MACHINE LEARNING APPLICATION IN LOW ENERGY LIQUID SCINTILLATOR NEUTRINO EXPERIMENT

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Yu Xu
aus
Jiangsu, China

Berichter: Prof. Dr. Livia Ludhova
. Prof. Dr. Achim Stahl

Tag der mündlichen Prüfung: Nov 30, 2020

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek verfügbar.

Abstract

Neutrinos are the keys for physics beyond the Standard Model, since neutrino oscillation is the directly observed new physics phenomena. We use PMNS matrix to describe this phenomena. In the matrix, there are six free parameters: three mixing angles θ_{12}, θ_{23}, θ_{13}; two difference of the squared neutrino masses Δm^2_{21}, Δm^2_{32}; and CP violation angle δ_{CP}. Nowadays, there are one and half unknown parameters: sign of Δm^2_{32} and δ_{CP}.

The Jiangmen Underground Neutrino Observatory (JUNO) Experiment was designed to measure the sign of Δm^2_{32}, and we aim to achieve 3 σ confidence level with 6 years data. To achieve this goal, we construct a 20 kt liquid scintillator (LS) detector to measure the $\bar{\nu}_e$ energy spectrum from Yangjiang and Taishan Nuclear Power Plant(NPP).

Machine Learning (ML) is becoming more and more popular in data analysis. It can simplify the process of reconstruction with faster speed and better performance. In this thesis I will discuss the application of ML application in JUNO experiment, including waveform reconstruction, particle identification, and vertex/energy reconstruction.

Waveform reconstruction is the base for other reconstructions. In this thesis, I show the potential of ML on waveform reconstruction: time reconstruction of each single hit, which is impossible with traditional method.

Particle identification (PID) can reduce background/signal (B/S) ratio and improve the sensitivity of the experiment. In this thesis, we study the potential of particle identification (PID) with machine learning method. We can receive 95% signal with 5% background for alpha/beta discrimination, and 50% signal with 5% background for electron/positron discrimination.

We also perform a research on vertex and energy reconstruction with machine learning method, and find that we can meet the requirement of 3% energy resolution.

Abstract

Neutrinos sind die Schlüssel für die Physik jenseits des Standardmodells, da die Neutrinoschwingung das direkt beobachtete neue physikalische Phänomen ist. Wir verwenden die PMNS-Matrix, um dieses Phänomen zu beschreiben. In der Matrix gibt es sechs freie Parameter: drei Mischwinkel θ_{12}, θ_{23}, θ_{13}; zwei Unterschiede der quadratischen Neutrinomassen Δm^2_{21}, Δm^2_{32}; und CP-Verletzungswinkel δ_{CP}. Heutzutage gibt es eineinhalb unbekannte Parameter: Vorzeichen von Δm^2_{32} und δ_{CP}.

Das Experiment des Jiangmen Underground Neutrino Observatory (JUNO) wurde entwickelt, um das Vorzeichen von Δm^2_{32} zu messen. Mit 6-Jahres-Daten wollen wir ein Konfidenzniveau von 3 σ erreichen. Erreichen Zu diesem Zweck konstruieren wir einen 20-kt-Flüssigszintillator (LS) -Detektor zur Messung des Energiespektrums von $\bar{\nu}_e$ aus dem Kernkraftwerk Yangjiang und Taishan (NPP).

Maschinelles Lernen (ML) wird in der Datenanalyse immer beliebter. Es kann den Rekonstruktionsprozess mit schnellerer Geschwindigkeit und besserer Leistung vereinfachen. In dieser Arbeit werde ich die Anwendung der ML-Anwendung im JUNO-Experiment diskutieren, einschließlich Waveform rekonstruktion, Partikelidentifikation und Vertex/Energie-Rekonstruktion.

Die Waveform rekonstruktion ist die Basis für andere Rekonstruktionen. In dieser Arbeit zeige ich das Potenzial von ML für die Waveform rekonstruktion: Zeitrekonstruktion jedes einzelnen Treffers, was mit herkömmlichen Methoden unmöglich ist.

Die Partikelidentifikation (PID) kann das Hintergrund/Signal (B/S) -Verhältnis verringern und die Empfindlichkeit des Experiments verbessern. In dieser Arbeit untersuchen wir das Potenzial der Partikelidentifikation (PID) mit der Methode des maschinellen Lernens. Wir können 95% Signal mit 5% Hintergrund für Alpha/Beta Diskriminierung und 50% Signal mit 5% Hintergrund für Elektronen/Positron Diskriminierung erreichen.

Wir führen auch eine Untersuchung zur Vertex- und Energierekonstruktion mit der Methode des maschinellen Lernens durch und stellen fest, dass wir die Anforderung einer Energieauflösung von 3% erfüllen können.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In the area of particle physics, the most successful theory is the Standard Model (SM). In 2012, Higgs boson, which is the last particle predicted by the Standard Model, was found by CMS and ATLAS experiments in LHC. Thus, the Standard Model is complete and successful to explain almost everything in the particle physics.

However, neutrinos are the key for physics beyond the Standard Model. In the Standard Model, neutrinos are the Dirac particles with 0 mass and charge, 1/2 spin. However, this is in conflict with multiple neutrino oscillation experiments. The first evidence of this conflict was discovered by Raymond Davis, in the Homestake experiment [17]. With Homestake's data, Davis found that the solar neutrino flux measured by the experiments was only 1/3 of the Stand Solar Model (SSM) prediction, which comes the famous "solar neutrino problem". Subsequently, this missing of solar neutrinos was validated by different neutrino experiments, including GALLEX[18]/GNO[19], SAGE[20], and Super Kamiokande(Super K)[21] experiment.

In the 1980s, Kamiokande and IMB [22] experiment found that the ν_e/ν_μ ratio in the atmospheric neutrino flux deviates from 1, which is the so called "atmospheric neutrino anomaly".

To explain the above experimental results, scientists proposed the neutrino oscillation theory, meaning that neutrinos can transfer from one flavor to another via the transition. In this case, neutrinos are not zero-mass and their flavor eigenstates are

not degenerate with their mass eigenstates, which are not in the framework of Standard Model. Similar to CKM matrix in the quark mixing, PMNS matrix is proposed to describe neutrino mixing. In the matrix, there are 6 parameters: 3 mixing angles (θ_{12}, θ_{23}, θ_{13}), 2 mass square differences (Δm^2_{21}, Δm^2_{32}) and one CP phase angle (δ_{CP}).

In 1998, Super K experiment published the result, showing the atmospheric neutrino oscillation spectrum with the zenith angle, which consists well with the neutrino oscillation theory. Later in 2001, SNO experiment [28] publish the result, which completely solve the solar neutrino problem. By measuring both the flux of only ν_e and all three flavor neutrinos in the same time, SNO found that the flux of the sum of all three flavor of neutrinos was consistent with the prediction of the Standard Solar Model, while the flux of ν_e is only one third of the prediction, which prove that the neutrinos are not really "missing" but just transfer to other flavors. In 2015, Super K and SNO won the Nobel Prize for the discovery of neutrino oscillation.

From the solar neutrino experiments and KamLAND experiment [29], scientists measure the value of θ_{12} and Δm^2_{21}; from the atmospheric neutrino experiments and accelerator neutrino experiments, scientists measure the value of θ_{23} and $|\Delta m^2_{32}|$. In 2012, Daya Bay experiment [24] published the measurement value of θ_{13}. Thus, we have known the value of 4.5 parameters in PMNS matrix, while the sign of Δm^2_{32} and CP phase angle δ_{CP} remain unknown. In the next decades, the aim of the neutrino oscillation experiments will be focus on the measurement of the remaining oscillation parameters.

JUNO experiment is also one of the next generation experiment aimed to measure the sign of Δm^2_{32}, while its huge target mass and high energy resolution allow it to explore much more physics topics. In this thesis, I will show how to do event reconstruction in the LS experiments, with the implementation of the fancy deep learning method.

This thesis follows the following outline:

Chapter 2 is the introduction of the theory of neutrino physics and neutrino experiments, as well as some open questions in neutrino physics.

Chapter 3 describes the details of the design of JUNO detector.

Chapter 4 introduces the basic theory of machine learning.

Chapter 5 starts with the status of traditional waveform reconstruction methods, and shows the limitation of them. Thus, we need machine learning methods to improve the performance, so as to meet the requirement of some physics topics.

Chapter 7 focuses on the implementation of machine learning method in the field of particle identification.

Chapter 6 shows the energy and vertex reconstruction with deep learning methods.

Chapter 8 shows how we can fit mass ordering based on Bayes theory, and discuss the possible factors which can affect the final results.

Chapter 9 is the conclusion and outlook of this thesis.

Chapter 2

Neutrino Physics

Neutrinos are among the elementary particles making up the world. There are about $330/cm^3$ neutrinos in the universe, making it second most abundant particle in our world, just behind photons. Thus, the properties of the neutrinos can significantly influence the evolution of the universe.

However, we still don't fully understand this ghost particles, even today. Although the Standard Model can successfully classify all known elementary particles, the properties of neutrinos are not fully in the framework of the Standard Model. In the Standard Model, the neutrinos should be zero mass while the neutrino oscillation experiments confirm that neutrinos have mass. Scientists establish neutrino oscillation theory to explain the results of experiments, while the theory is incompatible with the Standard Model. Thus, the study of the neutrinos is the key for the new physics beyond the Standard Model.

In this chapter, we will describe the properties neutrinos and the theory of neutrino oscillations, show scientists' effort to understand the neutrinos, and discuss the remaining questions in the current neutrino physics.

2.1 Neutrino in the Standard Model

In the Standard Model, neutrinos are particles with 0 mass and charge, 1/2 spin. The neutrinos only participate in weak and gravitational interactions. There are in total

3 flavors of neutrinos: ν_e , ν_μ, ν_τ, corresponding to 3 flavors of charged leptons e, μ, and τ. In the Standard Model, there are only left-handed neutrinos and right-handed anti-neutrinos, as observed in nature.

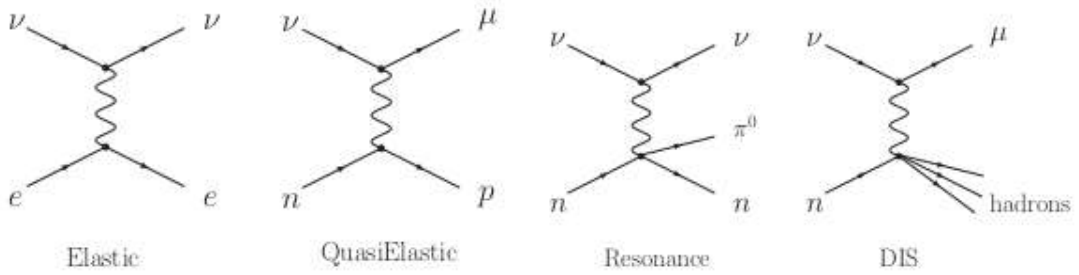


Figure 2.1: Four kinds of interaction between neutrino and matter: Elastic, QuasiElastic, Resonance, and Deep Inelastic Scattering (DIS)

In the experiments, we can detect neutrinos via their weak interaction with matter: Elastic Scattering (ES), Quasi-elastic scattering, single pion production and Deep Inelastic Scattering (DIS), as shown in Fig 2.1. With energy increase, the cross section of all interaction channel increase, while the domain interaction channel changes from ES to DIS, as shown in Fig 2.2 [4].

However, we need to point out that neutrinos are very hard to detect due to their small cross sections. For example, even for energy up to 100 GeV, the neutrino cross section is still very small, just at around 10^{-36} cm^2. Thus, we can only detect neutrinos with huge flux and/or high energy.

Due to the difficulty of neutrino detection, although neutrino was introduced to interpret beta decay in the 1930s, the observation of this particle was around 15 years later. In 1956, Clyde Cowan, Frederick Reines detected electron (anti)neutrino in the experiment. The other two flavors were observed even later. In 1962, Leon Lederman, Melvin Schwartz and Jack Steinberger found muon neutrino, and the last flavor (tau neutrino) was detected in 2000 in DUNOT project.

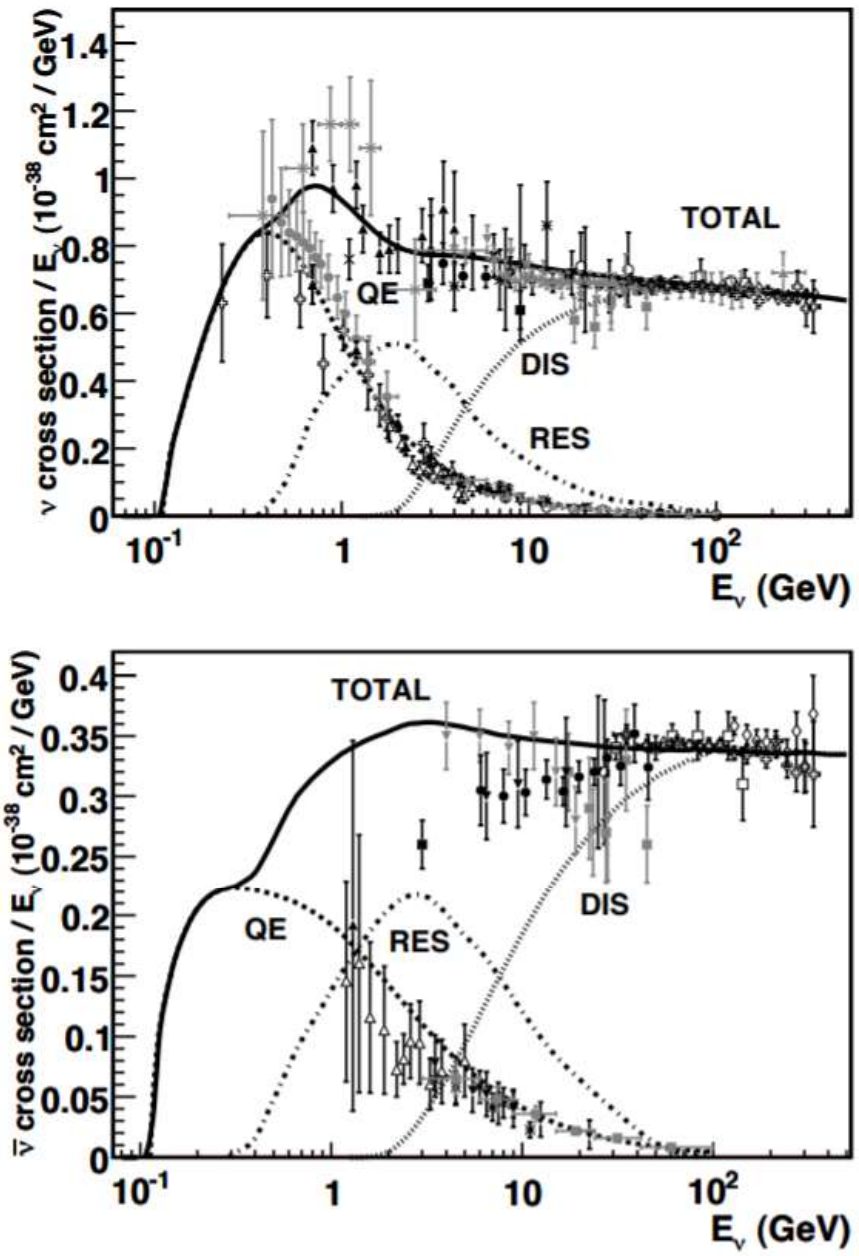


Figure 2.2: Total neutrino and antineutrino per nucleon CC cross sections (for an isoscalar target) divided by neutrino energy and plotted as a function of energy. These contributions include quasi-elastic scattering (dashed), resonance production (dot-dash), and deep inelastic scattering (dotted). Note that the quasi-elastic scattering data and predictions have been averaged over neutron and proton targets and hence have been divided by a factor of two for the purposes of this plot. [4]

2.2 Neutrino Oscillation

In the neutrino oscillation theory, the neutrinos have non-zero mass and their mass eigenstates and flavor eigenstates are not the same. Neutrinos are generated and detected as flavor eigenstate, while propagated as mass eigenstate. Since the neutrinos will propagate with different speed due to their different masses, the flavor of neutrinos detected should be the function of the distance from the neutrino source.

There are three generation of neutrinos, so each flavor of neutrinos can be treated as the superposition of three mass eigenstate [ν_1, ν_2, ν_3]:

$$|\nu_\alpha >= \Sigma_i U^*_{\alpha i}|\nu_i > \tag{2.1}$$

where ν_α is the flavor eigenstate [ν_e, ν_μ, ν_τ], ν_i is the mass eigenstate [ν_1, ν_2, ν_3], and U is the lepton mixing matrix (PMNS matrix). In the framework of three generation neutrino oscillation, the formula can be rewritten as

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \tag{2.2}$$

Thus, we can derive the neutrino oscillation probability:

$$\begin{aligned} P(\nu_\alpha \to \nu_\beta) =& | < \nu_\alpha | \nu_\beta > |^2 \\ =& \delta_{\alpha\beta} \\ & - 4\Sigma_{i>j} Re(U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j}) \sin^2(\frac{\Delta m^2_{ij} L}{4E}) \\ & + 2\Sigma_{i>j} Im(U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j}) \sin^2(\frac{\Delta m^2_{ij} L}{2E}) \end{aligned} \tag{2.3}$$

,

where $\Delta m^2_{ij} \equiv m^2_i - m^2_j$. Here we can take the unit of Δm^2_{ij} as eV^2, the unit of L as km, and the unit of E as GeV, then we can get:

$$\Delta m^2_{ij}(L/4E) \simeq 1.27 \Delta m^2_{ij}(eV^2) \frac{L(km)}{E(GeV)} \tag{2.4}$$

By parameterizing PMNS matrix U, we can get:

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & C_{23} & S_{23} \\ 0 & -S_{23} & C_{23} \end{pmatrix} \begin{pmatrix} C_{13} & 0 & S_{13}^{*} e^{i\delta_{CP}} \\ 0 & 1 & 0 \\ -S_{13}^{*} e^{i\delta_{CP}} & 0 & C_{13} \end{pmatrix} \begin{pmatrix} C_{12} & S_{12} & 0 \\ -S_{12} & C_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} e^{i\phi_1} & & \\ & e^{i\phi_2} & \\ & & 1 \end{pmatrix} \tag{2.5}$$

In the matrix $C_{ij} = \cos\theta_{ij}$, $S_{ij} = \sin\theta_{ij}$. The matrix is combined by four terms. The first three of them are two-generation oscillation matrices, while the last term is the phase term. The first term is also called "atmospheric neutrino oscillation" term, which is measured mainly via atmospheric neutrino experiments (Super K [23] et al.) . The second term is also called "reactor neutrino oscillation" term, which is first observed in Daya Bay experiment [24] and later confirmed by RENO experiment [25] and Double Chooz experiment [26]. The third term is also called "solar neutrino oscillation" term, which is measured by solar neutrino experiments (Super K [27], SNO [28]) and reactor neutrino experiments (KamLAND [29]). For the last term, ϕ_1 and ϕ_2 is Majorana phase angle, and we can measured them only if neutrinos are Majorana particles.

From the above functions, we know that there are six independent oscillation parameters: three mixing angles θ_{12}, θ_{23}, θ_{13}; two mass square differences $\Delta m_{21}^2 = m_2^2 - m_1^2$ and $\Delta m_{32}^2 = m_3^2 - m_2^2$; one CP phase angle δ_{CP}. The current values of neutrino oscillation parameters is shown in Fig 2.3.

In addition, the amplitude of oscillation is determined by the mixing angles, while the frequency of oscillation is determined by the mass square difference. Currently, we have measured 4.5 oscillation parameters (θ_{12}, θ_{23}, θ_{13}, Δm_{21}, $|\Delta m_{32}|$) from various neutrino oscillation experiments, and 1.5 oscillation parameters remain unknown (the sign of Δm_{32} and δ_{CP}). Currently various experiments are proposed to measure the oscillation last parameters. To measure the sign of Δm_{32}, there are reactor experiments (JUNO [1], RENO50 [30]), atmospheric neutrino experiment (Hyper K [31]), and accelerator neutrino experiment (DUNE [32]); To measure δ_{CP}, there are accelerator neutrino experiments (DUNE [32] and T2HK [31]) . We are able to know

all of the six neutrino oscillation parameters in the next 10 to 20 years.

NuFIT 4.1 (2019)

		Normal Ordering (best fit)		Inverted Ordering ($\Delta\chi^2 = 6.2$)	
		bfp $\pm 1\sigma$	3σ range	bfp $\pm 1\sigma$	3σ range
without SK atmospheric data	$\sin^2\theta_{12}$	$0.310^{+0.013}_{-0.012}$	$0.275 \to 0.350$	$0.310^{+0.013}_{-0.012}$	$0.275 \to 0.350$
	$\theta_{12}/^\circ$	$33.82^{+0.78}_{-0.76}$	$31.61 \to 36.27$	$33.82^{+0.78}_{-0.76}$	$31.61 \to 36.27$
	$\sin^2\theta_{23}$	$0.558^{+0.020}_{-0.033}$	$0.427 \to 0.609$	$0.563^{+0.019}_{-0.026}$	$0.430 \to 0.612$
	$\theta_{23}/^\circ$	$48.3^{+1.1}_{-1.9}$	$40.8 \to 51.3$	$48.6^{+1.1}_{-1.5}$	$41.0 \to 51.5$
	$\sin^2\theta_{13}$	$0.02241^{+0.00066}_{-0.00065}$	$0.02046 \to 0.02440$	$0.02261^{+0.00067}_{-0.00064}$	$0.02066 \to 0.02461$
	$\theta_{13}/^\circ$	$8.61^{+0.13}_{-0.13}$	$8.22 \to 8.99$	$8.65^{+0.13}_{-0.12}$	$8.26 \to 9.02$
	$\delta_{\rm CP}/^\circ$	222^{+38}_{-28}	$141 \to 370$	285^{+24}_{-26}	$205 \to 354$
	$\frac{\Delta m^2_{21}}{10^{-5}\ \text{eV}^2}$	$7.39^{+0.21}_{-0.20}$	$6.79 \to 8.01$	$7.39^{+0.21}_{-0.20}$	$6.79 \to 8.01$
	$\frac{\Delta m^2_{3\ell}}{10^{-3}\ \text{eV}^2}$	$+2.523^{+0.032}_{-0.030}$	$+2.432 \to +2.618$	$-2.509^{+0.032}_{-0.030}$	$-2.603 \to -2.416$
		Normal Ordering (best fit)		Inverted Ordering ($\Delta\chi^2 = 10.4$)	
		bfp $\pm 1\sigma$	3σ range	bfp $\pm 1\sigma$	3σ range
with SK atmospheric data	$\sin^2\theta_{12}$	$0.310^{+0.013}_{-0.012}$	$0.275 \to 0.350$	$0.310^{+0.013}_{-0.012}$	$0.275 \to 0.350$
	$\theta_{12}/^\circ$	$33.82^{+0.78}_{-0.76}$	$31.61 \to 36.27$	$33.82^{+0.78}_{-0.75}$	$31.61 \to 36.27$
	$\sin^2\theta_{23}$	$0.563^{+0.018}_{-0.024}$	$0.433 \to 0.609$	$0.565^{+0.017}_{-0.022}$	$0.436 \to 0.610$
	$\theta_{23}/^\circ$	$48.6^{+1.0}_{-1.4}$	$41.1 \to 51.3$	$48.8^{+1.0}_{-1.2}$	$41.4 \to 51.3$
	$\sin^2\theta_{13}$	$0.02237^{+0.00066}_{-0.00065}$	$0.02044 \to 0.02435$	$0.02259^{+0.00065}_{-0.00065}$	$0.02064 \to 0.02457$
	$\theta_{13}/^\circ$	$8.60^{+0.13}_{-0.13}$	$8.22 \to 8.98$	$8.64^{+0.12}_{-0.13}$	$8.26 \to 9.02$
	$\delta_{\rm CP}/^\circ$	221^{+39}_{-28}	$144 \to 357$	282^{+23}_{-25}	$205 \to 348$
	$\frac{\Delta m^2_{21}}{10^{-5}\ \text{eV}^2}$	$7.39^{+0.21}_{-0.20}$	$6.79 \to 8.01$	$7.39^{+0.21}_{-0.20}$	$6.79 \to 8.01$
	$\frac{\Delta m^2_{3\ell}}{10^{-3}\ \text{eV}^2}$	$+2.528^{+0.029}_{-0.031}$	$+2.436 \to +2.618$	$-2.510^{+0.030}_{-0.031}$	$-2.601 \to -2.419$

Figure 2.3: Three-flavour oscillation parameters from global fit [5]

In the reactor neutrino experiments, the neutrino sources are the reactors, and we

measure the survival probability of $\bar{\nu}_e$

$$P_{ee} = 1 - P_{21} - P_{31} - P_{32} \tag{2.6}$$

$$\begin{aligned} P_{21} &= \cos^4(\theta_{13}) \sin^2(2\theta_{12}) \sin^2(\Delta_{21}) \\ P_{31} &= \cos^2(\theta_{12}) \sin^2(2\theta_{13}) \sin^2(\Delta_{31}) \\ P_{32} &= \sin^2(\theta_{12}) \sin^2(2\theta_{13}) \sin^2(\Delta_{32}) \end{aligned} \tag{2.7}$$

here Δ_{ij} is

$$\Delta_{ij} = \frac{\Delta m_{ij} L}{4E} \simeq 1.27 \Delta m^2_{ij}(eV^2) \frac{L(km)}{E(GeV)} \tag{2.8}$$

From the formula, we can know that the survival probability is the function of L/E. For short baseline reactor neutrino experiments like Daya Bay [33], term P_{21} is almost zero and can be ignored, while for medium baseline neutrino experiments like JUNO [1], term P_{21} take the main effect and all three terms should be considered.

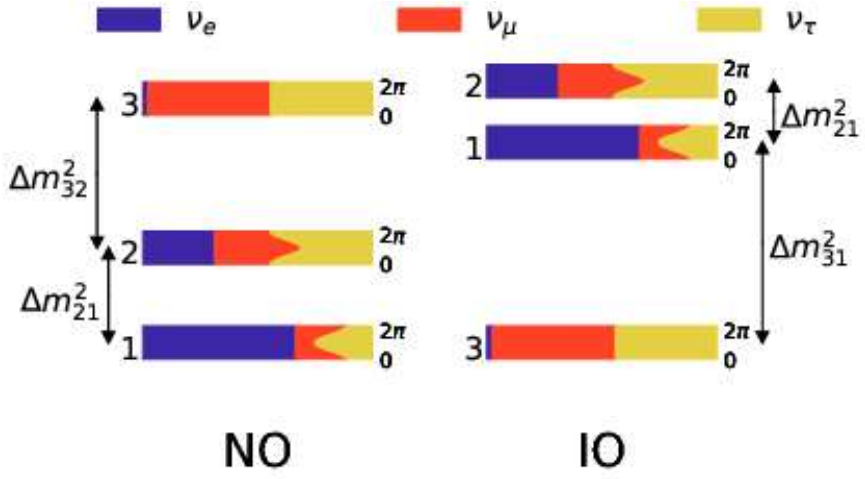


Figure 2.4: Neutrino Mass Ordering (MO) problem. Left: Normal Ordering (NO): $m_1 < m_2 < m_3$; Right: Invert Ordering (IO): $m_3 < m_1 < m_2$)

In addition, $\Delta m^2_{31} = \Delta m^2_{32} - \Delta m^2_{21}$, $|\Delta m^2_{32}| \gg \Delta m^2_{21}$. Since there the sign of Δm^2_{32} is unknown, there can be two different sign and value for the different sign of

Δm^2_{32}. Thus, there can be two possible neutrino spectrum observed by the detector. However, due to the smallness of Δm^2_{21}, the difference of the two spectrum is only about 3%, which require very high energy resolution to discriminate the difference.

JUNO will start taking data in early 2022, and promise to give the neutrino mass ordering result with the sensitivity of 3 σ after 6 years of data taking.

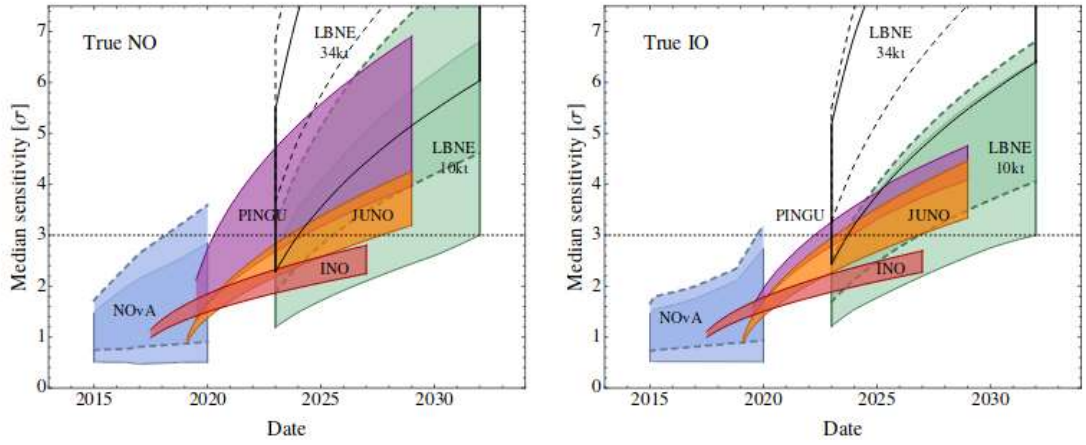


Figure 2.5: The median sensitivity in number of sigmas for rejecting the IO (NO) if the NO (IO) is true for different facilities as a function of the date. Left: NO is true; Right: IO is true [6]

2.3 Some Open Questions in the Neutrino Physics

There are still some questions remained to be answered in neutrino physics:

2.3.1 Dirac? Majorana?

Currently we still don't know the kinds of neutrinos: are they Dirac particles or Majorana particles? If they are Dirac particles, just as their corresponding leptons (e, μ, τ), they must be distinguishable from their antiparticle. However, they can also be Majorana particles, meaning that their anti-particles are themselves. The only way to solve this problem is to observe the neutrino-less double-beta decay $(0\nu\beta\beta)$ process.

The effective neutrino mass term in the $0\nu\beta\beta$ decay is defined as:

$$< m_{\beta\beta} >= |\Sigma_j m_j U_{ej}^2| = |u_{e1}^2 e^{i\alpha_1} m_1 + u_{e2}^2 e^{i\alpha_2} m_2 + u_{e3}^2 e^{i\alpha_3} m_3| \tag{2.9}$$

From the current experiments, we set the limit of the half life time of the neutrino less double beta decay process at the level of 10^{25} to 10^{26} years, see Tab 2.1 and 2.6. If the right neutrino mass ordering is Inverted Ordering, we are able to observe the $0\nu\beta\beta$ process with the next decade $0\nu\beta\beta$ experiments (nEXO [34], LEGEND [35], KamLAND2-Zen [36], PandaX-III [37]). Even if we don't see the neutrino less double beta decay process, we can exclude the inverted neutrino mass ordering at 90% CL.

Isotope	Experiment	$T_{1/2}^{0\nu\beta\beta}$[yr]	$< m_{\beta\beta} >$[meV]
^{136}Xe	EXO-200 [38]	$>1.6 \times 10^{25}$	<140-380
^{136}Xe	KamLAND-Zen [39]	$>1.07\times10^{26}$	<61-165
^{136}Xe	GERDA [40]	$>5.3\times10^{25}$	<150-330
^{130}Te	CUORE [7]	$>1.5\times10^{25}$	<110-520

Table 2.1: Summary of current $0\nu\beta\beta$ experiments

JUNO also proposed a plan to search $0\nu\beta\beta$ process. By building a balloon filled with enriched xenon gas (with ^{136}Xe up to 80%) dissolved in LS, inserted into the central region of the JUNO LS. The energy resolution can reach up to about 1.9% at the Q-value of ^{136}Xe $0\nu\beta\beta$ decay. JUNO has the potential to reach a sensitivity (at 90% C. L.) to of 1.8×10^{28} yr (5.6×10^{28} yr) with around 50 tons (5 tons) of fiducial (136)Xe and 5 years exposure, while in the 50-ton case the corresponding sensitivity to the effective neutrino mass, $m_{\beta\beta}$, could reach (5 - 12) meV, covering completely the allowed region of inverted neutrino mass ordering. [41]

2.3.2 Neutrino absolute mass

The observation of neutrino oscillation proves neutrinos to possess non-zero rest masses, which contradict the Standard Model. However, from neutrino oscillation experiments we can only know the neutrino mass squared differences Δm_{21}^2 and Δm_{32}^2, so the absolute value m_i of each mass state ν_i (i = 1,2,3) remains unknown. In addition, the possibility of the lightest neutrino mass to be 0 still exists.

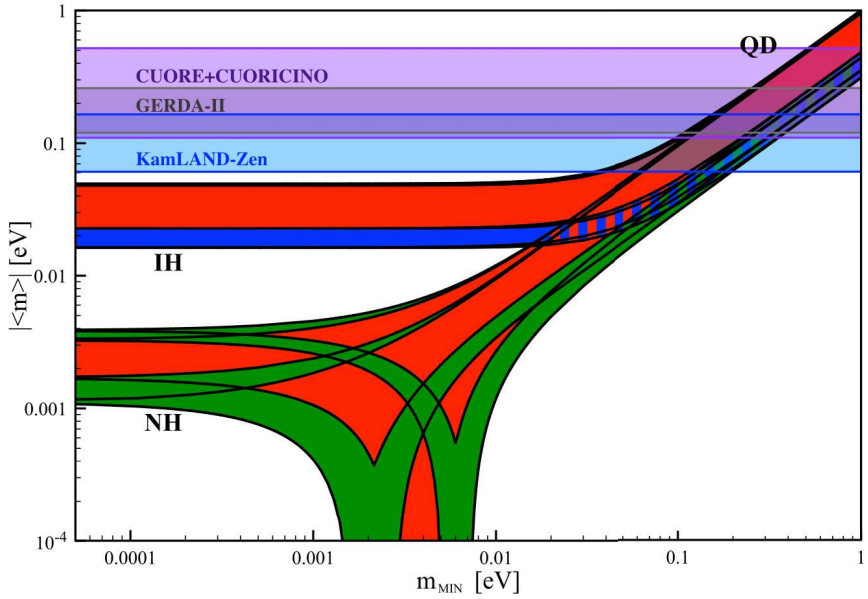


Figure 2.6: Current limit of the half-life of the neutrino-less double beta decay process [7].

Nowadays, from CMB data scientists give the tightest upper limit of Σm_i of 0.13 eV [42], and from $0\nu\beta\beta$ experiments we know in the case of Majorana neutrino the upper limit of effective Majorana mass $m_{\beta\beta}$ (eq. 2.9) is less than 0.2 eV (see Tab 2.1) . However, all above methods are all model-dependent.

The Karlsruhe Tritium Neutrino (KATRIN) experiment provides the model independent method to measure neutrino's mass m_ν and will reach the sensitivity of 200 meV with 5 years' data [43]. The principle of KATRIN is simple: by the precision measurement of tritium's β decay processes, scientists can get incoherent sum of spectra (or the "mass" of ν_e), which is called "m_{ν_e}". KATRIN starts taking data in early 2019 and gives the upper limit of m_ν of 1.1 eV with 90% confidence level (CL) [44] based on the first several months' data.

In conclusion, the precise measurement of neutrino mass is still challenging, and maybe we need some new ideas on this question.

2.3.3 Sterile neutrino

With the measurement of Z boson decay width [45], we know there are only 3 generations of active neutrinos. However, there may exist several kinds of "sterile" neutrinos, which only participate in gravitational interaction. The existence of sterile neutrinos could help us to answer several questions. Heavy sterile neutrino could explain the lightness weight of neutrino via seesaw mechanism, keV scale mass sterile neutrino is a promising candidate of the Warm Dark Matter (WDM), and sterile neutrinos at the eV or sub-eV scale can be the explanation of several experimental anomalies.

Since the "sterile" neutrinos do not participate in the weak interactions, they can not be detected directly and can only be observed via neutrino oscillation processes. There are several hints about the sterile neutrino at the eV or sub-eV scale, including the reactor anomaly [46], the gallium anomaly [47], the LSND anomaly [48], and the MiniBooNE anomaly [49]. However, these hints are in contradiction with other experiments. Nowadays, several experiments are running or proposed to search sterile neutrinos, including MicroBooNE [50], PROSPECT [51][52], SoLid [53]. Hopefully we are able to confirm or exclude the existence of sterile neutrino in the next decades.

If we finally find the sterile neutrino, it will be another path to probe the physics beyond the Standard Model.

2.3.4 CP phase

In our theory, matter and antimatter should be created equally at the beginning of the universe. However, this is not the truth, as we found that asymmetry between matter and antimatter is most apparent in the observation of the universe, which is composed of matter with little antimatter. One possible explanation is the violation of the so-called Charge-Parity (CP) symmetry. Until now, CP symmetry violation has only been observed in the physics of subatomic particles called quarks, but the magnitude of the CP symmetry violation is not large enough to explain the observed dominance of matter over antimatter in the universe.

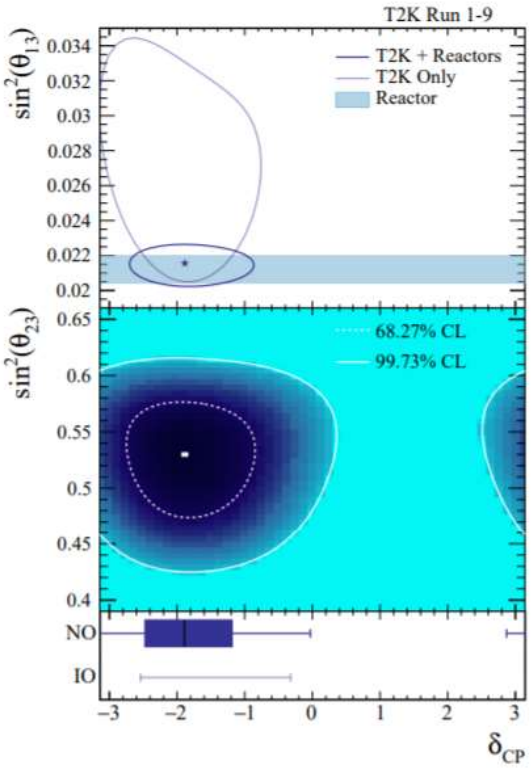


Figure 2.7: The upper panel shows 2D confidence intervals at the 68.27% confidence level for δCP vs $\sin^2\theta_{13}$ in the normal ordering. The intervals labelled T2K only indicate the measurement obtained without using the external constraint on $\sin^2\theta_{13}$, while the T2K + Reactor intervals do use the external constraint. The star shows the best-fit point of the T2K + Reactors fit in the preferred normal mass ordering. The middle panel shows 2D confidence intervals at the 68.27% and 99.73% confidence level for δCP vs $\sin^2\theta_{23}$ from the T2K + Reactors fit in the normal ordering, with the colour scale representing the value of the likelihood for each parameter value. The lower panel shows 1D confidence intervals on δCP from the T2K + Reactors fit in both the normal (NO) and inverted (IO) orderings. The vertical line in the shaded box shows the best-fit value of δCP , the shaded box itself shows the 68.27% confidence interval, and the error bar shows the 99.73% confidence interval. It is notable that there are no values in the inverted ordering inside the 68.27% interval. [8]

Recently the T2K Collaboration found some evidence of neutrino Neutrino CP symmetry violation, by measuring the difference of the oscillation probability for neutrinos and antineutrinos [8]. For the first time, T2K has disfavored almost half of the possible values at the 99.7% (3σ) confidence level, and is starting to reveal a basic property of neutrinos that has not been measured until now, as shown in Fig. 2.7.

2.3.5 Precise measurement of oscillation parameters

The precise measurement of neutrino oscillation parameters is a powerful tool to probe the physics beyond the Standard Model, for example, test the unitarity of PMNS matrix. However, the current results (See Table 2.3) can not meet the requirement. Thus, the precise measurement of neutrino oscillation parameters is one of the important goals of the next generation neutrino oscillation experiments. JUNO will improve the measurement of $\sin^2\theta_{12}$, Δm^2_{21}, Δm^2_{ee} to the world leading levels of 0.7%, 0.6% and 0.5% [1].

Chapter 3

JUNO Experiment

JUNO is a liquid scintillator neutrino experiments with 20 kt target mass. Its large target mass and high energy resolution allow it not only measure neutrino mass ordering but also perform lots of physics studies. In this chapter, I will show the structure of the detectors and the physics results(goals) of JUNO experiment.

3.1 JUNO Experiment

Jiangmen Underground Neutrino Experiment (JUNO) is a next generation liquid scintillator neutrino experiment. The main goal of JUNO is to measure the neutrino mass ordering, while its 20 kton target mass and excellent energy resolution of $3\%/\sqrt{\text{MeV}}$ will allow to study the neutrinos from multiple sources, including solar, geo, supernova, and atmospheric neutrinos.

3.1.1 Requirement and detector design

The main goal of JUNO and the principle to achieve the goal determines its location. JUNO determine the neutrino mass ordering by the fine measurement of the neutrino oscillation spectrum, thus it should be at the maximum oscillation amplitude location from the reactor site, as shown in the right plot in Fig 3.1. Besides, the distance from the detector to the main two nuclear power plant (NPP) should be the same, otherwise

the neutrino oscillation pattern from the different NPP would offset each other. Thus, the location of the detector is determined to be Jinji town, Kaiping city, Jiangmen city, Guangdong province, which is about 53 km both from Yangjiang Nuclear Power Plant (NPP) and Taishan power plant. In addition, JUNO requires low background, thus the detector is about 700 meter underground in order to reduce the effect of cosmic rays.

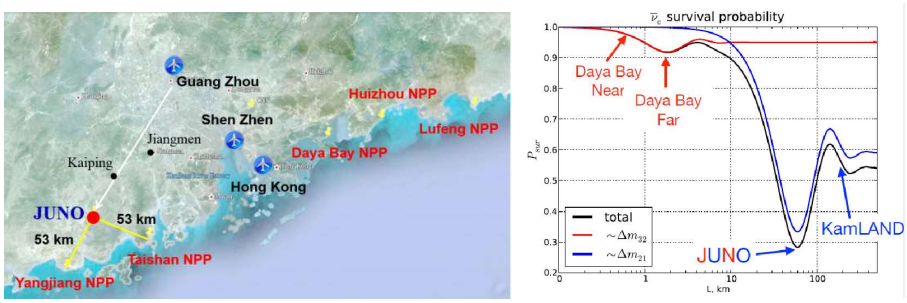


Figure 3.1: The location of JUNO experiment. It is the same distance from Yangjiang and Taishan NPP, and at the maximum oscillation amplitude location of θ_{12} oscillation.

As we discussed in the last chapter, one key point to determine the neutrino mass ordering in JUNO is the statistics, which determines the size of JUNO detector. Due to the long distance from the reactors to the detector, we have to construct the world's largest liquid scintillator detector in order to collect enough statistics. The central detector (CD) will be a huge acrylic sphere with the radius of 17.7 m, containing 20 kton liquid scintillator. In this case, there will be about 83 IBD events from the reactor neutrinos per day, and we can determine the mass ordering at at the confidence level of better than 3σ with 6 years' data taking.

