PDFs are referred as $\hat{S}(x)$ and $\hat{B}(x)$.

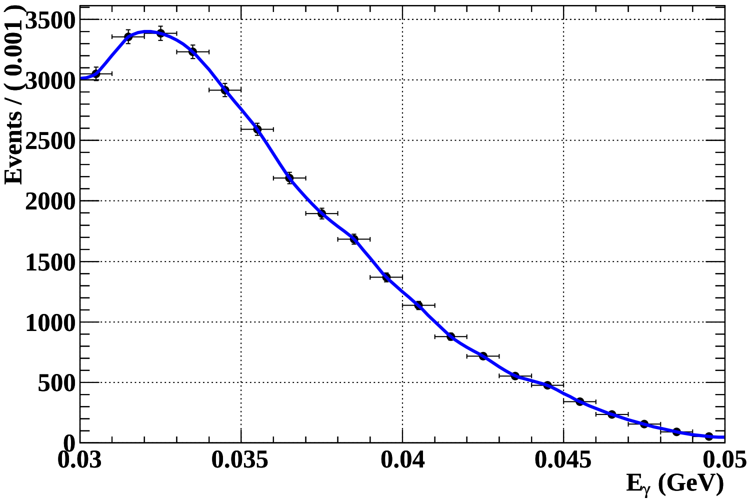
In the signal region, we fit the distribution of x with a function $f(x)$ which is written as a linear combination of the signal and background PDFs,

$$f(x) = N_{sig} \cdot \hat{S}(x) + N_{bkgd} \cdot \hat{B}(x)$$

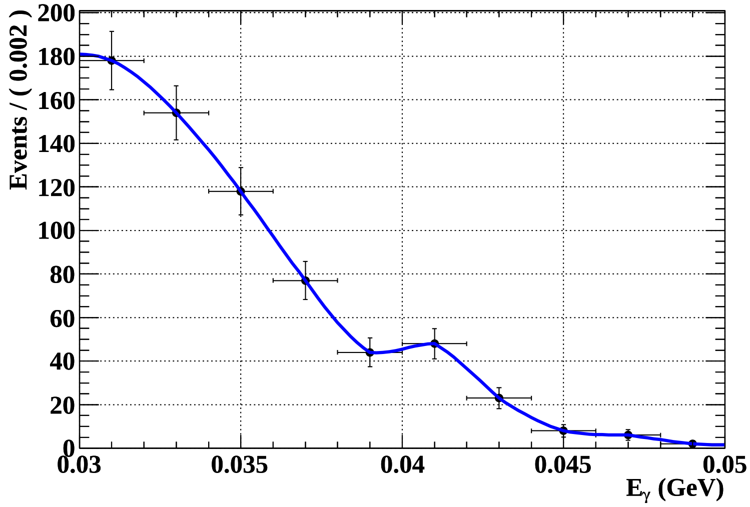
where N_{sig} and N_{bkgd} are the free fitting parameters and represent the expected number of signal and background events, respectively.

Photon Likelihood Fit

$S(E_\gamma)$ and $B(E_\gamma)$ distributions are shown in Figure 10.23.



(a) $S(E_\gamma)$, signal E_γ PDF



(b) $B(E_\gamma)$, background E_γ PDF

Figure 10.23: E_γ PDFs for the radiative decay signal and the background

The fitting result is shown in Figure 10.24. The fit gives,

$$\begin{aligned} N_{\gamma,sig} &= 315.0 \pm 46.2 \\ N_{\gamma,bkgd} &= 141.2 \pm 44.7 \end{aligned} \tag{10.31}$$

The detected number of signal events (∼324.2) in the experiment obtained from the cut-and-count approach agrees with the likelihood fitting result.

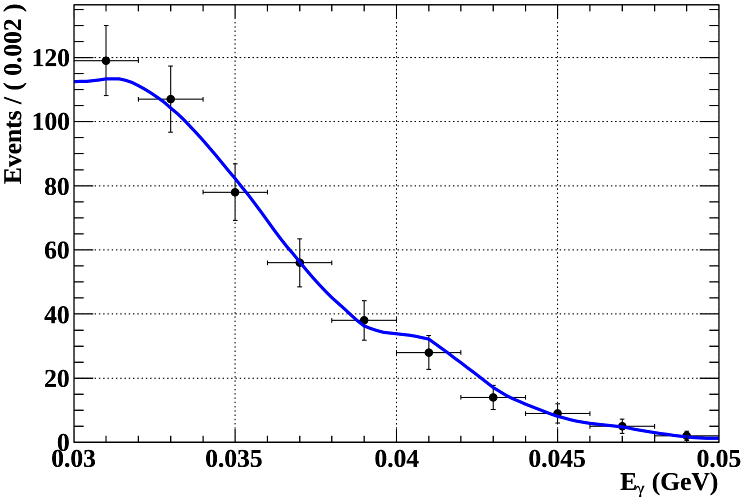
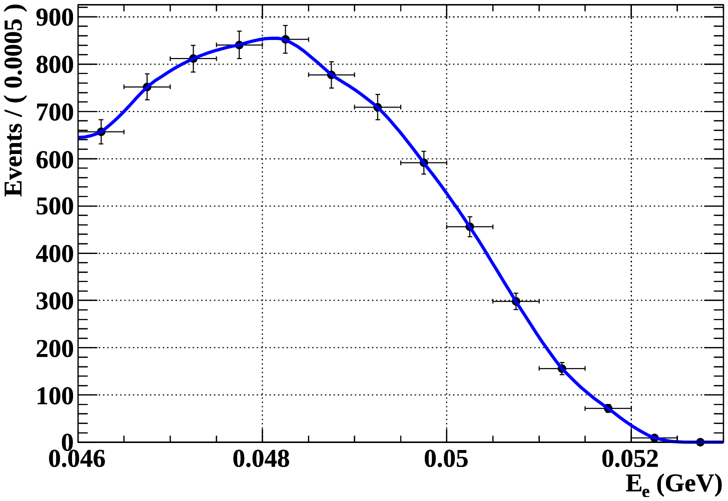


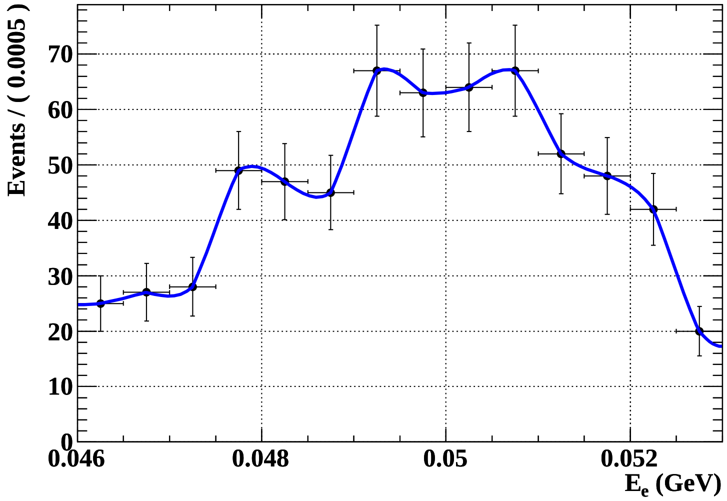
Figure 10.24: Fitting $f(E_\gamma)$ to E_γ distribution in the signal region

Positron Likelihood Fit

$S(E_e)$ and $B(E_e)$ distributions are shown in Figure 10.25.



