Doctoral Dissertation
博士論文

Suppression of γ-ray backgrounds for the highest sensitivity of $\mu^+ \to \mathrm{e}^+\gamma$ search in MEG II experiment

（MEG II 実験における最高感度での $\mu^+ \to \mathrm{e}^+\gamma$ 探索を可能にする背景 γ 線抑制）

A Dissertation Submitted for the Degree of Doctor of Philosophy
December 2021
令和3年12月博士（理学）申請

Department of Physics, Graduate School of Science,
The University of Tokyo
東京大学大学院理学系研究科物理学専攻

Rina Onda
恩田　理奈

Abstract

The $\mu \to e\gamma$ decay, which is one of the charged lepton flavor violating decays, can be strong evidence of physics beyond the standard model. In 2016, the MEG experiment set the most stringent upper limit for the branching ratio of the decay to be 4.2×10^{-13} (90% confidence level). The MEG II experiment aims to search for the decay with a higher sensitivity by one order of magnitude by updating the MEG detectors. The dominant background for the search is the accidental coincidence between uncorrelated positrons and γ-rays, and thus it is essential to improve detector resolutions in order to separate the signals from the backgrounds.

A liquid xenon γ-ray detector has been upgraded to measure γ-rays with better resolutions. A higher granularity and a better uniformity of the scintillation readout were realized by tiling silicon photomultipliers on the γ-ray entrance face, which results in better position and energy resolutions especially for the shallow events. The performance of the detector was measured in a series of pilot runs, and better position and energy resolution for the shallow events were proved while there still remains an uncertainty for the evaluation of the timing resolution. Further background suppression was found to be achievable by the methods developed in this work. Pileup γ-rays affect the γ-ray reconstruction especially for the energy measurement, and thus a series of algorithms to eliminate the pileup γ-rays was developed and its background suppression power was demonstrated.

A radiative decay counter (RDC) has been newly introduced in the MEG II experiment to tag the γ-ray backgrounds. It identifies a γ-ray from a radiative muon decay, which is one of the sources of the accidental γ-ray background, by detecting the positron emitted from the same muon decay. The expected performance of the detector was confirmed using data in the pilot runs. In addition, the algorithm of the $\mu \to e\gamma$ search analysis was updated to include the observables of the RDC.

The branching ratio sensitivity of the MEG II experiment was estimated based on the measured detector performance and the improvement of the analysis. The pileup elimination for the liquid xenon γ-ray detector was found to improve the sensitivity by 22–26%. A further improvement of 8% was achieved by reducing the γ-ray backgrounds with the RDC. With the updated performance and the algorithm, it was demonstrated that the sensitivity of $5.6\text{–}5.8 \times 10^{-14}$ is achievable with three years of data-taking.

Contents

Preface

Particle physics is a branch of physics that studies the elementary particles, which constitute the Universe, and the law rules them. It aims to find an ultimate theory that can explain everything in the Universe. The standard model is the most successful theory of today's particle physics though it is known that there are some phenomena the standard model cannot explain such as the origin of the dark matter. Therefore, new physics beyond the standard model is desired and physicists in the field have searched for it.

In the past, the conservation of the lepton flavor was believed because no process that violates the lepton flavor had been observed. In 1998, however, the violation of lepton flavor was discovered in the neutral lepton sector, which is called neutrino oscillation. This raised a doubt with respect to the flavor conservation of the charged lepton sector. Many models of new physics predict the charged lepton flavor violation can happen at experimentally reachable branching ratio. Therefore, various experiments have tried to find the violation.

The most stringent upper limit on the branching ratio of the $\mu \to \mathrm{e}\gamma$ decay, which is one of the charged lepton flavor violating decays, is 4.2×10^{-13} (90% confidence level), which was given by the MEG experiment [1]. The MEG II experiment is an upgraded experiment of the MEG experiment and aims to search for the decay with a higher sensitivity by one order of magnitude. To realize the target sensitivity, a suppression of the accidental backgrounds is essential since the sensitivity of MEG was limited by the number of the accidental backgrounds.

This thesis focuses on a suppression of γ-ray backgrounds in the MEG II experiment from the two aspects: improvement of the pileup elimination for a γ-ray detector and development of a method to tag the γ-ray backgrounds actively. Firstly, the physics motivation and the principle of the $\mu \to \mathrm{e}\gamma$ search are discussed in Chap. 1. Then, the MEG II experiment is explained in Chap. 2. The analysis and the performance of the γ-ray detector are explained in Chap. 3 followed by those of the detector for tagging the γ-ray backgrounds in Chap. 4. The pileup elimination algorithm and its performance for the γ-ray detector are described in Chap. 5. Chap. 6 explains the study on how to integrate information from the background detector in the physics analysis. Finally, the expected sensitivity of the MEG II experiment is discussed in Chap. 7, and the thesis is concluded in Chap. 8.

Chapter 1

Search for $\mu \to e\gamma$

In this chapter, an overview of the $\mu \to \mathrm{e}\gamma$ search is described. First of all, the physics motivation for the $\mu \to \mathrm{e}\gamma$ decay search is introduced followed by the descriptions of the principle and requirements for the search. Then, the MEG experiment is descried briefly as a preceding work and upgrade concepts to the MEG Ⅱ experiment are explained.

1.1 Physics motivation

1.1.1 The standard model

The standard model (SM) is the most successful theory in the particle physics, which describes 17 elementary particles and their interactions as summarized in Table 1.1. It is based on the gauge symmetry of $SU(3)_\mathrm{C} \times SU(2)_\mathrm{L} \times U(1)_\mathrm{Y}$, which gives three interactions of particles: the strong, weak, and electromagnetic interactions.

Table 1.1 Particles in the SM.

		Generation		
		I	Ⅱ	Ⅲ
Fermions	Quarks	u	c	t
		d	s	b
	Leptons	e	μ	τ
		ν_e	ν_μ	ν_τ

Bossons	Gauge Bossons	g
		γ
		Z
		W
	Scalar Boson	H

The interactions are propagated via gauge bosons. Gluons take a charge of the strong interaction. W and Z bosons correspond to the weak interactions. The electromagnetic interaction is propagated via photons. In addition to these bosons, there is a scaler boson called the Higgs boson, which provides the mass of particles via the Higgs mechanism.

Fermions are the particles that constitute matter. They are divided into two groups depending on how to interact with the gauge bosons: quarks and leptons. Quarks can interact via all the three interactions while leptons can only interact via the weak and electromagnetic interactions. The fermions have three generations which are distinguished with their masses and flavors.

A variety of experimental results has proved the SM succeeds at a relatively low energy region up to 100 GeV. Despite its success, the SM is not considered as the ultimate theory that can explain everything but as an approximation of a more fundamental theory at the low-energy scale due to the existence of several mysteries.

In the first place, the SM involves some theoretical problems. The three generations of the fermions are not explained; they are not theoretically required. We also have a lack of understanding for particle masses and mixing patterns, which results in the large number of underlying parameters in the SM. In addition, all of the three interactions described in the SM are not unified while the electromagnetic and the weak interactions have been unified into a single electroweak framework. Moreover, the gravitational interaction is not even mentioned in the SM. Furthermore, there is the hierarchy problem, which is related to the huge difference of energy scales between the weak scale of $\mathcal{O}(100\,\mathrm{GeV})$ and Planck scale of $\mathcal{O}(10^{19}\,\mathrm{GeV})$, where quantum effects are dominant. At the high energy scale close to the Planck scale, the quadratic divergence in the radiative corrections to the Higgs scalar mass becomes problematic, which requires a precise fine-tuning between the bare mass and the radiative corrections to keep the electroweak scale.

There are also some experimental results that cannot be explained by the SM. The discovery of the neutrino oscillation questioned the assumption that the neutrinos are massless. The observation of the dark matter also contradicts the SM.

To compensate the incompleteness of the SM, a variety of physics beyond the SM (BSM) has been proposed, and is under verification by various experiments.

1.1.2 Charged lepton flavor violation

Charged lepton flavor violation (cLFV) processes are good probes to search for BSM. In the SM, the flavor mixing is forbidden in the lepton sector while it is allowed in the quark sector via the CKM matrix. However, the observation of the neutrino oscillation, which is the lepton flavor violation in the neutral lepton sector, suggests the lepton flavor is also not conserved, and thus the violation in the charged flavor sector is supposed to happen as well though such processes have not been observed yet. Therefore, many experiments have searched for them.

Muon is a fermion in the second generation of charged leptons, which was discovered in 1937. The search for the cLFV processes in muon decay channels started in 1947, and they have been actively studied by many experiments improving their sensitivities. Fig. 1.1 shows the upper limits of the branching ratios for cLFV processes from muon decay. As shown in this plot, the three channels are often searched for in the field of cLFV search: $\mu \to \mathrm{e}\gamma$, $\mu N \to \mathrm{e}N$ and $\mu \to \mathrm{eee}$. The branching ratios of these channels depend on the model of new physics. Therefore, we can give constraints on the candidates by measuring the branching ratios of these channels.

1.1.3 $\mu \to \mathrm{e}\gamma$ decay in the SM

The $\mu \to \mathrm{e}\gamma$ decay is one of the cLFV muon decay channels. This decay is strictly prohibited in the SM, but it can have a non-zero branching ratio taking the neutrino oscillation into account

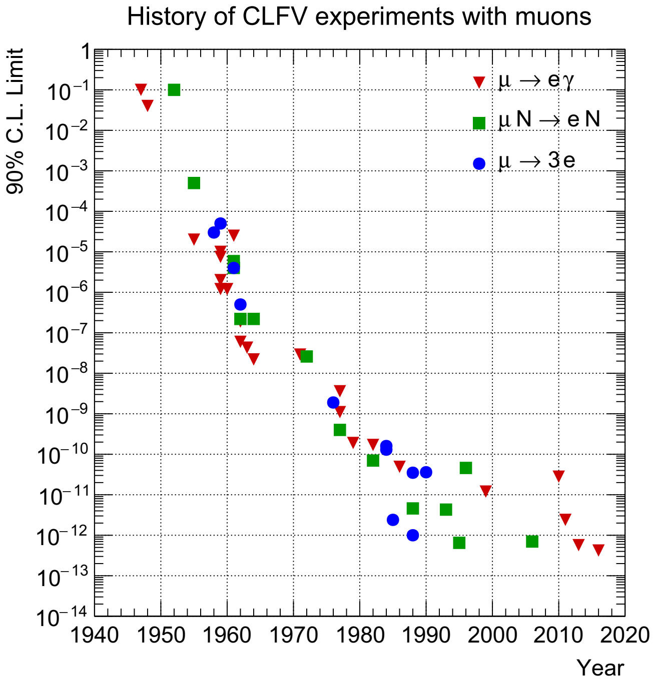


Figure 1.1 History of the cLFV search [2].

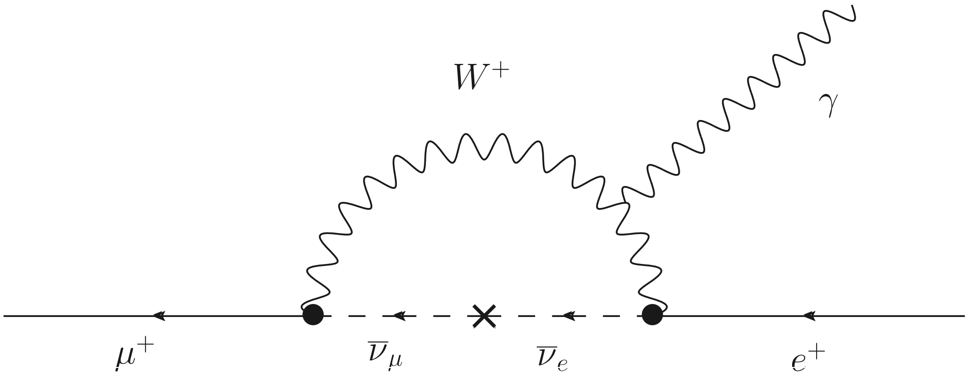


Figure 1.2 The $\mu \to \mathrm{e}\gamma$ process considering neutrino oscillation in the SM.

via the process shown in Fig. 1.2. Its branching ratio is calculated as follows [3]:

$$\mathcal{B}(\mu \to \mathrm{e}\gamma) = \frac{3\alpha}{32\pi} \left| \sum_{i=2,3} U^*_{\mu i} U_{\mathrm{e}i} \frac{\Delta m^2_{i1}}{M^2_W} \right| \sim 10^{-54},$$

where α is the fine-structure constant and U is the PMNS matrix. Since the mass difference of the neutrinos Δm is small compared to the mass of W boson M_W, the branching ratio is suppressed enough not to be observed experimentally in the near future.

1.1.4 $\mu \to \mathrm{e}\gamma$ decay in BSM models

In contrast to the SM, many BSM models predict the branching ratio of the $\mu \to \mathrm{e}\gamma$ decay is large enough to be observed experimentally.

A model-independent Lagrangian for the $\mu \to \mathrm{e}\gamma$ decay can be expressed as

$$\mathcal{L}_{\mu\to\mathrm{e}\gamma} = -\frac{G_\mathrm{F}}{\sqrt{2}}[m_\mu A_\mathrm{R}\bar{\mu}_\mathrm{R}\sigma^{\mu\nu}\mathrm{e}_\mathrm{L}F_{\mu\nu} + m_\mu A_\mathrm{L}\bar{\mu}_\mathrm{L}\sigma^{\mu\nu}\mathrm{e}_\mathrm{R}F_{\mu\nu} + (\mathrm{h.c.})],$$

where G_F is Fermi coupling constant, and A_R and A_L are coupling constants corresponding to $\mu \to e_\mathrm{R}\gamma$ and $\mu \to e_\mathrm{L}\gamma$ process, respectively. This Lagrangian results in the differential angular distribution given by

$$\frac{d\mathcal{B}(\mu \to \mathrm{e}\gamma)}{d\cos\theta_\mathrm{e}} = 192\pi^2[|A_\mathrm{R}|^2(1 - P_\mu\cos\theta_\mathrm{e}) + |A_\mathrm{L}|^2(1 + P_\mu\cos\theta_\mathrm{e})],$$

where θ_e is the angle between the muon polarization and the positron momentum in the muon rest frame, and P_μ is the magnitude of the muon polarization. The coupling constants, A_R and A_L, are found to give the dependence on the muon polarization from the equation. Since they depend on the model, it is possible to restrict models by measuring the distribution of the positron emission angle as well as by measuring the magnitude of the decay without relying on the angle distribution.

The Supersymmetry (SUSY) model is one of the promising candidates for BSM [4]. In this model, the symmetry under the super-transformation between bosons and fermions is introduced. This symmetry requires a partner for each SM particle, and the divergence in the radiative corrections of the Higgs mass is cancelled out thanks to these SUSY particles, which gives a natural solution for the hierarchy problem. In addition, the lightest SUSY particle can be a candidate for the dark matter since it is electrically neutral and weakly interacts with the matter.

In the SUSY scenario, a particle and its super-partner should be degenerated and have the same mass if the symmetry is exactly fulfilled. However, no SUSY particle with the same mass has ever been observed, which indicates the symmetry should be broken. In this case, the masses of the SUSY particles are given by a term which breaks the super-symmetry so that the particle pairs are disentangled. The LFV would be originated from the misalignment between particle and super-particle mass eigenstate, and can be induced by the presence of nonzero off-diagonal matrix elements in the mass matrix of slepton, which is super-partner of lepton, in the basis where the lepton mass matrix is diagonalized. However, some experimental observations suggest there are some mechanisms which suppress the matrix to be almost diagonal.

One of the possible mechanisms to explain the suppression is super gravity model (SUGRA). In this model, all sfermions have the same mass and the mass matrix is diagonal, which restricts the LFV completely at the Planck scale. This does not, however, necessarily imply that the LFV should not exist in this model. In fact, if there is an interaction which breaks lepton flavor conservation between the Plank and the electroweak scales, the LFV could be induced in the slepton mass matrices through radiative corrections. Hereafter, two kinds of sources which generate such effects are introduced.

SUSY seesaw

The seesaw mechanism provides a natural explanation of the tiny neutrino mass compared to other fermions by introducing the Majorana mass term of right-handed neutrinos. In this theory, left and right-handed neutrino masses are given by $M_{\mathrm{D}}^2/M_{\mathrm{R}}$ and M_{R}, where M_{D} is the Dirac mass and M_{R} is the Majorana mass of the right-handed neutrino. This formula suggests the left-handed neutrino mass is tiny when the right-handed neutrino mass is huge.

The SUSY seesaw model introduces the seesaw mechanism to the SUSY model. When the Majorana neutrinos are included, a new Yukawa coupling matrix for the neutrinos y_ν is additionally included in the lepton sector. Since there are both of Yukawa coupling matrices for the neutrino sector and for the charged lepton sector, the lepton flavor would no longer be conserved separately for each generation, which enhances the LFV decay. The Yukawa coupling for the neutrinos couples the right-handed neutrino multiplets with the left-handed lepton doublets. This leads to the off-diagonal term proportional to $(y_\nu)^*_{ki}(y_\nu)_{kj}$ in the mass matrix of the left-handed slepton through radiative corrections as

$$(m^2_{\tilde{l}L})_{ij} \simeq -\frac{1}{8\pi^2}(y_\nu)^*_{ki}(y_\nu)_{kj}m_0^2(3+|A_0|^2)\ln\frac{M_P}{M_R}, \tag{1.1}$$

where m_0 is the universal scaler mass, A_0 is the trilinear coupling, and M_P is the reduced Planck mass [5]. This mechanism is not efficient to generate off-diagonal terms in the right-handed slepton mass matrix, and thus only the left-handed slepton contributes to the LFV. If we assume that the neutrino mixing mostly originates from the neutrino Yukawa coupling constants, the branching ratios of $\mu \to \mathrm{e}\gamma$ and $\tau \to \mu\gamma$ decays can be evaluated by using the neutrino mixing parameters. Fig. 1.3 shows the correlation between those branching ratios in an example of this model. The neutrino mixing angle $\theta_{13} \sim 9^\circ$ [6] suggests a higher branching ratio for the $\mu \to \mathrm{e}\gamma$ decay.

SUSY GUT

The $SU(5)$ SUSY GUT is the model in which the gauge groups of $SU(3)_{\mathrm{C}}$, $SU(2)_{\mathrm{L}}$ and $U(1)_{\mathrm{Y}}$ are unified into a single $SU(5)$ group at the grand unification energy scale thanks to the SUSY particles. The Yukawa coupling constant corresponding to the top-quark mass is large enough to give the sizable soft SUSY-breaking mass terms. Since the right-handed sleptons are hosted in the **10** representation as well as the top quark, the off-diagonal elements for them arise as

$$(m^2_{\tilde{e}R})_{ij} \simeq -\frac{3}{8\pi^2}(V_R)_{i3}(V_R)^*_{j3}|y_u^{33}|^2m_0^2(3+|A_0|^2)\ln\frac{M_P}{M_G}, \tag{1.2}$$

where V_R is the mixing matrix of the right-handed slepton, y_u is the Yukawa coupling constant for up-type quarks, and M_G represents the GUT energy scale. Fig. 1.4 shows the processes induced by this off-diagonal elements. A higher branching ratio is expected with larger $\tan\beta$, where $\beta = \langle h_2\rangle / \langle h_1\rangle$ is the ratio of the expectation values of two Higgs fields: h_1 for down-type quarks and leptons, and h_2 for up-type quarks as shown in Fig. 1.5.

The $SO(10)$ GUT models typically relate the up-quark Yukawa couplings with the Dirac neutrino Yukawa couplings, and provide a natural setting for the seesaw mechanism with 16-

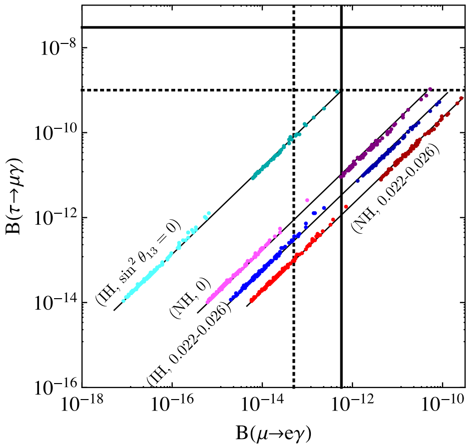


Figure 1.3 Correlation between the branching ratios of $\mu \to e\gamma$ and $\tau \to \mu\gamma$ decays in the MSSM (the minimum SUSY extension of the standard model) with seesaw mechanism [7]. The SUSY parameters of $\tan\beta = 30$ and $\mu > 0$ are taken, where β is the ratio of the expectation values of two Higgs fields and μ is the higgsino mass parameter. The results of the normal and inverted hierarchies (NH and IH) within $\sin^2\theta_{13} = 0.022$–$0.026$ and with $\sin^2\theta_{13} = 0$ are shown as indicated. The vertical and horizontal solid lines show the branching ratio upper bounds of each decay, and the dotted ones are expected sensitivities at SuperKEKB/Belle II and MEG II.

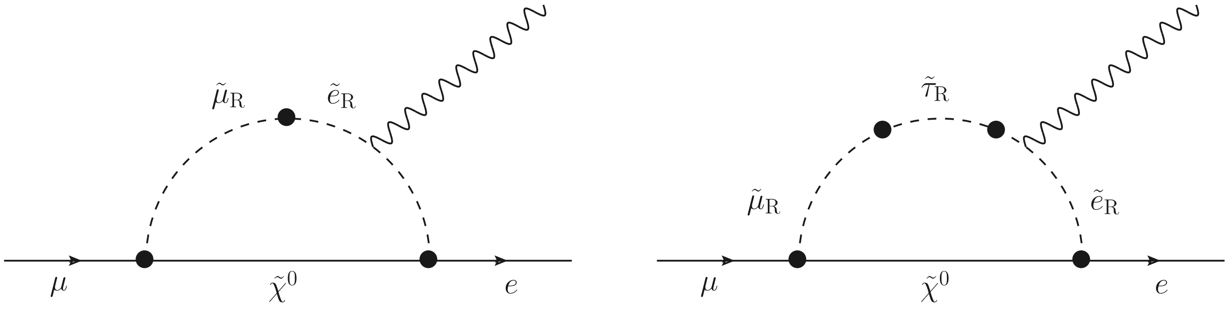


Figure 1.4 $\mu \to e\gamma$ in the $SU(5)$ SUSY GUT [9]

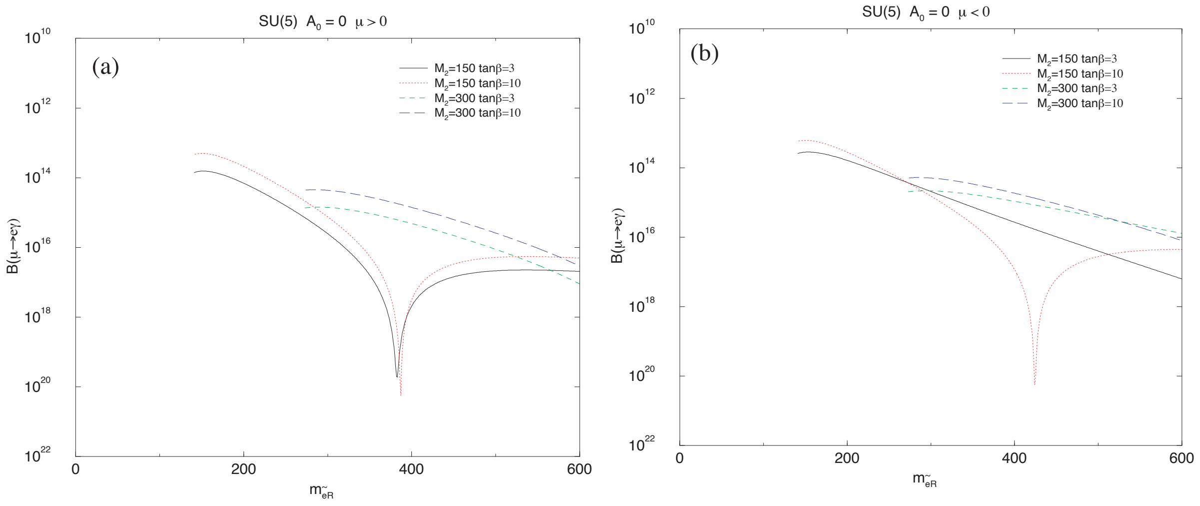
dimensional representation. In the case of the $SO(10)$ SUSY GUT, both left- and right-handed sleptons, i.e. $(m^2_{\tilde{e}L})_{ij}$ and $(m^2_{\tilde{e}R})_{ij}$, can contribute to the LFV. Fig. 1.6 shows the branching ratio of the $\mu \to e\gamma$ decay as a function of unified gaugino mass $M_{1/2}$ and the allowed parameter space [8].

Other models

In addition to the major models mentioned above, many theories predict observable rates of the $\mu \to e\gamma$ decay within the reach of MEG II such as extra-dimensions [11] [12] [13], left-right symmetry [14] [15] [16] [17], and little Higgs [18] [19] [20] [21].

Summary

In conclusion, the search for the $\mu \to e\gamma$ process is a powerful probe for BSM. Pushing down the sensitivity of $\mu \to e\gamma$ search to $\mathcal{O}(10^{-14})$ level is meaningful to validate many theories which



(a) (b)

Figure 1.5 Branching ratio of the $\mu \to e\gamma$ decay for (a) positive and (b) negative sign of the higgsino mass parameter μ as a function of right-handed selepton mass for four different sets of SUSY input parameters of the $SU(2)$ gaugino mass M_2 and $\tan\beta$ in the $SU(5)$ SUSY GUT [9]. The trilinear scalar coupling constant and the top-quark mass are fixed to be $A_0 = 0$ and $m_t = 175\,\mathrm{GeV}$, respectively.

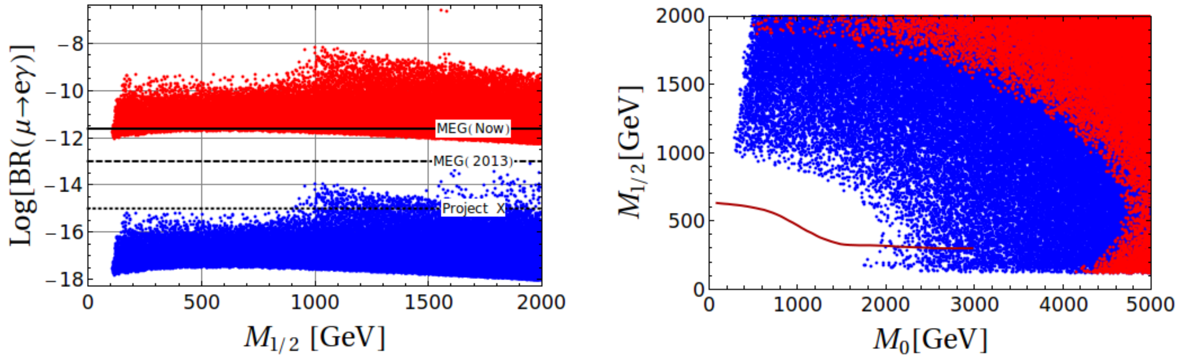


Figure 1.6 Branching ratio of the $\mu \to e\gamma$ decay obtained by scanning the mSUGRA (minimal SUperGRAvity) parameters for fixed $\tan\beta = 10$ and U_{e3} (left) and the allowed space in the M_0–$M_{1/2}$ plane which satisfy the MEG result in 2012 (right) [8]. The red and blue points correspond to the maximal-mixing scenario (PMNS) and the minimal-mixing scenario (CKM) case, respectively. The region below the red line is excluded by the LHC searches in 2012 [10].

predict the branching ratio of the decay at that level.

1.1.5 Relation with other experimental searches

As discussed in the previous section, the $\mu \to e\gamma$ decay is one of the interesting channels to search for BSM. There are, however, other muon decay channels which are sensitive to BSM as well as τ decay channels in addition to another approach of energy frontier experiments. In this section, these experimental searches are described with their relations to the $\mu \to \mathrm{e}\gamma$ search.

1.1.5.1 cLFV search with other μ processes

Other cLFV channels of muon, which are complementary to $\mu \to \mathrm{e}\gamma$ decay, such as $\mu N \to \mathrm{e}N$, $\mu \to \mathrm{eee}$, also have been searched for actively.

The effective Lagrangian[*1] for the cLFV processes via muon can be written as

$$\mathcal{L} = \frac{m_\mu}{(\kappa+1)\Lambda^2}\bar{\mu}_\mathrm{R}\sigma_{\mu\nu}\mathrm{e}_\mathrm{L}F^{\mu\nu} + \frac{\kappa}{(\kappa+1)\Lambda^2}\bar{\mu}_\mathrm{L}\gamma_\mu\mathrm{e}_\mathrm{L}(\bar{f}_\mathrm{L}\gamma^\mu f_\mathrm{L}), \tag{1.3}$$

where f is the fermionic fields [22]. The coefficients of the two types of operators are parametrized by two constants: the mass dimension-one Λ parameter, which represents the effective mass scale of the new degree of freedom, and the dimensionless parameter κ, which governs the relative size of the operators. The first term is the magnetic moment-type operator, which directly mediates $\mu \to \mathrm{e}\gamma$, $\mu \to \mathrm{eee}$ and $\mu N \to \mathrm{e}N$ conversion in nuclei at order α. The second term also mediates $\mu N \to \mathrm{e}N$ at leading order and $\mu \to \mathrm{eee}$ at tree level in the cases that the fermion is quark and electron, respectively, and $\mu \to \mathrm{e}\gamma$ is available at the one-loop level for the both cases. For $\kappa \ll 1$, the former dipole-type operator dominates cLFV phenomena while the latter four-fermion operators are dominant for $\kappa \gg 1$. Fig. 1.7 shows the sensitivity to Λ as a function of κ for the processes, which illustrates the sensitivities of them depend on the magnitude of κ.

Although the coefficients of the two operators depend on the model, many models including the SUSY-seesaw/GUT models mentioned above predict the former operator is dominant. In such a case, the branching ratios of the three muon processes have the following relations [9] [23]:

$$\begin{aligned}
\frac{\mathcal{B}(\mu \to \mathrm{eee})}{\mathcal{B}(\mu^+ \to \mathrm{e}^+\gamma)} &\approx 6 \times 10^{-3}, \\
\frac{\mathcal{B}(\mu^- N \to \mathrm{e}^- N)}{\mathcal{B}(\mu^+ \to \mathrm{e}^+\gamma)} &\approx 2.6 \times 10^{-3}\ (\text{for N} = \text{Al}).
\end{aligned}$$

This means a search for $\mu^+ \to \mathrm{e}^+\gamma$ decay with a sensitivity of $\sim 6 \times 10^{-14}$ is competitive to other experiments planned in the coming years, which are introduced below.

On the other hand, searches for different processes are important to restrict possible physics models. In the case of a positive signal, the amount of information from one experiment is limited; a positive signal in a $\mu N \to \mathrm{e}N$ conversion experiment cannot measure either Λ or κ but only a function of the two, for example. Therefore, a combination with information from other observables including those involving τ decays, studies of the electromagnetic properties of charged leptons, e.g. g−2, precision studies of neutrino processes, and searches for new physics at the energy frontier, is necessary.

$\mu^- N \to \mathrm{e}^- N$ conversion

In the $\mu N \to \mathrm{e}N$ conversion experiments, a negative muon is stopped in a thin target to form a muonic atom. In the SM, the muon decays in orbit as $\mu^- \to \mathrm{e}^-\bar{\nu}_\mathrm{e}\nu_\mu$ or is captured by a

---

[*1] The most general effective Lagrangian includes several other terms [9], but the subsets are sufficient for the discussion here following the original notation in [22].

nucleus of the atom as $\mu^- + N(A, Z) \to \nu_\mu + N(A, Z-1)$, where $N(A, Z)$ represents a nucleus of the atom whose mass number is A and atomic number is Z. The $\mu^- N \to \mathrm{e}^- N$ conversion is a conversion process from a muon to an electron expressed as

$$\mu^- + N(A, Z) \to \mathrm{e}^- + N(A, Z).$$

The signal is characterized by a single electron with a monochromatic energy around 105 MeV, which slightly depends on the target nucleus due to the difference of the binding energy of muon. The major backgrounds for the observation are the muon decay-in-orbit and the beam contamination such as pions. Therefore, use of a pulsed muon beam can reduce the backgrounds by searching for the conversion electrons in the bunch intervals since muonic atoms have lifetimes ranging from hundreds of nanoseconds up to the free muon lifetime at low Z.

The current experimental upper limit of 7×10^{-13} (90% C.L.) was given by the SINDRUM II experiment at PSI [24]. Further search for the $\mu^- N \to \mathrm{e}^- N$ conversion is planned by several collaborations such as Mu2e and COMET. The Mu2e experiment in Fermilab will finish construction in 2023 followed by the physics data-taking for three years to reach the sensitivity of 8×10^{-17} (90% C.L.) [25] [26]. The construction of the COMET experiment at J-PARC is also in progress. It aims at the sensitivity below 7×10^{-15} (90% C.L.) with 150 day data-taking at phase I in 2023 followed by one year data-taking to reach 6×10^{-17} (90% C.L.) at phase II, which is expected to start in 2026–2027 [27].

$\mu \to \mathrm{eee}$ **decay**

The $\mu \to \mathrm{eee}$ decay is another LFV decay process. For $\kappa \ll 1$, the $\mu \to \mathrm{e}\gamma$ search has an advantage while it is opposite for $\kappa \gg 1$.

The signal is characterized by three coincident electrons from the same muon. The background in the search is an accidental background which consists of electrons from different muon decays and radiative decays with internal conversion with a small energy fraction carried away by the neutrinos. Therefore, the good momentum, vertex and timing resolutions are required in addition to a high rate capability to cope with the high intensity muon beam.

An experimental search for the $\mu \to \mathrm{eee}$ decay is planned by Mu3e collaboration at PSI [28]. It is expected to start the physics data-taking for 300 days aiming at a single event sensitivity of 2×10^{-15} at the phase I [29], and 7×10^{-17} at the phase II with 260 day data-taking.

Muon anomalous magnetic moment

The magnetic moment by muon spin is written by

$$\mathbf{M}_\mu = \frac{g \cdot e}{2m_\mu}\mathbf{s},$$

where $\mathbf{s}$ is spin angular momentum and g is a factor to connect magnitudes of magnetic moment and spin. The factor g is exactly 2 when only tree-level process is considered, but it deviates from 2 with higher order processes. Some experimental results deviate from the SM prediction at the 3σ level.

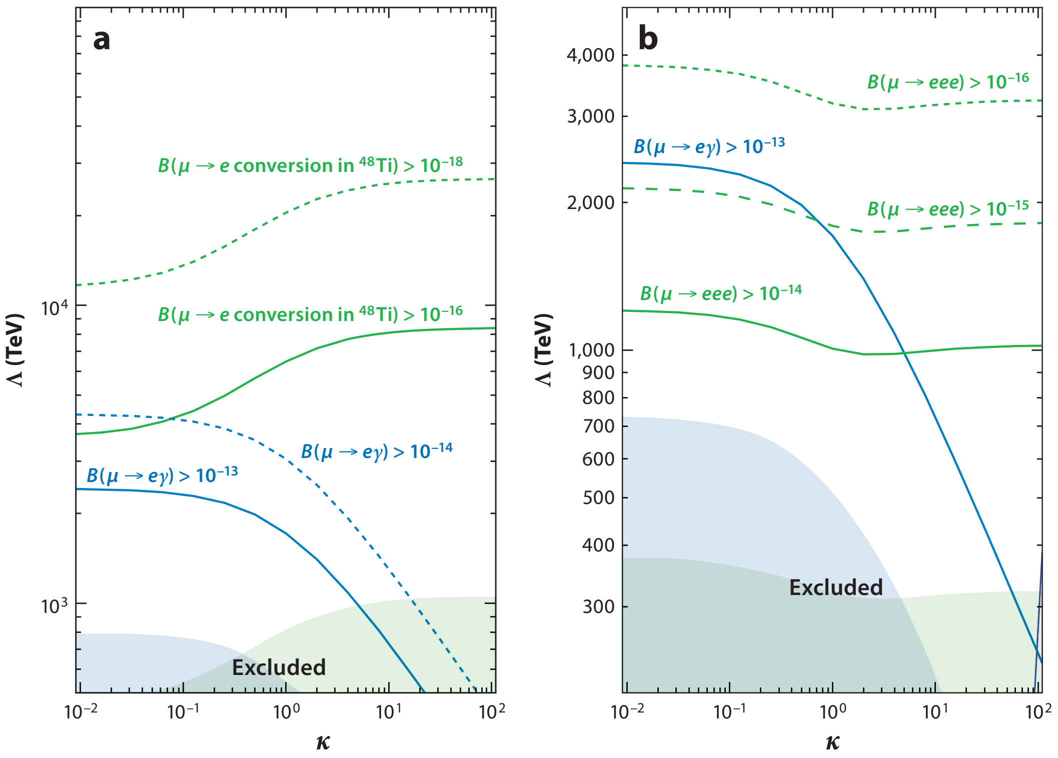


Figure 1.7 Sensitivity of (a) a $\mu N \to \mathrm{e}N$ conversion in ^{48}Ti experiment and (b) a $\mu \to \mathrm{eee}$ experiment with that of a $\mu \to \mathrm{e}\gamma$ search to the new physics scale Λ as a function of κ [22].

The deviation can be explained by introducing a contribution of new particles. The effective Lagrangian of the new physics contributions is given by [22]

$$\mathcal{L}_{g-2} \supset \frac{m_\mu}{\Lambda^2} \bar{\mu}_\mathrm{R} \sigma_{\mu\nu} \mu_\mathrm{L} F^{\mu\nu}.$$

Even though this Lagrangian is flavor conserving, it is similar to Eq. (1.3) in the limit of $\kappa \ll 1$. In a SUSY model discussed in [30], the branching ratio of $\mu \to \mathrm{e}\gamma$ and the discrepancy of g is related as

$$\mathcal{B}(\mu \to \mathrm{e}\gamma) \approx 10^{-4} (\frac{\delta a_\mu}{200 \times 10^{-11}})^2 |\delta_{LL}^{12}|^2,$$

where δa_μ is the difference between the theoretical value and the experimental observation of $a_\mu = (g-2)/2$, and $|\delta_{LL}^{12}|$ is a factor coming from cLFV coupling. Fig. 1.8 illustrates the relation, which clarifies the complementarity of the two observations.

The first result of E989 at Fermilab was published in 2021, which reports a greater value by 3.3σ than the SM prediction [31]. Combining this to the result of BNL E821 [32], the new experimental average became 4.2σ.

1.1.5.2 cLFV search with τ decay

The cLFV τ decay channels such as $\tau \to \ell\gamma$ and $\tau \to \ell\ell\ell$ ($\ell = \mathrm{e}$ or μ) have been searched for as well. Many BSM models predict high branching ratios enough to be detected experimentally

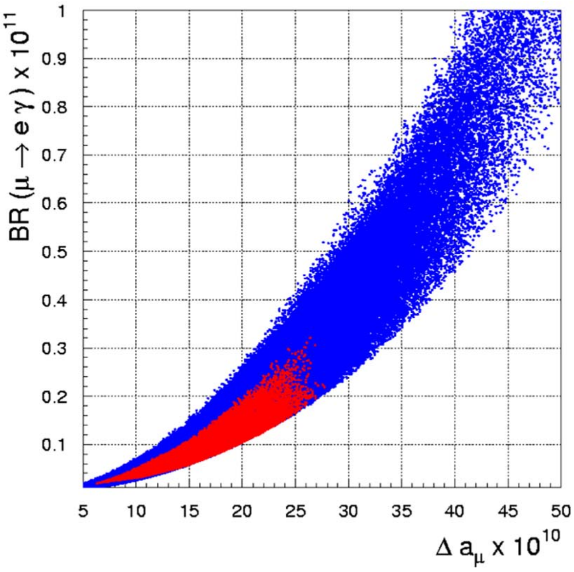
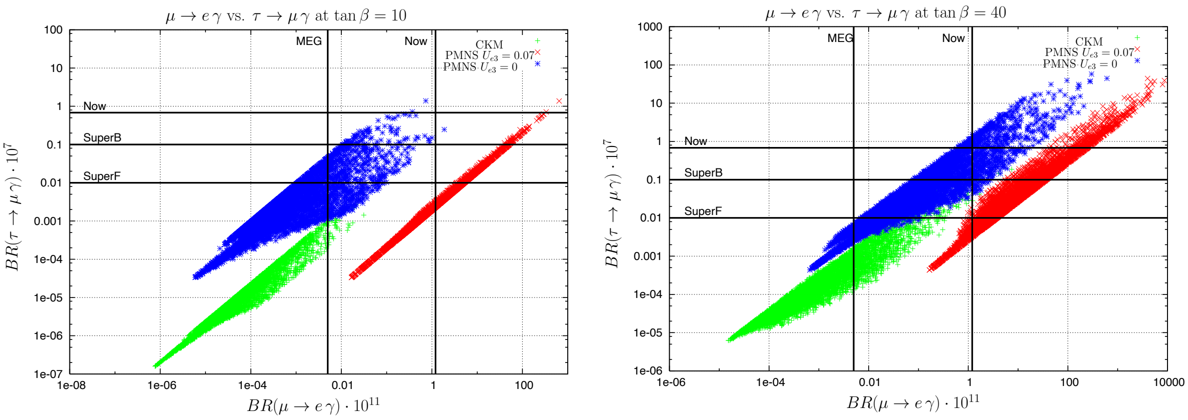


Figure 1.8 Relation between the branching ratio of $\mu \to e\gamma$ decay and deviation of muon anomalous magnetic moment assuming $|\delta^{12}_{LL}| = 10^{-4}$ obtained by scanning parameter spaces [30].



(a) (b)

Figure 1.9 Correlation between $\mu \to e\gamma$ and $\tau \to \mu\gamma$ in SUSY-GUT($SO(10)$) for (a) $\tan\beta = 10$ and (b) $\tan\beta = 40$ [35]. The plots are done by scanning the LHC accessible parameter space. The inner (red) areas correspond to points within the funnel region which satisfy the B-physics constraints.

around $\mathcal{O}(10^{-9})$–$\mathcal{O}(10^{-7})$ [33]. The τ decay channels are theoretically associated with the μ decay channels, and their branching ratios have correlations in many scenarios as shown in Fig. 1.3 and Fig. 1.9. In some parameter spaces, the τ channels have a higher sensitivity to BSM.

The Belle II experiment at SuperKEKB [34] plans to search for the τ decays with a sensitivity goal of $\mathcal{O}(10^{-9})$ (Fig. 1.10).

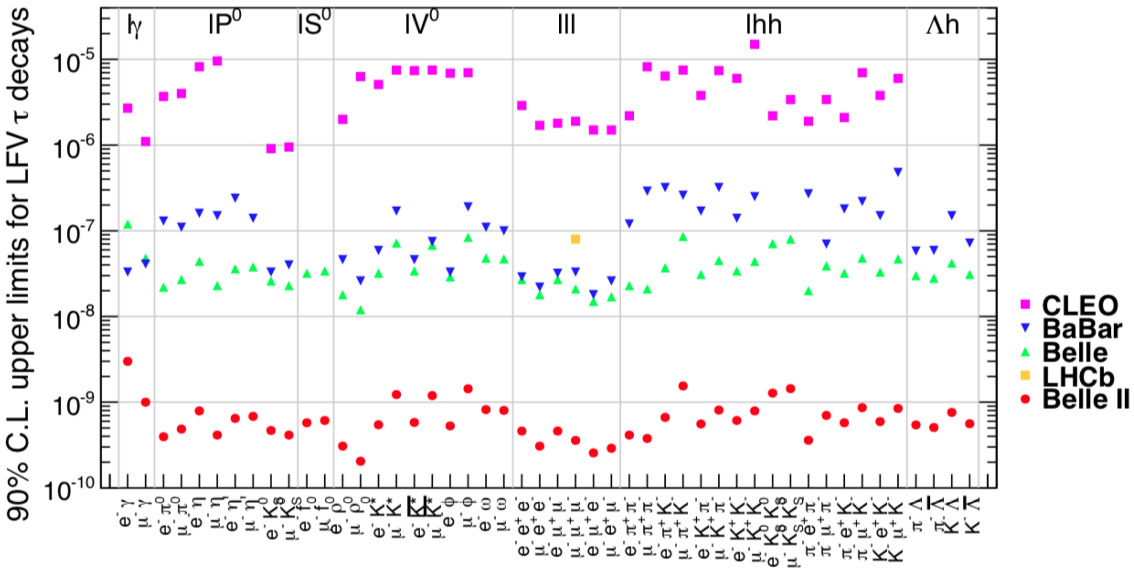


Figure 1.10 Upper limits for the branching ratio of LFV τ decays with limits expected to be imposed by Belle II assuming an integrated luminosity of 50 ab^{-1} [36].

1.1.5.3 Direct search in energy frontier experiments

The direct search for the SUSY particles is one of the attractive motivations for the energy frontier experiments such as the large hadron collider (LHC) experiment. However, its reachable energy region is limited by the beam energy. Fig. 1.11 shows the mass reach of the ATLAS experiment [37] for SUSY particles. It reaches 1–2 TeV scale. On the other hand, the MEG II experiment can indirectly access SUSY particles' masses around 10 TeV by searching for the $\mu \to e\gamma$ decay at the sensitivity of $\mathcal{O}(10^{-14})$ as shown in Fig. 1.12 [38]. In addition, cLFV searches are sensitive to particles which do not strongly interact such as sleptons, which are not strongly constrained by the LHC results. Therefore, the two approaches are complementary to explore new physics.

1.2 Principle of $\mu \to e\gamma$ search

In a $\mu \to e\gamma$ search, the decay mode is identified by detecting the decay products. Usually, positive muons are used not to form muonic atoms; negative muons can be captured by nucleus in target material. The decay from the muonic atom is no longer a two-body decay, which makes the identification difficult.

1.2.1 Signal

The signal $\mu^+ \to e^+\gamma$ decay is a two-body decay. In the rest frame of the muon, the muon emits a time coincident pair of a positron and a γ-ray with the energy of a half of the muon mass m_μ to the back-to-back direction. Therefore, the decay from stopped muons can be identified

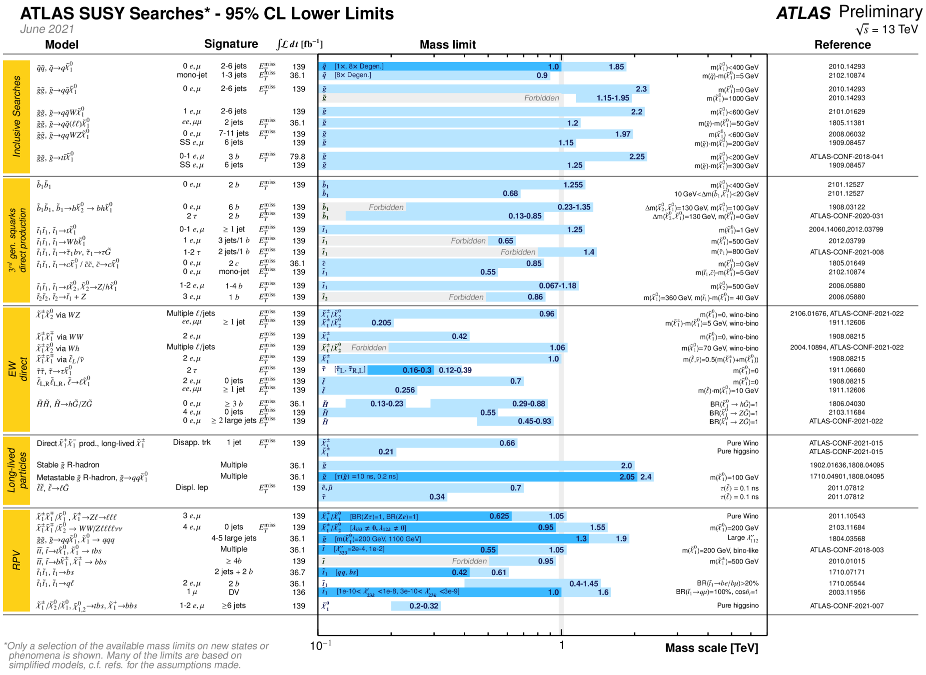


Figure 1.11 Mass reach of the ATLAS searches for SUSY in June 2021 [39]. A representative selection of the available search results is shown.

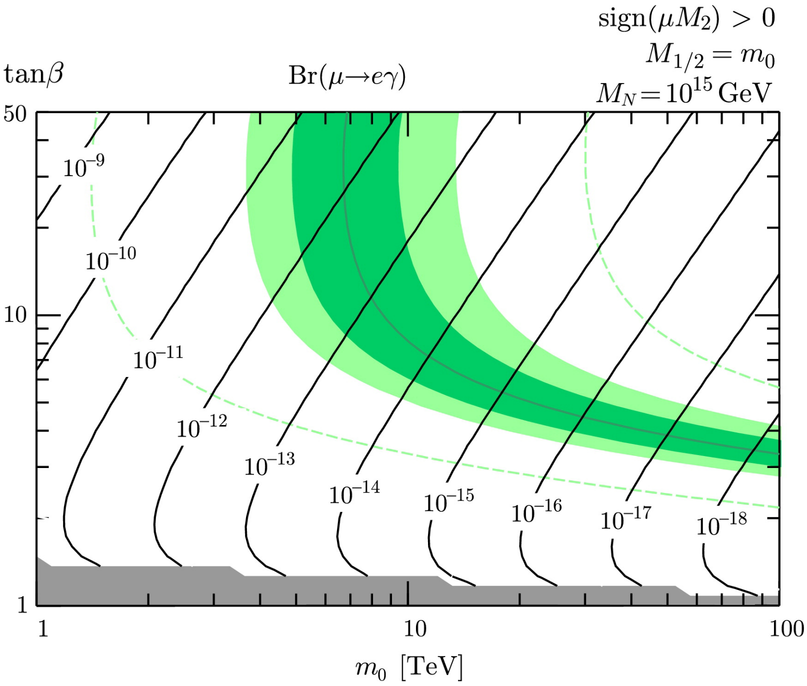


Figure 1.12 Branching ratio of $\mu \to e\gamma$ decay as a function of the universal scaler mass m_0 and $\tan\beta$ in the mSUGRA model [38]. Dark (light) green region satisfies $125\,\text{GeV} < m_h < 127\,\text{GeV}$ ($124\,\text{GeV} < m_h < 128\,\text{GeV}$) and dashed two lines show $120\,\text{GeV} < m_g < 130\,\text{GeV}$. The gray region is excluded by the non-perturbativity of the top Yukawa coupling.

by the following conditions:

- $E_{\mathrm{e}^+} = E_\gamma = m_\mu/2 \simeq 52.8\,\mathrm{MeV}$,
- $t_{\mathrm{e}^+\gamma} = 0$,
- $\Theta_{\mathrm{e}^+\gamma} = 180^\circ$,

where E_{e^+} and E_γ are the energies of the positron and the γ-ray. $t_{\mathrm{e}^+\gamma}$ and $\Theta_{\mathrm{e}^+\gamma}$ are the time difference and the opening angle between the positron and the γ-ray, respectively.

The expected number of signals is the product of the branching ratio of the signal decay $\mathcal{B}$ and the inverse of the single event sensitivity k, and thus it can be written as follows:

$$\begin{aligned} N_{\mathrm{sig}} &= \mathcal{B} \times k, \\ k &:= R_\mu \times T \times \Omega \times \epsilon_{\mathrm{e}^+} \times \epsilon_\gamma \times \epsilon_{\mathrm{cut}}, \end{aligned} \tag{1.4}$$

where R_μ is the muon beam intensity, T is the total data-acquisition live time, Ω is the geometrical acceptance of the detector, ϵ_{e^+} and ϵ_γ are the detection efficiencies of the positron and γ-ray respectively and ϵ_{cut} is the efficiency of the event selection. The increase of each of the factors except for the branching ratio given by the nature results in a higher sensitivity.

1.2.2 Backgrounds

In the SM, the μ^+ decays emitting a positron, neutrinos and a γ-ray as shown in Table 1.2. The products from the SM decay modes can mimic the signal decay by satisfying the signal event kinematics, and those events can be background events.

Table 1.2 Decay modes of muon [40].

Decay mode	Fraction
$\mu \to \mathrm{e}\nu\overline{\nu}$	~ 1
$\mu \to \mathrm{e}\nu\overline{\nu}\gamma$	$(6.0 \pm 0.5) \times 10^{-8}$ $(E_{\mathrm{e}} > 45\,\mathrm{MeV},\ E_\gamma > 40\,\mathrm{MeV})$
$\mu \to \mathrm{e}\nu\overline{\nu}\mathrm{e}^+\mathrm{e}^-$	$(3.4 \pm 0.4) \times 10^{-5}$

The background events in the $\mu^+ \to \mathrm{e}^+\gamma$ search can be classified into two groups: physics background and accidental background. The former one is from a radiative muon decay, which emits a positron and a γ-ray at the same time, and the latter is given by the accidental coincidence of a positron and a γ-ray from different muons. In principle, these background events can be distinguished from the true signal event since they can have slight differences from the signal event criteria. In reality, however, they are not necessarily distinguishable due to the finite detector resolutions.

To discuss the expected number of background events, let's introduce new quantities: the normalized positron and γ-ray energies to the signal energy, $x = 2E_{\mathrm{e}^+}/m_\mu$ and $y = 2E_\gamma/m_\mu$, and the opening angle, $z = \pi - \Theta_{\mathrm{e}^+\gamma}$. Using these quantities, the interesting region can be expressed as

$$\begin{cases} 1-\delta x \leq x \leq 1+\delta x, \\ 1-\delta y \leq y \leq 1+\delta y, \\ 0 \leq z \leq \delta z, \\ -\delta t_{\mathrm{e}^+\gamma} \leq t_{\mathrm{e}^+\gamma} \leq \delta t_{\mathrm{e}^+\gamma}, \end{cases}$$

where the variables with δ denote the ranges of the signal regions for the corresponding observables. The partial branching ratios in this region give the effective number of the background events.

Physics background

The radiative muon decay (RMD), $\mu^+ \to \mathrm{e}^+\nu_{\mathrm{e}}\overline{\nu}_\mu\gamma$, becomes a physics background. The positron and the γ-ray from this decay are emitted simultaneously. In addition, when the two neutrinos carry little momentum, the energies of the positron and the γ-ray are close to those of signals and they are emitted almost back-to-back: $x \approx 1$, $y \approx 1$ and $z \approx 0$.

The differential branching ratio $d\mathcal{B}$ of the RMD can be written within the framework of the V–A interaction as

$$\begin{aligned} d\mathcal{B}(\mu^+ \to \mathrm{e}^+\nu_{\mathrm{e}}\overline{\nu}_\mu\gamma) =& \frac{\alpha}{64\pi^3}\beta dx \frac{dy}{y} d\Omega_{\mathrm{e}^+} d\Omega_\gamma \\ & \times [F(x,y,d) - \beta \vec{P}_{\mu^+} \cdot \hat{p}_{\mathrm{e}^+} G(x,y,d) - \vec{P}_{\mu^+} \cdot \hat{p}_\gamma H(x,y,d)], \end{aligned} \tag{1.5}$$

where $d\Omega_{\mathrm{e}^+}$ and $d\Omega_\gamma$ are solid angles of the emitted positron and γ-ray whose energy is in the interval of dx and dy, respectively [9]. $\vec{P}_{\mu^+}$ is the muon polarization vector. $\hat{p}_{\mathrm{e}^+}$ and $\hat{p}_\gamma$ are unit vectors of positron and γ-ray momenta, $\vec{p}_{\mathrm{e}^+}$ and $\vec{p}_\gamma$, in the muon rest frame respectively. β is defined as $\beta \equiv |\vec{p}_{\mathrm{e}^+}|/E_{\mathrm{e}^+}$. d is given by $d \equiv 1 - \beta\hat{p}_{\mathrm{e}^+} \cdot \hat{p}_\gamma$. The functions, $F(x,y,d)$, $G(x,y,d)$ and $H(x,y,d)$, are given in Appendix A of [9]. From the four-body kinematics, the ranges of x and y are limited as

$$\begin{aligned} 2\sqrt{r} < x < 1+r \ (0 < y \leq 1-\sqrt{r}), \\ (1-y) + r/(1-y) \leq x \leq 1+r \ (1-\sqrt{r} \leq y \leq 1-r), \end{aligned} \tag{1.6}$$

where $r = (m_{\mathrm{e}^+}/m_{\mu^+})^2$.

In an approximation of the limit of $x \approx 1$, $y \approx 1$ and $z \approx 0$, the differential branching ratio is expressed as

$$d\mathcal{B}(\mu^+ \to \mathrm{e}^+\nu_{\mathrm{e}}\overline{\nu}_\mu\gamma) = \frac{\alpha}{16\pi}[J_1(1-P_{\mu^+}\cos\theta_{\mathrm{e}^+}) + J_2(1+P_{\mu^+}\cos\theta_{\mathrm{e}^+})]d(\cos\theta_{\mathrm{e}^+}).$$

In the condition of the experiment in which the angular resolution meets $\delta z \leq 2\sqrt{\delta x \delta y}$, J_1 and J_2 are given by

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

Fig. 1.13 shows a fraction of the $\mu^+ \to \mathrm{e}^+\nu_{\mathrm{e}}\overline{\nu}_\mu\gamma$ decay for the given δx and δy values assuming a polarization of -0.86 measured in the MEG experiment [41] and $\delta z \approx 20\,\mathrm{mrad}$. This figure

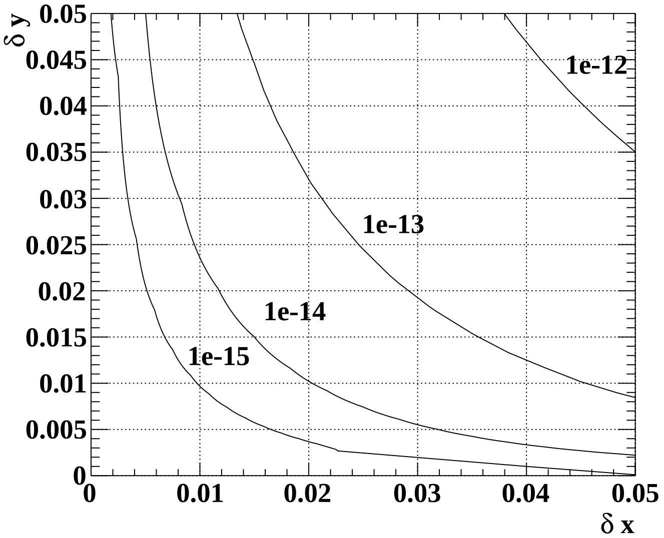


Figure 1.13 Effective branching ratio of the physics background from the $\mu^+ \to \mathrm{e}^+\nu_\mathrm{e}\overline{\nu}_\mu\gamma$ decay as a function of the positron energy resolution (δx) and γ-ray energy resolution (δy).

indicates that both δx and δy on the order of 0.01 are needed to achieve a sensitivity limit at the level of 10^{-15}.

The expected number of physics background events can be written by

$$N_\mathrm{RMD} = \mathcal{B}_\mathrm{RMD} \times R_\mu \times T \times \Omega \times \epsilon_{\mathrm{e}^+,\mathrm{RMD}} \times \epsilon_{\gamma,\mathrm{RMD}}, \tag{1.7}$$

where $\mathcal{B}_\mathrm{RMD}$ is the branching ratio of the background, which can be obtained by integrating the differential branching ratio, and $\epsilon_{\mathrm{e}^+,\mathrm{RMD}}$ and $\epsilon_{\gamma,\mathrm{RMD}}$ are the efficiencies to detect the RMD positrons and γ-rays, respectively.

Accidental background

The accidental background derives from a positron and a γ-ray originated from independent muon decays. If a pair of the positron and the γ-ray with energies near those of the signals is emitted simultaneously in the back-to-back direction, it can be regarded as a signal event. The source of such positrons is a Michel decay ($\mu^+ \to \mathrm{e}^+\nu_\mathrm{e}\overline{\nu}_\mu$). On the other hand, the γ-ray background comes from an RMD and an annihilation in flight of a positron (AIF).

The effective branching ratio of the accidental background $\mathcal{B}_\mathrm{acc}$ can be expressed as

$$\mathcal{B}_\mathrm{acc} = R_\mu \times f_{\mathrm{e}^+} \times f_\gamma \times \frac{(\delta z)^2}{4} \times (2\delta t_{\mathrm{e}^+\gamma}), \tag{1.8}$$

where f_{e^+} is the fraction of positrons which enter the signal energy region and f_γ is that of γ-rays.

f_{e^+} can be given by the Michel positron energy spectrum. When the positron polarization is neglected, the differential decay width of the Michel decay within the framework of the V–A interaction in the SM has a simple form of

$$d\Gamma(\mu^+ \to \mathrm{e}^+\nu_\mathrm{e}\overline{\nu}_\mu) = \frac{m_\mu^5 G_F^2}{192\pi^3} x^2 \left[(3-2x) + P_\mu \cos\theta_{\mathrm{e}^+}(2x-1)\right] dx d(\cos\theta_{\mathrm{e}^+}). \tag{1.9}$$

Fig. 1.14 shows the Michel positron energy spectrum obtained by integrating Eq. (1.9) over the angle assuming the polarization of -0.86. As the spectrum is almost flat in the signal region of $x \approx 1$, the integral can be approximated to be

$$f_{\mathrm{e}^+} \simeq 2\delta x.$$

f_γ can be obtained by the combination of RMD and AIF γ-ray spectra. Fig. 1.15 shows the differential branching ratio of RMD as a function of the positron and the γ-ray energy obtained by integrating Eq. (1.5) over the solid angles. The ratio is high for an energetic positron with a soft photon, where $x \approx 1$ and $y \approx 0$.

The γ-ray spectrum is obtained by integrating Eq. (1.5) over the positron energy and angles. With ignorance of the terms suppressed by $m_{\mathrm{e}^+}/m_{\mu^+}$, it is approximated by

$$\frac{d\mathcal{B}(\mu^+ \to \mathrm{e}^+\nu_\mathrm{e}\overline{\nu}_\mu\gamma)}{dyd(\cos\theta_\gamma)} = \frac{1}{y}[J_+(y)(1 + P_{\mu^+}\cos\theta_\gamma) + J_-(y)(1 - P_{\mu^+}\cos\theta_\gamma)], \tag{1.10}$$

where θ_γ is the angle between the muon spin polarization and the photon momentum, and $J_+(y)$ and $J_-(y)$ are defined by

$$J_+(y) = \frac{\alpha}{6\pi}(1-y)\left[(3\ln\frac{1-y}{r} - \frac{17}{2}) + (-3\ln\frac{1-y}{r} + 7)(1-y) + (2\ln\frac{1-y}{r} - \frac{13}{3})(1-y)^2\right],$$

$$J_-(y) = \frac{\alpha}{6\pi}(1-y)^2\left[(3\ln\frac{1-y}{r} - \frac{93}{12}) + (-4\ln\frac{1-y}{r} + \frac{29}{3})(1-y) + (2\ln\frac{1-y}{r} - \frac{55}{12})(1-y)^2\right].$$

Fig. 1.16 shows the γ-ray spectrum, which is obtained by integrating Eq. (1.10) over the angle assuming the polarization of -0.86. By integrating the spectrum over y, the contribution of the RMD $f_{\gamma,\mathrm{RMD}}$ can be given by the approximation in the signal region of $y \approx 1$ with unpolarized muons as

$$f_{\gamma,\mathrm{RMD}} \simeq \frac{\alpha}{2\pi}(\delta y)^2\{\ln(\delta y) + 7.33\}. \tag{1.11}$$

From the equation, $f_{\gamma,\mathrm{RMD}}$ is roughly proportional to $(\delta y)^2$.

Although Eq. (1.10) is obtained by integration over the positron energy, the transition probability is high for the low energy positrons when the γ-ray energy is high. Fig. 1.17 shows the positron energy spectrum obtained by integrating Eq. (1.5) over the solid angles and with respect to the γ-ray energy from $y = 0.9$ to $y = 1$. It peaks at the positron energy region lower than 8 MeV. This fact leads to an idea of the background suppression by identifying the RMD γ-rays with detection of the low energy positrons which derive from the same muons, which will be discussed in Sec. 2.1.3.3.

Some of the Michel positrons emit γ-rays via AIF and external Bremsstrahlung. The γ-rays from the AIF can have high energies while the contribution from the Bremsstrahlung is negligibly small in the signal region.

The energies of two γ-rays, E_{γ_1} and E_{γ_2}, generated by the AIF are given by

$$
\begin{aligned}
E_{\gamma 1} &= \frac{m_{\mathrm{e}^+}(E_{\mathrm{e}^+} + m_{\mathrm{e}^+})}{E_{\mathrm{e}^+} + m_{\mathrm{e}^+} - p_{\mathrm{e}^+} \cos\theta_{\mathrm{e}^+\gamma_1}}, \\
E_{\gamma 2} &= \frac{(E_{\mathrm{e}^+} - p_{\mathrm{e}^+} \cos\theta_{\mathrm{e}^+\gamma_1})(E_{\mathrm{e}^+} + m_{\mathrm{e}^+})}{E_{\mathrm{e}^+} + m_{\mathrm{e}^+} - p_{\mathrm{e}^+} \cos\theta_{\mathrm{e}^+\gamma_1}},
\end{aligned}
$$

where E_{e^+} and p_{e^+} are the energy and magnitude of the momentum of the positron respectively, and $\theta_{\mathrm{e}^+\gamma_1}$ is the direction of γ_1 momentum with respect to the positron momentum direction. Fig. 1.18(a) shows the energy distribution of the γ-rays from AIF as a function of $\theta_{\mathrm{e}^+\gamma_1}$ when $x = 1$. In most cases, either of the two γ-rays carries most of the energy in which the emission direction of the γ-ray with higher energy is closely aligned to that of the original positron while the energies of the two become equivalent otherwise.

The differential cross section for the annihilation of a positron is given by

$$
\begin{aligned}
\frac{d\sigma}{d\Omega} = -\frac{e^4(E_{\mathrm{e}^+} + m_{\mathrm{e}^+})}{p_{\mathrm{e}^+}} \Bigg[& \frac{1}{(E_{\mathrm{e}^+} + m_{\mathrm{e}^+} - p_{\mathrm{e}^+} \cos\theta_{\mathrm{e}^+\gamma_1})^2} \\
& - \frac{3m_{\mathrm{e}^+} E_{\mathrm{e}^+}}{2m_{\mathrm{e}^+}(E_{\mathrm{e}^+} + m_{\mathrm{e}^+})(E_{\mathrm{e}^+} - p_{\mathrm{e}^+} \cos\theta_{\mathrm{e}^+\gamma_1})} \\
& + \frac{E_{\mathrm{e}^+} + m_{\mathrm{e}^+} - p_{\mathrm{e}^+} \cos\theta_{\mathrm{e}^+\gamma_1}}{2(E_{\mathrm{e}^+} + m_{\mathrm{e}^+})^2(E_{\mathrm{e}^+} - p_{\mathrm{e}^+} \cos\theta_{\mathrm{e}^+\gamma_1})^2} \Bigg],
\end{aligned}
$$

in the laboratory coordinate system [42]. Fig. 1.18(b) shows the differential cross section as a function of $\theta_{\mathrm{e}^+\gamma_1}$ when $x = 1$. Combining Fig. 1.18(a) and Fig. 1.18(b), the energy spectrum of the γ-ray from AIF can be illustrated as Fig. 1.19, which indicates one of the γ-rays carries most of the energy with high frequency.

The contribution from AIF depends on the material distribution of the experiment. Fig. 1.20(a) shows the simulated energy deposit from a muon decay in the MEG II experiment. The AIF spectrum decreases moderately as the energy gets higher while the RMD spectrum decreases rapidly, which results in the high fraction of the AIF γ-ray in the signal region. For instance, the fraction of AIF for $E_\gamma > 48\,\mathrm{MeV}$ is 35%. Fig. 1.20(b) shows the integration of the energy spectrum of the AIF γ-ray as a function of δy. From this figure, the fraction of the AIF γ-ray $f_{\gamma,\mathrm{AIF}}$ can be approximately represented as a square function of δy as well, which is

$$
f_{\gamma,\mathrm{AIF}} \simeq r_{\gamma,\mathrm{AIF}}(\delta y)^2. \tag{1.12}
$$

The normalization $r_{\gamma,\mathrm{AIF}}$ is given by the fraction of the RMD and AIF contributions.

From the above results, the expected number of accidental background events can be written by

$$
\begin{aligned}
N_{\mathrm{acc}} &= \mathcal{B}_{\mathrm{acc}} \times R_\mu \times T \times \Omega \times \epsilon_{\mathrm{e}^+,\mathrm{acc}} \times \epsilon_{\gamma,\mathrm{acc}} \qquad (1.13) \\
&= R_\mu^2 \times T \times (2\delta x) \times \left[\frac{\alpha}{2\pi}(\delta y)^2 \{\ln(\delta y) + 7.33\} + r_{\gamma,\mathrm{AIF}}(\delta y)^2 \right] \times \frac{(\delta z)^2}{4} \times (2\delta t_{\mathrm{e}^+\gamma}) \\
&\propto R_\mu^2 \times T \times \delta x \times (\delta y)^2 \times (\delta z)^2 \times \delta t_{\mathrm{e}^+\gamma},
\end{aligned}
$$

where $\epsilon_{\mathrm{e}^+,\mathrm{acc}}$ and $\epsilon_{\gamma,\mathrm{acc}}$ are the detection efficiencies of background positrons and γ-rays, respectively.

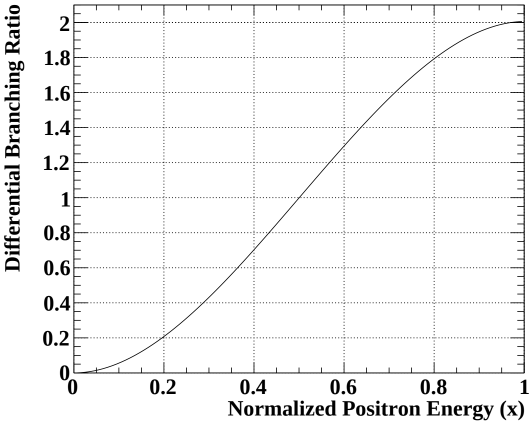


Figure 1.14 Energy spectrum of Michel positron. The muon spin polarization is set to −0.86.

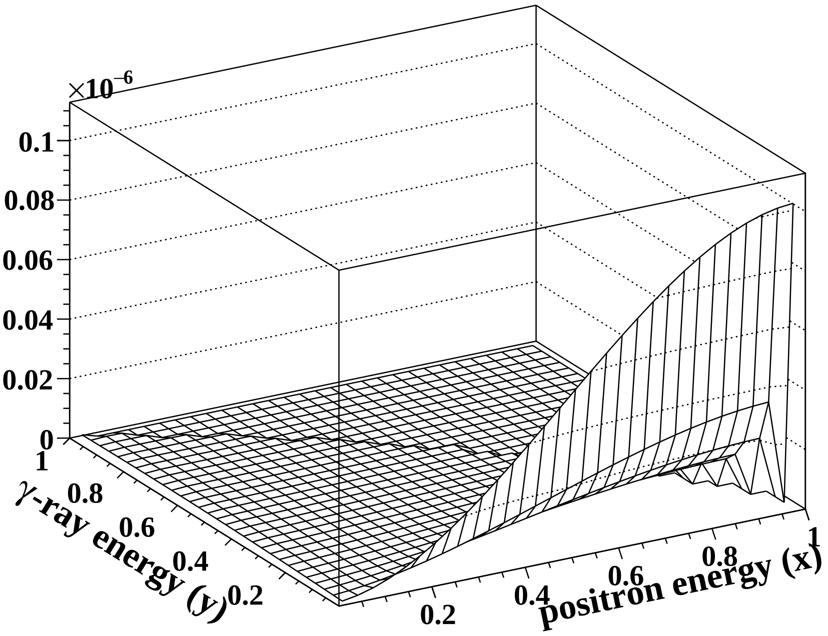


Figure 1.15 Differential branching ratio of the RMD as a function of the positron and γ-ray energies. The muon spin polarization is set to −0.86.

In addition to the RMD and AIF, there is a contribution from accidental pileup of multiple γ-rays when the muon beam intensity is high. The effect of the pileup will be discussed in Chap. 5.

Fig. 1.21 shows the effective branching ratio of the backgrounds as a function of the lower edge of the integration regions with respect to $E_{\mathrm{e}+}$ and E_γ in the MEG experiment. The effective branching ratio of the accidental background is larger than that of the RMD by one or two orders of magnitude, and so the accidental background is a dominant background in this experiment.

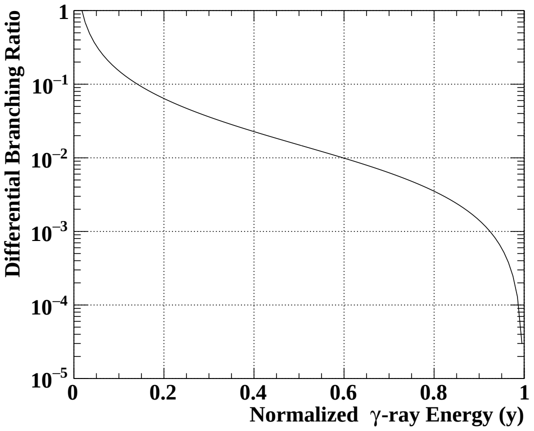


Figure 1.16 Energy spectrum of RMD γ-ray. The muon spin polarization is set to -0.86.

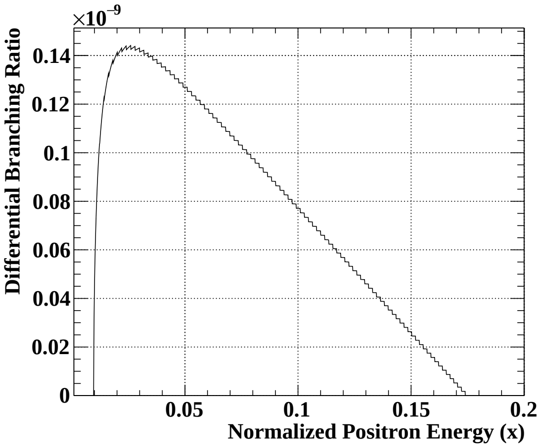
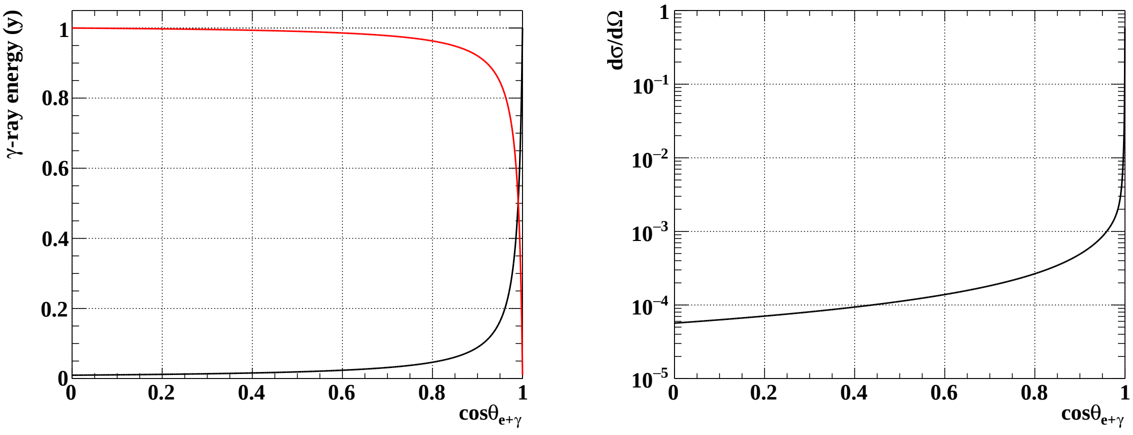


Figure 1.17 Differential branching ratio of the RMD as a function of the positron energy with γ-ray energy of $y = 0.9$–1. The muon spin polarization is set to -0.86. The lower edge of the distribution is given by the kinematics of $2\sqrt{r}$ in Eq. (1.6).



(a)

(b)

Figure 1.18 (a) Energies of γ-rays from AIF and (b) normalized differential cross section as a function of the angle between the positron and one γ-ray directions (black) when $x = 1$. That for the other γ-ray is also drawn in red.

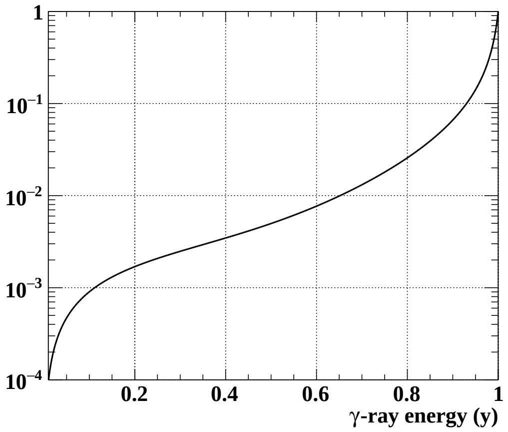
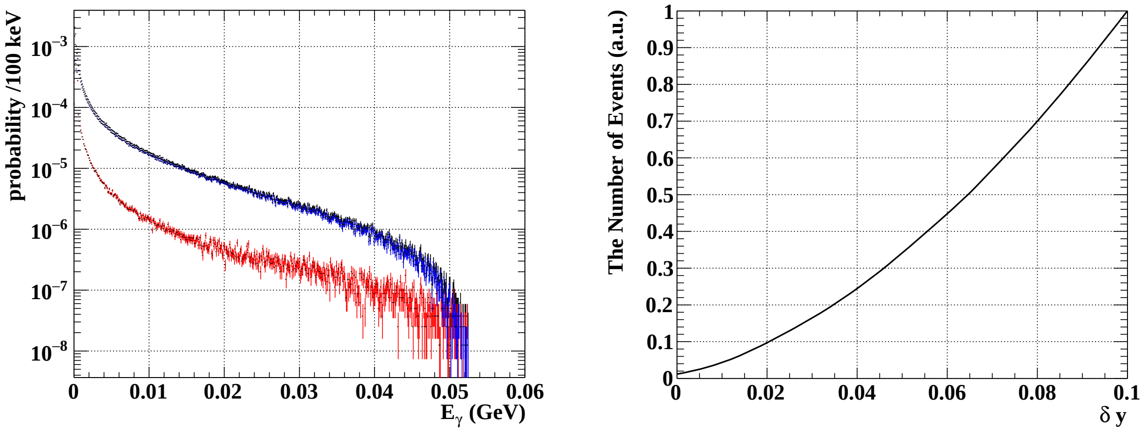


Figure 1.19 Energy spectrum of a γ-ray from AIF when $x = 1$.



