PH. D. THESIS

Analysis of the First MEG Physics Data to Search for the Decay $\mu^+ \rightarrow e^+\gamma$

by

Yusuke Uchiyama

DEPARTMENT OF PHYSICS, SCHOOL OF SCIENCE,
THE UNIVERISTY OF TOKYO

December 2009

Abstract

A search for the lepton-flavor-violating muon decay $\mu^+ \to \mathrm{e}^+\gamma$ was performed on the basis of the data from the initial three months of operation of the MEG experiment. We observed 5.2×10^{11} muon decays and explored the decay with a 90%-confidence sensitivity of

$$\mathcal{S}_{2008} = 1.3 \times 10^{-11}.$$

This sensitivity is comparable with the current best experimental limit of 1.2×10^{-11}.

This search was done in a blind analysis. The number of the $\mu^+ \to \mathrm{e}^+\gamma$ decay events in the data sample was extracted by a maximum likelihood fit. The best-estimated value is at $N_{sig} = 4.3$ and the 90%-confidence interval is evaluated to be $0 \leq N_{sig} \leq 14.7$ including systematic uncertainties. We set an upper limit on the branching ratio,

$$\mathcal{B}(\mu^+ \to \mathrm{e}^+\gamma) < 2.8 \times 10^{-11} \quad (90\ \%\ \text{C.L.}).$$

The large value compared to the sensitivity is considered due to a statistical fluctuation. We do not give a new record of the experimental limit but an independent result from a measurement with a comparable sensitivity.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Since when the muon was discovered, we have tried to understand why more than one generations of fundamental fermions exist and how the generations are related. The standard model (SM) of the elementary particle physics is one of the greatest successes of modern physics. Almost all kinds of phenomena can be explained by the SM. However, the SM is based on the three generations of leptons and quarks, but does not give the reason for the existence of the generations nor prediction of their masses and mixing patterns.

Recently, oscillations between types of neutrinos have been established by various experiments. The mixing patterns have been revealed progressively. Meanwhile, in the quark sector, the flavor-changing neutral currents (FCNC) have been studied for many years. The quark mixings are described by the quark-mixing matrix (known as CKM matrix). Understanding the mixings in charged-lepton sector will obviously give us a new insight for the generation mechanism of the fundamental particles and their interactions. So far, no evidence of charged-lepton-flavor violation (cLFV) is found. Within the SM, the conservation of lepton flavor is considered accidental; there is no explicit gauge symmetry for which lepton flavor is the conserved quantity, but, in the absence of neutrino mass, there is no mechanism for breaking this symmetry. Even if we introduce a tiny but finite neutrino masses in the SM, the rate of cLFV is too small to reach with possible experimental technique because of the GIM mechanism.

If there are either a new force mediated by a new gauge boson with non-diagonal lepton couplings or a new class of heavy particles with lepton-flavor mixings in this new sector, the cLFV processes would be induced. The possibility of cLFV exists in essentially all extensions to the SM. Because cLFV processes are free from the SM background, searches for them are among the most sensitive ways to look for physics beyond the SM. In particular, some models incorporating supersymmetry predict large branching ratios that can be reached with current or near-future experiments.

The cLFV processes have been investigated experimentally in various channels. The muon system gives the most stringent constraint to the new physics, while the kaon and tau systems also give important results. The $\mu^+ \to e^+\gamma$ decay is the simplest and the most famous process out of them. The current best limit on the branching ratio of the decay is given as $\mathcal{B}(\mu^+ \to e^+\gamma) < 1.2 \times 10^{-11}$ (90 % C.L.) by the MEGA experiment [1] in 1999. This sensitivity already entered into the phase space predicted by several interesting models, and started to set constraint to them.

In 1999, a proposal for a new $\mu^+ \to e^+\gamma$ decay search [2] by a group of Japanese physicists was approved by the research committee of the Paul Scherrer Institut (PSI). Since then, an experimental collaboration, which is now called the MEG (Muon to Electron and Gamma) collaboration, have been formed by approximately 60 physicists from Japan, Switzerland, Italy, Russia, and the United States. MEG aims at searching for the decay with a sensitivity two orders of magnitude below the current limit. The main features of the MEG experiment are the world's most intense DC muon beam from the 590 MeV ring cyclotron at PSI, a novel positron spectrometer with a specially graded magnetic field, and an innovative 900 liter liquid xenon gamma-ray detector. MEG started physics data taking in autumn 2008 after intense R&D and construction periods for 10 years. In 2008, it ran for three months and accumulated data of muon decay. We measured $\sim 1 \times 10^{14}$ muons stopping on the target, the largest data set so far.

The theme of this thesis is a search for the $\mu^+ \to \mathrm{e}^+\gamma$ decay using the first three months data of MEG taken in 2008. In Chapter 2, we summarize theoretical, phenomenological, and experimental features of the $\mu^+ \to \mathrm{e}^+\gamma$ search. In Chapter 3, we describe the experimental apparatus of MEG in detail. The run 2008 that is the full data set of this analysis is summarized in Chapter 4. The event reconstruction and calibration that give bases of the analysis are described in Chapter 5 and Chapter 6, respectively. Then the performance of the detector is evaluated in Chapter 7. The analysis to search for the $\mu^+ \to \mathrm{e}^+\gamma$ decay is described in Chapter 8 and the result is discussed. Finally, we conclude this thesis in Chapter 9.

Chapter 2

$\mu \rightarrow \mathrm{e}\gamma$ Decay

We discuss theoretical and experimental features of the $\mu \rightarrow \mathrm{e}\gamma$ decay in this chapter. First, we shortly summarize 'standard' muon decays. Theoretical frameworks and phenomenology of $\mu \rightarrow \mathrm{e}\gamma$ decay are briefly reviewed in the second section. In this section, we see that $\mu \rightarrow \mathrm{e}\gamma$ decay search can be a clear probe to new physics beyond the SM. In the next section, experimental features as well as the history of $\mu \rightarrow \mathrm{e}\gamma$ search are summarized.

2.1 Muon Decay in the Standard Model

The muon is the second-generation charged lepton which is a replication of the electron with a heavier mass. It interacts through the electromagnetic and weak interactions. It also couples to the Higgs boson. The Lagrangian for those interactions are given as

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

Muons decay through the charged-weak current interaction mediated by $W_{\mu}^{\pm}$ guage bosons, the second line of the Eq.2.1 ($V - A$ interaction). The decay modes and their branching ratios are summarized in Table 2.1. The dominant process, $\mu \rightarrow \mathrm{e}\nu_{\mu}\overline{\nu}_{\mathrm{e}}$, is called "Michel decay" [3].

2.2 Physics Motivations

2.2.1 Beyond the Standard Model

The SM is based on the fundamental fermions, quarks and leptons, and the gauge theory. It describes the interactions of quarks and leptons that are the constituents of all matter that we know about. The strong interactions are described by quantum chromodynamics (QCD) while the electromagnetic and the weak interactions have been unified into a single

Table 2.1: Decay modes and branching fractions of muon (Listed in PDG table [4]).

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) %	[5]
$\mu^- \to \mathrm{e}^-\nu_\mu\overline{\nu}_\mathrm{e}\mathrm{e}^+\mathrm{e}^-$	$(3.4 \pm 0.4) \times 10^{-5}$	[6]
$\mu^- \to \mathrm{e}^-\nu_\mathrm{e}\overline{\nu}_\mu$	< 1.2 %	[7]
$\mu^- \to \mathrm{e}^-\gamma$	$< 1.2 \times 10^{-11}$	[1]
$\mu^- \to \mathrm{e}^-\mathrm{e}^+\mathrm{e}^-$	$< 1.0 \times 10^{-12}$	[8]
$\mu^- \to \mathrm{e}^-\gamma\gamma$	$< 7.2 \times 10^{-11}$	[9]

electroweak framework. They are all gauge interactions, which are based on the principle of gauge symmetry. Quarks and leptons consist of three generations, each of which has identical gauge quantum numbers.

This theory has proven to be extremely successful in describing a tremendous variety of experimental data over an energy range from a fraction of an electron volt to about 100 GeV, a range of over 12 orders of magnitude. Nevertheless, it is thought as an effective theory in low-energy approximation, and more fundamental theory is thought to exist. The reason is that the SM contains some fundamental and theoretical problems. In addition, recently a few experiments and observations have indicated results which cannot be described in the SM. Clearly, it is desirable to have a more unified theory that can combine all these three interactions as components of a single force with one coupling constant.

The SM does not account for the existence of three generations of quarks and leptons. This problem is known as a flavor puzzle of the SM. We lack any understanding of particle masses and mixing patterns, which results in the large number of underlying parameters in the SM.

Another problem of the SM is the hierarchy problem. It is related to the huge difference of energy scales between the week scale of $\mathcal{O}(100$ GeV$)$ and the reduced Planck scale of $\mathcal{O}(10^{18}$ GeV$)$, where quantum gravitational effects become important. To understand this problem, let us think about the Higgs mass. The Higgs mass parameter receives enormous quantum corrections from the virtual effects of every particle which couples to the Higgs field,

$$m^2_{H_{SM}}(\mathrm{phys}) \approx m^2_{H_{SM}} + \frac{c}{16\pi^2}\Lambda^2, \tag{2.2}$$

where Λ is a cut-off energy interpreted as the scale at which the SM ceases to be valid. In the SM, incredible fine tuning is required: if we assume the validity of the SM as a low-energy effective theory below the GUT scale, and take $\Lambda = M_{GUT} \sim 10^{16}$ GeV, then the Lagrangian mass parameter $m^2_{H_{SM}}$ will have to be fine-tuned to 1 part in 10^{26} to provide the needed cancellation that will maintain a physical Higgs mass below its unitarity limit. To avoid excessive fine tuning between the two terms on the right-hand side, we would have to deduce that $\Lambda \leq \mathcal{O}(\mathrm{TeV})$. Thus it indicates that new degrees of freedom that are not included in the SM are there at TeV-scale. Those degrees of freedom could be a new force mediated by a new gauge boson or a new class of heavy particles.

Incompleteness of the SM is also given by experimental arguments such as neutrino oscillations and existence of cold dark matter in the Universe for which there is no candidate in the SM. In addition, the anomalous magnetic moment $(g-2)$ of the muon and some low-energy CP-violating observables measured at the B-factories could indicate the SM incompleteness, whereas the hadronic uncertainties as well as the limited experimental resolutions prevent any conclusive evidence of new physics in the quark sector.

Neutrino Mass and Mixing

In the (minimum) SM, neutrinos are strictly massless. No neutrino Dirac mass term can be introduced because of the absence of right-handed neutrinos, and no Majorana mass terms can be generated because of exact $B-L$ conservation. Owing to the massless neutrinos, there is no leptonic mixing in the SM, which in turn leads to separate lepton-number (lepton-flavor) conservation. Therefore, the observation of neutrino oscillations is evidence of physics beyond the SM. In another words, the lepton-flavor violation (LFV) is forbidden by an ‘accidental’ symmetry in the SM, and it can occur in all extensions of the SM.

If neutrinos are not massless, their mass matrix, just as in the case for quarks, will be non-diagonal and complex. The mass eigenstates are different from flavor ones

$$\nu_\alpha = \sum_i U_{\alpha i} \nu_i, \tag{2.3}$$

where $\nu_\alpha = \nu_{\mathrm{e}}, \nu_\mu, \nu_\tau$ are weak-flavor eigenstates and $\nu_i = \nu_1, \nu_2, \nu_3$ are mass eigenstates with mass eigen values m_1, m_2, m_3. U is a unitary matrix known as Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix. A mass eigenstate of i after a time interval t is given by

$$|\nu_i(t)) >= e^{-iE_i t} |\nu_i(0)\rangle. \tag{2.4}$$

Thus a neutrino of the flavor α after a time interval of t is given by

$$|\nu_\alpha(t) >= \sum_i U_{\alpha i} e^{-iE_i t} |\nu_i(0)\rangle. \tag{2.5}$$

The probability of finding flavor ν_β in ν_α beam at a distance x from the source is given by

$$P_{\nu_\alpha \to \nu_\beta} = \sum_i |U_{\alpha i}|^2 |U_{\beta i}|^2 + \sum_{i \neq j} U_{\alpha i} U^*_{\beta i} U^*_{\alpha j} U_{\beta j} \cos\left(\frac{2\pi x}{L_{ij}}\right), \tag{2.6}$$

where $L_{ij} = 2\pi/(E_i - E_j) \simeq 4\pi p/|m_i^2 - m_j^2|$ called the oscillation length. To have neutrino oscillation, we must have nonzero neutrino masses and mixing angles.

It is possible to accommodate the Dirac mass terms for the neutrinos if $SU(2)$ singlet right-handed neutrinos are added to the minimal SM

$$m_D \bar{\nu}_L \nu_R. \tag{2.7}$$

If Yukawa coupling for neutrinos is very small, the small masses of neutrinos can be explained. For example, the Yukawa coupling constant should be $\mathcal{O}(10^{-11})$ for a neutrino mass of 1 eV$/c^2$.

A more natural explanation for the small neutrino masses is given by the see-saw mechanism [10, 11]. Since neutrinos are electrically neutral, it is possible that their anti-particles are themselves. Then the Majorana mass term can be included in the Lagrangian,

$$m_L \nu_L \nu_L^c + M_R \nu_R \nu_R^c. \tag{2.8}$$

The complete mass term for the neutrino is written as

$$\begin{aligned} & m_D \bar{\nu}_L \nu_R + m_L \nu_L \nu_L^c + M_R \nu_R \nu_R^c + (h.c.) \\ &= (\bar{\nu}_L \bar{\nu}_R^c) \begin{pmatrix} m_L & m_D \\ m_D & M_R \end{pmatrix} \begin{pmatrix} \nu_L \\ \nu_R \end{pmatrix}. \end{aligned} \tag{2.9}$$

In the see-saw model, the left-handed Majorana mass is set to zero ($m_L = 0$) and the Dirac mass is required to be much smaller than the right-handed Majorana mass ($m_D \ll M_R$). Then the neutrino physical masses are

$$m_N \simeq M_R, \quad m_\nu \simeq \frac{m_D^2}{M_R}. \tag{2.10}$$

One neutrino remains super heavy while the other gets a tiny mass. For example, if M_R is 10^{15} GeV and the Dirac mass is of the order of 100 GeV, then neutrino mass becomes naturally $\mathcal{O}(10^{-2})$ eV. The Majorana mass terms violate lepton number by two units.

The neutrino mass can be written with Yukawa couplings Y_ν

$$m_\nu = -Y_\nu M_R^{-1} Y_\nu^T \langle H_u \rangle^2, \tag{2.11}$$

where $\langle H_u \rangle$ is the vacuum-expectation value (VEV) acquired by the up-sector Higgs field. The matrix Y_ν can be written in the general form

$$Y_\nu = U_{PMNS} \sqrt{m_\nu} R \sqrt{M_R} / \langle H_u \rangle, \tag{2.12}$$

where R is an arbitrary complex orthogonal matrix. A complete knowledge of Y_ν cannot be achieved by only low-energy observables from the neutrino sector.

$\mu \to e\gamma$ Decay through the Neutrino Oscillation

The neutrino oscillations show that lepton flavor is not conserved in neutrino sector. It also induces the transition between charged leptons at the one-loop level through diagrams such as shown in Figure 2.1. However, the LFV processes in charged lepton (cLFV) are severely suppressed because of the GIM mechanism. The $\mu \to e\gamma$ rate in the minimum extension of SM is given as

$$\begin{aligned} \mathcal{B}(\mu \to e\gamma) &= \frac{3\alpha}{32\pi} \sum_i \left| U_{\mu i}^* U_{ei} \frac{\Delta m_{1i}^2}{m_W^2} \right|^2 \\ &\approx \left(\frac{\alpha}{2\pi} \right) \sin^2 2\theta_{12} \left(\frac{\Delta m_{21}^2}{M_W^2} \right)^2 < 10^{-54}, \end{aligned} \tag{2.13}$$

where $\sin^2 2\theta_{12} = 0.86$ and $\Delta m_{21}^2 \sim 8 \times 10^{-5}$ eV^2 are used [4]. It is far from experimental reach.

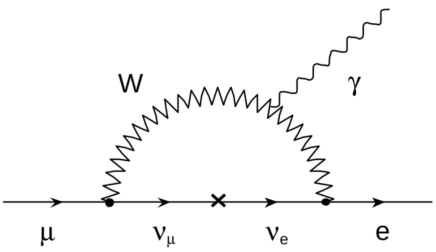


Figure 2.1: Example of Feynman diagram describing the $\mu \to e\gamma$ decay in the SM with a neutrino-mass extension.

In conclusion, the cLFV processes are free from the SM background. Searches for cLFV processes are among the most sensitive ways to investigate physics beyond the SM. In the following sub-sections, we review some models beyond the SM focused on the supersymmetric (SUSY) scenarios. In those models, the branching ratio of $\mu \to e\gamma$ decay is enhanced by the new physics and becomes accessible with present or near-future experimental techniques.

2.2.2 Supersymmetry and Lepton-Flavor Violation

The SUSY is a symmetry between bosons and fermions with the same quantum numbers apart from their spins. A SUSY transformation turns a bosonic state into a fermionic state, and vice versa. It predicts for every particle a supersymmetric partner with, in the limit of non-broken SUSY, the same mass. Those two particles belong to a super-multiplet. The Higgs mass is kept under control by the cancellation between those two since the contributions to the quantum correction of the two have opposite sign due to the difference in Fermi-Bose statistics. In this way, the hierarchy problem in the SM is solved naturally in SUSY extensions.

Table 2.2 lists the super-partners in the minimum SUSY extension of the SM (MSSM). The super-partners differ by 1/2 unit of spin from the corresponding particles. After electroweak symmetry breaking, the wino ($\tilde{W}$), the bino ($\tilde{B}$), and Higssino ($\tilde{H}$) mix one another and form two charged Dirac fermions called charginos ($\tilde{\chi}_i^{\pm}; i = 1, 2$), and four Majorana fermions called neutralinos ($\tilde{\chi}_i^0; i = 1-4$). In general, SUSY models contain at least two Higgs doublet fields to keep the SUSY invariance for three types of the Yukawa coupling constants: one Higgs field provides the mass terms for up-type quarks while the other provides mass terms for the down-type quarks and charged leptons. The ratio of the VEVs of the two is called $\tan\beta$.

SUSY breaking and SUSY flavor problem If the symmetry is exact, a particle and its super-partner are degenerated and have the same mass. However, no superparticles with the mass are not observed. Thus SUSY is broken. The LFV would originate from the misalignment between particle and superparticle mass eigenstates. In the basis where the lepton mass matrix is diagonalized, the presence of nonzero off-diagonal matrix elements in the slepton mass matrix would induce LFV. However, constrains from LFV and

Table 2.2: Particles content in the MSSM.

Particle	Spin	SUSY particles	Spin
quark (q)	1/2	squark ($\tilde{q}$)	0
lepton (l)	1/2	slepton ($\tilde{l}$)	0
gluon (g)	1	gluino ($\tilde{g}$)	1/2
$W^{\pm}, Z^0, \gamma$	1	chargino ($\tilde{\chi}_i^{\pm}$; $i = 1, 2$)	1/2
Higgs boson ($h, H, A, H^{\pm}$)	0	neutralino ($\tilde{\chi}_i^0$; $i = 1 - 4$)	1/2

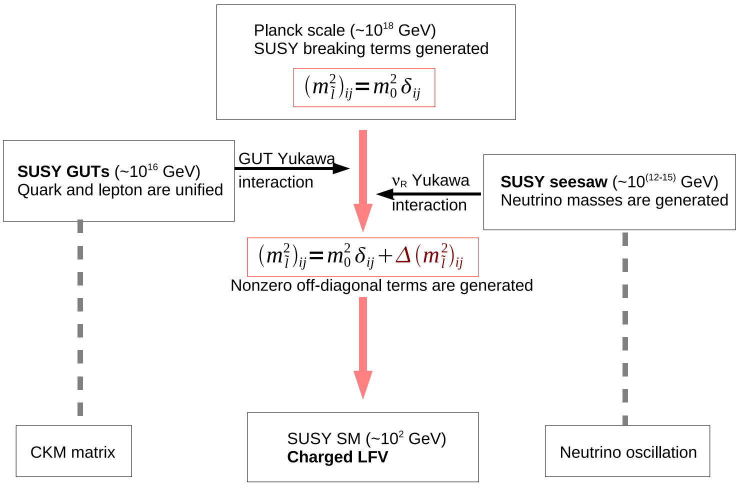


Figure 2.2: cLFV generation mechanism in mSUGRA scenario. The off-diagonal terms in the slepton mass matrix are induced by renormalization effects due to GUT and/or neutrino interactions.

FCNC in quark sector suggest the presence of a quite small amount of fermion-sfermion misalignment (SUSY flavor problem). There should be a special suppression mechanism for the flavor mixing of sfermions from the dynamics of SUSY breaking.

There are several scenarios that solve the SUSY flavor problem. One of them is the super gravity model (SUGRA). In SUGRA, all sleptons and squarks have the same mass (the universal structure) and the mass matrix is diagonal. Thus in this model, there is no cLFV at the Planck scale. However, if there are some interactions violating flavor between the Planck scale and the weak scale, cLFV can be induced by the slepton mass matrix through the renormalization group (RG) evolution. Here, we introduce so-called the Mass Insertion (MI) notation to denote the various flavor-violating entries of the slepton matrix. The MI parameters $\delta_{ij}^{L/R}$ for the left/right-handed sleptons are defined as

$$\delta_{ij}^{L/R} = \frac{(m^2_{\tilde{L}/\tilde{R}})_{ij}}{m^2_{\tilde{l}}}, \tag{2.14}$$

where $(m^2_{\tilde{L}/\tilde{R}})$ is the left/right-handed slepton mass matrix, whose off-diagonal entries are zero at the high scale but generated by RG evolution at the weak scale. This mechanism of cLFV generation is illustrated in Figure 2.2. Those interaction could be Yukawa interaction of neutrino and/or that of grand unified theory.

2.2.3 Right-Handed Neutrinos as a LFV source

MSSM with Right-Handed Neutrino

In a MSSM with right-handed neutrino model, the neutrino Yukawa couplings become large sources of LFV [12],

$$(m^2_{\tilde{L}})_{i\neq j} = -\frac{3m_0^2 + A_0^2}{8\pi^2}(Y_\nu)_{ik}(Y_\nu^\dagger)_{jk}\ln\Big(\frac{M_X}{M_{Rk}}\Big), \tag{2.15}$$

where M_X denotes the scale of SUSY-breaking mediation; m_0 and A_0 stand for the universal SUSY-breaking scalar mass and trilinear coupling at M_X, respectively. A complete determination of $(m^2_{\tilde{L}})_{i\neq j}$ would require a complete knowledge of the neutrino Yukawa matrix.

In general, there is no direct relationship between the neutrino mixing and the slepton mixing. If, however, we assume that the neutrino mixing mostly originates from the neutrino Yukawa coupling constants ($R \approx 1$), the information from neutrino oscillations can be related to the slepton mixing. Figure 2.3 shows the predicted branching ratio of the $\mu \to \mathrm{e}\gamma$ decay as a function of a Majorana mass scale. A large fraction of the range for M_R is already excluded by the current experimental limit.

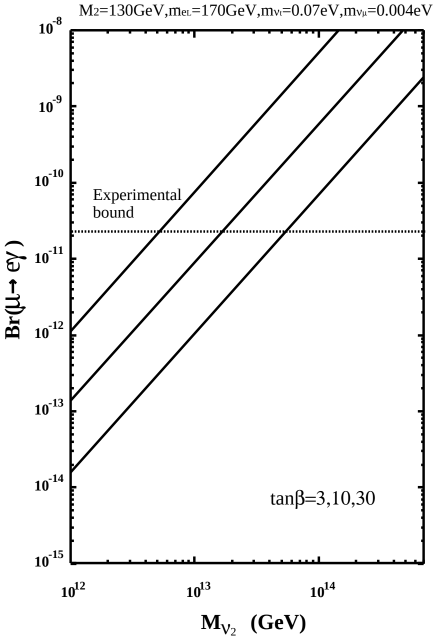


Figure 2.3: Dependence of the branching ratio of $\mu \to \mathrm{e}\gamma$ on the second-generation right-handed neutrino Majorana mass M_{ν_2} in the MSSM with right-handed neutrinos. The assumptions of the SUSY parameters are described in the figure. The larger $\tan\beta$ corresponds to the upper curve.

2.2.4 LFV in Grand-Unified Theories

Grand-unified theories (GUT) try to unify $SU(2) \times U(1)$ electroweak interaction and $SU(3)$ strong interaction in a single simple gauge group. The simplest GUT model is the minimal $SU(5)$ model. This model incorporates leptons and quarks into the same multiplets. Therefore, the coupling constants for the strong and electroweak interactions must be related to each other. As a consequence, there would be a new interaction in which leptons and quarks transform one another mediated by heavy bosons. This interaction brakes both baryon-number (B) and lepton-number (L) symmetries and leads to proton decay.

The proton decay rate predicted by the minimal $SU(5)$ is larger than the experimental limit by, for example, KAMIOKANDE [17]. Furthermore, the minimal $SU(5)$ fails to unify the gauge couplings of the SM at a single scale. Thus, the minimal $SU(5)$ model is excluded already. On the other hand, an extended version of $SU(5)$ with SUSY is found to unify them greatly.

It was first realized by Hall and Barbieri, that LFV will occur at experimentally accessible levels in a large class of SUSY-GUT models [13].

SUSY $SU(5)$

The $SU(5)$ running from the soft-braking scale M_X to the GUT scale already breaks the universality by generating LFV entries at M_{GUT}. Since both Q and e^c are hosted in the **10**, the CKM matrix mixing the left-handed quarks will give rise to off-diagonal entries in the running of the right-handed slepton masses due to the colored Higgs,

$$(m^2_{\tilde{R}})_{i \neq j} = -3 \frac{3m_0^2 + A_0^2}{8\pi^2} V_{ti} V_{tk} \ln\Big(\frac{M_X}{M_{GUT}}\Big). \tag{2.16}$$

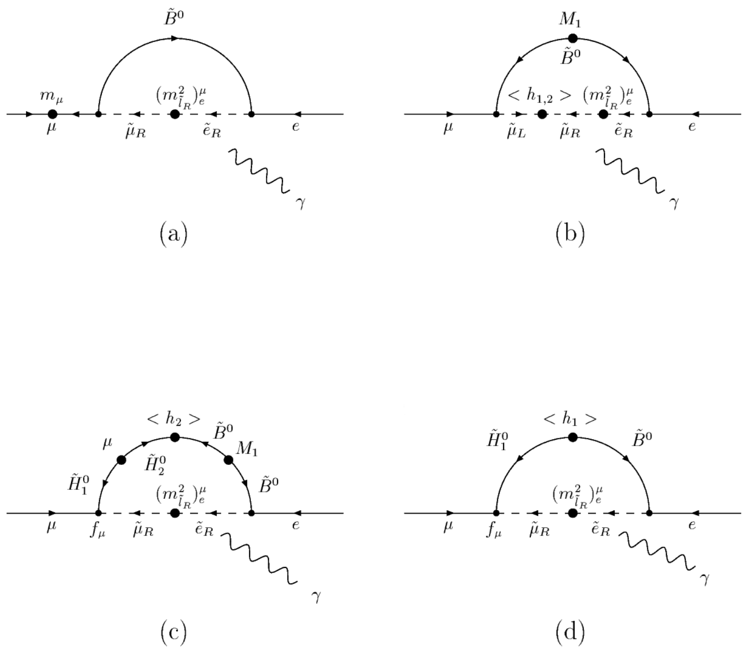
It is independent source of LFV from the neutrino Yukawa matrix. The LFV appears only in the right-handed slepton sector, in contrast to the neutrino Yukawa case, which has only left-handed slepton sector. The Feynman diagrams which can enhance the $\mu \to \mathrm{e}\gamma$ rate are shown in Figure 2.4.

However, except for few cases, the rate of $\mu \to \mathrm{e}\gamma$ turns out to be suppressed. The reasons of the suppression are the following: the relevant sources of LFV are CKM suppressed; the amplitude involving only δ^R do not have chargino contributions; and in certain regions of the parameter space, there could be cancellations between the bino and the Higgsino-bino-Higgsino contributions. Only for light SUSY particles and for moderate to large values of $\tan\beta$ and A_0, large values of $\mu \to \mathrm{e}\gamma$ branching ratio could be achieved. The predicted branching ratio calculated in [14] is shown in Figure 2.5.

SUSY $SU(5)$ with Right-handed Neutrinos

The main problems of pure SUSY $SU(5)$ model are mass-less neutrinos, unless one breaks R-parity. One way is going to $SO(10)$, where the the see-saw mechanism can naturally arise. The other way out includes adding singlets, right-handed neutrinos ($SU(5)_{RN}$).

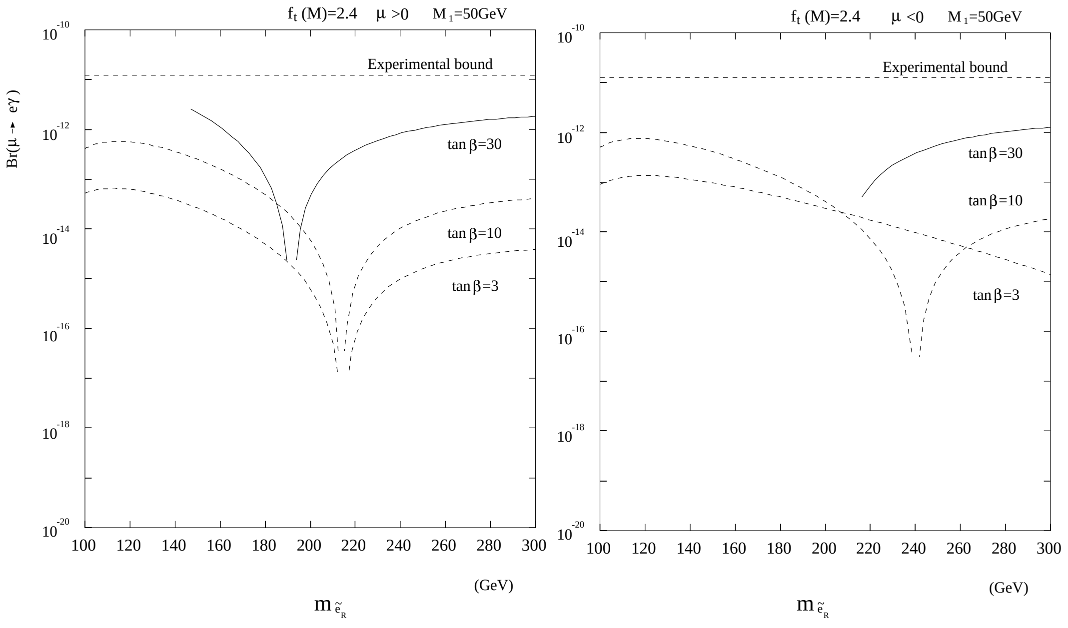
The situation for LFV can drastically change in the $SU(5)_{RN}$ [15]. In this case, in addition to $(m^2_{\tilde{R}})_{ij}$, we also have the $(m^2_{\tilde{L}})_{ij}$ MIs. In the $SU(5)_{RN}$, the dominant contributions to $\mu \to \mathrm{e}\gamma$ arise either from $\delta^L_{\mu e}$ (and $\delta^L_{\mu\tau}\delta^L_{\tau e}$) through the loop exchange of charginos/sneutrinos or from $\delta^L_{\mu\tau}\delta^R_{\tau e}$ through the loop exchange of a pure bino.



(a) (b)

(c) (d)

Figure 2.4: Feynman diagrams contributing to $\mu \rightarrow e\gamma$ in SUSY $SU(5)$ model.

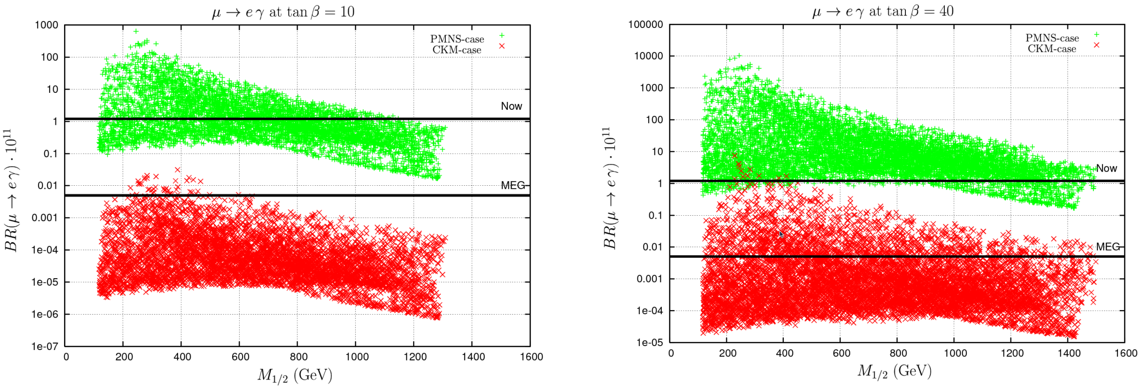


(a) (b)

Figure 2.5: Branching ratio of $\mu \rightarrow e\gamma$ as a function of right-handed selectron mass in SUSY $SU(5)$.

SUSY $SO(10)$

There are several features that make the $SO(10)$ appealing. In $SO(10)$ theories, the see-saw mechanism is naturally present, and the neutrino Yukawa couplings are related



(a) (b)

Figure 2.6: Branching ratio of $\mu \to e\gamma$ as a function of $M_{1/2}$ in SUSY $SO(10)$. Two colored plots show two extremal case of the mixing of neutrino Yukawa. Green ones show the case of maximal-mixing scenario where the mixing is given by U_{PMNS} with $|U_{e3}| = 0.07$, and red ones show the case of minimal-mixing scenario by U_{CKM}. Scatter plots are obtained by scanning SUSY parameter space in the LHC accessible region.

to those of the up quarks. Another interesting point is that the matter parity $M = (-1)^{3(B-L)}$ (equivalent to the R-parity) is a gauge transformation in SUSY $SO(10)$, resulting in the lightest SUSY partner (LSP) can be a natural candidate for the dark matter. Also in the SUSY $SO(10)$, both $(m^2_{\tilde{R}})_{ij}$ and $(m^2_{\tilde{L}})_{ij}$ can contribute to the source of LFV. The branching ratio of $\mu \to e\gamma$ calculated in [16] is shown in Figure 2.6 as a function of unified gaugino mass $M_{1/2}$. In these plots, two extremal case of the mixing of neutrino Yukawa are shown as benchmark cases. The green ones show the case of maximal-mixing scenario where the mixing is given by U_{PMNS}. In this case, the LFV is dominated by the contribution from neutrino Yukawa interaction. Since $R = 1$, the rate depends on the unknown neutrino mixing U_{e3}. The red ones show the case of minimal-mixing scenario by U_{CKM}. The rate is suppressed by the small mixing of the CKM angles. The existence of the large top Yukawa coupling would work to compensate such a suppression, but still about two order smaller than that in the U_{PMNS}-case. In this case, the contribution from the right-handed sleptons through $\delta^L_{\mu\tau}\delta^R_{\tau e}$ becomes sub-leading enhanced by a factor $(m_\tau/m_\mu)^2$.

2.2.5 Summary

So far we saw some examples of SUSY extensions, but in general, cLFV could be induced in all extensions to the SM (see, for example, [18, 19] as reviews). Searches for the LFV processes are very sensitive to high-energy physics beyond the SM such as SUSY and its breaking mechanism, origin of the neutrino masses, and grand unification. Pushing down the sensitivity of $\mu \to e\gamma$ search to 10^{-13} level can cover large part of the interesting region. Undoubtedly through the search, we will get a large amount of information for the new physics. A search for $\mu \to e\gamma$ decay with better sensitivity than current limit always has some possibility of discovering.

2.3 Experimental Search

2.3.1 History of $\mu^+ \to e^+\gamma$ Search

The muon was discovered in 1937 by Neddermeyer and Anderson in cosmic rays [20]. Since then for over a decade, the muon was thought to be the quantum mediating the strong nuclear force, as predicted by Yukawa [21]. In 1947, however, an experiment [22] clearly showed that the muon does not interact via the strong interaction, and thus the muon cannot be the π meson of Yukawa. Subsequent high-precision tests of the electromagnetic and weak couplings of muons and electrons have found no significant differences between these two particles except for their masses. It was believed that the muon decays into an electron and a neutral particle. In particular, the muon was thought to decay into a electron and a gamma ray if it is simply a heavy electron.

The first search for $\mu^+ \to e^+\gamma$ was made in 1947 using cosmic-ray muons [23]. Its negative result set an upper limit on the branching ratio ($\mathcal{B}$) of less than 10 %. This was the beginning of the search for LFV. In 1948, the continuous spectrum of electrons from muon decays was established [24]. This suggested a three-body decay with a final state of an electron accompanied by two neutral particles. The searches were significantly improved when muons became artificially produced at accelerators. In 1955, an upper limit of $\mathcal{B} < 2 \times 10^{-5}$ was set using the Nevis cyclotron at the Columbia University [25].

In the late 1950s, it was suggested that the weak interactions all arise from the coupling of a vector current with a heavy charged boson. One of the consequences of the existence of such an intermediate boson was the occurrence of $\mu^+ \to e^+\gamma$ decay with a branching ratio of order of 10^{-4} [26], because at that time there was thought to be only one type of neutrino. The puzzle of the absence of neutrino-less muon decays was phenomenologically solved by incorporating the separate electron-number (L_e) and muon-number (L_μ) conservation into the $V - A$ theory of weak interactions. This conservation law requires separate type of neutrinos for the muon and electron. The two-neutrino hypothesis was verified experimentally in 1962 at the Brookhaven National Laboratory (BNL) by observing neutrinos emitted in pion decay do not produce inverse-β decay [27]. With this discovery, interest in the search for neutrino-less decay modes of the muon waned and experiments essentially ended for about 15 years.

Now our understanding of modern elementary-particle physics is based on the SM. In the minimal version of the SM, where only an Higgs doublet is included and mass-less neutrinos are assumed, lepton-number conservation is an automatic consequence of gauge invariance and the renormalizability of the SM Lagrangian. It is the basis of a natural explanation for the smallness of lepton-flavor violation in charged-lepton processes.

In 1977, rumors circulated that an experiment at the Swiss Institute for Nuclear Research (SIN, currently PSI) had found a signal of the decay $\mu^+ \to e^+\gamma$. This underscored the fact that conservation of muon number is only empirical and without a fundamental basis. The rumors were later refuted, but motivated many theoretical and experimental activities. On the experimental side, besides the search at SIN [28], two other experiments were quickly assembled to search with improved sensitivity. One was located at the Tri-University Meson Facility (TRIUMF) [29], and the other was at the Los Alamos Meson Physics Facility (LAMPF) [30]. Results from these experiments showed no evidence for the decay at a level of 1.7×10^{-10}.

Now the interest of the search moved to the physics beyond the SM. The motivations of the search have been provoked by the various theoretical models which predict large values of the branching ratio reachable with current experimental techniques, as already discussed in the previous section.

In Figure 2.7, the historical progress of the upper limit on the branching ratio of $\mu^+ \rightarrow e^+\gamma$ decay is shown. The sensitivities have been improved primarily as the number of observed muons increases. Innovations in muon source have enabled breakthrough. The usage of accelerator achieved the first big progress in the middle of 1950s. Until 1964, muons were obtained from the π^+ beam stopped in a target. Next series of experiments from the middle of 1970s to present have used high-intensity μ^+ beam from meson factories. In particular in 1978, the experiment at LAMPF first used the surface muon [31] which was first exploited at the Lawrence Berkeley Laboratory 184" cyclotron in 1971. Owing to this technique, high-intensity μ^+ beam can be stopped in a very thin target. As the muon rate increases, the rate of background also increases. Therefore, $\mu^+ \rightarrow e^+\gamma$ search experiments have gradually improved detector resolutions. Table 2.3 lists the $\mu^+ \rightarrow e^+\gamma$ search experiments in the era of meson factories and their results with some experimental parameters. The MEGA experiment at LAMP set the current best limit of $\mathcal{B} < 1.2 \times 10^{-11}$ in 1999 [1]. A new experiment, MEG, has just started physics data taking since 2008 aiming at the sensitivity of two orders of magnitude below the current limits.

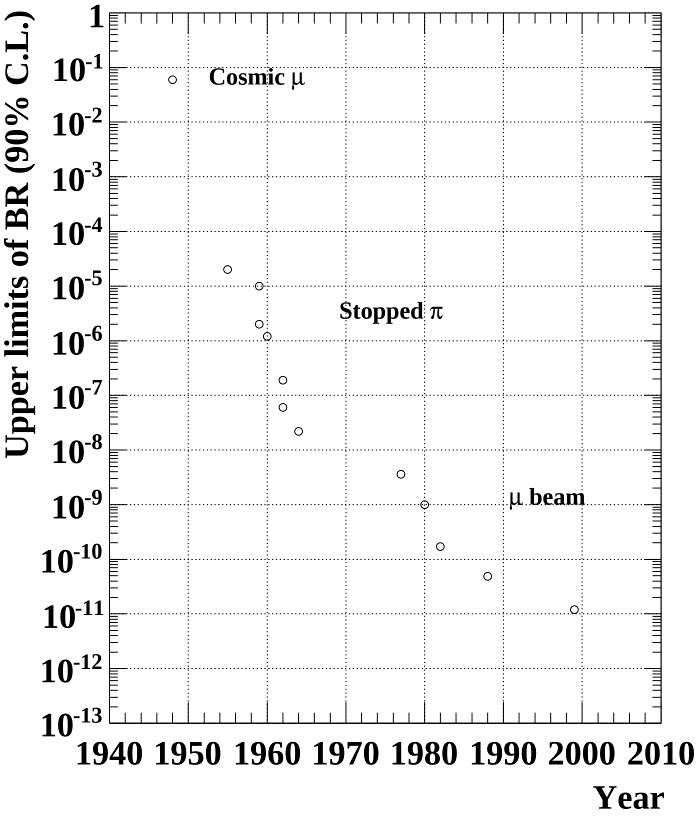


Figure 2.7: Historical progress of $\mu^+ \rightarrow e^+\gamma$ search.

Table 2.3: Historical progress of search for $\mu^+ \rightarrow \mathrm{e}^+\gamma$ since the era of meson factories with 90 % C.L. upper limits. The beam rates quoted are given as average beam intensities. The resolutions are given as full width at half maximum (FWHM).

Year	Site	Beam rate	ΔE_e	ΔE_γ	$\Delta t_{e\gamma}$	$\Delta\Theta_{e\gamma}$	Upper Limit	Ref.
1977	SIN	5×10^5 μ^+/s	10 %	8.7 %	6.7 ns		1.0×10^{-9}	[28]
1977	TRIUMF	2×10^5 π^+/s	8.7 %	9.3 %	1.4 ns		3.6×10^{-9}	[29]
1979	LAMPF	2.4×10^6 μ^+/s	8.8 %	8 %	1.9 ns	37 mrad	1.7×10^{-10}	[30]
1986	LAMPF	4×10^5 μ^+/s	8 %	8 %	1.8 ns	87 mrad	4.9×10^{-11}	[9]
1999	LAMPF	1.3×10^7 μ^+/s	1.2 %	4.5 %	1.6 ns	15 mrad	1.2×10^{-11}	[1]

2.3.2 Requirements of Muons

To improve the sensitivity, we have to measure a huge number of muons. Let us think of observing 3×10^{13} muons to achieve a sensitivity of 10^{-13} in a realistic period of data taking, let's say, 2 years ($\sim 6 \times 10^7$ sec). Taking a reasonable detection efficiency (~ 4 %) and efficiency of accelerator operation period (~ 50 %) into account, we need a muon intensity of $\sim 3 \times 10^7 \mu^+$/sec.

Negative muons (μ^-) form muonic atoms when they are captured by nuclei in material. It is not good for the $\mu \rightarrow \mathrm{e}\gamma$ search since bound states are formed and recoils of nuclei make the event signature complex. Therefore, experimentally positive muons (μ^+) have been used.

2.3.3 Event Signature

The $\mu^+ \rightarrow \mathrm{e}^+\gamma$ decay is a simple two-body decay from a muon at rest as shown in Figure 2.8(a), and the event signature has the following features:

- The positron has a monochromatic energy of 52.8 MeV, which is a half of the muon mass.
- The gamma ray has a monochromatic energy of 52.8 MeV.
- They are in time coincidence.
- They are emitted back-to-back from the same vertex.

Therefore, the following four variables are generally used in distinguishing signal events from a large amount of background events: the positron energy (E_e), the gamma-ray energy (E_γ), the relative timing between the positron and gamma ($t_{e\gamma}$), and the opening angle between them ($\Theta_{e\gamma}$).

2.3.4 Backgrounds

There are two kinds of major backgrounds mimicking the signature of $\mu^+ \rightarrow \mathrm{e}^+\gamma$ decay discussed above. One is a physics background from a radiative muon decay, $\mu^+ \rightarrow \mathrm{e}^+\nu_\mathrm{e}\overline{\nu}_\mu\gamma$. It becomes a prompt background when the two neutrinos carry off little energy as shown

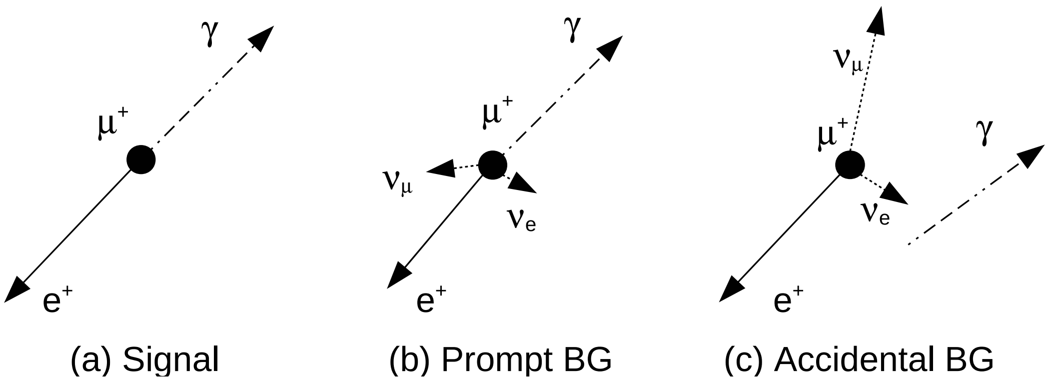


Figure 2.8: Schematics of $\mu^+ \to e^+\gamma$ event signature (a), and two types of backgrounds (b)(c).

in Figure 2.8(b). The other is an accidental overlap of an uncorrelated Michel positron and a high-energy gamma from any kinds of sources. We call this type of background accidental background (Figure 2.8(c)). The rates of backgrounds crucially depend on the detector performance. We discuss the details of those backgrounds in the following.

Prompt Background

Approximately 1 % of muon decays are accompanied by a radiation of photon $E_\gamma >$ 10 MeV. These radiative muon decay (RD) events can mimic a $\mu^+ \to e^+\gamma$ signal when the two neutrinos carry off little energy.

The differential branching ratio was calculated by several authors [32, 33]. Within the $V-A$ interaction, the differential branching ratio, where the final position and gamma are emitted at energy intervals of dx and dy with solid angles of $d\Omega_e$ and $d\Omega_\gamma$, respectively, in the muon rest frame can be write down as,

$$\begin{aligned} d\mathcal{B}_{RD}(\mu^+ \to e^+\nu\bar{\nu}\gamma) &= \frac{\alpha}{64\pi^3}\beta dx\frac{dy}{y}d\Omega_e d\Omega_\gamma[F(x,y,d) \\ &\quad -\beta\vec{P}_\mu \cdot \hat{p}_e G(x,y,d) \\ &\quad -\vec{P}_\mu \cdot \hat{p}_\gamma H(x,y,d)], \end{aligned} \tag{2.17}$$

where $\vec{P}_\mu$ is the muon polarization vector; $\hat{p}_e$ and $\hat{p}_\gamma$ are the unit vectors of momenta of the positron and the gamma, respectively; β is defined as $\beta \equiv |\vec{p}_e|/E_e$; d is given by $d \equiv 1 - \beta\hat{p}_e \cdot \hat{p}_\gamma$; and x and y are normalized positron and gamma energies, $x = 2E_e/m_\mu$ and $y = 2E_\gamma/m_\mu$. The formulas of $F(x,y,d)$, $G(x,y,d)$, and $H(x,y,d)$ in the SM are given in Appendix A.4. In addition, we define here z as $z \equiv \pi - \Theta_{e\gamma}$. In this notation, the RD looks similar to the $\mu^+ \to e^+\gamma$ decay when $x \approx 1$, $y \approx 1$, and $z \approx 0$. Let us consider here only unpolarized muons ($P_\mu = 0$). Then it becomes,

$$d\mathcal{B}_{RD}(\mu^+ \to e^+\nu\bar{\nu}\gamma) = \frac{\alpha}{64\pi^3}\beta dx\frac{dy}{y}d\Omega_e d\Omega_\gamma F(x,y,d)$$

$$= \frac{\alpha}{8\pi}\frac{\beta}{y}F(x,y,d)dxdyd(\cos\Theta_{e\gamma}). \tag{2.18}$$

When exactly $x = 1$ and $y = 1$, the differential decay width vanishes. In reality, the finite detector resolutions introduce background events. Given the detector resolutions, the sensitivity limitation from RD can be evaluated by integrating the differential branching ratio over the resolutions, more precisely, over the kinematic-box region of the signal which is determined by the resolutions. Let us take δx, δy, and δz to be the kinematic range of the signal region for positron energy ($1-\delta x \leq x \leq 1$), that for gamma energy ($1-\delta y \leq y \leq 1$), and that for the opening angle ($0 \leq z \leq \delta z$), respectively. The partial branching ratio after the integration is given by [34],

$$\begin{aligned} d\mathcal{B}_{RD}(\mu^+ \to \mathrm{e}^+\nu\bar{\nu}\gamma) &= \int_{1-\delta x}^{1} dx \int_{1-\delta y}^{1} dy \int_{0}^{min[\delta z, 2\sqrt{(1-x)(1-y)}]} dz \frac{dB(\mu^+ \to \mathrm{e}^+\nu\bar{\nu}\gamma)}{dxdydz} \\ &= \frac{\alpha}{8\pi}[J_1 + J_2], \end{aligned} \tag{2.19}$$

where J_1 and J_2 are given by,

$$\begin{aligned} J_1 &= \frac{8}{3}(\delta x)^3(\delta y)(\frac{\delta z}{2})^2 - (\delta x)^2(\frac{\delta z}{2})^4 + \frac{1}{3}\frac{1}{(\delta y)^2}(\frac{\delta z}{2})^8, \\ J_2 &= 8(\delta x)^2(\delta y)^2(\frac{\delta z}{2})^2 - 8(\delta x)(\delta y)(\frac{\delta z}{2})^4 + \frac{8}{3}(\frac{\delta z}{2})^6, \end{aligned} \tag{2.20}$$

when $\delta z < 2\sqrt{\delta x \delta y}$, which means an angular resolution better than the kinematic constraint of $2\sqrt{\delta x \delta y}$, and this assumption fits into our case.

To take some idea, let us take an example of MEGA detector resolutions. To keep 90 % of the signal coverage, we set here the width of the signal box to be 1.4 times FWHM. From the FWHM resolutions in Table 2.3, the integration ranges become,

$$\begin{aligned} \delta x &= 0.0084, \\ \delta y &= 0.023, \\ \delta z &= 0.0075, \\ \delta t_{e\gamma} &= 1.12 \text{ ns}^1. \end{aligned} \tag{2.21}$$

Put these number in the formula, then the effective branching ratio of the prompt background is given as,

$$\mathcal{B}_{RD}(\mu^+ \to \mathrm{e}^+\nu\bar{\nu}\gamma) \sim 4.4 \times 10^{-15}. \tag{2.22}$$

The prompt background is not found to be a serious problem as long as such detector resolutions are achieved.

Accidental Background

In a high intensity environment, an accidental overlap of uncorrelated positrons and gamma rays can also mimic the $\mu^+ \to \mathrm{e}^+\gamma$ signal with some probability. The higher the muon intensity becomes, the more serious the accidental background becomes.

---

[1] Time resolution does not work for discrimination of the prompt background, but this number is used for accidental background later.

The effective branching ratio of the accidental background can be estimated by

$$\mathcal{B}_{acc} = R_\mu \cdot f_e^0 \cdot f_\gamma^0 \cdot (\frac{\delta\omega_{e\gamma}}{4\pi}) \cdot (2\delta t_{e\gamma}), \tag{2.23}$$

where R_μ is an instant muon intensity; $\delta t_{e\gamma}$ is a half width of the signal region for time coincidence; $\delta\omega_{e\gamma}$ is that for back-to-back constraint. The terms f_e^0 and f_γ^0 are the fractions of the spectrum within the signal box of positron in the Michel decay and that of gamma from such as RD, respectively.

The positron energy spectrum of $\mu^+ \to e^+\nu_e\bar{\nu}_\mu$ decay (Michel spectrum) is shown in Figure 2.9. f_e^0 can be estimated by integrating the Michel spectrum over $1-\delta x \leq x \leq 1$, yielding $f_e^0 \approx 2(\delta x)$ since it is almost flat at $x \approx 1$.

To estimate f_γ^0, the differential branching ratio of RD (Eq.2.18) is integrated over positron energy (x) and the angle between positron and gamma ($\Theta_{e\gamma}$). By neglecting the terms suppressed by m_e/m_μ, the differential branching ratio is calculated to be [35]

$$d\mathcal{B}_{RD}(\mu^+ \to \mathrm{e}^+\nu\bar{\nu}\gamma) \;=\; \frac{2}{y}[J_+(y) + J_-(y)]dy, \tag{2.24}$$

where $J_+(y)$ and $J_-(y)$ are defined by

$$\begin{aligned} J_+(y) &= \frac{\alpha}{6\pi}(1-y)\Big[\Big(3\ln\frac{1-y}{r} - \frac{17}{2}\Big) \\ &\quad + \Big(-3\ln\frac{1-y}{r} + 7\Big)(1-y) \\ &\quad + \Big(2\ln\frac{1-y}{r} - \frac{13}{3}\Big)(1-y)^2\Big], \end{aligned} \tag{2.25}$$

$$\begin{aligned} J_-(y) &= \frac{\alpha}{6\pi}(1-y)^2\Big[\Big(3\ln\frac{1-y}{r} - \frac{93}{12}\Big) \\ &\quad + \Big(-4\ln\frac{1-y}{r} + \frac{29}{3}\Big)(1-y) \\ &\quad + \Big(2\ln\frac{1-y}{r} - \frac{55}{12}\Big)(1-y)^2\Big], \end{aligned} \tag{2.26}$$

where $r = (m_e/m_\mu)^2$. It is shown in Figure 2.10. The partial branching ratio integrated over the signal region ($1-\delta y \leq y \leq 1-r$) can be calculated with neglecting the higher order terms of $(1-y)$,

$$\begin{aligned} f_\gamma^0 &= \int_{1-\delta y}^{1-r} dy \frac{d\mathcal{B}_{RD}(\mu^+ \to e^+\nu\bar{\nu}\gamma)}{dy} \\ &\approx \frac{\alpha}{2\pi}(\delta y)^2[\ln(\delta y) + 7.33]. \end{aligned} \tag{2.27}$$

In this estimation, the other sources of high-energy gamma such as positron annihilation in flight (AIF), external bremsstrahlung are ignored. Those contributions depend on the actual design of the experiment. If we achieved extremely a good energy resolution, the AIF contribution becomes more important. In addition, accidental pileups of those gamma rays can be another source of background in high-energy region.

Given the angle resolution of δz, the size of signal box for back-to-back condition is given by $\delta\omega_{e\gamma} = \pi(\delta z)^2$.

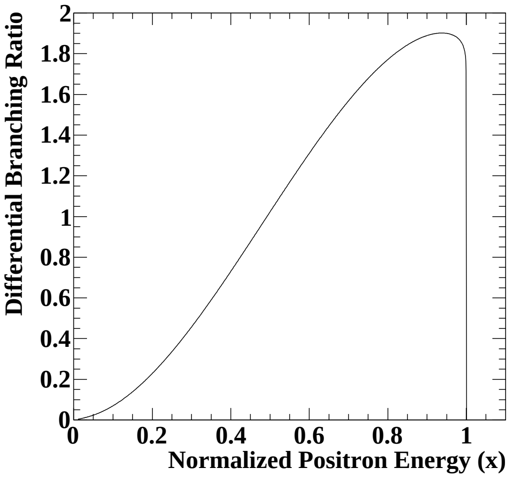


Figure 2.9: Positron energy spectrum of unpolarized $\mu^+ \rightarrow e^+\nu_e\bar{\nu}_\mu$ decay (Michel spectrum). A radiative correction due to the virtual photon emission and the inner bremsstrahlung is applied in the spectrum [36].

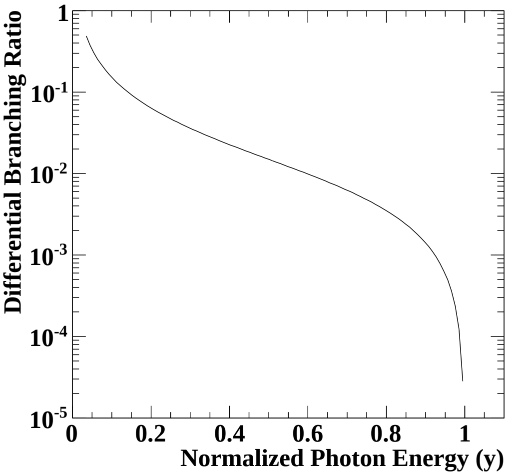


Figure 2.10: Photon energy spectrum of unpolarized $\mu^+ \rightarrow e^+\nu\bar{\nu}\gamma$ decay. This is obtained by integrating over the positron energy and the angle between a positron and a photon.

From the above, the effective branching ratio of accidental background is approximately given by

$$\mathcal{B}_{acc} \approx R_\mu \cdot (2\delta x) \cdot \left[\frac{\alpha}{2\pi}(\delta y)^2(\ln(\delta y) + 7.33)\right] \cdot \frac{(\delta z)^2}{4} \cdot (2\delta t_{e\gamma}) \tag{2.28}$$

Again, we here calculate an example of the effective branching ratio of the accidental background using numbers in Eq.2.21. The instant beam intensity was 2.6×10^8 in MEGA. It is higher than the average intensity listed in Table 2.3 because they used a pulsed beam with duty cycle 6 %. The effective branching ratio is then given as

$$\mathcal{B}_{acc} \sim 1.2 \times 10^{-12}. \tag{2.29}$$

This could be a serious problem. A new idea to suppress the background is necessary to go into the sensitivity of 10^{-13} level.

2.3.5 Requirements of $\mu^+ \rightarrow e^+\gamma$ Search

By the naive calculations of backgrounds above, the accidental background is found to be the dominant background source, and it will limit the experiment.

First, from Eq.2.23 we see the effective branching ratio of the accidental background is proportional to the instant muon beam intensity. Whereas we estimated that we need a $> 10^7$/sec muon intensity to get enough statistics. To achieve such an intensity with minimizing the background, using a direct current (DC) muon beam is the best solution. If a DC beam was used in MEGA, the background rate would be reduced by a factor 16 with same statistics.

We can see in Figure 2.10 that the background source of gamma ray is strongly suppressed at the signal region. It is also the case for gamma rays from AIF. With a good gamma-ray energy resolution, we can reduce background rate most efficiently. It is also important to reduce material which interacts with positrons to suppress additional gamma-ray yield.

In contrast, there are abundant positrons in the signal region. It is difficult to reduce background effectively by improving the positron energy measurement. It is more important to efficiently measure a huge amount of positrons because every muon of such a high intensity generates a positron.

Let us consider MEGA a little more. Briefly, they measured gamma rays using pair spectrometers, which consist of lead conversion foils, a multi-wire proportional chamber (MWPC), drift chambers, and plastic scintillators. They measured the $e^- + e^+$ generated from the gamma-ray conversion. This method gave a good energy and position resolution and also some information of direction. It, however, limited the gamma-ray detection efficiency and the timing resolution. It was a trade off between energy resolution and detection efficiency by the thickness of the conversion foils. If we develop a new type of gamma-ray detector with a high efficiency and a high time resolution, it becomes a great advantage of background suppression.

We can summarize the requirements of $\mu^+ \rightarrow e^+\gamma$ search in the following:

- high intensity DC μ^+ beam,
- high rate tolerable positron detector,
- high performance gamma-ray detector.

Chapter 3

Experimental Apparatus

This chapter gives detail descriptions of the experimental apparatus of MEG. In Figure 3.1, an overview of the experiment is shown. To summarize, the key elements of MEG are:

- The world's most intense DC muon beam,
- An innovative positron spectrometer,
- A new type of gamma-ray detector with liquid xenon.

With this unique apparatus, we cope with the unexplored region of the sensitivity.

All these components were constructed and commissioned by 2007 and an engineering run was conducted in 2007. We started full-scale physics data taking in 2008.

Coordinate System

Before entering the detail descriptions of the experiment, here we define the global coordinate system of MEG. It is used throughout this thesis. The origin is defined as the center of the positron spectrometer; it is also the center of the muon stopping target. The z-axis is assigned to the beam direction. The y-axis is vertical axis from bottom to top. Then the x-axis is defined as the other axis of right-handed rectangular coordinate system. The θ is defined as the polar angle from the z-axis, while the ϕ is the azimuthal one. In this coordinate system, the gamma-ray detector is located at the negative x side and trajectories of positrons are negative ϕ direction as shown in Figure 3.1.

In addition, we define a local coordinate system of the gamma-ray detector. It is used in the reconstruction of gamma rays. We define (u, v, w) coordinates. The u-axis is identical to the z-axis. Its center is at the center of the stopping target which is also the center of the gamma-ray detector. The v-axis is defined as a direction of negative ϕ along the surface of the inner face of the gamma-ray detector. Its center is at $y = 0$. The w-axis is defined as the depth in the detector. The direction is the same as the radial direction in cylindrical coordinates but the origin of the axis is at the surface of the inner face. The detailed design of the gamma-ray detector is described in Sec.3.2.2.

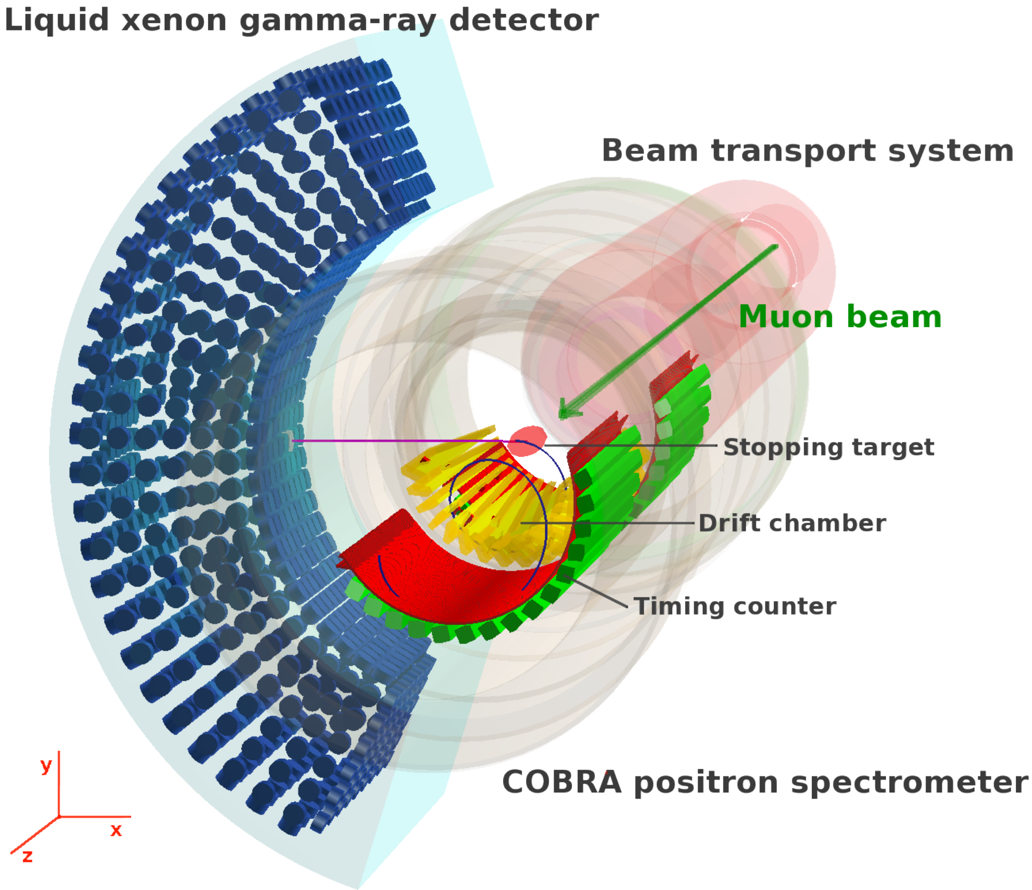


Figure 3.1: Overview of the MEG experiment.

3.1 μ^+ Beam

The sensitivity of the experiment improves primarily as the number of observed muons. Therefore, a high intensity muon beam is required. On the other hand, the accidental background rate increases as the instant intensity of the beam. Hence, a DC beam is preferable than a pulsed beam for $\mu^+ \to e^+\gamma$ search experiments.

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

3.1.1 PSI Proton Ring Cyclotron

PSI operates a 590 MeV proton ring cyclotron (Figure 3.2) with the maximum proton current of 2 mA and beam power of 1.2 MW. It is operated at an accelerator frequency of 51 MHz. Thus the primary proton beam has the RF pulse structure of 20 ns period. However, compared to the muon lifetime of 2.2 μsec, it is short enough and the muon decay rate is almost constant. It can offer $> 1.5 \times 10^8$/sec DC muon beam.

Furthermore, this cyclotron is currently being upgraded: its beam current is planed to reach 2.6 mA in a few years, and 3.0 mA some years thereafter [37]. Eventually, we will be able to get $> 2 \times 10^8 \mu^+$/sec.

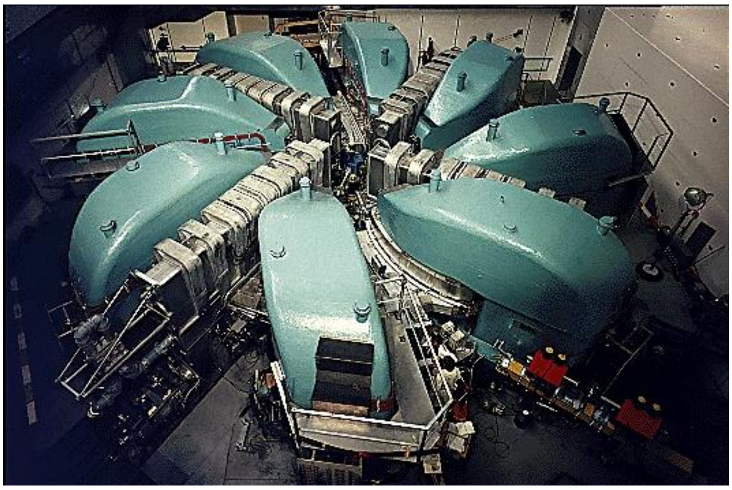


Figure 3.2: 590 MeV proton ring cyclotron at PSI.

