Doctoral Dissertation

博士論文

Measurement of neutrons produced in atmospheric neutrino interactions at Super-Kamiokande

(スーパーカミオカンデに於ける 大気ニュートリノ反応に伴う中性子の測定)

A Dissertation Submitted for the Degree of Doctor of Philosophy
December 2023

令和5年12月博士(理学)申請

Department of Physics, Graduate School of Science,
The University of Tokyo

東京大学大学院理学系研究科物理学専攻

Seungho Han

韓 昇澔

Abstract

Reconstructing recoil neutrons from neutrino-nucleus interactions poses a challenge, yet it holds significant merit due to its relevance to neutrino kinematics, nuclear effects, and potential for background rejection. However, adding to the difficulty in detection is the substantial uncertainty in neutron production. Several measurements suggest that the current default simulation models not only overpredict overall neutron production but also display considerable variability in predictions.

In this study, we measured the average neutron multiplicity and track length in atmospheric neutrino interactions at Super-Kamiokande as a function of visible energy. Using a neural network, we identified delayed neutron capture signals in a dataset spanning 4,270 days of livetime, including 561 days in a 0.011w% Gd-loaded phase featuring improved neutron tagging efficiency. The neutron tagging efficiency and its uncertainty were calibrated using an Am/Be neutron source, extensively studied in this thesis. Various hadron intranuclear cascade (INC) models were tested and compared with observations for the first time.

The observed average neutron multiplicity exhibited a clear linear increase with visible energy. While the results showed a deficit compared to the predictions of INC models widely used in the field — consistent with earlier studies — `INCL`, a model offering a more sophisticated description of the nucleus, demonstrated better agreement. This was particularly evident in sub-GeV neutrino events, where the deficit was more pronounced. This underscores the significance of accurately modeling nuclear effects for predicting nucleon production, especially at lower projectile energies where semi-classical assumptions of INC models are less applicable.

Acknowledgements

First and foremost, I want to express my sincere gratitude to the Super-Kamiokande collaborators. Their unwavering dedication over more than two decades, working day and night to maintain the detector and gather invaluable data, is the cornerstone of my work. Without their efforts, none of my achievements would have been possible. I stand on the shoulders of these giants, and the significant discoveries resulting from their consistent dedication paved the way for my contributions.

I would like to express my heartfelt gratitude to Okumura-san for introducing me to the realm of atmospheric neutrinos. His guidance served as a guiding light throughout my demanding five-year journey in graduate studies, steering through the challenges of a pandemic, and balancing a hectic personal life that often felt like navigating in the dark towards a mountain top. I am truly thankful for his patience and mentorship, which have been invaluable to me.

I extend my sincere gratitude to Hayato-san, whose wholehearted support has touched every aspect of my life, from personal matters to my career and research.

This work owes a great deal to the works of past pioneers, among them Akutsu-san, Ka Ming, Pablo, Okajima-san, Tristan, and Haibing.

I want to express special thanks to Yano-san for automating the deployment of calibration sources and overseeing the collection of calibration data. I learned a great deal and heavily relied on Yano-san's simulation work and Gd-neutron capture models, which were instrumental in my neutron calibration studies.

I want to express my gratitude to the Okayama group, with special thanks to Koshio-san, Hagiwara-san, Harada-san, Sakai-san, and Hino-san. The detector simulation developed by Koshio-san and Hagiwara-san serves as the foundation of my neutron calibration studies. The countless discussions with Harada-san on improving Am/Be simulation and `NTag` was something that kept me moving forward. Sakai-san's work on various INC models was a critical input to my model comparison studies. I am also grateful to Hino-san for resolving the Gd-capture ratio issue.

I would like to thank colleagues in the oscillation group, namely Roger, Thomas, and Miki-san. My work on atmospheric neutrino oscillation sensitivity studies with neutron tagging, as well as cosmic ray calibration, heavily relied on their work.

I would like to express my gratitude to colleagues in ATMPD, including Linyan, Noguchi-san, Ishitsuka-san, and a special acknowledgment to Ed. His questions played a crucial role in steering me in the right direction, pushing me forward when I felt stuck. Additionally, I want to thank Miura-san, Mine-san, and Kameda-san for their assistance in the simulation production and data reduction processes.

Appreciation is also due to colleagues in the LOWE group: Sekiya-san for overseeing Gd-loading, Izumiyama-san and Kanemura-san for improving `NTag`, Michael for offering insights into neutron vertex reconstruction, and H. Ito-san for measuring the Am/Be source intensity, critical in enhancing our calibration.

I would like to thank Tairafune-san and Nakajima-san for allowing me to contribute to the neutron tagging analysis in T2K.

Special thanks to those based in Kashiwa: Patrick, Junjie, Guille, Xubin, and Rhea, are among the few friends I made in Kashiwa, and they made my life enjoyable. Abe-san provided crucial advice that shaped my research, and Tashiro-san's assistance with simulation production and data reduction was invaluable. Mashima-san supported me through every administrative process for five years. My apologies to Tomiya-san and Yoshida-san for not being the best senpai.

I acknowledge the financial support from the MEXT scholarship, which sustained my five-year graduate studies in Japan.

I dedicate this thesis to my parents and family,
who have supported me throughout the entire journey.

Contents

Chapter 1

Introduction

Neutrinos were first theorized [1] in 1933 as part of the "weak" interaction theory to explain the prolonged lifetime of beta decay compared to electromagnetic interaction. The properties of neutrinos are defined by weak interaction characteristics and conservation laws. Neutrinos must be electrically neutral, possess an extremely light mass, and have a spin of 1/2. They exhibit an incredibly low interaction cross-section with matter when compared to other known particles, making them one of the most abundant elementary particles. However, despite their abundance, neutrinos remain mysterious, featuring various puzzling and unique aspects that set them apart from other particles.

One notable aspect is that, so far, only left-handed neutrinos and right-handed anti-neutrinos have been observed [2]. For decades, physicists had ample reasons to believe that neutrinos were massless, given the chiral nature [3] of weak interaction, an extremely small neutrino mass limit, and the apparent lack of right-handed neutrinos. The electroweak theory [4, 5, 6], which unifies weak interaction with electromagnetic interaction and predicts mediator boson masses, was formulated based on this assumption.

The surprising discovery of neutrino oscillation [7] suggests neutrinos have mass and hints at the possible existence of right-handed neutrinos, which do not interact weakly. While this finding sheds light on longstanding mysteries such as observed neutrino deficits [8], it also raises new questions. What is the origin of their small mass, and what are the properties of right-handed neutrinos and their relationship to left-handed neutrinos? Precision measurements of neutrino oscillation parameters, including mixing angles, differences in neutrino masses, and the correct "order" of the neutrino masses can provide valuable insights. Additionally, investigating whether the oscillation violates CP symmetry, as observed in quarks [9, 10], may offer an extra source of CP symmetry to explain the observed baryon-antibaryon asymmetry [11].

In this chapter, we briefly discuss the theory of standard three-flavor neutrino oscillation, outlining the remaining challenges and the current status of oscillation parameter measurements, as well as the obstacles faced by experiments. We then focus on how the detection of "neutrons" from atmospheric neutrino interactions in Super-Kamiokande can not only enhance future neutrino oscillation parameter measurements but also contribute to other rare physics searches where atmospheric neutrinos serve as backgrounds. Finally, we explore the motivation behind measuring the properties of neutrons produced in atmospheric neutrino interactions in water and conclude with an overview of this study.

1.1 Neutrino oscillation

The three known weak flavor eigenstates of neutrinos [12, 13, 14] form a complete and orthonormal basis for the three lepton generations in the Standard Model. With nonzero neutrino mass, we can establish another complete and orthonormal basis with three mass eigenstates, which may not necessarily align with the flavor basis. Let the flavor eigenstate be a coherent superposition of the mass eigenstates, linked by a 3×3 unitary matrix U:

$$|\nu_l\rangle = \sum_k U_{lk}^* |\nu_k\rangle \tag{1.1}$$

where $|\nu_l\rangle$ represents the flavor eigenstate of flavor $l \in \{e, \mu, \tau\}$, and $|\nu_k\rangle$ is the k^{th} mass eigenstate where $k \in \{1, 2, 3\}$.

Let $|\nu(x)\rangle$ denote a neutrino state produced via weak interaction at time $x^0 = 0$. Each mass eigenstate should acquire a distinct phase as it propagates in (vacuum) spacetime, so that:

$$\begin{aligned} |\nu(x)\rangle &= \sum_k U_{lk}^* |\nu_k\rangle \, e^{-i\phi_k(x)} \\ &= \sum_{l',k} U_{lk}^* U_{l'k} |\nu_{l'}\rangle \, e^{-i\phi_k(x)} \end{aligned} \tag{1.2}$$

where $\phi_k(x) \coloneqq p_k \cdot x$ represents the phase gain of $|\nu_k\rangle$ after propagation x.

Naturally, the probability that the state $|\nu\rangle$ which started with flavor l will be detected with another flavor l' is given by:

$$\begin{aligned} P(\nu_l \to \nu_{l'})(x) &= \left| \sum_k U_{lk}^* U_{l'k} e^{-i\phi_k(x)} \right|^2 \\ &= \delta_{ll'} + 2\text{Re} \sum_{k' \neq k} \left[U_{lk}^* U_{l'k} U_{lk'} U_{l'k'}^* e^{-i\Delta\phi_{kk'}(x)} \right] \end{aligned} \tag{1.3}$$

where $\Delta\phi_{kk'}(x) \equiv \phi_k(x) - \phi_{k'}(x)$.

If the neutrino masses are non-degenerate so that the mass eigenstates may fall out of phase, the probability in Equation 1.3 becomes non-zero for $l \neq l'$[1]. This flavor transition, permitted by the non-degeneracy of neutrino masses, is referred to as "oscillation," owing to the sinusoidal nature of the transition probability in Equation 1.3. Experimental evidence of neutrino oscillation has been established through experiments involving various neutrino sources: atmospheric [7], solar [8], reactor [16], and artificial [17].

In a scenario where the Dirac masses are considered, the unitary 3×3 neutrino mixing matrix U can be parameterized with four real free parameters. This considers the neutrino and lepton spinors of three generations absorbing the relative phases and eliminating $\binom{3}{2} - 1$ degrees of freedom from U, so that only the remaining four degrees of freedom can be considered physically meaningful. The standard approach to parameterizing the mixing matrix U is:

$$\begin{bmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{23} & \sin\theta_{23} \\ 0 & -\sin\theta_{23} & \cos\theta_{23} \end{bmatrix} \begin{bmatrix} \cos\theta_{13} & 0 & \sin\theta_{13}e^{-i\delta_{\mathrm{CP}}} \\ 0 & 1 & 0 \\ -\sin\theta_{13}e^{i\delta_{\mathrm{CP}}} & 0 & \cos\theta_{13} \end{bmatrix} \begin{bmatrix} \cos\theta_{12} & \sin\theta_{12} & 0 \\ -\sin\theta_{12} & \cos\theta_{12} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{bmatrix} \tag{1.4}$$

Here, θ_{ij}, where $i, j \in \{1, 2, 3\}$, denotes the mixing angles representing the internal mixing of mass eigenstates, and δ_{CP} signifies the overall complex phase of U. It is worth noting that the complex phase is not allowed for dimension smaller than 3. $\delta_{CP} \neq 0$ or π implies $P(\nu_l \to \nu_{l'}) \neq P(\bar{\nu}_l \to \bar{\nu}_{l'})$ and thus violation of CP symmetry.

[1] One might speculate about the oscillation of charged leptons produced through weak decay of a heavy particle. However, due to the substantially larger masses of charged leptons compared to neutrinos, verifying this experimentally with $O(1)$ m path length would necessitate an energy scale on the order of $O(10^{13})$ GeV. Refer to [15] for further details.

In the ultrarelativistic limit, the expression $\phi_k(x)$ in Equation 1.3 simplifies to $\phi_k(x) \approx m_k^2 L/2E$, where m_k represents the mass of the k^{th} mass eigenstate $|\nu_k\rangle$, $L = |\vec{x}|^2$ denotes the distance traveled by the neutrino, and $E = p_k^0 = E_\nu$ is assumed to be the energy of the neutrino, consistent across all k. Consequently, the oscillation probability $P(\nu_l \to \nu_{l'})$ is fully characterized by the traveled distance L, the neutrino energy E, and the so-called neutrino oscillation parameters: the three mixing angles θ_{ij}, the (Dirac) CP phase δ_{CP}, and the two independent mass-squared differences Δm^2_{21} and Δm^2_{32}, where $\Delta m^2_{ij} \equiv m_i^2 - m_j^2$.

It is important to note that, at the leading order, the (vacuum) oscillation probability as shown in Equation 1.3 is not sensitive to the relative ordering of neutrino masses m_i, as it depends on the mass-squared differences Δm^2_{ij}. Currently, two measured values of mass-squared differences exist, $\Delta m^2_{21} = 7.41^{+0.21}_{-0.20} \times 10^{-5}$ eV^2 and $\Delta m^2_{3\{1,2\}} = 2.507^{+0.026}_{-0.027} \times 10^{-3}$ eV^2. This information tells us that m_1 and m_2 are much closer, and m_3 is either much larger or smaller than the other two. The unknown direction of this mass difference is referred to as the mass ordering/hierarchy problem. In the normal mass ordering (NO) scenario, m_3 is larger than m_1 and m_2, while in the inverted mass ordering (IO) scenario, m_3 is smaller. This distinction has implications for the associated neutrino flavors with different mass ranges, as illustrated in Figure 1.1. Neutrino mass ordering is also believed to have implications for the origin of neutrino mass as well as the Dirac/Majorana nature of neutrinos.

Most of the oscillation parameters, except δ_{CP}, are relatively precisely measured as shown in Table 1.1. Several inferences can be drawn from Table 1.1: firstly, normal neutrino ordering is slightly favored by 2.5σ if atmospheric neutrino data is considered.[2] Secondly, whether θ_{23} is in the 0-45° or 45-90° is not yet known. This uncertainty arises because θ_{23} appears in the oscillation probability in the form of $\sin 2\theta_{23}$, leading to a degeneracy between the two angle regions. Finally, the current best fit δ_{CP} is close to 1.1π with normal mass ordering, allowing for CP conservation ($\delta_{\text{CP}} = \pi$) at 1-2σ confidence level. Figure 1.2 visually demonstrates the relatively larger mixing in neutrinos compared to quarks.

[2]The atmospheric neutrino's sensitivity to mass ordering is discussed in Section 1.2.

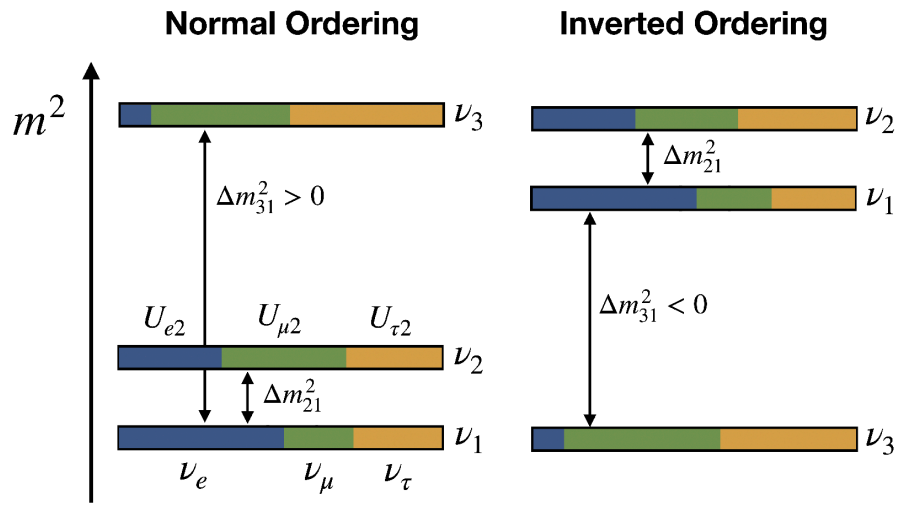


Figure 1.1: Illustration of neutrino mass ordering. (Reprinted from [18])

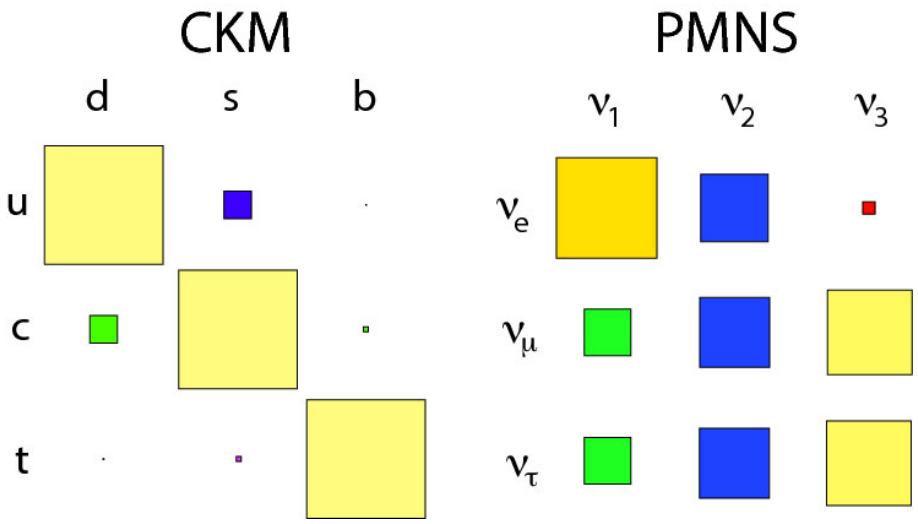


Figure 1.2: Visual comparison of quark (CKM, left) and neutrino (PMNS, right) mixing matrices. (Credit: S. Stone)

NuFIT 5.2 (2022)

		Normal Ordering (best fit)		Inverted Ordering ($\Delta\chi^2 = 2.3$)	
		bfp $\pm 1\sigma$	3σ range	bfp $\pm 1\sigma$	3σ range
without SK atmospheric data	$\sin^2\theta_{12}$	$0.303^{+0.012}_{-0.011}$	$0.270 \to 0.341$	$0.303^{+0.012}_{-0.011}$	$0.270 \to 0.341$
	$\theta_{12}/^\circ$	$33.41^{+0.75}_{-0.72}$	$31.31 \to 35.74$	$33.41^{+0.75}_{-0.72}$	$31.31 \to 35.74$
	$\sin^2\theta_{23}$	$0.572^{+0.018}_{-0.023}$	$0.406 \to 0.620$	$0.578^{+0.016}_{-0.021}$	$0.412 \to 0.623$
	$\theta_{23}/^\circ$	$49.1^{+1.0}_{-1.3}$	$39.6 \to 51.9$	$49.5^{+0.9}_{-1.2}$	$39.9 \to 52.1$
	$\sin^2\theta_{13}$	$0.02203^{+0.00056}_{-0.00059}$	$0.02029 \to 0.02391$	$0.02219^{+0.00060}_{-0.00057}$	$0.02047 \to 0.02396$
	$\theta_{13}/^\circ$	$8.54^{+0.11}_{-0.12}$	$8.19 \to 8.89$	$8.57^{+0.12}_{-0.11}$	$8.23 \to 8.90$
	$\delta_{\rm CP}/^\circ$	197^{+42}_{-25}	$108 \to 404$	286^{+27}_{-32}	$192 \to 360$
	$\frac{\Delta m^2_{21}}{10^{-5}\ \text{eV}^2}$	$7.41^{+0.21}_{-0.20}$	$6.82 \to 8.03$	$7.41^{+0.21}_{-0.20}$	$6.82 \to 8.03$
	$\frac{\Delta m^2_{3\ell}}{10^{-3}\ \text{eV}^2}$	$+2.511^{+0.028}_{-0.027}$	$+2.428 \to +2.597$	$-2.498^{+0.032}_{-0.025}$	$-2.581 \to -2.408$
		Normal Ordering (best fit)		Inverted Ordering ($\Delta\chi^2 = 6.4$)	
		bfp $\pm 1\sigma$	3σ range	bfp $\pm 1\sigma$	3σ range
with SK atmospheric data	$\sin^2\theta_{12}$	$0.303^{+0.012}_{-0.012}$	$0.270 \to 0.341$	$0.303^{+0.012}_{-0.011}$	$0.270 \to 0.341$
	$\theta_{12}/^\circ$	$33.41^{+0.75}_{-0.72}$	$31.31 \to 35.74$	$33.41^{+0.75}_{-0.72}$	$31.31 \to 35.74$
	$\sin^2\theta_{23}$	$0.451^{+0.019}_{-0.016}$	$0.408 \to 0.603$	$0.569^{+0.016}_{-0.021}$	$0.412 \to 0.613$
	$\theta_{23}/^\circ$	$42.2^{+1.1}_{-0.9}$	$39.7 \to 51.0$	$49.0^{+1.0}_{-1.2}$	$39.9 \to 51.5$
	$\sin^2\theta_{13}$	$0.02225^{+0.00056}_{-0.00059}$	$0.02052 \to 0.02398$	$0.02223^{+0.00058}_{-0.00058}$	$0.02048 \to 0.02416$
	$\theta_{13}/^\circ$	$8.58^{+0.11}_{-0.11}$	$8.23 \to 8.91$	$8.57^{+0.11}_{-0.11}$	$8.23 \to 8.94$
	$\delta_{\rm CP}/^\circ$	232^{+36}_{-26}	$144 \to 350$	276^{+22}_{-29}	$194 \to 344$
	$\frac{\Delta m^2_{21}}{10^{-5}\ \text{eV}^2}$	$7.41^{+0.21}_{-0.20}$	$6.82 \to 8.03$	$7.41^{+0.21}_{-0.20}$	$6.82 \to 8.03$
	$\frac{\Delta m^2_{3\ell}}{10^{-3}\ \text{eV}^2}$	$+2.507^{+0.026}_{-0.027}$	$+2.427 \to +2.590$	$-2.486^{+0.025}_{-0.028}$	$-2.570 \to -2.406$

Table 1.1: Three-flavor oscillation parameters obtained from different global analyses of neutrino data. (Reprinted from [19])

1.2 Parameter measurements and challenges

In Section 1.1, we discussed how, for a given neutrino (antineutrino) flavor, the three-flavor neutrino oscillation probability can be expressed as a function of oscillation parameters and the ratio between the neutrino energy E_ν and the path length L:

$$P(\nu_\alpha \to \nu_\beta) \approx f(\theta_{ij}, \Delta m^2_{ij}, \delta_{CP}; E_\nu/L) \tag{1.5}$$

When the neutrino (or antineutrino) α, β, E_ν, L, and the flux from the source are known, we can deduce the oscillation parameters from the observed event rates of each neutrino flavor. Different neutrino sources, spanning from MeV to GeV and including atmospheric, solar, reactor, and accelerator neutrinos, contribute to the measurements of oscillation parameters, as summarized in Table 1.1.

The dominant interaction channel in oscillation experiments is the neutrino-nucleus charged-current quasi-elastic (CCQE) interaction, which produces a single charged lepton (carrying the neutrino flavor) and scatters a single nucleon from the target nucleus, as in $\nu_l n \to l^- p$ or $\bar{\nu} l p \to l^+ n$. This channel is "clean," because the neutrino energy E_ν can be mostly inferred by reconstructing lepton kinematics only. In the ideal 2-body CCQE scenarios, such as $\nu_l + {}^{Z}_{A}\mathrm{X} \to l^- + {}^{Z-1}_{A}\mathrm{X}^* + p$ and $\bar{\nu_l} + {}^{Z}_{A}\mathrm{X} \to l^+ + {}^{Z-1}_{A-1}\mathrm{Y}^* + n$, where one outgoing lepton l and one nucleon n/p are involved, the energy of the incoming neutrino E_ν is determined as:

$$E_\nu \approx \frac{(m_N)^2 - (m_N^{\mathrm{eff}})^2 - m_l^2 + 2(m_N^{\mathrm{eff}})E_l}{2(m_N^{\mathrm{eff}} - E_l + p_l \cos\theta_l)} \tag{1.6}$$

Here, m_N is the mass of the incoming nucleon, $m_N^{\mathrm{eff}} \approx m_N - E_b$ is the effective mass of the outgoing nucleon, accounting for the nuclear binding energy E_b, and m_l, E_l, p_l, and θ_l represent the mass, energy, momentum, and scattering angle of the outgoing lepton, respectively.

In the context of δ_{CP}, the comparison between $P(\nu_\mu \to \nu_e)$ and $P(\bar{\nu}_\mu \to \bar{\nu}_e)$ with accelerator neutrinos is considered to be the most sensitive. This sensitivity arises from the well-characterized neutrino type, flavor, path lengths, and the reasonably accurate energy reconstruction performance of Equation 1.6, thanks to the well-known neutrino direction.

For determining the mass ordering, the most sensitive approach involves comparing the rates of ν_e and $\bar{\nu}_e$ in atmospheric neutrinos. This sensitivity is attributed to the Mikheyev-Smirnov-Wolfenstein (MSW) effect, also known as the "matter effect" [20, 21]. In contrast to the vacuum, ordinary matter, filled with bound e^-, introduces a matter effect through interactions such as $\nu_e e^- \to e^- \nu_e$ mediated by the charged current W^+. This alters the effective Hamiltonian experienced by propagating neutrinos, with distinct changes for neutrinos and antineutrinos based on different mass ordering scenarios, as illustrated in Figure 1.6.

Several challenges limit the precision of these measurements:

Lack of statistics

The smallness of neutrino interaction cross sections, on the order of 10^{-14} barns, makes it difficult to accumulate event statistics. Long-baseline accelerator neutrino experiments targeting δ_{CP} face this challenge, with oscillation path lengths typically reaching $O(100)$ km. For example, over the last decade, the T2K experiment has collected around 100 ν_e and $\bar{\nu}_e$ events at the far detector (Super-Kamiokande), which is $O(10^{-2})$ smaller than the events collected with atmospheric neutrinos. The small cross-section necessitates large, dense detectors with target materials like water, plastic, oil, or metal, often compromising event reconstruction capabilities.

Hadron reconstruction

Large, non-magnetized detectors face challenges in distinguishing between neutrino and antineutrino interactions, as the sign of the scattered lepton charge is lost. Additionally, these detectors struggle to accurately reconstruct recoil nucleon kinematics, posing a significant challenge to the neutrino reconstruction, especially for atmospheric neutrinos that are not well-characterized.

Nuclear effects

The use of composite nucleus targets introduces complications. Experiments relying on the "cleanliness" of CCQE events define their samples based on specific observed final states, e.g., one charged lepton and one proton, as expected by $\nu_l n \to l^- p$. However, events like $\bar{\nu}_l p \to l^- p \pi^-$ with π^- absorbed by the target nucleus cannot be distinguished from CCQE events. Accurate reconstruction of such events is crucial, demanding a comprehensive understanding of nuclear effects to estimate the fraction of non-identifiable backgrounds.

Degeneracy between δ_{CP} and mass ordering

In the first order, $P(\nu_\mu \to \nu_e)$ with the matter effect can be expressed [22]:

$$\begin{aligned} P(\nu_\mu \to \nu_e) \approx & \sin^2\theta_{23} \frac{\sin^2 2\theta_{12}}{(A-1)^2} \sin^2[(A-1)\Delta_{31}] \\ & \mp \alpha \frac{J_0 \sin\delta_{\mathrm{CP}}}{A(1-A)} \sin\Delta_{31} \sin(A\Delta_{31}) \sin[(1-A)\Delta_{31}] \\ & + \alpha \frac{J_0 \cos\delta_{\mathrm{CP}}}{A(1-A)} \cos\Delta_{31} \sin(A\Delta_{31}) \sin[(1-A)\Delta_{31}] \\ & + \alpha^2 \cos^2\theta_{23} \frac{\sin^2 2\theta_{12}}{A^2} \sin^2(A\Delta_{31}) \end{aligned} \tag{1.7}$$

assuming natural units, where:

$$\begin{aligned} \alpha &= \Delta m^2_{21}/\Delta m^2_{31} \\ \Delta_{ij} &= \Delta m^2_{ij}/4E_\nu \\ A &= \pm 2\sqrt{2} G_F n_e E_\nu / \Delta m^2_{31} \\ J_0 &= \sin 2\theta_{12} \sin 2\theta_{13} \sin 2\theta_{23} \cos\theta_{13} \end{aligned}$$

The opposite signs are for antineutrinos. The second term in the equation allows for CP violation for $\delta_{\mathrm{CP}} \neq 0$ or π, enhancing/suppressing $\nu_e/\bar{\nu}_e$ appearance for $\sin\delta_{\mathrm{CP}} > 0$. The third term changes sign for different mass orderings, enhancing/suppressing $\nu_e/\bar{\nu}_e$ appearance for normal mass ordering, potentially mimicking the δ_{CP} effect, as demonstrated in Figure 1.3.

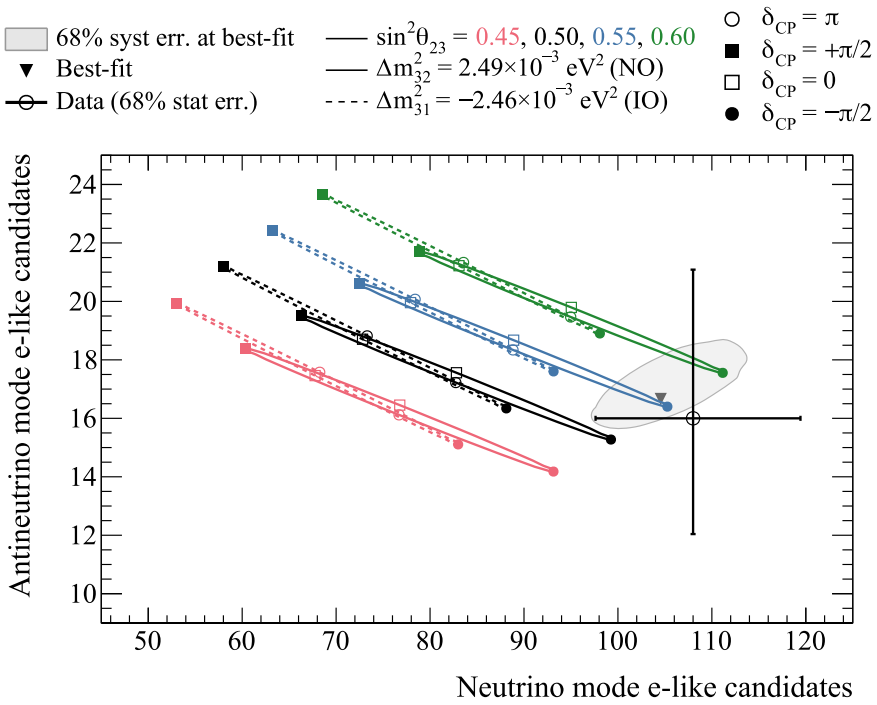


Figure 1.3: Comparison of observed e-like candidates in the T2K far detector (Super-Kamiokande) with ν_μ and $\bar{\nu}_\mu$ beam modes against predictions based on varying oscillation parameters. The assumed normal mass ordering mimics changes in the observed ν_e : $\bar{\nu}_e$ ratios corresponding to $\delta_{\mathrm{CP}} = -\pi/2$, while the inverted ordering mimics changes in the direction of $\delta_{\mathrm{CP}} = \pi/2$. T2K estimates that the δ_{CP}-mimicking effect (the relative size of the third term in Equation 1.7) of different mass ordering scenarios is ∼30% compared to the true δ_{CP} effect. (Reprinted from [23])

One proposed solution involves jointly analyzing data samples from accelerator and atmospheric neutrinos that share the same (far) detector. This approach leverages on the distinct sensitivities of the two data sets, as well as the common detector systematic uncertainties. Figures 1.4 and 1.5 illustrate the expected power to reject CP symmetry and incorrect mass ordering across various assumed true δ_{CP} values. The additional mass ordering sensitivity in atmospheric neutrino data restores the rejection power lost due to the degeneracy between the genuine CP effect and the matter effect demonstrated in Figure 1.3, particularly for the two scenarios: true $\sin\delta_{\mathrm{CP}} > 0$ with normal ordering and true $\sin\delta_{\mathrm{CP}} < 0$ with inverted ordering.

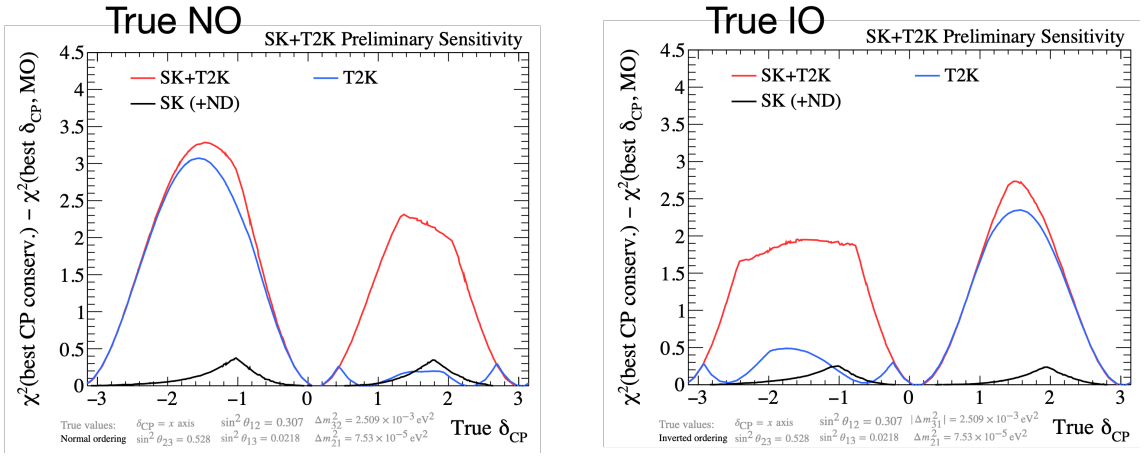


Figure 1.4: The expected power to reject CP symmetry compared across three different analyses: SK atmospheric neutrino data only with T2K near detector constraints (black), T2K accelerator neutrino data only (blue), and a joint fit of the two (red), for true normal (NO, left) and inverted (IO, right) mass ordering scenarios. The vertical axis represents the minimum χ^2 in the CP-symmetric scenario ($\delta_{CP} \in \{0, \pm\pi\}$) subtracted by the global minimum χ^2. (Adapted from [24])

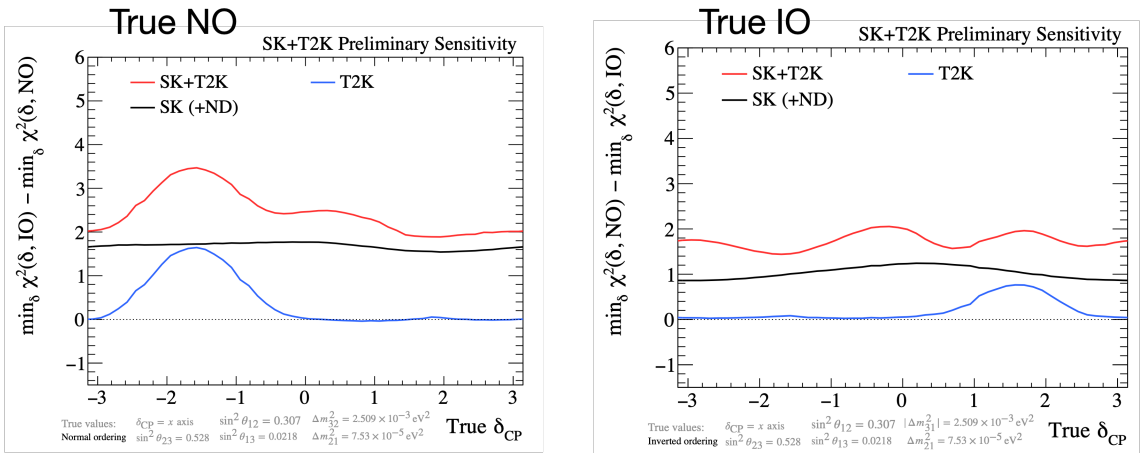


Figure 1.5: The expected power to reject incorrect mass ordering hypothesis. The vertical axis is the minimum χ^2 in the wrong mass ordering scenario subtracted by the minimum χ^2 in the correct mass ordering scenario (Adapted from [24])

The future determination of δ_{CP} and neutrino mass ordering can benefit significantly from the complementary strengths of accelerator neutrinos, known for their well-characterized properties, and atmospheric neutrinos, which offer abundant statistics, diverse path lengths, and a range of energies. While there is considerable potential for improvement on both fronts, the prospects for advancement appear particularly promising in the realm of atmospheric neutrinos, especially concerning the ability to differentiate between neutrinos and antineutrinos, and the reconstruction of neutrino path length and energy (L/E_ν). In the following section, we discuss how the detection of neutrons, especially with the Super-Kamiokande detector, can contribute to addressing these challenges.

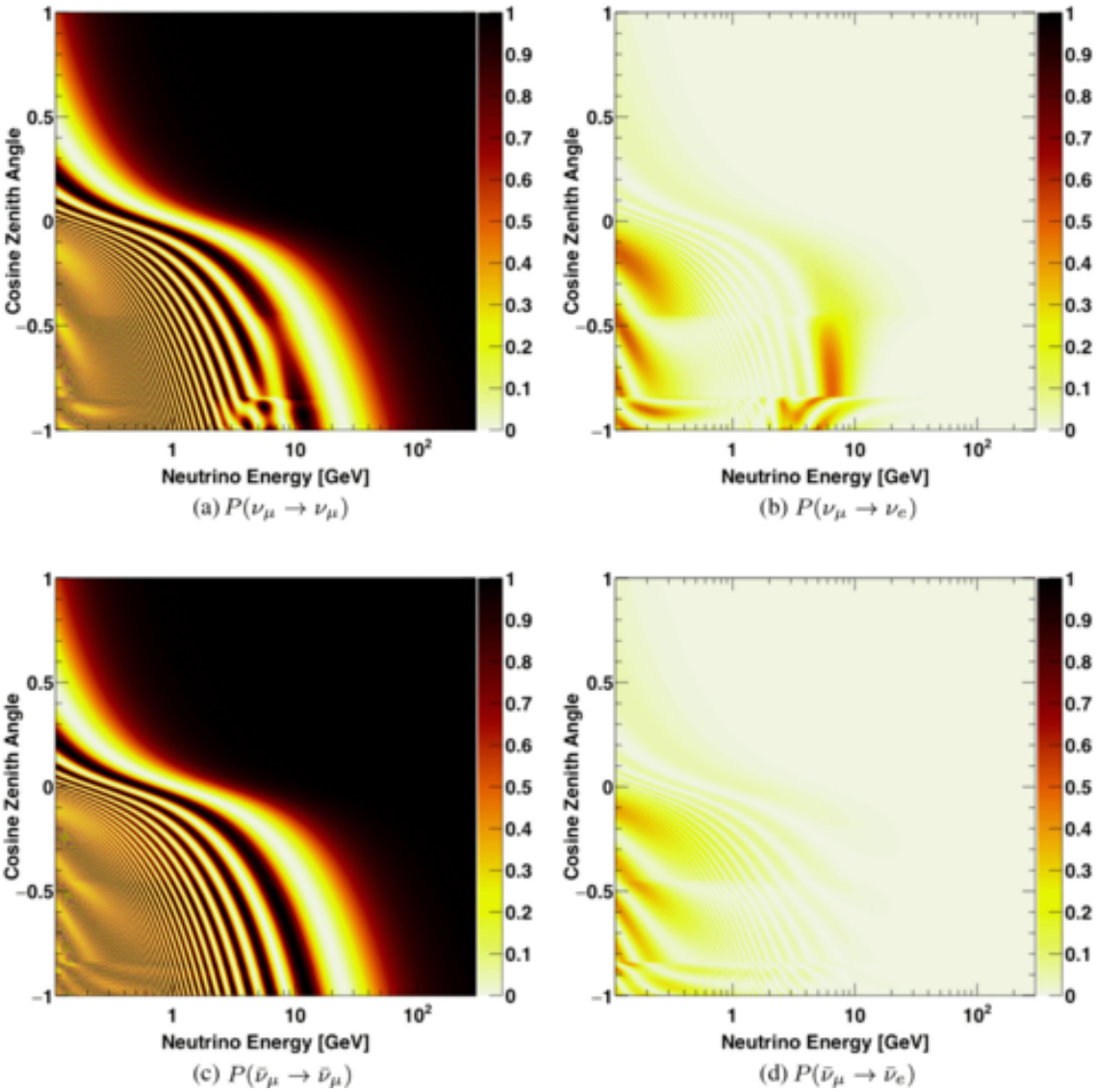


Figure 1.6: Oscillation probabilities for atmospheric ν_μ (top panel) and $\bar{\nu}_\mu$ (bottom panel) at Super-Kamiokande, as functions of true neutrino energy and zenith angle, assuming normal mass ordering. Distortions induced by matter effects in the Earth are observable only in the $O(1)$ GeV range for $\cos\theta_{\text{zenith}} < -0.5$ in the case of ν_μ, not for $\bar{\nu}_\mu$. This angular region corresponds to neutrino propagation through the Earth's core, mantle, and crust. Conversely, under the assumption of inverted mass ordering, matter effects manifest solely in the $\bar{\nu}_\mu$ oscillations, not in those of ν_μ. (Reprinted from [25])

1.3 Advantages of neutron-tagging

In this section, we discuss the advantages of "neutron-tagging" physical events in Super-Kamiokande (SK), with a particular focus on atmospheric neutrino events and oscillation analysis, though not limited solely to these.

Since 2008, SK has actively detected neutron signals in its atmospheric neutrino data. The primary objective has been to identify recoil neutrons from antineutrino CCQE interaction $\bar{\nu}_l p \to l^+ n$, and differentiate them from neutrino CCQE interaction $\nu_l n \to l^- p$. Most recoil neutrons undergo ${}^1\mathrm{H}(n,\gamma)$, resulting in the emission of a single 2.2 MeV gamma-ray. The detection efficiency of this gamma-ray at SK stands at approximately 20%. Figure 1.7 shows the difference in the expected neutron signal multiplicity between true neutrino and antineutrino events. It implies that requiring no neutron in the event allows the effective classification of antineutrinos from neutrinos. It is important to note that the neutron multiplicity in $\nu/\bar{\nu}$ CCQE events is not strictly confined to zero or one. This variability arises due to contamination from non-CCQE events, as well as the influence of nuclear effects and subsequent hadronic interactions within the detector medium. These factors contribute to the smearing of the neutron signal multiplicity, as evident in the tails displayed in Figure 1.7.

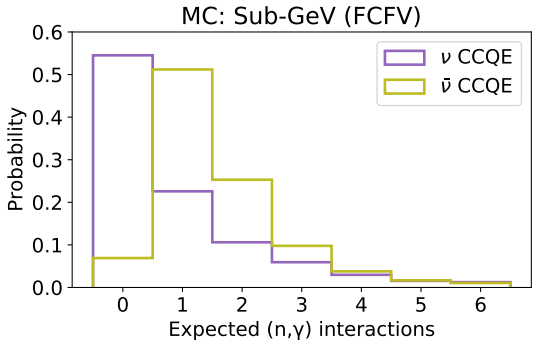


Figure 1.7: The probability density of expected neutron signal multiplicity for sub-GeV atmospheric neutrino events (purple) and antineutrino events (olive) with reconstructed visible energy below 1.33 GeV, and vertices within the fiducial volume (FCFV). Refer to Section 2.3 for a description of the simulation configuration.

By using the neutron-tagging technique for $\nu/\bar{\nu}$ classification, we can potentially leverage the sensitivity of δ_{CP} and mass ordering inherent in the expected $\nu_e/\bar{\nu}_e$ flux ratio at SK. When considering opposite-signed values of δ_{CP} or different mass ordering scenarios, they yield opposite-signed effects on $P(\nu_\mu \to \nu_e)$ and $P(\bar{\nu}_\mu \to \bar{\nu}_e)$, as indicated in Equation 1.7 and demonstrated in Figures 1.3 and 1.6. To magnify this effect, we can examine the ratio between observed ν_e and $\bar{\nu}_e$ events, particularly focusing on "up-going" events where neutrinos are expected to have traversed through the Earth.

Figure 1.8 shows the expected ratios of strictly up-going $\nu_e/\bar{\nu}_e$ fluxes observable at SK, assuming different values of δ_{CP} and mass ordering scenarios. The figure indicates that $\nu_e/\bar{\nu}_e$ flux ratios show sensitivity to δ_{CP} in the $O(0.1)$ GeV range and to mass ordering scenarios in the $O(1)$ GeV range.

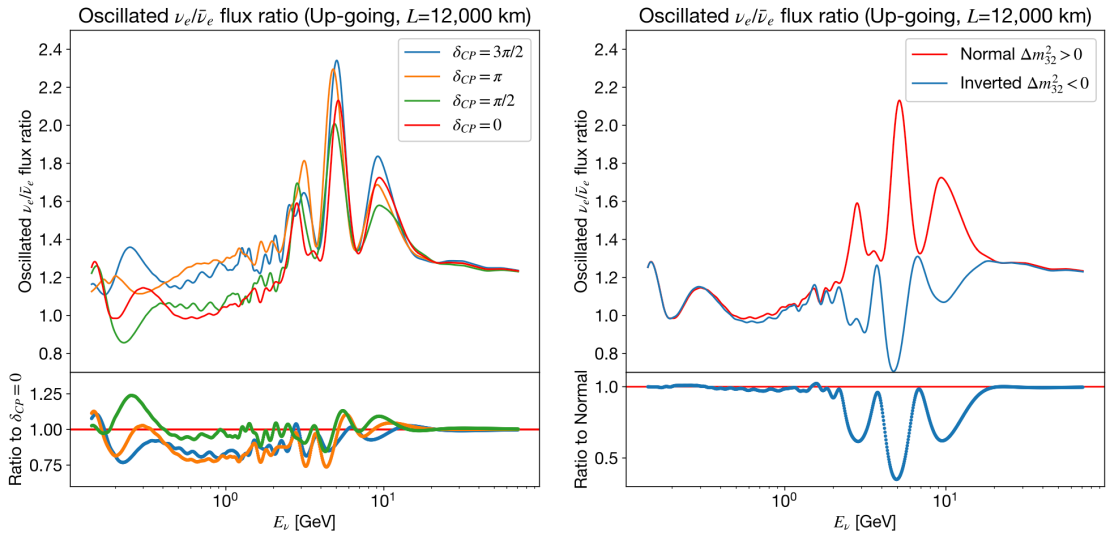


Figure 1.8: The predicted ratios of strictly up-going $\nu_e/\bar{\nu}_e$ fluxes, assuming different values of δ_{CP} (left) and mass ordering scenarios (right). Neutrinos were assumed to have traveled a path length of 12,000 km, with a zenith angle of 180°. In making these predictions, the HKKM 2011 flux model, detailed in Section 2.3.1, and the default oscillation parameters from Table 4.1 were used. To accommodate imperfect energy resolution, fast oscillations were averaged out in the figures.

Figure 1.9 shows the enhancement in the expected sensitivity to δ_{CP} and mass ordering due to the $\nu/\bar{\nu}$ classification through neutron-tagging. Notably, significant improvement in mass ordering sensitivity is expected regardless of the assumed δ_{CP}. Additionally, there is some extra sensitivity gain in δ_{CP}, particularly around the least probable value of δ_{CP}.

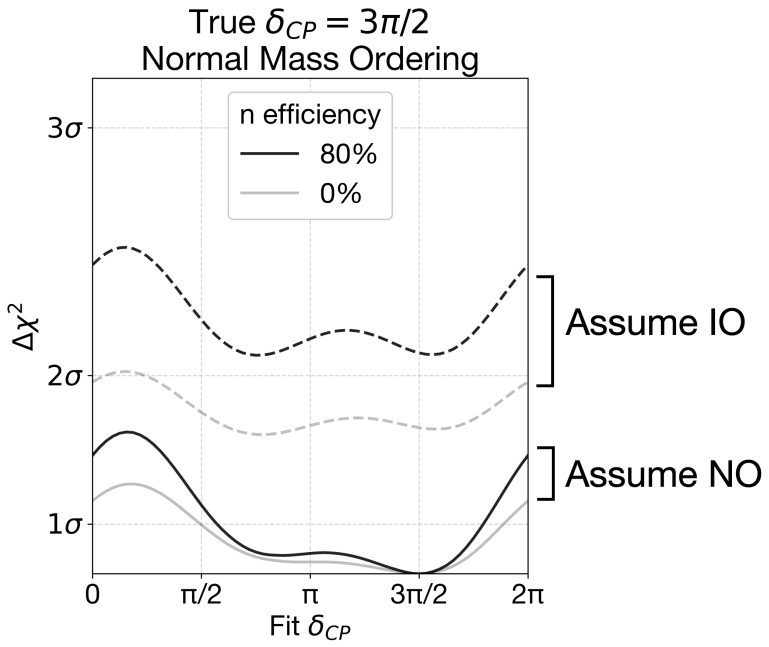


Figure 1.9: The expected Asimov (simulation-only) sensitivity of SK atmospheric neutrino observations to δ_{CP} and neutrino mass ordering, assuming an additional 10 years of operation from 2019, both with (thicker line) and without (lighter line) neutron-tagging. The oscillation parameters were adjusted to best describe fake data generated with a predefined set of ground truth parameters.
The ground truth assumes $\delta_{\mathrm{CP}} = 3\pi/2$ and normal mass ordering. The solid lines represent the fit results assuming the correct mass ordering, while the dashed lines show results under the assumption of the incorrect mass ordering. The analysis incorporates the default systematic uncertainties used in SK atmospheric neutrino oscillation studies [25].

Another advantage of detecting recoil neutrons from neutrino events lies in the potential to enhance the reconstruction of the incident neutrino kinematics. This becomes particularly valuable for atmospheric neutrino events, where the direction (indicative of path length) and energy are initially unknown. Currently, proxies such as the momentum and direction of scattered charged leptons are utilized, but their accuracy in representing incident neutrino kinematics is limited.

To achieve a more accurate reconstruction of neutrino energy, we can leverage the count of neutrons, expected to be proportionate to the energy transferred to the hadron system [26, 27]. Figure 1.10 illustrates the expected linear increase in the fraction of "missing neutrino energy" with respect to the detected neutron signals.

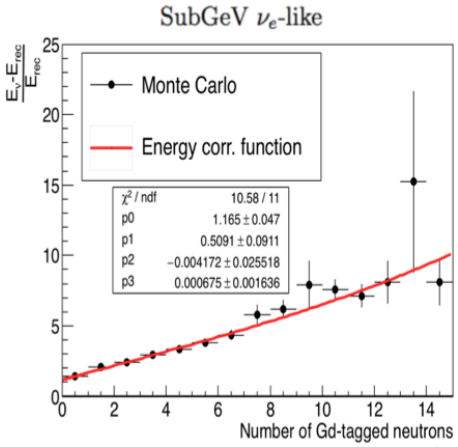


Figure 1.10: The expected fraction of energy reconstruction error $((E_\nu - E_{\rm rec})/E_{\rm rec})$, as a function of detected neutron signals. The simulation assumes 0.01w% Gd in the SK water volume. (Reprinted from [27])

For a more accurate reconstruction of the neutrino direction, we can leverage the fact that the expected neutron track length, from its production to capture, tends to be longer for more energetic neutrons. By reconstructing both the neutrino interaction vertex and the subsequent neutron capture vertices, we can obtain a set of vectors that positively correlate with the initial kinetic energy. Using these vectors, we can deduce the recoil neutron's kinetic energy and, consequently, the original neutrino direction, given the reconstructed lepton kinematics.

Figure 1.11 illustrates enhanced performance in the reconstruction of neutrino energy and direction [27, 28]. Notably, we observe a significant enhancement in neutrino energy reconstruction and a moderate improvement in neutrino direction reconstruction. The complete atmospheric neutrino oscillation analysis, incorporating all the mentioned advancements in $\nu/\bar{\nu}$ classification and energy/direction reconstruction, is currently in progress. Importantly, these techniques extend beyond atmospheric neutrinos and can be applied to accelerator neutrino events as well, offering further improvements in neutrino energy reconstruction and reducing the wrong sign ($\nu/\bar{\nu}$) background contamination.

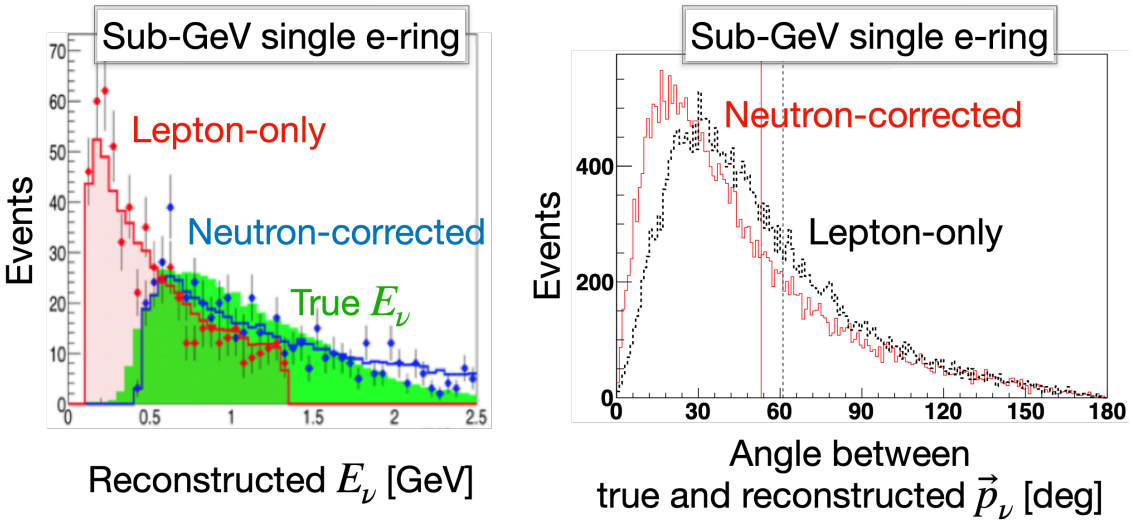


Figure 1.11: A comparison between true and reconstructed neutrino energy (left, [27]) and direction (right, [28]). Both simulations assume 0.01w% Gd in the SK water volume. The Sub-GeV sample consists of events with visible energy (lepton-only, highlighted in red) smaller than 1.33 GeV. (Reprinted from [27, 28])

There are also advantages beyond the scope of atmospheric neutrino oscillation analysis, particularly in analyses where signals are anticipated to exhibit a specific neutron count. For instance, in proton decays like $p \to e^+\pi^0$ or $p \to \bar{\nu}K^+$, as predicted by beyond-Standard-Model theories, no neutron is produced (unless nuclear de-excitation following a bound proton decay emits one). In the energy range of proton decay around ~ 1 GeV, neutrinos are expected to produce an average of 2-4 neutrons.

The diffuse supernova neutrino background (DSNB), expected to be in the range of $O(10)$ MeV, represents the postulated isotropic flux of neutrinos originating from all core-collapse supernova events throughout the universe's history. The most dominant interaction channel for DSNB is through $\bar{\nu}_e p \rightarrow e^+ n$, commonly known as inverse beta decay, which is expected to yield a single neutron.

Several typical backgrounds can be significantly reduced by requiring a single neutron in the detection process:

- β-radioactivity (0 neutron)
- Solar ν_e-e scattering (0 neutron)
- Michel electrons from cosmic ray muons[3] (0 neutrons)
- Atmospheric neutrino[4] inelastic scattering (0-2 neutrons)

Figure 1.12 visually demonstrates the variations in neutron counts among the aforementioned physics signals observable in SK.

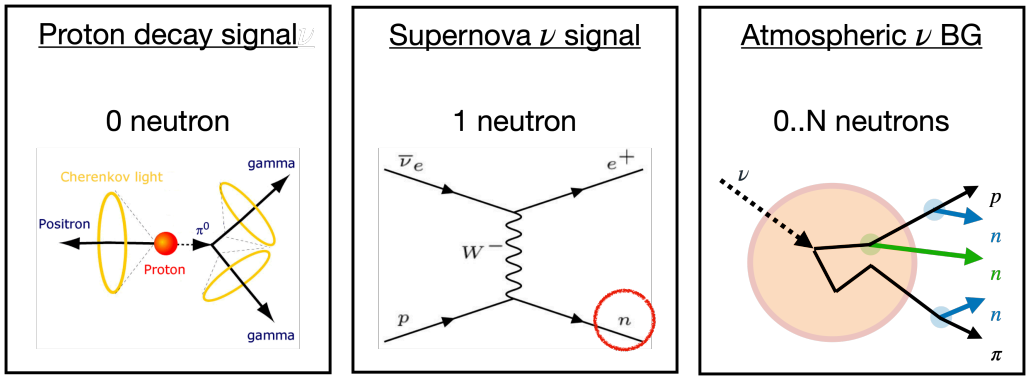


Figure 1.12: Illustration of the expected neutron counts for different physical events observable in Super-Kamiokande.

[3]Capture of a muon by ^{16}O can result in neutron emission through nuclear de-excitation, and such events may contaminate the signal sample.

[4]Neutral-current quasi-elastic (NCQE) interactions of atmospheric neutrinos, which knock out a single neutron, and the residual oxygen nucleus de-excites by emitting a $\sim$ 6 MeV gamma-ray. This becomes the dominant background in characterizing the DSNB flux.

1.4 SK-Gd

To enhance the efficiency of neutron detection, Super-Kamiokande (SK) has introduced Gadolinium (Gd), recognized as the most strongly neutron-absorbing element, into its inner water volume. The details of the SK detector and the Gd-loading process can be found in Section 2.2.2.

When neutrons are captured by $^{155/157}$Gd, the resulting compound nucleus $^{156/158}$Gd* undergoes de-excitation by emitting multiple gamma-rays with a combined energy of 8.54 and 7.94 MeV, respectively. These gamma-rays are more distinguishable than the 2.2 MeV signal from $^1\mathrm{H}(n,\gamma)$ in the pure water phase. In Figure 1.13, the typical energy deposition on the photosensors on the cylindrical detector wall is illustrated for both H and Gd(n,γ) signals. In contrast to H(n,γ) signals, the Gd(n,γ) vertex can be estimated with much greater resolution. This capability enables studies of recoil neutron kinematics and neutrino direction reconstruction [28].

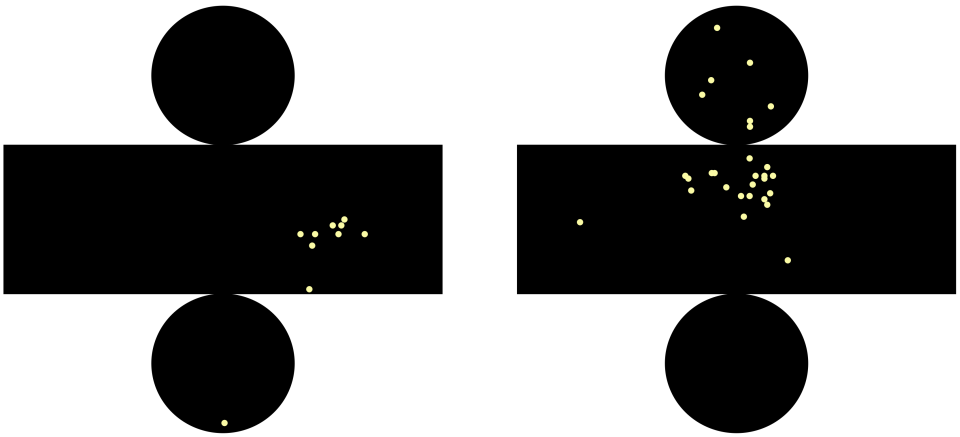


Figure 1.13: Event displays for $^1\mathrm{H}(n,\gamma)$ (left) and $^{157}\mathrm{Gd}(n,\gamma)$ (right). Yellow dots indicate the arrival of a single photon recorded on a photomultiplier tube (PMT). PMTs are not drawn to scale.

In pure water, more than 99.9% of neutron captures occur on Hydrogen (H). With Gd-loading, the fraction of neutrons captured by Gd increases, leading to a much faster rate of neutron captures. Figure 1.14 shows both the characteristic neutron capture time constant and the fraction of neutrons captured by Gd.

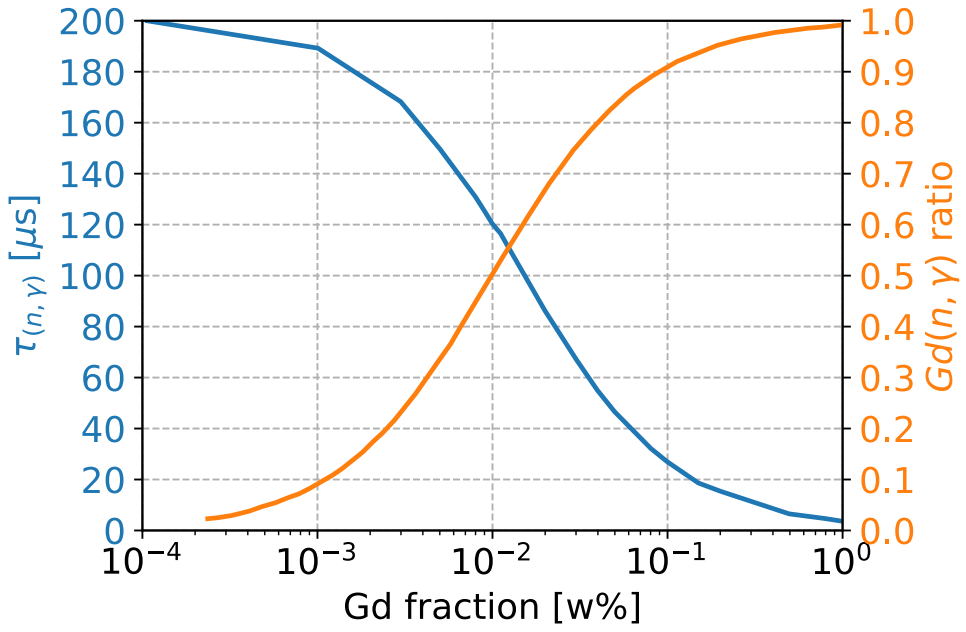


Figure 1.14: The expected neutron capture time constant $\tau_{(n,\gamma)}$ and the Gd(n,γ) fraction are plotted against the loaded Gd in mass fraction. For the SK-VI phase, which is within the scope of this thesis, the Gd mass fraction in the target volume is 0.0110w%, resulting in an expected capture time of around 120 μs and a Gd(n,γ) fraction of 50%. These predictions were generated using Geant4.9.5 [29, 30] with ENDF/B-VII.1 evaluated nuclear cross-section data [31].

1.5 Sources of neutron-related uncertainties

Unfortunately, there are many sources of uncertainties within the prediction of neutron production, and there exists no consensus on the exact magnitude of these uncertainties. Nevertheless, we can categorize these uncertainties into two main groups: a) modeling of neutrino-nucleus interactions, handled by neutrino event generators, and b) detector simulation. The details for both categories are provided in Sections 2.3.2 and 2.3.5. We outline the dominant sources of uncertainties here.

- **Primary neutrino-nucleon interaction (PI)**
  - Fermi motion and binding energy of target nucleon
  - Nucleon-level interaction cross-sections and kinematics
  - Hadronization
- **Hadronic final-state interaction (FSI)**
  - Hadron transport within target nucleus
- Nuclear de-excitation

- **Hadronic secondary interaction (SI)**
  - Hadron transport within detector volume
- Neutron capture detection efficiency

Numerous models, particularly those simulating hadronic interactions, contribute to a significant variability in their predictions. Furthermore, there are only a few measurements of nucleon-^{16}O cross-sections available in the 0.1-1 GeV range, which we can use to estimate uncertainties and constrain the models. For a detailed overview and visualization of the PI, FSI, and SI components of neutron production, refer to Section 2.3.2 and Figure 2.46.

1.6 Past neutron measurements

Over the past decade, efforts have been made to measure recoil and shower neutrons generated from neutrino interactions, aiming to evaluate the consistency between observed data and model predictions. We provide a brief overview of these neutron measurements, their implications, and potential avenues for future investigations.

The very first measurement of neutron multiplicity emerged from SK in 2011 [32]. The study presented results showing the mean neutron multiplicity as a function of visible energy (E_{vis}) in atmospheric neutrino events. The dataset spanned approximately 2 years, and the neutron detection efficiency stood at around 20%. Notably, the measurement lacked consideration of systematic uncertainty and was not compared with any model predictions.

A comparable study was published by the Sudbery Neutrino Observatory (SNO) in 2019 [33], using a Cherenkov detector similar to SK but with heavy water D_2O as the target. The dataset spanned about 4 years of atmospheric neutrino data, with an average neutron detection efficiency of 15% for $D(n,\gamma)$ signals and 45% for $Cl(n,\gamma)$ signals. The dominant uncertainties were from neutron detection efficiency and neutrino interaction modeling, with no consideration given to uncertainties in hadronic interactions within the detector volume. The SNO results, based on atmospheric neutrino data, showed reasonable agreement with the 2011 SK findings, as well as simulations from `GENIE` 2.10.02 and Geant4.10.0, as shown in Figures 1.15 and 1.16. In the sub-GeV visible energy range, the observed neutron multiplicity was slightly lower than the predictions, but the sizable statistical uncertainty in the observation makes it challenging to definitively label any model as inaccurate.

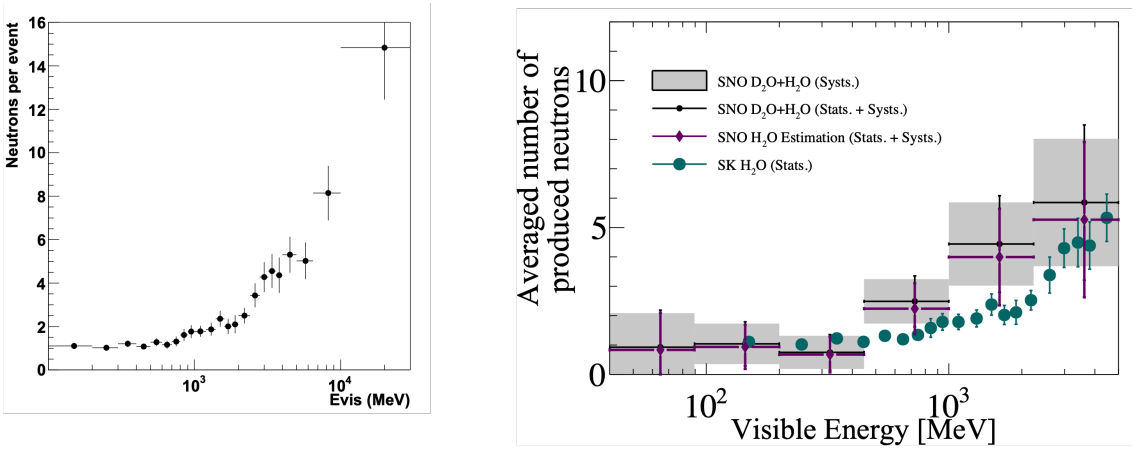


Figure 1.15: The average neutron signal multiplicity observed in atmospheric neutrino interactions for SK (left, [32]) and SNO (right, [33]), presented as a function of event visible energy E_{vis}. In the right figure, the SNO H_2O prediction is represented by purple points with error bars, including both statistical and prediction errors. showing reasonable agreement with the overlaid SK data represented by dark green points with statistical error bars.

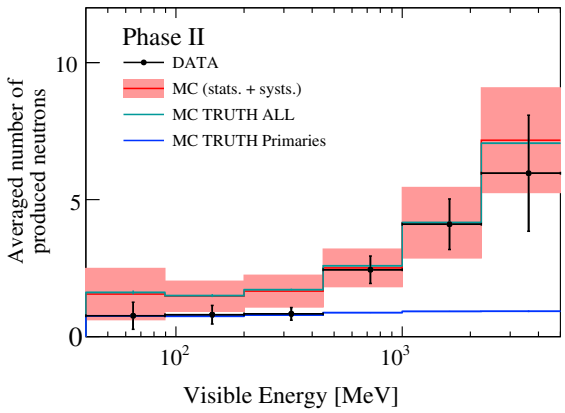


Figure 1.16: A comparison between predicted (red) and observed (black, with statistical errors) average neutron signal multiplicities in SNO is presented. The prediction was generated using neutrino event simulation with `GENIE` 2.12.02 and detector simulation with Geant4.10.0. The overall prediction error of 25% includes a 13% uncertainty due to PI/FSI within `GENIE` and a 16% uncertainty in signal efficiency. (Reprinted from [33])

T2K [34] conducted a measurement of neutron multiplicity in $\nu_\mu/\bar{\nu}_\mu$-CCQE-like interactions at its far detector (Super-Kamiokande), utilizing accelerator neutrinos. The measurement explored the dependence on the transverse muon momentum p_μ^T, which is believed to be positively correlated with the four-momentum transfer and the energy transferred to the hadronic system. This represents the first neutron production measurement employing well-controlled accelerator neutrino events.

The uncertainty in neutron detection efficiency in the far detector (SK) during the pure water phase was extensively examined. The observed neutron multiplicity was notably smaller than predictions, particularly for neutrino interactions with low energy transfer. Furthermore, the study revealed significant variability in the neutron multiplicity predictions among different neutrino event simulators. The results are shown in Figure 1.17.

Since this measurement utilized events observed at SK, the results currently serve as a basis for estimation of neutron production uncertainty in SK's atmospheric neutrino oscillation analysis. It is important to note that the result is dominated by statistical errors, and the uncertainty estimate is substantial, reaching up to 60% for events with high transverse momentum transfer.

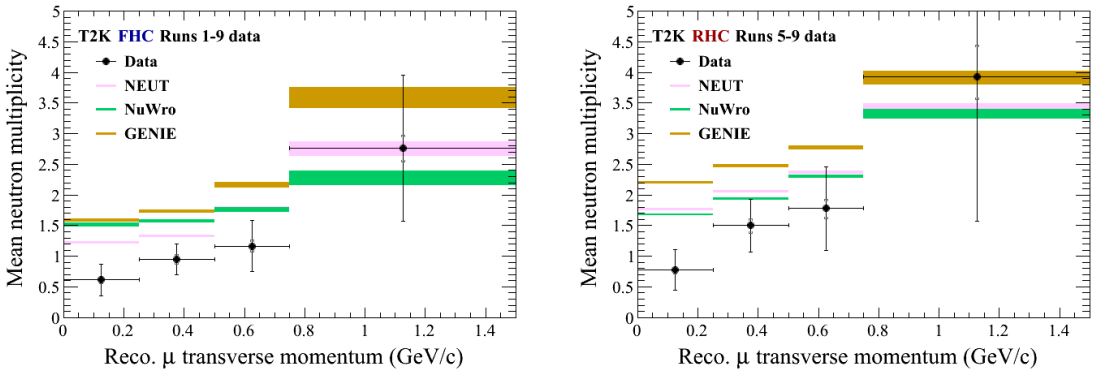


Figure 1.17: The average neutron signal multiplicity for ν_μ (left) and $\bar{\nu}_\mu$ (right) beam modes, observed at the far detector (SK), as a function of the reconstructed muon p_μ^T. The error bars in the data includes statistical errors (dominant) and systematic errors related to signal efficiency. The shaded regions are predictions from three distinct neutrino event generators, incorporating errors in signal efficiency.

In 2019, the MINERvA collaboration [35] presented the first measurement of neutron kinematic variables, such as $1/\beta$ ($\beta = v_n/c$). with a $\bar{\nu}_\mu$ beam with an average energy of 3 GeV and a fine-grained polystyrene $(C_8H_8)_n$ scintillator tracker as the target. Neutrons scatter protons and leaves a visible signal a small distance apart from the production vertex. Neutron kinematics can be deduced from the track length and time-of-flight of the signal. While the production of energetic neutrons aligned with predictions, there was a noticeable deficit of low-energy neutrons, surpassing the comprehensive uncertainty estimates.

In 2023, MINERvA published another measurement [36]. This time, they investigated the $\bar{\nu}_\mu$ multi-neutron production cross-section as a function of p_μ^T, using a higher average beam energy of 6 GeV. Once again, a deficit of low-energy neutrons was observed. The measured differential cross-section was compared with various PI models that considered the knock-out of multi-nucleon pairs. The measurement revealed a significantly lower cross-section compared to the expected values from the models. Figure 1.18 summarizes the observed deficit in neutron production in MINERvA.

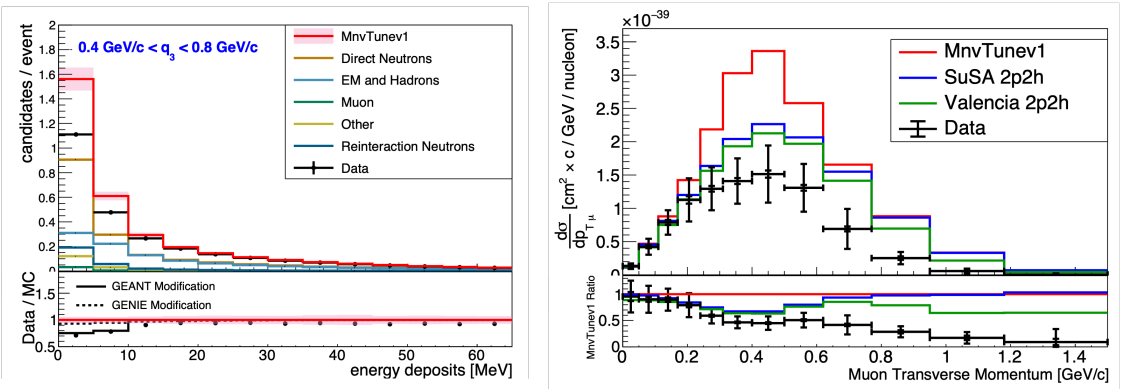


Figure 1.18: A comparison between predicted and observed distributions for the signal energy deposit (a proxy for neutron kinetic energy, left figure) and the multi-neutron production cross-section (as a function of p_μ^T, right figure) in the MINERvA hydrocarbon target. (Reprinted from [36])

Table 1.2 summarizes the past neutron measurements and this study.

	SK	SNO	T2K	MINERvA	SK
Year	2011 [32]	2019 [33]	2019 [34]	2019 [35] / 2023 [36]	2023
ν source	Atm.	Atm.	Acc. ($\nu_\mu/\bar{\nu}_\mu$)	Acc. ($\bar{\nu}_\mu$)	Atm.
E range [GeV]	0.1-20	0.1-5	0-5	2-6 / 2-12	0.1-10
Target	H_2O	D_2O	H_2O	C_8H_8	H_2O
ν events	$\sim$ 7000	512	311	16,129 / $\sim 3\times10^5$	39,376
n signal	H(n,γ)	D(n,γ) Cl(n,γ)	H(n,γ)	Recoil p	H(n,γ) Gd(n,γ)
Detected n	$\sim$ 2000	476	77	$\sim 10^4$ / $\sim 2\times10^5$	23,492
n efficiency	$\sim$ 20%	$\sim$ 15% $\sim$ 44%	$\sim$ 20%	$\sim$ 40-60%	$\sim$25% $\sim$50%
Observable (y)	Mean n multiplicity			n distribution / Multi-n xsec	Mean n multiplicity
Considered uncertainty (δy)	Stat only	PI/FSI+eff.		PI/FSI+SI+eff.	
Variable (x)	E_{vis}		p_μ^T	$1/\beta$ / p_μ^T	E_{vis}
PI/FSI modeling	`NEUT`	`GENIE`	`NEUT` `GENIE` `NuWro`	`GENIE`	`NEUT` `GENIE`
SI modeling	`Bertini`				`Bertini` `INCL`

Table 1.2: A simplified summary of previous neutron measurements and this study (rightmost column). The numbers marked with the $\sim$ symbol signify approximations or rough estimation by the author.

1.7 Thesis outline

This thesis focuses on measuring the average neutron capture multiplicity of atmospheric neutrino interactions observed in the Super-Kamiokande (SK) detector, as a function of neutrino event's visible energy. With the largest dataset of both neutrinos and neutron events in water target, this measurement also presents the first measurement of "track length" for neutrons in atmospheric neutrino interactions, made possible by recent Gd-loading. The results were compared with various models, some not included in any previous neutron production measurements.

Key motivations include:

- Quantifying systematic uncertainty in total neutron production in atmospheric neutrino interactions in water
- Comparing model predictions with data and providing constraints

Chapter 2 explains the Super-Kamiokande detector and the modeling of atmospheric neutrino interactions, as well as subsequent hadron and neutron interactions.

Chapters 3 and 4 outline the core contributions of this thesis. Chapter 3 covers the establishment and calibration of the neutron detection algorithm for use in SK, addressing both pure water and Gd-loaded phases. Various sources of uncertainty affecting neutron detection efficiency in a Gd-loaded water Cherenkov detector are studied in detail.

Chapter 4 explores data selection and quality for both atmospheric neutrino and neutron signals. It then discusses modeling neutron detection performance, considering variations with neutrino event properties, and contributing to the reduction of uncertainty. The chapter also details neutron multiplicity and uncertainty estimation, comparing results with various simulation models. Finally, Chapter 5 explores the implications and future prospects of the research.

1.8 Author's contributions

While the work presented in this thesis involved collaboration with members of the Super-Kamiokande collaboration, the primary content in Chapters 3, 4, and 5 is the result of the author's efforts. The key contributions are summarized below.

In Chapter 3, the author overhauled the neutron detection software, building upon the original algorithm for pure water phase data. Although the core algorithm remained mostly unchanged, a significant redesign of the neural network, including a streamlined feature set, followed modern machine learning standards. The introduction of a new algorithm (`RECO` vertex mode, detailed in Section 3.2) ensured consistent neutron detection performance across various capture distances, contributing to the software's robustness.

In the collaborative Am/Be neutron source calibration, the author made three key contributions. First, a comprehensive Am/Be simulation (explained in Section 3.3.3) was developed, including continuous source activity and fine-tuned scintillator optical parameters. Second, the author performed a complete extraction of signal efficiency from calibration data across SK-IV, SK-V, and SK-VI phases, comparing results and deriving scaling ratios for data/simulation efficiency (refer to Table 4.6). Third, systematic uncertainties were evaluated and an overprediction of the Gd neutron capture ratio in the Geant4 simulation was identified. These efforts significantly contributed to resolving the previously unexplained data-simulation discrepancy in estimated signal efficiencies.

Additionally, the author estimated potential contamination of Michel electrons and PMT afterpulses, as described in Sections 3.4 and 3.5.

The core content in Chapters 4 and 5 represents original work by the author, including the estimation of neutron detection efficiency through non-linear regression on multiple event variables, measurement of neutron multiplicity, systematic uncertainty evaluation, prediction of neutron multiplicity estimates for various hadronic interaction model combinations available in the detector simulator, linearity measurements, and model comparisons with observations.

Chapter 2

Atmospheric neutrino interactions at Super-Kamiokande

2.1 Atmospheric neutrinos

Atmospheric neutrinos originate from interactions between cosmic rays, mainly protons making up about 90%, and the Earth's atmosphere. The source of these cosmic rays is currently under investigation, but most are believed to come from outside the Solar System, from sources such as supernova remnants. When these cosmic rays collide with atoms in the Earth's atmosphere, predominantly nitrogen and oxygen nuclei, they create a cascade of secondary particles, primarily pions and kaons. These particles then undergo decay processes, resulting in lighter leptons, including muon neutrinos and electron neutrinos. The schematic is illustrated in Figure 2.1. Muon neutrinos result from both meson decays and subsequent muon decays, while electron neutrinos are dominantly produced in muon decays. As a result, the expected count of muon neutrinos is anticipated to be roughly twice that of electron neutrinos.

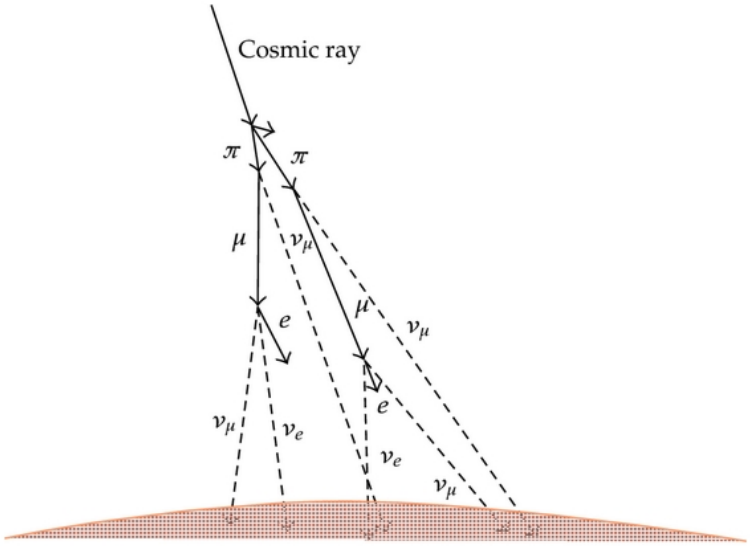


Figure 2.1: Schematic of atmospheric neutrino production. (reprinted from [37])

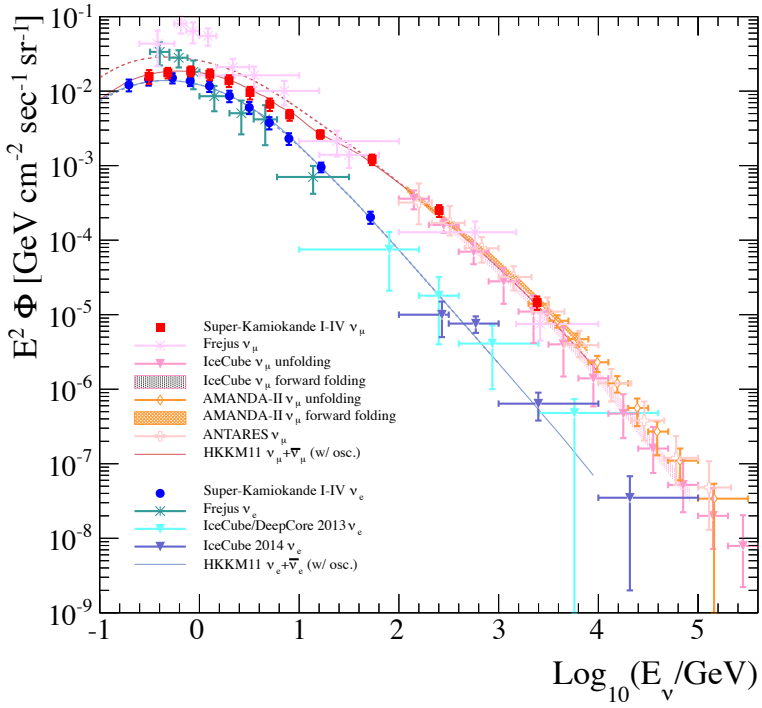


Figure 2.2: The measured energy spectra of the atmospheric ν_e and ν_μ fluxes by SK, shown with measurements by other experiments. (Reprinted from [38])

2.2 The Super-Kamiokande detector

The Super-Kamiokande (SK) detector is an underground water Cherenkov detector. Figure 2.3 illustrates schematic drawings of the setup. The detector consists of a cylindrical tank divided into two coaxial volumes: the inner detector (ID) and the outer detector (OD), both filled with water. The ID serves as the active target for physics analyses, detecting Cherenkov radiation through inward-facing photosensors mounted on the tank wall. Meanwhile, the OD, equipped with outward-facing photosensors, acts as a veto for cosmic rays and captures a portion of the energy from physical events escaping the ID. Our primary signal involves atmospheric neutrinos interacting within the ID volume, which typically show minimal activity in the OD. The distinct Cherenkov radiation signatures from scattered charged leptons (electrons or muons) with energies in the range of $O(0.1\text{-}1)$ GeV are observable within the ID. In cases where neutrons are produced during atmospheric neutrino interactions in water, faint neutron capture signals with energies around $O(1)$ MeV become visible in the detector within approximately $O(100)$ μs from the interaction. The following sections describe the operational principles of the detector, as well as description of hardware and software components used for signal detection, calibration procedures, and a recent Gd-loading campaign aimed at enhancing neutron capture detection efficiency.

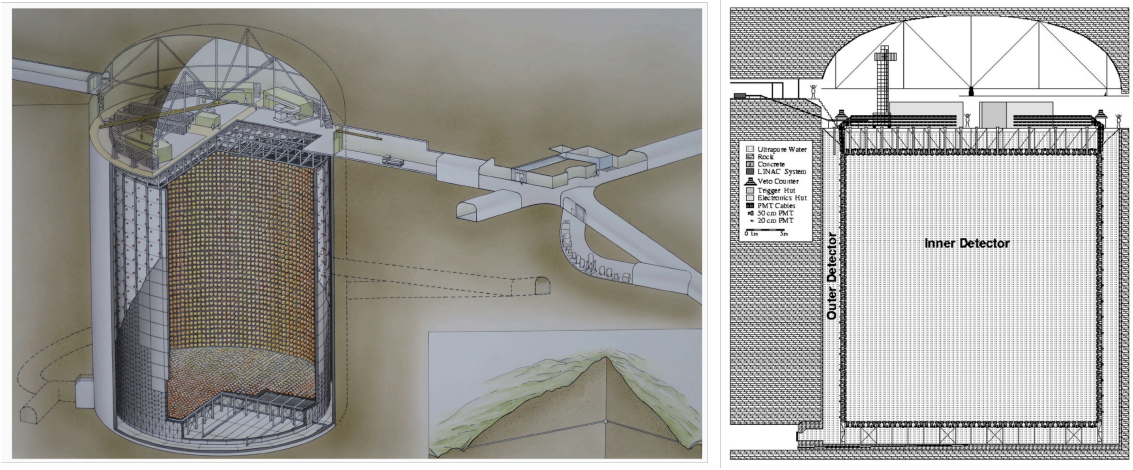


Figure 2.3: A schematic representation of the SK detector along with its relevant infrastructure (left [39]), and a sectional view of the detector system (right [40]).

2.2.1 Working principle

As a charged particle traverses a dielectric (polarizable) medium, it undergoes electromagnetic interactions with the surrounding molecules. This interaction excites the molecules, inducing radiation. Notably, if the particle's speed surpasses the phase velocity of the electromagnetic wave in the medium, radiation emitted simultaneously from different positions can combine on the wavefront, generating a kind of shockwave characterized by an angle of $\cos^{-1}(\beta/n)$. In this context, β denotes the ratio of the charged particle's speed to the speed of light in a vacuum, while n represents the medium refractive index which may vary with the wavelength of the radiated photons. These emitted photons, peaking at around 375 nm, are detectable using a photosensor.

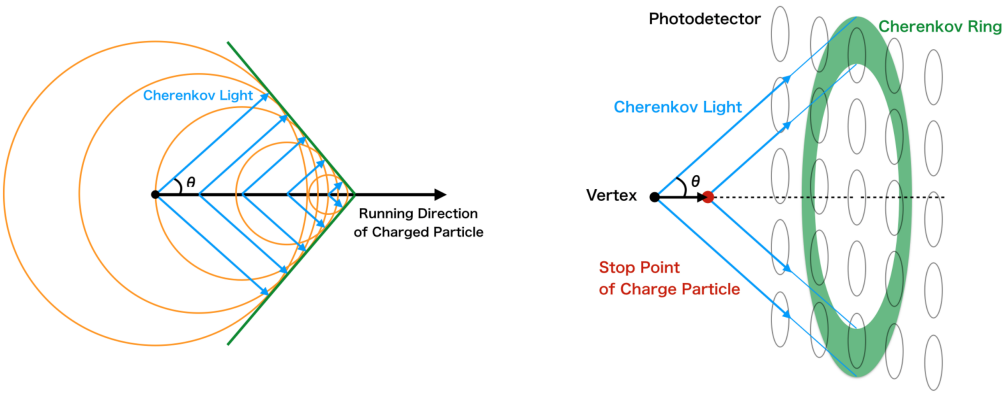


Figure 2.4: Illustration of Cherenkov radiation dynamics (left) and the resulting ring-shaped image on a photosensor plane (right). (Adapted from [41])

When a particle slows down below the Cherenkov threshold speed ($\beta_{\text{min}} = 1/n$), the particle ceases to radiate, resulting in a distinctive ring-shaped pattern, as illustrated in Figure 2.4. By setting up a photodetector plane (as shown in Figure 2.4) to measure the ring radius and estimate the number of emitted photons, we can technically reconstruct both the position and energy of the incoming charged particle. Figure 2.5 presents a typical Cherenkov ring generated by a muon-neutrino charged-current interaction observed in the detector.

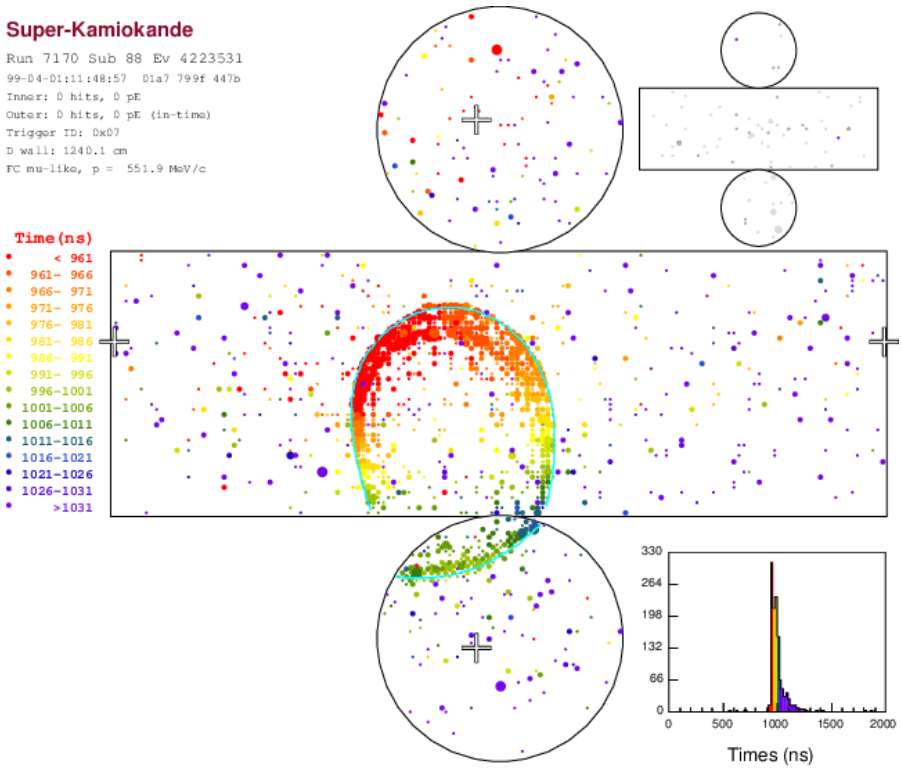


Figure 2.5: A sample display of 0.6 GeV muon-neutrino charged-current interaction.

The energy radiated per unit track length is described by the Frank-Tamm formula (in cgs units) [42]:

$$\frac{dE}{dx} = \frac{(ze)^2}{c^2} \int_{\epsilon(\omega)>1/\beta^2} \omega \left(1 - \frac{1}{\beta^2 \epsilon(\omega)}\right) d\omega \tag{2.1}$$

Here, ze represents the charge of the moving particle, ϵ is the permittivity of the medium, and ω is the frequency of the radiation. (We assume a magnetic permeability $\mu(\omega) = 1$) For an ultrarelativistic particle with $\beta \approx 1$, both the energy loss per path length and the Cherenkov angle remain nearly constant, given the particle's charge and the dielectric properties of the medium. The energy loss is in the range of O(1-2) keV/cm or O(200-1000) visible photons/cm, which is exceptionally small compared to the energy lost through ionization, typically on the order of $O(1)$ MeV/cm. Nevertheless, it proves to be a practical option for reconstructing a charged particle.

The water Cherenkov technique offers a distinct advantage in efficiently distinguishing electrons and muons. Electrons are more prone to induce electromagnetic interactions, such as pair production and bremsstrahlung[1]. As a result, Cherenkov rings produced by electrons often appear "fuzzier" compared to their muon counterparts. This "fuzziness" can be quantified and utilized as a particle identification (PID) discriminator, as illustrated in Figure 2.6. Further details on this discrimination technique are provided in Section 2.4.1.

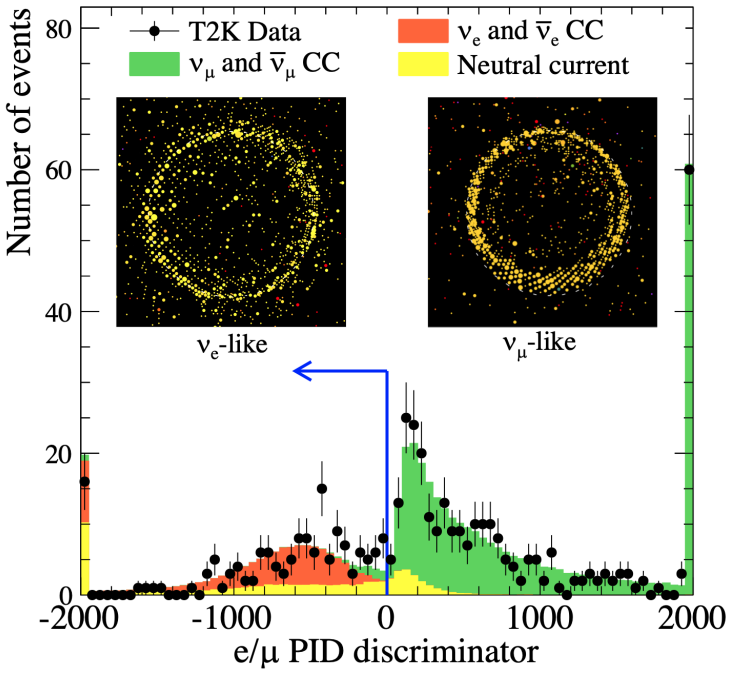


Figure 2.6: Illustration of the ν_e/ν_μ identification among beam (T2K) neutrino data collected in SK, based on the clarity of Cherenkov rings. (Reprinted from [43])

However, a drawback to this technique is that low-energy heavy particles are not visible, due to the energy threshold of Cherenkov radiation increasing with particle mass. The kinetic energy thresholds are 0.26 MeV for electrons, 54 MeV for muons, 71 MeV for charged pions, 255 MeV for charged kaons, and 481 MeV for protons. This is an inherent trade-off in the technique, prioritizing the neutrino event statistics.

[1]The critical energy for electrons in water is approximately 70 MeV, while for muons, it is around 1 TeV.

2.2.2 Detector components

Detector location

The detector is positioned at coordinates 36°25'32.6"N, 137°18'37.1"E, located about 1,000 meters below ground in the Kamioka Mine in Gifu, Japan, with an average overburden of 2,700 meters water equivalent (m.w.e.). This depth effectively diminishes the cosmic ray flux by a factor of roughly 10^{-5} in comparison to sea-level intensity. At this depth, the cosmic muon flux at SK reaches approximately 10^{-7} $cm^2 \cdot s^{-1} \cdot sr^{-1}$, resulting in an event rate of around 2 Hz at the detector.

