修士論文題目

Update of tau neutrino analysis with improved selection method in Super-Kamiokande

和訳：スーパーカミオカンデでのタウニュートリノ選別手法改良とデータ解析のアップデート

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

Research Center for Cosmic Neutrinos(RCCN)　OKUMURA Group
宇宙ニュートリノ観測情報融合センター　奥村研究室

学生証番号	35-176373
氏名	王儒
アルファベット	WANG　RU

ABSTRACT

A large water Cherenkov detector named Super-Kamiokande is located in Japan, which could observe some fully-contained and some partially-contained events per year and has been collecting data for more than 20 years. Tau neutrinos originates from $\nu_\mu \to \nu_\tau$ neutrino oscillations. And charged-current tau neutrino interaction could be observed in Super-Kamiokande because of wide range of neutrino energies and the large target mass of Super-Kamiokande. With the help of these huge volume data, this thesis explores the atmospheric neutrino appearance .

An improved event reconstruction algorithm call fiTQun has been used and also an aggrandize neural network has been implemented in selecting charged-current tau neutrino from background of CC muon neutrino, CC electron neutrino and neutral current. fiTQun reconstructs particle types and kinematics in the Super-Kamiokande detector based on maximum likelihood strategy. 5924.35 days of atmospheric neutrino data are involved in tau analysis. The final result is that tau neutrinos appearance has a significance of 5.4 σ which exceeds 5.0 σ currently. And the tau neutrino appearance is more confirmed now.

ACKNOWLEDGEMENT

First of all, no words could express my appreciation on my supervisor Prof.OKUMURA, for his accepting my application on the University of Tokyo, guidance and supervision both on my research and daily life in Japan, giving me the precious opportunity to join in SK group, offering me lots of flexibility on topic selection and helping me with applying for scholarship and dormitory. There are some misunderstandings because of different culture background. While Prof.OKUMURA is so nice and shows a great patience on me, listens to my explanation and helps me with overcoming the obstacles in front of me, which led me a fantastic and unforgotable days in Japan.

And then, I wish to thank Prof.Roger. I came to Kyoto frequently during my first year and learn reconstruction algorithm, tauNN with the help of him. He is humorous and patient. A thousand of thanks for his time and help.

I would also like to thank Miao and Zepeng who helped me a lot in letting me get in the right track in research. They teached me such necessary knowledge for my research like root, c++, data processing, apfit and fitqun.

What's more, I want to thank people in RCCN who gave me the family feel. In particular, Ms Chieko Mashima helped me a lot on apply for bussiness travel expense.

Last but not least, I want to show my gratitude to my girl friend Hu Hanyu who will be my wife in coming future and my family, for their continuous support and unconditional trust year after year.

Contents

Chapter 1

Introduction

1.1 Neutrino oscillation

There are flavors of neutrino ν_μ, ν_e and ν_τ and they are massless from the point view of the Standard Model. While there are many experiments which observe neutrino oscillations in different ways. These are the evidences that neutrino flavor eigenstates are the superposition of neutrino mass eigenstates.

The flavor eigenstates are expressed as:

$$|\nu_\alpha\rangle = \sum_{i=1}^{3} U_{\alpha i}|\nu_i\rangle \tag{1.1}$$

where $|\nu_\alpha\rangle$ is the flavor eigenstate: ν_μ, ν_e and ν_τ and $|\nu_i\rangle$ is the mass eigenstates: ν_1, ν_2 and ν_3. U is the Pontecorvo-Maki-Nakagawa-Sakata(PMNS) matrix which reveals the transformation relationship:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & cos\theta_{23} & sin\theta_{23} \\ 0 & -sin\theta_{23} & cos\theta_{23} \end{pmatrix} \times \begin{pmatrix} cos\theta_{13} & 0 & sin\theta_{13} \\ 0 & e^{-i\delta} & 0 \\ -sin\theta_{13} & 0 & cos\theta_{13} \end{pmatrix}$$
$$\times \begin{pmatrix} cos\theta_{12} & sin\theta_{12} & 0 \\ -sin\theta_{12} & cos\theta_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} e^{-i\rho} & 0 & 0 \\ 0 & e^{i\sigma} & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{1.2}$$

As for the time evolution:

$$\begin{aligned}|\nu_\alpha(t)\rangle &= \sum_i U_{\alpha i} e^{-iE_i t}|\nu_i(t=0)\rangle \\ &= \sum_i \sum_{\alpha'} U_{\alpha i} U^*_{\alpha' i} e^{-iE_i t}|\nu'_\alpha(t=0)\rangle\end{aligned} \tag{1.3}$$

where E_i is the energy eigenvalue of mass eigenstate $|\nu_i\rangle$. Then the probability between flavor eigenstate and mass eigen state is:

$$P(\nu_\alpha \to \nu_\beta) = |\langle\nu_\beta(t)|\nu_\alpha(0)\rangle|^2 \tag{1.4}$$

$$= |\sum_i U_{\beta i} U^*_{\alpha i} e^{-iE_i t}|^2 \tag{1.5}$$

$$= \sum_i |U_{\beta i} U_{\alpha i}|^2 + \sum_{i \neq j} U_{\alpha i} U^*_{\beta i} U^*_{\alpha j} U_{\beta j} e^{-i(E_i - E_j)t} \tag{1.6}$$

$$\begin{aligned}P(\nu_\alpha \to \nu_\beta) = \delta_{\alpha\beta} &- 4\sum_{i>j} Re(U_{\alpha i} U^*_{\beta i} U^*_{\alpha j} U_{\beta j}) sin^2(\triangle m^2_{ij} \frac{L}{4E})) \\ &+2\sum_{i>j} Im(U_{\alpha i} U^*_{\beta i} U^*_{\alpha j} U_{\beta j}) sin^2(\triangle m^2_{ij} \frac{L}{2E}))\end{aligned} \tag{1.7}$$

where $\triangle m^2_{ij} = m^2_i - m^2_j$ and L is the path length of the neutrino, and the difference of phase between two mass eigenstate,$\phi_i - \phi_j$ can be expressed as $L\frac{m^2_i - m^2_j}{2E}$ since the mass of neutrino ($< 1eV$) is too much tiny than the energies ($> 1MeV$) of current experiments. From these we could know that:

$$P(\nu_e \to \nu_e) = 1 - sin^2 2\theta_{13} sin^2(\frac{1.27\triangle m^2_{32} L}{E}) \tag{1.8}$$

$$P(\nu_\mu \to \nu_\mu) = 1 - sin^2 2\theta_{13} sin^2(\frac{1.27\triangle m^2_{32} L}{E}) \tag{1.9}$$

$$P(\nu_\mu \leftrightarrow \nu_e) = sin^2\theta_{23} sin^2 2\theta_{13} sin^2(\frac{1.27\triangle m^2_{32} L}{E}) \tag{1.10}$$

under the assumption that $\triangle m^2_{31} \approx \triangle m^2_{32}$ since the $\triangle m^2_{21}$ is small.

Additionally, if the neutrino doesn't propagate in vacuum but in matter, the oscillation probability will be modified because of the interactions with the electrons. Specifically, if neutrino travel in homogeneous matter, the oscillation parameters $sin^2\theta_{13}$ and $\triangle m^2_{32}$ will be changed to their matter-equivalents,

$$sin^2 2\theta_{13,M} = \frac{sin^2 2\theta_{13}}{(cos2\theta_{13} - A_{CC}/\triangle m^2_{32})^2 + +sin^2 2\theta_{13}} \tag{1.11}$$

$$\triangle m^2_{32}, M = \triangle m^2_{32}\sqrt{(cos2\theta_{13} - A_{CC}/\triangle m^2_{32})^2 + sin^2 2\theta_{13}} \tag{1.12}$$

where the $A_{CC} = \pm 2\sqrt{2}G_F N_e E_\nu$, N_e is the electron density which is assumed as a constant and G_F is the Fermi constant.

1.2 Atmospheric Neutrino

Atmospheric neutrinos come from the interaction between cosmic ray and the nucleis in air. Such interactions happen:

$$\begin{aligned} \pi^+ &\rightarrow \mu^+ + \nu_\mu \\ \pi^- &\rightarrow \mu^- + \overline{\nu_\mu} \\ \mu^+ &\rightarrow e^+ + \overline{\nu_\mu} + \nu_e \\ \mu^- &\rightarrow e^- + \overline{\nu_e} + \nu_\mu \end{aligned} \tag{1.13}$$

The ratio of muon and electron neutrinos is expected to be about 2 in energy range lower than 1 Gev based on this kinematic chain. It is need to be corrected at higher energies where neutrinos come from decays of parents which have more energies. Due to the extension of the Lorentz time, some of the muons generated by these decays reach the earth before they decay. Therefore, electron neutrinos are not generated, and the ratio of muons to electron neutrinos increases with energy.

Cosmic ray fluxes whose energies are more than 10 GeV are isotropic towards the Earth; hence, at equal angle with the zenith and nadir, the corresponding neutrino flux is expected to be equal. The path length of atmospheric neutrinos

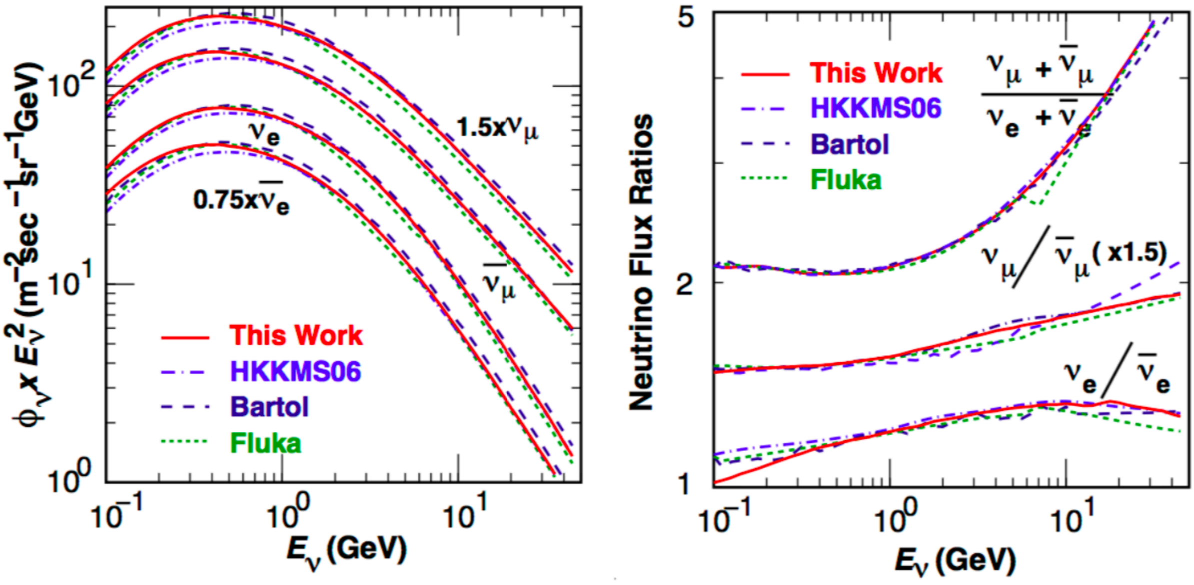


Figure 1.1: neutrino flux

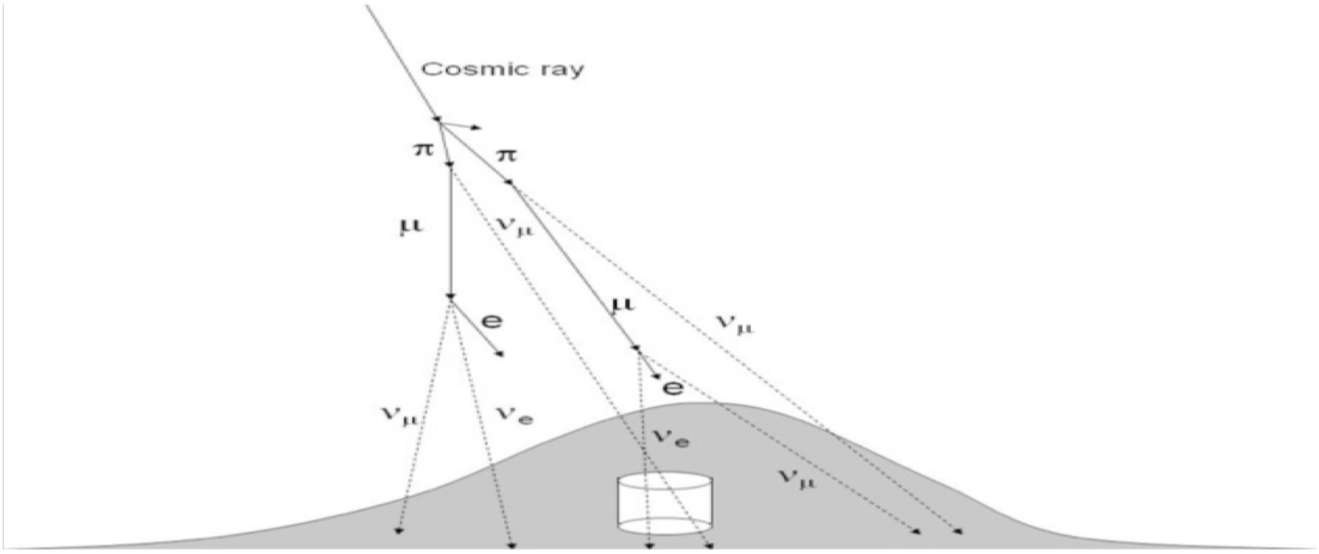


Figure 1.2: atmospheric neutrino

are in range between $O(10)$ km and $O(10^4)$ km. Because of the broad energy range of cosmic rays, the energy of atmospheric neutrinos is from $O(10)$ MeV to more than PeV. The varies of path length and energy makes atmospheric neutrinos a reliable source for neutrino oscillations researches.

1.3 CCQE Cross Section

The cross-section of charge current quasielastic scattering(CCQE) are obtained on the basis of Smith-Moniz CCQE model (Smith and Moniz, 1972) with a Relativistic Fermi gas (RFG) nuclear model. The difference between the model and the Nieves model (Nieves et al., 2004) was identified as the uncertainty of the whole cross section and the cross-sectional ratio $\bar{\nu}/\nu$ and $\frac{\nu_\mu+\bar{\nu_\mu}}{\nu_e+\bar{\nu_e}}$. the π^0 production to the charged pion production is estimated to be 40% and is taken as the systematic error.

The uncertainties on the parameters that characterize the Graczyk-Sobczyk form factor (Graczyk and Sobczyk, 2008) were given by bubble chamber experiments(Wilkinson et al., 2014).

1.4 Tau decays

Tau leptons produced by tau neutrinos will rapidly decay ($mean \quad life \quad time = (2.903 \pm 0.005) \times 10^{-13} s$)to secondary particles. We can split tau decays into leptonic decays

$$\begin{aligned} \tau^- \rightarrow e^- + \nu_\tau + \overline{\nu_e} \quad 17.82\% \\ \tau^- \rightarrow \mu^- + \nu_e + \overline{\nu_\mu} \quad 17.39\% \\ etc \end{aligned} \tag{1.14}$$

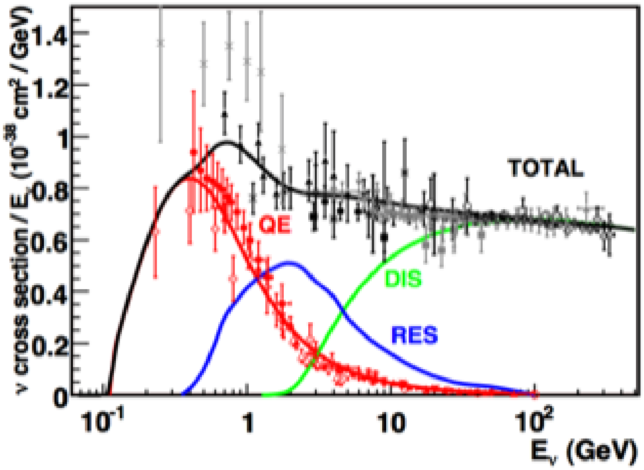


Figure 1.3: CC cross section

and hadronic decays

$$
\begin{aligned}
\tau^- \rightarrow \pi^- + \pi^0 + \nu_\tau \quad & 25.49\% \\
\tau^- \rightarrow \pi^- + \nu_\tau \quad & 10.82\% \qquad (1.15) \\
& etc
\end{aligned}
$$

It is hard to directly detect tau leptons in SK because of the short lifetime of tau leptons.

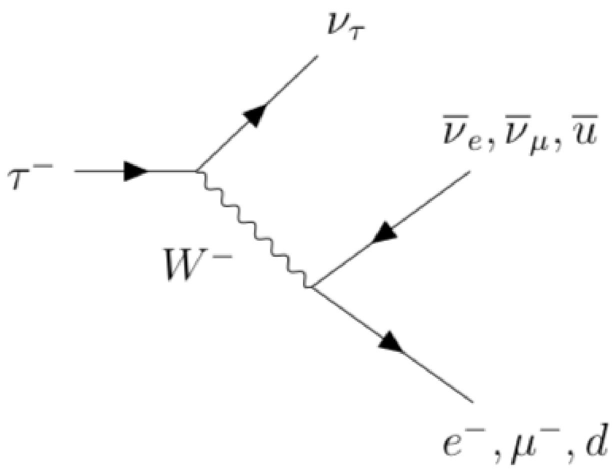


Figure 1.4: tau decay

1.5 Tau Neutrino Cross Section

There are some different types of weak interactions which include a ν_τ or a $\overline{\nu_\tau}$:

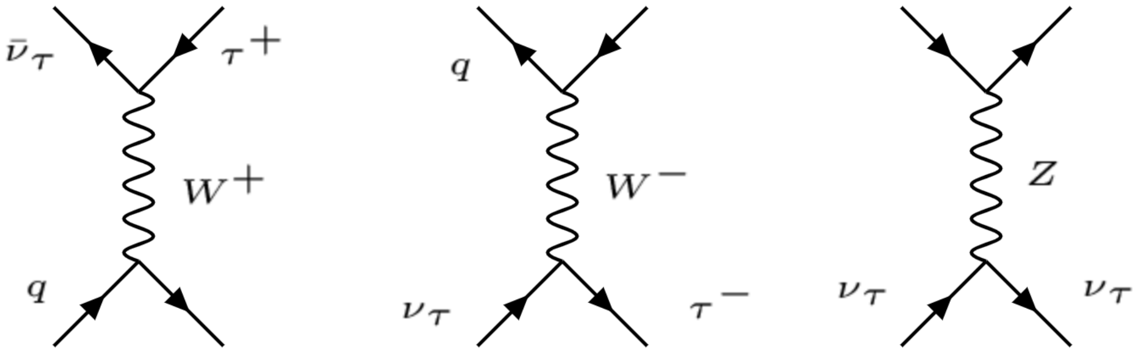


Figure 1.5: Weak interactions which contain a ν_τ or $\overline{\nu_\tau}$

Those weak interactions which are transferred by a charged W boson are classified as charged current interaction, such as:

$$\overline{\nu_\tau} + q \to \tau^+ + X$$

$$\nu_\tau + q \to \tau^- + X, \text{ etc.}$$

Similarly, those weak interactions which are transferred by a neutral Z boson are regarded as neutral current interaction, such as:

$$\nu_\tau + X \to \nu_\tau + X, \text{ etc.}$$

An intuitive interpretation of cross section is probability density for finding the scattered particles at a given angle. Hence, if the probability that the interaction happens is high, the corresponding cross section is large. Vice versa. And because the neutral current interaction is flavor blind, it is impossible to identify the flavor of a neutrino in such kinds of interactions. While charged

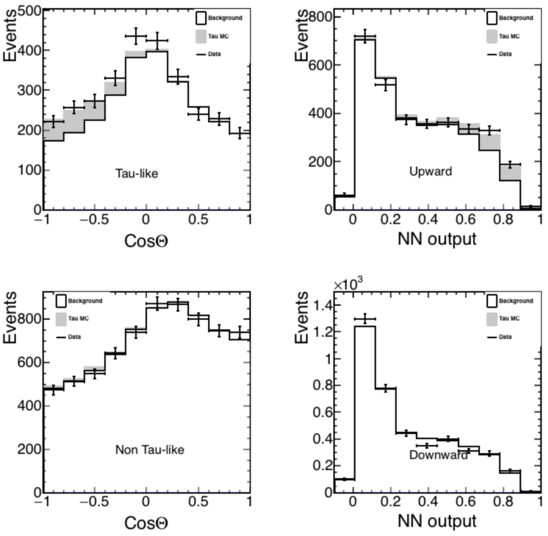


Figure 1.6: Fit results of Zepeng Li's

current interactions are approachable. So in this thesis, we only focus on the charged current tau neutrinos cross section. Moreover, the charged current tau neutrinos cross section is the flux-average of tau neutrinos cross section and anti-tau neutrinos cross section since there are no classification on ν_τ and $\overline{\nu_\tau}$. We will discuss these in Chapter 5.

