

Ph.D. Thesis

In situ high voltage generation in xenon gas
time projection chamber and track pattern recognition for
neutrinoless double beta decay search

(ニュートリノを伴わない二重ベータ崩壊探索のための、
キセノンガス TPC 内での高電圧生成とトラックパターン認識)

Department of Physics, Graduate School of Science, Tohoku University

Shinichi Akiyama

September 2025

Abstract

It is known that neutrinos have masses, but the reason for their extremely small masses compared to other charged leptons and quarks is unknown. The Majorana nature of neutrinos is considered to be an important key to understand the reason of the smallness of neutrino masses, and also to understand the origin of the matter – antimatter asymmetry of the universe. Experimentally, the only feasible way to confirm the Majorana nature of neutrinos is the observation of neutrinoless double beta ($0\nu\beta\beta$) decay.

A Xenon ElectroLuminescence detector (AXEL) is a high-pressure xenon gas time projection chamber (TPC) with the aim of observing the $0\nu\beta\beta$ decay of ^{136}Xe . The AXEL detector employs a cell-based readout structure called electroluminescence light collection cell (ELCC) to detect electroluminescence signals, enabling both good energy resolution and three-dimensional track reconstruction capability. The electron track in the TPC forms a “blob” at its endpoint due to increased energy deposition. As a result, $0\nu\beta\beta$ events, in which two electrons are emitted, are expected to exhibit two blobs, whereas single-electron tracks originating from gamma rays or beta decays would show only one blob. A 180 L prototype detector is currently being developed, focusing on the development of elemental technologies, verification of scaling-up techniques, and evaluation of energy resolution and track reconstruction near the Q value of $0\nu\beta\beta$.

An important technological element in the TPC is the generation of high voltage to form the drift electric field. A Cockcroft-Walton (CW) multiplier can be used for this purpose, as it converts a low voltage AC input to a high voltage DC output. The CW multiplier to supply high voltage to the AXEL detector was developed and installed at the 180 L prototype detector. The CW multiplier was implemented on a flexible printed circuit (FPC) board and coated with methyl silicone resin as a discharge countermeasure. Measurements with the 180 L prototype detector confirmed that the electrical pickup from the AC input of the CW multiplier to the signal line is sufficiently small, and stable operation of 40 days has been achieved. Data of the 2615 keV gamma rays from ^{208}Tl using thorium-doped tungsten rods as the source confirmed an energy resolution of $(0.672 \pm 0.083) \% \text{ FWHM}$. An interpolation based on the 2615 keV gamma rays from ^{208}Tl and other gamma ray peaks yielded an estimated energy resolution of $(0.678 \pm 0.010) \%$ at the Q value of $0\nu\beta\beta$, which is close to the design goal of 0.5 %. A tracks of the 2615 keV gamma-ray photoabsorption events (an electron track) and double escape events (an electron and a positron track) are reconstructed, confirming that they each have one and two blob structures, respectively.

In the search for $0\nu\beta\beta$ decay, the suppression of background events is of critical importance. To eliminate background events using track information, a discrimination model for $0\nu\beta\beta$ and background using machine learning was developed. The model is based on DenseNet and was trained using simulation data generated by Geant4. The model achieved a classification accuracy of 91.7 % on the validation dataset. However, when inputting the real 2615 keV gamma-ray data obtained by measurement, differences were observed in the shape of the signal likelihood distribution of gamma-ray backgrounds. This is likely due to the differences between the simulation data used for training the model and the real data. Several summary quantities of the simulation data and the real data were compared, and the cause of the observed differences in the signal likelihood distribution was investigated.

Using the performance of the model on the validation data, the sensitivity of a future 1-ton detector

is estimated. Three types of background sources were considered: ^{214}Bi , ^{208}Tl , and ^{137}Xe . ^{214}Bi and ^{208}Tl exist in materials and emit gamma rays with energies near the Q-value of xenon $0\nu\beta\beta$ decay, that is, 2458 keV. ^{137}Xe is produced via neutron capture on ^{136}Xe , induced by cosmic-ray neutrons, and undergoes beta decay with a Q-value of 4.16 MeV. We estimated an expected sensitivities for an 1 ton detector. Assuming a future improvement in energy resolution to 0.320 % FWHM and a pressure vessel made of radiopure oxygen-free copper, a lower limit on the $0\nu\beta\beta$ decay half-life of 1.12×10^{27} years at 90 % confidence level is expected for a 10-year measurement, which is 2.9 times better than the current world record.

List of Publications

- (1) Shinichi Akiyama, Junya Hikida, Masashi Yoshida, Kazuhiro Nakamura, Sei Ban, Masanori Hirose, Atsuko K Ichikawa, Yoshihisa Iwashita, Tatsuya Kikawa, Yasuhiro Nakajima, Kiseki D Nakamura, Tsuyoshi Nakaya, Shuhei Obara, Ken Sakashita, Hiroyuki Sekiya, Bungo Sugashima, Soki Urano, Sota Hatsumi, Sota Kobayashi, Hayato Sasaki, In Situ High-Voltage Generation with a Cockcroft – Walton Multiplier for a Xenon Gas Time Projection Chamber, Progress of Theoretical and Experimental Physics, Volume 2025, Issue 5, May 2025, 053H03, <https://doi.org/10.1093/ptep/ptaf066>
- (2) Masashi Yoshida, Kazuhiro Nakamura, Shinichi Akiyama, Sei Ban, Junya Hikida, Masanori Hirose, Atsuko K Ichikawa, Yoshihisa Iwashita, Yukimasa Kashino, Tatsuya Kikawa, Akihiro Minamino, Kentaro Miuchi, Yasuhiro Nakajima, Kiseki D Nakamura, Tsuyoshi Nakaya, Shuhei Obara, Ken Sakashita, Hiroyuki Sekiya, Hibiki Shinagawa, Bungo Sugashima, Soki Urano, High-pressure xenon gas time projection chamber with scalable design and its performance around the Q value of ^{136}Xe double-beta decay, Progress of Theoretical and Experimental Physics, Volume 2024, Issue 1, January 2024, 013H01, <https://doi.org/10.1093/ptep/ptad146>
- (3) J. Hikida, S. Akiyama, S. Ban, M. Hirose, A.K. Ichikawa, Y. Iwashita, T. Kikawa, Y. Nakajima, K.D. Nakamura, T. Nakaya, S. Obara, K. Sakashita, H. Sekiya, M. Yoshida, B. Sugashima, S. Urano, S. Hatsumi, S. Kobayashi, H. Sasaki, L. Evan, AXEL: High-pressure xenon gas TPC for neutrinoless double beta decay search, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2025, 170706, ISSN 0168-9002, <https://doi.org/10.1016/j.nima.2025.170706>.

Acknowledgment

I would like to express my sincere gratitude to all those who supported my doctoral research. First and foremost, I am deeply thankful to Professor Atsuko Ichikawa, my principal advisor. Her profound insights into physics consistently illuminated the path forward in my research, all while respecting my autonomy. I am also grateful for her generous support beyond the scope of academic matters.

I would like to extend my appreciation to Assistant Professor Kiseki Nakamura, whose guidance was invaluable in the daily progress of my research. His precise and robust engineering skills were indispensable to the development of my work, and having the opportunity to learn from him firsthand was an immensely valuable experience.

I am also grateful to Dr. Masashi Yoshida for sharing his extensive knowledge as I took over the research on the Cockcroft – Walton multiplier. It was a fortunate opportunity to engage with such a significant and practical topic in hardware development.

Furthermore, I would like to thank Mr. Shunsuke Tanaka. My research on machine learning would not have been possible without his prior work, which laid the essential foundation for my studies.

I would like to express my sincere appreciation to Mr. Bungo Sugashima, Mr. Junya Hikida and Mr. Hayato Sasaki for their extensive support in operating the detector in Kyoto. In particular, I am deeply grateful to Mr. Junya Hikida for his assistance throughout the entire process, from data acquisition to analysis.

Mr. Sota Kobayashi and Mr. Evan Lim Ee Hui have taken over and further developed the research on the Cockcroft – Walton multiplier. I am confident that their creative engineering efforts will lead to even greater advancements.

Mr. Soki Urano is working on the development of a novel method for detecting scintillation light, and I had the opportunity to contribute to his simulation work. Improving the detection efficiency of scintillation light is fundamentally important for the AXEL detector, and I look forward to the progress of his research.

I would also like to thank Mr. Sota Hatsumi for his generous support with the frontend board. In particular, his implementation of a method to stabilize the startup of the ADCH significantly reduced the effort required for data acquisition. I am sincerely grateful for his contributions.

Ms. Yuko Watanabe has taken over the detector simulation work that I was unable to complete. I am hopeful that her efforts will lead to a deeper understanding of the detector.

I would also like to express my gratitude to the other members of the AXEL collaboration for their valuable insights and support through meetings and discussions. Additionally, I am thankful to Ms. Saeko Daido for handling many administrative tasks on my behalf. Her meticulous work was greatly appreciated.

Finally, I would like to express my heartfelt gratitude to my family, who have always respected my choices and supported me throughout this journey.

*Shinichi Akiyama
Sendai, Japan
September 2025*

Contents

Part I	Introduction	9
Chapter 1	Introduction	11
1.1	Neutrino	11
1.2	Majorana mass and See-Saw mechanism	13
1.3	Leptogenesis	16
1.4	Neutrinoless double beta decay	17
1.5	The key factors in $0\nu\beta\beta$ search experiments	19
1.6	On-going $0\nu\beta\beta$ search experiments	20
1.7	Outline of this thesis	23
Chapter 2	AXEL experiment	25
2.1	Overview of the experiment	25
2.2	Roadmap of the AXEL experiment	27
2.3	180 L prototype detector	27
Part II	In-situ high voltage generation with Cockcroft-Walton multiplier	39
Chapter 3	Design and performance of the Cockcroft-Walton multiplier	41
3.1	Design of the CW multiplier	41
3.2	High voltage generation	43
3.3	CW surface insulation	45
3.4	Outgassing rate of the coated CW circuit board	45
3.5	Monitoring of the HV output voltage for the 180 L detector	46
3.6	Countermeasure for surface discharge on HDPE tube	47
Chapter 4	Data taking with CW multiplier, analysis, and performance evaluation	51
4.1	Data taking with CW multiplier	51
4.2	Analysis	54
4.3	Energy resolution	60
4.4	Breakdown of the energy resolution	61
4.5	Track topology	67
Part III	Signal vs. background discrimination by track pattern using machine learning	69
Chapter 5	Simulation for machine learning	71
5.1	Background assumptions	71
5.2	Creation of data set for training model	72

5.3	Data set characteristics	76
Chapter 6	Track pattern recognition by machine learning	85
6.1	Preprocessing	85
6.2	Neural network	87
6.3	Valuation index	92
6.4	Training the model	93
6.5	Likelihood histogram	96
6.6	Misclassified events	97
6.7	Dataset bias	98
Chapter 7	Future sensitivity of the AXEL experiment	103
7.1	Background source	103
7.2	Volume cut	105
7.3	Energy resolution cut	106
7.4	Topology cut using machine learning	106
7.5	Sensitivity for 1-ton detector	107
Part IV	Summary	111
Chapter 8	Future improvements	113
8.1	CW upgrade	113
8.2	Deep learning model upgrade	113
8.3	Energy resolution improvement	114
Chapter 9	Conclusion	115
	Bibliography	117

Part I

Introduction

Chapter 1

Introduction

1.1 Neutrino

The neutrino is a neutral lepton introduced by W.Pauli to explain the continuous spectrum of electron energy associated with beta decays. Since neutrinos interact very little with matter, it was not until the 1950s that (anti-)electron neutrinos were actually detected through an experiment using a nuclear reactor by Reines and Cowan[1]. Subsequently, an experiment using an accelerator lead by L.Lederman, M.Schwartz, and J.Steinberger observed muon neutrinos, demonstrating that neutrinos have flavors[2]. In 1975, Perl discovered the tau lepton[3], and its counterpart, the tau neutrino, was later detected through the DONUT experiment[4]. To date, no new generation of neutrinos have been discovered, and measurements of the Z boson decay width indicate that there are exactly three generations of active neutrinos with masses less than half that of the Z boson[5].

In the Standard Model, neutrinos are considered to be massless. However, the observation of atmospheric[6] and solar[7][8] neutrinos have revealed that neutrinos have mass and undergo neutrino oscillations, changing flavors over time. Neutrino oscillation occurs when the mass eigenstates of neutrinos $|\nu_i\rangle$ ($i = 1, 2, 3$) differ from their flavor eigenstates $|\nu_l\rangle$ ($l = e, \mu, \tau$). The relationship between these eigenstates is expressed by the following mixing matrix U_{li} .

$$|\nu_l\rangle = \sum_i U_{li} |\nu_i\rangle \quad (1.1)$$

$$\begin{aligned} U &= \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned} \quad (1.2)$$

Here, $s_{ij} = \sin \theta_{ij}$, $c_{ij} = \cos \theta_{ij}$, θ_{ij} are mixing angles, and δ is a CP violating phase. The matrix is called Pontecorvo – Maki – Nakagawa – Sakata (PMNS) matrix. Then, the probability of flavor oscillation from ν_α to ν_β is given by the following equation.

$$P_{\nu_\alpha \rightarrow \nu_\beta}(t) \simeq \delta_{\alpha\beta} - 4 \sum_{i>j} \text{Re}(\Lambda_{ij}^{\alpha\beta}) \sin^2 \left[\frac{\Delta m_{ij}^2 L}{4E} \right] + 2 \sum_{i>j} \text{Im}(\Lambda_{ij}^{\alpha\beta}) \sin \left[\frac{\Delta m_{ij}^2 L}{2E} \right] \quad (1.3)$$

$$\Lambda_{ij}^{\alpha\beta} = U_{\alpha i} U_{\beta i}^* U_{\alpha j}^* U_{\beta j} \quad (1.4)$$

Here, δ_{ij} is the Kronecker delta, $\Delta m_{ij}^2 = m_i^2 - m_j^2$ is the mass-squared difference between the mass

eigenstates, $|\nu_i\rangle$ and $|\nu_j\rangle$. As can be seen from Eq. 1.3, neutrino oscillation do not depends on the absolute values of the neutrino masses, but on the mass-squared differences. Table 1.1 shows the current best-fit values of the oscillation parameters. Various measurements have been conducted and the order of the

Table 1.1 Current best-fit values of the oscillation parameters. NO denotes normal ordering and IO denotes inverted ordering. Each parameter is presented with its best-fit value and the corresponding one sigma range. $\Delta m_{3l}^2 = \Delta m_{31}^2$ for NO and $\Delta m_{3l}^2 = \Delta m_{32}^2$ for IO. These values are taken from [9][10].

Parameter	NO	IO
$\sin^2 \theta_{12}$	$0.308^{+0.012}_{-0.011}$	$0.308^{+0.012}_{-0.011}$
θ_{12}	$33.68^{+0.73^\circ}_{-0.70^\circ}$	$33.68^{+0.73^\circ}_{-0.70^\circ}$
$\sin^2 \theta_{23}$	$0.470^{+0.017}_{-0.013}$	$0.562^{+0.012}_{-0.015}$
θ_{23}	$43.3^{+1.0^\circ}_{-0.8^\circ}$	$48.6^{+0.7^\circ}_{-0.9^\circ}$
$\sin^2 \theta_{13}$	$0.02215^{+0.00056}_{-0.00058}$	$0.02224^{+0.00056}_{-0.00057}$
θ_{13}	$8.56^{+0.11^\circ}_{-0.11^\circ}$	$8.58^{+0.11^\circ}_{-0.11^\circ}$
Δm_{21}^2	$7.49^{+0.19}_{-0.19} \times 10^{-5} \text{ eV}^2$	$7.49^{+0.19}_{-0.19} \times 10^{-5} \text{ eV}^2$
Δm_{3l}^2	$+2.513^{+0.021}_{-0.019} \times 10^{-3} \text{ eV}^2$	$-2.510^{+0.024}_{-0.025} \times 10^{-3} \text{ eV}^2$
δ	$212^{+26^\circ}_{-41^\circ}$	$285^{+25^\circ}_{-28^\circ}$

mass eigenvalues is not yet experimentally determined and there are two possibilities: $m_1 < m_2 \ll m_3$ and $m_3 \ll m_1 < m_2$. The former is called the normal ordering, and the latter is called the inverted ordering. The corresponding neutrino mixing matrices are given in Eq. 1.5 and 1.6.

$$U^{(\text{NO})} = \begin{pmatrix} 0.82 & 0.55 & -0.13 + 0.079i \\ -0.33 + 0.045i & 0.65 + 0.030i & 0.68 \\ 0.46 + 0.048i & -0.52 + 0.032i & 0.72 \end{pmatrix} \quad (1.5)$$

$$U^{(\text{IO})} = \begin{pmatrix} 0.82 & 0.55 & 0.039 + 0.14i \\ -0.39 + 0.090i & 0.53 + 0.060i & 0.74 \\ 0.39 + 0.079i & -0.64 + 0.053i & 0.65 \end{pmatrix} \quad (1.6)$$

Regarding the direct measurement of the absolute value of the neutrino mass, the KATRIN experiment measures tritium beta decay and provides the constraints

$$m_{\bar{\nu}_e} < 0.45 \text{ eV} \quad (1.7)$$

at 90 % confidence level (C.L.)[11]. Since $\bar{\nu}_e$ itself is not a mass eigenstate, the mass measured here is the weighted average of the mass eigenstates. From Eq. 1.1, $\bar{\nu}_e$ is expressed as a linear combination of the mass eigenstates $\bar{\nu}_i$, with coefficients U_{ei} . The probability that $\bar{\nu}_e$ corresponds to each mass eigenstate is given by $|U_{ei}|^2$. Therefore, the expected value of the mass is weighted by $|U_{ei}|^2$.

$$|m_{\bar{\nu}_e}|^2 = \sum_i |U_{ei}|^2 m_i^2 \quad (1.8)$$

Measurements of the cosmic microwave background provide constraints on the sum of mass eigenstate

values; observations from the Planck satellite have put the following constraints

$$\sum_i m_i < 0.26 \text{ eV} \quad (1.9)$$

at 95 % confidence level[12].

These measurements have revealed that neutrinos have mass, but it is more than six orders of magnitude lighter than that of charged leptons or quarks. The underlying reason for this is unknown.

1.2 Majorana mass and See-Saw mechanism

This section summarizes the mechanisms that give mass to neutrinos. The discussion in this section is based on references [13] and [14].

The equation governing the fermion field was derived by Dirac[15]. The Dirac equation for a free fermion field ψ with mass m_D is presented in Eq. 1.10.

$$i\gamma^\mu \partial_\mu \psi - m_D \psi = 0 \quad (1.10)$$

γ^μ denotes the gamma matrices, which, in the Dirac representation, are given as follows.

$$\gamma^0 = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} \mathbf{0} & \sigma^i \\ -\sigma^i & \mathbf{0} \end{pmatrix}, \quad \gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix} \quad (1.11)$$

σ^i ($i = 1, 2, 3$) denotes the Pauli matrix (Eq. 1.12), $\mathbf{1}$ represents the 2×2 identity matrix, and $\mathbf{0}$ denotes the 2×2 zero matrix.

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (1.12)$$

ψ is a four-component spinor. In this representation of the gamma matrices, ψ is expressed as a pair of two-component spinors u, v corresponding to positive and negative energy states.

$$\psi = \begin{pmatrix} u \\ v \end{pmatrix} \quad (1.13)$$

The Lagrangian corresponding to the Dirac equation is given in Eq. 1.14, where the bilinear term $m_D \bar{\psi} \psi$ is referred to as the mass term.

$$\mathcal{L} = i\bar{\psi}\gamma^\mu \partial_\mu \psi - m_D \bar{\psi} \psi \quad (1.14)$$

Let's decompose the fermion field into chirality eigenstates $\psi = \psi_R + \psi_L$.

$$\begin{aligned} \psi_R &\equiv \gamma_R \psi = \frac{1}{2}(1 + \gamma^5)\psi = \frac{1}{2} \begin{pmatrix} u + v \\ u + v \end{pmatrix} \equiv \frac{u + v}{\sqrt{2}} |\psi_R\rangle \\ \psi_L &\equiv \gamma_L \psi = \frac{1}{2}(1 - \gamma^5)\psi = \frac{1}{2} \begin{pmatrix} u - v \\ -(u - v) \end{pmatrix} \equiv \frac{u - v}{\sqrt{2}} |\psi_L\rangle \end{aligned} \quad (1.15)$$

The adjoint spinor $\bar{\psi} = \psi^\dagger \gamma^0$ is expressed as follows.

$$\begin{aligned} \overline{\psi_R} &= \overline{\gamma_R \psi} = (\gamma_R \psi)^\dagger \gamma^0 = \psi^\dagger \gamma_R^\dagger \gamma^0 = \psi^\dagger \gamma^0 \gamma_L = \bar{\psi} \gamma_L \\ \overline{\psi_L} &= \overline{\gamma_L \psi} = (\gamma_L \psi)^\dagger \gamma^0 = \psi^\dagger \gamma_L^\dagger \gamma^0 = \psi^\dagger \gamma^0 \gamma_R = \bar{\psi} \gamma_R \end{aligned} \quad (1.16)$$

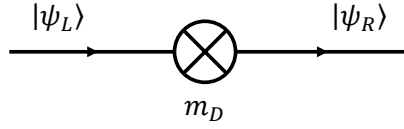


Fig. 1.1 Transition between left-handed particle and right-handed particle. The Dirac mass m_D arises from this type of transition.

Here, we used the relation $\gamma_{R,L}^\dagger = \gamma_{R,L}$ and the anticommutation relations of the gamma matrices $\{\gamma^5, \gamma^\mu\} = 0$. Thus, the Lagrangian in Eq. 1.14 is expressed as follows.

$$\begin{aligned}\mathcal{L} &= i(\overline{\psi_R} + \overline{\psi_L})\gamma^\mu\partial_\mu(\psi_R + \psi_L) - m_D(\overline{\psi_R} + \overline{\psi_L})(\psi_R + \psi_L) \\ &= i(\overline{\psi_R}\gamma^\mu\partial_\mu\psi_R + \overline{\psi_L}\gamma^\mu\partial_\mu\psi_L) - m_D(\overline{\psi_R}\psi_L + \overline{\psi_L}\psi_R)\end{aligned}\quad (1.17)$$

The orthogonality of the chiral projection matrices $\gamma_R\gamma_L = \gamma_L\gamma_R = 0$ was used in this step. From the above Lagrangian, we obtain the chirality-based Dirac equation.

$$\begin{cases} i\gamma^\mu\partial_\mu\psi_R - m_D\psi_L = 0 \\ i\gamma^\mu\partial_\mu\psi_L - m_D\psi_R = 0 \end{cases}\quad (1.18)$$

Similarly, by applying the particle-antiparticle transformation $\psi^c \equiv i\gamma_2\psi^*$, the Lagrangian and Dirac equation for the antiparticle can be expressed as follows.

$$\mathcal{L} = i(\overline{\psi_L^c}\gamma^\mu\partial_\mu\psi_L^c + \overline{\psi_R^c}\gamma^\mu\partial_\mu\psi_R^c) - m_D(\overline{\psi_L^c}\psi_R^c + \overline{\psi_R^c}\psi_L^c)\quad (1.19)$$

$$\begin{cases} i\gamma^\mu\partial_\mu\psi_L^c - m_D\psi_R^c = 0 \\ i\gamma^\mu\partial_\mu\psi_R^c - m_D\psi_L^c = 0 \end{cases}\quad (1.20)$$

We note that $(\psi_L)^c = i\gamma_2(\gamma_L\psi)^* = \gamma_R(i\gamma_2\psi^*) = (\psi^c)_R$, $(\psi_R)^c = (\psi^c)_L$ and we assume ψ_L^c means $(\psi^c)_L$.

For simplicity, assuming the fermion is at rest, that is, $\partial_i\psi = 0$ ($i = 1, 2, 3$), the Dirac equation in Eq. 1.18 becomes as follows.

$$\begin{cases} \dot{\psi}_R = -im_D\gamma_0\psi_L \\ \dot{\psi}_L = -im_D\gamma_0\psi_R \end{cases}\quad (1.21)$$

This means that the amplitude of the left-handed (right-handed) component changes in proportion to the amplitude of the right-handed (left-handed) component, with an intensity of m_D . This corresponds to the transition shown in Fig. 1.1, and the mass arising from such a transition is called the ‘‘Dirac mass’’.

Let’s consider the following term.

$$\mathcal{L} = -\frac{M_L}{2}(\overline{\psi_R^c}\psi_L + \overline{\psi_L}\psi_R^c) - \frac{M_R}{2}(\overline{\psi_L^c}\psi_R + \overline{\psi_R}\psi_L^c)\quad (1.22)$$

By adding these terms to the Lagrangian of Eq. 1.17 and Eq. 1.19, the Dirac equation is modified as

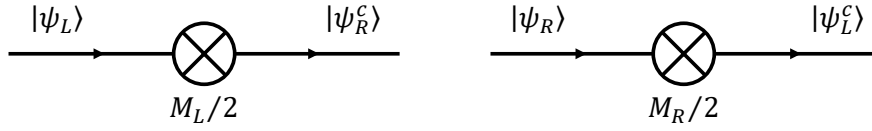


Fig. 1.2 Transition associated with Majorana mass.

follows:

$$\begin{cases} \dot{\psi}_R = -im_D\gamma_0\psi_L - i\frac{M_R}{2}\gamma_0\psi_L^c \\ \dot{\psi}_L = -im_D\gamma_0\psi_R - i\frac{M_L}{2}\gamma_0\psi_R^c \\ \dot{\psi}_L^c = -im_D\gamma_0\psi_R^c - i\frac{M_R}{2}\gamma_0\psi_R \\ \dot{\psi}_R^c = -im_D\gamma_0\psi_L^c - i\frac{M_L}{2}\gamma_0\psi_L \end{cases} \quad (1.23)$$

Here, too, the fermions are assumed to be at rest. This term induces additional transitions between particles and antiparticles as shown in Fig. 1.2. Such transitions are forbidden for charged leptons by charge conservation, but they can occur for neutrinos since neutrinos do not carry charge. The mass arising from these transitions is called the “Majorana mass”. In this case, the mass term of neutrino field ν in the Lagrangian can be expressed using the neutrino mass matrix M as follows:

$$\mathcal{L}_{\text{mass}} = -\frac{1}{2} \begin{pmatrix} \overline{\nu}_R^c & \overline{\nu}_R \end{pmatrix} M \begin{pmatrix} \nu_L \\ \nu_L^c \end{pmatrix} + h.c. \quad (1.24)$$

$$M = \begin{pmatrix} M_L & m_D \\ m_D & M_R \end{pmatrix} \quad (1.25)$$

By diagonalizing this mass matrix with a unitary matrix U , two mass eigenvalues, m_ν and m_N , are obtained,

$$U^T M U = \begin{pmatrix} m_\nu & 0 \\ 0 & m_N \end{pmatrix} \quad (1.26)$$

$$m_{N,\nu} = \frac{1}{2} \left(M_L + M_R \pm \sqrt{(M_L - M_R)^2 + 4m_D^2} \right) \quad (1.27)$$

Since the mass eigenvalue m_ν becomes negative when $m_D^2 > M_L M_R$, a orthogonal matrix O and a diagonal phase matrix ρ are introduced, leading to the following expression.

$$O = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}, \quad \rho = \begin{pmatrix} \rho_\nu & 0 \\ 0 & \rho_N \end{pmatrix} \quad (1.28)$$

$$U^T M U = \rho^T O^T M O \rho = \begin{pmatrix} \rho_\nu^2 m_\nu & 0 \\ 0 & \rho_N^2 m_N \end{pmatrix} \quad (1.29)$$

$$\tan 2\theta = \frac{2m_D}{M_R - M_L} \quad (1.30)$$

Here, ρ_N^2 is always equal to 1, while ρ_ν^2 is taken to be 1 when m_ν is positive, and -1 when m_ν is negative. In this way, the mass eigenvalues m_ν and m_N can always be made positive. The mass eigenstates

n_{kL} ($k = \nu, N$) satisfy the following relation:

$$\begin{pmatrix} \nu_L \\ \nu_L^c \end{pmatrix} = O\rho \begin{pmatrix} n_{\nu L} \\ n_{NL} \end{pmatrix} \quad (1.31)$$

The diagonalized mass term is expressed as follows.

$$\mathcal{L}_{\text{mass}} = -\frac{1}{2} \sum_{k=\nu, N} m_k \overline{n_{kR}^c} n_{kL} + h.c. \quad (1.32)$$

This mass term is a purely Majorana mass term, and it indicates that $n_k = n_{kL} + (n_{kL})^c$ is a Majorana particle.

The See-Saw mechanism[16][17][18] provides an explanation for the smallness of neutrino masses by taking into account both Dirac and Majorana mass terms. In the seesaw mechanism, it is assumed that $M_L = 0$ and $m_D \ll M_R$. In the Standard Model, a Majorana mass term for left-handed neutrinos is forbidden because the Higgs triplet required to construct a Lorentz-invariant term does not exist. While there is no guide to choose the right-handed Majorana mass M_R , assuming $m_D \ll M_R$ leads to the following result.

$$m_\nu \simeq \frac{m_D^2}{M_R} \ll m_D, \quad m_N \simeq M_R, \quad \tan \theta \simeq \frac{m_D}{M_R} \ll 1, \quad \rho_\nu^2 = -1 \quad (1.33)$$

Even if the Dirac mass of the neutrino is comparable to that of quarks or charged leptons, the presence of a large right-handed Majorana mass M_R can result in a very small neutrino mass m_ν . Assuming the Dirac mass of the neutrino is on the order of 100 GeV, comparable to that of charged leptons and quarks, and the observed neutrino mass m_ν is approximately 0.1 eV, the mass of the right-handed neutrino m_N corresponds to about 10^{14} GeV. Since the mixing angle θ is small, the mass eigenstates can be approximated as $n_{\nu L} \sim \nu_L$ and $n_{NL} \sim \nu_L^c$. This implies that the neutrino observed in weak interactions effectively corresponds to the light Majorana neutrino $n_{\nu L}$, whereas the non-interacting right-handed neutrino corresponds to the heavy Majorana neutrino n_{NL} .

1.3 Leptogenesis

Our world is composed of matter, and regions dominated by antimatter have not been observed in our universe. It is believed that in the early universe, baryons and antibaryons were created in equal amounts and some process led to a slight excess of baryons, resulting in the matter-dominated universe we observe today. This baryogenesis process is said to require the following three Sakharov conditions:

- (1) The existence of reactions that violate baryon number.
- (2) The violation of both C and CP symmetries.
- (3) Departure from thermal equilibrium.

In the Standard Model, the sphaleron process is a reaction that violates baryon number. In the sphaleron process, the difference between baryon number B and lepton number L , $B - L$, is conserved. Regarding (2), the relevant mechanism is described by the Kobayashi – Maskawa matrix[19]. This matrix describes the mixing between quark flavors and their weak interaction eigenstates, and its complex phase leads to the violation of CP symmetry. For (3), the electroweak phase transition could create the departure from the thermal equilibrium. However, this electroweak baryogenesis model is ruled out because the Kobayashi-Maskawa matrix does not provide sufficient asymmetry[20][21], and the large mass of the Higgs particle prevents the electroweak phase transition from being a first-order phase transition[22].

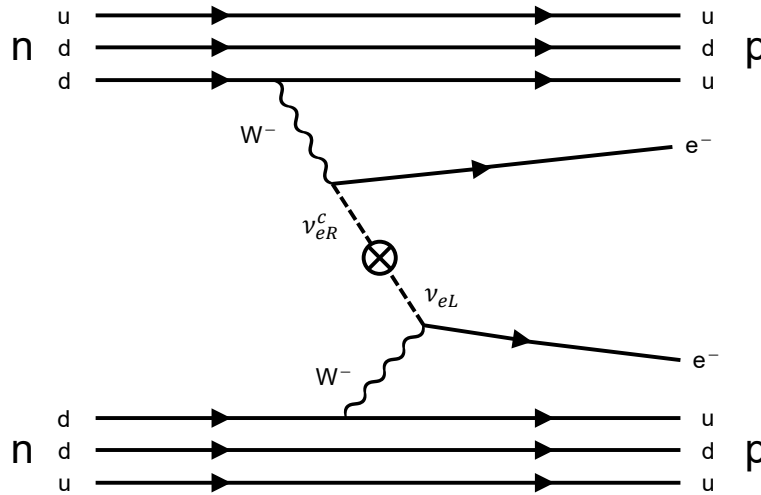


Fig. 1.3 $0\nu\beta\beta$ decay through light Majorana neutrino exchange.

If neutrinos are Majorana particles, the decay of possible heavy Majorana neutrinos can generate a net lepton number, violating CP symmetry, too. Leptogenesis is a scenario in which this net lepton number is converted into a net baryon number through the sphaleron processes, resulting in the matter-dominated universe[23].

1.4 Neutrinoless double beta decay

As mentioned above, it is important to confirm the Majorana nature of neutrinos, which is a prerequisite for the See-Saw mechanism that explains the extremely light mass of neutrinos and for leptogenesis that explains the matter-antimatter asymmetry of the universe. Neutrinoless double beta decay ($0\nu\beta\beta$) constitutes the only practically feasible experimental method for probing the Majorana nature of neutrinos.

Neutrinoless double beta decay is a phenomenon in which two beta decays occur simultaneously in a same nucleus and the nucleus emits only two electrons[24]. If neutrinos are Majorana particles, $\nu_R^c \rightarrow \nu_L$ transitions can occur, leading to $0\nu\beta\beta$ (Fig. 1.3). Such reaction is called the light Majorana neutrino exchange reaction. The $0\nu\beta\beta$ reaction can occur through mechanisms other than those depicted in Fig. 1.3. Regardless of the specific reaction pathway, the occurrence of $0\nu\beta\beta$ indicates that neutrinos have a finite Majorana mass[25]. Because, in general, the presence of the $0\nu\beta\beta$ process ($0 \rightarrow u\bar{u}d\bar{d}ee$) allows for the $\nu_{eR}^c \rightarrow \nu_{eL}$ transition as shown in Fig. 1.4 through the combination of weak interactions, which implies that neutrinos have a finite Majorana mass.

In the observation of $0\nu\beta\beta$, two electrons produced by the decay would be detected. In $2\nu\beta\beta$, which occurs as a second-order effect in the Standard Model, neutrinos carry part of the nuclear decay energy $Q_{\beta\beta}$. Since neutrinos cannot be detected, the sum of the kinetic energies of the two electrons forms a continuous spectrum with $Q_{\beta\beta}$ as the upper limit. In contrast, in $0\nu\beta\beta$, the sum of the kinetic energies of the two electrons produces a mono energy peak at $Q_{\beta\beta}$ in the spectrum (Fig. 1.5).

The half-life of $0\nu\beta\beta$ $T_{1/2}^{0\nu}$ is given by the following equation.

$$\frac{1}{T_{1/2}^{0\nu}} = G^{0\nu} |M^{0\nu}|^2 \langle m_{\beta\beta} \rangle^2 \quad (1.34)$$

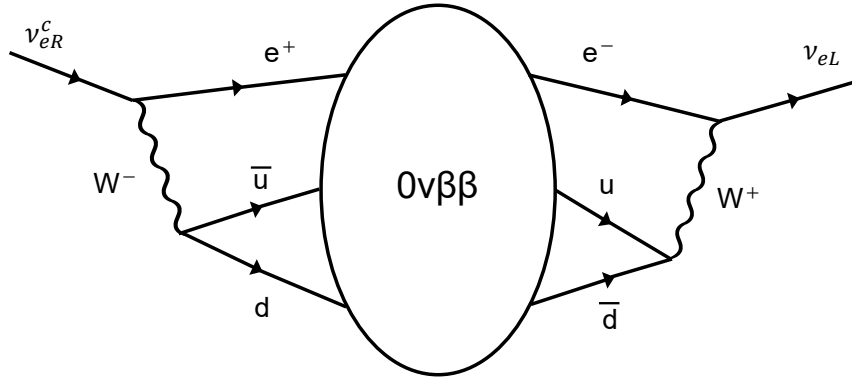


Fig. 1.4 The $\nu_{eR}^c \rightarrow \nu_{eL}$ transition occurring through any $0\nu\beta\beta$ process.

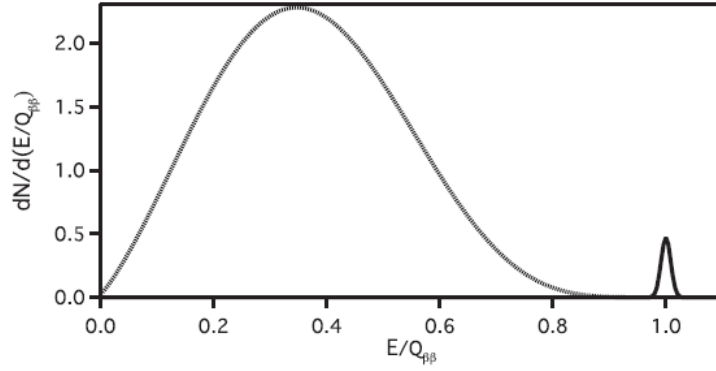


Fig. 1.5 Schematic spectrum of the sum of electron energies for $2\nu\beta\beta$ (dotted curve) and $0\nu\beta\beta$ (solid curve). Figure from [26].

Here, $G^{0\nu}$ is called the phase space factor and is determined by the decay energy $Q_{\beta\beta}$ and the atomic number Z of the double beta decay nucleus. $M^{0\nu}$ is the nuclear matrix element, which is estimated by theoretical models and has uncertainties. $\langle m_{\beta\beta} \rangle$ is called the $0\nu\beta\beta$ effective neutrino mass and is defined as follows. If neutrinos are Majorana particles, the neutrino mixing matrix has two additional complex phases, known as Majorana CP phases (α_{21}, α_{31}).

$$\begin{aligned}
 U &= \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \\
 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{i\alpha_{21}} & 0 \\ 0 & 0 & e^{i\alpha_{31}} \end{pmatrix}
 \end{aligned} \tag{1.35}$$

The effective neutrino mass is given by the following expression.

$$\begin{aligned}
 \langle m_{\beta\beta} \rangle &= \left| \sum_i U_{ei}^2 m_i \right| \\
 &= |c_{12}^2 c_{13}^2 m_1 + s_{12}^2 c_{13}^2 e^{2i\alpha_{21}} m_2 + s_{13}^2 e^{2i(\alpha_{31}-\delta)} m_3|
 \end{aligned} \tag{1.36}$$

From the neutrino mixing angles and mass squared differences obtained through neutrino oscillation

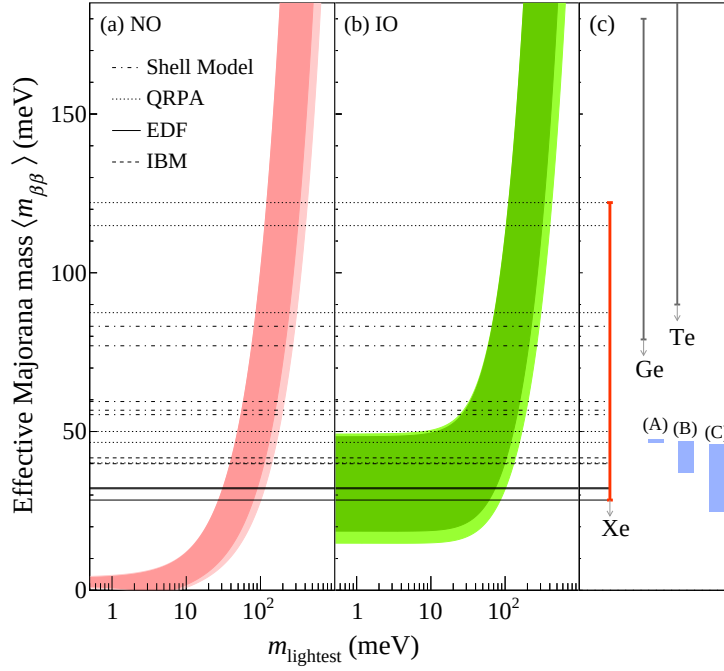


Fig. 1.6 Allowed region of the $0\nu\beta\beta$ effective neutrino mass $\langle m_{\beta\beta} \rangle$ as a function of the lightest neutrino mass m_{lightest} . The shaded regions are based on neutrino oscillation parameters for (a) the normal mass ordering (NO) and (b) the inverted mass ordering (IO). The horizontal lines indicate 90% C.L. upper limits on $\langle m_{\beta\beta} \rangle$ with ^{136}Xe from KamLAND-Zen for various nuclear matrix element calculations. (c) The corresponding limits for ^{136}Xe , ^{76}Ge , and ^{130}Te . (A)-(C) are theoretical predictions in the IO region. This figure is taken from [27].

experiments, the possible range of $\langle m_{\beta\beta} \rangle$ is determined, as shown in Fig. 1.6.

To date, $0\nu\beta\beta$ has not been observed, and only lower limits on its half-life have been provided. The next section will discuss the key factors in improving the sensitivity of $0\nu\beta\beta$ search experiments.

1.5 The key factors in $0\nu\beta\beta$ search experiments

The essence of $0\nu\beta\beta$ search lies in securing a large number of decay nuclei and thoroughly reducing background events.

Even if $0\nu\beta\beta$ decay occurs, it is extremely rare, with a half-life longer than 10^{26} years. To approach the effective neutrino mass beyond the inverted mass ordering region and to reach the normal region, ton-scale decay nuclei are required. Therefore, factors such as availability, the abundance ratio of the target isotopes, and the ease of isotope enrichment need to be considered.

Background reduction significantly affects the sensitivity of the experiment. For an exposure Mt , where M tons of double beta decay nuclei are monitored over t years, the sensitivity of the half-life $T_{1/2}^{0\nu}$ without background is proportional to the exposure,

$$T_{1/2}^{0\nu} \propto (\ln 2) \varepsilon \frac{N_A}{A} Mt \quad (1.37)$$

Here, ε is the signal detection efficiency, N_A is the Avogadro constant and A is the atomic number of target nuclei. However, in a presence of significant background, the sensitivity is proportional to the square root of the exposure,

$$T_{1/2}^{0\nu} \propto (\ln 2) \varepsilon \frac{N_A}{A} \sqrt{\frac{Mt}{b\Delta E}} \quad (1.38)$$

Here, $b\Delta E$ is the number of background events occurring within the signal region of ΔE width per unit time and unit mass. Therefore, it is necessary to minimize background contaminations. One major source is environmental radioactivity that generates background events near the $Q_{\beta\beta}$ energy, for example, in the case of ^{136}Xe , such as ^{214}Bi from the uranium series and ^{208}Tl from the thorium series. Cosmic rays could generate background events, so it is necessary to conduct experiments deep underground. If, for example, β -rays tracks or signal waveform shapes are available, it may further reduce background events beyond just energy selection. Even with these methods, the background that ultimately cannot be eliminated is $2\nu\beta\beta$. The high-energy tail of $2\nu\beta\beta$ cannot be distinguished from $0\nu\beta\beta$ by anything other than energy differences. The amount of $2\nu\beta\beta$ contamination in the region $-\Delta E$ from $Q_{\beta\beta}$ is approximately proportional to $(\Delta E/Q_{\beta\beta})^6$ [28]. Therefore, it is also important to improve the energy resolution.

1.6 On-going $0\nu\beta\beta$ search experiments

Various $0\nu\beta\beta$ experiments are being conducted with various nuclides. The Q values, natural abundances, and (lower limits of) half-lives of $2\nu\beta\beta$ and $0\nu\beta\beta$ for candidate nuclides are summarized in Tab. 1.2.

Table 1.2 Summary of the candidate isotopes used in $0\nu\beta\beta$ search experiments. $Q_{\beta\beta}$ values are from [29]. Natural abundances are from [30]. Lower limits of $T_{1/2}^{0\nu}$ are 90 % C.L..

nuclei	$Q_{\beta\beta}$ (keV)	natural abundance	$T_{1/2}^{2\nu}$ (years)	$T_{1/2}^{0\nu}$ lower limit (years)
^{48}Ca	4268	0.187 %	$(6.4_{-0.6}^{+0.7+1.2}) \times 10^{19}$ [31]	2.0×10^{22} [31]
^{76}Ge	2039	7.73 %	$(2.022 \pm 0.018 \pm 0.038) \times 10^{21}$ [32]	1.8×10^{26} [33]
^{82}Se	2998	8.73 %	$(9.39 \pm 0.17 \pm 0.58) \times 10^{19}$ [34]	2.5×10^{23} [34]
^{96}Zr	3356	2.80 %	$(2.35 \pm 0.14 \pm 0.16) \times 10^{19}$ [35]	9.2×10^{21} [35]
^{100}Mo	3034	9.82 %	$(6.90 \pm 0.15 \pm 0.37) \times 10^{18}$ [36]	1.5×10^{24} [37]
^{116}Cd	2813	7.49 %	$(2.74 \pm 0.04 \pm 0.18) \times 10^{19}$ [38]	1.0×10^{23} [38]
^{130}Te	2528	34.08 %	$(7.71_{-0.06}^{+0.08+0.12}) \times 10^{20}$ [39]	3.2×10^{25} [40]
^{136}Xe	2458	8.8573 %	$(2.34_{-0.46}^{+0.80+0.30}) \times 10^{21}$ [41]	3.8×10^{26} [27]
^{150}Nd	3371	5.638 %	$(9.34 \pm 0.22_{-0.60}^{+0.62}) \times 10^{18}$ [42]	2.0×10^{22} [42]

Currently, the KamLAND-Zen experiment has provided the most stringent lower limit on the half-life of ^{136}Xe $0\nu\beta\beta$ decay at 3.8×10^{26} years (90 % C.L.)[27]. Figure 1.7 shows a schematic view of the KamLAND-Zen experiment. They encapsulated 745 kg of 91 % enriched xenon, dissolved in liquid scintillator, inside a nylon inner balloon (IB) with a diameter of 3.8 meters. The IB is in an outer balloon with a diameter of 13 meters, which is in a stainless steel spherical tank contains 1 kiloton of liquid scintillator. The tank has a diameter of 18 meters and equipped with 1,325 17-inch photomultiplier tubes (PMTs) and 554 20-inch PMTs to detect scintillation light produced by xenon $0\nu\beta\beta$ decay. The exterior of the stainless steel tank functions as a water Cherenkov detector for shielding and cosmic-ray tagging. It consists of 3200 m³ of ultra-pure water and 140 20-inch PMTs. Since scintillation light is used as a signal, the estimated energy resolution is relatively modest i.e. $6.7\% / \sqrt{E(\text{MeV})}$ [27] and the significant background is the $2\nu\beta\beta$ of ^{136}Xe , which limits the sensitivity of the experiment. In addition, spallation of ^{136}Xe and ^{12}C in the liquid scintillator by cosmic-ray muons, as well as the subsequent neutron captures on ^1H , ^{12}C , and ^{136}Xe , produce unstable nuclei. Among these, long-lived isotopes that are difficult to eliminate through delayed coincidence techniques constitute a significant background.