Detector geometry and design

The detector is a cylindrical tank divided into two distinct, optically isolated volumes. Surrounding the tank are 26 Helmholtz coils to reduce the geomagentic field (∼450 mG) down to 50 mG, maximizing PMT photoelectron collection efficiencies.

The ID volume, measuring 33.8 meters in diameter and 36.2 meters in height, contains 32 kilotons of water. This space is currently equipped with 11,129 inward-facing 50 cm PMTs (Hamamatsu R3600) covering the entire inner surface. To minimize reflections, the inner wall (excluding the PMTs) is coated with black polyethylene terephthalate sheets. The effective photocoverage is approximately 40%. Due to potential background radioactivity near the tank wall, our analyses focus on events sufficiently separated from the tank wall. In recent atmospheric neutrino analyses, the Fiducial Volume (FV) of the ID is defined as a cylindrical space located 1 meter inward from the ID wall, with a total mass of 27.2 kilotons.

The OD wall is sparsely equipped with 1,885 outward-facing 20 cm PMTs (Hamamatsu R1408), serving as an active veto counter against incoming particles and providing a passive shield against neutrons and gamma-rays from the surrounding rocks. The OD region, with a width of about 2.5 meters, contains a total water mass of approximately 18 kilotons. To enhance reflection and photon detection efficiency, the OD wall is covered with a white Tyvek® sheet.

Water purification

To maintain optimal water purity in the SK detector, a continuous circulation and purification process is imperative. Unwanted radioactivity, particularly Radon and Radium, naturally present in air and water, increases noise hits and mimics low-energy PMT hits. The both the tank water and water sourced from the mine undergoes constant circulation through a purification system at a flow rate of 30 to 60 tons per hour, achieving full recirculation approximately every month.

The purification system, consisting of multiple stages, focuses on three key objectives: filtering particulates, controlling bacterial growth, and reducing Radon through degasification. Water is cooled to 13°C, aiding in maintaining uniform PMT dark noise levels and preventing bacterial growth. Water purity is assessed by measured resistivity, increasing from around 11 MΩ·cm pre-purification to 18.2 MΩ·cm post-purification, reaching the chemical limit for ultra-pure water. Particulates smaller than 0.2 μm entering the ultra-filter decrease from an estimated 1000 cm^{-3} to 6 cm^{-3}. The Radon concentration in water returning to the SK tank after filtration is estimated at 0.4 $mBq \cdot m^{-3}$. The additional features of the water purification system with Gd-loading is explained in Section 2.2.4.

Air purification

The mine's air, richer in radon gas due to radon-containing rocks, sees concentrations of about 1,000 to 2,000 Bq/m^3 in summer and 100 to 300 Bq/m^3 in winter due to varying air-flow patterns. An air purification system is employed to reduce radon levels, resulting in concentrations below 3 mBq/m^3, supplied between the water surface and the top of the SK tank.

Inner detector photomultiplier tubes

The ID utilizes 50 cm Hamamatsu R3600 photomultiplier tubes (PMTs), and their design is illustrated in Figure 2.7. The photocathode of the ID PMT is coated with bialkali (Sb–K–Cs) to match the wavelength of Cherenkov light, with a sensitivity range from 280 to 660 nm. At its peak (λ ~380 nm), the maximum quantum efficiency (QE) is approximately 21%. The average collection efficiency at the first dynode is around 70%, maintaining uniformity within ±7% across the entire PMT surface. The dynode structure follows the Venetian Blind style, and the gain is $O(10^7)$ at a high voltage of 2,000 V. The relative transit time spread for the single photoelectron signal is about 2.2 ns (1σ). An individual signal pulse with a charge deposit equivalent to or exceeding 0.25 photoelectrons is digitized as a hit. The average dark-hit rate at the 0.25-photoelectron threshold is approximately 3 kHz.

All ID PMTs are coated with an acrylic layer with over 96% transparency for a wavelength of 350 nm, and they feature a fiber-reinforced plastic (FRP) case to shield against pressure and prevent potential damage from spreading to other PMTs.

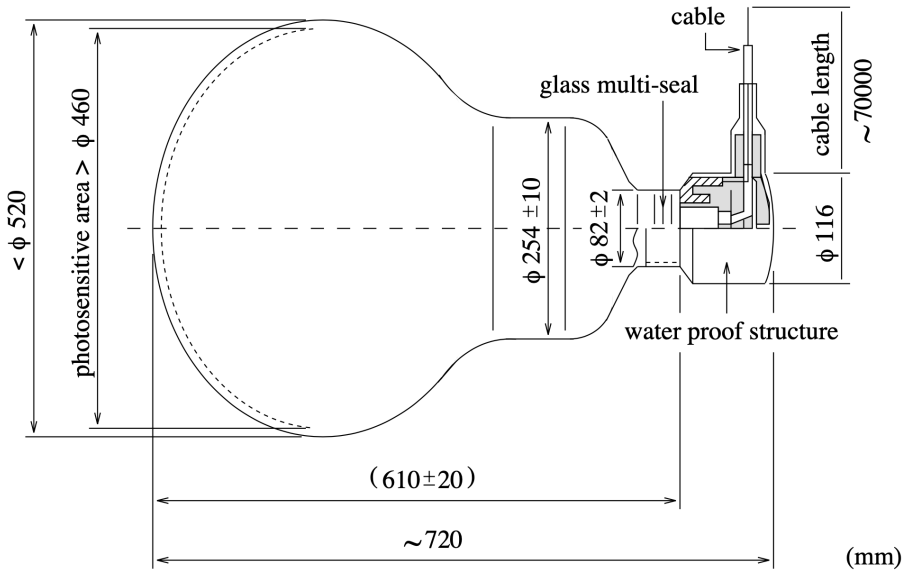


Figure 2.7: ID PMT design. (Reprinted from [40])

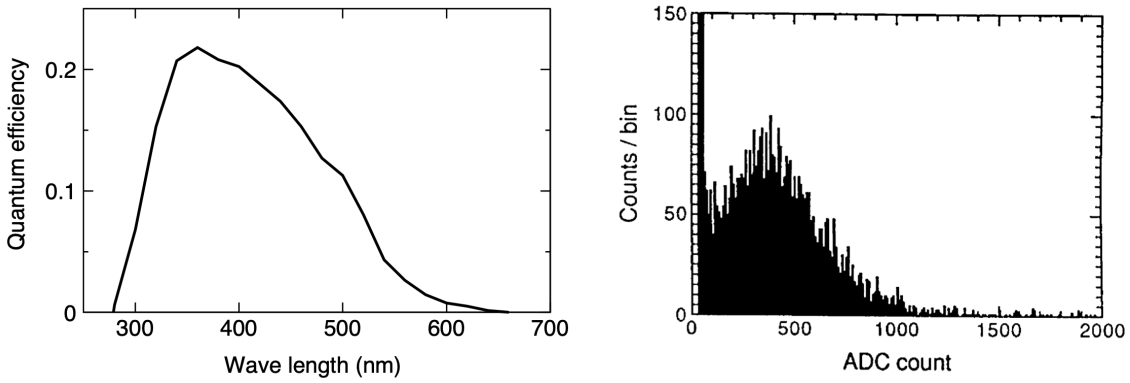


Figure 2.8: Quantum efficiency of the ID PMT photocathode plotted against incident photon wavelength (left) and the distribution of single photoelectron pulse heights observed with the ID PMT (right). (Reprinted from [40])

Frontend electronics, readout, online data processing

The signal from each ID PMT is directed to a module called "QTC-based Electronics with Ethernet" (QBEE). Each QTC (charge-to-time converter) has three input channels, and there are eight QTCs on each QBEE board, allowing it to manage input analog signals from 24 PMTs. The schematic of the QBEE module is shown in Figure 2.9.

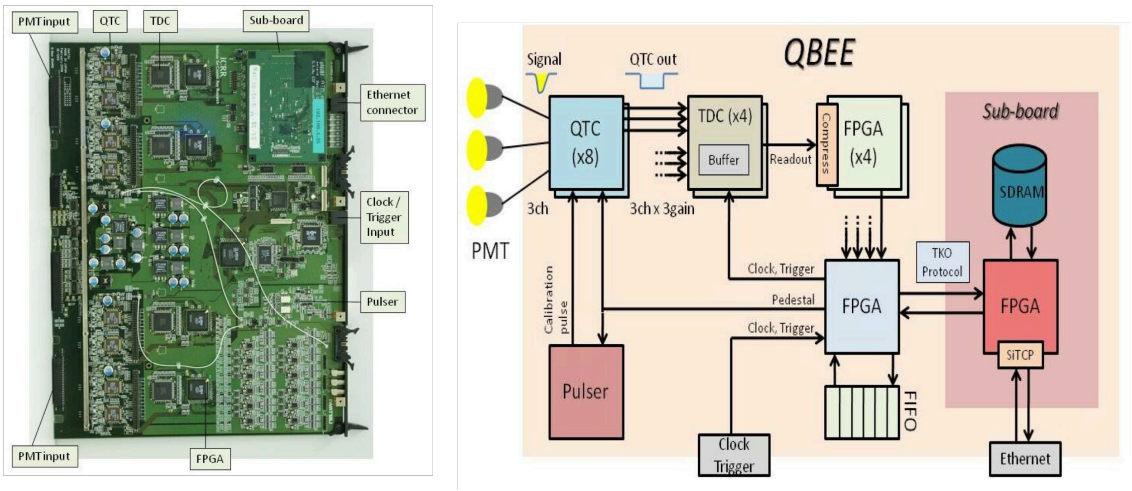


Figure 2.9: The actual QBEE board (left) and a schematic illustrating the data flow within the module (right). (Reprinted from [44])

A QTC takes an analog input signal and turns it into a square gate pulse, where the time width of the pulse corresponds to the integrated charge. If the input signal is above a certain threshold (for SK, -0.69 mV, which is equivalent to 0.25 photoelectrons), the signal's leading edge initiates the timing process. The charging and discharging gates, along with a VETO gate, remain open for 400 ns, 350 ns, and 150 ns, respectively. All signal currents are ignored while one of the gates is open. Consequently, the total time needed to process a single input current is approximately 900 ns. The time it takes for the capacitor to charge and discharge hit a certain voltage level is measured, which is proportional to the input charge. The linearity of this time width with respect to the input charge is within ±1%. This process is illustrated in Figure 2.10.

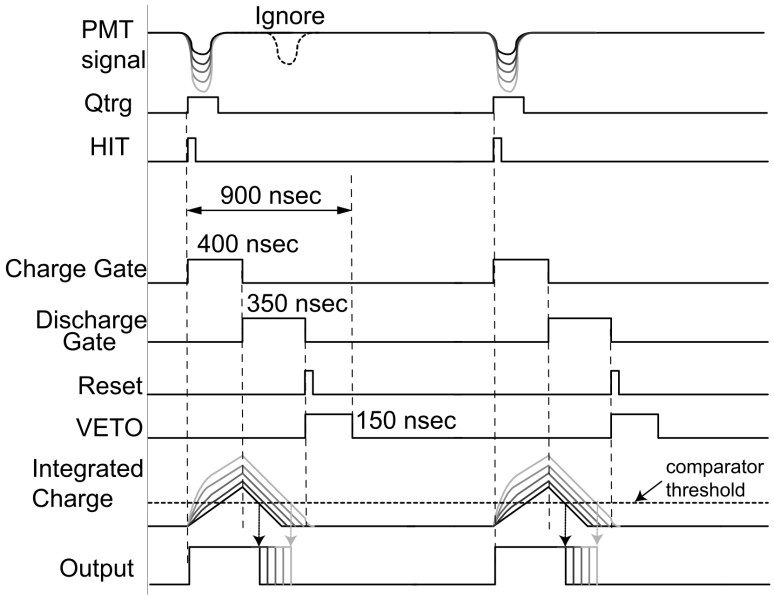


Figure 2.10: The schematic illustrating the gates and output issued following a PMT pulse above the discriminator threshold in the QTC. (Reprinted from [45])

The QTC incorporated in the QBEE modules features three gain stages: Small, Medium, and Large, with gain ratios set at $1 : 1/7 : 1/49$. The QTC application-specific integrated circuit (ASIC) performance is outlined in Table 2.1. Notably, the QTC exhibits a high dynamic range, ensuring the accurate reconstruction of light yield from GeV-scale events.

Input channels	3
Processing speed	∼900 ns / cycle
Gain ratios	1 : 1/7 : 1/49
Discriminator threshold	-0.3 ∼ -14 mV (Small)
Dynamic range	0.2 ∼ 51 pC (Small)
	∼ 357 pC (Medium)
	∼ 2500 pC (Large)
Charge resolution	∼0.1 pC (Small)
Output linearity	<1%
Timing resolution	0.3 ns (2 pC input)
	< 0.2 ns (>10 pC input)

Table 2.1: Performance of the QTC ASIC. (Adpated from [44], [46])

The QTC output pulse duration is digitized by the time-to-digital converter (TDC), which temporarily stores hit time and charge information within its internal buffer. A single 60 MHz master clock issues a trigger every 1,024 clocks (approximately 17 μs) to the field programmable gate arrays (FPGA) on each QBEE board. Care was taken to ensure that signal transfer cables for this master clock trigger have nearly equal lengths. Following the master clock trigger, each FPGA reads and compresses the hit information stored in the TDC buffer simultaneously during the 17 μs period, saving them in a first-in-first-out (FIFO) memory.

The FIFO memory content is read by a network interface card, and the information is transferred to online PCs via a Fast Ethernet connection, achieving a throughput of 11.8 MB/s per QBEE. All recorded PMT hits are sorted in time and segmented into 22-ms length blocks by a parallel processing approach involving 20 frontend PCs. The segmented data are then merged, and triggers are applied to construct physical events. The triggered events are reformatted into CERN ZEBRA and ROOT formats in real-time, serving various physics analyses. The data flow schematic within the online system is illustrated in Figure 2.11. Notably, this system can handle data flow from individual PMT hit rates of up to 10 kHz without loss of information.

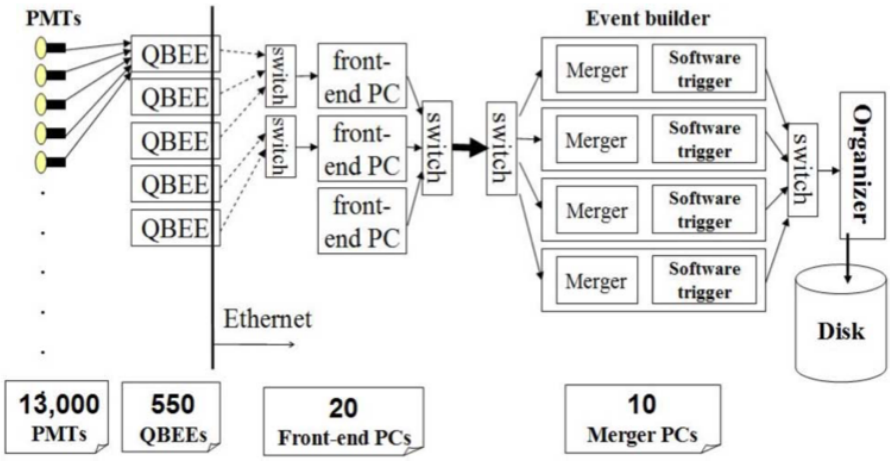


Figure 2.11: The schematic of the online data processing. (Reprinted from [47])

Software trigger

The Merger PCs, illustrated in Figure 2.11, use a "software" trigger to construct an event from time-sorted PMT hits. If the count of PMT hits within a 200-ns sliding window exceeds a predetermined threshold, the PMT hits within the corresponding event window (usually [-5, +35] μs from the trigger) are combined into a single event. The thresholds and event window details for each event trigger type can be found in Table 2.2. Originally configured to mimic trigger conditions and performances before the 2008 electronics upgrade introducing QBEEs, Fast Ethernet, and software triggers, these trigger definitions are continuously updated during data collection in response to changes in PMT dark rates and failed PMTs.

Trigger Type	Threshold [hits/200 ns]	Event window [μs]
Super High Energy (SHE)	70 $\rightarrow$ 58	[-5, +35]
High Energy (HE)	50	[-5, +35]
Low Energy (LE)	47	[-5, +35]
Super Low Energy (SLE)	34 $\rightarrow$ 31	[-0.5, +1.0]
Outer Detector (OD)	22 (in OD)	[-5, +35]

Table 2.2: Overview of software trigger types. Thresholds, indicated with arrows, may vary within the specified range.

For SHE triggers exclusively, a special attachment flag named AFT (after-trigger) is raised to capture neutron capture reactions occurring with a time scale of $O(100)$ μs. When this flag is attached, an extra 500-μs record of PMT hits is appended to the leading SHE-triggered event. If another event trigger is issued within 500 μs from the SHE event, the AFT attachment is omitted.

Another unique trigger type is the random (also known as "dummy") trigger. A "dummy" event is triggered by a clock with a frequency of $O(1)$ Hz. Initially designed to synchronize with the accelerator beam spill frequency for the T2K experiment, this trigger serves to extract background PMT hits when the beam is inactive. Dummy-triggered events have an event window of [-500, +500] μs, and the recorded PMT hits in these events are appended to our simulation to model both PMT dark noise hits and background radioactivity.

2.2.3 Calibration

This section discusses the calibration processes for both hardware and simulation. For additional information, refer to [46]. We focus specifically on the calibration related to the ID, as we do not incorporate OD information in our physics analyses, but rather use the OD solely for the purpose of cosmic ray rejection.

ID PMT and electronics calibration

To interpret the digitized output from the detector system in terms of physical quantities, it is crucial to calibrate the ID PMT and electronics. We focus on key parameters such as photon counts and the arrival time of each digitized PMT hit. The calibration process follows a specific sequence: we first calibrate the PMT charge output, and then, we fine-tune the timing response, given that the latter is dependent on the former.

High-voltage determination

Initially, 420 "reference PMTs" were pre-calibrated within a μ-metal shield, using a Xe lamp and scintillator ball combination. This pre-calibration aimed to maintain the variability of the output ADC counts for the reference PMTs below 1.3% RMS, given the same light input. These reference PMTs were symmetrically installed in the ID wall, as illustrated in Figure 2.12. Once all ID PMTs were in place within the tank, the scintillator ball used for pre-calibration was positioned at the tank center. This source then was used to adjust the high-voltage for each individual PMT, aligning it with the output charge of the reference PMTs in the same "group," i.e., in azimuthally symmetric positions as each tested PMT.

This adjustment successfully achieved an output charge variability below 1.3%, consistent with the earlier pre-calibration results. This adjustment took place only once in 2006, following an incident where approximately half of the PMTs were damaged. The scintillator ball used for the pre-calibration remains permanently positioned for real-time and long-term monitoring of the detector response.

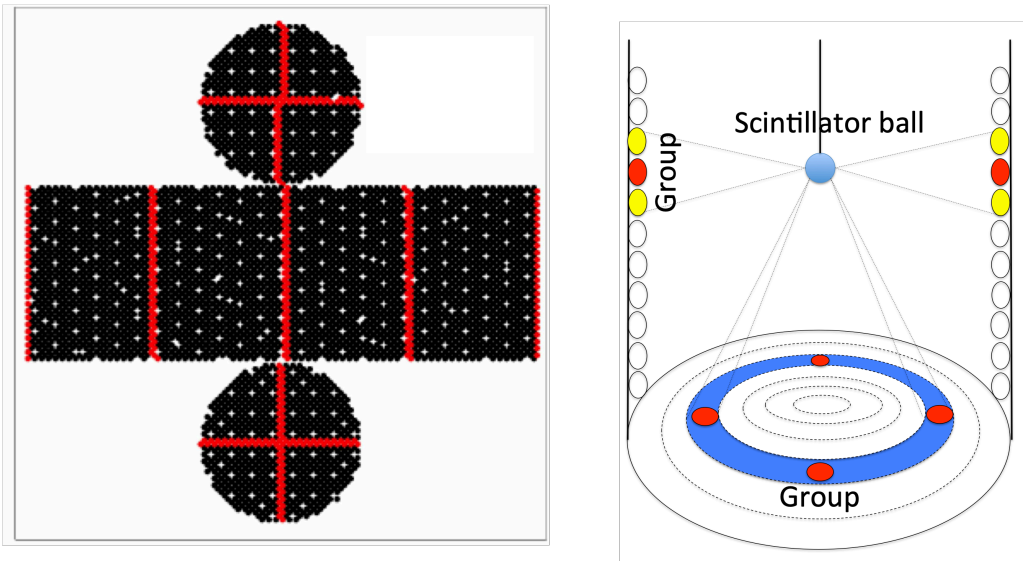


Figure 2.12: The positions of the reference PMTs (left) and a schematic illustrating the grouping of individual PMTs with reference PMTs based on their symmetrical geometrical relationship with the calibration source (right).

Gain measurement

The "gain" refers to the charge amplification rate of the PMT dynode. The process of measuring individual PMT gain involves two steps. Firstly, we determine the relative gain differences using a variable-intensity light source. Secondly, using a single-photoelectron level light source, we establish the average absolute gain scale applied to all ID PMTs, correcting for the earlier determined relative gain differences.

To evaluate relative gain differences, we utilize a scintillator ball coupled with an N-laser-driven dye laser capable of adjusting light flash intensities. By comparing the charge output during high-intensity flashes with the number of hits at a single-photoelectron level, we obtain a factor proportional to the gain of each PMT. The RMS of the distribution of this factor is 5.9%.

Absolute gain measurement involves a "Nickel ball," a spherical combination of 6.5 kg of NiO and 3.5 kg of polyethylene housing a Cf source at the center. Neutrons produced by Cf fission are thermalized and captured by ^{58}Ni, emitting 9 MeV energy through gamma-rays. Over 99% of PMT hits resulting from these signals are expected to be due to single-photon arrival. The absolute gain is derived from the peak of the gain-corrected single photoelectron charge peak, averaged across all ID PMTs, as illustrated in Figure 2.13.

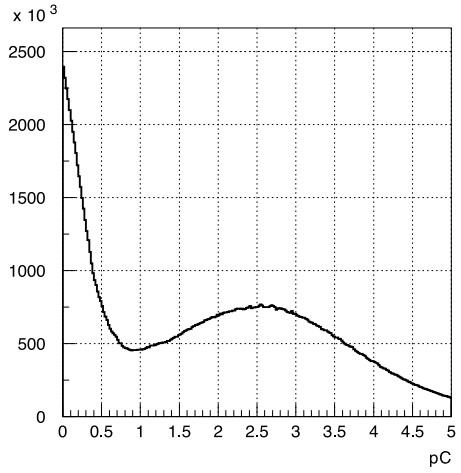


Figure 2.13: An example of combined gain-corrected charge distributions from all ID PMTs, displaying a peak at 2.5 pC. (Reprinted from [48])

Relative QE measurement

The "QE" of a PMT is defined as the product of its quantum efficiency (the conversion ratio of photons to photoelectrons at the photocathode) and the collection efficiency (the ratio of photoelectrons collected at the first dynode). Essentially, this quantifies the probability of a PMT recording a single photon arrival as a hit.

Prior to measurement, the water underwent convection, ensuring temperature variations below 0.01°C to guarantee uniform water transparency throughout the ID volume. Subsequently, the QE was measured using the same Nickel ball used for absolute gain measurement. This measurement is based on the number of hits recorded by individual PMTs, with corrections applied for photon attenuation and the angular acceptance of individual PMTs obtained from MC simulation, as illustrated in Figure 2.14. As the exact number of photon arrivals at each PMT is not known, the results provide relative differences in QE, which are then tabulated for integration into the MC simulation of the detector.

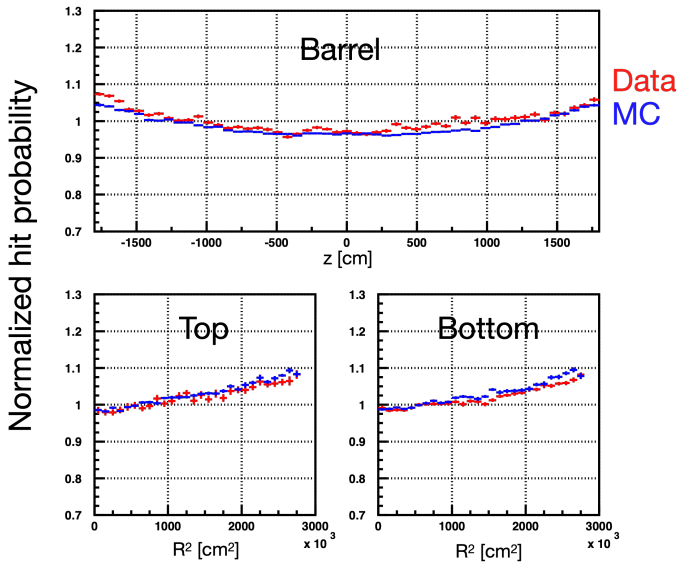


Figure 2.14: The normalized hit probabilities of PMTs, derived from single-photon level light emitted by the Nickel source, in both expected (blue) and observed (red) distributions. These distributions are shown in relation to the vertical positions of PMTs on the tank barrel (top figure) and the radial positions of PMTs on the tank top and bottom walls (bottom two figures). (Adapted from [46])

Charge linearity calibration

The ID PMTs are expected to show approximately 5% non-linearity for signals surpassing 250 photoelectrons. To assess the charge linearity of each PMT's response, a N-laser-driven dye laser connected to a MgO light diffuser ball was used. This setup was inserted into the ID volume at various positions and light intensities. Figure 2.15 shows the calibration configuration as well as the ratio between the observed and expected charge outputs as a function of the expected values. These results are then incorporated into the MC simulation of the detector response to address the apparent non-linearity observed in the PMT charge output response.

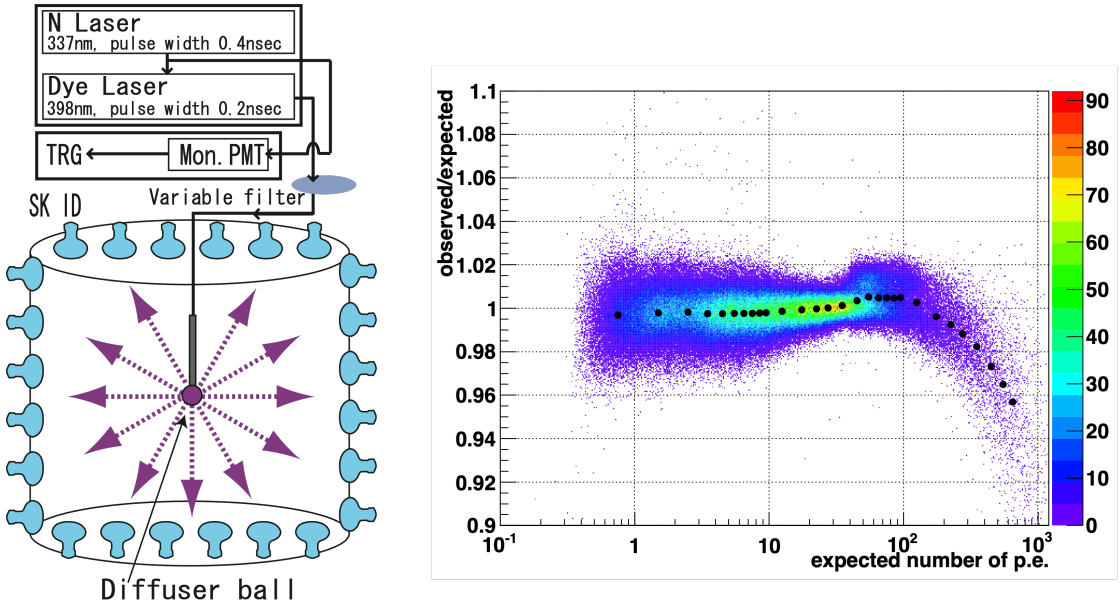


Figure 2.15: Setup featuring the N-dye laser diffuser ball for PMT charge linearity and timing response calibration (left), along with the observed charge linearity plotted against the expected charge output (right).

Timing

The configuration shown in Figure 2.15 is also used for calibrating the PMT timing response, which is dependent on factors of $O(1)$ ns, such as transit time, lengths of PMT signal cables, readout electronics, and the pulse height ("time-walk" effect). The timing of the laser signal is accurately known to the sub-ns level.

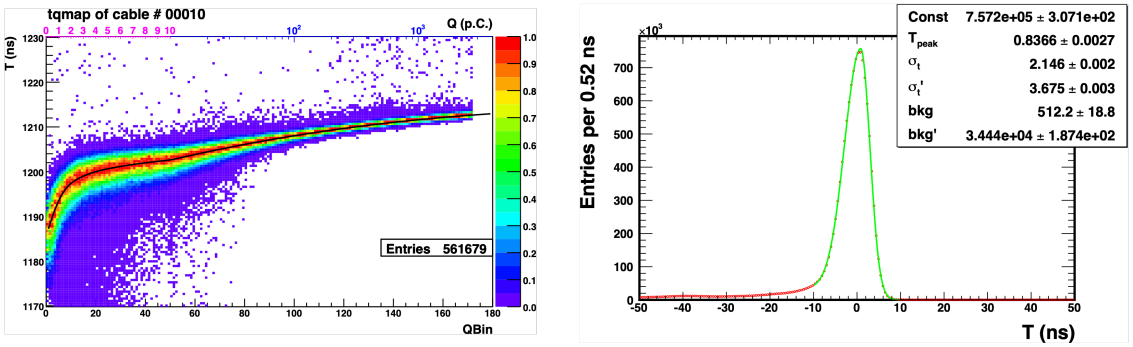


Figure 2.16: A 2D histogram representing the charge and relative timing of hits recorded by a PMT in the laser calibration setup (left), with a specific charge bin slice projected onto the time axis (right). A smaller value of T indicates later timing. The tail on the later timing side is attributed to light reflection and scattering.

Figure 2.16 displays a 2D histogram representing the charge and relative timing of hits recorded by a PMT in response to laser pulses with varying intensity and positions. Notably, the relative timing shows a dependence on signal charge, where a larger charge corresponds to faster timing, as expected due to the time-walk effect. This relationship is modeled using a segmented polynomial function with 15 free parameters, as detailed in [46]. Termed the TQ-map, the fitted function for each individual PMT is used to correct the time response of its respective readout channel. The corrected time distribution for each PMT can be fitted with an asymmetric Gaussian function, as shown in Figure 2.16, to estimate the PMT's timing resolution.

An N-laser light constantly flashes at a frequency of 0.03 Hz, resulting in a daily timing difference for each readout channel with a statistical accuracy of 0.05 ns. The timing response demonstrates a stability within 0.1 ns.

Photon tracking calibration

Photons undergo either scattering or absorption during their propagation in water, with the extent varying based on wavelength. The light scattering and absorption in water is studied using a monochromatic laser beam of different wavelengths, injected from the top of the tank to the bottom, as illustrated in Figure 2.17.

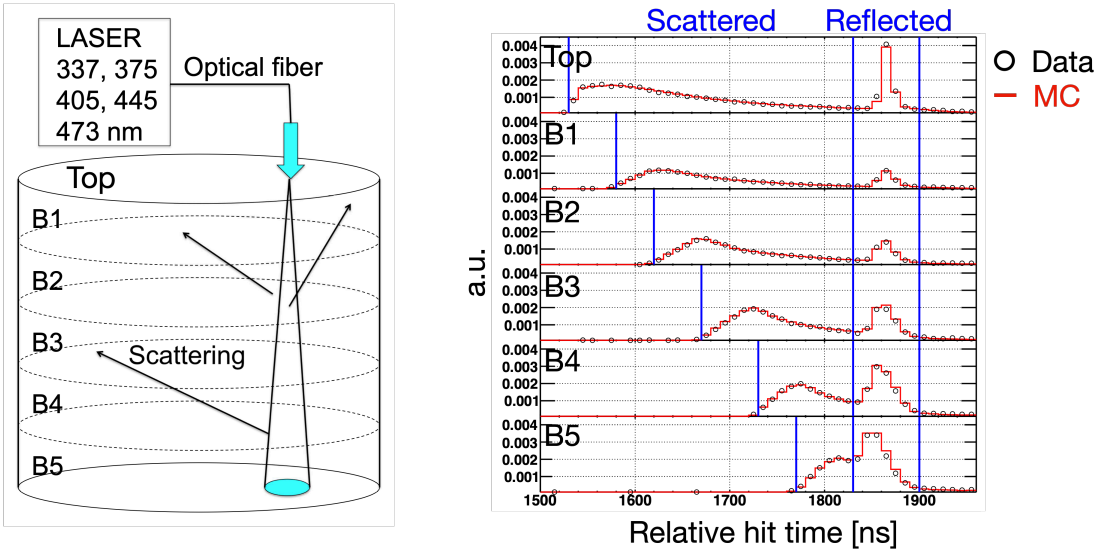


Figure 2.17: A schematic illustrating the photon tracking calibration through monochromatic laser injection (left) and the observed PMT hit time distributions divided into various tank regions (right). In the right figure, the milder peak on the left is attributed to scattered photons, while the sharper peak on the right is a result of photons reflected at the tank bottom. (Adapted from [46])

SK has its own empirical model for photon propagation in water. The model begins with an assumption of exponential light attenuation represented by $e^{-l/L(\lambda)}$, where l is the photon track length, and L is the characteristic attenuation length, also called water transparency. L is influenced by photon absorption or scattering, which is dependent on the photon wavelength λ. In the SK detector simulation, $L(\lambda)$ is defined as:

$$L(\lambda) = \frac{1}{\alpha_{abs}(\lambda) + \alpha_{sym}(\lambda) + \alpha_{asy}(\lambda)} \tag{2.2}$$

Here, α_{abs}, α_{sym}, and α_{asy} represent amplitudes (in m^{-1}) of absorption, symmetric scattering, and asymmetric scattering, respectively. To align with the laser data for each wavelength, we determine the amplitudes through simulation comparisons. The relationship between these amplitudes and the photon wavelength λ is parametrized, as described in [46].

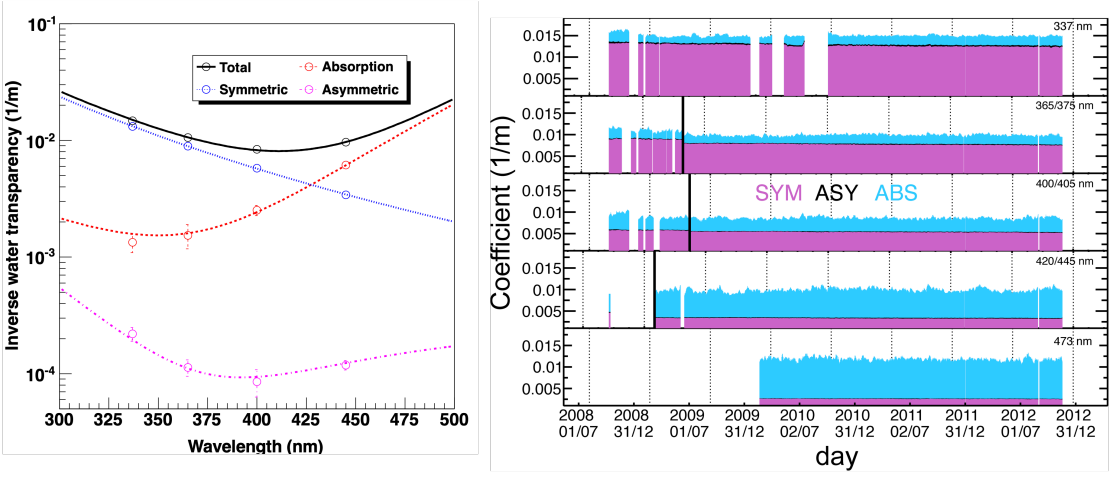


Figure 2.18: Illustration of fitted absorption, symmetric scattering, and asymmetric scattering amplitudes parametrized by photon wavelength (left), along with the temporal variation of these amplitudes (right). (Reprinted from [46])

These parameters, characterizing photon scattering and absorption rates in water, are referred to as "water parameters." Figure 2.18 shows α and L^{-1} as functions of λ with the continuously monitored and fitted water parameters. The fitted amplitudes in Figure 2.18 correspond to a water transparency L of approximately 120 m (for $\lambda = 440$nm).

Additionally, photons can be reflected when encountering a boundary between materials with different refractive indices. In SK simulation, four materials are considered: water, glass (PMT cover), bialkali (PMT photocathode), and vacuum (air). Reflection can occur on the PMT surface or the black sheet covering the tank wall. The complex refractive index of bialkali ($n_{\text{real}} + i \cdot n_{\text{imag}}$) is specifically tuned for various laser wavelengths, aligning with the steeper photon reflection peak in the laser calibration data in Figure 2.17.

Finally, a dedicated measurement of the reflectivity of the black sheet covering the wall was conducted. A laser light injector and a sample of the black sheet were submerged in the tank, and the light reflected off the black sheet sample was directed to the ID PMTs. The reflectivity was adjusted based on the ratio between the direct and reflected charge output with and without the black sheet sample.

Absolute QE measurement

Following the calibration of water parameters in each SK phase, the absolute QE scale (`COREPMT`), which normalizes the overall charge output response of the PMTs, is determined by comparing simulation results with calibration data. Distinct calibration data sets are used for the two analysis groups within SK.

The LOWE group, focusing on solar and astrophysical neutrinos with signal energies in the range of O(1-10) MeV, utilizes a linear accelerator (LINAC). This LINAC generates an electron beam with energies ranging from 5 to 18 MeV. The electron beam, originating from the klystron-based LINAC, is directed straight into the tank through three bending magnets and a beam pipe from the tank top. This setup is depicted in Figure 2.19.

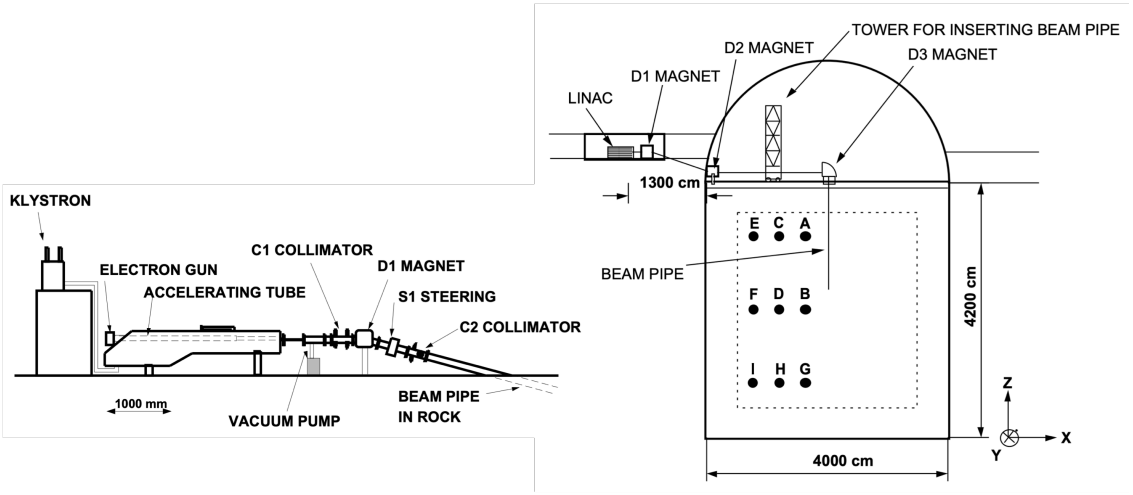


Figure 2.19: Illustration of the LINAC setup (left) and the calibration arrangement (right). (Reprinted from [49])

The ATMPD group, which focuses on GeV-level events like atmospheric neutrino interactions or proton decays, has used through-going cosmic ray muons. These muons traverse both the top and bottom of the detector with relatively consistent path lengths. The absolute QE value that best fits the observed through-going muon charge distribution is selected, as illustrated in Figure 2.20.

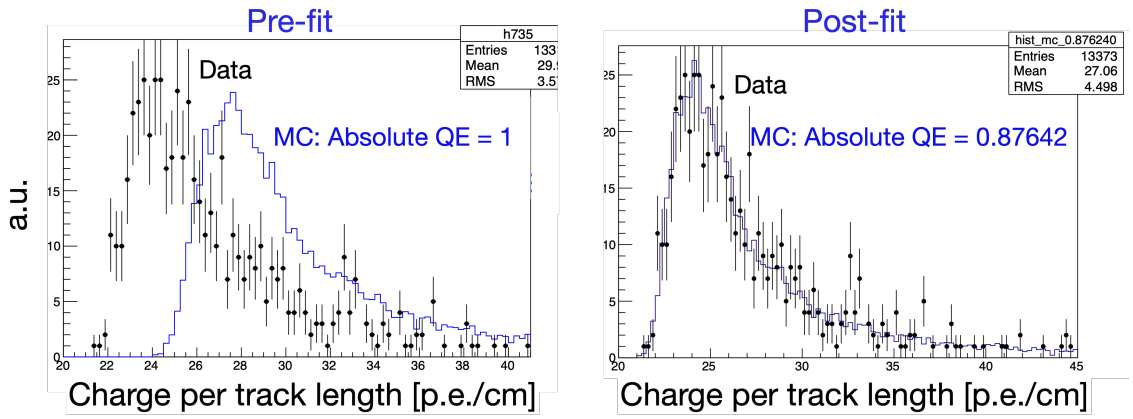


Figure 2.20: An illustration of how the absolute QE is obtained from the observed and expected charge per track length distributions of through-going cosmic muons. The left figure shows the simulated distribution with absolute QE of unity, while the right figure shows the distribution with the fitted value. (Adapted from [50])

The values obtained from two distinct calibration datasets show an agreement within 1%, owing to the high dynamic range and charge linearity of the ID PMTs.

Absolute energy scale calibration

Using a fully tuned simulation, we can reconstruct the momentum of charged particles based on their Cherenkov ring. The detailed reconstruction process itself is outlined in Section 2.4.1. The precision of the PMT charge response simulation significantly impacts the accuracy of the reconstructed momentum. Estimated errors in momentum reconstruction, spanning from $O(10^1\text{-}10^4)$ MeV/c, are determined using natural calibration sources specific to each SK phase. For the lower energy range around $O(10)$ MeV/c, Michel electrons resulting from cosmic-ray muons stopping inside the detector are used. In the 135 MeV range, neutral pion decays originating from atmospheric neutrino neutral-current interactions are used. For the high-energy range, up to 10 GeV/c, fully contained stopping muon events are used. Results up to SK-V demonstrate a reasonable agreement between simulation and data within 2%, as illustrated in Figure 2.21. As of the current writing, calibration for SK-VI has not been completed.

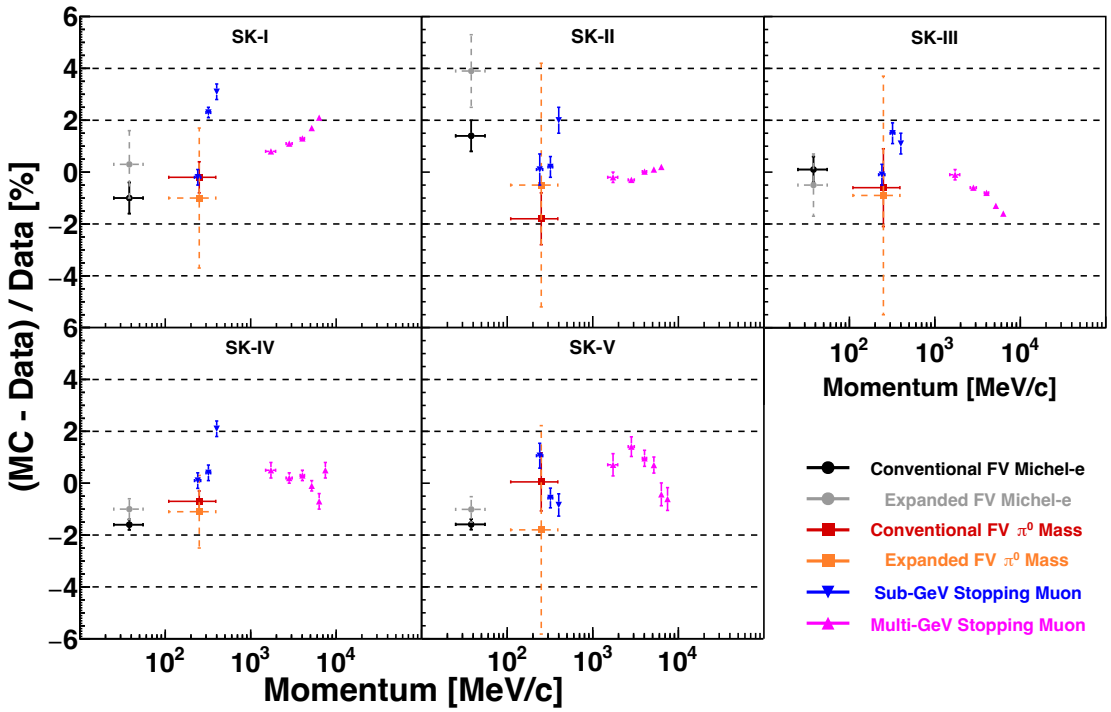


Figure 2.21: Absolute energy scale measurements from SK-I through V. "Conventional FV" refers to events with vertices more than 2 m away from the tank wall, while "Expanded FV" refers to events with vertices between 1 to 2 m from the wall. Both sets of events are utilized for analysis in this thesis. (Reprinted from [48])

Long-term monitoring of detector modeling parameters

The following parameters are monitored on a run-by-run basis, with each normal run typically spanning a full day:

- Water transparency (i.e., photon attenuation length)

  Cosmic ray muons, primarily minimum-ionizing particles with a 2 MeV/cm energy loss in water, serve as a natural calibration source. Their track lengths are directly proportional to the radiated energy, allowing us to discern the relative amount of energy dissipated through photon scattering and absorption in water. Through-going muons with reasonably fixed path lengths are used for monitoring water transparency.

- Dark rate

  The average hit rate per PMT is monitored using off-timing data, primarily consisting of dark noise and minimal background radioactivity.

- PMT relative gain

  The single photoelectron peaks of individual PMTs in off-timing data are utilized to determine the relative gains of each PMT.

- Top-Bottom Asymmetry

  The hit probability shows a vertical asymmetry, quantified by a parameter defined as $(\langle N_{\mathrm{Top}}\rangle - \langle N_{\mathrm{Bottom}}\rangle)/\langle N_{\mathrm{Barrel}}\rangle$, where N_{Top}, N_{Bottom}, and N_{Barrel} are the averages of the hit probabilities of top, bottom, and barrel PMTs, respectively. This rate is monitored using the Xe-lamp-driven scintillator ball that is permanently fixed at the tank center.

Figure 2.22 shows an example of the time variation of these parameters in SK-V.

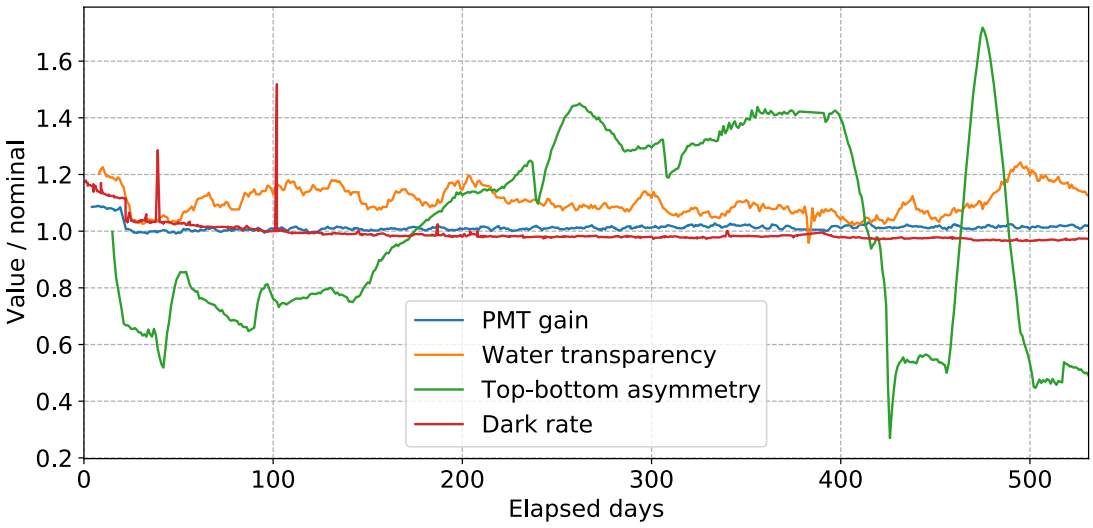


Figure 2.22: Temporal fluctuations (SK-V) in detector modeling parameters, continuously monitored on a run-by-run basis.

2.2.4 Operational phases

SK initiated its operations in 1996 and has been in continuous operation for over 26 years. The first phase, SK-I, experienced an interruption in 2001 due to an accident that damaged 60% of the ID PMTs. Following this setback, SK-II operated with approximately half of the ID PMTs starting in 2002. The successful re-installation of PMTs in 2005 marked the beginning of the third phase, SK-III, in 2006.

A significant electronics upgrade in 2008 marked the start of the fourth phase, SK-IV. This upgrade not only increased data processing speed but also facilitated a transition from hardware to software triggers. As a result, the recorded event length was expanded to 1 ms, allowing the tagging of neutron capture signals with a characteristic time constant of $O(100)$ μs following a neutrino interaction.

After a decade of operation in the fourth phase, the detector underwent comprehensive maintenance in 2018, including draining, cleaning, and overhaul, in preparation for Gd-loading. The fifth phase, SK-V, began in 2019, representing the final phase with an ultra-pure water target.

In July 2020, about 13 tons of Gd-sulfate were loaded into the ID volume to enhance neutron detection efficiency, marking the sixth phase, SK-VI. The Gd concentration of the ID water reached 0.011w% by August 2020.

The seventh phase, SK-VII, started in June 2022 with additional Gd-loading, reaching a Gd concentration of 0.033w% by July 2022. This phase is currently ongoing and actively collecting data.

A summary of the operational phases and their corresponding livetimes for atmospheric neutrino data is provided in Table 2.3. The data analysis in this thesis relies on fully-contained atmospheric neutrino events collected during the SK-IV, SK-V, and SK-VI phases, covering approximately 12 years of detector livetime. SK-I, SK-II, and SK-III data were excluded due to insufficient event lengths for neutron signal searches, while SK-VII was not considered in the thesis scope as it is still ongoing and not yet completed.

Phase	Dates	Livetime [days]	Gd concentration [w%]
SK-I-III	1996-2008	2805.9	-
SK-IV	2008-2018	3244.4	-
SK-V	2019-2020	461.0	-
SK-VI	2020-2022	564.4	0.0110 ± 0.0002 [51]
SK-VII	2022-present	-	0.0332 ± 0.0001

Table 2.3: A summary of SK operational phases. SK-IV, SK-V, and SK-VI data were used in this analysis.

0.01w% Gd-loading for SK-VI

From July 14$^{\text{th}}$ to August 17$^{\text{th}}$, 2020, 13 tons of $Gd_2(SO_4)_3 \cdot 8\,H_2O$ was introduced into the tank using a dedicated dissolving system (refer to Figure 2.24). This quantity corresponds to a Gd-sulfate concentration of 0.026% by weight, dissolved in 50 ktons of pure water within the detector system. The Gd mass fraction is 0.0110w%. Figure 2.23 illustrates the dissolved Gd-sulfate powder and the loaded amounts on each date.

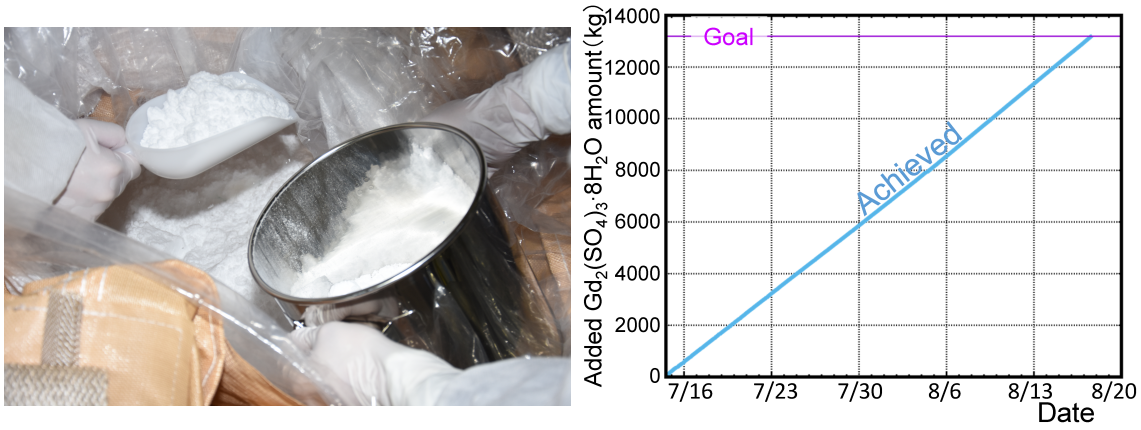


Figure 2.23: Gd-sulfate powder (left), along with the corresponding loading amounts into the detector by date (right). (Adapted from [52]).

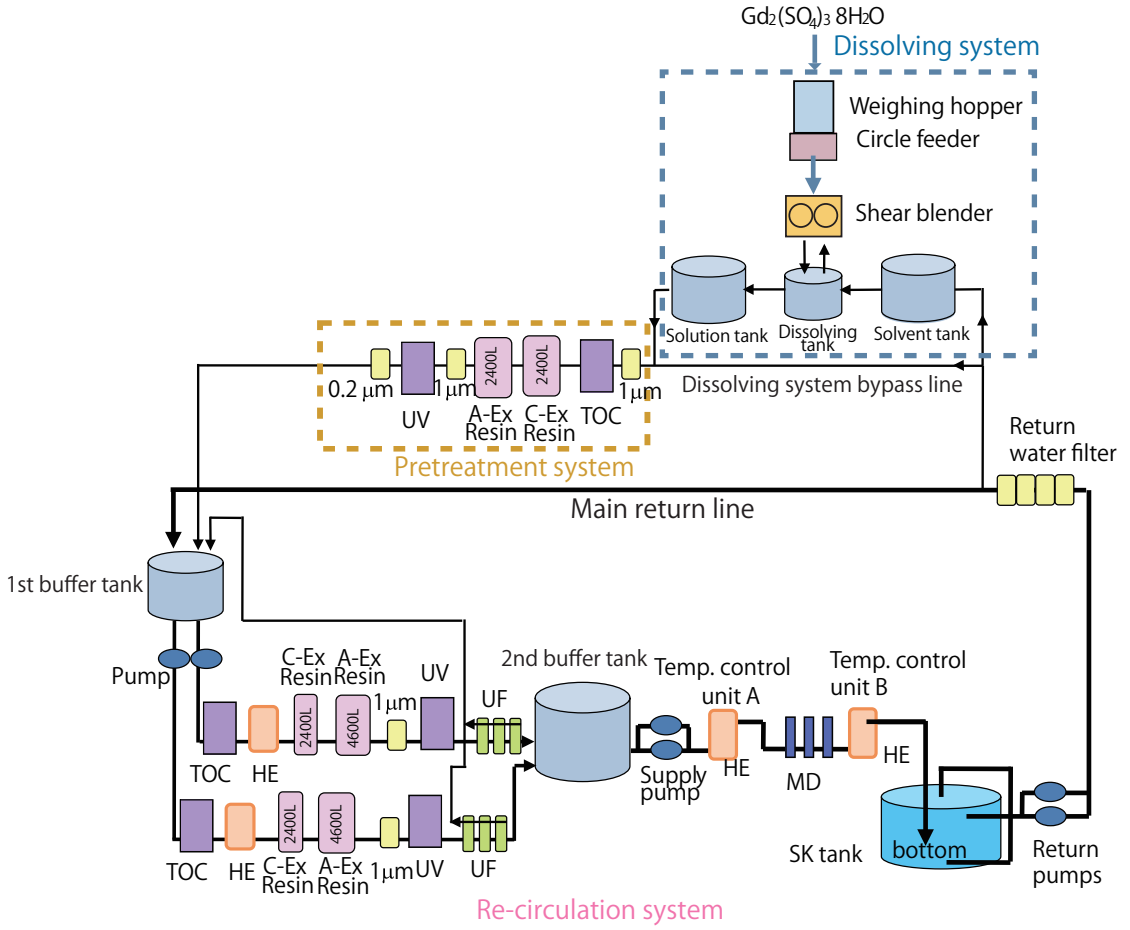


Figure 2.24: A schematic diagram of the SK-VI water circulation system with the Gd-dissolving system. (Reprinted from [51])

The $Gd_2(SO_4)_3 \cdot 8\,H_2O$ powder was dissolved in a small amount of water to create a solution. Before introducing the solution into the water circulation system, it was sent to the "pre-treatment" system to eliminate impurities and prevent bacterial growth. To maximize water purity, both the pre-treatment and water re-circulation systems incorporated cation and anion exchange resins featuring Gd and sulfate ions as ion exchange groups. This effectively removed all ionic impurities in the water, except for Gd^{3+} and SO_4^{2-}. The solution was further diluted with pure water to achieve the desired Gd concentration. The resulting solution was then directed to the SK tank via the "re-circulation" section of the system, which continued to circulate SK water even after the loading was completed. To facilitate a gradual bottom-up filling of the tank, the temperature of the supplied Gd water was set approximately 0.3°C lower than the tank water temperature just before the loading, as illustrated in Figure 2.25.

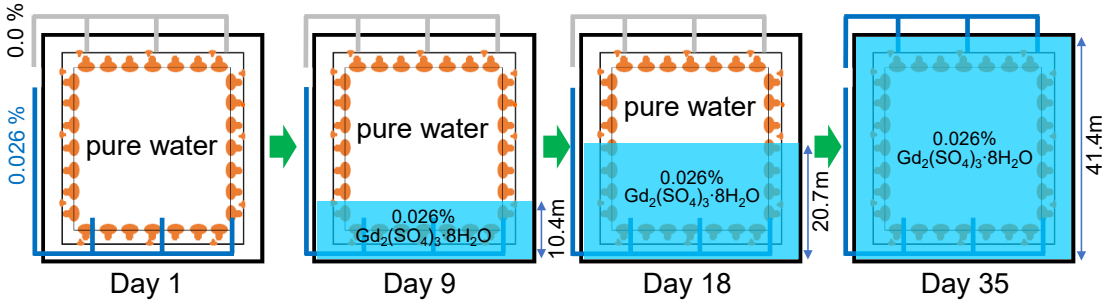


Figure 2.25: Illustration of the Gd-solution gradually ascending from the bottom to the top. (Adapted from [51]).

As the Gd-loaded water was introduced from the bottom, the concentration gradient by height was assessed using Atomic Absorption Spectroscopy (AAS). This technique measures the amount of specific wavelengths of light absorbed by vaporized Gd. The results, illustrated in Figure 2.27, demonstrate consistent uniformity across sampled positions. Simultaneously, water conductivity was measured by extracting samples from various heights in the Gd-loaded tank, as demonstrated in Figure 4.1. The water attenuation length, measured through cosmic ray muon data, decreased to 75 m during Gd-loading and subsequently returned to the baseline level observed in the pure water phases (approximately 90 m), as shown in Figure 2.26.

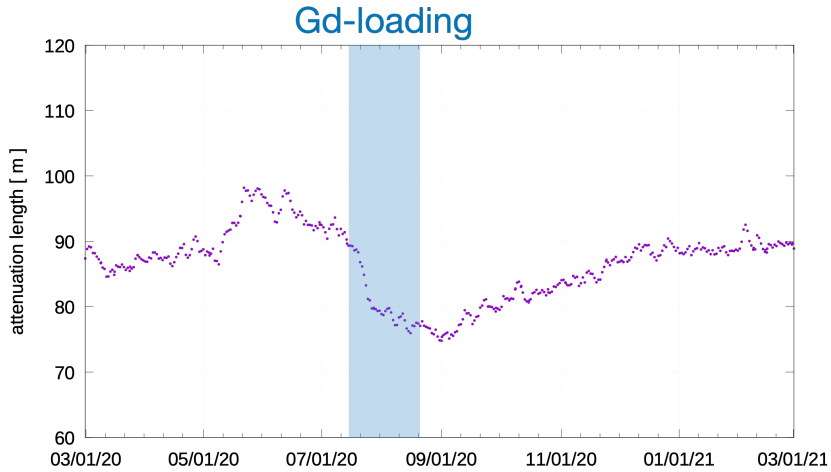


Figure 2.26: Attenuation length of water measured using through-going cosmic ray muons observed towards the end of SK-V and the first six months of SK-VI. (Adapted from [51]).

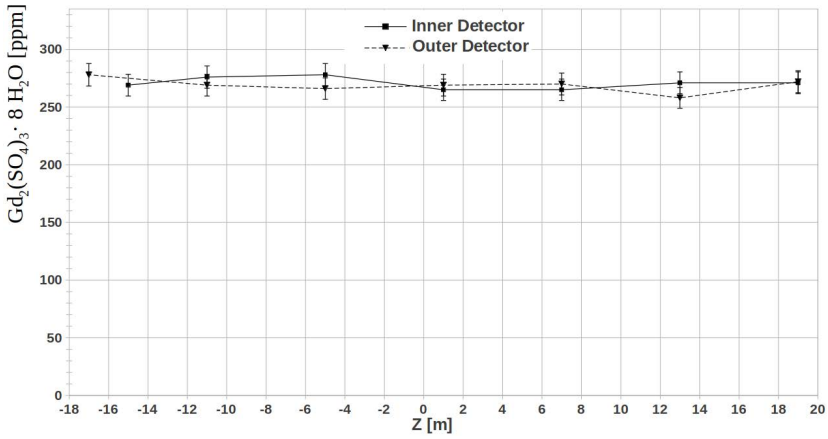


Figure 2.27: $Gd_2(SO_4)_3 \cdot 8\,H_2O$ concentration (260 ppm corresponding to 0.0011w% Gd concentration) in water sampled from various vertical positions, measured with AAS. (Reprinted from [51])

The primary source of background neutrons in the fiducial volume is expected to be the spontaneous fission of ^{238}U impurities in the loaded Gd-sulfate. With effective removal of radioactive impurities, the background neutron rate is expected to be below 600 neutrons per day [53]. This corresponds to an negligibly small $O(10^{-6})$ background neutron contribution per atmospheric neutrino event.

2.3 Event simulation

In this section, we discuss the default models used in SK for simulating atmospheric neutrino interactions and the subsequent hadronic interactions leading to neutron capture signals. We also explore the simulation of detector response to these signals and the methodology used to model backgrounds. Finally, we briefly touch on how we handle the simulation of neutrino oscillation.

2.3.1 Atmospheric neutrino flux model

As briefly outlined in Section 2.1, the primary production mechanism for atmospheric neutrinos is the collision of cosmic rays with nuclei in the atmosphere. Modeling the atmospheric neutrino flux is thus based on the measurement of the primary cosmic ray flux and the secondary hadronic interaction models predicting the outgoing hadron content from collisions between cosmic rays and air nuclei.

Additionally, the geomagnetic field deflects incoming charged particles, potentially preventing low-energy cosmic rays from reaching the atmosphere, as illustrated in Figures 2.28 and 2.29. This geomagnetic cutoff is a crucial consideration in modeling the cosmic ray flux.

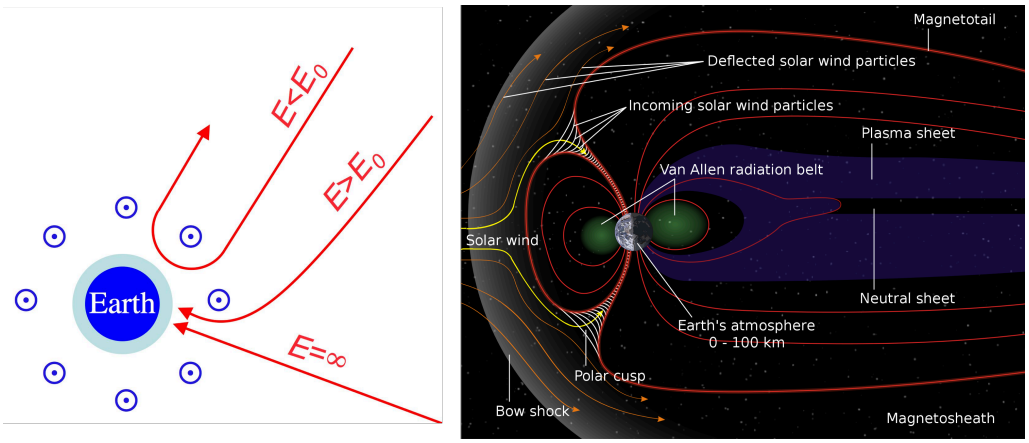


Figure 2.28: Illustration of the geomagnetic cutoff (left, credit: NMDB) and the Earth's magnetosphere blocking solar wind. (right, credit: NASA).

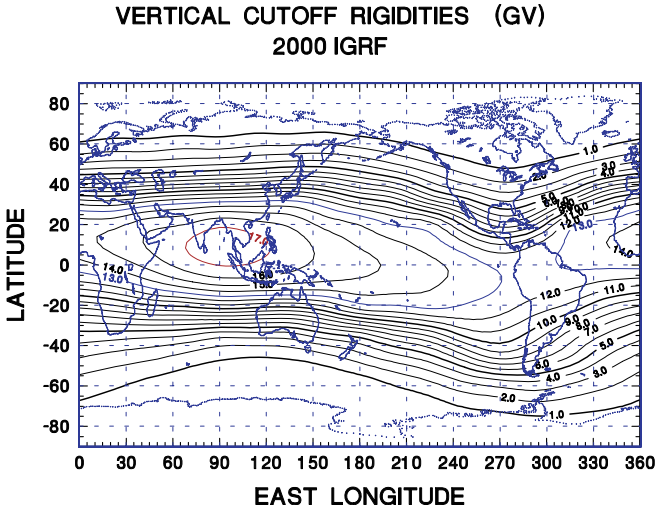


Figure 2.29: The geomagnetic field imposes an energy threshold, also known as cutoff rigidity (expressed in GeV per unit charge), which is higher at the geomagnetic equator and lower at the poles. Due to the gradual change in the geomagnetic field, the rigidity values exhibit seasonal variation. (Adapted from [54])

For this analysis, we used the Honda-Kajita-Kasahara-Midorikawa (HKKM) 2011 model [55] as the default flux model. In this model, both the energy and composition of primary cosmic rays are sampled from distributions based on satellite measurements (AMS [56, 57]) and high-altitude balloon measurements (BESS [58, 59]).

The composition and density of Earth's atmosphere nuclei are modeled using the US-standard '76 model [60], and the geomagnetic field is represented by the International Geomagnetic Reference Field (IGRF) 2005 model [61]. The uncertainty arising from potential time variations in these models is expected to be negligible when compared to other sources of uncertainty.

Hadronic interactions are handled by the `JAM` model, a component of `PHITS` (Particle and Heavy-Ion Transport System, [62]), for $E < 32$ GeV, and a modified `DPMJET-III` [63] for $E > 32$ GeV. Notably, uncertainties in pion and kaon production from these models, particularly in the sub-GeV and above 10 GeV ranges, are considered to carry a substantial (10-15%) uncertainty, as illustrated in Figure 2.30.

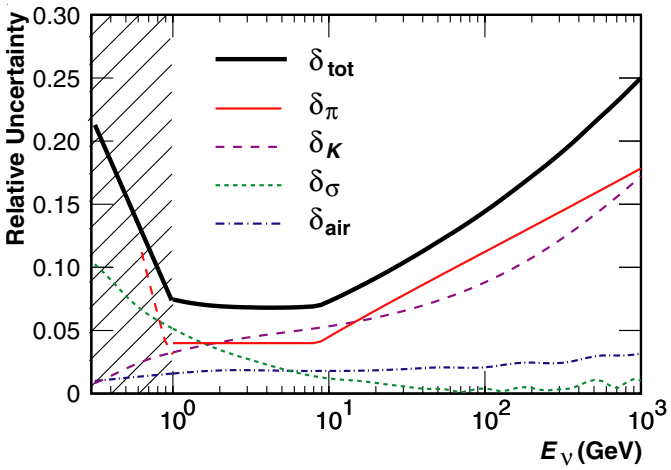


Figure 2.30: Uncertainty estimation in the HKKM 2007 model [64]. The sources of error include uncertainties in pion flux (δ_π), kaon flux (δ_K), mean free path (cross-section) uncertainties (δ_σ), and atmospheric density profile uncertainties (δ_{air}). (Adapted from [64]) In the HKKM 2011 model, updates in the hadronic interaction simulation are anticipated to reduce the uncertainty in the sub-GeV range by approximately 5%.

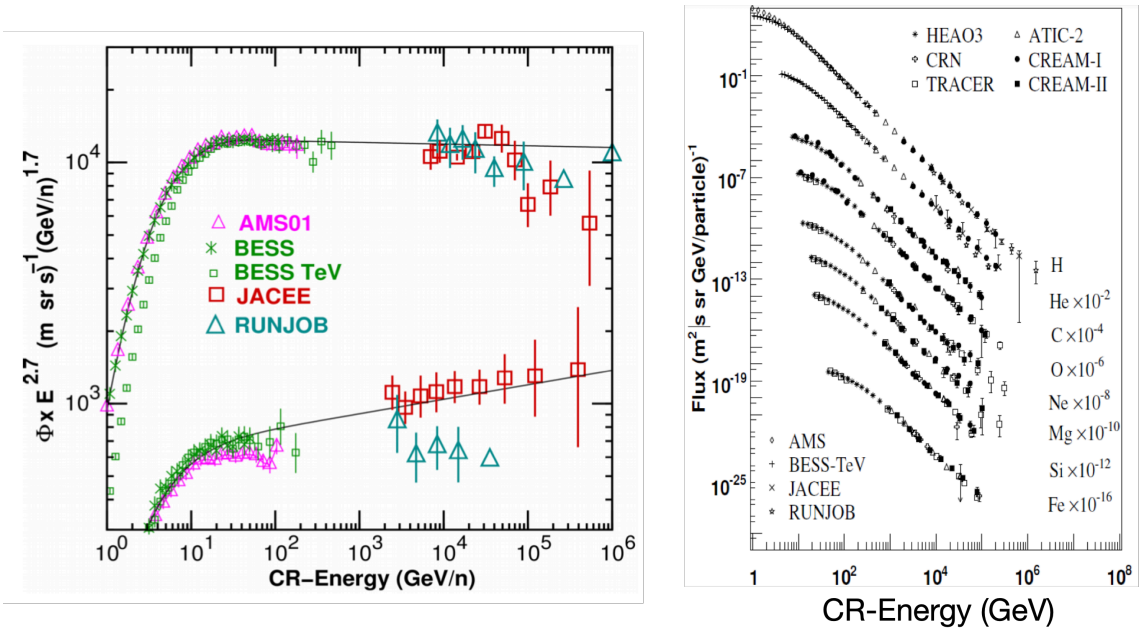


Figure 2.31: Energy distribution of cosmic rays (left) and their composition (right).

Flux calculation starts by tracking cosmic rays. Nuclei and flux are sampled from data-based distributions (see Figure 2.31), positioned randomly on a spherical surface 100 km above Earth. If they pass the geomagnetic cutoff and interact with the atmosphere, trajectories, and subsequent shower particles (pions, kaons, muons, and neutrinos) are tracked in a modeled atmosphere and geomagnetic field in three dimensions. Monitoring continues until particles reach the escape sphere (spherical surface with 10 times Earth's radius) or the Earth's surface. Neutrinos are traced within the Earth to check interactions from below. Overburden and terrain around the detector are considered.

The HKKM 2011 model provides flux predictions for various detectors, including SK. This is accomplished by counting neutrinos ($E_\nu > 100$ MeV) arriving on the Earth's surface within an azimuthal radius of 10° from the detector. The surface for the Kamioka site (SK) is shown in Figure 2.32. The use of such an extensive surface (approximately a 1,000 km radius circle) is essential for efficiently counting incoming neutrinos, as the actual size of the detector is significantly smaller compared to the Earth's dimensions. Any bias introduced by employing such a large counting surface and averaging the flux over a wide area is accounted for. For instance, larger weights are applied to horizontally arriving neutrinos than vertically arriving ones. The anticipated bias is at the 1% level. The uncertainty in the prediction is estimated to be roughly between 10-20%, as shown in Figures 2.33 and 2.34.

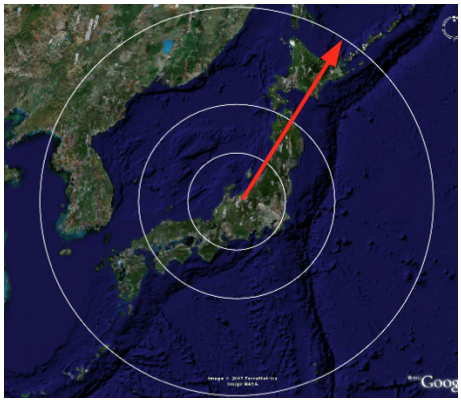


Figure 2.32: The surface designated for neutrino detection, also referred to as the "virtual detector," for Kamioka site (SK). The boundary of the surface is shown as the outermost white circle, with the radius indicated by the red arrow.

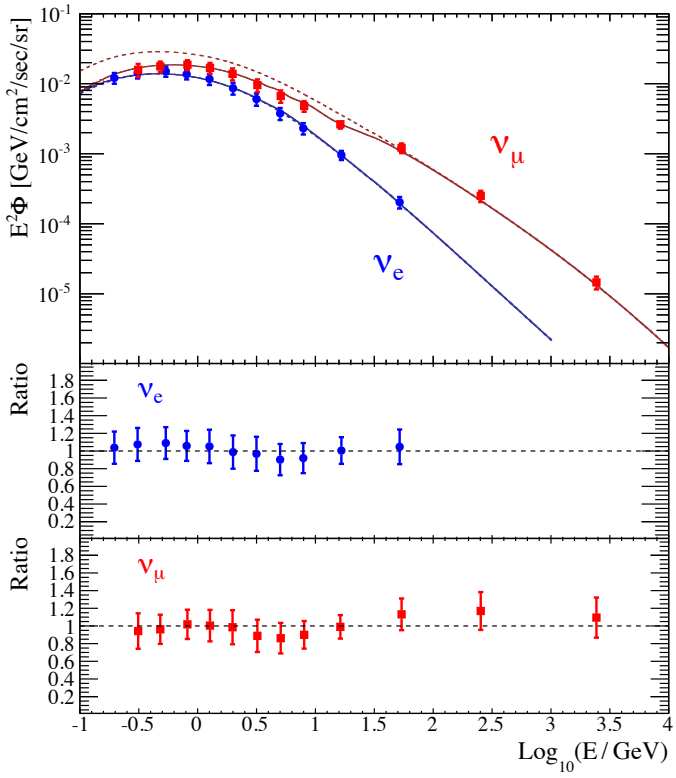


Figure 2.33: Atmospheric neutrino flux prediction at Kamioka by HKKM 2011 (solid lines: oscillated, dashed lines: unoscillated) compared with ν_e (blue) and ν_μ (red) flux measurements at SK [38]. (Adapted from [38]).

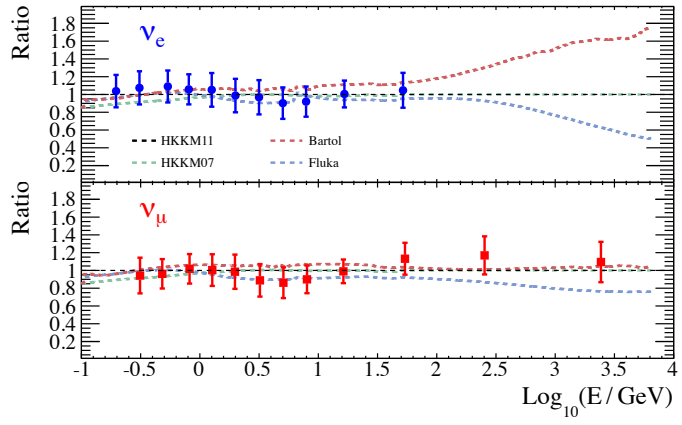


Figure 2.34: Comparison of different flux model predictions at Kamioka (oscillated) with ν_e (blue) and ν_μ (red) flux measurements at SK [38]. The vertical axis represents the ratio to the HKKM 2011 model. (Adapted from [38]).

It is important to recognize that the cosmic ray flux is influenced by solar activity, exhibiting an 11-year periodicity. During periods of maximum solar activity, the solar wind can disperse low-energy cosmic rays, diminishing their flux. The HKKM flux model offers atmospheric neutrino flux as a function of neutrino energy across three solar activity levels: minimum, maximum, and intermediate. The generation of simulated neutrino events assumes the cosmic ray flux at the intermediate solar activity (later adjusted by the livetime-weighted fractions spent in minimum and maximum solar activity, as explained in Section 2.3.7). These fractions are computed based on variations in sunspot numbers and neutron monitor data, which tracks the count of cosmic neutrons on the ground as a proxy for cosmic ray flux. The fluctuations in sunspot numbers and neutron monitor counts are shown in Figure 2.35.

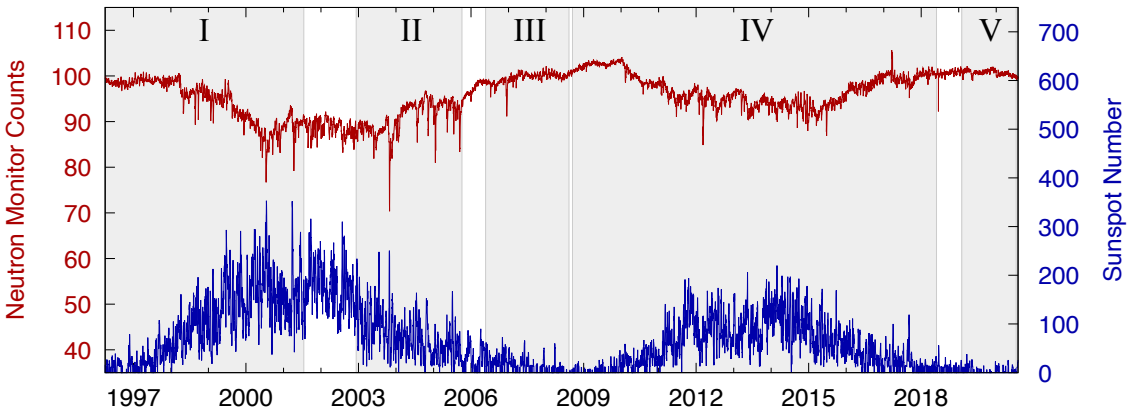


Figure 2.35: Neutron monitor counts [65] during the SK I-V data-taking phases. Each SK phase is represented by a shaded grey region. (Adapted from [48])

The proportions of minimum and maximum solar activity are calculated for each SK operational phase. The impact of solar activity on the overall atmospheric neutrino flux is expected to be negligible (at the ∼5% level, as indicated by cosmic neutron counts in Figure 2.35).

The HKKM 2011 flux model provides ν_e, $\bar{\nu}_e$, ν_μ, and $\bar{\nu}_\mu$ fluxes within the 100 MeV to 1 TeV range, which are considered reliably reconstructable with the detector.

2.3.2 Neutrino-nucleus interaction model

The interactions of neutrinos with the target nucleus in the detector medium are simulated using "neutrino event generators." These generators use theoretical models with parameters adjusted to match external measurements, allowing them to calculate cross-sections for various interaction channels. The simulated event rate is determined by sampling the fractions of these interaction channels based on the calculated cross-section ratios. Neutrino event generators also sample the kinematics of the outgoing final state based on the theoretical models.

In the energy range of 0.1-1 GeV, the dominant interaction is the quasi-elastic (QE) interaction (refer to Figures 2.37 and 2.38), involving the scattering of a bound nucleon. Since fully solving the neutrino-nucleus quantum many-body problem is intractable, event generator assumes that the neutrino interacts with one target nucleon, which is bound in the target nucleus. This approach, known as Impulse Approximation (IA), assumes that the interaction occurs on one nucleon while the other nucleons act as spectators. Subsequent nuclear effects, such as hadronic final-state interaction (FSI), are factorized out (illustrated in Figure 2.36).

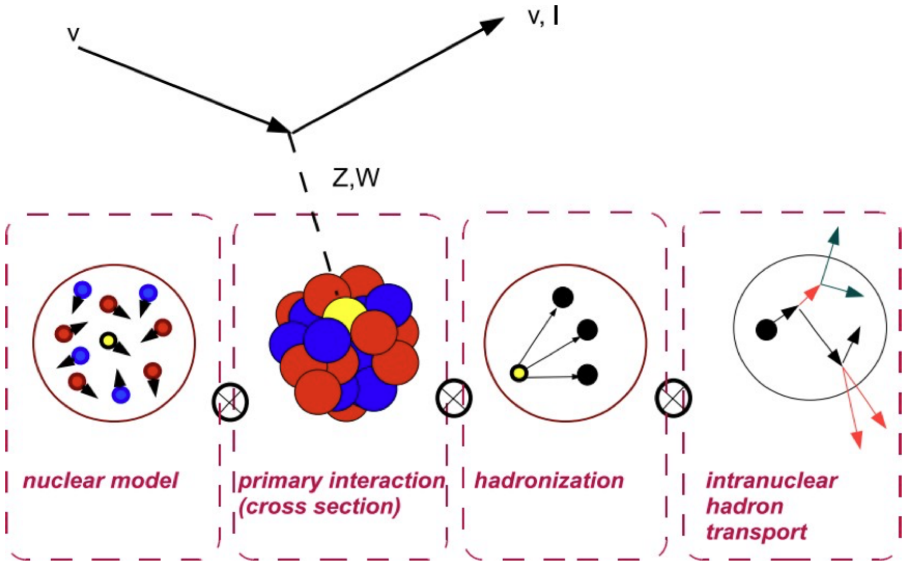


Figure 2.36: Illustration of the factorization process in modeling neutrino-nucleus interactions. (Credit: C. Andreopolous)

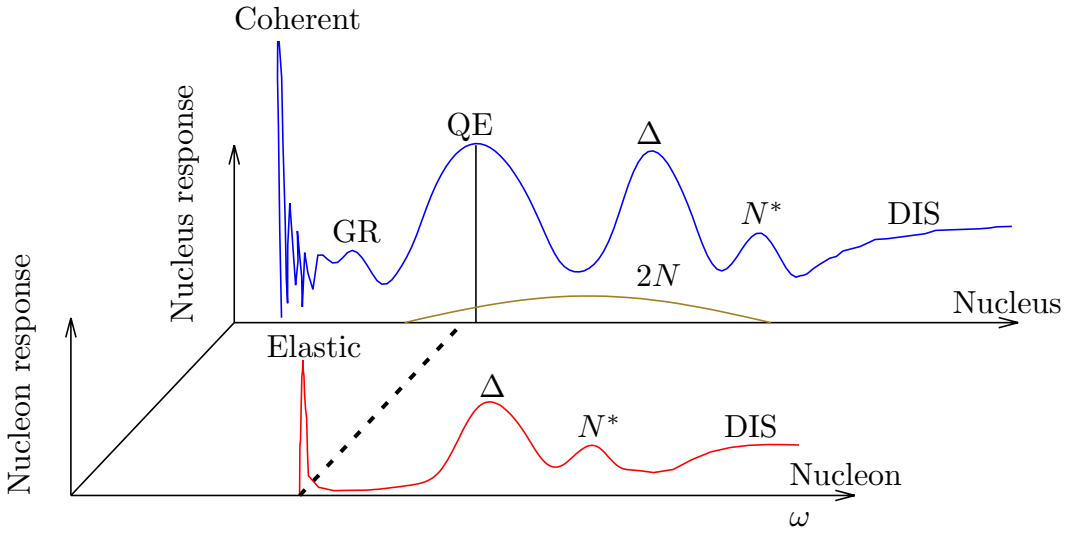


Figure 2.37: Illustration of the nucleon and nucleus responses to a leptonic probe. (Reprinted from [66])

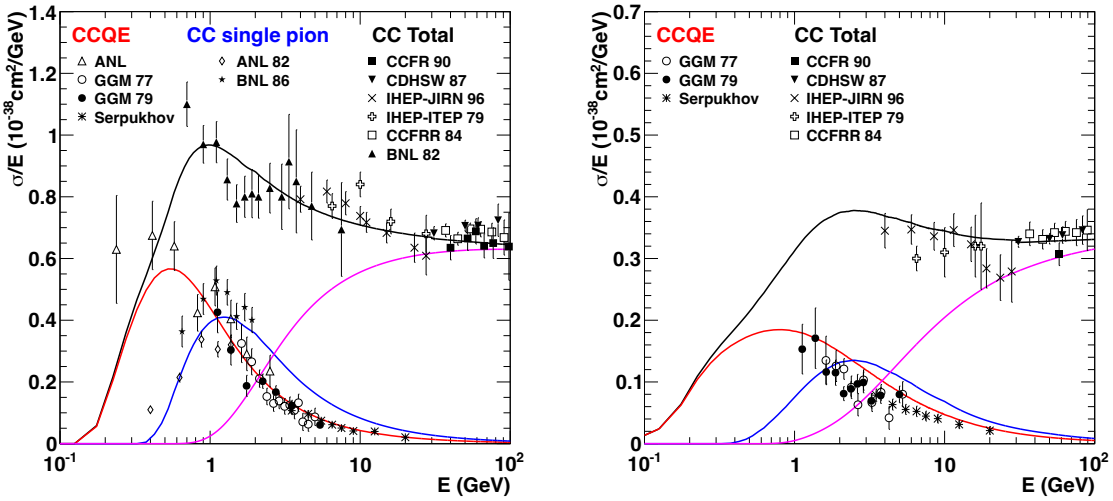


Figure 2.38: ν_μ (left) and $\bar{\nu}_\mu$ (right) cross-sections for CC interactions (per nucleon) normalized by and plotted against neutrino energy, alongside various measurements with error bars. The predictions are produced using the NEUT 5.3.6 generator with an assumed axial mass of $M_A = 1.21$ GeV and the global Fermi gas model for CCQE. (Reprinted from [67])

Nuclear ground state model

The Fermi gas model is used to describe nucleons within a target nucleus. In this model, the nucleus is regarded as a collection of non-interacting nucleons (fermions), each interacting with a general potential field. This approach allows for the straightforward summation of cross-section contributions from individual nucleons. The potential field can be represented as constant across the entire nuclear volume (global Fermi gas model), or it can be modeled based on the nucleon density that varies with position (local Fermi gas model). Neutrons and protons, being distinct particles, experience different forces (protons undergo additional Coulomb repulsion), leading to separate potential wells with distinct energy levels for neutrons and protons. The maximum energy that a nucleon can occupy is determined by the nucleon density and the size of the nucleus, referred to as the Fermi energy.

In the global Fermi gas model, the constant Fermi momentum is set to 225 MeV/c for ^{16}O, while in the local Fermi gas model, the position-dependent Fermi momentum is determined by the nuclear density profile $\rho(r)$ that best explains elastic electron-nucleus scattering data [68].

Figure 2.39 provides an illustration of the Fermi gas model and Fermi momentum as a function of relative position inside the nucleus.

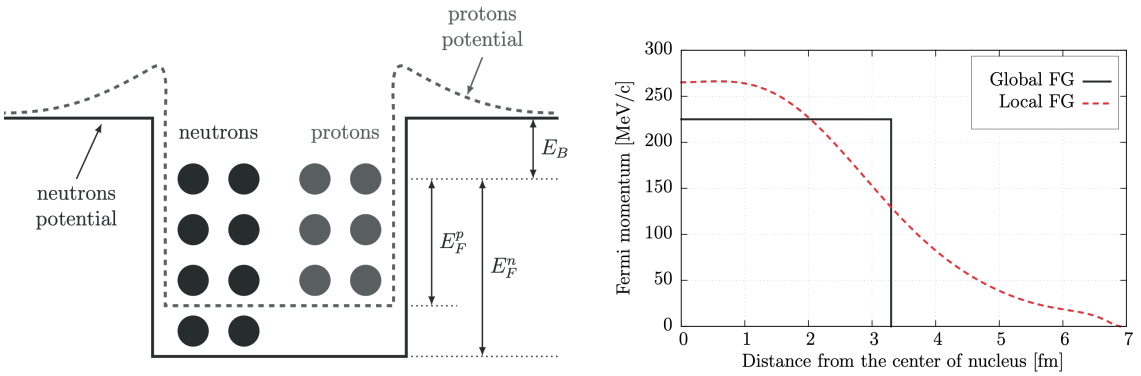


Figure 2.39: Illustration of the Fermi gas model for a nucleus (left) and Fermi momentum (right). $E_F^{p/n}$ represents the Fermi energy of protons/neutrons, and E_B denotes the minimum binding energy. (Adapted from [69])

Momentum of a nucleon undergoing free "Fermi" motion can be sampled from the Fermi momentum distribution. Interactions that result in a final momentum of a struck nucleon smaller than the Fermi momentum are prohibited (Pauli blocking). The ejection of a nucleon is further restricted by the nuclear binding energy, the difference of which from the Fermi energy per nucleon averages around 7-8 MeV. The key feature of the global Fermi gas model lies in the one-to-one correspondence between individual nucleon momentum and the energy required to remove the nucleon, owing to the constant potential depth. This is not the case for the local Fermi gas model, making it more suitable for describing low-energy interactions where nuclear effects are significant.

In `NEUT` 5.4.0, the default event generator for SK atmospheric neutrino simulation, CCQE simulation follows the local Fermi gas model.

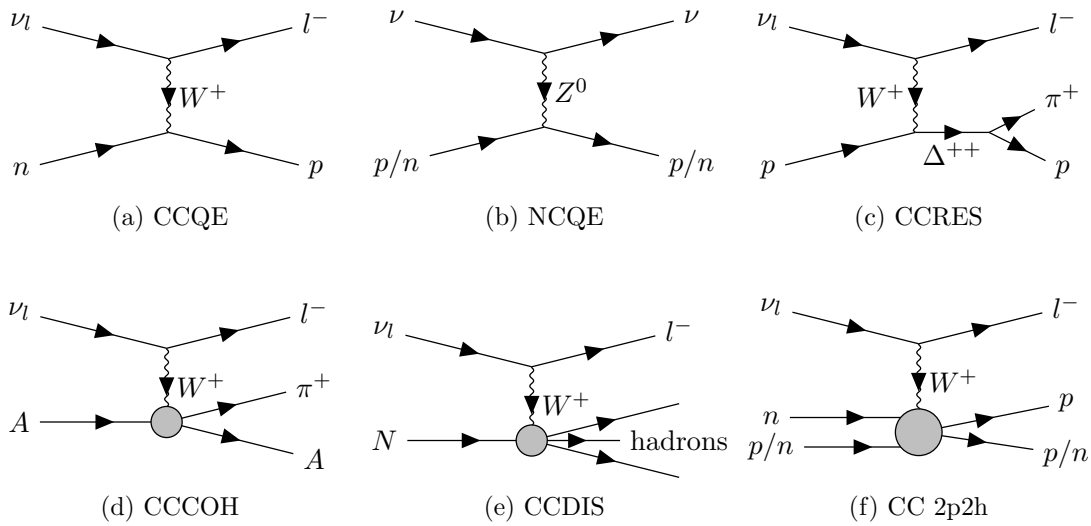
Interaction channels

The weak interaction between neutrinos and quarks occurs through the exchange of a charged $W^{\pm}$ boson or a neutral Z^0 boson. The former is termed Charged-Current (CC) interaction, while the latter is termed Neutral-Current (NC) interaction. CC interaction has an energy threshold equivalent to the lepton mass due to the production of a charged lepton, whereas NC interaction has no energy threshold and merely scatters a nucleon. CC interactions serve as target signals for neutrino oscillation experiments because the produced charged leptons carry the neutrino flavor.

In the energy range of interest (O(0.1-1) GeV) for neutrino event generators in accelerator/atmospheric neutrino oscillation experiments, the dominant interaction channels include quasi-elastic (QE or 1p1h), two-nucleon-correlated quasi-elastic (2N or 2p2h), Δ-resonance (RES), coherent pion production (COH), and deep inelastic scattering (DIS). The corresponding energy scales are illustrated in Figures 2.37 and 2.38. Figure 2.40 displays the Feynman diagrams for these major interaction channels. Both CC and NC interactions can undergo these types of interactions.

In NEUT, these interactions can be categorized based on the number of pions (mesons) produced:

- Quasi-elastic: 1p1h, 2p2h
- Single-pion production: RES, COH
- Multi-pion/nucleon production: DIS



(a) CCQE (b) NCQE (c) CCRES

(d) CCCOH (e) CCDIS (f) CC 2p2h

Figure 2.40: Feynman diagrams of the primary neutrino-nucleon/nucleus interaction channels in the O(0.1-1) GeV range. (Reprinted from [70])

Here, we provide a brief overview of the characteristics of each interaction channel, along with details about the default theoretical model employed in NEUT 5.4.0 for predicting cross sections and final state kinematics of these interactions.

Quasi-elastic (QE)

The differential cross section for neutrino quasi-elastic scattering on a "free" nucleon, per four-momentum transfer q^2, is expressed as [71]:

$$\frac{d\sigma^{\nu(\bar{\nu})}}{d|q^2|} = \frac{G_F^2 M^2 (f^{CC/NC})^2}{8\pi E_\nu^2}\left[A(q^2) \pm B(q^2)\frac{s-u}{M^2} + C(q^2)\frac{(s-u)^2}{M^4}\right] \tag{2.3}$$

Here, M denotes the nucleon mass, G_F is the Fermi coupling constant, $f^{CC} = \cos\theta_c$, where θ_c is the Cabibbo angle, and $f^{NC} = 1$. The variables s and u represent Mandelstam variables, and E_ν is the neutrino energy. The + sign is for neutrinos, and the − sign is for antineutrinos.

The functions $A(q^2)$, $B(q^2)$, and $C(q^2)$ are expressed in terms of vector and axial nucleon form factors. Vector form factors are related to nucleon electromagnetic form factors constrained by electron scattering experiments, while the axial form factor can only be studied through weak interactions with neutrinos, making it challenging to constrain. The axial form factor F_A is approximated in dipole form as:

$$F_A(q^2) = g_A\left(1 - \frac{q^2}{M_A}\right)^{-2} \tag{2.4}$$

Here, $g_A = F_A(0)$ is the axial coupling constant, well-constrained by β-decay measurements, and M_A is the "axial mass," determining the q^2-dependence of the form factor and carrying substantial uncertainty, as illustrated in Figure 2.41. The choice of vector form factor parametrization and the value of axial mass significantly influence the overall q^2-dependence of the inclusive QE cross sections. For instance, larger values of M_A are expected to increase both the total and differential cross sections with larger q^2 (Figure 2.41).

For CCQE interactions, NEUT 5.4.0 utilizes the Local Fermi Gas (LFG) model for the nuclear ground state supplemented by random phase approximation (RPA) correction, accounting for the long-range weak charge screening effect from other nucleons (not applicable to free nucleons). By default, NEUT 5.4.0 follows the so-called "Valencia model" [72] (provided by the group at IFIC, University of Valencia), naturally extending to describe neutrino interaction on strongly-correlated nucleon pairs, potentially knocking two nucleons (2p2h) instead of one (1p1h). The axial mass M_A is set to 1.05 GeV, and the BBBA05 [73] vector form factor parametrization is employed.