3.1.2 Surface Muon from πE5 Channel

The πE5 channel is one of the secondary beamlines which provide low-energy (10-120 MeV/c) pion and muon beam. We use a surface muon beam from this channel. The surface muons [31] are muons originating from pions stopping near the surface of the pion production target and decaying at rest. Before the surface muon method was developed, muon beams were generally produced by transporting a pion beam over a distance[1]. Only a small fraction of the decay muons were accepted for the beamline. To get higher intensity, a moderately high momentum ($\sim$ 100 MeV/c) was usually used and the momentum spread was large, which forced one to use a thick stopping target. In contrast, muons from the decay of π^+ at rest have an unique momentum of 29.8 MeV/c. During penetrating a thin layer of the surface of the target, it looses some energy, but still have low momentum spread. By tuning the beamline to accept $\sim$28 MeV/c, we can collect high intensity of positive muons.

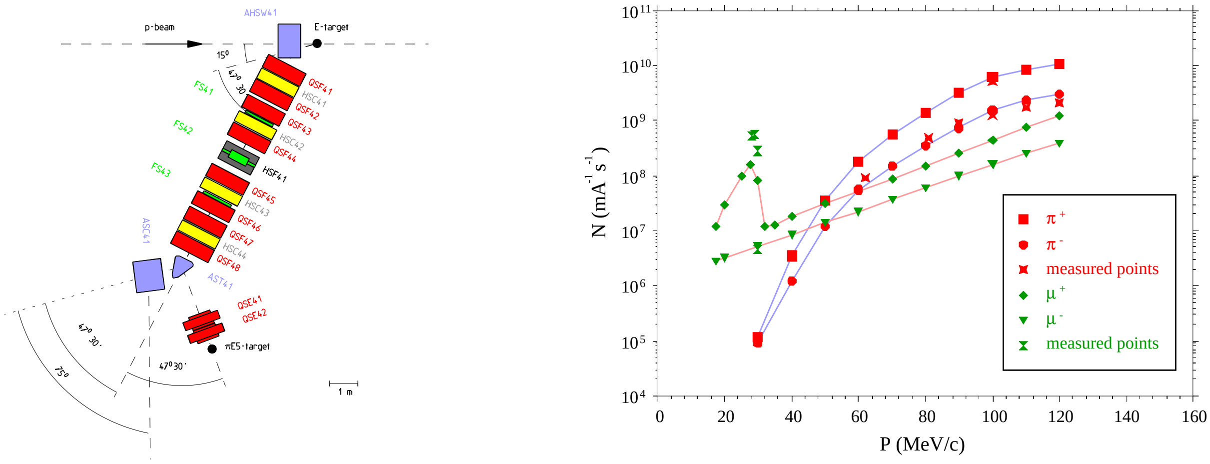
In the πE5 channel, a thick (4 cm) graphite π/μ production target (E-target) is placed in the primary proton beamline. The surface muons are extracted to the channel at an angle of 175° with respect to the primary proton beam. Figure 3.3(a) shows the πE5 beamline consisting of bending magnets, quadrupoles, hexapoles, and slits. Three sets of horizontal and a set of vertical slits define the momentum and acceptance. Figure 3.3(b) shows the expected and measured rates of muons and pions at the final focus of πE5 just behind the second bending magnet (AST). An enhancement due to surface muons is seen at momentum around 28 MeV/c. Because of the low momentum and the sharp momentum distribution, it is easy to degrade muon momentum and possible to stop them in a very thin target with a high efficiency.

3.1.3 Beam Transport System

The muon beam from the channel is transported to a stopping target in our detector system through several elements of the beam transport system to achieve high stopping

---

[1] It is still the case of negative muon beam because π^- stopped in a target are captured by nuclei and the surface muon method cannot be applied.



(a) (b)

Figure 3.3: (a) πE5 beamline. (b) Muon and pion flux at πE5 channel.

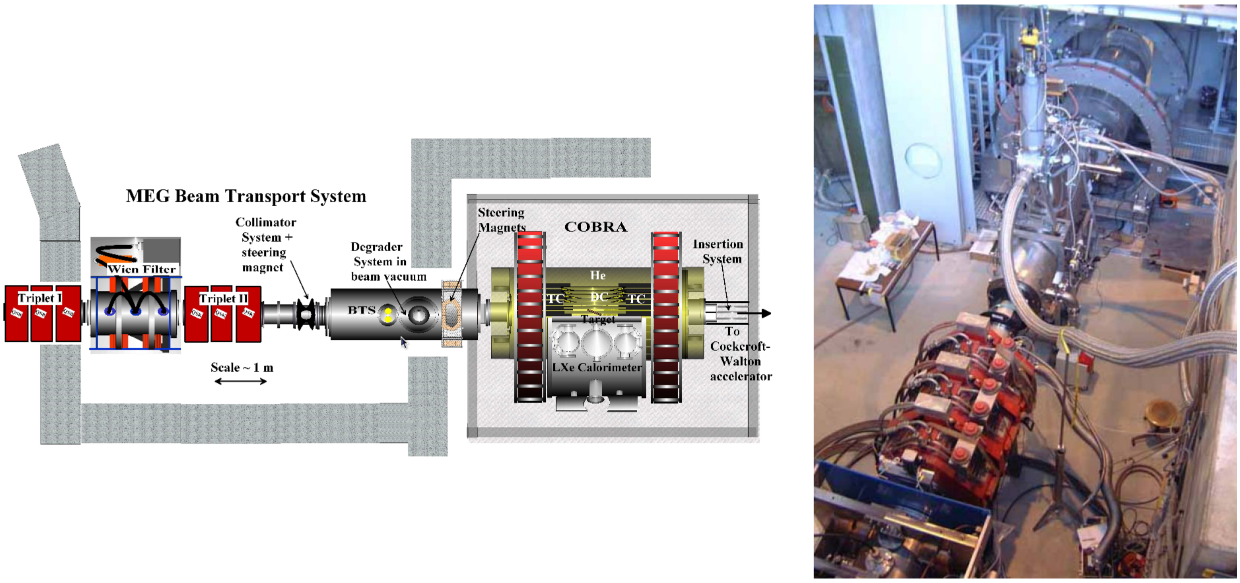
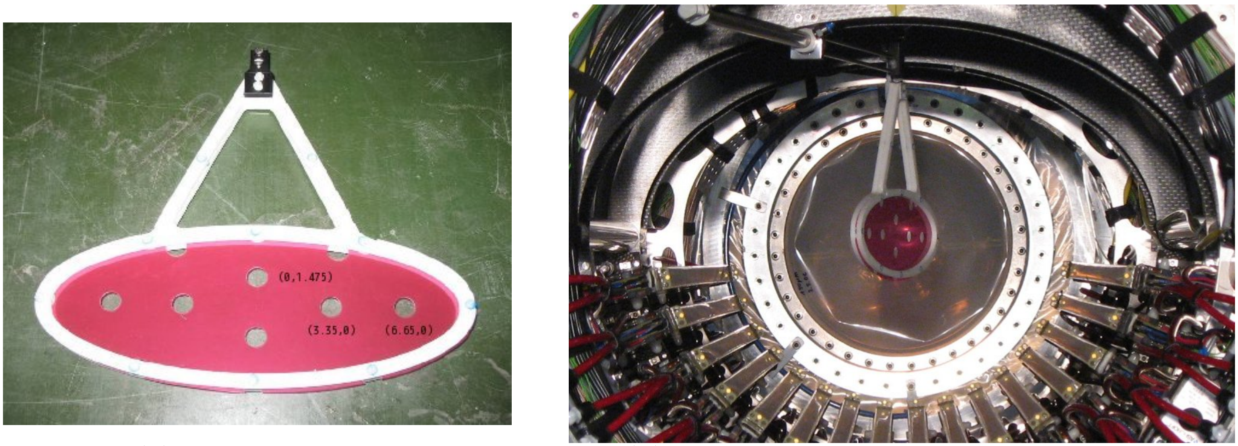


Figure 3.4: Schematic and picture of beam transport system. Surface muons enter from the left and go through the Wien filter, the BTS, and the COBRA magnet to reach the stopping target at the center of the magnet.

rate on a thin target with minimum beam-related background. The system consists of a quadrupole triplet (Triplet I), a crossed-field separator (Wien filter), a quadrupole triplet (Triplet II), and a beam transport solenoid (BTS) in the order of the beam flow as shown in Figure 3.4. The Wien filter cleanly separates eight times higher positron contamination by 7.5 σ by applying a horizontal magnetic field of 133 Gauss and a vertical electric field of 195 kV. The BTS adjusts the oscillation of the beam profile to minimize the spot size on the stopping target. We put a degrader made of 300 μm Mylar in the BTS to reduce muon momentum.



(a) Stopping target.

(b) Installation. View from downstream.

Figure 3.5: Pictures of MEG stopping target.

3.1.4 Stopping Target

Material of the stopping target itself can be causes of both scattering of positrons resulting in degradation of positron measurement and yielding undesired gamma rays from AIF of positrons becoming a source of accidental background. To minimize the thickness with keeping muon stopping power, we put the target with a slant angle with respect to the beam axis. We adopt a sheet of polyethylene/polyester with thickness of 205 μm (18 mg/cm^2) for the target material. Pictures of the target are shown in Figure 3.5. The sheet is supported by a Rohacell [38] frame whose density is 0.895 g/cm^3. The dimensions of ellipse are 79.8 and 200.5 mm for minor and major axises, respectively. The slant angle is optimized to 20.5°. We made six holes (10 mm diameter) on the sheet to study vertex reconstruction performance and to align the target position by using data.

3.1.5 Mode and Profile

We prepared several settings of beam modes with different intensities in 2008 run. So-called "normal" mode is nominal one for the physics data taking, and "ultra-low" mode is dedicated one for the study and calibration with RD. The others are prepared for the background studies. We directly measured the beam profile with an APD counter. The spot size at the center of the detector system is $\sigma_x = 9.5$, $\sigma_y = 10.2$ mm for the normal mode. The muon intensity of the normal mode was measured to be $(3.69 \pm 0.08) \times 10^7 \mu^+/\text{sec}$ at the center when the proton current is 2 mA. With stopping efficiency of 0.82 evaluated using MC simulation, the stopping rate of the normal mode is evaluated to be

$$R_{\mu,\text{stop}} = (3.0\ ^{-0.1}_{+0.2}) \times 10^7 \mu^+/\text{sec} \tag{3.1}$$

at 2 mA proton current. The stopping rate of the ultra-low mode is evaluated to be $1.2 \times 10^6 \mu^+/\text{sec}$.

Note that we set the beam rate lower than the maximum capability of the channel ($\sim 1.5 \times 10^8$) by the limitation from detector performance. From the beam point of view, there is a room for improvement.

3.2 Detector

An overview of the MEG detector is given in Figure 3.6. Positrons from muon decays are analyzed by the COBRA (COnstant-Bending-RAdius) spectrometer, which consists of a thin-walled superconducting solenoid magnet, a tracking system of low-mass drift chambers, and fast scintillator timing-counter arrays. Gamma rays are detected by a liquid xenon scintillation detector. It measures gamma-ray energy and first interaction point and time. We describe the concept and principle of each sub-detector in the following sections together with its actual design.

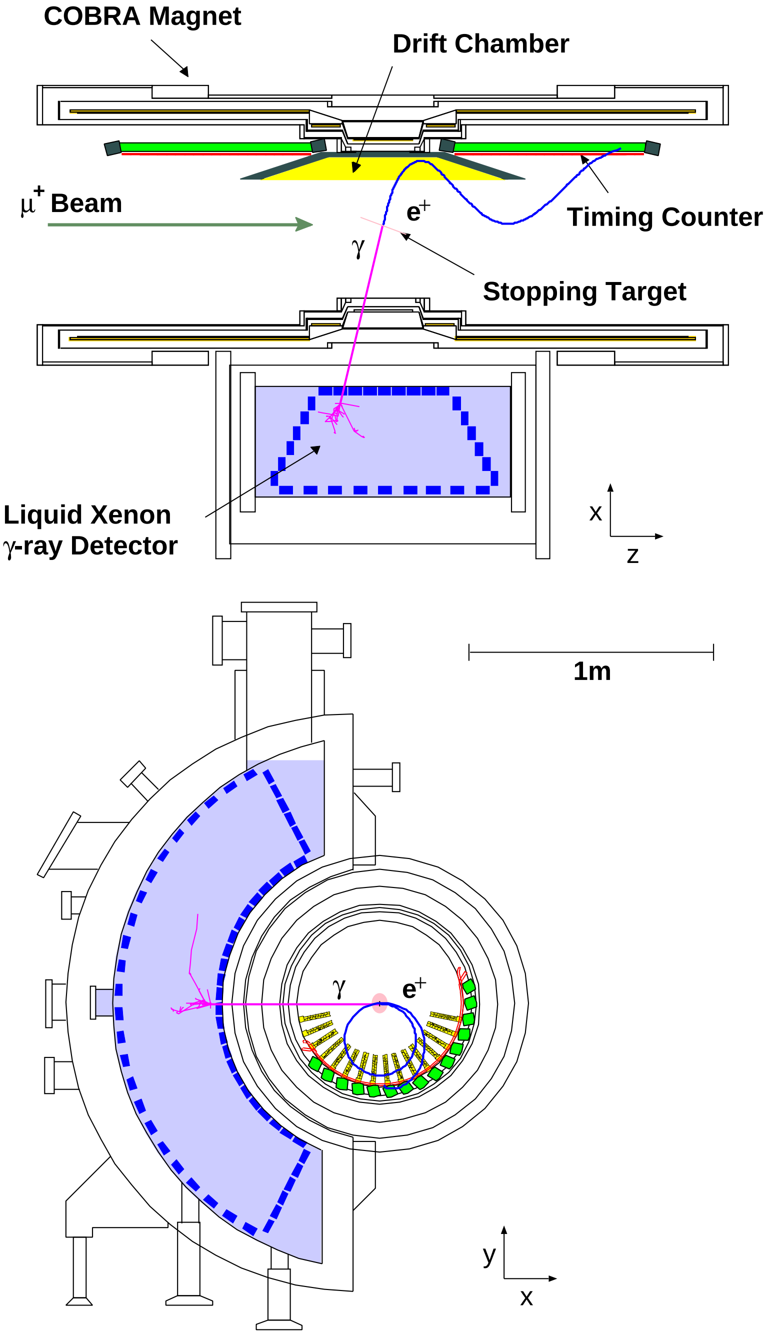


Figure 3.6: Schematic view of MEG detector.

3.2.1 The Positron Spectrometer

The requirements of the positron spectrometer are summarized in the following:

i) to have to work in very high rate environment (up to 1×10^8 s^{-1}),

ii) to be as low material as possible,

iii) to measure momentum, direction, and timing of positrons around 50 MeV with high resolutions.

We adopt a highly graded magnetic field to achieve i). The concepts of the field are schematically shown in Figure 3.7. Owing to the gradient fields, positrons with the same momentum follow trajectories with an almost constant projected bending radius independently of their emission angles. This allows a preferential acceptance of higher momentum positrons in tracking device as well as sweeping particles out of the detector more efficiently. Figure 3.8 shows the rate of positrons as a function of radius. By placing the drift chamber at a radius over 20 cm, we can reduce the hit rate to a level below the limit of stable operation.

The spectrometer is designed to be very thin to avoid undesirable gamma-ray generation from bremsstrahlung or AIF of positron, to minimize the probability that a gamma ray interacts with some material before reaching gamma-ray detector, and to suppress Coulomb multiple scattering of positron. The last point is crucial for the precise measurement of low-energy positron.

Momentum and emission angle of positron is measured by the trajectory in the drift chamber and the muon decay vertex is reconstructed by extrapolating the track back to the target plane. The track is also extrapolated down to the timing counter and gives us the length of the trajectory from the decay vertex. This track length is converted to the positron time-of-flight and, together with measurement on the timing counter, used to reconstruct the muon decay time.

COBRA Magnet

We constructed a superconducting magnet specially designed to form a highly graded magnetic field (Figure 3.9) [39]. A step-structure solenoid realizes such gradient field. It consists of five coils with three different radii: one central coil, two gradient coils and two end coils. The design and dimensions are shown in Figure 3.10. Figure 3.11 shows a profile of the magnetic field along the magnet axis. The magnetic field ranges from 1.27 at the center to 0.49 T at the edge.

As shown in Figure 3.6, gamma rays from muon decay on the target go through the magnet wall to reach the gamma-ray detector. If they interact with some material in front of the detector, it causes the signal inefficiency. Hence, the cable and wall of the magnet are designed to be very thin. The superconducting cable is made from NbTi multifilament embedded in copper matrix and high-strength aluminum stabilizer [40]. Nickel of 5000 ppm is added into the aluminum stabilizer to reinforce it mechanically. Owing to the high-strength superconductor, we can minimize the thickness of the support structure required for a given electromagnetic force acting on the coils. The total thickness of the magnet including its cryostat is suppressed to be 0.197 X_0. The transmission efficiency of gamma rays is 85 %.

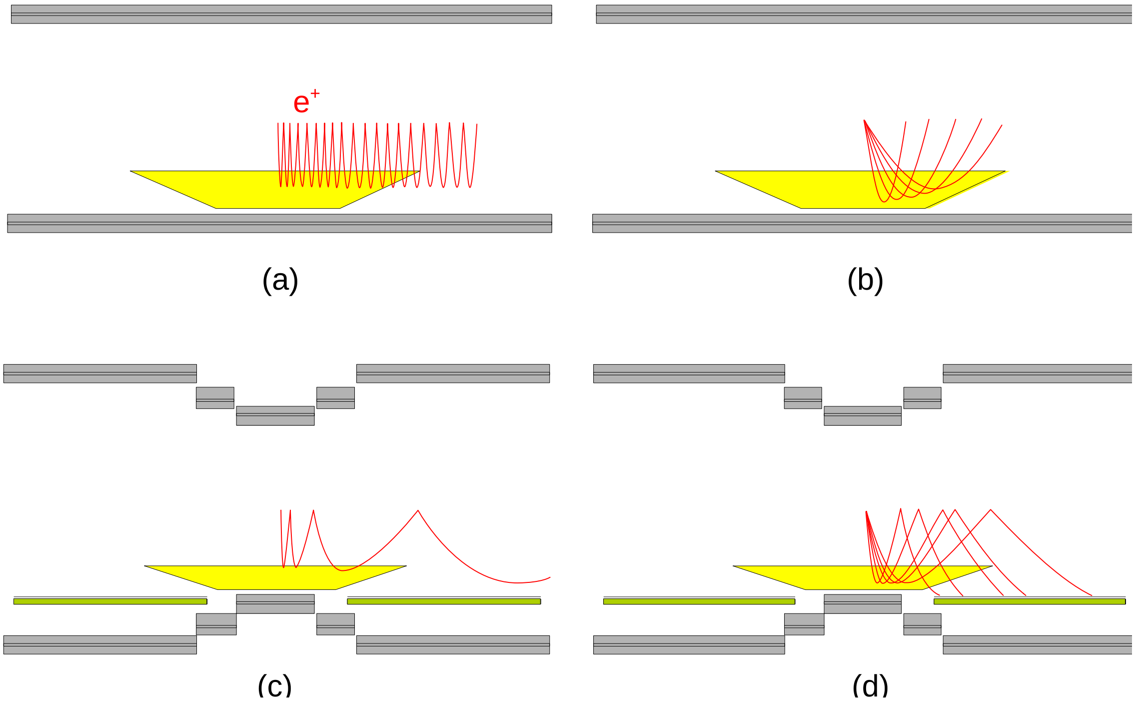


Figure 3.7: Conceptual illustrations of the COBRA spectrometer compared with one with a uniform magnetic field. (a) and (c) show trajectories of positrons emitted at 88°. The uniform field makes many turns inside the detector, whereas the gradient field sweep the positron out of the detector much more quickly. (b) and (d) show trajectories of mono-energetic positrons emitted at various angles. In the uniform field, the bending radius depends on the emission angle, whereas it is independent in the gradient field.

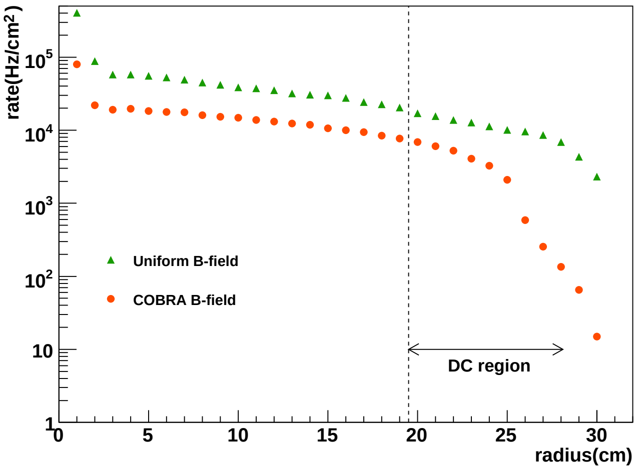


Figure 3.8: Rate of Michel positrons per cm^2 per second as a function of radius assuming muon decay rate of 3×10^7/sec.

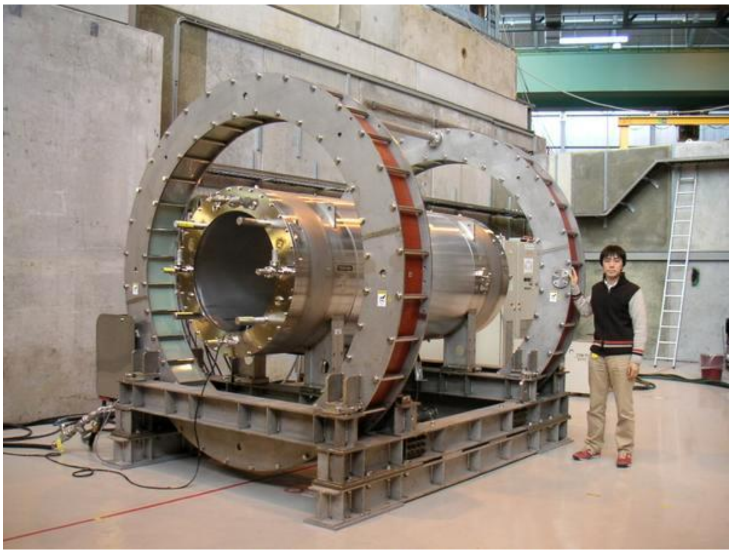


Figure 3.9: Picture of the COBRA magnet.

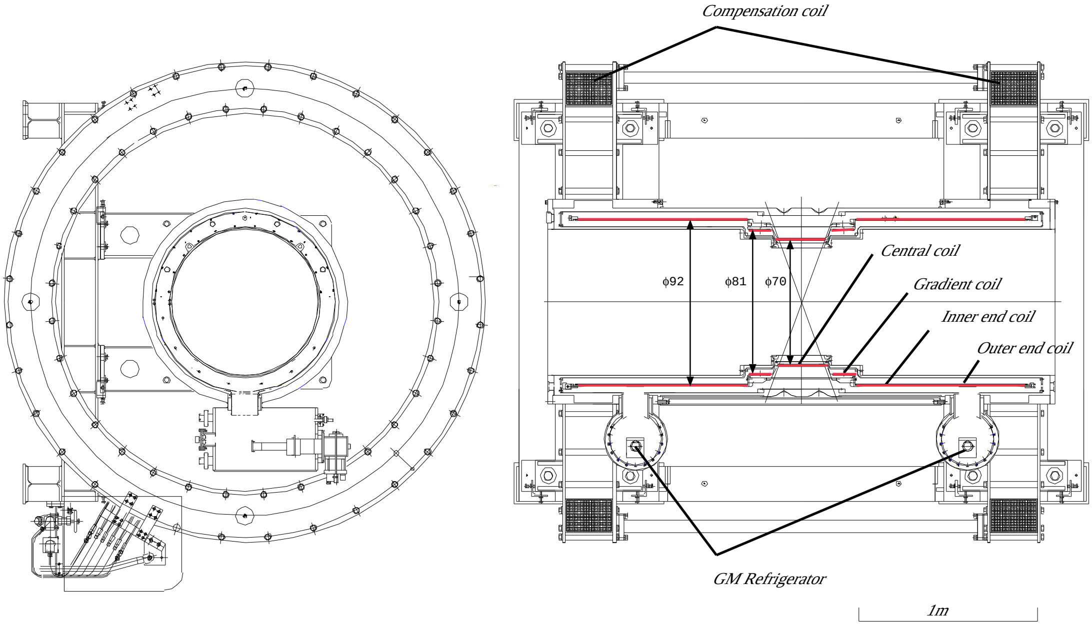


Figure 3.10: Schematic of the COBRA magnet.

The magnetic field produced by the COBRA magnet can significantly deteriorate the performance of PMTs used in the gamma-ray detector because their outputs are sensitive to magnetic field as shown in Figure 3.12. The dependence is different for different axes of the PMT because of the dynode structure. The PMTs are placed with different directions in accordance with the location in the detector. The requirement for tolerable magnetic field is estimated to be less than 50 Gauss. To reduce the fringe field, we use a pair of resistive coils placed at the both ends of the solenoid. The fringe field with the compensation coils around the gamma-ray detector is shown in Figure 3.13. Owing to the compensation coils, we successfully reduce the fringe field at the position of the gamma-ray detector.

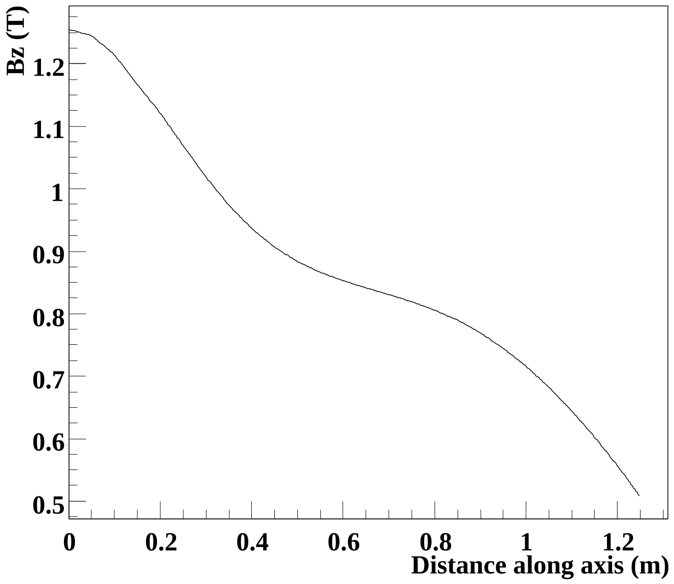


Figure 3.11: Profile of the magnetic field along the axis of the magnet.

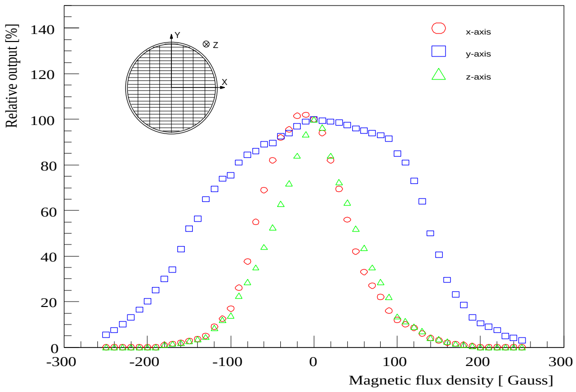


Figure 3.12: Response of PMT used in the gamma-ray detector in magnetic field.

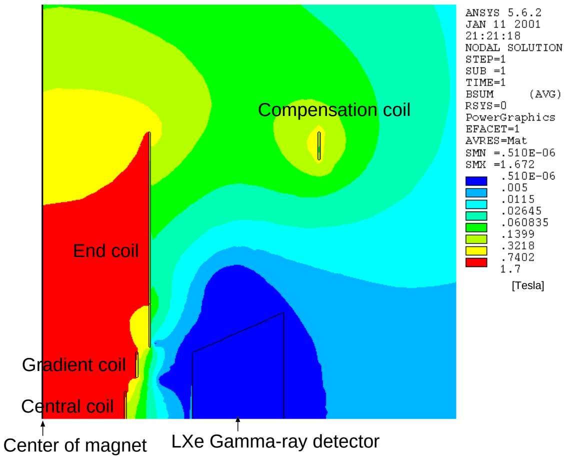


Figure 3.13: Contour plot of fringe field around the gamma-ray detector.

Drift Chamber System

The drift chamber system (DCH) measures trajectories of positrons and reconstructs their momentum, decay vertex, emission angle and time-of-flight. A picture of the whole system is shown in Figure 3.14. It consists of 16 radially aligned modules, spaced at 10.5° intervals in ϕ direction forming a half-circle around the target. The radial position ranges from 19.3 to 27.9 cm to measure only high momentum positron (>40 MeV). Design of a module is shown in Figure 3.15 with its dimensions. The inside of COBRA magnet is filled with pure helium to minimize material along positron trajectories. The drift chamber system is placed in the helium atmosphere.

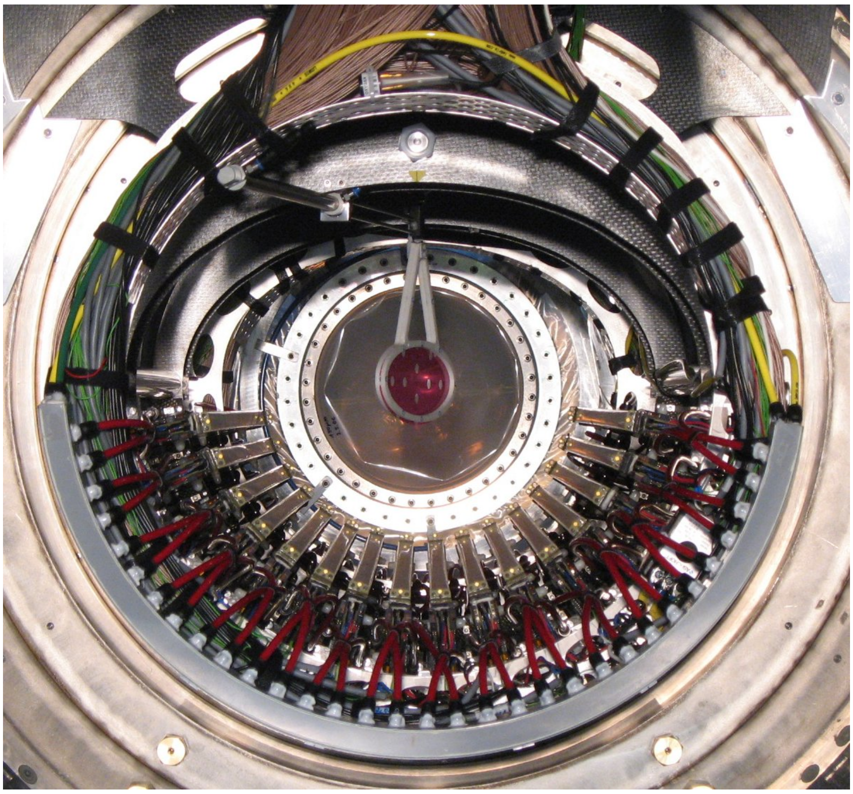


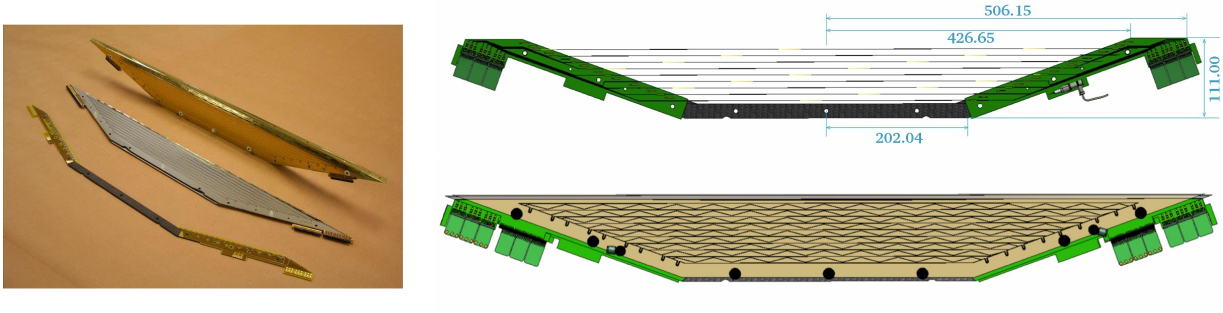
Figure 3.14: Picture of drift chamber system inside the COBRA magnet.

Each module contains two staggered layers of anode wire planes. Each layer containing nine drift cells is shifted by one-half cell each other to resolve left-right ambiguity. The two layers are separated by two inner cathode foils and also enclosed by a outer one. The cell configuration of a module is shown in Figure 3.16. The sense wires are made of Ni/Cr (80:20). Their diameter is 25 mm and resistance is 2200 Ω/m. Cathodes are made of 12.5 μm thick polyimide with aluminum deposition. They have vernier-pattern structure of 5 cm period for a precise z reconstruction. The principle of the vernier-pad method is illustrated in Figure 3.17 schematically. Owing to the vernier pattern, the number of cathode readouts is suppressed relatively low compared to that of conventional strip-pattern cathodes. It results in less amount of material of preamplifiers and cables which also take place in the region of positron tracks. All cathodes and potential wires are grounded. Positive high voltages are applied to sense wires; the nominal value is 1850 V.

The active gas of the chamber is composed of helium:ethane (50:50). We optimize it from the points of ionization loss in gas and multiple Coulomb scattering. The ionization loss for the minimum ionization particles is $\approx 65\ \mathrm{e}^-/\mathrm{cm}$. The field map and drift lines calculated by a GARFIELD simulation are shown in Figure 3.18.

A support frame made of carbon fiber supports the wires and foils. It is, however, designed to "open-frame" structure; there is no support frame at the target side (Figure 3.15). This structure helps to reduce amount of material in positron trajectories at a cost of a difficulty in the construction. Shape of each cell is formed by the thin cathode foil itself because of the open-structure frame. Hence, a precise pressure control between in (chamber gas) and out (helium atmosphere) of the chamber is necessary to maintain the cell spacing. We developed a dedicated gas-flow control system and it successfully controls the pressure better than 0.005 Pa stability while the required stability is less than 1 Pa.

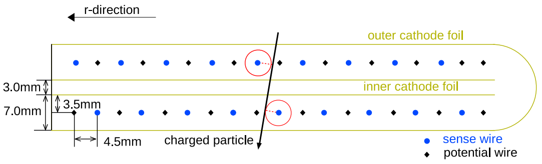
As a result of all efforts of material reduction, we constructed a tracking system with $2.0 \times 10^{-3}\ \mathrm{X}_0$ in total along a positron trajectory. Detailed description about the design, construction, and performance in 2007 engineering run are found in [41].



(a) Picture of a module.

(b) Design of a module.

Figure 3.15: A module of drift chamber.



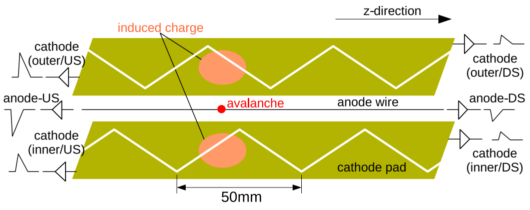
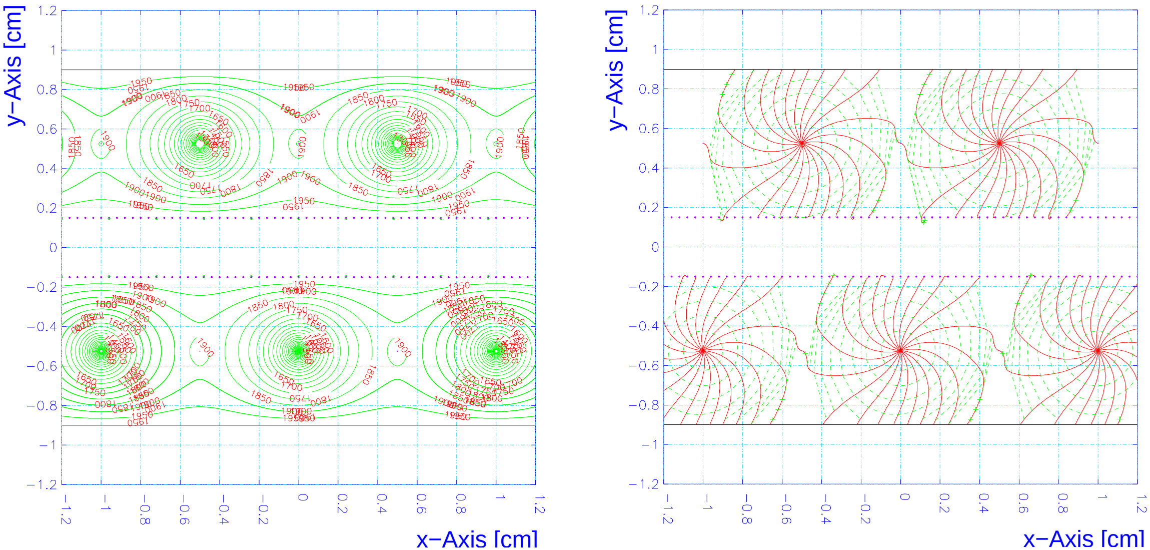


Figure 3.16: Cell configuration of a drift chamber module.

Figure 3.17: Schematic of vernier pattern on the cathode pads.



(a) Contour plot of the potential. (b) Drift lines with isochrone map.

Figure 3.18: Field map and drift lines of drift chamber calculated by the GARFIELD simulation.

Timing Counters

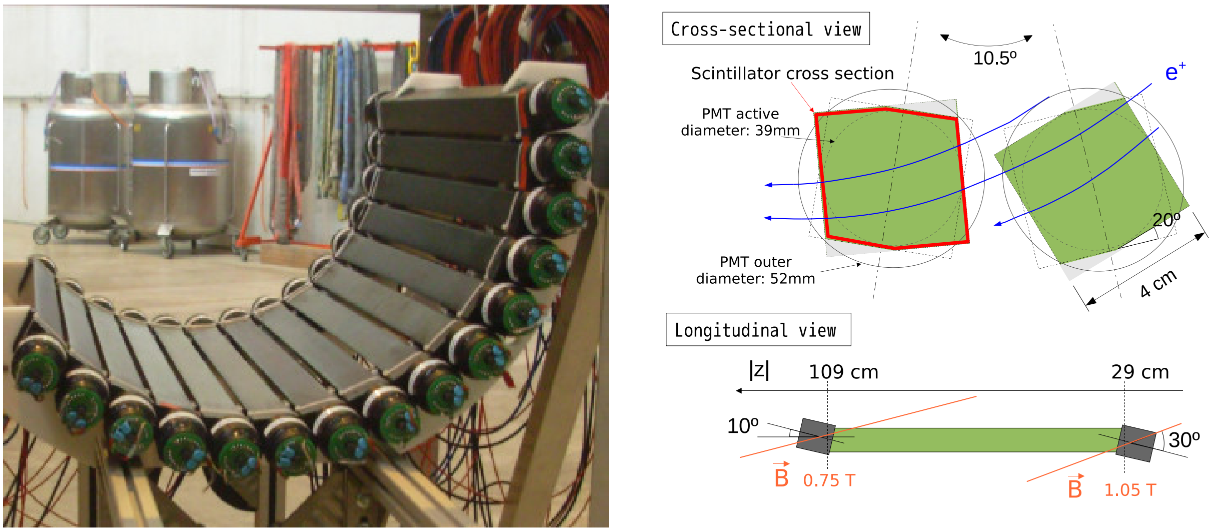
A pair of timing counters (TIC) are placed at the both sides of the drift chamber system. Positron tracked by the drift chamber finally hits the timing counter and the impact timing and position are measured. A module of the timing counter consists of two layers of scintillation counter arrays stacked orthogonally each other. The two layers have different structures and each has specific tasks. The outer layer is called ϕ-counter. Its main tasks are precise measurement of positron impact time and fast information of positron ϕ-emission angle. The inner one is called z-counter. Its main tasks are precise measurement of z impact position and fast information of positron θ-emission angle.

The ϕ-counter consists of 15 plastic scintillation bars ($4 \times 4 \times 80$ cm^3, BICRON BC-404 [42]). They are placed at 10.5° intervals in ϕ-direction at a radius of 32 cm and they cover 160° in total ($-150° < \phi < 10°$). Two 2-inch fine-mesh PMTs (HAMAMATSU R5924 [43]) are attached on both ends of each bar. A picture of the counter is shown in Figure 3.19(a), and its detail design is illustrated in (b).

The z-counter consists of 128 scintillating fibers (6×6 mm^2, SAINT-GOBAIN BCF-20 [42]). Each fiber is separated optically at the center and read out independently at the both ends by a 5×5 mm^2 silicon avalanche photo-diode (APD) (HAMAMATSU S8664-55 [43]). The z-counters are mounted on the ϕ-counters at a radius of 29 cm. A picture of the counter is shown in Figure 3.20.

Details of the design, construction, and performance in beam tests are described in [44][45].

In 2008 operation, we got some problems on the z-counter read out. We did not use the z-counter in trigger. Instead we used ϕ-counter also to get the z impact position with a slightly worse resolution. Also in the analysis of this thesis, the z-counter is not used, even though the fundamental functions were studied and confirmed with the data.



(a) (b)

Figure 3.19: Picture and design of timing ϕ-counter. The scintillation bars are mounted with a slant angle at 20° so that the path lengths of positrons inside the bar become uniform. The bars are hexagonally shaped to avoid geometrical conflict due to the rotation. PMTs are attached with a slant angle at 10° with respect to the z-axis (at $\sim 30°$ with respect to the magnetic field) to recover their operation in a magnetic field [46].

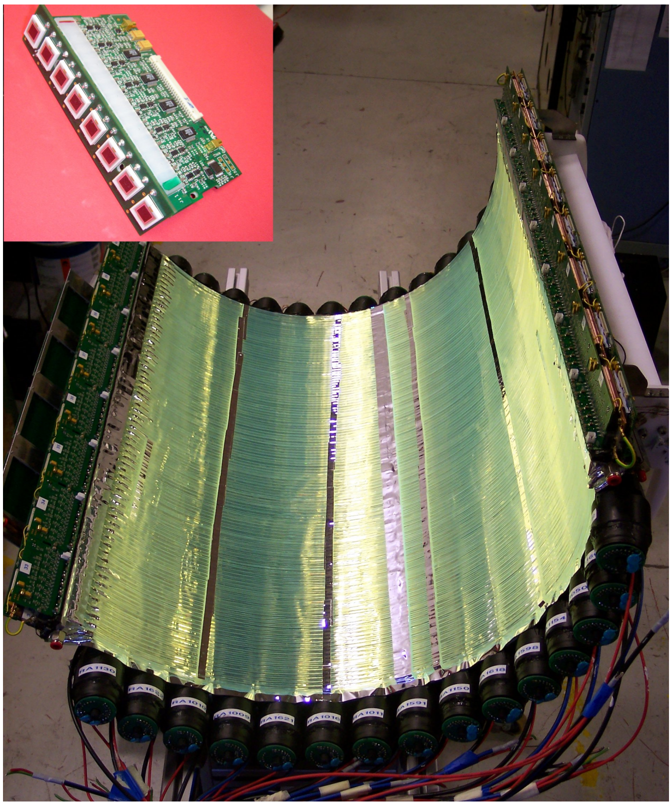


Figure 3.20: Picture of timing z-counter and APDs. It is mounted on the ϕ-counter.

3.2.2 The Gamma-ray Detector

The gamma-ray detector is undoubtedly the most innovative and challenging part of the experiment. Its performance is crucial for a successful search for the $\mu^+ \rightarrow e^+\gamma$ decay. We use a gamma-ray detector of a 900 liter homogeneous volume of liquid xenon (LXe). It is placed just outside of the COBRA magnet. Gamma rays that penetrated the COBRA spectrometer enter the detector. They interact with LXe and generate scintillation light. The scintillation light is collected by a number of photomultipliers (PMT) surrounding the active volume of LXe to measure the total energy released by the incident gamma ray as well as the position and time of its first interaction. A conceptual figure of the gamma-ray detector is shown in Figure 3.21. Sometimes multiple gamma rays enter the detector and are measured at the same time in a high rate of low-energy gamma-ray background because the detector consists of a large volume without any segmentation. Nevertheless, we can handle those pileup events correctly because the image of the light distribution from a large number of PMTs enables us to identify and unfold those multiple events. In addition, the time distribution and waveform can also be used to identify pileup events.

The R&D works, performance of prototype detector, design and construction of final detectors are described in detail in [47][48].

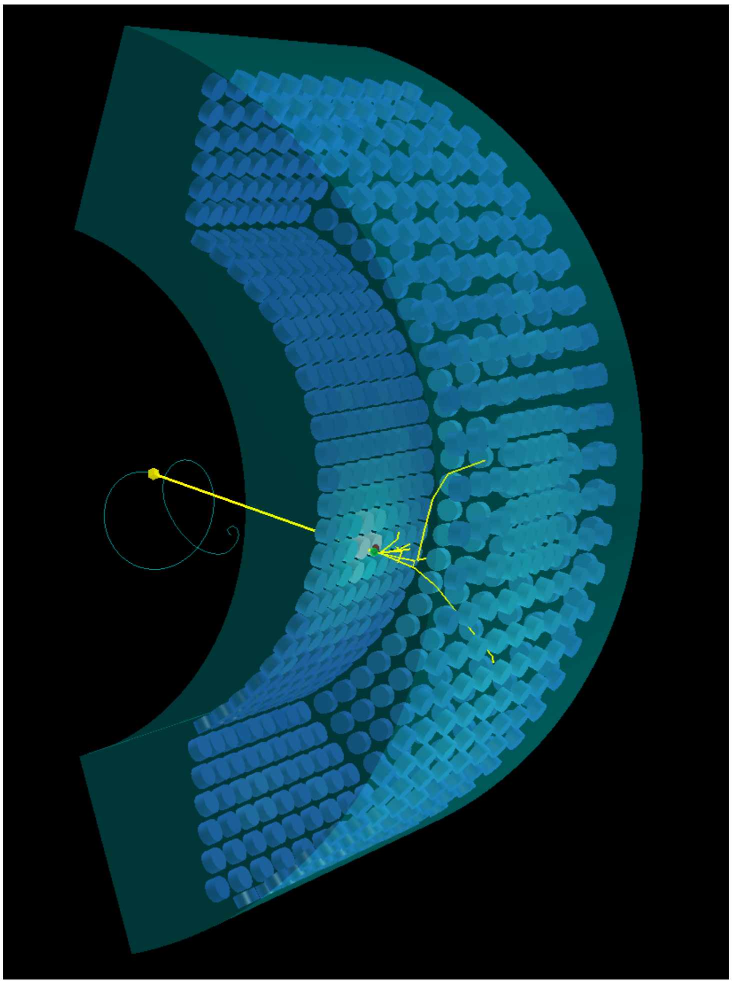


Figure 3.21: Conceptual figure of LXe gamma-ray detector.

Table 3.1: Properties of LXe.

Material Properties	Value & Unit	Ref.
Atomic Number	54	
Atomic Weight	131.293	[4]
Density at 161.4 K	2.978 g/cm^3	[53]
Boiling point	165.1 K	[4]
Melting point	161.4 K	[4]
Triple point(temperature)	161.3 K	[54]
Triple point(pressure)	0.805 atm	[54]
Radiation length	2.77 cm	[4]
Critical Energy	14.5 MeV	[55]
Mollier radius	4.2 cm	[55]
Scinti. wavelength (peak±FWHM)	(178 ± 14) nm	[56]
Refractive index at 175 nm	1.57 to 1.72	[57, 58, 59]
W_{ph} for electron	21.6 eV	[60]
W_{ph} for α particles	17.9 eV	[60]
Decay time (recombination)	45 ns	[61]
Decay time (fast components)	4.2 ns	[61]
Decay time (slow components)	22 ns	[61]
Absorption length	> 100 cm	[47]
Scattering length	29 cm to 50 cm	[59, 62, 63, 64]

Liquid Xenon as a Scintillation Material

Xenon is widely used as a detector material in different fields (see, for example, [49] for a recent review). It produces both charge carriers and scintillation photons in response to radiation. Therefore, it can be used as a scintillation or ionization medium. In some application, both of them are used simultaneously. In our detector, only the scintillation light is used.

The photon cross section of xenon atoms are shown in Figure 3.22. Photon around 50 MeV interacts with xenon mainly through the pair production. The scintillation mechanisms are given by the following two processes [51, 52]. One is a self-trapping process of excited xenon atom,

$$\begin{aligned} \mathrm{Xe}^* + \mathrm{Xe} + \mathrm{Xe} &\rightarrow \mathrm{Xe}_2^* + \mathrm{Xe}, && (3.2) \\ \mathrm{Xe}_2^* &\rightarrow 2\mathrm{Xe} + h\nu, && (3.3) \end{aligned}$$

where $h\nu$ is vacuum ultraviolet (VUV) scintillation photon. There are two states of the excited dimer Xe_2^*, singlet and triplet, which correspond to fast and slow components, respectively. The other process is a recombination process originating from a xenon ion,

$$\begin{aligned} \mathrm{Xe}^+ + \mathrm{Xe} &\rightarrow \mathrm{Xe}_2^+, && (3.4) \\ \mathrm{Xe}_2^+ + \mathrm{e} &\rightarrow \mathrm{Xe}^{**} + \mathrm{Xe}, && (3.5) \\ \mathrm{Xe}^{**} &\rightarrow \mathrm{Xe}^* + \mathrm{heat}, && (3.6) \\ \mathrm{Xe}^* + \mathrm{Xe} + \mathrm{Xe} &\rightarrow \mathrm{Xe}_2^* + \mathrm{Xe}, && (3.7) \end{aligned}$$

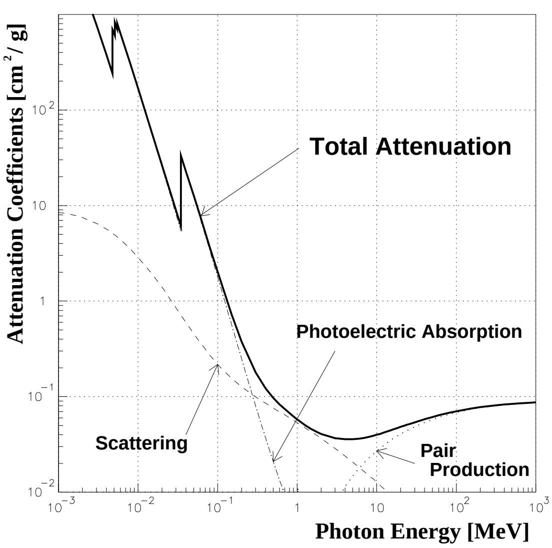


Figure 3.22: Photon cross section of xenon atoms [50].

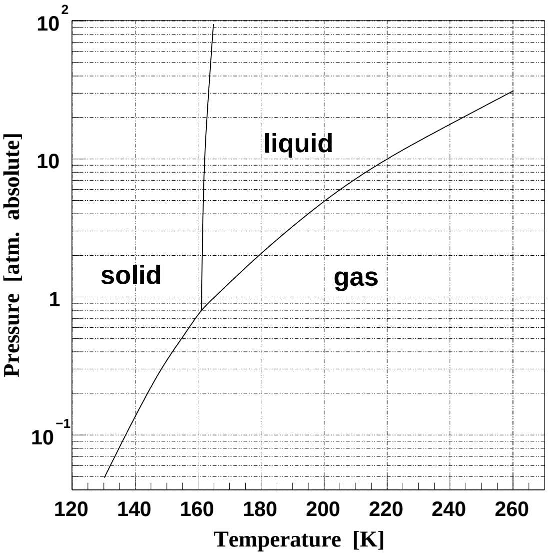


Figure 3.23: Xenon phase diagram.

$$\mathrm{Xe}_2^* \quad \rightarrow \quad 2\mathrm{Xe} + h\nu, \tag{3.8}$$

where the time dependence of the scintillation is dominated by the kinematics of the recombination process. In both processes, the scintillation photon comes from a de-excitation of a excited dimer to dissociative ground state.

Main properties of LXe and its scintillation are summarized in Table 3.1. Here, we can see several fascinating properties of LXe for a gamma-ray detector. The large stopping power for penetrating radiation due to its large atomic number and high density makes it efficient to detect gamma ray. It has large light yield which is necessary for the precise energy measurement. Its fast response enables high timing resolution. It is also suitable for the measurement in high rate environment to minimize pileup event. Those excellent properties become clear by comparing LXe with various other scintillators in Table 3.2. In addition, LXe has some other advantages. Since the scintillation light is not emitted by the excited state of xenon atom (Xe^*) but the excited dimer (Xe_2^*), xenon scintillation photons cannot be re-absorbed by xenon atoms themselves. This is an advantage to achieve high energy resolution with a large detector. We can make a single homogeneous detector with arbitrary size and shape because it is liquid. It is free from aging or damage by radioactivity. We can purify it at any point, which is not possible for crystal scintillators once they are constructed. We concluded LXe is the best scintillator for the gamma-ray detector of MEG.

However, we have to overcome some difficulties in practical usage of LXe. The scintillation is in VUV light. Those light can be easily absorbed by some contaminants like water or oxygen. The detector performance heavily depends on the purity of LXe. We also have to use VUV sensitive photon sensors. Another difficulty comes from the low temperature. The phase diagram of xenon is shown in Figure 3.23. We can see a narrow temperature range for liquid phase at around atmospheric pressure. Stable control of its temperature around 165 K is required to keep it in liquid.

Table 3.2: Comparison of various scintillators.

	LXe	LAr	NaI(Tl)	CsI(Tl)	BGO	LSO(Ce)	$PbWO_4$
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 output	75	90	100	165	21	83	0.083/0.29[1]

---

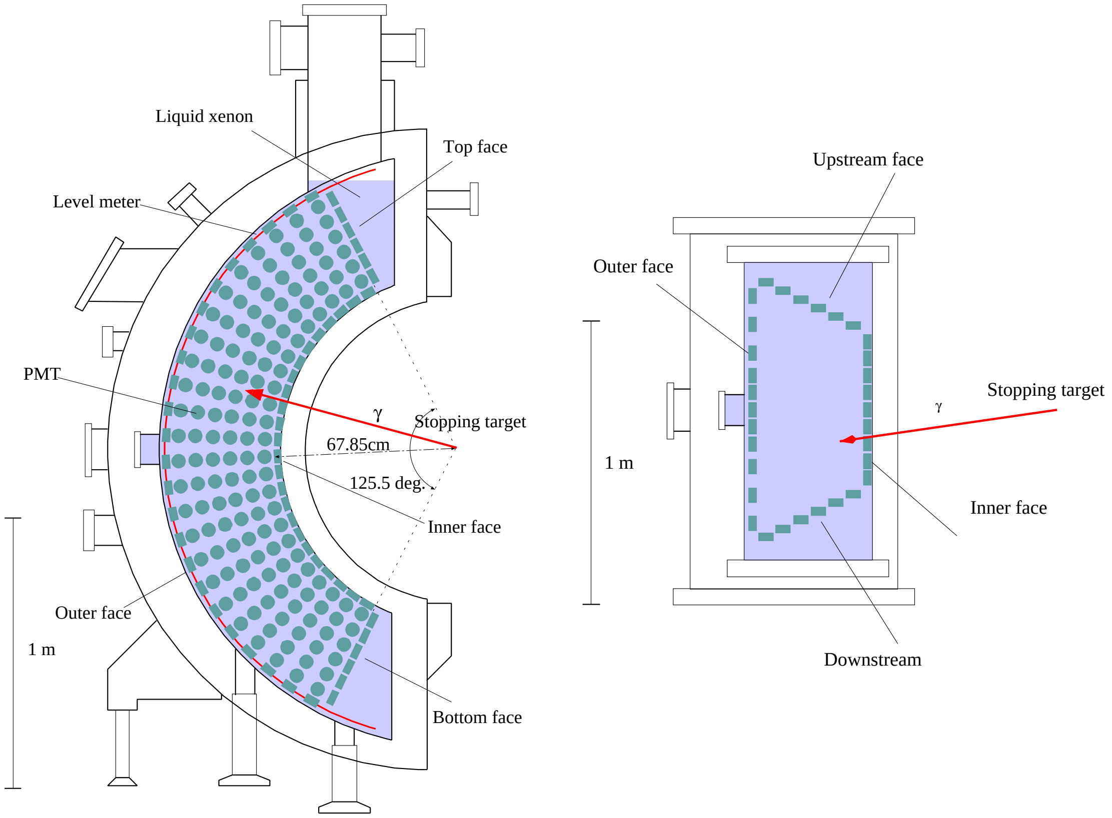
[1] slow/fast component

Design of the MEG LXe Detector

The schematic view of MEG LXe gamma-ray detector is shown in Figure 3.24. It is C-shaped to fit the outer radius of COBRA magnet. It consists of ∼900 liter LXe and 846 PMTs. The PMTs are placed on all of six faces of the detector and directly immersed in LXe surrounding the active volume of LXe. The definition of the six faces (inner, outer, upstream, downstream, top, and bottom) are also shown in the figure. The arrangement and density of PMTs are different for each face. They are shown in the development view of the detector in Figure 3.25. PMTs are most closely arranged on the inner face and its coverage of active photo-cathode is about 35 %. Part of the outer face has dense part whose density is the same as that of the inner face. This dense part is for the possibility of calibration with gamma ray impinging from the back of the detector. The active volume of the detector is ∼ 800 liter and it covers 11 % of the solid angle from the stopping target. The depth of the active volume is 38.5 cm, which corresponds to ∼ 14 X_0 and keeps the shower inside the detector.

Cryogenic System

A powerful and stable cryogenic system is indispensable to operate the large LXe detector since the temperature range of LXe is narrow as already shown. LXe is filled in a cryostat consisting of two layers of vacuum-tight vessels. Outer one makes a vacuum layer for thermal insulation. The gamma-ray entrance window is designed to be as thin as possible to maximize the probability of gamma-ray penetration. The window of outer vessel is made of a 0.7 mm thickness stainless steel plate, while that of inner vessel is made of aluminum honeycomb panels covered with carbon fiber plates because it requires mechanical strength up to ∼ 3 atm. The total thickness of the window is 0.075 X_0. A turbo-molecular pump is directly attached to each vessel to evacuate with high conductance. In addition, a cryo-pump is installed to the inner vessel to efficiently remove water vapor that is the main component of residual gas from the detector. A 200 W pulse-tube refrigerator [65], which was developed for this LXe detector, is mounted on top of the cryostat and controls the temperature of LXe. Its picture and cooling power are shown in Figure 3.26. In addition, cooling pipes of LN_2 are also available when a high cooling power is necessary such as in liquefaction. Several Pt-100 sensors are placed at different positions inside the cryostat and measure the temperatures. The level of LXe is measured



(a) Side view. (b) Top view.

Figure 3.24: Schematic view of LXe gamma-ray detector.

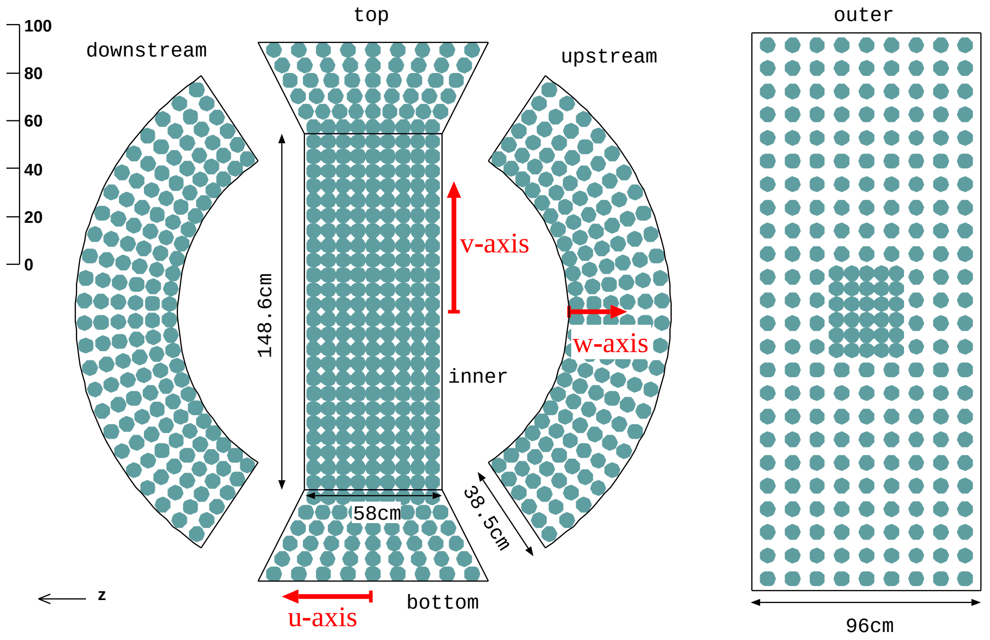
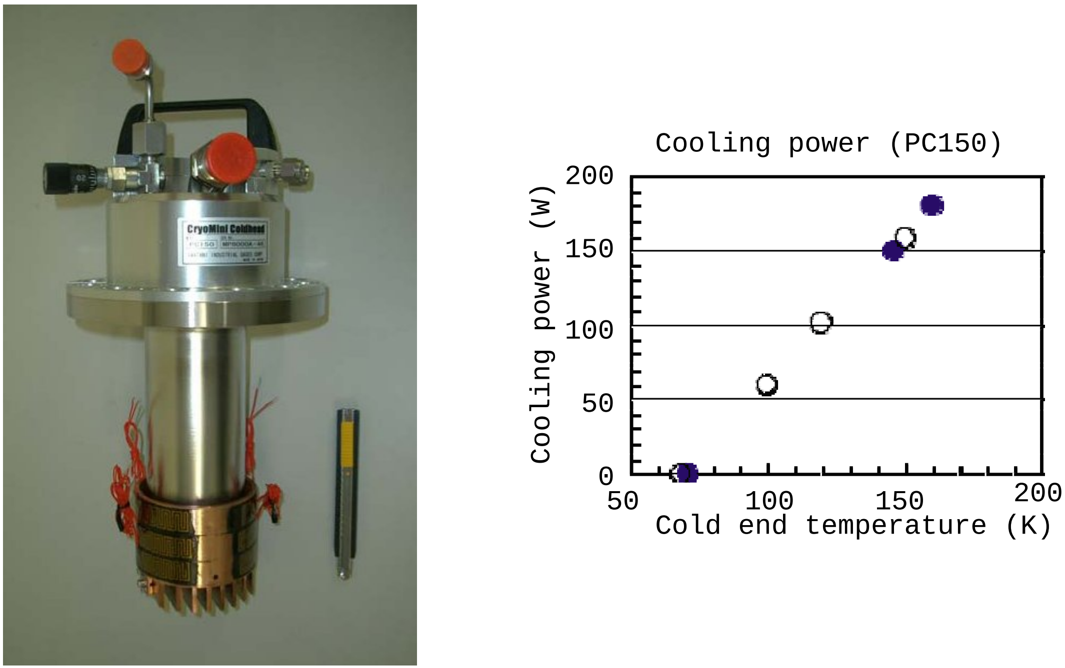


Figure 3.25: Development view of LXe gamma-ray detector. The local coordinate system of the detector is also shown as red arrows. Each light blue circle shows a PMT.

by a capacitance level meter. A picture of the detector is shown in Figure 3.27. This picture was taken just before closing covers. All components were mounted.

Besides the detector itself, two kinds of xenon storage systems and two types of circulation systems build the xenon system of MEG. One of the storage systems consists of eight high-pressure gas tanks (Figure 3.28(a)). The other is a dewar that can hold 1000 liter xenon in liquid phase [66] (Figure 3.28(b)). Both system can store the whole amount of xenon used in the experiment. The high-pressure gas tanks supply xenon directly to the detector or to the 1000-liter dewar by the pressure difference. Normally, we liquefy xenon in the 1000-liter dewar in parallel to the maintenance and preparation of the detector. The 1000-liter dewar and the detector are connected by a liquid-transfer line. Once the liquefaction is completed in the 1000-liter dewar, we can speedily transfer xenon to the detector in liquid phase. The recovery of LXe from the detector is also done in liquid using the liquid-phase circulation pump in the purification system.



(a) (b)

Figure 3.26: (a) Picture of 200 W pulse-tube refrigerator. (b) Its cooling power when operated with 6.5 kW GM helium compressor.

PMT

We developed a new model PMT suited to MEG in cooperation with HAMAMATSU PHOTONICS. The PMT, R9869, is shown in Figure 3.29 and its properties are summarized in Table 3.3. The PMTs must work in low-temperature LXe and be sensitive to the VUV LXe scintillation light. A quartz window is used to transmit VUV light. We chose Bialkali (K-Cs-Sb) for the VUV sensitive photo-cathode. Aluminum strips are attached on the photo-cathode to prevent from the increase in the sheet resistance at a low temperature. Thus, we achieved high sensitivity of ∼15 % quantum efficiency. In 2008, PMT gains were set to be about 1.7×10^6 with typical HV of 810 V.

Figure 3.27: Picture of LXe gamma-ray detector. It was just before closing covers. All components were mounted.

Calibration Devices

In the detector, several LEDs and radioactive sources are installed as constant light sources for PMT calibrations. Blue LEDs are placed at six positions on each lateral face. We can control the number and set of LEDs to be flashed, their intensity, and their flashing rate by remotely controlling LED drivers. Typically 10 LEDs are flashed at the same time in normal calibration to illuminate all the PMTs uniformly. Each LED is covered with a hand-made filter made of aluminum foil with some small pin-holes. It works as a light attenuator. Owing to the filter, we can apply higher voltage to the LED, resulting in less noise fluctuation in emitted light and also the stable operation of LED. LEDs are mainly used to calibrate PMT gains.

The LDE light has different wavelength from that of LXe scintillation. Therefore, it is not suitable for the absolute response calibration or quantum efficiency measurement of the PMT because the quantum efficiency depends on the wavelength. Instead, we can

(a) (b)

Figure 3.28: Pictures of xenon storage systems. (a) eight high-pressure gas storage tanks. (b) 1000 liter liquid storage dewar.

Figure 3.29: Picture of PMT for the LXe gamma-ray detector (R9869).

obtain the scintillation light using radioactive sources. We developed a calibration method with a radioactive point-source lattice [67]. We use ^{241}Am as an alpha source. It emits alpha ray at 5.485 MeV (84.5 %) and 5.443 MeV (13.0 %). Since the range of the alpha ray in LXe is as short as 40 μm, it can be used as a point-like light source from a known position. The half life of ^{241}Am is sufficiently long (432 years) so that we can regard it to be constant. The activity of each source is $\approx$ 200 Bq. The alpha sources are attached on thin (100 μm diameter) tungsten wires as shown in Figure 3.30. On each wire, alpha sources are attached at five positions with a constant interval of 12.4 cm. Five wires are installed in the detector. Therefore as a whole, a lattice of 25 alpha-source spots is formed in LXe active volume. In Figure 3.31, one of the wires installed in the detector can be

Table 3.3: Properties of the PMT (R9869). These are typical values.

Size	57 mm ϕ
Active area size	45 mm ϕ
PMT length	32 mm
Photo-cathode material	K-Cs-Sb
Dynode type	Metal channel
Number of dynode	12
Typical HV	900 V
Typical gain	1×10^6
Typical Q.E.	15 %
Rise time	2 nsec
Transit time	12.5 nsec Typ.
TTS	0.75 nsec Typ.

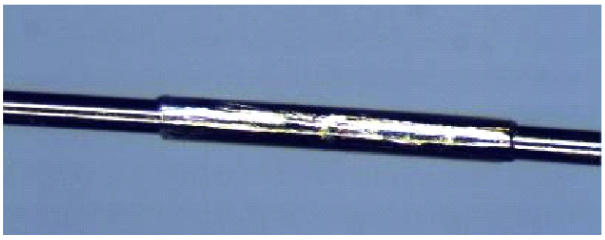


Figure 3.30: Micro-picture of a ^{241}Am source mounted on a 100 μm-diameter wire. The longitudinal dimension is ~2 mm

seen together with LEDs mounted on a lateral face.