Another key point for the determination of mass ordering is the energy resolution, which sets requirement to the transparency of liquid scintillator and the number of PMTs we use. As we see in Fig. 3.3, in order to determine the neutrino mass ordering at the confidence level of better than 3 σ, we need the energy resolution to be better than $3\%/\sqrt{\mathrm{MeV}}$, which is a big challenge for liquid scintillator experiment. The main factor to determine the energy resolution is the number of photonelectron (pe) detected for the event. Thus, we should use liquid scintillator with high light yield

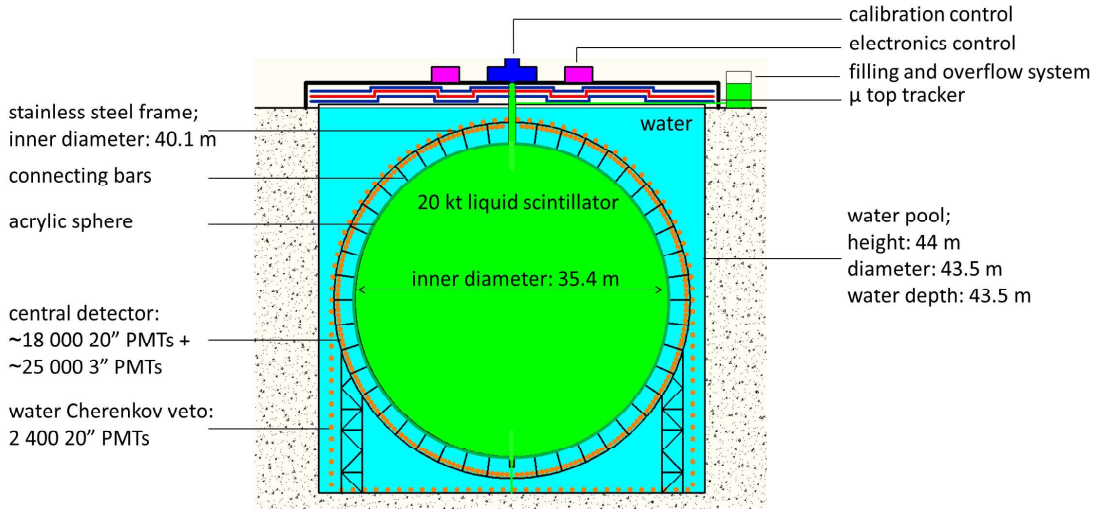


Figure 3.2: Schematic view of the JUNO detector

and long attenuation length, and the PMT photocathode coverage should be as high as possible.

For the liquid scintillator, we use Linear alkylbenzene (LAB) as target, 3 g/L 2,5-diphenyloxazole (PPO) as the fluor, and 15 mg/L p-bis-(o-methylstyryl)-benzene (bis-MSB) as the wavelength shifter. After the purification of liquid scintillator, the light yield can be up to more than 10,000 photons at 1 MeV, and the attenuation length of the liquid scintillator can be longer than 20 m at 430 nm. Thus, there will be about 5000 photons reaching the inner surface of the acrylic ball for 1 MeV event at the center of the detector.

In the experiment, we will use about 18, 000 large (20 inch) PMTs as optical sensor, which will cover about 75% of the detector's inner surface. In addition, we will also use 25, 000 3 inch PMTs which will additionally provide about 3% PMT photocathode coverage. The quantam efficiency of the PMTs will be larger than 35%. In this case, the detection efficiency of the PMTs in JUNO will be around 27%. Thus, there will be about 1400 photons transfer to photonelectrons for 1 MeV event at the center of the detector, which makes $3\%/\sqrt{E(MeV)}$ energy resolution possible in JUNO.

The requirement to tag cosmic muons determine the veto system. The central

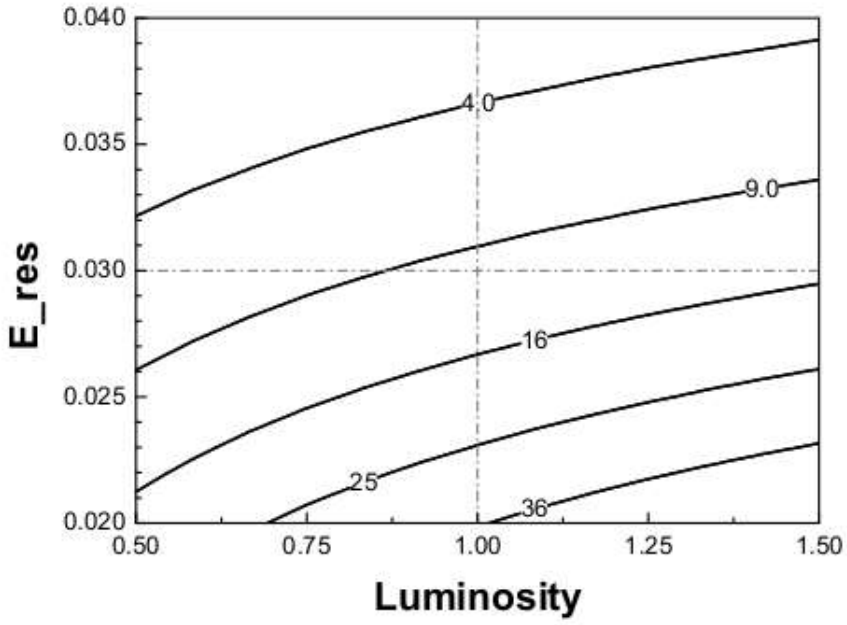


Figure 3.3: JUNO sensitivity vs energy resolution and luminosity. Luminosity =1 means 6 years' data taking. [1]

detector is submerged in the water pool (WP), which is equipped with about 2400 20 inch PMTs, acting as water Cherenkov detector. Besides, the water pool can also shield the central detector from the radioactivity and fast neutron from the surrounding rock. On the top of the water pool, there is another veto system called "Top Tracker (TT)", which can accurately measure the muon tracks.

3.1.2 Signals and Background

For the neutrinos from the reactors, the signals in the central detector are inverse beta decay (IBD) signals:

$$\bar{\nu}_e + p \rightarrow e^+ + n \tag{3.1}$$

The positron deposits its kinetic energy in liquid scintillator and annihilate with electron almost immediately, which forms the prompt signal with the energy of $E_{e^+} =$

E_{ν_e} - 0.782 MeV; The neutron first thermalizes and then is captured by a proton with the mean time of about 200 μs, which emits one 2.2 MeV γ and forms the delayed signal. Thus, this will generate a prompt (e^+) - delayed (γ) coincidence signal.

For IBD signals, there are 4 main backgrounds: accidental background, fast neutron background, ^{9}Li/^{8}He background and ^{13}C(α, n)^{16}O background. In addition, the geo neutrino signals are also IBD signals, which are indistinguishable with the reactor neutrino signals.

Accidental backgrounds

In addition to IBD signals, there are still several kinds of signals in JUNO detector, including radioactive signals, cosmogenic isotope signals, and spallation neutron signals. They happens harmony in time, and are possible to accidentally form prompt - delayed signals. The main part of the signals are radioactive signals, which mainly come from outside of LS region and decrease as radius decrease, while cosmogenic isotope signals and spallation neutron signals are uniformly distributed in the detector. Thus, we can reduce them by fiducial volume cut, energy selection, time coincidence, and vertex correlation of the prompt and delayed signals.

To measure the rate of the accidental backgrounds, we can first measure the rate of both prompt and delayed candidates, and then calculate the probability that two signals randomly satisfied the Δt required for IBD selection. Besides, the accidental backgrounds are uniform in time, thus we can also evaluate them with high accuracy via using prompt and delayed candidates separated by more than 2 ms or 2 meters, which in principle should provide consistent results. The uncertainty of the accidentals is dominated by the statistical uncertainty in the rate of delayed candidates.

Fast neutron background

When the cosmic muons pass the surrounding rock, or the corner of water pool, they cannot be tagged by the veto system. These muons can generate high energy neutrons, some of which pass through the water pool and come into the central detector. These neutrons are called fast neutrons. The neutrons can recoil off protons, which form

prompt signals, and then be captured on protons, which emit gammas and form delayed signals. Thus, they can mimic IBD signals in this way. Since fast neutrons come from outside of the detector, fiducial volume cut can significantly reduce the amount of fast neutron background.

In Daya Bay experiment, scientists found flat energy distribution from 12 MeV to 100 MeV of fast neutron signals [24]. Thus, we can estimate the fast neutron background by the extrapolation of fast neutron spectrum into the IBD energy region.

^{9}Li/^{8}He background

Cosmic muons can interact with isotopes in LS and generate lots of kinds of backgrounds. Two of them are ^{9}Li and ^{8}He, which will mimic IBD signals via β - n decays:

$$^{9}\mathrm{Li} \rightarrow^{8} \mathrm{Be} + n \tag{3.2}$$

$$^{8}\mathrm{He} \rightarrow^{7} \mathrm{Li} + n \tag{3.3}$$

The half life times of ^{9}Li and ^{8}He are 0.178 s and 0.119 s respectively, which are very close to the capture time of neutron in LS. However, they are time and space correlated to their mother muons, thus we can reduce them by optimizing the muon veto cut. The rate of ^{9}Li/^{8}He backgrounds can be evaluated from the distribution of the time since the last muon with the known decay times for the isotopes [54].

^{13}C(α, n)^{16}O background

In liquid scintillator experiments, alphas coming from the decay chain of ^{238}U and ^{232}Th can interact with ^{13}C and generate neutrons:

$$\alpha +^{13} \mathrm{C} \rightarrow^{16} \mathrm{O} + n \tag{3.4}$$

In this reaction, the ^{16}O can be also in ground state or excited state. When the ^{16}O is at ground state, the prompt signal is fast neutron. In JUNO experiment, more than 90% ^{16}O are at ground state [55], which can be discriminate with Pulse Shape

Discrimination (PSD) method, details see in Chapter 6.

IBD signal selection

According to the properties of the background, we can set the criteria to improve the data quality. The primary version is as follows [1]:

- fiducial volume cut R $<$ 17.2 m
- the prompt energy cut 0.7 MeV $< E_p <$ 12 MeV
- the delayed energy cut 1.9 MeV $< E_d <$ 2.5 MeV
- the time cut between the prompt and delayed signal $\Delta T <$ 1.0 ms
- the vertex cut between the prompt and delayed signal $R_{p-d} <$ 1.5 m
- Muon veto cut:
  - for muon taged by Water Pool, veto the whole LS volume for 1.5 ms
  - for good tracked muons in central detector and water Cherenkov detector, veto the detector volume via vertex cut between signal to muon $R_{d2\mu} <$ 3 m and time cut between signal to muon $T_{d2\mu} <$ 1.2s
  - for the tagged, non-trackable muons in central detector, veto the whole LS volume for 1.2s.

Based on MC studies [1], we can get the ratio of the signals and backgrounds passing the criteria. After preliminary calculations, we can know that JUNO will observe 60 IBD events per day, with about 6% backgrounds, as shown in Table 3.1

In addition, JUNO will use 1 GHz Flash ADC to sample and record PMT signals, so we can precisely reconstruct the hit time and charge from each PMT. Based on this detailed information, we are able to do Particle Identification (PID) and know the particle type. In this case, we can additionally remove the backgrounds and improve the data quality. Details see in Chapter 6.

Selection	IBD efficiency	IBD	Geo-νs	Accidental	^{9}Li/^{8}He	Fast n	(α, n)
-	-	83	1.5	5.7×10^4	84	-	-
Fiducial volume	91.8%	76	1.4		77		
Energy cut	97.8%						
Time cut	99.1%			410			
Vertex cut	98.7%	73	1.3	1.1	71		
Muon veto	83%	60	1.1	0.9	1.6	0.1	0.05
Combined	73%	60	3.8				

Table 3.1: The signal and background passing the cuts [1]

3.1.3 Other Physics Purpose

Due to JUNO's large target volume and high energy resolution, it can achieve a lot of physics goals in additon to mass ordering measurement.

Solar Neutrino

The requirement of radioactivity level of LS to measure mass ordering is 10^{-15} g/g for ^{238}U/^{232}Th. However, if we can improve the LS purity to the level of 10^{-16} g/g or even to KamLAND level (10^{-17} g/g), we are able to study solar neutrinos. With JUNO's large statistics, we are able to test the consistency of the standard LMA-MSW paradigm by measuring ^{8}B neutrino flux in the so called "up turn" region. In addition, JUNO could shed light on the metallicity problem with the precise measurement of ^{7}Be and ^{8}B neutrino flux.

Geo Neutrino

Currently we have a good knowledge of total internal Earth heat flow (46 ± 3 TW)[1]. However, the relative contribution of the two main sources of Earth's heat, radiogenic and primordial heat, are highly uncertain because their direct measurement is difficult. Chemical and physical models give estimated ranges of 15–41 TW and 12–30 TW for radiogenic heat and primordial heat, respectively [56]. By the precise measurement of geo neutrino, we can estimate the abundance of ^{238}U and ^{232}Th, and even their ratio by the analysis of geo neutrino energy spectrum. JUNO's large target mass

leads to its unprecedented statistics: 300 to 500 geo neutrinos per year, which allows for detailed study of geo neutrinos [57].

Supernova Burst Neutrino

Supernova burst neutrino is one of the most important goal for all of the neutrino experiments. We will have six detection channels in JUNO, reaching the total amount of o(10^3) to o(10^4) with a typical supernova burst at the distance of 10 kpc, as shown in Table 3.2

Channel	Type	Events for different $< E_\nu >$ values		
		12 MeV	14 MeV	16 MeV
$\nu_e + p \rightarrow e^+ + n$	CC	4.3×10^3	5.0×10^3	5.7×10^3
$\nu + p \rightarrow \nu + p$	NC	0.6×10^3	1.2×10^3	2.0×10^3
$\nu + e \rightarrow \nu + e$	ES	3.6×10^2	3.6×10^2	3.6×10^2
$\nu + {}^{12}\mathrm{C} \rightarrow \nu + {}^{12}\mathrm{C}^*$	NC	1.7×10^2	3.2×10^2	5.2×10^2
$\nu_e + {}^{12}\mathrm{C} \rightarrow e^- + {}^{12}\mathrm{N}$	CC	0.5×10^2	0.9×10^2	1.6×10^2
$\bar{\nu}_e + {}^{12}\mathrm{C} \rightarrow e^+ + {}^{12}\mathrm{B}$	CC	0.6×10^2	1.1×10^2	1.6×10^2

Table 3.2: The rate of supernova neutrinos for supernova at 10 kpc away from the earth [1]

Compared to water Cherenkov detectors (Super K [58], IceCube [59]) and liquid Ar detector (DUNE [60]), JUNO has its advantage detection channel due to its low energy threshold:

$$\nu + p \rightarrow \nu + p \tag{3.5}$$

In this channel, we can measure the total flux of all six flavor (anti)neutrinos with the same cross section and large statistics. Thus, combined the different advantage channel of different experiment, we will be able to study the supernova neutrinos in details.

Diffused Supernova Neutrino Background (DSNB)

Benefitting the large target mass and excellent $\bar{\nu}_e$ tagging and background rejection capabilities of liquid scintillator, JUNO will be able to detect DSNB signal at the

confidence level of 3σ with the typical DSNB parameters with 10 years data taking. Even if we don't detect DSNB signals, we can still improve current limits and exclude a significant range of DSNB parameter space.

Neutrinoless Double Beta Decay (JUNO Xe-LS)

Inspired by KamLAND-Zen [39], we have an idea to utilize JUNO's large LS volume for the search of $0\nu\beta\beta$ [41]. The basic idea is to build a balloon filled with enriched xenon gas dissolved in LS, and insert the balloon into the central region of JUNO LS. The predict sensitivity of JUNO Xe-LS experiment is shown in Fig. 3.4.

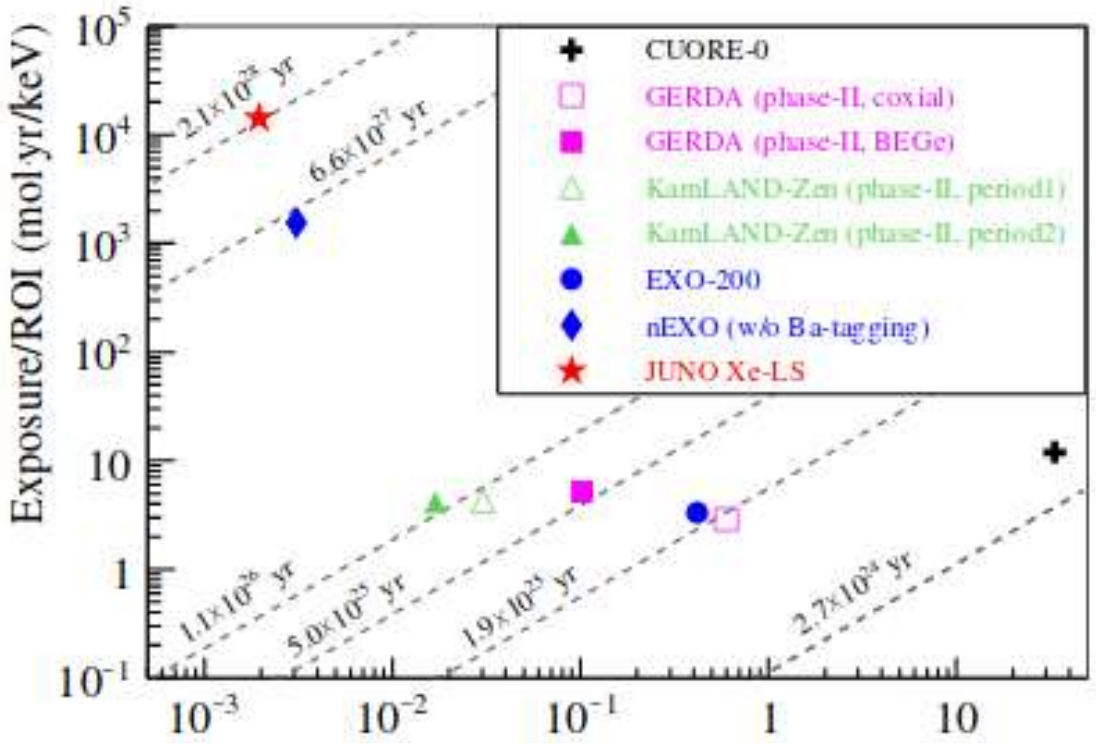


Figure 3.4: The potential Xe-LS detector at JUNO. The dashed lines are the contours of different sensitivities.

Chapter 4

Deep Learning

Deep Learning is becoming more and more popular not only in data science but also in experimental physics. Even in the field of neutrino physics, there are already some successful applications, including NOvA experiment [61][62][63] and MicroBOONE experiment [64][65][50]. Deep learning has show the potential in the following field [66]:

- Simulation
- Real Time Analysis and Triggering
- Object Reconstruction, Identification, and Calibration

The principle of deep learning is simple: we believe there should be some function between the origin information and the features. Mathematician prove that standard multilayer feedforward networks with as few as one hidden layer using arbitrary squashing functions are capable of approximating any Borel measurable function [1] from one finite dimensional space to another to any desired degree of accuracy, provided sufficiently many hidden units are available [67][68]. Thus, the question is, how to find the best groups of parameters.

[1]Definition of Borel measureable function: If $f : X \to Y$ is continuous mapping of X, where Y is any topological space, $(X, \mathcal{B})$ is measurable space and $f^{-1}(V) \in \mathcal{B}$ for every open set V in Y, then f is Borel measurable function.

Although the one-layer neural network with enough hidden units can approximate any Borel function, it is very hard to optimize. Thus, scientists choose neural networks which are "deeper" but "narrower", for the purpose of easier optimization. However, the deeper neural networks usually need larger computing power, and the so called "deep learning" methods become popular with the explosion of computer power.

In this chapter, I will first show the structure of a typical neural network: fully connected neural network. Then, I will interpret how neural network finds the best group of parameters. Last, I will show the popular convolutional neural network (CNN).

4.1 Neural Network

The regular Neural Networks are made of three parts, as shown in Fig. 4.1:

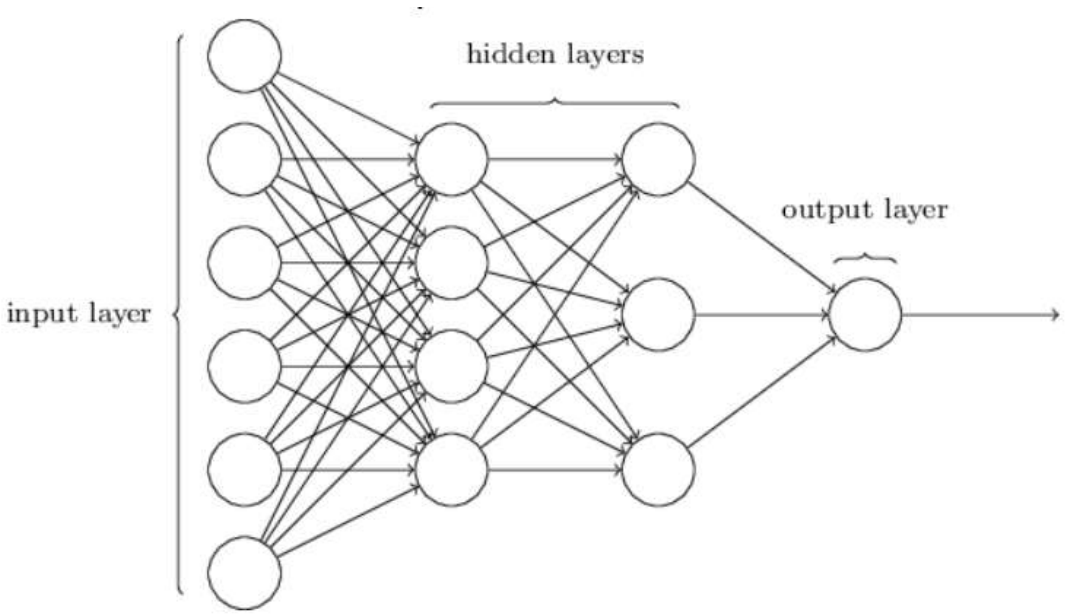


Figure 4.1: Schematic view of a regular Neural Network. It has 3 parts: input layer, hidden layer(s), and output layer
(from http://neuralnetworksanddeeplearning.com/chap1.html).

- Input layer: receive input information and transfer it into neuron network.

- Hidden layer(s): transfer, calculate, and analyze the input information from the previous layer and pass the results to the next layer. There can be more than one hidden layer.
- Output layer: print out the result of the neural network

4.1.1 Forward propagation

The basic unit of neural network is called "neuron", which calculates the inputs X_i from the previous layer, and then transfer the outputs Y to the neurons in the next layer, as shown in Fig. 4.2. The neuron is composed of weights w_i and bias b, performing the calculation as:

$$Y = \Sigma w_i * X_i + b \tag{4.1}$$

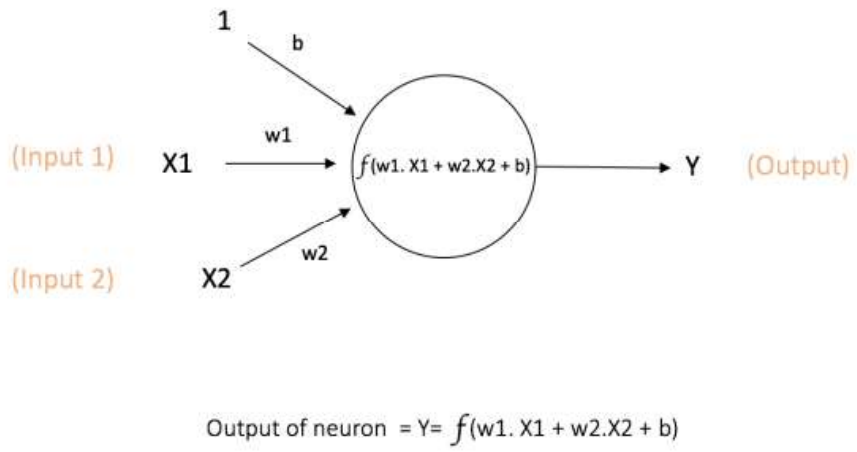


Figure 4.2: Structure of one neuron unit
(from https://ujjwalkarn.me/2016/08/09/quick-intro-neural-networks/).

Thus, the input information is processed layer by layer, and last the neural network gives an output, as shown in Fig. 4.3:

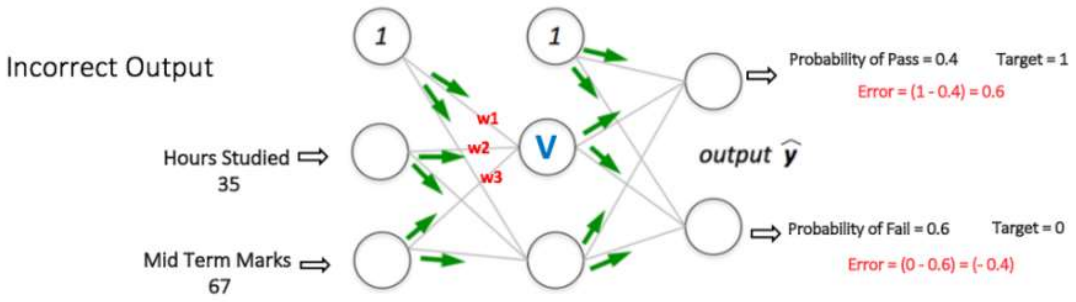


Figure 4.3: forward propagation step in a multi layer perceptron (from https://ujjwalkarn.me/2016/08/09/quick-intro-neural-networks/).

4.1.2 Activation function

For the purpose of non-linear transformation, we use an activation function on the output of each layer, thus the outputs of neuron networks are not simply the linear combination of inputs.

$$Y = \sigma(Y) \tag{4.2}$$

There are many kinds of activation functions, as shown in following parts:

- Sigmoid function:
$$\sigma(x) = \frac{1}{1 + e^{-x}}, \tag{4.3}$$
- Tanh function:
$$tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}, \tag{4.4}$$
- ReLU function:
$$\sigma(x) = max(0, x), \tag{4.5}$$

Due to the simple calculation and good non-linear effect, ReLU function is widely used nowadays. Usually, ReLU function is very effective in most cases. However, there are still some problems:

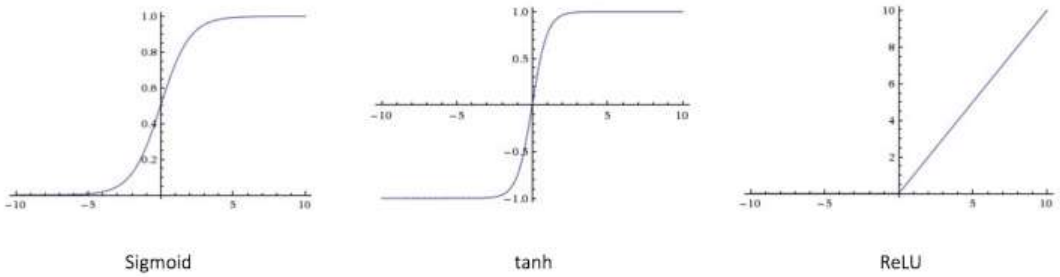


Figure 4.4: Three typical activation functions: Sigmoid function, tanh function, and ReLU function
(from https://ujjwalkarn.me/2016/08/09/quick-intro-neural-networks/).

- the output of ReLU is not zero-centered
- some neurons may never be activated, which is the so called "Dead ReLU problem".

Thus, scientists proposed some variants of ReLU function:

- Leakly ReLU

$$\sigma(x) = \begin{cases} x & x \geq 0 \\ \alpha x & x < 0, \end{cases} \tag{4.6}$$

  where α is determined by users, usually $\alpha = 0.01$

- ELU

$$\sigma(x) = \begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0, \end{cases} \tag{4.7}$$

  where α is learned from data.

4.1.3 Loss function

To evaluate the the performance of Neural Network, we define some loss function to calculate the "distance" between the outputs given by Neural Network and the true values. The goal of Neural Network is to reduce the "distance".

For different questions, there are different kinds of loss function. Here we will show two common loss functions.

One common loss function is "Mean Square Error (MSE)" function. It is often used in regression problems, for example, energy and vertex reconstructions. Giving the Neural Network output x, and true value t, the loss function is defined as:

$$\lambda(x) = (t - x)^2 \tag{4.8}$$

Another common loss function is the so called "Cross Entropy Error Function", which is usually used in classification problem, for example, particle identification. The function is defined as:

$$\lambda(x) = -[t * log(x) + (1 - t) * log(1 - x)] \tag{4.9}$$

Although the above two loss functions can suit for almost all cases, we still need to consider the most suitable loss function for each specific problem.

4.2 Back propagation

Neural Network can do the calculation on input data and give its results, while its output may far from the true values, as shown in Fig 4.5. Thus, we need to "train" the network by reducing the loss function values, so that the Neural Network outputs can be closer to the true values.

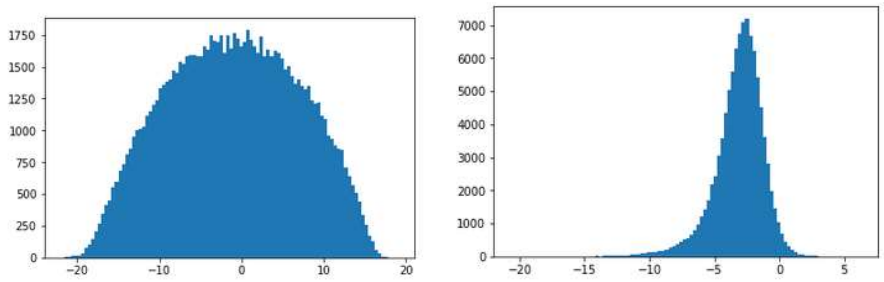


Figure 4.5: Example to show that Neural Network output can differ from true values. left: true values; right: the output of the neural network

The most popular algorithm to reduce the "distance" is backpropagation. The principle is as follows:

We can treat the network as the function between the input $\mathbf{X}$ and the true result $\mathbf{Y}$:

$$\mathbf{Y} = f(\mathbf{w}, \mathbf{b}; \mathbf{X}) \tag{4.10}$$

Usually the initial value of the parameters $\mathbf{w}$ and $\mathbf{b}$ can not lead to good result, as in Fig 4.5. We first find a suitable loss function to evaluate the distance between the neural network output and true result, and calculate the loss function output $loss$:

$$loss = ||Y_{pred} - Y_{true}|| \tag{4.11}$$

The function can be Function 4.8 or 4.9, or some other functions. For different questions, the most suitable loss function is different. Sometimes we need to construct a new loss function to solve the problem.

Thus, the problem is to reduce the variable $loss$. We can calculate the partial derivative of function with arguments, and then update the parameters with the partial derivative of function

$$w_i = w_i - \frac{\partial loss}{\partial w_i}, \tag{4.12}$$

$$b_i = b_i - \frac{\partial loss}{\partial b_i}. \tag{4.13}$$

The update is from the last layer to the first layer, so this process is called "back propagation", as shown in Fig. 4.6:

Each time we update the parameters, the loss will be smaller and the output of neural network will be closer to the true result. Thus, we get better and better groups of parameters. After we finish the training, the group of parameters can be used to predict the result on the new data sets.

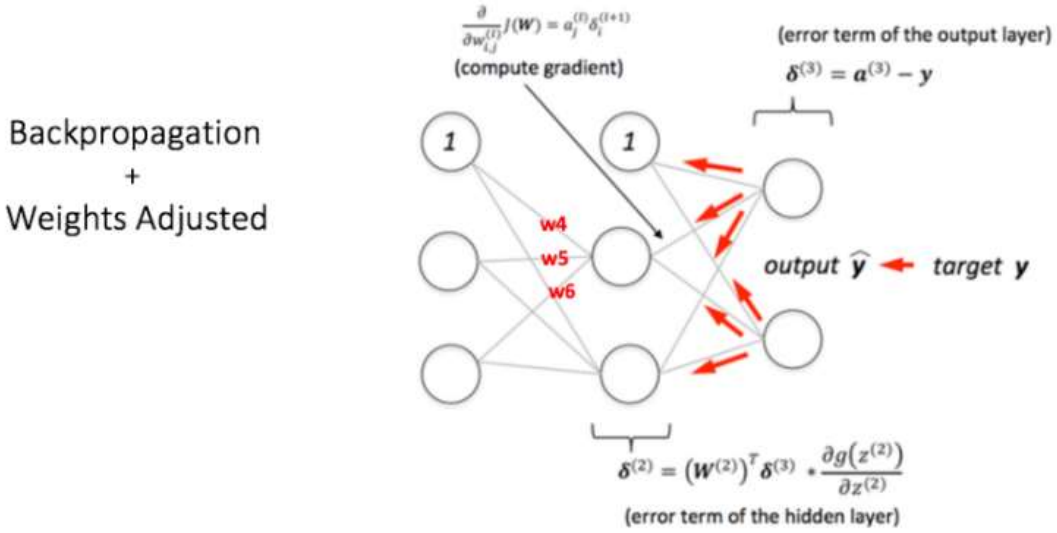


Figure 4.6: backward propagation and weight updation step in a multi layer perceptron
(from https://ujjwalkarn.me/2016/08/09/quick-intro-neural-networks/).

4.3 Convolutional Neural Network (CNN)

The normal fully connected neural networks are very powerful on the small size data, while they are not so suitable for the big size image data:

- The spatial information of input is lost when we transform the images to the vectors,
- There will be too many parameters so that the training will be very difficult,
- The big amount of parameters will lead to overfit.

Thus, the computer scientists developed another class of neural network: Convolutional Neural Network (CNN), which has a salient capability for extracting spatial features and avoid the above problems.

Convolutional Neural Networks are very similar to the ordinary Neural Networks. They are both made up of neurons with trainable weights and biases. The neurons

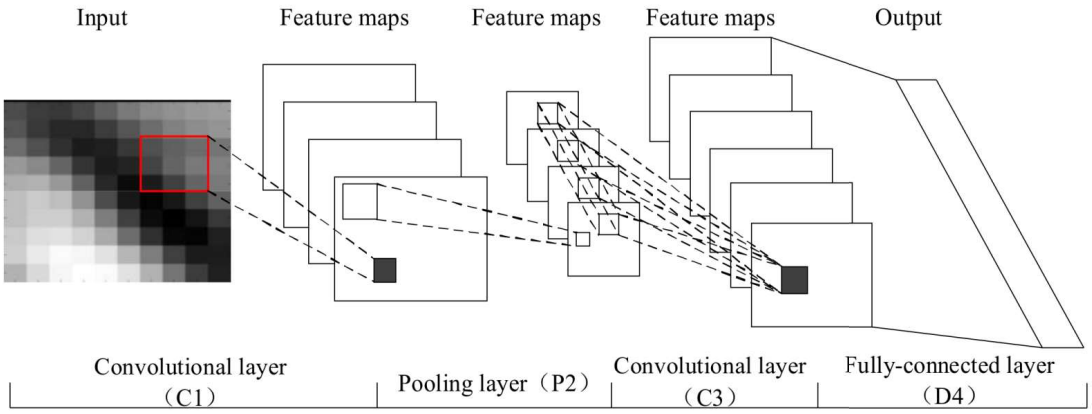


Figure 4.7: Structure of a typical convolutional neural network (CNN) with four layers, including a convolutional layer, a pooling layer, and a fully-connected layer. [10]

receive some inputs, perform dot products and pass the outputs to the next layer. The whole network expresses a single differentiable function and all the strategy we developed for regular Neural Network still apply.

So, what changed? In regular Neural Networks, each neuron is fully connected to all neurons in the previous layer. However, in Convolutional Neural Network, the few layers close to the input are always alternatively convolutional layers, each neuron of which is only connected with a small region of the previous layer, as shown in Fig. 4.7. Thus, the spatial information is utilized and the parameters are reduced.

4.3.1 Convolution layer

Convolution operation is the core operation of CNN, which means the operation of inner product the image and the filter matrix. The convolutional layers are made up of covolutional kernels (filter matrices), which can perform operations such as edge detection, blur, and sharpen. After passing through a convolutional layer, the input images become abstracted to a feature map, with the shape of height x width x depth (numer of channels of the feature map).

The convolutional layers should have three attributes:

- Convolutional kernels (filters) defined by a weight and height,
- The number of input channels and output channels,
- The depth of the Convolution filter (the input channels) must be equal to the number channels (depth) of the input feature map.

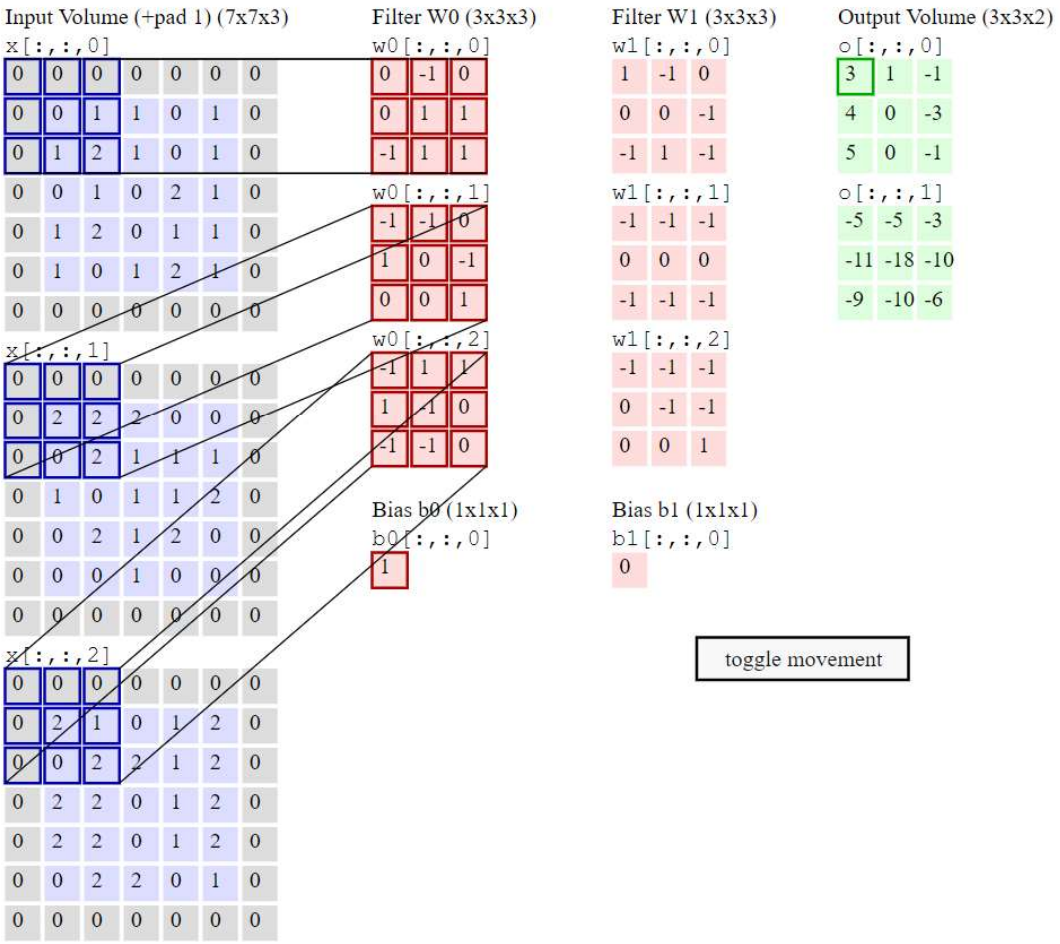


Figure 4.8: Example of convolution operation. Here we have two filters of size 3×3, and a padding of P=1 is applied to the input volume, making the outer border of the input volume zero.
(from https://cs231n.github.io/convolutional-networks/)

In the convolutional layers, each neuron only receives input from a subarea of the previous layer, which called its receptive field. Typically, each neuron can only "see" a very small part of the previous layer (e.g. size 3 x 3), while in a very deep network the neurons in the last layer can "see" very large or even whole field of the input layer.

4.3.2 Pooling layer

Pooling is another important concept of CNN, which plays the role as reducing the spatial size of the representation. Thus, the amount of parameters and the computation of the neural network are heavily reduced and the overfitting is controlled.

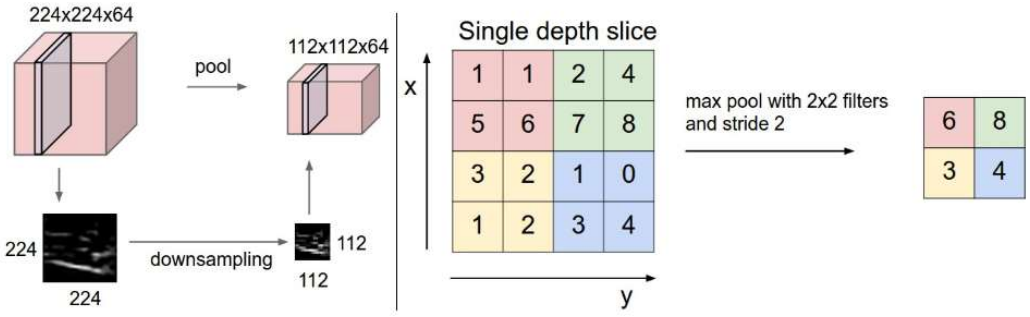


Figure 4.9: Pooling layer downsamples the volume spatially, independently in each depth slice of the input volume. Left: In this example, the input volume of size [224x224x64] is pooled with filter size 2, stride 2 into output volume of size [112x112x64]. Notice that the volume depth is preserved. Right: The most common downsampling operation is max, giving rise to max pooling, here shown with a stride of 2. That is, each max is taken over 4 numbers (little 2x2 square)
(from https://cs231n.github.io/convolutional-networks/)

The pooling layer operates independently on every depth slice of the input and resizes it spatially. In practice, there are two common pooling operations: max pooling and average pooling. The max pooling operation usually works with the filters of size 2x2 applied with a stride of 2 downsamples at every depth slice in the input by 2 along both width and height, discarding 75% of the activations:

$$f_{X,Y}(S) = max(S_{2X+a,2Y+b}), \tag{4.14}$$

where a=0 and b = 1.

The operation of average pooling is almost the same with max pooling, while the operation acts as:

$$f_{X,Y}(S) = mean(S_{2X+a,2Y+b}), \tag{4.15}$$

where a=0 and b = 1.

Currently, we usually use max pooling in our neural network, since it works better in most case.

4.4 Conclusion

In physics experiments, it is important and time consuming to extract the features from the data. Besides, features subtract by human may not comprehensive or even not correct. With the help of neural network, we can spend less time on this part and focus more on our physics purpose. Currently the structure of neural networks remains to be designed by human and the best structure of neural network for the specific problem is still unknown to us. However, there are already some works focusing on the self-designed of neural network. In the future, the network designed by the network can beat the artificial neural network, and we can only focus on the physics questions.

