

Master Thesis

Operation of the New T2K Near Detector SuperFGD
and Development of Electron-Neutrino Event
Selection
(新型T2K前置検出器SuperFGDの運用と電子
ニュートリノ事象選択の開発)

January 27, 2025

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

Tomochika Arai
新居 智将

Abstract

The T2K (Tokai to Kamioka) experiment is a long-baseline neutrino oscillation experiment aiming to search for CP violation in the neutrino sector. The systematic uncertainty on the ν_e and $\bar{\nu}_e$ cross-sections is one of the main source of systematic uncertainties in the oscillation measurements. The near detector ND280 is used to measure the neutrino cross sections. Recently, ND280 was upgraded to further reduce systematic uncertainties by addressing limitation of the original detectors. SuperFGD, which is a plastic scintillation tracker and a neutrino target, is one of the detectors newly installed in ND280.

In this thesis, we conducted the study to acquire and analyze data in SuperFGD. First, we developed an online gain monitor to keep taking good-quality data. The online gain monitor needs a new fast gain calculation method, and we developed a method that meets both time and accuracy requirements. The developed online monitor is currently being used in the detector operation. Next, the acquired data is used for event reconstruction as a first step of the analysis. We evaluated the impact of bad channels on track reconstruction. We revealed that clustered bad channels have a greater impact on track reconstruction than discrete bad channels. Additionally, this study enables an estimation of the effect on track reconstruction from bad channel fraction and distribution. Finally, based on the reconstructed data, the event selection is performed. We explored the method using an artificial neural network to separate an electromagnetic shower caused by an electron from that caused by a photon for the purpose of rejecting γ backgrounds in the ν_e event selection. The model we developed succeeded in the rejection of 95.5 % of γ background events with 80 % efficiency for electrons.

Contents

1	Neutrino Physics	4
1.1	Neutrino oscillation	4
1.1.1	Neutrino physics and neutrino mixing	4
1.1.2	Neutrino oscillation	5
1.1.3	CP-violation	6
1.2	Neutrino interaction	7
1.2.1	Charged-current (CC) interaction	7
1.2.2	Electron neutrino cross section	8
2	T2K Experiment	9
2.1	Overview	9
2.2	Neutrino Beamline at J-PARC	10
2.2.1	Neutrino Beamline	10
2.2.2	Beam upgrade	11
2.2.3	Off-axis method	11
2.3	Near Detectors	14
2.3.1	Components of classic ND280	14
2.3.2	ND280 Upgrade	17
2.4	Far Detector	18
2.5	Recent results	19
2.5.1	Measurement of neutrino oscillations	19
2.5.2	ν_e -CC cross-sections measurement	19
2.6	Upgrade and Operation Status of ND280 Upgrade	22
3	SuperFGD	23
3.1	Overview	23
3.2	Scintillators, WLS fibers, and MPPCs	24
3.3	Electronics	24
3.4	Calibration System	25
3.5	Commissioning, Installation, and Operation	26
3.6	Motivation of this thesis	27
4	Online Monitoring	28
4.1	Online Monitoring System for ND280	28
4.2	SuperFGD Online Monitor	28
4.2.1	Hit Summary	29

4.2.2	ADC Spectra	29
4.2.3	Event Summary	30
4.2.4	Hit Timing Summary	30
4.2.5	Gain monitor	30
4.3	Requirements for the online gain monitor	32
4.3.1	LED data for the gain monitoring	32
4.3.2	General method for gain calculation	32
4.3.3	Gain calculation for calibration and time constraint for the monitor	32
4.4	Gain calculation for Online Monitor	34
4.4.1	Ideas for the method	34
4.4.2	Procedure of online gain calculation	34
4.4.3	Validation by the comparison with offline method	37
4.4.4	Test with histogramming data	39
4.5	Operation of SuperFGD Online Monitor in the beam run	43
5	Evaluation of the effect of bad channels on track reconstruction	44
5.1	Bad channels during the beam run	44
5.2	Simulation and reconstruction	46
5.2.1	Monte Carlo simulation	46
5.2.2	Track reconstruction	46
5.3	Evaluation method	47
5.4	2D hit distributions with bad channels	49
5.5	3D hit reconstruction with bad channels	50
5.6	Track reconstruction with bad channels	55
5.6.1	Efficiency in the track reconstruction	55
5.6.2	Discussion	57
6	Separation of electrons from gamma background	62
6.1	e/γ separation	62
6.1.1	e/γ separation in the current ν_e event selection steps	62
6.1.2	γ backgrounds	63
6.1.3	Motivation	63
6.2	Convolutional Neural Networks	64
6.3	e/γ separation with CNN	64
6.3.1	Models	64
6.3.2	MC samples for the training, validation, and test	67
6.3.3	Training	69
6.4	Results	70
6.5	Discussion	73
7	Summary	76
A	Track Reconstruction	78
A.1	3D hit reconstruction	78
A.2	track reconstruction	79

Chapter 1

Neutrino Physics

1.1 Neutrino oscillation

1.1.1 Neutrino physics and neutrino mixing

The Standard Model (SM) is the theory describing electromagnetic, strong, and weak interactions. Neutrinos are the elementary particles in the SM which interact only via the weak interaction and the gravity. In the SM, there are neutrinos of three flavors: electron neutrino (ν_e), muon neutrino (ν_μ), and tau neutrino (ν_τ). There are also antineutrinos as antiparticles of them, and electron, muon, and tau antineutrinos are described as $\bar{\nu}_e$, $\bar{\nu}_\mu$, $\bar{\nu}_\tau$, respectively.

SM is based on the mathematical framework of quantum field theory. Therefore, the general state of a particle can be described as a superposition of eigenstates. In particular, the state of neutrino $|\nu\rangle$ can be written as a superposition of flavor eigenstates: $|\nu_e\rangle$, $|\nu_\mu\rangle$, and $|\nu_\tau\rangle$. Furthermore, the state of neutrino can be also written as a superposition of mass eigenstates. If the mass eigenvalues (m_1, m_2, m_3) are all different, it can be written as a superposition of mass eigenstates: $|\nu_1\rangle$, $|\nu_2\rangle$, and $|\nu_3\rangle$. The change of basis between these sets of eigenstates can be written using a 3×3 unitary matrix U as

$$\begin{pmatrix} |\nu_e\rangle \\ |\nu_\mu\rangle \\ |\nu_\tau\rangle \end{pmatrix} = U \begin{pmatrix} |\nu_1\rangle \\ |\nu_2\rangle \\ |\nu_3\rangle \end{pmatrix} \equiv \begin{pmatrix} U_{e,1} & U_{e,2} & U_{e,3} \\ U_{\mu,1} & U_{\mu,2} & U_{\mu,3} \\ U_{\tau,1} & U_{\tau,2} & U_{\tau,3} \end{pmatrix} \begin{pmatrix} |\nu_1\rangle \\ |\nu_2\rangle \\ |\nu_3\rangle \end{pmatrix}.$$

This unitary matrix U is called as Pontecorvo–Maki–Nakagawa–Sakata matrix (PMNS) matrix. U belongs to the unitary group $U(3)$, and the group has nine real degrees of freedom. Five of those are absorbed as relative phase parameters and physically irrelevant. U is generally written using four real parameters θ_{12} , θ_{13} , θ_{23} , and δ_{CP} as

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{CP}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{CP}} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (1.1)$$

$$= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{CP}} \\ -s_{12}c_{23} - c_{12}s_{13}s_{23}e^{i\delta_{CP}} & c_{12}c_{23} - s_{12}s_{13}s_{23}e^{i\delta_{CP}} & c_{13}s_{23} \\ s_{12}s_{23} - c_{12}s_{13}c_{23}e^{i\delta_{CP}} & -c_{12}s_{23} - s_{12}s_{13}c_{23}e^{i\delta_{CP}} & c_{13}c_{23} \end{pmatrix}, \quad (1.2)$$

where $s_{ij} = \sin \theta_{i,j}$ and $c_{ij} = \cos \theta_{i,j}$. The range of the parameters are $0 \leq \theta_{ij} \leq (\pi/2)$ and $-\pi \leq \delta_{CP} \leq \pi$. These θ_{12} , θ_{13} , and θ_{23} are called the mixing angles, and δ_{CP} is called the CP-violating phase.

1.1.2 Neutrino oscillation

Neutrino oscillation is the phenomenon that the flavors of neutrino change into each other while traveling, and it is explained by using quantum mechanics with the fact that neutrino has non-zero mass.

First, consider the neutrino as a free particle. Since the degrees of freedom in the mass (or flavors) are distinct from that in the spatial parameters, there are simultaneous eigenstates of the mass and the momentum. By restricting the statement that the state of neutrino is written as a superposition of mass eigenstates to the momentum eigenspace, the state of the neutrino in energy eigenstates can be written as a superposition of simultaneous eigenstates of the mass and the momentum. The following discussion is conducted in the momentum eigenspace with the eigenvalues $\mathbf{p} = (p_x, p_y, p_z)$. The free particle of mass eigenstate can consider as a plane wave, and the time evolution can be written as

$$|\nu_k(t)\rangle = \exp(-iE_k t) |\nu_k(0)\rangle \quad (k = 1, 2, 3),$$

where $E_k = \sqrt{p^2 + m_k^2}$ with $p = |\mathbf{p}|$.

Then, $P(\nu_\alpha \rightarrow \nu_\beta)$, which is the probability that the neutrino produced with flavor α at $t = 0$ is detected as a neutrino with flavor β at $t = t_0$ is calculated as below.

$$P(\nu_\alpha \rightarrow \nu_\beta) = |\langle \nu_\beta(0) | \nu_\alpha(t_0) \rangle|^2 \quad (1.3)$$

$$= \left| \sum_{j,k} U_{\alpha,j} U_{\beta,k}^* \langle \nu_k(0) | \nu_j(t_0) \rangle \right|^2 \quad (1.4)$$

$$= \left| \sum_k U_{\alpha,k} U_{\beta,k}^* \exp(-iE_k t) \right|^2 \quad (1.5)$$

$$= \delta_{\alpha,\beta} - 4 \sum_{j < k} \text{Re} [U_{\alpha,j} U_{\beta,j}^* U_{\alpha,k}^* U_{\beta,k}] \sin^2 \frac{(E_k - E_j)t}{2} \quad (1.6)$$

$$+ 2 \sum_{j < k} \text{Im} [U_{\alpha,j} U_{\beta,j}^* U_{\alpha,k}^* U_{\beta,k}] \sin(E_k - E_j)t. \quad (1.7)$$

Under the ultra-relativistic approximations, E_k is approximated to $E_k = \sqrt{p^2 + m_k^2} \simeq p + \frac{m_k^2}{2p}$. Since $E_1 \simeq E_2 \simeq E_3 \simeq p$, by setting this value as E , this formula is written as $E_k \simeq E + \frac{m_k^2}{2E}$. t can be also replaced by L , which is the distance the particle traveled, since the speed of the particle is almost equal to the speed of light.

Then, $P(\nu_\alpha \rightarrow \nu_\beta)$ is written as

$$P(\nu_\alpha \rightarrow \nu_\beta) = \delta_{\alpha,\beta} - 4 \sum_{j < k} \text{Re} [U_{\alpha,j} U_{\beta,j}^* U_{\alpha,k}^* U_{\beta,k}] \sin^2 \frac{\Delta m_{kj}^2 L}{4E} \quad (1.8)$$

$$+ 2 \sum_{j < k} \text{Im} [U_{\alpha,j} U_{\beta,j}^* U_{\alpha,k}^* U_{\beta,k}] \sin \frac{\Delta m_{kj}^2 L}{2E}, \quad (1.9)$$

where Δm_{kj}^2 is defined as $\Delta m_{kj}^2 \equiv m_k^2 - m_j^2$. The probability of the oscillation of antineutrino $P(\bar{\nu}_\alpha \rightarrow \bar{\nu}_\beta)$ is written as

$$P(\bar{\nu}_\alpha \rightarrow \bar{\nu}_\beta) = \delta_{\alpha,\beta} - 4 \sum_{j < k} \text{Re} [U_{\alpha,j} U_{\beta,j}^* U_{\alpha,k}^* U_{\beta,k}] \sin^2 \frac{\Delta m_{kj}^2 L}{4E} \quad (1.10)$$

$$- 2 \sum_{j < k} \text{Im} [U_{\alpha,j} U_{\beta,j}^* U_{\alpha,k}^* U_{\beta,k}] \sin \frac{\Delta m_{kj}^2 L}{2E}. \quad (1.11)$$

When considering the effect of matter, the behavior is similar under the adiabatic approximation.

1.1.3 CP-violation

In the probability of (anti-)neutrino oscillations, since $U_{\alpha,j} U_{\beta,j}^* U_{\alpha,k}^* U_{\beta,k}$ is a real number when $\alpha = \beta$, $P(\nu_\alpha \rightarrow \nu_\alpha) = P(\bar{\nu}_\alpha \rightarrow \bar{\nu}_\alpha)$ always stands for each $\alpha = e, \mu, \tau$ regardless of parameters. Then, the difference of the probability of (anti-)neutrino oscillations can be only seen in the probability of the oscillation between different flavor neutrinos. The experiment to measure the difference of the probability mostly uses the oscillation from ν_μ ($\bar{\nu}_\mu$) to ν_e ($\bar{\nu}_e$) for the reason of neutrino production process and the energy requirement for the interaction.

The difference between $P(\nu_\mu \rightarrow \nu_e)$ and $P(\bar{\nu}_\mu \rightarrow \bar{\nu}_e)$ is written as

$$P(\nu_\mu \rightarrow \nu_e) - P(\bar{\nu}_\mu \rightarrow \bar{\nu}_e) \quad (1.12)$$

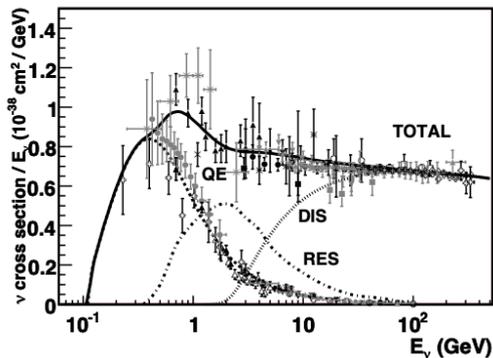
$$= 4 \sum_{j < k} \text{Im} [U_{\alpha,j} U_{\beta,j}^* U_{\alpha,k}^* U_{\beta,k}] \sin \frac{\Delta m_{kj}^2 L}{2E} \quad (1.13)$$

$$= -16 s_{12} c_{12} s_{23} c_{23} s_{13} c_{13}^2 \sin \delta_{CP} \sin \left(\frac{\Delta m_{21}^2 L}{4E} \right) \sin \left(\frac{\Delta m_{31}^2 L}{4E} \right) \sin \left(\frac{\Delta m_{32}^2 L}{4E} \right), \quad (1.14)$$

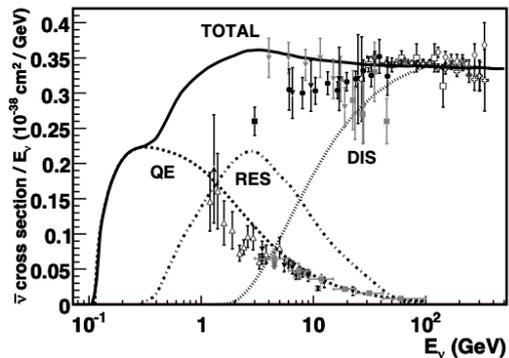
where s_{kj} , c_{kj} is same with that defined in Eq. 1.1. The value is non-zero only when all of the conditions are satisfied.

- $\theta_{jk} \neq 0, \frac{\pi}{2}$,
- $\Delta m_{kj} \neq 0$,
- $\delta_{CP} \neq 0, \pi$.

In particular, CP-symmetry is violated if the value is non-zero.



(a) ν_μ -CC cross sections



(b) $\bar{\nu}_\mu$ -CC cross sections

Figure 1.1. The measured cross sections of ν_μ -CC ($\bar{\nu}_\mu$ -CC) inclusive interaction and each type of ν_μ -CC ($\bar{\nu}_\mu$ -CC) interaction. [1] The marker displays the experiment result. The main neutrino source in this energy region is atmospheric neutrino and accelerator neutrino. The predictions line is by the NUANCE generator [2].

1.2 Neutrino interaction

The neutrino interacts with other particles only via weak interactions. There are two types of weak interaction. One is neutral-current (NC) interaction, which is mediated by the Z boson. The other is charged-current (CC) interaction, which is mediated by the W boson. NC interaction of neutrino with a nucleon does not produce the charged lepton, so it is impossible to determine the flavor of the interacted neutrino. In contrast to the NC interaction, the final state of CC interaction includes a charged lepton corresponding to the flavor of the neutrino in the initial state.

1.2.1 Charged-current (CC) interaction

There are some processes in the CC interaction of the neutrino with nucleus. In the energy region from sub-GeV to a few GeV, the dominant processes are CC quasi-elastic scattering (CCQE), CC resonance scattering (CC-RES), and CC deep inelastic scattering (CC-DIS). The final state depends on the process.

CCQE interaction is an elastic interaction between a neutrino and a nucleon, which is described as $\nu_l + n \rightarrow l^- + p$ or $\bar{\nu}_l + p \rightarrow l^+ + n$, where $l = e, \mu, \tau$. CCQE process have a charged lepton and a nucleon in the final state. In the CC-RES interaction process, the scattered nucleon is excited to a resonance state, and it decays to nucleons and mesons. The baryon resonance dominantly decay to a nucleon and a (charged or neutral) pion. A charged lepton is also included in the final state. In the CC-DIS interaction process, the scattered nucleon produce hadron jets. Then, the final state includes a charged lepton and several hadrons.

The dominant process depends on the energy of neutrino. The measured cross-section of ν_μ and $\bar{\nu}_\mu$ with each process are summarized in Fig. 1.1. The ν_μ -CC ($\bar{\nu}_\mu$ -CC) inclusive cross-section, which is the sum of the cross-sections with the processes of CC interaction, is also shown in Fig. 1.1.

1.2.2 Electron neutrino cross section

In the experiment that measure the neutrino oscillation of $\nu_\mu \rightarrow \nu_e$ ($\bar{\nu}_\mu \rightarrow \bar{\nu}_e$) to search for CP-violation, the uncertainty in the cross-section of ν_e -CC interactions contributes to the uncertainties in the measurement of neutrino oscillation. This oscillation is mainly measured by using neutrino produced by accelerators, and the accelerator neutrino has the energy of about 1 GeV. Thus, the precise ν_e -CC cross-section measurement in the momentum range around 1 GeV is necessary to measure the $\nu_\mu \rightarrow \nu_e$ oscillation precisely.

Currently, the interaction of the neutrino in the energy region from sub-GeV to several GeV is investigated by using accelerator neutrino. Due to the production process, neutrino beam from accelerator is dominated by ν_μ and $\bar{\nu}_\mu$, and ν_e and $\bar{\nu}_e$ account for only a small part of the neutrino beam flux. Thus, along with the small cross-section of neutrino, the ν_e cross-section measurement in the energy region is limited by small statistics. Therefore, high intensity neutrino beam and massive target is necessary for the measurement.

Additionally, the contamination of the background including photons causes a large uncertainty in the ν_e cross-section measurement. Hence, the ν_e event selection with high efficiency and purity is also necessary.

Chapter 2

T2K Experiment

2.1 Overview

The T2K (Tokai to Kamioka) experiment [3] is a long baseline neutrino oscillation experiment. The T2K experiment has a goal of determining CP-violation in the neutrino sector with more than 3σ confidence level by measuring the $\nu_\mu \rightarrow \nu_e$ ($\bar{\nu}_\mu \rightarrow \bar{\nu}_e$) oscillations precisely.

The overview of the T2K experiment is shown in Fig. 2.1. In the T2K experiment, ν_μ ($\bar{\nu}_\mu$) beam produced in the Japan Proton Accelerator Research Complex (J-PARC) at Tokai, Ibaraki is used as a neutrino source. The neutrinos are detected at detectors located in two different places: near detectors and the far detector. Near detectors are located 280 m downstream from the neutrino production target. The far detector, Super-Kamiokande (SK) at Kamioka, Gifu, is located 295 km downstream from the target.

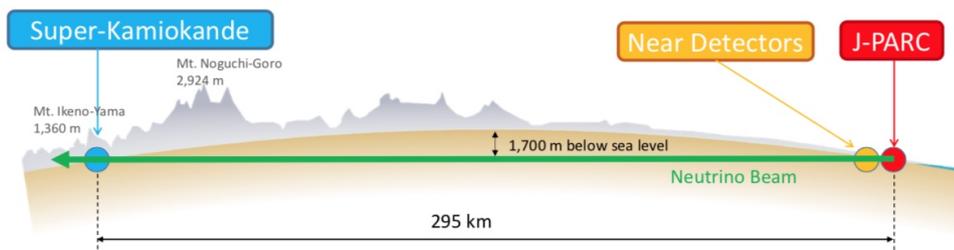


Figure 2.1. Overview of the T2K experiment.

2.2 Neutrino Beamline at J-PARC

2.2.1 Neutrino Beamline

J-PARC accelerator complex consists of three accelerators: the 400 MeV linear accelerator (LINAC), the 3 GeV Rapid-Cycling Synchrotron (RCS), and the 30 GeV Main Ring (MR). The overview of J-PARC accelerator complex is shown in Fig. 2.2 (a). The Hydrogen anions are accelerated up to 400 MeV by the LINAC, and converted to proton beam at the injection point to RCS. The protons are accelerated up to 3 GeV by the RCS, and finally, are accelerated up to 30 GeV by the MR. The protons are accelerated as bunches. There are eight bunches of the protons in a cycle (called “spill”) in the MR, and the bunches in a cycle are extracted to the neutrino beamline at intervals of around 580 ns. The next spill is injected from RCS after extract previous spill. The spill cycle is 1.36 s since 2023.

The protons from MR are injected into the neutrino beamline. The neutrino beamline consists of two parts: the primary beamline and the secondary beamline. In the primary beamline, the proton beam is bent and focused on the target. In the secondary beamline, the beam interacts with the target. The produced charged pions are focused by electromagnetic horns, and finally, the neutrino beam is produced by decaying charged pions. The overview of neutrino beamline is shown in Fig. 2.2 (b).

The target of the beam in the neutrino beamline is a graphite rod with 2.6 cm diameter and 91.4 cm long, which corresponds to 1.9 interaction length. As a result of the hadron interaction between the proton beam and graphite target, secondary pions are produced. Some amount of kaons are also produced.

The electromagnetic horns collect and focus the charged mesons to increase the neutrino flux at detectors. The electromagnetic horns can collect only positive or negative mesons, depending on the operation current. The charge of the meson decides whether the meson decays into neutrino or antineutrino, as described below. Thus, the neutrino beam can be either neutrino-dominant (called “Forward Horn Current” (FHC) or neutrino mode) or antineutrino-dominant (called “Reversed Horn Current” (RHC) or antineutrino mode) depending on the operation current of the horns.

The focused mesons decay in the decay volume and produce the neutrino beam. The major decay modes of charged pions and kaons are:

- $\pi^+ \rightarrow \mu^+ + \nu_\mu$, $\pi^- \rightarrow \mu^- + \bar{\nu}_\mu$,
- $K^+ \rightarrow \mu^+ + \nu_\mu$, $K^- \rightarrow \mu^- + \bar{\nu}_\mu$,
- $K^+ \rightarrow \pi^+ + \pi^0$, $K^- \rightarrow \pi^- + \pi^0$.

Third mode also produce neutrinos in the end through charged pions. These modes all produce muon (anti-)neutrino, then, the neutrino beam is dominated by ν_μ (in FHC) or $\bar{\nu}_\mu$ (in RHC). Additionally, some amounts of electron (anti-)neutrino exists in the neutrino beam. The major decay modes producing ν_e or $\bar{\nu}_e$ are:

- $K^+ \rightarrow \pi^0 + e^+ + \nu_e$, $K^- \rightarrow \pi^0 + e^- + \bar{\nu}_e$,
- $\mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu$, $\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu$.

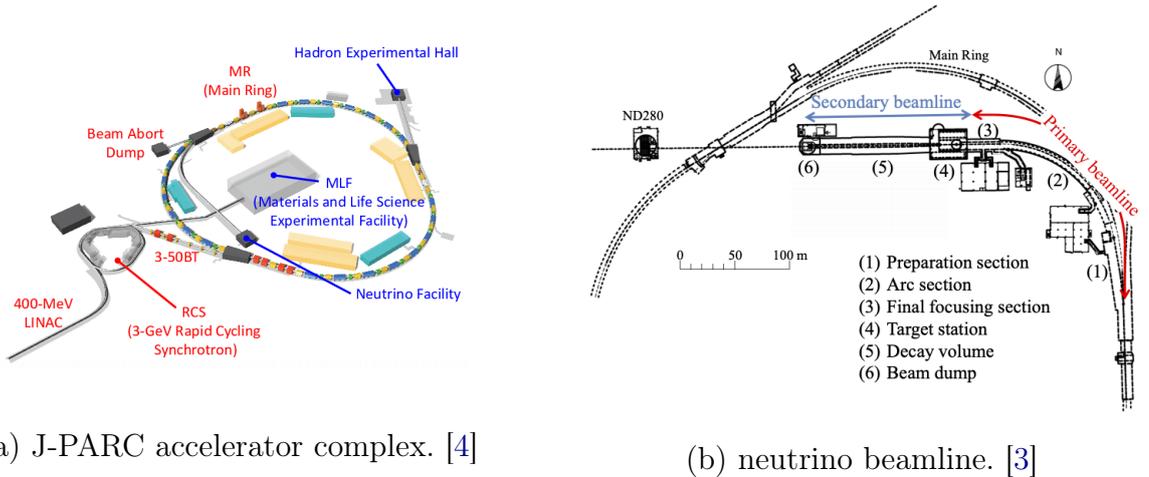


Figure 2.2. (a) is the overview of the J-PARC accelerator complex. (b) is the overview of the neutrino beamline at J-PARC. Neutrino beamline is also shown in the center of (a) without notation.

The fluxes of ν_μ , $\bar{\nu}_\mu$, ν_e , and $\bar{\nu}_e$ at ND280 in FHC and RHC mode are shown in Fig. 2.3. The flux of ν_e ($\bar{\nu}_e$) is less than 1 % of the flux of ν_μ ($\bar{\nu}_\mu$) in the neutrino (antineutrino) mode.

The number of protons shot to the target is called “protons on target” (POT). Under the assumption that the operation current of the horn is constant, the flux of the neutrino beam is in proportion to the number of charged mesons, and the number of charged mesons is in proportion to the number of protons shot to the target. Thus, the value of POT is used as an indicator of the statistics in T2K experiment. The accumulation of POT in the run is shown in Fig. 2.4. The data of 2.80×10^{21} POT with FHC mode and 1.77×10^{21} POT with RHC mode (the POTs with different horn current are simply summed up) has been acquired.

2.2.2 Beam upgrade

The beam power of MR reached to 515 kW in 2019. To increase the beam power further, the upgrade of MR aim to a beam power of 1.3 MW is underway. The plan to upgrade MR beam power is shown in Fig. 2.5.

Indeed, the beam power reached to 750 kW in 2023, and increased to 820 kW by the end of 2024. In addition to the upgrade of MR, the current applied to the electromagnetic horn was increased from ± 250 kA to ± 320 kA in 2023. The sign depends on the beam mode. The neutrino beam intensity was increased by about 10 % by the increase of the horn current.