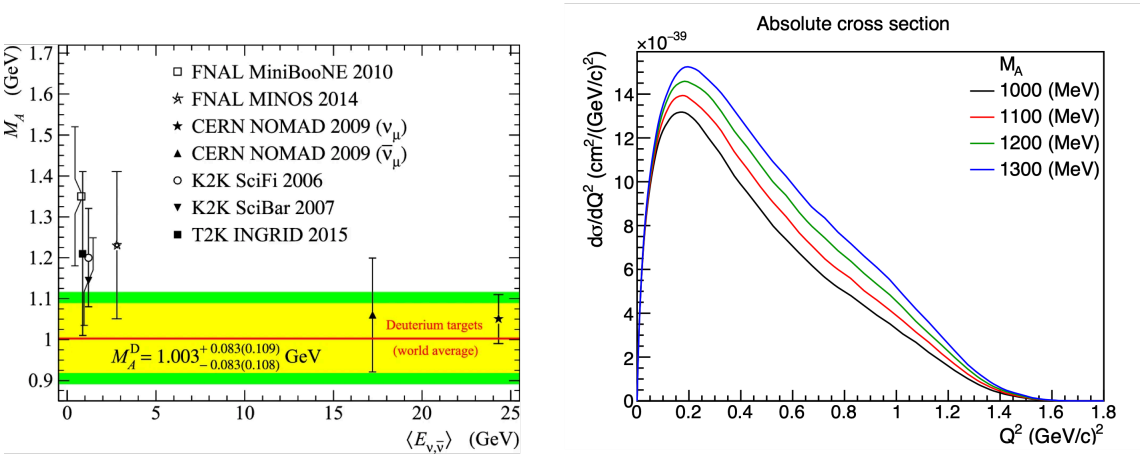


Figure 2.41: The measured axial mass values as a function of mean neutrino energy used in each measurement (left, reprinted from [74]) and the predicted differential cross section per target neutron for 1 GeV ν_μ-^{12}C CCQE interaction with various values of the axial mass M_A (right, reprinted from [70]). There is a slight tension between the axial mass measured with deuterium targets and measurements with composite nucleus targets.

Single-pion production (RES, COH)

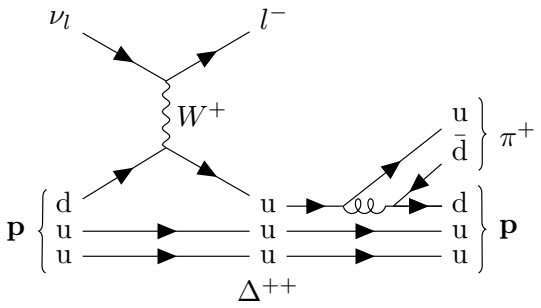


Figure 2.42: Feynman diagram illustrating the CC RES interaction producing a single π^+ through the Δ^{++} resonance (Reprinted from [48]).

An additional energy transfer of 300 MeV from the QE energy peak represents the resonance energy for nucleons excited to Δ-baryons with a mass of 1.2 GeV, which is the lowest mass resonance above the pion production threshold. Figure 2.42 provides an example Feynman diagram illustrating the Δ^{++} resonance. `NEUT` calculates the cross section for producing a baryon resonance, taking into account interference among multiple possible resonance states up to $W < 2$ GeV. Next, `NEUT` 5.4.0, by default, follows the formalism provided by [75], with corrections from [76]. This model calculates the differential cross sections as a product of the amplitude of each resonance production and the probability of the baryon resonance decay to pions. The differential cross section is given by:

$$\frac{d^2\sigma}{dq^2dW} = \frac{1}{32\pi ME_\nu^2} \cdot \frac{1}{2} \sum_{j,\text{spin}} |T(\nu N \to l N_j^*)|^2 \frac{1}{2\pi} \frac{\Gamma}{(W - M_j) + \Gamma^2/4} \tag{2.5}$$

where M is the mass of the target nucleon, W is the invariant mass of the hadronic system, M_j is the mass of the baryon resonance N_j^*, Γ is the decay width of the resonance. The first term in the summation computes amplitudes of intermediate baryon resonances, and the second term computes decays of the excited baryon into pions, considering branching fractions and the kinematics.

`NEUT` considers 28 baryon resonances in the $W < 2$ GeV range and extends the method to calculate single K, η, γ, Λ production cross sections as well.

Additionally, NEUT 5.4.0 adopts the Graczyk-Sobczyk formulation [77, 78] of form factors by default, with a dipole axial form factor $C_5^A(q^2) = C_5^A(0)(1 - q^2/M_A^2)^{-2}$. The parameters are set to $C_5^A(0) = 1.01$ and axial mass $M_A = 0.95$ GeV, tuned based on $\nu_\mu p \to \mu^- p \pi^+$ cross-section measurements with bubble chamber data [79].

Another interaction channel that can produce a single pion is through coherent interaction with the entire nucleus, represented by CC and NC coherent pion production. For CC:

$$\nu_l + A \to l^- + \pi^+ + A \tag{2.6}$$

$$\bar{\nu}_l + A \to l^+ + \pi^- + A \tag{2.7}$$

and for NC:

$$\nu(\bar{\nu}) + A \to \nu(\bar{\nu}) + \pi^0 + A \tag{2.8}$$

Coherent pion production occurs with low q^2, and little momentum is transferred to the nucleus, resulting in forward-scattered outgoing leptons and pions. Coherent pion production is expected to be very small compared to resonant pion production.

In NEUT, the Berger-Sehgal model [80] is used. The energy ranges were determined to qualitatively explain the recent measurement of coherent pion production [81].

Multi-pion/nucleon production (DIS)

DIS is the dominant interaction process for multi-GeV neutrinos. DIS interactions are characterized by a neutrino interacting at the quark scale, breaking apart the nucleon, and resulting in the production of multiple hadrons in the final state.

The calculation of the CC DIS cross section for the range $1.3 < W$ GeV involves integrating the following double-differential cross section (based on the formalism introduced in [82]):

$$\frac{d^2\sigma}{dxdy} = \frac{G_F^2 M E_\nu}{\pi}\left[\left(1 - y + \frac{y^2}{2} + C_1\right)F_2(x) \pm y\left(1 - \frac{y}{2} + C_2\right)xF_3(x)\right] \tag{2.9}$$

$$C_1 = \frac{m^2(y-2)}{4ME_\nu x} - \frac{Mxy}{2E_\nu} - \frac{m^2}{4E_\nu^2} \tag{2.10}$$

$$C_2 = -\frac{m^2}{4ME_\nu x} \tag{2.11}$$

Here, $x = -q^2/(2M\Delta E)$ and $y = \Delta E/E_\nu$ are Bjorken parameters where ΔE represents the difference between the incoming neutrino energy E_ν and the outgoing charged lepton energy E_l, M is the nucleon mass, and m is the lepton mass.

Three structure functions that are not shown are incorporated into F_2 and F_3 with modified structure function correlations based on Callan-Gross [83] and Albright-Jarlskog [84] relations. The calculation of F_2 and xF_3 relies on the GRV98_LO [85] parton distribution function (PDF) with Bodek-Yang correction [86, 87], which offers improved accuracy in the low-q^2 region of lepton inelastic scattering data.

The computed cross-section is normalized by the fraction of multi-pion events (two or more pions). This normalization is applied to exclude the contribution from resonant single-pion production, as explained earlier.

Using the sampled W, the number of produced hadrons (either a nucleon or a pion) is determined from a hadron multiplicity model. This model is a linear function of $\log(W^2)$ and is tuned to deuterium bubble chamber data [88] for accuracy.

For $W > 2$ GeV, PYTHIA v5.72 [89] is used. Similar to low-W DIS simulation, the overlap with single-pion production is excluded. The full event is simulated with PYTHIA based on the sampled W, and, unlike the low-W case, the production of heavier mesons such as K and η is permitted.

Table 2.4 summarizes the default models and options used for major channels of nucleon-level neutrino (CC) interactions in NEUT 5.4.0. Figures 2.44 and 2.45 show the fraction of simulated interaction channels in the atmospheric neutrino event simulation, plotted against the reconstructed event visible energy.

Channel	Cross section	Option	Value
CCQE	Valencia [72]	M_A	1.05 GeV
		Vector form factor	BBBA05
		Nuclear model	Local Fermi Gas
RES	Berger-Sehgal [76]	Form factor	Graczyk-Sobczyk (dipole) [78]
		M_A	0.95 GeV
		$C_5^A(0)$	1.01
		Max. W limit	2.0 GeV
DIS	Kretzer-Reno [82]	PDF	GRV98 + Bodek-Yang [85, 87]
		Low-W model	Custom model [88]
		High-W model	PYTHIA v5.72 [89]

Table 2.4: Overview of the default models and settings applied to the calculation of cross-sections for major neutrino-nucleon CC interaction channels in NEUT 5.4.0.

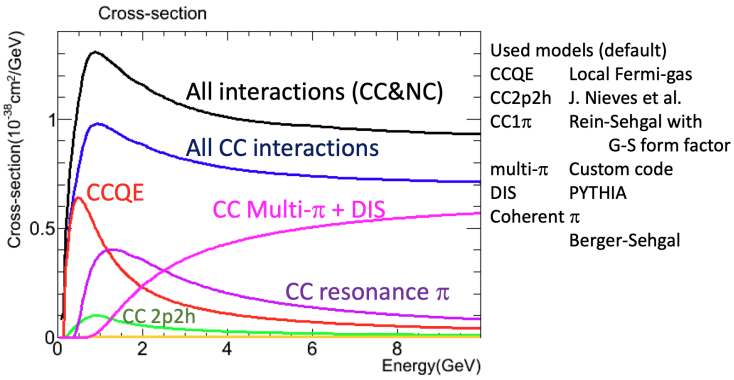


Figure 2.43: NEUT 5.4.0 default ν-^{12}C cross section averaged per nucleon (i.e,. $\{Z\sigma(\nu p) + (A-Z)\sigma(\nu n)\}/A$.) (Reprinted from [90])

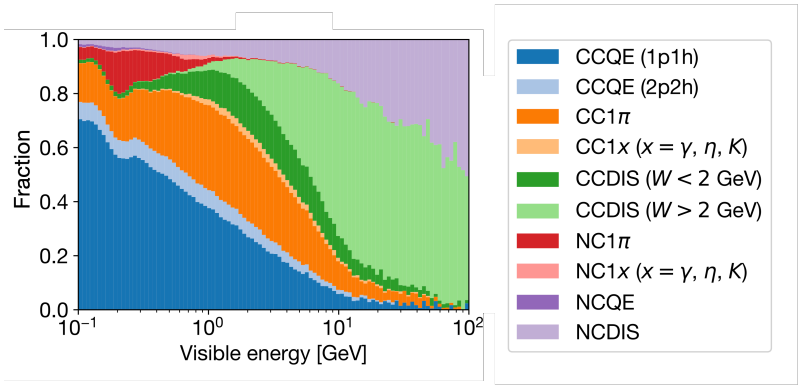


Figure 2.44: MC-simulated atmospheric neutrino interaction type with NEUT 5.4.0.

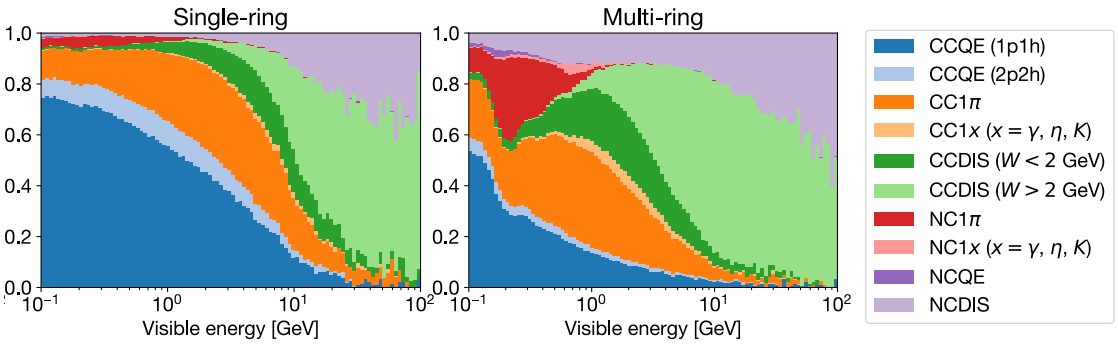


Figure 2.45: Distribution of MC-simulated atmospheric neutrino interaction types using NEUT 5.4.0, categorized by the number of reconstructed Cherenkov rings.

2.3.3 Hadronic interaction model

Through the neutrino-nucleon interaction described in the preceding subsection, various types of hadrons are generated. The most crucial ones for neutrino experiments to reconstruct are the recoil nucleons from CCQE scattering. Additionally, long-lived mesons, particularly charged pions, are abundantly produced in multi-GeV events. The energy of these hadrons can range from a few MeV to several GeV, depending on the energy of the incoming neutrinos. The transport of these hadrons within the target nucleus needs to be accounted for separately. The number of outgoing hadrons from the target nucleus, as well as their charge and kinematics, may differ from the expectations based on nucleon-level model predictions due to these nuclear effects. In the current suite of neutrino event generators, these effects are factored out and referred to as "final-state interaction" (FSI).

It is also essential to consider hadronic interactions within the detector target volume, which, in our case, is water. Particularly significant are inelastic hadron interactions on ^{16}O that can alter the hadron content. This type of hadronic interaction within the detector is often termed "secondary interaction" (SI). The schematic representation of the FSI and SI effects on the detectable hadron content is illustrated in Figure 2.46.

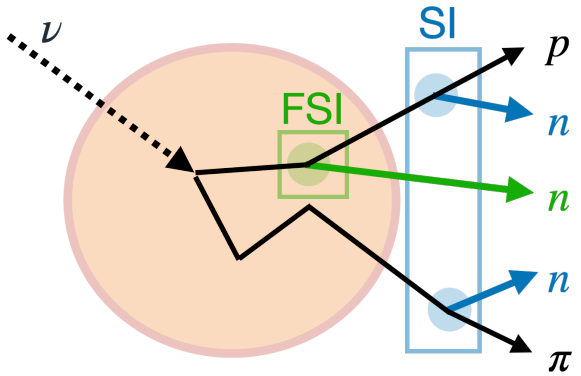


Figure 2.46: Illustration of the production of hadrons via neutrino-nucleon level scattering and the potential change of detectable hadron content due to FSI (green) within the target nucleus and SI (blue) within the detector volume.

As outlined in Section 1.7, these effects impact the reconstruction of assumed nucleon-level neutrino interactions. For instance, events reconstructed as QE-like, may actually include pion-producing events with pions absorbed through FSI/SI processes, introducing bias to our reconstruction. Similarly, neutrons intended for detection in SK to enhance neutrino reconstruction may experience "smearing" effects due to these re-scatterings. Therefore, it is crucial to simulate these effects with high accuracy.

Intranuclear cascade (INC) model

Most neutrino event generators and detector simulators used in neutrino oscillation experiments rely on"intranuclear cascade" (INC) models to simulate these hadronic re-scatterings. In the INC model of hadron-nucleus scattering, the nucleus is represented using the local Fermi gas model as an assembly of non-interacting free nucleons, with a continuous nuclear density function $\rho(r)$. The overall outcome of the hadron-nucleus scattering is approximated by the convolution of a "cascade" involving binary (two-body) collisions between a projectile (either the initial incoming hadron or subsequent recoil hadrons) and a nucleon. The schematic is illustrated in Figure 2.47. Theoretically, it has been shown that summing over the INC calculations corresponds to a solution to a Boltzmann transport equation [91]. In the energy range of O(0.1-1) GeV, which aligns with the size of nucleons, this model is expected to perform effectively, similar to the factorization of nucleon-level interactions in considering neutrino-nucleus interaction.

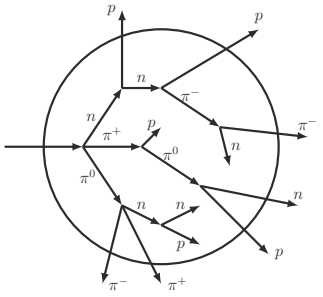


Figure 2.47: Schematic representation of the (space-like) INC model MC computation for hadron-nucleus interaction. (Reprinted from [69])

In INC models, particles are treated classically, moving in straight lines within a nucleus described by the Fermi gas model, as illustrated in Figure 2.47. Two approaches are used to MC-simulate the cascades of binary collisions:

- **Space-like approach**

  The discrete collision points of each projectile are sampled from the probability density function $P(\lambda) = e^{-\lambda/\langle\lambda\rangle}$, with the mean free path $\langle\lambda\rangle$ estimated as:

  $$\langle\lambda\rangle = \frac{1}{\rho\langle\sigma\rangle} = \frac{1}{\rho}\frac{A}{Z\sigma_p + (A-Z)\sigma_n} \tag{2.12}$$

  Here, $\langle\sigma\rangle$ is the averaged nucleon scattering cross section based on the atomic and mass numbers Z and A, and projectile-to-proton/neutron cross sections are denoted as $\sigma_{p/n}$. Cross sections are often either directly extracted from available external data or parametrized as a function of incident particle speed to fit external data. The iterations of collision steps continue until all hadrons exit the nucleus, determined by a pre-determined radius.

- **Time-like approach**

  In this approach, both the incoming projectile and each nucleon in the nucleus are assigned a position and momentum. Initial nucleon positions are sampled from the nuclear density $\rho(r)$. All particles, both projectiles and target nucleons, are propagated until any two of them get closer than the characteristic distance $d_{\text{int}} = \sqrt{\langle\sigma\rangle/\pi}$. The iterations of collision steps stop at a specific time, regardless of particle positions.

Final-state interaction (FSI) modeling

The modeling is separated based on whether the incident particle is a pion or a nucleon. For pions, four channels — scattering, absorption, production, and charge exchange — are considered. For low-momentum ($p < 500$ MeV/c) pions, cross-section calculations are based on [92], with outgoing pion kinematics modeled by [93, 94]. For nucleons and higher momentum pions, `NEUT` uses Bertini's implementation [95] of space-like INC model. The nuclear density profile is described by a Woods-Saxon type distribution for radial displacement r from the nucleus center:

$$\rho(r) = \frac{\rho_0}{1 + \exp\left(\frac{r - R_0}{a}\right)} \tag{2.13}$$

Here, R_0 represents the nucleus radius, and a is a parameter determining the "diffuseness" of the nuclear surface. For ^{16}O in `NEUT`, specific values [96] are assigned: $R_0 = 2.69$ fm, $a = 0.40961$ fm, and $\rho_0 \sim 10^{-3}$.

Higher momentum pion cross-sections are extracted from $\pi^{\pm}$-p scattering data, and pion multiplicities are tuned using bubble chamber data [97]. Nucleon multiplicity after internal pion absorption is adjusted using external π^+ measurements [98] under the assumption of isospin symmetry. Kinematics for two-body nucleon ejection are determined based on [99], while multi-body ejection kinematics are randomly sampled within the allowed phase space. Internal parameters governing pion interaction strengths are tuned based on pion scattering measurements [100], as shown in Figure 2.48.

For nucleon re-scattering within the target nucleus, `NEUT` considers elastic scattering, single and double pion production processes. The kinematics of the products are determined by the INC model, whose input nucleon-nucleon cross sections are based on [101]. Limited nucleon-^{16}O cross-section measurements, especially in the energy range of O(0.1-1) GeV, make it challenging to precisely characterize the nucleon scattering model. The uncertainty of the nucleon-nucleon cross section [101] for the O(0.1-1) GeV range is estimated to be 30% [34] based on `NEUT`-predicted cross sections for p-^{12}C compared with data [102].

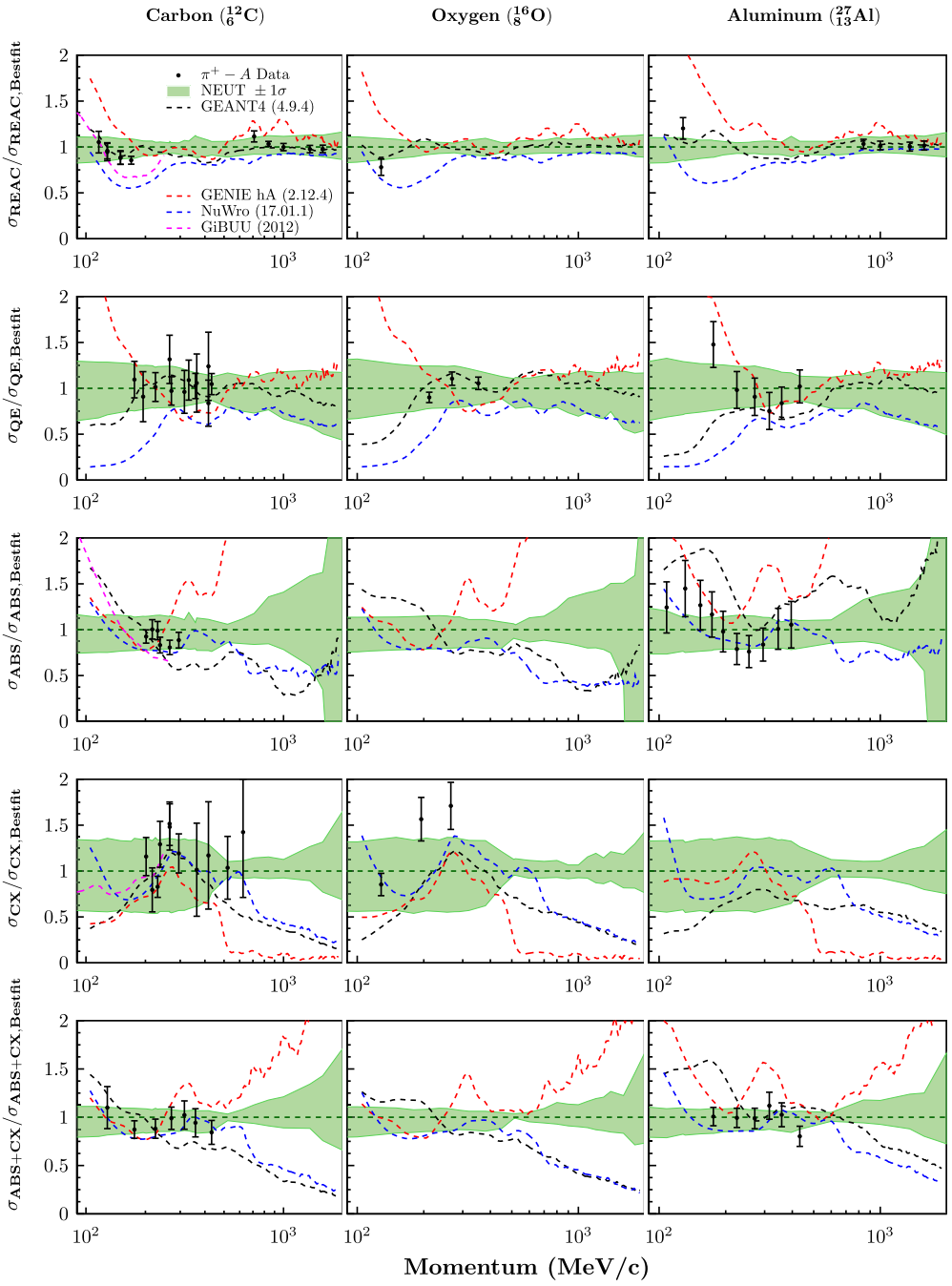


Figure 2.48: π^+-A cross section ratios between the NEUT 5.4.0 default and external data and other models. $\sigma_{\mathbf{REAC}}$ is the total reaction cross section, $\sigma_{\mathbf{QE}}$ is the quasi-elastic cross section, $\sigma_{\mathbf{ABS}}$ is the absorption cross section, $\sigma_{\mathbf{CX}}$ is the charge exchagnge cross section. (Reprinted from [100])

Figure 2.49 displays the sub-GeV proton interaction probability and scattering angles within a ^{16}O nucleus in NEUT.

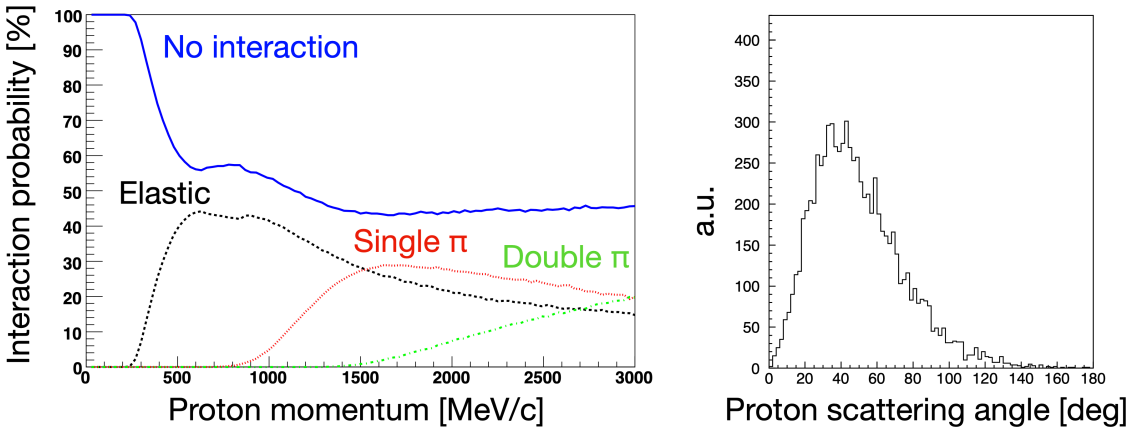


Figure 2.49: Probability of nucleon-nucleon interaction as a function of incoming nucleon momentum (left) and proton deflection angle (right) within ^{16}O. The incident proton momentum is uniformly sampled within the 0.5-1 GeV/c range. (Adapted from [103])

When the mean free path is smaller than the wavelength of the scattering particle, the generation of hadrons may be suppressed over a specific distance/time, known as the formation length/time [104]. This effectively decreases the hadron production cross-section. In NEUT, the formation length of pion production is parameterized to be proportional to the incident hadron momentum, with the coefficient determined by a neutrino-bubble chamber experiment [105]. For hadrons with a momentum of 1 GeV/c, the formation length is approximately 2.5 fm, which closely matches the size of the ^{16}O nucleus.

Nuclear de-excitation modeling

NEUT incorporates a straightforward nuclear de-excitation scheme for ^{16}O. Originally motivated by studies of proton decays in ^{16}O, this scheme gained further significance in the investigation of atmospheric neutrino NCQE interactions. NCQE interactions serve as a major background for inverse beta decay signals of diffuse supernova neutrinos. Figure 2.50 illustrates a neutrino knocking out a nucleon occupying the $p^{3/2}$ state of the ^{16}O nucleus.

In NEUT, the de-excitation of the residual nucleus after nucleon knockout can result in the emission of another nucleon or a gamma-ray. Deeper holes (such as ones in the $s^{1/2}$ state) in the residual nucleus with higher excitation energy are more likely to eject a nucleon or gamma-ray during the de-excitation process. The occupation probabilities of nucleon energy states are derived from [106], and the branching ratios for the knock-out of each state are based on [107].

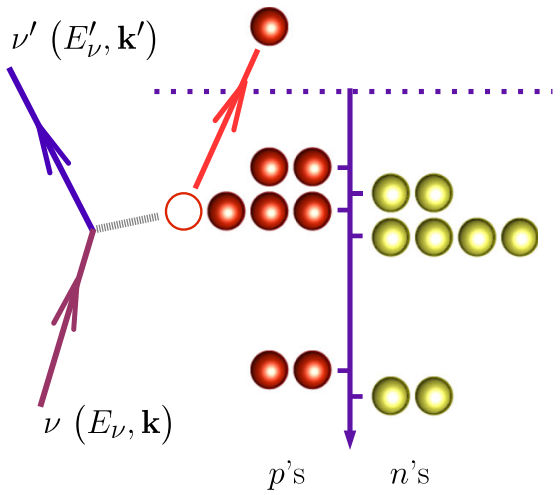


Figure 2.50: Illustration of a neutrino knocking out a proton from the $p^{3/2}$ state in a simple shell model of the ^{16}O nucleus. (Reprinted from [106])

Secondary interaction (SI) modeling

In the SK detector simulator, detailed in Section 2.3.5, the modeling of hadron interactions is designed to closely follow the approach used in `NEUT`. For low-momentum charged pions ($p < 500$ MeV), `NEUT` is directly called. Hadronic interactions involving higher momentum pions and nucleons are simulated using `GCALOR` 1.04/08. `GCALOR` provides interfaces to various MC simulation packages covering a range of hadron energies; for the relevant energy range, `HETC` [108], based on Bertini's INC model [95], handles hadronic scatterings, using mostly the same input cross sections implemented for `NEUT` nucleon interactions. Neutron interactions with kinetic energy below 20 MeV, including thermalization through elastic scatterings and neutron capture reactions, are modeled with `MICAP` [109]. The cross sections for `MICAP` are sourced from ENDF/B-V (1979). In SK simulation, `MICAP` with ENDF/B-V has demonstrated accurate reproduction of $^1\mathrm{H}(n,\gamma)^2\mathrm{H}$ signals, capturing both gamma-ray energy and the characteristic time constant (~ 200 μs) [34, 110]. These settings are used for simulating secondary interactions in pure water, corresponding to SK-IV and SK-V atmospheric neutrino MC simulation.

For the simulation of low-energy neutron interactions in Gd-loaded water, specifically targeting neutron captures on Gd, `MICAP` is replaced with the NeutronHP package [111] in Geant4.10.5.p01. This newer package relies on the updated evaluated nuclear cross sections of ENDF/B-VII.1 (2011, [31]). The validity of using Geant4 and ENDF/B-VII.1, especially regarding the simulation of the $\mathrm{Gd}(n,\gamma)$ reaction, will be discussed further in Section 3.3.6.

2.3.4 Neutron capture reaction model

For the $^{1}\mathrm{H}(n,\gamma)$ reaction, a single 2.2 MeV gamma-ray is emitted isotropically.

Regarding neutron captures on Gd, our detector simulator offers several models, including GLG4Sim [112] and GGARNET [113]. The default model is the ANNRI-Gd [113] model, which specifically considers the two most strongly neutron-absorbing Gd isotopes, $^{155/157}$Gd. In this model, the gamma-ray energy and multiplicity are sampled from a theoretical model fitted to data observed with the ANNRI Germanium spectrometer surrounding a natural Gd target irradiated with a neutron beam. The neutron energy was constrained to the thermal range by neutron time-of-flight. The uncertainties associated with using this model are further explored in Section 3.3.6.

2.3.5 Detector simulation

The outgoing particle kinematics from `NEUT` 5.4.0, the neutrino event generator with the HKKM 2011 atmospheric neutrino flux, were input into the SK detector simulator `skdetsim`, based on Geant 3.21, for the simulation of the detector response to atmospheric neutrino events. Table 2.5 provides a list of all the processes considered in particle tracking. The PMT charge/timing response, as well as optical photon propagation in water, are tuned with the detector calibration detailed in Section 2.2.3.

γ	Pair production
	Compton scattering
	Photoelectric effect
e	Multiple scattering
	Ionization / δ-ray production
	Bremsstrahlung
	Annihilation
	Cherenkov radiation
μ	Decay in flight
	Multiple scattering
	Ionization / δ-ray production
	Bremsstrahlung
	Pair production
	Nuclear interaction / μ^- capture
	Cherenkov radiation
Hadrons	Decay in flight
	Multiple scattering
	Ionization / δ-ray production
	Hadronic interactions / π^- capture
	Cherenkov radiation
Optical photons	Rayleigh scattering
	Mie scattering
	Absorption / Reflection

Table 2.5: Processes simulated in `skdetsim`, for each particle type.

2.3.6 Background modeling

The primary backgrounds for the neutron capture signals, typically ranging from 1 MeV with 1-10 PMT hits, are accidental coincidences attributed to PMT dark noise. Although a significant portion of dark noise is inherently random, there exist other types of background noise and radioactivity with characteristics that are not qualitatively characterized. To accurately simulate the background noise and minimize data-simulation bias, we extracted PMT hits recorded with random event triggers (also known as "dummy trigger" or "dark trigger," see "Software trigger" in Section 2.2.2). These dark events were pre-shuffled, and the PMT hits extracted from each of the dark events were then appended to the signal events simulated with the `skdetsim` detector simulator. The appended background hits did not overlap with the neutrino interaction signal following the event trigger, preventing interference with event reconstruction. At the same time, it was ensured that the background hits covered the entire neutron capture signal search time range, which is a few μs delayed from the event trigger.

2.3.7 Event weights

Once the full set of MC-simulated atmospheric neutrino events in the detector was generated, each event received a weight based on the three-flavor oscillation probability, calculated with a given true neutrino flavor and zenith angle. The neutrino production height was sampled from a height probability, a function of flavor and zenith angle provided by the flux model. The zenith angle and the production height fully characterizes the neutrino path length. For each neutrino event with a true lepton flavor $l \in \{e, \mu\}$, neutrino path length, and energy, oscillation probabilities $P(\nu_e \to \nu_l)$ and $P(\nu_\mu \to \nu_l)$ (the same applies for antineutrinos) were calculated using `Prob3++` [114] under the three-flavor standard oscillation scenario, including the Earth matter effect. The Earth matter density was modeled using the Preliminary Reference Earth Model (PREM, [115]). The three-flavor standard oscillation parameters used in this thesis are outlined in Table 4.1 [116]. The "oscillation" weight for each event ν_l ($l \in \{e, \mu\}$) was calculated as:

$$w_{\rm osc} = \frac{\phi_e(E_\nu, \theta_{\rm zenith}) P(\nu_e \to \nu_l) + \phi_\mu(E_\nu, \theta_{\rm zenith}) P(\nu_\mu \to \nu_l)}{\phi_l(E_\nu, \theta_{\rm zenith})} \tag{2.14}$$

where ϕ_l is the ν_l flux prediction with neutrino energy E_ν and zenith angle $\theta_{\rm zenith}$ (the same applies for antineutrinos). This weight considers contributions from both flavors, e and μ, oscillating to the observed flavor l.

In addition to oscillation, solar activity, as discussed in Section 2.3.1, was considered by normalizing the total event rate according to the average solar activity in each SK phase. (See Figure 2.35) If the livetime-averaged neutron monitor count is c, then the fraction spent in maximum solar activity for each SK phase is determined by $r_{\rm max} \equiv (c_{\rm max} - c_{\rm avg})/(c_{\rm max} - c_{\rm min})$, where $c_{\{\rm min/max\}}$ is the neutron monitor count at maximum and minimum solar activity. With $r_{\rm max}$, for each event, the "solar activity" weight is calculated as:

$$w_{\rm sol} = \frac{r_{\rm max}\phi_{\rm max}(E_\nu, \theta_{\rm zenith}) + (1 - r_{\rm max})\phi_{\rm min}(E_\nu, \theta_{\rm zenith})}{\phi_{\rm mid}(E_\nu, \theta_{\rm zenith})} \tag{2.15}$$

where $\phi_{\{\rm max/mid\}}$ is the flux prediction of the given neutrino type and flavor at maximum and intermediate solar activities, provided by the flux model.

In this analysis, the fractions determined for SK-IV (44% $r_{\rm max}$) were used for all phases for convenience, as most of our analyzed data is SK-IV ($\sim$ 10 years), which is much longer than the other two phases SK-V and SK-VI ($\sim$ 3 years combined). The final event weight assigned to each event is the product of both oscillation and solar activity weights, i.e., $w_{\rm osc} w_{\rm sol}$.

2.4 Event reconstruction

2.4.1 Atmospheric neutrino interaction

In this analysis, our goal is to characterize the properties of neutrons produced in atmospheric neutrino interactions within water, with a focus on understanding their dependence on neutrino interaction type and kinematics. For instance, we aim to investigate the increase in neutron production as a function of momentum transfer, and we seek to compare these observations with predictions from various models. Additionally, we want to examine the differences in neutron production between inelastic and elastic interactions and compare our findings with model predictions.

Also, there are inherent correlations between the variables in neutrino interactions and neutron detection efficiency. Neutrons produced in high-energy (on the order of O(1-10) GeV) neutrino interactions may travel farther from the interaction vertex, making them less likely to be detected compared to those produced in sub-GeV interactions. Moreover, neutrons generated close to the tank wall are difficult to distinguish from background radioactivity occurring around the wall.

While accurate reconstruction of the neutrino interaction is crucial, it remains challenging to precisely determine the interaction type or neutrino energy with atmospheric neutrinos given our current detector setup. Therefore, we rely on proxy variables that best capture the relevant information:

- **Visible energy**

  Visible energy is the sum of the reconstructed momenta from all rings in an event, assuming the ring was produced by an electron. This quantity provides a calorimetric measure of energy dissipated by the neutrino interaction, and, on average, exhibits a linear relationship with four-momentum transfer and the energy transferred to the hadronic system.

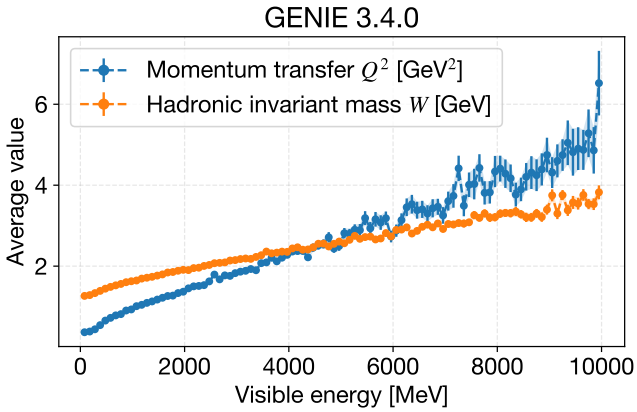


Figure 2.51: Average values of squared momentum transfer Q^2 (blue) and hadronic invariant mass W (orange) as a function of visible energy of SK atmospheric neutrino events simulated with `GENIE` 3.4.0.

- **Number of Cherenkov rings**

  The number of Cherenkov rings can indicate whether the interaction was quasi-elastic or inelastic, as implied in Figure 2.45. While the most energetic ring is likely associated with the charged lepton from the CC interaction, the presence of multiple rings may suggest the production of hadrons, particularly pions, indicating the inelasticity of the interaction and a larger energy transfer to the hadronic system, as shown in Figure 2.51.

- **Lepton type**

  Leveraging the excellent charged lepton PID performance of the detector, we can compare neutron properties between ν_e and ν_μ interactions.

- **Interaction vertex**

  The knowledge of the interaction vertex is essential for separating neutrino interactions from backgrounds such as cosmic rays or PMT flashers, which typically have vertices reconstructed on the tank wall.

SK uses `APFit` [117] to reconstruct observed Cherenkov rings in atmospheric neutrino events. We provide a brief overview of the Cherenkov reconstruction algorithm, covering aspects such as ring vertex determination, multiplicity, PID, and momentum in Appendix A. Further details can be found in [117, 118].

2.4.2 Neutron capture vertex

As discussed in Section 3.1, during the pure water phases of SK, 99.9% of neutron capture signals originated from neutron captures on free protons in water ($^1\mathrm{H}(n,\gamma)^2\mathrm{H}$), emitting a single 2.2 MeV gamma-ray. With 0.01w% Gd-loading in SK-VI or the current 0.03w% Gd in SK-VII, a significant portion, about half or more, of the neutron captures occur on Gd, primarily $^{155/157}\mathrm{Gd}(n,\gamma's)^{156/158}\mathrm{Gd}$. In such cases, multiple gamma-rays with a total energy of 7.94 MeV (for ^{157}Gd captures) or 8.54 MeV (for ^{155}Gd captures) are emitted.

The 2 MeV $^1\mathrm{H}(n,\gamma)$ signal results in approximately 6-7 PMT hits on average in our detector of size $O(10)$ m with approximately 12,000 ID PMTs. As a consequence, accurately estimating its vertex becomes extremely challenging. Therefore, we rely on the *a priori* knowledge of the neutron production vertex and assume that neutrons are captured in the vicinity of the production vertex.

On the other hand, with Gd-loading, the 8 MeV $\mathrm{Gd}(n,\gamma)$ signals typically deposit more than 20 PMT hits. In this energy range, we already have a vertex fitter that has been applied to reconstruct e-like rings of $O(1-10)$ MeV energy from solar and astrophysical neutrino interactions. Consequently, the neutron capture vertex can be independently reconstructed without relying on any *a priori* vertex information.

The primary motivation for neutron capture vertex reconstruction is to use the distance between neutron production and capture vertices as a proxy for the outgoing neutron momentum. This provides a handle on the outgoing neutron kinematics, which has been entirely lost during the 10 years of pure water phase data. In this thesis, we aim to investigate whether this new handle can offer additional constraints on the existing variability in hadronic interaction model predictions.

For neutron capture vertex reconstruction, we utilize the vertex fitter (`BONSAI` [119]) widely used in low-energy neutrino analyses (for example, [120]). This vertex fitter assumes that all light is emitted from a single vertex $\vec{x}$ at an instant time at T_0, given that $O(1)$ MeV electron tracks are $O(1)$ cm. It attempts to find them by minimizing the log likelihood $L(\vec{x}, T_0)$ of PMT hits, given by:

$$L(\vec{x}, T_0) = \sum_{i}^{N_{\mathrm{PMT}}} \log P(t_i - t_i^{\mathrm{ToF}}(\vec{x}) - T_0) \tag{2.16}$$

where t_i is the hit time of the i^{th} PMT, and $t_i^{\mathrm{ToF}}(\vec{x})$ is the photon time-of-flight (ToF) between the assumed vertex $\vec{x}$ and the i^{th} PMT. PMT charge is not considered, as most hits are due to single photon arrivals in the $O(1)$ MeV energy range. The probability P is determined by the distribution shown in Figure 2.52, empirically obtained from the $O(1)$ MeV electron beam data collected during LINAC calibration, as explained in Section 2.2.3.

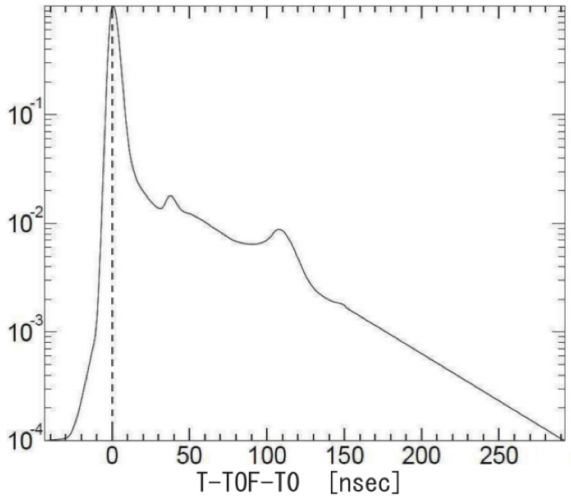


Figure 2.52: PMT hit probability density function as a function of hit time corrected by photon time-of-flight and the electron injection time T_0. The probability is highest at 0 as expected, while the peak width determines the timing resolution. The tails are due to reflected and scattered light, and the smaller peaks are due to PMT afterpulses whose time is characterized by the internal design of the PMT. (Reprinted from [121])

One challenge in optimization tasks involving maximization or minimization is the computational time and the risk of the computation getting trapped in local extrema. The `BONSAI` algorithm employs a distinctive strategy to efficiently explore the global extremum. Here, we provide a qualitative overview of the algorithm.

To begin with, the algorithm identifies "good PMT hits" based on "causality constraints." If two PMTs register hits from photons originating at a single vertex, the time difference between the two hits must be smaller than the distance between the two PMTs (assuming the speed of light, $c = 1$). Essentially, if two PMTs exhibit significantly different hit times, it is unlikely that they are detecting photons from the same source. We select the maximal set of "good PMT hits" satisfying the causality constraint between any two PMT hits. PMTs located more than 12.5 m or 35 ns apart from any other PMTs are rejected to prevent contamination from random PMT noise hits.

The goal is to determine the values of $\vec{x}$ and T_0, representing three vertex coordinates and a single signal time. If we have four equations related to these values, we can find a unique solution. From the "good PMT hits," we can randomly select four hits and solve four equations, stipulating that the distance between each PMT and the solution vertex must align with the expected photon time-of-flight.

Multiple combinations of four good PMT hits can yield a corresponding set of "candidate" vertices on a one-to-one basis. However, to avoid computational inefficiency (especially when having a large number of good PMT hits), we limit the number of possible combinations. Additionally, when candidate vertices are sufficiently close (less than 1.5 m apart), they are averaged.

Next, a list of vertices with significant log likelihoods is iteratively searched for and retained. Utilizing the selected candidate vertices and time, a dodecahedron (initial side length of 7.8 m) is drawn from each candidate vertex. Log likelihoods are computed for the 12 vertices of the dodecahedron, and only a fraction of vertices with the highest log likelihoods are preserved, while the rest are pruned. This process is repeated with a smaller ($\times 0.38$) dodecahedron for each remaining vertex until the size becomes sufficiently small (10 cm).

This method enables the exploration of the entire tank without being trapped in local extrema, and it efficiently selects likely vertices without the need to iterate over the entire grid points in the tank. The vertex resolution of this algorithm for Michel electrons, Gd(n, γ) signals, and p(n, γ) signals is shown in Figure 2.53. The vertex resolution for the target Gd(n, γ) signals is approximately 1 m, which falls between that for Michel electrons ($\sim$0.5 m) and ^{1}H(n, γ) signals ($>$ 3 m).

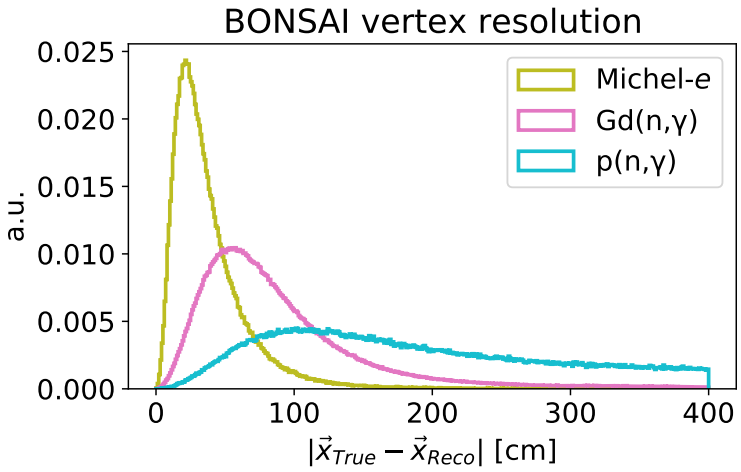


Figure 2.53: Comparison of vertex reconstruction performance for Michel electron, Gd(n, γ), and H(n, γ) signals. A smaller value indicates better performance.

2.5 Atmospheric neutrino event selection

Atmospheric neutrino events observed in the SK detector are classified into three distinct categories, as illustrated in Figure 2.54:

- **Fully contained (FC)**

  In FC events, there is no entry or exit of particles visible in the OD. Consequently, the majority of energy radiated in visible form is expected to be deposited within the ID.

- **Partially contained (PC)**

  In PC events, a portion of the radiated energy is deposited in the OD after the interaction, suggesting an exit of charged particles, mainly muons. However, there is no entry of particles visible in the OD, indicating that the interaction itself should have occurred within the ID.

- **Upward-going muon (UPMU)**

  UPMU events involve the entry of a charged muon visible in the OD but from the bottom of the tank. Cosmic rays are unlikely to come from the bottom, given their inability to penetrate the entire Earth radius. Therefore, these events are considered atmospheric muon neutrinos passing through the Earth and interacting within the rock surrounding the detector. Electron neutrinos are less likely as they are not as penetrative. UPMU events can be further classified into stopping muon events and through-going muon events.

The expected mean energies for atmospheric neutrinos are approximately $\sim$1 GeV for FC, $\sim$10 GeV for PC and stopping UPMU, and $\sim$100 GeV for through-going UPMU, as depicted in Figure 2.55. In this analysis, our focus is exclusively on FC events. This choice is to measure neutron properties as a function of the energy transferred by the neutrino interaction. FC events provide the complete calorimetric energy from the interaction deposited within the detector system, allowing us to minimize energy loss escaping the detector.

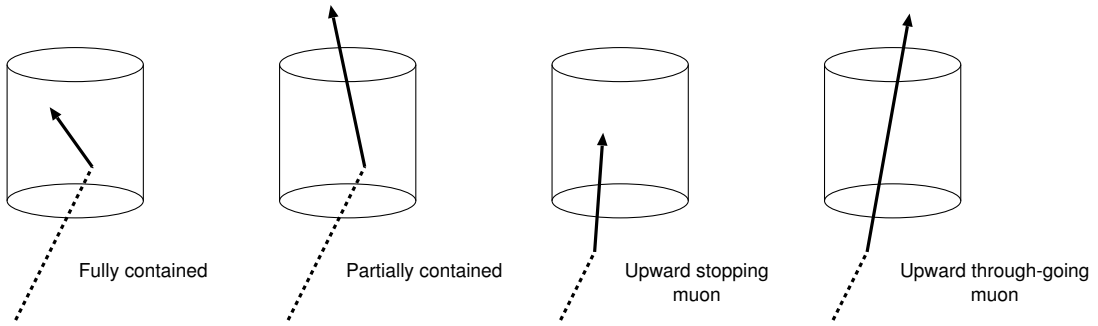


Figure 2.54: Illustration of SK atmospheric neutrino data event categories. (Reprinted from [122])

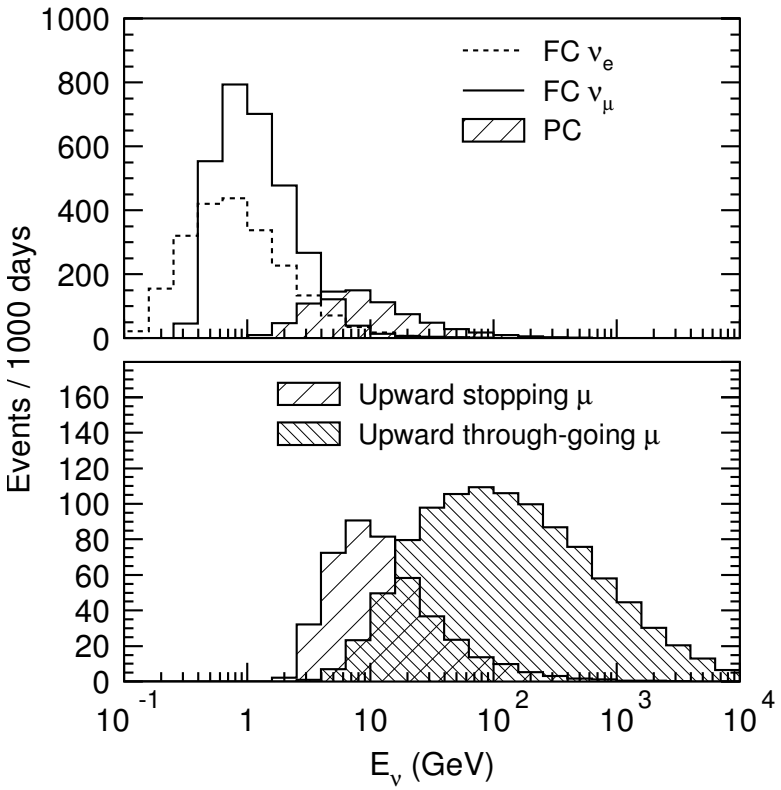


Figure 2.55: Expected true neutrino energy spectra of each event category. (Reprinted from [122])

Approximately 10^6 events are recorded daily in SK, predominantly consisting of background events such as cosmic rays and background radioactivity. To enhance the purity of atmospheric neutrino events, several data reduction cuts are applied. The primary backgrounds include:

- Cosmic ray muons and subsequent Michel electrons
- Low-energy gamma rays originating from radioactivity
- PMT "flashers," characterized by PMTs emitting light through internal discharge

We reduce the background level down to $O(0.1)\%$, using the following comprehensive five-stage data selection process:

(FC1) Initial cuts to reject evident backgrounds

(FC2) Subsequent cuts for further reject backgrounds

(FC3) Cosmic ray rejection based on muon reconstruction

(FC4) Flasher rejection based on PMT hit pattern likelihood

(FC5) Additional cuts to reject remaining backgrounds

We provide a description of the data selection criteria applied to event observables at each stage of data selection in Appendix B. For a detailed comparison between the data and simulation after passing these selection stages, refer to [118].

Chapter 3

Neutron detection

In this chapter, we begin by examining the characteristics of neutron capture signal and common background within the detector. Following this, we discuss the signal identification algorithm and evaluate its performance using calibration data obtained from a neutron point source, and compare the results with MC simulations. Sources of uncertainty in the evaluated performance are thoroughly examined.

3.1 Characteristics of signal and background

In pure water, signals primarily result from the $^{1}\mathrm{H}(n,\gamma)$ reaction, constituting over 99.9% of total neutron captures, emitting a single 2.2 MeV gamma-ray. For $\mathrm{Gd}(n,\gamma)$, which represents roughly half of neutron captures in 0.0110 w% Gd-loaded water, multiple gamma-rays are emitted, with a combined energy of either 7.94 MeV (for ^{157}Gd) or 8.54 MeV (for ^{155}Gd), as shown in Figures 3.1 and 3.2. The backgrounds primarily stem from coincidences of PMT noise, which includes dark noise, after-pulse, and scintillation within the PMT glass window. Another source of background are radioactive impurities found in water, PMT glass window, and the surrounding rock, notably ^{222}Rn, which produces $O(1)$ MeV electrons during its decay chain.

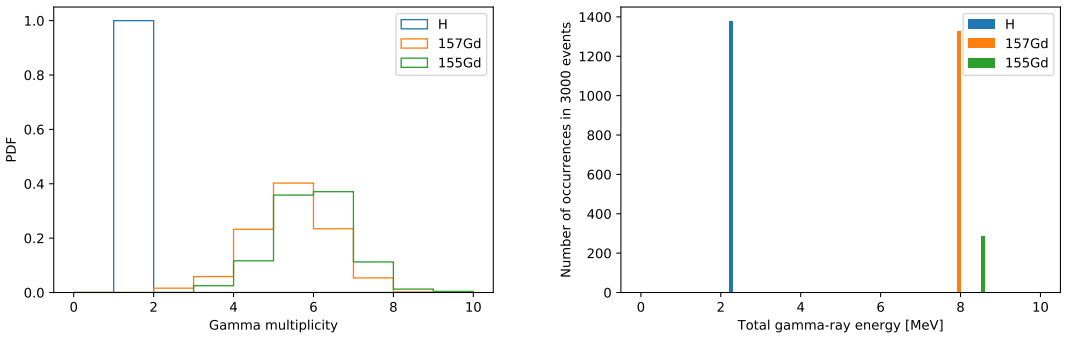


Figure 3.1: The gamma-ray multiplicity in neutron captures on H and Gd (left), and the energy sum of all emitted gamma-rays (right) obtained from simulation with GCALOR [123] and ANNRI-Gd model [113]. (area-normalized, reprinted from [124])

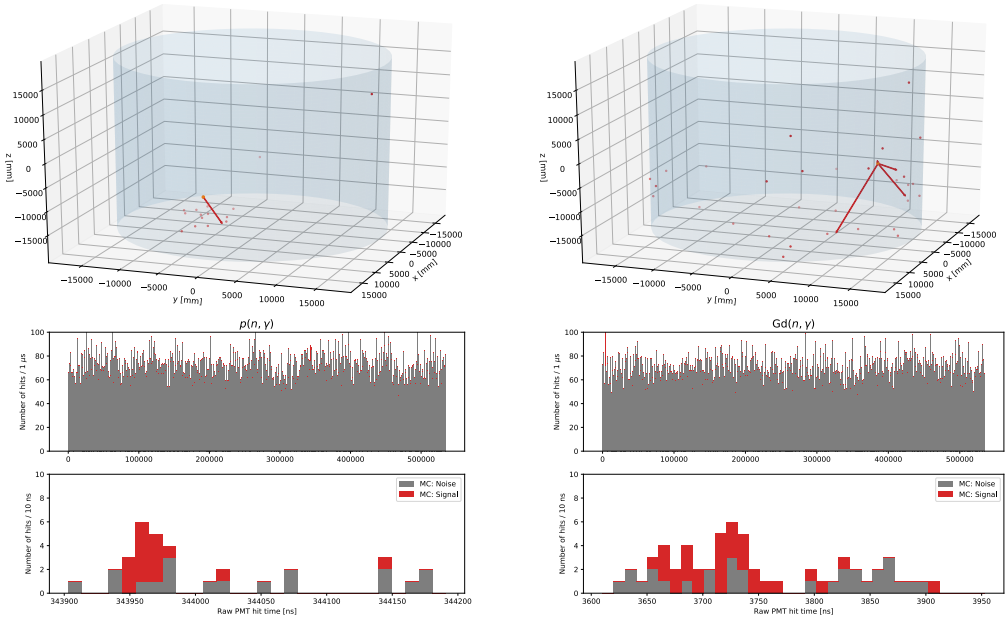


Figure 3.2: Sample MC event displays for a $^{1}\mathrm{H}(n,\gamma)$ (left) and a $^{157}\mathrm{Gd}(n,\gamma)$ event (right). The top figures are the 3D drawings of the inner tank dimensions with all neutron signal hits (red dots) and gamma-rays (red arrows) in each event. The bottom histograms are the raw PMT hit time distributions of each event with bin width being 1 μs and 10 ns, respectively. (reprinted from [124])

Following are the features that distinguish between signal and background:

- **Number of PMT hits**

  The likelihood of random coincidental PMT noise hits within a small fixed time window decreases exponentially with the total number of coincident PMTs. Conversely, neutron captures emit a consistent amount of energy (2-8 MeV per isotope) which typically triggers a specific number of PMTs. The deposited charge is less relevant, as most of the signal PMT hits result from a single Cherenkov photon.

- **Angular distribution of PMT hits**

  Gamma-rays from neutron captures scatter off bound electrons, generating Cherenkov light if their kinetic energy exceeds the Cherenkov threshold of 0.3 MeV. As a result, signals manifest as one or more faint electron-like Cherenkov rings with a distinctive angular distribution of roughly 42 degrees around the direction of the gamma-ray, and a total energy of 2-8 MeV. Conversely, coincidental random noise hits exhibit little angular correlation with each other. Radioactive decays near the tank wall typically have a highly localized angular distribution close to the tank wall.

- **Positional correlation of hit PMTs to signal vertex**

  Given that most neutrons in water can travel at most a few meters in a 36 m-high ID, the neutron capture vertex itself should be relatively close to the estimated neutron production vertex. Real neutron captures usually activate fewer PMTs that are farther from the capture vertex due to light attenuation in water. Fake signals caused by coincidental PMT noise display little correlation with the vertex. Furthermore, the reconstructed vertex for signals tends to be within the fiducial volume of the ID, while the reconstructed vertex for backgrounds is often on the tank wall.

- **Correlation to noisy PMTs**

  Certain PMTs that are relatively noisier are more prone to generating a false signal, whereas actual signal photons can trigger any PMT regardless of the noise rate of each individual PMT.

3.2 Signal identification algorithm

Here, we describe the algorithm implemented in the signal search software `NTag` [125]. Figure 3.3 shows a typical PMT hit time distribution.

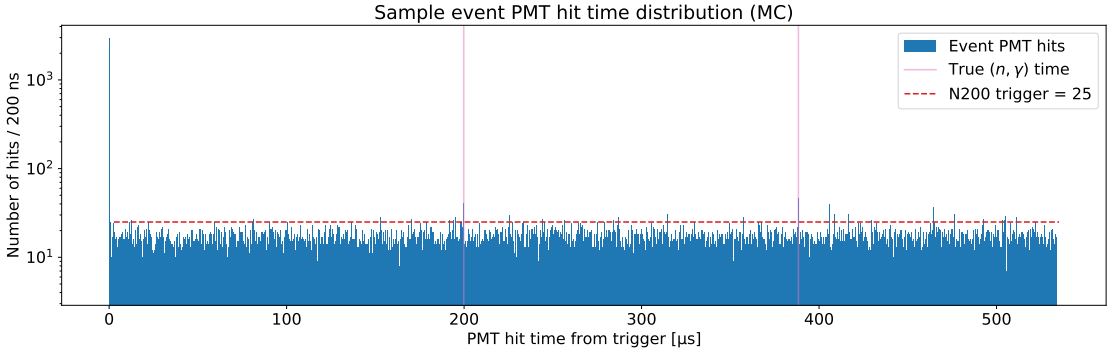


Figure 3.3: A sample hit time distribution for a simulated neutrino event. Events typically have a large number of triggering PMT hits and a small number of PMT hits due to neutron captures buried in background noise spanning over 500 μs.

Before initiating the signal search, we perform a preliminary cleaning of unnecessary PMT hits. We first eliminate PMT hits with charge deposits equivalent to over 10 photoelectrons, as the majority of signal PMT hits are expected to be caused by a single photon. Next, we discard the latter hit among two consecutive hits on the same PMT within a 1,000 ns interval to ensure uniform deadtime across all PMTs.

Then, the signal detection algorithm operates in two main stages:

Stage 1: Candidate Selection

We establish a threshold for the number of PMT hits to identify potential signal "candidates." A group of hits within a sliding time window is considered a candidate if the number of hits surpasses a predefined threshold.

Stage 2: Candidate Classification

Each candidate is then categorized as either signal or background based on its characteristics. In this particular analysis, we trained a neural network on simulated candidate features and then applied it to candidates observed in previously unseen data.

3.2.1 Stage 1: Candidate selection

For each event (which records PMT hits over a fixed period of time), we execute the following steps:

1. We begin by correcting for the photon time-of-flight (TOF) from the estimated signal vertex to each hit PMT. This correction reduces the variance of signal PMT hit times from the order of $O(100)$ to $O(10)$ ns, significantly enhancing the signal-to-noise ratio in a smaller time window as in Figure 3.4. The vertex information should be provided *a priori*. For instance, in the context of searching for neutrons produced in a neutrino interaction, a reconstructed neutrino vertex can be given. Alternatively, in the case of seeking neutrons originating from a neutron source, the position of the source can be used.

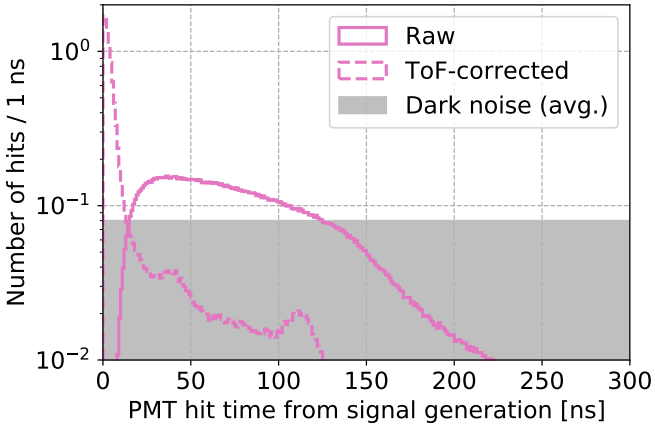


Figure 3.4: Gd(n, γ) signal PMT hit time distributions. True neutron vertex was used as the input for photon TOF correction. Dark noise represents the SK-VI average. The tails are mainly due to light reflection on the tank wall.

2. Next, we arrange all PMT hits chronologically and scan for sets of PMTs with a hit count exceeding a threshold of around 5-20 hits within a $O(10)$ ns sliding time window. In cases where there are adjacent sets (i.e., two distinct sets of PMTs separated by less than a predetermined time width), we discard the set with the smaller hit count and consider the set with the larger number of hits as a potential signal candidate, to prevent double counting of a signal.

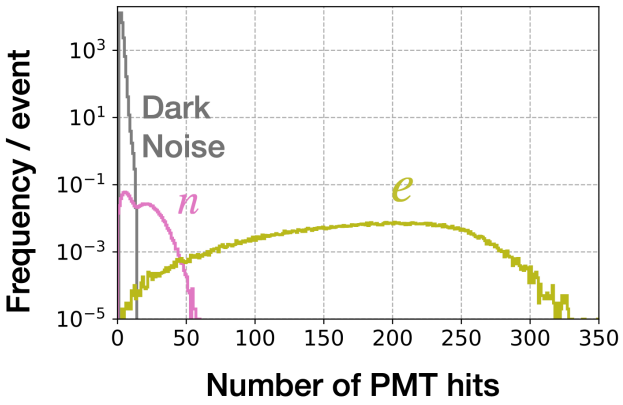


Figure 3.5: Number of PMT hits within 14 ns after photon TOF correction based on the true particle-gun simulation vertex, for PMT dark noise (gray), H and Gd neutron capture signals (pink), and Michel electron signals (olive).

Vertex options for photon TOF correction

The photon TOF correction's effectiveness in reducing hit time variance, as shown in Figures 3.4 and 3.5, depends on the proximity of the actual neutron capture vertex to the *a priori* vertex used for TOF correction. In most cases, the given vertex is the estimated position of neutron production, determined from prompt signals generated upon neutron production. This method is the most efficient way to approximate the neutron capture vertex, given that neutron capture gamma-rays, especially from $^1\mathrm{H}(n,\gamma)$, have low energy levels for accurate vertex reconstruction. The majority of neutron captures occur close to their production vertex.

However, this presents a challenge as signal detection performance starts to rely on the distance between the neutron capture and the provided input vertex. This is particularly relevant when detecting neutron captures far from the production vertex, indicating the production of highly energetic neutrons. To address this, we may independently reconstruct the neutron capture vertex and use it for TOF correction and subsequent feature extraction. This approach is feasible for 8 MeV $\mathrm{Gd}(n,\gamma)$ signals, with a vertex resolution of approximately 1 m. In the Gd-loaded phase, we can choose one of two strategies for signal detection:

- `INIT` **vertex mode** (default mode used in this study)

  No additional vertex reconstruction is performed; the selected candidates are used as they are. Features, such as signal vertex distance to the tank wall and hit timing distributions, are computed based on the initial provided vertex used for TOF correction in the candidate selection stage. This yields higher overall detection performance, though the performance does depend on the distance between the initial provided vertex and the actual neutron capture vertex. This approach is the only applicable method in the pure water phase.

- `RECO` **vertex mode** (used for vertex reconstruction only)

  The neutron capture vertex and time are independently reconstructed using only PMT hits contained in each candidate. Using the algorithm explained in Section 2.4.2, we find a vertex that maximizes the likelihood of hit PMTs in [-0.5, 1] μs range around each candidate. The TOF for each PMT is adjusted based on the distinctly reconstructed vertex for each candidate. A set of PMT hits that constitute a candidate are rearranged as the TOF changes. Due to the relatively lower vertex resolution of neutron capture signals compared to much stronger prompt signals, a larger sliding time window is necessary to include more signal hits. Features, such as signal vertex distance to the tank wall and hit timing distributions, are computed based on the reconstructed neutron capture vertex. A candidate is rejected if the reconstructed signal time deviates significantly (200 ns) from the central time of the original candidate.

Table 3.1 list the parameters used in this study. Figure 3.6 compares the performance of the `INIT` and `RECO` vertex modes.

Parameter	Description	Vertex mode	
		INIT	RECO
TWIDTH	Sliding time window width [ns]	14	14
TCANWIDTH	Hit selection time width for each candidate [ns]	14	30
NHITSTH	Minimum PMT hits to allow as a candidate	5	7
NHITSMX	Maximum PMT hits to allow as a candidate	400	400
TMINPEAKSEP	Minimum separation of two adjacent candidates [ns]	50	200
TMATCHWINDOW	Match window used for candidate labeling in MC [ns]	50	200

Table 3.1: Parameters used in the vertex modes for candidate selection. The parameter TWIDTH was adjusted to maximize the signal significance $S/\sqrt{S+B}$ for $^1\mathrm{H}(n,\gamma)$ signals, where S represents the number of signal PMT hits and B represents the number of background PMT hits in a candidate. Other parameters were determined heuristically. For more details, refer to [124].

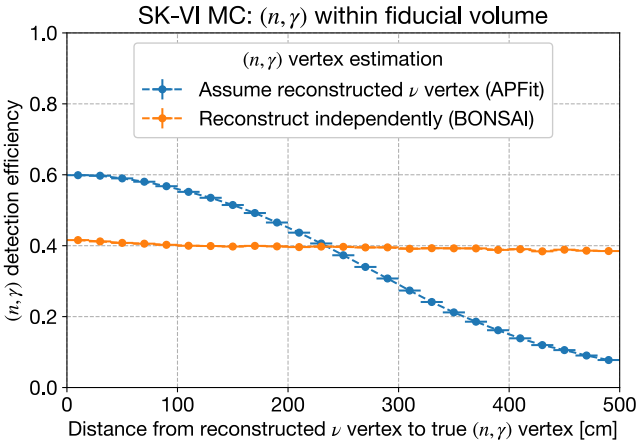


Figure 3.6: Comparison of overall signal detection efficiencies in SK-VI simulated atmospheric neutrino events, between the two vertex modes — INIT vertex mode (blue) and RECO vertex mode (orange), as a function of the distance from the reconstructed neutrino vertex to true neutron capture vertex. While the detection performance of the RECO vertex mode is generally not as robust as the INIT vertex mode due to the poorer vertex resolution, it is independent of the distance from the initially provided vertex to the neutron capture vertex.

3.2.2 Stage 2: Candidate classification

Extracting features

For each signal candidate selected in the steps described in 3.2.1, the following 14 features based on the discussion in 3.1 are extracted according to each vertex mode:

- `NHits`

  This represents the number of PMT hits within each candidate's time width (see `TCANWIDTH` in Table 3.1). Because the likelihood of accidental noise coincidences decreases exponentially with an increasing number of coincident PMT hits, this feature stands out as one of the most potent discriminators between signal and background.

- `NResHits`

  This denotes the count of "residual" hits, which is the number of PMT hits within a 200-ns time window ([-100, 100] ns, with 0 as the midpoint of the candidate time width) minus `NHits`. Signal hits typically have very small variance in time after TOF correction, resulting in a low count of residual hits. Conversely, background hits tend to have a larger time spread and a higher count of residual hits. This feature also aids in the rejection of PMT afterpulses that can lead to an increase of residual hits.

- `TRMS` [ns]

  This represents the root-mean-squared (RMS) value of the TOF-corrected PMT hit times. Signal hits tend to yield smaller values.

- `FitGoodness`

  This metric quantifies the goodness of hit timing based on the formula [119]:

$$g(t_{\text{test}}, \vec{x}_{\text{test}}) = \sum_{i=1}^{N} w_i e^{-\left(t_i - \frac{|\vec{x_i} - \vec{x}_{\text{test}}|^2}{c} - t_{\text{test}}\right)/(2\sigma_5^2)} \tag{3.1}$$

Here, $\vec{x}_{\text{test}}$ represents the tested vertex position, N is the number of hit PMTs in a candidate, and t_i, $\vec{x}_i$ are the recorded hit time and position of the i^{th} PMT respectively. c is the speed of light, σ_5 is the assumed timing resolution (5 ns) of the PMTs, and w_i, given by the equation below, is the Gaussian weight with a larger timing resolution σ_{60} (60 ns) to filter out outlier hits.

$$w_i = \frac{e^{-\left(t_i - \frac{|\vec{x}_i - \vec{x}_{\text{test}}|^2}{c} - t_{fit}\right)/(2\sigma_{60}^2)}}{\sum_{i=1}^{N} e^{-\left(t_i - \frac{|\vec{x}_i - \vec{x}_{\text{test}}|^2}{c} - t_{\text{test}}\right)/(2\sigma_{60}^2)}} \tag{3.2}$$

For the `INIT` vertex mode, t_{test} is defined as the central time of the candidate time window, and $\vec{x}_{\text{test}}$ is given by the initially provided vertex. For the `RECO` vertex mode, t_{test} is defined as the most likely signal generation time returned by `BONSAI`, and $\vec{x}_{\text{test}}$ is given by the `BONSAI`-reconstructed vertex. This metric is larger for signals and smaller for background.

- `DWall` [cm]

  This is the distance from the assumed neutron capture vertex to the nearest tank wall. In the `RECO` vertex mode, some background events originating from radioactivity near the tank wall or in the PMT glass window may have their vertices reconstructed very close to the tank wall.

- `DWallMeanDir` [cm]

  This represents the closest distance from the assumed neutron capture vertex to the tank wall in the average hit direction, or the mean of the unit vectors connecting the assumed neutron capture vertex to each hit PMT. Signal hits tend to have smaller values, as Cherenkov photons are more likely to hit PMTs closer to their point of origin due to light attenuation in water.

- `BurstRatio`

  This is the ratio of the number of "burst" PMTs to the total number of hit PMTs in a candidate, where a "burst" PMT is defined as a PMT hit with a preceding hit within 10 μs. When radioactive impurities trigger a PMT, scintillation within the PMT glass window occurs with a time constant of a

few μs, as shown in Figure 3.7. This feature captures the correlation between two consecutive noise hits.

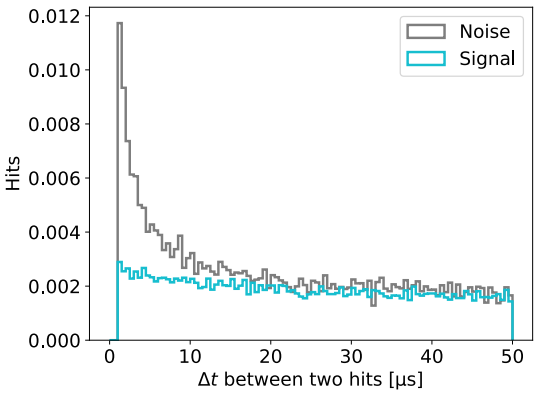


Figure 3.7: Time difference of two consecutive hits in a same PMT, for signal (cyan) and noise (gray) hits.

- `DarkLikelihood`

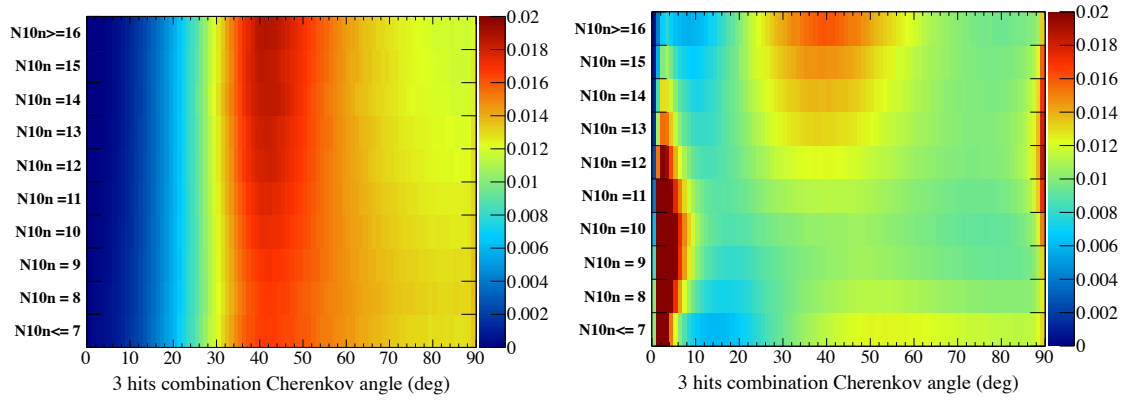
  This metric is related to the likelihood of the coincident noise scenario of hit PMTs, considering their individual average dark rates which are measured for each run in SK. The value is defined as:

$$\sigma\left(\log\prod_{i=1}^{N}\frac{r_i}{\langle r\rangle_{All}}\right) \tag{3.3}$$

  where σ represents the sigmoid function $\sigma(z) = 1/(1+e^{-z})$, N is the number of hit PMTs in a candidate, r_i is the measured average dark rate of the i^{th} PMT, and $\langle r\rangle_{All}$ is the average dark rate of all PMTs. The dark rates may vary from one run to another. When r_i is similar to $\langle r\rangle_{All}$, the value approaches 0.5. As r_i approaches ∞, the value approaches 1, and as r_i approaches 0, the value approaches 0. Signals tend to have smaller values compared to accidental background, which tends to include more noisy PMTs.

- `OpeningAngleStdev` [°]

  This is a measure of the variation in opening angles across all possible 3-hit-PMT combinations in the candidate. Opening angles are determined by



(a) Signal hits

(b) Background hits

Figure 3.8: Opening angle distributions of signal and background hits. `N10n` in vertical axes represents the number of PMT hits within 10-ns candidate time window. (Reprinted from [34])

creating a cone using the base circle formed by each 3-hit-PMT combination, with the apex placed at the initial provided vertex or the reconstructed neutron capture vertex, depending on the vertex mode. The half-angle of the cone is then taken as the opening angle. While the opening angle distribution of signal hits generally peaks at 42°, for background hits the spread can range from very small to large as illustrated in Figure 3.8b, particularly for a portion of hits that exhibit anisotropic behavior due to background radioactivity.

- `Beta1-5`

  The parameter `Beta(l)` is defined by the equation [126]:

$$\beta_l = \frac{2}{\texttt{NHits}(\texttt{NHits}-1)} \sum_{\substack{i,j \\ i \neq j}} P_l(\cos\theta_{ij}) \tag{3.4}$$

  where P_l represents the Legendre polynomial of degree l (with $1 \leq l \leq 5$), and θ_{ij} is the angle between vectors extending from the given vertex to the i^{th} and j^{th} hit PMTs. These parameters characterize the angular distribution of the PMT hits and have demonstrated significant discriminatory power between signal and background hits in the SNO experiment [126]. In general, lower values of the `Beta(l)` parameters indicate a higher degree of angular isotropy in hits. While β_l values for different l are correlated, including β_l for all five l

values in the classification process has demonstrated a slightly improved signal efficiency compared to using β_l for only two different l values.

For candidates selected in MC-simulated events, we assigned the label `signal` if the central time of the candidate fell within O(10-100) ns time window (specified by the parameter `TMATCHWINDOW` in `NTag`, refer to Table 3.1) of a simulated physics event (such as neutron capture, Michel electron, and non-capture gamma-ray with energy over 1 MeV). Candidates not meeting this criterion were labeled as `background`. Candidates labeled as `signal` were further classified into `H-capture` ($\sum E_\gamma \leq 6$ MeV), `Gd-capture` ($\sum E_\gamma > 6$ MeV), `Michel-e`, and `Gamma`, based on the coincident type of simulated physics event. We can use the terms 'true' background, ^{1}H(n,γ), and Gd(n,γ) candidates interchangeably with the corresponding labels in this context. Figure 3.9 shows the expected feature distributions.

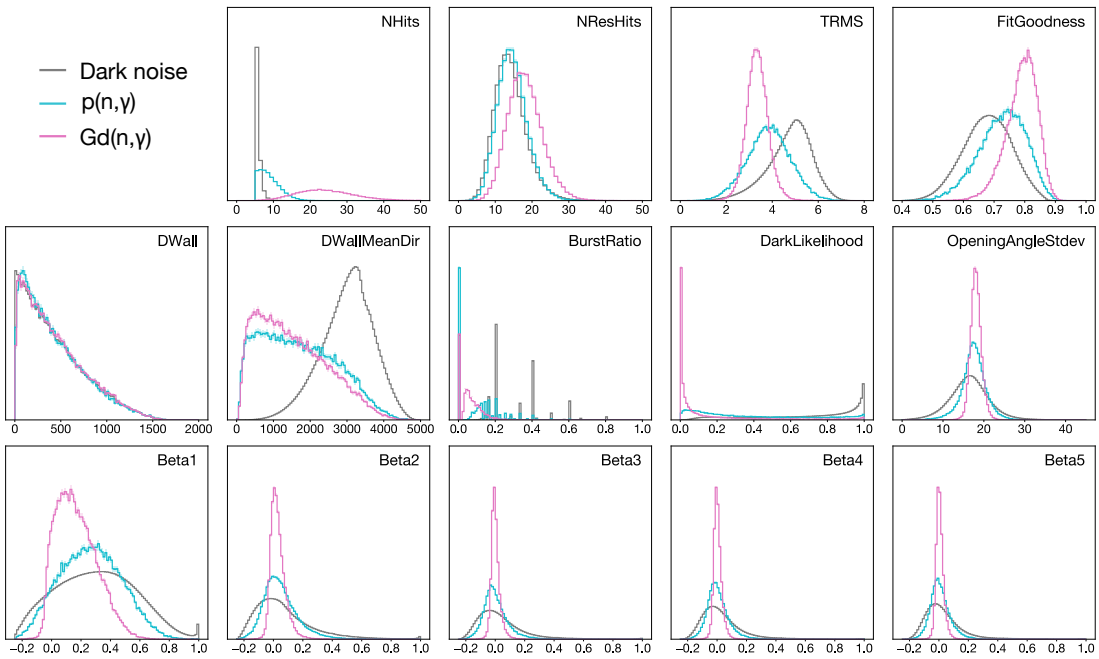


Figure 3.9: Comparison of 14 extracted features in the `INIT` vertex mode, for `background` (gray), `H-capture` (cyan), and `Gd-capture` (pink) candidates, which were selected from events of simulated PMT hits from SK-VI thermal neutron MC, along with background PMT hits extracted using random event triggers. (All histograms are normalized by area.)

Neural network for feature classification

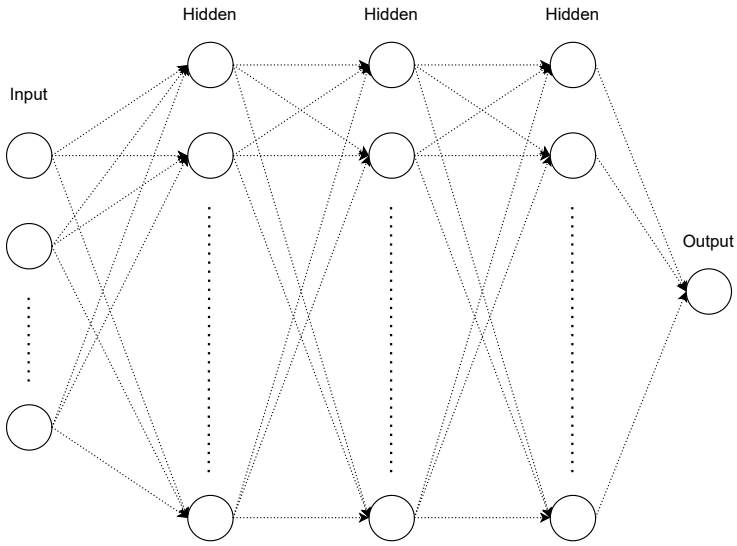


Figure 3.10: An illustration of a standard neural network architecture featuring multiple input nodes, three hidden dense layers, and a single output node.

A neural network is a computational model, similar to a multivariate function, that takes specific inputs and generates one or more outputs. In a feedforward, fully connected network (see Figure 3.10), nodes process incoming data from all nodes in the preceding layer using a function represented as:

$$\text{Node output} = \text{Activation}\left(\sum_{i}^{\text{All inputs}} (\text{Input}_i \times \text{Weight}_i) + \text{Bias}\right) \tag{3.5}$$

Each node is associated with a set of weights, which scale the inputs, and a bias term, which adjusts the output. Typically, activation functions like hyperbolic tangent tanh, sigmoid $f(x) = 1/(1+e^{-x})$, or Rectified Linear Unit (ReLU) given by $f(x) = \max(0, x)$ are chosen to introduce non-linearity and increase the complexity of the input-output relationship.

One of the most remarkable aspects of neural networks is that given a set of input data X and a loss function $L(X, W, B)$, we can find a set of weights W and biases B (the parameters of the neural network) that minimize the given loss function.

This process, often referred to as 'learning' or 'training', is carried out iteratively using the gradient descent algorithm for each batch of input features. This involves computing the gradients $\nabla_W L$ and $\nabla_B L$ and updating the weights and biases with $-\alpha\nabla_W L$ and $-\alpha\nabla_B L$ respectively. Here, the learning rate α is a hyperparameter that influences the learning process. The non-linearity introduced by the activation functions enables the network to learn complex relationships between the input and output. For more in-depth information, refer to [127].

For the task of signal and background classification in this analysis, we designed a feedforward, fully connected neural network with an input layer comprising 14 features, three hidden layers with 128 ReLU-activated nodes, a 50% dropout rate per each hidden node, and a single sigmoid output node. This network was implemented using Keras 2.6.0 [128]. The architecture which was optimized heuristically during training is summarized in Table 3.2.

Option	Settings
Layer architecture	14-128-128-128-1
Hidden layer activation	ReLU
Dropout rate	50%
Output layer activation	Sigmoid

Table 3.2: Neural network configuration settings.

For the network training, we utilized signal features generated from thermal neutron MC simulations using `skdetsim` v14 for the pure-water phases (SK-IV, SK-V) and v15 for the Gd-loaded phase (SK-VI). Detector parameters, including PMT gain as well as light attenuation and scattering parameters, were fine-tuned using cosmic ray muon and laser data specific to each phase. Background hits were derived from randomly triggered data events, which were shuffled in advance to average out the impact of varying dark rates over time. Additionally, background events with more than 60 ID PMT hits or 20 OD PMT hits within a 200-ns window were discarded to prevent the network from misidentifying a physical event as background. The background hits extracted were added to each simulated signal event within the range of [-1, 501] μs from T_0, the start of the simulation, without any duplicates.

The training simulation configurations are outlined in Table 3.3.

Option	Settings
Training data size	100,000 neutron particle-gun events
Signal search duration from T_0	[0, 500] μs
Neutron vertex	Random in ID
Neutron kinetic energy	0.025 eV
Neutron direction	Random in 4π sr
Background PMT hits	Extracted with random event triggers
(SK-VI only) Gd concentration	0.011w% [51]
(SK-VI only) Gd(n, γ) reaction model	ANNRI-Gd [113]

Table 3.3: Training simulation configuration settings.

Candidates were selected and features were extracted according to the steps outlined in Sections 3.2.1 and 3.2.2. True simulated neutron vertex was used for the initial TOF correction. Extracted features were standardized before being fed into the input layer.

For the loss function, we used binary cross entropy, a widely used metric for binary classification tasks. It is defined as:

$$L(y_i, y_i^{\text{true}}) = -\frac{1}{N}\sum_{i=1}^{N}\left(y_i^{\text{true}}\log y_i + \left(1 - y_i^{\text{true}}\right)\log\left(1 - y_i\right)\right) \tag{3.6}$$

Here, y_i represents the sigmoid output (ranging from 0 to 1) of the neural network, which is a function of input features $\vec{x}_i$, weights W, and biases B, while y_i^{true} is the true label, either 0 or 1 for the two classes. In our analysis, $y_i^{\text{true}} = 0$ signifies a candidate labeled as `background`, and $y_i^{\text{true}} = 1$ denotes a candidate labeled as `H-capture` or `Gd-capture`. It's important to note that as y_i approaches y_i^{true}, the value of L tends towards 0, regardless of whether y_i^{true} is 0 or 1. To prevent the weights from becoming overly complex, a small L2 regularization term was added, which is the sum of the squares of all weights with a coefficient of 0.00001.

Weights and biases were initialized using the He initialization technique [129], which helps mitigate the issues of vanishing or exploding gradients. The widely used stochastic gradient descent algorithm, Adam [130], was employed to update the weights using a minibatch of size 2048 per iteration. The initial learning rate was set to 0.0001. 80% of the training data was utilized for updating the weights, while the remaining 20% was reserved for validation. The hyperparameters, including the minibatch size, initial learning rate, and regularization term, were determined heuristically while closely monitoring the training process. The process of updating weights using the training data, continued until the signal efficiency in the validation set ceased to improve for 5 consecutive epochs. The training conditions are summarized in Table 3.4, and Figure 3.11 shows an example of loss versus epochs.

Four neural networks, all with the same architecture and training conditions, were trained on simulations tuned to each SK phase and vertex mode: SK-IV, SK-V, SK-VI with `INIT` vertex mode and SK-VI with `RECO` vertex mode.

Option	Settings
Train : validation split	80% : 20%
Minibatch size	2048
Input scaling	Standardized
Weight initialization	He [129]
Loss	Binary cross entropy
Regularization	L2: 0.00001
Loss optimizer	Adam [130]
Initial learning rate	0.0001

Table 3.4: Neural network training configuration settings.

Algorithm evaluation

We evaluated the signal efficiency and false positives per unit time, using a distinct test set of thermal neutron MC-simulated events for each SK phase (SK-IV, SK-V, and SK-VI). These simulations were generated with unique random seeds and unseen background PMT hits. The remaining configurations were aligned with the training data conditions outlined in Table 3.3. Moving forward, we will implicitly refer to the `INIT` vertex mode unless stated otherwise. The notation 'SK-VI (R)' denotes the results for SK-VI in the `RECO` vertex mode.

Stage 1 candidate selection performance

In the `INIT` vertex mode, the efficiency in selecting signals as candidates ranged from 71% to 80% for $^{1}\mathrm{H}(n,\gamma)$ signals, and while it reached 98% for $\mathrm{Gd}(n,\gamma)$ signals. The accidental background rate averaged 178, 454, and 486 background candidates per 500 μs interval (the typical search time range per neutrino event) for SK-IV, SK-V, and SK-VI, respectively. Notably, SK-V and SK-VI exhibited a higher background rate and $^{1}\mathrm{H}(n,\gamma)$ efficiency compared to SK-IV. This discrepancy is attributed to a larger average PMT dark rate (8 kHz) in SK-V and SK-VI in contrast to SK-IV (6 kHz).

In contrast, in the `RECO` vertex mode of SK-VI, a substantially lower count of 21 background candidates were selected per 500 μs. This reduction can be attributed to a higher threshold and an additional constraint on the reconstructed signal time. This mode also demonstrated a lower efficiency of 32% for $^{1}\mathrm{H}(n,\gamma)$ signals, while still maintaining a robust efficiency of 94% for $\mathrm{Gd}(n,\gamma)$ signals. Table 3.5 provides a summary of the candidate selection performance.

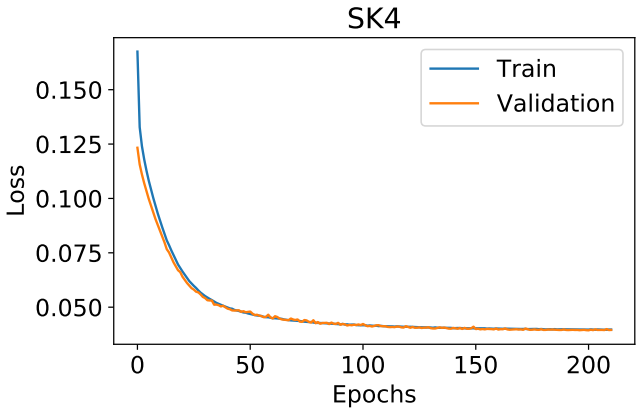


Figure 3.11: The evaluated loss against training epochs for both the training and validation sets (SK-IV `INIT-VERTEX` mode). Beyond 100 epochs, the loss stabilizes, showing minimal differences between the training and validation sets.

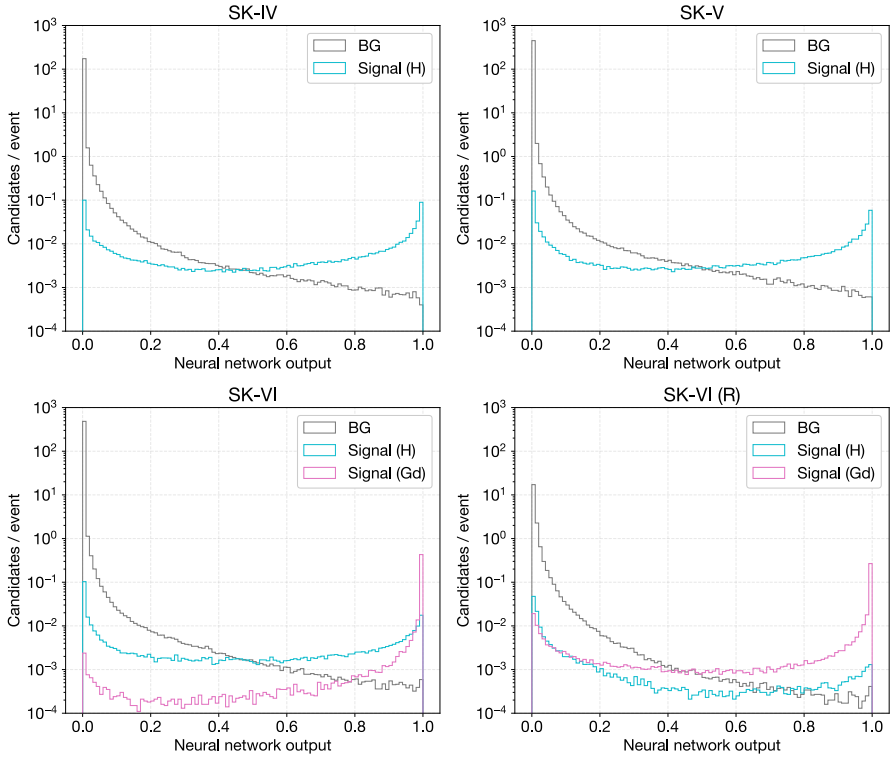


Figure 3.12: Distributions of neural network outputs for true signal and background candidates from the SK-IV, SK-V, and SK-VI thermal neutron MC simulations. SK-VI (R) shows the result in the `RECO` vertex mode.

Figure 3.12 shows the output of the trained neural networks for each SK phase. It is clear that the output for signals peaks at 1, while backgrounds peak at 0. Notably, the Gd(n, γ) signals show a much sharper peak at 1 compared to the ^{1}H(n, γ) signal, owing to the significantly better separability in `NHits` as opposed to the background. SK-V shows a slightly diminished signal peak compared to SK-IV. This is attributed to the larger dark rate, which causes signal features other than `NHits` to resemble those of the background, and vice versa. Furthermore, in the `RECO` vertex mode of SK-VI, there is a noticeable reduction in signal peaks compared to the `INIT` vertex mode. This is attributed to the degradation in vertex resolution and the resulting hit time resolution.

We can establish a cut threshold on the network output to distinguish signals from backgrounds. By decreasing the cut threshold from 1, we can enhance signal efficiency but at the expense of an increased number of false positives. This correlation is clearly illustrated in Figure 3.13. Therefore, selecting an appropriate cut threshold necessitates striking a balance between these two metrics.

To maintain the false positive rate per neutrino event at around $O(10^{-2})$, we have opted for a uniform cut threshold of 0.7 across all SK phases.

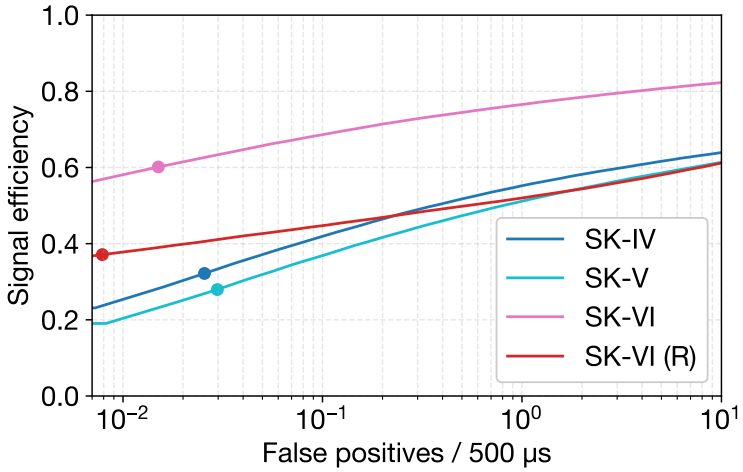


Figure 3.13: Performance of the overall algorithm plotted in terms of signal efficiency and false positives per 500 μs, with varying neural network output cut thresholds. The data points correspond to the performance achieved with a threshold of 0.7.