Chapter 5

Waveform Reconstruction

In JUNO experiment, the photons are detected by PMTs and recorded by 1 GHz Flash Analog Digital Convertor (FADC) readout system. Ideally, the precise reconstruction of charge and hit times of incident photons from Flash ADC waveforms would allow us to push the resolutions of energy and spatial reconstructions to their physical limits: a feature helpful to multiple physics purposes. In this chapter, I will discuss the waveform reconstruction with neural networks.

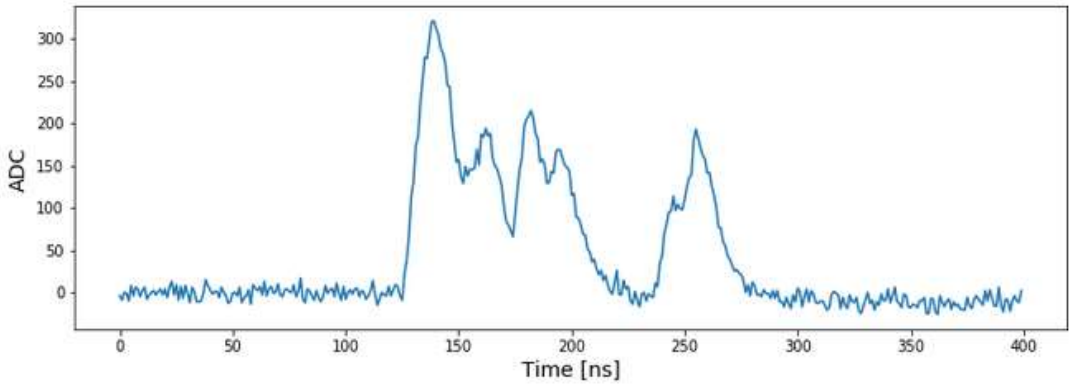


Figure 5.1: One PMT waveform with 12 p.e.

5.1 Traditional waveform reconstruction methods

In Daya Bay experiment, we develop several waveform reconstruction methods, including charge integration, waveform fitting, and deconvolution [2]. All these methods have comparable performance while deconvolution method achieves the best result. Besides, charge integration method could be a good choice for fast preliminary reconstruction due to its fast speed and stability. Since we use the same electronics system in JUNO, these methods can be also used in JUNO experiment.

5.1.1 Charge Integration

The simplest way for waveform reconstruction is charge intergration method. We can do this with three steps:

1) Find the baseline of the waveform: calculate the mean value of the waveform before the first pulse. Usually we take the first 100 ns;

2) Subtract the baseline from the waveform ;

3) Integral the region over zero

The advantage of charge method is obvious: it is fast, robust and simple to understand, which can be seen in Fig. 5.2. However, its advantage is also obvious: for the hits after the main pulse, their charge is under-estimated due to overshoot effect, which is also shown in Fig. 5.2. The residual non-linearity is estimated, with the result shown in Fig. 5.3. In the plot, the X axis is the MC true charge, and the Y axis is the ratio of the reconstructed charge over the MC one. With the increase of MC true charge, the number of hits after the main pulse also increases, inducing more charge under-estimation, which cause the decreasing trend of the non-linearity curve.

In conclusion, charge integration method would be suitable for fast reconstruction, while the residual non-linearity would be as large as 10% due to overshoot effect. If we can re-estimate the baseline of the hits after the main pulse, the residual non-linearity could improved to around 3%.

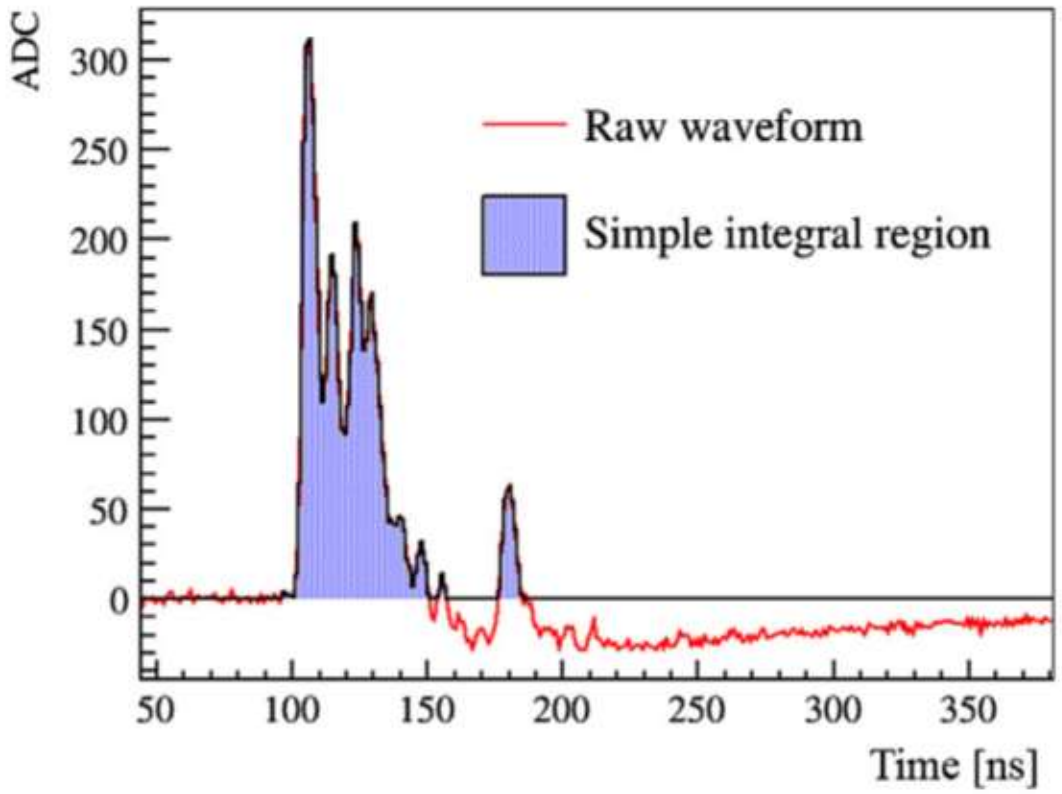


Figure 5.2: Example integral region of the charge integration method. The charge of the hits after the main peak is under-estimated due to overshoot [2].

5.1.2 Waveform Fitting

The waveform is composed of single p.e. pulse, thus we can fit the waveform with the calibrated single p.e. waveform as a template. One good fitting example is shown in Fig. 5.4. With this algorithm we can clearly get the charge and time of each peak, especially for the waveform with peaks well separated. The residual non-linearity of waveform fitting algorithm is about 2%, better than charge integration algorithm. However, this algorithm is not suitable in JUNO experiment due to the two shortages:

- The speed of waveform fitting algorithm is about 0.5 s per channel, and the speed is even slower for the channel with large number of hits. Considering

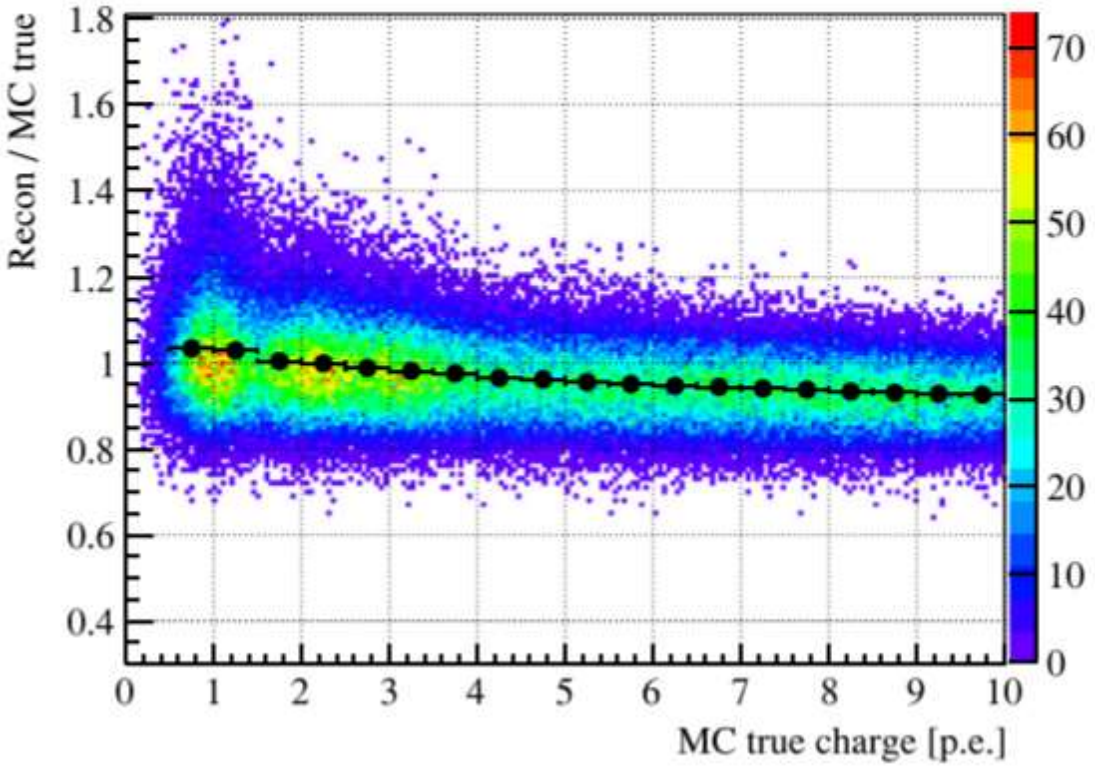


Figure 5.3: Residual non-linearity with the simple integral method in MC, which was about 10% due to the interplay of overshoot and pile-up hits. The color means event number in the certain bin. The X-axis is MC true charge, and the Y -axis is the ratio between the reconstructed and the MC one. The black points show the averaged ratio [2].

large number of channels in JUNO experiment (18000 channels), we need one to several hours to reconstruct one event, which is not affordable for us.

- In some cases the fitting can fail, and the failure rate would increase with the increase of the number of hits, which introduce a large uncertainty in the following reconstruction (energy reconstruction, vertex reconstruction, etc.),

For small detectors with less channels, the waveform fitting algorithm can be a good candidate. However, it is not suitable for experiments with large number of channels, such as JUNO experiment.

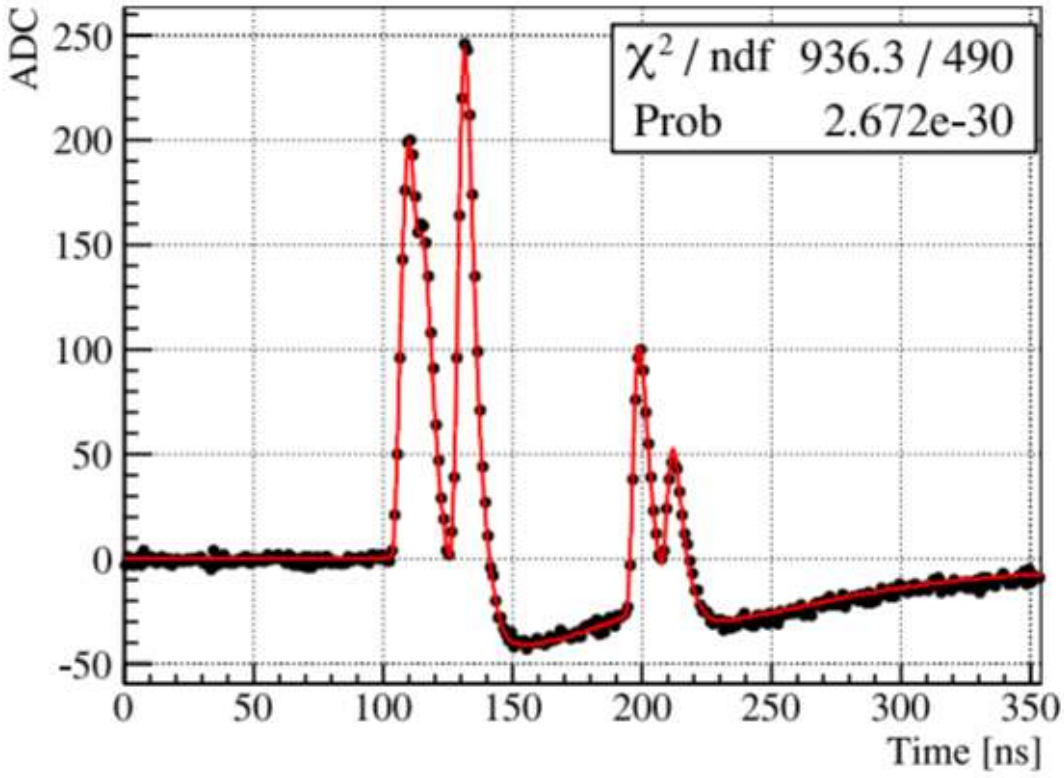


Figure 5.4: An example of waveform fitting. Late hits could be well reconstructed, but the fitting speed was a significant problem with 0.5 s per waveform [2].

5.1.3 Deconvolution Algorithm

Deconvolution is a popular and well developed tool in Digital signal processing (DSP) [69]. Since the PMT waveforms are also Digital signals, we can use this method to reconstruct the PMT charge. The charge is reconstructed with good linearity, and the effect of overlap is largely suppressed.

With this method, we need first convert the raw waveform to the frequency domain with Discrete Fourier Transform (DFT), which is easier to process. Then, we can multiply it with a noise filter, and divide it by the frequency response of a calibrated single p.e. waveform. Last, we convert it back to the time domain with Inverse DFT. One example of the raw waveform and its deconvolution result is shown in Fig. 5.5.

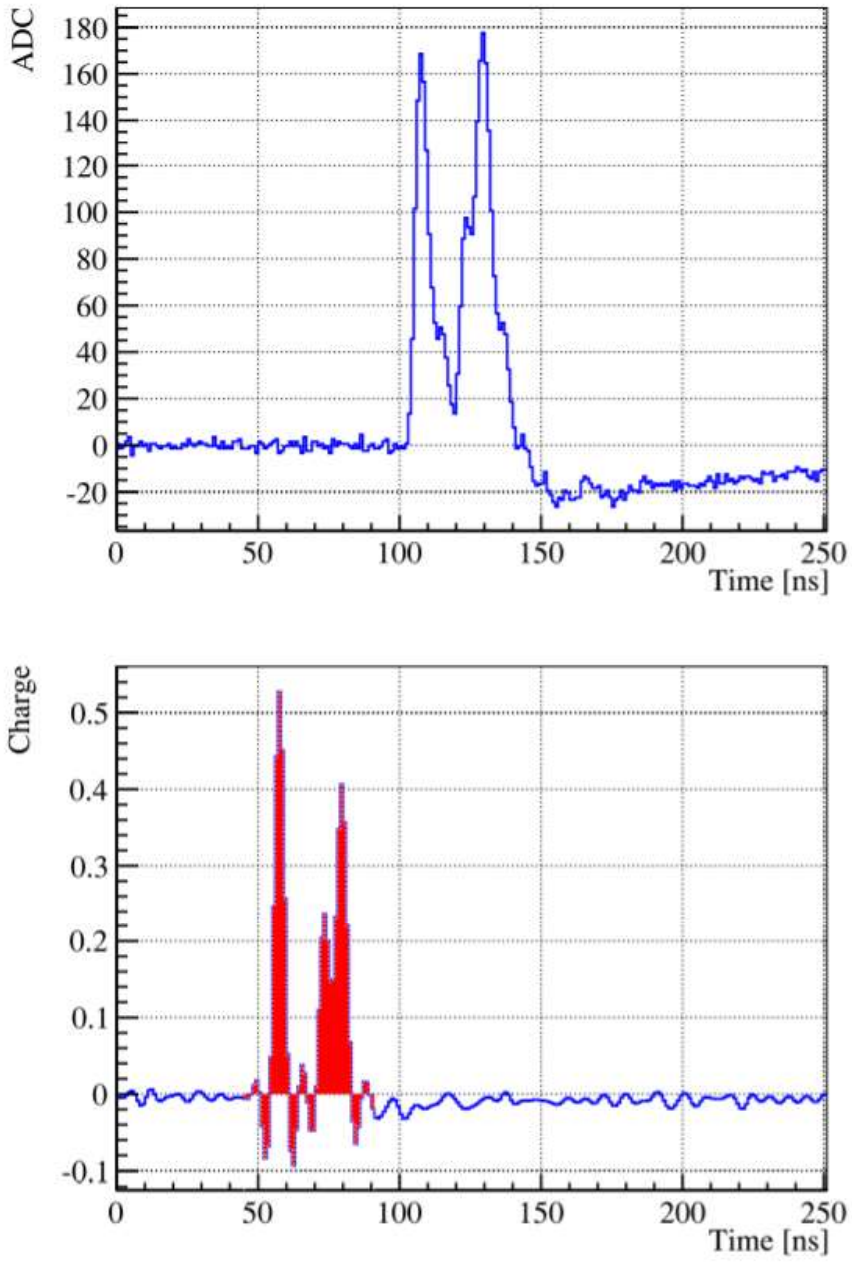


Figure 5.5: An example of a raw waveform (top) and its deconvolution result (bottom). It can be found that the overshoot has been well removed, but local ringing is introduced by the filter, known as Gibbs effects. To reconstruct the charge linearly, the red region is used to do charge integral, which covers the peak and its preceding and subsequent ringing for 9 ns [2].

To reconstruct the charge of the waveform, we just need to simply integral the peak region in the deconvolution result, as shown in Fig. 5.5. However, we see some rings around the main peak, thus the integral region x need to be carefully selected. By testing various x values, when x = 9 we could get the best performance, as shown in Fig. 5.6. In this case, the residual non-linearity is about 1%, which is better than other two methods and fulfill the requirement of JUNO experiment. Besides, the speed of deconvolution algorithm is about 0.5 ms per channel.

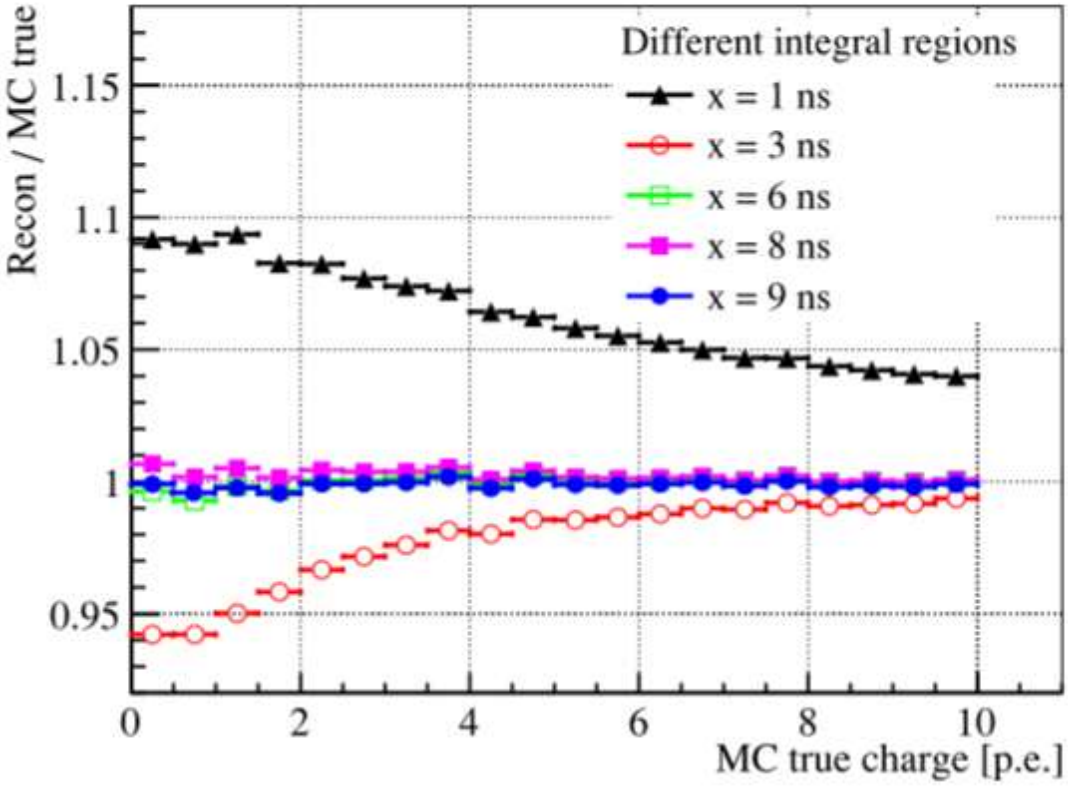


Figure 5.6: The residual non-linearity of the deconvolution method with different x. The blue open circle is x = 9 ns which has the best performance [2].

5.1.4 Summary of the reconstruction algorithms

For waveform reconstruction, several methods are developed to get the charge. The performance of these algorithms is summarized in Table 5.1. Deconvolution algorithm achieve 1% non-linearity, while charge integration algorithm is the fastest.

Algorithms	Speed/channel	Robustness	Non-linearity	Pile-up hits separation
Charge integration	<0.1 ms	No failure	3% to 10%	> 20 ns
Waveform fitting	0.5 s	sometimes fails	2%	>10 ns
Deconvolution	0.5 ms	No failure	1%	>10 ns

Table 5.1: The summary table for different charge reconstruction algorithms [2]

5.2 Reconstruction with Machine Learning

In the above section, we can see that there are several successful algorithms developed for charge reconstruction. However, all these methods can only reconstruct the time for well separated hits. F pile-up hits, they are not able to separate them. Thus, all these methods are not suitable for time reconstruction, especially for high energy events. In some physics studies, we need the time information of all hits for high energy events, for example, proton decay. Thus I develop a machine learning algorithm for this purpose.

Wasserstein distance

For the time reconstruction of pile-up hits, the first question is the standard to evaluate the result. For channel with one hit, the answer is simple: we calculate the distance between true hit time and reconstructed hit time $t_{true} - t_{rec}$. However, for channel with pile-up hits, the question is not so easy. For example, if there are two hits in one channel, first with one p.e. and second with two p.e., and with reconstruction algorithm you get three hits, as shown in Fig. 5.7, then how could you evaluate the reconstruction?

Actually, the real case can be even more complicated, since we usually hard to reconstruct to correct charge. For example, the true charge is 10 and you reconstructed

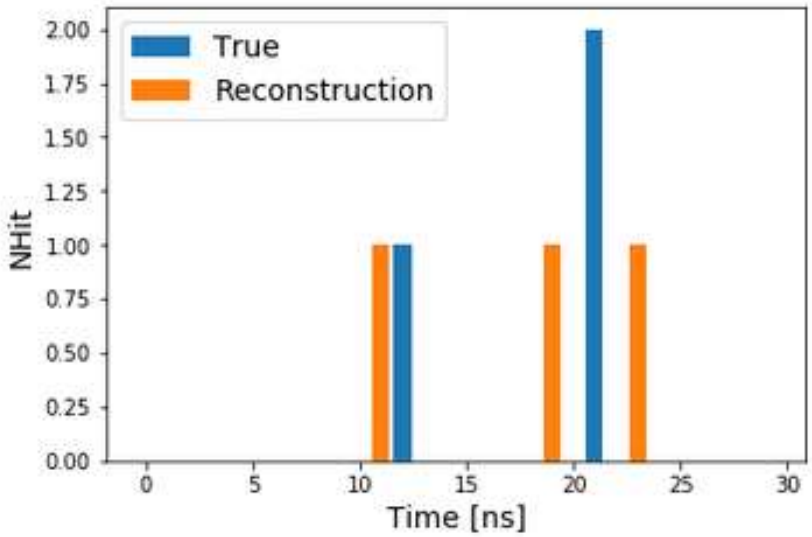


Figure 5.7: An example to show the difficulty to evaluate reconstruction performance on pile-up hits. In this channel there are two hits, first with one p.e. and second with two p.e.; after reconstruction you get three hits. It is hard to evaluate the performance.

charge is 11, and how to evaluate the effect of the extra charge?

Thus, the more reasonable solution is that we can treat the true(reconstructed) hits as a distribution, and then we evaluate the distance between two sets of distributions. Wasserstein distance [70] is the function evaluating the distance between two distributions, thus meets out requirement.

For two probability distributions μ and ν on Polish metric space $\mathbb{R}^d$, the Wasserstein distance of order p is defined by [71] :

$$W_p(\mu,\nu) := \inf_{\gamma\in\Gamma(\mu,\nu)}\left(\int_{\mathbb{R}^d\times\mathbb{R}^d} d(x-y)^p d\gamma(x,y)\right)^{1/p} = \left(\inf_{\varepsilon}\mathbb{E}[d(x,y)^p]\right)^{1/p} \tag{5.1}$$

Our FADC waveforms are 1D curves, and the true(reconstructed) hits are also 1D distributions. Thus, in the above Wasserstein distance function, the dimension $d = 1$, and the function can be simply to:

$$W_p(\mu,\nu) = ||F_\mu^{-1} - F_\nu^{-1}||_p = (\int_0^1 |F_\mu^{-1}(\alpha) - F_\nu^{-1}(\alpha)|^p d\alpha) \tag{5.2}$$

Moreover, in our problem, we need to calculate the Wasserstein distance of order p = 1 at one-dimensional time region t, and function can be simplified to:

$$W_1(\mu,\nu) = \int_{\mathbb{R}} |F_\mu(t) - F_\nu(t)| dt \tag{5.3}$$

In this case, it is easy to evaluate the performance of time reconstruction algorithms with Wasserstein distance, as shown in Fig. 5.8.

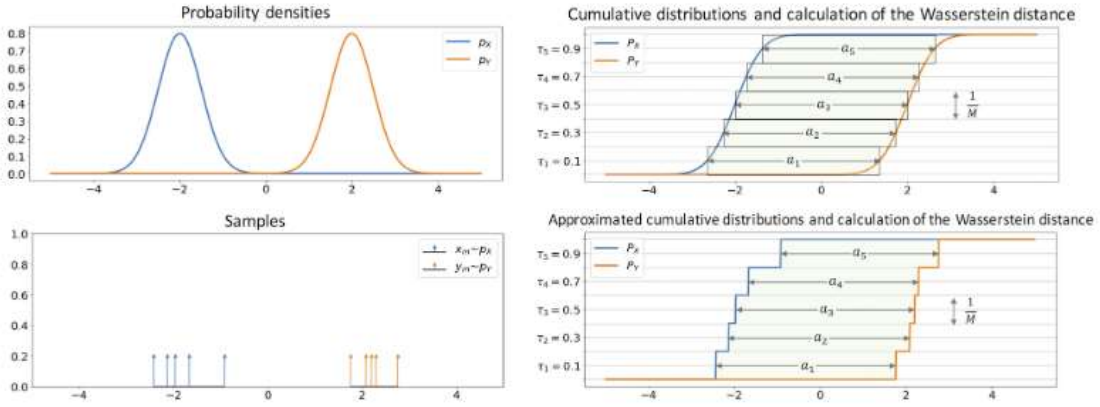


Figure 5.8: One example of the Wasserstein distance for one-dimensional probability distributions. In the top left plot, there are two continuous distribution, and their cumulative distributions are shown in the top right plot. We can calculate the Wasserstein distance by integral the shadow region. The case of discrete distribution is shown in the bottom two plots. [11]

In Fig. 5.8, we show an example of the Wasserstein distance for one-dimensional probability distributions. In the top left plot, there are two continuous distribution, and their cumulative distributions are shown in the top right plot. We can calculate the Wasserstein distance by integral the shadow region. The case of discrete distribution is shown in the bottom two plots.

In the following parts, I will use Wasserstein distance to evaluate the performance of time reconstruction algorithms.

Neural Network Model

When a photon hits PMT, the PMT will generate a photon electron and the photon electron will generate a pulse in the electronics system. Then, the FADC system will record this pulse with 1 GHz sampling rate. In DAQ time window, there will be one or several pulse recorded by the FADC system, and produce the waveform as shown in Fig. 5.9.

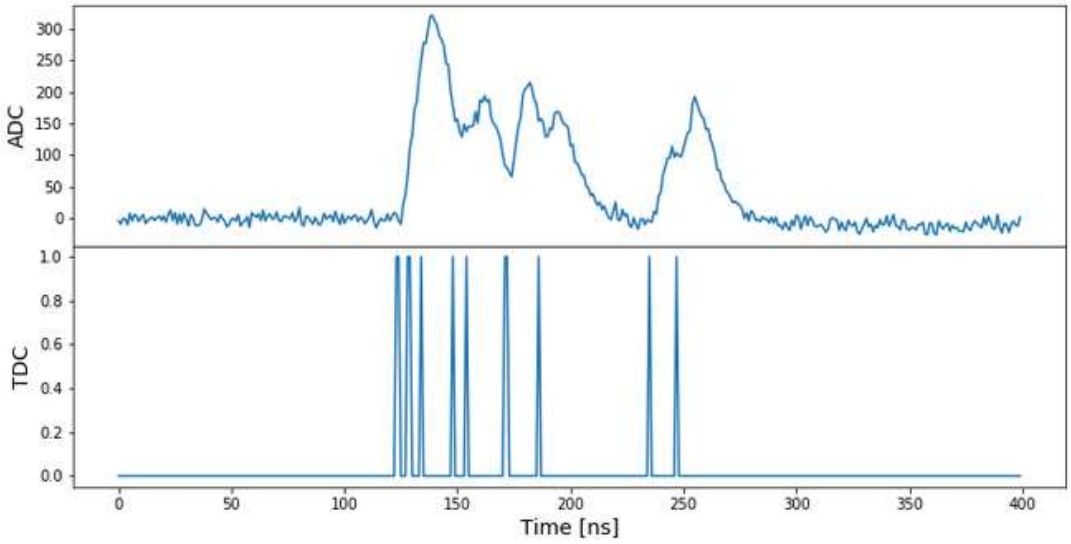


Figure 5.9: PMT waveform and hit time

In this process, I get the basic idea to design the neural network structure: the hits is widen to pulses in the waveform, and I need to reconstruct the pulses back to the hits. In the region of each pulse, there should be some form of transformation to convert the pulse to the hit.

Since the effective regions to reconstruct the hits are the regions of their corresponding pulses, I select one dimension convolutional layer (Conv1D) as the basic module of my neural network. Besides, due to the sparse of the input information, I use Leakly ReLU function instead ReLU function as my activation function. If I use ReLU as activation function most of the neurons will die after a few epoch of training, which is bad for the neural network. The core idea is that I fix the width

of each layer of the neural network, thus the initial information will not distort in the process, and each neuron of the outputs is some kind of convolution of the input waveforms.

```
----------------------------------------------------------------
        Layer (type)               Output Shape         Param #
================================================================
            Conv1d-1              [-1, 4, 1000]              32
         LeakyReLU-2              [-1, 4, 1000]               0
            Conv1d-3              [-1, 8, 1000]             232
         LeakyReLU-4              [-1, 8, 1000]               0
            Conv1d-5              [-1, 8, 1000]             456
         LeakyReLU-6              [-1, 8, 1000]               0
            Conv1d-7              [-1, 8, 1000]             456
         LeakyReLU-8              [-1, 8, 1000]               0
            Conv1d-9              [-1, 8, 1000]             456
        LeakyReLU-10              [-1, 8, 1000]               0
           Conv1d-11              [-1, 8, 1000]             456
        LeakyReLU-12              [-1, 8, 1000]               0
           Conv1d-13              [-1, 4, 1000]             228
        LeakyReLU-14              [-1, 4, 1000]               0
           Conv1d-15              [-1, 1, 1000]              29
             ReLU-16              [-1, 1, 1000]               0
================================================================
Total params: 2,345
Trainable params: 2,345
Non-trainable params: 0
----------------------------------------------------------------
Input size (MB): 0.00
Forward/backward pass size (MB): 0.75
Params size (MB): 0.01
Estimated Total Size (MB): 0.76
----------------------------------------------------------------
```

Figure 5.10: The structure of Neural Network

Based on this idea, I construct the neural network as shown in Fig. 5.10. For each Conv1D module, I set the width as 7. Since the width of single pulse is about 40 ns, I use 8 Conv1D layers in total, so that the receptive field is large enough to

cover one pulse. For the first seven layers I use Leakly ReLU with $\alpha = -0.05$ as the activation function. For the last layer I use ReLU as activation function, since the outputs a all non-negative values. I use Wasserstein distance as my loss function, with its mathematical expression the same with Eq. 5.3.

In General, this network is very light and easy to train. There are 2345 parameters in total and it cost less than 1 minute to train 300,000 channels for one epoch. After about 10 epoch, the network can output very good result. When the model is ready, it can reconstruct about 8000 channels per second on CPU (Intel(R) Xeon(R) CPU E5-1650 v4 @ 3.60GHz) .

Result

To evaluate the performance of the model, I generate another 30,000 channels as test data set. The model works well on the new samples, and the performance is shown in Fig. 5.11.

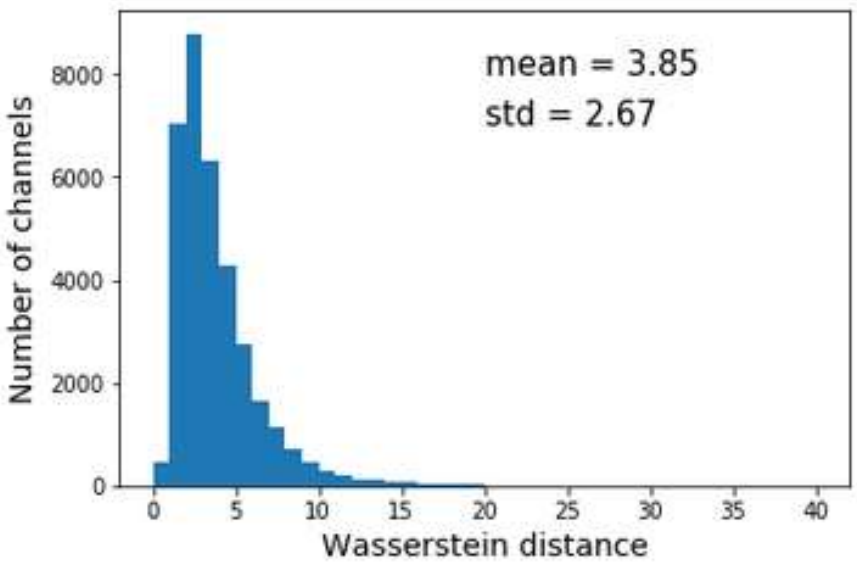


Figure 5.11: The distribution of Wasserstein distance between true hits and reconstructed hits. the mean value of wasserstein distance is 3.85 ns

In the test sample, the mean value of Wasserstein distance is 3.85 ns, which means

the offset of each hits is less than 4 ns. From Fig. 5.11, we can see for the most cases the Wasserstein distance is less than 10 ns, while there are some channels getting very large Wasserstein distance. In this samples, the ratio of the number of channels with Wasserstein distance larger than 20 and the number of all channels is 0.2%.

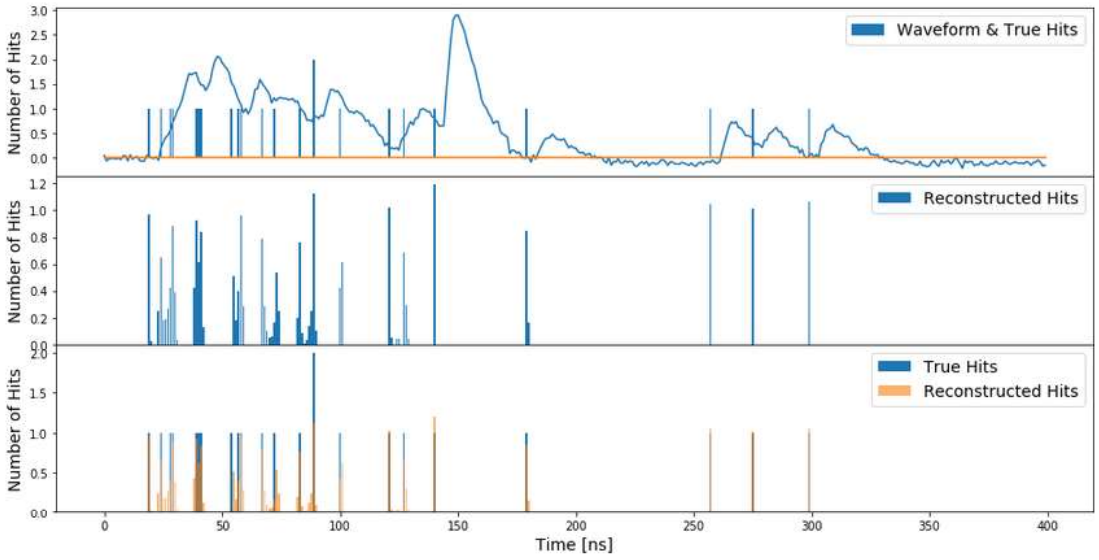


Figure 5.12: One typical reconstructed channel. In this channel the Wasserstein distance is 2.75, and most hits are correctly reconstructed, especially the separated hits.

In Fig. 5.12 I show an example of a typical reconstruction. In this channel there are in total 23 hits, and 20 of them are in the main pulse while the other hits in another pulse. Most of the hits are overlapped and there is obvious overshoot after the main pulse. In this channel the Wasserstein distance is 2.75. In Fig. 5.12 we can also see that for the separate pulses the reconstruction is quite good while for the overlapped pulses the reconstruction becomes bad. Additionally, we can also see that in the last three hits which are after the main pulse, the reconstruction is still very good, showing that this algorithm gets rid of the effect of overshoot.

In Fig. 5.13 I show an example of a failed reconstruction. In this channel there are in total 20 hits and 14 of them are in the main pulse. In this channel the Wasserstein distance is 41.08, which is a very bad result. We find that the overlapped pulses

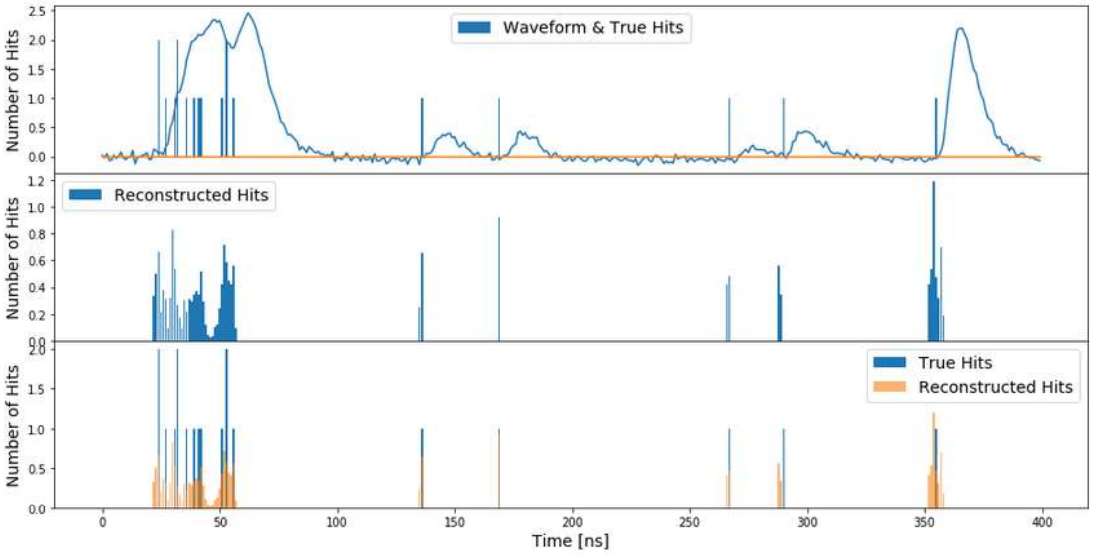


Figure 5.13: One example of a failed reconstruction. In this channel the Wasserstein distance is 41.08, the overlapped pulses are reconstructed to some small values, and even the 4th single peak is not correctly reconstructed.

are reconstructed to some small values, and even the 4th single peak is not correctly reconstructed.

With the above two examples, I think the neural network reconstruct the hit by researching the whole region of the peak and studying the trend of the waveform. The algorithm is not very sensitive to the height of the peak, but it is sensitive to the shape of the peak. In the current electronics samples we know the fluctuation of single p.e. peak is about 30 %, thus the insensitive of the height of pulse is a big advantage of the algorithm. However, considering the complexity of waveform in the real world, the algorithm can fail in some case. In addition, the algorithm will be sensitive with the difference between the MC samples and the real waveforms.

To overcome the shortage of the algorithm so that we can use it in the real case, a larger training sample and more accurate MC sampling will be useful. In the next step I will study to make the algorithm more robust.

I also study the relation between the value of Wasserstein distance and the total charge of the channel, as shown in Fig. 5.14. In the samples the total charge ranges

from around 5 p.e. to 45 p.e.. In the figure we can see that the Wasserstein distance keeps stable with the charge, which reveals the stability of the algorithm. In addition, the Wasserstein distance has the trend to become smaller with the increase of charge. I think the reason is simple: for the high charge case, there are more hits after the main pulse, and the reconstruction of these hits are usually easier and more accurate compared to the case of main pulse, thus lower the value of Wasserstein distance.

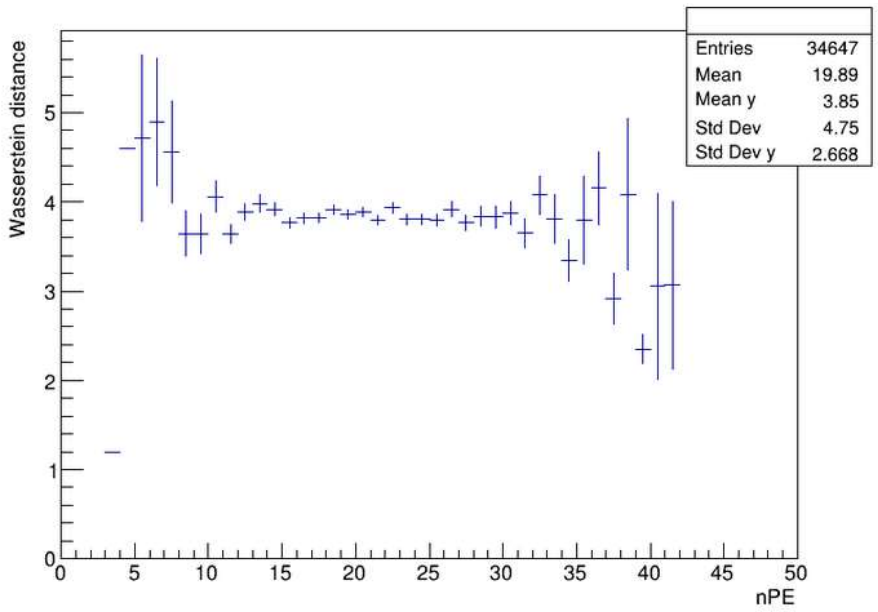


Figure 5.14: The relation between the value of Wasserstein distance and the total charge of the channel. The total charge ranges from around 5 p.e. to 45 p.e., and the value of Wasserstein distance keeps stable with the charge.

In some cases we only need to get the first hit time of the channel, thus I draw the resolution of the first hit time reconstruction, as shown in Fig. 5.15. More than half of the channels we can reconstruct the first hit time correctly, and most of the incorrect reconstructions have the offset of less than 3 ns. In the evaluating samples, the mean value of resolution is -0.49 ns and the standard deviation is 0.9 ns.