The data acquisition in the T2K with upgraded beam is described in Section 2.6.

2.2.3 Off-axis method

In the T2K experiment, the neutrino beam is directed at an angle of 2.5 degree with respect to the baseline connecting between the proton target and SK. By doing this, the

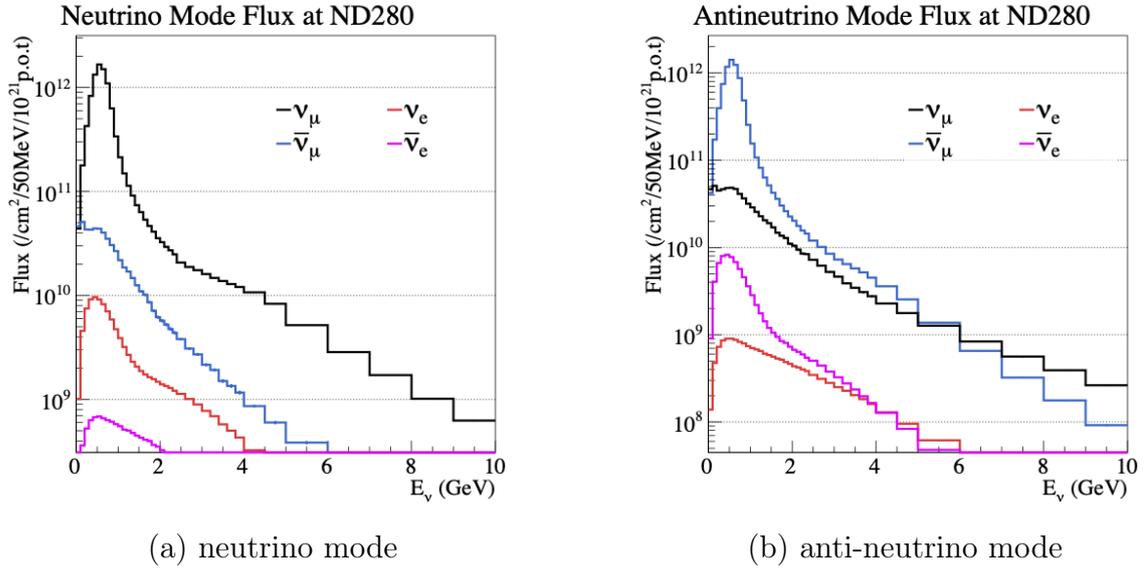


Figure 2.3. The fluxes of ν_μ , $\bar{\nu}_\mu$, ν_e , and $\bar{\nu}_e$ at near detectors. [5]

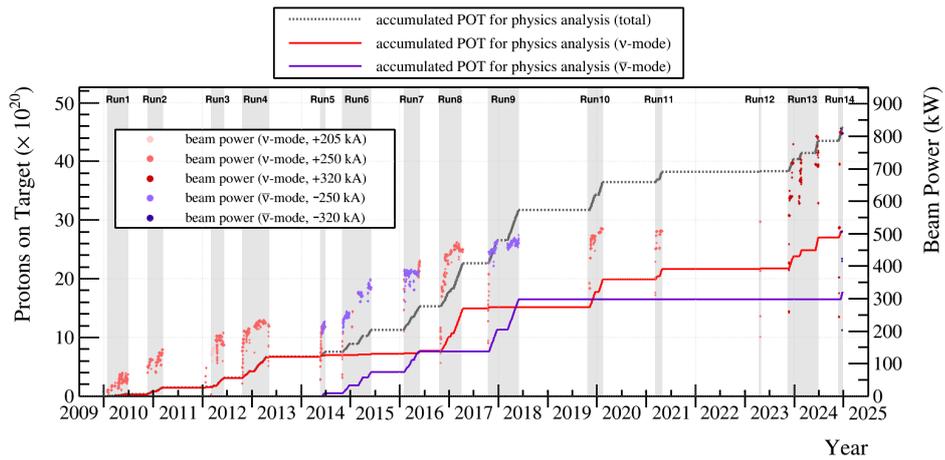


Figure 2.4. The accumulation of POT over time in the T2K experiment. The plot is produced by T2K beam group.

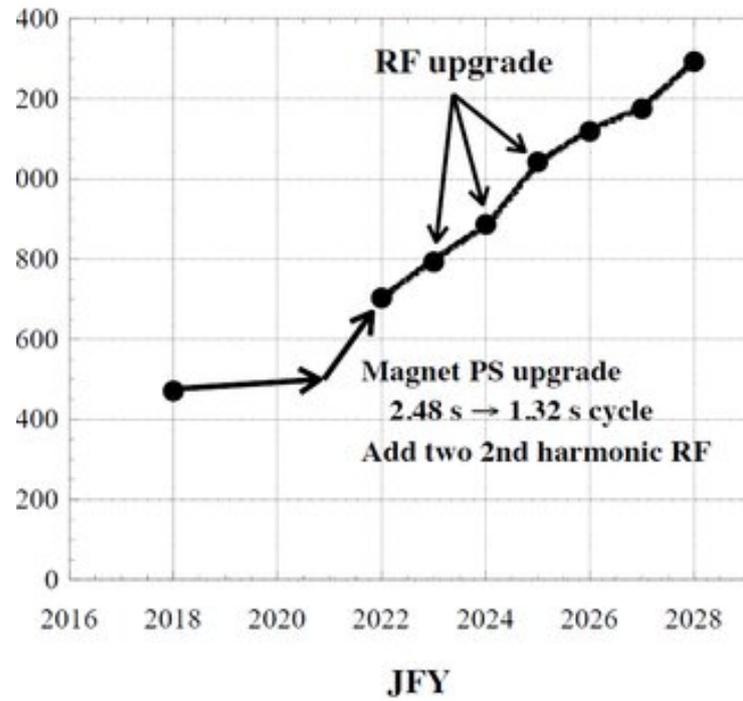


Figure 2.5. The plan to upgrade MR beam power. [4]

energy distribution of the neutrino delivered to the detectors becomes narrow and has a peak near the value, which maximizes the probability of the appearance of electron (anti-) neutrinos at SK, as shown in Fig. 2.6. This is called the off-axis method.

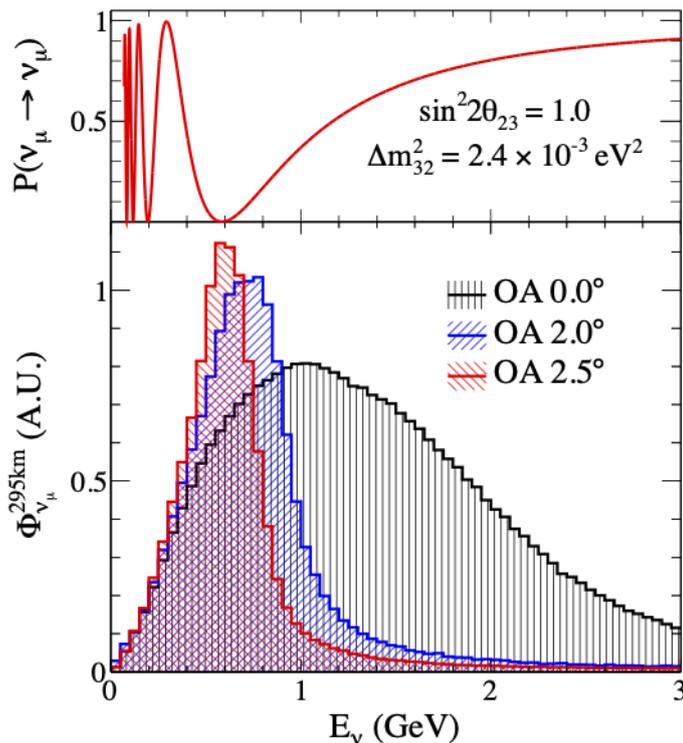


Figure 2.6. The estimated muon neutrino survival probability at the far detector and neutrino fluxes for different off-axis angles. [6]

2.3 Near Detectors

Near detectors consist of three detectors: INGRID, WAGASCI-BabyMIND detector, and ND280. Near detectors are all located 280 m downstream from the proton target while the angle from the beam direction is different.

INGRID is an on-axis detector located in the center of the beam. The detector is composed of an array of modules, and each module has iron target plates and tracking scintillator planes. The detector measures the profile and stability of the neutrino beam.

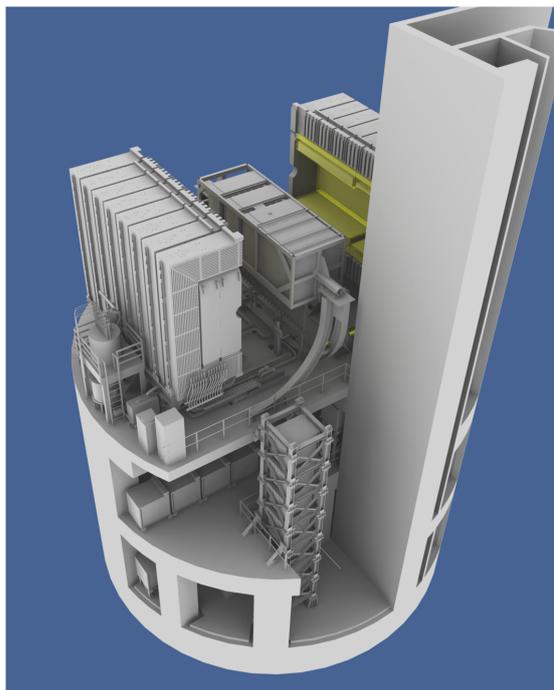
WAGASCI-BabyMIND detector is an off-axis detector with an angle of 1.5 degree. The detector consists of the neutrino detector with water and scintillator bars at the front and muon range detectors at the rear. The detector is designed to study neutrino-nucleus interactions.

The details of the ND280 are described below.

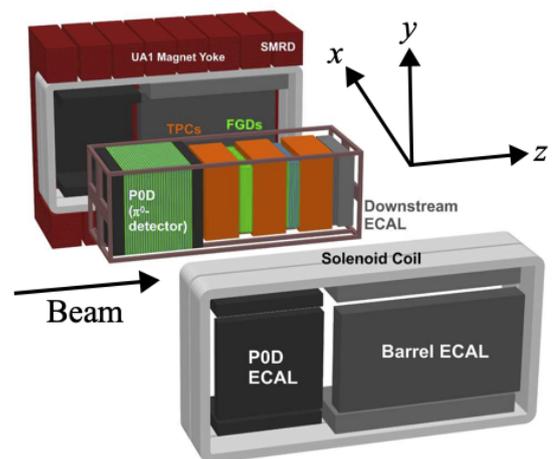
2.3.1 Components of classic ND280

ND280 is an off-axis detector with an angle of 2.5 degree, which is equal to the angle of SK. It is placed to constrain the uncertainties on the neutrino flux and interactions. It is surrounded by a dipole magnet, which produces a uniform horizontal magnetic field of 0.2 T.

ND280 was upgraded in 2024. However, most of the subdetectors before the upgrade are still used after the upgrade. The detector before the upgrade is called “classic ND280”.



(a) Overview of near detectors. [3]



(b) Overview of ND280. [3]

Figure 2.7. (a) is an illust of the near detectors in the pit. ND280 and INGRID are shown in the top and middle of the figure, respectively. (b) is an exploded view of the ND280. the magnet is closed when acquiring the data.

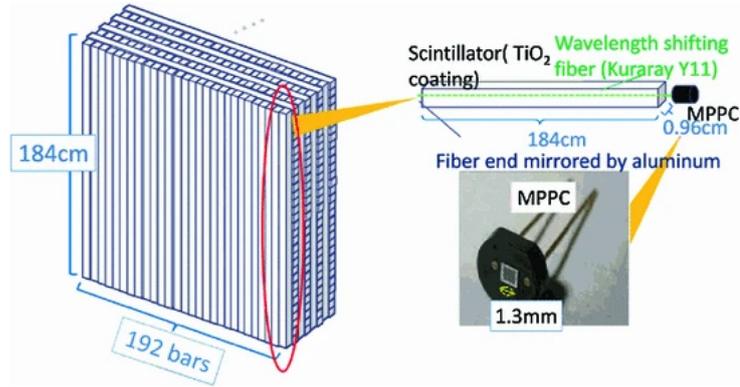


Figure 2.8. Overview of Fine grained detector (FGD). [7]

The components of the classic ND280 are below:

Fine grained detector (FGD)

Fine grained detector (FGD) is a neutrino target and scintillation tracker to reconstruct the trajectory of charged particle emitted by neutrino interaction. There are two FGDs in the ND280 as shown in Fig. 2.7 (b).

The overview of FGD 1 (located upstream in FGDs) is shown in Fig. 2.8. FGD 1 consists of 30 layers with 192 plastic scintillator bars per layer. FGD 2 (located downstream in FGDs) consists of 14 plastic layers of plastic scintillator bars and six water target modules. The water modules are inserted every two layers. There are two types of the layer depending on the direction of the scintillator bars. Every layer is in perpendicular to the beam direction, and different types of layers are lined up alternately. The bar is vertical direction in the layers of one type while horizontal direction in the layers of another type. The scintillation light from the bar is read by MPPC. The size of each FGD is 2300 mm \times 2300 mm \times 365 mm.

Time projection chamber (TPC)

Time projection chamber (TPC) is a tracker using the ionization of gas molecules by charged particles along the track. The precise tracking enables precise momentum reconstruction of the charged particle, and the distribution of the amount of ionization can be used to the particle identification. There are three TPCs in ND280. The active volume of TPC is filled by argon-based drift gas. Charged particles passing through the TPC produce ionization electrons and ionize gas molecules, and the electrons move to anode (readout) planes along the electric drift field. Anode plane is segmented with pads of 7.0 mm \times 9.8 mm, and the drift electrons are amplified and sampled in every pads. By combining timing information of the signal, TPC enables 3D tracking of charged particles. The relative momentum resolution for the direction in perpendicular to the magnetic field is less than 10 % in the momentum region below 1 GeV.

Pi-zero detector (PØD)

Pi-zero detector (PØD) is located at the most front part of the classic ND280. PØD aims to study about the neutral current process $\nu_\mu + N \rightarrow \nu_\mu + N + \pi^0 + X$ in the water. PØD consists of the two water target modules sandwiched by two electromagnetic calorimeters. Water target modules are composed of interleaved layers of scintillator bars, fillable water target bag, and brass, while calorimeter modules are composed of interleaved layers of scintillator bars and lead.

PØD detector was removed to install upgrade detectors as described in the next section. However, only the upstream electromagnetic calorimeter in PØD (called Upstream ECal) remains in the ND280 upgrade.

Electromagnetic calorimeter (ECal)

The electromagnetic calorimeter (ECal) is a sampling electromagnetic calorimeter that surrounds FGD, TPC, and PØD. ECal is used to complement event reconstruction by the detection and reconstruction of photons and by providing reconstruction information on charged particles.

Side muon range detector (SMRD)

Side muon range detector (SMRD) consists of the modules of plastic scintillation counters and they are inserted between the iron plates of the magnet surrounding ND280. SMRD has three roles in ND280: the measurement of the muon escaping to high angle with respect to the beam direction, the trigger for the cosmic muon event, and the detection of the interaction at the magnet and outside of the magnet.

2.3.2 ND280 Upgrade

Classic ND280 contributed to reduce systematic uncertainties. However, there were some limitations in the measurement due to structural reasons.

The first limitation is the low efficiency for the events with relatively large scattering angles with respect to the beam direction. This is caused by the poor reconstruction efficiency in FGD for the track of particles scattered to large angles. Even if the track is reconstructed, if the particle does not go through TPC, the particle is likely to be misidentified since the reconstruction information from TPC is largely contributed to the particle identification in ND280.

The second limitation is the low efficiency in the reconstruction and identification of the short tracks by low energy hadrons. The difficult point is that the track is short, so the track needs to be reconstructed in the detector, which is also a massive target.

To overcome these limitations, new detectors were installed where there was PØD central part. Upgrade detector consists of three detectors: SuperFGD, High-Angle TPC (HA-TPC), and Time-of-Flight Detector (TOF). The overview of upgraded ND280 is shown in Fig. 2.9. SuperFGD is located in the middle of the upgrade detector, and HA-TPCs are located in the above and below SuperFGD. These are surrounded by six ToF layers.

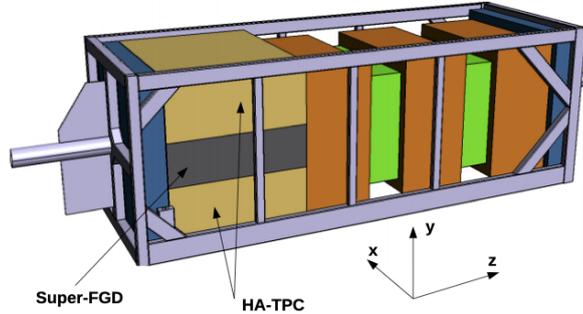


Figure 2.9. Overview of ND280 upgrade. [8]

SuperFGD

SuperFGD is a neutrino target and plastic scintillation tracker. The detail is described in Chapter 3.

HA-TPC

The design of HA-TPC is almost the same with TPC in classic ND280, but some points were updated to reduce the cost without performance degradation and increase active volume. The direction of HA-TPC is horizontal to reconstruct the tracks of particles scattered at high angles while it is vertical for the TPC in classic ND280. The purpose of HA-TPC is same with TPC in classic ND280, that is precise track reconstruction for the charge and momentum reconstruction and particle identification based on the distribution of the amount of ionization along the track.

ToF

ToF is a plate made by arranging plastic scintillator bars. It is designed to measure the precise time when the charged particle passes. It is used to reject backgrounds originating in the areas surrounding the detector, to improve particle identification, and to trigger cosmic muons.

2.4 Far Detector

Super-Kamiokande (SK) is used as the far detector in the T2K experiment. SK is a water Cherenkov detector located 295 km downstream from the neutrino production point and used to detect muon and electron (anti-)neutrinos. The overview of SK is shown in Fig. 2.10. SK consists of two major volumes: an inner detector (ID) and an outer detector (OD). ID has a cylindrical shape with 33.8 m in diameter and 36.2 m in height, and 11129

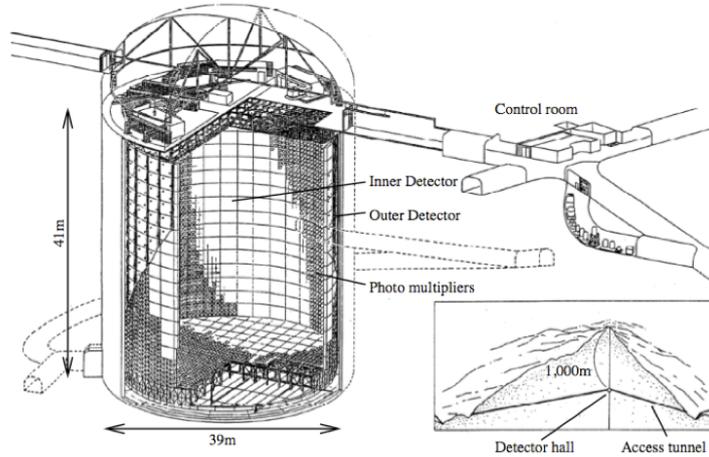


Figure 2.10. Overview of Super-Kamiokande. [3]

inward-facing 50 cm diameter PMTs are arranged on the inner walls. OD surrounds ID with about 2 m thickness, and 1885 outward-facing 20 cm diameter PMTs are arranged on the inner walls. Both were filled with pure water until 2019, and gadolinium was dissolved in the water to improve neutron detection efficiency in 2020. SK detects the ring of light on the wall as an intersection with the cone of Cherenkov radiation by charged particles. ID detects and identifies the leptons produced by the CC interactions of muon and electron neutrinos in water. OD is used to veto cosmic ray muons and other backgrounds.

2.5 Recent results

2.5.1 Measurement of neutrino oscillations

The latest measurement of neutrino oscillations [9] was published in 2023 with the data corresponding to 1.97×10^{21} POT in FHC mode and 1.63×10^{21} POT in RHC mode accumulated by 2020. As shown in Fig. 2.11, $\delta_{CP} = 0, \pi$, which conserve CP-symmetry, are excluded at 90 % confidence level. The statistical uncertainty is currently dominant. However, the significance of systematic uncertainties increases as the statistics increase.

2.5.2 ν_e -CC cross-sections measurement

The systematic uncertainty on the ν_e and $\bar{\nu}_e$ cross-sections is one of the main source of systematic uncertainties in the oscillations measurement. Currently, the value theoretically calculated based on the measured ν_μ ($\bar{\nu}_\mu$) cross-section is used as a ν_e ($\bar{\nu}_e$) cross-section, and it has systematic uncertainty of 3.0 % due to possible cross-section model differences. The direct measurement of ν_e ($\bar{\nu}_e$) cross-section is performed with ND280. The latest result of ν_e ($\bar{\nu}_e$) cross-sections [5] was published in 2020 with the data corresponding to 11.9×10^{20} POT in FHC mode and 6.3×10^{20} POT in RHC mode accumulated by 2017. These data are all collected by classic ND280. In this result, ν_e -CC ($\bar{\nu}_e$ -CC) inclusive cross-sections were measured. “ ν_e -CC inclusive” interaction consists of all type of charged-current interactions (e.g. CCQE, CC-Res) of electron neutrino. The result is

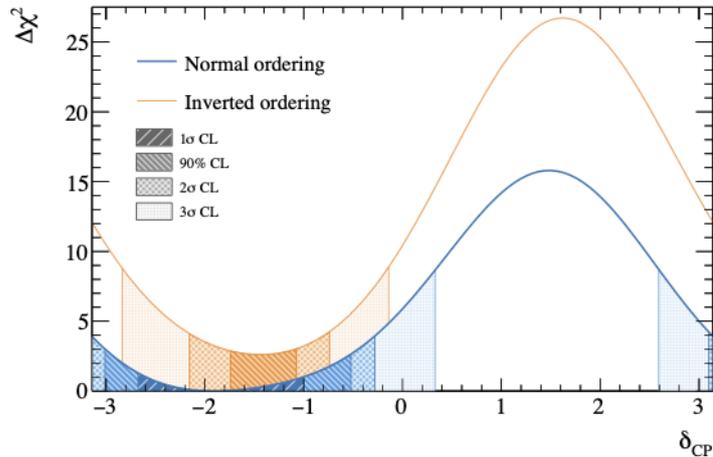
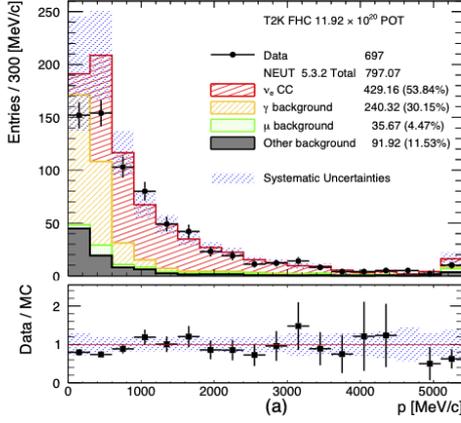


Figure 2.11. The $\Delta\chi^2$ distribution and the confidence intervals in δ_{CP} from fitting to the data. [9]. The result depends on the mass ordering.

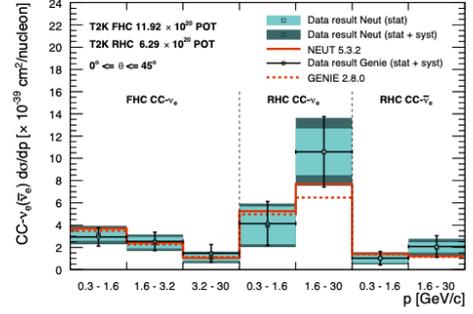
shown in Fig. 2.12 (a). The cross-section was measured for the ν_e in FHC mode and $\nu_e, \bar{\nu}_e$ in RHC mode. The results has about 20 % of statistical uncertainties and systematic uncertainties. This is caused by the limited statistics and the performance of the event selection.

Additionally, the measurement is conducted with the events of limited phase space. In particular, the electron (positron) produced in the interaction required that its momentum is larger than 300 MeV and its scattering angle is less than 45° . However, a part of ν_e events observed at SK, which is used to measure $\nu_\mu \rightarrow \nu_e$ oscillations, produces an electron (positron) being out of the phase space, as shown in Fig. 2.13. Therefore, ν_e cross-section measurements in the low momentum region and large scattering angle region are necessary.

The efficiency and purity of the event selection are about 30 % and 50 %, respectively. The photon is a dominant background source, particularly in the low momentum region. This is critical since the cross-section measurement in the low momentum region is essential, as discussed above.



(a) ν_e -C event selection results with the neutrino mode beam. [5]



(b) ν_e -CC cross-section results. [5]

Figure 2.12. (a) is the results of ν_e -CC events selection with the actual data by neutrino mode beam and the simulation data of the beam. The result with simulation data is scaled based on the value of POT. (b) is the results of the CC-crosssection of the ν_e in neutrino mode beam and the $\nu_e, \bar{\nu}_e$ in anti-neutrino mode beam.

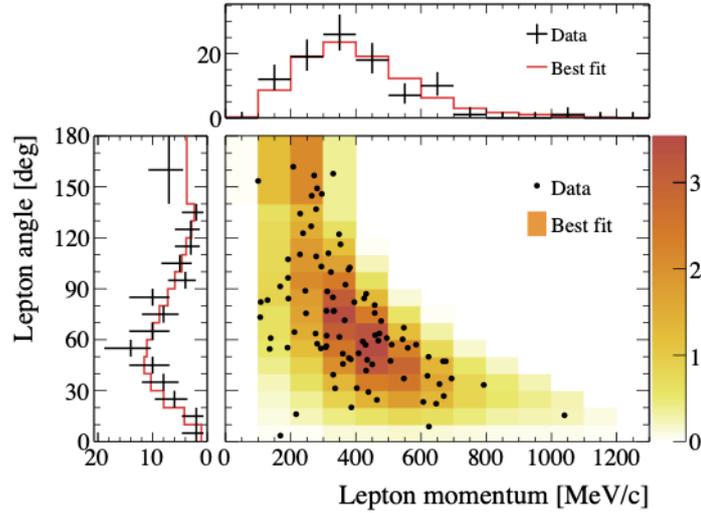


Figure 2.13. The distribution of the reconstructed momentum and scattered angle of the one electron-like ring events observed at SK in neutrino mode beam. Black points are observed events and the drawing is predicted distribution by the model with tuned parameters using real data. The top and left plots are the projection of the histogram. The black histogram and red line are data and predicted distribution, respectively. [9]

2.6 Upgrade and Operation Status of ND280 Upgrade

PØD detector was removed in 2022, and upgrade detectors were installed gradually from 2023 to 2024.

The assembly and commissioning were conducted prior to the installation of each upgrade detector. After the commissioning, the bottom TPC, SuperFGD, and four out of six TOF panels were installed. Only about 80 % of SuperFGD electronics were installed because the full electronics were not ready.

Partially installed upgrade detectors as described above took data of two continuous beam runs from Nov. 2023 to Feb. 2024. The summary of the beam runs are shown in Table 2.1. The operation of SuperFGD in these beam run was unstable. In particular, there were many issues in data acquisition and transfer.

After these beam runs, the maintenance of the detector with the feedback in the beam run and the installation of remaining detectors were conducted. The installation of upgrade detectors was completed in May 2024. First beam data with full upgraded ND280 was acquired in Jun. 2024. The operation of SuperFGD in this beam run was relatively stable. After the maintenance in the summer, upgraded ND280 took more data from Nov. 2024 to Dec. 2024. The beam mode was changed from FHC to RHC in the middle of run, while it was constantly FHC mode in other three runs. The summary of these beam runs are also shown in Table 2.1.