	SK-IV	SK-V	SK-VI	SK-VI (R)
Background candidates / 500 μs	177.7	453.8	485.6	21.2
^{1}H(n,γ) efficiency	70.52%	73.90%	80.13%	31.58%
Gd(n,γ) efficiency	-	-	98.40%	94.27%
Overall signal efficiency	70.52%	73.90%	89.52%	63.80%

Table 3.5: Summary of Stage-1-only candidate selection performance.

The performance of the entire algorithm, including both the candidate selection and the neural network classification stages with a 0.7 cut threshold, is summarized in Table 3.6. Specifically, the `INIT` vertex mode demonstrated an efficiency of 20-30% for ^{1}H(n,γ) signals and 94% for Gd(n,γ) signals.

Overall signal detection performance on thermal neutron MC

Stage 1 and 2 combined

	SK-IV	SK-V	SK-VI	SK-VI (R)
False positives / 500 μs	0.0267	0.0307	0.0156	0.0082
^{1}H(n,γ) efficiency	32.47%	28.24%	24.27%	3.07%
Gd(n,γ) efficiency	-	-	94.41%	69.50%
Overall signal efficiency	32.47%	28.24%	60.33%	37.22%

Table 3.6: Summary of signal detection performance of Stage 1 and 2 combined.

One might wonder which features contribute how much, on average, to the final network output. A widely-used metric in interpreting machine learning model predictions is the Shapley value [131], a concept from cooperative game theory that provides a solution for fairly distributing payoffs to players in a coalition.

The Shapley value ϕ_i is determined by the average marginal contribution of the i^{th} player to the total payoff, across all possible coalitions of N players:

$$\phi_i(v) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} [v(S \cup \{i\}) - v(S)] \tag{3.7}$$

Here, v represents the total expected sum of payoffs within a coalition S. This value possesses unique properties, such as efficiency ($\sum_i \phi_i(v) = v(N)$) and symmetry ($\phi_i(v) = \phi_j(v)$ if and only if $v(S \cup \{i\}) = v(S \cup \{j\})$ for all S that excludes both i and j), which ensure equitable distribution of payoffs.

Recently, this value has gained attention for fairly estimating feature contributions in individual model predictions. This value can be extended to models with multiple features, treating features as players, and v as the prediction for feature values in set S, marginalized over features not included in set S.

`SHAP` (SHapley Additive exPlanations) [132], one of the "additive" feature attribution methods, employs this approach and implements the computation of v as well as ϕ, which is referred to as a `SHAP` value. `SHAP` also defines "feature importance" as the mean of absolute `SHAP` values for all data instances. This can be interpreted as the average impact of each feature on output magnitude.

Figure 3.14 compares `SHAP` feature importance values calculated for randomly selected 200 candidates in each SK phase test data, normalized by the sum of the values. This allows us to compare the fraction of average impact a feature has on the network output among different SK phases. We can infer that `NHits`, `TRMS`, and `DWallMeanDir` are the top three features that contribute most to the network output in the `INIT` vertex mode, while `DWall` is equally important as `NHits` in the `RECO` vertex mode for SK-VI.

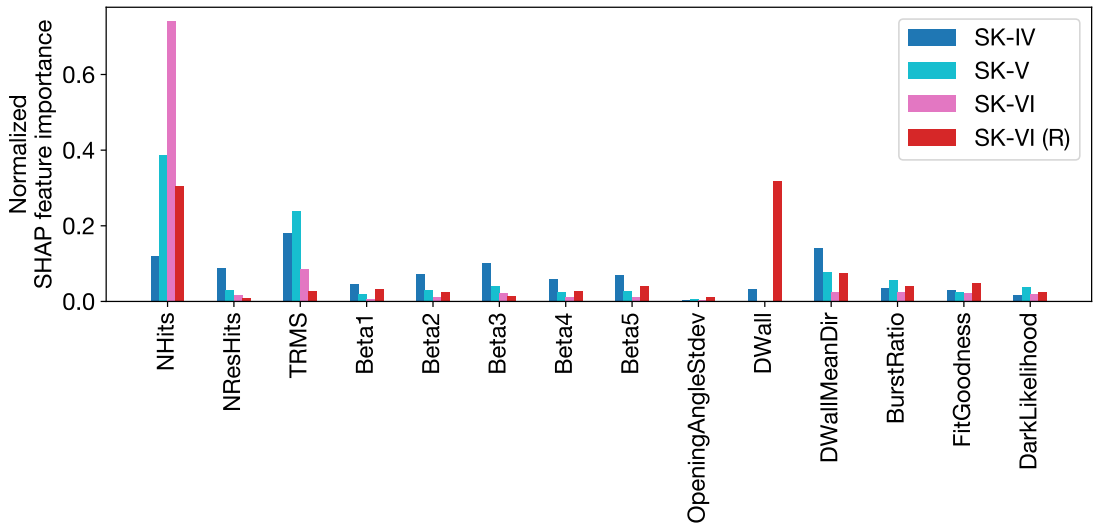


Figure 3.14: Comparison of normalized `SHAP` feature importance values across different SK phases.

We can also explore something similar to partial dependence plots. Figure 3.15 illustrates the partial dependence of the average neural network true positive rates (representing signal classification efficiencies) in relation to each feature. These plots indirectly reveal the features favored by the neural network, such as a high `NHits` count or a low `TRMS` value, for instance.

Most of the neural network preferences align with our expectations. Additionally, it is evident that neural networks tend to favor candidates with a larger fraction of signal hits (`SignalRatio`), even though they have not been explicitly trained on this. They also do not exhibit a preference for signal timing relative to the event trigger (`FitT`), a feature they have not been trained on either.

However, for certain features like `NResHits`, the neural network preference is entirely opposite between the pure water phases and the Gd-loaded phase. This discrepancy could be attributed to differences in gamma-ray multiplicity and total energy between $^1\mathrm{H}(n,\gamma)$ and $\mathrm{Gd}(n,\gamma)$ reactions.

From this point forward, "signal" candidates refer to candidates with a neural network output greater than 0.7.

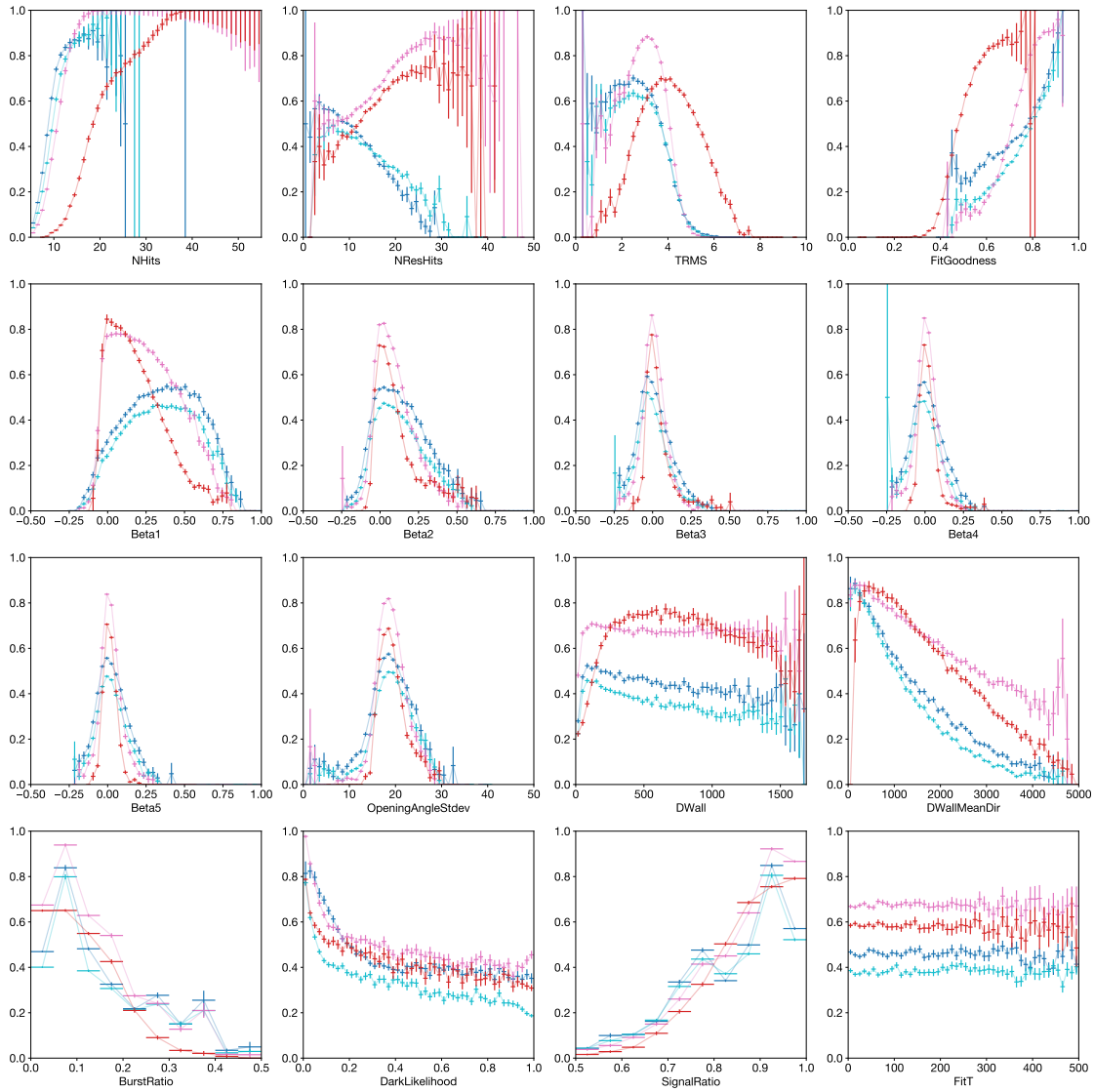


Figure 3.15: Partial dependence of neural network true positive rates (signal classification efficiency, vertical axes) for each feature, shown for SK-IV (blue), SK-V (cyan), SK-VI (pink), and SK-VI (R) (red). `SignalRatio` represents the ratio of signal PMT hits in each candidate, while `FitT` denotes the candidate time relative to the event trigger, both of which were not included in the training data.

3.3 Validation with Am/Be neutron source

3.3.1 Am/Be neutron source

The Am/Be (Americium/Beryllium) neutron source is a mixture of AmO_2 and Be powders enclosed in a thin stainless steel container, commonly used as a fast neutron point source. The radioactive isotope ^{241}Am undergoes alpha decay, a process with a half-life of 432 years. The resulting alpha particle, with a maximum kinetic energy of 5.5 MeV, collides with ^{9}Be, leading to the formation of a compound nucleus, $^{13}C^{*}$. This state quickly decomposes into $^{12}C^{*}$ and emits a neutron. Approximately half [133] of these emitted neutrons are accompanied by a 4.438 MeV gamma-ray. This detectable gamma-ray is produced during the de-excitation of the $^{12}C^{*}$ nucleus in its first excited state, enabling us to identify neutrons associated with this specific state. However, the remaining neutrons linked to other states (such as the ground state or higher excited states) either do not or rarely accompany detectable gamma-rays. This makes them an unwanted source of continuous pile-up. Figure 3.16 illustrates the expected kinetic energy distributions for neutrons in each excited state of $^{12}C^{*}$, ranging from approximately 1 to 8 MeV for gamma-ray associated neutrons.

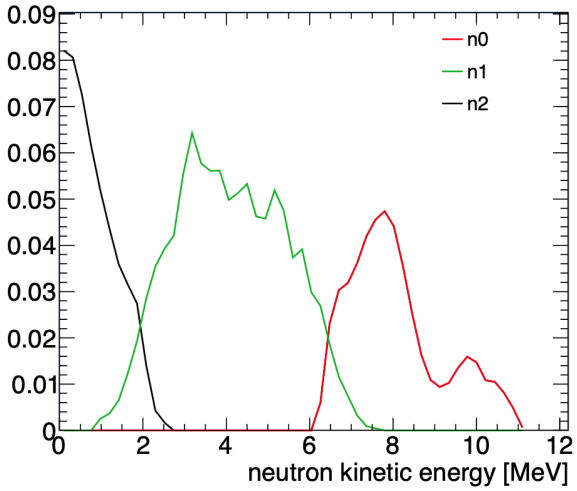
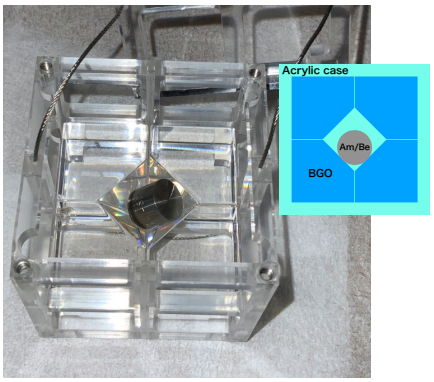


Figure 3.16: Kinetic energy distribution of neutrons from Am/Be source for each excited state of $^{12}C^{*}$, g.s. (red), 1st e.s. (green), and 2nd or higher e.s. (black) (Adapted from De Guarrini and Malaroda [134] and ISO8529-1 [135], reprinted from [136])

(a) Am/Be capsule



(b) Am/Be capsule surrounded by BGO scintillator crystals

Figure 3.17: Am/Be neutron source used for the calibration (left) and capsule surrounded by BGO scintillator crystals (right), which is the actual setup placed inside the tank for calibration (reprinted from [136]).

To assess the actual capability of our signal detection algorithm, as outlined in Section 3.2, and to confirm any differences between observations and our signal and detector simulation, we employed an Am/Be neutron source developed in the early 1990s, containing ^{241}Am with a strength of 97 μCi (3.6 MBq). This source was encapsulated within a stainless steel cylinder with an outer diameter and height of 12.5 mm, as illustrated in Figure 3.17a. The total neutron emission rate as well as the emission rates of neutrons corresponding to each excited state of ^{12}C* (including the ground state, 1$^{\text{st}}$ excited state, and 2$^{\text{nd}}$ or higher excited states) were thoroughly examined, as detailed in [137]:

- **Total neutron emission rate**: 236.8 $\pm$ 5.0 n/s

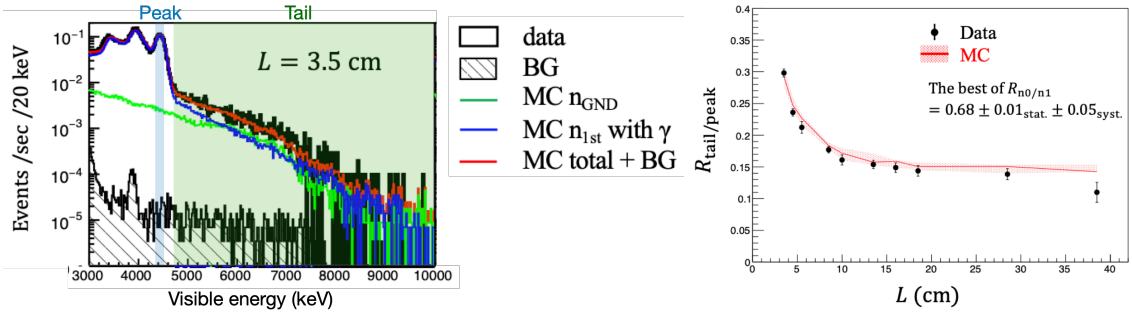
  This measurement was conducted at National Institute of Advanced Industrial Science and Technology (AIST) in Japan; the source was positioned at the center of a standard graphite pile, and the fully thermalized neutron fluence rate was measured using a ^{3}He proportional counter positioned 90 cm away from the center. The measured value was then compared with the reference value obtained with the standard Am/Be neutron source provided by AIST.

- **4.438 MeV γ / 1$^{\text{st}}$ e.s. n emission rate**: 110.1 $\pm$ 15.5 $\gamma(n)$/s

  The source was placed at distances of 10 mm and 50 mm from a high-purity Germanium (HPGe) detector. Energy spectra were measured and the gamma-ray peak was compared with Geant4 simulation. The 1$^{\text{st}}$ e.s. neutron emission rate was set to be equal to the measured 4.438 MeV gamma-ray emission rate.

- **Ground-state n emission rate**: 74.9 $\pm$ 11.9 n/s

  Deposited energy spectra in the NaI(Tl) crystal caused by the source activity as shown in Figure 3.18a were recorded using a PMT. The source was placed at various distances from the crystal, and the ratio $R_{\text{tail/peak}}$ of the high-energy tail (visible energy $>$ 5 MeV) to the 4.438 MeV gamma-ray peak was compared with Geant4 simulation to fit the neutron-to-gamma ratio. This ratio varies with distance L as shown in Figure 3.18b, due to lower likelihood of gamma-neutron coincidences which form a part of the high-energy tail. Since the contribution from neutrons from the 2$^{\text{nd}}$ or higher e.s. with kinetic energy below 3 MeV is negligible, the neutron-to-gamma ratio determines the ratio of the ground-state neutrons to the 1$^{\text{st}}$ e.s. neutrons of ^{12}C*.



(a) Observed energy spectrum

(b) $R_{\text{tail/peak}}$ vs. L

Figure 3.18: The energy spectrum of Am/Be source activity (left) and the tail-to-peak ratio $R_{\text{tail/peak}}$ as a function of source distance L (right). (reprinted from [137])

- **2$^{\text{nd}}$ or higher e.s. n emission rate**: 51.8 $\pm$ 17.2 n/s

  This value was estimated by subtracting the ground-state and 1$^{\text{st}}$ e.s. neutron emission rates from the total emission rate.

3.3.2 Calibration setup

The Am/Be neutron source was surrounded by Bismuth Germanate ($Bi_4Ge_3O_{12}$, BGO) scintillator crystals, as illustrated in Figure 3.17b. This configuration was designed to initiate scintillation responses induced by 4.438 MeV gamma-rays and to identify neutrons in the 1^{st} excited state of $^{12}C^*$ originating from the Am/Be neutron source. BGO is non-hygroscopic, allowing it to be safely immersed in water. The scintillation properties of BGO are outlined in Table 3.8. When a 4.438 MeV gamma-ray deposits all its energy within the scintillator, the ID PMTs register scintillation light equivalent to 1,000 - 2,000 photoelectrons in total.

The setup encased in Polymethyl methacrylate (PMMA, acrylic glass) was inserted into the tank through calibration holes located on the top. It was secured in a fixed position within the ID for an extended period, typically ranging from 30 minutes to 1 hour, during which PMT hits were recorded per each event triggers. While the triggering conditions have varied slightly over the course of more than 10 years of operation, PMT hits within 535 μs (or 835 μs for SK-IV) from the event trigger were logged for SHE-triggered events. This allowed for subsequent searches for neutron signals within the time frame.

It is important to consider that the calibration setup, including the use of BGO scintillators and the constant pile-up background neutrons stemming from the source, may introduce non-negligible effects on the signal efficiency estimate. For instance, BGO ($Bi_4Ge_3O_{12}$) has an effective neutron capture cross section of 1.47 barns [138]. The total excited energy resulting from neutron captures on ^{76}Ge or ^{74}Ge is 6.072 MeV and 6.505 MeV, respectively. These captures inside the scintillator, which can mimic the 4.438 MeV gamma-ray scintillation and consequently lead to false signal events, may account for a few percent of genuine signal events, considering the background source intensity and the relative capture rates on Ge, as illustrated in Figure 3.19. Furthermore, multiple gamma-rays from Gd(n, γ) reactions of piled-up neutrons in Gd-loaded water, in case roughly half of the energy is accidentally deposited in the scintillator, can also mimic the signal event trigger. We assessed these potential nuisance effects using MC simulations conducted in Geant4.

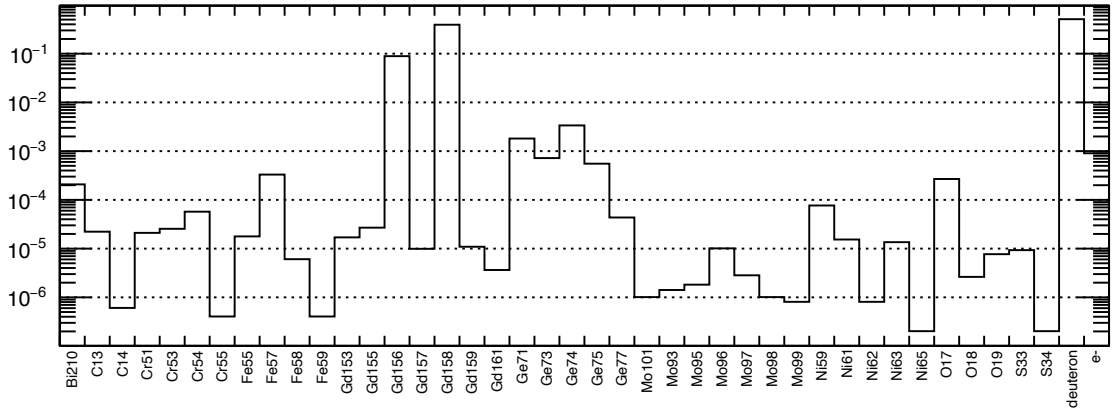


Figure 3.19: The thermal neutron capture probabilities on nuclides (shown are the final states) existent in the SK-VI Am/Be calibration setup, simulated with Geant4.10.5.p01 [139, 140] and ENDF/B-VII.1 [31]. (Reprinted from [136])

3.3.3 Simulation

Neutron particle-gun simulation with `skdetsim` in Geant3

To evaluate the neutron signal simulation capabilities of the detector simulator `skdetsim` (Geant3.21 [141]) used in atmospheric neutrino event simulations, the base MC simulation for comparison with Am/Be data was generated using `skdetsim` v15p2. Unfortunately, due to practical challenges in implementing the calibration setup in the no-longer-maintained Geant3 in FORTRAN, the simulation was conducted in a particle-gun setup. The configuration settings, including the detector parameters tuned for each SK phase, as well as the inclusion of background PMT hits extracted from randomly triggered data events, were similar to the thermal neutron simulation described in Section 3.2.2 and the ones outlined in Table 3.3 which were used for training and evaluating the signal detection algorithm.

The notable differences included the neutron kinetic energy, which followed the standard (the green distribution in Figure 3.16) corresponding to gamma-associated neutrons from Am/Be, and the fixed vertex position set to the source position of each calibration run. Additionally, no background event cuts were applied when appending background hits to the simulated signal events.

For this simulation, no consideration was given to factors such as the scintillator geometry, constant pile-up from the source, or resulting trigger dead time effects. The input Gd concentration for SK-VI was set to 0.0110 w%.

Full simulation of Am/Be calibration setup with `SKG4` in Geant4

We used `SKG4` v2.4.3, based on Geant4.10.5.p01 [139, 140], for the comprehensive simulation of the Am/Be calibration setup. This included modeling the stainless steel capsule, BGO scintillator crystals, and plastic casing, while considering the complete kinetic energy distributions of both gamma-ray associated and non-associated neutrons, as illustrated in Figure 3.16. To accurately replicate the BGO scintillation response to 4.438 MeV gamma-rays, we fine-tuned the input optical parameters for BGO by comparing the simulation results against Am/Be data.

For the tuning of BGO optical parameters, we used Am/Be data collected with the source at the center of the tank during the SK-V phase. Refractive indices and photon wavelength spectrum were sourced from [142] and [138], respectively. We generated MC simulations using various light yield parameters and estimated the peak value of the integrated charge spectrum. For each simulated event, the integration range was set to [-0.5, 5] μs from the simulation's T_0 to fully encompass all scintillation light with a decay constant of 300 ns. Then, we applied linear regression to establish the relationship between light yield and corresponding peak values. The regression was used to estimate the light yield corresponding to the peak value observed in the Am/Be data. The light yield value was further refined by modeling the peak as a combination of Gaussian and linear background components. Finally, we fine-tuned the Fano factor to match the observed peak width. This process is illustrated in Figure 3.20. The tuned light yield was 5,935 photons per MeV γ, which was significantly lower than the reference value of 8,400 photons [138].

The scintillation decay time was adjusted based on the pulse shape of 4.438 MeV gamma-ray scintillation, averaged among events with integrated charge within 1,250 to 2,000 photoelectrons. This involved applying TOF correction to the PMT hit time distribution, which was then weighted by deposited charge per event.

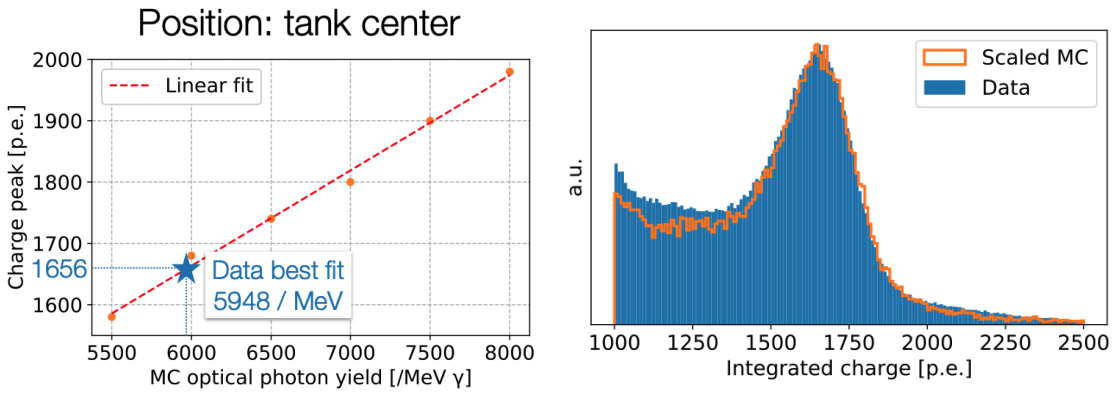


Figure 3.20: Correlation between the input BGO scintillation photon yield and the simulated peak position in the integrated charge spectrum (left) and a comparison between the integrated charge spectrum obtained from Am/Be data and the corresponding spectrum from the fine-tuned MC simulation (right). The minor variation observed in the lower tail is attributed to event pile-ups, a factor not yet incorporated in the MC simulation at this stage.

Following the approach outlined by Gundacker et al. [143], we fitted the average pulse shape with the following function:

$$f(t|\theta) = \Theta(t-\theta) \sum_{i\in\{\text{fast, slow}\}} \frac{e^{-(t-\theta)/\tau_{d,i}} - e^{-(t-\theta)/\tau_{r,i}}}{\tau_{d,i} - \tau_{r,i}} \cdot \rho_i \tag{3.8}$$

Here, t represents the PMT hit time, Θ is the Heaviside step function, θ is the start time of scintillation, τ_d is the decay time constant, τ_r is the rise time constant, and ρ_i is the fraction of the fast/slow component. The timing response of the PMTs on the order of $O(1)$ ns was neglected, as the pulse shape time scale was on the order of $O(1)$ μs.

The best-fit values were $\tau_{d,\text{slow}} = 433$ ns and $\tau_{d,\text{fast}} = 110$ ns, which aligned with the reference [144] for a water temperature of 13.5°C. The fitted fast component ratio was 12%, also in agreement with the reference value of 10% [145]. The fitted rise time values were on the order of $O(1)$ ns, which was negligibly small, so they were assumed to be 0 in the actual simulation input.

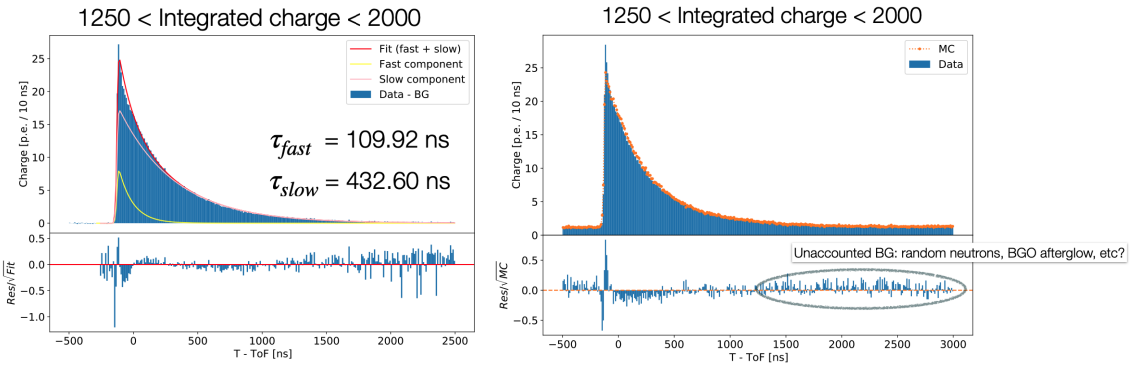


Figure 3.21: The average BGO scintillation pulse shape observed in the Am/Be data, along with the corresponding fitted pulse spectrum (left), and a comparison between the average scintillation pulse shapes from the data and the fine-tuned MC simulation (right).

SCINTILLATIONYIELD	5,935 photons / MeV γ
RESOLUTIONSCALE	0.86
FASTTIMECONSTANT	109.92 ns
SLOWTIMECONSTANT	432.60 ns
YIELDRATIO	0.1191

Table 3.7: Tuned BGO optical parameter inputs for Geant4.

Light yield	8,400 photons / MeV γ
Emission peak	480 nm
Rise time	$<$ 2 ns
Decay time	300 ns
Radiation length	1.1 cm
Temperature response	-1.2%/°C
Afterglow (after 20 ms)	150 ppm

Table 3.8: Reference values for BGO properties (Room temperature, [138])

The fit results and the data-MC comparison of the average BGO scintillation pulse shapes are displayed in Figure 3.21. Table 3.7 provides a summary of the adjusted BGO input optical parameters for Geant4.

At this point, the simulator can only simulate 'discrete' events of signals, such as a single instance of $n + \gamma$ (4.438 MeV) corresponding to the 1^{st} excited state of ^{12}C or just n corresponding to the other states of ^{12}C. In reality, these events occur continuously from the source, and a single triggered event may contain multiple pile-ups of gamma-rays and neutrons. Some of these continous emissions, through Ge(n, γ) or Gd(n, γ) reactions of 6-8 MeV for example, may eventually trigger an event that is indistinguishable from a genuine signal event — specifically, a 4.438 MeV gamma-ray scintillation-triggered event with a single neutron.

To accurately reproduce these effects, we developed a dedicated pile-up simulator, which reorganizes the 'discrete' simulated events of each neutron emission mode of Am/Be on a single global 'continuous' time axis, based on the emission rate of each mode as described in Section 3.3.1. The simulated PMT hits and randomly selected background PMT hits are rearranged on the global time axis, and event triggers are applied to generate events. This approach enables us to faithfully replicate both the constant source activity and the actual event triggering scheme used in our data. We performed a comprehensive simulation of the Am/Be calibration setup using a combination of `SKG4` v2.4.3 with finely-tuned BGO scintillation parameters, along with the pile-up simulator for each source position. The simulation was used to evaluate any potential adverse impacts of the calibration setup on the estimation of signal efficiency, thereby complementing the comparison of Am/Be data with the `skdetsim` neutron particle-gun simulation, which does not include such features.

As an additional point, the input Gd concentration was adjusted from the actual estimated value of 0.0110 w% to 0.0104 w% to align the simulated neutron capture time with the observations in the Am/Be data. This adjustment was made by another analysis group to compensate for the overestimation of Gd(n, γ) reactions in Geant4.10.5.p01 which takes neutron capture cross sections from ENDF/B-VII.1.

In the case of SK-IV/V, candidates were selected within the time range of [18, 534] μs from the event trigger, effectively eliminating the possibility of PMT afterpulse or Michel electron contamination. For SK-VI, which had a shorter capture time, a broader window of [3, 534] μs was used. The assessment of PMT afterpulse and Michel electron contamination for SK-VI is necessary and is detailed in Section 3.4 through 3.5. The rationale for the different search ranges arises from a technical issue (which could be addressed in the future), rather than a consideration related to physics. The initial TOF correction was determined based on the source position. Table 3.9 summarizes the signal search time range used for each SK phase.

				Signal search time range	
Target	Phase	Vertex mode	Trigger	Start	End
Pure water	SK-IV	INIT	0 μs	18 μs	534 μs
	SK-V				
Gd-water	SK-VI			3 μs	
		RECO			

Table 3.9: Summary of the signal search time range.

3.3.4 Data reduction

Data reduction involved three main steps: (1) selecting 4.438 MeV gamma-ray scintillation, (2) reducing cosmic rays, and (3) minimizing pile-ups. The following section explains the relevant 7 event cuts applied to the Am/Be data. As an example, figures comparing the cut variables of SK-VI data with the source positioned at the tank center for 1.1 hours (Run 85622, 4,002 seconds) and the corresponding MC simulation for a livetime of 1.7 hours (6,000 seconds) are provided.

In each figure, the legend label '(all)' denotes histograms before applying any cut, while '(cut)' indicates histograms after applying all previously mentioned cuts, excluding the one being discussed. The shaded areas in gray represents the data events that were rejected due to the discussed cut.

Cut #1: Selecting SHE and AFT event triggers

The full 4.438 MeV gamma-ray scintillation is expected to activate the SHE trigger which has the highest threshold in SK. Due to the deadtime resulting from electronic issues and pile-ups, SHE-triggered events are accompanied by AFT triggers — an additional 500-μs record of PMT hits — roughly half of the time. Events with both SHE and AFT triggers were selected, as shown in the figure below. This selection allows us to search for neutron capture signals with a characteristic decay time constant of $O(100)$ μs. Figure 3.22 shows the trigger type distribution expected and observed with Am/Be neutron source deployed near the tank center.

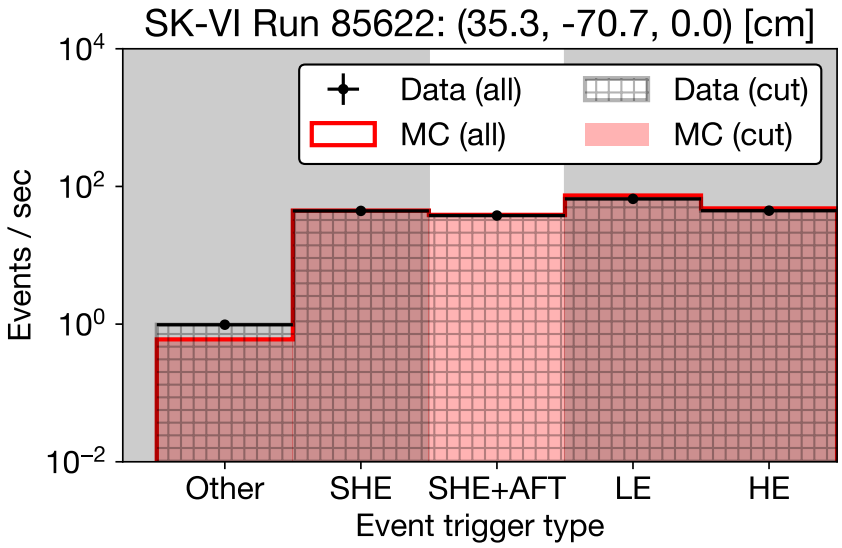


Figure 3.22: Event trigger type with Am/Be neutron source in SK tank.

Cut #2: Rejecting event triggers due to cosmic-rays

To identify the entry of a cosmic ray or the exit of an outgoing particle, we employ a clustering algorithm for the OD PMT hits. These hits are binned by the azimuthal and polar angles. For each bin, we calculate the gradient of total charge with respect to each neighboring grid. The bin is then paired with its neighboring bin that shows the highest charge gradient. Events with more than 10 OD PMT hits in the largest cluster (`NHITAC`) were rejected, as they may suggest the entry of a cosmic ray muon. The distribution of `NHITAC` is shown in Figure 3.23.

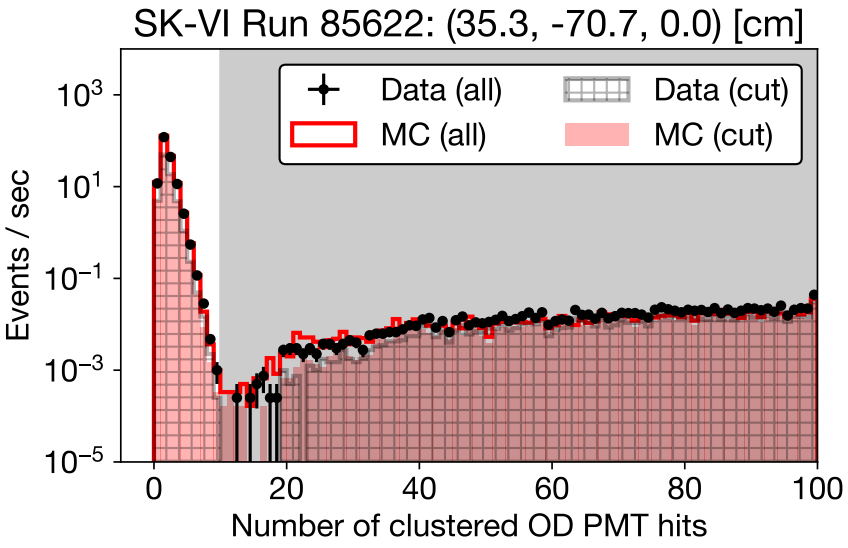


Figure 3.23: Number of clustered OD PMT hits.

Cut #3: Selecting event trigger from 4.438 MeV γ scintillation

The total deposited charge within the time window of [-0.5, 1.0] μs from the event trigger was integrated for each event to estimate the energy content of the triggering signal. The resulting distribution exhibits a distinct peak attributed to the 4.438 MeV gamma-ray scintillation, along with tails arising from neutron captures and inelastic scattering within the scintillator, as illustrated in Figure 3.25. The peak charge, measured in the number of photoelectrons, is determined from a histogram within the range of [500, 1500] photoelectrons, divided into 100 bins. Events falling within $\pm$200 photoelectrons around the estimated peak were subsequently selected. Figure 3.24 shows the distribution.

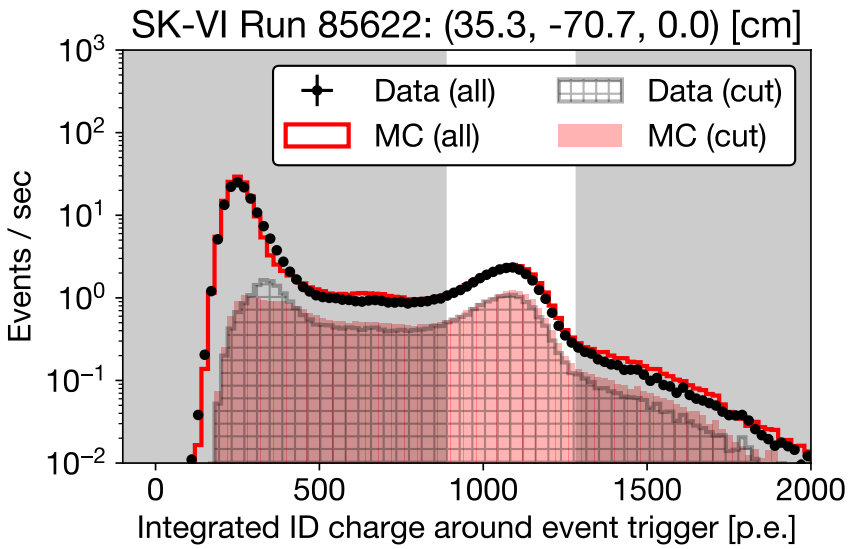


Figure 3.24: Integrated ID charge with Am/Be neutron source in SK tank.

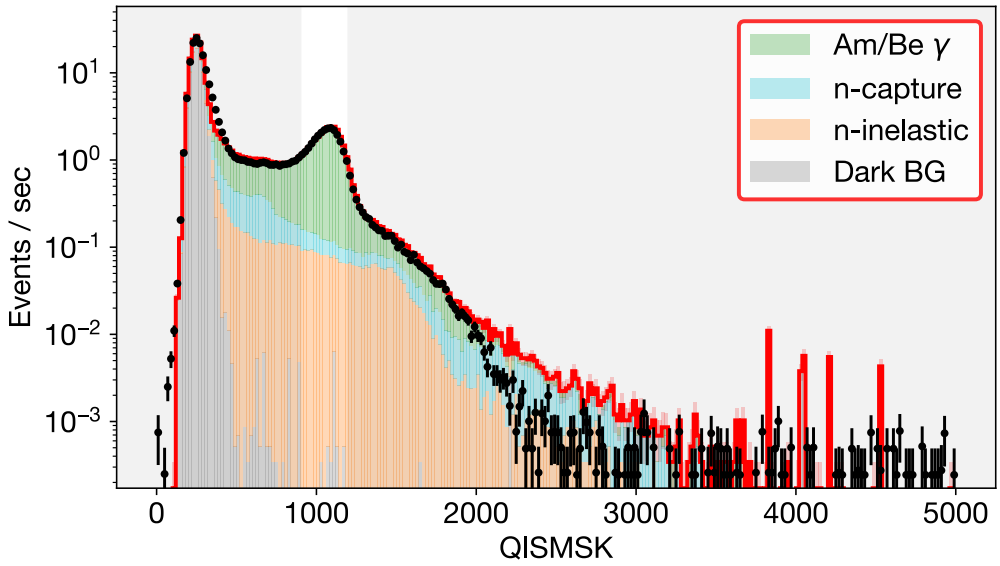


Figure 3.25: Integrated charge (`QISMSK`) distribution for SK-VI Am/Be data (black) and corresponding MC simulation (red) for Run 85622, prior to applying any cuts, along with a breakdown of triggering events. While the fraction of Am/Be γ triggers (green) is $\sim 95\%$, $\sim 5\%$ of the remaining events are expected to triggered by backgrounds, such as neutron capture gamma-rays (blue) or neutrons (orange) interacting within the scintillator.

Cut #4: Rejecting event triggers too close to the previous trigger

For two events in close proximity, there is a chance of late neutron captures from the prior trigger interfering with the subsequent event. To mitigate this, events triggered less than 1 ms after the previous trigger were rejected. Figure 3.26 shows the distribution of the time difference between two consecutive events.

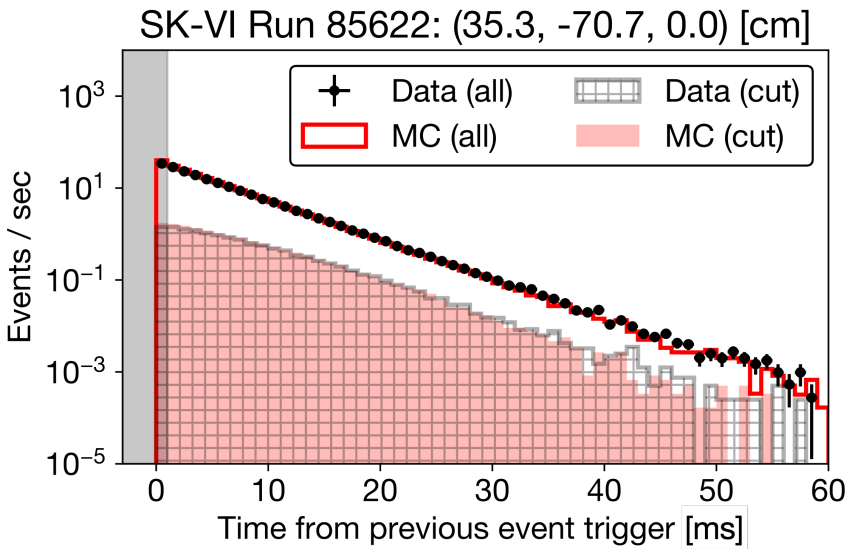


Figure 3.26: Time difference between two consecutive event triggers.

Cut #5: Rejecting cosmic rays within the signal search range

Cosmic rays can penetrate the detector even during events triggered by Am/Be gamma-rays and affect the signal efficiency estimate. Events with more than 14 OD PMT hits within a 200-ns window in the signal search range were rejected. Figure 3.27 shows the distribution of the expected and observed maximum OD PMT hits.

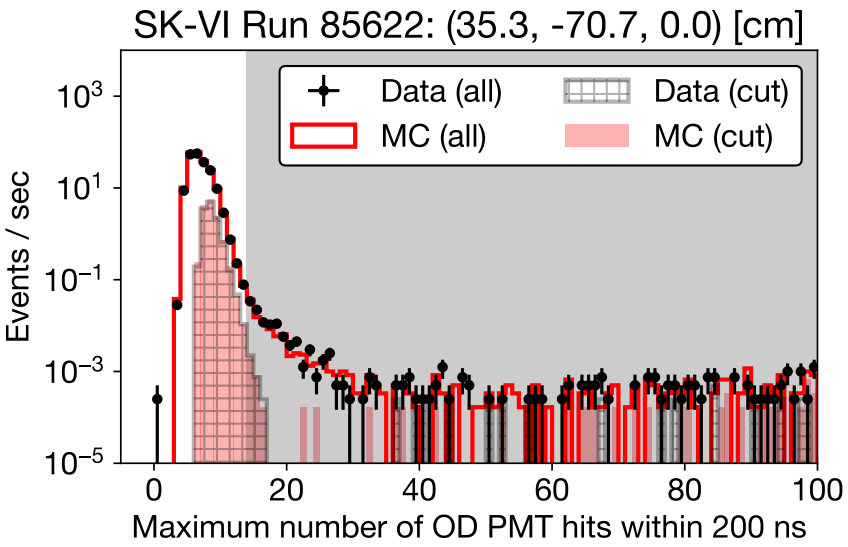


Figure 3.27: Maximum number of OD PMT hits within a 200-ns window in the signal search range. Events with a larger value implies a cosmic ray entry within the signal search range.

Cut #6: Rejecting events with pile-up scintillation

Scintillation resulting from pile-up gamma-rays or neutron captures can impact the signal efficiency estimate. For each event, the maximum number of ID PMT hits within a 200-ns window in the signal search range was determined, as shown in Figure 3.28. To find the most probable value, a histogram was generated in the range of [0, 500] PMT hits with 100 bins, and the peak bin value was derived. Minor peaks near 500 hits and 1,000 hits indicate scintillation caused by the 2.2 MeV gamma-ray from the $^1\mathrm{H}(n,\gamma)$ reaction and the 4.438 MeV gamma-ray from the Am/Be neutron source, respectively. Events with more than 50 hits beyond the estimated peak were rejected to mitigate such contamination.

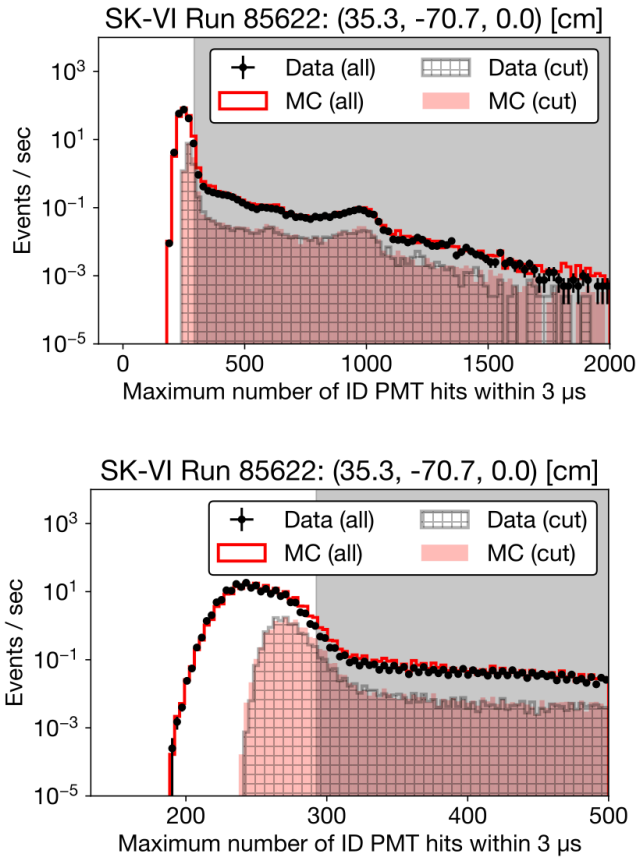


Figure 3.28: Maximum number of ID PMT hits within a 3-μs window in the signal search range. Events with a larger value implies a pile-up causing scintillation within the signal search range. Both figures show the same histograms, but with different horizontal axis ranges.

Cut #7: Selecting events with 535-μs duration

Although rare, some data events, even after the event trigger type selection, may have an event duration shorter than 535 μs, potentially due to electronics issues. The total number of PMT hits in the signal search range follows a normal distribution, with its mean determined by the average dark rate at the time of measurement. It is important to note that data from a specific run and the MC with background hits spanning an entire SK phase may exhibit different distributions. Outlier events with an unusually low number of hits may indicate abnormal events with shorter durations. The peak position was estimated in the histogram within the range of [0, 100,000] PMT hits, divided into 1,000 bins. Events falling within ±3,000 hits from the peak value were selected to ensure an event length of 535 μs.

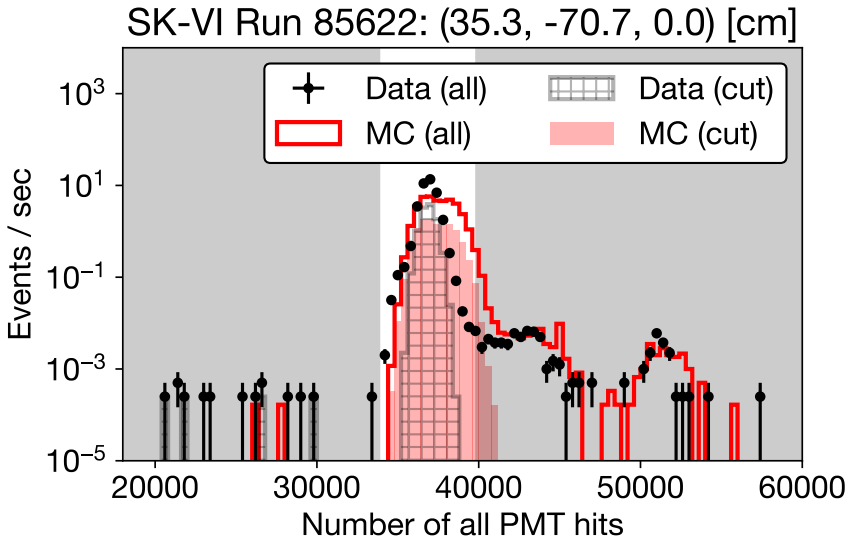


Figure 3.29: The distribution of all PMT hits within the signal search range, primarily influenced by the PMT dark rate. The disparity between data and MC results stems from the random selection of background PMT hits throughout each SK phase in the MC simulations. In contrast, the data specifically reflects the dark rate measured during the 0.5-1 hour timeframe. Outlier events resulting from data acquisition problems or background contamination were excluded from the analysis.

The event rates before and after applying each cut are outlined in Figure 3.30. The event rates in general show agreement between the data and the MC, while small difference in **Cut #3** reduction rates suggests there is potential for improvement in simulating both the Am/Be source activity and the scintillator response.

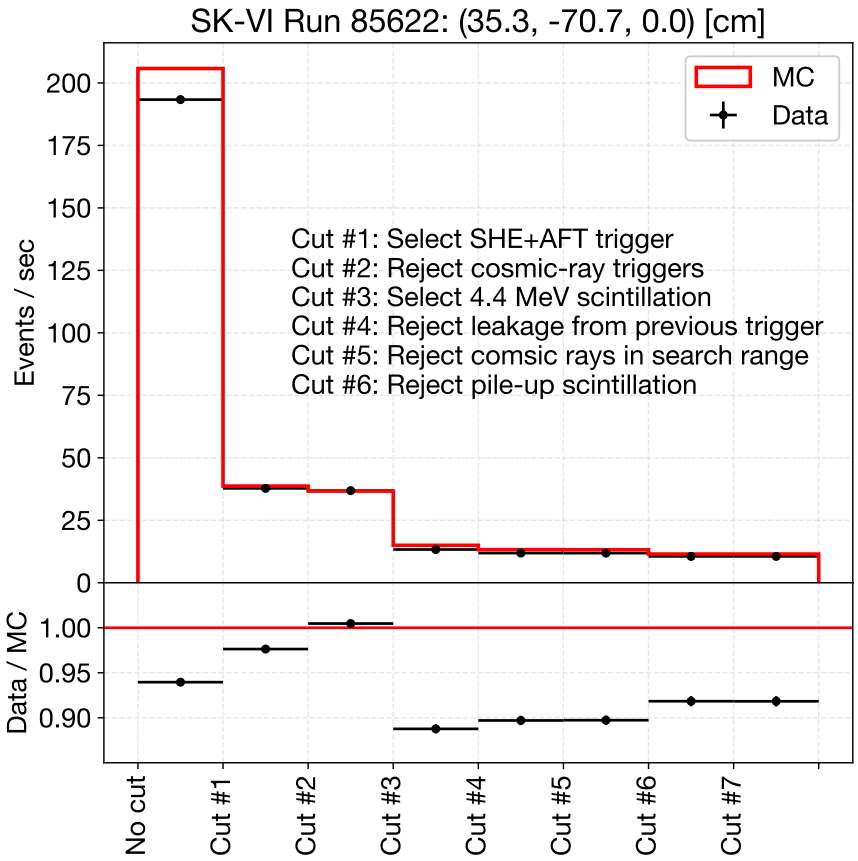


Figure 3.30: Am/Be event rates compared between data and MC.

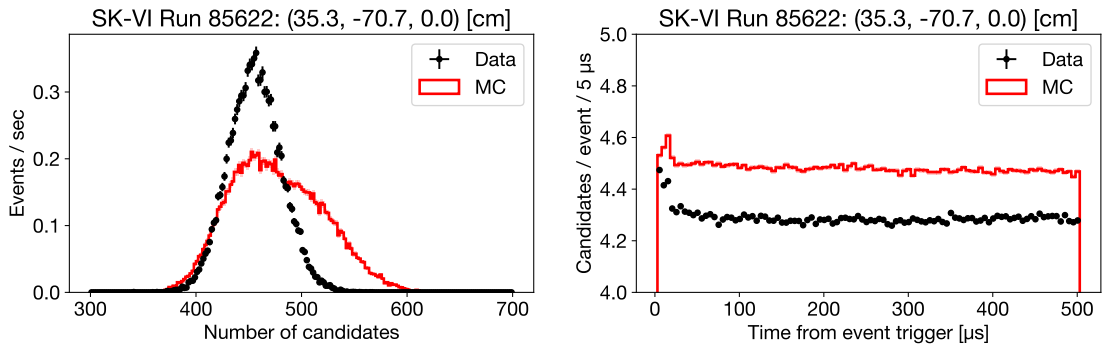


Figure 3.31: The distribution of total number of candidates in each event (left) and the candidate time distributions (right) of SK-VI Am/Be data and simulation.

Figure 3.31 illustrates the number of candidates and their corresponding time distributions in the selected Am/Be data and MC events. Specifically, for Run 85622, the overall number of candidates in the Am/Be data is 5% smaller compared to the MC, as the MC incorporates background hits from various times across the SK phase, each with differing dark rates. The candidate time distributions are in general flat over time. The 5% uptick in the number of candidates within 15 μs from the event trigger is attributed to PMT afterpulses, as also demonstrated in MC.

Figures 3.32 through 3.36 display the distributions of candidate features and network outputs from Am/Be data and MC simulations for the source positioned at the tank center, across various SK phases. Notably, the distributions from both the data and MC show excellent shape agreement.

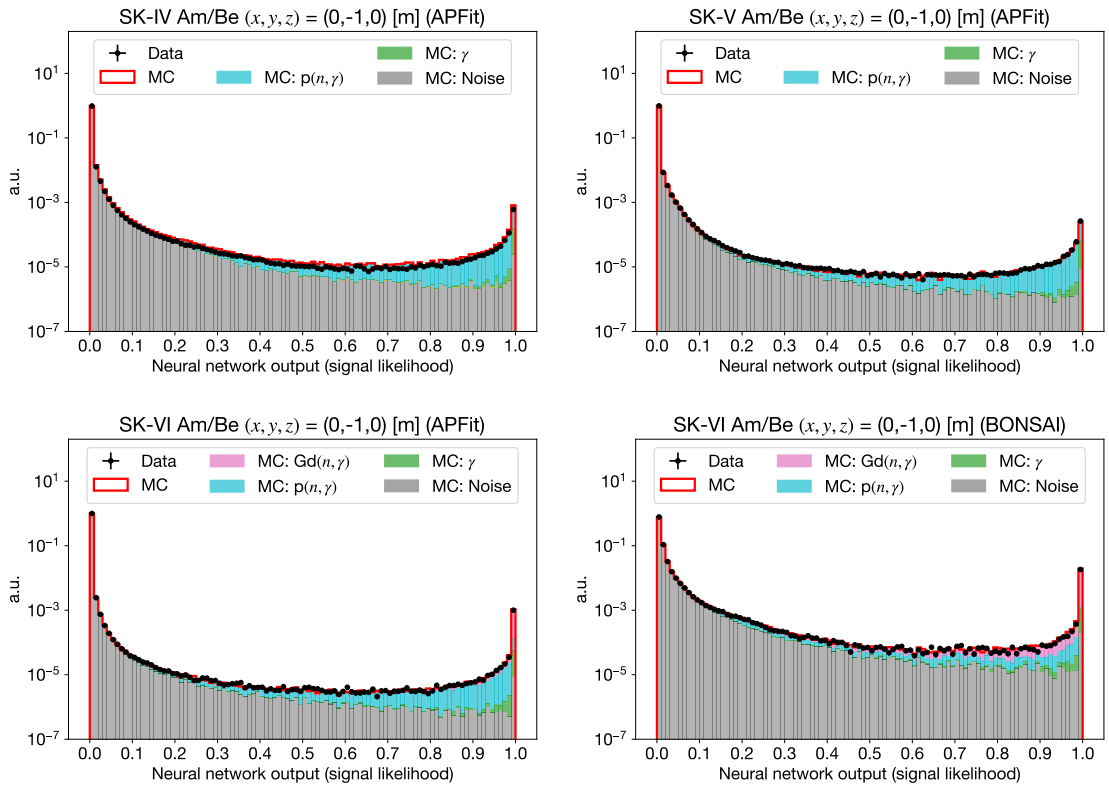


Figure 3.32: The neural network responses for candidates selected in both Am/Be data and full Am/Be MC simulation with `SKG4`, for SK-IV (top left), SK-V (top right), SK-VI (bottom center), and SK-VI (R) (bottom right). Both data and MC histograms are normalized by area. The colors in the legend indicate the fraction of labels derived from the full Am/Be calibration simulation. The label "MC: γ" corresponds to the 4.4-MeV gamma-ray emitted by the source.

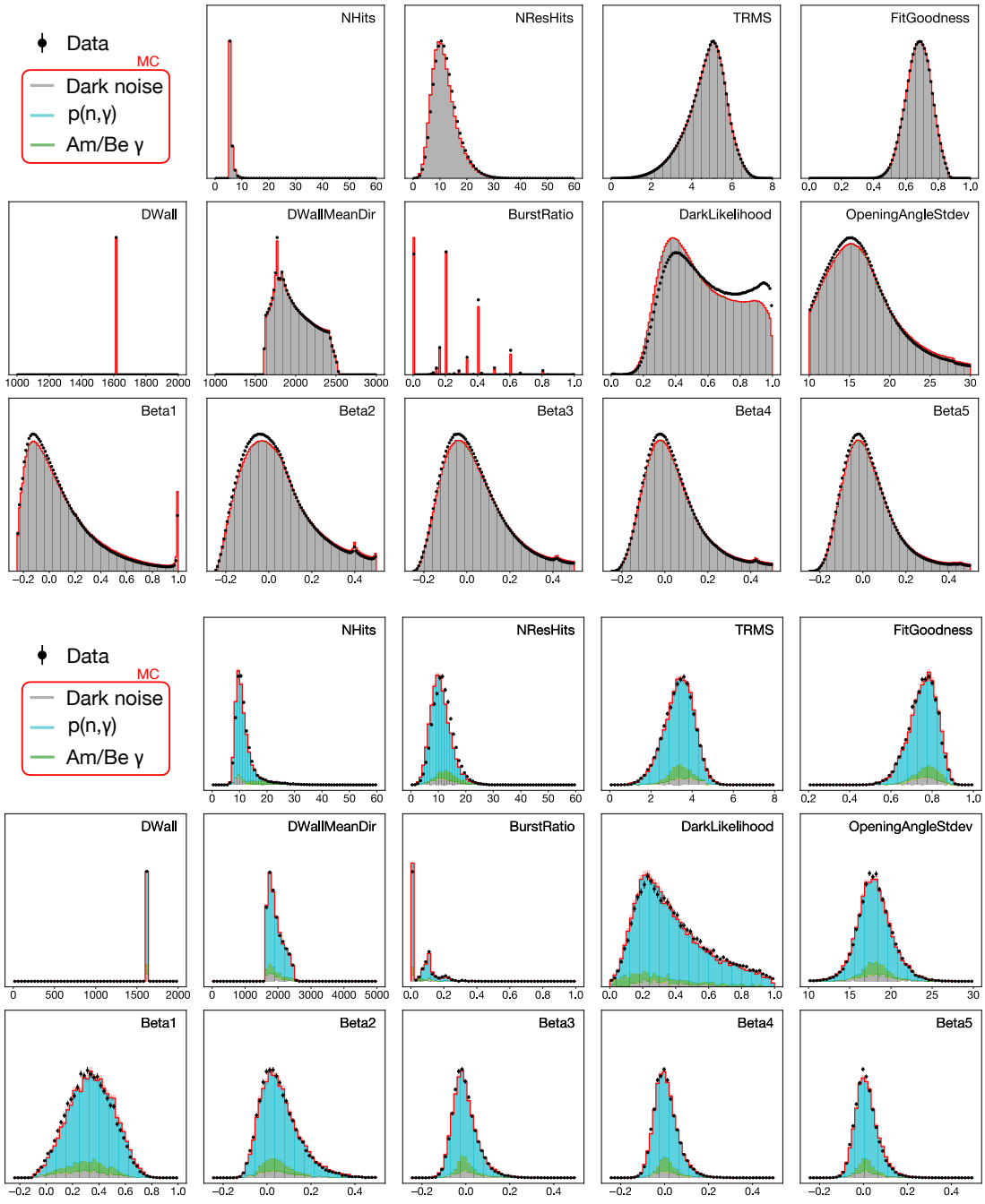


Figure 3.33: The feature distributions of signal candidates before (top) and after (bottom) the neural network selection for both the Am/Be data (Run 75502) and MC simulations, taken with the Am/Be neutron source at the tank center in SK-IV. Both data and MC histograms are normalized by area.

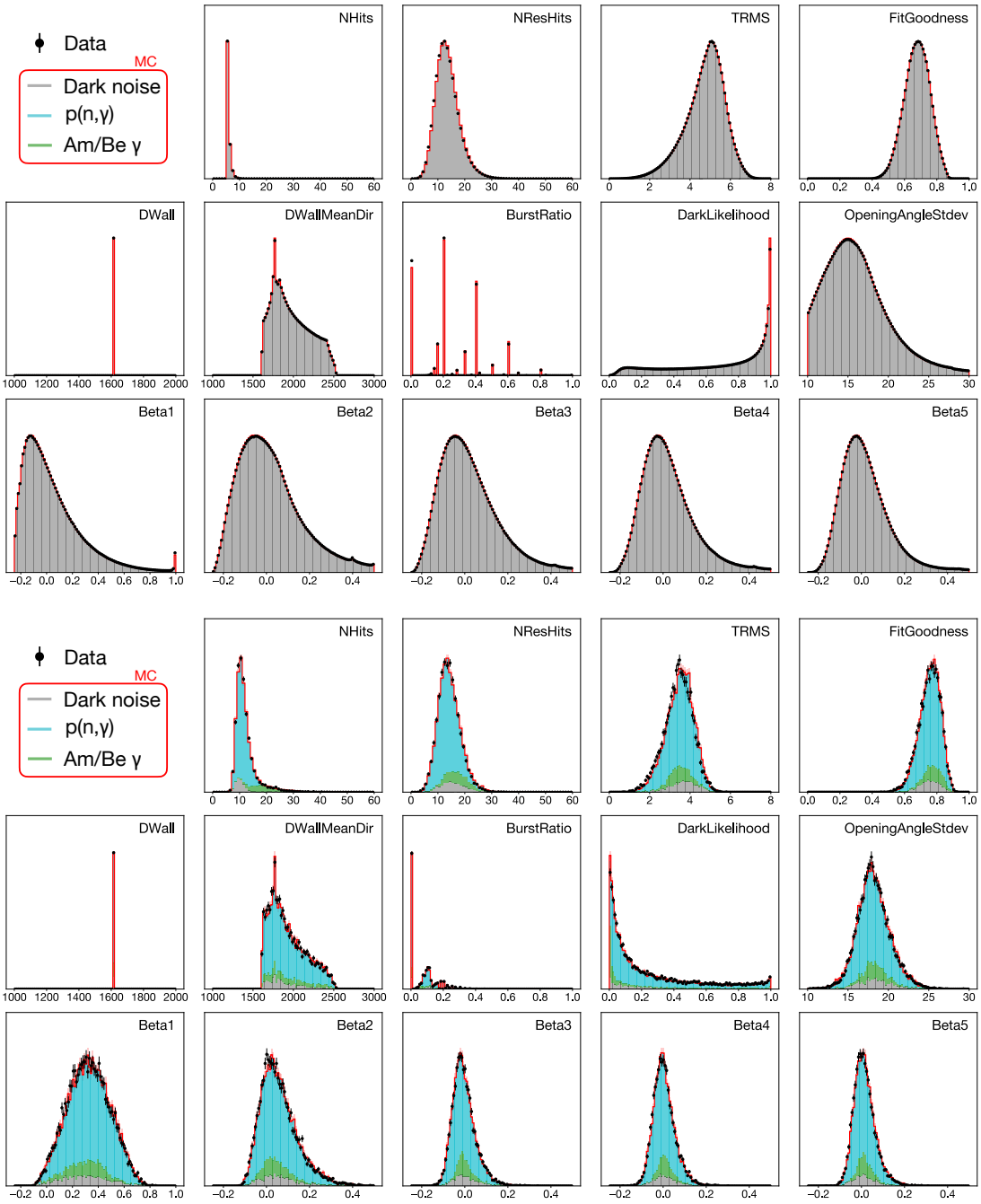


Figure 3.34: The feature distributions of signal candidates before (top) and after (bottom) the neural network selection for both the Am/Be data (Run 82406) and MC simulations, taken with the Am/Be neutron source at the tank center in SK-V. Both data and MC histograms are normalized by area.

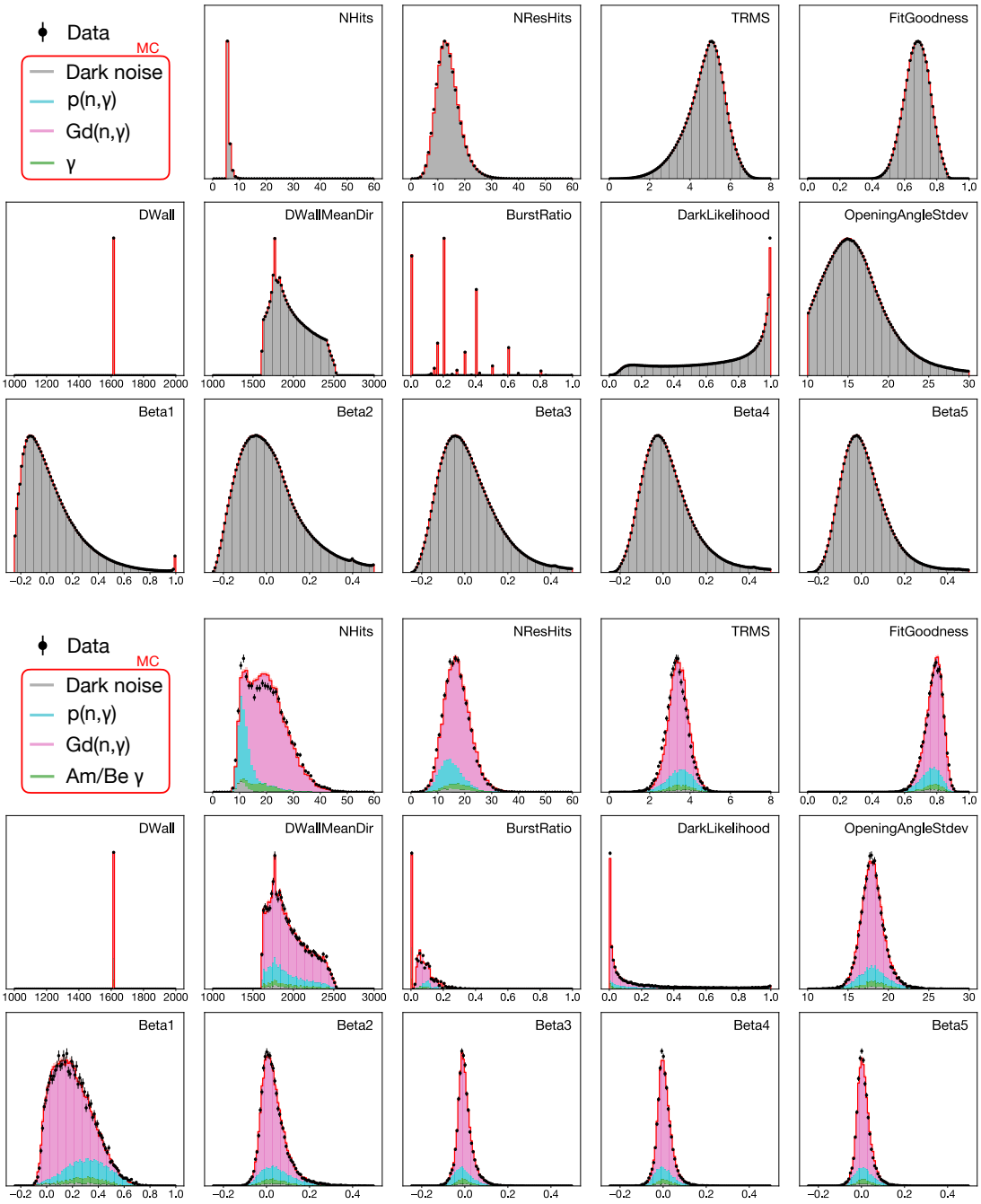


Figure 3.35: The feature distributions of signal candidates before (top) and after (bottom) the neural network selection for both the Am/Be data (Run 85622) and MC simulations, taken with the Am/Be neutron source at the tank center in SK-VI. Both data and MC histograms are normalized by area.

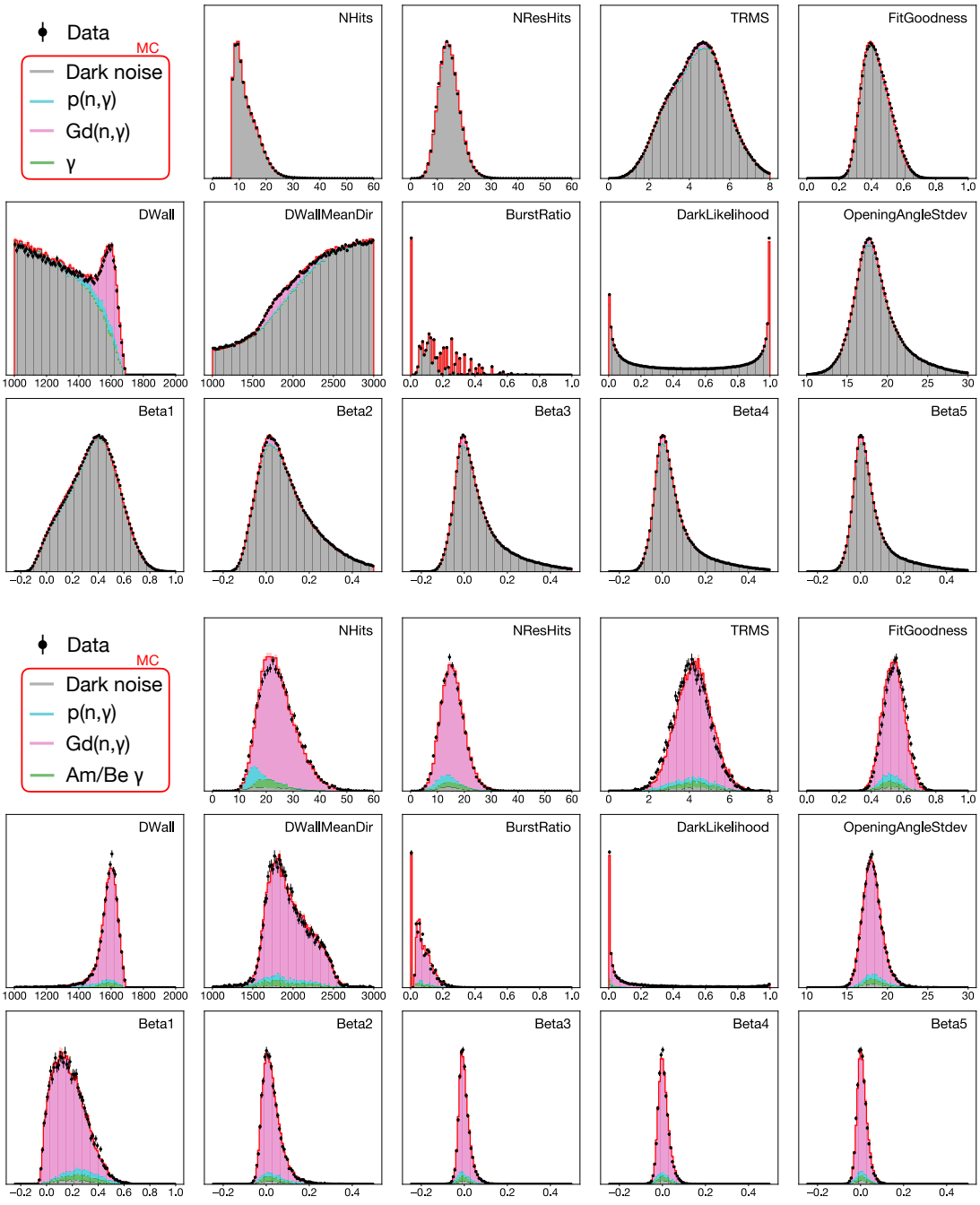


Figure 3.36: The feature distributions of signal candidates before (top) and after (bottom) the neural network selection for both the Am/Be data (Run 85622) and MC simulations, taken with the Am/Be neutron source at the tank center in SK-VI (R). Both data and MC histograms are normalized by area.

3.3.5 Neutron tagging efficiency estimation

When comparing Am/Be data and MC, we utilize the neutron tagging efficiency, which is calculated based on the total number of neutrons rather than the number of neutron captures. This approach simplifies the analysis by avoiding additional complexities. The neutron tagging efficiency ϵ represents the ratio of identified neutron capture signals to the overall number of neutrons. In the case of Am/Be data events, it can be approximated using the following formula:

$$\epsilon \approx \frac{\text{Candidates classified as signals} - \text{Estimated backgrounds}}{\text{Selected event triggers}} \tag{3.9}$$

For the denominator, we assume that each event trigger corresponds to the production of one neutron, accompanied by a 4.438 MeV gamma-ray scintillation.

To estimate the backgrounds in the numerator, we fit the time distribution of candidates classified as signals with an exponential curve given by:

$$f(t) = A(1 - e^{-t/\tau_{\text{thermal}}})e^{-t/\tau_{\text{capture}}} + B \tag{3.10}$$

Here, A, τ_{thermal}, τ_{capture}, and B are the fitted parameters. A represents the scaling coefficient, τ_{thermal} denotes the neutron thermalization time constant, τ_{capture} denotes the neutron capture time constant, and B is the constant offset. This offset can be interpreted as the constant background, resulting from pile-ups and dark background hits mis-identified as signals. The number of backgrounds is estimated by multiplying the number of bins by B and the number of selected event triggers.

Figure 3.37 provides an illustrative example of such a fit applied to the time distribution of signal candidates observed in the Am/Be data. For the pure water phases, we set τ_{thermal} to 0, since most neutrons are already thermalized within the search time window of [18, 535] μs from their production. Besides, the neutron capture probability in pure water is largely independent of neutron kinetic energy, due to the $1/v$-dependence of the $^1\text{H}(n, \gamma)$ cross section.

It is worth noting that pile-up backgrounds actually have time dependence in the form of $(1 - e^{-t/\tau_{\text{capture}}})$ as shown in Figure 3.38. This is because neutrons passing through the scintillator may initiate an event trigger, preventing their captures from leaking into the subsequent event. This leads to overestimation of B, which cannot be reproduced using a 'discrete' event generator alone.

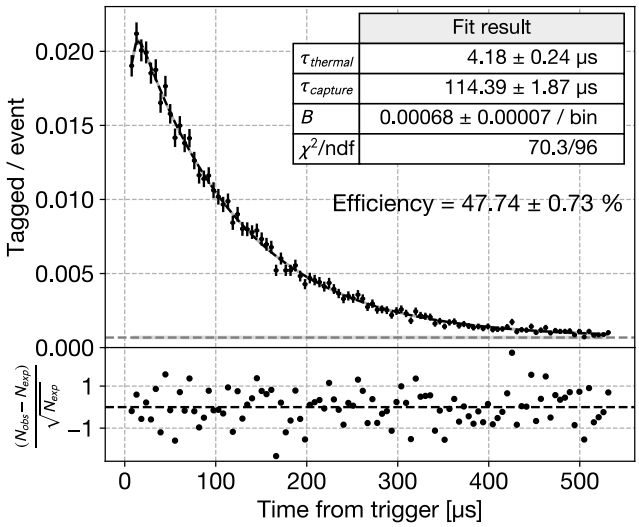


Figure 3.37: Fitted signal candidate time distribution for SK-VI Am/Be data, with source positioned at the tank center.

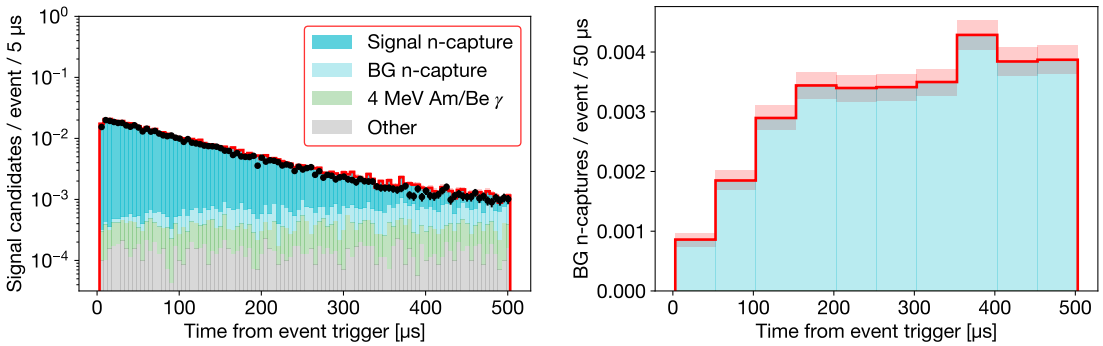


Figure 3.38: Comparison of time distributions for signal candidates between Am/Be data and MC (left), and the time distributions showing only the MC background neutron captures from pile-ups (right). The MC signal neutron captures correspond to neutrons accompanying a 4.438 MeV gamma-ray.

3.3.6 Systematic uncertainties in simulation model

We compared the estimated neutron tagging efficiency derived from both Am/Be data and MC simulation. For data estimation, only statistical uncertainty was considered. All major systematic uncertainties were propagated to our simulation model prediction. The propagated uncertainties were evaluated by contrasting the true neutron tagging efficiencies obtained from a nominal model, which is the same as neutron particle-gun setup in `skdetsim` described in Section 3.3.3, with those from a varied model. Below, we provide a description of each source of systematic uncertainty and its influence on the signal efficiency estimate.

Neutron kinetic energy

The Am/Be neutron source emits fast neutrons with kinetic energy on the order of O(1) MeV. In this energy range, the inelastic alpha emission process, specifically $^{16}\mathrm{O}(n,\alpha)^{13}\mathrm{C}$, becomes non-negligible as shown in Figure 3.39. This alpha emission is not visible in the detector and competes with the process of elastic collisions with free protons, which serves to moderate neutrons.

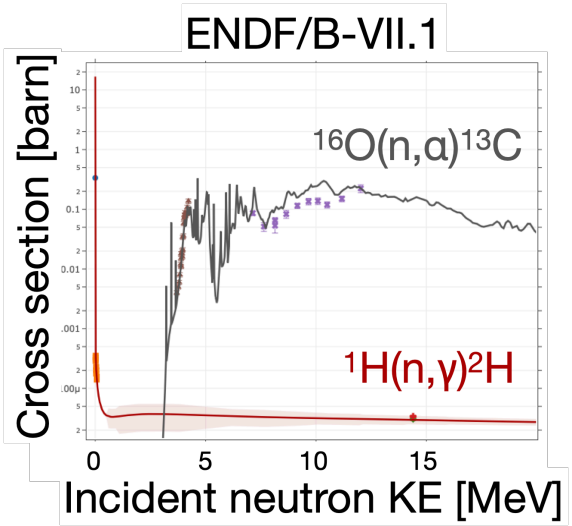


Figure 3.39: Cross sections of $^{16}\mathrm{O}(n,\alpha)^{13}\mathrm{C}$ and $^{1}\mathrm{H}(n,\gamma)^{2}\mathrm{H}$ reactions as a function of incident neutron kinetic energy, in the range of 0-20 MeV. The data is taken from ENDF/B-VII.1 [31] nuclear data.

We assume a ±0.5 MeV uncertainty in the standard kinetic energy distribution of Am/Be neutrons that accompany a 4.438 MeV gamma-ray. This is based on a comparison between the standard ISO distribution [135] and the recently measured spectrum by H. Ito et al. [137], as shown in Figure 3.40. This 0.5 MeV uncertainty in the neutron kinetic energy distribution propagates to the fraction of neutrons undergoing alpha emission and consequently to the neutron tagging efficiency estimate, as outlined in Table 3.10.

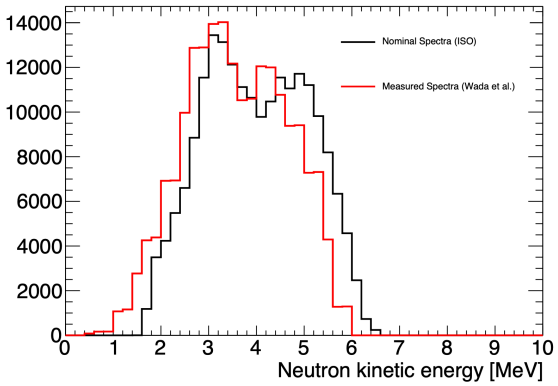


Figure 3.40: Comparison of Am/Be neutron kinetic energy distributions associated with the 4.438 MeV gamma-ray: Standard ISO distribution (black) and H. Ito et al. (red, [137]) (reprinted from [136])

MC neutron KE	SK-IV	SK-V	SK-VI	SK-VI (R)
Nominal (ISO-8529)	29.80 ± 0.16	26.09 ± 0.15	57.90 ± 0.17	39.55 ± 0.17
Shifted (-0.5 MeV)	30.09 ± 0.16	26.47 ± 0.15	58.66 ± 0.17	40.11 ± 0.17
Fractional change	1.0%	1.5%	1.3%	1.4%

Table 3.10: The true neutron tagging efficinecies in percentage for nominal MC and 'shifted' MC, where the neutron kinetic energy distribution is shifted by 0.5 MeV.

Detector modeling parameters

The detector modeling parameters including the relative gain and dark rates of individual PMTs, water transparency, and the top-bottom asymmetry (TBA) parameter reflecting the asymmetry in the mean PMT hit rate between the top and bottom of the tank, undergo gradual yet continuous changes over time. These parameters are automatically monitored on a run-by-run basis using randomly triggered data and the isotropic Xe light source fixed at the tank center. MC simulations were generated using these measured parameters, allowing us to observe how efficiency is affected by the varying parameters in comparison to the nominal MC. It is worth noting that not all runs were simulatable, as background PMT hits were only recorded during specific periods, typically correlated with the T2K accelerator neutrino schedule.

Figure 3.41 displays the true neutron tagging efficiency values obtained from run-by-run MC simulations as a function of the Run number. In Figure 3.42, histograms of the efficiency values are shown with varying detector modeling parameters measured in each run, categorized by each SK phase. The standard deviation of the efficiency distributions was 2.0% for SK-IV, 2.8% for SK-V, 1.0% for SK-VI, and 1.3% for SK-VI (R), relative to the average of each phase. These values were considered as the 1σ uncertainty resulting from the varying detector conditions.

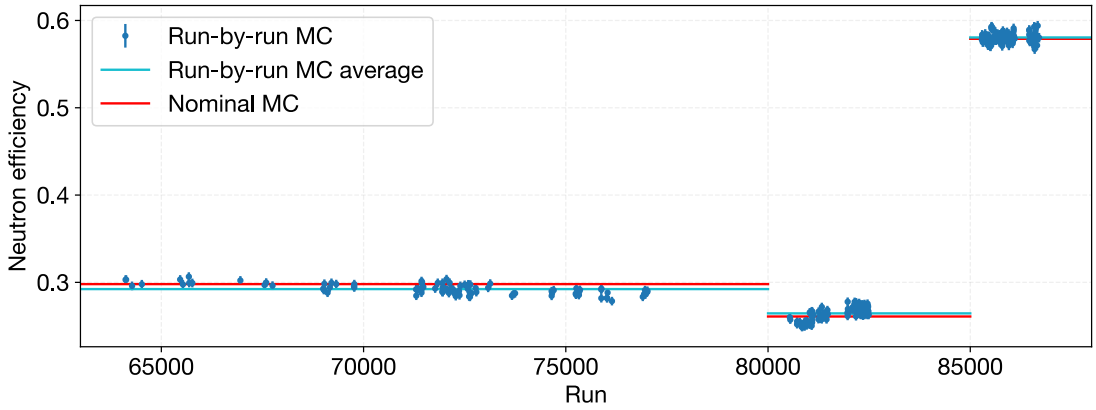


Figure 3.41: The true neutron tagging efficiencies obtained from run-by-run MC (blue, showing `INIT` vertex mode only), compared with the average for each SK phase (cyan) and the efficiency derived from the nominal MC (red).

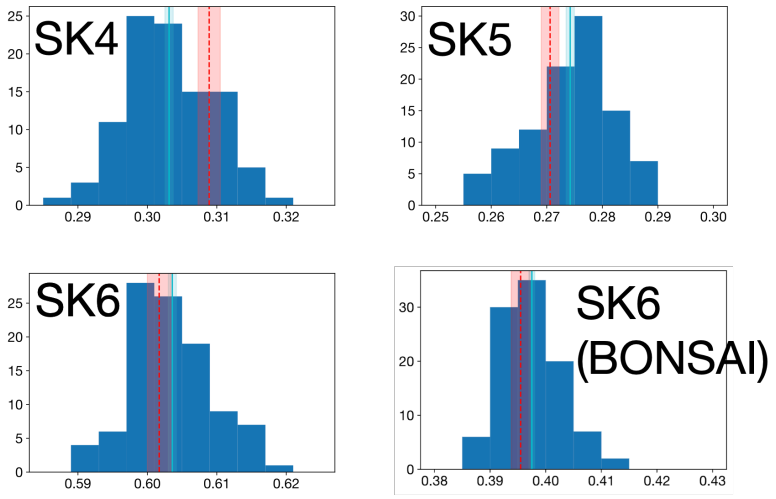


Figure 3.42: The distributions of true neutron tagging efficiencies derived from run-by-run MC simulations (blue histogram), categorized by each SK phase. The vertical axes represent the number of runs used in the simulation, with 100 runs in total for each phase. The cyan line represents the average of the blue histogram, while the red line represents the efficiency derived from the nominal MC.

Overall PMT response scaling factor (COREPMT)

The overall PMT response scaling factor is the final scaling factor to the PMT response that is introduced to match the simulated PMT charge with the observed. It is not measured run-by-run, but fitted to minimize the data-MC discrepancy in the charge per track length distributions of penetrating cosmic ray muons for each SK phase. The 1σ uncertainty of this factor was assumed to be 1% based on [34]. Table 3.11 summarizes the fractional changes due to 1σ shift in this factor, which are assumed as the propagated uncertainties to the neutron tagging efficiency estimate.

COREPMT	MC neutron tagging efficiency [%] SK-IV	SK-V	SK-VI	SK-VI (R)
+1σ	31.13 ± 0.16	27.73 ± 0.16	58.30 ± 0.17	39.94 ± 0.17
Nominal	30.89 ± 0.16	27.06 ± 0.16	57.90 ± 0.17	39.55 ± 0.17
-1σ	30.60 ± 0.16	26.73 ± 0.16	57.52 ± 0.17	39.17 ± 0.17
Fractional change	0.9%	1.8%	0.7%	1.0%

Table 3.11: Fractional changes in efficiency with shifted scaling factor.

Bias due to calibration setup

As discussed in Section 3.3.2, the calibration setup, including the pile-up backgrounds from the Am/Be neutron source and the BGO scintillator surrounding it, can impact the estimate of neutron tagging efficiency. Conducting the analysis without this setup is challenging, so we rely on simulations to estimate its effect. Additionally, generating multiple sets of simulations with different models — varied source neutron emission rates or different cross section datasets — using a complete Am/Be MC simulation with both `SKG4` and the pile-up simulator is an extremely time-consuming process. Therefore, we conservatively estimate the uncertainty by comparing only the existing nominal MC results with varying source positions.

We define the efficiency "bias factor" due to calibration setup as:

$$\text{Bias factor} = \frac{\text{Estimated signal efficiency in full Am/Be MC}}{\text{True signal efficiency in neutron particle-gun MC}} \tag{3.11}$$

For all source positions in each SK phase, the uncertainty of the bias factor was estimated as the quadratic sum of the mean of the MC statistical errors and the standard deviation of the bias factors for different source positions.

The two sets of simulations were all conducted in `SKG4` (Geant4.10.5.p01) with the same setup, including hardware and water modeling parameters, except for the calibration setup. The estimated bias factor (93.3%) for SK-VI was slightly larger than the value for the pure water phases (96.8% for SK-IV, 95.6% for SK-V), as gamma-rays from 8 MeV Gd(n, γ) reactions introduced an additional source of fake event triggers that mimic scintillation due to 4.438 MeV Am/Be gamma-rays. The `RECO` vertex mode for SK-VI showed even larger bias factor, possibly due to scintillator interference with the signals, affecting the performance of signal vertex reconstruction. These results are summarized in Figure 3.43. No significant vertex dependence was expected by the simulations, as shown in Figure 3.44. The estimated uncertainties of the bias factors for each SK phase were 6.9% for SK-IV, 4.6% for SK-V, 1.1% for SK-VI, and 1.2% in the SK-VI `RECO` vertex mode.

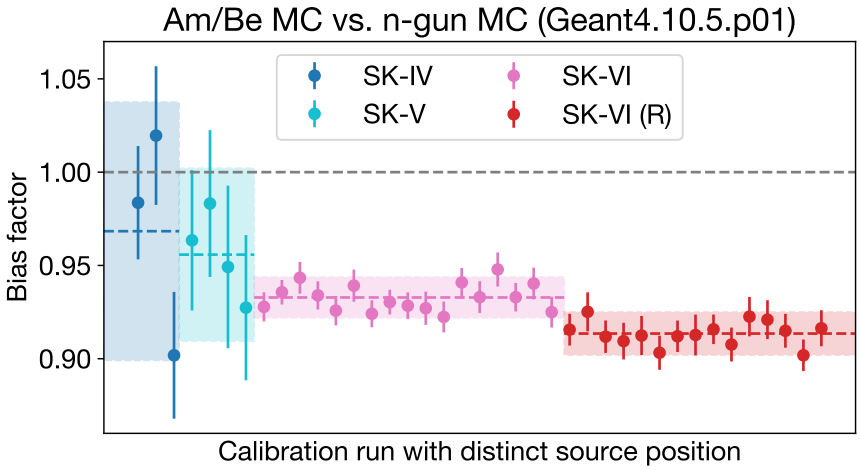


Figure 3.43: MC-predicted bias factors (points with MC-statistical errors) and the conservatively evaluated uncertainties (colored shade), resulting from the Am/Be calibration setup. The dashed lines represent the mean of the predicted bias factors for each SK phase.

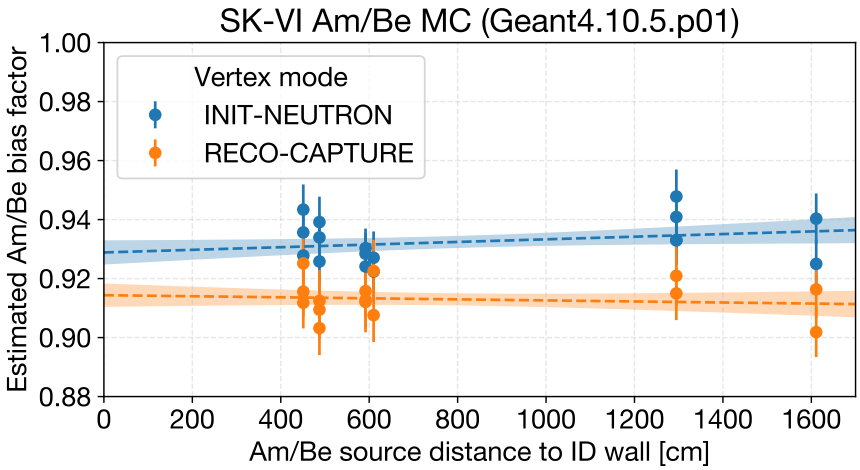


Figure 3.44: The MC-predicted bias factors plotted against the distance from the source to the ID tank wall, categorized by the two different vertex modes. The dashed lines correspond to the linear functions that best fit the simulations in each mode, while the shaded areas indicate the 1σ prediction intervals. The determined slopes aligned closely with zero, with a deviation of within 1.1σ.

Neutron capture ratios on H and Gd

The detection efficiencies for $^1\mathrm{H}(n,\gamma)$ and $\mathrm{Gd}(n,\gamma)$ signals show significant differences. Specifically, the efficiency for $^1\mathrm{H}(n,\gamma)$ signals is approximately 20%, while the $\mathrm{Gd}(n,\gamma)$ signal efficiency is around 95%. This makes the H/Gd-capture ratio a crucial factor in determining the overall signal detection efficiency. In the nominal MC simulation using `skdetsim` with 0.0110 w% input Gd concentration and neutron capture cross sections from ENDF/B-VII.1 [31], the thermal neutron capture ratios on Gd and H were found to be 52 : 48. To evaluate the uncertainty in this ratio, we compared the MC-simulated ratio with analytical predictions based on the input Gd concentration and cross sections. We also considered the ratios estimated based on the scales of $^1\mathrm{H}(n,\gamma)$ and $\mathrm{Gd}(n,\gamma)$ signal peaks in the signal energy distributions observed in Am/Be data and MC.

