

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

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

(The Super-Kamiokande Collaboration)

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

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

<sup>3</sup>Department of Theoretical Physics, University Autónoma Madrid, 28049 Madrid, Spain

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

<sup>5</sup>Department of Physics, Boston University, Boston, MA 02215, USA

<sup>6</sup>Department of Physics, British Columbia Institute of Technology, Burnaby, BC, V5G 3H2, Canada

<sup>7</sup>Department of Physics and Astronomy, University of California, Irvine, Irvine, CA 92697-4575, USA

<sup>8</sup>Department of Physics, California State University, Dominguez Hills, Carson, CA 90747, USA

<sup>9</sup>Institute for Universe and Elementary Particles, Chonnam National University, Gwangju 61186, Korea

<sup>10</sup>Department of Physics, Duke University, Durham NC 27708, USA

<sup>11</sup>Ecole Polytechnique, IN2P3-CNRS, Laboratoire Leprince-Ringuet, F-91120 Palaiseau, France

<sup>12</sup>Department of Physics, Gifu University, Gifu, Gifu 501-1193, Japan

<sup>13</sup>GIST College, Gwangju Institute of Science and Technology, Gwangju 500-712, Korea

<sup>14</sup>School of Physics and Astronomy, University of Glasgow, Glasgow, Scotland, G12 8QQ, United Kingdom

<sup>15</sup>Department of Physics and Astronomy, University of Hawaii, Honolulu, HI 96822, USA

<sup>16</sup>Center for Underground Physics, Institute for Basic Science (IBS), Daejeon, 34126, Korea

<sup>17</sup>Institute For Interdisciplinary Research in Science and Education, ICISE, Quy Nhon, 55121, Vietnam

<sup>18</sup>Department of Physics, Imperial College London , London, SW7 2AZ, United Kingdom

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

<sup>20</sup>Dipartimento di Fisica, INFN Sezione di Napoli and Università di Napoli, I-80126, Napoli, Italy

<sup>21</sup>Dipartimento di Fisica, INFN Sezione di Padova and Università di Padova, I-35131, Padova, Italy

<sup>22</sup>INFN Sezione di Roma and Università di Roma “La Sapienza”, I-00185, Roma, Italy

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

<sup>24</sup>Department of Physics, Keio University, Yokohama, Kanagawa, 223-8522, Japan

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

<sup>26</sup>Department of Physics, King’s College London, London, WC2R 2LS, UK

<sup>27</sup>Department of Physics, Kobe University, Kobe, Hyogo 657-8501, Japan

<sup>28</sup>Department of Physics, Kyoto University, Kyoto, Kyoto 606-8502, Japan

<sup>29</sup>Department of Physics, University of Liverpool, Liverpool, L69 7ZE, United Kingdom

<sup>30</sup>School of Physics and Astronomy, University of Minnesota, Minneapolis, MN 55455, USA

<sup>31</sup>Department of Physics, Miyagi University of Education, Sendai, Miyagi 980-0845, Japan

<sup>32</sup>Institute for Space-Earth Environmental Research, Nagoya University, Nagoya, Aichi 464-8602, Japan

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

<sup>34</sup>National Centre For Nuclear Research, 02-093 Warsaw, Poland

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

<sup>36</sup>Department of Physics, Okayama University, Okayama, Okayama 700-8530, Japan

<sup>37</sup>Media Communication Center, Osaka Electro-Communication University, Neyagawa, Osaka, 572-8530, Japan

<sup>38</sup>Department of Physics, Oxford University, Oxford, OX1 3PU, United Kingdom

<sup>39</sup>Rutherford Appleton Laboratory, Harwell, Oxford, OX11 0QX, UK

<sup>40</sup>Department of Physics and Astronomy, Seoul National University, Seoul 151-742, Korea

<sup>41</sup>Department of Physics and Astronomy, University of Sheffield, S3 7RH, Sheffield, United Kingdom

<sup>42</sup>Department of Informatics in Social Welfare, Shizuoka University of Welfare, Yaizu, Shizuoka, 425-8611, Japan

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

<sup>44</sup>STFC, Rutherford Appleton Laboratory, Harwell Oxford, and

Daresbury Laboratory, Warrington, OX11 0QX, United Kingdom

<sup>45</sup>Department of Physics, Sungkyunkwan University, Suwon 440-746, Korea

<sup>46</sup>Department of Physics, Faculty of Science, Tohoku University, Sendai, Miyagi, 980-8578, Japan

<sup>47</sup>Department of Physics, Tokai University, Hiratsuka, Kanagawa 259-1292, Japan

<sup>48</sup>Department of Physics, University of Tokyo, Bunkyo, Tokyo 113-0033, Japan

<sup>49</sup>Kavli 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

<sup>50</sup>Department of Physics, Institute of Science Tokyo, Meguro, Tokyo 152-8551, Japan

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

<sup>52</sup>Faculty of Science, University of Toyama, Toyama City, Toyama 930-8555, Japan

<sup>53</sup>TRIUMF, 4004 Wesbrook Mall, Vancouver, BC, V6T2A3, Canada

<sup>54</sup>Department of Engineering Physics, Tsinghua University, Beijing, 100084, China

<sup>55</sup>Faculty of Physics, University of Warsaw, Warsaw, 02-093, Poland

<sup>56</sup>Department of Physics, University of Warwick, Coventry, CV4 7AL, UK

<sup>57</sup>Department of Physics, University of Winnipeg, MB R3J 3L8, Canada

<sup>58</sup>Department of Physics, Yokohama National University, Yokohama, Kanagawa, 240-8501, Japan

(Dated: April 5, 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, 3]. 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 [4].

Neutrons can also be detected via the radiative neutron capture  $(n, \gamma)$  reaction, which has no practical energy threshold. Neutrons with kinetic energies of a few MeV<sup>173</sup> or lower are likely to thermalize, undergoing  $(n, \gamma)$  reactions with well-defined timescales and energy signatures<sup>175</sup> that enable clean signal selection. This characteristic historically made neutrons effective tags for antineutrino<sup>177</sup> charged-current (CC) interactions (e.g.,  $\bar{\nu}_e p \rightarrow e^+ n$ )<sup>178</sup> compared to neutrino interactions (e.g.,  $\nu_e n \rightarrow e^- p$ )<sup>179</sup>. Neutron tagging remains relevant today, for instance, in atmospheric neutrino oscillation analyses, where it enhances sensitivity to both neutrino mass ordering and CP<sup>182</sup> violation, by preventing the cancellation of opposite-sign<sup>183</sup> effects in neutrino and antineutrino oscillation probabilities<sup>184</sup> [6]. Additionally, neutron tagging helps suppress atmospheric neutrino backgrounds for rare events, such as<sup>186</sup> proton decay (e.g.,  $p \rightarrow e^+ \pi^0$ )<sup>187</sup>, which in many cases is<sup>188</sup> not expected to produce neutrons, or inverse beta decay<sup>189</sup> ( $\bar{\nu}_e p \rightarrow e^+ n$ ) induced by supernova  $\bar{\nu}_e$ , which typically<sup>189</sup> produces only one neutron. Accurately predicting the number of detectable neutrons from neutrino interactions<sup>191</sup> is crucial for these analyses and requires well-constrained<sup>192</sup> uncertainties in interaction models.<sup>193</sup>

The modeling approach commonly adopted by GeV-<sup>194</sup> scale neutrino experiments is illustrated in Figure 1. Neutrino event generators simulate outgoing hadrons from<sup>196</sup> the initial neutrino interaction, either at the nucleon or<sup>197</sup> quark level, and subsequently model intranuclear hadron<sup>198</sup> transport (often referred to as final-state interactions,<sup>199</sup> or FSI), followed by nuclear de-excitation. Similarly,<sup>200</sup> particle transport codes such as Geant4 [8–10] simulate<sup>201</sup> hadron transport and nuclear de-excitation to describe<sup>202</sup> downstream hadron-nucleus interactions within the detector.<sup>203</sup>

Accurate prediction of the observable neutron multi-<sup>205</sup> plicity requires careful consideration of secondary neu-<sup>206</sup> tron production. For hadron transport in the  $O(0.1\text{--}1)$ <sup>207</sup> GeV energy range, semi-classical Intranuclear Cascade<sup>208</sup> (INC) models [11] are commonly used. These models ap-<sup>209</sup>

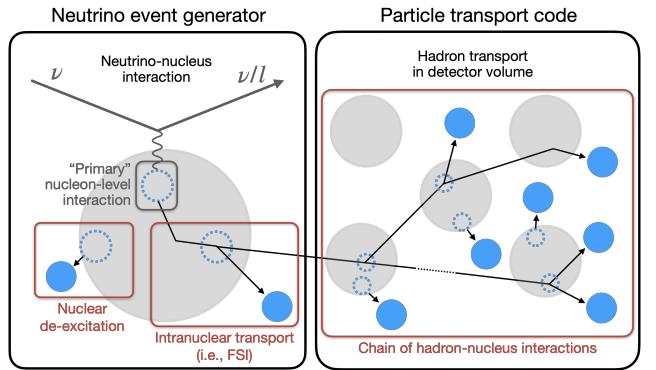


FIG. 1. Schematic of sources of nucleon production in a typical GeV-scale neutrino event simulation setup. Dashed blue circles represent nucleon holes, black arrows represent nucleon trajectories, and solid blue circles represent detectable nucleons ejected from nuclei.

proximate the process as a series of binary collisions between 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 [12]. Several studies have measured neutron production from atmospheric or artificial neutrino interactions using water (T2K [13]), heavy water (SNO [14]), and hydrocarbon (MINERvA [2, 15], KamLAND [16]) as target materials. Several of these studies [2, 13, 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  $(n, \gamma)$  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  $(n, \gamma)$  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 <sup>a</sup> [wt%] | Expected $(n, \gamma)$ fraction <sup>b</sup><br>$H(n, \gamma)$ [%] | $Gd(n, \gamma)$ [%] | $(n, \gamma)$ time constant <sup>c</sup> [ $\mu$ s] |
|-------------|--------------|--------------------|-------------------------------------|--|---------------------|---|
| SK-I-III    | 1996–2008    | 2805.9             | -                                   | >99.9  | -                   | No data   |
| SK-IV       | 2008–2018    | 3244.4             | -                                   | >99.9  | -                   | $204.8 \pm 9.8$                                     |
| SK-V        | 2019–2020    | 461.0              | -                                   | >99.9  | -                   | $199.8 \pm 10.2$                                    |
| SK-VI       | 2020–2022    | 564.4 <sup>d</sup> | $0.0110 \pm 0.0001$ [18]            | $56.1 \pm 1.5$   | $43.9 \pm 1.5$      | $116.2 \pm 2.3$                                     |
| SK-VII-VIII | 2022–present | -                  | $0.0332 \pm 0.0002$ [19]            | $29.7 \pm 0.7$   | $70.3 \pm 0.7$      | $61.8 \pm 0.1$ [19]                                 |

<sup>a</sup> Based on the amount of dissolved Gd.

<sup>b</sup> Based on the evaluated thermal  $(n, \gamma)$  reaction cross sections and uncertainties of ENDF/B-VII.1 [17].

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

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

## II. THE SUPER-KAMIOKANDE DETECTOR

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

Charged particles, namely electrons and muons produced by charged-current neutrino interactions, are identified through Cherenkov radiation. The radiation is projected onto the PMTs as a characteristic ring pattern that depends on the particle type and energy. This ring pattern serves as the basis for particle reconstruction. Neutrons are indirectly identified via Compton-scattered electrons resulting from  $(n, \gamma)$  reactions. In pure water, most occur on  $^1H$ , emitting a single 2.2 MeV gamma-ray. With the recent addition of gadolinium (Gd), a large fraction of neutrons are expected to be captured by Gd isotopes, resulting in a total gamma-radiated energy of around 8 MeV.

The  $O(1)$  MeV signal identification performance is significantly influenced by variations in detector characteristics. Parameters such as individual PMT gain, timing properties, quantum efficiency, and optical absorption and scattering in water are continuously monitored using cosmic-ray muons and light sources [21]. Additionally, the uncertainty in Cherenkov ring energy reconstruction (described in Section III B) is evaluated over a wide energy range using naturally occurring particles, including cosmic-ray muons, Michel electrons, and neutral pions produced in neutral-current (NC) atmospheric neutrino interactions in water. Figure 2 illustrates the agreement between data and simulation in energy reconstruction for the Gd-loaded SK-VI phase, which is mostly within 2% across the  $O(10\text{--}10^4)$  MeV range and consistent with the pure water phase results reported in [6].

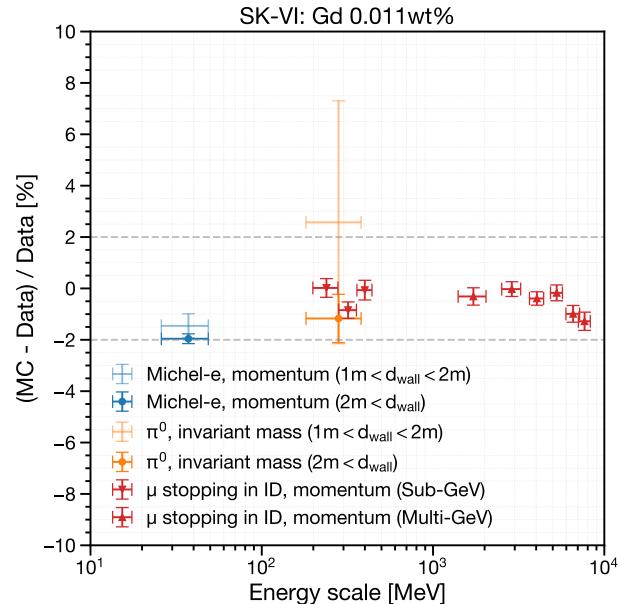


FIG. 2. Comparison of energy reconstruction performance between Monte Carlo (MC) simulation and observed data for naturally occurring particles in the Gd-loaded SK-VI phase. The variable  $d_{\text{wall}}$  represents the distance (in meters) from the reconstructed vertex to the nearest ID photodetector 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  $\mu$ s after certain ID triggers. This has enabled analysis of delayed neutron captures that occur with a time scale of  $O(10\text{--}100)$   $\mu$ s 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  $Gd_2(SO_4)_3$  into the water volume to enhance neutron detection efficiency [18, 19]. Table I summarizes the relevant operational conditions.