1.6 Previous Work

The tau excess plot is shown in Fig 1.6. And the hypothesis of no tau neutrinos appearance with a significance of 4.6σ [3], And a flux-averaged cross section is measured to be

$$(0.94 \pm 0.20) \times 10^{-38} cm^2$$

in the neutrino energy range between 3.5 Gev and 70 Gev. The work of this thesis will update this work with improved reconstruction algorithm and tau selection algorithm.

Chapter 2

Super-Kamiokande Experiment

Super-Kamiokande is a cylindrical 50 kton water Cherenkov detector located at Kamioka Observatory of Institute for Cosmic Ray Research in Gifu Prefecture, Japan. Because the detector is covered by rock whose mean thickness is approximately 1000 m, which is equivalent to the effect of 2700 m water covering, the flux of cosmic ray muon has to be reduced by 5 orders compared with that on the surface of the Earth. The main purpose of the Super-Kamiokande experiment is to detect nucleon decays and explore neutrinos from various natural sources including atmospheric neutrinos, solar neutrinos, supernova neutrinos and others. What's more, SK also played a far detector role in K2K experiment (Ahn et al., 2003) from 1999 to 2004. It has been used as far detector in T2K long-baseline neutrino experiment (Abe et al., 2011a) since 2009. Super-K has been running for more than 20 years since 1996. And the datasets of it are separated into five run periods: SK-I(1996-2001), SK-II(2002-2005), SK-III(2006-2008), SK-IV(2008-2018), SK-V(2018-now). In this thesis, I only used the data in first four run periods.

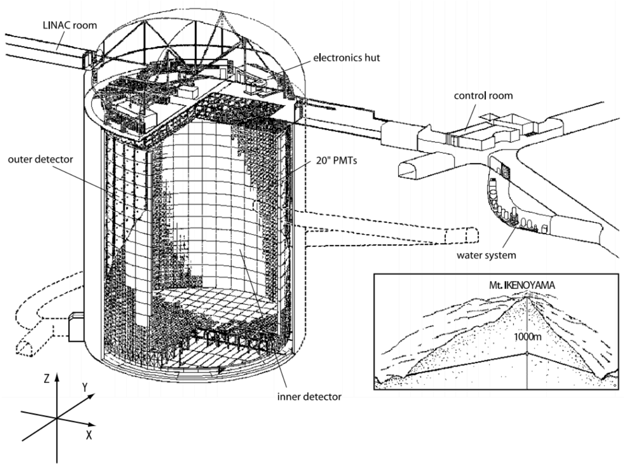


Figure 2.1: Schematic figure of Super-Kamiokande detector.

2.1 Cherenkov Radiation

Cherenkov radiation is produced by charged particles going through a dielectric medium under the condition that the speed of particle is faster than the speed of light in this medium. Hence every type of particles has its own momentum threshold to meet the requirement of Cherenkov radiation in pure water. Here is a table for those charged particles frequently observed in Super-K. All emitted photons are in a cone along the direction of the particle flies:

Particle type	electron	muon	charged pion	proton
Momentum threshold (Mev/c)	0.57	118	156	1051

Table 2.1: Momentum threshold of Cherenkov radiation in pure water. Take from Jiang Miao's thesis(2019)

$$\cos\theta_c = \frac{1}{n\beta} \tag{2.1}$$

where θ_c is the angle between the direction of photons emitted and the direction of the particle. n is the refractive index of the medium and β is the speed of the particle over the speed of light in the vacuum. And the relation between the number of photons created by Cherenkov radiation and the wavelength λ

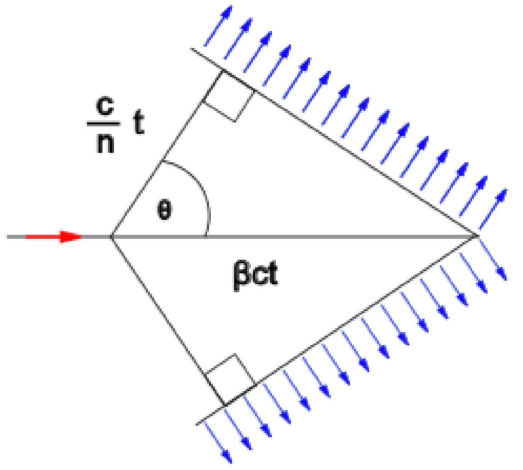


Figure 2.2: Cherenkov radiation. Take from Zepeng Li's thesis(2017)

is: All emitted photons are in a cone along the direction of the particle flies:

$$\frac{d^2N}{dxd\lambda} = \frac{2\pi\alpha}{\lambda^2}(1 - \frac{1}{n^2\beta^2}) \tag{2.2}$$

where the x denotes the path length of the particle and α denotes the fine structure constant. The range of sensitive wavelength in Super-Kamiokande PMTs is 300nm-600nm, 340 photons are produced per centimeter approximately.

2.2 Detector of Super-Kamiokande

The Super-K detector has a large cylindrical water tank which is filled with 50,000 ton of ultra pure water. The tank has a diameter of 39.3 m and a height of 41.4 m. The detector is constitute of two concentric cylindrical regions: the inner detector (ID) and the outer detector (OD). The OD region extends 2.2 m from its top and bottom, and 2.05 m inward the wall of the barrel. There is a

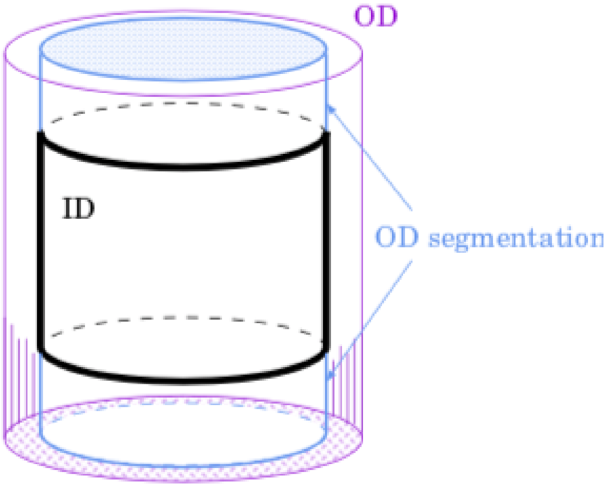


Figure 2.3: schematic view of SK detector

55-cm thick stainless steel support structure which is opaque to light between the ID and the OD. Photo-multipliers are placed on both sides of the stainless steel structures to detect the Cherenkov radiation. The diameter of ID is 33.8 meters and the height of ID is 36.2 meters. And there are 32,000 tons of ultra-pure water in ID. The 20 inch inward-facing PMTs are placed evenly on the inner surface of the ID. The space between the PMTs is separated by black sheets to prevent leaking of light into the OD. During SK-1, there were 11,146 PMTs, which corresponded to approximately 40% of the ID surface covered by the photocathode.

2.3 Photomultiplier Tube

Firstly, the 20-inch PMTs in the ID (Hamamatsu R3600) were developed by Hamamatsu Photonics K.K. and then the dynode structure and the bleeder circuit were changed later to improve timing response and photon-collection efficiency for Super-Kamiokande with the cooperation of Kamiokande collab-

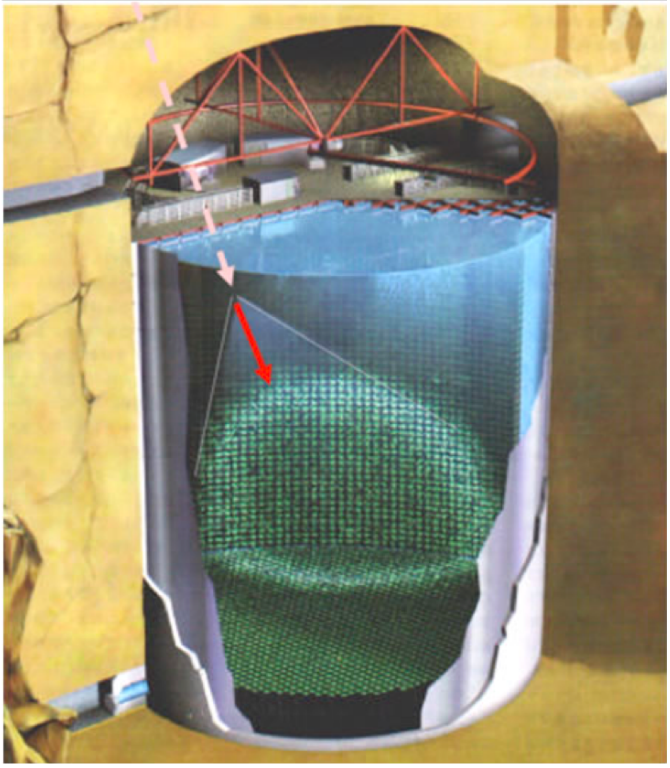


Figure 2.4: Mechansim of Super-Kamiokande detector.

orators.

The PMT photocathode is coated with a bi-alkali (Sb-K-Cs) with high quantum efficiency in the Cherenkov wavelength region (22% at a wavelength of 360-400 nm) and low thermal electron emission. The first dynode has a collection efficiency of approximately 70% and a photocathode uniformity of 7% at various incident locations. The ID PMT's gain is 10^{-7} with the high voltage is approximate 2000 V. A single photoelectron (p.e.) peak is clearly seen in a single p.e signal distribution shown in 2.7. The transit time spread of the single p.e. signal is approximately 2.2 ns, but the average rate of dark hits of 0.25 photoelectrons is approximately 3 kHz. After repair work, the PMT

Figure 2.5: PMT case for 20-inch ID PMT since SK-II period

located in the bottom wall of the ID tank imploded while refilling the SK tank On November 12, 2001. This implosion caused a cascade of PMT implosions that destroyed about 60% of ID and OD PMTs. To prevent such accidents, all internal PMTs have a cover consisting of a thickness of 12 mm clear UV transparent acrylic dome and fiber reinforced plastic (FRP) protected a photocathode and side area. For photons in water with wavelengths greater than 350 nm, the transparency of the acrylic cover layer exceeds 96%.

In the case of OD PMT (Hamamatsu R1408), an 8-inch photocathode was attached to a 60 cm x 60 cm x 1.3 cm wavelength shift plate as described before to improve light collection efficiency by 60%. However, the timing resolution wavelength shift plate of a single p.e. has been reduced from 13 ns to 15 ns, so OD is implemented as a calorimeter and veto counter rather than a particle tracker, higher light collection efficiency plays a more important comparing compared with the timing resolution.

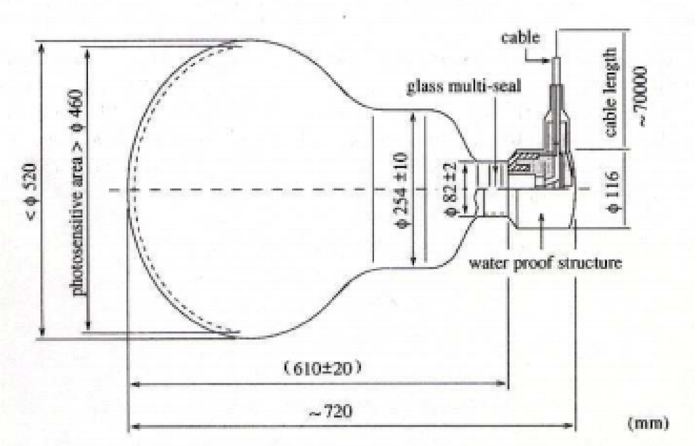


Figure 2.6: schematic view of PMT

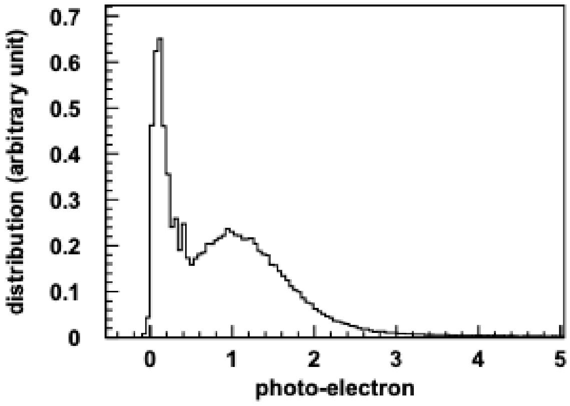


Figure 2.7: Single photoelectron distribution of the 20-inch PMT

shape	Hemispherical
Photocathode area	50 cm diameter
Window material	Bialkali (Sb-K-Cs)
Quantum efficiency	22% at $\lambda = 390$ nm
Dynodes	11 stage Venetian blind type
Gain	107 at 2000 V
Dark current	200 nA at 10^7 gain
Dark pulse rate	3 kHz at 10^7gain
Cathode non-uniformity	$< 10\%$
Anode non-uniformity	$< 40\%$
Transit time	90 nsec at 107 gain
Transit time spread	2.2 nsec (1σ) for 1 p.e. equivalent signals
Weight	13 kg
Pressure tolerance	6 kg/ cm^2 water proof

Table 2.2: Specification of the 20-inch ID PMT for Super-Kamiokande experiment

2.4 Electronics and Data Acquisition System for SK-IV

The front end electronics of Super-K were modified from Analog Timing Module in September 2008 to a dead-time free data acquisition system on the foundation of new electronics named QBEE (QTC-Based Electronics with Ethernet).

The QTC (charge-to-time converter) for PMT signal readout is used in QBEE and its surrounding are shown in Figure 2.8. Each QTC chip has three input channels with a charge dynamic range from 0.2 to 2500 pC. There are three gain ranges for each channel, whose relative gain ratio are 1/49: 1/7 : 1 respectively corresponding to small, medium and large gain. Each QTC channel has a built-in discriminator that can trigger itself and convert the integrated charge of the input signal into a timing signal. The dominant edge of the time signal represents the signal timing and the width represents the charge of the

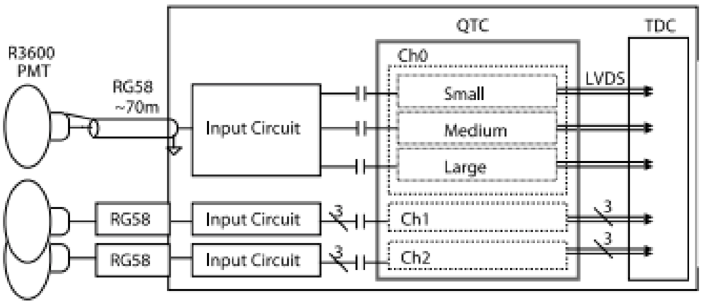


Figure 2.8: Block diagram and surroundings of QTC.

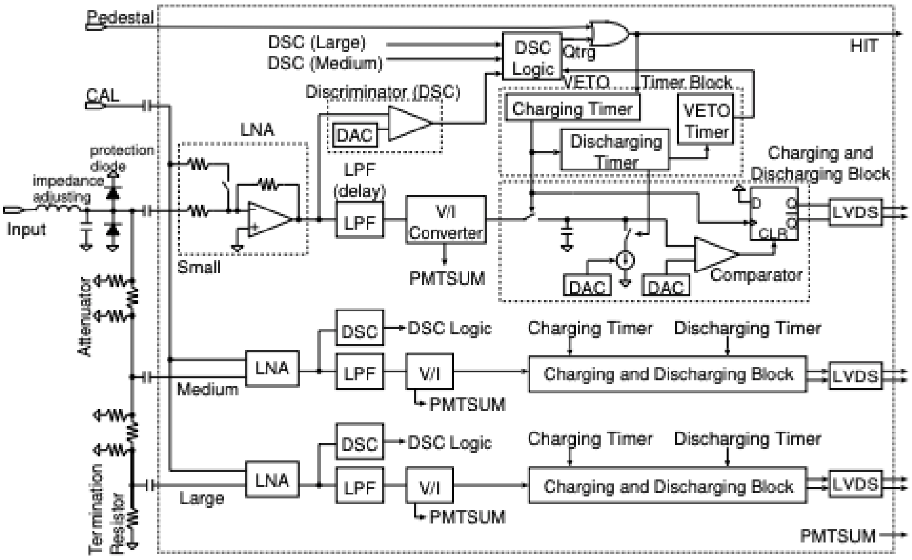


Figure 2.9: Block diagram of one QTC channel. There are three gain ranges for each channel

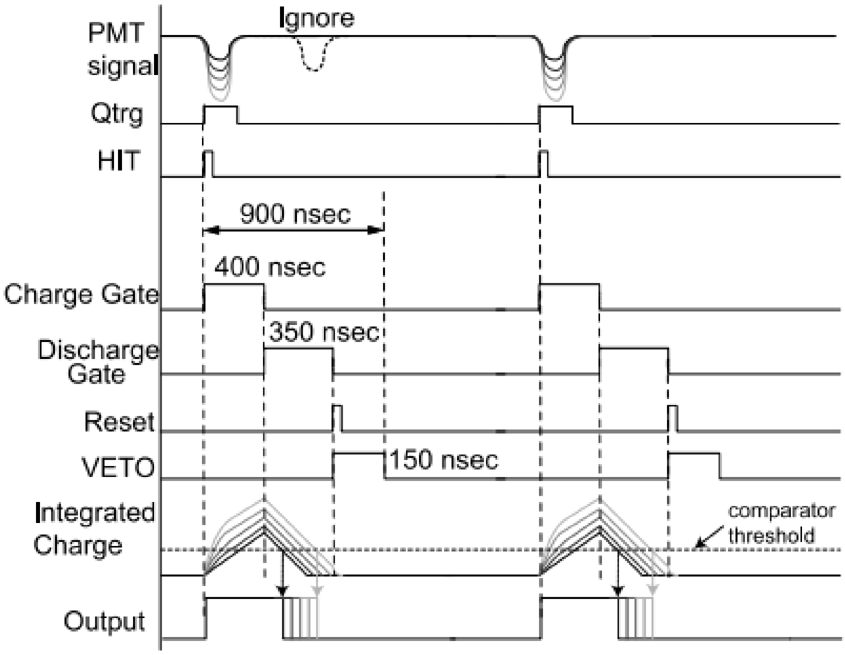


Figure 2.10: Timing chart for QTC operation.

input signal. Figure 2.9 shows a block diagram of a channel. The input signal from the PMT is amplified by a low noise amplifier (LNA) firstly and then delayed by a low pass filter (LPF). After being processed by the voltage-to-current converter (V/I), the signals are eventually integrated by the capacitor. All of the charge stored on the capacitor is monitored by an output signal named PMTSUM. In addition to the time information obtained by measuring the dominant edge of the input signal from the PMT, the QTC is also sensitive to charge. The input signal integration begins when the amplified input signal exceeds the threshold specified by the discriminator. At the same time, a signal is generated to record the trigger hit. When the integration of the charge controlled by the timer block ends, the charge stored in the capacitor begins to discharge at a constant current. The discharge time is also the width of the output signal and is proportional to the integrated charge of the input signal.