First, we estimated the neutron capture ratio on H analytically using the following equation:

$$r_{\mathrm{H}} \approx \frac{n_H g_w^H(T)\sigma_{H,2200\text{ m/s}}}{\sum_{i\neq H} n_i g_w^i(T)\sigma_{i,2200\text{ m/s}}} \tag{3.12}$$

where n_i is the number density of the i^{th} isotope, σ_i is the neutron capture cross section of the i^{th} isotope, v is the neutron speed, ρ is the neutron density, g_w is the Westcott g-factor (the ratio between the Maxwellian average cross section and the cross section evaluated at thermal neutron energy of 0.0253 eV), and T is the temperature. The $\mathrm{Gd}(n,\gamma)$ ratio r_{Gd} can be approximated as $1 - r_H$.

In the actual calculation, we only considered three isotopes: ^{1}H, ^{155}Gd, and ^{157}Gd, as the fraction of neutron captures on the other isotopes such as ^{16}O and ^{32}S was negligible (less than 0.1%). The number densities were estimated based on the actual loaded weight of $\mathrm{Gd_2(SO_4)_3 \cdot 8(H_2O)}$ powder for SK-VI [51]. The estimated number densities, along with the relevant cross section and the g-factors for these three isotopes, are summarized in Table 3.12.

	Number density [/cm^3]	Cross section [barns]	Westcott g-factor
^{1}H	6.68×10^{22}	0.332015	1.001
^{155}Gd	6.22×10^{16}	60740.1	0.8439
^{157}Gd	6.57×10^{16}	252928	0.8528

Table 3.12: Values used for estimating neutron capture ratios in SK-VI. Cross sections were obtained from ENDF/B-VII.1 [31] for a neutron kinetic energy of 0.0253 eV, and the Westcott g-factors ($T = 293.6$ K) were sourced from [146].

As for uncertainties, we assumed a 0.5% uncertainty for the number density of ^{1}H and 0.8% for those of the gadolinium isotopes. For cross section uncertainty, we adopted the standard values assigned in the ENDF/B-VII.1 [31] evaluation: 2.553% for $^1\mathrm{H}(n,\gamma)$, 3% for $^{155}\mathrm{Gd}(n,\gamma)$, and 4% for $^{157}\mathrm{Gd}(n,\gamma)$ reactions. Given recent measurements with smaller cross sections [147, 148] that conflict with the evaluated cross sections in ENDF/B-VII.1 [31], we may need to consider larger uncertainties, potentially up to around 10% for Gd isotopes. In addition, we accounted for a 1% uncertainty in the g-factor, which includes considerations of variations in nuclear data evaluation and the SK water temperature. The resulting estimated neutron capture ratio on ^{1}H with SK-VI Gd concentration was (56.2 ± 1.5)%.

The $^1\mathrm{H}(n,\gamma)$ ratio r_H can be experimentally determined using Am/Be data, by analyzing the distribution of the number of PMT hits (`NHits`), as illustrated in Figure 3.45. The chi-squared statistic $\chi^2(r_\mathrm{H}) = \sum_i^{n_i^\mathrm{Data}\geq5}(n_i^\mathrm{Data} - n_i^\mathrm{MC}(r_\mathrm{H}))^2/n_i^\mathrm{MC}(r_\mathrm{H})$ — n_i^Data represents the observed signal counts in the i^th data `NHits` bin, and $n_i^\mathrm{MC}(r_\mathrm{H})$ represents the expected signal counts in the i^th `NHits` bin scaled by r_H — was summed for all source positions. The χ^2-minimizing value of r_H was (56 ± 3)% (Figure 3.46). This result is consistent with the analytically predicted r_H of (56.2 ± 1.5)%. When fitting only the $\mathrm{H}(n,\gamma)$ peak (`NHits`$\in \{6,7,8,9\}$), which is less influenced by the shape uncertainty of the $\mathrm{Gd}(n,\gamma)$ `NHits` distribution, a similar result of (59 ± 5)% was obtained.

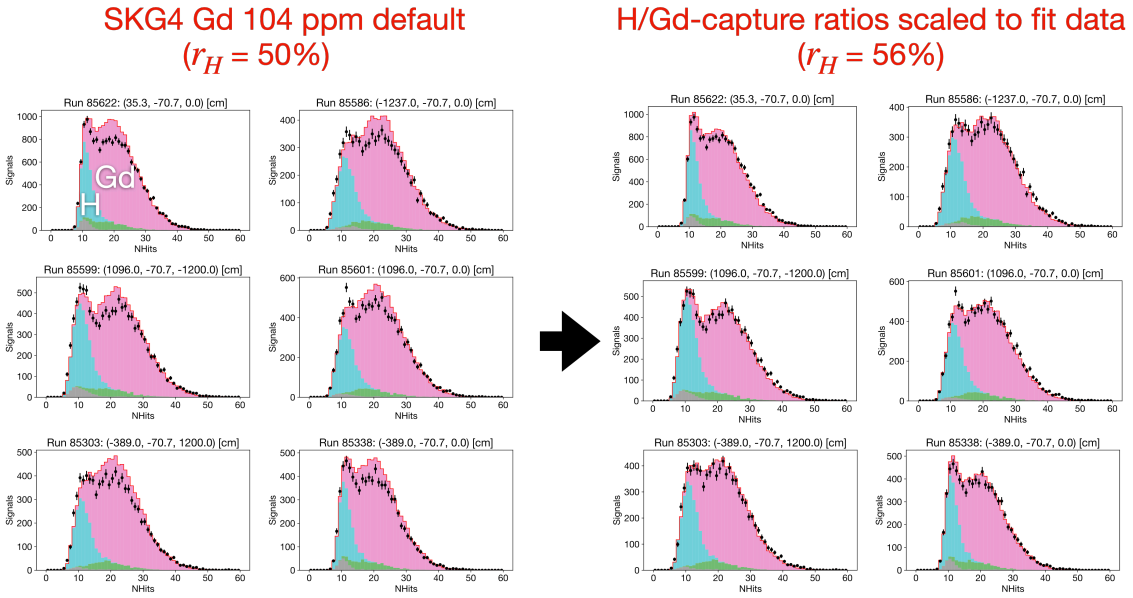


Figure 3.45: Comparison of `NHits` distributions for different source positions using Am/Be data (black) and MC simulations (red) with r_{H} set to 50% (left) and 56% (right). In the plots, colors correspond to various components: true Gd(n,γ) signals (pink), H(n,γ) signals (cyan), 4.438 MeV gamma-rays (green), and background PMT hits (gray). The histograms are normalized based on the number of events.

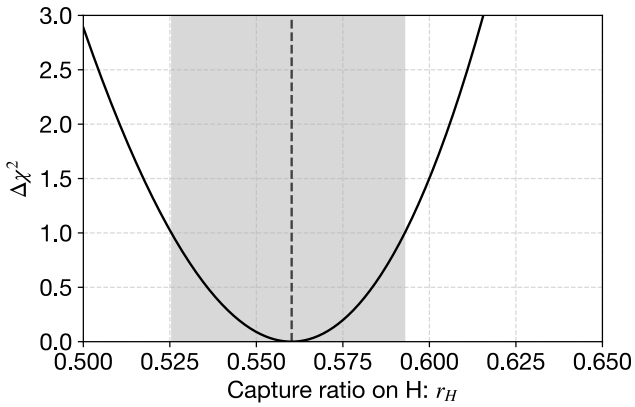


Figure 3.46: The distribution of $\Delta\chi^2$ as a function of r_{H}. The shaded gray area represents the 1σ uncertainty of the fitted r_{H}.

The MC-predicted r_H is 48%, showing a significant discrepancy with both the analytically predicted and measured r_{H} as shown in Figure 3.47. It seems like the MC simulation is giving an overestimated Gd-capture ratio, likely because the thermal motion of water molecules has not been accurately modeled[1]. Assuming an r_{H} value of 48% with uncertainties ranging from +8.2% (including both data-MC difference and the cross section uncertainties) to -1.5% (cross section uncertainties only), the resulting relative change in SK-VI neutron tagging efficiency is $^{+2.1}_{-9.6}$% (`INIT`) and $^{+3.1}_{-14.7}$% (`RECO`).

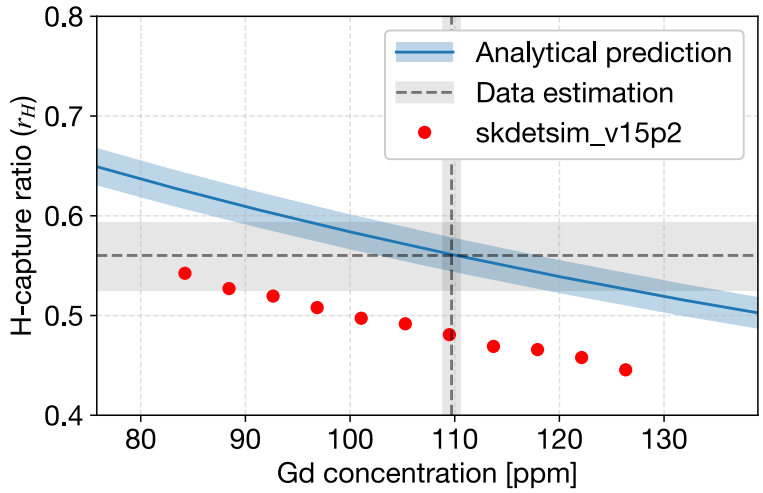


Figure 3.47: Comparison between the MC-predicted r_{H} (red) and the analytical predictions (blue) as a function of input Gd concentration. The values of Gd concentration and r_{H} estimated based on Am/Be data are indicated by dashed gray lines, with the shades representing the corresponding uncertainties.

[1]In Geant4, it seems that hydrogen is treated as a free proton without considering the molecular mass of water. This results in hydrogen moving 18 times faster in Geant4, decreasing the probability of neutron captures on hydrogen. Disabling the simulation of thermal motion in Geant4 changes the MC predictions to align with both the analytical prediction and observed results. As of the time of writing, this solution has not been officially adopted or integrated into the current analysis.

Gamma-ray emission model for Gd(n, γ) reaction

The ANNRI-Gd model which is used in the nominal MC demonstrates the closest agreement with our data, as illustrated in Figure 3.49, although it is not without its imperfections. To estimate the impact of the gamma-ray emission model to the neutron tagging efficiency, we compared the the nominal model with an alternative, GGARNET [113]. GGARNET uses a different photon strength function model used in simulating gamma-rays corresponding to de-excitation within quasi-continuum levels of $^{156/158}$Gd*.

Specifically, while ANNRI-Gd considers only the two most dominant electric dipole transition peaks in the de-excitation of $^{156/158}$Gd* isotopes using the standard Lorentzian (SLO) model — a sum of two Lorentzian peaks, the simplest model —GGARNET employs an extended model known as the enhanced generalized Lorentzian (EGLO) model [149]. GGARNET incorporates magnetic and higher order transitions, but it also introduces more empirical parameters that require dedicated tuning. The two models exhibit distinct gamma-ray energy spectra, while the ANNRI-Gd model aligns slightly better with the Ge detector measurements [113], as shown in Figure 3.48. Both models sample gamma-rays from discrete de-excitations in the same way, based on the observed multiplicities and discrete energy peaks.

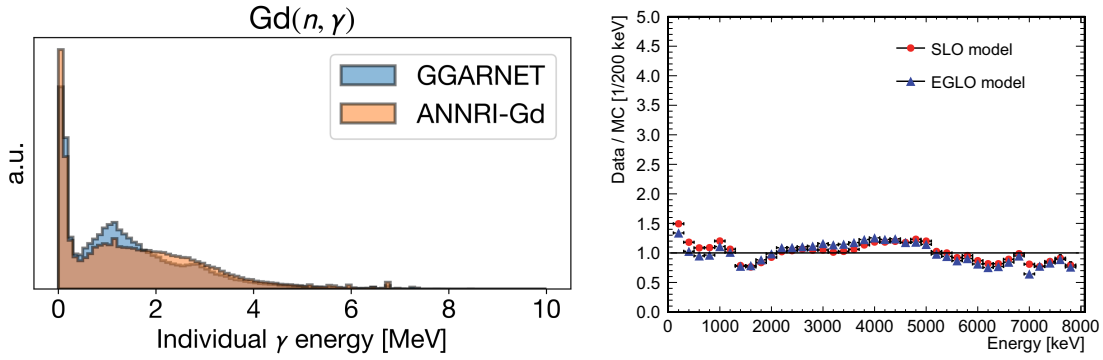


Figure 3.48: Individual gamma-ray energy spectra (left) of GGARNET (EGLO, blue) and ANNRI-Gd (SLO, orange) models, and the comparison with Ge detector array measurements of single gamma-ray events from ^{157}Gd(n, γ) reactions (right, reprinted from [113]).

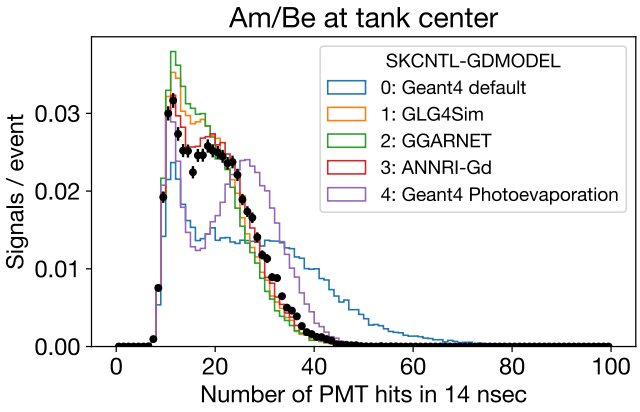


Figure 3.49: The `NHits` distribution of signal candidates taken from Am/Be data (black) with the source positioned at tank center compared with corresponding Am/Be MC with various Gd(n, γ) gamma-ray emission models. The neutron capture ratio on Gd was tuned to 44%.

As GGARNET tends to produce fewer energetic gamma-rays (>2 MeV, see Figure 3.48) that are more likely to scatter electrons over the Cherenkov threshold, the signal `NHits` distribution shifts to the left in comparison to the ANNRI-Gd model, as illustrated in Figure 3.49. This leads to a reduced true neutron tagging efficiency in SK-VI with GGARNET: (58.62 ± 0.17)% in `INIT` vertex mode and (36.75 ± 0.17)% in `RECO` vertex mode. In contrast, with ANNRI-Gd, the efficiencies were (60.17 ± 0.17)% (`INIT`) and (41.10 ± 0.17)% (`RECO`). We assigned fractional changes of 2.6% for the `INIT` vertex mode and 11.8% for the `RECO` vertex mode as uncertainties. The greater uncertainty associated with the `RECO` vertex mode signifies the fact that vertex reconstruction performance is heavily dependent on the gamma-ray emission model.

Neutron cross sections

The impact of cross section uncertainty appeared to be minimal. When comparing different cross section datasets, such as ENDF/B-VII.1 [31] with JEFF-3.3 [150], we observed a fractional difference on the order of 0.1% in the neutron tagging efficiency. As a result, it was not factored into the overall uncertainty estimate.

Neutron capture time constant

The neutron tagging efficiency estimate in Equation 3.10 relies on the width of the candidate selection time window ([18, 534] μs for SK-IV/V, [3, 534] μs for SK-VI, starting from the event trigger). The neutron capture time constant τ_{capture} in Equation 3.10, which depends on neutron capture cross sections, can influence the efficiency estimate. The observed τ_{capture} in Am/Be data was compared with an analytical prediction given by:

$$\tau_{\text{capture}} \approx \frac{1}{\sum_i n_i g_w^i(T)\sigma_{i,2200\ \text{m/s}}(2200\ \text{m/s})} \tag{3.13}$$

The parameters used were obtained from Table 3.12. The weighted mean of the observed time constants was 200.35 ± 3.74 μs for the pure water phases and 116.81 ± 0.27 μs for the Gd-loaded phase. The SK-VI result, obtained with the `RECO` vertex mode, was 116.85 ± 0.32 μs, consistent with the `INIT` vertex mode. These values aligned well with the predicted thermal neutron capture time constants 204.69 ± 5.33 μs for pure water and 114.87 ± 2.47 μs for 0.011 w% Gd-loaded water.

While the `skdetsim` simulation showed a consistent capture time for pure water phase, it suggested a slightly smaller capture time constant of 112.4 μs for SK-VI. However, the 4 μs discrepancy in neutron capture time constant led to a change in the efficiency estimate of less than 0.1%, and therefore was neglected.

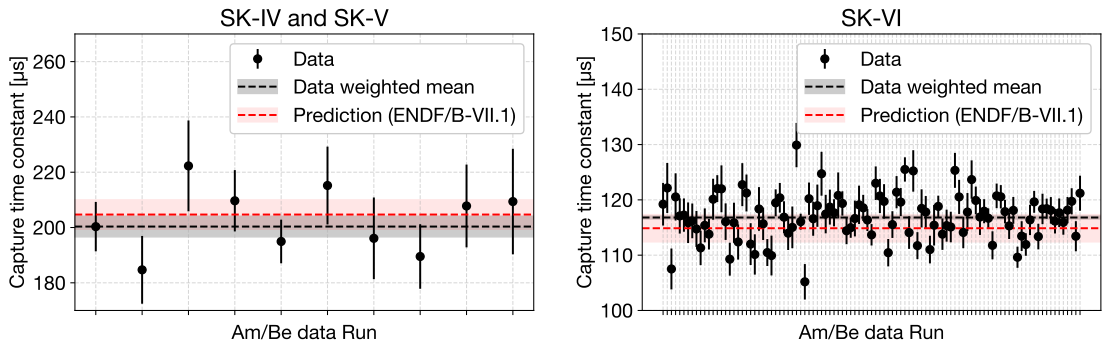


Figure 3.50: Neutron capture time constants obtained from the time distributions of signals processed with the `INIT` vertex mode, in SK-IV and SK-V (left), and SK-VI (right).

The major sources of uncertainty in estimating neutron tagging efficiency using the Am/Be neutron source discussed so far are outlined in Table 3.13. For the pure water phases, the most significant source of uncertainty arises from calibration setup bias, whereas in the Gd-loaded phase, the largest uncertainty stem from neutron capture ratios and the Gd(n, γ) gamma-ray emission model.

Source	SK-IV	SK-V	SK-VI	SK-VI (R)
Am/Be neutron kinetic energy	0.5%	0.9%	0.5%	0.6%
Detector modeling parameters	2.0%	2.8%	1.0%	1.3%
Overall PMT response scale	0.9%	1.8%	0.7%	1.0%
Bias due to calibration setup	6.9%	4.6%	1.1%	1.2%
Neutron capture ratios	-	-	$^{+2.1\%}_{-9.6\%}$	$^{+3.1\%}_{-14.7\%}$
Gd(n, γ) gamma-ray emission model	-	-	2.6%	11.8%
Total	7.3%	5.7%	$^{+3.8\%}_{-10.1\%}$	$^{+12.4\%}_{-18.9\%}$

Table 3.13: Evaluated major uncertainties to the neutron tagging efficiency calibration with the Am/Be neutron source.

3.3.7 Data-MC comparison

To validate the neutron capture signal detection efficiency on real data, we compared its estimates on Am/Be neutron source data with those obtained from MC simulations generated using `skdetsim` in a particle-gun setup. Since `skdetsim` operates in Geant3 and does not incorporate calibration setup (as implemented in Geant4), we did not apply Cuts #3 and #4 to `skdetsim` MC, as these cuts rely on calibration setup simulation. Instead, for each source position, we multiplied the estimated "bias factor", derived by comparing simple neutron particle-gun simulations with full Am/Be simulations using `SKG4` (as detailed in Section 3.3.6), with the true neutron tagging efficiency in `skdetsim`. This approach allowed us to account for the bias effect stemming from calibration setup and make a meaningful comparison between `skdetsim` and Am/Be data.

Figure 3.51 illustrates the estimated neutron tagging efficiency of `INIT` vertex mode for both Am/Be data and neutron simulation with the neutron source at the tank center. The data points are plotted against corresponding measurement dates. The observed efficiencies were consistent with the expected efficiencies within the total uncertainties for each SK phase, differing by approximately 10% at most. Notably, the evaluated efficiencies in data display variations over time due to fluctuations in detector modeling parameters.

We can also take an average of multiple measurements at the same position. Figure 3.52 (in the `INIT` vertex mode) and Figure 3.53 (in the `RECO` vertex mode) compare the weighted average of the efficiency estimates for each source position with the expected values from the simulation. SK-IV and SK-V demonstrate agreement between data and MC results, while for SK-VI in both vertex modes, a significant difference is observed across all source positions. This discrepancy is likely attributed to an approximately 10% larger Gd-capture ratio in our simulation. The average Data/MC ratios can be found in Table 3.14.

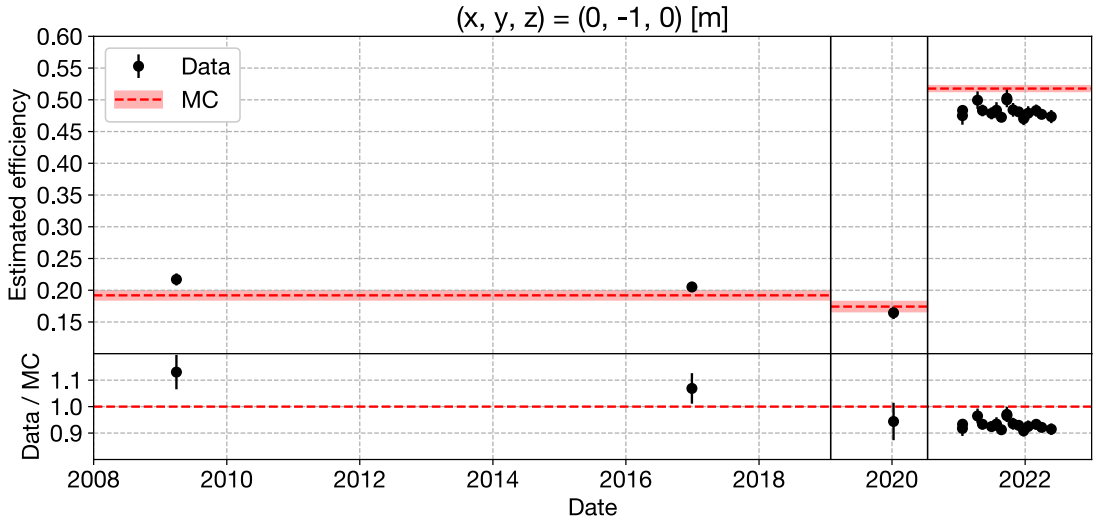


Figure 3.51: Neutron tagging efficiencies estimated for Am/Be neutron source data (in black) with the source positioned near the tank center, plotted against the measurement date. The expected efficiencies obtained from `skdetsim` neutron particle-gun MC simulations, adjusted by the corresponding calibration bias factors estimated with `SKG4`, are shown in red. The red shaded areas for the MC simulations incorporate major systematic uncertainties evaluated in Section 3.3.6.

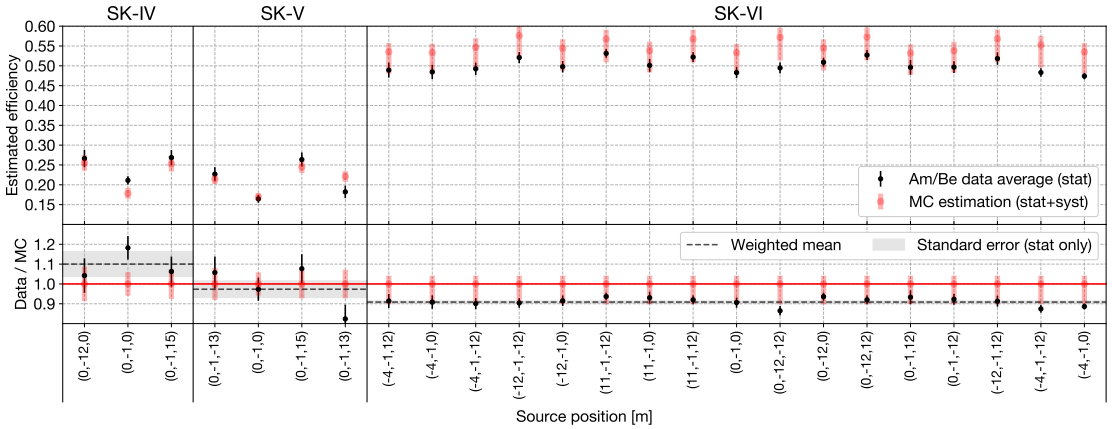


Figure 3.52: The estimated neutron tagging efficiencies of the `INIT` vertex mode for Am/Be data (black) and `skdetsim` neutron MC with `SKG4` calibration bias factors applied, plotted against the neutron source position.

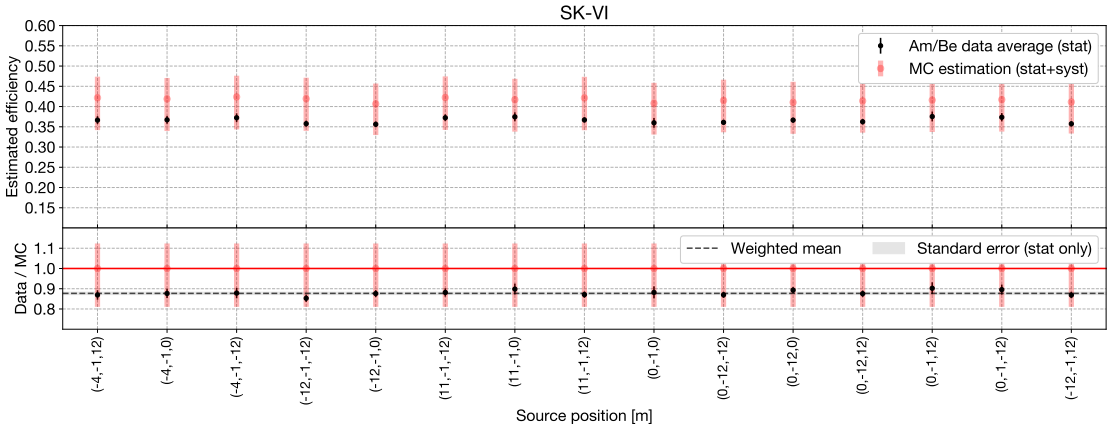


Figure 3.53: The estimated neutron tagging efficiencies of the `RECO` vertex mode for SK-VI Am/Be data (black) and `skdetsim` neutron MC with `SKG4` calibration bias factors applied, plotted against the neutron source position.

In our primary analysis, we utilize the obtained average data/MC efficiency ratios to account for the apparent disparity between calibration data and MC-estimated efficiencies. Especially for SK-VI, applying this correction allows us to disregard the substantial uncertainty in the Gd-capture ratio arising from simulation errors.

		Uncertainty		
SK phase	Data / MC efficiency ratio	Statistical	Systematic	Total
SK-IV	1.1000	0.0684	0.0803	0.1055
SK-V	0.9734	0.0443	0.0555	0.0710
SK-VI	0.9086	0.0069	0.0345	0.0352
SK-VI (R)	0.8770	0.0070	0.1087	0.1089

Table 3.14: Weighted mean and its statistical error for the data/MC efficiency ratios. These ratios and their uncertainties are utilized to adjust the signal efficiency, initially derived from the MC simulation, in our neutron multiplicity measurement (refer to Chapter 4). When incorporating systematic uncertainty into the analysis, we exclude the Gd-capture ratio uncertainty related to potential inaccuracies in modeling molecular thermal motion. Instead, we include uncertainties associated with neutron capture cross-section.

3.4 Michel electron separation

Michel electrons, the decay products of stopped muons, have an energy range of 0-50 MeV and occur within a few μs after muon production. In SK-VI, where the neutron capture time is faster, this time range overlaps with the neutron capture signal search window, which begins 3 μs after the event trigger. Michel electrons are more likely to be classified as signals because they tend to produce a much higher `NHits` count, typically in the hundreds, compared to typical background events.

To distinguish Michel electrons in SK-VI, a straightforward rectangular cut was employed based on the number of PMT hits and the time elapsed since the event trigger. Among the candidates with a neural network output greater than 0.7, those satisfying `NHits` $>$ 50 and a time interval of less than 20 μs from the event trigger were identified as Michel electrons. The remaining candidates were classified as neutron captures. Remarkably, this straightforward cut proved to be highly effective in cleanly distinguishing Michel electrons from neutron capture signals and vice versa, as illustrated in Figure 3.54.

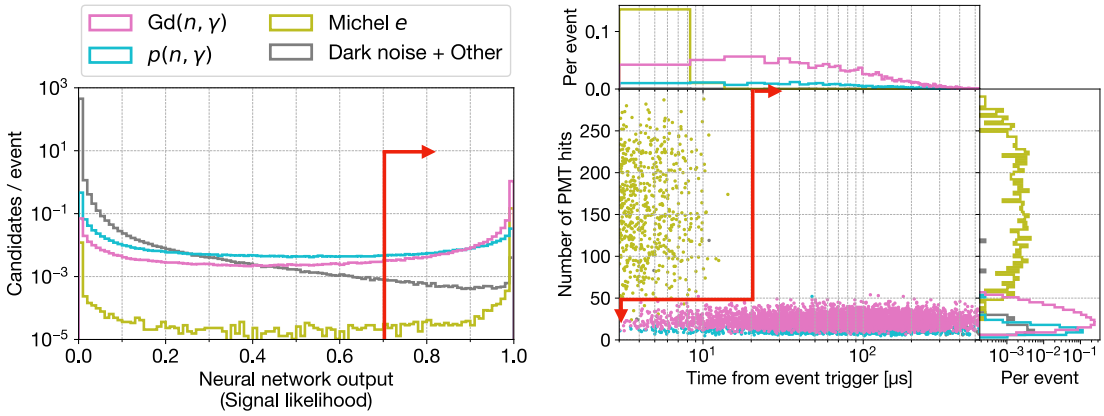


Figure 3.54: The neural network output distributions (left) for all candidates selected in MC-simulated atmospheric neutrino events, and the 2D scatter plot of the signal candidates in the plane of time and energy. The red arrows represent cuts for separating neutron captures from Michel electrons.

3.4.1 Validation with cosmic ray muon data

Cosmic ray muons continuously penetrate the overburden and reach the detector. Some of these muons come to a stop inside the ID and generate Michel electrons. This cosmic ray muon data serves as valuable calibration data for testing the Michel electron separation algorithm.

Stopping muon data was collected throughout SK-VI. The major reduction steps included requiring a large charge deposit in the ID to reject low-energy radioactivity, checking for OD PMT hits at the entrance and ensuring no other OD PMT hit clusters due to exiting particles to reject penetrating muons.

For each stopping muon event, the muon's entry point, momentum, and direction were reconstructed using an algorithm described in [28]. The muon range was determined by interpolating from tabulated data of muon Continuous Slowing Down Approximation (CSDA) ranges in water [151]. The stopping point of each muon was estimated using this information, and the estimated 1σ vertex resolution for this estimation was approximately 100 cm. The muon's estimated stopping vertex for each event was utilized for the initial TOF correction before the candidate selection stage. Candidates were then selected and classified according to the procedure outlined in Sections 3.2 and 3.4. MC simulations using `skdetsim` were generated based on the observed distribution of reconstructed muon momenta.

To validate the Michel electron separation cuts described in Section 3.4, events with an estimated muon stopping point farther than 3 m from the ID tank wall were selected. This ensured that Michel electrons were at least partially contained within the ID and prevented corner-clipping muons —penetrating muons that typically have a reconstructed stopping point close to the top corners of the tank wall — from contaminating the sample. Figure 3.55 displays the time distribution of candidates classified as Michel electrons, with the characteristic muon decay time accurately aligning with the expected value of 2.027 μs in water.

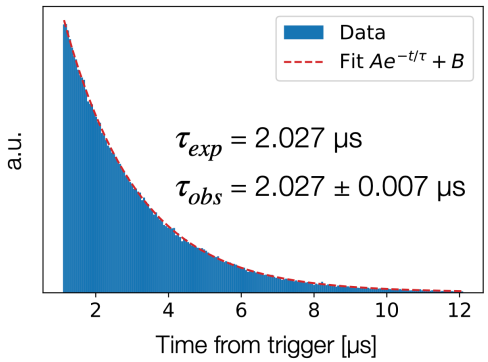


Figure 3.55: The observed muon decay time, fitted using a function of the form $Ae^{-t/\tau} + B$, where τ represents the muon decay time constant.

Figure 3.56 compares the features of Michel electron signals observed in stopping muon data with those from MC simulations. By examining the composition of MC-simulated signals within the data histograms, we can estimate the purity of Michel electrons in the selected signals. To find the purity in the selected candidates classified as Michel electrons, we minimized the following metric [152]:

$$-\ln L(a_e) \approx \sum_{\substack{\text{bins} \\ n_{\text{MC}}>0}} \left[n_{\text{MC}}(a_e) - n_{\text{Data}} + n_{\text{Data}} \ln \frac{n_{\text{Data}}}{n_{\text{MC}}(a_e)} \right] \tag{3.14}$$

Here, a_e is the Michel electron scaling parameter to be estimated, and $L(a_e)$ is the Poisson likelihood of the observation as a function of a_e. $n_{\text{MC}}(a_e)$ is the scaled MC bin value calculated as $a_e n^e_{\text{MC}} + (1 - a_e) n^{\text{BG}}_{\text{MC}}$, where n^e_{MC} is the MC true Michel electron count and $n^{\text{BG}}_{\text{MC}}$ is the MC true background count in each bin. n_{Data} is the observed count of candidates classified as Michel electrons in each bin.

The distribution of the metric as a function of a_e is shown in Figure 3.57. The minimizing a_e was multiplied by the MC true Michel electron purity to estimate the signal purity in the stopping muon data. The Michel electron detection efficiency was estimated by assuming the same Michel electron production ratio per event as in the MC simulation. The estimated Michel electron signal efficiency and purity for both stopping muon data and MC simulation are summarized in Table 3.15.

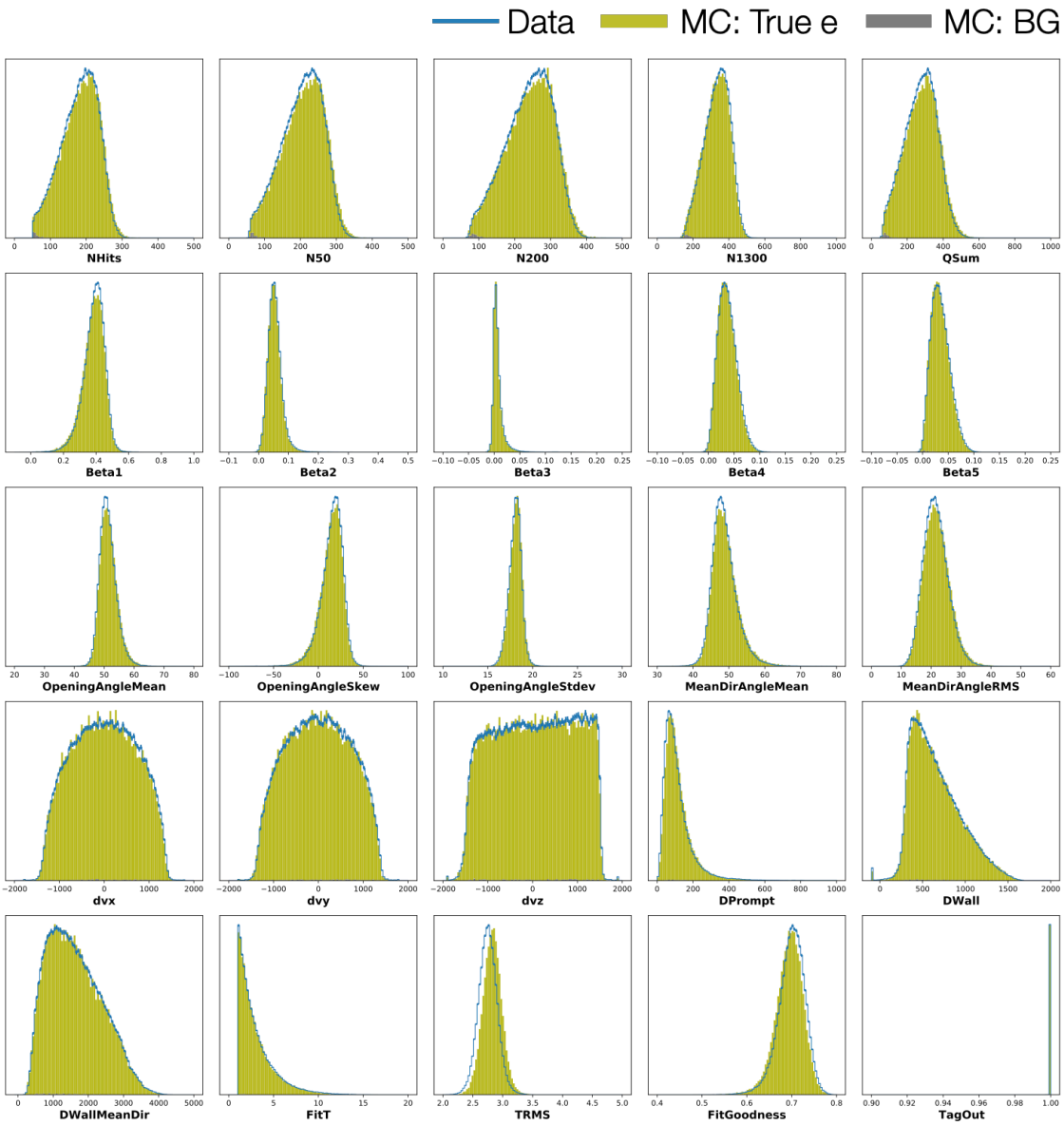


Figure 3.56: Feature distributions for candidates classified as Michel electrons, observed in the SK-VI stopping muon data (blue) and its corresponding MC simulation: MC true Michel electron signals (olive) and backgrounds (gray). Both the data and MC histograms are normalized by area.

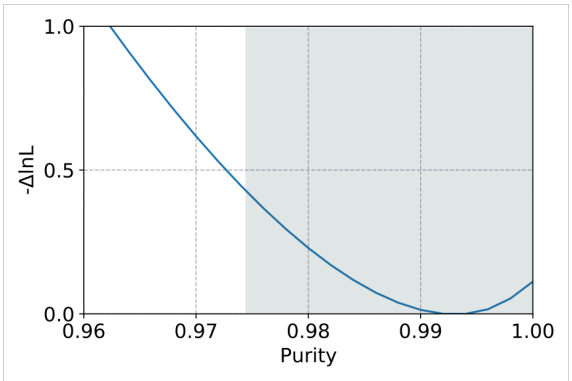


Figure 3.57: The change in the metric $L(a_e)$, defined in Equation 3.14, with respect to the Michel electron scaling parameter a_e. The shaded regions indicate the 1σ uncertainty in the fitted a_e.

	Data [%]	MC [%]
Efficiency	98.4 ± 1.3	98.8 ± 0.8
Purity	98.7 ± 0.5	99.3 ± 0.7

Table 3.15: Estimated Michel electron signal efficiency and purity in both the stopping muon data and its corresponding MC simulation, with statistical uncertainties.

The estimated signal efficiency and purity for Michel electrons in both the stopping muon data and its MC simulation were exceptionally high, exceeding 98%, and showed good agreement. This result demonstrates the effectiveness of the two-dimensional cut on signal energy and time in distinguishing Michel electrons from neutron captures. Consequently, we can confidently employ this technique to count neutron captures in neutrino events that involve both Michel electrons and neutron captures, with minimal concern for Michel electron contamination.

3.5 Effect of PMT afterpulsing on signal detection performance in SK-VI

Each hit PMT has a chance of experiencing an afterpulse, typically occurring approximately 10-20 μs after the initial PMT hit (main pulse), with a likelihood of around 0.1%. Due to the low probability, it is usually of minimal concern. However, in cases of highly energetic signals where multiple PMTs are triggered with significant charge deposits, this concern becomes more pronounced. The time frame of these afterpulses coincides with the widened signal search window in SK-VI, which starts 3 μs after the event trigger. We conducted a qualitative assessment of their impact on the signal detection efficiency and the false positive rate.

We compared the performance of signal detection on Multi-GeV (electron-equivalent visible energy larger than 1.33 GeV) atmospheric neutrino events, both fully contained and within the fiducial volume, with their corresponding MC simulation. We simulated afterpulses for each hit PMT, using a uniform afterpulse hit probability of 0.56%, along with a Gaussian-shaped timing probability with a mean of 14 μs after the initial hit and a standard deviation of 2.1 μs. Multi-GeV neutrino events typically has $O(10^3)$ PMT hits, or more than $O(10^4)$ photoelectrons in total. If afterpulsing does indeed impact signal detection performance, this event sample would be the most significantly affected.

Figure 3.58 displays the time distributions of all candidates along with those classified as neutron captures and Michel electrons, comparing the data to the MC. The afterpulse peak is clearly visible in the time distribution of all candidates, which are mostly due to accidental coincidences of background PMT hits including afterpulses. However, this peak disappears after the neural network's signal classification stage. This indicates that the neural network effectively filters out accidental candidates associated with PMT afterpulsing. Since the contamination from Michel electrons or PMT afterpulsing seems to be negligible, we can confidently start the signal search from 3 μs after the event trigger, as opposed to the 18 μs used in previous studies on neutron detection [34, 26, 28, 110, 153].

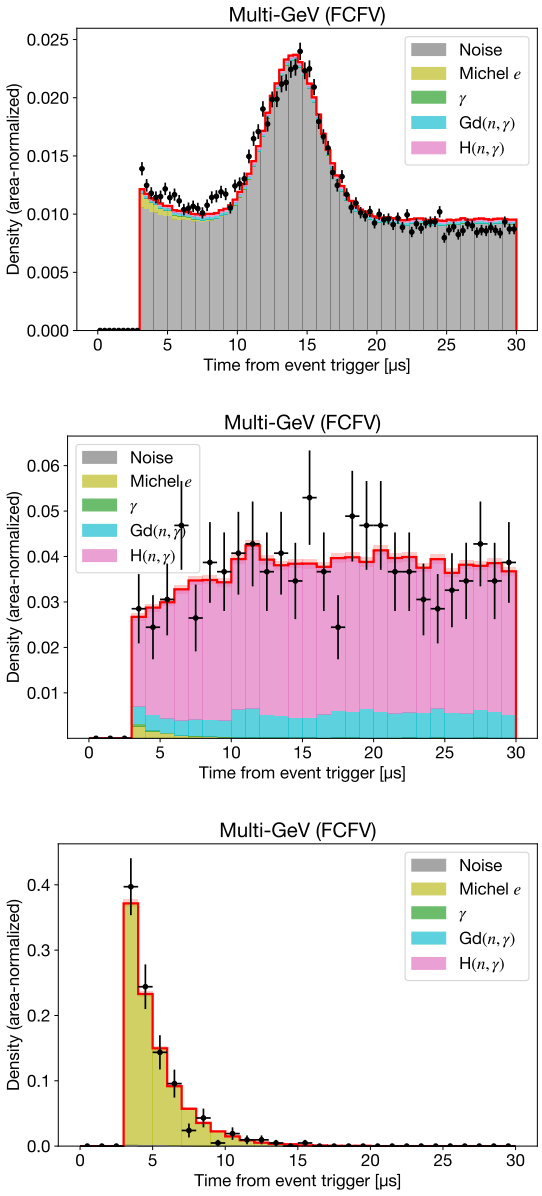


Figure 3.58: The time distributions of all selected candidates (top) and those classified as neutron captures (center) and Michel electrons (bottom), comparing the Multi-GeV atmospheric neutrino data (black) and the corresponding MC simulation (red). Both the data and the MC are normalized by area.

Chapter 4

Neutron multiplicity measurement

In this chapter, we delve into the specifics of data reduction, MC simulation, and neutron detection schemes tailored to this analysis. Next, we evaluate data quality by examining signals from both atmospheric neutrino interactions and delayed coincident neutron captures. We then review the method for estimating neutron capture detection efficiency, which is crucial for accurately evaluating the total produced neutron captures. Based on this method, we present measurements related to outgoing neutron multiplicity and kinematics in response to energy transferred during atmospheric neutrino interactions.

Specifically, we studied the average neutron capture multiplicity in atmospheric neutrino interactions by examining its dependence on the visible energy of the interaction. Neutron capture and visible energy serve as proxies for neutrons and transferred energy, respectively, as neither is directly observable in atmospheric neutrino interactions in SK. To assess the impact of uncertainties in hadronic re-interactions on our predictions, we compared the observed neutron capture multiplicity with different interaction models. This chapter presents a comprehensive overview of our findings, including an analysis of systematic uncertainties.

4.1 Data and simulation configurations

The data selected for analysis were obtained from Fully Contained (FC) events that passed all five FC data reduction stages, as described in Section 2.5.

4.1.1 Data run selection

For the pure water phases (SK-IV and SK-V), data from all "good" runs were used, excluding calibration or "bad" runs with unusual background rates in the reduction stages. In the Gd-loaded phase (SK-VI), data from the early phase, when Gd concentration may not have been uniform throughout the tank volume, were not included in the analysis, as such non-uniformity is not reliably simulatable.

The exact cut-off date was determined based on water conductivity and the measured neutron capture time constants with an Am/Be neutron source positioned at different locations within the tank. Water conductivity increases with Gd-loading due to an increase in the amount of free ions in the water. This serves as a reliable metric to assess ion concentration uniformity throughout the tank volume. Water conductivity at various vertical positions within the tank was measured throughout the Gd-loading process. Figure 4.1 displays the measured conductivity at different vertical positions and on different dates. Conductivity became maximal and uniform throughout the tank volume only after early September 2020, approximately 1.5 months after the start of Gd-loading (or the start of SK-VI on July 14, 2020).

Figure 4.2 illustrates the measured neutron capture time constants at three different vertical positions. A faster neutron capture time indicates a larger concentration of Gd. The position dependence is observed at the beginning of the Gd-loading. The capture time near the top of the tank becomes lower than the other two positions only after the mid-September measurement (September 15, 2020). Hence, to ensure Gd concentration uniformity in the SK tank, we used SK-VI atmospheric neutrino data taken after (including) September 15, 2020, excluding 2 months of data from the beginning of the Gd-loading in SK-VI.

The total detector livetime was 3244.4 days for SK-IV, 461.0 days for SK-V, and 564.4 days for SK-VI.

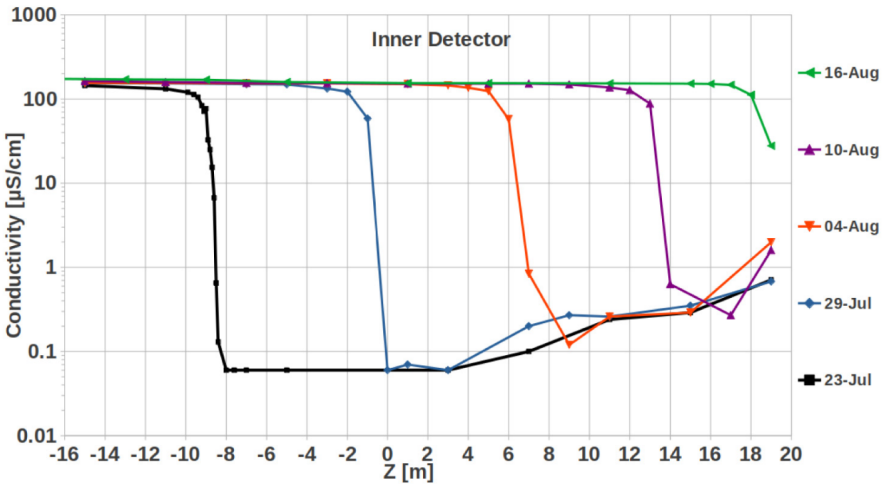


Figure 4.1: Water conductivity measurements taken during the Gd-loading process in SK-VI, plotted against vertical displacements of the measurement positions within the ID tank volume. (Reprinted from [51])

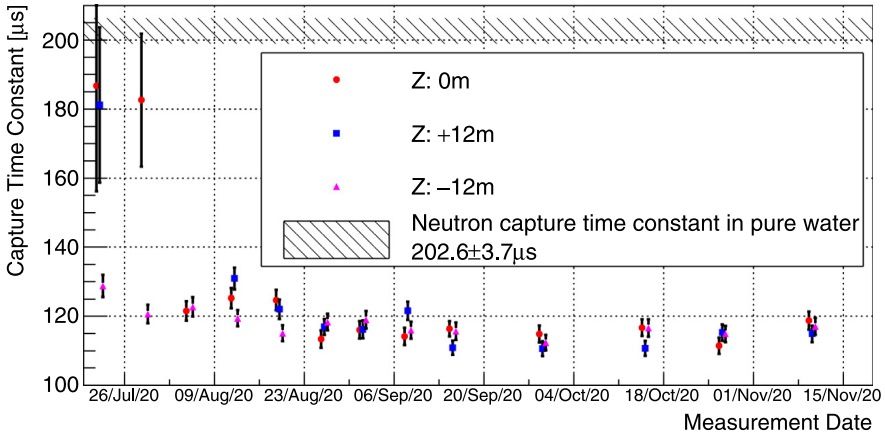


Figure 4.2: Neutron capture time constant measurements taken during and after the Gd-loading process in SK-VI, plotted against the measurement dates. (Reprinted from [51])

4.1.2 Simulation configuration

The MC simulation of atmospheric neutrino events in the ID volume, with a detector livetime of 500 years, was generated for each phase of SK, following the procedure outlined in Section 2.3. In terms of used software, `NEUT` 5.4.0 [154] was used for simulation of neutrino interaction in water based on the HKKM 2011 atmospheric neutrino flux model [55], without considering the oscillation. `skdetsim` v13p90 and v14 were used for detector simulation in SK-IV/VI and SK-V, respectively. The events were weighted based on the three-flavor oscillation probability following the details described in Section 2.3.7. The oscillation probability was calculated using the best-fit oscillation parameters obtained in the previous oscillation analysis [116], with constraints on θ_{13} with the value evaluated from reactor neutrino experiments [155]. The values used to calculate oscillation weights on an event-by-event basis in this analysis are outlined in Table 4.1.

Parameter	Value
$\sin^2\theta_{12}$	0.307
$\sin^2\theta_{13}$	0.021
$\sin^2\theta_{23}$	0.425
$\lvert\Delta m^2_{31,32}\rvert^2$	7.53×10^{-5} eV^2
$\lvert\Delta m^2_{23}\rvert^2$	2.53×10^{-3} eV^2
δ_{CP}	3.14

Table 4.1: Oscillation parameters [116] used in this analysis to calculate oscillation weights on an event-by-event basis. Normal mass ordering was assumed.

The parameters used to model the detector, such as the overall PMT response scaling (`COREPMT`) and the light attenuation factor in water, were fine-tuned for SK-IV and V primarily using cosmic-ray muon data, as explained in Section 2.2.3. As of the current writing, the official MC for SK-VI is not yet available. The simulation for SK-VI used the same detector modeling parameters as SK-IV. The effect of 0.011w% weight of Gd in water quality or atmospheric neutrino reconstruction capability is expected to be negligible, as discussed in Section 4.2.1 and Appendix A. Background PMT hits were properly sampled from randomly triggered events recorded in SK-VI.

4.1.3 Event reduction

We applied standard cuts to both the data and the MC simulation to identify fully-contained events with vertices in the fiducial volume (FCFV events). The selection stages are detailed in Appendix B. The efficiency of selecting true fully-contained neutrino events within the fiducial volume, with visible energy exceeding 30 MeV, during the data reduction process is estimated through simulation to be 96.2% (SK-VI MC) to 97.8% (SK-IV, [118]). The fraction of background events in the final data sample was maintained at the $O(0.1)\%$ level or lower, and this was monitored through visual inspection.

For this analysis, additional cuts were applied. The visible energy was restricted to below 10 GeV, as higher energy ranges are not expected to be well-suited for the "intranuclear cascade models" we would like to test. Additionally, an event with more than 15 OD PMT hits within a 200-ns time window in the signal search time range were excluded from the analysis, as such OD signatures may indicate the presence of cosmic-ray interference. This cut is similar to Cut #5 for Am/Be data reduction described in Section 3.3.4.

In addition, we excluded events that did not have a length of 535 μs. Specifically, events lacking AFT trigger flags (which ensure extended data recording) and events with a last candidate time earlier than 405 μs from the event trigger were excluded. The last cut was essential for SK-IV data, where roughly 2% of events with AFT trigger flags had a length less than 400 μs, possibly indicating an electronics issue. SK-V and SK-VI data did not have this problem.

The final data samples contained 29,942 events for SK-IV, 4,231 events for SK-V, and 5,203 events for SK-VI. These figures translate to livetime-normalized event rates of 9.23 $\pm$ 0.05 (SK-IV), 9.18 $\pm$ 0.14 (SK-V), and 9.22 $\pm$ 0.13 (SK-VI) atmospheric neutrino events per day, which are consistent across all phases. The dataset is summarized in Table 4.2.

	SK-IV	SK-V	SK-VI	
Gd	-		0.011w%	
Year	2008-2018	2019-2020	2020-2022	
Livetime [days]	3,244.4	461.0	564.4	
ν events	29,942	4,231	5,203	
ν event / day	9.23 ± 0.05	9.18 ± 0.14	9.22 ± 0.13	
Vertex mode	`INIT`			`RECO` (R)
Search range	[18, 534] μs from trigger		[3, 534] μs from trigger	
Detected n	15,705	2,035	5,752	4,359
n / ν event	0.525 ± 0.004	0.481 ± 0.011	1.106 ± 0.015	0.843 ± 0.013
Analysis	Multiplicity (used in combined data analysis)			Track length Multiplicity*

* We utilized SK-VI (R) in `RECO` vertex mode for comparisons with other datasets. However, it was not included in the final multiplicity data, which combines results from SK-IV through SK-VI. This exclusion is to avoid double-counting with `INIT` mode results.

Table 4.2: Summary of dataset statistics.

The same neutron capture signal detection scheme was applied to both data and simulated events. The `INIT` vertex mode was used across all SK-IV, V, and VI samples, while the `RECO` vertex mode was additionally applied to the SK-VI samples only. The signal search time range for each neutrino event was set to [18, 535] μs from the event trigger for the pure water phase (SK-IV and SK-V) samples and [3, 535] μs for the Gd-loaded phase (SK-VI) samples. For the neutron multiplicity analysis, we exclusively combined and compared results from the `INIT` vertex mode with the simulation. Results with the `RECO` mode were used for algorithmic comparison of multiplicity results and for measuring neutron track length.

4.2 Data quality

The final data sample was compared to the simulation to verify that there are no significant unexplained discrepancies. Figure 4.3 illustrates the simulated PMT hit time distribution for a typical neutrino event with a delayed coincident neutron capture signal. Both the prompt signals from atmospheric neutrino interactions (mainly involving e, μ, and π) and the delayed neutron capture signals were examined.

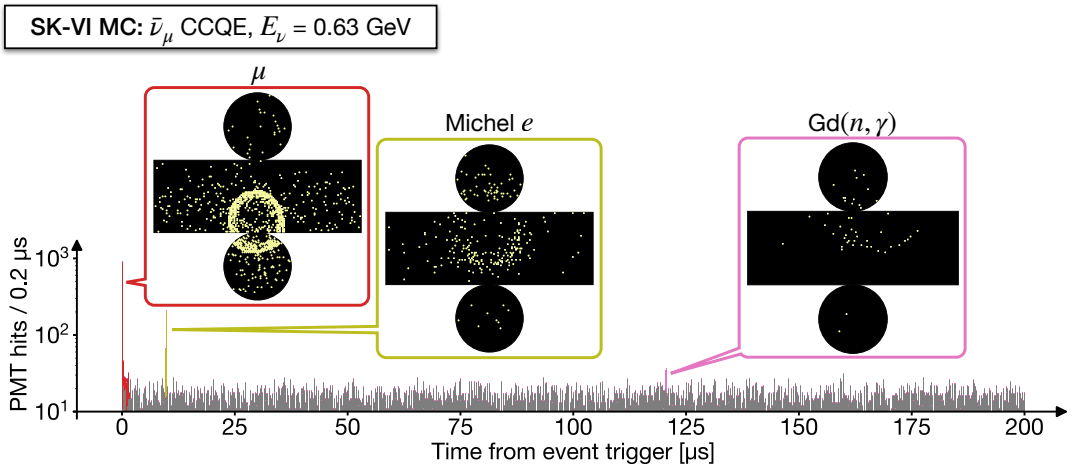


Figure 4.3: A PMT hit time distribution for a typical $\bar{\nu}_\mu$ charged-current quasi-elastic (CCQE) event, starting from the event trigger. The event displays feature the "prompt" muon signal (red) and the two types of "delayed" coincident signals — Michel electrons from muon decay (olive) and neutron captures on Gd (pink). This event was simulated using `skdetsim` in the SK-VI configuration.

4.2.1 Prompt atmospheric neutrino interaction signals

We focus on key factors, such as interaction types (CCQE, non-QE, NC), and transferred energy, which affect outgoing neutron multiplicity and kinematics. The number of reconstructed Cherenkov rings proves to be the optimal variable for distinguishing CCQE events (single-ring) from the others (multi-ring) (see Figure 4.9). The event rates and reconstructed energies in each sample are compared with the simulation. Additionally, we verify the validity of the simulation's oscillation weights by comparing the oscillated distributions of reconstructed lepton zenith angles with the corresponding observed data.

We examined reconstructed event variables affecting neutron capture detection efficiency. The precision of the initial photon TOF, crucial for signal detection, heavily depends on the base vertex quality. We assessed the reconstructed neutrino interaction vertex and other observables influencing vertex reconstruction performance, including particle type, vertex, momentum, and multiplicity.

Lastly, we show that the presence of Gd has minimal effect on event reconstruction. This is illustrated by comparing data from the pure water phase (SK-IV) with Gd-loaded SK-VI data, alongside SK-IV MC simulation. Specifically, we present distributions of ring counting likelihood and particle identification (PID) likelihood which affect event categorization. Further checks on additional distributions are outlined in Appendix C.

Unless specified otherwise, "data" refers to the combined final data sample (SK-IV + SK-V + SK-VI), and "MC" refers to combined MC events weighted by the livetime of each SK phase. Errors shown in this section are purely statistical.

Atmospheric neutrino event rate

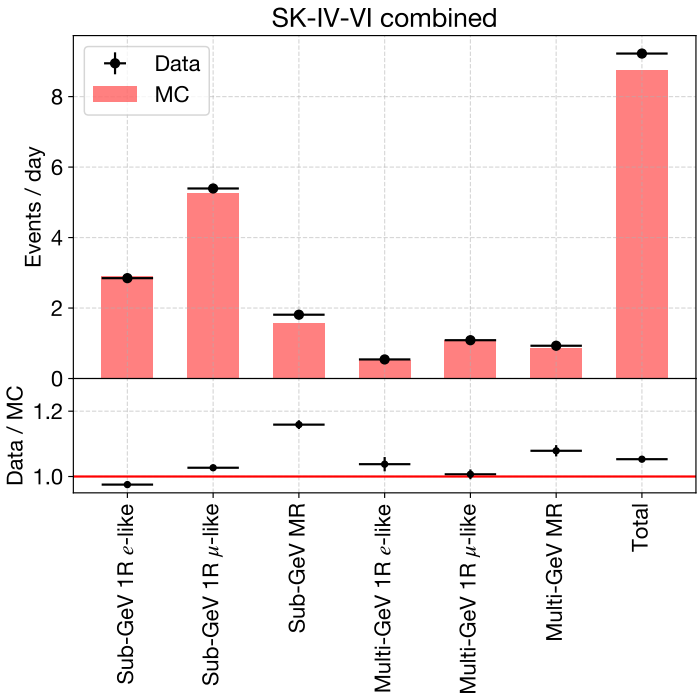


Figure 4.4: Comparison of the daily event rates in the final sample with the expected values. Events with visible energy below/above 1,330 MeV are categorized as sub-GeV/multi-GeV. Sub-GeV and multi-GeV samples were classified into single-ring (1R) and multi-ring (MR), with 1R further divided into e-like and μ-like.

Figure 4.4 presents a comparison between event rates in the data and simulation. The overall observed event rate is larger than the expected rate by approximately 5%, with the majority of this difference originating from the sub-GeV samples. Discrepancies in sub-GeV event rates fall within the uncertainty range of around 25% in the absolute sub-GeV atmospheric neutrino flux. Overall, the event rates align well with simulation expectations, though there appears to be a slightly higher multi-ring fraction in the sub-GeV sample than predicted. From this point, all displayed histograms are normalized based on the number of events.

Reconstructed neutrino interaction vertex

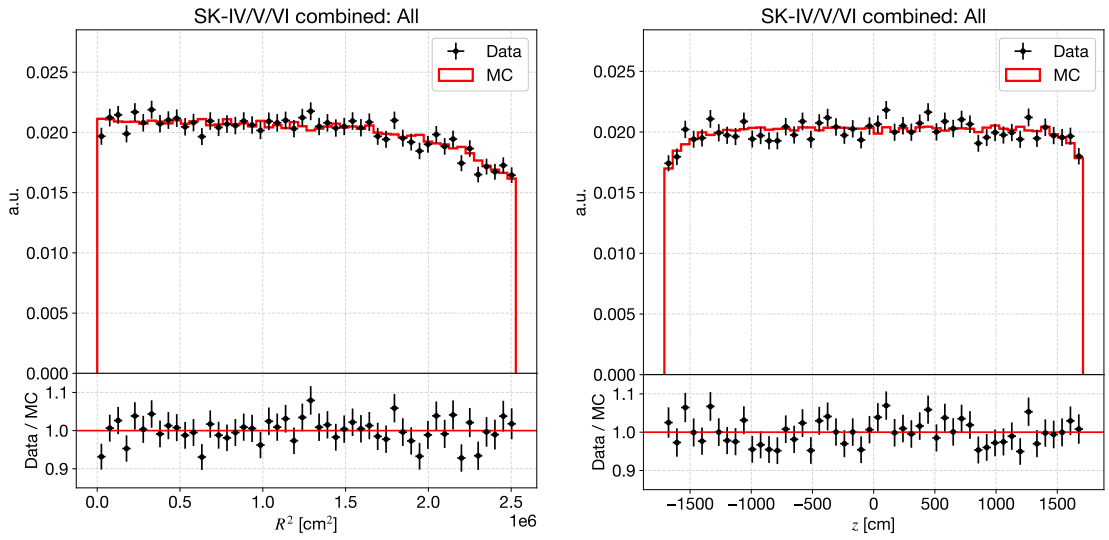


Figure 4.5: Distributions of the reconstructed atmospheric neutrino interaction vertex radial position R^2 (left) and vertical position z (right).

The distribution of reconstructed atmospheric neutrino interaction vertices appears mostly uniform across the entire detector fiducial volume. However, the event rate near the tank wall is slightly smaller, possibly due to the data reduction algorithm which tends to avoid vertices close to the tank wall in order to mitigate background radioactivity and cosmic rays. This characteristic is also manifest in the atmospheric neutrino simulation. The distributions weighted by detected neutron counts are shown in Figure 4.6, showing reasonable agreement with expectation.

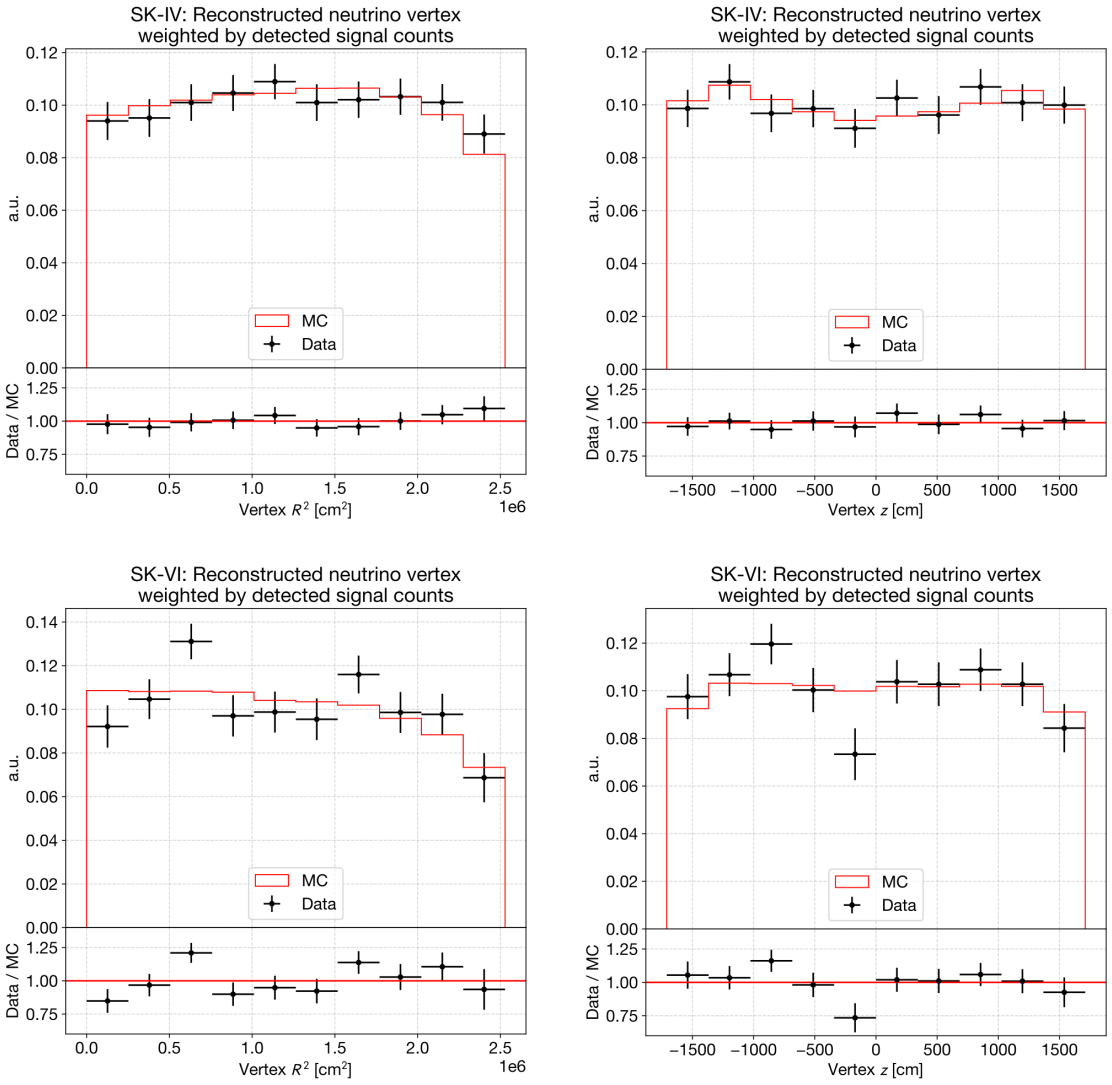


Figure 4.6: Distributions of the reconstructed atmospheric neutrino interaction vertex radial position R^2 (left figures) and vertical position z (right figures) for SK-IV (top figures) and SK-VI (bottom figures) weighted by the number of detected neutron counts.

Reconstructed Cherenkov ring multiplicity

The Cherenkov ring reconstruction algorithm is outlined in Section 2.4.1. Figure 4.7 shows the ring counting likelihood calculated for the first (most energetic) ring reconstruction, along with the reconstructed ring multiplicity distributions of both data and simulation. The ring counting likelihood, which determines whether an event is classified as single-ring or multi-ring, shows good agreement between data and simulation. Additionally, Gd-loaded SK-VI data shows no significant deviation from SK-IV data and simulation. The overall data indicates a few % larger multi-ring fraction, consistent with the observed sub-GeV multi-ring event rates. It is notable that single-ring events are expected to be predominantly CCQE events, while multi-ring events are mostly CC non-QE events or NC events.

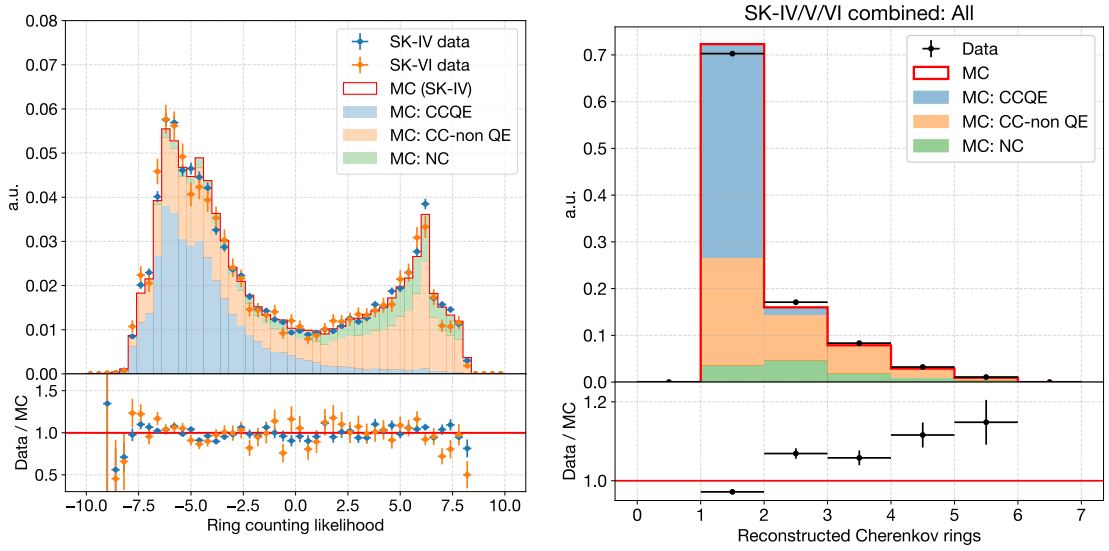


Figure 4.7: The distributions of the ring counting likelihood for the most energetic ring candidate (left) and the final reconstructed Cherenkov ring multiplicity (right).

Reconstructed visible energy

The total visible energy of an event is determined by summing the charge deposits of each reconstructed Cherenkov ring, converted into electron-equivalent energy. Figure 4.8 present the predicted and observed visible energy distributions. Shape disagreement below 200 MeV is associated with a bug in nuclear binding energy calculation, which is fixed in the newer versions of `NEUT`.

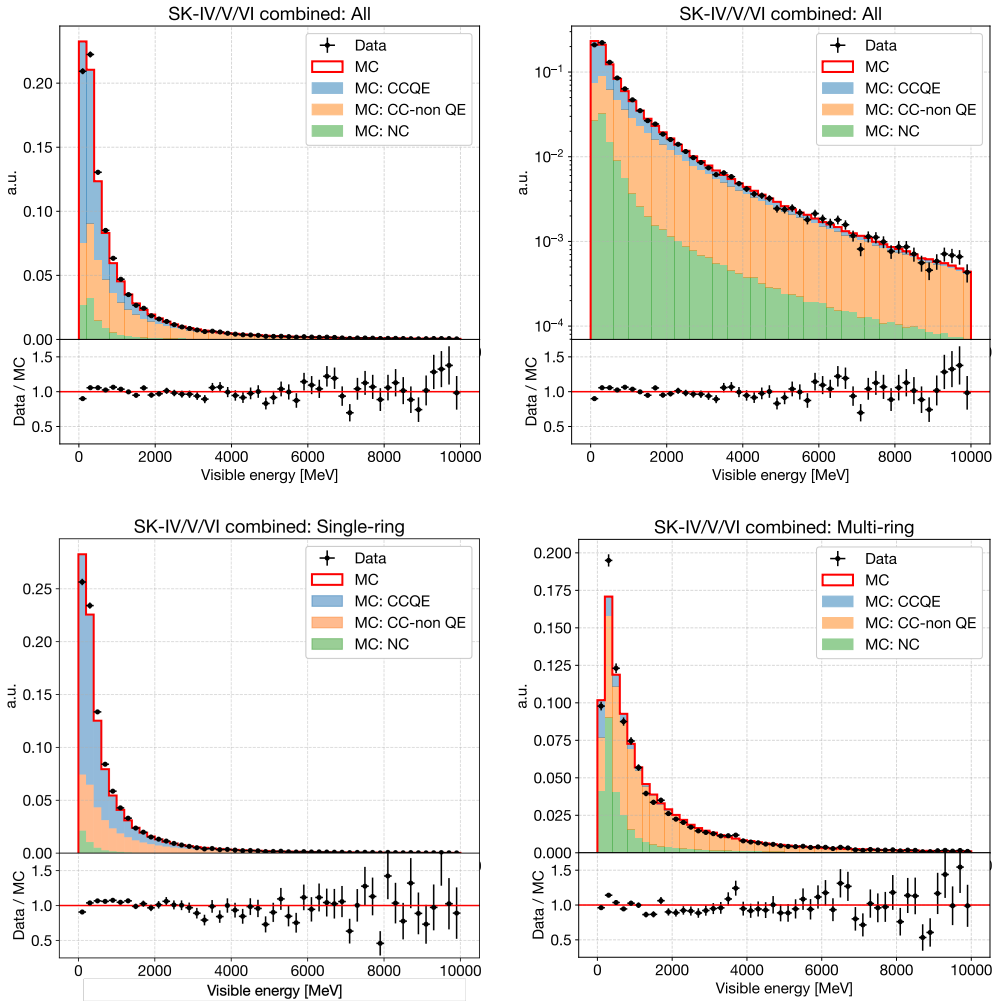


Figure 4.8: The reconstructed visible energy distribution of the entire final sample, with the vertical axis in linear scale (left) and in logarithmic scale (right). The reconstructed visible energy distributions of the single-ring (left) and multi-ring (right) samples.

Cherenkov ring particle identification

Figure 4.9 displays the distributions of particle identification (PID) likelihood as detailed in Section 2.4.1 for both the final data sample (SK-IV and SK-VI) and the simulation. The SK-VI distribution aligns well with SK-IV, indicating minimal impact of Gd on the ring PID.

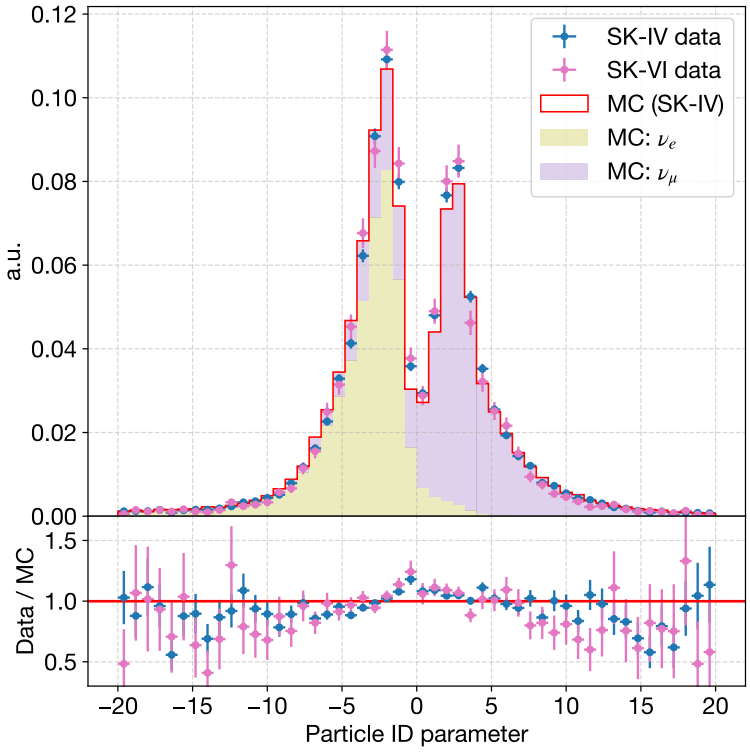


Figure 4.9: The PID likelihood distribution for all events in the final data sample. A Cherenkov ring is identified as "e-like" if the likelihood is less than 0, and as "μ-like" otherwise.

Reconstructed lepton momentum

The reconstructed lepton momentum distribution for single-ring events are shown in Figure 4.10. The overall shape agreement between data and simulation is good for both single-ring e-like and μ-like samples.

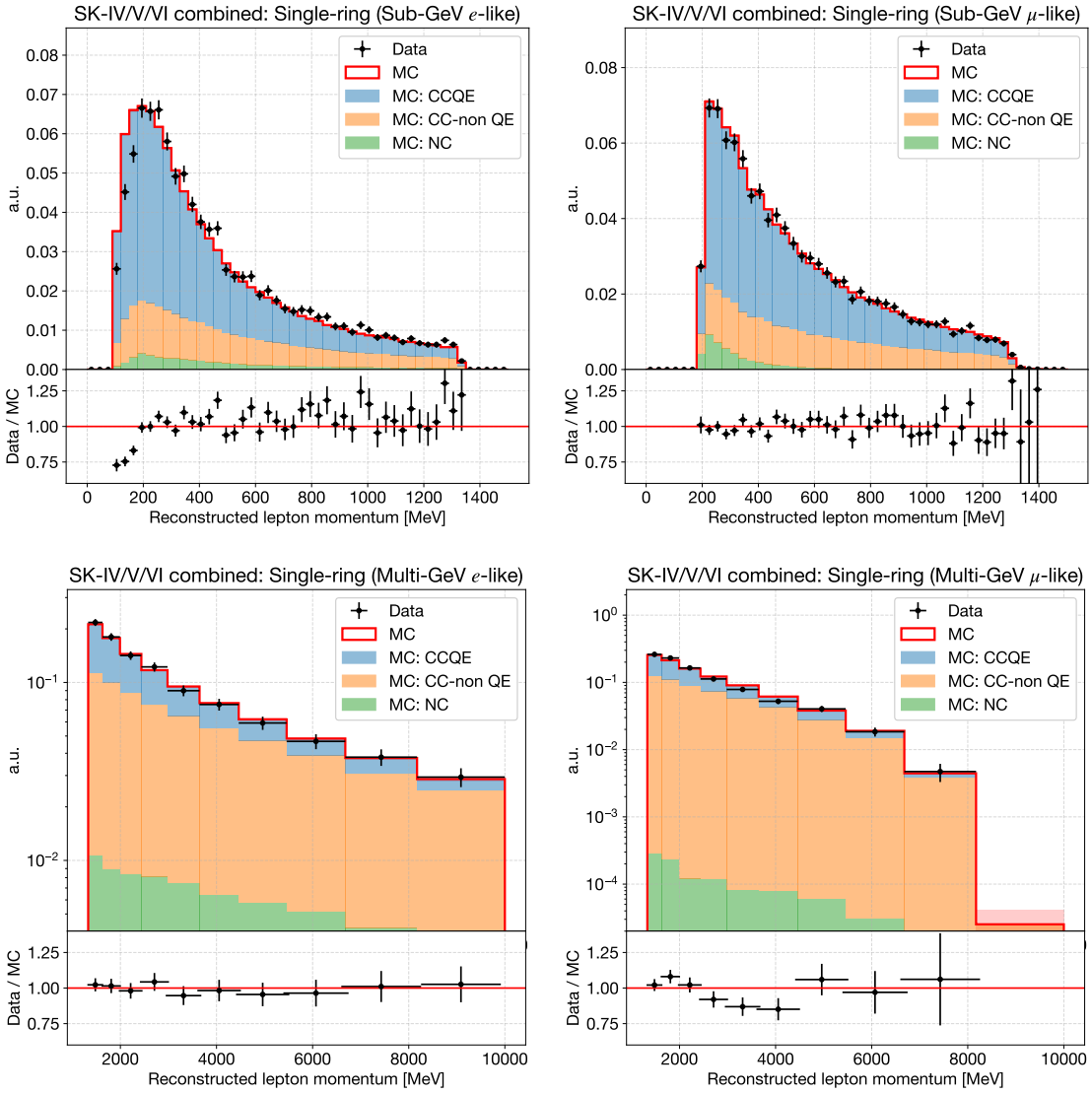


Figure 4.10: The reconstructed lepton momentum distributions for single-ring Sub-GeV e-like (top left), Sub-GeV μ-like (top right), Multi-GeV e-like (bottom left), and Multi-GeV μ-like (bottom right) samples. The disagreement in e-like events below 200 MeV/c is associated with a bug in nuclear binding energy calculation.

Reconstructed lepton zenith angle

Figure 4.11 displays the reconstructed distributions of lepton zenith angles.

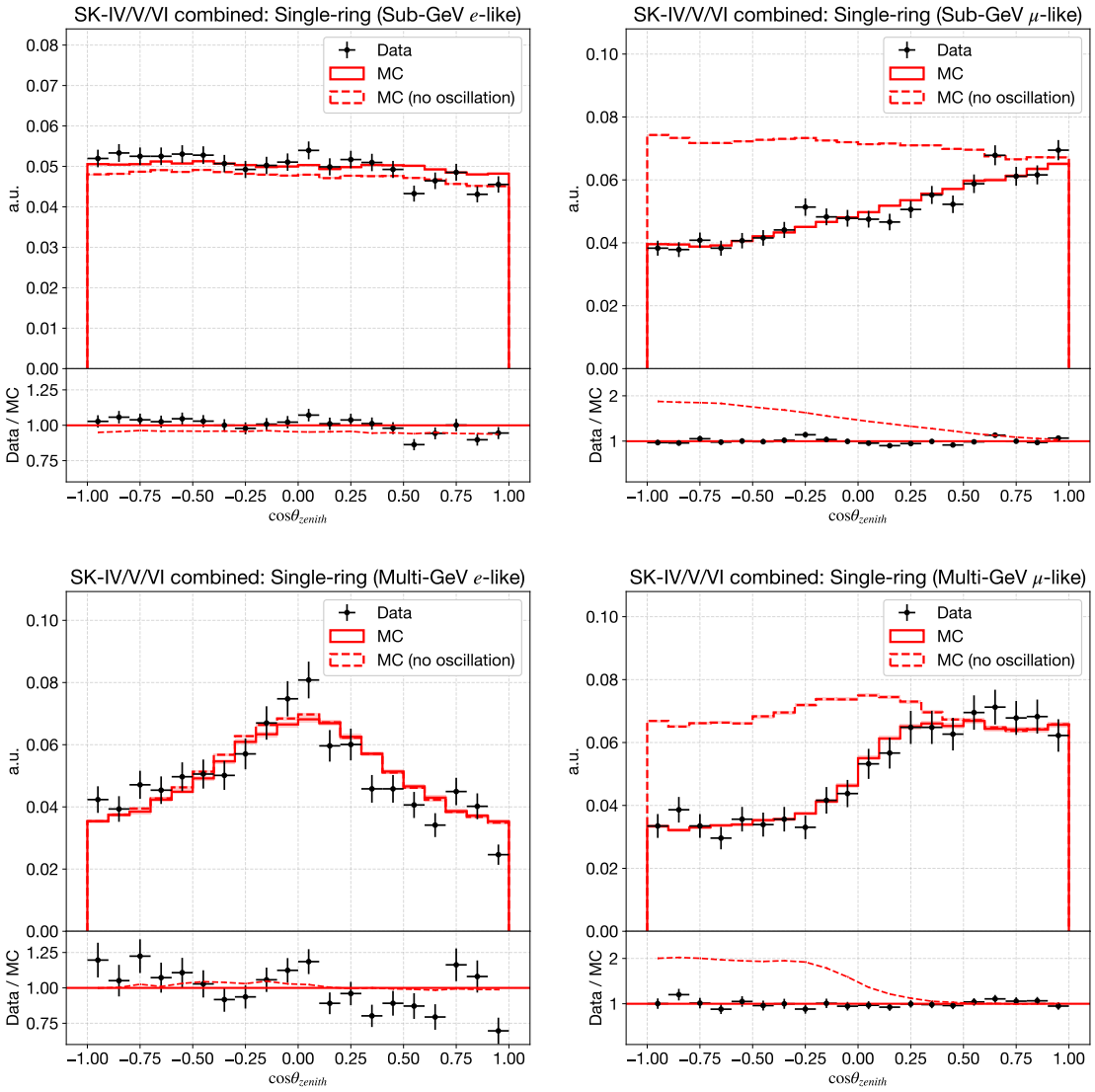


Figure 4.11: Cosine zenith angle distributions for single-ring, Sub-GeV e-like events (top left), μ-like events (top right), Multi-GeV e-like events (bottom left), and μ-like events (bottom right). $\cos\theta_{\text{zenith}}$ close to -1 indicates up-going neutrinos reaching the detector from below, indicating a longer path length. Both the data (black) and the MC simulation (red) are normalized based on the number of events. The red dashed line represents the simulation without applying oscillation weights, illustrating changes in the fraction of event rates due to neutrino oscillation.

The direction of the outgoing lepton serves as a proxy for the incoming neutrino direction, particularly for energetic neutrinos with $E_\nu > 500$ MeV, as leptons tend to be forward-scattered. The zenith angle not only determines the neutrino path length but also the oscillation probability together with E_ν. As illustrated in the figure, a noticeable decline in the expected event rates is evident for muon neutrinos, primarily due to the dominance of $\nu_\mu \to \nu_\tau$ transitions in atmospheric neutrinos, a consequence of the substantial θ_{23} angle. The oscillation parameters shown in Table 4.1 explains the observed shape of the muon zenith angle distributions very well.

4.2.2 Delayed neutron capture signals

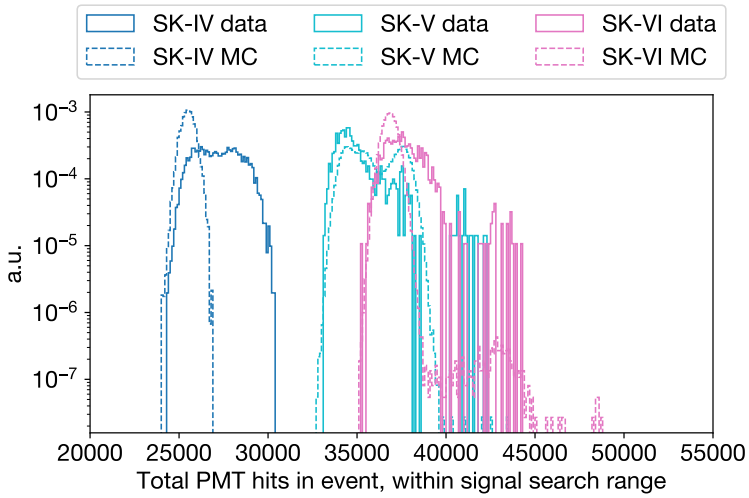


Figure 4.12: Distributions of the total number of PMT hits in the signal search range for each event, across different SK phases.

Before delving into the properties of the neutron capture signal, it is important to note the challenges in modeling background PMT hits over the entire data acquisition period. Figure 4.12 shows the distributions of the total number of PMT hits within the signal search time range for both data and simulation in SK-IV, V, and VI. It is evident that the modeling of background hits, which constitute most of the PMT hits within the given time range, does not fully describe data.

Table 4.3 provides a summary of the candidate rates selected in detection algorithm Stage 1 (PMT hit trigger). Following this initial stage, the candidates undergo classification into BG-like and signal-like categories in detection algorithm Stage 2, facilitated by a pre-trained neural network. In the case of SK-VI, where the signal search time range has been expanded closer to the neutrino event trigger to accommodate faster neutron capture time with Gd, we further categorize the signal-like candidates into e-like and n-like based on signal time and energy, taking into account the contamination from Michel electrons (refer to Section 3.2 for detailed information).

There is a noticeable difference between the expected and observed values. In the case of BG-like candidates, this difference is approximately 10-20%, and it is attributed to the imperfect sampling of background PMT hits included in our simulation, as illustrated in Figure 4.12. The influence of the time variation in background PMT hit rates is considered through the Am/Be neutron source measurements taken at different times and simulation with varying PMT dark rates.

The differences in the rates of expected and observed n-like candidates are approximately 20-30%. Evidently, the fluctuation across different time periods and algorithms is considerably larger than the statistical uncertainty, necessitating the consideration of signal efficiency variation across operational phases and signal detection algorithms. Naively, a phase-dependence of around 4% is expected when comparing SK-IV and SK-V, and an algorithm-dependence of around 7% when comparing SK-VI and SK-VI (R).

The observed e-like candidate ratio aligns with expectations from our simulation models. The `RECO` vertex mode demonstrates slightly higher efficiency in detecting Michel electrons, as muons tend to propagate much further than neutrons. The incorporation of signal vertex reconstruction aids in identifying Michel electrons produced by energetic muons.

Stage 1 candidate rate per atmospheric neutrino event

SK-IV	Data	MC	Data/MC
BG-like	186.52 ± 0.08	152.29 ± 0.02	1.2247 ± 0.0005
e-like	-	-	-
n-like	0.5245 ± 0.0042	0.6529 ± 0.0014	0.8033 ± 0.0067

SK-V	Data	MC	Data / MC
BG-like	414.70 ± 0.31	449.32 ± 0.03	0.9232 ± 0.0007
e-like	-	-	-
n-like	0.4810 ± 0.0107	0.5745 ± 0.0013	0.8372 ± 0.0187

SK-VI	Data	MC	Data / MC
BG-like	492.31 ± 0.31	452.53 ± 0.04	1.0879 ± 0.0007
e-like	0.1240 ± 0.0049	0.1238 ± 0.0006	1.0014 ± 0.0398
n-like	1.1055 ± 0.0146	1.4342 ± 0.0021	0.7708 ± 0.0102

SK-VI (R)	Data	MC	Data / MC
BG-like	21.803 ± 0.065	20.359 ± 0.008	1.0709 ± 0.003
e-like	0.1323 ± 0.0051	0.1330 ± 0.0006	0.9939 ± 0.0383
n-like	0.8433 ± 0.0128	1.1692 ± 0.0019	0.7213 ± 0.0110

Table 4.3: A breakdown of how candidates triggered at Stage 1 were classified into BG-like, e-like, and μ-like categories through the neural network and Michel-electron rejection cuts. (Statistical errors only)

Figure 4.13 shows the detected neutron capture signal multiplicity per neutrino event. In Figure 4.14, the average signal multiplicity per event is plotted against the date. Clearly, the data shows a larger fraction of events with no detected neutron capture signals and a smaller fraction with multiple detected signals. This indicates a smaller average signal multiplicity in the data compared to the simulation.

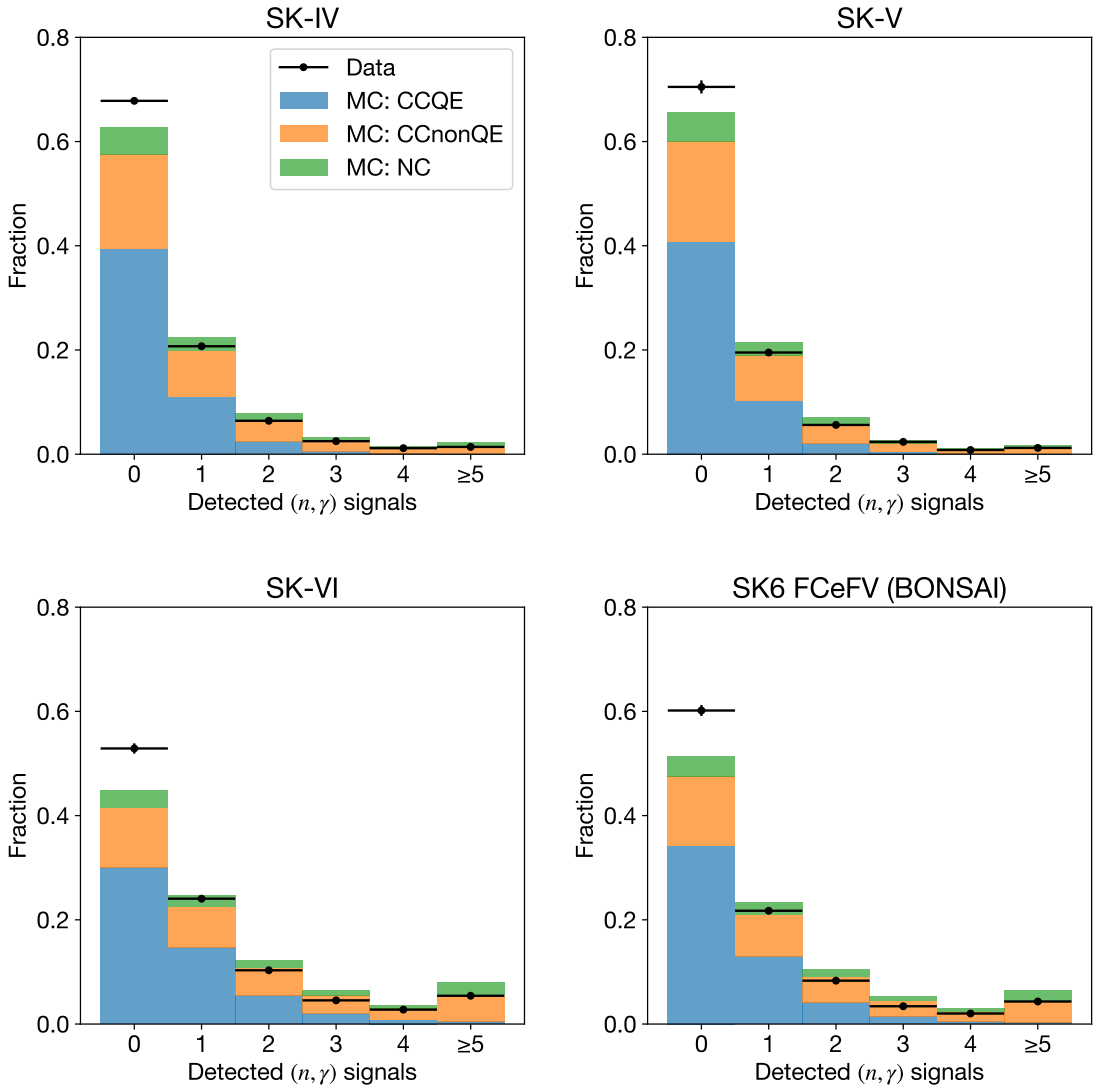


Figure 4.13: Fraction of detected neutron capture signal multiplicity per atmospheric neutrino event for each SK phase.

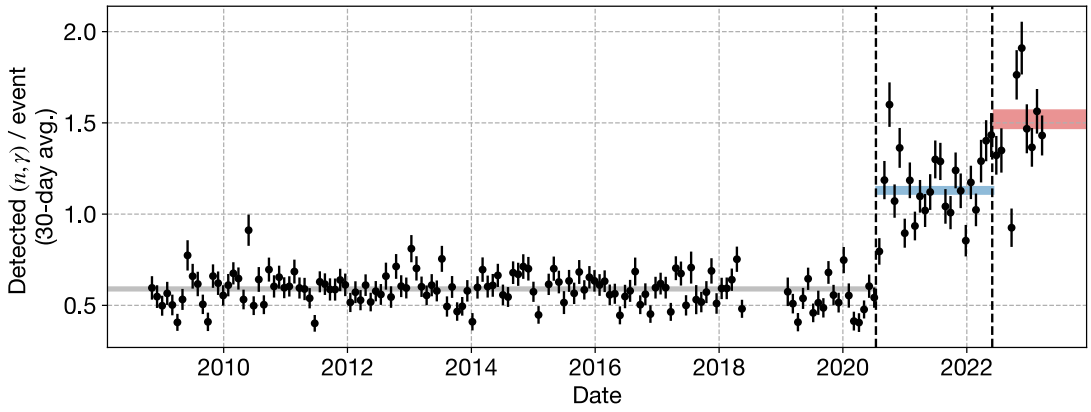


Figure 4.14: Neutron capture signal detection rate per event by date, for pure water phases (SK-IV, V: 2008-2020) and Gd-loaded phases (SK-VI, VII: 2020-present). Colored lines indicate the weighted mean and the statistical uncertainty of the mean.