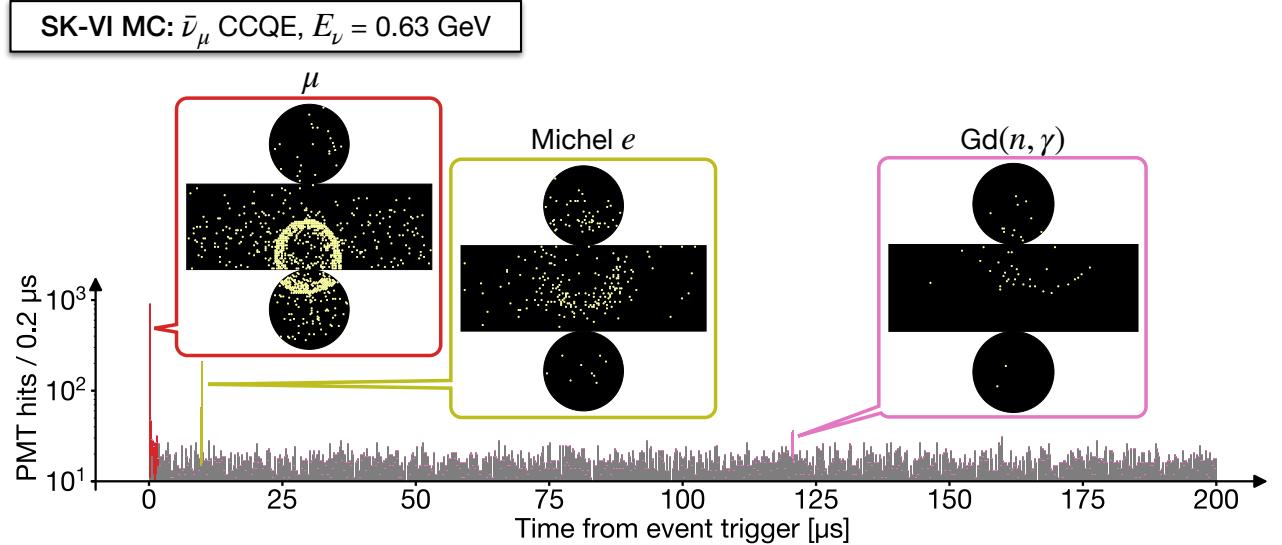


FIG. 3. A PMT hit time distribution for a typical  $\bar{\nu}_\mu$  charged-current (CC) quasi-elastic (QE) interaction with a muon and a neutron as final states. The zero is set at the event trigger. The event displays feature the “prompt” muon signal (red) and the two types of “delayed” coincident 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.

### 266 III. ATMOSPHERIC NEUTRINO EVENTS

#### 267 A. Data reduction

268 We followed a typical selection process for atmospheric  
269 neutrino interactions that are fully contained within the  
270 ID, similar to previous studies conducted at SK [6]. All  
271 events were required to pass the ID trigger with the  
272 threshold  $N_{200\text{-ns}} \geq 58$  PMT hits—roughly correspond-  
273 ing to a 10 MeV electron—followed by the extended event  
274 window of 535  $\mu\text{s}$  for neutron detection. Background  
275 events from cosmic-ray muons, radioactivity, and neu-  
276 trino interactions with exiting particles were reduced us-  
277 ing OD veto and ID charge cuts.

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  
281 m away from the ID tank wall (defining the fiducial vol-  
282 ume with 27.2 kton of water) and that the visible energy  
283 be larger than 30 MeV. The remaining background con-  
284 tamination, mainly due to cosmic-ray muons stopping in  
285 the ID and PMT discharges, was estimated to be below  
286 0.2%, based on visual inspection [7].

#### 287 B. Reconstruction of prompt Cherenkov rings

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

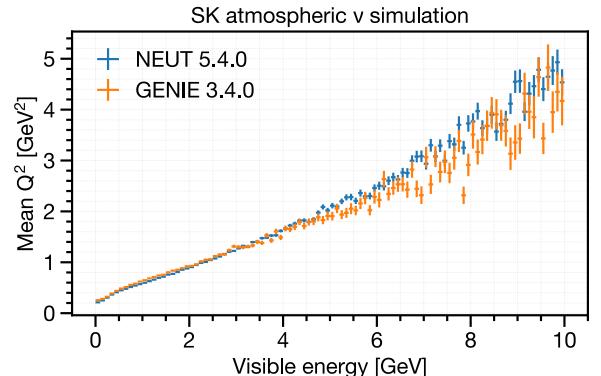


FIG. 4. Average squared momentum transfer ( $Q^2$ ) per visible energy bin for simulated atmospheric neutrino events at SK, compared between two different neutrino event genera-  
tors. NEUT 5.4.0 (blue) uses the baseline setup described in  
Section III C, while GENIE 3.4.0 [23–25] uses the “hN” setup  
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 [6, 7].

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-  
301 itive correlation shown in Figure 4.

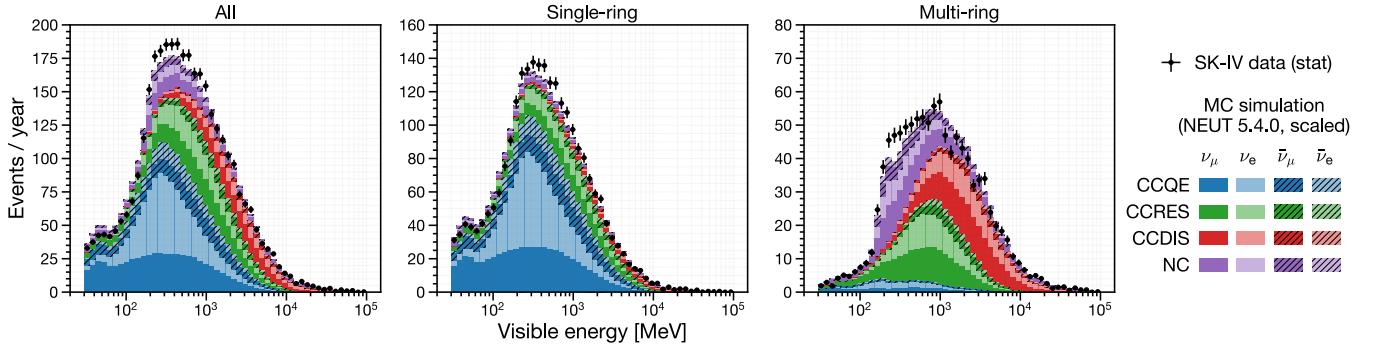


FIG. 5. Visible energy distributions for observed and simulated atmospheric neutrino interactions in water, fully contained within the fiducial volume: all events (left), events with one reconstructed Cherenkov ring (middle), and events with two or more rings (right). CCQE events dominate the single-ring 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 the uncertainty in absolute neutrino flux normalization.

### 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  $\nu_e$ ,  $\nu_\mu$ ,  $\bar{\nu}_e$ , and  $\bar{\nu}_\mu$  at the detector site, assuming no oscillations, as provided by Honda *et al.* [27] for  $[10^2, 10^7]$  MeV range.  $\nu_\tau$  and  $\bar{\nu}_\tau$  interactions were neglected as their overall presence in data is expected to be below 1%.

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  $\Delta$  resonance (RES), and multiple pion production via deep inelastic scattering (DIS). The models used are consistent with the latest SK atmospheric neutrino oscillation analysis [6]. 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] coupled with GEANT 3.21 [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 GEANT.

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 [8–10] 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  
 380 433

381 Neutrons produced by atmospheric neutrino interactions in the SK ID are primarily captured by  $^1\text{H}$  or 434  
 382  $^{155}/^{157}\text{Gd}$  isotopes within approximately  $O(100)$   $\mu\text{s}$ , with 435  
 383 expected capture ratios summarized in Table I. Electrons 436 that are scattered by the  $O(1)$  MeV gamma-rays generate 437  
 384 Cherenkov photons that form faint rings on the ID tank wall 438 (hereafter referred to as “neutron signal”), as 439  
 385 shown in Figure 3. 439

386 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 number of PMT hits, *assuming* it to be near another known vertex, such as the reconstructed neutrino interaction vertex, helps identify signals from random PMT coincidences caused by dark current. Remaining Michel electrons from muon decays can be effectively suppressed through cuts based on signal timing and energy. 448  
 397  
 449

398 **A. Signal selection algorithm** 450  
 399  
 451

400 The signal selection algorithm is based on Ref. [46], 453 and consists of two stages. 454

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

464 In the second candidate classification stage, we extract features of each candidate and use a neural network to 463 classify each candidate into signal and noise based on 464 input features. These features characterize the signal 465 energy, the background hit level, timing spread assuming 466 the vertex, correlation between the input vertex and the 467 hit PMT positions, correlation to the known properties 468 of PMT noise, and angular correlation among hit PMTs 469 relative to the Cherenkov cone opening angle.

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

471 Here, we provide the definition and unit of each feature 472 used for the classification of signal candidates, along with 473 their expected distributions as shown in Figure 6. 474

- **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, minus 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, calculated along the average direction of the unit vectors connecting the vertex to each hit PMT.

- **BurstRatio**

The ratio of the selected PMTs with a preceding hit within 10  $\mu\text{s}$ , 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  $\sigma$  represents the sigmoid function,  $r_i$  is the dark rate of the  $i^{\text{th}}$  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  $k^{\text{th}}$  Legendre polynomial and  $\theta_{ij}$  is the opening angle between the assumed signal vertex and the  $i^{\text{th}}$  and  $j^{\text{th}}$  hit PMTs.

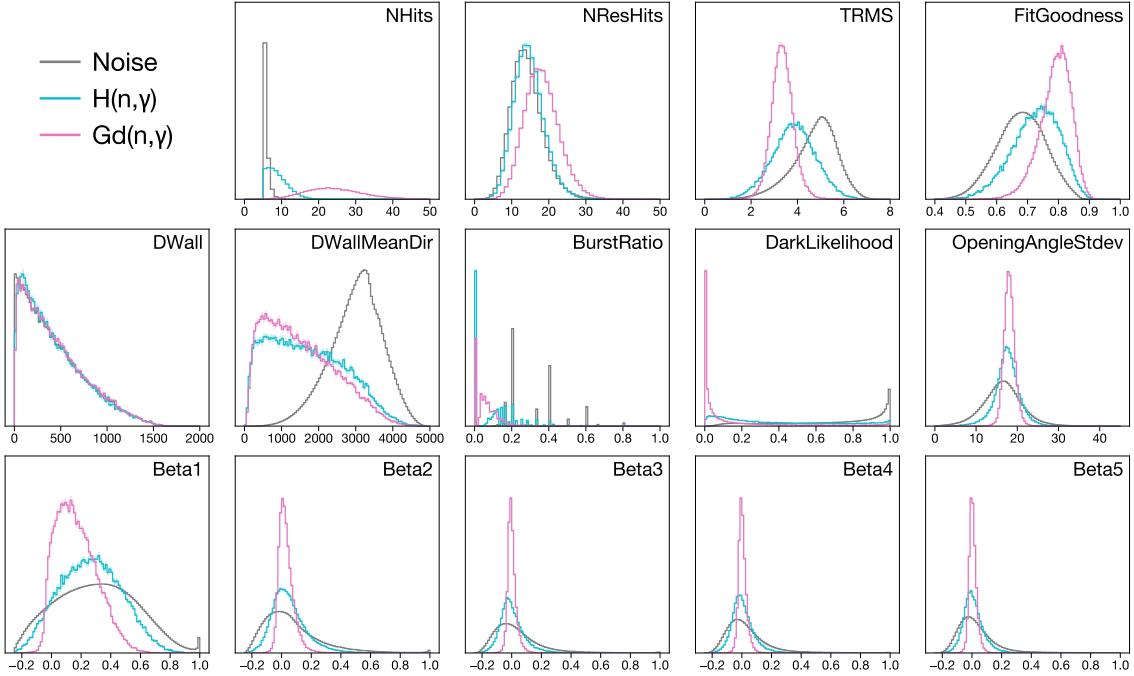


FIG. 6. The features (area-normalized) of neutron signals and noise from the thermal neutron MC simulation used for training the neural network for SK-VI phase.

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  $(n, \gamma)$  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  $98.7 \pm 0.5\%$ .