(a) $S(E_e)$, signal E_e PDF



(b) $B(E_e)$, background E_e PDF

Figure 10.25: E_e PDFs for the radiative decay signal and the background

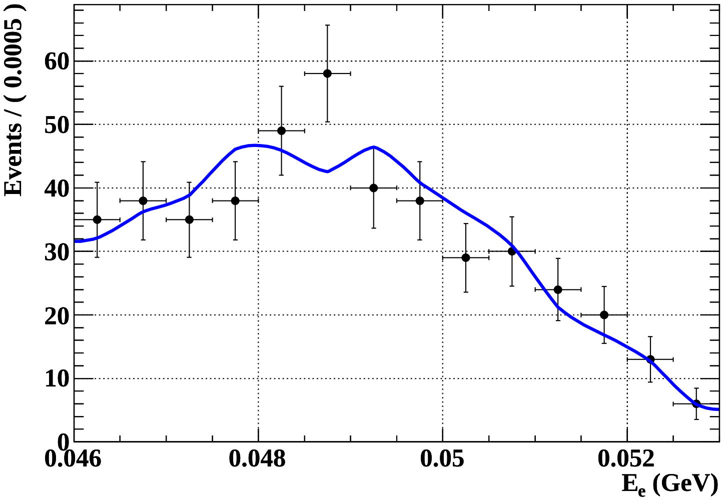


Figure 10.26: Fitting $f(E_e)$ to E_e distribution in the signal region

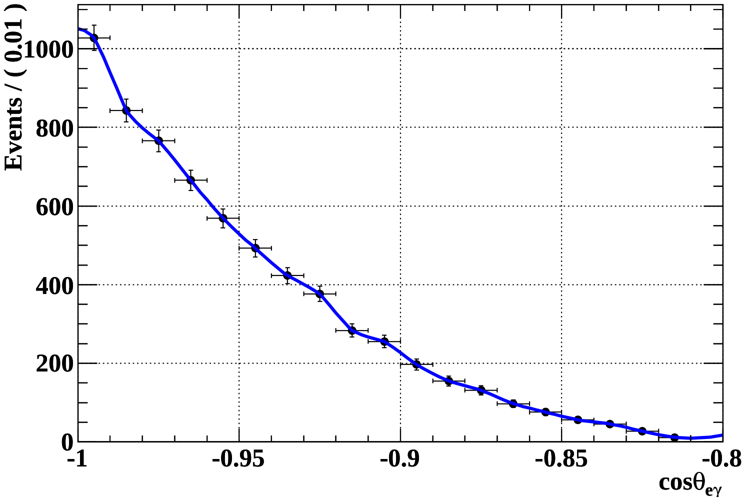
and the fitting result is shown in Figure 10.26. The fit gives,

$$\begin{aligned} N_{e,sig} &= 267.8 \pm 29.8 \\ N_{e,bkgd} &= 185.2 \pm 26.5 \end{aligned} \tag{10.32}$$

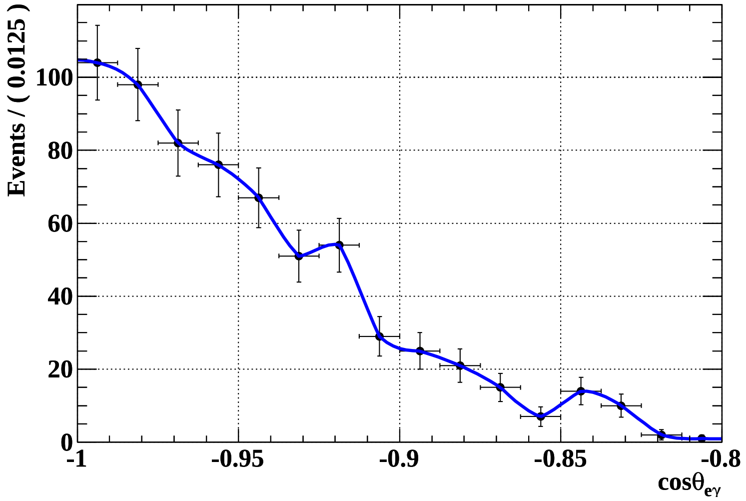
The experiment observation is within 2 standard divinations of the likelihood fitting result.

Opening Angle Likelihood Fit

$S(\cos\theta_{e\gamma})$ and $B(\cos\theta_{e\gamma})$ distributions are shown in Figure 10.27.



(a) $S(\cos\theta_{e\gamma})$, signal $\cos\theta_{e\gamma}$ PDF



(b) $B(\cos\theta_{e\gamma})$, background $\cos\theta_{e\gamma}$ PDF

Figure 10.27: $\cos\theta_{e\gamma}$ PDFs for the radiative decay signal and the background.

and the fitting result is shown in Figure 10.28. The fit gives,

$$\begin{aligned} N_{\cos\theta_{e\gamma},sig} &= 319.1 \pm 53.2 \\ N_{\cos\theta_{e\gamma},bkgd} &= 141.2 \pm 51.5 \end{aligned} \tag{10.33}$$

The detected number of signal events agrees very well with the likelihood fitting result.