Many of these issues can be effectively addressed by employing a xenon gas time projection chamber

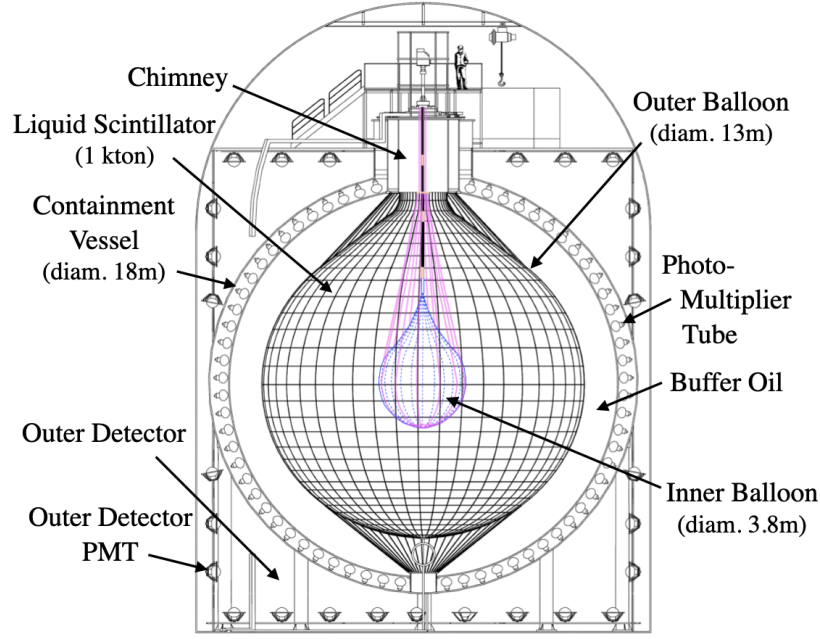


Fig. 1.7 Schematic view of the KamLAND-Zen experiment. This figure is taken from [43].

(TPC). The characteristics of the xenon gas TPC, a promising candidate for the search for $0\nu\beta\beta$, are described below.

^{136}Xe is an element with many advantages for $0\nu\beta\beta$ search experiments. As shown in Tab. 1.2, it has a relatively high natural abundance of 8.9%, and isotope separation technology is well-established, making enrichment easier. It is chemically stable and easy to purify. Additionally, being a gas, it allows for easier scalability to large detector masses compared to crystalline detectors such as germanium. Furthermore, the half-life of $2\nu\beta\beta$ decay is longer compared to other nuclides, resulting in less contamination in $0\nu\beta\beta$ signals. It can be a medium of a gas detector, for which the signal consists of ionization electrons and ions generated by the passage of charged particles. The average energy W required to produce one ionization electron in xenon gas is 22.1 eV[44], and the Fano factor is 0.13[45], resulting in an intrinsic energy resolution of 0.26% (FWHM) at the energy of $Q_{\beta\beta}$. This energy resolution is significantly superior to that of liquid scintillators, allowing for a substantial reduction in the contamination of $2\nu\beta\beta$.

TPC, proposed by D. R. Nygren[46], is a detector that enables three-dimensional track reconstruction by detecting ionization electrons generated by the passage of charged particles. An electric field is applied within the detector to drift the ionization electrons toward the readout plane. The energy of the charged particle is measured from the amount of the ionization electrons detected at the readout plane. By using the scintillation light produced by the charged particle as the event timing and measuring the time when the ionization electrons are detected, the position along the drift direction can be determined. By pixelating the readout plane, it is possible to reconstruct the three-dimensional track of the particle. Since a drift electric field is applied in the TPC, long-lived nuclear ions produced by spallation are collected to the cathode electrode. As a result, beta rays emitted from these ions can be effectively removed by applying a fiducial volume cut. Neutral ^{137}Xe , produced via neutron absorption, has a half-life of 3.8 minutes and a beta decay Q -value of 4.16 MeV. It is an unavoidable background source even in xenon gas TPCs, but its activation can be suppressed by mixing 0.1% of ^3He with the xenon[47]. In addition, three-dimensional track information can be utilized to identify reaction particles. Alpha and electron tracks can be distinguished based on their different ranges. Additionally, beta rays undergo multiple scattering and have a high energy loss only at low energies, resulting in a characteristic “blob” at the end

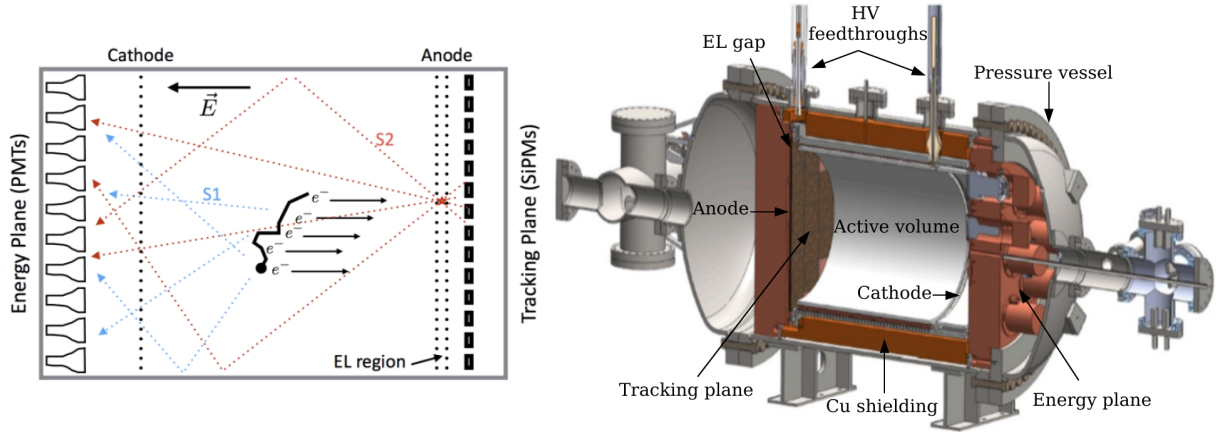


Fig. 1.8 Schematic diagram[50] (left) and cross-sectional view[51] (right) of the NEXT-White detector.

of the track. This allows for obtaining information on the number and direction of the electron tracks. This information can be used to eliminate background in $0\nu\beta\beta$ search experiments.

From the above perspectives, the xenon gas TPC can be considered to be one of the ideal detectors for $0\nu\beta\beta$ searches. In the following, experiments using the xenon gas TPCs for $0\nu\beta\beta$ searches are described.

1.6.1 NEXT

The concept of using a xenon gas TPC for $0\nu\beta\beta$ search was proposed by D.R Nygren[48] and the NEXT experiment[49] achieves high energy resolution and particle tracking performance with a xenon gas TPC by using a process called “electroluminescence (EL)”. The EL process refers to the excitation of gas atoms by electrons accelerated in a high electric field, followed by the emission of photons as the de-excitation of gas atoms. This process is different from the avalanche process in that it is a linear signal amplification process induced directly by the initial electrons, hence it has smaller fluctuations in the amplification.

They have completed measurements with a prototype detector NEXT-White, which uses about 5 kg of xenon. Figure 1.8 shows a schematic diagram and a cross-sectional view of the NEXT-white detector. When charged particles pass through, primary scintillation (S1) is emitted, and subsequently, ionization electrons are collected on the anode side and amplified in the EL region (S2). They have two detection planes: an energy plane using 12 photomultiplier tubes (PMTs) on the cathode side and a tracking plane using 1792 silicon photomultipliers (SiPMs) on the anode side[51]. The S1 and S2 signals are detected by the PMTs in the energy plane. The S1 signal is used to determine the event timing, and the S2 signal is used to measure the energy E of the charged particles. The track pattern is obtained by the tracking plane located behind the anode. They achieved a resolution of 1 % (FWHM) at 2.6 MeV with this detector[51] and conducted lifetime measurements of $2\nu\beta\beta$ [41]. They are now taking data with an 100 kg detector, NEXT-100.

They have developed an analysis method to further reduce background by deconvoluting diffused and blurred track images using the Richardson-Lucy algorithm[52]. Additionally, they are working on technology to achieve background-free experiment by directly detecting $^{136}\text{Ba}^{2+}$ produced from the decay of ^{136}Xe [53].

1.6.2 PandaX

PandaX-III experiment also uses a xenon gas TPC to search for $0\nu\beta\beta$ decay. A schematic diagram of the PandaX-III detector is shown in Fig. 1.9. Their working gas is a 10 bar Xe-(1%) trimethylamine gas mixture, and the detector contains about 140 kg 90 % enriched ^{136}Xe inside the TPC[54]. TMA

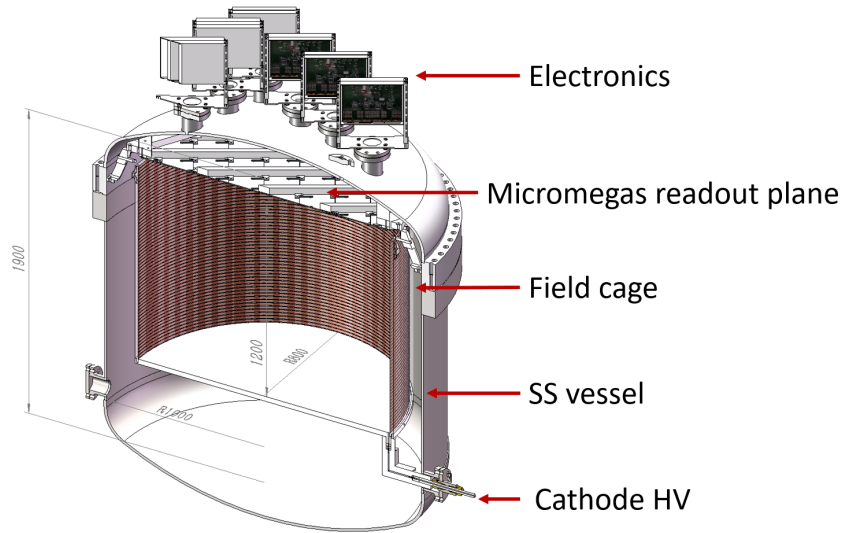


Fig. 1.9 Cross-sectional view of the PandaX-III detector. Figure from [54].

absorbs xenon scintillation light but has the effect of suppressing the diffusion of ionized electrons during drift. Therefore, it is possible to obtain more detailed track information. Additionally, the Penning effect promotes ionization, which can improve energy resolution. They use Micromegas for the ionization signal detection, measuring both energy and track. They achieved the energy resolution of 14 % (FWHM) at 60 keV in the 5 bar Xe-(1%)TMA gas mixture[55]. They developed background reduction techniques using the Kalman filter[56]. They are also developing techniques for background suppression[57] and event vertex reconstruction[58] using machine learning.

1.6.3 AXEL

The AXEL experiment adopts the xenon gas TPC with the EL process for the ionization signal detection. Its concept is similar to the NEXT experiment, but it features a unique cell structure in the ionization detection. Details of the AXEL experiment will be described in the following chapters.

1.7 Outline of this thesis

This paper aims to describe the development of two major technological elements for the AXEL experiment. One is related to the generation of high voltage to form an electric field that drifts ionization electrons in the TPC. The other concerns the development of a signal-background discrimination using track information to improve the efficiency of background elimination.

In Part I, including this chapter, we describe the research background and our $0\nu\beta\beta$ search experiment, the AXEL project. An overview of the AXEL experiment and details of the 180 L prototype detector currently under development are provided in Chapter 2.

In Part II, we describe the development of a Cockcroft-Walton (CW) multiplier for in-situ high voltage generation, data acquisition using the 180 L detector, and performance evaluation. The development of the CW multiplier for high voltage generation is discussed in Chapter 3. Data taking using the 180 L prototype detector equipped with the CW multiplier, the analysis of the obtained data, and the evaluation of the energy resolution and track reconstruction capabilities are provided in Chapter 4.

In Part III, we describe the application of machine learning to track pattern-based signal-background discrimination, and present a sensitivity study for future experiment utilizing this approach. The creation of simulation data used for training machine learning models is described in Chapter 5. The comparison

between the created simulation data and the real data obtained in Chapter 4 is also described, using various metrics. The construction of the machine learning models and their training results are described in Chapter 6. The developed model is applied to both the simulation data and the real data for comparison. The differences between the simulation and real data are also examined in Chapter 6. The achievable sensitivity for $0\nu\beta\beta$ half-life using the developed model is discussed in Chapter 7.

Part IV provides a summary of the study. In Chapter 8, we describe the future improvements derived from the insights gained in this study. Finally, the entire study is summarized in Chapter 9.

Chapter 2

AXEL experiment

AXEL (A Xenon ElectroLuminescence) is a experiment to search for $0\nu\beta\beta$ decay of ^{136}Xe using high pressure xenon gas TPC. In this chapter, we describe the overview of the experiment and roadmap to the large ton-scale detector. We also describe in detail the 180 L prototype detector which is currently developed.

2.1 Overview of the experiment

Figure 2.1 shows a schematic view of the AXEL detector. The AXEL detector is a high-pressure xenon gas TPC that operates by enclosing 10 bar of ^{136}Xe -enriched gas within a pressure vessel. The use of high-pressure xenon gas enables measurements with large-mass $0\nu\beta\beta$ nuclei. It is known that the intrinsic energy resolution based on ionization signals deteriorates when the xenon density exceeds 0.55 g/cm^3 , including in the case of liquid xenon[59]. However, assuming a temperature of 300 K (room temperature)

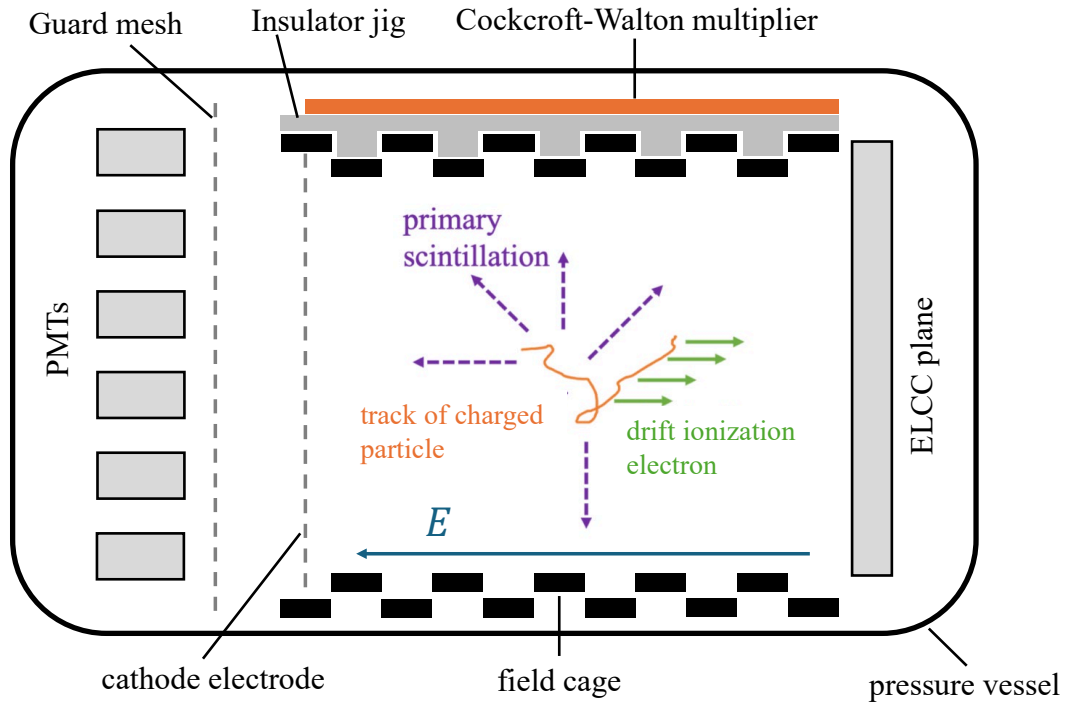


Fig. 2.1 Schematic view of the AXEL 180 L prototype detector.

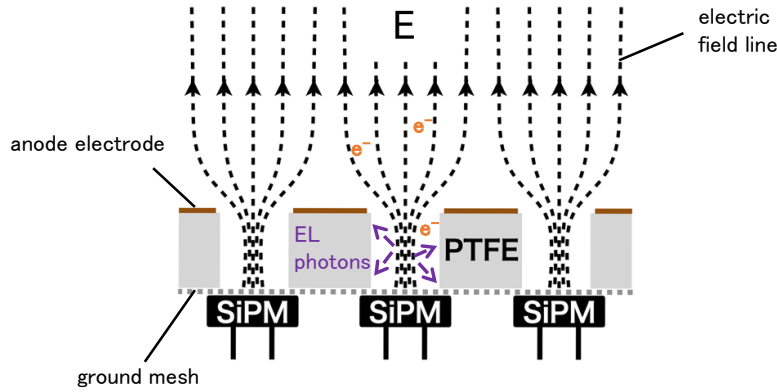


Fig. 2.2 Schematic cross-sectional view of the ELCC.

and a pressure of 10 bar, the xenon gas density is 0.056 g/cm^3 ^{*1}, and thus this issue does not arise. The AXEL experiment aims to conduct a measurement over approximately 10 ton-year of ^{136}Xe , with the goal of probing a effective neutrino mass of 20 meV, which would allow the inverted mass ordering to be nearly excluded. The active volume of the 1 ton detector corresponds to approximately 3 m in diameter and 2.5 m in length. Even with the installation of an external veto detector, the total size remains within approximately 10 m in diameter and 10 m in length, making it feasible within a space comparable to or smaller than that of the KamLAND-ZEN detector.

The detection principle is as follows. As a charged particle passes through the gas volume of the detector, it excites and ionizes xenon atoms along the track. The excited atoms result in primary scintillation light, whose wavelength is around 170 nm, at a timescale of a few tens of nanoseconds[61]. These scintillation photons are detected with vacuum-ultraviolet (VUV) sensitive PMTs installed behind a cathode electrode. This signal is used as the timing of events. Ionization electrons drift along the uniform electric field of 100 V/cm/bar formed by a field cage toward the detection plane called electroluminescence light collection cell (ELCC) plane. The ELCC is a cellular ionization-electron detection device utilizing the EL process. A schematic cross-sectional view of the ELCC is shown in Fig. 2.2. ELCC consists of the TPC anode and a ground mesh electrodes, with a 5mm-thick polytetrafluoroethylene (PTFE) plate between them. The plate has round holes arranged in a hexagonal pattern at 1 cm intervals, as the cells. Ionization electrons, following the electric field lines, enter the cells in the ELCC. The inside the cell is subjected to an electric field of 3 kV/cm/bar by the voltage difference between the anode and the ground mesh, where the field is strong enough to induce EL by the entering ionization electrons. The VUV photons emitted by the EL process in a cell are detected by the cell's photosensor. The number of detected photons at each timing and cell positions is proportional to the energy deposit of the corresponding position on the track. The detail of ELCC used in the 180 L prototype detector is described in Sec. 2.3.2.

The drift velocity of ionization electrons is approximately $1 \text{ mm}/\mu\text{s}$, resulting in a drift time on the order of 1000 μs for a 1-meter drift distance. In contrast, the scintillation light is detected within approximately 100 ns, which is sufficiently short. Therefore, the scintillation light can be used as timing information to determine the position along the drift direction. The EL photons are read out using dedicated electronics (Sec. 2.3.6) at a sampling rate of 5 MS/s. Three-dimensional tracks are reconstructed by combining the two-dimensional position information from the cells with the drift-time-based position along the drift

^{*1} The xenon gas density at the given temperature and pressure was obtained from the NIST Chemistry WebBook[60]

direction.

A Cockcroft-Walton (CW) multiplier is installed inside the pressure vessel to supply high voltage to the field cage. The details of the CW multiplier and the field cage are described in Chap. 3 and Sec. 2.3.4 respectively.

The AXEL experiment is similar to the NEXT experiment in many aspects, but it is characterized by ELCC, which reduces the position-dependent variation in detected photon counts. Additionally, the robust cell structure has the advantage of making it easier to scale up the detector compared to simple parallel mesh electrodes.

2.2 Roadmap of the AXEL experiment

We, the AXEL group, started with a small size detector and have made it larger step by step.

The demonstration of EL amplification using the ELCC was conducted with xenon gas at 1 bar in a small pressure vessel[62]. Signal readout was performed using a single VUV-sensitive PMT for all ELCC channels, resulting an energy resolution of 1.5 % in terms of $0\nu\beta\beta$ Q-value for the 59.5 keV gamma rays from an ^{241}Am source. In addition, a multi-anode PMT was employed to read out signals from individual cells, enabling the evaluation of tracking capabilities.

The first prototype was the 10 L size detector. The purpose of this prototype was to demonstrate the ELCC concept of energy measurement and track reconstruction. The FWHM resolution obtained with the 10 L size prototype detector with 4 bar xenon gas was $(4.00 \pm 0.30) \%$ at 122 keV, which corresponds to 0.9 % to 2.0 % at the $0\nu\beta\beta$ Q value, 2458 keV[63].

The second prototype is the 180 L size detector, which is currently being operated. The development of the 180 L prototype has two phases. In the first phase, the number of ELCC channels was 168 and the energy resolution of $(1.73 \pm 0.07) \%$ at 511 keV was obtained with 4 bar xenon gas, which corresponds to 0.79 % to 1.52 % at the Q value[64]. In the second phase, we upgraded the number of ELCC channels to 672, and obtained the energy resolution of $(0.73 \pm 0.11) \%$ at 1836 keV with 7.6 bar[65].

In these studies, a DC high-voltage power supply was used as the power source to form the drift electric field. But as discussed in Chap. 3, scaling up the detector requires higher voltages, which introduces several challenges. In this study, in-situ voltage generation by a Cockcroft-Walton (CW) multiplier is developed as an alternative and a long-term operation of 40 days is demonstrated.

Based on the know-how gained from these studies, we plan to build a third prototype detector of 1000 L size. The pressure vessel is already placed on the Kamioka Observatory and background study and demonstration of $0\nu\beta\beta$ decay search will be conducted.

After these prototype detectors, a 1-ton size detector will eventually be constructed underground for $0\nu\beta\beta$ decay search. The sensitivity estimation for the 1-ton size detector is discussed in Chap. 7.

In the next section, the detail of the current 180 L size detector is described.

2.3 180 L prototype detector

Figure 2.3 shows a 3D CAD rendering of the 180 L prototype detector. The 180 L prototype detector is installed in a cleanroom in Kyoto university. Inside the pressure vessel, the ELCC plane for detecting ionization electrons and the PMTs for detecting scintillation light are placed facing each other. A quasi-cylindrical field cage is installed between them to establish a drift electric field. The field cage has a flat section that provides space for installing the CW multiplier. Outside the pressure vessel, frontend electronics and the gas system are installed. The following subsections describe each of these components in detail.

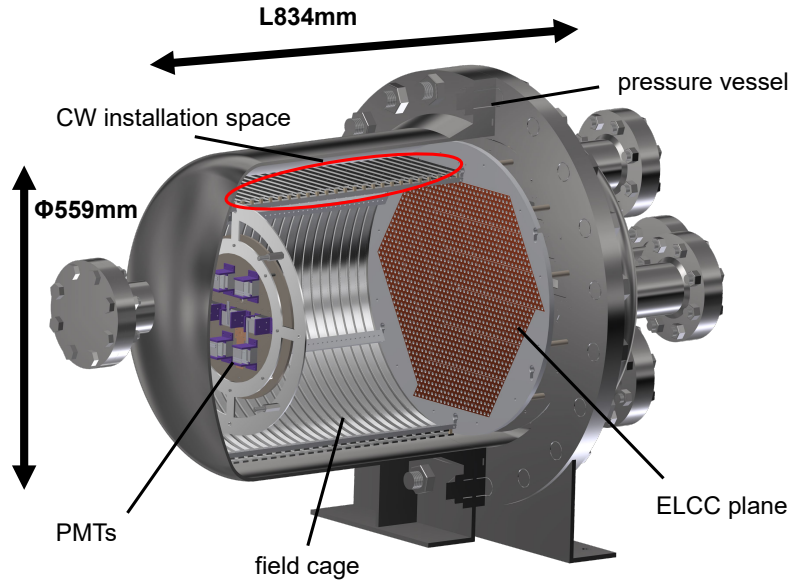


Fig. 2.3 A 3D CAD image of 180 L prototype detector.

2.3.1 Pressure vessel

The pressure vessel of the 180 L prototype detector is shown in Fig. 2.4.

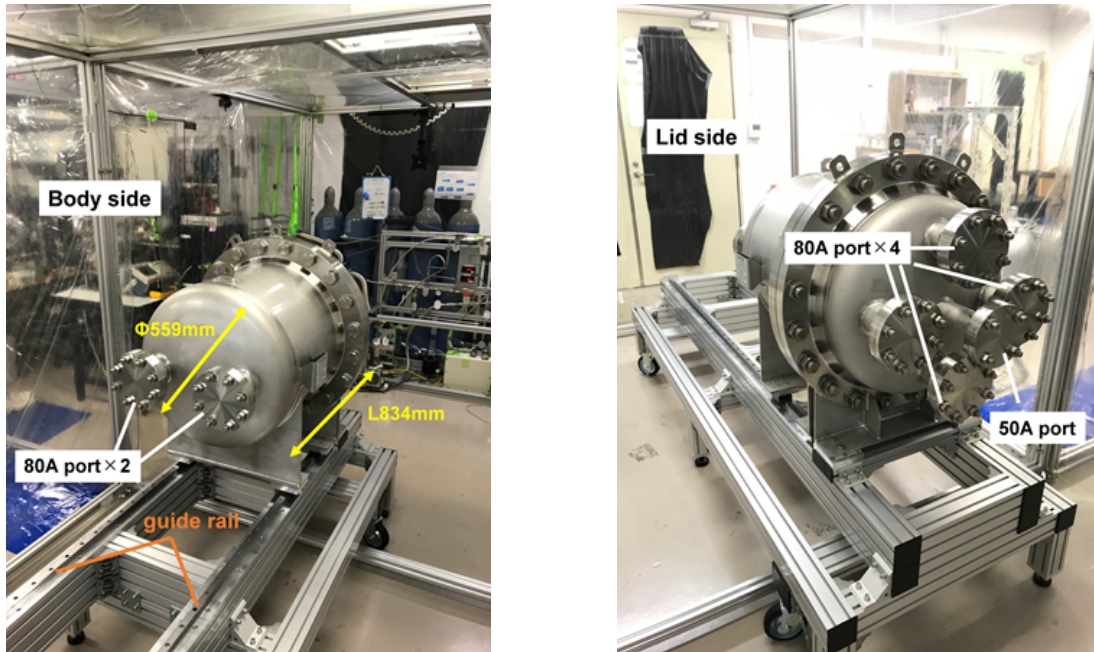


Fig. 2.4 Body side (left) and lid side (right) of the pressure vessel.

The pressure vessel consists of two parts, a lid and a body, which are coupled by flanges of JIS 10K550A. It is made of stainless steel and can withstand pressures up to 0.98 MPa. To enhance the pressure resistance, elliptical heads are used on both the lid and the body sides. The diameter is 559 mm and the length is 836 mm. The vessel has a wall thickness of 6 mm, and the total weight, including the

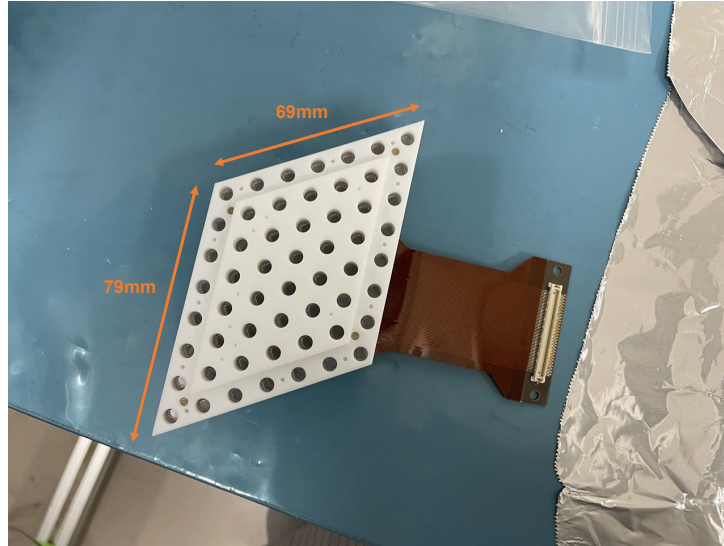


Fig. 2.5 A picture of an ELCC unit without the anode electrode. The PTFE body of the boundary cells is designed with half the standard thickness to fit the cover layer. One unit consists of 56(= 7×8) cells. This unit is laid out to form the ELCC plane.

lid and the body, is 350 kg. The lid side has four JIS 80A ports mounted in a circle and one JIS 50A port mounted in the center. The JIS 50A port is used as a gas line for evacuation, xenon filling, circulation, and extraction. One of the 80A ports on the lid side is used for a feedthrough of flexible printed circuit (FPC) cables to supply voltage to the photosensors of ELCC plane and to read out their signals. Another one of the 80A port is used for cables of high voltage and thermal sensors. One of the 80A ports on the body side is equipped with a feedthrough for the PMT's power and signal lines, while the other is used for gas circulation. The lid side is fixed, while the body slides open along a guide rail.

2.3.2 ELCC

The ELCC plane consists of multiple ELCC units, each composed of 56(= 7×8) cells (Fig. 2.5). Each cell is equipped with a SiPM, a Hamamatsu VUV sensitive S13370-3050CN multi-pixel photon counter (MPPC), mounted behind the ground mesh. The mesh is made of stainless steel wire with a diameter of 0.03 mm and an aperture of 77.8%. To prevent fraying, the non-cell portions of the mesh are sandwiched between two 100 μ m thick perfluoroalkoxy alkane (PFA) films and thermally welded[66]. The unit has a flexible printed cable, which is used to apply voltage to the MPPC and read out signals. The cable is connected to the front-end electronics (Sec. 2.3.6) via a feed-through.

The target value of the field strength in the cell is 3 kV/cm/bar but has not been achieved due to discharges between the anode and the ground electrodes^{*2}. In the previous study[64], discharge between the anode and the ground mesh electrodes occurred at the boundaries of the ELCC units and at the screw holes to fix the ELCC units. As a countermeasure against these discharges, the PTFE plate was divided into two layers and a cover layer was added to cover the gaps between units. Also, the screws fixing the ELCCs were modified so that they do not penetrate the PTFE plate[66]. Figure 2.6 top shows these modifications. However, electrical discharges had still occurred, so additional countermeasures are applied in this study. The two-layer structure in the previous study showed discharges presumably on the surface of the polyimide sheet sandwiched between the layers. Therefore, polyimide sheets are not used, and the two-layer structure is employed only at the perimeter cells of the unit. In addition, to prevent

^{*2} As of 2025, an electric field strength of 3 kV/cm/bar has been achieved through the implementation of a novel discharge mitigation technique for the ELCC.

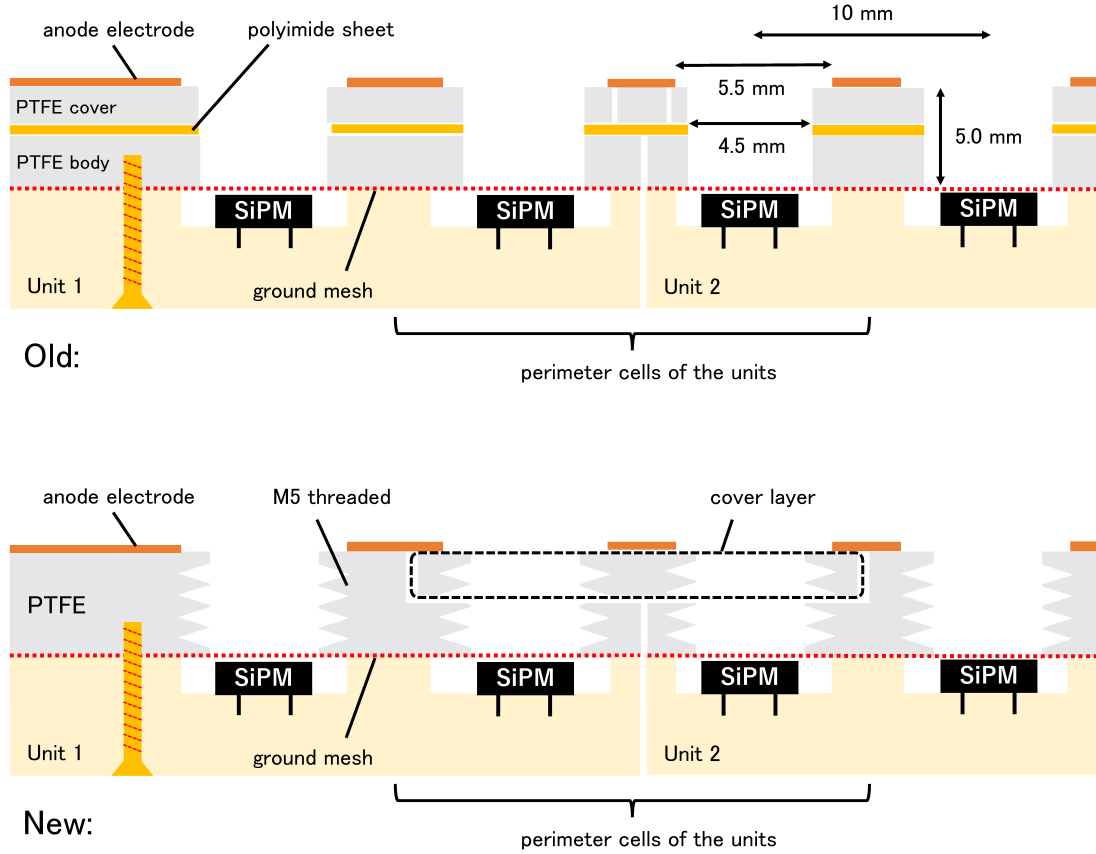


Fig. 2.6 Schematic cross-sectional view of the ELCC structures. The two-layer structure of the previous study[65] is shown on the top, and the tapped hole structure of the amplification section, the current upgrade, is shown on the bottom.

discharge along the holes of the cell, the holes are tapped using JIS M5 threads as shown in Fig. 2.6 bottom.

For the second phase of the 180L prototype detector, 12 ELCC units are used. Figure 2.7 shows the ELCC during its installation into the 180L prototype detector. The ELCC units are arranged in a hexagonal pattern. At each of the four boundaries of a unit, there is a half-thickness column that is shared with adjacent units and covered by a common cover layer to prevent discharges from the anode electrode mounted on top to the ground mesh through the gaps between the units.

The ELCC signals from each unit are transmitted to the front-end electronics (Sec. 2.3.6) located outside the pressure vessel via a 0.9 m-long FPC cable through the feedthrough. The FPC cable is divided into two segments: a 0.5 m inner cable located inside the pressure vessel, and a 0.4 m outer cable fixed at the feedthrough. Each cable is connected via HIROSE ELECTRIC FX11LA-116S-SV and FX11LA-116P-SV connectors mounted at each ends. To prevent warping during the heating process of surface mounting of the connector, a 0.5 mm thick polyimide sheet (Dupont Cirlex) is attached to the FPC around the connector. Out of the 116 pins of the connector, 112 are used for the high voltage supply and signal lines of 56 MPPCs, while the remaining 4 pins on the outermost periphery are used for grounding. The FPC is double-sided, with 100 μm line width and 500 μm pitch, and is routed such that the high-voltage and signal lines for each ELCC channel are arranged back-to-back. The four ground lines are also routed in a back-to-back configuration, with two lines on each side, running along both sides of the cable. The characteristic impedance of the cable, calculated based on the design parameters,

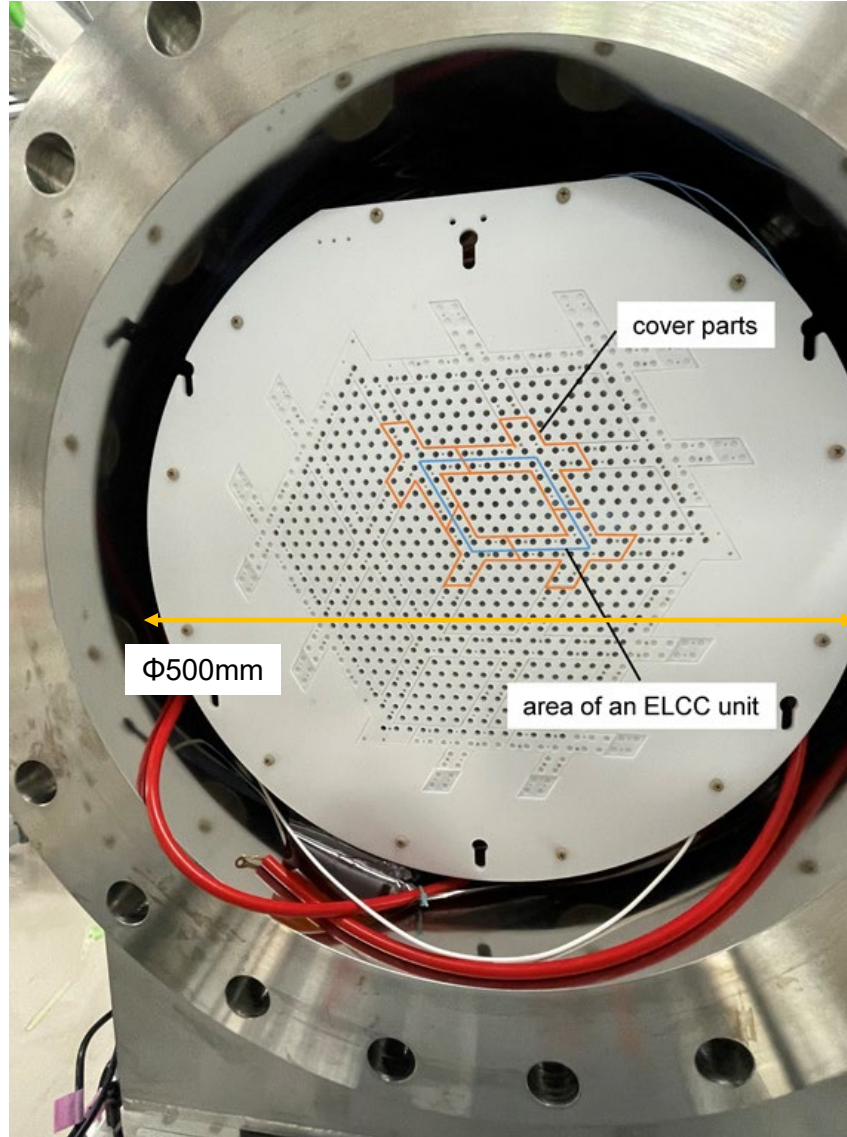


Fig. 2.7 ELCC during its installation into the 180 L prototype detector. A separate cover part is attached to cover the boundary of the unit.

is 70.4Ω .

2.3.3 PMT

PMTs are used for detection of primary scintillation light which determines the timing of the signal. Since the scintillation light of xenon is a VUV light around 170 nm, we used the VUV sensitive PMT, Hamamatsu R8520-406. The PMT has a $20.5 \times 20.5 \text{ mm}^2$ sensitive area and its quantum efficiency for 170 nm wavelength is about 30 %. Considering a photoelectron collection efficiency of 70 %, the resulting photon detection efficiency is 21 %. The PMT can be used under high pressure up to 10 bar, and by not using potting material, outgassing is also minimized. PMT installation in the 180 L prototype detector is shown in Fig 2.8. Seven PMTs are installed in the 180 L prototype detector.

A guard mesh connected to the ground is placed in front of the PMTs to protect the PMTs from the cathode voltage. An aperture of the guard mesh is 67 %. If the distance between the cathode and the guard mesh is too short, unintended EL amplification may occur. The threshold for the EL process is $E/p = 0.83 \text{ kV/cm/bar}$ [67]. As described in Sec. 4.1, a voltage of 34.3 kV was applied to the cathode

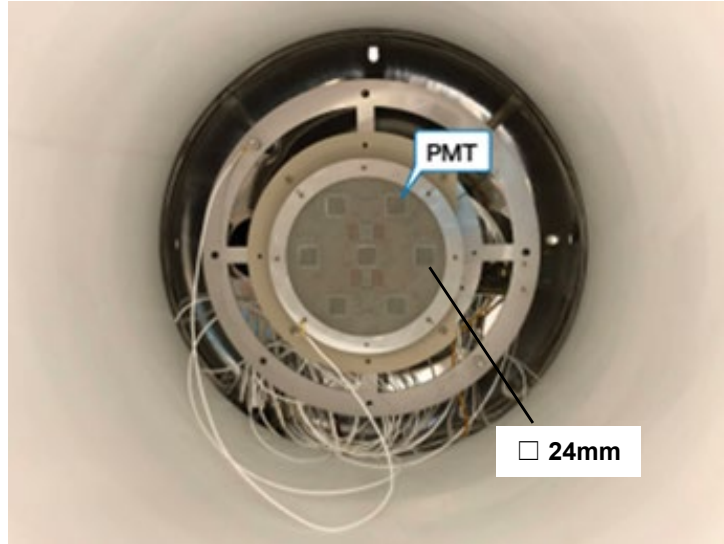


Fig. 2.8 PMT installed in the 180 L prototype detector. A guard mesh connected to the ground is installed in front of the PMTs to protect the PMTs from cathode voltage.

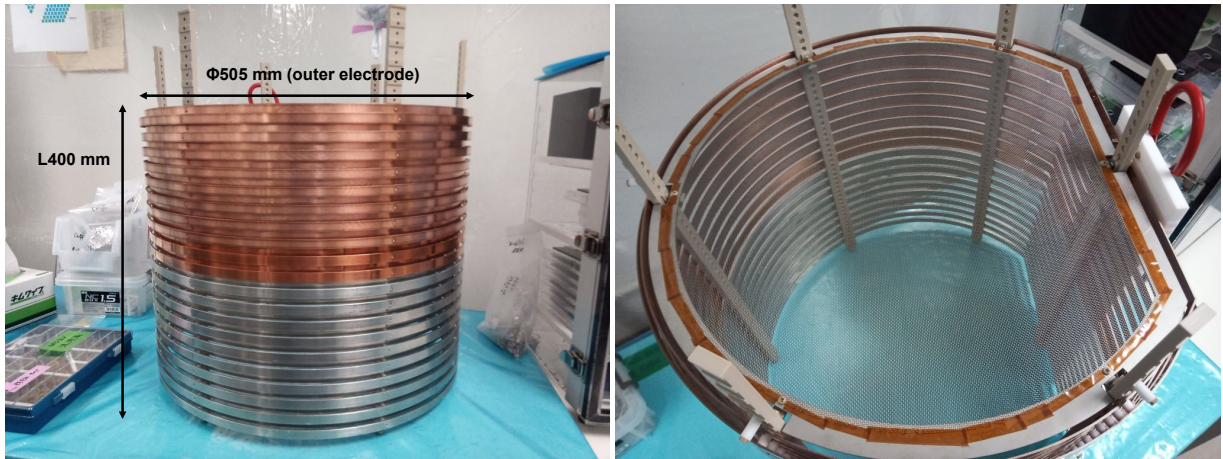


Fig. 2.9 Overall (left) and top view (right) pictures of the field cage.

during the measurement conducted in this study under 6.8 bar of xenon gas. Under these conditions, the reduced electric field between the cathode and the guard mesh was $E/p = 0.32 \text{ kV/cm/bar}$, which is below the EL threshold.

Power supply and signal transmission of the PMTs are carried out using coaxial cables with PTFE jacket via the feedthrough. The signal of PMTs is amplified 100 times by preamplifiers before being sent to the CAEN v1724 digitizer. This 8 channel digitizer operates with a 14-bit resolution and a dynamic range of 2.25 V, sampling at a rate of 100 MHz. The PMT waveforms are transmitted to a PC via an optical link of 80 MB/s transfer rate.

