

Measurement of neutron production in atmospheric neutrino interactions at Super-Kamiokande

S. Han,²⁸ K. Abe,^{1, 49} S. Abe,^{1, 49} Y. Asaoka,^{1, 49} C. Bronner,¹ M. Harada,¹ Y. Hayato,^{1, 49} K. Hiraide,^{1, 49} K. Hosokawa,¹ K. Ieki,^{1, 49} M. Ikeda,^{1, 49} J. Kameda,^{1, 49} Y. Kanemura,¹ R. Kaneshima,¹ Y. Kashiwagi,¹ Y. Kataoka,^{1, 49} S. Miki,¹ S. Mine,^{1, 7} M. Miura,^{1, 49} S. Moriyama,^{1, 49} M. Nakahata,^{1, 49, 1} S. Nakayama,^{1, 49} Y. Noguchi,¹ G. Pronost,¹ K. Sato,¹ K. Okamoto,¹ H. Sekiya,^{1, 49} H. Shiba,¹ K. Shimizu,¹ M. Shiozawa,^{1, 49} Y. Sonoda,¹ Y. Suzuki,¹ A. Takeda,^{1, 49} Y. Takemoto,^{1, 49} A. Takenaka,¹ H. Tanaka,^{1, 49} T. Yano,¹ T. Kajita,^{2, 49, 23} K. Okumura,^{2, 49} T. Tashiro,² T. Tomiya,² X. Wang,² S. Yoshida,² P. Fernandez,³ L. Labarga,³ N. Ospina,³ B. Zaldivar,³ B. W. Pointon,^{6, 53} C. Yanagisawa,^{4, 35} E. Kearns,^{5, 49} J. L. Raaf,⁵ L. Wan,⁵ T. Wester,⁵ J. Bian,⁷ N. J. Grishevich,⁷ M. B. Smy,^{7, 49} H. W. Sobel,^{7, 49} V. Takhistov,^{7, 25} A. Yankelevich,⁷ J. Hill,⁸ M. C. Jang,⁹ S. H. Lee,⁹ D. H. Moon,⁹ R. G. Park,⁹ B. S. Yang,⁹ B. Bodur,¹⁰ K. Scholberg,^{10, 49} C. W. Walter,^{10, 49} A. Beauchêne,¹¹ O. Drapier,¹¹ A. Giampaolo,¹¹ A. Ershova,¹¹ Th. A. Mueller,¹¹ A. D. Santos,¹¹ P. Paganini,¹¹ C. Quach,¹¹ R. Rogly,¹¹ T. Nakamura,¹² J. S. Jang,¹³ L. N. Machado,¹⁴ J. G. Learned,¹⁵ K. Choi,¹⁶ N. Iovine,¹⁶ S. Cao,¹⁷ L. H. V. Anthony,¹⁸ D. Martin,¹⁸ N. W. Prouse,¹⁸ M. Scott,¹⁸ Y. Uchida,¹⁸ A. A. Sztuc,¹⁸ V. Berardi,¹⁹ N. F. Calabria,¹⁹ M. G. Catanesi,¹⁹ E. Radicioni,¹⁹ A. Langella,²⁰ G. De Rosa,²⁰ G. Collazuol,²¹ M. Feltre,²¹ F. Iacob,²¹ M. Mattiazzi,²¹ L. Ludovici,²² M. Gonin,²³ L. Périsse,²³ B. Quilain,²³ C. Fujisawa,²⁴ S. Horiuchi,²⁴ M. Kobayashi,²⁴ Y. M. Liu,²⁴ Y. Maekawa,²⁴ Y. Nishimura,²⁴ R. Okazaki,²⁴ R. Akutsu,²⁵ M. Friend,²⁵ T. Hasegawa,²⁵ T. Ishida,²⁵ T. Kobayashi,²⁵ M. Jakkapu,²⁵ T. Matsubara,²⁵ T. Nakadaira,²⁵ K. Nakamura,^{25, 49} Y. Oyama,²⁵ A. Portocarrero Yrey,²⁵ K. Sakashita,²⁵ T. Sekiguchi,²⁵ T. Tsukamoto,²⁵ N. Bhuiyan,²⁶ G. T. Burton,²⁶ F. Di Lodovico,²⁶ J. Gao,²⁶ A. Goldsack,²⁶ T. Katori,²⁶ J. Migenda,²⁶ R. M. Ramsden,²⁶ Z. Xie,²⁶ S. Zsoldos,^{26, 49} A. T. Suzuki,²⁷ Y. Takagi,²⁷ Y. Takeuchi,^{27, 49} H. Zhong,²⁷ J. Feng,²⁸ L. Feng,²⁸ J. R. Hu,²⁸ Z. Hu,²⁸ M. Kawaue,²⁸ T. Kikawa,²⁸ M. Mori,²⁸ T. Nakaya,^{28, 49} R. A. Wendell,^{28, 49} K. Yasutome,²⁸ S. J. Jenkins,²⁹ N. McCauley,²⁹ A. Tarrant,²⁹ P. Mehta,²⁹ M. J. Wilking,³⁰ Y. Fukuda,³¹ Y. Itow,^{32, 33} H. Menjo,³² Y. Yoshioka,³² K. Ninomiya,³² J. Lagoda,³⁴ S. M. Lakshmi,³⁴ M. Mandal,³⁴ P. Mijakowski,³⁴ Y. S. Prabhu,³⁴ J. Zalipska,³⁴ M. Jia,³⁵ J. Jiang,³⁵ C. K. Jung,³⁵ W. Shi,³⁵ Y. Hino,³⁶ H. Ishino,³⁶ Y. Koshio,^{36, 49} F. Nakanishi,³⁶ S. Sakai,³⁶ T. Tada,³⁶ T. Tano,³⁶ T. Ishizuka,³⁷ G. Barr,³⁸ D. Barrow,³⁸ L. Cook,^{38, 49} S. Samani,³⁸ D. Wark,^{38, 44} A. Holin,³⁹ F. Nova,³⁹ S. Jung,⁴⁰ J. Y. Yang,⁴⁰ J. Yoo,⁴⁰ J. E. P. Fannon,⁴¹ L. Kneale,⁴¹ M. Malek,⁴¹ J. M. McElwee,⁴¹ T. Peacock,⁴¹ P. Stowell,⁴¹ M. D. Thiesse,⁴¹ L. F. Thompson,⁴¹ S. T. Wilson,⁴¹ H. Okazawa,⁴² S. M. Lakshmi,⁴³ S. B. Kim,⁴⁵ E. Kwon,⁴⁵ M. W. Lee,⁴⁵ J. W. Seo,⁴⁵ I. Yu,⁴⁵ A. K. Ichikawa,⁴⁶ K. D. Nakamura,⁴⁶ S. Tairafune,⁴⁶ K. Nishijima,⁴⁷ A. Eguchi,⁴⁸ S. Goto,⁴⁸ Y. Mizuno,⁴⁸ T. Muro,⁴⁸ K. Nakagiri,⁴⁸ Y. Nakajima,^{48, 49} S. Shima,⁴⁸ N. Taniuchi,⁴⁸ E. Watanabe,⁴⁸ M. Yokoyama,^{48, 49} P. de Perio,⁴⁹ S. Fujita,⁴⁹ C. Jesús-Valls,⁴⁹ K. Martens,⁴⁹ Ll. Martí,⁴⁹ K. M. Tsui,⁴⁹ M. R. Vagins,^{49, 7} J. Xia,⁴⁹ M. Kuze,⁵⁰ S. Izumiymama,⁵⁰ R. Matsumoto,⁵⁰ K. Terada,⁵⁰ R. Asaka,⁵¹ M. Ishitsuka,⁵¹ H. Ito,⁵¹ Y. Ommura,⁵¹ N. Shigeta,⁵¹ M. Shinoki,⁵¹ K. Yamauchi,⁵¹ T. Yoshida,⁵¹ Y. Nakano,⁵² F. Cormier,²⁸ R. Gaur,⁵³ V. Gousy-Leblanc,^{53, *} M. Hartz,⁵³ A. Konaka,⁵³ X. Li,⁵³ S. Chen,⁵⁴ Y. Wu,⁵⁴ B. D. Xu,⁵⁴ A. Q. Zhang,⁵⁴ B. Zhang,⁵⁴ M. Girgus,⁵⁵ M. Posiadala-Zezula,⁵⁵ S. B. Boyd,⁵⁶ R. Edwards,⁵⁶ D. Hadley,⁵⁶ M. Nicholson,⁵⁶ M. O'Flaherty,⁵⁶ B. Richards,⁵⁶ A. Ali,^{57, 53} B. Jamieson,⁵⁷ S. Amanai,⁵⁸ A. Minamino,⁵⁸ R. Shibayama,⁵⁸ R. Shimamura,⁵⁸ and S. Suzuki⁵⁸

(The Super-Kamiokande Collaboration)

¹Kamioka Observatory, Institute for Cosmic Ray Research, University of Tokyo, Kamioka, Gifu 506-1205, Japan

²Research Center for Cosmic Neutrinos, Institute for Cosmic Ray Research, University of Tokyo, Kashiwa, Chiba 277-8582, Japan

³Department of Theoretical Physics, University Autónoma Madrid, 28049 Madrid, Spain

⁴Science Department, Borough of Manhattan Community College / City University of New York, New York, New York, 1007, USA.

⁵Department of Physics, Boston University, Boston, MA 02215, USA

⁶Department of Physics, British Columbia Institute of Technology, Burnaby, BC, V5G 3H2, Canada

⁷Department of Physics and Astronomy, University of California, Irvine, Irvine, CA 92697-4575, USA

⁸Department of Physics, California State University, Dominguez Hills, Carson, CA 90747, USA

⁹Institute for Universe and Elementary Particles, Chonnam National University, Gwangju 61186, Korea

¹⁰Department of Physics, Duke University, Durham NC 27708, USA

¹¹Ecole Polytechnique, IN2P3-CNRS, Laboratoire Leprince-Ringuet, F-91120 Palaiseau, France

¹²Department of Physics, Gifu University, Gifu, Gifu 501-1193, Japan

¹³GIST College, Gwangju Institute of Science and Technology, Gwangju 500-712, Korea

¹⁴School of Physics and Astronomy, University of Glasgow, Glasgow, Scotland, G12 8QQ, United Kingdom

¹⁵Department of Physics and Astronomy, University of Hawaii, Honolulu, HI 96822, USA

¹⁶Center for Underground Physics, Institute for Basic Science (IBS), Daejeon, 34126, Korea

¹⁷Institute For Interdisciplinary Research in Science and Education, ICISE, Quy Nhon, 55121, Vietnam

¹⁸Department of Physics, Imperial College London , London, SW7 2AZ, United Kingdom

¹⁹Dipartimento Interuniversitario di Fisica, INFN Sezione di Bari and Università e Politecnico di Bari, I-70125, Bari, Italy

²⁰Dipartimento di Fisica, INFN Sezione di Napoli and Università di Napoli, I-80126, Napoli, Italy

²¹Dipartimento di Fisica, INFN Sezione di Padova and Università di Padova, I-35131, Padova, Italy

²²INFN Sezione di Roma and Università di Roma "La Sapienza", I-00185, Roma, Italy

²³ILANCE, CNRS - University of Tokyo International Research Laboratory, Kashiwa, Chiba 277-8582, Japan

²⁴Department of Physics, Keio University, Yokohama, Kanagawa, 223-8522, Japan

²⁵High Energy Accelerator Research Organization (KEK), Tsukuba, Ibaraki 305-0801, Japan

²⁶Department of Physics, King's College London, London, WC2R 2LS, UK

²⁷Department of Physics, Kobe University, Kobe, Hyogo 657-8501, Japan

²⁸Department of Physics, Kyoto University, Kyoto, Kyoto 606-8502, Japan

²⁹Department of Physics, University of Liverpool, Liverpool, L69 7ZE, United Kingdom

³⁰School of Physics and Astronomy, University of Minnesota, Minneapolis, MN 55455, USA

³¹Department of Physics, Miyagi University of Education, Sendai, Miyagi 980-0845, Japan

³²Institute for Space-Earth Environmental Research, Nagoya University, Nagoya, Aichi 464-8602, Japan

³³Kobayashi-Maskawa Institute for the Origin of Particles and the Universe, Nagoya University, Nagoya, Aichi 464-8602, Japan

³⁴National Centre For Nuclear Research, 02-093 Warsaw, Poland

³⁵Department of Physics and Astronomy, State University of New York at Stony Brook, NY 11794-3800, USA

³⁶Department of Physics, Okayama University, Okayama, Okayama 700-8530, Japan

³⁷Media Communication Center, Osaka Electro-Communication University, Neyagawa, Osaka, 572-8530, Japan

³⁸Department of Physics, Oxford University, Oxford, OX1 3PU, United Kingdom

³⁹Rutherford Appleton Laboratory, Harwell, Oxford, OX11 0QX, UK

⁴⁰Department of Physics and Astronomy, Seoul National University, Seoul 151-742, Korea

⁴¹Department of Physics and Astronomy, University of Sheffield, S3 7RH, Sheffield, United Kingdom

⁴²Department of Informatics in Social Welfare, Shizuoka University of Welfare, Yaizu, Shizuoka, 425-8611, Japan

⁴³August Cichkowski Institute of Physics, University of Silesia in Katowice, 75 Putku Piechoty 1, 41-500 Chorzów, Poland

⁴⁴STFC, Rutherford Appleton Laboratory, Harwell Oxford, and

Daresbury Laboratory, Warrington, OX11 0QX, United Kingdom

⁴⁵Department of Physics, Sungkyunkwan University, Suwon 440-746, Korea

⁴⁶Department of Physics, Faculty of Science, Tohoku University, Sendai, Miyagi, 980-8578, Japan

⁴⁷Department of Physics, Tokai University, Hiratsuka, Kanagawa 259-1292, Japan

⁴⁸Department of Physics, University of Tokyo, Bunkyo, Tokyo 113-0033, Japan

⁴⁹Kavli Institute for the Physics and Mathematics of the Universe (WPI), The University of
Tokyo Institutes for Advanced Study, University of Tokyo, Kashiwa, Chiba 277-8583, Japan

⁵⁰Department of Physics, Institute of Science Tokyo, Meguro, Tokyo 152-8551, Japan

⁵¹Department of Physics, Faculty of Science and Technology, Tokyo University of Science, Noda, Chiba 278-8510, Japan

⁵²Faculty of Science, University of Toyama, Toyama City, Toyama 930-8555, Japan

⁵³TRIUMF, 4004 Wesbrook Mall, Vancouver, BC, V6T2A3, Canada

⁵⁴Department of Engineering Physics, Tsinghua University, Beijing, 100084, China

⁵⁵Faculty of Physics, University of Warsaw, Warsaw, 02-093, Poland

⁵⁶Department of Physics, University of Warwick, Coventry, CV4 7AL, UK

⁵⁷Department of Physics, University of Winnipeg, MB R3J 3L8, Canada

⁵⁸Department of Physics, Yokohama National University, Yokohama, Kanagawa, 240-8501, Japan

(Dated: February 26, 2025)

We present measurements of total neutron production from atmospheric neutrino interactions in water, analyzed as a function of the electron-equivalent visible energy over a range of 30 MeV to 10 GeV. These results are based on 4,270 days of data collected by Super-Kamiokande, including 564 days with 0.011 wt% gadolinium added to enhance neutron detection. The measurements are compared to predictions from neutrino event generators combined with various hadron-nucleus interaction models, which consist of an intranuclear cascade model and a nuclear de-excitation model. We observe significant variations in the predictions depending on the choice of hadron-nucleus interaction model. We discuss key factors that contribute to describing our data, such as in-medium effects in the intranuclear cascade and the accuracy of statistical evaporation modeling.

teraction cross sections and particle kinematics. Modeling these “nuclear effects” is particularly critical for GeV-scale neutrino experiments aiming to measure neutrino oscillation parameters, including CP violation and mass ordering. Outgoing hadrons serve as valuable probes of these effects, and advancements in neutrino detectors have enabled precise measurements of hadron multiplicities and kinematics. For instance, measurements of proton production in neutrino-argon interactions have revealed several discrepancies between observed and predicted kinematic distributions [1].