Purification System

The performance of the gamma-ray detector severely depends on a purity of LXe. Owing to the scintillation mechanism through the excited dimer Xe_2^*, LXe itself is transparent to its own scintillation light. However, possible contaminants in LXe such as water and oxygen at ppm level considerably absorb the VUV light. Figure 3.32 shows absorption of the LXe scintillation light by water vapor and oxygen. The absorption spectrum of water vapor largely overlaps with the LXe scintillation spectrum. Water absorbs VUV light by the photo-dissociation process

$$H_2O + h\nu \rightarrow H_2 + O^*. \tag{3.9}$$

Water was found to be the dominant contaminant by investigating the residual gas in the inner vessel of prototype detector. The residual water vapor originates at out-gas from materials inside the detector such as detector wall and PMT support structures because we cannot heat the detector while evacuating the cryostat owing to the PMTs placed inside. Thus, water is the most dangerous contaminant for the LXe scintillation in our detector.

Figure 3.31: Picture of LEDs and alpha sources. LEDs are shown with orange circles. A wire of alpha source is indicated by a yellow arrow.

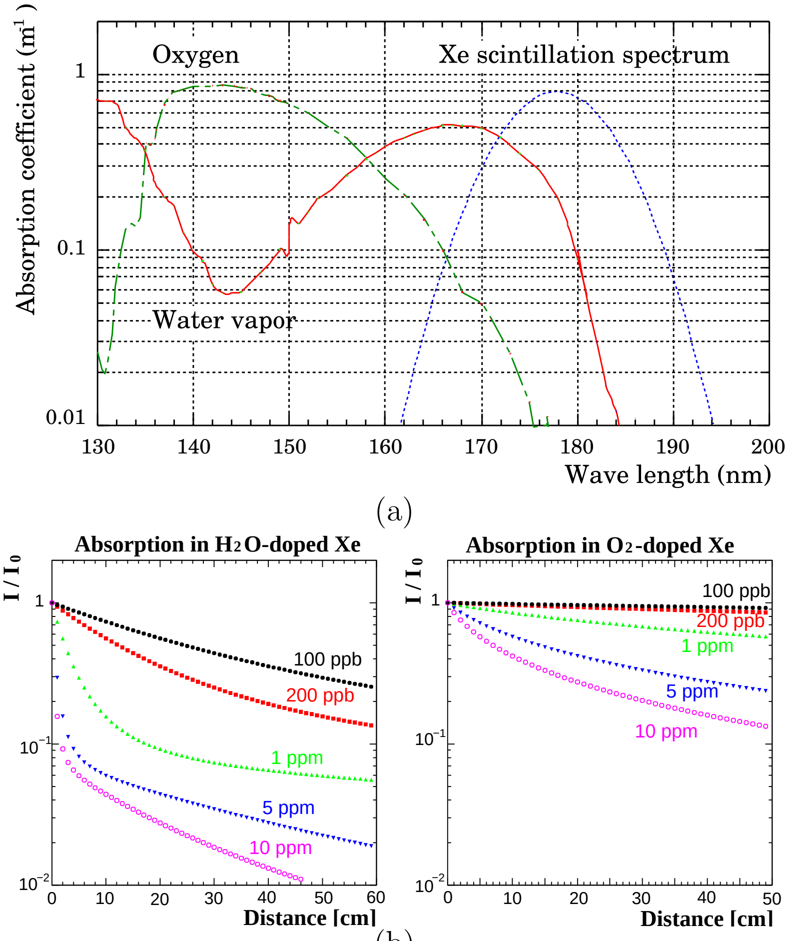
Besides the absorption, some contaminants cause quenching processes of LXe scintillation. Quenching processes of LXe scintillation are poorly known. In [69], the main reaction of quenching process of liquid argon (LAr) by N_2 contaminant was considered to be a two-body collision of the excimer states with the contaminant,

$$\mathrm{Ar}_2^* + \mathrm{N}_2 \rightarrow 2\mathrm{Ar} + \mathrm{N}_2. \tag{3.10}$$

Since it is a kinematic collision, the rate depends on both the concentration of the contaminant and the lifetime of the excimer. As a consequence, variation of the purity changes the time dependence of the scintillation light. The article reported a suppression of slow component of the LAr scintillation as nitrogen concentration increases. A similar quenching process can occur also in LXe.

To remove impurities in LXe, we developed two types of purification systems. One is a liquid-phase purification system consisting of a cryogenic centrifugal pump for a liquid circulation and a purifier cartridge of molecular sieves. We developed this system using the prototype detector [70]. It is dedicated to remove water contaminant. In addition, a purifier using copper beads, which was developed for LAr purification at CERN, was installed at the outlet of the circulation pump to remove oxygen contaminant. Owing to the circulation in liquid, the circulation speed is as high as ∼ 35 l/hour. The low conductance of the copper purifier limits the speed. We performed the liquid-phase purification before the data taking and during the beam-maintenance periods. We did not during data taking since the pump generates noise on detectors.

The other purification system is a gas-phase one using a metal-heated getter. A diaphragm pump circulates gas xenon taken from the detector. During the circulation, the getter removes H_2O, O_2, CO, CO_2, N_2, H_2, and hydro-carbon molecules from gas xenon down to 1.0 ppb level. This system was also developed using the prototype detector and the effectivity was confirmed [71]. The circulation speed is very slow compared to the liquid-phase purification (< 100 cm^3 liquid/hour). However, it can be used in parallel with data acquisition.



(a)

(b)

Figure 3.32: (a) Absorption coefficient of VUV light in 1 ppm water vapor and oxygen. Xenon scintillation spectra is superimposed [68]. (b) Light absorption for various concentrations of water and oxygen in LXe as a function of distance from a light source [47].

3.3 Electronics and Data Acquisition

3.3.1 Electronics Chain

The data flow from detectors to digitizers is schematically summarized in Figure 3.33. The outputs from the detectors pass through several devices with being processed adequately for each sub-detector, and are finally acquired as waveform digitized by a fast waveform digitizer. Owing to the waveform digitizer and a field programmable gate arrays (FPGA) trigger, the electronics chain becomes simple compared to that with conventional ADC/TDC and logical trigger circuit.

PMT outputs from the gamma-ray detector are put in active splitters via patch-panels, feed-through, and coaxial cables. The active splitter has three outputs: the first one is a wide-band (1.9 GHz) fully differential output, the second one is a 320 MHz-bandwidth differential output, and the third one is a four-to-one sum output. The first one goes to the waveform digitizer, the Domino Ring Sampler (DRS). The others go to the trigger. For the inner PMTs, the one-to-one outputs are used for the trigger while the sum outputs are used for the other PMTs. The splitter inverts negative PMT pulses to positive ones so that the signal will match the digitizer's dynamic ranges. The active splitter is connected to the DRS by a 2 m long high-density (0.68 cm pitch) twisted-pair cable, and to the trigger by a 2 m long lower-density (1.27 cm pitch) one.

The output from the timing ϕ-counter's PMT first goes to a passive splitter, and is

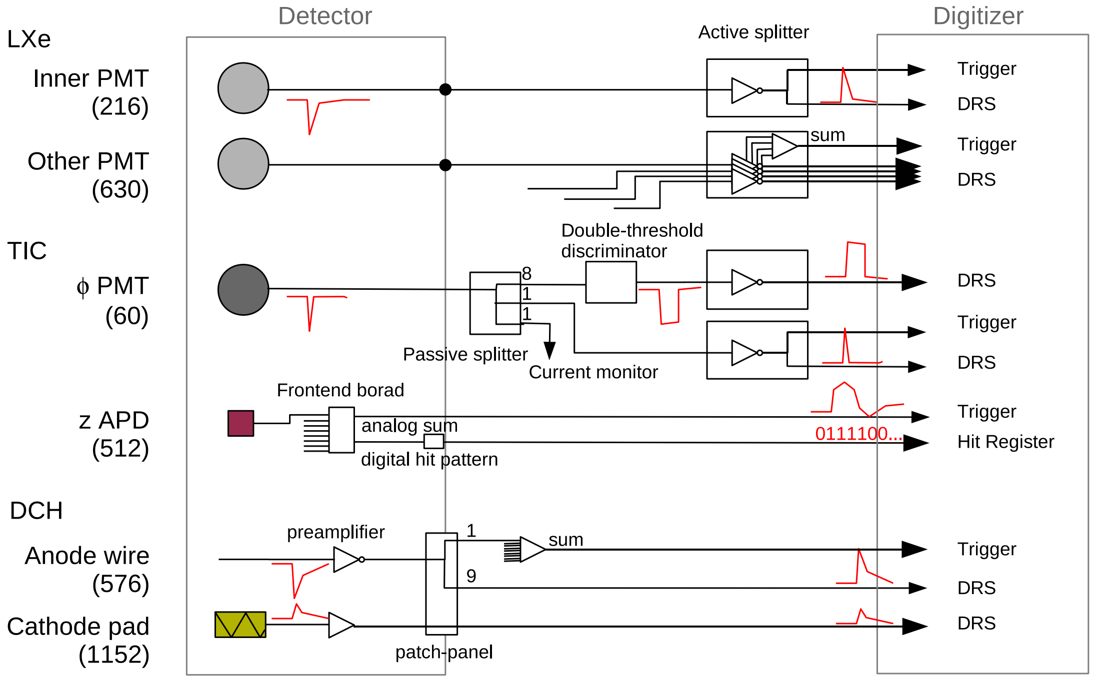


Figure 3.33: Schematic of data flow and electronics.

divided into three outputs with a fraction of 8:1:1. The largest output goes to a double-threshold discriminator (DTD), which has two levels of thresholds for discriminating a pulse of positron hit from noise or low-energy delta-ray hits. The lower level is used for the timing determination, while the higher is for the validation of pulse which is not a noise hit and has sufficient energy loss in the bar. Owing to this mechanism, we can lower the time-determining level as low as possible resulting in intrinsically good time resolution and minimum time-walk effect. In 2008, we set the lower level at 25 and the higher one at 800 mV. The discriminator outputs standard NIM pulses[2] of 50 ns width. The NIM pulse goes to the active splitter and the first output goes to the DRS. One of the smaller outputs of the passive splitter goes to the active splitter and the first output goes to the DRS and the second one goes to the trigger. The other output of the passive splitter goes to current monitor to check the PMT lifetime.

Eight APDs of the timing z-counter are mounted on a frontend board together with preamplifiers. This frontend board outputs an analog sum signal of the eight channels and a digital hit-pattern signal of individual channels. The analog output goes to the trigger while the digital hit-pattern goes to a dedicated hit register module.

We get six waveforms from a drift cell of the drift chamber: two from the both ends of a anode wire and four from the vernier-pattern cathode pads (two for outside pads and two for inside ones). The cathode pads are read out at one of the ends. Those outputs are first amplified by a preamplifier attached at the chamber frame structure. Here the anode output is inverted. At the patch-panel on the end-cap of the COBRA magnet, the wire signal is resistively divided into two outputs with a fraction of 9:1. The larger one and the cathode signals go to the DRS directly by coaxial cables. The smaller one is amplified and summed up over several wires. This sum signal goes to the trigger.

3.3.2 Data Acquisition System

An overview of the data acquisition (DAQ) system is shown in Figure 3.34. The MIDAS system [72] controls the whole system. There are mainly nine frontend sub-systems: four for trigger and five for DRS. Each of them, which corresponds to a VME crate, is controlled by a Linux PC. Each frontend is running with multi-threading process, and each event is stored in a ring buffer. In the main DAQ PC, an event builder is running and each event is re-built from the ring buffer of each frontend. The MIDAS system provides not only the control of frontend processes but also the logging system, the online database system for the parameters of DAQ and trigger, the slow control system, the alarm system, the history monitoring system, and web interfaces.

During the data taking, an online monitor is running and with its graphical display the quality of the data is checked in parallel with the DAQ. In addition, after each run is finished an offline process is run automatically and more detail information can be available within a few dozens of minutes.

---

[2] 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* correspond to the presence of a hit here. The discriminator itself is not a NIM standard module.

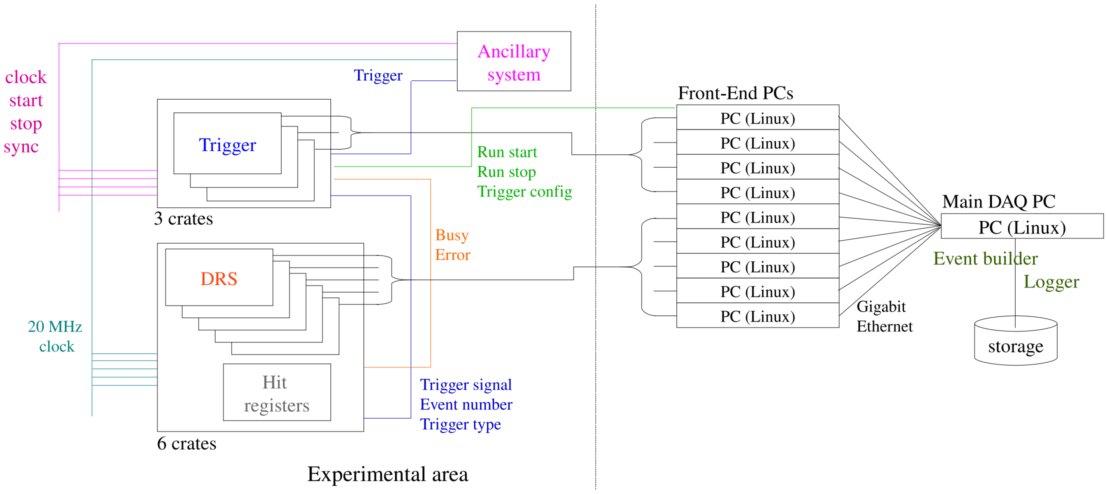


Figure 3.34: Schematic of DAQ system.

3.3.3 Trigger

The trigger system is based on a coupled use of flash analog to digital converters (FADC) and FPGA. The input signals are sampled by FADCs and the digitized information is analyzed by FPGAs.

The whole system is arranged in a tree structure on three layers with two different types of boards. The first layer consists of the Type1 boards compliant with the 6U VME standard. The Type1 board receives and digitizes analog input signals with FADCs (AD9218 [73]) at 100 MHz with 10 bits resolution , implements some reconstruction algorithms on a large size FPGA (Xilinx Virtex-IIpro [74]), and sends the information to the successive trigger layer. The two remaining trigger layers consist of a second type of boards (Type2) compliant with the 9U VME standard. The second layer determines trigger conditions of sub-systems, and finally the third layer makes a trigger decision.

In addition, an ancillary system was developed to ensure synchronous operation of the tree. It consists of ancillary boards of 9U VME boards, and distributes a reference clock and controls signals such as start and stop of the DAQ system. A master board hosts the reference clock oscillator (SaRonix SEL3935 [75]) generating a 19.44 MHz squared clock, and receives control signals from the third layer of the trigger tree. Those signals are fanned-out by three slave boards and distributed to all boards of the trigger and DRS. The jitter of distributed clock is measured to be less than 30 ps.

The trigger for the $\mu^+ \to e^+\gamma$ event candidate (MEG trigger) is decided by the following three conditions: the gamma energy, positron-gamma time coincidence, and positron-gamma direction match. Those kinematic conditions are checked with fast reconstruction algorithms implemented in the FPGA. A requirement of global trigger latency to be less than $\sim$500 ns prevents us from using information of the drift chambers that have slower signal due to the drift time. The pulse height of sum waveform of all PMTs gives an estimation of a gamma energy. A set of PMT-calibration factors can be incorporated in the online reconstruction. The threshold was set to around 40 MeV to guarantee the full efficiency at the signal energy. The online resolution is measured to be 13.8 % in FWHM. The time coincidence is checked using the time difference between the PMTs from the gamma-ray detector and the timing ϕ-counter. The PMT time is extracted by a parabolic interpolation of the waveform leading edge. We set the time window to 20 ns taking the worse time resolution and the spread of the positron time-of-flight into account. The online time resolution is measured to be 3.4 ns in sigma. The position of the inner face PMT which observes the maximum light gives an estimation of the gamma direction. There is a correlation between the emission angle of a positron on the target and the hit position on the timing counter for the signal energy positron. This correlation is used to estimate positron direction. In 2008, the charge ratio of ϕ-counter PMTs was used for the z hit position reconstruction instead of the APD signal from the z-counter. The resolution is measured to be 7.2 cm in sigma. The correlation is investigated using the MC simulation, and a look-up table was formed in advance. Collinearity of the gamma and positron directions is checked with this table and the index of the maximum output PMT in the gamma-ray detector. With these three conditions, the trigger rate was successfully reduced to $\sim$ 6.5 Hz when the muon stopping rate was 3×10^7/sec with the trigger latency of less than 400 ns.

In addition to the MEG trigger, various kinds of trigger settings were prepared for the calibration and the normalization. Table 3.4 lists those trigger settings. Multi-trigger settings can be used simultaneously with a set of pre-scale factors. In the normal data taking, total 11 settings were used with the pre-scale factors listed in the table. The trigger system also provides trigger rate scalers for each type of trigger setting which can be used to calculate efficiencies and the normalization factor.

3.3.4 Domino Ring Sampler

The waveform image gives us useful information such as pileup and particle identification, noise properties, and event-by-event baseline, as well as the precise charge and time information. Thus the waveform digitizing can be a superior to the conventional ADC/TDC. In high-energy physics, FADCs have been widely used as a waveform digitizer. They are typically used in the range of 50 to 250 MHz sampling and 10 to 12 bits resolution. However, with going to higher sampling frequencies, they suffer from the high power consumption and high cost. An alternative to the FADC is the usage of switched capacitor arrays (SCA). The DRS is one of the SCAs developed at PSI.

A channel of DRS contains 1024 capacitive sampling cells fabricated in a 0.25 μm CMOS and samples waveform at frequency raging from 0.5 to 5 GHz. The principle of the operation is illustrated in Figure 3.35. The sampling frequency is generated on chip by a series of inverters because it is very hard to generate and distribute clock signals in the GHz range, The sampling frequency is freely and continuously running on the inverter chain in circular fashion (domino wave). When an external trigger signal comes, the domino wave is stopped and the waveform stored in the sampling cells is kept. The stored signal is then read out by a shift resistor at lower frequency (33 MHz) and digitized externally by a 12-bits commercial FADC. Several channels can be multiplexed into a single FADC channels to reduce the cost. For the synchronization among chips, an external global clock can be sampled in each chip. A DRS chip consists of eight signal channels, a trigger signal channel and a clock signal channel. Two chips are mounted on

Table 3.4: List of trigger settings.

#	Name	Prescale in physics run	Description
0	MEG	1	$\mu \to e\gamma$. $(Q_{LXe} > Q_{high}) \wedge (\lvert\Delta T\rvert < T_N) \wedge$ direction-match
1	MEG-Q	150	$\mu \to e\gamma$. low Q_{LXe} threshold
2	MEG-D	500	$\mu \to e\gamma$. wide direction-match window
3	MEG-T	100	$\mu \to e\gamma$. wide time-coincidence window
4	RD	1000	RD, π^0-Dalitz. $(Q_{LXe} > Q_{high}) \wedge (\lvert\Delta T\rvert < T_N)$
5	RD-T	-	RD, CW-B. wide time-coincidence window
6	Pi0	-	π^0. $Q_{LXe} \wedge$ $Q_{NaI} \wedge$ time-coincidence b/w LXe and NaI
7	Pi0-T	-	π^0. $Q_{LXe} \wedge$ Q_{NaI}
8	NaI	-	NaI self. (Q_{NaI})
9	LXeHighQ	8000	LXe self. $(Q_{LXe} > Q_{high})$
10	LXeLowQ	-	LXe self, CW-Li, alpha. $(Q_{LXe} > Q_{low})$.
14	LED	10	LED. trigger signal from LED driver.
16	Michel	-	Michel. DCH hits $\wedge$ TIC hits
18	DCH	10^7	Michel (DCH self).
22	TIC	10^7	Michel (TIC self), Cosmic-ray on TIC.
24	TICPair	-	TIC test. A pair of TIC bar hits.
27	LXeCR	600	Cosmic-ray in LXe. $(Q_{LXe} > Q_{CR})$
31	Pedestal	20000	Pedestal.

a mezzanine board, and two mezzanine boards are mounted on a VME board. A picture of the boards and chips is shown in Figure 3.36.

We used mainly the second version of DRS (DRS2) [76] in 2008. However, since we had a few boards of the succession version, DRS3 [77] which were found to posses better timing performance, we used them for the digitizing the DTD-output NIM signals of the timing ϕ-counter. The DRS2 reads input signal in single-ended though the active splitter outputs differential signal. On the other hand, the DRS3 reads differentially. The dynamic range of DRS2 is from 0 to ~0.5 V, limited by the non-linear response. To meet this range, the PMT signal is inverted at the active splitter and the resistive attenuation is mounted on the mezzanine board. We put the attenuator of factor 1.5 for the readout of the gamma-ray detector. Together with the effect of single-end readout, the signal is attenuated by a factor 3.

We set the sampling frequency to 1.6 GHz for the outputs from the gamma-ray detector and the timing ϕ-counter. The total window of 640 ns works as a pipe-line for the trigger latency and eliminates any delay cables. For the drift chambers, 0.5 GHz was adopted since high time resolution is not necessary. Additional trigger delay of 500 ns was artificially inserted to compensate the signal delay due to the drift time.

3.3.5 Data Size

Since waveform of each channel is recorded, the data size becomes huge. If no data reduction is applied, the maximum data size is over 5 MB/event. Such large-size data

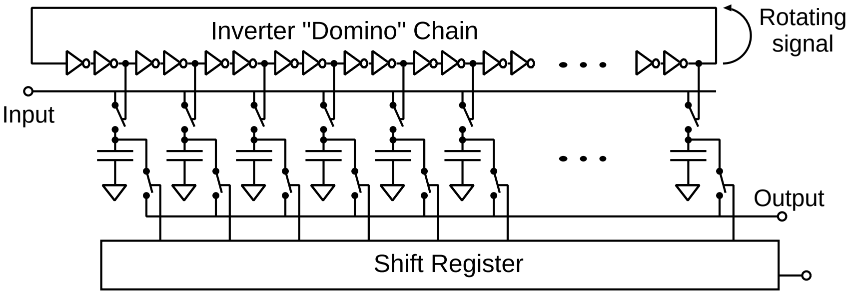


Figure 3.35: Schematic of DRS principle. The inverter chain generates the sampling frequency (Domino wave) and the waveform is sampled in the series of sampling cells (switched-capacitor array). Using the shift register, the waveform stored in the cells is read out at slower speed.

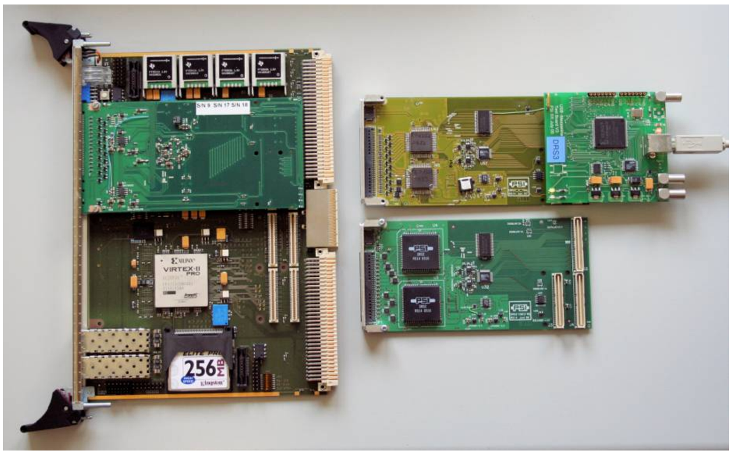


Figure 3.36: Picture of DRS chips and boards. Right top boards is a DRS3 mezzanine board with a USB adapter board. Right bottom board is a DRS2 mezzanine board. Each mezzanine board mounts two DRS chips. Left board is a VME board which mounts two mezzanine boards. In total, 32 channels can be digitized by a VME board.

cause several practical problems such as I/O speed and space of data storage. Thus, some data reductions are applied. For the raw DRS waveform, the following algorithms are implemented for each sub-detector:

- zero-suppression (DCH, TIC)
- region-of-interest (TIC)
- re-binning (DCH, LXe)

Those reductions are applied in online frontend. Then in offline, files are compressed with ZIP by factor 2. The typical data size is about 1.5 MB/event. In 2008, we took 31 TB data.

3.4 Calibration Apparatus

3.4.1 π^- Beam and Hydrogen Target

To calibrate the gamma-ray detector with high-energy gamma rays around and beyond 50 MeV, we use gamma rays from decays of neutral pion ($\pi^0 \to \gamma\gamma$). A neutral pion is produced by a charge exchange reaction of a negative pion ($\pi^- \mathrm{p} \to \pi^0 \mathrm{n}$). It has a momentum of 28 MeV/c in laboratory frame and decays immediately into two gamma rays with almost 100 % branching ratio. The spectrum of the gamma rays is continuous between 54.9 and 82.9 MeV. However, because of the decay kinematics, the energies and opening angle of the two gamma rays have strong correlation as illustrated in Figure 3.37. We can get almost monochromatic gamma rays at 54.9 and 82.9 MeV by selecting events with almost back-to-back gamma rays. For example, if we require the opening angle larger than 175° (170°), then the FWHM of the energy distribution becomes 0.5 % (2.3 %). In addition, a 129 MeV gamma-line from the radiative capture reaction of negative pion ($\pi^- \mathrm{p} \to \gamma \mathrm{n}$) is available. The Dalitz decay of neutral pion ($\pi^0 \to \gamma \mathrm{e}^+ \mathrm{e}^-$) is also an interesting process. We can study the time synchronization and the time resolution of the detector with a similar event topology to the $\mu^+ \to \mathrm{e}^+ \gamma$ decay.

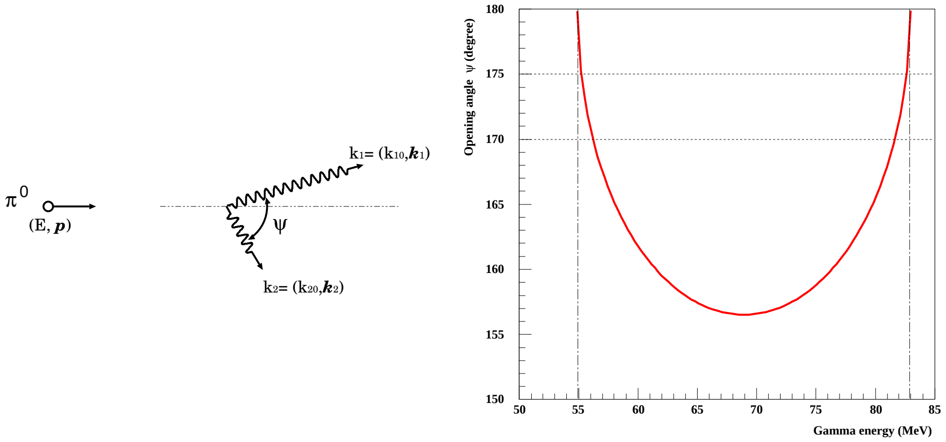


Figure 3.37: π^0 decay kinematics. The plot shows correlation between two gammas' energies and their opening angle.

The πE5 channel can provide also pion beam. Because of the limitation of the BTS current, momentum of the pion beam is set to 70.5 MeV/c, which is optimal within the available momentum to separate pion from contaminants (beam electrons and cloud muons). The spot size was measured to be $\sigma_x \sim 8.5$ and $\sigma_y \sim 7.5$ mm. The pion intensity at the center of COBRA is calculated to be $\sim$ 1.5 MHz at the proton current of 1.8 mA. We also prepared several settings with reduced beam intensity.

We use liquid hydrogen (LH$_2$) as a target of the charge exchange reaction. Hydrogen is liquefied in a cylindrical cell of 50 mm diameter, 75 mm length whose beam entrance side is made of a thin 135 μm Mylar window. A picture of the cell is shown in Figure 3.38. During the normal operation, the cell is filled with ~150 cc liquid hydrogen and kept lower than 20 K by a flow of liquid helium.

It takes ~5 days to set the beam and prepare the target.

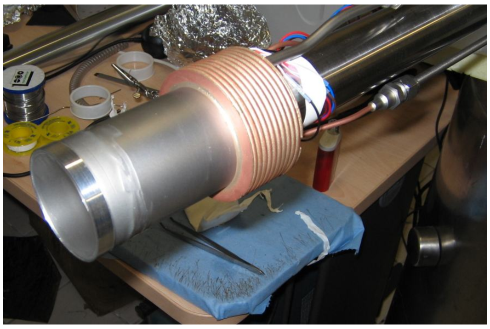


Figure 3.38: Picture of LH_2 target cell.

3.4.2 NaI Detector

To detect one of the gamma rays from π^0 decay, a NaI detector is placed at the opposite side of the LXe gamma-ray detector. The NaI detector consists of nine crystals of NaI ($62.5 \times 62.5 \times 300.5$ mm^3) each of which is viewed by an APD (10×10 mm^2, Hamamatsu S8664-1010), two plastic scintillation counters ($70 \times 70 \times 7$ mm^3) each of which is viewed by two fine-mesh PMTs (Hamamatsu H6152-70), and a lead plate (5 mm thickness) in front of the plastic counters as shown in Figure 3.39. This detector is mounted on a movable stage shown in Figure 3.40. We can obtain the back-to-back gamma rays at arbitrary regions of the LXe detector by moving the position of the NaI detector.

Figure 3.39: Picture of the NaI detector. It consists of a lead converter, two plastic scintillators, and nine NaI crystals.

Figure 3.40: Movable stage of the NaI detector to be placed at the opposite side of the LXe gamma-ray detector.

3.4.3 Cockcroft-Walton Proton Accelerator

The reaction ${}^{7}_{3}\mathrm{Li}(\mathrm{p},\gamma){}^{8}_{4}\mathrm{Be}$ produces a 17.6 MeV gamma-line [78]. This reaction is excitable with very low-energy protons, resonant at $E_p = 440$ keV with a resonance-width $\Gamma \approx 15$ keV. It is almost unique in providing high-energy gamma rays with a large peak cross section ($\sigma_{\mathrm{peak}} \approx 5$ mb). Therefore, it is very useful to monitor and calibrate the gamma-ray detector although the energy is three times smaller than that of gamma from $\mu^+ \to \mathrm{e}^+\gamma$ decay. Actually, it was used in a previous $\mu^+ \to \mathrm{e}^+\gamma$ search experiment [9]. In addition, the reaction ${}^{11}_{5}\mathrm{B}(\mathrm{p},\gamma){}^{12}_{6}\mathrm{C}$ excitable by protons, resonant at $E_p = 163$ keV is also interesting. It produces gamma-lines at 16.1, 11.7, 6.5, and 4.4 MeV. Since the 11.7 and 4.4 MeV gammas are produced in time coincidence with no angular correlation, they can be used to calibrate timing between the gamma-ray detector and the timing counters.

We use a dedicated 1 MeV Cockcroft-Walton (CW) accelerator [79] (Figure 3.41) to produce protons exciting those reactions. Table.3.5 summarizes the properties of the accelerator. The accelerator is placed at the downstream side of the detector system in a separate area. The proton beam goes to the center of COBRA in the opposite direction of the muon beam. The proton beam intensity is 10^{12}/sec. We use a lithium tetraborate ($\mathrm{Li_2B_4O_7}$) crystal disk target [80] which allows us to get the Li- and the B-lines simultaneously. The rate of 17.6 MeV line is $\approx 10^6$/sec, isotropically. Part of the proton beamline is made of bellows. This extendable beam pipe is remotely controlled and inserted into the COBRA volume to set the nuclear target at the center of COBRA with removing the muon stopping target to the parking place. Owing to this target exchange system, we can move from normal muon mode to the CW mode in ~20 minutes. The details of MEG CW accelerator and its setup, operation, and calibration methods are described in [81].

Table 3.5: Properties of MEG CW accelerator.

	Nominal	Measured at PSI
Terminal energy range (keV)	$300 - 900$	$200 - 1100$
Energy ripple(RMS eV)	< 500	< 50
Angular divergence (mrad×mrad)	5×5	$\sim 4 \times 4$
Spot size at 3 m (cm×cm)	$< 3 \times 3$	< 1
Energy setting reproducibility (%)	0.1	OK
Energy stability (FWHM %)	0.1	OK
Range of current (μA)	$1 - 100$	$0.1 - 135$
Current stability (%)	3	OK
Current reproducibility (%)	10	OK
Duty cycle (%)	100	OK
Start-up time (min)	< 20	< 15

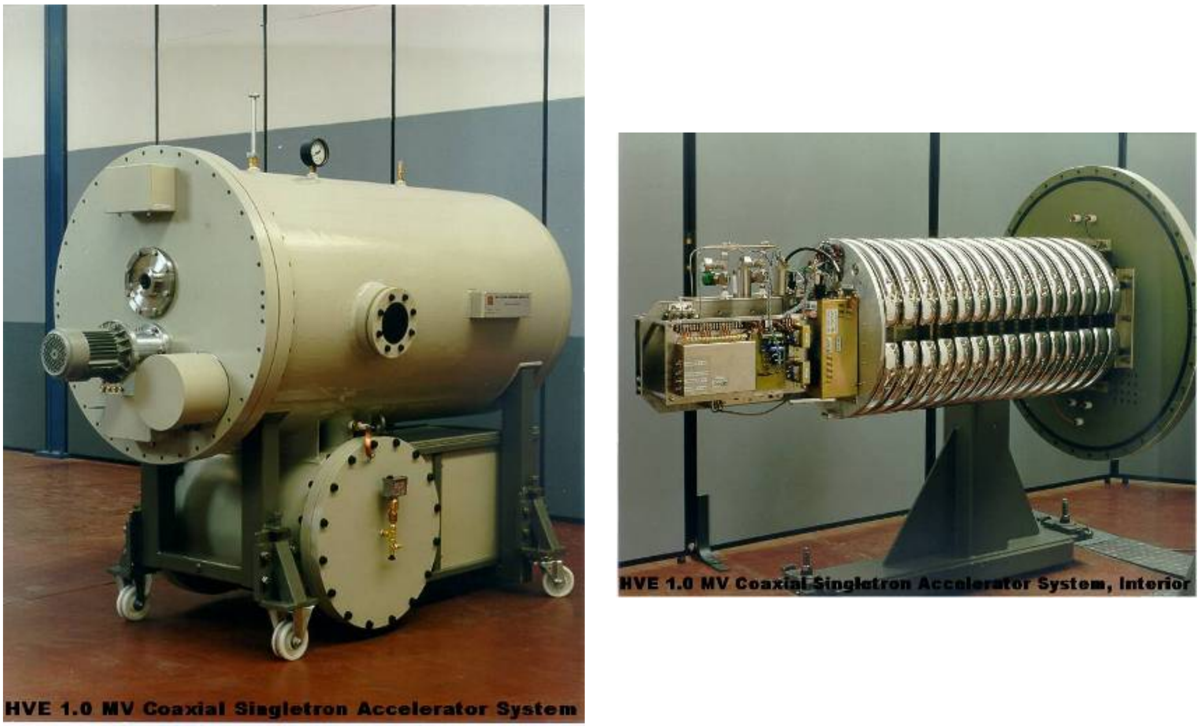


Figure 3.41: 1 MeV Cockcroft-Walton proton accelerator.

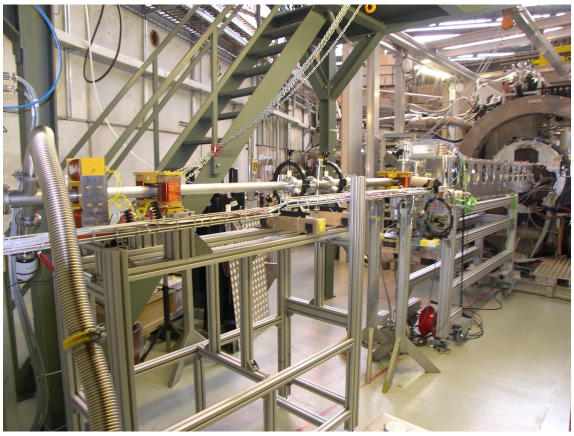


Figure 3.42: Proton beamline from the CW accelerator to the detector center. The CW accelerator is placed at the downstream of the πE5 area.

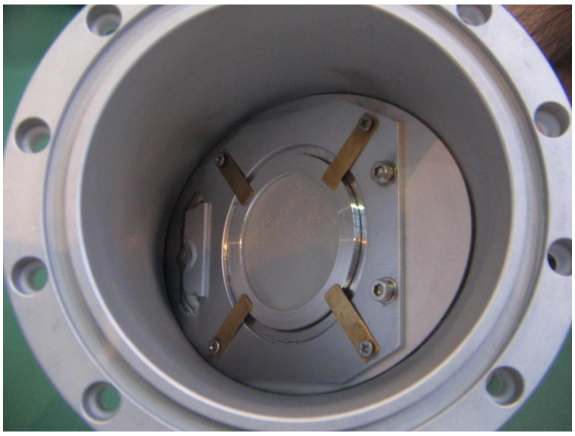


Figure 3.43: $Li_2B_4O_7$ target for the proton-nucleus reactions. The target is mounted at the edge of the beamline and inserted to the center of the COBRA magnet with the bellow system.

3.5 Simulation and Analysis Software

The MEG software is composed of three packages; MEGMC, MEGBartender, and MEGAnalyzer. The structure is shown in Figure 3.44. MEGMC and MEGBartender generate simulation data and MEGAnalyzer analyzes data from both experiment and simulation.

A detailed simulation of the full apparatus was developed and used throughout the experiment, from the design and optimization of all sub-systems to calculation of acceptances and efficiencies. The following sections describe a series of simulation processes.

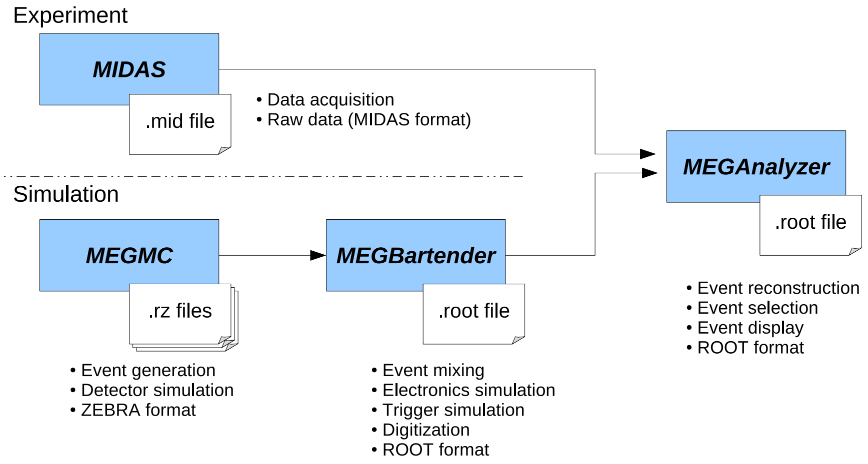


Figure 3.44: Structure of MEG software.

3.5.1 Event Generation

MEGMC is a GEANT3.21 [82] based Monte Calro simulation (MC). It processes the event generation and the detector simulation.

A muon decay on target can be generated both for normal SM decays and $\mu^+ \to e^+\gamma$ decay. It is also possible to generate a muon in beamline and let it propagate and decay. In addition to the event from muon decay, we prepared generations of several calibration events. Table 3.6 lists the event types implemented in MEGMC. For each event type, users can configure the kinematic conditions such as the phase space of the decay, momentum ranges of particles.

3.5.2 Detector Geometry and Response

All apparatus components are described in MEGMC. Detailed geometry and materials are implemented in the standard GEANT3 manner.

The transport of particles and their interactions with material are simulated by the GEANT-tracking. All relevant electromagnetic processes (pair production, bremsstrahlung, Compton scattering, photo-electric effect, ionization losses, multiple scattering and positron annihilation) are included, and the secondary particles are traced down to 10 keV. When

Table 3.6: List of event type implemented in the MC simulation.

$\mu^+ \to e^+\gamma$ signal event
Radiative decay (RD) event
Michel decay event
μ^+ beam
e^+ at 52.8 MeV
e^+ from flat spectrum
γ at 52.8 MeV
γ from RD
γ from e^+ annihilation-in-flight
γ from flat spectrum
Cosmic-ray μ
α from ^{241}Am source on wire
p^+ from CW accelerator + LiF target
p^+ from CW accelerator + B target
$\pi^0 \to \gamma\gamma$ decay

particles interact with the active volume, the hit information such as space-time and energy deposit are recorded. The response of each sub-detector is then precisely simulated.

Simulation of drift cell The precise isochrone map of the drift-chamber cell is calculated with GARFIELD [83] which is widely used to simulate wire chambers in high-energy experiments. It provides the field map, drift lines, drift time tables and arrival time distributions of electrons and ions. The details of energy loss of fast charged particles in gases are computed by the HEED [84] program. The electron transport parameters are provided by the MAGBOLTZ [85] program.

Scintillation photons in timing counter The propagation of scintillation photons in scintillation bar is calculated analytically taking into account attenuation in the bar and reflection on the surface. The photo-electric effect of PMT photo-cathode is simulated in accordance with the Poisson statistics.

Scintillation photons in liquid xenon The transportation of scintillation photons in LXe is simulated with our original code. The scintillation photons are isotropically generated at each point of interaction with LXe in accordance with the deposit energy and scintillation efficiency. Then every photon is traced. During the tracking, several optical processes are simulated such as absorption, Rayleigh scattering, reflection, transmittance of PMT window and photo-electric effect on photo-cathode.

3.5.3 Event Mixing

Simulating actual experiment requires to overlay several activities in a event. MEGBartender, which is the other simulation package developed by the MEG collaboration from scratch, does this. It can read multiple events generated by MEGMC and mix them to simulate accidental overlap. It is also possible to mix different type of events in different output files of MEGMC. Events are randomly or sequentially picked up from the input files and placed randomly in time in accordance with the Poisson distribution.

3.5.4 Electronics Simulation

Another task of MEGBartender is to simulate electronics and form digitized data with the same format as experimental data, namely DRS waveform data. Signals from the detector active volume pass through various devices such as PMT, wire/pad, cables, connectors, splitter, and finally reach DRS. It is difficult to simulate each element individually. However, if those devices are linear systems, we can regard the entire system as a device. Only necessary information is the impulse response of the system.

For the drift chamber, we obtained the impulse response using the signal of 5.4 keV soft X-ray from ^{54}Mn source. The soft X-ray usually generates a single-electron avalanche. Signal from those events at the DRS represents the impulse response of the system. For the gamma-ray detector, we used MC simulation for the input and the average waveform of real signal for the output. We can get the impulse response by deconvolution them. Once we obtained the impulse response, we can synthesize waveform by convolving hit information from MEGMC with the impulse response. We use a Fast Fourier Transform technique to calculate the convolution.

For the timing counter, we synthesize waveform in analytical way since the execution time is much faster than that by the convolution. We apply a series of digital filters with some parameters: a randomization with a Gaussian for the transit time spread of the PMT, a filter with a function $t^2 \exp(-(t/s)^2)$ as the single-electron response, and a RC-filter and reflection for the electronics response. The parameters were adjusted so that the output waveform can reproduce real data.

Before the digitization, electric noise is overlaid to the formed waveform. We generated noise with white and 1/f components, and applied a low-pass filter to simulate the frequency property. Then, we digitize the waveform to simulate DRS both in sampling and quantizing.

3.5.5 Analysis Framework

We analyze all processes, from handling of raw data to the physics analysis, in MEGAnalyzer. It can handle both experimental and simulation data in the same manner. It, and also MEGBartender, are organized by the analysis framework toolkit, ROME [86]. The ROME-generated software is highly modular. The analysis processes and information (such as data, parameters and analysis results) are treated as objects, tasks and folders respectively. A series of calculation codes is split up into tasks, which contain one or several calculation steps. We can exchange tasks arbitrarily as long as they access the same data folders. This modularity makes it very flexible to control the whole structure. Any user can easily add and write their own code with simple knowledge of C++ and ROOT [87].

Furthermore, the framework has features such as connection to DAQ systems, database access, socket connection, and Graphical User Interface (GUI) extension. Therefore, MEGAnalyzer also works as an online monitor and an event display.

Chapter 4

Run

4.1 Run 2008

We started experiment in 2008 with a conditioning and calibrating Michel run in June just after completing detector installation. From end of July to beginning of September, we conducted a π^0 calibration run. After a small period of trigger setup and background study, finally we started a physics run on 12 September. It continued until 17 December. Those data are the initial three month data of the MEG experiment and full data set of the analysis in this thesis. Run 2008 finished with another small π^0 calibration run and it enters in annual accelerator-showdown period of PSI.

4.1.1 Physics Run

The normal run period consisted of the MEG runs taking physics data with the MEG trigger, calibration runs (three times per week), dedicate RD runs (24 hours per week), and programmed beam-maintenance periods (three days per three weeks). In the MEG runs, we mixed 11 different types of trigger for the normalization, calibration, and monitoring (see Table 3.4). We took data with the normal mode of the muon beam ($3 \times 10^7 \mu^+$/sec, Sec.3.1.5). Figure 4.1 shows the integrated amount of physics data as a function of date. The livetime was 3.3 Msec with ∼85 % efficiency. In total, 9.5×10^{13} muons stopped on the target.

4.1.2 Dedicated Radiative Decay Run

The RD events, which were introduced as one of the background sources before, can be used as a calibration source of relative timing between positron and gamma. Moreover, it is very important to see the apparatus can detect RD events correctly to demonstrate the capability of observing $\mu^+ \to e^+\gamma$ events. To clearly observe RD events, we conducted data taking with reduced-intensity beam by a factor 25 (the ultra-low mode). The reduced-intensity beam suppresses the accidental background and makes signal-to-noise better. We used the RD triggers (Trigger#4,5), which relax the condition of angular correlation between positron and gamma. The energy threshold for gamma was also reduced to around 25 MeV. The energy spectrum of gamma from RD and angular correlation are studied as well as relative time. These special runs were carried out intermittently during run 2008 for one day per week. They are indicated in Figure 4.1 as green arrows.

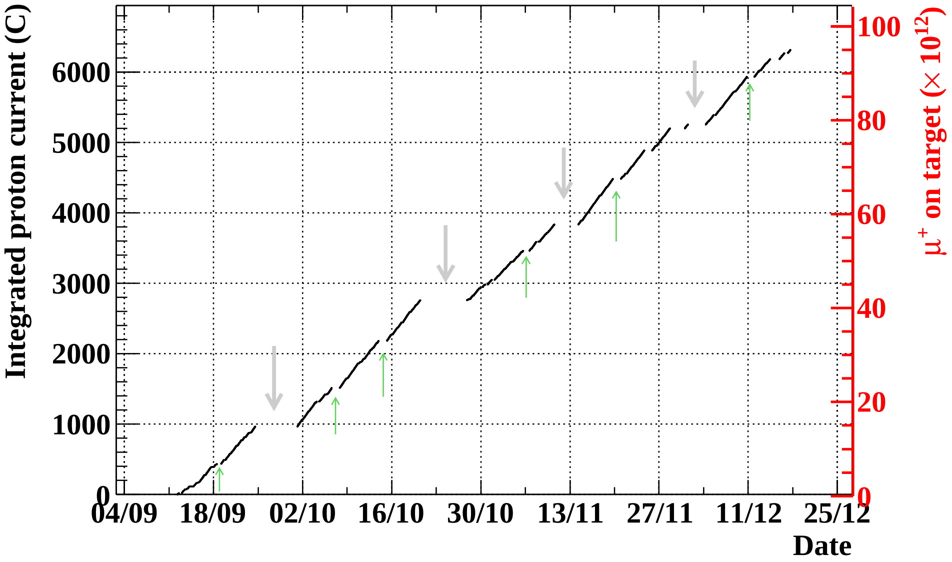


Figure 4.1: Integrated amount of physics data taken in 2008. In total, data of 6300 C proton current was acquired, which corresponds to 9.5×10^{13} μ^+ on target. Gray arrows show beam shutdown due to programmed beam maintenance, and green arrows show radiative decay run periods.

4.1.3 π^0 Calibration Run

We conducted two π^0 runs, one at the beginning (August) and the other at the end (December) of the data taking period. The purposes of the π^0 runs are the following:

- to evaluate energy, time, and position resolutions of the gamma-ray detector, and to get the responses,
- to calibrate energy scale of the gamma-ray detector,
- to study relative timing between positron and gamma and calibrate its t_0,
- to measure detection efficiency of the gamma-ray detector.

To achieve these purpose, we carried out several kinds of data taking.

$\pi^0 \to \gamma\gamma$ patch scan To evaluate energy and time response at different positions of the detector, a full scan of detector inner face was performed with $\pi^0 \to \gamma\gamma$ reaction. We divided the inner face into 24 parts and prepared patches of trigger for each part. Each of the patches contains 3×3 PMTs. The configuration of the patches is shown in Figure 4.2. Depending on the trigger patch, we moved the NaI detector to the opposite position of the patch to collect back-to-back events. The trigger is fired by a coincidence between LXe patch and reference counter (Trigger#6) or NaI (Trigger#7). We use the former for the time performance and the latter for the energy performance. The fraction of events is about 3:7. We collected 200 ~ 300k events for each patch.

$\pi^0 \to \gamma\gamma$ **reference run** For the reference of the patch scan, we everyday took $\pi^0 \to \gamma\gamma$ data at the same position (patch#8). We use them for corrections of light yield during the scan. We determine the energy scale of the gamma-ray detector with those run.

Pb brick run To measure the position resolution, we took special data with lead bricks mounted just in front of the gamma-ray detector entrance window. The lead bricks are 1.8 cm thick and have some slits of 1 cm width. Shadow of the bricks can be used to estimate the position resolution. Figure 4.3 shows the configuration of the bricks installation. The design and picture of the lead bricks are shown in Figure 4.4. We used the LXe-self trigger (Trigger #9) and collected $\sim$ 1M events for each brick position.

$\pi^0 \to \gamma e^+ e^-$ **Dalitz decay** To study relative time between positron gamma, we collected π^0 Dalitz decay data. We used the RD trigger (Trigger#4) with gamma energy window around 55 MeV. In total, we collected 500k events.

NaI single We collected NaI-self trigger data to evaluate the detection efficiency of gamma-ray detector.

4.1.4 Normal Calibration Runs

We took a full set of calibration data three times per week. It takes about three hours to take one set. A set of the calibration consists of the following runs,

- Pedestal run to check electronics and evaluate noise,
- LED run to calibrate the PMT gains; it was taken everyday with beam on and off conditions,
- Alpha run to calibrate the PMT quantum efficiencies and monitor LXe (30k events),
- Cosmic-ray run to monitor LXe (50k events),
- CW-Li run to monitor and calibrate LXe light yield and to correct non-uniformity of the gamma-ray detector (30k events),
- CW-B run to calibrate the timing counter and to monitor relative timing between the gamma-ray detector and the timing counters (10k events).

4.2 Run Condition

4.2.1 Discharge Problem on Drift Chamber

We suffered from a severe problem on drift-chamber operation in 2008. Several chambers frequently discharged. This problem is thought to be posed by helium gas permeated into HV lines. As described in Sec.3.2.1, the drift chamber is placed in the helium atmosphere. The outside of the chamber faced with pure-helium. To avoid connection of HV-to-GND lines via helium gas, we molded the HV lines with glue as a dielectric. However, slowly

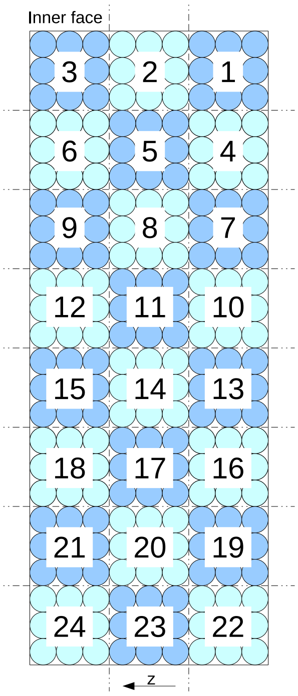


Figure 4.2: Configuration of patch and definition of patch number. The patch#8 was used as a reference.

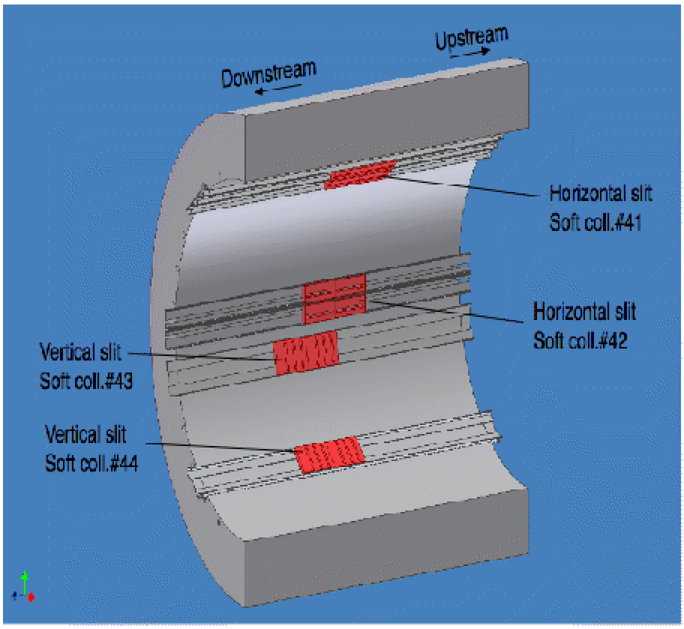
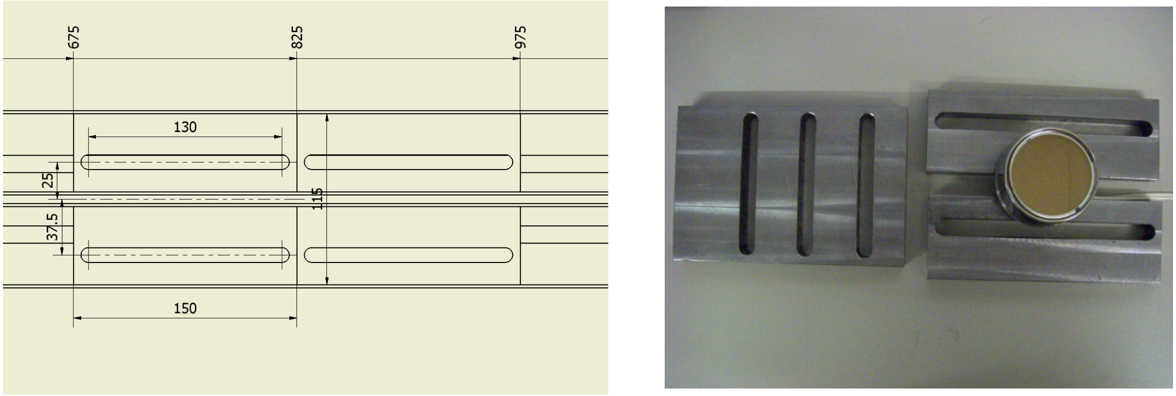


Figure 4.3: Configuration of lead bricks installation.



(a) Design of a lead brick.

(b) Picture of lead bricks.

Figure 4.4: Lead bricks for a study of position resolution.

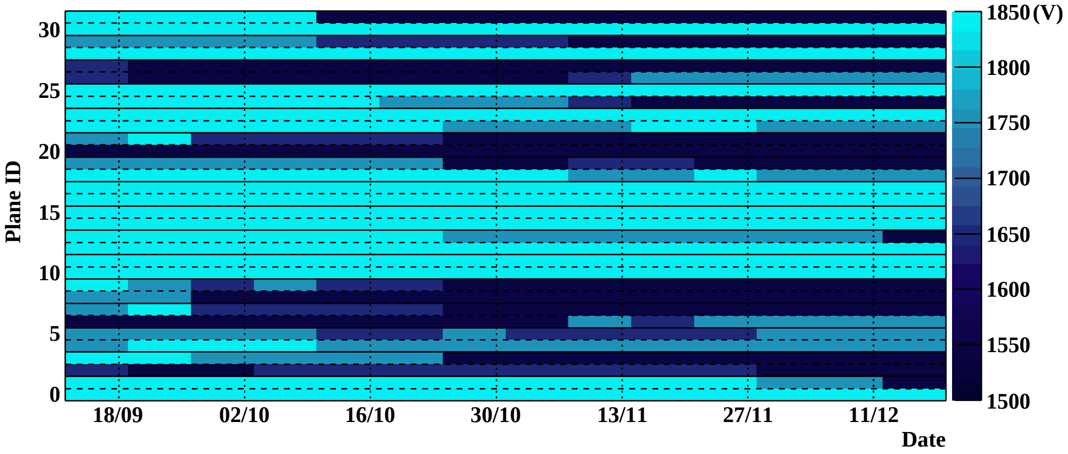


Figure 4.5: History of HVs applied to the drift chambers. HV is applied independently to each plane (in total 32 planes). The dark regions indicate planes to which we could not apply HV because of the discharge problem.

helium permeated into the HV lines. Since it took a time for helium to permeate, the chambers suffering from the discharge gradually increased. Finally, out of 32 planes, 18 planes were operational, and only 12 planes worked at the nominal voltage. Because of this problem, the drift chamber system did not work at the nominal performance. In particular, the efficiency was severely restricted in run 2008.

4.2.2 Helium Concentration inside the Magnet

To relax the potential of the discharge, we introduced small amount of air in the helium atmosphere as a temporary treatment. On average, 6 % air was doped. The beam tuning was done before the air-doping, and not corrected the effect of this additional air.

4.2.3 Problems on Timing-Counter APD

There were some problems on the z-counter read out. In the digital output of APD signal, 3/16 parts were out of control because of a hardware problem on the on-board electronics in addition to the several hot channels. On the analog output to be used in trigger, the noise level was too high to be used in the trigger decision. Therefore, we did not use the z-counter neither in trigger nor in the offline analysis. Instead, in trigger we reconstructed z-coordinate of hit position using the charge ratio of ϕ-counter PMTs, while in offline we do using the time difference between them. These alternative solutions have worse resolutions. In addition, it was found later that there was a bias in the trigger estimation of positron-emission angle from the z hit position reconstructed in this way. Those effects resulted in low trigger efficiency.

4.2.4 Purity and Light Yield of LXe

Since when we constructed the gamma-ray detector and first liquefied xenon in 2007, we have progressively improved the purity of LXe by the purification. Also during the data taking period, we continued purification, and the light yield of LXe scintillation was improving.

At the beginning of the physics run, the absorption length was estimated to be sufficiently long, and it seemed the improvement of observed light amount got saturated by the liquid-phase purification. Therefore, we decided to stop the purification and take data with stable condition, even though the light yield was estimated to be lower than the expected value calculated using the data obtained with the prototype detector.

We took data with this stable condition until the end of October. However, we found a slow decrease of the light yield after a while. It is thought due to small leak in liquid nitrogen cooling pipe in low temperature because there was correlation between the decrease of the light yield and the operation of liquid nitrogen. Or possible out-gas from material in the detector. From when we discovered this correlation, we quited to use liquid nitrogen pipe inside the detector for an additional cooling. To recover this small decrease, we performed a liquid-phase purification during a short beam-maintenance period. Then we found further improvement of the light yield, and newly found an effect of gas-phase purification which did not show a saturation of improvement. We think that previously the gas-phase purification had not worked efficiently because of a problem on the gas-phase purification pump, and after we fixed the pump it started effectively working. The improvement of the light yield during the data taking causes an instability in measurement point of view, but to understand the detector well and to extract a full performance of LXe scintillation, continuous purification was demanded. On the other hand, we had established the monitoring and calibrating methods using the various kinds of periodic calibration data. Therefore, we decided to continue purification and correct the energy scale using the calibration data in offline.

From the end of October to the end of the run, we continuously performed the gas-phase purification. In addition, we performed the liquid-phase in the periodic short beam-maintenance periods. The light yield kept increasing until the end of the run. Roughly, we observed 30 % of increase during the run. Yet at the end, it was still lower than the expectation by factor $\sim$1.5.[1] Not only the light yield but also the pulse shape were changed by the purification. This change caused variation of the time measurement. Monitoring and calibrating the light yield is described in Sec.6.4.3 and those of the time reference is in Sec.6.5.

4.2.5 Bad Channels

The timing ϕ-counter had a bad channel for the first half data. All of the PMTs were working fine. There was, however, a broken channel of the DTD, and it was exchanged at the end of October. After that, all channels were working.

There were three dead channels of the gamma-ray detector PMTs from the beginning. One lost connection in signal line, and the other two had short in HV line. A channel showed strange pulse shape and another channel showed unstable PMT gain. In addition, two PMTs showed severe decrease of gains during the run. The point is that all the PMTs on inner face worked. We did not lose any acceptance.

---

[1] We performed further purification during the long accelerator-shutdown period. At the beginning of 2009 operation, the light yield became the same level of the expectation.

Chapter 5

Event Reconstruction

5.1 Overview

In this chapter, we describe the reconstruction of each kinematic variable focusing on the methods and algorithms. The reconstruction of MEG starts from the analysis of waveform to extract physical quantities. Next, kinematic variables of each particle are reconstructed. In this stage, we treat positron and gamma separately. Finally, we reconstruct the relative angle and time between positron and gamma by combining each reconstruction. Figure 5.1 overviews the flow of the reconstruction. To extract full performance of those methods, precise calibrations are necessary in advance. They are described in the next chapter. The performance of the reconstruction is evaluated in Chapter 7 using obtained data.

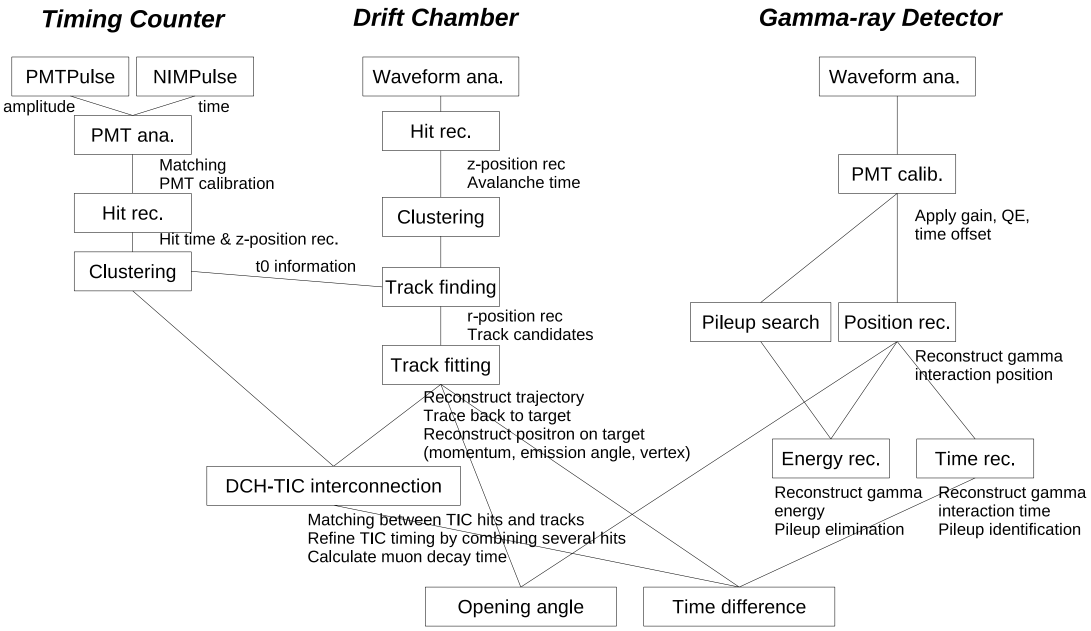


Figure 5.1: Flow chart of reconstruction.

5.2 Waveform Analysis

5.2.1 Waveform Analysis of Drift Chamber

We get six waveforms from a drift cell: two from both ends of an anode wire and four from cathode pads (inner and outer pads with two vernier-pattern pads). Figure 5.2 shows an example of six waveforms associated with a drift cell. We collectively treat those six. First, using the waveforms of anode wire, we search for pulses. A simple peak-search method cannot be used since the waveform of drift chamber has a multi-pulse structure of ionization clusters in single particle pass. Thus a pulse and its width is determined as the following procedure. A pulse is identified by the maximum peak over a given threshold. Then the width of the pulse is determined by pursuing the pulse in both directions until it goes down under another threshold. This pulse range is masked, and the pulse search is repeated.

Time extraction Time of each pulse is determined by a single-threshold crossing time.

Charge extraction The integration window of each pulse is determined by the anode waveform and common to the six waveforms. The window width is set to 50 ns, which was tuned to optimize the signal-to-noise ratio.

5.2.2 Waveform Analysis of Timing Counter

As described in Sec.3.3.1, the output of a ϕ-counter PMT is digitized in two ways. A PMT output is divided passively into three outputs with fraction of 1:1:8. One of the 10%-fraction outputs is digitized by DRS directly (PMT pulse). Together with the attenuation factor two on the DRS board, the digitized signal totally becomes 5 % of the original PMT output. The 80%-fraction one goes to the double-threshold discriminator (DTD) and its output (NIM pulse) is digitized. We use DRS3 for the NIM-pulse digitization as already mentioned in Sec.3.3.1. The other output of the splitter goes to a current monitor. Figure 5.3 shows typical waveforms associated with a PMT.

Time extraction We use the NIM pulse to get a pulse time. We prepared in advance a template waveform by averaging many pulses channel-by-channel. It is used as a fitting function of the NIM pulse. The free parameters of the fitting are baseline level and leading-edge time of the pulse.

Charge and amplitude A charge and an amplitude corresponding to the time-measured NIM pulse is estimated by using the PMT pulse. The charge is estimated by integrating pulse over 30 ns. The amplitude is measured as the voltage difference between the estimated baseline and the peak voltage of the pulse. We use the charge to measure the gains to equalize them. We use the amplitude to correct the time-walk effect of the NIM-pulse time.