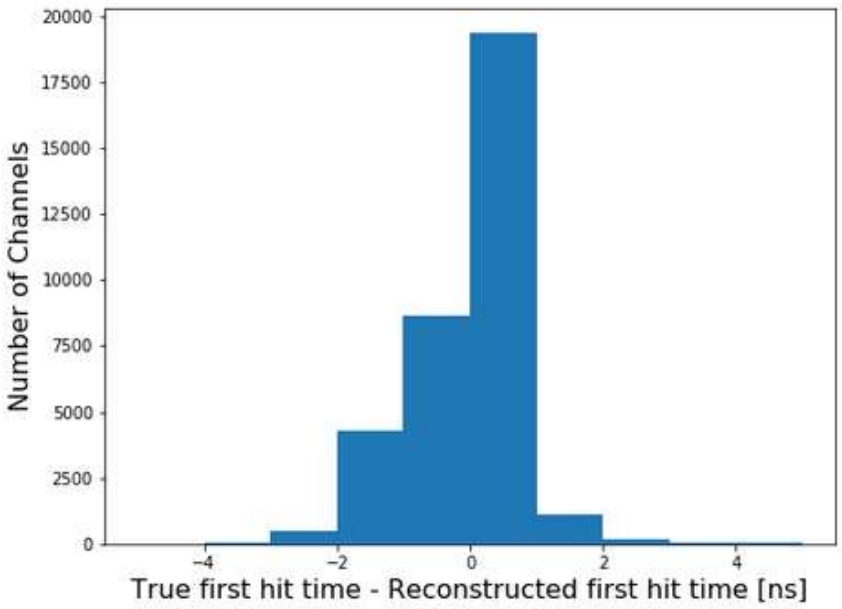


Figure 5.15: Accuracy of the first hit time reconstruction. More than half of the channels are reconstructed correctly, and most of the incorrect reconstructions have the offset of less than 3 ns. The mean value of resolution is -0.49 ns and the standard deviation is 0.9 ns.

5.3 Summary

Charge and time information is especially important in JUNO experiment, and we review some waveform reconstruction algorithms in this chapter. The traditional algorithms have good performance on charge reconstruction, and can also reconstruct the time for well separated hits. However, for pile-up hits they are not able to separate them, thus I develop a machine learning algorithm for this purpose.

JUNO experiment is a multi-purpose neutrino experiments, including the goal of studying solar neutrino, reactor neutrino, geo neutrino, atmospheric neutrino, supernova neutrino, proton decay, etc. In most of these fields, JUNO work in low energy range, and a large fraction of channels only get single hit. For these channels, we just count them as one p.e. and needn't to do reconstruction. For others, in some channels the hits are well separated, but in lots of channels there are pile-up peaks. Currently we only use the time of the first hit, the information of all hit times should

be helpful, especially for events with higher energy and close to the edge.

However, for high energy events such as atmospheric neutrino and possible proton decay events, almost all channels have large number of hits, and time information is important in this case. For these events, traditional methods would be no longer suitable and with machine learning algorithm we can reconstruct the charge and time of all hits with high accuracy.

Chapter 6

Particle Identification

In the physics experiments, the sensitivity is determined by statistics and background/signal (B/S) ratio. Thus, we need to carefully select cuts to increase B/S ratio. In the past liquid scintillator experiments like Daya Bay [24], the events are identified by their energy and vertex, and we cannot know the species of the particles. Although we know that the lighting processes of different kinds of particles are different, we are not able to utilize this information since we cannot get detailed time information of the events in these experiments. In the experiments like Borexino [72] and Double Chooz [73][74], FADC systems are equipped and we can record the detailed time information of each event in the detector with better electronics equipment. Thus, scientists manage to discriminate different kinds of particles.

In JUNO experiment, we use 1 GHz FADC in the electronics system, which allow us to record events' time information with the accuracy of 1 ns. Thus, we are able to study the details of the time profile when particles interact with LS, and do particle identification (PID) to identify the species of the particles with high efficiency.

Based on PID, we are able to select the physics signals with lower background in the analysis of physics data, which is crucial to improve the sensitivity of the experiment.

In this chapter, I will first discuss the principle to do the particle identification in the liquid scintillator, and list the applications of PID in the physics analysis. Then, I will review different methods to do PID, including Tail-to-Total method, Gatti

Particle type	Fast τ_1 / w_1 [ns] / [%]	Slow τ_2 / w_2 [ns] / [%]	Slower τ_3 / w_3 [ns] / [%]
γ, e^+, e^-	4.93 / 79.90	20.6 / 17.10	190 / 3.00
p	4.93 / 65.00	34.0 / 23.10	220 / 11.90
α	4.93 / 65.00	35.0 / 22.75	220 / 12.25

Table 6.1: Time constants τ_i and relative weights w_i assumed for the three exponential contributions to the light emission curves (Eq. 7.1) for different particle types assumed in the JUNO MC simulation.

method, neural network, and topological reconstruction method. I will use Gatti method as baseline in this chapter. Last, I will present and discuss the performance of PID with different methods, including Gatti, topological reconstruction, and neural network.

6.1 Principle

When a charged particles traverse liquid sicntillator, the liquid scintillator will be excited and then emit photons. The decay time distribution of liquid scintillator lighting process can be modeled as:

$$\phi_{\mathrm{em}}(t) = \sum_{i=1}^{n} \frac{w_i}{\tau_i} \mathrm{e}^{-\frac{t-t_0}{\tau_i}} \qquad \text{with} \qquad \sum_{i=1}^{n} w_i = 1, \tag{6.1}$$

In the above function, n is the number of fluorescence components in liquid scintillator, ω is the ratio of different components, τ is the time constant of different components, and t means the time after the start time of the event. With the experience of Daya Bay experiment, we find there are three fluorescence components in LAB, which means n=3. Usually we call the three component "fast", "slow", and "slower" based on the time constant of each component. The value of other parameters in Eq. 7.1 in the current version of JUNO MC code is shown in Table 6.1

With the parameters in Tab 6.1, we can draw the time profile of different kinds of particles and compare them, as shown in Fig. 6.1.

As we see in Table 6.1, the parameters between α's and electron's are very different,

which leads to their very different time profile, as shown in the left plot in Fig. 6.1. In the plots the x axis means the time when the photons hit the PMTs and y axis means the probability for the photons fall into the time bins. In this case, α/β discrimination is not difficult to us.

Regarding e^+ and e^-, the parameters are almost identical. However, positron will form a short-lived meta-state with a local electron called "positronium"(Ps) before annihilation, and then emits two 511 keV gammas. Depending on the spin configuration, the decay time in LS can be either 125 ps (para-Ps) or 3 ns(ortho-Ps). The fraction of ortho-Ps formation was reported to lie around 50% [75, 76, 77]. This effect is visible in particles' time profile, as shown in the right plot of Fig. 6.1. Thus, e^+/e^- discrimination is also possible in JUNO.

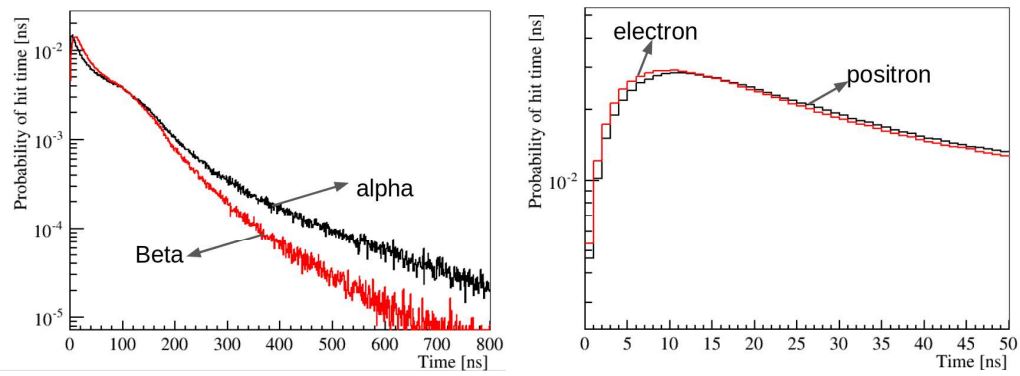


Figure 6.1: The difference of time profile of different kinds of particles. Left: The time profile of α and β. Right: The time profile of e^+ and e^-

6.2 particles in liquid scintillator experiments

Here let's see where the particles come from and the properties of the particles in low energy neutrino liquid scintillator experiments:

alpha

In the experiments the alphas come from the natural radioactivity, usually from the decay chain of U235, U238 and Th232, as shown in Fig 6.2, 6.3, and 6.4. The

energy of alpha events are below 10 MeV, while their visible energy are usually below 1.5 MeV due to the quenching effect in the liquid scintillator.

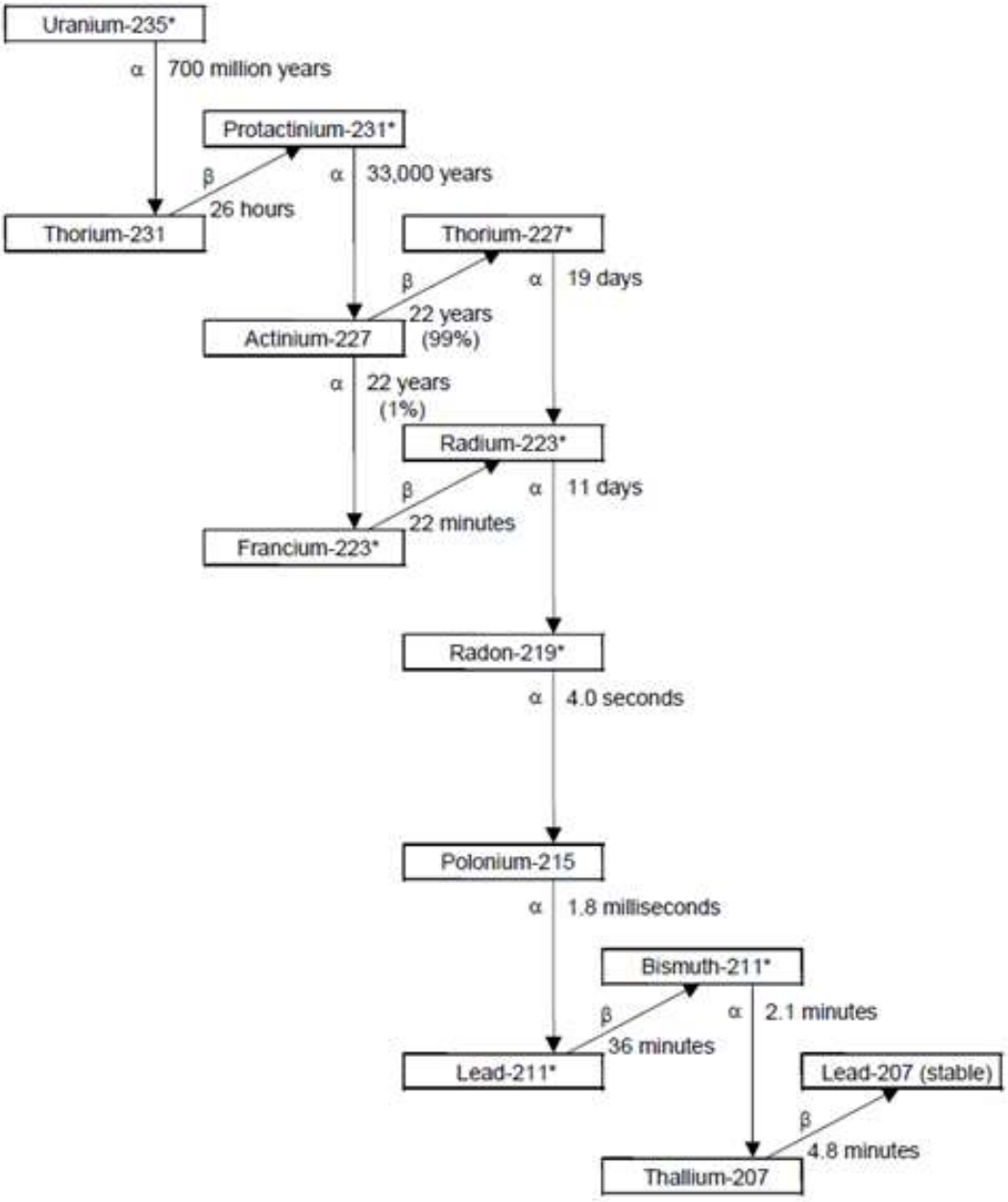


Figure 6.2: The uranium-235 decay chain. The symbols α and β indicate alpha and beta decay with the asterisk indicating if the radionuclide is also a significant gamma emitter. The times shown are the half-lives. Progeny that exist with less than 1% of their parent's activity, as a result of branching, are not shown for clarity. [9]

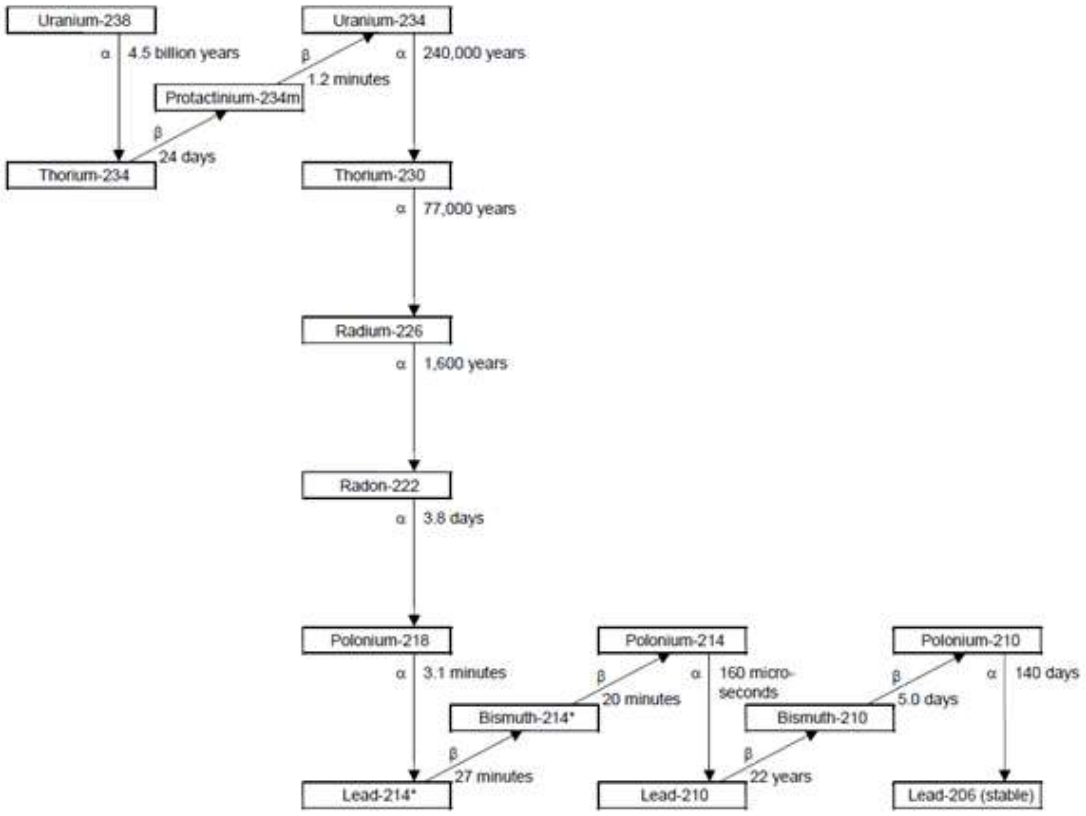


Figure 6.3: The uranium-238 decay chain. The symbols α and β indicate alpha and beta decay with the asterisk indicating if the radionuclide is also a significant gamma emitter. The times shown are the half-lives. Progeny that exist with less than 1% of their parent's activity, as a result of branching, are not shown for clarity. [9]

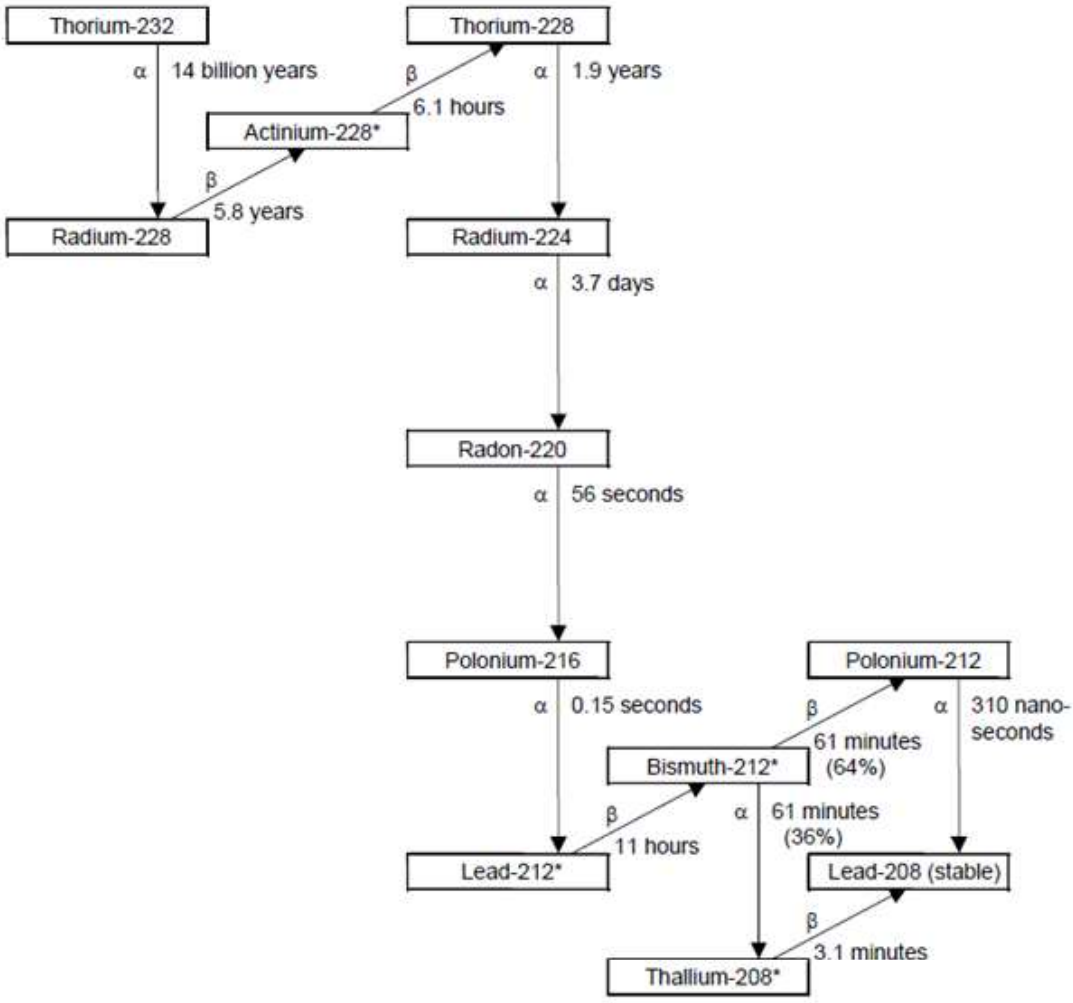


Figure 6.4: The thorium-232 decay chain. The symbols α and β indicate alpha and beta decay with the asterisk indicating if the radionuclide is also a significant gamma emitter. The times shown are the half-lives. Progeny that exist with less than 1% of their parent's activity, as a result of branching, are not shown for clarity. [9]

proton

The proton events in the experiment mainly come from three parts:

α can interact with ^{13}C and generate neutron

$$^{13}C + \alpha \rightarrow ^{16} O + n \tag{6.2}$$

The natural abundance of ^{13}C is 1.1 %. If ^{16}O is at ground state, the kenetic energy of neutron can be up to several MeV. The neutron can knock on proton and transfer its energy to proton. In this case the visible energy of proton events is mainly below 2.6 MeV.

When muons cross the rock outside the detector which cannot tagged by the veto system, they can generate neutrons. These neutrons usually have high energy and are called "fast neutrons". Fast neutrons can cross the water pool , enter the central detector, and scatter with proton which will then record by the detector. The spectrum of fast neutron events is flat, and their visible energy can be up to 100 MeV.

When supernova happens supernova neutrinos can scatter on proton

$$\nu + p \rightarrow \nu + p \tag{6.3}$$

This channel is the golden channel for supernova neutrino detection and is the advantage of liquid scintillator detector. With this channel we can detect all six flavor of neutrinos with large statistics. The visible energy will be below 2 MeV after quanching.

electron

The electron events are the main events in the experiments, which come from several sources:

Electrons can be scattered by neutrino events

$$\nu + e^- \rightarrow \nu + e^- \tag{6.4}$$

The energy of electron events depends on the incident neutrinos. For solar neutrino events the energy of electron is up to several MeV, for supernova neutrino events the energy of electron is up to several tens MeV, and for atmospheric neutrino events the energy of electron can be up to several hundreds MeV.

The radioactive isotopes also generate electron events via β decay, see also in Fig

isotope	half life time	decay mode	Q value
^{12}B	20.2 ms	β	13.4 MeV
^{12}N	11.0 ms	Electron-capture	17.3 MeV
^{11}Be	13.81 s	β	11.5 MeV
^{11}C	20.4 min	Electron-capture	1.98 MeV
^{10}C	19.3 s	Electron-capture	3.65 MeV
^{9}Li	178.3 ms	β/β-n	13.6 MeV / 11.9 MeV
^{9}C	126.5 ms	Electron-capture	16.468 MeV
^{8}He	119.0 ms	β/β-n	10.6 MeV / 8.6 MeV
^{8}Li	838 ms	β	16.0 MeV
^{8}B	770 ms	Electron-capture	18.0 MeV
^{7}B	3.3e-7 fs	proton emission	2.2 MeV
^{6}He	806.7 ms	β	3.5 MeV

Table 6.2: muon induced isotopes and their decay mode

6.2, 6.3, and 6.4.The energy of these electron events are mainly below three MeV.

The high energy muons will generate lots of radioactive isotopes when crossing liquid scintillator detector, as shown in Tab 6.2. Although we veto some time and volume after the muon events, there are still some long half life time isotopes survive.

positron

In liquid scintillator experiments the positron events are mainly come from IBD reaction:

$$\bar{\nu}_e + p \rightarrow e^+ + n \tag{6.5}$$

this is golden channel for $\bar{\nu}_e$ detection duo to the prompt-delayed coincidence. Since the positron will annihilate with electron at rest and emit two 511 keV gamma, the minimum energy of positron events is about 1 MeV. For reactor neutrino events the energy range is 1 to around 10 MeV, while for supernova and DSNB events the upper limit is much higher.

Sometimes radioactive isotopes can also generate e^+ events via $\beta+$ decay, such as ^{10}C and ^{11}C, as shown in Tab 6.2.

gamma

Similar to alpha events, the gamma events are mainly come from the decay chain of radioactivity isotopes. The energy of gamma events are mainly from 0 to 3 MeV.

muon

The muon events in the experiments come from two sources: one part of muon events is from cosmic ray, while the other part come from the CC interaction of muon neutrinos and proton/neutron. However, we will not discuss the identification of muon events in this thesis.

6.3 Possible Application

For the neutrino liquid scintillator experiments, usually the signals are electrons or positrons. However, there are several kinds of particles which can mimic the signals, including alphas, protons, gammas and muons. Thus, we need a good particle identification strategy so that we can get a pure data set for further analysis.

6.3.1 alpha/beta discrimination

It is mainly applied in in the field of low energy neutrino study, such as solar neutrino. Usually we need to select beta events and reject alpha events. In liquid scintillator, alphas mainly come from the decay chain of ^{238}U and ^{232}Th, and their visible energy in liquid scintillator range from 0 to below 1.5 MeV. Thus, we just need to discriminate alpha and beta events in this energy range. Due to the clear difference between alpha and beta time profile, the performance of alpha/beta discrimination is usually excellent.

6.3.2 e/p discrimination

Here there are two possibilities:

1. We want to select e^+/e^- events, and reject proton events. For example, the discrimination between IBD events and ^{13}C(α, n)^{16}O events.

2. We want to select proton events, and reject e^+/e^- events. For example, the discrimination between νp channel and νe channel of supernova neutrinos.

In the liquid scintillator, the visible energy of most proton events are mostly below 2 MeV, thus we set the energy range as 0 to 2 MeV in our study.

The performance of e/p discrimination is similar to alpha/beta discrimination, due to the similar time constants between alpha's and proton's.

6.3.3 e^+/e^- discrimination

Here we need to consider two possibilities:

1. We want to select e^+ events, and reject e^- events. The most important application is to discriminate IBD events and ^{9}Li/^{8}He events. Due to the large energy range of both these two particles, we need to do e^+/e^- discrimination at the energy range of 1 to 10 MeV.

2. We want to select e^- events, and reject e^+ events. his is the case of solar neutrino study, while we want to select electron elastic events and reject ^{10}C and ^{11}C events. Due to the low energetic of solar neutrino events, we just need to discriminate them in the energy range of 0 to 2 MeV.

e^+/e^- discrimination is more difficult than alpha/beta discrimination due to their similar time profile. However, it is still possible in JUNO.

6.4 PID Methods

There are already several methods for PID in the current liquid scintillator experiment.

6.4.1 Tail-To-Total method

Tail-To-Total method is the most simple method for the purpose of particle identification. The principle is simple: we select the time t, count the number of hits after t N_t and the total count of the event N, then we can get the ratio R:

$$R = \frac{N_t}{N} \tag{6.6}$$

Tail-to-Total method just need the number of hits and doesn't need further processing. Although the discrimination ability is not as good as the other PID methods, it is suitable for online classification due to its fast speed and stable performance.

6.4.2 Gatti Method

Gatti method is implemented in Borexino experiment [72], with the principle as follows: First we should get the time profile of each kinds of particle with large statistics, and calculate the probability that a photon electron is detected between time t_n and t_{n+1}, and then obtain the parameter $r_i(t_n)$ representing hit time distribution

$$r_i(t_n) = \int_{t_n}^{t_{n+1}} P_i(t)dt, (i = 1, 2) \tag{6.7}$$

where $i = 1, 2$ represent two kinds of particles. Then we can calculate the weights $w(t_n)$:

$$w(t_n) = \frac{r_1(t_n) - r_2(t_n)}{r_1(t_n) + r_2(t_n)}, \tag{6.8}$$

Now given the time profile of one event $r'_i(t_n)$ and weights $w(t_n)$, we can get the Gatti parameter as follows: which depend on the particle type:

$$G = \Sigma_n r'_i(t_n) w(t_n), \tag{6.9}$$

which depend on the particle type.

Generally Gatti filter has good performance on alpha/beta discrimination and e/p discrimination, while it is almost impossible to discriminate electron and positron. Thus I will use it as our baseline for alpha/beta discrimination and e/p discrimination in this thesis, due to its simplicity and stability .

6.4.3 Neural Network (NN)

Since the performance of Gatti filter is not good enough for the problems like e^+/e^- discrimination, we need to consider much more powerful methods. For this purpose, I introduce the fancy neural network method. Finally, neural network method achieve excellent performance on e^+/e^- discrimination. Besides, it works better on the alpha/beta discrimination and e/p discrimination.

In principle, Gatti filter is linear transformation of the time profile, while the utilization of time information is not perfect. In this case, if we can perform more complex transformation on input information, time information can be made the best use of. We can easily do this by adding several hidden layer in the neural network.

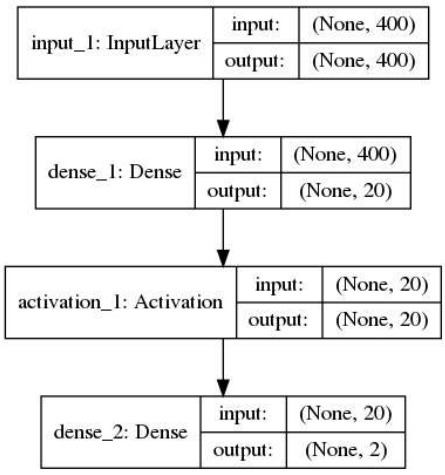


Figure 6.5: Structure of the neural network applied for the particle identification.

I construct the network with the structure shown as Fig 6.5. The input of the network is the 400 bin contents of the time profile, just the same as "Gatti" method, and it has an additional hidden layer with 400 neurons, then it give the output, which tells us what kind of particle it is. We also tried for more hidden layers and different number of neurons in the hidden layer, and find the results don't get better, thus we just keep the simple structure. In principle, "Gatti" filter can also be treated as one kind of neural network with no hidden layer, although the weights of this special "network" is not from training but from calculation.

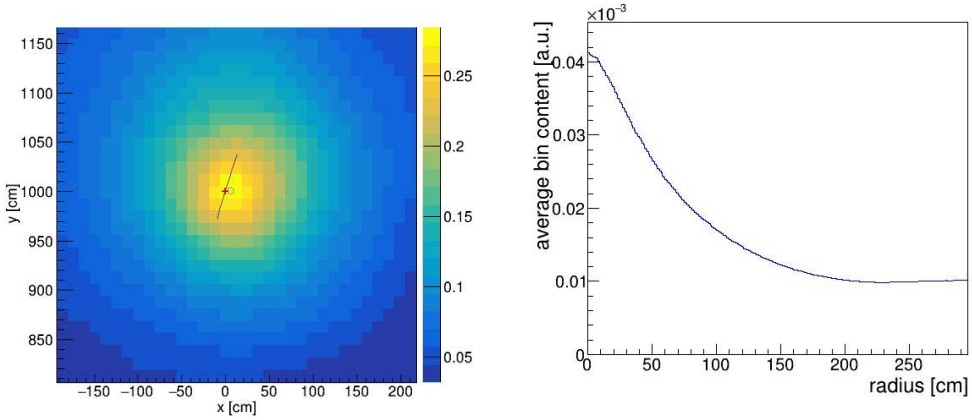
6.4.4 Topological Reconstruction (TR)

In a large, unsegmented LS detector, we can reconstructed the topology of an event with the PMT hit information using a method described in [78, 79]. Thus, we can also identify different kinds of particles with their topological information.

Compared to the previous methods, the topological reconstruction requires an additional knowledge of the reference parameters $\mathbf{r}_{\text{ref}}$ and t_{ref}, which denote one point in space and time, respectively, which the primary particle must have traversed. The parameters $\mathbf{r}_{\text{ref}}$ and t_{ref} can be obtained e.g. from an independent vertex reconstruction. The detection time t_{hit} of a scintillation photon produced at a position $\mathbf{r}$ along the particle track and observed as the k^{th} hit on the j^{th} PMT at position $\mathbf{r}_j$ can be expressed by

$$t_{\text{hit}} = t_{\text{ref}} \pm \frac{|\mathbf{r} - \mathbf{r}_{\text{ref}}|}{c_0} + \frac{|\mathbf{r}_j - \mathbf{r}|}{v_{\text{g}}} + t_{\text{s}}. \tag{6.10}$$

The second term represents the flight time of the particle under the assumption that it moves with vacuum speed of light c_0, being subtracted or added depending on the particle reaching $\mathbf{r}$ before or after traversing $\mathbf{r}_{\text{ref}}$, respectively [12].



(a) (b)

Figure 6.6: Topological reconstruction of a simulated positron event with 3 MeV kinetic energy: (a) projection of the emission density on the x-y-plane in arbitrary units and (b) its corresponding radial dependence around the reference point $\mathbf{r}_{\text{ref}}$ [12].

In this thesis all results from TR are taken from [12], and I put these results as a comparison. This method have excellent performance on e^+/e^- discrimination, while its performance on alpha/beta discrimination and e/p discrimination is not as good as Gatti method.

6.5 Event Samples

6.5.1 datasets

In this study, we simulate 120 k events for each group (α/β , e/p, e^+/e^-) uniformly in the detector. I set the energy range of 0 to 1.5 MeV for α/β , 0 to 2 MeV for e/p, 1 to 10 MeV for e^+/e^-, as shown in Table 6.3. I use 100 k of them to optimize the parameters, and the rest to validate the performance. Besides, to study the effect of TTS and dark noise, I prepare three distinct datasets:

- Dataset 1: the pure MC truth data. A full simulation was carried out with the official JUNO simulation software (J18v1r99-Pre1).
- Dataset 2: adding the smearing effect of vertex reconstruction and PMT's transit time spread (TTS). Gaussians were used to smear the vertex point, following a $\sigma/\sqrt{E(\mathrm{MeV})}$ rule, with E denoting the visible energy. We take the standard deviation value of 10 cm for 1 MeV events. Besides, the hit times were smeared with a Gaussian of the width corresponding to the actual TTS values of the JUNO PMTs.
- Dataset 3: adding dark noise based on Dataset 2. The rate of dark noise is set to 30 kHz for all large PMTs. The direct comparison between Datasets 2 and 3 can reveal the impact of dark noise on our discrimination methods.

6.5.2 figure of merit

We define some figures of merit to understand and evaluate the results:

Particle	α/β	e^-/p	e^+/e^-
Energy [MeV]	[0.2, 1.5]	[0, 2.0]	[1.0, 10.0]
Position	uniformly in the whole central detector		

Table 6.3: Energy range and position distribution of the simulated data samples.

1) discrimination efficiency ϵ_{sig}: the ratio of the number of signal events passing a cut N_{sig} and total amount of signal events N_{total_sig}

$$\epsilon_{sig} = \frac{N_{sig}}{N_{total\_sig}}, \tag{6.11}$$

2) impurity ϵ_{bkg} as the ratio of the number of remaining background events N_{bkg} after the cut and total background events N_{total_bkg}

$$\epsilon_{bkg} = \frac{N_{bkg}}{N_{total\_bkg}}, \tag{6.12}$$

In this thesis, I will prepare 3 kinds of plots to show the performance of different methods:

- Impurity vs efficiency in a fixed energy and radius range. I will show three plots, representing three different discrimination methods.
- Efficiency/impurity vs energy at a fixed level of ϵ_{bkg} and ϵ_{sig}. Here I will show three plots, each contains the performance of thee method on dataset 1/2/3.
- Efficiency/impurity vs radius at a fixed level of ϵ_{bkg} and ϵ_{sig}. Here I will also show three plots, each contains the performance of thee method on dataset 1/2/3.

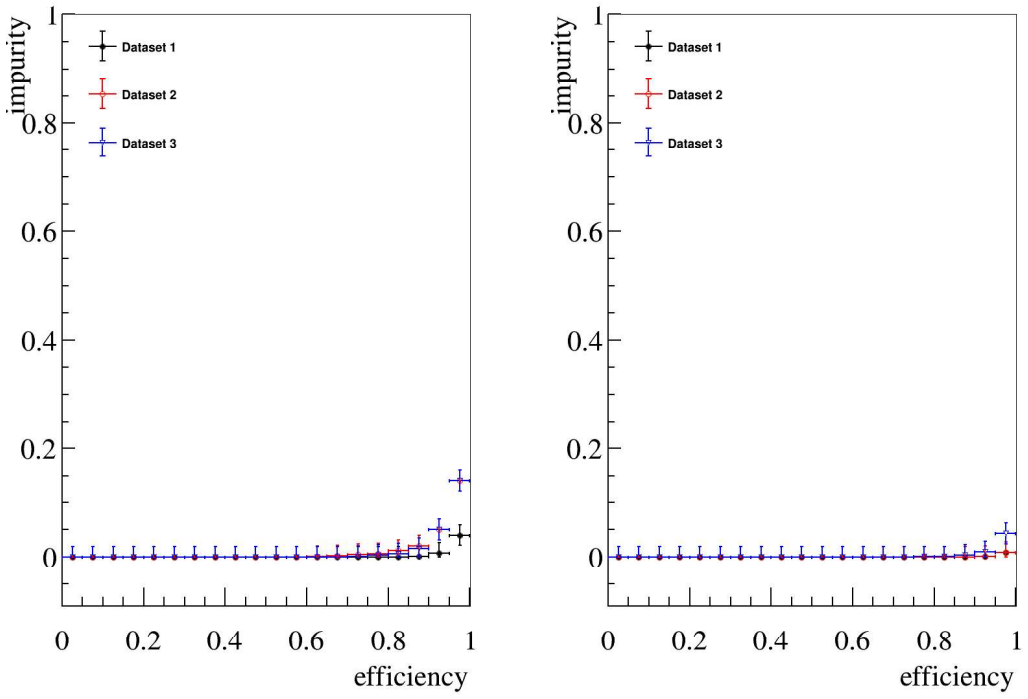
The discussion in this thesis is limited to the most realistic Dataset 3.

6.6 Performance in JUNO

6.6.1 α/β discrimination

Generally, α/β discrimination is not hard in JUNO experiment, due to the large difference between the time constant of α and β. We use the results of Gatti filter as our baseline, and find performances of neural network is better.

Here we obtain level of background impurity as a function of signal efficiency, as shown in Fig. 6.7 with the Gatti (a), NN (b):



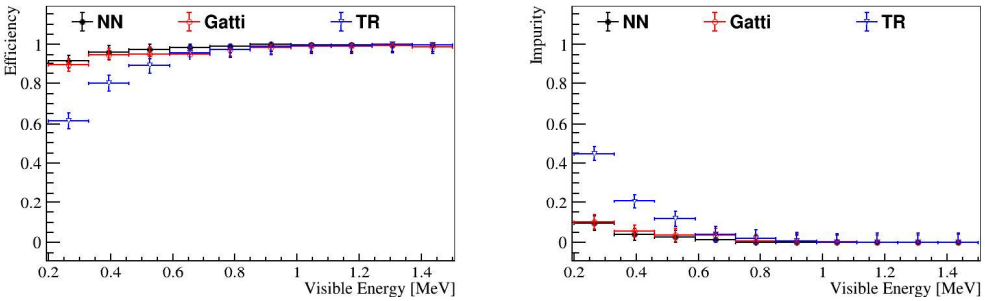
(a) Gatti (b) NN

Figure 6.7: Impurity as a function of efficiency for α/β. The results were obtained for visible energies between 1.0 MeV and 1.5 MeV.

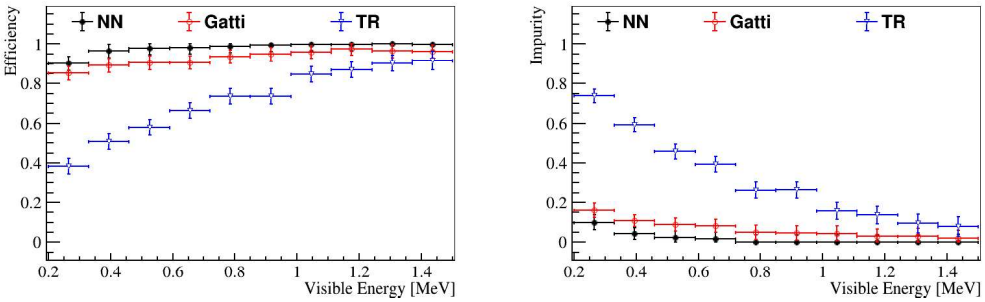
In Fig. 6.7, we select the events with visible energies between 1.5 MeV and 2.0 MeV. In each plot, there are three curves representing three datasets. We can find that both

Gatti and NN method achieve excellent performance on event classification. In the NN case, we can receive above 95% signal with less than 5% background contamination for Dataset 3.

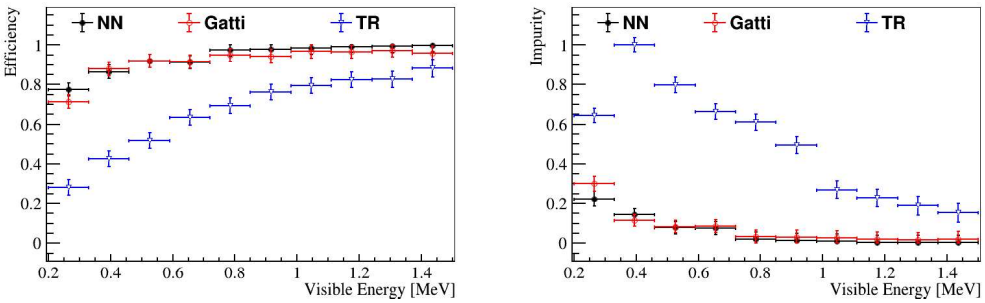
Additionally, we see the curve of dataset 1 and dataset 2 almost overlap, which means that TTS and vertex smear effect is not important in alpha/beta discrimination. Besides, the difference between curve of dataset 2 and dataset 3 is obvious, meaning that the crucial role of dark noise in alpha/beta discrimination.



(a) Dataset 1



(b) Dataset 2

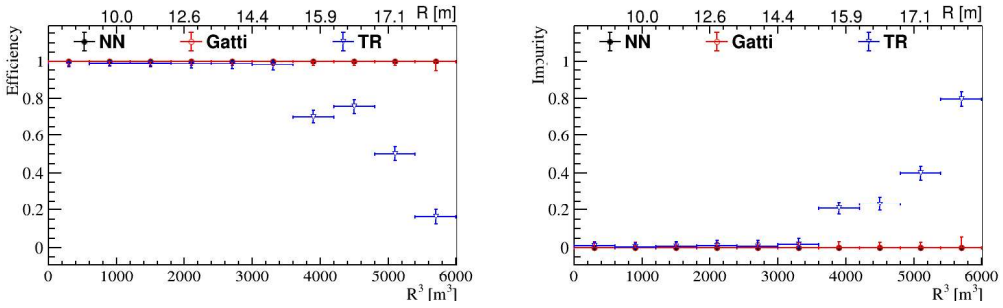


(c) Dataset 3

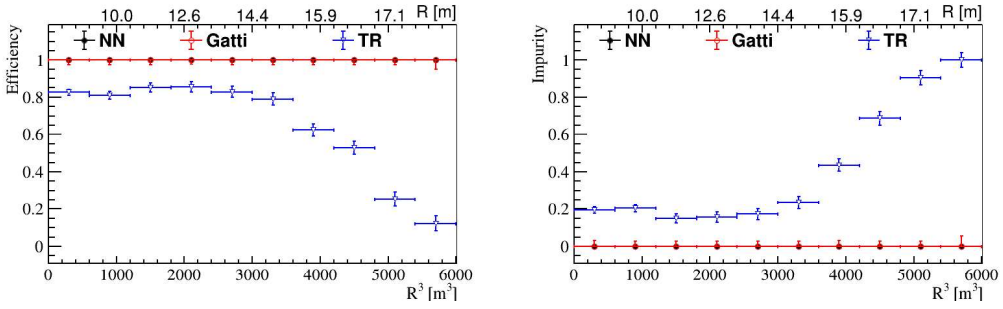
Figure 6.8: Performance of the α/β discrimination from all three methods. Impurity was obtained at efficiency fixed to 90% while efficiency was obtained at impurity fixed to 10 %. (a), (b), and (c) show results with Datasets 1, 2, and 3, respectively, as a function of visible energy.