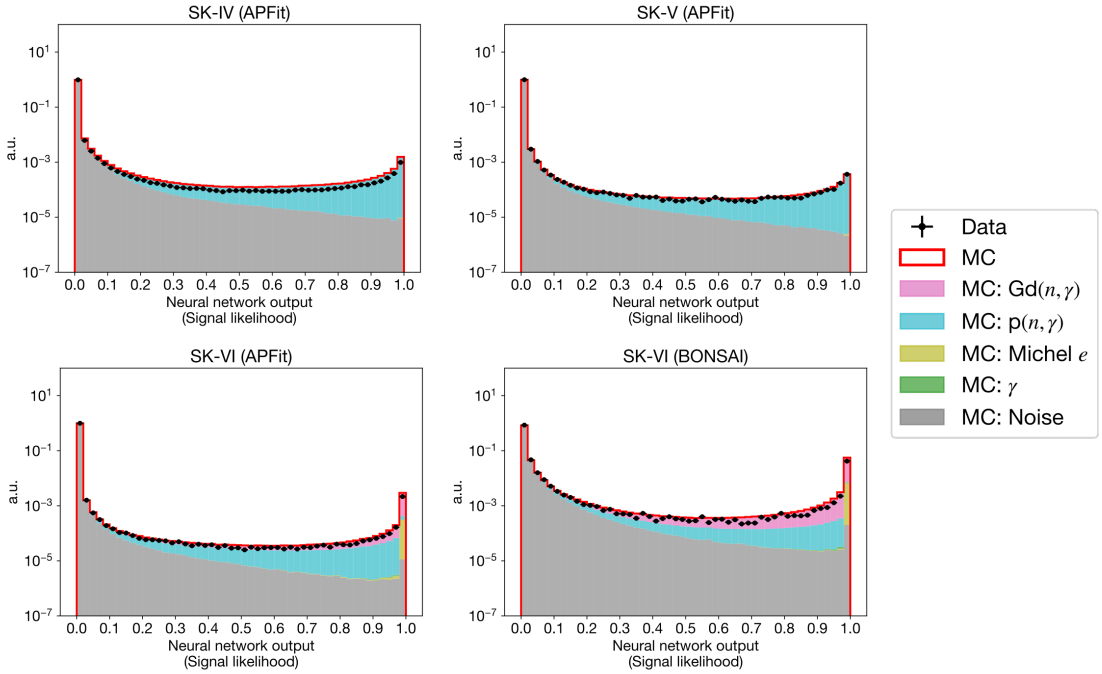


Figure 4.15: Neural network output distributions for all selected candidates in the final data sample and their corresponding simulations, for SK-IV (top left), SK-V (top right), SK-VI (bottom left), and SK-VI (R) (bottom right). All histograms are normalized by the total number of candidates.

To determine whether this discrepancy arises from reduced neutron production or signal detection efficiency, a precise estimation of signal detection efficiency and its uncertainty is essential. Section 4.3 provides a detailed account of this estimation.

Figure 4.15 shows the comparison of the neural network output of candidates selected in the actual data with those selected in simulation. The overall shape agreement is satisfactory, although there is a significant difference in the number of background candidates, particularly in SK-IV, where the background hit rates in data are noticeably different from those in simulation (refer to Figure 4.12).

In Figure 4.16, the false positives per event are displayed for different subsamples of simulated atmospheric neutrino events. The false positive rates per event observed in simulation were low, specifically 0.0249 for SK-IV, 0.0287 for SK-V, 0.0201 for SK-VI, and 0.0134 for SK-VI (R). Most false positives in the pure water phase were attributed to background PMT hits, resulting from accidental coincidences of PMT dark noise. No correlation between false positive rate and neutrino flavor or energy was observed. Since false positives due to background PMT hits are unrelated to neutrino interaction, their contamination can be estimated by fitting the signal time distribution with an exponential function plus a constant background term, as shown in Equation 3.10: $A(1 - e^{-t/\tau_{\mathrm{thermal}}})e^{-t/\tau_{\mathrm{capture}}} + B$.

Figures 4.17 to 4.20 illustrate the feature distributions for all candidates and those successfully passing the neural network selection. These figures show a reasonable shape agreement between the data and simulation.

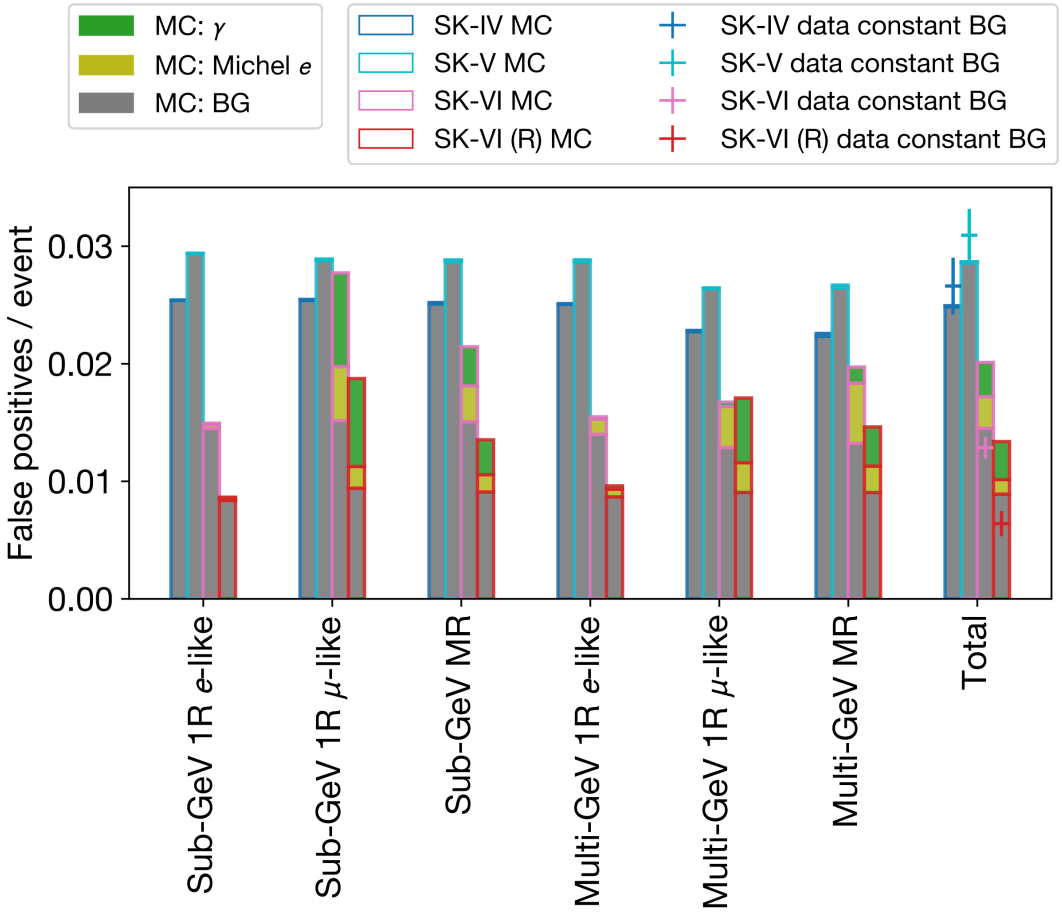


Figure 4.16: False positive rates per event, for each subsample. The stacked bars show the false positives per event in simulation, while the crosses are data estimation based on the constant background term B in Equation 3.10.

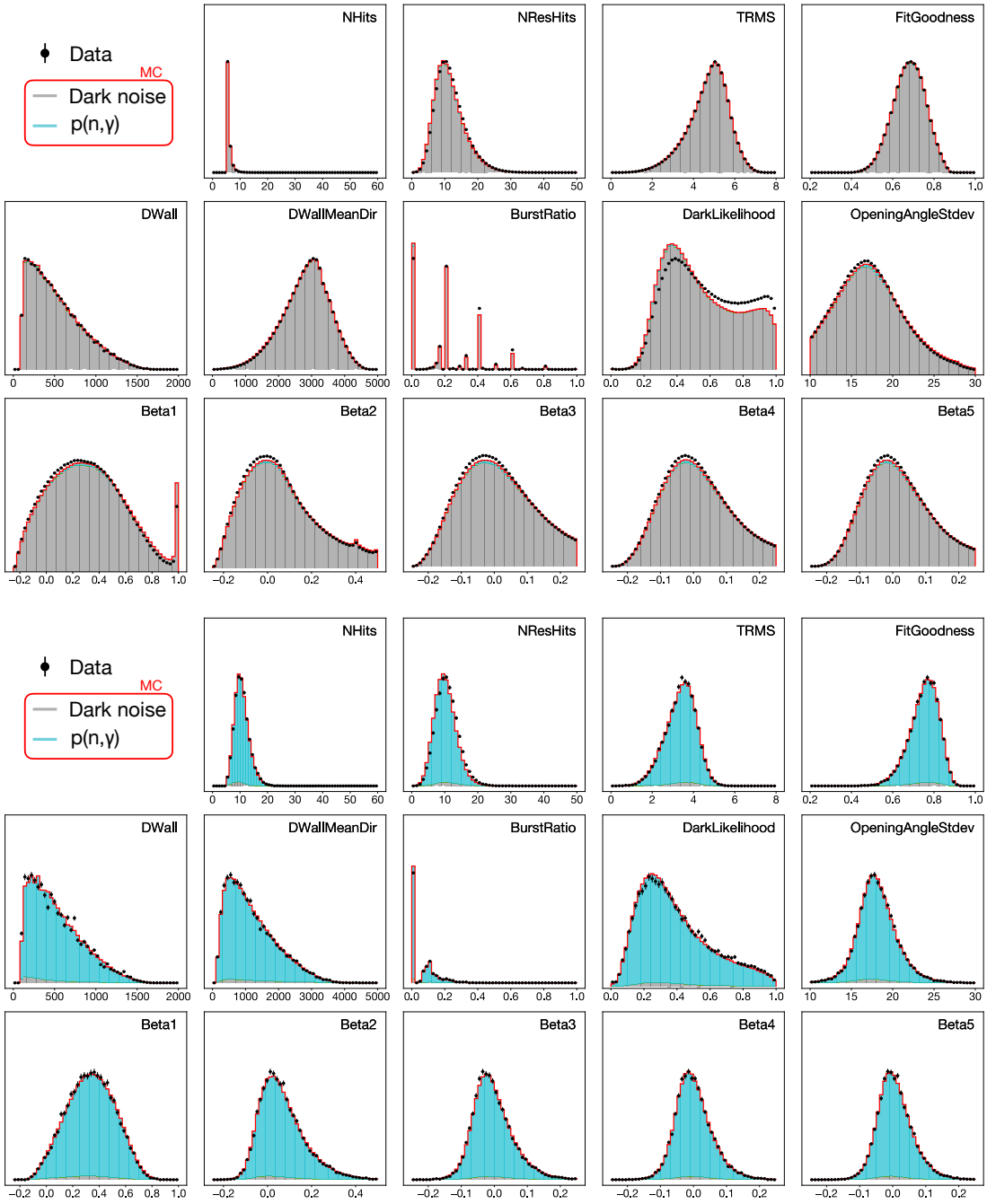


Figure 4.17: The feature distributions of signal candidates before (top) and after (bottom) the neural network selection for SK-IV atmospheric neutrino data and MC simulations. Both data and MC histograms are normalized by area.

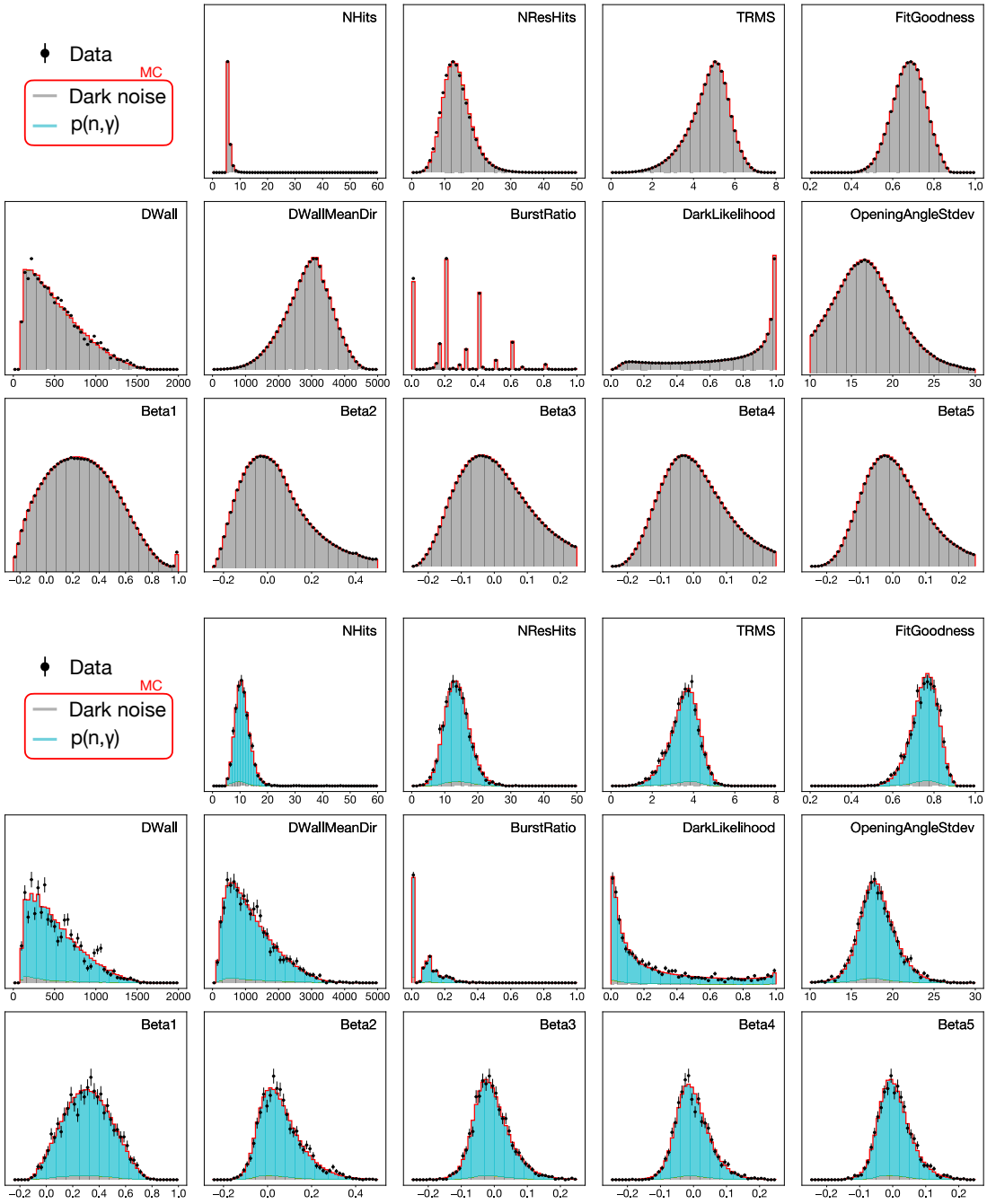


Figure 4.18: The feature distributions of signal candidates before (top) and after (bottom) the neural network selection for SK-V atmospheric neutrino data and MC simulations. Both data and MC histograms are normalized by area.

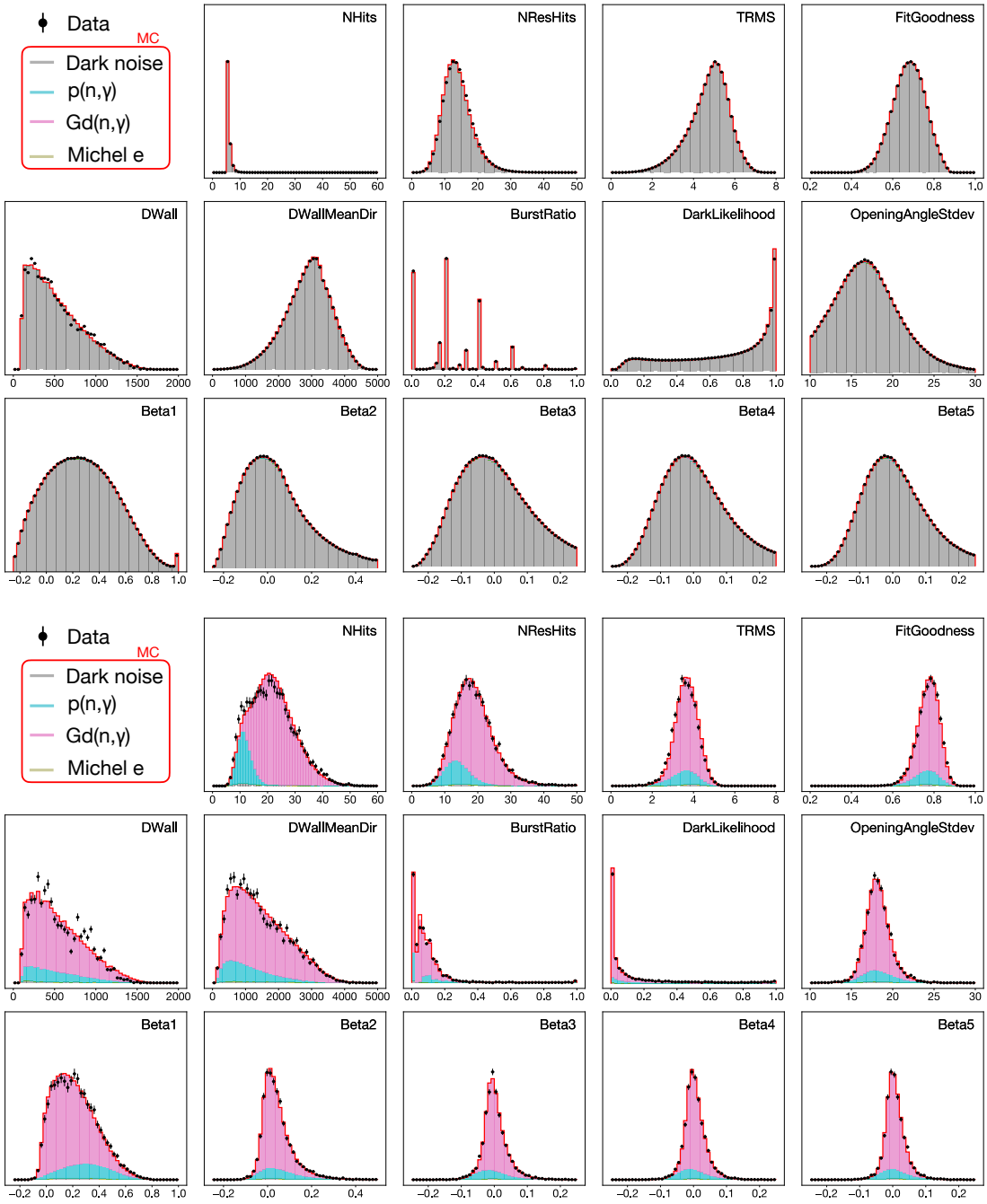


Figure 4.19: The feature distributions of signal candidates before (top) and after (bottom) the neural network selection for SK-VI atmospheric neutrino data and MC simulations. Both data and MC histograms are normalized by area.

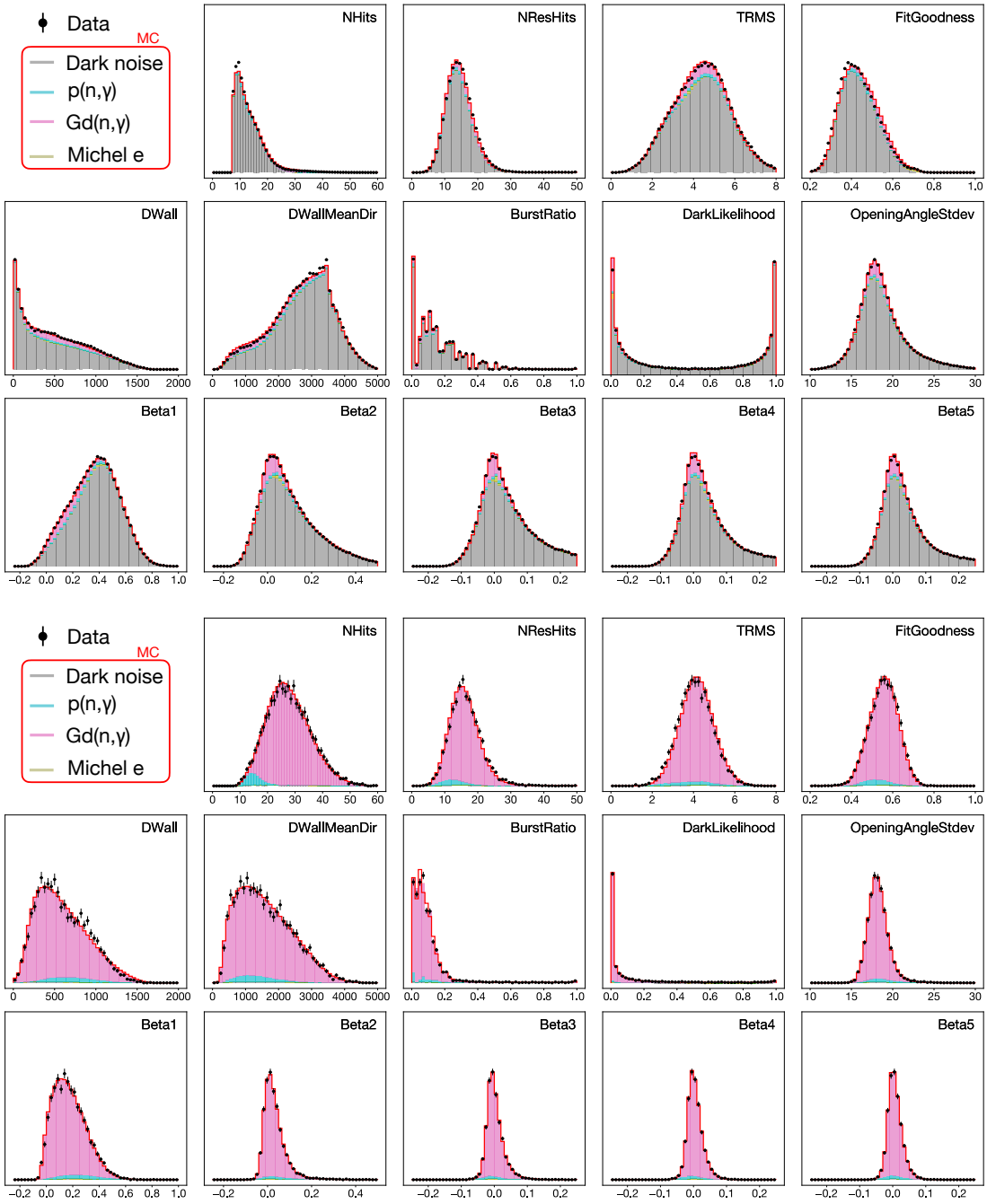


Figure 4.20: The feature distributions of signal candidates before (top) and after (bottom) the neural network selection for SK-VI (R) atmospheric neutrino data and MC simulations. Both data and MC histograms are normalized by area.

Figure 4.21 shows the time distributions of the selected neutron signal candidates. The distributions was fitted with a function of the form $A(1-e^{\tau_{\mathrm{thermal}}})e^{-t/\tau_{\mathrm{capture}}}+B$ (Equation 3.10). Table 4.4 compares the fitted parameters in data and MC, showing reasonable agreement in both time constants τ_{thermal} and τ_{capture} and background term B in all phases, while the normalization term A is in disagreement by 20-30%.

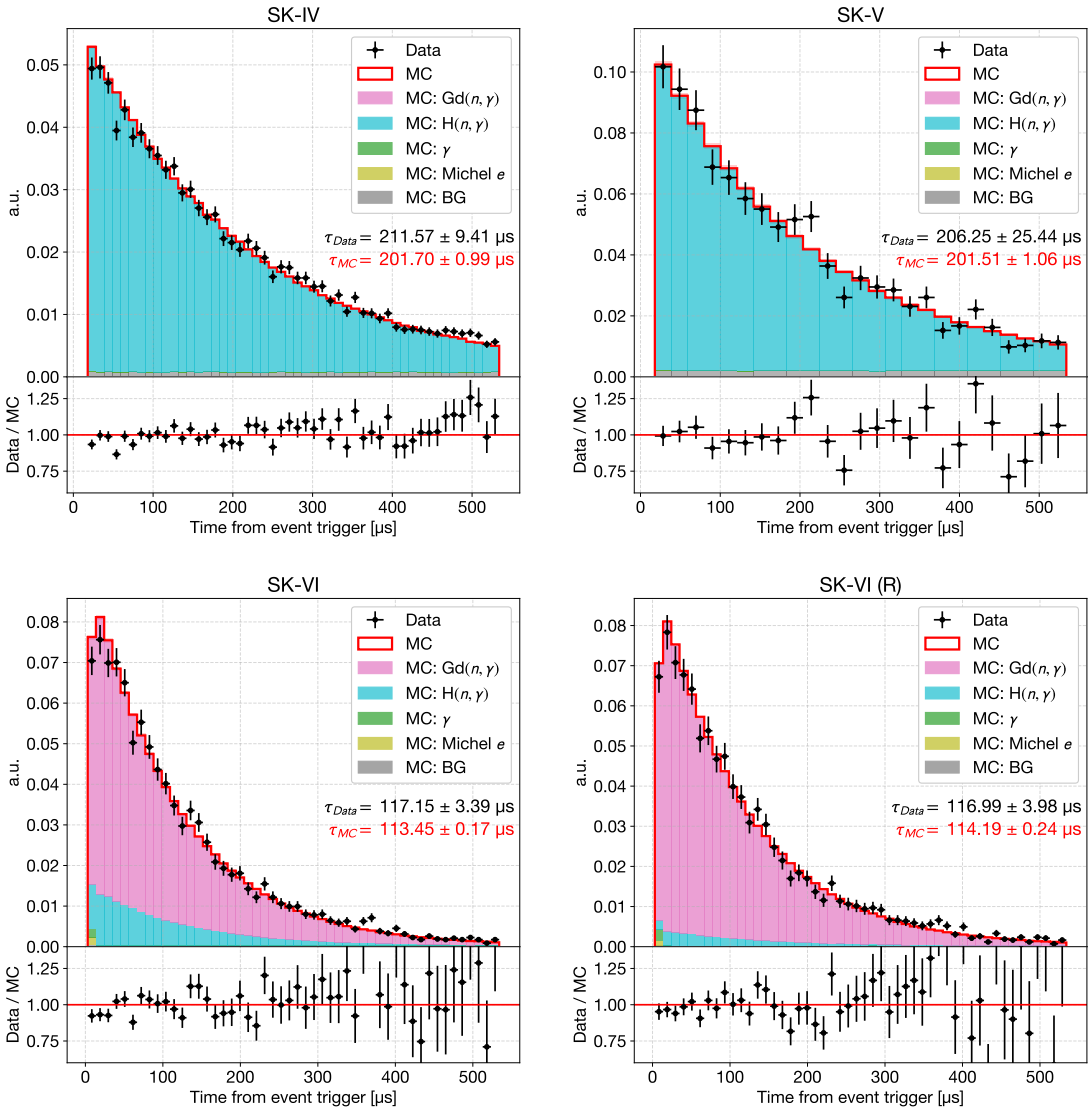


Figure 4.21: Signal candidate time distributions for each SK phase. In this figure, both data and MC distributions are normalized by the detected neutron counts.

Neutron capture time t: fit parameters

$$A(1 - e^{-t/\tau_{\text{capture}}})e^{1-\tau_{\text{thermal}}} + B$$

SK-IV	Data	MC	Data / MC
A $\times N_{\text{bins}}$	1.4539 ± 0.0212	1.9153 ± 0.0034	0.7591 ± 0.0111
τ_{thermal} [μs]	-	-	-
τ_{capture} [μs]	211.84 ± 9.42	201.68 ± 0.99	1.0504 ± 0.0470
B $\times N_{\text{bins}}$	0.0023 ± 0.0179	0.0265 ± 0.0024	0.8570 ± 0.6808

SK-V	Data	MC	Data / MC
A $\times N_{\text{bins}}$	1.3798 ± 0.0587	1.6551 ± 0.0029	0.8336 ± 0.0355
τ_{thermal} [μs]	-	-	-
τ_{capture} [μs]	206.28 ± 25.46	201.60 ± 1.00	1.0232 ± 0.1264
B $\times N_{\text{bins}}$	0.0110 ± 0.0450	0.0309 ± 0.0021	0.3559 ± 1.4575

SK-VI	Data	MC	Data / MC
A $\times N_{\text{bins}}$	5.2064 ± 0.1356	7.6744 ± 0.0095	0.7360 ± 0.0192
τ_{thermal} [μs]	5.3121 ± 0.7442	4.4432 ± 0.0379	1.1956 ± 0.1778
τ_{capture} [μs]	117.16 ± 3.39	113.47 ± 0.16	1.0324 ± 0.0299
B $\times N_{\text{bins}}$	0.0128 ± 0.0146	0.0128 ± 0.0009	1.0038 ± 1.1448

SK-VI (R)	Data	MC	Data / MC
A $\times N_{\text{bins}}$	3.9756 ± 0.1218	5.8000 ± 0.1089	0.6854 ± 0.0210
τ_{thermal} [μs]	5.6777 ± 0.8795	5.4440 ± 0.0531	1.0429 ± 0.1619
τ_{capture} [μs]	118.16 ± 3.99	114.22 ± 0.23	1.0345 ± 0.0350
B $\times N_{\text{bins}}$	0.0016 ± 0.0128	0.0060 ± 0.0010	0.2679 ± 2.1213

Table 4.4: Best-fit parameters and their 1σ uncertainties obtained from fitting the detected neutron capture time distribution using Equation 3.10.

In SK-VI with 8 MeV Gd(n, γ) signals, we can independently reconstruct the neutron capture vertex using the likelihood fitter `BONSAI`, as detailed in Section 2.4.2. The radial and vertical vertex distributions of candidates selected with the `RECO` vertex mode are shown in Figure 4.22. While the signals are expected to be uniformly distributed throughout the ID fiducial volume, the detected signals with the reconstructed neutron capture vertex close to the tank wall are significantly fewer. This tendency is consistent with the expectations from SK-VI simulation.

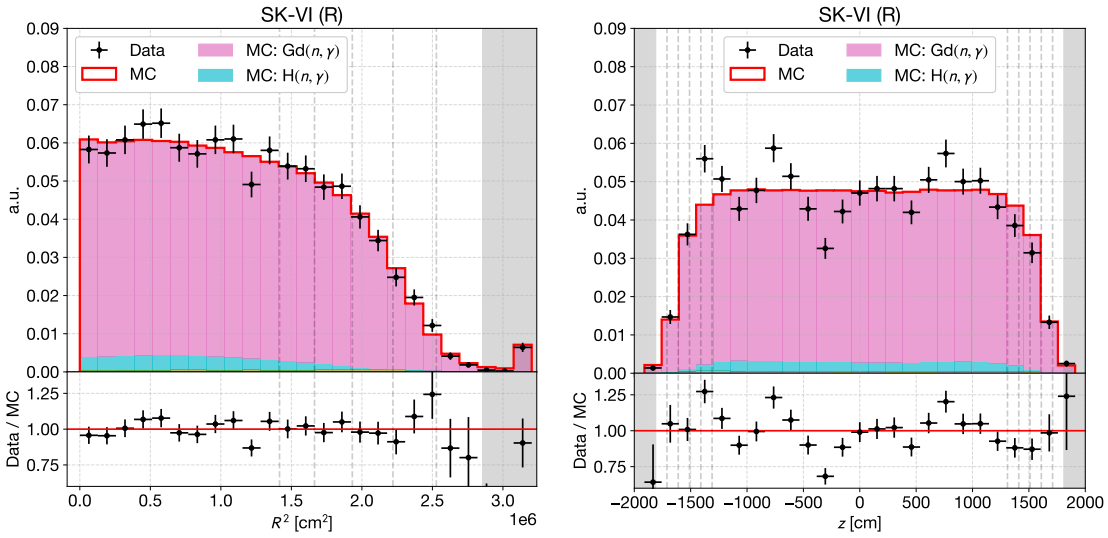


Figure 4.22: Radial (R^2, left) and vertical (z, right) vertex distributions of signal candidates selected in SK-VI with the `RECO` vertex mode. The gray shades indicate regions outside of the ID, and the dashed lines each indicate 1, 2, 3, 4, 5 m away from the ID wall. The fiducial volume for neutrino event selection is defined as 1 m away from the ID wall. There is no limits with regards to the reconstructed neutron capture vertex. Signals with reconstructed vertices outside the ID tank indicate the reconstruction algorithm failed to find a good vertex within the ID.

We examined the distance between the reconstructed neutron capture vertex and the reconstructed neutrino interaction vertex as a proxy for outgoing neutron kinematics and vertex reconstruction performances. The distributions of both data and simulation are compared in Figure 4.23, showing good agreement in shape. Candidates classified as signals but with a reconstructed distance larger than 10-20 m are mostly attributed to pure background PMT hits. The majority of neutrons were captured within 5 m from the reconstructed neutrino interaction vertex.

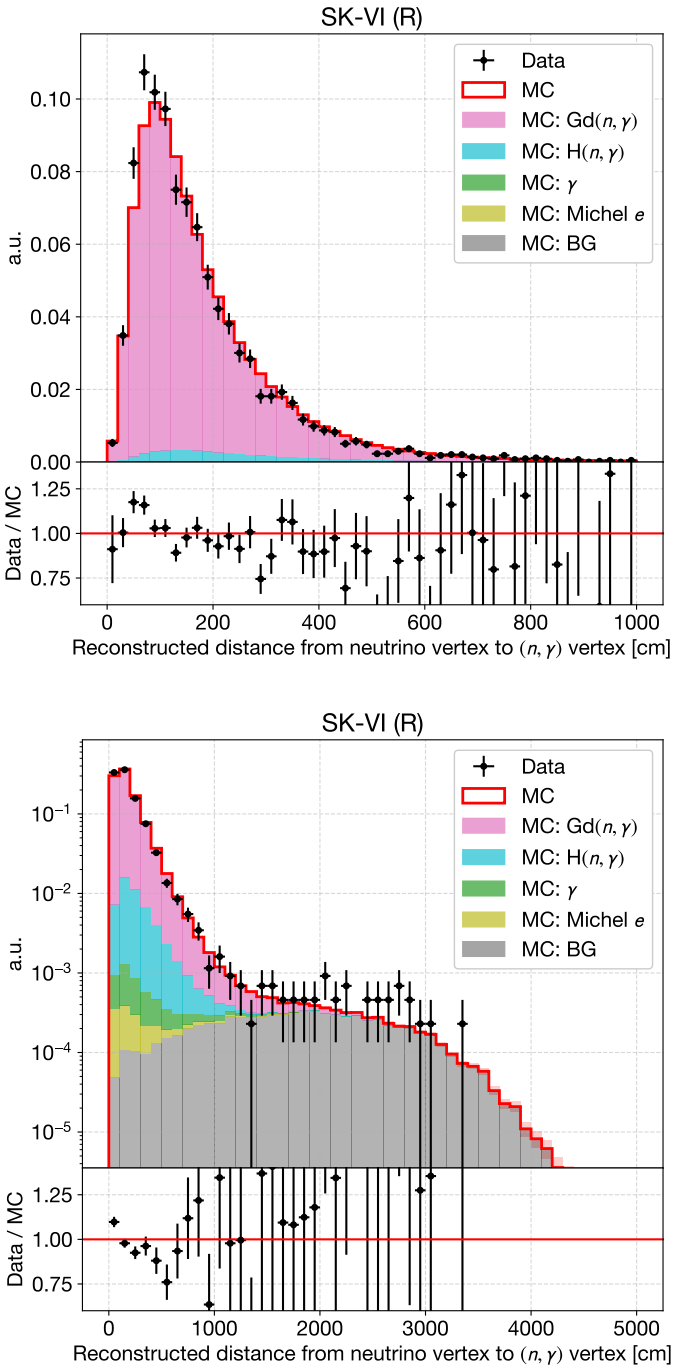


Figure 4.23: The reconstructed distance between neutron captures to neutrino interaction vertex, in linear scale (top) and in logarithmic scale (bottom).

Lastly, the angular correlation between neutron capture and lepton signals is investigated. In Figure 4.24, the cosine of angles between the reconstructed lepton direction and the reconstructed neutron direction was examined for single-ring events. Here, the reconstructed neutron direction is represented by a vector connecting the reconstructed neutrino interaction vertex and the reconstructed neutron capture vertex. The distribution appears mostly flat, with the data indicating a slight prevalence of forward-scattered neutrons, primarily attributed to neutrons from muon captures on ^{16}O.

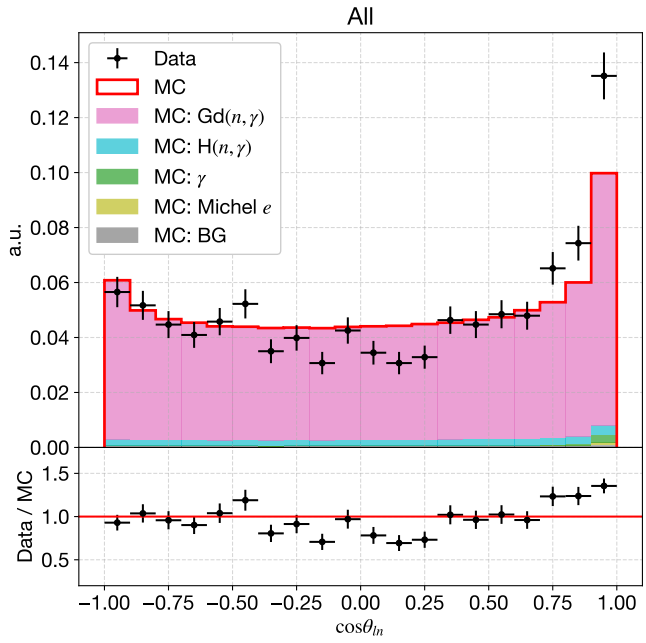


Figure 4.24: The distribution of cosine of angles between the reconstructed lepton direction and the reconstructed neutron direction, as observed across all single-ring events in SK-VI. The "forward-scattered" neutrons are indeed mostly emitted from "forward-going" muon captures on ^{16}O.

Figure 4.25 shows the angular correlation between charged lepton and neutron captures, divided into four single-ring subsamples: Sub-GeV/Multi-GeV e-like/μ-like. In the Sub-GeV e-like category, where neither muons nor pions are abundantly produced, we observe a diminished forward-going peak (near $\cos\theta_{ln} = 1$) compared to the other three subsamples, where muons or pions are more prevalent.

The data reveals a slightly higher fraction of neutrons captured in the forward direction (aligned with the lepton) compared to our simulation. This discrepancy may arise from various uncertainties in our simulation models, such as pion production, muon capture, and neutron scattering angle uncertainties.

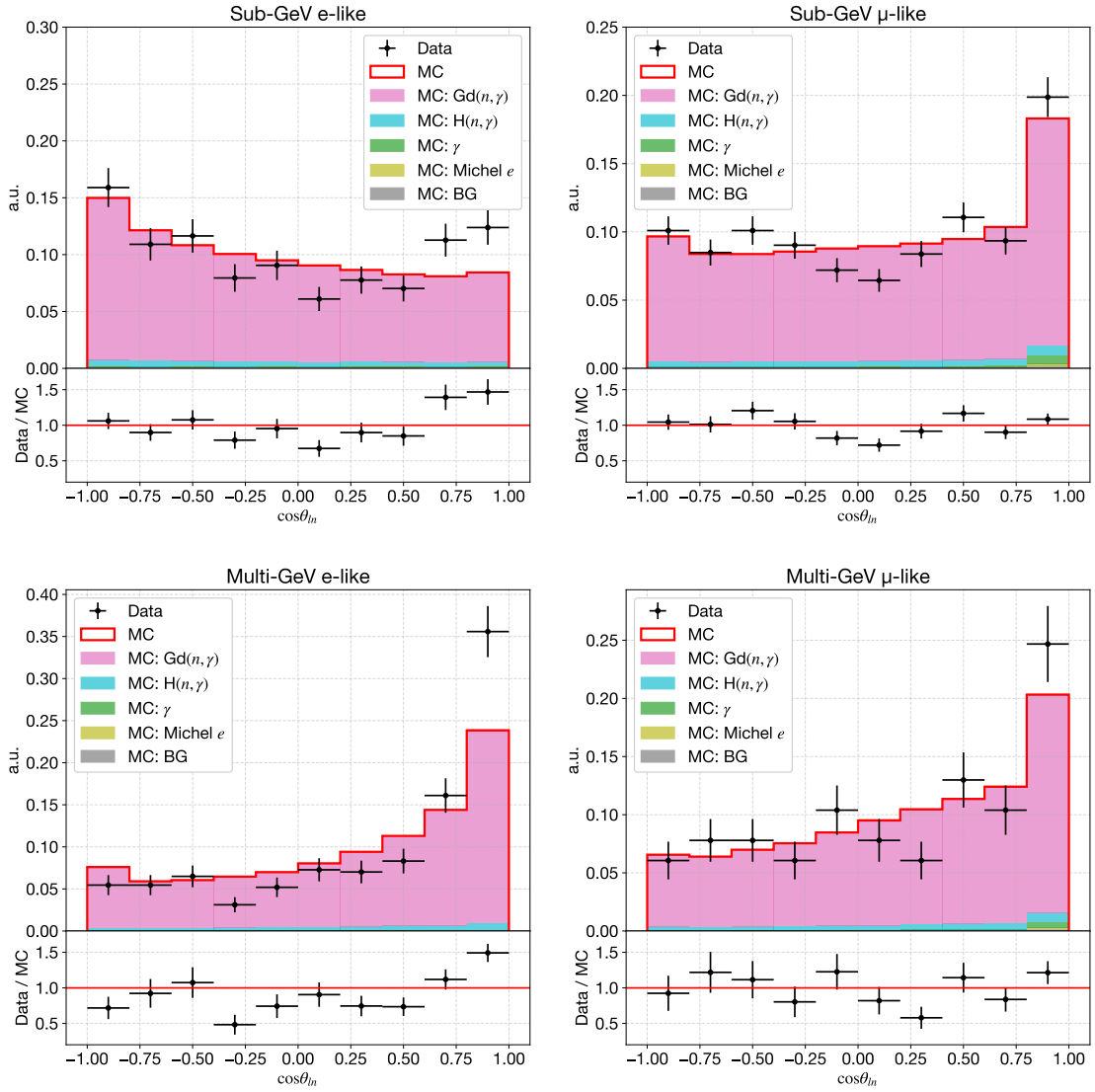


Figure 4.25: The distribution of cosine of angles between the reconstructed lepton direction and the reconstructed neutron direction in single-ring sub-GeV e-like (top left), μ-like (top right), multi-GeV e-like (bottom left), and μ-like events (bottom right) in SK-VI. The "forward-scattered" neutrons are indeed mostly emitted from "forward-going" muon captures on ^{16}O. Muons are either from muon neutrino CC interaction or from π/K decays.

4.3 Estimation of signal detection performance

It is crucial to accurately estimate the performance of neutron capture detection and its associated uncertainty when measuring neutron properties. For instance, the average neutron capture multiplicity per atmospheric neutrino event can be computed as follows:

$$\langle n \rangle = \left\langle \frac{N_i^{\text{detected}} - N_i^{\text{BG}}}{\epsilon_i} \right\rangle \tag{4.1}$$

Here, $\langle n \rangle$ represents the average neutron capture multiplicity per neutrino event, N_i^{detected} is the count of detected neutron captures in the i^{th} event, N_i^{BG} is the estimated number of backgrounds or false positives in the i^{th} event, and ϵ_i is the estimated signal detection efficiency for the i^{th} event. N_i^{BG} and ϵ_i are the two signal detection performance metrics that need to be accurately estimated.

We initially rely on MC simulation to estimate performance on an event-by-event basis. In terms of the MC-estimated signal efficiency, we have already observed non-negligible discrepancies between SK-VI Am/Be calibration data and its simulation (as discussed in Section 3.3.7), although these variations fall within our uncertainty estimates. To address this issue, we adjust the overall signal detection efficiency estimated through simulation by applying a scaling factor that best accounts for the efficiency values obtained from calibration data. In the context of neutron capture multiplicity, the uncertainty associated with this scaling factor directly influences the uncertainty in the final results.

Regarding the false positive rate per event, we base our estimates on values obtained from simulation without applying data-based corrections. This approach is justified because we utilize actual background PMT hits recorded by our detector to model backgrounds. The uncertainty in the false positive rate mostly arises from variations in PMT dark rates over time. Additionally, for SK-VI, uncertainties in the quantities of Michel electrons and gamma-rays resulting from nuclear de-excitation may also influence the false positive rate.

4.3.1 Dependence on neutrino event variables

Signal efficiency notably varies with neutrino energy due to deviations between the reconstructed and actual neutron capture vertices, especially for higher-energy neutrino events. The uncertainty in the reconstructed neutrino interaction vertex increases with higher energy, influenced by interference among multiple Cherenkov rings during reconstruction. Additionally, as neutrons gain more kinetic energy, their capture points shift farther from their production locations. Figure 4.26 illustrates the variation in signal efficiency between sub-GeV and multi-GeV events.

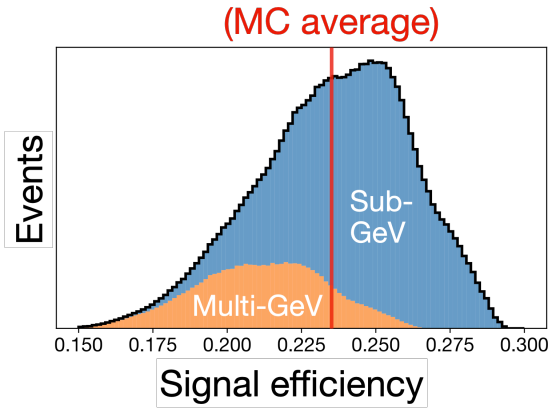


Figure 4.26: Distribution of signal efficiencies estimated for each SK-IV MC-simulated event. Details of the estimation method are provided in Section 4.3.2.

Assigning the average signal efficiency for all events in estimating the average signal multiplicity (Equation 4.1) could introduce opposite sign biases for sub-GeV and multi-GeV events. Therefore, minimizing systematic uncertainty requires a comprehensive understanding of factors influencing signal detection performance and modeling them to the best extent possible.

To understand factors influencing signal detection, we binned MC-simulated events based on five reconstructed observables in each neutrino event: visible energy, the number of Cherenkov rings, neutrino flavor (flavor of the most energetic ring), and the radial (R^2) and vertical (z) position of the reconstructed neutrino interaction vertex. This resulted in 19,404 bins across five dimensions, covering approximately 2 million MC-simulated atmospheric neutrino events meeting event selection criteria. Table 4.5 summarizes the binning strategy.

Reconstructed variable	Type	Range	Bins	Spacing
Visible energy [MeV]	Continuous	$[10^{-5}, 10^{5}]$	11	Log
Number of Cherenkov rings	Discrete	{Single, Multi}	2	Linear
Neutrino flavor	Discrete	$\{e, \mu\}$	2	Linear
Vertex R^2 [cm^2]	Continuous	$[0, 1690^2]$	21	Linear
Vertex z [cm]	Continuous	[-2200, 2200]	21	Linear

Table 4.5: Binning strategy used in MC simulation to explore how signal detection performance varies with reconstructed neutrino event variables.

We then estimated average signal efficiency and false positive rate per event for each bin, including associated MC statistical errors.

Projected onto a single feature axis (e.g., visible energy, radial and vertical coordinates of the reconstructed neutrino interaction vertex), binned performance metrics are illustrated in Figures 4.28 through 4.29. Signal efficiency notably depends on neutrino event visible energy across all SK phases and on the reconstructed neutrino interaction vertex, particularly its proximity to the tank wall. A substantial drop in signal efficiency is observed when the interaction vertex is too close to the tank wall, possibly due to difficulty distinguishing signals near the tank wall from background radioactivity in PMT glass or surrounding rock. In contrast, the false positive rate is mostly independent of neutrino visible energy but increases for a vertex close to the tank wall. Compared to visible energy and the interaction vertex, the number of rings and ring particle type have a smaller impact, though signal efficiency is slightly lower for multi-ring and e-like events, likely due to worse vertex resolution compared to single-ring and μ-like events.

Utilizing the estimated average performance metrics for each bin, we may evaluate signal detection performance for a given data event by extracting the five reconstructed observables and referencing the corresponding bin for average signal efficiency and false positive rate. However, this approach prompts questions about handling empty bins, events with no corresponding bin, determining an optimal binning scheme considering MC statistics, and addressing bin-to-bin correlation alongside statistical uncertainty.

4.3.2 Regression on neutrino event variables

To address these issues, we employed a non-linear multivariate regression model called the Generalized Additive Model (GAM) to obtain reliable performance estimates for each unseen data event, irrespective of the binning scheme or MC statistics. The simplest linear GAM assumes that the expected output $E[y(x)]$ for given features x is related by:

$$E[y] = b + \sum_{i=0}^{N} f_i(x_i) = b + \sum_{i=1}^{N} \sum_{j=1}^{M} a_i^j B_i^j(x_i) \tag{4.2}$$

Here, $f_i(x_i)$ represents the contribution from the i^{th} feature x_i, a_i^j are weights for each j^{th} basis spline $B_i^j(x_i)$ for x_i, b is a constant bias term, N is the number of features, and M is the number of basis splines per feature. For simplicity, the features x_i are considered independent. Once the form and number of splines are determined, the parameters a_i^j and b can be fitted to best describe (x, y) data points using a straightforward least-square method, as illustrated in Figure 4.27.

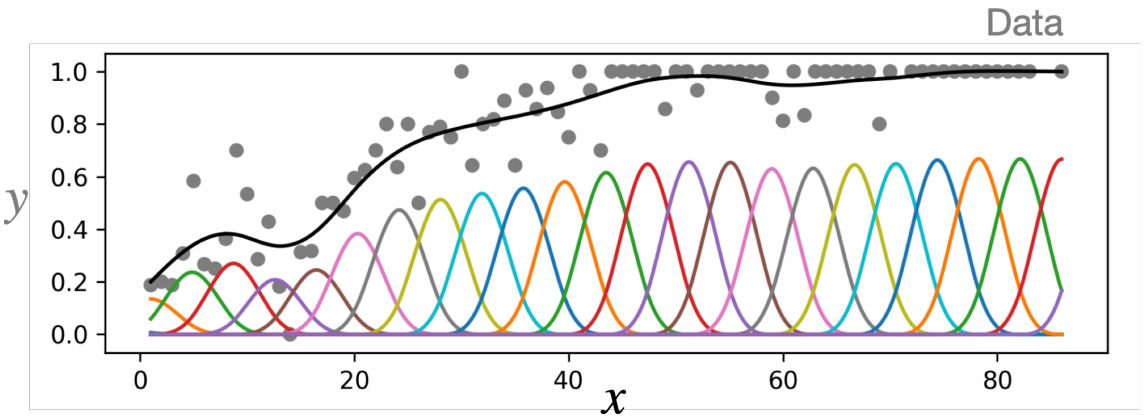


Figure 4.27: Illustration of non-linear regression on a single feature x using GAM. The weights assigned to basis splines are adjusted so that the sum of the splines best fit the gray data points, minimizing the sum of squared residuals.

The GAM with basis splines can capture any non-linear relationships between input and output using a simple least-square fitting method. The contributions of each output, $f_i(x_i)$, are straightforwardly added (hence the term "generalized additive") to predict the final output. This simplicity provides greater interpretability. Using this technique, we can construct regression models that model the signal detection performance as a function of neutrino event's reconstructed variables.

We model two signal detection performance metrics, the efficiency and false positive (background) rate. So we have two regression models as in Equations 4.4, for SK-IV, SK-V, SK-VI, and SK-VI (R), resulting in total 8 GAM models.

$$\text{Expected signal efficiency}_{\text{SK-}x} = f_{\text{SK-}x}(E_{\text{vis}}, N_{\text{ring}}, \text{Flavor}, R^2, z) \tag{4.3}$$

$$\text{Expected false positive rate}_{\text{SK-}x} = g_{\text{SK-}x}(E_{\text{vis}}, N_{\text{ring}}, \text{Flavor}, R^2, z) \tag{4.4}$$

Here $x \in \{\text{IV, V, VI, VI(R)}\}$. The distinct models $f_{\text{SK}-x}$ and $g_{\text{SK}-x}$ are the GAM models that are five-dimensional extension of the function described in Equation 4.2, and the parameters such as the weights and biases need to be fitted to some input using the least-square method as mentioned earlier. The input we can use to fit the parameters are the binned average performance metrics from our simulation.

We used `LinearGAM` implemented in `pyGAM` 0.9.0 [156] to construct each of the 8 GAMs. To handle continuous variables like visible energy and vertex coordinates, we applied 20 cubic splines per feature. Regression on binary variables, such as the number of rings (single-ring and multi-ring) or particle type (e-like and μ-like), were modeled with two Heaviside step functions.

Smoothing was applied by imposing a penalty on the second derivative of the fitted function to prevent overfitting. The hyperparameter λ_i, determining the size of this penalty for each feature x_i, was heuristically determined by monitoring the fit with different sets of λ. A default value of $\lambda = 0.001$ was used, except for the case of fitting false positive rate against continuous features. For example, λ =1,000 was applied for visible energy, and $\lambda = 0.1$ for vertex coordinates, as smaller values of λ led to overfitting — unphysical cubic spikes showed up in the fitted curve.

Central bin values were used as feature values x during the fitting process. The average signal efficiency and the logarithm (to ensure the fit result is always positive) of the average false positive rate in each bin served as the response variable y. The inverse of the squared ratio between the average performance metric and its statistical error in each bin was used as the weighting factor for each data point. It is important to note that only data points with non-zero signal efficiency and false positive rate were included in the fitting procedure.

Figures 4.28 and 4.29 showcase the GAM fit — a 1-dimensional projection of the best-fit results, along with the overlaid 1σ prediction interval onto the binned performance metrics. The partial dependence on feature x_i ($f_i(x_i)$ in Equation 4.2) is distinctive for each GAM fitted in different SK phases but is shared across all subsamples within the same phase without considering feature correlations. Despite not incorporating feature correlations, our GAMs effectively capture trends in binned performance metrics across various subsamples.

Figure 4.30 illustrates how well the fitted GAM estimates the average signal multiplicity in each SK phase using MC simulation data, spanning the visible energy range from 30 MeV to 100 GeV. The GAM-derived signal multiplicities match the true values well, deviating less than 3% for visible energy greater than 100 MeV. However, for lower energy events, the GAM tends to underestimate signal efficiency by 5-10%, leading to an overestimation of signal multiplicity by the same margin.

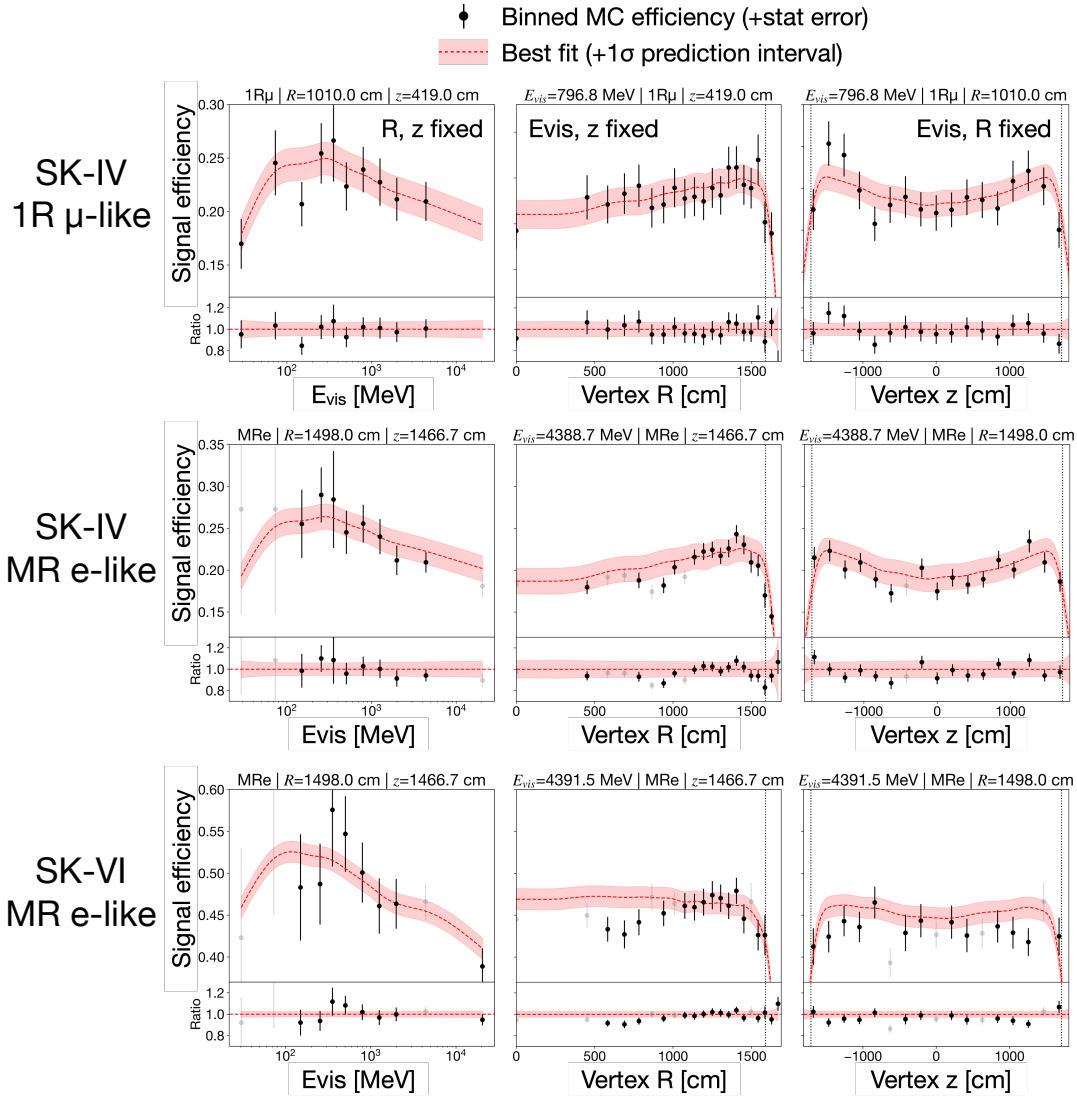


Figure 4.28: Partial dependence plots illustrating the average signal efficiency in the MC-simulated SK-IV single-ring μ-like (top), SK-IV multi-ring e-like (center), and SK-VI multi-ring e-like sample (bottom). The plots vary with visible energy (left), radial (R^2, center), and vertical position (z, right) of the reconstructed neutrino interaction vertex. Each black dot on the plots represents the average signal efficiency of MC events in its corresponding bin, with error bars indicating statistical uncertainty. In each plot, variables other than the one on the horizontal axis are held constant at the values specified above. The red dashed line represents the GAM that best fits the binned signal efficiencies, and the red shaded area depicts the 1σ prediction interval of the fitted GAM. Gray dots denote binned values not used in the fit due to either efficiency or false positive rate being zero.

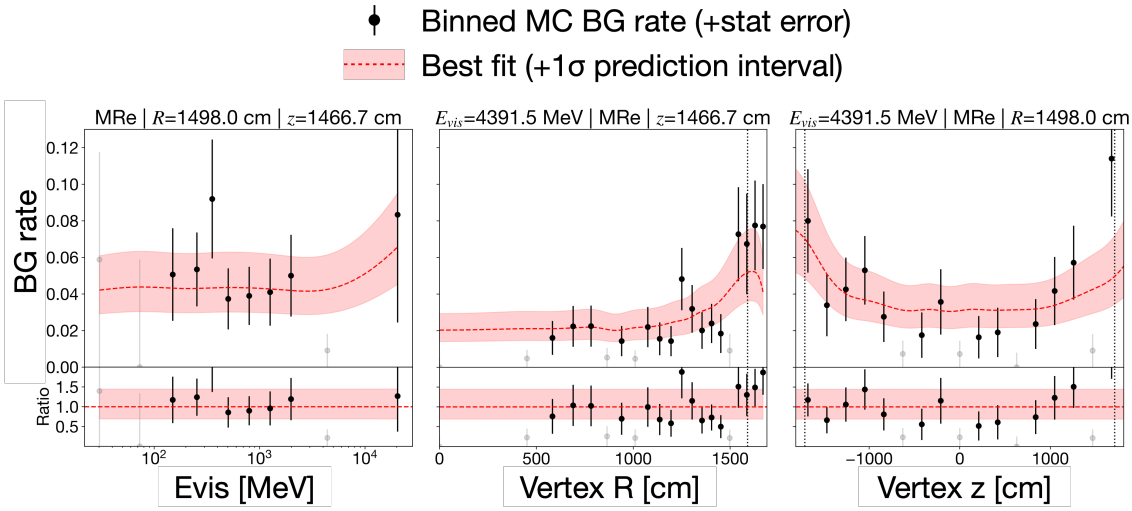


Figure 4.29: Partial dependence plots for the average MC false positive rate per event in the MC-simulated SK-VI multi-ring e-like sample.

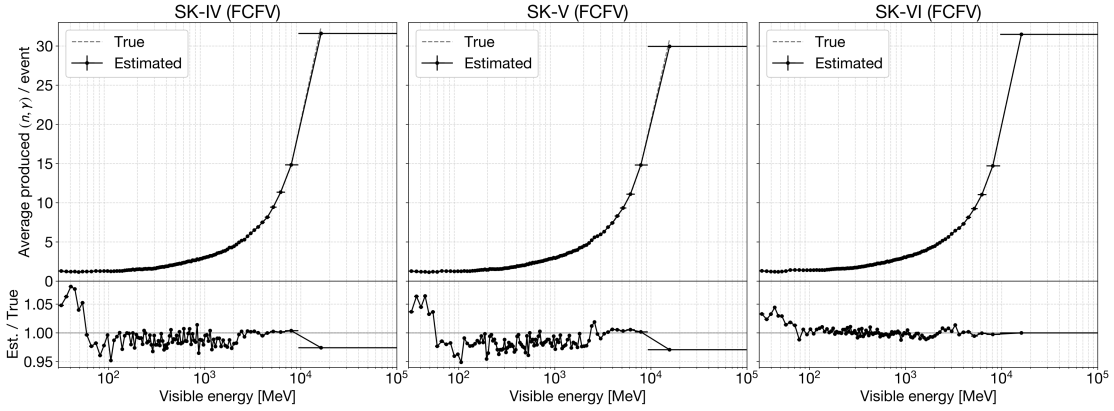


Figure 4.30: Comparison of true and estimated average signal multiplicity as a function of visible energy, for each SK phase simulation.

Our GAM does not account for feature correlation, potentially leading to less precise capture of trends between response variables and features across subsamples. To assess GAM-based measurements, we applied this method to estimate the average true neutron capture multiplicity in an unseen MC simulation generated with an older version of `NEUT` 5.1.3. It predicts about 20% less neutron production than the nominal simulation against which the GAM was fitted. Reconstructed event variables from the unseen simulated events were used to regress performance metrics, plugged into Equation 4.1 for estimating signal detection performance —efficiency, false positive rate, and their 1σ uncertainty — on an event-by-event basis.

Figure 4.31 compares the performance of three different methods in estimating the average neutron capture multiplicity as a function of neutrino event visible energy. This evaluation was conducted on the aforementioned unseen MC simulation with $\sim 20\%$ smaller neutron multiplicity and varied detector modeling compared to our nominal MC. As anticipated (see Figure 4.26), relying on a single value for all events significantly skews the curve. Utilizing binned performance metrics for each event yields improved accuracy compared to using a single value. However, in bins with limited MC statistics, results become less reliable and more erratic due to statistical errors. The fitted GAM, offering a smooth prediction of performance metrics, shows the most reliable multiplicity estimation performance. The slight deviation between true and estimated multiplicities for each visible energy bin should be considered a systematic uncertainty.

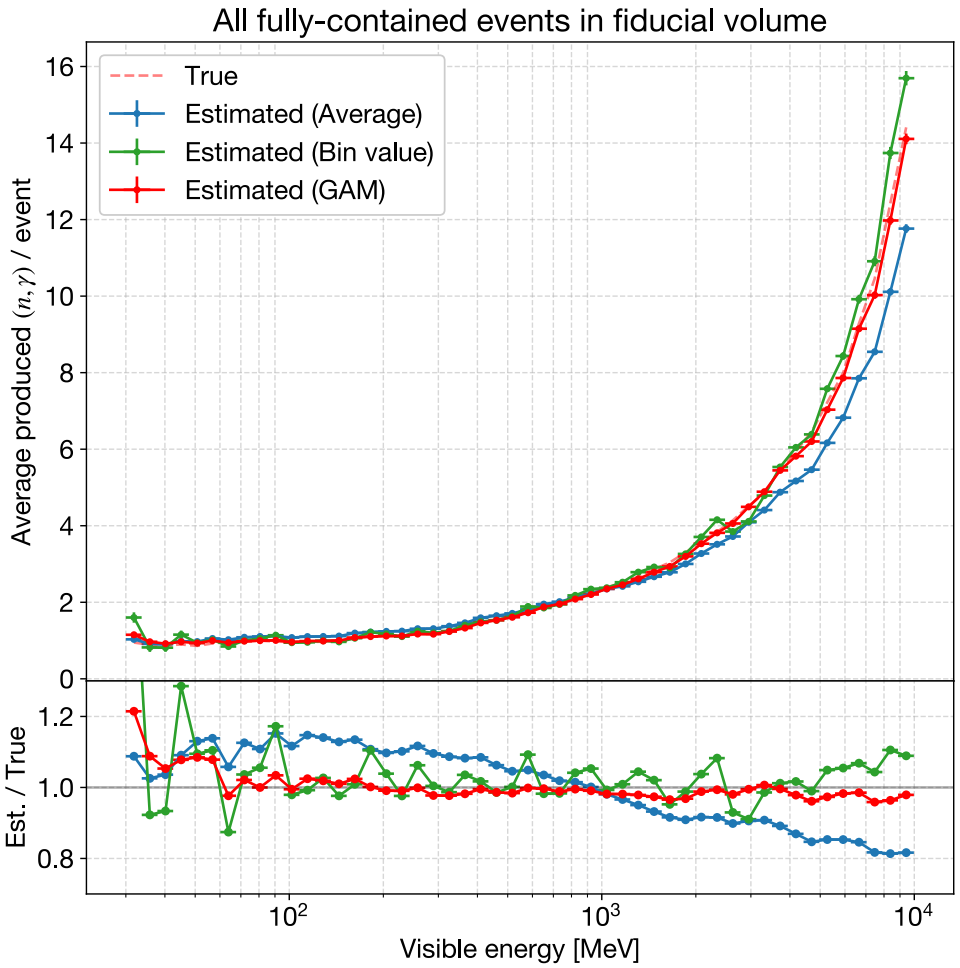


Figure 4.31: Comparison of three methods for estimating signal detection performance in an unseen SK-IV MC simulation with distinct neutron multiplicity and detector modeling from our nominal MC. The average signal multiplicity is plotted against visible energy, with signal efficiency ϵ and expected false positives N^{BG} from Equation 4.1 estimated on an event-by-event basis. The blue curve assumes a single average signal efficiency and false positive rate for all events. The green curve involves looking up average signal detection performance metrics binned by reconstructed event variables. The red curve uses non-linear regression on binned average performance metrics with a GAM. For the red and green curves, binned performance metrics were extracted from the nominal SK-IV MC simulation.

Calibration-based correction

One final step remains before incorporating the estimated performances into our analysis. The MC signal detection efficiency, as estimated by the fitted GAM, is adjusted using the overall scaling factor shown in Table 3.14. This adjustment is implemented to address the discrepancy in the observed and expected signal efficiencies in the Am/Be neutron source calibration.

4.4 Average neutron capture multiplicity

We measured the average neutron capture multiplicity per atmospheric neutrino event using 39,376 selected atmospheric neutrino events observed over 4269.8 days of detector livetime from 2008 to 2022. The number of detected neutron capture signals was adjusted by accounting for the expected false positives and signal detection efficiency, as described by Equation 4.1. Neutrino events were binned by their visible energy, and the average signal multiplicity was calculated for each bin to characterize neutron production as a function of neutrino kinematics.

Statistical uncertainties are discussed in Appendix D.

4.4.1 Systematic uncertainties

Systematic uncertainties are classified into three types:

(1) **Signal efficiency mean**

This uncertainty is related to the mean (or scale) of the estimated signal efficiency. Notably, this uncertainty is independent of event statistics, The uncertainty due to signal efficiency "spread", which is averaged down to $O(0.1)\%$ due to large statistics, is discussed in Appendix D.2.

(2) **Signal efficiency regression with GAM**

This uncertainty arises from systematic imperfections in the performance regression using GAM. Factors contributing to this uncertainty include assumptions of no correlation of features, deviations between the fitted GAM and the binned performance metrics.

(3) **Neutrino event reconstruction**

This uncertainty arises from the visible energy reconstruction, representing the horizontal axis of our measurement. As we conduct measurements for both single-ring and multi-ring subsamples, we also account for the uncertainty arising from the classification of single-ring vs. multi-ring events.

(1) Signal efficiency mean

We adjust the signal efficiency from simulation using a scaling factor, calculated as the ratio between estimated efficiencies from Am/Be data and simulation (details in Table 3.14). The uncertainties associated with these correction factors are treated as systematic uncertainty in the signal efficiency mean. Notably, these uncertainties consider differences in the background PMT hit rate between data and simulation.

Comparing signal multiplicities across SK phases, the top figure in Figure 4.32 displays average signal multiplicity distributions with GAM-estimated efficiency correction applied. The bottom left figure shows the result of additionally applying efficiency scaling factors from Am/Be calibration (Table 3.14). While the agreement between SK-V and SK-VI improves, there is still room for better agreement between SK-IV and the others[1].

To address this discrepancy, we found a set of correction factors that ensures the best consistency among different phase datasets, minimizing the χ^2 between any two datasets. The results of SK-VI (R) were fixed and used as a reference. The obtained scaling factors are listed in Table 4.6. The result of mutual χ^2 minimization is shown in the bottom right figure of Figure 4.32.

Differences between calibration-based and consistency-based scaling factors are considered as an additional source of uncertainty. These fractional differences were added to the calibration systematic uncertainties in Table 3.14 in quadrature. The total fractional uncertainty for each dataset was set to 15.5% for SK-IV, 9.6% for SK-V, 5.9% for SK-VI, and 12.4% for SK-VI (R).

[1]This issue for SK-IV is speculated to be coming from using different versions of the simulator between Am/Be simulation and atmospheric neutrino simulation, which is currently under investigation.

Efficiency scaling	SK-IV	SK-V	SK-VI	SK-VI (R)	Reduced χ^2
No scaling	1	1	1	1	0.8843
Calibration-based	1.1000	0.9784	0.9086	0.8770	1.3747
Consistency-based	0.9534	1.0387	0.9452	0.8770 (fixed)	0.7610

Table 4.6: The signal efficiency scaling factors used to plot Figure 4.32, and the resulting reduced sum of mutual χ^2.

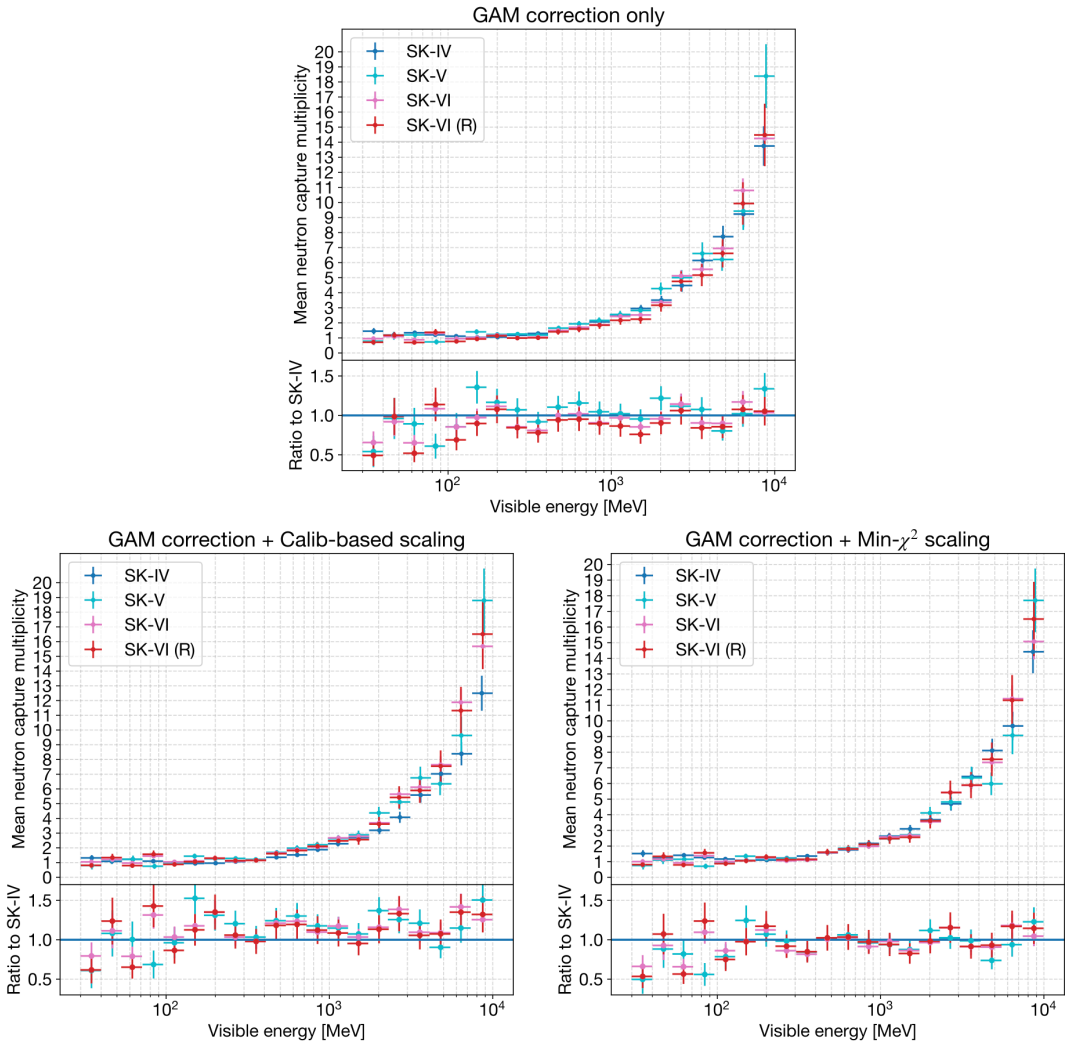


Figure 4.32: Comparison of average neutron capture multiplicity across datasets, using GAM correction only (top), with additional calibration-based scaling (bottom left), and consistency-based scaling (bottom right).

(2) GAM regression

We assigned uncertainty based on the difference between the actual signal multiplicity and the signal multiplicity estimated through the fitted GAM, shown in Figure 4.31. This assignment was carried out on a bin-by-bin basis, as bins at lower and higher energy ranges exhibit a larger bias compared to mid-range bins.

(3) Neutrino event reconstruction

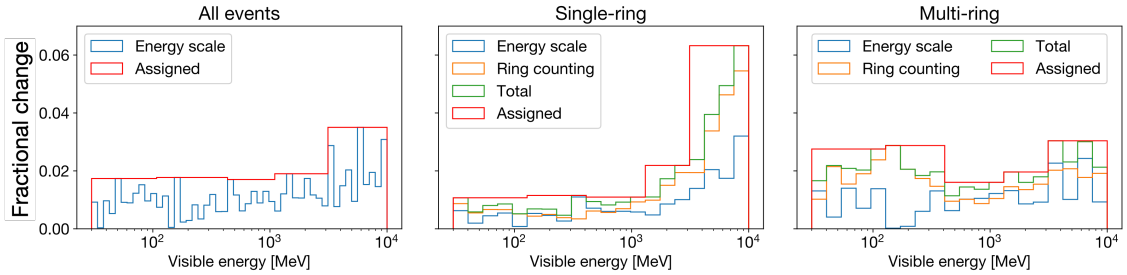


Figure 4.33: Assigned event reconstruction uncertainties in each subsample.

The cosmic ray calibration results up to SK-V indicate an energy scale error of less than 2% [48]. To account for this uncertainty, we shifted the visible energy of all simulated neutrino events by ±2%, and the resulting fractional changes in bin values were designated as the uncertainty in each bin due to the energy scale.

Additionally, single vs. multi-ring classification uncertainty was estimated. Assuming a 5% uncertainty in the multi-GeV e-like event rate, derived from a past oscillation analysis [116], we propagated the corresponding ±0.5 uncertainty in the ring counting likelihood near the cutoff at zero. The fractional changes in bin values were assigned on a bin-by-bin basis.

Uncertainties arising from other reconstructed variables, such as ring PID likelihood and non-neutrino background events, were estimated to be less than 0.1% and thus were omitted in this analysis.

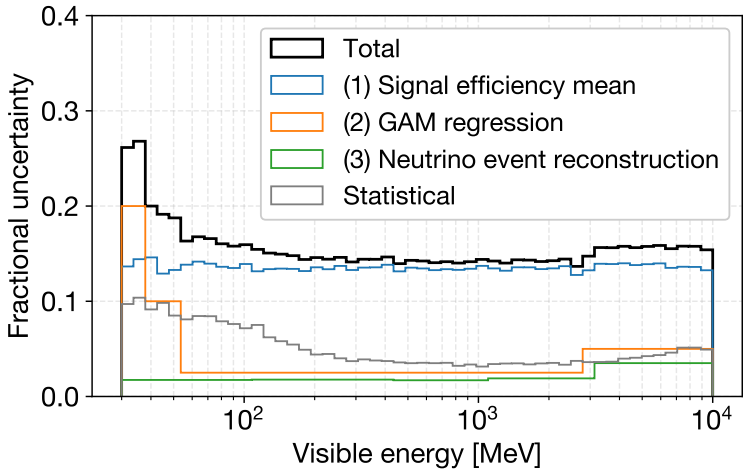


Figure 4.34: Bin-by-bin assigned systematic uncertainties to all final sample events.

With the exception of extremely low-energy events, the most prominent source of uncertainty is the uncertainty in the signal efficiency mean, approximately $\sim 15\%$.

Figure 4.35 displays the mean neutron capture multiplicity distributions of combined data with systematic errors. A clear linear relationship between the mean neutron multiplicity and the event visible energy is evident.

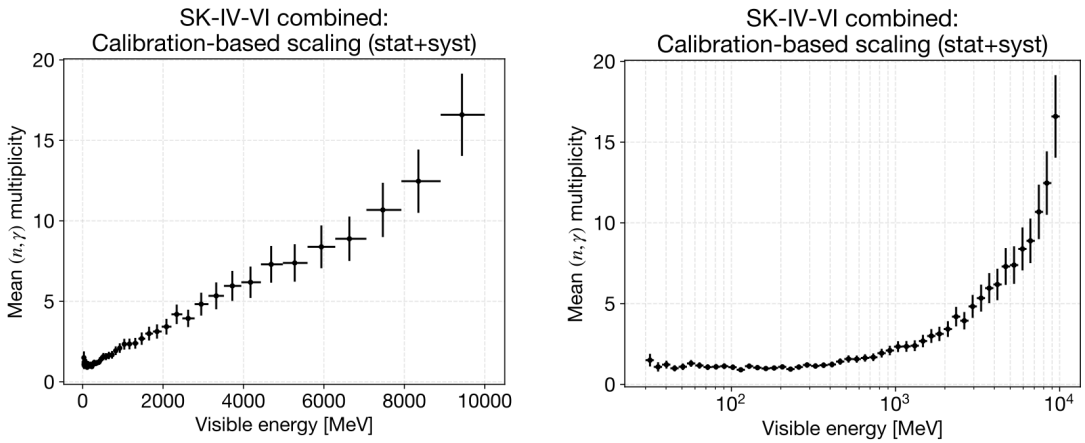


Figure 4.35: Mean neutron capture multiplicity distributions from combined data, plotted against event visible energy. The distributions are shown with both linear (left) and logarithmic (right) horizontal axis scales.

Figure 4.36 presents a comparison between the combined data, the default MC simulation, and the pure-water estimation from the SNO collaboration in 2019 [33]. While the agreement with the SNO estimation is reasonably good, a noticeable deficit is observed across the 0.1-10 GeV range when compared to the default MC.

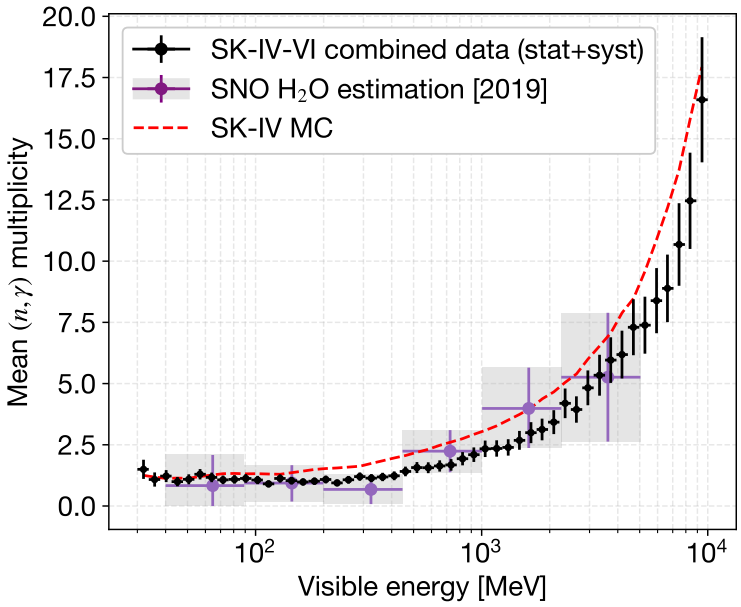


Figure 4.36: Comparison of SK data with default MC and SNO results [33].

Table 4.7 provides an summary of the average neutron capture multiplicity observed in atmospheric neutrino events with visible energy deposits ranging from 30 MeV to 10 GeV in the SK detector, including systematic and statistical uncertainties. The results from SK-V, VI, and VI (R) exhibit consistency, with a slight tension observed with SK-IV and MC results by approximately 1σ.

	Mean	Syst	Stat
SK-IV	2.0385	0.3421	0.0137
SK-V	2.3331	0.2714	0.0462
SK-VI	2.3597	0.2080	0.0223
SK-VI (R)	2.3016	0.3229	0.0287
SK-IV + V + VI	2.1134	0.2967	0.0119
MC	2.6117	-	-

Table 4.7: Observed mean (n,γ) multiplicity and uncertainties in each dataset.

4.4.2 Linear regression on visible energy

We applied the same linear fit to the full dataset and the two subsamples, single-ring and multi-ring datasets. The fitted functions are shown in Figures 4.37 through 4.39, and the fitted slopes and y-intercepts are listed in Table 4.8 and illustrated in Figure 4.40. The fitted slopes and intercepts among different SK phases and the detection algorithms (`INIT` and `RECO` vertex modes for SK-VI) showed good consistency.

Across all subsamples, we observed reduced neutron production compared to our predictions. Comparing single-ring and multi-ring subsamples, neutron production in multi-ring events was nearly twice as much as in single-ring events, which is due to multi-ring events are dominantly inelastic events with a significant amount of energy transferred to hadron production, as predicted by our simulation (refer to Figure 2.44).

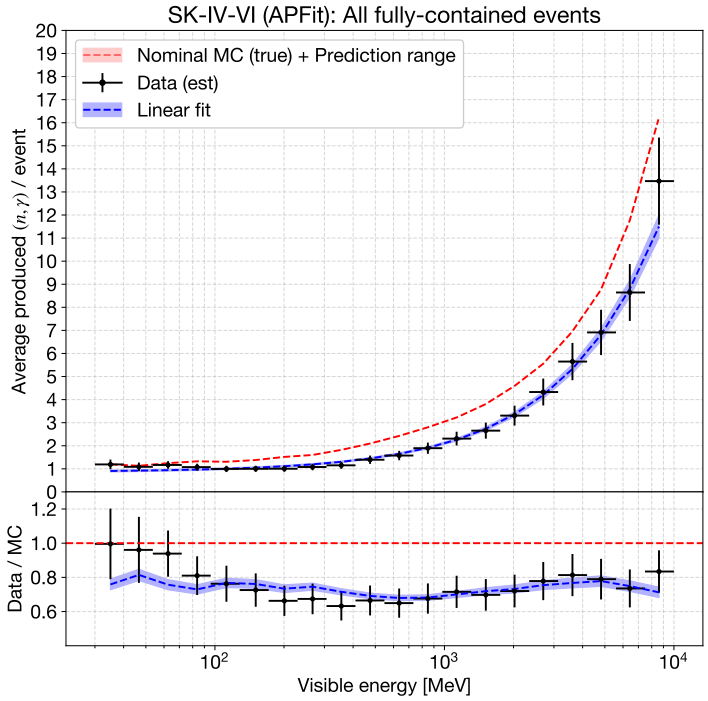


Figure 4.37: Average neutron capture multiplicity (black) distribution with a fitted linear function (blue).

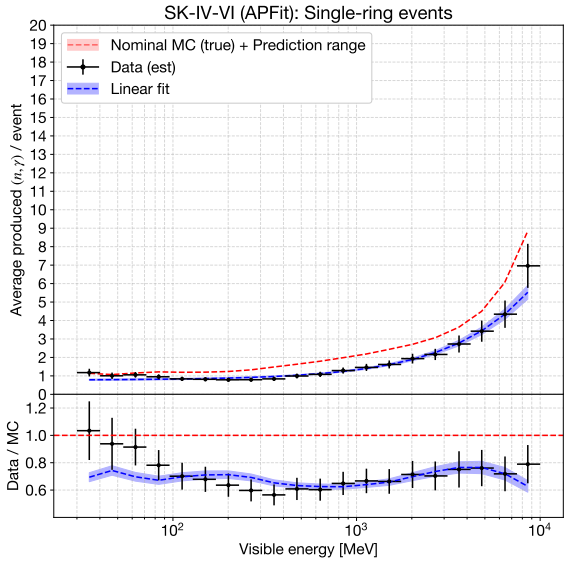


Figure 4.38: Average neutron capture multiplicity (black) distribution in single-ring events with a fitted linear function (blue).

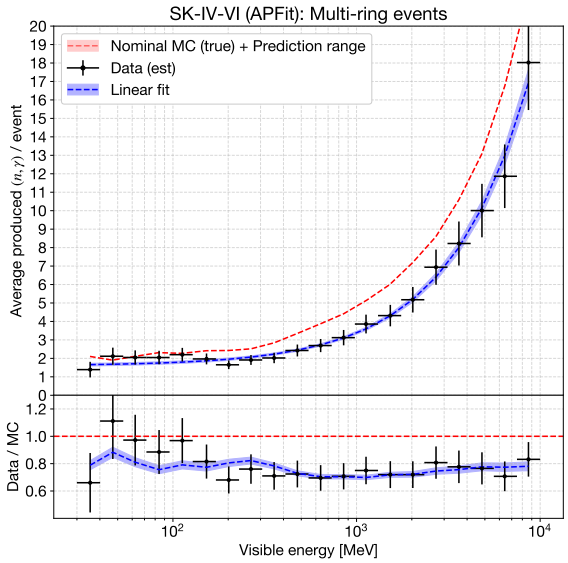


Figure 4.39: Average neutron capture multiplicity (black) distribution in multi-ring events with a fitted linear function (blue).

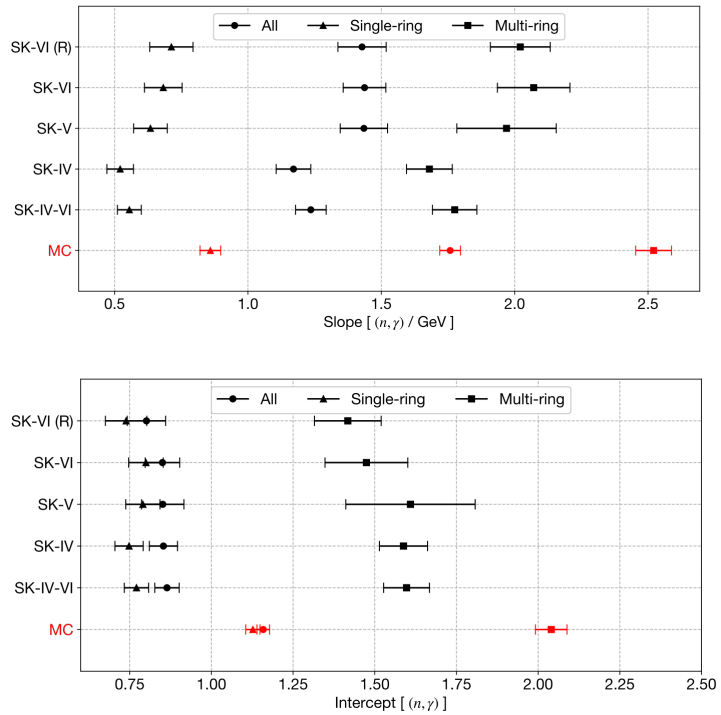


Figure 4.40: A visual summary of fitted slopes and intercepts. (stat + syst)

All events	Slope $[(n,\gamma)$ / GeV$]$	Intercept $[(n,\gamma)]$	Reduced χ^2
Data	1.2359 ± 0.0574	0.8642 ± 0.0373	0.51
MC	1.7581 ± 0.0389	1.1587 ± 0.0192	1.32
Data/MC	0.7029 ± 0.0362	0.7458 ± 0.0345	-

Single-ring	Slope $[(n,\gamma)$ / GeV$]$	Intercept $[(n,\gamma)]$	Reduced χ^2
Data	0.5546 ± 0.0447	0.7708 ± 0.0373	0.78
MC	0.8591 ± 0.0388	1.1270 ± 0.0218	0.80
Data/MC	0.6465 ± 0.0597	0.6839 ± 0.0356	-

Multi-ring	Slope $[(n,\gamma)$ / GeV$]$	Intercept $[(n,\gamma)]$	Reduced χ^2
Data	1.7752 ± 0.0829	1.5973 ± 0.0701	0.44
MC	2.5209 ± 0.0672	2.0401 ± 0.0483	1.18
Data/MC	0.7042 ± 0.0380	0.7830 ± 0.0391	-

Table 4.8: Fitted slopes and intercepts of combined data and MC. (stat + syst)

Additionally, we performed the fit on single-ring e-like and μ-like subsamples, with the reconstructed lepton momentum as the horizontal axis, as displayed in Figure 4.41. Comparison between single-ring e-like and μ-like subsamples reveals that μ-like events exhibit a larger intercept. This observation is attributed to low-energy muons captured by ^{16}O before decaying into electrons, especially within the proximity of the neutrino interaction point.

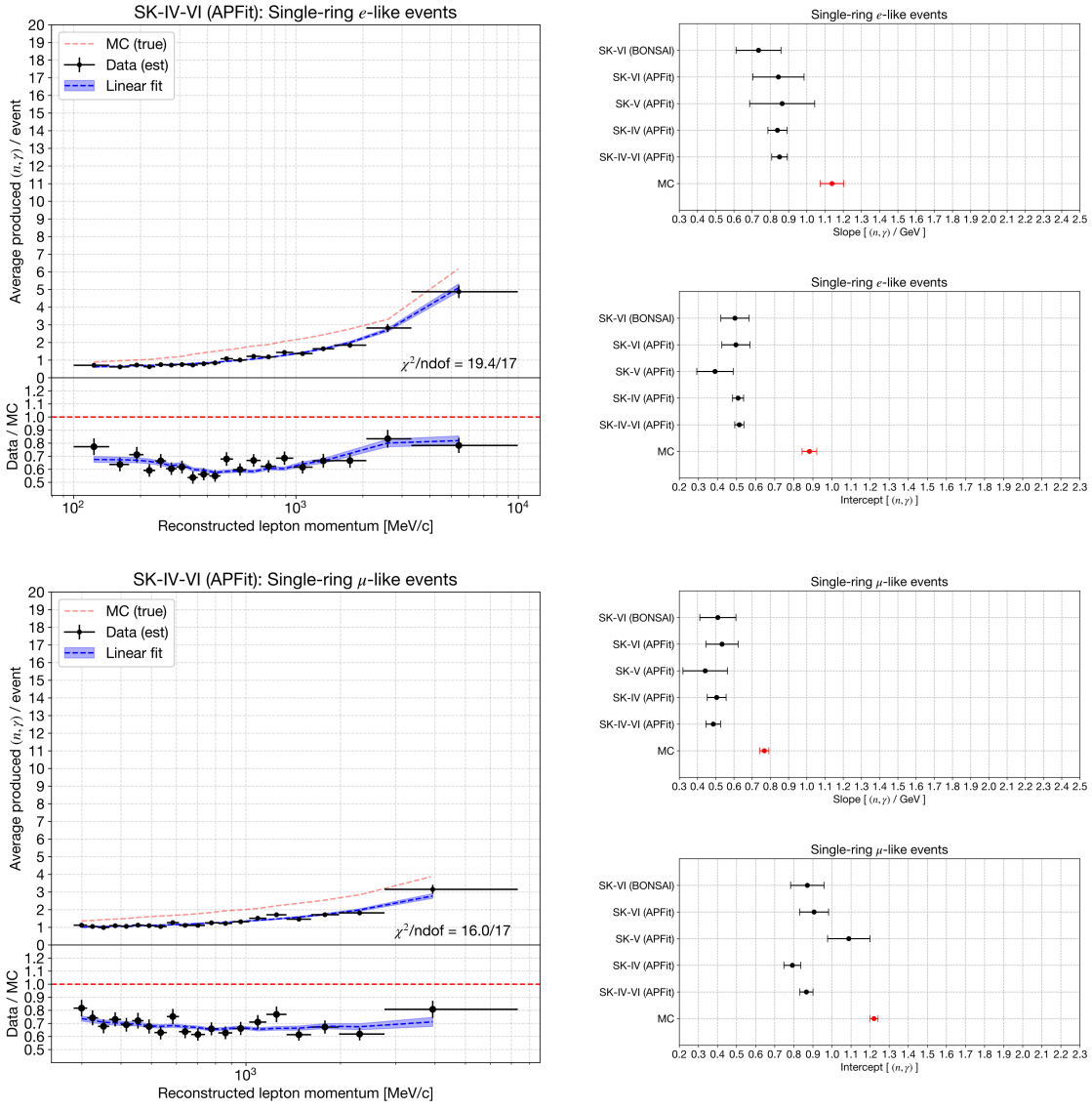


Figure 4.41: Average neutron capture multiplicity observed in single-ring e-like (top) and μ-like (bottom) events, as a function of reconstructed lepton momentum.