Therefore, both timing information and charge information are encoded into the timing output signal. There are three timers for charging, discharging and VETO in the timer block. The length of time of the gate is controlled by a digital-to-analog converters (DAC) by adjusting the discharge current and the comparator threshold.

A timer that is charged after being triggered by the discriminator output signal opens the charge gate lasting for 400ns. The charging gate is closed by the switch between the charging capacitor and the V/I converter, and the input signal is stored in the capacitor. After the charge gate, the discharge timer opens the discharge gate for about 350 ns. After the switch between the capacitor and the discharge current source closed, the input signal is ignored by turning on the switch near the V/I converter. The time for the integrated signal voltage t dropping to the comparator threshold level is indicated by the trailing edge of the output signal from the QTC. The output signal from QTC is positively relevant to the charge of the input signal. Reset and VETO signals are issued when the the discharge gate to reinitialize other QTC circuits ends. The total time spent processing one input signal is about 900 ns.

Hardware triggers are used for SK-I, SK-II and SK-III trigger systems, while software triggers are used new DAQ readout systems in SK-IV rather than recording and analyzing all hits. Seventeen microseconds worth of PMT data is read out at a frequency of 60 kHz and particle interactions are identified via using the software trigger therein, which lead to a dead-time free system in SK-IV and the decay electron tagging efficiency is therefore improved.

Additionally, the modified electronics also has a better linearity and a wider dynamic range, which improves the precision of high energy event reconstruction.

2.5 Background Reduction

2.5.1 Water Purification

It is important to let the water as pure as possible because the light collected by the PMTs has to travel through the water. Water taken from the mine is first passed through a 1 μm filter to prevent large particles. It then passed through a heat exchanger which makes the water cool to about 13°C which is approximate the limit temperature in which the growth of bacteria hardly happens. And then a cartridge polisher clears away heavy ions, and a UV sterilizer gets rid of the remaining bacteria. The water then flows through a reverse osmosis and vacuum degasser system that filter radon and other dissolved gas. In the final stage ultra filters kills particles larger than about 10 nm in diameter. The water is then returned to the tank. Water in the tank circulates through the system at approximate 60 tons/hour.

2.5.2 Air Purification System

The air in the mine housing Super-K detector has a high concentration (10–5000 Bq/m^3) of radon. Fresh air is unstopped pumped into the control room and cavern area of the SK experiment. Radon concentrations stays about 20-30 Bq/m^3 inside the top of the Super-K tank and lower than 100 Bq/m^3 at the cavern.

2.6 Calibration

2.6.1 Relative Gain Calibration

In order to uniformly charge all PMTs in the detector, the high voltage of each PMT is set one by one. Relative gain calibration is performed by scintillating the laser using a device similar to timing calibration measuring and correcting the remaining individual variations in PMT gain. A pulsed laser near the ID center repeats isotropic flicker at two different intensities: high intensity and low intensity. For high intensities, each PMT detects the appropriate number of photons and creates an average charge Q (i) for each ID PMT i. For the other case, the laser is flashed at a low intensity, only a few PMTs get a single p.e. hit at a time. The number of hits observed by i-th PMT is counted and recorded as N(i). The complicating factors in measuring those two intensities Q(i) and N(i) are closed due to the same position of the light source:

$$\begin{aligned} Q(i) &\propto I_H \times a(i) \times \epsilon_{qe}(i) \times G(i) \\ N(i) &\propto I_L \times a(i)_{qe}(i) \end{aligned} \tag{2.3}$$

where $I_H(I_L)$ denotes the average intensities of the high(low) intensity flashes. The value of a(i), ϵ_{qe} and G(i) denote the acceptance, quantum efficiency and gain of i-th ID PMT respectively. The relative gain of each PMT G(i) can be obtained by taking the ratio of Q(i) over N(i) as:

$$G(i) \propto \frac{Q(i)}{N(i)} \tag{2.4}$$

The standard deviation of all PMT gains is 5.9%. The respective relative gain factors gotten above are used to correct the coefficients and convert the output charge into the number of photoelectrons observed at each PMT.

2.6.2 Absolute Gain Calibration

The charge recorded by a PMT in pico Coulomb (pC) is converted into the number of incident photoelectrons via using the absolute gain. The absolute gain was estimated by observing the charge distribution of single p.e. signals with a spherical gamma-emitting nickel source, which emits 9 MeV gamma rays isotropically by capturing neutrons emitted from a 252Cf source in it. The nickel source was put near the center of the ID tank and triggered 0.004 p.e/event for each PMT, which makes more than 99% hits single p.e. hits.

Figure 2.11 shows the observed charge distribution for the single p.e. signals. All hits from the PMT are corrected for relative gain changes and accumulated. The conversion factor from the observed charge to the number of p.e.s has a value of 2.658 pC/p.e, which is the peak of the single p.e. distribution. Furthermore, Figure 2.12 shows the the peak of charge distribution for the dark hits, which is proportional to the absolute gain of PMT. And the peak increases with time passing and the increasing rate is relevant to the production year of PMT. The reason of such an increase of the gain is still under investigation, while a correction is applied on the gain to reduce the effect.

2.6.3 Quantum Efficiency Calibration

Including the gain, the relative difference of quantum efficiency (Q.E.) is also estimated for each PMT because it effects the charge response with a small number of incident photons. The measurement uses the same nickel source method as the absolute gain calibration.

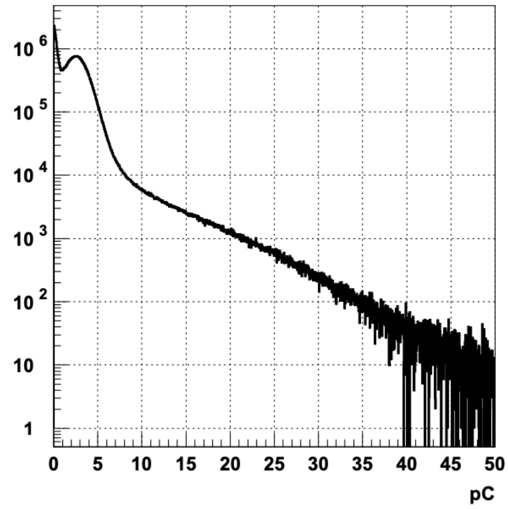


Figure 2.11: Distribution of the observed charge for single p.e. signals obtained from the calibration data with nickel source

The average number of hits on a PMT with a low intensity of incident light is proportional to the quantum efficiency $\epsilon_{qe}(i)$. The relative Q.E. for each PMT $\epsilon_{qe}(i)$ is obtained by comparing the simulated events, which contribute into the acceptance factor a(i) excluding the variations in $\epsilon_{qe}(i)$ and the real nickel source data. The detector simulation uses the resulting q.e. for each PMT.

2.6.4 Relative Timing Calibration

The time response is important for reconstructing the event vertices and track directions precisely. It changes among PMTs because of many factors such as the difference on length of cable, the process time of electronics etc. Moreover, the time response is also relevant to the observed charge, since hits with less charge exceed threshold of TDC discriminator later than those with larger charge. This is time-walk effect. In order to make a correction table for the time-walk effect for each PMT in the detector with taking the overall process time into consideration, the timing calibration is necessary.

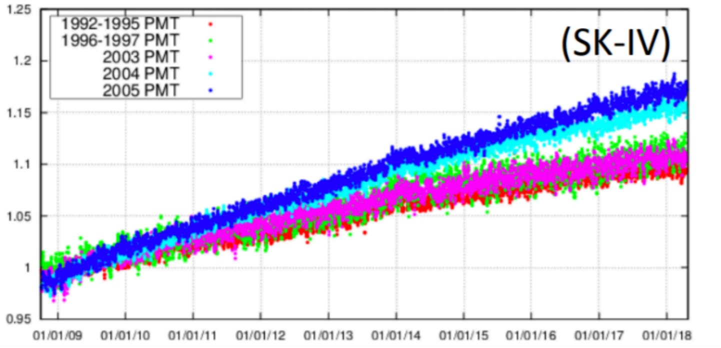


Figure 2.12: Time variation of the peak of the charge distribution for the PMT dark hits. Different color denote the different production year of PMT. A correction on PMT gain is applied based on this measurement.

A schematic diagram of the timing calibration system is as follows: A FWHM 0.4 ns fast pulse with a wavelength of 337 nm is generated by a nitrogen laser and monitored by a 2 inch PMT with fast trigger response. The laser light pulse is changed to 398 nm through a dye containing all the response in which Cherenkov emission spectrum is included, and light absorption in water and the quantum efficiency of the PMT almost becomes maximized. The light intensity can be varied by adjusting the filter, which allows you to measure the time response at different pulse heights. The TQ distributions are gotten separately for each PMT, and then is determined TQ map which is a polynomial as a function of charge to extract the calibration constants. It is ensured that the PMT time response across the detector is uniforming.

2.6.5 Water Property Calibration

The photon propagation in water is modeled empirically with the light scattering and absorption taken into account. The attenuation of light in water

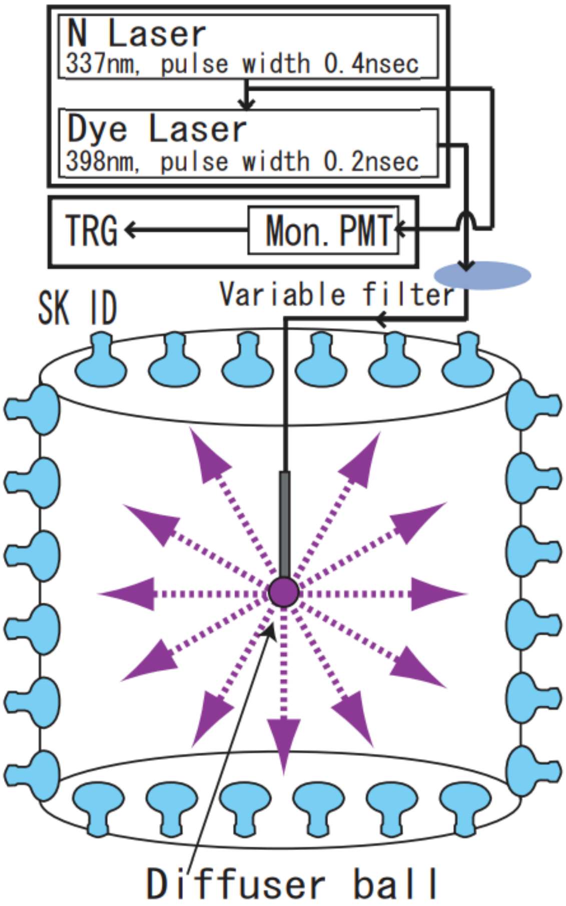


Figure 2.13: Schematic view of the timing measurement system using a laser.

can be represented as exp(− l/L(λ)), where l is the distance the light traveled and L(λ) denotes the attenuation length as a function of wavelength λ.

In Super-K simulation, L(λ) is defined as

$$L(\lambda) = \frac{1}{\alpha_{abs}(\lambda) + \alpha_{asym}(\lambda) + \alpha_{sym}(\lambda)} \tag{2.5}$$

where α_{abs}, α_{asym}, α_{sym} are coefficients for absorbing, asymmetric scattering and symmetric scattering respectively. The asymmetric term α_{asym} accounts for forward Mie scattering, while the effect of symmetric scattering α_{sym} takes the Rayleigh scattering and symmetric Mie scattering into consideration and

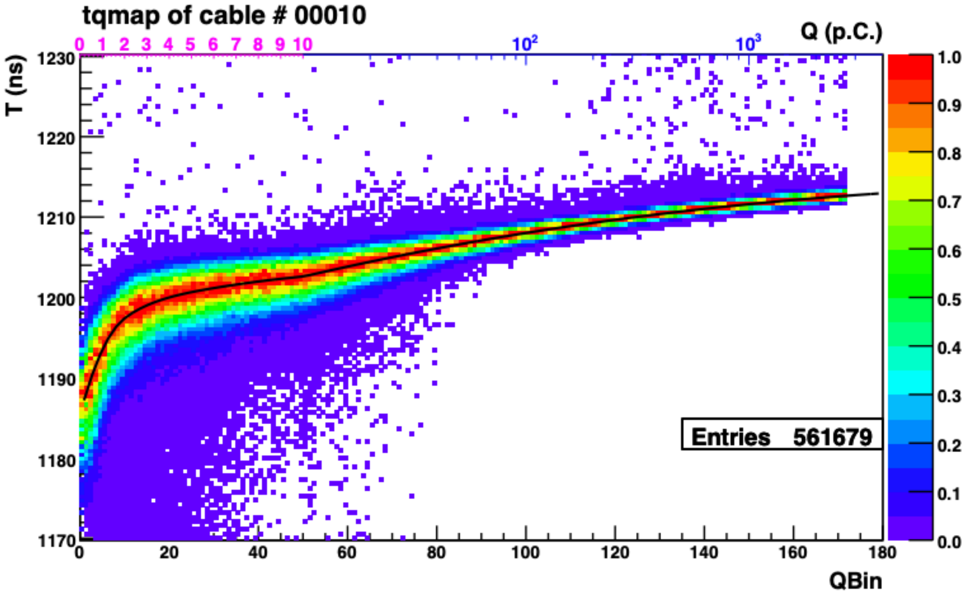


Figure 2.14: A typical TQ distribution for ID PMT

can be expressed by $1+cos^2\theta$, where θ denotes the direction of the scattered photon. The laser injector system to measure the water constants is shown in Figure 2.15. An adjustable wavelength collimated laser beam is injected vertically downward into the top of the Super-K can. Scattered and reflected light is detected by the PMT in the five partitions belonging to the barrel region marked from B1 to B5 and the top wall of the detector. The hit time distribution after subtracting the PMT flight time for each detector area is displayed in Figure 2.16. The sharp peaks on the right between 1830 to 1900 ns represent photons reflected by the PMTs and black sheets on the bottom wall, while those earlier hits are because of the photon scattered in water. And the water calibration constants, α_{abs}, α_{asym}, α_{sym} are tuned to achieve the best agreement between simulation and data as shown in Figure 2.17.

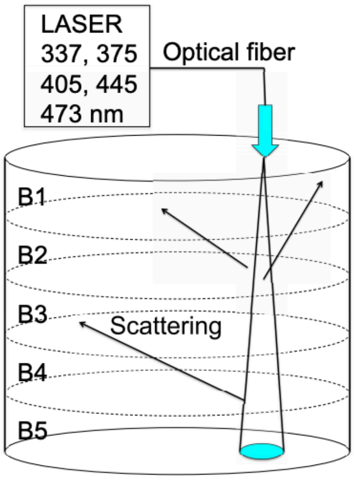


Figure 2.15: The laser injector system for the water property measurement

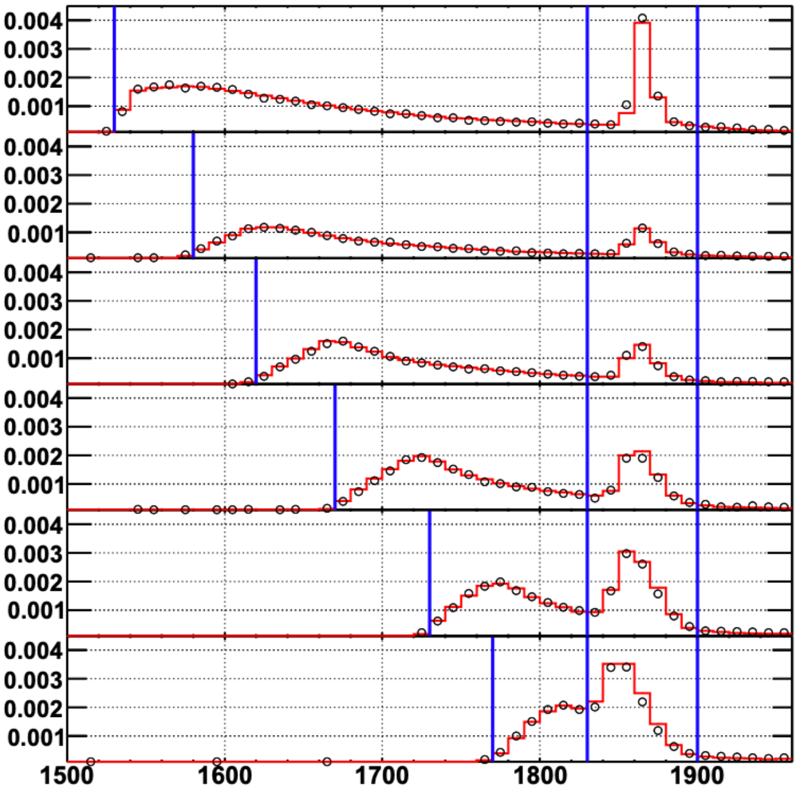


Figure 2.16: PMT hit time distribution in different detector regions for data (open circle) and Monte Carlo (red line) after tuning by scattering and absorption parameters

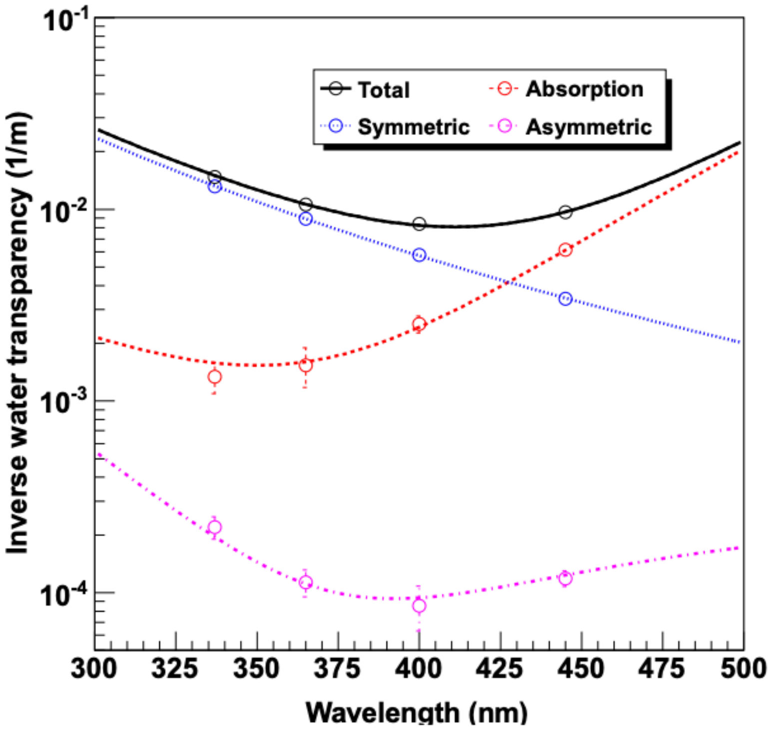


Figure 2.17: An example of water parameter fitting. The data used here is taken in Apr 2009. Points represent the average value of coefficients

Chapter 3

Reconstruction Algorithm

Neutrinos themselves can not directly be seen by our detector, while they rarely interact with nucleons or electron in the Super-K tank and produce charged particles and emits Cherenkov light in water. And Cherenkov light travels in water and is detected by PMTs. Original PMT hit information such as amount of photons and arrival time is translated into physical properties or quantities of events, which is called events reconstruction. What's more, the algorithm which is used to do these things is named as reconstruction algorithm. And there are two different types of algorithms: APFit and fiTQun.

3.1 APFit

APFit determines the physics quantities step by step. There are four major steps for reconstruction:

- Event vertex

  We assume that all rings are generated at one event vertex and we search for a point in which the time distribution becomes sharpest after time of flight(TOF) subtraction.