Table 2.1. The beam runs with upgrade detectors

Period	Upgrade	POT in FHC	POT in RHC
Nov. 23rd-Dec. 25th, 2023	partial	2.04×10^{20}	0
Jan. 16th-Feb. 23rd, 2024	partial	1.07×10^{20}	0
Jun. 6th-Jun. 28th, 2024	full	2.08×10^{20}	0
Nov. 25th-Dec. 23rd, 2024	full	1.04×10^{20}	1.23×10^{20}

Chapter 3

SuperFGD

3.1 Overview

SuperFGD is one of the ND280 upgrade detectors, which is a plastic scintillation tracker and also a neutrino target. Conceptual sketch of SuperFGD is shown in Fig. 3.1. It consists of $1 \times 1 \times 1 \text{ cm}^3$ plastic-scintillator cubes, wavelength-shifting (WLS) optical fibers, and silicon multi-pixel photon counters (MPPCs).

When a charged particle passes through the scintillator cube, it produces scintillation light. The number of photons of scintillation light is in proportion to the energy deposit of the particle in the cube, which depends on the type and energy of the particle and path length of the track in the cube. The light produced in the cube is propagated in three orthogonal directions by three WLS fibers which go through the cube. Finally, the propagated light is read out by an MPPC.

SuperFGD has 192, 56, and 182 cubes in x , y , and z directions, respectively, where the y direction is the vertical direction, the z direction is the direction derived by projecting the beam direction onto the horizontal plane, and the x direction is the horizontal direction, which is perpendicular to the z direction. Thus, SuperFGD has $192 \times 56 \times 182 = 1,956,864$ cubes, and $(192 \times 56) + (192 \times 182) + (56 \times 182) = 55888$ fibers and MPPC

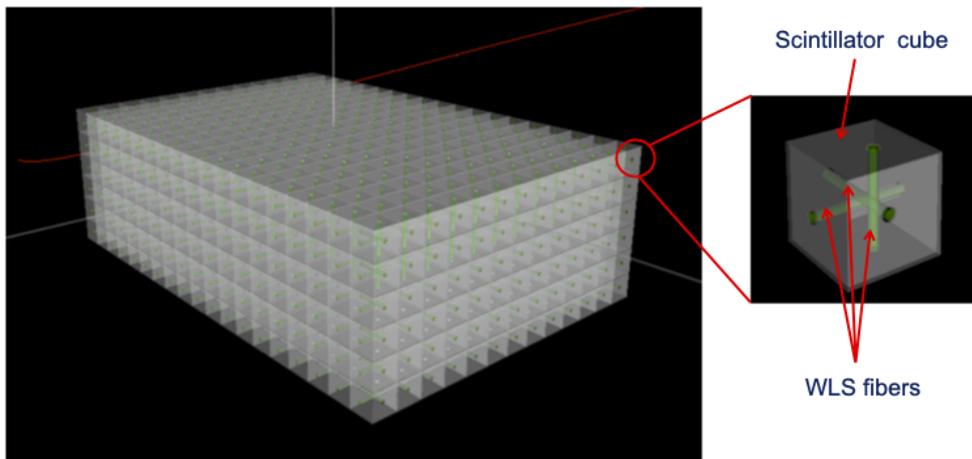


Figure 3.1. Conceptual sketch of SuperFGD. [8]

channels.

3.2 Scintillators, WLS fibers, and MPPCs

Scintillator cubes are made of polystyrene doped with 1.5 % of paraterphenyl and 0.01 % of POPOP. The surface of the cube was etched with a chemical agent to form a reflecting layer covering the cube. Three orthogonal holes of 1.5 mm diameter are drilled in the cubes to pass through the fibers.

As a WLS fiber, Y-11 (200) produced by KURARAY Co., Ltd. [10], which is also used in the other scintillator detectors in ND280, is used. The diameter of the fiber is 1 mm. The absorption wavelength has a peak around 430 nm, and it is in good agreement with the wavelength of the scintillation light by the scintillator cubes. The emission wavelength of the fiber has a peak around 476 nm, and the light with the wavelength repeats the total internal reflection in the fiber and travels to MPPC .

MPPCs are mounted onto 881 printed circuit boards (PCBs). Each PCB accommodates 64 MPPCs and MPPCs are arranged on a grid of 8×8 . The number of MPPCs is larger than the number of fibers, and some of MPPCs are not used.

3.3 Electronics

SuperFGD uses the Cherenkov Imaging Telescope Integrated Readout Chip (CITIROC [11]) ASIC to process signals from MPPCs. The block scheme of CITIROC is shown in Fig. 3.2 (a). Each CITIROC chip has 32 channels. The signal is split and amplified with two different gains: high gain (HG) and low gain (LG), to handle a wide range of signal. The analog signal is then sent to an external ADC. In addition, a fast (15 ns) shaper and a constant-threshold discriminator generate timing signals corresponding to the rising and falling edges of the pulse, allowing for the measurement of timing and time over threshold information. Eight CITIROC chips are mounted on a frontend board (FEB), which controls and reads out signals from 256 MPPCs.

The overview of the readout electronics is shown in Fig. 3.2 (b). Fourteen FEBs and one Optical Concentrator Board (OCB) are housed in a crate. The OCB manages the communication of clock and slow control signals between the FEBs and the data acquisition (DAQ) system. It also performs local event building for the FEBs within the crate. There are 16 crates in total, with eight located on each side of the detector. Synchronization within the SuperFGD electronics and with other detector systems is managed by a Master Clock Board (MCB).

Usually, the data is sent from OCB to DAQ-PC event by event. However, in some cases, such as LED data taking for the calibration, a high event rate causes a data transfer issue. Since calibration is performed channel by channel, there is no need to consider correlations between hits within the same event or to record timing information. Another method for sending data from OCB to DAQ-PC was developed using this characteristic. In the method, a histogram is prepared in the OCB for each channel that belongs to the OCB. For each event, only the value of HG-ADC at each channel is filled in the histogram, and other information is discarded. (The value of LG-ADC is not used for the

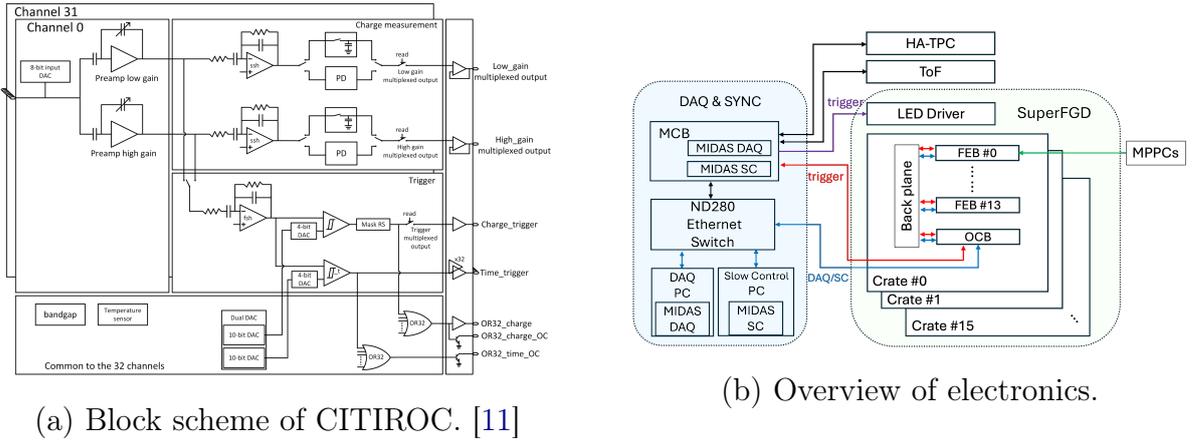


Figure 3.2. (a) is a block scheme of CITIROC located in FEBs. (b) is an overview of readout electronics of SuperFGD.

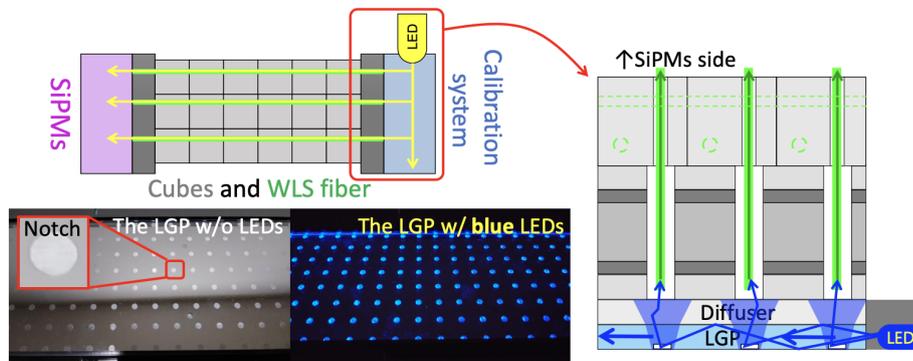
calibration.) The data is not sent to DAQ-PC event by event, and only the histograms are sent periodically. This method is called the “histogramming” system.

3.4 Calibration System

The calibration of MPPCs is conducted using LED light. The LED calibration system [12] is also used for the channel checks in the commissioning and monitoring of the stability of MPPC in regular operation.

The LED calibration system needs to distribute the light by LED to every channel. However, due to spatial limitation and experimental requirements, it had to achieve in the confined space with minimal material. Then, Light Guide Plates (LGP) were used to distribute light instead of arranging LED channel by channel. The picture and conceptional sketch is shown in Fig. 3.3.

LGP modules are installed on the surface of SuperFGD, and LGP has notches at the position corresponding to the endpoint of fibers. The light from LEDs enter to the LGP from the side and travel in the LGP. While the light travels through the LGP, it scatters



by a notch and reaches a WLS fiber. The light is propagated by the fiber and reach an MPPC. In this way, the LED flash for the calibration is delivered to MPPCs and an MPPCs are calibrated with the signals for the light.

There are two sizes of LGP modules. One is the module with the notches on a grid of 56×8 . 47 modules of this size are installed on the surface on the xy -side and yz -side. The other is the module with the notches on a grid of 96×8 . 46 modules of this size are installed on the surface on the xz -side. There are $47 + 46 = 93$ LGP modules in total.

For the LED data taking, MCB send LED trigger signals to OCBs, and also send signals to the another electronics board called “LED driver”. LED driver control LEDs by generating the signal. When LED driver receive the signal from MCB, the FPGA in the LED driver makes digital signal. The signal is amplified by an operational amplifier (also in LED driver) and sent to LEDs to flash them. The pulse height and width are programmable by changing the configuration of FPGA.

3.5 Commissioning, Installation, and Operation

The timeline of SuperFGD is summarized in Table 3.1.

The SuperFGD was assembled from October 2022 to April 2023. After the assembly, the commissioning was conducted before the installation on ND280. Since access to the detector is limited after installation, every MPPC channel was checked with LED and cosmic rays in the commissioning. During this commissioning period, only a part of the electronics was delivered at that time, so the entire DAQ system could not be tested. Therefore, the check of all MPPC channels with limited electronics was conducted by reconnecting the cables between the MPPC-PCB and the FEB.

After the commissioning, SuperFGD was installed to ND280 on October 12th, 2023. Since the electronics was still incomplete at that time, the electronics of 10 crates out of 16 crates were installed, and the first operation with the beam was conducted with those electronics. SuperFGD had taken data of two continuous beam time from November 2023 to February 2024.

Full electronics was finally installed in March 2024. The first data acquisition of SuperFGD with full electronics started in June 2024, and beam data was taken for around two months by the end of 2024.

Table 3.1. The timeline of the SuperFGD installation and operation

Period	
Oct. 2022-Apr. 2023	Assembly
Apr.-Oct. 2023	Commissioning on the ground
Oct. 12th, 2023	Installation
Nov. 23rd-Dec. 25th, 2023	Beam time
Jan. 16th-Feb. 23rd, 2024	Beam time
Mar. 2024	Electronics Full installation
Jun. 6th-Jun. 28th, 2024	Beam time
Nov. 25th-Dec. 23rd, 2024	Beam time

3.6 Motivation of this thesis

This thesis presents research on the SuperFGD detector. In the experiment, ensuring the acquisition of high-quality data is the first critical step. To this end, the development of an online gain monitor is described in Chapter 4. Then, event reconstruction plays a key role in data analysis. Chapter 5 evaluates the impact of bad channels on reconstruction performance. Finally, Chapter 6 explores the use of an artificial neural network to reject photon background in preparation for measurement of electron neutrino interactions.

Chapter 4

Online Monitoring

4.1 Online Monitoring System for ND280

During the detector operation, it is important to detect and fix anomalies as soon as possible to keep taking good-quality data as long as possible. Even if it is difficult to fix immediately, it is essential to know when the issue occurred precisely. The status of the hardware, including temperature and voltage, is monitored by the slow control system. However, some anomalies can be detected only after analyzing the acquired data.

In order to monitor the detector performance in real time, we have developed the ND280 Online Monitor system. It analyzes the acquired data immediately and summarizes information about the detector performance. The detector status is then monitored by the experimental shifters.

Online monitor receives the data from the global DAQ system and makes plots to be checked by shift persons. A period during which the data acquisition continues is called DAQ run. Typically, DAQ run continues for from a few hours to a few days. The plots are reset at the beginning of each run and every two hours, and they are updated regularly using the data received since the last reset. The update times can be set for each plot. While data accumulation is necessary for the right decision, the online monitor needs to detect anomalies as soon as possible, so it is requested that the update time for each plot be set to no longer than 10 minutes.

4.2 SuperFGD Online Monitor

The plots for the online monitor should be designed to detect potential anomalies in the detector. Moreover, they should be prepared in a way that allows the shifter to identify anomalies easily with minimal instructions, even if they are not experts.

For SuperFGD Online Monitor, the following possible anomalies are considered:

- The anomalies in the electronics.
- The anomalies in timing synchronization.
- The anomalies in the gain of MPPC.

First, the anomalies in the electronics can be caused by either the issues in DAQ system or hardware issue including cable connections. When this anomaly occurs, the signal from a specific channel or channels that belong to the specific DAQ unit can be lost. It can also happen that the derived data from the specific channel is broken and the value of hit information is unnatural. Second, when some issue was occurred in timing synchronization, timing information will be affected. For example, the recorded hit time in the same event can be widely different depending on the clock board to which the channel belongs. Finally, the anomalies in the gain of MPPC are considered to be caused by hardware or environmental issues in most cases. However, some analysis is needed to calculate the gain value of each channel, so the gain value cannot be directly monitored by the slow control system. Then, the online monitor is used to monitor gain values. To detect these anomalies, the following plots are prepared.

4.2.1 Hit Summary

The ADC value of HG, LG, and timing information at each channel is recorded when the value is above the channel's threshold and the event trigger signal (e.g. beam, cosmic, and LED trigger) is issued. The number of hits and the sum of the acquired value during the run at each channel are plotted on the channel 2D map for each item. The map of the number of hits is shown in Fig. 4.1 (a).

These plots are mainly for the detection of anomalies in the electronics. These plots can be used to find dead and noisy channels by monitoring channels that have no (few) events or many more events than other channels. A sum of the item's value over the events during the run at each channel has no physical meaning. However, strange behavior at specific channels and breakage or loss of binary data can be found by comparing them with other channels.

There is a histogram for each crate, which records the number of hits against the channel index in the crate. This information can be used to check the details.

4.2.2 ADC Spectra

The ADC spectra of HG and LG recorded at the channel level are also prepared. There are so many channels that it is not the histogram for each channel but 2D histograms summarizing the spectra of the channels belonging to each crate, as shown in Fig. 4.1 (b). The x and y axes of the histogram represent the channel index in the crate and ADC bins, respectively, and the 2D bins are colored according to the number of events in each bin.

These plots are also for the detection of anomalies in the electronics. These spectra plots can be used to check the problematic channels further. It is possible to estimate the cause of the issue to some degree from the characteristics of the spectrum. For example, if there are many events with a specific amplitude value, it is likely an issue in the DAQ system. On the other hand, if the histograms totally shift to the larger value, it is likely due to a large gain and an issue in the hardware like MPPC or the environment, including temperature.

4.2.3 Event Summary

The distribution of event data size is recorded for each trigger type as shown in Fig. 4.1 (c). The distribution of the number of hits in each event is also recorded.

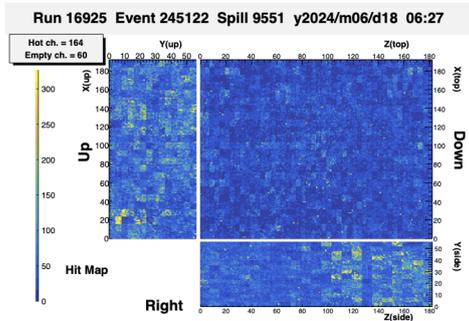
These plots are also for the detection of anomalies in the electronics. However, in contrast to the above two types of plots, these plots are for the monitoring of the events overall but not the monitoring of each channel. These plots has a sensitivity for the probabilistic anomalies.

4.2.4 Hit Timing Summary

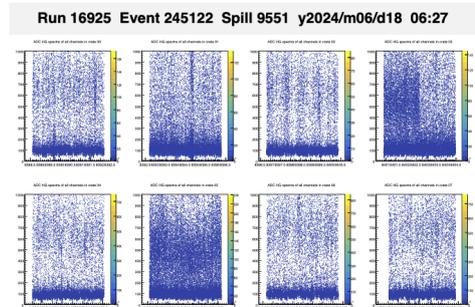
The hit time distribution throughout events is recorded as shown in Fig. 4.1 (d). Hit time is defined by the elapsed time since the external trigger is received until the signal at the channel exceeds the threshold. Each beam spill has eight bunches, and a trigger is sent for each spill. Therefore, the hit time distribution is expected to have eight equidistant peaks. This plot is to monitor the anomalies in the timing synchronization.

4.2.5 Gain monitor

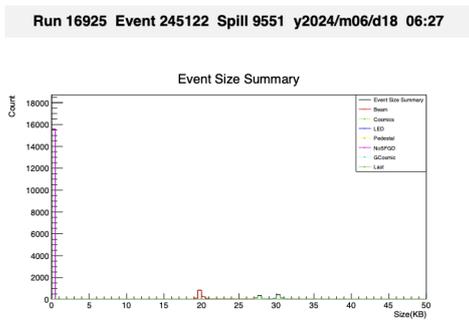
The gain of MPPC can be estimated from the interval between the photoelectron peaks of signal. The gain value varies per channel, and it is a good indicator to monitor the channel status. The map of gain value for the HG channels is monitored as shown in Fig. 4.1 (e). The plot of averaged gain value in each ASIC is also recorded. For calculating the gain during beam time, the LED calibration system injects light between beam spills. The gain of each channel is calculated from these data, and anomalies are detected from the change of the value. The following sections in this chapter describe the study for the development of HG monitor.



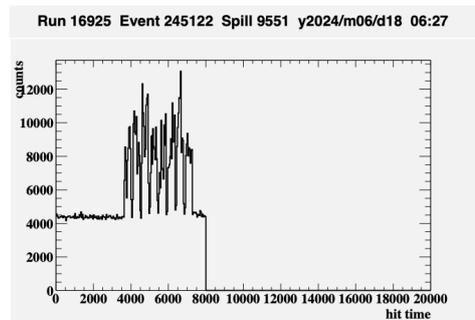
(a) Hit Summary.



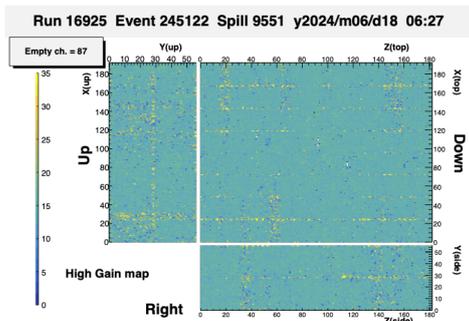
(b) Spectra.



(c) Event Summary.



(d) Hit time distribution.



(e) Gain monitor

Figure 4.1. Examples of the plots for SuperFGD Online Monitor. (a) is a map of the number of hits at each channel. (b) is a spectra of ADC values with HG at each of channel in crate 0 – 7. The histogram for each crate is lined up. (c) is a distribution of event data size with each trigger type. (d) is the hit time distribution throughout events (e) is a map of Gain value of each channel.

4.3 Requirements for the online gain monitor

4.3.1 LED data for the gain monitoring

LED data for the gain monitoring is recorded by using histogramming system (see Sec. 3.3). LED trigger rate for SuperFGD is ~ 1 Hz, and the data is taken at every channel. The histograms are sent sequentially at a rate of about 82 histograms per second, then a histogram of each channel is sent every ~ 680 seconds. Since only the events which cross the threshold is recorded at each channel, the histogram typically contains around 500 events (out of ~ 680 triggered events) in total. The received histograms of the same channel are accumulated on the monitoring system until monitoring plots are reset (every two hours).

4.3.2 General method for gain calculation

LED intensity for gain calibration is adjusted so that several photons reach an MPPC. For most channels, around five photons are delivered on average. The ADC value distribution with LED trigger has several discrete peaks corresponding to different number of detected photons, as shown in Fig. 4.2. Typical ADC value by a certain number of photons, which is the center of the peak is considered linear to the number of photons. The gain value is calculated by the slope coefficient of the typical ADC value against the number of photons.

The threshold for recording events at each channel is set to the ADC value corresponding to 3.5 photons based on the gain and pedestal obtained by calibration. Therefore, the derived ADC distribution should have no peak corresponding to less than four photons, and the four-photon peak is considered to be distorted. Therefore, the gain is calculated with the peaks corresponding to more than four photons.

4.3.3 Gain calculation for calibration and time constraint for the monitor

The method for gain calibration (hereinafter, this is called “offline method”) is performed with:

1. Estimating the center of peaks by fitting the histogram by multiple Gaussian distributions.
2. Fitting the derived center position of the peak against the number of photons by a linear function. Gain is derived as a fitted value of the slope.

Since this offline method provides a precise gain value, it is always used when the time restriction is loose enough. However, it takes ~ 50 minutes to calculate the gain for all channels due to the time-consuming fitting process.

The time constraint in the gain calculation for the online monitor depends on the timing when to calculate the gain value. There are two choices for that: to calculate the gain for all channels when the online monitor plots are updated, and to calculate the gain for each channel when receiving histograms of the channel. Under the current situation, the latter choice provides a looser time restriction. In the latter case, the requirement is

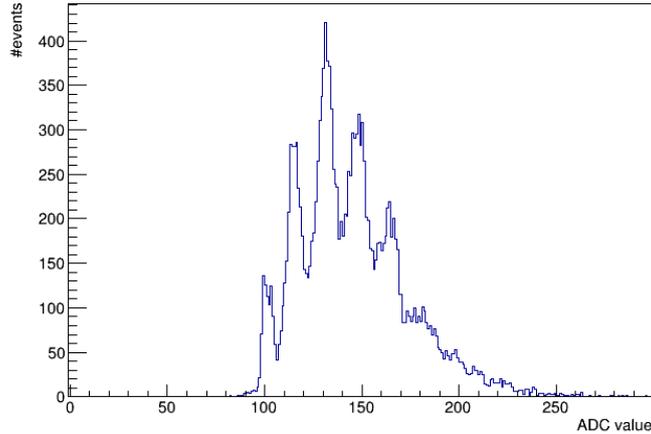


Figure 4.2. Typical distribution of ADC value with LED trigger. There are peaks by the events of a certain number of photons. The peaks are more evident than in the histogram for the online monitor because this is a histogram for calibration.

that the gain need to be calculated in the time shorter than the histogram receiving cycle. In other words, the gain calculation for all channels needs to be processed in 680 seconds in total, which is the time it takes for the histogramming system to send the histograms of all channels. The offline method does not meet this condition.

Online monitor is used to detect anomaly channels, so the accurate gain value is not necessary. As one of the criteria to detect anomaly channels, in the (offline) data quality check, the channels are considered anomaly when the gain value is out of 3σ of the gain distribution. The typical mean and standard deviation of the distribution is around 15 and 1.5. Therefore, a deviation of about 30 % from the mean value is allowed to the gain value in the anomaly channel detection. Then, it is important to calculate gain within less than 30 % relative error from true value. When using the calibration value as a true value, typically about 96 % of channels achieve less than 30 % relative error in the data quality check. The accuracy in the gain monitor should not get worse so much than that in the data quality check. Thus, the requirement is set to calculate gain within less than 30 % relative error from calibration value for more than 95 % of channels in good runs, which are judged based on the data quality check.

In conclusion, the requirements for the gain calculation method for the online monitor are below.

- Calculate the gain for all channels in 680 seconds.
- Calculate the gain for more than 95 % of channels within less than 30 % relative error from calibration value (in the good run).

4.4 Gain calculation for Online Monitor

4.4.1 Ideas for the method

Peak positions in the histograms can be determined without fitting by identifying the local maximum bins. However, in the online monitor histograms, each bin typically contains at most 50 events, leading to significant statistical fluctuations. This could result in many local maxima that do not correspond to true peak positions. Thus, additional techniques to reduce statistical uncertainties and eliminate false peaks are necessary.

4.4.2 Procedure of online gain calculation

The new method (hereinafter, this is called “online method”) for gain calculation consists of seven steps.

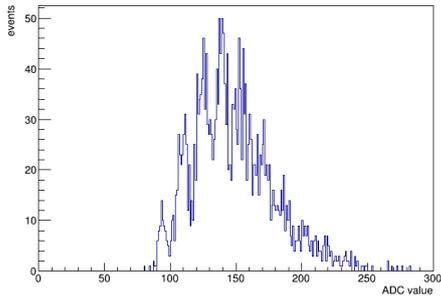
1. Take the moving average of the bin contents.
2. Set the bin range for the calculation.
3. Select local maximum bins.
4. Reject suspicious peaks.
5. Calculate the precise position for each peak.
6. Extract peaks that are approximately equidistant.
7. Calculate the gain value.

The example illustration of the method is shown in Fig. 4.3. Fig. 4.3 (a) is an example of original histogram used to calculate the gain, and (b)-(h) are corresponding to each step.

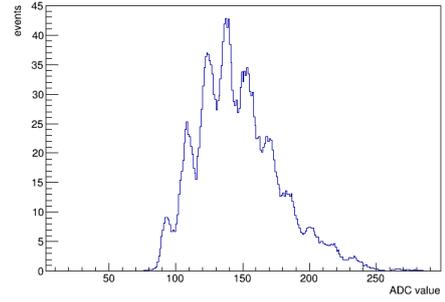
(b) First step is to take the moving average of the bin’s contents. This is the step to reduce the effect of statistical fluctuations. The average is taken in seven consecutive bins. In other words, the contents of bin i of the histogram before the process was x_i , and the contents of bin i change to $\frac{1}{7} \sum_{j=-3}^3 x_{i+j}$ in this step.

(c) Next, the range of bins for the gain calculation is set to exclude the peaks by noise apart from the typical signal size. This is the step to reduce the time for calculation and to prevent the results from being affected by accidental entries. The leftmost and rightmost bins whose heights are more than 1/100 of the highest bin are picked, and only the bins between these bins are used in the following steps.