(a) (b)

Figure 1.20 (a) Simulated energy deposit from a muon decay on target for RMD decay (blue) and Michel decay (red) in the MEG II. (b) Integrated energy spectrum of AIF as a function of δy obtained by integrating the spectrum of (a) with respect to y from $1 - \delta y$ to 1. The y-axis is normalized at $\delta y = 0.1$.

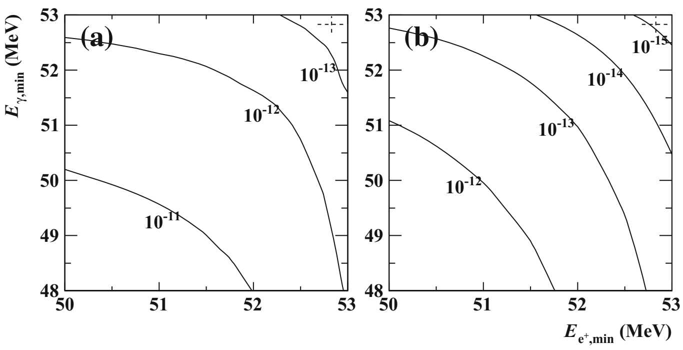


Figure 1.21 Effective branching ratios of (a) accidental background and (b) RMD background in $E_{\mathrm{e^+},\mathrm{min}} < E_{\mathrm{e^+}} < 53.5\,\mathrm{MeV}$, $E_{\gamma,\mathrm{min}} < E_\gamma < 53.5\,\mathrm{MeV}$, $|t_{\mathrm{e^+}\gamma}| < 0.24\,\mathrm{ns}$ and $\cos\Theta_{\mathrm{e^+}\gamma} <$ -0.9996. The former one was evaluated from the timing side-band data and the latter was calculated with theoretical formula folded with detector responses [1].

1.2.3 Experimental requirements

In order to achieve a high sensitivity, N_{sig} should be maximized while N_{RMD} and N_{acc} should be minimized. This leads to the following requirements.

Firstly, a high-intensity direct-current (DC) muon beam is required. High muon statistics results in the large number of signals from Eq. (1.4). However, the number of the accidental backgrounds increases quadratically with R_{μ} while that of the signals increases linearly. Thus, a DC muon beam is a better choice to acquire the same statistics with a lower beam intensity than a pulsed beam.

Secondly, the improvement of the detector resolutions is essential for the high sensitivity. As shown in Eq. (1.7) and Eq. (1.13), both of the number of the backgrounds depend on the energy and opening angle resolutions, and that of the accidental backgrounds depends on the timing resolutions as well.

Furthermore, the reduction of γ-ray backgrounds from AIF can be realized by minimizing the material throughout the positron path.

1.3 MEG experiment

The MEG experiment searched for the $\mu^{+} \to \mathrm{e}^{+}\gamma$ decay with the highest sensitivity in the world. Fig. 1.22 shows the schematic of the MEG experiment. The world's most intense DC muon beam available in the Paul Sherrer Institut (PSI) in Switzerland was used in the experiment. Muons were stopped in a target, and then decayed into positrons and γ-rays. The tracks of positrons were bent into spirals in the magnet produced by a superconducting solenoid and their momenta, vertices and timings were measured with drift chambers and timing counters. The energies, positions and timings of γ-rays were measured with a liquid xenon detector. Thanks to the high performance of these detectors, the MEG experiment searched for the decay with the sensitivity of 5.3×10^{-13}. Although the positive signal was not observed, the experiment succeeded in setting the current upper limit on the branching ratio, $\mathcal{B}(\mu \to \mathrm{e}\gamma) < 4.2 \times 10^{-13}$ (90% confidence level) [1].

Further enhancement of the sensitivity cannot be expected in the MEG experiment even with more statistics. Fig. 1.23 shows the improvement of the sensitivities of the MEG thanks to the statistics used for each analysis. The figure indicates the degree of the sensitivity enhancement by an increased statistics becomes less with accumulation of the data. This is because the increase of statistics contributes to the increase of the backgrounds as well. As shown in Fig. 1.21, the effective branching ratio of the accidental backgrounds in $52.5\,\mathrm{MeV} < E_{\mathrm{e}^{+}} < 53.5\,\mathrm{MeV}$ and $52.0\,\mathrm{MeV} < E_{\gamma} < 53.5\,\mathrm{MeV}$ is in the order of 10^{-13} in the MEG. Considering the energy resolution of 330 keV (core) for positrons and 2.3% for γ-rays in the MEG, a fundamental improvement of the detector performance is unavoidable to reach the sensitivity of $\mathcal{O}(10^{-14})$.

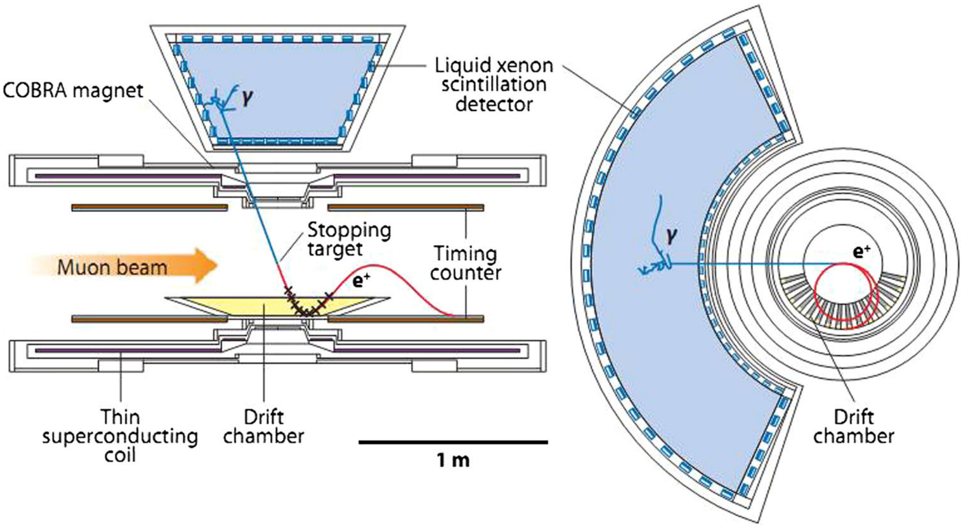


Figure 1.22 Schematic of the MEG experiment.

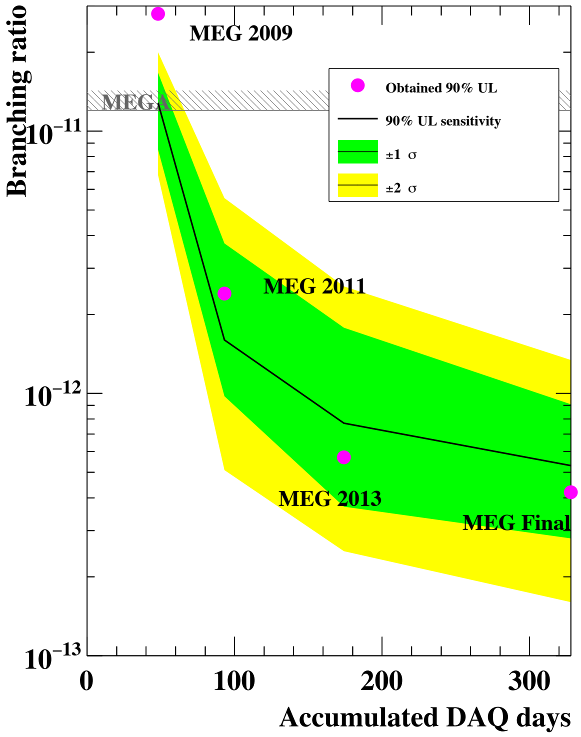


Figure 1.23 Branching ratio sensitivities (black line) and upper limits (purple markers) of MEG results.

1.4 MEG II experiment

The MEG II experiment, an upgrade of the MEG experiment, was proposed to improve the sensitivity by one oder of magnitude [43]. Fig. 1.24 shows the schematic view of the MEG II experiment. Muons are stopped in the target and emit positrons and γ-rays. The positrons are bent in a magnetic field produced by COBRA magnet, and then detected with the cylindrical drift chamber and the pixelated timing counter while the γ-rays are detected with the liquid

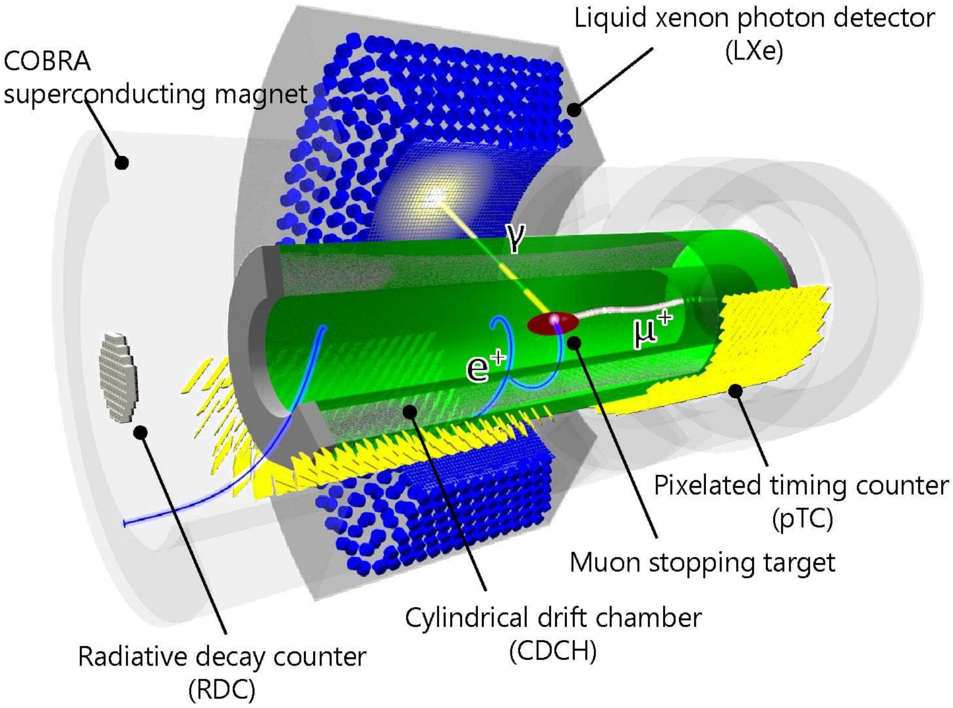


Figure 1.24 Schematic of the MEG II experiment [2].

xenon detector. The radiative decay counter is newly installed for tagging γ-ray backgrounds. The improvement in the sensitivity is achieved by upgrades of the detectors in two aspects.

The first strategy is to improve the detector resolutions, which is effective to reduce the accidental backgrounds. As shown in Eq. (1.13), the number of accidental backgrounds increases with the resolutions. Therefore, the MEG II detectors were designed to achieve better resolutions by a factor of two than the MEG detectors.

The second is an increase of the accumulated statistics. The muon beam rate in the MEG experiment was limited to 3×10^7 μ^+stops/s, which is lower than a half of the maximum intensity, due to the limited performance of the detectors; the drift chamber was not operational under a high intensity environment causing discharge, and the performance of the MEG detectors can significantly deteriorate at the increased background level. Thus, an improvement of the detectors to be tolerant toward an increase of the beam intensity enables a gain in the total number of muons. In addition, the efficiency of the positron spectrometer is doubled by reducing the material between the drift chamber and the timing counter, which causes positron scattering.

Chapter 2

MEG II Experiment

This chapter begins with explanations about apparatus for the MEG II experiment. Then, the configurations of the pilot runs for the detector commissioning are summarized followed by the descriptions about the detector simulation. Finally, the procedure for the $\mu^+ \to \mathrm{e}^+\gamma$ analysis is described.

2.1 Apparatus

In this section, the apparatus for the MEG II experiment is explained such as the beamline, the detectors, the DAQ system.

Here we define a global coordinate system (x, y, z) as follows. The origin is defined at the center of the COBRA magnet and the muon stopping target is located there. The z-axis is set parallel to the muon beam direction, and the y-axis is set vertical upward. The x-axis is set horizontal, so that (x, y, z) forms a right-handed system. A cylindrical coordinate system (r, θ, ϕ) is also defined as $r := \sqrt{x^2 + y^2}$, $\theta := \tan^{-1}(z/r)$, $\phi := \tan^{-1}(y/x)$.

2.1.1 Beamline

The MEG II experiment uses the muon beam at PSI. Fig. 2.1 shows an overview of the beam line. The beam is transported along the arrows to the MEG II detectors placed in the πE5 beamline.

Proton acceleration

The muon beam at PSI originates from protons. The protons from a hydrogen source are accelerated in three steps. At the first step, a Cockcroft–Walton pre-accelerator shown in Fig. 2.2 accelerates them up to 870 keV [45]. The pre-accelerated protons are injected into a isochronous cyclotron called Injector 2, which consists of four magnets, two radio-frequency (RF) cavities operated at $\sim$ 400 kV and two third-harmonic cavities operated at $\sim$ 30 kV shown in Fig. 2.3. This first cyclotron accelerates the protons up to 72 MeV and the beam is delivered to the second cyclotron, the main ring cyclotron. The second cyclotron is composed of four main RF cavities and eight magnets as shown in Fig. 2.4, and accelerates the beam up to 590 MeV. Table 2.1

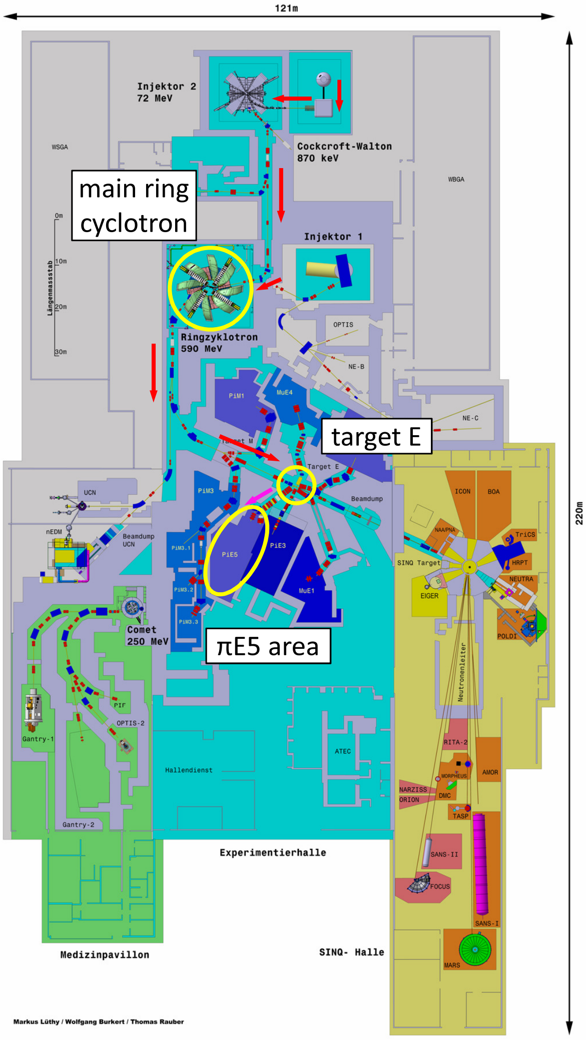


Figure 2.1 Experimental hall in PSI [44]. The beam path is shown in arrows.

summarizes the characteristics of the main ring cyclotron.

Muon production

The accelerated protons are injected into two graphite targets called target M and target E (Fig. 2.5), which have 5 mm and 40 mm thickness along beam axis, respectively. In these targets, pions are produced and subsequently muons are produced by $\pi^+ \to \mu^+\nu_\mu$ decay. Pions have a lifetime of $\sim$ 26 ns at the rest, which is long enough to regard the extracted muon beam as a DC beam effectively.

πE5 beamline

πE5 beamline, where the MEG II experiment is conducted, provides the highest intensity muon beam in the seven beamlines stretching from target E. This beamline has an angle of 166° with respect to the proton beam direction in order to collect surface muons with magnets, which originate from pions stopped in a thin region of the target surface. The surface muons can be stopped on a thin target thanks to small momenta of $\sim$ 28 MeV/c with 5–7% (FWHM).

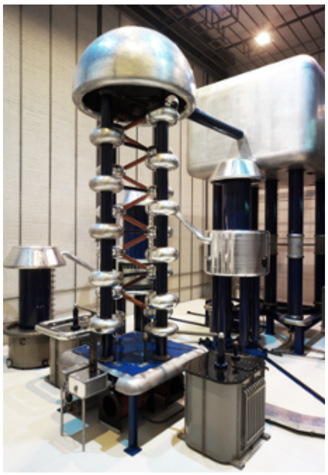


Figure 2.2 Cockcroft–Walton pre-accelerator [46].

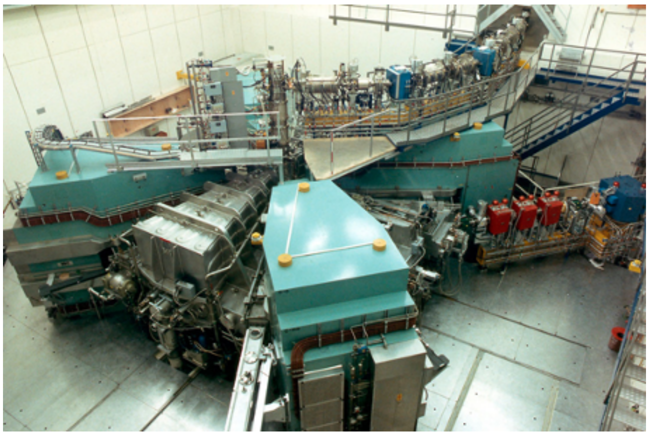


Figure 2.3 Injector 2 [47].

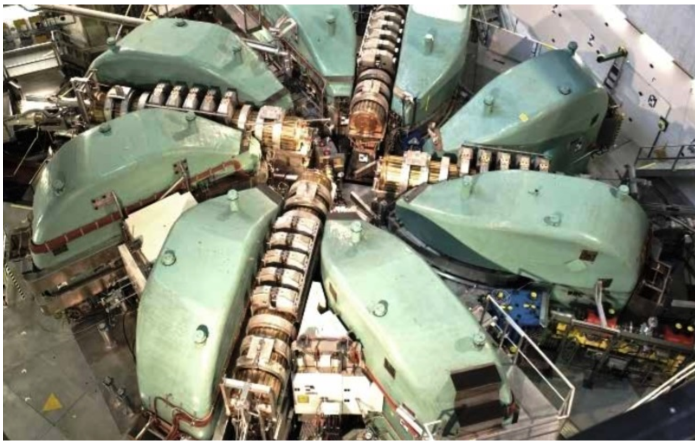


Figure 2.4 Main ring cyclotron [44].

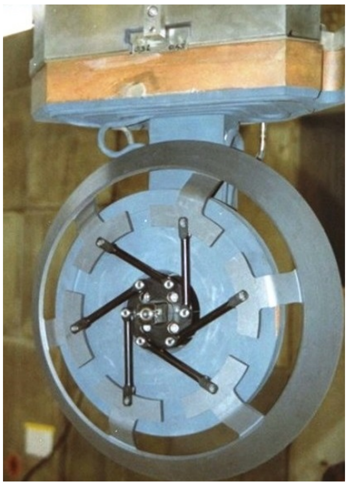


Figure 2.5 Pion production target [44].

Table 2.1 Characteristics of the main ring cyclotron [48].

Injection energy	72 MeV
Extraction energy	590 MeV
Extraction momentum	1.2 GeV/c
Beam current	2.2–2.4 mA DC
Accelerator frequency	50.63 MHz
Time between pulses	19.75 ns
Bunch width	0.3 ns

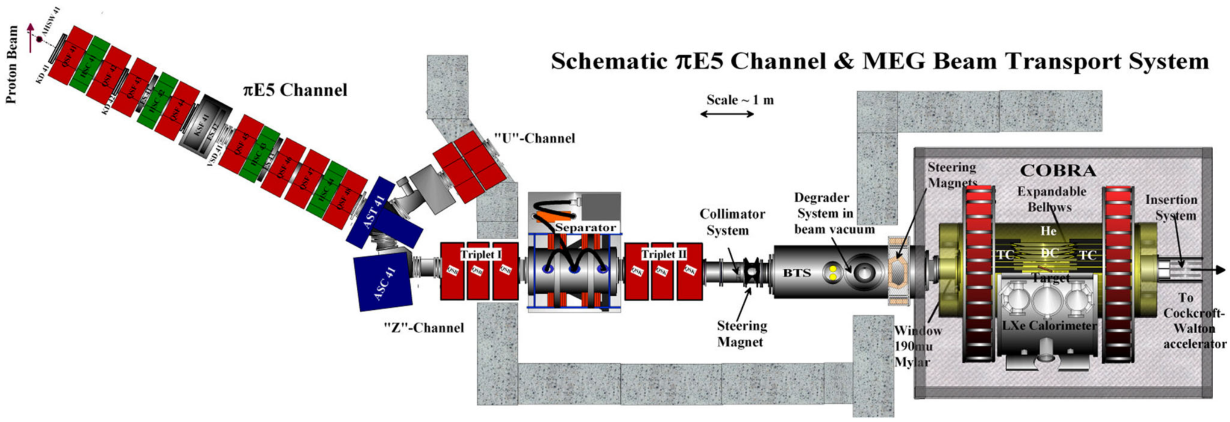


Figure 2.6 Schematic view of the πE5 beamline and the beam transport system in the MEG Ⅱ experiment [2].

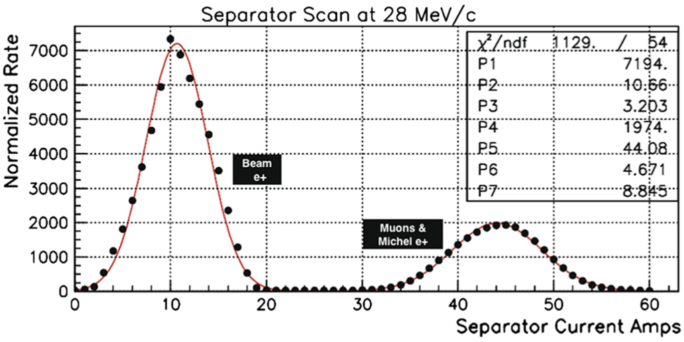


Figure 2.7 Measurement of the separation quality with the Wien-filter [2].

Fig. 2.6 shows a schematic view of the πE5 beamline and the beam transport system in the MEG Ⅱ experiment. The muon beam is delivered to the MEG Ⅱ beam transport system adjusting its momentum with bending magnets, quadrupoles, hexapoles and slits through the πE5 beamline. The transport system consists of the first quadrupole triplet magnet (Triplet1), a Wien-filter (Separator), the second quadrupole triplet magnet (Triplet2), and the beam transport solenoid (BTS) in this order. The muon beam contains positron contamination whose amount is eight times higher than the muons, which is derived from the Michel decay of the muon and the π^0 decay in the target. It is separated with the separator after being squeezed with the first magnet, which results in positrons less than 1% of muons (Fig. 2.7). After that, the muon beam is squeezed again with the second magnet. The BTS has a role of focusing the beam on a muon stopping target. The beam is collimated with a collimator, and then the momenta of muons are adjusted with a 300 μm-thick Mylar degrader to maximize a stopping efficiency in the target.

2.1.2 Muon stopping target

The muon beam is injected into the muon stopping target placed at the center of the MEG Ⅱ detectors. The target needs to have a high muon stopping efficiency with low material

Figure 2.8 Muon stopping target in the MEG II experiment [2].

to avoid multiple Coulomb scattering of the emitted positrons, and the production of γ-ray backgrounds from positron bremsstrahlung or annihilation. In the MEG experiment, the deformation of the target caused a large systematic uncertainty in the reconstructed positron direction, which resulted in the largest systematics of 13%. Therefore, the MEG II target must be mechanically stable. In addition, a monitoring system of the target deformation is introduced.

In the pilot runs for the MEG II experiment, a $174\,\mu$m thick plastic scintillating film was used as a target (Fig. 2.8). It was placed 15° slanted from the beam axis to make an effective thickness for the muon beam longer and that for the outgoing positron shorter. Optical markings are printed on the target and their positions were measured with a CCD camera to monitor the deformation.

2.1.3 Detectors

The MEG II detectors were upgraded from the MEG as discussed in Sec. 2.1. In addition, a new detector for further background suppression is installed.

2.1.3.1 Positron spectrometer

The positron spectrometer in the MEG II experiment consists of COnstant Bending RAdius (COBRA) superconducting magnet to bend positron tracks, a cylindrical drift chamber (CDCH) to track positrons and a pixelated timing counter (pTC) to measure positron timings. They were designed to satisfy the requirements to have high resolutions and a high detection efficiency, and to be tolerant of a high background environment from Michel decays.

COBRA magnet

The COBRA magnet is a superconducting solenoid developed for the MEG experiment and reused in the MEG II experiment. It forms a gradient magnetic field by placing coils with different radii and wire winding densities (Fig. 2.9). The field strength is 1.27 T at the center and decreasing to 0.49 T at the both ends. This gradient magnetic field is designed so that positrons with the same momenta emitted from the target to have the same projected bending radii independent of their emission angles (Fig. 2.10(a)). This enables us to detect positrons with the high momenta selectively. Moreover, positrons can be swept away quickly even if they are emitted almost perpendicularly to the beam axis because of the decreasing magnetic field

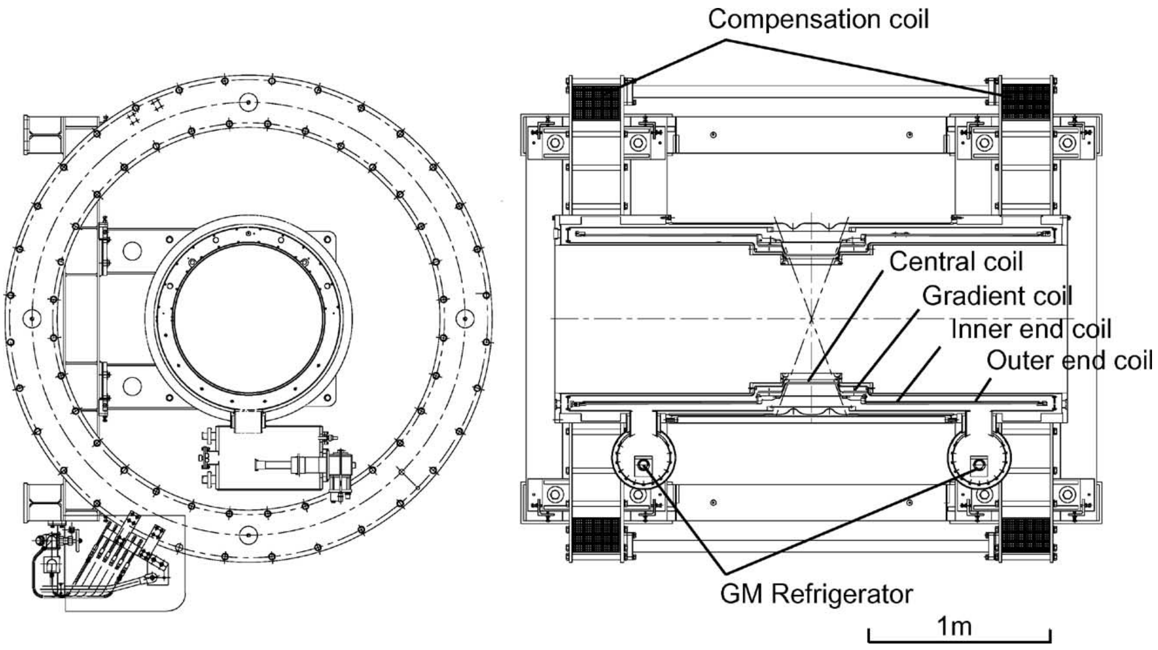
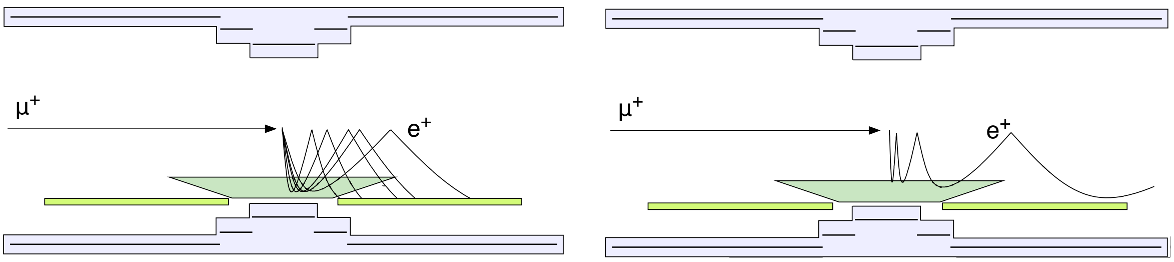


Figure 2.9 Schematic view of the COBRA magnet [49].



(a)

(b)

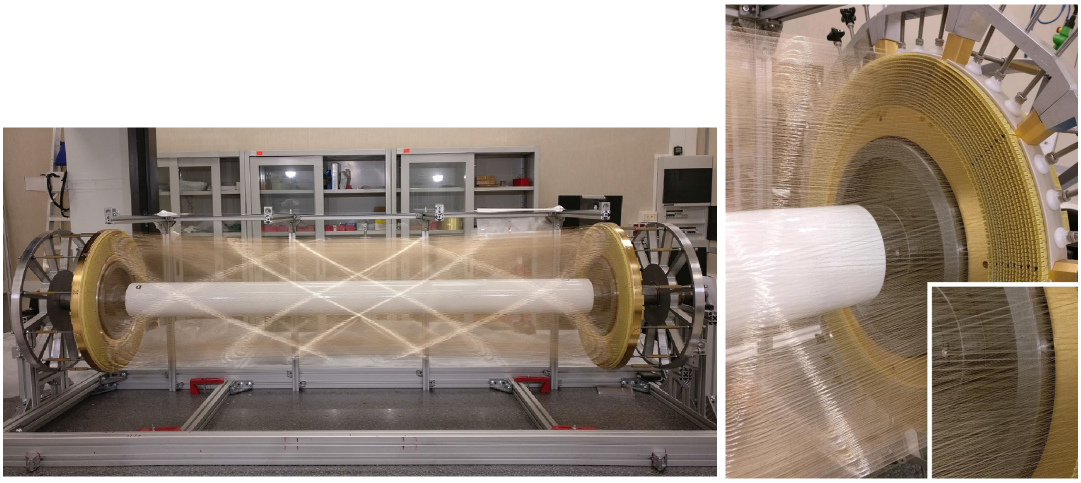
Figure 2.10 Principle of COBRA magnet [1]. (a) Positrons with the same momenta have the same radius. (b) Positrons are swept away quickly.

(Fig. 2.10(b)). Therefore, the hit rate at the positron detectors can be suppressed.

The COBRA magnet was designed not to interfere γ-ray detection. It has a thin window with $0.197X_0$ thickness at the acceptance region of the γ-rays so that the signal γ-rays can penetrate it. The PMTs used for the γ-ray detector are not operational in magnetic field, and thus the compensation coils are placed to reduce the stray field at the location of the PMTs from the COBRA magnet down to 50 G.

Cylindrical drift chamber

A new single-volume wire drift chamber was designed to have a high transparency for the MEG II experiment since the positron detection efficiency in the MEG experiment was low due to the multiple Coulomb scattering at mechanical supports and electronics of the drift chamber. It has a cylindrical shape with the inner radius of 17 cm, outer of 29 cm and the length of 191 cm (Fig. 2.11). A low-density gas mixture of He and C_4H_{10} is enclosed in the volume with the



(a) (b)

Figure 2.11 (a) Overview and (b) an enlarged view of the end cap of the MEG II drift chamber [2].

proportion of 90 : 10. The total radiation length is $1.58 \times 10^{-3} X_0$ in a single turn while it was $2.0 \times 10^{-3} X_0$ in the MEG drift chamber. The low material budget also reduces generation of the γ-ray backgrounds from the positron annihilation in flight.

The drift chamber consists of nine drift cell layers as shown in Fig. 2.12[*1]. Sense wires made of gold plated tungsten with the diameter of 20 μm are centered in each cells, and silver plated aluminum field wires with the diameter of 40/50 μm surround the sense wires. Wire planes are slanted at alternating signs stereo angles raging from 6.0° to 8.5°, which allows to reconstruct the z-longitudinal coordinate precisely. The drift cells are almost square from 6.6 mm in the innermost layer to 8.7 mm in the outermost layer. Each drift cell layer consists of two criss-crossing field wire planes enclosing a sense wire plane.

Pixelated timing counter

A pixelated timing counter is placed outside the drift chamber and measures the positron timings. It is composed of an upstream and a downstream sector. The timing counter for the MEG II experiment is a highly segmented scintillation counter made of BC422 fast plastic scintillator (Saint-Gobain). It consists of 512 small scintillator tiles (Fig. 2.13(b)), 256 tiles for each sector, while 30 scintillator bars with the size of $40 \times 40 \times 800\,\mathrm{mm}^3$ were used for the MEG experiment (Fig. 2.13(a)). The size of the scintillation tiles is $40 \times 120 \times 5\,\mathrm{mm}^3$ or $50 \times 120 \times 5\,\mathrm{mm}^3$ depending on the position as shown in Fig. 2.14. The pileup probability can be significantly reduced, which enables operation in a high-rate environment.

This segmented design enhances the timing resolution. The resolution of the MEG timing counter was limited by a large variation in the optical photon path lengths due to the large scintillator size, and thus the small dimensions of the single counter can make the timing resolution better. In addition, since each particle can hit multiple counters, nine hits on average

---

[*1] This figure includes ten layers though the most outer layer was removed in the real detector.

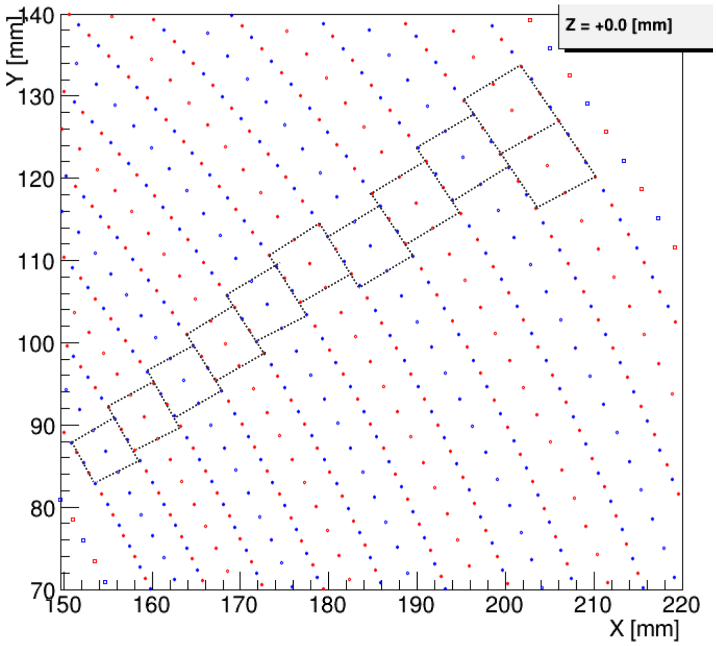


Figure 2.12 Drift cells configuration at the center of the drift chamber [2].

(Fig. 2.15(a)), the combined timing resolutions can be greatly better than the resolution of the single counter; the resolution improves as $\sigma_t^{\mathrm{single}}/\sqrt{N_{\mathrm{hit}}}$, where $\sigma_t^{\mathrm{single}}$ is the timing resolution of a single counter, which includes the intrinsic resolution of a single counter, the error in time alignment over the counters, and the electronics jitter. Fig. 2.15(b) shows the dependence of the combined timing resolution on the number of hits N_{hit}, which was measured for Michel positrons in the pilot run. The resolution of 31 ps is achieved at $N_{\mathrm{hit}} = 9$.

The silicon photomultiplier (SiPM) from AdvanSiD (ASD-NUV3S-P High-Gain (MEG)) is used for the MEG II timing counter instead of the photomultiplier tube (PMT) for the MEG because it is superior for the timing counter in the two aspects: the compactness and the high immunity to magnetic fields; the degradation of the performance of PMTs in the magnetic field deteriorated the timing resolution in the MEG. Each tile is readout by six series-connected SiPMs attached to each end.

The scintillation tile is wrapped with an enhanced specular reflector (ESR) from 3M (Fig. 2.16). An optical fiber is inserted to the scintillator tile in order to inject a synchronous laser signal for the timing alignment between tiles.

Expected performance

The performance of the spectrometer is evaluated with MC and the expectation is summarized in Table 2.2.

2.1.3.2 γ-ray detector

The positions, timings and energies of γ-rays were measured with a liquid xenon (LXe) γ-ray detector in the MEG experiment. LXe of 900 L was used as a scintillator and its scintillation

(a) (b)

Figure 2.13 (a) The MEG positron timing counter [1]. (b) The MEG Ⅱ positron timing counter [2].

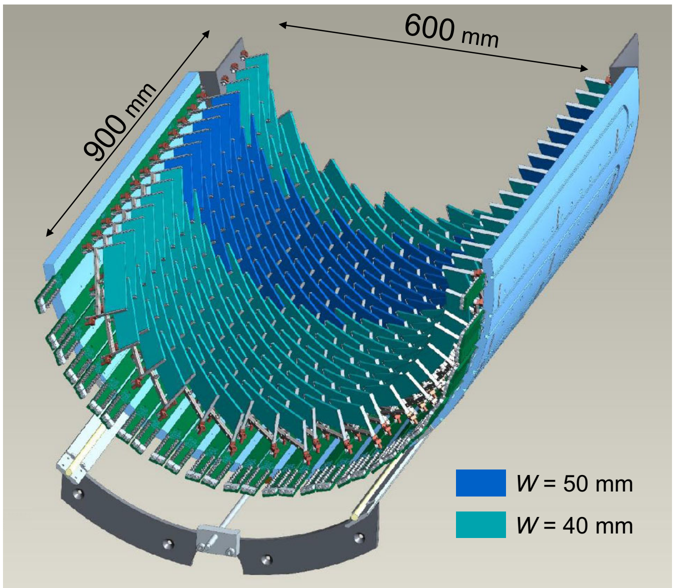
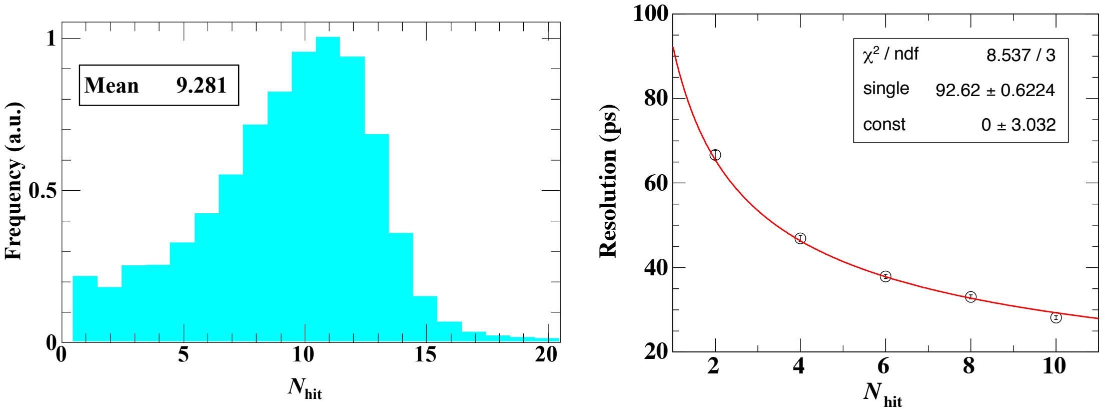


Figure 2.14 Design of the downstream timing counter. Each scintillation counter is 40 mm or 50 mm wide depending on its position.

light was detected by surrounding photosensors as shown in Fig. 2.17. PMTs were used as photosensors and those on the γ-ray entrance face were replaced with SiPMs in the MEG Ⅱ.

In the MEG experiment, 846 PMTs were installed on faces of the detector. The inner face was covered by 216 PMTs with a minimum spacing between adjacent PMTs. However, since the photo-cathode of the PMT has a round-shape, the interval among the PMTs is 63 mm which is much larger than the diameter of the active area of the PMT, 46 mm (Fig. 2.18(a)). This non-uniform PMT coverage caused an even-by-event fluctuation in collection efficiency of scintillation light. Fig. 2.18(b) shows the efficiency of scintillation light collection as a function of the depth of the first interaction for the signal photon. The collection efficiency depends on the incident position: a high efficiency if γ-rays enter in front of the PMTs (case A) and a low efficiency otherwise (case B). This deteriorates the energy and position resolutions, especially for shallow events. Since many γ-rays interact in the shallow region as shown in Fig. 2.19, this



(a) (b)

Figure 2.15 (a) Expected number of hits for a signal positron from the simulation [2]. (b) The timing resolution measured in the pilot run 2016 as a function of the number of hits N_{hit} [2]. The red curve shows the best fit by a function of $\sigma_{\text{single}}/\sqrt{N_{\text{hit}}} \oplus \sigma_{\text{const}}$.

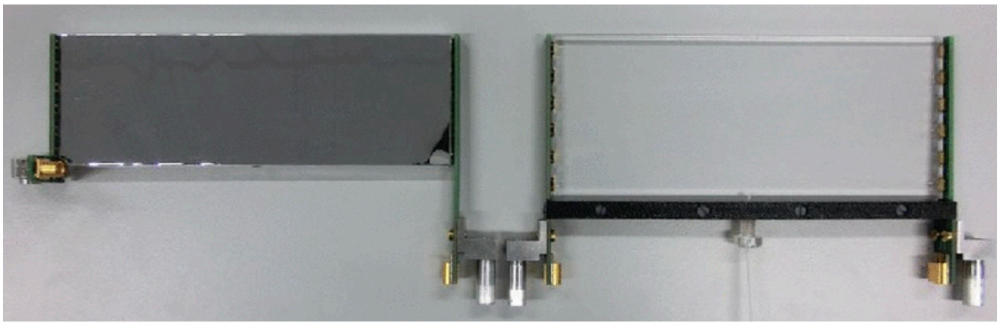


Figure 2.16 Scintillation counters of the MEG II timing counter [2]. Each scintillator has the size of $40, 50 \times 120 \times 5\,\text{mm}^3$. Scintillation light is read out by the six series connected SiPMs attached to the both sides. The fiber inserted in the right scintillator is used for the timing alignment.

fluctuation is a crucial problem.

The main upgrade of the LXe detector for the MEG II experiment is to improve this non-uniform response by replacing the 216 PMTs on the inner face with 4092 SiPMs as shown in Fig. 2.20. Thanks to the square shape of the photosensor, the inner face can be covered uniformly by the sensitive area. Moreover, the higher granularity (Fig. 2.21) leads to better energy and position resolutions of the shallow events.

The SiPM has further advantages compared to the PMTs. Its insensitivity to magnetic fields and photoelectron counting capability enable an easier and more reliable calibration of the detector. Less material budget in the entrance face increases the detection efficiency by 9%. Its lower bias voltage, typically under 100 V, also makes the operation easier.

In addition to the photosensor replacement, the layout of the PMTs is modified to make the best use of them based on MC. The inner face is extended horizontally by 10% to each side (Fig. 2.22). The PMTs on the lateral faces are tilted so that all the photo-cathodes lie in the same plane. The extended volume and the sensor configuration reduce the effect of leakage due

Table 2.2 Positron detector performance [50].

	MEG (measured)	MEG II (simulated)
E_{e^+} resolution (keV)	380	100
θ_{e^+} resolution (mrad)	9.4	6.7
ϕ_{e^+} resolution (mrad)	8.7	4.9
t_{e^+} resolution (ps)	107	40
vertex resolution ($z_{\mathrm{e}^+}/y_{\mathrm{e}^+}$ mm)	2.4/1.2	1.7/0.8
efficiency (%)	30	65

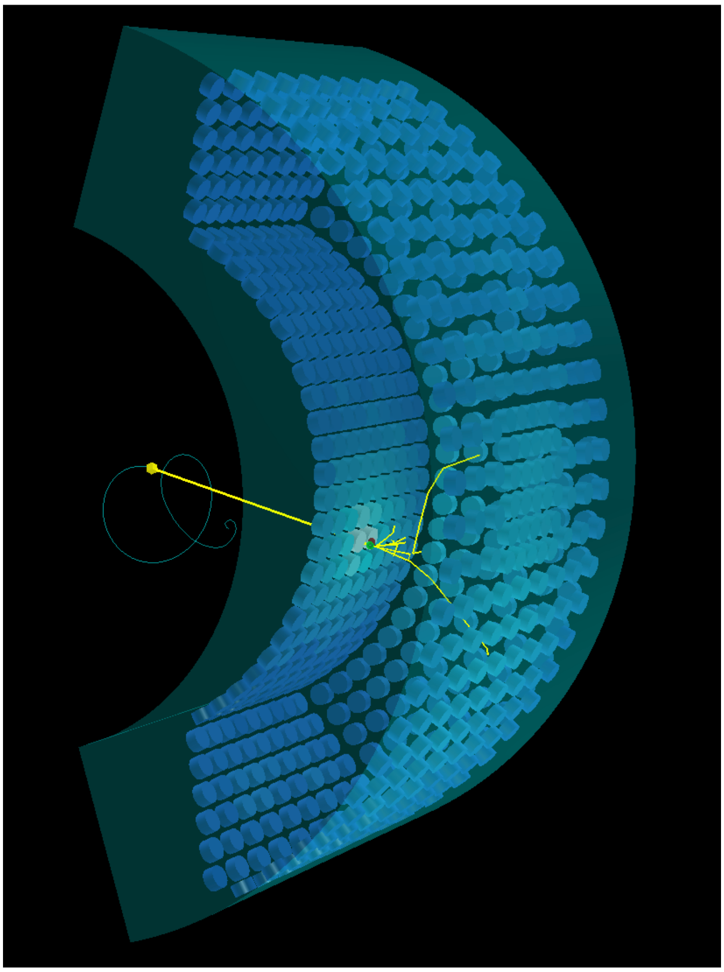


Figure 2.17 Overview of the LXe detector in MEG. A γ-ray enters to the inner face of the detector and scintillation light induced by the γ-ray is detected by photosensors.

to shower fluctuations for events near the lateral walls, which improves the energy resolution especially for those events. The PMTs on the top and the bottom faces are also rearranged to gain uniform response for the events near the top or the bottom face (Fig. 2.23).

Geometry

Fig. 2.24 shows the overview of the LXe detector. The detector has a C-shape with six faces called inner, outer, upstream, downstream, top and bottom face. The distance between the

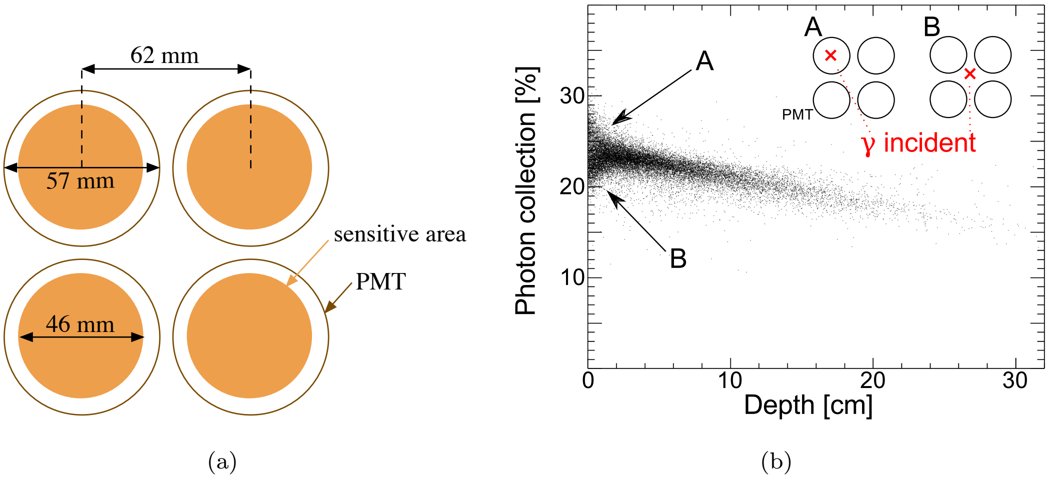


Figure 2.18 (a) Layout of the PMTs on the inner face in MEG. (b) Simulated photon collection efficiency as a function of the conversion depth of signal γ-rays in MEG [2].

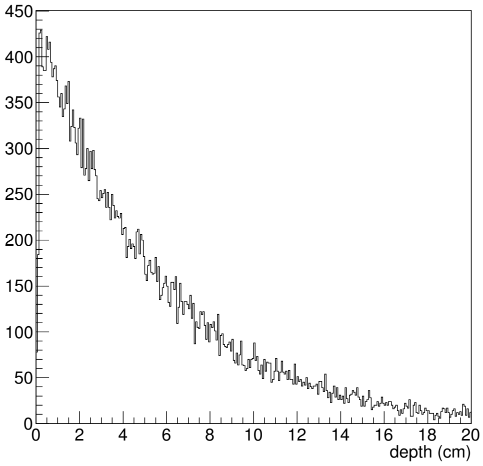


Figure 2.19 Simulated distribution of the conversion depth for the signal γ-rays [2].

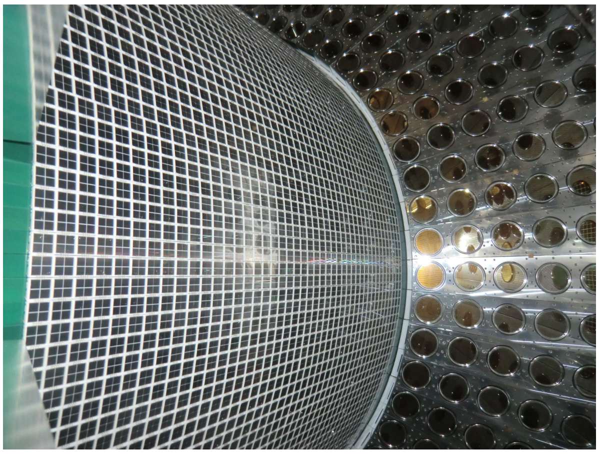


Figure 2.20 The inner face of the LXe detector for the MEG Ⅱ experiment.

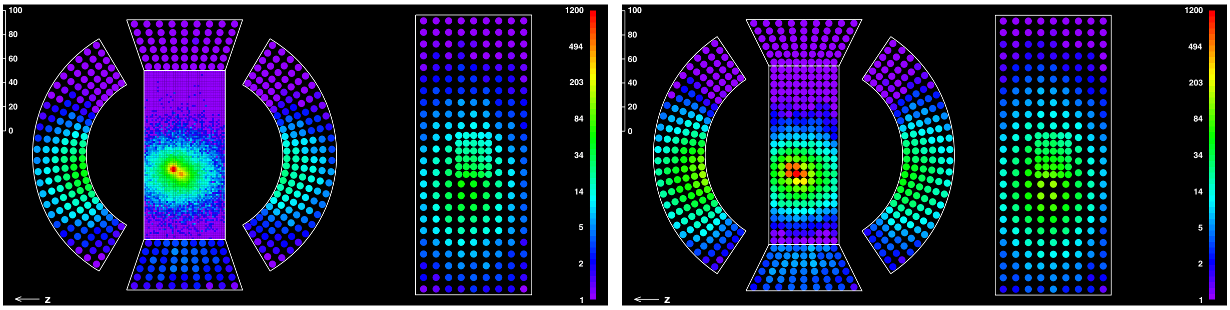


Figure 2.21 Example of scintillating light distributions in MEG Ⅱ (left) and MEG (right) for the same MC event [2].

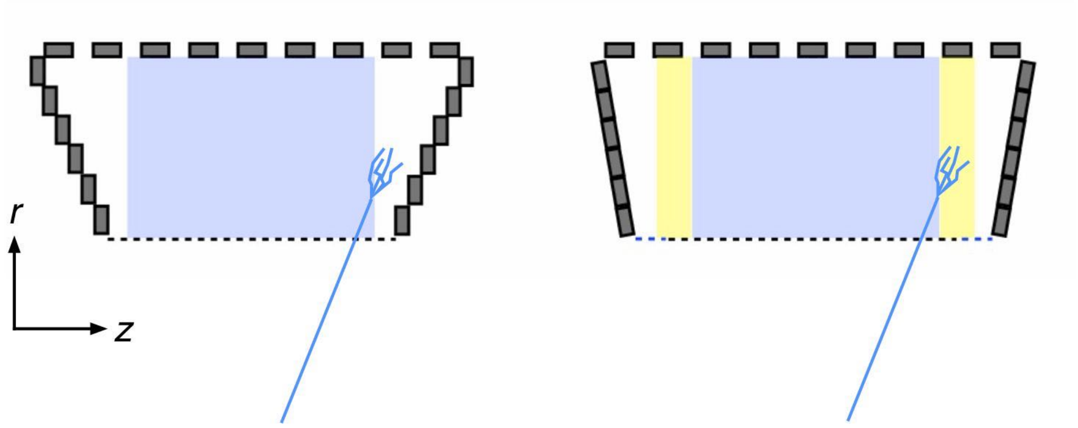
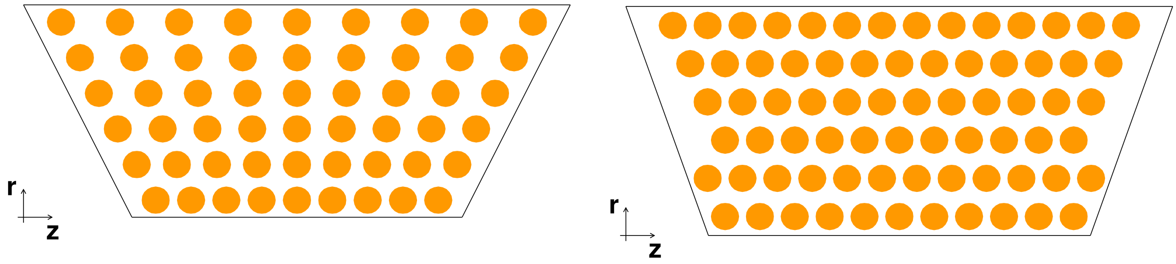


Figure 2.22 Layouts of the PMTs on the lateral faces of the LXe detector for MEG (left) and MEG Ⅱ (right) [2].



(a)

(b)

Figure 2.23 Layouts of the PMTs on the top and the bottom faces of the LXe detector in (a) MEG and (b) MEG Ⅱ.

inner face and the beam axis R_{inner} is 64.84 cm for MEG Ⅱ such that the detector subtends $\sim$11% of the solid angle. The thickness of the active volume is 38.5 cm, which corresponds to $13.8X_0$, to capture the whole electromagnetic shower from the signal γ-ray.

The development view with the definition of the local coordinate system is illustrated in Fig. 2.25. A local coordinate (u, v, w) is defined to describe a position of a γ-ray vertex in the

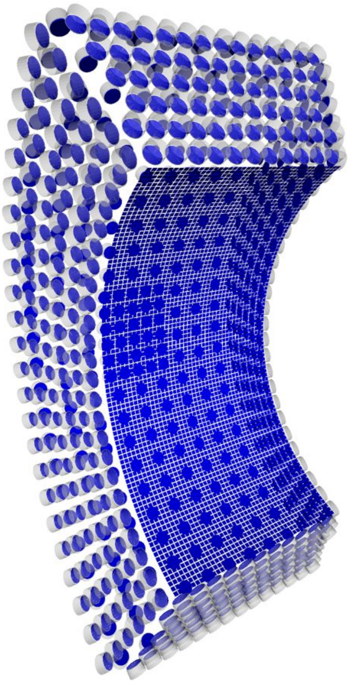


Figure 2.24 Overview of the LXe detector for the MEG II.

detector as

$$\begin{aligned} u &= z, \\ v &= \tan\left(-y/x\right) \times R_{\text{inner}}, \\ w &= \sqrt{x^2 + y^2} - R_{\text{inner}}, \end{aligned}$$

where u and v correspond to the horizontal and vertical coordinates on the inner face, respectively, and w is the depth from the inner face. The active volume of this detector, where γ-ray hits are analyzed, is defined as $|u| < 23.9\,\text{cm}$, $|v| < 67.9\,\text{cm}$ and $0 < w < 38.5\,\text{cm}$.

The detector cryostat consists of an inner and an outer vessel to keep xenon in liquid state. The layer between the two vessels is in vacuum state to thermally insulate the inner vessel filled with liquid xenon, and super insulation sheets are installed between the vessels to prevent a radiant heat. An entrance window γ-rays go through is required to be thin to suppress their conversions before the active volume. It is composed of a honeycomb-structure aluminum window covered with a carbon fiber plate resulting in material thickness of $0.081X_0$ radiation length. The total material budget of the γ-ray entrance window is $0.110X_0$ including that of the photosensors, $0.029X_0$.

Liquid xenon

Liquid xenon is used as a scintillator for several advantages. Table 2.3 summarizes its properties. Thanks to the large atomic number and high density, LXe has a high stopping power, which allows a compact detector with a reasonable detection efficiency. The high light yield provides a good detector resolution. The fast decay time of the scintillation is suitable for an operation in a high pileup environment. In addition, a liquid scintillator is easier to obtain a

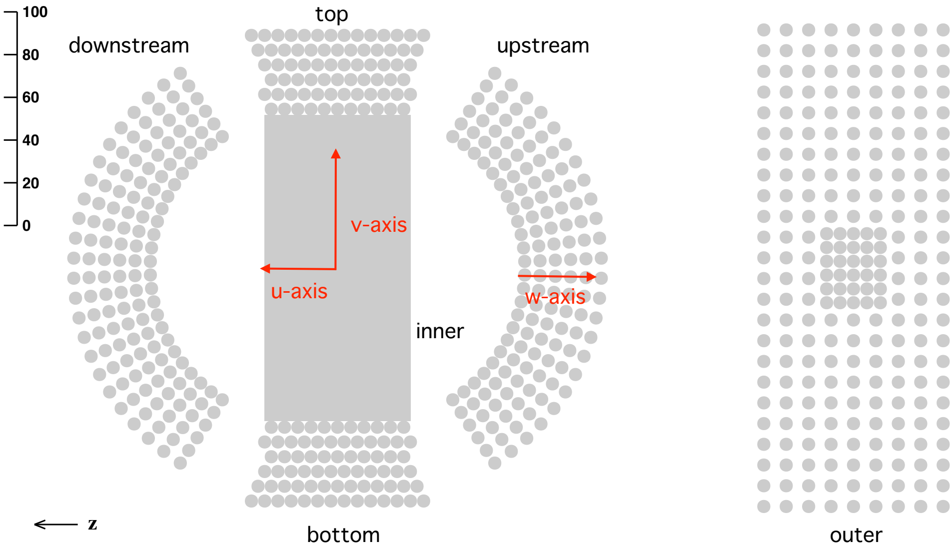


Figure 2.25 Development view of the LXe detector for the MEG II. The red axises show directions of the local coordinate on the detector.

uniform response than a crystal scintillator.

Table 2.3 Properties of LXe.

Item	Value
Atomic Number	54
Density	2.953 g/cm^3 [51]
Radiation length	2.872 cm [51]
Moliere radius	5.224 cm
Scintillation Wavelength (mean)	$174.8 \pm 0.1(\text{stat.}) \pm 0.1(\text{syst.})$ nm [52]
Scintillation Wavelength (FWHM)	$10.2 \pm 0.2(\text{stat.}) \pm 0.2(\text{syst.})$ nm [52]
Decay time (fast)	4.2 ns [53]
Decay time (slow)	22 ns [53]
Decay time (recombination)	45 ns [53]
W-value for electron	21.6 eV [54]
W-value for alpha	17.9 eV [54], 19.6 eV [55]
Refractive index (for $\lambda = 175$ nm)	1.65
Critical energy	14.5 MeV [56]

Another advantage is its transparency against its own scintillation light. When a particle deposits its energy in LXe, scintillation light is produced by the self-trapping of excited xenon atoms (Xe^*) in two processes: direct excitation (Eq. (2.1)) and recombination of electron-ion pair (Eq. (2.2)). They create excited atoms, which combine with a neutral ground-state Xe atom to form the molecular dimer Xe^{**}. The dimer decays to the monatomic ground state via emission of a vacuum ultraviolet (VUV) photon. Since the scintillation light is generated via the excimer not via the excited xenon atom itself in LXe, the yielded light is not absorbed by

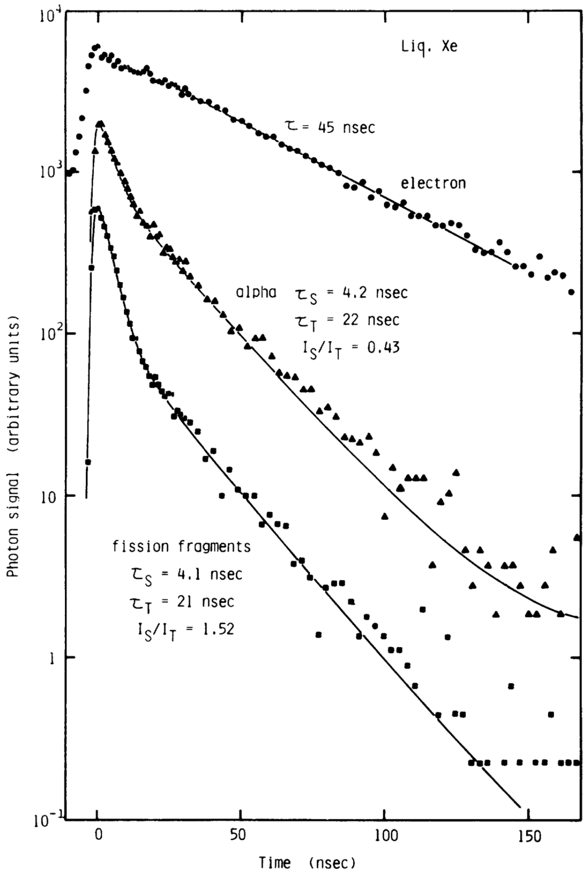


Figure 2.26 LXe scintillation decay time for each particle [53].

LXe.

$$\begin{aligned} \mathrm{Xe}^* + \mathrm{Xe} + \mathrm{Xe} &\rightarrow \mathrm{Xe}_2^* + \mathrm{Xe} \\ \mathrm{Xe}_2^* &\rightarrow 2\mathrm{Xe} + h\nu \end{aligned} \tag{2.1}$$

$$\begin{aligned} \mathrm{Xe}^+ + \mathrm{Xe} \quad &\rightarrow \mathrm{Xe}_2^+ \\ \mathrm{Xe}_2^+ + \mathrm{e}^- \ &\rightarrow \mathrm{Xe}^{**} + \mathrm{Xe} \\ \mathrm{Xe}^{**} \ &\rightarrow \mathrm{Xe}^* + \mathrm{heat} \\ \mathrm{Xe}^* + \mathrm{Xe} + \mathrm{Xe} \ &\rightarrow \mathrm{Xe}_2^* + \mathrm{Xe} \\ \mathrm{Xe}_2^* \ &\rightarrow 2\mathrm{Xe} + h\nu \end{aligned} \tag{2.2}$$

The decay of the excimer, $\mathrm{Xe}_2^* \rightarrow 2\mathrm{Xe} + h\nu$, has two decay time constants since the excimer has mainly two excited molecular states of singlet and triplet; the decay time of the singlet state is much shorter than that of the triplet state. As shown in Fig. 2.26, the scintillation decay times are different depending on particles, which can be used for a particle identification.

On the other hand, LXe has disadvantages as well. Firstly, LXe is expensive. Secondly, since the wavelength of its scintillation light is in the VUV region, photosensors sensitive to light in this region are required. Thirdly, the detector operation is difficult because LXe must be kept

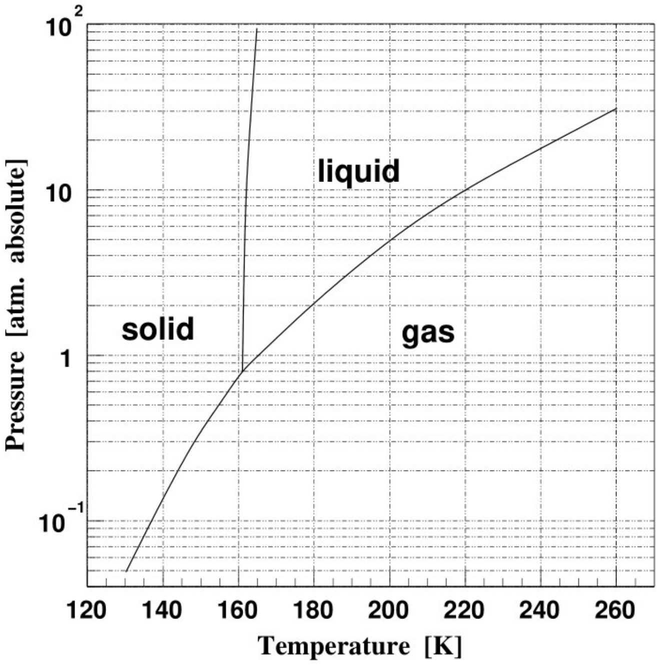


Figure 2.27 Phase diagram of xenon.

free from contaminations in liquid state. Fig. 2.27 shows a xenon phase diagram. To keep liquid state, only a small temperature region is allowed around the atmospheric pressure. Therefore, a precise temperature control is essential. Moreover, impurities in LXe results in a reduced light yield since a oxygen or a water contamination can absorb scintillation light and a nitrogen contamination can quench the xenon scintillation by the following process:

$$\mathrm{Xe}_2^* + \mathrm{N}_2 \rightarrow 2\mathrm{Xe} + \mathrm{N}_2.$$

When a γ-ray is injected into material, three types of reactions occur, i.e. photoelectric absorption, Compton scattering and pair creation. Fig. 2.28 shows the cross section of xenon, which indicates pair creation is dominant process around the energy of the γ-ray from the $\mu \to e\gamma$. Fig. 2.29 illustrates the development of the electromagnetic shower when a signal γ-ray enters the LXe detector; the γ-ray firstly converts, and a positron and an electron are emitted by the pair creation. They fly emitting γ-rays by bremsstrahlung, and the γ-rays travel in the detector depositing their energies by Compton scattering. The cross section of the Compton scattering for the γ-ray of 1 MeV is $0.052\,\mathrm{cm}^2/\mathrm{g}$ resulting in mean free path of 6.5 cm, where most of the generated γ-rays has the energy below 1 MeV. Fig. 2.30 shows the averaged shower profile along with the transverse direction and the longitudinal direction. The shower spread is limited within $\mathcal{O}(1)$ cm, which can be understood by the mean free path.

Photosensors

The photosensors for the LXe detector must be sensitive to the VUV light and operational at the low temperature. A new type of PMT was developed for the MEG and that of SiPM for the MEG II which fulfill these requirements. Hereafter, the properties of the photosensors are

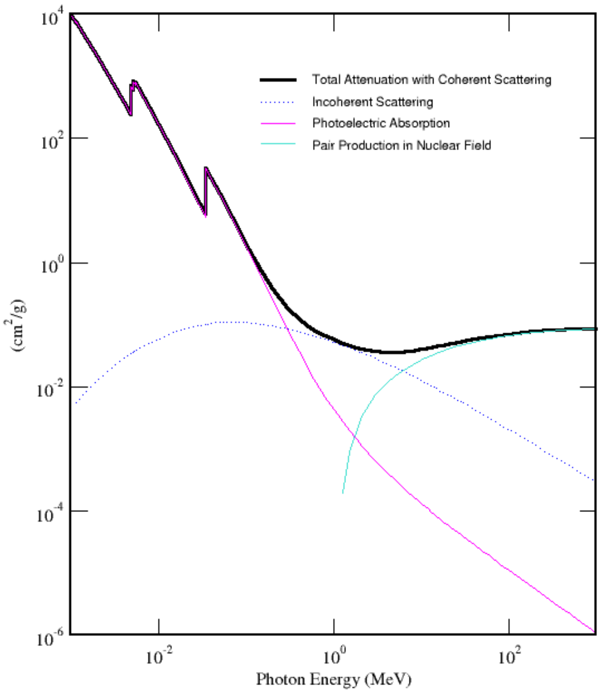


Figure 2.28 Photon cross section of xenon [57].

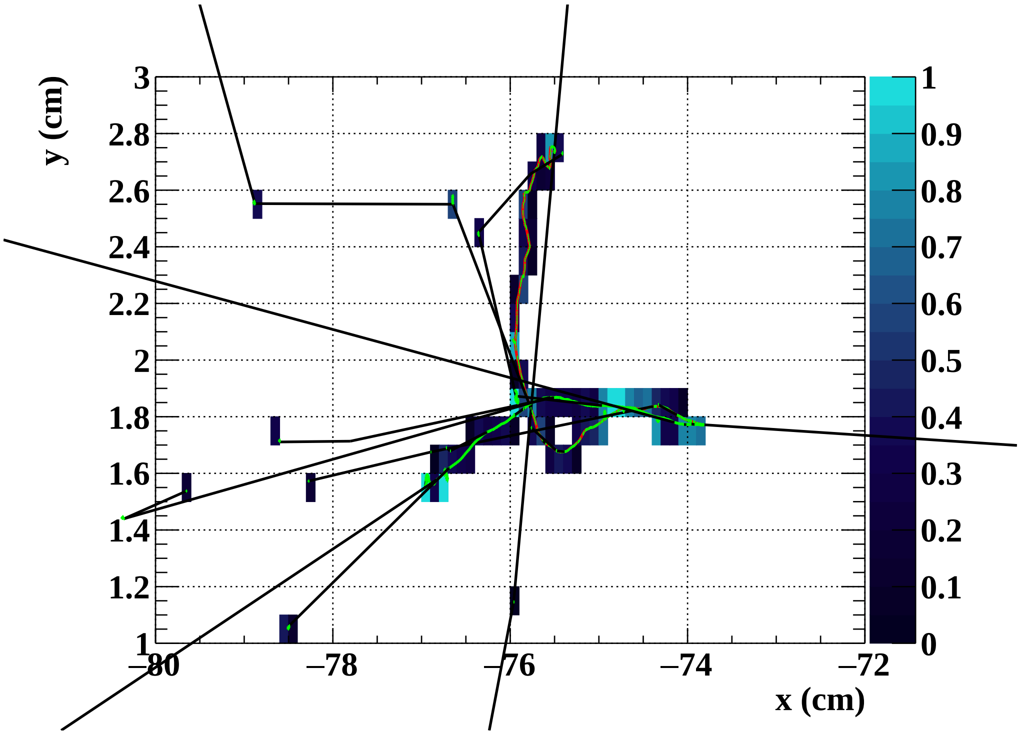
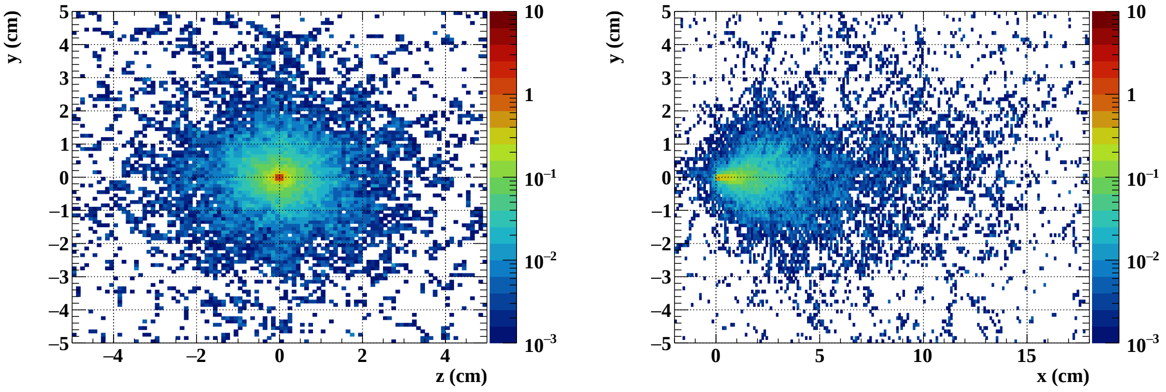


Figure 2.29 Development of electromagnetic shower in LXe simulated with GEANT4 (version 10.4 [58]). Particle tracks of γ-rays, positrons and electrons are shown as lines in black, red and green, respectively. The original signal γ-ray from a muon comes from the right. The energy deposit on each point is overlaid. The scale of the bar is in the unit of MeV.



(a) (b)

Figure 2.30 Shower profile from a signal γ-ray along with (a) the transverse direction and (b) the longitudinal direction simulated with GEANT4 (version 10.4 [58]). Mean of energy deposits in the area of 0.1 cm square over $\mathcal{O}(100)$ events is shown, where the first conversion points of each event are aligned to be the origin. The scale of the bar is in the unit of MeV.

Figure 2.31 VUV-sensitive PMT (R9869) [59].

discussed.

PMT

A new PMT which satisfies the requirements was developed in collaboration with Hamamatsu Photonics (Fig. 2.31). Table 2.4 summarizes the properties of the PMT (R9869). For the sensitivity to the VUV light, Bialkali (K-Cs-Sb) is used as a photo-cathode and a quarts window, which is 80% transparent for $\lambda \sim 178\,\text{nm}$, is adopted. For the operation at the low temperature, the photo-cathode has aluminum strips to avoid the increase of the sheet resistance. In order to stabilize the gain at high-rate environment, Zener diodes are inserted in parallel to the last two stages of the resistive divider for twelve-step dynodes and keep a voltage around 85 V (Fig. 2.32).

In total, 668 PMTs are installed in the detector: 234 on the outer face, 144 on each of the upstream and the downstream faces, and 73 on each of the top and the bottom faces, respectively.

Table 2.4 Properties of VUV-sensitive PMT (R9869).

Size	57 mm ϕ
Active Area	45 mm ϕ
PMT length	32 mm ϕ
Photo-cathode material	K-Cs-Sb
Dynode type	Metal channel
Number of dynode	12
Typical HV	900 V
Typical gain	1×10^6
Typical QE	15%
Rise time	2 ns
Typical Transit time	12.5 ns
Typical Transit Time Spread	0.75 ns

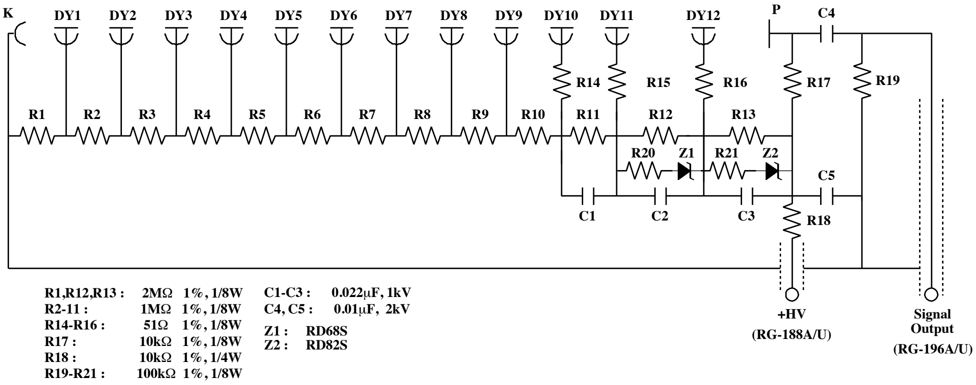


Figure 2.32 Bleeder circuit of PMT (R9869) [60].

VUV-sensitive MPPC

Since standard SiPMs are not sensitive to VUV light, a VUV-sensitive MPPC (Multi-Pixel Photon Counter, S10943-4372) was developed[*2] for the LXe detector of the MEG II experiment (Fig. 2.33). One MPPC package has a size of $15 \times 15\,\mathrm{mm}^2$. It consists of distinct four bare MPPC chips with the active area of $5.95 \times 5.85\,\mathrm{mm}^2$ for each. Each chip contains 13925 micro-cells with pixel pitch of $50\,\mu$m. The four chips are glued on a ceramic base with conductive adhesive.

A high sensitivity to VUV light is achieved by optimizing the design: removing the protection coating, optimizing the optical matching between LXe and the chip surface, and thinning down the top contact layer. In the standard SiPMs, there is a protection layer of epoxy resin at the surface of the sensor. This layer is removed since it disturbs VUV photons to enter the sensitive region due to the absorption, and a 0.5 mm thick VUV-transparent quartz window is mounted

---

[*2] MPPC is the product name of SiPM device produced by Hamamatsu Photonics.