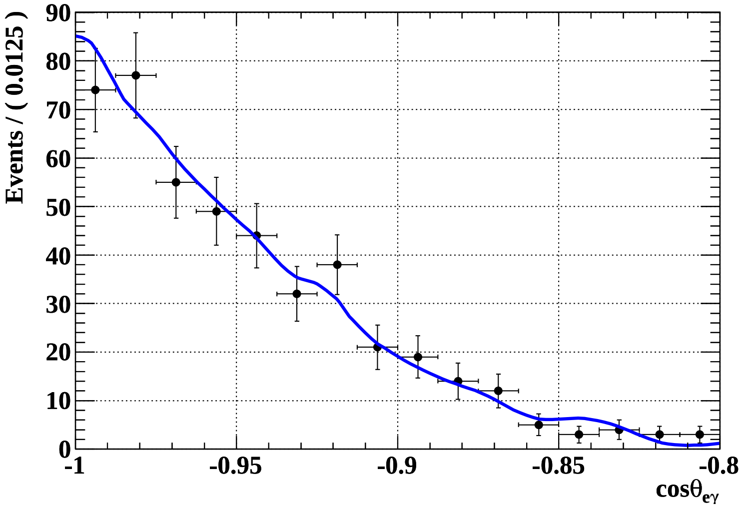
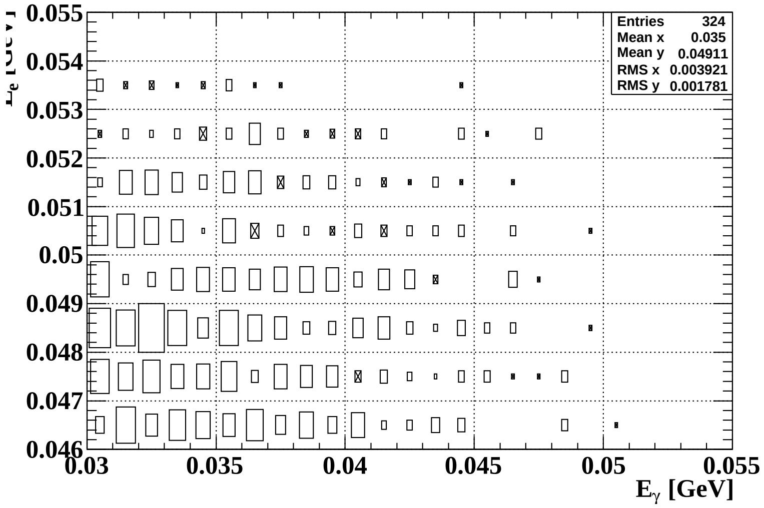


Figure 10.28: Fitting $f(\cos\theta_{e\gamma})$ to $\cos\theta_{e\gamma}$ distribution in the signal region

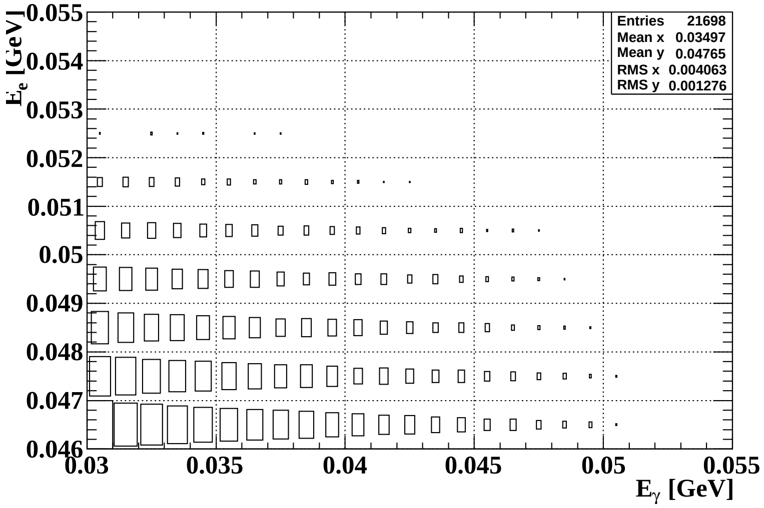
Correlations

The radiative decay is a four-body decay. Three observables have specific kinematic relationships. In order to further verify that the observed events are consistent with being radiative decay events, we have to check that the kinematic correlations of the observables are consistent with radiative decay as well.

The distributions of E_e vs. E_γ, $\cos\theta_{e\gamma}$ vs. E_γ, and $\cos\theta_{e\gamma}$ vs. E_e obtained by the data after background subtraction and the MC are shown in Figures 10.29 10.30 and 10.31. The shapes of theory and experiment plots are similar, and they both have the same kinematic boundaries. If we compare the data plot to the MC plot, we can calculate the normalized $\hat{\chi}^2$ to be ∼4.4, ∼0.65 and ∼3.2.