Detecting neutrons is challenging since they do not leave ionization tracks. Tracking detectors identify neutrons via displaced proton tracks, with detection efficiency limited by the small (n, p) reaction cross section [2]. This issue is critical with GeV-scale neutrino interactions, where neutrons often carry substantial energy. Misestimating this “missing energy” can bias key measurements, such as the Dirac CP phase [3].

Neutrons can also be detected via the radiative neutron capture (n, γ) reaction, which has no practical energy threshold. Neutrons with kinetic energies of a few MeV or lower are likely to thermalize, undergoing (n, γ) reactions with well-defined timescales and energy signatures that enable clean signal selection. This characteristic historically made neutrons effective tags for antineutrino charged-current (CC) interactions (e.g., $\bar{\nu}_\mu p \rightarrow \mu^+ n$) compared to neutrino interactions (e.g., $\nu_\mu n \rightarrow \mu^- p$) [4]. Neutron tagging remains relevant today, for instance, in atmospheric neutrino oscillation analyses, where it enhances sensitivity to both neutrino mass ordering and CP violation, by preventing the cancellation of opposite-sign effects in neutrino and antineutrino oscillation probabilities [5]. Additionally, neutron tagging helps suppress atmospheric neutrino backgrounds for rare events, such as proton decay (e.g., $p \rightarrow e^+ \pi^0$ [6]), which in many cases is not expected to produce neutrons, or inverse beta decay ($\bar{\nu}_e p \rightarrow e^+ n$) induced by supernova $\bar{\nu}_e$, which typically produces only one neutron. Accurately predicting the number of detectable neutrons from neutrino interactions is crucial for these analyses and requires well-constrained uncertainties in interaction models.

The modeling approach commonly adopted by GeV-scale neutrino experiments is illustrated in Figure 1. Neutrino event generators simulate outgoing hadrons from the initial neutrino interaction, either at the nucleon or quark level, and subsequently model intranuclear hadron transport (often referred to as final-state interactions, or FSI), followed by nuclear de-excitation. Similarly, particle transport codes such as Geant4 [7–9] simulate hadron transport and nuclear de-excitation to describe downstream hadron-nucleus interactions within the detector.

Accurate prediction of the observable neutron multiplicity requires careful consideration of secondary neutron production. For hadron transport in the $O(0.1\text{--}1)$ GeV energy range, semi-classical Intranuclear Cascade (INC) models [10] are commonly used. These models approximate the process as a series of binary collisions be-

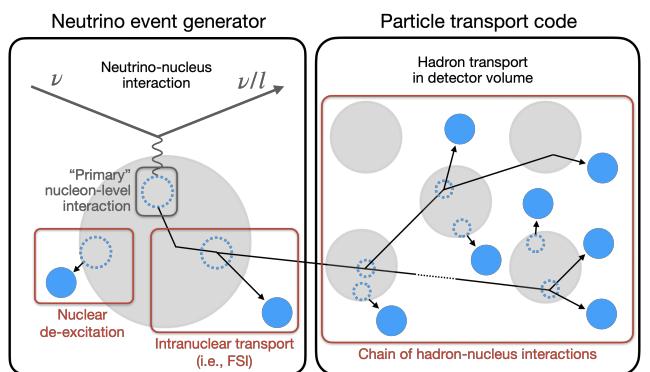


FIG. 1. Schematic of sources of nucleon production in a typical GeV-scale neutrino event simulation setup. Dashed blue circles represent nucleon holes, black arrows represent nucleon trajectories, and solid blue circles represent detectable nucleons ejected from nuclei.

tween a hadron projectile and a quasi-free nucleon within the target nucleus. Subsequent nuclear de-excitation involves an evaporation process that releases neutrons with kinetic energies of a few MeV and also contributes significantly to neutron production at lower energies. Notably, variations in how these models account for nuclear effects often lead to significant discrepancies in predictions [11]. Several studies have measured neutron production from atmospheric or artificial neutrino interactions using water (T2K [12]), heavy water (SNO [13]), and hydrocarbon (MINERvA [14, 15], KamLAND [16]) as target materials. Several of these studies [12, 14, 15] reported deficits in observed neutron signals compared to predictions from neutrino event generators, with the discrepancies often attributed to the inaccuracy of hadron transport models.

In this paper, we present a measurement of neutron production following atmospheric neutrino interactions in water. Using data from events fully contained within the Super-Kamiokande detector from 2008 to 2022, we evaluated the average multiplicity of (n, γ) reactions (or “neutron captures”) as a function of electron-equivalent visible energy, a calorimetric proxy for neutrino momentum transfer. These observations were compared with predictions generated using various models relevant to secondary neutron production.

This paper is structured as follows. Section II provides a brief overview of the SK detector. The selection process for atmospheric neutrino events and neutron signals, along with the estimation of selection performance, are detailed in Sections III and IV. Section V outlines the methodology for determining the average (n, γ) multiplicity per visible energy bin and the associated systematic uncertainties. Section VI introduces the interaction models used for generating predictions. Finally, Sections VII and VIII present a comparison of observations with predictions and discuss the implications of the results.

TABLE I. SK operational phases and neutron-related characteristics. SK-IV, V, VI data were used in this analysis.

Phase	Dates	Livetime [days]	Gd concentration ^a [wt%]	H(n, γ) [%]	Expected (n, γ) ratio ^b H(n, γ) [%]	Gd(n, γ) [%]	(n, γ) time constant ^c [μs]
SK-I-III	1996-2008	2805.9	-	>99.9	-	-	No data
SK-IV	2008-2018	3244.4	-	>99.9	-	-	204.8 ± 9.8
SK-V	2019-2020	461.0	-	>99.9	-	-	199.8 ± 10.2
SK-VI	2020-2022	564.4 ^d	0.0110 ± 0.0001 [18]	56.1 ± 1.5	43.9 ± 1.5	116.2 ± 2.3	
SK-VII-VIII	2022-present	-	0.0332 ± 0.0002 [19]	29.7 ± 0.7	70.3 ± 0.7	61.8 ± 0.1 [19]	

^a Based on the amount of dissolved Gd.

^b Based on the evaluated thermal (n, γ) reaction cross sections and uncertainties of ENDF/B-VII.1 [17].

^c Weighted mean of all Am/Be neutron source measurements, explained in Section IV B.

^d Excludes earlier runs which showed signs of non-uniform Gd concentration, i.e., varying time constant by position.

209 II. THE SUPER-KAMIOKANDE DETECTOR

210 Super-Kamiokande (SK) [20] is an underground water
 211 Cherenkov detector located in Gifu, Japan. It consists of
 212 two optically separated, concentric cylindrical volumes:
 213 the inner detector (ID) containing 32.5 ktons of water
 214 and equipped with 11,129 inward-facing photomultiplier
 215 tubes (PMTs) and the outer detector serving as a cosmic-
 216 ray veto. The detector registers a PMT signal with a
 217 pulse height greater than 0.25 photoelectron-equivalent
 218 charge as a “hit”. If the number of ID or OD PMT hits
 219 within a 200-ns sliding time window ($N_{200\text{-ns}}$) exceeds a
 220 given threshold, an event trigger is issued. The details of
 221 the detector can be found in [20, 21].

222 Charged particles, namely electrons and muons pro-
 223 duced by charged-current neutrino interactions, are iden-
 224 tified through Cherenkov radiation. The radiation is pro-
 225 jected onto the PMTs as a characteristic ring pattern
 226 that depends on the particle type and energy. This ring
 227 pattern serves as the basis for particle reconstruction.
 228 Neutrons are indirectly identified via Compton-scattered
 229 electrons resulting from (n, γ) reactions. In pure water,
 230 most occur on ^1H , emitting a single 2.2 MeV gamma-
 231 ray. With the recent addition of gadolinium (Gd), the
 232 majority of neutrons are expected to be captured by Gd
 233 isotopes, resulting in a total gamma-radiated energy of
 234 around 8 MeV.

235 The $O(1)$ MeV signal identification performance is sig-
 236 nificantly influenced by variations in detector character-
 237 istics. Parameters such as individual PMT gain, tim-
 238 ing properties, quantum efficiency, and optical absorption
 239 and scattering in water are continuously monitored using
 240 cosmic-ray muons and light sources [21]. Additionally,
 241 the uncertainty in Cherenkov ring energy reconstruc-
 242 tion (described in Section III B) is evaluated over a wide en-
 243 ergy range using naturally occurring particles, including
 244 cosmic-ray muons, Michel electrons, and neutral pions
 245 produced in neutral-current (NC) atmospheric neutrino
 246 interactions in water. Figure 2 illustrates the agreement
 247 between data and simulation in energy reconstruction for
 248 the Gd-loaded SK-VI phase, which is mostly within 2%
 249 across the $O(10\text{-}10^4)$ MeV range and consistent with the
 250 pure water phase results reported in [5].

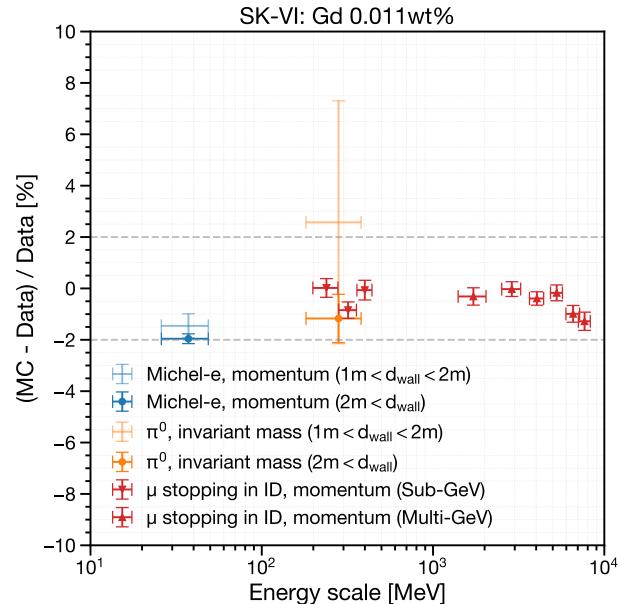


FIG. 2. Comparison of energy reconstruction performance between Monte Carlo (MC) simulation and observed data for naturally occurring particles in the Gd-loaded SK-VI phase. The variable d_{wall} represents the distance (in meters) from the reconstructed vertex to the nearest ID photodetector wall. The fiducial volume for this study is defined as $d_{\text{wall}} > 1$ m.

The detector has operated through eight different phases. Neutron detection began with the fourth phase SK-IV, following the electronics upgrade [22] that allowed extended event recording up to 535 μs after certain ID triggers. This has enabled analysis of delayed neutron captures that occur with a time scale of $O(10\text{-}100)$ μs following an atmospheric neutrino interaction. Between SK-IV and SK-V, in 2018, the detector underwent refurbishment, during which malfunctioning PMTs were replaced. The later phases, SK-VI, SK-VII, and SK-VIII involved the dissolution of $\text{Gd}_2(\text{SO}_4)_3$ into the water volume to enhance neutron detection efficiency [18, 19]. Table I summarizes the relevant operational conditions.

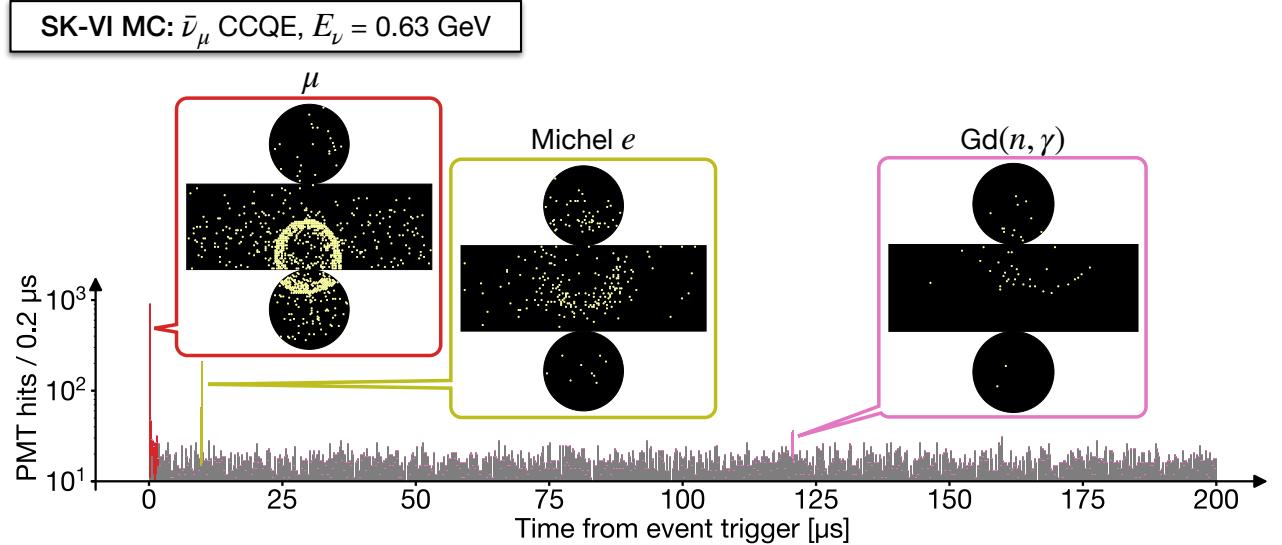


FIG. 3. A PMT hit time distribution for a typical $\bar{\nu}_\mu$ charged-current (CC) quasi-elastic (QE) interaction with a muon and a neutron as final states. The zero is set at the event trigger. The event displays feature the “prompt” muon signal (red) and the two types of “delayed” coincident signals — Michel electrons from muon decay (olive) and neutron captures on Gd (pink). The gray bars represent randomly recorded background PMT hits. This event was simulated in the SK-VI configuration.

265 III. ATMOSPHERIC NEUTRINO EVENTS

266 A. Data reduction

267 We followed a typical selection process for atmospheric
268 neutrino interactions that are fully contained within the
269 ID, similar to previous studies conducted at SK [5]. All
270 events were required to pass the ID trigger with the
271 threshold $N_{200\text{-ns}} \geq 58$ PMT hits—roughly correspond-
272 ing to a 10 MeV electron—followed by the extended 535
273 μs event window for neutron detection. Background
274 events arising from cosmic-ray muons, low-energy ra-
275 dioactivity, and neutrino events with particles exiting the
276 tank were reduced through OD veto and ID charge cuts.
277

278 Selected events were reconstructed as described in Sec-
279 tion III B. To further reject low-energy backgrounds, we
280 required that the reconstructed vertex be more than 1 m
281 away from the ID tank wall (defining the fiducial volume
282 with 23.2 kton of water) and that the visible energy be
283 larger than 30 MeV. The remaining background contam-
284 ination, mainly due to corner-clipping cosmic-ray muons
285 and PMT discharges, was estimated to be below 0.2%,
286 based on visual inspection [6].

286 B. Reconstruction of prompt Cherenkov rings

287 Figure 3 shows a PMT hit time distribution of a typi-
288 cal $\bar{\nu}_\mu$ charged-current (CC) quasi-elastic (QE) event fol-
289 lowed by a Michel electron and a $\text{Gd}(n, \gamma)$ reaction, along
290 with the corresponding event displays. For the “prompt”
291 radiation due to charged particles (namely, electrons and
292

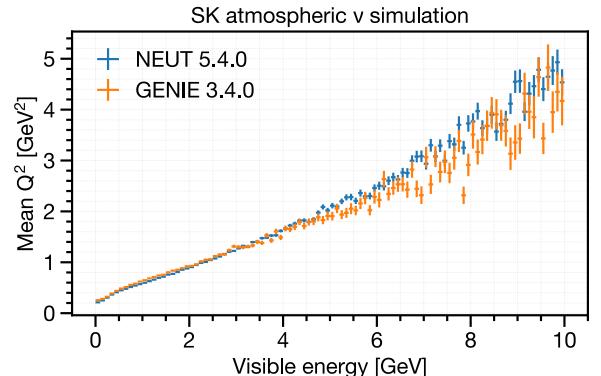


FIG. 4. Average squared momentum transfer (Q^2) per visible energy bin for simulated atmospheric neutrino events at SK, compared between two different neutrino event genera-
tors. NEUT 5.4.0 (blue) uses the baseline setup described in
Section III C, while GENIE 3.4.0 [23–25] uses the “hN” setup
as described in Section VI.

