

## Measurement of muon-induced high-energy neutrons from rock in an underground Gd-doped water detector

F. Sutanto <sup>1,2,\*</sup>, O. A. Akindele <sup>2</sup>, M. Askins <sup>3,†</sup>, M. Bergevin <sup>2,‡</sup>, A. Bernstein,<sup>2</sup> N. S. Bowden <sup>2</sup>, S. Dazeley <sup>2,§</sup>, P. Jaffke <sup>4</sup>, I. Jovanovic <sup>1</sup>, S. Quillin <sup>5</sup>, C. Roecker,<sup>6,||</sup> and S. D. Rountree <sup>4</sup>

<sup>1</sup>*Department of Nuclear Engineering and Radiological Sciences, University of Michigan, Ann Arbor, Michigan 48109, USA*

<sup>2</sup>*Lawrence Livermore National Laboratory, Livermore, California 94550, USA*

<sup>3</sup>*Physics Department, University of California, Davis, California 95616, USA*

<sup>4</sup>*Center for Neutrino Physics, Virginia Tech, Blacksburg, Virginia 24061, USA*

<sup>5</sup>*Atomic Weapons Establishment, Aldermaston, Reading RG7 4PR, United Kingdom*

<sup>6</sup>*Department of Nuclear Engineering, University of California, Berkeley, California 94720, USA*



(Received 4 December 2019; revised 17 July 2020; accepted 18 August 2020; published 28 September 2020)

We present a measurement of the rate of correlated neutron captures in the WATCHBOY detector, deployed at a depth of approximately 390 meters water equivalent (m.w.e.) in the Kimballton Underground Research Facility (KURF). WATCHBOY consists of a cylindrical 2 ton water target doped with 0.1% gadolinium, surrounded by a 40 ton undoped water hermetic shield. We present a comparison of our results with the expected rate of correlated neutron captures arising from high-energy neutrons incident on the outside of the WATCHBOY shield, predicted by a hybrid FLUKA-GEANT4-based simulation. The incident neutron energy distribution used in the simulation was measured by a fast neutron spectrometer, the 1.8 ton Multiplicity and Recoil Spectrometer (MARS) detector, at the same depth. We find that the measured detection rate of two correlated neutrons is consistent with that predicted by simulation. The result lends additional confidence in the detection technique used by MARS, and therefore in the MARS spectra as measured at three different depths. Confirmation of the fast neutron flux and spectrum is important as it helps validate the scaling models used to predict the fast neutron fluxes at different overburdens.

DOI: [10.1103/PhysRevC.102.034616](https://doi.org/10.1103/PhysRevC.102.034616)

### I. INTRODUCTION

Muon-induced fast neutrons are an important background for many rare event measurements in underground laboratories. Fast neutrons were a possible source of the discrepancy between the CDMS [1] and the DAMA/LIBRA dark matter experiments [2–4]. High-energy neutrons were also shown to be the dominant background in the relatively shallow reactor neutrino oscillation experiment such as Double Chooz [5,6], Daya Bay [7], and RENO [8]. Despite their importance, it is difficult to unambiguously measure the neutron flux and energy distribution at depth, because the flux is low and generally rare event detectors are designed to moderate and shield themselves from the fast neutron flux rather than measure it.

In order to plan and design experiments at a range of depths, existing models used to scale high-energy neutron flux in rock must be better validated than current data allows

[9]. Motivating the present work, a kiloton scale gadolinium-doped light water Cherenkov detector WATCHMAN (WATER Cherenkov Monitor of ANtineutrinos) is planned in the near future in order to demonstrate remote monitoring of the antineutrino flux arising from the Hartlepool reactor in Northern England [10]. In such detectors, the antineutrino signal appears as two interactions (a positron followed by a neutron capture) correlated in time and position [11]. Backgrounds include accidental coincidences of uncorrelated single events, antineutrinos from other reactors, and correlated signals arising from muon induced radionuclides and fast neutrons. Two experiments, MARS [12] and WATCHBOY [13], were built to measure fast neutrons and long-lived radionuclides respectively, with the goal of improving the predictive capability for large detectors such as WATCHMAN.