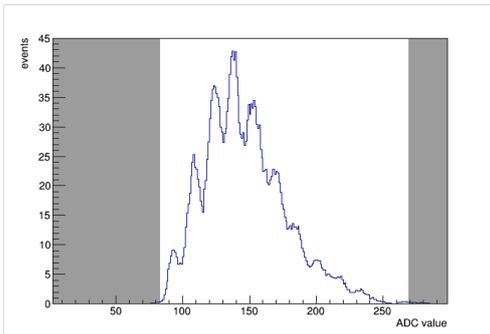
(d) After these preparation steps, we pick up all of the local maximum bins. The local maximum bin is defined as the bin whose height equals or exceeds both the previous and next bin. If the local maximum bin is consecutive, which can occur when the bins of the same height are in a row, we only pick the middle bin of the consecutive local maximum bin. If the number of consecutive local maximum bins is even, we pick the left bin of the two middle bins. The peak position is adjusted in the following step, so this choice has no effect when the number of consecutive local maximum bins is less than four. If four or



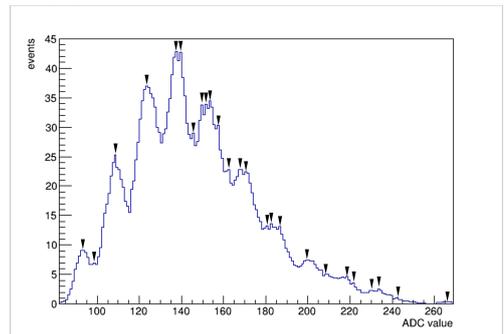
(a) ADC distribution.



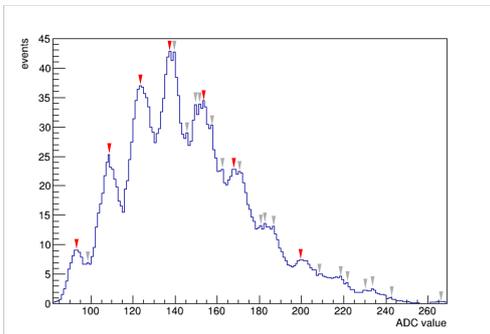
(b) After take moving average.



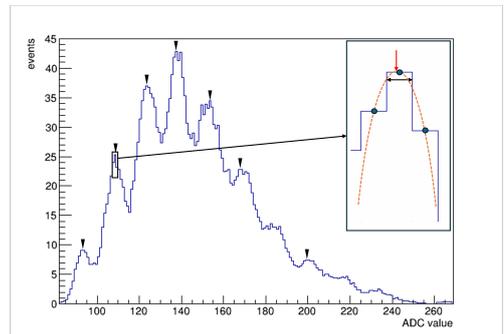
(c) Set range for calculation.



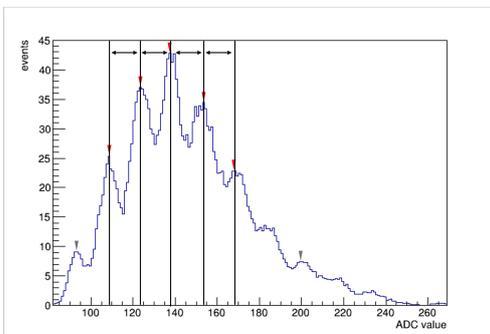
(d) Select local maximum bins.



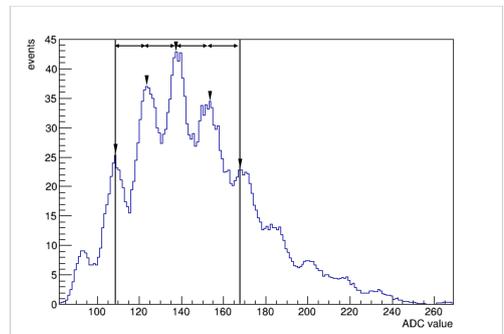
(e) Reject suspicious peak.



(f) Calculate the precise position for each peak. Orange dotted line is parabola and the red arrow is the center of the parabola.



(g) Extract approximately equidistant peaks.



(h) Calculate gain value.

Figure 4.3. The online method for gain calculation

more than the same height bins are consecutive, the following adjustment does not work well, but it rarely happens, and also almost all of such a peak was rejected in step 4 or 6 and not used in the final gain calculation.

(e) The selected local maximum bins include fake peaks due to statistical fluctuation although the number is decreased by taking the moving average. Then, significant local minimum bin is required between neighboring local maximum bins to reject fake peaks. The criterion of significant local minimum bin is set to a value that can reject the hypothesis that there is no bottom between the local maximum bins with a confidence level of 1σ . As a formula, the upper limit for significant local minimum bin is described as $\min(h_1, h_2) - 3\sqrt{\min(h_1, h_2)}/7$ where h_1 and h_2 are the heights of the local maximum bins. The effect of taking moving averages is also considered in the above criterion. The two neighboring local maximum bins without significant local minimum bin in between are considered to belong to the same peak, and only the higher bin remains as the center of the peak.

(f) Next, the precise position of the center of the peak is calculated from the selected bin and both sides of the bin. For the calculation, the parabola passing through the three points, which are the center of the selected bin and the center of the immediate left and immediate right of the selected bin, are used. This parabola is uniquely determined from the given three points. The center position of the parabola is used as the precise peak position.

Specifically, the precise position is calculated as $c + \{(h_r - h_l)/2(2h_c - h_l - h_r)\}$, where c, h_c is the index and the height of the selected local maximum bin, and h_l, h_r are the heights of the bins which are immediate left and immediate right of the selected bin.

(g)(h) If the number of the remaining peaks is less than two, the calculation is failed because the gain calculation needs at least two peaks. If the number of remaining peaks equals two, gain is calculated as the distance between the peaks. If the number of peaks is equal to three, the first peak is ignored due to the threshold effect as described above, and the gain is calculated as the distance between the second and third peaks.

When the number of the peaks is more than three, the first peak is ignored. We define the range of peaks used for the gain calculation, l -th peak to r -th peak, in the following way.

First, The leftmost peak is set to $l = 2$. Let the number of the peaks be N_{peak} , and repeat the following process for $i = 3, 4, \dots, N_{peak} - 1$.

- If the distance between i -th peak and $(i + 1)$ -th peak is equal to the value which is obtained by dividing the distance between l -th peak and i -th peak by $(i - l)$ within 10 % relative difference, $(i + 1)$ -th peak is decided to be used.
- Otherwise, l -th peak and $(i + 1)$ -th peak are considered incompatible. The leftmost peak for calculation is only updated when the change increases the number of usable peaks. Specifically, for every $j = i + 1, i + 2, \dots, i + (i - l)$, if the distance between the j -th peak and the $(j + 1)$ -th peak matches the value obtained by dividing the distance between the i -th peak and the j -th peak by $(j - i)$, within a 10 % relative difference, l is updated to i , and i is skipped to $i + (i - l) + 1$. This process excludes peaks up to and including the $(i - 1)$ -th peak from calculation. If the condition is

Table 4.1. The number of success/fail channels in gain calculation

		offline method		
		Success	Fail	Total
online method	Success	42148	24	42172
	Fail	47	33	80
	Total	42195	57	42252

not met, or if the number of peaks after the i -th peak is fewer than $(i - l) + 1$ (i.e., $i + (i - l) + 1 > N_{\text{peak}}$), the rightmost peak used for calculation is set to $r = i$.

Finally, an average of the distance between l th peak and r th peak is calculated as a gain value.

The time complexity of each step of this method is at most $O(N_{\text{bins}})$. The first and second steps ((b) and (c) in Fig. 4.3) dominate the calculation time, but this is very simple. The latter steps are conducted for each local maximum bin or peak, so the number is much decreased from the number of bins. Also, the calculation in those steps is not so much complicated, so this takes much less time than the first two steps.

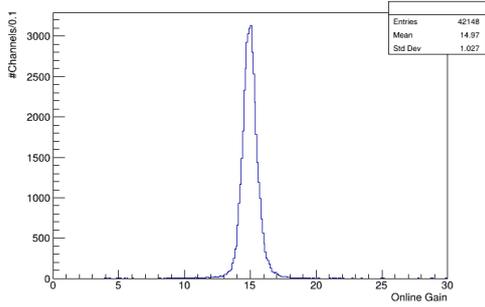
4.4.3 Validation by the comparison with offline method

The method was validated by comparing the results against the results of the offline method. There are two purposes for the validation. One is to check the online method works well as a gain calculation method. Since the histogram for the online gain monitor has small statistics, the result with the data is affected by the small statistics. Therefore, the validity of the online method as a gain calculation method should be checked with large statistics data. The other is to estimate the time required for the computation. The computing time is expected to depend on the number of bins but not the statistics. Therefore, the computing time is expected to be almost the same for the histogram in the validation and the histogram for the online gain monitor. Thus, the computing time of the online method in the online gain monitor is estimated from the elapsed time in the validation.

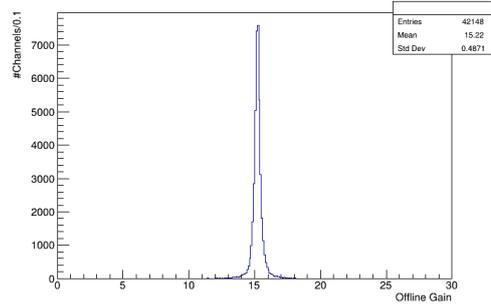
The data acquired for gain calibration was used for validation. The histogram for the calibration has about $\sim 20,000$ entries, while the histogram for online monitor has $500 - 5,000$ entries depending on the time since the last reset. The calibration run data taken in Feb. 2024 was used. All of SuperFGD's electronics were not installed at that time, and the calibration data was taken for 42252 channels, which were prepared for data acquisition in February.

First, the number of channels that could calculate the gain with each method is summarized in table 4.1. The online method has succeeded in calculating 42172 channels out of 42252 channels. The success channel accounts for more than 99.8% of the tried channels.

The distribution of the gain value calculated by the online method is shown in Fig. 4.4 (a). The distribution of the gain value calculated by the offline method is shown in Fig. 4.4 (b) as a reference. The mean value is slightly small but almost consistent. The standard deviation in the online method is around twice larger than that in the offline method.

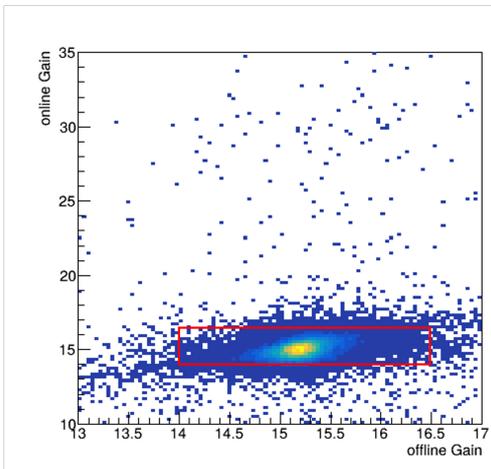


(a) online method

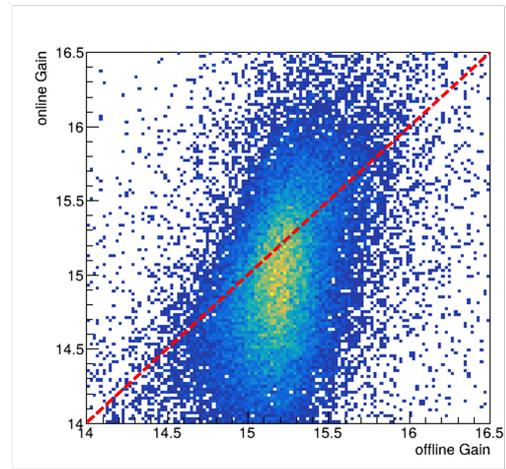


(b) offline method

Figure 4.4. The distribution of gain value calculated with each method.



(a) overall



(b) enlarged

Figure 4.5. The comparison of gain value with methods. (b) is a plot enlarged in the rectangle region surrounded by a red line in (a). The red dot line in (b) represents the line that satisfies the condition that online gain equals offline gain.

The correlation of gain value between the online method and offline method is shown in Fig. 4.5. Typical miscalculation of gain with the online method is caused by missing a peak between two peaks. Then, the calculated gain value becomes double. There is a small part of such channels as shown in the area where gain with online method is between 25 and 35 in Fig. 4.5 (a). However, the fraction of such channels is less than 0.2 %. It can be confirmed by that the peak around two in Fig. 4.6 is hardly seen.

The estimation of the peak position in the online method is not as precise as the offline method, which uses multiple Gaussian fittings. However, the calculated gains of 98.6% of the all 42252 channels are within less than 30 % relative error from the calibration value, and also that of 95.5% of the channels are within less than 10 % relative error from the calibration value. Therefore, the online method achieves the accuracy requirement if the statistics are large enough. Thus, the online method is considered to work well as a gain calculation method.

The computing time is less than 10 seconds in total for 42252 histograms. Thus, the

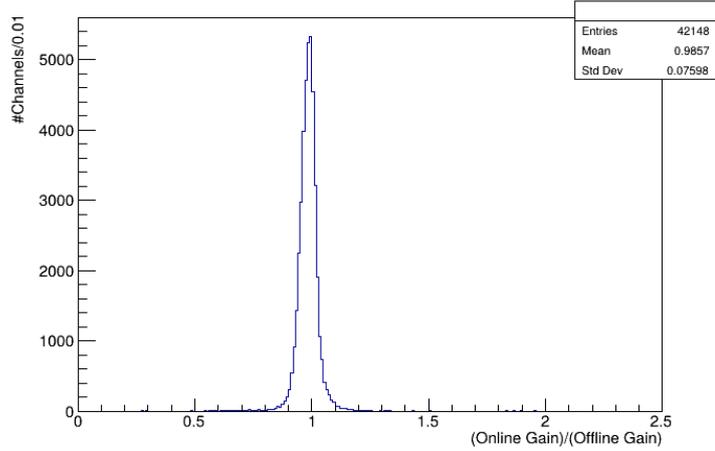


Figure 4.6. The ratio of gain with online method to the value with offline method.

online method is expected to calculate the gain for all channels in less than 15 seconds. It is enough faster than the requirement (680 seconds).

In conclusion, the online method is considered to be valid as a gain calculation method and to meet the time requirement.

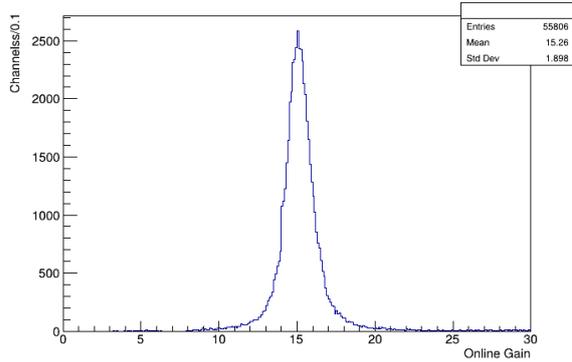
4.4.4 Test with histogramming data

The online method was implemented for the online monitoring system and checked the performance with histogramming data acquired in Jun. 2024. The used run is guaranteed to be a good run by data quality check. DAQ run is divided into “subrun”s. Subrun switches every time 2 GB of data is acquired, and it is about 45 minutes for upgraded ND280. One subrun data was used for the test.

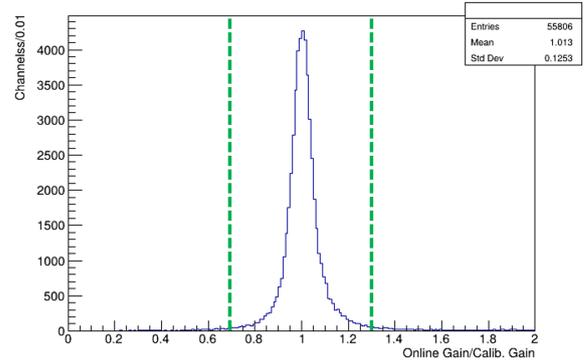
The check of requirements

Since the histograms sent in 45 minutes are included in one subrun data, the upper limit for the calculation time of gain with one subrun data is also 45 minutes. In one subrun, histograms are sent, and gain is recalculated around $(45 \times 60)/680 \sim 4$ times for each channel. The calculation using the online method with one subrun data was executed in 400 seconds. This corresponds to the speed of calculating gain for all channels in 100 seconds. Therefore, the online method meets the time requirement. This is longer than expected time in the validation, but it is considered due to the processing time for other monitoring plots.

The online method successfully calculated gain for 55825 channels out of 55888 channels with histogramming data. The distribution of the calculated gain value and the ratio to the value in calibration is shown in Fig. 4.7 (a) and (b), respectively. Online method calculate the gain for 95.4 % of channels within less than 30 % relative error from calibration value. Therefore, the online method also meets the accuracy requirement. Thus, the online method meet both requirements.



(a) Gain distribution



(b) Rate distribution

Figure 4.7. (a) is a distribution of gain values with the online method and histogramming data. (b) is a distribution of the ratio of the gain value with the online method to the gain value in the calibration. The green dot line represents 0.7 and 1.3, which are the lower and upper limits of 30 % relative error.

Gain monitoring plots

These calculated gain of each channel are displayed for monitoring by mapping on the channel position map (Fig. 4.8) and the grid with DAQ board and channel index (Fig. 4.9 (a)). The structural issues can be found in the gain position map. For example, in Fig. 4.8, there are some vertical or horizontal lines and some rectangle regions of the channels whose gain value is higher than other channels. It was understood that the line of high gain channels is caused by the light leakage from the unused screw hole for fixing LED guide panel and that gain in the rectangle region was not correctly calculated due to the too-high LED intensity of the LED guide panel for the area.

Gain plot on the grid with DAQ board and channel index can be used for the frontend board level and ASIC level performance check. For example, there is a segment colored orange at the region surrounded by a purple circle in Fig. 4.9 (a). This means the gain value of the consecutive channels from channel 0 to channel 63 in a specific frontend board is higher than others. The channels that belong to the same ASIC have consecutive channel indexes, and each ASIC has 32 channels. Therefore, this region corresponds to the channels belonging to either of two specific ASICs. Thus, the issue of those ASICs can be suspected from this plot. Also, there are some channels that failed in gain calculation in the region surrounded by the red circle in Fig. 4.9 (a). MPPC is surely allocated to those channels as shown in Fig. 4.9 (b), the ASICs including these channels are considered dead-suspicious ASICs.

The plot of gain averaged in each ASIC is shown in Fig. 4.9 (c). This plot is sorted based on the crate and ASIC index. This plot is easy to check due to the reduction in the number of bins. On the other hand, the noticeability of the anomaly of a single channel is worse than the channel-level gain plot due to being averaged. The features mentioned in the channel-level gain plot can also be seen in this plot. The purple and red circle indicates the same problematic ASICs as in Fig. 4.9 (a).

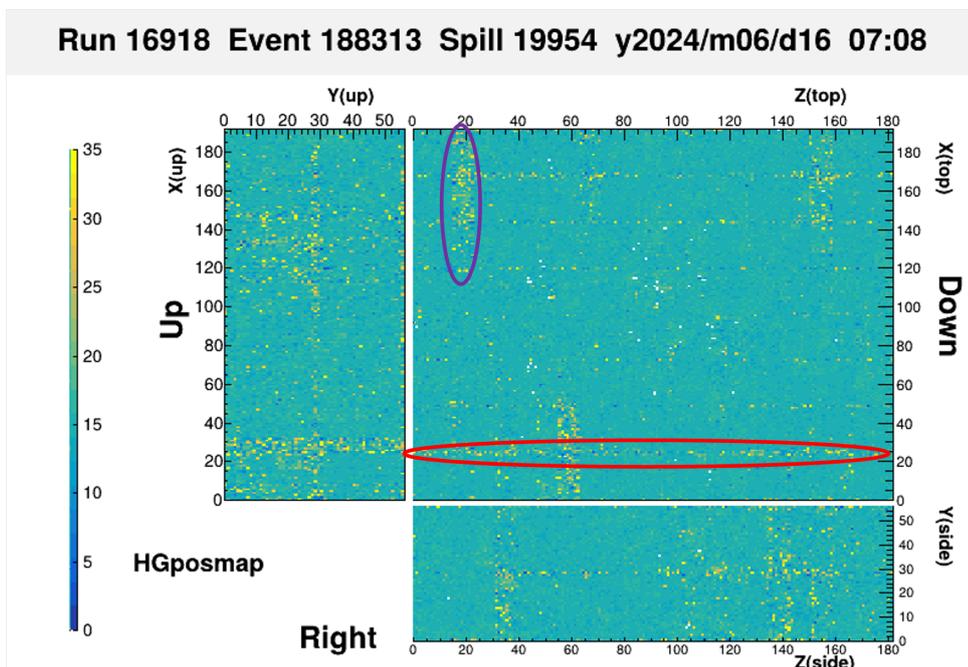
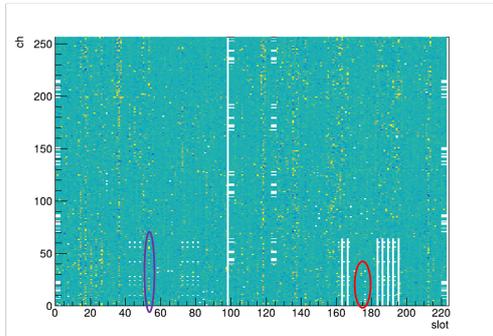
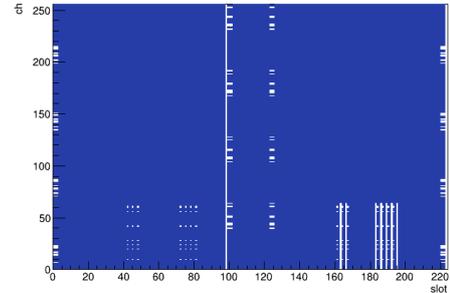


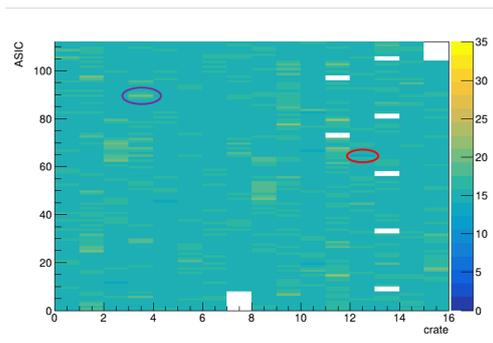
Figure 4.8. The plot which filled gain values onto a 2D map of channel positions. The red and purple circle indicates the area of the suspected issue. It was found that the red region is due to the light leakage from the screw hole and that the purple region is due to too-high LED intensity at the guide panel for calibration.



(a) Gain distribution plot with the DAQ board and channel index



(b) Summary of the channels allocated on the DAQ board



(c) The distribution plot of averaged gain in each ASIC against the crate and ASIC index

Figure 4.9. (a) A gain distribution mapped on the grid with a frontend board index and channel index. The red and purple circles indicate the area of dead and noisy suspicious ASICs, respectively. Each ASIC has 32 channels, and the indexes are consecutive. In particular, the channel from index 0 to index 31 belong to same ASIC in each frontend board. (b) The channel map on the DAQ board. For each pair of frontend board index and channel index, there is an allocated MPPC if the cell in the histogram is blue and no allocated MPPC if the cell is white. If the cell in (a) is white, though blue in (b), it means that the gain calculation for the channel failed. (c) The plot that maps the distribution of averaged gain in each ASIC on the grid with a crate index and ASIC index. The ASICs indicated by red and purple circles correspond to the ones in (a).

4.5 Operation of SuperFGD Online Monitor in the beam run

The developed online gain monitor has been implemented in the on-site monitoring system of the experiment. The gain position map (Fig. 4.8), the averaged gain plot (Fig. 4.9 (c)), and the gain distribution (Fig. 4.7 (a)) are used for online monitoring during the beam run from November to December 2024 and will be used for the beam run in 2025. The gain position map is mainly used to find too-high gain regions and to monitor dead-like anomaly channels by displaying the number of empty channels in the map, which is equal to the number of channels that failed in the gain calculation. The averaged gain plot is mainly used to detect the issue in the readout electronics units. The gain distribution is used to watch the overall trends by monitoring the mean value and Gaussian-like shape of the distribution. The snapshots of the gain position map and the averaged gain plot in the actual Online Monitor are shown in Fig. 4.10.

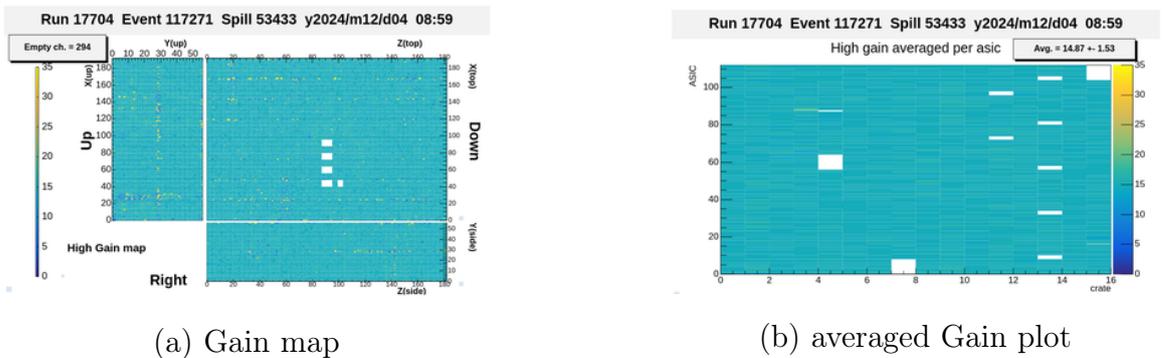


Figure 4.10. The snapshot of the Online Monitor using in the actual beam run

Chapter 5

Evaluation of the effect of bad channels on track reconstruction

The track reconstruction for SuperFGD was originally developed and evaluated under the assumption that all channels function properly. However, some channels have exhibited abnormal behavior during operation. To address this, data from these problematic channels is excluded by ignoring their hit information during analysis. This chapter investigates the impact of such masking on track reconstruction, comparing it to the ideal scenario where all channels are fully operational.

5.1 Bad channels during the beam run

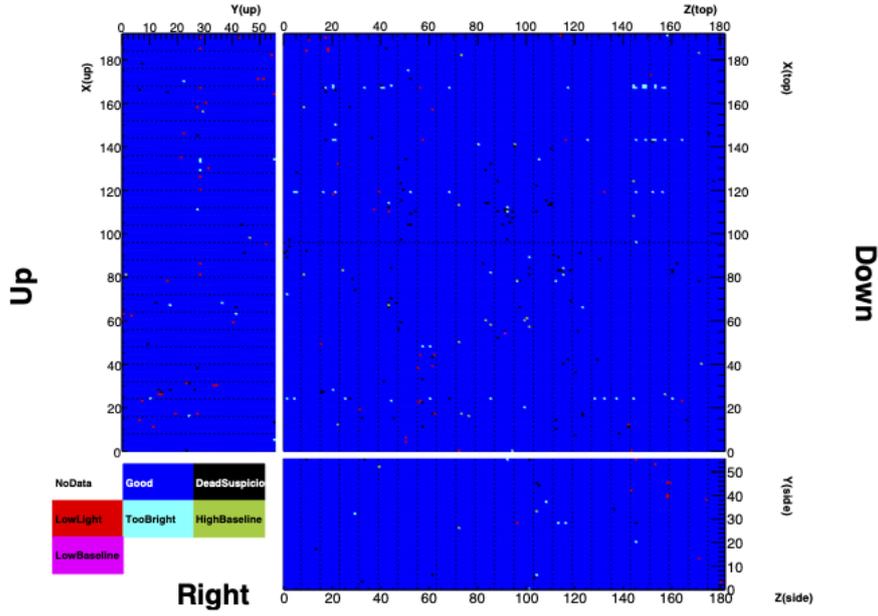
In the superFGD, anomaly channel detection is conducted as a part of calibration. The detected anomaly channels are collectively called as “bad” channels while other channels are called as “good” channels. Since only well-calibrated channels should be used for the analysis, the channels which failed in gain calibration are also included in bad channels.