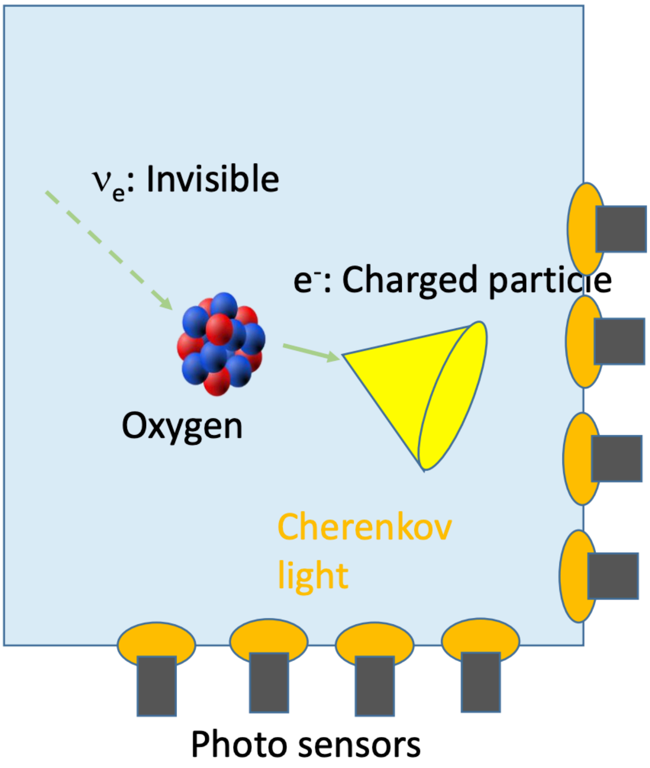


Figure 3.1: Schematic figure of how neutrinos are detected

- Count number of rings

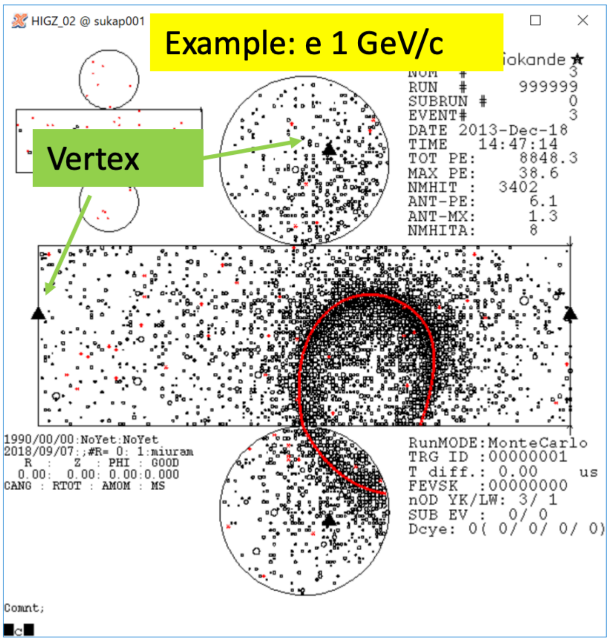
  We assume that spherical coordinate locates around the vertex. And then we draw 42° from hitted PMTS. After that, the direction of Cherenkov ring could be identified as intersection of these virtual circles. Finally, we apply a likelihood method to determine how many ring there are.

- Particle type

  The PID could be identified by hit pattern and opening angle: electron has fuzzy edge due to EM shower while muon has sharp edge and smaller opening angle in low momentum.

- momentum of each ring

  By taking observed charge in 70° cone and [-50 ns,+250 ns] event time range into consideration, we could reconstruct momentum of each ring.



```
NUM  #          3
RUN  #     999999
SUBRUN #        0
EVENT#          3
DATE 2013-Dec-18
TIME   14:47:14
TOT PE:    8848.3
MAX PE:    38.6
NMHIT :  3402
ANT-PE:     6.1
ANT-MX:     1.3
NMHITA:     8
```

```
1990/00/00:NoYet:NoYet
2018/09/07:;#R= 0: 1:miuram
   R  :   Z  : PHI : GOOD
  0.00:  0.00: 0.00:0.000
CANG : RTOT : AMOM : MS
```

```
RunMODE:MonteCarlo
TRG ID :00000001
T diff.: 0.00     us
FEVSK  :00000000
nOD YK/LW: 3/ 1
SUB EV :   0/ 0
Dcye: 0( 0/ 0/ 0/ 0)
```

```
Comnt;
```

Figure 3.2: Vertex construction of APFit

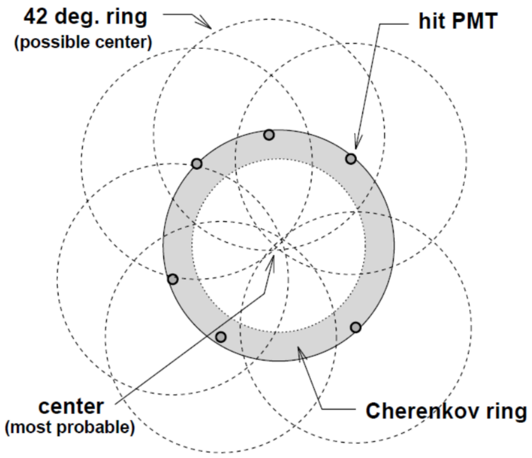


Figure 3.3: ring counting construction of APFit

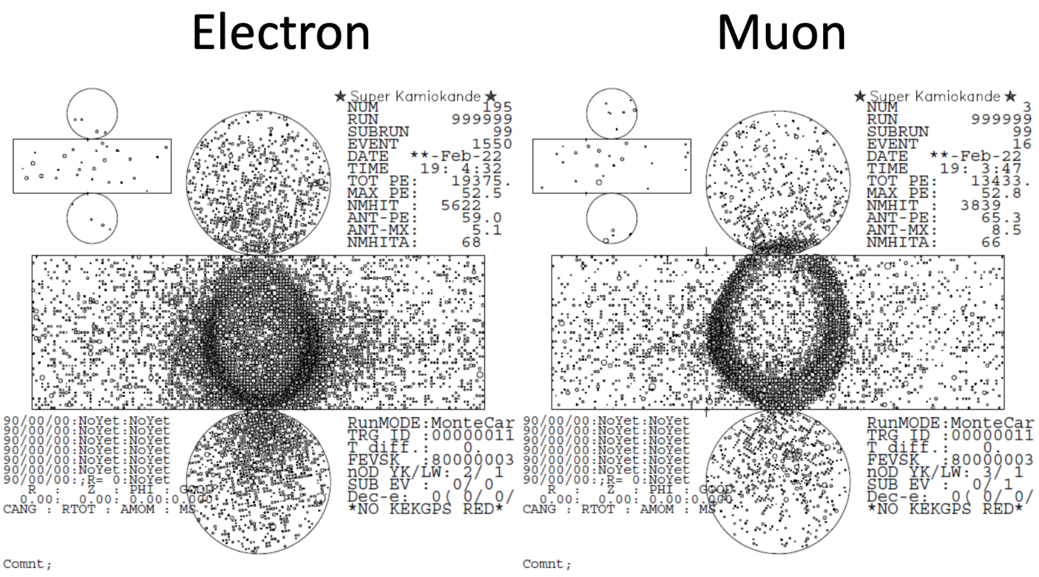
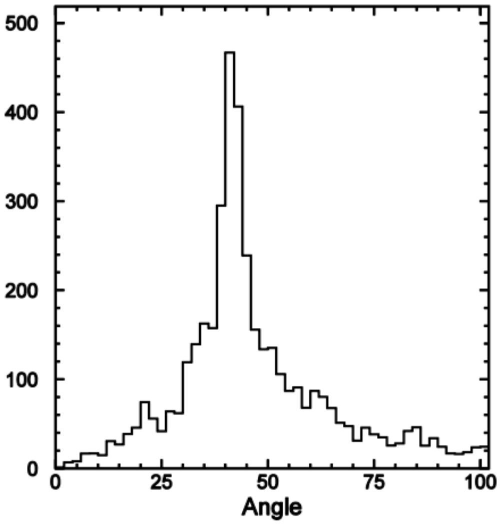
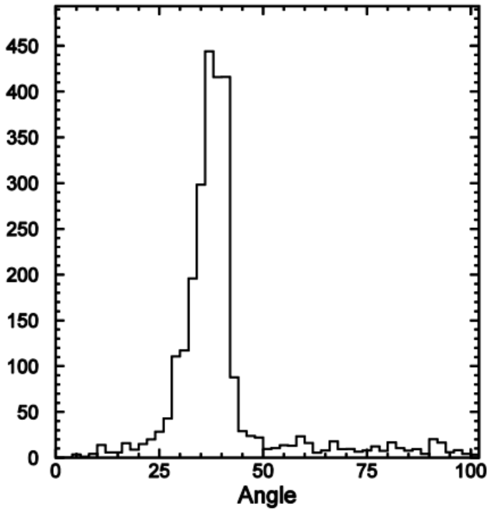


Figure 3.4: ring counting construction of APFit



(a) opening angle of electrons



(b) opening angle of muons

Corrections and adjustments would be applied if two rings were overlapped.

It is worth noting that in APFit the hit time information is only used in the first step to find the initial vertex candidate and not used in the following steps such as particle identification or ring counting.

3.2 fiTQun

A new algorithm called fiTQun uses maximum likelihood to reconstruct particle types while determining the detector kinematics. This algorithm is based on experiments developed for MiniBooNE. As we have seen so far-with additional features like rebuilding multi-ring events. More information such as PMT hit in the suburbs of Cherenkov and taken over to fiTQun compared to APFit. FiTQun's fitting process runs multiple times, finding the best kinetic parameters (vertex, momentum, direction etc) for each possible particle composition hypothesis (particle type etc) and finally getting the best Choose. Since Super-Kamiokande starts to detect in 1996, the remarkable development of computing power has made it possible to fit and determine every single movement parameter or particle configuration. Currently, fiTQun is used in T2k analyses to extend the fiducial volume from the that of APFit. We only used the fiTQun with SK4 since the systematic analysis for SK1-SK3 are not finished yet.

3.3 Likelihood Function

Taking An event topology hypothesis Γ (e.g. single-ring μ-like) into consideration in the fitting along with its own associated kinematic parameters θ, such as vertex positions, particle generation times, particle directions and their momenta. The likelihood function which is used to estimate the kinematic variables in fiTQun for a given hypothesis is on the foundation of the observed charge and hit time of each PMT is defined as:

$$L(\Gamma,\theta) = \prod_{j}^{unhit} P_j(unhit|\Gamma,\theta) \prod_{i}^{hit} 1 - P_i(unhit|\Gamma,\theta) f_q(q_i|\Gamma,\theta) f_t(t_i|\Gamma,\theta) \tag{3.1}$$

In this expression, index j runs on all PMTs (also called unhit PMTs) that are not hitten. For each unhit PMT, the probability of not registering a hit with a fitting hypothesis (Γ, θ) is obtained as $P_j(unhit|\Gamma,\theta)$. Index i represents the PMT registering a hit. The charge likelihood $fq(q_i|\Gamma,\theta)$ indicates the likelihood density of q_i for these PMTs, under the hypothesis (Γ,θ). Similarly, the likelihood density which generates a hit at the observed time t_i is defined as $f_t(t_i|\Gamma,\theta)$.

In fact, the particle and photon propagation processes in the detector are independent from PMT and electronic response. Thus, given the hypothesis $\mu_i(\Gamma,\theta)$, the charge probability can be overwritten with respect to the expected number of photoelectrons (the predicted charge) produced by the ith PMT. Likelihood function is

$$L(\Gamma,\theta) = \prod_{j}^{unhit} P_j(unhit|\mu_j) \prod_{i}^{hit} 1 - P_i(unhit|\mu_j) f_q(q_i|\mu_j) f_t(t_i|\Gamma,\theta) \tag{3.2}$$

Here, the unhit probability $P_j(unhit|\mu_j)$ and the charge likelihood density

$f_q(q_i|\mu_i)$ are just characteristics of PMT. The processes of Cherenkov photon emission and propagation in water is taken into consideration when calculating the predicted charge $\mu_i(\Gamma, \theta)$ and is decoupled from the detector electronics.

In the process of calculating the predicted chargeμ_i, the observed photons fall into two categories: One is produced by the Cherenkov process and directly impacts the PMT through water in the detector, also known as unscattered light or direct light. The other, called indirect light, is a photon that is scattered or reflected before being observed by the PMT. The two types of photons are considered separately and summed to build the final predicted charge.

The predicted charge of the direct light reaching the PMT is obtained by integrating the Cherenkov radiation distribution along the trajectory with taking the light transmission from water, the distance between the light source and the PMT, and the PMT angular acceptance into account. The charge is generated by indirect light arriving at the PMT is calculated by integrating the product of the direct light emission profile and a scattering function. This function is generated in advance in consideration of the effects of the position of the PMT, the light source and the wall. For multi-ring case, the predicted charge of each ring is first calculated respectively and then summed to calculate the total expected charge. The final charge likelihood $fq(q_i|\mu_i)$ is obtained by comparing the photoelectron generation prediction of the Poisson distribution with the charge observed in the PMT.

The time likelihood term is written as $f_t(t_i|t_i^{exp}, \Gamma, p, \mu_i)$, where p is the particle momentum under the topology hypothesis Γ, and t_i^{exp} represents the expected hit time, which is defined as the arrival time of unscattered photons emitted from the midpoint of track and arriving at the PMT directly as

$$t_i^{exp} = t + s_{mid}/c + |\mathbf{R}_i^{PMT} - \mathbf{x} - s_{mid}\mathbf{d}|/c_n \tag{3.3}$$

Where t and x are the creation time and vertex of the particle, d is the direction of the particle, $RPMT_i$ is the position of the i-th PMT, and s_{mid} is half the length of the track. The parameters c and c_n are the group velocities of the Cherenkov light in vacuum and water, respectively. The likelihood of time depends on the predicted charge because the first photon arriving at the PMT is recorded as a hit time, making the time likelihood density distribution of more incident photons narrower, i.e. more predicted charge. It will be brought. The track length s of the particles determined by the topology Γ and the momentum p also affects the shape of the likelihood density, since not all photons are produced in the middle of the track. The shape with the time likelihood of direct and indirect photon hits is determined based on particle gun simulation. During the fitting process, the contributions of these two likelihoods are calculated separately in the same way as the charge likelihood. The final time likelihood of a single particle is gotten by combining these two likelihoods according to their relative intensities. In the case of a multi-particle topology hypothesis, the time likelihood is calculated ring-by-ring and then merged into a final likelihood function under the assumption that the photons from a particle with earlier t_i^{exp} always arrive earlier than the photons from any other particles whose t_i^{exp} values are later.

The best set of the kinematic parameters $\hat{\theta}$ for a given event topology hypothesis Γ is obtained by maximizing likelihood value $L(\theta|\Gamma)$. The best estimation of a given event is determined by comparison of $L(\Gamma|\hat{\theta})$ among all hypotheses, Γ. It worth noting that the best estimation is not always the hypothesis with the largest likelihood value. A hypothesis with more rings, i.e. more parameters, always fits the observation better and therefore has a higher likelihood value due to the increased number of degrees of freedom. Therefore care is needed when determining the best hypothesis. The detail of the topology selection, including ring counting and particle identification, will be introduced in next section.

3.4 Fitting Procedure

The fiTQun reconstruction process can be divided into four steps.

- Vertex pre-fitting

  Vertex Pre-fitter is a fast algorithm that estimates the initial vertex position using only hit time information of the PMT around the primary event trigger. In this step, it is assumed that all observed light is emitted from a single location. The estimation is done by searching time t and vertex position x under which the goodness function is maximized.

  The vertex pre-fitter is a fast algorithm which estimates an initial vertex position with only the hit time information from PMTs around the primary event trigger. It is assumed that all observed light is emitted from a single position in this step. The estimation is done by searching for time t and the vertex position x which maximize the goodness function:

$$G(\mathbf{x},t) = \sum_{i}^{hit} e^{(-(T_i^{res}/\sigma)^2/2)} \qquad where \qquad T_i^{res} = t_i - t - |\mathbf{R}_i^{PMT} - \mathbf{x}|/c_n \tag{3.4}$$

is the residual hit time calculated assuming a point-like light source at the interaction vertex and accounting for the photon ' s time-of-flight. The position of the i-th PMT is denoted by $\mathbf{R}_i^{PMT}$. As the time and vertex approach their true values during a grid search process, the T_i^{res} distribute near zero, resulting in a large goodness value. The vertex returned during this step is just a rough estimation, called the pre-fit vertex. This vertex will be fitted again with higher precision during the minimization of $L(\Gamma,\theta)$ in Step single-ring reconstruction and Step multi-ring reconstruction.

- Hit clustering

  The Super-K event is defined by the detector activity within the $O(10\mu s)$time window centered on the event trigger. A plenty of sub-events representing clusters that are temporally distant from the primary triggered PMT hit may be included in one event. Muon decay is such an example. The primary event trigger is generated by a muon, which is generated by an extra hit of its decaying electrons. In order to separate the hits from the particles created at different times for further fitting, the main trigger-centric activities and other sub-events (if any) are searched and selected by the hit clustering algorithm.

The algorithm first searches for sub-event activity near the event trigger

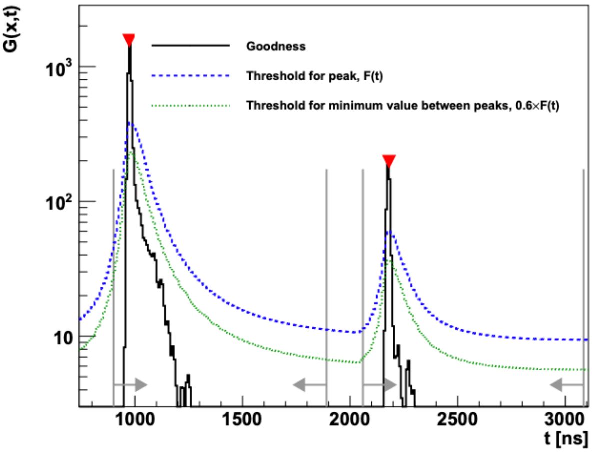


Figure 3.6: An example of the goodness function with respect to hit time

time by the peak-finding algorithm. The goodness of vertices G(x, t) from Equation 24 is scanned at t while vertex x is fixed to the pre-fit vertex for additional delay peak searching. Figure 4.6 shows a good distribution of muon decay. Take an event as an example. The parent mu and its Michel electron are the two main peaks. Scattered or reflected light can result in false peaks in or near the true hit activity cluster. Thus, the peak must exceed the minimum threshold F(t). The definition of F(t) is

$$F(t) := 0.25 \arg\max_{i \in M} \frac{G(\mathbf{x}, t_i)}{(1 + ((t - t_i)/\gamma)^2)} + \eta. \tag{3.5}$$

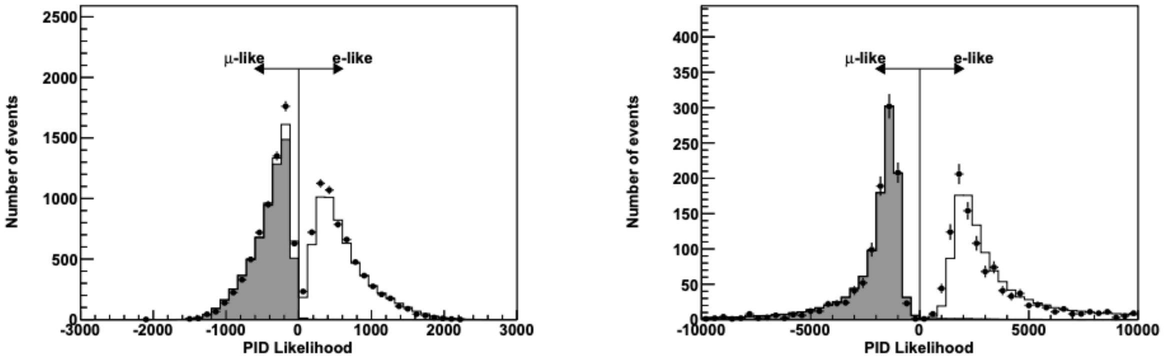
where M represents all local maxima of the goodness function G(x, t). The time constant γ is assigned as 25 ns (70 ns) when $t < t_i (t > t_i)$. In order to suppress hits from dark noise of the PMTs, the threshold function is offset by $\eta = 9$. The minimum goodness between any two peaks is required to be below 0.6 × F(t), which is shown as green dashed curve. The figure shows only two peaks that pass these criteria (marked by red triangles) which will be subject to further fitting.