In Fig. 6.8 we demonstrate a direct comparison between all three methods. Panels (a), (b), and (c) show impurity and efficiency as a function of energy for Datasets 1,

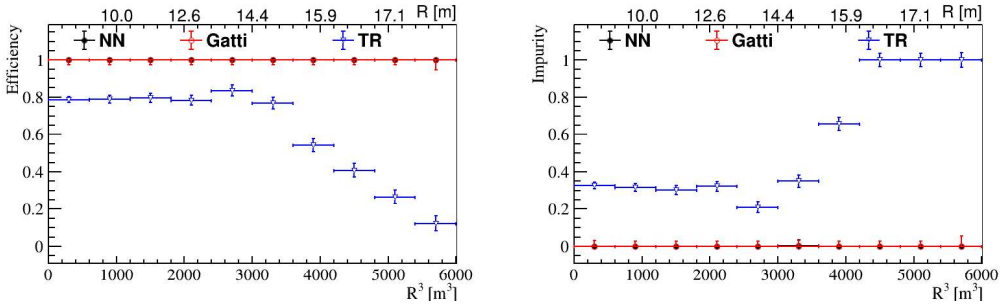
2, and 3. Efficiency was determined at a fixed impurity level of 10%. Impurity was determined with the required efficiency set to 90%. It is clear that as the performance become better as energy increases, due to the increase of p.e. statistics. For events above 0.8 MeV, 100% efficiency is achieved given 10% impurity. Even for energy as low as 0.2 MeV, we can still achieve nearly 80% efficiency given 10% impurity.



(a) Dataset 1



(b) Dataset 2



(c) Dataset 3

Figure 6.9: Performance of the α/β discrimination from all three methods. Impurity was obtained at efficiency fixed to 90% while efficiency was obtained at impurity fixed to 10 %. (a), (b), and (c) show results with Datasets 1, 2, and 3, respectively, as a function of R^3.

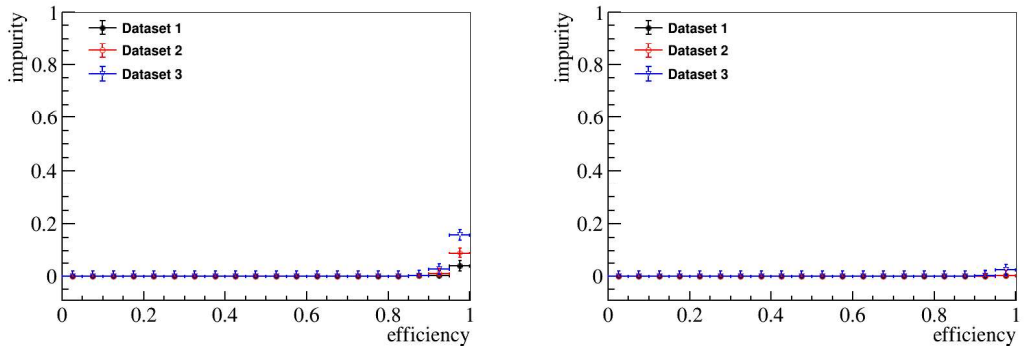
In Fig 6.9 we show impurity and efficiency as a function of R^3 for Datasets 1, 2, and 3. Energy is fixed to 1 to 1.5 MeV. We are glad to see that Gatti and NN method

both have excellent performance in the whole detector.

6.6.2 e/p discrimination

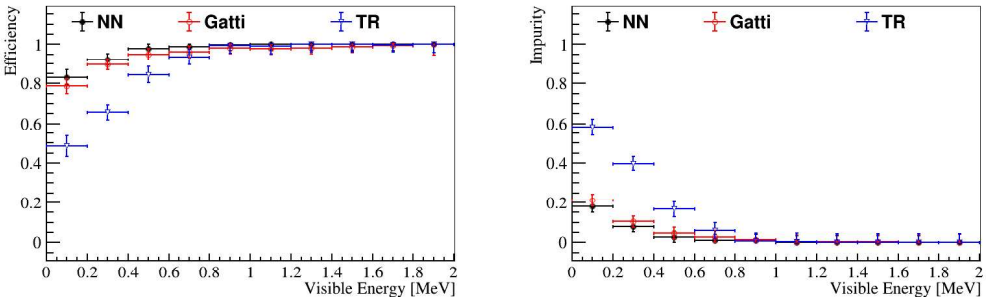
In general, the results of e/p discrimination is very similar to the case of α/β discrimination. In the NN case, we can receive above 95% signal with less than 5% background contamination for Dataset 3.

The performances of e/p discrimination are put in Fig 6.10, 6.11, and 6.12. We have good performance even at the energy just above JUNO's energy threshold (default 0.2 MeV), which is import for this topics since most signals in supernova neutrinos' pES channel are at the energy range of below 0.5 MeV.

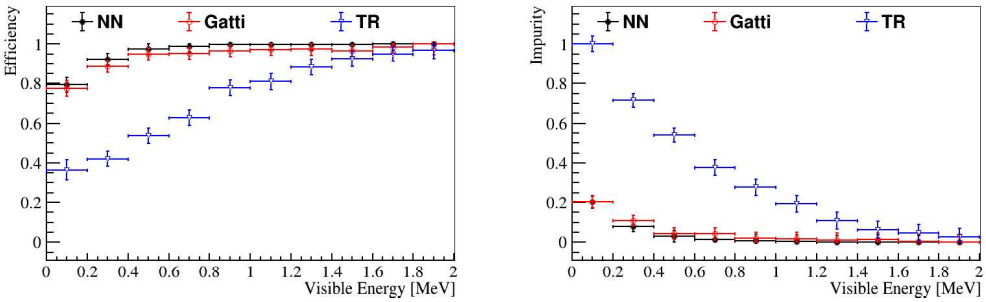


(a) Gatti (b) NN

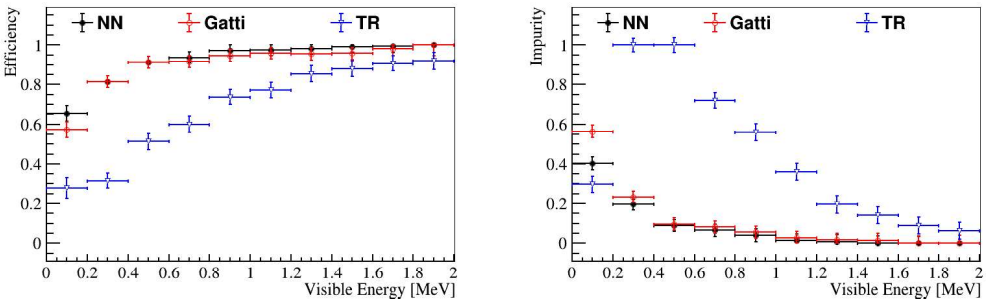
Figure 6.10: Impurity as a function of efficiency for e/p discrimination. The results were obtained for visible energies between 1.5 MeV and 2.0 MeV.



(a) Dataset 1

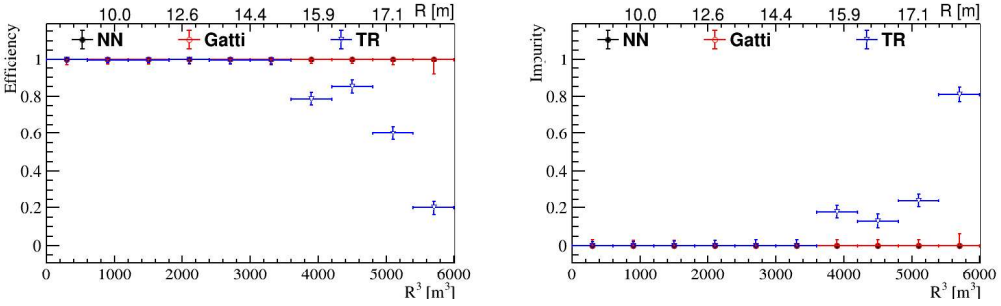


(b) Dataset 2

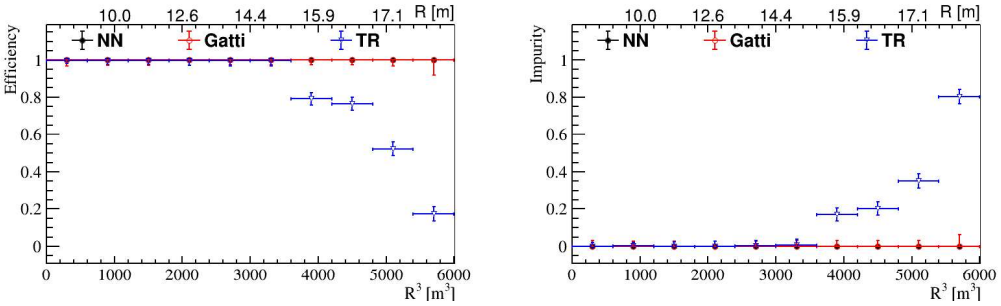


(c) Dataset 3

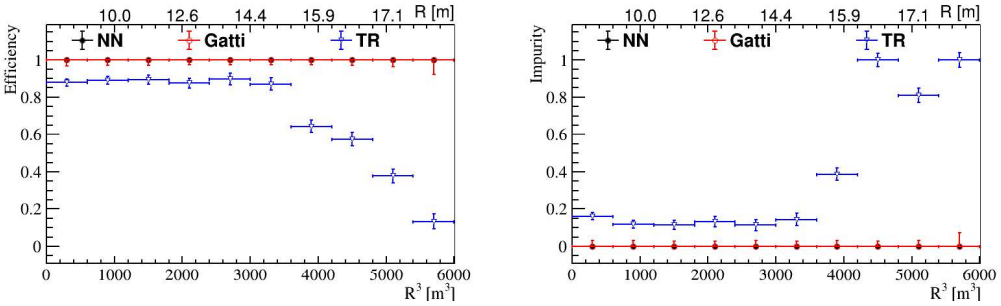
Figure 6.11: Performance of the e/p discrimination from all three methods. Impurity was obtained at efficiency fixed to 90% while efficiency was obtained at impurity fixed to 10 %. (a), (b), and (c) show results with Datasets 1, 2, and 3, respectively, as a function of visible energy.



(a) Dataset 1



(b) Dataset 2



(c) Dataset 3

Figure 6.12: Performance of the e/p discrimination from all three methods. Impurity was obtained at efficiency fixed to 90% while efficiency was obtained at impurity fixed to 10 %. (a), (b), and (c) show results with Datasets 1, 2, and 3, respectively, as a function of R^3.

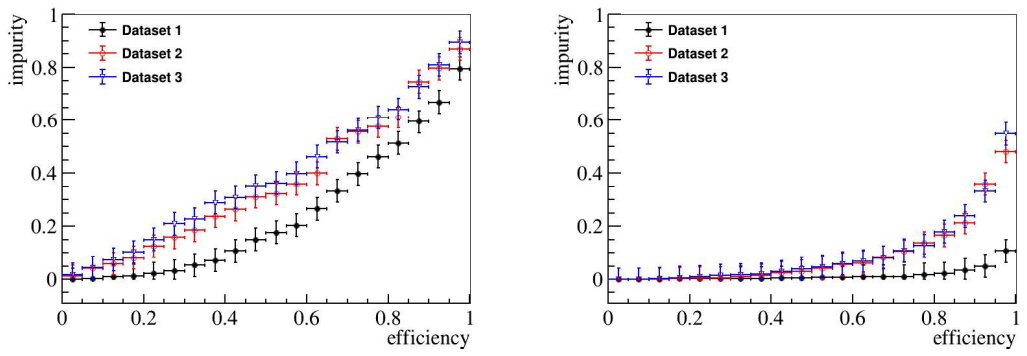
6.6.3 e^+/e^- discrimination

The performance of e^+/e^- discrimination show some different behavior with that of α/β discrimination, since the difference of the time profile between positron's and electron's is small, and the difference mainly located in a narrow region. Here we will discuss this part.

Firstly, Gatti filter is not suitable for this topic anymore, as shown in Panel (a) of Fig. 6.13. Actually, we are almost impossible to discriminate electron and positron with Gatti filter. However, for NN, we are still able to get 50% events with background contamination at about 5%, as shown in Panel (b) of Fig. 6.13.

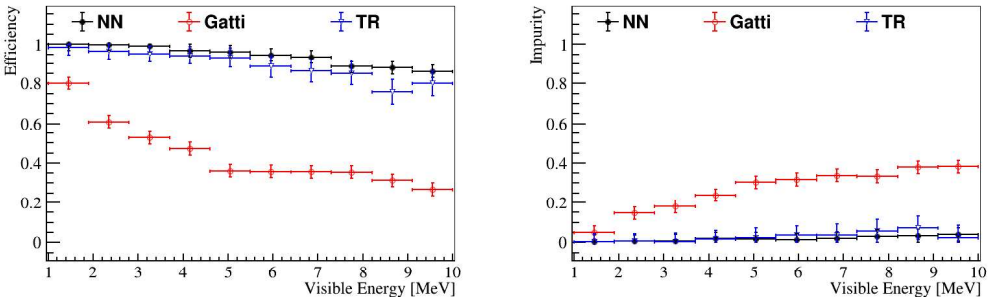
From Fig. 6.13, we see a clear difference between the performance of dataset 1 and dataset 2. This shows that the ability to discriminate electron and positron is determined by the TTS effect of PMT. In the future experiments, if we can reduce timing uncertainty, we can have better potential for electron/positron discrimination.

In addition, the difference between curve of dataset 2 and dataset 3 is small, meaning that the rate of dark noise is not so important in alpha/beta discrimination. This is because that the difference of electron's and positron's time profile is located at their peak region. In this case, the effective statistics of dark noise is largely reduced.

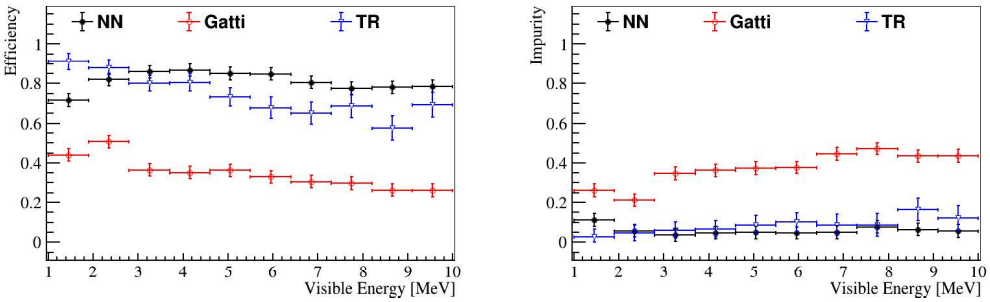


(a) Gatti (b) NN

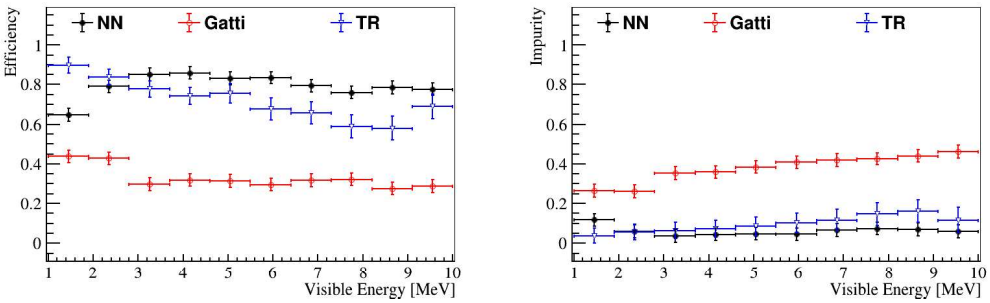
Figure 6.13: Impurity as a function of efficiency for e^+/e^- discrimination. The results were obtained for visible energies between 2.75 MeV and 3.25 MeV.



(a) Dataset 1



(b) Dataset 2

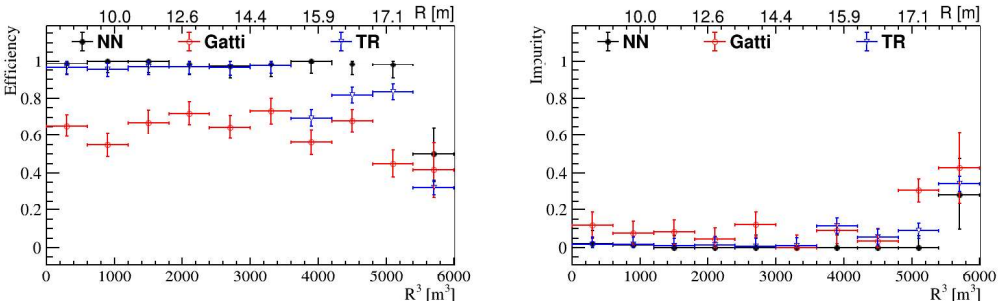


(c) Dataset 3

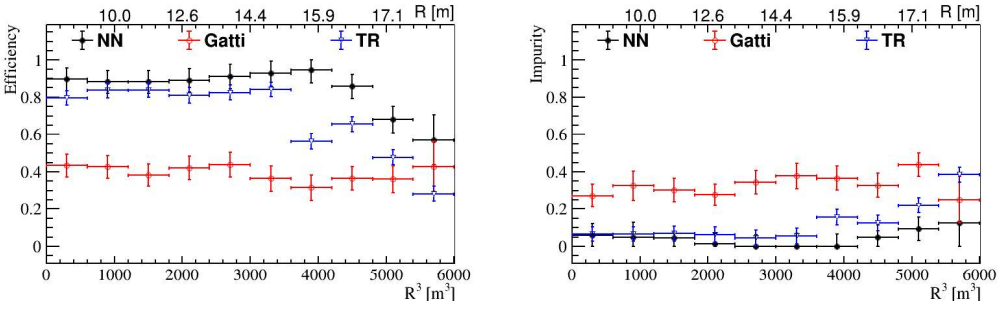
Figure 6.14: Performance of the e^+/e^- discrimination from all three methods. Impurity was obtained at efficiency fixed to 50%, while efficiency was obtained at impurity fixed to 20 %. (a), (b), and (c) show results with Datasets 1, 2, and 3, respectively, as a function of visible energy.

In Fig. 6.14 we show impurity and efficiency as a function of energy for Datasets 1, 2, and 3. Efficiency was determined at a fixed impurity level of 20%. Impurity was

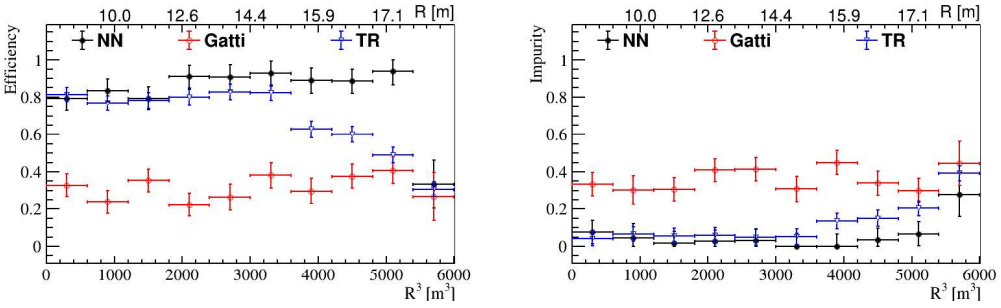
determined with the required efficiency set to 80%. There is an interesting thing that for dataset 1, the performance become worse as energy increase, just as we expect since the difference of time profile would become smaller with the increase of enengy; however, for dataset2 and dataset 3, the best performance is not at the low energy region but at the energy region between 3 and 5 MeV. We have some speculation here: for the low energy samples, the photon statistics is limited compared to high energy samples, and the TTS effect will be much more significant; in this case, the difference of the time profiles for the different particles is totally smeared out. And as the photon statistics increase with energy increase, the effect of TTS become smaller and smaller, so the performance increase first and then decrease. This may be not bad in JUNO experiment, since the events are mainly in the energy range of 3 to 5 MeV.



(a) Dataset 1



(b) Dataset 2



(c) Dataset 3

Figure 6.15: Performance of the e^+/e^- discrimination from all three methods. Impurity was obtained at efficiency fixed to 50%, while efficiency was obtained at impurity fixed to 20 %. (a), (b), and (c) show results with Datasets 1, 2, and 3, respectively, as a function of R^3.

The performance of impurity and efficiency as a function of energy is shown in 6.15. For events in R<16 m, the performance is stable; however, for events in R>16

m, the performance become worse. This is because total reflection start to take effect at around R=16 m, thus p.e. statistics become smaller.

6.7 Summary and Outlook

In this chapter we study the potential of particle identification in JUNO. We can have excellent discrimination ability for alpha/beta and e/p, while for electron/positron the performance is not as good. Neural network shows its advantage in this field.

For electron/positron discrimination, we find that TTS is the main factor that affect the performance of discrimination. Thus, we can consider only use Hammamastu PMTs. In this case, we have only 25% statistics, while TTS can be reduced from 12 ns to 3 ns. Further study need to be performed to see if it helps.

When JUNO experiment starts to run, we can have real data to continue our study. At that time, we should perform out study on real data. To train the neural network, we need to find the natural pure data sources. For alpha and beta sources, we can select Bi-Po cascade decay events; for proton, we can search the spallation neutron sources; for e^+ sources, we can managed to select pure ^{11}C events. Besides, we should check if we can still have good performance when we train on these sources and perform on other datasets with MC data.

Chapter 7

Vertex and Energy Reconstruction

The vertex and energy reconstruction is the base for physics analysis. JUNO experiment requires the resolution of energy reconstruction at 3%@1MeV level, which is a big challenge for liquid scintillator experiments. Besides, the performance of energy reconstruction is dependent on vertex reconstruction, and we require the resolution of vertex reconstruction at 10 cm@1MeV level.

Compared with current running liquid scintillation detectors, the detector volume of JUNO is much larger and the absorption or scattering of optical photons is more significant. In addition, there are several mediums with different optical properties in the detector, which make our work more complex. For example, the refractive index of the liquid scintillator is 1.49, while that of water is 1.33 (more details discussed in Chapter 3). This difference brings refraction and total reflection which will affect the time of flight (TOF) of scintillation photons. As a result, the optical model is a challenge for JUNO. The performances of the vertex reconstruction also relies on the fluctuation of the generation of the scintillation light, as well as the timing precision of the PMTs.

In this chapter, I will firstly show the available information for reconstruction in the experiment. Then, I discuss the factors which lead to bias the reconstruction. Last, I show the structure of the neural network model and give the performance of the reconstruction in my study.

7.1 Input Information: time and charge of hits

To achieve our goal, we need to first see what is available in our hand. In JUNO experiment, the information of photons are collected by the PMTs and then recorded by the FADCs. Thus, all we can know is the collected charge on each PMT and the hit time of each photon. Due to the limitation of current waveform reconstruction algorithm, we can only know the accurate time of the first hit photon on each PMT, and the following photons' hit time information is lost.

The position of the event has a significant impact on the charge and the time response map of PMTs. Here is an example: in Fig. 7.1 and Fig. 7.2, I show the time and charge map of PMTs of several event at different vertexes. We can find obvious clusters in the plots, especially for events closer to the edge. The position and patter of the clusters is determined by the vertex of events.

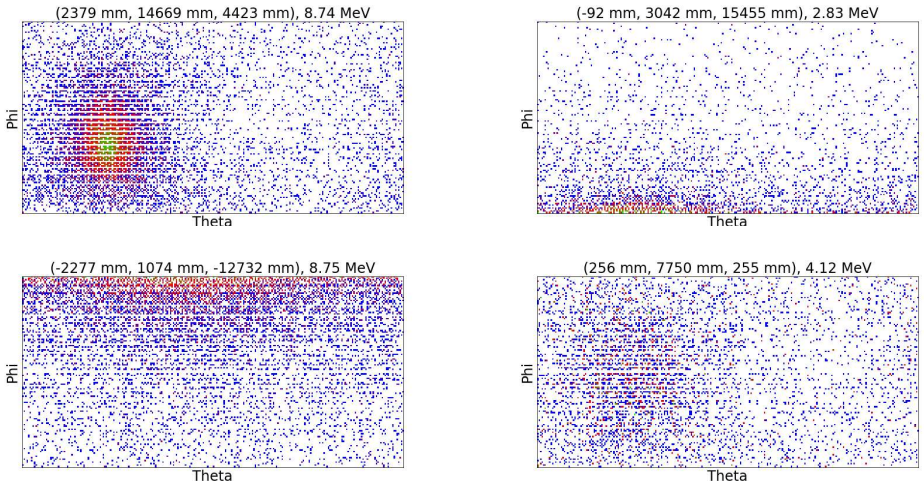


Figure 7.1: Examples of 2D projection of the charge distribution for events with different vertexes and energies.

7.2 Factors affecting vertex reconstruction

Supposing when one event happens, all photons are generated Immediately and there is no decay for light from the vertex of the event and PMT system, we can reconstruct

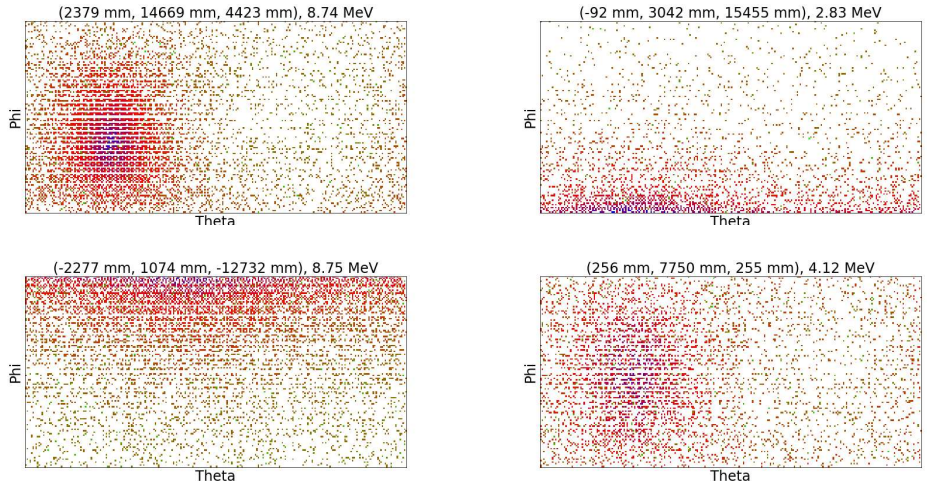


Figure 7.2: Examples of 2D projection of the first hit time distribution for events with different vertexes and energies.

the vertex of the event with almost no bias by calculating the time of flight from the vertex to the PMTs. Besides, the light can experience absorption and re-emission, refraction, total reflection, and scattering in the path to PMTs. Thus, the number of photons and their hit time on each PMT are changed.

7.2.1 Light generation in liquid scintillator

When the charged particles cross the liquid scintillator, they can deposit energy and ionize or excite the molecules of the liquid scintillator. Usually the excited electrons are π electrons. Then the excited molecules can de-excite and emit a certain wavelength of UV or visible light. The light is not emitted immediately, and the number of emitted photons follows function:

$$\phi_{\mathrm{em}}(t) = \sum_{i=1}^{n} \frac{w_i}{\tau_i} \mathrm{e}^{-\frac{t-t_0}{\tau_i}} \qquad \text{with} \qquad \sum_{i=1}^{n} w_i = 1, \tag{7.1}$$

In the above function, n is the number of fluorescence components in liquid scintillator, ω is the ratio of different components, τ is the time constant of different components, and t means the time after the start time of the event.

The excited π electrons can be at singlet state or triplet state, as shown in Fig. 7.3, while the decay time constants τ of the de-excitation processes are not the same. In this case, the light process has fast and slow components.

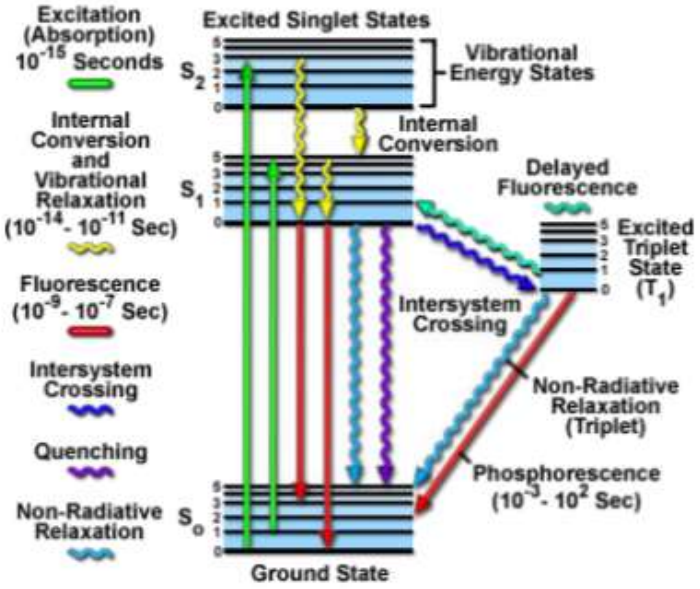


Figure 7.3: The energy structure of π electron. S_0 is ground state, S_1, S_2, S_3 are singlet state, T_1 is triplet state [13].

7.2.2 Light propagation in the detector

Absorption and Re-emission

In JUNO's scintillator, we add PPO as primary fluor and bis-MSB as wavelength shifter. When an optical photon propagates in LS, there exists a competition for light absorption and scattering between all these components. Once a photon is absorbed by a fluorescent molecule, a new photon may be re-emitted with a certain probability. These processes will repeat until the photon is absorbed by a molecule and no new photon is produced, or the photon reaches the PMT. The diagram in Fig. 7.5 schematically describes the model of light propagation in a step-wise way [14] :

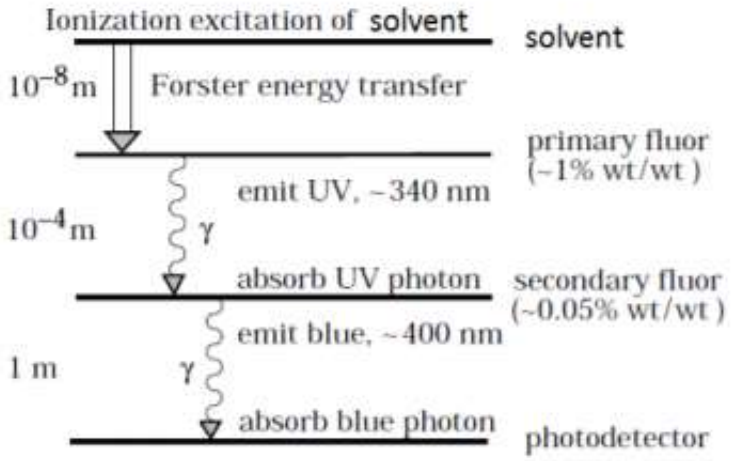


Figure 7.4: Energy transfer between solvent, primary fluor, and secondary fluor/wavelength shifter. [13]

- If a photon is absorbed by a solvent molecule, it vanishes. However, if it is a UV photon, the solvent may still originate a re-emission.
- If a photon is absorbed by a primary fluor or wavelength shifter molecule, it vanishes. A new photon may be emitted with a probability corresponding to the wavelength-dependent fluorescence quantum yield of the primary fluor or wavelength shifter, respectively.
- If a photon is not absorbed by any component of LS, it continues to propagate.
- If photon scattering (Rayleigh) happens, the photon changes its direction according to the scattering cross section and continues propagation.

The absorption and re-emission effect is an important source of the non-uniform and non-linear energy response in a LS detector. Besides, in this process, the photon change its wavelength and velocity while effective velocity changes less than 1% as travel distance changes which can be neglected safely. [15]

For an initial intensity I_0 of an incident light at position x = 0 propagating with a wavelength λ in a liquid material comprising of a mixture of N attenuating species,

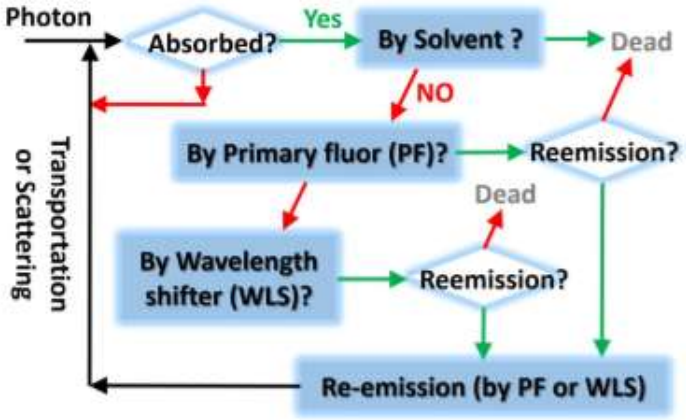


Figure 7.5: The schematic diagram of light propagation in the JUNO optical model [14].

with the intensity reduced to I after propagating an x distance through the liquid, the expression for the intensity can be written as follows

$$I = I_0 e^{-\frac{x}{L_\lambda}} \tag{7.2}$$

where L_λ is intrinsic attenuation length.

Even for a highly-purified LAB sample, there are still impurities at the order of several ppb and ppt, and hence the attenuation length of the LAB samples will still be influenced by the impurities.

In the studies in [80], we can achieve a light attenuation length of 25.8 m, comparable to the diameter of the JUNO detector [80]

Rayleigh scattering

The photons can elastic scatter on the molecules in liquid scintillator, which is called "Rayleigh scattering". In this process, the energies of the photons are not changed, while their directions are changed, which will spread the time of flight of the photons. The intensity of scattering light follows the function

$$I = I_0 \frac{1+\cos^2\theta}{2R^2}(\frac{2\pi}{\lambda})^4(\frac{n^2-1}{n^2+2})^2(\frac{d}{2})^6 \tag{7.3}$$

where I_0 is the intensity of incident light, θ is the scattering angle, R is the distance to the molecule, λ is the wavelength of light, n is the refractive index of the material, and d is diameter of the molecules. Averaging this over all angles gives the Rayleigh scattering cross-section [81]:

$$\sigma_s = \frac{2\pi^5}{3}\frac{d^6}{\lambda^4}(\frac{n^2-1}{n^2+2})^2 \tag{7.4}$$

The Rayleigh scattering length is different for different kinds of liquid scintillators. For LAB which will be used in JUNO experiment, the Reyleigh scattering length is 45 m [82].

Edge effect

When the light travels to the edge of the two kinds of medium, it can refract or total reflect at the interface due to the different optical properties of the mediums. For light from the vertex to PMTs, it travels through liquid scintillator, acrylic ball, and water. The optical properties of liquid scintillator and acrylic are similar, thus we don't need to consider the effects when light pass through the interface of them.

However, when light enters water, refraction will happen due to the different refractive index of liquid scintillator and water (1.49 for liquid scintillator and 1.33 for water).

Besides, according to the total reflection formula on the interface between liquid scintillator and water:

$$r_c = R_{LS} \times \frac{n_{water}}{n_{LS}} \tag{7.5}$$

where r_C is the radius where the total reflection will start to happen, R_{LS} is the radius of the central detector, n_{water} is the refractive index of water, n_{LS} is the refractive index of liquid scintillator.

Thus, we can know that the total reflection will start to happen at about 16 m from the center of the detector, and part of the photons emitted in this region will

not reach PMTs. Besides, Rayleigh scattering will make the situation more complex.

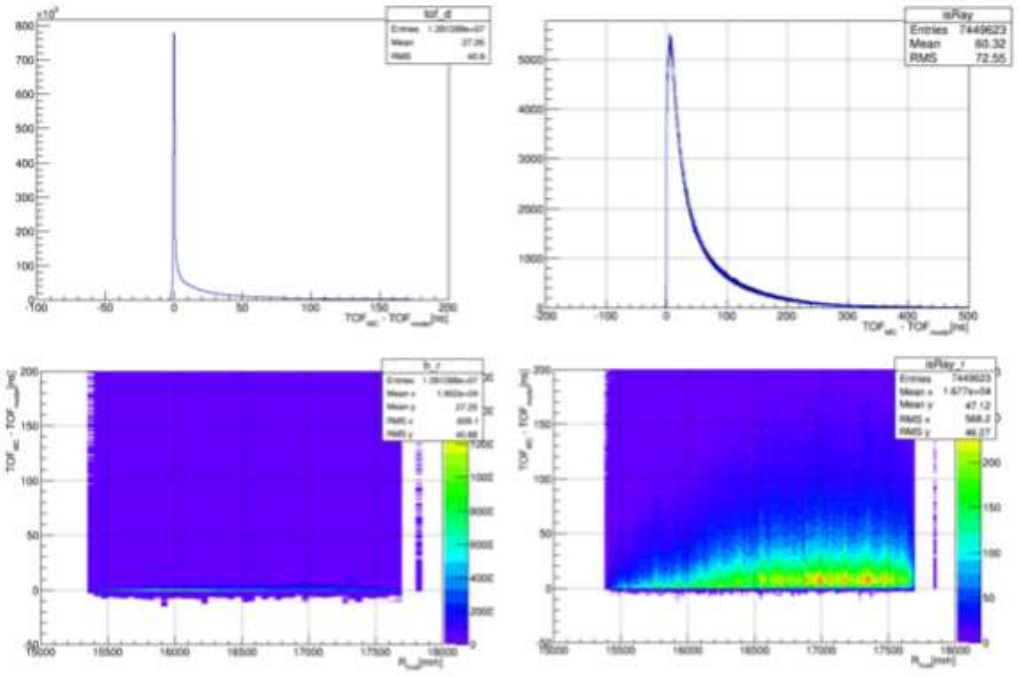


Figure 7.6: TOF in dark zone comparison between MC and model. Most estimations are consistent with simulation. For Rayleigh scattering photons(second column), the bias becomes large and is broaden widely. The closer to the edge, the worsen the estimation is. [15]

7.3 Vertex Reconstruction

7.3.1 Charge Center Method

The closer the PMTs to the source, the more photons will be collected. By considering the collected charge as a weight for each PMT, the source position can be estimated through formula:

$$\bar{r_0} = \frac{\sum_i q_i \bar{r}_i}{\sum_i q_i}, \tag{7.6}$$

where r_i is the position of the ith PMT and q_i is the number of photoelectrons detected by the ith PMT. Only the PMTs that actually have been hit are included in the sum. The charge q_i can be given by adding up all detected hits of the ith PMT, which is described by the following equation,

$$q_i = \sum_{j}^{k} h_j, \tag{7.7}$$

where h_j is the j-th hit on the i-th PMT. The sum j to k counts over all hits for the i-th PMT. For real event data, the collected charge of a PMT is related to the photoelectron hits and can be known by us. This can be used instead, as long as the PMT does not come close to saturation.

Suppose the true vertex is z_0, one can mathematically calculate the vertex z from the charge center method:

$$\begin{aligned} < z > &= \frac{1}{4\pi} \int z d\Omega \\ &= \frac{1}{4\pi} \int_0^{2\pi} d\phi \int_0^{\pi} (z_0 + r \cdot \cos\theta) \sin\theta d\theta \\ &= \frac{1}{2} \int_0^{\pi} -(z_0 + (\sqrt{R^2 - z_0^2 \sin^2\theta} - z_0 \cos\theta)) \cos\theta d\cos\theta \\ &= \frac{1}{2} \int_{-1}^{1} (z_0 + x\sqrt{R^2 - z_0^2 x^2} - z_0 x^2) dx \\ &= \frac{2}{3} z_0 \end{aligned} \tag{7.8}$$

where r is the radius of the event, θ, ϕ, x is the coordinate of the event, and R is the radius of the central detector.

Thus, to get the true vertex, the charge center result need to be multiplied by a factor 1.5. Considering the real optical processes during the photon propagation, the factor is set to 1.2 in JUNO. This result is derived from mathematics purely, regardless of the geometry of detector and any physical processes in experiment, which means its reliability in any case. However, the result of charge center method is quite rough, thus it can only be an auxiliary tool.

7.3.2 Reconstruction with Time information

Usually people reconstruct vertex of the event by constructing likelihood function and minimize it. Since our goal is the reconstruction of low energy events, and these events can be treated as point-like events, then we can define the residual hit time variable $t_{i,res}$ as:

$$t_{i,res}(\vec{r_0}, t_0) = t_i - tof_i - t_0, \tag{7.9}$$

where t_i is first hit time of i^{th} PMT, t_0 is event start time, and tof_i is the time of flight for the scintillation photons.

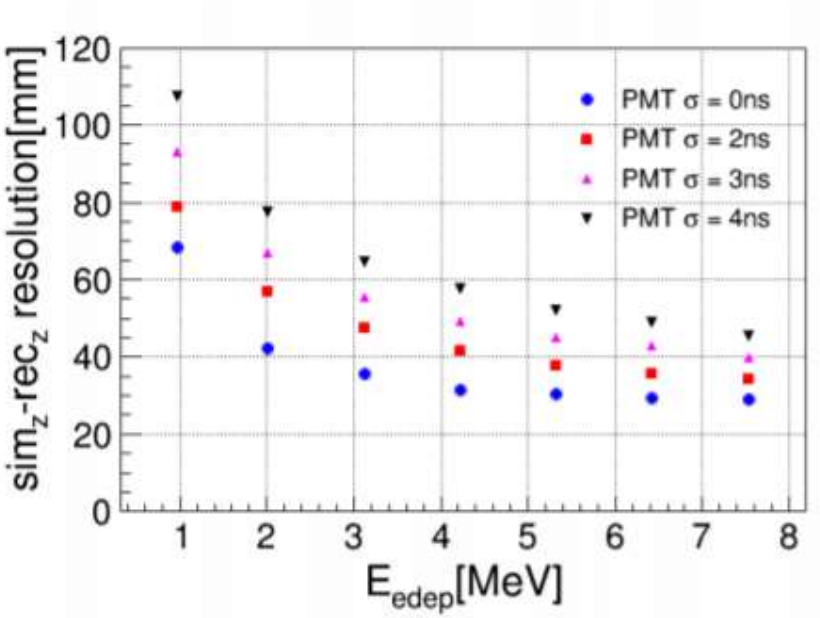


Figure 7.7: The performance of vertex resolution and the effect of TTS . If no TTS is considered, the vertex resolution is about 7cm@1MeV. If TTS is about 10 ns, vertex resolution increase to 11cm@1MeV [15].

The probability density function $f(t_{res})$ reflects the dispersion time of the first photon to reach a PMT out of n p.e.s, and could be calculated as:

$$f_n(t_{res}) = nf(t_{res})(\int_{t_{res}}^{\infty} f(x)dx)^{n-1}. \tag{7.10}$$

Then we can define a likelihood function as follows:

$$\mathcal{L}(\bar{r_0}, t_0) = -\sum_i (f_n(t_{res}). \tag{7.11}$$

By minimizing the likelihood function, we can reconstructed the best vertex position, with the performance shown in Fig. 7.7 [15].

Due to the powerful fitting ability of neural network, I construct a neural network to utilize time and charge information to reconstruct vertex. The model is based on the famous Resnet-50 model, with the structure shown in Fig. 7.8. I use ResNet module as the basic module, and overlap them one by one. In total the model has 50 layers and about 2 million parameters.

For the model, I project 3D time and charge map on 2D map, as shown in Fig. 7.1 and Fig. 7.2 and use them as the input of neural network. Then, the model calculates with the input information and gives us the output, which is the reconstructed vertex.

The setting of the neural work is as follows:

- For loss function: I use mean square error function, which means the square of distance between reconstructed vertex and true vertex.
- For the optimizer, I use Adam, with the default setting
- For the activation function, I use no activation function in the last layer, while I use ReLU function in all other layers
- For batch size, I use 64 events in one batch
- I train 80 epochs in total
- For learning rate, I set the initial learning rate as 0.001, and divide the number by 2 every 10 epoch.