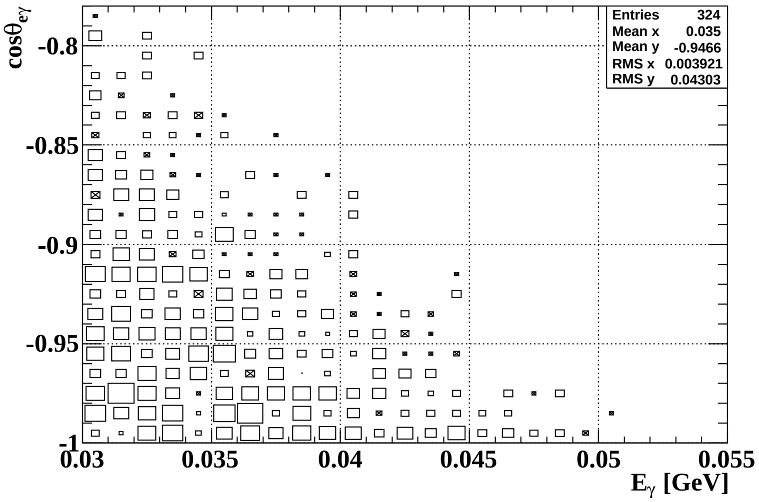


(a) E_e vs. E_γ in data

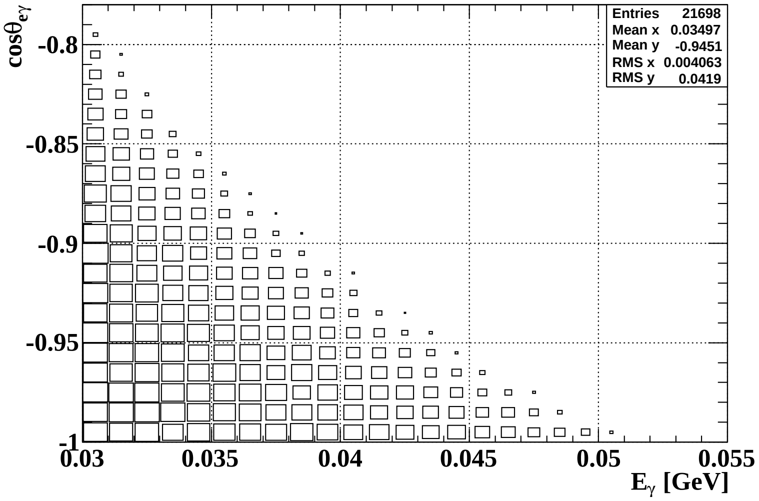


(b) E_e vs. E_γ from the MC

Figure 10.29: E_e vs. E_γ

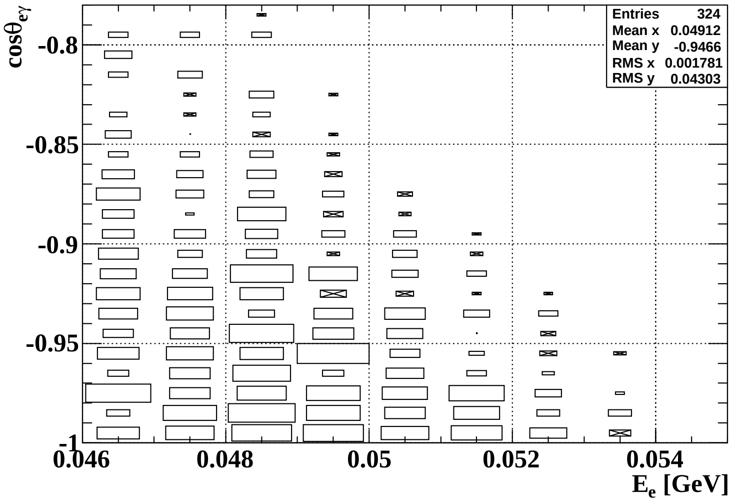


(a) $\cos\theta_{e\gamma}$ vs. E_γ in data

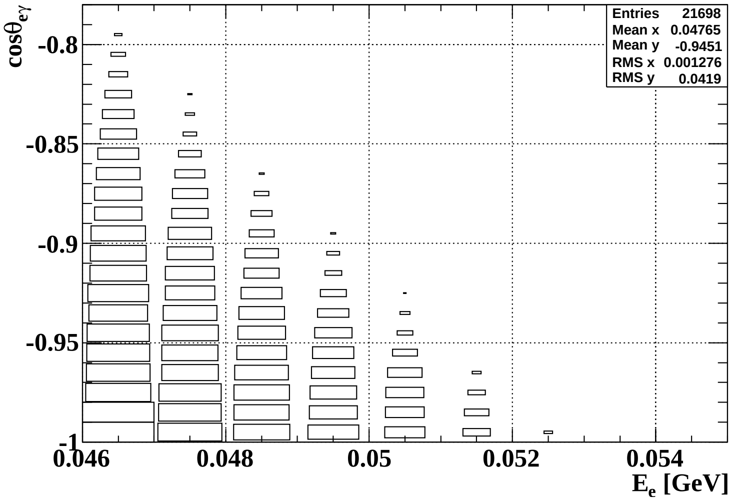


(b) $\cos\theta_{e\gamma}$ vs. E_γ from the MC

Figure 10.30: $\cos\theta_{e\gamma}$ vs. E_γ



(a) $\cos\theta_{e\gamma}$ vs. E_e in data



(b) $\cos\theta_{e\gamma}$ vs. E_e from the MC

Figure 10.31: $\cos\theta_{e\gamma}$ vs. E_e

The kinematic correlations among the observables are qualitatively consistent with theoretical distribution. The normalized $\hat{\chi}^2$ is very good for the $\cos\theta_{e\gamma}$ vs. E_γ plot. In both other cases when E_e is involved, the normalized $\hat{\chi}^2$ is noticeably bigger. This is also consistent with the fact that the E_e distribution has a big statistical fluctuation in the data, as we found earlier in this section.

Summary

Distributions of all three kinematic variables are compared with those of theory by likelihood fitting. We find the signal region E_γ and $\cos\theta_{e\gamma}$ distributions are well fitted and the experiment result of the number of signal events obtained from the cut-and-count approach agrees with both fitting results very well. The fit to the signal region E_e distribution is, however, not as satisfactory. The experiment observation is within 2 standard deviations of the fitting result. This is mostly probably due to statistical fluctuation. In general, distributions of all three kinematic variables are consistent with theory.

The uncertainties of the fitting parameters for $f(E_\gamma)$ and $f(\cos\theta_{e\gamma})$ are much bigger than those for $f(E_e)$, even though the first two are much better fitted. The expected signal distributions for both E_γ and $\cos\theta_{e\gamma}$ are very similar to their background distributions, therefore, the discriminating power of the fit is not as significant as in case of E_e, where the expected signal and the background distributions are greatly different.

Finally, we also have verified that the kinematic correlations among the observables are consistent with the radiative decays.

Chapter 11

Conclusion

In this thesis, we have presented a measurement of the branching fraction for radiative muon decays. The data were recorded in the MEG experiment at PSI during 2008. The data were taken periodically one day per week with a reduced beam intensity of $1.2 \times 10^6\ \mu^+/s$. The positron was measured by a spectrometer with gradient magnetic field. The photon was detected by an innovative ∼900 liter liquid xenon scintillation detector.

We developed precise event reconstruction algorithms and various calibration techniques. The selection criteria were established and the detector performances evaluated. We adopted a normalization scheme using Michel decays. We elaborated the energy dependent detector efficiencies and acceptances to obtain the effective number of observed muons.

Our result for the radiative decay branching fraction calculated by the cut-and-count analysis approach is,

$$B(\mu \to e\nu\bar{\nu}\gamma)|_{E_\gamma > 30 \text{ MeV } E_e > 46 \text{ MeV}} = (2.84 \pm 0.20(stat) \pm 0.05(syst)) \times 10^{-7}$$

This value is in excellent agreement with the prediction of V-A interaction in the Standard Model.

For the purpose of cross check, the distributions in three kinematic variables, E_e, E_γ, and $\theta_{e\gamma}$, are compared to their theoretical distributions by the likelihood fit method. The fitted numbers of signal events agree with the experimentally detected number of events obtained by the cut-and-count approach within 1 to 2 standard deviations. The distributions of E_γ and $\cos\theta_{e\gamma}$ in the signal regions are very well fitted and have very good agreement with the theory. The distribution in E_e agrees less well and is consistent with a statistical fluctuation. Kinematic correlations among the observables are also verified to be consistent with the radiative decays.

The uncertainty in $B(\mu \to e\nu\bar{\nu}\gamma)$ is dominated by statistics. Even though the MEG experiment, which is optimized for two-body decay of back-to-back products with monochromatic energies, has very low acceptance and efficiencies for low energy photons and positrons, systematic uncertainties are well controlled by careful treatment of the energy dependent efficiency calculations.

We verify that in this previously unmeasured region of radiative muon decay, the data is well described by the V-A theory of muon decay.

Finally, a summary plot of the ratio of the measured radiative muon decay branching fraction to the Standard Model prediction of all existing experimental results is shown in Figure 11.1. It shows the overall consistency of all measurements of this branching fraction with the Standard Model.

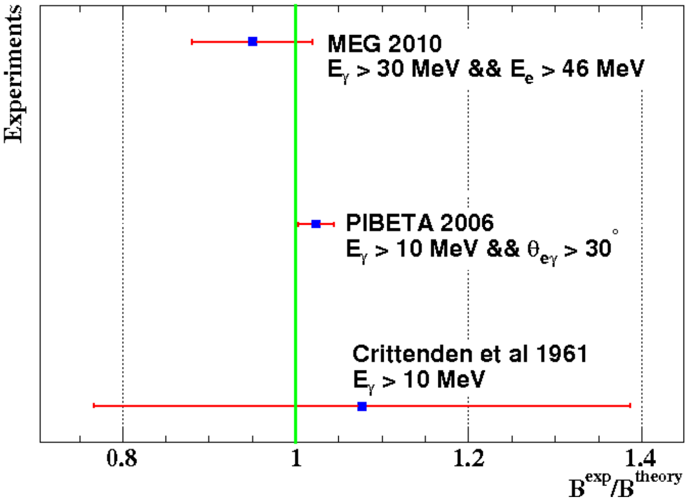


Figure 11.1: B^{exp}/B^{theory} of all existing experimental results

Appendices

A $\mu \to e\gamma$ Decay

A.1 Phenomenology

The $\mu \to e\gamma$ decay is forbidden in the SM due to the accidental conservation of lepton flavor. With a minimum modification of the SM to introduce non-zero neutrino mass, $\mu \to e\gamma$ is induced at one-loop level through the neutrino oscillation, whose Feynman diagram is shown in Figure A.2.