In the first step of hit clustering, all vertex positions are considered to be close to pre-fit vertex. while his assumption is broken when the primary particle(e.g large momentum muons) travels a long distance from the interaction vertices. In this case, the distribution of goodness may be blurred and may be below the threshold, resulting in loss of sub-events and reduced attenuation of electronic marking efficiency. In order to improve the marking efficiency, vertex pre-fitting runs again after masking the hit caused by the primary particles to the vertices of the vertices close to the secondary particles. The peak-finding algorithm then uses the new vertex as the location of the Goodness function x to search for missing goodness peaks.

For each found peak, its associated hits are contained by a time window defined as $-180ns < T_i^{res} < 800ns$. Those hits will be input to the vertex pre-fitter and peak finder once again to get the final sub-event candidates with a higher precision for the full event reconstruction. The interaction time and vertex fitted in each window will be used as the seeds for fitting in following steps.

- Single-ring reconstruction

  A single-ring filter, whose topology hypothesis Γ in Equation 4.2 is single-particle, is first applied to each time window determined in the previous



(a) Sub-GeV events

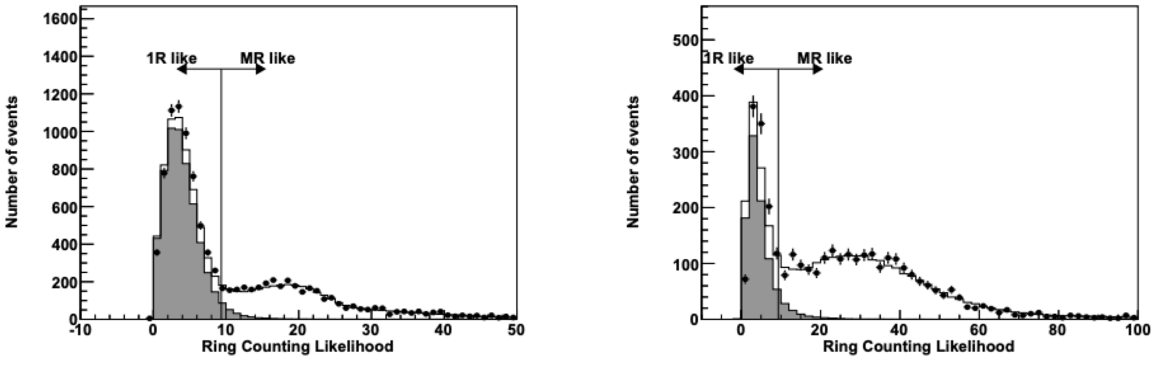
(b) Multi-GeV events

Figure 3.7: PID likelihood distribution of FC single-ring events

step. By changing the kinematics parameters of the event (including created vertex and time, particle momentum and direction), it is possible for the likelihood function to maximize the distribution of charge and time observed in the PMT. For each hypothesis considered by the single ring fit, the process runs iteratively with single-ring fitter: electron, muon, and charge-pion. Particle identification (PID) is then determined by comparing their best fit likelihood values. For instance, the difference between the electron and muon of an event is based on the ratio of the logarithm of likelihood value to the optimal electron and the muon hypothesis: $ln(L_e/L_\mu)$.

- Multi-ring reconstruction

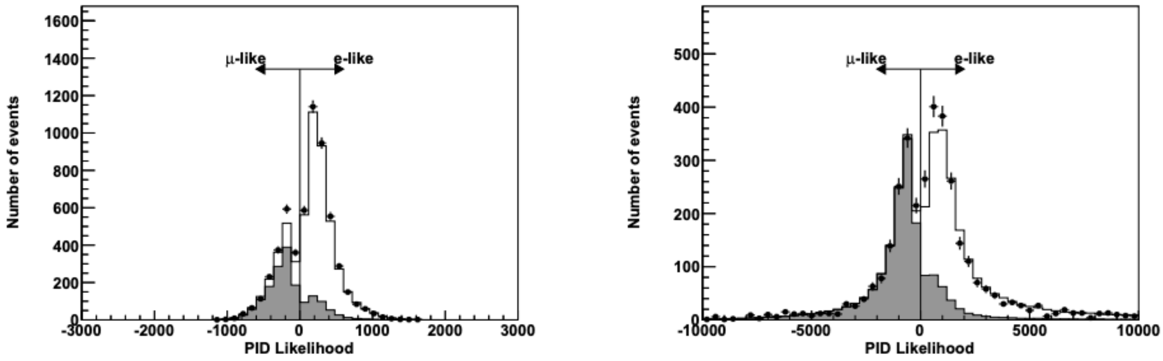
  Since most atmospheric neutrino events with multiple GeV energies have high sensitivity to mass hierarchy and multi-particle final states, the reconstruction performance of light-producing particles is critical for the



(a) Sub-GeV events

(b) Multi-GeV events

Figure 3.8: The distribution of the likelihood ratio between the best-fit result of single-ring hypothesis and multi-ring hypothesis of FC atmospheric neutrino events



(a) Sub-GeV events

(b) Multi-GeV events

Figure 3.9: PID likelihood distributions of the most energetic ring for fully contained multi-ring events

analysis of atmospheric neutrino oscillations . In fiTQun, the multi-ring fitter is only used in the first time window. This reduces the computation time because it is just approximately equivalent to the main event trigger and not any delay time. The multi-ring hypothesis for fitting is constructed by repeatedly adding a new ring in the previous fitting results. After adding a new ring, the kinematic parameters of the likelihood function are changed to find their best fit. Check all three hypotheses that $e-like$, $\mu-like$ and $\pi^{+}-$ are used in single-ring reconstruction. The new ring is verified based on the likelihood ratio of the updated fit result to the original result of each hypothesis, and the new ring is accepted only when the likelihood ratio passes the ring count criterion. The process of searching for a new ring is repeated until the hypothesis of the new ring is rejected by the standard or until there are six rings in the hypothesis. The likelihood ratio between the 2-ring and 1-ring hypotheses of the sub-GeV and multi-GeV events. This likelihood ratio, determined based on the MC study at 9.35 (11.83), reduces the single-ring and multi-ring separations used for the $e-like(\mu-like)$ rings hypothesis. In addition, according to the MC study, ring candidates with other angular distances less than 20° are usually due to particle scattering rather than new particles, so these candidates are false and will be destroyed. When this happens, create a new topology hypothesis with a ring at the original two closed rings' positions and group them together. Merging is a modification of all particle hypotheses, keeping all other rings fixed. This process of ring merging and rearranging is performed in descending order of energy for all ring candidates of the event.

3.5 Comparison

In this section, we will compare these two different types of reconstruction algorithm.

3.5.1 Single-ring reconstruction

The vertex resolution of fiTQun for CCQE ν_e events is stable at 20.6 cm with the visible energy in the range from 100 MeV to 1330 MeV, while the vertex resolution of APFit varies from 34.6 cm to 25.3 cm in the same energy range shown in the Fig 3.11. The vertex resolution of fiTQun for CCQE ν_μ events, which varies from 29.2 cm to 15.9 cm, is better than that of APFit, which is in range from 34.2 cm to 18.2 cm. And The performance of fiTQun for direction reconstruction is basically the same as that of APFit shown in the Fig.4.13. The momentum resolution of fiTQun for CCQE ν_e events gets better from 5.39% to 2.58% as the visible energy increases, while the resolution of APFit is in the range from 7.04% to 3.32% revealed in the Fig 3.15. The momentum resolution for CCQE ν_μ events is stable across the energy range, being lower than 2.5% for fiTQun and a little bit worse for APFit. In addition, fiTQun also shows a better ability to identify between electrons and muons with less than a 1% mis-identification rate in the same energy range compared to APFit revealed in Fig 3.16.

3.5.2 Multi-ring reconstruction

FiTQun shows a better performance on reconstructing multi-ring events than APFit and the better ability to correctly identify single-ring topologies is at the same level for both algorithms. The fake ring ratio of fiTQun is a little bit

Type	fiTQun	APFit
True CCQE ν_e sample		
Vertex Resolution	20.6cm	24.9cm
Direction Resolution	1.48°	1.68°
Momentum Resolution	2.90%	3.56%
Momentum Bias	0.43%	0.63%
Mis-PID rate	0.02%	0.50%
True CCQE ν_μ sample		
Vertex Resolution	15.8cm	17.3cm
Direction Resolution	1.00°	1.28°
Momentum Resolution	2.26%	2.60%
Momentum Bias	-0.18%	0.54%
Mis-PID rate	0.05%	0.91%

Table 3.1: Visible energy is 1Gev here

larger than that of APFit for true single-ring or two-ring event reconstruction, but the increase of fake ring ratio is not large enough to neutralize the benefit from the higher fraction of the events reconstructed correctly.

Type	fiTQun			APFit		
True Number of Rings	1 ring	2 rings	≥3 rings	1 ring	2 rings	≥3 rings
True 1 ring	95.0%	4.64%	0.41%	95.9%	3.85%	0.29%
True 2 rings	27.8%	66.7%	5.56%	42.5%	52.8%	4.63%
True ≥ 3 rings	7.04%	25.5%	67.5%	20.2%	33.0%	46.8%

Table 3.2: a ring whose final state particles with energy 30 MeV higher than the Cherenkov threshold is defined as true rings

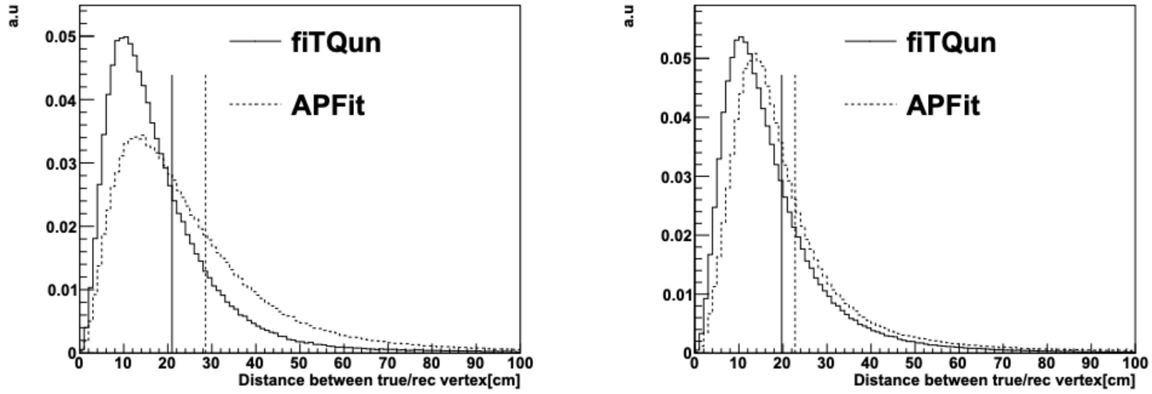


Figure 3.10: Single-ring electron(left) and muon(right) vertex resolution

Type	fiTQun		APFit	
	Event Rate	Purity	Event Rate	Purity
eee events(target: $\nu_\mu CC1\pi^0$)				
Sub-GeV	278.5	52.1%	175.9	42.9%
Multi-GeV	112.1	47.0%	32.8	36.2%
μee events(target: $\nu_\mu CC1\pi^0$)				
Sub-GeV	384.5	54.6%	201.2	38.0%
Multi-GeV	143.5	64.6%	51.6	32.0%

Table 3.3: Comparing the performance of $CC1\pi^0$ events reconstruction and selection in atmospheric neutrino MC events with three-ring final state.

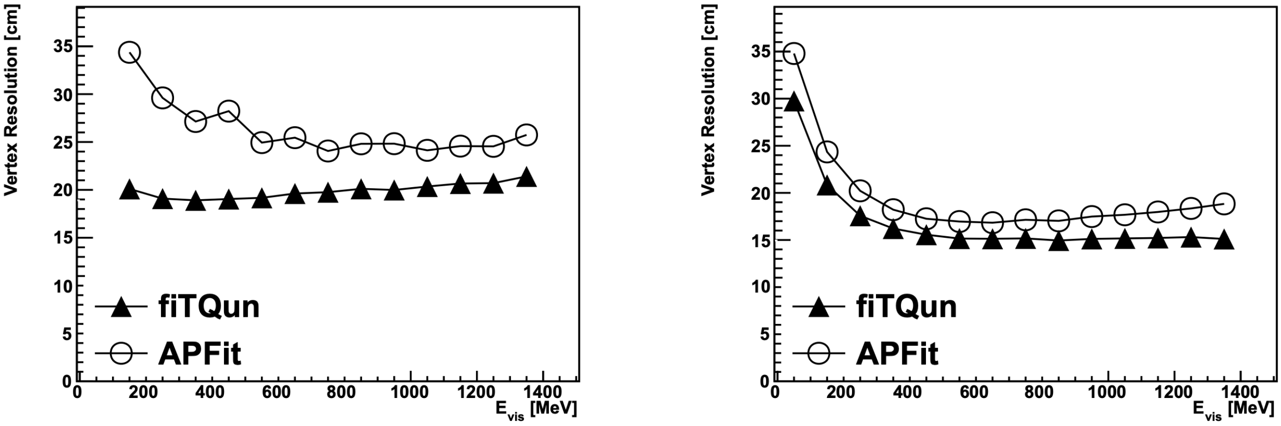


Figure 3.11: Vertex resolution

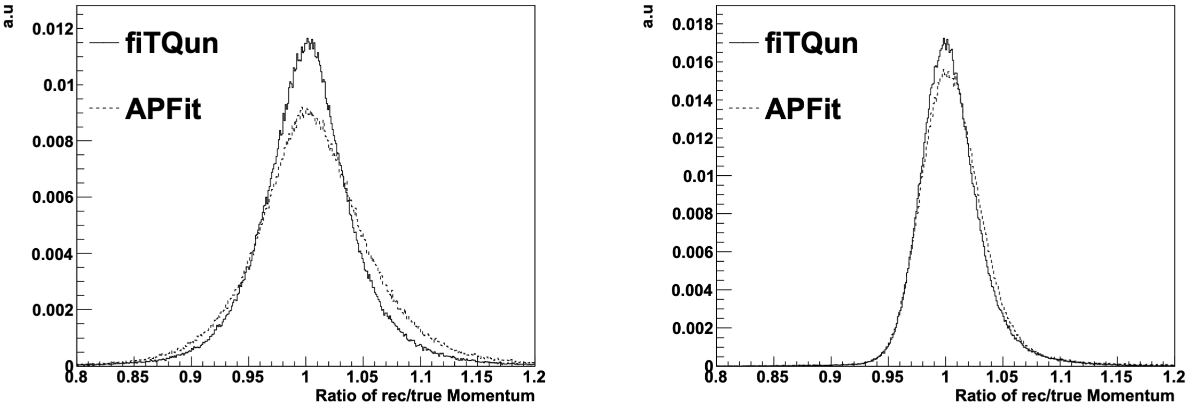


Figure 3.12: Single-ring electron (left) and muon (right) direction resolution

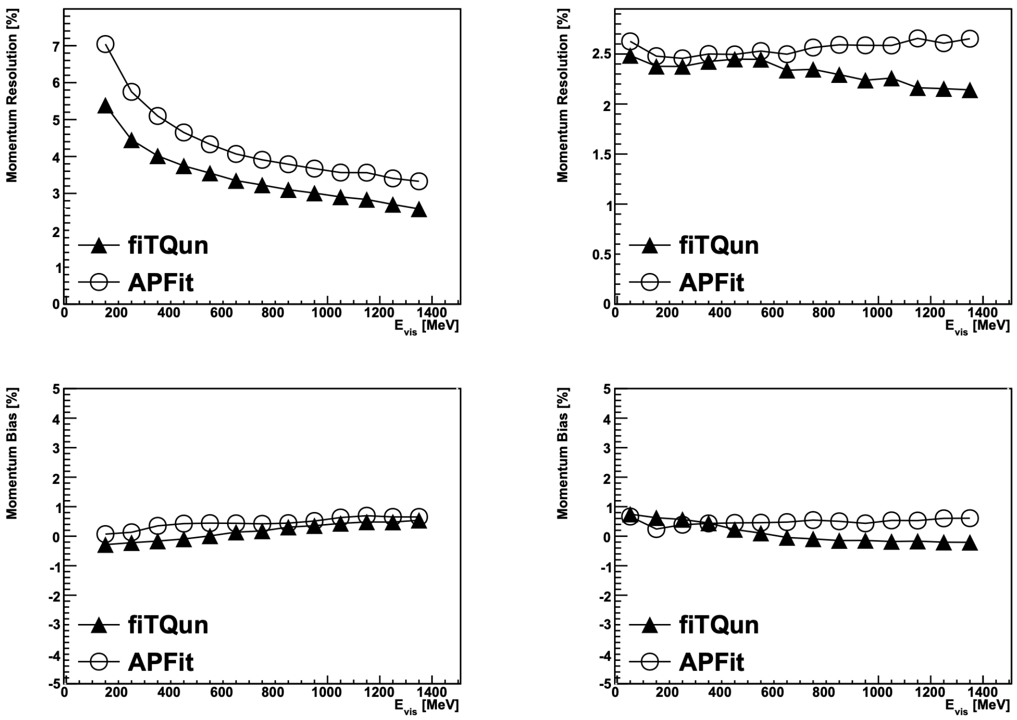


Figure 3.13: Direction resolution of single-ring electron (left) and muon (right) events