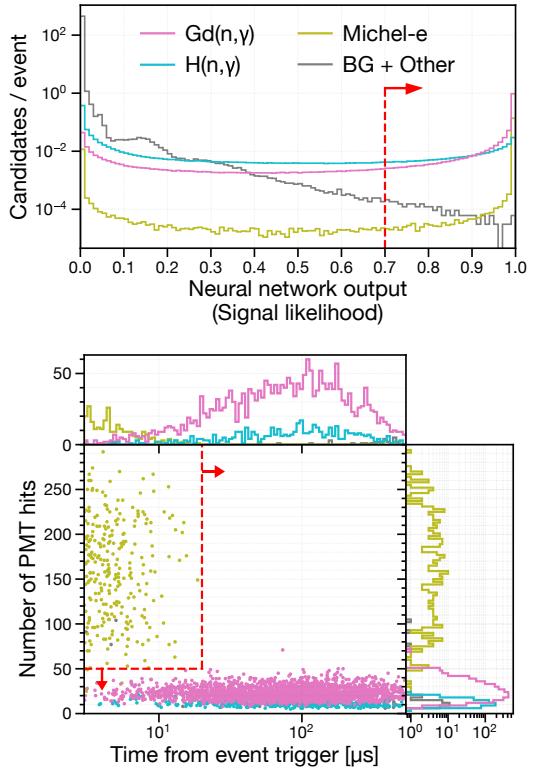


FIG. 7. The neural network response to neutron signals and backgrounds in the SK-VI atmospheric neutrino simulation (top), and the time versus energy distribution of signal candidates passing the neural network selection (bottom). Red arrows indicate the corresponding cut points.

## 506 B. Signal selection performance on calibration data

507 An Am/Be neutron source with a measured total intensity  
 508 of  $236.8 \pm 5.0$  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 scin-  
 530 tillator 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.

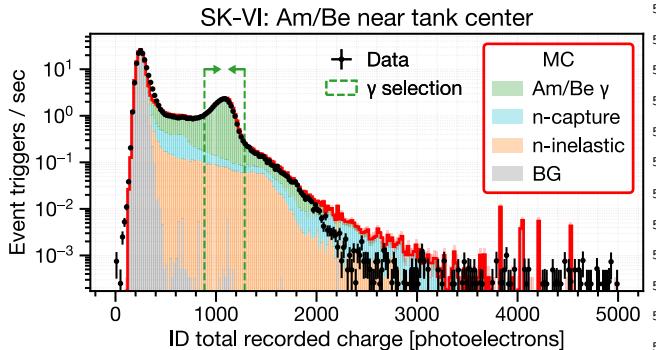


FIG. 8. Distribution of recorded charge within the time win-  
 556 dow  $[-0.5, 1.0]$   $\mu\text{s}$  relative to event triggers, measured with  
 557 an Am/Be neutron source positioned near the ID tank center.  
 558 The black points represent SK-VI data, while the red line shows  
 559 the simulated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

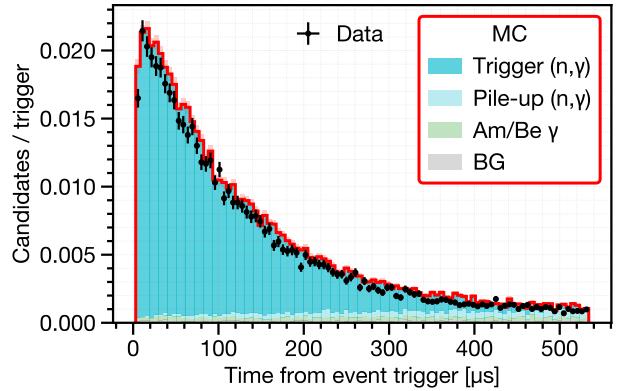


FIG. 9. Exponential decrease of the selected neutron signal candidates as a function of the time from the selected event triggers with the Am/Be neutron source positioned near the ID tank center, in the SK-VI phase. The label “Trigger ( $n, \gamma$ )” indicates captures of neutrons produced within 350 ns from the event trigger, while the label “Pile-up ( $n, \gamma$ )” indicates captures of piled-up neutrons without such correlation to the event trigger. The pink shades indicate MC statistical errors.

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

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

where the normalization constant  $A$ , the background constant  $B$ , the neutron thermalization time scale  $\tau_{\text{thermal}}$ , and the neutron capture time constant  $\tau_{\text{capture}}$  are free parameters. The signal efficiency was evaluated as the number of identified signals per selected event triggers, corrected by the constant background term  $B$ . Figure 10 shows the estimated neutron detection efficiencies for various source positions in the ID.

The major sources of systematic uncertainty are summarized in Table II. In the pure water phases (SK-IV and V), the dominant source of uncertainty lies in the potential bias caused by the calibration setup, such as the unwanted event triggers or time correlation of false positives to event triggers, often leading to an overestimation of the background constant  $B$ . The size of this uncertainty was conservatively estimated by comparing the true and estimated signal efficiencies from the simulations and quantifying the fluctuation within each SK phase. In the Gd-loaded phase (SK-VI), the dominant source of uncertainty is in the fraction and the  $\gamma$  emission model of the  $Gd(n, \gamma)$  reaction. The size of these uncertainties is estimated based on the evaluated thermal neutron capture cross section uncertainties in ENDF/B-VII.1, as well as variations in the estimated signal efficiency when using an alternative photon strength function to describe the Gd continuum gamma cascade in the ANNRI-Gd model [41].

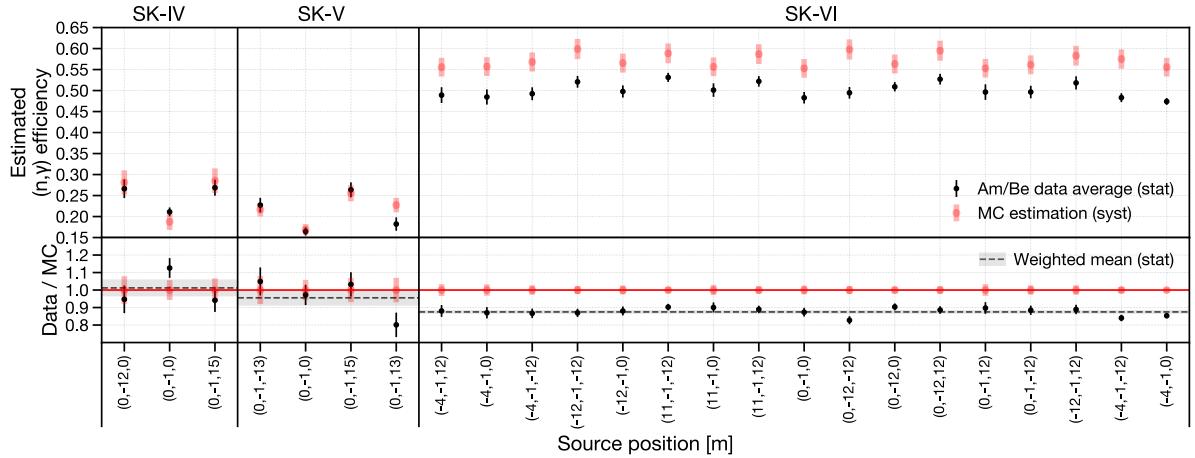


FIG. 10. Estimated  $(n, \gamma)$  signal selection efficiency for each calibration position within the tank. For positions in SK-VI where there are multiple measurements at different dates, the estimated efficiencies were averaged. The systematic uncertainties considered for MC simulation are summarized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

TABLE II. Major sources of systematic uncertainty in the neutron detection efficiency estimated with the Am/Be neutron source. <sup>593</sup>

| Source                                    | SK-IV | SK-V | SK-VI               |
|---|-------|------|---------------------|
| Am/Be neutron characterization            | 0.5%  | 0.9% | 0.5% <sup>596</sup> |
| Detector response                         | 2.2%  | 3.3% | 1.2% <sup>597</sup> |
| Bias due to calibration setup             | 6.9%  | 4.6% | 1.1%                |
| Gd( $n, \gamma$ ) fraction                | -     | -    | 2.1%                |
| Gd( $n, \gamma$ ) $\gamma$ emission model | -     | -    | 2.6% <sup>598</sup> |
| Total                                     | 7.3%  | 5.7% | 3.8% <sup>599</sup> |

## V. $(n, \gamma)$ MULTIPLICITY ESTIMATION

The average multiplicity of total  $(n, \gamma)$  reactions is computed as the average of the expected number of  $(n, \gamma)$  reactions estimated on an event-by-event basis, as follows:

$$\langle N \rangle = \left\langle \frac{N_i^{\text{detected}} - N_i^{\text{BG}}}{\epsilon_i} \right\rangle \quad (4)$$

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

Accurate estimation of  $N_i^{\text{BG}}$  and  $\epsilon_i$  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  $N_i^{\text{BG}}$  and  $\epsilon_i$ , we trained Generalized Additive Models (GAMs) [54] on the baseline atmospheric neutrino simulations. Using pyGAM 0.9.0 [55], a total of six linear GAMs were constructed across the three SK phases, for two output metrics: signal selection efficiency and false positive rate.

Using the baseline MC simulation generated with NEUT 5.4.0, GAMs were fitted to the simulated distributions of each metric, averaged within bins in a five-dimensional feature space. The features consist of reconstructed neutrino event variables, including visible energy, Cherenkov ring multiplicity, the particle type of the most energetic ring, and the radial and vertical displacements of the neutrino interaction vertex. No assumptions were made regarding feature correlations, and appropriate smoothing was applied to mitigate overfitting. The  $1\sigma$  prediction interval was determined based on the MC statistical uncertainties within each bin.

The discrepancy between observed and predicted signal efficiencies in SK-VI (Figure 10) likely arises from an overestimation of the Gd( $n, \gamma$ ) fraction in Geant4 [51, 52]. For example, the Gd( $n, \gamma$ ) fraction estimated using SK-VI Am/Be data was  $(44 \pm 3)\%$  [51], consistent with the  $(43.9 \pm 1.5)\%$  evaluated with thermal neutron capture cross sections in ENDF/B-VII.1 [53], yet both were lower than the 52% predicted by Geant4.10.5.p01 NeutronHP using the same cross sections. This problem with Geant4 is attributed to its treatment of hydrogen as free rather than bound in water molecules, leading to an underestimation of the competing  ${}^1\text{H}(n, \gamma)$  reaction [52].

To address this discrepancy, we applied a correction based on the weighted mean of Am/Be data-to-MC efficiency ratios across all source positions. Signal efficiencies in the pure water phases with limited calibration data were further corrected to ensure consistency in the overall neutron counts with the well-calibrated Gd-loaded phase (see Appendix A for details). The final correction factors were  $0.90 \pm 0.12$  (SK-IV),  $0.94 \pm 0.04$  (SK-V), and  $0.88 \pm 0.01$  (SK-VI).

The measured  $\tau_{\text{capture}}$  in Equation 3 were  $200.4 \pm 3.7$   $\mu\text{s}$  (pure water, SK-IV and SK-V combined) and  $116.9 \pm 0.3$   $\mu\text{s}$  (Gd-loaded, SK-VI), consistent with ENDF/B-VII.1 predictions of  $204.7 \pm 5.3$   $\mu\text{s}$  and  $114.9 \pm 2.5$   $\mu\text{s}$ .

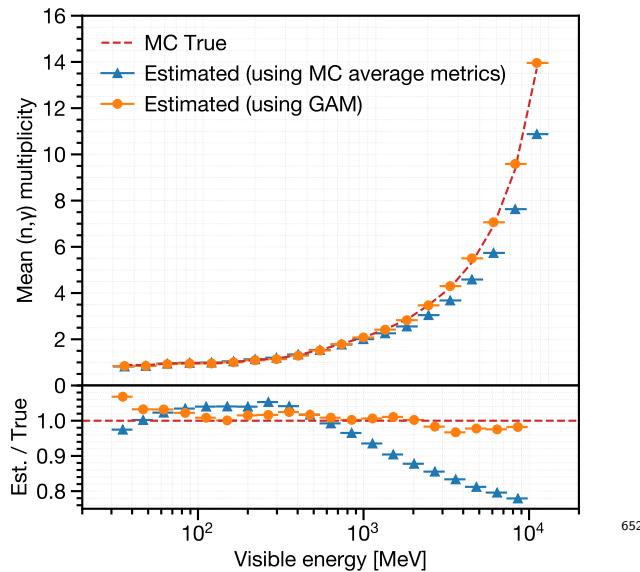


FIG. 11. The true (dashed lines) and estimated average ( $n, \gamma$ ) multiplicity (circle markers: using GAM, triangle markers: using the overall average signal efficiency and false positive rate obtained from the given simulation) as a function of visible energy, for the test simulations of atmospheric neutrino events produced with NEUT 5.1.4.

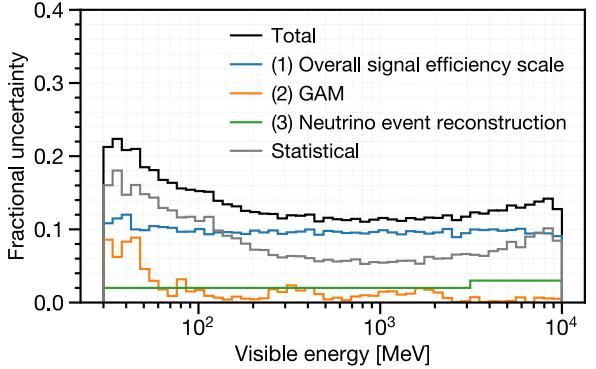


FIG. 12. Fractional uncertainties assigned to the estimated average ( $n, \gamma$ ) multiplicity per bin for the full data sample.