Bad channels can be caused by a bad cable connection or an issue in a unit of readout electronics. In such cases, the channels that belong to the specific DAQ unit can fall in bad altogether. Indeed, the issues in ASIC and FEB have happened so far, and they caused 32 and 256 bad channels which belong to them, respectively.

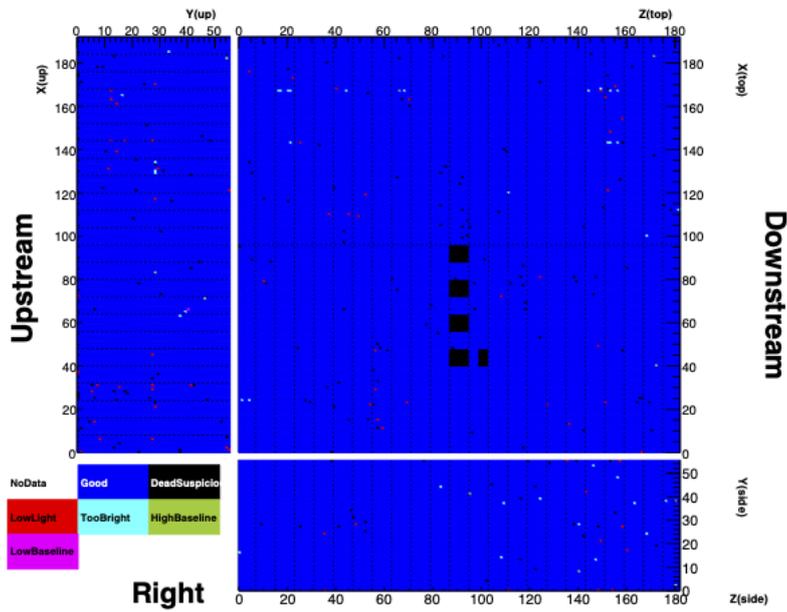
The tables of the number and fraction of bad channels in the recent runs are shown in Table 5.1, and the distribution on the 2D map is shown in Fig. 5.1. The numbers of problematic FEBs and ASICs are also listed in Table 5.1, and the value in the column of “channel” is the number of bad channels other than that caused by the issue in FEB or ASIC.

Table 5.1. Bad channels in the recent beam run. The value in the column of “channel” is the number of channel-level bad channels. The bad channels which belong to the problematic FEBs or ASICs are not counted.

Beam time	Calibration date	FEB	ASIC	channel	total	fraction (%)
Jun. 2024	Jun. 26th, 2024	0	0	331	331	0.59
Nov.-Dec. 2024	Dec. 4th, 2024	1	1	290	578	1.03



(a) Jun. 2024 Run



(b) Nov.-Dec. 2024 Run

Figure 5.1. Map of the flag of channels based on HG calibration.

5.2 Simulation and reconstruction

5.2.1 Monte Carlo simulation

The performance of reconstruction is evaluated by processing the reconstruction steps with the samples generated by Monte Carlo simulation. In this study, particle gun sample, which is the sample simulating the behavior of specific well-known type of particle, is used.

The passage of particles through the detector material is simulated based on the Geant4 [13] software library. The geometry of upgraded ND280 including SuperFGD is implemented. While the result of sampling measurement is reflected to the parameters of the components (e.g. side lengths of cubes), the misalignments in the SuperFGD and between SuperFGD and other subdetectors are not reflected since the position calibration studies are ongoing.

The generation of scintillation light, propagation through WLS fibers, and detection with MPPC are simulated based on the measured performance of components. A part of parameters used in the simulation is shown in Table 5.2. The response of readout electronics is also simulated. The charge and timing information is recorded in a format similar to the real detector data.

Table 5.2. Major parameters used in the detector response simulation

component	parameter	value
Scintillation cube	Photons per energy deposit	225 photons/MeV
	Cross talk fraction	3.0 %/face
WLS fiber	Attenuation length of long component	280 cm
	Attenuation length of short component	32 cm
	The fraction of long component	0.75
MPPC	Gain	5.5×10^5
	Dark noise rate	3000 Hz

5.2.2 Track reconstruction

The reconstruction starts with the set of hits, which consist of charge and timing information, from triggered channels as input. Hereinafter, the pair of charge and timing information at each triggered MPPC channel is called “2D hit” to distinguish it from the reconstructed hit described later. The number of 2D hits in each event is the number of triggered MPPC channels.

The workflow of the reconstruction is shown in Fig. 5.2. The process to reconstruct the track consists of two major steps: 3D hit reconstruction and track reconstruction.

3D hit reconstruction starts with dividing 2D hits into groups based on the hit timing. Next, in each group of 2D hits, the number of photons generated in each cube is estimated to be consistent with the charge information of the 2D hits in the group. Maximum likelihood estimation is used to estimate the number of photons generated in each cube. The cubes in which one or more photons are estimated to be generated are called “3D hits”. The 3D hit information also consists of the charge and timing information. The

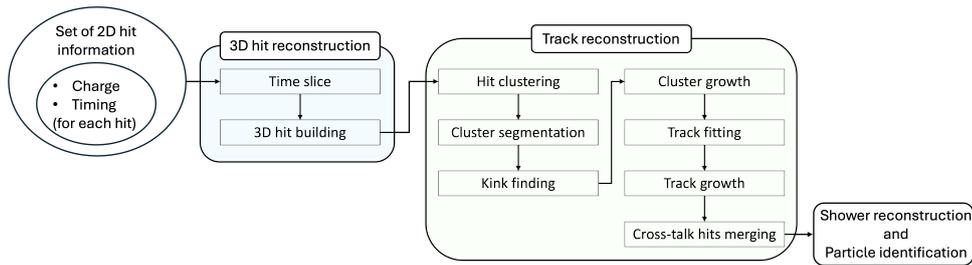


Figure 5.2. the workflow of reconstruction in SuperFGD, which focuses on the steps until track reconstruction.

charge information is the number of photons estimated to be generated in the cube, and the timing information is calculated from the timing information of 2D hits, which correspond to the fibers passing through the cube.

The track is reconstructed from the set of 3D hits reconstructed in the previous step. First, nearby 3D hits are connected by the edge. Next, 3D hits are segmented so that only hits that are part of the confidently same track belong to the same group by cutting the connection around the points where interaction or scattering is suspected to have occurred. Finally, the track is reconstructed by fitting the segmented hits. The reconnection of the segments is also tried before and after track fitting to reconnect the segments that were accidentally separated, though it was a single track. The reconstructed track is used in the shower reconstruction, particle identification, and further analysis. More detail of the track reconstruction method is described in Appendix A.

As described above, the track reconstruction consists of many steps. Therefore, the effect of the mask on channels cannot be calculated simply and has to be evaluated using simulation samples.

5.3 Evaluation method

The effect on the single track reconstruction is evaluated with a sample of muon events generated in SuperFGD. The profile details of the samples are shown in Table 5.3.

Table 5.3. Simulation Samples Profile

	conditions
particle	muon (μ^-)
number of events	10000
initial point	uniform random in SuperFGD
initial direction	uniform random in 4π solid angle
initial momentum	uniform random in 0.1 – 2.5 GeV

To evaluate reconstruction performance with bad channels, specific channels are masked, and their hit information is discarded in the simulation. Reconstruction is then performed with the masked data and compared to results without masking. The list of masked channels is predefined, and it does not depend on the events.

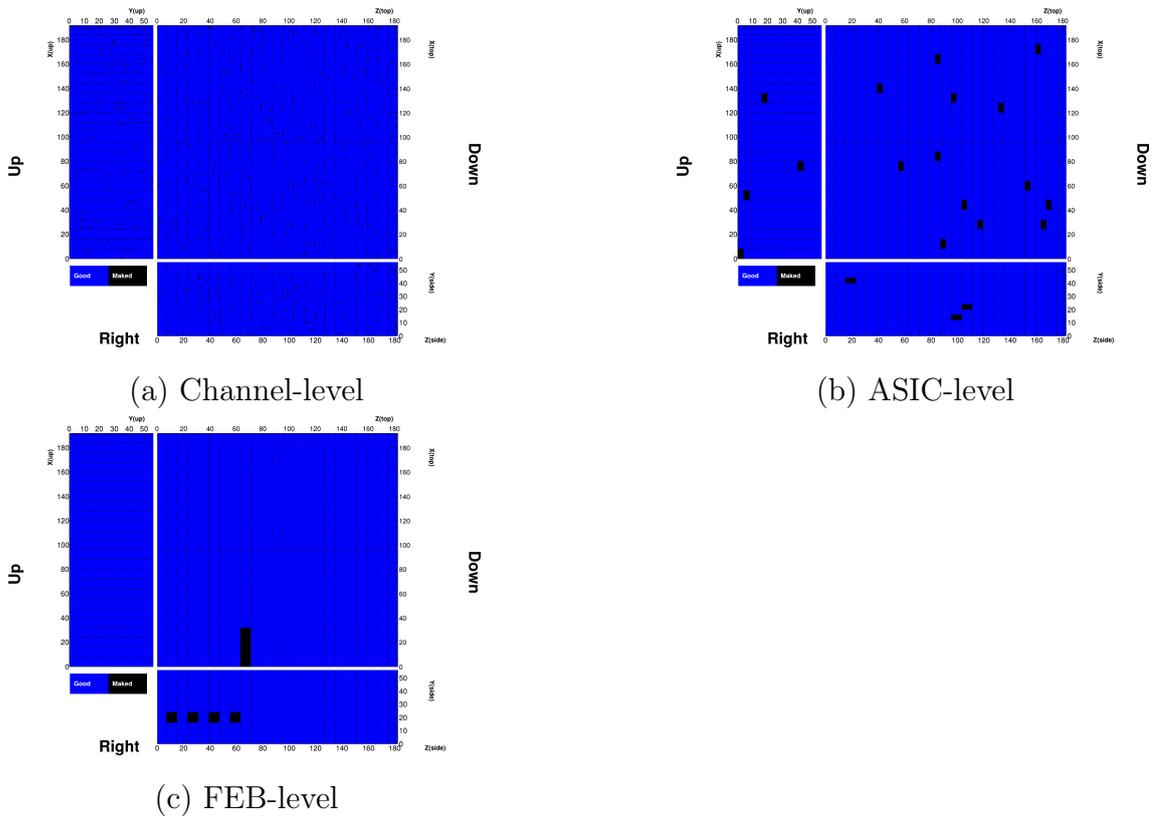


Figure 5.3. Examples of the mask.

Considering the actual situation, the impact of the defect of each level of the units in DAQ was investigated. Channel, ASIC, and FEB were used as a unit in DAQ. When an ASIC is in trouble, a block of 4×8 channels will be missed. When an FEB is in trouble, four blocks of 8×8 channels which are covered by eight ASICS will be missed.

In this investigation, five mask conditions were used. The conditions are summarized in Table 5.4. The set of masked channels was determined by randomly choosing units with a certain probability. Three masks were prepared for each condition to reduce the effect of position dependence. Examples of the mask of condition type 1, 2, 3, which are channel, ASIC, and FEB-level mask with 1% fraction, is shown in Fig. 5.3.

Table 5.4. Mask conditions

type	unit	fraction
1	channel	1 %
2	ASIC	1 %
3	FEB	1 %
4	channel	0.1 %
5	ASIC	0.1 %

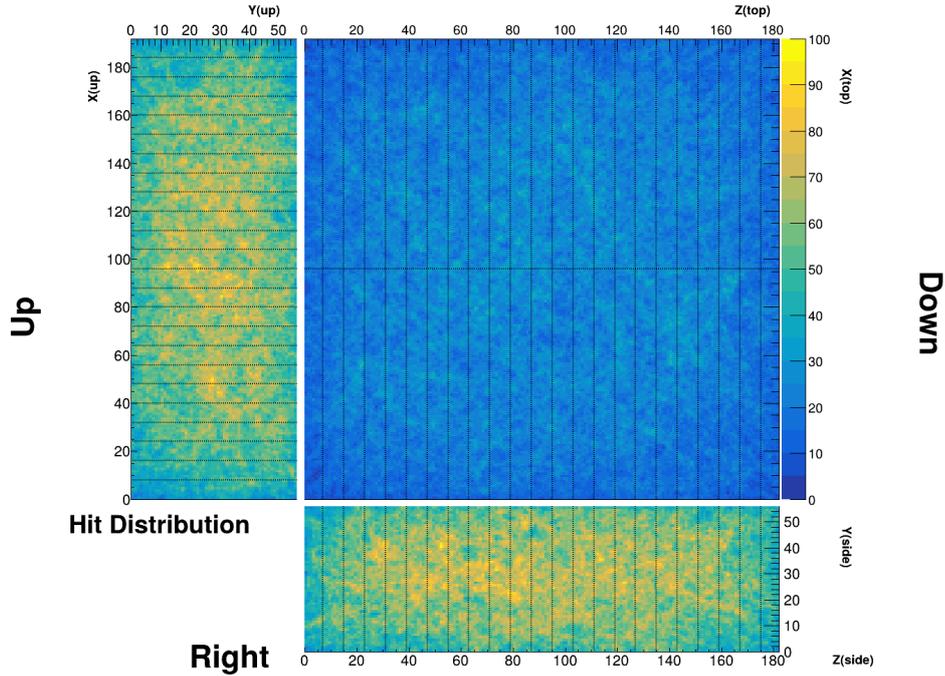


Figure 5.4. Hit distribution of samples before mask. This displays that how many events have hits at each channel.

5.4 2D hit distributions with bad channels

First, the fraction of the lost hits by the mask was investigated. The fraction is expected to be almost the same as that of masked channels. However, the hit distribution is not uniform, as shown in Fig. 5.4. Thus, the fraction of the masked hits also depends on which view the mask is in and where it is located within the view. For example, the XY- and YZ-view channels have more hits than XZ-view due to the depth, which means the number of cubes skewered by the corresponding fiber. Also, the channels near the center have more hits than those near the surface.

9672 events that had at least one hit were used for this investigation. In the remaining 328 events, the muon was generated on the surface of SuperFGD and moved away from the detector, so no hit was recorded. The number and the fraction of masked channels and the fraction of masked hits for each mask are summarized in Table 5.5. The number of events that lost at least one hit by the mask and the average number of lost hits in such events are also shown in the table.

The fraction of masked hits is close to the fraction of masked channels for most masks. However, 1% of FEBs and 0.1% of ASICs correspond to a few units, which means the mask consists of a few blocks of channels. Thus, the fraction of masked hits is heavily dependent on the position of the mask for those masks, and in some masks of those conditions, the fraction of masked hits was different from the fraction of masked hits.

When compared between different unit masks with the same 1% fraction, around 70% of events were affected by channel-level mask, and only a few hits were lost on average, though only around 20% of events were affected by FEB-level mask, and more than hits are lost in average. This is interpreted as being due to the difference in the dispersion and

the size of the block of masked channels. The masked channels are widely scattered over the whole channel map in the channel-level mask, but each is a single masked channel. Therefore, most tracks are across at least one masked channel, but only a few hits are lost by the mask. In the FEB-level mask, the masked channels are located locally, but each one composes 8×8 blocks of masked channels, and three more such blocks exist in the neighborhood. Therefore, the number of tracks across masked channels is limited, but relatively many hits are lost if the track is across the block of masked channels. The ASIC-level mask has features that are in the middle of those of the FEB-level mask and those of the channel-level mask. These differences cause a difference in the track reconstruction efficiency, as discussed in the following sections.

Table 5.5. Masked channel and Masked hits

fraction	unit	set	ch_{mask}	$R_{\text{ch. mask}} (\%)$	$\text{Evt}_{\text{masked}}$	\bar{h}_{mask}	$R_{2\text{Dhit mask}} (\%)$
1%	channel	1	524	0.94	6597	2.7	0.93
		2	555	0.99	6833	2.8	0.98
		3	529	0.95	6598	2.7	0.91
	ASIC	1	512	0.92	2035	7.7	0.82
		2	480	0.86	2225	7.2	0.84
		3	640	1.15	2795	8.0	1.17
	FEB	1	728	1.30	2011	11.3	1.19
		2	512	0.92	1652	12.4	1.07
		3	256	0.46	483	11.3	0.28
0.1%	channel	1	58	0.104	1651	1.1	0.098
		2	45	0.081	1285	1.1	0.076
		3	54	0.097	1620	1.2	0.098
	ASIC	1	32	0.057	247	6.0	0.077
		2	32	0.057	126	7.0	0.046
		3	64	0.115	211	6.3	0.069

ch_{mask} : The number of masked channels.

$R_{\text{ch. mask}}$: The fraction of masked channel to overall 55888 channels.

$\text{Evt}_{\text{masked}}$: The number of events which lost at least one hit by the mask.

\bar{h}_{mask} : The average number of lost hits in the events lost at least one hit.

$R_{2\text{Dhit mask}}$: The fraction of lost hits to a total of 1920662 hits in 9672 events.

5.5 3D hit reconstruction with bad channels

The loss of 2D hits by mask always affects the results of 3D hit reconstruction. For each mask, at least one reconstructed 3D hit was lost in more than 90% of the events that lost at least one hit by the mask. Even if the set of the position of the reconstructed 3D hit was the same, the estimation of the intensity of scintillation light at each cube hit differed.

The number of events that reconstructed 3D hit was lost, the average of the number of the lost 3D hits, and the 3D hit reconstruction efficiency is shown in Table 5.6. The

efficiency of the reconstruction of 3D hits is essential because the track is reconstructed from 3D hits. A plot of the loss fraction of reconstructed 3D hits, which is complementary value of efficiency, against the fraction of masked channel is shown in Fig. 5.5. 3D hit required corresponding 2D hits in all of XZ-, YZ-, and XY-view for reconstruction. This means that to reconstruct the 3D hit at (x_0, y_0, z_0) , the 2D hit located at (x_0, z_0) in the XZ-view, the hit located at (y_0, z_0) in the YZ-view, and the hit located at (x_0, y_0) in the XY-view are all needed. Therefore, the loss fraction of reconstructed 3D hits is expected to be around three times larger than the fraction of masked channel. The existence of the 2D hit in every view is only a requirement, so 3D hits that satisfy this condition are not always reconstructed. However, the results are in good agreement with this prediction. In particular, the results of the channel-level mask, which is robust to position dependence, are in line with the prediction.

Table 5.6. 3D hit reconstruction

fraction	unit	set	$\text{Evt}_{\text{masked}}$	$\text{Evt}_{3\text{Dhit lost}}$	\bar{h}_{lost}	$\text{Eff.}_{3\text{Dhit}} (\%)$
1%	channel	1	6597	6374	4.6	97.17
		2	6833	6614	4.7	97.00
		3	6598	6378	4.6	97.21
	ASIC	1	2035	1972	12.8	97.58
		2	2225	2132	12.1	97.53
		3	2795	2716	13.1	96.58
	FEB	1	2011	1953	18.3	96.58
		2	1652	1619	22.0	96.58
		3	483	471	15.8	99.29
0.1%	channel	1	1651	1553	2.0	99.70
		2	1285	1220	2.0	99.77
		3	1620	1516	2.0	99.71
	ASIC	1	247	240	11.5	99.74
		2	126	119	10.2	99.88
		3	211	206	9.0	99.82

$\text{Evt}_{3\text{Dhit lost}}$: The number of events which lost at least one 3D hit.

\bar{h}_{lost} : The average number of lost 3D hits in the events lost at least one 3D hit.

$R_{3\text{Dhit lost}}$: The fraction of lost hits to a total of 1042994 3D hits in 9672 events.

The distribution of the loss number of the reconstructed 3D hits in each event is displayed in Fig. 5.6. In channel-level masks, only a few hits are lost in each event, and the average of the lost hits becomes less than half when the fraction of masked channel changes from 1% to 0.1%. Around ten hits are lost in each event by ASIC-level mask, and it is almost identical for both masks with 1% and 0.1%. More hits, typically 15 – 20 hits, are lost in each event by FEB-level mask.

Fake 3D hits were reconstructed when some channels are masked. This is explained as a secondary effect of the loss of 3D hits by the mask. An example is shown in Fig. 5.7. In the example, the 3D hit at $(x, y, z) = (75, 14, 97)$ was lost due to the mask for the channel at $(x, y) = (75, 14)$ in XY-view. However, the 2D hit at $(y, z) = (14, 97)$ in YZ-view was still exist, so the 3D hit at $(x, y, z) = (x, 14, 97)$ for any of $0 \leq x \leq 191$ should be exist.

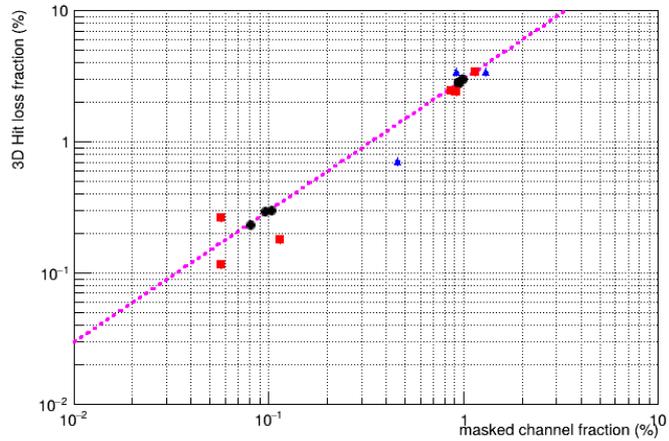


Figure 5.5. A plot of the loss fraction in 3D hit reconstruction against the fraction of the masked channel. The results by the channel, ASIC, and FEB-level mask are presented by the points colored black, red, and blue, respectively. The pink dot line represents that the loss fraction in 3D hit reconstruction is three times as large as the fraction of the masked channel.

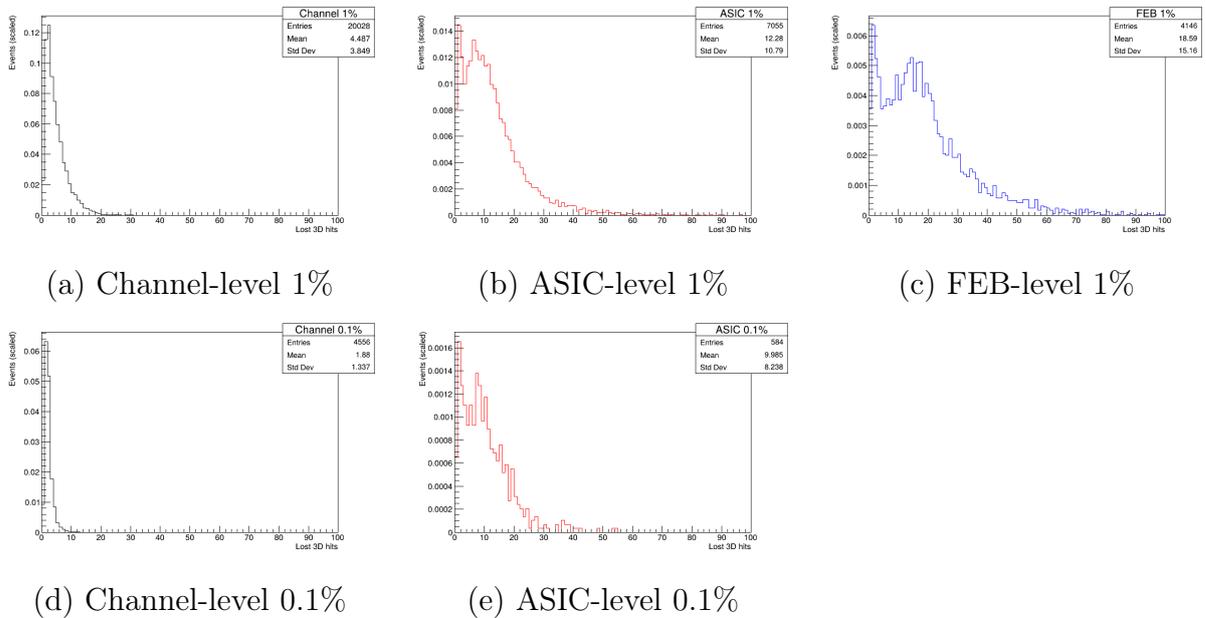
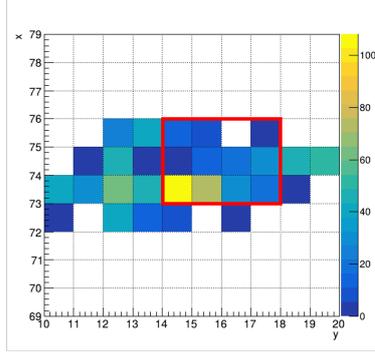


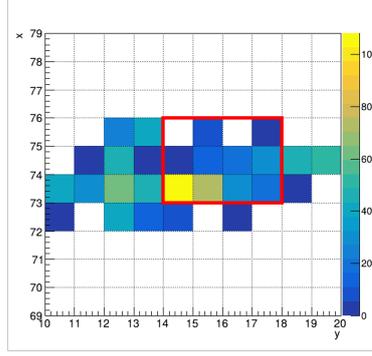
Figure 5.6. The distribution of the loss number of the reconstructed 3D hits.

As a result, the hit at $(x, y, z) = (73, 14, 97)$ was reconstructed, though the hit did not exist in the event. This is a typical mechanism for the wrong additional reconstruction of the 3D hits with the mask.

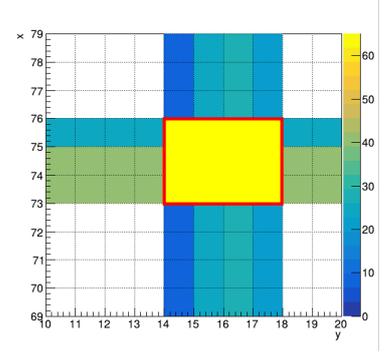
Although it was rare, there were cases where fake 3D hits are reconstructed. In this case, the intensities of scintillation light at other cubes were underestimated due to the effect of masked hits, and fake hits were reconstructed to fill the gap between the intensity of actually received light at the channel and the intensity expected from underestimated light at 3D hits. A case like that was at most one case for each mask in this test.



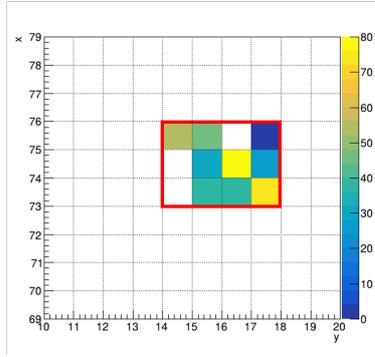
(a) 2D hits in XY-view without mask.



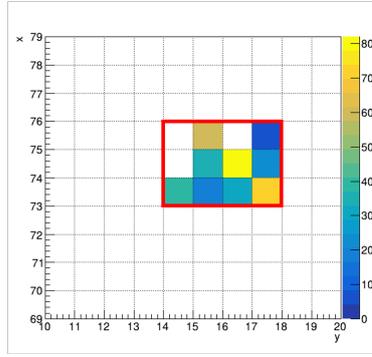
(b) 2D hits in XY-view with mask.