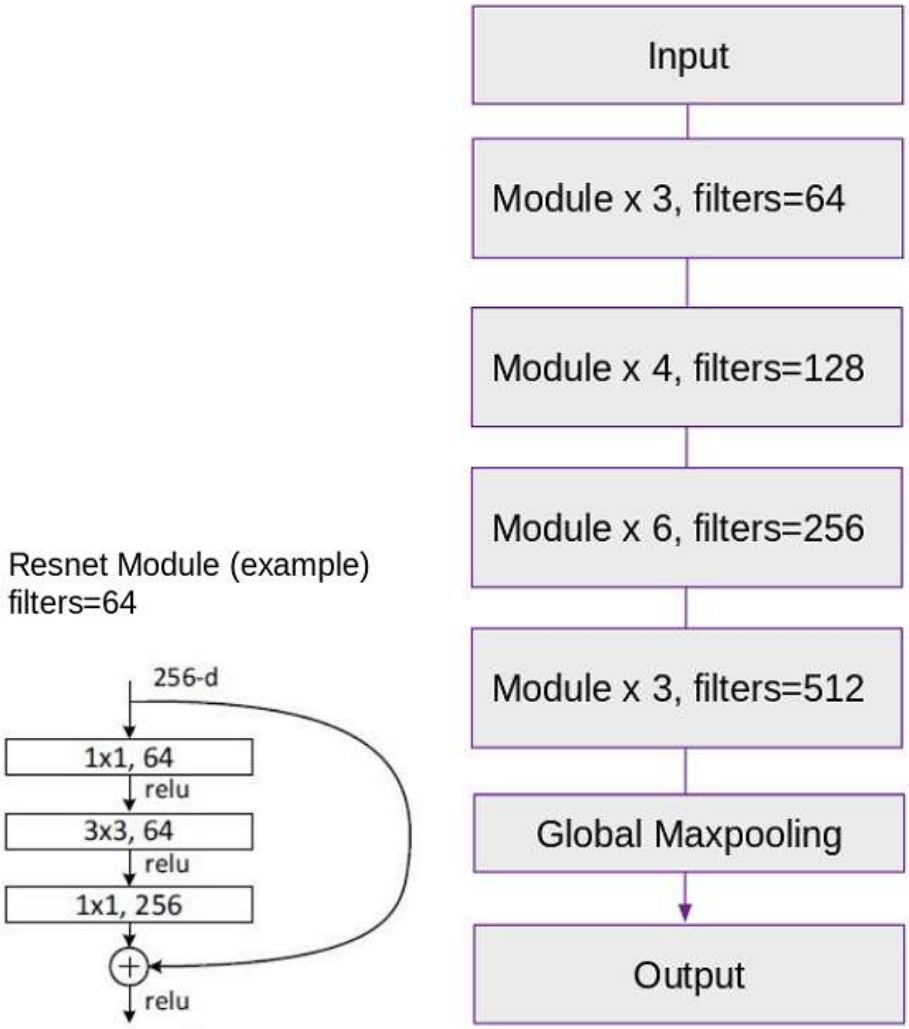


Figure 7.8: The structure of the neural network. Left: the structure of one Resnet module; Right: the strurcture of whole network based on the module

To train the neural network, I generate 1 million positron events as the sample. The visible energy range of the events is from 1 to 10 MeV, and the events are uniformly distributed in the detector.

I split the dataset into two parts: training dataset and validation dataset. I put 90% of the samples in the training dataset and 10% in the validation dataset.

During training, I monitor the loss value of the validation dataset and record the

minim value of the whole process. In the end of every epoch, I record the model and compare the loss value of validation dataset in this epoch and the recorded minim value. If the loss value is smaller than the recorded minim value, then I save the parameters of the model and update the minim value, else I do nothing. Thus, when I finish the training, the best group of parameters during training is saved.

To evaluate the performance of the model, I generate another dataset. Mono-energetic samples for positron events with 1, 2, 3, 4, 5, 6, 7, and 8 MeV. The performance of the model on this dataset is shown in Fig. 7.9. From the plot we can see that the vertex resolution is about 6.3 cm at 1 MeV.

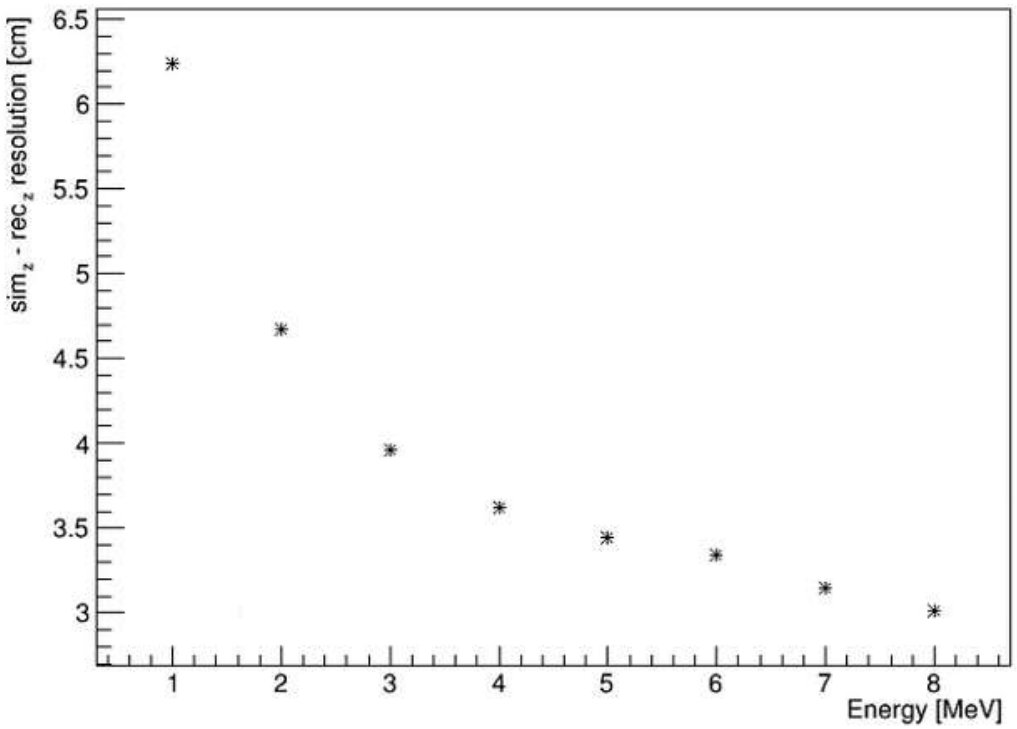


Figure 7.9: The performance of the vertex resolution by neural network, the resolution about 6.3 cm at 1 MeV

Besides, we can also study the stability of the performance with energy and radius, as shown in Fig. 7.10. In the top plot of Fig. 7.10, we can see the bias of reconstruction with energy is less then 0.2 cm in the whole energy range of reactor neutrino spectrum.

In the bottom plot, we see the bias of reconstruction with radius. The absolute value of bias is also less than 0.2 cm, and the performance remains stable in the whole detector volume, and the edge effect is well corrected.

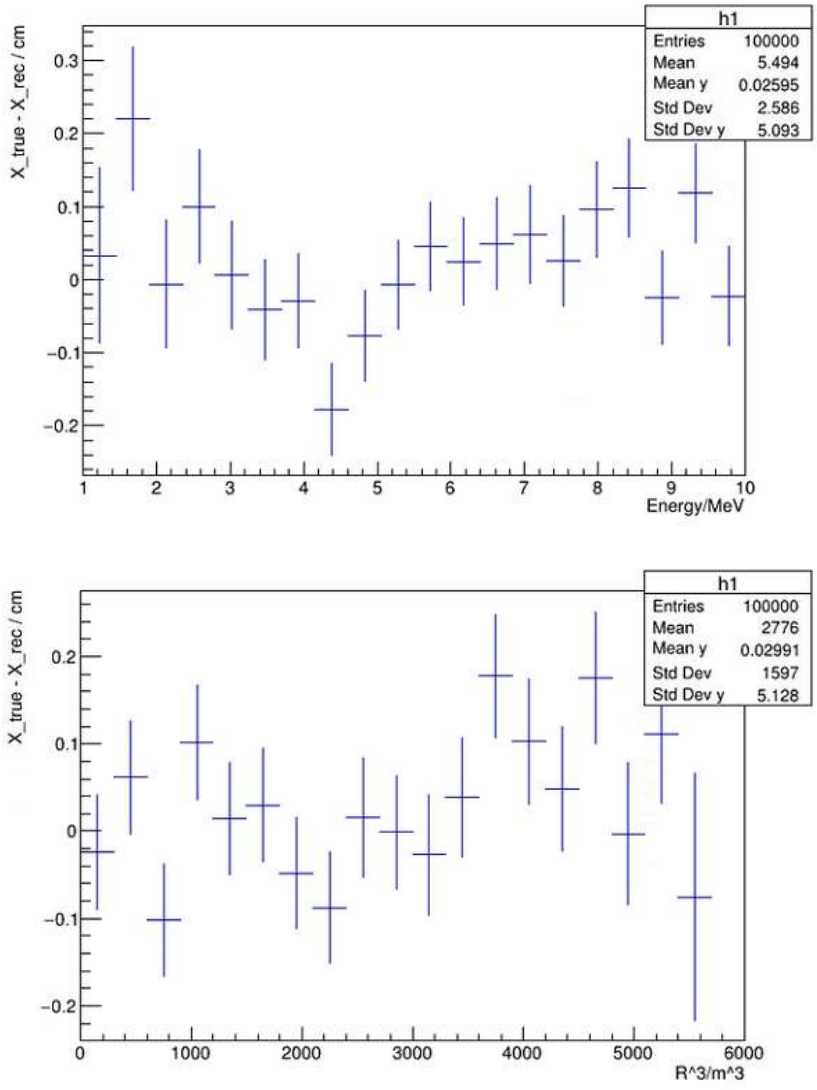


Figure 7.10: Top: the bias of vertex reconstruction with energy; Bottom: the bias of vertex reconstruction with radius.

7.4 Energy Reconstruction

7.4.1 Factors affecting energy reconstruction

energy resolution function

$$\frac{\sigma_E}{E} = \sqrt{(\frac{a}{\sqrt{E}})^2 + b^2 + (\frac{c}{E})^2} \tag{7.12}$$

with the visible energy E in the unit of MeV.

- a: statistical, true nPEs
- b: constant, non-uniformity, 0.81%
- c: noise, dark noise, 1.21%

In the liquid scintillator, the charged particles can deposit energy and emit light. The light output is approximately linear to the particle's energy, thus we can evaluate the energy of incident particle by counting the collected photons on the PMTs.

In Fig. 7.11, we see that there are about 1300 photons collected by PMTs per MeV. However, since we require the accuracy of energy reconstruction of 3% / $\sqrt{MeV}$, we need to consider the second order effect. For the same energy, there is a wide spread of collected photons, which is mainly related to the position of the particles, and called "non-uniformity" effect. For different energies, the mean value of collected photons is increasing with energy, which is called "non-linearity" effect.

Thus, for energy reconstruction, the question is: giving the total charge and the vertex of the event, what is its true energy?

Non-uniformity effect

The photons in the liquid scintillator can be absorbed, scattered, and refracted. In the different positions of the detector, the mean distance from the event to the PMTs is different, thus the ratio of lost photons is also different, which leads to the non-uniformity effect of energy reconstruction.

To show the non-uniformity effect, I simulate 100,000 zero energy positron events uniformly in the whole detector, and the relation between true energy and charge is

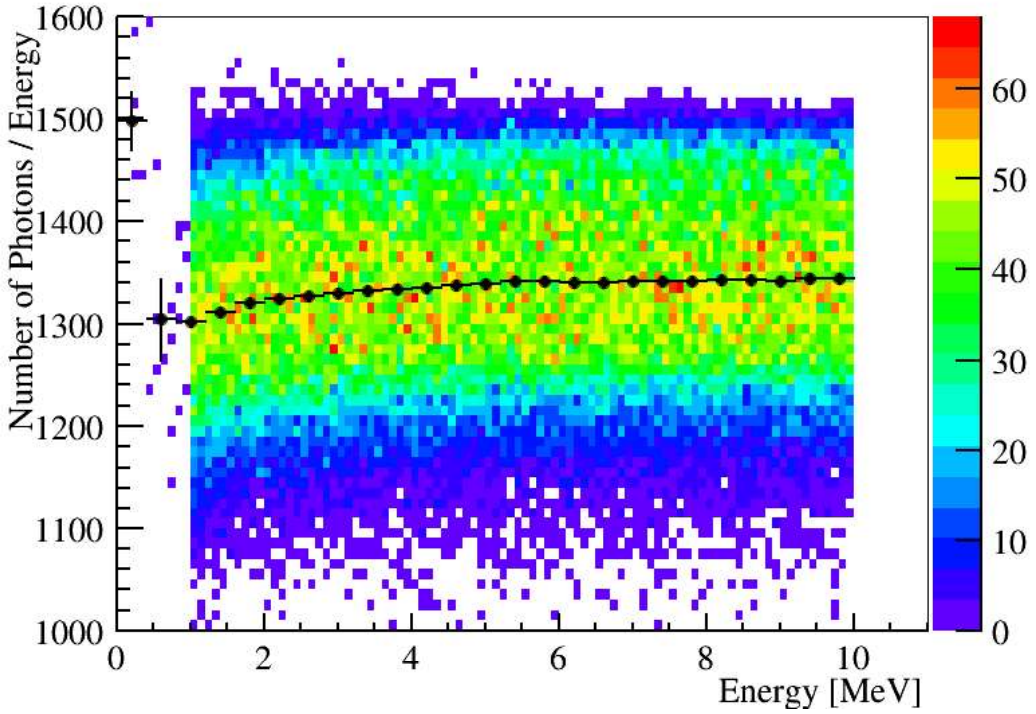


Figure 7.11: The correlation of charge and energy for positron events uniformly distributed in the detector. The energy of positron events is from 1 to 10 MeV.

shown in Fig. 7.12. In the plot, the spread of charge at same radius is quite narrow, while the mean value of charge is changed with radius. We see the mean value increase first, since when the event becomes closer to the boundary, the sum of path to all PMTs becomes shorter and more photons can reach PMT. However, the value starts to decrease at the radius at around 16 m, where total refraction start to take effect.

In the experiments, we usually calibrate the detector with mono-energetic sources placed at the different positions inside the detector and get the response map to fix the effect. Thus, the resolution of energy reconstruction is highly dependent on the resolution of vertex reconstruction.

Non-linearity effect

In the liquid scintillator, the light output from the particles is quenched. We have an empirical model to describe the quenching mechanism, which was first discussed

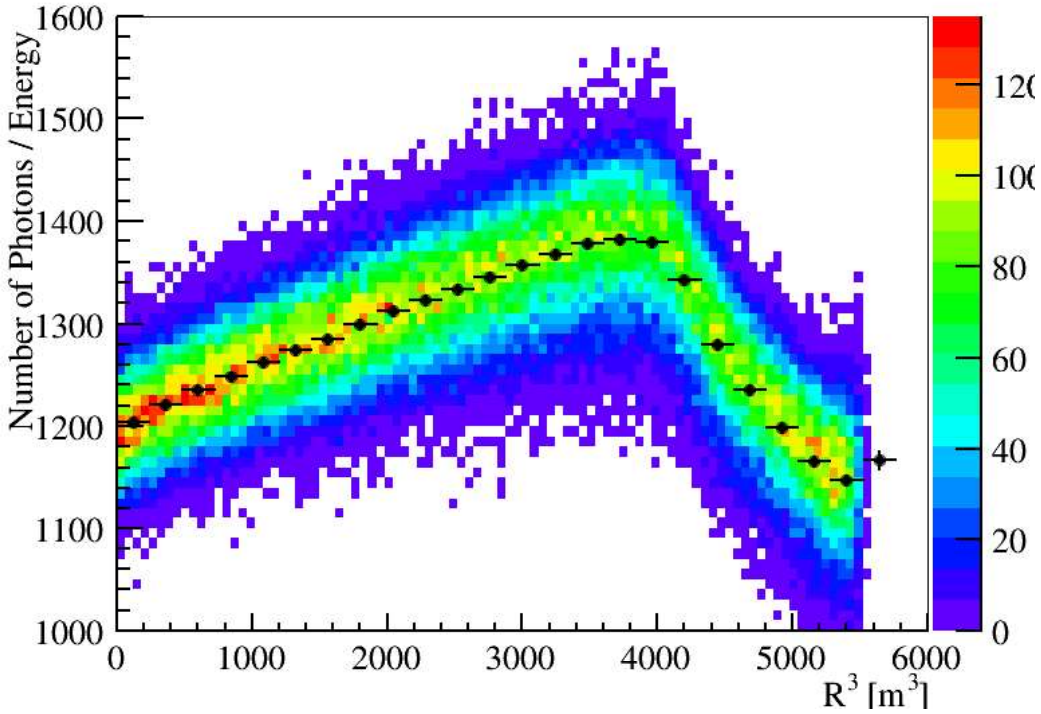


Figure 7.12: The correlation of charge and energy. positron events with 0 kinetic energy, uniformly distributed in the detector

by Birks and is named "Birks law". The light output of the liquid scintillator is related to the energy deposition density dE/dr :

$$\frac{dL}{dr} = S\frac{\frac{dE}{dr}}{1 + k_B\frac{dE}{dr} + k_C(\frac{dE}{dr})^2} \tag{7.13}$$

where dL/dr is the scintillation light yield per unit path length r, S is the scintillation light yield per MeV, k_B is the Birks' constant, and k_C is the second order parameter. From the function, we know that the quenching effect is increasing with larger energy deposition density dE/dr. Usually, the particles' dE/dx increase with the decrease of energy, meaning that the particles' visible energy is not linear proportional to their true energy.

Additionally, the particles in liquid scintillator can emit Cheronkov light, while

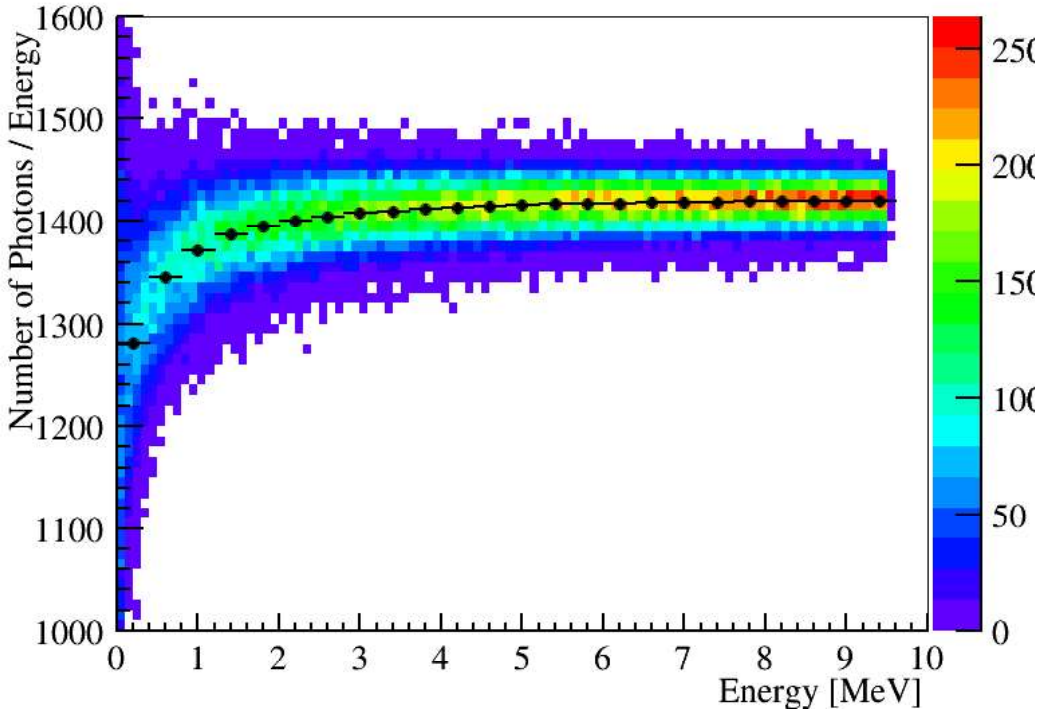


Figure 7.13: The correlation of measured charge and energy. The momentum of the electron events are from 0 to 10 MeV, and all electron events are distributed in the center of the detector

the ratio of number of photons and the energy of particles is changing with energy, which also introduce non-linearity.

The performance of non-linearity effect is shown in Fig. 7.13. Here I simulated 100,000 electron events all in the detector center, while their momentum range from 0 to 10 MeV. We can see clearly the non-linearity curve in the whole energy range.

7.4.2 Reconstruction with simple information

Generally, when we reconstruct energy, we start at the total charge of the event, then we correct energy non-uniformity and energy non-linearity step by step. Thus, I construct a neural network model, with the structure shown in Fig. 7.14.

The input information of the model is the events' position and charge. If the

detector is a perfect sphere, we only need the radius of the event since charge is symmetric at the axis of θ and ϕ. However, with the existence of chimney, the detector is not symmetric at θ axis, thus we also need θ information of the event.

The model works as follows: first, we input the r and θ information, which is followed by two layers. Then, we multiply the output with the charge, and add another layer to deal with it. Up to now, what I try to do is to correct the energy non-uniformity effect. Then, we add another four layers and give the output. In this part of the model, I try to correct the energy non-linearity effect.

For the hyper parameter of the network, I use the following setting:

- For loss function: I use mean square error function, which means the square of distance between reconstructed energy and true energy.
- For the optimizer, I use Adam, with the default setting.
- For the activation function, I use no activation function in the last layer, while I use ReLU function in all other layers.
- For batch size, I use 64 events in one batch.
- For learning rate, I set the initial learning rate as 0.001, and divide the number by 2 every 5 epoch.

For the training of the model, I use a special strategy: I train the two parts of the model separately. Firstly, I fix the parameters of the second part of the model, and train the network to optimize the parameters of the first part of the model. After the loss value becomes stable and stop decreasing, I break the training, fix the parameters of the first part, activate the parameters of the second part, and then start training. Again, when the loss value becomes stable, I stop training. Last, I activate all of the parameters and start training, then give the final result.

I train the model like this because the correlation between input and output is not so strong. When energy non-linearity effect couples together with energy non-uniformity effect, the situation will become complex, and when we train the model it is hard for the neural network to limit the parameters in a reasonable range. Thus,

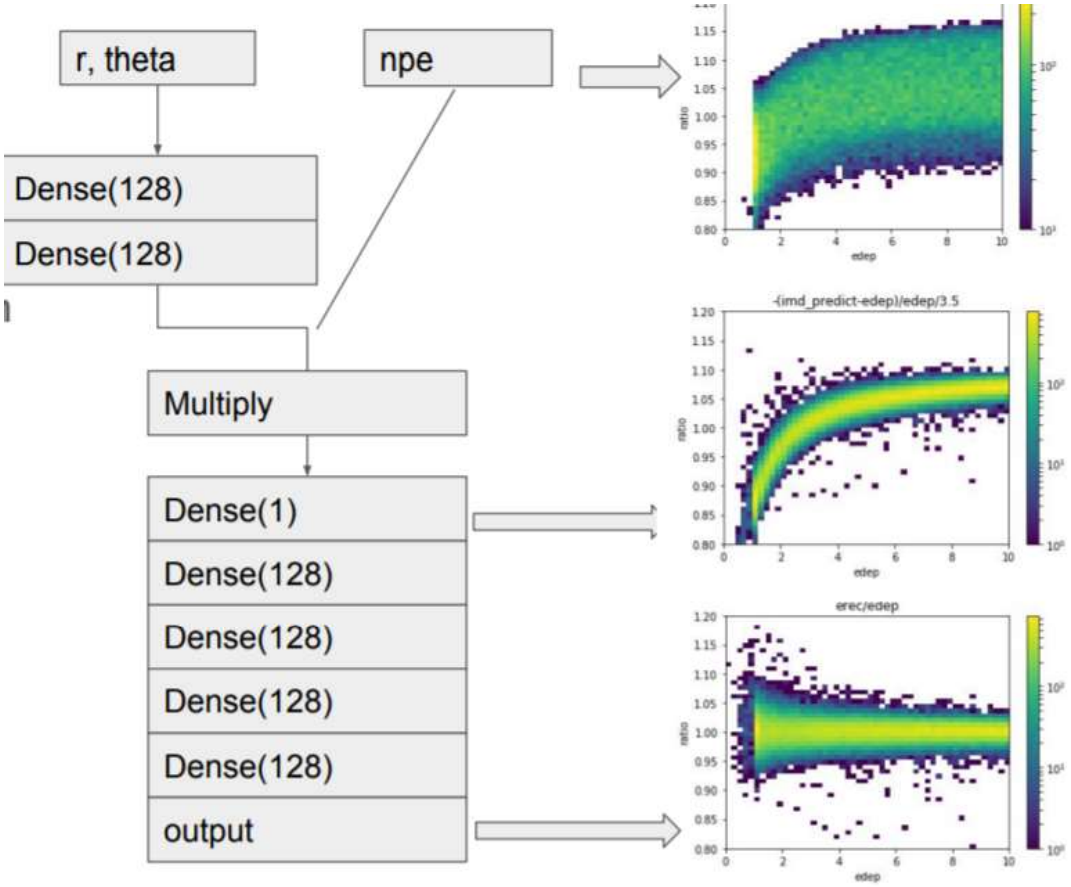


Figure 7.14: The principle of energy reconstruction

I train the model separately, and the neural network can find some not so good but reasonable values, and then I can optimize the parameters based on a good point.

Actually we can see how neural network works after finishing training. In the right column of Fig. 7.14, I show three plots: the top one is the correlation map of energy and charge, the middle one is the output after the network finish the first part of calculation, while the bottom plot shows the result of output. We see clearly that the wide spread curve becomes narrow in the middle plot, and become a line in the bottom one. So the model really works!

Last, let's evaluate the performance of the model. In Fig. 7.15 is the resolution of the reconstructed energy. The resolution is 3.03% at 1MeV, which meet the

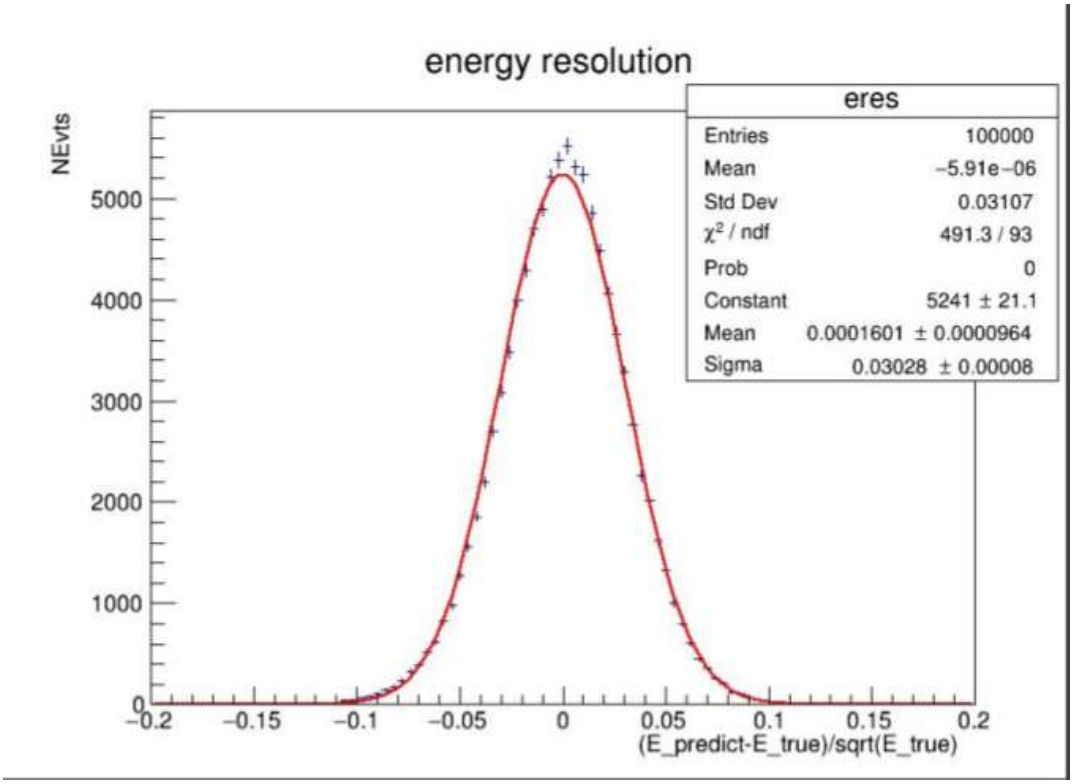


Figure 7.15: The principle of energy reconstruction. We correct energy non-uniformity effect in the first step and correct energy non-linearity effect in the second step.

requirement of JUNO experiment.

Fig. 7.16 shows the bias of the reconstruction with radius. For events in the radius of 17.2 m, the bias is less than 0.1% and after 17.2 m, the bias increases.

Fig. 7.17 shows the bias of the reconstruction with energy. In the whole energy range of reactor neutrino events, the bias remains less than 0.1%, which means the energy non-linearity is corrected quite well.

7.5 Outlook

Currently the reconstruction of vertex is only based on the hit time of the first hit photon on each PMT, due to the limitation of the waveform reconstruction algorithm. With new waveform reconstruction algorithm, we are able to discriminate the

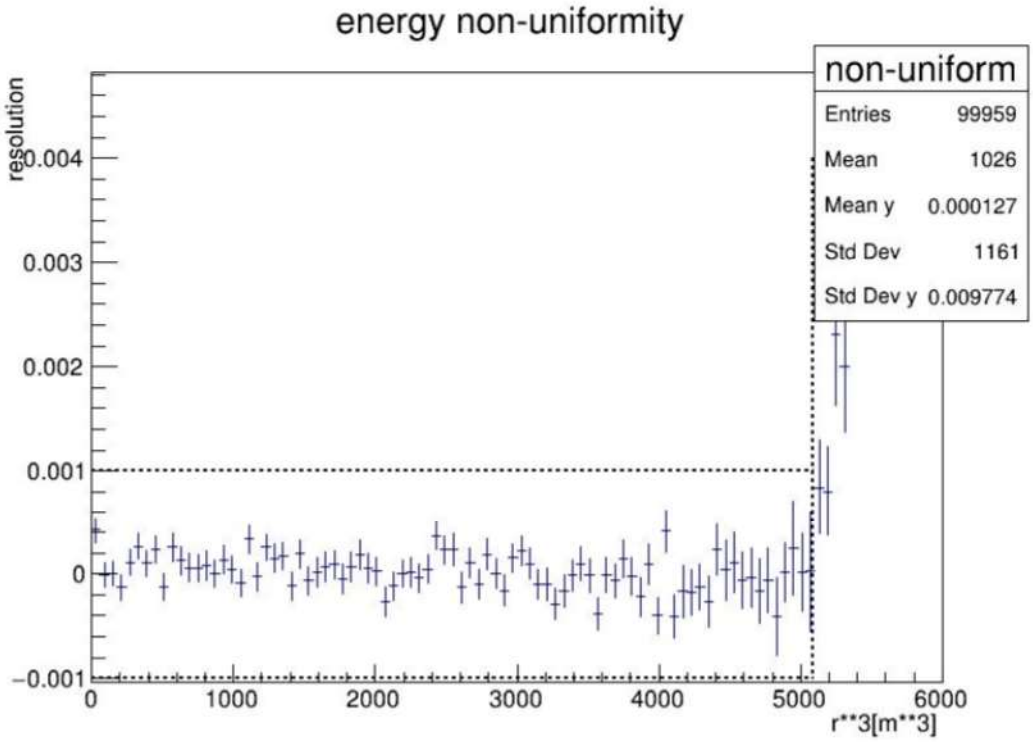


Figure 7.16: The uniformity of energy reconstruction. Bias less than 0.1% in fiducial volume.

overlapped peaks of FADC and reconstruct the hit time of each single photon, thus additionally improve result the vertex reconstruction.

For machine learning models, there are usually tens of millions parameters. In current studies, computer scientists find than we can find a way to remove 90% to 95% of parameters with the performance not decrease. Although the reconstruction speed is hundreds times faster than traditional method, we can still accelerate the calculation and save the resources.

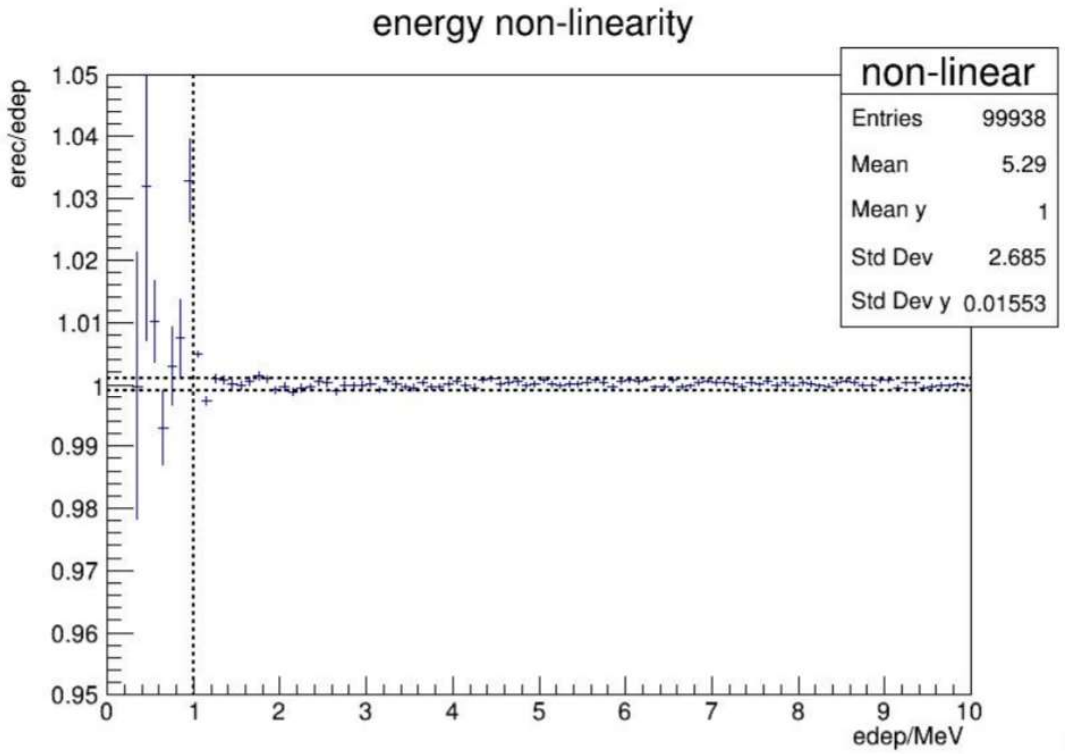


Figure 7.17: Correction of energy non-linearity effect. Bias less than 0.1%.

Chapter 8

The fit of Mass Ordering

Traditionally we fit the mass ordering based on the principle of frequency statistics, with the construction of chi2 function. In this chapter I try to do the fit based on the Bayes theory.

In this chapter, first I will show the principle to determine the neutrino mass ordering, then is how we get the expected neutrino energy spectrum with different assumption of neutrino mass ordering. Last I introduce the principle of Bayes theorem, and show how we determine the neutrino mass ordering based on this.

8.1 Principle

For reactor neutrino oscillation, the survival probability of $\bar{\nu}_e$ can be broken down into three terms:

$$P_{ee} = 1 - P_{21} - P_{31} - P_{32} \tag{8.1}$$

$$\begin{aligned} P_{21} &= \cos^4(\theta_{13}) \sin^2(2\theta_{12}) \sin^2(\Delta_{21}) \\ P_{31} &= \cos^2(\theta_{12}) \sin^2(2\theta_{13}) \sin^2(\Delta_{31}) \\ P_{32} &= \sin^2(\theta_{12}) \sin^2(2\theta_{13}) \sin^2(\Delta_{32}) \end{aligned} \tag{8.2}$$

where Δ_{ij} is

$$\Delta_{ij} = \frac{\Delta m^2_{ij} L}{4E} \simeq 1.27 \Delta m^2_{ij}(eV^2) \frac{L(km)}{E(GeV)} \tag{8.3}$$

From the formula, what can we know? First, for the different mass orderings, we have Δm^2_{31} with different value and sign, since $\Delta m^2_{31} = \Delta m^2_{32} + \Delta m^2_{21}$, and $\Delta m^2_{21} \ll |\Delta m^2_{32}|$. In this case, P_{31} and P_{32} is dependent on mass orderings, which lead to different neutrino energy spectrum for different mass ordering. However, due to the smallness of Δm^2_{21}, the difference of the two spectra is only about 3%, which requires 3% energy resolution to discriminate the difference. In addition, it is not possible to get the absolute value of Δm^2_{32}, since $\sin^2$ function is a periodic function, and we can only know the sign of Δm^2_{32} while the absolute value is measured by other experiments. Last, terms P_{31} and P_{32} are suppressed by $\sin^2 2\theta_{13}$, which requires larger value of $\sin^2 2\theta_{13}$ for better discrimination ability ;

8.2 Reactor neutrino energy spectrum

The reactor neutrinos are emitted via β-decays of unstable fission fragments, with the spectrum as follows:

$$S(E) = \Sigma_i F_i S_i(E) \tag{8.4}$$

In this study, the fraction F_i for isotopes ^{235}U, ^{239}Pu, ^{238}U, ^{241}Pu is 0.58, 0.30, 0.07, 0.05 respectively [1].

For the spectrum of the isotopes $S_i(E)$ we use the Huber [83] and Mueller [84] flux models, which can be seen in Fig. 8.1.

IBD cross section is taken from [85], which is also shown in Fig. 8.1. At the energy range of several MeV, the IBD cross function can also be simplified to [86],

$$\sigma(\bar{\nu_e}p) = 9.45 p_e E_e \times 10^{-44} \mathrm{cm}^2, \tag{8.5}$$

where p_e is the momentum of positron and E_e is the kinetic energy of positron.

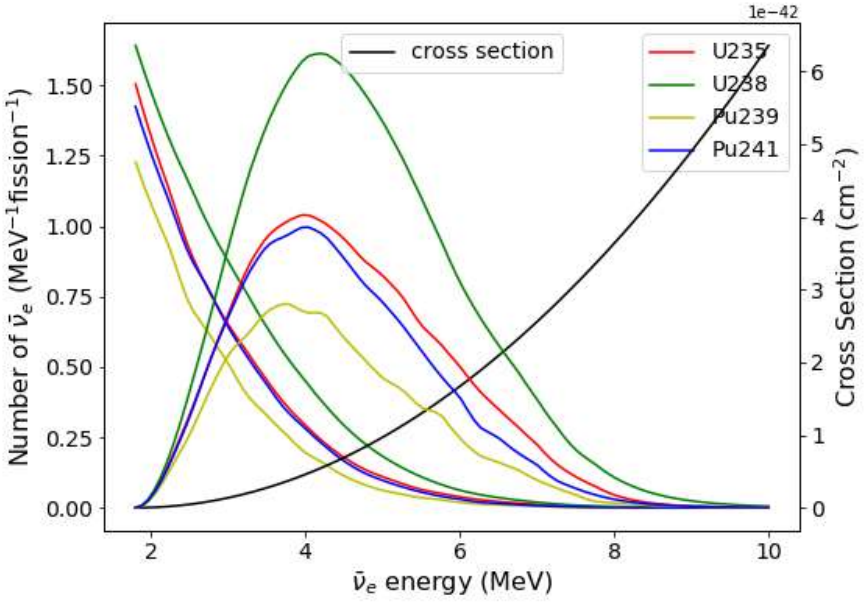


Figure 8.1: Exp colored curves represent Reactor neutrino spectrum of 4 isotopes (^{235}U, ^{239}Pu, ^{238}U, ^{241}Pu), and black line represents the cross section of inverse beta decay. Their products are colored "gaussian like" curves, represent the spectrum observed by the detector

With the reactor neutrino flux and IBD cross section, we can calculate the neutrino energy spectrum under non-oscillation case in the detector. In JUNO experiment, the hydrogen content in the liquid scintillator is 12.01%, thus for the detector at the

distance of 1 km from the reactor, there should be 0.943 events/day/ton/GW, with the spectrum shown in Fig. 8.2.

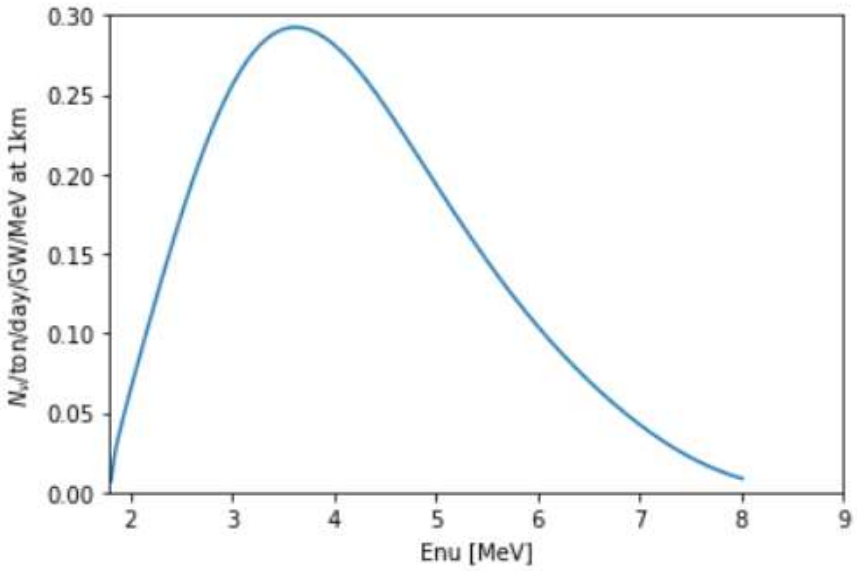


Figure 8.2: The non-oscillated neutrino energy spectrum for detector at the distance of 1 km from the reactor. Event rate is 0.943 events/day/ton/GW

Assuming the detector with 20 kton liquid scintillator located at equal baselines of 52.5 km away from two reactor complexes (36 GW in total), we get the energy spectrum in Fig. 8.3. We select the events range from 1.8 to 8 MeV in 200 bins. We can see clear difference between the normal order case and inverted order case.