2.3.4 Field cage

The field cage is a frame for creating a uniform electric field that guides ionization electrons to the detection plane. Figure 2.9 shows the assembled field cage. The field cage of the 180 L prototype consists of D-shaped electrodes. Each D-shaped electrode is formed by wrapping a 12 mm-wide band. Two types, differing in the diameter of the rounded part, approximately 500 mm, are arranged alternately with a 10 mm pitch. Two types of electrodes are arranged in an alternating pattern of large and small sizes with

overlapping sections, effectively shielding against the intrusion of a ground potential from the pressure vessel. The electric field strength inside the field cage was simulated, and it was confirmed to be uniform within $\pm 5\%$ over the region up to a radius of 229.3 mm, including the ELCC plane[68]. Each electrode is held in place by fixing it to six pillars made of poly ether ether ketone (PEEK). The flat section of the electrodes allows the CW multiplier to be placed between the pressure vessel and the electrode. On top of the field cage, the cathode made of a mesh is attached. The aperture of the mesh is 71 %, allowing scintillation light to pass through while preventing leakage of the electric field. Forty bands of electrode are used for the field cage in this study. The first 20 bands are made of 3 mm thick aluminum and the rest is made of 1.5 mm thick oxygen-free copper. Aluminum was initially adopted due to its low cost and its ability to reflect VUV light, which was expected to enhance the detection efficiency of primary scintillation light. Although oxygen-free copper does not reflect scintillation light, it was adopted later because it offers advantages such as lower outgassing and reduced radioactive impurities compared to aluminum. A 15 mm thick high-density polyethylene (HDPE) tube is inserted between the field cage and the pressure vessel for insulation. The voltage to the cathode electrode and each stage of the field cage is supplied by a the CW multiplier described in Chap. 3.

2.3.5 Gas system

Gas system has functions of vacuuming air, filling, circulating, and extracting xenon. A schematic diagram of the gas system is shown in Fig. 2.10.

An evacuation pump unit consists of a turbo molecular pump (Osaka Vacuum, TG350FCAB) and a dry scroll pump (Anest Iwata, ISP250C). The exhaust speed of the turbo molecular pump is 330 L/s. A Pirani gauge and a cold cathode gauge are installed upstream of the pump. These vacuum gauges are used not only for monitoring the vacuum level, but also for evaluating the outgassing rate—including potential leaks— by observing the rate of pressure increase after isolating the vacuum pump by closing the valve. Additionally, another Pirani gauge is installed near the pressure vessel to evaluate the outgassing rate of the 180 L prototype detector.

A circulation pump (IBS Inc, MB-601HPAL) with a maximum flow rate of 70 L/min is used to fill and circulate gas. It is also used to recover xenon gas from the pressure vessel into storage cylinders. This circulation pump is a sealed metal bellows pump capable of operating at a maximum discharge pressure of 690 kPaG. The flow rate of the circulation pump is monitored by a mass flow meter (Bronkhorst, F-111CM-40K-AAD-88-K) installed on the downstream side of the pump. The circulation pump requires the downstream pressure to be below about 50 kPaG at startup. If the circulation pump stops due to a power outage or an interlock during circulation, the pressure in the piping may exceed 50 kPaG, making it impossible to restart the circulation pump as it is. Therefore, a buffer chamber with a size of 3 L is installed and, in such situations, the IN-V2 and the CC-V1 valves leading to the 180 L pressure vessel are closed, and the BF-V1 valve is opened to lower the pressure in the circulation pipeline, allowing the circulation pump to be restarted.

During gas circulation, purification is carried out using a molecular sieve (Applied Energy Systems, 250C-V04-I-FP) and a getter (API, API-GETTER-I-RE). The molecular sieve is used to remove impurities other than N_2 , such as O_2 , H_2O , and CO_2 , while the getter is used to remove N_2 . The gas purity is monitored by a dew point meter (Michell Instruments, PURA) installed on the upstream side of the circulation pump, with a measurement range from $-120^\circ C$ to $-40^\circ C$.

Xenon gas is stored in four 47 L cylinders. During gas recovery, xenon gas is transferred from the 180 L pressure vessel to the cylinders using a circulation pump, but about 0.3 bar of xenon gas remains on the upstream side of the pump. This residual gas is recovered by solidifying it using a cylinder immersed in liquid nitrogen. This cylinder is connected to an emergency recovery 47 L cylinder via a 0.2 MPaG

rupture disk to prepare for local pressure increases due to depletion of liquid nitrogen or other causes.

After recovering the xenon gas, the SP-Ar and SP-V1 valves are opened and the argon gas is introduced into the pressure vessel at 1 barA to eliminate the pressure difference between the inside and outside of the pressure vessel before opening the pressure vessel to the atmosphere.

2.3.6 Front-end electronics and data acquisition system

Signals from the MPPCs of each ELCC unit are connected to the front-end electronics (AxFEB)[69] outside the pressure vessel through the FPC cable. Figure 2.11 shows the block diagram of the AxFEB. AxFEB is responsible for both supplying voltage to the MPPCs and reading out signals. A gain of MPPC can be adjusted by a DAC connected to the anode of MPPC that can be fine-tuned for each channel separately from the bias voltage applied to the cathode of the entire MPPC. Since the MPPC signal must be read out by DC coupling to suppress waveform distortion, the amplifier circuit is configured to cancel the baseline variations caused by the DAC[69]. The MPPC signals are converted to digital waveforms by two types of ADCs, ADCL and ADCH. ADCL is 12-bit with a dynamic range of 2 V peak to peak and a sampling rate of 5 MS/s. Before the signal is input to ADCL, it is amplified by a factor of 5 and further shaped by a second-order Sallen-key filter with a time constant of 220 ns. ADCL is used to record EL signals with a time width of about 100 μ s. The number of photons incident on the MPPC is determined from the integral of the waveform. The time structure of the waveform is used both to reconstruct the position along the drift direction of the track and to correct for the non-linearity in the output of the MPPC when a large number of photons is incident in a short period of time. ADCH is 12-bit with a dynamic range of 2 V peak to peak and a sampling rate of 40 MS/s. It is used to measure dark pulses whose width is about tens of ns. Before the signal is input to ADCH, it is amplified by a factor of 165. The ADCH waveform is used to evaluate the magnitude of dark pulses and to equalize the gain of each MPPC via the DAC adjustment. In ADCH, 8 channels are aggregated into one ADC using a multiplexer, and data is acquired while switching.

A block diagram of the data acquisition system is shown in Fig. 2.12. The waveform data from the ELCC and PMT are stored on separate PCs. To issue the trigger, a trigger board, Hadron Universal Logic module (HUL) [70] is used. The 180 L prototype detector uses 12 AxFEBs, each holding waveform data for 56 channels of MPPCs. The HUL receives the sum of the 56-channel MPPC waveforms from each AxFEB via LVDS and calculates the sum of all boards. This sum is then averaged over a set width, and a trigger is issued when a certain threshold is exceeded. The signal sum of the channels set as veto is also sent from the AxFEB to the HUL, and a veto signal is issued when the threshold exceeds. HUL outputs two NIM signals, the send-trigger signal and the send-header signal. The send-header signal is used as the trigger for both digitizers, AxFEBs and v1724. The send-trigger signal is recorded by the digitizer for the PMTs and is used to indicate the trigger timing within the PMT waveforms. The event matching of ELCC waveforms and PMT waveforms is performed during the analysis phase using the timing information. There are two types of triggers for ADCL: the fiducial trigger and the whole trigger, and one type of trigger for ADCH: the high gain trigger, issued by the HUL. The whole trigger is issued when the summed signal height of all channels exceeds a threshold. The fiducial trigger is issued similarly except that the veto channels must have no hit signals. The whole trigger is to collect the data of xenon K_{α} characteristic X-ray (29.68 keV) for calibration, while the fiducial trigger is to accumulate higher energy events contained in the fiducial volume. The use of all channels in the whole trigger is motivated by the calibration method based on xenon K_{α} X-rays[65]. Ionization electrons produced by K_{α} X-ray interactions diffuse during drift and are typically detected across multiple adjacent channels. Therefore, to calibrate channels located at the boundaries of the fiducial region, signals from neighboring channels, including those configured as veto channels, are also required. Since the whole trigger is configured with

a low threshold to take xenon K_α X-rays, it is issued more frequently than the fiducial trigger of interest. Therefore, a downsampling ratio can be set to control the actual frequency at which whole triggers are issued. The high gain trigger is issued periodically during intervals when the ADCL triggers are inactive, in order to acquire ADCH waveforms. However, as described in Chap. 4, temporal variations in photon counts during data acquisition are corrected using xenon K_α , and therefore the ADCH data obtained while data taking are currently not used.

2.3.7 High voltage

Figure 2.13 shows a schematic diagram of the high-voltage system in the 180 L prototype detector. A MATSUSADA PRECISION Inc. HFR10-20N is used as the power supply for applying the anode voltage. Its rated voltage is -20 kV and the rated current is $500\text{ }\mu\text{A}$. The cathode power supply was initially MATSUSADA PRECISION Inc. HFR10-30N, which is replaced by the CW multiplier in this study. Details of the CW multiplier are described in Chap. 3. The anode power supply can control the output voltage with a control voltage of $0\text{--}5\text{ V}$, and it allows for smooth variation of the output voltage using a potentiometer for resistance division. The output voltage of the anode power supply is delivered to the ELCC anode inside the pressure vessel through a feedthrough using a single-wire cable with a silicone-rubber jacket. When the design voltage of 3 kV/cm/bar EL electric field and 100 V/cm/bar drift electric field are applied at 8 bar, the current flowing in the anode power supply is $56\text{ }\mu\text{A}$. A protection circuit with diodes and resistors are configured to prevent the anode power supply from being damaged by reverse currents associated with discharge. In addition, an interlock is implemented to stop the anode power supply when an excessive current flows through it, which is regarded as a discharge. The anode power supply has a monitor voltage with 5 V output for rated voltage and current. The current value is monitored by the Graphtec GL820 data logger, and an interlock is triggered when a current over $450\text{ }\mu\text{A}$ flows.



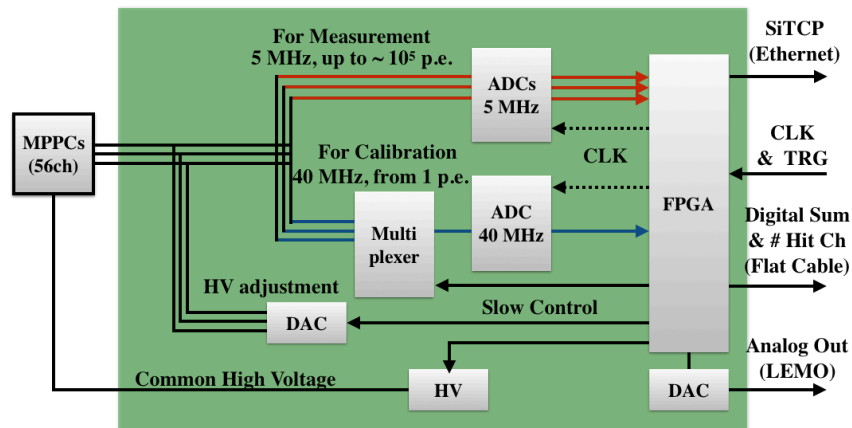


Fig. 2.11 The block diagram of the AxFEB. Figure from [69]

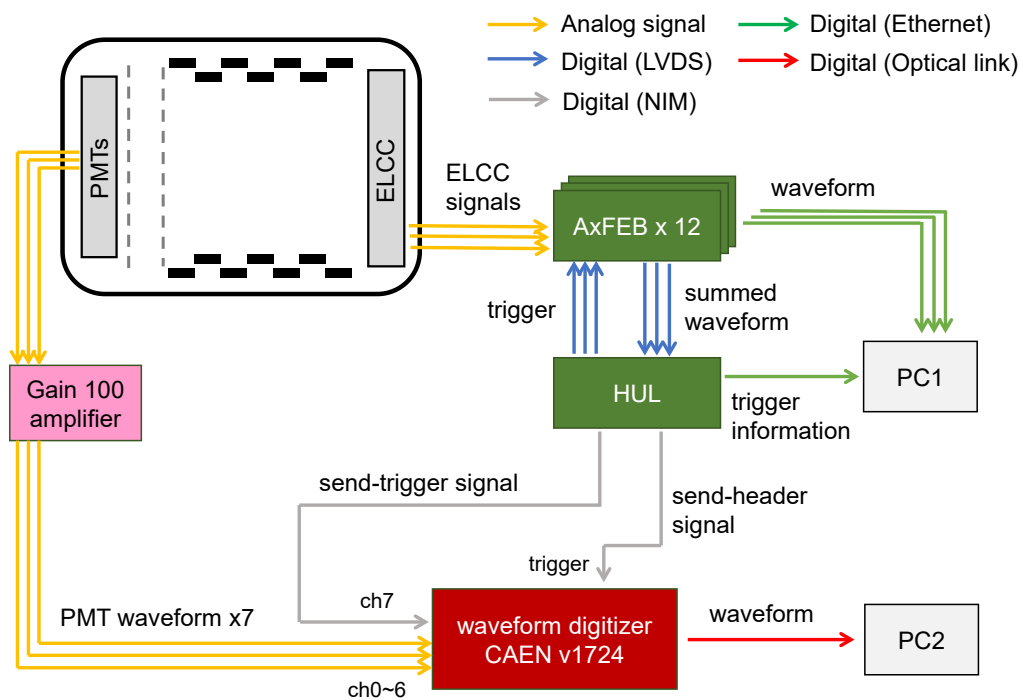


Fig. 2.12 Block diagram of the data acquisition system.

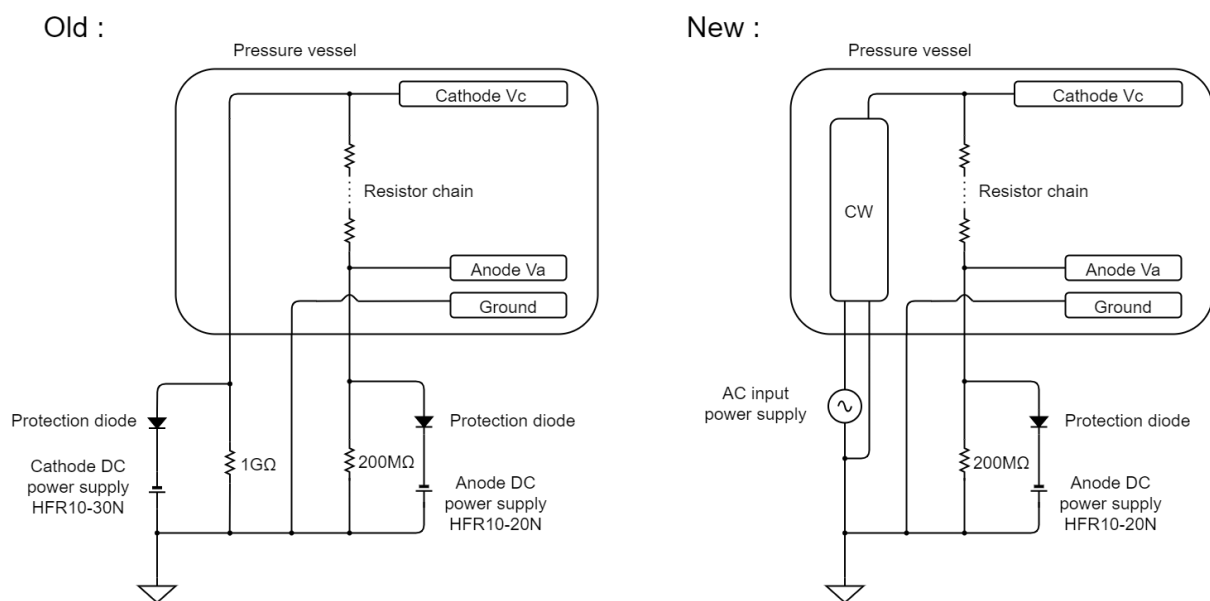


Fig. 2.13 Schematic diagram of the high voltage system in the 180 L prototype detector.

Part II

In-situ high voltage generation with
Cockcroft-Walton multiplier

Chapter 3

Design and performance of the Cockcroft-Walton multiplier

A high voltage is applied to the field cage to generate the electric field necessary for drifting ionization electrons. In case of the 180L prototype and the 1000L detector, the target voltage is -44.8 kV and -79.8 kV at 8 bar respectively. To feed such a high voltage from outside the pressure vessel, high voltage feedthroughs compatible with high pressures are needed. For example, in the NEXT-White detector, specially developed low-radioactivity feedthroughs are used to apply potentials to both the cathode electrode and the EL amplification gate electrode[71]. Another approach is to introduce a relatively low voltage from outside the pressure vessel and boost it inside the pressure vessel. The CW multiplier[72] can be used to convert a low voltage AC input to a high voltage DC output. This approach was proposed for liquid argon TPC's[73], but has not been realized in actual operation. One difficulty comes from the large baseline variation on the signal line caused by the AC input, which makes signal readout and analysis difficult. If the AC input is turned off after charging the capacitors in the CW multiplier, this baseline shift disappears, but it is difficult to monitor that the voltage is properly maintained.

In the AXEL experiment, the EL photons induced by the ionization electrons are used as signals. Since the ionization signal is converted to light to be read out and the light signal is amplified with quite high efficiency by photon counters, it is highly resistant to electronic noise. We have developed a CW multiplier to supply high voltage to the AXEL detector and installed to the 180L prototype detector. In this chapter, the developed CW multiplier is detailed. The content of this chapter and the following one is structured based on the previously published paper[74].

3.1 Design of the CW multiplier

A schematic diagram of the CW multiplier is shown in Fig.3.1. The CW multiplier is composed of multiple stages of a basic circuit, consisting of two capacitors and two diodes, connected in series. The AC voltage causes the potential of the capacitor on the upper side of the figure to oscillate. The rectifying action of the diode causes a charge to accumulate in the capacitor on the lower side of the figure, resulting in

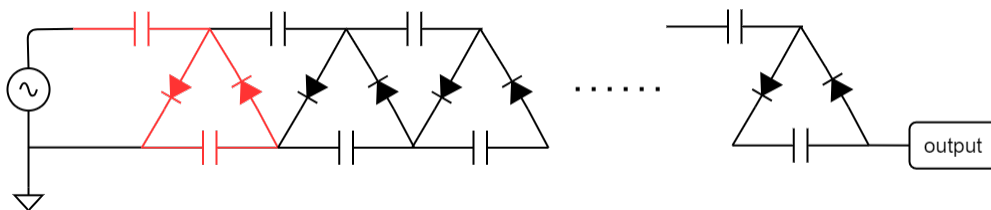


Fig. 3.1 Schematic diagram of the CW multiplier. The red section corresponds to one stage of the circuit.

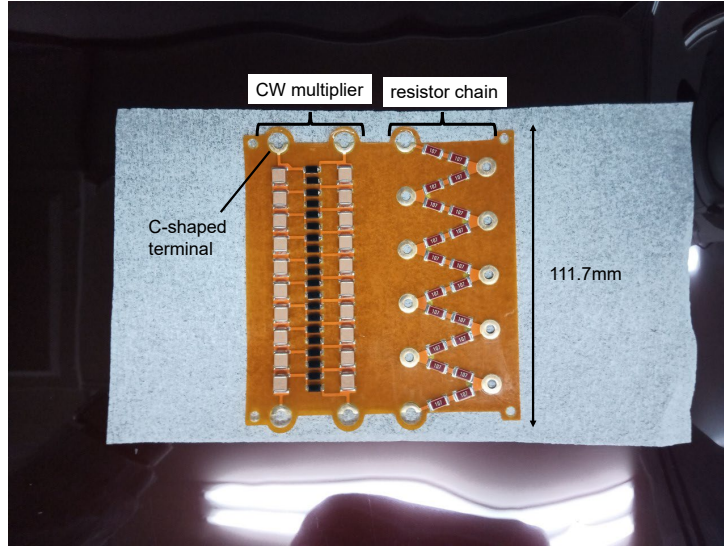


Fig. 3.2 CW multiplier and resistor chain board implemented as a flexible printed circuit.

a DC high voltage at the output. The ideal output voltage V of an N -stage CW multiplier is given as $V = 2NU$, where U is the amplitude of the AC input voltage, but the actual output voltage of a CW multiplier is lower than the ideal voltage due to the parasitic capacitance of diodes and load resistance[75].

To install the CW multiplier inside the pressure vessel, there are several constraints regarding size and material. The multiplier has to be installed in a narrow space between the HDPE tube and the field cage. In case of the 180 L prototype, the width is limited to about 20 cm and height to 3 cm. The length is also limited to 40 cm to fit to the length of the field cage. Since the voltage required for the drift electric field is proportional to the drift length, the voltage must be increased within the length of the field cage.

The CW multiplier has to be composed of low outgassing materials. This is because electronegative impurities such as oxygen capture ionization electrons, leading to a reduction in EL photon counts and a consequent degradation of energy resolution. The typical outgassing rate of the prototype detector was $1.23 \times 10^{-4} \text{ Pa m}^3/\text{s}$ [65], with which xenon gas purity was kept sufficiently high by continuous purification using the molecular sieve and the getter. The outgassing of the CW multiplier should be well below this rate.

To achieve compact dimensions and minimize outgassing, a flexible printed circuit (FPC) board populated with surface-mount devices was adopted. FPC made in this study consists of an $18 \mu\text{m}$ -thick copper electrode sandwiched between a $25 \mu\text{m}$ -thick polyimide base and coverlay, NIKKAN INDUSTRIES F-30V and CISV respectively. One board is 111.7 mm long and contains 10 CW stages, and a resistor chain to evenly divide the electrical potential to be applied to the electrodes of the field cage. The FPC board has C-shaped terminals at both ends, allowing them to be connected to each other (Fig. 3.2). This enables the construction of a CW circuit with more stages. The right side of the board in Fig. 3.2 shows a resistor chain, which divides the CW output voltage at the top and supplies it to the TPC electrodes spaced at 1 cm intervals. Each stage has a resistance of $200 \text{ M}\Omega$, resulting in a total of $2 \text{ G}\Omega$ per board.

The deviation from the ideal output voltage of the CW multiplier can be reduced by using large capacitors. However, capacitors with higher voltage ratings tend to have smaller capacitance. The circuit was designed to operate at 2 kV peak-to-peak input and capacitors with a voltage rating of 2 kV and as large a capacitance as possible were selected. Consequently, a Knowles Syfer's $0.1 \mu\text{F}$ 2220 chip size ceramic capacitor was adopted. Diodes are also subjected to a maximum input peak-to-peak voltage, those with a voltage rating of 2 kV and fast response characteristics, Micro Commercial Components FM2000GP with a reverse recovery time of 500 ns and typical junction capacitance of 8 pF were selected.

While a higher resistance value helps suppress deviations from the ideal output of the CW multiplier, the resistance must also be low enough to supply the induced charge generated by the drift of ionization electrons in the TPC. This trade-off is evaluated based on the time constant of the circuit formed by the top-stage capacitor and the resistor chain. When the rated voltage is applied to the top-stage capacitor, the stored charge is $0.1 \mu\text{F} \times 2 \text{ kV} = 0.2 \text{ mC}$. A reaction on the order of MeV typically generates approximately 10^5 ionization electrons. Assuming the maximum operation rating as 10 kHz, motivated by the rate of the 180 L detector with sources, if the time constant is set to 1000 seconds, the total induced charge during that period would be approximately 0.1 % of the capacitor's stored charge and sufficiently small. The corresponding resistance value is $10 \text{ G}\Omega$. Given a drift length of 40 cm, this translates to $250 \text{ M}\Omega$ per centimeter. Therefore, Bourns Inc. CHV2512-JW-107ELF $100 \text{ M}\Omega$ resistors were selected, with two used per stage.

Since each component has a voltage rating of 2 kV, achieving the required voltage of -44.8 kV for the drift length of 40 cm would ideally require only 23 stages in the CW multiplier. However, in practice, the circuit was designed with 40 stages, taking into account the operation at a AC voltage lower than the rated voltage of the circuit elements and the deviation from the ideal gain. As a result, each board was equipped with 10 stages, with one stage per centimeter.

3.2 High voltage generation

The actual output voltage of the CW multiplier is lower than the ideal output voltage[75]. One of the causes is that a portion of the current used to charge the CW multiplier is consumed by the charging and discharging of the parasitic capacitances of components such as diodes and input cables. The effect from diodes is described by the following equation,

$$V_{\text{CW}} = 2NU \left[\frac{\tanh(2N(C_d/C)^{1/2})}{2N(C_d/C)^{1/2}} \right] \quad (3.1)$$

where N is the number of stages, U is the amplitude of the AC input voltage, C_d is the parasitic capacitance of the diode, and C is the capacitance used in the CW multiplier. This expression is valid under the assumption that the load resistance is sufficiently large. As shown in Eq. 3.1, the effect of the parasitic capacitance is independent of the input voltage and frequency. Another factor is the continuous removal of charge from the CW multiplier due to the load resistance. When the parasitic capacitance C_d is small, this effect can be described by the following equation,

$$V_{\text{CW}} = 2NU \left\{ 1 - \frac{1}{fRC} \left[\frac{2}{3}N^2 + \frac{1}{2}(1 - \epsilon_1)N + \left(\frac{1}{2}\epsilon_1 - \frac{1}{6} \right) \right] \right\} \quad (3.2)$$

where R is the resistance per stage, f is the frequency of the AC input. ϵ_1 represents the fraction of the AC input cycle during which the smoothing column capacitors are charged. From Eq. 3.2, it can be seen that the reduction in output voltage due to the load resistance can be mitigated by increasing the frequency of the AC input voltage.

The frequency dependence of the output voltage was measured with the setup shown in Fig. 3.3. Seven FPC boards and a PTFE plate to secure them were prepared. To prevent electrical discharge, two grooves, oriented perpendicular to the electric field, were machined at 1 cm intervals in front, back, and side of the PTFE plate. Before making the frequency measurement, a test was performed with an input peak-to-peak voltage of 1800 V, near the voltage rating, to the CW multiplier in a same setup that in Fig. 3.3. Because discharges happened on two boards during these tests, frequency measurements were only made on the remaining five boards. To feed the AC input power to the CW multiplier, a sine wave from a

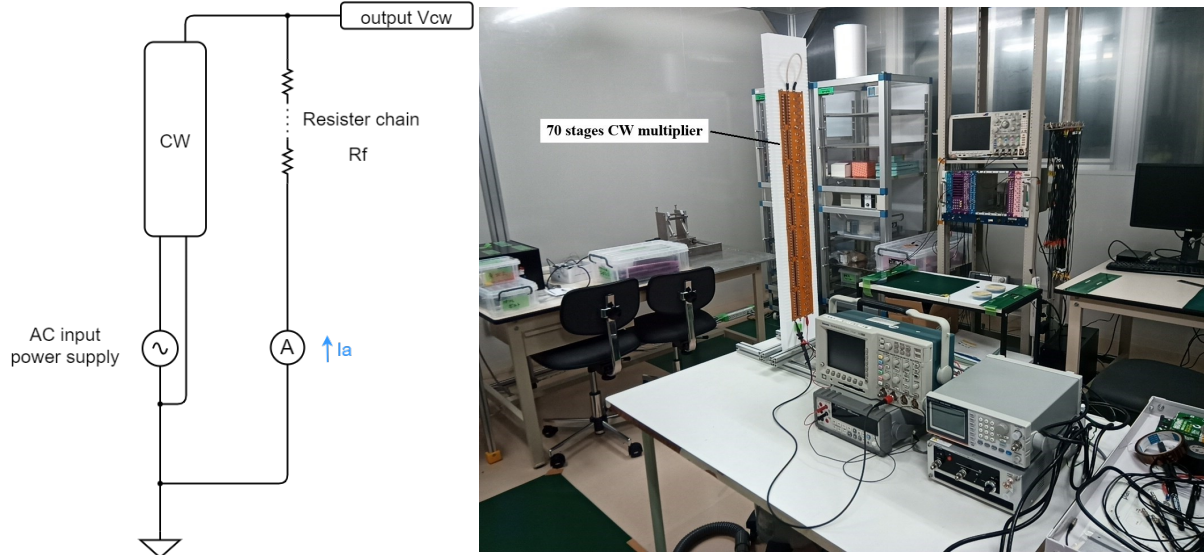


Fig. 3.3 Schematic diagram (left) and a picture (right) of the measurement of the CW multiplier high voltage generation.

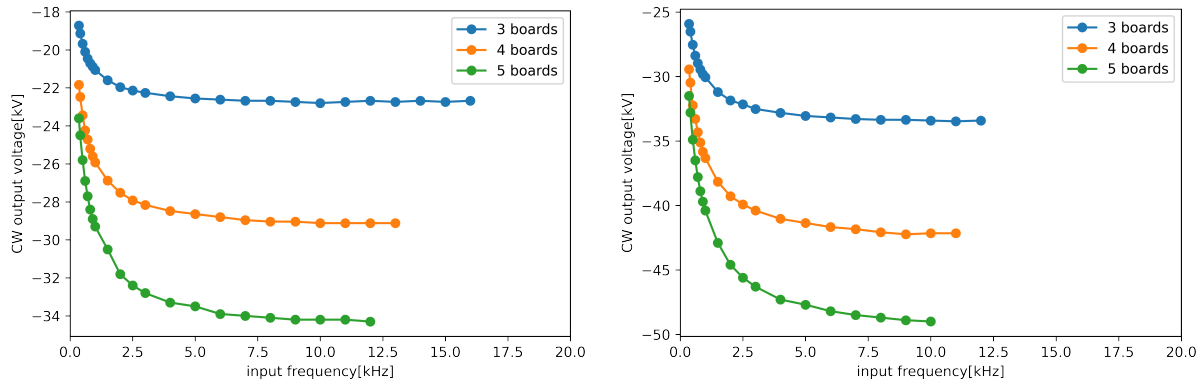


Fig. 3.4 CW output voltages as a function of input frequency at 800 V peak-to-peak input(left) and at 1200 V(right).

function generator GW Instek AFG-2005v was amplified by an AC amplifier Matsusada HJOPS-2B10. The maximum output voltage and current of the AC amplifier is ± 2 kV, 10 mA respectively, and the slew rate is 150 V/ μ s and bandwidth is 18 kHz (-3 dB). The output of the CW multiplier was connected to the ground through a resistor chain R_f on the boards and an ampere meter Agilent Technologies U3401A. The output voltage V_{CW} is obtained from the measured current value I_a as $V_{CW} = -I_a \times R_f$. Here the resistance value R_f is 2 G Ω per 10 stages of the resistor chain on the boards.

The measurement results are shown in Fig. 3.4. The input peak-to-peak voltages are 800 V and 1200 V. Data points missing at high frequencies are due to instabilities in the amplifier output, which caused oscillations. As more boards were added, the frequency at which that happened became progressively lower. This is thought to be due to the increase in capacitive load with more boards, which leads to the amplifier's output current reaching its limit.^{*1}

^{*1} Operation at higher frequencies has been confirmed to be feasible with the use of higher power audio amplifiers and transformers, and implementation is planned for future development.

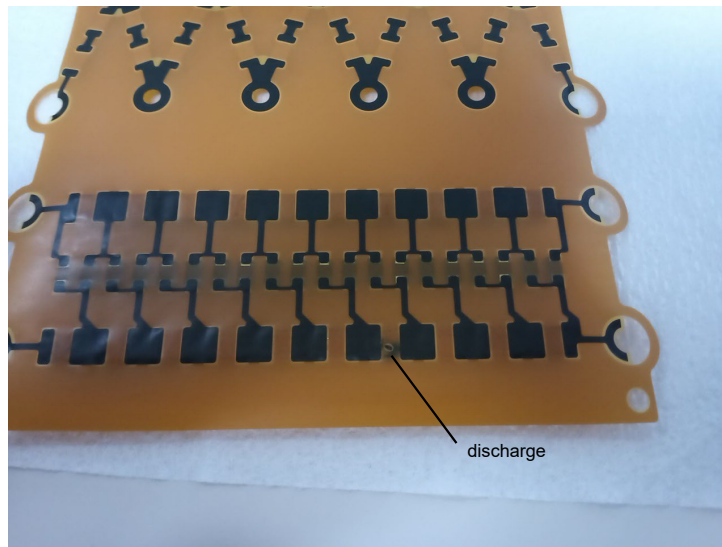


Fig. 3.5 Trace of surface discharge on the FPC of CW circuit. A discharge occurred across a 3 mm gap between the pads on the FPC board, resulting in the board being burned.

3.3 CW surface insulation

During the test operation of the CW multiplier at a peak-to-peak voltage of 1800 V, a surface discharge occurred on the capacitor and FPC surface, and the output voltage of the CW multiplier deteriorated (Fig. 3.5). To allow for higher voltage in- and output, the circuit was coated with the methyl silicone resin, Shin-Etsu Chemical Co. Ltd. KR-251, which was applied as follows:

- (1) Ultrasonically clean the circuit in ethanol for 15 minutes.
- (2) Submerge the circuit in a vat containing KR-251 and run it through ultrasonic cleaning for 15 minutes to defoam.
- (3) Degas the circuit under vacuum for one hour while submerged in KR-251.
- (4) Lift the circuit out of the KR-251 bath and dry it under vacuum for 4 hours.

Since the connecting electrodes are also coated in this method, the contacts were restored by gently scraping off the coating with tweezers after drying. A test was conducted in air using a setup with a single coated FPC board with a setup described in Sec.3.2. With an input of 5 kHz and 1900 V peak-to-peak, no discharge was observed for 10 minutes, confirming an improvement in withstand voltage. Even with the above procedure, air bubbles around the circuit elements could not be completely avoided (Fig. 3.6). To estimate the impact of outgassing from these bubbles on gas purity, the outgassing rate of a coated FPC was measured.

3.4 Outgassing rate of the coated CW circuit board

A vacuum test was performed using an FPC board that was cut into about one-fourth and then coated. The cut board was placed in a NW50 pipe connected to a turbo molecular pump, evacuated for about 5 days. The outgassing rate was estimated from the pressure rise as a function of time after closing off the vacuum connection to the pump. The pressure changed from 1.7×10^{-4} Pa to 91 Pa over about 14 hours. The estimated outgassing rate is 8.9×10^{-7} Pa m³/s including leaks in the vacuum system. Since the 180 L prototype detector uses four circuit boards, the outgassing is estimated to be 1.4×10^{-5} Pa m³/s,



Fig. 3.6 Example of air bubbles remaining in gaps between circuit elements.



Fig. 3.7 CW and resistor chain are mounted on a PTFE jig and installed on the field cage. The field cage and anode electrode are connected via a resistor FPD4D200MOHMF (black plate seen against the white PTFE at the end of the field cage).

which is about an order of magnitude less than the operational outgassing rate $1.2 \times 10^{-4} \text{ Pa m}^3/\text{s}$ for the prototype without CW multiplier[65].

3.5 Monitoring of the HV output voltage for the 180 L detector

To actually supply HV to the 180 L detector, four CW multiplier and resistor chain boards were serially connected and fixed on a PTFE plate, which was installed on the flattened outside of the field cage as shown in Fig. 3.7. To prevent electrical discharges, two grooves, oriented perpendicular to the drift direction, were machined at 1 cm intervals on the front, back, and side surfaces of the PTFE plate. A stainless steel screw connected each electrode of the field cage to its corresponding stage in the resistor chain. Although it is preferable to use coaxial cables to suppress AC pickup on the MPPC signal line, a silicon-rubber insulated single-wire cable was used for the input to the CW multiplier because the use of coaxial cables caused oscillation of the AC amplifier due to the capacitive load of the cable.*²

*² The integration of a high-power audio amplifier and a transformer as the power supply enables voltage application through coaxial cables.

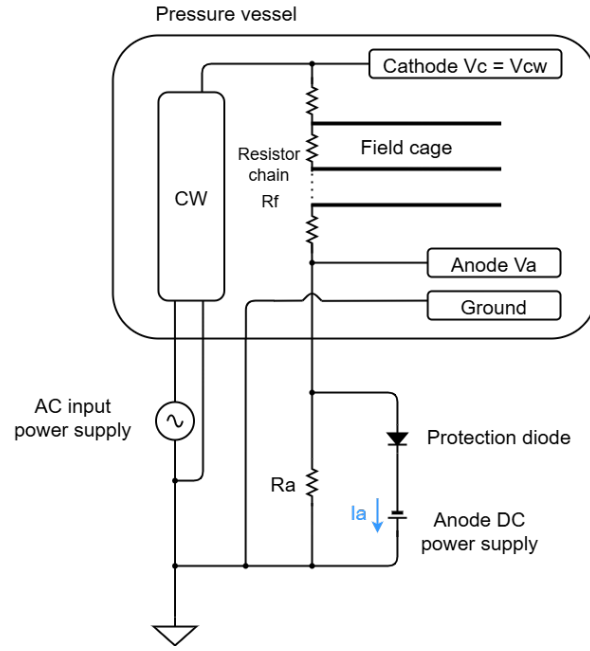


Fig. 3.8 Schematic diagram of the HV supply to the 180 L prototype detector

Unlike the setup shown in Fig. 3.3, the 180 L detector has the end of the resistor chain connected to the anode electrode, making it impossible to directly monitor the output voltage of the CW multiplier. To estimate the output voltage of the CW multiplier, the monitoring voltages of the anode HV power supply were used. Figure 3.8 shows the schematic diagram of the HV supplies to the 180 L prototype detector. V_a and V_c means the anode and cathode voltage, respectively. The output voltage V_{CW} of the CW multiplier is given by,

$$V_{CW} = V_c = \left(1 + \frac{R_f}{R_a}\right)V_a + R_f I_a, \quad (3.3)$$

Here, R_a is the additional load resistor (Fig. 3.8) for the anode power supply, which is $200 \text{ M}\Omega$, and R_f is the resistance of the entire TPC resistance chain, which is $8.1 \text{ G}\Omega$. The anode voltage V_a and current I_a are - as indicated above - taken from the monitoring outputs of the anode power supply.

3.6 Countermeasure for surface discharge on HDPE tube

Before acquiring data with the 180 L detector, a high voltage test of the CW multiplier was conducted in 6.9 bar xenon gas, resulting in a discharge with a loud noise before reaching 30 kV. The discharge was captured by a USB camera module ELP-USB13MAF-V75J and the motion detection program “motion”[76]. The camera was installed adjacent to the PMT, facing the ELCC surface. The discharge image stored by motion is shown in Fig. 3.9. Upon opening the pressure vessel, discharge marks were observed on the surfaces of the HDPE tube and the FPC board. As shown in Fig. 3.10, on the FPC board, the marks are located between resistors. From this, it was inferred that the discharge occurred along the path from the field cage electrode to the pressure vessel via the surface of the HDPE tube (Fig. 3.11).

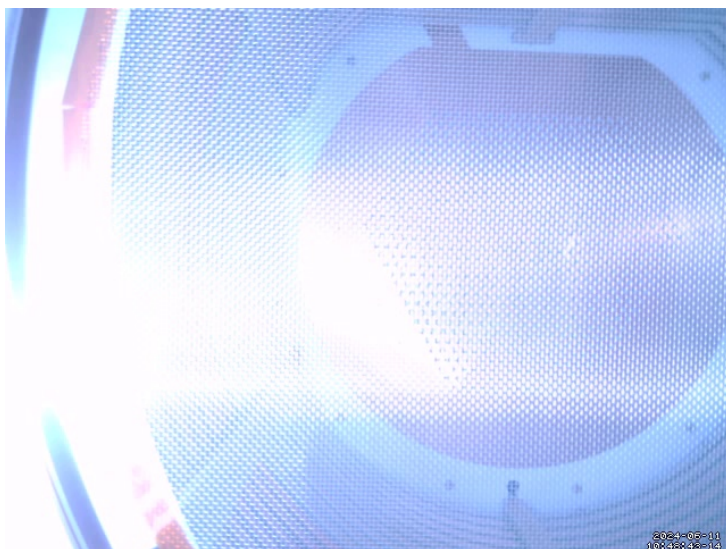


Fig. 3.9 The discharge image stored by motion. The cathode mesh electrode is visible in the foreground, with the ELCC surface in the background, and the discharge occurs near the cathode.

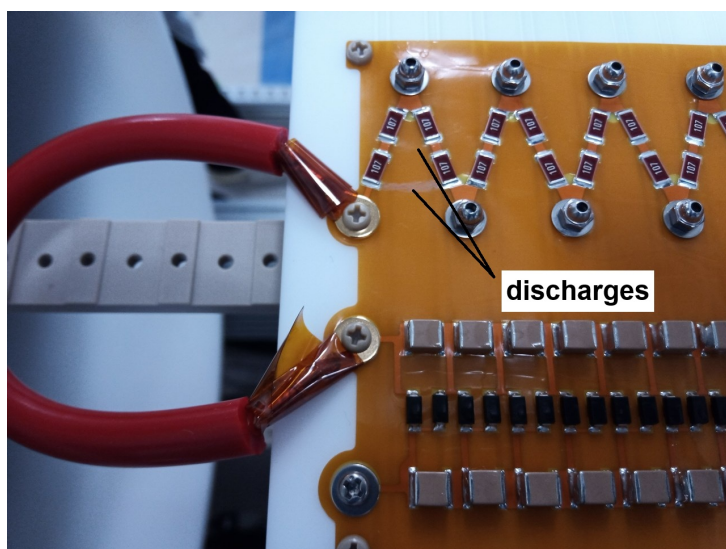


Fig. 3.10 The discharge on FPC board. Traces of surface discharge bypassing between the resistor chain are visible.

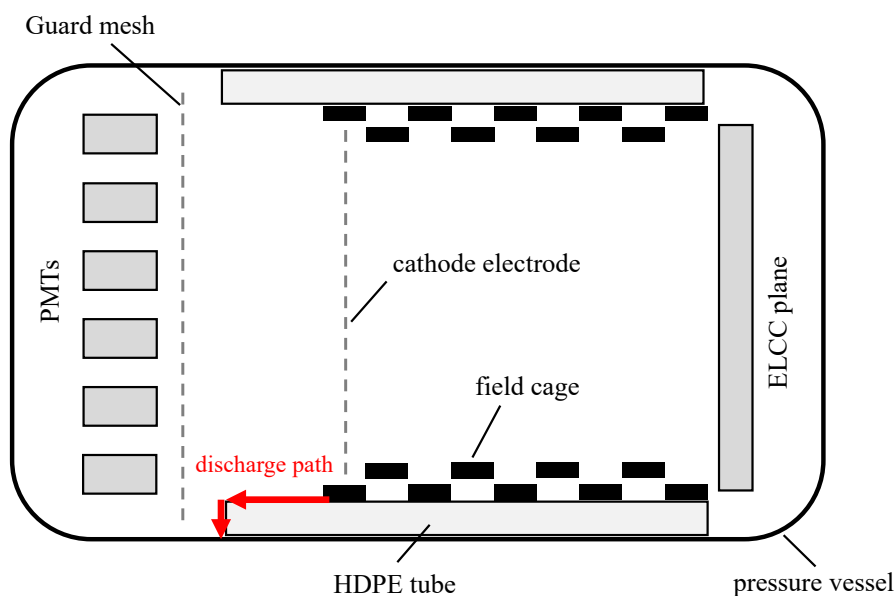


Fig. 3.11 The presumed discharge path. It was considered that the discharge traveled from the field cage electrode along the surface of the HDPE tube towards the pressure vessel.



Fig. 3.12 Grooves carved into the HDPE tube as a countermeasure against the discharge. In the background, the PMTs and the guard mesh can be seen.

The surface discharge on the resistor chain was thought to be due to the instantaneous large potential fluctuations caused by the electrode discharge, resulting in dielectric breakdown of the FPC surface. As a countermeasure against the discharge, 12 grooves with a depth of approximately 1 mm were created at about 1 cm intervals on the inside surface of the HDPE tube using a router (Fig. 3.12). In addition, improvements were made to the voltage input to the CW multiplier. The boosting of the CW multiplier is performed by gradually increasing the amplitude of the sine wave from the function generator on the input side. Previously, this was manually controlled using a rotary knob, but a program to automatically boost at a constant rate using serial communication was developed. To prevent damage to the CW circuit from continuous discharges, an interlock was implemented to detect discharges and stop the input to the CW multiplier. The discharges were detected by the motion program and saved as video files on a PC. The folder in which the videos were created was monitored and when a new file was detected, an alarm signal from the data logger was triggered to stop the input to the CW multiplier. Since the condition for issuing the interlock is the generation of a video file, the input to the CW multiplier is stopped even if the discharge is at the ELCC. However, during the data taking starting from July 8 to September 13 in 2024, there were only two interlocks, so there were no practical problems in acquiring the source data.

Chapter 4

Data taking with CW multiplier, analysis, and performance evaluation

In this chapter, we describe the data acquisition using the 180 L detector equipped with the CW multiplier, the analysis of the obtained data, and the evaluation of the detector's performance.