293 muons) produced via the primary neutrino interaction,
294 we followed the typical Cherenkov ring reconstruction
295 process [26] as applied in previous SK analyses [5, 6].

296 The visible energy of an event is defined as the sum of
297 the reconstructed kinetic energies of all Cherenkov rings,
298 assuming each ring originates from an electron. This
299 calorimetric measure serves as a reliable proxy for neu-
300 trino momentum transfer, as demonstrated by the pos-
itive correlation shown in Figure 4.

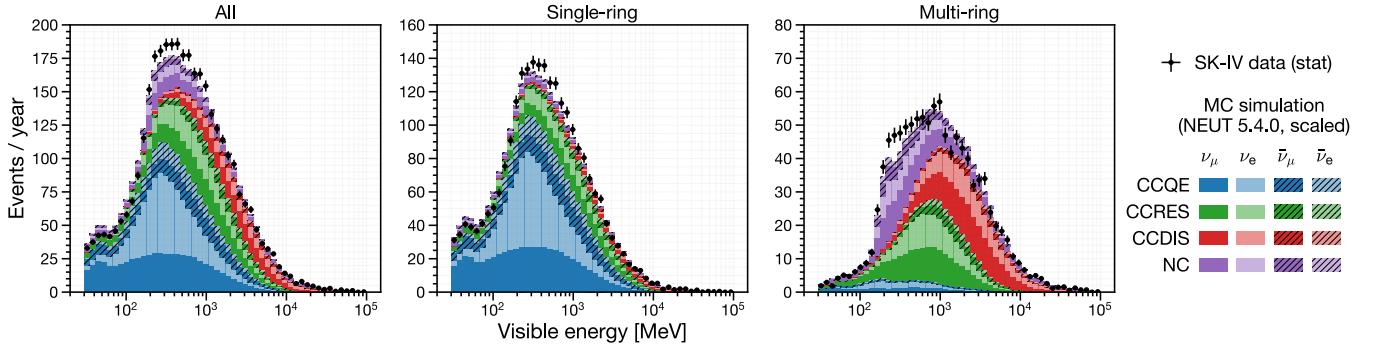


FIG. 5. Visible energy distributions for observed and simulated atmospheric neutrino interactions in water, fully contained within the fiducial volume: all events (left), events with one reconstructed Cherenkov ring (middle), and events with two or more rings (right). CCQE events dominate the single-ring sample, while RES and DIS events dominate the multi-ring sample. The simulation setup is detailed in Section III C. Simulated distributions are scaled to match the SK-IV data to account for neutrino flux uncertainties.

C. Baseline simulation setup

The simulation of atmospheric neutrino events involves a convolution of the atmospheric neutrino flux, neutrino event generator, and detector simulation.

To determine the incoming neutrino kinematics and event rate, we used atmospheric flux calculations for ν_e , ν_μ , $\bar{\nu}_e$, and $\bar{\nu}_\mu$ at the detector site, assuming no oscillations, as provided by Honda et al. [27] for neutrino energies ranging from 100 MeV to 10 TeV. Interactions of ν_τ and $\bar{\nu}_\tau$ were not simulated.

Neutrino interactions in water were modeled using neutrino event generators [28], which compute the cross sections for each interaction channel and sample outgoing particle kinematics. NEUT 5.4.0 [29] was used as the baseline event generator. It covers major interaction channels, including QE scattering with single (1p1h) or double (2p2h) nucleon knockout, single pion production due to Δ resonance (RES), and multiple pion production via deep inelastic scattering (DIS). The models used are consistent with the latest SK atmospheric neutrino oscillation analysis [5].

The transport of hadronic final states within the target nucleus (FSI) was modeled separately from the primary neutrino interaction. The transport of low-momentum (< 500 MeV/c) pions was based on Salcedo et al. [30] with parameters tuned to fit external pion-nucleus scattering data [31]. For higher energy nucleons and pions, FSIs were modeled using the INC approach, which includes elastic scattering and single/double pion production. The cross sections for reactions with free target nucleons were sourced from Bertini [32] for nucleon projectiles and from external pion scattering data for high-momentum pion projectiles [33]. The de-excitation of an oxygen target was modeled using tabulated occupation probabilities for nucleon energy states [34] and branching ratios for knockout of each state [35].

The subsequent particle transport in water and detector responses were simulated using GCALOR [36] cou-

pled with Geant3.21 [37]. Propagation of nucleons and charged pions below 10 GeV was simulated based on the Bertini cascade model [38]. For cross sections, GCALOR uses ENDF/B-VI [39] for low-energy (below 20 MeV, or 195 MeV/c in momentum) neutrons, and Bertini-based tabulation [38] for the remaining nucleons and charged pions. The transport of low-momentum pions was separately modeled using the NEUT pion FSI routine to ensure consistency between NEUT and Geant3.

For the transport of low-energy neutrons in the Gd-loaded SK-VI phase, we imported reaction cross sections from the Geant4.10.5.p01 [7–9] NeutronHP database [40] based on ENDF/B-VII.1 [17], replacing ENDF/B-VI for neutron energies below 20 MeV. Additionally, we modeled the gamma-cascade resulting from neutron captures on $^{155}/^{157}\text{Gd}$ using the ANNRI-Gd model [41, 42].

The characteristics of individual PMTs and the optical parameters in water were adjusted to align with calibration data obtained from light sources and through-going cosmic-ray muons. To accurately account for detector noise, randomly recorded PMT hits were included as background, represented as gray bars in Figure 3.

500 years of atmospheric neutrino events were simulated for each SK phase and processed as described in Sections III B and III A. The events in the final sample were weighted based on the standard three-flavor oscillation probability in matter [43], using the oscillation parameters fitted with reactor constraints in the previous SK analysis [44] and the PREM model [45] for Earth’s matter density. Corrections for atmospheric neutrino flux accounting for solar activity were also applied.

Figure 5 shows the visible energy distributions of the simulated event sample, demonstrating reasonable agreement with the observations. Single-ring events exhibit a higher fraction of QE interactions and a smaller fraction of DIS compared to multi-ring events. For the same visible energy, multi-ring events are expected to produce more secondary neutrons than single-ring events, due to a higher fraction of DIS interactions.

378 **IV. NEUTRON SIGNAL SELECTION** 432

379 Neutrons produced by atmospheric neutrino interactions in the SK ID are primarily captured by ^1H or 434
 380 $^{155}/^{157}\text{Gd}$ isotopes within approximately $O(100)$ μs , with 435
 381 expected capture ratios summarized in Table I. Electrons 436 that are scattered by the $O(1)$ MeV gamma-rays generate 437 Cherenkov photons that form faint rings on the ID tank wall 438 (hereafter referred to as “neutron signal”), as shown in Figure 3. 439

387 Fast neutrons, along with the subsequent gamma-rays 440 and scattered electrons, typically travel only a few tens 441 of centimeters in water. Thus, signal photons can be approximated as originating from a single vertex. While 442 reconstructing this vertex is challenging due to the limited 443 number of PMT hits, *assuming* it to be near another known 444 vertex, such as the reconstructed neutrino interaction 445 vertex, helps identify signals from random PMT coincidences 446 caused by dark current. Remaining Michel electrons from muon decays can be effectively suppressed 447 through cuts based on signal timing and energy. 448

450 **A. Signal selection algorithm** 451

399 The signal selection algorithm is based on Ref. [46], 452 and consists of two stages. 453

401 In the first candidate search stage, we initially subtract 455 the expected photon time-of-flight (ToF) from the individual 456 PMT hit times for a given *assumed* signal vertex, 457 which was provided by the reconstructed neutrino event 458 vertex. Then, we slide a time window of 14 ns width 459 on the ToF-corrected PMT hit times to trigger on the 460 number of included PMT hits. The threshold was 5 for 461 pure water phases SK-IV and V, and 7 for the Gd-loaded 462 phase SK-VI. The search span for each event was [18, 534] 463 μs from the event trigger for SK-IV and V, and [3, 534] 464 μs for SK-VI with a shorter neutron capture time. For 465 overlapping candidates within 50 ns, only the candidate 466 with the largest number of PMT hits is selected. 467

468 In the second candidate classification stage, we extract 469 features of each candidate and use a neural network to 470 classify each candidate into signal and noise based on 471 input features. These features characterize the signal 472 energy, the background hit level, timing spread assuming 473 the vertex, correlation between the input vertex and the 474 hit PMT positions, correlation to the known properties 475 of PMT noise, and angular correlation among hit PMTs 476 relative to the Cherenkov cone opening angle. 477

478 The major changes from the original algorithm [46] include 479 a simplified algorithm, a reduced set of features, 480 and a heuristically tuned neural network architecture. 481 These modifications aim to reduce performance bias between 482 the data and the simulation that is used to train 483 the neural networks. 484

485 Here, we provide the definition and unit of each feature 486 used for the classification of signal candidates, along with 487 their expected distributions as shown in Figure 6. 488

- **NHits**

The number of selected PMT hits within the 14-ns sliding time window.

- **NResHits**

The number of PMT hits within [-100,+100] ns from the center of the 14-ns sliding time window, subtracted from NHits.

- **TRMS** [ns]

The root mean square (RMS) of the ToF-corrected time distribution of the selected PMT hits.

- **FitGoodness**

The normalized likelihood of the ToF-corrected time distribution of the selected PMT hits, given the *assumed* signal vertex and the Gaussian PMT timing resolution of 5 ns.

- **DWall** [cm]

The distance from the *assumed* signal vertex to the nearest tank wall.

- **DWallMeanDir** [cm]

The shorter of the radial and vertical distances from the *assumed* signal vertex to the tank wall, weighted by the mean of the unit vectors connecting the vertex to each hit PMT.

- **BurstRatio**

The ratio of the selected PMTs with a preceding hit within 10 μs , which are likely caused by scintillation within the irradiated PMT glass.

- **DarkLikelihood**

The normalized log likelihood ratio based on measured individual PMT dark rates, given by:

$$\text{DarkLikelihood} = \sigma \left(\log \prod_{i=1}^{\text{NHits}} \frac{r_i}{\langle r \rangle} \right) \quad (1)$$

where σ represents the sigmoid function, r_i is the dark rate of the i^{th} PMT, and $\langle r \rangle$ is the average dark rate of all ID PMTs.

- **OpeningAngleStdev** [deg]

The standard deviation of the opening angles of cones formed by every possible combination of three hit PMTs and the *assumed* signal vertex.

- **Beta(k)**, $k \in \{1, 2, 3, 4, 5\}$

$$\text{Beta}(k) = \frac{2}{\text{NHits}(\text{NHits} - 1)} \sum_{i \neq j} P_k(\cos \theta_{ij}) \quad (2)$$

where P_k is the k^{th} Legendre polynomial and θ_{ij} is the opening angle between the assumed signal vertex and the i^{th} and j^{th} hit PMTs.

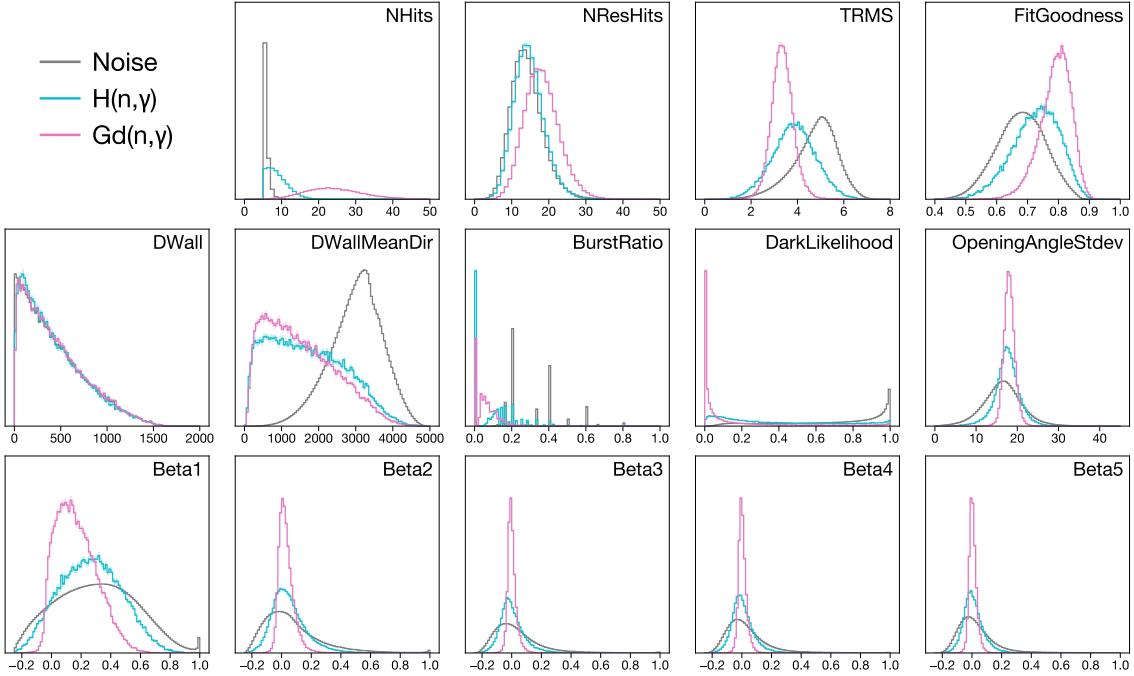


FIG. 6. The features (area-normalized) of neutron signals and noise from the thermal neutron MC simulation used for training the neural network for SK-VI phase.

For training the neural network, we utilized the features and labels of first-stage candidates from a particle-gun simulation of thermal neutrons, with vertices randomized within the ID. Each first-stage candidate, as shown in Figs. 6 and 7, was labeled as signal if it was triggered within 50 ns of the simulated (n, γ) reaction, and otherwise as background. Of the training dataset, 80% was used for updating the neural network weights, while the remaining 20% was reserved for validation.

We implemented a feed-forward fully connected neural network using Keras 2.6.0 [47]. The network consisted of an input layer with 14 features, followed by three dense layers, each comprising 128 ReLU-activated nodes with a 50% dropout rate, and a single sigmoid output node. Weights and biases were initialized following He et al. [48] and optimized by minimizing the binary cross-entropy loss iteratively on minibatches of size 2,048 using the Adam optimizer [49]. The initial learning rate was set to 0.0001. Training was stopped when signal efficiency on the validation set showed no improvement for 5 consecutive epochs. A neural network was trained for each SK phase: SK-IV, SK-V, and SK-VI.

Candidates with a neural network output greater than 0.7 were classified as signals, while those with a large number of PMT hits ($\text{NHits} > 50$) and occurring earlier than the typical neutron capture timescale ($< 20 \mu\text{s}$) were identified as Michel electrons and excluded. The effectiveness of this Michel electron rejection is illustrated in Figure 7. When applied to cosmic-ray muons decaying within the ID, the selection achieved an efficiency of $98.4 \pm 1.3\%$ and a purity of $98.7 \pm 0.5\%$.