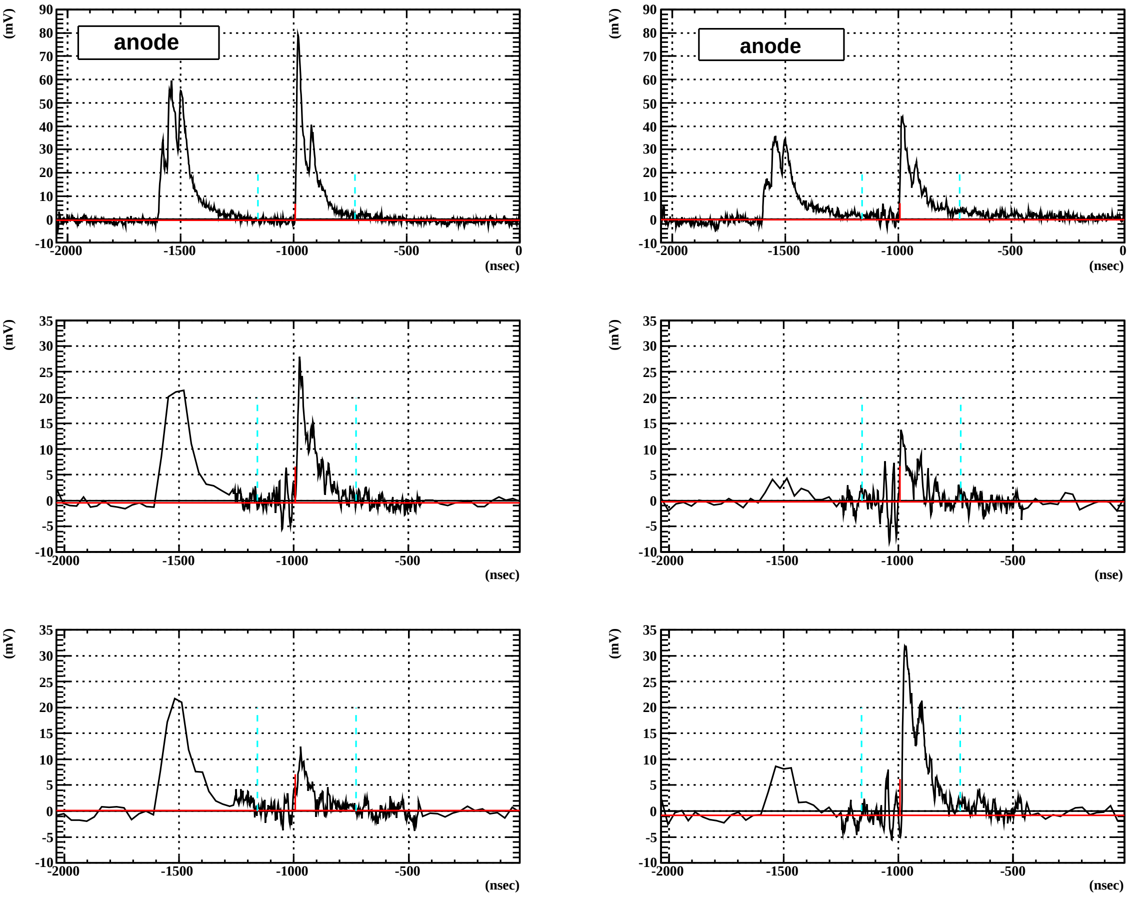
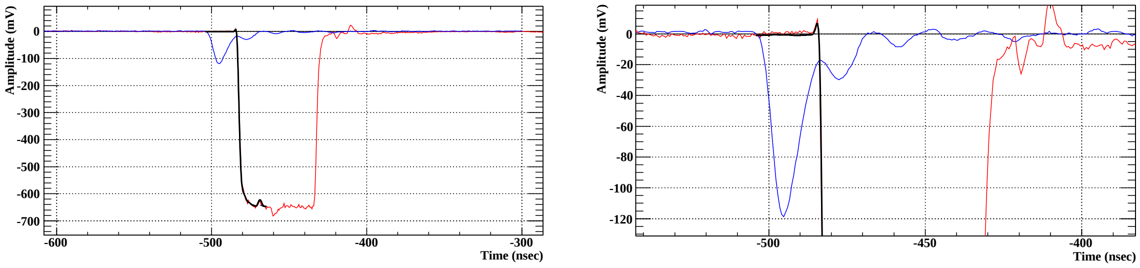


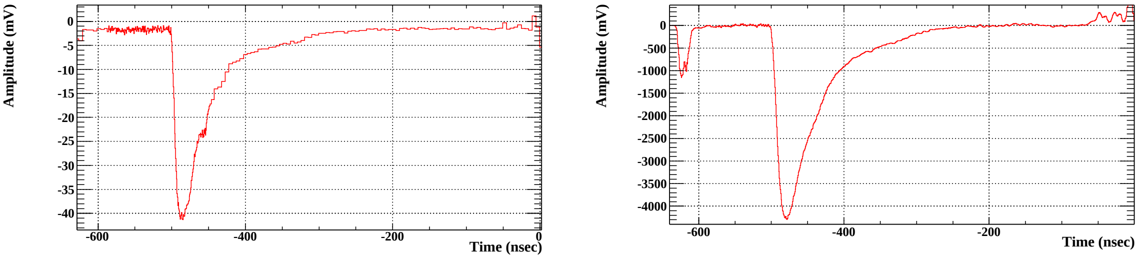
Figure 5.2: Typical waveforms of drift chamber. Those six belong to a drift cell. The red horizontal line shows the calculated baseline, and vertical one shows the pulse timing. The region between light blue dashed lines shows the expected region where triggered signal appear. Two sides of the cathode waveforms are re-binned for data reduction.



(a) Waveforms of a PMT

(b) Close-up view

Figure 5.3: Typical waveforms from a TIC PMT. Blue lines show waveform of attenuated PMT signal (PMT pulse). Red lines show output pulse of double-threshold discriminator (NIM pulse) digitized by DRS3. A template waveform fitted to the NIM pulse is overlaid as black lines. The NIM pulses come with a delay of $\sim$ 20 ns.



(a) Waveform of a PMT (b) Sum waveform over all PMTs

Figure 5.4: Typical waveform of LXe interacting with a gamma ray. (a) shows waveform of a PMT. (b) shows sum of waveform over all PMTs. The peak at the beginning of the window is related to the stop signal.

5.2.3 Waveform Analysis of Gamma-ray Detector

Typical waveform from LXe interacted with a gamma ray is shown in Figure 5.4. For the data reduction, every eight sampling points are averaged and recorded as a point except for the interesting regions; a leading-edge part and a baseline part just in front of the leading edge. The baseline is estimated by averaging the points in the baseline region event-by-event for each channel.

Time extraction For the time extraction, we do not apply any digital filters to retain maximum information of pulse shape and band width. We extract the time of a pulse with the digital-constant-fraction method. It determines the pulse time as a time at which the signal reaches a given fraction (here 30 %) of the full pulse height (Figure 5.5). By this method, we can determine the pulse time independently of the amplitude (no time-walk effect). The pulse height is estimated by using the charge of the pulse not by the direct measurement of the amplitude so that the noise and the photoelectron statistical fluctuation in the pulse shape are eliminated. The measured pulse charge is converted to a voltage with a factor calculated by assuming the fixed pulse shape. The level crossing time is extracted with linear interpolation of the adjacent two points.

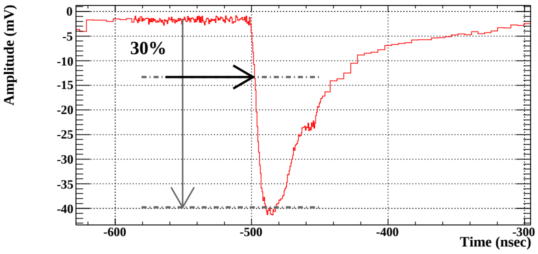


Figure 5.5: Time extraction by the digital-constant-fraction method. The fraction is set to 30% of the full amplitude.

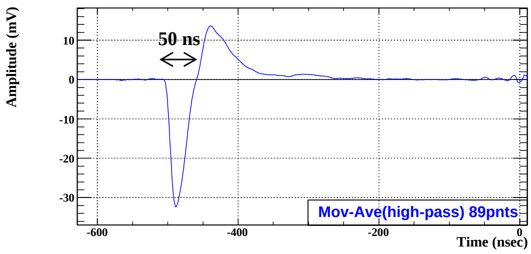


Figure 5.6: LXe waveform shaped with a high-pass filter.

Charge estimation We estimate a pulse charge, which is later converted to the number of photoelectrons observed by the PMT, by integrating the pulse. Since we have full waveform, digital filters can be used in offline process to maximize the signal-to-noise ratio. We observed some slow-component noise of $\sim$ 1 MHz. To eliminate this kind of noise, we apply a high-pass filter. On the other hand, we do not need filtering the high-frequency components of noise since the integration itself works as a powerful low-pass filter, and charge estimation is not influenced so much. We form a high-pass filter based on the moving-average method, which is usually used in a low-pass filter. The high-pass filter works as subtraction of waveforms between raw waveform and (low-pass) moving-averaged one. The number of averaging points is set to 89, which is corresponding to a cut-off frequency of 11 MHz. It was adjusted with a trade off between gain and noise reduction power. The filtered waveform is shown in Figure 5.6. The gain is about 0.35. The baseline becomes flat and at zero. The pulse shape becomes sharper, resulting in shorter integration width. The integration range is determined by the sum waveform and a common window is used for all individual channels. It is determined as a window of 48 ns width just in front of the zero-crossing time. By applying the filter, we successfully reduce the noise contribution to the energy resolution from 0.6 % to 0.4 % of 52.8 MeV. Figure 5.7 shows reconstructed energy of pedestal events with and without applying the filter. The shorter window also works effectively to reduce pileup events.

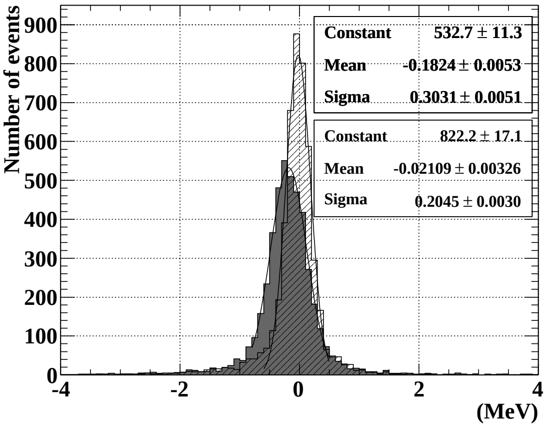
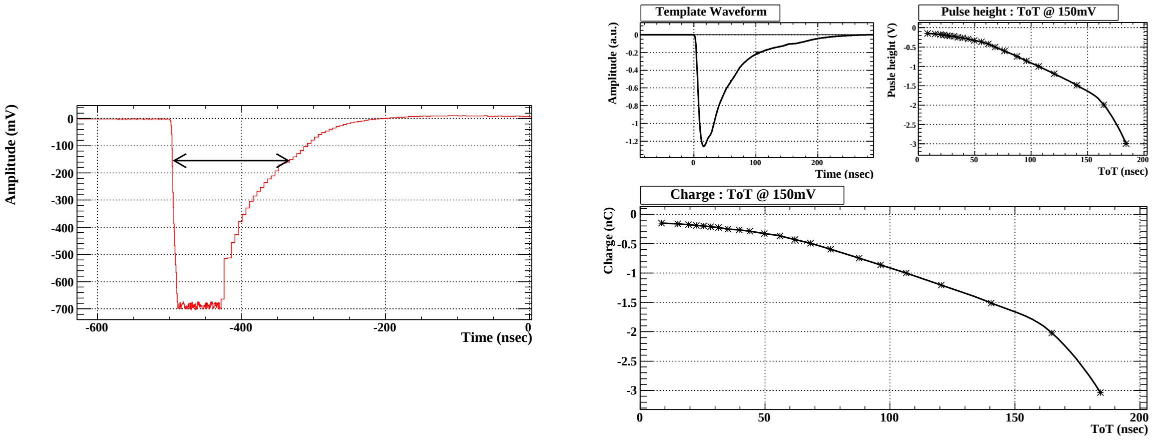


Figure 5.7: Reconstructed energy of pedestal events in normal data taking. The result when we use normal waveform to estimate PMT charges is shown in filled histogram and that with high-pass filtered waveform is shown in hatched one.

When a gamma ray interacts at a very close point to a PMT, the PMT observes very large pulse and the signal gets saturated with a limited dynamic range of the electronics. Figure 5.8(a) shows an example of saturated signal. To recover these saturated channels, we use the time-over-threshold (ToT) method in estimating the charge. The ToT is a duration of the pulse over a given threshold. The threshold was set to 150 mV. We know the average pulse shape of gamma-ray interaction events (shown in Figure 5.8(b) as a “template waveform”). Hence, we can convert the ToT value to the charge. The recovering is important to achieve a high efficiency of gamma-ray detection because the probability of gamma-ray interaction at shallow region is high; about 15 % events interact



(a) An example of saturated channel.

(b) ToT - charge relation.

Figure 5.8: Saturated signal and charge estimation with ToT. When a pulse saturates, the ToT method is used for charge estimation instead of charge integration. Pulse shape of template waveform is used to get the conversion factor from ToT to charge.

within 1 cm and most of them have at least one saturated channel.

5.3 Positron Reconstruction

5.3.1 Hit Reconstruction and Track Finding

Drift Chamber Hit Reconstruction

We measure the z-coordinate of drift-chamber hit by a combination of anode-charge-division and cathode-vernier-pad. At first, we roughly reconstructed it by the ratio of charges measured at both ends of the hit wire. We define the charge-division

$$\epsilon_a = \frac{Q_u - Q_d}{Q_u + Q_d}, \tag{5.1}$$

where $Q_{u(d)}$ is charge on the upstream (downstream) end. With this charge-division, the z is reconstructed by

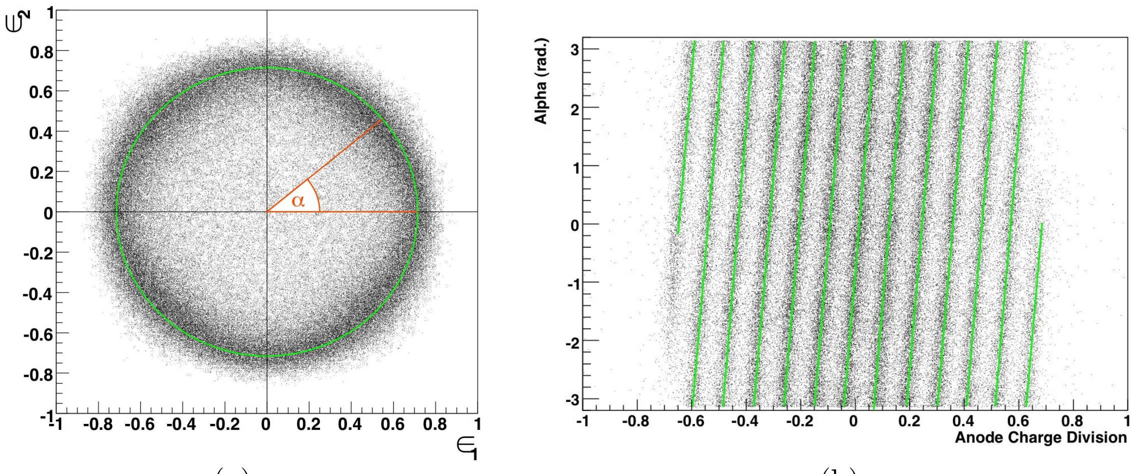
$$z = \left(\frac{Z}{\rho} + \frac{L}{2}\right) \cdot \epsilon_a, \tag{5.2}$$

where Z is input impedance, L is wire length, and ρ is wire resistance per unit length. With this method, we determine the turns of the vernier-pattern period (n).

Then, we reconstruct it more precisely using the vernier pattern. The charge-division for the vernier pads are similarly defined (ϵ_1, ϵ_2) for inner side cathode pad and outer one, respectively. Using the phase of the vernier pattern, $\alpha = \tan^{-1} \epsilon_2/\epsilon_1$, z is reconstructed by

$$z = \frac{l}{2\pi} \cdot \alpha + \mathrm{n} \cdot l, \tag{5.3}$$

where l is the pattern pitch, equal to 5 cm. Figure 5.9 shows the vernier circle and relation between anode charge-division and vernier circle.



(a) (b)

Figure 5.9: z reconstruction of the drift-chamber hit. (a) shows the vernier circle. One turn corresponds to a period of vernier pattern (5 cm). (b) shows relation of the vernier phase with anode-charge-division.

Track Finding

To find tracks, we apply topological pattern recognition. First, using the z and time information, we cluster hits associated with a positron pass within each module. Since the two layers of a module is staggered by a half cell, we can solve the "left-right ambiguity" in most case. Next, we connect the clusters. If we find a "seed" of track with three clusters, then we roughly reconstruct the trajectory with a circle. Using this swim function and hit coordinates, clusters associated with a positron are connected progressively, resulting in a track candidate. During this process, we can refine the left-right ambiguity. We can also get information of incident angle to each cell.

For a found track candidate, we can estimate the time of the track, t_{track}, by using the timing of all hits belonging to the track. Using this drift-chamber self-contained t_{track} and the track direction, the timing-counter hit associated with the track is searched for, and the t_{track} is refined by the timing-counter hit time. The drift time of each hit is given by $t_{wire,i} - t_{track}$, where $t_{wire,i}$ is a time extracted from the i-th wire waveforms.

Given drift time, incident angle, left-right solution, and B-field strength, we can calculate the drift distance uniquely. Since the relation has strong angular dependence, we prepared the time-to-distance functions for different incident angle as shown in Figure 5.10. Using these functions, precise hit coordinate is determined.

5.3.2 Tracking

Track Fitting

Next, a trajectory of the track candidate is precisely analyzed by a track fitting and the state of positron on the target is reconstructed. We use the Kalman filter technique in track fitting. The Kalman filter [88] was originally developed as a linear estimation for the state of a dynamic system from a series of incomplete and noisy measurements. In this 20 years, it has been extensively used for track fitting in high-energy physics (for

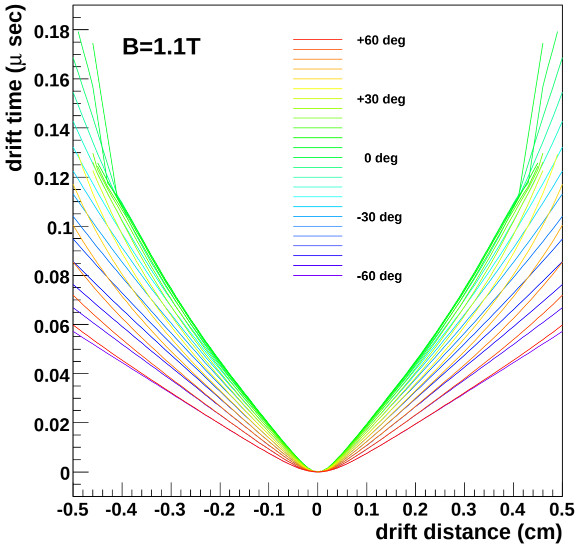


Figure 5.10: Time-to-distance functions for various incident angles. In the practical usage, the functions are implemented as two-dimensional isochrone map for different angles.

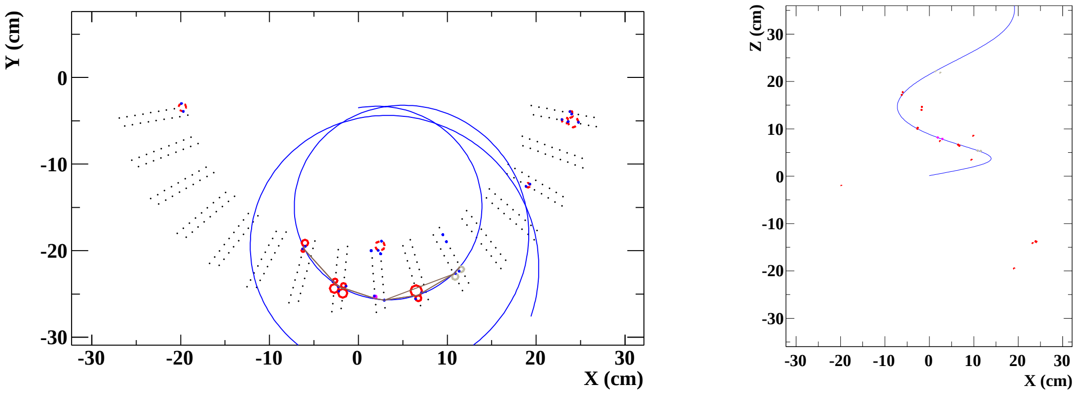
example [89]). It possesses the following features for effective track fitting:

- multiple scattering and energy losses are included in natural way;
- a 3-dimensional trajectory is restored that approximates closely the real one;
- complex tracker geometries and non-uniform magnetic field are handled in a simple way;
- N×N matrix inversion required in the global least squares fit (where N is the total number of measurements) is avoided;
- control of error propagation is provided.

Those features are suitable for the tracking of the MEG positron spectrometer.

The track of a charged particle in a magnetic field at each point can be uniquely described by a vector of five parameters: two for the position, two for the direction and one for the momentum. The trajectory is reconstructed recursively by adding measurements. The operations of track fitting with Kalman filter consist of the following three processes:

- **Prediction** ($\mathbf{x}_k^{k-1}$) is estimation of 'present' k-th point state vector using all 'past' information. The extrapolation takes into account all possible effects that affects the system (such as multiple scattering and energy loss).
- **Filtering** ($\mathbf{x}_k^k$) is estimation of 'present' k-th point state vector with all measurements including k-th measurement. The state vector is updated taking into account, with the appropriate weight, the present measurement and prediction by the previous step.
- **Smoothing** ($\mathbf{x}_k^N$) is backward estimation of state vector with all measurements.



(a) $x-y$ view (b) $x-z$ view

Figure 5.11: An example of track fitting using Kalman filter.

The evolution of the state vectors ($\mathbf{x}$) are described with a system equation.

$$\mathbf{x}_k = \mathbf{f}_{k-1}(\mathbf{x}_{k-1}) + \mathbf{w}_{k-1}, \tag{5.4}$$

where $\mathbf{f}_{k-1}$ is track propagator from 'k-1'-th measurement to 'k'-th measurement; $\mathbf{w}$ represent a random disturbance. The multiple scattering and energy loss are fetched in the system via this term and its covariance matrix. In the presence of a magnetic filed the track propagator is nonlinear. It is approximated by a linear function, for example by the two terms of its Taylor expansion (extended Kalman filter) so that the concept of linear filtering can be applied. The system equation is then given by

$$\mathbf{x}_k \quad = \quad \mathbf{f}_{k-1}(\mathbf{x}_{k-1}^{k-1}) + \mathbf{F}_{k-1}(\mathbf{x}_{k-1} - \mathbf{x}_{k-1}^{k-1}) + \mathbf{w}_{k-1}, \tag{5.5}$$

where $\mathbf{F}$ is the Jacobian matrix of $\mathbf{f}$. The actual measurement ($\mathbf{m}$) is brought in the evolution by converted to the state vector with linear relation of,

$$\mathbf{m}_k = \mathbf{H}_k \mathbf{x}_k + \epsilon_k, \tag{5.6}$$

where matrix $\mathbf{H}$ is called projection matrix; ϵ is the measurement uncertainty.

The accurate reconstruction of trajectory requires a precise track model for the trajectory of the charged particle in a magnetic field, namely the definite description of the track propagator. Because of the specially graded magnetic field, the equation of motion of charged particles in the COBRA field must be solved by numerical integration. We use the fifth order Cash-Karp method [90], which is one of the Runge-Kutta methods and can maneuver the adaptive step-size effectively to approximate the highly graded field sufficiently. A measured field map is incorporated in the system. A correct and efficient description of the multiple scattering and energy loss effects requires the knowledge both of the detector resolutions and of the amount of material traversed by the particle.

In the process of filtering, the local χ_k^2 is calculated as a residual between the prediction and measurement. It can be used as a powerful test of outlier hits. Thus the Kalman filter can correct the result of track finding by adaptive detection of outliers. If the χ_k^2 is larger than a given threshold value, the hit is removed from the track candidate. It is also possible to correct the left-right ambiguity in the process. On the other hand, a global χ^2_{track} which is the sum of χ^2_k can be used as a test of ghost track or the track quality.

In the final smoothing process, the state vector is back-propagated to the muon stopping target plane. An example of reconstructed trajectory is shown in Figure 5.11. The state vector on the target gives the muon vertex point $\vec{x}_\mu = (x_\mu, y_\mu, z_\mu)$, emission angle of positron $\hat{p}_e = (\theta_e, \phi_e)$ and its momentum $|\vec{p}_e|$ (thus its energy E_e) with their covariance matrix. At the same time, the track is extrapolated to the timing counter. The total track length from the target to the timing counter gives the positron time-of-flight. The hit information of the timing counters are not incorporated in the prediction of state vector on the target because the extrapolation from the last measurement in the drift chamber to the first timing-counter hit is so long that the it does not give a useful information.

5.3.3 Time

Timing Counter Hit Reconstruction

We reconstruct a hit time of a ϕ-counter by averaging the two PMTs time after correcting the time-walk effect as a linear function of $1/\sqrt{A_i}$, where A_i is the measured amplitude of i-th PMT. The hit position in z-coordinate is reconstructed by the time difference between the two PMTs.

High momentum ($\sim$ 50 MeV/c) positrons from muon decays on target often penetrate a couple of bars like an example shown in Figure 5.12. For those events, hits associated with a positron track are clustered by their time and z hit position. Then the timing-counter hit time t_{TIC} is given by the time of first bar hit.

DCH-TIC Interconnection and Track Length

To reconstruct the time of muon decay, the time-of-flight of positron t_{eToF} has to be subtracted from the timing-counter hit time. It is measured by the track length of the reconstructed Kalman track extrapolated to the surface plane of the ϕ-counter bar. The connection is schematically explained in Figure 5.12. Matching between them is decided by looking at the residual between track extrapolation and timing-counter hit ($\Delta z_{\text{DCH}-\text{TIC}}$ and $\Delta r_{\text{DCH}-\text{TIC}}$ in z and $r = \sqrt{x^2 + y^2}$ directions respectively) and also the $\chi^2_{\text{DCH}-\text{TIC}}$. With the long projection from the last measurement of the drift chamber to the hit on the timing counter, the projected track has large uncertainty due to the possible scattering. To correct it, we use the value of $|\Delta z_{\text{DCH}-\text{TIC}}|$. We extracted the correction function from the π^0 Dalitz decay data. In addition to the multiple scattering, sometimes hard scattering occurs on the materials placed along the trajectory like cables and preamplifier. In such a case, the trajectory is completely changed and projection does not work, which results in large value of $|\Delta z_{\text{DCH}-\text{TIC}}|$ and tail of time measurement.

Refinement of the Timing-Counter Hit Time

For positrons with several bar hits, we can improve the time measurement by combining the hits. Even though we sometimes get more than three consecutive hits in a cluster, we only use first two hits since it is difficult to estimate the pass of positron after hitting a few bars. The propagation time between the two adjacent bars is estimated by the extrapolated track, and by subtracting it we calculate the time of positron hitting the first bar with another bar measurement. The time on the first bar t_{TIC} is modified as the average of the two measurements.

Finally, we reconstruct positron emission time

$$t_e = t_{TIC} - t_{eToF}. \tag{5.7}$$

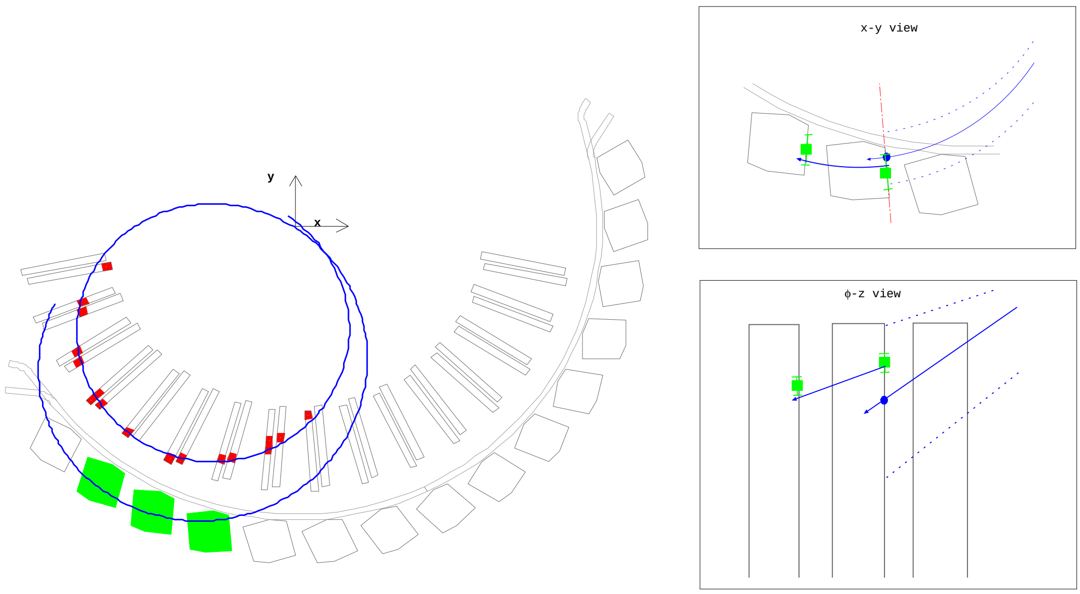


Figure 5.12: Schematic view of connection between track and timing-counter hits. Track fitted to the drift-chamber hits is extrapolated to the surface plane of the first hit bar. Blue circles in the close-up views show the predicted hit position and green squares show timing counter measurements. Because of the long distance projection, the predicted position on the timing counter has large uncertainty especially in z direction. The first two hits associated with the track is combined by the track to improve timing measurement.

5.4 Gamma-ray Reconstruction

5.4.1 General Remarks

The response of LXe to a gamma ray around 50 MeV is in the transient region from single process to electro-magnetic shower development. It can be understood noticing that the xenon critical energy is ∼14 MeV. However, we call here this process a "shower" even though the gamma ray losses its energy in a few steps. Some examples of the shower development in response to 52.8 MeV gamma in the MC simulation are shown in Figure 5.13. Note that there is large fluctuation in the shower development. The keys of gamma-ray reconstruction are there in the treatments of the shower fluctuation.

Before reconstructing gamma-ray hit, we convert the PMT charge to more physical variables, number of photoelectrons (N_{pe}) detected by the PMT,

$$N_{pe,i} = Q_i/(e \cdot G_i), \tag{5.8}$$

where Q_i is a pulse charge of the i-th PMT calculated by the charge integration of waveform, e is the elementary electric charge, and G_i is the PMT's gain. Next, it is converted to estimated value of the number of scintillation photons (N_{pho}) hitting the PMT photocathode,

$$N_{pho,i} = N_{pe,i}/QE_i, \tag{5.9}$$

where QE_i is the PMT's quantum efficiency including the collection efficiency of the first dynode. The calibration of PMTs is described in Sec.6.4.1. In reconstruction, we usually use the calibrated charge, N_{pho}, as an output of PMT. We use N_{pe} when we estimate the statistical fluctuation in the observed number of photoelectrons.

5.4.2 Position

A gamma-ray interaction point is reconstructed by fitting the PMT-output distribution. The method is schematically shown in Figure 5.14. We calculate an expected light distribution so that each PMT output is proportional to the solid angle of the photo-cathode viewed from the reconstructed position. The solid angle of a photo-cathode (a circle) from an arbitrary point, $\Omega_i(x, y, z)$, is calculated numerically. The three dimensional position of gamma-ray interaction, $(x_\gamma, y_\gamma, z_\gamma)$, is calculated by minimizing

$$\chi^2_{\rm pos} = \sum_i \frac{N_{pho,i} - c \times \Omega_i(x_\gamma, y_\gamma, z_\gamma)}{\sigma_{pho,i}(N_{pho,i})}, \tag{5.10}$$

where c is a constant factor and one of the free parameters of the fitting, and $\sigma_{pho,i}(N_{pho,i})$ is statistical uncertainty of the PMT's charge calculated by,

$$\begin{aligned} \sigma_{pho,i}(N_{pho,i}) &= \frac{1}{QE_i} \times \sigma_{pe,i}(N_{pe,i}) && (5.11) \\ &= \frac{1}{QE_i} \times \sqrt{N_{pe,i}} && (5.12) \\ &= \sqrt{\frac{N_{pe,i}}{QE_i}}. && (5.13) \end{aligned}$$

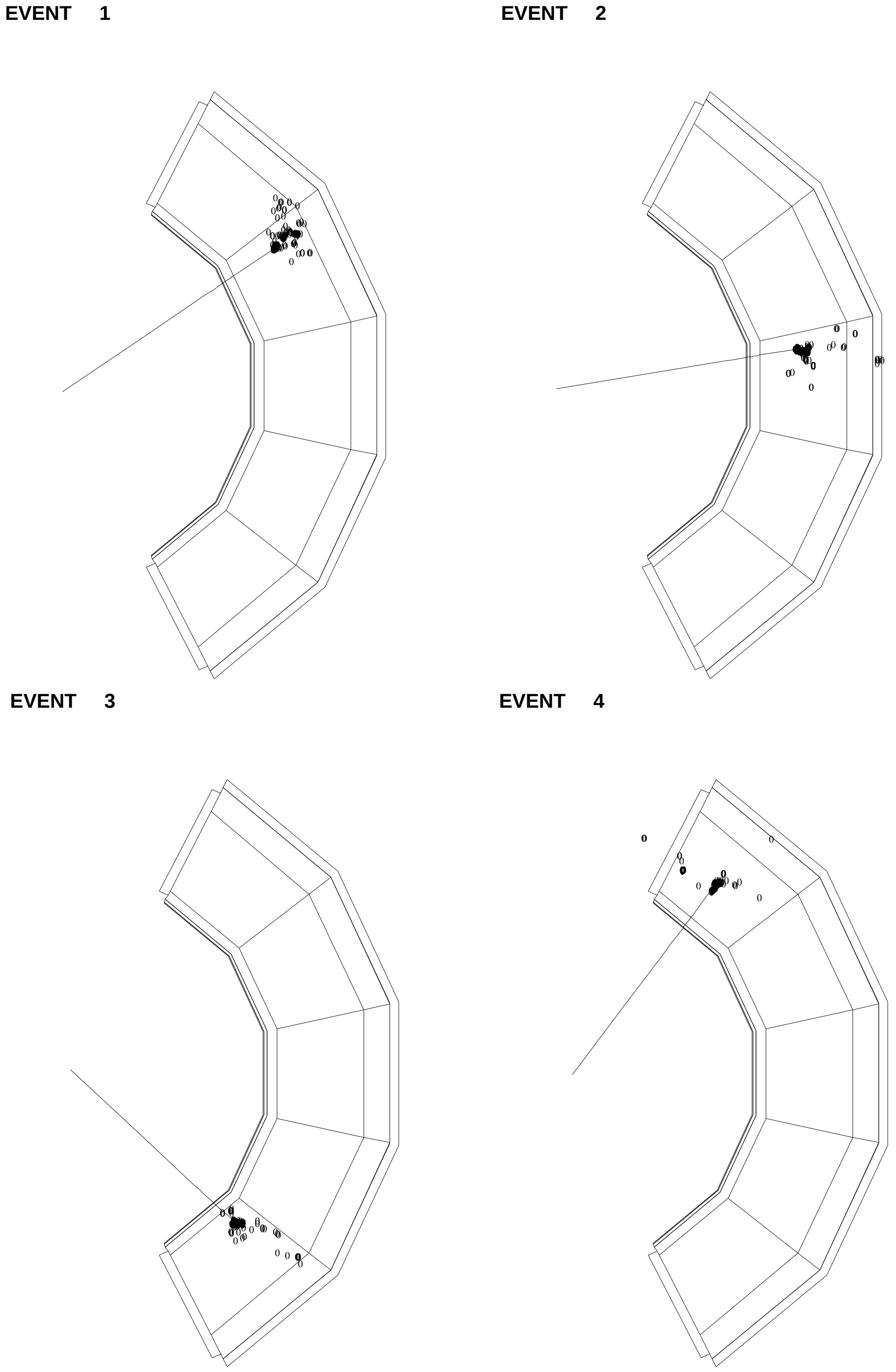


Figure 5.13: Examples of electromagnetic shower in LXe in response to 52.8 MeV gamma simulated with MC. Each small circle shows an energy deposit.

This fitting estimates three-dimensional position. Therefore we can reconstruct not only the two-dimensional position projected to the inner face (u_γ and v_γ) but also the depth of the interaction (w_γ).

The precision of this method will be limited by the fluctuation in shower developing behind the first interaction point since we assume here the scintillation light comes from a point-like source. To minimize the effect, we perform the fitting twice only using PMTs in restricted region on the inner face. First fitting uses typically 45 PMTs around one observed maximum light. Second fitting is then performed with fewer, typically 15 PMTs, around the result of the first fitting. The way of PMT selection is explained in the Figure 5.15.

Nevertheless, the reconstructed position by the fitting has bias in u- and w-coordinates, while there is no bias in v-coordinate. In u-coordinate, the gamma ray comes into the detector with incident angle respect to the PMT direction. This oblique incidence causes a bias of reconstruction to outside direction because the shower usually develops in the original gamma ray direction. In w-coordinate, the bias is in deeper direction by the same reason. These biases are studied with MC simulation, and corrections are applied.

In addition, the event-by-event difference of shower shape can be estimated by the difference between the two fitting results. It gives us the direction of the shower development. It is also used to correct the remaining influence of the shower fluctuation.

5.4.3 Energy

Gamma rays around 50 MeV converted in LXe usually deposit all of their energy in the LXe active volume. Almost all of the deposit energy is converted to the emission of scintillation. Our basic idea of energy reconstruction is to collect as much scintillation photons as possible. If we can collect all of the photons, of course it is not possible, and if the absorption length is sufficiently long, then the total number of collected photons is proportional to the original gamma energy regardless of the position and shape of the shower.

On the basis of the idea, we reconstruct gamma energy by summing up outputs of all PMTs. In summation, we put a weight for each PMT to take into account the different coverage of photo-cathode for different location of the PMT. We define the following quantity,

$$N_{sum} = \sum_i w_i \times N_{pho,i}, \tag{5.14}$$

where a weight factor w_i is the inverse of photo-cathode coverage of i-th PMT.

This quantity can be a good estimator of the deposit energy when the first interaction point in LXe (w_γ) is deeper than a couple of centimeter regardless of the interaction point and the event-by-event fluctuation in shower development. However, when the gamma conversion occurs very close to the inner surface, the N_{sum} becomes very sensitive to the position of conversion point relative to the PMT alignment. Efficiency of photon collection changes event-by-event. For those events, a solid angle of photo-cathode from a conversion point can be a better variable to correct the collection efficiency than the coverage. We investigated the dependence of N_{sum} on conversion depth and solid angle for the response to the 55 MeV gamma ray, and found a good correlation of sum of N_{pho} over inner face PMTs with the solid angle of the maximum-output PMT on inner face ($\Omega_{in,max}$) as shown in Figure 5.16. We apply a correction of N_{sum} with $\Omega_{in,max}$ for events with $w_\gamma < 3$ cm. The correction function is given as a linear function, whose coefficients are obtained by the 55 MeV-gamma events in the π^0 calibration.

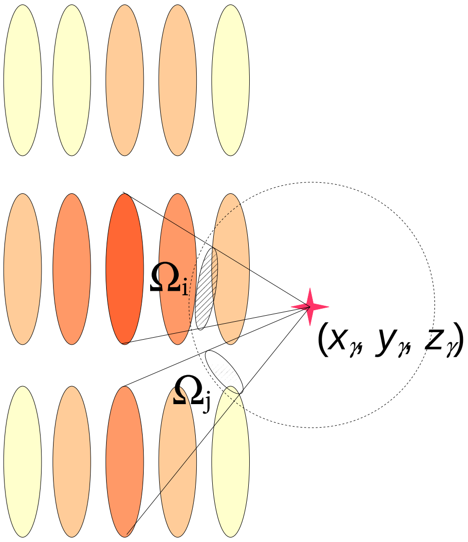


Figure 5.14: Schematic explanation of position fitting. Red star marker shows the first interaction point of an incident gamma. Orange ellipses show PMT photo-cathodes, and their color indicate the amount of observed light (the deeper color, the larger amount of light). The solid angle of each PMT from the interaction point can be estimated by the relative amount of light observed by the PMT.

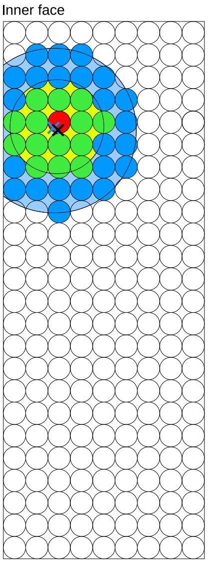


Figure 5.15: PMT selection used in the position fitting. The blue cross marker shows the initial estimate of interaction point calculated by the weighted mean around the maximum output PMT shown as red color. The PMTs used in the first fitting are those within 3.5-PMT distance from the point. The second fitting is done in more restricted region, 2-PMT radius shown as green PMTs from the result of the first fitting shown as black cross.

The remaining position dependence is corrected globally after the N_{sum} reconstruction. The detail of the calibration of the position dependence and the correction are described in Section 6.4.2.

The linearity of N_{sum} to gamma energy E_γ was checked using several energy points available from our calibration methods as shown in Figure 5.17. A conversion factor from N_{sum} to E_γ is given from the 55 MeV peak.

5.4.4 Time

With the reconstructed interaction point, we can reconstruct the time of first interaction by individual PMT,

$$t_{hit,i} = t_{PMT,i} - t_{delay,i} - t_{offset,i}, \tag{5.15}$$

where $t_{PMT,i}$ is time of i-th PMT measured by the constant-fraction method described in Sec.5.2.3; $t_{delay,i}$ is time delay during the scintillation light propagation in LXe; and

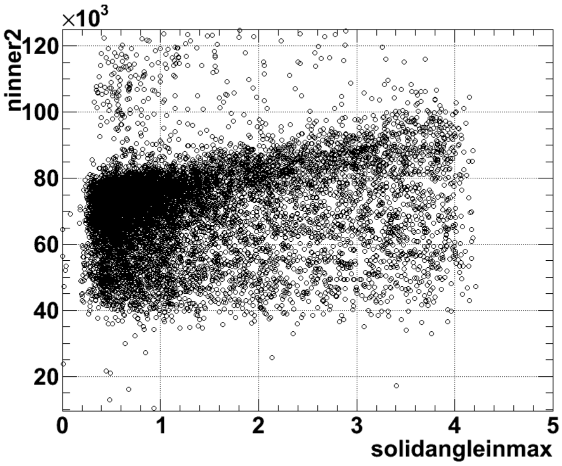


Figure 5.16: Correlation between $\Omega_{in,max}$ and sum of inner PMT N_{pho} for $w_\gamma < 3$ cm.

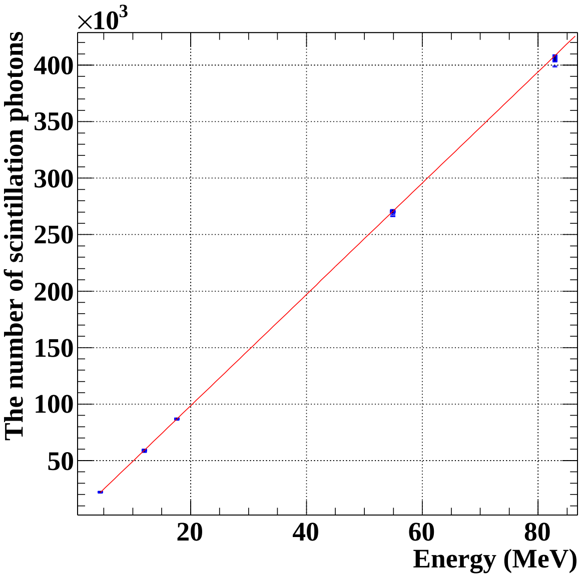


Figure 5.17: Linearity plot of E_γ v.s. N_{sum}. The five points are given from 4.4 and 11.7 MeV from CW-B run, 17.6 MeV from CW-Li run, and 54.9 and 82.9 MeV from π^0 run. Red line is the best-fit linear function whose intercept is fixed to zero.

$t_{offset,i}$ is a constant time offset of the channel. Let us think about the second term. We can think of possible cause of the delay as the following three: time of scintillation light propagating in LXe from light source to the PMT, t_{prop}; that of light which reaches the PMT indirectly from the light source, t_{indir}; and the time-walk effect of time-extraction method, t_{walk}. The t_{prop} can be a function of distance between light source and PMT d and effective light speed in LXe v_{eff}. We could model t_{indir} as a function of incident angle to the PMT η because as η gets larger, the fraction of scattered or reflected photons in the observed light increases. The t_{walk} would be a function of N_{pe}. Those are illustrated in Figure 5.18. Thus the t_{delay} can be write down as

$$t_{delay} = t_{prop}(d, v_{eff}) + t_{indir}(\eta) + t_{walk}(N_{pe}). \tag{5.16}$$

Details of the extraction of those function forms and coefficients, and calibration of time offset are described in Sec 6.4.4.

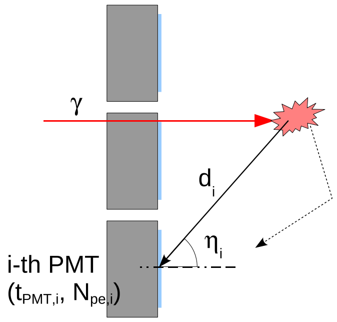


Figure 5.18: Illustration of time reconstruction.

Then, we reconstruct the hit time, t_{LXe}, by combining those measurements. The χ^2_{time} is calculated with PMTs which collect more than 50 photoelectrons,

$$\chi^2_{\text{time}} = \sum_i \frac{(t_{hit,i} - t_{LXe})^2}{\sigma_{t,i}(N_{pe})^2}, \tag{5.17}$$

where $\sigma_{t,i}(N_{pe})^2$ is time resolution of each PMT as a function of the number of photoelectrons. The t_{LXe} is determined so that the χ^2_{time} becomes minimum. Typically about 150 PMTs are used for around 50 MeV gamma. During this fitting process, large χ^2 channels are rejected to remove pileup effect and the minimization is iterated.

5.4.5 Pileup

Pileup Identification

We identify pileup events by

i) light distribution in the inner surface PMTs,

ii) time distribution in all PMTs.

The first item separates multiple gamma rays spatially. We perform a peak search on the light distribution of inner surface PMTs. We can find the position of each gamma ray. The power and mis-identification probability depend on the threshold for the peak search. In 2008 analysis, the threshold was set to 500 photons, which corresponds to ~80 photoelectrons and ~8 mV of pulse-height.

The second item uses temporal separation of different gamma rays. For this, we can use the normalized χ^2 value of the time fitting, $\hat{\chi}^2_{\mathrm{time}}$. For a single gamma event, all PMTs have hits in coincidence. On the other hand when pileup occurs by the accidental overlap, the two gammas have different conversion time. Even though the overlapping gamma ray has small energy, the closest PMT observe relatively large amount of light and the time is measured as the overlapping hit time. Those events results in large χ^2 values. The distribution of $\hat{\chi}^2_{\mathrm{time}}$ of signal MC and data are shown in Figure 5.19. We set a threshold for the identification at 3.

These two methods are complementary each other.

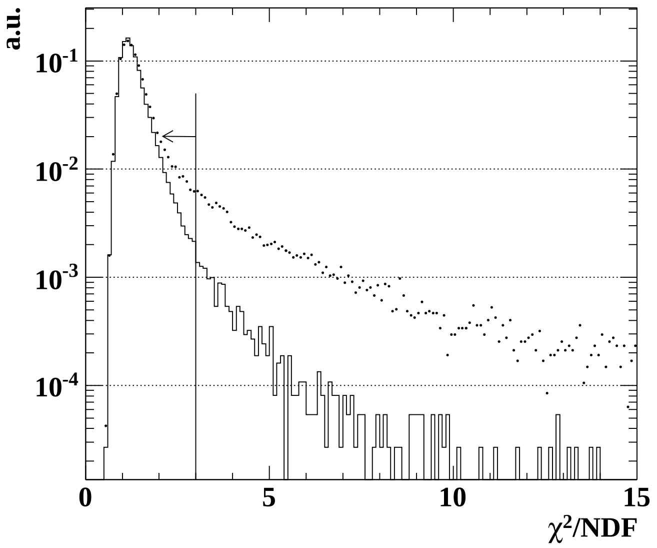
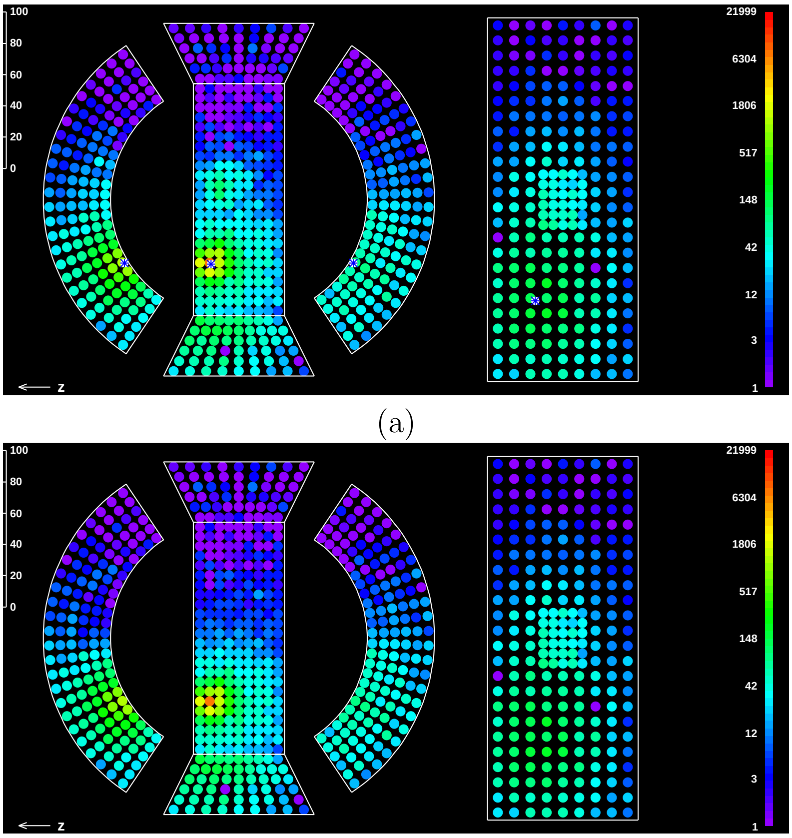


Figure 5.19: Distribution of normalized χ^2 of the time fitting. Dotted plot shows that of data and solid line shows that of signal MC.

Pileup Elimination

By the MC study, in 3×10^7/sec beam rate, about 8.5 % of events suffer from pileups, and it causes an inefficiency for the signal detection. Even if an event is overlapped by another gamma ray, we can use those events by unfolding or eliminating the overlapping one. We developed an algorithm for the pileup elimination to recover the efficiency. Elimination of contribution of energy by pileup gamma ray is done by the following steps.

i) Estimate energy by fitting PMT charges except for those around the pileup gamma. The fitting is done on the basis of a table obtained from CW data. In the table, average outputs of each PMT for each position mesh(1.55 × 1.55 × 1.55 cm) are written.



(a)

(b)

Figure 5.20: Light distribution before(a) and after(b) eliminating pileup contribution. The color axis shows the amount of observed light on each PMT.

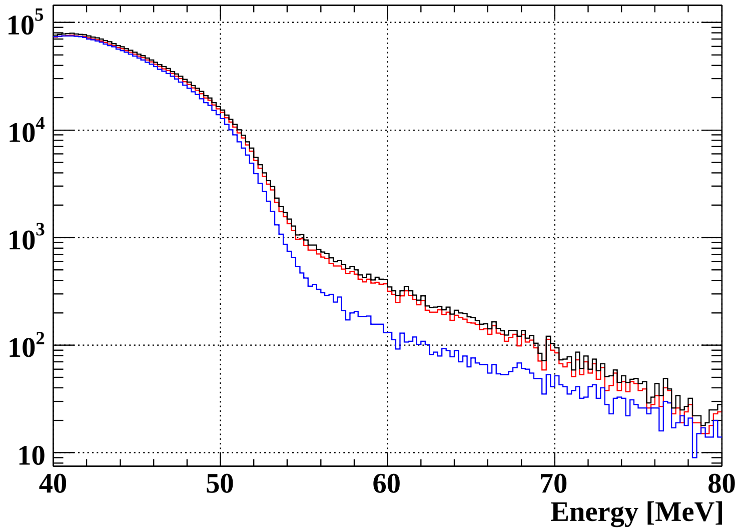


Figure 5.21: Energy spectrum in the sideband data. Red line is that after subtracting pileup events identified by time but not by light. Blue line is that after pileup elimination (the same event set as the red line).

ii) Calculate the expectation of outputs of PMTs around the pileup gamma ray using the table, and replace them with the expectation.

iii) Reconstruct the energy in the usual way using the replaced PMT charges.

Figure 5.20 shows PMT output distribution "before" and "after" eliminating contribution of a pileup gamma ray.

The left plot of Figure 5.21 shows gamma energy spectrum of the sideband in 2008 muon runs with various cuts. For the blue histogram, the same cuts of gamma ray for physics analysis are applied.

5.5 Combined Analysis

5.5.1 Relative Angle

The gamma-ray direction $\hat{p}_\gamma(\theta_\gamma, \phi_\gamma)$ is reconstructed as the direction from the muon decay vertex $\vec{x}_\mu$ reconstructed by the positron tracking to the reconstructed position of the gamma ray $\vec{x}_\gamma$; $\hat{p}_\gamma = (\vec{x}_\gamma - \vec{x}_\mu)/|\vec{x}_\gamma - \vec{x}_\mu|$. Then the opening angle between positron and gamma ray $\Theta_{e\gamma}$ is given by

$$\cos\Theta_{e\gamma} = \hat{p}_e \cdot \hat{p}_\gamma. \tag{5.18}$$

The relative angles in θ -and ϕ-directions are reconstructed as

$$\theta_{e\gamma} = \theta_\gamma - (\pi - \theta_e), \tag{5.19}$$

$$\phi_{e\gamma} = \phi_\gamma - (\pi - \phi_e). \tag{5.20}$$

5.5.2 Relative Time

The time-of-flight of gamma $t_{\gamma ToF}$ is calculated by a line segment between the reconstructed interaction point and muon decay vertex reconstructed by the positron tracking. Then the gamma emission time t_γ is reconstructed as,

$$t_\gamma = t_{LXe} - t_{\gamma ToF}. \tag{5.21}$$

Then the relative time difference between positron and gamma ray is given by

$$t_{e\gamma} = t_\gamma - t_e. \tag{5.22}$$

Chapter 6

Calibration

Needless to say, calibrations of detectors are crucial in a precise measurement. In addition, for a long-term experiment, continuous monitoring of detectors becomes a key to achieve high performance and to validate the results. We developed various calibration and monitoring methods. Some of them are complementary and some are duplicative for the cross check. Some instabilities and deviations of observables are corrected by the calibration data.

6.1 DRS Calibration

6.1.1 Voltage Calibration

Figure 6.1 shows a typical response curve of a DRS cell. The DRS has non-linear response. We see worse resolution over 0.5 V input from the curve. To use good resolution range, we adjusted inputs from MEG detectors to be smaller than 0.5 V. We prepared these response curves in advance by looking at the response for different input voltages from internal DC voltage generator on DRS boards. The online frontend load and apply the prepared calibration tables. Therefore, the data written in the disk is already calibrated and in unit of volt.

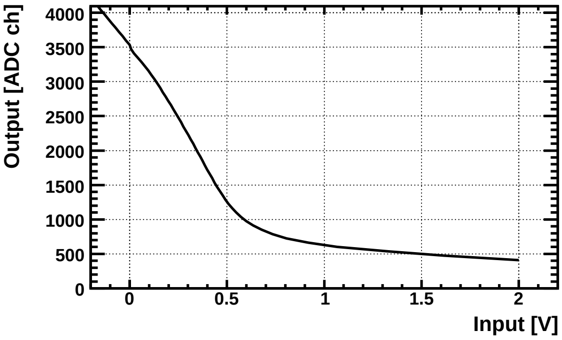


Figure 6.1: Typical response curve of a DRS cell.

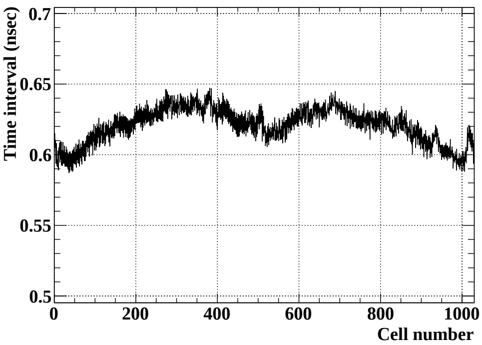


Figure 6.2: Typical sampling intervals of a DRS channel.

6.1.2 Time Calibration

A sampling frequency is generated by a series of inverters in a DRS chip. The speed of this “domino wave” is controlled by an analog voltage. Thus the speed varies if the voltage varies, for example, because of a change of temperature. The speed is regulated once at the beginning of each run but there is no feed-back system to regulate it during the run. The sampling interval between each point reflects a characteristic of individual inverter. As a result, each sampling point has different sampling interval. It is, however, a hardware-related characteristic, and thus stable over time relatively to the speed of domino wave.

Therefore we first apply the time interval calibration for the time of each sample point. A dynamic event-by-event calibration is then performed by analyzing a global clock signal. During this clock analysis, the phases of domino wave of all chips are synchronized to get a common time zero.

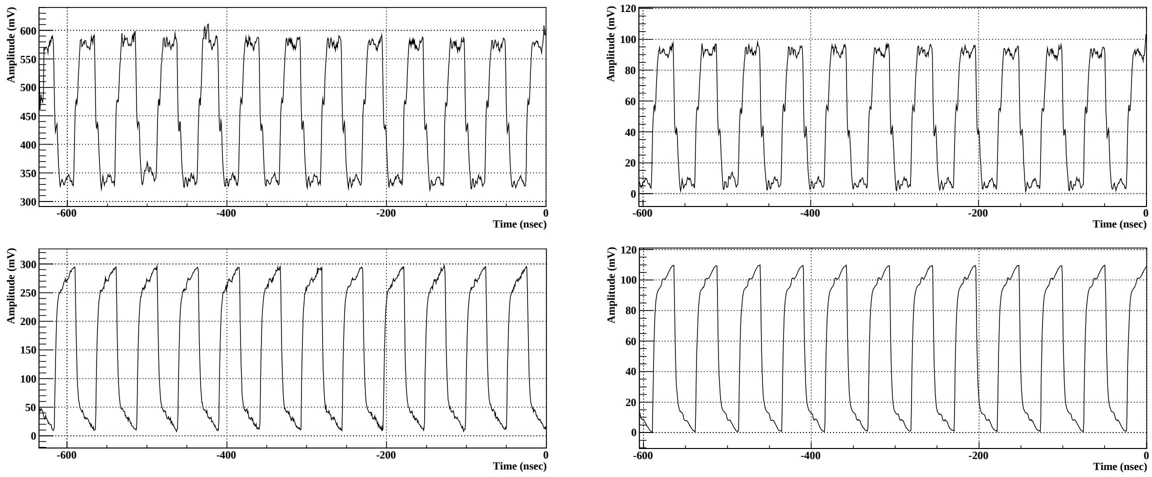
Static Calibration of Sampling Intervals

A sampling interval corresponding to a given inverter was calculated using a periodic wave from an external function generator. Sine waves at 240 (60) MHz were used for chips used at 1.6 (0.5) GHz sampling. We took data of the sine waves with random trigger and analyzed them in offline. On each event, periods between zero-crossing points of the sine waves were measured. Sampling intervals of cells between the zero-crossing period were adjusted so that the measured zero-crossing period become that of input sine wave. This regulations were repeated until sampling intervals of all cells got converged. The calculated sampling intervals of all cells of all channels are stored in a database, and they are applied when the waveform data are read in offline analysis. Figure 6.2 shows an example of sampling intervals of a chip calculated by this procedure.

Clock Analysis

A global common clock (square wave of 19.44 MHz) is distributed to all DRS boards from the trigger system. It is used as the time reference of the experiment. The clock signal is digitized and recorded in an additional channel of DRS and analyzed i) to calibrate the domino wave event-by-event, and ii) to synchronize the phase among chips. The domino wave of each chip is running freely and independently. It stops when an external trigger signal comes. Although the trigger signal from the trigger system is also common, because of the independent circulation of domino wave on each chip, actual stop timing of each chip fluctuates as large as one sampling interval, namely $\sim$600 ps for 1.6 GHz sampling. Therefore, the synchronization of chip timings is quite important to achieve the time resolution of $<$100 ps. On the other hand, for the DCH analysis, we do not need such a high time resolution. Thus we analyze the clock only for the data of 1.6 GHz sampling.

The clock signal taken by DRS2 and DRS3 are shown in Figure 6.3. It is clear that DRS3 has better quality of the signal: higher bandwidth and much less distortion. It is also known that DRS3 has less cross talk of the clock signal onto the normal data channels. Taking those better timing performance and the available number of channels into account, we adopted DRS3 to digitize timing-counter NIM pulses, which are one of the most important parts of timing analysis. While DRS2 chips were used for all of LXe



(a) Raw data. (b) After analysis.

Figure 6.3: Examples of the global clock signal digitized by DRS2 (top) and DRS3 (bottom). You can see the clock signals are out of phase before the clock analysis (a). We analyze clock signal so that they get in phase (b).

channels, which are also important but whose timing performance can be recovered by using many channels on many chips.

The procedure of clock analysis is the following:

i) Measure roughly the domino speed by a peak search with differential signal of the clock, and correct the speed.

ii) Search for a edge of the clock closest to the expected timing where the pulse of triggered signal should be around. A reference time of the channel is defined as the edge time and a global time zero is defined as the reference time of a reference channel defined in advance.

iii) Fit several edges around the reference time with a template clock waveform.

iv) Adjust the domino wave locally around the reference time using the results of fittings in step iii).

v) Perform a fitting with wider range and determine the reference time precisely.

vi) Shift the time of all sampling points of the channel to synchronize the reference time with the global time zero.

We estimate the precision of the synchronization of chips by the difference of time resolutions of the timing counter evaluated by looking at the time difference of two consecutive bar's hits between the case that two bars are connected to a common DRS chip and the case using different two chips. Average values are estimated to be σ_{clock} ∼47 and ∼37 ps for DRS2 and DRS3, respectively.

6.2 Calibration of Drift Chamber

6.2.1 z-coordinate Calibration

The z-coordinate calibration of drift-chamber hit, which means the relative gain correction of each ends, is done using the known pattern of vernier pad. We calibrate the anode relative gains so that the reconstructed length corresponding to a turn of vernier circle becomes the pitch length of 5 cm. We calibrate the cathode gains using the fact that the mean charge collected on each pad should be a quarter of total charge.

6.2.2 Time Offset Calibration

We calibrate the time offset of each channel by finding the leading edge of hit-time distribution corrected by t_{track}. We fit the distribution with a polynomial function, empirically given as fifth-order one, and the offset is determined by the time that the height of the distribution reaches 15 % of the maximum. An example of the fitted distribution is shown in Figure 6.4.

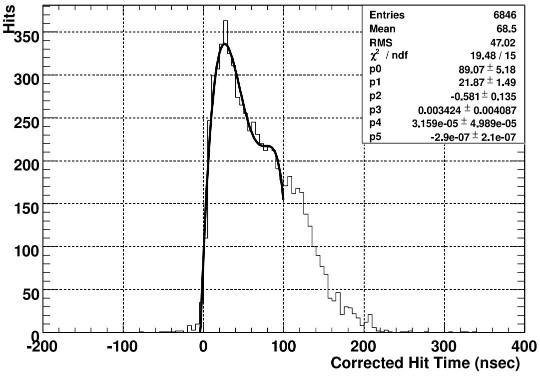


Figure 6.4: An example of corrected hit time distribution.

6.2.3 Time-to-Distance Calibration

The time-to-distance functions are calculated by GARFIELD program. If deviation from data is observed, the function is modified so that the deviation disappear. This calibration is done with simple iterative procedure.

i) Calculate the residual between obtained drift distance and the closest approach from the wire to the reconstructed track.

ii) Put a corrective offset to the drift distance so that the residual is minimized.

iii) Perform above for many events with similar incident angle and closed field strength, and build a new time-to-distance plot.

iv) Fit the plot and form a new time-to-distance function.

Repeat this procedure until the functions are converged.

6.3 Calibration of Timing Counter

6.3.1 z-coordinate Calibration

The z-coordinate calibration means adjustment of time offset between the two PMTs because z hit position is reconstructed by the time difference between the two. We use the reconstructed positron tracks as a reference of z hit position. The relative time offsets were calibrated so that the distribution of $\Delta z_{DCH-TIC}$ has mean at zero. This calibration was confirmed by looking at the distribution of z hit position differences between adjacent bars hits for multiple-bar hits.

6.3.2 Inter-Bar Time Offset Calibration

We use CW-B data to calibrate the relative time offsets among bars. The ${}^{11}_{5}\mathrm{B}(\mathrm{p},\gamma){}^{12}_{6}\mathrm{C}$ reaction produces simultaneous gamma rays at 11.7 and 4.4 MeV with no angular correlation. We use the 4.4 MeV gamma ray at the gamma-ray detector as reference time, and measure the 11.7 MeV one on each bar. After the time-of-flights of two gamma rays are corrected, the distributions of time difference of the two give the relative offsets.

6.4 Calibration of Gamma-ray Detector

6.4.1 PMT Response Calibration

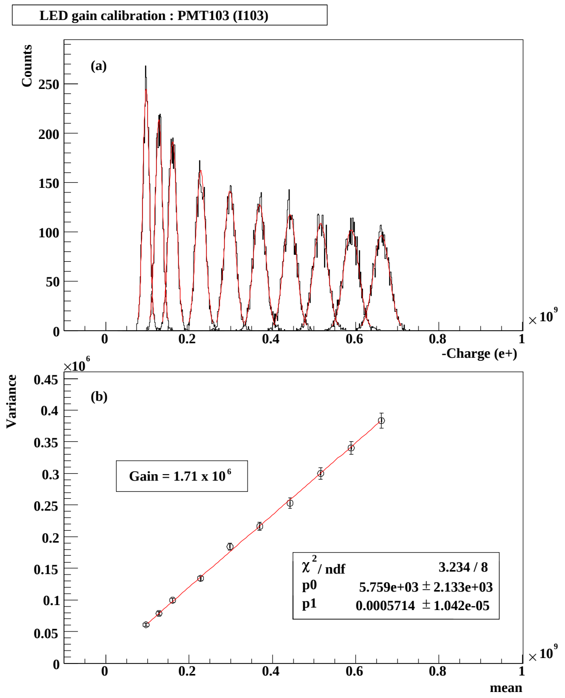
PMT Gain

We calibrate the PMT gains using LEDs. We use LEDs as stable light sources to monitor PMT outputs and also to calculate absolute value of PMT gains.

Gain Calculation We calculate the gain by means of the photoelectron statistics. Here we use a simple model. We assume LED output is constant and statistics of N_{pe} obeys the Poisson distribution. Let N_{pho} be the number of photon from LED to the PMT, which is constant in our assumption. Then the average number of photoelectrons will be $\overline{N}_{pe} = N_{pho} \cdot QE$, and its standard deviation will be $\sigma_{pe} = \sqrt{\overline{N}_{pe}}$, while the observed charge is given by $Q = Ge \cdot N_{pe}$. If we take many LED events, the mean and variance of the distribution become $\overline{Q} = Ge \cdot \overline{N}_{pe}$ and $\sigma_Q^2 = (Ge \cdot \sigma_{pe})^2 + \sigma_{noise}^2$ respectively. Therefore we obtain the relation,

$$\sigma_Q^2 = Ge \cdot \overline{Q} + \sigma_{noise}^2. \tag{6.1}$$

By taking LED events at different intensities, we can extract the gain, G, as a slope of the variance-mean plot. This method is insensitive to the noise contribution, σ_{noise}^2. Figure 6.5(a) shows the charge distribution of LED runs. In the normal calibration, we took data at 10 different LED intensities as one set of gain calibration, and each step contains 3000 events. The number of photoelectrons in each steps is sufficiently large to regard the distribution to be a Gaussian. We fit the charge distribution of each step with a Gaussian and estimate the mean and variance. Figure 6.5(b) shows the obtained mean-variance plot. It shows that those plots are well on a line. By fitting them with a linear function, we get the PMT's gain. The χ^2 distribution of the mean-variance fit is



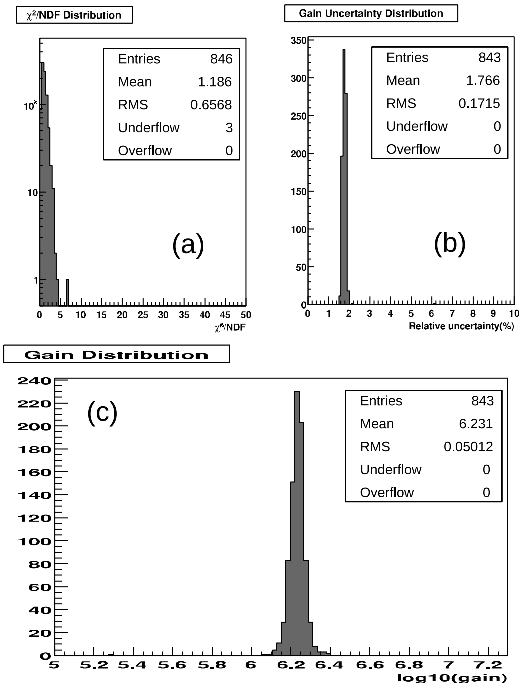


Figure 6.5: An example of gain calculation with LED data. (a) shows charge spectra of different intensity LED runs. (b) is the mean-variance plot of (a). The gain is extracted from the slope of the fitted line.