4.1 Data taking with CW multiplier

In preparation for data acquisition using the 180 L detector equipped with the CW multiplier, the pressure vessel was closed and evacuated for about 5 days before introducing xenon gas. The pressure reached 4.5×10^{-2} Pa and the outgassing rate was 1.5×10^{-4} Pa m³/s. Then, xenon gas was filled to about 6.8 barA. The active volume of the detector is 19.6 L, which corresponds to a xenon mass of approximately 0.73 kg*¹ assuming a xenon gas temperature and pressure of approximately 300 K and 6.8 bar during the measurement. Before applying HV, gas purification was performed for 9 days. The detector operating voltages were then adjusted while taking data on a trial basis. After another 9 days to optimize detector operating conditions, data was taken from July 8, 2024 to September 13, 2024, with intervals in between. The data used in the analysis were acquired for a total of 40 days. The measurement period was determined with the aim of acquiring approximately 1000 photoelectric absorption events of 2615 keV gamma rays from the source described below, in order to compare them with simulation data.

At 6.8 bar xenon gas, an anode voltage of -10.2 kV and a cathode voltage of -38.1 kV are required to achieve an EL amplification field of 3 kV/cm/bar and a drift field of 100 V/cm/bar. As a discharge happened around -38 kV during detector optimization, data-taking was conducted at the 90 % of the design value, -34.3 kV, which gives a 90 V/cm/bar drift electric field. The ELCC is designed to achieve a 100 % collection efficiency of ionization electrons when the ratio between the EL amplification field and the drift electric field is 3 : 0.1. Therefore, the applied voltage to the ELCC was also set at the 90 % of the design value, too. The applied voltages were hence -9.18 kV for the anode and -34.3 kV for the cathode.

A test measurement was conducted to ensure that the ELCC signal is not affected by pickup noise from the AC input of the CW multiplier. A sample waveform of an ELCC channel without signal hits is shown in Fig 4.1. The applied AC frequency of the CW multiplier is 6 kHz and the corresponding cycle is 167 μ s. The baseline is stable within 1 to 2 ADC counts. The distribution of the baseline standard deviations σ_{bl} for all the ELCC channels is shown in Fig. 4.2. The mean standard deviation is 0.46, while it is 0.45 without high voltage applied. The effect of baseline fluctuations within one ADC count on the energy resolution is small enough compared to other factors. Therefore, the effect of AC pickup due to

*¹ The xenon gas density at the given temperature and pressure was obtained from the NIST Chemistry WebBook[60]

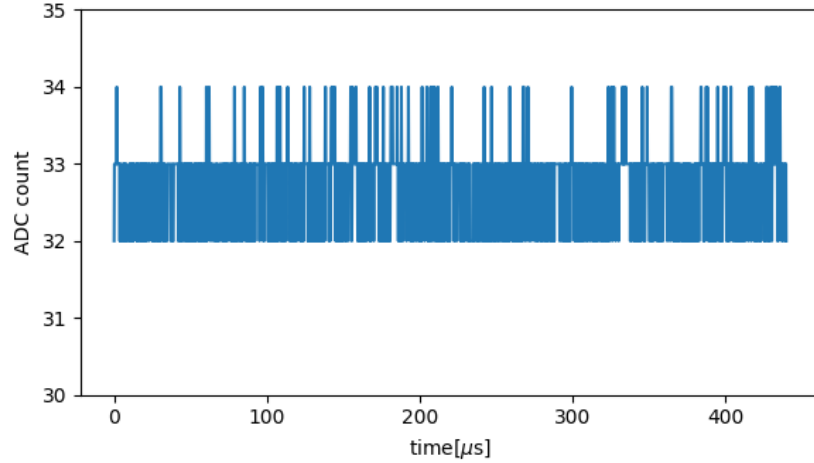


Fig. 4.1 Example waveform of an ELCC channel without signal hits. Sampling rate is 5×10^6 samples/s ($0.2 \mu\text{s}/\text{sample}$). One ADC count corresponds to 0.488 mV . The input frequency of 6 kHz for the CW multiplier corresponds to $167 \mu\text{s}$.

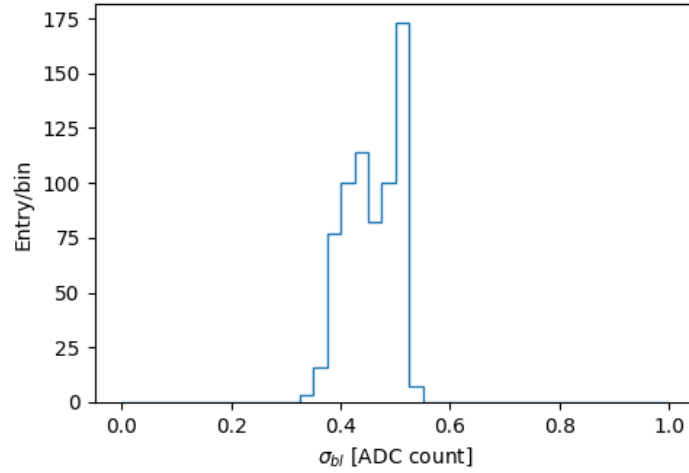


Fig. 4.2 Distribution of the standard deviation of the baseline of the ELCC channels.

the CW multiplier is not a problem.

To evaluate the actual energy resolution and track reconstruction performance at the Q value of xenon $0\nu\beta\beta$ decay, measurements were conducted using a gamma-ray source. Thorium-doped tungsten rods were used as a gamma-ray source. This is a commercial product for welding and contains 2% thorium and hence contains ^{208}Tl in the thorium series. A ^{208}Tl nucleus emits a gamma ray of 2615 keV , which is close to the energy of β -rays from the ^{136}Xe $0\nu\beta\beta$, 2458 keV . The weight of the thorium-doped tungsten rods used was 1 kg at the beginning and later doubled. The intensity is 80 kBq and 160 kBq respectively. The rods are attached to the upper side of the cylindrical part of the pressure vessel (Fig. 4.3).

The outer cells of the ELCC plane were set as veto to ensure that the energy of the particle track is fully captured without leakage. Two dead channels and two high dark current channels, shown in Fig. 4.4, were found. In the dead channels, although the AxFEB was supplying voltage normally, no signal was observed from the MPPC. During installation checks, reconnecting the FPC cable occasionally restored the signal, suggesting that the issue was likely due to poor contact in the connector or cable. The high dark current channels cannot be used for energy measurements but can still be utilized as veto. Then, the high dark current channels were also added to the veto, as were the channels surrounding dead channels.

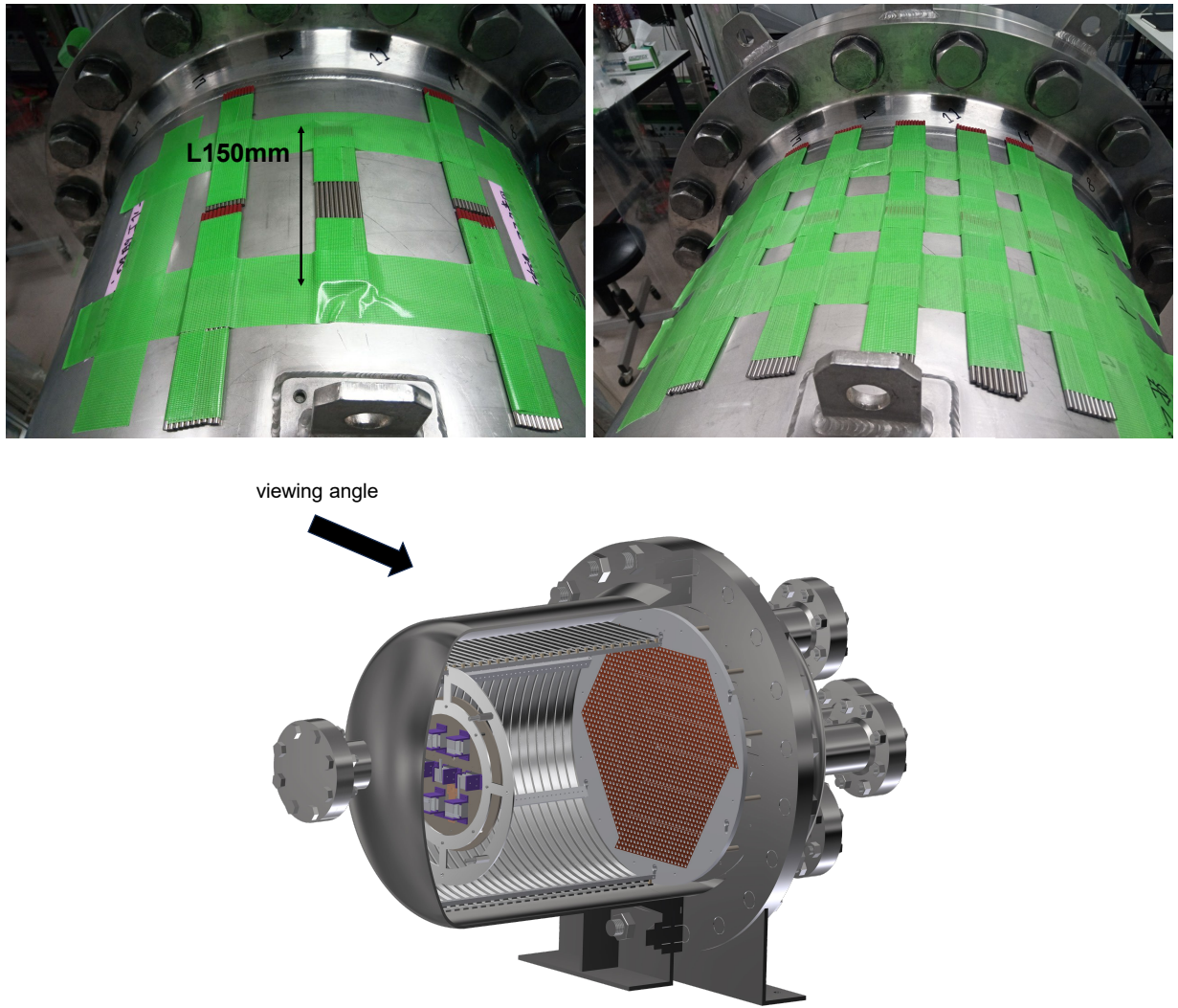


Fig. 4.3 Installation of 1 kg (left) and 2 kg (right) of thorium-doped tungsten rods on the pressure vessel. Ten rods (200 g) are bound together with curing tape and attached to the upper side of the cylindrical part of the pressure vessel. The sectional view of the 180 L prototype detector on the bottom figure shows the viewing angle.

Various quantities monitored during the data taking period are shown in Fig. 4.5. The sampling interval is 33 seconds. Moisture content was below the lower limit of the meter's range for most of the measurement period. This means that the moisture content was less than 0.05 ppm. The cathode voltage has a temporal variation of about 1 kV and can drop several kV momentarily. An example case on July 20 is shown in Fig. 4.6. The cathode voltage fluctuated by 0.4 kV around 34.2 kV during this period. As the cathode voltage is calculated from the monitored anode current and voltage (Eq. 3.3), the sharp cathode voltage drop of 7 kV observed in one bin near noon in Fig. 4.6 likely is an artifact from a discharge in the CW or ELCC. The frequency of such a momentary voltage drop was less than a few times per day and did not stop the data taking. In another incident on July 27, the cathode voltage dropped to about 2 kV. This was caused by an interlock triggered by a discharge in the ELCC, which stopped the input to the CW multiplier; however, the cathode voltage did not fall to 0 kV. At that time, the anode voltage of 9.18 kV remained applied, and it is presumed that a reverse current through the diode in the CW multiplier caused a voltage drop across the resistor chain, resulting in a residual voltage of 2 kV at the cathode.

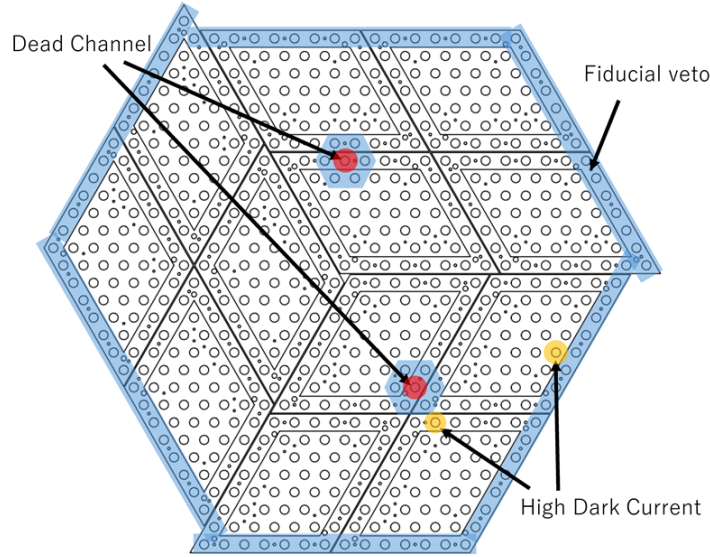


Fig. 4.4 Configuration of veto channels. Dead channels are indicated in red, while high dark current channels are shown in yellow. The colored channels including these channels were assigned to veto.

As described in Sec. 2.3.6, there are two types of triggers for ADCL: a fiducial trigger for acquiring high energy events such as 2.615 MeV gamma-rays from the thorium-doped tungsten rods in this measurement, and a whole trigger for collecting calibration data of xenon K_α X-rays. In this study, the threshold of the fiducial trigger was set to about 500 keV and the whole trigger to slightly above the baseline. The fiducial trigger rate at this threshold was 10.5 Hz. To reduce data size, the whole trigger was set to be issued only once per 50 000 times. The measurement conditions are summarized in Tab. 4.1.

Table 4.1 Summary of the measurement conditions.

Gas pressure	(6.77 ± 0.04) bar
Anode voltage	(9.200 ± 0.003) kV
Cathode voltage	(34.1 ± 1.3) kV
EL field	(2.72 ± 0.02) kV/cm/bar
Drift field	(89.9 ± 4.8) V/cm/bar
Gas temperature	$27.6^\circ\text{C} - 30.0^\circ\text{C}$
Total measurement time	39.8 days
Average fiducial trigger rate	10.5 Hz

AxFEB is acquired ELCC waveforms with a time width of 440 μs , including 40 μs before the trigger, and for PMT signals, the waveform digitizer stored PMT waveforms with a window of 900 μs with pre-trigger region of 95 % of the recording window. The ELCC waveforms and the PMT waveforms obtained from a single trigger constitute data for one event.

4.2 Analysis

Here, the method for estimating the photon count of each event from the obtained data is briefly described. The analysis method was based on the previous study; for details, refer to [65]. In the following, the origin is defined at the center of the ELCC plane. The z-axis is taken along the direction from the anode to the

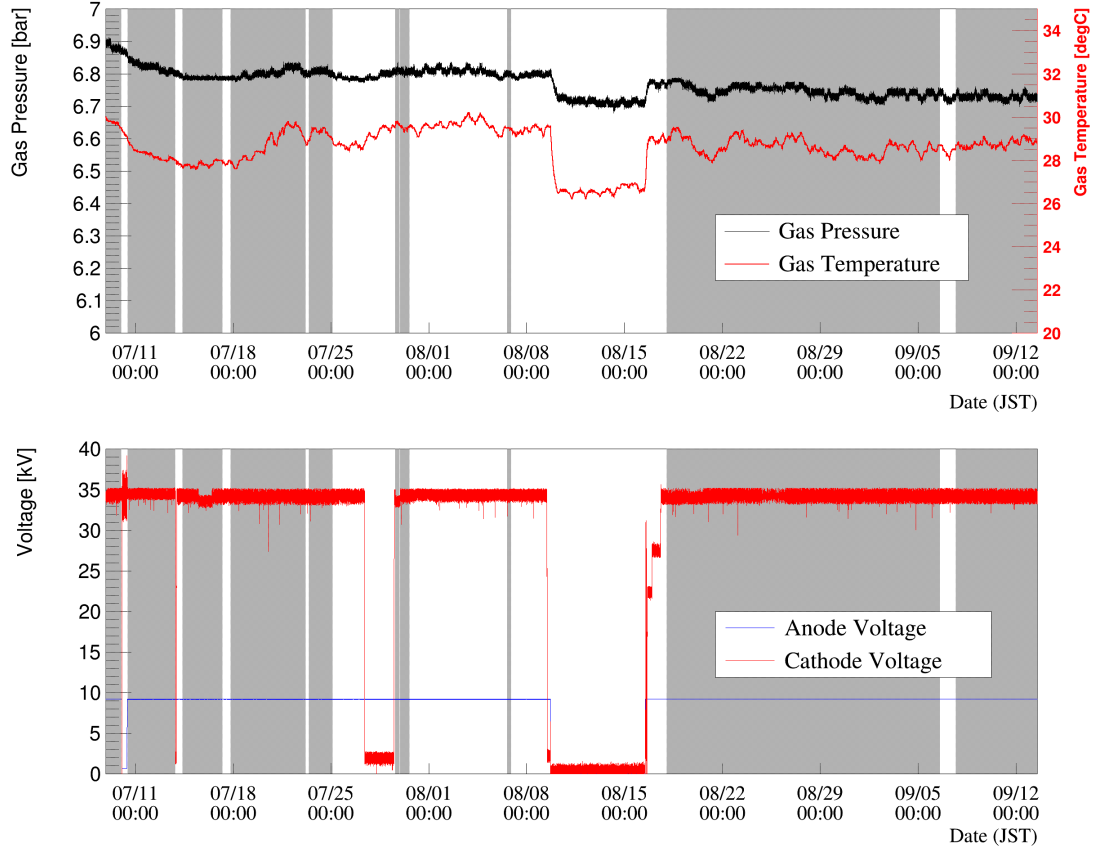


Fig. 4.5 Trends of the gas conditions (upper), and absolute value of high voltages (lower). The gray-shaded areas are data-taking periods. The drops in the anode and cathode voltage from August 9 to 16 correspond to manual shutdowns, while other voltage drops (including drops on July 10) correspond to discharges.

cathode, the y-axis points vertically upward, and the x-axis is defined horizontally to form a right-handed coordinate system.

4.2.1 ELCC waveform analysis

In ELCC, EL photons emitted by drifted ionization electrons is measured individually for each channel using MPPCs and recorded as waveform data. From the waveform of each ELCC channel, hits are identified with a certain threshold from the baseline. Photon counts of each hit are calculated using the MPPC gain. Since MPPC pixels have a recovery time after the charge is released by photon detection to restore the bias voltage, MPPCs have a non-linear output for high incident light intensity. The MPPC we use has 3600 pixels and a recovery time of approximately 70 ns. In comparison, the instantaneous maximum photon count of a $0\nu\beta\beta$ event can reach approximately 20 000 photons/(MPPC · μ s)[66], which is a non-negligible number relative to the pixel count of the MPPC. Therefore, this MPPC non-linearity is corrected by following equation using the recovery time which is measured for each MPPC[64]:

$$N_{\text{cor}} = \frac{N_{\text{obs}}}{1 - \frac{\tau}{\Delta t \cdot N_{\text{pixel}}} N_{\text{obs}}} \quad (4.1)$$

where N_{obs} and N_{cor} are the photon counts before and after the correction, τ is the recovery time, Δt is the time width of 200 ns corresponding to the sampling rate of the ELCC signals. $N_{\text{pixel}} = 3600$ is the number of pixels of the MPPC.

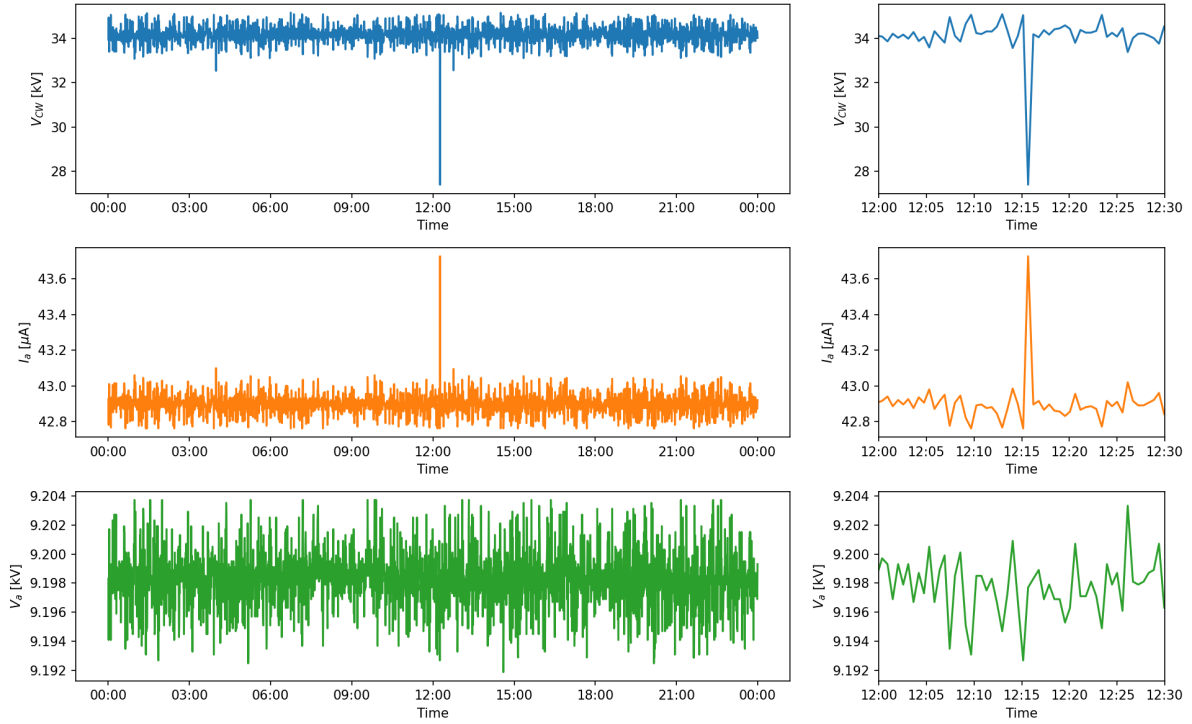


Fig. 4.6 Zoom up views of the trend of the cathode voltage (top), anode current (middle), and anode voltage (bottom) on July 20.

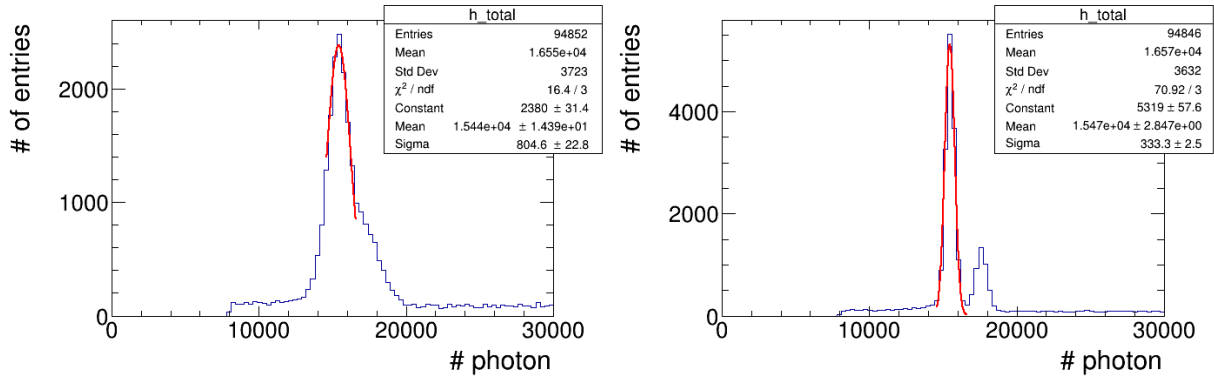


Fig. 4.7 Photon count spectrum of the xenon K_{α} characteristic X-ray before and after the correction.

From the hits, clusters are formed by grouping those that are spatially and temporally adjacent. Photon counts of each cluster or each event are obtained by summing up those of hits after the gain of the EL process in each ELCC cell (hereafter, EL gain) is corrected for. The EL gain is defined as the average detected photon count for one ionization electron. The EL gains are different cell-by-cell due to small difference of ELCC structure (machining inaccuracy, placement inaccuracy, surface condition difference etc.) and photon detection efficiency differences of MPPCs. To determine the correction factor, clusters by xenon K_{α} characteristic X-ray detected throughout the measurement are used. Figure 4.7 shows an example of the photon count spectrum of the xenon K_{α} characteristic X-ray before and after the correction. After the correction, it is evident that the peaks of the K_{α} (29.68 keV) and the K_{β} (33.62 keV) characteristic X-rays are clearly separated.

The average EL gain in this measurement was found to be 11.5, which is smaller than the 12.5 in the previous study[65]. This is presumably because the number of photons reaching the MPPC decreased due

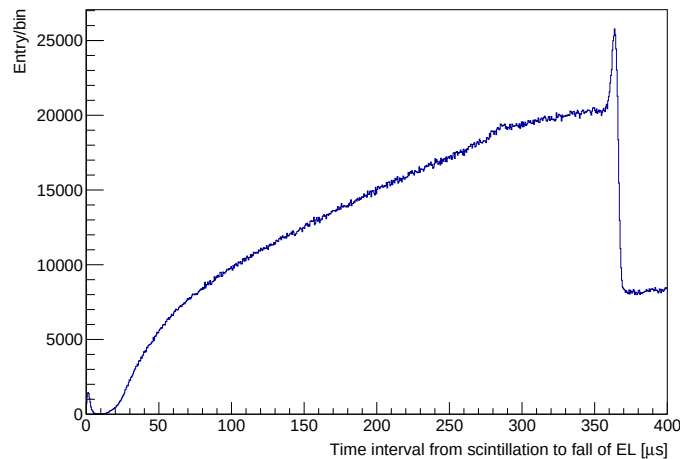


Fig. 4.8 Distribution of the time intervals between scintillation and the end of the ELCC signal. The peak at $363\ \mu\text{s}$ is formed by the events across the cathode plane.

to the change of the geometry of ELCC cell; 4.5 mm open hole in the previous study to 4.1 mm tapped hole in this study (Sec. 2.3.2). Therefore, alternative methods to mitigate discharges in the cells will be employed in the future.

4.2.2 PMT waveform analysis

To select primary scintillation photon, hits are identified from the PMT waveform with a certain threshold from the baseline. Since EL light from ELCC cells as well as primary scintillation light makes hits, those with a width of less than 400 ns and at least $1\ \mu\text{s}$ away from other hits are selected as primary scintillation candidates. We refer to hits of these candidates as “scintillation-like hits”, and other hits as “EL-like hits”. Of these scintillation-like hits, when two or more PMTs has hits within same 100 ns time window, they are considered as primary scintillation light hit clusters. This criterion is based on the fact that the signal from primary scintillation light typically a single photon per PMT, making it necessary to distinguish them from randomly occurring PMT dark pulses. Events with multiple clusters are rejected in this study, as it is unclear which cluster is truly associated with the event timing. The distribution of the time intervals between the selected primary scintillation candidate and the end of the ELCC signal is shown in Fig. 4.8. The peak at $363\ \mu\text{s}$ is considered to be formed by events crossing the cathode, $z = 40\ \text{cm}$. From this, the drift velocity of ionization electrons was derived to be $1.10\ \text{mm}/\mu\text{s}$. The z -positions of ELCC hits are reconstructed using this drift velocity. As can be seen in Fig. 4.8, more events were detected at larger time intervals (i.e., longer drift length). This is likely because, for events with longer drift length, the scintillation light is generated closer to the PMTs, resulting in improved detection efficiency due to a larger solid angle coverage. Entries with time interval exceeding $363\ \mu\text{s}$ are presumably events in which EL light induced photoelectrons on the cathode mesh, which then drift and are subsequently detected by ELCC. A small peak is also observed around $0\ \mu\text{s}$. This likely reflects events in which the primary scintillation light was not detected, and the EL light from the ELCC was mistakenly identified as primary scintillation light by the PMTs.

There are additional selections to avoid timing mismatches. Events in which multiple send-trigger signals are accidentally detected are excluded, as they can cause timing mismatches depending on which signal corresponds to the actual onset of the ELCC event. In addition, there were events in which EL-like hits were observed in the region where the scintillation-like hits were expected. Since EL-like hits are set of narrow hits, misidentifying some of them as scintillation-like hits can lead to timing mismatches.

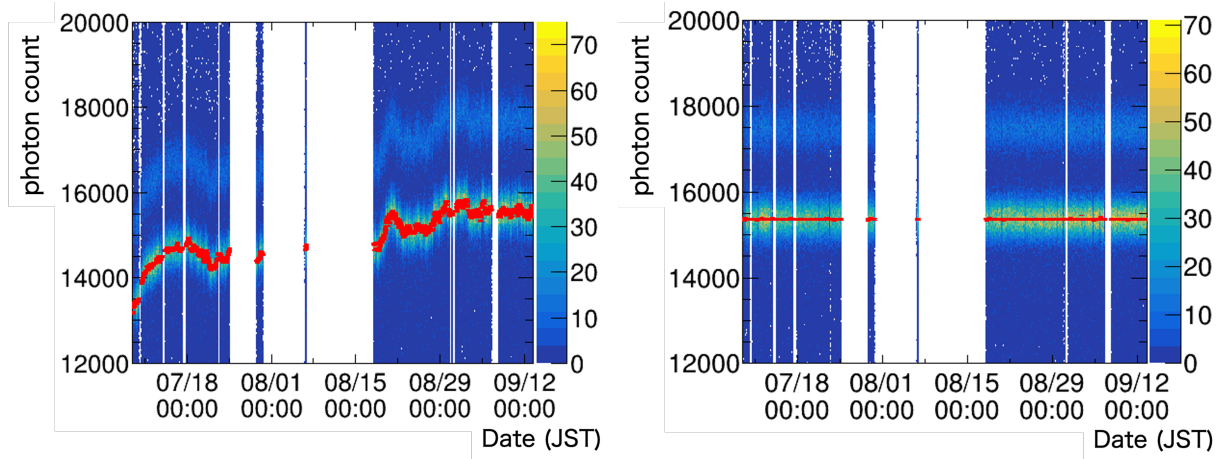


Fig. 4.9 Temporal variation of the photon count of xenon K_{α} characteristic X-rays before (left) and after (right) correction. The energy peak of the xenon K_{α} is determined by performing a Gaussian fit to the histogram of the photon counts. Red lines represent the peak position and its fitting error.

These events were also removed.

4.2.3 Fiducial volume cut and overall corrections

The fiducial volume cut is applied to reject events with hits in the veto channel and restrict the z -position to the $8\text{ cm} < z < 36\text{ cm}$ region. As will be described later, the restriction in the z -direction was determined based on the range in which the z -dependence of the photon count remains linear. There are small photon-count clusters generated by a few electrons and they disturb the fiducial volume cut. These electrons are considered to be produced by the photoelectric effect by EL light or by the release of electrons attached to impurities. After eliminating these small clusters, from events before the fiducial volume cut is applied by using the information from ELCC and PMT. Then, following overall corrections are applied.

Correction for temporal variation of photon counts

A temporal variation in EL light intensity was observed. One possible cause is the temporal variation in gas purity, which likely led to changes in the amount of impurities that absorb EL light. Therefore, corrections were applied for 30 minute intervals in this study, using the xenon K_{α} energy peaks. Figure 4.9 shows the time evolution of the photon counts from xenon K_{α} before and after correction. The photon counts corresponding to the xenon K_{α} were corrected to be uniform in time.

Correction of the z -dependence of the photon count

Some of ionization electrons are lost during drift due to attachment to impurities. The attenuation length is determined from the z -dependence of the photon count of xenon K_{α} clusters. Figure 4.10 shows the dependence in this study, from which the attenuation length was determined to be $(27\,500 \pm 1020)\text{ mm}$. This corresponds to an electron lifetime of $(25.0 \pm 0.9)\text{ ms}$. The photon counts for every waveform sample are corrected using this attenuation length. In Fig. 4.10, deviations are observed for the smallest and the largest drifts, but the cause is not known. A possible reason is the mis-reconstruction of the z -position. The mis-reconstruction originates from the low detection efficiency of scintillation light. Currently, a new module with higher detection efficiency is under development, which utilizes an acrylic plate coated with a wavelength-shifting material and an MPPC glued behind it. Another possible reason for the deviation on the smallest drifts is insufficient MPPC non-linearity correction. On the smallest drifts, diffusion of

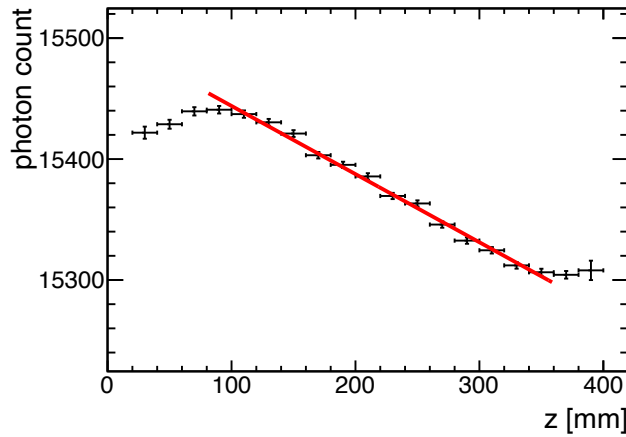


Fig. 4.10 Dependence of the photon counts of K_{α} clusters on the z -position. The red line shows a linear fit.

ionization electrons during drift is small, resulting in a relatively higher photon counts per unit time and very large MPPC non-linearity. The MPPC signal is shaped by an analog filter with a time constant of 220 ns and digitized at a finite sampling rate of 5 MHz [69], and the correction may be inaccurate for such a case. This effect may be validated by comparing simulation data with and without the inclusion of MPPC non-linearity.

Overall fine-tuning for the non-linearity of MPPCs

The effective recovery time of an MPPC can vary from the pre-measured one due to temperature changes, etc. If there remains an overall difference between the effective and the independently calibrated MPPC non-linearity, the effect appears as a linear relation between the photon counts and corrected squared sum, CSS: $\sum_i r^i (N_{\text{rec}}^i)^2$ of events as follows [65].

$$\sum_i r^i N_{\text{rec}}^i - N_{\text{true}} = \sum_i \frac{r^i N_{\text{obs}}^i}{1 - k' N_{\text{obs}}^i} - N_{\text{true}} \simeq \Delta k \sum_i r^i (N_{\text{rec}}^i)^2 \quad (4.2)$$

Here, i runs over every sampling of the waveforms of every hit channel, r^i is the correction factor other than the MPPC non-linearity, N_{obs}^i and N_{rec}^i are the photon counts before and after the MPPC non-linearity correction (Sec. 4.2.1), and N_{true} is the true total photon count of the event. $\Delta k = k - k'$ is the difference between the true and calibrated MPPC non-linearity. The true (calibrated) non-linearity parameter $k^{(\prime)}$ is expressed as $k = \tau^{(\prime)} / (\Delta t \cdot N_{\text{pixel}})$ using the true (calibrated) MPPC recovery time τ , sampling interval Δt , and the number of MPPC pixels N_{pixel} . In this study, the photopeaks of the 911 keV gamma rays from ^{228}Ac , 1461 keV gamma rays from ^{40}K , and the double escape peak of 2615 keV gamma rays from ^{208}Tl were used to evaluate this effect. Distributions of photon counts vs. CSS for these peaks are shown in Fig. 4.11. As described in [65], the observed slopes indicate the overall bias in the MPPC non-linearity correction through the recovery time of MPPCs. The obtained overall bias $\Delta\tau = \tau - \tau'$ was 6.11 ns, compared to the mean calibration value of 72.9 ns. The MPPC non-linearity correction, EL gain correction, correction for the temporal variation of the photon counts and correction of the z -dependence of the photon counts are then repeated with the recovery times shifted by this overall bias.

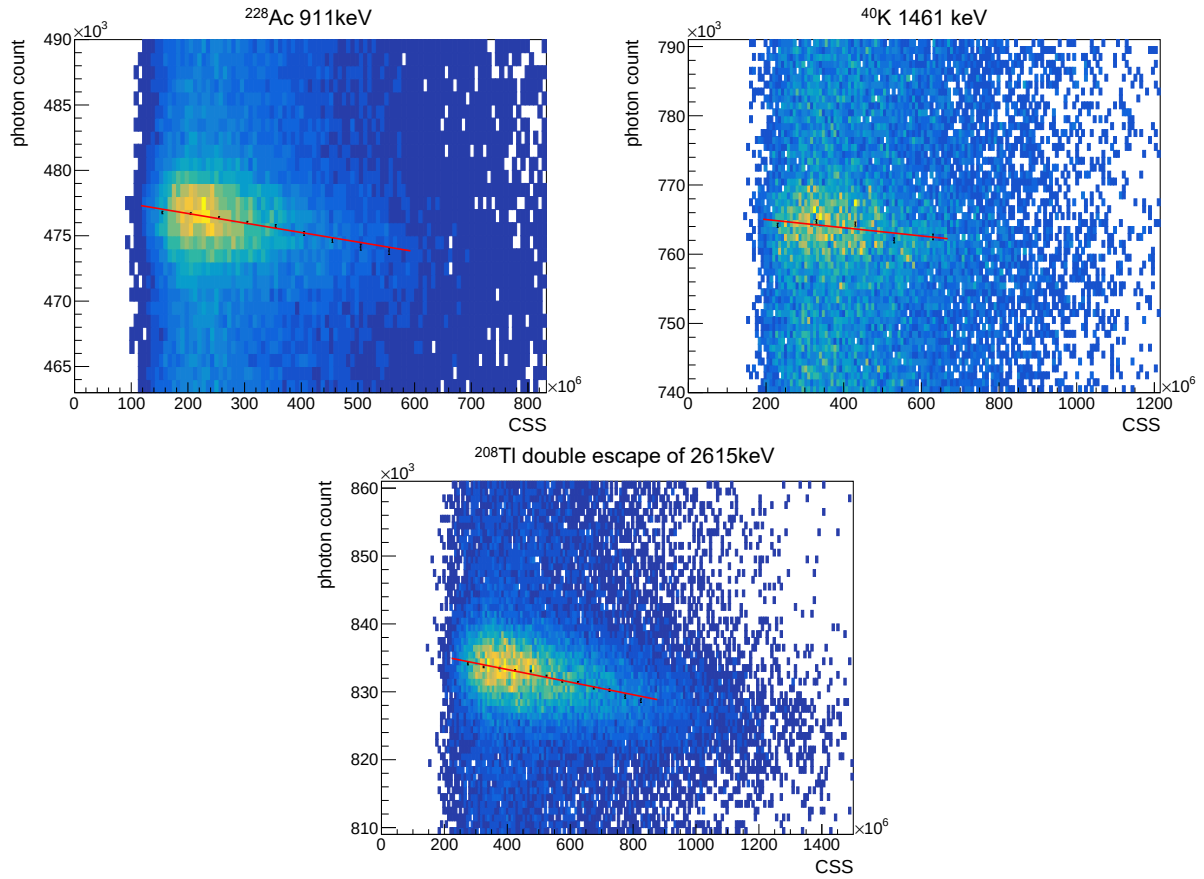


Fig. 4.11 Photon counts vs. CSS of the photopeak of 911 keV gamma rays from ^{228}Ac (top left), 1461 keV gamma rays from ^{40}K (top right) and double escape peak of 2615 keV gamma rays from ^{208}Tl (bottom).

Table 4.2 Summary of the mean photon counts and resolutions for peaks in Fig. 4.12.

	Energy	photon counts	resolution [FWHM]
xenon K_α	29.68 keV	$(1.549\,64 \pm 0.000\,08) \times 10^4$	$(4.431 \pm 0.097) \%$
xenon K_β	33.62 keV	$(1.761\,33 \pm 0.000\,19) \times 10^4$	$(4.364 \pm 0.023) \%$
^{228}Ac	911.2 keV	$(4.804\,52 \pm 0.000\,31) \times 10^5$	$(1.103 \pm 0.017) \%$
environmental ^{40}K	1461 keV	$(7.704\,36 \pm 0.001\,82) \times 10^5$	$(1.065 \pm 0.066) \%$
Double escape of ^{208}Tl 2615 keV	1593 keV	$(8.400\,14 \pm 0.000\,34) \times 10^5$	$(0.980 \pm 0.009) \%$
^{208}Tl	2615 keV	$(1.378\,68 \pm 0.000\,59) \times 10^6$	$(0.672 \pm 0.083) \%$

4.3 Energy resolution

The obtained EL photon count spectrum is presented in Fig. 4.12. The ^{208}Tl 2615 keV gamma ray peak and the single escape peak are clearly seen. Each peak of the spectrum was fitted with a combination of a Gaussian and a linear function, and the results are summarized in Table 4.2. By interpolating these results, the energy resolution at the ^{136}Xe $0\nu\beta\beta$ Q value of 2458 keV was estimated. Two types of energy dependence were considered: a case in which statistical fluctuation dominate ($a\sqrt{E}$) and a case in which systematic errors proportional to the energy exist ($a\sqrt{E} + bE^2$). Figure 4.13 shows the result of the interpolation to the $0\nu\beta\beta$ Q value. Characteristic X-rays are used for EL gain correction in the

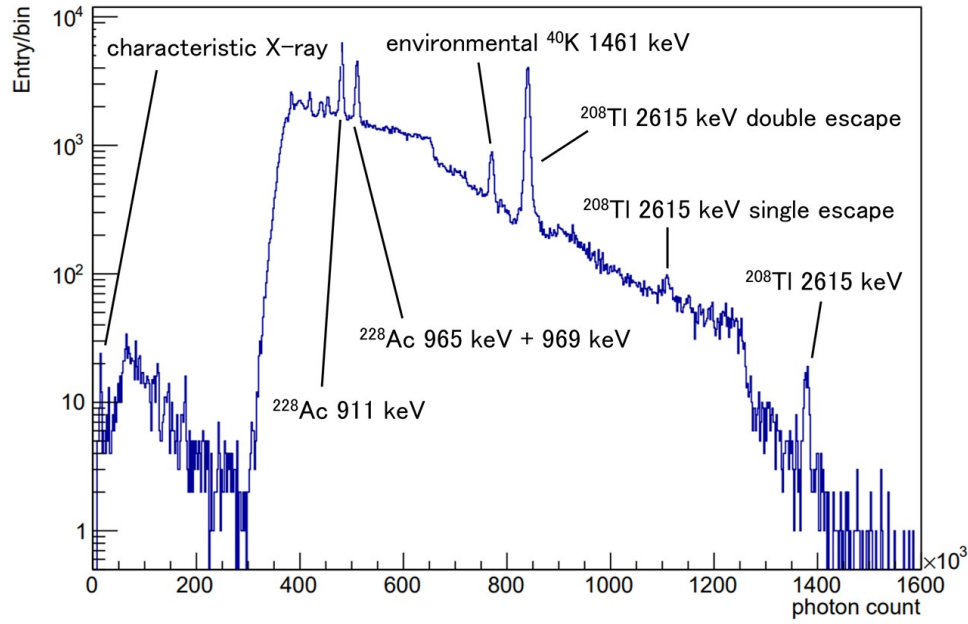


Fig. 4.12 Photon count spectrum after applying all corrections and cuts. The drop in the spectrum below 3×10^5 photons is due to the downsampling of the whole trigger. To reduce the data size and acquire high-energy events efficiently, the whole trigger is issued only once every 50 000 occurrences.

Table 4.3 Summary of the fitted functions to the energy resolutions.

Function	a	b	resolution at the $0\nu\beta\beta$ Q value [FWHM]
$a\sqrt{E}$	0.336 ± 0.005	-	$(0.678 \pm 0.010) \%$
$a\sqrt{E + bE^2}$	0.297 ± 0.029	$(2.99 \pm 2.59) \times 10^{-4}$	$(0.788 \pm 0.077) \%$

analysis, which may lead to an apparently better energy resolution than the actual one. Double escape events involve the emission of two charged particles — an electron and a positron — and may exhibit different energy resolution behavior compared to single-electron photoabsorption events. Therefore, the interpolation was performed using only the data points corresponding to photoabsorption events of gamma rays (indicated by black dots in the figure). The Fit results are summarized in Tab. 4.3. The estimated energy resolution at the $0\nu\beta\beta$ Q value is $(0.678 \pm 0.010) \%$ for $a\sqrt{E}$ and $(0.788 \pm 0.077) \%$ for $a\sqrt{E + bE^2}$. As shown in Tab. 4.3, the value of b in the form $a\sqrt{E + bE^2}$ is small and generally consistent with zero, suggesting that the contribution to the energy resolution is predominantly due to statistical fluctuations.

4.4 Breakdown of the energy resolution

The energy resolution of the 2615 keV γ -ray events obtained in the previous section was analyzed in terms of its contributing factors based on [65][68]. Table 4.4 summarizes the breakdown. In the following, we describe the details of each estimate according to the source of error.

4.4.1 Fluctuation in the signal generation process

This section discusses the following four factors contributing to fluctuations in signal generation.

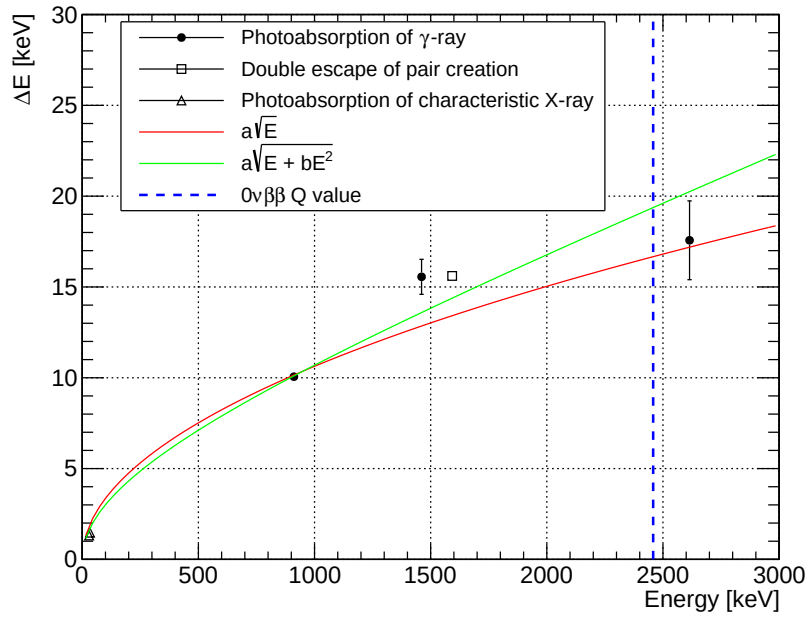


Fig. 4.13 Measured FWHM energy resolutions at various energies and interpolation to the ^{136}Xe Q value.