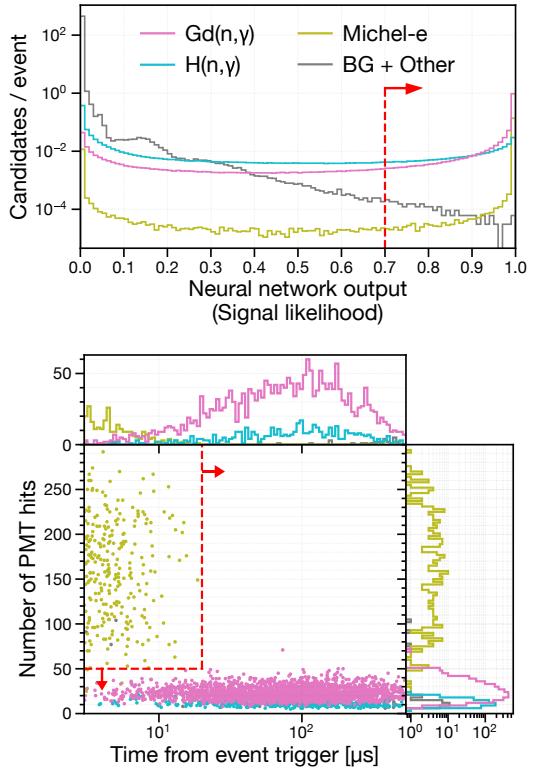


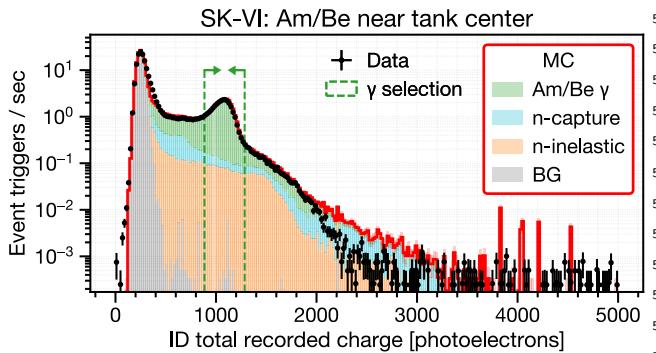
FIG. 7. The neural network response to neutron signals and backgrounds in the SK-VI atmospheric neutrino simulation (top), and the time versus energy distribution of signal candidates passing the neural network selection (bottom). Red arrows indicate the corresponding cut points.

506 B. Signal selection performance on calibration data

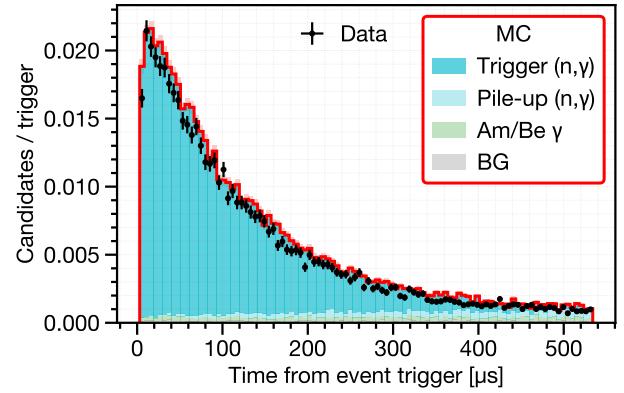
507 An Am/Be neutron source with a measured total intensity
 508 of 236.8 ± 5.0 neutrons/s [50] was used to obtain cal-
 509 ibration data for estimating the signal detection perfor-
 510 mance. The first excited state of the alpha-absorbed ${}^9\text{Be}$,
 511 with a roughly 60% branching ratio, emits a fast neutron
 512 and a 4.44 MeV gamma-ray simultaneously. This source
 513 was encapsulated with Bismuth Germanate ($\text{Bi}_4\text{Ge}_3\text{O}_{12}$,
 514 BGO) crystals so that the 4.44 MeV gamma-rays can
 515 induce scintillation. The setup was deployed in various
 516 positions within the ID, and events were recorded for 30
 517 minutes to 1 hour. Events with a trigger charge yield
 518 corresponding to the 4.44 MeV gamma-ray scintillation
 519 were regarded as the single neutron control sample.

520 The observed light yield distribution was compared
 521 with dedicated simulation, as shown in Figure 8. The
 522 simulation accounts for continuous source activity and
 523 pile-up, by reorganizing the simulated detector response
 524 to Am/Be neutron emission on a single global time axis,
 525 based on the measured total neutron intensity and the es-
 526 timated branching ratios to each excited state of alpha-
 527 absorbed ${}^9\text{Be}$. As shown in Figure 8, this simulation
 528 accurately models event triggers due to ambient neutron
 529 captures and neutron inelastic interactions within scintil-
 530 lator crystal. The contamination of such unwanted event
 531 triggers in the 4.44 MeV gamma-ray event selection was
 532 estimated to be at a few percent level.

533 Within the selected events in the single neutron con-
 534 trol sample, signal candidates were obtained following the
 535 algorithm described in Section IV A, with the *assumed*
 536 signal vertex set at the source position.



556 FIG. 8. Distribution of recorded charge within the time
 557 window $[-0.5, 1.0]$ μs relative to event triggers, measured with
 558 an Am/Be neutron source positioned near the ID tank cen-
 559 ter. The black points represent SK-VI data, while the red
 560 line shows the simulated prediction. The green dashed ar-
 561 rows indicate the selection window for 4.44 MeV gamma-ray-
 562 induced scintillation. Roughly 95% of the selected events are
 563 attributed to gamma-rays from the Am/Be source (green),
 564 with the remaining 5% arising from ambient neutron cap-
 565 tures (blue) and neutrons inelastically producing charged par-
 566 ticles in the scintillator (orange). The contribution from back-
 567 ground events (labeled “BG,” gray) is negligible.



568 FIG. 9. Exponential decrease of the selected neutron signal
 569 candidates as a function of the time from the selected event
 570 triggers with the Am/Be neutron source positioned near the
 571 ID tank center, in the SK-VI phase. The label “Trigger (n, γ)”
 572 indicates captures of neutrons produced within 350 ns from
 573 the event trigger, while the label “Pile-up (n, γ)” indicates
 574 captures of piled-up neutrons without such correlation to the
 575 event trigger.

576 Figure 9 shows an example time distribution of the
 577 selected signal candidates. Such distributions of the time
 578 t were fitted with a function f of the form:

$$f(t) = A(1 - e^{-t/\tau_{\text{thermal}}})e^{-t/\tau_{\text{capture}}} + B \quad (3)$$

579 where the normalization constant A , the background
 580 constant B , the neutron thermalization time scale
 581 τ_{thermal} , and the neutron capture time constant τ_{capture}
 582 are free parameters. The signal efficiency was evaluated
 583 as the number of identified signals per selected event trig-
 584 gers, corrected by the constant background term B . Figure
 585 10 shows the estimated neutron detection efficiencies
 586 for various source positions in the ID.

587 The major sources of systematic uncertainty are sum-
 588 marized in Table II. In the pure water phases (SK-IV
 589 and V), the dominant source of uncertainty lies in the
 590 potential bias caused by the calibration setup, such as
 591 the unwanted event triggers or time correlation of false
 592 positives to event triggers, often leading to an overesti-
 593 mation of the background constant B . The size of this
 594 uncertainty was conservatively estimated by comparing
 595 the true and estimated signal efficiencies from the sim-
 596 ulations and quantifying the fluctuation within each SK
 597 phase. In the Gd-loaded phase (SK-VI), the dominant
 598 source of uncertainty is in the fraction and the γ emis-
 599 sion model of the $Gd(n, \gamma)$ reaction. The size of these un-
 600 certainties is estimated based on the evaluated thermal
 601 neutron capture cross-section uncertainties in ENDF/B-
 602 VII.1, as well as variations in the estimated signal effi-
 603 ciency when using an alternative photon strength func-
 604 tion to describe the Gd continuum gamma cascade in the
 605 ANNRI-Gd model [41].

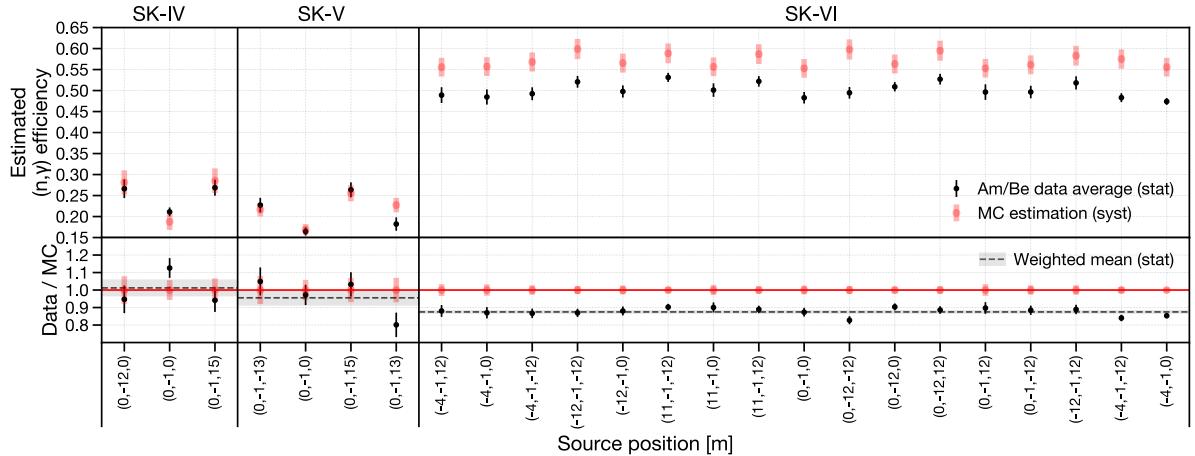


FIG. 10. Estimated (n, γ) signal selection efficiency for each calibration position within the tank. For positions in SK-VI where there are multiple measurements at different dates, the estimated efficiencies were averaged. The systematic uncertainties considered for MC simulation are summarized in Table II.

TABLE II. Major sources of systematic uncertainty in the neutron detection efficiency estimated with the Am/Be neutron source.

Source	SK-IV	SK-V	SK-VI
Am/Be neutron characterization	0.5%	0.9%	0.5%
Detector response	2.2%	3.3%	1.2%
Bias due to calibration setup	6.9%	4.6%	1.1%
Gd(n, γ) fraction	-	-	2.1%
Gd(n, γ) γ emission model	-	-	2.6%
Total	7.3%	5.7%	3.8%

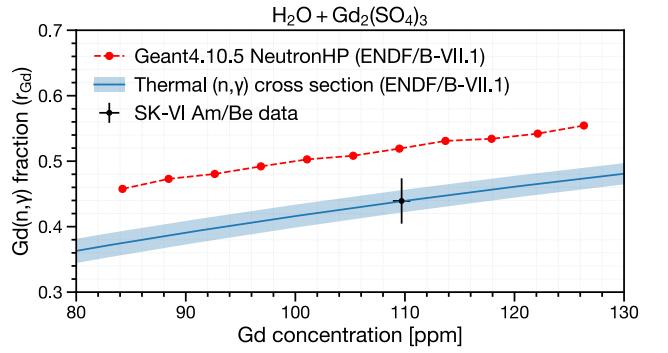


FIG. 11. Predicted Gd(n, γ) fraction in Gd-loaded water as a function of Gd concentration, compared with SK-VI Am/Be data (black). The red dots represent predictions by MC simulation using the NeutronHP model in Geant4.10.5.p01 with neutron cross sections from ENDF/B-VII.1, while the blue line and shades represent evaluation of Equation 4 based on the evaluated thermal (n, γ) cross sections and uncertainties in ENDF/B-VII.1.

The noticeable difference between the observed and predicted signal efficiencies in SK-VI is thought to stem from the overestimation of the Gd(n, γ) fraction in our MC simulation setup. Figure 11 compares the MC-simulated Gd(n, γ) fraction r_{Gd} with the analytically evaluated fraction, assuming completely thermalized neutrons:

$$r_{\text{Gd}} \approx 1 - r_{\text{H}} \approx 1 - \frac{n_{\text{H}} g_{\text{H}}(T) \sigma_{\text{H}}(v_{\text{thermal}})}{\sum_i n_i g_i(T) \sigma_i(v_{\text{thermal}})} \quad (4)$$

Here, for the i^{th} isotope, n_i is the number density, $g_i(T)$ is the Westcott g -factor for temperature T , and $\sigma_i(v_{\text{thermal}})$ is the neutron capture cross section evaluated at thermal neutron speed $v_{\text{thermal}} = 2200$ m/s.

The value estimated with SK-VI Am/Be data agrees well with the analytical evaluation based on g_i and σ_i from ENDF/B-VII.1 [51], while both differ from the MC simulation using Geant4.10.5.p01 NeutronHP and ENDF/B-VII.1. The reason seems to be that it considers the thermal motion of hydrogen as free rather than bound in a water molecule [52], thus underestimating the fraction of ${}^1\text{H}(n, \gamma)$ that competes with the Gd(n, γ) reaction.

To account for such a difference between signal efficiencies evaluated with Am/Be data and MC simulation, we used the weighted mean of the ratios of the two evaluated at all Am/Be source positions as the correction factor. The obtained efficiency correction factor was 1.01 ± 0.04 for SK-IV, 0.96 ± 0.04 for SK-V, and 0.88 ± 0.01 for SK-VI, including both statistical and systematic errors.

The average neutron capture time constants were measured as 200.4 ± 3.7 μs for the pure water phase (SK-IV, V) and 116.9 ± 0.3 μs for the Gd-loaded phase (SK-VI). These results are consistent with ENDF/B-VII.1 predictions of 204.7 ± 5.3 μs and 114.9 ± 2.5 μs , where the errors are derived from evaluated cross section uncertainties. Relative to this, the Geant4 prediction for the Gd-loaded phase was 112.4 μs .

604 V. (n, γ) MULTIPLICITY ESTIMATION

605 The average multiplicity of (n, γ) reactions, or neutron captures, is computed as the average of the expected number of (n, γ) reactions estimated on an event-by-event basis, as follows:

610 Here, N_i^{detected} is the count of detected signals, N_i^{BG} is the estimated number of false positives, and ϵ_i is the estimated signal detection efficiency of the i^{th} event.

615 Accurate estimation of N_i^{BG} and ϵ_i is crucial. While Am/Be calibration data provides a basis for these estimates, additional factors in atmospheric neutrino events—such as neutrino vertex reconstruction accuracy and larger neutron kinetic energy—may significantly impact the signal selection performance. To better account for these effects in the calculation of N_i^{BG} and ϵ_i , we trained Generalized Additive Models (GAMs) [53] on the baseline atmospheric neutrino simulations. Using pyGAM 0.9.0 [54], a total of six linear GAMs were constructed across the three SK phases, for two output metrics: signal efficiency and false positive rate.

620 The input features for the analysis included five reconstructed neutrino event variables: visible energy, Cherenkov ring multiplicity, the particle type of the most energetic ring, and the radial and vertical displacements of the neutrino interaction vertex. GAMs were fitted

625 against the distributions of the average simulated output metric per bin in the five-dimensional feature space. No assumptions were made regarding feature correlations, and appropriate smoothing was applied to prevent overfitting. The 1σ prediction interval was derived based on the statistical uncertainties within each bin.

630 GAMs were trained on the baseline MC simulation generated with NEUT 5.4.0, and tested on a simulation using NEUT 5.1.4, which predicts roughly 10% lower (n, γ) multiplicity overall. Figure 12 shows the performance of the trained GAM in estimating the true average signal multiplicity per visible energy bin. By reconstructing N_i^{BG} and ϵ_i on an event-by-event basis, the GAM helps reduce potential biases in the results, particularly in the multi-GeV bins.

635 The following major systematic uncertainties affecting signal counting were evaluated on a bin-by-bin basis for each operational phase and data subsample:

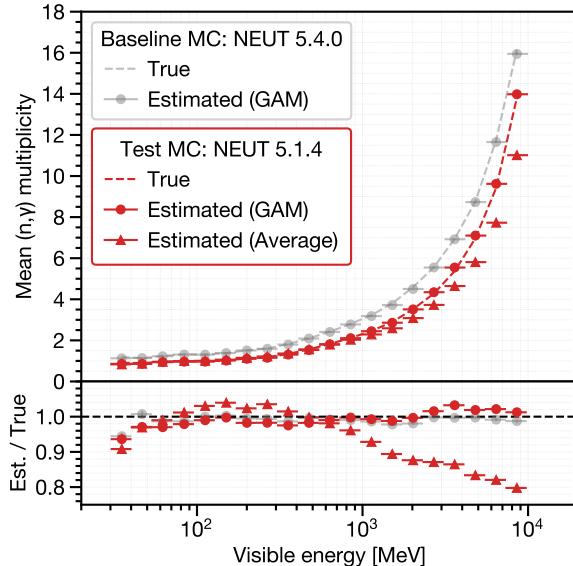


FIG. 12. The true (dashed lines) and estimated average (n, γ) multiplicity (circle markers: using GAM, triangle markers: using the overall average signal efficiency and false positive rate) as a function of visible energy, for baseline and test simulations of atmospheric neutrino events.