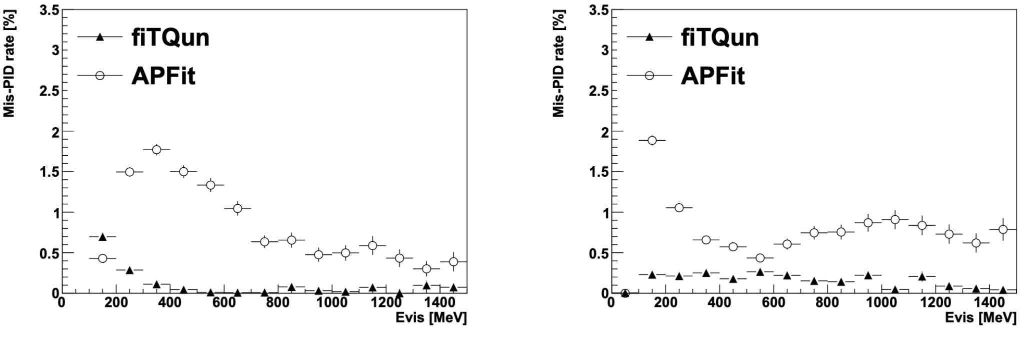


Figure 3.14: Single-ring electron (left) and muon (right) momentum resolution

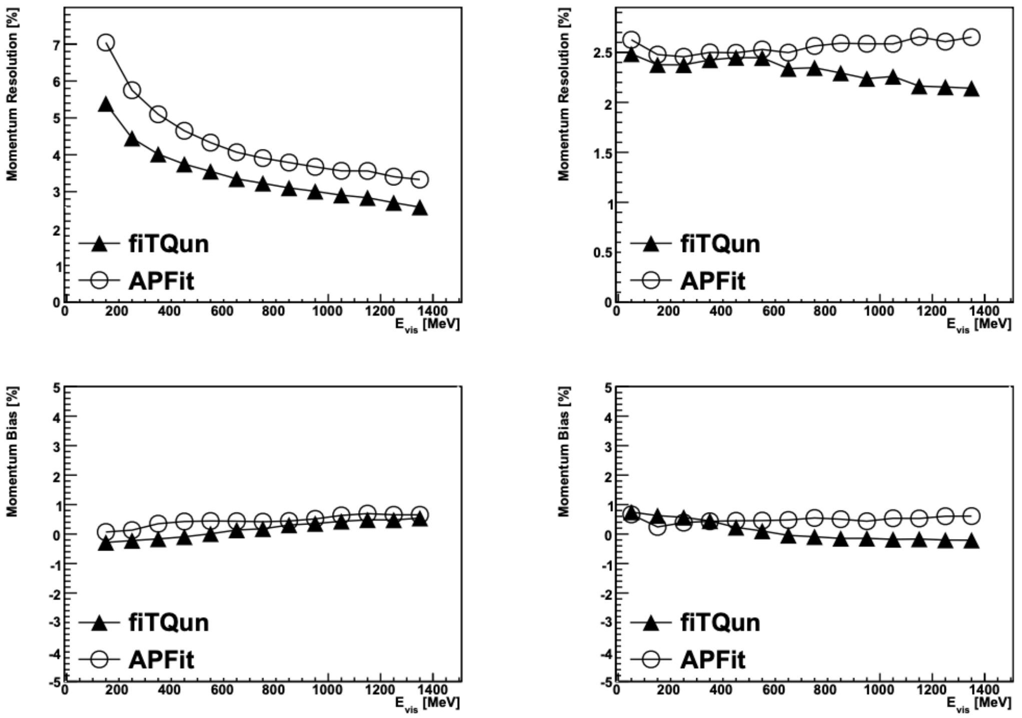


Figure 3.15: Momentum resolution and bias as a function of visible energy

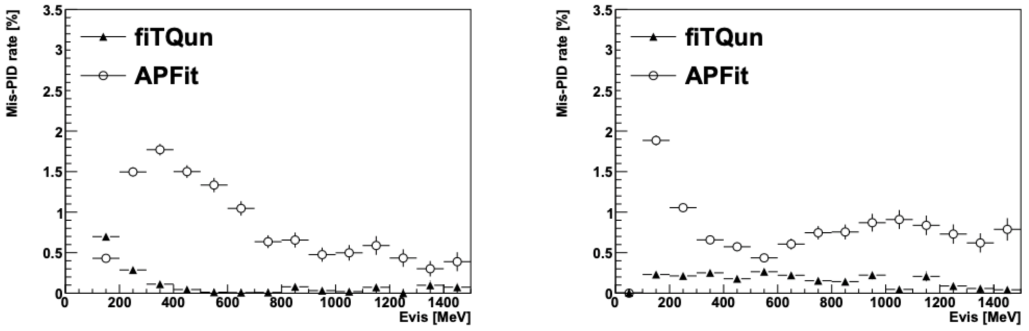


Figure 3.16: Mis-identification rate as a function of visible energy

Chapter 4

Tau Analysis with Improved Neural Network

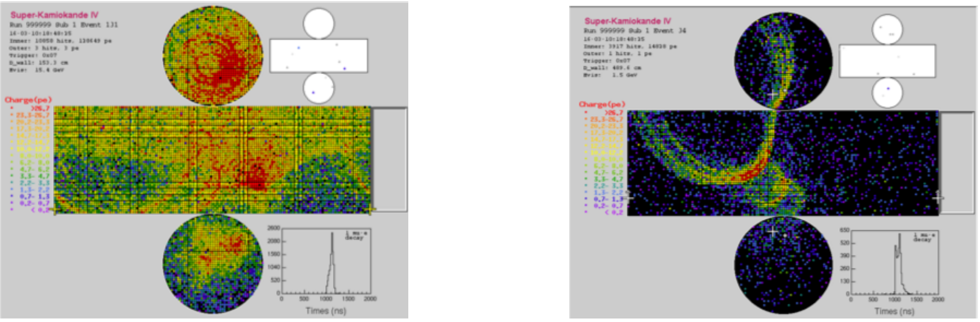
4.1 Tau selection with Neural Network

Tau neutrinos are requested to be discriminated from backgrounds induced by electron neutrino and muon neutrino. And neural network is built for this classification. There are some events display of SuperKamiokande:

This is a binary classification between signal and background events, deep inelastic scattering events and neutral current events are regarded as background events. While hadronic tau decay and leptonic tau decay are tau events. In fact, these events could be split into more subclass: Charged-Current quasi elastic(CCQE) ν_μ, CCqe ν_e, CCqe ν_τ and etc. The reason why we just did a binary classification rather than a multi-class classification is that we found binary classification is basically the same as multi-class classification. I preserve the binary classification considering the same level accuracy and unbalanced data.

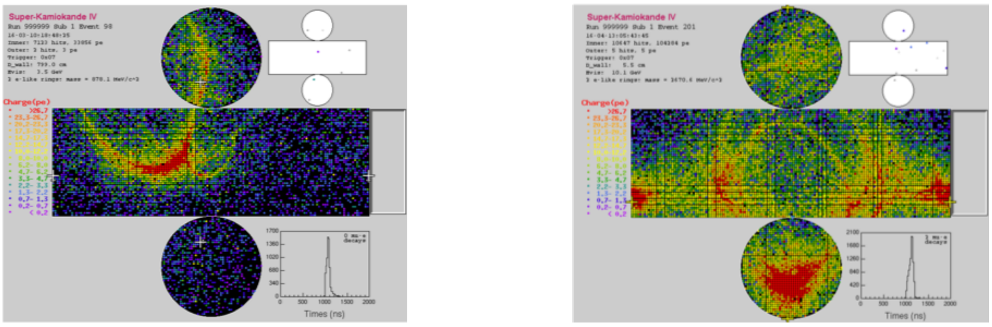
This time, the tau neural network is improved in these parts:

- Use fiTQun reconstruction algorithm instead of APfit.



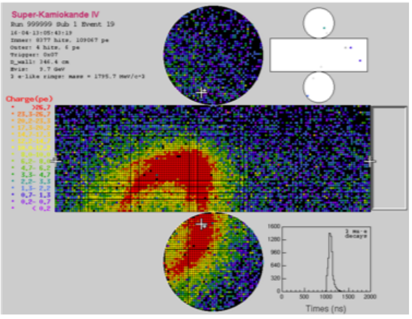
(a) DIS

(b) cc bkg events except DIS



(c) Neutral current

(d) hadronic tau decay



(e) leptonic tau decay

Figure 4.1: Several event displays

- Use improved tau neural-network(tau NN) algorithm.

As is discussed in previous chapter, a better performance is expected because fiTQun is more accurate than APfit in some reconstruction variables. This is the motivation why we want to substitute APfit with fiTQun shown in the Fig 4.2 and 4.3. As for the different structure, this is because we have seven input variables, and the possible nonlinear combinations of those variables are more than:
$\sum_{i=1}^{8} \binom{8}{i} = 2^7 = 128$, which is just the combination of linear term.
And previous network only contains one layer with 10 nodes which much less than our expectation. So I change the structure of neural network.

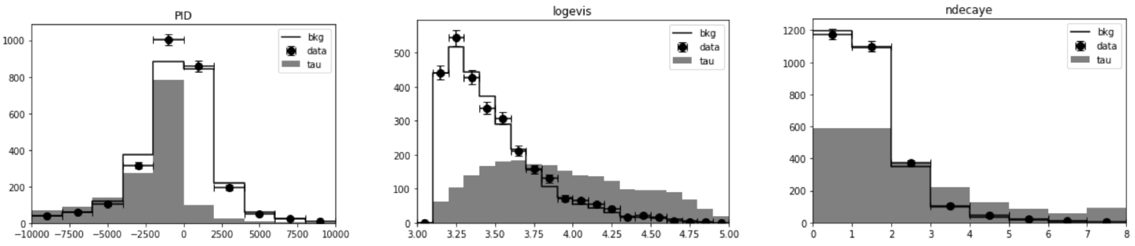
4.2 Input Variables

Seven basic input variables and a critical input variables are used. Those input variables come from the previous work.

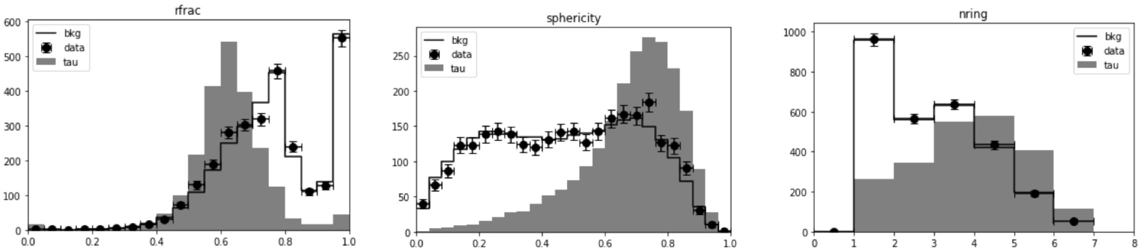
All input variables except possible Cherenkov ring candidates are built correctly by comparing the distribution of fiTQun version and APfit version.

Let's explain those input variables one by one, shown in Fig 4.2 and 4.3:

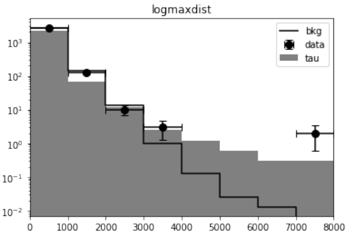
- The particle identification of the ring with maximum energy.(a)
  Tau lepton rapidly decays into daughter particles through lepton decay and hadron decay after formation. Most decay channels have at least one non-shower particle, except for the leptonic decay. Shower particles have smaller values in the definition of particle recognition possibilities than non-showering particles. The particle identification of the highest



(a) PID (b) logevis (c) ndecaye



(d) rfrac (e) sphericity (f) nring

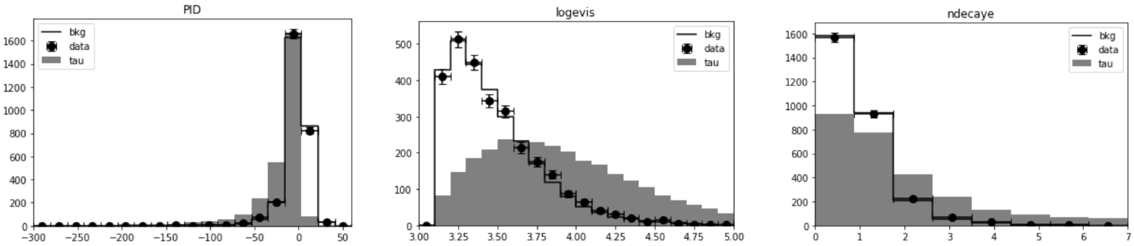


(g) logmaxdist

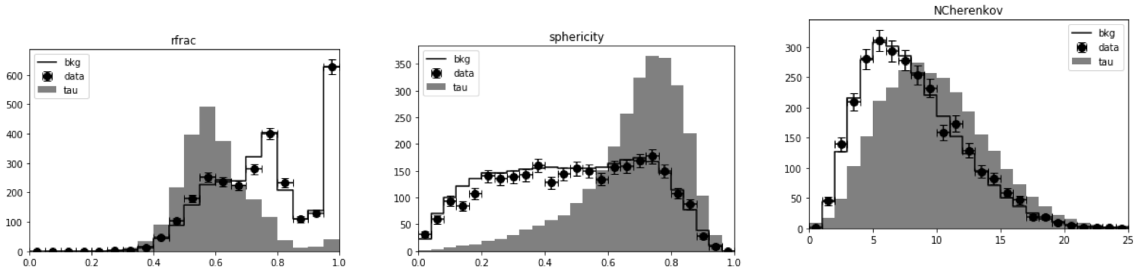
Figure 4.2: distribution of inputs(fitqun version), black points denotes the downward data events, grey histogram denotes the tau MC events and black histogram denotes the downward bkg events

energy ring of the tau is mainly distributed in the negative range, but the background is widely distributed in the both negative and positive areas.

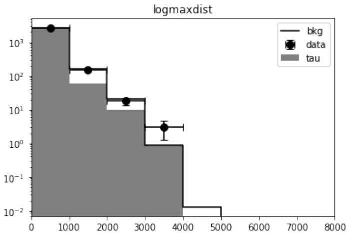
- The log base 10 of the total visible energy (Evis) of the event.(b)
  The energy of the shower electron, which produces the same amount of observed Cherenkov light, is defined as Evis. Tau events have higher average visible energy than background events. What's more, the background events decrease with increasing Evis while tau events have a peak near



(a) PID (b) logevis (c) ndecaye



(d) rfrac (e) sphericity (f) possible Cherenkov ring candidates



(g) logmaxdist

Figure 4.3: distribution of inputs(apfit version), black points denotes the downward data events, grep histogram denotes the tau MC events and black histogram denotes the downward bkg events

3.60 Gev.

- The number of decay electron candidates in the event.(c)

  Tau events are expected more decay electrons from hadronic tau decays than background events.

- The maximum distance between the primary interaction point and any decay electron from a pion or muon decay.(g)

As is known to us that a higher energy muon can travel a longer distance in water. So CC ν_μ background which involves a energetic neutrino is expected to have a large distance between the point where the primary interaction happens and the point where the decay electron comes into being from a muon. However, those pions from hadronic tau decay are expected to have smaller momentum which leads to a short distance.

- The clustered sphericity of the event in the center of mass system.(e)
  Intuitively, sphericity is a feature evaluating how spherical an object is. Sphericity 1 means perfectly isotropic while sphericity 0 means a purely one-directional. Quantitatively, we have to introduce a sphericity tensor.

  $$S^{\alpha\beta} = \frac{\sum_i p_i^\alpha p_i^\beta}{\sum_i p_i^2} \tag{4.1}$$

  where $\alpha, \beta = 1, 2, 3$ are the three Cartesian directions of the momentum, the sum is over the bins in geodesic space. Assuming this matrix has three eigenvalues: $\lambda_1 > \lambda_2 > \lambda_3$. Then the sphericity can be expressed as:

  $$S = \lambda_2 + \lambda_3 \tag{4.2}$$

  Peak of sphericity of tau is near 1 because hadronic decay of heavy tau is more isotropic than background events. While background events flattly distribute in range [0.0,1.0]. DIS usually carries high energy which is the reason why background events have larger sphericity events.

- The number of Cherenkov ring.(f)
  This variable is only the variable which is different from the APfit. This is because there is no such possible Cherenkov ring candidates in fiTQun. Because tau events have more product particles, tau events are expected

to contain more rings than how many rings background events contain.

- The fraction of total number of photoelectrons in the event carried by the photoelectrons locates in an angle range.(d)

  This variables is calculated as:

$$rfrac = \frac{\sum_{\theta_i < 48^\circ} q_i}{\sum q_i} \tag{4.3}$$

  where θ_i is the angle between the direction of the most energy ring and the direction of vertex to a PMT. The variables is the ratio of charged contained within a solid angle with 48° over the whole charge. Tau hadronic decays are expected to have smaller value of this variable because the energy is carried by multiple particles.

4.3 Change of the tauNN

New tauNN's structure is described in Fig 4.4 and Fig 4.5. Because previous tauNN have too less nodes compared to the number of input variables, we decide to increase the number of layers and number of nodes. The reason why I chose this specific structure is because I tried some other structures like 1 layer, 2 layer with more or less nodes, this structures is the best. And also I didn't try all of the possible structure.

We use softmax function as the activation function of last layer instead of identity function in order to constrain output of tauNN in the range of [0,1]. Addition, we change the lost function from mean of root square(MSE) to the cross entropy.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \widehat{Y_i})^2 \tag{4.4}$$

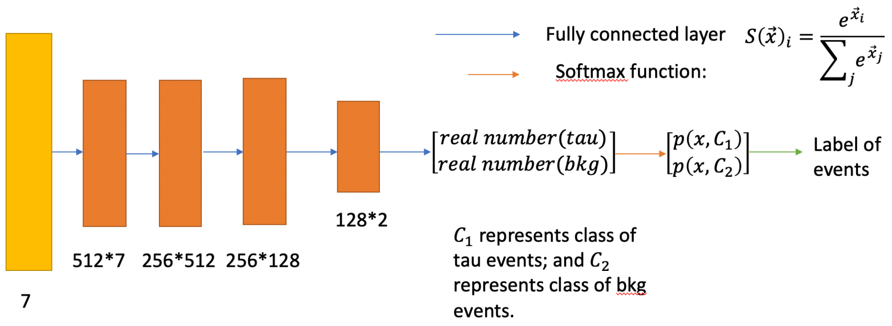
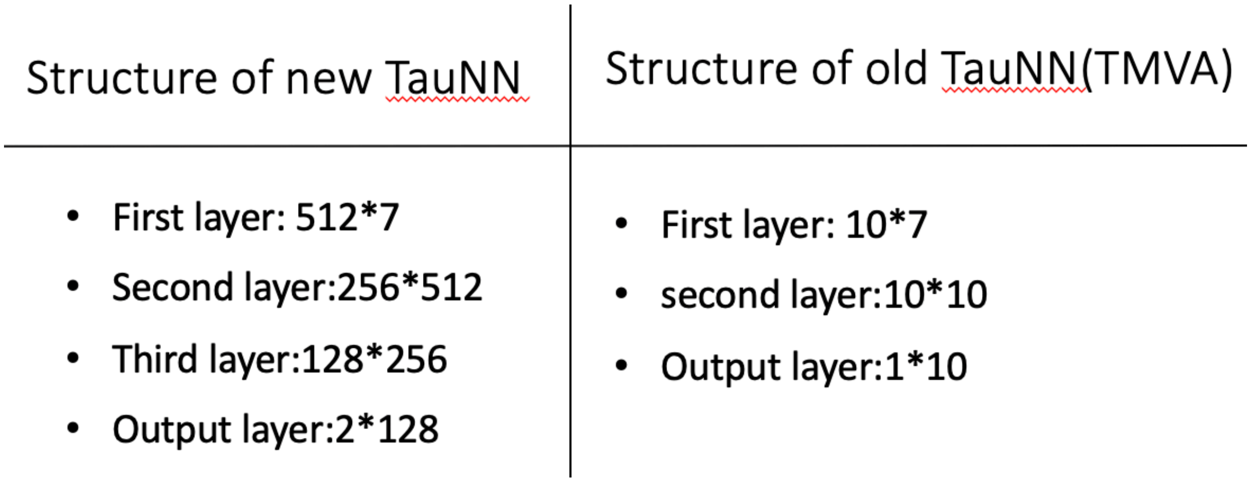


Figure 4.4: Schematic figure of tauNN



Structure of new TauNN	Structure of old TauNN(TMVA)
• First layer: 512*7 • Second layer:256*512 • Third layer:128*256 • Output layer:2*128	• First layer: 10*7 • second layer:10*10 • Output layer:1*10

Figure 4.5: details of new tauNN and old tauNN

where Y_i is the label gotten from our tauNN and $\widehat{Y_i}$ is the true label from Monte Carlo data.