The MARS detector was deployed at depths of 390, 540, and 1450 m.w.e. within the Kimballton mine near the Kimballton Underground Research Facility (KURF) in Virginia. The purpose of the deployment was to provide a set of neutron spectra measurements at different depths, measured with the same detector and systematic effects. MARS neutron spectra were published elsewhere, along with a detailed description of the detector [14]. Briefly, MARS was a transportable fast neutron detection system that consisted of two plastic scintillator-based neutron detectors mounted above and below a lead spallation target. The dimensions of the lead

\*fsutanto@umich.edu

†Present address: Physics Department, University of California, Berkeley, California 94720, USA.

‡bergevin1@llnl.gov

§dazeley2@llnl.gov

||Present address: Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA.

target were 101 cm long by 71 cm wide by 20 cm high (1.8 tons). The dimensions of both neutron detectors were similar: 100 cm  $\times$  75 cm  $\times$  25 cm. Each neutron detector consisted of twelve 2 cm thick plastic scintillator layers interleaved with gadolinium coated Mylar sheets. The detection method was as follows: High-energy neutrons initiate a spallation reaction in the lead target and generate secondary neutrons. The number of secondary neutrons produced is correlated with the energy of the initiating fast neutron. A significant fraction of these secondary neutrons down-scatter in the scintillator layer, thermalize and eventually capture on a gadolinium nucleus an average of  $18.7 \pm 3.0 \mu\text{s}$  later [12]. The de-excitation of the gadolinium nucleus upon capture generates several gamma rays with total energy of 7.9 and 8.5 MeV for  $^{157}\text{Gd}$  and  $^{155}\text{Gd}$ , respectively. These gamma rays interact and partially deposit their energy in the plastic scintillator.

The primary aim of WATCHBOY was to measure the rate of radionuclide production in water due to the passage of high energy cosmic ray muons through its 2 ton water target [13]. The target was surrounded by a hermetic 40 ton undoped water muon detector and neutron shield, the purpose of which was to stop high energy neutrons from reaching the target, a potential source of backgrounds for the radionuclide measurement. Nevertheless, a small fraction of the highest energy neutrons were able to penetrate the shield and make it to the target, producing correlated groups of neutron captures via spallation in the target volume. Since WATCHBOY was a large and immobile detector, it was deployed at a single depth, which was estimated to be 390 m.w.e., the same depth as one of the MARS deployments.

WATCHBOY was conceived as a prototype for WATCHMAN. The radionuclide considered to be of greatest importance to WATCHMAN was  $^9\text{Li}$  ( $\tau = 257 \text{ ms}$ ,  $Q$  value = 11.9 MeV), which decays via the simultaneous emission of a multi-MeV beta particle and a neutron, similar to the antineutrino inverse beta decay (IBD) reaction. While WATCHBOY was intended for measuring radionuclides, its dominant background was caused by energetic neutrons produced in the surrounding rock by muons passing near the detector. These neutrons can enter the WATCHBOY target undetected by the outer shield, and initiate spallation reactions and nuclear recoils in the detector [15] which lead to correlated neutron capture signals in the WATCHBOY target.

Also of great interest to WATCHMAN was the development of a simulation code capable of predicting the rate, via spallation, of free neutrons in water-based detectors at any depth. To address this need, this work has two aims: First, to measure the rate of correlated neutron capture candidates in the (Gd-water) WATCHBOY target, and second, to determine the accuracy of two candidate simulation codes (GEANT4 and FLUKA) at predicting the observed correlated event rates using the independent MARS measurement of the neutron energy spectrum as input.

The WATCHBOY detector and its data acquisition system (DAQ) are described in Sec. II. Data selection and analysis are described in Sec. III. The simulation model is presented in Sec. IV. Section V describes our measurements of the multiplicity of correlated neutron-like events. Finally, conclusions are presented in Sec. VI.