4.4.3 Various hadronic interaction models

In neutrino event simulation, discussed in Sections 2.3.2 and 2.3.3, and illustrated in Figure 2.46, the process is divided into two main components. The neutrino event generator models outgoing particles from the nucleus, while the detector simulator handles the propagation of these particles and the subsequent detector response.

Observed neutrons can originate from direct recoil in neutrino CCQE scattering, hadronic re-interactions within the target nucleus (final-state interactions, FSI), or hadronic showers of energetic outgoing nucleons within water (secondary interactions, SI). For neutrino reconstruction, primary interaction (PI) neutrons are of primary interest, while, for background rejection using neutrons, all three sources are significant. Neutrino event generators handle PI and FSI, while the detector simulator manages SI.

The contribution of these components varies with interaction kinematics. The number of neutrons from PI and FSI is relatively independent of transferred energy, especially dominant in low-energy events where outgoing hadrons lack the energy to knock out multiple neutrons. Therefore, the intercept measured in Section 4.4.2 is particularly relevant to PI and FSI.

On the other hand, the number of neutrons from SI is highly dependent on transferred energy, as energetic hadrons can produce more hadrons through showers until they fully deposit their energy in the large detector. This quantity of neutrons can serve as a calorimetric measure of the hadronic invariant mass, as demonstrated in Section 4.4.2. The expected proportion of PI/FSI and SI in total observed signals is illustrated in Figure 4.42.

There exist several models in the market that address each part of the simulation, and a significant prediction variability was observed in the choice of the neutrino event generator in an earlier study [34]. In this study, we particularly focus on the FSI and SI parts, where uncertainties are not quantified, and differences in neutron production are not well understood.

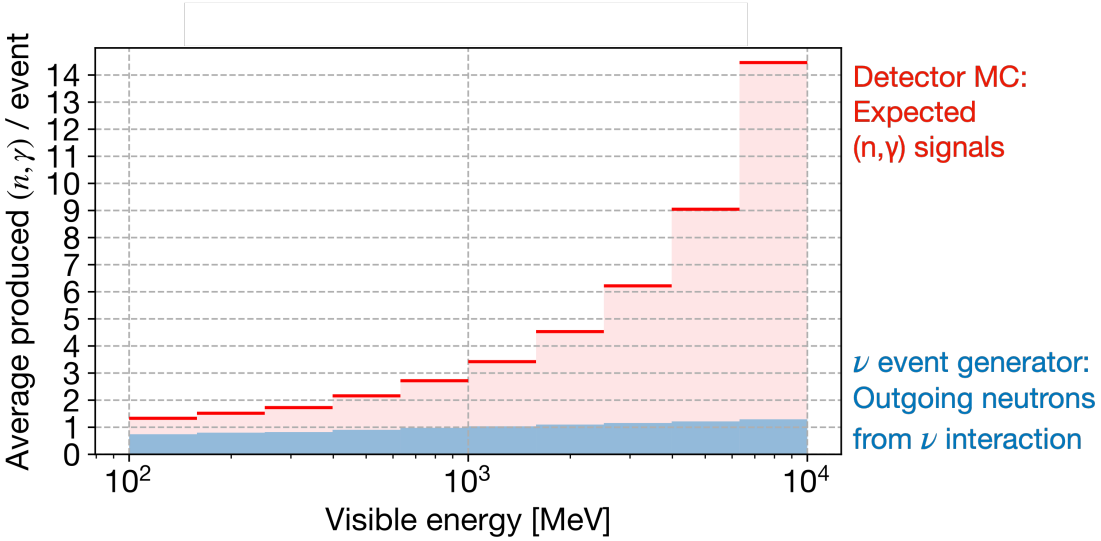


Figure 4.42: Mean neutron capture multiplicity plotted against visible energy, extracted from default simulation models (`NEUT` 5.4.0 for PI/FSI and SK detector simulator `skdetsim` in Geant3.21 with SK-IV default model configuration for SI). The blue histogram displays the number of neutrons outgoing as a result of PI and FSI simulation, while the red bars represent the average of total neutron captures per bin, including all three contributions: PI, FSI, and SI.

In this analysis, we examined six different neutrino event generator options, each with varying neutrino interaction and FSI models, along with five sets of SI models implemented in our detector simulator. In total, we compared predictions for the average neutron multiplicity across 30 different combinations of neutrino event generator and detector simulator options.

The atmospheric neutrino flux model remained fixed; specifically, the HKKM 2011 atmospheric neutrino flux model [55] was used for all model combinations.

Neutrino event generator options

- NEUT 5.4.0

  This neutrino event generator was utilized in our default atmospheric neutrino event simulation. Details on the modeling of neutrino interaction and final state interactions within NEUT 5.4.0 are provided in Section 2.3.2.

- NEUT 5.6.3

  This version of NEUT includes a bug fix in nucleon binding energy, resulting in a deeper nuclear potential that reduces low-energy cross sections.

- GENIE 3.4.0 (specific to G_18a_10x_02_11b tune, x∈{a,b,c,d})

  GENIE [157] stands out as a widely used neutrino event generator beyond SK and T2K. This version of GENIE has incorporated various FSI model options, enabling a comparison of FSI models. The neutrino interaction modeling is largely similar to the aforementioned NEUT versions, including the modeling of the nucleus ground state using the Local Fermi Gas (LFG) model and almost identical treatment of QE and 2p2h interactions. Differences lie in resonant pion production as well as inelastic scattering involving hadronization.

  GENIE also provides a wealth of FSI models so that comparing various FSI options within GENIE allows us to disentangle FSI from the PI/FSI convolution in the outcome.

The differences among the primary neutrino event models are summarized in Table 4.9. Figure 4.43 provides a comparison of the fraction of simulated atmospheric neutrino interactions in water, for each neutrino event generator. NEUT 5.4.0 predicts a slightly higher fraction of CCQE events, attributed to a bug in nuclear binding energy calculation. This is also evident in Figures 4.44 and 4.45, showing a more pronounced low-energy peak below 200 MeV with NEUT 5.4.0. On the other hand, NEUT 5.6.3 and GENIE 3.4.0 exhibit a similar CCQE fraction. NEUT 5.6.3 predicts a slightly larger fraction of CC inelastic interactions, while GENIE predicts a larger fraction of NC interactions. Generally, there is reasonable agreement among the three primary neutrino event models.

		NEUT 5.4.0	NEUT 5.6.3	GENIE 3.4.0
Nucleus ground state		LFG	LFG	LFG
Cross sections	**QE**	Nieves (2013)	+ Binding E correction	Nieves (2004)
	2p2h		Nieves (2013)	Nieves (2012)
	RES	Berger-Sehgal (+Graczyk-Sobczyk)		Berger-Sehgal
	COH	Berger-Sehgal		Berger-Sehgal
	SIS/DIS	Custom/ Pythia/ GRV98LO (+Bodek-Yang)		Pythia/ GRV98LO (+Bodek-Yang)

Table 4.9: Summary of neutrino interaction model differences.

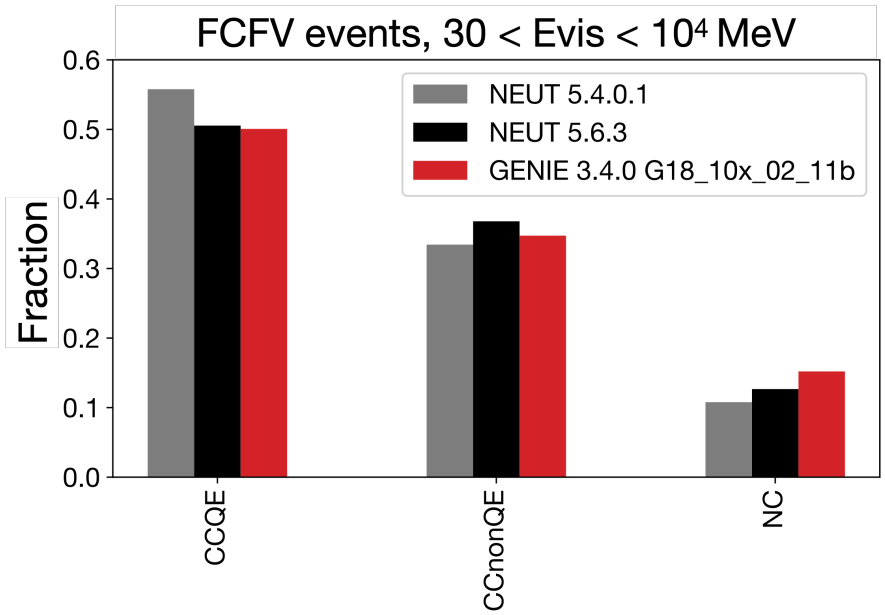


Figure 4.43: Comparison of the predicted fractions of atmospheric neutrino interactions in water among various neutrino event generators.

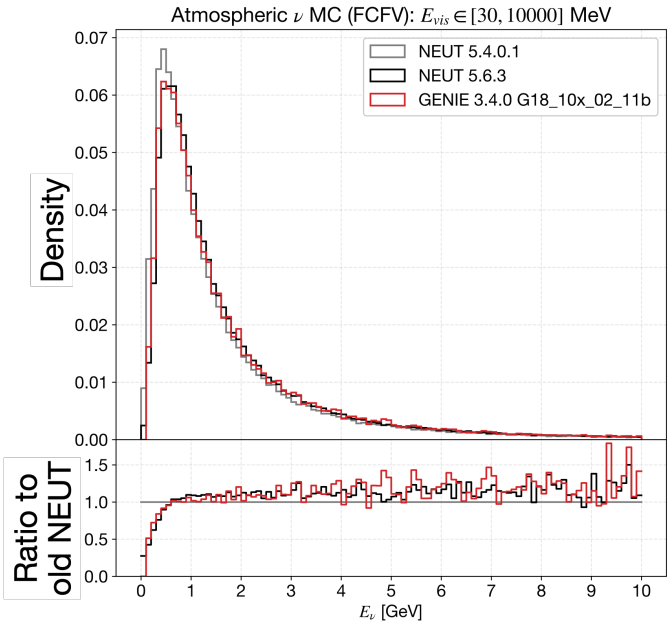


Figure 4.44: True neutrino energy distributions in fully contained atmospheric neutrino events in water, simulated using different neutrino event generators.

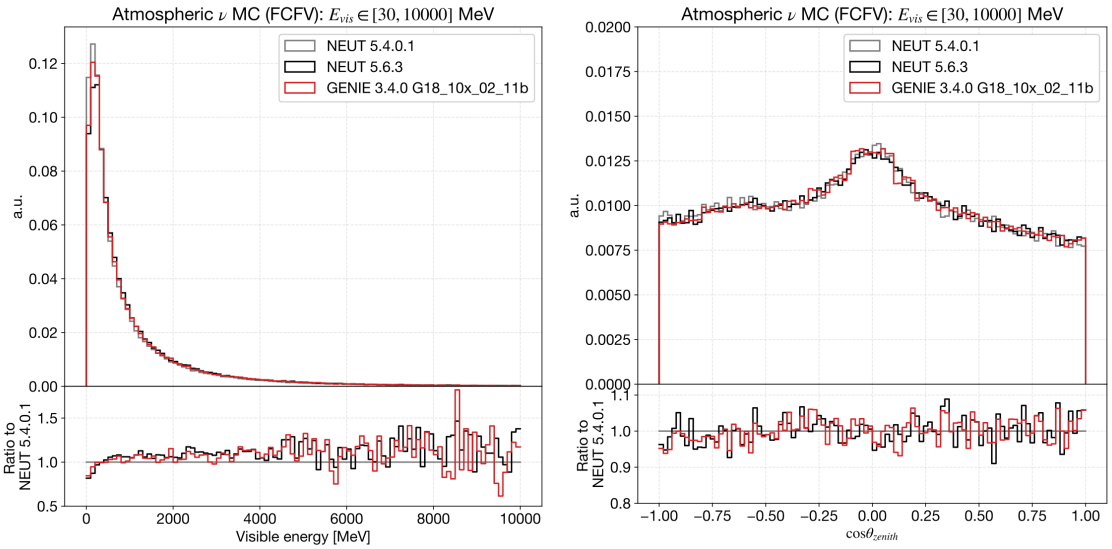


Figure 4.45: Comparison of the distributions of reconstructed neutrino visible energy (left) and lepton zenith angle (right) among different neutrino event generators.

The following are descriptions of the FSI models available with GENIE 3.4.0:

- INTRANUKE/hA 2018 [158]

  This model, while not a cascade model, effectively calculates outgoing secondaries using pion/nucleon-to-nucleus cross-section data up to 1.2 GeV. Above 1.2 GeV, cross section data points are extrapolated.

- INTRANUKE/hN 2018 [158]

  This is a space-like intranuclear cascade (INC) model (see Section 2.3.3 for description) capable of working on most hadrons. This model was validated against external data.

- Geant4 Bertini Cascade model [159]

  Based on the Geant4 adaptation of Bertini [95], this is also an INC model that relies on CERN-HERA compilations of hadron–nucleon interaction data [160]. Associated cross section errors range from 10-30%. The model includes a native evaporation and break-up model.

- INCL [161]

  This cascade model stands out from those mentioned earlier in a few ways. Firstly, it's a time-like INC model. All participants (incident projectile and target nucleons) and spectators (non-target nucleons) are treated as point particles and tracked in both space and time. Additionally, each nucleon is situated in a potential well whose depth depends on its momentum and position. Collisions occur when two particles approach within a certain distance, determined by parametrized cross sections specific to the type of hadron and collision. Pauli blocking of a collision with a nucleon having less than Fermi energy is determined based on the probabilistic occupancy of energy states in the Fermi sphere. Tracking stops after a predetermined time, and the residual nucleus at that point is passed to the de-excitation routine. In GENIE, it is coupled with ABLA07 [162], known to provide a good explanation for spallation data when combined with INCL [163]. The model has undergone validation across various observables in a wide energy range [164].

The differences among the aforementioned FSI models, including the one native to `NEUT`, are summarized in Table 4.10. Figure 4.46 shows an example of FSI model predictions (`G4Bertini` not shown) in the estimated p-^{12}C reaction cross section as a function of proton kinetic energy. While the agreement above 150 MeV is good, at lower energies the model predictions diverge.

Item / Model	`NEUT`	`GENIE`			
		`hA`	`hN`	`G4Bertini`	`INCL`
Nucleus model	Continuum				Particles
Collision	Mean free path				Distance
Propagation	Space				Time
Cross section	Tabulated			Parametrized	
Stopping criterion	Nucleus radius				Time
Pauli blocking	Strict	None		Strict	Probabilistic
Medium effect	Yes (π only [165])	Yes		None	Yes
Formation zone	Yes (π only [166])	None			
Nuclear de-excitation	Yes (^{16}O only)	None		Yes [167]	Yes [162]

Table 4.10: Summary of FSI model differences in neutrino event generators.

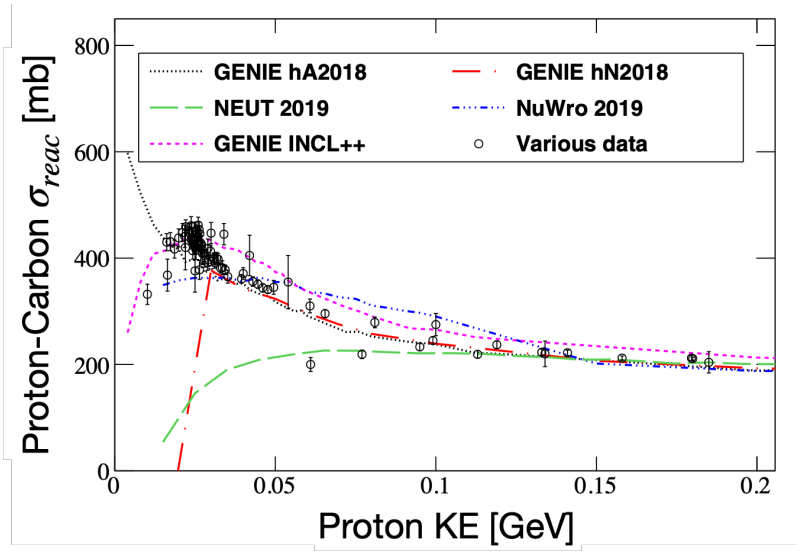


Figure 4.46: Comparison of p-^{12}C reaction cross section predictions as a function of proton kinetic energy. (Reprinted from [168])

Figure 4.47 illustrates the predicted outgoing neutron momentum distributions for atmospheric neutrino interactions in water, using the previously mentioned neutrino event generator options. Notably, the proton and neutron multiplicities, as well as the kinematics at low energies, showed significant variations among the different model options. Another factor contributing to these differences, in addition to the scattering models themselves, was the distinct nuclear de-excitation (pre-equilibrium/compound) routines that come into play after all secondary particles were tracked in the FSI models. These de-excitation models are responsible for the emission of very low-energy nucleons.

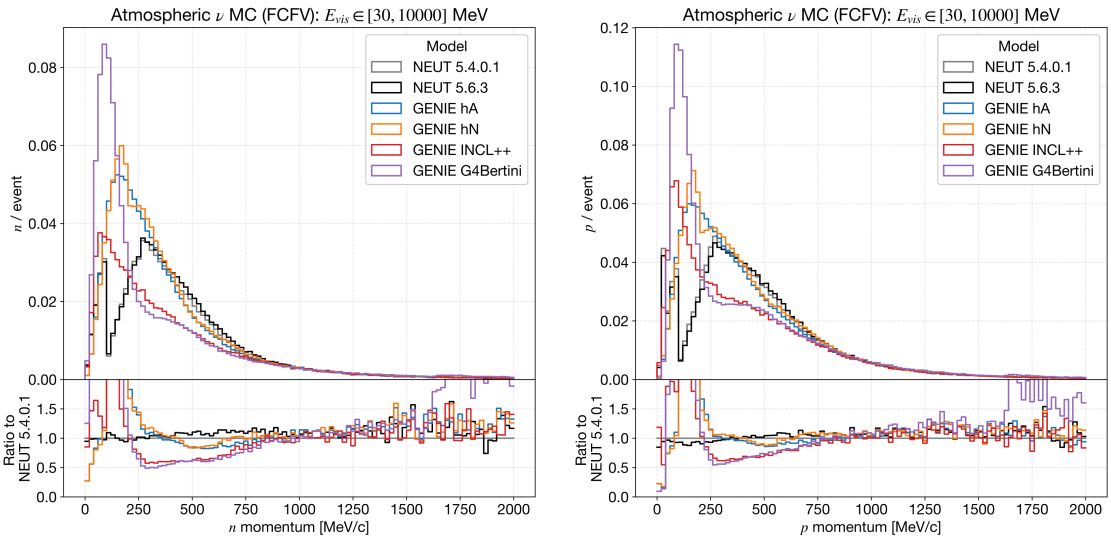


Figure 4.47: Comparison of FSI model prediction of outgoing neutron (left) and proton (right) momentum distributions for atmospheric neutrino interactions in water.

Our detector simulators, `skdetsim` in Geant3.21 and `SKG4` in Geant4.10.5.p01, offer a range of hadronic secondary interaction models:

Detector simulator options

- `GCALOR`

  This model is available in `skdetsim`. For propagation of hadron with kinetic energy below 10 GeV, `GCALOR` uses `HETC` [108] implemented based on Bertini [95]. For thermalization and captures of neutrons with kinetic energy lower than 20 MeV, `GCALOR` uses `MICAP` which estimates outgoing particle multiplicity and kinematics based on the evaluated nuclear data ENDF/B-V.

- SK-IV/V default

  This model is set as the default in `skdetsim` for simulating the pure water phase. In the default settings of SK-IV/V, pions with momentum lower than 500 MeV/c (which is relevant to the Δ resonance) are processed using the `NEUT` pion FSI model. This model incorporates medium corrections, as outlined in [165], with parameters fine-tuned to pion-nucleus scattering data [100].

- SK-VI default

  This model is set as the default in `skdetsim` for the Gd-loaded phase simulation. In addition to the SK-IV/V default, NeutronHP package based on ENDF/B-VII.1 [31] from Geant4.10.5.p01 is used to propagate neutrons below 20 MeV kinetic energy.

- Geant4 Bertini cascade model

  This model is available in `SKG4`. This is basically the same as Geant4 Bertini cascade model implemented in `GENIE`.

- Geant4 `INCL`

  This model is available in `SKG4`. In Geant4.10.5.p01, it is coupled to the default Geant4 precompound model [169] rather than `ABLA` [162] as in `GENIE`.

To determine the average neutron capture multiplicity per outgoing hadron produced via neutrino-nucleus interaction within the ID water volume, we conducted simulations using our detector simulators, `skdetsim` and `SKG4`. The simulations employed a particle-gun setup, wherein 10,000 single nucleons (hadrons) were randomly fired at the tank center, covering a kinetic energy range of 0-10 GeV. Subsequently, for each interaction model and momentum bin, we calculated the average number of neutron captures.

Figure 4.48 shows a comparison of the predictions made by different secondary interaction models for the average neutron capture multiplicity in water, plotted against the incident nucleon momentum. Model variations were observed in the range of approximately 10% to 50%, with the `INCL` model notably predicting 20-30% fewer neutron captures than the other models based on the Bertini cascade model.

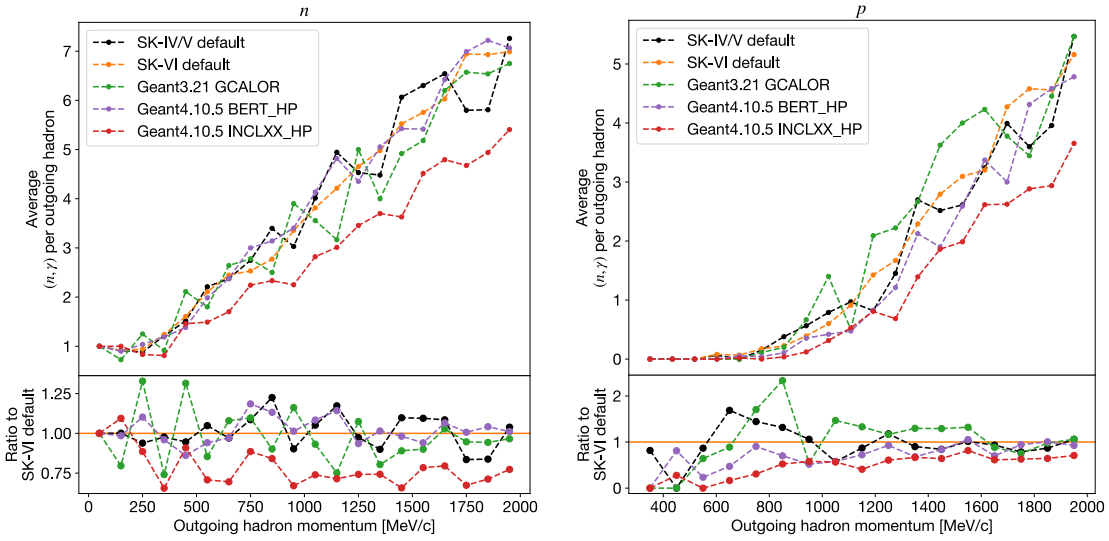


Figure 4.48: Comparison of predictions from different secondary interaction models for the average neutron capture multiplicity as a function of incoming neutron (left) and proton momentum (right) within a water volume.

4.4.4 Generation of model predictions

Conducting a full detector simulation for all 30 model combinations is computationally expensive, and it is unnecessary for obtaining neutron capture counts. Instead, we utilize tabulated average neutron capture multiplicity per incident hadron as visually represented in Figure 4.48. We assign a weight to each outgoing particle $\in n, p, \pi^{\pm}, \mu^{-}$ from the neutrino event generator, based on the tabulated average neutron capture multiplicity to calculate the total expected signals for each event.

This approach is justified for two main reasons: firstly, the "average" neutron capture multiplicity from secondary interactions should primarily depend on the outgoing particle multiplicity and kinematics, independent of the neutrino interaction itself. Secondly, most secondary neutrons are produced by inelastic interactions or nuclear captures of one of $\{n, p, \pi^{\pm}, \mu^{-}\}$, as illustrated in Figure 4.49. However, it is important to note that for higher energy events, kaon inelastic scattering becomes non-negligible, reaching the order of $O(1)\%$.

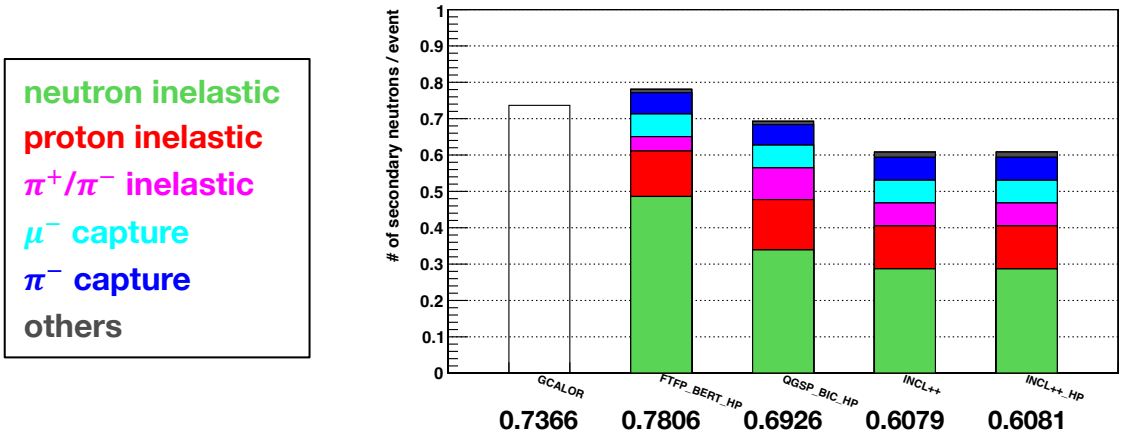


Figure 4.49: Breakdown of neutron production sources in atmospheric neutrino interactions (below 2 GeV) in water, categorized by interactions and analyzed using five distinct hadronic inelastic scattering models. The stacked bars are color-coded to represent the specific neutron production source interaction. (Note: `GCALOR` does not offer a breakdown by interactions.) The category labeled as "Others" includes interactions such as kaon inelastic interactions, with a larger fraction in higher energy events. (Reprinted from [170])

Furthermore, we conducted full detector simulations exclusively using the default secondary interaction model for each neutrino event generator option. The purpose was to obtain the reconstructed visible energy, which serves as the horizontal axis in our measurements. The reconstructed visible energy for every event, spanning all secondary interaction model choices, was derived from this comprehensive detector simulation using the default secondary interaction model.

In this context, we made the assumption that reconstructed visible energy is independent of the secondary interaction model. This assumption holds true to some extent since the reconstructed visible energy is primarily influenced by charged leptons resulting from neutrino interactions, and the variability in pion production among different SI models is rather small (see Appendix E). Impact of large variability in nucleon production of SI models on visible energy is expected to be negligible as nucleons are mostly invisible in the detector.

Figure 4.50 presents a comparison of predictions for the average neutron capture multiplicity across all 30 model combinations as a function of event visible energy. The variability in model predictions hovers around 30-40%, particularly noticeable for higher event visible energy. Notably, a clear distinction exists between model combinations featuring `INCL` as the secondary interaction model and those with alternative secondary interaction models. As observed in Figure 4.48, `INCL` resulted in fewer neutron captures, translating to a 30-40% reduction in signal multiplicity compared to other secondary interaction models such as the Bertini cascade model.

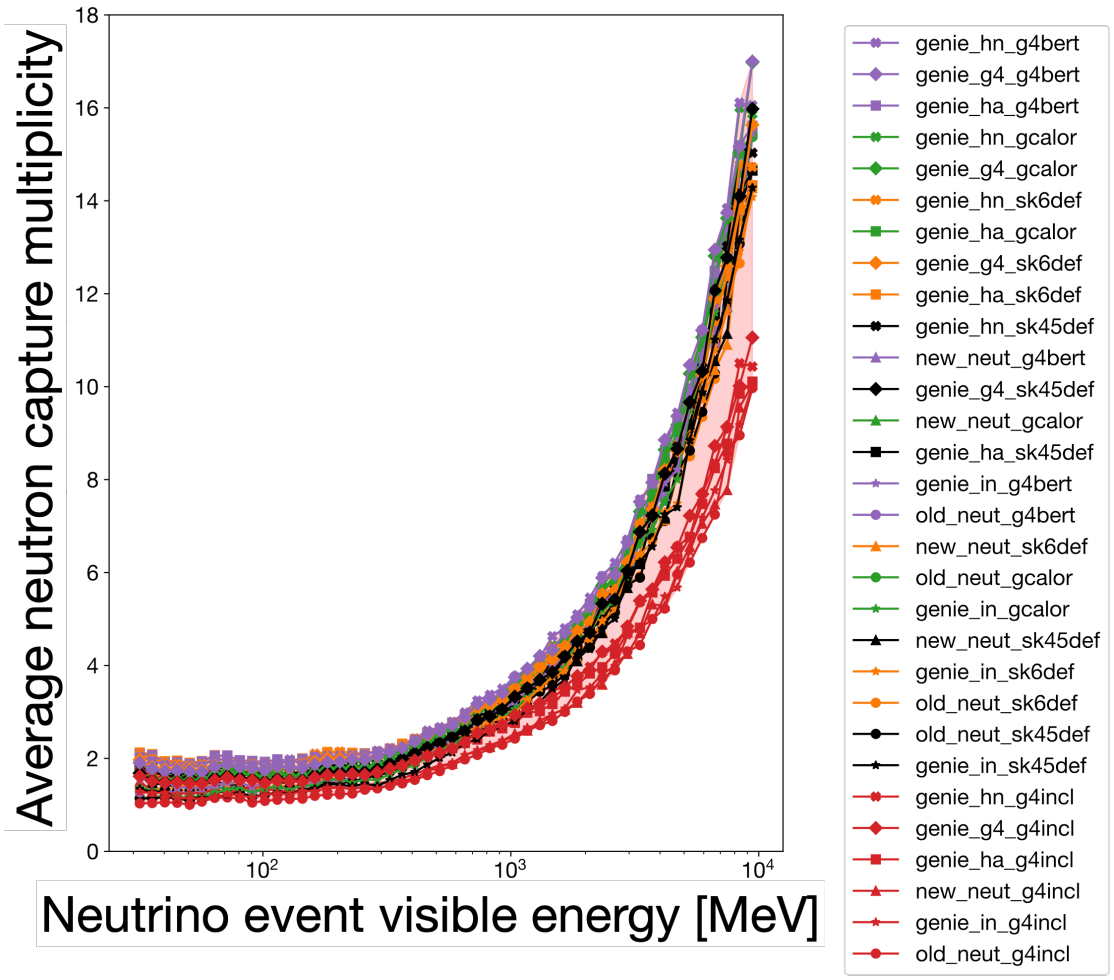


Figure 4.50: Comparison of the average neutron capture multiplicity predictions for each model combination. Contributions from neutron, proton, and pion interactions in water, along with negative muon captures on ^{16}O, were taken into account.

In Figure 4.51, we break down the overall model prediction variability into two sources: neutrino event generation (PI/FSI) and detector simulation (SI). When comparing model choices within each origin while keeping the choice fixed for the other origin, we observe that PI/FSI model prediction variability is pronounced at lower energy events, while SI model prediction variability becomes prominent at higher energy events. This aligns with expectations, given that FSI model prediction variability is associated with low-energy nucleon kinematics and multiplicity, whereas higher energy nucleons play a more substantial role in SI model predictions. Additionally, SI occurs more frequently at higher energies, obscuring the original hadron content more strongly in higher energy events. The default neutrino event generator option (NEUT 5.4.0) showed the smallest neutron production, while the default detector simulator option (SK-VI default) showed relatively larger neutron production.

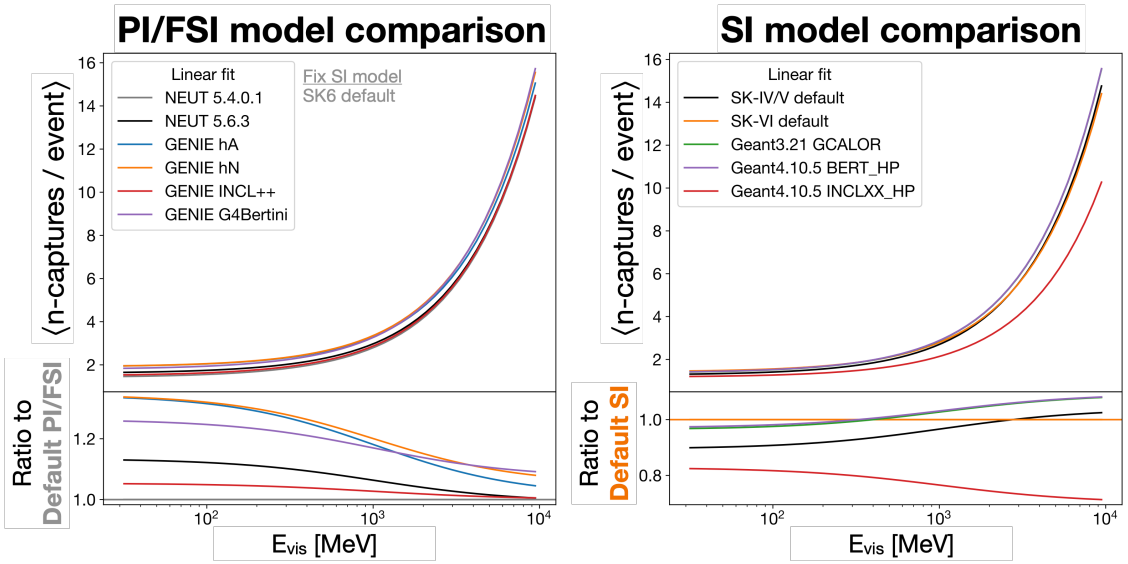


Figure 4.51: Comparison of the impact of different neutrino event generator options (left) and detector simulator options (right) on the predicted average neutron capture multiplicity. In the left figure, each neutrino event generator option is paired with the nominal secondary interaction model (SK-VI default). In the right figure, each detector simulator option is paired with the nominal neutrino event generator option (NEUT 5.4.0).

It is important to note that the derived model predictions take into account tabulated contributions from a specific set of particle interactions only, omitting a small contribution from kaon interactions, which contribute around $O(1)\%$ for higher energy events as mentioned earlier. Instead of directly comparing the derived model predictions presented in Figures 4.50 and 4.51 with the data outlined in Section 4.4.2, we calculated the ratio between each model combination prediction and the prediction from the default model (`NEUT 5.4.0` and `SK-VI default` options). We then multiplied this ratio by the prediction based on full detector MC simulation with the default model combination, to compensate for the absence of kaon contributions in the model predictions at higher energies.

4.4.5 Comparison of model predictions with data

Figure 4.52 shows the comparison of the observed average neutron capture multiplicity with the prediction range, spanning the minimal to maximal predictions for each visible energy bin. We observe agreement between the data and the predictions of the nominal simulation in the low-energy range below 0.1 GeV and the higher energy range above 1 GeV. However, in the intermediate energy range of $0.1 < E < 1$ GeV, the observed average neutron capture multiplicity is slightly smaller than our most conservative prediction (`NEUT` 5.4.0 and `INCL`). This particular combination yielded the best agreement in the fitted slope and intercept (Figure 4.53, Table 4.11). Figure 4.54 provides a visual comparison of the data with the linear model, assuming fitted slope and intercept from the best model combination.

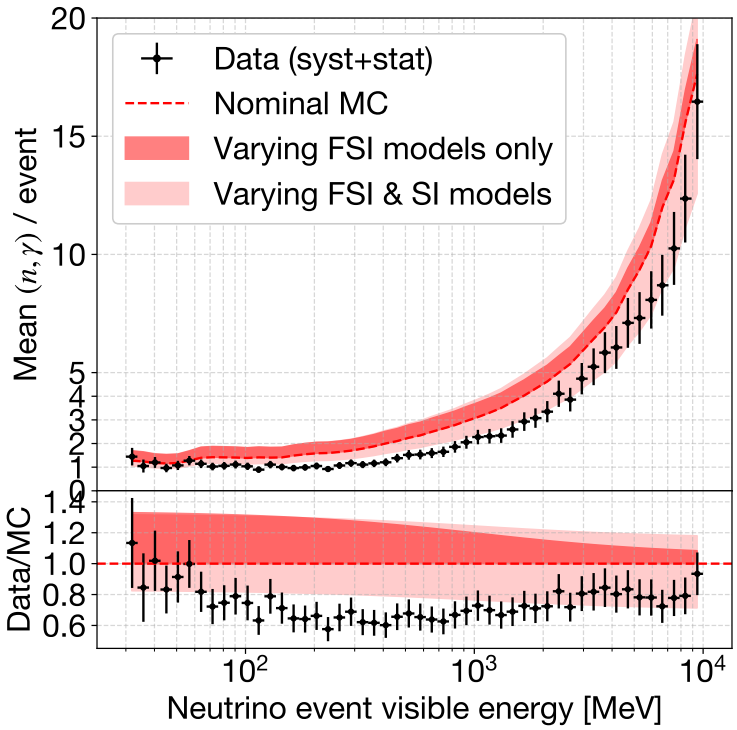


Figure 4.52: Comparison of predicted average neutron capture multiplicities for different model combinations. The nominal MC results shown in red dashed line is from the full detector MC simulation of SK-VI atmospheric neutrino events (corresponding to `NEUT` 5.4.0 with `SK-VI default`).

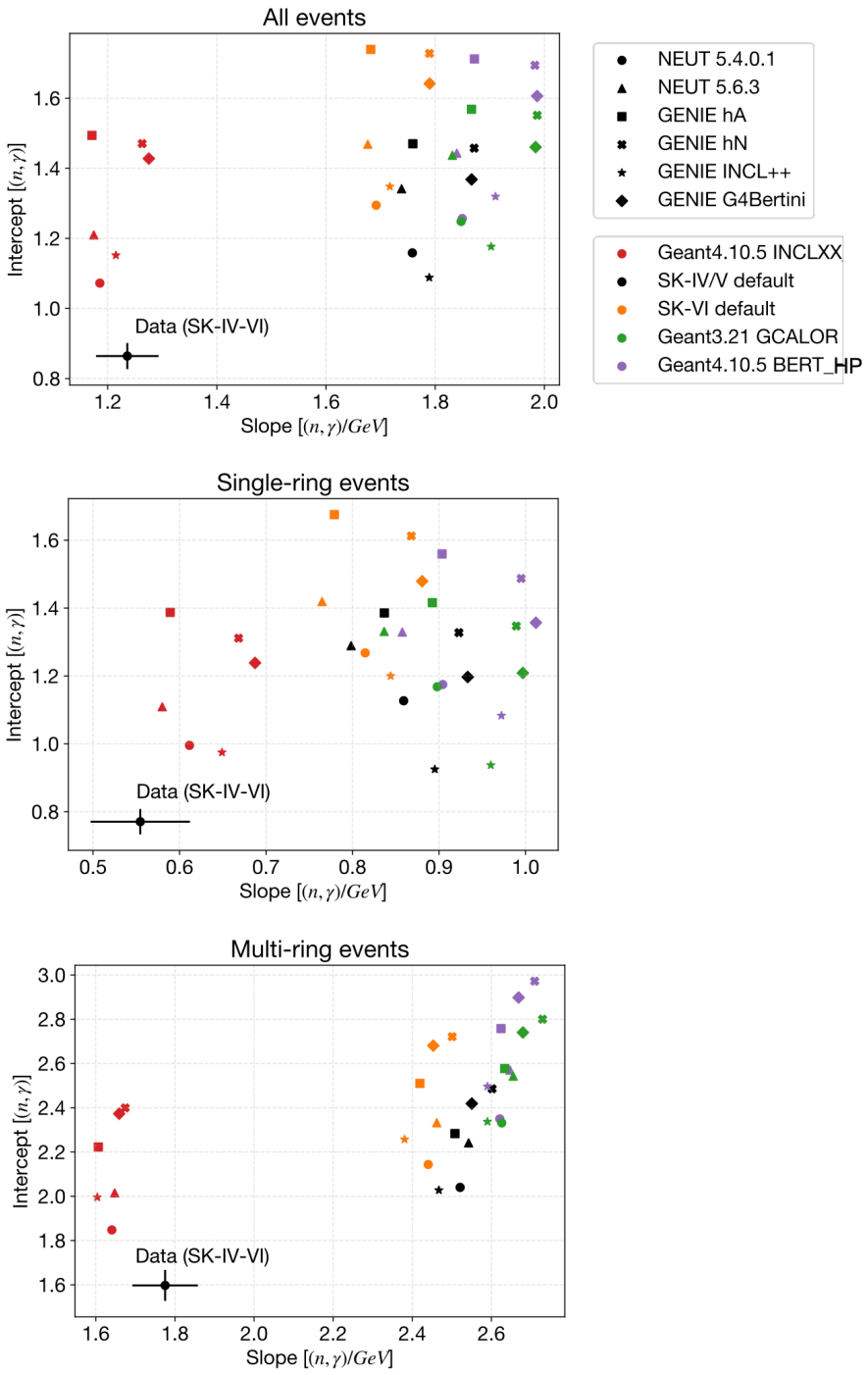


Figure 4.53: Comparison of fitted linear slope and intercepts between combined data and model combinations.

Rank	PI/FSI	SI	Reduced χ^2	p-value
1	NEUT 5.4.0	INCL	0.394461	1.00
2	GENIE INCL	INCL	0.490170	1.00
3	NEUT 5.6.3	INCL	0.614911	0.99
4	GENIE Bertini	INCL	1.403655	0.05
5	GENIE hA	INCL	1.527829	0.02
6	GENIE hN	INCL	1.570021	0.01
7	GENIE INCL	SK-IV/V default	2.526251	0.0
8	NEUT 5.4.0	SK-IV/V default	2.599332	0.0
9	NEUT 5.4.0	SK-VI default	2.778316	0.0
10	GENIE INCL	SK-VI default	3.282918	0.0
11	NEUT 5.6.3	SK-IV/V default	3.423120	0.0
12	NEUT 5.6.3	SK-VI default	3.722445	0.0
13	NEUT 5.4.0	GCALOR	3.853385	0.0
14	NEUT 5.4.0	G4Bertini	3.923076	0.0
15	GENIE INCL	GCALOR	3.981559	0.0
16	GENIE hA	SK-IV/V default	4.463219	0.0
17	GENIE Bertini	SK-IV/V default	4.816538	0.0
18	NEUT 5.6.3	GCALOR	4.942095	0.0
19	GENIE INCL	G4Bertini	4.949126	0.0
20	NEUT 5.6.3	G4Bertini	5.069997	0.0
21	GENIE hN	SK-IV/V default	5.522949	0.0
22	GENIE hA	SK-VI default	5.909420	0.0
23	GENIE Bertini	SK-VI default	6.151070	0.0
24	GENIE hA	GCALOR	6.371587	0.0
25	GENIE Bertini	GCALOR	6.883173	0.0
26	GENIE hN	SK-VI default	6.931615	0.0
27	GENIE hN	GCALOR	7.712574	0.0
28	GENIE hA	G4Bertini	7.755188	0.0
29	GENIE Bertini	G4Bertini	8.219881	0.0
30	GENIE hN	G4Bertini	9.031177	0.0

Table 4.11: Model combinations ranked by the reduced-χ^2 between data (all events) and predictions assuming linear model based on the fitted slope and intercepts shown in Figure 4.53. The χ^2 and p-values are roughly estimated based on the assigned total errors in our data.

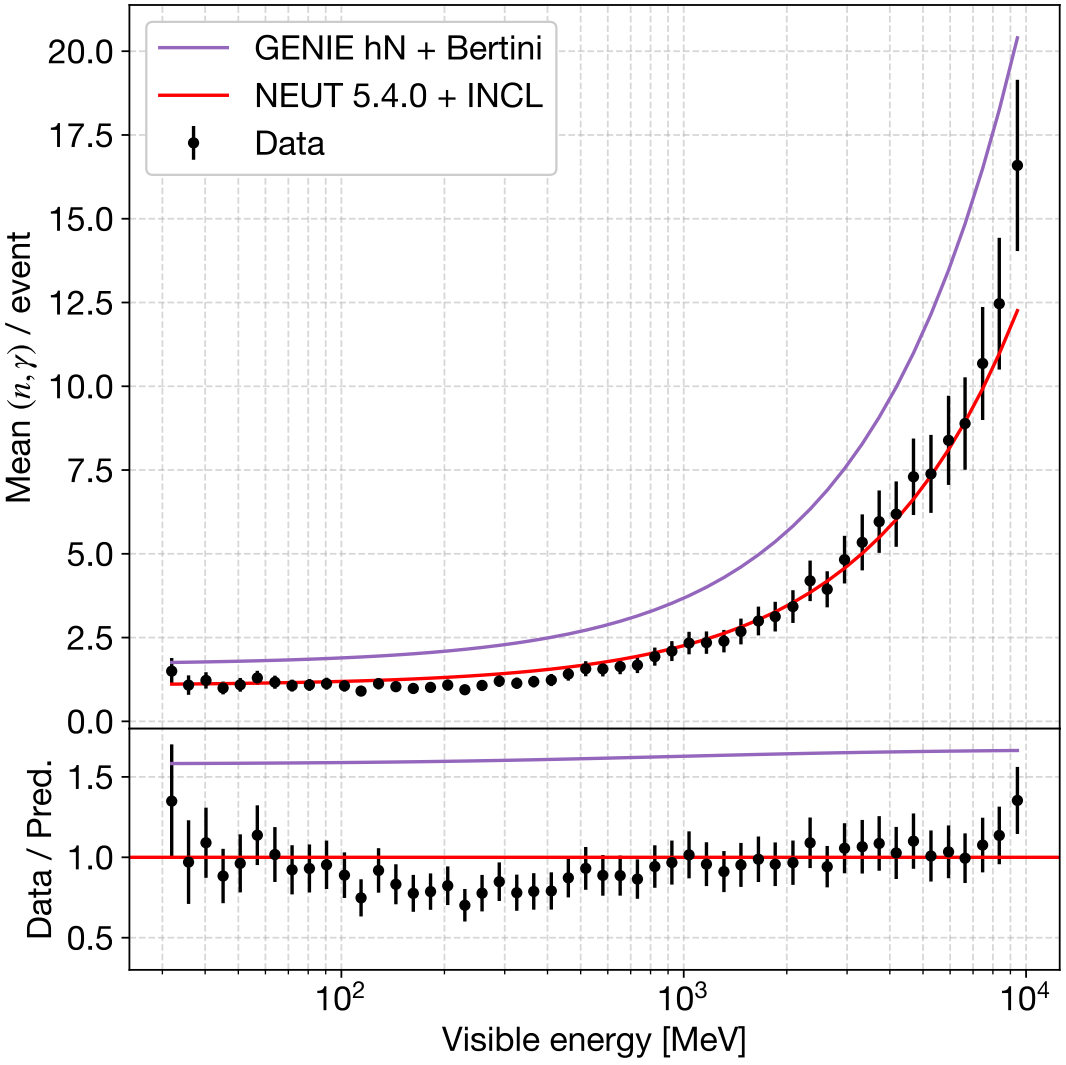


Figure 4.54: Data comparison with the linear model predictions assuming fitted slope and intercept from (`NEUT` 5.4.0 + `INCL`) and (`GENIE` hN + `Bertini`) model combinations.

Let us summarize the implications of our findings:

- Slope

  As shown in Figures 4.51 and 4.52, the slope of `INCL` was notably lower than that of other SI models, aligning well with the observed slope in all combinations. In contrast, other SI models showed less compatibility with our data, with Geant4 Bertini, widely used in detector simulation, showing the poorest performance.

- Intercept

  While the distinction is not as pronounced as with SI models, different PI/FSI model choices also exhibit differences in the intercept. For instance, `GENIE` hN and hA models generally show larger intercepts compared to other models, making them incompatible with the data. The use of `INCL` as an FSI model reduces the intercept by 20-30%.

  `NEUT` 5.4.0 generally shows the smallest neutron production among all PI/FSI options. This is attributed to a larger fraction of CCQE events compared to non-QE events in the lower energy end, as demonstrated in Figure 4.43 and 4.45. The almost identical nature of `NEUT` 5.6.3 and `NEUT` 5.4.0, except for the nuclear binding energy correction, allows us to attribute the 10% difference between the two `NEUT` versions to the difference in the CCQE fraction.

Finally, we compared the average neutron "track length" or "travel distance" – the reconstructed distance between the neutrino interaction vertex and the subsequent neutron capture vertex – observed in the SK-VI (R) data with model predictions. The model predictions for different neutrino event generator options were obtained from full detector simulation. However, due to the absence of full detector simulation for all model combinations involving different secondary interaction model options, we conducted full detector simulations for the `NEUT` 5.4.0 option and four distinct secondary interaction models – `GCALOR`, `Bertini`, `INCL`, and Geant4 Binary Cascade model (`BIC`) [171].

Figure 4.55 illustrates the average neutron track length as a function of event visible energy for the four secondary interaction models. Generally, for an energy range over 100 MeV, the average neutron track length increases with the event visible energy, aligning with expectations. In the lower energy range, the length is larger due to the poorer resolution of the neutrino interaction vertex, which competes with the actual distance traveled by the neutrons. Bertini-like models (GCALOR, Bertini) predict approximately 5% larger average distances compared to more sophisticated models (`INCL`, `BIC`) across all energy ranges. Based on this observation, we assumed a uniform 5% uncertainty in the average distance, accounting for the uncertainty in the secondary interaction model, for all energy ranges.

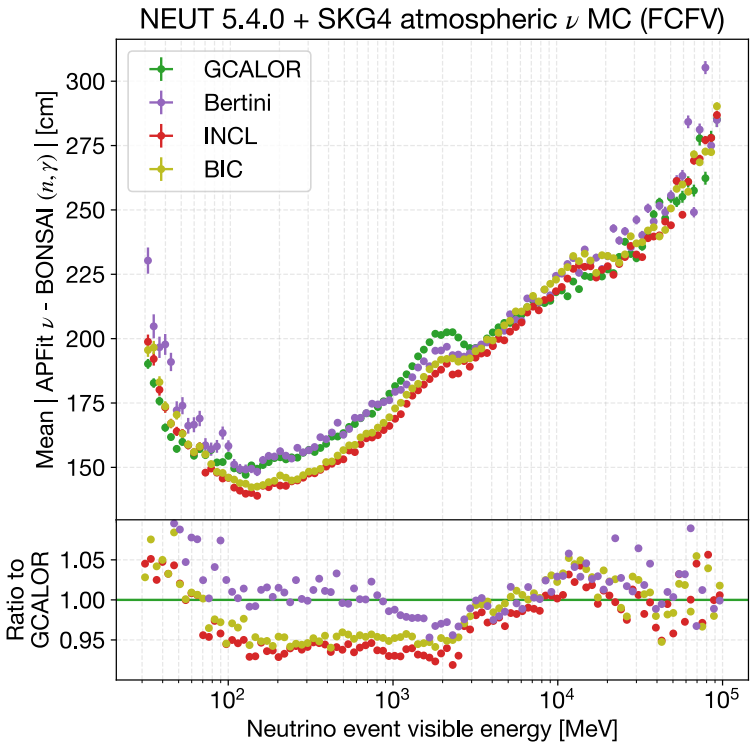


Figure 4.55: Comparison of average reconstructed neutron track length as a function of event visible energy among different sets of full atmospheric neutrino detector simulations with four different SI models. Only neutron captures with a distance of less than 10 m are considered.

Figure 4.56 presents a comparison of the observed average reconstructed distance between the neutrino interaction vertex and the subsequent neutron capture vertex with the corresponding model predictions using different neutrino event generator options. The uncertainty in these predictions was added by a 5% uniform uncertainty accounting for the variability in the secondary interaction model.

The observed average neutron track length was slightly smaller than the full MC prediction of the default model combination, slightly favoring more sophisticated models such as `INCL` and `BIC` over Bertini-variants. Overall, the observation was in agreement with the model predictions within the assumed 5% secondary interaction model uncertainty and 5-10% statistical uncertainties. The variability in predictions within different neutrino event generator options (FSI models) had a smaller impact ($\sim 3\%$) compared to the secondary interaction model.

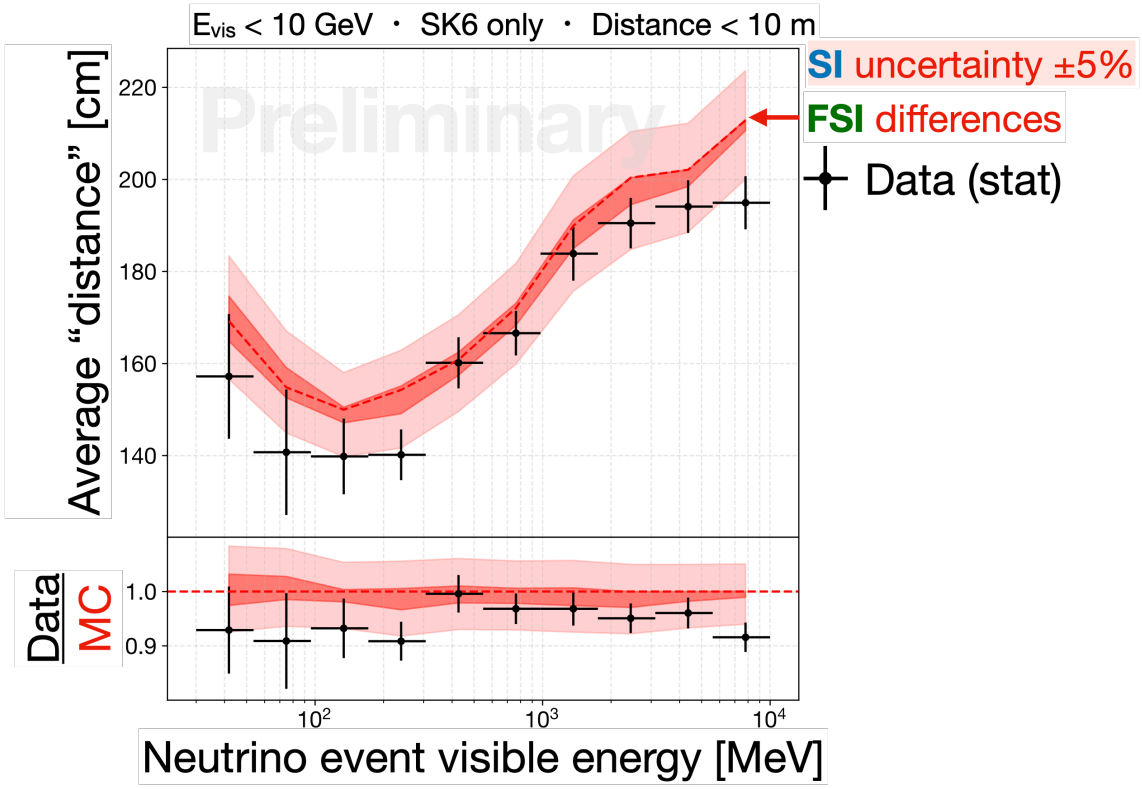


Figure 4.56: Comparison between SK-VI (R) observations and the predicted range for the average neutron "travel" distance, which is the reconstructed distance between the neutrino interaction vertex and the subsequent neutron capture vertex. Only neutron captures with a distance of less than 10 m are considered.
The red dashed line represents the full MC prediction with the default model combination (`NEUT` 5.4.0 and SK-VI default detector simulation), and the thicker red shaded area represents the range of varying predictions due to changes in the neutrino event generator option, reflecting the variability in predictions due to PI and FSI model uncertainty. The lighter red shaded area represents the uniform 5% uncertainty due to SI model uncertainty, as observed in Figure 4.55.

Chapter 5

Discussion and conclusions

5.1 Discussion

What are the key factors that cause variations in model predictions?

The key distinctions among the models lie in their use of different INC models for the simulation of hadron inelastic interactions. The inelastic interaction of nucleons, especially those of neutrons, is expected to be the predominant source of neutron signal production (see Figure 4.49).

There is a notable discrepancy in the model predictions for outgoing nucleon multiplicity and kinematics, particularly in the prediction of low-energy nucleons (specifically, below 300 MeV in kinetic energy or 800 MeV/c in momentum, as shown in Figure 4.47). In this energy range, incident nucleons near the ^{16}O Fermi momentum are likely to knock out another nucleon from the nucleus, as demonstrated by the $^{16}\mathrm{O}(n, 2n)^{15}\mathrm{O}$ cross-section in Figure 5.1. Therefore, the INC model's prediction of low-energy nucleon multiplicity as well as the target nucleon's momentum distribution should be the primary determinant of the final neutron signal multiplicity.

The semi-classical approach of INC models is generally not expected to perform well for incident hadrons with kinetic energy less than 100 MeV (or momentum less than 450 MeV/c), where the incident nucleon wavelength becomes comparable to the size of a nucleon. However, certain INC models, such as `INCL`, incorporate more quantum mechanical considerations compared to others, and it is worth exploring how these considerations impact predictions. Here, we highlight unique features of `INCL` that differentiate it from other models:

- **Tracking motion of all target nucleons**

  Besides the projectile, all nucleons in the target nucleus are tracked in both space and time within a specific interaction time. The cascade stops even when some particles remain within the nucleus.

- **Position and momentum-dependent target nucleon potential**

  Nucleons are placed in a potential well whose size and depth depend on nucleon position and energy, resulting in variable escape probability.

- **Probabilistic Pauli blocking**

  Not every outgoing nucleon below the Fermi momentum p_F is blocked; instead, it is determined by taking into account the probabilistic occupancy of nucleon energy states. This plays a crucial role in shaping `INCL`'s prediction of low-energy nucleon production[1], placing it in a middle ground between models that completely disregard Pauli blocking and those that strictly apply it for all nucleons below p_F.

- **Coulomb repulsion**

  Coulomb repulsion of the nucleus affects the trajectories of incident protons (reducing effective cross section), but does not affect neutrons.

[1]Recent studies indicate that INC models can effectively extend their validity to levels as low as a few MeV. For instance, incorporating the "discrete-level-constraint (DLC)" instead of the continuous distribution of nucleon energy states over the Fermi sea has been shown to be particularly promising in this context [172, 173].

- **Nuclear de-excitation after the end of cascade**

  Nucleons or gamma-rays are emitted as a result of the residual nucleus de-excitation.

These characteristics of `INCL` can be succinctly described as capturing a more realistic nuclear model that takes into account nucleon correlations, albeit in a semi-classical manner. The improved nuclear model is likely a contributing factor to the enhanced reproducibility of data, particularly in the low-energy region where simpler versions of INC model without nucleon correlations are expected to fail.

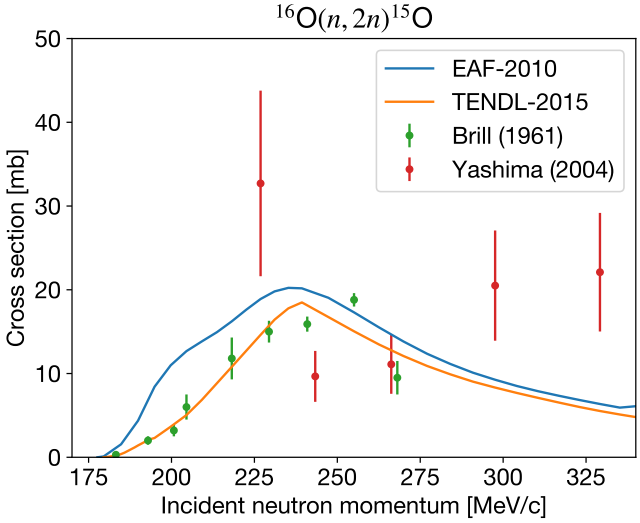


Figure 5.1: $^{16}O(n,2n)^{15}O$ cross-section as a function of incident neutron momentum.

What can be said about PI or FSI models?

The impact of PI/FSI models on the observed neutron multiplicity in atmospheric neutrino events at SK is not as pronounced as the impact of SI models. Therefore, it is hard to make definitive statements about the PI/FSI models. However, we can infer that the native neutron production in `GENIE` is not favored by our measurement results unless it is supported by `INCL`. Achieving a reliable estimation of neutron detection efficiency in the low-energy range with less than a 10% uncertainty could provide a basis for more definitive conclusions in the future.

Implications on the neutron deficit observed in earlier studies

Previous studies conducted by SNO [33], T2K [34], and MINERvA [35, 36] all employed the Geant4 Bertini cascade model (SNO, MINERvA) or Geant3 GCALOR (T2K), both of which were shown to overpredict neutron production in this study. Utilizing `INCL` as the secondary hadron inelastic scattering model could potentially resolve this issue.

The overprediction issue observed in earlier studies was particularly pronounced at lower energy transfer or visible energy, aligning with the observations in this study. MINERvA also noted a deficit in low-energy recoil proton signals compared to Geant4 predictions, which might be linked to the Geant4 Bertini model overpredicting low-energy neutrons compared to other models, as illustrated in Figure 4.47.

Unaccounted sources of model prediction variability

In addition to the discussed factors, there might be additional variability in hadronization models that determine outgoing hadron kinematics and multiplicity between neutrino interaction and FSI. Moreover, potential differences between INC and non-INC models could contribute to the overall variability in model predictions.

5.2 Summary

Precisely predicting the number of neutrons produced in neutrino interactions is crucial for advancing future projects involving large water Cherenkov detectors, such as Super-Kamiokande (and potentially Hyper-Kamiokande). These projects include the analysis of neutrino oscillation and the search for rare events like proton decay and the diffuse supernova neutrino background (DSNB).

This thesis presents the measurement of the average neutron multiplicity observed in atmospheric neutrino interactions within the water volume of the Super-Kamiokande detector. The measurement is provided as a function of the event's electron-equivalent "visible" energy, a semi-calorimetric measure of the neutrino momentum transfer. The primary motivation behind this study was to validate the observed neutron deficit from previous studies and evaluate the uncertainty in total neutron production relative to kinematic observables.

The key advancements from previous studies include:

- Extensive dataset spanning 12 years of atmospheric neutrino events in water
- Broad energy range spanning 30 MeV to 10 GeV,
  made possible by accurate modeling of neutron detection efficiency.
- Measurement of neutron track length as a proxy for neutron kinematics,
  made possible through recent Gd-loading.
- Inclusion of `INCL`, an intranuclear cascade model with a realistic nuclear model,
  which has not been previously compared in neutron production measurements.

For this purpose, a neutron detection algorithm was established based on a simple low-energy trigger and a neural network binary classification, calibrated with an Am/Be neutron source. The thesis presents an extensive study on the uncertainty of neutron detection performance, revealing a noteworthy bias stemming from the calibration setup and thermal neutron capture simulation errors in Geant4. These findings may prove beneficial for future water Cherenkov detectors.

Neutron multiplicity was estimated through an event-by-event prediction of neutron detection efficiency using a multivariate non-linear regression technique applied to simulation data. The observed neutron multiplicity showed a linear correlation with the visible energy of the neutrino event. Multi-ring events, characterized by a higher fraction of inelastic interactions, exhibited larger multiplicities compared to single-ring events. The results were compared with predictions from various hadronic interaction models. While a deficit of neutrons was observed compared to traditional, default semi-classical hadronic interaction models, better agreement was achieved when using the `INCL` — a similar semi-classical model but with more sophisticated description of nucleus. This highlights the significance of nuclear modeling in predicting low-energy nucleon production and kinematics.

5.3 Outlook

In the coming years, the ANNIE experiment [174] is set to measure neutron production and relevant cross-sections using a Gd-loaded water Cherenkov detector situated near low-energy neutrino beamline at Fermilab. T2K has recently integrated a new fine-grained scintillator tracker into their near detector complex, allowing for the observation of neutron signals [175], similar to MINERvA. The WCTE experiment [176], a fine-grained compact water Cherenkov detector near the secondary hadron test beamline at CERN, aims to measure neutron production from hadron-water interaction. This measurement will directly constrain hadronic interaction models and disentangle FSI and SI effects.

This study, combined with these forthcoming efforts, has the potential to advance future investigations. A notable impact is expected in improving the analysis of atmospheric neutrino oscillations targeting neutrino mass ordering and CP violation. The key lies in the differences in oscillation probabilities between ν_e and $\bar{\nu}_e$ as illustrated in Figure 1.8. The classification of ν_e and $\bar{\nu}_e$ is imperative, and for atmospheric neutrino analysis, this can be accomplished by tagging recoil neutrons from $\bar{\nu}_e$. Thus, precise estimation of neutron production in neutrino events and associated uncertainties is crucial. This study is expected to significantly reduce these uncertainties, thereby enhancing the overall analysis. To illustrate, we provide a simplified demonstration of how a more precise measurement of neutron production, as achieved in this study, can enhance the sensitivity — for δ_{CP} as an example — in the analysis of atmospheric neutrino oscillations at SK.

We focus on the impact of δ_{CP} sensitivity primarily in low-energy (<500 MeV) $\nu_e/\bar{\nu}_e$ events. For simplicity, we only examine the sub-GeV single-ring e-like sample, and compare predictions of the following three neutron production models: a) a "no-neutron" model (i.e., there is no use of neutron information in the analysis), b) the nominal model based on our default simulation setup, and c) an updated model with a 20% reduction in neutron production.

We assigned 1σ uncertainties in the predicted neutron multiplicity as 0%, 50%, and 10% for each respective model. Figure 5.2 illustrates the predicted mean neutron multiplicity as a function of reconstructed electron momentum.

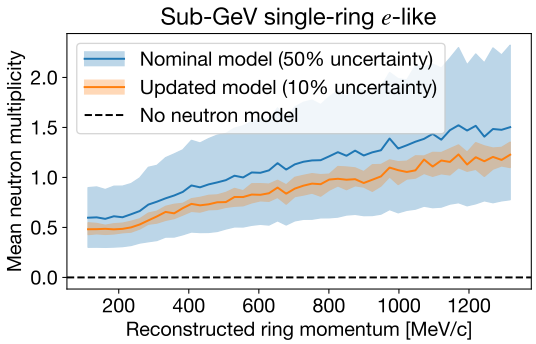


Figure 5.2: Predicted mean neutron multiplicity as a function of reconstructed electron ring momentum in the sub-GeV single-ring e-like sample, for three different neutron production models described in the text.

To enhance the purity of ν_e CC interactions and eliminate $\bar{\nu}_e$ events, we may reject events with one or more identified neutron signals. We assume a consistent 80% neutron detection efficiency across all models. Figure 5.3 compares the distributions of detected neutrons and the fractions of true ν_e events within 0-neutron subsamples for three different models. The figure illustrates that the fraction of true ν_e events improves with an increased number of detected neutrons.

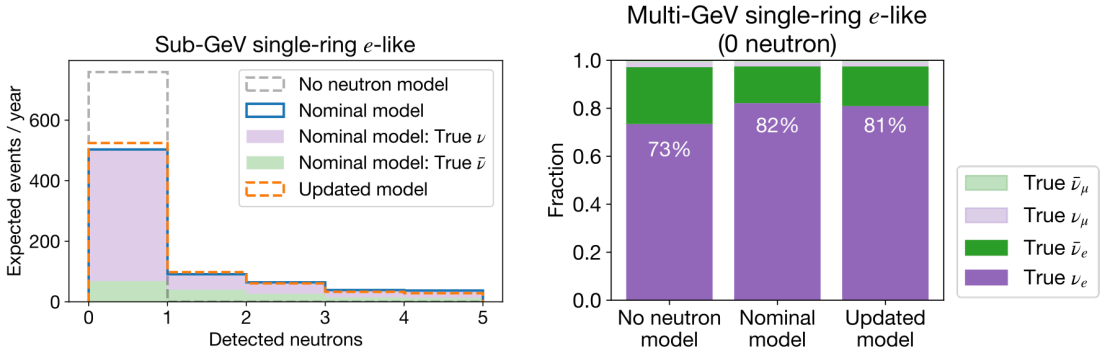


Figure 5.3: Expected Sub-GeV single-ring detected neutron distribution (left) and the fraction of true ν_e events in the 0-neutron subsample (right), comparing the three models described in the text.

Then, we can compare the expected number of events in each momentum and zenith angle bin by varying oscillation parameters, as is typically done in the SK atmospheric neutrino oscillation analysis. In this example, we compare the expected event counts with varying δ_{CP} (other parameters taken from Table 4.1). Figure 5.4 compares the event counts in momentum and zenith angle bins expected by the nominal model between CP conserving and violating scenarios. We see that the number of expected low-energy ν_e events is enhanced with assuming CP violation.[2]

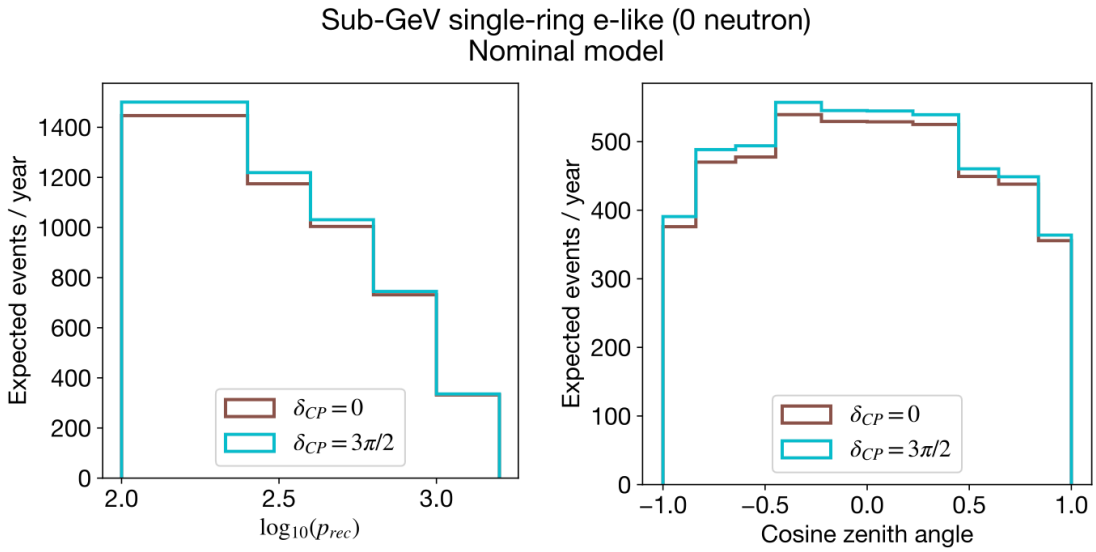


Figure 5.4: Sub-GeV single-ring e-like 0-neutron events per year in each reconstructed electron momentum (p_{rec}) bin (left) and zenith angle bin (right) expected by the nominal model, for CP conserving ($\delta_{CP} = 0$, brown) and violating ($\delta_{CP} = 3\pi/2$, cyan) neutrino oscillation.

To quantify the sensitivity to δ_{CP}, we may calculate the $\sum \chi^2(\delta_{CP}) \equiv 2\sum_i (e_i(\delta_{CP}) - o_i + o_i \ln(o_i/e_i(\delta_{CP})))$ based on the likelihood of the observed (in this example, expected by ground truth) bin count o_i in the i^{th} bin, based on the statistical uncertainties of the expected bin count $e_i(\delta_{CP})$ that varies with assumed δ_{CP}.

[2]CP violation has opposite-signed effects on $P(\nu_\mu \to \nu_e)$ and $P(\bar{\nu}_\mu \to \bar{\nu}_e)$, and the low-energy $\bar{\nu}_e$ events are expected to decrease with CP violation in the normal mass ordering scenario. We may leverage this effect further by purifying $\bar{\nu}_e$ events, however, purifying $\bar{\nu}_e$ is more challenging than purifying ν_e. This is because ν_e CC events also produce neutrons through hadronic re-scatterings, requiring additional considerations beyond a simple cut on neutron counts for purifying $\bar{\nu}_e$.

Comparing with the ground truth histograms, assuming CP violation ($\delta_{\text{CP}} = 3\pi/2$), Figure 5.5 shows the distribution of $\sum \chi^2$ as a function of the assumed δ_{CP} for the three models with distinct uncertainties. We observe that purifying ν_e events by requiring no tagged neutrons has an impact on the CP-symmetry ($\delta_{CP} = 0$) rejection power. Additionally, the size of the uncertainty in neutron production assumed by the models is directly translated to the variation in the expected sensitivity. While the updated model assumes reduced neutron production compared to the nominal model, the uncertainty in the nominal model is large, so the negative-1σ-varied sensitivity of the updated model actually becomes larger than that of the nominal model, as shown in Figure 5.6, albeit by a small amount ($\sim$5%).

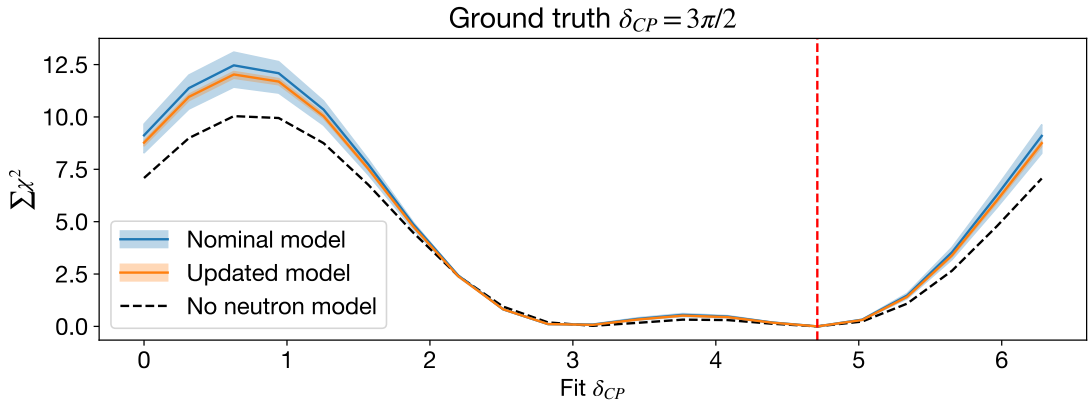


Figure 5.5: $\sum \chi^2$ for the three neutron production models assuming distinct uncertainties, compared with the ground truth histograms assuming CP violation, as a function of assumed δ_{CP} values. The red dashed line indicates the ground truth $\delta_{CP} = 3\pi/2$. The $\sum \chi^2$ is calculated for the momentum and zenith angle bins of sub-GeV single-ring e-like 0-neutron events as shown in Figure 5.4. The error band for each model was calculated based on model predictions of neutron production varied by 1σ uncertainty.

This simplified demonstration provides one example of how a reduced neutron production uncertainty can enhance the sensitivity of neutrino oscillation analysis using neutron counts. In the actual analysis, we should utilize the other subsamples and employ more sophisticated techniques to enhance neutrino reconstruction (e.g., $\nu/\bar{\nu}$ purification and missing neutrino momentum carried away by hadrons) based on observed neutron counts and kinematics. Additionally, we need to consider other systematic uncertainties that may wash out some of the sensitivity.

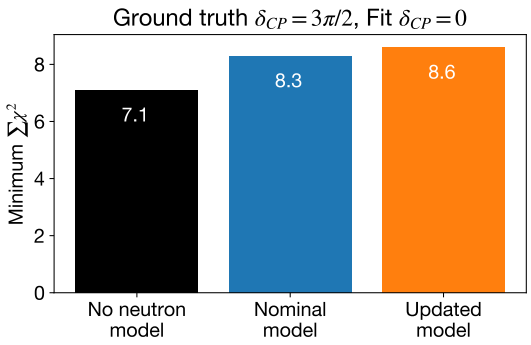


Figure 5.6: Minimum $\sum \chi^2$ between the CP violating and conserving scenarios for each model, considering neutron production uncertainties.

Apart from the analysis on atmospheric neutrino oscillation, we anticipate advancements in the following research areas due to reduced uncertainty in neutron production and improved selection of the hadron re-scattering model:

- **Accelerator neutrino oscillation analysis**

  Optimal choices of hadronic interaction models are expected to improve the capability of neutrino event generators in estimating outgoing particle kinematics and addressing contamination from hadronic, non-CCQE background interactions that is crucial for the test of neutrino CP symmetry.

- **Rare event searches in water Cherenkov detectors**

  This study may expedite rare event searches in water Cherenkov detectors that utilize neutron-tagging to reject atmospheric neutrino backgrounds. This study establishes a robust neutron detection method, reducing uncertainty in neutron multiplicity, potentially improving planned rare event searches, especially the search for proton decay.

Looking ahead, the future Hyper-Kamiokande, featuring an eight times larger fiducial volume, aims to provide a definitive test of the neutrino mass ordering and CP symmetry, as well as the evidence for the aforementioned rare events. This study provides a robust foundation for such future studies that rely on our understanding of secondary hadronic neutron production following neutrino interactions.

Appendix A

Reconstruction algorithm of neutrino interaction

A.1 Vertex Reconstruction

We first find the vertex $\vec{x}$ that maximizes the following estimator:

$$g(\vec{x}, T_0) = \frac{1}{\sum_i (1/\sigma_i^2)} \sum_i \exp\left\{ - \frac{(t_i - \Delta T_i(\vec{x}) - T_0)^2}{2(1.5\sigma_i)^2} \right\} \tag{A.1}$$

where σ_i is the timing resolution, t_i is the hit time of the i^{th} PMT, and $\Delta T_i(\vec{x})$ is the estimated photon time-of-flight from the assumed vertex $\vec{x}$ to the i^{th} PMT. T_0 is the assumed signal generation time.

The initial Cherenkov ring direction is determined by the normalized sum of vectors connecting the initial guess vertex to each hit PMT, weighted by the corrected hit charge $q_i^{\text{corr}}(\vec{x})$, given by:

$$q_i^{\text{corr}}(\vec{x}) = q_i^{\text{obs}} \exp\left(-\frac{|\vec{x}_i^{\text{PMT}} - \vec{x}|}{L} \frac{\cos\theta(\vec{x})}{f(\theta(\vec{x}))} \right)$$

Here, q_i^{obs} is the observed charge, $\vec{x}_i^{\text{PMT}}$ is the position of the i^{th} PMT, L is the light attenuation length in water, $\theta(\vec{x})$ is the incident angle assuming photon origin at $\vec{x}$, and $f(\theta)$ is the PMT angular acceptance.

The ring direction undergoes further refinement by testing various Cherenkov angles. The Cherenkov angle, typically 42° for particles moving close to the speed of light in water, may vary for particles with lower momenta. A new goodness-of-fit parameter, $G(\vec{d}, \theta_c)$, is introduced to achieve three goals: (i) penalize small, non-Cherenkov-like angles, (ii) match the opening angle to the expected value, and (iii) maximize the charge contained in a ring of the assumed angle.

$$G(\vec{d}, \theta_c) = \frac{1}{\sin\theta_c} \exp\left[-\frac{(\theta_c - \theta_{\text{exp}})}{2\sigma_\theta^2}\right] \left(\frac{\partial Q^{\text{corr}}(\vec{d}, \theta)}{\partial\theta}\Big|_{\theta=\theta_c}\right)^2 \int_0^{\theta_c} Q^{\text{corr}}(\vec{d}, \theta) d\theta$$

Here, $\vec{d}$ is the assumed ring direction, θ_c is the tested Cherenkov angle, and θ_{exp} is the expected Cherenkov angle, chosen from 42° assuming near-speed-of-light particles, the angle assuming a muon, or the angle assuming an electron with energies to produce the observed charge. σ_θ is the expected resolution of the Cherenkov angle. The first term addresses goal (i), the second term addresses goal (ii), and the third term addresses goal (iii).

Minimizing $G(\vec{d}, \theta_c)$ refines the ring direction and Cherenkov angle. The initial guess vertex is then further refined by minimizing the estimator in Equation A.1 with a test vertex constrained to be on the line defined by the fitted ring direction and the Cherenkov angle.

A.2 Ring counting

Hough transform

For ring counting, we employ the Hough transformation technique [177]. This technique aims to identify the optimal set of parameters of a function, typically forming a geometrical shape in the coordinate space, based on a limited set of available coordinates modeled by the function with a given parameter set. For instance, when dealing with a few points on a line in a 2D Cartesian coordinate system, the Hough transform helps identify the most suitable set of slope and intercept to best characterize the given set of points. Similar to conventional regression tasks, the key elements for estimating parameters using the Hough transform include the form of the function and the available data points.

In this technique, each data point is treated equivalently, representing all possible value sets of parameters that can model the given data point. Essentially, each data point in the coordinate space is transformed into all possible sets of parameters in the parameter space. Subsequently, the set of parameters that receives the most "votes" from the data points is likely the optimal parameter set for describing the data points. Figure A.1 provides an illustrative example of circle center estimation given a set of points on a circle.

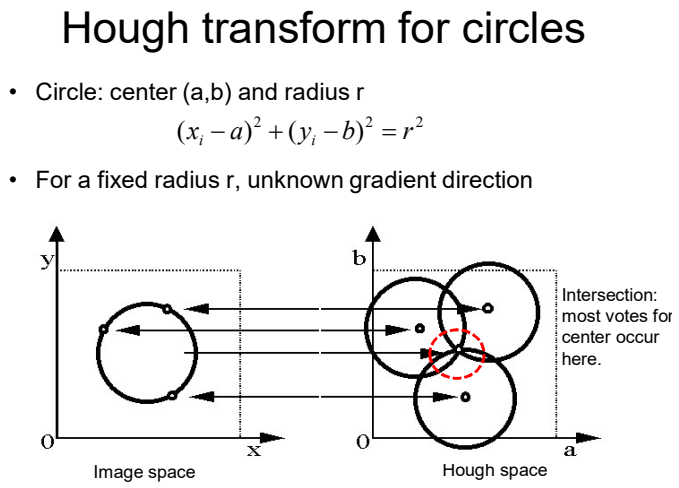


Figure A.1: Locating the center of a circle. (Reprinted from [178])

This method for estimating circle centers can be readily extended to the estimation of multiple circle centers. The distribution of weighted parameters in Hough space will exhibit multiple modes if there are several underlying circles. In the context of Cherenkov ring counting at SK, we can initially project all PMT hits onto a 2D plane represented by polar-azimuthal coordinates (Θ, Φ) calculated from the reconstructed vertex, given the cylindrical shape of our detector. Subsequently, the Cherenkov rings in this plane will appear as circles with a known radius (determined by the Cherenkov angle and the reconstructed vertex). The PMT hits then become our data points, each of which can be transformed into circles in the Hough space coordinates (Θ', Φ'). The circles in Hough space are further weighted by the observed charge of the corresponding PMT hit. An illustration of charge-weighted PMT hits in Hough space is presented in Figure A.2.

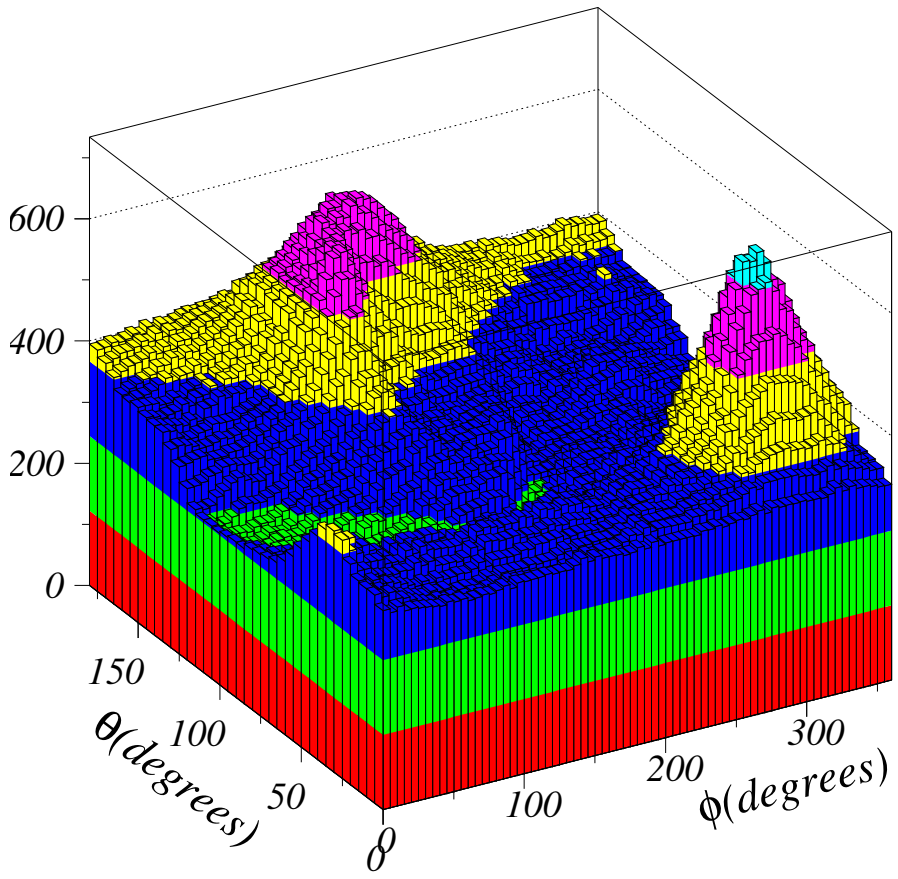


Figure A.2: Charge-weighted PMT hits in Hough space (Reprinted from [118])

The highest peak in the Hough space, as illustrated in Figure A.2, is identified as a Cherenkov ring. Consequently, every event is guaranteed to have at least one recognized ring, regardless of its faintness or potential absence. Then, rings are counted iteratively. We assess whether considering the next ring candidate (the next highest peak in the Hough space) enhances the likelihood of PMT hits, defined by:

$$L(N_{\text{ring}}) = \sum_{i,\theta_{i,j}<1.2\theta_{c,j}}^{N_{\text{PMT}}^{\text{ring}}} \log\left[P(q_i, \sum_{j}^{N_{\text{ring}}} q_{i,j}^{\text{exp}})\right] \tag{A.2}$$

where N_{ring} is the number of rings, $N_{\text{PMT}}^{\text{ring}}$ is the number of PMTs within (or near the edge of) the assumed Cherenkov rings. The index i iterates over such PMTs, and the index j iterates over the assumed Cherenkov rings. q_i is the observed charge of the i^{th} PMT, and $q_{i,j}^{\text{exp}}$ is the expected charge of the i^{th} PMT due to the j^{th} Cherenkov ring. $P(q_1, q_2)$ is the probability of observing q_1 when q_2 is expected. For low expected charge ($q_2 < 20$ photoelectrons), P is tabulated based on the observed single photoelectron charge peak of individual PMTs and a Poisson likelihood of observing multiple photons. For higher expected charge, P is sampled from a Gaussian distribution. Although the actual implementation of Equation A.2 is intricate, the fundamental concept is to iteratively include Cherenkov ring candidates from the Hough space if $L(N_{\text{ring}} + 1) > L(N_{\text{ring}})$.

A.3 Particle identification

Cherenkov rings are categorized as either e-like or μ-like. Electrons, with a lower critical energy of 70 MeV in water, exhibit a distinct behavior. Beyond this energy threshold, electrons tend to lose a significant amount of energy through bremsstrahlung, resulting in the production of gamma-rays and subsequent electrons. This process leads to an electromagnetic shower that makes the ring appear fuzzier. In contrast, muons, being much heavier, primarily lose a small amount of energy through ionization, preserving the sharpness of the ring edges. Figure 2.6 provides an illustration of e-like and μ-like rings.

For each identified Cherenkov ring, a likelihood of the PMT hit pattern is established based on both electron and muon scenarios. The lepton $l \in e, \mu$ scenario pattern likelihood for the j^{th} ring, denoted as L_j^{pattern}, is given by:

$$L^{\text{pattern}}j(l) = \prod^{N\text{PMT}^{\text{ring}}} i, \theta_i < 1.5\theta_c P\left(q_i, qi, j^{\text{exp}}(l) + \sum_{k \neq j} q_{i,k}^{\text{exp}}\right) \tag{A.3}$$

In this equation, the product is taken over all PMTs contained within the j^{th} ring with an opening angle given by 1.5 times the Cherenkov angle θ_c. The probability $P(q_1, q_2)$ retains the same definition as in Equation A.2. The last summation term adds the contribution to the total charge from other rings in the event without assuming a particle type. The expected charge $q_{i,j}^{\text{exp}}(l)$, assuming lepton type l for the j^{th} ring, is calculated based on tabulated values of MC-simulated charge for each PMT. This calculation considers the lepton momentum and the distance between the lepton vertex and the PMT. In the actual implementation, Equation A.3 and the expected charge $q_{i,j}^{\text{exp}}(l)$ take into account the contribution from scattered and reflected light, with the detailed process outlined in [118].

From Equation A.3, we can construct a probability assuming a χ^2 distribution, given by:

$$\chi_j^2(l) = -2 \log L_j^{\text{pattern}}(l) \tag{A.4}$$

$$P_j^{\text{pattern}}(l) = \exp\left[-\frac{\chi_j^2(l) - \min(\chi_j^2(e), \chi_j^2(\mu))}{4N_{\text{PMT}}^{\text{ring}}}\right] \tag{A.5}$$

We can also consider the potential differences in Cherenkov angles (for single-ring events only) between electron and muon scenarios by introducing $P^{\text{angle}}j(l)$:

$$P^{\text{angle}}j(l) = \exp\left[-\frac{(\theta c, j - \theta_j^{\text{exp}}(l))^2}{2\sigma\theta}\right] \tag{A.6}$$

where $\theta_{c,j}$ is the reconstructed Cherenkov angle of the j^{th} Cherenkov ring, and $\theta_j^{\text{exp}}(l)$ is the expected Cherenkov angle under the lepton type l scenario. σ_θ is the

uncertainty in the Cherenkov angle reconstruction. For multi-ring events, we simply set $P_j^{\text{angle}}(l) = 1$. Then, the total probability for each lepton type l scenario for the j^{th} Cherenkov ring is given by $P_j(l) = P_j^{\text{pattern}}(l)P_j^{\text{angle}}(l)$. The preference for the scenarios can be quantified by the following single parameter, which we call "PID likelihood" or "PID parameter":

$$L_j^{\text{PID}} = \sqrt{-\log P_j(\mu)} - \sqrt{-\log P_j(e)} \tag{A.7}$$

If this parameter is positive, the ring is classified as e-like, and if negative, μ-like. Scaling the parameter by a constant does not affect the classification performance. The distribution of this parameter is shown in Figure 4.9 in Section 4.2.1 (For T2K data and simulation, it is shown in Figure 2.6, although using a different algorithm [43]). The overall mis-identification probability is at $O(1)\%$ level.

A.4 Momentum reconstruction

Each Cherenkov ring corresponds to a particle with a particular momentum, computed from the sum of the observed charge within the Cherenkov ring. In the case of multi-ring events, rings may overlap, so the observed charge must be divided among each ring. The observed charge on the i^{th} PMT from the j^{th} ring is defined using the fraction of the expected charge from the j^{th} ring ($q_{i,j}^{\text{exp}}$), such that:

$$q_{i,j} = q_i \left(\frac{q_{i,j}^{\text{exp}}}{\sum_j q_{i,j}^{\text{exp}}} \right) \tag{A.8}$$

Then, the total charge contribution Q_j^{total} (also known as RTOT in SK) from the j^{th} Cherenkov ring is quantified as:

$$Q_j^{\text{total}} = \frac{G_{\text{MC}}}{G_{\text{Data}}} \left(\sum_i q_{i,j}^{\text{corr}} - \sum_j Q_j^{\text{scat}} \right) \tag{A.9}$$

Here, G_{MC} and G_{Data} are the overall gain scaling factors assumed in MC and observed in data, respectively. The first term in the parentheses is the sum of the charge contribution of the individual PMT hits of the j^{th} Cherenkov ring, corrected by the PMT angular acceptance and the light attenuation in water. The second term is the contribution by the expected scattered light due to all Cherenkov rings. Specific conditions imposed on PMTs when calculating Equation A.9 can be found in [118]. Q_j^{total} is converted into lepton momentum using a lookup table based on MC simulation.

Appendix B

Data selection stages

B.1 FC1: Initial cuts to reject evident backgrounds

Low-energy radioactivity rejection

- $PE_{300} > 200$ p.e.

  PE_{300}, the maximum total ID PMT charge in a sliding 300-ns window, must be greater than 200 photoelectrons (equivalent to 22 MeV electron).

Cosmic ray rejection

- The OD trigger must be off, OR
- $NHITA_{800} \leq 50$

  $NHITA_{800}$, the number of OD PMT hits within [-500, +300] ns window from the event trigger, must be equal to or smaller than 50 hits.

The first cut removes $O(1)$ MeV radioactivity from our data. These cuts reduce the event rate from $O(10^6)$ to $O(10^3)$ per day.

B.2 FC2: Subsequent cuts to further reject backgrounds

Flasher and radioactivity rejection

- $PE_{max}/PE_{300} < 0.5$

  The maximum charge in any single PMT (PE_{max}) must be smaller than than half the maximum total ID PMT charge in a sliding 300-ns window (PE_{300}).

Further cosmic ray rejection

- The OD trigger must be off, OR
- $NHITA_{800} \leq 30$, OR

  The number of OD PMT hits within [-500, +300] ns window from the event trigger must be equal to or smaller than 30 hits.

- $PE_{total} > 100{,}000$ p.e.

  The total charge deposited in the ID must be greater than 100,000 photoelectrons. For such events, OD PMT can register fake hits due to crosstalk with nearby ID PMTs, so they should be exempt from the above OD selections.

These cuts reduce the event rate down to $O(10^2)$ events per day.

B.3 FC3: Cosmic ray rejection based on fit

Even at this point, a significant portion of the remaining data events consists of cosmic ray muons and low-energy noise events that resemble signals. These events typically manifest as having relatively small OD activity and non-negligible ID activity.

To further reject cosmic ray muons, we use a specialized muon fitter for events with over 1,000 ID PMT hits, where more than 230 photoelectrons are deposited. This fitter determines the entry time (T_0) and the entry and exit (or stop) vertices ($\vec{x}$entry and $\vec{x}$exit/stop) that maximize the "goodness" parameter g according to the following expression:

$$g(\vec{x}_{\text{entry}}, \vec{x}_{\text{exit/stop}}, T_0) = \frac{1}{\sum_i (1/\sigma_i^2)} \sum_i \exp\left\{ -\frac{(t_i - \Delta T_i(\vec{x}_{\text{entry}}, \vec{x}_{\text{exit/stop}}) - T_0)^2}{2(1.5\sigma_i)^2} \right\} \tag{B.1}$$

Here, σ_i represents the timing resolution, t_i is the hit time of the i^{th} PMT, and $\Delta T_i(\vec{x}\text{entry}, \vec{x}\text{exit/stop})$ is the estimated photon arrival time on the i^{th} PMT since T_0. The estimation of ΔT_i involves determining the photon production vertex on the muon track based on the initial and final muon vertices, as well as the Cherenkov angle, and then summing the muon and photon time-of-flight.

Through-going muon rejection

- $g < 0.75$, OR

  The maximal through-going muon goodness g must be less than 0.75.