Figure 6.6: Distributions of gain calculation. (a) χ^2 distribution of mean-variance fit. (b) Uncertainties of gain calculation. (c) Gain distribution.

shown in Figure 6.6(a). By monitoring these χ^2 values, we checked bad channels. The precision of this method is about 1.8 % as shown in Figure 6.6(b). Gains were adjusted around 1.7×10^6 as shown in Figure 6.6(c).

Stability and Correction We checked the long-term stability of PMT gains by looking at the LED event peak. Figure 6.7 shows two typical examples of time evolutions of PMT gains. We found several points from those plots. First, we observed two kinds of instabilities of PMT outputs. One is a gradual gain decrease. The gradual gain decreases were severe particularly in the π^0 run (in August and last few points in December), while in physics run gains were stable. These decreases are considered to be an aging effect of dynode surface material due to the load from large current. When PMT sees a large amount of light, a huge amount of electrons hit dynodes (especially one at the last stage), and cause local heat and material change. It causes the change of the second-emission-ratio of the dynode resulting in the change of gain.

The other instability is a rate-dependent gain shift. For example, in Figure 6.7(b) we see two bands of gain plots. Those two correspond to the gain values measured in beam on and off. Figure 6.8 clearly shows the shift of gain value after opening the beam blocker. The shift values are different for each PMT, and dependent on the beam rate.

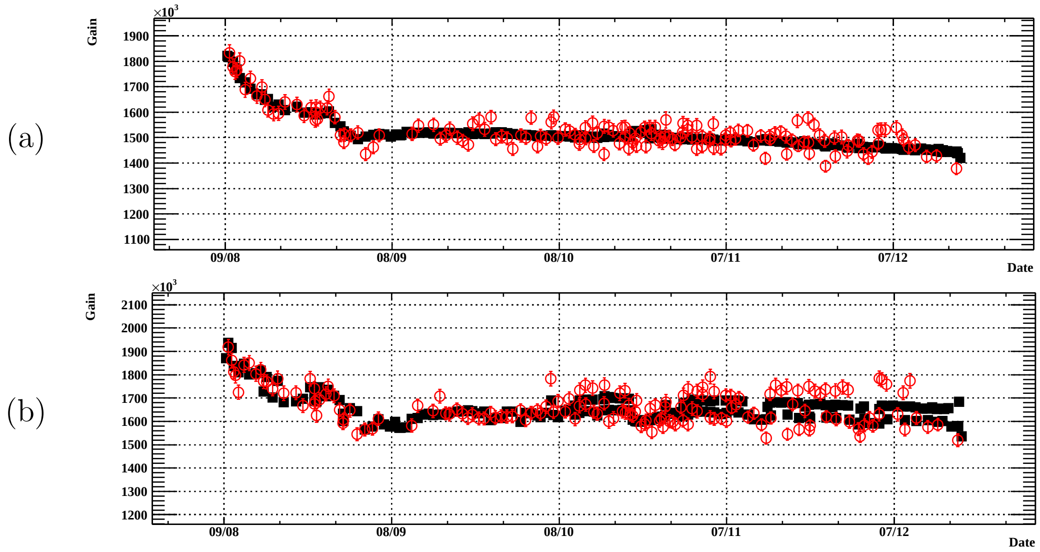


Figure 6.7: Examples of evolution of PMT gains. Red blank circle plots show absolute gain values calculated by the photoelectron statistical method. Black square plots show relative gain traced with the LED peak. (a) typical example of a stable PMT. (b) typical example showing rate dependent gain shift.

Typical shift value is ∼ 2 % to higher gain in the normal beam. In [91], similar kind of time dependence of rate-dependent gain shift was explained by a hole-trapping model on the dynode surface.

To correct these effects, we took the LED run every day. In particular, to measure the response under the conditions of physics run, we took LED data under the same beam conditions, while normal calibration sets including LED runs described previous paragraph are taken without beam. From Figure 6.7, we found that the light intensities of LEDs are quite stable and we can use them as constant light sources over time. Owing to this stability, the relative gain correction can be done much precisely by using the absolute measured peak value of the LED events, instead of measuring absolute gain value each time with the statistical method. Thus, we correct the gains relatively by using the variation of the peak value of LED events.

We checked the stability of the correction using LED events mixed in the physics run taken at 0.5 Hz. Figure 6.9 shows the total number of photoelectrons of the LED events as a function of time. After the gain correction, LED peaks are quite stable over time within ± 0.5 %.

PMT Quantum Efficiency

The PMTs' quantum efficiencies, or in another words equalization factors, can be calibrated using alpha events and 17.6 MeV line from ${}^{7}_{3}\mathrm{Li}(\mathrm{p},\gamma){}^{8}_{4}\mathrm{Be}$ reaction. The LED is not suitable for this purpose because wavelength of the light is different from that of xenon scintillation, and the response of the PMT is sensitive to the wavelength. The overall PMT response factor can be decomposed into

$$f = 1/(G \times QE \times CE), \tag{6.2}$$

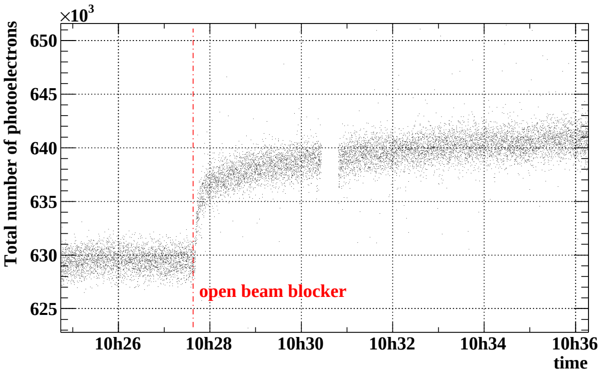


Figure 6.8: Time dependence of the gain shift when counting rate is changed. Red dot-dashed line shows the time when the beam blocker was opened.

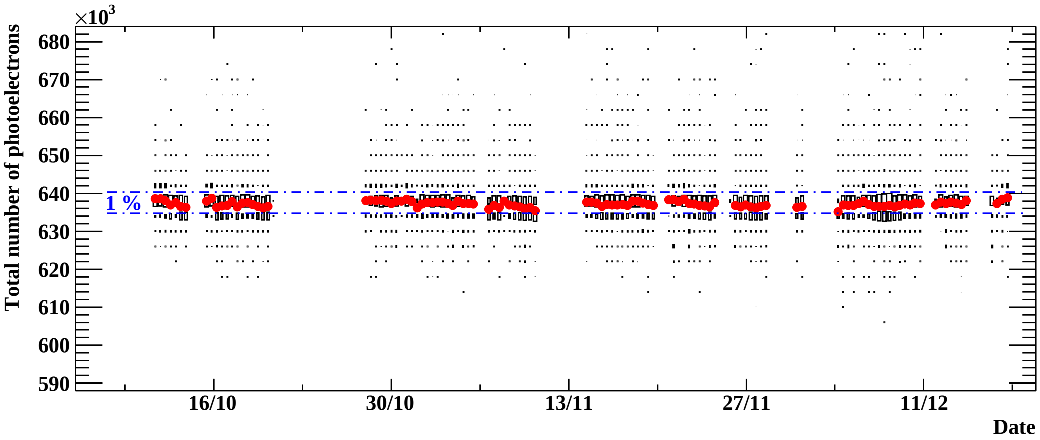


Figure 6.9: Plot of total number of photoelectrons of LED events vs. date. LED events were taken during the physics run. Red plots show peak value for each slice. After the gain correction, LED peaks are quite stable over time.

where CE is a collection efficiency of photoelectrons at the first dynode. The PMT gain G is calculated in the previous sub-section. Here, we want to calculate remaining part. According to our convention, in this thesis we call the multiplication of the quantum efficiency and the correction efficiency as "QE" because we do not calculate them separately.

The point-like alpha source on wires are mounted in the detector as described in Section 3.2.2. Since we know the positions of alpha sources and spectrum of the alpha particles, we can estimate the amount of light observed by given PMT. We use the MC simulation for the estimation. The QE is measured by comparing the observed charge spectrum with that of the MC estimation. In this process, we can use two types of alpha data. One is alpha data taken in LXe in normal calibration. The other is data taken in gas xenon (GXe). We took those data after the physics data taking while we were recovering xenon from the detector to the storage tank. These two kinds of data have advantages and disadvantages, and they are complementary. The LXe-alpha data are real-time data under the same conditions, while GXe-alpha data are less sensitive to the parameters of scintillation optics used in the MC simulation.

We took the LXe-alpha data in normal calibration run three times per week. For every calibration set, we can calculate a QE of each PMT. The precision of the measurements was estimated to be 3 %. We monitored the stability throughout the data taking, and confirmed the stability better than the each measurement precision. Therefore, we calculated average value using seven calibration sets taken when the gains and light yield of LXe were stable. In this way, we can reduce statistical uncertainty of the measurement.

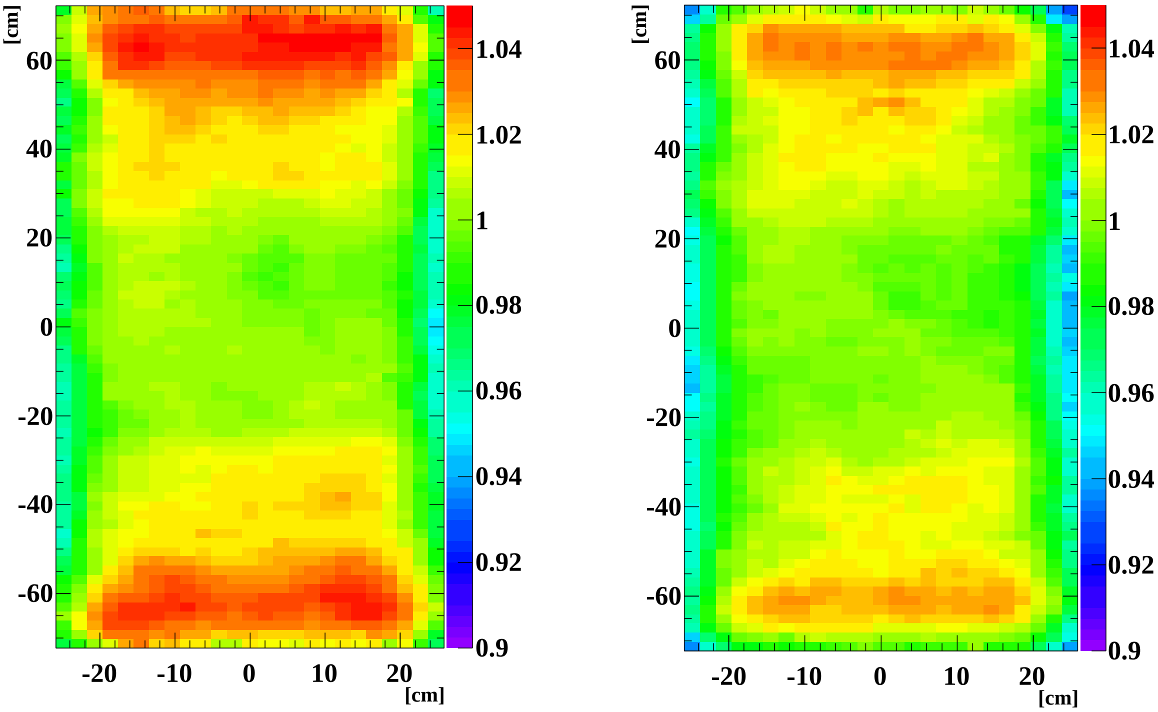
In GXe data, we observed upstream-downstream asymmetry of the measured QEs, which can be explained by the absence of BTS magnetic field in the measurement. This asymmetry is corrected with respect to the QEs measured with LXe data which does not show such asymmetry. On the other hand, there is systematic dependence of calculated QEs in LXe data. PMTs placed close to the corner of the detector have higher QEs, while this kind of dependence is not seen in ones measured with GXe data. The reason can be thought by the mismatch of optical parameters like scatter length and reflection coefficients. Therefore, the dependence is corrected globally, and systematic uncertainty due to the mismatch of MC parameters can be reduced. Finally, the measurement of both are averaged to reduce statistical uncertainty of measurement.

If the QEs measured in this way are free from systematic error, then the response to the 17.6 MeV peak becomes symmetry to the detector geometry. However, we observed some asymmetry: top-bottom asymmetry and a little upstream-downstream asymmetry. The reason is not fully understood yet. Displacement of PMTs and alpha wires is a possible cause. We can correct these asymmetries, for example, by applying a global position correction after reconstructing the energy. However, it is better to correct the PMT responses in advance. Thus, we corrected QE values to cancel out the asymmetry in 17.6 MeV line. Note that we do not correct them so that the response to 17.6 MeV become uniform, but correct only the asymmetry in order not to introduce systematic uncertainty again by the unknown optical parameters. A set of QEs calculated in this way is used throughout the data in 2008.

Another set of QEs was calculated by another analysis group of our collaboration with independent method. We carried out a study to estimate the effect of QE-measurement error in the reconstruction. The difference of the two sets of QEs was about 8 % in RMS. This difference was used as possible systematic errors of the measurement. The impact of wrong QEs on the reconstruction performance was investigated using the MC simulation with wrong QE values. This study shows that this level of error has little effect on the energy resolution if the position dependence is corrected. More effect can be seen in position resolution, but still sufficiently small respect to the current-achieved position resolution.

6.4.2 Non-uniformity

The gamma-ray detector has small non-uniformity of energy response with the N_{sum}-reconstruction described in Section 5.4.3, even though we try to cancel out coverage difference. The non-uniformity stems from a geometrical effect of different effective coverage and also the finite absorption length. To correct it, we use the actual response to the 17.6 MeV-monochromatic gamma rays. The data are divided into a mesh of (u_γ, v_γ) plane, and at each point, we measure the peak of the spectrum. Figure 6.10 shows the response maps. We observed a change of the dependence during data taking. The change is due to the change of optical properties by the purification. Therefore, we prepared two



(a) (b)

Figure 6.10: Relative energy response map in (u, v)-plane. These maps were made using the response to the 17.6 MeV line. The position dependence changed before (a) and after (b) the purification. The correction factors are extracted as inverse of those maps.

sets of correction tables for the data before and after the large jump of the LXe light yield (see the next sub-section).

For the w-coordinate, we calculated the correction factor using the central part of the detector. The w-correction is applied independently of the (u, v)-correction.

6.4.3 Energy Scale and Light Yield

We calibrate the energy scale using the 55 MeV gamma ray from π^0 decay. The uncertainty of the peak determination is evaluated to be 0.8 % by the error of the fitting.

As described in Sec 4.2.4, during the physics run in 2008, the light yield was changing as a result of purification. Therefore, we had to monitor the light yield continuously and correct the energy scale. We use several calibration sources to follow the variation of the light yield. Figure 6.11 shows the time evolution of amount of the observed light relative to the point when we determined the energy scale. We use mainly 17.6 MeV line during the physics run, and use 55 MeV peak during the π^0 run. We use cosmic-ray Landau peak for the connections between the π^0 run and the physics run, because we cannot use CW nor π^- beam during the exchange of beamline. Several jumps in the plot were caused by the liquid-phase purification during the beam-maintenance periods. The continuous increase after 30 October was due to the gas-phase purification.

This history is used for the correction of energy scale. The uncertainty of energy scale

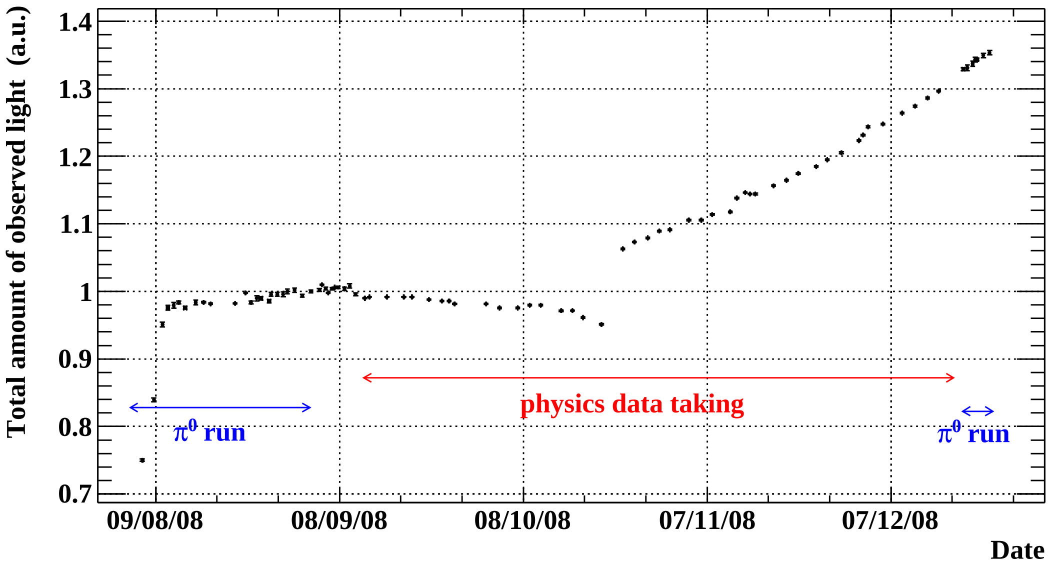


Figure 6.11: Evolution of the total amount of observed light. In this plot, results from several calibration sources are combined. The y-axis is relative to the point when we determined energy scale.

from the correction is estimated to be 0.18 %, mainly stemming from the uncertainty of the cosmic-ray Landau peak fitting.

6.4.4 Timing Parameters

Here, we calibrate parameters for the gamma-ray detector time reconstruction. The time reconstruction is based on the hit time reconstructed by individual PMT given by Eq.5.15 and Eq.5.16. The t_{delay} extracted by the π^0 data is shown in Figure 6.12 as a function of $1/\sqrt{N_{pe}}$. Here, the time reference is given by the reference counter in the NaI detector. What we have to do are to extract function form of each term and to calibrate coefficients of it. The difficulty stems from the fact that those three terms are correlated one another. For example, if we look at the delay time as a function of N_{pe}, then d is also effectively changed. We cannot disentangle the effect of light speed and time walk easily.

To disentangle each effect, we apply the following steps. First, we determine v_{eff}, and assume t_{prop} is equal to d/v_{eff}. By subtracting d/v_{eff}, we can remove the contribution of t_{prop} at least in first order. Then, we look at the dependence of t_{delay} as a function of η on the data in narrow region of N_{pe} as shown in Figure 6.13. We found that the dependence is well modeled with a function of $(1-\cos\eta)^2$. We adopt the empirical parametrization of t_{indir} and subtract it. Then, we look at the remaining dependence of N_{pe}, and we consider this dependence is the time-walk effect. In this way, hit times by individual PMT are calibrated. Now, we can reconstruct the hit time, t_{LXe}. Finally, we see t_{LXe} as a function of the total number of photoelectrons $N_{pe,sum}$. If the extracted t_{walk} is wrong, it would show some dependence. In this way, we disentangle the effect of d and N_{pe}.

We tried this procedure with v_{eff} = 10 cm/ns, which is a group velocity in LXe

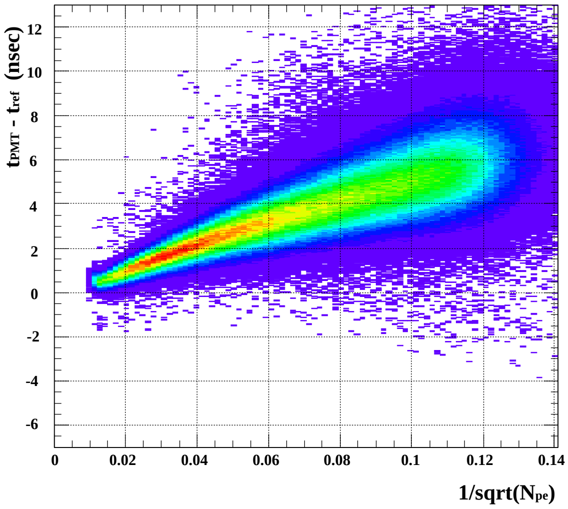


Figure 6.12: t_{delay} as a function of $1/\sqrt{N_{pe}}$. It is not clear this dependence shows time-walk effect as a function of $1/\sqrt{N_{pe}}$ or light propagation time as a function of d.

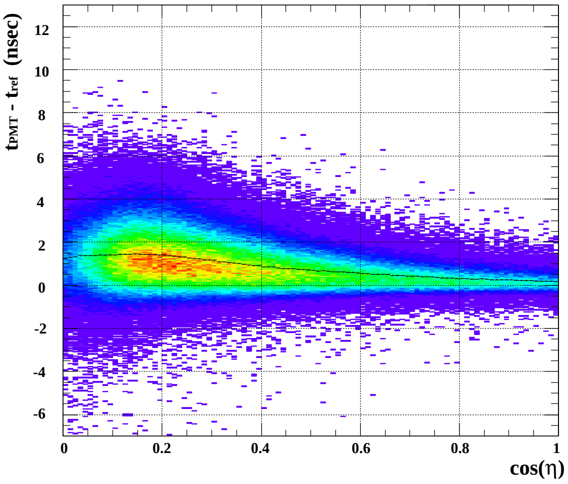


Figure 6.13: t_{delay} as a function of cosine of PMT incident angle η after t_{prop} is subtracted.

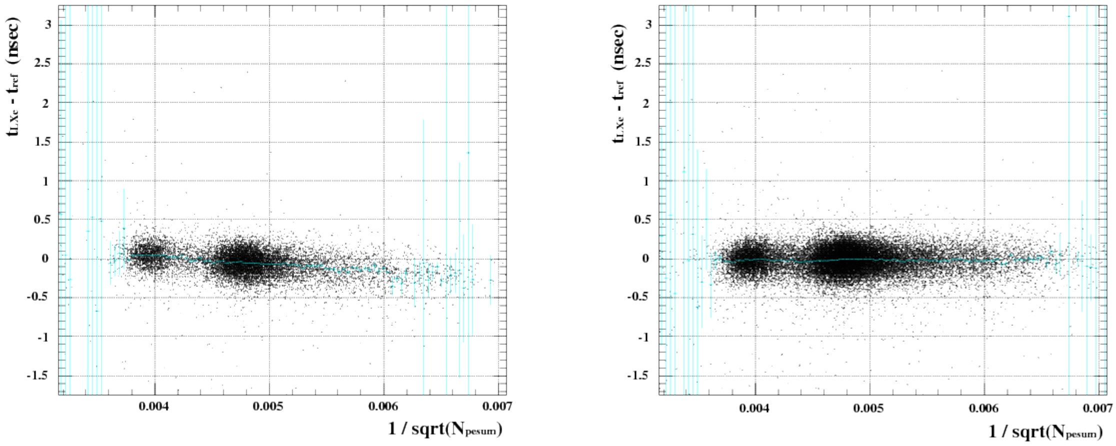
expected for the scintillation spectrum with peak value at 178 nm.[1] Then, we found strong correlation between t_{LXe} and $N_{pe,sum}$ shown in Figure 6.14(a). This indicates slower velocity of scintillation-photon in LXe. Then we tuned v_{eff} so that the dependence on $N_{pe,sum}$ disappear (Figure 6.14(b)). We found that v_{eff} around 8 cm/ns, the dependence disappear. At the same time, we found there is no time-walk effect in the t_{PMT}. The fact means that the digital-constant-fraction method described in Sec.5.2.3 works well and eliminates time-walk effect. This measured v_{eff} would partly include the effect of indirect photons, resulting in slower value. Or it may indicating shorter wavelength of LXe scintillation, possibly by the deformation of spectrum due to absorption by some impurities.

Once we calibrated the t_{delay} and subtract it from the PMT time, we can reconstruct t_{LXe} by individual PMT. Therefore, the systematic differences of reconstructed time by each PMT, $t_{hit,i}$, among the channels show the time offsets of channels. We calibrate the time offset so that $t_{hit,i}$ are synchronized. To avoid possible bias of position reconstruction and time parameter calibration, we use all the scanned data over the inner face to get uniform event distribution. Figure 6.15 shows t_{hit} distribution for all inner face PMTs after the calibrated t_{delay} and t_{offset} are subtracted.

Finally, the time resolutions of a single PMT are evaluated by looking at the time difference between the reference counter as a function of $1/\sqrt{N_{pe}}$. They are well linear to that quantity as shown in Figure 6.15(d). Those resolutions are used in the chi-square fitting.

---

[1] The propagation of light in media would obey the group velocity v_g instead of the phase velocity of $v_p = c/n = 18$ cm/ns (for LXe). In case of the known wavelength spectrum of LXe around 178 nm, v_g become 10 cm/ns.



(a) v_{eff} = 10 cm/ns (b) v_{eff} = 8 cm/ns

Figure 6.14: Dependence of reconstructed time on the total number of photoelectrons with different value of v_{eff} used in the reconstruction.

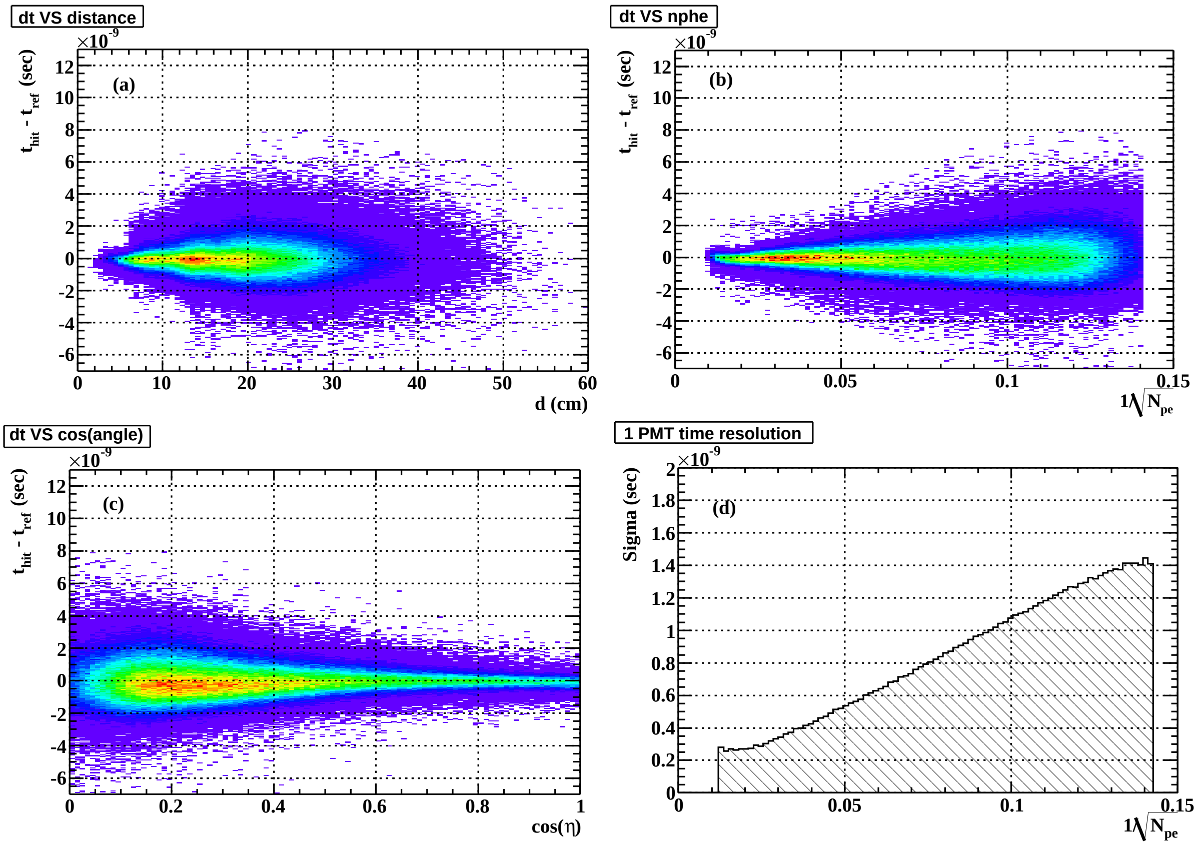


Figure 6.15: Distributions of t_{hit} reconstructed by individual PMT after calibrated t_{delay} is subtracted as a function of d (a), $1/\sqrt{N_{pe}}$ (b), and $\cos(\eta)$ (c). Calibration works fine and no dependence is seen. (d) shows the spread of the distribution (b).

6.5 Calibration of Relative Time

We discussed the variation of light yield as a result of the purification and its correction in Section 6.4.3. During the purification, not only the light yield but also the pulse shape of LXe scintillation were changing. The average pulse shape of LXe scintillation pulse in response to a gamma-ray interaction at the beginning and at the end of run 2008 are shown in Figure 6.16. It was getting wider. As a result, the reference time of the gamma-ray detector was also changed because the constant-fraction method is based on the constant pulse shape.

To monitor and correct the effect, we used the RD peak in the dedicated RD run. The dedicated RD run, which is described in Section 4.1.2, were taken in reduced-intensity beam. Therefore, we can calibrate the timing precisely with a good signal-to-noise ratio. The run were taken once per week for about 24 hours. We can calibrate time reference (t_0) of $t_{e\gamma}$ with each set. Figure 6.17 shows the plot of the t_0 as a function of time. Before we started the continuous purification, it was stable within 15 ps. However after we started the purification, it started drifting. We used a fitted linear function to correct the t_0. After the correction, the stability of t_0 over the whole period is evaluated to be better than 20 ps by looking at the RD peak in normal MEG run.

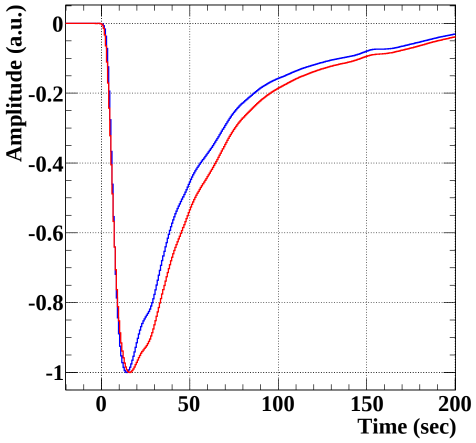


Figure 6.16: Change of LXe pulse shape. Blue line shows the average pulse shape at the beginning of run 2008 (in September), and red one shows that at the end of the run (in December). Those two are normalized by the pulse heights.

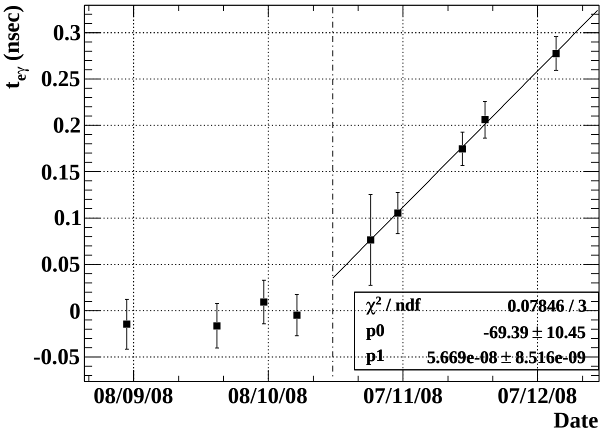


Figure 6.17: Drift of $t_{e\gamma}$ time reference. Each point shows the t_0 evaluated by a set of RD run. After we started the continuous purification, the time reference started drifting. This time drift was corrected by the fitted linear function shown in this figure.

Chapter 7

Performance

We evaluate performance of the detector in this chapter. First, we describe selection criteria for the performance evaluation. These criteria are also applied in the final analysis. Note that the selection is not so critical in our analysis because finally we perform the likelihood analysis with event-by-event PDFs. This method naturally takes the quality of each event into account. In an extreme case, we can use all the reconstructed events as long as we know the response to those events. Thus, our policy of event selection is to select events with minimum quality cuts and evaluate performance in detail. The resolutions of detector are evaluated using actual data, and response functions are extracted. Those responses are used directly or with a small correction as a probability density functions for the signal event in the likelihood analysis that will be described in Chapter 8. The detection efficiencies of both particles are also evaluated.

7.1 Event Selection

7.1.1 Definition of Fiducial Volume and Acceptance

First, we define the fiducial volume of the gamma-ray detector. At the very edge of the detector, the performance is extremely low in energy and position measurements and efficiency. Therefore, the detector was designed so that the half region of the outer-most PMTs are out of the acceptance. We define the fiducial volume with this original design, and we use full volume in the depth direction:

$$|u_\gamma| < 25, \quad |v_\gamma| < 71, \quad 0 < w_\gamma < 38.5 \text{ cm}. \tag{7.1}$$

The positron acceptance is then defined so that the inverse of the positron direction goes into the gamma-ray detector fiducial volume. With this definition, the acceptance of angular range is not defined uniquely because of the spread of muon decay vertex on the target. The acceptance is shown in Figure 7.1. In the original design, the center part in θ direction ($|\cos\theta| < 0.08$) was out of acceptance because a positron emitted in this direction turns many times in the spectrometer resulting in the difficulty of tracking and also the low probability that the positron reaches the timing counter. Actually, the event distribution shows low efficiency in this region. However, we do not intentionally exclude these events.

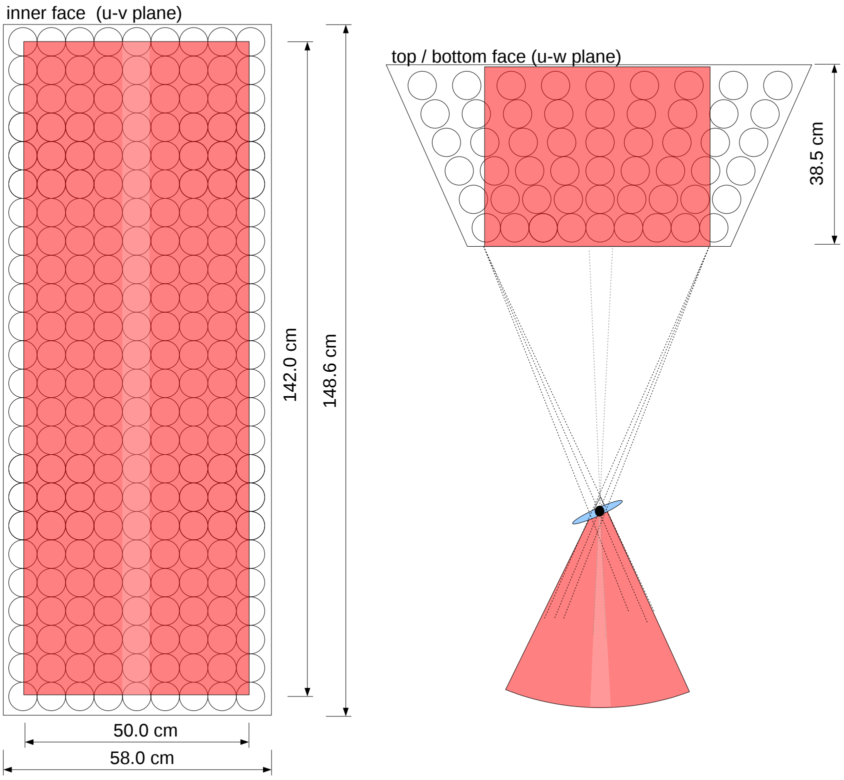


Figure 7.1: Definition of the fiducial volume. The red shaded part is the fiducial volume. The corresponding positron acceptance is defined so that the inverse of the direction goes to the fiducial volume. The light shaded part around center is positron low efficiency part due to the multiple turns, but it is also in the acceptance.

7.1.2 Positron Analysis Cut

Number of Hits

To select positrons reconstructed with reliable tracking, we required the following: the total number of hits ≥ 7; the number of chambers that have hits > 3; the span of the track > 4 chambers; the number of multi-hit chambers > 1. These are really minimum criteria, just to exclude abnormal tracks. The first and second criteria are actually full efficient. The third and fourth selections are shown in Figure 7.2.

Track-Fitting Quality Cut

We require the following criteria on the fitted tracks to select events with good quality of the fitting: $\delta E_e \leq 0.7$ MeV $\wedge$ $\delta\theta_e \leq 0.6°$ $\wedge$ $\delta\phi_e \leq 1.5°$, where δ means the uncertainty evaluated by the covariance matrix of the Kalman filter. In addition, we require the normalized χ^2 of the Kalman filter to be $\hat{\chi}^2_{\mathrm{track}} \leq 12$.

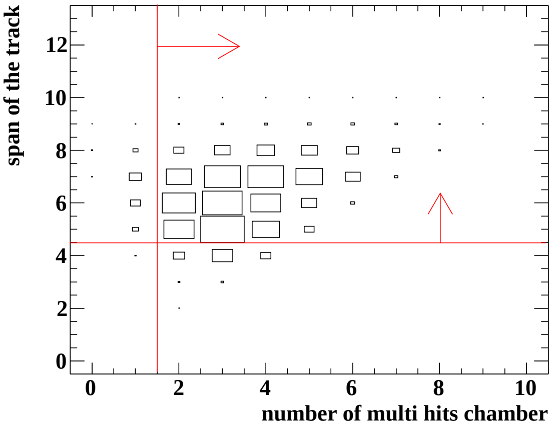
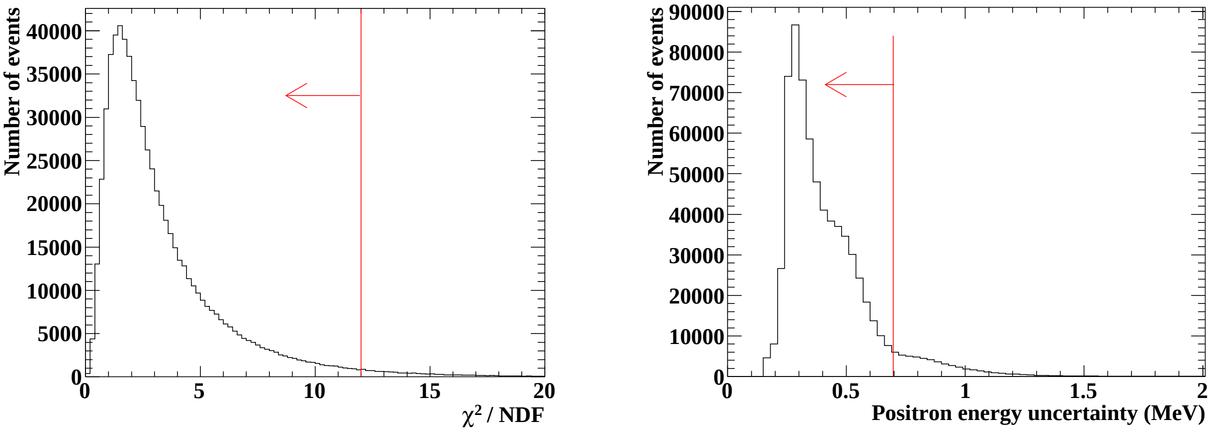
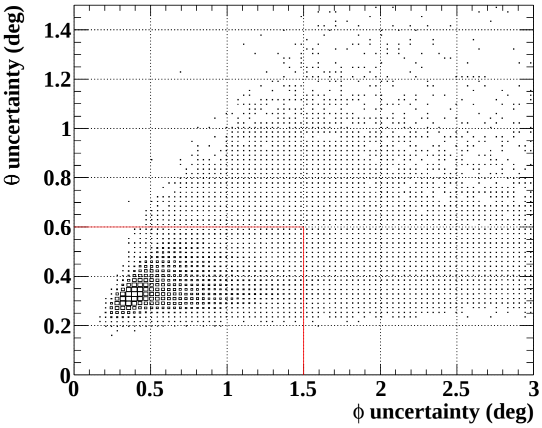


Figure 7.2: Distribution of track span v.s. number of multi-hit chambers.



(a) Distribution of normalized χ^2.

(b) Distribution of E_e uncertainty.



(c) Distribution of uncertainties of angular measurements.

Figure 7.3: Track quality selections.

Vertex Cut

To guarantee that the positron comes from a muon decay on target, we require the reconstructed muon decay vertex is on the target. Because of the additional air doped inside the spectrometer, the beam distribution on the target was slightly off center to upstream. The two-dimensional distribution of muon decay vertex and the selection criteria are shown in Figure 7.4.

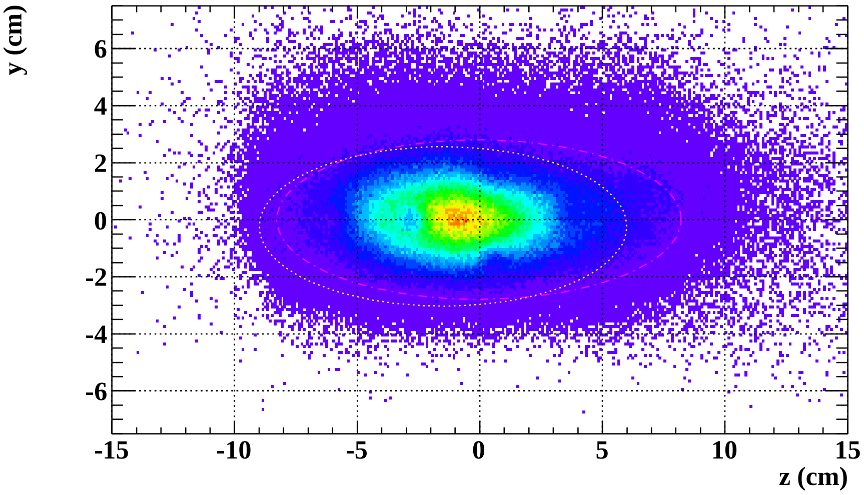
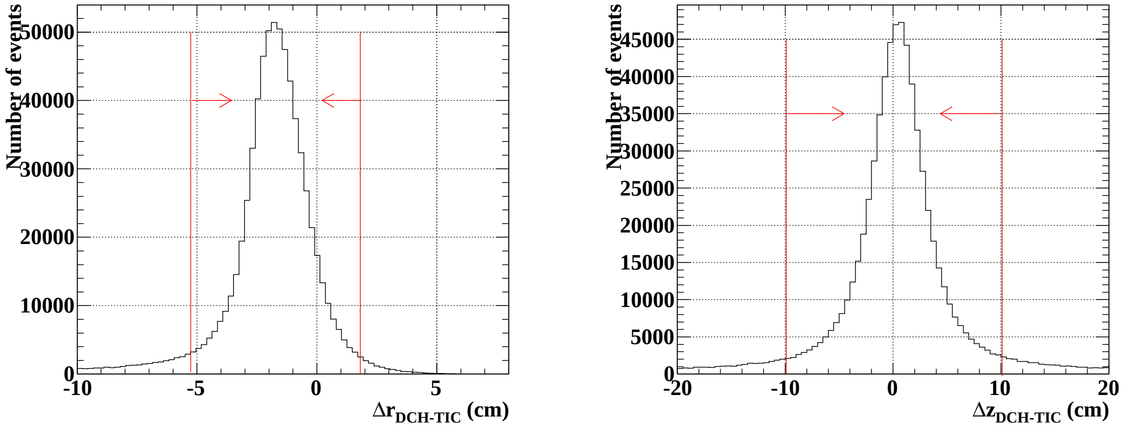


Figure 7.4: Two-dimensional distribution of muon decay vertex. Magenta dot-dashed line ellipse shows the target position selection and white dashed line ellipse shows beam distribution selection. The beam distribution was slightly off center to upstream side because of the additional air doping inside the spectrometer. A shadow of hole is clearly seen at $z = -3$ cm. Events inside of both ellipses are selected.

DCH-TIC Matching Cut

To select events well matched between the drift-chamber track and the timing-counter hit, we apply cuts on the quantities $|\Delta z_{DCH-TIC}|$ and $|\Delta r_{DCH-TIC}|$. As will be shown in Sec.7.5, the time resolution becomes drastically worse as these values increase. Figure 7.5 shows the distributions of those values. There are large tail components, which consist of two kinds of events. One is events with uncorrelated positrons for track and timing-counter hit. It makes a flat distribution in these plots because it is accidental. The other is positrons hardly scattered by some materials between the drift chamber and the timing counter. These hard scatterings cannot be predicted in the tracking and the projection becomes worse, resulting in the worse timing resolution. We required $|\Delta z_{DCH-TIC} - 0.1| < 10$ and $|\Delta r_{DCH-TIC} + 2.2| < 3$ cm. The systematic bias in $|\Delta r_{DCH-TIC}|$ distribution stems from that the r-position of timing-counter hit is just given as the center of the bar.



(a) (b)

Figure 7.5: Distributions of DCH-TIC matching quantities in (a) r- and (b) z- directions.

Ghost-Track Selection

Sometimes different tracks share some hits in the track finding process in high rate environment. The Kalman filter attempts to fit them separately. On the other hand, sometimes track finding process recognizes a track as different multiple tracks with some combinations of the hits connections. Even in such a case, the Kalman filter tries to fit them as different tracks. We call these tracks associating with a common positron ghost tracks.

Now we get two separate problems. One is that we have to decide the tracks which share some hits are really different tracks or ghosts. The other is if they are ghosts, then we have to choose the best measured track among them.

We put the following conditions on ghost tracks:

- ghosts must share a common timing-counter hit;
- the number of chambers which has at least one hit in common must be more than the half number of hit chambers on shorter of the two tracks.

If two tracks satisfy these conditions, then those two are recognized as ghosts, otherwise as different tracks.

We then rank those ghost tracks with the number of chambers spanned (S) and the normalized χ^2 of track fitting, $\hat{\chi}^2_{\text{track}}$. We defined the following quantity which represents the measurement quality;

$$G_{\text{rank}} = \frac{1}{S} + \alpha \hat{\chi}^2_{\text{track}}, \tag{7.2}$$

with which tracks are ranked. The smaller value a track has, the higher position it is ranked in. The coefficient α was tuned to 0.025 so that the second term dominates the ranking when ghosts have very different $\hat{\chi}^2_{\text{track}}$ and the first term dominates when the ghosts have similar $\hat{\chi}^2_{\text{track}}$. We select the highest rank track out of the ghosts passing the other selections.

7.1.3 Gamma-ray Analysis Cut

Pileup Cut

As described in Sec 5.4.5, we do not discard but use pileup events after eliminating the overlapping gamma. In the following two case, however, we reject the events:

- event on which the elimination does not seem working correctly;
- event identified as a pileup by timing distribution but not by light distribution, thus the elimination is not applied.

For the first item, we require $0 \leq (E_\gamma^{\text{normal}} - E_\gamma^{\text{eliminate}})/E_\gamma^{\text{normal}} < 0.1$, where E_γ^{normal} gamma energy by normal reconstruction without pileup elimination and $E_\gamma^{\text{eliminate}}$ is that after pileup elimination. Figure 7.6 shows the effect of the cut. Inefficiency due to this cut is 1.7 %. For the second item, we require $\hat{\chi}^2_{\text{time}} < 3 \vee N_{peak}^{\text{light}} > 1$. Inefficiency due to this cut is 3.9 %. In total, inefficiency by pileup rejection is (4.5 ± 2.6) %.

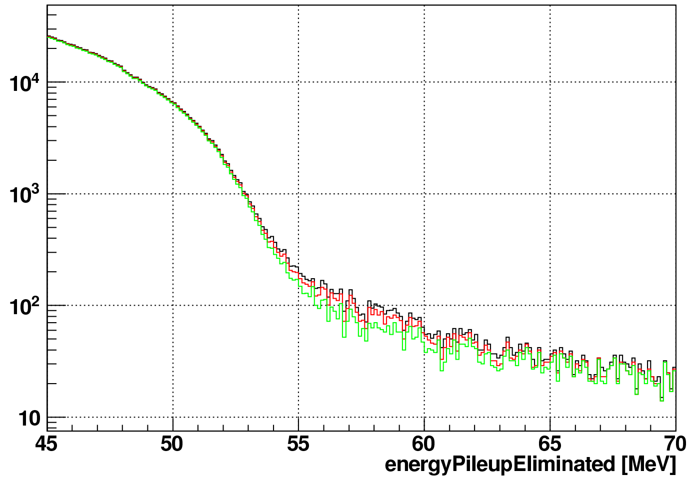


Figure 7.6: Gamma energy spectrum with additional pileup cut. Black line is without any cut, red is after cutting events for which the elimination works to increase energy, green is after cutting large elimination events additionally.

Cosmic-ray Cut

To cut cosmic-ray events, we use the event topology. Since cosmic-ray events deposit large energy in deep part of the detector, the ratio of inner PMTs sum to outer ones sum becomes different from that of gamma-ray events originating on the target. We require $N_{inner}/N_{outer} > 0.3$. Figure 7.7 shows cosmic-ray spectrum taken with the LXe-self trigger without beam. With the cut, the rate of cosmic-ray background is successfully reduced well below that from gamma from muon decays. The rate in $51 < E_\gamma < 54.6$ MeV is 0.35 Hz, which is about 1 % of gamma-ray background from muon decays. This cut also rejects gamma rays from the target when their interactions occur very at deep position. The signal inefficiency by this cut is estimated to be 1.1 % by using the MC simulation.

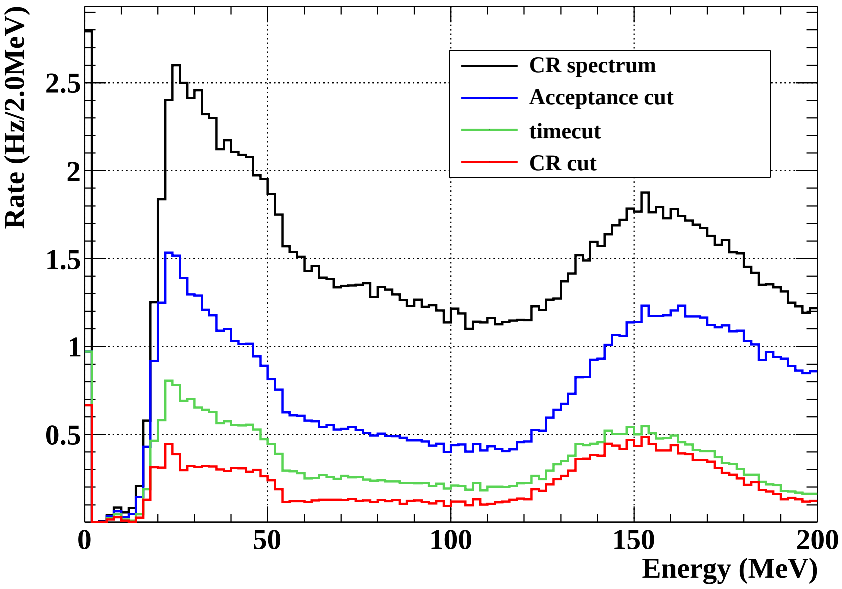


Figure 7.7: Cosmic-ray spectrum with several cuts. Black line shows measured spectrum, blue one shows that in fiducial volume, green one requires $\hat{\chi}^2_{\text{time}} < 3$, and red one requires also $N_{inner}/N_{outer} > 0.3$.

7.2 Positron Energy, E_e

We evaluated the positron energy scale and resolution by fitting the kinematic edge of the measured Michel-positron energy spectrum. We formed the fitting function by folding the theoretical Michel spectrum with the energy-dependent detector efficiency, and the response function for mono-energetic positrons. The former is empirically given by an error function ϵ_{acc}. The latter is well described in the MC simulation by a triple-Gaussian g_{tri}. Thus fitting function is given by

$$f(x) = (f_{\text{theo}}(x) \times \epsilon_{acc}(x)) \otimes g_{tri}(x), \tag{7.3}$$

where $f_{\text{theo}}(x)$ is the theoretical Michel spectrum. Figure 7.8 shows the fitting on reconstructed Michel spectrum for a range of $51 < E_e < 55$ MeV. The extracted response function is also shown in the figure. The resolutions extracted from the data are 0.374, 1.06, and 2.00 MeV in sigma for the core component and two tails, with corresponding fractions of 60, 33, and 7 %, respectively. The complete set of the spectrum parameters is given in Table 7.1.

To suppress statistical uncertainty, the response function was estimated on a large data sample: no cut on gamma energy, relative angles, and relative time. We evaluate the impact of the cuts defining the analysis window by repeating the Michel fitting on the $t_{e\gamma}$-sideband with gamma-energy and relative-angle cuts. The results are given in Table 7.2. Parameters not listed in the table were fixed to the nominal value. These differences indicate the systematic uncertainty of the positron response function.

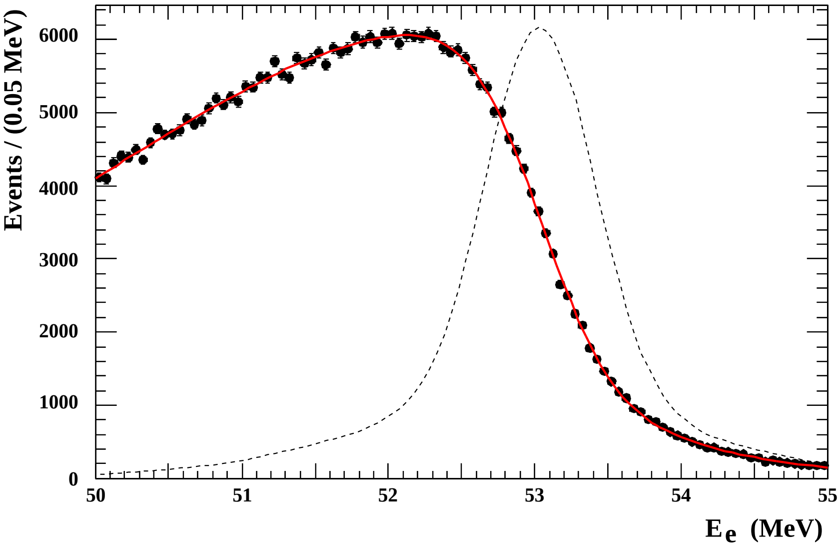


Figure 7.8: Fitting of Michel spectrum. Fitting range is $51 < E_e < 55$ MeV. Dashed line shows obtained response function.

Table 7.1: Results of the Michel fit.

Parameter	Value
μ_{acc} (MeV)	51.3
σ_{acc} (MeV)	3.32
f_{core}	0.607
μ_{core} (MeV)	0.206
σ_{core} (MeV)	0.371
μ_{tail} (MeV)	0.11
σ_{tail} (MeV)	1.07
f_{out}	0.07
μ_{out} (MeV)	0.0
σ_{out} (MeV)	2.0
χ^2/NDF	88/95

Table 7.2: Alternative parameter set estimated with analysis window cuts.

Parameter	Value
f_{core}	0.81
μ_{core} (MeV)	0.23
σ_{core} (MeV)	0.4
μ_{tail} (MeV)	1.4
σ_{tail} (MeV)	0.4

7.3 Gamma-ray Energy, E_γ

We measured the energy resolution of the gamma-ray detector using a response to the 55 MeV gamma from π^0 decay. To get 55 MeV gamma ray, we selected the π^0 decay events where 83 MeV gamma ray were detected in the NaI detector with an opening angle larger than 170°. A correction with the measured opening angle was then applied. Figure 7.9 shows the reconstructed energy spectrum at a point around center of the detector. The lower tail comes from mainly two reasons. One is the interaction of gamma ray on some material in front of the LXe active volume. The other is the shower escape from the inner face. You also see higher tail in the distribution. It is particular to the π^0 data since much more backgrounds from electrons in beam and high-energy gammas from π^0 decays are there in the environment with π^- beam.

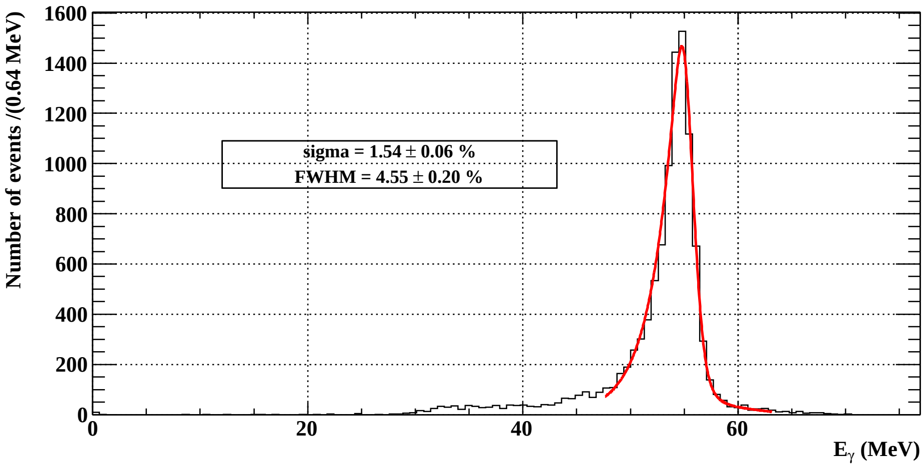


Figure 7.9: Reconstructed energy distribution of 55 MeV gamma events from π^0 decays. The numbers quoted in the figure is the intrinsic resolution after the pedestal distribution is unfolded.

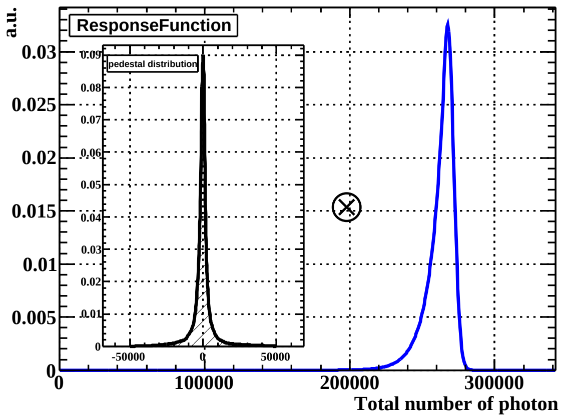


Figure 7.10: Fitting function.

We fitted the distribution with an asymmetric function with lower-side tail, $f_0(x)$, convolved with the pedestal distribution in the π^0 run, $h_\pi(x)$ (Figure 7.10),

$$f_0(x) \otimes h_\pi(x), \tag{7.4}$$

and $f_0(x)$ is given as

$$f_0(x) \quad = \quad \begin{cases} A\exp\left(\frac{t}{\sigma_{\mathrm{up}_0}^2}\left\{\frac{t}{2}-(x-x_0)\right\}\right) & x \leq x_0+t, \\ A\exp\left(\frac{(x-x_0)^2}{-2\sigma_{\mathrm{up}_0}^2}\right) & x > x_0+t, \end{cases} \tag{7.5}$$

where A is a scale parameter; x_0 is a peak-position parameter; t is a transition parameter; and σ_{up_0} is a resolution parameter that indicates the spread of the distribution in higher side. For this function $f_0(x)$, the FWHM is given as

$$\mathrm{FWHM} = \sqrt{2\ln 2}\sigma_{\mathrm{up}_0} - \frac{\ln 2}{t}\sigma_{\mathrm{up}_0}^2 - \frac{t}{2}. \tag{7.6}$$

By deconvolving the pedestal contribution in the spectrum, we can extract the intrinsic resolution of the detector that is not affected by a particular condition. To evaluate the practical resolution in normal muon beam, we then convolved the response function $f_0(x)$ with the pedestal distribution in normal muon run, $h_\mu(x)$. For the resolution, we quote here the effective sigma of higher energy side, $\sigma_{\rm up}$, of the response function $f_0(x) \otimes h_\mu(x)$ because it is the most important quantity for the background suppression, while the FWHM value is important when we think about the efficiency.

Since the resolution is position dependent, the response was extracted for each position. Figure 7.11 shows resolution maps for various positions of the gamma-ray interaction for $w_\gamma \geq 2$ cm. The maps show better resolution around the center of the detector, and the $\sigma_{\rm up}$-distribution shows a peak at around $\sigma_{\rm up} = 1.8$ %. However, since actual event distribution shows fewer events around the center because of the low efficiency of accompanied positron, the averaged resolution over the actual event distribution is $\sigma_{\rm up} = 2.0$ % for $w_\gamma \geq 2$ cm. The relative uncertainties of resolution evaluation are estimated to be $10 \sim 14$ % from the errors of fitting and the variation of the resolutions in close positions. The resolutions are measured separately for the events with $0 \leq w_\gamma < 1$, and $1 \leq w_\gamma < 2$ cm because the resolution heavily depends on the depth of interaction when interaction occurs in shallow region. The average resolutions are 3.0 and 4.2 % for $1 \leq w_\gamma < 2$ and $0 \leq w_\gamma < 1$ cm, respectively. Note that those numbers are representative ones and in the likelihood analysis position-dependent responses are used.

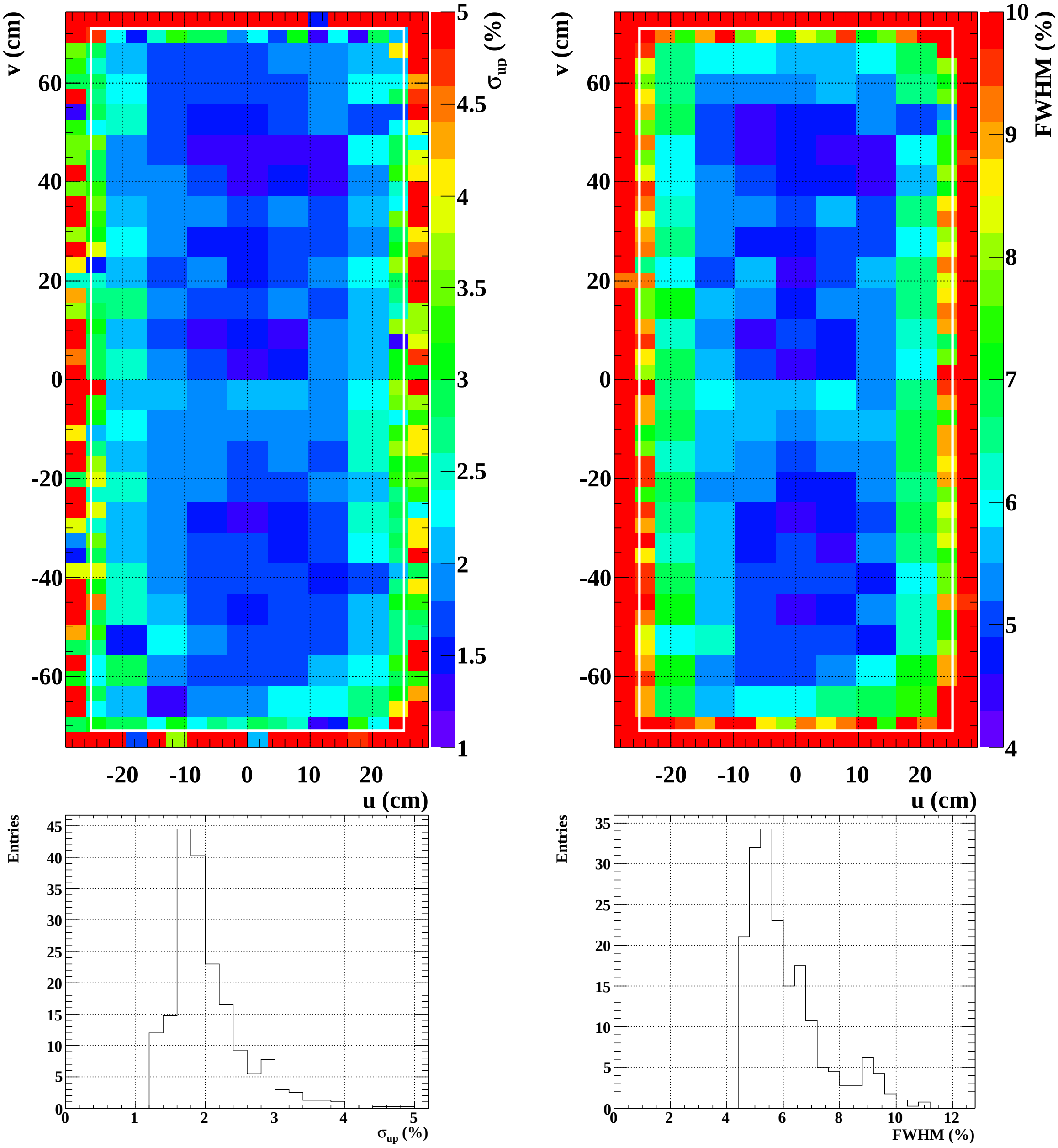
Validation Check of the Analysis with Spectrum from Muon Decays

We can also use the gamma-ray background spectrum to evaluate the energy resolution and the energy scale. The gamma-ray background comes mainly from RD and from positron AIF especially in high-energy region. Those rates are rapidly suppressed as their energy increase as discussed in Sec 2.3.4. The steepness of the background suppression indicates the energy resolution. The background spectrum was calculated with the MC simulation. It was formed with detector resolution, pileup component, and a scale factor. The measured spectrum was fitted with the MC spectrum with taking the resolution, pileup component fraction, and scale parameters as free parameters. We performed this fitting separately for the different positions. Figure 7.12(a) shows an example of the fitting. Comparisons between the parameters estimated with the π^0 data and those with the fitting are shown in Figure 7.12(b)(c). The consistent results validate the analysis of energy measurement. The fit results of energy-scale parameter shown in (c) shows small bias of ~0.5 %. When the scale estimated in this way is plotted as a function of the time, we observe slight increase of 0.25 % for whole period. This bias is used when we consider the systematic uncertainty of the energy scale.