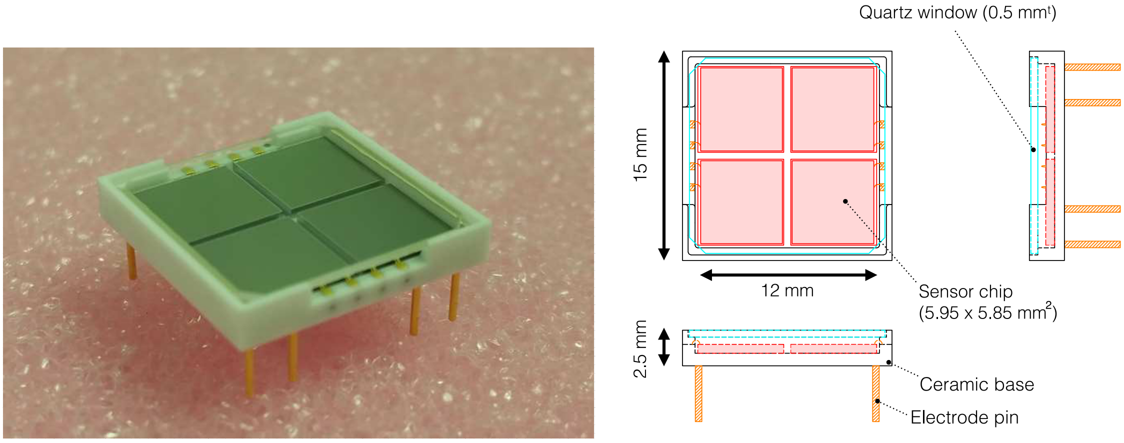
on the package instead. There is a thin gap of 0.5 mm between the window and the chips. It is filled with LXe when the MPPC is operated in the LXe detector. The refractive index of the quartz, 1.60, is close to that of LXe, 1.65, which reduces the reflection at the boundary. In addition to the removement of the protection coating, the top contact later of the chip is thinned down to reduce absorption. Even with these optimizations, VUV light can be absorbed in the top contact later since the attenuation length of VUV light in silicon is only 5 nm, which is still shorter than the thickness of the layer. Nevertheless, VUV light can be detected in the way that a part of charge carriers generated in the contact later drifts to the active layer and triggers an avalanche. To enhance this process, further improvements are applied. The dopant concentration in the contact layer is adjusted to have a low electric field, which helps the carriers to drift to the active layer. The processing of the contact layer is optimized to minimize the lattice defects not to trap the carriers.

To suppress the number of readout channels, an enlargement of a sensitive area is also required since the sensitive areas of the standard SiPMs are typically $3 \times 3\,\mathrm{mm}^2$. This is realized by an enlargement of the MPPC chip and the design that the MPPC package includes four chips. In original, the four chips can be read out individually, but they are connected inside a PCB, which results in the active area of $12 \times 12\,\mathrm{mm}^2$ per channel. There are, however, two issues caused by the enlargement of the chip: an increase of dark noise and a longer decay time due to a larger capacitance. The first issue is not a problem for the LXe detector since it is caused by a thermal excitation, and thus it can be greatly suppressed at LXe temperature. The second issue can be solved by connecting them in series, which reduces the capacitance (Fig. 2.34). Thanks to the large sensitive area, the inner face of the LXe detector of $0.92\,\mathrm{m}^2$ is tiled with 4092 MPPCs with a total active area of $0.57\,\mathrm{m}^2$.

The performance of the VUV-sensitive MPPC was measured in LXe and proved to have an excellent performance such as a high photon detection efficiency, a high gain, a low probability of the optical crosstalk and the after-pulsing and a low dark count rate as summarized in Table 2.5.

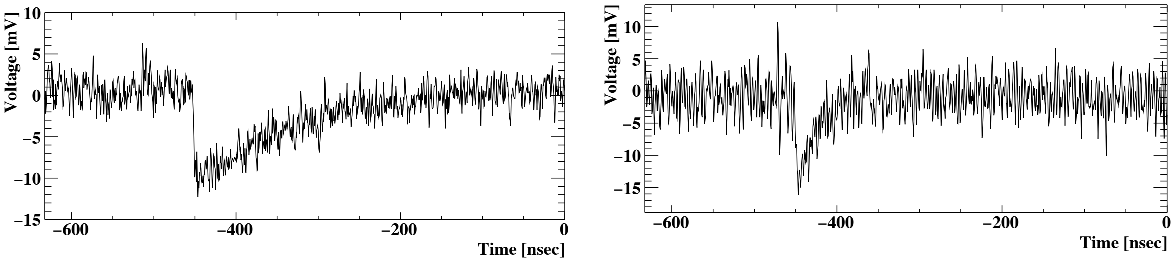
The mass production of 4200 MPPCs including spares was performed by Hamamatsu Photonics, and they were produced in four batches called lot A, B, C and D. Before the installation, all MPPCs were tested at room temperature. I-V curves were obtained to identify bad MPPCs, and some MPPCs were removed due to strange current values. The waveforms were also measured with a pulsed LED light. Fig. 2.35 shows the measured decay time of each MPPC. The production lot dependence was observed due to different after-pulsing probabilities. This kind of lot dependence is also observed in the LXe detector, and so it is corrected for the reconstruction.

The MPPCs are mounted on PCB strips as shown in Fig. 2.36. Each PCB strip has 22 MPPCs, and two PCBs are installed in a line along the z-axis with 93 lines covering the inner face of the detector. The electrode pins of the MPPCs are inserted into sockets on the PCBs (Fig. 2.38(a)). The four MPPC chips in each package are connected in series inside a circuit of the PCB. Instead of a simple series connection, the chips are connected with so-called hybrid connection shown in Fig. 2.37. There are a decoupling capacitor and resistors in-between, which enable to connect the signal line in series and the bias line in parallel. Thanks to this connection



(a) (b)

Figure 2.33 Overview of VUV-sensitive MPPC developed for MEG II (S10943-4372, Hamamatsu Photonics) [61].



(a) (b)

Figure 2.34 Typical waveforms of the MPPC with two photoelectrons, i.e. primary single photoelectron signal from LED overlapped with optical crosstalk. The four chips in the MPPC are connected in (a) parallel and (b) series [61].

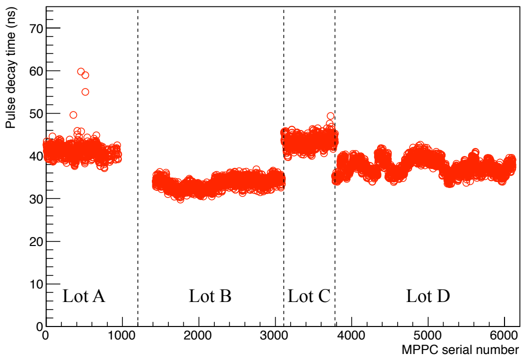


Figure 2.35 Pulse decay time of each MPPC as a function of serial number [61].

Table 2.5 Properties of VUV-sensitive MPPC (S10943-4372) [61].

Size	$15 \times 15\,\mathrm{mm}^2$
Effective photosensitive area	$12 \times 12\,\mathrm{mm}^2$
Pixel pitch	$50\,\mu\mathrm{m}$
Typical HV	55 V per chip
Typical gain	5×10^5
Typical PDE	15%
Typical dark count rate	$5\,\mathrm{Hz/mm}^2$
Correlated noise probability	$< 30\%$

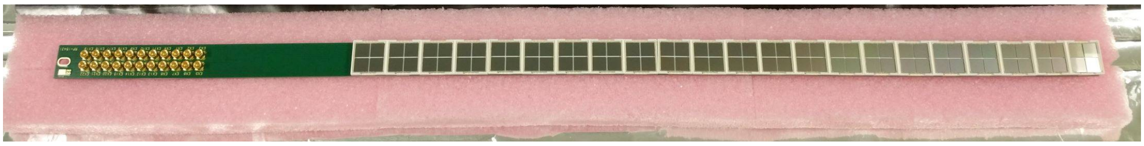


Figure 2.36 PCB for the MPPCs.

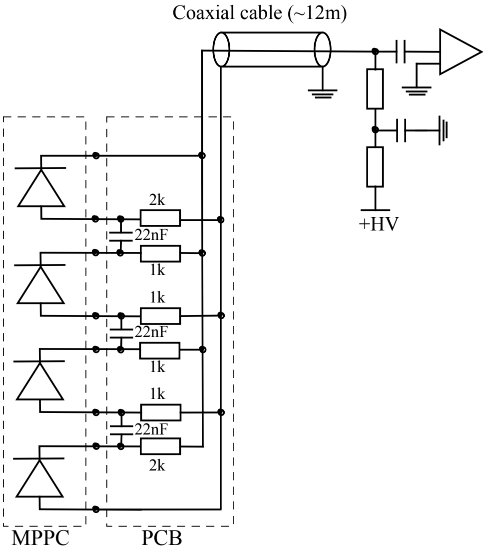
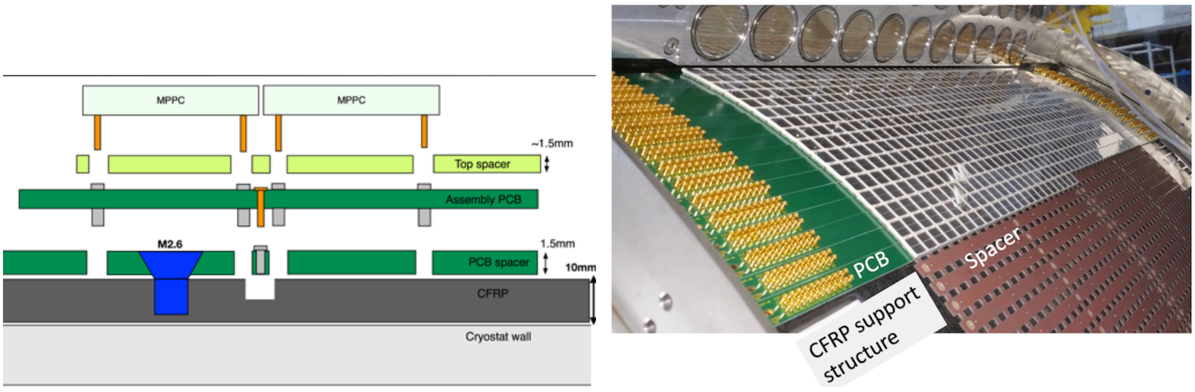


Figure 2.37 Hybrid connection used for the MPPC [61].

method, the applied high voltage is smaller than the case of a simple series connection. To avoid electric noise, a coaxial-like signal line structure is implemented in the PCBs.

The PCBs are aligned on structures made of Carbon Fiber Reinforced Plastic (CFRP), and they are fixed on the cryostat as shown in Fig. 2.38(b). The PCBs are mounted on four separate CFRP structures. The temperature coefficient of the CFRP is adjusted to be consistent with that of the PCBs to meet the thermal contraction at LXe temperature. A spacer made of FR4 is inserted in the gap between the MPPCs and the PCBs to prevent LXe from filling the gap, and another spacer is inserted between the PCB and the CFRP since the CFRP is conductive.



(a) (b)

Figure 2.38 MPPC support structure [62].

Xenon handling system

Fig. 2.39 shows a whole xenon control system. Besides the LXe detector cryostat, auxiliary components are connected to control, store and purify xenon.

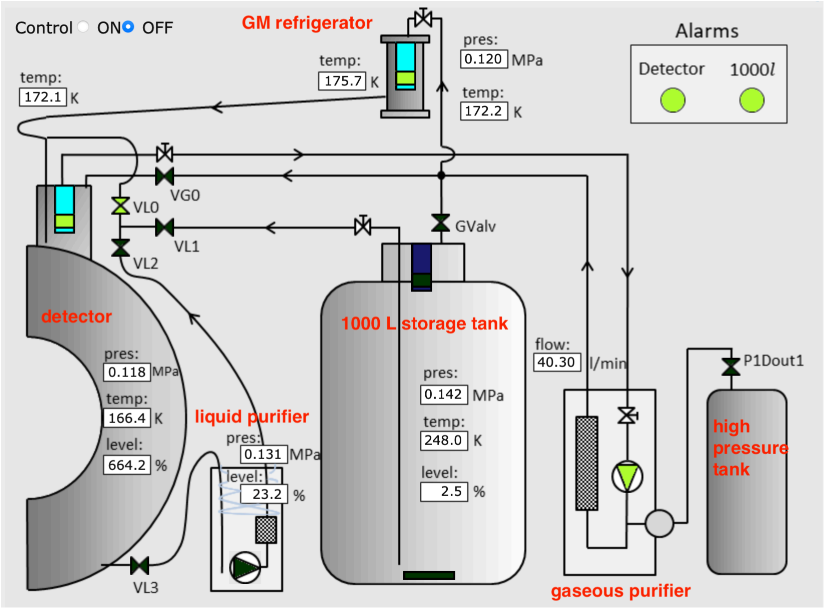
Since xenon has a narrow liquid state range at the operation pressure, the temperature and the pressure of LXe must be well controlled for the safe and stable operation. A 200 W pulse-tube refrigerator [63] is mounted on the top chimney of the detector, and a heater is installed in the detector. A 400 W Gifford-McMahon (GM) refrigerator [64] is newly introduced to the MEG II to compensate the heat from the increased number of cables. Cooling pipes with liquid nitrogen attached to the outside of the inner vessel also contribute to cooling. The temperature is kept with a precision of 0.1 K, and the pressure is within 0.8 kPa.

There are two types of xenon storage. The 1000 L storage tank stores LXe during short-term maintenance of the LXe detector. The pressure and temperature can be controlled with an independent liquid nitrogen cooling system. The high pressure tanks keep gaseous xenon during long-term detector maintenance. The storage consists of eight tanks with 250 L capacity and tolerance to 8 MPa pressure.

To keep the high purity of xenon, purification systems in liquid phase and gaseous phase are used. The liquid purifier filters water with molecular sieves. The gaseous purifier can absorb most kinds of molecules except for rare gas with a metal-heated getter.

2.1.3.3 Background detector

A radiative decay counter (RDC) is newly installed in the MEG II experiment to reduce γ-ray backgrounds. As described in Sec. 1.2.2, the dominant background is the accidental coincidence between a positron from Michel decay and a γ-ray from AIF and RMD. Thanks to the reduction of the CDCH material in MEG II, γ-rays deriving from AIF is expected to be suppressed compared to the MEG experiment. On the other hand, the number of RMD backgrounds does not change, and the fraction of background γ-rays from RMD can be 65% for



Control ○ ON ● OFF

temp: 172.1 K

temp: 175.7 K

pres: 0.120 MPa

temp: 172.2 K

detector pres: 0.118 MPa

temp: 166.4 K

level: 664.2 %

liquid purifier pres: 0.131 MPa

level: 23.2 %

1000 L storage tank pres: 0.142 MPa

temp: 248.0 K

level: 2.5 %

flow: 40.30 l/min

Figure 2.39 Xenon control system in MEG II.

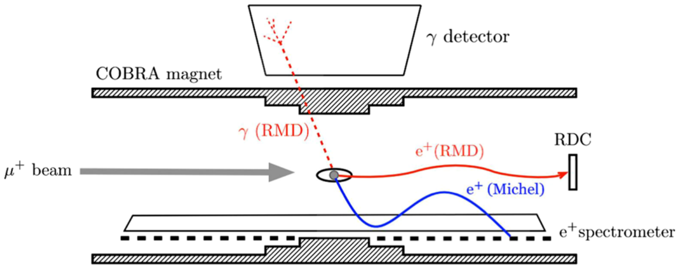


Figure 2.40 Schematic view of the RMD detection with the downstream RDC [2].

$E_\gamma > 48\,\mathrm{MeV}$.

Fig. 2.40 describes a concept of the RMD detection with the RDC. Since the RMD events with high energy γ-rays emit low energy positrons simultaneously, those events can be identified by detecting the positrons. The low energy positrons follow an almost helical trajectories with small radii along the beamline due to the COBRA magnetic field, and thus the RDC is placed on the beam axis. The RDC can be installed both upstream and downstream of the target, where 52% and 48% of the positrons are emitted, respectively.

The RMD events can be identified by taking the time coincidence between the RDC and the LXe detector. The red histogram in Fig. 2.41 shows the simulated distribution of the time differences between the detectors for accidental background events while the blue one is for $\mu^+ \to e^+\gamma$ signal events. The peak in the red histogram corresponds to the RMD events. There are also flat regions in both histograms, which derive from background Michel positrons. They can be distinguished, however, by measuring their energies since the Michel positrons have relatively high energies as shown in Fig. 2.42.

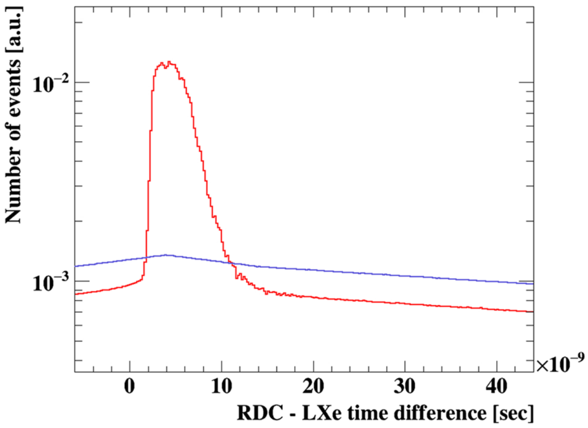


Figure 2.41 Simulated time differences between the RDC and the LXe detector for accidental background events (red) and $\mu^+ \to e^+\gamma$ signal events (blue) [2].

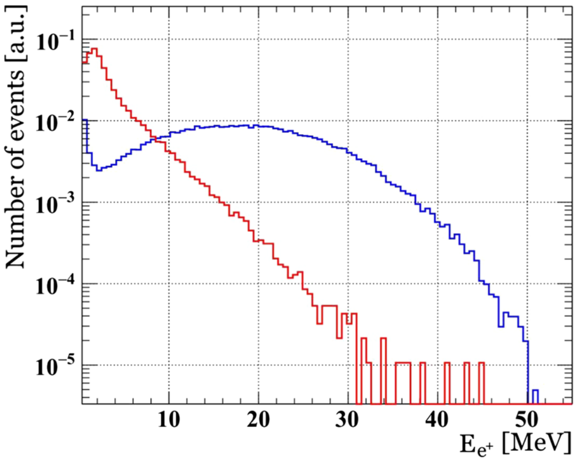


Figure 2.42 Simulated energy distributions of RMD for $E_\gamma \geq 48\,\mathrm{MeV}$ (red) and Michel (blue) positrons detected by the RDC [2].

Downstream RDC

The RDC installed in the downstream side consists of a timing counter and a calorimeter as shown in Fig. 2.43. Its geometrical acceptance to the RMD positrons emitted to downstream side is 88%. Since the Michel positrons hit the RDC at the frequency of about 10 MHz, both the timing counter and the calorimeter are finely segmented: 12 plastic scintillator bars for the timing counter and 76 LYSO (Cerium-doped Lutetium Yttrium Orthosilicate) crystals for the calorimeter.

Hereafter, the designs of the timing counter and the calorimeter are described. The measurements in lab for the performance evaluation are explained in Appendix A.

Timing counter

The timing counter is composed of 12 fast plastic scintillators (BC-418, Saint-Gobain) with different sizes as shown in Fig. 2.44. The scintillators have characteristics for a high timing

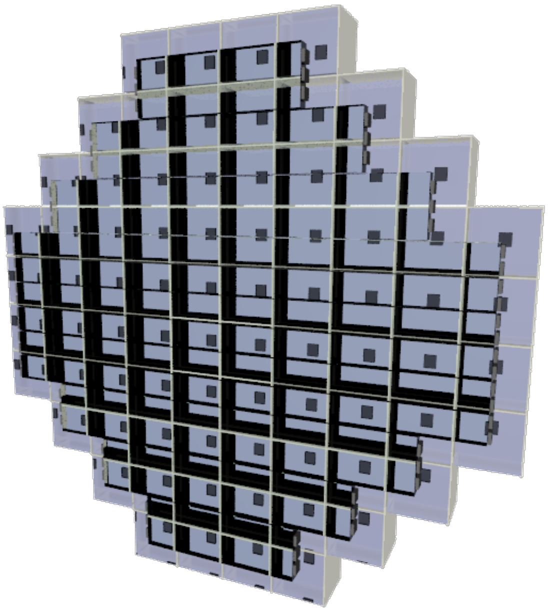


Figure 2.43 Design of the downstream RDC.

Table 2.6 Properties of BC-418 (Saint-Gobain) [66].

Light Output, %Anthracene	67
Rise Time	0.5 ns
Decay Time	1.4 ns
Wavelength of Max. Emission	391 ns

resolution with a high light yield and fast rise time as described in Table 2.6. The thicknesses of all the scintillators are 5 mm. The widths of the six scintillators at the center are 1 cm because of a higher hit rate while those for the others are 2 cm. The lengths are 7, 11, 15 and 19 cm in this order from an outer side. The scintillation light is read out with SiPMs (S13360-3050PE, Hamamatsu Photonics) at both ends of each scintillator. Two or three SiPMs are used for scintillators with 1 cm or 2 cm widths, respectively. The SiPMs have a high detection efficiency at the wavelength of the scintillation light (Table 2.7). The SiPMs on each face are connected in series to reduce the number of readout channels. The series connection is superior to the parallel connection for the timing counter since a rise time of the waveform is shorter as the capacitance of the sensor gets smaller. They are mounted on PCBs by soldering. Six SiPMs on the central three scintillators are mounted on a single PCB together due to the limited space while others are mounted every scintillator (Fig. 2.45). The SiPMs are coupled to the scintillators with optical cement. The scintillators are wrapped with reflectors of aluminized mylar, and then shielded with black sheets from Tedlar (Fig. 2.46). The timing resolution was measured to be better than 90 ps with a ^{90}Sr source, which is good enough for the RMD detection [65].

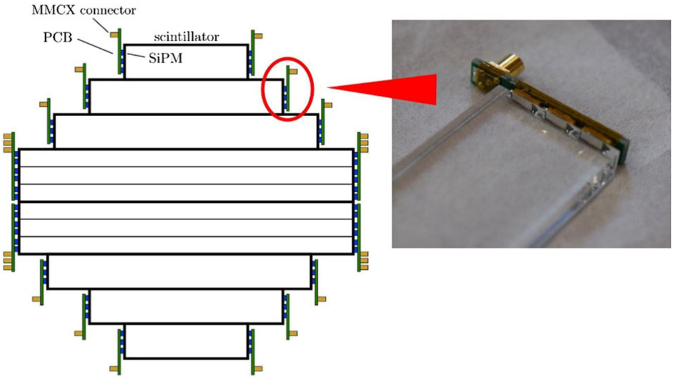


Figure 2.44 RDC timing counter [2].

Table 2.7 Properties of SiPM (S13360-3050PE) [67].

Pixel pitch	50 μm
Effective photosensitive area	$3.0 \times 3.0\,\mathrm{mm}^2$
Number of pixels	3600
Fill factor	74%
Peak sensitive wave length	450 nm

Figure 2.45 PCB for scintillators with a 2 cm width (left) and a 1 cm width (right).

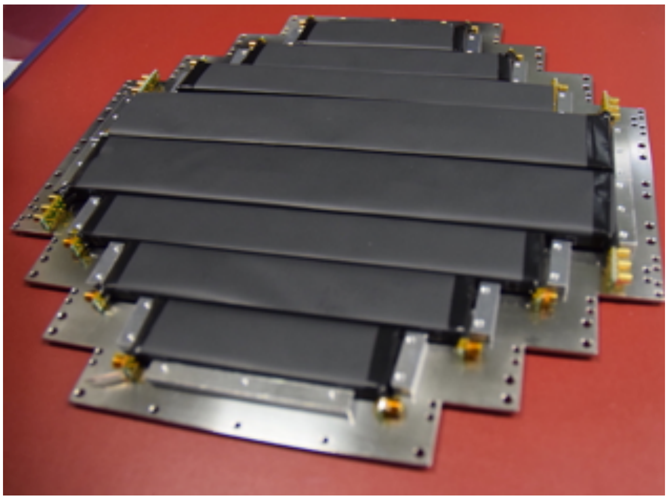


Figure 2.46 RDC timing counter after shielding [2].

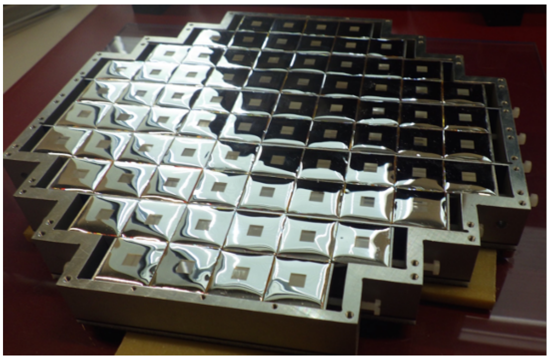


Figure 2.47 RDC calorimeter.

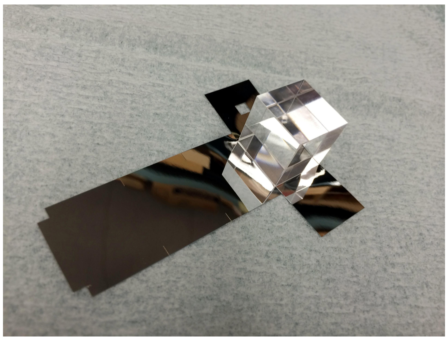


Figure 2.48 LYSO crystal. Each crystal is wrapped with the reflector.

Calorimeter

The calorimeter is composed of 76 LYSO crystals (Shanghai Institute of Ceramics) in the shape of a 2 cm cube as shown in Fig. 2.47. Each crystal is wrapped with a 65 μm thick reflector (ESR, 3M) [68]. Table 2.8 summarizes the properties of the LYSO crystal. The LYSO crystal has a high light yield and a short decay time constant, which are suitable for the measurement of energy in a high rate environment. LYSO contains the radio isotope ^{176}Lu (2.59% natural abundance), which decays to ^{176}Hf with emission of a β^-, followed by a cascade of 307, 202 and 88 keV γ-rays (Fig. 2.49). The decay rate per crystal was measured to be small enough ($\sim$ 2 kHz) not to affect the RMD positron detection. The energy spectrum has a peak at 596 keV as shown in Fig. 2.50, which can be used for the energy calibration.

Each crystal is read out with a single SiPM (S12572-25P, Hamamatsu Photonics). SiPMs with many pixels are desirable in order to avoid the saturation of the pixels due to the high light yield of the crystals. On the other hand, the pixel size of such SiPMs is small, and thus they have small capacitance. This results in a worse energy resolution due to a low gain. SiPMs with 15, 25 or 50 μm pitch were compared in terms of the energy resolution and linearity for incident scintillation light, and those of 25 μm pitch were selected [65]. In addition, the effect of the saturation can be suppressed to some extent by reducing the operation voltage resulting in the low after-pulsing and crosstalk probability. The SiPMs are mounted on PCBs with soldering. As shown in Fig. 2.51, the PCBs have flexible branches where the SiPMs are placed. The SiPMs are pressed against each LYSO crystal with springs and coupled using optical grease. This removable attachment enables a replacement of the SiPMs or the crystals for maintenance. The energy resolutions of most crystals were measured to be $\sim$ 6% at 1 MeV with a ^{60}Co source, which is good enough for the background separation [69].

Support structures

The support structures for the RDC are constructed with non-magnetic materials (Fig. 2.52). To fix the timing counter, a stainless steel plate is placed in the front of the timing counter. The center of the timing counter is not covered to minimize the amount of material. In order

Table 2.8 Properties of the LYSO crystal [70].

Density	7.4 g/cm^3
Radiation length	1.14 cm
Decay constant	42 ns
Emission peak	420 ns
Light yield	30000 p.h./MeV

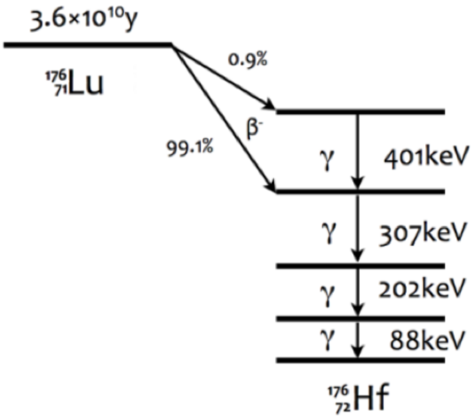


Figure 2.49 Decay scheme of ^{176}Lu.

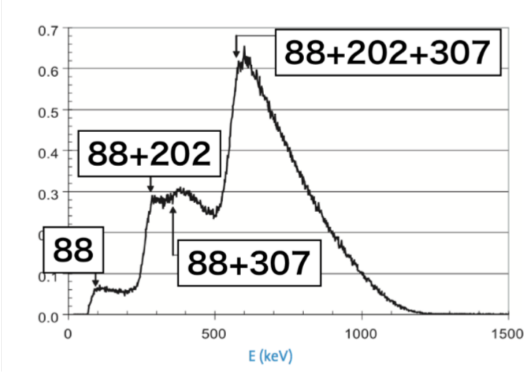


Figure 2.50 Energy spectrum of the intrinsic radioactivity of LYSO [71].

Table 2.9 Properties of S12572-25P [72].

Pixel pitch	25 μm
Effective photosensitive area	3.0×3.0 mm^2
Number of pixels	14400
Fill factor	65%

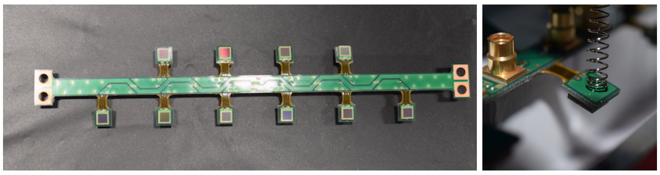


Figure 2.51 PCB for the RDC calorimeter. The SiPMs are mounted on the each branch (left). The SiPMs are pressed from the backside with springs (right).

to absorb the stress of the springs ($\sim$ 2.5 kg in total) with the minimum amount of material, a 2 mm thick Rohacell plate sandwiched with CFRP layers with a thickness of 0.2 mm each is inserted between the timing counter and the calorimeter. In addition, a 0.1 mm thin aluminum plate is inserted behind the Rohacell plate (Fig. 2.53) and the side is covered with an aluminum frame for light shielding of the calorimeter (Fig. 2.55). Two 0.2 mm plates made of Delrin® and one CFRP plate are placed behind the calorimeter (Fig.2.54). The springs are fixed on a chassis made of brass (Fig. 2.56). The cables are fixed on the chassis not to touch the springs (Fig. 2.57).

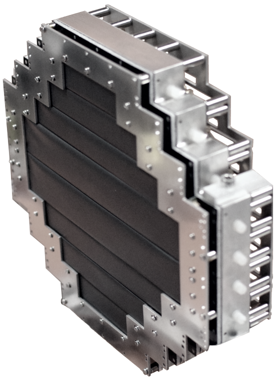


Figure 2.52 Outlook of the RDC.

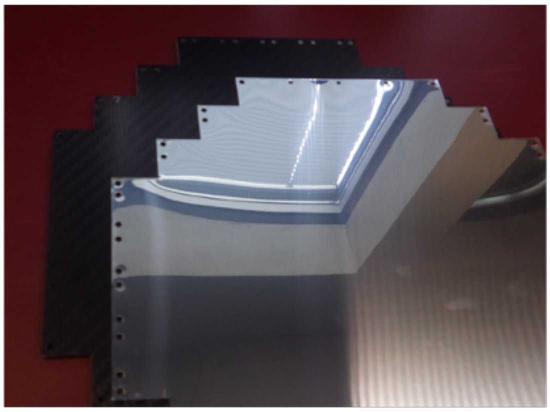


Figure 2.53 Rohacell plate and aluminum plate inserted between the timing counter and the calorimeter.

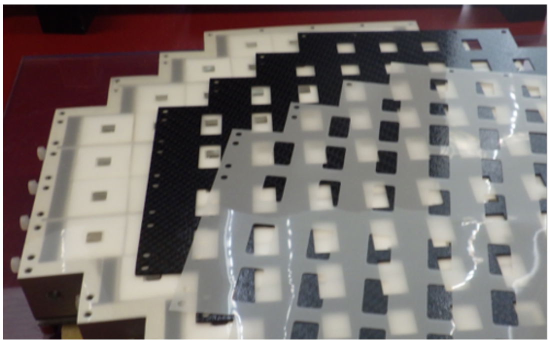


Figure 2.54 Delrin and CFRP plates inserted behind the RDC calorimeter.

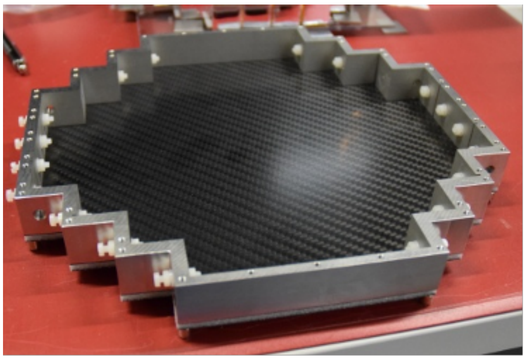


Figure 2.55 Aluminum frame for the RDC calorimeter.

The backside of the calorimeter is covered with an aluminum holder for light shielding (Fig. 2.58).

The RDC is mounted on a moving arm system attached to the end-cap of the COBRA magnet as shown in Fig. 2.59. The moving arm enables the RDC to be moved away from the beam-axis when a calibration target for the LXe detector is inserted from the downstream side (Fig. 2.60). It can be controlled remotely by water pistons made of plastic. The shaft is made of titanium to work under heavy loads and other parts are made of aluminum. The position of the detector can be monitored with two end-switches. Since the end-cap of the COBRA separates the inner volume from the outside, SiPM signals are transmitted via feed-through PCBs attached to the end-cap. The design of the PCB is the same as that used for the LXe detector, which is described in Sec. 2.1.3.2.

Control system

Two thermometers (PT100) are installed in the RDC; one is near the SiPM of LYSO crystal around the center and the other is placed outside the main frame. Signals from the sensors are read out via the feed-through PCB.

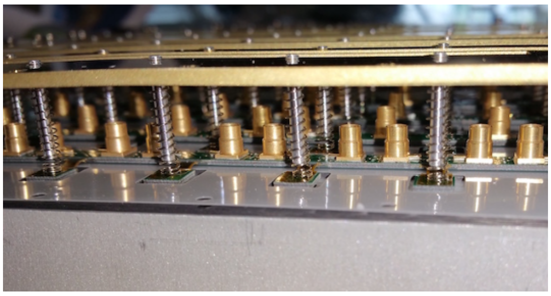


Figure 2.56 Springs to fix the SiPMs for the RDC calorimeter.

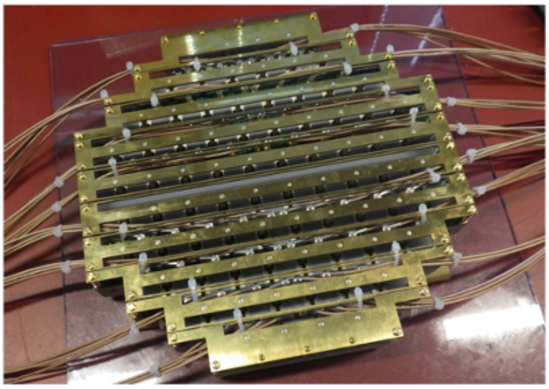


Figure 2.57 Brass chassis to fix the springs.

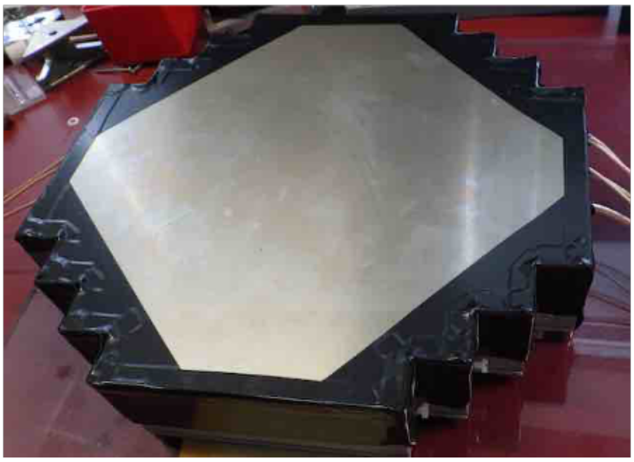
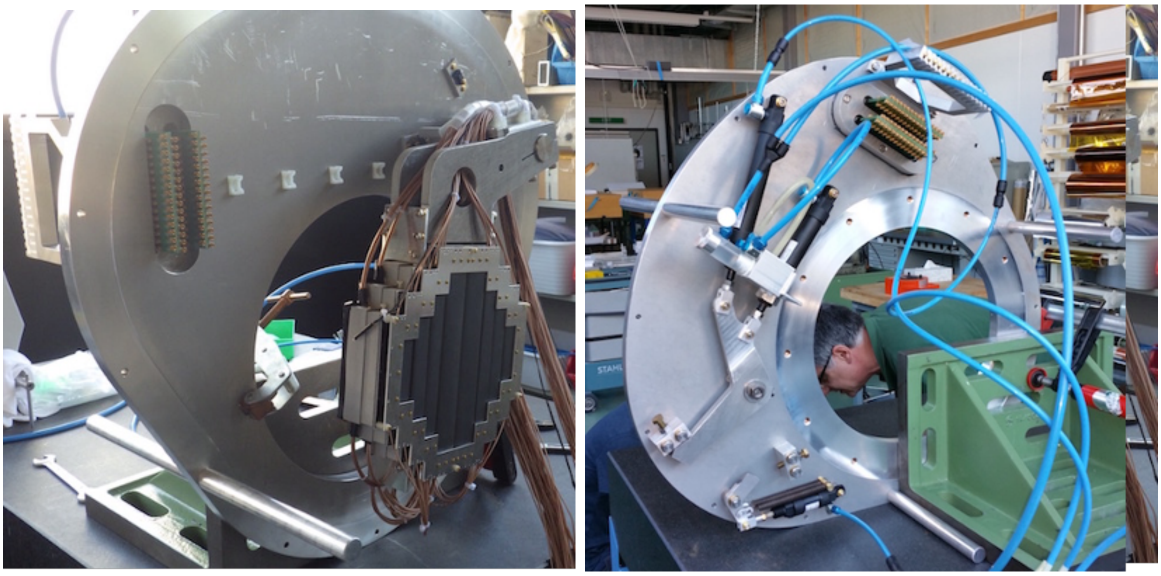


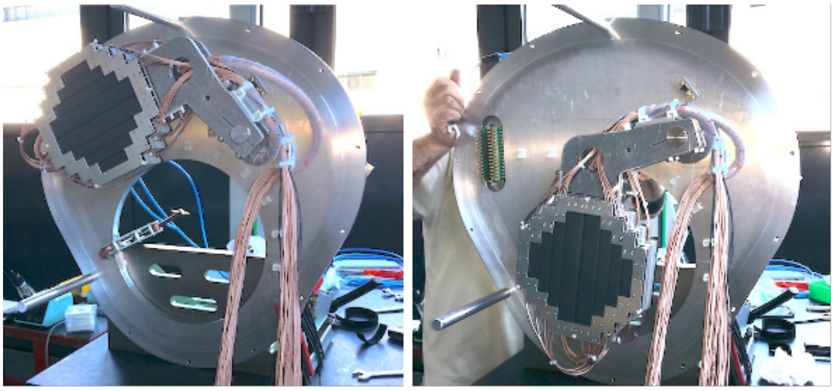
Figure 2.58 Aluminum holder for the light shielding of the RDC calorimeter.



(a)

(b)

Figure 2.59 RDC mounted on a moving arm viewed from (a) front and (b) back sides. Compressed air is provided through the blue tube.



(a) (b)

Figure 2.60 RDC at (a) parking position and (b) measurement position.

The moving arm is controlled by pneumatic system. When pressurized gas of 7–8 bar is supplied to the pipes connected to the water piston, the RDC is inserted to or moved out from the beam-axis. Compressed air is supplied as the pressurized gas source and its flow is controlled with remote valves.

The two end-switches are the physical switches installed to monitor the position of the RDC. One of them is placed to be on when the RDC is at the measurement position and the other is at the parking position. Current is supplied and read out via the feed-through PCB. When the switches are not pressed, the electric path is disconnected.

Upstream RDC

The RDC for the upstream is under development. The requirement for the detector is more stringent than the downstream since the μ beam needs to pass the upstream RDC. Firstly, the material thickness must be small enough not to affect the beam, and thus a calorimeter cannot be placed. Secondly, the detector must be finely segmented and have a fast response to distinguish the RMD positrons from beam muons. Thirdly, radiation hardness is required for the operation under the muon beam.

Various detectors have been tested such as scintillating fibers, a diamond detector and a silicon detector, but they cannot satisfy the requirements. A resistive plate chamber, a type of gaseous detectors, has been investigated as a candidate, which will be described in Appendix B.

Since this thesis focuses on the downstream RDC and the performance of the upstream RDC has not been fixed yet, only the downstream RDC is taken into account hereafter. The upstream RDC will be installed after the development, and a further improvement of the sensitivity is expected.

2.1.4 Trigger and data acquisition

The DAQ system is also upgraded for the MEG II experiment. In the MEG experiment, the detector signals were actively split and then sent to the trigger and DAQ systems as shown in

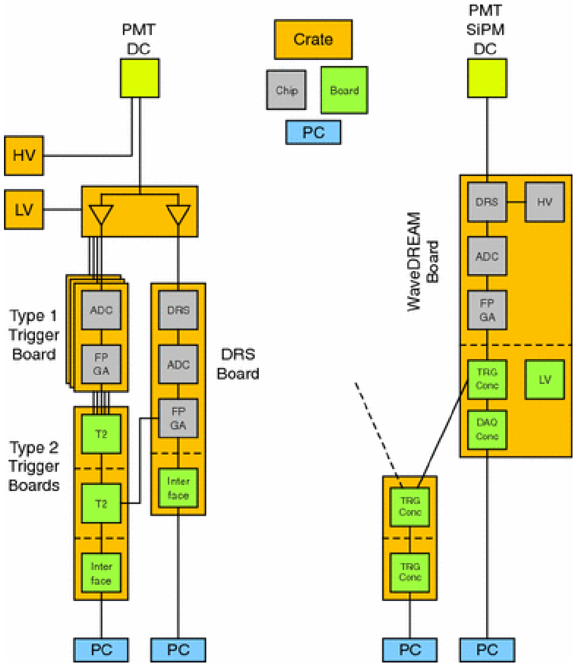


Figure 2.61 Designs of electronics for the MEG (left) and the MEG Ⅱ (right) [74].

Fig. 2.61. However, due to the limited space for the electronics in the experimental area, adopting such a scheme is impossible with the increased number of channels in the MEG Ⅱ experiment by a factor of three. Therefore, a new simplified and compact system called WaveDAQ was developed which integrates the trigger and DAQ functionalities onto the same electronics board, the WaveDREAM board (WDB).

The WaveDAQ crate contains the crate management board (CMB), 16 slots for WDBs and two slots for so-called concentrator boards, the trigger concentrator board (TCB) and the data concentrator board (DCB) as shown in Fig. 2.62. The CMB supplies power to all the boards in the crate and manage the slow control system. The TCB generates a trigger. The DCB configures all boards inside the crate, distributes the master clock and trigger signals, and reads out waveform data. The whole of the system is controlled by Maximum Integration Data Acquisition System (MIDAS) [73].

WaveDREAM boards

The WDB has multiple functions such as a signal amplification and shaping, a HV supply and a signal digitization. Fig. 2.63 shows the schematic view of the WDB. Two DRS4 (Domino Ring Sampling) chips [74] are connected to two 8-channel ADCs, which are read out by a Field-Programmable Gate Array (FPGA). The output stream of the ADCs is used in the FPGA to perform complex trigger algorithms.

The DRS4 chip digitizes input waveforms with a maximum sampling speed of 5 GSPS. In the pilot runs, the sampling frequency of 1.2 GHz was used for the LXe detector and the RDC instead of 1.4 GHz planned in the physics data-taking. This is because of the trigger latency

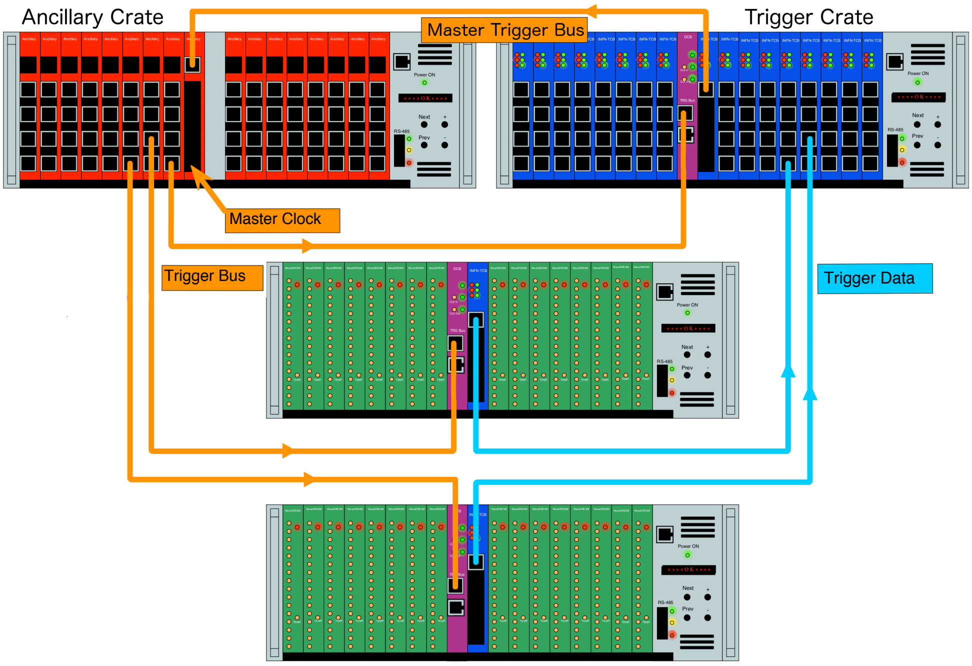


Figure 2.62 Sketch of WaveDAQ crate with WDBs (green), DCBs (violet), TCBs (blue) and Ancillary board (red) [75].

of ∼ 680 ns, which is reduced by replacing the ADC for the final system [75]. Fig. 2.64 shows a readout mechanism of the DRS chip. The amplitudes of the waveforms are stored on the 1024 capacitors called DRS cells. A logic wave, so called Domino wave, opens analogue switches at the input of the DRS cells of each channel. An on-chip phase locked loop (PLL) locks the generation of the Domino wave to an external reference clock. The sampling is stopped when a trigger is fired, and then the stored charges are read out by the ADCs in the order of the DRS cells via a shift register. One DRS chip includes eight readout channels and one clock channel as shown in Fig. 2.65.

A reference clock of 80 MHz sine wave is distributed to each clock channel on DRS chips in order to synchronize channels, which enables the precise timing alignment with offline analysis. A sine wave is fitted to each clock waveform, and the timings of the DRS chips are aligned based on zero-crossing times extracted by the fitting. The precision of the synchronization between the chips was evaluated as follows; pulses were input into channels in different chips which were divided from a single signal generated with a function generator. The pulse timings were extracted by waveform analysis. The time jitter between the chips was given by the standard deviation of their timing difference, which was 45 ps.

The WDB provides variable amplification and flexible shaping of a waveform. Two stage amplifiers and a programmable attenuator are implemented on the analog frontend of the WDB, which enable to select the gain from 0.5 to 100. The shaping is available with a programmable pole-zero cancelation. In the pilot runs, an electronics amplifier of 2.5 was adopted for the photosensors of the LXe detector. The signals of the RDC timing counter were amplified by a factor of 100 and shaped with pole-zero cancellation while neither amplification nor shaping was

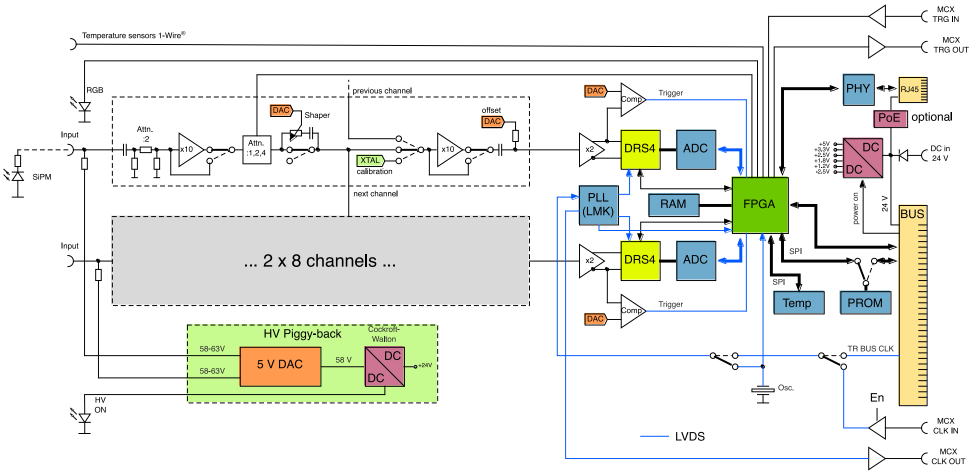


Figure 2.63 Schematic view of WaveDREAM board [2].

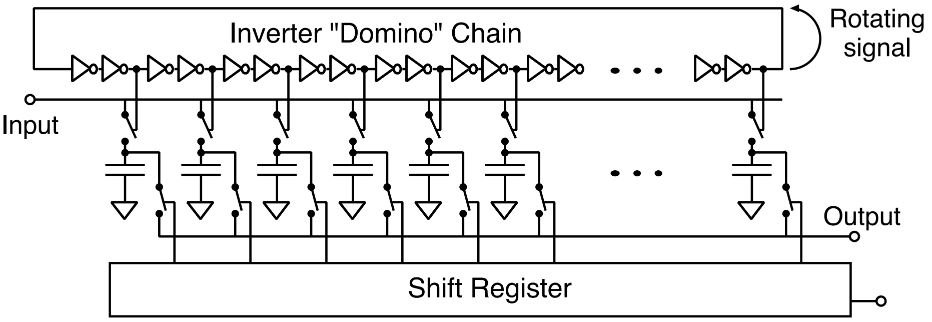


Figure 2.64 Mechanism of DRS4 [74].

applied to the RDC calorimeter.

The voltages for sensors up to $\sim 240\,\mathrm{V}$ can be supplied with a Cockcroft–Walton (CW) voltage multiplier on the WDB. Channel by channel tuning of the applied voltages is possible thanks to a 5 V digital-to-analog converter (DAC), which is mounted after the CW voltage multiplier. The currents are measured with a precision of 1 nA.

Trigger

The trigger in the MEG II experiment is required to collect $\mu^+ \to \mathrm{e}^+\gamma$-like events suppressing the backgrounds by almost six orders of magnitude for the effective DAQ and the reduction of the overall data size. The information from the pTC and the LXe detector is used for the on-line trigger, while the signals from CDCH is not used due to its slow response. Hence, the observables of E_γ, $t_{\mathrm{e}^+\gamma}$ and $\Theta_{\mathrm{e}^+\gamma}$ are reconstructed on-line on the FPGA.

E_γ is reconstructed by summing up weighted waveforms over MPPCs and PMTs in the LXe detector. The weights are defined to compensate the difference of the waveform shape between the MPPCs and the PMTs, and the photosensor performance such as the gain and the photon detection efficiency. Events whose amplitudes of the summed waveforms exceed a certain threshold are selected as signal-like events. A veto threshold is also available to reject high energy

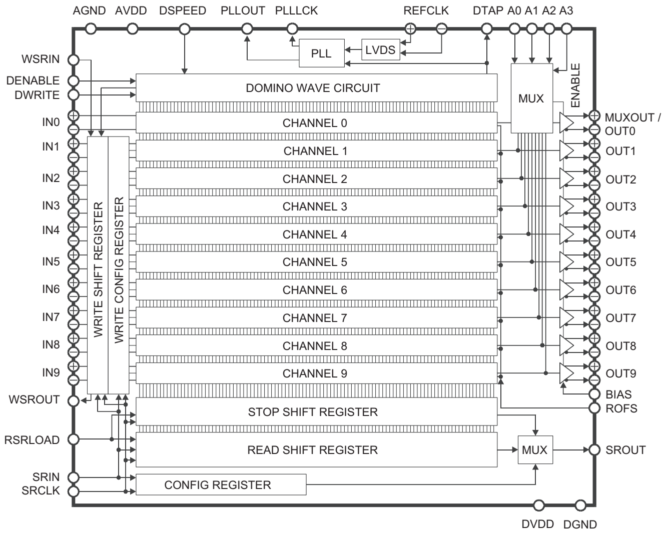


Figure 2.65 Brock diagram of DRS4 chip. Eight readout channels and one clock channel are mounted in a chip [74].

background events originating from the cosmic rays.

$t_{\mathrm{e}^+\gamma}$ is reconstructed from the on-line timing of the pTC and the LXe detector using comparators on the WDBs coupled to each input signal with the time-of-flight correction from the target to each detector. While being straightforward in the case of the LXe detector, the curved tracks of positrons make it more complicated. The timing of the pTC is extracted from a single counter placed at the most inner z of all hit counters. The correction of the track length is applied based on the position of the counter. Since the CDCH information is unavailable, there is an ambiguity of the number of positron turns; the time of flight can be different depending on the number of positron turns even though the selected counter is the same as shown in Fig. 2.66. This results in a relatively wide time window.

The pixelated design of the pTC and the higher granularity of the LXe detector give a tighter angle constraint than that of MEG. $\Theta_{\mathrm{e}^+\gamma}$ is estimated by the first hit pixel of the pTC and the peak position on the inner face of the LXe detector using a look-up table created based on a simulation.

Thanks to the improvements in the on-line reconstruction resolutions, the trigger rate is expected to be $\sim 10\,\mathrm{Hz}$, which is comparable with that of the MEG in spite of the increased beam rate.

Not only the $\mu^+ \to \mathrm{e}^+\gamma$ trigger but also a variety of trigger logics is implemented for the detector calibrations. Hereafter, the triggers used for this thesis are described.

EGamma trigger

The background γ-ray and the 17.6 MeV γ-ray data are taken based on the on-line reconstructed energy of the LXe detector. This trigger is fired when the amplitude of the sum waveform exceeds a given threshold, which is generated by summing up the waveforms of each channel with a weight given by the calibrated sensor performance.

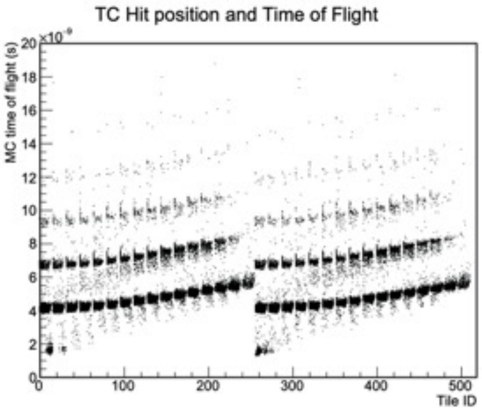


Figure 2.66 Time of flight of positrons to each pTC tile [76]. Clusters which derive from the difference of the number of positron turns can be seen.

CEX trigger

The trigger for calibration data taken with the charge exchange reaction of a π^- beam (Sec. 3.1) requires the coincidence hits in the LXe detector and a reference counter. For the LXe side, an excess of the amplitude of the sum waveform is required to obtain high energy γ-ray events as well as the EGamma trigger. For the reference counter, an excess of the amplitude of the sum waveform of the calorimeter part is required for the energy measurements as well. Coincidence hits whose pulse amplitudes are higher than a threshold in two timing counters are required for the timing measurement.

Alpha trigger

The trigger to take α events from the calibration source (Sec. 3.1) is also fired based on the sum of the sensor waveforms, but only the PMTs around each α source wire shown in Fig. 2.67 are used to have a better signal-to-noise ratio. To remove a contamination from γ-rays and cosmic rays, a particle identification based on charge-to-height ratio is also implemented.

RDC LYSO trigger

Calibration data of the RDC calorimeter uses a self-trigger of each channel, and the trigger is fired based on the waveform amplitude of a single channel (Sec. 4.1).

Random trigger

Data can be acquired by firing the trigger at random timings, which are used to monitor the performance of electronics such as noise level and event-by-event fluctuation of the baseline. They are also useful to take γ-ray data without trigger bias.

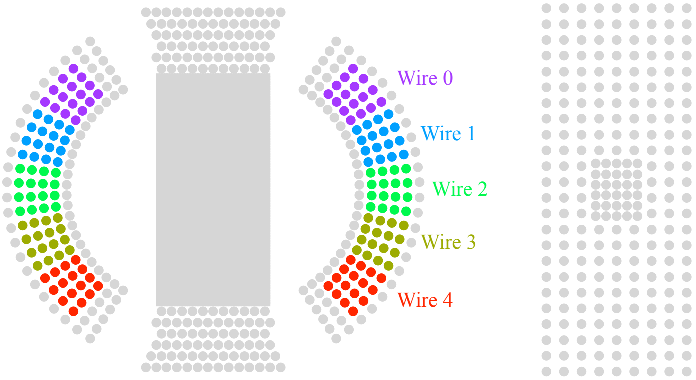


Figure 2.67 PMTs used for the alpha trigger. Only 32 PMTs around each wire are used for the trigger [62].

2.2 Pilot runs

The commissioning of the LXe detector and the RDC have been carried out since 2017 after the detector construction and the installation in the MEG experimental area. The configurations of each year data-taking are summarized in this section.

2.2.1 Run configuration

Pilot run 2017

The pilot run 2017 is the first operation and data-taking with the muon beam for the LXe detector. Data of γ-rays from the muon beam stopped on the target placed at center of the MEG II detector was accumulated for the LXe detector and the RDC. The beam intensity was adjusted to 3.2×10^7 μ^+stops/s to keep the induced current on the RDC in the dynamic range of the readout, which has been extended since the run 2018. The construction of the CDCH had not been finished, and so a mechanical mockup was installed instead, which had no wire, no electronics and a thinner aluminum end-cap of 2 mm thickness.

Pilot run 2018

In the pilot run 2018, the background γ-ray data from the muon beam were taken at a reduced beam intensity of 0.7×10^7 μ^+stops/s and the so-called MEG II beam intensity of 7×10^7 μ^+stops/s for the study on the background γ-ray spectrum. An unexpected hit rate and hit position distribution were observed with the RDC, and it was found that the hit timings were correlated to the cyclotron acceleration. This implies the beam was contaminated with positrons which should be eliminated with the separator (Sec. 2.1.1). In order to avoid the high current and the radiation damage, data for the RDC were taken only with the lower beam intensity.

The data suffers from an instability in terms of the monitoring the LXe detector performance because the DAQ configuration and the beam intensity were often switched. There is another reason that the BTS magnet was not stable due to an instability of the liquid helium supply from the PSI.

In addition to the background γ-rays, the 17.6 MeV γ-rays from the nuclear reaction of Li were measured to evaluate the energy resolution of the LXe detector. The position resolution was also measured by placing a lead collimator in front of the detector.

Pilot run 2019

There was an indication of the degradation of the sensitivity to the VUV light of the SiPMs installed in the LXe detector in the run 2018 data. Since it was difficult to conclude the degradation due to the instability of the 2018 data, dedicated data were taken for the LXe detector in 2019. The detector was operated under the MEG II intensity muon beam, and the photosensor performance was often measured to monitor the degradation.

It was found that the muon beam intensity seemed to be 10–20% higher than expected, which was suggested from the observation of the detector hit rates. This was likely due to a problem on the beam blocker during a beam intensity measurement.

Pilot run 2020

In the pilot run 2020, the background γ-rays were measured with the LXe detector. Higher hit rates than expected from the measured beam intensity were observed again with all the detectors, and thus the beam intensity was tuned referring to the detector hit rates. The background γ-ray data were accumulated at beam intensities of $1.6 \times 10^7\ \mu^+$stops/s and $7.7 \times 10^7\ \mu^+$stops/s, which were calculated by the comparisons between the detector hit rates and the beam intensity measured in the run 2021.

The quasi-monochromatic γ-ray of 54.9 MeV using the charge exchange reaction from a π^- beam was also measured to evaluate the energy resolution and the timing resolution of the LXe detector near the signal energy.

Run 2021

In the run 2021, all the detectors were installed and they were read out with the full electronic channels. Although data under the muon beam were accumulated, only the noise data are used in this thesis.

2.2.2 DAQ system

Since the full DAQ system was not available due to the delay of the development, a prototype system was used for the data acquisition in the pilot runs [75].

The number of readout channels were limited in the pilot runs due to the lack of the WDBs, and only a quarter of the LXe channels was read out (Table 2.10) while the RDC channels were fully read out until the run 2020. Fig. 2.68 shows the channel assignment of the LXe detector for each run. The readout area on the inner face was concentrated while that of the PMTs was widely distributed. This is because local information around the γ-ray hit position on the entrance face is important for the reconstruction of the position and the timing, and the wide coverage of the PMTs is required to achieve a reasonable uniformity of reconstructed energy. Only the events in which γ-rays hit the central part of the readout area are selected in the analysis.

Table 2.10 Number of readout channels of the LXe detector for each year.

Year	number of SiPM channel	number of PMT channel	Typical DAQ rate
2017	704	192	$\sim 3\,\mathrm{Hz}$
2018	640	378	$\sim 3\,\mathrm{Hz}$
2019	640	378	~ 10–$25\,\mathrm{Hz}$
2020	720	378	$\sim 20\,\mathrm{Hz}$
2021	4092	668	~ 6–$25\,\mathrm{Hz}$

2.3 Detector simulation

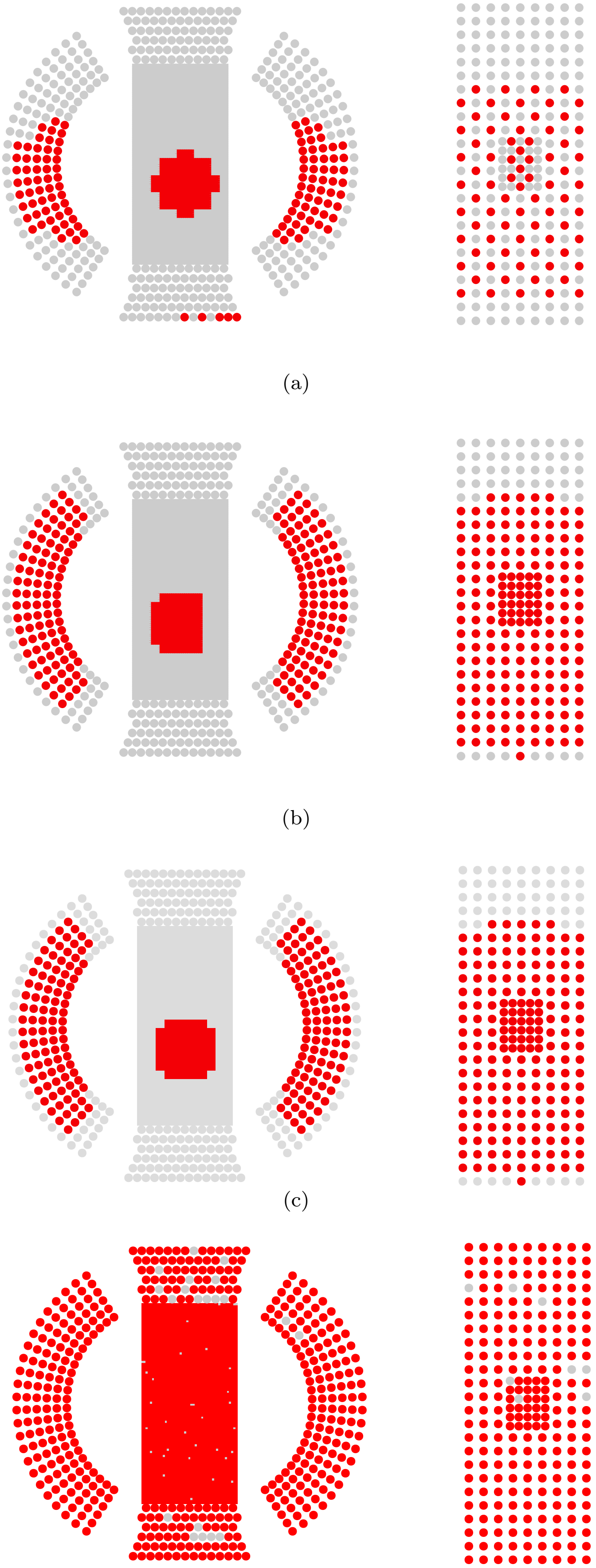
A framework for the detector simulation is developed to gain more insight into the detector performance and backgrounds. It consists of several steps as shown in Fig. 2.69. Firstly, events in the simulation are generated in "gem4" based on Monte Carlo simulation software, Geant4 (version 10.4 [58]), and interactions of particles in the detectors are simulated and recorded in "sev" files. Then, "bartender" mixes the events of multiple event types at a certain event rate to simulate pileup effects and the results are recorded in "sim" files. A electronics simulation is also performed and the simulated waveforms are recorded in "raw" files with the same format as the real data. Finally, "analyzer" analyzes the waveforms with the same algorithm as data. The bartender and the analyzer are based on the ROME framework [77].

2.3.1 Event generation and interaction

At the event generation stage of gem4, the primary particles and their kinematics can be specified. The event types used in this thesis are listed in Table 2.11.

The MEG II apparatus is implemented in the gem4. Physics interactions of generated particles with materials are simulated in the framework of the Geant4 and the hit information such as energy deposits is recorded once the particles interact with a detector sensitive volume.

When a particle enters the scintillator part of the LXe detector, it produces electromagnetic showers, which are also simulated by Geant4. Based on the property of LXe (Table 2.3), scintillation photons are generated at each energy deposit point of the electromagnetic shower. The number of generated scintillation photons is calculated with the W-value considering a Poisson fluctuation and the generation timings are given by scintillation time constants. Propagation of scintillation photons is simulated considering an absorption by impurities, the Rayleigh scattering and a reflection on the detector material. If scintillation photons reach a sensitive volume of a photosensor such as a photo-cathode of the PMT and a silicon surface of the MPPC, the



(a)

(b)

(c)

(d)

Figure 2.68 Readout channels of the LXe detector in (a) the run 2017, (b) the run 2018/2019, the run 2020 and (d) the run 2021. The non-readout channels and the dead channels are shown in gray.

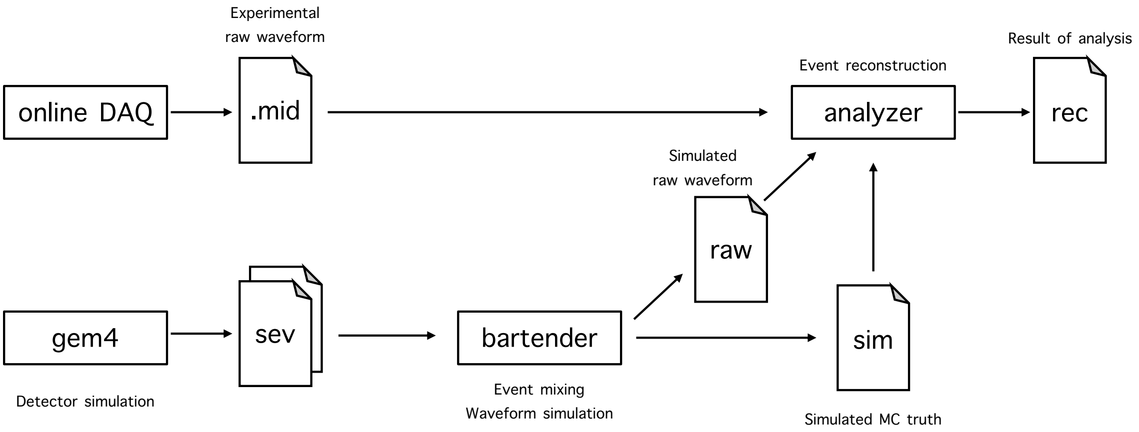


Figure 2.69 Framework of MEG II simulation software.

timings and the total number of incident photons in the volume are recorded. The fired pixel and the incident angle of the photon are also recorded for the MPPCs to be used in the following waveform simulation.

In the case of the RDC, such a dedicated simulation is skipped and only the hit information in the scintillator volume is recorded.

Table 2.11 Event types in gem4.

ID	Description	Generated particle(s)
2	SM decay of μ	e^+ and γ-ray from the SM decay on the target
21	Signal γ	52.8 MeV γ on the target
12	Michel decay of μ	e^+ from the Michel decay on the target
22	Radiative decay of μ	e^+ and γ from the RMD on the target
30	Muon beam	μ^+ injected from 564.70 cm upstream of the target [*3]
63	Alpha	5.5 MeV α from ^{241}Am inside LXe detector
64	CW-Li	17.6 MeV γ from the target
67	π^0 2γ	two γ-rays from π^0 decay ($p_{\pi^0} = 27.62$ MeV) on the target
76	Uniform γ	γ-rays from the target whose energies are uniformly selected from a certain energy range

2.3.2 Event mixing

Since the gem4 simulates a single event of a specific event type, mixing of events is necessary to know the pileup effects. In the bartender, events of different event types can be mixed at a certain rate. An event with pileups was generated by mixing pileup events at a random timing assuming a Poisson distribution to a main event whose timing was fixed to zero.

[*3] The generation point corresponds to just after the last magnet of Triplet2.

2.3.3 Waveform simulation

The bartender simulates waveforms of photosensors after the event mixing based on the recorded information in gem4.

For the waveform simulation of the LXe detector, the waveforms are generated from the single photoelectron response of each sensor as shown in Fig. 2.70. The waveform of the MPPC is obtained from the single photoelectron events acquired using an LED while that of the PMT is obtained from deconvolution of γ-ray waveforms with an exponential function of the scintillation time constant since the single photoelectron peak cannot be resolved in the charge distribution.

The MPPC characteristics are also simulated. When one pixel is fired, the crosstalk is simulated by firing another pixel with a certain probability and the after-pulsing is by firing the same pixel with certain probabilities with given time constants. Two time constants with different probabilities are introduced for the better description of the after-pulsing [78]. The dark noise is added with a given rate. To simulate the saturation, when more than one photon enter the same pixel, waveforms from the second photon are reduced according to the elapsed time from the prior photons; the posterior photons are ignored if they enter in a dead time, otherwise amplitudes of their waveforms are reduced with a time constant. The parameters used for the waveform simulation are summarized in Table 2.12.

For the simulation of the RDC waveforms, the waveforms are generated from the energy deposits of the hits since optical photons are not simulated in the previous step. Both for the timing counter and the calorimeter of the RDC, template waveforms and typical energies for the templates are obtained from μ beam data (Fig. 2.71). The template waveforms are scaled based on the energy deposits of the hits.

At the end of the simulations, electronics responses are simulated. A white noise is added to reproduce the noise observed in data. The simulated waveforms are digitized by the sampling frequency of 1.2 GHz, and the waveforms outside the dynamic range of the electronics are cut.

To make the noise simulation more realistic, a method which directly uses noise data is also implemented. The noise waveforms obtained with the random trigger without signals are added to the simulated waveforms without the white noise. Since the noise waveforms were taken at the sampling frequency of 1.4 GHz in the run 2021, the digitization is carried out at the same frequency for the simulation with the full readout channels.

Table 2.12 Parameters used to simulate MPPC characteristics for the LXe detector [61].

Parameters	Values
Dead Time	1 ns
Recovery Time Constant	38 ns
Crosstalk Probability	0.15
After-pulsing Probability	0.1 / 0.05
After-pulsing Time Constant	5 ns / 200 ns
Dark Current Rate	500 Hz

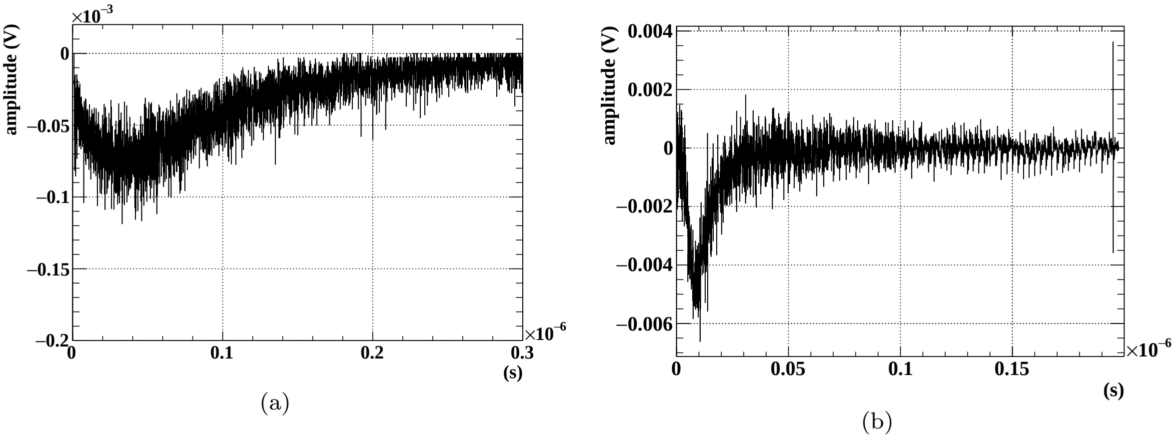


Figure 2.70 Single photoelectron response used for the waveform simulation of the LXe (a) MPPCs and (b) PMTs.

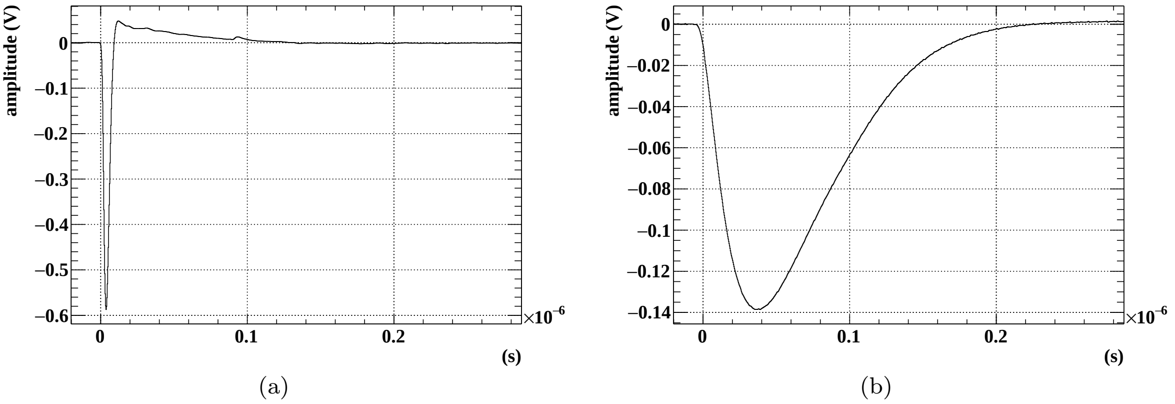


Figure 2.71 Template waveforms used for the waveform simulation of the RDC (a) timing counter and (b) calorimeter.

2.3.4 Background γ-ray

For the simulation of the background γ-rays from the muon beam, a dedicated operation is performed since its computational cost is too high to generate enough γ-ray events from muon decays due to a very small probability of having a high energy γ-ray from a muon decay. Firstly, background γ-ray events are classified into three groups by their sources as follow, and events of each group are simulated separately.

- AIF : γ-ray originating from a Michel muon decay stopped on the target. High energy γ-rays mainly come from the annihilation in flight of Michel positrons.
- RMD : γ-ray from a radiative muon decay stopped on the target.
- DIF : γ-ray from a muon decay in flight (i.e. not on the target).

Then, the generated events are mixed with each event rate calculated from the probability to have a hit in the LXe detector to produce the original background γ-ray event set.

Since only γ-rays with high energy above 20 MeV are important, simulations of particles below 20 MeV are skipped. The phase spaces of the muon decay for the AIF and the RMD event generation are restricted to $P_{\mathrm{e}+} > 20\,\mathrm{MeV}$ and $E_\gamma > 20\,\mathrm{MeV}$, respectively. For the further reduction, once positrons in the AIF reach more than 16 cm away from the target along beam axis, their simulations are stopped because such positrons are found not to produce the high energy γ-rays arriving at the LXe detector. Thanks to these reduction methods, events roughly corresponding to one day data-taking were simulated for 50 days with 200 physical cores of the recent CPUs as summarized in Table 2.13.

Table 2.13 Simulated number of events.

Event type	The number of simulated events
AIF	$\sim 1.3 \times 10^5$
RMD	$\sim 1.3 \times 10^6$
DIF	$\sim 4.8 \times 10^5$

2.3.5 Pileup mixing for uniform γ-ray

The uniform γ-ray events were used to train a deep learning model for the pileup elimination (Chap. 5). The pileup events of the uniform γ-rays were generated by mixing only a light distribution measured with the photosensors in the LXe detector. Waveforms of an event was digitized as a single γ-ray event, and they were converted to the number of photons of each photosensor with waveform analysis. For the pileup events, the light distributions of the pileup γ-rays were added to those of the uniform γ-rays.

2.4 Statistical method for $\mu^+ \to e^+\gamma$ search

In the $\mu^+ \to \mathrm{e}^+\gamma$ search, the number of signal events N_{sig} and its confidence interval are estimated with a maximum likelihood method. The likelihood function is composed of probability density functions (PDFs) of signal, RMD and accidental background. The estimated N_{sig} is converted to the branching ratio $\mathcal{B}$ with the normalization factor, which is given by the accumulated number of muon decays. The branching ratio sensitivity is defined as the median of the 90% confidence upper limits of pseudo experiments.

2.4.1 Likelihood function

The likelihood function $\mathcal{L}$ for target parameters $\boldsymbol{\theta}$ is given by

$$\mathcal{L}(\boldsymbol{\theta}|\boldsymbol{X}) = \prod_{i}^{N_{\mathrm{obs}}} p(\boldsymbol{x}_i|\boldsymbol{\theta}), \tag{2.3}$$

where N_{obs} is the total number of events, $\boldsymbol{x}_i$ is a set of observables of i-th event, and $\boldsymbol{X}$ is an array of the observable sets ($\boldsymbol{x}_1$, $\boldsymbol{x}_2$, ..., $\boldsymbol{x}_N$). The probability density function $p(\boldsymbol{x}_i|\boldsymbol{\theta})$ shows the probability to find $\boldsymbol{x}_i$ under a condition of $\boldsymbol{\theta}$. The best estimate $\hat{\boldsymbol{\theta}}$ is defined such that maximizes the likelihood function with a given set of observables.