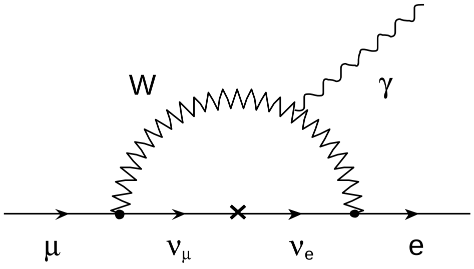


Figure A.2: Feynman diagram of $\mu \to e\gamma$ decay in the SM induced through the mixings of neutrinos.

However, the rate of this process is strongly suppressed by the fourth power of the ratio of the neutrino masses m_{1i} to the W boson mass m_W as,

$$\begin{aligned} B(\mu \to e\gamma) &= \frac{3\alpha}{32\pi} \sum_i \left| U^*_{\mu i} U_{ei} \frac{\Delta m^2_{1i}}{m^2_W} \right|^2 \\ &\approx \left(\frac{\alpha}{2\pi}\right) \sin^2 2\theta_{12} \left(\frac{\Delta m^2_{21}}{M^2_W}\right)^2 < 10^{-54}, \end{aligned} \tag{A.1}$$

where $U_{\alpha i}(\alpha = e, \mu, \tau)$ is the Pontecorvo-Maki-Nakagawa-Sakata matrix element which diagonalizes the neutrino mass matrix by $\nu_\alpha = U_{\alpha i}\nu_i$[65]. In the calculation, $\sin^2 2\theta_{12} = 0.86$ and $\Delta m^2_{21} \sim 8 \times 10^{-5}$ eV^2 are used [9]. This rate is tens of orders of magnitude below the experimental detection of the present day.

Therefore, $\mu \to e\gamma$ decay is free from any SM background, searches of which is among the most sensitive and direct ways to reveal physics beyond the SM, where some models in the supersymmetric grand unification theories (SUSY GUT) and supersymmetric standard models predict $\mu \to e\gamma$ decay with much higher rates[66, 67, 68, 69, 70, 71, 72]. Predictions of some models are described below.

The minimal SUSY extension of the SM (MSSM) with heavy right-handed neutrinos introduced by the seesaw mechanism predicts large $B(\mu \to e\gamma)$, where the neutrino mixings, mostly originate from the neutrino Yukawa couplings, induce cLFV with large rates[72]. Figure A.3 shows the predicted $B(\mu \to e\gamma)$ as a function of the second-generation right-handed neutrino Majorana mass, M_{ν_2}.

In $SO(10)$ SUSY GUT, both left-handed and right-handed sleptons have Yukawa coupling hence, contribute to LFV. $\mu \to e\gamma$ takes place at the level of loop diagrams involving the exchange of a third generation slepton, which enhances the rate. Feynman diagrams of $\mu \to e\gamma$ and predicted values of $B(\mu \to e\gamma)$ are shown in Figure A.4. The predictions are close to the current experimental bound.

More models with predictions of $B(\mu \to e\gamma)$ close to the proposed MEG sensitivity ($\sim 10^{-13}$) can be found in [73]. The search for MEG experiment can therefore make constraints of these models.

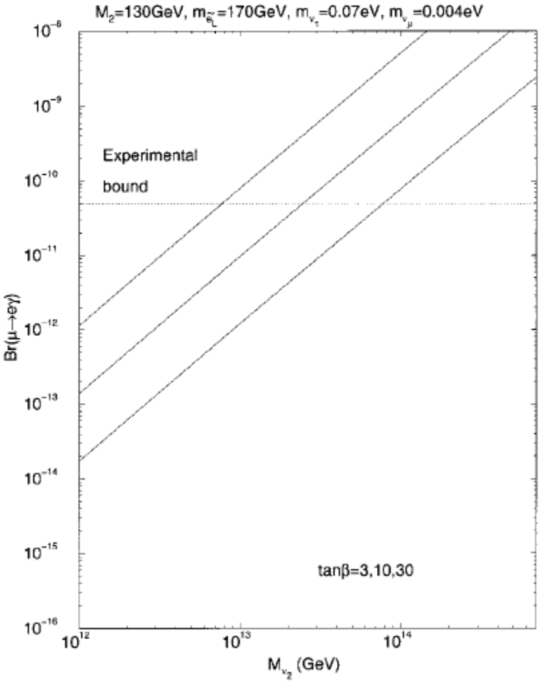
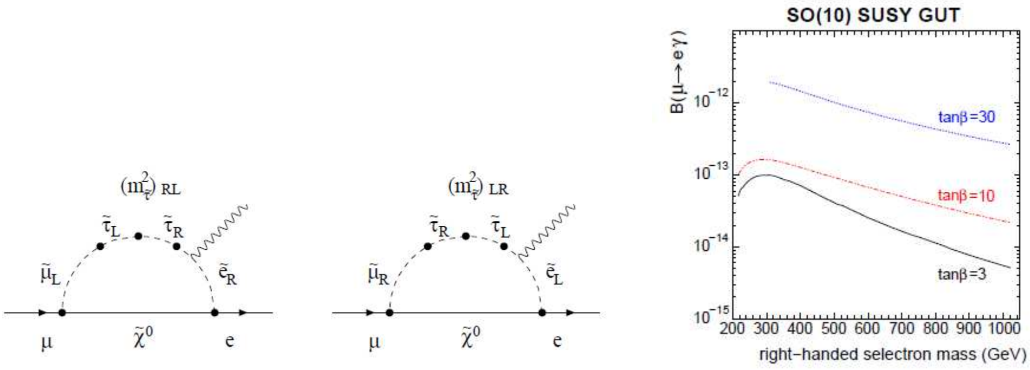


Figure A.3: Dependence of the branching ratio of $\mu \to e\gamma$ on the second-generation right-handed neutrino Majorana mass, M_{ν_2}, in the MSSM with right-handed neutrinos, where M_2 is the $SU(2)$ gaugino (superpartner of gauge boson) mass, $m_{\tilde{e}_L}$ is the left-handed selectron (supersymmetric partner of an electron) mass. The three curves corresponds to $\tan\beta$ (the ratio of the two Higgs vacuum expectation values) = 30, 10 and 3 from top to bottom.