Systematic Uncertainty on Energy Scale

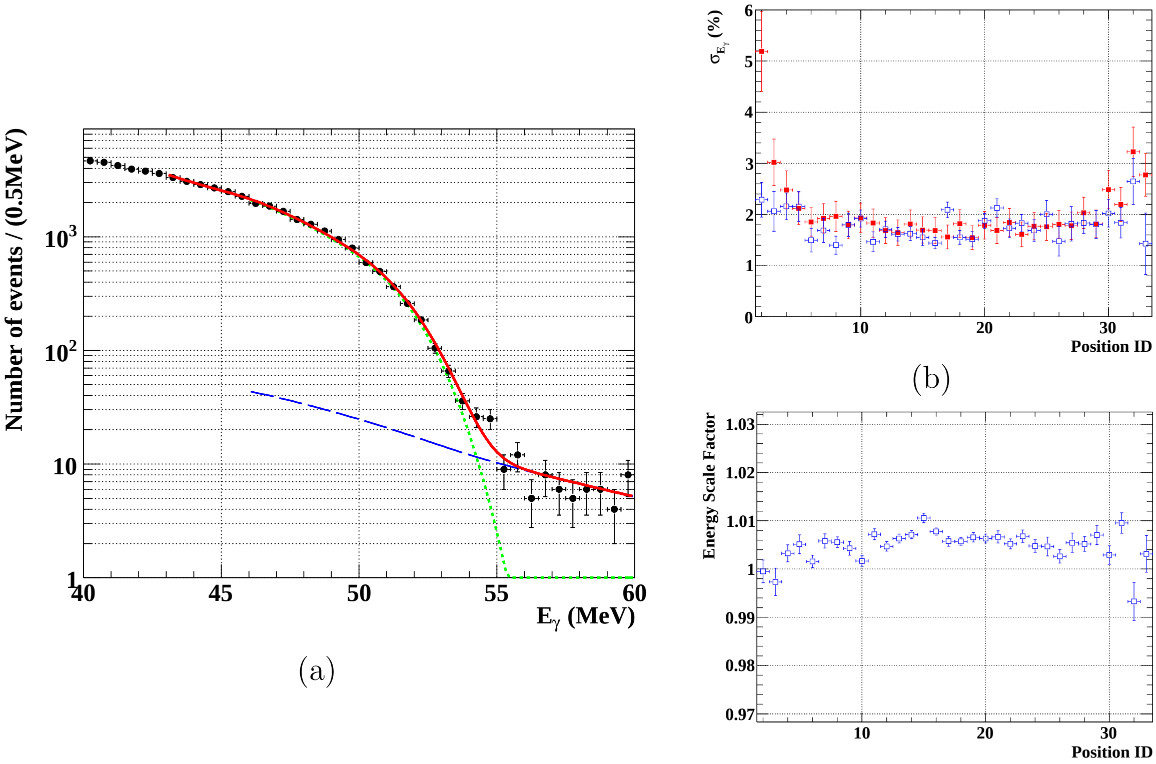
The factors of the systematic uncertainties on E_γ scale are listed as the following:

- uncertainty of 55 MeV-peak estimation in the π^0 run: 0.08 %;
- uncertainty of the gain shift during the the π^0 run: ~ 0.2 %;
- uncertainty of the light-yield connection between π^0 run and physics run, mainly stemming from one of cosmic-ray peak estimation: 0.18 %;



(a) (b)

Figure 7.11: Distribution of the gamma-ray detector energy resolution for $w_\gamma \geq 2$ cm, (a) high energy side sigma and (b) FWHM. Upper two-dimensional plots show the resolution map as a function of interaction position. White boxes in the maps indicate the acceptance. Lower histograms shows the distribution of resolution at each point inside the acceptance. One entry corresponds to an area of one PMT.



(a)

(b)

(c)

Figure 7.12: Gamma background spectrum fit. Fitting function is formed by folding MC spectrum (green line for single gamma and blue line for pileup component) and detector resolution. Energy scale, resolution and pileup fraction are free parameters and estimated by the fitting. (b) and (c) show the results. (b) Comparison of energy resolution estimated by π^0 data (red filled squares) and background fitting (blue blank squares). (c) Energy scale relative to that determined by the 55 MeV peak in π^0 run.

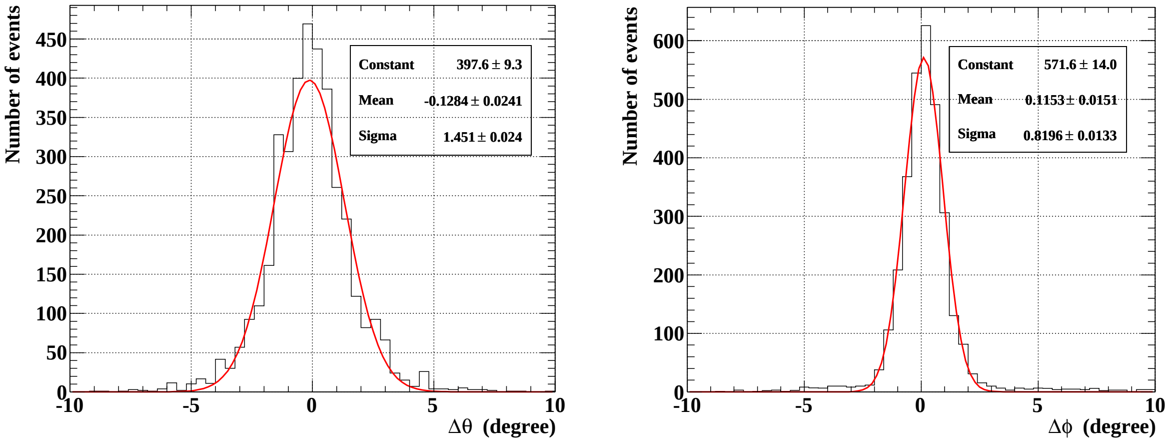
- variation of energy scale estimated by the fitting of sideband data: 0.25 %;

In total, we assign 0.4 % uncertainty. This estimation is consistent with the systematic difference in the sideband fit result of $\sim$ 0.5 %. The systematic error of the uniformity corrections with three times lower energy gamma (17.6 MeV) and the possible time variation due to the purification are included in those systematic uncertainties.

7.4 Relative Angle, $\theta_{e\gamma}$ and $\phi_{e\gamma}$

7.4.1 Positron Emission Angle

The positron angular resolution was evaluated by exploiting tracks that make two turns in the spectrometer, where each turn is treated as an independent track. The θ- and ϕ-resolutions were extracted separately from the differences of the two track-segments at the point of closest approach to the beam-axis. The distributions of the differences are plotted in Figure 7.13. The resolutions were estimated with the sigmas of fitted Gaussians divided by $\sqrt{2}$. They are $\sigma_{\theta_e} = 18$, $\sigma_{\phi_e} = 10$ mrad. Those estimations do not include the effect of scattering in the target. However, it is found to be negligible compared to the



(a) (b)

Figure 7.13: Difference of positron directions reconstructed with each turn in (a) θ and (b) ϕ directions.

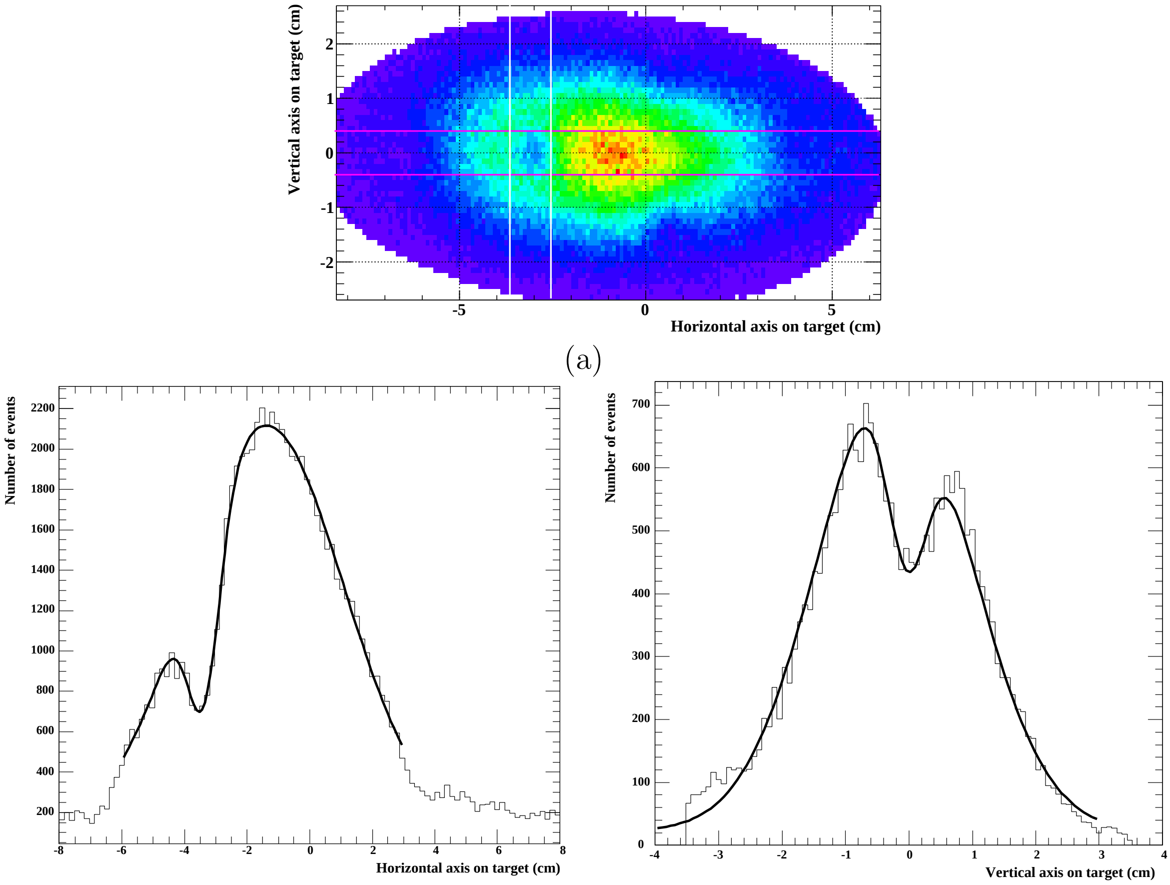
obtained resolutions. Uncertainty of the resolutions are estimated to be ~ 1 mrad by the comparison with the MC simulation and variation of the results by changing the cuts.

7.4.2 Muon Decay Vertex

The vertex position resolutions were measured by looking at the reconstructed edges of holes on the target. Figure 7.14 shows distribution of reconstructed vertex. We selected the region including a hole in horizontal and vertical directions as shown in the figure by the magenta and white lines. Projections of the slices are shown in (b) and (c) for horizontal and vertical directions, respectively. The shape of the hole image was fitted with a function formed with beam profile and vertex resolution. The vertex position resolutions are measured to be ~ 3.2 and ~ 4.5 mm in the vertical and the horizontal directions on the target plane, respectively. Those results were validated by another estimation using the two turn events in the same way as for the angular resolutions.

7.4.3 Gamma-ray Position

We measured the gamma-ray position resolution in the dedicated π^0 run with lead bricks placed just in front of the gamma-ray entrance window of the detector. The details of the brick are shown in Figure 4.4. Figure 7.15(a) shows the distribution of reconstructed positions in the run. It shows the image of the lead brick. We used peaks of the slits due to the shadow of the brick to evaluate the resolution. The projection of the distribution on v-axis was fitted with superimposition of three Gaussian (for slits peaks) , two Error functions (for brick edges), and a floor component (for the events penetrated the brick) as shown in (b). The average of the sigmas is $\sigma_{\text{data}} = 6.9$ mm. The width of the fitted Gaussian contains the width of slit itself and the contribution from the spread of π^0 decay point. A dedicated MC simulation with the same configuration was conducted to evaluate these effect. The same fit on the MC is shown in (c). The average of the sigmas in MC is $\sigma_{\text{MC}} = 6.5$ mm. We consider the difference $\sqrt{\sigma_{\text{data}}^2 - \sigma_{\text{MC}}^2} = 1.8$ mm gives an additional



(a)

(b) Projection on horizontal direction

(c) Projection on vertical direction

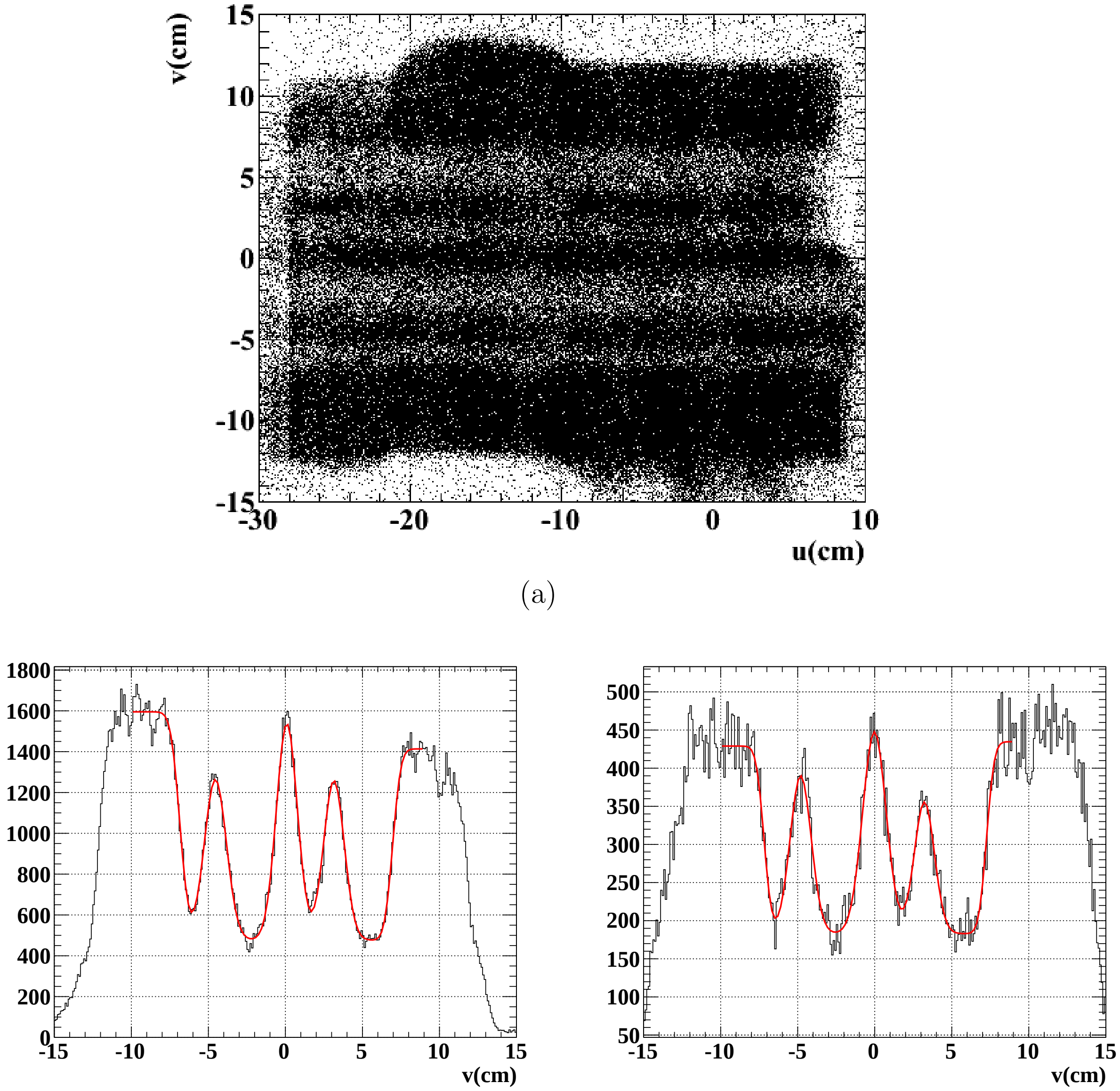
Figure 7.14: Target hole fit to evaluate vertex resolutions. (a) shows the distribution of reconstructed vertex. (b) and (c) show the projections of slices given by magenta and white lines in (a). Solid lines show the fitted function formed with beam profile and vertex resolution.

contribution to the resolution stemming from errors of the PMT QEs' measurements.

The bias of reconstructed position was estimated by looking at the absolute positions of reconstructed slits. The bias is 0.7 mm in standard deviation. It also stems from the errors of QEs. We take this bias into the response as an additional smearing because we could not measure the bias for all position and thus we cannot correct them.

We investigated detail response of the position measurement to the signal events using MC simulation with taking into account the difference from the data. Responses obtained from the MC are shown in Figure 7.16. In u- and v-direction, the responses can be expressed by double-Gaussians. The tail component comes from events with large spread of shower or back scattered events. The core component has sigma of $\sim$ 4 mm with about 70% fraction, while the tail component has sigma of $\sim$ 8 mm with about 30% fraction. The resolutions depend on the relative position to the PMT alignment. Figure 7.17 shows the dependence of the resolutions on distance from the center position of the nearest PMT. We can measure the position with better resolution when the interaction occurs at the mid of PMTs, because the light is effectively distributed to several PMTs. The effective resolutions are $\sim$ 5 cm for both u- and v-directions on average. The uncertainties are estimated to be 0.3 mm for both directions.

The response in w-direction has asymmetric shape. We quote here the half width of the distribution which contains 68 % of events as a resolution. The effective resolution is $\sim$ 6 mm. The uncertainty is estimated to be 0.7 mm.

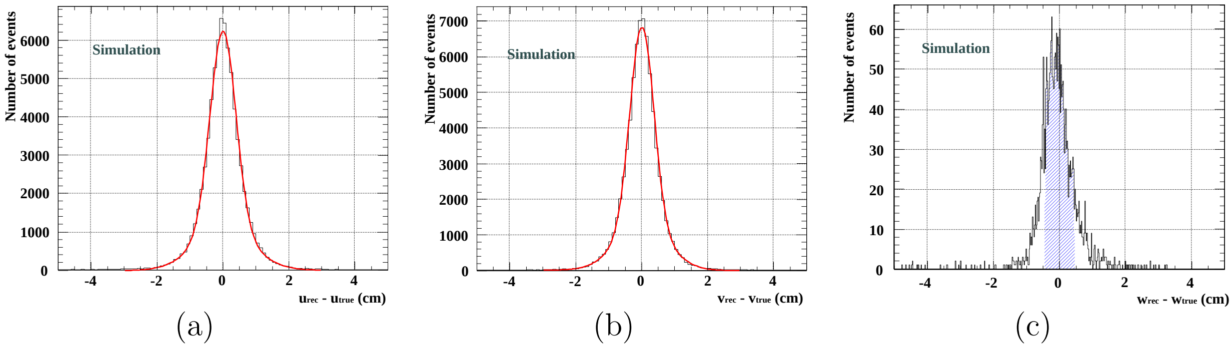


(a)

(b) data

(c) MC simulation

Figure 7.15: Event distribution of π^0 run with a Pb brick. (a) shows scatter plot of the reconstructed position of the data. (b) shows the projection on v-axis for the data, and (c) shows that of the MC simulation.



(a) (b) (c)

Figure 7.16: Position response of the gamma-ray detector obtained from the MC simulation in (a) u-, (b) v-, and (c) w-directions.

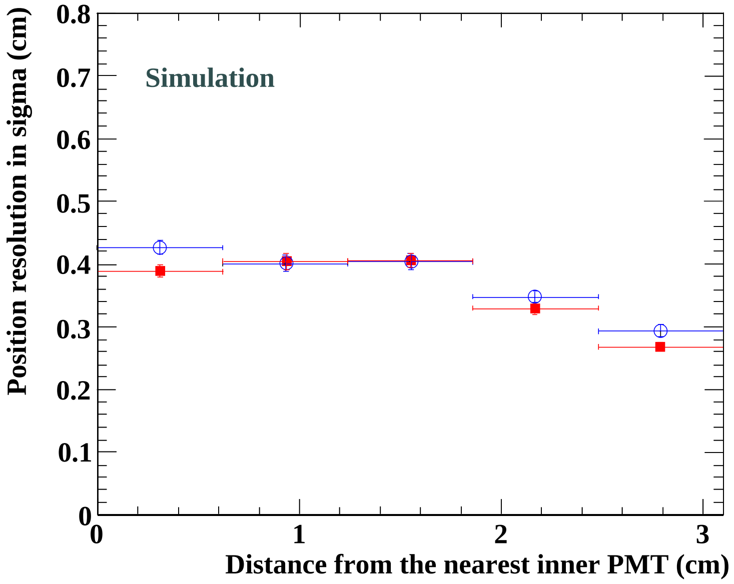


Figure 7.17: Position resolution as a function of the distance from the nearest inner PMT center in u (red square), and v (blue circle) directions. Resolutions quoted here are value of sigma of core component Gaussian.

7.4.4 Combined Resolution

The angular resolutions of the gamma direction were evaluated to be $\sigma_{\theta_\gamma} = 9.9$ and $\sigma_{\phi_\gamma} = 9.2$ mrad on average including the uncertainty of vertex position. We see here worse resolution in θ-direction because both u and w measurements contribute to the θ-direction whereas only v measurement does for ϕ-direction. The resolutions of relative angle in θ- and ϕ-directions were then obtained by combining the angular resolutions of the two particles. They are evaluated to be

$$\sigma_{\theta_{e\gamma}} = 21, \quad \sigma_{\phi_{e\gamma}} = 14 \text{ mrad}. \tag{7.7}$$

There is a significant difference in the responses between the two directions. Therefore, we separately treat the two directions in the likelihood analysis.

7.5 Relative Time, $t_{e\gamma}$

7.5.1 Positron Time

We measured the resolutions of the timing-counter bars by looking at the distributions of time difference between two consecutive bar's hits. The results are shown in Figure 7.18. Resolutions indicated by arrow markers were estimated using the two bars read out with different DRS chips. Thus those resolutions include the chip-synchronization accuracy. On the other hand, the others were read out with a single chip for the two bars so that we can measure the bar's intrinsic resolutions. The resolutions were measured at two different period of the run, and found to be stable. One point with worse resolution on bar 15 is due to the defect of a DTD channel described in Sec 4.2.5. After exchanging the DTD channel, the resolution was recovered (from black marker to red one). The average resolution of single bar is estimated to be $\sigma_{\mathrm{1bar}} = 67$ ps, where the contribution of clock synchronization ($\sigma_{\mathrm{clock(DRS3)}} = 37$ps) is subtracted (intrinsic resolution). If it is included, the resolution becomes 76 ps. For the positron with multiple hits we can improve the intrinsic resolution to 54 ps by combining the hits.

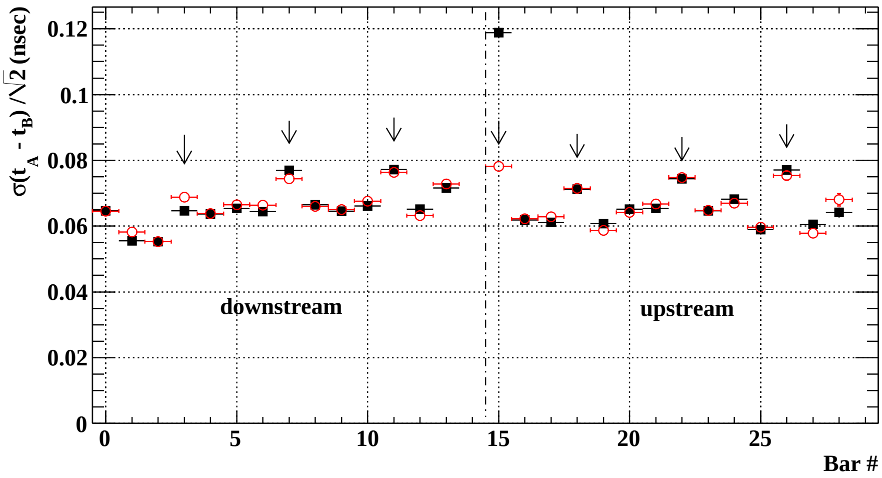


Figure 7.18: Time resolution of timing-counter bars. They are evaluated at two different period of the run; September (black square) and November (red blank circle).

The timing resolution strongly depends on the quality of the DCH-TIC interconnection. We investigated the dependence using the π^0 Dalitz decay data. Figure 7.19 shows the result. As the value of $|\Delta z_{DCH-TIC}|$ increases, the resolution becomes worse. Positrons with large values of $|\Delta z_{DCH-TIC}|$ were probably scattered between the drift chamber and the timing counter. Thus the measurements of the track length were screwed up. The Dalitz decay has a similar topology to the $\mu^+ \to \mathrm{e}^+\gamma$ decay but the conditions such as target material, vertex distribution, positron and gamma-ray energy, and angular correlation are different. Here, we assume those differences contribute independently as a quadratic sum in the resolution, and the dependence on $|\Delta z_{DCH-TIC}|$ is common to positrons from muon decays in the physics data taking. With this assumption, we extracted the $|\Delta z_{DCH-TIC}|$ dependence of positron time resolution. It is shown as a red plot in the figure. At the edge of $|\Delta z_{DCH-TIC}| = 0$, the positron time resolution is

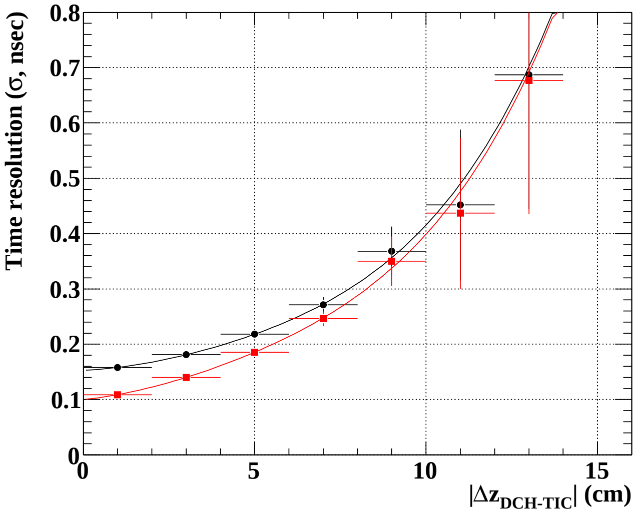


Figure 7.19: Dependence of time resolution on $|\Delta z_{DCH-TIC}|$. Black plot shows the measured dependence of the Dalitz data. Red one shows estimated resolution for the signal positrons. Those plots are fitted with fourth order polynomial functions.

(100 ± 13) ps.

The accuracy of the time-of-flight measurement cannot be measured independently. Hence, we estimated it with the MC simulation tuned to reproduce the conditions of the drift chamber. It is evaluated to be $\sigma_{eToF} = 88$ ps for signal positron with $|\Delta z_{DCH-TIC}| < 6$ cm.

7.5.2 Gamma-ray Time

Intrinsic Resolution

To investigate 'intrinsic' time resolution of the gamma-ray detector, we performed a divided analysis. We define two PMT groups, even- and odd-groups. PMTs are assigned to the two alternately, and each group contains half number of PMTs as shown in Figure 7.20. The gamma-ray interaction time was reconstructed independently for each group only using PMTs belonging to the group. The intrinsic time resolution was evaluated by the distribution of half of time difference between the two reconstructed time,

$$\frac{\Delta t^{\mathrm{int}}}{2} = \frac{t_{\mathrm{odd}} - t_{\mathrm{even}}}{2}, \tag{7.8}$$

whose uncertainty is statistically same as that of average of the two, $(t_{\mathrm{odd}} + t_{\mathrm{even}})/2$, which is equivalent to t_{LXe}. The distribution for 55 MeV gamma events in August π^0 run is shown in Figure 7.21. It shows the intrinsic time resolution of $\sigma_{t\gamma}^{\mathrm{int}} = 45$ ps at 55 MeV.

The position dependence of the resolution was investigated along u, v, and w. The results are shown in Figure 7.22. The resolution is gradually improved as the interaction occurs closer to lateral faces and suddenly becomes worse. It can be understood that as closing in one of lateral faces, the total number of photoelectrons increases owing to the PMTs on the lateral face. When it gets too close to the edge, then the shower escape dominantly works to reduce the number of photoelectrons resulting in worse resolution.

Part of the worse resolution at the shallow part is also explained by the decrease of the photoelectrons. In addition, there is another contribution. When interaction occurs near the inner face, then the observed number of photoelectrons in each group heavily depends on the interaction position. This disproportion of the observed number of photoelectrons deteriorates the evaluation by the distribution of $(t_{\mathrm{odd}} - t_{\mathrm{even}})/2$ because this evaluation assumes equal resolution of the two groups.

The intrinsic resolution becomes 36 ps at 83 MeV. Figure 7.23 shows the dependence of the intrinsic resolution on the total number of photoelectrons and incident gamma energy. They shows improvement as the number of photoelectrons increases. The intrinsic time resolution is completely dominated by the photoelectron statistics.

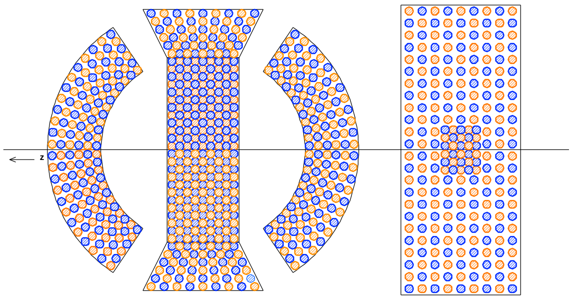


Figure 7.20: PMT grouping scheme for the intrinsic time resolution study. PMTs indicated with blue circles belong to odd group, and those with orange circles belong to even group.

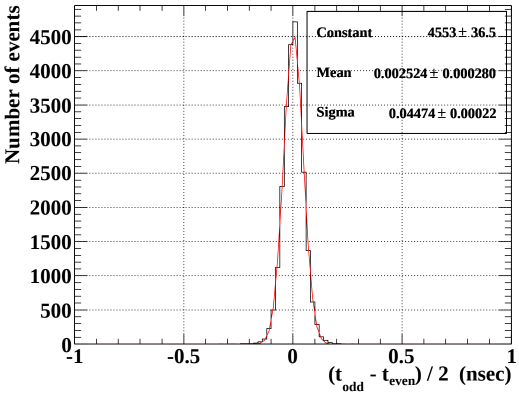
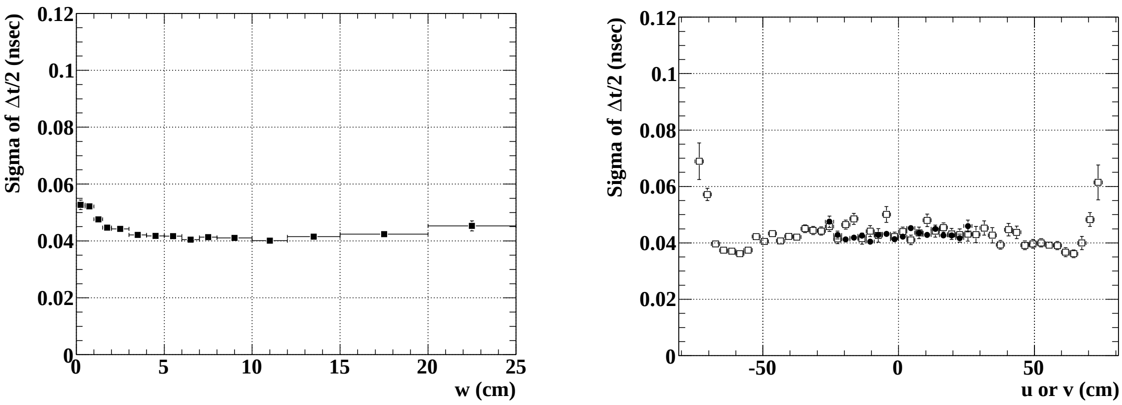


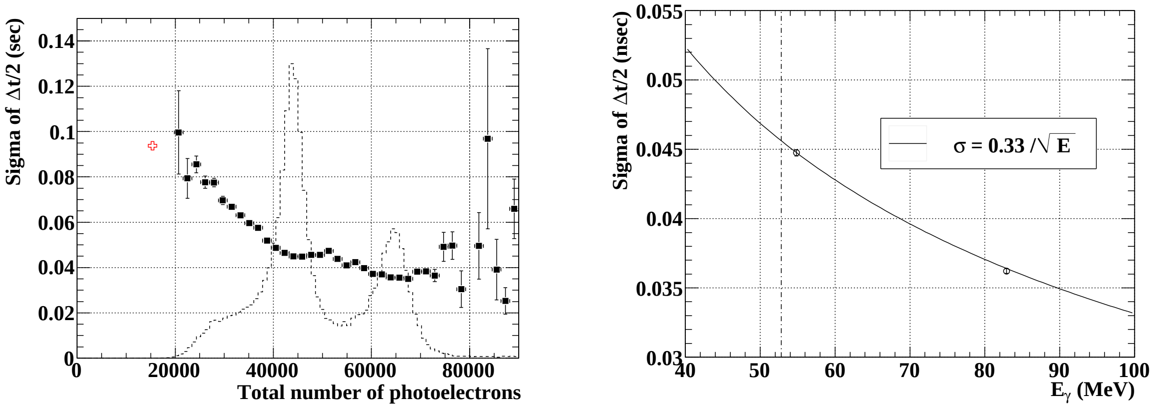
Figure 7.21: Distribution of half of time difference of two group reconstructed time. The intrinsic time is estimated to be 45 ps in sigma at 55 MeV.



(a)

(b)

Figure 7.22: Intrinsic time resolution at different position, (a) as a function of w_γ and (b) as a function of u_γ (black circle) and v_γ (blank square).



(a) (b)

Figure 7.23: Intrinsic time resolution as a function of total number of photoelectrons (a) and energy (b). The red cross mark in (a) shows the resolution estimated in the CW run at 17.6 MeV.

Practical Resolution

We measured the practical time resolution of the gamma-ray detector using $\pi^0 \rightarrow \gamma\gamma$ data. We used events where one of the two gamma rays was measured by the plastic scintillation counters mounted on the NaI detector as a reference counter.

Figure 7.24 shows the distribution of the time difference between the 55 MeV gamma ray at the gamma-ray detector and the 83 MeV one at the reference counter,

$$\Delta t^{\text{abs}} = t_\gamma - t_{ref}, \tag{7.9}$$

where the time-of-flight of the two gamma rays are already subtracted with an assumption of the π^0 decay vertex at the origin. The spread of the distribution includes contributions from the uncertainty of the vertex position and the time resolution of the reference counter. The contribution from the vertex uncertainty is evaluated to be 58 ± 2 ps from the spread of the π^- beam on the target. The reference counter resolution is evaluated to be 93 ± 7 ps by looking at the time difference between the two PMTs on a plate. By subtracting these contributions, the practical time resolution is evaluated to be $\sigma_{t\gamma}^{\text{abs}} = 135 \ominus 58 \ominus 93 = 78$ ps at 55 MeV.

At 83 MeV, the sigma of Δt^{abs} distribution is 127 ps. In the second π^0 run in December, the sigma of distribution for 55 MeV gamma ray is improved to 130 ps owing to the increase in the LXe light yield. Using these measurement points, the energy (more precisely photoelectron) dependence is extracted. Figure 7.25 shows the plot and extracted dependence-curve. Here the December data is plotted at an energy to which the mean number of photoelectrons corresponds in August. The dashed curve shows the practical time resolution after subtracting the other contributions. From this curve, the time resolution for the signal is evaluated to be

$$\sigma_{t\gamma}^{\text{abs}} = 80 \pm 6 \text{ ps}. \tag{7.10}$$

We consider the difference from the intrinsic resolution stems from errors of position reconstruction and the spread of shower which can be effectively canceled out in the evaluation of the intrinsic resolution. In particular, the remaining constant term of 24 ps would be contribution of the depth reconstruction error which directly deteriorates the time-of-flight measurement.

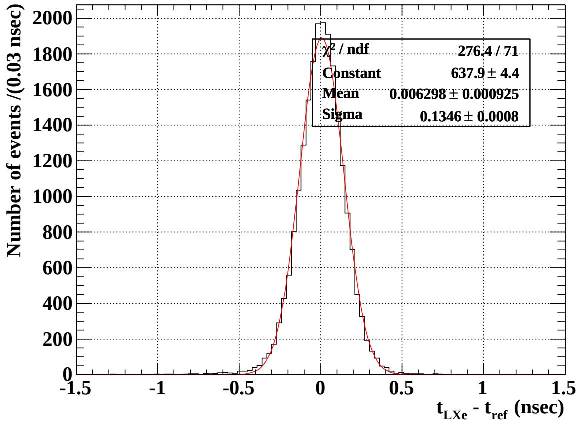


Figure 7.24: Distribution of time difference between the gamma-ray detector and the reference counter.

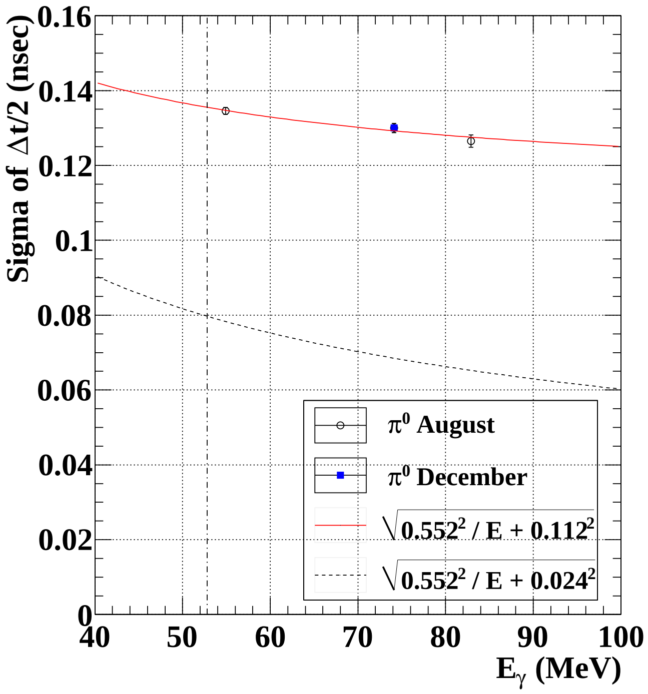


Figure 7.25: Energy dependence of the gamma-ray detector timing resolution. Plot with blank circle and blue filled square show the sigma of the time difference distribution in August and December π^0 runs, respectively. Dashed line shows the gamma-ray detector resolution after the other contributions are subtracted. Dot-dashed vertical line shows the signal energy.

7.5.3 Gamma-Positron Relative Time

The overall time resolution between positron and gamma ray was measured with the RD events. Since the RDs are time coincident events, they make a peak in time distribution on top of the flat distribution of accidental background. We used the E_γ-sideband data of physics run. The event topology is very close to that of $\mu^+ \to e^+\gamma$ event, and the analysis and run conditions are common to those for signal. The time distribution is shown in Figure 7.26. It was fitted with a Gaussian plus flat floor. With taking into account the different gamma-ray energy, the time resolution for the signal is evaluated to be

$$\sigma_{t_{e\gamma}} = 148 \pm 17 \text{ ps}, \tag{7.11}$$

where the uncertainty includes the errors of fitting and extrapolation to the signal energy. This resolution is consistent with the combination of individual components of resolutions.

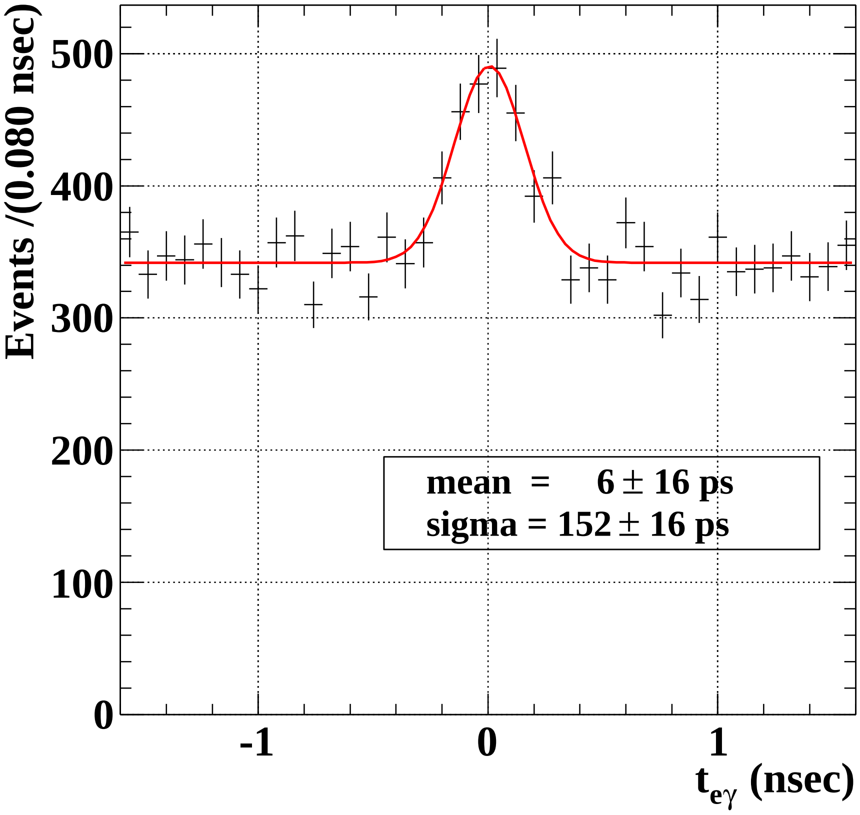


Figure 7.26: Distribution of $t_\gamma - t_e$ in physics run for $40 < E_\gamma < 45$ MeV. It shows clear radiative decay peak on top of the accidental background floor.

7.6 Positron Detection Efficiency, ϵ_e

The probability that a positron from $\mu^+ \to \mathrm{e}^+\gamma$ decay would be measured is given by the following joint probability

$$p(\text{acceptance}, \text{trigger}, \text{tracking}, \text{TIC hit}, \text{DCH} - \text{TIC match}, \text{selection}). \tag{7.12}$$

Here, we investigate the positron detection efficiency ϵ_e which is defined as the probability of measuring a positron emitted in the acceptance defined in Sec.7.1.1. The positron detection efficiency can be decomposed into the two component; the tracking efficiency in the drift chamber and the probability of getting a hit on the timing counter. We define them more precisely with conditional probabilities,

$$\begin{aligned} \epsilon_e &= p(\text{good track} \mid \text{acceptance}) \times p(\text{matched TIC hit} \mid \text{good track}, \text{acceptance}) \\ &= \epsilon_{DCH} \times A_{TIC}, \end{aligned} \tag{7.13}$$

where 'good track' requires the selection criteria for the final analysis; 'matched TIC hit' requires at least one hit on the timing counter fired by the trigger and matched with a reconstructed good track.

We use Michel positrons taken in the physics run as a control sample. Only the difference from positrons from $\mu^+ \to \mathrm{e}^+\gamma$ decays is their energy. By selecting high-energy positrons ($E_e > 50$ MeV), we can reduce the difference, and finally a correction will be applied using the measured energy dependence.

7.6.1 Conditional Timing-Counter Efficiency

The conditional timing-counter efficiency A_{TIC} was measured with data sample of the DCH-self trigger. We selected the good-track events in the sample without any reference to the timing-counter information. Here, we only applied a t_{track} cut to remove tracks which are out of timing-counter digitization time because the digitization window of the drift chamber is wider than that of the timing counter. Then the ratio between the number of matched good tracks and the number of good tracks gives the probability. This ratio is plotted in Figure 7.27 as a function of the positron energy. The average probability for Michel positrons with $E_e > 50$ MeV is 34 %. By estimating the end-point of the probability curve, the probability for the signal positron is evaluated to be $A_{TIC} = 38$ %. The inefficiency stems from positrons which do not hit the timing counter by changing their trajectories from the expectations because of the energy loss, multiple scattering, hard scattering, or annihilation by interacting with material in the spectrometer such as chamber frame, preamplifiers, and cables.

7.6.2 Drift-Chamber Tracking Efficiency

The drift-chamber tracking efficiency ϵ_{DCH} was evaluated with data sample of the TIC-self trigger. The number of observed Michel positrons can be written as

$$N_{obs} = N_\mu \times f(E_e > 50\ \text{MeV}) \times \frac{4\pi}{\Omega} \times \epsilon_{DCH} \times A_{TIC} \times \frac{1}{P}, \tag{7.14}$$

where N_μ is the number of muons stopping on the target, and we can calculate the number with the stopping rate $R_{\mu,\text{stop}}$ given by Eq.3.1 and the total amount of proton current in

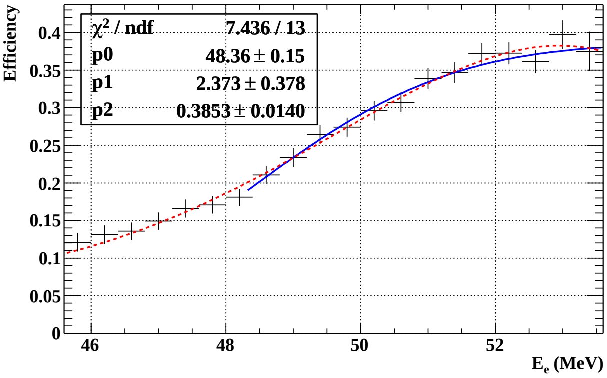


Figure 7.27: Conditional timing-counter efficiency curve. The efficiency for the signal is estimated by the end-point of the curve. Blue solid line shows the fitted error function, and red dashed one shows the fitted 4-th polynomial function.

the livetime; $f(E_e > 50\text{ MeV})$ is a fraction of Michel spectrum over 50 MeV calculated to be 0.101; the solid angle of acceptance is $4\pi/\Omega = 0.1$; and P is the pre-scale factor of the TIC-self trigger of 10^7. With this relation, the positron detection efficiency is evaluated to be 13.5 % for Michel positrons over 50 MeV. The energy dependence is almost flat at the high-energy region and the correction factor is extracted to 1.02 for the signal energy. Thus the efficiency of signal positron is evaluated to be

$$\epsilon_e \approx 14\ \%. \tag{7.15}$$

With the evaluated A_{TIC}, the drift-chamber tracking efficiency is evaluated to be $\epsilon_{DCH} \approx 37\%$. The inefficiency stems from the dead layers and low voltages due to the discharge problem.

7.7 Gamma-ray Detection Efficiency, ϵ_γ

The gamma-ray detection efficiency was estimated using the MC simulation, and confirmed by the measurement of π^0 events and gamma-ray background rate in physics run.

Figure 7.28(a)(b) shows the distribution of energy deposit in LXe for 52.8 MeV monochromatic gamma ray. The tail comes from interactions with material in front of the active volume and shower escape from the inner face. Figure 7.28(c) shows the radius distribution of first interaction point of the gamma rays. The contribution of the inefficiency is summarized in Table 7.3. The detection efficiency is given in Figure 7.28(d) a function of the energy cut threshold. For the analysis window ($E_\gamma > 46$ MeV), the efficiency becomes 67 %. The efficiency was evaluated at each position of the detector and average efficiency was calculated with a weight of actual event distribution of positron direction. Then it is evaluated to be 66 %.

The efficiency was measured using NaI-self trigger data in the π^0 run. By tagging a gamma around 83 MeV from π^0 decay, we can get the condition that a gamma around 55 MeV is emitted to the detector. We can calculate the efficiency by looking at the measured spectrum shown in Figure 7.29. The efficiency is estimated to be (68 ± 2.5) % for $E_\gamma > 46$ MeV, where the uncertainty comes from the evaluation of the radiative

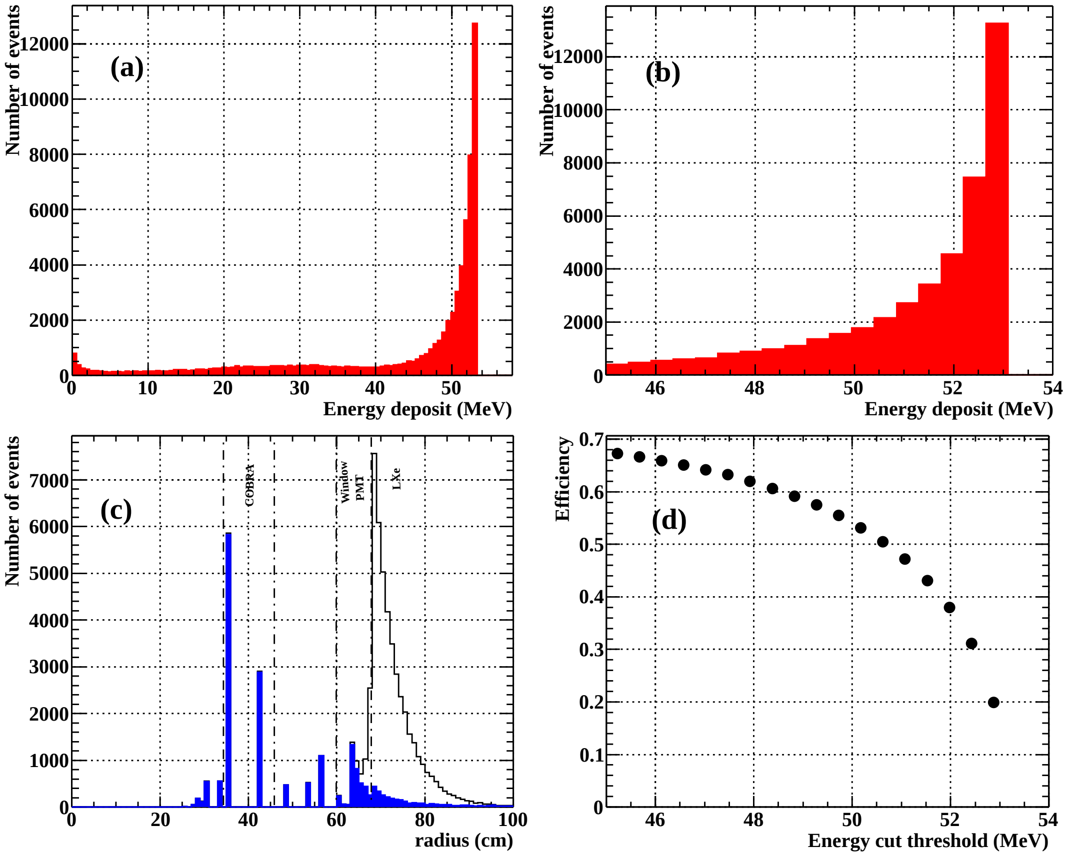


Figure 7.28: Evaluation of gamma-ray detection efficiency with the MC simulation. (a) Distribution of the energy deposit of 52.8 MeV gamma generated in the acceptance direction. (b) Close-up view for the analysis region. (c) Distribution of radius position of the first interaction. Blue filled histogram shows that for the event $E_{\mathrm{dep}} < 46$ MeV, and blank one for all events. (d) Efficiency curve as a function of the energy cut threshold value.

Table 7.3: Inefficiency of gamma-ray measurement.

	Inefficiency (%))
Drift-chamber frame, support structure	2.5
COBRA magnet	14.2
Entrance window, PMT, support holder	7.3
Other materials	4.0
Shower escape	6.0
Selection	5.6 [1]
Total	37

---

[1] For the detected events.

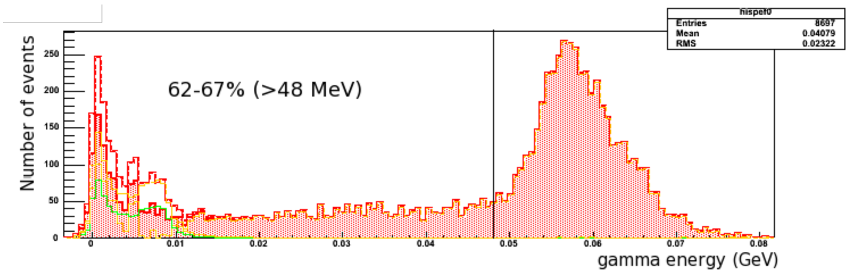


Figure 7.29: Evaluation of gamma-ray detection efficiency with π^0 data. Different spectra at low energy show different estimations of neutron events.

capture reaction, $\pi^- \mathrm{p} \rightarrow \gamma \mathrm{n}$, which emit a 129 MeV gamma. This consistent results validates the efficiency evaluation with the MC above.

In addition, the gamma-ray background rate was measured and compared to the expectation from the MC simulation. As will be discussed in Sec.8.2.1, the rate was agreed in ∼5 % relatively, while the expectation has a relative uncertainty of ∼7 % coming from the beam rate uncertainty and evaluation of the contribution from the muon decay off target.

The cuts of pileup and cosmic-ray events give 5.6 % inefficiency. Thus in total, the gamma-ray measurement efficiency is given as

$$\epsilon_\gamma = (63 \pm 4)\ \% \tag{7.16}$$

7.8 Summary of Performance in 2008

The summary of the detector performance is given in Table 7.4.

Table 7.4: Summary of detector performance. Resolutions are given in sigma.

	Resolution (σ)
Positron energy	0.70 %
Positron emission angle (θ, ϕ)	(18, 10) mrad
Muon decay vertex	3-4.5 mm
Timing-counter time	55-70 ps
Gamma energy	2.0 %
Gamma position (u, v and w)	(5, 5, 6) mm
Gamma timing	80 ps
$t_{e\gamma}$	148 ps
($\theta_{e\gamma}, \phi_{e\gamma}$)	(21, 14) mrad
	efficiency
Positron efficiency	14 %
Gamma efficiency	63 %

Chapter 8

$\mu^+ \to e^+\gamma$ Search Analysis

8.1 Overview of Analysis

The goals of this chapter are

- to estimate the best value of the number of $\mu^+ \to e^+\gamma$ decay events in the data;
- to estimate the interval of that value;
- to calculate the observed number of muon decays;
- and finally to calculate the best-estimated value of the branching ratio and its interval.

To measure the branching ratio on $\mu^+ \to e^+\gamma$ decay, we need to extract the number of $\mu^+ \to e^+\gamma$ events in the data sample and the total number of muon decays that we observed. The latter is measured by counting Michel decays in the data sample. For the former, we perform a likelihood analysis. We actively detect the signal events instead of passively counting the remaining events by cutting background candidates. A confidence interval of the measured branching ratio is then assigned with a frequentist approach.

In addition to the likelihood analysis, we also perform cut-counting analyses to understand background and to check the analysis. It is difficult to see the background level in the likelihood analysis because the analysis window is set relatively wide so that the background distribution is fitted. Therefore, we conveniently define signal boxes to see the background level.

To avoid introducing bias in analysis, we adopted a blind analysis. It is often used in rare-decay searches in high-energy physics (for example, see a review [92]). We blinded ourselves by a simple hidden-box method. We masked events in the region where $\mu^+ \to e^+\gamma$ events might appear in $(t_{e\gamma}, E_\gamma)$ two-dimensional plane. We calibrated the detectors, selected events, evaluated the performance, and estimated the background without looking at events inside the hidden box.

8.1.1 Pre-selection

To reduce the data size and to analyze efficiently, first we applied a pre-selection for all events of the MEG trigger data. Full reconstruction was not done in this stage, but some fundamental kinematics were reconstructed. The point of pre-selection is not to discard any good events.

The selection criteria are the following,

i) $-6.9 < t_\gamma - t_{TIC} < 4.4$ nsec

ii) $|t_{track} - t_{TIC}| < 50$ nsec,

where t_γ is the gamma emission time reconstructed with an assumption that the muon decay vertex is at the origin; t_{TIC} is timing-counter hit timing and the time-of-flight is not subtracted; Thus we did not use precise tracking information for the pre-selection. The reason of the asymmetric window for i) is to acquire the multi-turn events; because of the absence of the tracking, we do not know the number of turns before the hit on the timing counter. The second criterion requires at least one track associated with the trigger is found. Reduction power of the pre-selection is about factor 6.

8.1.2 Blinding

For the pre-selected event, we apply a blinding process, namely masking events in a pre-defined hidden box for the blind analysis. We defined the hidden parameters as a set of $(t_{e\gamma}, E_\gamma)$, and the hidden box as

i) $48 < E_\gamma < 57.6$ MeV

ii) $|t_{e\gamma}| < 1$ nsec,

Full reconstruction is required to efficiently hide the possible signal events. The box size was determined with the detector resolutions to fully cover the signal region. Selected events go to a separated data stream, and user cannot touch them. Only the unmasked events are opened to the user. We call these unmasked events sideband data. Thus, the following event selection criteria were tuned without looking at the data in the signal region but only with the sideband data. Event distributions on $(t_{e\gamma}, E_\gamma)$ plane after the pre-selection and blinding process is shown in Figure 8.1. No cuts are applied for events in the plot.

8.1.3 Definition of Analysis Window

We defined the analysis window to which the likelihood analysis is applied as the following:

- $50 < E_e < 56$ MeV,
- $46 < E_\gamma < 60$ MeV,
- $|\theta_{e\gamma}| < 100$ and $|\phi_{e\gamma}| < 100$ mrad,
- $|t_{e\gamma}| < 1$ nsec,

This window is not a so-called signal window. It includes background events to estimate the background distribution in a fitting. It is set wider than the hidden box so that we can study the fitting and the background before opening the box.

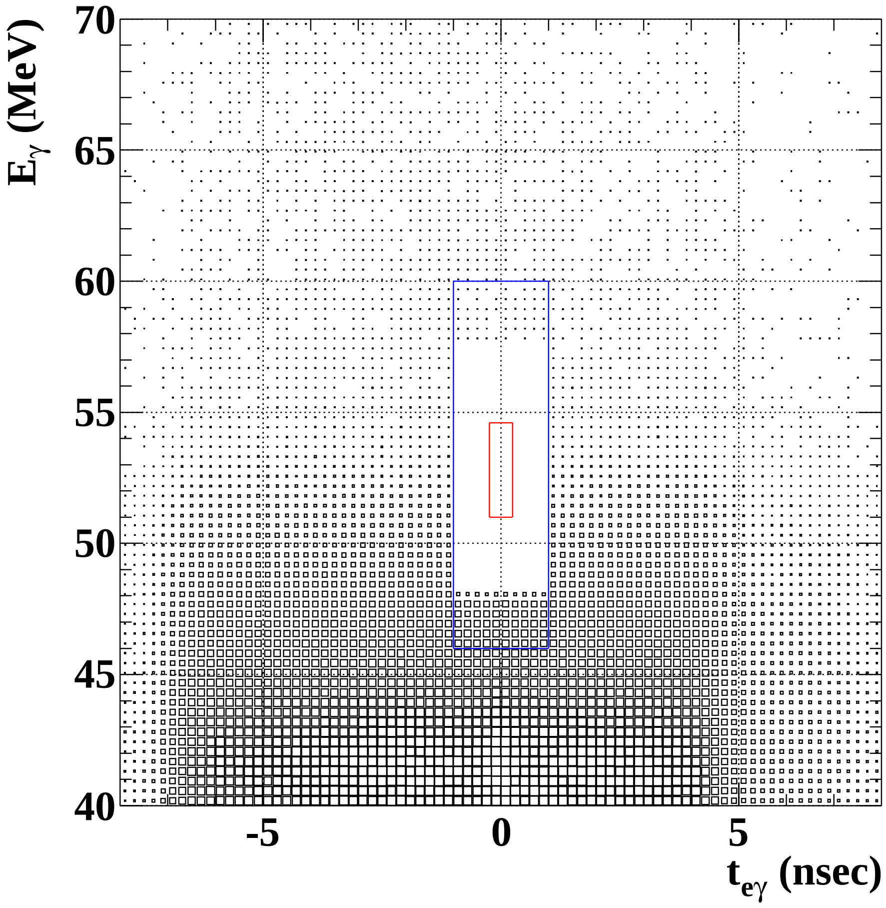


Figure 8.1: Distribution of $t_{e\gamma}$ v.s. E_γ after the pre-selection and blinding process. No cuts are applied. Events outside the pre-selection time window come from accidental overlap positrons. Analysis window for the likelihood analysis and signal box (A) are also shown as blue and red boxes, respectively.

8.1.4 Definition of Signal Boxes

To know our status and to understand backgrounds, we need to estimate the level of background in our data. For that purpose, here we define two signal boxes for convenience. One is a box based on the evaluated resolutions; the half width of each variable is basically given by 1.64 σ[1].

- $52.2 < E_e < 53.8$ MeV,
- $51.0 < E_\gamma < 54.6$ MeV,
- $|t_{e\gamma}| < 242$ ps,
- $\pi - \Theta_{e\gamma} < 33$ mrad,

where we use the quantity of $\pi - \Theta_{e\gamma}$ for the back-to-back condition instead of separated $\theta_{e\gamma}$ and $\phi_{e\gamma}$ for an efficient cut. We call this box "signal box A".

---

[1] For a Gaussian distribution, this width retains 90% efficiency.

The other is more optimized to the sensitivity,

- $52.0 < E_e < 56.0$ MeV,
- $51.0 < E_\gamma < 56.0$ MeV,
- $|t_{e\gamma}| < 273$ ps,
- $\pi - \Theta_{e\gamma} < 42$ mrad.

Asymmetric windows on E_e and E_γ are adopted because the background distributions have large asymmetries for the signal energy. The window for time and angles are then broadened so that the sensitivity becomes the highest considering the signal efficiencies and the number of expected backgrounds. We call this box "signal box B".

8.2 Background Estimation

8.2.1 Rate of Gamma-ray Background

In Sec 2.3.4, we roughly estimated the background and found the accidental background will be a dominant source. The rate of the accidental background depends on the beam intensity, time resolution, angular resolution, positron background spectrum, and gamma-ray background spectrum. In particular, the gamma-ray spectrum is the most important part for the suppression of the background. The theoretical spectrum of gamma from RD was used in the previous estimation. In reality, it depends on the gamma-ray yield depending on the actual design of the apparatus and its energy resolution.

A LXe-self trigger was mixed in the data taking. Using this data, we can measure the actual rate of the gamma-ray background in the detector. It is shown in Figure 8.2. The rate in signal box A is 34 Hz.

The expected spectrum by the MC simulation is superimposed in this plot. For the absolute rate, the MC spectrum has about 7 % uncertainty mainly from the uncertainty of beam rate and contribution from moun decay off target. In this plot, we used the muon beam rate measured at the center of COBRA, $R_{\mu,\mathrm{center}} = (3.69 \pm 0.08) \times 10^7 \mu^+/\mathrm{sec}$. Data and MC match well both in shape and in absolute rate. AIF is found to be a large source of high-energy gamma yield comparable with RD. For example, in signal box A, the fraction of RD is 51 %, that of AIF is 42 %, and that of pileup component is 8 %. The cosmic-ray contribution is measured to be 1 % from the data with no beam (see Figure 7.7). The origins of the AIF process are mainly the drift-chamber material (> 60 %) and that of the target.

8.2.2 Estimation of Background Level

It is difficult to define the expected number of backgrounds in the signal region in the likelihood analysis because we use the background distribution positively in the fitting and we use also worse measured events, that is, the event selection is not optimized to eliminate background events. However, to know our status and to understand backgrounds, we need to estimate the level of background in our data. We estimated the number of background events in the signal boxes defined in Sec.8.1.4 using the sideband data.

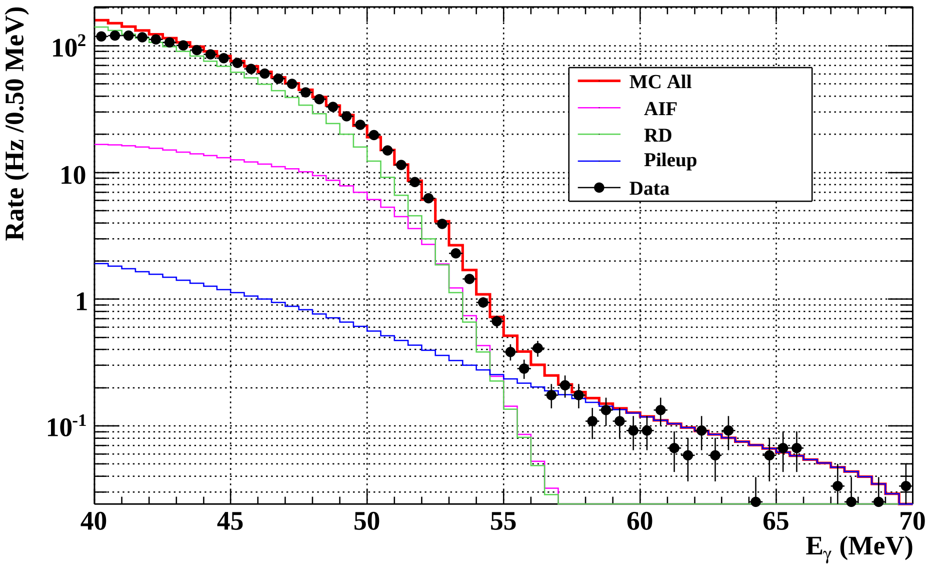


Figure 8.2: Absolute rate of gamma-ray background. It is compared with the MC expectation (green: gamma from radiative decay, magenta: that from positron annihilation-in-flight, blue: pileup component, and red: total). Discrepancy at the low-energy is due to a trigger effect.

Accidental Background Events

We use the $t_{e\gamma}$-sideband data to estimate the accidental background. The range is selected to $1 < |t_{e\gamma}| < 4$ ns not to be affected by the trigger. In addition, to increase statistics, we also use a wider range of $\pi - \Theta_{e\gamma}$ of 60 mrad with a correction of small trigger effect.

The event distribution in the sideband is shown in Figure 8.3. The number of events observed in the sideband window is 38 events. Taking the wider windows into account, the number of the accidental backgrounds in signal box A is estimated to be 0.95 ± 0.15, where the error is statistical uncertainty. That in signal box B is estimated to be 2.08 ± 0.03 in the same way.

The background level was estimated in another way. A maximum likelihood fit was performed on another sideband window that includes lower range of E_γ (> 46 MeV). From the probability ratio between the signal box and the sideband window, the expected number of the accidental backgrounds is estimated to be 0.8 ± 0.02, where statistical uncertainty is smaller owing to larger amount of statistics in low-energy region.

Radiative Decay Events

The expected number of RD events in signal box A (B) is estimated to be 0.02 ± 0.004 (0.03 ± 0.006) by calculating the partial branching ratios in the boxes and scaling the observed number of RD events in the E_γ-sideband by the ratio of the branching ratios. The detail description is given in Appendix A. The RD events is not a severe background source in this analysis.

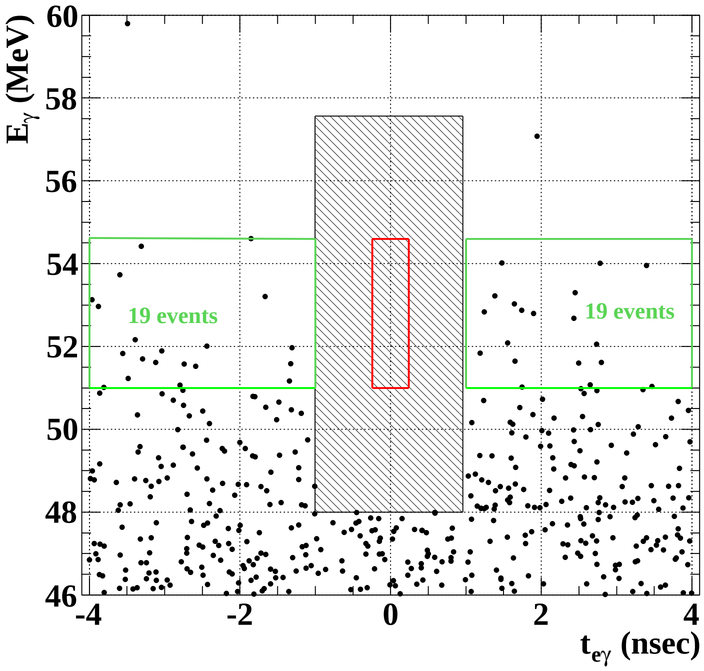


Figure 8.3: Event distribution in the $t_{e\gamma}$-sideband. A looser cut on $\pi - \Theta_{e\gamma}$ than that for the signal box is applied. Green boxes show the windows used for the background estimation. Red one shows the signal box. Hatched one shows the hidden box.

8.2.3 Distribution of Background Events from Sideband Data

Since the background is dominated by the accidental background, we can estimate distributions of the background in the analysis box by looking at those in the $t_{e\gamma}$-sideband. Figure 8.4 shows the $t_{e\gamma}$ distribution when we applied the analysis window cuts on variables except for $t_{e\gamma}$. The scatter plot of $t_{e\gamma}$ v.s. E_γ is shown in Figure 8.5. The time distribution is flat within the statistical fluctuation in this range. For a wider range, the time distribution is deformed by the trigger resolution. From this plot, we estimated the number of events in the analysis box to be 1192 events.

We obtained the distribution of E_e and E_γ from the sideband data where the cuts on the other energy and angles are relaxed. The modelings of the distributions are already described in Sec.7.2 and Sec.7.3 for E_e and E_γ, respectively.