This likelihood function can be extended to take the statistical fluctuation of N_{obs} in a Poisson distribution into account, that is,

$$\mathcal{L}(\boldsymbol{\theta}|\boldsymbol{X}) = \frac{N^{N_{\mathrm{obs}}}e^{-N}}{N_{\mathrm{obs}}!}\prod_{i}^{N_{\mathrm{obs}}} p(\boldsymbol{x}_i|\boldsymbol{\theta}), \tag{2.4}$$

where N is the sum of the expected numbers of events [79].

In this context, the parameters to be estimated are the expected numbers of the signal, the RMD and the accidental background events, namely

$$\boldsymbol{\theta} = (N_{\mathrm{sig}}, N_{\mathrm{RMD}}, N_{\mathrm{acc}}). \tag{2.5}$$

Then, N can be written by the sum of the numbers of the three event types:

$$N = N_{\mathrm{sig}} + N_{\mathrm{RMD}} + N_{\mathrm{acc}}. \tag{2.6}$$

From the definition of PDF, there are relations bellow:

$$p(\mathrm{sig}) + p(\mathrm{RMD}) + p(\mathrm{acc}) = 1, \tag{2.7}$$

$$p(\mathrm{sig}|\boldsymbol{x}_i) + p(\mathrm{RMD}|\boldsymbol{x}_i) + p(\mathrm{acc}|\boldsymbol{x}_i) = 1, \tag{2.8}$$

where $p(\mathrm{sig})$, $p(\mathrm{RMD})$ and $p(\mathrm{acc})$ are the PDF of signal, RMD or accidental background event, and $p(\mathrm{sig}|\boldsymbol{x}_i)$, $p(\mathrm{RMD}|\boldsymbol{x}_i)$ and $p(\mathrm{acc}|\boldsymbol{x}_i)$ are the PDF of signal, RMD or accidental background event after the observation of $\boldsymbol{x}_i$. From the Beyes' theorem, the following relation is completed:

$$p(\mathrm{x}|\boldsymbol{x}_i) = \frac{p(\boldsymbol{x}_i|\mathrm{x})p(\mathrm{x})}{p(\boldsymbol{x}_i)}, \tag{2.9}$$

where x = sig, RMD or acc. Thus, the PDF can be written as

$$p(\boldsymbol{x}_i|N_{\mathrm{sig}}, N_{\mathrm{RMD}}, N_{\mathrm{acc}}) = p(\boldsymbol{x}_i|\mathrm{sig})p(\mathrm{sig}) + p(\boldsymbol{x}_i|\mathrm{RMD})p(\mathrm{RMD}) + p(\boldsymbol{x}_i|\mathrm{acc})p(\mathrm{acc}). \tag{2.10}$$

Since $p(\mathrm{x})$ is the probability to be the event type x, they are given by

$$p(\mathrm{x}) = \frac{N_{\mathrm{x}}}{N}. \tag{2.11}$$

Therefore, the PDF can be

$$p(\boldsymbol{x}_i|N_{\mathrm{sig}}, N_{\mathrm{RMD}}, N_{\mathrm{acc}}) = \frac{N_{\mathrm{sig}}}{N}\cdot S(\boldsymbol{x}_i) + \frac{N_{\mathrm{RMD}}}{N}\cdot R(\boldsymbol{x}_i) + \frac{N_{\mathrm{acc}}}{N}\cdot A(\boldsymbol{x}_i), \tag{2.12}$$

where $S(\boldsymbol{x}_i)$, $R(\boldsymbol{x}_i)$ and $A(\boldsymbol{x}_i)$ are the PDF for signal, RMD and accidental background event, respectively.

Since the N_{RMD} and N_{acc} can be estimated also from the sideband of the fit region, constraints can be added as

$$\begin{aligned}\mathcal{L}(N_{\mathrm{sig}}, N_{\mathrm{RMD}}, N_{\mathrm{acc}}|\boldsymbol{X}) = {} & \frac{e^{-N}}{N_{\mathrm{obs}}!}e^{-\frac{(N_{\mathrm{RMD}}-\mu_{\mathrm{RMD}})^2}{2\sigma_{\mathrm{RMD}}^2}}e^{-\frac{(N_{\mathrm{acc}}-\mu_{\mathrm{acc}})^2}{2\sigma_{\mathrm{acc}}^2}} \\ & \times\prod_{i=1}^{N_{\mathrm{obs}}}(N_{\mathrm{sig}}S(\boldsymbol{x}_i) + N_{\mathrm{RMD}}R(\boldsymbol{x}_i) + N_{\mathrm{acc}}A(\boldsymbol{x}_i)),\end{aligned} \tag{2.13}$$

where μ_x and σ_x (x = RMD or acc) are the means and the uncertainties of the estimation.

2.4.2 PDF

The PDFs are defined as functions of the observables $\boldsymbol{x}_i$ as

$$\boldsymbol{x}_i = (E_\gamma, E_{\mathrm{e}^+}, t_{\mathrm{e}^+\gamma}, \phi_{\mathrm{e}^+\gamma}, \theta_{\mathrm{e}^+\gamma}, t_\mathrm{rdc}^\mathrm{us}, t_\mathrm{rdc}^\mathrm{ds}, E_\mathrm{rdc}^\mathrm{ds}). \tag{2.14}$$

E_γ and E_{e^+} are the reconstructed energy of the γ-ray and the positron, respectively. The time difference of the positron and the γ-ray $t_{\mathrm{e}^+\gamma}$ is calculated from the measured timings considering the propagation times from the decay vertex. The relative azimuthal and polar angles of the flipped positron direction with reference to the γ direction $(\phi_{\mathrm{e}^+\gamma}, \theta_{\mathrm{e}^+\gamma})$ are calculated using the reconstructed positions and the decay vertex. The positron timing and energy measured with the RDC ($t_\mathrm{rdc}^\mathrm{us}$, $t_\mathrm{rdc}^\mathrm{ds}, E_\mathrm{rdc}^\mathrm{ds}$) are used to enhance the discriminant power of accidental background.

As discussed above, three types of PDF are considered: S, R and A representing signal, RMD and accidental background, respectively.

Signal PDF

The PDF for the signal event is defined as

$$\begin{aligned} S(E_\gamma, E_{\mathrm{e}^+}, t_{\mathrm{e}^+\gamma}, \phi_{\mathrm{e}^+\gamma}, \theta_{\mathrm{e}^+\gamma}, t_\mathrm{rdc}^\mathrm{us}, t_\mathrm{rdc}^\mathrm{ds}, E_\mathrm{rdc}^\mathrm{ds} | \boldsymbol{r}_\gamma, \boldsymbol{y}_i) &= S(t_{\mathrm{e}^+\gamma} | E_\gamma, E_{\mathrm{e}^+}, \boldsymbol{y}_i) \\ &\times S(E_\gamma | \boldsymbol{r}_\gamma) \\ &\times S(E_{\mathrm{e}^+} | \boldsymbol{y}_i) \\ &\times S(\phi_{\mathrm{e}^+\gamma} | \boldsymbol{r}_\gamma, \boldsymbol{y}_i) \\ &\times S(\theta_{\mathrm{e}^+\gamma} | \boldsymbol{r}_\gamma, \boldsymbol{y}_i) \\ &\times S(t_\mathrm{rdc}^\mathrm{us}, t_\mathrm{rdc}^\mathrm{ds}, E_\mathrm{rdc}^\mathrm{ds}), \end{aligned}$$

where $\boldsymbol{r}_\gamma$ is the first conversion point of a γ-ray and $\boldsymbol{y}_i$ is a set of reconstructed positron variables such as direction, decay vertex, and tracking quality.

The signal PDF of the γ-ray energy $S(E_\gamma)$ is defined by the detector response to the signal γ-ray. It is created from the simulated energy distribution for the signal γ-ray smeared with an additional Gaussian to match the measured energy resolution[*4]. The E_{e^+} PDF is given by the detector response to the the signal positron. The signal PDF for the relative timing of a positron and a γ-ray is defined by a combination of the the timing resolutions of them[*5]. The angular PDFs are given by the γ-ray position resolution, the positron vertex resolution and the positron direction resolution. The signal PDF of the RDC observables is created by the timing and energy distribution of Michel positrons as will be discussed in Sec. 6.2.1.

[*4] The signal E_γ PDF will be generated from the energy distribution for 54.9 MeV γ-ray in CEX calibration in the analysis of real data.

[*5] The $t_{\mathrm{e}^+\gamma}$ resolution will be evaluated with the timing coincidence hits to the detectors from the RMD in the analysis of real data.

RMD PDF

The PDF for the RMD event is written as

$$\begin{aligned} R(E_\gamma, E_{\mathrm{e}^+}, t_{\mathrm{e}^+\gamma}, \phi_{\mathrm{e}^+\gamma}, \theta_{\mathrm{e}^+\gamma}, t^{\mathrm{us}}_{\mathrm{rdc}}, t^{\mathrm{ds}}_{\mathrm{rdc}}, E^{\mathrm{ds}}_{\mathrm{rdc}} | \boldsymbol{r}_\gamma, \boldsymbol{y}_i) = &\, R(t_{\mathrm{e}^+\gamma} | \boldsymbol{y}_i) \\ &\times R(E_\gamma, E_{\mathrm{e}^+}, \phi_{\mathrm{e}^+\gamma}, \theta_{\mathrm{e}^+\gamma}) \\ &\times R(t^{\mathrm{us}}_{\mathrm{rdc}}, t^{\mathrm{ds}}_{\mathrm{rdc}}, E^{\mathrm{ds}}_{\mathrm{rdc}}). \end{aligned}$$

The RMD PDF for $t_{\mathrm{e}^+\gamma}$ is also given by the combination of the timing resolutions. The variables, $E_\gamma, E_{\mathrm{e}^+}, \phi_{\mathrm{e}^+\gamma}, \theta_{\mathrm{e}^+\gamma}$, are correlated with each other, and the relation can be calculated theoretically in the standard model. Therefore, the RMD PDF of the variables is obtained by convoluting the theoretical spectrum with the detector response functions for each observable. The RMD PDF for the RDC observables is the same as the signal PDF since only accidental hits from Michel decays are expected.

Accidental background PDF

The PDF for the accidental background is defined as

$$\begin{aligned} A(E_\gamma, E_{\mathrm{e}^+}, t_{\mathrm{e}^+\gamma}, \phi_{\mathrm{e}^+\gamma}, \theta_{\mathrm{e}^+\gamma}, t^{\mathrm{us}}_{\mathrm{rdc}}, t^{\mathrm{ds}}_{\mathrm{rdc}}, E^{\mathrm{ds}}_{\mathrm{rdc}} | \boldsymbol{r}_\gamma, \boldsymbol{y}_i) = &\, A(t_{\mathrm{e}^+\gamma} | E_\gamma, E_{\mathrm{e}^+}, \boldsymbol{y}_i) \\ &\times A(E_\gamma | \boldsymbol{r}_\gamma) \\ &\times A(E_{\mathrm{e}^+} | \boldsymbol{y}_i) \\ &\times A(\phi_{\mathrm{e}^+\gamma} | \boldsymbol{r}_\gamma) \\ &\times A(\theta_{\mathrm{e}^+\gamma} | \boldsymbol{r}_\gamma) \\ &\times A(t^{\mathrm{us}}_{\mathrm{rdc}}, t^{\mathrm{ds}}_{\mathrm{rdc}}, E^{\mathrm{ds}}_{\mathrm{rdc}} | E_\gamma). \end{aligned}$$

The $t_{\mathrm{e}^+\gamma}$ distribution in accidental background events should be flat by definition, and thus its PDF is given by a uniform distribution. The E_γ is generated from the simulated background spectrum with the Gaussian smearing[*6]. At the same time, the accidental background PDF for the RDC observables is created as a function of E_γ with the smeared spectrum to include its energy dependence as will be discussed in Sec. 6.1.2. The other PDFs are generated based on the resolutions of each variable.

2.4.3 Analysis Region

In this thesis, the likelihood fitting is performed in the analysis region defined as follows:

- $48\,\mathrm{MeV} < E_\gamma < 58\,\mathrm{MeV}$,
- $52.2\,\mathrm{MeV} < E_{\mathrm{e}^+} < 53.5\,\mathrm{MeV}$,
- $|t_{\mathrm{e}^+\gamma}| < 0.25\,\mathrm{ns}$,
- $|\theta_{\mathrm{e}^+\gamma}| < 40\,\mathrm{mrad}$,
- $|\phi_{\mathrm{e}^+\gamma}| < 40\,\mathrm{mrad}$,
- $-20\,\mathrm{ns} < t^{\mathrm{us}}_{\mathrm{rdc}} < 28\,\mathrm{ns}$,

---

[*6] The accidental background E_γ PDF will be generated using the timing sideband data in the analysis of real data.

- $-20\,\mathrm{ns} < t^{\mathrm{ds}}_{\mathrm{rdc}} < 28\,\mathrm{ns}$,
- $0\,\mathrm{MeV} < E^{\mathrm{ds}}_{\mathrm{rdc}} < 50\,\mathrm{MeV}$.

2.4.4 Normalization

Since the likelihood fitting provides the number of the signal events, it needs to be converted with a normalization factor to obtain the branching ratio of the $\mu^+ \to \mathrm{e}^+\gamma$ decay. The normalization factor k is given by the accumulated number of muon decays as Eq. (1.4). In the analysis for the $\mu^+ \to \mathrm{e}^+\gamma$ search, it is computed from a combination of the detected number of events for two independent datasets: Michel decay events and RMD events.

In this thesis, the nominal value of the normalization factor is estimated based on the original DAQ plan of the MEG II experiment and the simulation at the design stage except for ϵ_{e^+}, which was updated to be more realistic based on the observation in the pilot runs [50]. Table 2.14 summarizes the values used for the calculation. As a result, the total number of muon decays is $k = 9.38 \times 10^{13}$. All of the sensitivity estimations are performed by scaling these values as necessary.

Table 2.14 Nominal settings for the normalization of the MEG II experiment.

Variables	Values
R_μ	7×10^7 μ^+stops/s
T	Three years with 20 week data-taking per year including 84% live fraction
Ω	10.8%
ϵ_{e^+}	65%
ϵ_γ	69%
ϵ_{cut}	91% (same as the first half of MEG)

2.4.5 Sensitivity calculation

The branching ratio sensitivity is defined as the median of the upper limits at 90% confidence level for pseudo experiments with a background-only hypothesis. In this thesis, the upper limit is calculated by asymptotic formulae in which the profile likelihood ratio is approximated by the chi-square distribution [80].

An ensemble of pseudo experiments are generated from given PDFs assuming a background-only hypothesis; N_{sig} is fixed to be zero. The pseudo experiment is a statistical simulation, which simulates a statistical behavior of observables based on their distributions and correlations including the detector response. A fluctuation of the number of events is also simulated following Poisson distribution with mean values of expected numbers of N_{RMD} and N_{acc}. The projected sensitivity is calculated from $\mathcal{O}(1000)$ pseudo experiments.

2.5 Subject of this thesis

This thesis focuses on a suppression of γ-ray backgrounds in the MEG II experiment with two methods: the elimination of pileup γ-rays and the identification of the background γ-rays. The former one was achieved by developing the pileup elimination algorithms for the LXe detector, and the latter was by introducing the RDC. Thus, the analysis and the performance of the detectors and the background suppression methods are discussed in the following chapters. Hereafter, statuses and updates for the two detectors are summarized, and the MEG II projected sensitivity without any updates in this thesis is described.

2.5.1 Status and update for LXe detector

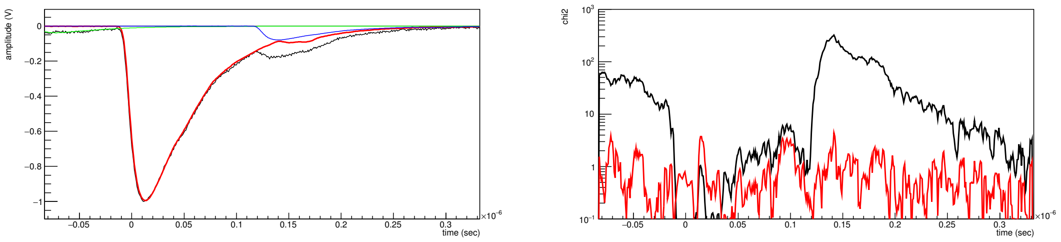
The LXe detector had been developed by the liquid xenon group of the MEG II collaboration. Its construction was finished in 2017, and installed to the MEG experimental area followed by the commissioning of the detector. The detector performance was measured with the muon beam and other γ-ray sources, which are explained in Sec. 3.1.2. The position and energy resolutions were evaluated using the data taken in 2018 and 2019, and were reported in [62], which are summarized in Sec. 3.4 and Sec. 3.6, respectively. The timing resolution was also measured with γ-rays from the muon beam and reported in [62]. It was updated in this thesis using the other γ-ray source. The author contributed to the development of a reference counter for the measurement, the timing calibration of the detector and the evaluation of the resolution as described in Sec. 3.5.

Since an unexpected degradation of the MPPC performance was observed under the muon beam, the radiation damage on the MPPC was investigated. The author was responsible for the lab test, which was carried out to understand the observed degradation in the detector. The measurements and the results are explained in Appendix C.

Since pileup γ-rays increases the number of background events in the signal energy region, the further background suppression can be achieved by eliminating them. The algorithm used in [62] tries to find out the pileup γ-rays by searching for the excess of a measured waveform from a template waveform (Fig. 2.72). This method is sensitive only to the pileup γ-rays whose hit timings are far enough from the main γ-ray since their waveforms cannot be separated from that of the main γ-ray. The author developed a series of the pileup elimination algorithms to deal with all of the pileups, and succeeded in achieving better performance, which is discussed in Sec. 5.

2.5.2 Status and update for RDC

The RDC had been developed by the RDC group of the MEG II collaboration. It was tested with the muon beam in 2016 after the construction followed by the performance measurement with the LXe detector from 2017. The author has taken responsibility all of the operation, calibration and performance evaluation of the RDC from 2017, which are described in Chap. 4,



(a) (b)

Figure 2.72 Pileup elimination by waveform for a typical waveform with pileups [62]. (a) The original PMT waveform (black) is unfolded to three identified pulses (red, blue, and green). (b) Chi-square of this event assuming (black) signal γ-ray, and (red) three γ-rays. The pileups are found by searching for the excess in the chi-square.

in addition to the development of the procedure to introduce the RDC information to the $\mu^+ \to e^+\gamma$ analysis as discussed in Chap. 6. These works confirmed the background suppression power with the RDC and the improvement of the sensitivity.

2.5.3 MEG II projected sensitivity without updates in this thesis

Here, we define the parameters used for the sensitivity calculation as the nominal setting, which are summarized in Table 2.15. In addition to the designed beam intensity of 7×10^7 μ^+stops/s, the case at a half intensity, 3.5×10^7 μ^+stops/s, is calculated as a possible scenario to cope with the PDE degradation of the LXe MPPCs. The positron resolutions are updated from the design values based on the observations [50]. The γ-ray resolutions at the MPPC PDE of 6% are used, and the timing resolution is assumed to be consistent with the MC expectation. The nominal branching ratio sensitivity before applying any updates in this thesis was calculated with the updated nominal parameters at 7×10^7 μ^+stops/s; the γ-ray pileup elimination used in [62] was applied, and the RDC was excluded. For the generation of the E_γ PDF, simulated data with a real noise taken in 2021 were used. The sensitivity was calculated to be 7.8×10^{-14} for three-year data-taking.

This value is, indeed, worse than the previously reported values in [62], i.e. 6×10^{-14}. The discrepancy partly comes from an increase of the number of high energy γ-rays resulting from the consideration of the real noise situation; only a Gaussian white noise was included for the E_γ PDF generation in the previous estimation[*7] though there are some new observations such as a coherent noise in reality as will be discussed in Sec. 5.3.2. Fig. 2.73 shows the energy spectra generated with the simulated noise and the real noise. The events in 51.5–54 MeV increases by 28% due to the real noise from the simulated one. Another cause is an increase of positron inefficiency due to the higher noise level than the original expectation: $k = 1.03 \times 10^{14}$ in the

---

[*7] The performance was validated with data in the run 2019, and its consistency with the simulated noise dataset was proved. However, only the limited channels were used in the run, which led an underestimation of the effect of the real noise.

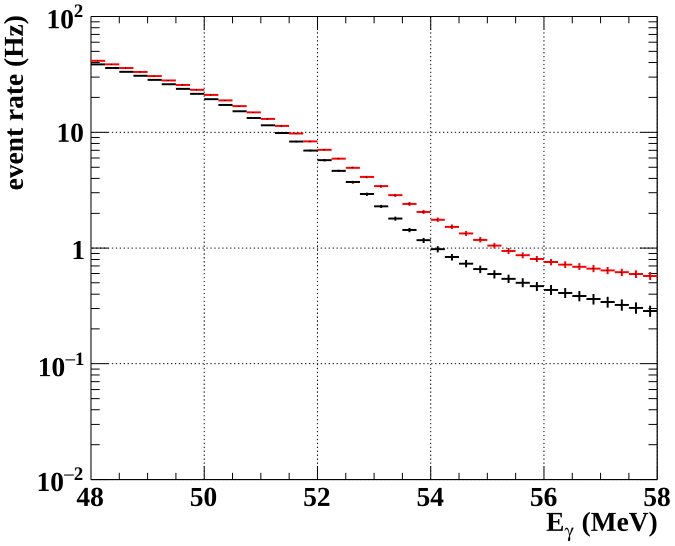


Figure 2.73 Energy spectra of the simulated background γ-rays with the simulated noise (black) and the real noise (red) after applying the elimination algorithm in [62].

past.

Table 2.15 Parameters for sensitivity calculation as the nominal setting.

Variables	Values	
	7×10^7 μ^+stops/s	3.5×10^7 μ^+stops/s
$\sigma_{p_{e^+}}$	100 keV/c	90 keV/c
$\sigma_{\theta_{e^+}}$	6.7 mrad	6.2 mrad
$\sigma_{\phi_{e^+}}$	4.9 mrad	4.7 mrad
σ_{E_γ}	1.7%	
σ_{u_γ}	2.5–8.1 mm (for different w_γ)	
σ_{v_γ}	2.5–7.4 mm (for different w_γ)	
σ_{w_γ}	2.4–13.9 mm (for different w_γ)	
$\sigma_{t_{e^+\gamma}}$	70 ps	68 ps
ϵ_{e^+}	65%	74%
ϵ_γ	69%	
k	9.38×10^{13} (for three years)	5.34×10^{13} (for three years)

Chapter 3

Performance Evaluation of LXe Detector

The LXe γ-ray detector measures γ-ray energy, position and timing. This detector was developed for the MEG experiment and upgraded for the MEG II experiment in order to achieve better performance as mentioned in Sec. 2.1.3.2. In this chapter, the descriptions of several calibration methods and sensor alignment are given firstly. Then, the performance of the detector in terms of the position, timing and energy resolution is also discussed in this order.

3.1 Calibration

The LXe detector requires precise calibration to achieve the aimed resolutions. Therefore, many calibration methods are developed as summarized in Table 3.1.

Table 3.1 Calibration tools for the LXe detector [2].

Category	Process	Energy
Charge exchange reaction	$\pi^- \mathrm{p} \to \pi^0 \mathrm{n}$ $\pi^0 \to \gamma\gamma$	$55, 83\,\mathrm{MeV}$ photons
Proton accelerator	$^{7}\mathrm{Li}(\mathrm{p},\gamma)^{8}\mathrm{Be}$ $^{11}\mathrm{B}(\mathrm{p},\gamma)^{12}\mathrm{C}$	$14.8, 17.6\,\mathrm{MeV}$ photons $4.4, 11.6, 16.1\,\mathrm{MeV}$ photons
Neutron generator	$^{58}\mathrm{Ni}(\mathrm{n},\gamma)^{59}\mathrm{Ni}$	$9\,\mathrm{MeV}$ photons
Am source	$^{241}\mathrm{Am}(\alpha,\gamma)^{237}\mathrm{Np}$	$5.5\,\mathrm{MeV}$ αs
LED		blue-UV region

3.1.1 Sensor calibration

For the event reconstruction, the number of photons N_{pho} on each photosensor is calculated from the observed charge q as

$$N_{\mathrm{pho}} = q/R_{\mathrm{sensor}}/e, \tag{3.1}$$

where R_{sensor} is a sensor response, which must be measured correctly for the precise reconstruction.

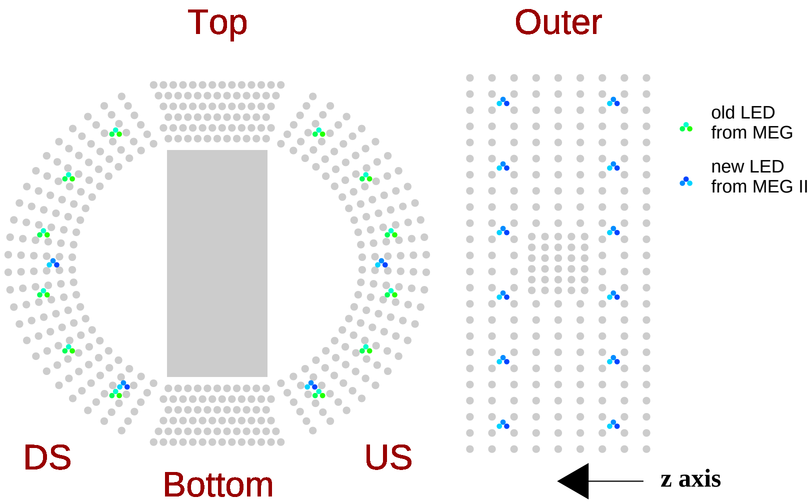


Figure 3.1 Installed positions of the LEDs [84].

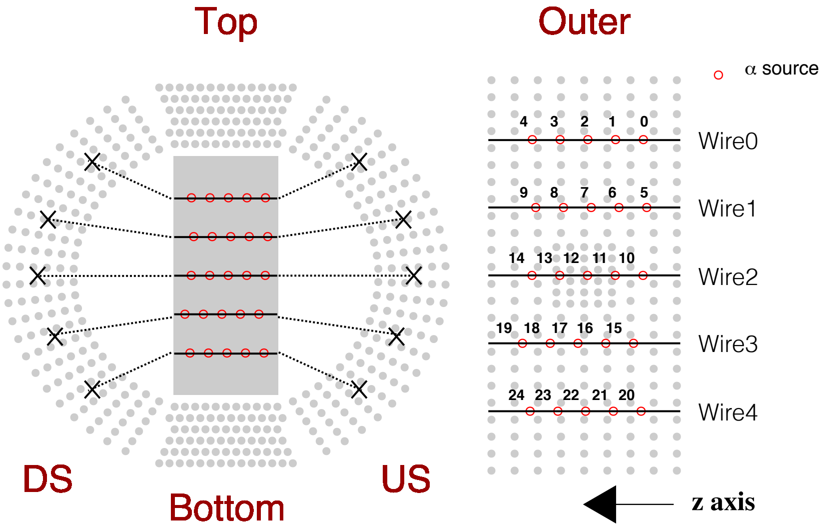


Figure 3.2 Installed positions of the alpha sources [84].

The response of the photosensors is calibrated using blue-light LEDs and α sources installed in the detector as shown in Fig. 3.1 and Fig. 3.2. The LEDs on the lateral face (Toyoda Gosei E1L49–3B1A–02. [81]) are reused from the MEG, and used for the PMTs (Fig. 3.3(a)). The LEDs (Kingbright KA-3021QBS-D. [82]) on the outer face are newly installed for the MEG II for the MPPCs, which are covered with a Teflon sheet to diffuse the light for a uniform light distribution (Fig. 3.3(b)). The α sources, ^{241}Am, are installed by being crimped on tungsten wires with 100 μm diameter as shown in Fig. 3.3(c) [83]. Five wires are stretched between the upstream and the downstream faces, and each wire includes five sources at intervals of 12.4 cm. The activity of the sources is 200 Bq at most, which is low enough not to affect the beam data-taking.

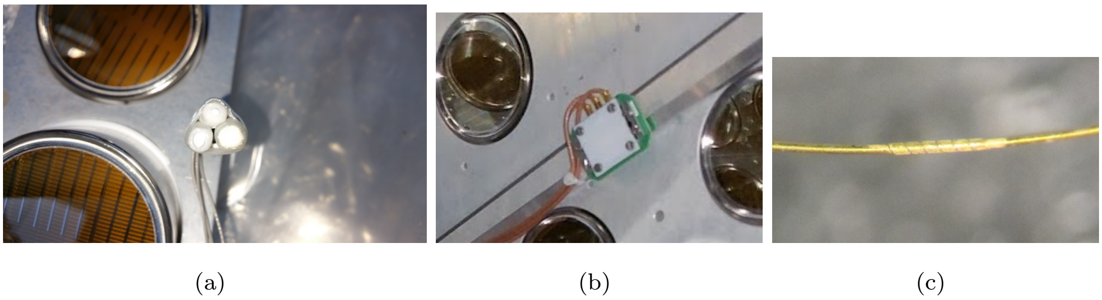

(a) (b) (c)

Figure 3.3 Sensor calibration sources. (a) LEDs on the lateral face. (b) LEDs on the outer face. (c) α source crimped on a wire.

PMT

The sensor response term of the PMTs R_{PMT} can be written as

$$R_{\mathrm{PMT}} = Gain \times CE \times QE. \tag{3.2}$$

The $Gain$ includes the sensor gain and the amplifier gain of the electronics. The QE is the quantum efficiency, which is a probability of the incident photon conversion, and the CE is the collection efficiency, which is a probability where the converted photoelectrons reach the active region of the sensor.

The PMT gain is calculated using the Poisson statistics of the number of photoelectrons observed with a PMT N_{phe} for a constant intensity light as

$$\sigma^2_{N_{\mathrm{phe}}} = \mu_{N_{\mathrm{phe}}} + \sigma^2_{\mathrm{noise}}, \tag{3.3}$$

where $\sigma_{N_{\mathrm{phe}}}$ and $\mu_{N_{\mathrm{phe}}}$ are the standard deviation and the mean of N_{phe}, respectively, and σ_{noise} is a noise contribution. Using this equation, the relation between the standard deviation σ_q and the mean μ_q of the measured charge can be written as

$$\sigma^2_q = Gain \times (\mu_q + Gain \times \sigma^2_{\mathrm{noise}}), \tag{3.4}$$

since the charge can be converted from the number of photoelectrons with $q = Gain \times N_{\mathrm{phe}} \times e$.

The calibration data are taken with the LEDs changing the intensities as shown in Fig. 3.4. The PMT gain is extracted by fitting a linear function to the mean–variance relation of the measured charge as its slope corresponds to the objective value. The PMT gain needs to be calibrated periodically because the value continuously decreases during the data-taking with the μ beam probably due to the degradation of the dynode material.

The other terms of the PMT response, i.e. the QE and the CE, can be calculated by comparing the observed charge for the scintillation light from the α sources with that of MC simulation. The observed charge q can be written with the expected number of photons $N^{\mathrm{MC}}_{\mathrm{pho}}$ as

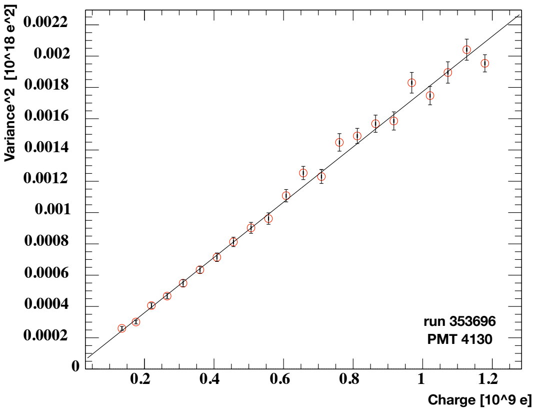


Figure 3.4 Relation between charge mean μ_q and charge variance σ_q^2 measured with different LED intensities [62]. The gain is calculated from the fitted slope.

$$\begin{aligned} q &= Gain \times CE \times QE \times N_{\text{pho}} \times e, \\ &= Gain \times CE \times QE \times N_{\text{pho}}^{\text{MC}} \times LY \times e, \end{aligned} \tag{3.5}$$

where LY is a relative level of the light yield normalized to 1 for the MC. From the direct observation, we can know the product of the response parameters as

$$CE \times QE \times LY = q/Gain/N_{\text{pho}}^{\text{MC}}/e. \tag{3.6}$$

The precision of the measurement was estimated to be 4% [85].

MPPC

Similarly to the PMTs, the sensor response term of the MPPCs R_{MPPC} can be written as

$$R_{\text{MPPC}} = Gain \times PDE \times ECF, \tag{3.7}$$

where PDE is the photon detection efficiency. The PDE is a product of the quantum efficiency of silicon, the avalanche initiation probability and the geometrical fill factor, which is a fraction of the sensitive microcells over all cells. ECF is the excess charge factor.

The MPPC gain is calculated with a single photoelectron peak obtained using an LED light. Fig. 3.5 shows an example of the measured charge distribution. The gain is extracted from the difference between the mean of zero photoelectron and single photoelectron peaks obtained by fitting two Gaussians. Fig. 3.6(a) shows the measured MPPC gains. The distribution is sufficiently uniform though there is a slight lot dependence. The precision of the measurement was estimated to be 2.5% [85].

The ECF, which is a factor to correct the measured charge for the effects of correlated noises, is also calculated with the LED data assuming a Poisson distribution. If there is no correlated noise, the number of photoelectrons follows a Poisson distribution $P_\lambda(x)$, where λ is the Poisson mean. Since the number of the zero photoelectron events is not affected by the correlated noise, λ can be obtained by a fraction of such events f_0 as

$$f_0 = P_\lambda(0) = e^{-\lambda}. \tag{3.8}$$

If there is the correlated noise, the mean of the distribution μ deviates from the expected mean, and so the ECF can be calculated as $ECF = \mu/\lambda$. Fig. 3.6(b) shows the ECFs of the MPPCs. A large production lot dependence is observed. The precision of the measurement was estimated to be 1.4%.

The PDE, which is defined by $N_{\text{phe}} = N_{\text{pho}} \times PDE$, is measured using the α sources by comparing the measured number of photoelectrons with that of MC similar to the PMT as

$$PDE \times LY = q/Gain/ECF/N_{\text{pho}}^{\text{MC}}/e. \tag{3.9}$$

The PDEs measured at the beginning of the run 2019 are shown in Fig. 3.7(a). The precision of the measurement was estimated to be 3%. The average PDE is 7%, which is much lower than that measured at the lab (∼ 15%). It is suspected that this discrepancy is caused by PDE decrease due to the exposure to the μ beam. Fig. 3.7(b) shows the PDEs as a function of the incident angle from each α source to each MPPC. A larger angular dependence was observed than that expected from the reflection, which is consistent with the observation at the lab [62].

Fig. 3.8 shows the MPPC response, i.e. $R_{\text{MPPC}} \times LY$, to the VUV light and the visible light averaged over all MPPCs during the run 2019. The response to the VUV light was found to decrease by 11% correlated to the beam exposure. If the decrease comes from the gain or the ECF, the response to visible light should decrease to the same level. However, it decreased only by 1%, and thus that is not the case. To test the contribution of LY, the PMT response to the VUV light except for the gain, i.e. $CE \times QE \times LY$, was also measured and the result is shown in Fig. 3.9. It also decreased, but only by 4%, which means the degradation of the MPPC response cannot be explained by the decrease of the xenon light yield. Therefore, it is concluded that the degradation of the MPPC response mainly comes from the PDE. This degradation was confirmed with other variables such as charge for background γ-rays and current measured with the MPPCs under the μ beam [62].

Since the degradation is correlated with the beam usage, the PDE decrease can be caused by radiation damage. The possible irradiation sources are γ-rays, neutrons and VUV photons, and their fluence during the run 2019 are summarized in Table 3.2. Since the dose levels of the γ-ray and the neutron are much less than the levels at which the performance of the MPPC can be affected according to previous studies [86] [87] [88] [89] and the damage due to the VUV photon has not been reported, the PDE degradation was unexpected.

A possible cause of the degradation is a surface damage by the VUV photon irradiation. In the surface damage, the oxide charges are accumulated near the interface between the silicon layer and the passivation layer, and they can distort the electric field around it. Since the collection efficiency for the VUV light is enhanced by the electric field as described in Sec. 2.1.3.2, it can be affected by the distortion while that for the visible light is hardly affected.

The effect of the radiation damage was investigated at the lab. Although the degradation observed for the MPPCs in the LXe detector was not reproduced completely, the PDE degradation due to the VUV photon irradiation was observed with much slower decreasing speed as will be described in Appendix C. The result suggests that the PDE can be saturated around 10% of the original value, which corresponds to the PDE of 2%.

If the degradation is due to the surface damage, a thermal annealing can recover the degraded PDE by the release of the accumulated charges by the thermal excitation. The effect of the annealing was tested for a small number of MPPCs in the LXe detector after the run 2018. LXe was transferred to the storage tank, and the detector was filled with gaseous xenon at room temperature during the annealing. The MPPCs were heated up by the sensor current when they were illuminated by intense LED light. The annealing was performed for six MPPCs with different conditions: the induced current from 17 to 24 mA and the duration of 23 or 38 hours. The temperature of the MPPCs was not directly measured, but it is expected be about 70°C according to a lab test performed under the same condition. Fig. 3.10(a) shows the degree of PDE recovery for VUV light after the annealing. It was significantly recovered to the same level observed at the lab test before installation, i.e. $\sim 17\%$. The dependence on the annealing strength of the recovery was also observed, which supports it resulted from the annealing. Not only the PDE for VUV light but also that for visible light recovered as shown in Fig. 3.10(b) though the recovery was much smaller than that of VUV light since the PDE for visible light was not reduced so much.

The signal amplification is useful to compensate the decreasing PDE during data-taking since the annealing is not possible so frequently. In the original plan, the PDE of $\sim 15\%$ is assumed, and the amplifier is not used to avoid the signal amplitude of the MPPCs exceeding the dynamic range of the readout, but it can be increased depending on the PDE as shown in Table 3.3.

Although the MPPC PDE can be recovered by the thermal annealing, the degradation limits the continuous operational days without any annealing in the middle of the beam time. Fig. 3.11 shows the history of the averaged PDE on the readout MPPCs measured in the pilot runs from 2017 to 2019, and its extrapolation to the future. Since the speed of the degradation were measured to be gradually slowing down, there is a large uncertainty in the extrapolation.

In the worst case, the PDE will be zero after a 70–100 day beam usage at the MEG II beam intensity (7×10^7 μ^+stops/s) while the original plan is 120 day data-taking per year. Therefore, detector operation to survive with reasonable performance must be considered. The following three scenarios were considered in [62]. The first scenario is just to shorten the DAQ time per year. The LXe detector can be operated for 60 days with a full PDE recovery at the beginning of the year's DAQ. The second scenario is to reduce the beam intensity by a factor of two, which enables the detector operation for 120 days. Since the accidental backgrounds increase according to the beam intensity, the significance $N_{\mathrm{sig}}/\sqrt{N_{\mathrm{BG}}}$ gets better than the first scenario. In addition, pileup conditions are improved by the reduced beam intensity resulting in a better

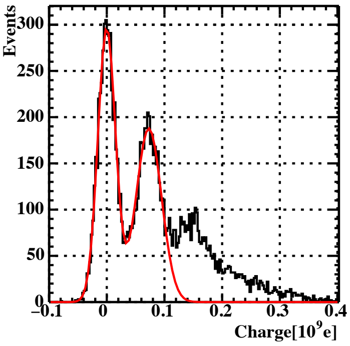


Figure 3.5 Charge distribution of an MPPC obtained using a weak LED light. The red line is a sum of fitted Gaussians [62].

background suppression performance. The third scenario is to have an annealing not only at the begging of each year, but also during the beam time. In total, it will take about 60 days to finish all the annealing processes, and thus the remaining 80 days can be used for the DAQ. This results in a higher statistics and a higher PDE than the former two scenarios. These three scenarios were compared in terms of the branching ratio sensitivity, and it was found that the highest sensitivity is achievable with the second scenario.

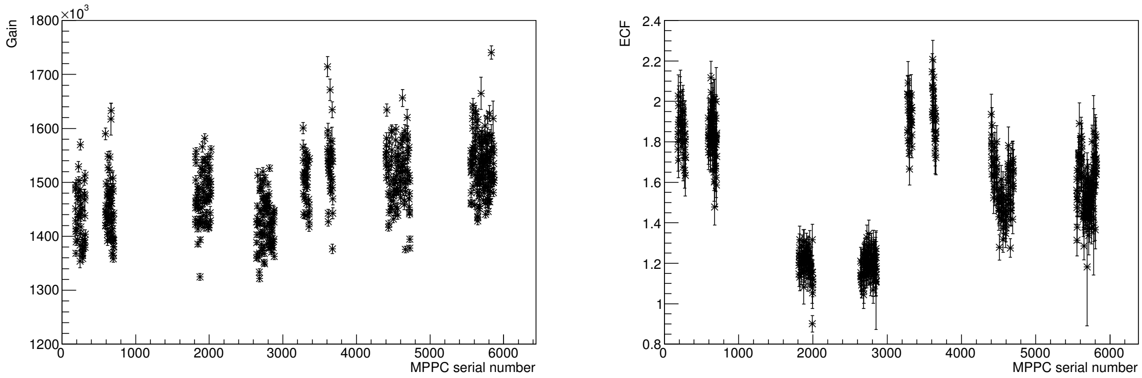
To be optimistic, the PDE can be saturated around 2% based on the observation at the lab[*1]. In this case, the detector is operational even at the MEG II intensity for 120 days though the PDE is kept low after the saturation. This enables to accumulate the highest statistics in the scenarios.

The two scenarios, the halved intensity in the pessimistic case and the MEG II intensity in the optimistic case, are discussed in Sec. 7.3. Note that other scenarios with the intermediate beam intensity are also possible. It can be optimized to maximize the accumulated statistics and detector performance.

Table 3.2 Expected irradiation fluence in run 2019 corresponding to 160 hours in the MEG II beam intensity [62].

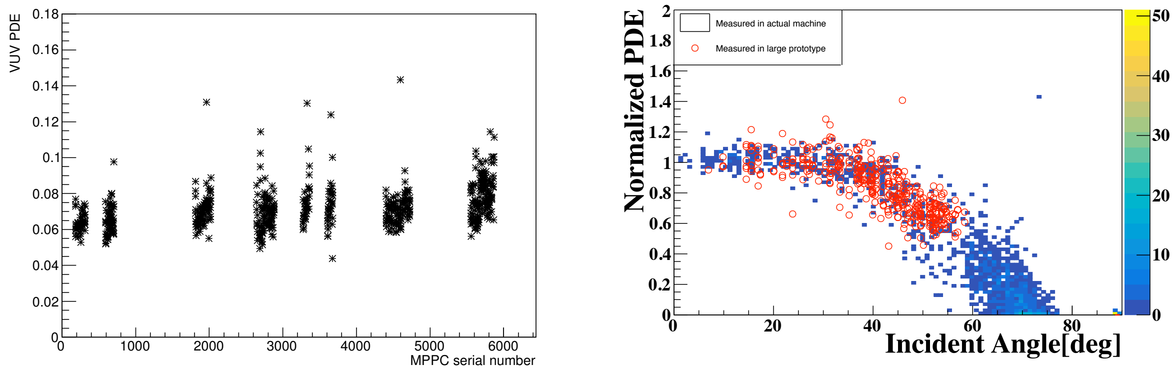
irradiation source	dose/fluence
γ	0.01 Gy
VUV photon	4.6–5.8 $\times 10^{10}$ /mm^2
neutron	2.9 $\times 10^6$ n/cm^2

[*1] The PDE degradation was supposed to be saturated around 6% in an optimistic scenario in [62] based on the observation of a lab test; saturation around 35% for the wavelength of 190 nm had been observed. However, it was found that the saturation level is lower for the wavelength of 150–180 nm with the subsequent measurement (Appendix C).



(a) (b)

Figure 3.6 (a) Gain and (b) ECF of the MPPCs at the over voltage 7 V [62].



(a) (b)

Figure 3.7 (a) PDE of each MPPC at over voltage 7 V measured at the beginning of the run 2019 [62]. (b) MPPC VUV PDE as a function of the incident angle [84]. The individual difference of PDE of each MPPC is corrected using the data point of the smallest incident angle.

3.1.2 Detector calibration

Proton accelerator

A CW proton accelerator is located at the downstream side of the $\pi E5$ beamline (Fig. 3.12). Accelerated protons are injected into a target made of $\mathrm{Li_2B_4O_7}$ placed at the center of the COBRA magnet [48]. The nuclear reactions both of Li and B produce γ-rays: 14.8 MeV and 17.6 MeVγ-ray from ${}^{7}_{3}\mathrm{Li}(p,\gamma)^{8}_{4}\mathrm{Be}$ at $E_p = 440\,\mathrm{keV}$, and 4.4 MeV and 11.7 MeVγ-ray from ${}^{11}_{5}\mathrm{B}(p,\gamma)^{12}_{6}\mathrm{C}$ at $E_p = 163\,\mathrm{keV}$.

The γ-ray with 17.6 MeV from the lithium excitation is used not only for the energy calibration but also for the evaluation energy resolution of the LXe detector since its energy spread is small enough (12 keV). The other γ-ray with 14.8 MeV is not used due to the large decay width and the overlap with a lower energy tail of the 17.6 MeV peak. From the boron excitation, γ-rays of

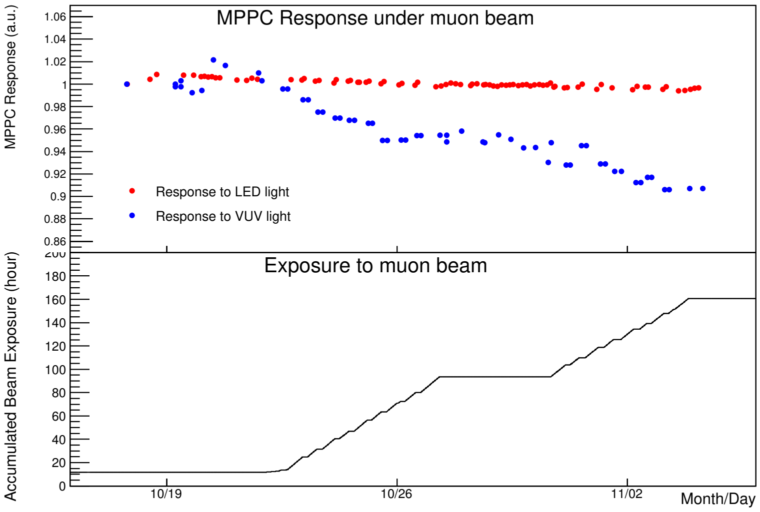


Figure 3.8 Averaged MPPC response to the VUV light from the α sources (blue) and the visible light from the LEDs (red) during the run 2019 (upper), and the accumulated beam usage (lower) [62]. The values are normalized by the first measurement points.

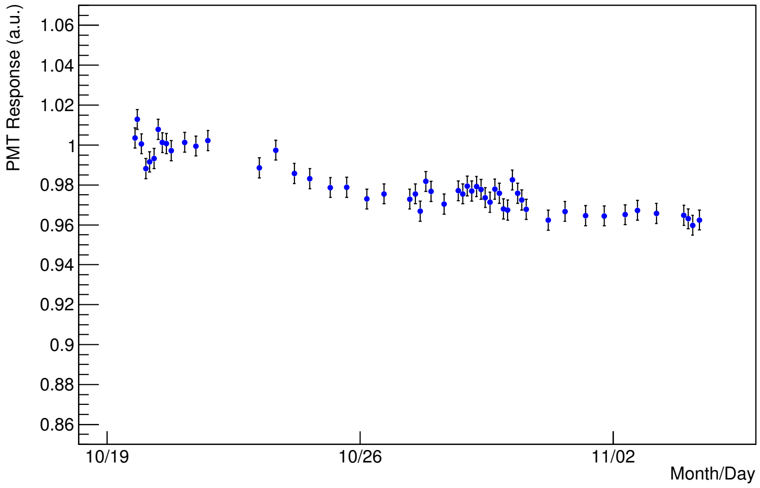
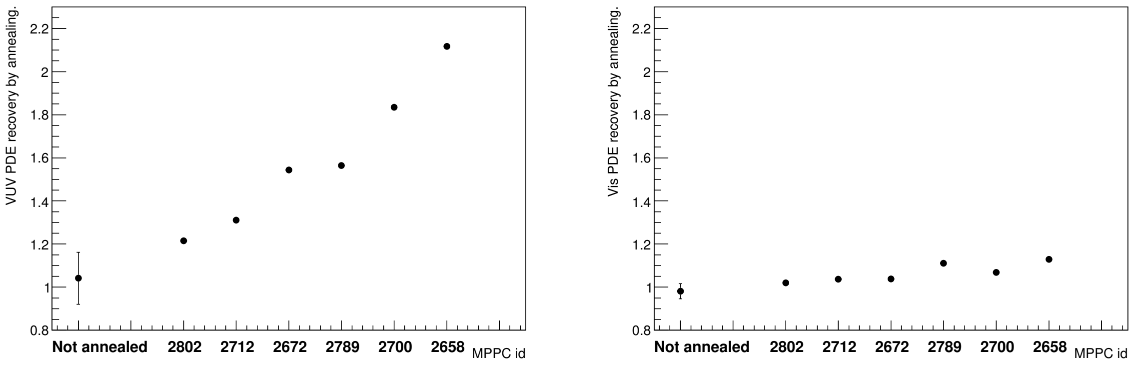


Figure 3.9 Averaged PMT response to the VUV light from the α sources divided by the gain, i.e. $CE \times QE \times LY$, during the run 2019 [62].

4.4 MeV and 11.7 MeV are emitted simultaneously, which can be used for timing alignment of the LXe detector and the pTC.

CEX

The Charge EXchange reaction (CEX) of π^- on proton is used to calibrate energy and timing and to evaluate the performance with quasi-monochromatic γ-rays with energies near the signal γ-ray energy. A π^- beam with the momentum of 70.5 MeV/c is injected into a liquid hydrogen target, and then π^0 produced by the CEX immediately decays to two γ-rays as



(a) (b)

Figure 3.10 Recovery of the PDE (a) for VUV light, (b) for blue LED light due to the annealing [62]. Non-annealed MPPCs are also shown for comparison. The MPPC ids in the x-axis are in an ascending order of the annealing strength. The error bar of the non-annealed MPPCs shows the 1σ spread of the PDE ratio distribution of the non-annealed MPPCs.

Table 3.3 Amplifier gain to be used for different PDEs.

PDE	amplifier gain
$8\% < PDE \leq 22\%$	1
$4\% < PDE \leq 8\%$	2.5
$2\% < PDE \leq 4\%$	5
$0\% < PDE \leq 2\%$	10

$$\pi^- + p \rightarrow \pi^0 + n \rightarrow \gamma + \gamma + n. \tag{3.10}$$

The γ-rays are emitted back-to-back with the energy of 67.5 MeV in the rest frame of the π^0:

$$E_\gamma^{\text{rest}} = \frac{m_{\pi^0}}{2} \simeq 67.5\,\text{MeV}. \tag{3.11}$$

In the laboratory frame, they are boosted by π^0 momentum and their energies can be written by

$$E_\gamma = \frac{m_{\pi^0}}{2}\gamma(1 \pm \beta\cos\theta_{\text{rest}}), \tag{3.12}$$

where β is the velocity of the π^0 ($\simeq 0.2$), γ is the Lorentz factor and θ_{rest} is the emission angle in the rest frame. Therefore, the γ-ray energy ranges from 54.9 MeV to 82.9 MeV depending on the angle. The quasi-monochromatic γ-rays of 54.9 MeV and 82.9 MeV are available by selecting back-to-back events. The energy spread can be narrower than 0.2% if the events with opening angle above 175° are selected, which is sufficient to measure the energy resolution of 1%.

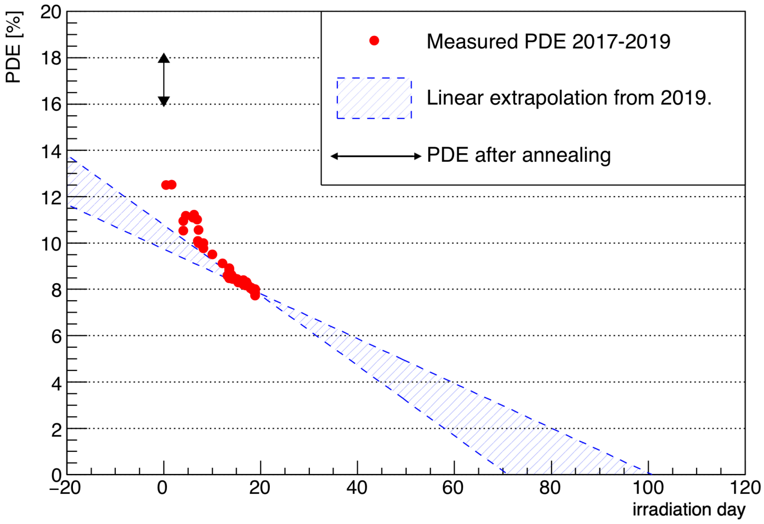


Figure 3.11 Measured PDE history measured in the pilot runs from 2017 to 2019, and its extrapolation. The x-axis is the usage time of the MEG II beam. Blue band shows a linear extrapolation of the $9 \pm 2\%$ degradation observed in the run 2019 [62].

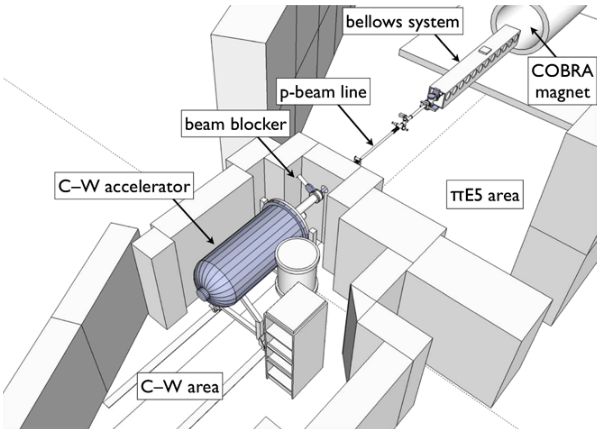


Figure 3.12 Location of the CW proton accelerator [90].

In order to select such events, a reference counter is placed in the opposite side of the LXe detector (Fig. 3.13). It is composed of a timing counter and a calorimeter. The calorimeter consists of a 4 × 4 arrayed Bismuth Germanium Oxide (BGO) crystals sized $46 \times 46 \times 200\,\mathrm{mm}^3$ (Fig. 3.14(a)). Each crystal is read out with a PMT (H8409-70, Hamamatsu Photonics) attached on the back (Fig. 3.14(b)). The timing counter is placed in front of the calorimeter (Fig. 3.15). It consists of two plastic scintillator counters, which are placed to be orthogonal to each other, and a lead converter with 4 mm thickness placed in front of them. The thickness of the converter was optimized to maximize the efficiency based on the simulation. The scintillators (EJ-230) are sized $80 \times 175 \times 5\,\mathrm{mm}^3$, and read out with 16 SiPMs (S13360-3050PE, Hamamatsu Photonics) coupled to each side. To reduce the number of readout channels, the adjacent four SiPMs are connected in series, and thus 16 channels are read out in total.

The reference counter is mounted on a mover to scan the whole LXe detector acceptance

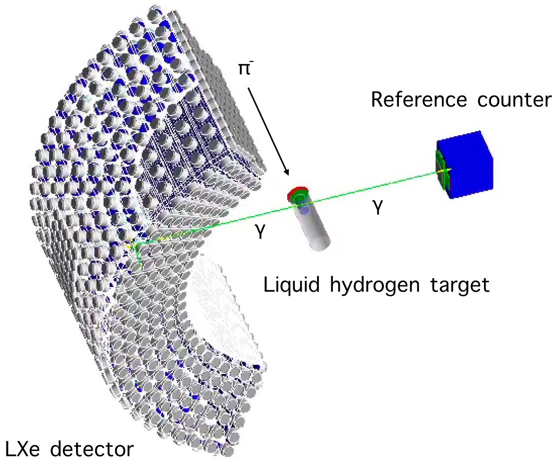


Figure 3.13 Overview of the CEX measurement.

(a) (b)

Figure 3.14 BGO calorimeter for CEX calibration [44]. (a) 16 crystals are installed in a detector holder. (b) Each BGO crystal is read out with a PMT.

(Fig. 3.16). It can be moved in ϕ (vertical) and z (horizontal) directions, and rotated in order to face the liquid hydrogen target.

In the run 2020, the CEX data-taking was performed. The LH_2 target was installed to $z \sim 7\,\mathrm{cm}$ due to a mechanical conflict though it would be the center. The position of the reference counter was shifted by 7 cm accordingly. In this thesis, only the timing calibration and resolution are discussed with the CEX data.

3.2 Alignment of MPPCs

The γ-ray hit position is reconstructed using the detected charge distribution and the positions of MPPCs, and thus the MPPCs must be well aligned. In order to achieve the aimed position

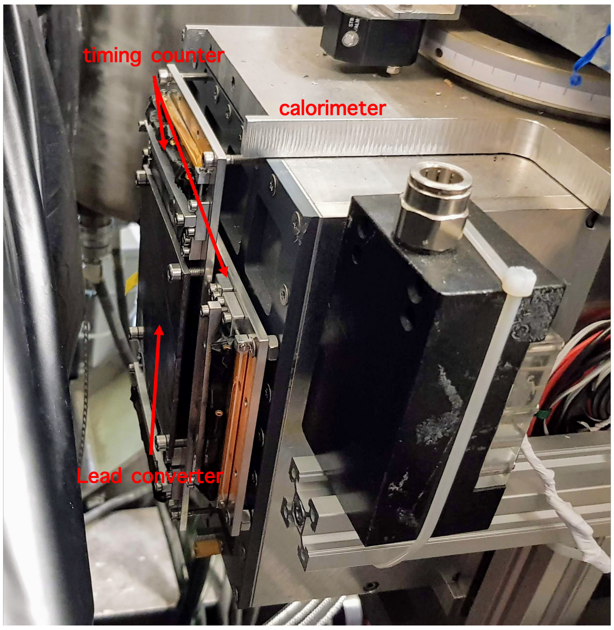


Figure 3.15 Timing counter for CEX calibration.

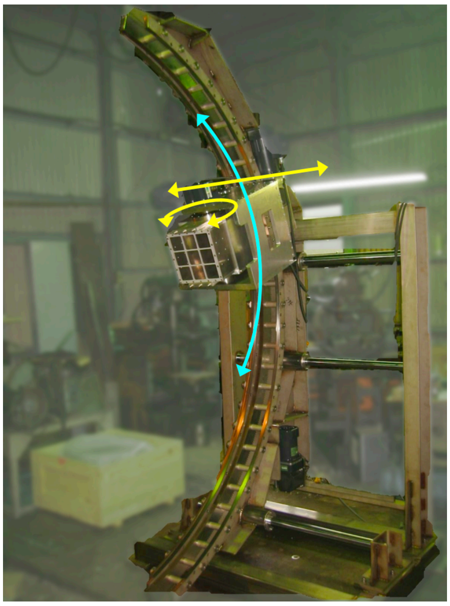


Figure 3.16 Mover for the reference counter. The arrows shows the degree of freedom [44].

resolution of 2 mm, a precision of 0.5 mm is required for the measurement of the MPPC positions. This was realized by combining two independent measurements.

3.2.1 Measurement with laser scanner

The surfaces of the installed MPPCs were surveyed with a three-dimensional laser scanner (FARO Edge ScanArm) during the detector construction (Fig. 3.17(a)). Edges of MPPC chips and MPPC packages can be seen in the data as shown in Fig. 3.17(b). The gaps between the four chips were fitted, and the position and the normal vector of each MPPC were reconstructed. Since the space between the inner and the outer faces of the cryostat is limited, the data quality only for ∼10% of MPPCs was good enough to reconstruct the positions. Therefore, those of the other MPPCs were interpolated by the reconstructed ones assuming the MPPCs on the same CFRP are well aligned with a good precision.

3.2.2 Measurement with γ-rays

Since the positions of the MPPCs are affected by the thermal contraction of the PCBs and the CFRPs, a complementary measurement was performed at the LXe temperature. The γ-rays of 120 keV from a ^{57}Co (Fig. 3.18(a)) were injected into the LXe detector and interact with the LXe within 3 mm. The γ beam was collimated with a brass collimator to be a size of $1.5 \times 40\,\mathrm{mm}^2$ on the MPPC surface. The whole inner face of the LXe detector was scanned by changing the position and the direction of the beam with moving stages. The movement was monitored with an optical laser and a bubble level [91], which enables to align the beam with precisions of 30 μm in z and 80 μm in ϕ direction. Since the scintillation light is localized thanks to the short attenuation length for 120 keV γ-ray, an excess of the event rate can be observed when a γ-ray hits above the MPPC (Fig. 3.18(b)), from which the MPPC position is reconstructed.

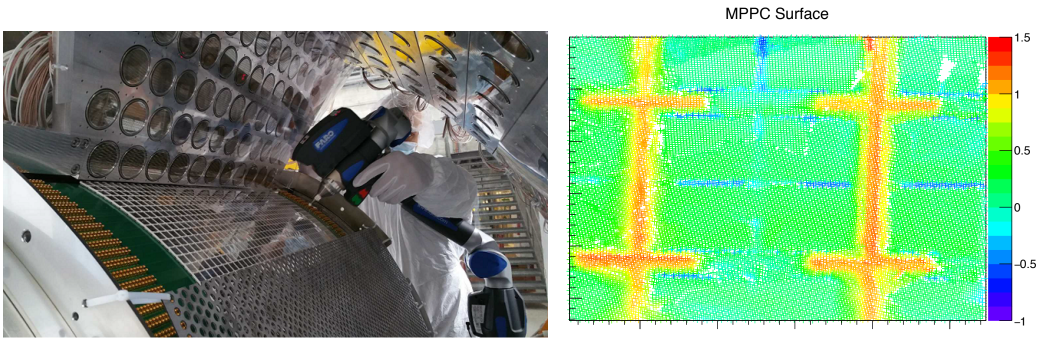
3.2.3 Combined analysis

A three-dimensional position of each MPPC at the LXe temperature was calculated by combining the two measurements; the laser scanner measurement provides a three-dimensional position at room temperature, and the γ-ray measurement provides a two-dimensional position at the LXe temperature. The reconstructed position in the γ-ray measurement was fitted by the measured position with the laser scanner $\vec{x}_{\mathrm{laser}}$ with a transformation, Euler rotation $R(\alpha, \beta, \gamma)$ and offset $\vec{c}_{\mathrm{offset}}$, and a thermal contraction a as

$$\vec{x'}_{\mathrm{laser}} = (1-a)R(\alpha, \beta, \gamma)\vec{x}_{\mathrm{laser}} + \vec{c}_{\mathrm{offset}}.$$

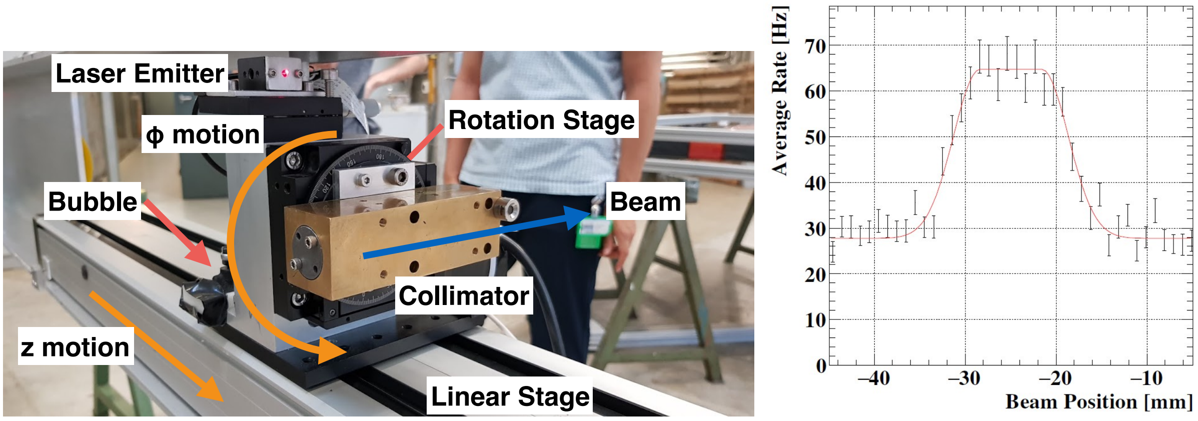
The fitted thermal contraction rate was found to be consistent with that of the detector material (16 ppm/K). The fitted laser positions and the γ-ray positions were in agreement within an accuracy of 0.3 mm, which fulfills the requirement of 0.5 mm.

The results suggest that the MPPCs deviate from the designed positions from −4.1 mm to +0.3 mm along with the r direction and −3.1 mm to +0.5 mm along with the z direction mainly due to a curvature of the CFRPs (Fig. 3.19). In this thesis, the correction was not applied, but the effect is limited since only a part of channels, which concentrate in the center of one CFRP, is used.



(a) (b)

Figure 3.17 (a) MPPC position measurement by a laser scanner [91]. (b) Example of measured MPPC surface by laser scanner [84]. Color shows the coordinate of perpendicular direction. The red region corresponds to the ceramic package of each MPPC, and the blue region corresponds to the gap between four chips on a package (Fig. 2.33).



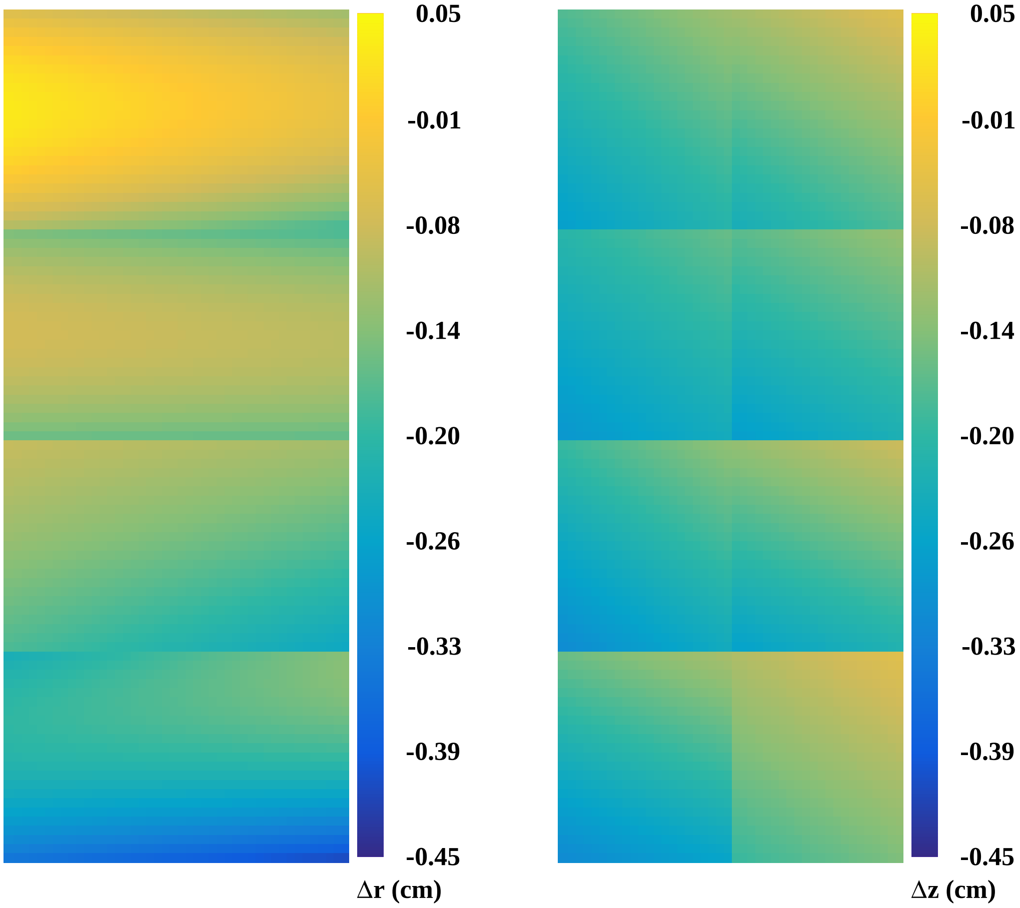
(a) (b)

Figure 3.18 (a) Apparatus for MPPC alignment with γ-rays [84]. (b) The γ-ray event rate of an MPPC as a function of z position of the beam [91].

3.3 Waveform analysis

Signals of the photosensors are recorded with the waveform digitizer. Fig. 3.20 shows typical waveforms of the LXe detector for a γ-ray event. The information of the signals such as timings and charges is extracted by analyzing them.

The pulse timing is defined by the timing when the waveform crosses a certain threshold voltage given by the product of the signal amplitude and a constant fraction. This method has an advantage of the time walk suppression compared to using a constant threshold. The constant fraction is optimized to have the best timing resolution for the MPPCs and the PMTs, individually. The resolution can be better with a lower threshold due to the smaller statistical fluctuation for the first arriving photon timing while it can be more sensitive to the noise and the transit time spread. The optimal fraction was measured to be 5% for the MPPCs and 10% for the PMTs.

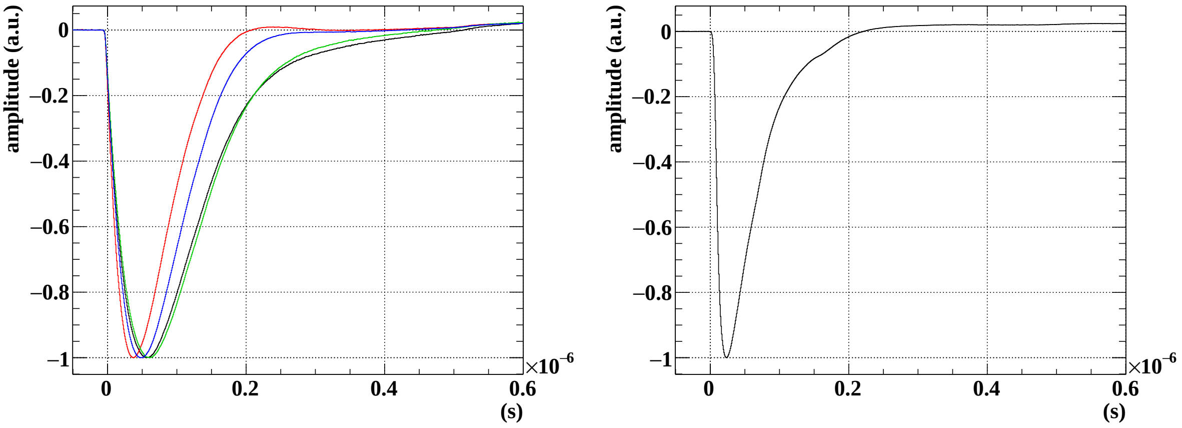


(a) (b)

Figure 3.19 Deviation of MPPC positions from the design along with (a) r and (b) z directions. The four gaps along with the x-axis are due to the different CFRPs. The central slit along with the y-axis in (b) is due to the different PCBs.

Since the timing extraction is sensitive to high frequency noises, a noise reduction algorithm is appleid to the original waveform. Such noises come from the system clock of the readout electronics, they are synchronized to the phase of the clock. Therefore, they can be reduced by subtracting template waveforms generated using random trigger events. For further reduction, a moving average filter of adjacent three points is also applied for the MPPC waveforms. The application of the digital lowpass filter is a trade-off between a reduction of the high frequency noise and a smearing of the sensor waveforms, and thus the strength of this filter was optimized to get the best timing resolution.

The charge of each photosensor is obtained by integrating each waveform over 300 ns around a common reference timing. The reference timing is extracted from a sum waveform of the MPPCs based on the constant fraction method. The baseline is calculated by taking a mean of amplitudes before the charge integration region. The charges are converted to the number of photons using the calibration parameters.



(a) (b)

Figure 3.20 Typical waveforms of (a) MPPC and (b) PMT. Each color in (a) corresponds to each production lot: Lot A (black), Lot B (red), Lot C (green) and Lot D (blue). The variety of the MPPC waveforms comes from the different correlated noise probability (Sec. 2.1.3.2).

3.4 Position resolution

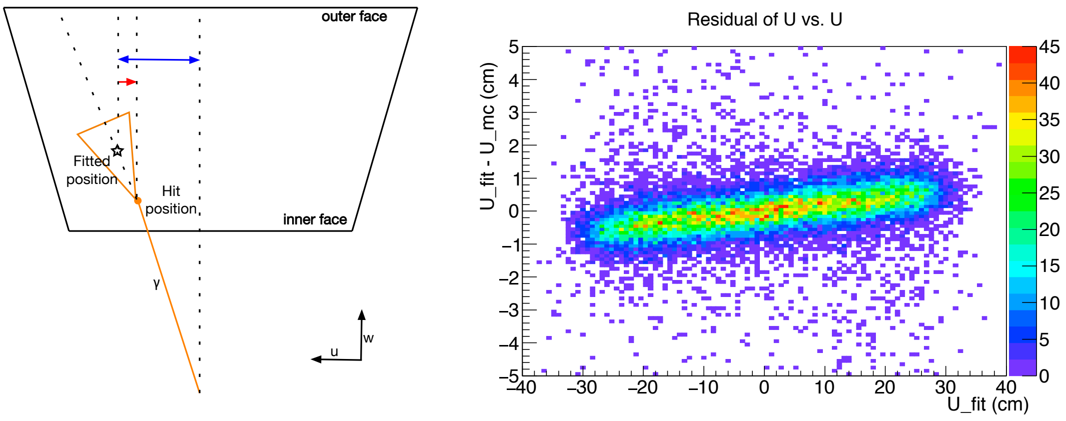
The hit position of the γ-ray is reconstructed using the light distribution on the inner face, and thus the improved granularity thanks to the replacement of the PMTs on the inner face to the MPPCs is expected to make the position resolution for the shallow events better. This was verified by taking γ-ray data with a collimator and comparing the reconstructed position distribution with MC.

3.4.1 Position reconstruction

The position of the γ-ray conversion is reconstructed from the light distribution on the inner face. Firstly, initial position is computed from the mean of MPPC positions weighted by the detected number of photons N_{pho}. Then, a fitting is performed to minimize the difference between the detected number of photons and the expected one from the solid angle Ω of each MPPC from the γ-ray hit position $\vec{x}$ assuming the scintillation photons are isotropically emitted from one point as

$$\chi^2 := \sum_{\mathrm{MPPC}} \left(\frac{N_{\mathrm{pho}} - C \times \Omega(\vec{x})}{\sigma(N_{\mathrm{pho}})} \right)^2 ,$$

where C is a normalization of the light distribution, which is estimated in the fitting as a nuisance parameter. $\sigma(N_{\mathrm{pho}})$ is the uncertainty of the N_{pho} of each channel, which is given by the statistical fluctuation of the number of photoelectrons as



(a) (b)

Figure 3.21 (a) Schematic of the u correction [62]. A bias in the reconstruction (red arrow) is corrected by the reconstructed γ-ray position (blue arrow). (b) Correlation between the bias and the reconstructed position in MC.

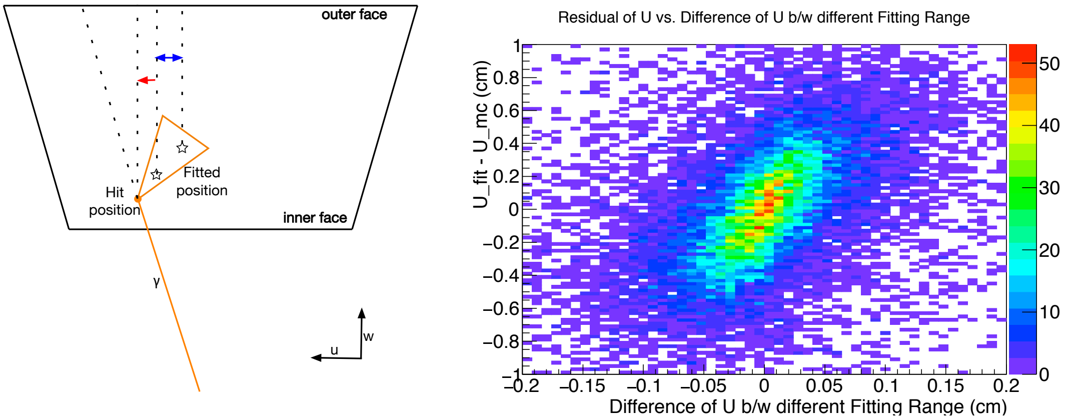
$$\begin{aligned}\sigma(N_{\mathrm{pho}}) :&= \frac{\sigma(N_{\mathrm{phe}})}{PDE},\\ &= \frac{\sqrt{N_{\mathrm{phe}}}}{PDE}.\end{aligned}$$

The estimated position is biased to the direction of shower development due to the assumption that all the scintillation photons come from one point since they are emitted from each point of the electromagnetic shower in reality. To reduce this bias, only the limited number of MPPCs around the hit position is used in the fitting. In addition, two corrections are applied whose parameters are derived from the MC simulation. The first type corrects a global bias with respect to u direction. The direction of the shower development tends to be similar to the γ-ray incident direction. This causes the reconstructed u position to be larger when a γ-ray hits at large $|u|$ position as shown in Fig. 3.21. This dependence is corrected as a function of the reconstructed u position. The second type corrects an effect of even-by-event fluctuation of the shower direction. Even for the events with the same hit position, the shower direction can be different due to the fluctuation. In order to correct the dependence on the shower direction, it is reconstructed by performing the position fitting for several fitting regions utilizing the fact that the reconstruction bias becomes larger for a wider fitting region (Fig. 3.22).