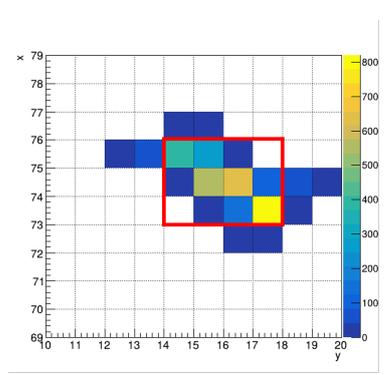
(c) 2D hits in XZ- and YZ-view at $z = 97$.



(d) reconstructed 3D hits without mask at $z = 97$.



(e) reconstructed 3D hits with mask at $z = 97$.



(f) true 3D hits at $z = 97$.

Figure 5.7. An example of how fake 3D hit is reconstructed with the mask for 2D hits. In this example, the fake hit was incorrectly reconstructed at $(x, y, z) = (73, 14, 97)$. (a) and (b) are the acquired 2D hits in XY-view without and with the mask. The 2D hit in XY-view at $(x, y) = (75, 14)$ is lost by the mask in (b). In the following, we consider reconstructing the 3D hits at $z = 97$. The candidate for 3D hits is required corresponding 2D hit in other views. Then, x and y coordinates of the candidates are restricted by XZ-view and YZ-view, respectively. XZ-view does not restrict the y coordinate of candidates since the fiber extends along the y -axis. Similarly, YZ-view does not restrict the x coordinate. (c) displays the hits in XZ- and YZ-views at $z = 97$. As a result of restriction by XZ- and YZ-views, the candidate of the 3D hits at $z = 97$ are narrowed down to the region colored yellow. This restriction is common regardless of the mask. 3D hit candidates are determined by combining XY-view and this restriction. Then, 3D hits are selected and reconstructed from among candidates. The detail of the process is described in Appendix A. (d) and (e) are the reconstructed 3D hits at $z = 97$ from the 2D hits without and with the mask. The 3D hit at $(x, y, z) = (75, 14, 97)$ is lost in the situation with the mask, and the 3D hits at $(x, y, z) = (73, 14, 97)$ is reconstructed instead to make the reconstruction results consistent with the 2D hit in YZ-view at $(y, z) = (14, 97)$. (f) displays true hits in the simulation. It can be seen that the hit at $(x, y, z) = (73, 14, 97)$ did not exist in actuality. Please note that the number of reconstructed hits can be less than the number of true hits since the hits with small charge are not reconstructed.

5.6 Track reconstruction with bad channels

5.6.1 Efficiency in the track reconstruction

Only “good single-track” events were used for the investigation in track reconstruction. The conditions for the “good single-track” events are as follows.

- Muon generation point is in SuperFGD active volume and more than 5.145 cm away from the perimeter.
- No track of the particle in the event except for the primary particle (muon) has intersection with SuperFGD physical volume.

First requirement is to guarantee the reconstruction of the muon track. 5.145 cm is corresponding to the length of five times of the side of the scintillator cube. Thus, the muon should pass at least five cubes otherwise decay in SuperFGD, and the number of 2D hits from those cubes are considered enough large to reconstruct the muon track. Second requirement is to reject the events including extra tracks other than the muon track. In the event, electrons are sometimes kicked out by the muon via the electron-muon scattering and it also make a track and sometimes cause another interaction. In addition, an electron from the decay of a muon can be recorded other than the muon track. In this investigation, these events are excluded and the events which have single track was used for the evaluation. 3908 events were met these conditions and used for the evaluation.

For each mask (including no mask as a reference), the events are classified into three categories based on the track reconstruction results, which are “exactly one track” reconstructed events, “more than one track” reconstructed events, and “no track” reconstructed events. The number of the events in each category and the track reconstruction efficiency, which is the fraction of exactly one track reconstructed events over the 3908 “good single-track” events, are summarized in Table 5.7. The plot of the track reconstruction efficiency against the fraction of masked channel is shown in Fig. 5.8. Only the statistical uncertainty was included in the error of the efficiency. Track reconstruction efficiency was $96.0 \pm 0.3\%$ when no channel was masked. It decreased to around 93% with 1% channel-level mask and to around 90% with 1% ASIC- and FEB-level mask. The decrease caused by the 0.1% channel-level mask was within the error, while the 0.1% ASIC-level mask reduced the efficiency slightly to around 95%.

The reconstruction results for the events that was correctly reconstructed exact one track without mask and at least one hit was lost by the mask are summarized in Table 5.8. Around 2400 events, which are corresponding to 65% of the events that was correctly reconstructed exact one track without mask, lost hits by the 1% channel-level mask, but the exact one track was successfully reconstructed in more than 90% of those events. This is because the reconstruction process is tolerant to the noise and the hits can be considered to belong same track even if the hits are not connected as shown in Fig. 5.9. There are two main cases for the failure in track reconstruction due to single channel loss. One is the case that the track is almost parallel to the fiber that corresponding to the masked channel and many 3D hits are lost by the loss of only one 2D hit. An example of this case is shown in Fig. 5.10 The other is the case that the track was divided as a

Table 5.7. Track reconstruction Efficiency

fraction	unit	set	$\text{Evt}_{N_{\text{track}}=1}$	$\text{Evt}_{N_{\text{track}}>1}$	$\text{Evt}_{N_{\text{track}}=0}$	$\text{Eff}_{\text{track}} (\%)$
	no mask		3752	156	0	96.0 ± 0.3
1%	channel	1	3637	268	3	93.1 ± 0.4
		2	3623	281	4	92.7 ± 0.4
		3	3642	263	3	93.2 ± 0.4
	ASIC	1	3488	415	5	89.3 ± 0.5
		2	3527	376	5	90.3 ± 0.5
		3	3330	573	4	85.2 ± 0.6
	FEB	1	3475	416	8	88.9 ± 0.5
		2	3476	422	2	88.9 ± 0.5
		3	3670	235	3	93.9 ± 0.4
0.1%	channel	1	3747	161	0	95.9 ± 0.3
		2	3754	154	0	96.1 ± 0.3
		3	3740	167	1	95.7 ± 0.3
	ASIC	1	3714	194	0	95.0 ± 0.3
		2	3731	177	0	95.5 ± 0.3
		3	3725	183	0	95.3 ± 0.3

$\text{Evt}_{N_{\text{track}}=1}$: The number of events that exactly one track was reconstructed.

$\text{Evt}_{N_{\text{track}}>1}$: The number of events that more than one track were reconstructed.

$\text{Evt}_{N_{\text{track}}=0}$: The number of events that no track was reconstructed.

$\text{Eff}_{\text{track}}$: The fraction of events that exactly one track was reconstructed.

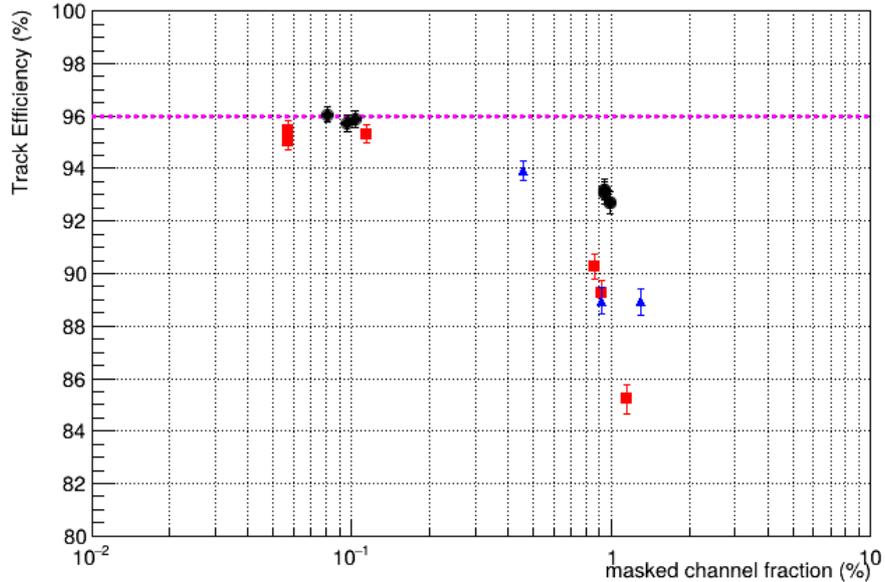


Figure 5.8. Track reconstruction efficiency against the fraction masked channel. Only the statistical uncertainty was included in the error bar and the effect of position dependency was ignored.

kink in reconstruction steps. The shape of reconstructed track was slightly changed by the mask for single channel and the set of hits around the missing hit become unnatural as a single track. An example of this case is shown in Fig. 5.11. With regard to ASIC- and FEB-level mask, the track reconstruction was failed in the events more than half of the events that lost hits by the mask. This is because the track lost adjacent multiple hits when it across the ASIC- or the FEB-level mask, and it is enough to divide the track as shown in Fig. 5.12.

Table 5.8. Reconstruction for good event

fraction	unit	set	$\text{Evt}'_{N_{\text{track}}=1}$	$\text{Evt}'_{N_{\text{track}}>1}$	$\text{Evt}'_{N_{\text{track}}=0}$	$\text{Evt}'_{\text{total}}$
1%	channel	1	2263	130	3	2396
		2	2332	146	4	2482
		3	2284	126	3	2413
	ASIC	1	330	260	5	595
		2	462	223	5	690
		3	465	419	5	889
	FEB	1	451	262	17	730
		2	282	267	10	559
		3	83	78	3	164
0.1%	channel	1	496	10	0	506
		2	384	4	0	388
		3	460	12	1	473
	ASIC	1	55	40	0	95
		2	25	20	0	45
		3	37	26	0	63

The number and fraction is among the events that was correctly reconstructed exact one track without mask and at least one hit was lost by the mask.

$\text{Evt}_{N_{\text{track}}=1}$: The number of events that exactly one track was reconstructed.

$\text{Evt}_{N_{\text{track}}>1}$: The number of events that more than one track were reconstructed.

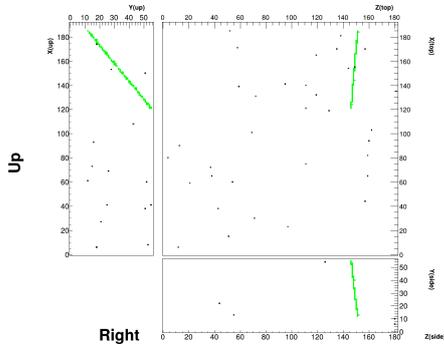
$\text{Evt}_{N_{\text{track}}=0}$: The number of events that no track was reconstructed.

$\text{Eff.}_{\text{track}}$: The fraction of events that exactly one track was reconstructed.

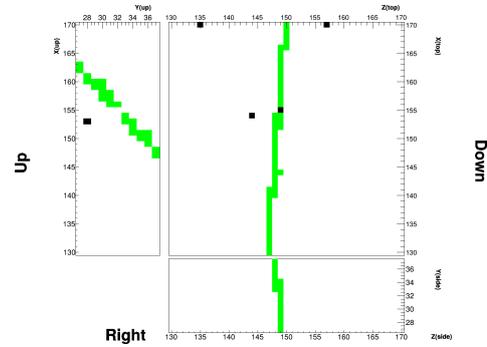
5.6.2 Discussion

The plot of failure fraction in track reconstruction against the masked channel fraction is shown in Fig. 5.13. In this point, the failure fraction is defined as the fraction of the events that failed in reconstruction of exact one track with mask among the events that succeeded in reconstruction of exact one without mask. This can also be represented as $(\text{Evt}_{N_{\text{track}}>1} + \text{Evt}_{N_{\text{track}}=0})/3752$, where $\text{Evt}_{N_{\text{track}}>1}$ and $\text{Evt}_{N_{\text{track}}=0}$ are the numbers as defined in Table 5.8, and 3752 is the number of events that were correctly reconstructed exactly one track without mask as shown in Table 5.7.

The plot was fitted by the line under the assumption that the failure fraction is in proportion to masked channel fraction where the masked channel fraction is relative small, and the results are shown in Fig. 5.13. The fraction of bad channels are also shown in

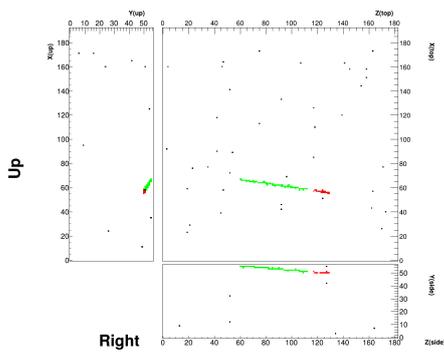


(a) The projection of reconstructed 3D track for each plane.

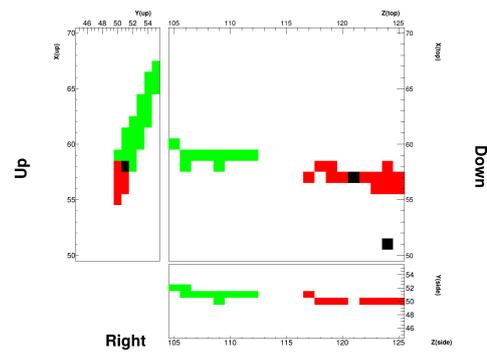


(b) The enlarged picture of (a) nearby the divided point by the mask

Figure 5.9. An example of the event succeeded in track reconstruction even though the track was divided by the mask for single channel.*Legend of the plots will be added later. The detail of description will be also added.

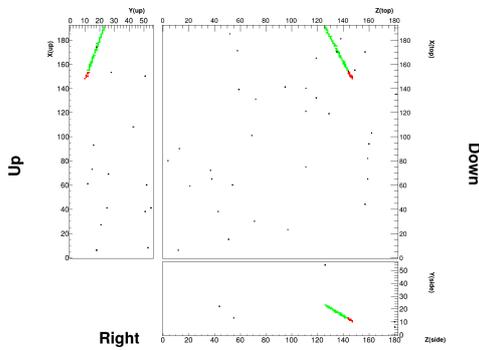


(a) The projection of reconstructed 3D track for each plane.

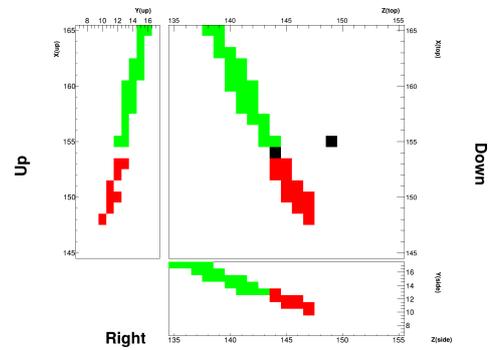


(b) The enlarged picture of (a) nearby the divided point by the mask

Figure 5.10. An example of the event failed in track reconstruction due to the loss of several 3D hits by the mask for single channel.*Legend of the plots will be added later. The detail of description will be also added.

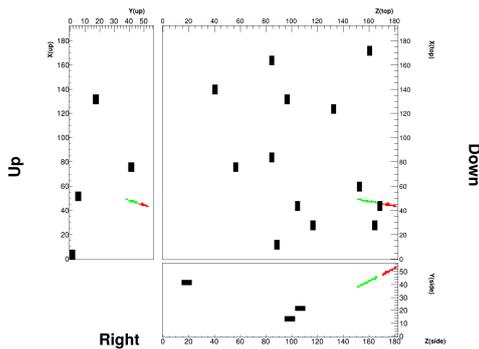


(a) The projection of reconstructed 3D track for each plane.

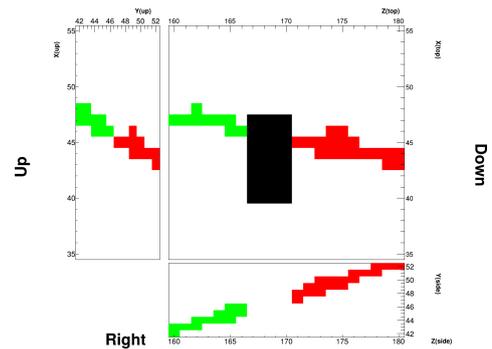


(b) The enlarged picture of (a) nearby the divided point by the mask

Figure 5.11. An example of the event failed in track reconstruction because the track was divided as a kink in reconstruction steps. The shape of reconstructed track was slightly changed by the mask of single channel and the set of hits around the missing hit become unnatural as a single track.*Legend of the plots will be added later. The detail of description will be also added.



(a) The projection of reconstructed 3D track for each plane.



(b) The enlarged picture of (a) nearby the divided point by the mask

Figure 5.12. An example of the event failed in track reconstruction due to the loss of several 3D hits by the masks for ASICs.

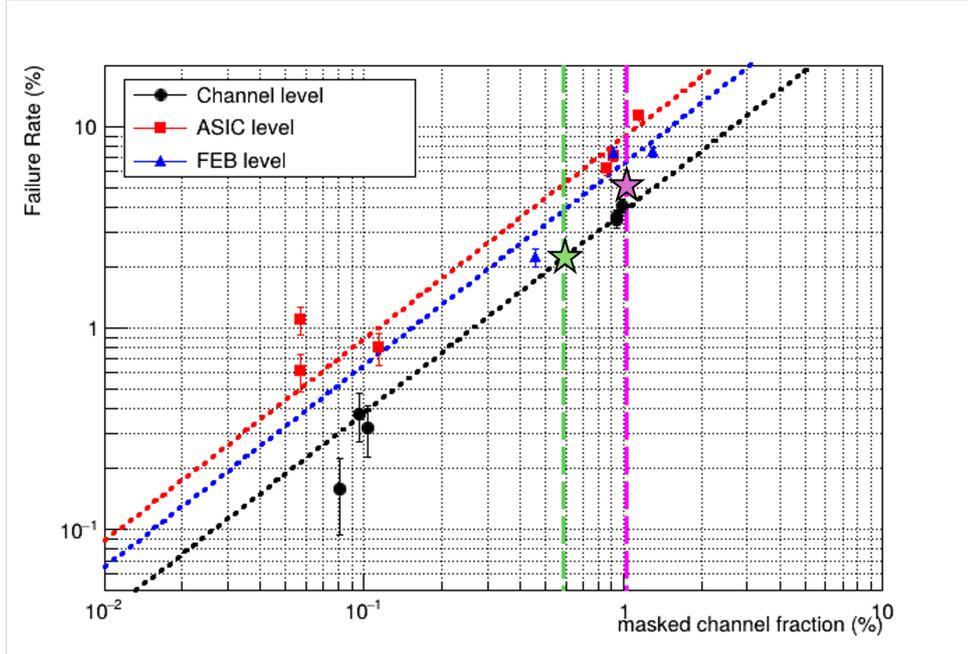


Figure 5.13. The plot of the failure fraction in track reconstruction against the fraction of masked channels. The marker represents the result with each mask. The fit was conducted with the assumption that the failure fraction is in proportion to the masked channel fraction. The dotted line of each color shows the fit results of the markers of the corresponding color. Dashed lines represent the fraction of bad channels in the recent runs. The green one is the fraction in the beam run in Jun. 2024, and the magenta one is the fraction in the beam run from Nov. 2024 to Dec. 2024. The star markers represent the estimated failure fraction in the actual situations, considering the distribution.

the plots. From this plot, the efficiency in the track reconstruction with bad channels in the actual situation can be roughly estimated. In the beam run in Jun. 2024, there were 0.59 % bad channels and no problematic FEB and ASIC. This means that the bad channels were distributed like the masked channels in channel-level masks. Under this condition, the failure fraction is estimated to be around 2.2 %, and then the track reconstruction efficiency will be decreased from 96% to $96 \times (1 - 0.022) \simeq 94$ %. In the beam run in Nov.-Dec. 2024, there were 1.03 % bad channels in total, and there was one problematic FEB and one problematic ASIC. $256 + 32 = 288$ bad channels were due to problematic FEB and ASIC, and the remaining 290 bad channels were distributed like the masked channels in channel-level masks. Then, this bad channel distribution is reproduced as the sum of 0.5 % FEB-level mask and 0.5 % channel-level mask, and the result is expected to be that in the middle of that with 1 % FEB-level mask and that with 1 % channel-level mask. Under this condition, the failure fraction is estimated to be around 5 %, and then the track reconstruction efficiency will be decreased from 96% to 91 %.

Since the 5 % decrease in the track reconstruction efficiency is not negligible, it is necessary to develop a method of reconstruction with bad (masked) channels information. One of the idea is to use the bad channels information in the 3D hit reconstruction step. As described above, currently only the cubes which have corresponding three 2D hits are

selected as 3D hit candidates. In addition to that, including the cubes with two 2D hits and one bad channel in 3D hit candidates is planned. This change is expected to be effective in preventing track dividing due to the loss of adjacent multiple hits by ASIC- or FEB-level bad channels at least. The efficacy of correcting single channel loss of this method needs to be evaluated.

Chapter 6

Separation of electrons from gamma background

In the ν_e -CC cross-section measurement, the misidentification of γ 's as electrons is a significant source of background. To reduce the background, we have attempted to improve e/γ separation using an artificial neural network.

6.1 e/γ separation

6.1.1 e/γ separation in the current ν_e event selection steps

The selection of “ ν_e -CC” events is processed following the reconstruction steps, which are summarized in Sec. 5.2.2. This selection aims to separate “ ν_e -CC” events from “ ν_μ -CC”, “NC background”, “ γ background”, and “other” events.

The selection is composed of several steps. The overview of the selection is shown in Fig. 6.1. e/γ separation is one of the selection steps to reject γ backgrounds from ν_e -CC event. In the e/γ separation step, whether the electromagnetic shower (reconstructed in the previous step) is caused by an electron or a photon is determined.

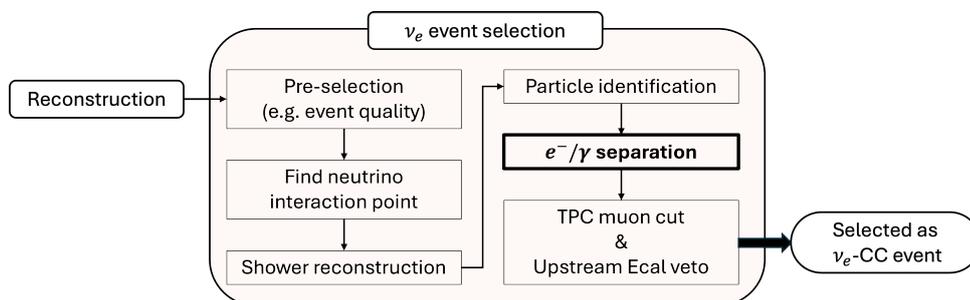


Figure 6.1. Overview of the current ν_e -CC event selection flow. The e/γ separation step is shown on the right in the figure.

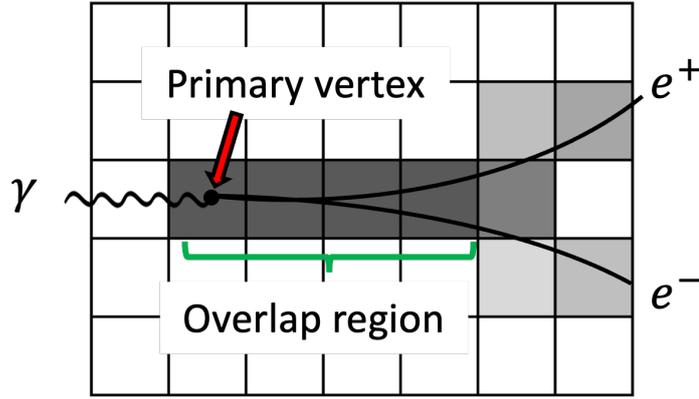


Figure 6.2. A sketch of the hits in the event by electron-positron pair production by the photon.

6.1.2 γ backgrounds

In a ν_μ -CC event involving π^0 production, a photon from the decay of π^0 can be misidentified as an electron, leading to its misclassification as a ν_e -CC event. Although the fraction of such events is small in ν_μ -CC event, they are non-negligible backgrounds since the ν_μ flux is about 100 times larger than that of ν_e ,

Separation of electrons and photons is difficult because both generate electromagnetic showers. A key difference between electron and photon events is that photons produce an electron-positron pair in their interaction. With the typical photon energy, the pair is emitted in the forward direction, initially with nearly overlapping tracks (see Fig. 6.2). This results in energy deposits twice as large as those from a single electron near the neutrino interaction point, allowing separation based on energy deposits in this region.

6.1.3 Motivation

Based on this characteristic of the events, the separation in the current method is conducted by BDT with four parameters: the energy deposits in the first region of the track starting with the neutrino interaction point (such a track is called “primary track”), the distance and the average energy deposit between the neutrino interaction point and the first junction point in the primary track, and the number of tracks branching off from the primary track. This method rejects around 92 % of γ background events with 80 % of e^- efficiency. However, better discrimination capability is needed to reject a large amount of γ background.

The hit map in SuperFGD for each event can be considered as an image data. Specifically, the hit map in each of XZ-, YZ-, and XY-view is regarded as an image, and the channels and charge information of each hit corresponds to the pixels and color of each pixel. γ background rejection is conducted by using hit information in SuperFGD to determine whether an event is caused by an electron or by a photon. This process can be considered as the image classification that tags images as events by “electron” or “photon”. Thus, we have tried to identify electrons from gammas by using an image classification technique.

6.2 Convolutional Neural Networks

The artificial neural network is a machine learning model, which is inspired by the learning mechanism of animals. It consists of calculation units called artificial neurons, and some of them are connected by weighted edges. These calculation units and weighted edges are modeled after neurons and synapses in animal brains, respectively. For most of the model, the artificial neurons are grouped into layers, and the process are conducted layer by layer. Signals are entered into the input layer and propagate from layer to layer with weighted edges. Finally, the signals reach the output layer, and the outputs from the units in the output layer are the outputs of the model.

Convolutional neural network (CNN) is the neural network, which uses convolution in the propagation between layers. CNN is applied to many tasks including image classification. In particular, residual neural network [14], which is one of architectures of CNN model, is representative architecture used for image recognition and classification.

Generally, residual neural network is the architecture for supervised learning. In the supervised learning, the model needs to be trained by samples for training. During training, the model is provided with input samples along with their correct answers. The model is trained by optimizing the parameters (e.g. weights of edges) to move the outputs closer to the correct answers. In the training of the model for image classification, the input sample is an image and the answer is a tag which is a name of the group the picture belongs to. Typically, the training is monitored with validation samples. The samples for the validation are not included in the samples for training, and the progress of the learning is monitored by evaluating the performance of the model with validation samples. The correct answers are used only for the evaluation and not known to the model. Hence, the parameters in the model are not changed by validation samples, and the samples can be used to the validation again and again. After the model is enough trained, the performance of the model is finally evaluated with test samples. Both of the validation samples and the test samples are the samples for the evaluation of the model. However, they differ in that the validation samples are used many times to monitor the training and thus the number of the samples is relatively small, while the test samples are used only at the end of training for the precise evaluation and thus the number of the samples is relatively large.

As described in the end of Sec. 6.1.3, γ background rejection can be considered as the image classification problem. CNN is a typical method for image classification. Therefore, we have tried rejection of γ backgrounds by CNN image classification.

6.3 e/γ separation with CNN

6.3.1 Models

The three hit images in each of XZ-, YZ-, and XY-view from SuperFGD in the event are used as inputs for the classification. Only charge information of each hit is used, and timing information is ignored.