(a) (b)

Figure A.4: Feynman diagrams of $\mu \to e\gamma$ and predicted values of $B(\mu \to e\gamma)$

A.2 The history of $\mu \to e\gamma$ Decay Search

Experimental search for $\mu \to e\gamma$ decay started in 1947, 10 years after the discovery of the muon. Many experiments had carried out and improved the upper limits on $B(\mu \to e\gamma)$ from the first measurement of less than 10% by some ten orders of magnitude. The progress is based on the innovations of morden accelarator technology, which is shown in Figure A.5. Table A.1 summerizes the $\mu \to e\gamma$ search experiments in the last thirty years facilitated by the modern meson facotries.

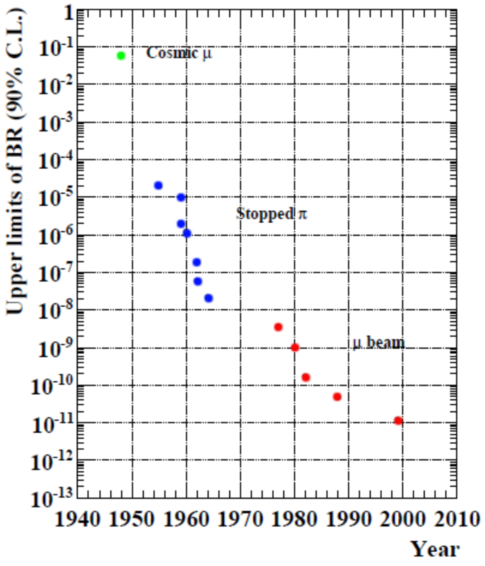


Figure A.5: Progress of the upper limit on $B(\mu \to e\gamma)$ (90% C.L.)

A.3 $\mu \to e\gamma$ Event Signature and Background

$\mu \to e\gamma$ has a very simple event signature that is characterized by 2-body decay. In the muon rest frame $\mu \to e\gamma$ decay has the following characteristics, as shown in Figure A.6(a),

Year	Site	Beam rate	Upper Limit	Ref.
1977	TRIUMF[1]	$2 \times 10^5\ \pi^+/\mathrm{s}$	3.6×10^{-9}	[74]
1980	SIN[2]	$5 \times 10^5\ \mu^+/\mathrm{s}$	1.0×10^{-9}	[75]
1982	LAMPF[3]	$2.4 \times 10^6\ \mu^+/\mathrm{s}$	1.7×10^{-10}	[76]
1988	LAMPF	$4 \times 10^5\ \mu^+/\mathrm{s}$	4.9×10^{-11}	[8]
1999	LAMPF	$1.3 \times 10^7\ \mu^+/\mathrm{s}$	1.2×10^{-11}	[1]

Table A.1: Search for $\mu^+ \rightarrow e^+\gamma$ with 90 % C.L. upper limits on the branching ratio in the last thirty years.

---

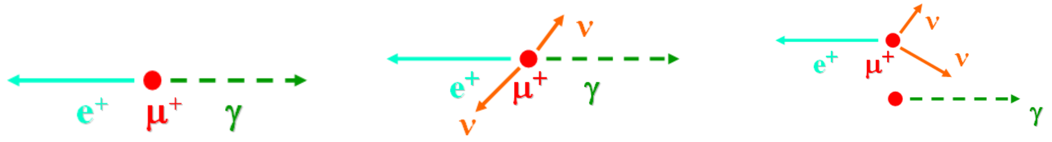
[1] Tri-University Meson Facility

[2] Swiss Institute for Nuclear Research, now Paul Scherrer Institut

[3] Los Alamos Meson Physics Facility

1. Both the produced positron and photon have the same energy of 52.8 MeV, which is half of the muon rest energy.
2. The positron and photon are emitted back to back from the decay vertex.
3. And they are in time coincidence.

Therefore, there are essentially four observables in the experiment to distinguish one $\mu \rightarrow e\gamma$ event from those of the background. They are positron energy, E_e, photon energy, E_γ, relative timing between the two, $t_{e\gamma}$, and the opening angle between the two, $\theta_{e\gamma}$.



(a) $\mu \rightarrow e\gamma$ Signature (b) Physics Background (c) Accidental Background

Figure A.6: $\mu \rightarrow e\gamma$ signature and two types of backgournds

There are two major types of backgrounds. One is the physics background from the radiative muon decay, $\mu \rightarrow e\overline{\nu}\nu\gamma$, with both positron and photon having high energies and coming nearly back-to-back, as shown in Figure A.6(b). The other background is an accidental coincidence both in time and direction of a positron from a normal

Michel decay, $\mu \rightarrow e\overline{\nu}\nu$, with a high energy photon from either a radiative decay, a positron annihilation-in-flight or an external bremsstrahlung of a positron, as shown in Figure A.6(c).

Clearly, the background rates depends on the muon beam rate, R_μ, and the detector resolutions. For the radiative decay background, to first order, the rate is written as,

$$B_{RD}^{background} \propto R_\mu \delta E_e (\delta E_\gamma)^2 (\delta\theta_{e\gamma})^2, \tag{A.2}$$

where δE_e is the detector positron energy resolution, and so forth. $B_{RD}^{background}$ is proportional to δE_e and $(\delta E_\gamma)^2$ due to the features of the end point limit, e.g. $x \rightarrow 1$ and $y \rightarrow 1$, of Equation 2.2 in case of unpolarized muon.

The accidental background rate, to the first order, is written as,

$$B_{acc}^{background} \propto R_\mu^2 \delta E_e (\delta E_\gamma)^2 (\delta\theta_{e\gamma})^2 \delta t_{e\gamma}, \tag{A.3}$$

where its proportionality to R_μ^2 is because that positron and photon in an accidental event are from different sources (muon decays).

Therefore, it is crucial for the MEG experiment to obtain fine resolutions to achieve the aiming limit. Table A.3 lists the desired detector resolutions in sigma. More detailed estimation[28, 64] shows that at the muon beam rate of $3 \times 10^7 s^{-1}$ and with these resolutions, the accidental coincidence is the dominant background.

	resolutions (σ)
δE_e	150 - 200 keV
δE_γ	900 keV
$\delta \theta_{e\gamma}$	7-8.5 mrad
$\delta t_{e\gamma}$	65ps

Table A.2: Desired values of MEG detector resolutions

B Radiative Decay Formula

The differential branching ratio of the radiative muon decay, $\mu \to e\overline{\nu}\nu\gamma$, is given in Equation 2.2, in which the functions, $F(x, y, d)$, $G(x, y, d)$, and $H(x, y, d)$, in the SM are given as,

$$F(x, y, d) = F^{(0)}(x, y, d) + rF^{(1)}(x, y, d) + r^2F^{(2)}(x, y, d), \tag{B.4}$$

$$G(x, y, d) = G^{(0)}(x, y, d) + rG^{(1)}(x, y, d) + r^2G^{(2)}(x, y, d), \tag{B.5}$$

$$H(x, y, d) = H^{(0)}(x, y, d) + rH^{(1)}(x, y, d) + r^2H^{(2)}(x, y, d), \tag{B.6}$$

where $r = (m_e/m_\mu)^2$; x and y are the normalized electron and photon energies, $x = 2E_e/m_\mu$ and $y = 2E_\gamma/m_\mu$; d is given by $d \equiv 1 - \beta\hat{p}_e \cdot \hat{p}_\gamma$; β is defined as $\beta \equiv |\vec{p_e}| / E_e$; $\hat{p}_e$ and $\hat{p}_\gamma$ are the unit vectors the momenta of the positron and photon in the muon rest frame.