## VI. TESTED MODELS

To compare with data, we generated predictions for the average ( $n, \gamma$ ) multiplicity as a function of neutrino event visible energy using various neutrino event generators and hadron-nucleus interaction models. Six generator configurations were tested: NEUT 5.4.0, NEUT 5.6.3, GENIE hA, GENIE hN, GENIE BERT, and GENIE INCL. The GENIE setups use GENIE 3.4.0 with G\_18a\_10x\_02\_11b physics tunes [25], where  $x \in \{a, b, c, d\}$  corresponds to the FSI model: INTRANUKE/hA [24] (hA), INTRANUKE/hN [24] (hN), the Geant4 Bertini cascade model [56] (BERT), and the Liège INC model [57] (INCL).

NEUT 5.4.0 follows the setup in Section III C, while NEUT 5.6.3 includes a modified nuclear binding energy, slightly reducing CCQE interactions. The GENIE setups share QE and single-pion production models with NEUT but differs in FSI and hadronization. Except for GENIE hA, all FSI models use the full INC approach based on free nucleon cross sections. The Liège INC model is coupled with ABLA07 [58] for nuclear de-excitation.

Secondary hadron-nucleus interaction models were tested using SK detector simulations with GEANT 3.21 and Geant4.10.5.p01. Six configurations were evaluated: SK-IV/V default, SK-VI default, G3 GCALOR, G4 BERT, G4 BERT\_PC, and G4 INCL\_PC. The SK-IV/V and SK-VI setups differ mainly in neutron reaction cross sections below 20 MeV, using ENDF/B-VI and ENDF/B-VII.1, respectively. G3 GCALOR relies entirely on GCALOR within GEANT 3.21, while the baseline setups use the NEUT pion FSI routine. Geant4-based models adopt ENDF/B-VII.1 for low-energy neutron transport, with G4 BERT and G4 BERT\_PC using the Bertini cascade, and G4 INCL\_PC using the Liège INC model. Configurations with the PC suffix employ the Geant4 Pre-compound model [59] for nuclear de-excitation, while G4 BERT uses a simpler native model [56]. GEANT-3-based models rely on Bertini-based tabulation [38] for hadron-nucleus cross sections (excluding low-energy neutrons), whereas Geant4-based models use Glauber parameterization [60].

Trained GAMs were tested on a simulation using NEUT 5.1.4, which predicts roughly 10% lower ( $n, \gamma$ ) multiplicity overall. Figure 11 shows the performance of the trained GAM in estimating the true average signal multiplicity per visible energy bin. By reconstructing  $N_i^{\text{BG}}$  and  $\epsilon_i$  on an event-by-event basis, the GAM helps reduce potential biases in the results, particularly in the multi-GeV bins.

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

### (1) Overall signal efficiency scale

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

### (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 12 shows fractional uncertainties per visible energy bin for the full data sample. The largest uncertainty lies in the overall signal efficiency scale and statistics.

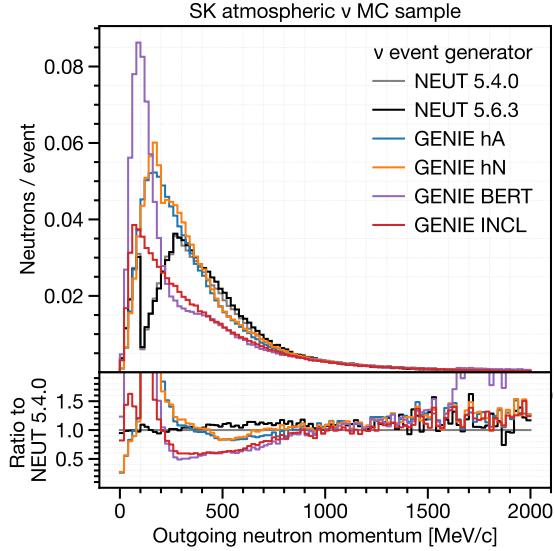


FIG. 13. Comparison of neutrino event generator options with different FSI models: Outgoing neutron momentum distribution per selected atmospheric neutrino event.

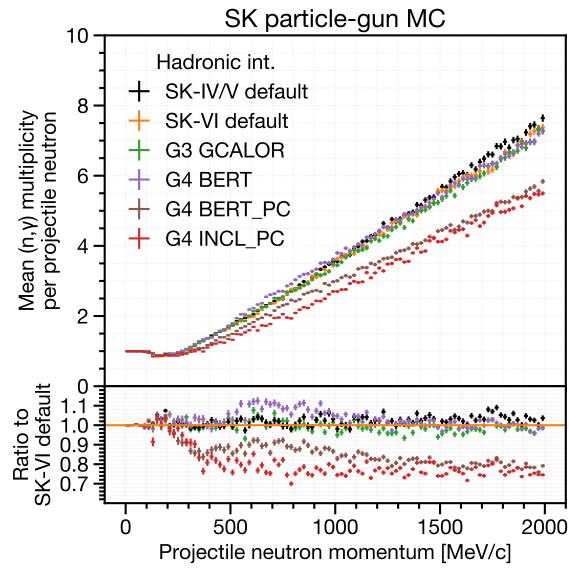


FIG. 14. Comparison of secondary hadron-nucleus interaction models: Average  $(n, \gamma)$  multiplicity per projectile neutron momentum bin, based on simulations using a single-neutron particle-gun setup in water.

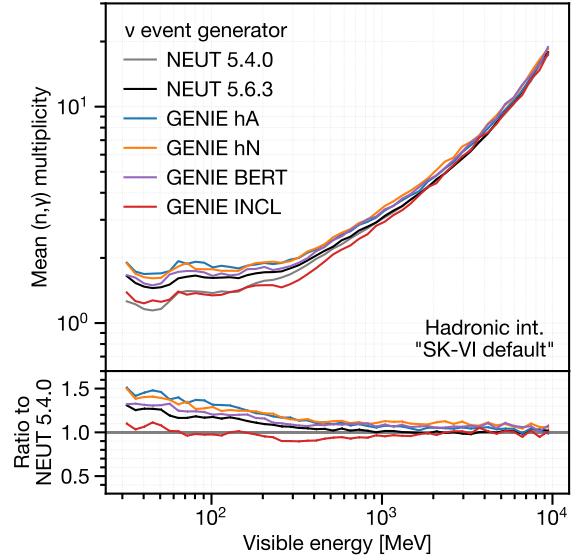


FIG. 15. Comparison of the predicted average  $(n, \gamma)$  multiplicity predicted by the neutrino event generator options paired with **SK-VI default**.

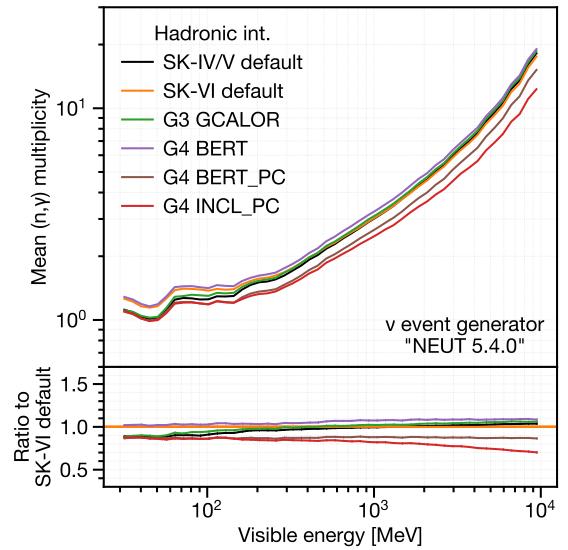


FIG. 16. Comparison of the predicted average  $(n, \gamma)$  multiplicity predicted by the secondary hadron-nucleus interaction models paired with **NEUT 5.4.0**.

For each neutrino event generator, a full MC simulation equivalent to 50 years of atmospheric neutrino exposure was generated using the baseline detector model (**SK-VI default**). To account for variations in secondary hadron-nucleus interactions without additional full simulations, we precomputed average signal multiplicities for neutron-producing projectiles ( $n, p, \pi^\pm$ , and  $\mu^-$ ) in water up to 10 GeV/c, based on particle-gun MC simulations.

Figure 13 compares outgoing neutron momentum distributions predicted by the six neutrino event generator options, highlighting notable differences below 1 GeV/c, which are further discussed in Section VIII.

Figure 14 shows the average signal multiplicities as a function of projectile neutron momentum for the six secondary hadron-nucleus interaction models, with variations of up to 40%.

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.

Figures 15 and 16 show the model predictions. In visible energy range above a few hundred MeV, the average  $(n, \gamma)$  multiplicities are expected to increase linearly with visible energy. Differences across neutrino event generator configurations are more pronounced at visible energies below 1 GeV, while variability among secondary hadron-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**).

## VII. RESULTS

Table III summarizes the number of atmospheric neutrino events and detected neutron signals in the final data sample. The overall average  $(n, \gamma)$  multiplicity per neutrino event,  $\langle N \rangle_{\text{overall}}$ , is estimated using Equation 4 for all selected atmospheric neutrino events without binning.

The combined data was compared with the baseline simulation and neutron production estimates in water from SNO [14], as shown in Figure 17. For visible energies above 200 MeV—where outgoing lepton Cherenkov rings are well reconstructed—a linear relationship between visible energy and the average  $(n, \gamma)$  multiplicity was observed, as expected. While the data aligned well with the SNO estimate, it was 10–30% lower than the baseline predictions in the sub-GeV energy range. A similar deficit was reported by SNO in heavy water [14], where their data was compared to **GENIE 2.10.2 (hA FSI model)** coupled with the **Geant4 Bertini cascade model**.

Predictions from various model combinations, as described in Section VI, were compared with the combined data across different subsamples, as shown in Figure 18. The bottom panel of Figure 18 contrasts the data with predictions from selected FSI models (**GENIE hN** and **INCL**) and secondary interaction models (**G4 BERT**, **G4\_BERT\_PC**, and **G4 INCL\_PC**). The observed deficit in the  $[0.1, 0.4]$  GeV range was primarily found in the single-ring sample, which is expected to be dominated by CCQE interactions (see Figure 5). This deficit was only reproduced when using models that predict fewer neutrons than the baseline for both FSI and secondary interactions, such as **GENIE INCL** and **G4 BERT\_PC**.

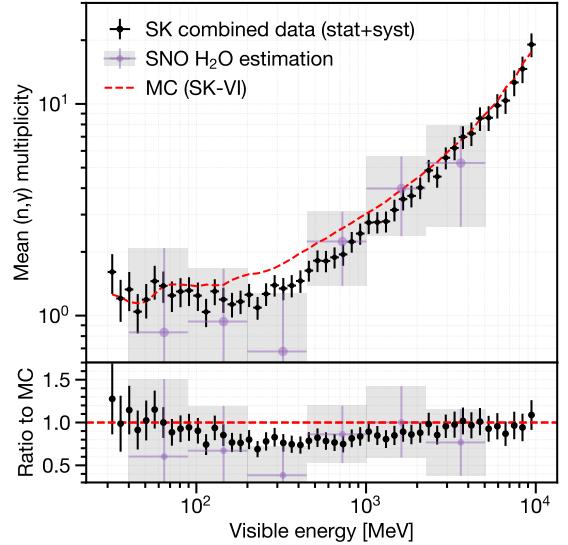


FIG. 17. Average  $(n, \gamma)$  multiplicity in atmospheric neutrino events as a function of their visible energy. Black data points represent combined SK data with statistical and systematic uncertainties. Purple crosses with shades are estimates for pure water based on SNO measurements [14] using a D<sub>2</sub>O target. The red dashed line represents the true average  $(n, \gamma)$  multiplicity from the baseline SK-IV MC simulation.

In the multi-ring sample at a few-GeV energies, there was significant variation in the predicted “slope” of the average  $(n, \gamma)$  multiplicity as a function of visible energy. In particular, **G4 INCL\_PC** predicted a smaller slope compared to **G4 BERT** and **G4\_BERT\_PC**.

The goodness of fit between model predictions and data for the entire visible energy range of  $[0.03, 10]$  GeV was quantified using the following chi-square ( $\chi^2$ ) definition, considering only the two most dominant sources of uncertainty: statistical uncertainty and signal efficiency scale uncertainty.

$$\chi^2 = \sum_i^{N_{\text{bins}}} \frac{(sO_i - E_i)^2}{\sigma_{\text{stat},i}^2} \quad (5)$$

Here,  $i$  is the bin number,  $O_i$  and  $E_i$  represent the observed and expected average  $(n, \gamma)$  multiplicity, respectively, and  $\sigma_{\text{stat},i}$  is the statistical uncertainty in the  $i$ -th bin. The total number of bins,  $N_{\text{bins}}$ , 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  $(n, \gamma)$  multiplicity follows a normal distribution. The normalization scale  $s$  was allowed to float within  $[0, 2]$ . Two  $\chi^2$  calculations were performed: (1) a normalization-free  $\chi^2$ , where  $s$  was minimized without constraint, and (2) a constrained  $\chi^2$ , 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  $\sigma_{\text{scale}}$  set to 10%.