3.4.2 Performance Measurement

Measurement with 17.6 MeV γ-ray

The position resolution was evaluated with the CW-Li γ-rays in the run 2018. A 25 mm thick lead collimator was placed in front of the LXe detector, which had several slits of 5 mm width at



(a) (b)

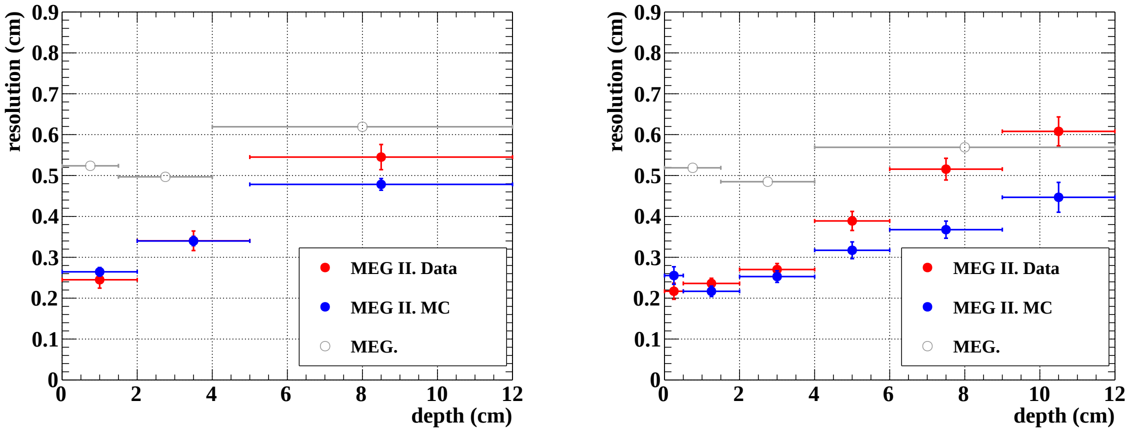
Figure 3.22 (a) Schematic of the shower correction [62]. A bias in the reconstruction (red arrow) is corrected by the difference of the reconstructed position using different fitting ranges (blue arrow). (b) Correlation between the bias and the difference of the reconstructed position in MC.

intervals of 50 mm. The resolution was evaluated with the width of the reconstructed position distribution of the slits. An excess of the event rate corresponding to each slit was fitted with a simulated hit position distribution smeared by a Gaussian. The resolution was defined by the sigma of the Gaussian.

The γ-rays from the RMD were not used because the spot size of the muon beam is 1.2 cm for the vertical direction and 4 cm for the horizontal direction, which can smear the position distribution to the extent being not negligible. That of the proton beam is small enough, which is 0.6 mm for the vertical direction and 0.8 mm for the horizontal direction. In addition, the position resolution does not depend on the energy since the resolution is not dominated by the statistical fluctuation according to MC.

The estimated resolution is shown in Fig. 3.23 as a function of the reconstructed conversion depth. The improvement of the resolution in the shallow region is observed as expected. On the other hand, that in the deep region is measured to be slightly worse than the expectation. This can be because of a contribution of the noise, but it has not yet been understood quantitatively.

An angular dependence of the MPPC PDE which cannot be explained by the Fresnel reflection was found as described in Sec. 3.1.1, and the PDE for the larger incident angle is lower than that for the smaller incident angle. This leads to a bias of the position reconstruction; the conversion depth is biased to be shallower since the light distribution on the MPPCs becomes sharper. Although the bias can be corrected by the measured angular dependence, it can become systematic uncertainty if we misunderstood the angular dependence, and thus its effect was investigated. As a result, it was found that even if the angular dependence is totally misunderstood, the angular resolution of $\Theta_{e^+\gamma}$ increases only by 1.2%, which results in 0.6% degradation of the branching ratio sensitivity at most.



(a) (b)

Figure 3.23 Position resolution for 17.6 MeV γ-ray as a function of the conversion depth for (a) horizontal u-direction, and (b) vertical v-direction. The resolutions in MEG are also shown for comparison.

The effect of the position resolution on the branching ratio sensitivity was investigated. For the calculation, the resolutions except for the position resolution were fixed to the measured resolution in the pilot runs. The improvement in the shallow region leads to a 30% better sensitivity than MEG. On the other hand, the sensitivity is worse by 4% compared to the MC expectation due to the worse resolution in the deep region. In total, therefore, a 26% sensitivity improvement is achieved compared to MEG.

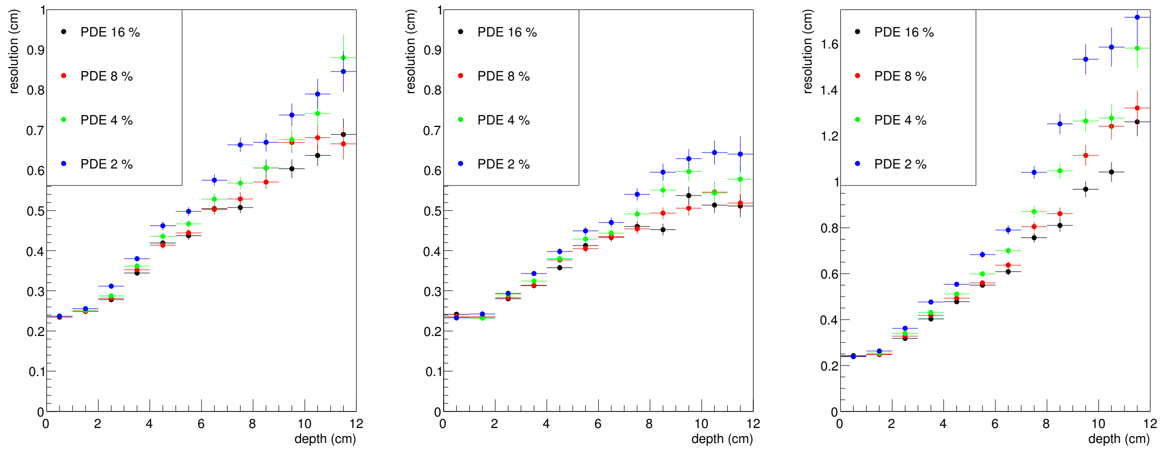
Effect of PDE degradation

The increase of the statistical fluctuation due to the reduced PDE can increase the fitting uncertainty for the position reconstruction, which results in the worse position resolution. Fig. 3.24 shows the position resolution simulated assuming various MPPC PDEs. The degradation of the resolution was observed in the deep region, where the fitting uncertainty can be larger due to the smaller number of photoelectrons on the inner face. The noise effect is negligible for the position resolution since it was proved that the position resolution measured with the 17.6 MeV γ-ray from the CW-Li at the 7% PDE is not affected, whose signal-to-noise ratio is equivalent to that at the 2.3% PDE for the signal γ-ray.

Fig. 3.25 shows the effect on the branching ratio sensitivity from the degradation of the position resolution. It was found to be less than a few percent.

3.5 Timing resolution

The hit timing of a γ-ray is reconstructed using the extracted timings of each channel. An improvement of the timing resolution is also expected thanks to a better estimation of a correction parameter resulting from the better position resolution. The timing resolution was measured with a high energy γ-ray from the CEX reaction.



(a) (b) (c)

Figure 3.24 Position resolution in the simulation as a function of the conversion depth at various PDE for (a) u, (b) v, and (c) w direction [62].

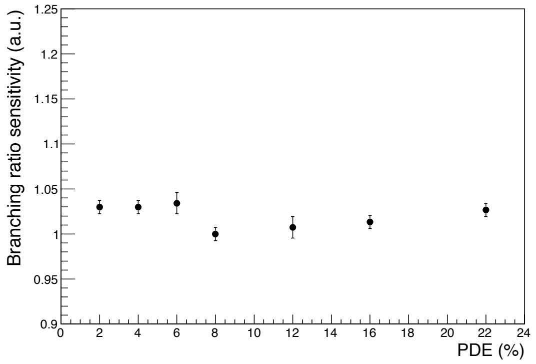


Figure 3.25 Effect on the branching ratio sensitivity from the position resolution degradation at different MPPC PDEs [62]. The values are normalized by that of 8% PDE.

3.5.1 Timing reconstruction

The hit timing of a γ-ray is reconstructed based on the weighted mean of the sensor timings. Firstly, initial timing is extracted from the channel whose detected number of photons is the maximum on the inner face. Then, a fitting is performed to minimize the difference between the hit timing t_γ and each photosensor timing as

$$\begin{aligned}
\chi^2 &:= \sum_{\mathrm{pm}} \chi^2_{\mathrm{pm},i}, \\
\chi^2_{\mathrm{pm},i} &:= \frac{(t_{\mathrm{pm},i} - t_{\mathrm{corr},i} - t_\gamma)^2}{\sigma^2_{\mathrm{pm},i}}, \\
t_{\mathrm{corr},i} &:= t_{\mathrm{prop},i} + t_{\mathrm{walk},i} + t_{\mathrm{offset},i}.
\end{aligned}$$

Here, $t_{\mathrm{pm},i}$ is the extracted timing of each photosensor. $t_{\mathrm{corr},i}$ is a correction term, which consists of the propagation time from the reconstructed γ-ray hit position to each photosensor $t_{\mathrm{prop},i}$, the time walk effect $t_{\mathrm{walk},i}$ and the time offset $t_{\mathrm{offset},i}$ coming from the difference of the cable length and electronics. $\sigma_{\mathrm{pm},i}$ is the precision of the timing information. The fitting is iterated until χ^2/ndf gets lower than 10. To eliminate the pileup effect, the photosensors whose $\chi^2_{\mathrm{pm},i}$ are more than 10 are removed from the reconstruction. The rests can contribute to the reconstruction, but only a part of them has a sufficient timing resolution to affect the reconstruction in practice. Typically, contributions from channels below 100 photoelectrons are negligible though channels above 50 photoelectrons are included.

3.5.2 Performance Measurement

The timing resolution was evaluated using the γ-rays from the CEX reaction (Sec. 3.1.2) in 2020. The reference detector was placed at the opposite side of the readout area of the LXe detector.

Performance of timing counter

The reference timing was measured with the timing counter. It is defined by the mean of the timings measured with the two scintillator counters, where the timing of each counter is defined by the mean of the timings of the readout channels as

$$\begin{aligned} t_{\mathrm{ref}} &= \frac{t_{\mathrm{front}} + t_{\mathrm{back}}}{2}, \\ t_{\mathrm{front,back}} &= \frac{1}{N_{\mathrm{ch}}} \sum_{i \in \mathrm{front,back}} t_i, \end{aligned} \tag{3.13}$$

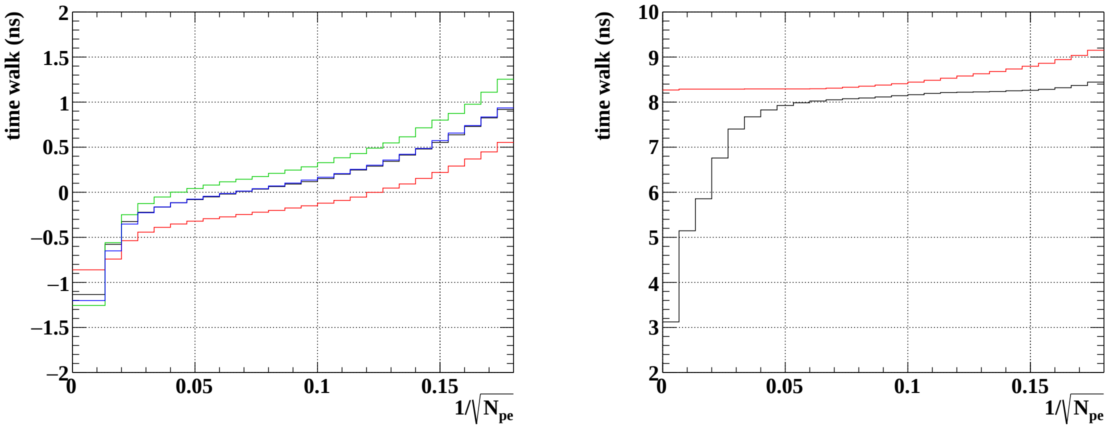
where t_{ref} is the reference timing, t_{front} and t_{back} is the timing of the scintillator counter in the front and the back side, respectively, and t_i is the timing of each readout channel. The number of the channels N_{ch} is eight for each scintillator.

The hit position x_{ref} on the timing counter is also computed by the timing difference of the channels attached on each side, t_{left} and t_{right}, as

$$\begin{aligned} x_{\mathrm{ref}} &= (t_{\mathrm{left}} - t_{\mathrm{right}}) \times v_{\mathrm{eff}}, \\ t_{\mathrm{left,right}} &= \frac{2}{N_{\mathrm{ch}}} \sum_{i \in \mathrm{left,right}} t_i, \end{aligned} \tag{3.14}$$

where v_{eff} is the effective velocity of scintillation light, which is measured to be $\sim 12\,\mathrm{cm/ns}$. Since the two scintillators are arranged to be orthogonal to each other, the positions along with the vertical and horizontal axis can be obtained. The events in which the reconstructed position satisfies $|x_{\mathrm{ref}}| < 4\,\mathrm{cm}$ for both is used for the analysis. More than a half of the total energy deposit in the center 2×2 crystals was also required for the BGO calorimeter.

The timing resolution of the timing counter $\sigma_{t_{\mathrm{ref}}}$ was obtained during the CEX data-taking



(a) (b)

Figure 3.26 Measured time walk of (a) MPPC and (b) PMT. Each color corresponds to each group; MPPC Lot A (black), Lot B (red), Lot C (green) and Lot D (blue) in (a), and PMT outer (black) and lateral (red) in (b).

by taking the timing difference of the two counters as

$$\sigma_{t_{\mathrm{ref}}} = \frac{\sigma(t_{\mathrm{front}} - t_{\mathrm{back}})}{2}. \tag{3.15}$$

It was estimated to be 38(0.4) ps.

Calculation of calibration parameters

The calibration parameters were calculated using the reference timing. The timing resolution of each channel σ_{pm} and the time walk t_{walk} were estimated from the spread and the mean of $t_{\mathrm{pm}} - t_{\mathrm{prop}} - t_{\gamma,\mathrm{exp}}$, where $t_{\gamma,\mathrm{exp}}$ is the γ-ray hit time on the LXe detector expected from the timing counter and the time-of-flight, i.e. $t_{\gamma,\mathrm{exp}} = t_{\mathrm{ref}} - TOF_{\mathrm{ref}} + TOF_{\gamma}$. These TOF are calculated from the reconstructed positions on each detector assuming the vertex is the center of the target. The propagation time t_{prop} is calculated from the distance between the reconstructed position and each photosensor divided by the effective velocity of light in LXe.

Fig. 3.26 and Fig. 3.27 show the obtained t_{walk} and σ_{pm} as a function of an inverse of square-root of the number of photoelectrons, respectively. These functions are extracted for each production lot for the MPPCs because they have different shapes of a waveform due to the different after-pulsing probabilities. The PMTs are grouped by those on the outer face and the others since their time walks and resolutions were found to be different. The mean of this distribution has a channel-by-channel difference, which is interpreted as a time offset of each channel t_{offset}. The distribution of the measured time offset for the PMTs was found to be wider than that for the MPPCs as shown in Fig. 3.28. This is likely due to the individual difference of the cable length since the cables for the PMTs are reused from MEG with some modifications.

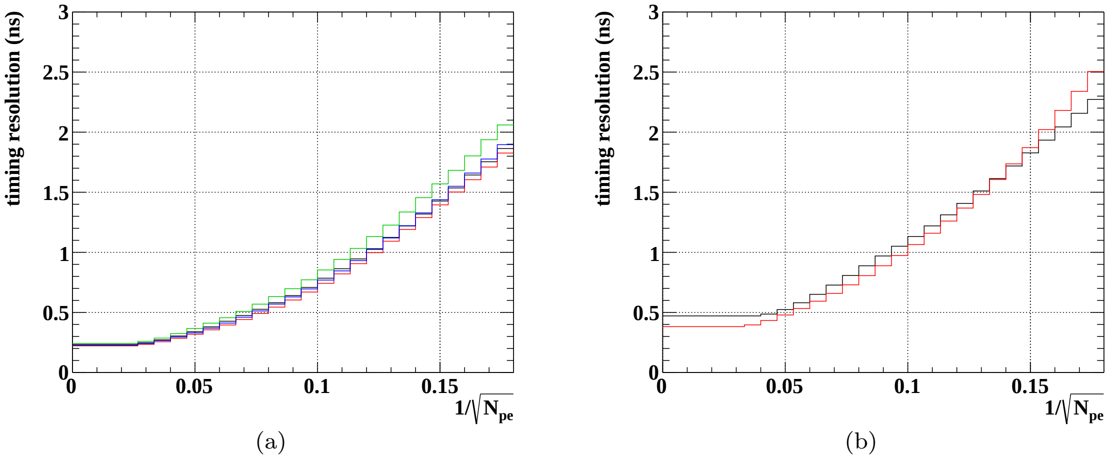


Figure 3.27 Measured timing resolution of (a) MPPC and (b) PMT. Each color corresponds to each group; MPPC Lot A (black), Lot B (red), Lot C (green) and Lot D (blue) in (a), and PMT outer (black) and lateral (red) in (b).

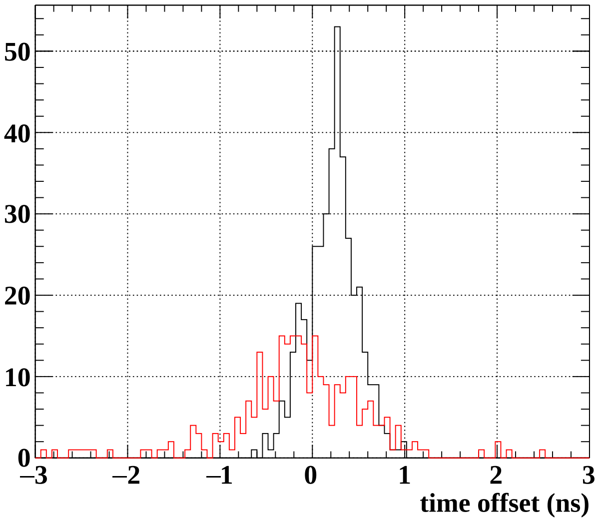


Figure 3.28 Measured time offset of the MPPCs (black) and the PMTs (red).

Detector resolution

There are two variables which describe the detector timing resolution. One is the absolute resolution which is a spread of the difference between the reconstructed γ-ray timing and the reference timing. This is the resolution directly related to the physics analysis, and can be measured in the CEX run using the reference timing as the true value. The other is the intrinsic timing resolution estimated from an even-odd analysis. In the even-odd analysis, the photosensors are divided into two groups, e.g. sensors with even channel numbers and odd channel numbers (Fig. 3.29), and the timing is reconstructed independently for both groups. The intrinsic resolution is estimated from the difference of them. This resolution can be better than the absolute one since some effects fluctuating the reconstructed timing are canceled by the even-odd subtraction, such as a precision of the propagation time given by the hit position resolution and a coherent noise from the electronics.

The absolute timing resolution is calculated by subtracting possible uncertainties from the timing difference between the LXe detector and the timing counter as

$$\begin{aligned}\sigma_{\mathrm{t}_{\gamma}} &= \sigma(t'_{\gamma} - t'_{\mathrm{ref}}) \ominus \sigma_{\mathrm{t}_{\mathrm{ref}}} \ominus \sigma_{\mathrm{t}_{\mathrm{vertex}}},\\ t'_{\gamma} &= t_{\gamma} - TOF_{\gamma},\\ t'_{\mathrm{ref}} &= t_{\mathrm{ref}} - TOF_{\mathrm{ref}}. \end{aligned} \tag{3.16}$$

The contribution from the vertex uncertainty $\sigma_{\mathrm{t}_{\mathrm{vertex}}}$ is estimated to be 48(2) ps using MC from the measured beam size of 6.5 mm and 5.5 mm along with the x and y direction. Fig. 3.30(a) shows the timing difference between the two detectors after the TOF correction. The absolute timing resolution was evaluated to be 81(3) ps for 54.9 MeV. The intrinsic timing resolution was also evaluated using the same data to be 38(0.7) ps (Fig. 3.30(b)), which is consistent with that measured using background γ-rays in the pilot run 2019 [62].

The measured resolutions and their MC expectation are summarized in Table 3.4. To make the simulation realistic, the noise obtained by firing trigger randomly was added to the simulated waveforms. In addition, precisions of the timing calibration of the electronics and the synchronization between the WDBs were added based on the measured values, 25 ps and 45 ps, respectively. The measured absolute resolution was found to be much worse than the MC expectation while the intrinsic resolution was consistent with each other. As a result, the discrepancy between the absolute and intrinsic resolution, 72(3) ps, was found to be much larger than the expectation, 30 ps.

The timing resolution for 82.9 MeV γ-rays is also evaluated. The absolute timing resolution was 80(3) ps while the intrinsic resolution was 31(0.4) ps, which is slightly better resolution than that for 54.9 MeV due to the higher photoelectron statistics. The difference between the absolute and intrinsic resolution was 73(3) ps, which is consistent with that observed for 54.9 MeV. This indicates the term causing the additional deviation does not depend on the photoelectron statistics.

Fig. 3.31 shows the timing resolutions as a function of the reconstructed conversion position. Both of them have almost flat distributions and the discrepancy remains independent of the position. Fig. 3.32 shows the timing resolution as a function of the number of photoelectrons. The improvement of the resolution due to the increased number of photoelectrons was observed for both resolutions.

There are two types of candidates to explain the discrepancy: the effect canceled by the even-odd subtraction and the effect added to the calculation of the absolute resolution. In the former type, the precision of the propagation time, the coherent noise effect, the precision of synchronization are included. Indeed, however, these effects are included in the simulation, and thus they are almost excluded. There is also possibility that the analysis causes some kinds of bias which can be reduced by the subtraction. An overestimation of the timing resolution of the timing counter belongs to the latter case. However, it is not likely since its resolution was also measured with a ^{90}Sr source in advance, and the result was consistent with each other. There is another possibility that the estimation of the vertex uncertainty is wrong. Since its estimation totally relies on the measured beam size and the simulation, it can be underestimated due to an error of either. Fig. 3.33 shows a relation between the beam size and the simulated vertex contribution. It is obvious that the vertex contribution can be easily misunderstood if there is an error in the measured beam size. Indeed, the measured beam size in 2020, $(\sigma_x, \sigma_y) = (6.5\,\mathrm{mm}, 5.5\,\mathrm{mm})$, is much smaller than that measured in MEG, $(\sigma_x, \sigma_y) = (1\,\mathrm{cm}, 1\,\mathrm{cm})$, and if the measured value is wrong for some reasons and the true size is the same as that in MEG, the observed discrepancy can be almost explained. Therefore, it turned out to be important to estimate the vertex contribution in another way in order to certify the estimation.

A dedicated data-taking is planned to measure the vertex contribution directly. The measurement uses γ-rays from the CEX. An additional timing counter whose design is the same as the reference timing counter is placed in front of the LXe detector and the timing of one of the γ-rays is measured with it instead of the LXe detector. Since we know the timing resolutions of both timing counters, the vertex contribution can be estimated in the same way as Eq. (3.16). This measurement will be performed in the CEX run 2021, and the absolute timing resolution will be re-evaluated using the measured vertex uncertainty.

Note that the $t_{\mathrm{e}^+\gamma}$ resolution for the $\mu^+ \to \mathrm{e}^+\gamma$ analysis will be evaluated from the RMD data; a peak in the timing difference between the positron timing measured with the pTC and the γ-ray timing measured with the LXe detector is fitted with a double Gaussian. This allows us to check the consistency of the measured resolutions and gain information about the cause of the unknown contribution: whether it derives from the systematics of the CEX measurement or whether it is related to the detector performance itself.

The effect on the branching ratio sensitivity is evaluated. If the unknown contribution is explained by the error of the vertex uncertainty estimation, a timing resolution of $\sim 50\,\mathrm{ps}$ is achieved, which results in the sensitivity improvement of a few percent from MEG. If the unknown contribution remains even with the improved estimation of the vertex contribution, the timing resolution is worse than the MC expectation by $27\,\mathrm{ps}$ resulting in a 13% worse sensitivity.

Table 3.4 Comparison of the detector timing resolution between MC and data.

Variable	MC	data
Absolute resolution	55 ps	81 ps
Intrinsic resolution	40 ps	38 ps

Effect of PDE degradation

The timing resolution can be worsen both by the lower signal-to-noise ratio and the larger statistical fluctuation, which are expected at lower PDE. The former contribution can be com-

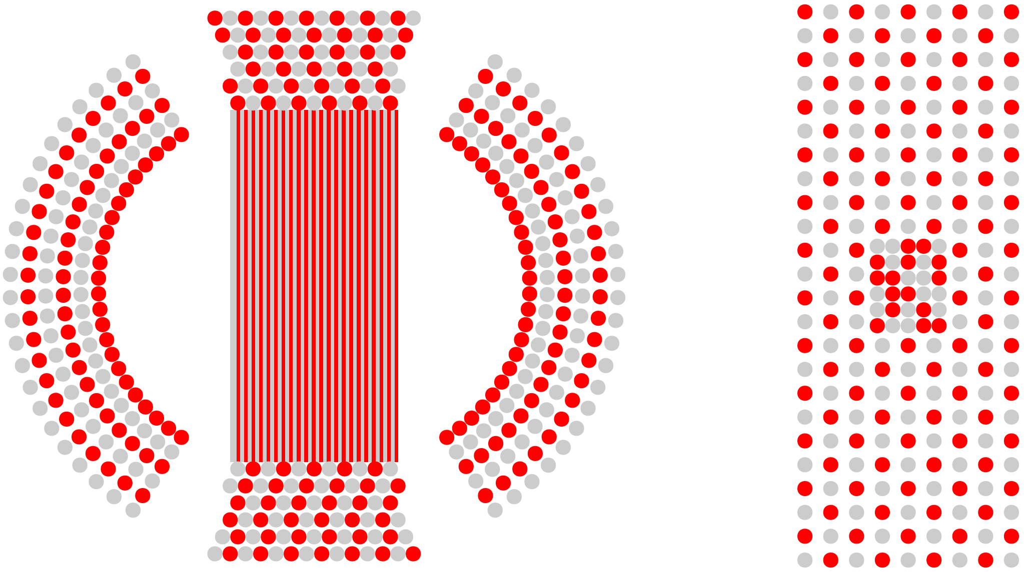
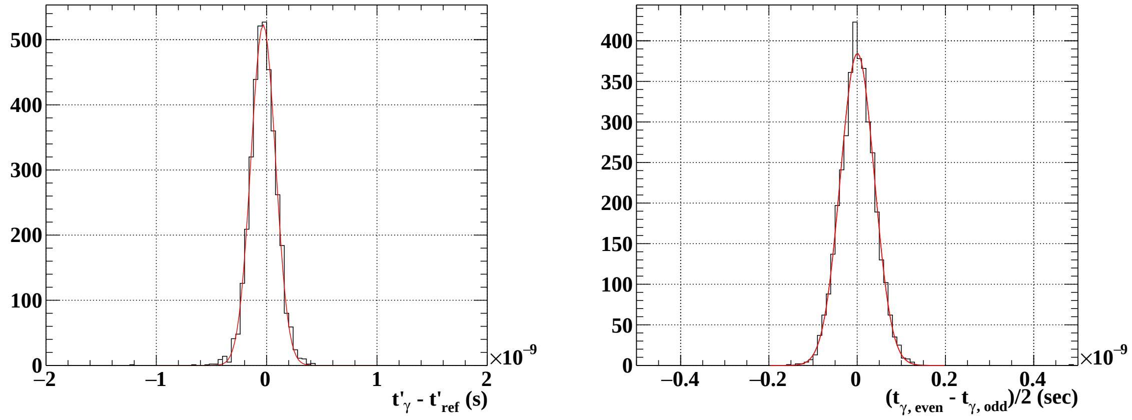
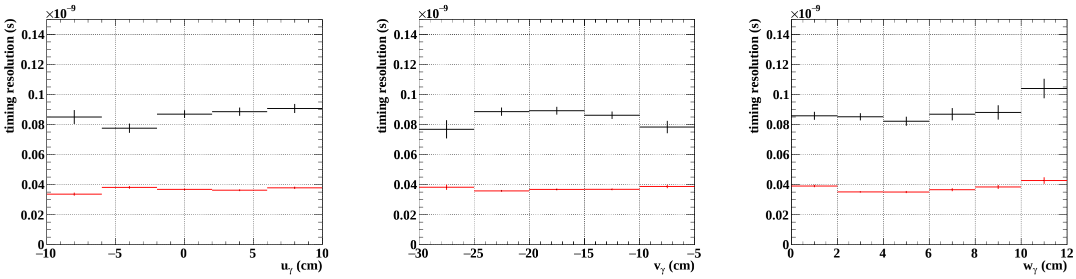


Figure 3.29 Assignment for even (gray) and odd (red) channels in the even-odd analysis.



(a) (b)

Figure 3.30 Measured timing difference between (a) the LXe detector and the timing counter and (b) the even and odd timings. σ of the fitted Gaussian is 102 ps and 38 ps, respectively.



(a) (b) (c)

Figure 3.31 Position dependence of absolute (black) and intrinsic (red) timing resolution for (a) u direction, (b) v direction and (c) w direction.

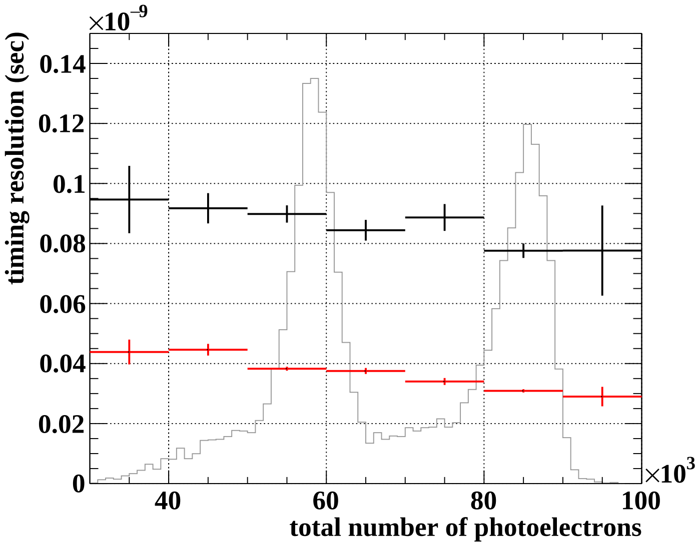


Figure 3.32 Absolute (black) and intrinsic (red) timing resolution as a number of total number of photoelectrons. The spectrum of the total number of photoelectrons is also shown (gray). The two peaks correspond to the energies of 54.9 MeV and 82.9 MeV, respectively.

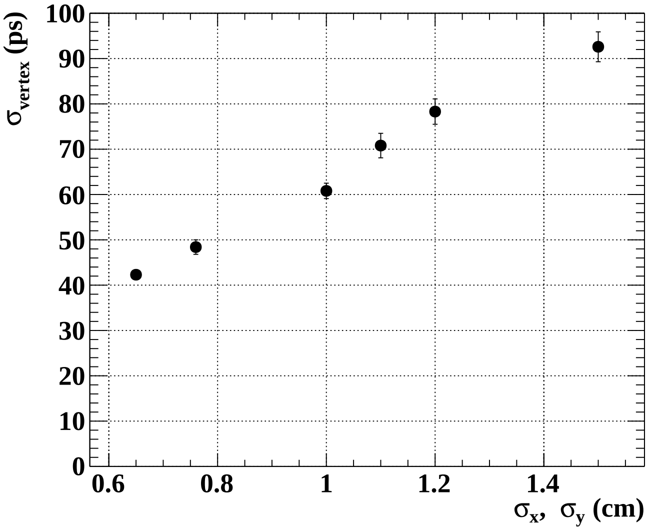
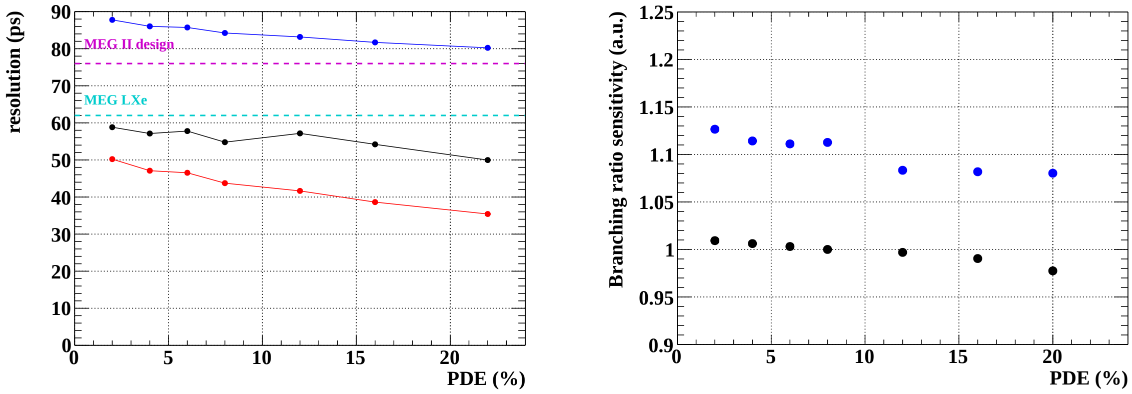


Figure 3.33 Simulated vertex contribution as a function of a beam size. The sizes along with x-axis and y-axis are assumed to be the same.

pensated to some extent by amplification since the noise level is not proportional to the gain of the amplifier as shown in Table 3.5. The timing resolution at lower PDE was simulated by adding the pedestal data taken at each gain to the simulated waveform. The precisions of the timing calibration and synchronization of the electronics were added as well. Fig. 3.34 shows the simulated timing resolution and the effect on the sensitivity. Since energy dependence of the unknown term causing the discrepancy is not observed, its effect is constantly added to the estimated resolutions. The effect of the PDE degradation was found to be up to 5% for the



(a) (b)

Figure 3.34 (a) Simulated absolute (black), intrinsic resolution (red) and absolute resolution including the unknown term of 72 ps (blue). (b) Effect on the branching ratio sensitivity as a function of the MPPC PDE assuming the simulated absolute timing resolution (black) and that with the unknown contribution (blue). The values are normalized by that of 8% PDE in the former case.

expected resolution.

Table 3.5 Noise level (RMS of pedestal waveform) at each amplifier gain in the run 2019.

amplifier gain	1	2.5	5	10
Noise level	0.7 mV	0.8 mV	0.9 mV	1.3 mV

3.6 Energy resolution

The energy of a γ-ray is reconstructed from the detected number of photons. In MEG, the energy resolution in the shallow region was worse than that in the deep region due to the non-uniformity of scintillation light collection, which can be improved thanks to use of the MPPCs. The energy resolution measured in MEG was worse than the expectation by 1.5%. The reason of the disagreement is not understood yet, and it can also remain in MEG II. Since the energy resolution in MEG II is expected to be 0.7% according to the simulation, the unknown contribution can be dominant. The resolution was measured using several γ-ray sources.

3.6.1 Energy reconstruction

The energy of a γ-ray is reconstructed from the weighted sum of the detected number of photons on each sensor as

$$\begin{aligned} E_\gamma &= C \times \sum_{\mathrm{pm}} N_{\mathrm{pho},i} \times w_i, \\ N_{\mathrm{pho},i} &= q_i / Gain_i / ECF_i / QE_i. \end{aligned} \tag{3.17}$$

The number of photons $N_{\mathrm{pho},i}$ detected with each photosensor is computed from the integrated charge q_i. The weight w_i is the correction for the sensor coverage, which is given by the size of the insensitive area around each sensor divided by the size of the sensitive area in order to realize uniform response in the whole detector. The factor C converts the sum of the number of photons to the energy, which is calculated using the monochromatic γ-ray sources. A further correction is applied to reduce the position dependence of the reconstructed energy, which reflects possible errors in the photosensor calibration and the position dependence of photon collection efficiency. This is also measured using the monochromatic γ-ray sources.

3.6.2 Performance Measurement

The energy resolution was evaluated in the pilot run 2019. In spite of the limited number of readout channels, it was found in simulation that the same energy resolution as with full electronics was achievable as shown in Table 3.6 by applying the event selection as

- $|u| < 9\,\mathrm{cm}$,
- $|v + 17| < 9\,\mathrm{cm}$,
- $|w| < 12\,\mathrm{cm}$.

The third selection was applied to reject the very deep events since scintillation light from the deep conversion depth distributes wider than the MPPC readout area. In addition, a selection of $E_{\gamma,\mathrm{MPPC}} > E_{\gamma,\mathrm{PMT}}/6$ was used to reject cosmic-ray events since the cosmic rays do not enter the detector from the inner face and thus a fraction of the energy measured with the MPPCs is smaller than that of γ-ray.

Table 3.6 Energy resolution estimated with MC simulations in MEG II for the signal 52.8 MeV γ-ray

	Full channel readout	Limited channel readout
$w < 2\,\mathrm{cm}$	0.72(1)%	0.73(1)%
$w > 2\,\mathrm{cm}$	0.70(1)%	0.76(1)%

Measurement with 17.6 MeV γ-ray

The energy resolution at 17.6 MeV was measured using CW-Li. Fig. 3.35 shows the reconstructed energy distribution after the non-uniformity correction is applied. The energy resolution was evaluated to be 3.1% at 17.6 MeV, which is much worse than the expected resolution of 1.0% from the MC. A part of the difference between the MC and the data comes from an electronics noise, which is estimated to be 1.2% from the reconstructed energy distribution of pedestal data. The rest part can result from the unknown contribution observed in MEG since the energy resolution excluding the noise effect is 2.8%, which is consistent with the resolution measured in MEG for the deep events ($w > 2\,\mathrm{cm}$). This term worsens the energy resolution independent of the conversion depth. The reconstructed energies only with the MPPCs and the PMTs were compared between the MC and the data, and the discrepancy is consistently observed for both.

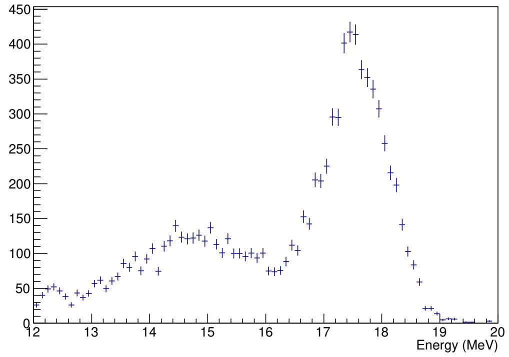


Figure 3.35 Reconstructed energy spectrum of 17.6 MeV γ-ray from CW-Li [62].

Measurement with background γ-ray

The energy resolution near the signal energy was evaluated using the background γ-rays from the muon beam. It was extracted by fitting an expected spectrum to the measured one with a binned-likelihood method. The likelihood $\mathcal{L}(\boldsymbol{p})$, and its negative log likelihood nll are defined as

$$\mathcal{L}(\boldsymbol{p}) := \prod_{b=0}^{N-1} \frac{f_b(\boldsymbol{p})^{h_b} e^{-f_b(\boldsymbol{p})}}{h_b!},$$

$$nll := -\ln \mathcal{L}(\boldsymbol{p}) = -\sum_{b=0}^{N-1} \{h_b \ln f_b(\boldsymbol{p}) - f_b(\boldsymbol{p})\} + const.\ ,$$

where $\{h_b | 0 \le b < N\}$ is a N-bin histogram of the measured spectrum, and $\{f_b(\boldsymbol{p}) | 0 \le b < N\}$ is that of the expected spectrum given by convoluting the simulated spectrum f_{MC} with an additional resolution σ as

$$f(x|\boldsymbol{p}) = R \times \int f_{\mathrm{MC}}(x/S + s) e^{-s^2/2\sigma^2} ds + f_{\mathrm{CR}}(x).$$

An energy scale S and an overall rate R are also considered as nuisance parameters. A non-beam contribution f_{CR} is also included based on the spectrum measured without muon beam[*2].

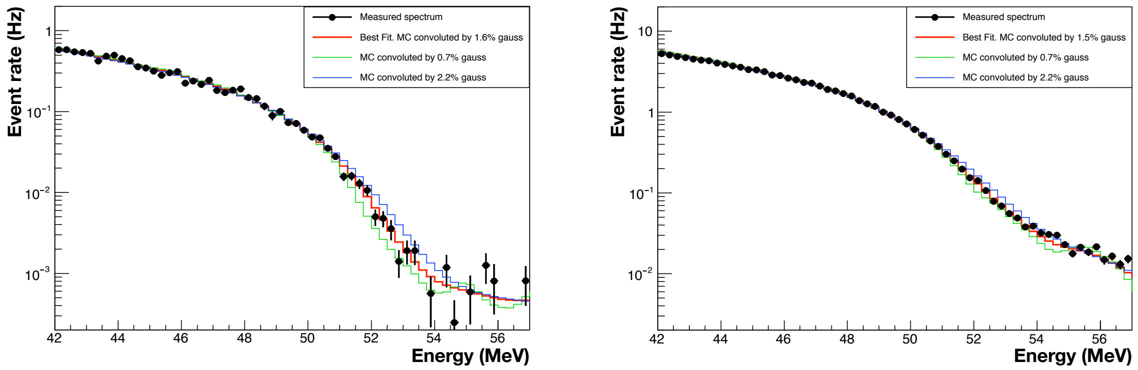
The measured background spectra at the reduced intensity (0.7×10^7 μ^+stops/s) and at the MEG II intensity (7×10^7 μ^+stops/s) are shown in Fig. 3.36, together with the expected spectra. A pileup elimination was applied in which sum waveforms of the MPPCs and the PMTs are unfolded by fitting template waveforms [62]. The best estimate of the fitting parameters are

$$\begin{aligned} S &= 1.013(3),\ R = 1.16,\ \sigma = 1.6(2)\% \quad \text{(for reduced intensity)}, \\ S &= 1.007(1),\ R = 1.13,\ \sigma = 1.5(1)\% \quad \text{(for MEG II intensity)}. \end{aligned}$$

The larger uncertainty of the reduced intensity is due to the limited statistic. The energy scales agree with that of the CW-Li γ-ray within 1% precision. The event rate was 10–20% higher

---

[*2] It mainly comes from cosmic rays, and neutrons from the accelerator captured by the LXe.



(a) (b)

Figure 3.36 Measured energy spectra of the background γ-ray and the best fit result (a) at the reduced intensity, (b) at the MEG II intensity [62]. The expected spectrum at better and worse resolution than best fit are also shown for comparison.

than that expected from the beam intensity. This is likely due to the higher beam intensity than the setting as mentioned in Sec. 2.2.1. The energy resolution can be obtained by combining the fitted σ and the 0.7% energy resolution in the simulation, and it was estimated to be 1.7(2)% for the reduced intensity and 1.7(1)% for the MEG II intensity. The noise contribution was only 0.4% for the signal energy γ-ray, and thus it is negligible.

Fig. 3.37 shows the depth dependence of the energy resolution. An improvement in the shallow region, where many γ-rays interact as shown in Fig. 2.19, was observed though it was worse than the expectation due to the unknown contribution.

The expected sensitivity was evaluated with the measured energy resolution. The improvement for the shallow events results in the 15% improvement compared to MEG. The sensitivity improvement from MEG is 10% smaller than the design value since it was calculated assuming only a half of the unknown contribution remains though it was found to completely remain.

Although the cause of the unknown term has not been understood yet, there are several possibilities to explain it. One possibility is a problem of the detector system such as a convection of LXe in the cryostat. Another possibility is an event-by-event fluctuation of the shower development which is not correctly simulated for some reasons. An intrinsic property of the LXe scintillation is also a candidate; a part of the energy deposit in LXe can be used for a generation of the charged carries or an emission in infrared, which may cause an event-by-event fluctuation of the energy used for an emission of the VUV light. Further studies are necessary for a better understanding.

Effect of PDE degradation

The energy resolution can be affected both by the statistical fluctuation and the noise on the signals. Fig. 3.38(a) shows the simulated energy resolution of the signal γ-ray with the contribution for the statistical fluctuation. It was found to increase the resolution by 0.15% at

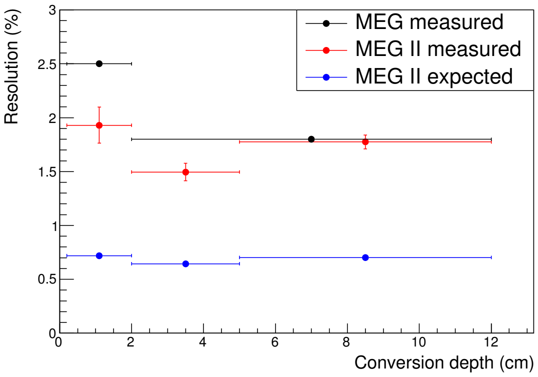
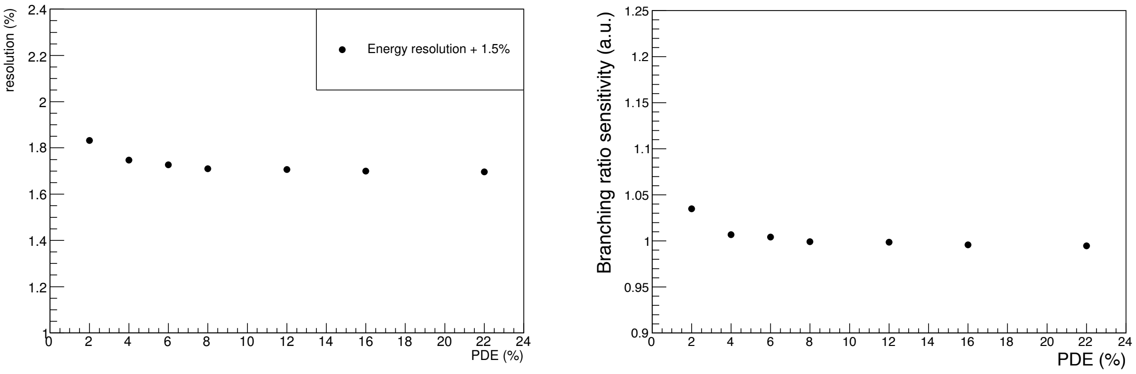


Figure 3.37 Energy resolution as a function of the conversion depth estimated from background γ-rays at MEG II intensity [62].



(a)

(b)

Figure 3.38 (a) Expected energy resolution for the signal γ-ray as a function of the PDE including the unknown contribution of 1.5%. (b) Effect on the branching ratio sensitivity from the energy resolution degradation at different MPPC PDEs [62]. The values are normalized by that of 8% PDE.

the PDE of 2%. The noise contribution was simulated using the noise level measured in the pilot run 2019, and it is found to be less than 0.6% at 2% PDE by applying the amplification. As a result, the sensitivity can be degraded by 4% as shown in Fig. 3.38(b).

Chapter 4

Performance Evaluation of Radiative Decay Counter

The downstream RDC has been tested with a μ beam. In this chapter, the calibration is described followed by the discussion on the performance measured with the μ beam.

4.1 Calibration

The energy of the RDC calorimeter is converted from charge as

$$E = q \times S_{\mathrm{calo}}, \tag{4.1}$$

$$S_{\mathrm{calo}} = (S_{\mathrm{calib}} + f_{\mathrm{temp}} \times \Delta T)/f_{\mathrm{bias}}, \tag{4.2}$$

where q is the charge, S_{calo} is an energy scale conversion factor of the calorimeter, S_{calib} is the factor obtained from calibration, f_{temp} and f_{bias} are correction factors of temperature and bias, ΔT is a difference between temperatures at the time of calibration and the measurement.

Fitting function

The energy of the RDC calorimeter is calibrated using the intrinsic radioactivity of LYSO. The highest energy peak of 596 keV (88 + 202 + 307 keV) is fitted with a function of the theoretical energy spectrum including γ-rays and β decays convoluted with a gaussian function (Fig. 4.1), that is,

$$f(x) = \int \mathrm{Gaus}(x/p_7 - E, \sigma(E, p_8, p_9)) \times \mathrm{Decay}(E)dE, \tag{4.3}$$

$$\mathrm{Gaus}(x, \sigma) = \exp(-x^2/2\sigma^2), \tag{4.4}$$

$$\sigma(x, \sigma_{\mathrm{energy}}, \sigma_{\mathrm{noise}}) = x \times \sigma_{\mathrm{energy}} + \sigma^2_{\mathrm{noise}}, \tag{4.5}$$

$$\mathrm{Decay}(x) = \sum_{i=0}^{6} p_i \times \mathrm{BetaDecay}(x/p_7 - E_i), \tag{4.6}$$

$$\mathrm{BetaDecay}(x) = \mathrm{Norm} \times \sqrt{x^2 + 2x \times m_{\mathrm{e}}c^2} \times (Q - x)^2 \times (x + m_{\mathrm{e}}c^2), \tag{4.7}$$

where Norm is normalization constant, m_{e} is the mass of an electron, c is the velocity of light, and Q is a Q-value of β decay. $p_0 - p_9$ are fitting parameters; $p_0 - p_6$ are amplitudes of the

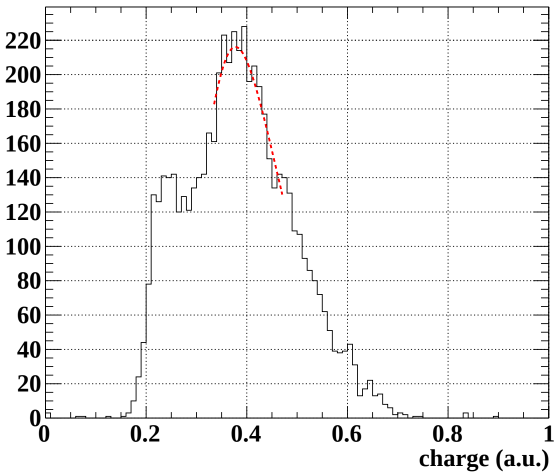


Figure 4.1 Example of the measured energy spectrum of the LYSO intrinsic radioactivity and the fitting function.

Table 4.1 Parameters of the fitting function for the LYSO intrinsic radioactivity spectrum.

p_0	Amplitude of the γ-ray peak of 88 keV
p_1	Amplitude of the γ-ray peak of 202 keV
p_2	Amplitude of the γ-ray peak of 307 keV
p_3	Amplitude of the γ-ray peak of 88 + 202 keV
p_4	Amplitude of the γ-ray peak of 88 + 307 keV
p_5	Amplitude of the γ-ray peak of 202 + 307 keV
p_6	Amplitude of the γ-ray peak of 88 + 202 + 307 keV
p_7	Conversion factor from charge to energy
p_8	Energy resolution
p_9	Noise

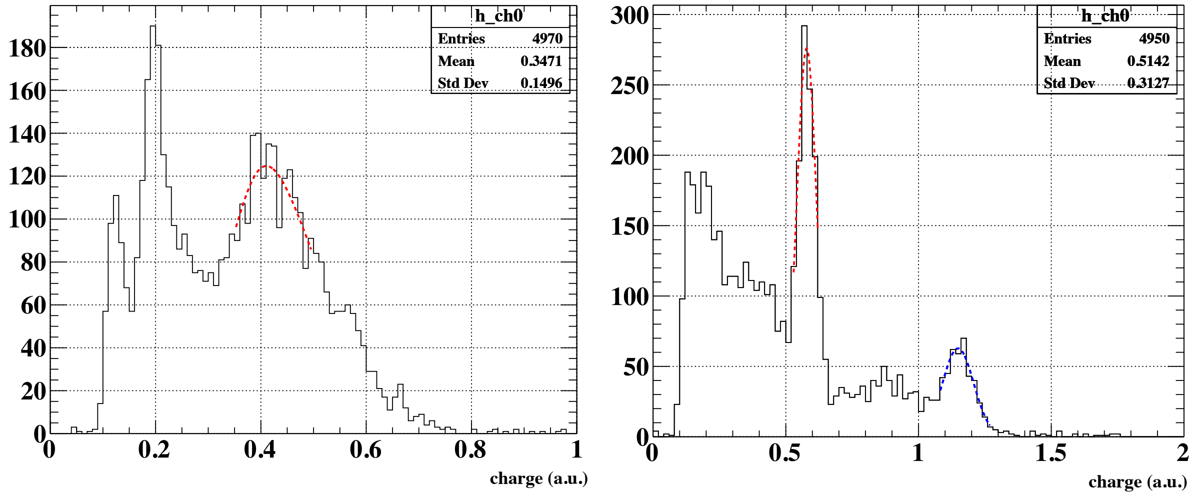
γ-ray peaks, p_7 is a conversion factor from charge to energy, p_8 is the energy resolution, and p_9 corresponds to noise as summarized in Table 4.1. The energy is calibrated using the conversion factor p_7 obtained by fitting, i.e. $S_{\text{calib}} = p_7$.

Since the fitting function has many parameters, the fitting results was found to be unstable; the results can easily change by fluctuation of data and a fitting range. Therefore, some fitting parameters are fixed to expected values in order to reduce the degree of freedom. The parameter settings are written in Table 4.2. The amplitudes of γ-ray peaks except for the highest one are set to reproduce the measured spectrum by hand due to a difficulty of estimation by the fitting because the fitting is performed only around the highest peak. The energy resolution is set to the measured value. The improvement reduces the instability.

The uncertainty from the fitting was evaluated by repeating the calculation with three different datasets obtained under the same condition, which was 1.2%.

Table 4.2 Parameter settings of the fitting function for the LYSO intrinsic radioactivity spectrum.

Parameter	Value
p_0	0
p_1	0
p_2	0
p_3	0.5
p_4	0.1
p_5	0
p_6	free
p_7	free
p_8	8%
p_9	free



(a) (b)

Figure 4.2 Energy spectra of (a) intrinsic radioactivity and (b) ^{88}Y with fitting results of each peak.

Bias correction

The calibration using the intrinsic radioactivity was verified by the measurement with ^{88}Y. The source was attached to each crystal and its energy spectrum was obtained. Fig. 4.2 shows the energy spectra of the same crystal. The peaks of the ^{88}Y were fitted with gaussian functions and the conversion factors were calculated with the peak positions. The relation between the factors obtained with the two different methods is shown in Fig. 4.3. The factors of the intrinsic radioactivity are consistently higher than those of ^{88}Y, which can be because of a bias from the fitting. Here, the factors calculated with the peak of 0.898 MeV are used as the result of ^{88}Y to suppress an effect from SiPM saturation. Therefore, the bias is corrected based on the fitting result of a liner function: $f_{\text{bias}} = 1.122 \pm 0.021$.

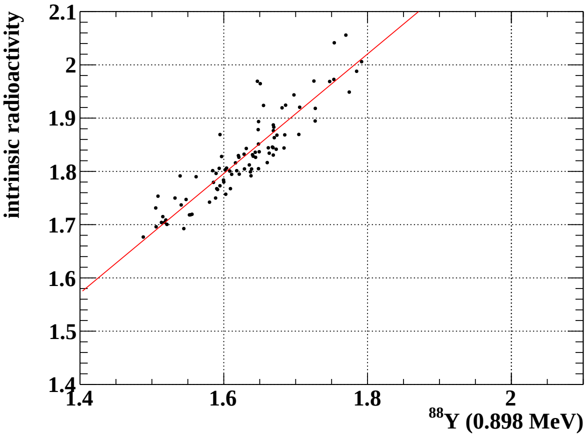
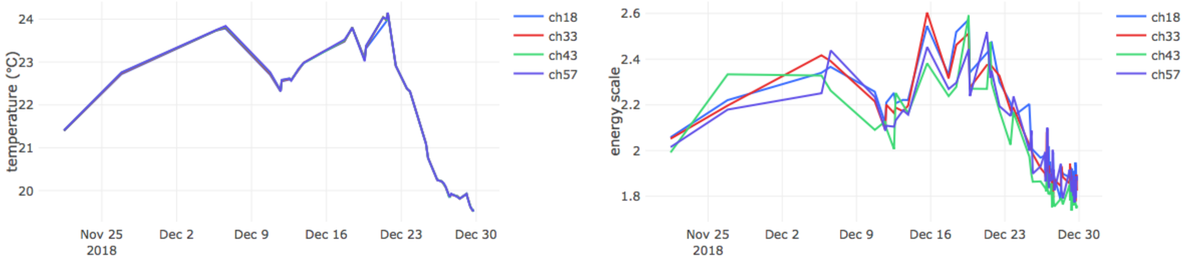


Figure 4.3 Comparison between the conversion factors obtained from the intrinsic radioactivity (vertical) and ^{88}Y (horizontal) with a fitted linear function. The result of the 0.898 MeV peak is used as ^{88}Y.



(a) (b)

Figure 4.4 (a) Measured temperature and (b) conversion factors during the monitoring of five weeks. The results of four channels are shown.

Temperature correction

Temperature correction is applied since a SiPM gain can change depending on the temperature resulting from a change of its breakdown voltage. The temperature dependence was measured by monitoring the conversion factors of four channels for five weeks as shown in Fig. 4.4. The temperature coefficient was obtained by fitting the relation with a linear function (Fig. 4.5), that is $f_{\text{temp}} = 0.114 \pm 0.03$.

Effect on energy resolution

The uncertainties from the calibration parameters affect the energy resolution. Their effects were estimated using the measured values and their uncertainties as summarized in Table 4.3. The measured energy resolution of 6.5% with ^{88}Y was used as the energy resolution before the conversion to the energy. The mean of the measured energy scales with different channels was used as a representative. The temperature change was assumed to be 1%, which would be expected in the experimental environment with an air conditioner. Including these uncertainties, the energy resolution was 6.9%, which is good enough not to worsen the branching ratio

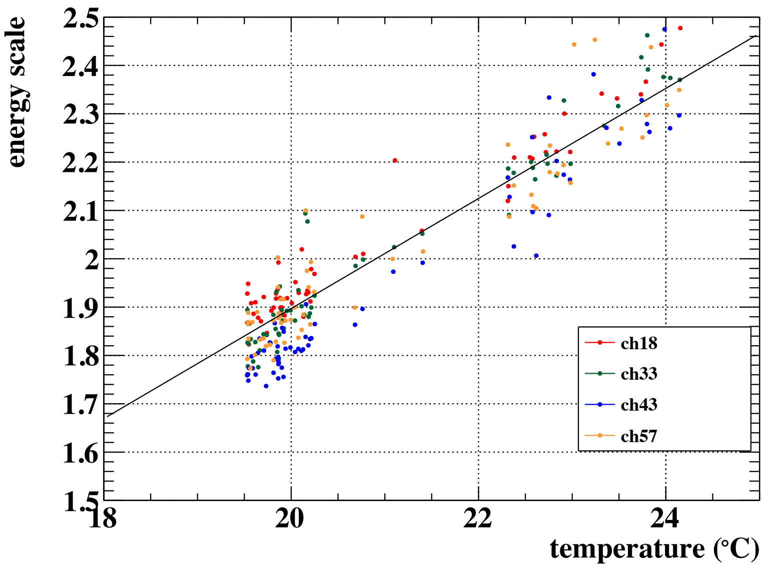


Figure 4.5 Temperature dependence of the conversion factor. The results of four channels are shown.

sensitivity; it is not deteriorated at least up to 10% energy resolution [65].

Table 4.3 Parameter values and uncertainties used for the estimation of the energy resolution.

Parameter	Value
Energy resolution before conversion	6.5% (1.836 MeV)
S_{self}	1.83 ± 0.022
f_{temp}	0.114 ± 0.003
f_{bias}	1.122 ± 0.021
δT	1°C

4.2 Detector performance

The performance of the RDC was evaluated with the data taken under the muon beam at 7.7×10^7 μ^+stops/s in the run 2020. In this section, the reconstruction algorithm for the RDC is described followed by the discussion on the performance.

4.2.1 Reconstruction

Fig. 4.6 shows an overview of the RDC reconstruction chain. Firstly, the waveforms of the timing counter are analyzed. Then, those of the calorimeters are analyzed with information of the timings extracted in the timing counter. Hit reconstruction is performed at each detector, and the reconstructed hits are clustered based on timings.

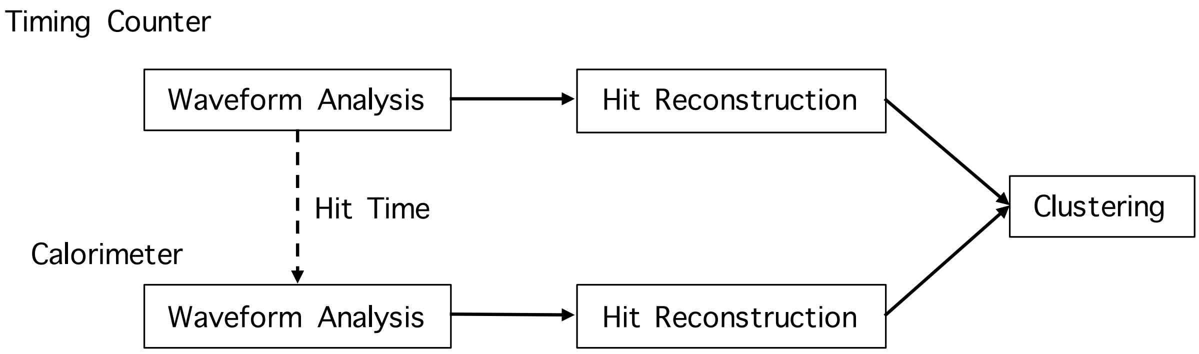


Figure 4.6 Overview of RDC reconstruction chain.

Waveform Analysis

Reconstruction starts from analyzing the raw waveform recorded with the waveform digitizer. Fig. 4.7 shows typical waveforms of the RDC. The information of signals such as the timings and charges is extracted by analyzing them.

To find the signals of the timing counter, a peak search is performed with a given threshold. The baseline is calculated by taking the mean of the amplitudes before the found pulse timing. The timing is extracted with the constant fraction method. The signal charge is obtained by integrating the waveform for 20 ns around the pulse. Fig. 4.8(a) shows the hit rate of each plastic scintillator. The maximum hit rate is about 1.2 MHz, which means one hit is expected in 830 ns on average. This results in an overlap of pulses with the probability of 0.9% over all incident hits since the pulses within 10 ns cannot be separated.

In the case of the calorimeter, the pulse timing is given based on the timings of found pulses in the timing counter, and the charge integration is performed around the timing since some signals are too small that they cannot be found by peak search. Fig. 4.8(b) shows the hit rates of each crystal. The center part has the highest hit rate of 0.6 MHz, which results in an overlap of pulses with the probability of 10% considering the charge integration region of 250 ns. The effect of the overlapping pulses is eliminated as follows. If another pulse whose timing is in the baseline calculation region of the pulse is found, the region is shifted to an earlier timing where no pulse is found. If another pulse whose timing is in the charge integration range of the pulse is found, the charge is estimated from the pulse height by multiplying a conversion factor extracted from a template waveform. If the adjacent pulse is so close that it affects the calculation of the pulse height of the focusing pulse, i.e. the adjacent pulse timing is within the charge integration region of the focusing pulse, the amplitude of the adjacent pulse at the timing of the height calculation is estimated from the template waveform, and it is subtracted from the original height of the focusing pulse. Then, the charge estimation is performed from the modified height.

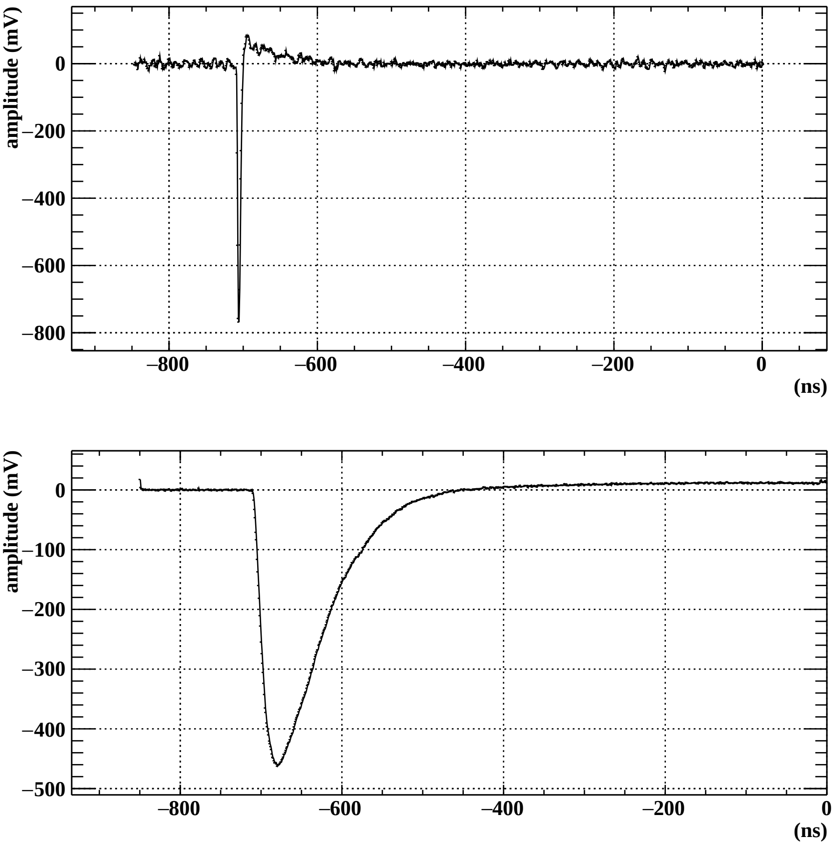
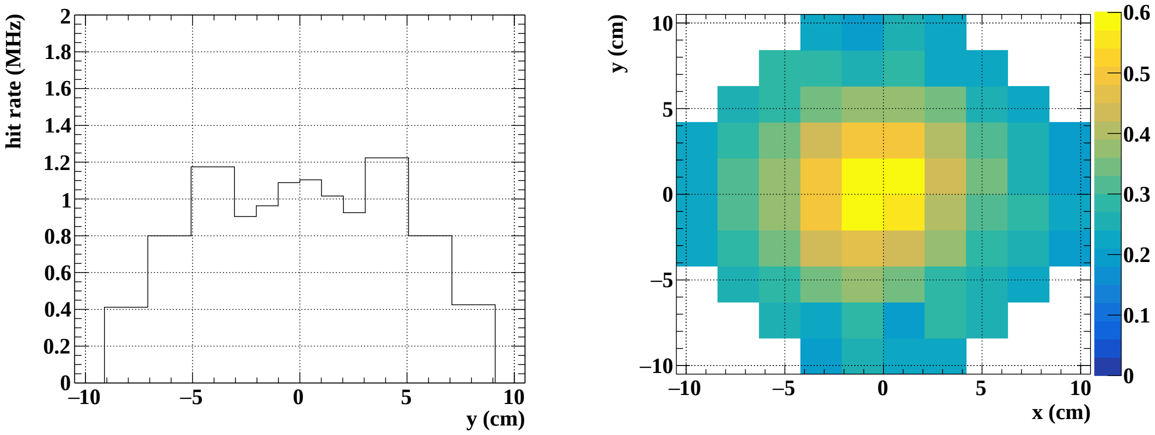


Figure 4.7 Typical waveforms of the RDC. The timing counter (top) and the calorimeter (bottom).



(a)

(b)

Figure 4.8 Measured hit rates of each scintillator for (a) the timing counter and (b) the calorimeter. The scale of the bar in (b) is in the unit of MHz. The depression at the center in (a) is due to the different scintillator sizes.

Hit reconstruction

A hit in the timing counter is reconstructed by combining the information of two channels of the same scintillator. The hit timing is defined as the average of the two channels as

$$t = \frac{t_{\mathrm{left}} + t_{\mathrm{right}}}{2}, \tag{4.8}$$

where t_{left} and t_{right} are the extracted timings of the left and right channels. The energy is given by the geometric mean of the channels:

$$E = \sqrt{q_{\mathrm{left}} \times q_{\mathrm{right}}} \times S_{\mathrm{tc}}, \tag{4.9}$$

where q_{left} and q_{right} are the integrated charges of the left and right channels. S_{tc} is an energy scale conversion factor of the timing counter, which is calculated by comparing the peak energies of Michel positrons of MC and data.

A hit in the calorimeter is reconstructed channel by channel. The energy of the hit is converted from the charge, which is discussed in Sec. 4.1.

Clustering

Since a positron leaves several hits in the RDC, the reconstructed hits belonging to the same positron should be combined. This clustering is performed based on the hit timings. If the hits on the timing counter are within 5 ns, they are regarded as the identical positron hits. Since the hits of the calorimeter are already reconstructed based on the timings of the timing counter, they are automatically assigned to the same cluster with the seed hits. The fastest timing of the clustered timing counter hits is adopted as the cluster timing while the energy is calculated by summing up the energies of the members.

4.2.2 Performance

The performance of the RDC is characterized by the timing and energy distributions and extracted parameters from these distributions. Firstly, the event selection criteria are explained. Then, the raw distributions are shown followed by the evaluation of the parameters; the RMD fraction and the hit rate of accidental positrons. In addition, the effect of the SiPM saturation observed for the calorimeter is discussed.

Event Selection

To discuss the performance of the RDC, it must be ensured that the γ-ray energy is measured correctly with the LXe detector since the detection performance of the RMD positrons relies on it. The same selection as that for the evaluation of the γ-ray energy resolution (Sec. 3.6.2) was applied. To reduce the pileup effect, the pileup elimination based on the peak search and the waveform unfolding is applied[*1], which will be explained in Chap. 5. The events whose γ-ray energies are from 48 MeV to 58 MeV were used for the analysis. Fig. 4.9 shows the energy spectra measured with the LXe detector. The data and MC were consistent with each other.

[*1] The deep learning-based rejection is not used due to the limited number of channels.

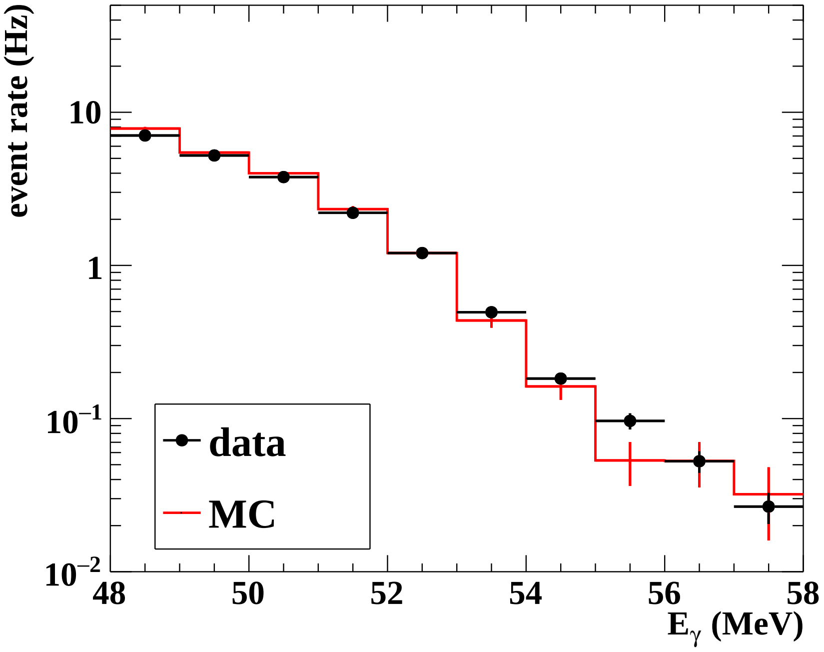


Figure 4.9 Energy spectra of the background γ-ray measured with the LXe detector for data at 7.7×10^7 μ^+ stops/s in the run 2020 (black) and MC (red).

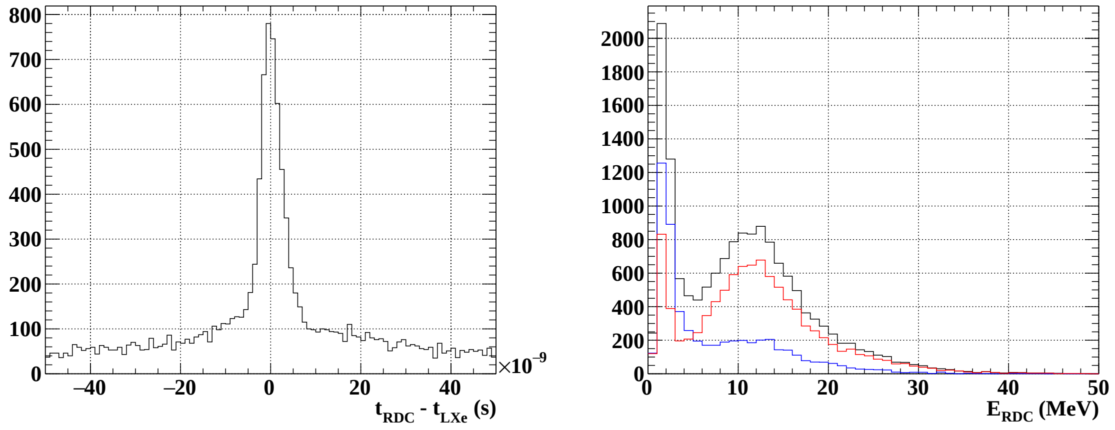
Measured timing and energy distributions

The distribution of the timing difference between the RDC and the LXe detector of the RMD-like hit, which is defined as a particle that hits the RDC at the closest timing to the center of the RDC and the LXe timing difference, is shown in Fig. 4.10(a). A clear peak together with a low tail component was observed as expected, which correspond to the RMD positrons and the Michel positrons, respectively. The spread of the peak results from different positron paths.

Fig. 4.10(b) shows the energy distribution. The low energy peak corresponding to the RMD positron was observed. By selecting the events within ±8 ns timing window, the events coming from the RMD positrons and the Michel positrons can be roughly separated as demonstrated in the figure. This means the RDC successfully detects both type of positrons, and they can be distinguished by the RDC observables. Note that the low energy peak in the off-timing events derives from γ-rays generated from Michel positrons.

RMD fraction

The RMD fraction f_{RMD} is the most important parameter to describe the RDC performance, which is the product of the detection efficiency and the fraction of the RMD events over the all events. It is defined as the number of RMD positrons detected with the RDC N_{RMD} divided by the number of γ-ray events triggered with the LXe detector N_γ. The number of RMD positrons can be calculated by counting excess of events over the expectation from the accidental hit rate of the background R_{acc} assuming a Poisson distribution as



(a) (b)

Figure 4.10 (a) Timing difference between the RDC and the LXe detector under the muon beam at $7.7 \times 10^7\ \mu^+$ stops/s. (b) Energy distributions measured with the RDC for all events (black), off-timing (red) and on-timing events (blue).

$$\begin{aligned} f_{\mathrm{RMD}} &= \frac{N_{\mathrm{RMD}}}{N_\gamma}, \\ N_{\mathrm{RMD}} &= N_{\text{on-timing}} - N_{\text{off-timing}} \times \frac{N_{\mathrm{BG}}(t_{\text{min, on-timing}}, t_{\text{max, on-timing}})}{N_{\mathrm{BG}}(t_{\text{min, off-timing}}, t_{\text{max, off-timing}})}, \\ N_{\mathrm{BG}}(t_{\min}, t_{\max}) &= A \times \int_{t_{\min}}^{t_{\max}} e^{-2R_{\mathrm{acc}}|t|} dt, \end{aligned} \tag{4.10}$$

where $N_{\text{on/off-timing}}$ is the numbers of events in the on/off-timing region from $t_{\text{min, on/off-timing}}$ to $t_{\text{max, on/off-timing}}$ and A is normalization.

Overall fraction was evaluated to be 26.6(4)% with data, which is consistent with the MC expectation of 27.2(4)%. To estimate the dependence on the γ-ray energy, the events were grouped into several energy regions and the efficiency was calculated in the same way. As shown in Fig. 4.11, the values of data and MC were consistent with each other except for the high energy region where statistics is limited. Therefore, it is confirmed that the RDC can detect the RMD positrons as expected.

Hit rate of Michel positron

The positron from the Michel decay is a background source for the RMD detection. The increase of the backgrounds reduces the power of the RMD identification since the probability to have a hit around the timing center in the non-RMD event gets higher.

The expected hit rate was estimated with MC. Fig. 4.12 shows the muon decay vertex along with the z-axis. The probabilities of Michel positron hit at the RDC with different vertex positions are summarized in Table 4.4. The hit probability becomes higher as the muon decays at further downstream due to less interfering materials. Since the beam rate is usually defined as the stopping rate at the target, the probabilities over the total muon incidence from the

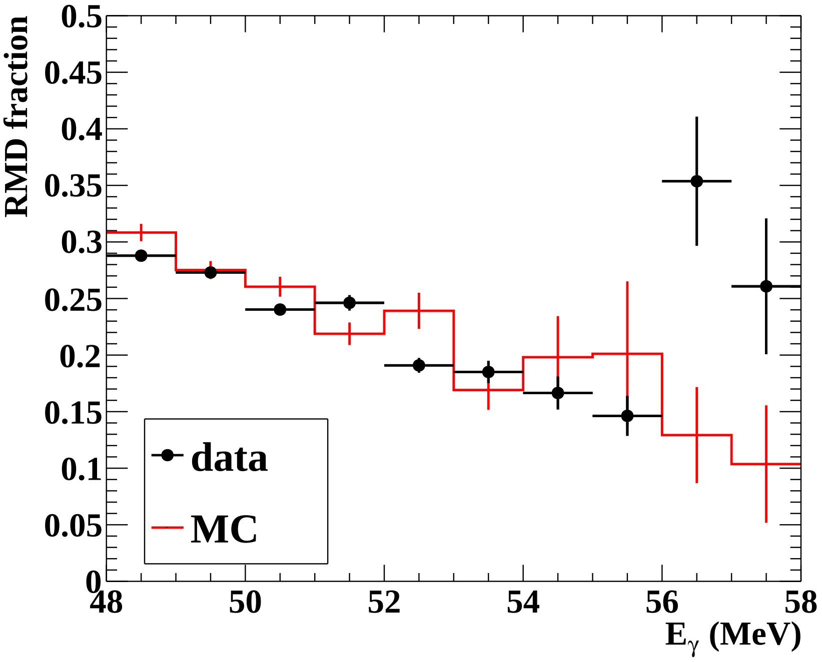


Figure 4.11 Dependence of RMD fraction on the γ-ray energy for data (black) and MC (red).

upstream should be corrected with the stopping efficiency at the target of 90.2(4)%. From the number of hits over all muon decay, the hit rate can be 10.4(1)% of the beam rate. Note that it is 14.0(2)% for the run 2017 due to the less materials since the mockup was installed instead of the real CDCH.