The overview of the process of classification is shown in Fig. 6.3. Three classifications are conducted separately from the image of each view by different models, and the final

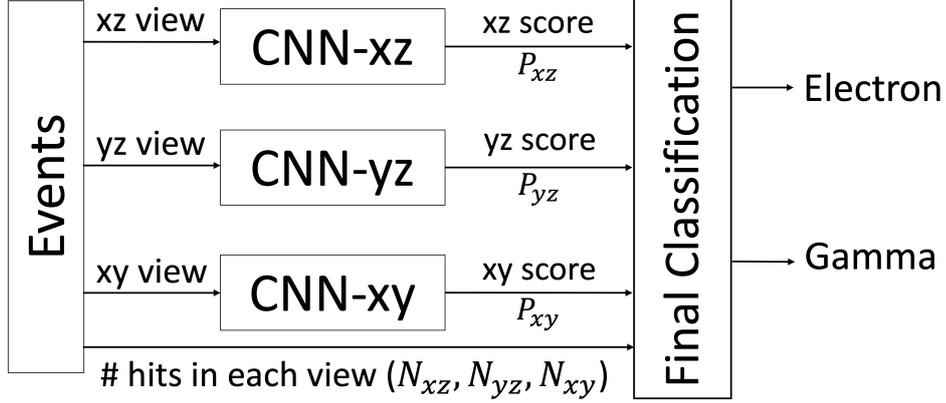


Figure 6.3. Overview of the process of classification.

Table 6.1. Architectures of ResNet. [14] The middle column shows the architecture of ResNet-50. ResNet-50 consists of five major processes and postprocessing. Each major process consists of multiple blocks, and the block is composed of convolution layers with residual connection. An example of the block is shown in Fig. 6.4. Postprocessing is conducted by average pooling, fully-connected, and softmax layers.

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
		3×3 max pool, stride 2				
conv2_x	56×56	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
conv4_x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
conv5_x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

classification for the event is decided by summarizing those outputs. The three models have same structure, but the parameter is different, since each model is optimized for the purpose of the classification from the image of each view.

Regarding the details of the methods for the classification from each view, the model based on ResNet-50 [14] is used. It is one of the original architectures of the Residual Neural Networks. The architectures of residual networks, including ResNet-50, are shown in Table 6.1. The residual neural networks have characteristic structures called “Residual blocks”. One of the residual block structures included in ResNet-50 is shown in Fig. 6.4. In the block, the state before the convolution is added to the result of the convolution. Then, the result of the convolution corresponds to the difference, or “residual”, between the input and output of the block. This structure enables the model to remain the feature of the original input to the model even after processing many layers. Therefore, all layers work effectively when the number of layers is increased, and it enables the model to learn the complicated features. The first and the last layers in the model are changed according to the purpose.

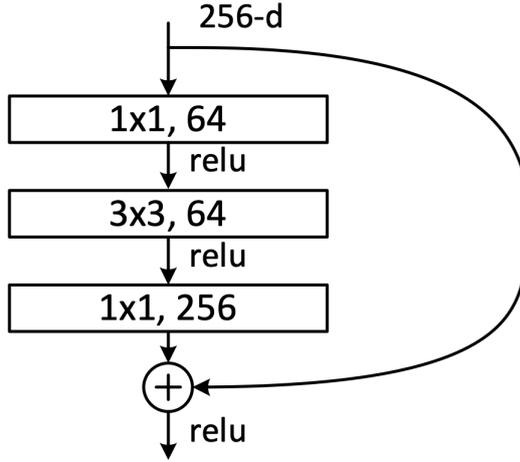


Figure 6.4. Residual block structure included in ResNet-50. [14] It corresponds to the second convolutional block shown in Table 6.1.

The input size for original ResNet-50 is $224 \times 224 \times 3$. This means that the image consists of 224×224 pixels and each pixel has three values that describe the color of the pixel in RGB format. On the other hand, the original size of the SuperFGD hit image of XZ-, YZ-, and XY-view is $192 \times 182 \times 1$, $56 \times 182 \times 1$, and $192 \times 56 \times 1$, respectively. Both the model and the image are fitted to a shape of $224 \times 224 \times 1$. In the original model, the input is convoluted with kernels of the size of $7 \times 7 \times 3$ in the first step. This step is changed to the convolution with kernels of the size of $7 \times 7 \times 1$ to fit the size of the input. As for the image, the hit image in each view is embedded in the center of $224 \times 224 \times 1$ image with actual scale. Other pixels are filled with zeros.

Original ResNet-50 is designed to classify into 1000 object categories, and the output of the model is an array $\mathbf{p} = (p_1, p_2, \dots, p_{1000})$ of length 1000, which satisfy $p_1 + p_2 + \dots + p_{1000} = 1$. This output means that the probability that the input image belongs to category i is p_i . In this study, there are only two categories: electron and photon. Therefore, the last layer of the model is modified so that the output of the model becomes an array of length two whose sum is one. Hereinafter, the probability that the input image is caused by an electron is called as the score for the image.

The final score for the classification is decided by taking a weighted average of the log-likelihoods from the scores in each view. The number of hits in each view is taken as the weight. The formula to calculate the final score is expressed as

$$\log \frac{p_{\text{total}}}{1 - p_{\text{total}}} = \frac{1}{N_{\text{XZ}} + N_{\text{YZ}} + N_{\text{XY}}} \sum_{i=\text{XZ, YZ, XY}} N_i \log \frac{p_i}{1 - p_i},$$

where p_{total} is the total score, p_{XZ} , p_{YZ} , p_{XY} are the scores for the image from the models for XZ-, YZ-, and XY-view, respectively, and N_{XZ} , N_{YZ} , N_{XY} are the number of hits in XZ-, YZ-, and XY-view, respectively. Since the total score is calculated as weighted average, it also takes a value between 0 and 1, and represents the probability that the input image is caused by an electron. The classification is conducted in the view whose image has at least 10 hits. Then, the weight for the view of less than 10 hits is set to zero regardless of the number of hits, and the view is ignored in the final classification.

6.3.2 MC samples for the training, validation, and test

Event samples by electrons and photons for the training, validation, and test of CNN are prepared by Monte Carlo simulation in a similar way described in Sec. 5.2.1. This sample set is referred to as “Sample Set 1” in the following sections. The sample profile is shown in the column of “Sample Set 1” of Table 6.2.

In addition to the simulation samples generated specifically for this study, a common simulation sample in T2K is also used. These samples are made before installing SuperFGD. Therefore, the detector response simulation is slightly inaccurate because the latest parameter from the actual detector performance is not reflected. However, increasing the images for training is essential, so the images are used for the training. This sample set is referred to as “Sample Set 2” in the following sections. The sample profile is shown in the column of “Sample Set 2” of Table 6.2.

The samples are separated into groups of images for the training, validation, and test. The numbers of samples in each group are summarized in Table 6.3. These sample sets are common in the models for each view.

Three models are trained separately. For each view, the model for the classification from the view is trained by only the samples which has at least 10 hits in the image of the view. The number of the samples which passed this precut for each view in training, validation, and test group is summarized in Table 6.4, Table 6.5, and Table 6.6, respectively. Among the samples that passed this precut, 211456 samples from sample set 1 and 864000 samples from sample set 2 of each particle are used to train the model.

For the validation and test, 8000 and 25120 samples of each particle are used, respectively.

Table 6.2. Simulation samples profile.

particle	Sample Set 1		Sample Set 2	
	electron	photon	electron	photon
initial point	uniform random in SuperFGD			
initial direction	uniform random in 4π solid angle			
initial momentum range	[0, 1.5] GeV	[0, 1.5] GeV	[0, 3.5] GeV	[0, 1.5] GeV
number of events	300000	600000	2000000	2000000

Table 6.3. Number of samples for the training, validation, and test.

group	Sample Set 1		Sample Set 2	
	electron	photon	electron	photon
training	260,000	520,000	2,000,000	2,000,000
validation	10,000	20,000	0	0
test	30,000	60,000	0	0

Table 6.4. Samples for training

precut	Sample Set 1		Sample Set 2	
	electron	photon	electron	photon
no cut	260,000	520,000	2,000,000	2,000,000
$N_{XZ} \geq 10$	227,449	227,328	1,864,470	921,827
$N_{YZ} \geq 10$	224,478	222,021	1,848,713	906,698
$N_{XY} \geq 10$	224,456	222,445	1,847,444	906,331

Table 6.5. Samples for the validation

precut	Sample Set 1	
	electron	photon
no cut	10000	20000
$N_{XZ} \geq 10$	8833	8646
$N_{YZ} \geq 10$	8735	8458
$N_{XY} \geq 10$	8734	8481

Table 6.6. Samples for the test

precut	Sample Set 1	
	electron	photon
no cut	30000	60000
$N_{XZ} \geq 10$	26169	26196
$N_{YZ} \geq 10$	25799	25591
$N_{XY} \geq 10$	25801	25639

Table 6.7. Used view in samples for the test.

particle	only XZ	only YZ	only XY	XZ and YZ	XZ and XY	YZ and XY	all
electron	298	96	394	180	310	512	24118
photon	390	91	387	133	345	644	23998

Table 6.8. Hardware Environment and OS.

Items	Parameters
CPU	Intel Core i9-14900KF
GPU	NVIDA GeForce RTX 4090 24 GB GDDR6X
Memory Size	64 GB
OS	Ubuntu 24.04.1 LTS

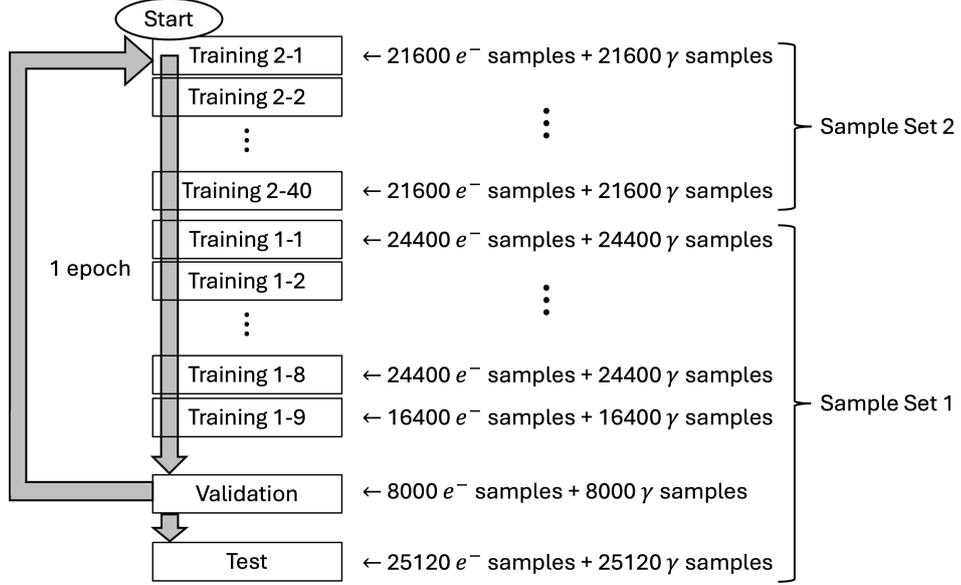


Figure 6.5. Workflow of the training.

6.3.3 Training

The hardware environment and operating system for the training and evaluation of the model are shown in Table 6.8.

Due to working memory limit of the computer, the samples in training group from Sample Set 1 are further divided into 9 subgroups: group 1 – 1, 1 – 2, . . . , 1 – 9. The groups from group 1 – 1 to group 1 – 8 consist of 24384 electron samples and 24384 photon samples. The group 1 – 9 consists of 16384 electron samples and 16384 photon samples. The samples in training group from Sample Set 2 are further divided into 40 subgroups: group 2 – 1, 2 – 2, . . . , 2 – 40. Every group consists of 21600 electron samples and 21600 photon samples.

The overview of the workflow is shown in Fig. 6.5. The training started with the training by samples in Sample Set 2. After that, the model is trained by samples in Sample Set 1. At the end of one cycle, the performance of the model is checked quickly by validation samples. The model is trained by iterating this process. The number of iterations is often called “epoch number”, and this phrase will be used in the following sections. At last, the optimized model is evaluated by the test samples.

In the training by each subgroup, samples in the subgroup are shuffled and grouped into batches of 32. The parameters in the model are updated for each batch. In particular,

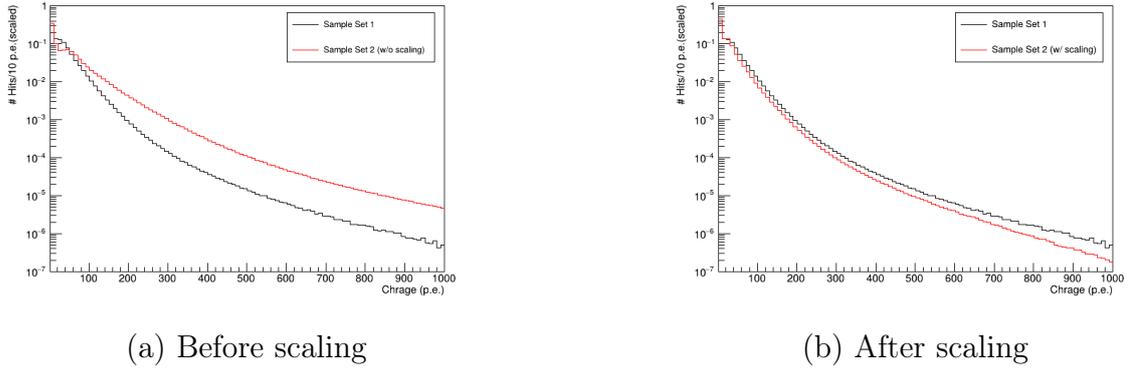


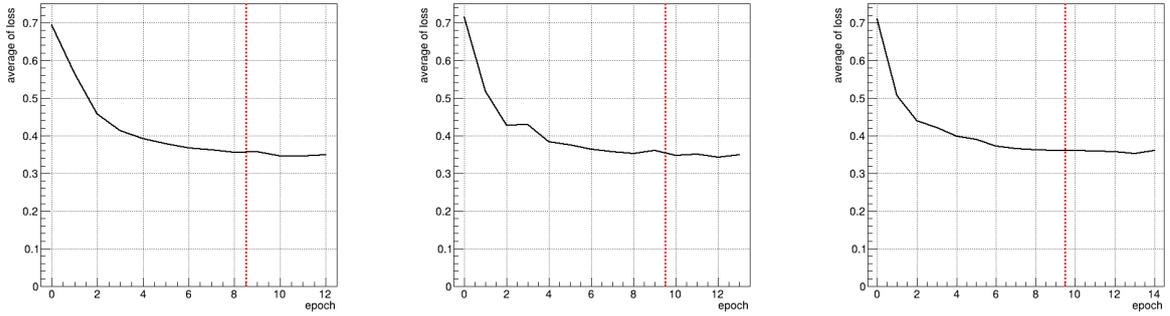
Figure 6.6. The distribution of the charge value of hits in the samples for training. The y -axis is scaled by dividing it by the total number of hits.

the parameters are moved to reduce the value of cross-entropy loss. In the classification problem, the value of cross-entropy loss is calculated as $(-\log p)$ where p is the probability of the image belongs to the correct category. Specifically, when the output from the model is (p_1, p_2) , $p = p_1, p_2$ when the correct category is category 1, 2, respectively. This value is equal to 0 when the model correctly classified the category with a probability of 1, and get increase when the probability for the correct category goes down. Stochastic Gradient Descent (SGD) method is used to update parameters to reduce the loss, and the learning rate is set to 0.001. The cross-entropy loss is also used for the validation. The performance of the model in the middle of training is evaluated by an average of the values of cross-entropy loss for validation samples. Each model took around 25 hours per epoch, and is trained 10 epochs.

Because of the difference in the detector response simulation, mainly in the scintillation efficiency, the charge distribution of hits are different between the Sample Set 1 and the Sample Set 2 as shown in Fig. 6.6 (a). To reduce the difference, the charge of the hits in Sample Set 2 is halved. As a result, the charge distribution of the Sample Set 1 and the Sample Set 2 became close to each other as shown in Fig. 6.6 (b). This change is introduced in the middle of training. The change is reflected immediately before epoch 9, 10, and 10 in the training of the model for XZ-, YZ-, and XY-view.

6.4 Results

The change of the average of cross-entropy loss for validation samples during the training for each model is shown in Fig. 6.7. The loss is decreased by training and converged to certain value. This typical and desirable change in the model training could be seen in the all three models. Moreover, charge scaling for sample set 2 which described in Sec. 6.3.3 is effective. As shown in Fig. 6.8, the averaged loss is converged to a certain value around 0.36 once, but it started to decrease again after scaling, and dropped to around 0.35. The epoch that recorded the minimum averaged loss in training is after scaling the charge in every model. The parameters of the model which recorded the minimum averaged loss in the training are selected as the best parameter for the model and the test is conducted



(a) the model for XZ-view image (b) the model for YZ-view image (c) the model for XY-view image

Figure 6.7. The change of averaged loss for validation samples. Red dotted line in each plot represents the timing of started charge scaling.

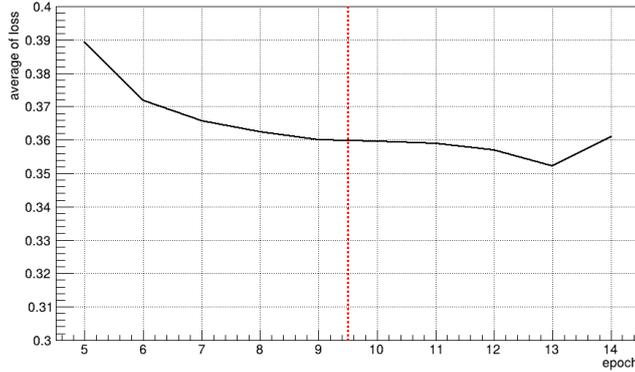


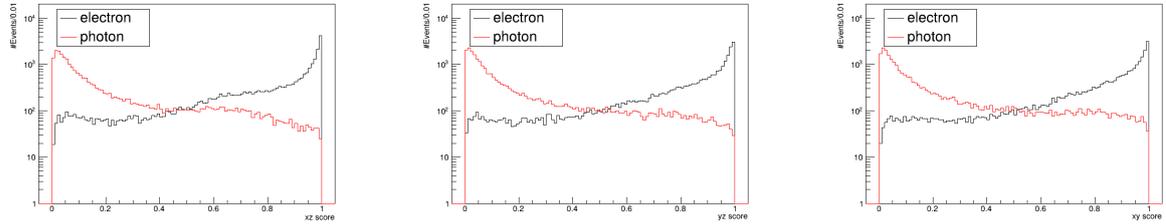
Figure 6.8. The enlarged plots of the change of averaged loss of the CNN-XY after epoch 5. Red dotted line in each plot represents the timing of started charge scaling.

with the parameters. The parameters after epoch 10, 12, 13 are used in the model for XZ-, YZ-, and XY-view, respectively.

The classification results of the models with selected best parameters for test samples are shown in Fig. 6.9. In every plot, the distribution of the score, which indicates the probability of the events is made by electron, has a peak around 1 for electron samples and has a peak around 0 for photon samples. Thus, separation is considered to be successful at a certain precision.

The classification result is shown in Fig. 6.10. The peak of score distribution, which exist around 1 for electron samples and around 0 for photon samples, became sharper than the one in the distribution by each model alone.

The selection of the electron event is conducted by setting a threshold value. Two quantities are defined and used hereinafter: the efficiency and the background rejection ratio. The efficiency is defined as the fraction of events classified as electron events among the true electron events. The background rejection ratio is defined as the fraction of events rejected as non-electron events among true photon events. The background rejection ratio



(a) the model for XZ-view image (b) the model for YZ-view image (c) the model for XY-view image

Figure 6.9. The classification results of the model for each view. Prediction value is the probability of the events is made by electron. Then, the value should be close to 1 for the electron samples and close to 0 for the photon samples. The y-axis is set to log-scale.

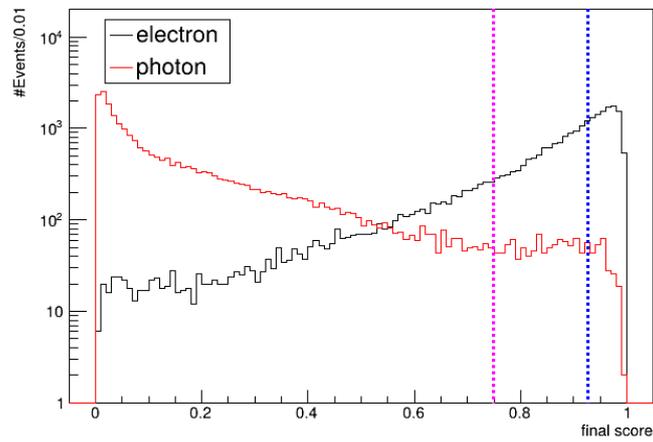
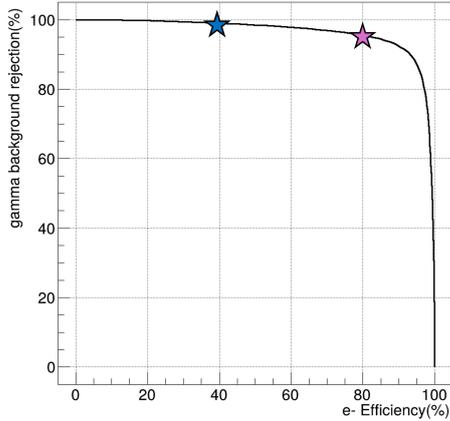
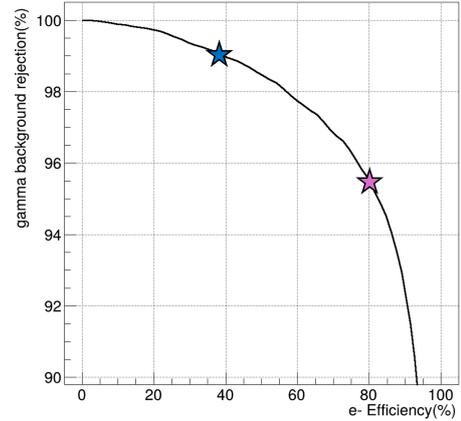


Figure 6.10. The total classification results for electron and photon samples. The score is the probability that the events are made by an electron. The threshold value for the selection with the efficiency of 80% and the selection with the background rejection ratio of 99% are shown by magenta and blue dotted lines, respectively.



(a) overall



(b) enlarged

Figure 6.11. (a) The change of the efficiency and the ratio of background rejection when move the threshold value.

(b) The enlarged plot of (a) from 90 % to 100% on the y-axis. The point of the selection with the efficiency of 80% and the selection with the background rejection ratio of 99% are shown by magenta and blue star markers, respectively.

vs. the efficiency for different threshold value is shown in Fig. 6.11 (a). Fig. 6.11(b) shows a magnified view of the vertical axis from 90 % to 100 %. The threshold value for the efficiency of 80 % is around 0.75, and the background rejection ratio for the threshold value is 95.5%. The corresponding point is indicated by magenta star markers in Fig. 6.11 (a) and (b), and the corresponding threshold value is also displayed by a magenta dotted line in Fig. 6.10. The threshold value for the background rejection ratio of 99 % is around 0.93, and the efficiency for the threshold value is 39.2%. The corresponding point and threshold are displayed by blue star markers and a blue dotted line.

6.5 Discussion

The performance of the separation for each momentum band is shown in Figure 6.12. The performance is decreased in the low momentum region.

As discussed in Sec. 2.5.2, the cross-section measurement in the low momentum region is essential, and the contamination by γ background is significant in the region. Therefore, the poor separation performance in the low momentum region is undesirable. To increase the performance, the additional training with new samples in the region is considered to be effective.

The architecture can also be improved. Currently, the three models calculate the score separately, and the results are finalized by the predetermined formula. By merging the models in the middle, the final classification becomes to be calculated by using neural network.

Once the method is finalized, the performance needs to be compared with the current method. Since the precut is currently different between the current method and CNN

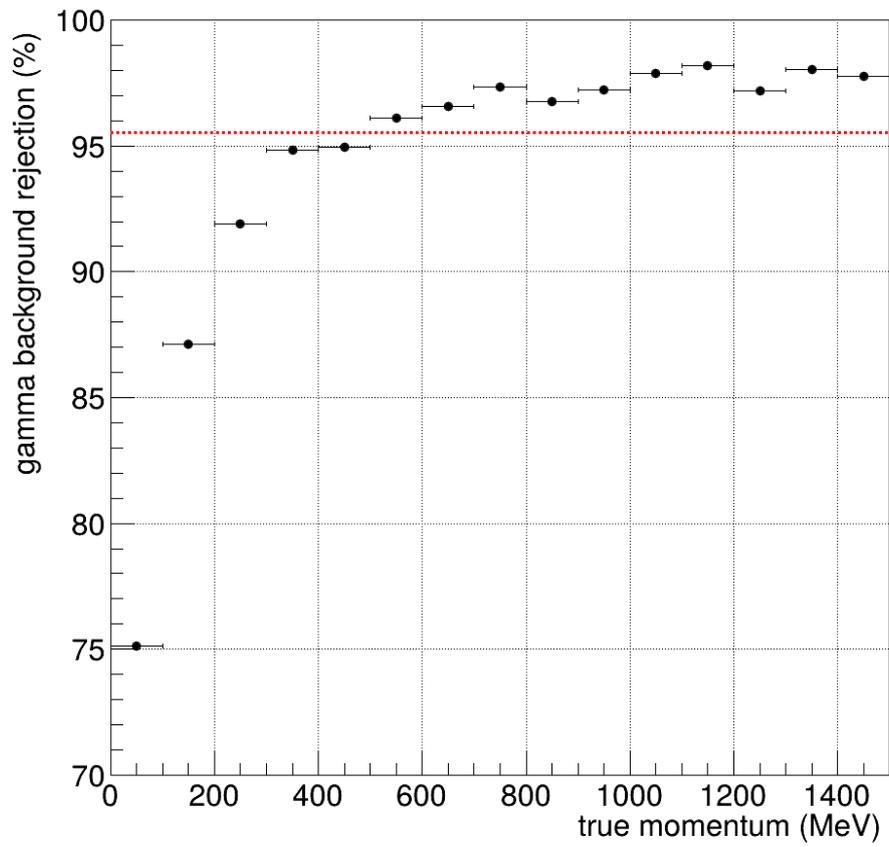


Figure 6.12. The background rejection ratio for each momentum band. The threshold is decided for each momentum bin to achieve 80 % of e^- efficiency. The red dotted line shows the value for all of test samples (95.5 %).

method, the performance of the separation cannot be compared directly. The performance with the samples which passed the selection steps before e^-/γ separation should be evaluated. For the practical use, the performance with the beam samples also should be investigated.

Chapter 7

Summary

The T2K (Tokai to Kamioka) experiment is a long baseline neutrino oscillation experiment. The T2K experiment has a goal of determining CP-violation in the neutrino sector with more than 3σ confidence level. T2K excluded CP conservation at 90 % confidence level so far. The systematic uncertainty on the ν_e and $\bar{\nu}_e$ cross-sections is one of the main source of systematic uncertainties in the oscillation measurements. The direct measurement of ν_e ($\bar{\nu}_e$) cross-section is performed with the T2K near detector ND280. The measurement is limited due to a photon background in the low momentum region and a low reconstruction efficiency for the events with the particle scattering to relatively large angles. To reduce systematic uncertainties further by overcoming the limitation of the reconstruction in ND280 for the track of particles scattered to large angles and short tracks, new detectors were installed in ND280. SuperFGD is one of them, which is a plastic scintillation tracker and also a neutrino target.