(1) Overall signal efficiency scale

This includes calibration uncertainties (Section IV B) and phase-dependent variations (to be explained in Section VII).

(2) Signal selection performance modeling (GAM)

This is quantified as the difference between true and estimated signal multiplicities in simulations. The ratio is also used to correct the estimated distribution.

(3) Neutrino event reconstruction

This assumes 2% visible energy resolution (Figure 2), with ring-counting errors accounted for single- and multi-ring events.

Figure 13 shows fractional uncertainties per visible energy bin for the full data sample. The most dominant contributions are signal efficiency scaling and statistical errors.

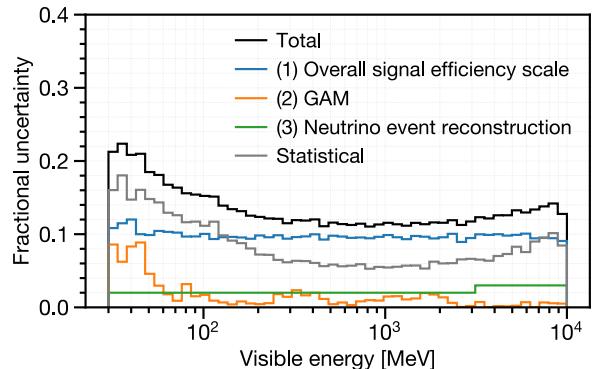


FIG. 13. Fractional uncertainties assigned to the estimated average (n, γ) multiplicity per bin for the full data sample.

664 VI. TESTED INTERACTION MODELS

665 To compare with data, we generated predictions for the
 666 average (n, γ) multiplicity as a function of visible neutrino
 667 energy using various neutrino event generators and
 668 hadron-nucleus interaction models. Six generator config-
 669 urations were tested: NEUT 5.4.0, NEUT 5.6.3, GENIE
 670 hA, GENIE hN, GENIE BERT, and GENIE INCL. The GENIE
 671 setups use GENIE 3.4.0 with G_18a_10x_02_11b physics
 672 tunes [25], where $x \in \{a, b, c, d\}$ corresponds to the FSI
 673 model: INTRANUKE/hA [24] (hA), INTRANUKE/hN
 674 [24] (hN), the Geant4 Bertini cascade model [55] (BERT),
 675 and the Liège INC model [56] (INCL).

676 NEUT 5.4.0 follows the setup in Section III C, while
 677 NEUT 5.6.3 includes a modified nuclear binding energy,
 678 slightly reducing CCQE interactions. The GENIE setups
 679 share QE and single-pion production models with NEUT
 680 but differs in FSI and hadronization. Except for GENIE
 681 hA, all FSI models use the full INC approach based on
 682 free nucleon cross sections. The Liège INC model is cou-
 683 pled with ABLA07 [57] for nuclear de-excitation. Figure
 684 14 compares outgoing neutron momentum distributions,
 685 revealing notable differences below 1 GeV/c, which are
 686 further discussed in Section VIII.

687 Secondary hadron-nucleus interaction models were
 688 tested using SK detector simulations built with
 689 Geant3.21 and Geant4.10.5.p01. Six configurations were
 690 tested: SK-IV/V default, SK-VI default, G3 GCALOR,
 691 G4 BERT, G4 BERT_PC, and G4 INCL_PC. The SK-IV/V
 692 default and SK-VI default setups primarily differ in
 693 their neutron reaction cross sections below 20 MeV, us-
 694 ing ENDF/B-VI and ENDF/B-VII.1, respectively. G3
 695 GCALOR relies entirely on GCALOR within Geant3.21,
 696 unlike the baseline setups that use the NEUT pion FSI
 697 routine. Geant4-based models (G4 BERT, G4 BERT_PC,
 698 and G4 INCL_PC) adopt ENDF/B-VII.1 for low-energy
 699 neutron transport. For intranuclear hadron transport,
 700 G4 BERT and G4 BERT_PC use the Geant4 Bertini cascade,
 701 while G4 INCL_PC uses the Liège INC model. Configura-
 702 tions with the PC suffix uses the Geant4 Precompound
 703 model [58] for nuclear de-excitation, whereas G4 BERT
 704 uses a simpler native model [55]. For hadron-nucleus
 705 reaction cross sections (excluding low-energy neutrons),
 706 Geant3-based models rely on Bertini-based tabulation
 707 [38], while Geant4-based models use Glauber parameter-
 708 ization [59]. Figure 15 compares the predicted average
 709 signal multiplicities as a function of projectile neutron
 710 momentum, showing up to 40% variation.

711 For each neutrino event generator, a full MC simula-
 712 tion equivalent to 50 years of atmospheric neutrino ex-
 713 posure was generated using the baseline detector model
 714 (SK-VI default). To account for variations in secondary
 715 hadron-nucleus interactions without additional full simu-
 716 lations, we precomputed average signal multiplicities for
 717 neutron-producing projectiles (n, p, π^\pm , and μ^-) in water
 718 up to 10 GeV/c, based on particle-gun MC simulations.
 719 Figure 15 shows an example tabulation for neutron pro-
 720 jectiles.
 730

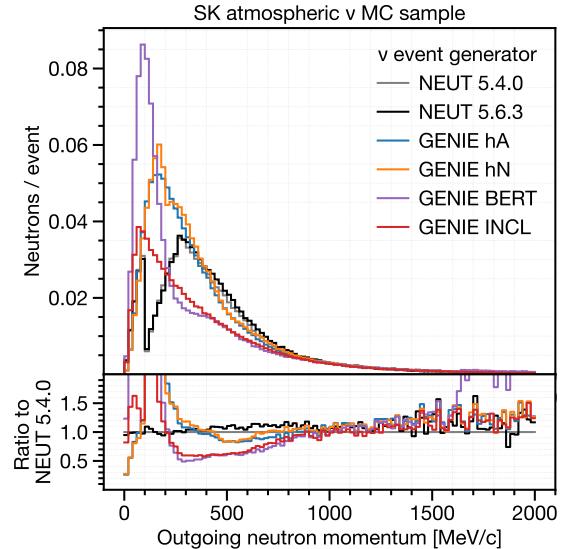


FIG. 14. Model comparison: Outgoing neutron momentum distribution per atmospheric neutrino event in the sample.

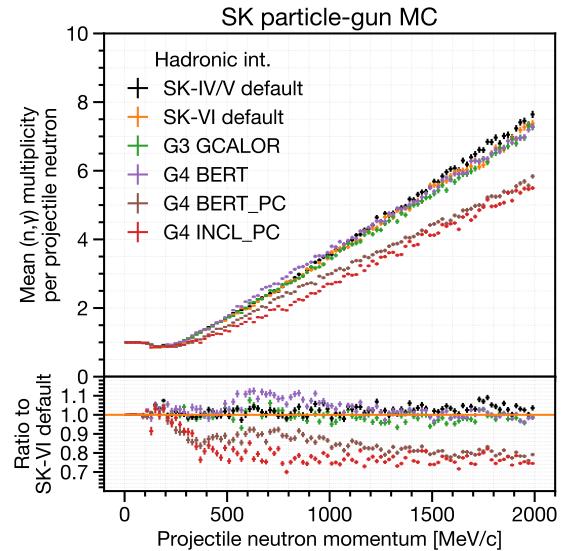


FIG. 15. Model comparison: Average (n, γ) multiplicity per projectile neutron momentum bin, based on simulations using a single-neutron particle-gun setup in water.

Predictions were made by convolving hadron momentum distributions with momentum-to-multiplicity tables. For each event generator, ratios relative to the baseline detector simulation model (SK-VI default) were used to scale the signal multiplicities obtained from the full MC simulation. In configurations using either NEUT 5.4.0 or NEUT 5.6.3 with SK-IV/V default, the discrepancy between this prediction method and the 500-year exposure full MC simulation was within 5% across all visible energy range and data subsamples.

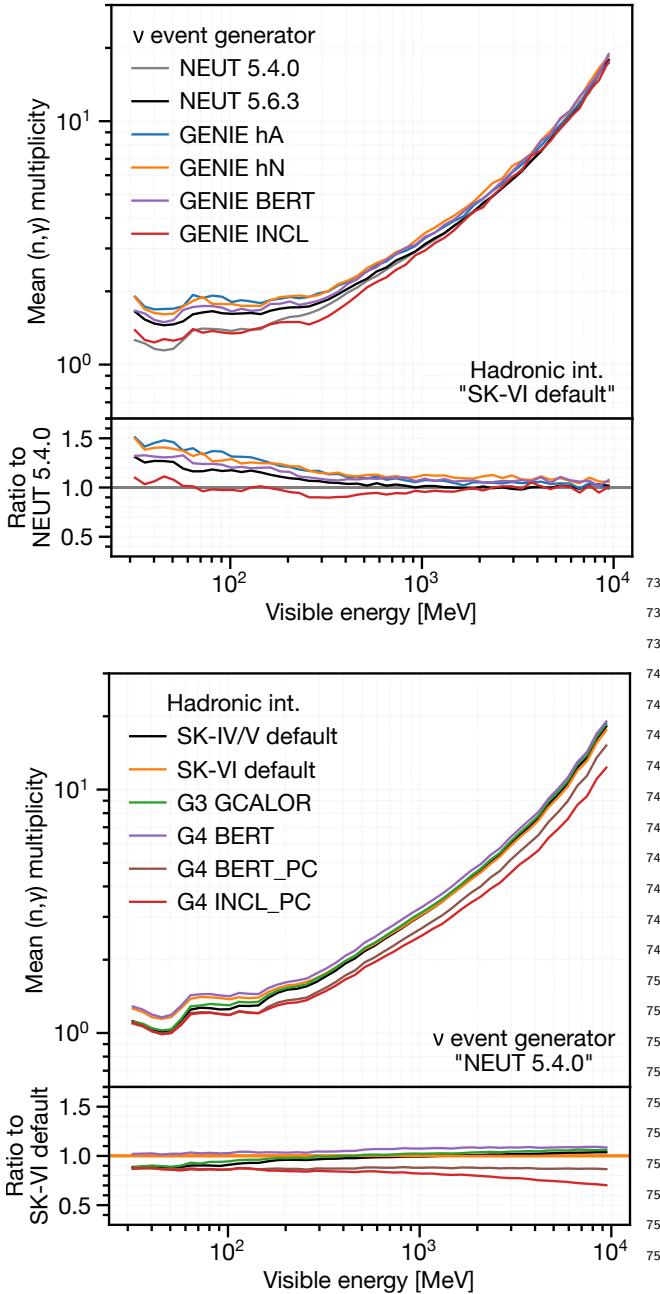


FIG. 16. Comparison of the neutrino event generator options (top) and secondary hadron-nucleus interaction models (bottom) on the predicted average (n, γ) multiplicity. In the top figure, each neutrino event generator option is paired with the baseline secondary interaction model (SK-VI default). In the bottom figure, each model is paired with the baseline neutrino event generator option (NEUT 5.4.0).

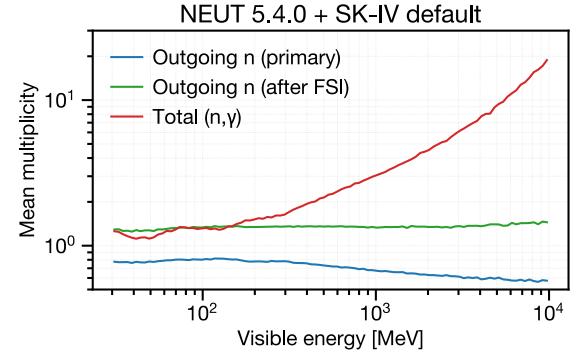


FIG. 17. Mean multiplicity of outgoing neutrons from the primary neutrino interaction (blue) and subsequent FSIs (green) and resulting total (n, γ) reactions (red), in the baseline SK-IV simulation setup.

nucleus interaction models remains relatively constant. GENIE INCL predicts fewer neutrons than the other FSI models. NEUT 5.4.0 predicts fewer neutrons than NEUT 5.6.3, due to a larger QE fraction. Additionally, the Geant4 Bertini cascade model with the Precompound model (G4 BERT_PC) predicts fewer neutrons than when using the native de-excitation model (G4 BERT).

Figure 17 shows the sources of neutron production in the baseline SK-IV simulation as a function of visible energy, offering insight into the origins of neutron production across different neutrino energy regions. The number of outgoing neutrons from neutrino-nucleus interactions is expected to be nearly constant regardless of visible energy. In events with lower visible energies, the momentum of the outgoing neutrons is small, so each outgoing neutron corresponds almost directly to a single (n, γ) reaction. At higher energies, the outgoing hadrons gain more energy, leading to a linear increase in secondary neutron production. This suggests that the (n, γ) multiplicity in lower visible energies is sensitive to the modeling of low-energy nucleon transport within the nucleus, while the linear increase is primarily determined by the choice of secondary hadron-nucleus interaction model.

VII. RESULTS

Table III summarizes the number of atmospheric neutrino events (" ν events") and detected neutron signals (" n signals") in the final data sample. The overall average (n, γ) multiplicity per neutrino event, $\langle N \rangle_{\text{overall}}$, is estimated using the signal efficiency correction from Equation 5 for all events. To account for uncertainties in outgoing neutron kinematics that affect detection efficiency, a separate signal detection algorithm was applied to SK-VI data. This algorithm uses a likelihood-based vertex fitter [60] to reconstruct the $\text{Gd}(n, \gamma)$ vertex, independently from the neutrino interaction vertex. The results from this method serve as a reference for the baseline neutron detection algorithm described in Section IV A.

Figure 16 shows the model predictions. In visible energy range above a few hundred MeV, the average (n, γ) multiplicities are expected to linearly increase with visible energy. Differences across neutrino event generator configurations are more pronounced at visible energies below 1 GeV, while variability among secondary hadron-

TABLE III. Summary of atmospheric neutrino events and detected neutron signals in the final sample. $\langle N \rangle_{\text{overall}}$ is the unbinned application of the signal efficiency correction given by Equation 5, estimated using signal efficiency scales from Am/Be calibration (Section IV B). The “SK-VI (Reference)” column shows SK-VI results using a neutron-energy-independent algorithm. Errors for the observed $\langle N \rangle_{\text{total}}$ are listed as statistical first, followed by systematic uncertainty, which mainly arises from signal efficiency calibration. Other errors are statistical only. The bottom two rows show the expected $\langle N \rangle_{\text{overall}}$ and the true overall (n, γ) multiplicity extracted from the baseline full MC simulations.