- $\text{NHITA}_{\text{entry}} < 10$ and $\text{NHITA}_{\text{exit}} < 10$

  The number of hit OD PMTs located within 8 m from the entry (exit) point in a [-500, +300] window $\text{NHITA}_{\text{entry}}$ ($\text{NHITA}_{\text{exit}}$) must be smaller than 10 PMTs.

Stopping muon rejection

- The goodness fitter fails to find the appropriate set of muon entry and stopping vertices, OR
- $\mathrm{NHITA_{entry}} < 10$

  The number of hit OD PMTs located within 8 m from the entry point in a [-500, +300] window $\mathrm{NHITA_{entry}}$ must be smaller than 10 PMTs.

Cable-hole muon rejection

On top of the SK detector tank, there are twelve "cable-holes" designed for the passage of signal and high voltage supply cables. As illustrated in Figure B.1, no OD PMT is positioned at these hole locations. When cosmic ray muons traverse these openings, they enter the ID region without hitting OD PMTs. Consequently, these events might be mis-identified as contained neutrino events due to the absence of an OD signal. To prevent such mis-identifications, veto counters (consisting of a 2×2.5 m plastic scintillation counter and a PMT) were strategically installed at four out of the twelve cable-holes. Figure B.1 provides a schematic representation of the cable hole and the associated veto counter. Events with no veto counter hits are selected for further analysis.

Low-energy event rejection

- $\mathrm{NHIT_{total}} \geq 500$ hits, OR

  The total number of ID PMT hits must be greater than or equal to 500 hits.
- $\mathrm{NHIT_{50}} \geq 50$ hits

  The maximum number of ID PMT hits in a sliding 50-ns window ($\mathrm{NHIT_{50}}$) must be greater than or equal to 50 hits. Before counting the number of hits in the sliding window, the photon time-of-flight (TOF) is corrected for based on the vertex that minimizes the variance of the TOF-corrected hit times. $\mathrm{NHIT_{50}}$ of 50 hits corresponds to visible energy of 9 MeV.

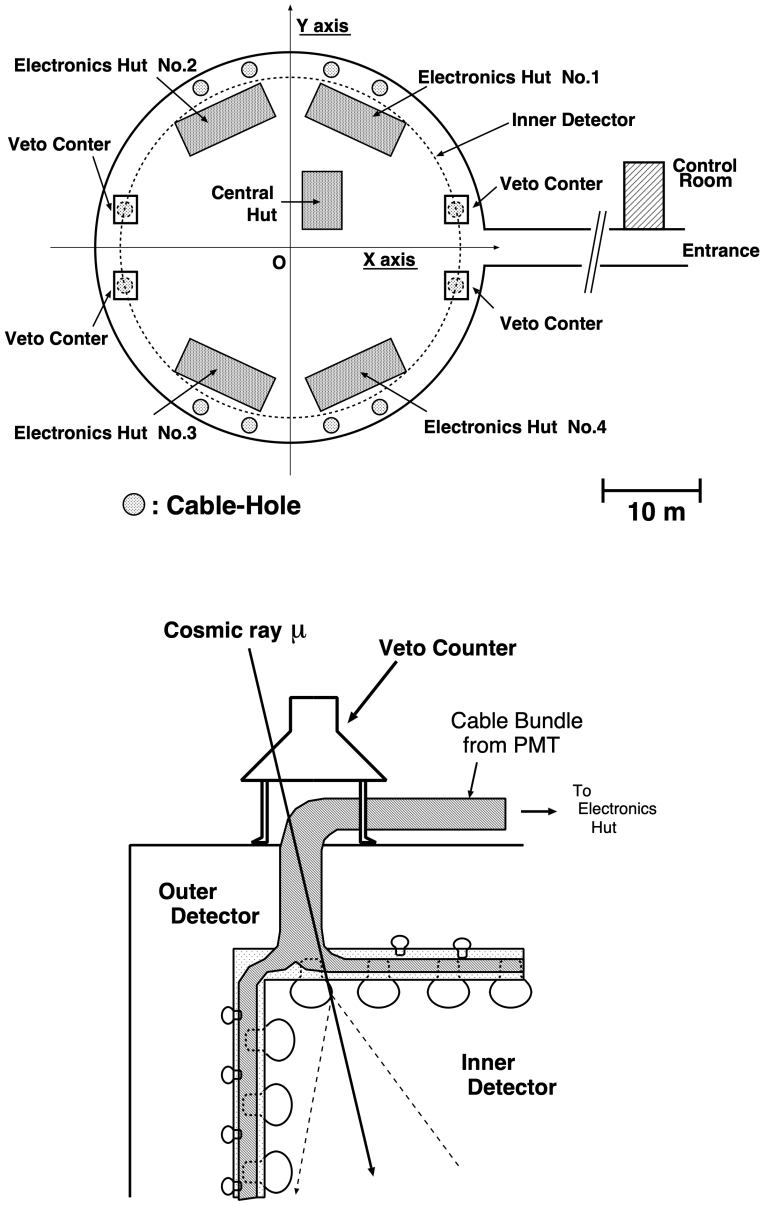


Figure B.1: A bird's-eye view of the SK tank indicating the locations of cable holes (top) and a sectional view showing a cable hole equipped with the veto counter (bottom).

Flasher rejection

Flasher events often have a broad timing distribution compared to events produced by charged particles. The hit time distributions of typical flasher events and neutrino events are shown in Figure B.2.

- $\mathrm{NHIT}_{100} < 20$

  The minimum number of ID PMT hits in a sliding 100-ns window within [+300, +800] μs range from the event trigger must be less than 20 hits.

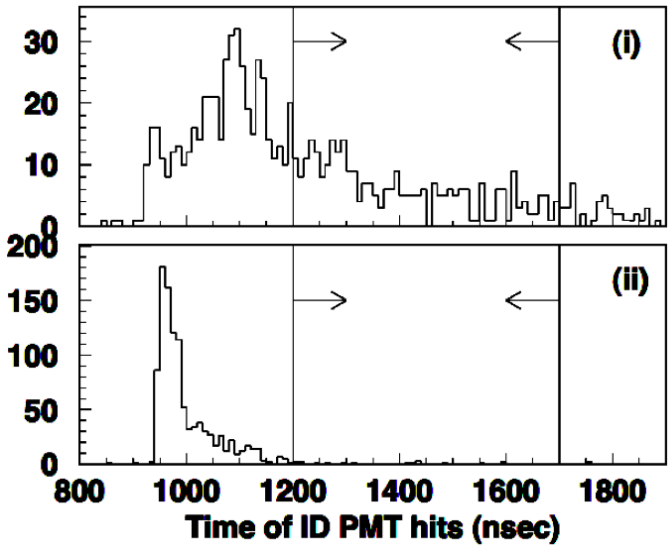


Figure B.2: PMT hit timing of flasher (top) and neutrinos (bottom)

Accidental coincidence rejection

If a low-energy radioactivity triggers an event without OD signature and a cosmic ray accidentally comes in and fires both the OD and ID PMTs within $O(100)$ ns from the event trigger, such events can elude the previous OD cuts designed to reject OD PMT hits near the event trigger.

- $\mathrm{PE}_{\mathrm{late}} < 5{,}000$ p.e., OR

The OD PMT charge within [+300, +800] ns from the event trigger ($\mathrm{NHITA}_{\mathrm{late}}$) must be smaller than 5,000 photoelectrons.

- $\mathrm{NHITA}_{\mathrm{late}} < 20$

  The number of OD PMT hits within [+300, +800] ns from the event trigger ($\mathrm{NHITA}_{\mathrm{late}}$) must be smaller than 20 hits.

The event rate at this point is around 45 events per day.

B.4 FC4: Flasher rejection based on likelihood

Events triggered by PMT flashers, i.e., the internal discharge of PMTs, tend to exhibit a single PMT hit with a very large charge deposit and small, isotropically distributed nearby PMT hits due to the light emitted by the discharge. One of the major characteristics of these events is that a PMT that "flashes" tends to flash repeatedly for a certain period of time. As a result, multiple events with very similar hit patterns — both in charge and positional distributions — may pass the three aforementioned data selection stages. Typical hit patterns of flasher events are shown in Figure B.3.

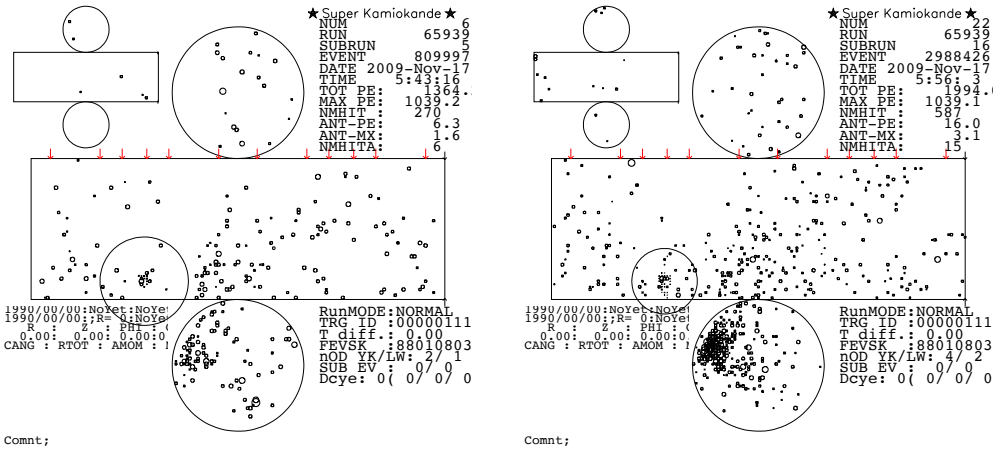


Figure B.3: Typical recurring flasher events.

We take advantage of this characteristic to reject flasher events. In simple terms, we reject events that have multiple occurrences with recurring hit patterns. To achieve this, we calculate two metrics that quantify the "similarity" of a tested event with the other compared events.

One metric is the "patch-by-patch charge correlation" r given by:

$$r = \frac{1}{N} \sum_i \frac{(Q_i^A - \langle Q^A \rangle) \times (Q_i^B - \langle Q^B \rangle)}{\sigma^A \times \sigma^B} \tag{B.2}$$

where we divide the ID wall into 1,450 patches of 2 m × 2 m squares, and N is the number of patches (1,450). $Q_i^{A(B)}$ represents the total charge in the i^{th} patch of the

two compared events A and B, $\langle Q^{A(B)} \rangle$ is the average, and $\sigma^{A(B)}$ are the standard deviation of $Q_i^{A(B)}$.

The other metric is the "overall PMT-by-PMT charge discrepancy" d_{KS}, which is the Kolmogorov-Smirnov test statistic for the accumulated PMT-by-PMT charge distributions of the two compared events.

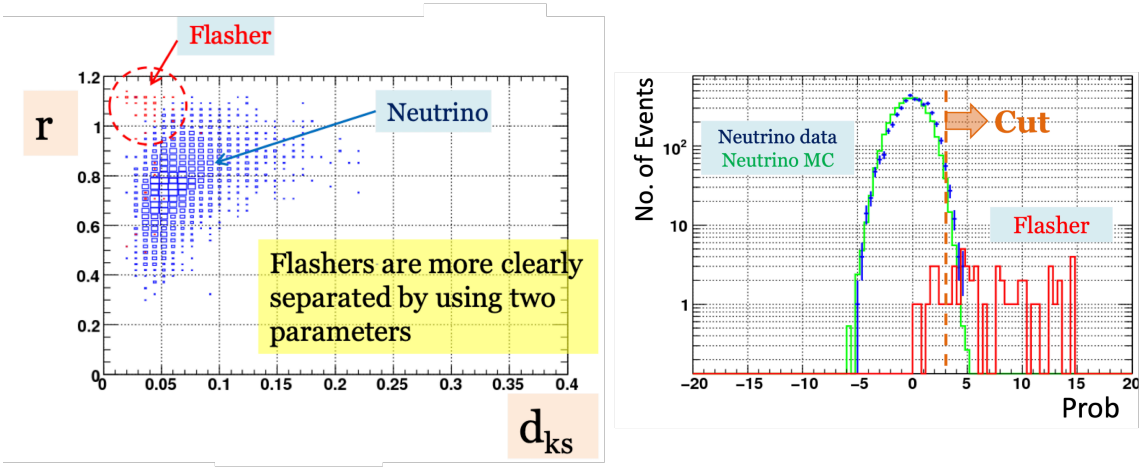


Figure B.4: r and d_{KS} distributions (left) and flasher likelihood distributions (Equation B.3, right) of flasher and neutrino events.

Figure B.4 illustrates the distinguishability of flasher events from neutrino events through two metrics capturing the similarity in hit patterns between two events. When flasher events are compared with a recurring event, they typically show a high value of r and a low value of d_{KS} — a scenario rarely observed when comparing two neutrino events.

To exploit this distinguishability, the probability density of obtaining r and d_{KS} for atmospheric neutrino events was established based on values calculated with MC-simulated atmospheric neutrino events. Utilizing this probability density, we can quantify how "unlikely" it is to obtain a specific set of r and d_{KS} if an event under test were a neutrino event. This is accomplished by defining the "unlikeliness" metric $L(r, d_{\mathrm{KS}})$:

$$L(r, d_{\mathrm{KS}}) = \frac{1}{2}\left[\frac{1}{10}\sum_{x\in r, d_{\mathrm{KS}}}\left(-2\sum_{i}^{\text{Top 10}}\ln p(x, q_i) - 2\sum_{i}^{10}\ln\left(\frac{i}{N^{\alpha_x}}\right)\right) + C\right] \quad \text{(B.3)}$$

Here, $p(r, Q)$ and $p(d_{\mathrm{KS}}, Q)$ denote the probability density of two atmospheric neutrino events having r and d_{KS}, respectively, when the average total charge of the two events is Q. Only the top 10 events with the largest r values and the smallest d_{KS} values are considered for a given event. The coefficients α_r and $\alpha_{d_{\mathrm{KS}}}$ and C are tuned heuristically to align the peak of the L values for MC-simulated atmospheric neutrino events to zero. Each event in the database is assigned a single value of L. Figure B.4 compares the distribution of L between atmospheric neutrinos (peaking close to zero) and PMT flashers (with larger values of L).

Events with $L < 3$ are selected, while events with $L > 3$ are considered flashers and subsequently rejected.

The event rate at this stage is approximately 18 events per day.

B.5 FC5: Additional cuts to reject remaining backgrounds

At this stage, additional intricate criteria are applied to further eliminate cosmic rays, flashers, and low-energy-cosmic-ray coincidences. Refer to [118] for details.

Another category of background events, "invisible muons," involves cosmic-ray muons losing energy below the Cherenkov threshold, becoming visible only when stopping in the ID to produce Michel electrons. To distinguish these events from low-energy electron neutrino events, we rely on OD hits preceding the ID event trigger, represented by the variable $\mathrm{NHITA_{early}}$. Specifically, invisible muon events are rejected using following complex data selection cuts:

- $\mathrm{NHITA_{early}} < 50$ hits, AND
- $\mathrm{PE_{total}} > 1{,}000$ p.e., OR

  Michel electrons have maximum energy of around 50 MeV, so that events with the total ID PMT charge ($\mathrm{PE_{total}}$) larger than 1,000 photoelectrons are very unlikely to be originated from Michel electrons.
- $\mathrm{NHITA_{early}} < 5$ hits, OR
- $d_{\mathrm{OD}} > 5$ m AND $\mathrm{NHITA_{early}} < 10$ hits, OR

  Here, d_{OD} is the distance between the position of the OD PMT hit cluster around the event trigger and the earlier time range of [-9, -0.2] μs from the trigger.
- $d_{\mathrm{OD}} < 5$ m AND $\mathrm{NHITA_{early}} + \mathrm{NHITA_{500}} < 10$ hits

  Here, $\mathrm{NHITA_{500}}$ is the number of OD PMT hits within [-0.2, +0.3] μs range from the event trigger.

The event rate at this stage is around 16 events per day.

B.6 FCFV: Events fully-contained in fiducial volume

Following are the standard cuts to select fully-contained events in fiducial volume (FCFV) and further reduce the following sources of background:

- **No particle exit from the ID**

  The number of OD PMT hits in the largest cluster (`NHITAC`, explained in Section 3.3.4), must not exceed 15.

- **Low-energy backgrounds**

  The visible energy must be greater than 30 MeV.

- **Background radioactivity near the tank wall**

  The reconstructed interaction vertex must be more than 1 m away from the ID tank wall.

  The event rate at this stage is around 9 events per day. Over 99% of the selected events are expected to be atmospheric neutrino interactions.

Appendix C

Comparison of atmospheric neutrino data: SK-IV with SK-VI

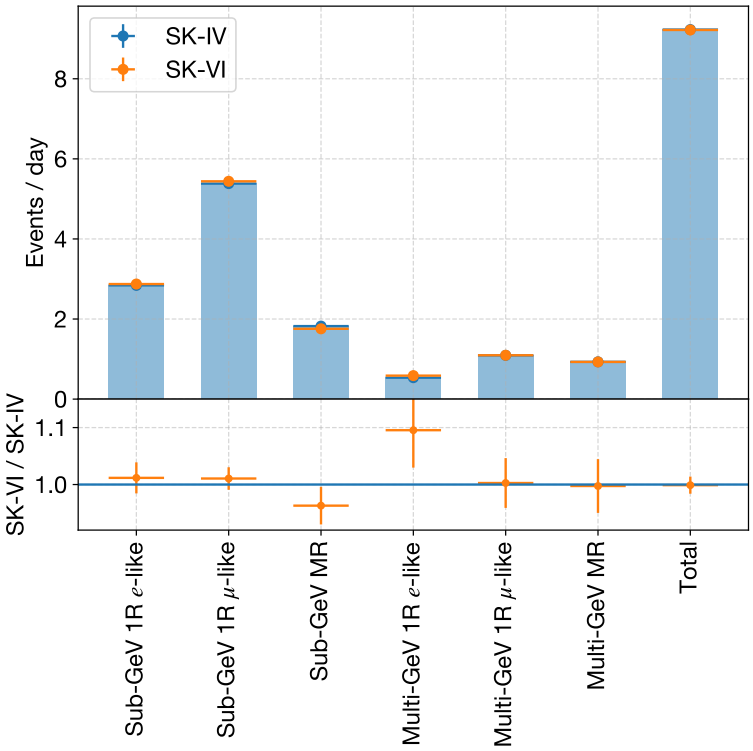


Figure C.1: Event rates.

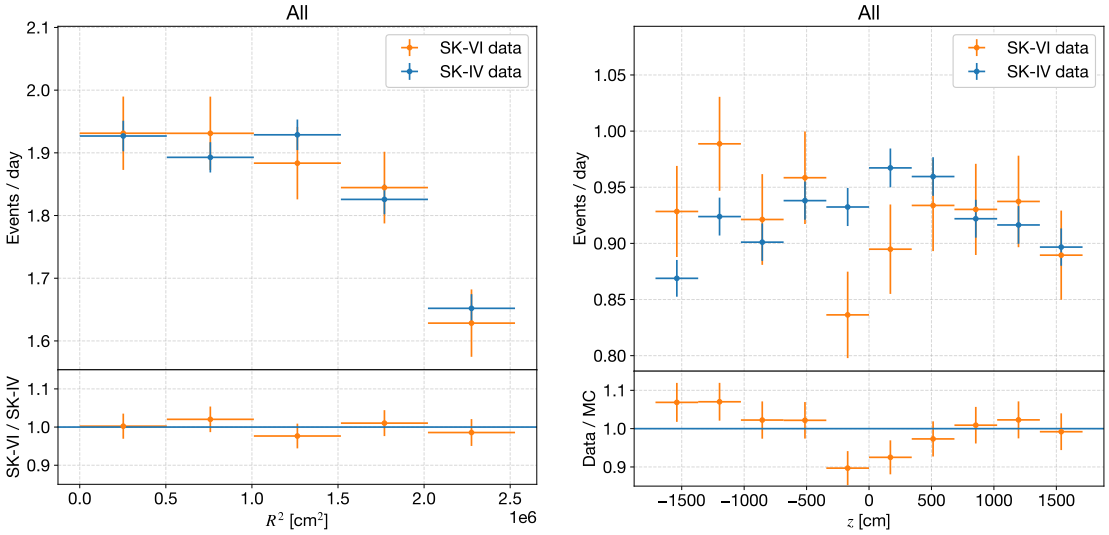


Figure C.2: Reconstructed vertex distributions R2 (left), z (right)

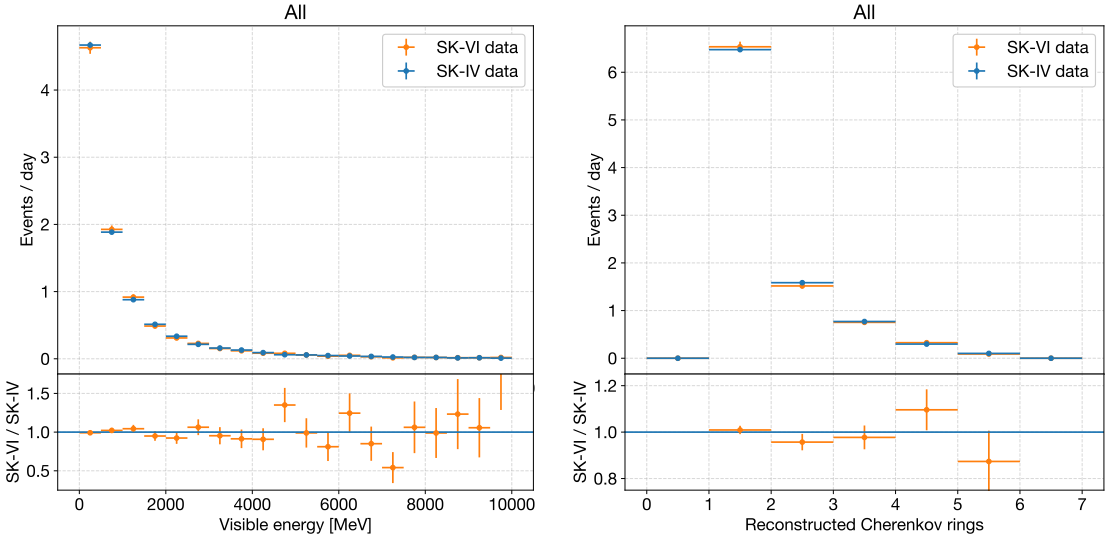


Figure C.3: Event visible energy (left) and number of reconstructed Cherenkov rings (right).

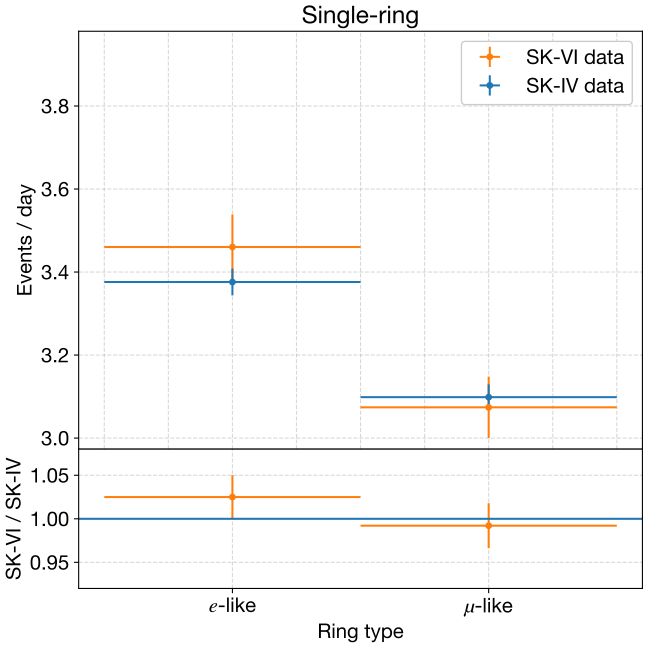


Figure C.4: Single-ring event visible energy (left) and ring PID (right).

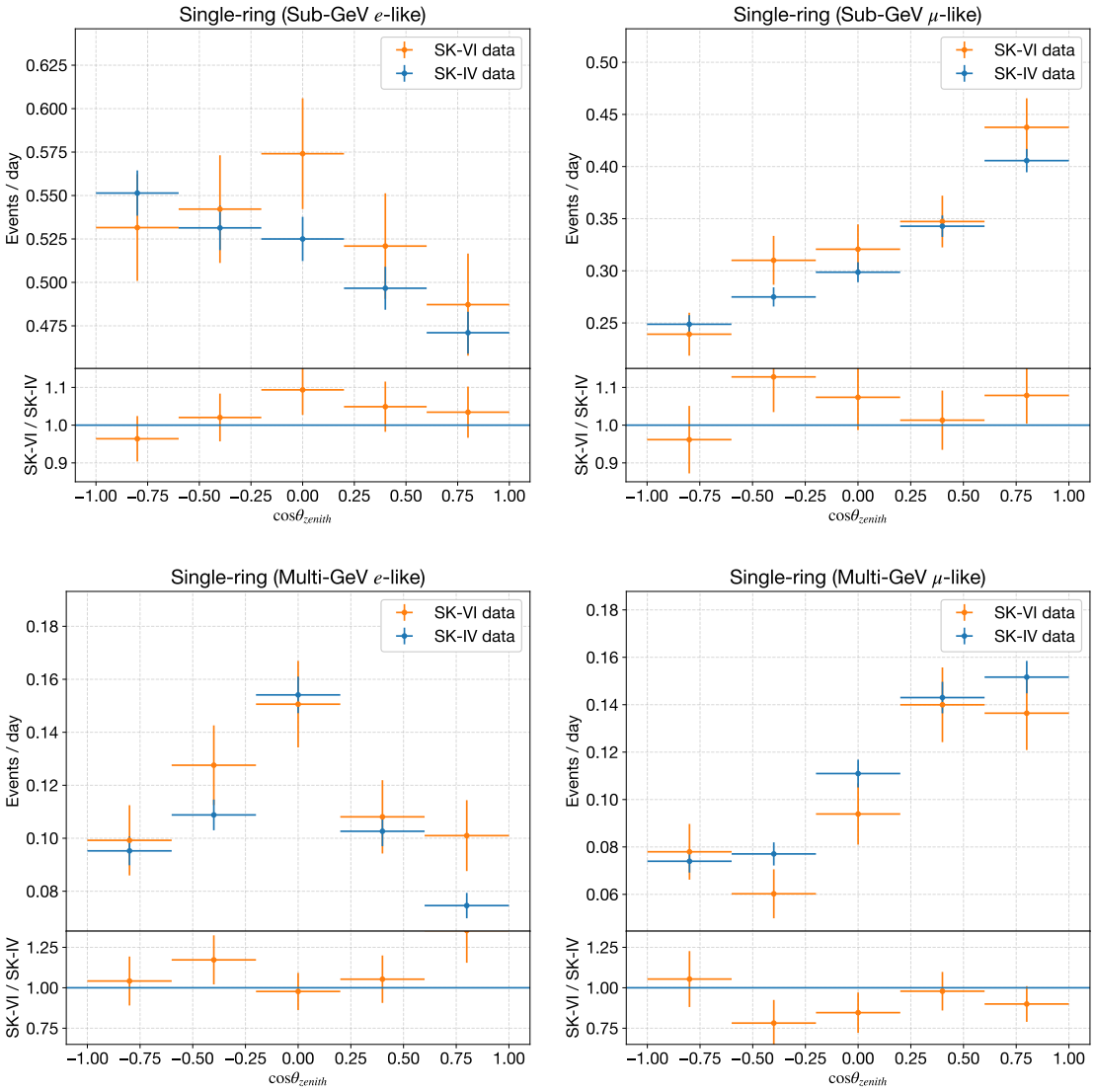


Figure C.5: Cosine zenith angle distributions

Appendix D

Statistical uncertainties

D.1 Statistical inference

Statistical uncertainty stems from N_i^{detected} — the number of detected signals in the i^{th} event — in Equation 4.1. We assume that N_i^{detected} follows a binomial distribution with N_i^{signal} trials and the binomial probability ϵ_i, where N_i^{signal} is the total number of produced signals, and ϵ_i is the signal efficiency of the i^{th} event.

If we assume that, for sufficiently small visible energy bins, both the number of produced signals (N^{signal}) and the signal efficiency (ϵ) remain nearly constant across various events within the bin, then the standard deviation of N^{detected} is:

$$\Delta N^{\text{detected}} \sim \sqrt{N^{\text{signal}}\epsilon(1-\epsilon)} \sim \sqrt{N^{\text{detected}}(1-\epsilon)} \tag{D.1}$$

Thus, the uncertainty in the average of $N^{\text{detected}}/\epsilon$ across all events in the bin is given by:

$$\delta\left\langle \frac{N^{\text{detected}}}{\epsilon} \right\rangle_{\text{event}} \sim \frac{1}{\sqrt{N_{\text{event}}}} \frac{\sqrt{N^{\text{detected}}(1-\epsilon)}}{\epsilon} \tag{D.2}$$

The reliability of the expression $\langle N^{\text{detected}}/\epsilon \rangle$ itself, as defined in Equation 4.1, is rooted in the Central Limit Theorem. This is further supported by Figure 4.31,

where the signal multiplicity estimated using Equation 4.1 aligns closely with the actual signal multiplicity in the simulation.

The statistical uncertainty, as given by Equation D.2, holds true even in the more realistic scenario where N_i^{signal} follows a Poisson distribution, and ϵ_i is sampled from a normal distribution for each i^{th} event.

Our event statistics are sufficiently large, so that we can easily adjust each visible energy bin to contain at least 20 neutrino events. Therefore, it is reasonable to use Equation D.2 as an estimate for the statistical uncertainty across all bins.

D.2 Signal detection performance spread

The spread or fluctuation of the signal detection performance around the mean is influenced by several factors. As previously discussed, the effect of these factors on the "average" signal multiplicity per event is mitigated by the factor of $1/\sqrt{N_{\text{event}}}$.

Detector modeling parameters

The temporal variations in the estimated signal efficiencies derived from Am/Be data were at a maximum of 9.5% for SK-IV, 9.4% for SK-V, and 4.6% for SK-VI.

MC regression

The systematic uncertainty in the efficiency spread is derived from the 1σ prediction intervals of the signal efficiency obtained through GAM regression. On average, the fraction of this interval relative to the GAM predictions was 6.2% for SK-IV, 7.5% for SK-V, and 2.3% for SK-VI. The 1σ prediction interval of the false positive rate was 40.7%, 50.4%, and 37.6% for SK-IV, SK-V, and SK-VI, respectively.

Signal vertex

The proximity of the signal vertex to the vertex used for photon TOF correction plays a crucial role in determining the efficiency of signal detection. There should be a certain level of uncertainty in our simulation regarding this distance. To assess the impact, we compared the reconstructed distance from the neutrino interaction vertex to the neutron capture vertex in both SK-VI (R) data and simulation, as illustrated in Figure 4.23. Understanding how the detection efficiency in the `INIT-NEUTRON` vertex mode is influenced by distance, as shown in Figure 3.15, allows us to extrapolate the differences in the distance distributions between data and MC simulation to the signal efficiency. The expected fractional changes were 3.5% for SK-IV, 4.0% for SK-V, and 2.3% for SK-VI.

Neural network credibility

The credibility of a neural network is inherently influenced by the diffuse boundary between class labels used during training, leading to inherent uncertainties in both the weights and outputs of the trained network. Consider a scenario where training data is randomly sampled from a larger dataset with a vague boundary between two binary classes, and a network of the same architecture is trained multiple times. In such cases, one can anticipate a distribution of weights for the network.

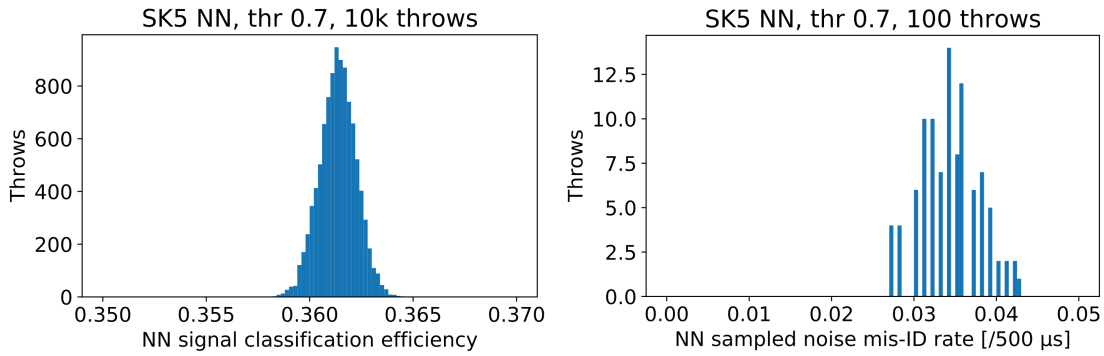


Figure D.1: Variance in dropout-sampled neural network signal classification efficiency (left) and false positive rate per 500 μs (right).

For each unseen candidate, sampling from this ensemble of trained networks produces a distribution of network outputs. Gal and Ghahramani [179] demonstrated that for networks trained with dropout, as used in this analysis, network outputs can be sampled by activating dropout during inference without the need to train network ensembles. Figure D.1 illustrates the distributions of signal efficiency and false positive rates sampled by activating dropout during inference.

The evaluated uncertainty in signal efficiency was 0.2%, 0.3%, and 0.2% for SK-IV, SK-V, and SK-VI, respectively, while the uncertainty in false positive rates was 4.5%, 4.8%, and 0.9% for the same SK phases.

Event reconstruction

The uncertainties associated with the reconstructed event variables used as input for the fitted GAM can influence the expected output performances. In this analysis, the 2% uncertainty in the visible energy scale was the only non-negligible factor impacting the final estimates.

For signal efficiency, the assigned uncertainties were 0.3%, 0.2%, and 0.3% for SK-IV, SK-V, and SK-VI, respectively. Also, the uncertainties for false positive rates were 1.2%, 1.1%, and 1.3% for the corresponding SK phases.

Muon-induced backgrounds

The quantity of Michel electrons and nuclear de-excitation gamma-rays resulting from negative muon capture on nuclei is dependent on the number of muons generated in an event. This quantity introduces a degree of uncertainty due to the variability in pion production, a precursor to muon generation. By analyzing the time distribution of both Michel electron and neutron capture candidates, we determined the maximum and minimum potential levels of muon-induced backgrounds. The uncertainty in false positives due to the uncertainty in muon-induced backgrounds was estimated to be 2.5% for SK-VI.

The total systematic uncertainty in the signal detection performance spread is summarized in Table D.1. The total uncertainty was scaled by the factor of $1/\sqrt{N_{\text{event}}}$ where N_{event} represents the number of neutrino events in each visible energy bin.

Source	**Signal efficiency**			**False positive rate**		
	SK-IV	**SK-V**	**SK-VI**	**SK-IV**	**SK-V**	**SK-VI**
Detector response	9.5%	9.4%	4.6%	40.7%	50.4%	37.6%
MC regression	6.2%	7.5%	2.3%			
Signal vertex	3.5%	4.0%	2.0%			
Neural network	0.7%	0.8%	0.3%	4.5%	4.8%	0.9%
Event reconstruction	0.2%	0.3%	0.2%	1.2%	1.1%	1.3%
μ-induced BG						2.5%
Total	11.9%	12.7%	5.5%	41.0%	50.6%	37.7%

Table D.1: Systematic uncertainties relevant to signal detection performance spread.

Appendix E

Comparison of neutrino event generators

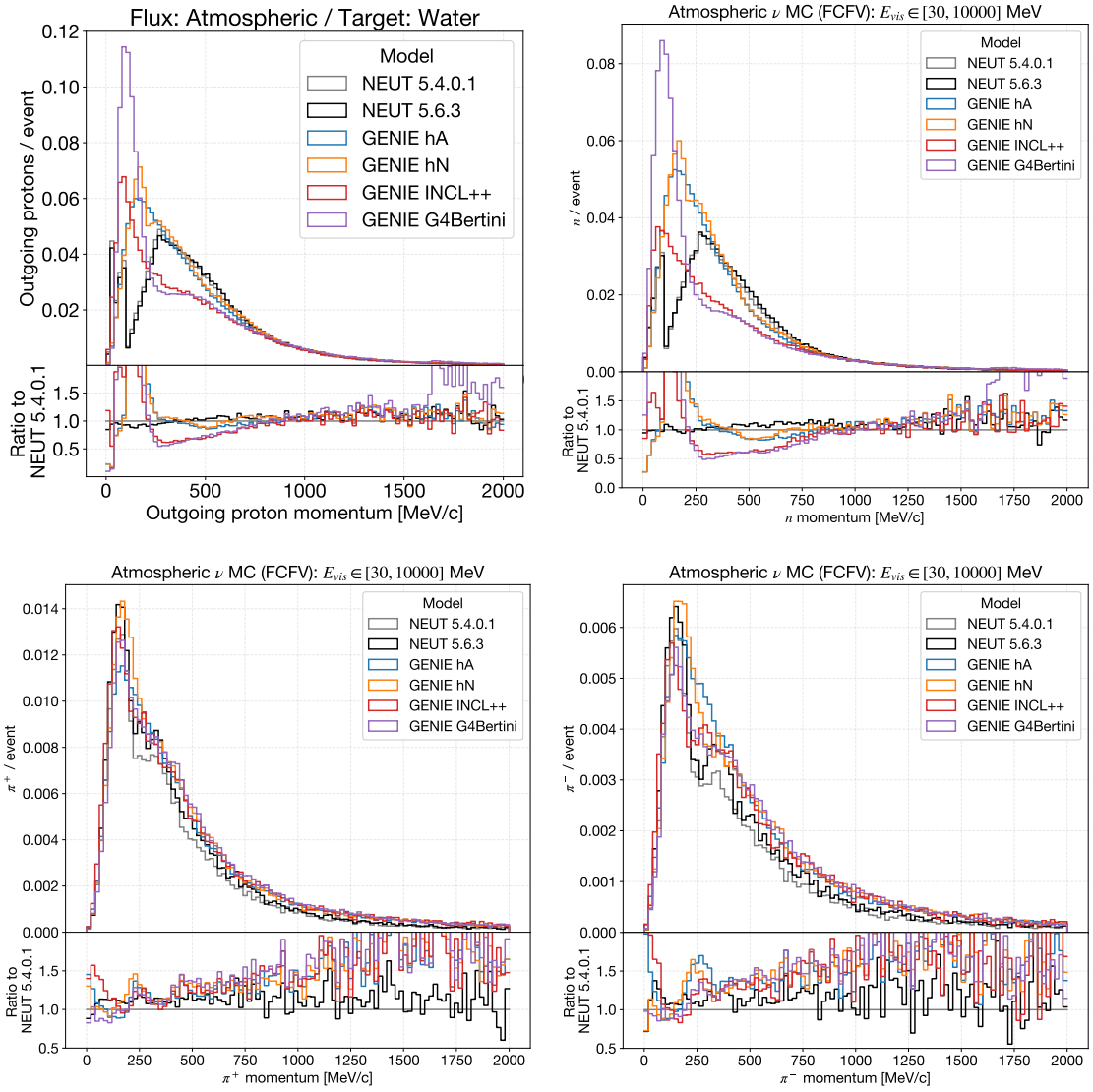


Figure E.1: Outgoing hadron multiplicity and kinematics comparison of different neutrino event generator options.

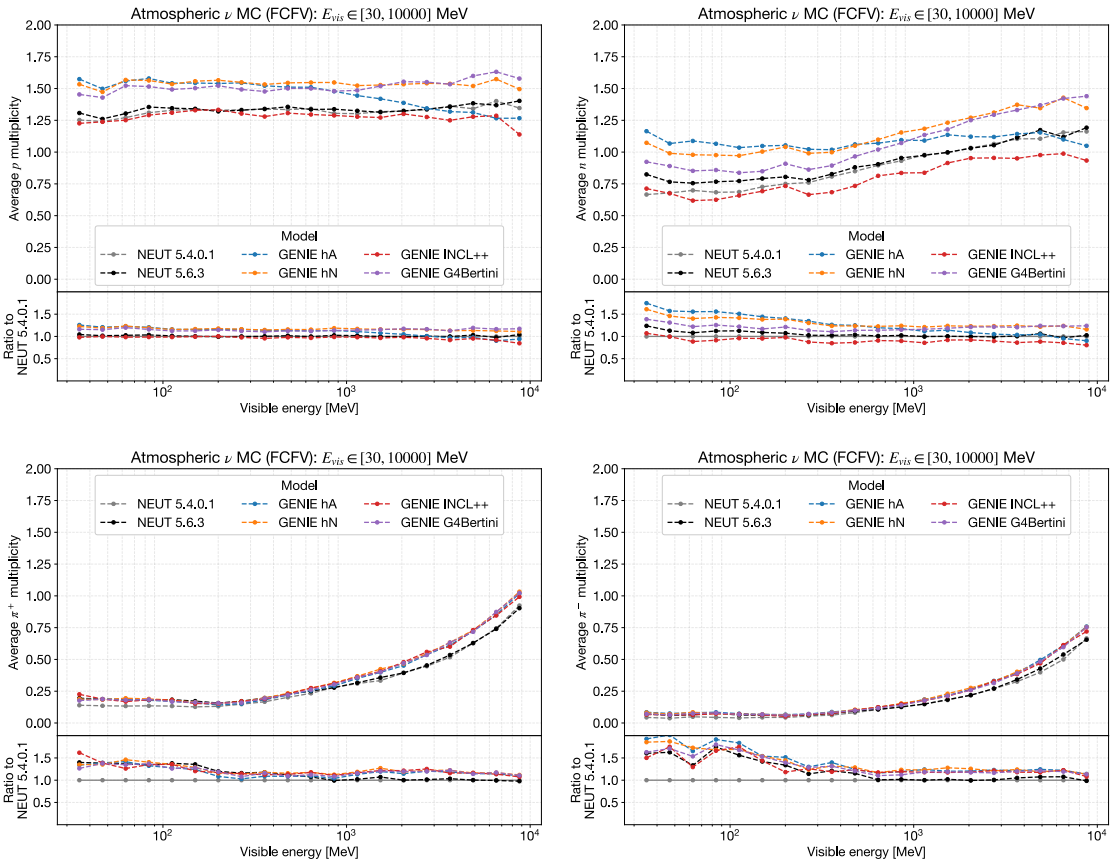


Figure E.2: Outgoing hadron multiplicity and kinematics comparison of different neutrino event generator options.

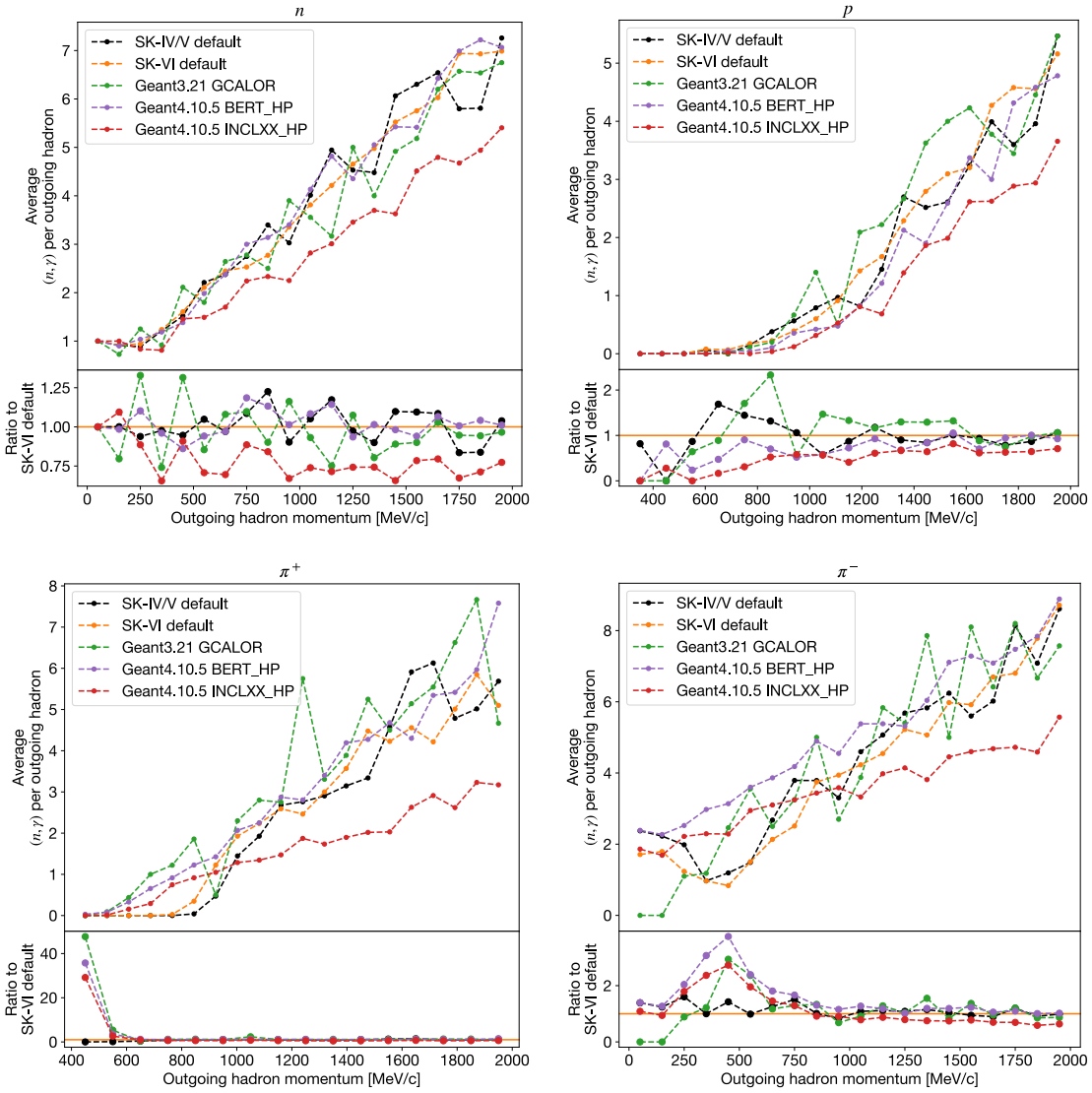


Figure E.3: Outgoing hadron multiplicity and kinematics comparison of different neutrino event generator options.

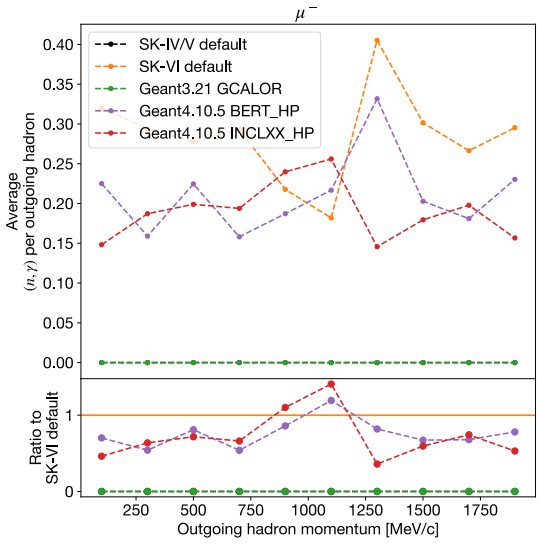


Figure E.4: Muon multiplicity

Bibliography

[1] E. Fermi. "Tentativo di una teoria dell'emissione dei raggi beta". In: *Ric. Sci.* 4 (1933), pp. 491–495.

[2] M. Goldhaber, L. Grodzins, and A. W. Sunyar. "Helicity of Neutrinos". In: *Phys. Rev.* 109 (3 Feb. 1958), pp. 1015–1017. DOI: 10.1103/PhysRev.109.1015. URL: https://link.aps.org/doi/10.1103/PhysRev.109.1015.

[3] C. S. Wu et al. "Experimental Test of Parity Conservation in Beta Decay". In: *Phys. Rev.* 105 (4 Feb. 1957), pp. 1413–1415. DOI: 10.1103/PhysRev.105.1413. URL: https://link.aps.org/doi/10.1103/PhysRev.105.1413.

[4] S. L. Glashow. "The renormalizability of vector meson interactions". In: *Nuclear Physics* 10 (1959), pp. 107–117. ISSN: 0029-5582. DOI: https://doi.org/10.1016/0029-5582(59)90196-8. URL: https://www.sciencedirect.com/science/article/pii/0029558259901968.

[5] A. Salam and J. C. Ward. "Weak and electromagnetic interactions". In: *Il Nuovo Cimento (1955-1965)* 11.4 (Feb. 1959), pp. 568–577. ISSN: 1827-6121. DOI: 10.1007/BF02726525. URL: https://doi.org/10.1007/BF02726525.

[6] S. Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (21 Nov. 1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264. URL: https://link.aps.org/doi/10.1103/PhysRevLett.19.1264.

[7] Y. Fukuda et al. "Evidence for Oscillation of Atmospheric Neutrinos". In: *Phys. Rev. Lett.* 81 (8 Aug. 1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562. URL: https://link.aps.org/doi/10.1103/PhysRevLett.81.1562.

[8] Q. R. Ahmad et al. "Measurement of the Rate of $\nu_e + d \rightarrow p + p + e^-$ Interactions Produced by ^8B Solar Neutrinos at the Sudbury Neutrino Observatory". In: *Phys. Rev. Lett.* 87 (7 July 2001), p. 071301. DOI: 10.1103/PhysRevLett.87.071301. URL: https://link.aps.org/doi/10.1103/PhysRevLett.87.071301.

[9] J. H. Christenson et al. "Evidence for the 2π Decay of the K_2^0 Meson". In: *Phys. Rev. Lett.* 13 (4 July 1964), pp. 138–140. DOI: 10.1103/PhysRevLett.13.138. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.138.

[10] K. Abe et al. "Observation of Large *CP* Violation in the Neutral B Meson System". In: *Phys. Rev. Lett.* 87 (9 Aug. 2001), p. 091802. DOI: 10.1103/PhysRevLett.87.091802. URL: https://link.aps.org/doi/10.1103/PhysRevLett.87.091802.

[11] N. Aghanim et al. "Planck2018 results: VI. Cosmological parameters". In: *Astronomy and Astrophysics* 641 (Sept. 2020), A6. ISSN: 1432-0746. DOI: 10.1051/0004-6361/201833910. URL: http://dx.doi.org/10.1051/0004-6361/201833910.

[12] F. Reines and C. Cowan. "The Neutrino". In: *Nature* 178.4531 (Sept. 1956), pp. 446–449. ISSN: 1476-4687. DOI: 10.1038/178446a0. URL: https://doi.org/10.1038/178446a0.

[13] G. Danby et al. "Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos". In: *Phys. Rev. Lett.* 9 (1 July 1962), pp. 36–44. DOI: 10.1103/PhysRevLett.9.36. URL: https://link.aps.org/doi/10.1103/PhysRevLett.9.36.

[14] K. Kodama et al. "Observation of tau neutrino interactions". In: *Physics Letters B* 504.3 (2001), pp. 218–224. ISSN: 0370-2693. DOI: https://doi.org/10.1016/S0370-2693(01)00307-0. URL: https://www.sciencedirect.com/science/article/pii/S0370269301003070.

[15] E. K. Akhmedov. "Do charged leptons oscillate?" In: *Journal of High Energy Physics* 2007.09 (Sept. 2007), pp. 116–116. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2007/09/116. URL: http://dx.doi.org/10.1088/1126-6708/2007/09/116.

[16] F. P. An et al. “Observation of Electron-Antineutrino Disappearance at Daya Bay”. In: *Phys. Rev. Lett.* 108 (17 Apr. 2012), p. 171803. DOI: 10.1103/PhysRevLett.108.171803. URL: https://link.aps.org/doi/10.1103/PhysRevLett.108.171803.

[17] K. Abe et al. “Indication of Electron Neutrino Appearance from an Accelerator-Produced Off-Axis Muon Neutrino Beam”. In: *Phys. Rev. Lett.* 107 (4 July 2011), p. 041801. DOI: 10.1103/PhysRevLett.107.041801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.107.041801.

[18] M. Prado Rodriguez. “Neutrino Mass Ordering with IceCube DeepCore”. In: *Physical Sciences Forum* 8.1 (2023). ISSN: 2673-9984. DOI: 10.3390/psf2023008007. URL: https://www.mdpi.com/2673-9984/8/1/7.

[19] I. Esteban et al. “The fate of hints: updated global analysis of three-flavor neutrino oscillations”. In: *Journal of High Energy Physics* 2020.9 (Sept. 2020). ISSN: 1029-8479. DOI: 10.1007/jhep09(2020)178. URL: http://dx.doi.org/10.1007/JHEP09(2020)178.

[20] L. Wolfenstein. “Neutrino oscillations in matter”. In: *Phys. Rev. D* 17 (9 May 1978), pp. 2369–2374. DOI: 10.1103/PhysRevD.17.2369. URL: https://link.aps.org/doi/10.1103/PhysRevD.17.2369.

[21] S. P. Mikheyev and A. Y. Smirnov. “Resonant amplification of ν oscillations in matter and solar-neutrino spectroscopy”. In: *Il Nuovo Cimento C* 9.1 (Jan. 1986), pp. 17–26. ISSN: 0390-5551. DOI: 10.1007/BF02508049. URL: https://doi.org/10.1007/BF02508049.

[22] J. Arafune, M. Koike, and J. Sato. “CP violation and matter effect in long baseline neutrino oscillation experiments”. In: *Phys. Rev. D* 56 (5 Sept. 1997), pp. 3093–3099. DOI: 10.1103/PhysRevD.56.3093. URL: https://link.aps.org/doi/10.1103/PhysRevD.56.3093.

[23] K. Abe et al. “Measurements of neutrino oscillation parameters from the T2K experiment using 3.6×10^{21} protons on target”. In: *Eur. Phys. J. C* 83.9 (2023), p. 782. DOI: 10.1140/epjc/s10052-023-11819-x. URL: https://doi.org/10.1140/epjc/s10052-023-11819-x.

[24] J. Xia. "T2K-Super-Kamiokande Joint Neutrino Oscillation Sensitivity". Neutrino Oscillation Workshop (NOW). 2022. URL: https://agenda.infn.it/event/30418/contributions/170689/attachments/95729/131800/NOW2022_SKT2K_09.05.22.pdf.

[25] K. Abe et al. "Atmospheric neutrino oscillation analysis with external constraints in Super-Kamiokande I-IV". In: *Phys. Rev. D* 97 (7 Apr. 2018), p. 072001. DOI: 10.1103/PhysRevD.97.072001. URL: https://link.aps.org/doi/10.1103/PhysRevD.97.072001.

[26] Y. Okajima. "Development of Neutrino Energy Reconstruction Algorithm for Neutrino Mass Hierarchy Determination". PhD thesis. Tokyo Institute of Technology, 2020.

[27] P. Fernández Menéndez. "Neutrino Physics in Present and Future Kamioka Water-Čerenkov Detectors with Neutron Tagging". PhD thesis. University Autonomous of Madrid, 2017.

[28] S. Miki. "Neutrino event reconstruction algorithm with neutron detection in SK-Gd". MA thesis. The University of Tokyo, 2022.

[29] S. Agostinelli et al. "Geant4—a simulation toolkit". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(03)01368-8. URL: https://www.sciencedirect.com/science/article/pii/S0168900203013688.

[30] E. Mendoza and D. Cano-Ott. *Update of the Evaluated Neutron Cross Section Libraries for the Geant4 Code (see also INDC(NDS)-0612)*. Tech. rep. NUCLEAR PHYSICS AND RADIATION PHYSICS. International Atomic Energy Agency (IAEA), 2018, p. 34. URL: http://inis.iaea.org/search/search.aspx?orig_q=RN:50006622.

[31] M. B. Chadwick et al. "ENDF/B-VII.1 Nuclear Data for Science and Technology: Cross Sections, Covariances, Fission Product Yields and Decay Data". In: *Nuclear Data Sheets* 112.12 (2011). Special Issue on ENDF/B-VII.1 Library, pp. 2887–2996. ISSN: 0090-3752. DOI: 10.1016/j.nds.2011.11.

002. URL: https://www.sciencedirect.com/science/article/pii/S009037521100113X.

[32] H. Zhang. "Neutron tagging and its application to physics in Super-Kamiokande IV". In: *32nd International Cosmic Ray Conference*. Vol. 4. 2011, p. 71. DOI: 10.7529/ICRC2011/V04/0353.

[33] B. Aharmim et al. "Measurement of neutron production in atmospheric neutrino interactions at the Sudbury Neutrino Observatory". In: *Physical Review D* 99.11 (June 2019). ISSN: 2470-0029. DOI: 10.1103/physrevd.99.112007. URL: http://dx.doi.org/10.1103/PhysRevD.99.112007.

[34] R. Akutsu. "A Study of Neutrons Associated with Neutrino and Antineutrino Interactions on the Water Target at the T2K Far Detector". PhD thesis. The University of Tokyo, 2019.

[35] M. Elkins et al. "Neutron measurements from antineutrino hydrocarbon reactions". In: *Phys. Rev. D* 100 (5 Sept. 2019), p. 052002. DOI: 10.1103/PhysRevD.100.052002. URL: https://link.aps.org/doi/10.1103/PhysRevD.100.052002.

[36] A. Olivier et al. *Measurement of the Multi-Neutron $\bar{\nu}_\mu$ Charged Current Differential Cross Section at Low Available Energy on Hydrocarbon*. Tech. rep. 25 pages, 11 figures; Added ancillary files with cross section values as .csv Matches preprint accepted by publisher. 2023. arXiv: 2310.17014. URL: https://cds.cern.ch/record/2878377.

[37] T. Kajita. "Atmospheric Neutrinos". In: *Advances in High Energy Physics* 2012 (Nov. 2012), p. 504715. ISSN: 1687-7357. DOI: 10.1155/2012/504715. URL: https://doi.org/10.1155/2012/504715.

[38] E. Richard et al. "Measurements of the atmospheric neutrino flux by Super-Kamiokande: Energy spectra, geomagnetic effects, and solar modulation". In: *Physical Review D* 94.5 (Sept. 2016). ISSN: 2470-0029. DOI: 10.1103/physrevd.94.052001. URL: http://dx.doi.org/10.1103/PhysRevD.94.052001.

[39] Y. Suzuki. "The Super-Kamiokande experiment". In: *The European Physical Journal C* 79.4 (Apr. 2019), p. 298. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-6796-2. URL: https://doi.org/10.1140/epjc/s10052-019-6796-2.

[40] Y. Fukuda et al. "The Super-Kamiokande detector". In: *Nucl. Instrum. Meth. A* 501 (2003). Ed. by V. A. Ilyin, V. V. Korenkov, and D. Perret-Gallix, pp. 418–462. DOI: 10.1016/S0168-9002(03)00425-X.

[41] T. Lou. "Development of 50 cm Diameter Hybrid Photo-Detector for Hyper-Kamiokande". MA thesis. The University of Tokyo, 2017.

[42] J. D. Jackson. *Classical electrodynamics*. Third edition. New York : Wiley, [1999] ©1999. URL: https://search.library.wisc.edu/catalog/999849741702121.

[43] K. Abe et al. "Constraint on the matter–antimatter symmetry-violating phase in neutrino oscillations". In: *Nature* 580.7803 (2020), pp. 339–344. ISSN: 1476-4687. DOI: 10.1038/s41586-020-2177-0. URL: https://doi.org/10.1038/s41586-020-2177-0.

[44] K. Okumura et al. *Upgrade of the Super-Kamiokande data acquisition system (in Japanese)*. May 2009.

[45] H. Nishino et al. "High-speed charge-to-time converter ASIC for the Super-Kamiokande detector". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 610.3 (Nov. 2009), pp. 710–717. ISSN: 0168-9002. DOI: 10.1016/j.nima.2009.09.026. URL: http://dx.doi.org/10.1016/j.nima.2009.09.026.

[46] K. Abe et al. "Calibration of the Super-Kamiokande detector". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 737 (Feb. 2014), pp. 253–272. ISSN: 0168-9002. DOI: 10.1016/j.nima.2013.11.081. URL: http://dx.doi.org/10.1016/j.nima.2013.11.081.

[47] S. Yamada et al. “Commissioning of the New Electronics and Online System for the Super-Kamiokande Experiment”. In: *IEEE Transactions on Nuclear Science* 57.2 (2010), pp. 428–432. DOI: 10.1109/TNS.2009.2034854.

[48] Kamiokande Collaboration et al. *Atmospheric neutrino oscillation analysis with neutron tagging and an expanded fiducial volume in Super-Kamiokande I-V*. 2023. arXiv: 2311.05105 [hep-ex].

[49] M. Nakahata et al. “Calibration of Super-Kamiokande using an electron LINAC”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 421.1–2 (Jan. 1999), pp. 113–129. ISSN: 0168-9002. DOI: 10.1016/s0168-9002(98)01200-5. URL: http://dx.doi.org/10.1016/S0168-9002(98)01200-5.

[50] Y. Noguchi. *ATMPD-related tunings for SKG4*. May 2023.

[51] K. Abe et al. “First gadolinium loading to Super-Kamiokande”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1027 (Mar. 2022), p. 166248. DOI: 10.1016/j.nima.2021.166248. URL: https://doi.org/10.1016%2Fj.nima.2021.166248.

[52] Super-Kamiokande Collaboration. *Introduction of Gadolinium into Super-Kamiokande and the Start of New Observations*. Aug. 21, 2020. URL: http://www-sk.icrr.u-tokyo.ac.jp/sk/news/2020/08/sk-gd-detail-e.html.

[53] L. Martí Magro. *Evaluating gadolinium for use in Super-Kamiokande*. 2011.

[54] M. A. Shea and D. F. Smart. “Geomagnetic cutoff rigidities and geomagnetic coordinates appropriate for the Carrington flare Epoch”. In: *Advances in Space Research* 38.2 (2006). The Great Historical Geomagnetic Storm of 1859: A Modern Look, pp. 209–214. ISSN: 0273-1177. DOI: https://doi.org/10.1016/j.asr.2005.03.156. URL: https://www.sciencedirect.com/science/article/pii/S0273117705008410.

[55] M. Honda et al. “Improvement of low energy atmospheric neutrino flux calculation using the JAM nuclear interaction model”. In: *Physical Review D* 83.12 (June 2011). DOI: 10.1103/physrevd.83.123001. URL: https://doi.org/10.1103%2Fphysrevd.83.123001.

[56] J. Alcaraz et al. “Cosmic protons”. In: *Physics Letters B* 490.1 (2000), pp. 27–35. ISSN: 0370-2693. DOI: https://doi.org/10.1016/S0370-2693(00)00970-9. URL: https://www.sciencedirect.com/science/article/pii/S0370269300009709.

[57] J. Alcaraz et al. “Helium in near Earth orbit”. In: *Physics Letters B* 494.3 (2000), pp. 193–202. ISSN: 0370-2693. DOI: https://doi.org/10.1016/S0370-2693(00)01193-X. URL: https://www.sciencedirect.com/science/article/pii/S037026930001193X.

[58] T. Sanuki et al. “Precise Measurement of Cosmic-Ray Proton and Helium Spectra with the BESS Spectrometer”. In: *The Astrophysical Journal* 545.2 (Dec. 2000), pp. 1135–1142. ISSN: 1538-4357. DOI: 10.1086/317873. URL: http://dx.doi.org/10.1086/317873.

[59] S. Haino et al. “Measurements of primary and atmospheric cosmic-ray spectra with the BESS-TeV spectrometer”. In: *Physics Letters B* 594.1–2 (July 2004), pp. 35–46. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2004.05.019. URL: http://dx.doi.org/10.1016/j.physletb.2004.05.019.

[60] URL: http://nssdc.gsfc.nasa.gov/space/model/atmos/us_standard.html.

[61] URL: http://www.ngdc.noaa.gov/IAGA/vmod/igrf.html.

[62] K. Niita et al. “PHITS: A particle and heavy ion transport code system”. In: *Radiat. Meas.* 41 (2006), pp. 1080–1090. DOI: 10.1016/j.radmeas.2006.07.013.

[63] S. Roesler, R. Engel, and J. Ranft. “The Monte Carlo Event Generator DPMJET-III”. In: *Advanced Monte Carlo for Radiation Physics, Particle Transport Simulation and Applications*. Springer Berlin Heidelberg, 2001, pp. 1033–1038. ISBN: 9783642182112. DOI: 10.1007/978-3-642-18211-2_166. URL: http://dx.doi.org/10.1007/978-3-642-18211-2_166.

[64] M. Honda et al. "Calculation of atmospheric neutrino flux using the interaction model calibrated with atmospheric muon data". In: *Phys. Rev. D* 75 (4 Feb. 2007), p. 043006. DOI: 10.1103/PhysRevD.75.043006. URL: https://link.aps.org/doi/10.1103/PhysRevD.75.043006.

[65] Neutron Monitor Database. "Newark/Swarthmore, USA/Bartol Neutron Monitor Database". URL: http://www01.nmdb.eu/station/newk/.

[66] T. Van Cuyck. "Short-range correlations and meson-exchange currents in neutrino-nucleus scattering". In: 2017. URL: https://api.semanticscholar.org/CorpusID:126395148.

[67] Y. Koshio. "Observation of Atmospheric Neutrinos". In: *Universe* 6.6 (2020). ISSN: 2218-1997. DOI: 10.3390/universe6060080. URL: https://www.mdpi.com/2218-1997/6/6/80.

[68] H. De Vries, C.W. De Jager, and C. De Vries. "Nuclear charge-density-distribution parameters from elastic electron scattering". In: *Atomic Data and Nuclear Data Tables* 36.3 (1987), pp. 495–536. ISSN: 0092-640X. DOI: https://doi.org/10.1016/0092-640X(87)90013-1. URL: https://www.sciencedirect.com/science/article/pii/0092640X87900131.

[69] T. Golan. "Modeling nuclear effects in NuWro Monte Carlo neutrino event generator". PhD thesis. University of Wroclaw, 2014.

[70] S. Abe. "First measurement of the strangeness axial coupling constant using neutral current quasi-elastic interactions of atmospheric neutinos at KamLAND". PhD thesis. Tohoku University, 2022.

[71] C. H. Llewellyn Smith. "Neutrino reactions at accelerator energies". In: *Physics Reports* 3.5 (1972), pp. 261–379. ISSN: 0370-1573. DOI: https://doi.org/10.1016/0370-1573(72)90010-5. URL: https://www.sciencedirect.com/science/article/pii/0370157372900105.

[72] R. Gran et al. "Neutrino-nucleus quasi-elastic and 2p2h interactions up to 10 GeV". In: *Phys. Rev. D* 88 (11 Dec. 2013), p. 113007. DOI: 10.1103/PhysRevD.88.113007. URL: https://link.aps.org/doi/10.1103/PhysRevD.88.113007.

[73] R. Bradford et al. "A New Parameterization of the Nucleon Elastic Form Factors". In: *Nuclear Physics B - Proceedings Supplements* 159 (2006). Proceedings of the 4th International Workshop on Neutrino-Nucleus Interactions in the Few-GeV Region, pp. 127–132. ISSN: 0920-5632. DOI: https://doi.org/10.1016/j.nuclphysbps.2006.08.028. URL: https://www.sciencedirect.com/science/article/pii/S0920563206005184.

[74] I. D. Kakorin, K. S. Kuzmin, and V. A. Naumov. "Running axial mass of the nucleon as a phenomenological tool for calculating quasielastic neutrino–nucleus cross sections". In: *The European Physical Journal C* 81.12 (Dec. 2021), p. 1142. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-021-09945-5. URL: https://doi.org/10.1140/epjc/s10052-021-09945-5.

[75] D. Rein and M. Sehgal. "Neutrino-excitation of baryon resonances and single pion production". In: *Annals of Physics* 133.1 (1981), pp. 79–153. ISSN: 0003-4916. DOI: https://doi.org/10.1016/0003-4916(81)90242-6. URL: https://www.sciencedirect.com/science/article/pii/0003491681902426.

[76] Ch. Berger and L. M. Sehgal. "Lepton mass effects in single pion production by neutrinos". In: *Phys. Rev. D* 76 (11 Dec. 2007), p. 113004. DOI: 10.1103/PhysRevD.76.113004. URL: https://link.aps.org/doi/10.1103/PhysRevD.76.113004.

[77] K. M. Graczyk and J. T. Sobczyk. "Form factors in the quark resonance model". In: *Phys. Rev. D* 77 (5 Mar. 2008), p. 053001. DOI: 10.1103/PhysRevD.77.053001. URL: https://link.aps.org/doi/10.1103/PhysRevD.77.053001.

[78] K. M. Graczyk et al. "C_5^A axial form factor from bubble chamber experiments". In: *Phys. Rev. D* 80 (9 Nov. 2009), p. 093001. DOI: 10.1103/PhysRevD.80.093001. URL: https://link.aps.org/doi/10.1103/PhysRevD.80.093001.

[79] C. Wilkinson et al. "Reanalysis of bubble chamber measurements of muon-neutrino induced single pion production". In: *Phys. Rev. D* 90 (11 Dec. 2014),

p. 112017. DOI: 10.1103/PhysRevD.90.112017. URL: https://link.aps.org/doi/10.1103/PhysRevD.90.112017.

[80] Ch. Berger and L. M. Sehgal. "Partially conserved axial vector current and coherent pion production by low energy neutrinos". In: *Phys. Rev. D* 79 (5 Mar. 2009), p. 053003. DOI: 10.1103/PhysRevD.79.053003. URL: https://link.aps.org/doi/10.1103/PhysRevD.79.053003.

[81] A. Higuera et al. "Measurement of Coherent Production of $\pi^{\pm}$ in Neutrino and Antineutrino Beams on Carbon from E_{ν} of 1.5 to 20 GeV". In: *Phys. Rev. Lett.* 113 (26 Dec. 2014), p. 261802. DOI: 10.1103/PhysRevLett.113.261802. URL: https://link.aps.org/doi/10.1103/PhysRevLett.113.261802.

[82] S. Kretzer and M. H. Reno. "Tau neutrino deep inelastic charged current interactions". In: *Physical Review D* 66.11 (Dec. 2002). ISSN: 1089-4918. DOI: 10.1103/physrevd.66.113007. URL: http://dx.doi.org/10.1103/PhysRevD.66.113007.

[83] C. G. Callan and David J. Gross. "High-Energy Electroproduction and the Constitution of the Electric Current". In: *Phys. Rev. Lett.* 22 (4 Jan. 1969), pp. 156–159. DOI: 10.1103/PhysRevLett.22.156. URL: https://link.aps.org/doi/10.1103/PhysRevLett.22.156.

[84] C. H. Albright and C. Jarlskog. "Neutrino production of M+ and E+ heavy leptons (I)". In: *Nuclear Physics B* 84.2 (1975), pp. 467–492. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(75)90318-1. URL: https://www.sciencedirect.com/science/article/pii/0550321375903181.

[85] M. Glück, E. Reya, and A. Vogt. "Dynamical parton distributions revisited". In: *The European Physical Journal C* 5.3 (Sept. 1998), pp. 461–470. ISSN: 1434-6052. DOI: 10.1007/s100529800978. URL: http://dx.doi.org/10.1007/s100529800978.

[86] A. Bodek, I. Park, and U. K. Yang. "Improved low Q2 model for neutrino and electron nucleon cross sections in few GeV region". In: *Nuclear Physics B - Proceedings Supplements* 139 (Feb. 2005), pp. 113–118. ISSN: 0920-5632.

DOI: 10.1016/j.nuclphysbps.2004.11.208. URL: http://dx.doi.org/10.1016/j.nuclphysbps.2004.11.208.

[87] A. Bodek. "A Unified Model for inelastic e-N and ν-N cross sections at all Q2". In: *AIP Conference Proceedings*. AIP, 2005. DOI: 10.1063/1.2122031. URL: http://dx.doi.org/10.1063/1.2122031.

[88] C. Bronner and M. Hartz. "Tuning of the Charged Hadrons Multiplicities for Deep Inelastic Interactions in NEUT". In: *Proceedings of the 10th International Workshop on Neutrino-Nucleus Interactions in Few-GeV Region (NuInt15)*. Journal of the Physical Society of Japan, Dec. 2016. DOI: 10.7566/jpscp.12.010041. URL: http://dx.doi.org/10.7566/JPSCP.12.010041.

[89] T. Sjostrand. *PYTHIA 5.7 and JETSET 7.4: Physics and manual*. Tech. rep. Long version of publication in Comput.Phys.Commun. Geneva: CERN, 1993. URL: https://cds.cern.ch/record/287231.

[90] Y. Hayato. "NEUT Status, plan and new demands". 2018. URL: https://indico.cern.ch/event/703880/contributions/3157411/attachments/1734393/2804568/20181016_NEUT.pdf.

[91] V. E. Bunakov and G. V. Matvejev. "The physical and mathematical foundations of the intranuclear cascade model algorithm". In: *Zeitschrift für Physik A Atoms and Nuclei* 322.3 (Sept. 1985), pp. 511–521. ISSN: 0939-7922. DOI: 10.1007/BF01412089. URL: https://doi.org/10.1007/BF01412089.

[92] L.L. Salcedo et al. "Computer simulation of inclusive pion nuclear reactions". In: *Nuclear Physics A* 484.3 (1988), pp. 557–592. ISSN: 0375-9474. DOI: https://doi.org/10.1016/0375-9474(88)90310-7. URL: https://www.sciencedirect.com/science/article/pii/0375947488903107.

[93] G. Rowe, M. Salomon, and R. H. Landau. "Energy-dependent phase shift analysis of pion-nucleon scattering below 400 MeV". In: *Phys. Rev. C* 18 (1 July 1978), pp. 584–589. DOI: 10.1103/PhysRevC.18.584. URL: https://link.aps.org/doi/10.1103/PhysRevC.18.584.

[94] R. Seki, K. Masutani, and K. Yazaki. "Unified analysis of pionic atoms and low-energy pion-nuclear scattering: Hybrid analysis". In: *Phys. Rev. C* 27 (6 June 1983), pp. 2817–2832. DOI: 10.1103/PhysRevC.27.2817. URL: https://link.aps.org/doi/10.1103/PhysRevC.27.2817.

[95] H. W. Bertini. "Low-Energy Intranuclear Cascade Calculation". In: *Phys. Rev.* 131 (4 Aug. 1963), pp. 1801–1821. DOI: 10.1103/PhysRev.131.1801. URL: https://link.aps.org/doi/10.1103/PhysRev.131.1801.

[96] C. W. de Jager, H. J. de Vries, and C. P. de Vries. "Nuclear charge and magnetization density distribution parameters from elastic electron scattering". In: *Atomic Data and Nuclear Data Tables* 14 (1974), pp. 479–508. URL: https://api.semanticscholar.org/CorpusID:123291324.

[97] J. Whitmore. "Multiparticle production in the Fermilab bubble chambers". In: *Physics Reports* 27.5 (1976), pp. 187–273. ISSN: 0370-1573. DOI: https://doi.org/10.1016/0370-1573(76)90004-1. URL: https://www.sciencedirect.com/science/article/pii/0370157376900041.

[98] D. Rowntree et al. "π^+ absorption on N and Ar". In: *Phys. Rev. C* 60 (5 Oct. 1999), p. 054610. DOI: 10.1103/PhysRevC.60.054610. URL: https://link.aps.org/doi/10.1103/PhysRevC.60.054610.

[99] B. G. Ritchie. "Parametrization of total and differential cross sections for πd→pp below 1 GeV". In: *Phys. Rev. C* 44 (1 July 1991), pp. 533–536. DOI: 10.1103/PhysRevC.44.533. URL: https://link.aps.org/doi/10.1103/PhysRevC.44.533.

[100] E. S. Pinzon Guerra et al. "Using world $\pi^{\pm}$-nucleus scattering data to constrain an intranuclear cascade model". In: *Phys. Rev. D* 99 (5 Mar. 2019), p. 052007. DOI: 10.1103/PhysRevD.99.052007. URL: https://link.aps.org/doi/10.1103/PhysRevD.99.052007.

[101] H. W. Bertini. "Nonelastic Interactions of Nucleons and π Mesons with Complex Nuclei at Energies Below 3 GeV". In: *Phys. Rev. C* 6 (2 Aug. 1972), pp. 631–659. DOI: 10.1103/PhysRevC.6.631. URL: https://link.aps.org/doi/10.1103/PhysRevC.6.631.

[102] W. Y. Ma et al. "Current status of final-state interaction models and their impact on neutrino-nucleus interactions". In: *Journal of Physics: Conference Series* 888.1 (Sept. 2017), p. 012171. DOI: 10.1088/1742-6596/888/1/012171. URL: https://dx.doi.org/10.1088/1742-6596/888/1/012171.

[103] G. Mitsuka. "Study of Non-Standard Neutrino Interactions with Atmospheric Neutrino Data in Super-Kamiokande". PhD thesis. The University of Tokyo, 2009.

[104] S. M. Eliseev. "Deep inelastic scattering of leptons on nuclei: Hadron formation, cumulative particles production". In: *Refereed and selected contributions from International Conference on Quark Nuclear Physics*. Ed. by Charlotte Elster, Josef Speth, and Thomas Walcher. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 433–435. ISBN: 978-3-662-09712-0.

[105] D. S. Baranov et al. "Estimation of the hadron formation length in neutrino interactions". In: *Sov. J. Nucl. Phys. (Engl. Transl.); (United States)* 40:6 (Dec. 1984). URL: https://www.osti.gov/biblio/5600893.

[106] A. M. Ankowski et al. "Analysis of γ-Ray Production in Neutral-Current Neutrino-Oxygen Interactions at Energies above 200 MeV". In: *Phys. Rev. Lett.* 108 (5 Feb. 2012), p. 052505. DOI: 10.1103/PhysRevLett.108.052505. URL: https://link.aps.org/doi/10.1103/PhysRevLett.108.052505.

[107] H. Ejiri. "Nuclear deexcitations of nucleon holes associated with nucleon decays in nuclei". In: *Phys. Rev. C* 48 (3 Sept. 1993), pp. 1442–1444. DOI: 10.1103/PhysRevC.48.1442. URL: https://link.aps.org/doi/10.1103/PhysRevC.48.1442.

[108] T. A. Gabriel. *THE HIGH ENERGY TRANSPORT CODE HETC*. 1985.

[109] T. A. Gabriel, J. O. Johnson, and J. Brau. *MICAP: A program for low energy neutron, ion and gamma-ray transport and one of its applications in calorimeter design*. Tech. rep. CONF-8708186–7. United States, 1987, p. 19. URL: http://inis.iaea.org/search/search.aspx?orig_q=RN:21015198.

[110] T. J. Irvine. "Development of Neutron-Tagging Techniques and Application to Atmospheric Neutrino Oscillation Analysis in Super-Kamiokande". PhD thesis. The University of Tokyo, 2014.

[111] E. Mendoza and D. Cano-Ott. "Update of the Evaluated Neutron Cross Section Libraries for the Geant4 Code". In: *IAEA technical report INDC(NDS)-0758, Vienna* (June 2018).

[112] G. Horton-Smith. *GLG4sim, Generic liquid-scintillator anti-neutrino detector Geant4 simulation*. URL: http://neutrino.phys.ksu.edu/~GLG4sim/ (visited on 01/06/2021).

[113] M. Sakuda et al. *Website of the ANNRI-Gd model*. URL: http://www.physics.okayama-u.ac.jp/~sakuda/ANNRI-Gd_ver1.html (visited on 12/14/2020).

[114] R. A. Wendell. "Prob3++". URL: https://github.com/rogerwendell/Prob3plusplus.

[115] A. M. Dziewonski and D. L. Anderson. "Preliminary reference Earth model". In: *Physics of the Earth and Planetary Interiors* 25.4 (1981), pp. 297–356. ISSN: 0031-9201. DOI: https://doi.org/10.1016/0031-9201(81)90046-7. URL: https://www.sciencedirect.com/science/article/pii/0031920181900467.

[116] M. Jiang et al. "Atmospheric neutrino oscillation analysis with improved event reconstruction in Super-Kamiokande IV". In: *Progress of Theoretical and Experimental Physics* 2019.5 (May 2019). DOI: 10.1093/ptep/ptz015. URL: https://doi.org/10.1093%2Fptep%2Fptz015.

[117] M. Shiozawa. "Reconstruction algorithms in the Super-Kamiokande large water Cherenkov detector". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 433.1 (1999), pp. 240–246. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(99)00359-9. URL: https://www.sciencedirect.com/science/article/pii/S0168900299003599.

[118] A. Takenaka. "Search for Proton Decay via $p \to e + \pi^0$ and $p \to \mu + \pi^0$ with an Enlarged Fiducial Mass of the Super-Kamiokande Detector". PhD thesis. The University of Tokyo, 2020.

[119] M. Smy. "Low Energy Event Reconstruction and Selection in Super-Kamiokande-III". In: *International Cosmic Ray Conference*. Vol. 5. International Cosmic Ray Conference. Jan. 2008, pp. 1279–1282.

[120] K. Abe et al. "Solar neutrino results in Super-Kamiokande-III". In: *Phys. Rev. D* 83 (5 Mar. 2011), p. 052010. DOI: 10.1103/PhysRevD.83.052010. URL: https://link.aps.org/doi/10.1103/PhysRevD.83.052010.

[121] K. Huang. "Measurement of the Neutrino-Oxygen Neutral Current Quasi-elastic Interaction Cross-section by Observing Nuclear De-excitation γ-rays in the T2K Experiment". PhD thesis. Kyoto University, 2015.

[122] K. P. Lee. "Study of the neutrino mass hierarchy with the atmospheric neutrino data observed in Super-Kamiokande". PhD thesis. The University of Tokyo, 2012.

[123] C. Zeitnitz and T. A. Gabriel. "The GEANT-CALOR interface and benchmark calculations of ZEUS test calorimeters". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 349.1 (1994), pp. 106–111. ISSN: 0168-9002. DOI: https://doi.org/10.1016/0168-9002(94)90613-0. URL: http://www.sciencedirect.com/science/article/pii/0168900294906130.

[124] S. Han. "A study of neutron tagging algorithm for SK-Gd". MA thesis. The University of Tokyo, 2021.

[125] S. Han. *NTag*. https://github.com/seunghophysics/NTag. 2020.

[126] A. Bellerive et al. "The Sudbury Neutrino Observatory". In: *Nuclear Physics B* 908 (July 2016), pp. 30–51. ISSN: 0550-3213. DOI: 10.1016/j.nuclphysb.2016.04.035. URL: http://dx.doi.org/10.1016/j.nuclphysb.2016.04.035.

[127] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016.

[128] F. Chollet et al. *Keras*. https://keras.io. 2015.

[129] K. He et al. *Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification*. 2015. arXiv: 1502.01852 [cs.CV].

[130] D. P. Kingma and J. Ba. *Adam: A Method for Stochastic Optimization*. 2017. arXiv: 1412.6980 [cs.LG].

[131] L. S. Shapley. *Notes on the n-Person Game – II: The Value of an n-Person Game*. 1951.

[132] S. Lundberg and S. I. Lee. *A Unified Approach to Interpreting Model Predictions*. 2017. arXiv: 1705.07874 [cs.AI].

[133] "The 4.438MeV gamma to neutron ratio for the Am–Be neutron source". In: *Applied Radiation and Isotopes* 65.12 (2007), pp. 1318–1321. ISSN: 0969-8043. DOI: https://doi.org/10.1016/j.apradiso.2007.04.007. URL: http://www.sciencedirect.com/science/article/pii/S0969804307001200.

[134] F. De Guarrini and R. Malaroda. "Two different technique measurements of the neutron spectrum of an Am-Be source". In: *Nuclear Instruments and Methods* 92.2 (1971), pp. 277–284. ISSN: 0029-554X. DOI: https://doi.org/10.1016/0029-554X(71)90205-9. URL: https://www.sciencedirect.com/science/article/pii/0029554X71902059.

[135] *Neutron reference radiations fields — Part 1: Characteristics and methods of production*. Standard. Geneva, CH: International Organization for Standardization, Mar. 2021.

[136] M. Harada. "Development of Neutron Tagging Algorithm and Search for Supernova Relic Neutrino in SK-Gd Experiment". PhD thesis. Okayama University, 2023.

[137] H. Ito et al. "Analyzing the neutron and γ-ray emission properties of an americium–beryllium tagged neutron source". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1057 (Dec. 2023), p. 168701. DOI: 10.1016/j.nima.2023.168701. URL: https://doi.org/10.1016%2Fj.nima.2023.168701.

[138] *BGO Bismuth Germanate*. URL: https://www.crystals.saint-gobain.com/products/bgo (visited on 12/27/2020).

[139] S. Agostinelli et al. "Geant4—a simulation toolkit". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(03)01368-8. URL: https://www.sciencedirect.com/science/article/pii/S0168900203013688.

[140] J. Allison et al. "Geant4 developments and applications". In: *IEEE Transactions on Nuclear Science* 53.1 (2006), pp. 270–278. DOI: 10.1109/TNS.2006.869826.

[141] R. Brun et al. "GEANT3". In: (Sept. 1987).

[142] P. Williams et al. "Optical, Thermo-optic, Electro-optic, and Photoelastic Properties of Bismuth Germanate (Bi4Ge3O12)". In: *Applied Optics - APPL OPT* 35 (July 1996), pp. 3562–3569. DOI: 10.1364/AO.35.003562.

[143] S. Gundacker et al. "Precise rise and decay time measurements of inorganic scintillators by means of X-ray and 511 keV excitation". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 891 (2018), pp. 42–52. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2018.02.074. URL: https://www.sciencedirect.com/science/article/pii/S0168900218302286.

[144] W. Wolszczak et al. "Temperature properties of scintillators for PET detectors: A comparative study". In: *2014 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. 2014, pp. 1–4. DOI: 10.1109/NSSMIC.2014.7430768.

[145] G. F. Knoll. *Radiation Detection and Measurement, 3rd ed.* 3rd edition. New York: John Wiley and Sons, 2000.

[146] B. Pritychenko and S. F. Mughabghab. "Neutron Thermal Cross Sections, Westcott Factors, Resonance Integrals, Maxwellian Averaged Cross Sections and Astrophysical Reaction Rates Calculated from the ENDF/B-VII.1, JEFF-3.1.2, JENDL-4.0, ROSFOND-2010, CENDL-3.1 and EAF-2010 Evaluated Data Libraries". In: *Nuclear Data Sheets* 113.12 (2012), pp. 3120–3144. ISSN: 0090-3752. DOI: https://doi.org/10.1016/j.nds.2012.11.007. URL: https://www.sciencedirect.com/science/article/pii/S0090375212000932.

[147] M S Basunia A Hurst B Sleaford N Summers J E Escher Zs. Révay L Szentmiklósi T Belgya H. D. Choi R. B. Firestone and M Krtička. "Radiative Capture Cross Sections of 155,157Gd for Thermal Neutrons". In: *Nuclear Science and Engineering* 177.2 (2014), pp. 219–232. DOI: 10.13182/NSE13-49. URL: https://doi.org/10.13182/NSE13-49.

[148] M. J. Trbovich J A Burke N J Drindak H D Knox R V Ballad R C Block Y Danon G. Leinweber D. P. Barry and L I Severnyak. "Neutron Capture and Total Cross-Section Measurements and Resonance Parameters of Gadolinium". In: *Nuclear Science and Engineering* 154.3 (2006), pp. 261–279. DOI: 10.13182/NSE05-64. URL: https://doi.org/10.13182/NSE05-64.

[149] J. Kopecky, M. Uhl, and R. E. Chrien. "Radiative strength in the compound nucleus ^{157}Gd". In: *Phys. Rev. C* 47 (1 Jan. 1993), pp. 312–322. DOI: 10.1103/PhysRevC.47.312. URL: https://link.aps.org/doi/10.1103/PhysRevC.47.312.

[150] A. J. M. Plompen et al. "The joint evaluated fission and fusion nuclear data library, JEFF-3.3". In: *The European Physical Journal A* 56.7 (2020), p. 181. ISSN: 1434-601X. DOI: 10.1140/epja/s10050-020-00141-9. URL: https://doi.org/10.1140/epja/s10050-020-00141-9.

[151] D. E. Groom, N. V. Mokhov, and S. Striganov. "MUON STOPPING POWER AND RANGE TABLES 10 MeV–100 TeV". In: *Atomic Data and Nuclear Data Tables* 78.2 (2001), pp. 183–356. ISSN: 0092-640X. DOI: https://doi.org/10.1006/adnd.2001.0861. URL: https://www.sciencedirect.com/science/article/pii/S0092640X01908617.

[152] S. Baker and R. D. Cousins. "Clarification of the Use of Chi Square and Likelihood Functions in Fits to Histograms". In: *Nucl. Instrum. Meth.* 221 (1984), pp. 437–442. DOI: 10.1016/0167-5087(84)90016-4.

[153] H. Zhang. "Study of Low Energy Electron Anti-neutrinos at Super-Kamiokande lV". PhD thesis. Tsinghua University, 2012.

[154] Y. Hayato and L. Pickering. "The NEUT neutrino interaction simulation program library". In: *The European Physical Journal Special Topics* 230.24

(Oct. 2021), pp. 4469–4481. DOI: 10.1140/epjs/s11734-021-00287-7. URL: https://doi.org/10.1140%2Fepjs%2Fs11734-021-00287-7.

[155] C. Patrignani. "Review of Particle Physics". In: *Chinese Physics C* 40.10 (Oct. 2016), p. 100001. DOI: 10.1088/1674-1137/40/10/100001. URL: https://dx.doi.org/10.1088/1674-1137/40/10/100001.

[156] D. Servén and C. Brummitt. *pyGAM*. https://github.com/dswah/pyGAM. 2018.

[157] C. Andreopoulos et al. "The GENIE Neutrino Monte Carlo Generator". In: *Nucl. Instrum. Meth. A* 614 (2010), pp. 87–104. DOI: 10.1016/j.nima.2009.12.009. arXiv: 0905.2517 [hep-ph].

[158] C. Andreopoulos et al. "The GENIE Neutrino Monte Carlo Generator: Physics and User Manual". In: (Oct. 2015). arXiv: 1510.05494 [hep-ph].

[159] D. H. Wright and M. H. Kelsey. "The Geant4 Bertini Cascade". In: *Nucl. Instrum. Meth. A* 804 (2015), pp. 175–188. DOI: 10.1016/j.nima.2015.09.058.

[160] V. Flaminio et al. "Compilation of Cross-sections. 3. P and $\bar{P}$ Induced Reactions". In: (Apr. 1984).

[161] A. Boudard et al. "New potentialities of the Liège intranuclear cascade model for reactions induced by nucleons and light charged particles". In: *Phys. Rev. C* 87 (1 Jan. 2013), p. 014606. DOI: 10.1103/PhysRevC.87.014606. URL: https://link.aps.org/doi/10.1103/PhysRevC.87.014606.

[162] A. Kelic, M. Valentina Ricciardi, and K. H. Schmidt. *ABLA07 - towards a complete description of the decay channels of a nuclear system from spontaneous fission to multifragmentation*. 2009. arXiv: 0906.4193 [nucl-th].

[163] J. C. David. "Spallation reactions: A successful interplay between modeling and applications". In: *The European Physical Journal A* 51.6 (June 2015). ISSN: 1434-601X. DOI: 10.1140/epja/i2015-15068-1. URL: http://dx.doi.org/10.1140/epja/i2015-15068-1.

[164] A. Boudard et al. “Intranuclear cascade model for a comprehensive description of spallation reaction data”. In: *Phys. Rev. C* 66 (4 Oct. 2002), p. 044615. DOI: 10.1103/PhysRevC.66.044615. URL: https://link.aps.org/doi/10.1103/PhysRevC.66.044615.

[165] L. L. Salcedo et al. “Computer simulation of inclusive pion nuclear reactions”. In: *Nuclear Physics A* 484.3 (1988), pp. 557–592. ISSN: 0375-9474. DOI: https://doi.org/10.1016/0375-9474(88)90310-7. URL: https://www.sciencedirect.com/science/article/pii/0375947488903107.

[166] V. Ammosov. *Low multiplicity final states in neutrino interaction at low energy—SKAT results*. 2001.

[167] V. Weisskopf. “Statistics and Nuclear Reactions”. In: *Phys. Rev.* 52 (4 Aug. 1937), pp. 295–303. DOI: 10.1103/PhysRev.52.295. URL: https://link.aps.org/doi/10.1103/PhysRev.52.295.

[168] S. Dytman et al. “Comparison of validation methods of simulations for final state interactions in hadron production experiments”. In: *Phys. Rev. D* 104 (5 Sept. 2021), p. 053006. DOI: 10.1103/PhysRevD.104.053006. URL: https://link.aps.org/doi/10.1103/PhysRevD.104.053006.

[169] J. M. Quesada et al. In: *Progress in Nuclear Science and Technology* (2011).

[170] S. Sakai. *Neutron multiplicity by atmospheric neutrino events*. Dec. 2022.

[171] G. Folger, V. N. Ivanchenko, and J. P. Wellisch. “The Binary Cascade”. In: *The European Physical Journal A - Hadrons and Nuclei* 21.3 (Sept. 2004), pp. 407–417. ISSN: 1434-601X. DOI: 10.1140/epja/i2003-10219-7. URL: https://doi.org/10.1140/epja/i2003-10219-7.

[172] M. Nakano and Y. Uozumi. “Extension of the intranuclear cascade model to neutron-induced nonelastic cross sections in the low-energy region”. In: *Phys. Rev. C* 100 (3 Sept. 2019), p. 034619. DOI: 10.1103/PhysRevC.100.034619. URL: https://link.aps.org/doi/10.1103/PhysRevC.100.034619.

[173] M. Nakano, Y. Yamaguchi, and Y. Uozumi. “Two effects in nucleon-induced nonelastic cross sections based on the intranuclear cascade model”. In: *Phys. Rev. C* 102 (2 Aug. 2020), p. 024608. DOI: 10.1103/PhysRevC.102.024608. URL: https://link.aps.org/doi/10.1103/PhysRevC.102.024608.

[174] A. R. Back et al. *Accelerator Neutrino Neutron Interaction Experiment (ANNIE): Preliminary Results and Physics Phase Proposal*. 2017. arXiv: 1707.08222 [physics.ins-det].

[175] S. Gwon et al. "Neutron detection and application with a novel 3D-projection scintillator tracker in the future long-baseline neutrino oscillation experiments". In: *Phys. Rev. D* 107 (3 Feb. 2023), p. 032012. DOI: 10.1103/PhysRevD.107.032012. URL: https://link.aps.org/doi/10.1103/PhysRevD.107.032012.

[176] S. V. Garode et al. "Mechanical design of water cherenkov test experiment (WCTE) at CERN". In: *J. Phys. Conf. Ser.* 2374.1 (2022), p. 012035. DOI: 10.1088/1742-6596/2374/1/012035.

[177] P. V. C. Hough. "Machine Analysis of Bubble Chamber Pictures". In: *Conf. Proc. C* 590914 (1959). Ed. by L. Kowarski, pp. 554–558.

[178] K. Grauman. "Fitting: Voting and the Hough Transform". URL: https://vision.cs.utexas.edu/376-spring2018/slides/lecture7-spring2018.pdf.

[179] Y. Gal and Z. Ghahramani. *Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning*. 2016. arXiv: 1506.02142 [stat.ML].