TABLE III. A summary of the number of atmospheric neutrino events (“ $\nu$  events”) and detected neutron signals (“ $n$  signals”) in the final sample.  $\langle N \rangle_{\text{overall}}$  is the unbinned application of Equation 4, followed by signal efficiency scale corrections with the factors given in Section IV B. These factors are specifically chosen to ensure consistency of  $\langle N \rangle_{\text{overall}}$  across the three SK phases (see Appendix A). Errors for the observed  $\langle N \rangle_{\text{overall}}$  are listed as statistical first, followed by systematic uncertainty. Other errors are statistical only. The bottom two rows show the expected  $\langle N \rangle_{\text{overall}}$  and the true overall ( $n, \gamma$ ) multiplicity extracted from the baseline full MC simulations.

|   | SK-IV                    | SK-V                     | SK-VI                    |
|---|--------------------------|--------------------------|--------------------------|
| $\nu$ events                                  | 29,942                   | 4,231                    | 5,203                    |
| Events/day                                    | $9.23 \pm 0.05$          | $9.18 \pm 0.14$          | $9.22 \pm 0.13$          |
| $n$ signals                                   | 15,705                   | 2,035                    | 5,752                    |
| $n$ signals/event                             | $0.525 \pm 0.004$        | $0.481 \pm 0.011$        | $1.106 \pm 0.015$        |
| Observed $\langle N \rangle_{\text{overall}}$ | $2.49 \pm 0.03 \pm 0.23$ | $2.49 \pm 0.10 \pm 0.11$ | $2.49 \pm 0.06 \pm 0.05$ |
| Expected $\langle N \rangle_{\text{overall}}$ | 2.83                     | 2.84                     | 2.85                     |
| True ( $n, \gamma$ ) multiplicity             | 2.85                     | 2.85                     | 2.86                     |

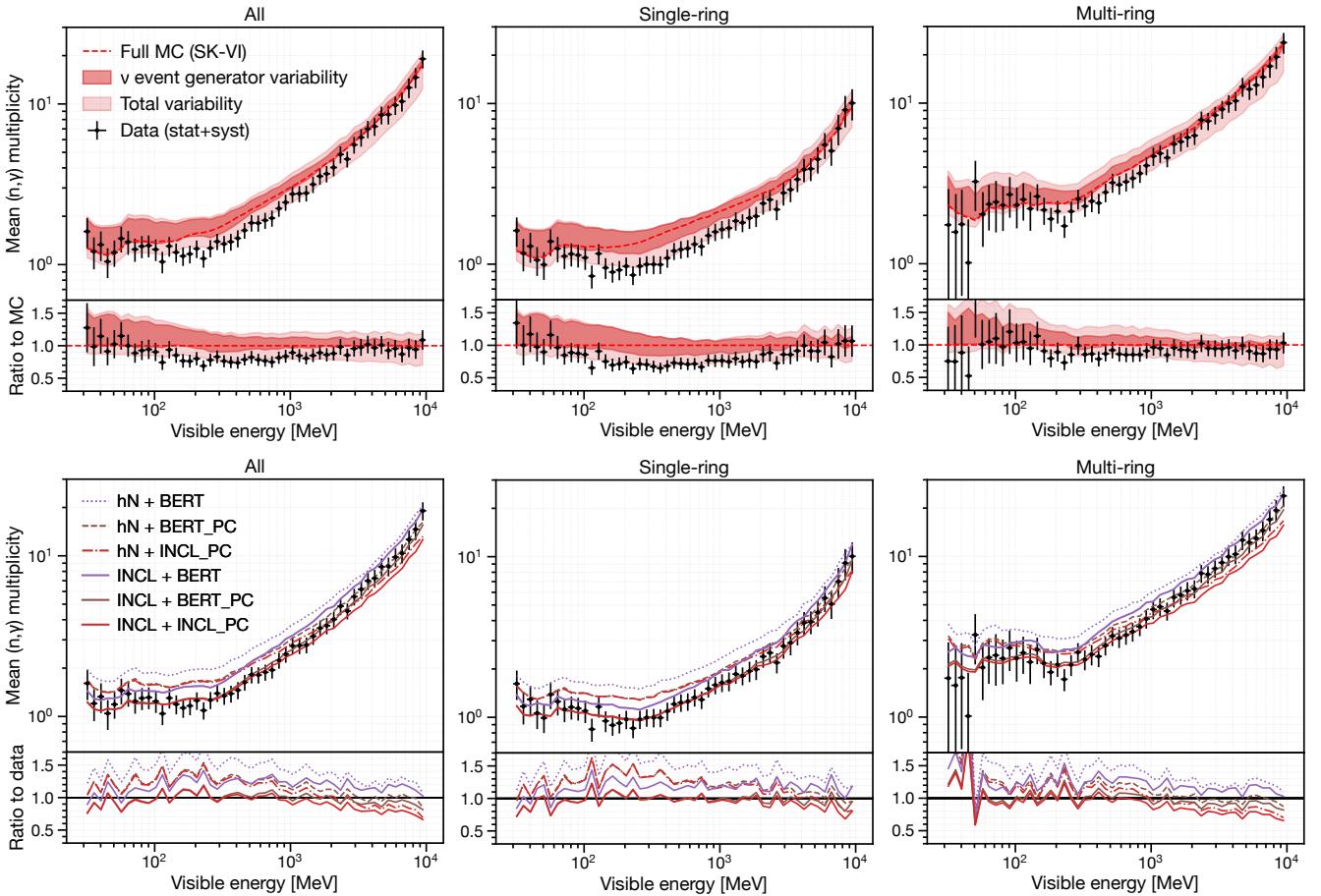


FIG. 18. Comparison of various model predictions with the combined data, showing the average ( $n, \gamma$ ) multiplicity as a function of atmospheric neutrino event visible energy, for all events (left), single-ring events (middle), and multi-ring events (right). The black data points represent the 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 good agreement with the data (see Table IV).

| $\nu$ generator | Secondary int. | Minimizing $s$  |                 | Reduced $\chi^2$ |                 | $p$ -value |                 |
|-----------------|----------------|-----------------|-----------------|------------------|-----------------|------------|-----------------|
|                 |                | Free $s$        | Constrained $s$ | Free $s$         | Constrained $s$ | Free $s$   | Constrained $s$ |
| 1               | NEUT 5.4.0     | SK-IV/V default | 0.75            | 0.90             | 2.76            | 11.66      | $< 10^{-4}$     |
| 2               | NEUT 5.4.0     | SK-VI default   | 0.74            | 0.90             | 3.64            | 14.38      | $< 10^{-4}$     |
| 3               | NEUT 5.4.0     | G3 GCALOR       | 0.73            | 0.90             | 2.14            | 13.47      | $< 10^{-4}$     |
| 4               | NEUT 5.4.0     | G4 BERT         | 0.70            | 0.90             | 2.14            | 20.76      | $< 10^{-4}$     |
| 5               | NEUT 5.4.0     | G4 BERT_PC      | 0.84            | 0.90             | 2.73            | 3.68       | $< 10^{-4}$     |
| 6               | NEUT 5.4.0     | G4 INCL_PC      | 0.89            | 0.90             | 5.08            | 5.11       | $< 10^{-4}$     |
| 7               | NEUT 5.6.3     | SK-IV/V default | 0.72            | 0.90             | 2.92            | 15.94      | $< 10^{-4}$     |
| 8               | NEUT 5.6.3     | SK-VI default   | 0.71            | 0.90             | 4.40            | 19.97      | $< 10^{-4}$     |
| 9               | NEUT 5.6.3     | G3 GCALOR       | 0.71            | 0.90             | 2.24            | 18.30      | $< 10^{-4}$     |
| 10              | NEUT 5.6.3     | G4 BERT         | 0.67            | 0.90             | 2.53            | 27.88      | $< 10^{-4}$     |
| 11              | NEUT 5.6.3     | G4 BERT_PC      | 0.81            | 0.90             | 3.38            | 5.87       | $< 10^{-4}$     |
| 12              | NEUT 5.6.3     | G4 INCL_PC      | 0.86            | 0.90             | 6.38            | 6.86       | $< 10^{-4}$     |
| 13              | GENIE hA       | SK-IV/V default | 0.73            | 0.90             | 3.17            | 14.83      | $< 10^{-4}$     |
| 14              | GENIE hA       | SK-VI default   | 0.67            | 0.90             | 6.71            | 30.43      | $< 10^{-4}$     |
| 15              | GENIE hA       | G3 GCALOR       | 0.71            | 0.90             | 2.47            | 17.69      | $< 10^{-4}$     |
| 16              | GENIE hA       | G4 BERT         | 0.65            | 0.90             | 3.73            | 36.46      | $< 10^{-4}$     |
| 17              | GENIE hA       | G4 BERT_PC      | 0.77            | 0.90             | 5.29            | 10.91      | $< 10^{-4}$     |
| 18              | GENIE hA       | G4 INCL_PC      | 0.80            | 0.90             | 9.64            | 12.93      | $< 10^{-4}$     |
| 19              | GENIE hN       | SK-IV/V default | 0.71            | 0.90             | 1.75            | 16.50      | 0.0030          |
| 20              | GENIE hN       | SK-VI default   | 0.66            | 0.90             | 4.32            | 31.29      | $< 10^{-4}$     |
| 21              | GENIE hN       | G3 GCALOR       | 0.69            | 0.90             | 1.48            | 20.77      | 0.028           |
| 22              | GENIE hN       | G4 BERT         | 0.63            | 0.90             | 2.29            | 40.19      | $< 10^{-4}$     |
| 23              | GENIE hN       | G4 BERT_PC      | 0.76            | 0.90             | 3.13            | 10.10      | $< 10^{-4}$     |
| 24              | GENIE hN       | G4 INCL_PC      | 0.79            | 0.90             | 6.48            | 10.34      | $< 10^{-4}$     |
| 25              | GENIE BERT     | SK-IV/V default | 0.75            | 0.90             | 1.22            | 9.79       | 0.17            |
| 26              | GENIE BERT     | SK-VI default   | 0.70            | 0.90             | 3.31            | 21.35      | $< 10^{-4}$     |
| 27              | GENIE BERT     | G3 GCALOR       | 0.73            | 0.90             | 1.21            | 13.69      | 0.17            |
| 28              | GENIE BERT     | G4 BERT         | 0.67            | 0.90             | 1.79            | 27.70      | 0.0020          |
| 29              | GENIE BERT     | G4 BERT_PC      | 0.80            | 0.90             | 2.56            | 6.06       | $< 10^{-4}$     |
| 30              | GENIE BERT     | G4 INCL_PC      | 0.83            | 0.90             | 5.60            | 7.12       | $< 10^{-4}$     |
| 31              | GENIE INCL     | SK-IV/V default | 0.87            | 0.90             | 0.84            | 1.21       | 0.74            |
| 32              | GENIE INCL     | SK-VI default   | 0.80            | 0.90             | 1.33            | 4.59       | 0.082           |
| 33              | GENIE INCL     | G3 GCALOR       | 0.83            | 0.90             | 1.36            | 2.75       | 0.067           |
| 34              | GENIE INCL     | G4 BERT         | 0.76            | 0.90             | 0.87            | 8.55       | 0.71            |
| 35              | GENIE INCL     | G4 BERT_PC      | 0.92            | 0.92             | 0.95            | 0.96       | 0.56            |
| 36              | GENIE INCL     | G4 INCL_PC      | 0.97            | 0.97             | 2.74            | 2.74       | $< 10^{-4}$     |

TABLE IV. Reduced  $\chi^2$  (see Equation 5) and  $p$ -values for each model combination. The signal efficiency scale  $s$  was either freely minimized or constrained to 10%. The fit has 38 fixed degrees of freedom.  $p$ -values assume Gaussian bin errors without bin correlations.

Table IV compares the reduced  $\chi^2$  values across different model combinations. Among the neutrino event generator options, GENIE INCL showed the lowest  $\chi^2$  for both free and constrained  $s$ . Among the secondary interaction models, G4 BERT\_PC provided the best agreement with data under the 10% normalization constraint. The combination of these two models yielded the best overall agreement under the same constraint. Other secondary interaction models such as SK-IV/V default showed reasonable agreement when coupled with GENIE INCL or when normalization was allowed to float. In contrast, G4 INCL\_PC showed higher unconstrained  $\chi^2$  values, due to the shallower slope prediction for the multi-ring events.

## 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 13, which illustrates differences in neutron momentum distributions across the tested FSI models. By qualitatively comparing the features of these models, we discuss potential factors contributing to the observed discrepancies in predictions.

### A. Nuclear de-excitation

813 867      814 868      815 869      816 870      817 871      818 872      819 873      820 874      821 875      822 876      823 877      824 878      825 879      826 880      827 881      828 882      829 883      830 884      831 885      832 886      833 887      834 888      835 889      836 890      837 891      838 892      839 893      840 894      841 895      842 896      843 897      844 898      845 899      846 900      847 901      848 902      849 903      850 904      851 905      852 906      853 907      854 908      855 909      856 910      857 911      858 912      859 913      860 914      861 915      862 916      863 917      864 918      865 919      866 920      867 921      868 922      869 923      870 924      871 925      872 926      873 927      874 928      875 929      876 930      877 931      878 932      879 933      880 934      881 935      882 936      883 937      884 938      885 939      886 940      887 941      888 942      889 943      890 944      891 945      892 946      893 947      894 948      895 949      896 950      897 951      898 952      899 953      900 954      901 955      902 956      903 957      904 958      905 959      906 960      907 961      908 962      909 963      910 964      911 965      912 966      913 967      914 968      915 969      916 970      917 971      918 972      919 973      920 974      921 975      922 976      923 977      924 978      925 979      926 980      927 981      928 982      929 983      930 984      931 985      932 986      933 987      934 988      935 989      936 990      937 991      938 992      939 993      940 994      941 995      942 996      943 997      944 998      945 999      946 999

Nuclear de-excitation models are crucial for predicting total neutron production, as they govern low-energy neutron emission per nuclear interaction. This is evident when comparing G4 BERT and G4 BERT\\_PC, which differ only in their de-excitation treatment. For instance, the inaccuracy of the Geant4 Bertini cascade model prediction is indicated by the distinct large neutron peak below 250 MeV/c in GENIE BERT (Figure 13), compared to the other models such as ABLA07 in GENIE INCL or the native de-excitation model in NEUT. (The GENIE hA and hN options omit nuclear de-excitation entirely.) Coupling the Geant4 Precompound model with the Geant4 Bertini cascade model significantly reduces neutron production across all visible energy ranges, leading to much better agreement with data (Figure 18).