	SK-IV	SK-V	SK-VI	SK-VI (Reference)
ν events	29,942	4,231	5,203	Same as SK-VI
Events/day	9.23 ± 0.05	9.18 ± 0.14	9.22 ± 0.13	Same as SK-VI
n signals	15,705	2,035	5,752	4,412
n signals/event	0.525 ± 0.004	0.481 ± 0.011	1.106 ± 0.015	0.848 ± 0.017
Observed $\langle N \rangle_{\text{overall}}$	$2.21 \pm 0.03 \pm 0.11$	$2.46 \pm 0.10 \pm 0.11$	$2.50 \pm 0.06 \pm 0.05$	$2.49 \pm 0.06 \pm 0.05$
Expected $\langle N \rangle_{\text{overall}}$	2.83	2.84	2.85	2.87
True (n, γ) multiplicity	2.85	2.85	2.86	Same as SK-VI

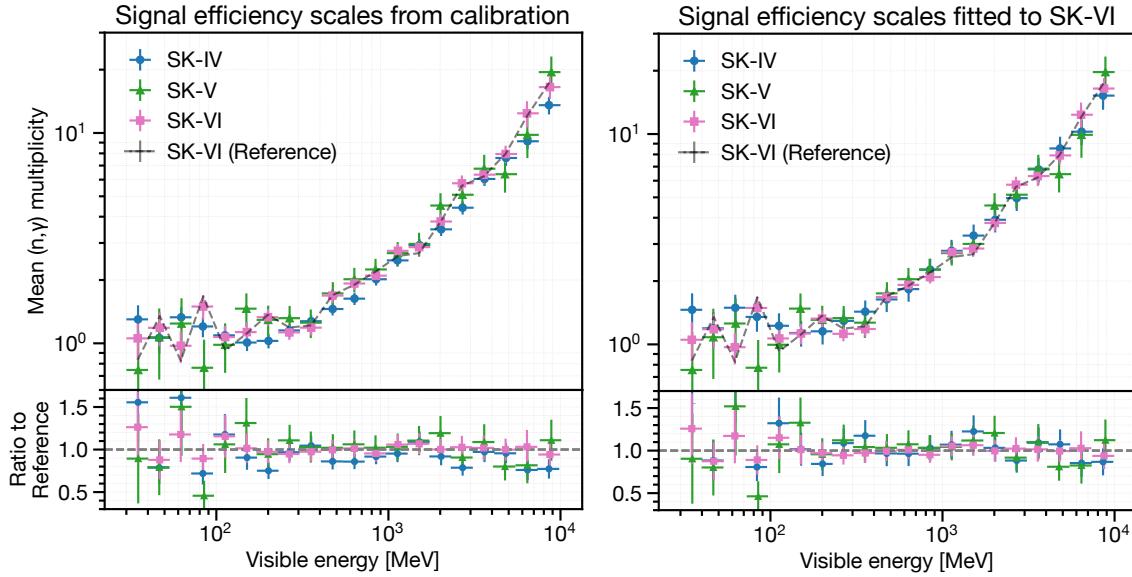


FIG. 18. Average (n, γ) multiplicities plotted against neutrino event visible energy across SK phases. The left plot uses signal efficiency scales from Am/Be calibration (Section IV B), while the right plot adjusts SK-IV and SK-V scales to match $\langle N \rangle_{\text{total}}$ of SK-VI. Error bars include statistical and systematic uncertainties, with the right plot also accounting for scale differences. The dashed “SK-VI (Reference)” line shows SK-VI data using a neutron-energy-independent reference algorithm.

We find that the observed $\langle N \rangle_{\text{overall}}$ values are smaller than expected by full MC simulation. Also, $\langle N \rangle_{\text{overall}}$ is approximately 2.5 in SK-V and SK-VI but significantly lower in SK-IV at 2.21 ± 0.03 (stat) ± 0.11 (syst). While this systematic uncertainty takes into account the difference between Am/Be calibration data and simulations as shown in Figure 10, there may exist unaccounted systematic error sources in SK-IV, as it has the longest data-taking period (10 years) but has only a limited amount of calibration data to constrain the entire period. In contrast, SK-VI with a shorter runtime (2 years) periodically took calibration data across the full detector volume. Thus, the signal efficiency scale for SK-VI is considered more reliable than that of SK-IV.

In subsequent figures, we adjust the signal efficiency correction factors to align $\langle N \rangle_{\text{overall}}$ with the SK-VI reference value ($\langle N \rangle_{\text{overall}} = 2.49$) that is derived using the aforementioned neutron-energy-independent reference signal detection algorithm. The difference between calibration-based and phase-consistency-based correction factors is treated as an independent systematic uncertainty in the signal efficiency scale, as shown in Figure 13. This results in a final $\langle N \rangle_{\text{overall}}$ of 2.49 ± 0.03 (stat) ± 0.26 (syst), corresponding to a systematic uncertainty of approximately 10%. Figure 18 illustrates the impact of the two signal efficiency correction methods: one based on Am/Be neutron source calibration and the other on the overall phase consistency.

The combined data was compared with the SK-VI MC simulation and neutron production estimates in water from SNO [13], as shown in Figure 19. While the combined data matched well with the SNO estimate, it was 10-30% lower than the SK-VI simulation results in the sub-GeV energy range. For neutrino events with visible energy greater than a few hundred MeV, where the Cherenkov rings of outgoing leptons are accurately reconstructed, a linear relationship was observed between the visible energy and the average (n, γ) multiplicity as expected.

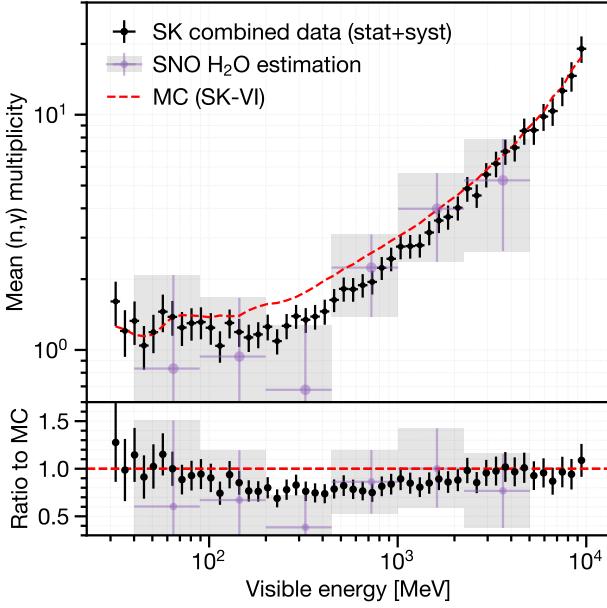


FIG. 19. Average (n, γ) multiplicity in atmospheric neutrino events as a function of their visible energy. Black data points represent combined SK data with statistical and systematic uncertainties. Purple crosses with shades are estimates for pure water based on SNO measurements [13] using a D₂O target. The red dashed line represents the true average (n, γ) multiplicity from the baseline SK-IV MC simulation.

A. Effective metrics for model evaluation

The “slope” of the linear increase in the average (n, γ) multiplicity as a function of visible energy was determined through a linear fit over the energy range [0.3, 10] GeV, as shown in Figure 20. At lower visible energies, sensitivity to low-energy nucleon FSI modeling becomes significant. To quantify neutron production in low energy neutrino events, the “low-energy (low-E) multiplicity” was defined as the average (n, γ) multiplicity observed in [0.1, 0.3] GeV visible energy range. Events below 0.1 GeV were excluded due to large systematic uncertainties.

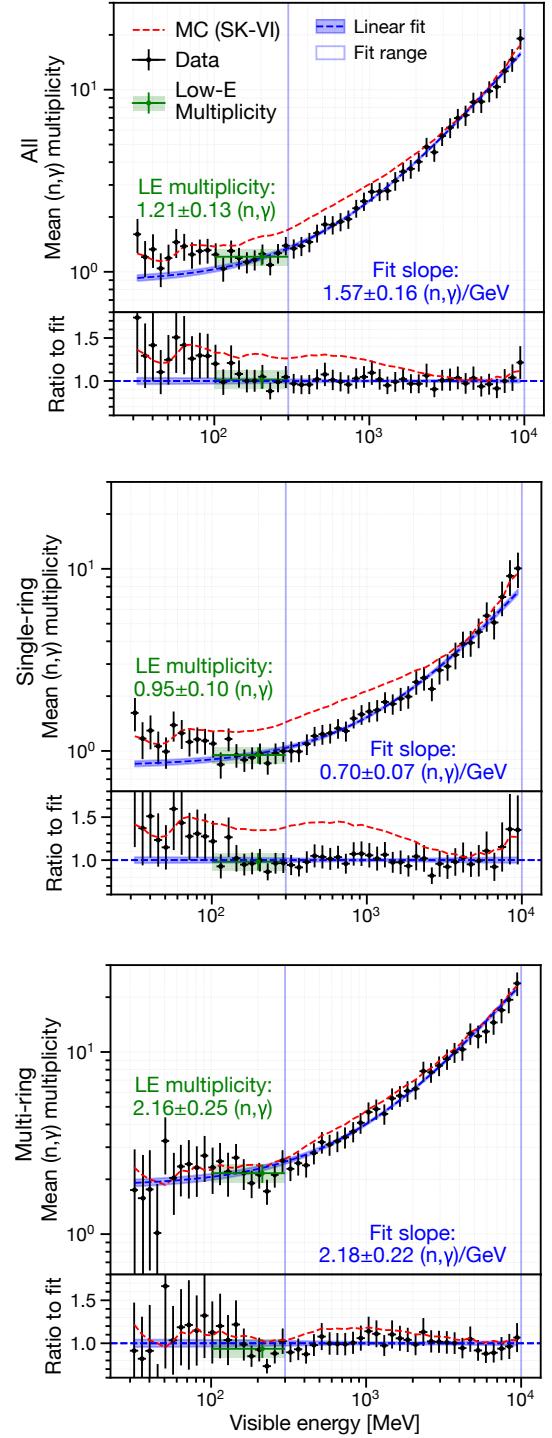


FIG. 20. Average (n, γ) multiplicity as a function of visible energy for neutrino events: all events (top), single-ring events (middle), and multi-ring events (bottom). Red dashed lines show predictions from the full MC simulation in the SK-VI baseline setup. Blue dashed lines represent fitted linear functions with 1 σ prediction intervals shaded in blue, with the fit covering the energy range [0.3, 10] GeV. Green points represent low-energy multiplicity in the [0.1, 0.3] GeV range. Error bars include both statistical and systematic uncertainties.

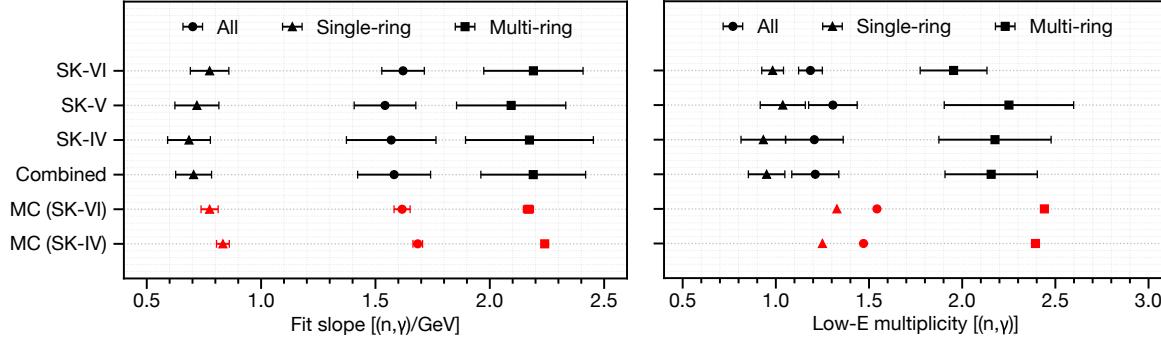


FIG. 21. Measured slopes (left) and low-energy multiplicities (right) across different SK phases and event types. Data error bars (black) represent both statistical and systematic uncertainties, while MC error bars in the left panel indicate statistical uncertainties. In the right panel, MC predictions correspond to the ‘true’ average (n, γ) multiplicity.

Figure 21 compares the measured slopes and low-energy multiplicities across different SK phases and event types. For the slope fit shown in the figure, the binning scheme was adjusted to ensure sufficient number of neutrino events ($\gtrsim 30$) in each bin. The measured values were consistent across SK phases within the assumed uncertainties. Both metrics were higher for multi-ring events than for single-ring events. This difference is expected and due to the higher fraction of DIS events in multi-ring samples, as shown in Figure 5. Additionally, while the observed slopes were consistent with baseline predictions, the low-energy multiplicities were lower than expected.

Figure 21 shows three distinct groups of slopes: the G4 INCL_PC model gives the smallest slopes, G4 BERT_PC predicts intermediate slopes, and other variants of the Bertini cascade model produce the largest slopes. The slope observed in the single-ring data better matches the G4 INCL_PC predictions, while the slope in the multi-ring data is closer to the G4 BERT_PC predictions. This distinction is most evident in the bottom panel of Figure 22, where G4 INCL_PC matches well with sub-GeV single-ring data, and G4 BERT_PC is a better fit for multi-GeV multi-ring data. For low-energy multiplicities, NEUT 5.4.0, NEUT 5.6.3, and GENIE INCL are preferred, as they predict lower neutron production. In contrast, models such as GENIE hA, hN, and BERT overestimate low-energy multiplicities, producing values significantly higher than the estimated 1σ uncertainty.

The goodness of fit between model predictions and data was quantified using the following chi-square (χ^2) definition, considering only the two most dominant sources of uncertainty: statistical uncertainty and signal efficiency scale uncertainty.

$$\chi^2 = \sum_i^{N_{\text{bins}}} \frac{(sO_i - E_i)^2}{\sigma_{\text{stat},i}^2} \quad (6)$$

Here, i is the bin number, O_i and E_i represent the observed and expected average (n, γ) multiplicity, respectively, and $\sigma_{\text{stat},i}$ is the statistical uncertainty in the i -th bin. The total number of bins, N_{bins} , is 20 per subsample for both single-ring and multi-ring events, resulting in 40 bins in total. The binning was optimized to ensure a sufficient number of events per bin so that the average (n, γ) multiplicity follows a normal distribution. The normalization scale s was allowed to float within $[0, 2]$. Two χ^2 calculations were performed: (1) a normalization-free χ^2 , where s was minimized without constraint, and (2) a constrained χ^2 , where s was minimized with a penalty term $(s-1)^2/\sigma_{\text{scale}}^2$ and an additional constraint $|s-1| < \sigma_{\text{scale}}$, with σ_{scale} set to 10%.

B. Comparison with model predictions

Predictions from various model combinations (explained in Section VI) were compared with the combined data, as shown in Figure 22. A notable feature in the data is a “valley” in the $[0.1, 0.3]$ GeV region, which is less pronounced in the baseline simulation. This valley is only reproduced when using models that predict fewer neutrons than the baseline for both FSI and secondary interaction models, such as GENIE INCL and G4 BERT_PC. In multi-GeV events, especially in the multi-ring sample, there was a large variation in the predicted slope.

The bottom panel of Figure 22 highlights these variations, by comparing the data with predictions from specific FSI models (GENIE hN, which predicts more neutrons, and GENIE INCL, which predicts fewer neutrons) and secondary interaction models (G4 BERT, which predicts more neutrons, and G4 BERT_PC and G4 INCL_PC, which predict fewer neutrons). The valley in the $[0.1, 0.3]$ GeV region was only reproduced for configurations of GENIE INCL with either G4 BERT_PC or G4 INCL_PC. G4 INCL_PC predicted smaller slopes in multi-ring multi-GeV events compared to G4 BERT or G4 BERT_PC.