This loss function is usually used in regression problem where we assume that error uniformly distributes in a range. As for the cross-entropy loss function, it's under the assumption that errors locate nearly from some discrete points. And it widely used in classification problem.

$$crossentropy = -\frac{1}{n}\sum_{i=1}^{n}(Y_i \cdot log(\widehat{Y_i}) - \widehat{Y_i} \cdot log(Y_i)) \tag{4.5}$$

The output of NN is defined as:

$$output = \frac{p(x, C_1)}{p(x, C_1) + p(x, C_2)} \tag{4.6}$$

where the $p(x, C_1)$ and $p(x, C_2)$ shown in the Fig 4.4, which represents how the event looks like tau. The more larger value is, the more likely as tau the event is. In addition, these two values are normalized which means the are in the range of [0,1]. How can we choose the cut for classification? Because the label should be binary which is tau or background, we find a cut which equal to 0.5 to separate the range into two parts. The idea is as follows:

- R_1 denotes the decision boundary for tau and R_2 denotes the decision boundary for bkg in Fig 4.6.
- Green area and red area denote the events which should be bkg but is misclassified as tau.
- Blue area denotes the events which should be tau but is misclassified as bkg.
- As the cut $\widehat{x}$ moves, the green area and blue area stay. Only the red increases or decreases.
- When Red area disappears, $p(x, C_1) = p(x, C_2)$. where $p(x, C_i) = \frac{e^{x_i}}{\sum_j e^{x_j}}$, x_j is the real number in Fig 4.4.
- We have $p(x, C_1) + p(x, C_2) = 1$ where $p(x, C_1) = p(x, C_2) = 0.5$

By comparing the output distributions of training set and test set, we know the tauNN is not overtraining because training set(29000 tau events and 29000

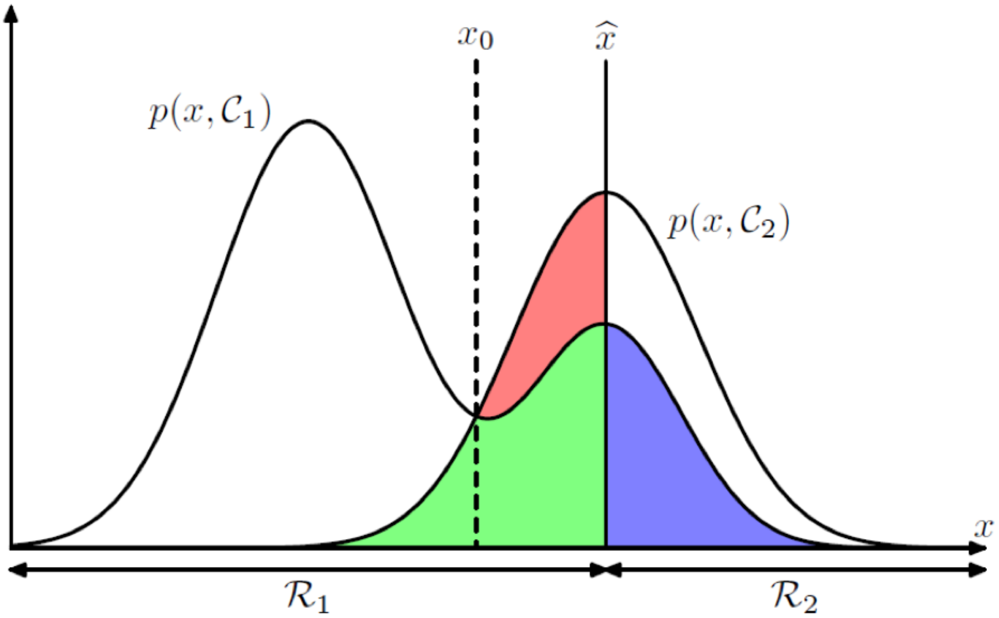


Figure 4.6: conceptional plot for signal and bkg separation

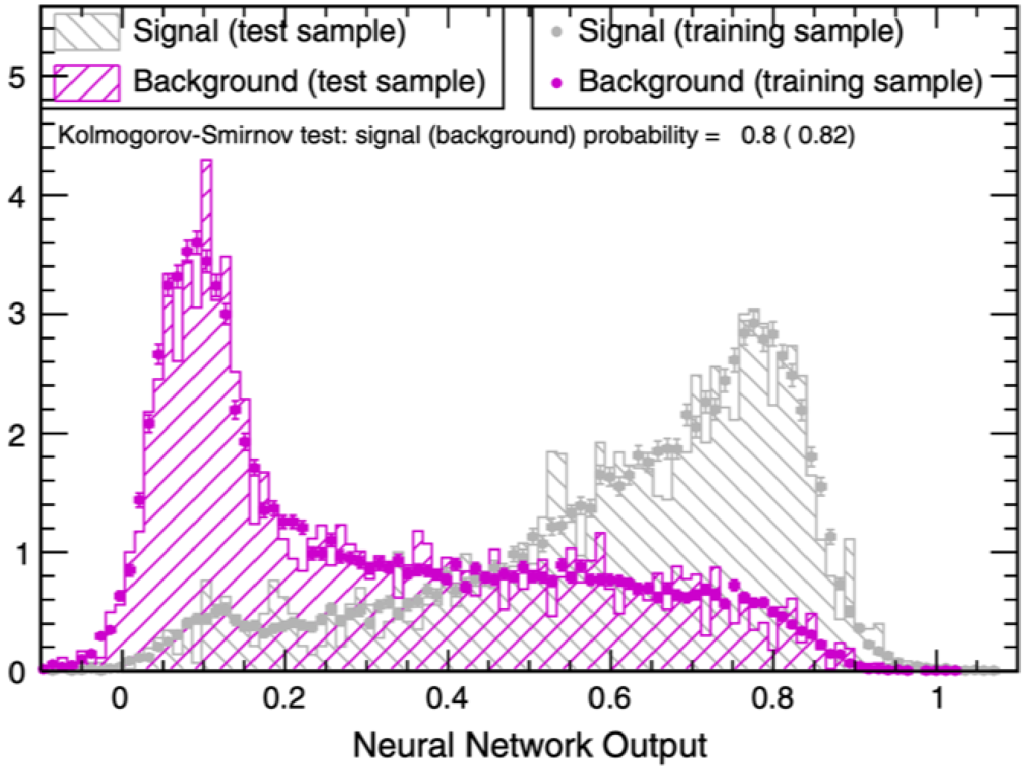


Figure 4.7: Output of previous work

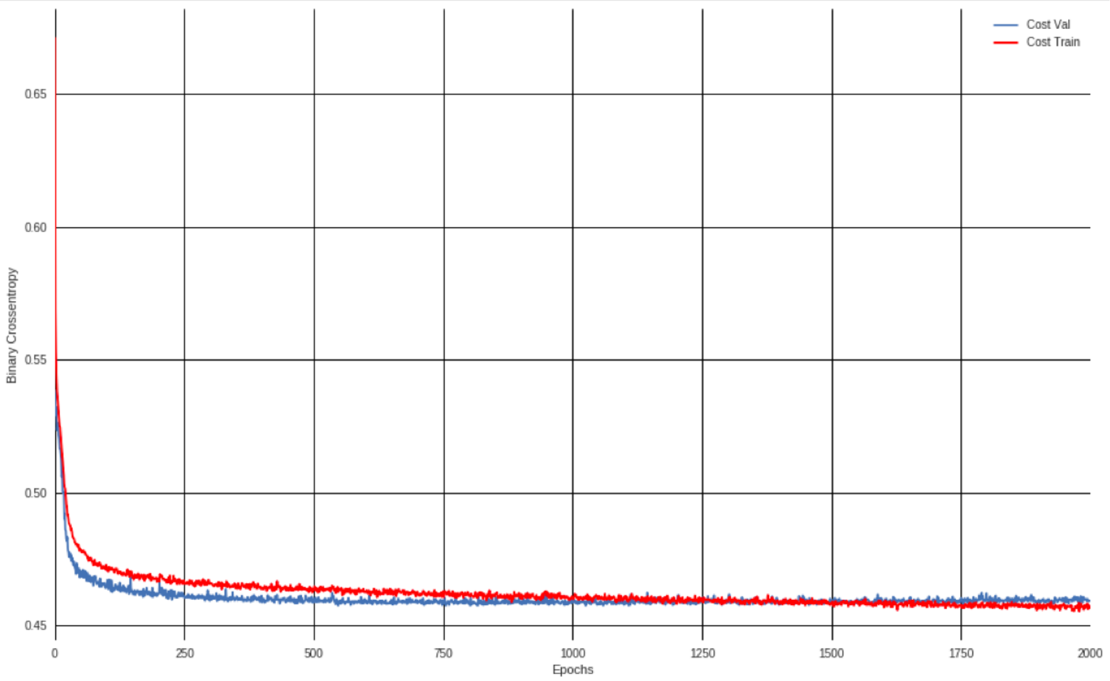


Figure 4.8: cost per epoch for training set(red) and test set(blue)

bkg events) and test set(3000 tau events and 3000 bkg events) are independent and identical distribution and tauNN is just implemented on training set. Same hyper-parameter are used: learning rate=0.02, optimizer=adam, epoch=1000. So the consistency between training set and test set reveals that tauNN is not overfitting. The selection power of new tauNN stems from new NN algorithm by comparing the output of previous work, new tauNN with apfit variables and new tauNN with fitqun variables.

Looking at Fig 4.7, 4.8 and 4.9, the new NN improves the accuracy of events identification(the overlapped area of new tauNN with apfit is less than the overlapped area of old tauNN with apfit) while fitqun doesn't strengthen the power of tau selection compared with apfit.

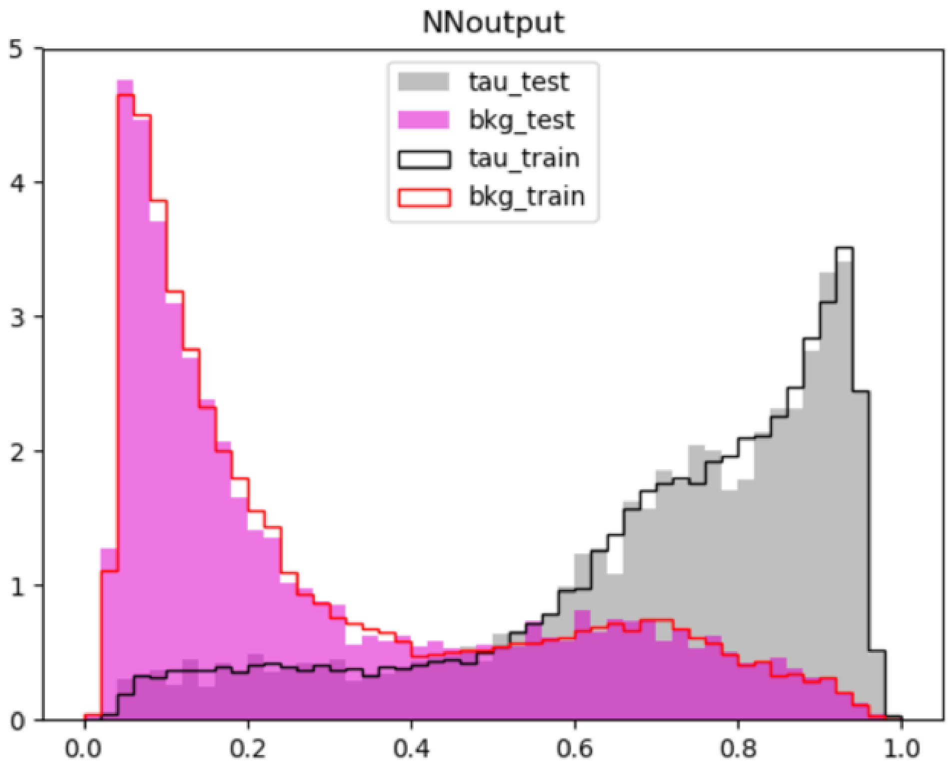


Figure 4.9: Output of new tauNN with apfit variables, new tau NN makes contribution in increasing the accuracy of tau selection compared with Fig 4.7

The terminology called accuracy is defined as:

$$accuracy = \frac{N'_{tau} + N'_{bkg}}{N_{tau} + N_{bkg}} \tag{4.7}$$

$$tau \quad accuracy = \frac{N'_{tau}}{N_{tau}} \tag{4.8}$$

$$bkg \quad accuracy = \frac{N'_{bkg}}{N_{bkg}} \tag{4.9}$$

From table 5.1 and table 5.1, we can know that new tau NN improves that selection power in each subclass, while fitqun doesn't improve the selection power compared with apfit.

Type	acc	tau acc	bkg acc
Current apfit with old NN	0.740	0.760	0.720
All apfit with new NN	0.795	0.822	0.767
taunn_maxenergyID	0.792	0.828	0.755
taunn_evis	0.796	0.835	0.757
taunn_decaye	0.799	0.843	0.756
taunn_distmax	0.797	0.831	0.763
taunn_gmsph	0.799	0.836	0.761
taunn_numseeds	0.795	0.829	0.762
taunn_rfrac	0.798	0.840	0.756
All fitqun with new NN	0.795	0.819	0.770

Table 4.1: Replacing fitqun variables for apfit variables one by one. Compared the first line, second line and last line, we could know that the improvement of accuracy comes from the new tauNN, while fitqun variables don't have large effect on accuracy

Type	Old tauNN with apfit	New tauNN with apfit	New tauNN with fitqun
$CC\nu_\mu$	0.846	0.855	0.851
$CC\nu_e$	0.656	0.690	0.689
$CC\nu_\tau$	0.759	0.822	0.819
NC	0.323	0.389	0.401

Table 4.2: Comparison in subclass. New tau NN improves the accuracy of all subclass compared to the previous tau NN

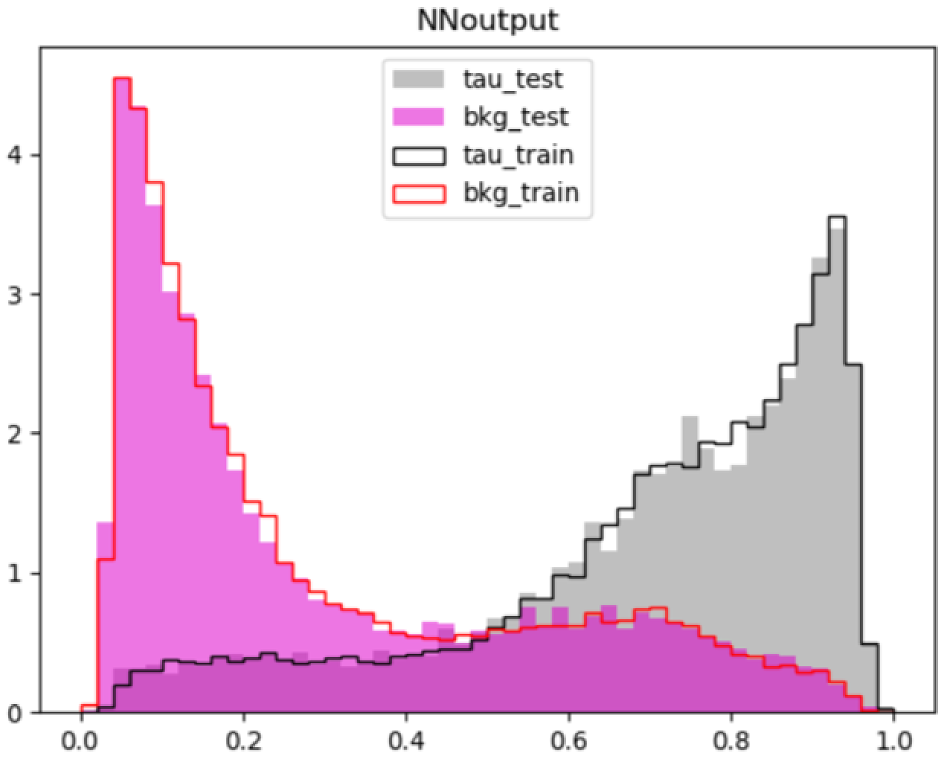


Figure 4.10: Output of new tauNN with fitqun variables, fiTQun variables doesn't improve the sensitivity of tau selection compared with Fig 4.8

Chapter 5

Tau analysis

This chapter will discuss the atmospheric tau neutrino appearance significance with the neural network mentioned in previous chapter.

5.1 Fit the Tau and Background Normalization

To fully extract information from the sample, this analysis apples a 2 - dimensional unbinned likelihood fit method. With the output of neural network and the reconstructed zenith angle of the events, two-dimensional particle distribution functions (PDFs) are made for background and tau signal. The horizontal(vectical) axis of the plots(Fig 5.1) is the cosine of the reconstructed zenith angle(NN output). For the NN output, tau-like events are close to 1 and non-tau-like events are close to 0. The signal events (left panel of Fig 5.1) are basically tau-like and come from below ($cos(\theta)$ near -1), while the background events are more non-tau-like and come from all directions. The amount of signal and back-ground events can be adjusted by varying the normalization of the distributions.

The fit of the data is done to the sum of the signal and background PDFs using different the normalization, and we want to let fit model as closed to data as possible:

$$fit \quad model = \alpha \times Signal + \beta \times BKG, \tag{5.1}$$

where α is the normalization of signal and β the normalization of background. Separate PDFS into Super-K run period for both signal events and bkg events. And the data are fitted into fit model individually for each run period jointly. The likelihood of this joint fit is as follows:

$$L(\alpha, \beta) = \prod_{SK\,I}^{SK\,IV} \prod_{event\,i} f(NN, cos\theta; \alpha, \beta) \tag{5.2}$$

where $f(NN, cos\theta; \alpha, \beta)$ is the normalized probability density function induced by the likelihood from simulation MC shown in Fig 5.2. The task is to maximize the PDF:

$$\max_{\alpha,\beta}\{L(\alpha, \beta)\} = \prod_{SK\,I}^{SK\,IV} \prod_{event\,i} f(NN, cos\theta; \alpha, \beta) \tag{5.3}$$

which is equivalent to :

$$\min_{\alpha,\beta}\{-log(L(\alpha, \beta))\} = -\sum_{SK\,I}^{SK\,IV} \sum_{event\,i} log(f(NN, cos\theta; \alpha, \beta)) \tag{5.4}$$

To show the results of the combined fit graphically which is easy for us to check, I plotted binned projections in zenith for both tau-like ($NNoutput < 0.5$) and non-tau-like ($NNoutput > 0.5$) events and the projections in NN output for both upward-going ($cos(\theta) > 0.1$) and downward-going ($cos(\theta) < 0.1$) events in Fig. 5.1. The 2D PDFs of signal and background are projected into 1D with respect to different ways and have been rescaled to the fitted normalization parameters α and β. All distributions show good agreements

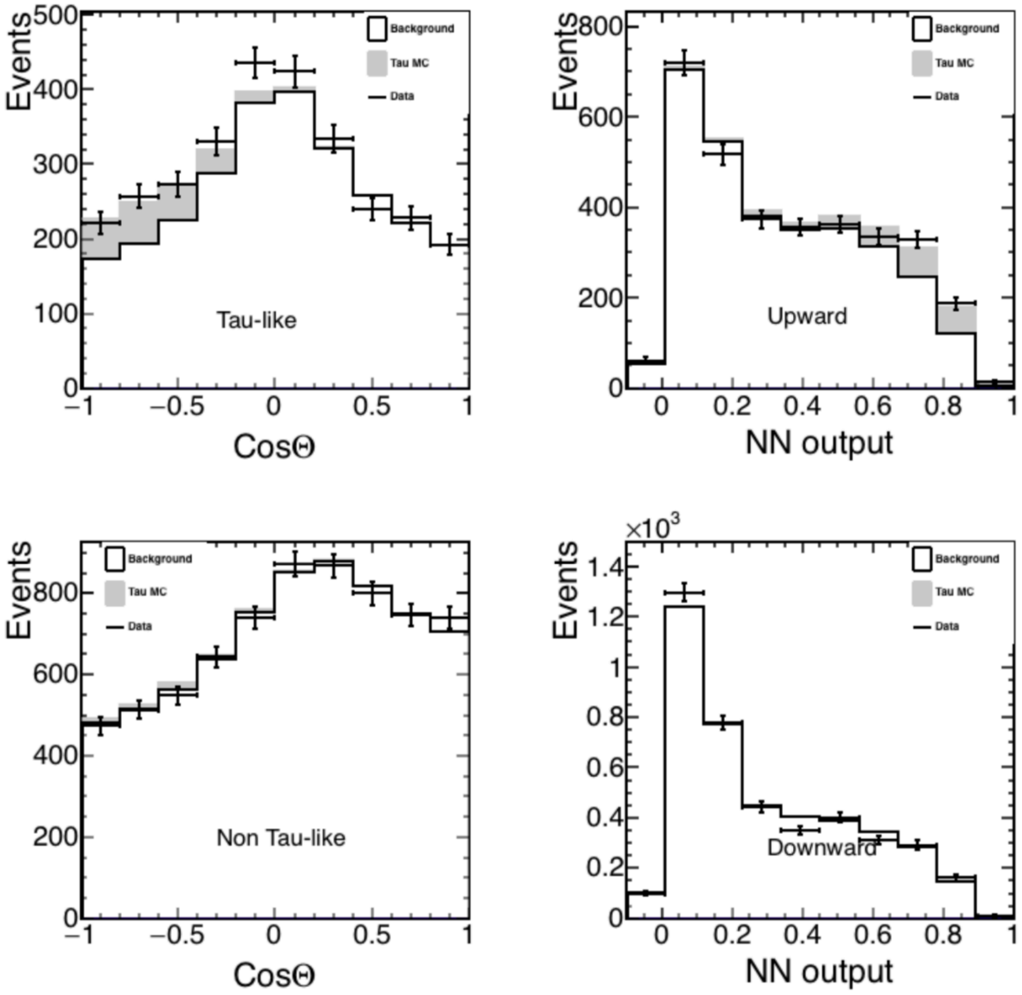


Figure 5.1: Fit results. Take from Zepeng Li's thesis(2017)

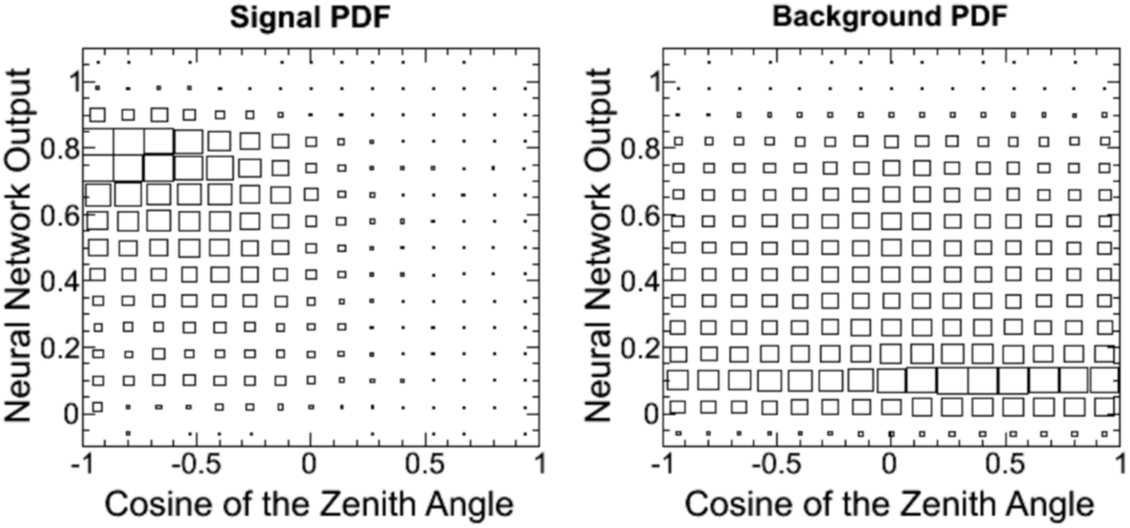


Figure 5.2: 2-dimensional probability distribution likelihood. Take from Zepeng Li's thesis(2017)

between data and MC. The tau neutrino signal mainly comes from the upward-going direction as expected. The PDFs and data from all SK1 to SK4 periods are stacked in these plots.