We would expect flat distributions also for the angular distributions. However, acceptance and trigger effects deform the distributions. Figure 8.8 shows the observed angle distributions in the sideband. In particular, there was a bias in trigger conditions in ϕ-direction. It is considered due to the mismatch in the look up table, which was prepared using the timing z-counter while the ϕ-counter was actually used. We lost some trigger efficiency because of this bias.

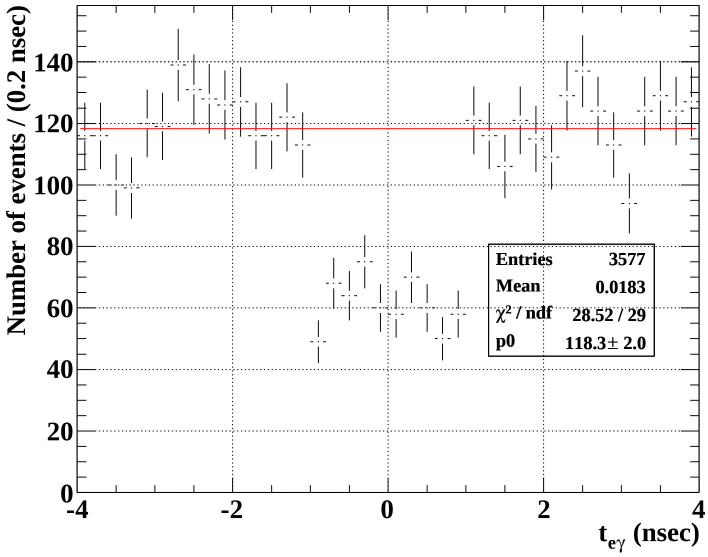


Figure 8.4: $t_{e\gamma}$ distribution in the sideband. The analysis window cuts are applied on variables except for $t_{e\gamma}$. Fewer events at the center part is due to the blinding.

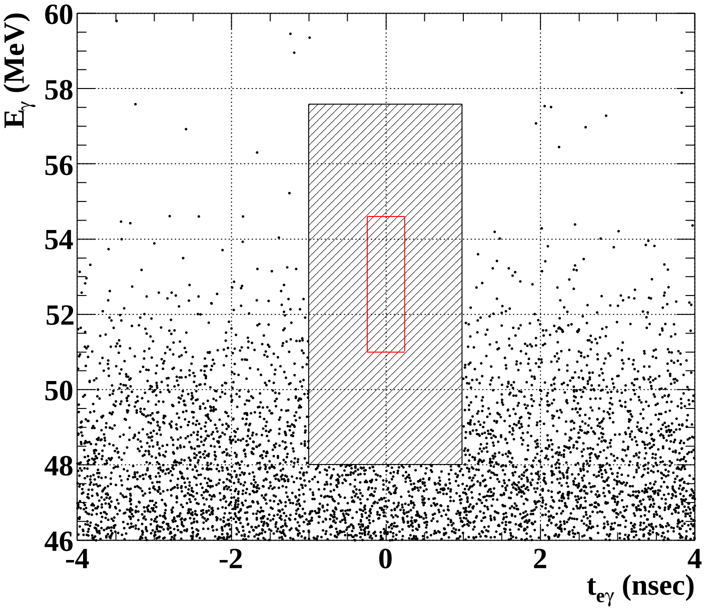


Figure 8.5: Scatter plot of $t_{e\gamma}$ v.s. E_γ. The analysis window cuts are applied on variables except for the two.

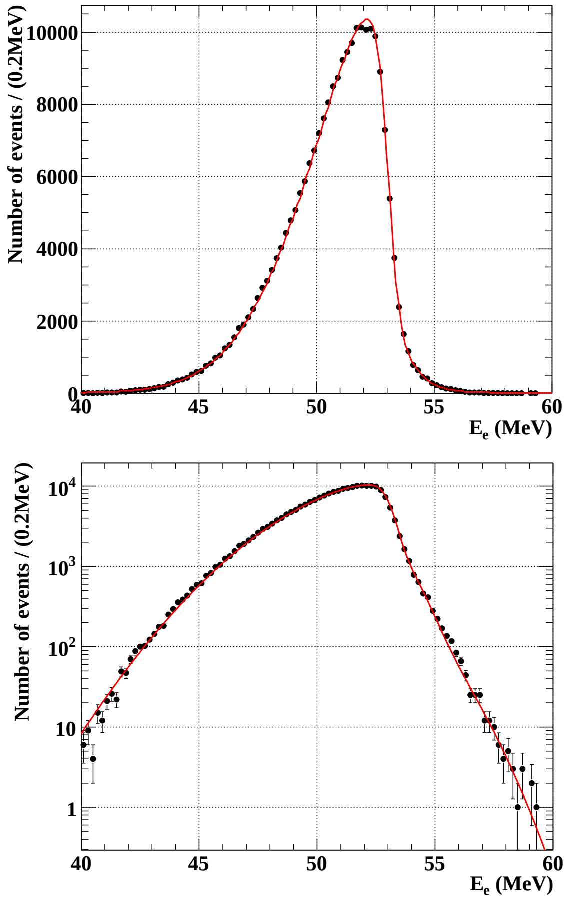


Figure 8.6: Positron background spectrum. Black circle plot shows data spectrum measured in the sideband. Red solid line shows Michel spectrum folded with acceptance and efficiency function and detector response fitted to the data.

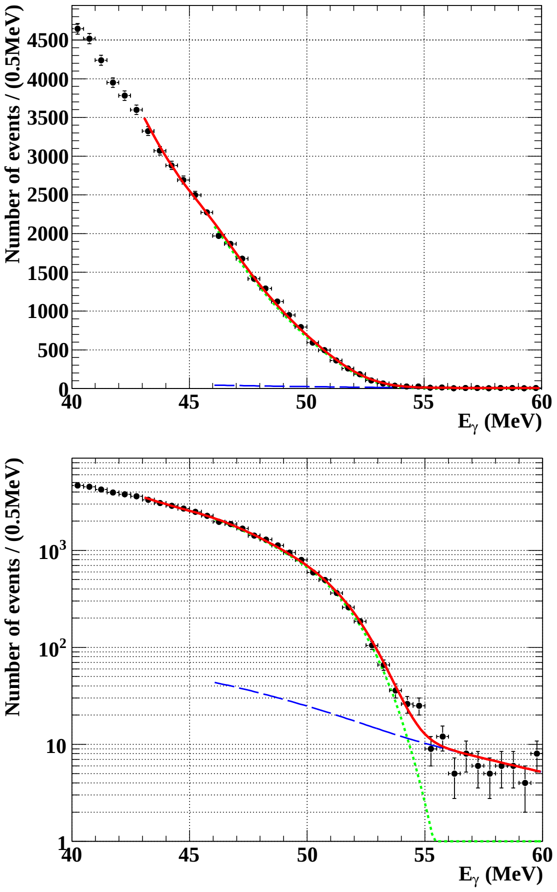


Figure 8.7: Gamma background spectrum. Black circle plot shows data spectrum measured in the sideband data. Red solid line shows the MC spectrum folded with detector resolution fitted to the data. The MC spectrum is composed of single-gamma component (green dotted line) and pileup component (blue dashed line).

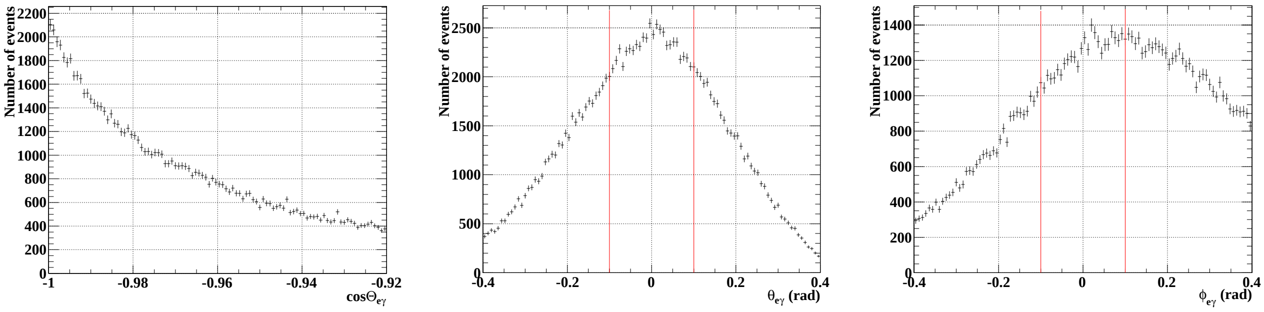


Figure 8.8: Angular distributions of background in the sideband.

8.3 Maximum Likelihood Fit

8.3.1 Formalism

Here, we define the likelihood function which is used to estimate the best values of the number of $\mu^+ \to \mathrm{e}^+\gamma$ events (N_{sig}), that of RD ($\mu^+ \to \mathrm{e}^+\nu_\mathrm{e}\overline{\nu}_\mu\gamma$) events (N_{RD}), and that of accidental background events (N_{BG}).

Let us define the following variables:

- $\vec{\mathbf{x}}$: a set of observed variables of a observation (observation vector);
- $\vec{\mathbf{X}}$: a data set of N_o independent observations $\{\vec{\mathbf{x}}_1, \vec{\mathbf{x}}_2, \cdots, \vec{\mathbf{x}}_{N_o}\}$;
- $\vec{\theta}$: a set of unknown parameters to be estimated.

Then a likelihood function is normally defined as

$$\mathcal{L}(\vec{\mathbf{X}} \mid \vec{\theta}) = \prod_{i=1}^{N_o} p(\vec{\mathbf{x}}_i \mid \vec{\theta}), \tag{8.1}$$

where $p(\vec{\mathbf{x}}_i \mid \vec{\theta})$ is a conditional probability density function (PDF). It gives the probability density of the observation $\vec{\mathbf{x}}_i$ when the parameter set is equal to $\vec{\theta}$. It is a function of $\vec{\mathbf{x}}$ and normalized to 1. The left-hand side, $\mathcal{L}(\vec{\mathbf{X}} \mid \vec{\theta})$, is a likelihood if we regard it as a function only of $\vec{\theta}$. Then it does not have a meaning of probability, and thus it is not normalized. The maximum likelihood fit is a method of estimating unknown parameter set $\vec{\theta}$ by searching for the value $\vec{\theta}^{\mathrm{best}}$ for which $\mathcal{L}(\vec{\mathbf{X}} \mid \vec{\theta})$ has its maximum, given a particular set of observations $\vec{\mathbf{X}}$.

In high-energy physics, experiments are usually conducted with limitations of time or beam, and we do not fix the number of measurements in advance. In such case, the observed number of events statistically fluctuates. When we regard the observed number of events as one of parameters to be estimated, the likelihood function is extended to

$$\mathcal{L}_{ext}(\vec{\mathbf{X}} \mid \vec{\theta}) = \frac{N^{N_o} e^{-N}}{N_o!} \prod_{i=1}^{N_o} p(\vec{\mathbf{x}}_i \mid \vec{\theta}), \tag{8.2}$$

where N is the estimated or fitted number of events. This is called the extended likelihood function. By the extended maximum likelihood fit, we can estimate the number of events to be observed in the experiment using the ‘shape’ of event distribution, and this estimated number can be thought a better estimator of the ‘true’ number of events than the actual-observed number.

We use the extended maximum likelihood fit to estimate the number of events of each event type. In our case, the parameters to be estimated are

$$\vec{\theta} = (N_{sig}, N_{RD}, N_{BG}) \ \text{ and } \ N, \tag{8.3}$$

and they are connected by a relation,

$$N = N_{sig} + N_{RD} + N_{BG}. \tag{8.4}$$

Thus the total number of free parameters of the fitting is three. The number of each event type obeys the Poisson distribution and the errors are independent one anther in the extended maximum likelihood fit, whereas the numbers of events between event types are fully anti-correlated and the errors become smaller in the normal maximum likelihood fit. After this, we describe $\mathcal{L}_{ext}(\vec{\mathbf{X}} \mid \vec{\theta})$ as $\mathcal{L}(N_{sig}, N_{RD}, N_{BG})$ or simply $\mathcal{L}$.

As discussed in Sec.2.3, the four kinematic variables, $(E_e, E_\gamma, \Theta_{e\gamma}, t_{e\gamma})$, can be used to discriminate backgrounds. We saw, however, different detector responses for the angular measurement in θ and ϕ directions. Therefore, we separately treat these two directions for the back-to-back condition. Thus, the set of kinematic variables becomes

$$(E_e, E_\gamma, \theta_{e\gamma}, \phi_{e\gamma}, t_{e\gamma}). \tag{8.5}$$

We observed that the performance of the detector is position dependent. This means the PDF is different for each event. One way to cope with this situation is using an average PDF. If we do not care where and when the event was detected, we can form a single PDF. Measurements would obey this distribution. This is not biased. However, we could lose some sensitivity because we abandon some useful information. A better way is using event-by-event PDFs. Then, we can use a good-measured event with a large weight or large significance while we can also use worse-measured events without discarding them.

To take into account the event-by-event difference of PDFs, we incorporate also the uncertainties or resolutions of measurements into the observation vector. Thus, our observation vector is given as

$$\vec{\mathbf{x}}_i = (E_e, E_\gamma, \theta_{e\gamma}, \phi_{e\gamma}, t_{e\gamma},\ \delta E_e, \delta E_\gamma, \delta\theta_{e\gamma}, \delta\phi_{e\gamma}, \delta t_{e\gamma})_i = (\vec{x}_i, \delta\vec{x}_i), \tag{8.6}$$

Using Bayes theorem, the probability that an event is observed at $\vec{\mathbf{x}} = \vec{\mathbf{x}}_i$ is given by

$$p(\vec{\mathbf{x}}_i) = P(sig) \cdot p(\vec{\mathbf{x}}_i \mid sig) + P(RD) \cdot p(\vec{\mathbf{x}}_i \mid RD) + P(BG) \cdot p(\vec{\mathbf{x}}_i \mid BG), \tag{8.7}$$

because $P(sig|\vec{\mathbf{x}}_i) + P(RD|\vec{\mathbf{x}}_i) + P(BG|\vec{\mathbf{x}}_i) = 1$. Thus our PDF can be written

$$p(\vec{\mathbf{x}}_i|N_{sig}, N_{RD}, N_{BG}) = \frac{N_{sig}}{N} \cdot p(\vec{\mathbf{x}}_i \mid sig) + \frac{N_{RD}}{N} \cdot p(\mathbf{x}_i \mid RD) + \frac{N_{BG}}{N} \cdot p(\vec{\mathbf{x}}_i \mid BG). \tag{8.8}$$

Let us write the PDFs like

$$\begin{aligned} p(\vec{\mathbf{x}}_i \mid sig) = S(\vec{\mathbf{x}}_i) = s(\vec{x}_i|\delta\vec{x}_i)p(\delta\vec{x}_i), \\ p(\vec{\mathbf{x}}_i \mid RD) = R(\vec{\mathbf{x}}_i) = r(\vec{x}_i|\delta\vec{x}_i)p(\delta\vec{x}_i), \\ p(\vec{\mathbf{x}}_i \mid BG) = B(\vec{\mathbf{x}}_i) = b(\vec{x}_i|\delta\vec{x}_i)p(\delta\vec{x}_i), \end{aligned}$$

then our likelihood can be formed as

$$\begin{aligned} &\mathcal{L}(N_{sig}, N_{RD}, N_{BG}) \\ &= \frac{N^{N_o} e^{-N}}{N_o!} \prod_{i=1}^{N_o} \Big(\frac{N_{sig}}{N} \cdot S(\vec{\mathbf{x}}_i) + \frac{N_{RD}}{N} \cdot R(\mathbf{x}_i) + \frac{N_{BG}}{N} \cdot B(\vec{\mathbf{x}}_i) \Big) \\ &= \frac{N^{N_o} e^{-N}}{N_o!} \prod_{i=1}^{N_o} \Big(\frac{N_{sig}}{N} s(\vec{x}_i|\delta\vec{x}_i)p(\delta\vec{x}_i) + \frac{N_{RD}}{N} r(\vec{x}_i|\delta\vec{x}_i)p(\delta\vec{x}_i) + \frac{N_{BG}}{N} b(\vec{x}_i|\delta\vec{x}_i)p(\delta\vec{x}_i) \Big) \\ &= \frac{N^{N_o} e^{-N}}{N_o!} \prod_{i=1}^{N_o} p(\delta\vec{x}_i) \prod_{i=1}^{N_o} \Big(\frac{N_{sig}}{N} s(\vec{x}_i|\delta\vec{x}_i) + \frac{N_{RD}}{N} r(\vec{x}_i|\delta\vec{x}_i) + \frac{N_{BG}}{N} b(\vec{x}_i|\delta\vec{x}_i) \Big), \end{aligned} \tag{8.9}$$

where from the second line to the third, we used an assumption that $p(\delta\vec{x}_i)$ is common to all the event types[2]. The term $\prod_{i=1}^{N_o} p(\delta\vec{x}_i)$ is then omitted in the likelihood fit because it is independent of the fit parameters, while $p(\delta\vec{x}_i)$ is defined by the measured event distribution for the toy-MC simulation, described later in Sec.8.4.1.

The maximization of the likelihood, or more precisely the minimization of the negative log likelihood (NLL) is done using the MINUIT package [93].

8.3.2 Probability Density Functions

PDFs for $\mu^+ \to \mathrm{e}^+\gamma$ Decay

The $\mu^+ \to \mathrm{e}^+\gamma$ signal PDF, S (or s), can be written as a product of statistically independent PDFs for the five kinematic variables.

$$S(E_e, E_\gamma, \theta_{e\gamma}, \phi_{e\gamma}, t_{e\gamma}) = S_1(E_e)S_2(E_\gamma)S_3(\theta_{e\gamma})S_4(\phi_{e\gamma})S_5(t_{e\gamma}), \tag{8.10}$$

each of the components is defined by the detector response function with measured resolutions. In Chapter 7, we investigated the detector performance and obtained all of the response functions. $S_1(E_e)$ is defined by the positron energy response evaluated by the fitting of Michel-spectrum edge. It is given as a triple Gaussian. We use a common PDF over all events in the analysis region. $S_2(E_\gamma)$ is defined by the gamma energy response evaluated by the 55 MeV gamma in π^0 run. It is asymmetry with lower tail. Since the resolution is position dependent, S_2 is different for the reconstructed gamma position. $S_3(\theta_{e\gamma})$ and $S_4(\phi_{e\gamma})$ are defined by the angular responses formed by combining the evaluated resolutions of gamma-ray position, positron emission angles and muon decay vertex. Since the gamma position resolution is evaluated position dependently, they are also dependent on the reconstructed gamma position. $S_5(t_{e\gamma})$ is defined by the timing response evaluated by the RD peak in normal data taking. The resolution depends on the quantity $|\Delta z_{DCH-TIC}|$, which is the difference between the z-coordinate of positron hit position on the timing ϕ-counter bar surface reconstructed by the timing counter and the projected track. It also depends on the reconstructed gamma energy to take into account the dependence on deposit energy and collected amount of light. The average PDFs are shown in Figure 8.9.

PDFs for Radiative Decay

The RD ($\mu^+ \to \mathrm{e}^+\nu_\mathrm{e}\overline{\nu}_\mu\gamma$) PDF, R (or r), is more complicated because the kinematic variables are correlated one another except for $t_{e\gamma}$. It is defined as

$$R(E_e, E_\gamma, \theta_{e\gamma}, \phi_{e\gamma}, t_{e\gamma}) = R_1(E_e, E_\gamma, \theta_{e\gamma}, \phi_{e\gamma})R_2(t_{e\gamma}). \tag{8.11}$$

$R_2(t_{e\gamma})$ is the same function as for the signal. $R_1(E_e, E_\gamma, \theta_{e\gamma}, \phi_{e\gamma})$ is formed with the theoretical correlated spectrum of RD folded with the detector response functions and

---

[2] We may have different event distribution between the signal and the background if the polarization of the muon beam is partly preserved in the target since upstream-downstream asymmetry would be expected for the background in this case. The asymmetry of the event distribution due to the polarization is, however, expected to be almost anti-symmetric, while the detector performance is symmetric for both the gamma-ray detector and the positron spectrometer. Therefore, the distribution of the resolution for the background is still expected to be the same as for the signal

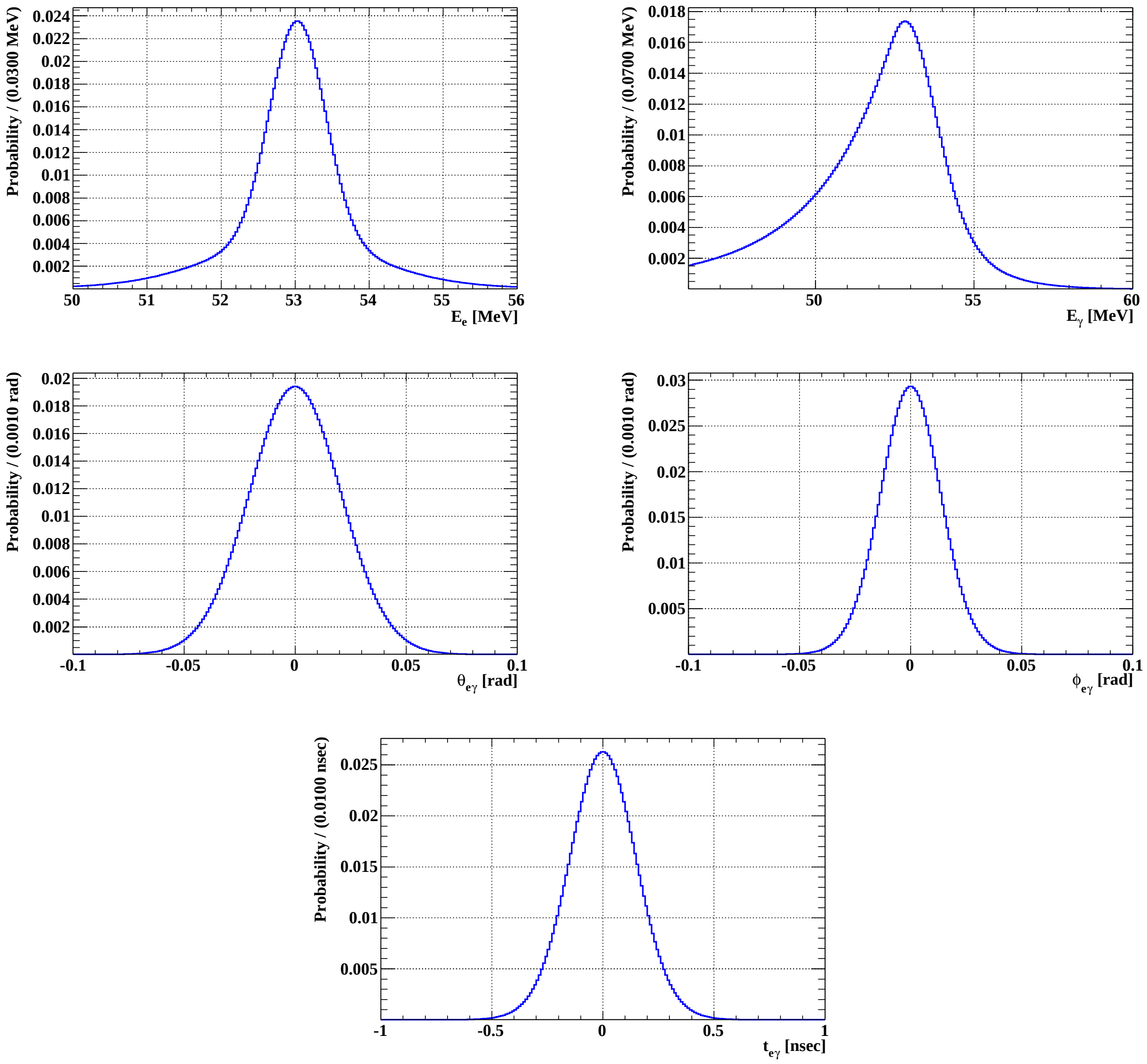


Figure 8.9: PDFs for $\mu^+ \to e^+\gamma$ signal event. The average PDFs over events in analysis window are shown here. In fitting, event-by-event PDFs are used.

acceptance functions. The theoretical spectrum given by Eq.2.18 is a three-dimensional correlated function of E_e, E_γ, and $\cos\Theta_{e\gamma}$. A three-dimensional convolution is calculated event-by-event. The projections of the RD PDF on each kinematic variable are shown if Figure 8.10.

PDFs for Accidental Background

The accidental background PDF, B (or b), can be written as a product of statistically independent PDFs for the five kinematic variables

$$B(E_e, E_\gamma, \theta_{e\gamma}, \phi_{e\gamma}, t_{e\gamma}) = B_1(E_e)B_2(E_\gamma)B_3(\theta_{e\gamma})B_4(\phi_{e\gamma})B_5(t_{e\gamma}), \tag{8.12}$$

similar to the signal PDF while each component is defined by the measured background spectrum. $B_1(E_e)$ is obtained by a fitting of the measured Michel spectrum. $B_2(E_{e\gamma})$ is

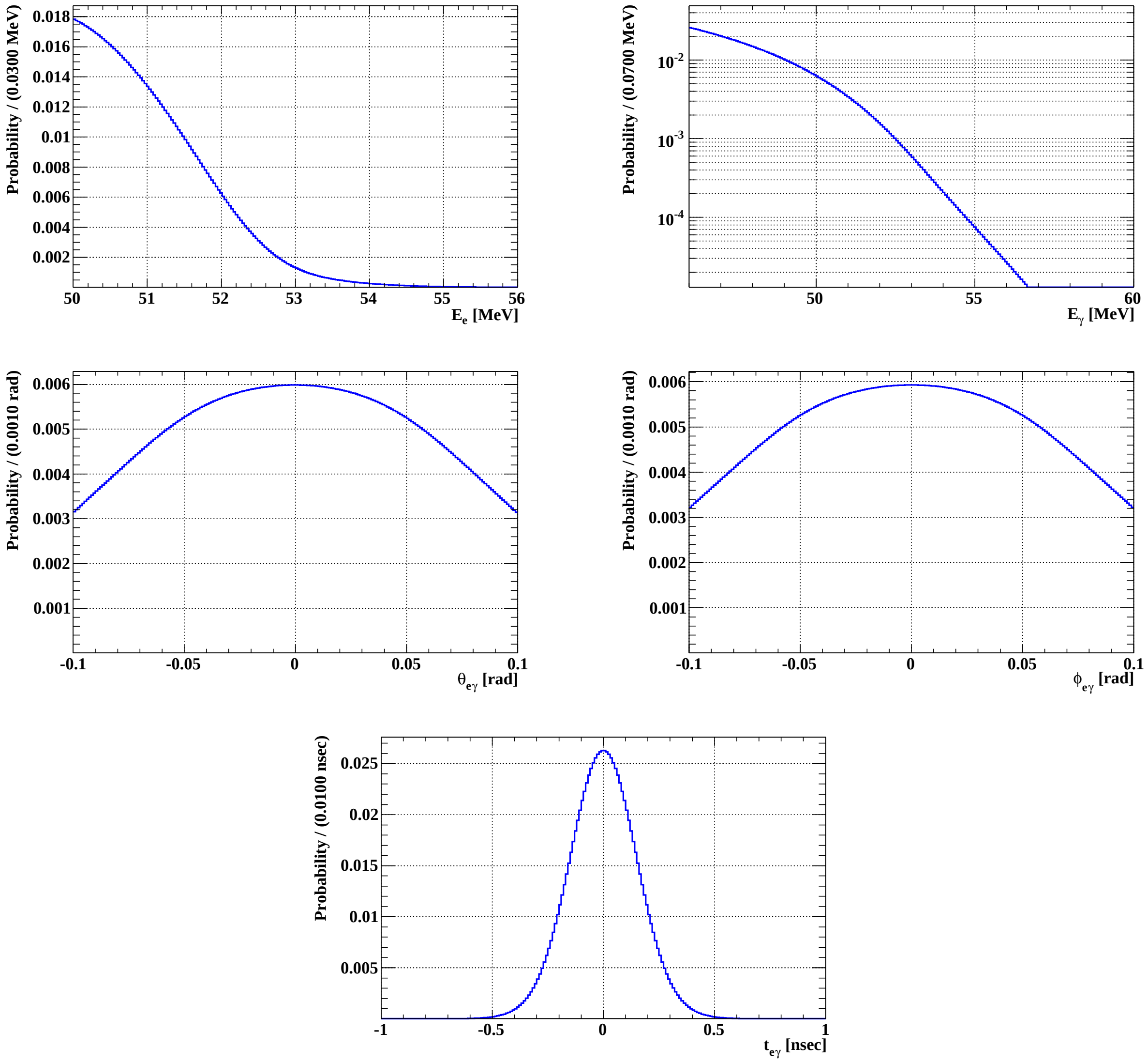


Figure 8.10: PDFs for $\mu^+ \to \mathrm{e}^+\nu_\mathrm{e}\overline{\nu}_\mu\gamma$ RD event. Here the projections on each variable are shown, but (a)-(d) are correlated each other.

defined by the background gamma energy spectrum measured in the $t_{e\gamma}$-sideband. The spectrum is separately measured in 38 sections in $(u_\gamma, v_\gamma, w_\gamma)$. Each spectrum is fitted with the MC one smeared by the detector response. $B_3(\theta_{e\gamma})$ and $B_4(\phi_{e\gamma})$ are obtained by fitting third-order polynomials to the spectrum measured in eight slices along u_γ and v_γ directions, respectively. The asymmetries comes from the trigger bias. $B_5(t_{e\gamma})$ is defined by a flat distribution because the positron and gamma are uncorrelated. The average PDFs are shown in Figure 8.11.

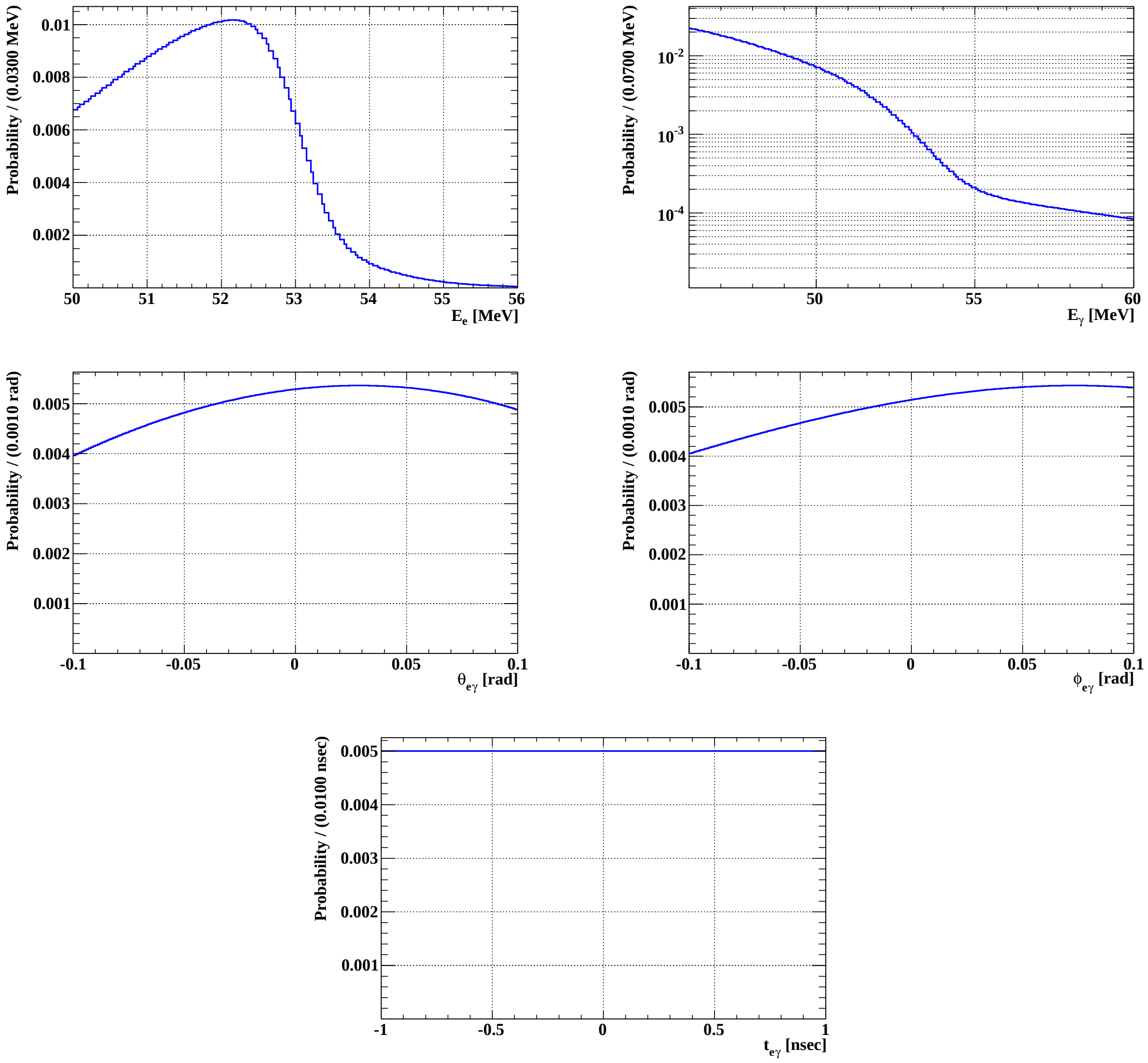


Figure 8.11: PDFs for accidental background event. The averaged PDFs are shown here.

8.4 Confidence Region

In past experiments performing likelihood analyses, the shape of likelihood functions were used to estimate the upper limits on the number of signal events. Namely, the 90%-confidence contour is constructed as a contour of NLL where the value of NLL becomes 1.35 larger than its minimum value. This is known as the MINOS (profile likelihood) method [94]. It is, however, now known that likelihood-base confidence intervals are in general only approximations (asymptotic approximations in the large number of observation), and they are not good approximations for small number. Therefore, we calculate a confidence region with more statistically accurate method.

8.4.1 Toy MC

We can simulate the experiment using the measured PDFs because they completely describe the statistical behavior of the measurements. We call this simulation 'toy MC' and the simulated experiment 'toy experiment' to distinguish it from the general full MC simulation. This toy MC is a powerful tool to study statistical behavior of our experiment and analysis even though the word 'toy' is named.

In the toy MC, the distribution of resolutions in the likelihood function, $\prod p(\delta\vec{x}_i)$, is also necessary, whereas it is omitted in the fitting process. We use the resolution distribution based on the measured event distribution. The simulation scheme is the following:

i) Generate a list of events observed in the analysis window and the $t_{e\gamma}$-sidebands. This list contains values of the parameters to describe the PDFs of the likelihood function associated with each event.

ii) Select an event randomly from the list for every generation of the toy-MC event.

iii) Form PDFs with the parameters (resolutions) associated with the selected event.

iv) Generate the kinematic variables in accordance with the PDFs.

Using the measured event distribution allows to take into account the correlation of the resolution and the efficiency between the gamma-ray detector and the positron spectrometer which is expected for the events selected in the analysis window.

8.4.2 Construction of Confidence Region

We construct a confidence region on the (N_{sig}, N_{RD})-plane by the Feldman-Cousins approach [95]. The confidence region is constructed in the following procedure:

i) Perform a likelihood fit on the data with a constraint on physically allowed region ($N_{sig} > 0$, $N_{RD} > 0$, $N_{BG} > 0$) and obtain the best-fit point $(N_{sig}^{\mathrm{best}_0}, N_{RD}^{\mathrm{best}_0}, N_{BG}^{\mathrm{best}_0})$. Here, the subscript 0 means the measurement of the real experiment.

ii) Pick a point $(N_{sig}^{i}, N_{RD}^{i})$ as a sample point in the plane. The picked sample point is denoted by a superscript i.

iii) Calculate a likelihood ratio R^i_{data} at the sample point as

$$R^i_{data} = \frac{\mathcal{L}(\vec{\mathbf{X}}_0 \mid N^i_{sig}, N^i_{RD}, N^{\text{best}_0}_{BG})}{\mathcal{L}(\vec{\mathbf{X}}_0 \mid N^{\text{best}_0}_{sig}, N^{\text{best}_0}_{RD}, N^{\text{best}_0}_{BG})}. \tag{8.13}$$

iv) Simulate many experiments assuming N^i_{sig}, N^i_{RD}, and $N^{\text{best}_0}_{BG}$ as the true expectations. The number of simulated events for each event type in each experiment is randomly fluctuated in accordance with the Poisson distribution. Each toy experiment is denoted by a subscript j.

v) Perform a likelihood fit for each toy experiment with a constraint on physically allowed region, obtain the best-fit point $(N^{\text{best}_j}_{sig}, N^{\text{best}_j}_{RD}, N^{\text{best}_j}_{BG})$, and calculate a likelihood ratio $R^i_{MC_j}$ as

$$R^i_{MC_j} = \frac{\mathcal{L}(\vec{\mathbf{X}}_j \mid N^i_{sig}, N^i_{RD}, N^{\text{best}_0}_{BG})}{\mathcal{L}(\vec{\mathbf{X}}_j \mid N^{\text{best}_j}_{sig}, N^{\text{best}_j}_{RD}, N^{\text{best}_j}_{BG})}. \tag{8.14}$$

vi) Calculate the confidence level at the sample point from the probability

$$P(R^i_{data} < R^i_{MC}) \tag{8.15}$$

over the simulated experiments. If the probability is 0.9, for example, the sample point (N^i_{sig}, N^i_{RD}) should be on the contour of 90 % confidence level (C.L.).

vii) Repeat the procedure ii)-vi) for many choices of (N^i_{sig}, N^i_{RD}) and scan the contour on the plane. All points with the probability less than or equal to 0.9 form the acceptance region of 90% C.L.

This procedure is illustrated in Figure 8.12. In this way, we can construct a two-dimensional confidence region at any confidence level. The point is the ordering principle of the acceptance region by the likelihood ratio. The one-parameter confidence interval is given by the edges of the contour in the direction.

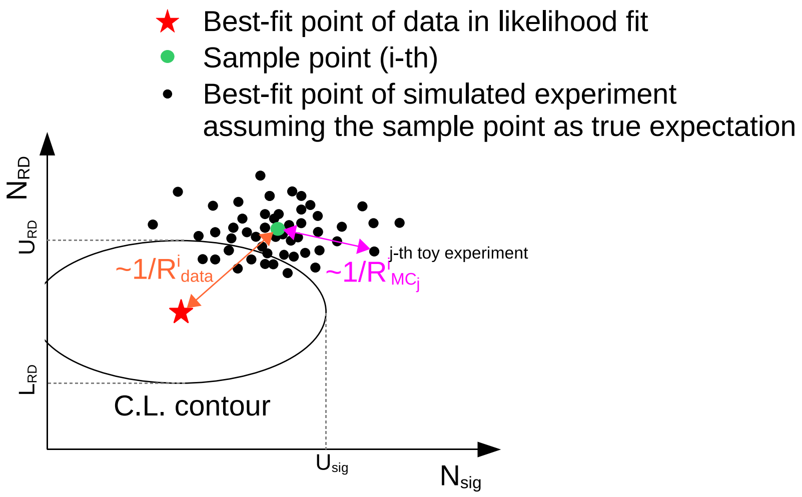


Figure 8.12: Illustration of confidence region construction. The inverse of the likelihood ratio corresponds to the distance between the estimated point and the assumed true point.

8.5 Normalization

A branching ratio is defined as a ratio of the decay rate for a particular process, for example $\Gamma(\mu \to e\gamma)$, to the total one Γ_{tot}: $\mathcal{B}(\mu \to e\gamma) \equiv \Gamma(\mu \to e\gamma)/\Gamma_{tot}$. For the muon decay, the total decay rate is nearly saturated by that for Michel decay $\Gamma(\mu \to e\nu\overline{\nu})$ with a small contribution from RD. Thus, we use the Michel decay for the normalization channel. The number of Michel positrons is counted simultaneously with the signal with the same analysis cuts. In this way, we can calculate the branching ratio independently of the instant beam intensity and nearly insensitively to the positron acceptance and efficiency factors which differ only for small momentum-dependent effects between the two modes. Thus it is an effective method especially for run 2008 where the conditions of the drift chamber were drastically changing.

We use the TIC-self trigger data mixed in the physics data taking as the Michel data sample. The number of detected events in the sample after the selection can be written as the following product of factors

$$N_{MD} = N^{\mu}_{e\nu\overline{\nu}} \times \mathcal{B}_{e\nu\overline{\nu}} \times f^{E}_{e\nu\overline{\nu}} \times T_{e\nu\overline{\nu}} \times \frac{1}{P_{e\nu\overline{\nu}}} \times \epsilon^{tirg}_{e\nu\overline{\nu}} \times A^{TIC}_{e\nu\overline{\nu}} \times \epsilon^{DCH}_{e\nu\overline{\nu}} \times A^{DCH}_{e\nu\overline{\nu}}, \tag{8.16}$$

where N_{μ} is the number of stopped muons during the data taking; $\mathcal{B}_{e\nu\overline{\nu}}$ is the branching ratio of Michel decay, and $f^{E}_{e\nu\overline{\nu}}$ is the fraction of Michel spectrum above 50 MeV; T is the livetime; P is the pre-scale factor of the TIC-self trigger; $\epsilon^{tirg}_{e\nu\overline{\nu}}$ is the conditional trigger efficiency; $A^{TIC}_{e\nu\overline{\nu}}$ is the conditional acceptance of timing counter including the probability of DCH-TIC matching; $\epsilon^{DCH}_{e\nu\overline{\nu}}$ is the conditional tracking efficiency including the selection criteria; and $A^{DCH}_{e\nu\overline{\nu}}$ is the geometrical acceptance of the drift chamber.

A similar equation can be written down for the number of detected $\mu^{+} \to \mathrm{e}^{+}\gamma$ events from the MEG trigger data:

$$N_{sig} = N^{\mu}_{sig} \times \mathcal{B}_{e\gamma} \times T_{e\gamma} \times \times \epsilon^{tirg}_{e\gamma} \times A^{TIC}_{e\gamma} \times \epsilon^{DCH}_{e\gamma} \times A^{DCH}_{e\gamma} \times \epsilon^{\gamma}_{e\gamma} \times A^{\gamma}_{e\gamma}, \tag{8.17}$$

where the factors are defined in the same way as for Michel decay. Additional factors are $\epsilon^{\gamma}_{e\gamma}$ for gamma-ray detection and reconstruction efficiency including selection criteria, and $A^{\gamma}_{e\gamma}$ for the conditional acceptance of the gamma ray from $\mu^{+} \to \mathrm{e}^{+}\gamma$ decay.

Here, we explain what we mean by the word 'conditional' for each factor. First, we define the positron geometrical acceptance A^{DCH} so that the inverted direction of positron goes to the fiducial volume of the gamma-ray detector defined in Sec. 7.1. Therefore, it is common to the two modes by definition. $\epsilon^{DCH}_{e\nu\overline{\nu}}$ is the tracking efficiency for the positron inside the acceptance. A^{TIC} is the probability of getting at least one timing-counter hit when the positron is emitted inside the acceptance and the track is reconstructed. ϵ^{tirg} is the probability that the trigger conditions are satisfied if the event satisfies all the analysis selection criteria above. $A^{\gamma}_{e\gamma}$ is the probability that the gamma-ray from $\mu^{+} \to \mathrm{e}^{+}\gamma$ decay detected in the fiducial volume of the gamma-ray detector when the accompanied positron is detected in the acceptance. It is close to 1 but the finite detector resolution affects it. $\epsilon^{\gamma}_{e\gamma}$ is the efficiency for the gamma inside the acceptance.

The number of stopped muon and livetime is common to the two modes since we measure those two modes simultaneously. Then the branching ratio on $\mu^{+} \to \mathrm{e}^{+}\gamma$ decay can be written as

$$\frac{\mathcal{B}_{e\gamma}}{\mathcal{B}_{e\nu\overline{\nu}}} = \frac{N_{sig}}{N_{MD}} \times \frac{f^{E}_{e\nu\overline{\nu}}}{P_{e\nu\overline{\nu}}} \times \frac{\epsilon^{trig}_{e\nu\overline{\nu}}}{\epsilon^{trig}_{e\gamma}} \times \frac{A^{TIC}_{e\nu\overline{\nu}}}{A^{TIC}_{e\gamma}} \times \frac{\epsilon^{DCH}_{e\nu\overline{\nu}}}{\epsilon^{DCH}_{e\gamma}} \times \frac{1}{\epsilon^{\gamma}_{e\gamma}} \times \frac{1}{A^{\gamma}_{e\gamma}}. \tag{8.18}$$

Number of Michel positron The number of detected Michel positron is $N_{MD} = 11414$ in the analysis range $50 < E_e < 56$ MeV.

Pre-scale factor The pre-scale factor of the TIC-self trigger is $P_{e\nu\overline{\nu}} = 10^7$, while that of the MEG trigger is 1.

Fraction of Michel spectrum The fraction of Michel spectrum included in the analysis range is calculated to be $f^E_{e\nu\overline{\nu}} = 0.101 \pm 0.006$ by integrating the theoretical Michel spectrum over 50 MeV. The systematic uncertainty stems from that of positron energy scale of 200 keV in the Michel spectrum fitting.

Trigger-efficiency ratio The trigger efficiency consists of three components corresponding to the three trigger criteria described in Sec.3.3.3: gamma energy, time coincidence and direction match. The trigger efficiency by the timing-counter hit is canceled out in the ratio. The efficiency of gamma energy is estimated by the online energy resolution and evaluated to > 99 % for $E_\gamma > 45$ MeV. It is confirmed by looking at the ratio of the trigger rate in the LXe-self trigger with normal and lower thresholds. The timing efficiency is also estimated by checking the online time resolution and confirmed the full efficiency (> 99 %). The direction-match efficiency is evaluated to be 0.66 by the MC simulation for the signal. This number is confirmed by analyzing unbiased data of the dedicated RD run which were taken without the direction-match conditions in the trigger. In total, the trigger-efficiency ratio is evaluated to be $\epsilon^{trig}_{e\gamma}/\epsilon^{trig}_{e\nu\overline{\nu}} = 0.66 \pm 0.03$.

DCH-TIC matching-probability ratio It is calculated using the DCH-self trigger data. The probability is evaluated as a function of positron momentum (Figure 7.27). The ratio is estimated to $A^{TIC}_{e\gamma}/A^{TIC}_{e\nu\overline{\nu}} = 1.11 \pm 0.02$ by comparing the end point of the plot and the average above 50 MeV, where the uncertainty is estimated from different fittings to the plot.

Tracking-efficiency ratio It is calculated using random trigger data samples. The ratio of detected events to the full Michel spectrum is calculated as a function of positron momentum. The efficiency ratio is evaluated to $\epsilon^{DCH}_{e\gamma}/\epsilon^{DCH}_{e\nu\overline{\nu}} = 1.02 \pm 0.005$ by comparing the end-point of the plot and the integral above 50 MeV, where the uncertainty is estimated from different fittings to the plot.

Gamma-ray efficiency It is evaluated in Sec.7.7 to be $\epsilon^{\gamma}_{e\gamma} = 0.63 \pm 0.04$.

Gamma-ray acceptance The effect that the gamma from $\mu^+ \to e^+\gamma$ decay slips from the fiducial volume when the accompanied positron is detected in the acceptance consists of two opposite effects; missing and gaining. These effects are simulated using measured angle and position resolutions and evaluated to be $A^{\gamma}_{e\gamma} = 0.98 \pm 0.005$.

Normalization Factor

Altogether from the above, the normalization factor[3] k , which is the effective number of observed Michel decay, is given as

$$k = (5.2 \pm 0.5) \times 10^{11}. \tag{8.19}$$

With this number, the branching ratio is given by $\mathcal{B}(\mu^+ \to e^+\gamma) = N_{sig}$ / k. The summary of the normalization is given in Table 8.1.

This result was confirmed by alternative methods for the normalization. The normalization with radiative decay is described in Appendix A.2, and that with accidental background is in Appendix B. Table 8.2 summarizes those results.

Table 8.1: List of factors in normalization calculation. Here ratio is given as that of signal to Michel.

factor	estimated value	relative uncertainty
Number of Michel events	11414	–
Pre-scale factor	10^7	–
Michel spectrum fraction	0.101 ± 0.006	5.9 %
Trigger efficiency ratio	0.66 ± 0.03	4.5 %
DCH-TIC matching probability ratio	1.11 ± 0.02	1.8 %
Tracking efficiency ratio	1.02 ± 0.005	0.5 %
Gamma-ray efficiency	0.63 ± 0.04	5.0 %
Gamma-ray acceptance	0.98 ± 0.005	0.5 %
Normalization factor	$(5.2 \pm 0.5) \times 10^{11}$	10.0 %

Table 8.2: Summary of normalization.

	Michel	RD	BG
k (10^{11})	5.2 ± 0.5	4.4 ± 1.1	5.2 ± 0.8

---

[3] The inverse of this number, $(1.9 \pm 0.2) \times 10^{-12}$, can be regarded as the single-event sensitivity (SES) of this measurement. However, we do not define it as SES here because the meaning of this number is slightly different from SES in conventional cut-counting analysis. In the likelihood analysis, the analysis region is set to wider and including background events.

8.6 Results and Discussion

8.6.1 Analysis on Sideband Data

Before analyzing the analysis window, we applied the maximum likelihood fit on the sideband fictitious windows. We defined two windows in positive and negative $t_{e\gamma}$-sideband: $1.5 < t_{e\gamma} < 3.5$ and $-3.5 < t_{e\gamma} < -1.5$ ns. Ranges of the other variables are identical to those of the analysis window. We believe there are no signal (and no RD events) in these windows. Thus, these analyses can be a kind of estimation of our sensitivity as well as a check of the maximum likelihood fit. Table 8.3 gives the results.

Table 8.3: Results of likelihood analysis in the sidebands.

Fit region (ns)	N_o	Best-fit values $(N_{sig}, N_{RD}, N_{BG})$	N_{sig} upper limit (90% C.L.)	N_{sig} lower limit
$1.5 < t_{e\gamma} < 3.5$	1197	$\left(1.6^{+3.4},\ 26^{+17}_{-16},\ 1169^{+38}_{-37}\right)$	10.4	zero-consistent
$-3.5 < t_{e\gamma} < -1.5$	1209	$\left(0^{+1.2},\ 9.6^{+16},\ 1199^{+38}_{-38}\right)$	4.6	zero-consistent

8.6.2 The Maximum Likelihood Fit on the Data

We opened the hidden box and applied the maximum likelihood fit on the data in the analysis window. In the analysis window, $N_o = 1189$ events are observed. In Figure 8.13, the event distributions in the analysis window and the fitted spectra are shown as the projection on each kinematic variable. The best-fit values of $\vec{\theta} = (N_{sig}, N_{RD}, N_{BG})$ are at

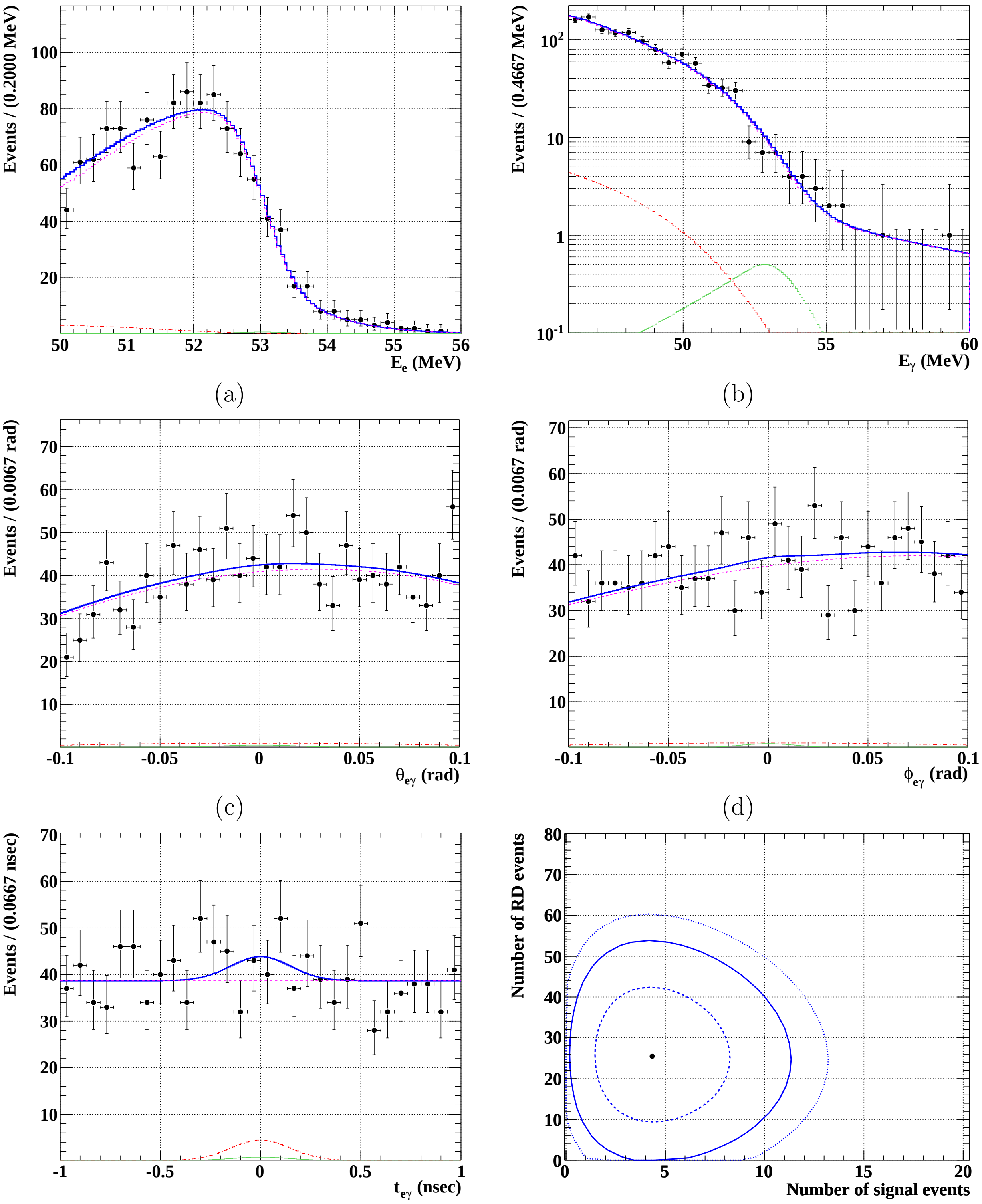
$$\vec{\theta}^{\text{best}} = \left(4.3^{+3.9}_{-2.9},\ 25^{+17}_{-16},\ 1159^{+38}_{-37}\right), \tag{8.20}$$

where the asymmetric errors are those from MINUIT (so-called MINOS errors at one sigma). The likelihood function is shown in Figure 8.13(f) in two-dimensional (N_{sig}, N_{RD}) plane.

The expected number of RD events in the analysis window is calculated to be 41 ± 8. See Appendix A for the detail calculation. The best-fit value of $N_{RD}^{\text{best}} = 25^{+17}_{-16}$ is consistent with the expectation. We will discuss the result of N_{sig} best-fit value later on.

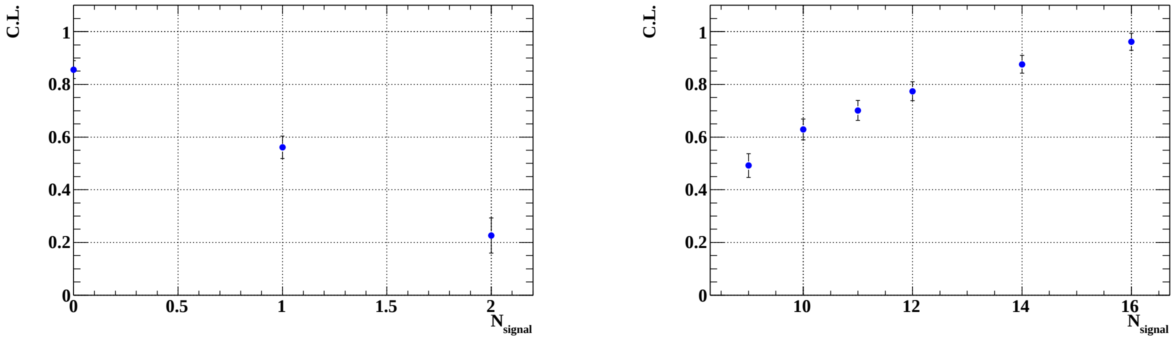
8.6.3 Confidence Region from the Data

We applied the procedure of constructing confidence region described in Sec.8.4.2 to the data. In particular, we calculated the confidence region at 90 % level. First, we scanned the confidence levels at points on a rough grid of (N_{sig}, N_{RD})-plane because it is very time consuming to cover all the plane. By this rough sampling as well as by the shape of the likelihood function, we found that there is no correlation between N_{sig} and N_{RD}. As described in Sec.8.4.2, the one-parameter interval of N_{sig} is calculated by taking the projection of the contour to the N_{sig}-axis. Thus under the no-correlated confidence region, the limits on N_{sig} are on the line of N_{RD} best-fit value. We scanned the confidence levels more finely on the line of N_{RD}^{best}. The distribution of the confidence levels on the N_{RD}^{best}



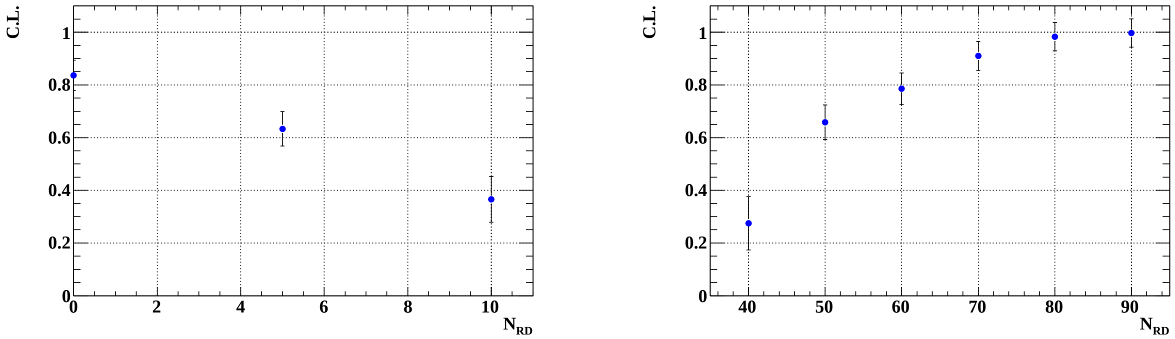
(a) (b) (c) (d) (e) (f)

Figure 8.13: Result of the maximum likelihood fit. Projected event distributions in the analysis window are shown in (a)-(e). Green solid line, red dot-dashed one, and magenta dashed one show the fitted signal, RD, and BG PDFs, respectively. Blue solid line is the total spectrum. (f) Contour plot of the two-dimensional likelihood function as a function of the number of signal and that of RD events. The black dot mark shows the best-fit value. The dashed-, solid-, dotted-lines show the 1, 1.645, 2 sigma or more precisely 0.5, 1.353, 2 ΔNLL contours, respectively.



(a) Lower side (b) Upper side

Figure 8.14: Distribution of N_{sig} confidence level on the N_{RD}^{best}-line.



(a) Lower side (b) Upper side

Figure 8.15: Distribution of N_{RD} confidence level on the N_{sig}^{best}-line.

is shown Figure 8.14. The lower side of the interval is zero-consistent, that is, the 90%-confidence interval contains $N_{sig} = 0$. The result of this measurement is included in the region over which the probability $p(\vec{\mathbf{X}} \mid N_{sig} = 0)$ integrated becomes 90 %. In this case, not a two-sided interval but an upper limit is set in the Feldman-Cousins unified approach. The 90%-confidence upper limit on N_{sig} is calculated to be at 14.5 by an interpolation of the plot. The same calculation is done for N_{RD}, yielding a 90%-confidence upper limit of 68 (Figure 8.15).

8.6.4 Systematic Uncertainties

Summary of Systematic Uncertainties

The systematic uncertainties are summarized in Table 8.4.

Incorporating Systematic Uncertainties

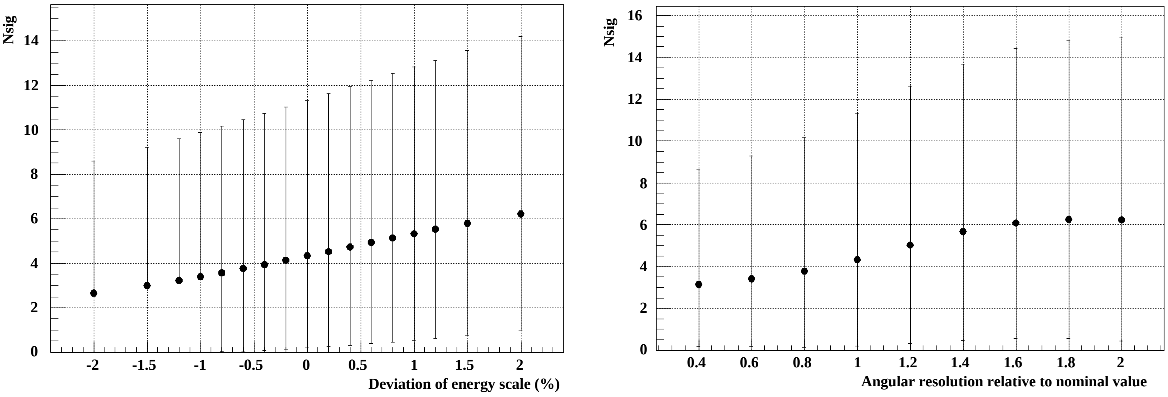
Both the numerator and the denominator could have systematic uncertainties in the likelihood analysis because we actively measure the number of signal using PDFs. The

Table 8.4: Summary of systematic uncertainties.

factor	estimated value	impact on N_{sig}
Gamma-ray energy scale	0.4%	0.4
Gamma-ray energy resolution	10-15%	negligible
Positron spectrum	–[1]	1.14
Angular resolution	1 mrad	0.35
Time resolution	17 ps	negligible
Time center	16 ps	negligible
Normalization factor	10%	–

---

[1] given as alternative parameter sets.



(a) Gamma-ray energy scale (b) Angular resolution

Figure 8.16: Scan of the deviations of the N_{sig} best-fit value. The error bars shows 90%-level upper limit by the MINOS method.

impacts of the systematic uncertainties on the number of signals were determined by looking at the deviation of the best-fit value $\vec{\theta}_{\text{best}}$ when we changed each parameter for its uncertainty value. Figure 8.16 shows examples of scans of N_{sig} deviations by changing of the parameters. The uncertainty of positron-spectrum shape is found to have the largest impact (1.14 events); the next contribution comes from that of gamma-ray energy scale (0.4 events) and then that of angular resolution (0.35 events). The others are found to have negligible effects. The systematic uncertainty of N_{sig} is evaluated to be 1.26 events by adding the individual contribution in quadrature.

In our case, the statistical uncertainty is large, and we accept an approximation to incorporate systematic uncertainties in the calculation of the upper limit. Here, we decided to use a method which takes the statistics-only contour and scales it uniformly along rays from the best-fit value. The scaling factor is given by adding systematic uncertainty of N_{sig} and that of the normalization in quadrature to the statistical uncertainty. When we incorporate the systematic uncertainties, the upper limit on N_{sig} becomes 14.7. Generally, this approximation over-estimates the impact, but the effect is small. The impact of the systematic uncertainties is less than 2 %.

8.6.5 Upper Limit on the Branching Ratio $\mu^+ \to e^+\gamma$

Since the fit result is zero-consistent, an upper limit on $\mu^+ \to e^+\gamma$ branching ratio is set. Together with the normalization factor, the upper limit is calculated to be

$$\mathcal{B}(\mu^+ \to e^+\gamma) < 2.8 \times 10^{-11} \tag{8.21}$$

at 90 % C.L.

8.6.6 Sensitivity of Run 2008

The sensitivity of an experiment is a measure of the result that the experiment is expected to give for the limit on a parameter in the absence of signal. It describes accuracy of the experiment, and is independent of the actual data that could fluctuate.

Here, we define the 90%-confidence sensitivity, $\mathcal{S}$, by the average upper limit at 90 % C.L. over the ensemble of the simulated toy experiments. In each experiment, the expected or true numbers (not actual-generated ones) of RD and BG are given by the best-fit values of the real experiment, $(N_{RD}, N_{BG}) = (25, 1159)$, while that of signal is zero, $N_{sig} = 0$. We simulated 800 experiments, and calculated a 90%-confidence upper limit on N_{sig} for each experiment. The distribution of N_{sig} upper limits is shown in Figure 8.17. The mean value of the distribution is at 6.5. The branching-ratio sensitivity of run 2008 is thus calculated to be

$$\mathcal{S}_{2008} = 1.3 \times 10^{-11}. \tag{8.22}$$

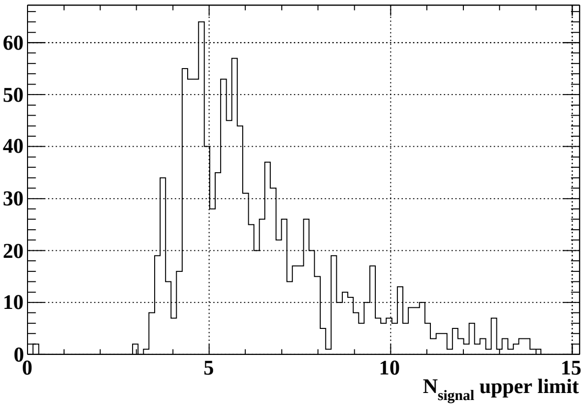


Figure 8.17: Distribution of upper limits of toy experiments.

8.6.7 Cut Analysis

For a cross check and better understanding of the result, here we perform conventional cut-counting analyses. They were also done as blind analyses.

We already defined signal boxes before opening the hidden box to see the background level. They are a box based on 1.64 σ (A) and an optimized one (B). The event selection was however not changed, thus not optimized for the cut analyses, but the same one as that for the likelihood was applied.

The corresponding signal efficiencies for the signal boxes are evaluated using the signal PDFs as ratios to those for likelihood analysis window. We calculated the single-event sensitivities (SES) by applying the relative efficiencies to the normalization factor k. Those numbers are summarized in Table 8.5.

We observed one event in signal box A and two events in signal box B. The event distributions in signal box B projected on each variable are shown in Figure 8.18, where the signal PDFs normalized to 10 events are superimposed. Two-dimensional plots of E_e v.s. E_γ are also shown in Figure 8.19. Then we calculate the confidence intervals in accordance with the Feldman-Cousins method based on the Poisson statistics

$$P(n \mid \mu) = \frac{(\mu + b)^n \exp(-(\mu + b))}{n!}, \tag{8.23}$$

where n is observed number, μ is expected mean number of signal events, and b is the known number of background events. The results are zero-consistent for both signal boxes, and upper limits are given by 3.39 and 3.81 for signal box A and B, respectively. Together with the SESs, upper limits on the $\mu^+ \to \mathrm{e}^+\gamma$ branching ratio,

$$\mathcal{B}_{\mathrm{boxA}} < 1.7 \times 10^{-11}, \tag{8.24}$$

$$\mathcal{B}_{\mathrm{boxB}} < 1.3 \times 10^{-11} \tag{8.25}$$

are obtained.

Table 8.5: Summary of cut analyses.