In the detector, what we can observe is the positron signals from IBD reaction, with energy:

$$E_{e^+} = E_\nu + m_p - m_n + 2m_{e^+} = E_\nu - 0.782 \tag{8.6}$$

8.3 Bayes theory

According to Bayes theory, given event A and B, we know that

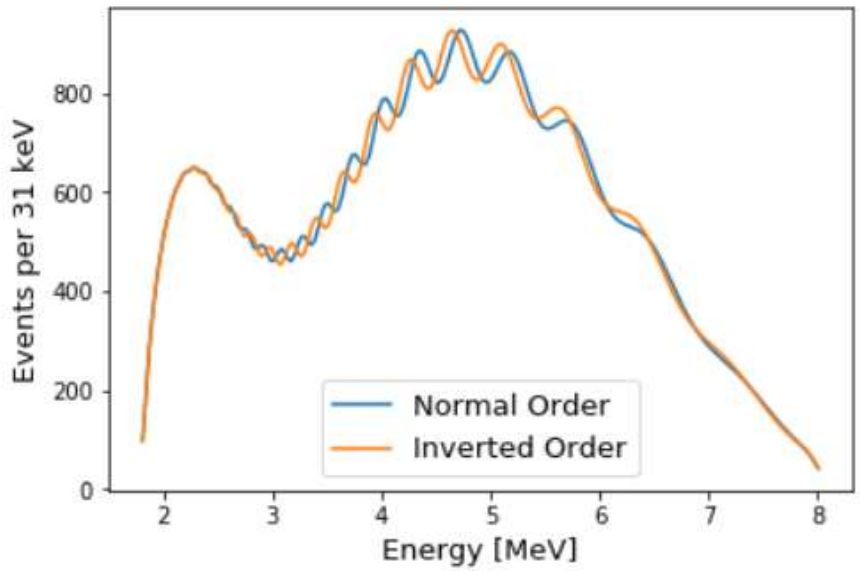


Figure 8.3: The oscillated neutrino energy spectrum assuming 20 kton liquid scintillator located at equal baselines of 52.5 km away from two reactor complexes (36 GW in total). Both NMO and IMO case are considered. Energy range from 1.8 MeV to 8 MeV, 200 bins in total.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{8.7}$$

Here, we can take A as the mass ordering (NH or IH), B as the observed spectrum (x), thus we have

$$\begin{aligned} P(NH|x) &= \frac{P(x|NH)P(NH)}{P(x)} \\ &= \frac{P(x|NH)P(NH)}{P(x|NH)P(NH) + P(x|IH)P(IH)} \end{aligned} \tag{8.8}$$

Here, P(NH) and P(IH) = 1 - P(NH) should reflect one's knowledge of NH and IH prior to the experiment. In the MH problem, it is reasonable to assume that NH and IH are equally likely, that is P(NH) = P(IH) = 50%. We will make this assumption throughout the paper. Consequently, Eq. reduces to

$$P(NH|x) = \frac{P(x|NH)}{P(x|NH) + P(x|IH)} \tag{8.9}$$

For the probability of IH, given the observed spectrum x:

$$P(IH|x) = \frac{P(x|IH)}{P(x|NH) + P(x|IH)} \tag{8.10}$$

For N bins, we have observed events in each bin $\mu_1, \mu_2 ... \mu_N$. For the NH assumption, we have predicted events in each bin $\mu_1^{NH}, \mu_2^{NH} ... \mu_N^{NH}$, for the IH assumption, we have predicted events in each bin $\mu_1^{IH}, \mu_2^{IH} ... \mu_N^{IH}$, thus we have

$$P(x|NH) = \prod_i^N P(\mu_i|\mu_i^{NH}) \tag{8.11}$$

$$P(x|IH) = \prod_i^N P(\mu_i|\mu_i^{IH}) \tag{8.12}$$

in order to simplify the calculation, we use the log values:

$$log(P(x|NH)) = \sum_i^N log(P(\mu_i|\mu_i^{NH})) \tag{8.13}$$

$$log(P(x|IH)) = \sum_i^N log(P(\mu_i|\mu_i^{IH})) \tag{8.14}$$

We assume the number of events in each bin to follow gauss distribution. For Gaussian random variable with mean m and standard deviation σ, evaluated at t, we denote its pdf as

$$G(t; m, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-m)^2}{2\sigma^2}} \tag{8.15}$$

This approximation is often presented in terms of an approximation to a one-to-one transformation of P(x—NH), namely

$$\Delta\chi^2(x) = -2log(P(IH|x)/P(NH|x) \tag{8.16}$$

With the above equations Eq. 8.9, 8.10, and8.16, we get [87]

$$\Delta\chi^2(x) = -2log(P(x|IH)/P(x|NH) \tag{8.17}$$

8.4 The fit of Mass Ordering (MO)

8.4.1 Ideal case

Assuming 20 kton liquid scintillator located at equal baselines of 52.5 km away from two reactor complexes (36 GW in total). DAQ time is 2000 live days and a detector energy resolution is set to 3% at 1 MeV. The relevant oscillation parameters are taken from the global analysis [28] as $\Delta m^2_{21} = 7.54 \times 10^{-5}$ eV^2, $(\Delta m^2_{31} + \Delta m^2_{32})/2 = 2.43 \times 10^{-3}$ eV^2, $\sin^2\theta_{13} = 0.024$, $\sin^2\theta_{12} = 0.307$.

The ability of discrimination in the ideal case is shown in Fig. 8.4. A sensitivity of $\Delta\chi^2 = 16$ is obtained for this case.

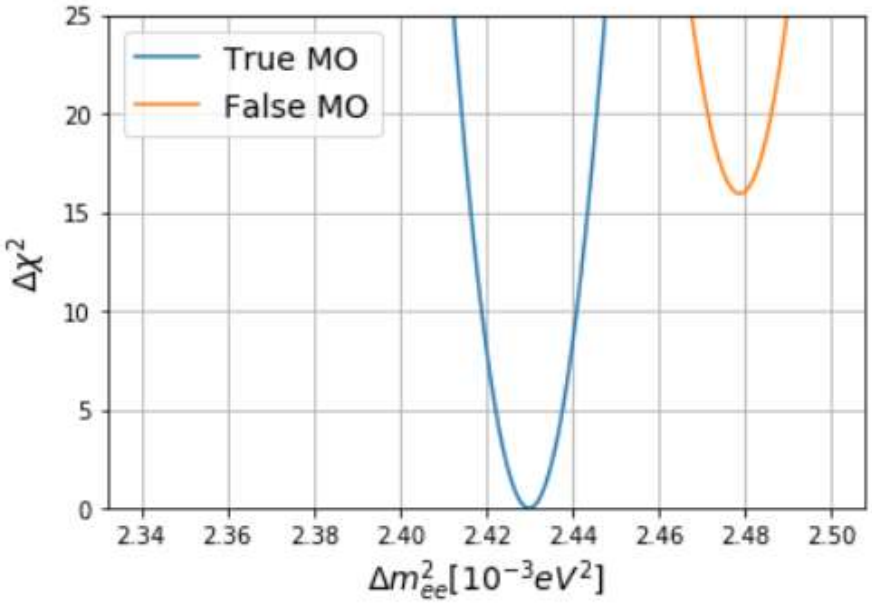


Figure 8.4: MO sensitivity for the ideal distributions of the reactor cores

8.4.2 Actual distributions

In this step we consider the actual power and baseline distributions, which are shown in Tab. 8.1. The main neutrino source for JUNO experiment is the nearby Yangjiang (YJ) and Taishan (TS) NPPs, the effect of remote reactors in the Daya Bay (DYB) and the possible Huizhou (HZ) NPP are also not negligible in the analysis. The effect of other NPPs is neglected in this study, since they are much further away from the experimental site (more than 400 km).

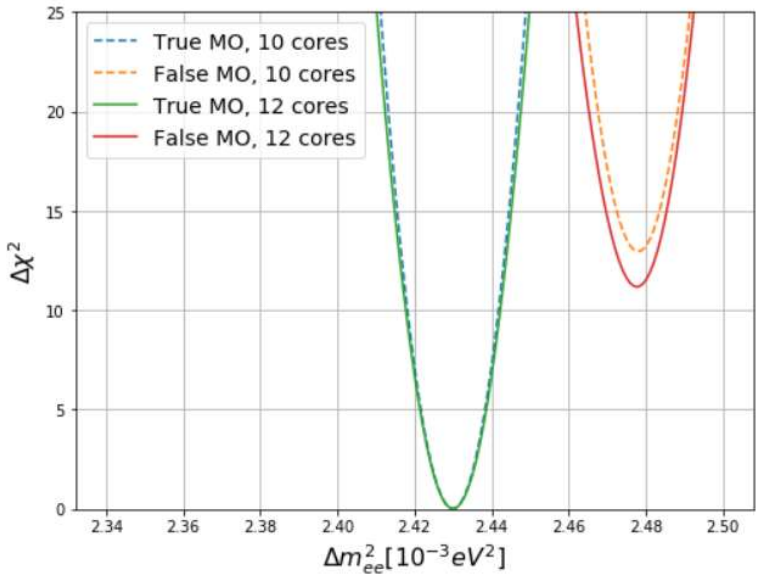


Figure 8.5: the comparison of MO sensitivity for the actual distributions of the reactor cores: dash is w/o DYB and HZ reactor cores; solid is w/ DYB and HZ reactor cores

The reduction of sensitivity due to the actual distribution of reactor cores is shown in Fig. 8.5, which gives a degradation of $\Delta\chi^2 = 4.7$. The degradation includes$\Delta\chi^2 = 3$ due to the baseline differences of Taishan and Yangjiang NPP, and$\Delta\chi^2 = 1.7$ with the inclusion of Daya Bay and Huizhou NPPs.

In all the following, the actual spacial distribution of reactor cores as shown in Tab. 8.1 is taken into account.

Cores	YJ-C1	YJ-C2	YJ-C3	YJ-C4	YJ-C5	YJ-C6
Power (GW)	2.9	2.9	2.9	2.9	2.9	2.9
Baseline (km)	52.75	52.84	52.42	52.51	52.12	52.21
Cores	TS-C1	TS-C2	TS-C3	TS-C4	DYB	HZ
Power (GW)	4.6	4.6	4.6	4.6	17.4	17.4
Baseline (km)	52.76	52.63	52.32	52.20	215	265

Table 8.1: Summary of the power and baseline distribution for the Yangjiang and Taishan reactor complexes, as well as the remote reactors of Daya Bay and Huizhou. [3]

8.4.3 Shape Uncertainty

In this step we add 1% bin-to-bin uncertainty, e.g., adding 1% uncertainty on each bin, which gives a degradation of $\Delta\chi^2 = 0.6$, as shown in Fig. 8.6.

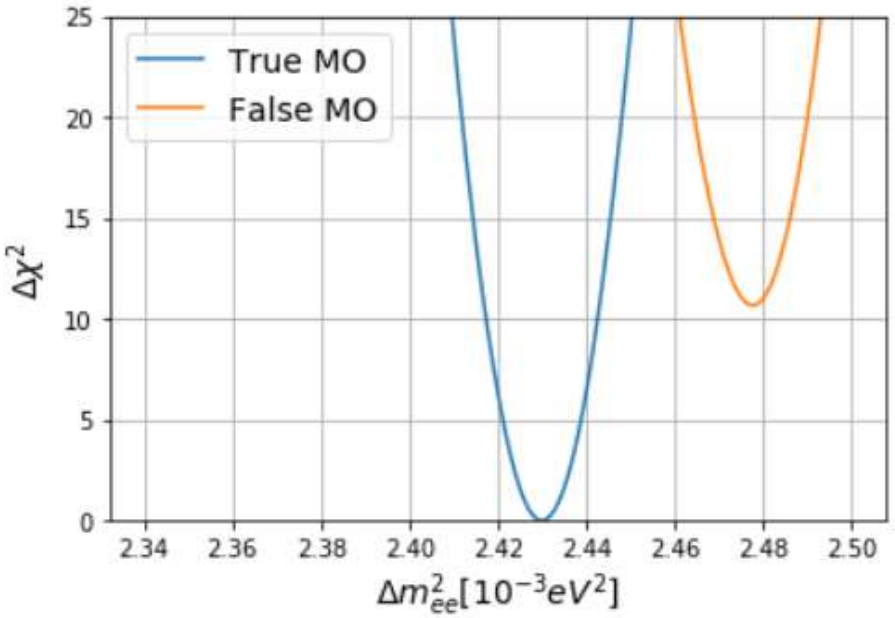


Figure 8.6: MO sensitivity for the case adding 1% bin-to-bin uncertainty

8.4.4 Background

In this step we add backgrounds including geo-neutrino background, accidental background, ^{9}Li/^{8}He background, ^{13}C$(\alpha, n)^{16}$O background, and fast neutron background, as shown in Tab. 8.2. The shape of the backgrounds are shown in the left panel of Fig. 8.7.

Event type	Rate (per day)	Rate uncertainty (relative)	Shape uncertainty
IBD candidates	60	-	-
Geo-νs	1.1	30%	5%
Accidental signals	0.9	1%	negligible
Fast-n	0.1	100%	20%
^{9}Li/^{8}He	1.6	20%	10%
^{13}C$(\alpha, n)^{16}$O	0.05	50%	50%

Table 8.2: The background summary table for the analysis of reactor antineutrinos [1]

From the above Table, the total background to signal (B/S) ratio is 6.3%, which reduces the MH sensitivity by $\Delta\chi^2 = 0.6$. The shape uncertainties of backgrounds contribute to a 0.4% bin-to-bin uncertainty, which further gives a reduction of $\Delta\chi^2 = 0.1$

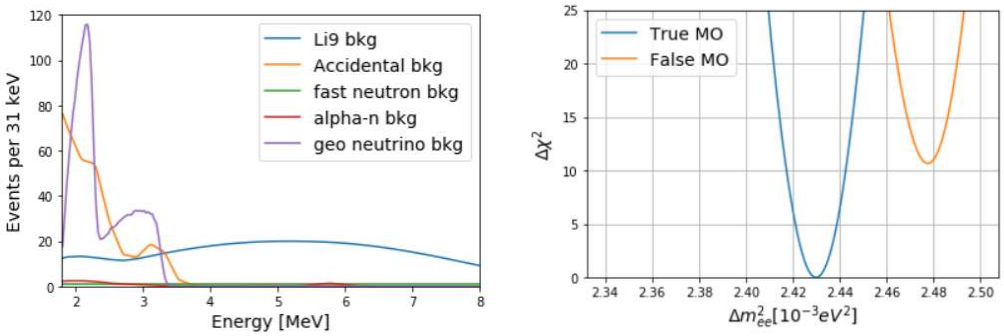


Figure 8.7: Left pannel: Spectra for the antineutrino signal and five kinds of main backgrounds, including the accidental, $^{8}He/^{9}Li$, fast neutron, and $^{13}C(\alpha, n)^{16}O$ and geo-neutrinos. Rate taken from [1].
Right Pannel: MO sensitivity for the case adding backgrounds. Both rate and shape uncertainty considered.

8.4.5 Update with PDG2020

In PDG2020, oscillation parameters change to:

$$\sin^2\theta_{12} = 0.307 \pm 0.013 \tag{8.18}$$
$$\Delta m^2_{21} = (7.53 \pm 0.18) \times 10^{-5} eV^2 \tag{8.19}$$
$$\sin^2\theta_{13} = (2.18 \pm 0.07) \times 10^{-2} \tag{8.20}$$
$$\tag{8.21}$$

With the new group of parameters, we get delta $\Delta\chi^2 = 7.6$. The degradation is due to the reduction of Δm^2_{21} and θ_{13} value.

8.4.6 Effect of energy resolution

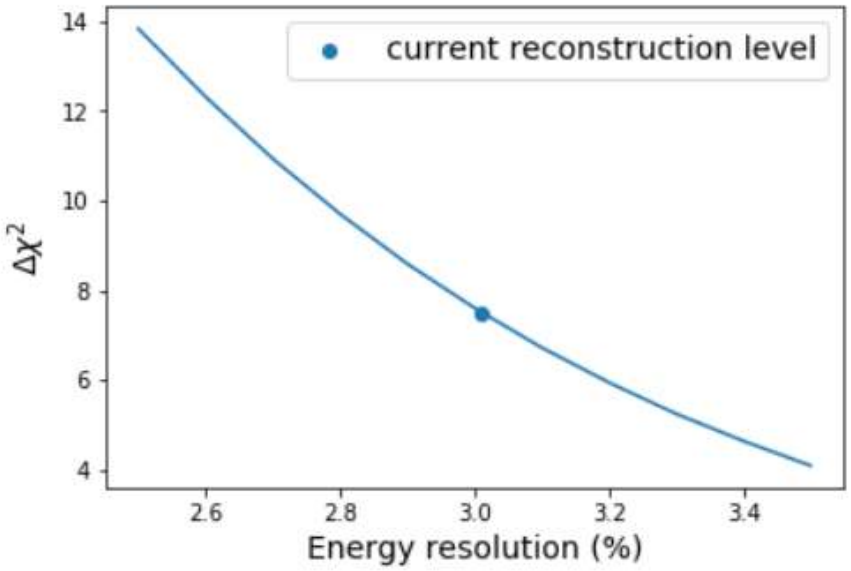


Figure 8.8: The MO discrimination ability of JUNO as functions of the detector energy resolution. $\Delta\chi^2$ is 7.5 at expected reconstruction level (3%)

In order to precisely measure the fast oscillations and discriminate the different

oscillation mode (IO/MO), which is driven by Δm^2_{31} and Δm^2_{21}, the energy resolution should be better than the size of $\Delta m^2_{21}/|\Delta m^2_{31}|$. The limit of energy resolution is mainly defined from the photon-electron statistics (1200 p.e./MeV), which corespond to 3%@1MeV. With the increase of photon-electron statistics and better reconstruction, the discrimination ability can increase significantly, as shown in Fig. 8.8. $\Delta\chi^2$ is 7.5 at expected reconstruction level (3%)

8.4.7 ^{210}Po contamination

In default, we think all α sources are from ^{238}U and ^{232}Th chain. However, ^{210}Po contamination can be out of equilibrium with ^{238}U chain, as found in Borexino experiment. The contamination value can reach to 70000 cpd/kton at the beginning of data taking, which is 26 times higher than the value in JUNO report. Thus, the count of ^{13}C$(\alpha, n)^{16}$O will also increase in this case.

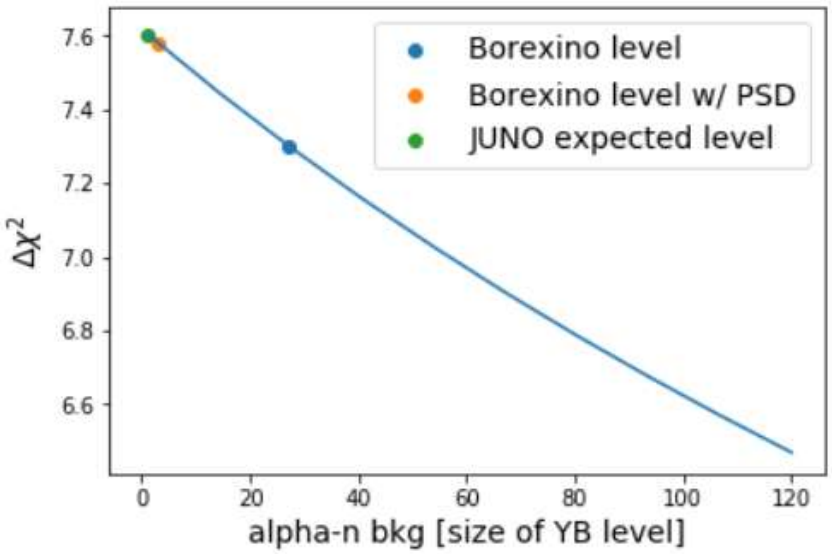


Figure 8.9: The MO discrimination ability of JUNO as functions of level of ^{13}C$(\alpha, n)^{16}$O background. The point of Borexino level and after PSD pointed in the curve, with $\Delta\chi^2$ value corresponding to 7.3 and 7.58 respectively.

In ^{13}C$(\alpha, n)^{16}$O reaction, the ^{16}O can be also in ground state or excited state.

When the ^{16}O is at ground state, the prompt signal is fast neutron, which can be discriminate with IBD prompt signal (positron) with high efficiency, as we discussed in Chapter. 7. In the experiment, more than 90% ^{16}O are at ground state [55], which can be discriminate with PSD.

Here we evaluate the effect of $^{13}\mathrm{C}(\alpha, n)^{16}\mathrm{O}$ background level to MO sensitivity, as shown in Fig. 8.9. When the value reach Borexino level, the $\Delta\chi^2$ value is degraded to 7.3; However, after PSD, the $\Delta\chi^2$ value can reach to 7.58.

8.5 Summary

In this chapter we introduce Bayes theory and the fit of mass ordering. With 6 years data taking, we are able to gain 3.16 σ's sensitivity to determine mass ordering with parameters in [1]. However, with the update neutrino oscillation parameter, the sensitivity is decreased by $\delta\chi^2 = 2.5$, corresponding to 0.42 σ. The degrade of the sensitivity is due to the decrease of the parameters θ_{13} and Δm^2_{21}. Here we show the effect of $\Delta\chi^2$ from each factors:

Optimal	+Baseline	+DYB&HZ	+Shape uncertainty	+Background
13.1	10.6	9.1	8.4	7.6

Table 8.3: Summary of effect of $\Delta\chi^2$ from each factors

Chapter 9

Conclusions

9.1 Summary

In this thesis, we study the potential of machine learning in different fields in JUNO experiment.

Waveform Reconstruction: we can reconstruct hit time of each single p.e., with wasserstain distance less than 4 ns.

Particle Identification: We can receive 95% signal with 5% background for alpha/beta and e/p discrimination, and 50% signal with 5% background for electron/positron discrimination.

Vertex/Energy reconstruction: 6 cm for vertex and 3% for energy at 1MeV.

9.2 Outlook

For machine learning models, there are usually tens of millions parameters. In current studies, computer scientists find than we can find a way to remove 90% to 95% of parameters with the performance not decrease. Although the reconstruction speed is hundreds times faster than traditional method, we can still accelerate the calculation and save the resources.

Now the reconstruction for different parts is independent. We should built a complete reconstruction workflow with machine learning method in the future. Particle identification and vertex reconstruction could based on waveform reconstruction results with machine learning method.

Appendix A

Borexino Experiment

Borexino aims to measure the neutrino flux from the sun. By now, Borexino is the only experiment which can measure the solar neutrino flux from different process (^{7}Be, ^{8}B, pp, pep, CNO) due to its low energy threshold, high energy resolution and low background level. In principle, Borexino detects all three flavor of solar neutrinos via their elastic scattering (ES) off electrons:

$$\nu_{e,\mu,\tau} + e^- \rightarrow \nu_{e,\mu,\tau} + e^- \tag{A.1}$$

According to the Standard Solar Model, we can calculate the solar neutrino flux from the different process at Borexino location, see in Fig A.1. Combined with neutrino oscillation theory, we can get the corresponding event rate in Borexino detector. Actually, there are only about tens to hundreds events per day for most neutrino species in Borexino detector. Thus, Borexino has extremely high requirement on the control of background level so as to measure the neutrino flux from the different process.

Generally, the backgrounds in the Borexino experiment come from two part: cosmic ray induced backgrounds and natural radioactivity. Borexino located in the Hall C of Laboratori Nazionali del Gran Sasso (LNGS), Italy, which is shielded by the mountain with about 1400 m rock cover (3800 m.w.e. (=meter water equivalent)). In this case, the muon flux is reduced by a factor of 10^6, corresponding to 1.2 muons

Solar-ν	Φ_ν(GS98) high metallicity [cm^{-2} s^{-1}]	Φ_ν(AGSS09) low metallicity [cm^{-2} s^{-1}]	R_ν(GS98) high metallicity [cpd/100 ton]	R_ν(AGSS09)low metallicity [cpd/100 ton]	Main background
pp	$5.98\ (1 \pm 0.006)$	$6.03\ (1 \pm 0.006)$	130.8 ± 2.4	131.9 ± 2.4	^{14}C
^{7}Be* (384 keV)	$0.53\ (1 \pm 0.07)$	$0.48\ (1 \pm 0.07)$	1.90 ± 0.14	1.73 ± 0.12	^{85}Kr, ^{210}Bi
^{7}Be* (862 keV)	$4.47\ (1 \pm 0.07)$	$4.08\ (1 \pm 0.07)$	46.48 ± 3.35	42.39 ± 3.05	^{85}Kr, ^{210}Bi
pep	$1.44\ (1 \pm 0.012)$	$1.47\ (1 \pm 0.012)$	2.73 ± 0.05	2.79 ± 0.06	^{11}C, ^{210}Bi
^{13}N	$2.96\ (1 \pm 0.14)$	$2.17\ (1 \pm 0.14)$	2.42 ± 0.34	1.78 ± 0.23	
^{15}O	$2.23\ (1 \pm 0.15)$	$1.56\ (1 \pm 0.15)$	2.75 ± 0.42	1.92 ± 0.29	
^{17}F	$5.52\ (1 \pm 0.17)$	$3.40\ (1 \pm 0.16)$	0.068 ± 0.012	0.042 ± 0.007	
CNO	$5.24\ (1 \pm 0.21)$	$3.76\ (1 \pm 0.21)$	5.24 ± 0.54	3.74 ± 0.37	^{11}C, ^{210}Bi
^{8}B	$5.58\ (1 \pm 0.14)$	$4.59\ (1 \pm 0.14)$	0.44 ± 0.07	0.37 ± 0.05	^{208}Tl, ext γ

Figure A.1: The event rate of different kinds of solar neutrino flux calculated by Standard Solar Model and their main background(s)[16]

m^{-2} hour^{-1} [16]. In Borexino experiment, we veto all the detector after each muon for 300 ms, making the residual rate of most cosmic isotopes negligible for the analysis.

The radioactive backgrounds in Borexino consist of two parts: inside the liquid scintillator or from the outside (Stainless Steel Sphere, PMT glass, etc). In Borexino, we use the most pure liquid scintillator in the world, with the contamination of 10^{-18} g/g for ^{238}U and ^{232}Th. To prevent the external backgrounds, Borexino constructs the detector with its unique structure, as follows:

The Borexino detector is consists of three parts: water tank, non-scintillating buffer and scintillator. The rigion in radius 4.5m is the scintillator part containing 278 ton liquid scintillator in the 125 μm thick spherical nylon ball. The active medium is a solution of PPO (2,5-diphenyloxazole, a fluorescent dye) in pseudo-cumene (PC, 1,2,4-trimethylbenzene) at a concentration of 1.5 g/l. [88] The nylon ball is called "Inner Vessel (IV)", seperating the scintillator part and non-scintillating buffer part. The rigion from 4.5 m to 6.85 m is the non-scintillating buffer part. At radius 5.5 m there is another nylon vessel called outer vessel (OV), which serving as a barrier against the backgrounds from outside. At radius 6.85 m is the stainless steel sphere (SSS), which encloses the central part of the detector. Besides, the stainless steel sphere is also served as the support of 2212 8" internal PMTs and 208 muon PMTs equipped on the outer side. The rigion out of 6.85 m is water in the 9 m base radius and 16.9 m height water tank, which tags the muons crossing the detector and shields the neutron and gamma from the rock.

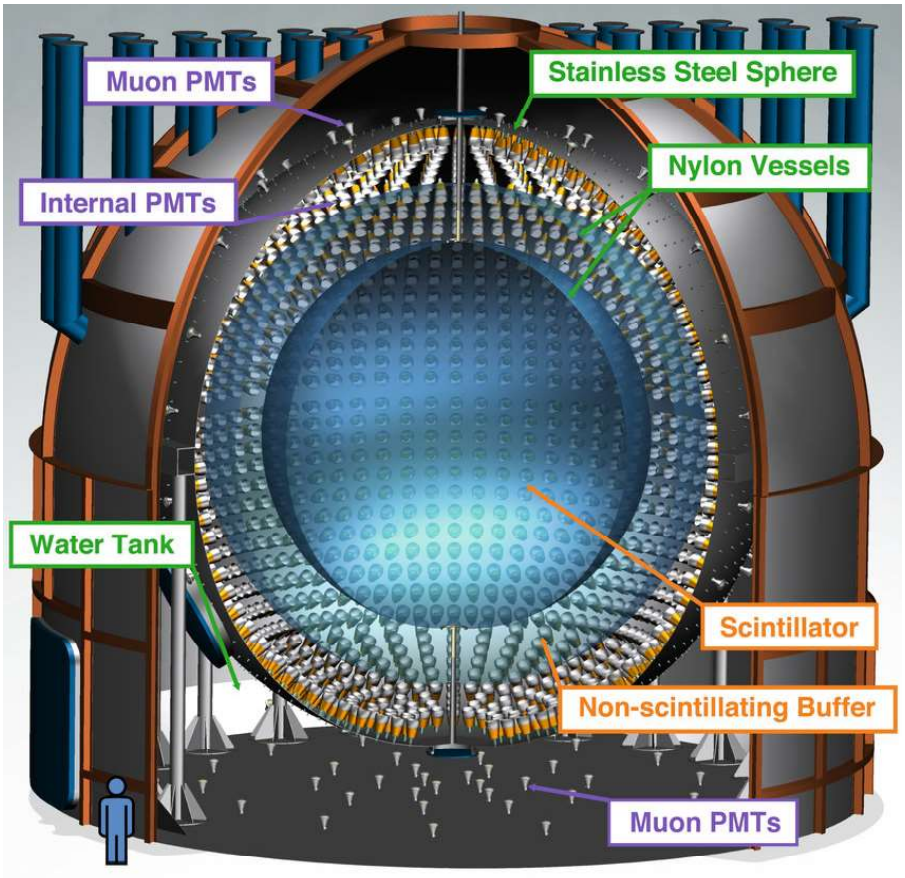


Figure A.2: Borexino detector

As we discussed above, the purity of Borexino's liquid scintillator is the key point for the solar neutrino study, thus additional to the effort on the liquid scintillator purification and detector design, scientists develop pulse shape discrimination (PSD) method to discriminate signals and backgrounds. Details see in Chapter 6.

Since 2007 when Borexino started taking data, Borexino has made lots of prograss on the understanding of solar and solar neutrinos:

2007: first real-time spectral measurement of sub-MeV (^{7}Be) solar neutrinos [89]. The measurement is consist with the Standard Solar Model

2010: Observe the geo neutrinos with the rate $3.9^{+1.6+5.8}_{-1.3-3.2}$/(100 ton $\cdot$ year). This measurement rejects the hypothesis of an active geo-reactor in the Earth's core with

a power above 3 TW at 95% C.L [90]

2010: The survival probabilities for ^{7}Be and ^{8}B neutrinos as measured by Borexino differ by 1.9 σ. These results are consistent with the prediction of the MSW-LMA solution of a transition in the solar ν_e survival probability Pee between the low energy vacuum-driven and the high-energy matter-enhanced solar neutrino oscillation regimes.

2012: observe the pep solar neutrino for the first time. [91]

2014: report spectral observations of pp neutrinos, demonstrating that about 99% of the power of the Sun is generated by the proton–proton fusion process. [92]

2020: 5 σ observation of CNO neutrinos [93].

A.1 Pulse Shape Discrimination (PSD)

For Borexino experiment, we also try machine learning method, same as we discussed in JUNO.

In Borexino, we take event samples from Bi214-Po214 cascade decay.

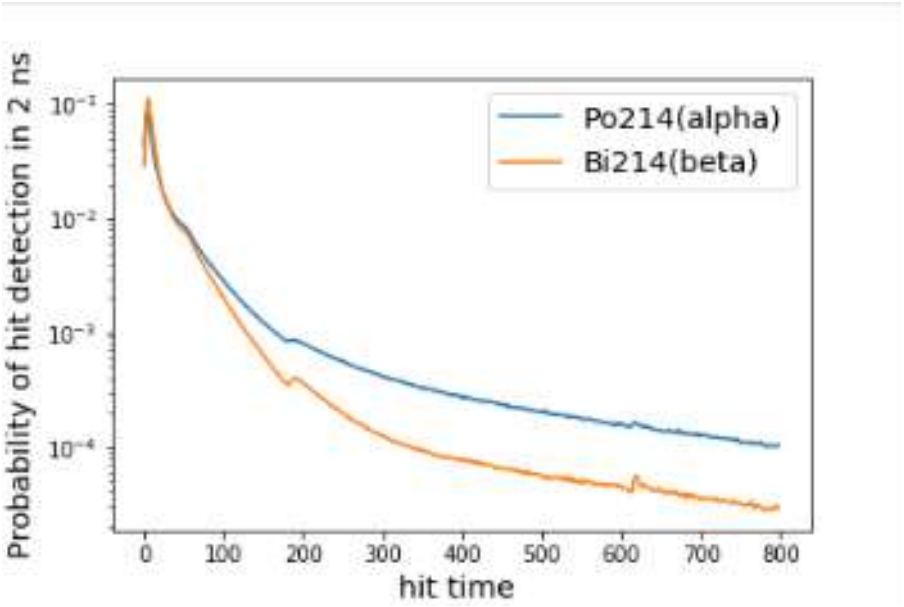


Figure A.3: Time profile of Borexino's alpha source and beta source. We take event samples from Bi214-Po214 cascade decay.

Neural Network model implemented on Borexino's data.

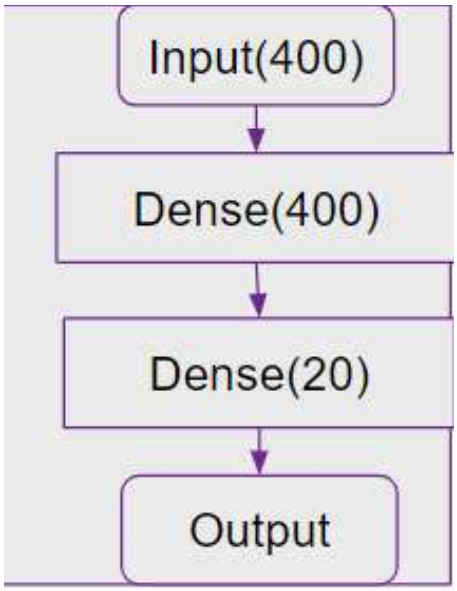


Figure A.4: Neural Network model implemented on Borexino's data.

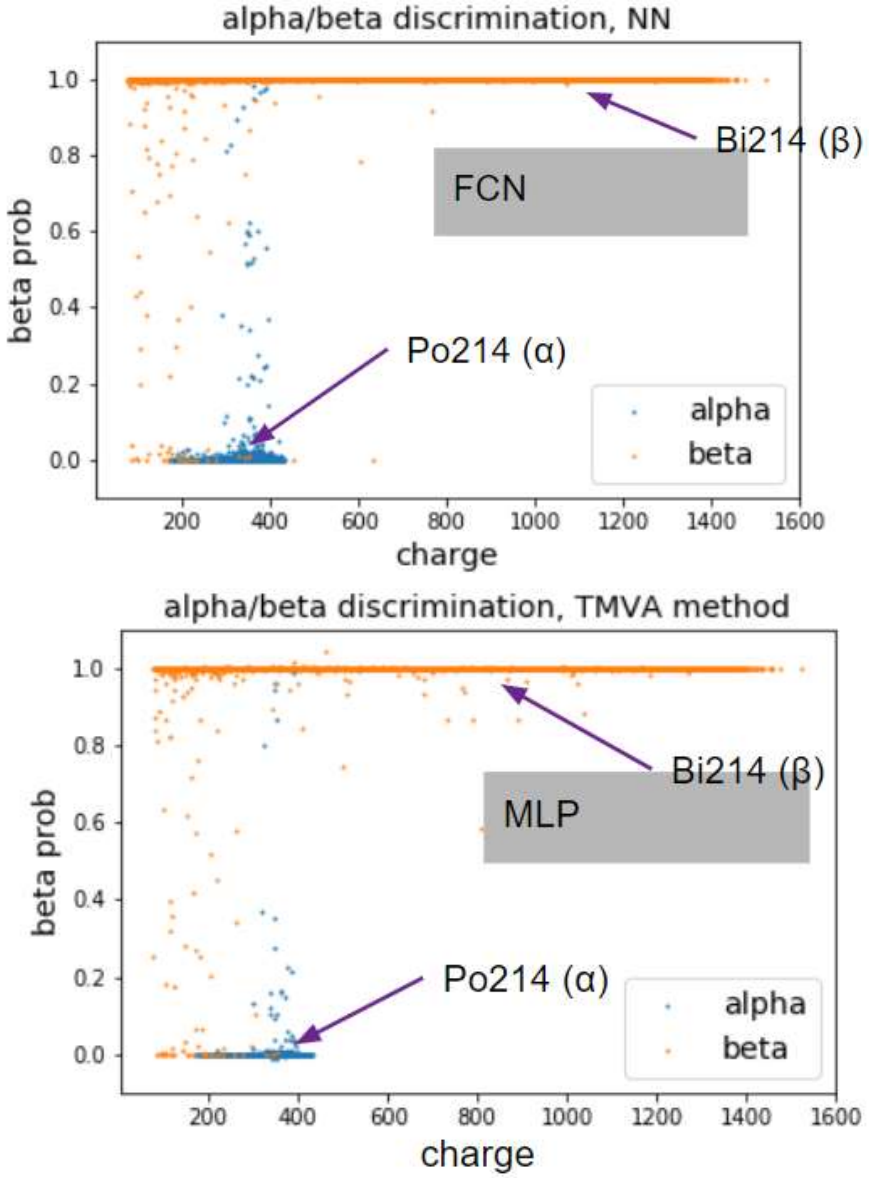


Figure A.5: Performance On Bi214 - Po214 test samples, both two methods accieve better than 99.5% accuracy. Top: performance of FCN filter; Bottom: performance of MLP filter.

We try to train on MC data and validate on real data. There is no much difference on it.

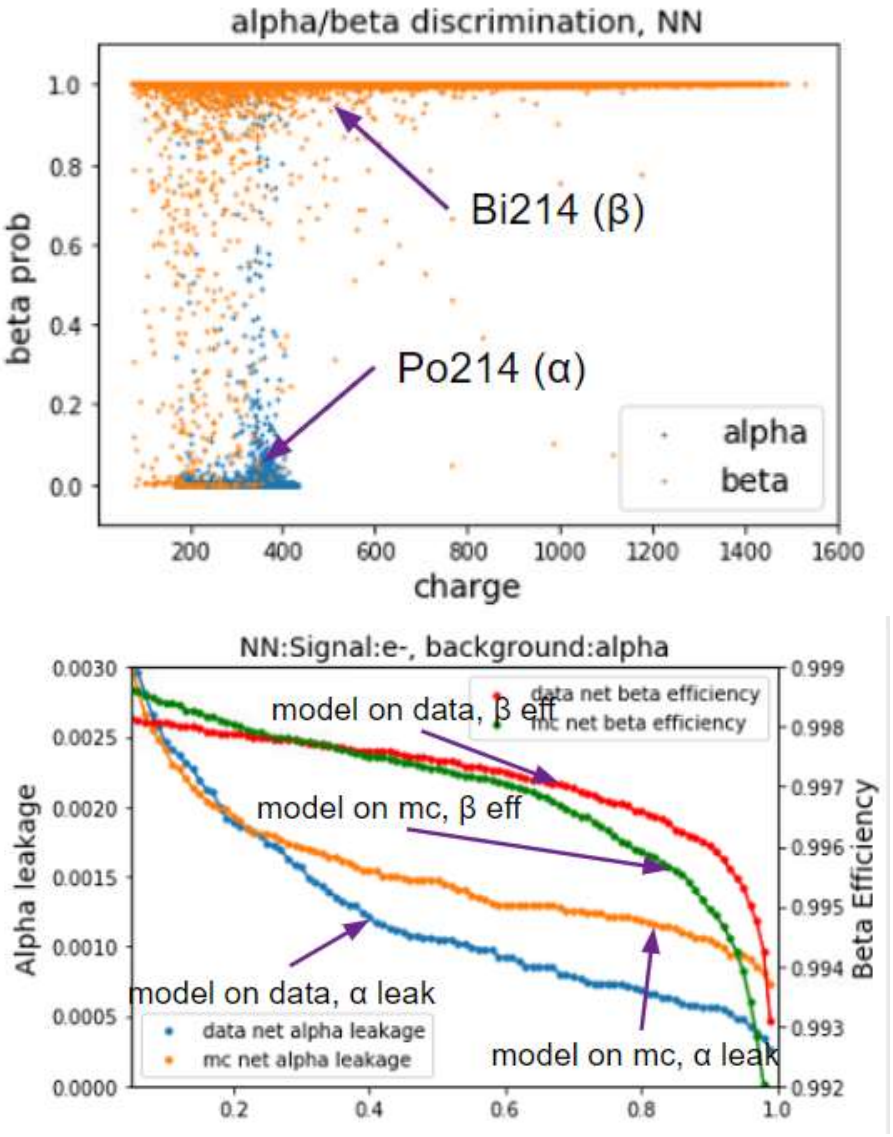


Figure A.6: Train on MC data and validate on real data. Top: performance of neural network trained on MC data. Bottom: The difference of performance of neural work trained on MC data and real data

A.2 Vertex Reconstruction

We also try to use neural network on vertex reconstruction with Borexino's data. Model is based on MC data, and the structure is shown as A.8.