Fluctuation of the number of initial ionization electrons

The fluctuation in the initial number of ionization electrons $\varepsilon_{\text{init}}$ is given by the following equation.

$$\varepsilon_{\text{init}} = 2.355 \times \sqrt{\frac{F}{N}} = 2.355 \times \sqrt{\frac{FW}{E}} \quad (4.3)$$

where E is the energy of the event. F and W is the Fano factor and the average energy to produce one ionization electron (Sec. 1.6). Consequently, the fluctuation at 2615 keV is estimated to be 0.25 %.

Recombination

When the drift electric field is weak, ionization electrons are lost due to recombination, resulting in a deterioration of the energy resolution. Figure 4.14 shows the relationship between the electric field and energy resolution when measuring 662 keV gamma rays with an ionization chamber[59]. At drift fields equivalent to ≥ 100 V/cm/bar, the energy resolution reaches 0.5 %, whereas in an applied field of 89.9 V/cm/bar, the resolution deteriorates to 0.6 %. This corresponds to 0.17 % when scaled to 2615 keV.

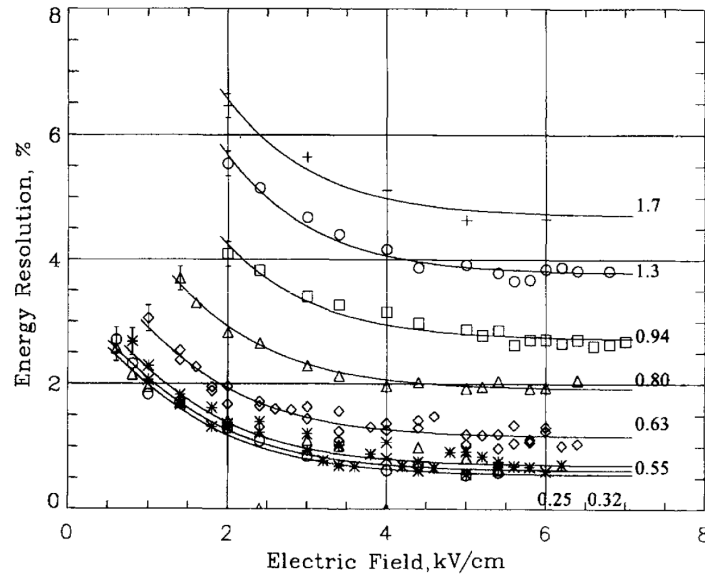
Fluctuation of the attachment

The attenuation length of ionization electrons in this measurement was found to be 27 500 mm, and the variation in the photon counts within the fiducial region $8 \text{ cm} < z < 36 \text{ cm}$ was at most 1.02 %. Therefore, using the average number of photons for 2615 keV events, $N_{\text{ph}} = 1.379 \times 10^6$, the fluctuation due to attachment ε_{att} is expressed by the following equation and is estimated to be at most 0.02 %.

$$\varepsilon_{\text{att}} = 2.355 \times \frac{\sqrt{0.0102 \times N_{\text{ph}}}}{N_{\text{ph}}} \quad (4.4)$$

Table 4.4 Breakdown of the energy resolution of 2615 keV γ -ray events.

Fluctuation of the number of initial ionization electrons (Sec. 4.4.1)	0.25 %
Mis-reconstruction of z -position (Sec. 4.4.4)	0.24 %
Fluctuation of the EL generation and detection (Sec. 4.4.1)	0.20 %
Error in EL gain calibration (Sec. 4.4.2)	0.18 %
Error in time dependence correction (Sec. 4.4.2)	0.18 %
Recombination (Sec. 4.4.1)	0.17 %
Variation in time bin of time variation correction (Sec. 4.4.2)	0.12 %
Offset of the baseline (Sec. 4.4.3)	≤ 0.11 %
Error in z dependence correction (Sec. 4.4.2)	≤ 0.06 %
Fluctuation of the attachment (Sec. 4.4.1)	≤ 0.02 %
Accuracy of the MPPC recovery time (Sec. 4.4.2)	≤ 0.03 %
Position dependence of the EL gain (Sec. 4.4.3)	~ 0 %
Waveform filtering in the AxFEB (Sec. 4.4.3)	~ 0 %
Estimation total	0.52 % to 0.54 %
Data total	(0.67 ± 0.08) %

Fig. 4.14 Drift electric field dependence of the intrinsic energy resolution (%FWHM) measured for 662 keV γ -rays. Gas densities are shown at right of the curves. This figure is taken from [59].

Fluctuation of the EL generation and detection

The fluctuation of the EL generation and detection $\varepsilon_{\text{ELgain}}$ is given by the following equation, using the average number of ionization electrons $N_e = E/W$ and the average EL gain $g = 11.5$ (Sec. 4.2.1).

$$\varepsilon_{\text{ELgain}} = 2.355 \times \sqrt{\frac{1}{gN_e}} \quad (4.5)$$

Consequently, the fluctuation at 2615 keV is estimated to be 0.20 %.

4.4.2 Correction error

This section discusses the following five factors contributing to errors in the correction process.

EL gain correction

The error in EL gain correction originates from the accuracy in determining the photon count of the K_α peak for each channel used in the correction, and is given by the following equation,

$$\varepsilon_{\text{ELcor}} = 2.355 \times \frac{\sqrt{\sum_{\text{ch}} (\epsilon_{\text{ch}} N_{\text{ch}})^2}}{N} \sim 2.355 \times \bar{\epsilon} \sqrt{\frac{\sum_{\text{ch}} N_{\text{ch}}^2}{N^2}} \quad (4.6)$$

where ϵ_{ch} is the error of the peak photon count for each channel, $\bar{\epsilon}$ is the mean error, N_{ch} is the mean photon count for each channel, and the N is the mean photon count at 2615 keV. $\bar{\epsilon} = 0.53\%$ is derived from the mean of the fit errors for the K_α peak for each channel. $\sum_{\text{ch}} N_{\text{ch}}^2/N^2$ is also calculated for each event, and 0.021 is obtained as the mean of its distribution. Therefore, the contribution is estimated to be $\varepsilon_{\text{ELcor}} = 0.18\%$. This contribution can be interpreted as $2.355 \times \bar{\epsilon}/\sqrt{n_{\text{eff}}}$, where n_{eff} denotes the effective number of hit channels. From this relation, $n_{\text{eff}} = 47.2$ is obtained.

Accuracy of the MPPC recovery times

The contribution of the MPPC recovery time error can be separated into the error associated with the recovery time of individual MPPCs and the overall systematic bias. The recovery time of individual MPPCs has been measured with a precision about 0.5 ns. Simulations have shown that this contribution is negligible[66]. The effect of the overall bias is estimated from Eq. 4.2. After the overall fine-tuning for the non-linearity of MPPCs, the difference of the non-linearity parameter $\Delta k = (8.33 \pm 2.08) \times 10^{-7}$ is obtained. The standard deviation of the distribution of CSS for 2615 keV events ΔCSS is 1.96×10^8 , and the mean photon count N is 1.379×10^6 . Therefore, the contribution is expressed by the following equation and is estimated to be at most 0.03 %.

$$\varepsilon_{\text{MPPC}} = 2.355 \times \frac{\Delta k \times \Delta \text{CSS}}{N} \quad (4.7)$$

Error in time dependence correction

The temporal variation in photon count is corrected using the peak photon count of the K_α characteristic X-ray in each time bin (Sec. 4.2.3). The relative error of the peak photon count is, on average, 0.076 %, and based on the error propagation, the relative error of the correction factor is also 0.076 %. Therefore, the error associated with the time dependence correction is estimated to be $\varepsilon_{\text{tcor}} = 2.355 \times 0.076 \% = 0.18 \%$.

Variation in time bin of time variation correction

The impact of temporal variation within a time bin was estimated from the peak values of adjacent time bins. Over the 30-minute interval between time bins, the average temporal variation in relative error was 0.18 %. Assuming uniform variation within each bin, the resulting error is estimated to be $\varepsilon_{\text{tbin}} = 2.355 \times 0.18 \% / \sqrt{12} = 0.12 \%$.

Error in z dependence correction

The error in the z dependence correction originates from the error in determining the attenuation length. The variation in photon count ΔN_{cor} due to deviations in the attenuation length is given by the following equation.

$$\Delta N_{\text{cor}} = \sum_i r^i N_{\text{obs}}^i \left(1 + \frac{z^i}{\lambda'}\right) - \sum_i r^i N_{\text{obs}}^i \left(1 + \frac{z^i}{\lambda}\right) \simeq \left(\frac{1}{\lambda'} - \frac{1}{\lambda}\right) \bar{z} N_{\text{cor}} \quad (4.8)$$

Here i runs over every sampling of the waveforms of every hit channel, r^i is the correction factor other than the z dependence correction, z^i is the z -position of each sampling of the waveform, $\lambda^{(i)}$ is the true (estimated) attenuation length, and $\bar{z}$ is the photon-count weighted centroid position of event in z direction. Therefore, the slope in photon counts vs. $\bar{z} \times$ photon counts distribution shows the difference of the attenuation length. In this study, the photopeaks of the 911 keV gamma rays from ^{228}Ac , 1461 keV gamma rays from ^{40}K , and the double escape peak of 2615 keV gamma rays from ^{208}Tl are used to evaluate this effect and $1/\lambda' - 1/\lambda = (4.24 \pm 0.69) \times 10^{-6} \text{ mm}^{-1}$ is obtained. Since the standard deviation $\Delta(\bar{z}N_{\text{cor}})$ at 2615 keV is obtained as $6.80 \times 10^7 \text{ mm}$ and the mean photon count $N = 1.379 \times 10^6$, the contribution is estimated to be at most 0.06 % by the following equation.

$$\varepsilon_{\text{zcor}} = 2.355 \times \left(\frac{1}{\lambda'} - \frac{1}{\lambda} \right) \times \frac{\Delta(\bar{z}N_{\text{cor}})}{N} \quad (4.9)$$

4.4.3 Hardware-origin error

This section discusses contributions originating from hardware components.

The EL gain varies depending on the relative position within the cell in which the ionization electron enters[77]. This effect has been taken into account in the simulation and a comparison with the case assuming uniform gain within the cell has shown that the impact is negligible.

The signal from the ELCC is shaped by a Sallen-Key filter on the AxFEB and subsequently digitized. The effects of this filtering and digitization have also been evaluated through simulation and found to be negligible.

Offset of the baseline

Since the waveform is digitized, the baseline cannot be determined with a precision better than 1 ADC count. Therefore, when the total number of hit clocks (hereafter referred to as hit volume) fluctuates, the integrated value can vary by up to the fluctuation multiplied by 1 ADC count. The standard deviation of the hit volume distribution ΔV at 2615 keV is 2296 clocks. Using the mean MPPC gain of $g = 0.954 \text{ ADC/1 p.e.}$ and the mean photon count $N = 1.379 \times 10^6$, the variation caused by a maximum offset of 1 ADC count across all channels is calculated to be at most 0.41 %, as given by the following equation.

$$\Delta N = 2.355 \times \frac{\Delta V}{g \times N} \quad (4.10)$$

The contribution of this variation to the energy resolution $\varepsilon_{\text{offset1}}$ is estimated using the effective number of hit channels n_{eff} obtained in Sec. 4.4.2 (EL gain correction), yielding $\varepsilon_{\text{offset1}} = 0.41 \% / \sqrt{n_{\text{eff}}} = 0.06 \%$.

In addition, even when the hit volume is identical, variations in the offset values can also contribute to fluctuations. The mean of the hit volume distribution $\bar{V}$ at 2615 keV is 1.165×10^4 clocks. Therefore, if the offset uniformly distributes from 0 to 1, the variation caused by this effect is calculated to be 0.60 %, as given by the following equation.

$$\Delta N = 2.355 \times \frac{\bar{V}}{\sqrt{12} \times g \times N} \quad (4.11)$$

The contribution of this variation to the energy resolution $\varepsilon_{\text{offset2}}$ is estimated as $\varepsilon_{\text{offset2}} = 0.60 \% / \sqrt{n_{\text{eff}}} = 0.09 \%$. By combining these two contributions, the impact of the baseline offset is estimated to be at most 0.11 %.

4.4.4 Mis-reconstruction of z -position

This section discusses the error arising from mis-reconstruction of the z -position due to incorrect detection of scintillation light.

Let ϵ denote the detection efficiency of scintillation light, and μ_{acc} the average number of accidental scintillation hits. Then, the probabilities that the number of scintillation hit clusters n_{sci} is 0 or 1 are given by the following expressions.

$$P(n_{\text{sci}} = 0) = (1 - \epsilon)e^{-\mu_{\text{acc}}} \quad (4.12)$$

$$P(n_{\text{sci}} = 1) = (1 - \epsilon)\mu_{\text{acc}}e^{-\mu_{\text{acc}}} + \epsilon e^{-\mu_{\text{acc}}} \quad (4.13)$$

From the measured distribution of number of scintillation hit clusters, solving the above equations simultaneously yields $\epsilon = 0.701$ and $\mu_{\text{acc}} = 0.402$. Under these conditions, the probability of mis-reconstruction due to failure in detecting scintillation light and the presence of accidental scintillation hit is calculated to be 14.6 %, as given by the following equation.

$$P_{\text{miss}} = \frac{(1 - \epsilon)\mu_{\text{acc}}e^{-\mu_{\text{acc}}}}{P(n_{\text{sci}} = 1)} \quad (4.14)$$

Assuming that the peak photon count at 2615 keV follows a Gaussian distribution with mean N and standard deviation σ , and that a fraction p of events fluctuates by $N \times d$, the resulting increase in variance is given by $N^2pd^2(1 - p)$, and the corresponding coefficient of variation is $d\sqrt{p(1 - p)}$. If the distribution of z position due to mis-reconstruction is uniform within the fiducial region $8 \text{ cm} < z < 36 \text{ cm}$, the variation of photon count is uniformly distributed within a $d_{\text{max}} = 1.02 \%$ range (Sec. 4.4.1, Fluctuation of the attachment). Based on this, the contribution to the resolution $\varepsilon_{\text{zmis}}$ is estimated to be 0.24 %, as given by the following equation.

$$\varepsilon_{\text{zmis}} = 2.355 \times \frac{d_{\text{max}} \times \sqrt{P_{\text{miss}} \times (1 - P_{\text{miss}})}}{\sqrt{12}} \quad (4.15)$$

4.4.5 Summary

The energy resolution of 2615 keV γ -ray events was measured to be $(0.67 \pm 0.08) \%$, while the estimated resolution ranges from 0.52 % to 0.54 %, indicating an unknown contribution of 0.25 %. One unaccounted factor is the fluctuation due to MPPC non-linearity. As described in Sec. 4.2.1, the MPPCs have a non-linear output under high light intensity. Whether multiple photons simultaneously hit a given pixel is governed by a statistical process, which introduces an associated error. However, this error has been provisionally estimated at 0.09 % based on simulation, and it does not fully account for the total unknown contribution of 0.25 %. Therefore, further investigation into other contributing factors is necessary.

In this measurement, the drift electric field strength was set to 89.9 V/cm/bar, and the EL field was reduced to 2.72 kV/cm/bar. These conditions affect recombination and the fluctuation of EL generation and detection, respectively. If the electric fields had been applied according to the design specifications, their contributions would be estimated as 0.19 % and $\sim 0 \%$, respectively. Therefore, the degradation in resolution due to reduced electric field strength is estimated to be 0.18 %, suggesting that its impact on the overall energy resolution is limited.

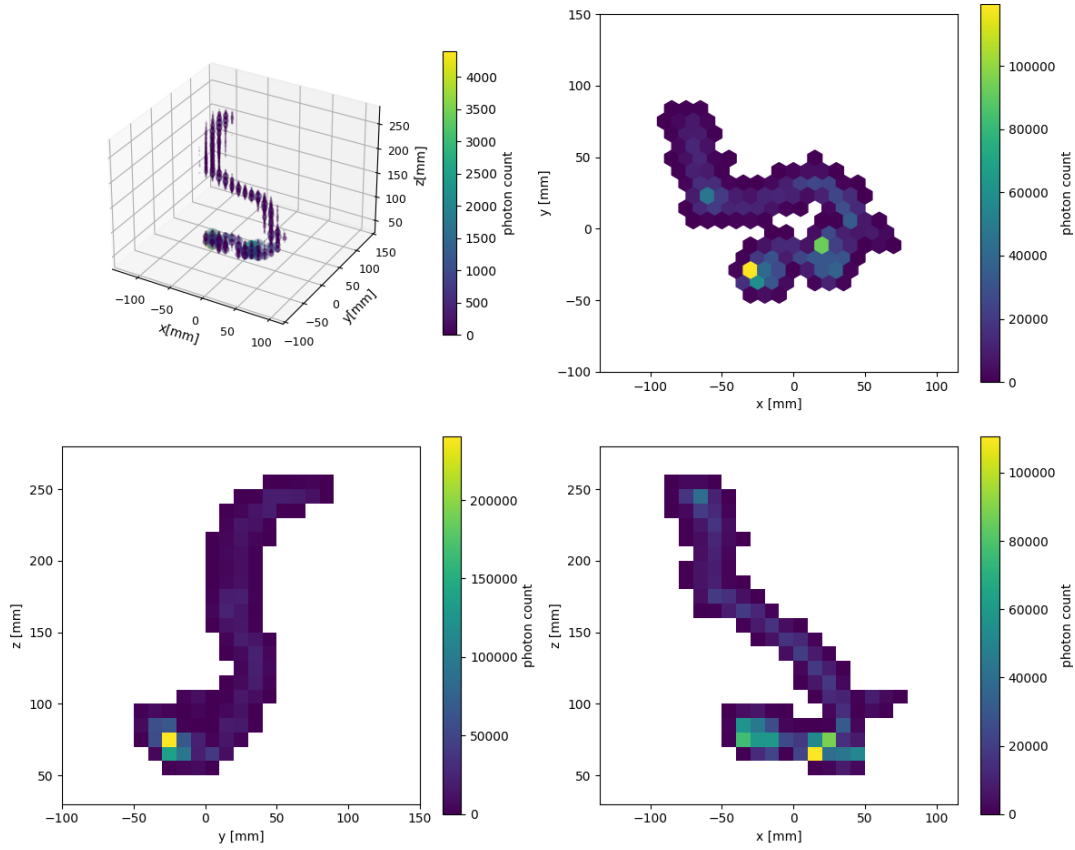


Fig. 4.15 An example of reconstructed track of 2615 keV events. In the 3D plot, the size of the points is proportional to the photon counts. One blob can be seen at the end of the track.

4.5 Track topology

Reconstructed tracks of a 2615 keV and a 1593 keV of gamma induced electrons are shown in Fig. 4.15 and Fig. 4.16, respectively. The 2615 keV corresponds to the photoelectric absorption peak of gamma rays from ^{208}Tl , and the event looks like a single electron track. The 1593 keV corresponds to the double-escape peak of the 2615 keV gamma rays from ^{208}Tl , and the event presumably consists of two tracks, one electron and one positron, originating from a single vertex due to pair creation. At the end of the each track, there is a blob associated with high energy loss and multiple scattering. In the 2615 keV event, there is one blob produced at the electron stop. The 1593 keV event produces two blobs at the end of electron and positron track, respectively. Upon scanning multiple one cluster events we observe that most of 2615 keV events appeared as single electron track as that in Fig. 4.15, while at 1593 keV, they appeared as tracks with one electron and one positron as is in Fig. 4.16. Multi-cluster events were also checked by eye and they are likely caused either by bremsstrahlung photons emitted by electrons or by simultaneous interactions involving multiple photons. Since two blobs are expected in the $0\nu\beta\beta$ event as well, this track information can be used to distinguish the single-electron events from gamma ray backgrounds when the blobs are clearly seen as in these figures.

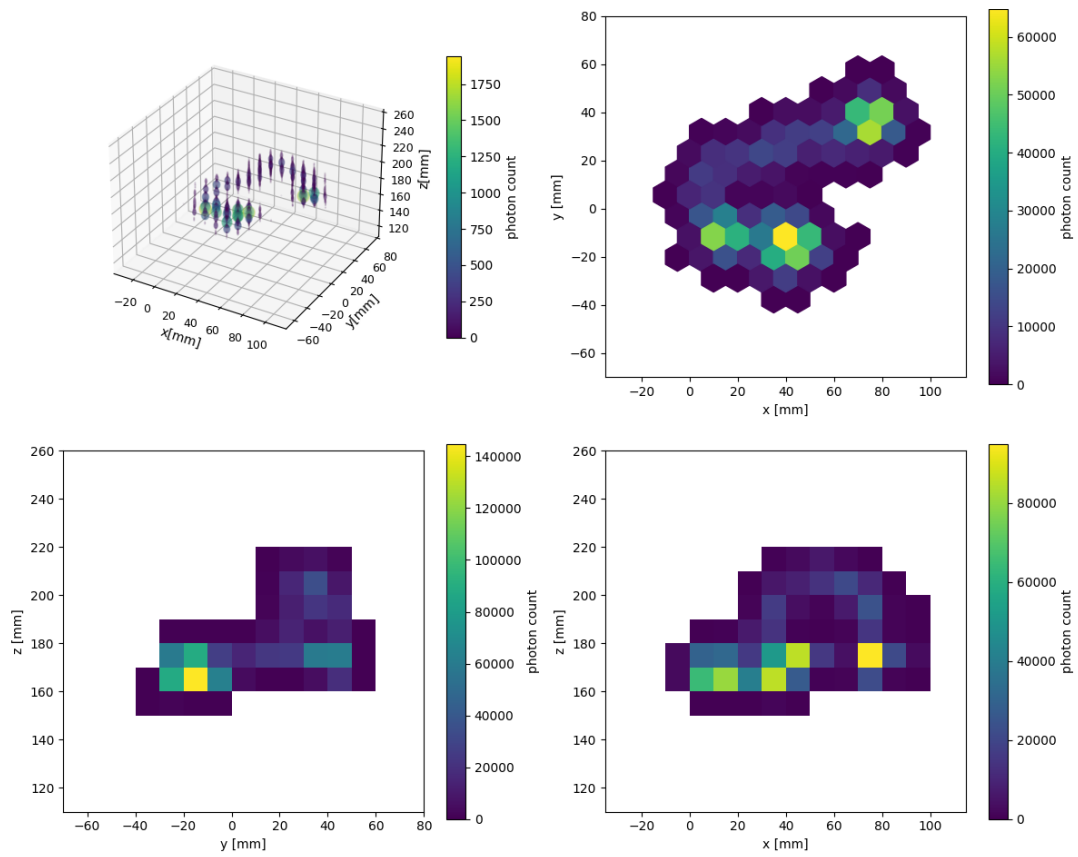


Fig. 4.16 An example of reconstructed track of 1593 keV events. In the 3D plot, the size of the points is proportional to the photon counts. There are 2 blobs corresponds to the end of the track of an electron and a positron.

Part III

Signal vs. background discrimination by track
pattern using machine learning

Chapter 5

Simulation for machine learning

5.1 Background assumptions

The purpose of using machine learning in the AXEL experiment is to discriminate between signals and background events using the 3D track information obtained by TPC. As described in Sec. 1.6, the high energy resolution of the xenon gas TPC utilizing EL amplification makes the contamination from $2\nu\beta\beta$ negligible. According to [28], the contamination fraction F from $2\nu\beta\beta$ approximately scales as $F = 7Q\delta^6/m_e$, where $\delta = \Delta E/Q$ is the FWHM energy resolution expressed as a fraction, Q is the Q value of $0\nu\beta\beta$ and m_e is the electron mass. For 1 ton of ^{136}Xe , the annual number of $2\nu\beta\beta$ decays is approximately 1.3×10^6 , with the half-life of 2.34×10^{21} years[41]. Even without considering the detector efficiency, the number of events leaking into the region of interest due to the energy resolution is less than 6.9×10^{-7} per year when $\delta = 0.5\%$ and it is negligible. Prompt gamma rays resulting from the absorption of cosmic-ray-induced neutrons can be eliminated by surrounding the detector with veto detectors and applying coincidence techniques. Long-lived nuclei produced by the nuclear spallation of ^{136}Xe are collected on the cathode of the TPC. Electron-ion recombination in gas is weak, and a study investigating the decay chain of ^{222}Rn in a xenon gas TPC has shown that the resulting ions are collected at the cathode with nearly 100 % efficiency over a 530 mm drift region[78]. Therefore, β -rays and associated γ -rays from these long-lived nuclides can be removed by applying a fiducial cut near the cathode, and coincidence techniques. α tracks are very short and can be easily rejected thanks to the track reconstruction capability of the TPC. A potential issue with beta sources is the production of ^{137}Xe through neutron absorption by ^{136}Xe [51]. ^{137}Xe has a half-life of 3.8 minutes and undergoes β -decay with a Q -value of 4.16 MeV. The continuous β -spectrum includes components with energies near the Q -value of $0\nu\beta\beta$, which can contribute to the background. Due to its relatively long lifetime, it cannot be vetoed by coincidence with cosmic-ray muon events. This background is unavoidable when using ^{136}Xe ; however, it can be mitigated by adding ^3He , which has a large neutron absorption cross section[47].

Consequently, the potential background sources include electrons produced by photoelectric absorption of gamma rays with energies near the Q -value of xenon $0\nu\beta\beta$, as well as electrons from the decay of cosmogenically produced ^{137}Xe , which acts as a β -emitter and cannot be vetoed by coincidence techniques. In this study, three types of background sources are considered: ^{214}Bi , ^{208}Tl , and ^{137}Xe . ^{214}Bi is a nucleus in the uranium series and undergoes beta decay to ^{214}Po , emitting a 2448 keV gamma ray with a probability of approximately 1.5 %, close to the Q -value of ^{136}Xe $0\nu\beta\beta$. The energy difference is 0.69 %, which cannot be completely separated by the designed FWHM energy resolution of 0.5 % in the AXEL experiment. ^{208}Tl is a nucleus in the uranium series and undergoes beta decay to ^{208}Pb , emitting a 2615 keV gamma ray with a probability of approximately 99.8 %. Although this energy is significantly higher than the Q -value, it can contribute to events near the Q -value if it undergoes Compton scattering in components outside the TPC, such as the pressure vessel, losing energy before being photoelectrically

absorbed within the TPC. As previously mentioned, ^{137}Xe is produced via neutron capture induced by cosmic rays, and the continuous component of its beta spectrum includes events near the $0\nu\beta\beta$ Q-value, which contribute to the background. To distinguish these backgrounds, the track information will be used. In the $0\nu\beta\beta$ event, multiple scattering of two electrons at the stop produces two blobs, whereas only one blob is produced in the γ ray background from ^{214}Bi and ^{208}Tl , or β ray background from ^{137}Xe . These characteristics can be used for discrimination.

In this study, the following procedures were used to build and validate the machine learning model. Simulations of $0\nu\beta\beta$ events and γ ray background events were performed using Geant4[79]. The generated ionization electrons are converted into signal waveforms considering the detector response; diffusion in the drift, EL light generation at the ELCC, analog filtering and digitization in the frontend board. The waveform is reconstructed into a 3D track by the same analysis as Chap. 4. These track data are split into datasets for training and validation of machine learning models. The machine learning model is trained to discriminate tracks into $0\nu\beta\beta$ signal and background based on training data, and its performance is evaluated on a validation data set. The photoelectric absorption events of the 2615 keV gamma rays from ^{208}Tl , observed in the measurements described in Chap. 4, were also input into the model. The behavior of these events was then compared with that of the simulated background data.

5.2 Creation of data set for training model

5.2.1 Event generation using Geant4

The particles produced by the $0\nu\beta\beta$ decay and the γ -ray reaction were simulated using Geant4. Differences in beta-ray energy manifest as variations in track length. Although the Q-value of $0\nu\beta\beta$ in xenon is 2458 keV and the gamma ray from ^{214}Bi has an energy of 2448 keV, the simulation energy was set to 2615 keV for both $0\nu\beta\beta$ and gamma rays in order to enable a direct comparison of track shapes with the 2615 keV events obtained in Chap. 4.

The $0\nu\beta\beta$ events were generated uniformly within the xenon gas volume in the geometry. The energy and opening angle of the electrons emitted in the $0\nu\beta\beta$ decay were sampled according to the following distribution[80],

$$\frac{d\Gamma}{dE_1 d\cos\theta} \propto 1 - \beta_1\beta_2 \cos\theta \quad (5.1)$$

where Γ denotes the decay rate, E_1 is the energy of one of the emitted electrons (with $E_2 = Q_{\beta\beta} + 2m_e - E_1$), β_1 and β_2 represent the velocities of the respective electrons normalized by the speed of light and θ is the opening angle between them.

γ -ray events were generated vertically downward from an area above the gas volume, 260 mm wide and 370 mm long, which corresponds to the sensitive area of the TPC. However, as will be discussed later in this section, it was found that electrons emitted via the photoelectric effect are strongly biased in the direction of incident γ -rays. Consequently, using these data for model training could lead to incorrect learning, where vertically downward tracks are mistakenly classified as background. To mitigate this issue, we applied an additional process that randomly rotates each track in 4π steradians around the midpoint of its spatial coordinates, thereby eliminating the directional bias of electron emission.

To reduce data size, only events that dropped all energy into xenon gas were stored. In this study, 1.2×10^6 $0\nu\beta\beta$ events and 3.6×10^9 γ -ray events were generated.

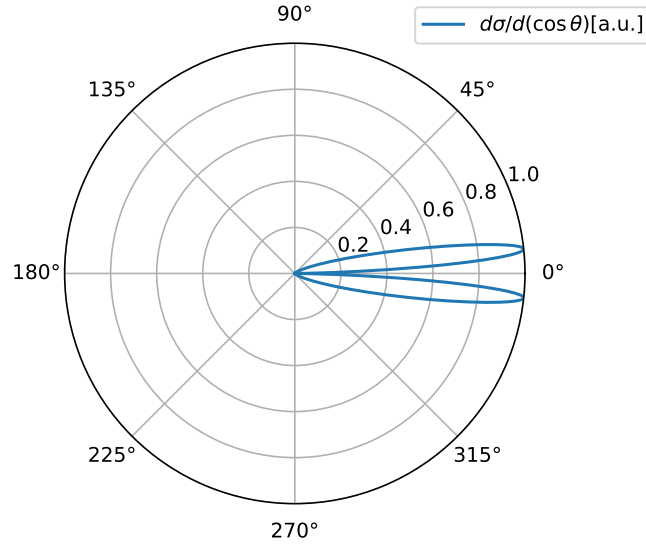


Fig. 5.1 Sauter-Gavrila distribution at the energy of photoelectron $E = 2580$ keV, $\beta = 0.986$ and $\gamma = 6.05$

Emission direction of electrons in the photoelectric effect

The direction of the electron emitted by photoelectric absorption is considered to follow the Sauter-Gavrila distribution with respect to the incident direction of the gamma ray[81][82],

$$\frac{d\sigma}{d(\cos\theta)} \sim \frac{\sin^2\theta}{(1 - \beta \cos\theta)^4} \left\{ 1 + \frac{1}{2}\gamma(\gamma - 1)(\gamma - 2)(1 - \beta \cos\theta) \right\} \quad (5.2)$$

where β and γ are the Lorents factors of the photoelectron. Figure 5.1 shows the angular distribution described above for an electron with energy corresponding to a 2615 keV gamma ray minus the xenon binding energy of 34.57 keV ($\beta = 0.986$, $\gamma = 6.05$). It can be seen that the emission direction of the electron is highly biased toward the forward direction.

5.2.2 Waveform generation

In this section, we describe a method for generating ADC waveforms from Geant4 simulation data that resemble those obtained from actual measurements. Since the simulation data were produced in parallel with the acquisition of real data, some of the parameters used in the waveform generation were derived from the analysis results of measurement data collected up to July 25.

Generation of ionization electron

Ionization electrons are randomly generated along the track using the information of energy deposit from Geant4 simulation. A W value of 22.1[44] and a Fano factor F of 0.13[45] are used to determine the total number of electrons generated.

Diffusion and attachment

The position of the electrons are then randomly displaced by the amount of diffusion expected after the drift from the generation point to the point 2 cm before ELCC. The diffusion constants were calculated using Garfield++[83] for pure xenon at 6.8 bar, and the obtained values are 0.118 cm/ $\sqrt{\text{cm}}$ for transverse diffusion, 0.033 cm/ $\sqrt{\text{cm}}$ for longitudinal diffusion, respectively. Considering the effect of electron

attachment, the electrons were eliminated with the following probabilities according to the z -position.

$$P_{\text{attach}} = 1 - \exp(-z/L_{\text{att}}) \quad (5.3)$$

Here, z is the z -position of the electron, L_{att} is the attenuation length due to the attachment, respectively. As the attenuation length, we used $L_{\text{att}} = 36\,000$ mm estimated from measurements up to July 25. For the full dataset, the estimated attenuation length is $(27\,500 \pm 1020)$ mm.

EL photon conversion

For each electron, the number of photons generated in the EL process is calculated according to a Poisson distribution with mean EL gain g . The mean EL gain is set to $g = 10.8$, estimated from measurements up to July 25. In order to incorporate the effect of the EL gain being different for each ELCC channel, we fluctuate it channel-by-channel by the standard deviation of 5.9%, which were determined from the distribution of the photon counts of the K_α peak of each channel.

MPPC non-linearity

The recovery time τ of each MPPC, which determines the linearity characteristic is randomly set according to the nominal distribution of mean 73.2 ns and standard deviation 2.5 ns based on the recovery times measured in the calibration. In the model, the gain of one MPPC pixel with a hit decreases by a factor of $1 - \exp(-t/\tau)$ at time t after the hit, the effective gain of the MPPC $\bar{g}$ is given by,

$$\bar{g}[i] = g_0 \left(1 - \sum_{j=0}^{\infty} \frac{N_{\text{obs}}[i-j]}{N_{\text{pix}}} \exp(-j\Delta t/\tau) \right) \quad (5.4)$$

where i, j are time indices with a calculation time step Δt , g_0 is the gain of MPPC under conditions with no prior hits and N_{pix} is the number of pixels of the MPPC. N_{obs} is the number of photons observed and is related to the true number of photons incident on the MPPC, N_{true} , by $N_{\text{obs}} = \bar{g}/g_0 \times N_{\text{true}}$. The calculation time step Δt is set to 20 ns. Thus, the mean number of photons observed $\overline{N_{\text{obs}}}$ is given by

$$\overline{N_{\text{obs}}}[i] = N_{\text{true}}[i] - \frac{N_{\text{true}}[i]}{N_{\text{pix}}} \sum_{j=0}^{\infty} N_{\text{obs}}[i-j] \exp(-j\Delta t/\tau) \quad (5.5)$$

The number of photons at each time index was determined by a binomially distributed random number $B(n, p)$ with $n = N_{\text{true}}$ and $p = \overline{N_{\text{obs}}}/N_{\text{true}}$.

Analog filter

An analog filter was applied to the photon count at each clock. The schematic diagram of second-order Sallen-key filter used in AxFEB is shown in Fig. 5.2. The transfer function $G(s) = V_{\text{out}}/V_{\text{in}}$ of this filter is given by

$$G(s) = \frac{1}{\overline{R_1 R_2 C_1 C_2} \left(s^2 + s \left(\frac{1}{R_1 C_1} + \frac{1}{R_2 C_1} \right) + \frac{1}{R_1 R_2 C_1 C_2} \right)} \quad (5.6)$$

where $s = j\omega$, j is the imaginary unit and ω is the frequency of the input signal. For AxFEB, $R_1 = R_2 = 1\text{ k}\Omega$ and $C_1 = C_2 = 220\text{ pF}$. With $R_1 C_1 = R_2 C_2 = \tau$, the inverse Laplace transform of Eq. 5.6 yields the following impulse response function,

$$y(t) = \frac{t}{\tau^2} \exp(-t/\tau) \quad (5.7)$$

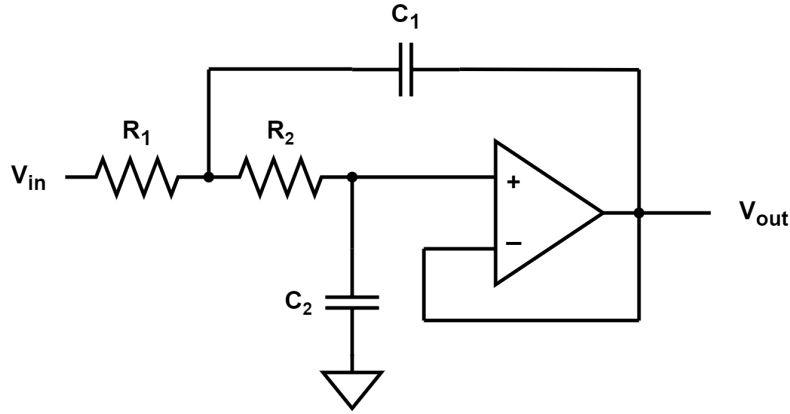


Fig. 5.2 The schematic diagram of second-order Sallen-key filter.

The filtered photon count was obtained considering the contribution of past clocks using the response in Eq. 5.7. The filtered photon count was then converted to an ADC count by multiplying it by the MPPC gain and rounding to the nearest integer. MPPC gain was set to the 0.9577 ADC/1 p.e. according to the previous study[66].

5.2.3 Analysis

The sequential array of the ADC data obtained are used as ELCC waveforms and analyzed in the same way as the previous chapter. In this study, the scintillation is not simulated and the timing information is created with the event occurrence timing set to $t = 0$. The values given in the previous section, the MPPC gain, MPPC non-linearity and the decay length, are also used for this analysis, too. A fiducial volume cut was also applied, but with the cut range set to $2\text{ cm} < z < 39\text{ cm}$ in order to maximize the usable volume. Fine-tuning (Sec. 4.2.3) was not performed because the non-linear parameters of MPPC are known.

For the real data, the cut conditions described in Chap. 4 were slightly relaxed as follows to increase the amount of data available for comparison with the simulation. The fiducial volume cut range was modified to $2\text{ cm} < z < 39\text{ cm}$, consistent with the simulation. Additionally, the extra selection criteria for avoiding timing mismatches described at the end of Sec. 4.2.2 were not applied.

Then, events within one FWHM of the peak energy were selected. The simulation data were further divided into training and validation datasets in a 3:1 ratio. The number of selected events in the simulation and real data is summarized in Tab. 5.1.

Table 5.1 The number of events after applying the analysis.

data	$0\nu\beta\beta$ events	background events
simulation train data	17680	8383
simulation validation data	5894	2794
real data	-	835

5.2.4 Elimination of the Compton-scattering events

From the post-analysis data set, we further cut events with multi clusters, which is mainly due to the Compton effect. Since multiple clusters are also produced by electron bremsstrahlung, which we do

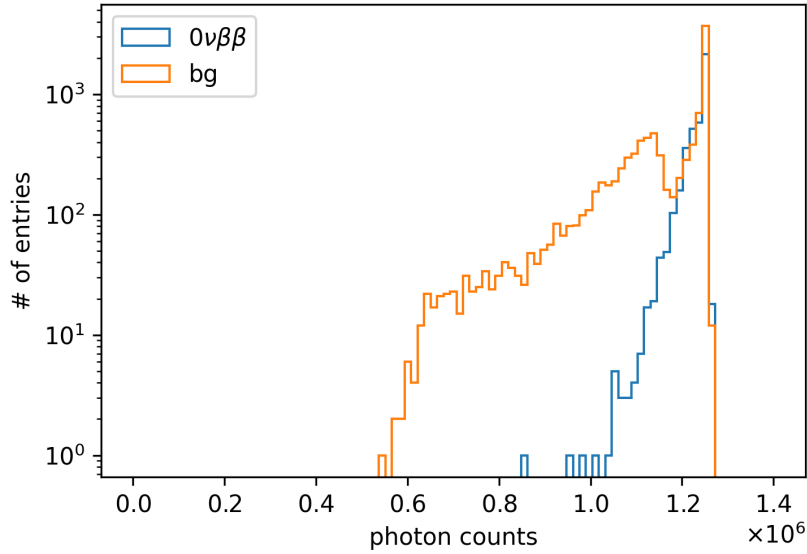


Fig. 5.3 Distribution of photon counts of the cluster with the largest photon count in an event. In the background data (bg), a dip that corresponds to the Compton edge is observed at around 1.2×10^6 photon count.

not want to cut, the cutting method was studied by focusing on the photon counts of the clusters. The distribution of photon counts of the cluster with the largest photon count in the event is shown in Fig. 5.3. In the background data, a dip that corresponds to the Compton edge is observed at around 1.2×10^6 photon count. The Compton edge is given by Eq. 5.8 is 2382 keV for the 2615 keV γ ray, corresponding to 91.1 % of the total energy.

$$E_{ce} = \frac{2E^2}{m_e + 2E} \quad (5.8)$$

Here, E is the energy of initial γ ray, and m_e is the mass of electron. We set the threshold to cut the Compton effect at 92 % of the peak. The survival efficiency of $0\nu\beta\beta$ and background events by this cut is 99.3 % and 75.1 %, respectively.

5.3 Data set characteristics

In this section, we compare the characteristics of the simulation data with the real data obtained in Chap. 4. In addition to comparisons between simulation datasets, comparisons between real data and training data were also performed for the background. We focused on the overall and local photon count, the number of clusters, and the volume and the position the tracks as factors influencing the model's output. The number of events is 17 554 for the simulated $0\nu\beta\beta$ data, 6230 for the simulated background data, and 503 for the real ^{208}Tl data after the Compton-scattering cut. In the following, the error bars of each histogram bin represent the standard deviation according to the Poisson distribution. Given the differences in the amount of data, each bin of the histograms was normalized to the proportion of the total number of events.

5.3.1 Overall photon count

Figure 5.4 shows the distribution of the photon counts of events. The photon count of the simulated data is 7.4 % smaller than that of the real data. This is mainly due to the ELgain in the simulation was configured based on measurement results obtained up to July 25 (Sec. 5.2.2), resulting in a value that

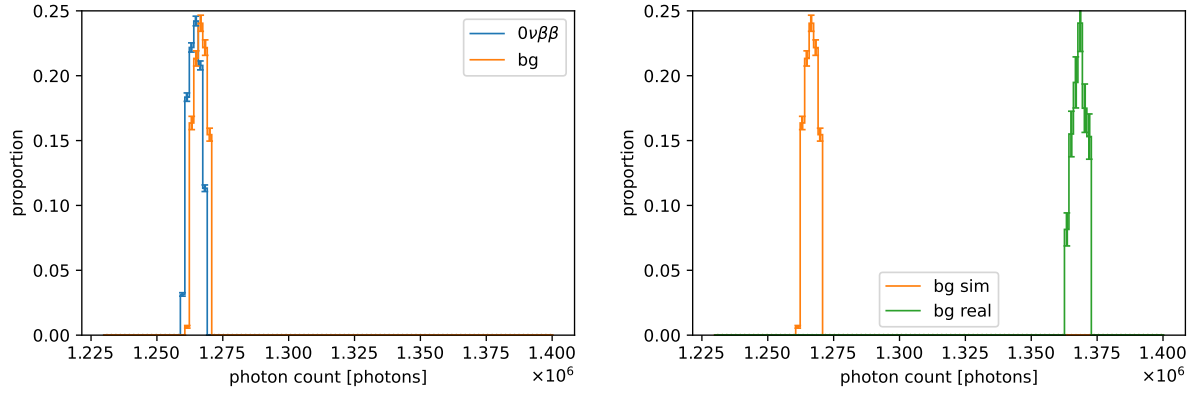


Fig. 5.4 Comparison of photon count distributions. The left panel shows the comparison with simulated $0\nu\beta\beta$ and background, while the right panel shows the comparison between simulated background and real data.

was 6.1 % smaller compared to the real data. A possible explanation for the remaining difference is that, in the real data, EL light is generated randomly during the drift within the ELCC cell, whereas in the simulation data, it is generated all at once upon reaching the ELCC. This discrepancy may have led to an excessive application of the MPPC non-linearity in the simulation, resulting in insufficient correction during the analysis. Differences in overall photon count between real and simulation data could affect machine learning; however, in our model, normalization is applied, and thus such differences are not expected to influence the results.

5.3.2 Local photon count

In the analysis of ELCC waveforms, the photon count for each hit within an event is calculated based on the pulse height and the time width. The distribution of the photon count of the highest-photon-count hit among these is shown in Fig. 5.5. This quantity is related to the spatial concentration of ionization electrons that generate the EL signal. The distribution of $0\nu\beta\beta$ is slightly larger than the background in the simulation data. In the comparison between real data and simulation, the real data tends to be larger and has a wider distribution. One of the factors is the difference in EL gain discussed in Sec. 5.3.1.