5.2 Tau Neutrino Significance

The results of new fit are shown in Fig 5.3 ,Fig 5.4, Fig 5.5 ,Fig 5.6 and Table 5.1, Table 5.2 where fiTQun is only used with SK4 apr16 data.

With the same Apr16 dataset, new tauNN led a stronger tau selection power and stronger tau selection power could pick more tau events as shown in Table 5.1. With more tau events, the tau neutrinos appearance result gets better than previous work with old tauNN. The tau neutrinos appearance result is

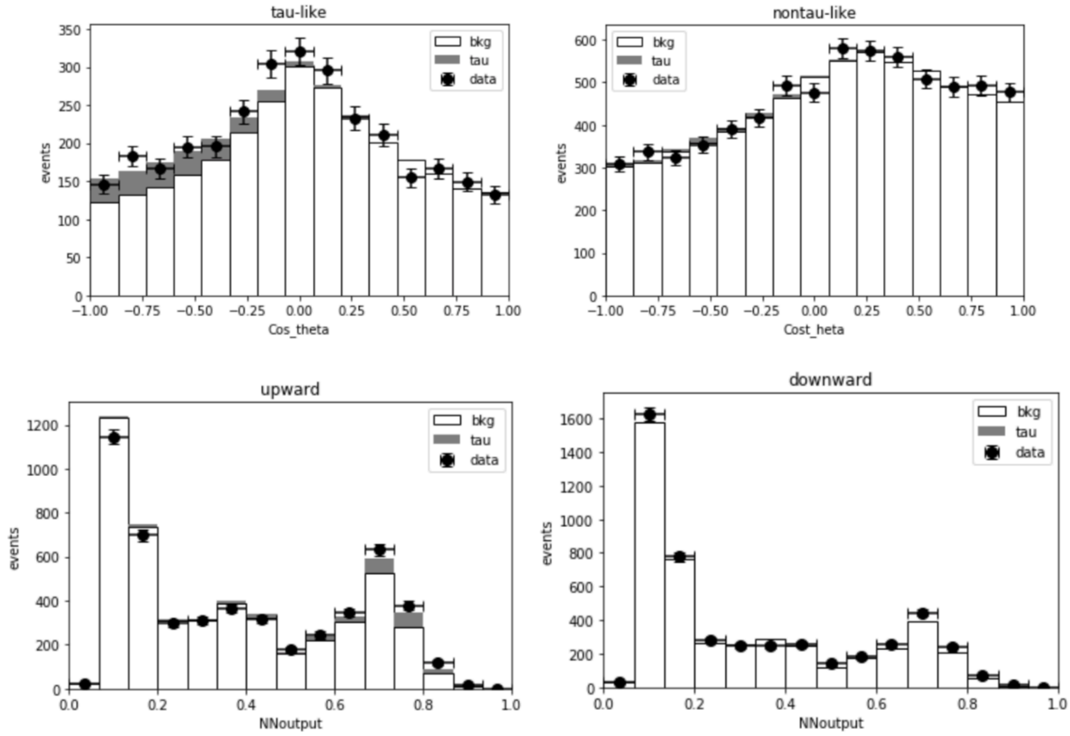


Figure 5.3: Fit results of new tauNN with apfit and apr16 sk1-sk4

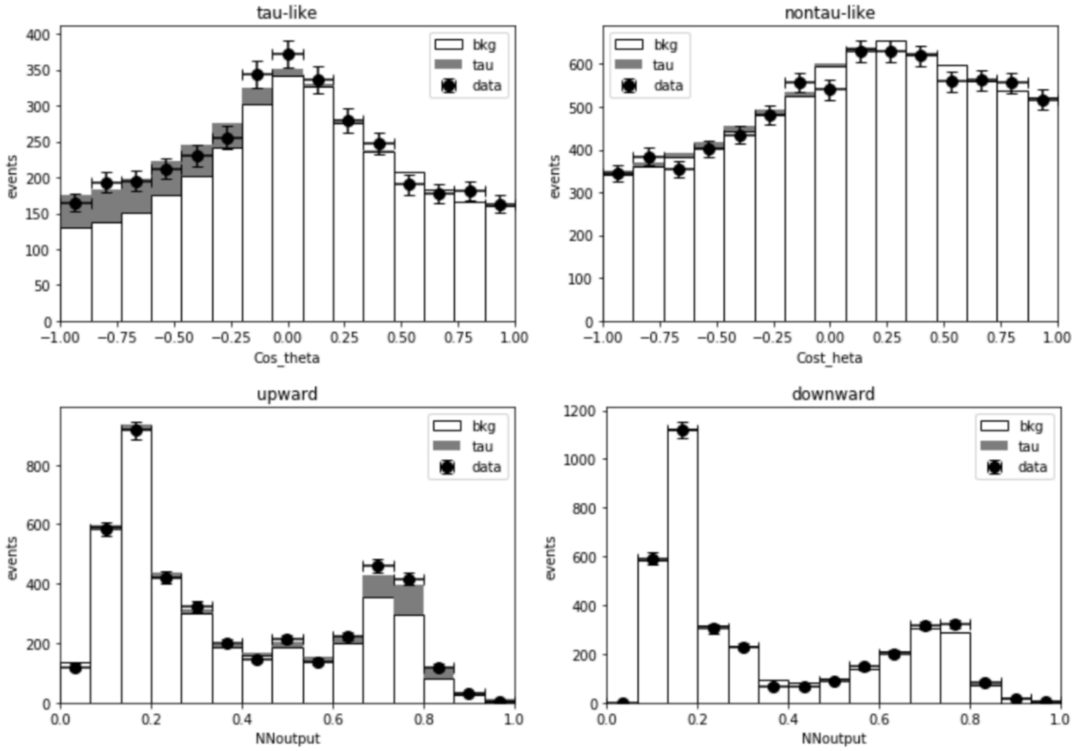


Figure 5.4: Fit results of new tauNN with apfit and apr18 sk1-sk4

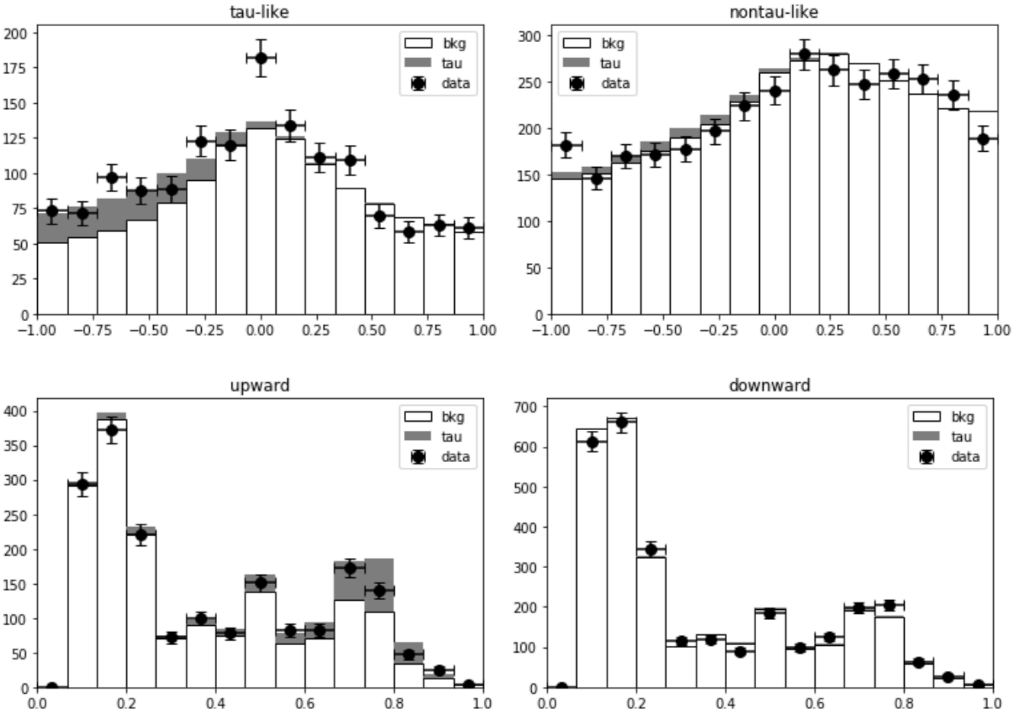


Figure 5.5: Fit results of new tau NN with fitqun and apr16 sk4

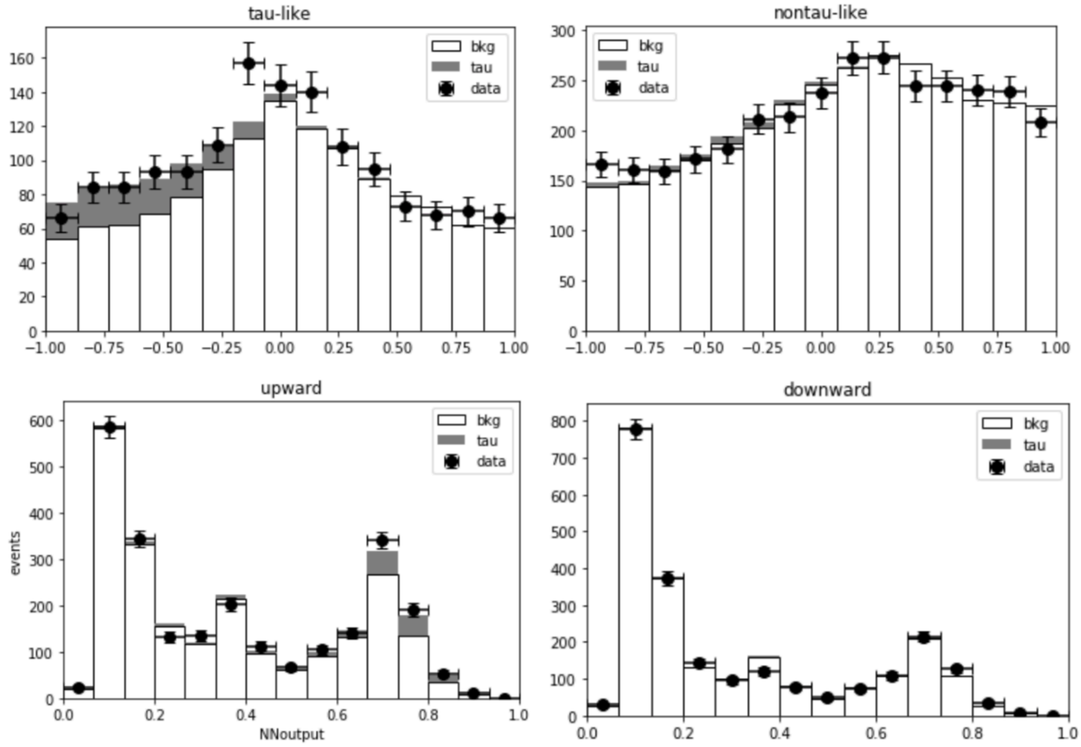


Figure 5.6: Fit results of new tau NN with apfit and apr16 sk4

NN type	old tauNN	new tauNN	new tauNN	new tauNN
reconstruction and data	apfit, apr16 sk1-sk4	apfit, apr16 sk1-sk4	fitqun, apr16 sk4	apfit, apr16 sk4
live time	5325.8	5325.80	2519.89	2519.89
α value	1.47 ± 0.32	1.96 ± 0.37	1.78 ± 0.41	1.86 ± 0.42
β value	1.00	0.92	0.93	0.92
statistic uncertainty	16%	13%	18%	18%
systematic error	14%	14%	14%	14%
total error	22%	19%	23%	23%
expected CCQE tau events	224.5 ± 57.2	232.7 ± 44.2	99.1 ± 22.8	103.5 ± 23.8
selected CCQE tau events	338.1 ± 72.7	312.8 ± 59.4	136.9 ± 31.5	143.1 ± 32.9
Significance (σ)	4.6	5.3	4.4	4.4

Table 5.1: Results of Tau neutrinos Appearance Significance. Second column is the result of previous work in which Apr16 sk1-sk4 data are used with old tau NN, third column is the result of new tauNN with apfit and Apr16 sk1-sk4 data, fourth column is the result of new tauNN with fitqun and Apr16 sk4 data, and fifth column is the result of new tauNN with apfit and Apr16 sk4 data.

NN type	new tauNN	new tauNN	new tauNN	new tauNN
Neut version	new	old	new	old
reconstruction and data	apfit, apr18 sk1-sk4	apfit, apr16 sk1-sk4	apfit, apr18 sk4	apfit, apr16 sk4
live time	5924.36	5325.80	3118.45	2519.89
α value	1.72 ± 0.32	1.96 ± 0.37	1.73 ± 0.39	1.86 ± 0.42
β value	1.02	0.92	1.02	0.92
statistic uncertainty	12%	13%	17%	18%
systematic error	14%	14%	14%	14%
total error	18%	19%	22%	23%
expected CCQE tau events	214.8 ± 39.6	232.7 ± 44.2	114.3 ± 25.6	103.5 ± 23.8
selected CCQE tau events	297.1 ± 54.9	312.8 ± 59.4	161.1 ± 36.1	143.1 ± 32.9
Significance (σ)	5.4	5.3	4.5	4.4

Table 5.2: Results of Tau neutrinos Appearance Significance. Second column is the result of APFit Apr18 sk1-sk4 data are used with new tau NN, third column is the result of new tauNN with apfit and Apr16 sk1-sk4 data, fourth column is the result of new tauNN with apfit and Apr18 sk4 data, and fifth column is the result of new tauNN with apfit and Apr16 sk4 data.

improved from 4.6 σ to 5.3 σ which exceeds 5.0 σ, so the tau neutrinos appearance result is more confirmed with new tauNN and Super-Kamiokande experiments.

And also the dataset is updated from Apr16 to Apr18 shown in Table 5.2. The different number expected CCQE tau events and selected CCQE tau events between Apr18 and Apr16 is probably because they used different neut library and MC data are different in each SK run period for Apr18 and Apr16. The new tauNN is trained on Apr18 and Apr16 independently. The statistical uncertainty of normalization parameters in Apr18 is less than the statistical uncertainty of normalization parameters in Apr16, which is because the live time is increased from 5325.80 days to 5924.36 days, and the tau neutrinos appearance result is improved from 5.3 σ to 5.4 σ.

Because the selection power of fitqun and apfit are basically the same in the dataset Apr16 SK4, the the tau events picked are approximately equally as expected in the previous chapter. The statistic uncertainty will be the same, while we use the same systematical error. This is the reason why fitqun and apfit have the same level of tau neutrinos appearance significance.

5.3 Discussion

The difference between fiTQun version input variables and APFit version input variables is the ring candidates. So is ring candidate the reason why fitqun doesn ’ t improve accuracy? To answer this question, I’ve done some experiments and got some results shown Table 5.2.

By comparing the accuracy with number of rings and accuracy with number of possible Cherenkov ring candidates, we can know that this input variable just effects the accuracy a little bit. Hence it can ' t explain why fitqun doesn ' t get the accuracy better.

The reason why fitqun doesn ' t improve accuracy needs further investigation.

reconstruction type	NN type	acc	tau acc	bkg acc
apfit with number of rings	new NN	0.790	0.840	0.740
apfit with number of ring candidates	new NN	0.795	0.822	0.767
fitqun with number of rings	new NN	0.795	0.819	0.770

Table 5.3: Comparison amount apfit with number of rings or number of ring candidates

Chapter 6

Summary

I study the different reconstruction algorithm and explore how they effect the tau selection power and final tau neutrinos appearance significance. I also research on how to improve the tauNN selection power by changing the structure of tauNN and then contribute in the final tau neutrinos appearance significance.

This thesis reveals fitqun variables don't contribute in tau neutrino appearance significance compared with apfit. And the new tau NN with apfit and apr16 sk1-sk4 data excludes the hypothesis of no-tau-appearance with a significance level of 5.3σ, which has improved from old tauNN whose significance level is 4.6σ, and new NN improves discrimination power of tau events from background events. The dataset is also updated from Apr16 to Apr18, and the corresponding tau neutrinos appearance significance is improved from 5.3 σ to 5.4 σ which exceeds to 5.0 σ. The tau neutrinos appearance is more confirmed with new tauNN.

Reference

[1] Bishop. Pattern recognition and machine learning. 2006.

[2] K. Abe et al(Super Kamiokande Collaboration). Search for differences in oscillation parameters for atmospheric neutrinos and antineutrinos at super-kamiokande. *Phys. Rev. Lett.*, 107:241801, Dec 2011.

[3] Z. Li et al(Super Kamiokande Collaboration). Measurement of the tau neutrino cross section in atmospheric neutrino oscillations with super-kamiokande. *Phys. Rev. D*, 98:052006, Sep 2018.

[4] Miao Jiang. Atmospheric neutrino oscillation analysis with improved event reconstruction in super-kamiokande iv, doctor thesis of kyoto university. 2019.

[5] Zepeng Li. Measurement of tau neutrino appearance and charged-current tau neutrino cross section with atmospheric neutrinos in super-kamiokande, doctor thesis of duke university. 2017.

[6] Lee Ka Pik. Study of the neutrino mass hierarchy with the atmospheric neutrino data observed in super-kamiokande, doctor thesis of tokyo university. 2012.