Figure 23 compares the measured slopes and low-energy multiplicities with model predictions. The pre-

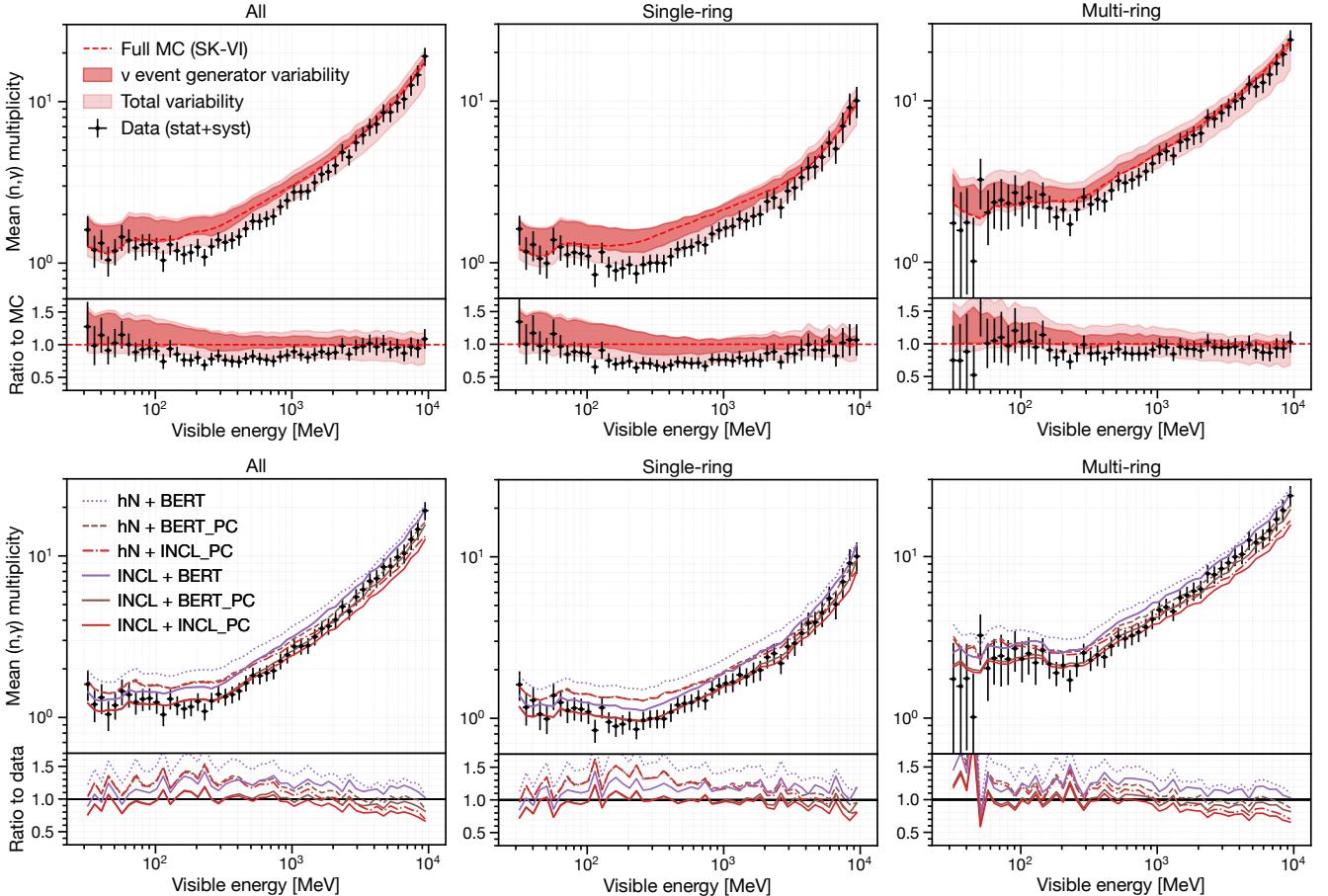


FIG. 22. Comparison of various model predictions with the combined data, showing the average (n, γ) multiplicity as a function of atmospheric neutrino event visible energy, for all events (left), single-ring events (middle), and multi-ring events (right). The black data points represent the combined data, including both statistical and systematic uncertainties. In the top panel, the combined data are compared to the prediction intervals of the models. Thicker shaded regions represent the range of predictions obtained by varying only the neutrino event generator options while keeping the secondary hadron-nucleus interaction model fixed to the baseline (SK-VI default). Lighter shaded regions indicate the broader range of predictions from all model combinations. The bottom panel shows corresponding plots for specific model choices that are based on GENIE and Geant4, including the one (GENIE INCL + G4 BERT_PC) that shows the overall best agreement with the data (See Figure 24).

Figure 24 compares the combined reduced χ^2 values⁹¹⁵ for the single-ring and multi-ring data across different model predictions. Among the neutrino event generators, GENIE INCL showed the lowest χ^2 for both free and constrained s . For secondary hadron-nucleus interaction models, G4 BERT_PC provided the best agreement with data under the 10% normalization scale constraint. The⁹¹⁶ combination of these two models yielded the best over-⁹¹⁷ all agreement under the same constraint. Other Bertini-⁹¹⁸ based secondary hadron interaction models also showed⁹¹⁹ reasonable agreement when coupled with GENIE INCL,⁹²⁰ particularly when normalization was allowed to float. In⁹²¹ contrast, G4 INCL_PC showed higher unconstrained χ^2 ⁹²² values, especially in the multi-ring sample. This trend⁹²³ is consistent with the shallower slopes observed in the⁹²⁴ multi-GeV region in Figures 22 and 23.⁹²⁵

VIII. DISCUSSION

Overall, our data favor models that predict relatively lower neutron production for single-ring sub-GeV events and higher neutron production for multi-ring multi-GeV events. In this section, we examine key features of these models that influence neutron production predictions, focusing on Figure 14, which illustrates differences in nucleon momentum distributions across various neutrino event generators. By qualitatively comparing the variations among different INC models and nuclear de-excitation models used in these generators, we discuss potential sources of the observed differences in their predictions.

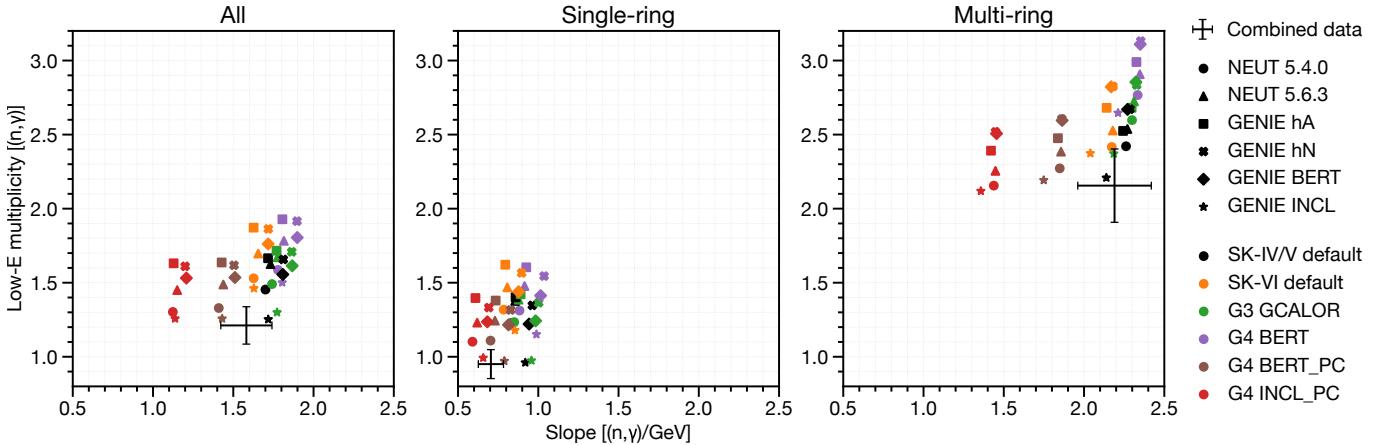


FIG. 23. Scatter plots of fitted slopes and low-energy multiplicities in model predictions and data (black crosses, error bars include both statistical and systematic uncertainties), shown for all events (left), single-ring events (middle), and multi-ring events (right). Shapes represent neutrino event generator options used, while colors indicate the secondary hadron-nucleus interaction model options used.

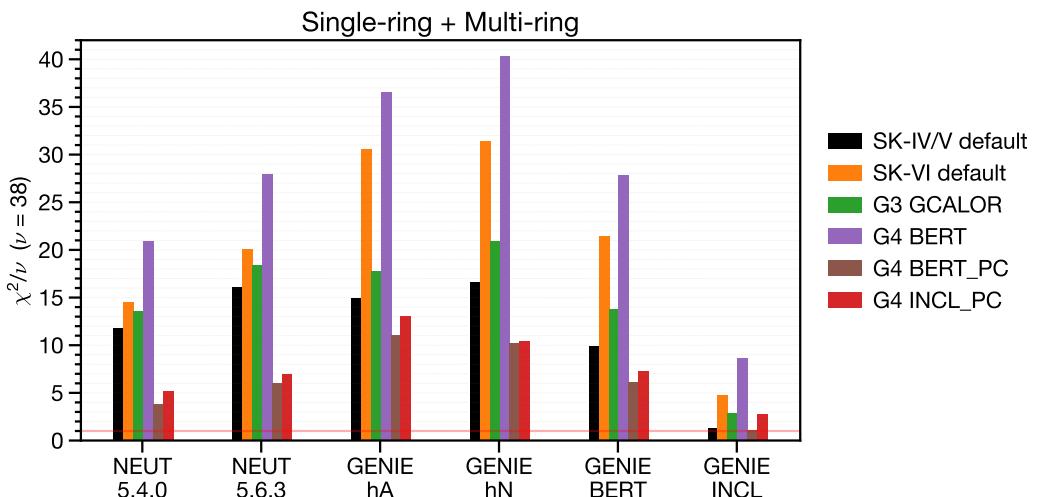


FIG. 24. Comparison of combined reduced χ^2 values (Equation 6) for single-ring and multi-ring data across various model predictions. Colors indicate different secondary hadron-nucleus interaction models. The unhatched area represents the normalization-free χ^2 (s minimized within $[0, 2]$ without penalty), while the total height, including the hatched area, represents χ^2 with a 10% normalization scale constraint. The hatched area shows the contribution from this constraint.

928

A. Nuclear de-excitation

Nuclear de-excitation models are essential for predicting total (n, γ) multiplicity, as they determine the production of very-low-energy neutrons per nuclear interaction. This is evident when comparing G4 BERT and G4 BERT_{PC} in Figures 15, 16, 22, and 23. In the models used in this study, neutron emission is primarily governed by the Weisskopf statistical evaporation model [61], applied after intranuclear cascade and nuclear fragmentation. Here, nucleon emission probability is proportional to the inverse reaction (nucleon absorption) cross section. Compared to the Dostrovsky parameterization [62]

used in the Geant4 Bertini cascade model (see the GENIE BERT neutron peak below 250 MeV/c in Figure 14), the Geant4 Precompound model features tuned level density parameters [58] and improved cross section parameterization [63] fitted to a broader dataset [64, 65], reducing neutron emission probabilities. In addition, unlike the Geant4 Bertini cascade model, the Geant4 Precompound model allows Fermi breakup in nuclei up to mass number 16, including oxygen, which further reduces isolated neutrons contributing to (n, γ) signals. The ABLA07 model used in GENIE INCL predicts neutron rates similar to NEUT 5.4.0 and NEUT 5.6.3, while the GENIE hA and hN options omit nuclear de-excitation entirely.

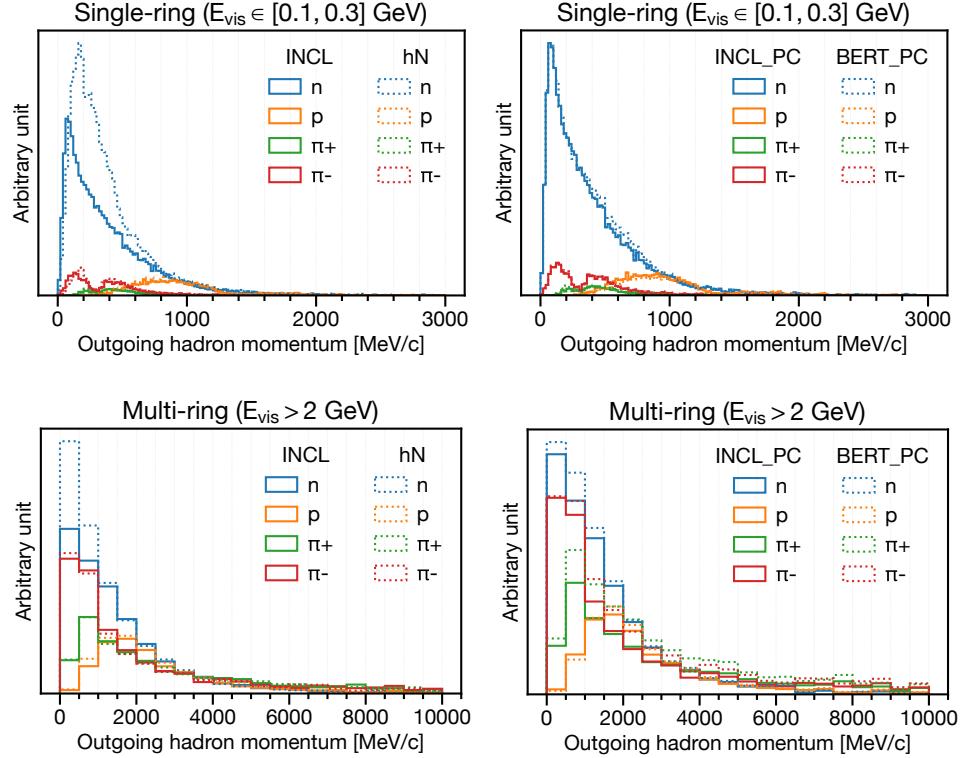


FIG. 25. Momentum distributions of outgoing hadrons predicted by neutrino event generators, weighted by the average (n, γ) multiplicity per projectile hadron momentum bin as predicted by secondary hadron-nucleus interaction models. The area under each histogram bin represents the average contribution of a specific hadron at a given momentum to the observable (n, γ) multiplicity. The top panels show predictions for single-ring events with visible energy (E_{vis}) in the range [0.1, 0.3] GeV, while the bottom panels correspond to multi-ring events with visible energy above 2 GeV. The left plots compare two different cascade models for final-state interactions (GENIE INCL and GENIE hN) coupled with the secondary hadron-nucleus interaction model G4 INCL_{PC}, whereas the right plots compare two secondary interaction models (G4 INCL_{PC} and G4 BERT_{PC}) coupled with the neutrino event generator option GENIE hN.

953

B. Pauli blocking

974

C. Other nuclear in-medium effects

954 A key distinction among the FSI cascade models is
 955 their treatment of Pauli blocking, which prevents nu-
 956 cleons from occupying the same quantum state, as dic-
 957 cated by the Pauli exclusion principle. For instance, NEUT
 958 5.4.0, NEUT 5.6.3 and GENIE BERT implement strict
 959 Pauli blocking, which forbids all nucleon scattering below
 960 the oxygen Fermi momentum of 225 MeV/c, modeling
 961 the nucleus as a degenerate Fermi gas. This strict ap-
 962 proach results in a characteristic dip in the outgoing neu-
 963 tron momentum distributions for NEUT 5.4.0 and NEUT
 964 5.6.3 within the corresponding energy range, as shown
 965 in Figure 14. In contrast, the Liège INC model applies a
 966 probabilistic approach to Pauli blocking during the cas-
 967 cade, accounting for nucleon holes created by prior cas-
 968 cade steps. This results in smoother neutron distribu-
 969 tions near the Fermi momentum, as seen in Figure 14.
 970 In comparison, the GENIE hA and hN models omit Pauli
 971 blocking entirely, leading to significantly higher predicted
 972 neutron production than NEUT 5.4.0, NEUT 5.6.3, and
 973 GENIE INCL.

974 Compared to other secondary hadron-nucleus interaction
 975 models in this study, the Liège INC and Geant4
 976 Bertini cascade models account for hard-core nuclear re-
 977 pulsion (preventing closely overlapping nucleons from in-
 978 teracting) and cascade step correlations (accounting for
 979 nucleon removal from preceding steps), and thus pre-
 980 dict lower neutron production at intermediate energies
 981 (250–1000 MeV/c), as shown in Figure 14. The dif-
 982 ferences in the outgoing hadron momentum distribution
 983 translate into variations in neutron production. The im-
 984 pact on the observed neutron multiplicity is particularly
 985 evident in low-energy single-ring events. For example,
 986 Figure 25 shows the predicted average contribution of
 987 outgoing nucleons and pions with varying momenta from
 988 neutrino interactions to the observed (n, γ) multiplicities.
 989 We observe that GENIE hN, which neglects in-medium ef-
 990 fects, and GENIE INCL, which includes them, predict sig-
 991 nificantly different contributions to the signal multipli-
 992 city due to varying predictions of the outgoing neutron
 993 momentum distribution at intermediate energies.