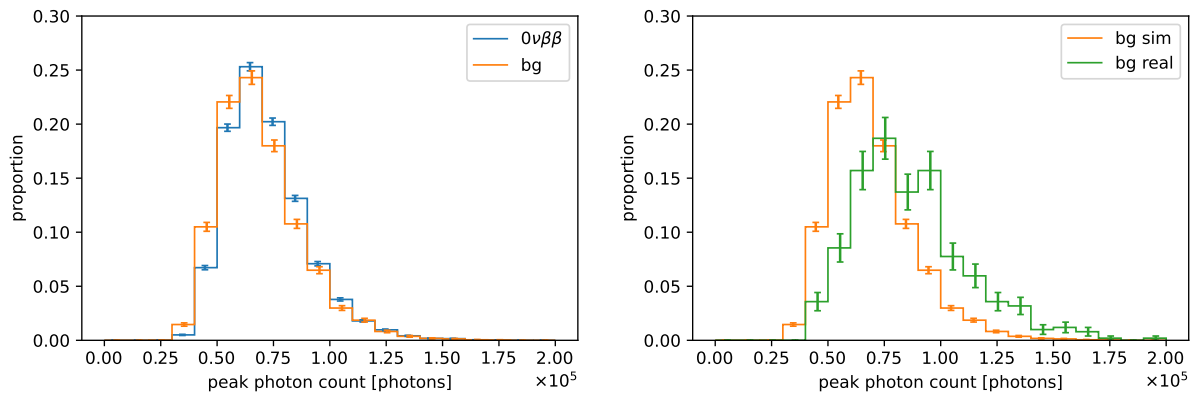


Fig. 5.5 Comparison of the photon count distributions of the highest-photon-count hits in each event. Comparison with simulated $0\nu\beta\beta$ and background are shown in left, simulated background and real data are shown in right.

Another factor is the bias introduced by the veto around the dead or high dark current channel (Fig. 4.4)

in the real data, which results in tracks being concentrated in a narrow region.

5.3.3 Number of clusters

Figure 5.6 shows the distribution of the number of clusters (Sec. 4.2.1) of events. The simulation data

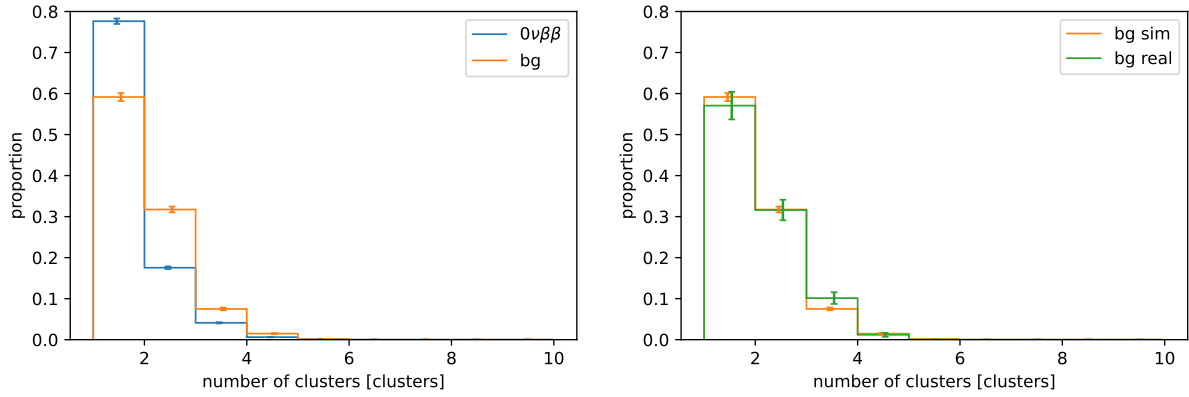


Fig. 5.6 Comparison of cluster number distributions. The left panel shows the comparison with simulated $0\nu\beta\beta$ and background, while the right panel shows the comparison between simulated background and real data.

tends to show more clusters in the background than in the $0\nu\beta\beta$ events. This is likely due to the gamma-ray background causing Compton scattering followed by photoelectric absorption, resulting in clusters forming at distant locations. Additionally, photons from bremsstrahlung create clusters at distant locations. In the $0\nu\beta\beta$, the decay energy is shared between two electrons, whereas in background events, the electron generated by photoelectric absorption carries the entire energy, making it more energetic and more likely to cause bremsstrahlung. The distribution of the number of clusters in the simulation and in the real data is relatively consistent. Figure 5.7 shows a comparison of simulation and real data without the Compton-scattering cut described in Sec. 5.2.4. Compared to Fig. 5.6, the number of clusters

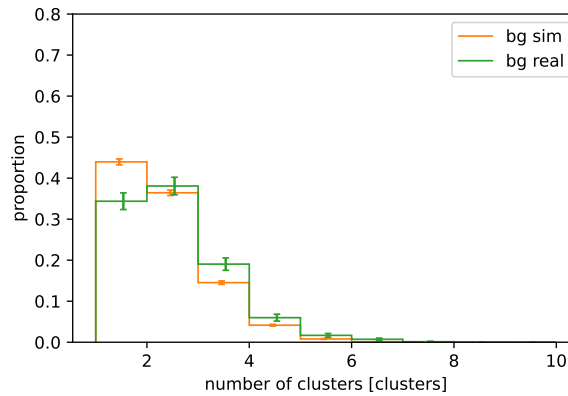


Fig. 5.7 Comparison of cluster number distributions for simulated background and real data without the Compton-scattering cut.

has increased for both simulation and real data, but the increase is greater for real data. This may be because the real data include accidental coincidence events such as those shown in Fig. 5.8. Although such events are not considered in the simulation, they are eliminated by the Compton-scattering cut.

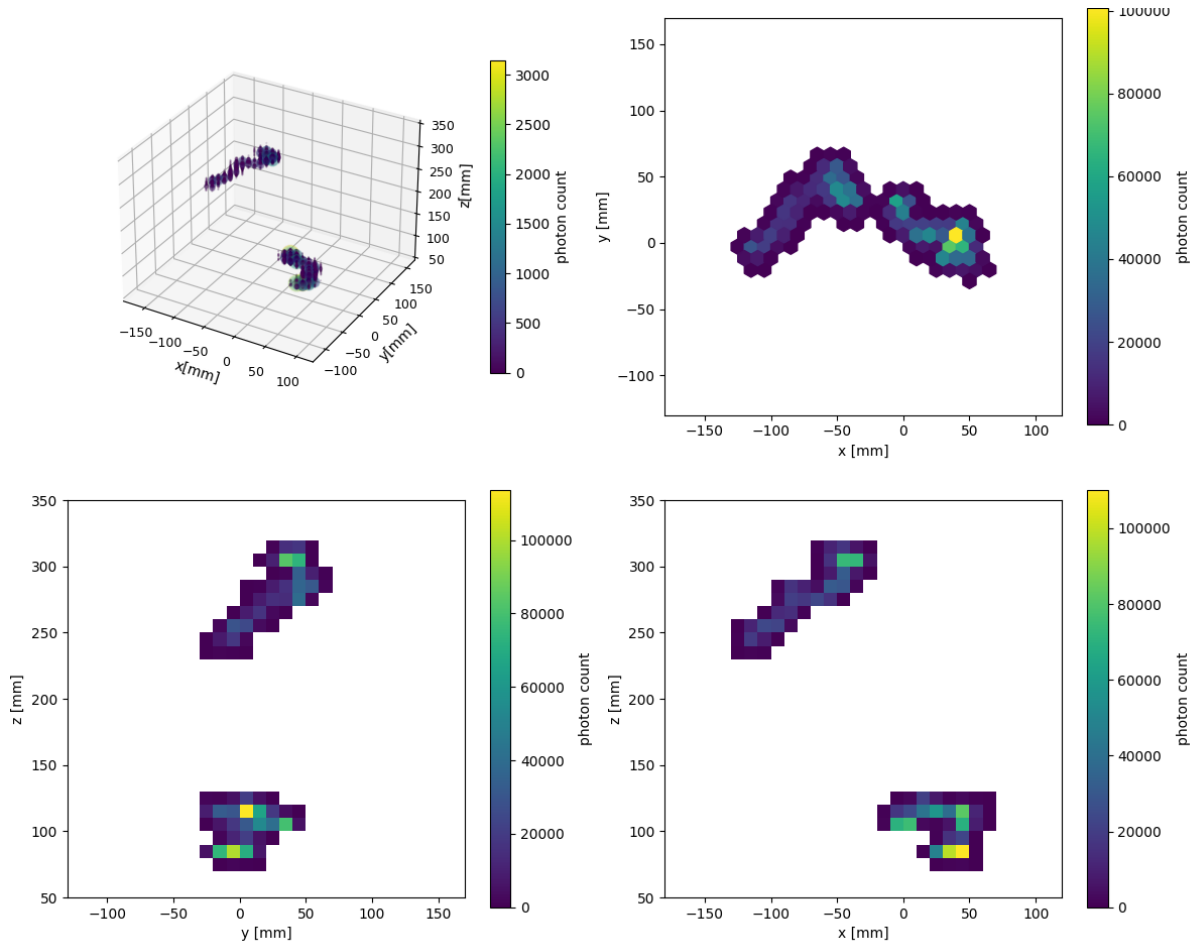


Fig. 5.8 An example track of accidental coincidence event. It is likely that the total energy of the events occurring at the same time coincidentally matches 2615 keV.

5.3.4 Track volume

Figure 5.9 shows the distribution of the number of hit channels of each event. This corresponds to the degree of spread of the tracks in the xy plane. In the simulation data, the $0\nu\beta\beta$ tends to have fewer hit

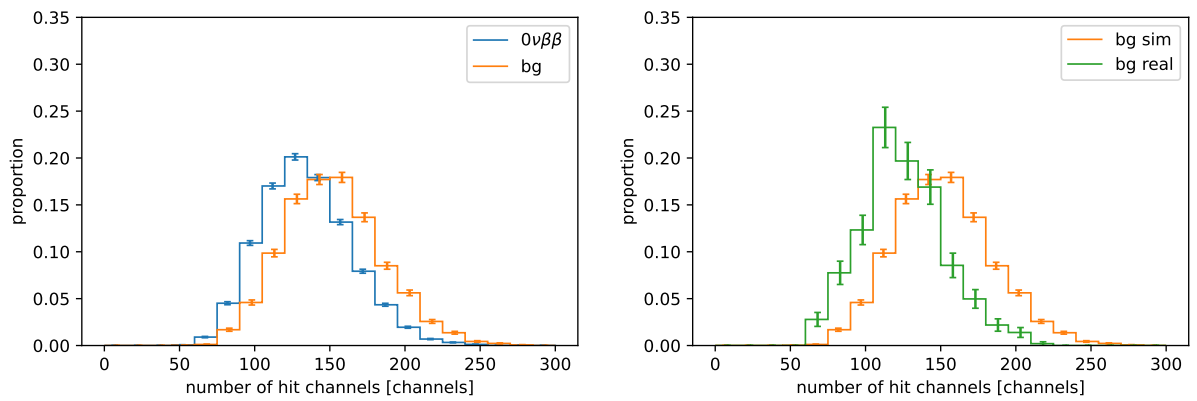


Fig. 5.9 Comparison of number of hit-channels distributions. The left panel shows the comparison with simulated $0\nu\beta\beta$ and background, while the right panel shows the comparison between simulated background and real data.

channels compared to the background. This is likely because, in the $0\nu\beta\beta$, the decay energy is divided between two electrons, resulting in shorter ranges for each electron. In the comparison between the background simulation data and the real data, the real data tends to have fewer hit channels. This is likely caused by the dead or high dark current channels as discussed in Sec. 5.3.2, which result in tracks being concentrated in a narrow region. The distribution of the real data more closely resembles that of the $0\nu\beta\beta$ simulation data, which may influence the machine learning model's classification. This issue is discussed in detail in Sec. 6.6. A similar trend can be seen in the distribution of the hit volume which is the total number of ADC clocks for all hits for an event (Fig. 5.10).

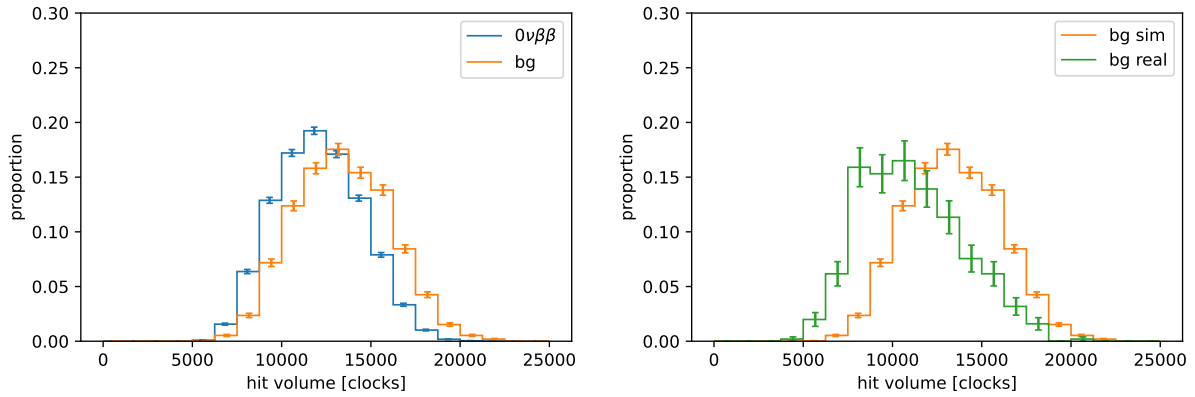


Fig. 5.10 Comparison of hit volume distributions. The left panel shows the comparison with simulated $0\nu\beta\beta$ and background, while the right panel shows the comparison between simulated background and real data.

5.3.5 Track position

For the comparison of track position, following two quantities are considered.

Photon-count weighted centroid position

Figure 5.11, 5.12, 5.13 show the distributions of the photon-count weighted centroid position of hits in the x, y, z direction, respectively. The difference in the x direction between the real and simulated data

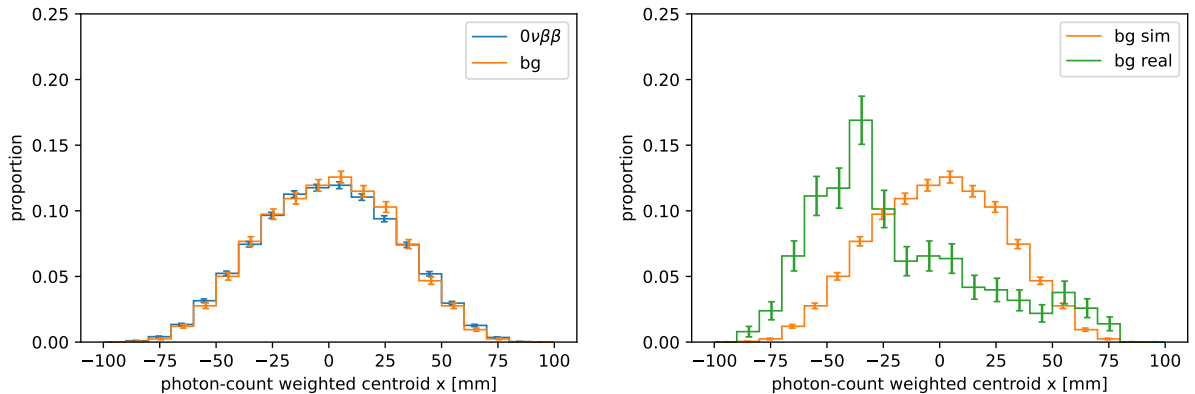


Fig. 5.11 Comparison of distributions for photon-count weighted centroid of hits in the x-direction. The left panel shows the comparison with simulated $0\nu\beta\beta$ and background, while the right panel shows the comparison between simulated background and real data.

is due to positional constraints imposed by the veto pattern settings around dead channels and high dark current channels, as described in Sec. 5.3.2. The slight positive bias of the y direction of the background

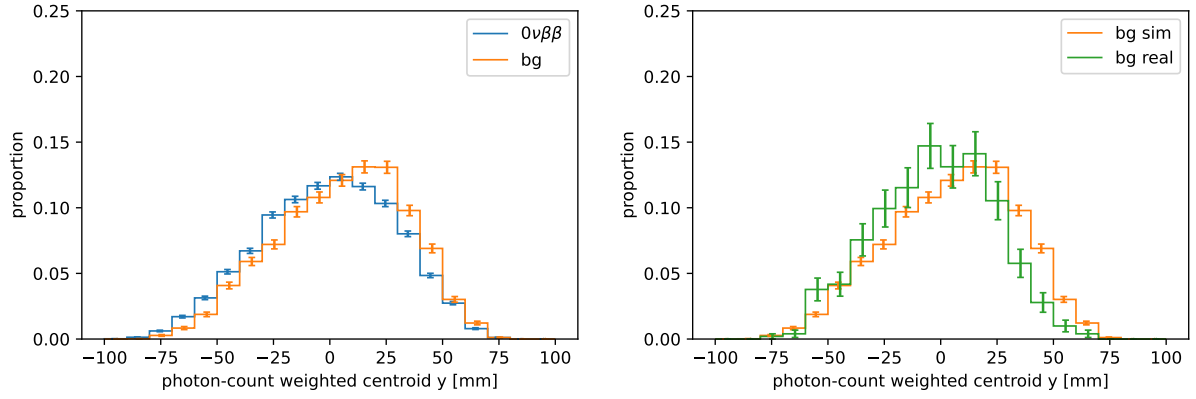


Fig. 5.12 Comparison of distributions for photon-count weighted centroid of hits in the y-direction. The left panel shows the comparison with simulated $0\nu\beta\beta$ and background, while the right panel shows the comparison between simulated background and real data.

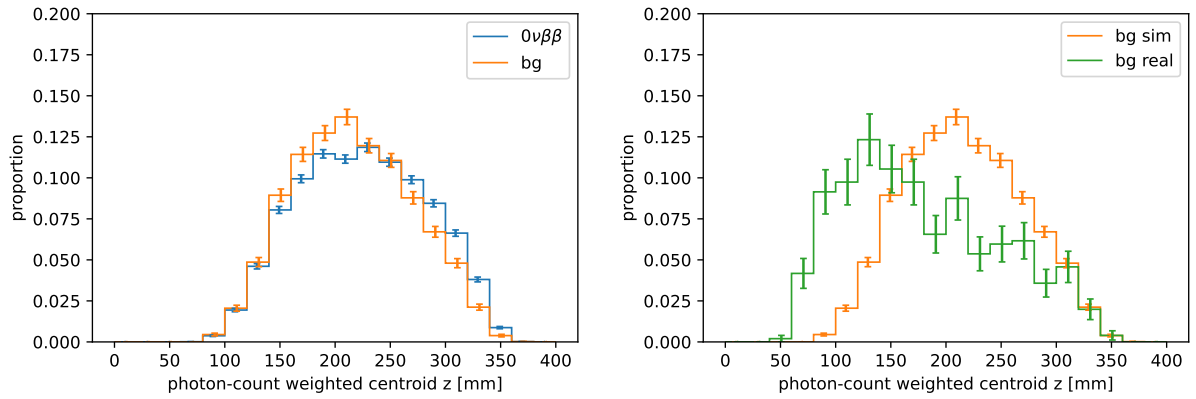


Fig. 5.13 Comparison of distributions for photon-count weighted centroid of hits in the z-direction. The left panel shows the comparison with simulated $0\nu\beta\beta$ and background, while the right panel shows the comparison between simulated background and real data.

in the simulation data may be due to the fact that the generation of γ -ray events is not isotropic, but rather vertically downward (Sec. 5.2.1). The z direction tends to be smaller in the real data than in the simulated data. This may be due to the low detection efficiency of scintillation photon in the real data, which may lead to a miss reconstruction of the z position.

Rise and fall z position

Rise and fall z -positions refer to the z -position of the earliest and latest hits in the event, respectively. Figure 5.14, 5.15, 5.16 show the distribution of the rise, fall position and its difference, respectively. The difference between $0\nu\beta\beta$ and background in the simulation data may reflect the difference in track volume described in Sec. 5.3.4. In comparison with the real data, both rise and fall positions tend to be smaller in the real data, but the difference between rise and fall is not so large. This also suggests the z -misreconstruction in the real data. The difference between fall and rise position is possibly due to the track size bias caused by veto. The distribution of the real data is closer to that of the $0\nu\beta\beta$ rather than the simulation background, which could potentially affect the machine learning model's discrimination.

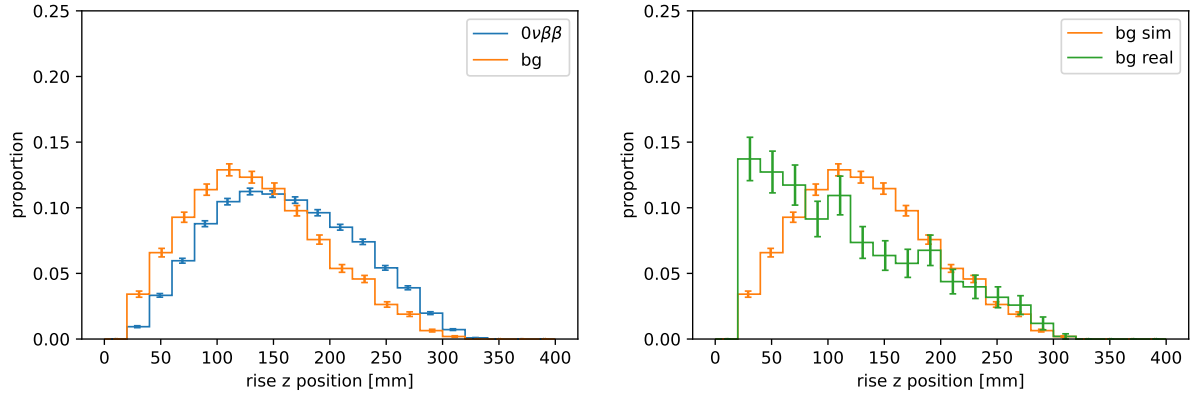


Fig. 5.14 Comparison of rise positions in z direction. The left panel shows the comparison with simulated $0\nu\beta\beta$ and background, while the right panel shows the comparison between simulated background and real data.

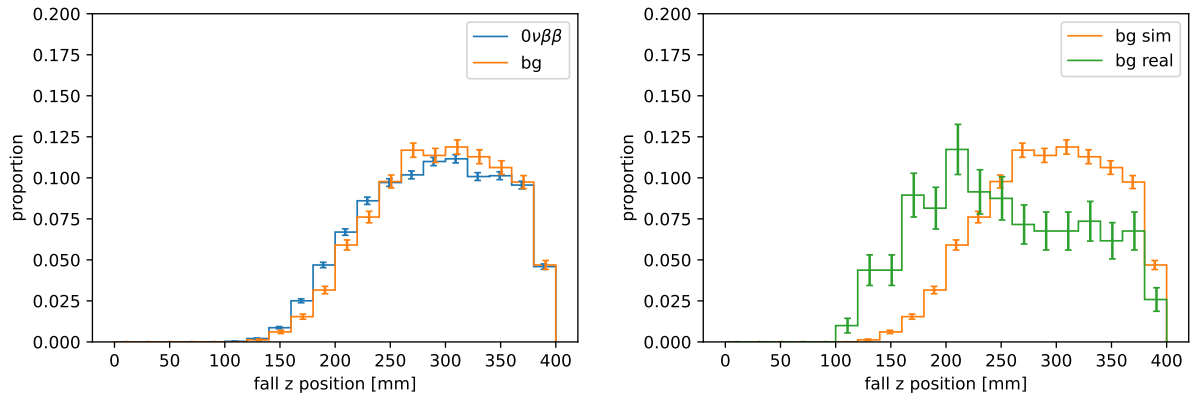


Fig. 5.15 Comparison of fall positions in z direction. The left panel shows the comparison with simulated $0\nu\beta\beta$ and background, while the right panel shows the comparison between simulated background and real data.

5.3.6 Summary

The comparison between the $0\nu\beta\beta$ and background in the simulation data is summarized as follows. The $0\nu\beta\beta$ data showed a tendency to have fewer clusters compared to the background. This is likely due to the Compton effect and bremsstrahlung. In the $0\nu\beta\beta$ data, the energy is divided between two electrons, resulting in shorter ranges for each electron compared to the single electron in the background, leading to a tendency for smaller track sizes.

The comparison between the simulated background data and the real ^{208}Tl data is summarized as follows. Due to differences in simulation settings, particularly EL gain, there is a 7.4% difference in the overall photon counts. Accidental coincidence events, which are not considered in the simulation, are eliminated by the Compton-scattering cut, resulting in a good match between the cluster number distributions in the simulation and real data. However, the track sizes tend to be smaller in the real data, which is likely due to the bias introduced by the veto around the dead or high dark current channels in the actual data. Regarding the z-position of the tracks, the actual data shows a tendency to be smaller, which may be influenced by mis-reconstruction in z-position due to the low scintillation detection efficiency.

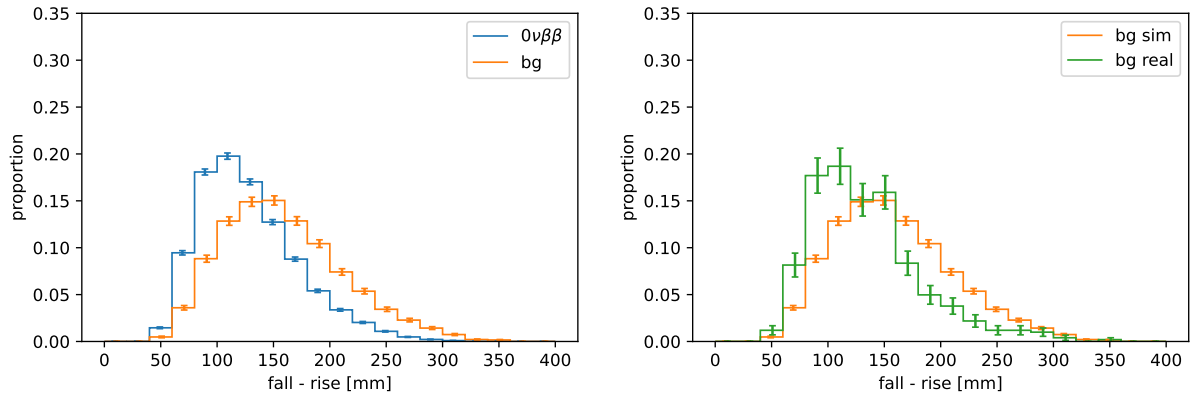


Fig. 5.16 Comparison of difference of fall and rise positions in z direction. The left panel shows the comparison with simulated $0\nu\beta\beta$ and background, while the right panel shows the comparison between simulated background and real data.

Chapter 6

Track pattern recognition by machine learning

In this chapter, we construct a machine learning model using the simulation data generated in the previous chapter. We begin by describing the preprocessing steps applied to the data before feeding it into the model. Next, we explain the model architecture used in this study, along with the training procedure and evaluation metrics. Finally, we present the training results for several model variations. The characteristics of events where the model's classification does not work well are also investigated, and the reasons for the differences between the actual data and the validation data are considered.

6.1 Preprocessing

For each of the selected events, the rectangular volume containing the track is divided into appropriate lattice voxels and the corresponding photon count, or energy deposit of the track in other words is assigned to each voxel. This information is fed into the machine learning model. To assign the photon counts to the voxels, additional transformations are performed as follows.

6.1.1 Geometrical transformation

ELCC cells are arranged in a hexagonal lattice, so voxelization with a square lattice would result in unequal contributions to each voxel. Therefore, the layout of the ELCC cell is converted to a square lattice shape by a skew transform A_{skew} and y-directional expansion $A_{y-\text{exp}}$ in advance.

$$A_{\text{skew}} = \begin{pmatrix} 1 & \tan \frac{\pi}{6} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad A_{y-\text{exp}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2/\sqrt{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (6.1)$$

Figure 6.1 shows an example of a track projected in the x-y plane (left) before and (right) after the transformation, respectively. This transformation allows the original track information to be preserved by voxelizing it into a grid with 10 mm intervals in the xy direction. Although this transformation alters the distances between individual track points, the relative positions of adjacent cells are preserved. As described in Sec. 6.2.2, the model used in this study employs operations that incorporate spatially adjacent information but do not rely on the actual distances between points. Therefore, applying this transformation is considered to be unproblematic.

6.1.2 voxel binning

The original data set has a resolution of 10 mm, the pitch of the ELCC, in the xy direction. The resolution in the z direction is determined by the ADCL sampling interval and the drift velocity of the ionization electrons, which is approximately 0.2 mm. The ionization electrons diffuse during drift, with diffusion

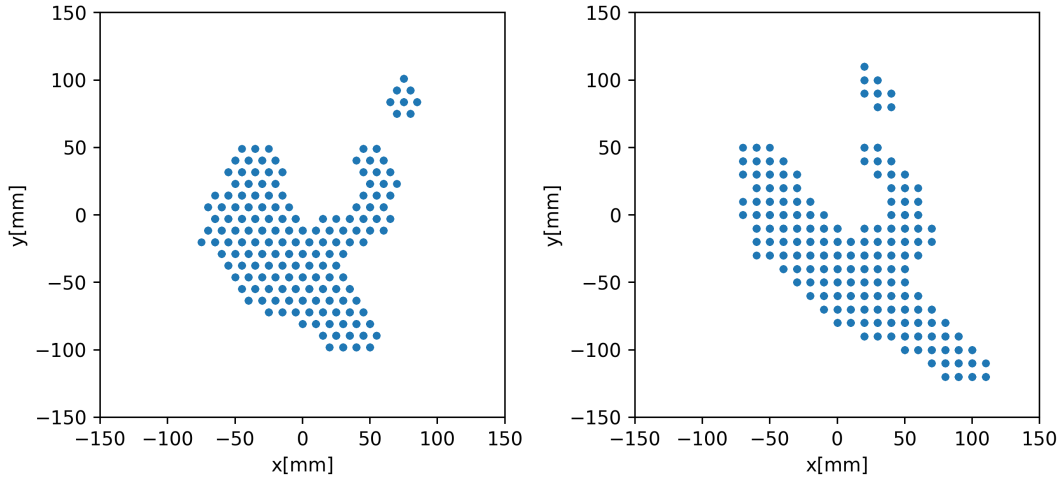


Fig. 6.1 An example of a track projected to the x-y plane before (left) and after (right) transformation respectively.

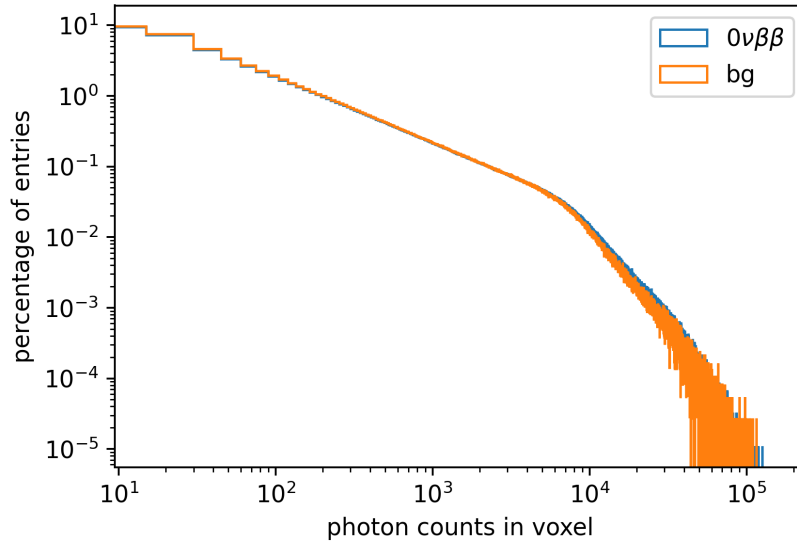


Fig. 6.2 PDF of photon counts in voxels with at least one photon.

coefficients of $0.118 \text{ cm}/\sqrt{\text{cm}}$ in the xy direction and $0.033 \text{ cm}/\sqrt{\text{cm}}$ in the z direction. Due to the lack of GPU memory, all information in the z direction cannot be used and must be merged at coarser intervals. In this study, segmentation into voxels was performed at a pitch of 10 mm in the xy direction and in the z direction in a range of 2.5 mm to 10 mm, depending on the model.

6.1.3 Photon count transformation

Figure 6.2 shows the probability density function (PDF) of the photon counts in voxels with at least one photon, based on training data voxelized at 10 mm intervals for z direction. Considering the ELCC plane and the 40 cm drift region of the 180 L prototype detector, the number of voxels in the x , y and z directions was set to $(x, y, z) = (30, 36, 40)$. In both $0\nu\beta\beta$ and background events, the maximum photon count of a voxel is about 10^5 photons. Photon counts less than 1/1000 of the maximum value occur with

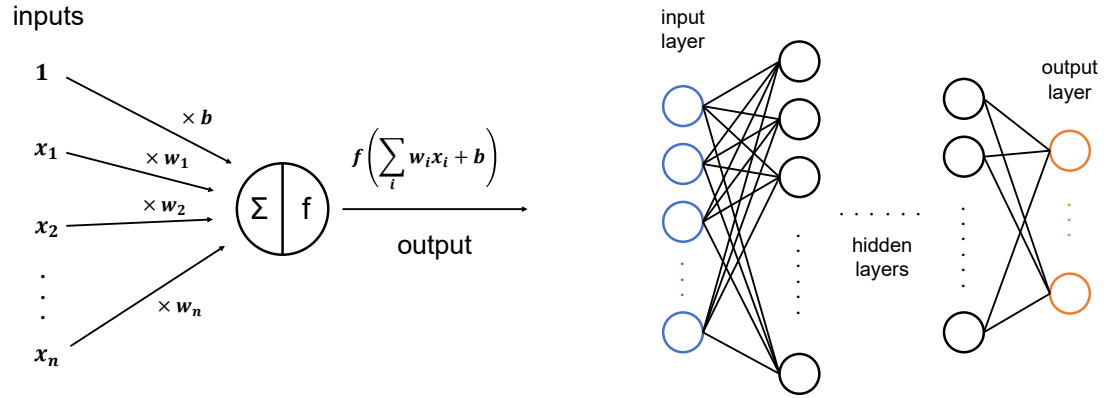


Fig. 6.3 A diagram of a neuron (left) and multilayer perceptron (right).

an approximately 100 000 times higher rate. We scaled the photon counts to account for the possibility that these relatively small values may be difficult to maintain in the neural network. The scaling pattern is as follows,

- (1) identical
Does not scale photon counts.
- (2) normalize
Normalize to the range $[0,1]$ by dividing the photon counts by the maximum value in the event.
- (3) digitize
Divide from 1 photon to the maximum number of photons in the event by 10 on the linear scale. The value of each voxel is converted to $[0, 0.1, \dots, 1.0]$.
- (4) log-digitize
Divide from 1 photon to the maximum number of photons in the event by 10 on the log scale. The value of each voxel is converted to $[0, 0.1, \dots, 1.0]$.
- (5) exponential
Let the (i,j,k) -th voxel value be P_{ijk} . For each non-zero voxel, convert P_{ijk} to $\exp(P_{ijk}/\max(P))/e$.

The latter three patterns scale small values: pattern 3 and 4 discretize the values, while pattern 5 smoothly scales the values to the range $(1/e, 1]$.

6.2 Neural network

In this section, we describe the learning principles of the neural network and the settings for the base model, loss function, and optimizer.

6.2.1 Principal of neural network

Neural networks are networks that mimic the neurons of the brain. A diagram of a neuron and a multilayer perceptron (MLP), a basic neural network, are shown in Fig. 6.3. Given an input x_i the neuron computes the sum of its inputs, each with a different weight w_i and a bias b . Then, it outputs $f(\sum_i w_i x_i + b)$ with a nonlinear “activation function” f . MLPs are composed by combining many such neurons. A certain number of neurons make up one layer, and the output of each neuron is the input for the next layer. The layer between the input and output is called the “hidden layer”. The hidden layers and the connections between each neuron allow for complex representations. By adjusting the weights appropriately, the

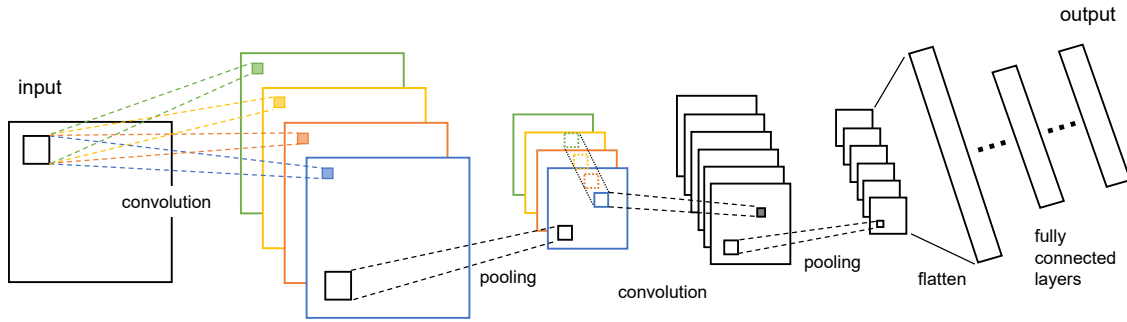


Fig. 6.4 A conceptual diagram of a CNN. The first convolution and pooling layer is colored corresponding to each filter.

mapping F between various patterns of the input vector $\vec{x}$ and the output vector $\vec{y}$ can be approximated.

Model learning is done by iteratively adjusting weights. Some data reflecting the distribution of the inputs and target values to be approximated are used as the training data. The weights are adjusted to minimize the “loss”. The loss is calculated by comparing the output and the known target value using a dedicated function. Since it is generally impossible to input all training data into the network at once due to lack of memory capacity, a certain number of randomly selected data is input as a “mini-batch”. The weights are updated for each mini-batch, and one round of training using all the data is called an “epoch”. If training is performed correctly, the loss on the training data decreases with each epoch. Since the model is trained to reproduce the inputs and target values of the training data, it must be checked to ensure that it approximates the correct target values for data other than the training data. For this reason, validation data are prepared separately from the training data, and performance is evaluated using this validation data. If the model is trained to merely reproduce the specific input-target relationships of the training data, the loss on the validation data will not decrease and may even increase with each epoch. This phenomenon is known as overfitting. Common countermeasures include increasing the size of the training dataset, as well as applying techniques such as dropout[84].

6.2.2 Convolutional Neural Network

In a MLP, each neuron is independent. In contrast, a network can be created that takes into account the relationships that neighboring pixels have with each other, such as in an image. The convolutional neural network (CNN) is one of its kind. A conceptual diagram of a CNN for a two-dimensional image is shown in Fig. 6.4. In the convolutional layer, weights with size $n \times m$, called a filter or a kernel, are multiplied and summed with the input while shifting its positions. This allows the information from adjacent pixels to be incorporated. The pooling layer reduces the spatial dimensions of the data by computing either the average or the maximum value within a specified input window. This makes the output robust to minute changes in position. The features extracted through repeated convolution and pooling operations are eventually flattened into a one-dimensional vector and fed into a fully connected network, similar to an MLP.

6.2.3 DenseNet

The performance of the network improves by increasing the depth of the layers, but if it becomes too deep, the gradient magnitude for updating the weights diminishes, making learning impossible. ResNet[85] solves this problem by introducing a direct skip connection between the layers, and has achieved a

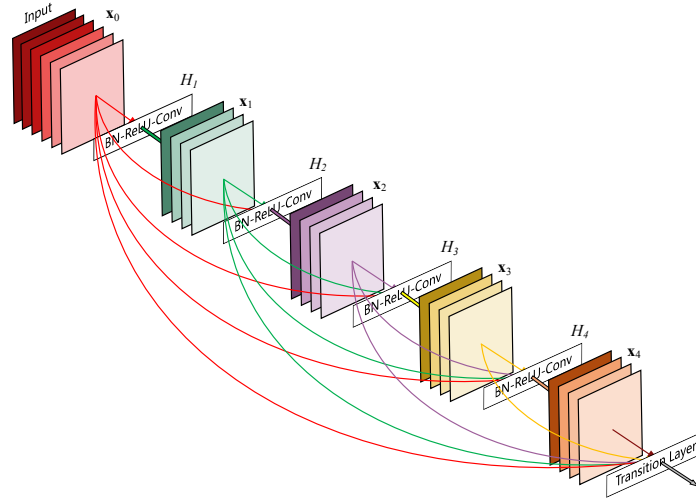


Fig. 6.5 A 5-layer dense block with a growth rate of $k = 4$. Figure from [86].

performance improvement of more than 100 layers deep. DenseNet[86] is a type of convolutional neural network that further develops ResNet and introduces a structure called “Dense Block” that connects a layer to all other layers (Fig. 6.5).

$$\mathbf{x}_l = H_l([\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{l-1}]) \quad (6.2)$$

Here, $H_l(\cdot)$ is a non-linear transformation to l -th layer, $\mathbf{x}_i$ represents the feature-maps of the i -th layer, and in particular $\mathbf{x}_0$ represents the input image. $[\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{l-1}]$ represents the concatenation of the feature-maps produced in layers $0, \dots, l-1$. If the input has k_0 features and k features are generated at each layer, then the input at the l -th layer has $k_0 + k \times (l-1)$ feature maps. This parameter k is called “growth rate” of the network. To facilitate down-sampling for the feature maps, there are “transition layers” between dense blocks. The transition layer consists of a batch normalization layer (Sec. 6.2.4), followed by a convolutional layer with a kernel size of 1×1 , and an average pooling layer with a 2×2 window. The output size of feature maps is reduced to $\lfloor \theta m \rfloor$, where m is the number of input feature maps and $0 < \theta < 1$ is the reducing parameter. We followed Densenet-121 and set $\theta = 0.5$.

In [86], several networks are proposed depending on the number of layers and growth rate. In this study, the lightest model, DenseNet-121, was used as a base model. The changes from Densenet-121 are as follows.

- (1) 2-D convolution changed to 3-D convolution.
- (2) Removed the first 7×7 convolution and 3×3 max pooling layers.
- (3) Change growth rate from 32 to 12.
- (4) In the dense block and transition layer, the BatchNormalization[87] layer was changed to the InstanceNormalization[88] layer.

The removal of the first convolution and pooling layer, (2), was taken considering the difference in size of the input data. In [86], image of 224×224 size is assumed as input. In this study, since the xy direction has only 30×36 resolution, we decided not to reduce the size by convolution and pooling in the input part. The lower growth rate, (3), is due to GPU memory limitations. The change in the normalization method, (4), is related to the stability of the training and will be discussed in detail in the next section.

6.2.4 Normalization

During training, the distribution of inputs to each layer changes as a result of updates to the model's weight parameters. This is called the internal covariance shift and makes the training process slower. The normalization layer is inserted before an activation function. The normalization layer normalizes the mean and the variance according to a specific axis of the input data. This is said to improve learning speed and also behave as regularization[87].

Batch normalization is performed by the following formula,

$$y = \frac{x - E[x]}{\sqrt{\text{Var}[x] + \epsilon}} * \gamma + \beta \quad (6.3)$$

where E and Var are the mean and the variance of the distribution of the input along the spatial and mini-batch axes. In other words, if the input mini-batch data shape are represented by a vector (n, c, d, h, w) where n is the index of the events, c of the feature maps, d, h, w of each spatial axes respectively,

$$E[x]_c = \frac{1}{NDHW} \sum_{n=1}^N \sum_{d=1}^D \sum_{h=1}^H \sum_{w=1}^W x_{ncdhw}, \quad \text{Var}[x]_c = \frac{1}{NDHW} \sum_{n=1}^N \sum_{d=1}^D \sum_{h=1}^H \sum_{w=1}^W (x_{ncdhw} - E[x]_c)^2 \quad (6.4)$$

where N, D, H and W are the size of the corresponding quantity. β and γ are used to scale and shift the normalized output and have the size of the feature maps C . They were set to $\beta = 0$ and $\gamma = 1$ for this study. ϵ is set to small value 1×10^{-5} to prevent dividing by zero. Batch normalization includes the events axis n in the mean and variance calculations, thus requiring the use of mean and variance information across the dataset rather than single data. Then, moving averages of the mean and variance obtained from the input data during training are used as an alternative to the mean and variance for the entire dataset.

We initially used batch normalization according to the Densenet-121 implementation, but the accuracy for validation data varied by several tens of percent from epoch to epoch, making stable learning impossible. This may be due to the fact that the mean and variance stored in the batch normalization layer are not properly representative of the data set because the mini-batch size was relatively small (8 to 32) in this study (Sec. 6.4.2) due to lack of GPU memory.

To avoid this learning instability, we employed instance normalization. The formula for instance normalization is the same as for batch normalization (Eq. 6.3), but processing in the events axis n is not performed.

$$E[x]_{nc} = \frac{1}{DHW} \sum_{d=1}^D \sum_{h=1}^H \sum_{w=1}^W x_{ncdhw}, \quad \text{Var}[x]_{nc} = \frac{1}{DHW} \sum_{d=1}^D \sum_{h=1}^H \sum_{w=1}^W (x_{ncdhw} - E[x]_{nc})^2 \quad (6.5)$$

As a result, the mean and variance are computed per event, eliminating the need to estimate these statistics over the entire dataset. Therefore, the moving average of the mean or variance of the data set is no longer calculated, but is obtained from the input data each time, both during training and evaluation.