	1.64σ box (A)		optimized box (B)	
	range	(ϵ_{rel})	range	(ϵ_{rel})
E_e (MeV)	$[52.2, 53.8]$	(0.85)	$[52.0, 56.0]$	(0.92)
E_γ (MeV)	$[51.0, 54.6]$	(0.64)	$[51.0, 56.0]$	(0.70)
$\lvert t_{e\gamma}\rvert$ (ps)	< 242	(0.90)	< 273	(0.94)
$\phi - \Theta_{e\gamma}$ (mrad)	< 33	(0.80)	< 42	(0.94)
SES (10^{-12})	5.0	(0.39)	3.5	(0.56)
$N_{BG}^{\mathrm{exp}} + N_{RD}^{\mathrm{exp}}$	0.95+0.02		2.08+0.03	
N^{obs}	1		2	
Upper limit (10^{-11})	1.68		1.32	

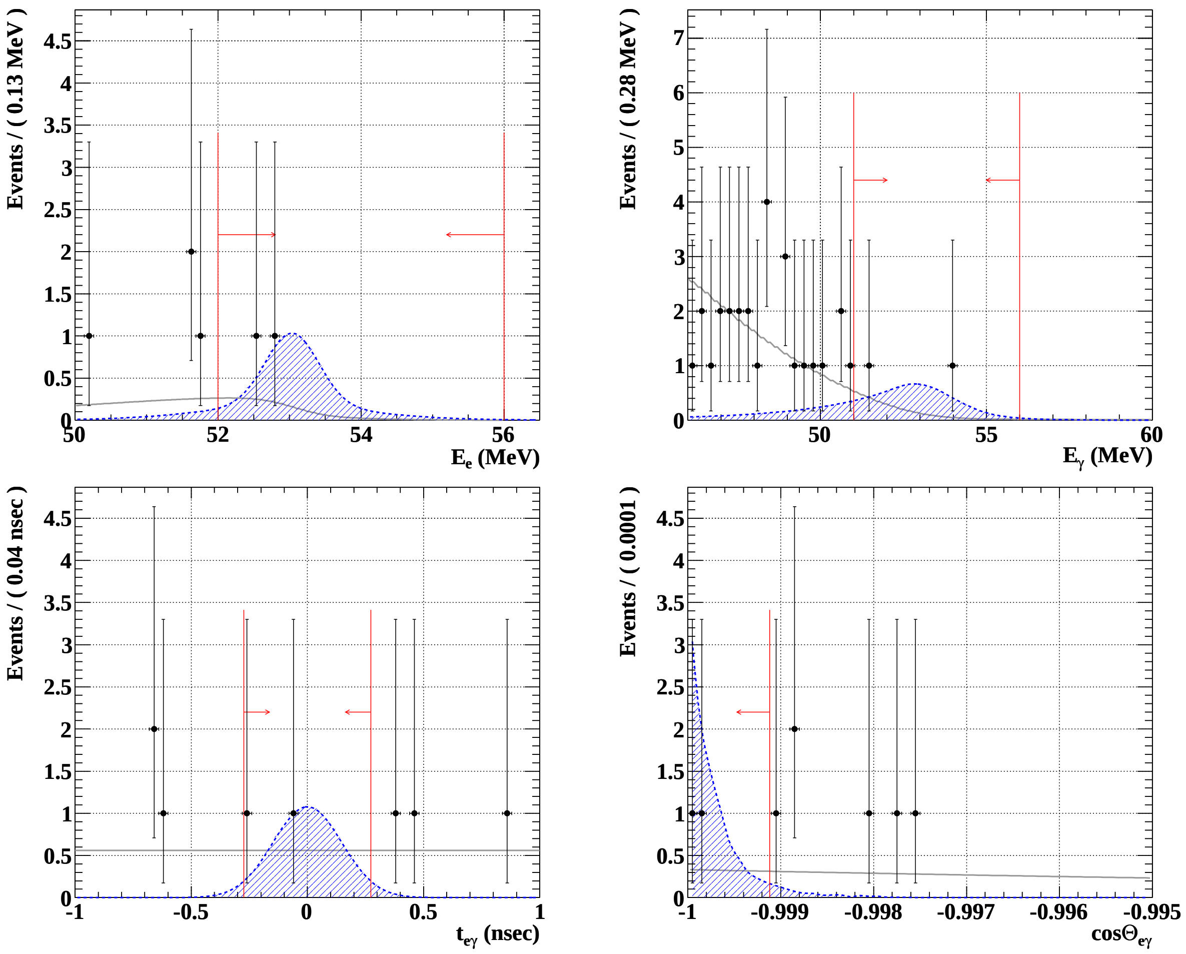
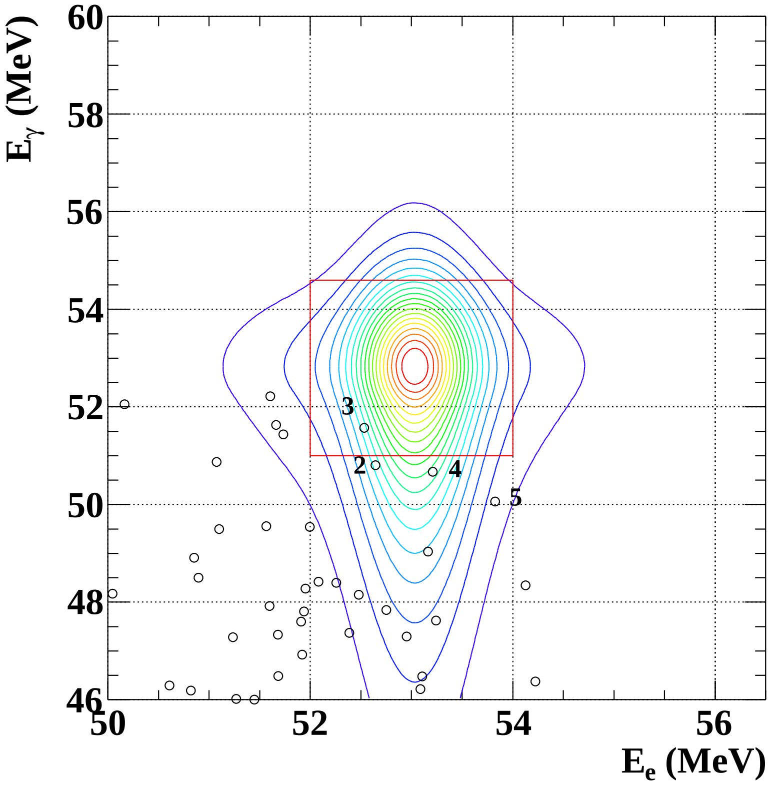
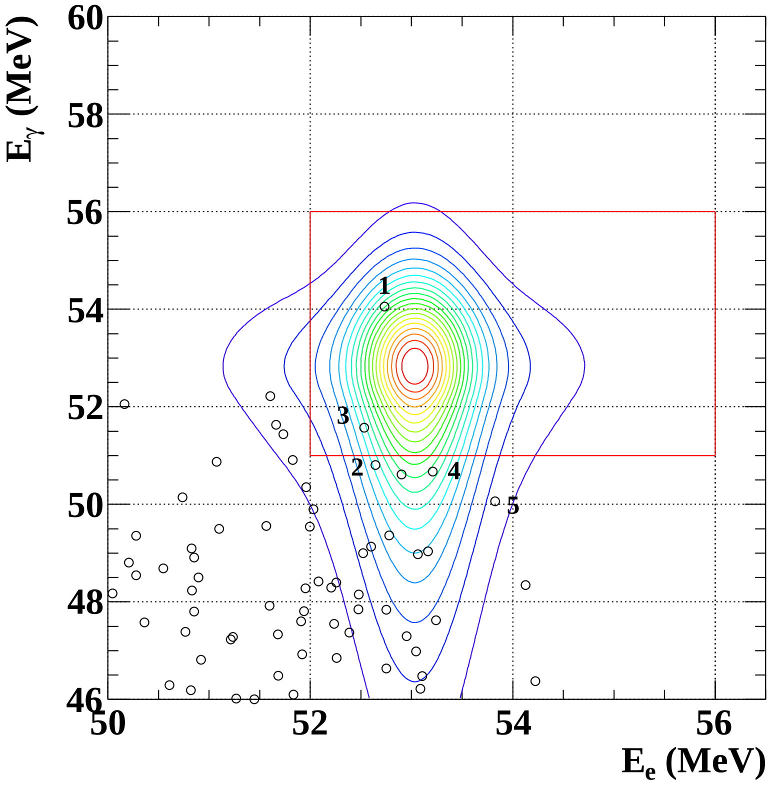


Figure 8.18: Event distributions in signal box B projected on each variable. Blue hatched plots show the signal PDF normalized to 10 events.



(a) Signal box A



(b) Signal box B

Figure 8.19: E_e vs E_γ two-dimensional plots of the signal boxes. Two-dimensional signal PDF is superimposed. The color axis is arbitrary. Numbers in the plots show the rank of S/B listed in the Table 8.6.

8.6.8 Discussion

Figure 8.17 shows that the result of upper limit $N_{sig} < 14.5$ is quite rare. The probability of having an upper limit greater than 14.5 is less than 1 % in the toy experiments. Even though the result is zero-consistent at 90 % level, let us check that the rare result is just due to a statistical fluctuation or indicating something.

Candidate Events

We list up events that are likely to be $\mu^+ \to e^+\gamma$ events by the ratio of event-type likelihoods, S/B. The distribution of S/B is shown in Figure 8.20 together with that for signal-MC events and background-MC ones. Table 8.6 lists the best five signal-like events.

One event is observed at a large value of S/B significantly separated from the distribution. An event display of the event is shown in Figure 8.21. It is found to be a double-pileup event on which the pileup elimination was failed. We successfully eliminated the first pileup, while the second pileup was not detected with the nominal threshold because the conversion of the second-pileup gamma occurred at deep position of the detector and thus the light distribution was smeared. If we set lower threshold for the pileup search, then we can eliminate also the second pileup. Then, E_γ of this event becomes 47.7 MeV and S/B becomes 3.6. The event gets buried in the background distribution. Therefore, most likely this event is not a signal but a background accidentally getting a large signal likelihood because of the reconstructed energy of $E_\gamma = 54.1$ MeV.

This is a very special event. First it is a double-pileup event. Moreover, one of them which has relatively large energy (> 5 MeV) is missed because of the deep conversion point. Such gamma event is overlapped by a positron collinearly, while the timing is not in coincidence so much. The probability that an event is identified to a double pileup in $51 < E_\gamma < 54.5$ MeV is 2 % in the sideband data while that of single pileup is 18 %. The probability of the deep conversion point is about 2 % if the gamma ray came from the target. It is possible that the second pileup is a cosmic-ray event with a probability of ~ 3 %. Thus, the probability of having this kind of events is roughly 0.1 % of normal accidental background in signal box A, that is, 0.001 events are expected.

Impact of the Best Rank Events

To see the effect of such kind of events, we repeated the analysis with a lower threshold for the pileup search so that the second pileup is also eliminated. The PDFs are modified in accordance with the new threshold.

Table 8.6: List of events with large likelihood ratio S/B.

Run/Event	E_e (MeV)	E_γ (MeV)	$t_{e\gamma}$ (ps)	$\theta_{e\gamma}$ (mrad)	$\phi_{e\gamma}$ (mrad)	S/B
35909/1908	52.7	54.1	-262	9.5	6.3	1206.6
34221/2058	52.6	50.8	60.2	3.3	1.8	303.2
30109/1371	52.5	51.6	-50.8	-11.9	11.3	271.9
40330/ 853	53.2	50.7	82.0	-28.6	-2.3	250.4
40077/1210	53.8	50.1	143.6	3.0	14.4	128.6

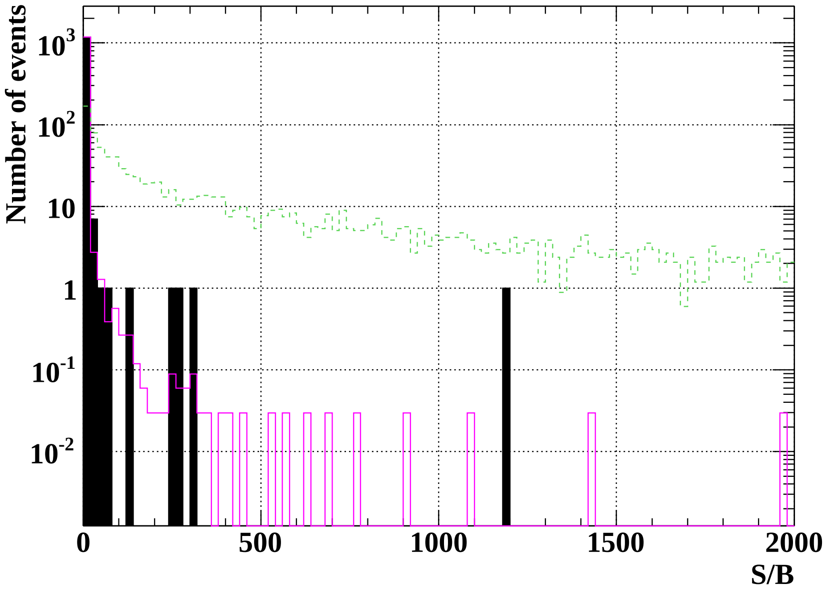


Figure 8.20: Distribution of the ratio of event-type likelihoods, S/B for data (black filled histogram), MC signal (green dashed one), and MC background (magenta blank one). The MC distributions are normalized to the event number of data.

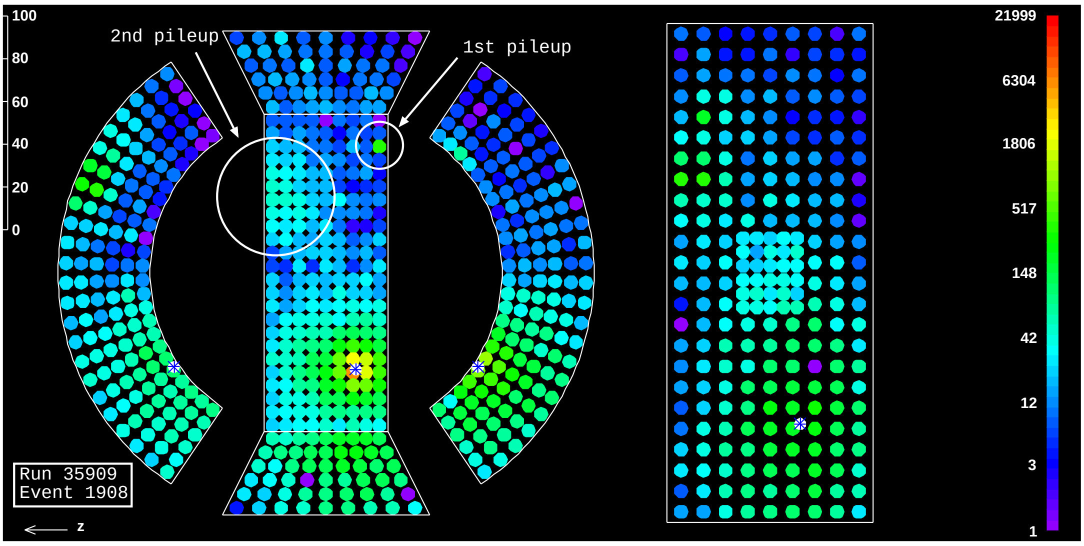


Figure 8.21: Event display of the first-rank event. It found to be a double-pileup event. The color axis shows the amount of observed light on each PMT.

The best-fit values of the maximum likelihood fit becomes $\vec{\theta}^{\text{best}} = (2.0^{+3.8}, 24^{+17}_{16}, 1119^{+37}_{-37})$. The 90%-confidence upper limit on N_{sig} is then calculated to be at 11.4. Change of the likelihood function is shown in Figure 8.22. This event is found to have an impact larger than one event for signal. Now the probability of having an upper limit greater tha 11.4 is ≈ 5 %. The upper limit on the branching ratio becomes 2.2×10^{-11}, where the inefficiency for the signal with the new threshold that amounts to $\sim$5 % is taken into account.

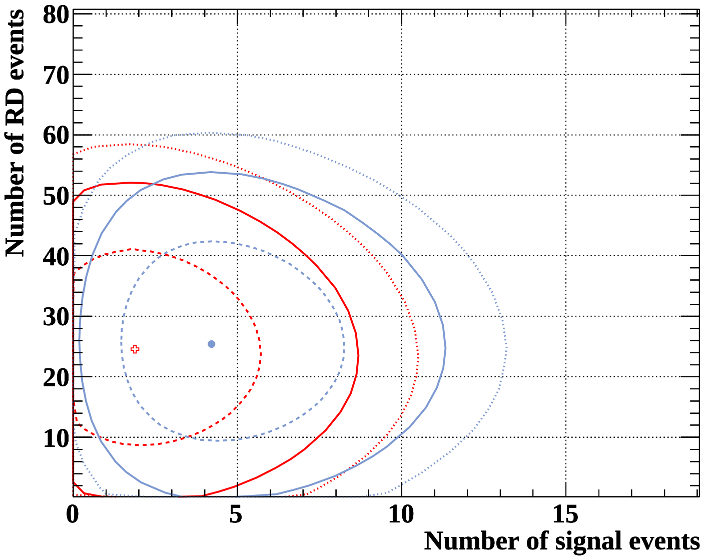


Figure 8.22: Change of the likelihood function by the change of the pileup search threshold. Red contours show new ones while blue ones are ones of original likelihood function.

Relationship with Cut Analysis

The sensitivity of cut analysis B is comparable with that of the likelihood analysis; the difference is less than 10 %.

The cut analyses show well consistent results with their background expectations, and no indication of $\mu^+ \to e^+\gamma$ decay is found. When looking at the event distribution in Figure 8.19(b), we see some events around the edge of the box. Those events are almost overlapped with the signal-like candidates in the likelihood analysis. Potentially, some of them could be observed inside the box by fluctuations in the measurement or the decision of the box size. If the closest two or three events were additionally observed in the box, the upper limit on the branching ratio of cut analysis B would be 2.3×10^{-11} or 2.8×10^{-11}, respectively. They would be comparable with the likelihood-analysis result. Conversely, if we removed the double-pileup event, it would be 8.8×10^{-12}.

The upper limit obtained by cut analysis B is 2.2 times better than that by the likelihood analysis. However, we adopt the result of the likelihood analysis as our result, because we decided to carry on with the likelihood analysis before opening the hidden box. If we decide the analysis after looking at their results, it just introduces bias in the result.

In conclusion, we consider the result of large N_{sig} that is about to exclude null signal at 90 % confidence level is due to a statistical fluctuation by a very rare accidental background event. The result is statistically true, but we cannot say there is some indication of signal from those data. Since the background PDF is formed by the actual measured spectra, it takes into account those rare events in principle. We analyzed fully in blinded. Therefore, the obtained upper limit is valid without any bias from the statistical point of view.

We can also consider that this fluctuation is one of the systematic uncertainty from

Table 8.7: Summary of the upper limit and sensitivity.

	$\mathcal{B}$ upper limit[1] (10^{-11})
Analysis window	2.8
Analysis window (tighter pileup cut)	2.2
Cut analysis (optimized box)	1.3
Sideband ($t_{e\gamma} > 0$)	2.0
Sideband ($t_{e\gamma} < 0$)	0.89
Sensitivity	1.3

---

[1] before incorporating systematic uncertainties.

the pileup cut criteria. If we take in the systematic uncertainty, the confidence region is broadened and the zero point is naturally included in the 90%-level confidence region as we saw in the change of likelihood function.

The obtained limit is not as stringent as the current best limit of 1.2×10^{-11} set by MEGA. However, when we think of the relation or comparison with other experimental results, it is more important to consider sensitivities than actual-obtained upper limits because the latter are subject to statistical fluctuation as see in our result. Our sensitivity is 1.3×10^{-11}. Therefore, through this analysis we searched for the $\mu^+ \rightarrow e^+\gamma$ decay with a comparable sensitivity to the current limit, and we set an independent upper limit.

The sensitivity is mainly limited by the data statistics. By continuing data taking, we can progressively improve the sensitivity. From here on, new data start to search for the decay in unexplored region. We discuss prospects of the experiment in the following section.

8.7 Future Prospects

MEG resumed data taking at the beginning of November 2009, and will accumulate data at least for two years. Since PSI has long accelerator-shutdown period (from end of December to April) every year, the run is divided into sub data sets for every calender year. First, let us think of the prospect of the run 2009.

The largest improvement achieved in 2009 is the full operation of the drift chamber. We found the causes of the problem and solved them during the long shutdown period after run 2008. All the chambers were modified. Almost all chambers are now operated at the nominal voltage.

First, consider the expected data statistics of run 2009. The positron efficiency will improve to $\epsilon_e = 0.4$ ($\epsilon_{DCH} \times A_{TIC} = 0.8 \times 0.5$), factor 3.0 improvement. The trigger efficiency is also expected to be improved by a better look-up table of direction match and looser conditions, from 0.66 to 0.88, factor 1.3. We optimized calibration time, which recover dead time by factor 1.3. The beam rate was adjusted to a slightly higher intensity, factor 1.1. On the other hand, the beam time of run 2009 is factor 0.56 of that of run 2008. Taking all the factors into account, we expect 3.2 times larger statistics than that of run 2008.

Owing to the full operation of the drift chamber, the resolutions are also expected to improve. In [41], a realistic evaluation of the positron spectrometer performance was conducted. Here, we assume the following resolutions in reference to the study:

- E_e resolution with a single Gaussian whose sigma is the same value as that of core Gaussian in 2008 ($\sigma_{E_e} = 374$ keV).
- Angular resolutions of $\sigma_{\theta_e} = 5.2$, $\sigma_{\phi_e} = 9.0$ mrad.
- Time resolution of $t_{e\gamma} = 122$ ps improved in track length measurement.

PDFs for signal and backgrounds are re-formed using these values.

Now, we calculate the expected sensitivity of run 2009. We use a similar signal box to that of B in 2008. The window is narrowed in accordance with the resolution improvement with keeping the same signal efficiency. Thus, the expected single-event sensitivity is SES $= (3.6 \times 10^{-12})/3.2 = 1.1 \times 10^{-12}$. The expected number of background events in the signal box is estimated using the toy MC. In this calculation, we take the higher beam intensity into account assuming $N_{BG} \propto R_\mu^2$. The number of background events is estimated to be $N_{BG}^{\mathrm{exp}} = 0.63$. The expected 90%-confidence sensitivity is estimated to be

$$\mathcal{S}_{2009}^{\mathrm{exp}} = 3.2 \times 10^{-12}. \tag{8.26}$$

We are still limited by the statistics. Note we here only assumed improvements in the drift chamber.

Finally, let us estimate the expected sensitivity of MEG with current knowledge. MEG will run at least additional two years toward the sensitivity of $\mathcal{O}(10^{-13})$. Let assume beam time of 100 weeks (8.7 times longer than run 2008). We expect further improvement of efficiencies: trigger efficiency (to 95 %), DAQ livetime (to 90 %), and positron efficiency (to 50 %). In total, 7.8 times higher efficiency can be expected. Together with the beam time, 68 times statistics will be accumulated.

Under the conditions assumed in the estimation of run 2009, we expect to observe background events as large as 13 events. Thus, we have to suppress background further.

We think there is large possibility of improvement in the gamma-ray energy measurement. We are now considering three directions of improvements. First, we will be able to improve the resolution at least at the center of the detector to 1.3 % level, because we verified such resolution using the prototype detector. The difference is size of the detector and geometrical arrangement. We think more precise calibration of PMT QE enables to go to the level. Next, we can improve the worse resolution at the edge part of the detector. For example, by optimizing the weighting factor of PMT summation for each position, we can effectively reconstruct the total number of scintillation light using the anti-correlation between faces of the detector. Preliminary study shows this kind of analysis can improve resolutions at large part of the detector acceptance to the same level as that at the center. In the analysis of run 2008, we could not confirm the improvement because of unstable conditions of the calibration run. However, further study with more stable data in 2009 will achieve a resolution of 1.7 % for all acceptance. Last point is improvement of the pileup elimination. It was, for the first time, introduced in the analysis of run 2008. We think we can improve it with further study. Therefore, here we assume the following:

- the energy resolution will become 1.7 % for $w_\gamma \geq 2$, and 2.5 % for $0 \leq w_\gamma < 2$ cm, uniformly over the acceptance.
- the pileup level after elimination will go down to the level where we achieved when we rejected all events identified as pileup in the 2008 analysis.

In addition, we expect the time resolution of the gamma-ray detector will be improved to $\sigma_{t_\gamma} = 65$ ps by the increase in the light yield. According to [41], the positron energy resolution will improve to $\sigma_{E_e} = 335$ keV.

Using these assumptions, corresponding PDFs are formed. Then we optimize the signal box. We can calculate the expected number of background and efficiency for the signal using the PDFs. The two-dimensional PDFs and the optimized box are shown in Figure 8.23. The corresponding signal efficiency is 0.45 relative to that of the analysis window. The SES would be 6.6×10^{-14}. The expected number of background events in the box is estimated to be 4.0. Then the expected 90%-confidence sensitivity would be given as

$$\mathcal{S}^{\text{exp}} = 3.0 \times 10^{-13}. \tag{8.27}$$

If we pose ourselves that the background level should be ≤ 1, we can tune the beam intensity to suppress background. From the relation $N_{BG} \propto R_\mu^2$, the expected sensitivity as a function of R_μ can be calculated as shown in Figure 8.24 together with the expected number of backgrounds. At a beam intensity of $R_\mu = 1.5 \times 10^7$, half of current nominal value, the background would be 1. In this case, the sensitivity becomes $S^{\text{exp}} = 4.4 \times 10^{-13}$. The 90%-confidence belts in Feldman-Cousins method are shown in Figure 8.25. With this sensitivity, we can set a two-sided interval at this C.L. when we observed at least three events. In this case, the confidence interval would be set to

$$1.3 \times 10^{-14} < \mathcal{B}(\mu^+ \to \mathrm{e}^+\gamma) < 8.4 \times 10^{-13}. \tag{8.28}$$

at 90 % C.L.

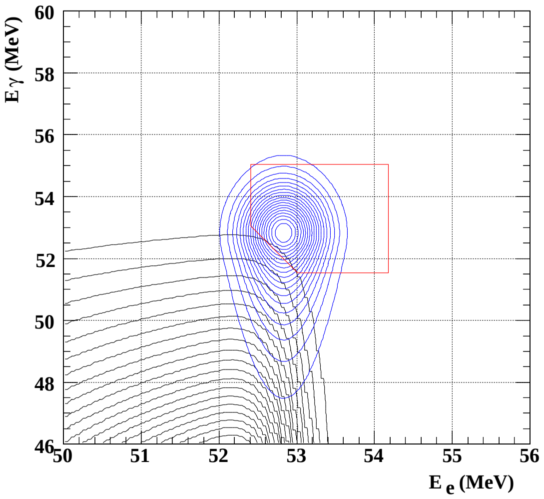


Figure 8.23: Two-dimensional PDFs with expected performance. Blue contour shows signal PDF and black one shows BG PDF. The normalization of the two are arbitrary. Red box shows the optimized signal box.

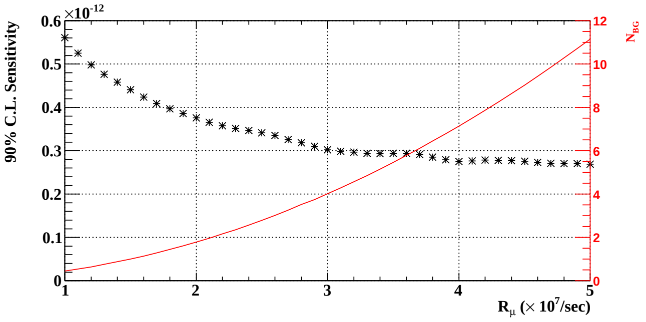


Figure 8.24: Expected 90%-confidence sensitivity as a function of muon stopping rate.

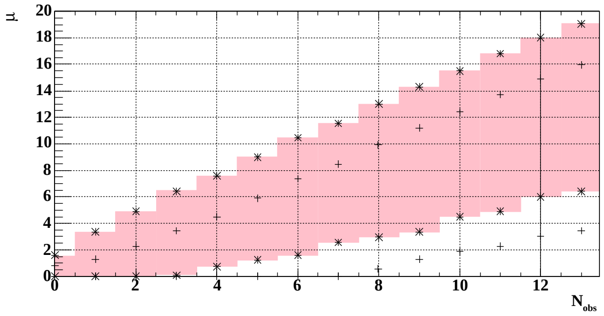


Figure 8.25: 90%-confidence belt in Feldman-Cousins method with the Poisson statistics in presence of background of 1 (colored belt) and 4 events (interval between cross markers).

Chapter 9

Conclusion

The MEG experiment started physics data taking in September 2008 to search for the lepton-flavor-violating muon decay, $\mu^+ \to \mathrm{e}^+\gamma$. We searched for the decay on the basis of the initial three months data of MEG. The accumulated number of 9.5×10^{13} muons were stopped on the target. Positrons from the muon decays were measured by a spectrometer which forms a specially graded magnetic field. Gamma rays were detected by a 900 liter liquid xenon scintillation detector.

To avoid introducing any artificial bias in the search, we adopted a blind analysis with a hidden box. We established precise reconstruction and calibration methods, and evaluated detector performance using calibration data samples. The backgrounds were evaluated using the sideband data.

The 90%-confidence sensitivity of this search is

$$\mathcal{S}_{2008} = 1.3 \times 10^{-11} \tag{9.1}$$

with the number of observed muon decays of $(5.2 \pm 0.5) \times 10^{12}$. It is comparable with the current best upper limit of 1.2×10^{-11} set by MEGA.

We performed a maximum likelihood fit to detect $\mu^+ \to \mathrm{e}^+\gamma$ signal. The best-estimated value of the number of signal events is 4.3. The 90%-confidence interval includes $N_{sig} = 0$, and the upper limit is at 14.7 including systematic uncertainties. Therefore, we set an upper limit on the branching ratio

$$\mathcal{B}(\mu^+ \to \mathrm{e}^+\gamma) < 2.8 \times 10^{-11} \quad (90\ \%\mathrm{C.L.}). \tag{9.2}$$

We also performed an independent analysis by cutting background events and counting remaining events inside a signal box for a cross check of the analysis. We expect 2.1 background events, and we observed 2 events in the box. This result yields an upper limit of $\mathcal{B} < 1.3 \times 10^{-11}$ at 90 % C.L.

We conclude the reason of the large number of N_{sig} and the large value of upper limit compared to the sensitivity in the likelihood analysis is a statistical fluctuation. We do not give a new record of the experimental limit but an independent result from a measurement with a comparable sensitivity.

The sensitivity is limited by the statistics. The observed number of muon decays was less than expected mainly because of a problem on the drift chamber system. In 2009, MEG runs with a fixed drift chamber system, and accumulates further data. The sensitivity is going to improve as the data statistics. It is promising that MEG explores unexplored region of the branching ratio. The sensitivity will go down to $\sim 3 \times 10^{-12}$ with the data 2009, and finally to $\sim 3 \times 10^{-13}$.

Appendix A

Radiative Decay Analysis

In chapter 2, we introduced the radiative muon decays, $\mu^+ \to e^+\nu_e\overline{\nu}_\mu\gamma$, as one of background sources. On the other hand, we used them as as a powerful timing-calibration source in chapter 6 and 7. In addition, we can use them to check our analysis. Detecting RD events as expected demonstrates the capability of our experiment to really observe $\mu^+ \to e^+\gamma$ events if there are. For example in MEGA, they could not see the RD peak in time distribution because of the worse signal-to-noise ratio as a consequence of their lower time resolution and higher instant beam intensity[1]. We see a clear peak as shown in Figure 7.26.

A.1 E_γ-Sideband Data

We analyzed the RD events in the E_γ-sideband data. Here we define a sideband wide box

- $40 < E_\gamma < 48$ MeV,
- $48 < E_e < 56$ MeV,
- $|\theta_{e\gamma}| < 100$ and $|\phi_{e\gamma}| < 100$ mrad.

The partial branching ratio of this box is calculated to be 25.5×10^{-10} by numerically integrating the formula of the differential branching ratio.

To calculate the expected number of RD events in the box, detector resolutions and efficiencies have to be considered. We put the following inputs for the calculation:

- E_γ efficiency: An efficiency curve as a function of reconstructed energy obtained from the ratio between the gamma-ray background rate in normal-threshold MEG trigger (trigger0) and that in lower-threshold MEG trigger (trigger1). It is given as an error function.
- E_e acceptance: An acceptance curve as a function of the real positron energy estimated by the Michel spectrum fit. It is given as an error function.

---

[1] We do not say they had no capability of detecting RD events. By using all kinematic variables, they detected RD events in the likelihood fit as expected.

- Direction-match efficiency: Probability that an event passes the direction-match trigger condition is estimated by two methods relatively to the signal (back-to-back) efficiencies. One is using unbiased RD trigger data in the dedicated RD run with ultra-low beam intensity. The other is by looking at the angle distribution of accidental background events.
- E_γ response: A representative response function obtained by averaging the position-dependent response function over actual event distribution.
- E_e response: The same response function as for $\mu^+ \to e^+\gamma$ analysis (a triple Gaussian) obtained by the Michel-spectrum fitting.
- Angular response: The same response function as for $\mu^+ \to e^+\gamma$ analysis. Treat $\theta_{e\gamma}$ and $\phi_{e\gamma}$ separately.

Those efficiency curves are shown in Figure A.1. We applied these effects to the theoretical three-dimensional distribution, that is, multiplied efficiency curves and convolved response functions. Then we integrated the formed distribution over the box. The effective branching ratio is calculated to be $\mathcal{B}_{RD}^{\mathrm{wide}} = (5.6 \pm 0.8) \times 10^{-10}$, where the uncertainty comes from those of efficiencies and resolutions, mainly from the E_e-acceptance curve.

We performed a maximum likelihood fit on the data in the box. The best-fit value is at $N_{RD} = 244 \pm 47$, as shown in Figure A.2.

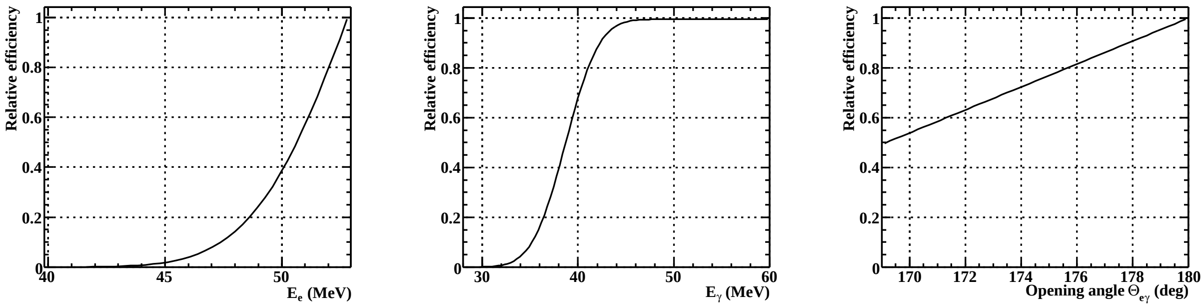


Figure A.1: Efficiency curves applied to the calculation of RD branching ratio. They are relative to those for the signal. The E_e efficiency is mainly due to the design of the spectrometer, and the other two are due to trigger conditions.

A.2 Normalization by Radiative Decay

From the relation

$$k \cdot \mathcal{B}_{RD} = N_{RD}, \tag{A.1}$$

we calculate the normalization factor k from the radiative decay,

$$k_{RD} = 244/(5.6 \times 10^{-10}) = (4.4 \pm 1.1) \times 10^{11}. \tag{A.2}$$

The relative uncertainty is 24 %. The number is consistent with that by Michel decays in Sec.8.5.

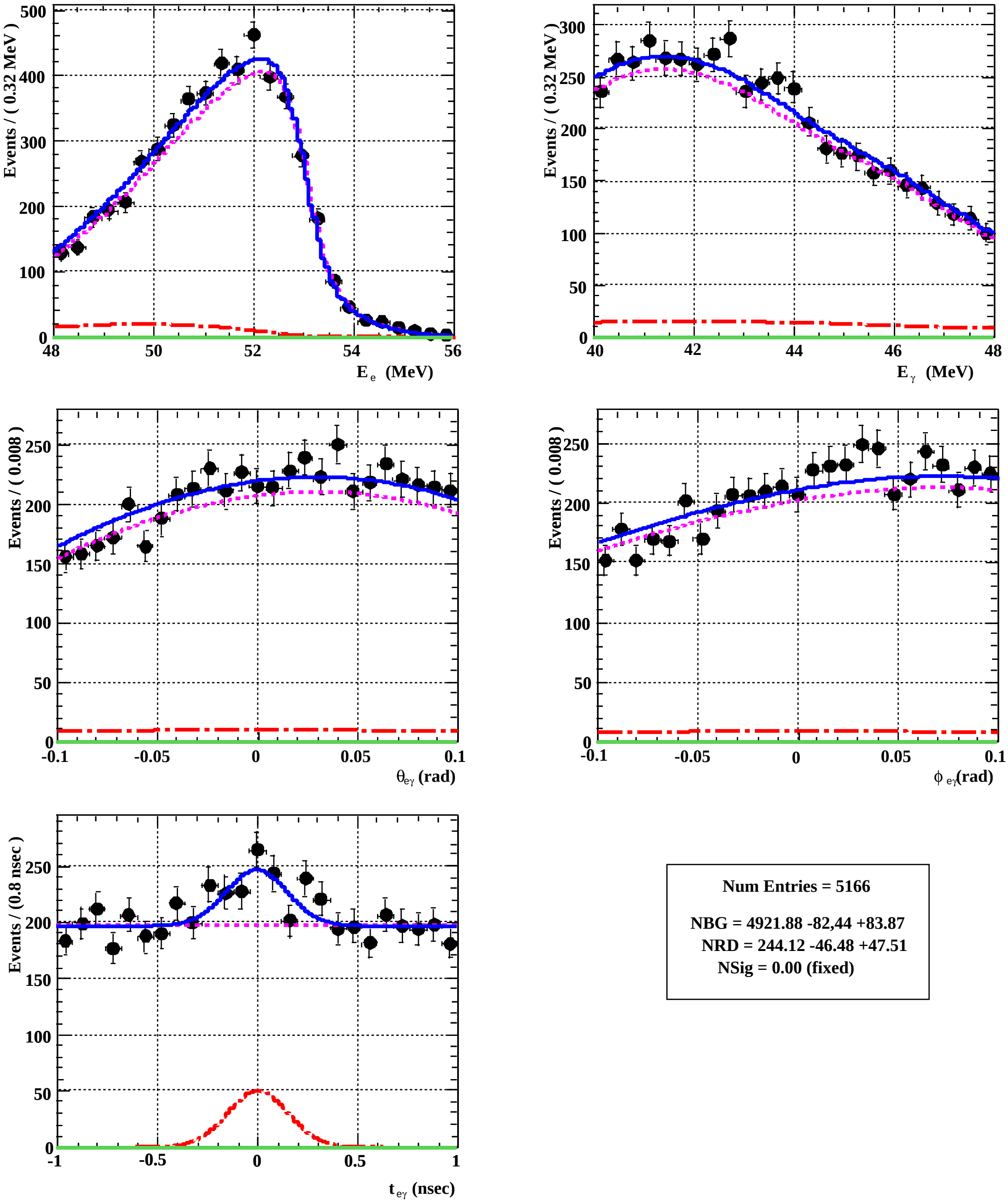


Figure A.2: Result of a maximum likelihood fit on the E_γ-sideband data. Projected event distributions are shown. Red dot-dashed line and magenta dash line show the fitted RD and BG PDFs, respectively. The number of signal was fixed to zero in this fit. Blue solid line is the total fitted spectrum.

A.3 Calculation of Expected Number of Radiative-Decay Events

We calculate the expected number of RD events in several boxes used in the analyses by the ratio of the calculated effective branching ratios. The numbers are summarized in Table A.1. We expect 41 ± 8 RD events in the analysis window.

Table A.1: Effective branching ratio of radiative decay in several boxes.

Window	E_γ (MeV)	E_e (MeV)	$\theta_{e\gamma}$,$\phi_{e\gamma}$ (mrad)	$\mathcal{B}_{RD}$	N_{RD}
Analysis window	$[46, 60]$	$[50, 56]$	100	0.95×10^{-10}	41
Sideband wide window	$[40, 48]$	$[48, 56]$	100	5.6×10^{-10}	244
Signal box A (1.64σ cuts)	$[51, 54.6]$	$[52.2, 53.8]$	33[1]	4.4×10^{-14}	0.019
Signal box B (optimized)	$[51, 56]$	$[52, 56]$	42[1]	7.6×10^{-14}	0.033

---

[1] for $\pi - \Theta_{e\gamma}$

A.4 Detail Description of Radiative-Decay Differential Branching Ratio

The differential branching ratio of RD is given by Eq.2.17. The functions appearing in Eq.2.17, $F(x, y, d)$, $G(x, y, d)$, and $H(x, y, d)$ in the SM, are given as follows:

$$\begin{aligned} F &= F^{(0)} + rF^{(1)} + r^2F^{(2)}, \\ G &= G^{(0)} + rG^{(1)} + r^2G^{(2)}, \\ H &= H^{(0)} + rH^{(1)} + r^2H^{(2)}, \end{aligned} \tag{A.3}$$

where $r = (m_e/m_\mu)^2$.

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

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

$$F^{(2)}(x,y,d) = \frac{32}{d^2}\left\{\frac{(4-3y)}{x} - 3\right\} + \frac{48y}{d}, \tag{A.6}$$

$$\begin{aligned} G^{(0)}(x,y,d) &= \frac{8}{d}\{xy(1-2y) + 2x^2(1-3y) - 4x^3\} \\ &\quad +4\{-x^2(2-3y-4y^2) + 2^3(2+3y)\} \end{aligned}$$

$$-4dx^3y(2+y), \tag{A.7}$$

$$G^{(1)}(x,y,d) = \frac{32}{d^2}(-1+2y+2x)+\frac{8}{d}(-xy+6x^2)-12x^2(2+y), \tag{A.8}$$

$$G^{(2)}(x,y,d) = -\frac{96}{d^2}, \tag{A.9}$$

$$\begin{aligned} H^{(0)}(x,y,d) &= \frac{8}{d}\{y^2(1-2y)+xy(1-4y)-2x^2y\} \\ &\quad +4\{2xy^2(1+y)-x^2y(1-4y)+2x^3y\} \\ &\quad 2d\{x^2y^2(1-2y)-4x^3y^2\} \\ &\quad 2d^2x^3y^3, \end{aligned} \tag{A.10}$$

$$\begin{aligned} H^{(1)}(x,y,d) &= \frac{32}{d^2}\left\{-\frac{y(1-2y)}{x}+2y\right\} \\ &\quad +\frac{8}{d}\{y(2-5y)-xy\} \\ &\quad +4xy(2y-3x)+6dx^2y^2, \end{aligned} \tag{A.11}$$

$$H^{(2)}(x,y,d) = -\frac{96y}{d^2x}+\frac{48y}{d}. \tag{A.12}$$

Appendix B

Normalization with Background

B.1 Overall Efficiency

We estimated the overall efficiency of our measurement with $\mu^+ \rightarrow \mathrm{e}^+\gamma$ like background events. By counting the number of accidental background events in the $t_e\gamma$ sideband, the overall efficiency can be calculated.

The number of accidental background events is given by

$$\begin{aligned} N_{BG} &= (R_\mu \cdot f_e^0 \cdot \frac{\Omega}{4\pi} \cdot \epsilon_e) \times (R_\mu \cdot f_\gamma^0 \cdot \frac{\Omega}{4\pi} \cdot \epsilon_\gamma) \times \frac{\Delta\omega}{\Omega} \times \Delta t \times T \\ &= N_\mu^2 \cdot f_e^0 \cdot f_\gamma^0 \cdot \frac{\Delta t}{T} \cdot \frac{\Delta\omega}{4\pi} \cdot \epsilon, \end{aligned} \tag{B.1}$$

where R_μ is a muon intensity and $N_\mu(= R_\mu T)$ is the total number of muon; T is the total livetime; f_e^0 and f_γ^0 are the fractions within the signal window of Michel spectrum and background-gamma spectrum, respectively; $\Delta\omega$ and Δt is full width of a signal window for solid angle and time; Ω is the geometrical acceptance; ϵ_e and ϵ_γ are detection (including selection and analysis) efficiencies of positron and gamma; and ϵ is the overall efficiency including the acceptance.

Window We set a window in which we count the number of background events as

- $50 < E_e < 56$ MeV,
- $46 < E_\gamma < 60$ MeV,
- $1 < |t_{e\gamma}| < 4$ ns,
- $|\theta_{e\gamma}| < 100$ and $|\phi_{e\gamma}| < 100$ mrad.

The solid angle of the window is calculated to be $\Delta\omega/4\pi = 3.19 \times 10^{-3}$.

Number of background events $N_{BG} = 3577$ events are observed in the window.

Number of muons The number of muons contributing to the background is slightly different from that of muons stopping on the target. For the gamma-ray background, we assume all muons reaching the target contributes; the average rate is 3.54×10^7/sec. Positrons from muon decay off target less contributes to the background owing to the tracking, but it has large uncertainty; $(3.2 \pm 0.32) \times 10^7$. The livetime is $T = 3.3 \times 10^6$ sec. The effective number of muons is calculated to $N_\mu^{eff} = 11.1 \times 10^{13}$.

Michel spectrum fraction This is the same as that calculated in Sec 8.5, $f_e^0 = 0.101 \pm 0.006$ for $E_e > 50$ MeV. Correction factors of efficiency ratio between signal and Michel are $1.11 \times 1.02 = 1.13 \pm 0.023$.

Gamma-ray background spectrum fraction The fraction of RD-gamma spectrum over 46 MeV is calculated to be 1.20×10^{-4}. The effect of detector resolution is evaluated to be a factor 1.09. Thus, $f_\gamma^0 = (1.31 \pm 0.085) \times 10^{-4}$, where the error comes from the uncertainty of energy scale. The fraction of gamma from AIF is estimated to be (24.6 ± 1.0) % for $46 < E_\gamma < 60$ MeV. Thus, $f_\gamma^{eff} = (1.63 \pm 0.13) \times 10^{-4}$.

Effect of low-energy tail in E_γ response The difference of survival rate of gamma rays between signal and background needs to be considered. The fraction of the RD events over 46 MeV surviving in the window is 50 %, while that of the signal events is 66 %. Thus, we correct the efficiency by a factor $66/50 = 1.32$.

Direction match efficiency ratio In Eq.B.1, a uniform efficiency over the window is assumed, which is not the case because of the dependence of the trigger efficiency on the opening angle. The average efficiency relative to that at $\Theta_{e\gamma} = 180°$ is evaluated to be 83 ± 2 %.

Putting numbers above in Eq.B.1, the overall efficiency is calculated to be

$$\epsilon = (5.4 \pm 0.8) \times 10^{-3}. \tag{B.2}$$

This number is consistent with the multiplication of individual components estimated separately,

$$\begin{aligned} \epsilon &= \Omega/4\pi \times \epsilon_e \times \epsilon_\gamma \times \epsilon_{trigger} \\ &= 0.1 \times 0.14 \times 0.63 \times 0.66 \\ &= 5.8 \times 10^{-3}. \end{aligned} \tag{B.3}$$

B.2 Normalization Factor

Using the total number of stopping muon, $N_\mu = 9.5 \times 10^{13}$, the normalization factor is calculated to be

$$k_{BG} = (5.2 \pm 0.8) \times 10^{11}, \tag{B.4}$$

which is consistent with Eq.8.19.

Acknowledgment

First of all, I would like to express my great appreciation to my supervisor Prof. Toshinori Mori.

I would like to express my gratitude to Dr. Wataru Ootani, Dr. Toshiyuki Iwamoto, Dr. Ryu Sawada, Prof. Satoshi Mihara, Dr. Hajime Nishiguchi and Dr. Shuei Yamada for their helpful advice. I would like to thank Mr. Hiroaki Natori, Mr. Yasuhiro Nishimura, Dr. Kenji Ozone, Ms. Yasuko Hisamatsu, Ms. Bai Xue, Mr. Daisuke Kaneko, and Mr. Yuki Fuji for their kind cooperation and joyful research time with them.

I appreciate all the members of the MEG collaboration. I learned a lot from Prof. Akihiro Maki, Prof. Tom Haruyama, and Mr. Katsuyu Kasami; Dr. Stefan Ritt, Dr. Peter R. Kettle, Dr. Malte Hildebrandt, Dr. Oleg Kiselev, Dr. Matthias Schneebeli, and Ms. Jeanine Adam; Prof. William Molzon, Mr. Feng Xiao, Mr. Ben Golden, Ms. Christina Topchyan, and Dr. Elisabetta Baracchini; Dr. Dmitry N. Grigoriev, Dr. Boris I. Khazin, Dr. Fedor Ignatov, and Dr. Yuri V. Yudin; Prof. Flavio Gatti, Dr. Simeone Dussoni, Dr. Paolo W. Cattaneo, Prof. Gianluigi Boca, Prof. Giancarlo Piredda, Dr. Gianluca Cavoto, Dr. Cecilia Voena, Dr. Francesco Renga, Dr. Giovanni Signorelli, Dr. Fabrizio Cei, Mr. Giovanni Gallucci, Prof. Carlo Bemporad, Dr. Angela Papa, Dr. Marco Panareo, Dr. Marco Grassi, Dr. Donato Nicolò, Mr. Luca Galli, and Dr. Alessandro Baldini, who, together with Prof. Mori, shares the spokes-person of the MEG experiment. I enjoyed discussions with them. I really appreciate an exciting experience with them.

I would like to thank PSI staff for their support of the experiment and also my Swiss life.

I am deeply grateful to all of ICEPP members and secretaries.

I got financial support from Japan Society for the Promotion of Science. Part of my travel cost was also supported by PSI.

Finally, I gratefully acknowledge my family for a lot of encouragement and support.

December 2009,
Tokyo,
Yusuke Uchiyama

Bibliography

[1] M. Ahmed *et al.*, "Search for the lepton-family-number nonconserving decay $\mu^+ \rightarrow e^+\gamma$ ", Phys. Rev. D **65**, 112002 (2002); M. L. Brooks *et al.*, "New Limit for the Lepton-Family-Number Nonconserving Decay $\mu^+ \rightarrow e^+\gamma$ ", Phys. Rev. Lett. **83**, 1521-1524 (1999).

[2] T. Mori *et al.*, "Search for $\mu^+ \rightarrow e^+\gamma$ down to 10^{-14} branching ratio", Research Proposal to Paul Scherrer Institut, R-99-05 (1999), `http://meg.web.psi.ch/docs/prop_psi/index.html/`

[3] L. Michel, "Interaction between Four Half-Spin Particles and the Decay of the μ-Meson", Proc. Phys. Soc. A **63**, 514-531 (1950).

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

[5] R. R. Crittenden *et al.*, "Radiative Decay Modes of the Muon", Phys. Rev. **121**, 1823-1832 (1961).

[6] W. Bertl *et al.*, "Search for the decay $\mu^- \rightarrow e^-e^+e^-$", Nucl. Phys. B **260**, 1.31 (1985).

[7] S. Freedman *et al.*, "Limits on neutrino oscillations from $\bar{\nu}_e$ appearance", Phys. Rev. D **47**, 811-829 (1993).

[8] U. Bellgardt *et al.*, "Search for the decay $\mu^- \rightarrow e^-e^+e^-$", Nucl. Phys. B **299**, 1-6 (1988).

[9] R. D. Bolton *et al.*, "Search for rare muon decays with the Crystal Box detector", Phys. Rev. D **38**, 2077-2101 (1988); R. D. Bolton *et al.*, "Search for the Decay $\mu^+ \rightarrow e^+\gamma$ ", Phys. Rev. Lett. **56**, 2461-2464 (1986).

[10] M. Gell-Mann, P. Ramond, and R. Slansky, in "Supergravity", edited by D. Freedman and P. van Nieuwenhuizen (North-Holland, Amsterdam), 315, (1979).

[11] T. Yanagida, in Proceedings of Workshop on Unified Theory and Baryon Number in the Universe, New York, KEK Report No. 79-18, 95.

[12] J. Hisano and D. Nomura, "Solar and atmospheric neutrino oscillations and lepton flavor violation in supersymmetric models with right-handed neutrinos", Phys. Rev. D **59**, 116005, (1999).

[13] R. Barbieri and L. J. Hall, "Signals for supersymmetric unification", Phys. Lett. B **338**, 212-218 (1994); R. Barbieri, L. J. Hall, and A. Strumia, "Violations of lepton flavour and CP in supersymmetric unified theories", Nucl. Phys. B **445**, 219-251 (1995).

[14] J. Hisano *et al.*, “Exact event rates of lepton flavor violating processes in supersymmetric $SU(5)$ model”, Phys. Lett. B **391**, 341-350 (1997).

[15] J. Hisano *et al.*, “Waiting for $\mu \rightarrow e\gamma$ from the MEG experiment”, arXiv:0904.2080 [hep-ph] (2009).

[16] L. Calibbi *et al.*, “Lepton flavour violation from supersymmetric grand unified theories: Where do we stand for MEG, PRISM/PRIME and a super flavour factory”, Phys. Rev. D **74**, 116002 (2006).

[17] K. S. Hirata *et al.*, “Experimental limits on nucleon lifetime for lepton+meson decay modes”, Phys. Lett. B **220**, 308-316 (1989).

[18] Y. Kuno and Y. Okada, “Muon decay and physics beyond the standard model”, Rev. Mod. Phys. **73**, 151-202 (2001).

[19] M. Raidal *et al.*, ”Flavour physics of leptons and dipole moments”, report of the CERN workshop “Flavour in the era of the LHC”, arXiv:0801.1826 [hep-ph] (2008).

[20] S. H. Neddermeyer and C. D. Anderson, “Note On The Nature Of Cosmic Ray Particles”, Pyhs. Rev. **51**, 884-886 (1937).

[21] H. Yukawa, “On the Interaction of Elementary Particles”, Proc. Phys. - Math. Soc. Jap, **17** 48-57 (1935).

[22] M. Conversi, E. Pancini, and O. Piccioni, “On the Disintegration of Negative Mesons”, Phys. Rev. **71** 209-210 (1947).

[23] E. P. Hincks and B. Pontecorve, “Search for Gamma-Radiation in the 2.2-Microsecond Meson Decay Process”, Phys. Rev. **73**, 257-258 (1947).

[24] J. Steinberger, “On the Range of the Electrons in Meson Decay”, Phys. Rev **74**, 500-501 (1948); *ibid* **75**, 1136-1143 (1949).

[25] Lokonathan, S. and J. Steingerger, “Search for Improbable Muon Decay”, Phys. Rev. **98**, 240 (1955).

[26] G. Feinberg, “Decays of the μ Meson in the Intermediate-Meson Theory”, Phys. Rev. **110**, 1482-1483 (1958).

[27] G. Danby *et al.*, “Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos”, Phys. Rev. Lett. **9**, 36-44 (1962).

[28] A. van der Schaaf *et al.*, “A SEARCH FOR THE DECAY $\mu^+ \rightarrow e^+\gamma$ ”, Nucl. Phys. **A340**, 249-270 (1980); H. P. Povel *et al.*, “A NEW UPPER LIMIT FOR THE DECAY $\mu^+ \rightarrow e^+\gamma$ ”, Pyhs. Lett. B **72**, 183-186 (1977).

[29] P. Depommier *et al.*, “New Limit on the Decay $\mu^+ \rightarrow e^+\gamma$ ”, Phys. Rev. Lett. **39**, 1113-1116 (1977); G. Azuelos *et al.*, “New Upper Limit on the Decay $\mu \rightarrow e\gamma\gamma$”, *ibid.* **51**, 164-167 (1983).

[30] W. W. Kinnison *et al.*, "Search for $\mu^+ \rightarrow e^+\gamma$ ", Phys. Rev. D **25**, 2846-2868 (1982); J. D. Bowman *et al.*, "Upper Limit for the Decay $\mu^+ \rightarrow e^+\gamma$ ", Phys. Rev. Lett. **42**, 556-560 (1979).

[31] A. E. Pifer, T. Bowen, and K. R. Kendall, "A HIGH STOPPING DENSITY μ^+ BEAM", Nucl. Instr. Meth. **135**, 39-46 (1976); H. -W. Reist *et al.*, "DEVELOPMENT OF A LOW-MOMENTUM "SURFACE" MUON BEAM FOR LAMPF", *ibid* **153**, 61-64 (1978).

[32] T. Kinoshita and A. Sirlin, "RADIATIVE DECAY OF THE MUON", Phys. Rev. Lett. **2**, 177-178 (1959).

[33] C. Fronsdal and H. Überall, "μ-Meson Decay with Inner Bremsstrahlung", Phys. Rev. **113**, 654-657 (1959).

[34] Y. Kuno and Y. Okada, "Proposed $\mu \rightarrow e\gamma$ Search with Polarized Muons", Phys. Rev. Lett. **77**, 434-437 (1996).

[35] Y. Kuno, A. Maki, and Y. Okada, "Background suppression for $\mu \rightarrow e\gamma$ with polarized muons", Phys. Rev. D **55**, R2517-R2520 (1997).

[36] T. Kinoshita and A. Sirlin, "Radiative Corrections to Fermi Interactions", Phys. Rev. **113**, 1652-1660 (1959).

[37] W. Wagner *et al.*, "PSI status 2008 - Developments at the 590 MeV proton accelerator facility", Nucl. Instr. Meth. A **600**, 5-7 (2009).

[38] ROHACELL®, Evonik Röhm GmbH, http://www.rohacell.com/.

[39] W. Ootani *et al.*, "Development of a Thin-wall Superconducting Magnet for the Positron Spectrometer in the MEG Experiment", IEEE Trans. Appl. Superconduct. **14**, 568-571 (2005).

[40] A. Yamamoto *et al.*, "Development towards Ultra-thin Superconducting Solenoid Magnets for High Energy Particle Detectors", Nucl. Phys. B (Proc. Suppl.) **78**, 565-570 (1999); A. Yamamoto *et al.*, "A Thin Superconducting Solenoid Magnet for Particle Astrophysics", IEEE Trans. Appl. Superconduct. **12**, 438-442 (2002).

[41] H. Nishiguchi, "An Innovative Positron Spectrometer to Search for the Lepton Flavour Violating Muon Decay with a Sensitivity of 10^{-13}", Ph.D. Thesis, the University of Tokyo (2008), http://meg.web.psi.ch/docs/theses/nishiguchi_phd.pdf/.

[42] Saint-Gobain Ceramics & Plastics, http://www.detectors.saint-gobain.com/.

[43] Hamamatsu Photonics, http://www.hamamatsu.com/.

[44] S. Dussoni, "Search for Lepton Flavour Violation: The MEG experiment", Ph.D. Thesis, Università degli Studi di Genova (2006).

[45] R. Valle, "Development of the Timing Counter", Ph.D. Thesis, Università degli Studi di Genova (2006), http://meg.web.psi.ch/docs/theses/valle_phd.pdf/.

[46] M. Rossela *et al.*, “Behaviour in high magnetic fields of fine-mesh photodetectors for fast time-of-flight detectors”, Nucl. Instr. Meth. A, **567**, 200-204 (2006).

[47] K. Ozone, “Liquid Xenon Scintillation Detector for the New $\mu \to e\gamma$ Search Experiment”, Ph.D. Thesis, the University of Tokyo (2005), `http://meg.web.psi.ch/docs/theses/ozoned.pdf/`.

[48] R. Sawada, “A Liquid Xenon Scintillation Detector to Search for the Lepton Flavour Violating Muon Decay with a Sensitivity of 10^{-13}”, Ph.D. Thesis, the University of Tokyo (2008), `http://meg.web.psi.ch/docs/theses/sawada_phd.pdf/`.

[49] E. Aprile and T. Doke, “Liquid Xenon Detectors for Particle Physics and Astrophysics”, arXiv:0910.4956v1 [hep-ex] (2009).

[50] Photon Cross Sections Data Base, `http://physics.nist.gov/PhysRefData/`.

[51] S. Kubota, M. Hishida, and J. Raun, “Evidence for a triplet state of the self-trapped exciton states in liquid argon, krypton and xenon”, J. Phys. C **11**, 2645-2651 (1978); S. Kubota *et al.*, “Dynamical behavior of free electrons in the recombination process in liquid argon, krypton, and xenon”, Phys. Rev. B **20**, 3486-3496 (1979).

[52] T. Doke, “FUNDAMENTAL PROPERTIES OF LIQUID ARGON, KRYPTON AND XENON AS RADIATION DETECTOR MEDIA”, Portgal. Phys. **12**, 9-48 (1981).

[53] D. Lide ed., “Handbook of Chemistry & Physics 2004-2005”, Crc Pr. I Llc. (2004).

[54] A.C. H. Hallett, “ARGON, HELIUM AND THE RARE GASES” Vol **I**, Chap IX, Interscience Publishers, Inc., Yew Nork/London (1961).

[55] C. W. Fabjan and F. Gianotti, “Calorimetry for particle physics”, Rev. Mod. Phys. **75**, 1243-1286 (2003).

[56] J. Jortner, L. Mever, S. A. Rice, and E. G. Wilson, “Localized Excitations in Condensed Ne, Ar, Kr, and Xe”, J. Chem. Phys. **42**, 4250-4253 (1965); N. Schwenter, E.-E. Koch, and J. Jortner, “Electronic Excitations in Condensed Rare Gases”, Springer-Verlag, Berlin (1985).

[57] L. M. Barkov, A. A. Grebenuk, N. M. Ryskulov, P. Yu. Stepanov, and S. G. Zverev, “Measurement of the refractive index of liquid xenon for intrinsic scintillation light”, Nucl. Instr. Meth. A, **379**, 482-483 (1996).

[58] V. N. Solovov *et al.*,”Measurement of the refractive index and attenuation length of liquid xenon for its scintillation light”, Nucl. Instr. Meth. A, **516**, 462-474 (2004).

[59] G. M. Seidel, R. E. Lanou, and W. Yao, “Rayleigh scattering in rare-gas liquids”, Nucl. Instr. Meth. A, **489**, 189-194 (2002).

[60] T. Doke *et al.*,”Absolute Scintillation Yields in Liquid Argon and Xenon for Various Particles”, Jpn. J. Appl. Phys. **41**, 1538-1545 (2002).

[61] A. Hitachi *et al.*, “Effect of ionization density on the time dependence of luminescence from liquid argon and xenon”, Phys. Rev. B **27**, 5279-5285 (1983).

[62] A. Braem *et al.*,”Observation of the UV scintillation light from high energy electron showers in liquid xenon”, Nucl. Instr. Meth. A, **320**, 228-237 (1992).

[63] V. Y. Chapel, M. I. Lopes, R. Ferreira Marques, and A.J.P.L. Policarpo, “Purification of liquid xenon and impurity monitoring for a PET detector”, Nucl. Instr. Meth. A, **349**, 500-505 (1994).

[64] N. Ishida *et al.*, “Attenuation length measurements of scintillation light in liquid rare gases and their mixtures using an improved reflection suppresser”, Nucl. Instr. Meth. A, **384**, 380-386 (1997).

[65] T. Haruyama *et al.*, “Development of a high-power coaxial pulse tube refrigerator for a liquid xenon calorimeter”, Adv. Cryo. Eng. **49**, 1459-1466 (2004).

[66] T. Iwamoto *et al.*, “Development of a large volume zero boil-off liquid xenon storage system for muon rare decay experiment (MEG)”, Cryogenics **49**, 254-258 (2009).

[67] A. Baldini *et al.*, “A radioactive point-source lattice for calibrating and monitoring the liquid xenon calorimeter for the MEG experiment”, Nucl. Instr. Meth. A, **565**, 589-598 (2006).

[68] Y. Yoshino *et al.*, “Absorption cross section measurement of water vapor in the wavelength region 120 to 188mm”, Chem. Phys. **211** 387 (1996).

[69] S. Himi *et al.*, “Liquid and Solid Argon, and Nitrogen-doped liquid and solid Argon Scintillators”, Nucl. Inst. Meth. **203**, 153 (1982).

[70] S. Mihara *et al.*, “Development of a method for liquid xenon purification using a cryogenic centrifugal pump”, Cryogenics **46** 688-693 (2006).

[71] S. Mihara *et al.*, “Development of a liquid-xenon photon detector - toward the search for a muon rare decay mode at Paul Scherrer Institut”, Cryogenics **44** 223-228 (2004).

[72] MIDAS, Maximum Integration Data Acquisition System, `https://midas.psi.ch/`.

[73] Analog Devices, AD9218, 10 bit 3V Dual A/D converter, `http://www.analog.com/`.

[74] Xilinx, Virtex II Pro, `http://www.xilinx.com/`.

[75] SaRonix, Crystal Clock Oscillator, `http://www.pericom.com/saronix/`.

[76] S. Ritt, “The DRS chip: cheap waveform digitizing in the GHz range”, Nucl. Instr. Meth. A, **518**, 470-471 (2004).

[77] S. Ritt, “Design and Performance of the 5 GHz Waveform Digitizing Chip DRS3”, Nucli. Sci. Symmp. Conf. Rec. IEEE **4**, 2485-2488 (2007).

[78] NACRE - European Compilation of Reaction Rates for Astrophysics, `http://pntpm.ulb.ac.be/Nacre/nacre.htm/`.

[79] High Voltage Engineering Europe, http://www.highvolteng.com/.

[80] Jinan Jingyi Photonics Technology Co., http://www.jnjydkdp.com/.

[81] A. Papa, "Search for the lepton flavour violation in $\mu^+ \to e^+\gamma$. The calibration methods for the MEG experiment", Ph.D. Thesis, the University of Pisa (2009), http://meg.web.psi.ch/docs/theses/Angela.pdf/.

[82] GEANT, Detector Description and Simulation Tool version 3.21, http://wwwasd.web.cern.ch/wwwasd/geant/.

[83] Garfield, Simulation of gaseous detectors, http://garfield.web.cern.ch/garfield/.

[84] HEED, Interactions of particles with gases, http://consult.cern.ch/writeup/heed/.

[85] Magboltz, Transport of electrons in gas mixtures, http://consult.cern.ch/writeup/magboltz/.

[86] ROME, Root based Object oriented Midas Extension, https://midas.psi.ch/rome/.

[87] ROOT, An Object-Oriented Data Analysis Framework, http://root.cern.ch/.

[88] R. E. Kalman, "A New Approach to Linear Filtering and Pbrediction Problems", Transaction of the AMSE-J. Basic Engineering **D82**, 35-45, (1985).

[89] F. Rfühwirth, "APPLICATION OF KALMAN FILTERING TO TRACK AND VERTEX FITTING", Nucl. Instr. Meth. A, **262**, 444-450 (1987).

[90] J. R. Cash and A. H. Karp, "A variable order Runge-Kutta method for initial value problems with rapidly varying right-hand sides", ACM Trans. Math. Software **16**, 201-222 (1990).

[91] M. Yamashita, "Time dependence of rate-dependent photomultiplier gain and its implications", Rev. Sci. Instrum. **51(6)**, 768-775 (1980).

[92] J. R. Klein and A. Roodman, "Blind Analysis in Nuclear and Particle Physics", Annu. Rev. Nucl. Part. Sci. **55**, 141-163 (2005).

[93] MINUIT, Function Minimization and Error Analysis, http://lcgapp.cern.ch/project/cls/work-packages/mathlibs/minuit/home.html/.

[94] F. James, "Statistical Methods in Experimental Physics (2nd Edition)", World Scientific Publishing Co. Pte. Ltd. (2006).

[95] G. J. Feldman and R. D. Cousins, "Unified approach to the classical statistical analysis of small signals", Phys. Rev. D **57**, 3873-3889 (1998).