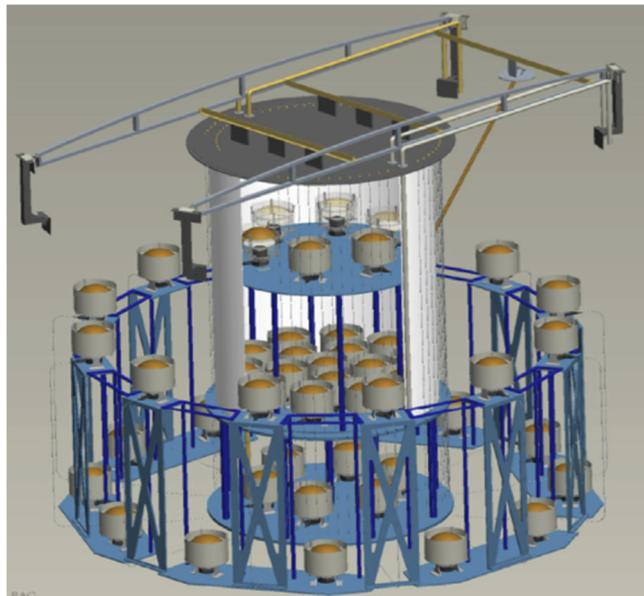


FIG. 1. The PMT arrangement and supporting structure inside the WATCHBOY detector [13]. The containment bag was filled with Gd-doped water and consisted of two optically separated regions—a lower 2-ton target region instrumented with 16 PMTs and an upper veto region instrumented with 6 PMTs. The containment bag was surrounded by 40-ton pure water veto volume containing 30 PMTs. WATCHBOY was deployed at a depth of  $\approx 390$  m.w.e. within KURF.

## II. WATCHBOY DETECTOR

Figure 1 shows a schematic of the detector components inside WATCHBOY, with the outer tank walls not displayed [13]. A cutaway view of the gadolinium doped inner containment bag is shown in the center. The containment bag was filled with Gd-doped water (0.2%  $\text{GdCl}_3$ ) and consisted of two optically separated volumes: A lower 2 ton “target” region, equipped with 16 tightly packed upward-facing 10 in. Hamamatsu R7081 photomultiplier tubes (PMTs), and an upper region which formed part of the muon veto. Gadolinium was chosen for its high neutron absorption cross section (49 000 b for natural Gd). Cherenkov light produced following neutron capture on Gd in the target volume was collected on the 16 upward-facing PMTs. The target walls were coated with a highly diffusely reflective Teflon to enhance light collection. Due to the reflectivity of the walls and the small number of PMTs, most of the detected light followed multiple reflections. The effect was to distribute ( $\approx 9 \text{ ns}$  time spread) the light fairly evenly among the target PMTs whenever Cherenkov light was produced by a physics event. As mentioned earlier, above the target and within the same containment bag was an optically separated top volume which formed part of the muon veto equipped with six upward-facing 10 in. Hamamatsu R7081 PMTs. Since it existed within the same containment bag, it was also filled with Gd-doped water. Surrounding the bag was a  $\approx 40$  ton pure water volume containing 30 additional 10 in. PMTs that also formed part of the muon veto.

Four 16-channel 250-MHz 14-bit SIS3316 Struck analog-to-digital converter (ADC) digitizer boards were used to

collect the PMT signals. PMT signals were summed into groups of four to create a trigger. Triggers were formed whenever any summed group reached a threshold of  $\approx 1$  photoelectron (PE). If any of the discriminators were triggered, the pulse integral of all 52 channels were then read out to disk.

### III. DATA SELECTION AND ANALYSIS

*Ad hoc* neutron selection criteria were developed and discussed in detail in the previous WATCHBOY paper [13]. Here we discuss only the pertinent highlights.

One of the most prevalent backgrounds in WATCHBOY was caused by the PMTs: Flashers and other noise triggers, which predominantly produced signals within only a single PMT. A straightforward way to remove these noise triggers was to define a parameter sensitive to the “evenness” of the distribution of light among the PMTs, called “charge balance” ( $Q_b$ ). This parameter was defined as follows:

$$Q_b = \sqrt{\frac{\sum Q_i^2}{Q_{\text{sum}}^2} - \frac{1}{N}}, \quad (1)$$

where  $N$  is the total number of target PMTs,  $Q_i$  is the charge collected by the  $i$ th PMT, and  $Q_{\text{sum}}$  is the summed charge of all 16 target PMTs. A charge balance value close to zero indicates evenly distributed signal among all 16 target PMTs, and close to one indicates the opposite extreme (most of the light detected was by a single PMT).