In this thesis, we conducted the study to acquire and analyze data in SuperFGD. First, we developed an online gain monitor to keep taking good-quality data. The online gain monitor needs a new fast gain calculation method, and we developed a method that meets both time and accuracy requirements. The developed online monitor is currently being used in the detector operation.

Next, the acquired data is used for event reconstruction as a first step of the analysis. We evaluated the impact of bad channels on track reconstruction. The impact was evaluated in some scenarios based on the past and current situations. This study revealed that clustered bad channels have a greater impact on track reconstruction than discrete bad channels. Additionally, this study enables an estimation of the effect on track reconstruction from bad channel fraction and distribution. This study also asserted the need for reconstruction method with bad channels information based on the estimation result from the current bad channels distribution.

Finally, based on the reconstructed data, the event selection is performed. We explored the method using an artificial neural network to separate an electromagnetic shower caused by an electron from that caused by a photon for the purpose of rejecting γ backgrounds in the ν_e event selection. The model we developed succeeded in the rejection of 95.5 % of γ background events with 80 % efficiency for electrons.

Acknowledgment

I would like to express my deepest gratitude to my supervisor, Prof. Masashi Yokoyama for giving me a special opportunity to conduct my research and excellent continuous guidance. I am also grateful to Prof. Yasuhiro Nakajima and Prof. Kota Nakagiri for their kind support during my master's course and a lot of helpful advice. I am also thankful to the members of the Yokoyama-Nakajima group: Daniel, Xiaodong, Yoshimi, Kodama-san, Okinaga-san, Kobayashi-san, Goto-kun, Mizuno-kun, Muro-kun, Hayasaki-kun, Masaki-kun, Yufei, and Kono-san for the fruitful discussion and great help in my research life. I would especially like to say many thanks to Daniel, Kodama-san, Okinaga-san, and Kobayashi-san for giving me many valuable advice and suggestions. I would also like to give a special thanks to Eguchi-san for giving me insightful advice.

I appreciate the support and feedback from the members of SuperFGD group: T.A. Doyle, T. Matsubara, L. Berns, H. Tanigawa, S. Abe, J. McKean, Arimoto-san, Furui-san, and Tsushima-san. Thanks to you, I enjoyed my work and living at J-PARC. I also thank to the support from Simulation & Reconstruction group and ND280 AI/ML Working Group. The discussion was always interesting and informative for me. I would especially like to extend my sincere thanks to S.A. Monsalve for the worthwhile advice related to machine learning. I would also like to thank the members of the T2K collaboration: A.K. Ichikawa, T. Kikawa, Arihara-san, Kawaue-san, Nishimori-san, Onda-san, and Otani-kun.

Moreover, I would like to express my gratitude to Prof. Tsutomu Mibe and Prof. Taikan Suehara for the insightful comments and suggestions.

Finally, I am extremely grateful to my family for all the kind support and help they gave me.

Appendix A

Track Reconstruction

The workflow of the reconstruction is shown in Fig. 5.2. The process of reconstructing the track consists of two major steps: 3D hit reconstruction and track reconstruction.

A.1 3D hit reconstruction

As described in Sec. 5.2.2, the triggered MPPC channels are called “2D hits”, and the cubes in which one or more photons are estimated to be generated are called “3D hits”. At first, 3D hits are reconstructed from 2D hits. It consists of two detailed steps:

- Time slice
- 3D hit building

Time slice

Since the track goes through SuperFGD in several nanoseconds, the 2D hits from the same track should be recorded within a few tens of nanoseconds. Therefore, the 2D hits in the event are separated based on the timing information. 2D hits are arranged in chronological order. The timing of adjacent hits in the array is compared, and if they are more than 50 ns apart, the array is sliced between those hits. The hits in the segment between slices (or the ends of the array) are grouped as the hits can be from the same track. The hits from different groups will never be used to reconstruct the same track. Typically, multiple groups in the event are caused by the coincidence (e.g. a cosmic ray during the beam trigger) and the noise.

3D hit building

Next, the cubes that satisfy the below condition are listed as candidates for 3D hits.

- There are hits at all three channels corresponding to the fibers that go through the cube, and the hits belong to the same group.

Then, the charge value, which represents the estimated number of photons generated, is allocated to each candidate cube. The value is determined to maximize the likelihood \mathcal{L} ,

whose logarithm is defined by the following formula:

$$\ln \mathcal{L} = \sum_{i \in S} (\ln \mathcal{L}_{i,\text{charge}} + \ln \mathcal{L}_{i,\text{entropy}}), \quad (\text{A.1})$$

where S is a set of candidate cubes, and $\mathcal{L}_{i,\text{charge}}$ and $\mathcal{L}_{i,\text{entropy}}$ are the logarithm of charge term and entropy term of the likelihood of each cube. The charge term of each cube is defined by

$$\ln \mathcal{L}_{i,\text{charge}} = \sum_{j=1}^3 \frac{(Q_{i,j}^{\text{est}} - Q_{i,j}^{\text{obs}})^2}{Q_{i,j}^{\text{obs}}}, \quad (\text{A.2})$$

where $Q_{i,j}^{\text{obs}}$ and $Q_{i,j}^{\text{est}}$ are the observed charge and estimated charge at the channel corresponding to the fiber going through cube i . Estimated charge $Q_{i,j}^{\text{est}}$ at each channel is calculated by summing up the product of the allocated charge and attenuation effect over the cubes penetrated by the fiber. The entropy term, which is introduced to resolve a degeneracy, is written as $-q_i \log q_i$, where q_i is allocated charge to the cube. The optimization is conducted using the gradient descent method. The allocated value is constrained to be positive, but it is not constrained to be integer. The timing of 3D hits is determined independently for each cube by minimizing the chi-square value calculated from the timing of corresponding 2D hits, the propagation time of the light in the fiber, and the uncertainty for timing information of 2D hits. Finally, the candidates whose allocated charge value (=estimated number of generated photons) is less than 2.5 are rejected. The 3D hits with an allocated charge value of less than 50 and have at least two neighbor 3D hits with a charge of more than 50 are rejected.

A.2 track reconstruction

After reconstruct 3D hits, track is reconstructed by segmenting 3D hits and fitting the segmented hits. The step to segment 3D hits consists of three detailed steps:

- Hit clustering
- Cluster segmentation
- Kink finding

Additionally, the step to reconstruct track from segmented 3D hits consists of four detailed steps:

- Cluster growth
- Track fitting
- Track growth
- Cross-talk hits merging

This step includes the steps trying to reunify segmented 3D hits.

Hit clustering

First, the set of 3D hits is divided into connected components called “clusters”. Two 3D hits are considered adjacent to each other when the hits are at at most 1 cube length apart in the x , y , and z directions and the timings differ by at most 15 ns. The coordinate of hit (x, y, z) is a tuple of integer which satisfies $0 \leq x < 192$, $0 \leq y < 56$, and $0 \leq z < 182$, and it is defined based on the position of the cube from left to right, from bottom to top, and from upstream to downstream. A graph is constructed by regarding hits as vertices and the pairs of adjacent hits as edges. Each connected component of the graph corresponds to the connected components of the hits.

Cluster segmentation

Next, the cluster is further divided into paths. This step starts with the graph mentioned above with weighted edges. The weight of edge between hit i and hit j is defined by

$$w_{i,j} = \max(1, |t_i - t_j|/10 \text{ ns}) + \frac{0.01}{1 + c_i + c_j}, \quad (\text{A.3})$$

where (t_i, c_i) and (t_j, c_j) are the timing and charge information of the hit i and j , respectively. Then, a minimum spanning tree (MST) [15] is taken on the graph. After that, the vertex which is one of the endpoints of the longest (simple) path on the tree is selected as a root of the tree. Finally, the paths are collected by iterating the following process.

- Find the farthest vertex from the root among unvisited vertices and take a path to the root, which is uniquely determined.
- Discard the overlap region of the selected path with the paths already collected.
- Separate the path at junctions, which have an adjacent vertex, whose subtree rooted at that vertex has a size of four or more, other than the vertices on the path.
- Add separated paths to the collection, and mark the vertices on the paths as visited.

Kink finding

In this step, the path is cut on the kink point. The kink point is found in the following steps. The path is written as an array of hits, and the i -th component of the array is referred to as hit i . The hit i is judged as a kink point when there is a hit l such that $l < i < l + d - 1$ and hit i is more than the threshold distance away from the segment between hit l and hit $(l + d - 1)$. The distance of hit index d for the baseline segment is set to seven or the path length if it is less than seven. The threshold distance for the kink is set to a smaller of 1.7 cm and 0.7 times the length of the segment between the endpoints of the path.

The path is cut on the largest kink point, which is determined based on the distance of the hit from the segment. Then, this process is repeated to each of the divided tracks until such a kink point does not exist or the path length becomes less than four.

Cluster growth

In this step, the pair of paths which meet certain conditions are merged. First, the average charge value recorded at the hits on the path must exceed 15.0 photoelectron. Both paths have to meet this condition. Second, the paths have to share one of the endpoints. Then, the seven closest hits to the shared endpoint are extracted from each path. The set of seven hits from each track and the merged set of 13 hits are fitted by the line, respectively. Please note that the one hit is shared with the tracks. Then, the pair of paths which satisfy the following conditions are considered to be possibly merged.

- The angle between the fitted (oriented) line is less than 90 degree. The orientation is the direction from the shared endpoint to the other endpoint in one track and conversely in another track.
- The value of $\Delta\chi^2$ is less than eight. $\Delta\chi^2$ is defined by $\Delta\chi^2 \equiv \chi_{\text{merge}}^2 - \chi_1^2 - \chi_2^2$, where χ_{merge}^2 , χ_1^2 , χ_2^2 are the chi-square values of the fitting of the set of merged hit set and of the set of hit set from single tracks, respectively.

For each of the possibly merged pairs, the priority is calculated as $(n_1 \times n_2) + (c_1 \times c_2)/400 - \Delta\chi^2$, where n_1, n_2 and c_1, c_2 are the number of hits and the sum of charge value among hits of the each track of the pair, respectively. Only the highest priority pair is merged, and the process is repeated until the possibly merged pair does not exist.

Track fitting

The path with more than three hits and more than 15.0 photoelectron charge per hit on average is fitted to reconstruct the track. The fitting is conducted in the following processes. First, $n_{\text{sample}} = 1000$ samples of the initial state of the track are prepared. The initial position is sampled in the volume of cubes of the first five hits on the path. The initial direction is sampled as the direction from the initial position to another selected position in the volume of cubes of the first five hits. The curvature is sampled by fitting the sampled consecutive hits on the path. The sampled track seeds are propagated sequentially from the point to the point nearest to the hit on the path. During the propagation, the effects of multiple scattering and hard scattering are included stochastically. The fitting track is derived by averaging sampled tracks with the weight of likelihood at each point. The likelihood of the sampled track is calculated as a product of the likelihood of each hit on the path so far. The likelihood for each hit is constant when the distance from the track to the hit is less than half the length of the cube, and it is exponentially attenuated with an attenuation constant based on the uncertainty of the track position when the distance exceeds half the length of the cube. If the likelihood of the track is too biased at the last point, the fitting process is repeated from the sampling. The process with the track seeds using the last five hits of the path is also conducted, and the track is smoothed by using both forward and backward tracking results. After that, the estimated value of hit timing and energy deposit at each hit cube from the fitting result is calculated.

Track growth

After the fitting, the pairs of the track are tried to merge. First, the distance between the pair of endpoints (among four possible pairs of the endpoints of two tracks) has to be less than 5 cm. Hereinafter, we refer to the endpoints of the nearest pair simply as endpoints of tracks. The tracks in the pair are called the leading track and the following track. Let the position of the endpoints and the (normalized) direction of the tracks at the endpoint of the leading track and following track be $\mathbf{p}_1, \mathbf{p}_2$ and $\mathbf{d}_1, \mathbf{d}_2$, respectively. The vector from the endpoint of following track to the endpoint of leading track $\Delta\mathbf{p} \equiv \mathbf{p}_1 - \mathbf{p}_2$ is decomposed to the component parallel to leading track $\Delta\mathbf{p}_{\parallel}$ and the component perpendicular to the leading track $\Delta\mathbf{p}_{\perp}$.

Second requirement is the length of overlap of the tracks has to be less than 1 cm. The length of the overlap is calculated as $-(\Delta\mathbf{p}_{\parallel}) \cdot \mathbf{d}_1$. Then, the pairs which meets either of following two conditions are merged.

- The angle between tracks is less than 30 degree and the matching “good ness” is larger than 16.
- The angle between tracks is less than 15 degree and $|\Delta\mathbf{p}_{\perp}|$ is less than 1.5 cm.

The matching goodness is calculated by

$$(\Delta\mathbf{p}_{\perp})^T (\sigma_{p,1} + \sigma_{p,2})^{-1} (\Delta\mathbf{p}_{\perp}) + (\Delta\mathbf{d})^T (\sigma_{d,1} + \sigma_{d,2})^{-1} (\Delta\mathbf{d}), \quad (\text{A.4})$$

where $\Delta\mathbf{d} \equiv \mathbf{d}_1 - \mathbf{d}_2$, and $\sigma_{p,1}, \sigma_{p,2}$ and $\sigma_{d,1}, \sigma_{d,2}$ are the covariance matrices for the position of the endpoints and the (normalized) direction of the tracks at the endpoint of leading track and following track, respectively. When the pair which meets the condition find, they are merged immediately, and repeat this process until no more pairs can be merged.

Cross-talk hits merging

Paths with fewer than four hits are not used to reconstruct the track. Then, those hits are tried to merge with the existing track as cross-talk hits. Every hit in the path should be adjacent to the hits in the track to be merged. Otherwise, the path is left without being merged. When the path meets the condition for the merger, the path is decomposed, and the hits are treated individually. Each hit is merged with the hit with the largest charge among hits adjacent to that hit and on the track.

After this process, the position, charge, and timing information of the hits in the track are updated by taking cross-talk hits information into account. Then, the track is refitted, and we finally get the reconstructed track. The particle identification and interaction vertex reconstruction processes follow after the track reconstruction.

List of Figures

1.1	The measured cross sections of ν_μ -CC ($\bar{\nu}_\mu$ -CC) inclusive interaction and each type of ν_μ -CC ($\bar{\nu}_\mu$ -CC) interaction. [1] The marker displays the experiment result. The main neutrino source in this energy region is atmospheric neutrino and accelerator neutrino. The predictions line is by the NUANCE generator [2].	7
2.1	Overview of the T2K experiment.	9
2.2	(a) is the overview of the J-PARC accelerator complex. (b) is the overview of the neutrino beamline at J-PARC. Neutrino beamline is also shown in the center of (a) without notation.	11
2.3	The fluxes of ν_μ , $\bar{\nu}_\mu$, ν_e , and $\bar{\nu}_e$ at near detectors. [5]	12
2.4	The accumulation of POT over time in the T2K experiment. The plot is produced by T2K beam group.	12
2.5	The plan to upgrade MR beam power. [4]	13
2.6	The estimated muon neutrino survival probability at the far detector and neutrino fluxes for different off-axis angles. [6]	14
2.7	(a) is an illust of the near detectors in the pit. ND280 and INGRID are shown in the top and middle of the figure, respectively. (b) is an exploded view of the ND280. the magnet is closed when acquiring the data.	15
2.8	Overview of Fine grained detector (FGD). [7]	16
2.9	Overview of ND280 upgrade. [8]	18
2.10	Overview of Super-Kamiokande. [3]	19
2.11	The $\Delta\chi^2$ distribution and the confidence intervals in δ_{CP} from fitting to the data. [9]. The result depends on the mass ordering.	20
2.12	(a) is the results of ν_e -CC events selection with the actual data by neutrino mode beam and the simulation data of the beam. The result with simulation data is scaled based on the value of POT. (b) is the results of the CC-crosssection of the ν_e in neutrino mode beam and the $\nu_e, \bar{\nu}_e$ in anti-neutrino mode beam.	21
2.13	The distribution of the reconstructed momentum and scattered angle of the one electron-like ring events observed at SK in neutrino mode beam. Black points are observed events and the drawing is predicted distribution by the model with tuned parameters using real data. The top and left plots are the projection of the histogram. The black histogram and red line are data and predicted distribution, respectively. [9]	21
3.1	Conceptual sketch of SuperFGD. [8]	23

3.2	(a) is a block scheme of CITIROC located in FEBs. (b) is an overview of readout electronics of SuperFGD.	25
3.3	The picture and Conceptual sketch of Light Guide Plate (LGP). [12] . . .	25
4.1	Examples of the plots for SuperFGD Online Monitor. (a) is a map of the number of hits at each channel. (b) is a spectra of ADC values with HG at each of channel in crate 0 – 7. The histogram for each crate is lined up. (c) is a distribution of event data size with each trigger type. (d) is the hit time distribution throughout events (e) is a map of Gain value of each channel.	31
4.2	Typical distribution of ADC value with LED trigger. There are peaks by the events of a certain number of photons. The peaks are more evident than in the histogram for the online monitor because this is a histogram for calibration.	33
4.3	The online method for gain calculation	35
4.4	The distribution of gain value calculated with each method.	38
4.5	The comparison of gain value with methods. (b) is a plot enlarged in the rectangle region surrounded by a red line in (a). The red dot line in (b) represents the line that satisfies the condition that online gain equals offline gain.	38
4.6	The ratio of gain with online method to the value with offline method. . .	39
4.7	(a) is a distribution of gain values with the online method and histogramming data. (b) is a distribution of the ratio of the gain value with the online method to the gain value in the calibration. The green dot line represents 0.7 and 1.3, which are the lower and upper limits of 30 % relative error. . .	40
4.8	The plot which filled gain values onto a 2D map of channel positions. The red and purple circle indicates the area of the suspected issue. It was found that the red region is due to the light leakage from the screw hole and that the purple region is due to too-high LED intensity at the guide panel for calibration.	41
4.9	(a) A gain distribution mapped on the grid with a frontend board index and channel index. The red and purple circles indicate the area of dead and noisy suspicious ASICs, respectively. Each ASIC has 32 channels, and the indexes are consecutive. In particular, the channel from index 0 to index 31 belong to same ASIC in each frontend board. (b) The channel map on the DAQ board. For each pair of frontend board index and channel index, there is an allocated MPPC if the cell in the histogram is blue and no allocated MPPC if the cell is white. If the cell in (a) is white, though blue in (b), it means that the gain calculation for the channel failed. (c) The plot that maps the distribution of averaged gain in each ASIC on the grid with a crate index and ASIC index. The ASICs indicated by red and purple circles correspond to the ones in (a).	42
4.10	The snapshot of the Online Monitor using in the actual beam run	43
5.1	Map of the flag of channels based on HG calibration.	45

5.2	the workflow of reconstruction in SuperFGD, which focuses on the steps until track reconstruction.	47
5.3	Examples of the mask.	48
5.4	Hit distribution of samples before mask. This displays that how many events have hits at each channel.	49
5.5	A plot of the loss fraction in 3D hit reconstruction against the fraction of the masked channel. The results by the channel, ASIC, and FEB-level mask are presented by the points colored black, red, and blue, respectively. The pink dot line represents that the loss fraction in 3D hit reconstruction is three times as large as the fraction of the masked channel.	52
5.6	The distribution of the loss number of the reconstructed 3D hits.	52
5.7	An example of how fake 3D hit is reconstructed with the mask for 2D hits. In this example, the fake hit was incorrectly reconstructed at $(x, y, z) = (73, 14, 97)$. (a) and (b) are the acquired 2D hits in XY-view without and with the mask. The 2D hit in XY-view at $(x, y) = (75, 14)$ is lost by the mask in (b). In the following, we consider reconstructing the 3D hits at $z = 97$. The candidate for 3D hits is required corresponding 2D hit in other views. Then, x and y coordinates of the candidates are restricted by XZ-view and YZ-view, respectively. XZ-view does not restrict the y coordinate of candidates since the fiber extends along the y -axis. Similarly, YZ-view does not restrict the x coordinate. (c) displays the hits in XZ- and YZ-views at $z = 97$. As a result of restriction by XZ- and YZ-views, the candidate of the 3D hits at $z = 97$ are narrowed down to the region colored yellow. This restriction is common regardless of the mask. 3D hit candidates are determined by combining XY-view and this restriction. Then, 3D hits are selected and reconstructed from among candidates. The detail of the process is described in Appendix A. (d) and (e) are the reconstructed 3D hits at $z = 97$ from the 2D hits without and with the mask. The 3D hit at $(x, y, z) = (75, 14, 97)$ is lost in the situation with the mask, and the 3D hits at $(x, y, z) = (73, 14, 97)$ is reconstructed instead to make the reconstruction results consistent with the 2D hit in YZ-view at $(y, z) = (14, 97)$. (f) displays true hits in the simulation. It can be seen that the hit at $(x, y, z) = (73, 14, 97)$ did not exist in actuality. Please note that the number of reconstructed hits can be less than the number of true hits since the hits with small charge are not reconstructed.	54
5.8	Track reconstruction efficiency against the fraction masked channel. Only the statistical uncertainty was included in the error bar and the effect of position dependency was ignored.	56
5.9	An example of the event succeeded in track reconstruction even though the track was divided by the mask for single channel.*Legend of the plots will be added later. The detail of description will be also added.	58
5.10	An example of the event failed in track reconstruction due to the loss of several 3D hits by the mask for single channel.*Legend of the plots will be added later. The detail of description will be also added.	58

5.11	An example of the event failed in track reconstruction because the track was divided as a kink in reconstruction steps. The shape of reconstructed track was slightly changed by the mask of single channel and the set of hits around the missing hit become unnatural as a single track.*Legend of the plots will be added later. The detail of description will be also added. . . .	59
5.12	An example of the event failed in track reconstruction due to the loss of several 3D hits by the masks for ASICs.	59
5.13	The plot of the failure fraction in track reconstruction against the fraction of masked channels. The marker represents the result with each mask. The fit was conducted with the assumption that the failure fraction is in proportion to the masked channel fraction. The dotted line of each color shows the fit results of the markers of the corresponding color. Dashed lines represent the fraction of bad channels in the recent runs. The green one is the fraction in the beam run in Jun. 2024, and the magenta one is the fraction in the beam run from Nov. 2024 to Dec. 2024. The star markers represent the estimated failure fraction in the actual situations, considering the distribution.	60
6.1	Overview of the current ν_e -CC event selection flow. The e/γ separation step is shown on the right in the figure.	62
6.2	A sketch of the hits in the event by electron-positron pair production by the photon.	63
6.3	Overview of the process of classification.	65
6.4	Residual block structure included in ResNet-50. [14] It corresponds to the second convolutional block shown in Table 6.1.	66
6.5	Workflow of the training.	69
6.6	The distribution of the charge value of hits in the samples for training. The y -axis is scaled by dividing it by the total number of hits.	70
6.7	The change of averaged loss for validation samples. Red dotted line in each plot represents the timing of started charge scaling.	71
6.8	The enlarged plots of the change of averaged loss of the CNN-XY after epoch 5. Red dotted line in each plot represents the timing of started charge scaling.	71
6.9	The classification results of the model for each view. Prediction value is the probability of the events is made by electron. Then, the value should be close to 1 for the electron samples and close to 0 for the photon samples. The y -axis is set to log-scale.	72
6.10	The total classification results for electron and photon samples. The score is the probability that the events are made by an electron. The threshold value for the selection with the efficiency of 80% and the selection with the background rejection ratio of 99% are shown by magenta and blue dotted lines, respectively.	72

6.11 (a) The change of the efficiency and the ratio of background rejection when move the threshold value. (b) The enlarged plot of (a) from 90 % to 100% on the y-axis. The point of the selection with the efficiency of 80% and the selection with the background rejection ratio of 99% are shown by magenta and blue star markers, respectively.	73
6.12 The background rejection ratio for each momentum band. The threshold is decided for each momentum bin to achieve 80 % of e^- efficiency. The red dotted line shows the value for all of test samples (95.5 %).	74

List of Tables

2.1	The beam runs with upgrade detectors	22
3.1	The timeline of the SuperFGD installation and operation	26
4.1	The number of success/fail channels in gain calculation	37
5.1	Bad channels in the recent beam run. The value in the column of “channel” is the number of channel-level bad channels. The bad channels which belong to the problematic FEBs or ASICs are not counted.	44
5.2	Major parameters used in the detector response simulation	46
5.3	Simulation Samples Profile	47
5.4	Mask conditions	48
5.5	Masked channel and Masked hits	50
5.6	3D hit reconstruction	51
5.7	Track reconstruction Efficiency	56
5.8	Reconstruction for good event	57
6.1	Architectures of ResNet. [14] The middle column shows the architecture of ResNet-50. ResNet-50 consists of five major processes and postprocessing. Each major process consists of multiple blocks, and the block is composed of convolution layers with residual connection. An example of the block is shown in Fig. 6.4. Postprocessing is conducted by average pooling, fully-connected, and softmax layers.	65
6.2	Simulation samples profile.	67
6.3	Number of samples for the training, validation, and test.	67
6.4	Samples for training	68
6.5	Samples for the validation	68
6.6	Samples for the test	68
6.7	Used view in samples for the test.	68
6.8	Hardware Environment and OS.	69

Bibliography

- [1] J. A. Formaggio and G. P. Zeller. From eV to EeV: Neutrino cross sections across energy scales. *Reviews of Modern Physics*, 84:1307–1341, Sep 2012. (pages 7, 83).
- [2] D. Casper. The nuance neutrino physics simulation, and the future. *Nuclear Physics B - Proceedings Supplements*, 112(1):161–170, 2002. (pages 7, 83).
- [3] K. Abe et al. The T2K experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 659(1):106–135, 2011. (pages 9, 11, 15, 19, 83).
- [4] S. Igarashi et al. Accelerator design for 1.3-MW beam power operation of the J-PARC Main Ring. *Progress of Theoretical and Experimental Physics*, 2021(3):033G01, 02 2021. (pages 11, 13, 83).
- [5] K. Abe et al. Measurement of the charged-current electron (anti-)neutrino inclusive cross-sections at the T2K off-axis near detector ND280. *Journal of High Energy Physics*, 2020(10):114, Oct 2020. (pages 12, 19, 21, 83).
- [6] K. Abe et al. T2K neutrino flux prediction. *Physical Review D*, 87:012001, Jan 2013. (pages 14, 83).
- [7] K. Ieki. *The Fine-Grained Detector*, pages 35–50. Springer Japan, Tokyo, 2016. (pages 16, 83).
- [8] K. Abe et al. T2K ND280 Upgrade – Technical Design Report, 2020. (pages 18, 23, 83).
- [9] K. Abe et al. Measurements of neutrino oscillation parameters from the T2K experiment using 3.6×10^{21} protons on target. *The European Physical Journal C*, 83(9):782, Sep 2023. (pages 19, 20, 21, 83).
- [10] LTD. KURARAY CO. Plastic Scintillating Fibers. Accessed on Jan. 9th, 2025. (page 24).
- [11] Weeroc. CITIROC1A - Datasheet V2.53, 2019. Accessed on Jan. 9th, 2025. (pages 24, 25).
- [12] T. Arihara et al. Development of the in-situ Calibration System using LEDs and Light Guide Plates for the SuperFGD. *Journal of Physics: Conference Series*, 2374(1):012118, 2022. (pages 25, 84).

- [13] Geant4 collaboration. Geant4. Accessed on Jan. 9th, 2025. (page 46).
- [14] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep Residual Learning for Image Recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016. (pages 64, 65, 66, 86, 88).
- [15] R. C. Prim. Shortest connection networks and some generalizations. *The Bell System Technical Journal*, 36(6):1389–1401, 1957. (page 80).