The hit rate was calculated by counting the number of events per unit time. In order to avoid including the RMD positron, the timing sideband from −40 ns to −20 ns in the timing difference distribution was used for the estimation. The results measured in each year are listed in Table 4.5. There is still a large uncertainty for the estimation mainly due to the uncertainty of the beam rate as discussed in Sec. 2.2, but they imply the hit rate can be higher than the MC expectation (Table 4.4) with up to 40% excess, which possibly because the material budget from the target to the RDC is not correctly simulated; the less materials for the CDCH result in the higher hit rate as seen in the difference between the mockup and the real CDCH cases, for example.

Table 4.4 Hit probability to the RDC over the incident muons decaying at each position.

vertex	hit probability to RDC
upstream of target	3.0(2)%
on target	9.4(1)%
downstream of target	16.0(6)%
overall	9.4(1)%

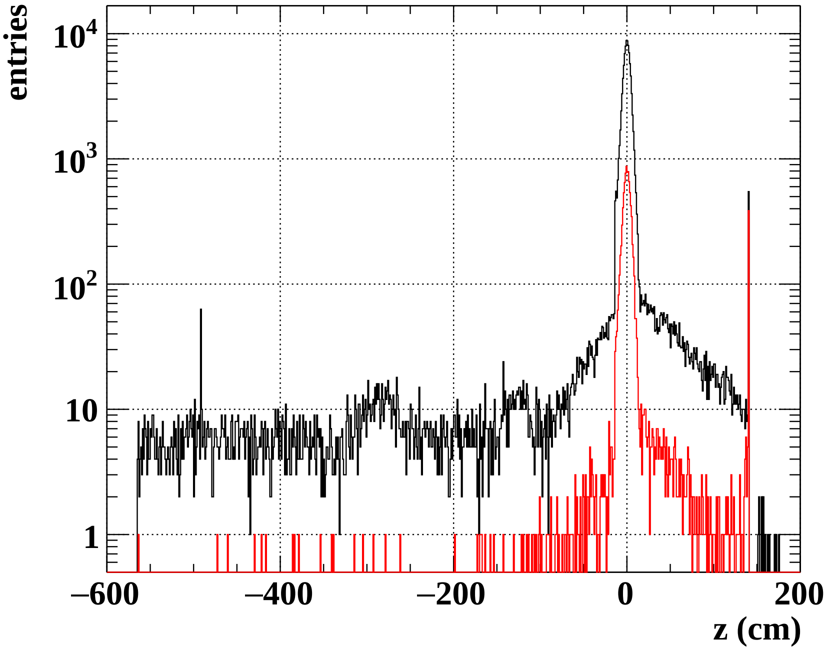


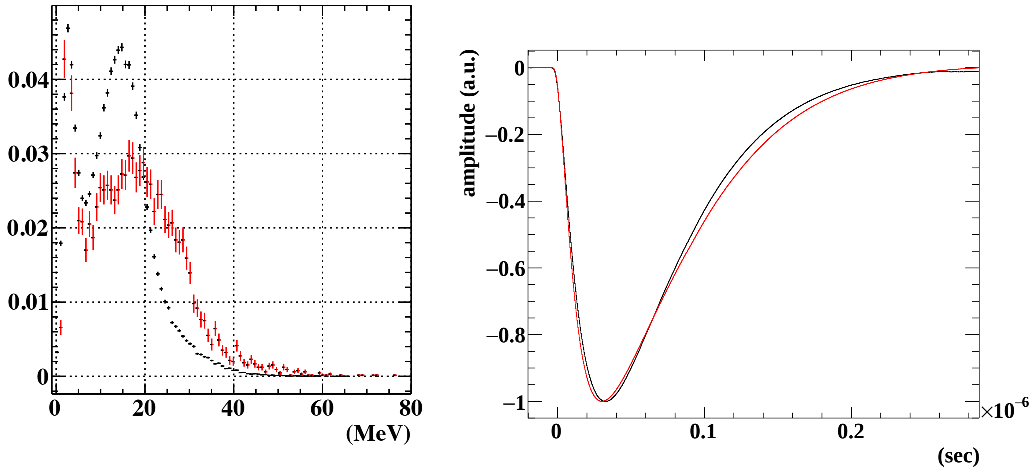
Figure 4.12 Simulated muon decay vertex along with the z-axis for all events (black) and events with a RDC hit. Most muons decay on the target around zero. The sharp peak at the downstream side corresponds to the surface of the RDC.

Table 4.5 Hit rate measured with the RDC. The data in 2018 is not included due to the serious beam contamination.

year	beam rate (μ^+stops/s)	hit rate (MHz)	hit probability (%)
2017	3.2×10^7	5.7	18
2019	0.7×10^7	1.0	14
	7×10^7	10	14
2020	1.6×10^7	2.1	13
	7.7×10^7	10	13

Saturation of SiPM for calorimeter

The SiPM used to read out the scintillation light of each LYSO crystal has a small pixel pitch in order to avoid the saturation of the pixels as discussed in Sec. 2.1.3.3. Nevertheless, it was found that the energy distribution measured with the calorimeter is greatly affected by the saturation effect. Fig. 4.13(a) compares the energy distribution measured under the muon beam and the MC expectation. The shape in the high energy is deformed, which can be explained by the underestimation of the energy due to the saturation. The SiPM has 14400 pixels for each while about 1300 photons are generated per energy deposit of 1 MeV, and so the amount of generated photons reaches a half of the number of the pixels with the energy deposit of 6 MeV. In fact, the deformation of the waveform was observed as shown in Fig. 4.13(b). The difference of the shape can reflect the fact that the amplitude around the peak becomes lower due to the



(a) (b)

Figure 4.13 (a) Energy distribution under the muon beam of data (black) and MC (red). (b) Template waveforms of the SiPM attached to the LYSO crystal averaged over the events with energy deposits below (black) and over (red) 10 MeV normalized by the pulse height.

saturation while that in the other region is not affected since the amount of light is small enough.

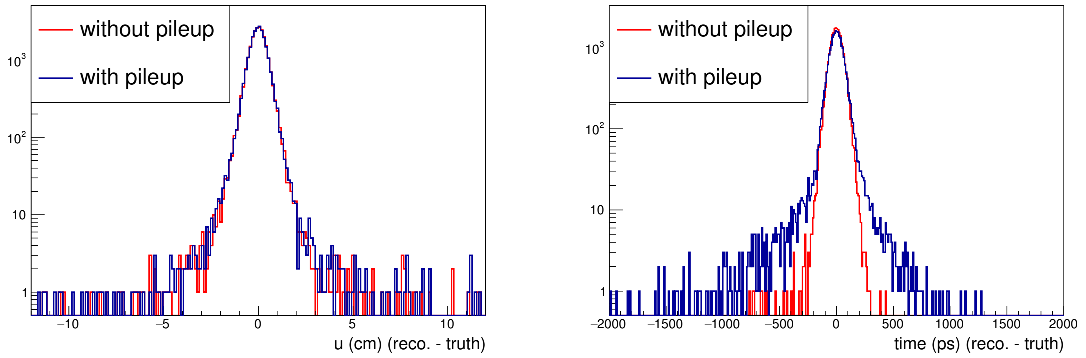
No correction is applied for the energy reconstruction since it was found that the saturation does not affect the branching ratio sensitivity as will be discussed in Sec. 7.2.

Chapter 5

γ-ray Background Reduction by Pileup Elimination

The energy, position and timing of γ-ray are reconstructed using the photosensor waveforms as discussed in Chap. 3. In practice, however, multiple γ-rays can hit the LXe detector in the analysis time window, which can affect the reconstruction of γ-rays. The position and the timing resolution are not affected so much because they are reconstructed from the local information only around the hit position as shown in Fig. 5.1. The reconstructed timings of some γ-rays are biased by the pileup γ-rays as shown in the long tail of Fig. 5.1(b). This leads to an analysis inefficiency for the signal events up to 3%. The most significant effect from the pileup γ-rays is seen in the energy reconstruction because it is performed using global information. Thus, this chapter focuses on the pileup effect on the reconstructed energy.

In Sec. 5.1, sources of pileup γ-rays are introduced. The pileup elimination algorithm is explained in Sec. 5.2 followed by the performance evaluation in Sec. 5.3.



(a)

(b)

Figure 5.1 Pileup effect on the (a) position (u-direction) and (b) timing reconstruction of the signal γ-rays in MC [62]. The differences between the reconstructed variable and its truth are shown.

5.1 Sources of pileup

There are mainly two types of pileup γ-ray sources separated from the aspect of the origin; the pileup γ-ray originates from a different μ or the same μ as the main γ-ray. Their characteristics as a pileup γ-ray, i.e. the incident timings, are different due to this difference of the origin, which requires different approaches to eliminate them. Hereafter, the detailed explanations for these pileup γ-rays follow.

5.1.1 Pileup γ-ray from same μ decay

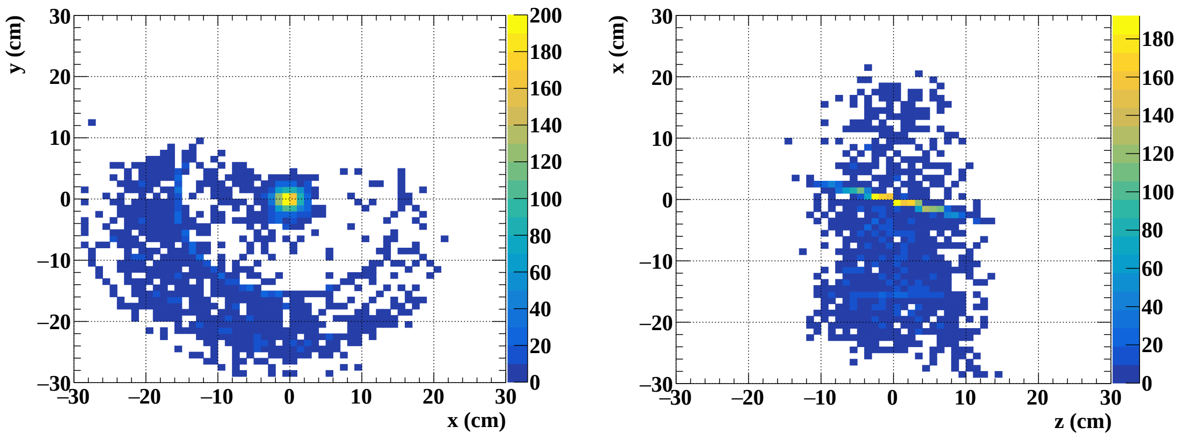
The first type of pileup γ-rays is a pileup γ-ray from the same muon decay. Such a γ-ray originates from AIF.

Most of the Michel positrons are swept away by the gradient magnetic field of the COBRA, but some of them interact with materials in flight, and two γ-rays are emitted being boosted in the direction of the flight. Fig. 5.2 shows the interaction points of AIF from which γ-rays deposit their energies to the LXe detector ($E_\gamma > 20\,\mathrm{MeV}$). About 42% of the AIF γ-rays are emitted on the target, which is a central part in the figure, and most of the rests are emitted in the drift chamber.

The background events with the AIF γ-rays reconstructed to be signal-like energy can be classified into two event types shown in Fig. 5.3. The first type is the event in which only one of the emitted γ-rays hits to the LXe detector with a high energy around the signal energy. Since such a γ-ray carries most of the parental positron energy, the other γ-ray is emitted nearly backward, and thus it cannot come to the detector. This type of event is called AIF 1γ event. The second type is the event in which both of the emitted γ-rays hit the LXe detector, which is called AIF 2γ event. If their hit positions are too close to identify them, they look like a single high energy γ-ray.

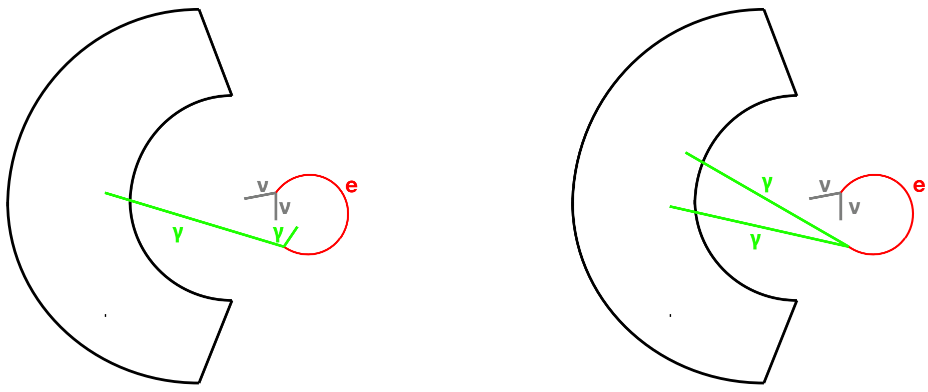
The latter type is the source of simultaneous pileup γ-rays. This AIF 2γ event has a large contribution near the signal energy. Fig. 5.4 shows the energy spectra for different sources of the background γ-rays from a single muon decay. The number of the γ-rays from RMD rapidly decreases as the energy gets close to the signal energy while that from AIF decreases gradually resulting from the Michel positron spectrum (Fig. 1.14). As a result, the fraction of AIF 2γ events becomes dominant in the high energy region though this is moderated by the energy resolution as shown in Fig. 5.5. At the currently achieved energy resolution of 1.7%, the AIF 2γ contribution is 34% near the signal energy ($51.5\,\mathrm{MeV} < E_\gamma < 54\,\mathrm{MeV}$).

All of this type of pileup γ-rays hit the LXe detector at the same timing with each other. Such γ-rays, namely on-timing pileups, cannot be found by a temporal search, and thus a spatial search is the only effective way.



(a) (b)

Figure 5.2 Interaction points of AIF projected to (a) x–y axis and (b) z–x axis. The hot region around the center corresponds to the target.



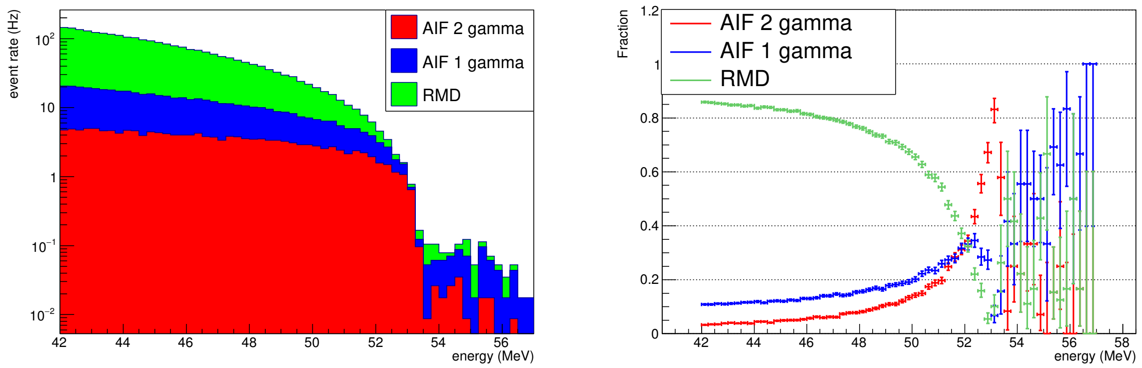
(a) (b)

Figure 5.3 Two types of high energy background γ from AIF. (a) AIF 1γ and (b) AIF 2γ [62].

5.1.2 Accidental pileup γ-ray from a different μ decay

The second type of pileup γ-rays is an accidental pileup γ-ray from a different muon decay. In the MEG II experiment, 7×10^7 muons decay on the target per second. Some of them emit γ-rays in the radiative muon decay, and some of the Michel positrons emit γ-rays via the interactions with materials. The energy spectra of these γ-rays are given in Fig. 1.20(a). A single muon decay emits a γ-ray which deposits energy in the LXe detector with a probability of 1.3%, and thus one pileup γ-ray enters the LXe detector at 0.7 MHz on average ($E_\gamma > 0.2\,\mathrm{MeV}$[*1]).

[*1] Only the γ-rays whose energy deposits are larger than 0.2 MeV are included since the energy lower than the threshold is too small to affect the reconstruction.



(a) (b)

Figure 5.4 (a) Simulated energy deposits of the background γ-rays in the LXe detector for each event type. Events above $m_\mu/2$ come from muon decays before stopped on the target. Muon momentum can boost decay products in these events. (b) Fraction of the background γ-rays of each type calculated from (a) [62].

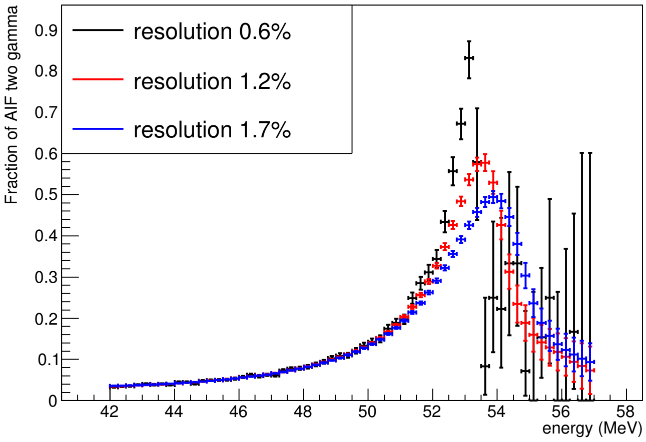


Figure 5.5 Expected fraction of the AIF 2γ events with the different energy resolutions [62].

Fig. 5.6 shows simulated energy spectra of background γ-rays without accidental pileups and with full pileups expected at the MEG II intensity. The pileup γ-rays bias the reconstructed energy to be higher, and the number of events near the signal energy increases from 64 Hz to 178 Hz (51.5 MeV $< E_\gamma <$ 54 MeV). This greatly deteriorates the sensitivity to signal γ-rays, and thus they must be eliminated.

Most of this type of pileup γ-rays hit the LXe detector at a different timing from the main γ-ray, and such γ-rays, namely off-timing pileups, can be found by a temporal search. A search in space using the light distribution of photosensors is also effective if their hit timings are close enough to be included in the charge calculation.

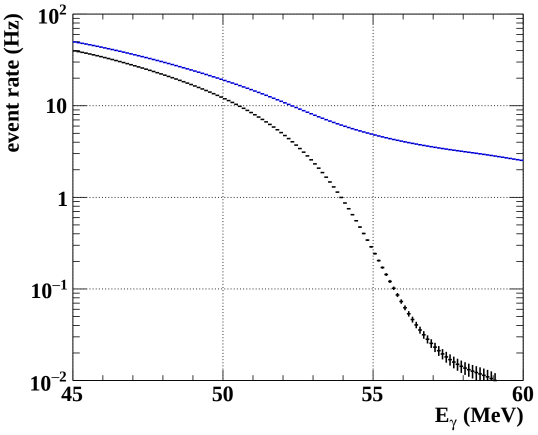


Figure 5.6 Reconstructed energy spectra of background γ-rays in the simulation without accidental pileups (black) and pileups from the MEG II beam without pileup elimination (blue).

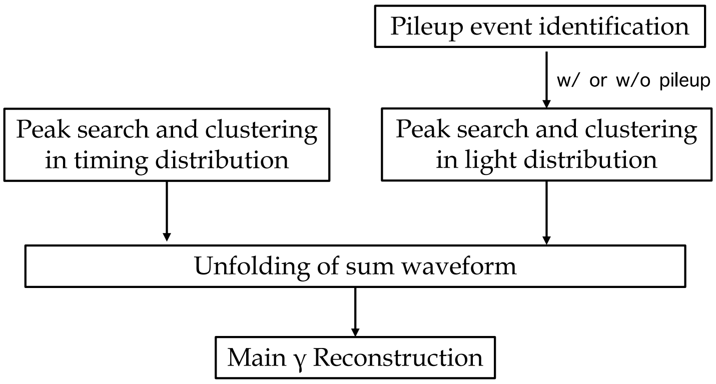


Figure 5.7 Flow of the pileup elimination algorithm.

5.2 Elimination algorithm

A pileup elimination algorithm combines methods to search for pileup γ-rays temporally and spatially in order to find both of off-timing and on-timing pileup γ-rays. Fig. 5.7 shows an overview of the algorithm flow. There are mainly three steps. The first step judges whether the event is likely to be a pileup event or not. The second step searches for peaks both in a light and timing distribution. In the peak search flow in the light distribution, the result from the first step is used. The third step unfolds the pileup γ-rays in the sum waveforms of photosensors using the information from the previous steps. At the end, an event status is assigned combining the results of the three steps and the main γ-ray energy is reconstructed or the event is discarded depending on the status. Hereafter, these algorithms are explained in detail.

5.2.1 Identification of pileup events with deep learning

The first step is a pileup event identification. At this step, whether the event is likely to include pileup γ-rays or not is judged using a deep learning technique. In this study, the light distribution was used for an input of the deep learning model as the first step to examine its usefulness though the timing distribution would be also useful to search especially for the off-timing pileup γ-rays.

Deep learning

The deep learning is a type of machine learning, which consists of a neural network with more than two layers. These neural networks mimic the behavior of the human brain to learn from a huge amount of data. The network has a layered structure in which dozens of computation nodes are included. The nodes in one layer are connected to the nodes in the following layers, and the computational results are transmitted to the connected nodes. The transmitted values from the previous layers are summed up with given weights and biases which are assigned to each connection, and the sum is used for the calculation at the node. The final layer outputs certain values, which is often likenesses with which the input sample belongs to classes for the classification or expected values for the regression.

Before its use, the deep learning model must be trained how to estimate a correct answer from data. This is achieved by optimizing the weights to minimize the difference between the current output and the correct value using a training dataset which is usually a set of input data and expected output values.

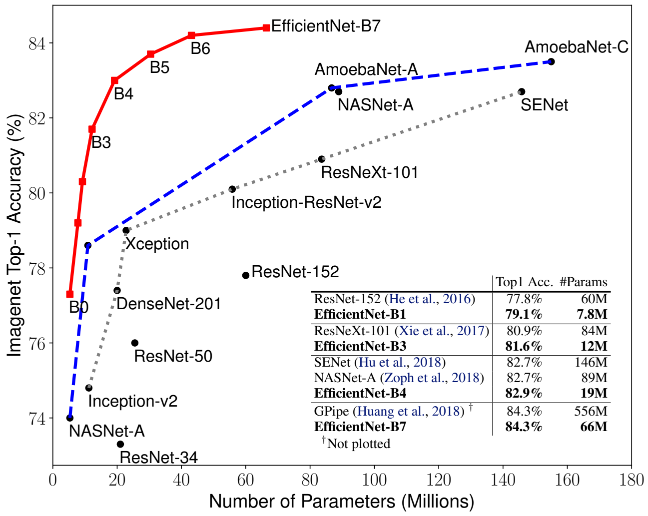
Model implementation

The deep learning model used for the pileup elimination is based on a two-dimensional convolutional neural network (CNN). The CNN is a type of neural networks which is often used to treat an image as the input such as image recognition. The input data are convoluted using a series of filters with trainable weights. In the training stage, the weights are optimized to extract useful characteristics so that the model can give a correct answer. Since the convolution is carried out by scanning the whole inputs with the common filter, the CNN has a shift invariance. The extracted features are input to fully connected layers in which a calculation of the final output is performed using the features.

EfficientNet

The model is implemented referring to EfficientNet [92], which is a type of the CNNs. The EfficientNet model succeeded in achieving a state-of-the-art performance with a reduced amount of parameters by scaling up its architecture in an optimal way. Fig. 5.8 compares the accuracy of the EffcientNets, which is a family of models developed by scaling up a baseline network obtained with neural architecture search, to the accuracies of other famous models. EfficientNets achieved better performance with smaller model sizes.

In general, the performance of the CNN can be improved by being scaled up if more computational resources are available. There are commonly three scalable dimensions in the architecture: depth, width and resolution (Fig. 5.9). The width scaling is a way to increase the image sizes of intermediate layers. Wider networks tend to be able to capture more fine-grained features and



	Top1 Acc.	#Params
ResNet-152 (He et al., 2016)	77.8%	60M
EfficientNet-B1	**79.1%**	**7.8M**
ResNeXt-101 (Xie et al., 2017)	80.9%	84M
EfficientNet-B3	**81.6%**	**12M**
SENet (Hu et al., 2018)	82.7%	146M
NASNet-A (Zoph et al., 2018)	82.7%	89M
EfficientNet-B4	**82.9%**	**19M**
GPipe (Huang et al., 2018) †	84.3%	556M
EfficientNet-B7	**84.3%**	**66M**

†Not plotted

Figure 5.8 Accuracies of models as a function of model size measured with ImageNet [92].

are easier to train. The depth scaling is the most common way in which the number of layers is gained. This allows the network to capture richer and more complex features though too deep networks are difficult to be trained. In the resolution scaling, channels of intermediate layers are increased to extract a variety of information. They focused on an efficiency of the model scaling, and systematically studied the scaling procedure.

A new compound scaling method was introduced in [92]. It uses a compound coefficient ϕ to uniformly scale network width w, depth d and resolution r in a principled was as

$$\begin{aligned} d &:= \alpha^{\phi}, \\ w &:= \beta^{\phi}, \\ r &:= \gamma^{\phi}, \end{aligned}$$

where α, β, γ are constants that can be determined by a small grid search. The compound scaling was demonstrated by scaling up the widely-used models, MobileNets [93] and ResNet [94] with different ϕ, and better accuracy was achieved than that of other single-dimension scaling methods. They also provided new models, EfficientNets, which were produced by scaling up a baseline model, EfficientNet-B0. As a result, they overcame any other models in terms of the accuracy with smaller parameter sizes.

Model for pileup identification

The model developed for the pileup identification is based on EfficientNet architecture. Table 5.1 summarizes its architecture. It consists of a series of convolution layers for feature extraction followed by a two-layer prediction part. The convolution part is composed of mobile inverted bottleneck blocks MBConv [95] [96] with sqeeze-and-excitation optimization.

The smallest model in the EfficientNet family, EfficientNet-B0, was adopted due to the limited computation resource. The structure of the convolution part is automatically defined by its operators and the input resolution while the two fully connected layers were chosen in comparison

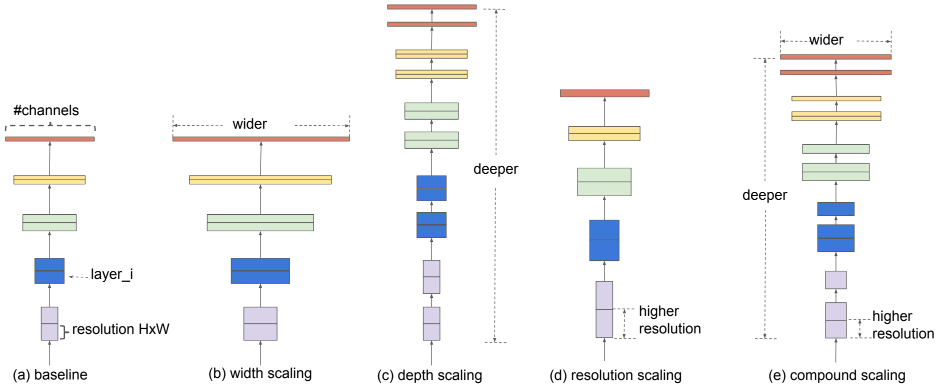


Figure 5.9 Schematic of model scaling: (a) baseline network example, (b)-(d) one dimensional scaling, and (e) compound scaling [92].

to the model with the single layer in terms of the performance.

Table 5.1 Model architecture for pileup identification. The operator MBConv$n$ means the MBConv block with an expansion factor of n.

Stage	Operator	Resolution	#Channels	#Layers
1	Conv3 × 3	93 × 44	32	1
2	MBConv1, k3 × 3	46 × 21	16	1
3	MBConv6, k3 × 3	46 × 21	24	2
4	MBConv6, k3 × 3	23 × 11	40	2
5	MBConv6, k3 × 3	12 × 6	80	3
6	MBConv6, k3 × 3	6 × 3	112	3
7	MBConv6, k3 × 3	6 × 3	192	4
8	MBConv6, k3 × 3	3 × 2	320	4
9	Conv1 × 1 & Pooling & FC	3 × 2	1280	1
10	FC		256	1

Model input

The input of the model is a distribution of the number of photons detected with each photosensor on the inner face. Since the MPPCs are aligned in a matrix, the light distribution can simply be regarded as a two-demential grayscale image of 93 × 44 pixels. The shift invariance of the CNN is suitable for this application because the pileup γ-rays can be anywhere on the image.

Although the model adopted for the pileup elimination uses only the light distribution on the inner face as its input, it is also possible to use the light distributions on the other faces. Indeed, the model which uses the light distribution on the inner and the outer faces was tried in expectation of the improvement of the identification performance for the pileup γ-rays with deep conversion points. However, it could not outperform the model only with the inner face. This can be because it fails to benefit from the increased information due to its poor architecture or training procedure, and thus there is room for improvement.

Model output

The output of the model denotes how likely the event includes pileup γ-rays. If the event is a single γ-ray event, it should be close to zero while if the event has multiple γ-rays, it should be close to one.

Training dataset

The model was trained using a dataset generated with MC. It consists of two types of data: for main γ-rays and for pileup γ-rays. The main γ-ray dataset includes only one γ-ray in each event, and it was generated from the uniform energy spectrum from 20 MeV to 100 MeV. The uniformity and the wide energy range make the model independent of deposited energy. The energy independence is important since a fraction of pileup events in the background events depends on the energy. If the model learns to exploit information of the energy, it may just predict a high energy event as a pileup event, for example. The pileup γ-ray dataset was resampled from the original pileup γ-ray data from muon decays in order to moderate its energy dependence of the frequency as shown in Fig. 5.10. Otherwise, the training becomes much more difficult as too many low energy pileup γ-rays appear, which are hardly found. The energy cut of $E_\gamma > 0.2\,\mathrm{MeV}$ was also applied to ignore the pileup γ-rays which do not affect the reconstruction.

Approximately, 1.6×10^5 and 1.2×10^5 events were prepared for two types of the event, respectively. A part of the samples was randomly selected at a fraction of 10% for each event type. The selected samples were isolated from the others to be used to validate whether the model is over-fitted to the training dataset.

For the training, events belonging to each class, i.e. single γ-ray events and pileup events, and their class labels are required. The former events was simply prepared by sampling from the main γ-ray dataset, which were labeled zero. The latter events were generated by mixing events sampled both from the main γ-ray and the pileup γ-ray dataset assuming their hit timings are the same. These events were given labels of one.

To be realistic, the noise data taken with the random trigger without the muon beam were added to the simulated waveforms, and the light distribution was extracted by analyzing them. In addition, the input pixels corresponding to the dead channels observed in the run 2021 were masked (Fig. 2.68(d)).

Data pre-processing

Data pre-processing was applied to the inputs. Since there are some dead channels in the LXe detector, the values of pixels corresponding to such channels were estimated by a mean of surrounding pixel values. Normalization was applied with the maximum value to suppress an energy dependent performance. Furthermore, the negative values were cut off and zero was set instead to avoid unexpectedly large negative inputs; the number of photons calculated from the

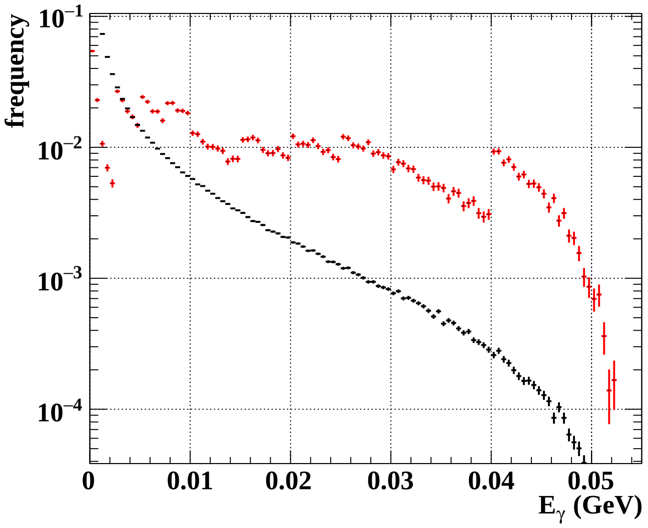


Figure 5.10 Energy spectra of the pileup γ-rays for the original (black) and the resampled (red) dataset.

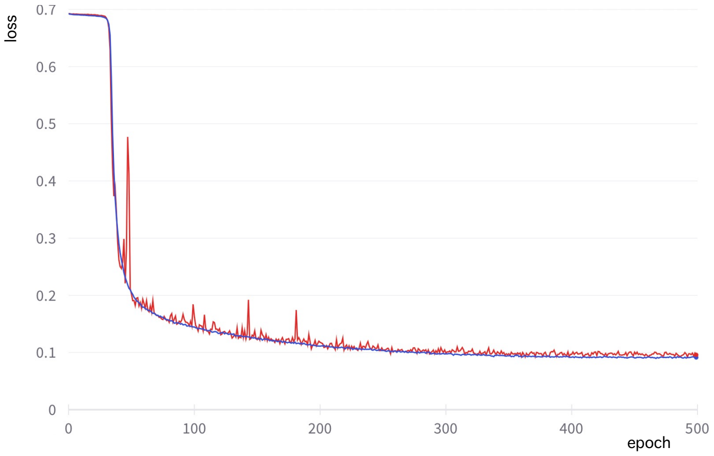


Figure 5.11 Transition of the loss for the training (blue) and the validation (red) dataset.

integrated charge can be estimated to be negative if accidental pileup γ-rays are in a baseline calculation region. These two operations assure the input values range from zero to one.

Training procedure

The model was trained to minimize a loss function of binary cross entropy using a stochastic gradient decent with an initial learning rate of 0.01. The learning rate was reduced with cosine annealing scheduler. The batch size was set to 200, and the training was stopped after 500 epochs. It took about 15 hours with Tesla P100-PCIE-16GB. The training successfully converged and the loss for the validation dataset was reduced to the equivalent level for the training dataset as shown in Fig. 5.11.

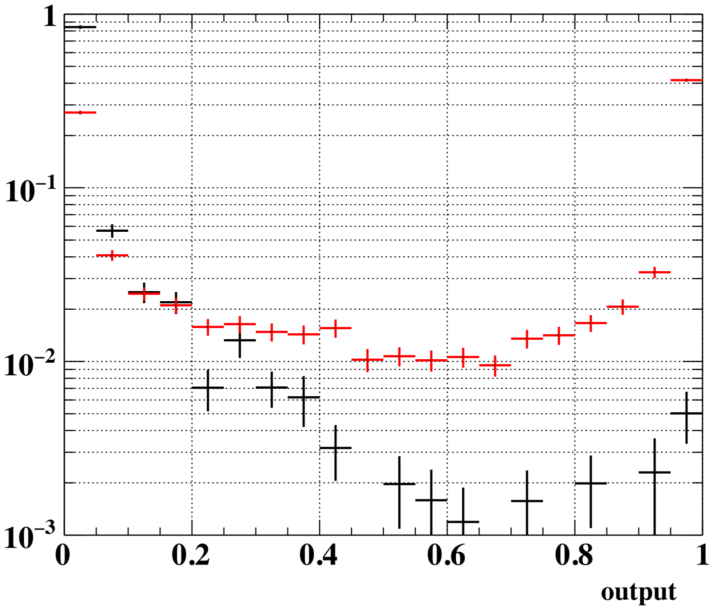


Figure 5.12 Model outputs for the background events with a single γ-ray (black) and multiple γ-rays (red).

Prediction

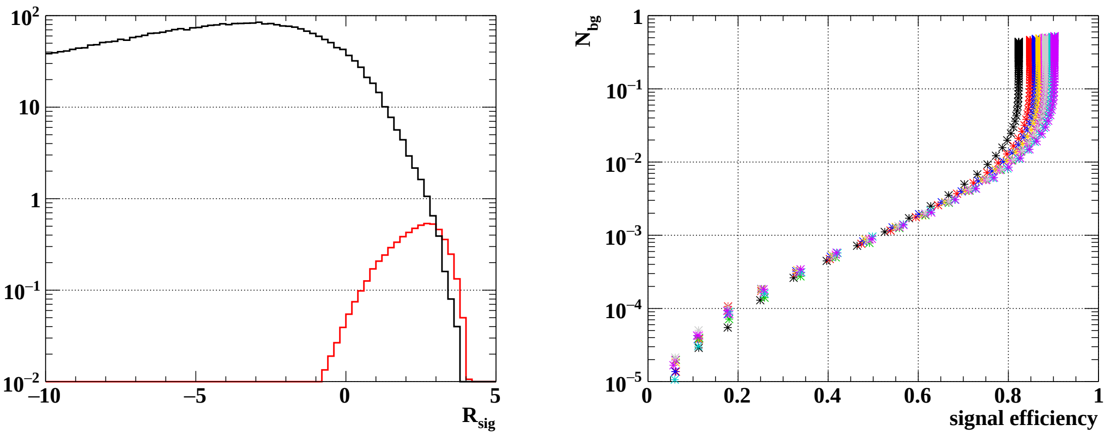
At the prediction stage, the model estimates how likely each event includes pileup γ-rays. Fig. 5.12 shows distributions of the model outputs for the background events with a single γ-ray or multiple γ-rays. Almost all single γ-ray events are estimated to be close to zero. On the other hand, the distribution of multiple γ-ray events has two peaks around zero and one and a tail in-between. The peak around zero derives from low energy pileup γ-rays, which are too difficult to be identified, and the events with moderate energy pileup γ-rays distributes in the middle. Otherwise, the events are successfully predicted to be near one.

Threshold scan

To decide whether an event has pileup γ-rays, the continuous output value given by the model needs to be translated to a binary flag. This can be done by setting a certain threshold for the value. A threshold scan was performed to maximize the signal-to-background ratio of the signal likelihood R_{sig} defined as

$$R_{\text{sig}} := \log_{10}\left(\frac{S(\boldsymbol{x}_i)}{f_R R(\boldsymbol{x}_i) + f_A A(\boldsymbol{x}_i)}\right),$$

where f_R and f_A are fractions of RMD and accidental background, which are set to 0.1 and 0.9, respectively. Fig. 5.13(a) shows the R_{sig} distributions for signal and background events. A signal box was defined by setting a lower threshold in these distributions, and the number of background events in the box was counted. Fig. 5.13(b) shows the relation between the signal efficiency and the number of background events with different thresholds of the model output,



(a)

(b)

Figure 5.13 Threshold scan to decide the threshold for the model output. (a) R_{sig} distributions for the background (black) and the signal events (red) for the threshold of 0.40. The number of the backgrounds is normalized by the expected value for three years of data-taking, and that of the signals is normalized assuming the branching ratio of 6×10^{-14}. (b) Relation between the number of backgrounds and the signal efficiency with the threshold of 0.10 (black), 0.20 (red), 0.30 (blue), 0.35 (orange), 0.40 (green), 0.45 (magenta), 0.50 (gray), 0.6 (light blue) and 0.7 (violet). The number of backgrounds is normalized by the maximum value.

and some values are summarized in Table 5.2. The threshold of 0.40 was chosen since the least number of backgrounds was achieved at most signal efficiencies.

Table 5.2 The number of backgrounds at different signal efficiencies with the different thresholds for the model output. The least value at each signal efficiency is written in boldface.

Signal efficiency	The number of backgrounds								
	0.10	0.20	0.30	0.35	0.40	0.45	0.50	0.60	0.70
0.3	0.00022	0.00025	0.00027	0.00027	**0.00021**	0.00025	0.00026	0.00024	0.00027
0.4	0.00046	0.00047	0.00050	0.00050	**0.00045**	0.00050	0.00049	0.00050	0.00052
0.5	0.00095	0.00092	0.00098	0.00099	**0.00084**	0.00093	0.00093	0.00097	0.00091
0.6	0.00204	0.00186	0.00192	0.00178	**0.00171**	0.00177	0.00182	0.00182	0.00173
0.7	0.00488	0.00427	0.00411	0.00392	0.00368	0.00377	0.00378	0.00360	**0.00355**

5.2.2 Peak search and clustering of photosensors

The second step consists of peak searches and clusterings of photosensors. The pileup γ-rays are searched for both in the light distribution and the timing distribution. The searches in the two different distributions are complementary to each other since the former one mainly focuses on the on-timing pileups while the latter focuses on the off-timing pileups. The information from the deep learning model is used for the search in the light distribution on the inner face.

Search in light distribution

The peak search in the light distribution finds local maxima of the number of detected photons on the inner face and the outer faces separately. It is performed after applying two filters so as not to pick up fake peaks, which are generated by the shower fluctuation or the noise. One is re-binning to reduce the granularity of the MPPC readout in which 2×2 channel values are summed up while the re-binning is not applied to the PMTs. The other is a low-pass filter with a moving average of 3×3 bins. They are applied both to the MPPC and PMT readout after the re-binning. The channel whose excess of the number of photons over the expectation from the surroundings is larger than a threshold of 200 photons as a nominal value is identified as a peak center. A two-dimensional Gaussian is fitted to the found peak, and the energy is roughly reconstructed from the integrated volume of the Gaussian.

The information from the deep learning model is used to select the peak search flow on the inner face as shown in Fig. 5.14 in the belief that the model correctly tells us the existence of the pileup γ-rays. If the model judges the event has only one γ-ray, the peak search is performed with the nominal threshold. In the case that more than one peak is found by the peak search against with the prediction, only the peaks whose reconstructed energies are larger than 10 MeV are regarded as true peak; otherwise the found peaks are discarded. This leads to avoiding picking up fake peaks. If the model judges the event is not a single γ-ray event, the peak search is repeated until more than one γ-ray is found lowering the threshold down to 30% of the nominal one. This treatment helps it to find pileup γ-rays whose conversion positions are deep, which cannot be found by the nominal threshold since the detected number of photons on the inner face can be small. There is the case that no additional peak is found even with the minimum threshold. In such a case, the peak search is given up, but the information that the deep learning model regards the event as a pileup-like event is recorded for the next step.

In addition to the search in the normal light distribution, the search in the "inverted" light distribution is carried out. This aims to identify the pileup γ-ray entering in the baseline calculation region of the waveforms. The number of photons calculated from the integrated charge is biased to be low if the baseline is calculated with the signals, and it can be negative value with a large magnitude (Fig. 5.15). Since the magnitude can be larger as the pulse height in the baseline region becomes higher, there can be a local minimum near the conversion point of the pileup γ-ray. It is searched for in the light distribution after being multiplied by -1. This peak search is performed only with the nominal threshold because the deep learning model is insensitive to such pileup γ-rays, which are not included in the training samples.

The channels around the found peaks are clustered. The clustering starts from the peak channel, and the neighboring eight channels whose detected number of photons are more than the greater of either 50% of the peak channel value or 100 photons are added to the cluster. It is repeated until the update of the members stops. If hit positions of γ-rays are close to each other, some channels can be assigned to several clusters. Such channels are reassigned to the cluster to which the distance from them is the shortest. The cluster timing and position are defined by the pulse timing and the position of the peak channel.

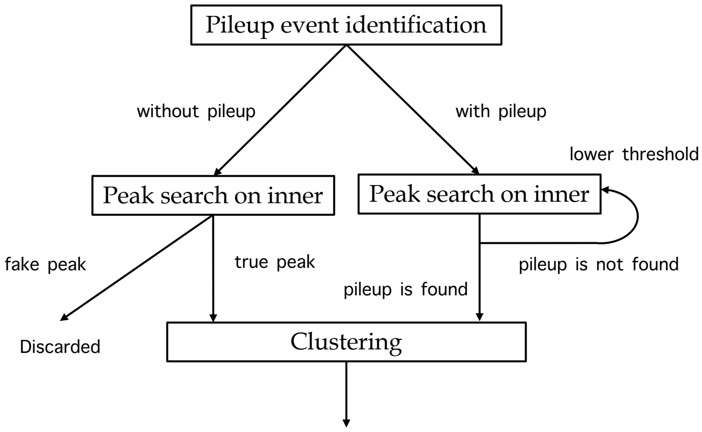


Figure 5.14 Flow of the peak search algorithm in the light distribution on the inner face.

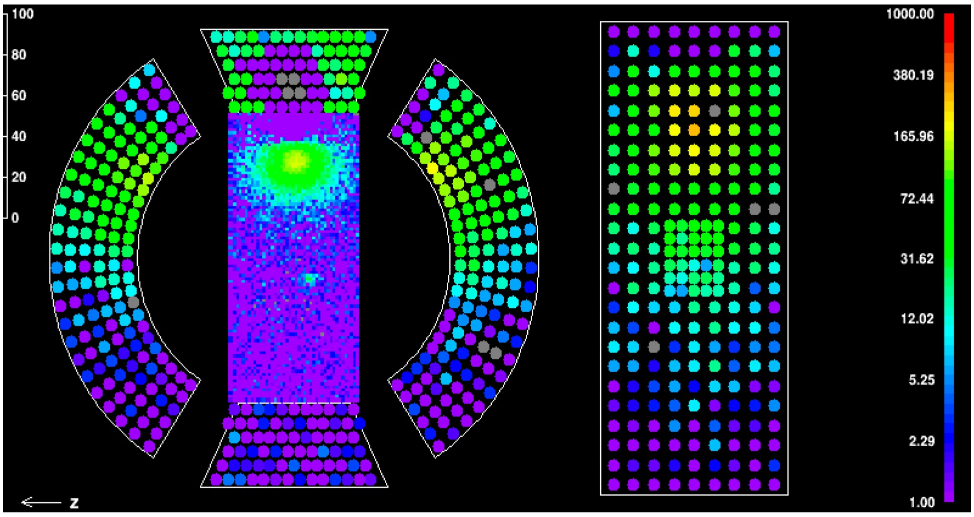


Figure 5.15 Example of the light distribution of the event with a pileup γ-ray in the baseline region. The purple region around the top face reflects the negative value due to the pileup.

Search in timing distribution

The search in the timing distribution is performed on all the faces with the timing χ^2 value of each channel $\chi^2_{\mathrm{pm},i}$ defined by 3.13 because the channels that are affected by the off-timing pileups have large values (Fig. 5.16). The channels with $\chi^2_{\mathrm{pm},i} > 50$ are clustered face by face. To avoid picking up the noise, channels with less than 50 photons or amplitude smaller than 5 mV are ignored. The center of the cluster is defined by the mean of sensor positions in the cluster, and the cluster timing is defined by the sensor timing of the central channel.

The clusters are merged based on the distance between the cluster centers in the u, v and θ directions. The criteria are defined depending on the combination of the faces of the clusters as summarized in Table 5.3. The thresholds, d_{thr} and θ_{thr}, are set to 18.6 cm and 16.75°, respectively, which are defined based on the distance between the sensors.

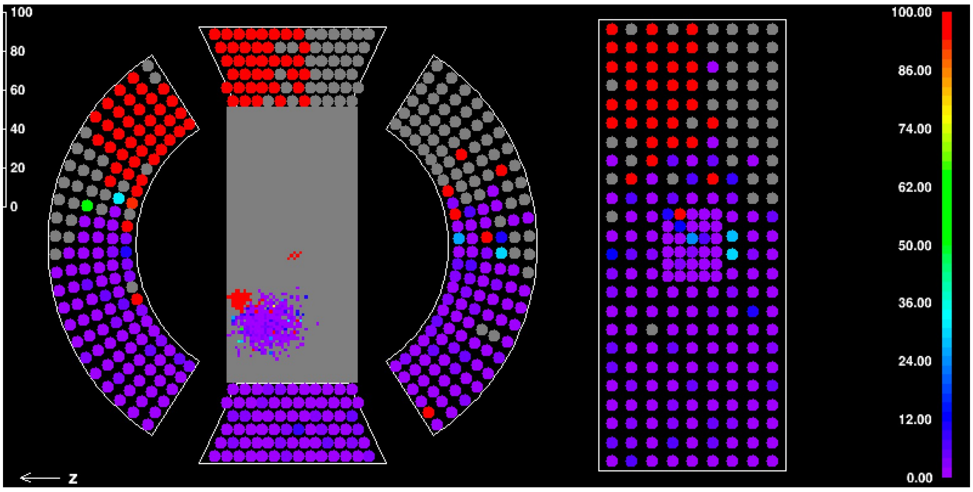


Figure 5.16 Example of the $\chi^2_{\mathrm{pm},i}$ distribution of the event with an off-timing pileup γ-ray. The channels in red have large values.

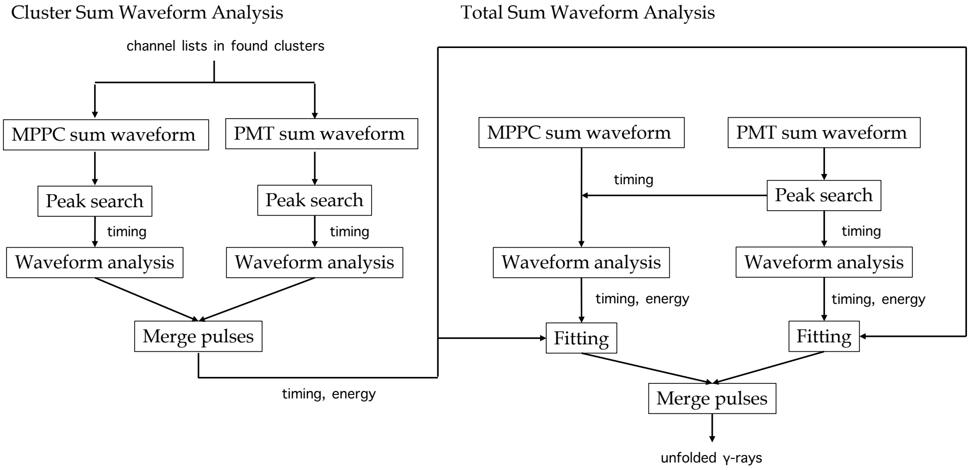


Figure 5.17 Flow of the unfolding algorithm of sum waveforms. It consists of two parts: cluster sum waveform analysis (left part) and total sum waveform analysis (right part).

5.2.3 Unfolding of sum waveform

Finally, the multiple γ-rays are unfolded using sum waveforms to reconstruct the energy of the main γ-ray. The algorithm is composed of mainly two parts as shown in Fig. 5.17. The first part analyzes the sum waveform of channels in each cluster. The second part analyzes the sum waveform of all channels. The results are combined, and they are used for fitting of template waveforms.

Cluster sum waveform analysis

The sum waveforms in each cluster are analyzed to extract the timings and energies of the cluster. The sum waveform is generated for MPPCs and PMTs separately by summing up the individual waveforms weighted by the sensor responses used in the energy reconstruction (Eq. (3.17)), which enables the reconstruction of the γ-ray energy directly from the area of the

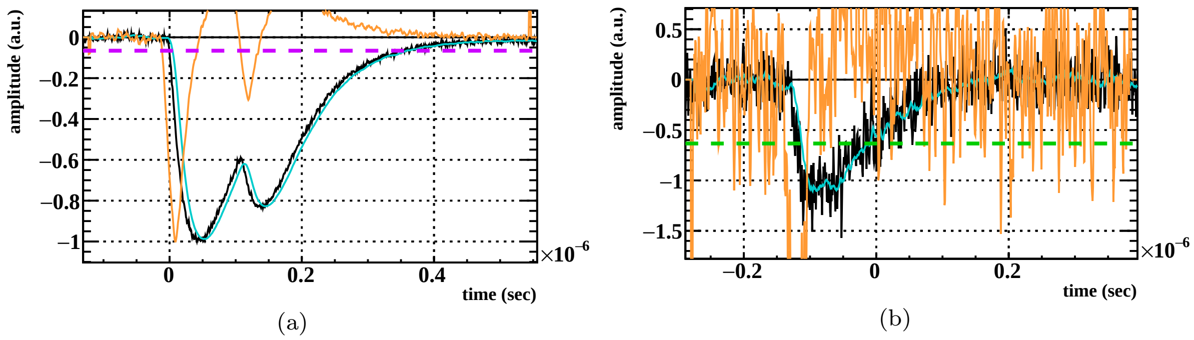
Table 5.3 Criteria to merge clusters found by the timing based clustering. dX is given by the magnitude of the difference of X between the two clusters.

Face		Criterion
Cluster 1	Cluster 2	
Inner	Inner	$\sqrt{du^2 + dv^2} < d_{\mathrm{thr}}$
	Outer	$d\theta < \theta_{\mathrm{thr}} \cap dv < d_{\mathrm{thr}}$
	Upstream Downstream	$(d\theta < \theta_{\mathrm{thr}} \times 1.5 \cup du < d_{\mathrm{thr}} \times 1.5) \cap dv < d_{\mathrm{thr}}$
	Top Bottom	$(d\theta < \theta_{\mathrm{thr}} \cup du < d_{\mathrm{thr}}) \cap dv < d_{\mathrm{thr}} \times 1.5$
Outer	Outer	$d\theta < \theta_{\mathrm{thr}} \cap du < d_{\mathrm{thr}}$
	Upstream Downstream	$dv < d_{\mathrm{thr}}$
	Top Bottom	$d\theta < \theta_{\mathrm{thr}}$
Upstream Downstream	Upstream Downstream	$dv < d_{\mathrm{thr}}$
	Top Bottom	$(d\theta < \theta_{\mathrm{thr}} \cup du < d_{\mathrm{thr}}) \cap dv < d_{\mathrm{thr}}$
Top Bottom	Top Bottom	$(d\theta < \theta_{\mathrm{thr}} \cup du < d_{\mathrm{thr}}) \cap dv < d_{\mathrm{thr}}$

sum waveform. The timings of waveforms are shifted by time offsets and propagation times from the reconstructed γ-ray hit position of each channel, which are used for the timing reconstruction (Sec. 3.5).

A peak search is applied to two shaped waveforms: a differential waveform and a moving-averaged waveform. The differentiation of the waveform is calculated by taking the difference between amplitudes of two points given points apart each other. Before taking the differentiation, a low-pass filter of the moving average is applied to reduce zig-zag structures due to the noise. The peak search tries to find local minima below a given threshold. The threshold is defined as RMS on the baseline multiplied by a certain factor. The parameters for the shaping and the peak search are summarized in Table 5.4. The peak search algorithm in the two waveforms is performed in two steps (Fig. 5.18). First of all, it is applied to the differential waveform. The threshold for the differential waveform becomes too low to find any pulses when the number of channels used for the sum is small. In such a case, pulses are searched for in the moving-averaged waveform. Note that the pulses which are close to each other can be detected better in the differential waveform since they can be separated by the differentiation while the moving average smears them.

The pulse timings and energies are extracted from the original sum waveforms based on the



(a) (b)

Figure 5.18 Two examples of the cluster sum waveforms: the original sum waveforms in black, the differential waveforms in orange, the moving-averaged waveforms in light blue. (a) Peak search is applied to the differential waveform and (b) to the moving-averaged waveform. The peak search thresholds are given by the dashed line in violet and green, respectively.

found peak timings. The pulse timings are defined by the constant fraction method. The energies are calculated by integrating the waveforms. The pulses found in the different clusters whose extracted timings are within 15 ns are merged. Note that the on-timing pileups are merged to the main γ-ray here since they cannot be unfolded by the waveform.

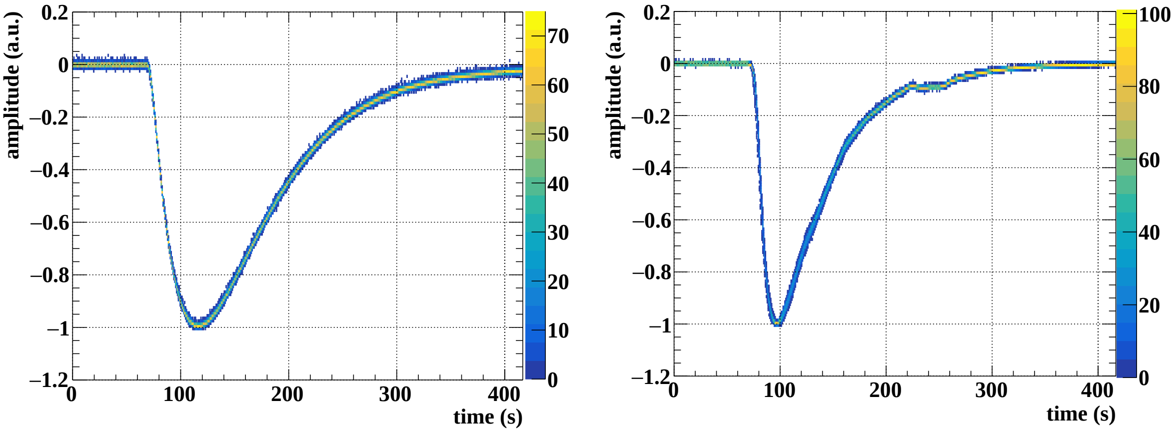
Table 5.4 Parameters for the waveform shaping and the peak search.

Waveform	Parameter	Value
Differential	Averaging points	29 points
	Differential points	5 points
	Threshold	10× RMS on baseline
Moving-averaged	Averaging points	21 points
	Threshold	15× RMS on baseline

Total sum waveform analysis

The sum waveforms of all MPPCs and PMTs are generated in the same way. The peak search is applied only to the differential waveform of the PMTs since a larger statistical fluctuation and noise fluctuation can be expected when the MPPC PDE gets smaller. The pulse timings and energies are extracted from the sum waveforms both of the MPPCs and the PMTs referring to the peak timings in the PMT waveform. For each of the MPPC and the PMT, the pulses in the sum waveforms are merged with the pulses found in the cluster sum waveforms based on the timings.

Now, we have a list of pulses, each of which corresponds to a temporally separated γ-ray. The pulses are unfolded by fitting the superposition of N_{pulse} template waveforms to the total sum waveforms independently, where N_{pulse} is the number of pulses in the list. The template waveforms are created from the sum waveforms of many single γ-ray events as shown in Fig. 5.19. A relative event-by-event fluctuation of the pulse shape is also computed from the standard



(a)

(b)

Figure 5.19 Superimposition of sum waveforms of (a) the MPPCs and (b) the PMTs. The amplitude is normalized and the pulse timing is aligned event by event. The second peak in the PMT waveform at 250 ns is due to a reflection of the signal.

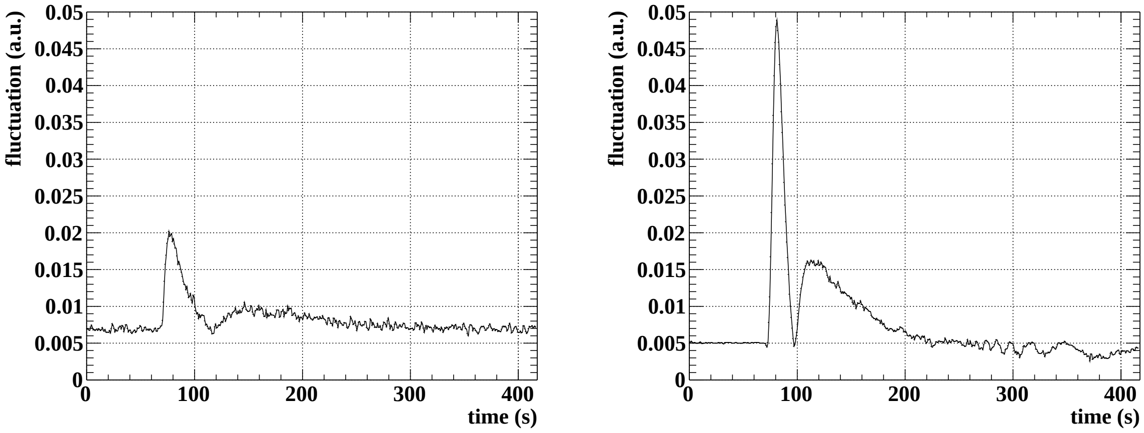
deviation of the amplitude at each timing as shown in Fig. 5.20, which is used to weight the difference of amplitudes between the raw waveform and the estimated waveform at each timing in the fitting. The fitting parameters are the timing and amplitude of each pulse, and the baseline. The extracted pulse timings and energies from the waveform analyses are used as initial values for them. In the fitting, the deviation of the raw waveform and the sum of the template waveforms is minimized. Fig. 5.21 shows a typical sum waveform with pileups. The sum waveforms are unfolded to two pulses. The energy of each γ-ray is reconstructed from the fitting results. The pulses in the MPPC waveform and the PMT waveform are merged by the timings. A pair of pulses whose timings are closest to the trigger timing is regarded as the main γ-ray.

5.2.4 Event status assignment

At the end, an event status is assigned based on the results of the three steps. There are five categories as follows:

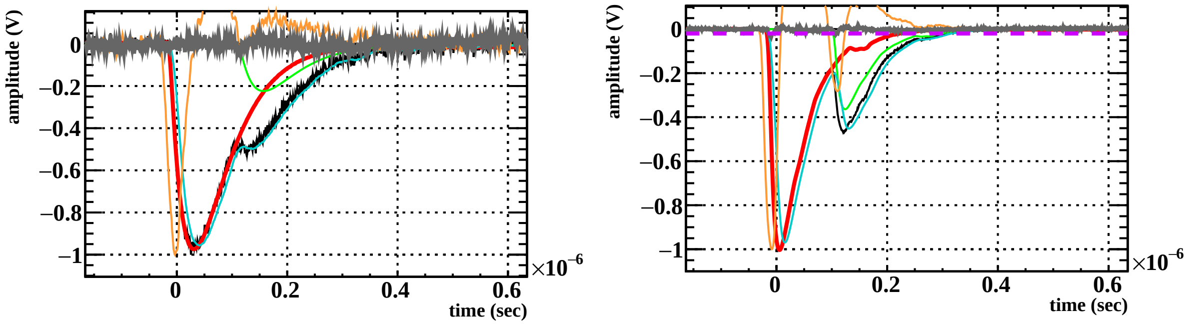
- **NoPileup** : No pileup is found, and the fitting in the total sum waveform successfully converges with a single pulse.
- **Unfolded** : More than one γ-ray is found, and they are successfully unfolded.
- **Coincidence** : On-timing pileup γ-rays are found, which cannot be unfolded.
- **DLRejected** : The deep learning model identifies pileups, but no pileup γ-ray is found with the others.
- **NotConverged** : The fitting in the total sum waveforms fails to converge.

The third status is assigned to the events in which a pileup γ-ray found with the peak search at the second step are merged to the main as a result of the cluster sum waveform analysis.



(a) (b)

Figure 5.20 Event-by-event fluctuation of the sum waveform of (a) the MPPCs and (b) the PMTs defined by the standard deviation of amplitudes at each bin in Fig. 5.19. The local minima at 120 ns (MPPC) and 100 ns (PMT) are due to the normalization at the peak of the waveform.



(a) (b)

Figure 5.21 Sum waveform unfolding for a typical waveform with a pileup γ-ray of (a) the MPPC sum waveform and (b) the PMT sum waveform. The original waveform is shown in black with the differential waveform (orange) and the moving-averaged waveform (light blue). The waveform is deconvoluted to two pulses (red and green). The difference between the original waveform and the sum of deconvoluted waveforms is shown in gray. The two pulses are found by the peak search in the total PMT waveform with the threshold (dashed line in violet).

There is the case that the main γ-ray is found in the sum waveform of the pileup cluster in addition to the pileup pulse when their hit positions are close to each other. The found main γ-ray pulse is merged to the pulse found in the main γ-ray cluster since they have exactly the same timings. In order to avoid regarding such a case as the coincidence event, only the pulse whose extracted timing is the closest to the cluster timing is thought to be an on-timing pileup γ-ray pulse. Since the cluster timing is defined by the sensor timing of the representative, i.e. the peak/central channel of the cluster found in the light/timing distribution, the main pulse in the cluster has the closest timing to the cluster timing. The main γ-ray cannot be the main pulse in the pileup cluster, and thus the misidentification is suppressed.

The fourth category corresponds to the case the peak search in the light distribution fails to find more than one γ-ray even with the lower threshold. This can be caused both by the misidentification of the deep learning model or the incompleteness of the peak search algorithm. There is also the case the second peak is found in the light distribution, but no pulse is found with the waveform analysis when the second peak results from the noise.

The fifth status relies on the fitting result of the total sum waveforms. If there are pileup γ-rays which are not found by the former analysis, the sum of the fitted waveforms can differ from the raw waveform. Even if it is not the case, the fitting can fail when the noise is too large or too many pulses are in a short time range. The failure is detected by the deviation of the waveforms; the fitting is thought not to converge if the deviation is larger than a given threshold. The threshold is defined by ten times of the mean of the baseline deviation in order to consider the noise effect.

The first and the second statuses are the cases that γ-ray energies succeed in being reconstructed while the others are not. Therefore, the events having any of the latter three statuses are discarded from the analysis.

5.3 Performance evaluation with MC

The expected performance of the series of algorithms was evaluated with the MC for the full channel readout configuration. The MPPC PDE was set to 13% for all. To make the simulation realistic, noise data at the amplifier gain of 1 taken in 2021 were added to the simulated waveforms, and the signals of dead channels were discarded based on the knowledge in 2021. Hereafter, the performance of each algorithm is explained step by step followed by discussions on the contribution from the noise and the dead channels, and the effects of smaller MPPC PDEs and the precision of the sensor calibration.

5.3.1 Performance of each algorithm

Identification of pileup events with deep learning

In order to show how well the deep-learning-based identification works, the performance of the deep learning algorithm is compared to that of the peak search in the light distribution on the inner face with the nominal threshold, which uses the same distribution as inputs.

Fig. 5.22 compares the fractions of the pileup events correctly identified as a function of the conversion depth of the pileup γ-ray. In the shallow region, the performance of the two methods is equivalent. In the deep region, however, the fraction of the peak search method decreases rapidly while the other keeps the high detection efficiency. Fig. 5.23 shows an event example in which a pileup γ-ray converts in the deep region. Since no clear peak is made on the inner face by the deep pileup γ-ray hit, the peak search on the inner face cannot identify the second peak. On the other hand, the deep learning model predicts the event is likely to include pileup γ-rays with the output of 0.98.

Another event example is shown in Fig. 5.24 which has a single γ-ray. The peak search method finds the very local peak that derives from the conversion of a γ-ray escaped from the shower in addition to the main peak. On the other hand, the deep learning model predicts the event should be a single γ-ray event with output of 0.03.

Fig. 5.25 shows an example of event in which γ-rays enter the edge of the detector. A part of the light distribution originating from each γ-ray on the inner face is broken. However, it does not affect the prediction, and the model successfully outputs the value of 0.99. Fig. 5.26 shows the fractions of the pileup events correctly identified as a function of the u position of the main γ-ray. The fraction is kept high even near the edge of the fiducial volume.

Fig. 5.27 shows the relations between the number of backgrounds and the signal efficiency of the two methods. The number of backgrounds is normalized by that after applying the pileup elimination method in the previous study [62]. The smaller number of background events at the same signal efficiency is achieved by the deep-learning-based rejection because of the higher detection efficiency in the deep region and the tolerance to the fake peaks.

Although the deep learning model succeeds in achieving the high performance, it has a weak point that its prediction lacks the information of the found pileup γ-rays; it can only give whether the event is a pileup event or not while the peak search can tell us which peak is identified. The disadvantage makes it difficult to utilize its result in the following waveform unfolding. On the other hand, the low performance of the peak search in a deep region partly comes from the fact that the peak search threshold is set to a value high enough not to pick up the fake peaks. In other words, the peak search can find more pileup γ-rays if a lower threshold is used though this cannot be realized because of the existence of the fake peaks. Therefore, the operation to select the peak search flow based on the deep learning prediction is adapted to make use of each strength as explained in Sec. 5.2.2.

The success of the deep learning model can result from the optimization of the peak search procedure in the model. As explained in Sec. 5.2.2, the re-binning and the low-pass filter are explicitly applied in the peak search method to avoid picking up the fake peaks. This kind of process is done automatically in the CNN layers and the strength of the filtering and the threshold are optimized during the training. In fact, the two strengths of the deep learning discussed above imply that the deep learning model recognizes the peak-like structure not from a local peak structure but from more global one, which can be accomplished by the optimization.

Peak search and clustering of photosensors

Fig. 5.28 shows the energy spectra of the background γ-rays. The events in which more than one γ-ray is found with the peak search method are rejected. A pileup identification by waveform [62] is also applied to reduce the off-timing pileups, which searches for pileup γ-rays in the sum waveform by detecting an excess from the template waveform. The algorithms reduce 59% of the background events between 51.5 MeV and 54 MeV in total (red), which is dominated by the contribution from the peak search in the light distribution on the inner face (blue).

Not only the background events but also the signal events with pileup γ-rays or fake peaks are rejected, which results in the signal inefficiency. If we simply discard all the pileup identified

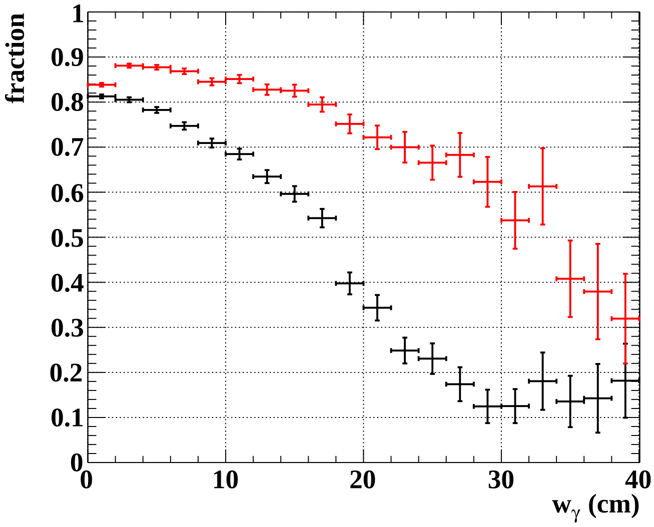


Figure 5.22 Detection efficiency of the pileup γ-rays for the peak search method (black) and the deep learning model (red) as a function of the conversion depth of the pileup γ-ray.

events, the signal inefficiency is estimated to be 27% including 10% rejection of single γ-ray events.

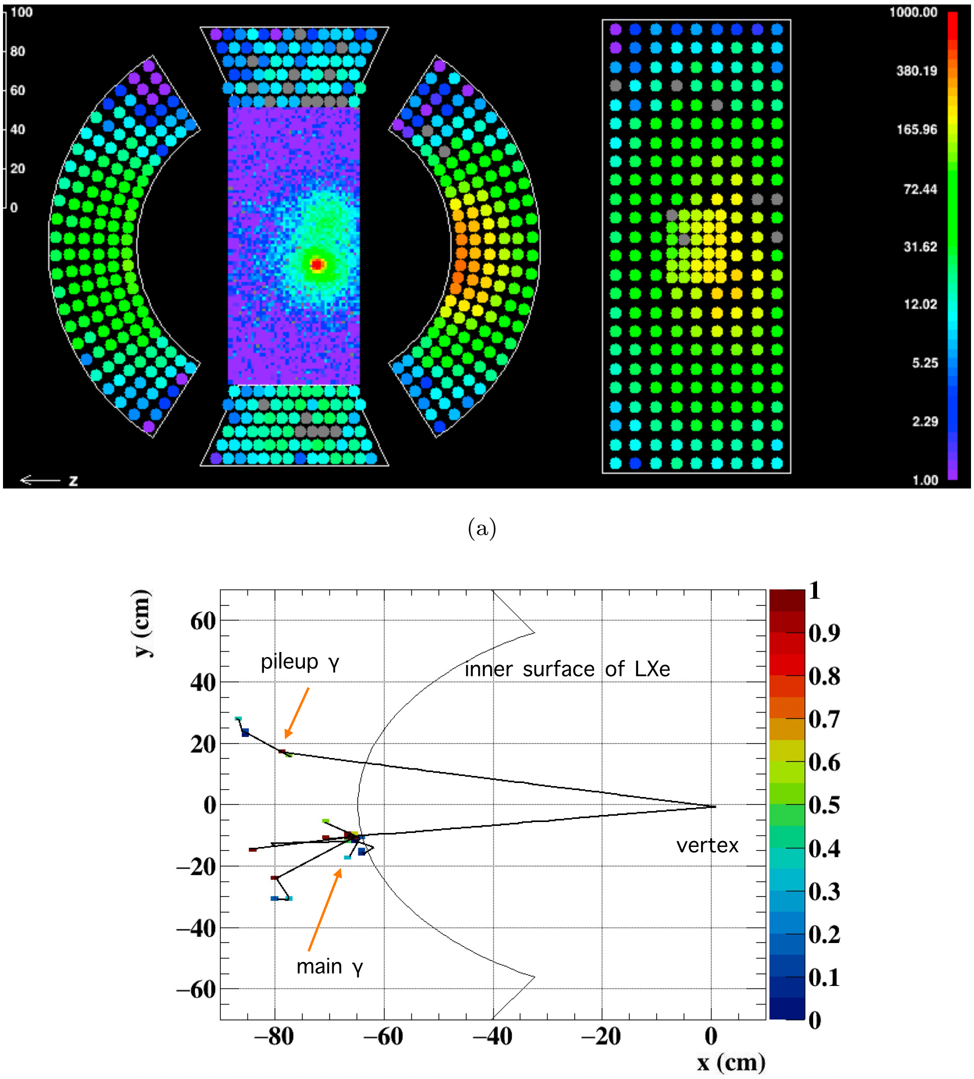
Unfolding with sum waveform

Fig. 5.29 shows the simulated energy spectrum of the background γ-rays with the sum waveform unfolding (blue). As shown in Fig. 5.30, the signal efficiency is recovered by 19% with an increase in the background events of 8% (green) compared to the rejection based on the peak search result (light blue). This would result in a better branching ratio sensitivity. The signal inefficiency of 7% coming from 4% of **Coincidence**, 2% of **DLRejected** and 1% of **NotConverged** event statues includes 5% rejection of single γ-ray events.

5.3.2 Effect of noise

The performance described above was evaluated by mixing the noise data to simulated waveforms in order to include non-Gaussian effects such as a correlated noise among channels and a spike. Here, the performance difference between a simulated Gaussian white noise and a real noise is discussed. Note that the noise level of the simulated noise is the same as the observation: 0.7 mV RMS.

Fig. 5.31(a) shows the predictions of the deep learning model for a single γ-ray with the real noise data. The distribution of the model trained with the real noise peaks at zero (black). On the other hand, that trained with the simulated noise tends to output higher values and forms a peak at one (red) while it correctly predict most of events to be zero for the simulated noise events (blue). This result implies that the real noise has an effect to create fake peaks, which deceives the model, and the deterioration can be moderated by training the model with the real noise. The predictions for multiple γ-ray events with the real noise are also shown in



(a)

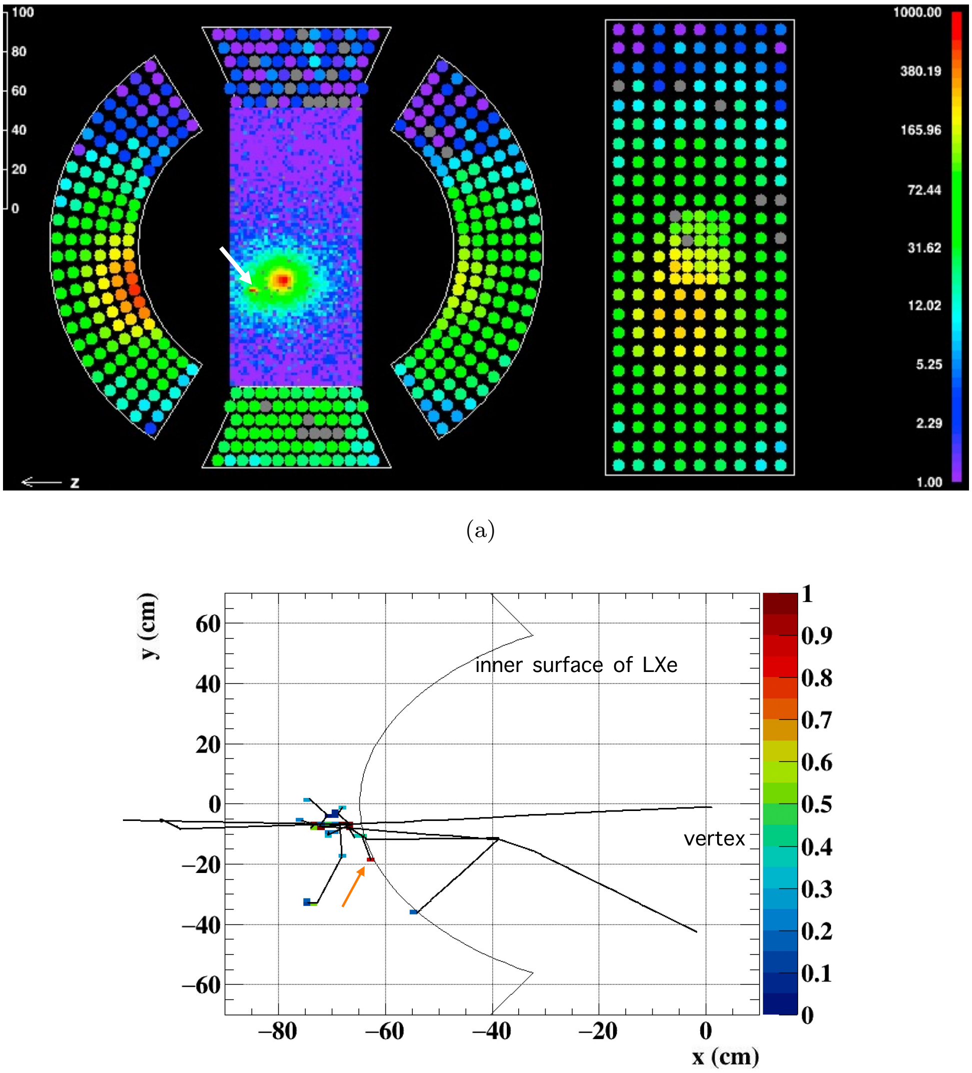
(b)

Figure 5.23 Example of an event with a pileup γ-ray that converts at a deep position. (a) The light distribution and (b) the particle track. The deep learning model predicts the event as a pileup event while the peak search on the inner face does not find the second peak. The scale of the bar is in the unit of MeV.