In the Geant4 Bertini cascade model, neutron emission follows the Weisskopf statistical evaporation model [61]. A key input determining neutron emission probability is the inverse reaction (neutron absorption) cross section. The inaccuracy of the Bertini model and its variants in predicting neutron evaporation may stem from its use of the simplified cross section parameterization by Dostrovsky [62]. In comparison, the Geant4 Precompound model uses a more refined parameterization based on the nuclear optical model potential [63], fitted to a broader dataset [64], resulting in lower evaporation rates particularly for low-energy neutrons. Additionally, the Geant4 Precompound model features tuned level density parameters and allows Fermi breakup in nuclei up to mass 16 [59], including oxygen, further reducing isolated neutrons that contribute to  $(n, \gamma)$  signals.

### B. Nuclear in-medium considerations

A key distinction among the FSI cascade models is their treatment of Pauli blocking, which prevents nucleons from occupying the same quantum state, as dictated by the Pauli exclusion principle. For instance, NEUT 5.4.0, NEUT 5.6.3 and GENIE BERT implement “strict” Pauli blocking, which forbids all nucleon scattering below the predefined oxygen Fermi momentum around 225 MeV/c, modeling the nucleus as a degenerate Fermi gas. This approach results in a characteristic dip in the outgoing neutron momentum distributions for NEUT 5.4.0 and NEUT 5.6.3 within the corresponding energy range, as shown in Figure 13. The Liège INC model applies a probabilistic approach for nucleon collisions throughout the cascade (except for the initial collision), accounting for nucleon holes in the surrounding phase space volume [65, 66]. This results in smoother neutron distributions near the Fermi momentum, as seen in Figure 13. In contrast, the GENIE hA and hN models omit Pauli blocking entirely, leading to significantly higher predicted neutron production than NEUT 5.4.0, NEUT 5.6.3, and GENIE INCL.

In the 0.2–1 GeV/c range, the number of outgoing neutrons predicted by the Liège INC and Geant4 Bertini cascade models is lower than that from the GENIE hA, hN, and NEUT nucleon FSI models, as illustrated in Figure 13. One contributing factor may be the inclusion of light cluster formation—such as deuterons, tritons, and alpha particles—among cascade products, which reduces the number of isolated neutrons contributing to the observable signal. Ref. [67] also identifies cluster formation as a major mechanism behind the reduction of outgoing protons in the Liège INC model. Additional in-medium effects, considered either explicitly or implicitly in the two models—such as short-range nucleon repulsion and local nuclear density reduction after each collision—further reduce the number of nucleon-nucleon collisions during the cascade.

### C. $\pi^\pm$ production

Accurate modeling of pion-nucleus interactions is crucial for predicting neutron production in multi-ring events with energetic pions (see Figures 24 and 25 in Appendix C). The difference in predicted neutron production in the few-GeV multi-ring sample between the Liège INC model (G4 INCL\\_PC) and the Geant4 Bertini cascade model (G4 BERT\\_PC), as observed in Figure 18, can be attributed to differences in pion production cross sections for pion projectiles. Our results are consistent with those of Ref. [68], which report that the Liège INC model tends to predict lower inclusive  $\pi^\pm$  production cross sections on light nuclei (mass number  $A < 20$ ) compared to the Geant4 Bertini cascade model. In addition, the pion potential within the nucleus differs between the two models. In the Liège INC model, the pion potential is generally deeper [69] than the constant 7 MeV potential used in the Geant4 Bertini cascade model [56], which impacts pion transparency. For instance, Ref. [70] demonstrated that the Liège INC model shows lower pion transparency for pions with momenta below 300 MeV/c compared to the GENIE hA, hN, or NEUT pion FSI models.

### D. Other considerations

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) result in a 5–10% difference in the average  $(n, \gamma)$  multiplicities for visible energy below 300 MeV, as shown in Figure 16. Our model predictions do not account for interactions of  $\nu_\tau$  and  $\bar{\nu}_\tau$  arising from oscillations; however, their contribution is expected to be smaller than our uncertainty budget in the multi-GeV visible energy region, where their effects become relevant.

916

## IX. CONCLUSIONS AND PROSPECTS

969

## ACKNOWLEDGMENTS

We gratefully acknowledge the cooperation of the Kamioka Mining and Smelting Company. The Super-Kamiokande experiment has been built and operated from funding by the Japanese Ministry of Education, Culture, Sports, Science and Technology; the U.S. Department of Energy; and the U.S. National Science Foundation. Some of us have been supported by funds from the National Research Foundation of Korea (NRF-2009-0083526, NRF-2022R1A5A1030700, NRF-2202R1A3B1078756) funded by the Ministry of Science, Information and Communication Technology (ICT); the Institute for Basic Science (IBS-R016-Y2); and the Ministry of Education (2018R1D1A1B07049158, 2021R1I1A1A01042256, 2021R1I1A1A01059559, RS-2024-00442775); the Japan Society for the Promotion of Science; the National Natural Science Foundation of China under Grants No.12375100; the Spanish Ministry of Science, Universities and Innovation (grant PID2021-124050NB-C31); the Natural Sciences and Engineering Research Council (NSERC) of Canada; the Scinet and Westgrid consortia of Compute Canada; the National Science Centre (UMO-2018/30/E/ST2/00441 and UMO-2022/46/E/ST2/00336) and the Ministry of Science and Higher Education (2023/WK/04), Poland; the Science and Technology Facilities Council (STFC) and Grid for Particle Physics (GridPP), UK; the European Union’s Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant agreement no.754496; H2020-MSCA-RISE-2018 JENNIFER2 grant agreement no.822070, H2020-MSCA-RISE-2019 SK2HK grant agreement no. 872549; and European Union’s Next Generation EU/PRTR grant CA3/RSUE2021-00559; the National Institute for Nuclear Physics (INFN), Italy.

917 Accurately modeling neutron production in neutrino<sup>971</sup>  
 918 interactions is essential for characterizing incoming neu<sup>972</sup>  
 919 trinos, which is crucial for advancing precision measure<sup>973</sup>  
 920 ments of neutrino oscillation parameters and rare event<sup>974</sup>  
 921 searches involving neutron tagging. Recent studies have<sup>975</sup>  
 922 suggested potential inaccuracies in the modeling of sec<sup>976</sup>  
 923 ondary hadron-nucleus interactions.<sup>977</sup>

924 This paper reports a measurement of total neutron<sup>979</sup>  
 925 production following atmospheric neutrino interactions<sup>980</sup>  
 926 within the water volume of the Super-Kamiokande (SK)<sup>981</sup>  
 927 detector. Atmospheric neutrino events were binned<sup>982</sup>  
 928 by their electron-equivalent “visible energy,” a semi-<sup>983</sup>  
 929 calorimetric proxy for neutrino momentum transfer, in<sup>984</sup>  
 930 the range of [0.03, 10] GeV. The dominant systematic<sup>985</sup>  
 931 uncertainty was the roughly 10% uncertainty in the over-<sup>986</sup>  
 932 all signal efficiency scale. The average neutron cap-<sup>987</sup>  
 933 ture multiplicity in each visible energy bin was compared<sup>988</sup>  
 934 against predictions from various combinations of neutrino<sup>989</sup>  
 935 event generators and secondary hadron-nucleus interac-<sup>990</sup>  
 936 tion models.<sup>991</sup>

937 Our data provides strong discriminative power for eval<sup>993</sup>  
 938 uating these models, with predictions varying by up to<sup>994</sup>  
 939 50%. Two key observations were made: a reduction in<sup>995</sup>  
 940 neutron production for sub-GeV single-ring events, where<sup>996</sup>  
 941 secondary hadron interactions are minimal, and a nearly<sup>997</sup>  
 942 linear increase in neutron production with visible energy,<sup>998</sup>  
 943 especially in the multi-GeV region. The first observa-<sup>999</sup>  
 944 tion highlights the need for moderate neutron evapora-<sup>1000</sup>  
 945 tion and consideration of nuclear in-medium effects. The<sup>1001</sup>  
 946 second is more sensitive to pion production in cascade<sup>1002</sup>  
 947 models, which become important at higher energies.<sup>1003</sup>

948 This study highlights the crucial role of hadron-nucleus  
 949 interaction models in accurately predicting total neu<sup>1004</sup>  
 950 tron production from neutrino interactions. The ob-<sup>1005</sup>  
 951 served discrepancy with the evaporation model [62] com-<sup>1005</sup>  
 952 monly used in Bertini-based cascade models aligns with<sup>1006</sup>  
 953 previous studies, including neutron measurements from<sup>1007</sup>  
 954 SK neutral-current quasielastic selections [71] and sec-<sup>1008</sup>  
 955 ondary gamma-ray measurements using Germanium de-<sup>1009</sup>  
 956 tectors with a neutron beam on a water target [72]. Our<sup>1010</sup>  
 957 data also supports the selection of specific models, such<sup>1011</sup>  
 958 as GENIE INCL and G4 BERT\_PC, which reduce the un-<sup>1012</sup>  
 959 certainty in total neutron production from atmospheric<sup>1013</sup>  
 960 neutrino events to around 10%. These models may help<sup>1014</sup>  
 961 explain the deficits in total neutron capture signals re-<sup>1015</sup>  
 962 ported by T2K [13] and SNO [14], and offer insights<sup>1016</sup>  
 963 into the low-energy neutron signal deficits observed in<sup>1017</sup>  
 964 MINERvA [2]. Lastly, this work will improve the use of<sup>1018</sup>  
 965 neutron tagging for identifying atmospheric antineutrino<sup>1019</sup>  
 966 events at SK, enhancing sensitivity to neutrino mass or<sup>1020</sup>  
 967 dering and searches for proton decays and the diffuse<sup>1021</sup>  
 968 supernova neutrino background.<sup>1022</sup>

### Appendix A: $\langle N \rangle_{\text{overall}}$ consistency across SK phases

When applying a signal efficiency correction based on the weighted mean of Am/Be data-to-MC efficiency ratios—as shown in Figure 10 and described in Section IV B—the efficiency-corrected estimate of total neutron production  $\langle N \rangle_{\text{overall}}$  was  $2.21 \pm 0.03 \text{ (stat)} \pm 0.11 \text{ (syst)}$  in SK-IV, lower than the values observed in the other two phases:  $2.46 \pm 0.10 \pm 0.11$  for SK-V and  $2.50 \pm 0.06 \pm 0.05$  for SK-VI. This discrepancy may be due to unaccounted systematic uncertainties in SK-IV, which had the longest data-taking period (nearly 10 years) but only a limited amount of calibration data to constrain the entire duration. In contrast, SK-VI—with a much shorter runtime (about 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.

To test the robustness of the signal efficiency correction derived from the Am/Be calibration data and simulation,

we applied an alternative neutron signal detection algorithm to both the calibration and atmospheric neutrino datasets in SK-VI. This algorithm uses a likelihood-based vertex fitter [73] to reconstruct the  $Gd(n, \gamma)$  vertex independently of the neutrino interaction vertex, making it more resilient to uncertainties in outgoing neutron kinematics that affect the detection efficiency in the baseline algorithm described in Section IV A. Using this reference algorithm, we obtained  $\langle N \rangle_{\text{overall}} = 2.49 \pm 0.06 \pm 0.05$  for SK-VI, consistent with the baseline result.

Relying on the signal efficiency scale obtained for SK-VI, we adjust the correction factors for all phases such that  $\langle N \rangle_{\text{overall}}$  aligns with the SK-VI reference value of 2.49. The difference between the calibration-based and phase-consistency-based correction factors is treated as an independent systematic uncertainty in the signal efficiency scale and is included in the uncertainty budget shown in Figure 12. This yields a final  $\langle N \rangle_{\text{overall}}$  of  $2.49 \pm 0.03 \text{ (stat)} \pm 0.26 \text{ (syst)}$ , corresponding to a total systematic uncertainty of approximately 10%. Figure 19 illustrates the impact of the two signal efficiency correction approaches: one based on Am/Be neutron source calibration and the other based on phase consistency.

## Appendix B: Effective metrics for model evaluation

To evaluate model performance, we define two effective metrics. The first is the slope of the linear increase in average  $(n, \gamma)$  multiplicity as a function of visible energy, obtained by performing a linear fit over the range [0.3, 10] GeV, as shown in Figure 20. At lower visible energies, the modeling of low-energy nucleon FSI becomes more important (see Figure 23 in Appendix C). To quantify neutron production in this regime, we define the low-energy (low-E) multiplicity as the average  $(n, \gamma)$  multiplicity in the [0.1, 0.3] GeV visible energy range. Events below 0.1 GeV are excluded due to large systematic uncertainties.

Figure 21 compares the measured slopes and low-energy multiplicities across different SK phases and event topologies. For the slope fits shown in the figure, the binning scheme was adjusted to ensure sufficient statistics (typically  $\gtrsim 30$  neutrino events per bin). The measured values are consistent across SK phases within assigned uncertainties. Both metrics are higher for multi-ring events than for single-ring events, as expected due to the higher fraction of deep inelastic scattering (DIS) interactions in the multi-ring sample (see Figure 5). While the observed slopes agree with baseline model predictions, the measured low-energy multiplicities were lower than expected.