First we project PMT hits on 2D image, as the examples shown in A.7:

Train strategy:

- Loss function: MSE
- Optimizer: Adam
- Init learning rate: 0.003
  - train 50 epoch
  - learning rate divide by 3 every 10 epoch
- Save the best group of parameters in all 50 epoch

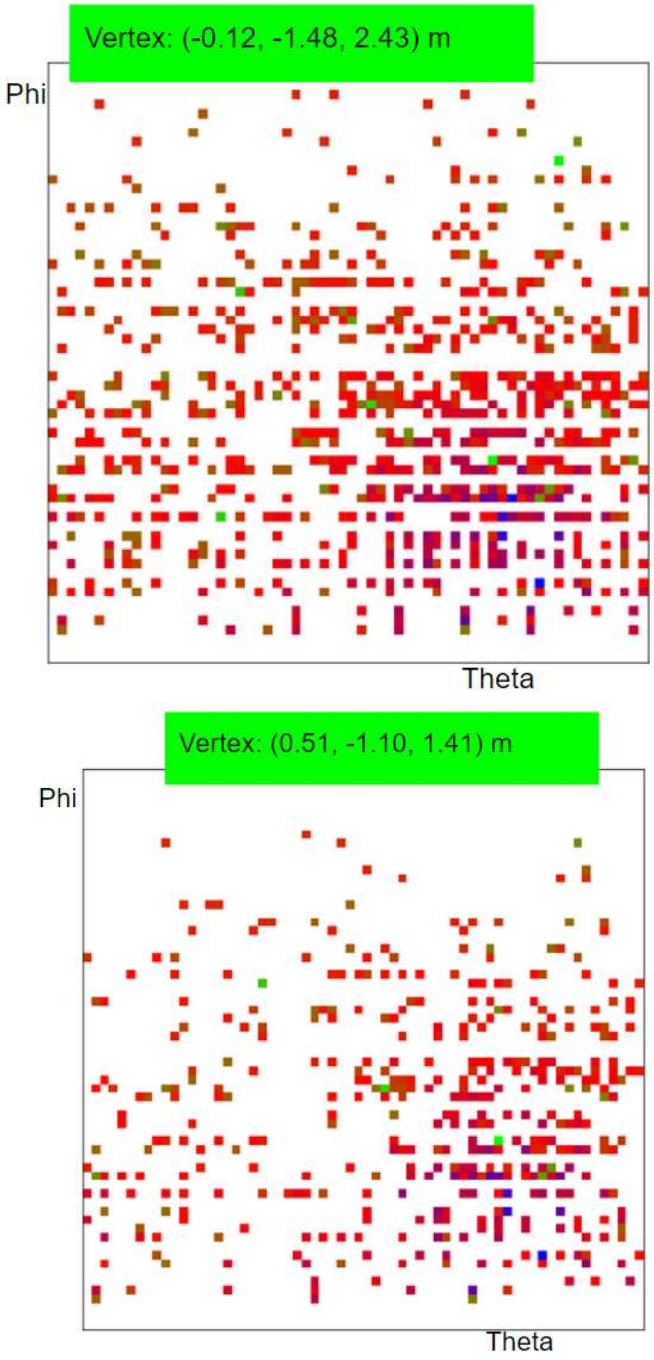


Figure A.7: Examples of 2D projection of Borexino data. Top: vertex at (-0.12, -1.48, 2.43) m; Bottom: vertex at (0.51, -1.10, 1.41) m

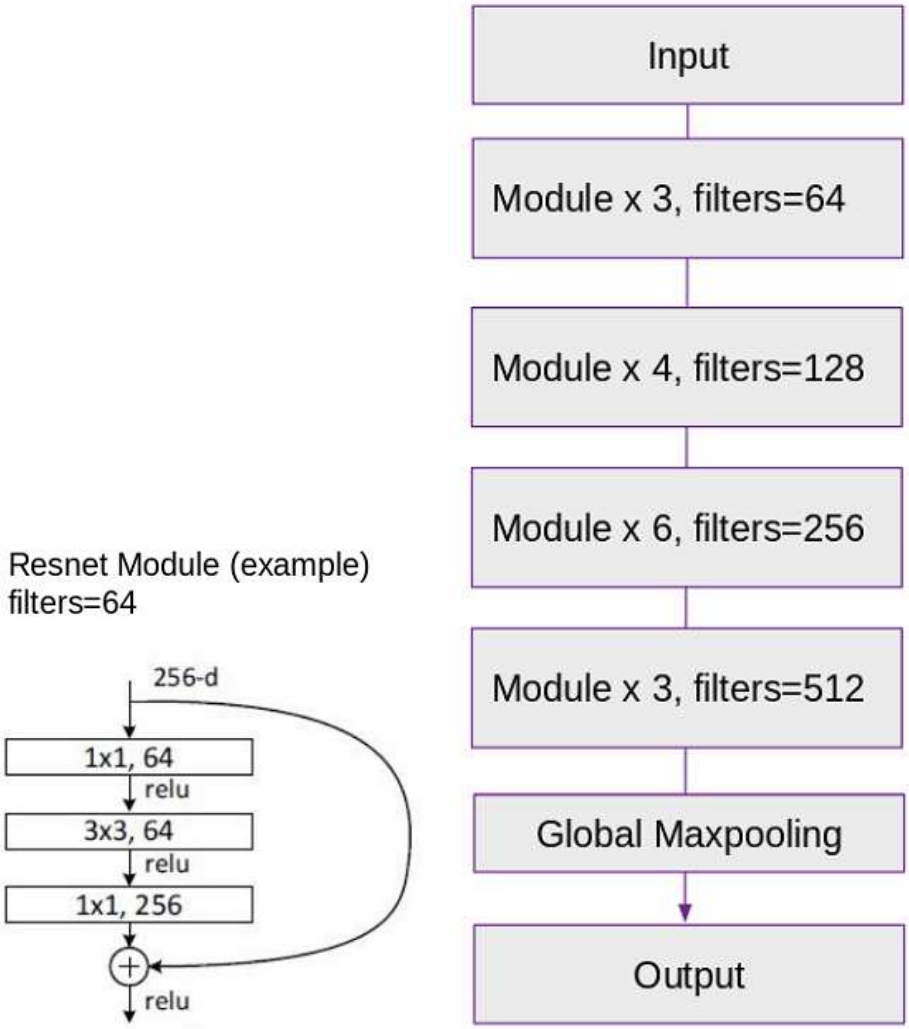


Figure A.8: The structure of the neural network. Left: the structure of one Resnet module; Right: the strurcture of whole network based on the module

Deep Learning (DL) can reconstruct similar result with Borexino's traditional reconstruction (LNGS)

	Deep Learning		LNGS	
	mean (cm)	std (cm)	mean (cm)	std (cm)
x	0.48	7.42	-0.32	7.72
y	0.38	7.35	-0.016	7.68
z	-0.37	7.58	-0.41	8.12

Figure A.9: Comparison of performance of vertex reconstruction on x, y, z axis

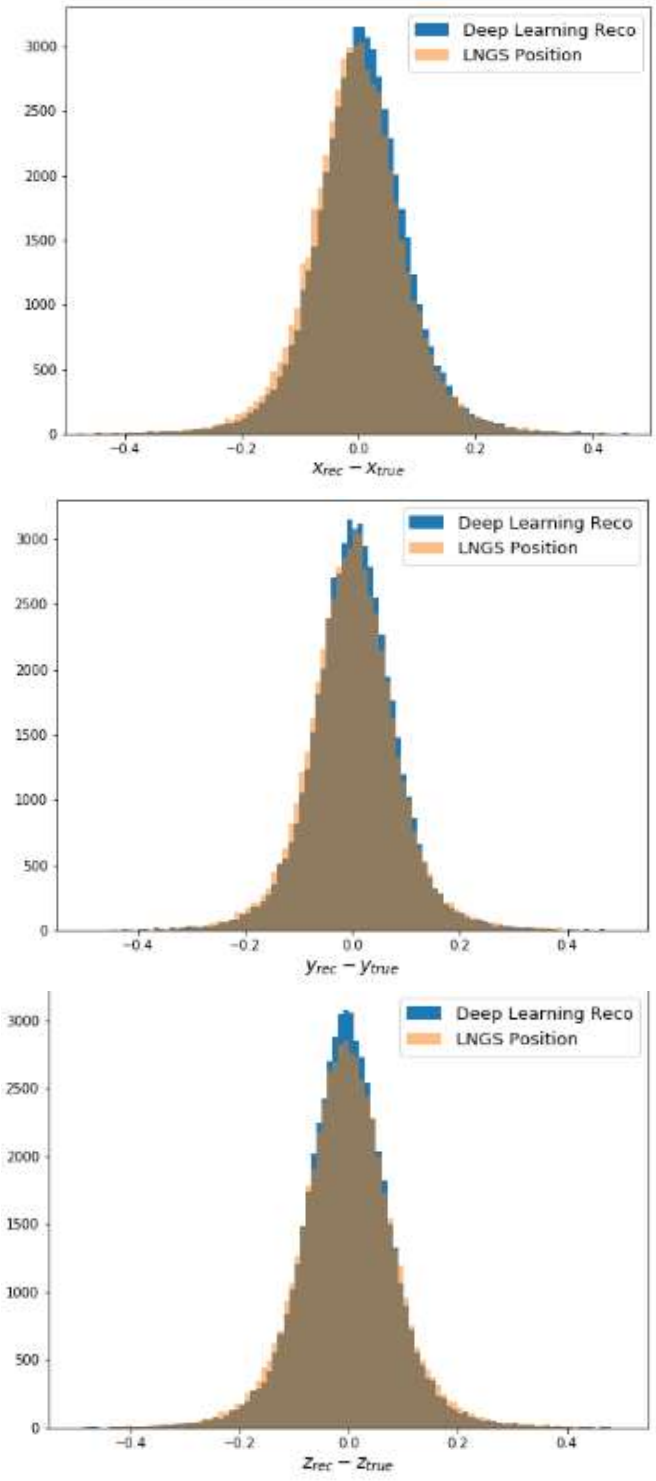


Figure A.10: Comparison of performance of vertex reconstruction on x, y, z axis. Blues represent Neural Network results, and gray represent traditional results.

Performance of bias vs radius:

- Bias increase as radius increase
- DL result and LNGS have similar performance on z axis
- DL result and LNGS have opposite trend on x axis

Performance of resolution vs radius:

- Resolution becomes better as radius increase
- DL has better performance for R>3 m, especially at z axis

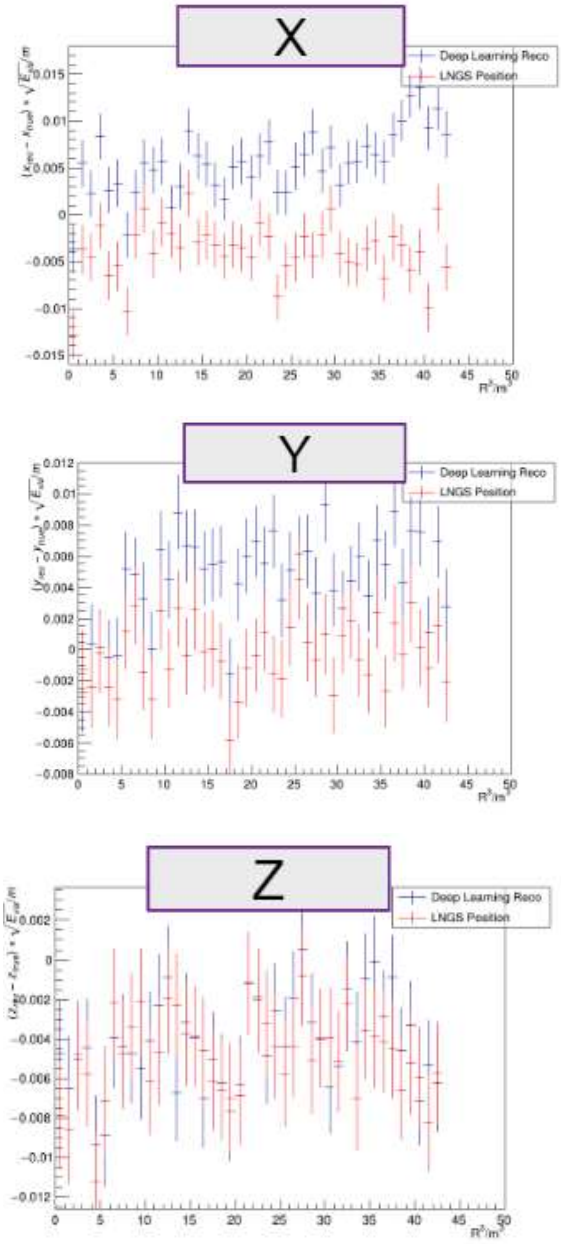


Figure A.11: Comparison of bias of vertex reconstruction vs radius on x, y, z axis

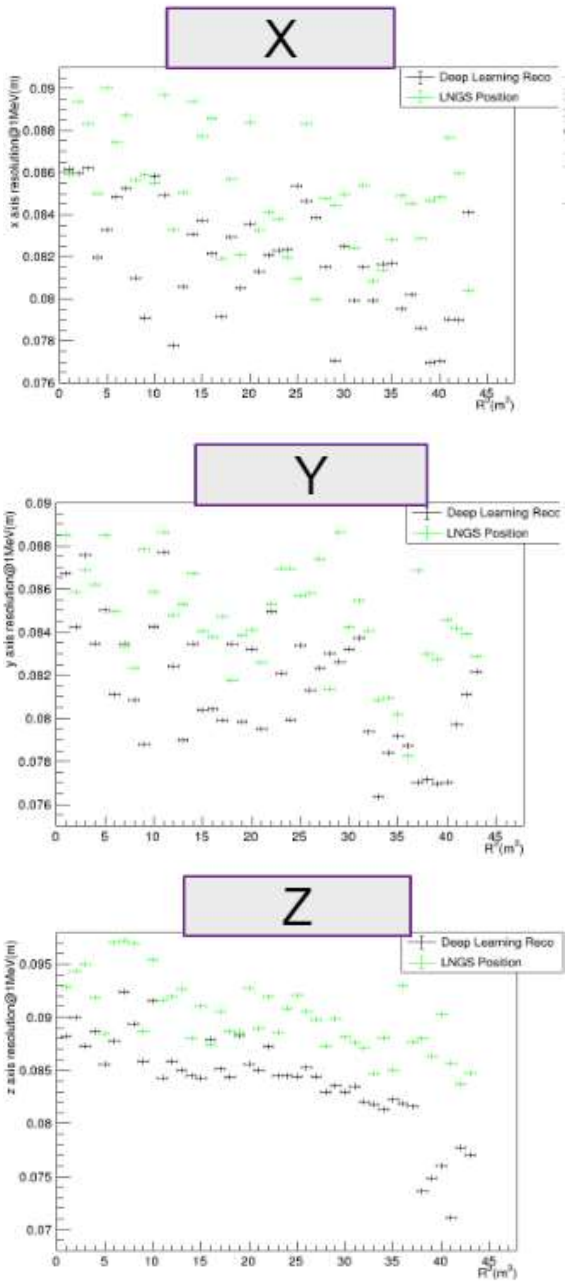


Figure A.12: Comparison of resolution of vertex reconstruction vs radius on x, y, z axis

Performance of bias of vertex reconstruction vs energy:

- Bias seems to be stable with energy
- DL result and LNGS have similar performance on z axis
- DL result and LNGS have opposite trend on x axis

Performance of resolution of vertex reconstruction vs energy:

- For energy < 0.5 MeV, LNGS has better performance
- For energy > 0.5 MeV, DL has better performance

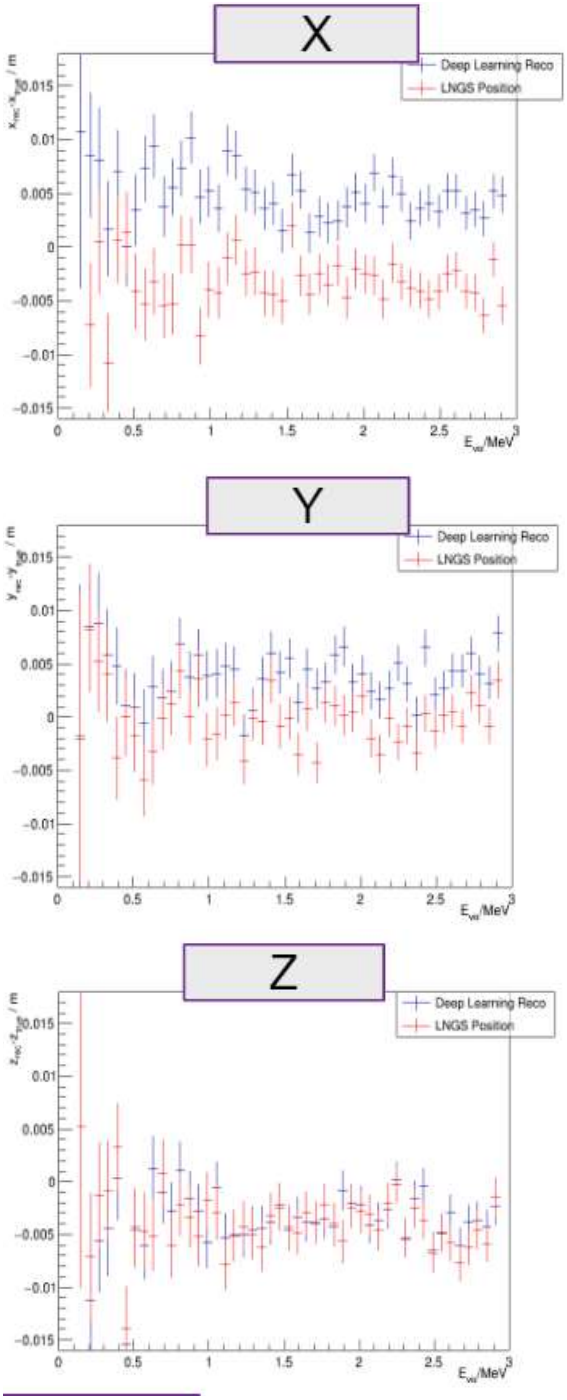
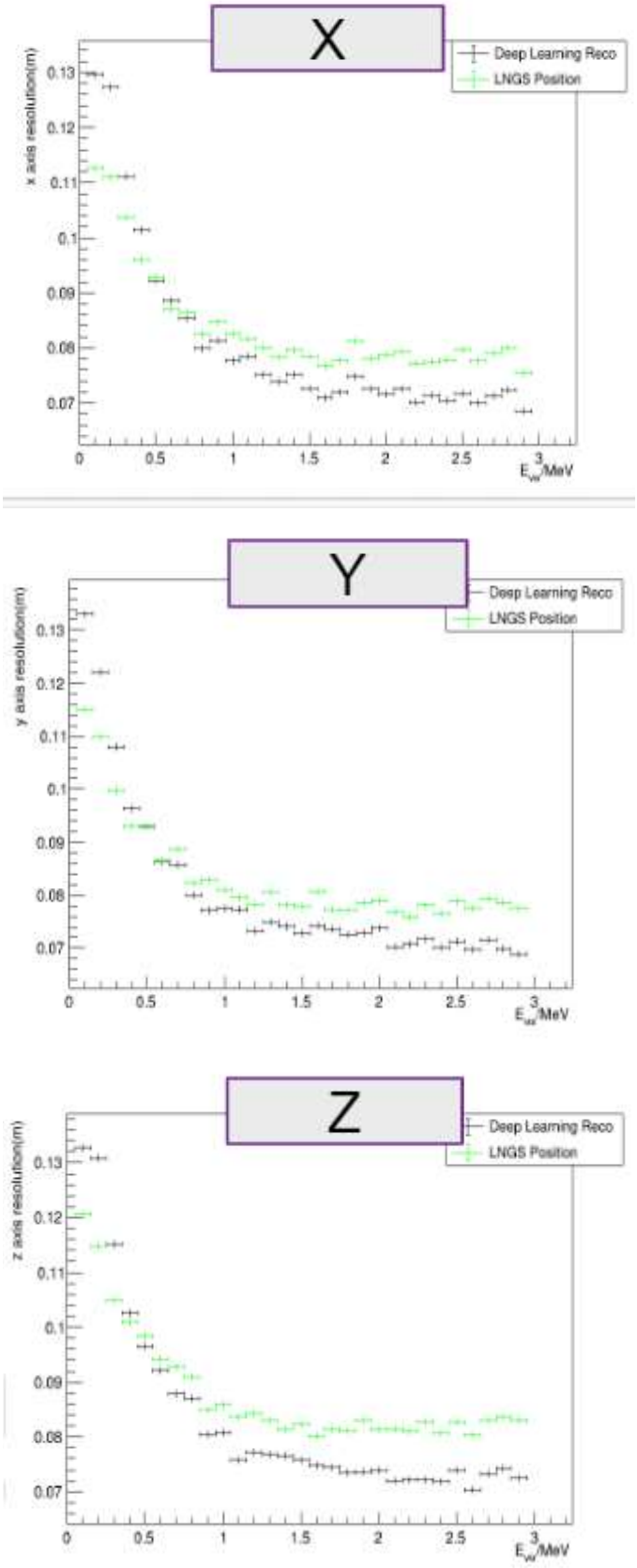


Figure A.13: Comparison of bias of vertex reconstruction vs energy on x, y, z axis


X
Deep Learning Reco
LNGS Position
x axis resolution(m)
0.13
0.12
0.11
0.1
0.09
0.08
0.07
0
0.5
1
1.5
2
2.5
3
E_vis/MeV
Y
Deep Learning Reco
LNGS Position
y axis resolution(m)
0.13
0.12
0.11
0.1
0.09
0.08
0.07
0
0.5
1
1.5
2
2.5
3
E_vis/MeV
Z
Deep Learning Reco
LNGS Position
z axis resolution(m)
0.13
0.12
0.11
0.1
0.09
0.08
0.07
0
0.5
1
1.5
2
2.5
3
E_vis/MeV

Bibliography

[1] Fengpeng An, Guangpeng An, Qi An, Vito Antonelli, Eric Baussan, John Beacom, Leonid Bezrukov, Simon Blyth, Riccardo Brugnera, Margherita Buizza Avanzini, and et al. Neutrino physics with juno. *Journal of Physics G: Nuclear and Particle Physics*, 43(3):030401, Feb 2016.

[2] Yongbo Huang, Jinfan Chang, Yaping Cheng, Zhang Chen, Jun Hu, Xiaolu Ji, Fei Li, Jin Li, Qiuju Li, Xin Qian, and et al. The flash adc system and pmt waveform reconstruction for the daya bay experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 895:48–55, Jul 2018.

[3] Yu-Feng Li, Jun Cao, Yifang Wang, and Liang Zhan. Unambiguous determination of the neutrino mass hierarchy using reactor neutrinos. *Physical Review D*, 88(1), Jul 2013.

[4] J. A. Formaggio and G. P. Zeller. From ev to eev: Neutrino cross sections across energy scales. *Reviews of Modern Physics*, 84(3):1307–1341, Sep 2012.

[5] Ivan Esteban, M. C. Gonzalez-Garcia, Alvaro Hernandez-Cabezudo, Michele Maltoni, and Thomas Schwetz. Global analysis of three-flavour neutrino oscillations: synergies and tensions in the determination of 23, cp, and the mass ordering. *Journal of High Energy Physics*, 2019(1), Jan 2019.

[6] Mattias Blennow, Pilar Coloma, Patrick Huber, and Thomas Schwetz. Quantifying the sensitivity of oscillation experiments to the neutrino mass ordering. *Journal of High Energy Physics*, 2014(3), Mar 2014.

[7] S. T. Petcov. The nature of the neutrino (dirac/majorana) and double beta decay with or without neutrinos, 2019.

[8] K. Abe and et al. Constraint on the matter-antimatter symmetry-violating phase in neutrino oscillations, 2019.

[9] WB Oatway, Kelly Jones, S Holmes, George Etherington, J Marsh, RGE Haylock, and Colin Muirhead. Background material on nature of radioactive contamination at the university of manchester and possible health risks: Task 1 report. *HPA Radiological Protection Division RPD-EA-6-2009*, 01 2009.

[10] Qiaomu Zhu, Jinfu Chen, Lin Zhu, Xianzhong Duan, and Yilu Liu. Wind speed prediction with spatio–temporal correlation: A deep learning approach. *Energies (Basel)*, 11(4), 3 2018.

[11] Soheil Kolouri, Phillip E. Pope, Charles E. Martin, and Gustavo K. Rohde. Sliced-wasserstein autoencoder: An embarrassingly simple generative model, 2018.

[12] Livia Ludhova, Henning Rebber, Björn Soenke Wonsak, and Yu Xu. Particle identification at mev energies in juno, 2020.

[13] Xiaobo Li. *PhD Thesis*. PhD thesis, IHEP, 5 2011.

[14] Yan Zhang, Ze-Yuan Yu, Xin-Ying Li, Zi-Yan Deng, and Liang-Jian Wen. A complete optical model for liquid-scintillator detectors, 2020.

[15] Qin Liu, Miao He, Xuefeng Ding, Liangjian Wen, Weidong Li, and Haiping Peng. A vertex reconstruction algorithm in the central detector of juno. *Journal of Instrumentation*, 13, 03 2018.

[16] G. Bellini, J. Benziger, D. Bick, G. Bonfini, D. Bravo, M. Buizza Avanzini, B. Caccianiga, L. Cadonati, F. Calaprice, P. Cavalcante, and et al. Final results of borexino phase-i on low-energy solar neutrino spectroscopy. *Physical Review D*, 89(11), Jun 2014.

[17] Raymond Davis. A review of the homestake solar neutrino experiment. *Progress in Particle and Nuclear Physics*, 32:13 – 32, 1994.

[18] D. Vignaud. The gallex solar neutrino experiment. *Nuclear Physics B - Proceedings Supplements*, 60(3):20 – 29, 1998.

[19] M. Altmann, M. Balata, P. Belli, E. Bellotti, R. Bernabei, E. Burkert, C. Cattadori, R. Cerulli, M. Chiarini, M. Cribier, S. d'Angelo, G. Del Re, K.H. Ebert, F. von Feilitzsch, N. Ferrari, W. Hampel, F.X. Hartmann, E. Henrich, G. Heusser, F. Kaether, J. Kiko, T. Kirsten, T. Lachenmaier, J. Lanfranchi, M. Laubenstein, K. Lützenkirchen, K. Mayer, P. Moegel, D. Motta, S. Nisi, J. Oehm, L. Pandola, F. Petricca, W. Potzel, H. Richter, S. Schoenert, M. Wallenius, M. Wojcik, and L. Zanotti. Complete results for five years of gno solar neutrino observations. *Physics Letters B*, 616(3):174 – 190, 2005.

[20] J.N. Abdurashitov, E.L. Faizov, V.N. Gavrin, A.O. Gusev, A.V. Kalikhov, T.V. Knodel, I.I. Knyshenko, V.N. Kornoukhov, I.N. Mirmov, A.M. Pshukov, A.M. Shalagin, A.A. Shikhin, P.V. Timofeyev, E.P. Veretenkin, V.M. Vermul, G.T. Zatsepin, T.J. Bowles, J.S. Nico, W.A. Teasdale, D.L. Wark, J.F. Wilkerson, B.T. Cleveland, T. Daily, R. Davis, K. Lande, C.K. Lee, P.W. Wildenhain, S.R. Elliott, M.L. Cherry, and R.T. Kouzes. Results from sage (the russian-american gallium solar neutrino experiment). *Physics Letters B*, 328(1):234 – 248, 1994.

[21] Christopher W. Walter. The super-kamiokande experiment. *Neutrino Oscillations*, page 19–43, Mar 2008.

[22] John M. LoSecco. The history of "anomalous" atmospheric neutrino events: A first person account. *Physics in Perspective*, 18(2):209–241, Jul 2016.

[23] Ashie and et al. Evidence for an oscillatory signature in atmospheric neutrino oscillations. *Phys. Rev. Lett.*, 93:101801, Sep 2004.

[24] F. P. An, J. Z. Bai, A. B. Balantekin, H. R. Band, D. Beavis, W. Beriguete, M. Bishai, S. Blyth, K. Boddy, R. L. Brown, and et al. Observation of electron-antineutrino disappearance at daya bay. *Physical Review Letters*, 108(17), Apr 2012.

[25] J. K. Ahn, S. Chebotaryov, J. H. Choi, S. Choi, W. Choi, Y. Choi, H. I. Jang, J. S. Jang, E. J. Jeon, I. S. Jeong, and et al. Observation of reactor electron antineutrinos disappearance in the reno experiment. *Physical Review Letters*, 108(19), May 2012.

[26] Y. Abe, J. C. dos Anjos, J. C. Barriere, E. Baussan, I. Bekman, M. Bergevin, T. J. C. Bezerra, L. Bezrukov, E. Blucher, and et al. Improved measurements of the neutrino mixing angle 13 with the double chooz detector. *Journal of High Energy Physics*, 2014(10), Oct 2014.

[27] S. Fukuda et al. Solar B-8 and hep neutrino measurements from 1258 days of Super-Kamiokande data. *Phys. Rev. Lett.*, 86:5651–5655, 2001.

[28] J.J. Simpson. First results from SNO. *Nuovo Cim. B*, 117:1227–1236, 2002.

[29] Eguchi et al. First results from kamland: Evidence for reactor antineutrino disappearance. *Phys. Rev. Lett.*, 90:021802, Jan 2003.

[30] Hyunkwan Seo. Status of RENO-50. *PoS*, NEUTEL2015:083, 2015.

[31] K. Abe, H. Aihara, C. Andreopoulos, I. Anghel, A. Ariga, T. Ariga, R. Asfandiyarov, M. Askins, J. J. Back, and et al. Physics potential of a long-baseline neutrino oscillation experiment using a j-parc neutrino beam and hyper-kamiokande. *Progress of Theoretical and Experimental Physics*, 2015(5):53C02–0, May 2015.

[32] DUNE Collaboration. The dune far detector interim design report volume 1: Physics, technology and strategies, 2018.

[33] Daya Bay Collaboration. A precision measurement of the neutrino mixing angle $\mathrm{theta}_1 3 using reactor antineutrinos at daya bay, 2007.$

[34] J. B. Albert, G. Anton, I. J. Arnquist, I. Badhrees, P. Barbeau, D. Beck, V. Belov, F. Bourque, J. P. Brodsky, E. Brown, and et al. Sensitivity and discovery potential of the proposed nexo experiment to neutrinoless double- decay. *Physical Review C*, 97(6), Jun 2018.

[35] Jordan Myslik. Legend: The large enriched germanium experiment for neutrinoless double-beta decay, 2018.

[36] R. Nakamura, H. Sambonsugi, K. Shiraishi, and Y. Wada. Research and development toward KamLAND2-zen. *Journal of Physics: Conference Series*, 1468:012256, feb 2020.

[37] Xun Chen et al. Pandax-iii: Searching for neutrinoless double beta decay with high pressure ^{136}xe gas time projection chambers, 2016.

[38] Delia Tosi. The search for neutrino-less double-beta decay: summary of current experiments, 2014.

[39] A. Gando, Y. Gando, T. Hachiya, A. Hayashi, S. Hayashida, H. Ikeda, K. Inoue, K. Ishidoshiro, Y. Karino, M. Koga, and et al. Search for majorana neutrinos near the inverted mass hierarchy region with kamland-zen. *Physical Review Letters*, 117(8), Aug 2016.

[40] Background-free search for neutrinoless double- decay of 76ge with gerda. *Nature*, 544(7648):47–52, Apr 2017.

[41] Jie Zhao, Liang-Jian Wen, Yi-Fang Wang, and Jun Cao. Physics potential of searching for 0 decays in juno. *Chinese Physics C*, 41(5):053001, May 2017.

[42] Antonio J. Cuesta, Viviana Niro, and Licia Verde. Neutrino mass limits: robust information from the power spectrum of galaxy surveys, 2015.

[43] J. Angrik et al. KATRIN design report 2004. 2005.

[44] M. Aker. et al. An improved upper limit on the neutrino mass from a direct kinematic method by katrin, 2019.

[45] M. Tanabashi et al. Review of particle physics. *Phys. Rev. D*, 98:030001, Aug 2018.

[46] G. Mention, M. Fechner, Th. Lasserre, Th. A. Mueller, D. Lhuillier, M. Cribier, and A. Letourneau. Reactor antineutrino anomaly. *Physical Review D*, 83(7), Apr 2011.

[47] Carlo Giunti and Marco Laveder. Statistical significance of the gallium anomaly. *Physical Review C*, 83(6), Jun 2011.

[48] A. Aguilar, L. B. Auerbach, R. L. Burman, D. O. Caldwell, E. D. Church, A. K. Cochran, J. B. Donahue, A. Fazely, G. T. Garvey, R. M. Gunasingha, and et al. Evidence for neutrino oscillations from the observation of ν_e appearance in a $\bar{}$beam. *Physical Review D*, 64(11), Nov 2001.

[49] A.A. Aguilar-Arevalo, B.C. Brown, L. Bugel, G. Cheng, J.M. Conrad, R.L. Cooper, R. Dharmapalan, A. Diaz, Z. Djurcic, D.A. Finley, and et al. Significant excess of electronlike events in the miniboone short-baseline neutrino experiment. *Physical Review Letters*, 121(22), Nov 2018.

[50] R. Acciarri, C. Adams, R. An, A. Aparicio, S. Aponte, J. Asaadi, M. Auger, N. Ayoub, L. Bagby, B. Baller, and et al. Design and construction of the microboone detector. *Journal of Instrumentation*, 12(02):P02017–P02017, Feb 2017.

[51] J. Ashenfelter, A.B. Balantekin, C. Baldenegro, H.R. Band, C.D. Bass, D.E. Bergeron, D. Berish, L.J. Bignell, N.S. Bowden, J. Boyle, and et al. The prospect reactor antineutrino experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 922:287–309, Apr 2019.

[52] M. Andriamirado et al. Improved short-baseline neutrino oscillation search and energy spectrum measurement with the prospect experiment at hfir, 2020.

[53] Y. Abreu, Y. Amhis, L. Arnold, G. Ban, W. Beaumont, M. Bongrand, D. Boursette, J.M. Buhour, B.C. Castle, K. Clark, and et al. A novel segmented-scintillator antineutrino detector. *Journal of Instrumentation*, 12(04):P04024–P04024, Apr 2017.

[54] Liang-jian Wen, Jun Cao, Kam-Biu Luk, Yu-qian Ma, Yi-fang Wang, and Chang-gen Yang. Measuring cosmogenic Li-9 background in a reactor neutrino experiment. *Nucl. Instrum. Meth. A*, 564:471–474, 2006.

[55] Jie Zhao, Zeyuan Yu, Jianglai Liu, Xiaobo Li, Feihong Zhang, and Dongmei Xia. 13c(alpha,n)16o background in a liquid scintillator based neutrino experiment, 2013.

[56] S. T. Dye. Geoneutrinos and the radioactive power of the earth. *Reviews of Geophysics*, 50(3), 2012.

[57] Ran Han, Yu-Feng Li, Liang Zhan, William F McDonough, Jun Cao, and Livia Ludhova. Potential of geo-neutrino measurements at juno. *Chinese Physics C*, 40(3):033003, Mar 2016.

[58] K. Abe, Y. Haga, Y. Hayato, M. Ikeda, K. Iyogi, J. Kameda, Y. Kishimoto, M. Miura, S. Moriyama, M. Nakahata, and et al. Real-time supernova neutrino burst monitor at super-kamiokande. *Astroparticle Physics*, 81:39–48, Aug 2016.

[59] R. Abbasi, Y. Abdou, T. Abu-Zayyad, M. Ackermann, J. Adams, J. A. Aguilar, M. Ahlers, M. M. Allen, D. Altmann, and et al. Icecube sensitivity for low-energy neutrinos from nearby supernovae. *Astronomy Astrophysics*, 535:A109, Nov 2011.

[60] Artur Ankowski, John Beacom, Omar Benhar, Sun Chen, John Cherry, Yanou Cui, Alexander Friedland, Ines Gil-Botella, Alireza Haghighat, Shunsaku Horiuchi, Patrick Huber, James Kneller, Ranjan Laha, Shirley Li, Jonathan Link, Alessandro Lovato, Oscar Macias, Camillo Mariani, Anthony Mezzacappa, Evan O'Connor, Erin O'Sullivan, Andre Rubbia, Kate Scholberg, and Tatsu Takeuchi. Supernova physics at dune, 2016.

[61] A. Aurisano, A. Radovic, D. Rocco, A. Himmel, M.D. Messier, E. Niner, G. Pawloski, F. Psihas, A. Sousa, and P. Vahle. A convolutional neural network neutrino event classifier. *Journal of Instrumentation*, 11(09):P09001–P09001, Sep 2016.

[62] Pierre Baldi, Jianming Bian, Lars Hertel, and Lingge Li. Improved energy reconstruction in nova with regression convolutional neural networks. *Physical Review D*, 99(1), Jan 2019.

[63] Shiqi Yu. Electron neutrino energy reconstruction in nova using cnn particle ids, 2019.

[64] C. Adams, M. Alrashed, R. An, J. Anthony, J. Asaadi, A. Ashkenazi, M. Auger, S. Balasubramanian, B. Baller, C. Barnes, and et al. Deep neural network for pixel-level electromagnetic particle identification in the microboone liquid argon time projection chamber. *Physical Review D*, 99(9), May 2019.

[65] Lauren E. Yates. Microboone investigation of low-energy excess using deep learning algorithms, 2017.

[66] Kim Albertsson and et al. Machine learning in high energy physics community white paper, 2018.

[67] G. Cybenko. Approximation by superpositions of a sigmoidal function, 1989.

[68] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5):359 – 366, 1989.

[69] Steven W. Smith. *The Scientist and Engineer's Guide to Digital Signal Processing*. California Technical Publishing, USA, 1997.

[70] Victor M. Panaretos and Yoav Zemel. Statistical aspects of wasserstein distances. *Annual Review of Statistics and Its Application*, 6(1):405–431, Mar 2019.

[71] Cédric Villani. *The Wasserstein distances*, pages 93–111. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.

[72] Borexino Collaboration. Final results of borexino phase-i on low-energy solar neutrino spectroscopy. *Phys. Rev. D*, 89:112007, Jun 2014.

[73] Y. Abe and et al. Ortho-positronium observation in the double chooz experiment. *Journal of High Energy Physics*, 2014(10), Oct 2014.

[74] Stefan Wagner. *Energy non-linearity studies and pulse shape analysis of liquid scintillator signals in the Double Chooz experiment*. PhD thesis, Heidelberg U., 2014.

[75] C. A. Parker and C. G. Hatchard. Triplet-singlet emission in fluid solutions. phosphorescence of eosin. *Trans. Faraday Soc.*, 57:1894–1904, 1961.

[76] C. A. Parker and C. G. Hatchard. Delayed fluorescence from solutions of anthracene and phenanthrene. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 269(1339):574–584, 1962.

[77] G. Laustriat. The luminescence decay of organic scintillators. *Molecular Crystals*, 4(1-4):127–145, 1968.

[78] B.S. Wonsak, C.I. Hagner, D.A. Hellgartner, K. Loo, S. Lorenz, D.J. Meyhöfer, L. Oberauer, H. Rebber, W.H. Trzaska, and M. Wurm. Topological track reconstruction in unsegmented, large-volume liquid scintillator detectors. *Journal of Instrumentation*, 13(07):P07005–P07005, jul 2018.

[79] Henning Rebber. *Event Discrimination with Topological 3D Reconstruction at MeV Energies in the JUNO Experiment.* PhD thesis, University of Hamburg, 11 2019.

[80] H. Yang, D. Cao, Z. Qian, X. Zhu, C. Loh, A. Huang, R. Zhang, Y. Yang, Y. Liu, B. Xu, and M. Qi. Light attenuation length of high quality linear alkyl benzene as liquid scintillator solvent for the JUNO experiment. *Journal of Instrumentation*, 12(11):T11004–T11004, nov 2017.

[81]

[82] M. Wurm, F. von Feilitzsch, M. Göger-Neff, M. Hofmann, T. Lachenmaier, T. Lewke, T. Marrodán Undagoitia, Q. Meindl, R. Möllenberg, L. Oberauer, and et al. Optical scattering lengths in large liquid-scintillator neutrino detectors. *Review of Scientific Instruments*, 81(5):053301, May 2010.

[83] Patrick Huber. Determination of antineutrino spectra from nuclear reactors. *Physical Review C*, 84(2), Aug 2011.

[84] Th. A. Mueller, D. Lhuillier, M. Fallot, A. Letourneau, S. Cormon, M. Fechner, L. Giot, T. Lasserre, J. Martino, G. Mention, and et al. Improved predictions of reactor antineutrino spectra. *Physical Review C*, 83(5), May 2011.

[85] A Oralbaev, M Skorokhvatov, and O Titov. The inverse beta decay: a study of cross section. *Journal of Physics: Conference Series*, 675(1):012003, feb 2016.

[86] Alessandro Strumia and Francesco Vissani. Precise quasielastic neutrino/nucleon cross-section. *Physics Letters B*, 564(1-2):42–54, Jul 2003.

[87] X. Qian, A. Tan, W. Wang, J. J. Ling, R. D. McKeown, and C. Zhang. Statistical evaluation of experimental determinations of neutrino mass hierarchy. *Physical Review D*, 86(11), Dec 2012.

[88] M. Chen, F. Elisei, F. Masetti, U. Mazzucato, C. Salvo, and G. Testera. Quenching of undesired fluorescence in a liquid scintillator particle detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 420(1):189 – 201, 1999.

[89] C. Arpesella, G. Bellini, J. Benziger, S. Bonetti, B. Caccianiga, F. Calaprice, F. Dalnoki-Veress, D. D'Angelo, H. de Kerret, A. Derbin, and et al. First real time detection of 7be solar neutrinos by borexino. *Physics Letters B*, 658(4):101–108, Jan 2008.

[90] G. Bellini, J. Benziger, S. Bonetti, M. Buizza Avanzini, B. Caccianiga, L. Cadonati, F. Calaprice, C. Carraro, A. Chavarria, F. Dalnoki-Veress, and et al. Observation of geo-neutrinos. *Physics Letters B*, 687(4-5):299–304, Apr 2010.

[91] G. Bellini, J. Benziger, D. Bick, S. Bonetti, G. Bonfini, D. Bravo, M. Buizza Avanzini, B. Caccianiga, L. Cadonati, F. Calaprice, and et al. First evidence ofpepsolar neutrinos by direct detection in borexino. *Physical Review Letters*, 108(5), Feb 2012.

[92] G. Bellini et al. Neutrinos from the primary proton–proton fusion process in the Sun. *Nature*, 512(7515):383–386, 2014.

[93] M. Agostini and et al. Sensitivity to neutrinos from the solar cno cycle in borexino, 2020.

Eidesstattliche Erklärung

Ich, Yu Xu

erkläre hiermit, dass diese Dissertation und die darin dargelegten Inhalte die eigenen sind und selbstständig, als Ergebnis der eigenen originären Forschung, generiert wurden.

Hiermit erkläre ich an Eides statt

1. Diese Arbeit wurde vollständig oder größtenteils in der Phase als Doktorand dieser Fakultät und Universität angefertigt;
2. Sofern irgendein Bestandteil dieser Dissertation zuvor für einen akademischen Abschluss oder eine andere Qualifikation an dieser oder einer anderen Institution verwendet wurde, wurde dies klar angezeigt;
3. Wenn immer andere eigene- oder Veröffentlichungen Dritter herangezogen wurden, wurden diese klar benannt;
4. Wenn aus anderen eigenen- oder Veröffentlichungen Dritter zitiert wurde, wurde stets die Quelle hierfür angegeben. Diese Dissertation ist vollständig meine eigene Arbeit, mit der Ausnahme solcher Zitate;
5. Alle wesentlichen Quellen von Unterstützung wurden benannt;
6. Wenn immer ein Teil dieser Dissertation auf der Zusammenarbeit mit anderen basiert, wurde von mir klar gekennzeichnet, was von anderen und was von mir selbst erarbeitet wurde;
7. Ein Teil oder Teile dieser Arbeit wurden zuvor veröffentlicht und zwar in:
    - Livia Ludhova and Henning Rebber and Björn Soenke Wonsak and Yu Xu "Particle Identification at MeV Energies in JUNO", submit to arxiv:2007.02687 .

date:_____ Signature:_____

Acknowledgements

While the thesis is done I realize that few more lines are needed to express my great thanks to those that, in many different ways, helped me to reach this end.

My first thanks should go to Prof. Dr. Livia Ludhova for giving me the opportunity to write work. I'm stupid on many things and I would specially thanks for your patience and tolerance. I would also thank to Prof. Dr. Achim Stahl and Prof. Dr. Christopher Wiebusch. Thank you for your discussion on my work.

A sincere thankfulness I have to express to the whole IKP neutrino group! We stay together not only for work but also for life. Discussion and work with you help me a lot. Specially thank to Philipp Kampmann and Ömer Penek for your help on my document work. Thank to Yuhang Guo and Runxuan Liu for your meal.

I would like also thank to JUNO collaboration. Thanks Zuyuan Yu for your kindly advice to my work and life. Thanks Benda Xu for discussion and cooperation on waveform reconstruction and thanks for your GPU cluster; Thanks for Dr. Björn Wonsak and Henning Rebber for our good cooperation in the context of particle identification. And thanks for all collaboration members, I'm happy to be with you for four years!

Last I would thank to my family. Thanks for all your support.