Fig. 5.31(b). In this case, no apparent difference is observed between the outputs of the model trained with the simulated noise for the two types of events though the peak at one of the model trained with the real noise is higher than the others. This indicates the noise difference does not affect the detection of pileup γ-rays so much.

A possible cause of the misidentified peak in the single γ-ray distribution is a correlated noise, which is not included in the simulated noise. As shown in Fig. 5.32, adjacent channels in the LXe detector are connected to the same WDB. Fig. 5.33 shows the noise distribution of the MPPCs added to the simulated waveforms. There are noises which coherently arise in the channels of the same WDB, which results in a cluster of channels whose energies are consistently bias to be positive or negative, and this cluster can be misidentified as a contribution of a pileup γ-ray.

Fig. 5.34 shows an example of a single γ-ray event with the simulated noise or the real noise.



(a)

(b)

Figure 5.24 Example of an event only with a main γ-ray. (a) The light distribution and (b) the particle track. The arrows point to a fake peak generated by a γ-ray escaped from the shower. The deep learning model predicts the event as a single γ-ray event correctly while the peak search on the inner face picks up the fake peak. The scale of the bar is in the unit of MeV.

The correlated noise forms a peak-like structure resulting in a misidentification of the pileup event. The deep learning model trained with the simulated noise successfully predicts the event is not likely a pileup event with the output of 0.0050 for the simulated noise event (Fig. 5.34(a)) while it incorrectly outputs the value of 0.92 for the real noise event (Fig. 5.34(b)). This misidentification, however, can be corrected by training the model with the real noise, and the output value decreases to 0.0031.

To see the effect of the coherent noise, the noise data are added shuffling the channel assignment on the inner face. This operation simulates with the same noise level, but ignoring the clustered noise structure (Fig. 5.34(c)). Fig. 5.35(a) shows the predictions of the model trained with the simulated model for the different noise datasets. The shuffle reduces the fraction of the

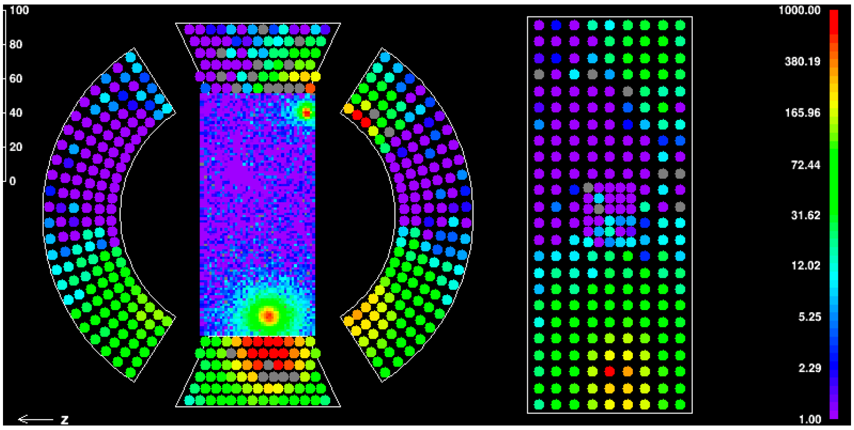


Figure 5.25 Example of an event in which γ-rays enter the edge of the detector. The main and the secondary γ-rays hit at $(u, v, w) = (-5.63, -60.24, 3.46), (-27.92, 59.51, 1.61)$ cm, respectively. There is another γ-ray hit in the baseline calculation region, i.e. 76.12 ns before of the main, at $(u, v, w) = (10.11, 23.53, 2.40)$ cm, which is a source of the depression on the outer face.

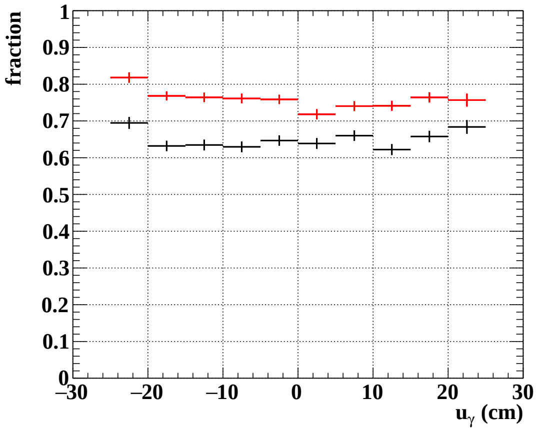


Figure 5.26 Detection efficiency of the pileup γ-rays for the peak search method (black) and the deep learning model (red) as a function of the u position of the main γ-ray.

misidentification to some extent (red), but there is still a peak around one, which can come from the other effects such as larger noise channels than the simulation. Fig. 5.35(b) shows the predictions of the models trained with the real noise with and without the shuffle for the shuffled noise datasets. The fraction of the model trained without the shuffle around one (black) is higher than that with the shuffle (red), which also implies the coherent noise contributes to form a peak-like structure. However, the difference is found to vanish if the models are trained in the same noise situation with the prediction (Fig. 5.35(c)). These facts tell us the sensitivity to the noise situation of the model and the importance to tune the model to the noise situations in order to maximize the model performance.

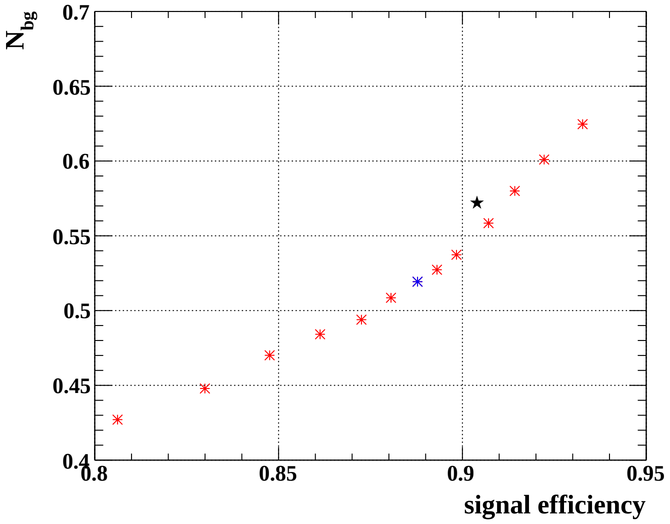


Figure 5.27 Relation between the number of background events and the signal efficiency ($E_\gamma = 51.5$–$54\,\mathrm{MeV}$) for the peak search method (black) and the deep learning model (red) when the events in which more than one γ-ray is found are rejected. The different points for the deep learning model correspond to the different thresholds for the rejection. The optimal point (threshold = 0.40) is in blue.

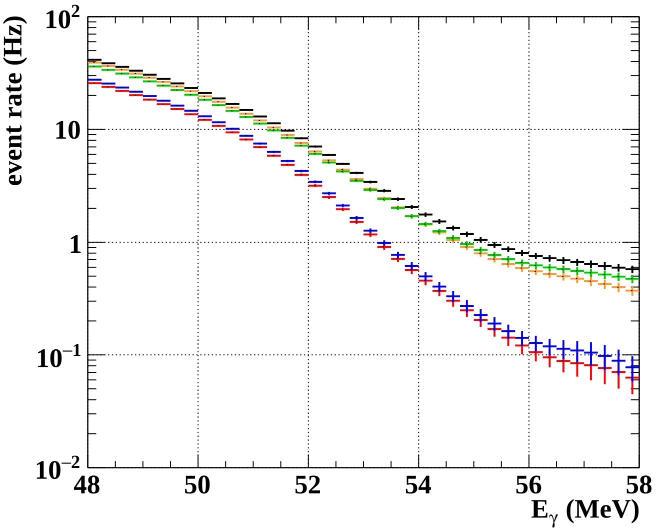


Figure 5.28 Energy spectra of the simulated background γ-rays without the peak search and clustering method (black) and with all the elimination methods (red), the search with the light distribution on the inner face (blue) and on the outer face (orange), and the search in timing distribution (green). A pileup identification by waveform [62] is applied to reduce the off-timing pileups.

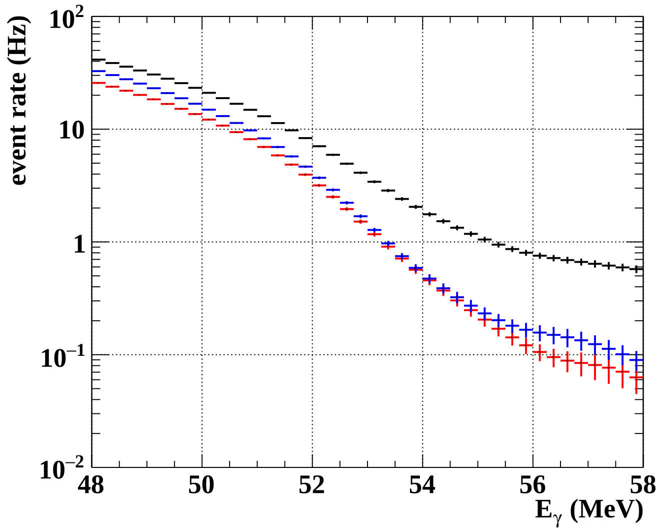


Figure 5.29 Energy spectra of the simulated background γ-rays without the peak search and clustering method (black), with the peak search elimination methods (red) in addition to the pileup identification by waveform [62], and with the waveform unfolding (blue).

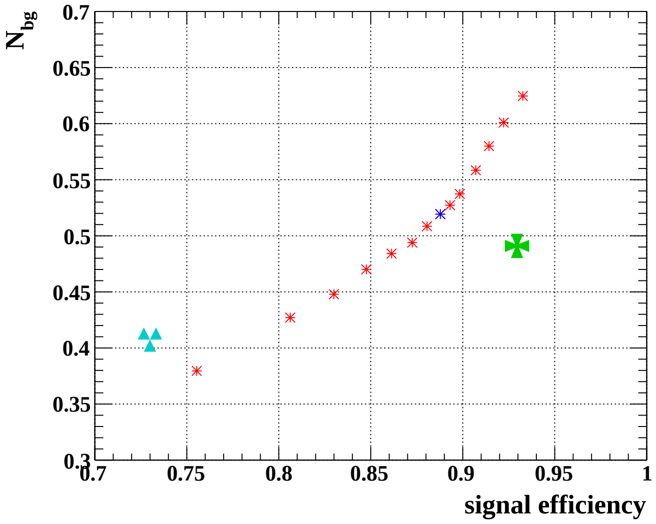
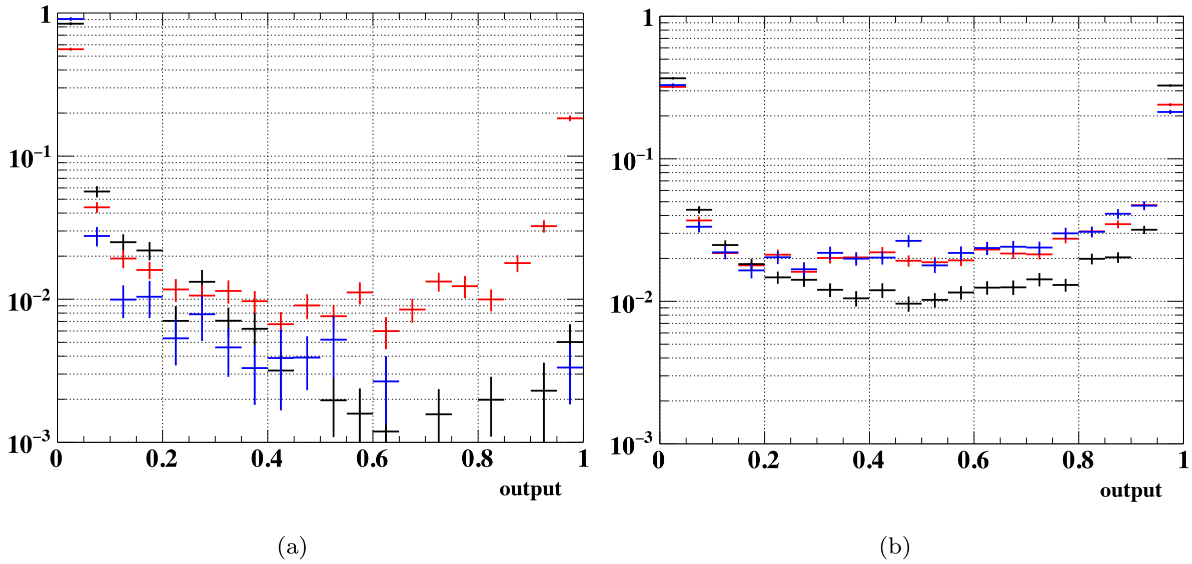


Figure 5.30 Relation between the number of background events and the signal efficiency (E_γ = 51.5–54 MeV) for the unfolding method (green) in addition to those for the peak search method (light blue) and the deep learning model (red) when the events in which more than one γ-ray is found are rejected. The different points for the deep learning model correspond to the different thresholds for the rejection. The optimal point (threshold = 0.40) is in blue.



(a)

(b)

Figure 5.31 Predictions of models for (a) single γ-ray events and (b) multiple γ-ray background events to which real noises are added. The models are trained with a real noise (black) and simulated noise (red). The predictions of the model trained with a simulated noise for simulated noise events are also shown (blue).

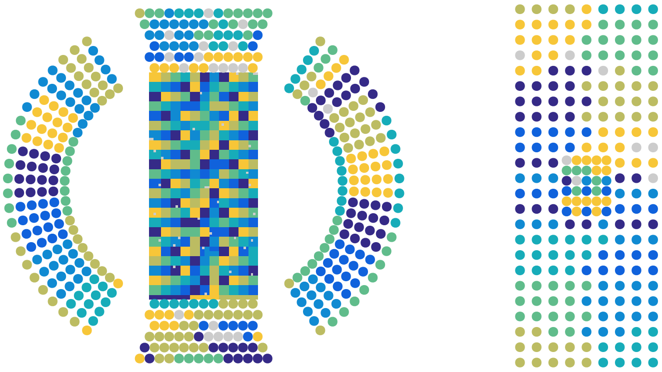
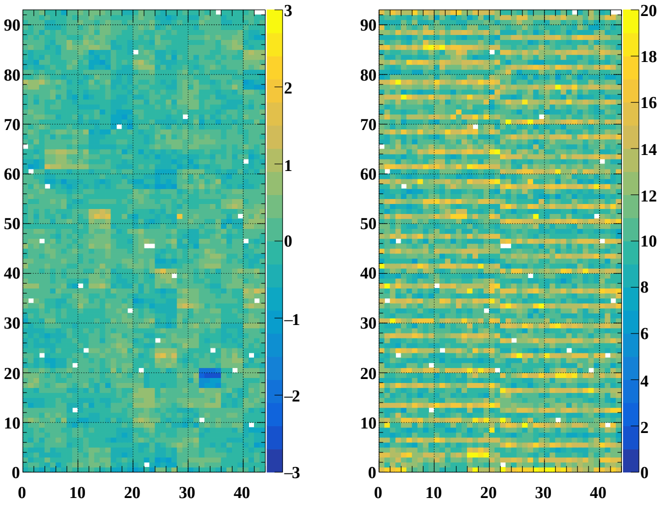


Figure 5.32 Assignment of readout electronics channels in run 2021. The adjacent channels in the same color are connected to the same WDB.

5.3.3 Effect of dead channels

The existence of dead channels can worsen the performance of the pileup elimination by loosing their information. Its effect was evaluated by comparing the case in which no dead channel is assumed to the case in which the dead channels found in the run 2021 are considered: 30 channels for the MPPCs and 28 channels for the PMTs.

Fig. 5.36 compares the predictions of the deep learning models trained with and without the dead channels for the background events under the same dead channel assumption with the training. No significant difference is observed in the distribution. The distribution of the model trained without dead channels for the background events with dead channels is also shown in the same plot (blue). There is little difference compared to the others as well. It is also found that the overall elimination power of a series of the algorithms is consistent within 1% precision.



(a) (b)

Figure 5.33 Noise distribution measured in run 2021 at gain 1. (a) Mean and (b) standard deviation of the number of photons over $\mathcal{O}(10^3)$ events for each channel are shown. Different offsets among the WDB are seen in (a). The striped pattern in (b) is due to the MPPC lot dependence of the correlated noise.

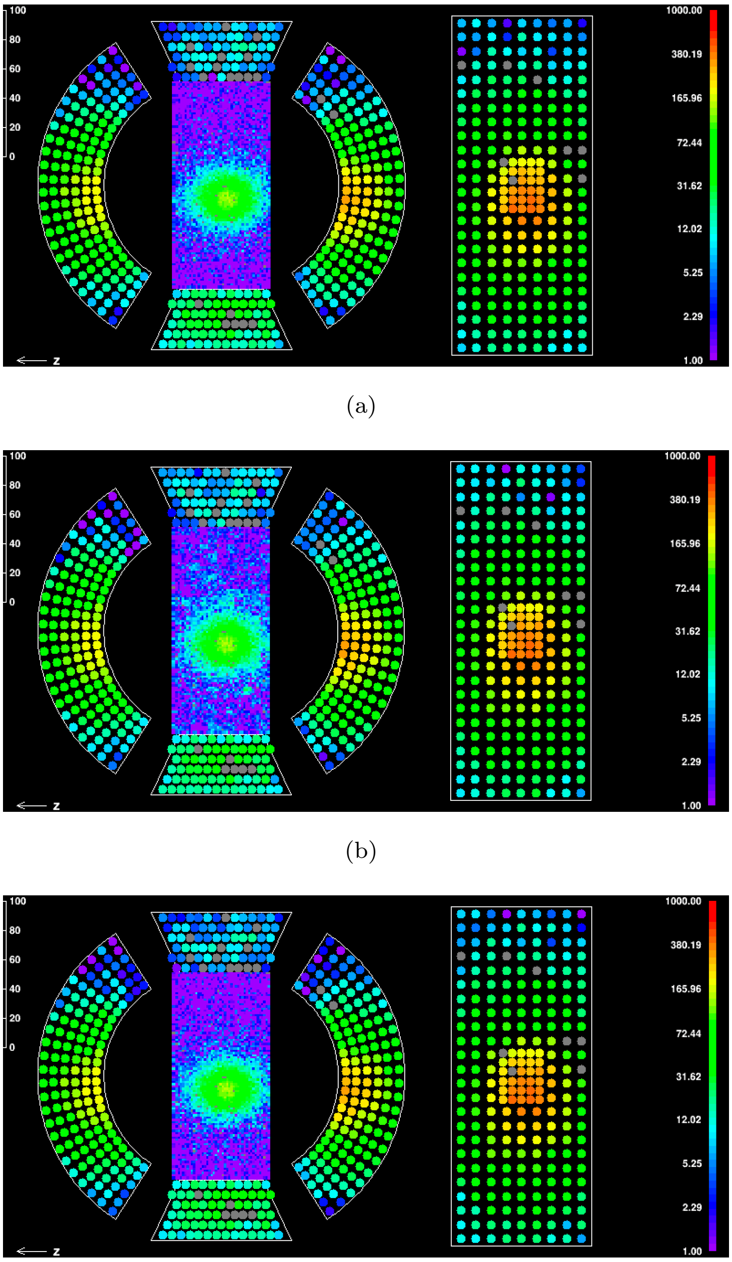
These observations tell us that the dead channels at the level observed in the run 2021 hardly affect the performance of the elimination.

5.3.4 Effect of PDE degradation

As discussed in Sec. 3.1.1, the LXe detector will be operated with the annual annealing to keep the reasonable MPPC PDE from 16% down to 2%. The effect of the reduced PDE can be partly compensated by increasing the electronics gain, but the lower signal-to-noise ratio can affect the pileup elimination performance. Therefore, the performance was also evaluated at lower PDEs. The amplifier gain was also changed accordingly following Table 3.3.

Fig. 5.37 shows the relations between the number of backgrounds and the signal efficiency when the MPPC PDE is 13% and 2%, where the deep learning model is trained with data at PDE of 13%. Both of the number of backgrounds and the signal efficiency increase due to the lower PDE. Fig. 5.38 compares those for 2% PDE, where the deep learning model is trained with data at PDE of 13% and 2%. The number of backgrounds for the model trained at 2% PDE is less at the optimal point while the signal efficiency is equivalent. As a result, that after being applied all the algorithms is slightly smaller with the similar signal efficiency.

Table 5.5 summarizes the number of backgrounds and the signal efficiency for the different MPPC PDEs. The number of backgrounds and the signal efficiency tend to increase as the PDE gets lower, which results from the gained insensitivity to both of true and fake peaks with small sizes. The difference between the models trained assuming the fixed PDE and the corresponding



(a)

(b)

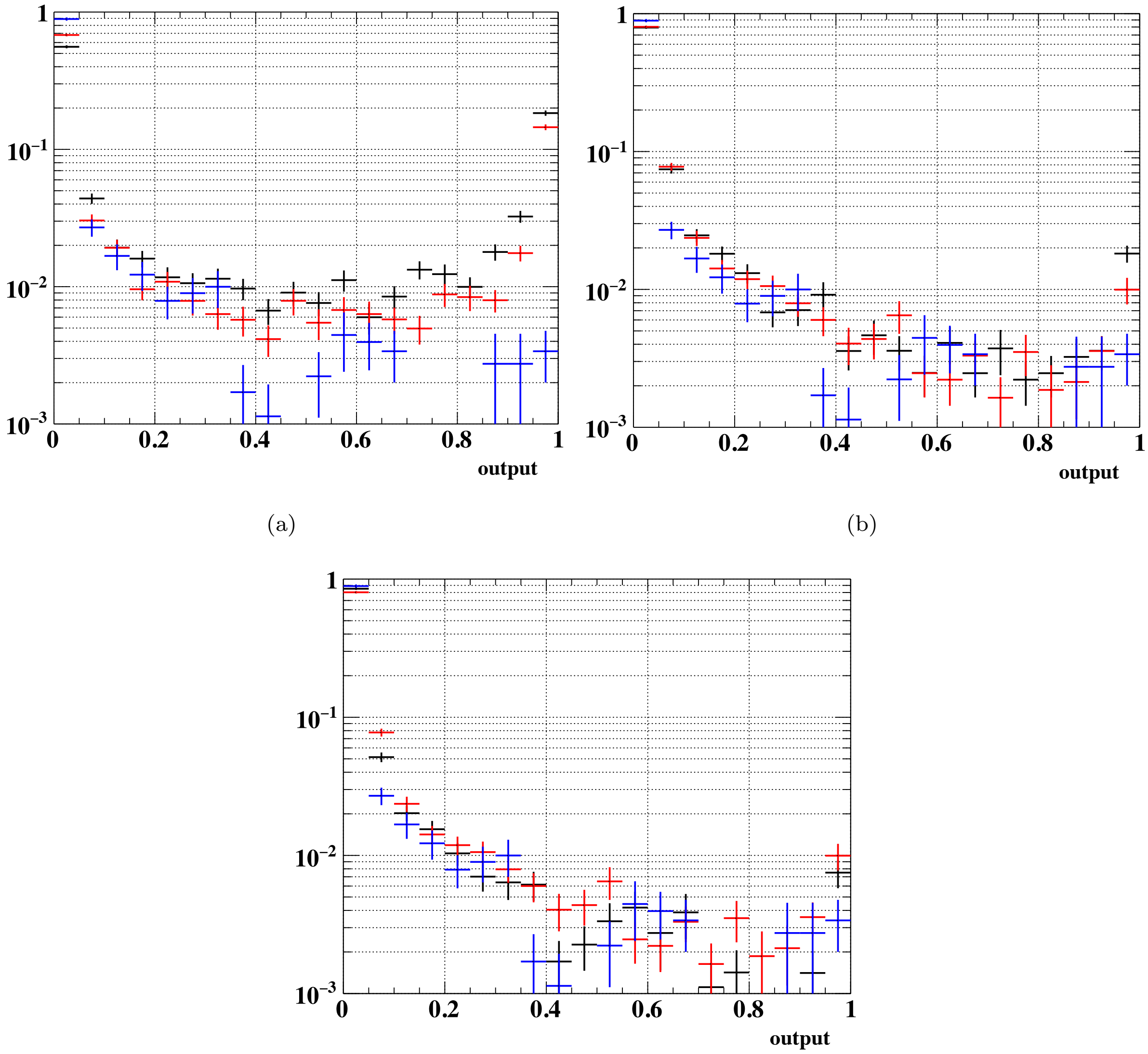
(c)

Figure 5.34 Example of an event without a pileup γ-ray when (a) simulated noise or (b) real noise is added. For (c) real noise is added like (b), but the channel assignment of the MPPCs is randomly shuffled.

PDEs is not significant.

5.3.5 Effect of uncertainty from calibration

The reconstructed number of photons can fluctuate and be biased due to the statistical and systematic errors in the calibration, which can affect the performance. The statistical uncertainty of the calibration factor was estimated to be 4% both for the MPPCs and the PMTs (Sec. 3.1.1). The systematic uncertainty was estimated by a half of the difference between the measured



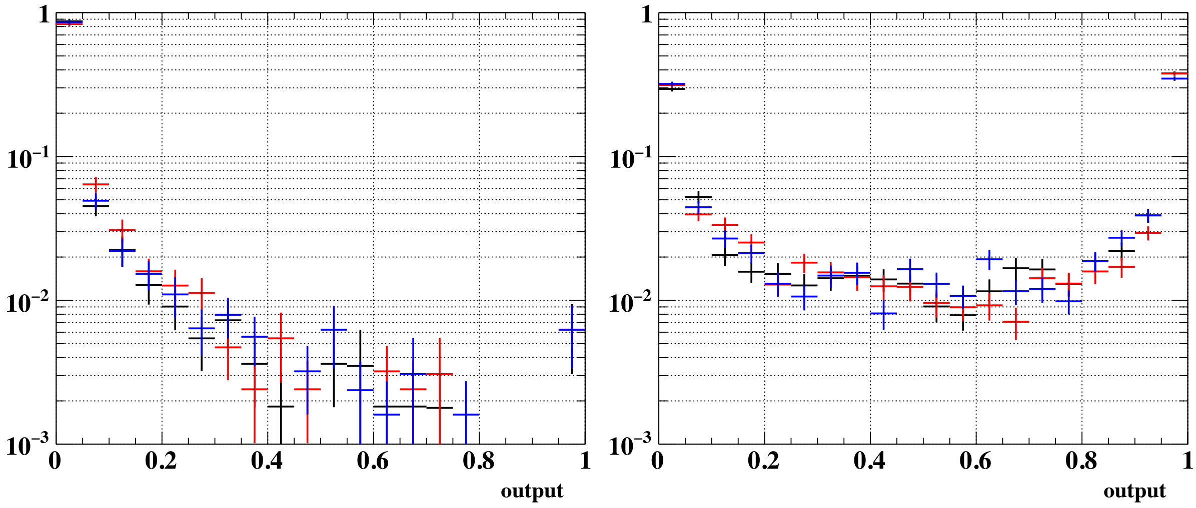
(a)

(b)

(c)

Figure 5.35 Predictions of models for single γ-ray background events. (a) The results for the real noise before (black) and after (red) the shuffle, and the simulated noise data (blue). The models are trained with the simulated noise. (b) The results for the shuffled real noise data. The models are trained with the real noise without (black) and with the shuffle (red). (c) The results for the real noise data without/with the shuffle predicted by the models trained with the real noise without/with the shuffle (black/red). The predictions of the model trained with a simulated noise for simulated noise events are also shown in all (blue).

calibration factors and those after applying the correction to get a uniform sensor response for CW γ-ray data as shown in Fig. 5.39. The effects of these uncertainties were investigated by artificially changing the calibration factors in the prediction stage while keeping the original values in the training stage. The factors were first systematically changed for the systematic uncertainty and then randomly fluctuated with 5% or 10% standard deviation. As a result, no deterioration of the performance was observed.



(a) (b)

Figure 5.36 Prediction of models for (a) single γ-ray events and (b) multiple γ-ray background events without (black) and with (red, blue) masking dead channels. The models are trained without (black, blue) and with (red) masking dead channels.

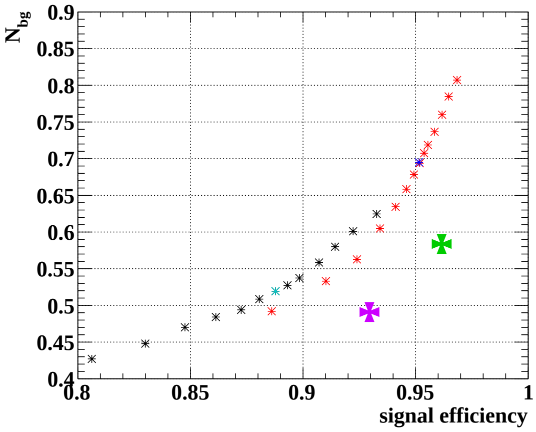


Figure 5.37 Relation between the number of background events and the signal efficiency ($E_\gamma = 51.5$–$54\,\text{MeV}$) for the unfolding method (violet/green) and the deep learning model (black/red) when the events in which more than one γ-ray is found are rejected when MPPC PDE is 13%/2%. The model is trained with dataset at MPPC PDE of 13%. The different points for the deep learning model correspond to the different thresholds for the rejection. The optimal point (threshold = 0.40) is in light blue/blue.

5.3.6 Summary of performance

The performance of the pileup elimination algorithms were evaluated taking the real noise situation and the dead channels into account. In total, the number of backgrounds was reduced by 51% keeping the signal efficiency of 93% for 51.5–54 MeV. The effects of the PDE degradation and the accuracy of the calibration parameters were also discussed. The lower PDE was found to increase both of the number of backgrounds and the signal efficiency by 9% and

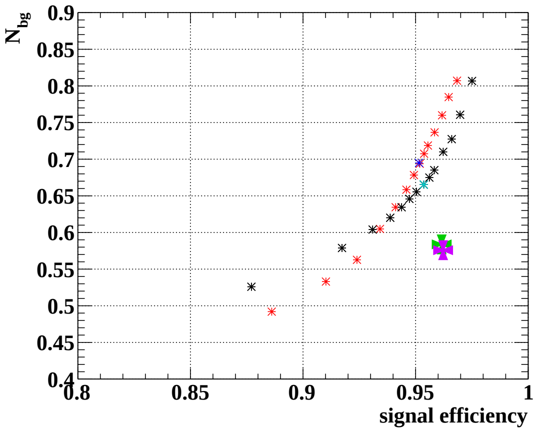
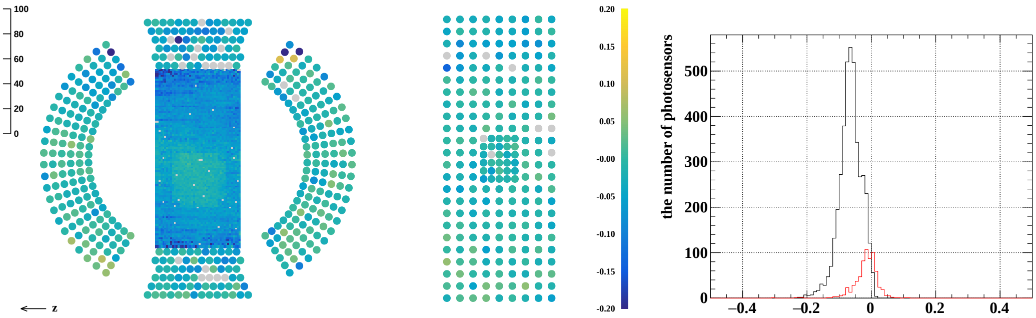


Figure 5.38 Relation between the number of background events and the signal efficiency ($E_\gamma = 51.5$–$54\,\mathrm{MeV}$) for the unfolding method (green/violet) and the deep learning model (red/black) when the events in which more than one γ-ray is found are rejected when MPPC PDE is 2%. The model is trained with dataset at MPPC PDE of 13%/2%.The different points for the deep learning model correspond to the different thresholds for the rejection. The optimal point (threshold = 0.40) is in blue/light blue.

Table 5.5 Number of backgrounds and signal efficiency for different MPPC PDEs ($E_\gamma = 51.5$–$54\,\mathrm{MeV}$) when the deep learning model is trained at a fixed PDE of 13% or each PDE.

PDE	Trained at PDE 13%		Trained at each PDE	
	N_{bg}	signal efficiency	N_{bg}	signal efficiency
13%	0.491	0.930	0.491	0.930
8%	0.468	0.950	0.486	0.956
4%	0.489	0.944	0.485	0.942
2%	0.584	0.962	0.576	0.962

3% at most, respectively. The increase of the background events was much smaller compared to the total amount of the reduction power, and thus the algorithms work even at the lower PDE. The possible error of the calibration turned out to be negligible in terms of the pileup elimination. Therefore, it is concluded that the pileup elimination algorithms can greatly suppress the background γ-rays with the high signal efficiency even under the realistic experimental condition.



(a)

(b)

Figure 5.39 Systematic uncertainty for calibration (a) in a 2D map and (b) in histograms for MPPCs (black) and PMTs (red).

Chapter 6

γ-ray Background Reduction with RDC

In the $\mu^+ \to e^+\gamma$ analysis of the MEG II experiment, the number of signal events is estimated with the maximum likelihood analysis, which is discussed in Sec. 2.4. The information from the RDC can be used to reject background-like events, but this decreases the signal efficiency due to the accidental coincidence of a signal γ-ray hit at the LXe detector and a positron hit at the RDC. Instead, we use the RDC information in the likelihood function by introducing the PDFs of the RDC observables. This allows the maximum use of the RDC information in terms of the sensitivity. In this chapter, the implementation of the RDC PDFs and the procedure to make them are discussed.

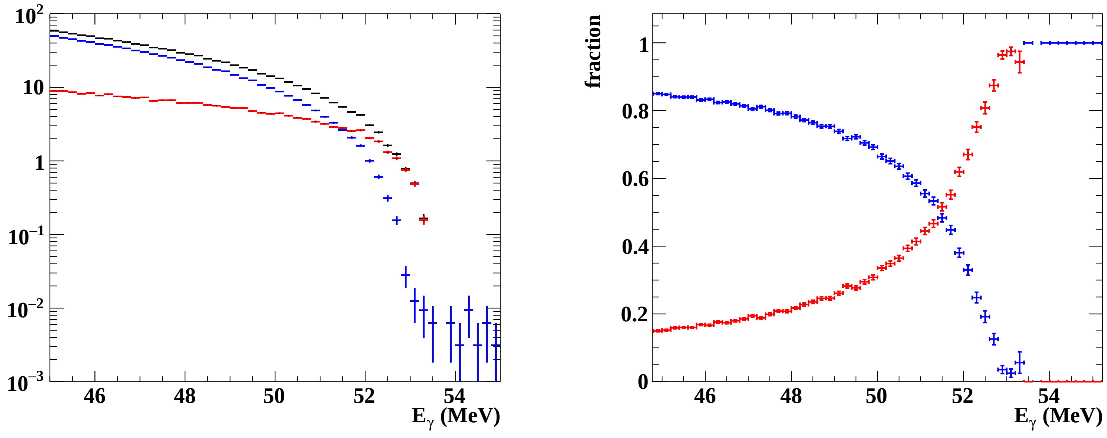
6.1 Concept

The likelihood function consists of PDFs of signal, RMD and accidental background. In the MEG experiment, the likelihood function was defined for the five observables $(E_\gamma, E_{e^+}, t_{e^+\gamma}, \phi_{e^+\gamma}, \theta_{e^+\gamma})$. Here, we extend it to include the three RDC observables $(t_{\mathrm{rdc}}^{\mathrm{us}}, t_{\mathrm{rdc}}^{\mathrm{ds}}, E_{\mathrm{rdc}}^{\mathrm{ds}})$ to improve the background discriminant power in the MEG II experiment. Note that the timing measured with the upstream RDC, $t_{\mathrm{rdc}}^{\mathrm{us}}$, is included in the following discussion for the future application though it is excluded for the sensitivity calculation in Chap. 7.

6.1.1 Correlations between observables

Dependence of relevant variables was investigated to decide how to generate the PDFs because correct treatment of them is important in order to perform the physics analysis accurately[*1]. From now on, the correlations among the parameters are described.

[*1] The effect of the RDC on the branching ratio sensitivity in [2] was overestimated due to the ignorance of E_γ dependence. The detection efficiency and the fraction of the RMD events were assumed to be constant in the calculation. In fact, however, they decrease as E_γ gets close to the signal energy as is discussed in this section, which results in a smaller sensitivity improvement.



(a) (b)

Figure 6.1 (a) Simulated E_γ spectra and fractions of total (black), RMD (blue) and AIF (red) events.

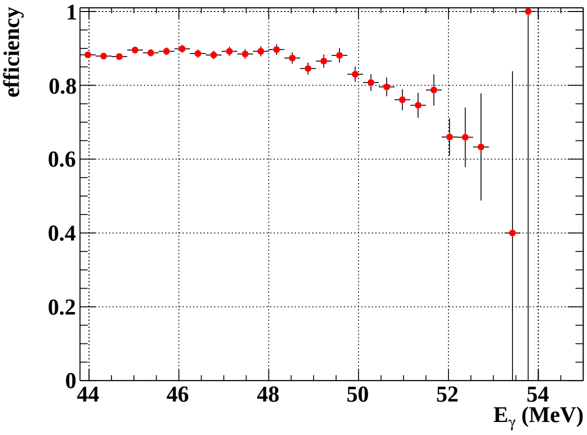


Figure 6.2 Simulated detection efficiency of the RMD positrons with the downstream RDC, which is defined by the number of events with a hit at the RDC over the number of events in which the RMD positrons are emitted toward downstream. The inefficiency mainly comes from the geometrical acceptance.

E_γ vs. $(t^{\mathrm{us}}_{\mathrm{rdc}}, t^{\mathrm{ds}}_{\mathrm{rdc}}, E^{\mathrm{ds}}_{\mathrm{rdc}})$

The main sources of accidental background γ-rays are RMD and AIF. Fig. 6.1(a) shows the E_γ spectra of each event type. The number of the RMD events decreases rapidly as E_γ gets close to the signal energy while that of the AIF decreases gently. As a result, the fraction of RMD events drops in the higher energy region as shown in Fig. 6.1(b), which indicates that the fraction of events detected by the RDC decreases accordingly. Moreover, the detection efficiency of RMD positrons is lower for high energy γ-rays since the energies of positrons accompanied by such γ-rays are low (Fig. 6.2). Therefore, all the RDC observables depend on E_γ.

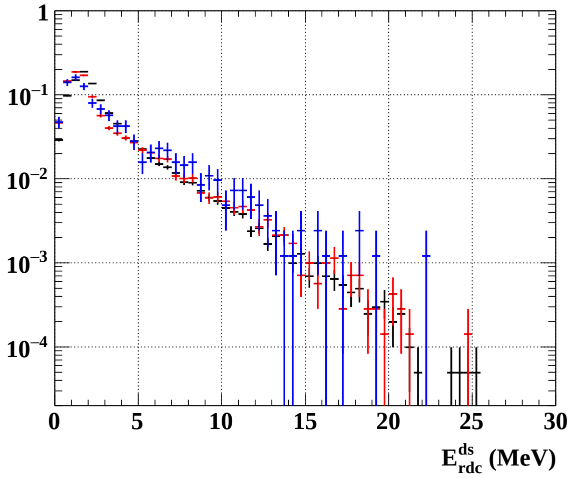


Figure 6.3 Simulated energy spectra of RMD positrons detected by the downstream RDC with E_γ of 48–49.75 MeV (black), 49.75–51.50 MeV (red) and 51.50–53.25 MeV (blue).

E_γ vs. $E^{\mathrm{ds}}_{\mathrm{rdc}}$

In addition to the probability to detect RMD positrons, the $E^{\mathrm{ds}}_{\mathrm{rdc}}$ distribution itself has a kinematical correlation with E_γ since the γ-ray detected by the LXe detector and the positron detected by the RDC originate from the same muon decay. However, the dependence is small enough to be negligible (Fig. 6.3).

$t^{\mathrm{us}}_{\mathrm{rdc}}$ vs. $(t^{\mathrm{ds}}_{\mathrm{rdc}}, E^{\mathrm{ds}}_{\mathrm{rdc}})$

The positrons from RMD events with high energy γ-rays are swept away either to the upstream or downstream of the muon stopping target. This means that when an RMD positron is detected by the downstream RDC, it must not be detected with the upstream RDC, and vice versa. Therefore, the probabilities of the upstream and downstream RDC to detect the RMD positrons correlate with each other.

$t^{\mathrm{ds}}_{\mathrm{rdc}}$ vs. $E^{\mathrm{ds}}_{\mathrm{rdc}}$

The RMD positrons fly along different paths depending on their energies. Fig. 6.4 shows the correlation between $t^{\mathrm{ds}}_{\mathrm{rdc}}$ and $E^{\mathrm{ds}}_{\mathrm{rdc}}$ of the RMD positrons under the condition of $E_\gamma > 48\,\mathrm{MeV}$. Not only the difference in the width of the first peak but also the existence of the second peak only for low energy positrons can be seen in the timing distribution. The second peak comes from the positrons which rebound from the upstream after hitting some materials.

6.1.2 Format

The PDFs should return the probability density to observe a given observable set, which is usually realized by a function with some parameters. As shown in the previous section, however, the distributions of the RDC observables have too complex forms to be parametrized. Therefore, we decided to use a non-parametric implementation with four dimensional histograms instead

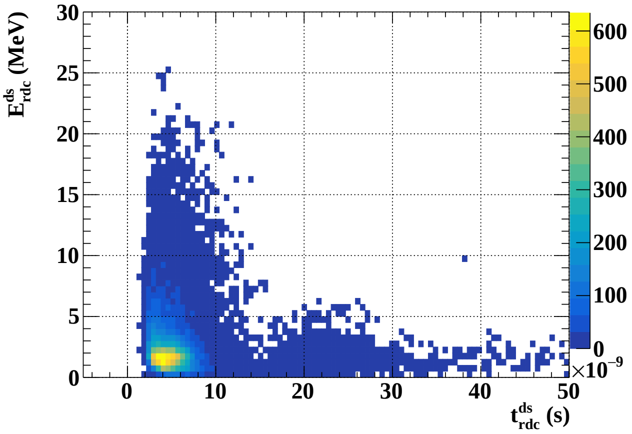


Figure 6.4 Correlation between the timing and energy of RMD positrons detected by the downstream RDC ($E_\gamma > 48\,\text{MeV}$).

of certain functions.

The RDC PDFs are defined for the analysis region of E_γ. The timing PDFs range from $-20\,\text{ns}$ to $28\,\text{ns}$. Their last bins contain entries where there is no hit observed in the time range. The $E^{\text{ds}}_{\text{rdc}}$ PDFs are defined in 0–50 MeV. The first bins corresponding to negative energy are filled with the no hit entries.

6.2 Procedure to make PDFs

Fig. 6.5 shows the projections of the signal and accidental background PDFs. Since the RDC detects the positrons from the RMD, which is a source of the accidental background γ-rays, there is a peak deriving from such positrons in the accidental background PDF while only accidental hits of Michel positrons are in the signal PDF. The procedures to make these PDFs are as follows.

6.2.1 Signal PDF

In the signal events, the positrons from signal muon decays follow trajectories with large radii such that they hit the positron spectrometer, and thus they cannot hit the RDC since its acceptance is small enough. Therefore, only accidental Michel positrons leave hits in the signal events.

The timing PDFs for the signal events are created assuming a certain hit rate, which is defined from the muon beam rate and the hit probability. The hit rate can be calculated directly from data by counting the number of events in the timing sideband as demonstrated in Sec. 4.2.2. The $E^{\text{ds}}_{\text{rdc}}$ distribution, i.e. the Michel positron energy distribution, is also obtained from data in the timing sideband.

Since there are no upstream-downstream correlation and no dependence on the γ-ray energy, the signal PDF can be simplified as

$$\begin{aligned} S_{\mathrm{RDC}}(t^{\mathrm{us}}_{\mathrm{rdc}}, t^{\mathrm{ds}}_{\mathrm{rdc}}, E^{\mathrm{ds}}_{\mathrm{rdc}}) = {} & S_{\mathrm{RDC}}(t^{\mathrm{us}}_{\mathrm{rdc}}) \\ & \times S_{\mathrm{RDC}}(t^{\mathrm{ds}}_{\mathrm{rdc}} | E^{\mathrm{ds}}_{\mathrm{rdc}}) \\ & \times S_{\mathrm{RDC}}(E^{\mathrm{ds}}_{\mathrm{rdc}}). \end{aligned} \tag{6.1}$$

Note that the timing-energy dependence of the downstream RDC remains since they are dependent on each other in terms of whether there is a hit or not.

6.2.2 RMD PDF

The positrons in the RMD events which are physics backgrounds do not hit the RDC as well, and so the RMD PDF can be created in the same way as the signal PDF.

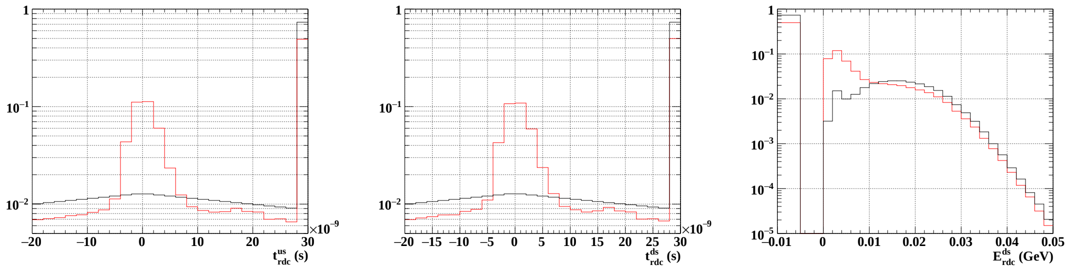
6.2.3 Accidental background PDF

As described in Sec. 6.1.2, there are correlations among the RDC observables and the energy of γ-ray in the accidental background events. They can be included in the accidental background RDC PDF as the conditional probability:

$$\begin{aligned} A_{\mathrm{RDC}}(t^{\mathrm{us}}_{\mathrm{rdc}}, t^{\mathrm{ds}}_{\mathrm{rdc}}, E^{\mathrm{ds}}_{\mathrm{rdc}} | E_\gamma) = {} & A_{\mathrm{RDC}}(t^{\mathrm{us}}_{\mathrm{rdc}} | E_\gamma, t^{\mathrm{ds}}_{\mathrm{rdc}}, E^{\mathrm{ds}}_{\mathrm{rdc}}) \\ & \times A_{\mathrm{RDC}}(t^{\mathrm{ds}}_{\mathrm{rdc}} | E_\gamma, E^{\mathrm{ds}}_{\mathrm{rdc}}) \\ & \times A_{\mathrm{RDC}}(E^{\mathrm{ds}}_{\mathrm{rdc}} | E_\gamma). \end{aligned} \tag{6.2}$$

The accidental background RDC PDF can be created from data just by filling the histogram with the observed parameter sets. It was, however, found to require a huge amount of statistics, and thus a dedicated approach is taken, which will be discussed in the next section.

For simulation studies using MC, the accidental background PDF is generated by mixing the simulated distributions of RMD and Michel positrons taking the dependence of the parameters described above into account. Fig. 6.6 depicts the process for the generation. Firstly, an event type of the γ-ray trigger is selected from RMD or AIF based on their trigger fractions. Then, E_γ is decided with the energy spectrum of the event type. If RMD is selected, the direction of the RMD positron and whether it is detected by either RDC are decided. If the positron is detected by the upstream RDC, the hit timing is chosen from the simulated timing distribution. In the case the positron is detected by the downstream RDC, the positron energy is chosen from the energy distribution in addition to the hit timing. Then, contribution from Michel positrons is simulated based on the hit rate and the energy is chosen from the energy distribution of Michel positrons. If AIF is selected, only accidental hits of Michel positrons should be considered. Finally, one hit whose timing is the closest to the timing center is chosen from a set of hits for the upstream and downstream RDC, respectively.



(a) (b) (c)

Figure 6.5 Projection of RDC PDFs for signal (black) and accidental background (red) to (a) $t^{\mathrm{us}}_{\mathrm{rdc}}$, (b) $t^{\mathrm{ds}}_{\mathrm{rdc}}$ and (c) $E^{\mathrm{ds}}_{\mathrm{rdc}}$.

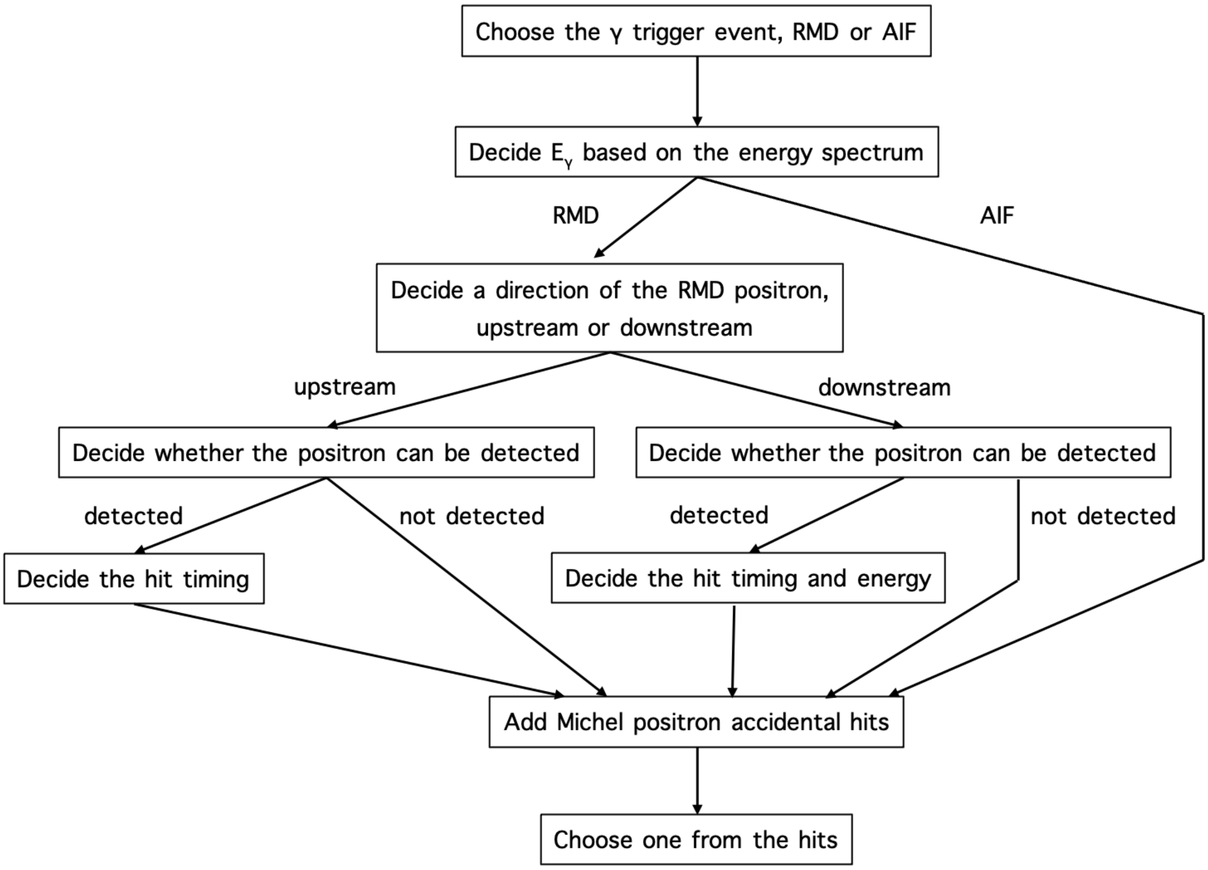
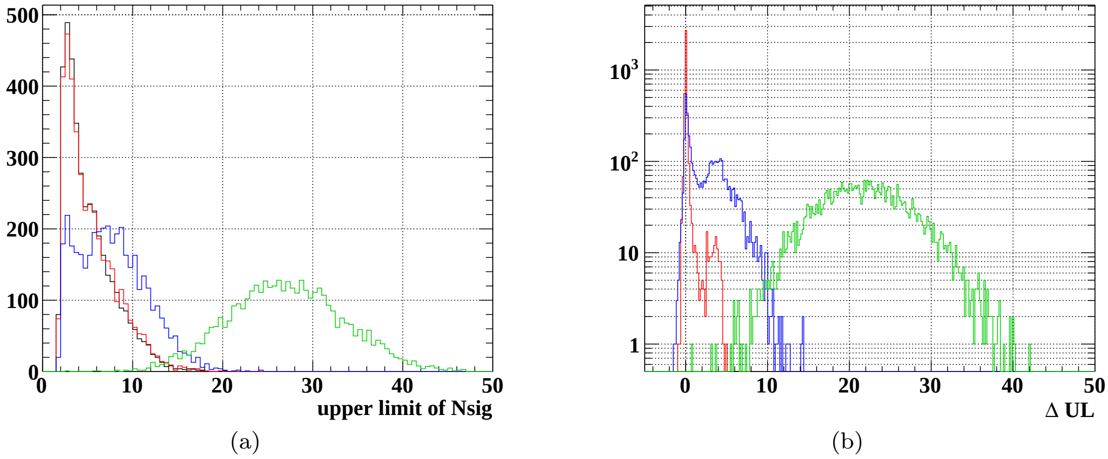


Figure 6.6 Process to create accidental background RDC PDF from MC simulation.

6.3 Statistical limitation and its solution

As mentioned in the previous section, the accidental background RDC PDF can be obtained directly from data in principle. This implementation requires sufficient entries in each bin for stable estimation of the number of signal events. In this section, the problem of statistical limitation and its solution are explained.



(a) (b)

Figure 6.7 (a) Upper limits on the number of signals for 4000 pseudo experiments and (b) deviation of upper limits from those calculated using the accidental background PDF generated with 1×10^9 events. Each color corresponds to the results using the PDF generated with 1×10^9 (black), 1×10^8 (red), 1×10^7 (blue) and 1×10^6 (green) events.

6.3.1 Statistical limitation

A lack of entries in each bin can cause an error of the analysis. Assume, for example, the number of entries in a certain bin of the accidental background PDF is accidentally much less than the expectation. If data is fit with such a PDF, the excess of the entries in the bin will be regarded as signal-like though it is not true. Fig. 6.7(a) shows the upper limits on the number of signals when accidental background PDFs generated with the different number of events are used to analyze the common pseudo experiments. The distribution is deformed when the statistics is small resulting from an increase of the experiments in which the upper limits are calculated to be higher as shown in Fig. 6.7(b). Therefore, it should be ensured that the statistical fluctuation is small enough not to affect the result.

The expected number of events which can be used for the accidental background PDF is calculated to be $\sim 1 \times 10^6$ events assuming a DAQ rate of 10 Hz, 120 DAQ days (1 year)[*2] and 1% event selection efficiency[*3]. Apparently, this cannot give sufficient entries with the PDF configuration used for the sensitivity calculation at the design stage such as bin widths of 0.5 MeV for E_γ, 2 ns for $t^{\mathrm{us}}_{\mathrm{rdc}}$ and $t^{\mathrm{ds}}_{\mathrm{rdc}}$, and 2 MeV for $E^{\mathrm{ds}}_{\mathrm{rdc}}$: the green case in Fig. 6.7(a).

Note that the statistics for the signal and RMD PDFs is not so severe since only one dimensional histogram of the Michel energy distribution is necessary for their generation, and so the following discussion focuses on the accidental background PDF.

[*2] The E_γ PDF for the accidental background events would be prepared year by year since it can be different due to the condition of the LXe detector.

[*3] A series of selections is applied to create the E_γ PDF for the accidental background events [44], and its efficiency was a few percent in the MEG.

6.3.2 Solution

The problem mainly comes from the sparsity of the four dimensional histogram. To solve this, two methods are developed; one is on the implementation of the histogram, and the other is on how to generate it.

Optimization of bin widths

The first idea is to optimize bin widths. Originally, the bin widths are set to be narrow enough not to affect the sensitivity. This is, however, not necessarily required for all bins since the importance of the bins is different depending on their regions; the timing coincidence region is more important than the timing sideband, and the low energy region of $E_{\mathrm{rdc}}^{\mathrm{ds}}$ is more important than the high energy region for the signal and background identification. Moreover, the sparsity is striking in such less important regions. Therefore, there is room to improve the setting of the bin widths.

Variable bin widths are adopted for the optimization to increase the entries of the sparse regions maintaining the information in the important regions. The result of the optimization is summarized in Table 6.1.

Table 6.1 Optimized bin edges of the RDC PDF. The bins for no hit events are not included.

Variables	Bin edges
E_γ	$48, 48.5, 49, 49.5, 50, 50.5, 51, 51.5, 52, 52.5, 53, 58\,\mathrm{MeV}$
$t_{\mathrm{rdc}}^{\mathrm{us}}$	$-20, -8, -4, 0, 4, 8, 28\,\mathrm{ns}$
$t_{\mathrm{rdc}}^{\mathrm{ds}}$	$-20, -8, -4, 0, 4, 8, 28\,\mathrm{ns}$
$E_{\mathrm{rdc}}^{\mathrm{ds}}$	$0, 2, 4, 6, 8, 10, 50\,\mathrm{MeV}$

Extraction of distributions in reduced dimension

In the simplest way, the four dimensional histogram is directly filled with obtained events, but this can waste statistics in the context of the RDC PDF; the E_γ dependences of the timing and energy distribution of the RMD positrons are not so important, and the energy distribution of the Michel positrons does not depend on E_γ and timing. Therefore, the extraction of those distributions in reduced dimension and reconstruction of the PDF with them can increase the effective statistics.

Fig. 6.8 summarizes how to extract the distributions. The original four dimensional PDF is projected to proper dimensions for each purpose. The RMD fraction detected by the upstream or downstream RDC is extracted from the projection to (E_γ, $t_{\mathrm{rdc}}^{\mathrm{us}}$) or ($E_\gamma$, $t_{\mathrm{rdc}}^{\mathrm{ds}}$). The number of RMD events can be calculated by subtracting the expected number of Michel background events from the total number of events, and the fraction is given by dividing it with the total number. This calculation is repeated for all the E_γ bins. The $t_{\mathrm{rdc}}^{\mathrm{us}}$ distribution is extracted from the projection to $t_{\mathrm{rdc}}^{\mathrm{us}}$ by subtracting the expected timing distribution of Michel positrons. The $t_{\mathrm{rdc}}^{\mathrm{ds}}$

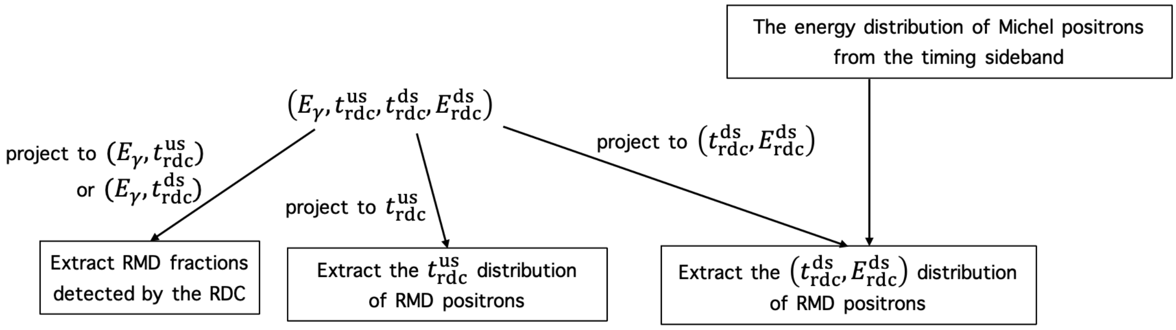


Figure 6.8 Overview of the process to extract distributions.

and $E_{\mathrm{rdc}}^{\mathrm{ds}}$ distributions are extracted from the two dimensional histogram of them since ignoring their correlation was found to result in sensitivity deterioration. The extraction can be achieved by subtracting the energy distribution of Michel positrons normalized with the expected number of Michel events in the timing region.

The extracted distributions are mixed by a similar method to the simulation-based generation; the RMD and Michel distribution are mixed with the RMD fractions.

Effect of improvements

The effect of the improvements was investigated using MC. Fig. 6.9 shows the median upper limits and the coverages as a function of the number of events used for the accidental background PDF. A common set of null signal pseudo-experiments was analyzed with accidental background PDFs generated with the different numbers of events. Thanks to the two methods, both of the median upper limit and the coverage are kept stable[*4] and no deterioration is observed with the statistics which can be achieved with one year data-taking, i.e. 1×10^6 events.

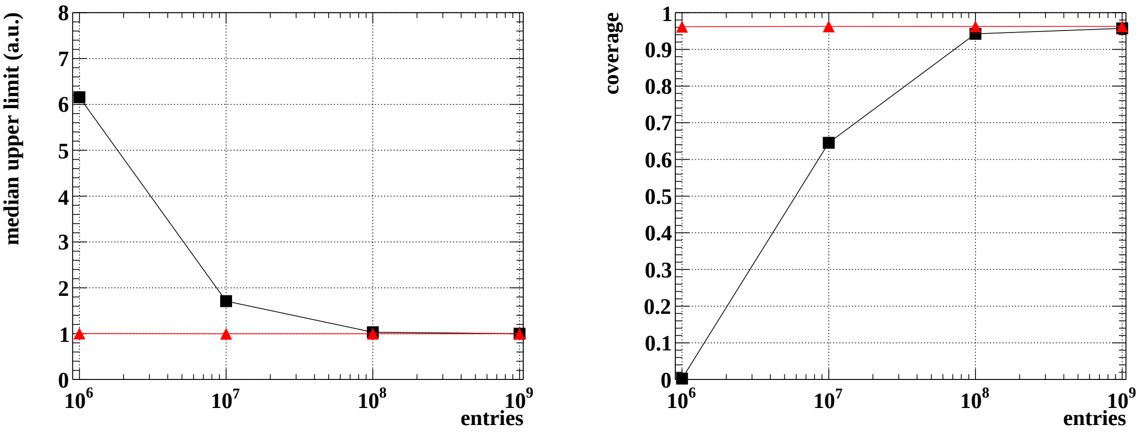
The test above was performed for the null signal experiments. To test that the analysis can correctly detect the signal events with the improved RDC PDF, the statistical power and significance were also studied under the assumptions of finite branching ratios. Common sets of pseudo-experiments under the assumption of different branching ratios were generated and analyzed with PDFs generated with 1×10^9 and 1×10^6 statistics. The p-value against the null hypothesis was calculated for each experiment. Fig. 6.10 shows the p-values at 95%-quantile as a function of the assumed branching ratio. The equivalent signal detection power is achieved with the improved RDC PDF generated with 1×10^6 statistics. The branching ratio down to 1.4×10^{-13} can be detected at the significance level of 1% with the power of 95%.

6.4 Uncertainty from PDF

The uncertainty from the RDC PDF can be divided into the three components due to the generation procedure: the RMD fraction, the RMD timing and energy distribution and the

---

[*4] The coverage tends to be 5% higher than the confidence level, which probably comes from the asymptotic method.



(a) (b)

Figure 6.9 (a) Median upper limits and (b) coverages as a function of the number of events used for the accidental background PDF without (black) and with (red) the improvements.

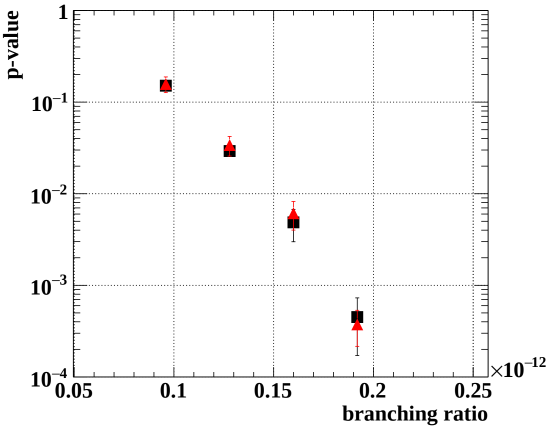


Figure 6.10 P-values at 95% quantile as a function of the assumed branching ratio without (black) and with (red) the improvements.

Michel energy distribution. The uncertainty from them can come from the statistics used to generate the PDF. Since the Michel energy distribution can be extracted from the off-timing region independent of the γ-ray detection, a sufficient amount of statistics can be accumulated for the generation. On the other hand, the others require a γ-ray hit at the LXe detector, and thus the available statistics is limited by the event selection. In particular, the uncertainty from the RMD fraction can be the largest as it needs to be extracted for each E_γ region.

The RMD fraction is affected by the statistical fluctuation of the number of events used for the calculation. The fluctuation is simulated by extracting the fraction from different datasets, and it increases as the energy becomes higher due to the lower statistics as shown in Fig. 6.11.

The RMD timing and energy distribution can fluctuate resulting from the statistics of the original distribution. Fig. 6.12 shows the uncertainty of each bin calculated from different PDFs which are generated using the different RMD timing and energy distributions. The fluctuation

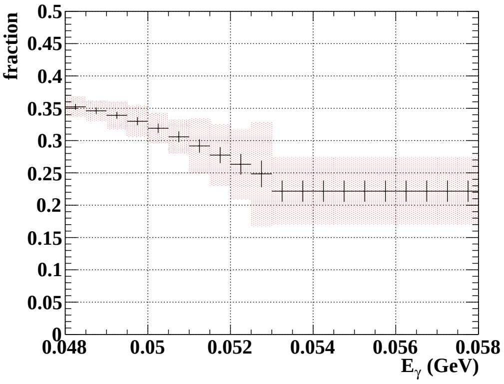


Figure 6.11 Uncertainty of the RMD fraction defined by the standard deviation (black bar) and peak-to-peak (red).

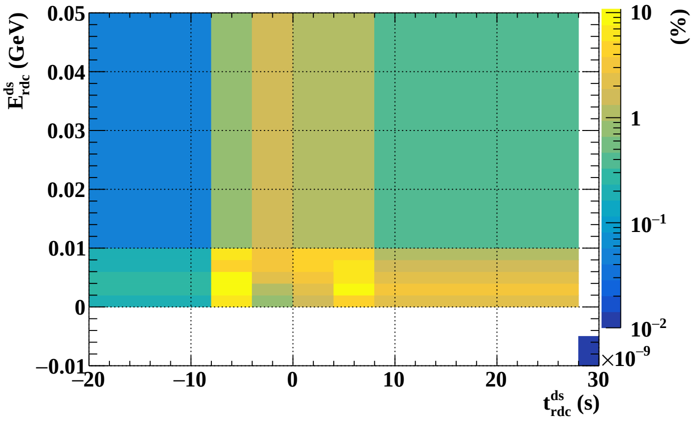


Figure 6.12 Uncertainty of the RMD timing and energy distribution.

is higher in the RMD-like region. Note that only the distribution of the downstream RDC is shown here, but it is obvious that the uncertainty from the timing distribution of the upstream RDC can be lower since it can be almost the same with the projection of the two-dimensional distribution.

Fig. 6.13 shows the uncertainty of the median upper limit calculated by analyzing with ten different RDC PDFs generated from different datasets. As the statistics for the RDC PDF generation gets smaller, the uncertainty becomes larger due to the statistical fluctuation. It is, however, limited to 1.2% even with the statistics of one year data-taking, i.e. 1×10^6 events, which is small enough.

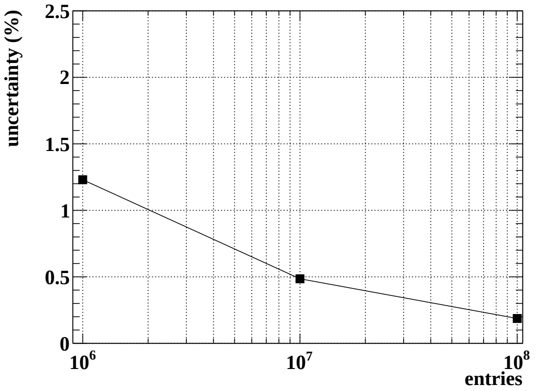


Figure 6.13 Uncertainty of the median upper limit from the RDC PDF as a function of the number of events used for the generation of the accidental background PDF.

Chapter 7

MEG II Projected Sensitivity

In this chapter, the branching ratio sensitivity is evaluated based on the measured performance of the MEG II detectors through the past pilot runs and the analysis update. As discussed in Sec. 2.5, the nominal value for the branching ratio sensitivity without any updates in this thesis was evaluated to be 7.8×10^{-14} for three years of data-taking. Hereafter, the effects of the updates in this thesis are described followed by the discussion on the projected sensitivity of the MEG II experiment.

7.1 Effect of γ-ray pileup elimination

Fig. 7.1 compares the branching ratio sensitivity after applying each algorithm for the pileup elimination. The event rejection based on the deep learning model improves the sensitivity by 18% resulting from the reduction of on-timing pileups. The peak search and the waveform unfolding results in further sensitivity improvement by 4%. In total, the sensitivity improvement from the nominal configuration is 22%.

In Fig. 7.1, the sensitivities at the halved intensity are also shown. To see the effect of the γ-ray pileups, the variables for positrons which change depending on the beam intensity (Table 2.15) were fixed to those at the nominal intensity. The pileup elimination enables to achieve 15% better sensitivity by doubling the beam intensity. It is also illustrated that the pileup elimination leads to a better sensitivity than that of the previous algorithm even with the halved statistics thanks to the elimination of the on-timing pileups.

Effect of PDE degradation

As discussed in Sec. 5.3.4, the performance of the pileup elimination depends on the MPPC PDE: the larger amount of backgrounds and the higher signal efficiency at the lower PDE. Fig. 7.2 shows the branching ratio sensitivity as a function of the PDE. There is no apparent tendency depending on the PDE, but the variation of the sensitivity fluctuates from -4% to $+7\%$, which depends on the balance between the increase of the backgrounds and the signal efficiency. The difference between the PDEs at which the models are trained is not significant as well. Therefore, we can use a single model independent of the actual PDE of data.

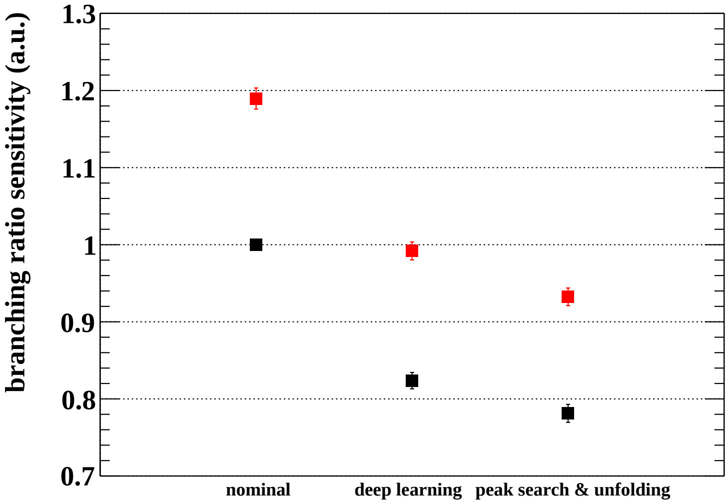


Figure 7.1 Relative size of the branching ratio sensitivity after applying the elimination algorithms at the beam intensity of 7 × 10^7 μ^+stops/s (black) and 3.5 × 10^7 μ^+stops/s (red). It is defined by the branching ratio sensitivity with the corresponding update over that under the nominal condition. Thus the better sensitivity is, the smaller value in the plot.

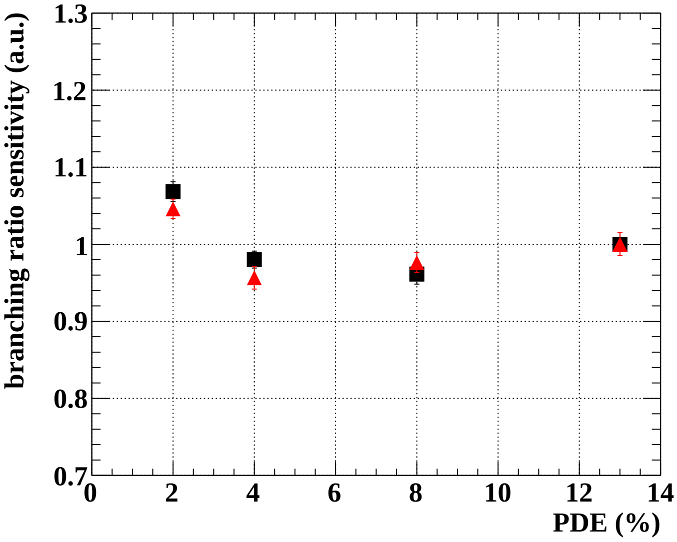


Figure 7.2 Relative size of the branching ratio sensitivity at the beam intensity of 7 × 10^7 μ^+stops/s as a function of MPPC PDE. The deep learning model is trained at a fixed PDE of 13% (black) and each PDE (red). It is defined by the branching ratio sensitivity with the corresponding update over that under the nominal condition. Thus the better sensitivity is, the smaller value in the plot.

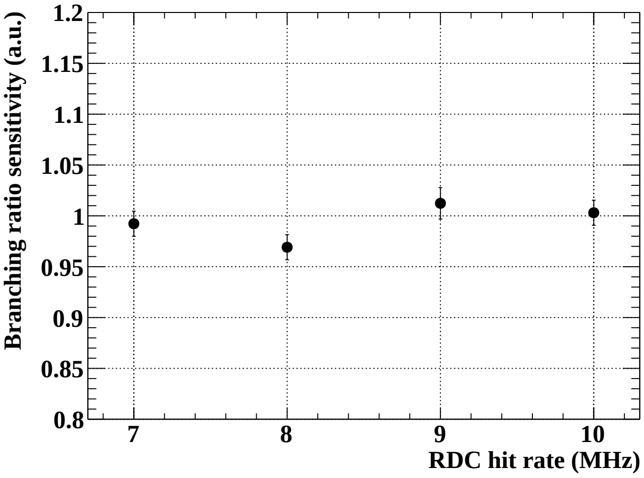


Figure 7.3 Relative size of the branching ratio sensitivity as a function of RDC background hit rate. The MC expectation is 7 MHz. It is defined by the branching ratio sensitivity with the corresponding update over that under the nominal condition. Thus the better sensitivity is, the smaller value in the plot.

7.2 Effect of RDC

The branching ratio sensitivity was evaluated by introducing the RDC observables to the analysis. The improvement due to the downstream RDC on the sensitivity was evaluated to be 8%.

Hit rate of Michel positron

The measurements carried out since 2017 suggests the hit rate of the Michel positron at the RDC is higher than the expectation by up to 40% (Sec. 4.2.2). The effect on the sensitivity was evaluated by assuming different hit rates and a constant beam intensity of 7×10^7 μ^+stops/s for γ-ray. Fig. 7.3 shows the sensitivity as a function of the background hit rate. The sensitivity does not change from the MC expectation of 7 MHz up to 42% excess (10 MHz). Therefore, the possible excess expected from the measurements has no effect on the sensitivity.

Saturation of SiPM for calorimeter

The saturation of the SiPM for the calorimeter was observed, and it deforms the energy distribution as discussed in Sec. 4.2.2. Fortunately, the most important part in the energy distribution is the low energy region around a few MeV, which acts as a decisive factor to discriminate the RMD positrons from the Michel positrons, and thus the deformation in the high energy region hardly affects the sensitivity. Indeed, the detailed shape ranging from 10 MeV to 50 MeV is decided to be ignored for the physics analysis as discussed in Sec. 6. The effect on the sensitivity was estimated using the measured energy distribution, and no degradation was observed.

7.3 MEG II projected sensitivity

The branching ratio sensitivity of the MEG II experiment was calculated including the effects from the pileup elimination algorithms and the RDC. As discussed in Sec. 3.1.1, the effect of the PDE degradation was taken into account assuming the optimistic and the pessimistic cases: the beam intensity of 7×10^7 μ^+stops/s with the PDE saturation at 2%, and 3.5×10^7 μ^+stops/s without the saturation.

Fig. 7.4 shows the branching ratio sensitivity as a function of calendar year assuming DAQ time of 20 weeks per year. The sensitivity with three years of data-taking will be 5.6×10^{-14} and 5.8×10^{-14} with the beam intensity of 7×10^7 μ^+stops/s and 3.5×10^7 μ^+stops/s, respectively. Thanks to the better performance of the positron reconstruction at the lower beam intensity, the sensitivity at the halved intensity is equivalent though the total statistics becomes half. The reduced beam intensity has an advantage in terms of the MPPC PDE degradation of the LXe detector as well; the PDE can be zero or be saturated around 2% after 70–100 DAQ day in the case of the 7×10^7 μ^+stops/s beam intensity. Other scenarios are also possible, and the beam intensity will be optimized to maximize the sensitivity considering the accumulated statistics and the detector performance.

The branching ratio sensitivity of the MEG II experiment is roughly one order of magnitude better than that of MEG. This is realized by the reduction of the background events and the increase of statistics. Table 7.1 summarizes the sensitivities and their factors, i.e. the normalizations and the median upper limits, for the MEG and the MEG II experiment[*1]. In the case of the beam intensity of 7×10^7 μ^+stops/s, the statistics is 5.1 times higher that of MEG, which comes from the doubled intensity and the improved detection efficiency. On the other hand, the N_{sig} upper limit is suppressed to 54% of that of MEG thanks to the improved detector resolutions, the pileup elimination and the background identification with the RDC. Even in the case of the halved intensity, which is similar to that of MEG, the statistics is 3.0 times higher while the N_{sig} upper limit is suppressed to one third of that of MEG.

Thanks to the suppressed background level, the further sensitivity improvement is possible by extending DAQ time; for example, it reaches 4×10^{-14} for five years of data-taking[*2]. This also suggests that even lower beam intensity enables the achievement of better sensitivity if DAQ time is extended by extra years.

[*1] The nominal values of k in Table 2.15 are reduced taking into account the inefficiencies from the pileup elimination, 6.4% and 4.6% for $48\,\mathrm{MeV} < E_\gamma < 58\,\mathrm{MeV}$, for the two cases in MEG II, respectively.