Figure 22 compares the measured slopes and low-energy multiplicities with model predictions. The predictions show 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 18, 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  $1\sigma$  uncertainty.

## Appendix C: Sources of neutron production

Here, we clarify the sources of neutron production in our simulation. Figure 23 shows the baseline SK-IV neutron production as a function of visible energy. The number of outgoing neutrons from neutrino-nucleus interactions remains nearly constant across visible energies. At lower energies, the outgoing neutrons have small momenta, with each neutron corresponding closely to a single  $(n, \gamma)$  reaction. At higher energies, larger hadron momentum results in a linear increase in secondary neutron production. This indicates that, for the two metrics defined in Appendix B, the low-energy multiplicity is sensitive to nucleon transport modeling, while the slope is primarily determined by the secondary hadron-nucleus interaction model.

Figure 24 shows the contributions of nucleons and pions to the  $(n, \gamma)$  multiplicity as a function of visible energy, for different secondary hadron-nucleus interaction models. A general trend across all models is that, in the single-ring sample, most  $(n, \gamma)$  reactions are induced by outgoing nucleons from neutrino interactions, while in the multi-ring sample, outgoing pions dominate in the few-GeV range. Figure 25 further breaks down the contributions to the  $(n, \gamma)$  multiplicity by outgoing particle momentum, for two FSI models (GENIE INCL and GENIE hN) and two secondary interaction models (G4 INCL\_PC and G4 BERT\_PC). In single-ring atmospheric neutrino events with visible energy in the range [0.1, 0.3] GeV, the total  $(n, \gamma)$  multiplicity is primarily driven by neutrons below 1 GeV/c (from neutrino-nucleus interactions), with negligible variation across secondary interaction models. However, in multi-ring events with visible energies above 2 GeV, the total contribution from outgoing pions at all momenta becomes comparable to that from neutrons.

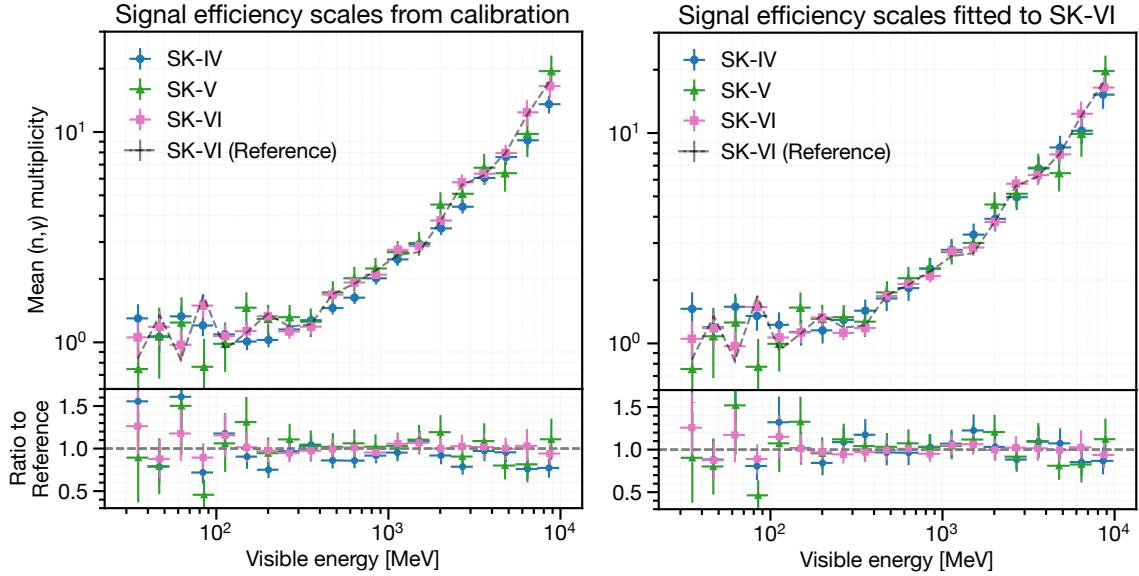


FIG. 19. Average  $(n, \gamma)$  multiplicities plotted against neutrino event visible energy across SK phases. The left plot uses signal efficiency scales from Am/Be calibration (Section IV B), while the right plot adjusts SK-IV and SK-V scales to match  $\langle N \rangle_{\text{total}}$  of SK-VI. 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.

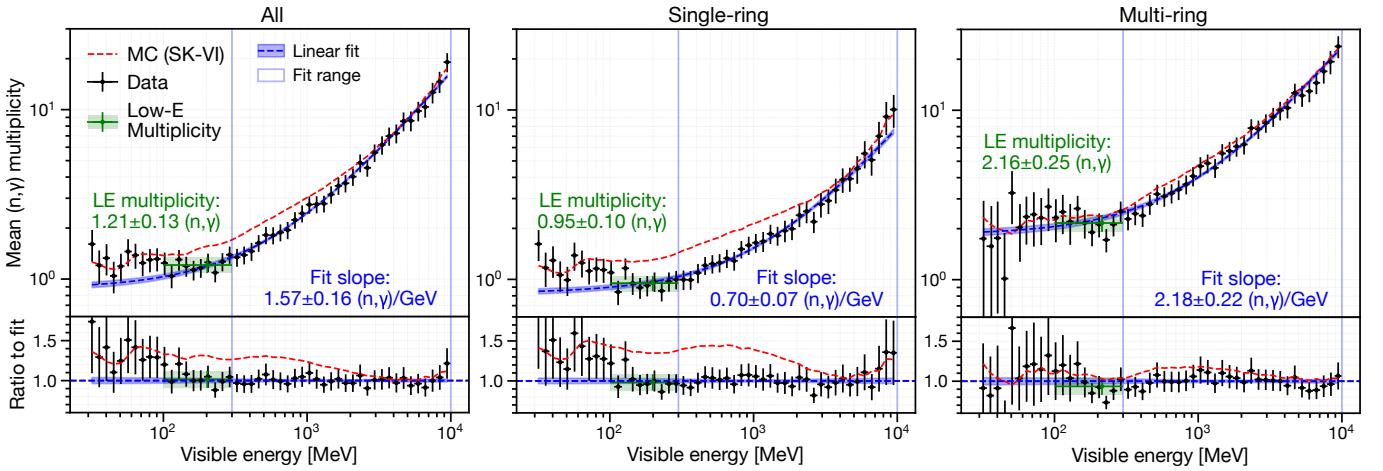


FIG. 20. Average  $(n, \gamma)$  multiplicity as a function of visible energy for neutrino events: all events (left), single-ring events (center), and multi-ring events (right). Red dashed lines show predictions from the full MC simulation in the SK-VI baseline setup. Blue dashed lines represent fitted linear functions with  $1\sigma$  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.

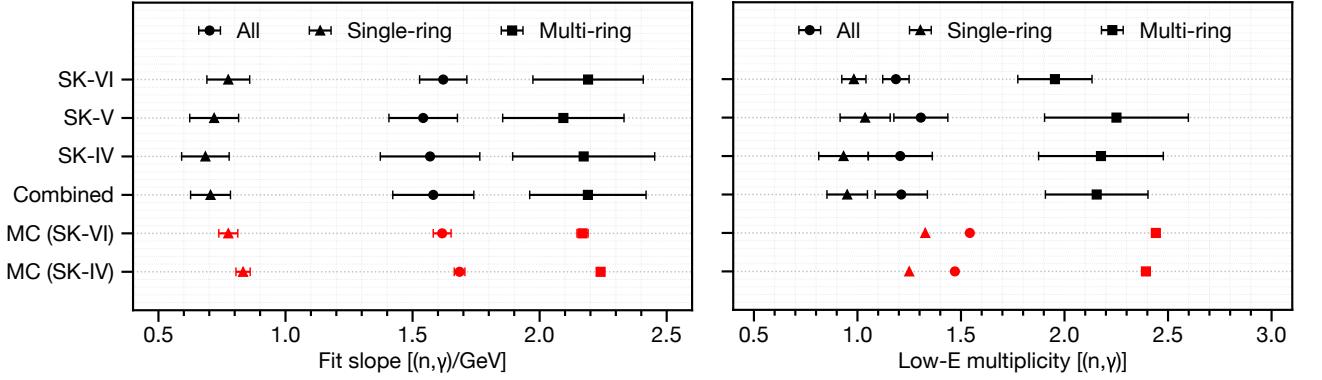


FIG. 21. Measured slopes (left) and low-energy multiplicities (right) across different SK phases and event types. Data error bars (black) represent both statistical and systematic uncertainties, while MC error bars in the left panel indicate statistical uncertainties. In the right panel, MC predictions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

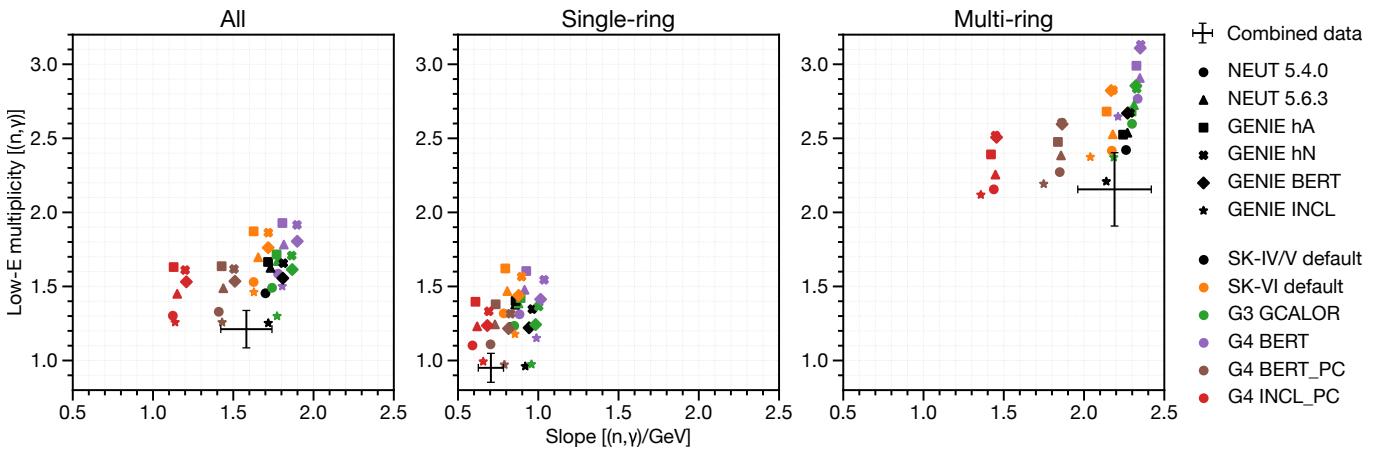


FIG. 22. Scatter plots of fitted slopes and low-energy multiplicities in model predictions and data (black crosses, error bars include both statistical and systematic uncertainties), shown for all events (left), single-ring events (middle), and multi-ring events (right). Shapes represent neutrino event generator options used, while colors indicate the secondary hadron-nucleus interaction model options used.

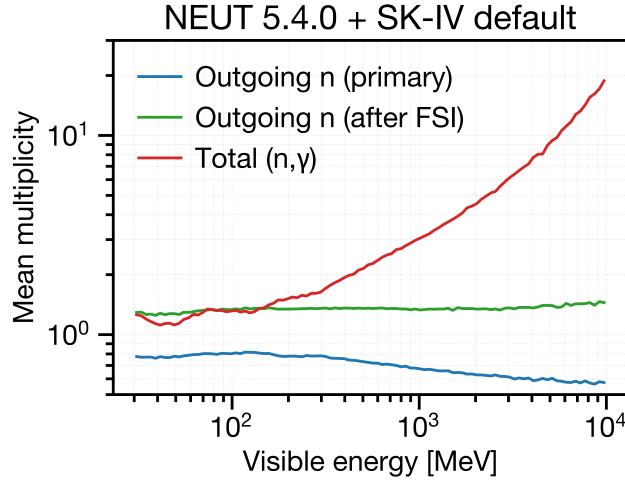


FIG. 23. Mean multiplicity of outgoing neutrons from the primary neutrino interaction (blue) and subsequent FSIs (green) and resulting total ( $n, \gamma$ ) reactions (red), in the baseline SK-IV simulation setup.

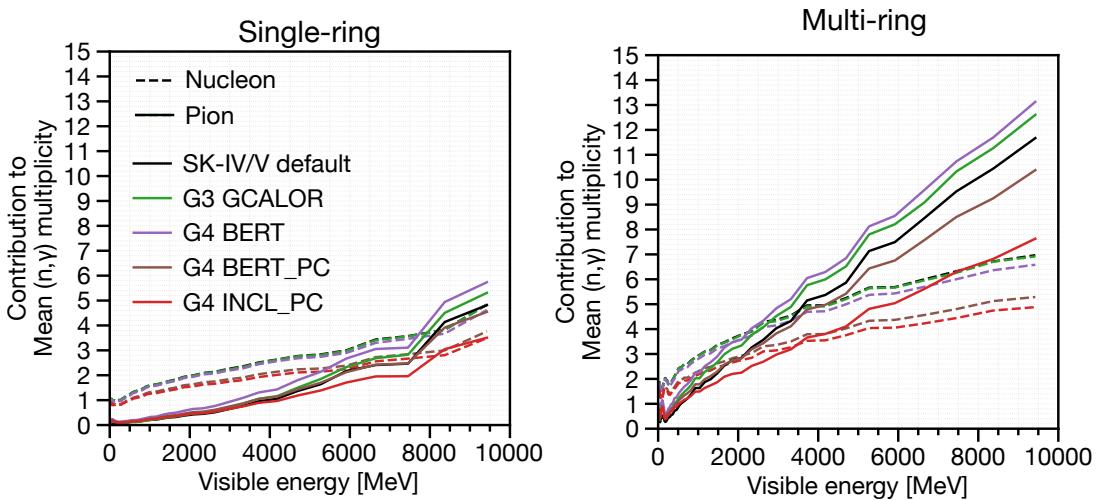


FIG. 24. Average contribution of outgoing nucleons (dashed lines) and pions (solid lines) to the observed ( $n, \gamma$ ) multiplicities as a function of visible energy, shown separately for single-ring events (top) and multi-ring events (bottom). The line colors represent different secondary hadron-nucleus interaction models used in the predictions.