6.2.5 Loss function

A loss function represents the difference between the model output and the target value, and the model is trained to minimize the loss for the training data. In classification problems, the target value represents the label of the correct class. In this study, we used cross entropy loss, which is commonly used for

classification problems. The loss for a single data of index n is given by,

$$l_n = -w_{y_n} \log \left(\frac{\exp(x_{n,y_n})}{\sum_{c=0}^{C-1} \exp(x_{n,c})} \right) \quad (6.6)$$

Here, C is the number of classes to be identified, in this case, $C = 2$; single track (background) or two tracks ($0\nu\beta\beta$). x_n is the output of the model, whose size is equal to the number of classes. y_n is the target value, in this case, 0 for the single track and 1 for the two tracks. Hence, for example, $x_{n,0}$ means the output value of the model for the single track class on the n -th event. The weight w_{y_n} is used to compensate for the contribution to the loss function in cases where there is a significant imbalance in the amount of data between classes. The term in the log function is a “softmax function”, which converts the output of the model to probability values for the class. This loss is calculated for each mini-batch of size N , and mean of the losses are used to update the model parameters.

$$l(x, y) = \sum_{n=1}^N \frac{1}{\sum_{n=1}^N w_{y_n}} l_n \quad (6.7)$$

6.2.6 Optimizer

An optimizer updates the model parameters according to a specific algorithm to reduce losses. A number of parameter update algorithms have been proposed. Here, some of the optimizers are described below according to the implementation of PyTorch[89].

The simplest algorithm is stochastic gradient descent (SGD), whose algorithm is as follows,

Algorithm 1 SGD

Input: $\gamma, \theta_0, f(\theta)$

Output: θ_t

$t \leftarrow 0$

while θ_t not converged **do**

$t \leftarrow t + 1$

$g_t \leftarrow \nabla_{\theta} f_t(\theta_{t-1})$

$\theta_t \leftarrow \theta_{t-1} - \gamma g_t$

end while

return θ_t

where γ is the learning rate and determines the scale at which the parameters are updated, θ_0 is the initial model parameters, and $f(\theta)$ is the function to be minimized and corresponds to the loss function.

SGD is simple, but convergence is slow if the gradient g_t is not oriented toward the true minimum of $f(\theta)$. One method to solve this problem is “momentum”[90], which adds a weighted past gradient to the current gradient.

Algorithm 2 SGD+momentum**Input:** $\gamma, \mu, \theta_0, f(\theta)$ **Output:** θ_t

```

 $t \leftarrow 0$ 
while  $\theta_t$  not converged do
   $t \leftarrow t + 1$ 
   $g_t \leftarrow \nabla_{\theta} f_t(\theta_{t-1})$ 
  if  $t > 1$  then
     $b_t \leftarrow \mu b_{t-1} + g_t$ 
  else
     $b_t \leftarrow g_t$ 
  end if
   $g_t \leftarrow b_t (*)$ 
   $\theta_t \leftarrow \theta_{t-1} - \gamma g_t$ 
end while
return  $\theta_t$ 

```

A slightly different version of momentum is the Nesterov's Accelerated Gradient[91] (NAG). The algorithm of NAG is a transformation of the $(*)$ line of Alg. 2 into $g_t \leftarrow g_t + \mu b_t$.

In addition to these algorithms, methods that adaptively vary the amount of parameter updates have been proposed, one of which is Adam[92]. The algorithm of the Adam optimizer is as follows,

Algorithm 3 Adam**Input:** $\gamma, \beta_1, \beta_2, \theta_0, f(\theta)$ **Output:** θ_t

```

 $m_0 \leftarrow 0$ 
 $v_0 \leftarrow 0$ 
 $t \leftarrow 0$ 
while  $\theta_t$  not converged do
   $t \leftarrow t + 1$ 
   $g_t \leftarrow \nabla_{\theta} f_t(\theta_{t-1})$ 
   $m_t \leftarrow \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t$ 
   $v_t \leftarrow \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot g_t^2$ 
   $\widehat{m}_t \leftarrow m_t / (1 - \beta_1^t)$ 
   $\widehat{v}_t \leftarrow v_t / (1 - \beta_2^t)$ 
   $\theta_t \leftarrow \theta_{t-1} - \gamma \cdot \widehat{m}_t / (\sqrt{\widehat{v}_t} + \epsilon)$ 
end while
return  $\theta_t$ 

```

Here m_t and v_t is a exponential moving averages of the gradient and the squared gradient, respectively. These moving averages are estimates of the 1st moment and the 2nd raw moment of the gradient. The moving averages are initialized as 0, so moments are biased towards zero. $\widehat{m}_t$ and $\widehat{v}_t$ are the bias-corrected estimates. $|\widehat{m}_t / \widehat{v}_t|$ is a quantity called *signal-to-noise* ratio (SNR) that indicates the indeterminacy of the direction of the gradient and approaches 0 near the optima. Therefore, scaling the learning rate γ by SNR facilitates convergence.

We used Adam optimizer with parameters $\gamma = 1 \times 10^{-3}$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\epsilon = 1 \times 10^{-8}$. These values are set to those of the original paper. We also tried SGD+momentum and NAG, but did not adopt them due to slow learning convergence.

6.3 Valuation index

The combination of model predictions and true labels is represented by the confusion matrix in Tab. 6.1.

Positive and Negative represent $0\nu\beta\beta$ and background, respectively. We use the following two metrics,

Table 6.1 Confusion matrix. The abbreviations used in the table are as follows: TP (true positive), FN (false negative), FP (false positive), and TN (true negative).

		Prediction	
		Positive	Negative
Label	Positive	TP	FN
	Negative	FP	TN

accuracy and receiver operating characteristic curve, to compare model performance.

6.3.1 Accuracy

The accuracy is the percentage of correctly classified events and is expressed by the following formula,

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \quad (6.8)$$

Prediction of the model is positive if the probability of being a signal exceeds 0.5.

6.3.2 Receiver operating characteristic (ROC) curve

The accuracy is simple, but we need to fix the threshold at which we determine a signal. We would like to set a threshold value that makes the best trade-off between signal efficiency and background event rejection efficiency. Therefore, we introduce a comprehensive evaluation metric for the threshold value. A receiver operating characteristics (ROC) is a curve plotting false positive rate (FPR) and true positive rate (TPR) shown below, varying the signal threshold from 0 to 1.

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}} = 1 - \frac{\text{TN}}{\text{FP} + \text{TN}} \quad (6.9)$$

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (6.10)$$

FPR is also interpreted as (1 - background rejection ratio), as can be seen from Eq. 6.9. TPR means signal efficiency. The ROC curve is a monotonically increasing curve from (0,0) to (1,1), with the curve moving toward the upper left indicating better model performance. Therefore, the area under the ROC curve is used as an evaluation index as Area Under the Curve (AUC) (Fig. 6.6).

6.4 Training the model

6.4.1 Photon scaling

Fig. 6.7 shows the loss and accuracy as functions of epochs for each photon scaling method in Sec. 6.1.3. The solid line represents the values for the training data, while the dashed line corresponds to those for the validation data. After around the 10th epoch, the loss on the validation data begins to increase with each epoch, indicating the onset of overfitting. Table 6.2 shows the result of training for each method of photon scaling. In this study, number of z division is 40 and size of mini-batch is fixed for 32. The accuracy and ROC AUC are for validation data and evaluated at epoch with minimum loss. Scaling to compensate for small photon counts (digitize, log-digitize, exponential) are all less accurate than

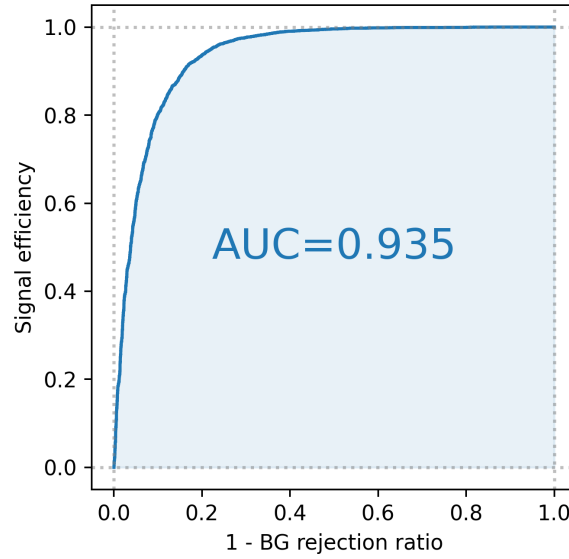


Fig. 6.6 An example of ROC curve.

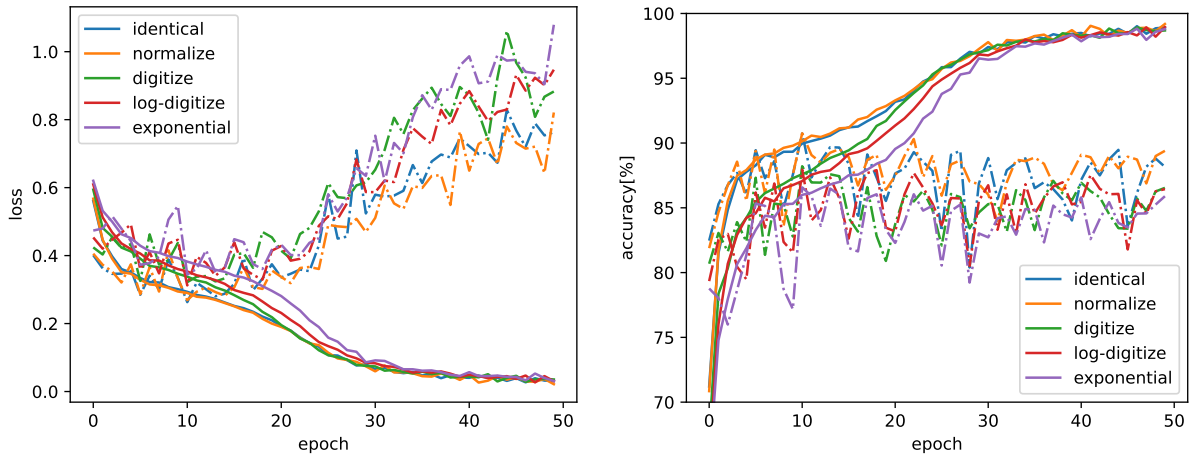


Fig. 6.7 Loss (left) and accuracy (right) as functions of epochs for each photon scaling method. The solid line represents the values for the training data, while the dashed line corresponds to those for the validation data.

simple normalization or identical. Although no significant difference is observed between the identical and normalization methods, applying normalization removes information about the total photon count in each track. As a result, the energy-based cut and the topological cut performed by the machine learning model become independent. Therefore, the scaling method is fixed to “normalize” and the number of z divisions are varied for further optimization.

6.4.2 Number of divisions of z

Fig. 6.8 shows the loss and accuracy as functions of epochs for each z division. As in the previous section, the loss on the validation data begins to increase after around the 10th epoch. However, the increase in loss becomes more gradual as the number of divisions in the z direction increases. This suggests that increasing the number of divisions allows the model to learn more detailed information about the tracks, thereby extending the training cycle before overfitting occurs and enabling the acquisition of more

Table 6.2 Training result for each scaling method.

		epoch	loss	accuracy	ROC AUC
Scaling method	identical	10	0.263	90.7 %	0.925
	normalize	10	0.269	90.8 %	0.925
	digitize	10	0.322	88.2 %	0.901
	log-digitize	10	0.312	87.9 %	0.904
	exponential	10	0.344	86.6 %	0.884

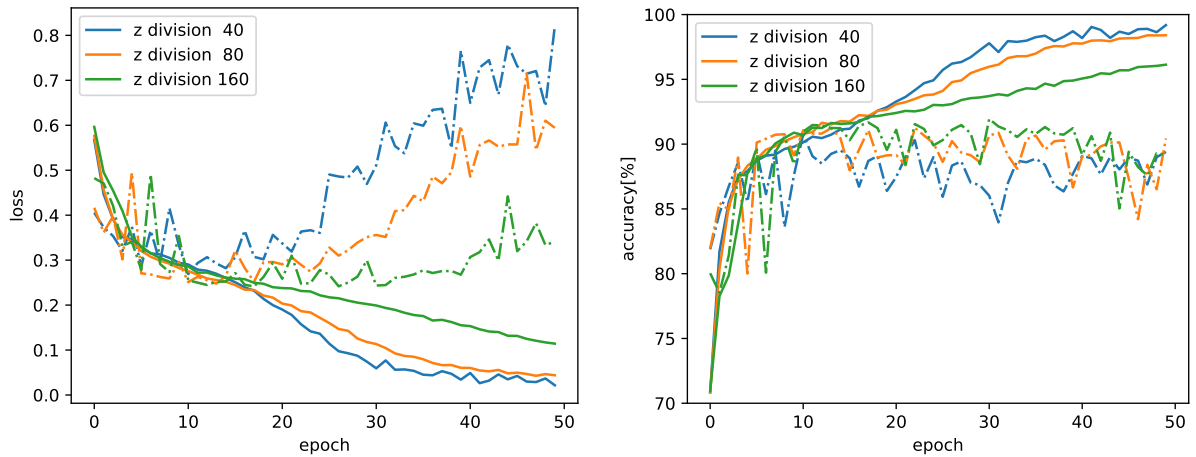


Fig. 6.8 Loss (left) and accuracy (right) as functions of epochs for each z division. The solid line represents the values for the training data, while the dashed line corresponds to those for the validation data.

generalizable features. Table 6.3 shows the result of training for each number of z divisions. In this study,

Table 6.3 Training result for each number of z division.

		epoch	loss	accuracy	ROC AUC
Number of z division	40	10	0.269	90.8 %	0.925
	80	13	0.248	91.2 %	0.933
	160	17	0.241	91.7 %	0.936

three patterns of division are considered: 40 (1 cm), 80 (0.5 cm), and 160 (0.25 cm). Further division was not performed, as it resulted in a significant increase in computation time and the calculations could not be continued due to hardware failure. The model with 40 divisions is the normalize model of the previous section. Since increasing the number of divisions increases the size of voxels, the mini-batch sizes are set to 32, 16, and 8, respectively, due to memory limitations. Increasing the number of divisions increases the number of epochs required for training, but also improves accuracy: 91.7% with 160 divisions. Given the minimal difference between the 80-division and 160-division models, a DeLong's test[93][94] was conducted to compare their AUCs. The resulting p-value was 0.228, indicating no statistically significant

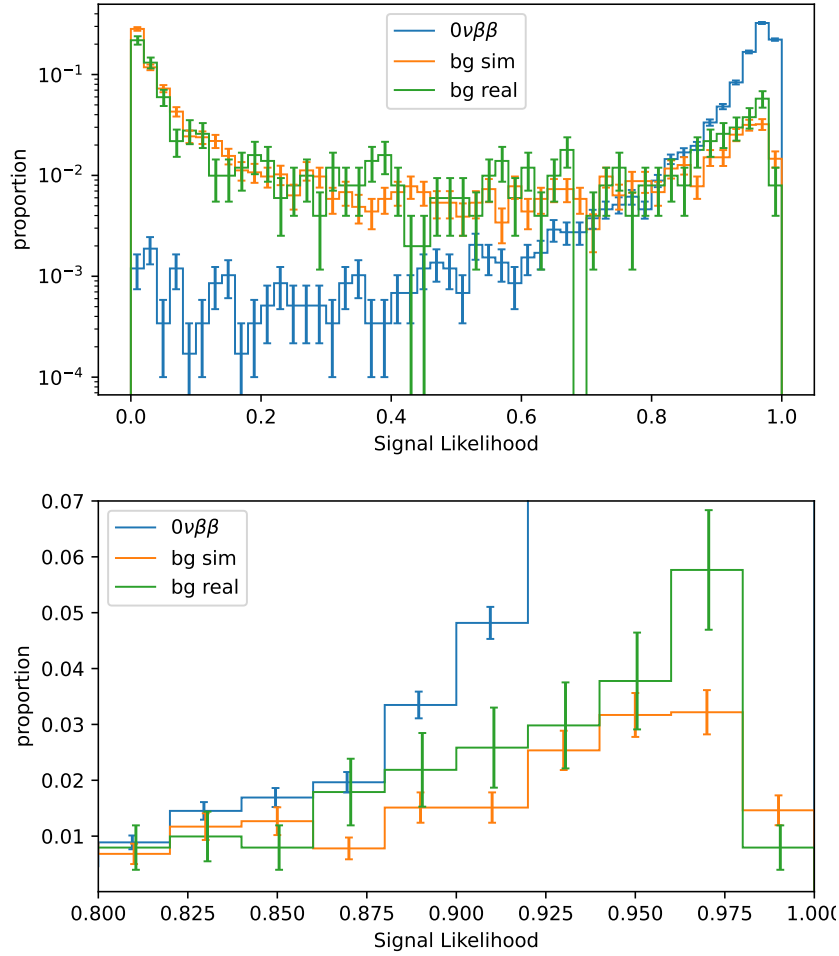


Fig. 6.9 The likelihood distribution of simulation and real data (top) and its zoom up view (bottom). Histograms are normalized so that the total of all bins equals one. The blue and orange lines are for the simulation $0\nu\beta\beta$ and background, respectively, while the green line for the real data.

difference and suggesting that increasing the number of divisions does not provide a clear advantage. In the following evaluations, the 160-division model with normalization for scaling was used as the best-case.

6.5 Likelihood histogram

We applied the best model obtained in the previous section to the validation data and real data. The softmax function (Sec. 6.2.5) was applied to the model outputs to obtain the signal likelihood for each event. Figure 6.9 shows the histogram of the obtained signal likelihood distribution. Ideally, the signal likelihood should be close to 1 for $0\nu\beta\beta$ and close to 0 for the background. For the simulation data, the $0\nu\beta\beta$ and background are well separated, but some of the background data shows high signal likelihood. This trend is more pronounced in the real data, indicating that the expected performance from the validation data is not being achieved in the real data. In the following sections, we investigate the characteristics of misclassified events to infer what information the model utilizes. We also discuss how the distribution of summary quantities differs between simulation and real data, and how these differences influence the likelihood distribution.

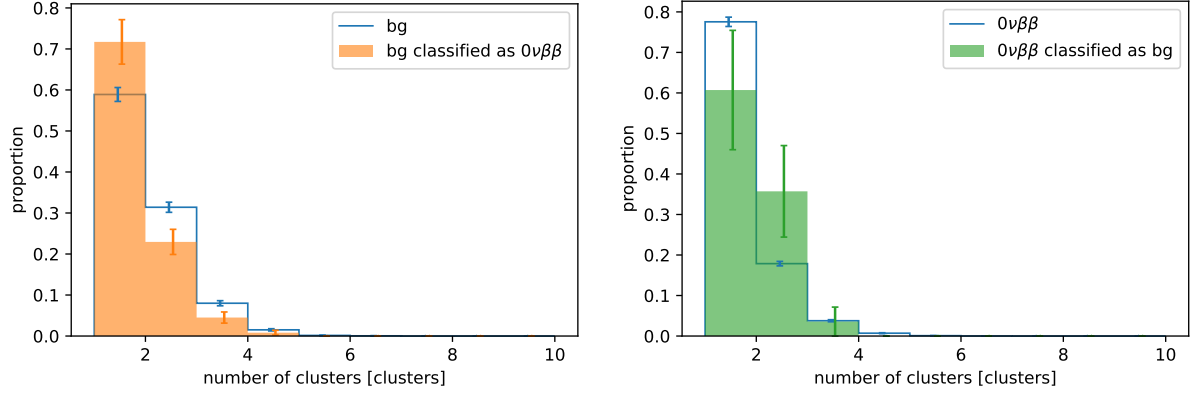


Fig. 6.10 Overall-misclassified comparison of the cluster number distributions of background events (left), and $0\nu\beta\beta$ events (right).

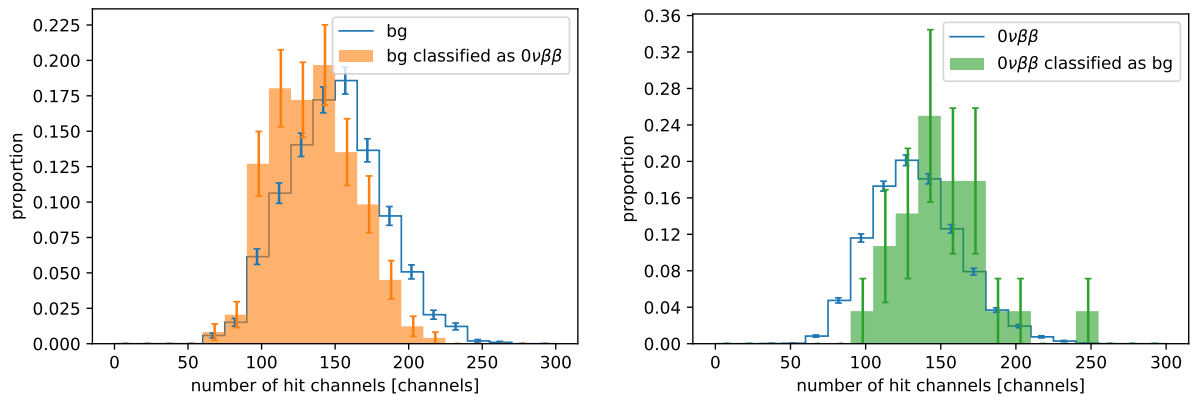


Fig. 6.11 Overall-misclassified comparison of the distributions of the number of hit channels of background events (left), and $0\nu\beta\beta$ events (right).

6.6 Misclassified events

The machine learning model outputs the posterior probability that input events belong to each class ($0\nu\beta\beta$ or background) when equal prior probabilities are assigned. In this section, the characteristics of events with a probability of 90% or higher for the incorrect class, are investigated. To understand the characteristics of misclassified events, distributions of summary quantities similar to those in Chap. 5 were created for the validation data.

Differences are observed in the distribution concerning the number of clusters and the track volume, while no significant differences were observed in the other distributions. Figure 6.10 shows the overall-misclassified comparison of the cluster number distributions of background events and $0\nu\beta\beta$ events. Compared to the overall distributions, misclassified events in background tend to have fewer clusters, while misclassified events in $0\nu\beta\beta$ tend to have more clusters. Considering that the background tends to have a higher number of clusters, as mentioned in Sec. 5.3.3, this trend of misclassification is natural. Figure 6.11 shows the comparison of the distributions of the number of hit channels. Compared to the overall distributions, misclassified events in background tend to have fewer hit channels, while misclassified events in $0\nu\beta\beta$ tend to have more hit channels. A similar trend is observed in the hit volume (Fig. 6.12). These observations are also considered to reflect the tendency for the track volume of $0\nu\beta\beta$ to be smaller than that of the background in the simulation (Sec. 5.3.4).

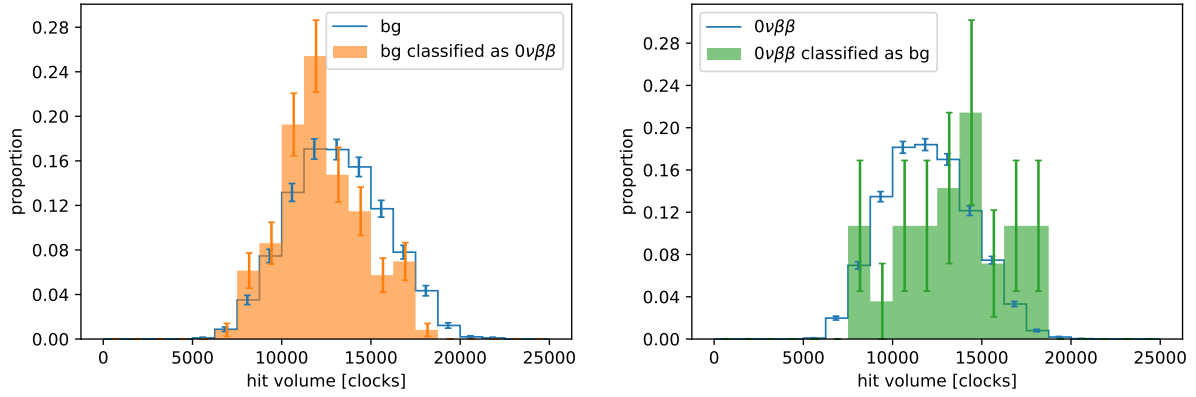


Fig. 6.12 Overall-misclassified comparison of the distributions of the hit volume of background events (left), and $0\nu\beta\beta$ events (right).

Based on the comparisons, it is inferred that the model uses information such as the number of clusters and track volume as one of the criteria for discrimination. In the real data, as mentioned in Sec. 5.3.4, the track volume of the real data is distributed on the smaller side compared to the simulation background and is closer to the distribution of the simulation $0\nu\beta\beta$. This may be causing the shift in the likelihood distribution. In the next section, the causes of the differences in the distribution of track volumes are examined.

6.7 Dataset bias

6.7.1 Veto by dead channels and high dark current channels

The difference in signal likelihood between the validation data and the real data, as mentioned in Sec 6.5, is due to the characteristics of the datasets. The most apparent difference is that the real data has a veto set as shown in Fig. 4.4, which biases the track shapes. To evaluate the effect, a dataset was created from the validation data by removing events with hits at the veto channels in Fig. 4.4, and this dataset was re-evaluated by inputting it into the model. By applying the veto, the data size is reduced to approximately one-sixth to one-eighth. The number of $0\nu\beta\beta$ events decreases from 5854 to 926, and the background events from 2051 to 254. Changes are observed in the distribution of summary quantities. Figure 6.13 shows the comparison of the distribution of photon-count weighted centroid position between the validation data and the real data, with and without the application of the veto. By applying the veto, the distribution of the photon-count weighted centroid position in the x direction becomes closer to that of the real data. In the y-direction, the application of the veto slightly narrows the distribution width. However, the distribution of the centroid position in the z direction does not show significant changes. Figure 6.14 shows a similar comparison for the number of hit channels and the hit volume. Both tend to be smaller in the real data compared to the simulation, and when the veto is applied to the simulation data, the distribution tends to asymptotically approach that of the real data.

Figure 6.15 shows a comparison of the signal likelihood of the data after applying the veto. The high-side tail of the background's signal likelihood is closer to that of the real data. This suggests that the discrepancy in the likelihood distribution discussed in Sec. 6.5 is primarily due to biases in the track distribution introduced by vetoes on dead channels and high-dark channels. In the absence of these vetoes, the model is expected to perform comparable to its performance on the validation data.

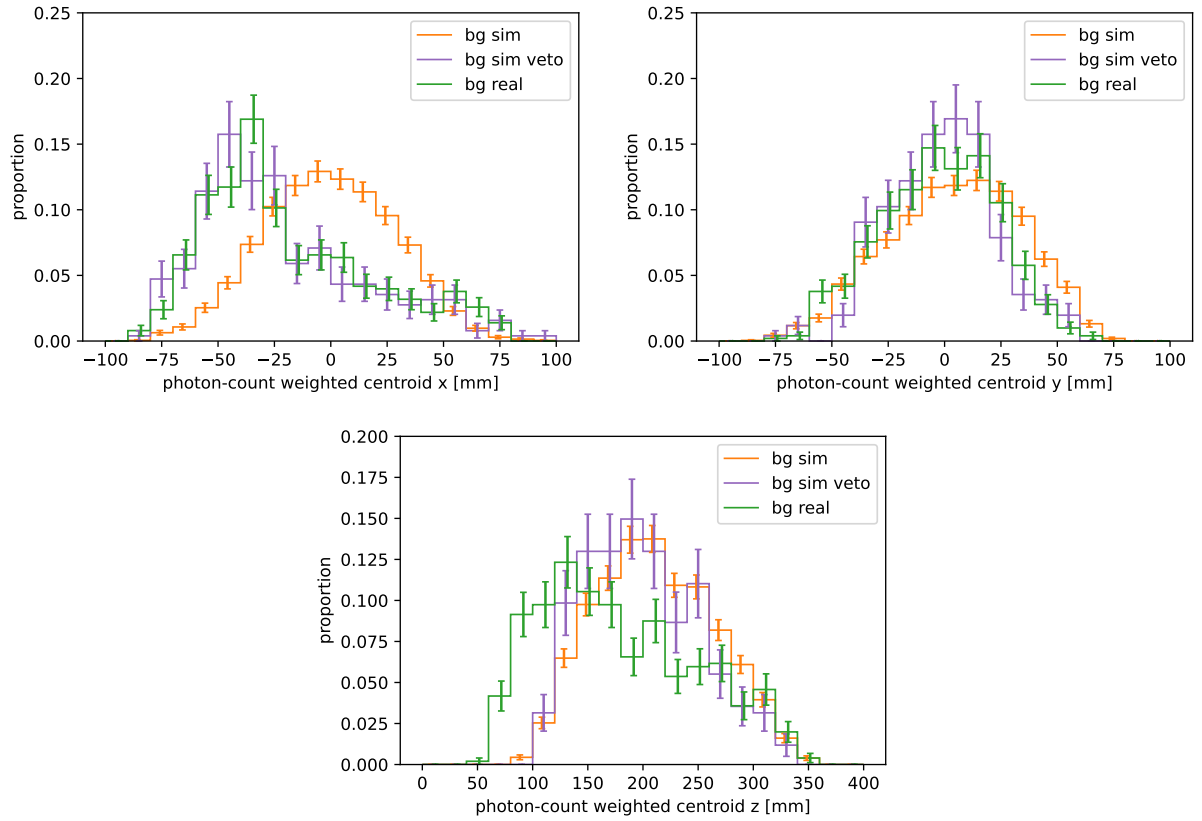


Fig. 6.13 Comparison of the distribution of photon-count weighted centroid in x (upper left), y (upper right), z (bottom) direction between the validation data and the real data, with and without the application of the veto. Orange and purple lines are the validation data with and without the application of the veto, while green is the real data.

6.7.2 Z position

Applying a veto to the simulation data brings the track volume distribution closer to the real data, but it still does not match perfectly. Another possible cause is the difference in the distribution of the track's z -position. Ionization electrons diffuse as they drift toward the ELCC, with less diffusion occurring at smaller z -positions due to the shorter drift distance. As shown in Fig. 6.13, the distribution of the centroid position in the z direction in the real data is skewed towards the smaller side compared to the simulation data. This could be due to z mis-reconstruction, but it is also possible that the tracks are actually distributed towards the smallest drift side. Figure 6.16 shows a scatter plot of the centroid position in the z direction and hit volume in the real data. There is a positive correlation between the two, with a correlation coefficient of 0.60. In contrast, the correlation with the number of hit channels is low at 0.32, indicating no significant correlation. This is because changes in the number of hit channels require diffusion of about 10 mm, the pitch of the ELCC, whereas changes in hit volume can be observed with diffusion corresponding to the 200 ns clock of the ADCL, which is about 0.2 mm.

To compare track volumes while excluding the influence of z -position differences, the z -direction distribution shown in Fig. 6.13 was flattened by applying a weighting based on its inverse. The same weighting was applied to the track volume, and the resulting distribution is shown in Fig. 6.17. Compared to Fig. 6.14, the difference in distribution between the real data and the simulation data has decreased. This suggests that there is a difference in the distribution of the track's z -position between the real data and the simulation data, which in turn causes the difference in track volume and subsequently the

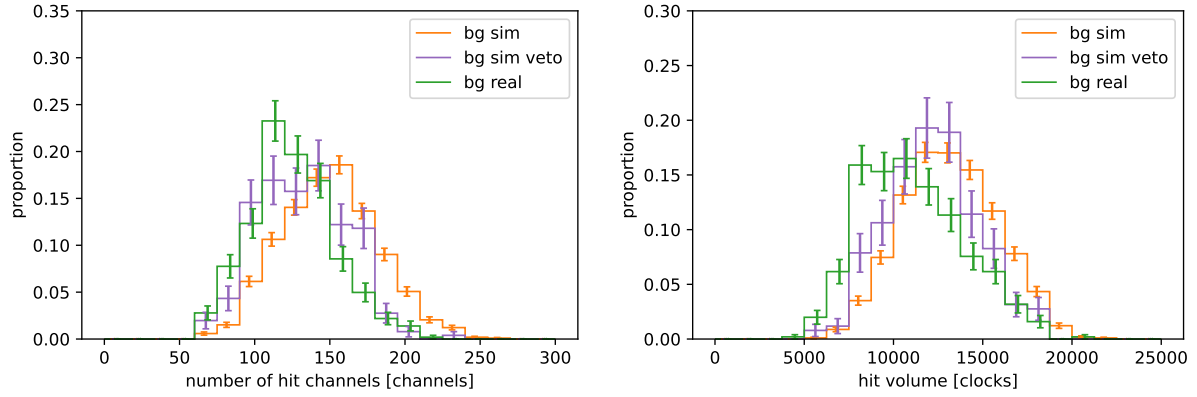


Fig. 6.14 Comparison of the distribution of number of hit channels (left), hit volume (right) between the validation data and the real data, with and without the application of the veto. Orange and purple lines are the validation data with and without the application of the veto, while green is the real data.

difference in the likelihood distribution.

The reason of the difference in the z-position has not been identified, but one possible cause is the position of the source. In the real data, as shown in Fig. 4.3, the source is attached to the pressure vessel, and its z-position corresponds to a location 174 mm to 474 mm from the ELCC surface. In contrast, in the simulation, as described in Sec.5.2.1, gamma rays are emitted from directly above the sensitive volume towards the bottom, corresponding to a z-position of 20 mm to 390 mm. This difference suggests that the incident direction of the gamma rays differs between the real data and the simulation. In particular, the real data indicate that gamma rays are more frequently incident from the PMT side toward the ELCC side than in the opposite direction. Considering that the emission direction of photoelectrons tends to be biased toward the direction of incident photons (Sec. 5.2.1), as shown in Fig. 6.18, in the real data, the initial direction of the photoelectrons can be biased toward the ELCC when the fully contained condition is applied. Given that vetoes are applied around the center of the ELCC surface due to dead channels and high dark channels, it is possible that the track endpoints are biased toward the smallest drift side, resulting in the centroid position in the z direction being biased towards the smaller side.

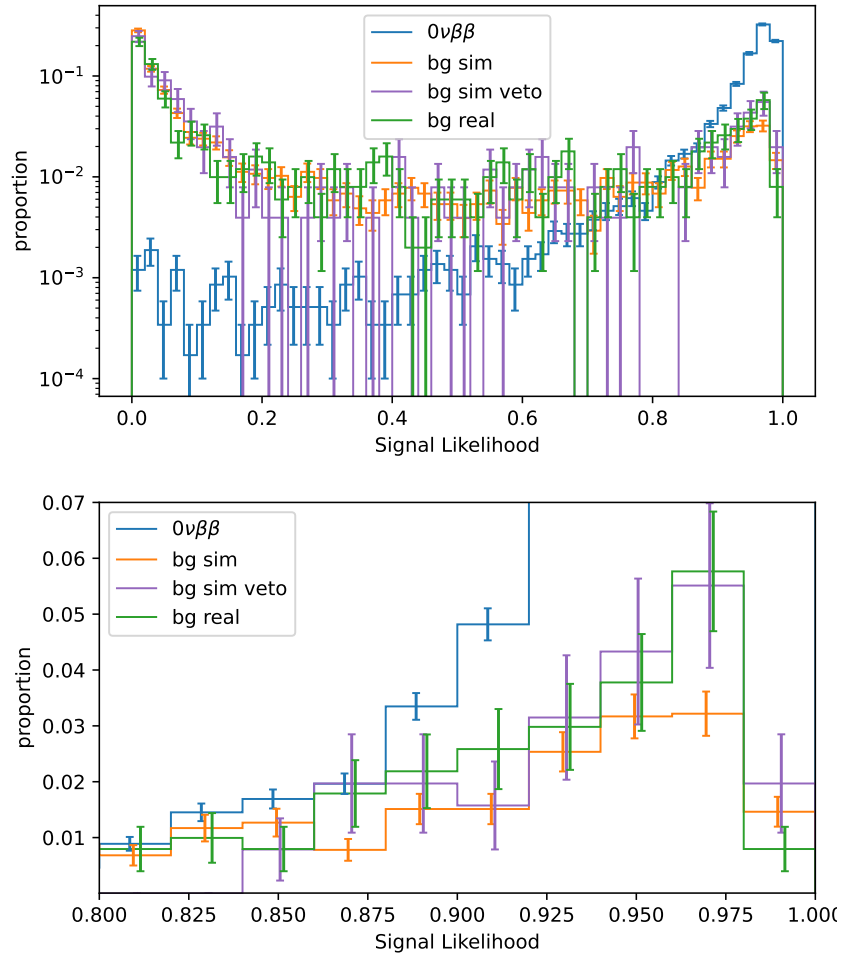


Fig. 6.15 The likelihood histogram for simulation and real data (top) and its zoom up view (bottom). The blue, orange and green lines are the same as lines in Fig. 6.9. The purple line is the simulation data after applying the veto. Applying a veto to the simulation data brings the high-side tail of the signal likelihood closer to that of the real data.

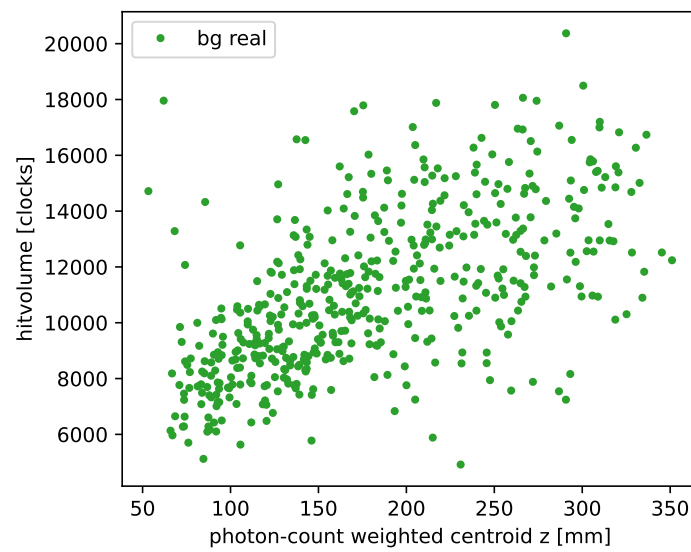


Fig. 6.16 Scatter plot of the centroid position in the z direction and hitvolume in the real data.

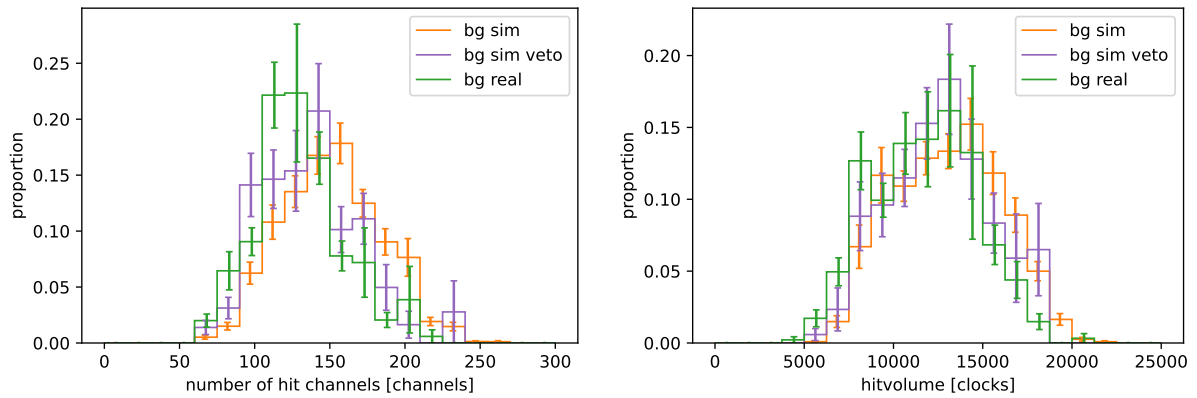


Fig. 6.17 Track volume distribution weighted by the inverse of the proportion of the centroid position in the z direction.

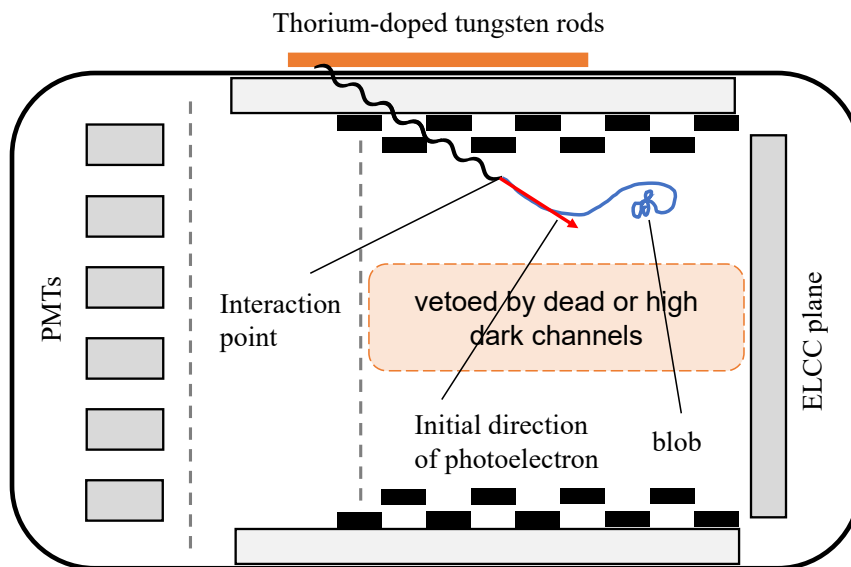


Fig. 6.18 Schematic diagram of the track associated with gamma rays from the source. When the incident gamma rays are directed toward the ELCC, the initial velocity of the photoelectron also directed toward the ELCC, causing the end of the track to be biased toward the smallest drift side.

Chapter 7

Future sensitivity of the AXEL experiment

In this chapter, the sensitivity of the AXEL experiment that can be achieved when applying the background event exclusion using the CNN model constructed in Chap. 6 is evaluated. The sensitivity evaluation is carried out according to the procedure described below.

(1) Background source

The number of occurrences is evaluated for each type of background.

(2) Volume cut

The efficiency of tracks being contained within the sensitive volume of the detector and the size of the machine learning model (180 L detector size) is applied.

(3) Energy resolution cut

The efficiency of tracks falling within a specific region of interest (ROI) around the $0\nu\beta\beta$ Q value is applied.

(4) Topology cut using machine learning

The selection efficiency corresponding to a threshold applied to the signal likelihood output by the machine learning model is incorporated.

(5) Sensitivity

Based on the background rate after applying the above selection efficiencies, the upper limit on the signal is derived using the Feldman-Cousins method[95] and converted into sensitivity.

The following sections describe the details of each procedure.

7.1 Background source

In this section, we estimate the amount of the activity of the backgrounds discussed in Sec. 5.1. The ^{214}Bi and ^{208}Tl isotopes are present in various components of the detector; however, since the pressure vessel accounts for the majority of the detector's mass, it is expected to become the dominant source of background as the radioactivity of other components is reduced. Therefore, in this study, the background is assumed to originate solely from ^{214}Bi and ^{208}Tl present in the material of the pressure vessel.

A pressure vessel with a ϕ 4m $\times$ L 4m is assumed to accommodate the sensitive volume of the 1-ton detector, which has dimensions of, for example, ϕ 3m $\times$ L 2.5m at 10 barA. Both ends of the vessel are closed with ellipsoidal caps, each 0.4 meters in length. Two materials are considered for the pressure vessel: SUS304 stainless steel and oxygen-free copper. Although oxygen-free copper has lower mechanical strength, it exhibits lower levels of intrinsic radioactivity. According to JIS B 8265, the minimum wall thickness t (mm) of a cylindrical pressure vessel is given by the following formula:

$$t = \frac{PD_i}{2\sigma_a\eta - 1.2P} \quad (7.1)$$

where P is the design pressure (MPa), D_i is the inner diameter of the cylindrical shell (mm), σ_a is the allowable tensile stress of the material (N/mm²), and η is the weld joint efficiency. Assuming a design pressure of $P = 1$ MPa for safety, an inner diameter $D_i = 3960$ mm, and a weld joint efficiency $\eta = 1$, and using the allowable tensile stress for SUS304 as $\sigma_a = 129$ N/mm² according to JIS standards, the required wall thickness is calculated to be $t = 15.4$ mm. Based on this result, a wall thickness of 2 cm is assumed for the SUS304 pressure vessel in this study. For oxygen-free copper, the allowable tensile stress is $\sigma_a = 46$ N/mm². Applying the same calculation method yields a required wall thickness of $t = 43$ mm. Accordingly, a wall thickness of 5 cm is assumed for the oxygen-free copper vessel. Under these assumptions, the mass of the pressure vessel is estimated to be 10.0 tons for SUS304 and 27.8 tons for oxygen-free copper.

²¹⁴Bi is a nuclide in the uranium series. The half-life of ²³⁸U, 4.5×10^9 yr, is much longer than that of other nuclides in the series, so that the radioactivity of ²¹⁴Bi is equal to that of ²³⁸U. We assume the ²³⁸U content in SUS304 to 3.7 mBq/kg, based on the pressure vessels reported in [96]. For 10.0 tons of SUS304 pressure vessel, this corresponds to an activity of 37 Bq. For oxygen-free copper, the ²³⁸U concentration C_U of 2.9×10^{-12} g/g is assumed based on the data reported in [97]. The corresponding radioactivity R_{Bi} is given by the following expression, using the decay constant of ²³⁸U, $\lambda_U = 1.55 \times 10^{-10}$ yr⁻¹ = 4.92×10^{-18} s⁻¹:

$$R_{Bi} = N_A \times \frac{C_U}{238} \times \lambda_U \quad (7.2)$$

From this, the radioactivity of this clean oxygen-free copper is calculated to be 0.036 mBq/kg, which corresponds to a total activity of approximately 1.0 Bq for the 27.8-ton pressure vessel. Multiplying these radioactivities by the 2448 keV gamma ray intensity of 1.545 % [98] results in the 2448 keV gamma ray flux of 1.8×10^7 yr⁻¹ for SUS304 pressure vessel and 4.9×10^5 yr⁻¹ for oxygen-free copper, respectively.