[*2] The sensitivity with the halved intensity gets equivalent to the other after five years of data-taking as shown in Fig. 7.4. This is because the number of backgrounds for the halved intensity is suppressed enough not to saturate the sensitivity yet while the other has started to be saturated due to the higher background level.

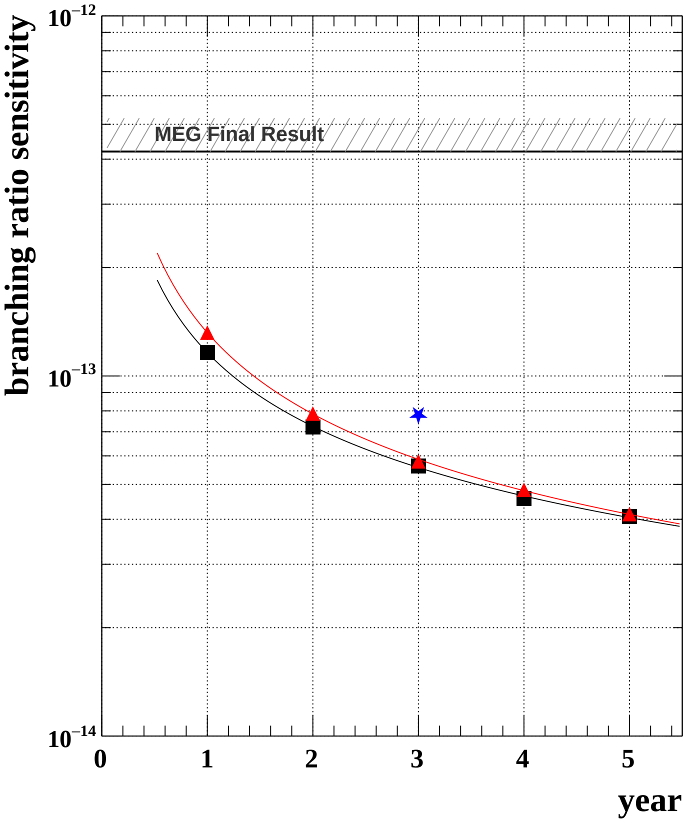


Figure 7.4 Branching ratio sensitivity as a function of calendar year assuming DAQ time 20 weeks per year at the intensity of 7×10^7 μ^+stops/s (black) and 3.5×10^7 μ^+stops/s (red). The blue point shows that without including any updates in this thesis (Sec. 2.5).

Table 7.1 Sensitivity and its factors for the MEG and MEG II experiment.

Variables	MEG (2009–2013)	MEG II (for three years)	
	3×10^7 μ^+stops/s	7×10^7 μ^+stops/s	3.5×10^7 μ^+stops/s
k	1.71×10^{13}	8.78×10^{13}	5.09×10^{13}
N_{sig} upper limit (90% C.L.)	9.1	4.9	3.0
sensitivity	5.3×10^{-13}	5.6×10^{-14}	5.8×10^{-14}

Chapter 8

Conclusions

The MEG Ⅱ experiment is planned to search for the $\mu \to \mathrm{e}\gamma$ decay with the higher sensitivity than the previous MEG experiment by one order of magnitude. Since the sensitivity of the MEG experiment was limited by the number of accidental backgrounds, a reduction of such events is mandatory to realize a deeper search for the decay in MEG Ⅱ. This can be achieved by improving detector resolutions, reducing pileup events and identifying the backgrounds. This thesis focuses on the suppression of the γ-ray backgrounds especially from the latter two aspects.

The sources of the γ-ray backgrounds are the RMD and AIF of a positron from the Michel decay. The γ-ray from the RMD can be tagged by detecting the positron emitted from the same muon decay. There are two types of AIF γ-ray events in which one or two γ-rays from the AIF enter the γ-ray detector. In the latter case, the events can be rejected by finding the two coincident hits. The higher beam intensity significantly increases accidental pileup γ-ray hits, which increases the number of events in the signal energy region. The contribution from such γ-rays can be reduced by reconstructing the energy of each hit.

The liquid xenon γ-ray detector has been upgraded by replacing the PMTs with the SiPMs on the γ-ray entrance face to realize a high granularity and a good uniformity of the scintillation readout. The performance of the detector was evaluated from the data obtained in the pilot runs. Improvements of the position and energy resolutions for the shallow events were confirmed while the energy resolution in the deep events ($w > 2\,\mathrm{cm}$) was worse than expected, which is probably due to the unknown contribution also observed in MEG. The timing resolution was estimated with two different methods. One of them was found to be consistent with the MC expectation while the other to be much worse than the expected value. This can come from a systematic error of the measurement, and thus a dedicated data-taking was proposed to reduce it, which is planned to be performed in December 2021.

A series of algorithms to eliminate the pileup γ-rays was developed which utilizes the high granular readout channels and waveforms digitized at a high sampling frequency. The pileup γ-rays are searched for spatially and temporally using the information of each channel. For the spatial search on the inner face, the prediction by a deep learning model is used to increase the detection efficiency of the pileup γ-rays and suppress misidentification of fake peaks. This spatial search is sensitive to the AIF background with two incident γ-rays, which have a large contribution near the signal energy. The energies of the found γ-rays are reconstructed separately by unfolding waveforms, which allows us to use the pileup events resulting in the high signal efficiency of 93% while the number of the backgrounds around the signal energy is reduced by 51% compared to the previous study.

Although the light distribution was used as the model input in this thesis, the timing distribution would be also useful to identify off-timing pileups as used for the timing-based search (Sec. 5.2.2). The combination of the light and timing distributions is also possible, which can improve the performance especially for the pileups whose timings are slightly different from the main γ-ray though they are in the charge integration region. In order to correctly train the model and benefit from the timing information, the model architecture, data pre-processing and training procedure must be investigated.

The RDC was newly introduced in the MEG II experiment to tag the RMD γ-ray backgrounds. The expected performance of the detector was confirmed using the data in the pilot runs. The RDC identifies 27% of the γ-ray backgrounds in the analysis region. The accidental hit rate of the Michel positron exceeded the expectation by up to 40%, but it was found not to affect the branching ratio sensitivity. In addition, a procedure to introduce the observables of the detector to the $\mu^+ \to e^+\gamma$ analysis was developed, and the correlations among them and the γ-ray energy are correctly taken into account in the analysis.

The measured detector performance and the improvement in the analysis lead to the projected sensitivity of the MEG II experiment of 5.6–5.8 $\times 10^{-14}$ for the branching ratio of $\mu^+ \to e^+\gamma$ with three years of data-taking, which is 10 times higher than that of the MEG experiment. This was achieved by the reduction of the γ-ray backgrounds thanks to the development of the pileup elimination algorithms and the introduction of the RDC, which provide a 22–26% and 8% improvement, respectively. The unprecedented sensitivity of the $\mu^+ \to e^+\gamma$ search of the MEG II experiment would enable us to test many of the well-motivated BSM theories.

The engineering run with the full detector was successfully completed and a part of physics data was taken in the run 2021. A full-blown accumulation of the physics data to search for the $\mu^+ \to e^+\gamma$ decay will start in 2022.

Appendix A

Lab Test for Downstream RDC

The downstream RDC had been developed toward the installation to the MEG II experiment. In this chapter, the measurements at the lab during the construction are described.

A.1 Timing counter

The timing resolution of the timing counter was measured with a ^{90}Sr source. Fig. A.1 shows the setup for the measurement. A reference counter whose timing resolution was 30 ps was placed under the test counters to trigger events. The reference counter was made of a plastic scintillator with a size of $5 \times 5 \times 4\,\mathrm{mm}^3$ and its scintillation light was read out with a SiPM. The waveforms of the signals were recorded with a DRS4 evaluation board. The timings of the signals were extracted using the constant fraction method. The timing resolutions were calculated by

$$\sigma_{\mathrm{counter}} = \sigma\left(\frac{t_{\mathrm{left}} + t_{\mathrm{right}}}{2} - t_{\mathrm{ref}}\right), \tag{A.1}$$

where t_{left} and t_{right} are the extracted timings of the left and the right channels of the test counter, respectively, and t_{ref} is the timing of the reference counter. The ^{90}Sr source was placed at several positions along the direction of the length and timing resolutions were measured for each. Fig. A.2 shows results of the measurements. All timing resolutions are better than 90 ps, which are good enough for the RMD detection.

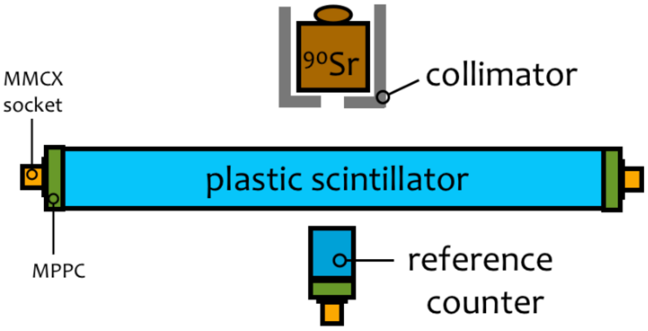


Figure A.1 Schematic view of the setup for the timing counter [65].

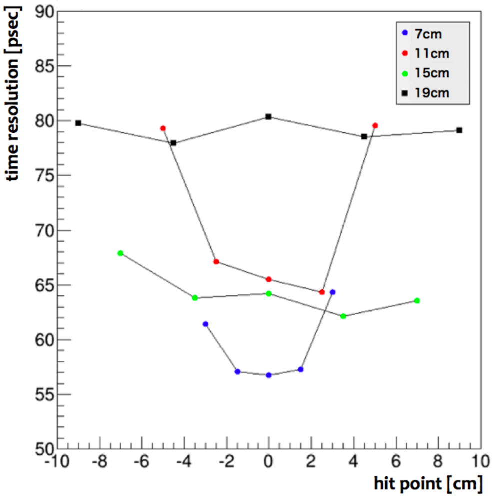


Figure A.2 Measured timing resolutions [65]. The x-axis corresponds to the position of the radiative source along the direction of the length where the origin is the center of scintillators.

A.2 Calorimeter

The LYSO crystals were tested and the energy resolution was evaluated. Fig. A.3 shows the setup for the measurement. The energy of γ-rays from a ^{60}Co source was measured with each LYSO crystal attached to the SiPM using spring pins. The waveform was recorded with a DRS4 evaluation board. The charge was calculated by integrating the waveform. Fig. A.4 shows an example of the observed energy spectrum. Since the energy tail of the intrinsic radioactivity overlaps the energy peaks at 1.17 MeV and 1.33 MeV of the ^{60}Co source, double Gaussian functions with an exponential components were used as a fitting function. The energy resolutions at each peak energy are defined by σ/E_γ, where E_γ and σ are the mean of sigma of the corresponding Gaussian. Fig. A.5 shows the result for 76 crystals. The energy resolutions of most crystals were measured to be ~6% at 1 MeV.

In a high rate environment, energy resolution can be worsened by the afterglow effect of LYSO. Afterglow is a delayed light emission of crystals with very long time constant (typically few hours). When electrons are excited, some of them are trapped in lattice defects such as oxygen vacancies. Then, these trapped electrons induce scintillation light. Since the high hit rate of positrons up to 600 kHz at the central crystal is expected, the effect was studied by exposing the crystals to a ^{90}Sr source. The increase of the current due to afterglow was measured with the SiPMs attached to the crystals. The increase of the current in the MEG II beam environment was estimated to be $\sim 10\,\mu$A at maximum, which is small enough; the deterioration of the energy resolution is less than 1% at 1 MeV.

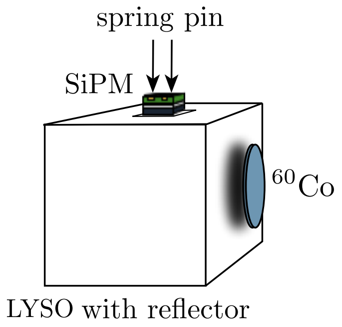


Figure A.3 Schematic view of the setup to measure the energy resolution [69].

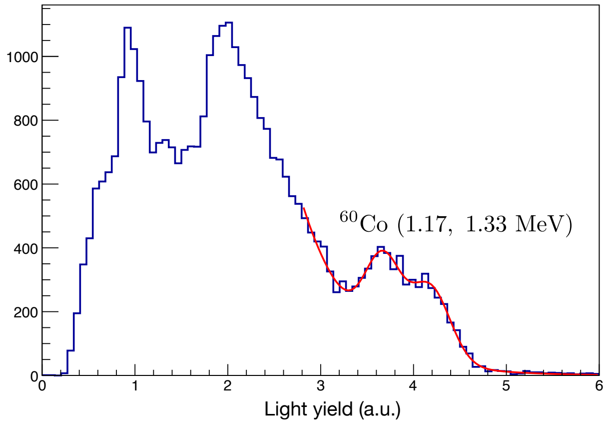


Figure A.4 Observed energy spectrum including ^{60}Co and the intrinsic radioactivity. The two peaks at 1.17 MeV and 1.33 MeV and the background floor are fitted with double gaussian functions and an exponential function [69].

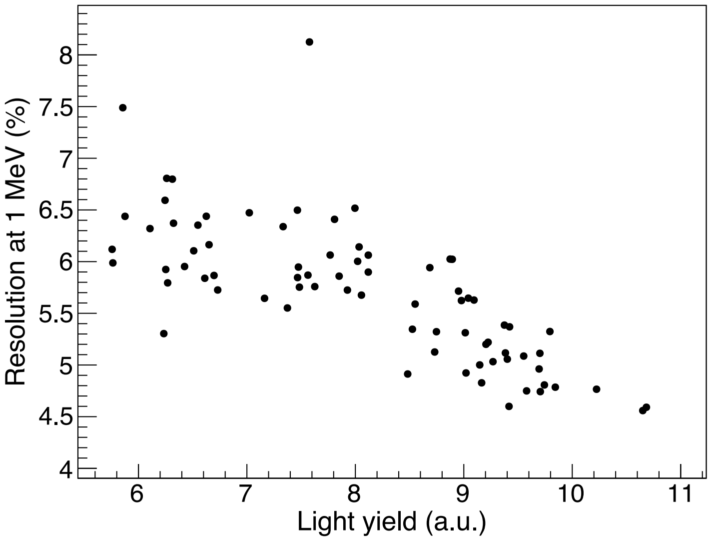


Figure A.5 Measured energy resolutions of the LYSO crystals as a function of the obtained light yields [69].

Appendix B

Resistive Plate Chamber

A resistive plate chamber (RPC) has been investigated as a candidate of the upstream RDC. In this chapter, an overview of the RPC as the upstream RDC is given.

Fig. B.1 shows a schematic view of a typical RPC. The RPC consists of parallel plates, positively-charged anodes and negatively-charged cathodes. They are made of a very high resistivity material. The gap between the plates is filled with a gas and electrons are knocked out of gas atoms when charged particles pass through the chamber. These electrons are accelerated with the electric field and hit other atoms causing an avalanche of electrons. The gained signal is read out with metallic strips.

The high resistive plates realize a stable operation to suppress the electrical discharge. They are usually made of glass or bakelite. For the upstream RDC, however, the plates made of such materials can be so thick that they cannot avoid affecting the beam. Therefore, a 50 μm thick polyimid film coated with diamond-like carbon (DLC) is adopted for the resistive plates. DLC is made of carbon having a combination of sp^2 and sp^3 bonds. As a result, a DLC coating has a combination of the properties of both diamond (sp^3 bonding) and graphite (sp^2 bonding) such as electric insulation and conductivity. By changing a fraction of the two bonds, the resistivity of DLC can be controlled. Since the high resistivity can be realized with the DLC coat thinner than 100 μm, the material budget of DLC is negligible.

A R134a based gas mixture with iso-C_4H_{10} and SF_6 is often used for the RPC. Thanks to the high electronegativity of R134a and SF_6, a multiplication of electrons can be suppressed resulting in a stable operation while iso-C_4H_{10} plays a role of a quencher preventing discharges by absorbing UV light.

The detection efficiency of the RPC can get higher as the gap becomes thicker while the timing resolution can be worsen due to a longer drift distance. The multiple-layer structure can realize a high detection efficiency keeping the sufficient timing resolution. The detection efficiency of the n-layer RPC can be written by that of the single layer $\epsilon_{\mathrm{single}}$ as follows [97]:

$$\epsilon_n = 1 - (1 - \epsilon_{\mathrm{single}})^n. \tag{B.1}$$

Therefore, the high detection efficiency and the good timing resolution can be achieved by stacking thin layers.

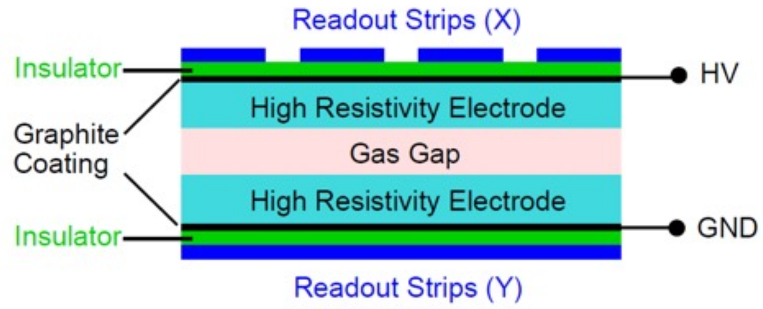


Figure B.1 Schematic view of the RPC [99].

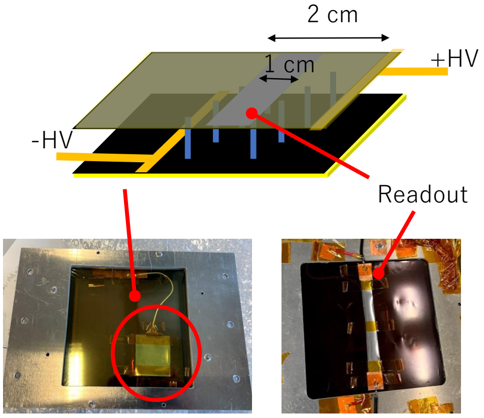


Figure B.2 Overview of the prototype RPC [98].

A single-layer prototype RPC was tested with a μ beam at the πE5 beamline [98]. Fig. B.2 shows the overview of the prototype. It sizes $2 \times 2\,\mathrm{cm}^2$ with a gap of $384\,\mu$m thickness. A gas mixture of 94% Freon (R134a), 5% iso-C4H10 and 1% SF6 was filled in the gap. The high voltage was supplied via conductive tapes attached on the edge of the DLC surface, and the signal was read out from both ends of the strips. The performance of the RPC was measured with positrons and muons at a rate up to $1\,\mathrm{MHz/cm}^2$. The detection efficiency of 60% was achieved, which implies the aimed efficiency of 90% is achievable by stacking four layers. The timing resolution, 170 ps, was much better than the requirement of 1 ns. A voltage drop due to the high hit rate was observed, which was consistent with the expectation from the design. The results tell us the use of the RPC as the upstream RDC is possible though there still remain some technical challenges such as a suppression of a voltage drop and a difficulty to control the resistivity.

Appendix C

Study on Radiation Damage to VUV-MPPC

As discussed in Sec. 3.1.1, a degradation of MPPC PDE was observed under the muon beam. Since the degradation is correlated with the beam exposure and its magnitude is larger for VUV light than visible light, it is likely because of a surface damage by the particle irradiation. This chapter begins with a summary of radiation damage to the SiPM followed by explanations of lab test results.

C.1 Radiation damage to SiPM

There are two types of radiation damage known for the SiPM: bulk damage due to Non Ionizing Energy Loss (NIEL), and surface damage due to Ionizing Energy Loss (IEL) [100].

The bulk damage is primarily caused by high energy particles such as protons, pions, electrons and photons, and by neutrons. They can displace atoms out of their lattice site and generate crystal defects leading to changes in the sensor performance related to the newly introduced energy levels in the energy gap between the valence and the conduction band, $E_{\text{gap}} \sim 1.12$. A typical phenomenon caused by the introduced energy level is an increase of the leakage current; defects with energy levels close to the middle of the band gap ($\sim 0.56\,\text{eV}$) facilitate the thermal excitation of electrons and holes, increasing the dark current. This is reported to occur above $1 \times 10^{8}\,/\text{cm}^{2}$ ($1\,\text{MeV}$ neutron equivalent) by many studies [86] [87] [88]. Some studies also report a decrease of the signal size possibly caused by an increase of pixel occupancy due to the increased dark signals [88]; if a photon hits a pixel which is not fully recovered from the previous discharge, the effective gain can be smaller.

The surface damage is primarily produced by photons and charged particles, generating charges in the oxide (SiO_2) and at the Si–SiO_2 interface, and interface traps at the Si–SiO_2 interface. The energy deposit above the energy gap generates electron-hole pairs in the high resistivity passivation layer. Most of the electrons leave the passivation layer due to a relatively high mobility and low trapping probability. On the other hand, some of the holes, which move via polaron hopping, are captured by deep traps in the passivation layer or the interface traps between silicon layers resulting in an accumulation of the stationary charges. A large increase of the dark current is reported above 200 Gy due to an increase of surface current [89]. If a fraction of the surface current reaches the amplification region, it also increases the dark-count rate.

Fig. C.1 shows the radiation environment of the LXe detector. When the muon beam is used, γ-rays mainly from the RMDs hit the detector through the MPPCs on the entrance face. The xenon scintillation light (7.1 (mean) ± 0.1 (FWHM) eV) generated from these γ-rays hit the MPPCs. Neutrons coming from the accelerator also hit the detector.

The expected irradiation dose in the run 2019 is summarized in Table 3.2. The neutron fluence was estimated to be less than $2.9 \times 10^6\,/\mathrm{cm}^2$ (1 MeV neutron equivalent) from the measurement at the MEG II experimental area [43], which is much smaller than the level where the NIEL damage is reported. The γ-ray dose was simulated to be 0.01 Gy [43], which is also small enough. The VUV fluence was measured to be $4.6 \times 10^{10}\,/\mathrm{mm}^2$ from the induced current on the MPPCs in the run 2019 while the estimation from the simulation was $5.8 \times 10^{10}\,/\mathrm{mm}^2$. A radiation damage to SiPMs by the VUV irradiation was not reported. Thus, a series of irradiation tests was performed at the lab to understand the observed PDE degradation.

C.2 VUV photon irradiation

VUV-MPPCs were irradiated with the intense VUV light from a xenon lamp to study the effect of the VUV light. The irradiation was performed both at room temperature and at low temperature since the accumulation of the stationary charges can be enhanced at lower temperature due to lower mobility.

C.2.1 Irradiation at room temperature

Fig. C.2(a) shows a schematic of the setup for the VUV photon irradiation at room temperature. A VUV-MPPC was irradiated with VUV photons and its response to VUV light was

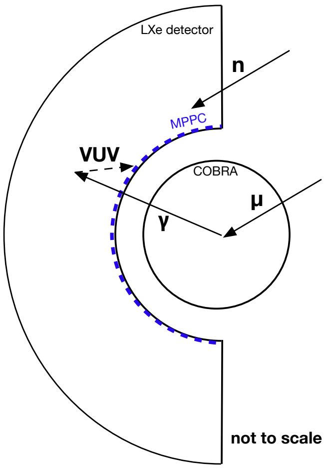


Figure C.1 Radiation environment of the LXe detector [62].

monitored together with that of another VUV-MPPC without irradiation. All of the irradiation and measurements was performed at 25°C.

A VUV-MPPC was irradiated by a xenon lamp (Hamamatsu L9456-03 [101]) placed at a distance of 5 cm, and the light emitted from the lamp directly entered the irradiation sample without any filters. The spectrum of the emitted light from the lamp is shown in Fig. C.2(b).

The VUV component of the irradiation light was measured to estimate irradiation flux beforehand. Two bandpass filters (ACTON FB180-B-.5D and Edmund optics #33-026, Fig. C.3) were used to select VUV light peaked at 190 nm. The amount of light was measured placing the lamp at different distances ranging from 15 cm to 53 cm, and the results were extrapolated to that at the measurement distance to avoid saturation the MPPC. As a result, the irradiation flux of the VUV light was estimated to be 5×10^{13} photons/mm^2/hour.

The response to VUV light was monitored by periodically measuring the charge with another xenon lamp (Hamamatsu L9455-13 [101]) placed at 35 cm away from the MPPC. The two bandpass filters were used to monitor the response to the VUV light peaked at 190 nm in addition to one ND filter to avoid saturation. The response of another MPPC without irradiation was also measured for the comparison. In order to ignore the position dependence of the light distribution from the lamp, the irradiation sample was displaced and the reference sample was put on the same position during the charge measurement. The uncertainty of the measured charge was estimated to be ~5% by repeating the measurement including this displacement, which is enough small.

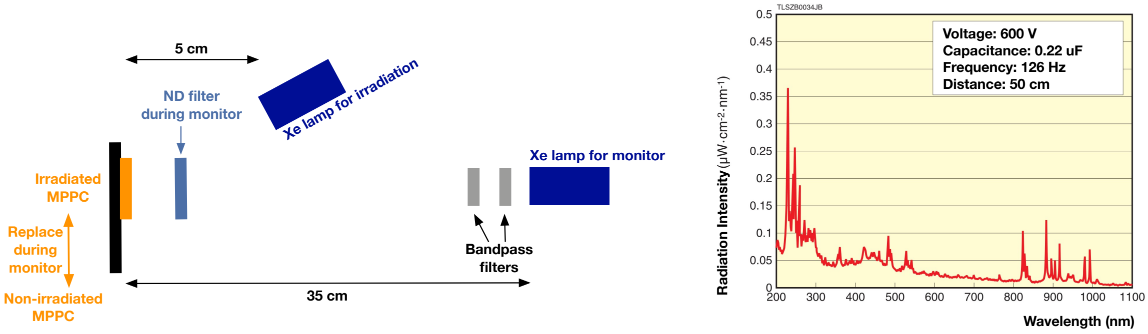
Fig. C.4(a) shows a history of the VUV response for the irradiated and the non-irradiated MPPCs. The VUV response of the irradiated MPPC was degraded by 65% in total after 3×10^{16} /mm^2 VUV photon irradiation while that of the non-irradiated sample was stable. The degradation speed was decreasing and saturated when it reaches 35% of the original value. The observed degradation is much slower by a factor of $\mathcal{O}(10^4)$ than that in the LXe detector: 9% degradation at the expected fluence of 4.6×10^{10} /mm^2 VUV photon.

The irradiated MPPC was annealed after the irradiation. It was illuminated with room light being applied bias voltage, and was kept at 70°C for 24 hours. The annealing increased the charge for VUV light from 35% to 105% of that before the irradiation, and thus the degradation was recovered completely. The recovered MPPC was irradiated again, and the degradation was observed at slightly slower speed (Fig. C.4(b)).

C.2.2 Irradiation at low temperature

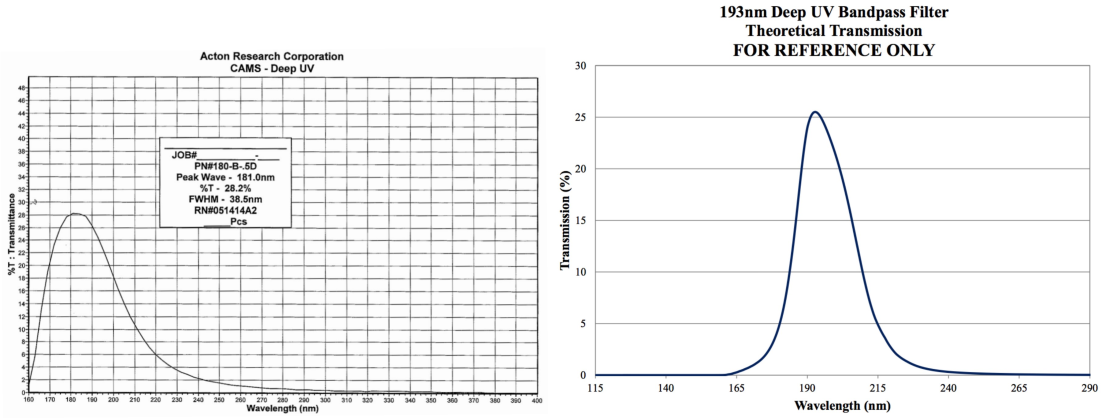
The PDE degradation was observed, but its speed was found to be much slower than the observation of the LXe detector. This might be because of the difference of temperature; more charges can be accumulated at lower temperature, which results in faster degradation speed. Therefore, the irradiation at LXe temperature was also tried [102].

Fig. C.5(b) shows the setup for the irradiation. It was covered with a dewar for thermal insulation, and purged with gaseous nitrogen to avoid condensation. The MPPC was cooled



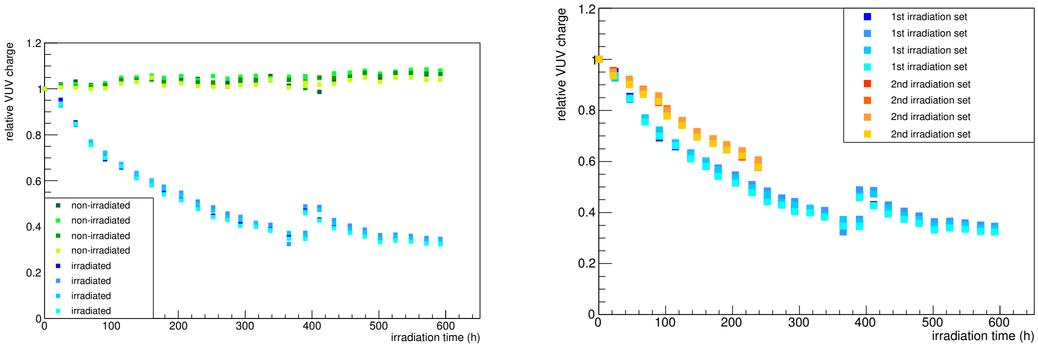
(a) (b)

Figure C.2 (a) Setup of VUV photon irradiation and (b) spectral radiation intensity of the Xe lamp [62]. This lamp can emit VUV light down to $\lambda = 185\,\mathrm{nm}$.



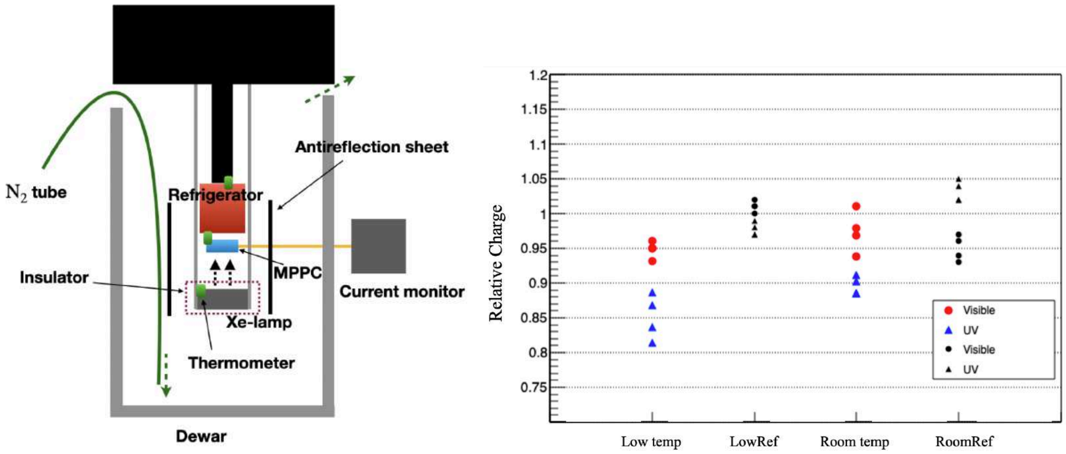
(a) (b)

Figure C.3 Spectrum of the bandpass filters.



(a) (b)

Figure C.4 (a) Response to VUV light from a fixed intensity xenon lamp. Different data series with similar colors show four chips on the same MPPC package [62]. The increase around 390 hours of irradiated samples is due to the annealing at room temperature caused by 12 day intermission. (b) VUV response in the second irradiation after one cycle of annealing [62].



(a) (b)

Figure C.5 (a) Setup for the VUV photon irradiation at low temperature and (b) measured charges after 24 hour irradiation normalized by those before irradiation [102]. The results of four chips on a MPPC were shown.

down to LXe temperature with a pulse tube refrigerator (Aisin TAC 101J), and was irradiated with the xenon lamp placed at the distance of 3 cm without any filters. The temperature around the MPPC was monitored with a thermometer and kept around 170 K during the irradiation.

The irradiation flux was estimated by measuring the charge with the xenon lamp used for the irradiation. During the measurement, the two bandpass filters and a collimator with a radius of 1 mm was used, and the lamp was placed at the distance of 42 cm to reduce the amount of photons. As a result, the irradiation flux of the VUV light was estimated to be 5.4×10^{13} photons/mm^2/hour. The response to VUV light was also measured with the filters and the collimator.

Fig. C.5(b) shows the relative charges of the irradiated and the non-irradiated samples after 24 hour irradiation corresponding to the dose level of 1.3×10^{15} photon/mm^2 in total. The results of the irradiation at room temperature are also shown, which was performed with the same setup. The charges decreased by 10–20% at low temperature while the decrease was about 10% at room temperature. The charges of the non-irradiated sample also changed up to 7%, which implies the uncertainty of the measurement.

Although larger degradation at low temperature was observed, it was still too slow compared to the observation in the LXe detector by a factor of $\mathcal{O}(10^4)$. Therefore, it cannot be concluded that the PDE degradation in the LXe results from the VUV light irradiation.

C.2.3 Dependence on wavelength

If the PDE degradation is caused by the surface damage, there should be wavelength dependence of the magnitude of the degradation; light with a shorter wavelength should be easier to be affected since it tends to be absorbed near the surface. Indeed, the response for visible light was not degraded according to the observation in the LXe detector, which supports the surface damage hypothesis. To make the effect of the VUV light irradiation clear, the dependence of the PDE degradation on the wavelength was also investigated at the lab.

Fig. C.6 shows the setup for the measurement. In this measurement, there are several improvements to gain the amount of VUV light compared to the measurements above. Two xenon lamps with $\mathrm{MgF_2}$ windows (Hamamatsu L13651-01-3 and L12745-01-3 [101]) were used instead of those used for the previously mentioned lamps with UV glass windows (Fig. C.7). Bandpass filters which select light with shorter wavelengths ($\lambda \sim 172\,\mathrm{nm}$) were also used (Fig. C.8). The whole setup was placed in a thermal chamber, and it was purged with gaseous nitrogen to avoid the absorption of light with short wavelengths. In order to study the wavelength dependence, five LEDs with different wavelengths raging from 280 nm to 645 nm were also installed.

The irradiation flux was estimated in the similar way; the charge from the 2 W xenon lamp placed at 45 cm away from a VUV-MPPC was measured using a ND filter and a collimator with a radius of 3 mm, and it was scaled. As a result, the irradiation flux of the VUV light was estimated to be 2.9×10^{14}–1.2×10^{15} photons/mm^2/hour depending on the sensor positions.

For the PDE monitoring, the charge for the 20 W xenon lamp was measured with the bandpass filter, a ND filter and the collimator, which was placed at the distance of 45 cm. The response for wavelength of 150–180 nm was extracted by comparing the charges measured in the atmosphere and gaseous nitrogen. The responses for other wavelengths were measured by illuminating the LEDs.

Fig. C.9 shows histories of the sensor response for different wavelengths. The charges for the VUV light decreased to 10% of the origin and were saturated at the same dose level as the irradiation test at room temperature (Sec. C.2.1). The difference of the saturation level can comes from the different wavelengths. The decrease was also observed for the wavelengths of 280 nm and 380 nm, and the saturation levels are 15% and 25%, respectively. A large variation of the decreasing speed among the chips was also observed for these wavelengths, which can be related to the difference of the irradiation speed. Meanwhile, the responses for longer wavelengths were almost stable during the irradiation. These results confirmed the effect of the VUV light irradiation on the PDE depends on the wavelength.

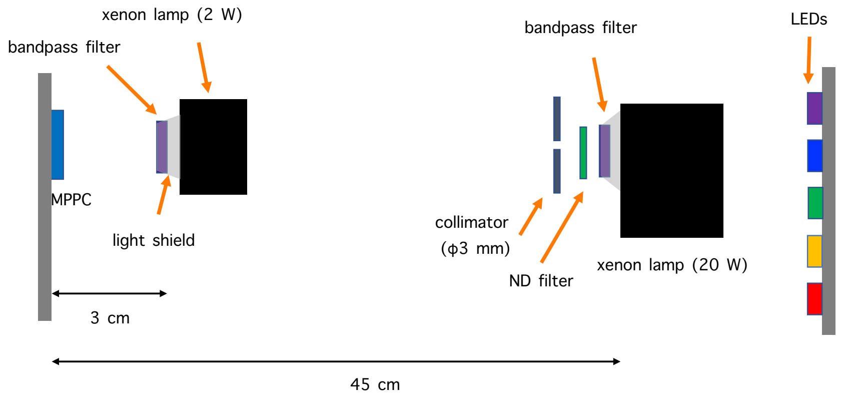
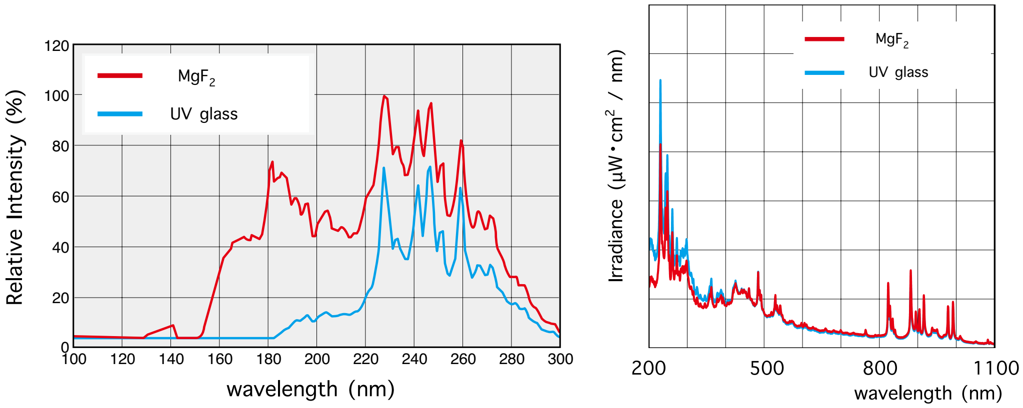
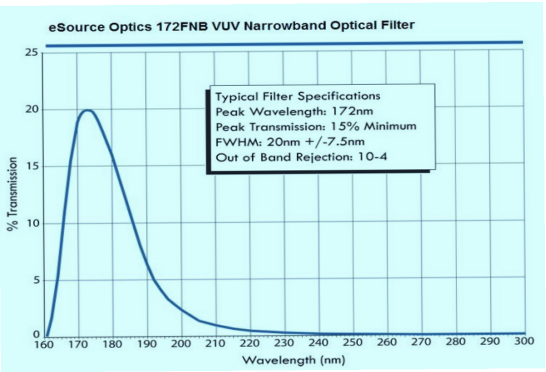


Figure C.6 Setup for the VUV photon irradiation to study on wavelength dependence.



(a) (b)

Figure C.7 Spectral radiation intensities of (a) 20 W xenon lamp (L12745-01-3) and (b) 2 W xenon lamp (L13651-01-3) (from [101], translated by the author). (a) was measured in gaseous nitrogen while (b) was in the atmosphere.



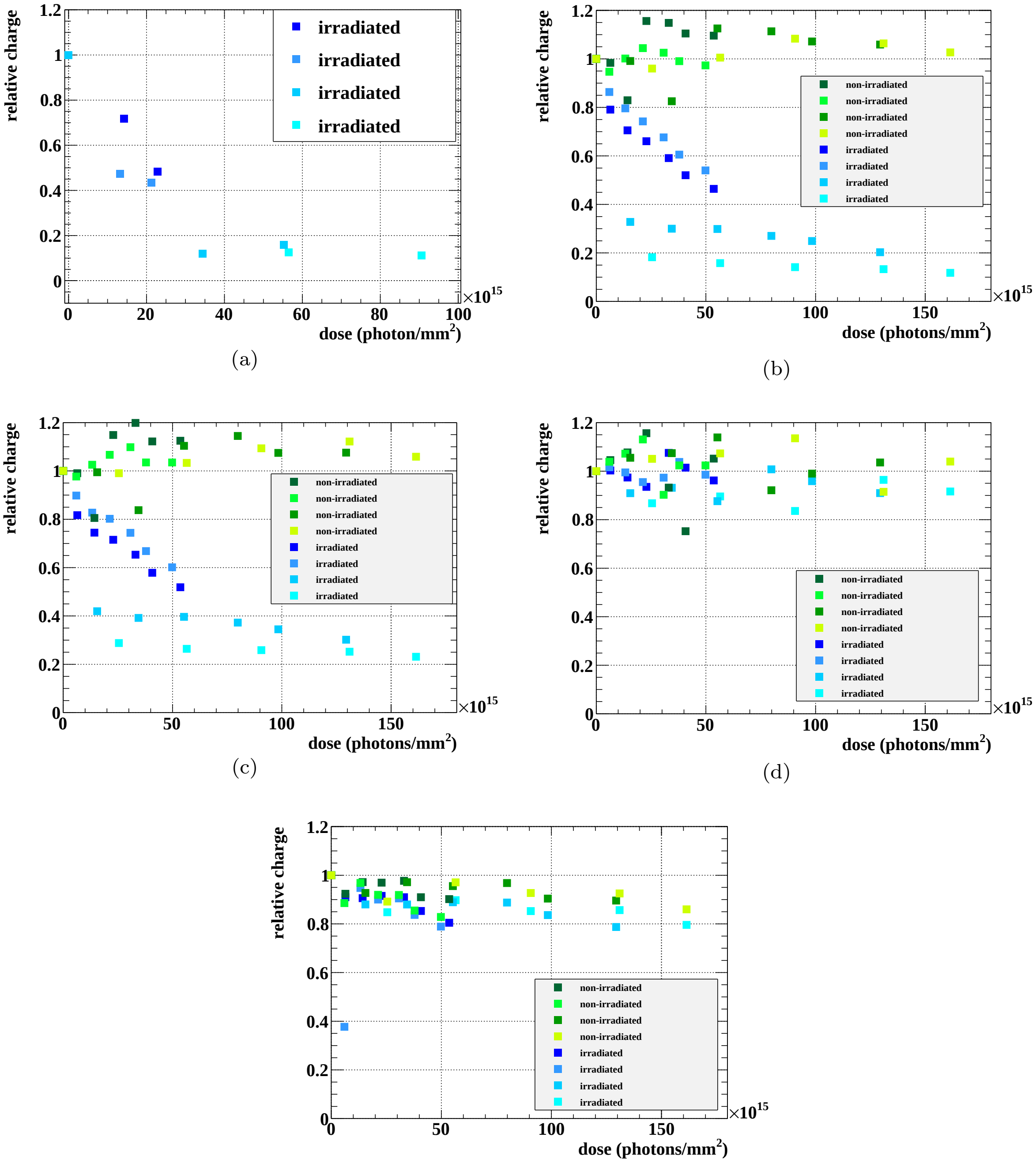
(a)

Figure C.8 Spectrum of the bandpass filter [103].

C.3 Gamma irradiation

The VUV responses of several VUV-MPPC samples irradiated with γ-rays were measured. The MPPC were irradiated using a ^{60}Co source ($E_\gamma = 1.17, 1.33\,\mathrm{MeV}$ at the Takasaki Advanced Radiation Research Institute in 2015. The total irradiation dose of each MPPC was 1–4 $\times\, 10^3$ Gy [104], which is much higher than the expectation of 0.6 Gy in the whole MEG II experiment.

In 2019, the VUV PDEs were measured with the setup shown in Fig. C.10. The MPPCs were placed on a circuit board and the signals from one or two chips on an MPPC were read out individually. Two blue light LEDs were installed to measure single photoelectron signals for gain calibration. An α source was attached on a tungsten wire and installed in the middle. The signals were amplified and their waveforms were acquired with the DRS evaluation board. The PDE was measured by comparing the expected number of photons from the α source and the measured number of photoelectrons.



(a)

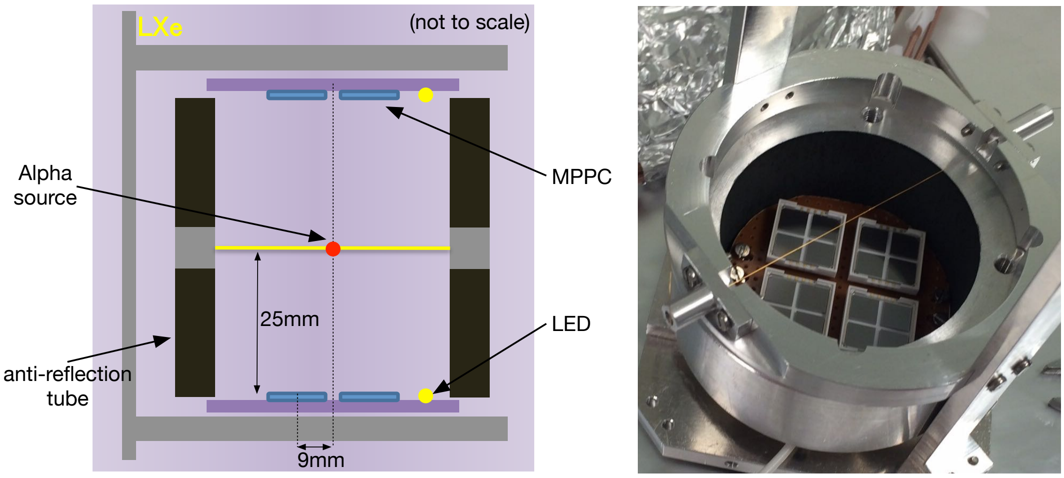
(b)

(c)

(d)

(e)

Figure C.9 Response for a fixed intensity of light with wavelengths of (a) 150–180 nm, (b) 280 nm, (c) 380 nm, (d) 465 nm, (e) 569 nm and (f) 645 nm. Different data series with similar colors show four chips on the same MPPC package. The *i*-th points of different chips were measured at the same irradiation time, and the variety of the dose levels at the same time is due to the position dependence of the light distribution. The data points in (a) are normalized by charges measured by the non-irradiated samples to compensate a large fluctuation due to the instability of the lamp.



(a) (b)

Figure C.10 Small chamber setup [62].

Fig. C.11(a) shows the measured PDEs of the γ-ray irradiated samples with that of a non-irradiated MPPC [1]. The PDEs of them were consistent with each other except for one irradiated MPPC, which may be due to some systematics or mistakes in the measurement. In any case, no significant degradation was observed even at the much higher irradiation fluence, and thus the PDE degradation observed at the LXe detector was not reproduced by this γ-ray irradiation. It must be, however, noted that the irradiated MPPCs were kept at room temperature for four years after the irradiation and a possibility of the recovery by an annealing at room temperature cannot be excluded.

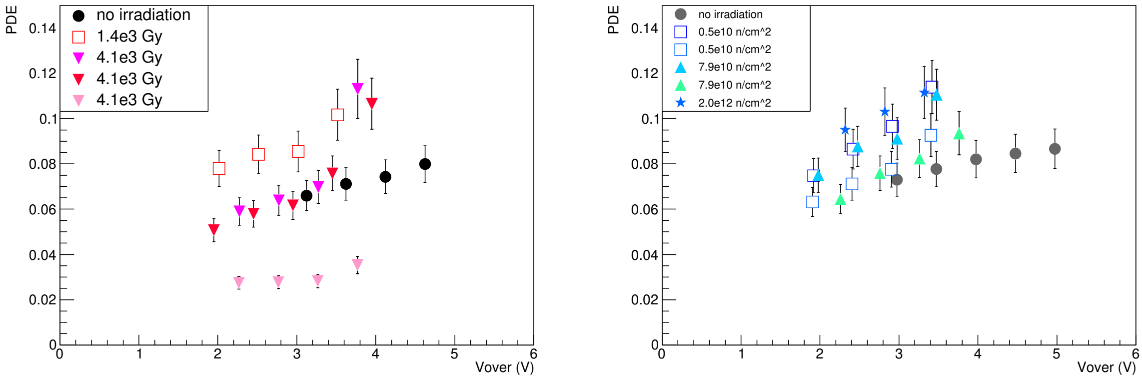
C.4 Neutron irradiation

The VUV responses of several VUV-MPPC samples irradiated with neutrons were also measured. The MPPCs were irradiated at the Tandem electrostatic accelerator at Kobe University in 2015. Neutrons from $^{9}\mathrm{Be} + \mathrm{d} \rightarrow {}^{10}\mathrm{B} + \mathrm{n}$ where deuterons were accelerated to 3 MeV were used for irradiation. The total fluence of each MPPC was measured by a ELMA diode. It ranges from 5×10^{9}–$2 \times 10^{12}\,\mathrm{n/cm^2}$ (1 MeV neutron equivalent) depending on the irradiation time and the geometrical distance from the target [104].

The VUV PDEs were measured with the same setup used for the γ-ray irradiated samples in 2019. As shown in Fig. C.11(b), no effect of irradiation was observed. Therefore, neutrons were excluded from the candidates though a possibility of recovery due to the annealing at room temperature still remains.

---

[1] The VUV PDE was not measured before the irradiation since its degradation was not expected.



(a)

(b)

Figure C.11 VUV PDEs of the irradiated MPPCs as a function of the over voltage in (a) gamma irradiation test, and (b) neutron irradiation test. Different colors show different MPPC chips. Black markers in (a) and gray markers in (b) are the MPPCs without irradiation, and the others are the MPPCs with different dose levels [62].

Acknowledgements

First of all, I would like to thank Prof. Toshinori Mori and Prof. Wataru Ootani for giving me a great opportunity to join the MEG II experiment. Especially, I received generous support from my academic supervisor, Prof. Wataru Ootani. I am grateful to Dr. Toshiyuki Iwamoto for giving me his kind supports not only to my research but also to my daily life in Switzerland. In addition, I would like to offer my special thanks for Dr. Yusuke Uchiyama, Dr. Kei Ieki and Dr. Shinji Ogawa for their insightful advices for my study.

Last but not least, I would like to express my appreciation for all the members of the MEG II collaboration for their generous help and work to progress the experiment. I am particularly grateful to my Japanese colleagues, Dr. Miki Nishimura, Dr. Mitsutaka Nakao, Mr. Masashi Usami, Mr. Satoru Kobayashi, Mr. Kazuki Toyoda, Mr. Atsushi Oya and Mr. Kosuke Yanai to share the wonderful time.

Bibliography

[1] A.M. Baldini et al. (MEG collaboration). Search for the lepton flavour violating decay $\mu^+ \to e^+\gamma$ with the full dataset of the MEG experiment. Eur. Phys. J. C, 76(8):434, 2016. arXiv:1605.05081, doi:10.1140/epjc/s10052-016-4271-x.

[2] A.M. Baldini et al. (MEG II collaboration). The design of the MEG II experiment. Eur. Phys. J. C, 78(5), May 2018. URL: http://dx.doi.org/10.1140/epjc/s10052-018-5845-6, doi:10.1140/epjc/s10052-018-5845-6.

[3] T.P. Cheng and L. Li. $\mu \to e\gamma$ in Theories with Dirac and Majorana Neutrino-Mass Terms. Phys. Rev. Lett., 45:1908–1911, Dec 1980. URL: https://link.aps.org/doi/10.1103/PhysRevLett.45.1908, doi:10.1103/PhysRevLett.45.1908.

[4] S.P. Martin. A supersymmetry primer. Advanced Series on Directions in High Energy Physics, page 198, Jul 1998. URL: http://dx.doi.org/10.1142/9789812839657_0001, doi:10.1142/9789812839657_0001.

[5] J. Hisano, T. Moroi, K. Tobe, and M. Yamaguchi. Lepton-flavor violation via right-handed neutrino Yukawa couplings in the supersymmetric standard model. Physical Review D, 53(5):24422459, Mar 1996. URL: http://dx.doi.org/10.1103/PhysRevD.53.2442, doi:10.1103/physrevd.53.2442.

[6] I. Esteban, M.C. Gonzalez-Garcia, M. Maltoni, I. Martinez-Soler, and T. Schwetz. Updated fit to three neutrino mixing: exploring the accelerator-reactor complementarity. Journal of High Energy Physics, 2017(87):090–090, Jan 2017. doi:https://doi.org/10.1007/JHEP01(2017)087.

[7] T. Goto, Y. Okada, T. Shindou, M. Tanaka, and R. Watanabe. Lepton flavor violation in the supersymmetric seesaw model after the LHC 8 TeV run. Phys. Rev. D, 91:033007, Feb 2015. URL: https://link.aps.org/doi/10.1103/PhysRevD.91.033007, doi:10.1103/PhysRevD.91.033007.

[8] L. Calibbi, D. Chowdhury, A. Masiero, K.M. Patel, and S.K. Vempati. Status of supersymmetric type-I seesaw in SO(10) inspired models. Journal of High Energy Physics, 2012(11):40, 2012. URL: https://doi.org/10.1007/JHEP11(2012)040, doi:10.1007/JHEP11(2012)040.

[9] Y. Kuno and Y. Okada. Muon decay and physics beyond the standard model. Rev. Mod. Phys., 73:151–202, 2001. arXiv:hep-ph/9909265, doi:10.1103/RevModPhys.73.151.

[10] A. Parker. SUSY searches (ATLAS/CMS): the lady vanishes, 2012. URL: https://indico.cern.ch/contributionDisplay.py?contribId=11&confId=181298.

[11] K. Agashe, A.E. Blechman, and F. Petriello. Probing the Randall-Sundrum geometric origin of flavor with lepton flavor violation. Phys. Rev. D, 74(5), Sep 2006. URL: http://dx.doi.org/10.1103/PhysRevD.74.053011, doi:10.1103/physrevd.74.053011.

[12] A.M. Iyer and S.K. Vempati. Lepton masses and flavor violation in the Randall-Sundrum model. Phys. Rev. D, 86:056005, Sep 2012. URL: https://link.aps.org/doi/10.1103/PhysRevD.86.056005, doi:10.1103/PhysRevD.86.056005.

[13] M. Beneke, P. Moch, and J. Rohrwild. Lepton flavour violation in RS models with a brane- or nearly brane-localized Higgs. Nuclear Physics B, 906:561–614, 2016. URL: https://www.sciencedirect.com/science/article/pii/S055032131600081X, doi:https://doi.org/10.1016/j.nuclphysb.2016.02.037.

[14] V. Cirigliano, A. Kurylov, M.J. Ramsey-Musolf, and P. Vogel. Lepton flavor violation without supersymmetry. Phys. Rev. D, 70:075007, Oct 2004. URL: https://link.aps.org/doi/10.1103/PhysRevD.70.075007, doi:10.1103/PhysRevD.70.075007.

[15] A.G. Akeroyd, M. Aoki, and Y. Okada. Lepton flavor violating τ decays in the left-right symmetric model. Phys. Rev. D, 76:013004, Jul 2007. URL: https://link.aps.org/doi/10.1103/PhysRevD.76.013004, doi:10.1103/PhysRevD.76.013004.

[16] C. Lee, P.S. Bhupal Dev, and R.N. Mohapatra. Natural TeV-scale left-right seesaw mechanism for neutrinos and experimental tests. Phys. Rev. D, 88:093010, Nov 2013. URL: https://link.aps.org/doi/10.1103/PhysRevD.88.093010, doi:10.1103/PhysRevD.88.093010.

[17] P.F. Pérez and C. Murgui. Lepton flavor violation in left-right theory. Phys. Rev. D, 95:075010, Apr 2017. URL: https://link.aps.org/doi/10.1103/PhysRevD.95.075010, doi:10.1103/PhysRevD.95.075010.

[18] S.R. Choudhury, A.S. Cornell, A. Deandrea, N. Gaur, and A. Goyal. Lepton flavor violation in the little Higgs model with T parity. Phys. Rev. D, 75:055011, Mar 2007. URL: https://link.aps.org/doi/10.1103/PhysRevD.75.055011, doi:10.1103/PhysRevD.75.055011.

[19] L. Wang and X. Han. Lepton flavor-violating processes in the simplest little Higgs model: $e^+e^-(\gamma\gamma) \to l_i\bar{l}_j$ under new bound from $l_i \to l_j\gamma$. Phys. Rev. D, 85:013011, Jan 2012. URL: https://link.aps.org/doi/10.1103/PhysRevD.85.013011, doi:10.1103/PhysRevD.85.013011.

[20] M. Blanke, A.J. Buras, B. Duling, S. Recksiegel, and C. Tarantino. FCNC Processes in the Littlest Higgs Model with T-Parity: an Update, 2010. arXiv:0906.5454.

[21] F. del Águila, J.I. Illana, and M.D. Jenkins. Lepton flavor violation in the Simplest Little Higgs model. Journal of High Energy Physics, 2011(3):80, 2011. URL: https://doi.org/10.1007/JHEP03(2011)080, doi:10.1007/JHEP03(2011)080.

[22] A. de Gouvéa and N. Saoulidou. Fermilab's intensity frontier. Annual Review of Nuclear and Particle Science, 60(1):513–538, 2010. URL: https://doi.org/10.1146/annurev-nucl-100809-131949, arXiv:https://doi.org/10.1146/annurev-nucl-100809-131949, doi:10.1146/annurev-nucl-100809-131949.

[23] A. Czarnecki, W.J. Marciano, and K. Melnikov. Coherent muon-electron conversion in muonic atoms. AIP Conference Proceedings, 435(1):409–418, 1998. URL: `https://aip.scitation.org/doi/abs/10.1063/1.56214`, `arXiv:https://aip.scitation.org/doi/pdf/10.1063/1.56214`, `doi:10.1063/1.56214`.

[24] W. Bertl et al. A search for μ-e conversion in muonic gold. The European Physical Journal C - Particles and Fields, 47(2):337–346, 2006. URL: `https://doi.org/10.1140/epjc/s2006-02582-x`, `doi:10.1140/epjc/s2006-02582-x`.

[25] L. Bartoszek et al. Mu2e Technical Design Report, 2015. `arXiv:1501.05241`.

[26] R.H. Bernstein. The Mu2e Experiment. Frontiers in Physics, 7, Jan 2019. URL: `http://dx.doi.org/10.3389/fphy.2019.00001`, `doi:10.3389/fphy.2019.00001`.

[27] M. Moritsu. The COMET Experiment: Search for Muon-to-Electron Conversion. URL: `https://journals.jps.jp/doi/abs/10.7566/JPSCP.33.011111`, `arXiv:https://journals.jps.jp/doi/pdf/10.7566/JPSCP.33.011111`, `doi:10.7566/JPSCP.33.011111`.

[28] N. Berger. The Mu3e Experiment. Nuclear Physics B - Proceedings Supplements, 248-250:35–40, 2014. 1st Conference on Charged Lepton Flavor Violation. URL: `https://www.sciencedirect.com/science/article/pii/S0920563214000085`, `doi:https://doi.org/10.1016/j.nuclphysbps.2014.02.007`.

[29] K. Arndt et al. Technical design of the phase I Mu3e experiment. Nucl. Instrum. Methods Phys. Res., A, 1014:165679, 2021. URL: `https://www.sciencedirect.com/science/article/pii/S0168900221006641`, `doi:https://doi.org/10.1016/j.nima.2021.165679`.

[30] G. Isidori, F. Mescia, P. Paradisi, and D. Temes. Flavor physics at large $\tan\beta$ with a binolike lightest supersymmetric particle. Phys. Rev. D, 75:115019, Jun 2007. URL: `https://link.aps.org/doi/10.1103/PhysRevD.75.115019`, `doi:10.1103/PhysRevD.75.115019`.

[31] B. Abi et al. Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm. Phys. Rev. Lett., 126:141801, Apr 2021. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.126.141801`, `doi:10.1103/PhysRevLett.126.141801`.

[32] G.W. Bennett et al. Final report of the E821 muon anomalous magnetic moment measurement at BNL. Phys. Rev. D, 73:072003, Apr 2006. URL: `https://link.aps.org/doi/10.1103/PhysRevD.73.072003`, `doi:10.1103/PhysRevD.73.072003`.

[33] T. Aushev et al. Physics at Super B Factory, 2010. `arXiv:1002.5012`.

[34] T. Abe et al. Belle II Technical Design Report, 2010. `arXiv:1011.0352`.

[35] L. Calibbi, A. Faccia, A. Masiero, and S.K. Vempati. Lepton flavour violation from SUSY-GUTs: Where do we stand for MEG, PRISM/PRIME and a super flavour factory. Phys. Rev. D, 74:116002, 2006. `arXiv:hep-ph/0605139`, `doi:10.1103/PhysRevD.74.116002`.

[36] M. Villanueva. Prospects for τ lepton physics at Belle II. SciPost Physics Proceedings, Feb 2019. `doi:10.21468/SciPostPhysProc.1.003`.

[37] The ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider.

Journal of Instrumentation, 3(08):S08003–S08003, Aug 2008. URL: https://doi.org/10.1088/1748-0221/3/08/s08003, doi:10.1088/1748-0221/3/08/s08003.

[38] T. Moroi, M. Nagai, and T.T. Yanagida. Lepton-flavor violations in high-scale SUSY with right-handed neutrinos. Physics Letters B, 728:342–346, 2014. URL: https://www.sciencedirect.com/science/article/pii/S0370269313009647, doi:https://doi.org/10.1016/j.physletb.2013.11.058.

[39] A. Parker. SUSY Summary Plot of ATLAS in June 2021. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2021-019/.

[40] K.A. Olive. Review of particle physics. Chinese Phys. C, 40(10):100001, oct 2016. URL: https://doi.org/10.1088/1674-1137/40/10/100001, doi:10.1088/1674-1137/40/10/100001.

[41] A.M. Baldini, Y. Bao, E. Baracchini, C. Bemporad, F. Berg, M. Biasotti, G. Boca, P.W. Cattaneo, G. Cavoto, F. Cei, and et al. Muon polarization in the MEG experiment: predictions and measurements. The European Physical Journal C, 76(4), Apr 2016. URL: http://dx.doi.org/10.1140/epjc/s10052-016-4047-3, doi:10.1140/epjc/s10052-016-4047-3.

[42] H.W. Kendall and M. Deutsch. Annihilation of positrons in flight. Phys. Rev., 101:20–26, Jan 1956. URL: https://link.aps.org/doi/10.1103/PhysRev.101.20, doi:10.1103/PhysRev.101.20.

[43] A.M. Baldini et al. (MEG collaboration). MEG Upgrade Proposal. arXiv:1301.7225, 2013. URL: https://arxiv.org/abs/1301.7225, arXiv:1301.7225.

[44] D. Kaneko. The final result of $\mu^+ \to e^+\gamma$ search with the MEG experiment. PhD thesis, The University of Tokyo, 2016. URL: https://www.icepp.s.u-tokyo.ac.jp/download/doctor/phD2016_kaneko.pdf.

[45] A. Kolano, A. Adelmann, R. Barlow, and C. Baumgarten. Intensity limits of the PSI Injector II cyclotron. Nucl. Instrum. Methods Phys. Res., A, 885:5459, Mar 2018. URL: http://dx.doi.org/10.1016/j.nima.2017.12.045, doi:10.1016/j.nima.2017.12.045.

[46] The proton accelerator at the Paul Scherrer Institute: forty years of top-flight research. URL: https://www.psi.ch/media/the-psi-proton-accelerator.

[47] Injector 2: a pre-accelerator for protons. URL: https://www.psi.ch/media/injector-2-a-pre-accelerator-for-protons.

[48] A. Papa. Search for the lepton flavour violation in $\mu^+ \to e^+\gamma$. The calibration methods for the MEG experiment. PhD thesis, The University of Pisa, 2009. URL: http://meg.icepp.s.u-tokyo.ac.jp/docs/theses/Angela.pdf.

[49] W. Ootani, W. Odashima, S. Kimura, T. Kobayashi, Y. Makida, T. Mitsuhashi, S. Mizumaki, R. Ruber, and A. Yamamoto. Development of a thin-wall superconducting magnet for the positron spectrometer in the MEG experiment. IEEE Trans. Appl. Supercond., 14(2):568–571, 2004. doi:10.1109/TASC.2004.829721.

[50] A.M. Baldini et al. The Search for $\mu^+ \to e^+\gamma$ with 10^{-14} Sensitivity: The upgrade of the MEG Experiment. Symmetry, 13(9), 2021. URL: https://www.mdpi.com/2073-8994/

13/9/1591, doi:10.3390/sym13091591.

[51] Particle Data Group. The Review of Particle Physics. Atomic and Nuclear Properties of Materials, 2019. URL: http://pdg.lbl.gov/2020/reviews/rpp2020-rev-atomic-nuclear-prop.pdf.

[52] K. Fujii, Y. Endo, Y. Torigoe, S. Nakamura, T. Haruyama, K. Kasami, S. Mihara, K. Saito, S. Sasaki, and H. Tawara. High-accuracy measurement of the emission spectrum of liquid xenon in the vacuum ultraviolet region. Nucl. Instrum. Meth. A, 795:293 – 297, 2015. URL: http://www.sciencedirect.com/science/article/pii/S016890021500724X, doi:https://doi.org/10.1016/j.nima.2015.05.065.

[53] A. Hitachi, T. Takahashi, N. Funayama, K. Masuda, J. Kikuchi, and T. Doke. Effect of ionization density on the time dependence of luminescence from liquid argon and xenon. Phys. Rev. B, 27:5279–5285, May 1983. URL: https://link.aps.org/doi/10.1103/PhysRevB.27.5279, doi:10.1103/PhysRevB.27.5279.

[54] T. Doke, A. Hitachi, J. Kikuchi, K. Masuda, H. Okada, and E. Shibamura. Absolute Scintillation Yields in Liquid Argon and Xenon for Various Particles. Jpn. J. Appl. Phys, 41(Part 1, No. 3A):1538–1545, Mar 2002. URL: https://doi.org/10.1143/jjap.41.1538, doi:10.1143/jjap.41.1538.

[55] T. Doke and K. Masuda. Present status of liquid rare gas scintillation detectors and their new application to gamma-ray calorimeters. Nucl. Instrum. Meth. A, 420(1):62 – 80, 1999. URL: http://www.sciencedirect.com/science/article/pii/S0168900298009334, doi:https://doi.org/10.1016/S0168-9002(98)00933-4.

[56] C.W. Fabjan and F. Gianotti. Calorimetry for particle physics. Rev. Mod. Phys., 75:1243–1286, Oct 2003. URL: https://link.aps.org/doi/10.1103/RevModPhys.75.1243, doi:10.1103/RevModPhys.75.1243.

[57] M.J. Berger et al. Xcom photon cross sections database. URL: hhttps://physics.nist.gov/PhysRefData/Xcom/html/xcom1.html, doi:https://dx.doi.org/10.18434/T48G6X.

[58] S. Agostinelli et al. GEANT4–a simulation toolkit. Nucl. Instrum. Meth. A, 506:250–303, 2003. doi:10.1016/S0168-9002(03)01368-8.

[59] R. Sawada. A Liquid Xenon Scintillation Detector to Search for the Lepton Flavor Violating Muon Decay with a Sensitivity of 10^{-13}. PhD thesis, The University of Tokyo, 2008. URL: https://www.icepp.s.u-tokyo.ac.jp/download/doctor/phD2008_sawada.pdf.

[60] Y. Nishimura. A Search for the Decay $\mu^+ \to e^+\gamma$ Using a High-Resolution Liquid Xenon Gamma-Ray Detector. PhD thesis, The University of Tokyo, 2010. URL: https://www.icepp.s.u-tokyo.ac.jp/download/doctor/phD2010_nishimura.pdf.

[61] K. Ieki, T. Iwamoto, D. Kaneko, S. Kobayashi, N. Matsuzawa, T. Mori, S. Ogawa, R. Onda, W. Ootani, R. Sawada, K. Sato, and R. Yamada. Large-area MPPC with enhanced VUV sensitivity for liquid xenon scintillation detector. Nucl. Instrum. Meth. A, 925:148 – 155, 2019. URL: http://www.sciencedirect.com/science/article/pii/

S0168900219301858, doi:https://doi.org/10.1016/j.nima.2019.02.010.

[62] S. Ogawa. Liquid xenon detector with highly granular scintillation readout to search for $\mu^+ \to e^+\gamma$ with sensitivity of 5×10^{-14} in the MEG experiment. PhD thesis, The University of Tokyo, 2020. URL: https://meg.web.psi.ch/docs/theses/ogawa_phd.pdf.

[63] T. Haruyama, K. Kasami, H. Inoue, S. Mihara, and Y. Matsubara. Development of a high-power coaxial pulse tube refrigerator for a liquid xenon calorimeter. AIP Conference Proceedings, 710, 06 2004. doi:10.1063/1.1774839.

[64] CRYOMECH AL300 CRYOCOOLERS. https://www.cryomech.com/products/al300/.

[65] S. Nakaura. Development of Radiative Decay Counter for ultimate sensitivity of MEG II experiment. Master's thesis, The University of Tokyo, 2016. URL: https://meg.web.psi.ch/docs/theses/nakaura_master.pdf.

[66] Saint-Gobain. BC-418, BC-420, BC-422 Premium Plastic Scintillators data sheet. URL: https://www.crystals.saint-gobain.com/sites/imdf.crystals.com/files/documents/bc418-420-422-data-sheet.pdf.

[67] Hamamatsu Photonics. S13360-3050PE data sheet. URL: http://www.hamamatsu.com/jp/en/S13360-3050PE.html.

[68] 3M. Enhanced Specular Reflector. URL: https://www.3m.com/3M/en_US/company-us/all-3m-products/~/3M-Enhanced-Specular-Reflector-3M-ESR-/?N=5002385+3293061534&rt=rud.

[69] R. Iwai. Development and commissioning of MEG II Radiative Decay Counter. Master's thesis, The University of Tokyo, 2017. URL: https://meg.web.psi.ch/docs/theses/iwai_master_final.pdf.

[70] S.I. of Ceramics. Lyso crystal data sheet. URL: http://www.siccas.com/LYSOScintillatorCrystal.htm.

[71] Saint-Gobain. Preludetm 420 data sheet. URL: https://www.crystals.saint-gobain.com/products/prelude-420-LYSO.

[72] Hamamatsu Photonics. S12572-025P data sheet. URL: http://www.hamamatsu.com.cn/UserFiles/DownFile/Product/20140117212906284.pdf.

[73] MIDAS system. URL: https://midas.triumf.ca/.

[74] S. Ritt, R. Dinapoli, and U. Hartmann. Application of the DRS chip for fast waveform digitizing. Nucl. Instrum. Meth. A, 623(1):486 – 488, 2010. URL: http://www.sciencedirect.com/science/article/pii/S0168900210006091, doi:https://doi.org/10.1016/j.nima.2010.03.045.

[75] M. Francesconi. The new trigger and data acquisition system for LFV searches in the MEG II experiment. PhD thesis, The University of Pisa, 2020. URL: https://etd.adm.unipi.it/theses/available/etd-04072020-171648/unrestricted/francesconi.pdf.

[76] M. Francesconi. The MEG II Trigger and Data Acquisition System. Master's thesis, The University of Pisa, 2017. URL: https://core.ac.uk/download/pdf/79622567.pdf.

[77] ROOT based Object oriented Midas Extension. https://elog.psi.ch/rome/.

[78] Y. Du and F. Retiére. After-pulsing and cross-talk in multi-pixel photon counters. Nucl. Instrum. Meth. A, 596:396–401, 2008. URL: https://www.sciencedirect.com/science/article/pii/S0168900208012643?via%3Dihub, doi:10.1016/j.nima.2008.08.130.

[79] R. Barlow. Extended maximum likelihood. Nucl. Instrum. Methods Phys. Res., A, 297(3):496–506, December 1990. doi:10.1016/0168-9002(90)91334-8.

[80] G. Cowan, K. Cranmer, E. Gross, and O. Vitells. Asymptotic formulae for likelihood-based tests of new physics. Eur. Phys. J. C, 71(2), Feb 2011. URL: http://dx.doi.org/10.1140/epjc/s10052-011-1554-0, doi:10.1140/epjc/s10052-011-1554-0.

[81] Toyoda Gosei E1L49–3B1A–02. https://www-jlc.kek.jp/~tauchi/index/LXeTPC/meetings/d070524-components/LED/E1L49_xxxxx_JEA.pdf.

[82] Kingbright KA-3021QBS-D. https://www.kingbright.com/attachments/file/psearch/000/00/00/KA-3021QBS-D(Ver.9B).pdf.

[83] A. M. Baldini et al. (MEG collaboration). A radioactive point-source lattice for calibrating and monitoring the liquid xenon calorimeter of the MEG experiment. Nucl. Instrum. Meth. A, 565(2):589 – 598, 2006. URL: http://www.sciencedirect.com/science/article/pii/S0168900206011685, doi:https://doi.org/10.1016/j.nima.2006.06.055.

[84] S. Kobayashi. Research on precise gamma-ray position measurement with MEG II liquid xenon detector. Master's thesis, The University of Tokyo, 2019. (in Japanese). URL: https://www.icepp.s.u-tokyo.ac.jp/download/master/m2018_kobayashi.pdf.

[85] K. Toyoda. Research and Development on Calibration of MEG II Gamma Ray Detector. Master's thesis, The University of Tokyo, 2020. URL: https://www.icepp.s.u-tokyo.ac.jp/download/master/m2019_toyoda.pdf.

[86] I. Nakamura. Radiation damage of pixelated photon detector by neutron irradiation. Nucl. Instrum. Meth. A, 610(1):110 – 113, 2009. New Developments In Photodetection NDIP08. URL: http://www.sciencedirect.com/science/article/pii/S0168900209010407, doi:https://doi.org/10.1016/j.nima.2009.05.086.

[87] Y. Qiang, C. Zorn, F. Barbosa, and E. Smith. Radiation hardness tests of SiPMs for the JLab Hall D Barrel calorimeter. Nucl. Instrum. Meth. A, 698:234 – 241, 2013. URL: http://www.sciencedirect.com/science/article/pii/S0168900212011473, doi:https://doi.org/10.1016/j.nima.2012.10.015.

[88] M.C. Vignali, E. Garutti, R. Klanner, D. Lomidze, and J. Schwandt. Neutron irradiation effect on SiPMs up to $\Phi_{neq} = 5 \times 10^{14}$ cm^{-2}. Nucl. Instrum. Meth. A, 912:137 – 139, 2018. doi:10.1016/j.nima.2017.11.003.

[89] T. Matsubara. Radiation damage of MPPC by gamma-ray irradiation with Co 60. PoS, PD07:032, 2008. doi:10.22323/1.051.0032.

[90] J. Adam et al. The MEG detector for $\mu^+ \to e^+\gamma$ decay search. Eur. Phys. J. C, 73(4), Apr 2013. URL: http://dx.doi.org/10.1140/epjc/s10052-013-2365-2, doi:10.1140/epjc/s10052-013-2365-2.

[91] S. Kobayashi, M. Francesconi, L. Galli, K. Ieki, T. Iwamoto, T. Libeiro, N. Matsuzawa, W. Molzon, T. Mori, M. Nakao, S. Ogawa, R. Onda, and W. Ootani. Pre-

cise measurement of 3D-position of SiPMs in the liquid xenon gamma-ray detector for the MEG II experiment. Nucl. Instrum. Meth. A, 936:189 – 191, 2019. Frontier Detectors for Frontier Physics: 14th Pisa Meeting on Advanced Detectors. URL: http://www.sciencedirect.com/science/article/pii/S0168900218314840, doi:https://doi.org/10.1016/j.nima.2018.10.170.

[92] M. Tan and Q.V. Le. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, 2020. arXiv:1905.11946.

[93] A.G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications, 2017. arXiv:1704.04861.

[94] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 770–778, Los Alamitos, CA, USA, jun 2016. IEEE Computer Society. URL: https://doi.ieeecomputersociety.org/10.1109/CVPR.2016.90, doi:10.1109/CVPR.2016.90.

[95] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. Chen. MobileNetV2: Inverted Residuals and Linear Bottlenecks, 2019. arXiv:1801.04381.

[96] M. Tan, B. Chen, R. Pang, V. Vasudevan, M. Sandler, A. Howard, and Q.V. Le. MnasNet: Platform-Aware Neural Architecture Search for Mobile, 2019. arXiv:1807.11626.

[97] W. Riegler, C. Lippmann, and R. Veenhof. Detector Physics and Simulation of Resistive Plate Chambers. Nucl. Instrum. Methods Phys. Res., A, 500:144–162. 27 p, Jun 2002. URL: http://cds.cern.ch/record/570462, doi:10.1016/S0168-9002(03)00337-1.

[98] A. Oya, K. Ieki, A. Ochi, R. Onda, W. Ootani, and K. Yamamoto. Development of high-rate capable and ultra-low mass Resistive Plate Chamber with Diamond-Like Carbon, 2021. arXiv:2109.13525.

[99] R.R. Shinde and E. Yuvaraj. Fabrication and Characterization of Glass Resistive Plate Chamber (RPC). URL: https://www.tifr.res.in/~ehep2019/pdfs/exp04_rpc.pdf.

[100] E. Garutti and Y. Musienko. Radiation damage of SiPMs. Nucl. Instrum. Meth. A, 926:69–84, 2019. arXiv:1809.06361, doi:10.1016/j.nima.2018.10.191.

[101] Hamamatsu xenon flash lump catalog. https://www.hamamatsu.com/resources/pdf/etd/Xe-F_TLS1022J.pdf.

[102] K. Shimada. Research on photon detection efficiency of VUV-MPPC in MEG II liquid xenon detector. Master's thesis, The University of Tokyo, 2021. (in Japanese). URL: https://meg.web.psi.ch/docs/theses/shimada_master.pdf.

[103] eSource optics P/N 12172FNB. https://www.esourceoptics.com/catalog/item/8090484/8726585.htm.

[104] N. Shibata. Development of HCAL scintillator cell unit for ILC detector. Master's thesis, The University of Tokyo, 2016. (in Japanese). URL: https://www.icepp.s.u-tokyo.ac.jp/download/master/m2015_shibata.pdf.