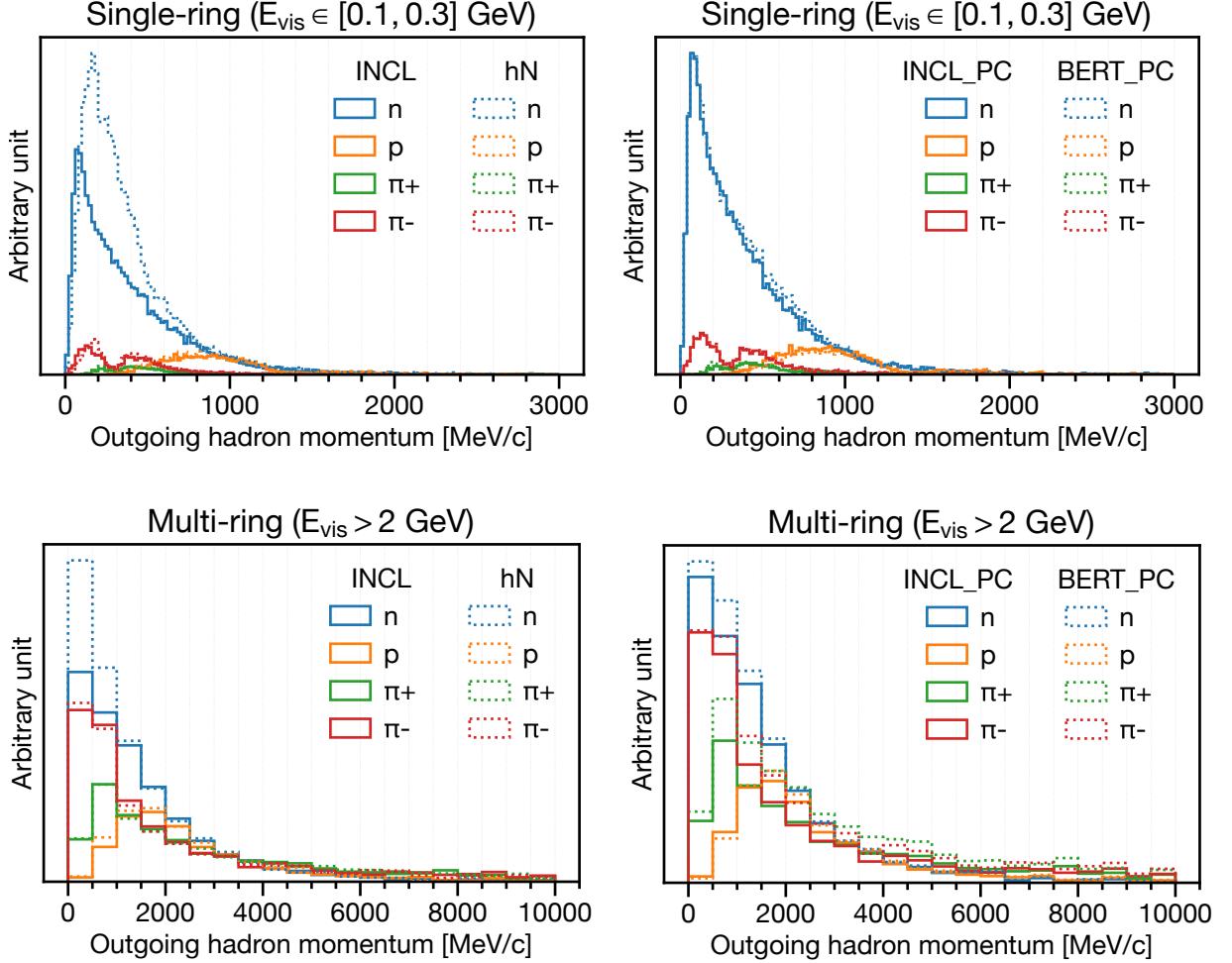


FIG. 25. Momentum distributions of outgoing hadrons predicted by neutrino event generators, weighted by the average  $(n, \gamma)$  multiplicity per projectile hadron momentum bin as predicted by secondary hadron-nucleus interaction models. The area under each histogram bin represents the average contribution of a specific hadron at a given momentum to the observable  $(n, \gamma)$  multiplicity. The top panels show predictions for single-ring events with visible energy ( $E_{\text{vis}}$ ) in the range  $[0.1, 0.3] \text{ GeV}$ , while the bottom panels correspond to multi-ring events with visible energy above  $2 \text{ GeV}$ . 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.

- [1] P. Abratenko *et al.* (MicroBooNE collaboration), Phys. Rev. D **110**, 013006 (2024). 1185  
1186
- [2] M. Elkins *et al.* (MINERvA collaboration), Phys. Rev. D **100**, 052002 (2019). 1187  
1188
- [3] P. Abratenko *et al.* (MicroBooNE collaboration), The European Physical Journal C **84**, 1052 (2024). 1189  
1190
- [4] A. M. Ankowski, P. Coloma, P. Huber, C. Mariani, and E. Vagnoni, Phys. Rev. D **92**, 091301 (2015). 1191  
1192
- [5] C. L. Cowan, F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire, Science **124**, 103 (1956) 1193  
1194  
<https://www.science.org/doi/pdf/10.1126/science.124.3212.103>.
- [6] T. Wester *et al.* (Super-Kamiokande collaboration), Phys. Rev. D **109**, 072014 (2024). 1195  
1196
- [7] A. Takenaka *et al.* (Super-Kamiokande collaboration), Phys. Rev. D **102**, 112011 (2020). 1198  
1199
- [8] S. Agostinelli *et al.* (Geant4 collaboration), Nucl. Instrum. Methods A **506**, 250 (2003). 1200  
1201
- [9] J. Allison *et al.*, IEEE Transactions on Nuclear Science **53**, 270 (2006). 1202  
1203
- [10] J. Allison *et al.*, Nucl. Instrum. Methods A **835**, 186 (2016). 1204  
1205
- [11] J. Cugnon, Nucl. Phys. A **387**, 191 (1982). 1206  
1207
- [12] J.-C. David *et al.*, Progress In Nuclear Science and Technology **2**, 942 (2011). 1208  
1209
- [13] 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). 1210  
1211
- [14] B. Aharmim *et al.* (SNO collaboration), Phys. Rev. D **99**, 112007 (2019). 1212  
1213
- [15] A. Olivier *et al.* (MINERvA collaboration), Phys. Rev. D **108**, 112010 (2023). 1214  
1215
- [16] S. Abe *et al.* (KamLAND Collaboration), Phys. Rev. D **107**, 072006 (2023). 1216  
1217
- [17] M. B. Chadwick *et al.*, Nuclear Data Sheets **112**, 2887 (2011), special Issue on ENDF/B-VII.1 Library. 1218  
1219
- [18] K. Abe *et al.* (Super-Kamiokande collaboration), Nucl. Instrum. Methods A **1027**, 166248 (2022). 1220  
1221
- [19] K. Abe *et al.*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **1065**, 169480 (2024). 1222  
1223
- [20] S. Fukuda *et al.* (Super-Kamiokande collaboration), Nucl. Instrum. Methods A **501**, 418 (2003). 1224  
1225
- [21] K. Abe *et al.*, Nucl. Instrum. Methods A **737**, 253–272 (2014). 1226  
1227
- [22] S. Yamada *et al.*, IEEE Trans. Nucl. Sci. **57**, 428 (2010) 1228  
1229
- [23] C. Andreopoulos *et al.*, Nucl. Instrum. Methods A **614**, 87 (2010), arXiv:0905.2517 [hep-ph]. 1230  
1231
- [24] C. Andreopoulos *et al.*, The GENIE Neutrino Monte Carlo Generator: Physics and User Manual (2015) 1232  
1233  
arXiv:1510.05494 [hep-ph].
- [25] J. Tena-Vidal *et al.* (GENIE), Phys. Rev. D **104**, 072009 (2021), arXiv:2104.09179 [hep-ph]. 1234  
1235
- [26] M. Shiozawa, Nucl. Instrum. Methods A **433**, 240 (1999) 1236  
1237
- [27] M. Honda, T. Kajita, K. Kasahara, and S. Midorikawa, Phys. Rev. D **83**, 123001 (2011). 1238  
1239
- [28] R. L. Workman *et al.* (Particle Data Group), PTEP **083C01**, 729 (2022). 1240  
1241
- [29] Y. Hayato and L. Pickering, Eur. Phys. J Spec. Top. **230**, 4469–4481 (2021). 1242  
1243
- [30] L. L. Salcedo, E. Oset, M. J. Vicente-Vacas, and C. Garcia-Recio, Nucl. Phys. A **484**, 557 (1988). 1244  
1245
- [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) iINDC(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] S. Han, Measurement of neutrons produced in atmospheric neutrino interactions at super-kamiokande, Ph.D. thesis, The Univ. of Tokyo (2024).
- [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] B. Pritychenko and S. F. Mughabghab, Nuclear Data Sheets **113**, 3120 (2012).
- [54] T. Hastie and R. Tibshirani, *Generalized Additive Models*, Chapman & Hall/CRC Monographs on Statistics & Applied Probability (Taylor & Francis, 1990).
- [55] D. Servén, C. Brummitt, and H. Abedi, dswah/pygam: v0.8.0 (2018).
- [56] D. Wright and M. Kelsey, Nucl. Instrum. Methods A **804**,

- 1246 175 (2015).  
 1247 [57] D. Mancusi, A. Boudard, J. Cugnon, J.-C. David, P. Kai<sub>1267</sub>  
 1248 taniemi, and S. Leray, Phys. Rev. C **90**, 054602 (2014).<sub>1268</sub>
- 1249 [58] A. Kelic, M. V. Ricciardi, and K.-H. Schmidt, Abla07<sub>1269</sub>  
 1250 towards a complete description of the decay channels of<sub>270</sub>  
 1251 a nuclear system from spontaneous fission to multifrag<sub>271</sub>  
 1252 mentation (2009), arXiv:0906.4193 [nucl-th].<sub>1272</sub>
- 1253 [59] J. Quesada Molina *et al.*, Progress in Nuclear Science and<sub>273</sub>  
 1254 Technology **2**, 936 (2011).<sub>1274</sub>
- 1255 [60] V. M. Grichine, The European Physical Journal C **62**<sub>1275</sub>  
 1256 399 (2009).<sub>1276</sub>
- 1257 [61] V. Weisskopf, Phys. Rev. **52**, 295 (1937).<sub>1277</sub>
- 1258 [62] I. Dostrovsky, Z. Fraenkel, and G. Friedlander, Phys<sub>1278</sub>  
 1259 Rev. **116**, 683 (1959).<sub>1279</sub>
- 1260 [63] A. Chatterjee, K. H. N. Murthy, and S. K. Gupta, Pra<sub>1280</sub>  
 1261 mana **16**, 391 (1981).<sub>1281</sub>
- 1262 [64] C. Kalbach, Preco-2000, exciton model preequilib<sub>1282</sub>  
 1263 rium code system with direct reactions. preco-d2, pre<sub>1283</sub>  
 1264 equilibrium and direct reaction double differential cross<sub>1284</sub>  
 1265 sections (2002).<sub>1285</sub>
- [65] A. Boudard and J. Cugnon, Incl4 - the liège inc model  
 for high-energy hadron-nucleus reactions: A short de-  
 scription of the incl4.2 and incl4.4 versions (2008).
- [66] A. Boudard, J. Cugnon, J.-C. David, S. Leray, and  
 D. Mancusi, Phys. Rev. C **87**, 014606 (2013).
- [67] A. Ershova *et al.*, Phys. Rev. D **106**, 032009 (2022).
- [68] D. Mancusi *et al.*, The European Physical Journal A **53**,  
 80 (2017).
- [69] T. Aoust and J. Cugnon, Phys. Rev. C **74**, 064607 (2006).
- [70] S. Dytman, Y. Hayato, R. Raboanary, J. T. Sobczyk,  
 J. Tena-Vidal, and N. Vololoniaina, Phys. Rev. D **104**,  
 053006 (2021).
- [71] S. Sakai *et al.* (Super-Kamiokande Collaboration), Phys.  
 Rev. D **109**, L011101 (2024).
- [72] Y. Hino, A comparison of n-o inelastic scattering between  
 the experimental and simulation models towards under-  
 standing neutrino reaction (2024).
- [73] M. Smy, in *International Cosmic Ray Conference*, Interna-  
 tional Cosmic Ray Conference, Vol. 5 (2008) pp.  
 1279–1282.