$F(x, y, d)$, $G(x, y, d)$, and $H(x, y, d)$ at each order are given as,

$$\begin{aligned} F^{(0)}(x, y, d) \quad = \quad & \frac{8}{d}\left[y^2(3 - 2y) + 6xy(1 - y) + 2x^2(3 - 4y) - 4x^3\right] \\ & +8\left[-xy(3 - y - y^2) - x^2(3 - y - 4y^2) + 2x^3(1 + 2y)\right] \\ & +2d\left[x^2y(6 - 5y - 2y^2) - 2x^3y(4 + 3y)\right] \\ & +2d^2x^3y^2(2 + y), \end{aligned} \tag{B.7}$$

$$
\begin{aligned}
F^{(1)}(x,y,d) \quad &= \quad \frac{32}{d^2}\left[-\frac{y(3-2y)}{x} - (3-4y) + 2x\right] \\
&\quad +\frac{8}{d}\left[y(6-5y) - 2x(4+y) + 6x^2\right] \\
&\quad +8\left[x(4-3y+y^2) - 3x^2(1+y)\right] \\
&\quad +6dx^2y(2+y)
\end{aligned}
\tag{B.8}
$$

$$
F^{(2)}(x,y,d) \quad = \quad \frac{32}{d^2}\left[\frac{(4-3y)}{x} - 3\right] + \frac{48y}{d},
\tag{B.9}
$$

$$
\begin{aligned}
G^{(0)}(x,y,d) \quad &= \quad \frac{8}{d}\left[xy(1-2y) + 2x^2(1-3y) - 4x^3\right] \\
&\quad +4\left[-x^2(2-3y-4y^2) + 2x^3(2+3y)\right] \\
&\quad -4dx^3y(2+y),
\end{aligned}
\tag{B.10}
$$

$$
\begin{aligned}
G^{(1)}(x,y,d) \quad &= \quad \frac{32}{d^2}(-1+2y+2x) \\
&\quad +\frac{8}{d}(-xy+6x^2) \\
&\quad -12x^2(2+y)
\end{aligned}
\tag{B.11}
$$

$$
G^{(2)}(x,y,d) \quad = \quad -\frac{96}{d^2},
\tag{B.12}
$$

$$
\begin{aligned}
H^{(0)}(x,y,d) &= \frac{8}{d}\left[y^2(1-2y)+xy(1-4y)-2x^2y\right] \\
&\quad +4\left[2xy^2(1+y)-x^2y(1-4y)+2x^3y\right] \\
&\quad +2d\left[x^2y^2(1-2y)-4x^3y^2\right] \\
&\quad +2d^2x^3y^3, \qquad \text{(B.13)}
\end{aligned}
$$

$$
\begin{aligned}
H^{(1)}(x,y,d) &= \frac{32}{d^2}\left[-\frac{y(1-2y)}{x}+2y\right] \\
&\quad +\frac{8}{d}\left[y(2-5y)-xy\right] \\
&\quad +4xy(2y-3x) \\
&\quad +6dx^2y^2, \qquad \text{(B.14)}
\end{aligned}
$$

$$
H^{(2)}(x,y,d) = -\frac{96y}{d^2x}+\frac{48y}{d}. \qquad \text{(B.15)}
$$

Bibliography

[1] M. L. Brooks *et al.*, Phys. Rev. Lett. **83**, 1521 (1999).

[2] M. Ahmed *et al.*, Phys. Rev. D **65**, 112002 (2002).

[3] L. Michel, Proceedings of the Physical Society A **63**, 514 (1950).

[4] R. R. Crittenden, W. D. Walker, and J. Ballam, Physical Review **121**, 1823 (1961).

[5] W. Bertl *et al.*, Nuclear Physics B **260**, 1 (1985).

[6] S. J. Freedman *et al.*, Phys. Rev. D **47**, 811 (1993).

[7] U. Bellgardt and G. Otter, Nucl. Phys. B **299**, 1 (1988).

[8] R. D. Bolton *et al.*, Physical Review D **38**, 2077 (1988).

[9] Particle Data Group, Review of Particle Physics, `http://pdg.lbl.gov/`.

[10] A. Lenard, Physical Review **90**, 968 (1953).

[11] C. Fronsdal and H. Überall, Physical Review **113**, 654 (1959).

[12] T. Kinoshita and A. Sirlin, Phys. Rev. Lett. **2**, 177 (1959).

[13] S. G. Eckstein and R. H. Pratt, Annals of Physics **8**, 297 (1959).

[14] Y. Kuno and Y. Okada, Rev. Mod. Phys. **73**, 151 (2001).

[15] D. Pocanic, Study of pion structure through precise measurements of the $\pi^+ \to e^+\nu\gamma$ decay, presented in the 5th International Workshop on Chiral Dynamics (CD2006), Durham, USA, 2006.

[16] B. A. VanDevender, *An Experimental Study of Radiative Muon Decay*, PhD thesis, University of Virginia, 2006, `https://pibeta.web.psi.ch/docs/publications/brent_diss.pdf`.

[17] A. Pifer, Nuclear Instruments and Methods **135**, 39 (1976).

[18] H. Reist, Nuclear Instruments and Methods **153**, 61 (1978).

[19] ROHACELL®, Evonik Rohm GmbH, `http://www.rohacell.com/`.

[20] W. Ootani *et al.*, IEEE Trans. Appl. Supercond. **14**, 568 (2004).

[21] D. Anderson, Nuclear Instruments and Methods in Physics Research **224**, 315 (1984).

[22] D. Green, H. Haggerty, S. Hansen, and M. Takasaki, Nuclear Instruments and Methods in Physics Research A **256**, 305 (1987).

[23] J. Allison *et al.*, Nuclear Instruments and Methods in Physics Research A **310**, 527 (1991).

[24] Y. Ajima *et al.*, Nuclear Instruments and Methods in Physics Research A **443**, 71 (2000).

[25] S. Uno *et al.*, Nuclear Instruments and Methods in Physics Research A **330**, 55 (1993).

[26] H. Hirano *et al.*, Nuclear Instruments and Methods in Physics Research A **455**, 294 (2000).