995

D. Cross sections

1014

IX. SUMMARY

Modeling pion-nucleus interactions is important for neutron production in multi-ring events, which are likely to involve multiple energetic pions, as suggested by Figures 25 and 26. The Liège INC model (G4 INCL_PC) underprediction of the multi-ring slope observed in Figures 22 and 23 can be attributed to the internal pion production cross sections used for pion projectiles. This is supported by Ref. [66], which shows that the Liège INC model underpredicts the inclusive pion production cross sections for pion projectiles compared to the Geant4 Bertini cascade model, especially for lighter nuclei with mass numbers below 20.

In addition, variations in low-energy neutron reaction cross section datasets (e.g., ENDF/B-VI used in SK-IV/V default vs. ENDF/B-VII.1 used in SK-VI default) make roughly 5–10% difference in the average (n, γ) multiplicities for visible energy below 300 MeV, whose effect is visible in Figures 16, 23, and 24.

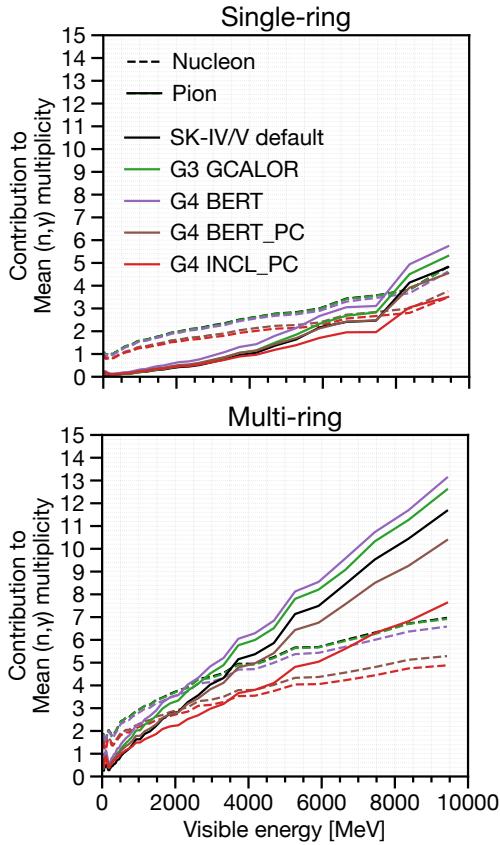


FIG. 26. Average contribution of outgoing nucleons (dashed lines) and pions (solid lines) to the observed (n, γ) multiplicities as a function of visible energy, shown separately for single-ring events (top) and multi-ring events (bottom). The line colors represent different secondary hadron-nucleus interaction models used in the predictions.

Accurately modeling neutron production in neutrino interactions is essential for characterizing incoming neutrinos, which is crucial for advancing precision measurements of neutrino oscillation parameters and rare event searches involving neutron tagging. Recent studies have suggested potential inaccuracies in the modeling of secondary hadron-nucleus interactions.

This paper reports a measurement of total neutron production following atmospheric neutrino interactions within the water volume of the Super-Kamiokande (SK) detector. In SK, neutrons with kinetic energies below a few MeV are typically captured by surrounding nuclei, such as ^1H or $^{155}/^{157}\text{Gd}$, releasing 2–8 MeV of energy through gamma-ray emissions. These radiative neutron captures ((n, γ) reactions) were identified using a detection algorithm that combines a simple low-energy trigger with a neural network binary classifier, calibrated with an Am/Be neutron source. Event-by-event neutron capture multiplicity was estimated using a multivariate non-linear regression technique. Atmospheric neutrino events were binned by their electron-equivalent “visible energy,” a semi-calorimetric proxy for neutrino momentum transfer, in the range of [0.03, 10] GeV. The dominant systematic uncertainty was the roughly 10% uncertainty in the overall signal efficiency scale. The average neutron capture multiplicity in each visible energy bin was compared against predictions from various combinations of neutrino event generators and secondary hadron-nucleus interaction models.

Our data offers strong discriminative power for evaluating hadron-nucleus interaction models, with model predictions varying by up to 50% depending on the energy range. From comparing the data with various model predictions, two key observations were made: first, a reduction in average neutron production in sub-GeV single-ring events, where the impact of secondary hadron interactions is minimal; and second, a nearly linear increase in neutron production as a function of visible energy, especially in the multi-GeV region. The first observation is sensitive to the FSI models used in neutrino event generators, suggesting the need for moderate neutron evaporation through nuclear de-excitation and proper consideration of nuclear in-medium effects, which reduce outgoing nucleons at intermediate momenta below 1 GeV/c. The second observation is more sensitive to secondary hadron-nucleus interaction models, with the impact of internal pion production cross sections for pion projectiles becoming more significant. This study highlights the crucial role of hadron-nucleus interaction models in accurately predicting total neutron production from neutrino interactions.

1066

ACKNOWLEDGMENTS

1067 We gratefully acknowledge the cooperation of the
1068 Kamioka Mining and Smelting Company. The Super-
1069 Kamiokande experiment has been built and operated
1070 from funding by the Japanese Ministry of Education,
1071 Culture, Sports, Science and Technology; the U.S.
1072 Department of Energy; and the U.S. National Sci-
1073 ence Foundation. Some of us have been supported
1074 by funds from the National Research Foundation of
1075 Korea (NRF-2009-0083526, NRF-2022R1A5A1030700,
1076 NRF-2202R1A3B1078756) funded by the Ministry of
1077 Science, Information and Communication Technology
1078 (ICT); the Institute for Basic Science (IBS-R016-Y2);
1079 and the Ministry of Education (2018R1D1A1B07049158,
1080 2021R1I1A1A01042256, 2021R1I1A1A01059559, RS-
1081 2024-00442775); the Japan Society for the Promotion
1082 of Science; the National Natural Science Foundation of
1083 China under Grants No.12375100; the Spanish Ministry
1084 of Science, Universities and Innovation (grant PID2021-
1085 124050NB-C31); the Natural Sciences and Engineering
1086 Research Council (NSERC) of Canada; the Scinet
1087 and Westgrid consortia of Compute Canada; the Na-
1088 tional Science Centre (UMO-2018/30/E/ST2/00441 and
1089 UMO-2022/46/E/ST2/00336) and the Ministry of Sci-
1090 ence and Higher Education (2023/WK/04), Poland;
1091 the Science and Technology Facilities Council (STFC)
1092 and Grid for Particle Physics (GridPP), UK; the Eu-
1093 ropean Union's Horizon 2020 Research and Innova-
1094 tion Programme under the Marie Skłodowska-Curie
1095 grant agreement no.754496; H2020-MSCA-RISE-2018
1096 JENNIFER2 grant agreement no.822070, H2020-MSCA-
1097 RISE-2019 SK2HK grant agreement no. 872549; and
1098 European Union's Next Generation EU/PRTR grant
1099 CA3/RSUE2021-00559; the National Institute for Nu-
1100 clear Physics (INFN), Italy.

- [1] P. Abratenko *et al.* (MicroBooNE collaboration), Phys. Rev. D **110**, 013006 (2024).
[2] P. Abratenko *et al.* (MicroBooNE collaboration), The European Physical Journal C **84**, 1052 (2024).
[3] A. M. Ankowski, P. Coloma, P. Huber, C. Mariani, and E. Vagnoni, Phys. Rev. D **92**, 091301 (2015).
[4] C. L. Cowan, F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire, Science **124**, 103 (1956), <https://www.science.org/doi/pdf/10.1126/science.124.3212.103>.
[5] T. Wester *et al.* (Super-Kamiokande collaboration), Phys. Rev. D **109**, 072014 (2024).
[6] A. Takenaka *et al.* (Super-Kamiokande collaboration), Phys. Rev. D **102**, 112011 (2020).
[7] S. Agostinelli *et al.* (Geant4 collaboration), Nucl. Instrum. Methods A **506**, 250 (2003).
[8] J. Allison *et al.*, IEEE Transactions on Nuclear Science **53**, 270 (2006).
[9] J. Allison *et al.*, Nucl. Instrum. Methods A **835**, 186 (2016).
[10] J. Cugnon, Nucl. Phys. A **387**, 191 (1982).
[11] J.-C. David *et al.*, Progress In Nuclear Science and Technology **2**, 942 (2011).
[12] R. Akutsu, A study of neutrons associated with neutrino and antineutrino interactions on the water target at the T2K far detector, Ph.D. thesis, The Univ. of Tokyo (2019).
[13] B. Aharmim *et al.* (SNO collaboration), Phys. Rev. D **99**, 112007 (2019).
[14] M. Elkins *et al.* (MINERvA collaboration), Phys. Rev. D **100**, 052002 (2019).
[15] A. Olivier *et al.* (MINERvA collaboration), Phys. Rev. D **108**, 112010 (2023).
[16] S. Abe *et al.* (KamLAND Collaboration), Phys. Rev. D **107**, 072006 (2023).
[17] M. B. Chadwick *et al.*, Nuclear Data Sheets **112**, 2887 (2011), special Issue on ENDF/B-VII.1 Library.
[18] K. Abe *et al.* (Super-Kamiokande collaboration), Nucl. Instrum. Methods A **1027**, 166248 (2022).
[19] K. Abe *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **1065**, 169480 (2024).
[20] S. Fukuda *et al.* (Super-Kamiokande collaboration), Nucl. Instrum. Methods A **501**, 418 (2003).
[21] K. Abe *et al.*, Nucl. Instrum. Methods A **737**, 253–272 (2014).
[22] S. Yamada *et al.*, IEEE Trans. Nucl. Sci. **57**, 428 (2010).
[23] C. Andreopoulos *et al.*, Nucl. Instrum. Methods A **614**, 87 (2010), arXiv:0905.2517 [hep-ph].
[24] C. Andreopoulos *et al.*, The GENIE Neutrino Monte Carlo Generator: Physics and User Manual (2015), arXiv:1510.05494 [hep-ph].
[25] J. Tena-Vidal *et al.* (GENIE), Phys. Rev. D **104**, 072009 (2021), arXiv:2104.09179 [hep-ph].
[26] M. Shiozawa, Nucl. Instrum. Methods A **433**, 240 (1999).
[27] M. Honda, T. Kajita, K. Kasahara, and S. Midorikawa, Phys. Rev. D **83**, 123001 (2011).
[28] R. L. Workman *et al.* (Particle Data Group), PTEP **083C01**, 729 (2022).
[29] Y. Hayato and L. Pickering, Eur. Phys. J Spec. Top. **230**, 4469–4481 (2021).
[30] L. L. Salcedo, E. Oset, M. J. Vicente-Vacas, and C. Garcia-Recio, Nucl. Phys. A **484**, 557 (1988).
[31] E. S. P. Guerra *et al.*, Phys. Rev. D **99**, 052007 (2019).
[32] H. W. Bertini, Phys. Rev. **131**, 1801 (1963).
[33] P. de Perio, AIP Conf. Proc. **1405**, 223 (2011), arXiv:1405.3973 [nucl-ex].
[34] A. M. Ankowski, O. Benhar, T. Mori, R. Yamaguchi, and M. Sakuda, Phys. Rev. Lett. **108**, 052505 (2012).
[35] H. Ejiri, Phys. Rev. C **48**, 1442 (1993).
[36] C. Zeitnitz and T. A. Gabriel, Nucl. Instrum. Meth. A **349**, 106 (1994).
[37] R. Brun, F. Bruyant, M. Maire, A. C. McPherson, and P. Zanarini, *GEANT3*, Tech. Rep. (1987).
[38] H. W. Bertini, Phys. Rev. **188**, 1711 (1969).
[39] P. F. Rose, *ENDF-201: ENDF/B-VI summary documentation*, Tech. Rep. (Brookhaven National Lab. (BNL), Upton, NY (United States), 1991).
[40] 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. (International Atomic Energy Agency (IAEA), 2018) iNDc(NDS)-0758.
[41] K. Hagiwara *et al.*, Prog. Theor. Exp. Phys. **2019**, 023D01 (2019), <https://academic.oup.com/ptep/article-pdf/2019/2/023D01/27970473/ptz002.pdf>.
[42] T. Tanaka *et al.*, Prog. Theor. Exp. Phys. **2020**, 043D02 (2020), <https://academic.oup.com/ptep/article-pdf/2020/4/043D02/33040537/ptaa015.pdf>.
[43] V. Barger, K. Whisnant, S. Pakvasa, and R. J. N. Phillips, Phys. Rev. D **22**, 2718 (1980).
[44] M. Jiang *et al.* (Super-Kamiokande collaboration), Progress of Theoretical and Experimental Physics **2019**, 053F01 (2019), <https://academic.oup.com/ptep/article-pdf/2019/5/053F01/28638877/ptz015.pdf>.
[45] A. M. Dziewonski and D. L. Anderson, Physics of the Earth and Planetary Interiors **25**, 297 (1981).
[46] K. Abe *et al.* (Super-Kamiokande collaboration), J. Instrum. **17** (10), P10029.
[47] F. Chollet *et al.*, Keras, <https://keras.io> (2015).
[48] K. He, X. Zhang, S. Ren, and J. Sun, Delving deep into rectifiers: Surpassing human-level performance on imagenet classification (2015), arXiv:1502.01852 [cs.CV].
[49] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization (2017), arXiv:1412.6980 [cs.LG].
[50] H. Ito, K. Wada, T. Yano, Y. Hino, Y. Ommura, M. Harada, A. Minamino, and M. Ishitsuka, Nucl. Instrum. Methods A **1057**, 168701 (2023).
[51] B. Pritychenko and S. F. Mughabghab, Nuclear Data Sheets **113**, 3120 (2012).
[52] Y. Hino *et al.*, Modification on thermal motion in geant4 for neutron capture simulation in gadolinium loaded water (2024), arXiv:2412.04186 [hep-ex].
[53] T. Hastie and R. Tibshirani, *Generalized Additive Models*, Chapman & Hall/CRC Monographs on Statistics & Applied Probability (Taylor & Francis, 1990).
[54] D. Servén, C. Brummitt, and H. Abedi, dswah/pygam: v0.8.0 (2018).
[55] D. Wright and M. Kelsey, Nucl. Instrum. Methods A **804**, 175 (2015).
[56] S. Leray, D. Mancusi, P. Kaitaniemi, J. C. David, A. Boudard, B. Braunn, and J. Cugnon, Journal of

- 1223 Physics: Conference Series **420**, 012065 (2013). 1235
- 1224 [57] A. Kelic, M. V. Ricciardi, and K.-H. Schmidt, Abla07 1236
towards a complete description of the decay channels of¹²³⁷
a nuclear system from spontaneous fission to multifrag¹²³⁸
mentation (2009), arXiv:0906.4193 [nucl-th]. 1239
- 1225 [58] J. Quesada Molina *et al.*, Progress in Nuclear Science and¹²⁴⁰
Technology **2**, 936 (2011). 1241
- 1226 [59] V. M. Grichine, The European Physical Journal C **62**¹²⁴²
399 (2009). 1243
- 1227 [60] M. Smy, in *International Cosmic Ray Conference*, In¹²⁴⁴
ternational Cosmic Ray Conference, Vol. 5 (2008) pp¹²⁴⁵
1279–1282. 1246
- 1228 [61] V. Weisskopf, Phys. Rev. **52**, 295 (1937).
- 1229 [62] I. Dostrovsky, Z. Fraenkel, and G. Friedlander, Phys.
Rev. **116**, 683 (1959).
- 1230 [63] A. Chatterjee, K. H. N. Murthy, and S. K. Gupta, Pra-
mana **16**, 391 (1981).
- 1231 [64] C. Kalbach, Preco-2000, exciton model preequilibrium
code system with direct reactions. preco-d2, pre-
equilibrium and direct reaction double differential cross-
sections (2002).
- 1232 [65] H. P. Wellisch and D. Axen, Phys. Rev. C **54**, 1329
(1996).
- 1233 [66] D. Mancusi *et al.*, The European Physical Journal A **53**,
80 (2017). 1247