²⁰⁸Tl belongs to the thorium series and is in radiative equilibrium with ²³²Th, which has a half-life of 1.4×10^{10} years. We assume the ²³²Th content in SUS304 to 0.10 mBq/kg based on [96] and the corresponding activity is 1.0 Bq for 10.0 tons of SUS304 pressure vessel. For oxygen-free copper, the ²³²Th concentration C_{Th} is taken to be 2.4×10^{-12} g/g based on [97]. The radioactivity for this clean oxygen-free copper is calculated to be 9.7×10^{-3} mBq/kg, which corresponds to a total activity of approximately 0.27 Bq for the 27.8-ton pressure vessel. Multiplying these radioactivities by the 2615 keV gamma ray intensity of 99.754 % [98] results in the 2615 keV gamma ray flux of 3.1×10^7 yr⁻¹ for SUS304 pressure vessel and 8.5×10^6 yr⁻¹ for oxygen-free copper, respectively.

For ¹³⁷Xe, we take the rate of 7.91×10^{-8} keV⁻¹kg⁻¹year⁻¹ at around 2.5 MeV, as assumed for a detector using one ton of ¹³⁶Xe in the NEXT experiment [47], as the basis. Since this value includes the application of the fiducial cut efficiency (0.829) and two types of topological cut efficiencies (0.548 and 0.064) [99], it is corrected by dividing by these factors. In addition, this rate was evaluated for the Gran Sasso National Laboratory (LNGS). We convert this to the expected rate at the Kamioka Observatory, which is the assumed experimental site. The conversion is performed by calculating the cosmic-ray muon flux I contributing to the production of ¹³⁷Xe using the equation provided in [100], and applying the ratio of the fluxes as a correction factor.

$$I(X) \sim A \left(\frac{X_0}{X} \right)^\eta \exp(-X/X_0) \quad (7.3)$$

where $A = (2.15 \pm 0.08) \times 10^{-6}$ cm⁻²s⁻¹sr⁻¹, $\eta = 1.93^{+0.20}_{-0.12}$, and $X_0 = 1155^{+60}_{-30}$ meters of water equivalent. By substituting the meters of water equivalent values of 3800 for LNGS and 2700 for the Kamioka Observatory into the equation, and taking the ratio of the resulting muon fluxes, a correction factor of 5.01 from LNGS to Kamioka is obtained. As a result, we assume that the ¹³⁷Xe rate is 1.36×10^{-5} keV⁻¹kg⁻¹year⁻¹.

7.2 Volume cut

This section discusses the proportion of events that are generated by the background source and detected as the fiducial volume contained events.

Gamma rays from ^{214}Bi and ^{208}Tl are emitted isotropically from the pressure vessel. For ^{208}Tl in particular, it is necessary to evaluate the fraction of events that undergo Compton scattering within the pressure vessel, lose energy, and are subsequently detected by the TPC with energies near the $0\nu\beta\beta$ Q-value. Beta particles from ^{137}Xe are generated randomly within the xenon volume of the detector. Using the Geant4 simulation data conducted for 2448 keV gamma rays with the geometry of the 1-ton detector with SUS304 pressure vessel, the proportion of tracks contained within the sensitive volume was examined^{*1}.

The self-shielding effects was roughly estimated by assuming that gamma rays travel perpendicularly through the vessel wall. This is a conservative assumption for shielding estimation. Then, the average attenuation factor F for gamma rays emitted uniformly within the thickness T can be obtained by taking the expected value of $\exp(-\mu t)$:

$$F = \int_0^T \frac{dt}{T} \exp\left(-\frac{\mu_m}{\rho} t\right) = \frac{\rho}{\mu_m T} \left(1 - \exp\left(-\frac{\mu_m}{\rho} T\right)\right) \quad (7.4)$$

where μ , mass attenuation coefficient for 2448 keV gamma rays is $3.961 \times 10^{-2} \text{ cm}^2/\text{g}$ for SUS304 and $3.916 \text{ cm}^2/\text{g}$ for oxygen-free copper[101]. SUS304 is treated as a mixture composed of 74 % iron, 18 % chromium, and 8 % nickel by mass. Using the densities of SUS304 (7.93 g/cm^3) and copper (8.94 g/cm^3), the attenuation factors are calculated to be $F_{\text{SUS304},2\text{cm}} = 0.743$ and $F_{\text{Cu},5\text{cm}} = 0.472$, respectively. Therefore, in the case of oxygen-free copper, a self-shielding correction factor of $F_{\text{Cu},5\text{cm}}/F_{\text{SUS304},2\text{cm}} = 0.636$ was applied to the number of gamma rays simulated for SUS304.

The CNN model developed in this study requires that the track is contained within the 180 L detector fiducial region. In addition, the model was trained for the 6.8 bar pressure while the simulation data for the 1-ton detector was generated at 10 bar. Therefore, the track size in the 1-ton detector simulation was scaled by 10/6.8 and then volume cut was applied. As a result of this cut, 3333 out of 10 000 events for $0\nu\beta\beta$ and 2630 out of 1×10^7 2448 keV gamma ray events remained. For ^{214}Bi , this ratio of 2.63×10^{-4} is the cut efficiency. For ^{208}Tl , it is additionally necessary to evaluate the fraction of events that, after undergoing Compton scattering, are photoelectrically absorbed with energies near the Q-value. To achieve this, we have made a 2615 keV simulation data set by scaling the energy of the available 2448 keV gamma simulation data set. Figure 7.1 shows the energy spectrum near the $0\nu\beta\beta$ Q-value. The number of events within the defined ROI was estimated using the linear fit function around the Q-value. The ratio of this number to the number of photoelectric absorption events, was then calculated and used as a correction factor. Exact number depends on the ROI, but 0.47 % for a $Q_{\beta\beta} \pm 5 \text{ keV}$ energy window. Finally, the Compton-scattering cut described in Sec. 5.2.4, was applied, resulting in the survival probability of 99.3 % for $0\nu\beta\beta$ and 75.1 % for the background for both ^{214}Bi and ^{208}Tl .

For ^{137}Xe , since it emits beta particles, only events in the simulation data where electrons were emitted via photoelectric absorption were considered. Among these, the fraction of events that were fully contained within a region equivalent to a volume of 180 L detector region was calculated to be 0.291, and this value is used as the cut efficiency. Since the emission of xenon characteristic X-rays is not taken into account, the cut efficiency obtained here is considered to be an underestimate. A more conservative

^{*1} Although the gamma-ray energy from ^{208}Tl is 2615 keV, we evaluated the proportion of tracks using the simulation data at 2448 keV, which was already available for the 1-ton detector.

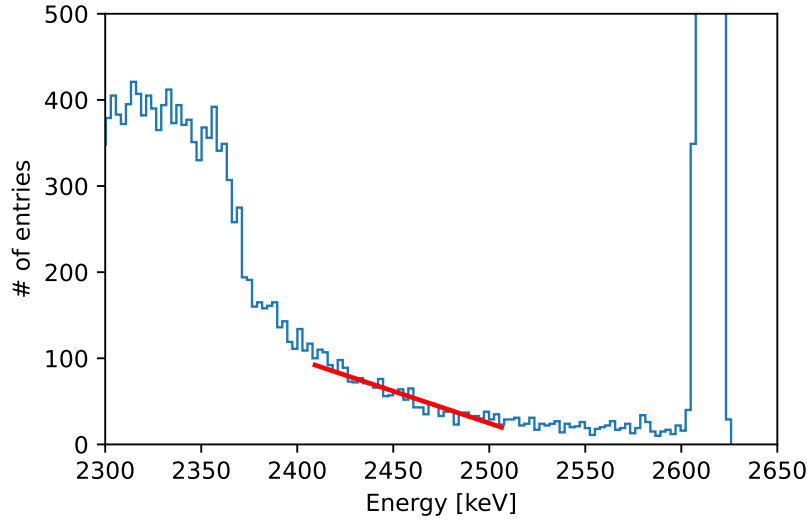


Fig. 7.1 Simulated energy spectrum of events from the ^{208}Tl 's in gas vessel. It is obtained by scaling the simulation result of the 2448 keV gamma. The red line represents a linear fit around the $0\nu\beta\beta$ Q-value region.

approach would be to use the efficiency of 3333/10000 for the electron-only $0\nu\beta\beta$ case. However, as discussed in Sec. 7.5, the contribution from ^{137}Xe is negligible, and thus the impact of this difference in cut efficiency is expected to be small.

7.3 Energy resolution cut

Events are selected whose reconstructed energy is within the region of interest (ROI) centered on the $0\nu\beta\beta$ Q value. For ^{214}Bi , the contamination level is significantly affected by the energy reconstruction resolution. Figure 7.2 shows the probability distribution of the reconstructed energy for $0\nu\beta\beta$ and ^{214}Bi background with an energy resolution of 0.678 % (FWHM) and an ROI of the $0\nu\beta\beta$ Q value ± 10 keV. Here, Gaussian distribution is assumed. In this case, the survival probability of $0\nu\beta\beta$ and ^{214}Bi background remaining after the cut is 84.2 % and 49.8 %, respectively. Although narrowing the ROI range can reduce background contamination, it also decreases the number of signals. Therefore, the range is optimized to maximize the sensitivity.

7.4 Topology cut using machine learning

A cut is applied based on the CNN model described in Sec. 6.5 using the track information of each event. The model is trained to distinguish between $0\nu\beta\beta$ and background events with the same energy. Therefore energy-based cut described in the previous section and the topology-based cut presented here is independent. True positive rate (TPR) and false positive rate (FPR) (Sec. 6.3.2) at each threshold value of the receiver operating characteristic (ROC) curve for the validation data are taken as the signal efficiency and background efficiency, respectively. The signal threshold is optimized to maximize the sensitivity.

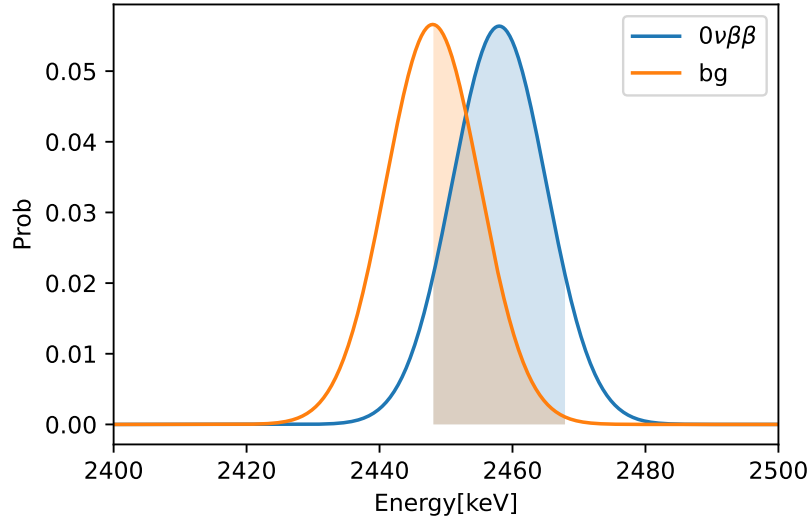


Fig. 7.2 Probability distribution of the energy for $0\nu\beta\beta$ and ^{214}Bi background with an energy resolution of 0.678 % (FWHM) and an ROI of the $0\nu\beta\beta$ Q value ± 10 keV. The area of the shaded region represents the efficiency remaining within the ROI.

7.5 Sensitivity for 1-ton detector

With the t years observation of M kg $0\nu\beta\beta$ neucli, the expected number of $0\nu\beta\beta$ events is given as follows,

$$N = \log 2 \cdot \frac{N_A}{W} \cdot \varepsilon_{\text{sig}} \cdot \frac{M \cdot t}{T_{1/2}^{0\nu}} \quad (7.5)$$

Here, N_A is the Avogadro constant, W is the atomic weight of the neucli, ε_{sig} is the signal detection efficiency, and $T_{1/2}^{0\nu}$ is the half life of $0\nu\beta\beta$ decay. If one obtain the upper limit of $0\nu\beta\beta$ events μ_{up} , one obtain the lower limit of half life $T_{1/2}^{0\nu}$ from Eq. 7.5,

$$T_{1/2}^{0\nu} = \log 2 \cdot \frac{N_A}{W} \cdot \varepsilon_{\text{sig}} \cdot \frac{M \cdot t}{\mu_{\text{up}}} = \log 2 \cdot N_{\text{Xe}} \cdot t \cdot \frac{\varepsilon_{\text{sig}}}{\mu_{\text{up}}} \quad (7.6)$$

To estimate μ_{up} from the observation, the Feldman-Cousins method[95] is used in this study. It gives the upper limit of signals $\mathcal{U}(n|b)$ for a given observation n and a mean background level b with certain confidence interval. To estimate the sensitivity as the expected upper limit for no signal case, the mean of the upper limit is calculated using a Poisson distribution[102],

$$\mu_{\text{up}} \equiv \text{E}[\mathcal{U}(n|b)] = \sum_{n=0}^{\infty} \text{Po}(n|b) \mathcal{U}(n|b) \quad (7.7)$$

where $\text{Po}(n|b)$ is a Poisson pdf with variable n and mean b . When the average background level b is large, the computation time increases significantly. Therefore, for $b > 50$, evaluation is approximated as,

$$\mu_{\text{up}} \sim \mathcal{U}(b|b) \quad (7.8)$$

Sections 7.1 to 7.4 determine the efficiency of signal events and the number of background events. Using this, we evaluate the sensitivity of a 1-ton detector over a 10-year measurement period. The evaluation is conducted for two energy resolutions: 0.678 % (FWHM) obtained in this study and an anticipated

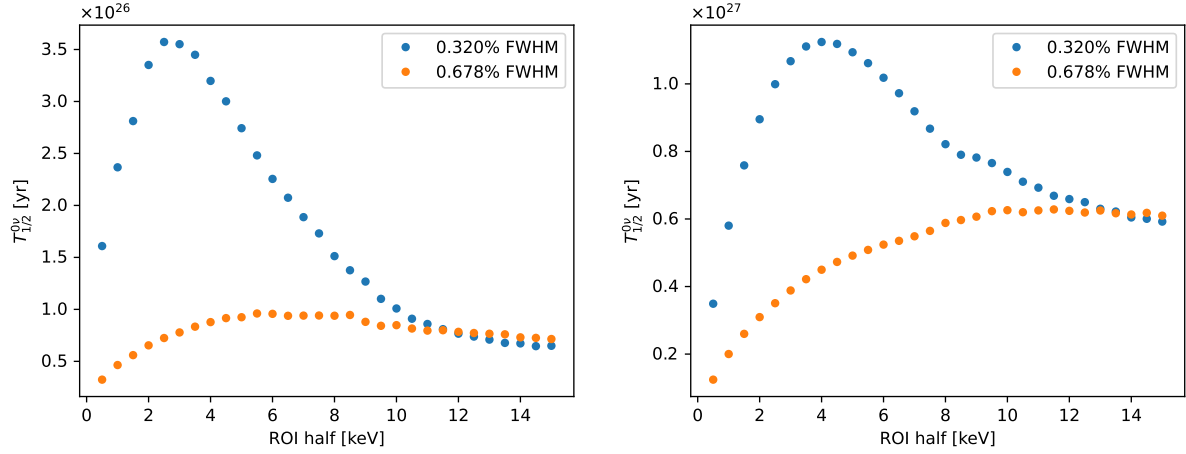


Fig. 7.3 Sensitivity in each ROI range. Left is for the SUS304 pressure vessel and right is for the oxygen-free copper pressure vessel.

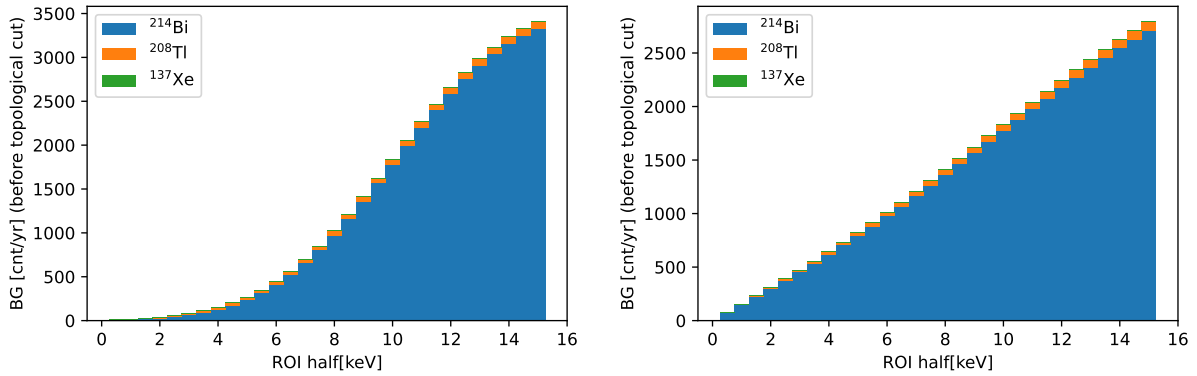


Fig. 7.4 Breakdown of background contributions for SUS304 pressure vessel. The left panel corresponds to a energy resolution of 0.320 % FWHM, while the right panel corresponds to 0.678 % FWHM.

improved energy resolution of 0.32 % [68].

Figure 7.3 shows the sensitivities at each ROI half width for various energy resolutions, comparing the cases of a SUS304 pressure vessel and an oxygen-free copper pressure vessel. For both materials, better energy resolution leads to more effective background suppression through energy resolution cuts, resulting in a higher lower limit on $T_{1/2}^{0\nu}$. However, the SUS304 pressure vessel exhibits radioactivity higher than that of oxygen-free copper, yielding only about one-fourth the lower limit compared to the oxygen-free copper case. Figure 7.4 and 7.5 shows the breakdown of background contributions for the case of SUS304 and oxygen-free copper pressure vessel, respectively. In both cases, the contribution from ^{137}Xe is negligible. For the case of SUS304 pressure vessel, ^{214}Bi accounts for the majority of the background, whereas in the case of oxygen-free copper pressure vessel, the contribution from ^{208}Tl is more significant compared to SUS304. Furthermore, the contribution from ^{208}Tl increases as the ROI becomes narrower.

Table 7.1 summarizes the efficiency of the topology cut, the number of background events, the upper limit on the signal, and the resulting lower limit on $T_{1/2}^{0\nu}$. The best result was obtained with a 0.320 % energy resolution using an oxygen-free copper pressure vessel, yielding a lower limit on $T_{1/2}^{0\nu}$ of 1.12×10^{27} years.

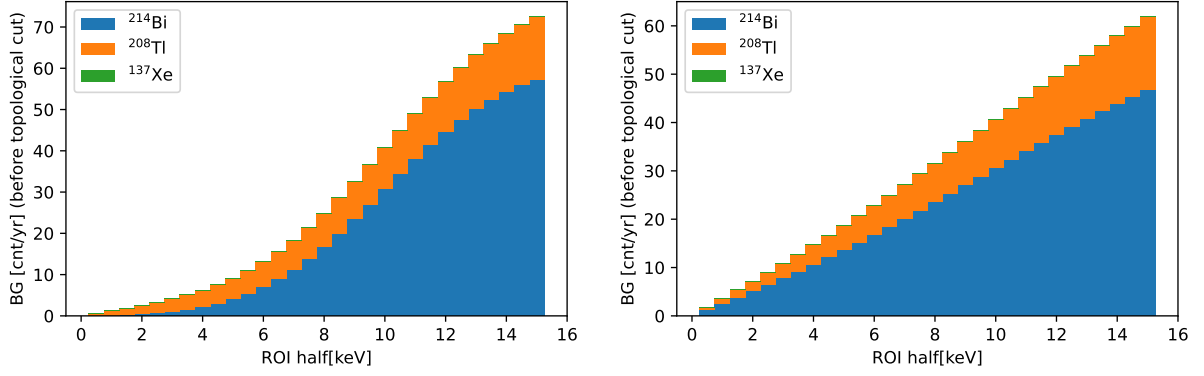


Fig. 7.5 Breakdown of background contributions for oxygen-free copper pressure vessel. The left panel corresponds to a energy resolution of 0.320 % FWHM, while the right panel corresponds to 0.678 % FWHM.

Table 7.1 Summary of the optimized ROI half width, threshold and efficiency of topology cut, the number of background events, the upper limit on the signal, and the resulting lower limit on $T_{1/2}^{0\nu}$.

FWHM energy resolution		SUS304		oxygen-free copper	
		0.320%	0.678%	0.320%	0.678%
optimized ROI half width		2.5 keV	5.5 keV	4.0 keV	11.5 keV
Topology cut threshold		0.926	0.956	0.931	0.909
Topology cut efficiency	signal	0.778	0.595	0.757	0.825
	background	0.0941	0.0512	0.0863	0.112
Background count	^{214}Bi	42.7	882	2.19	35.8
	^{208}Tl	14.6	32.0	4.02	11.6
	^{137}Xe	0.0203	0.0447	0.0325	0.0935
Signal upper limit		12.2	35.8	5.31	12.1
$T_{1/2}^{0\nu}$		3.57×10^{26} yr	9.61×10^{25} yr	1.12×10^{27} yr	6.28×10^{26} yr

Part IV

Summary

Chapter 8

Future improvements

In this chapter, we discuss potential improvements aimed at enhancing the sensitivity of the AXEL experiment.

8.1 CW upgrade

We have been conducting measurements of radiation level of detector components with a germanium semiconductor detector at the Kamioka Observatory. The ceramic capacitor used in the CW multiplier was found to have very high radioactivity around 1000 mBq/kg in the uranium series. In a 1-ton detector, roughly 500 capacitors would be used corresponding to a total of about 300 mBq. Its activity is close to that of the potential oxygen-free copper pressure vessel (Sec. 7.1). Therefore, the capacitors used in this study would significantly degrade the sensitivity of a 1-ton detector experiment. As an alternative to the ceramic capacitors, we are developing a new low-activity capacitor made of a double-sided FPC board stacked together.

8.2 Deep learning model upgrade

The model employed in this study is based on the CNN. As the spatial resolution of the input data increases, memory consumption correspondingly rises. The TPC data are sparse, with values at limited xyz positions and most points having values of zero. Therefore, it will be possible to improve memory efficiency by using techniques such as sparse convolution[103] and graph neural networks[104].

As verified in Sec. 6.6, the distribution of track sizes differs between $0\nu\beta\beta$ and background, with the tails of the distribution tending to result in misclassified events. The distribution is broadened due to the diffusion effect during drift within the TPC. Therefore, by using methods such as Richardson-Lucy deconvolution[105] to reconstruct the tracks before diffusion and then performing discrimination, it may be possible to reduce misclassified events. Although still under development, Fig. 8.1 shows the track of a 2615 keV event from ^{208}Tl before and after applying the Richardson – Lucy deconvolution. To deconvolute the track, a point spread function (PSF) is required that characterizes the diffusion of the track. In the AXEL experiment, this corresponds to the diffusion of ionization electrons during drift and the binning effect introduced by the ELCC. However, in Fig. 8.1, only diffusion during drift is considered. Nevertheless, the result demonstrates that the track becomes noticeably sharper after deconvolution.

In this study, the model was developed for tracks within the size of the 180 L detector. This design choice was made both to facilitate comparison with real data and to accommodate limitations in memory capacity. As a result, a tight volume cut is required in the sensitivity evaluation of the 1-ton detector, leading to the loss of approximately 67% of signal events (Sec. 7.2). It can be improved by the usage of sparse convolution or graph neural network and deconvolution.

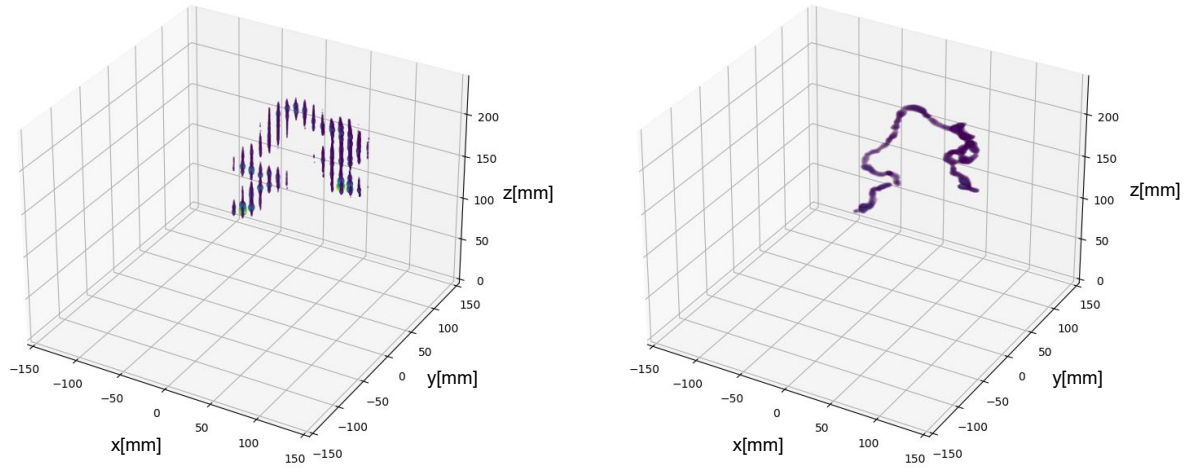


Fig. 8.1 Track of a 2615 keV event from ^{208}Tl before (left) and after (right) applying the Richardson-Lucy deconvolution.

8.3 Energy resolution improvement

We consider that there are rooms for improvement in the resolution of 0.678% FWHM at the $0\nu\beta\beta$ Q value obtained in Sec. 4.3. In a previous study, the breakdown of energy resolution was evaluated using measurements with 1836 keV gamma rays[65], and it was suggested that various improvements could enhance the energy resolution to 0.32% (FWHM) at the $0\nu\beta\beta$ Q value[68]. Improved energy resolution enables a narrower ROI width in the signal region, which is expected to enhance sensitivity by reducing background contamination.

Chapter 9

Conclusion

The discovery of neutrino oscillations revealed that neutrinos have masses, but the reason for their extremely small masses compared to other charged leptons and quarks remains unknown. If neutrinos are Majorana particles, the See-saw mechanism can explain their light masses, and it is also a premise for leptogenesis, a leading mechanism explaining the matter-antimatter asymmetry. Experimentally, the only feasible way to confirm the Majorana nature of neutrinos is the observation of neutrinoless double beta ($0\nu\beta\beta$) decay, and various candidate nuclei are being experimented on.

We are conducting the AXEL project using a high-pressure xenon gas TPC with the aim of observing the $0\nu\beta\beta$ decay of ^{136}Xe . By using EL signals for readout, we can suppress amplification fluctuations and achieve good energy resolution. Additionally, by using the ELCC for reading out signals, we can reconstruct three-dimensional track information. The electron tracks in the TPC deposit greater energy at their endpoints and form “blobs”. Therefore, by using track information, we can distinguish between two-track events from $0\nu\beta\beta$ and one-track background events originating from gamma rays. We are currently developing a 180 L prototype detector, focusing on the development of elemental technologies, verification of scaling-up techniques, evaluation of energy resolution near the Q value of $0\nu\beta\beta$ and demonstration of background rejection capability using track information.

An important technological element in the TPC is the generation of high voltage to form the drift electric field. The 180 L prototype detector and the next 1000 L detector require high voltages of -44.8 kV and -79.8 kV , respectively. To apply such high voltages from outside the pressure vessel, high voltage feedthroughs compatible with high pressure are necessary. Another approach is to introduce relatively low voltage from outside the pressure vessel and boost it inside the vessel. The CW multiplier can be used for this purpose, as it converts a low voltage AC input to a high voltage DC output. In the AXEL experiment, the EL photons induced by the ionization electrons are used as signals. Since the ionization signal is converted to light to be read out and the light signal is amplified with quite high efficiency by photon counters, it is highly resistant to electronic noise. We developed a CW multiplier to supply high voltage to the AXEL detector and installed it at the 180 L prototype detector. The CW multiplier was implemented on a flexible printed circuit board and coated with methyl silicone resin as a discharge countermeasure. Measurements with the 180 L prototype detector confirmed that the pickup from the AC input of the CW multiplier to the signal line is sufficiently small, and stable operation of 40 days has been achieved. Data of the 2615 keV gamma rays from ^{208}Tl using thorium-doped tungsten rods as the source confirmed an energy resolution of $(0.672 \pm 0.083)\%$ FWHM. An interpolation based on the 2615 keV gamma rays from ^{208}Tl and other gamma ray peaks yielded an estimated energy resolution of $(0.678 \pm 0.010)\%$ at the Q value of $0\nu\beta\beta$, which is close to the design goal of 0.5 %. We also reconstructed the tracks of the 2615 keV gamma-ray photoabsorption events and double escape events, confirming that they each have one and two blob structures, respectively.

To eliminate background events using track information, we have developed a machine learning model to discriminate $0\nu\beta\beta$ and background. The model is constructed with DenseNet which is a convolutional

neural network (CNN) and was trained using simulation data generated by Geant4. We optimized the model by varying the normalization and voxelization resolution when inputting the tracks into the model. The model that normalized the voxel light intensity and divided the z -direction into 160 bins performed the best, achieving an accuracy of 91.7 % and an area under the curve (AUC) of 0.936 on the validation data. However, when inputting the real 2615 keV gamma-ray data from ^{208}Tl , differences were observed in the shape of the signal likelihood distribution of gamma-ray backgrounds. This indicates there are differences between the simulation data used for training the model and the real data. Comparing the distribution of summary quantities between the training data and the real data, differences were observed in the position of the centroid and the track volume distribution. The difference in the centroid position is thought to be due to the application of vetoes for high dark and dead channels in the real data. When applying vetoes to the simulation data, the distribution of the centroid position in x direction became closer. However, those of the track volume and the centroid position in z direction did not match, with the real data tending to be smaller. The cause of this is unknown, but one possible explanation is the angular dependence of the photoelectrons emitted by photoelectric absorption, combined with the position of the source. Using the performance of the model on the validation data, we estimated the sensitivity of a future 1-ton detector. Three types of background sources were considered: ^{214}Bi , ^{208}Tl , and ^{137}Xe . Among these, ^{214}Bi and ^{208}Tl were assumed to originate from the pressure vessel. Two types of materials for the pressure vessel were considered: SUS304 and oxygen-free copper. After selecting tracks at the size of the 180 L detector, which is the model training condition, we estimated the sensitivity by optimizing the cut range based on the energy ROI and the cut threshold of the CNN model. Using the energy resolution of 0.678 % obtained in this study and also an anticipated improvement to 0.32 %, we evaluated the sensitivity. With a 10-year measurement period, we obtained the best sensitivity: 90 % C.L. lower limits on the half-life of $0\nu\beta\beta$ of 1.12×10^{27} years for the case of 0.320 % FWHM energy resolution and oxygen-free copper pressure vessel. This is 2.9 times better than the current world record, and the corresponding effective neutrino mass $\langle m_{\beta\beta} \rangle$ is (16-71) meV using the same nuclear matrix elements as the result of KamLAND-Zen[27].

To further improve sensitivity in future studies, it is essential to reduce background and enhance signal efficiency. In particular, background originating from the pressure vessel can be mitigated by employing materials with lower radioactivity or by using active materials such as plastic scintillators to enable vetoing. From an analysis perspective, track deconvolution may allow for more detailed reconstruction of track shapes and improved energy resolution, potentially leading to higher signal efficiency. Additionally, the use of memory-efficient machine learning models, such as graph networks, may enable optimization by expanding the volume cut region, thereby contributing to improved signal efficiency.

Bibliography

- [1] F. Reines *et al.*, Phys. Rev., **117**, 159–173 (Jan 1960).
- [2] G. Danby *et al.*, Phys. Rev. Lett., **9**, 36–44 (Jul 1962).
- [3] M. L. Perl *et al.*, Phys. Rev. Lett., **35**, 1489–1492 (Dec 1975).
- [4] K. Kodama *et al.*, Physics Letters B, **504**(3), 218–224 (2001).
- [5] Physics Reports, **427**(5), 257–454 (2006).
- [6] Y. Fukuda *et al.*, Phys. Rev. Lett., **81**, 1562–1567 (Aug 1998).
- [7] Q. R. Ahmad *et al.*, Phys. Rev. Lett., **87**, 071301 (Jul 2001).
- [8] Q. R. Ahmad *et al.*, Phys. Rev. Lett., **89**, 011301 (Jun 2002).
- [9] I. Esteban *et al.*, Journal of High Energy Physics, **2024**(12), 216 (Dec 2024).
- [10] I. Esteban *et al.*, Nufit 6.0 (2024) (), <http://www.nu-fit.org> [Access date: May 13th, 2025].
- [11] M. Aker *et al.*, Direct neutrino-mass measurement based on 259 days of katrin data (2024), arXiv:2406.13516.
- [12] N. Aghanim *et al.*, Astronomy and Astrophysics, **641**, A6 (September 2020).
- [13] J. Shirai, F. Suekane, *Neutrino physics (in japanese)*, (Asakura Publishing Co., Ltd., 2021).
- [14] C. Giunti, M. Laveder, Neutrino mixing (2004), arXiv:hep-ph/0310238.
- [15] P. A. M. Dirac, R. H. Fowler, Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character, **117**(778), 610–624 (1928), <https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1928.0023>.
- [16] T. Yanagida, Conf. Proc. C, **7902131**, 95–99 (1979).
- [17] M. Gell-Mann, P. Ramond, R. Slansky, *Complex spinors and unified theories*, (North-Holland., 1979).
- [18] R. N. Mohapatra, G. Senjanović, Phys. Rev. Lett., **44**, 912–915 (Apr 1980).
- [19] M. Kobayashi, T. Maskawa, Progress of Theoretical Physics, **49**(2), 652–657 (02 1973), <https://academic.oup.com/ptp/article-pdf/49/2/652/5257692/49-2-652.pdf>.
- [20] M. Gavela *et al.*, Nuclear Physics B, **430**(2), 382–426 (1994).
- [21] P. Huet, E. Sather, Phys. Rev. D, **51**, 379–394 (Jan 1995).
- [22] K. Kajantie *et al.*, Phys. Rev. Lett., **77**, 2887–2890 (Sep 1996).
- [23] M. Fukugita, T. Yanagida, Physics Letters B, **174**(1), 45–47 (1986).
- [24] W. H. Furry, Phys. Rev., **56**, 1184–1193 (Dec 1939).
- [25] J. Schechter, J. W. F. Valle, Phys. Rev. D, **25**, 2951–2954 (Jun 1982).
- [26] F. T. Avignone, S. R. Elliott, J. Engel, Rev. Mod. Phys., **80**, 481–516 (Apr 2008).
- [27] S. Abe *et al.*, Search for majorana neutrinos with the complete kamland-zen dataset (2024), arXiv:2406.11438.
- [28] S. R. Elliott, P. Vogel, Annual Review of Nuclear and Particle Science, **52**(Volume 52, 2002), 115–151 (2002).
- [29] B. Pritychenko, Nuclear Physics A, **1033**, 122628 (2023).
- [30] JAEA, Tables of nuclear data (Access date: April 17th, 2025), <https://www.ndc.jaea.go.jp/NuC/>.
- [31] R. Arnold *et al.*, Phys. Rev. D, **93**, 112008 (Jun 2016).
- [32] M. Agostini *et al.*, Phys. Rev. Lett., **131**, 142501 (Oct 2023).

- [33] M. Agostini *et al.*, Phys. Rev. Lett., **125**, 252502 (Dec 2020).
- [34] R. Arnold *et al.*, The European Physical Journal C, **78**(10), 821 (Oct 2018).
- [35] J. Argyriades *et al.*, Nuclear Physics A, **847**(3), 168–179 (2010).
- [36] E. Armengaud *et al.*, The European Physical Journal C, **77**(11), 785 (Nov 2017).
- [37] E. Armengaud *et al.*, Phys. Rev. Lett., **126**, 181802 (May 2021).
- [38] R. Arnold *et al.*, Phys. Rev. D, **95**, 012007 (Jan 2017).
- [39] D. Q. Adams *et al.*, Phys. Rev. Lett., **126**, 171801 (Apr 2021).
- [40] D. Q. Adams *et al.*, Phys. Rev. Lett., **124**, 122501 (Mar 2020).
- [41] P. Novella *et al.*, Phys. Rev. C, **105**, 055501 (May 2022).
- [42] R. Arnold *et al.*, Phys. Rev. D, **94**, 072003 (Oct 2016).
- [43] T. K.-Z. collaboration *et al.*, Journal of Instrumentation, **16**(08), P08023 (aug 2021).
- [44] S. P. Ahlen, Rev. Mod. Phys., **52**, 121–173 (Jan 1980).
- [45] D. Anderson *et al.*, Nuclear Instruments and Methods, **163**(1), 125–134 (1979).
- [46] D. R. Nygren, eConf, **C740805**, 58 (1974).
- [47] L. Rogers *et al.*, Journal of Physics G: Nuclear and Particle Physics, **47**(7), 075001 (may 2020).
- [48] D. R. Nygren, The optimal detectors for wimp and 0-neutrino double beta decay searches: identical high pressure xenon gas tpc, In *2007 IEEE Nuclear Science Symposium Conference Record*, volume 2, pages 1051–1055 (2007).
- [49] V. Álvarez *et al.*, Journal of Instrumentation, **8**(09), P09011 (sep 2013).
- [50] A. Simón, on behalf of the NEXT Collaboration, Journal of Physics: Conference Series, **2374**(1), 012033 (nov 2022).
- [51] P. Novella *et al.*, Journal of High Energy Physics, **2023**(9), 190 (Sep 2023).
- [52] A. Simón *et al.*, Journal of High Energy Physics, **2021**(7), 146 (Jul 2021).
- [53] N. K. Byrnes *et al.*, Nature Communications, **15**(1), 10595 (Dec 2024).
- [54] W. Zhang *et al.*, Journal of Instrumentation, **18**(12), C12001 (dec 2023).
- [55] H. Lin, on behalf of the PandaX-III collaboration, Journal of Physics: Conference Series, **2502**(1), 012007 (may 2023).
- [56] T. Li *et al.*, Journal of High Energy Physics, **2021**(6), 106 (Jun 2021).
- [57] S. Xia *et al.*, Journal of Physics G: Nuclear and Particle Physics, **50**(12), 125103 (oct 2023).
- [58] T. Li *et al.*, Journal of High Energy Physics, **2023**(5), 200 (May 2023).
- [59] A. Bolotnikov, B. Ramsey, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, **396**(3), 360–370 (1997).
- [60] P. J. Linstrom, W. G. Mallard, Nist chemistry webbook, <https://doi.org/10.18434/T4D303> (2025), Retrieved May 23, 2025.
- [61] H. A. Koehler *et al.*, Phys. Rev. A, **9**, 768–781 (Feb 1974).
- [62] S. Akiyama, Master’s thesis, Kyoto University (2014).
- [63] S. Ban *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, **875**, 185–192 (2017).
- [64] S. Ban *et al.*, Progress of Theoretical and Experimental Physics, **2020**(3), 033H01 (03 2020), <https://academic.oup.com/ptep/article-pdf/2020/3/033H01/32970446/ptaa030.pdf>.
- [65] M. Yoshida *et al.*, Progress of Theoretical and Experimental Physics, **2024**(1), 013H01 (12 2023), <https://academic.oup.com/ptep/article-pdf/2024/1/013H01/55371817/ptad146.pdf>.
- [66] K. Nakamura, *Development of a large-sized high-pressure xenon gas time projection chamber for neutrinoless double beta decay search*, PhD thesis, 京都大学 (5 2022).
- [67] C. M. B. Monteiro *et al.*, Journal of Instrumentation, **2**(05), P05001 (may 2007).
- [68] M. Yoshida, *Development of a high-pressure xenon gas time projection chamber and evaluation of its performance at around the Q value of ^{136}Xe double-beta decay*, PhD thesis, 京都大学 (3 2024).

-
- [69] K. Z. Nakamura *et al.*, IEEE Transactions on Nuclear Science, **67**(7), 1772–1776 (2020).
 - [70] R. Honda *et al.*, Meeting Abstracts of the Physical Society of Japan, **71.2**, 223–223 (2016).
 - [71] F. Monrabal *et al.*, Journal of Instrumentation, **13**(12), P12010 – P12010 (December 2018).
 - [72] J. D. Cockcroft, E. T. S. Walton, E. Rutherford, Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character, **136**(830), 619–630 (1932), <https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1932.0107>.
 - [73] A. Marchionni *et al.*, Journal of Physics: Conference Series, **308**(1), 012006 (jul 2011).
 - [74] S. Akiyama *et al.*, Progress of Theoretical and Experimental Physics, **2025**(5), 053H03 (05 2025), <https://academic.oup.com/ptep/article-pdf/2025/5/053H03/63140723/ptaf066.pdf>.
 - [75] M. M. Weiner, Review of Scientific Instruments, **40**(2), 330–333 (02 1969), https://pubs.aip.org/aip/rsi/article-pdf/40/2/330/19117236/330_1_online.pdf.
 - [76] Motion-Project, motion (Access date: April 29th, 2025), <https://motion-project.github.io/index.html>.
 - [77] S. Pan, *Development of a high pressure xenon gas time projection chamber with a unique cellular readout structure to search for neutrinoless double beta decay*, PhD thesis, 京都大学 (5 2020).
 - [78] P. Novella *et al.*, Journal of High Energy Physics, **2018**(10), 112 (Oct 2018).
 - [79] S. Agostinelli *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, **506**(3), 250–303 (2003).
 - [80] M. Doi, T. Kotani, E. Takasugi, Progress of Theoretical Physics Supplement, **83**, 1–175 (03 1985), <https://academic.oup.com/ptps/article-pdf/doi/10.1143/PTPS.83.1/5227078/83-1.pdf>.
 - [81] Geant4Collaboration, Geant4 physics reference manual (Access date: June 7th, 2025), <https://geant4-userdoc.web.cern.ch/UsersGuides/PhysicsReferenceManual/BackupVersions/V10.5-2.0/html/index.html>.
 - [82] M. Gavrila, Phys. Rev., **113**, 514–526 (Jan 1959).
 - [83] H. Schindler, Garfield++ (Access date: April 30th, 2025), <https://garfieldpp.web.cern.ch/garfieldpp/>.
 - [84] N. Srivastava *et al.*, Journal of Machine Learning Research, **15**(56), 1929–1958 (2014).
 - [85] K. He *et al.*, Deep residual learning for image recognition (2015), arXiv:1512.03385.
 - [86] G. Huang *et al.*, Densely connected convolutional networks (2018), arXiv:1608.06993.
 - [87] S. Ioffe, C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift (2015), arXiv:1502.03167.
 - [88] D. Ulyanov, A. Vedaldi, V. Lempitsky, Instance normalization: The missing ingredient for fast stylization (2017), arXiv:1607.08022.
 - [89] A. Paszke *et al.*, Pytorch: An imperative style, high-performance deep learning library, In H. Wallach *et al.*, editors, *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc. (2019).
 - [90] B. Polyak, Ussr Computational Mathematics and Mathematical Physics, **4**, 1–17 (12 1964).
 - [91] N. Y. E., Dokl. Akad. Nauk SSSR, **269**, 543–547 (1983).
 - [92] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization (2017), arXiv:1412.6980.
 - [93] E. R. DeLong, D. M. DeLong, D. L. Clarke-Pearson, Biometrics, **44**(3), 837–845 (1988).
 - [94] X. Sun, W. Xu, IEEE Signal Processing Letters, **21**(11), 1389–1393 (2014).
 - [95] G. J. Feldman, R. D. Cousins, Phys. Rev. D, **57**, 3873–3889 (Apr 1998).
 - [96] E. Aprile *et al.*, The European Physical Journal C, **82**(7), 599 (Jul 2022).
 - [97] D. Leonard *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, **591**(3), 490–509 (2008).
 - [98] NNDC, Nudat3 decay radiation search (Access date: June 2nd, 2025), https://www.nndc.bnl.gov/nudat3/indx_dec.jsp.

- [99] J. I. Muñoz Vidal, *The NEXT path to neutrino inverse hierarchy*, PhD thesis, U. Valencia (main) (2018).
- [100] J. A. Formaggio, C. Martoff, Annual Review of Nuclear and Particle Science, **54**(Volume 54, 2004), 361–412 (2004).
- [101] NIST, Xcom (Access date: June 2nd, 2025), <https://physics.nist.gov/PhysRefData/Xcom/html/xcom1.html>.
- [102] J. Gómez-Cadenas *et al.*, Journal of Cosmology and Astroparticle Physics, **2011**(06), 007 (jun 2011).
- [103] B. Graham, L. van der Maaten, Submanifold sparse convolutional networks (2017), arXiv:1706.01307.
- [104] R. Abbasi *et al.*, Journal of Instrumentation, **17**(11), P11003 (nov 2022).
- [105] L. B. Lucy, Astron. J., **79**, 745–754 (1974).