[27] Garfield, Simulation of gaseous detectors Version 8.01., `http://garfield.web.cern.ch/garfield/`.

[28] H. Nishiguchi, *An Innovative Positron Spectrometer to Search for the Lepton Flavour Violating Muon Decay with a Sensitivity of* 10^{-13}, PhD thesis, University of Tokyo, 2008, `http://meg.web.psi.ch/docs/theses/nishiguchi_phd.pdf/`.

[29] Saint-Gobain Ceramics & Plastics, `http://www.detectors.saint-gobain.com/`.

[30] Hamamatsu Photonics, `http://www.hamamatsu.com/`.

[31] S. Dussoni, *Search for Lepton Flavour Violation: The MEG experiment*, PhD thesis, Universita degli Studi di Genova, 2006.

[32] R. Valle, *Development of the Timing Counter*, PhD thesis, Universita degli Studi di Genova, 2006, `http://meg.web.psi.ch/docs/theses/valle_phd.pdf/`.

[33] E. Aprile and T. Doke, ArXiv e-prints (2009), 0910.4956.

[34] T. Doke, Portugal Physics **12**, 9 (1981).

[35] S. Kubota, M. Hishida, and J. Raun, Journal of Physics C Solid State Physics **11**, 2645 (1978).

[36] S. Kubota, M. Hishida, M. Suzuki, and J. Ruan, Physics Review B **20**, 3486 (1979).

[37] D. L. ed., *Handbook of Chemistry* & *Physics 2004-2005* (Crc Pr. I Llc., 2004).

[38] C. W. Fabjan and F. Gianotti, Reviews of Modern Physics **75**, 1243 (2003).

[39] J. Jortner, L. Meyer, S. A. Rice, and E. G. Wilson, J. Chem. Phys. **42**, 4250 (1965).

[40] L. M. Barkov, A. A. Grebenuk, N. M. Ryskulov, P. Y. Stepanov, and S. G. Zverev, Nuclear Instruments and Methods in Physics Research A **379**, 482 (1996).

[41] A. Hitachi *et al.*, Physics Review B **27**, 5279 (1983).

[42] K. Ozone, *Liquid Xenon Scintillation Detector for the New $\mu \rightarrow e + \gamma$ Search Experiment*, PhD thesis, University of Tokyo, 2005, `http://meg.web.psi.ch/docs/theses/ozoned.pdf/`.

[43] G. M. Seidel, R. E. Lanou, and W. Yao, Nuclear Instruments and Methods in Physics Research A **489**, 189 (2002), arXiv:hep-ex/0111054.

[44] R. Sawada, *A Liquid Xenon Scintillation Detector to Search for the Lepton Flavor Violating Muon Decay with a Sensitivity of* 10^{-13}, PhD thesis, University of Tokyo, 2008, `http://meg.web.psi.ch/docs/theses/sawada_phd.pdf/`.

[45] Telfon®, DuPont, `http://www.DuPont.com/Teflon/`.

[46] A. Baldini *et al.*, Nuclear Instruments and Methods in Physics Research A **565**, 589 (2006).

[47] High Voltage Engineering Europe, `http://www.highvolteng.com/`.

[48] NACRE - European Compilation of Reaction Rates for Astrophysics, `http://pntpm.ulb.ac.be/Nacre/nacre.htm/`.

[49] R. D. Bolton *et al.*, Physical Review Letters **56**, 2461 (1986).

[50] TEXAS INSTRUMENTS, `http://www.ti.com/`.

[51] ANALOG DEVICES, `http://www.analog.com/`.

[52] XILINX Virtex II Pro, `http://www.xilinx.com/`.

[53] SARONIX, Crystal Clock Oscillator, `http://www.pericom.com/saronix/`.

[54] S. Ritt, Nucl. Instr. Meth. A , 470 (2004).

[55] MIDAS, Maximum Integration Data Acquisition System, `https://midas.psi.ch/`.

[56] GEANT, Detector Description and Simulation Tool Version 3.21, `http://wwwasd.web.cern.ch/wwwasd/geant/`.

[57] ZEBRA, A data structure management package. CERN Program Library Long Writeups Q100/Q101.

[58] ROOT, An Object-Oriented Data Analysis Framework, `http://root.cern.ch`.

[59] ROME, Root based Object oriented Midas Extension, `https://midas.psi.ch/rome/`.

[60] HEED, Interactions of particles with gases Version 1.01, `http://consult.cern.ch/writeup/heed`.

[61] Magboltz, Transport of electrons in gas mixture Version 7.1., `http://consult.cern.ch/writeup/magboltz`.

[62] E. C. Ifeachor and B. W. Jervis, *Digital Signal Processing, chapter 7, Finite Impulse Response (FIR) Filter Design.* (Prentice Hall, Inc., 2nd edition, 2002).

[63] R. E. Kalman, Transactions of the ASME–Journal of Basic Engineering **82**, 35 (1960).

[64] Y. Uchiyama, *Analysis of the First MEG Physics Data to Search for the Decay* $\mu^+ \to e^+\gamma$, PhD thesis, University of Tokyo, 2009, `http://meg.web.psi.ch/docs/theses/uchiyama_phd.pdf`.

[65] S. M. Bilenky, S. T. Petcov, and B. Pontecorvo, Physics Letters B **67**, 309 (1977).

[66] R. Barbieri and L. J. Hall, Physics Letters B **338**, 212 (1994), arXiv:hep-ph/9408406.

[67] R. Barbieri, L. Hall, and A. Strumia, Nuclear Physics B **445**, 219 (1995), arXiv:hep-ph/9501334.

[68] J. Hisano, T. Moroi, K. Tobe, M. Yamaguchi, and T. Yanagida, Physics Letters B **357**, 579 (1995), arXiv:hep-ph/9501407.

[69] J. Hisano, T. Moroi, K. Tobe, and M. Yamaguchi, Physics Review D **53**, 2442 (1996), arXiv:hep-ph/9510309.

[70] J. Hisano, T. Moroi, K. Tobe, and M. Yamaguchi, Physics Letters B **397**, 357 (1997), arXiv:hep-ph/9605296.

[71] J. Hisano, D. Nomura, Y. Okada, Y. Shimizu, and M. Tanaka, Physics Review D **58**, 116010 (1998), arXiv:hep-ph/9805367.

[72] J. Hisano and D. Nomura, Physics Review D **59**, 116005 (1999), arXiv:hep-ph/9810479.

[73] J. Hisano, M. Nagai, P. Paradisi, and Y. Shimizu, Journal of High Energy Physics **12**, 30 (2009), 0904.2080.

[74] P. Depommier *et al.*, Phys. Rev. Lett. **39**, 1113 (1977).

[75] van der Schaaf *et.al.*, Nucl. Phys. A **340**, 249 (1980).

[76] W. W. Kinnison *et al.*, Phys. Rev. D **25**, 2846 (1982).