During calibrations, a  $^{252}\text{Cf}$  source was inserted through the yellow calibration tube, as shown in Fig. 1, and placed approximately 4 cm from the side wall of the inner detector. Figure 2 shows charge balance as a function of detected photoelectrons for the  $^{252}\text{Cf}$  calibration data. From the calibration runs, neutron capture candidates were found to preferentially have charge-balance values less than 0.6. The summed number of photoelectrons were also larger on average than for background data (Figs. 2 and 3). The background data consisted primarily of the PMT flasher-based events described earlier, and also events caused by low energy gamma rays from the PMT glass, which is known to contain potassium,

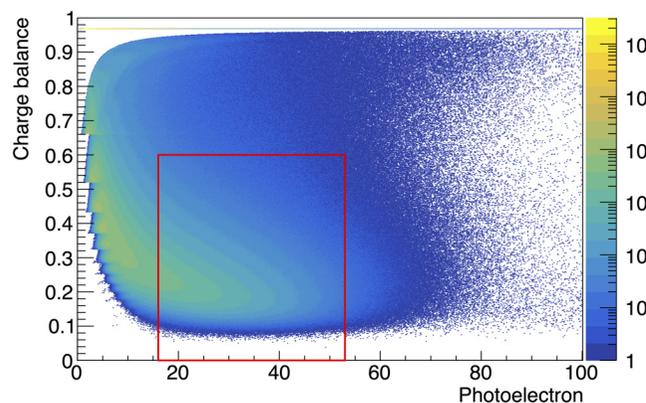


FIG. 2. The detector response to a  $^{252}\text{Cf}$  source in terms of the total detected photoelectrons and the evenness of the light distribution as defined in Eq. (1). The analysis cuts that select for neutron capture candidate [13] are shown in red.

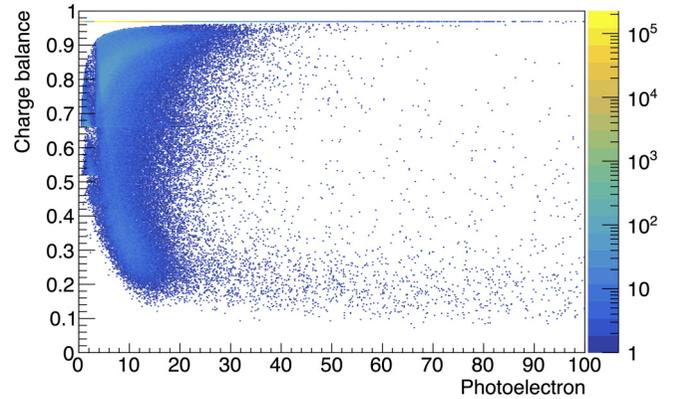


FIG. 3. The detector response to background in terms of the total detected photoelectrons and the evenness of the light distribution as defined in Eq. (1).

thorium, and uranium. Using the same neutron selection cuts as previously described in [13], we required the summed charge of the target PMTs to be in the range 16 to 53 photoelectrons. The upper limit helps to reject Michel electrons and muons that clip the target. To minimize background cosmogenic neutrons caused by muons traversing the detector, all events within 1 ms of a muon candidate event were rejected [13]. Muon events were defined as any event with summed charge in the veto greater than 30 photoelectrons or summed charge in the target region greater than 100 photoelectrons.

Using  $^{252}\text{Cf}$  data, we isolated a subset of events highly likely to be neutron captures by selecting events correlated with a previous event [16]. The technique takes advantage of the fact that multiple neutrons are generally emitted simultaneously in  $^{252}\text{Cf}$  spontaneous fission events (3.76 prompt neutrons/fission on average [17,18]). Figure 4 shows a two-dimensional picture of the interevent time distribution observed in  $^{252}\text{Cf}$  calibration data, as explained below:

- (i) Uncorrelated events appear in region A. These events have a long time interval from the prior event ( $\Delta t$

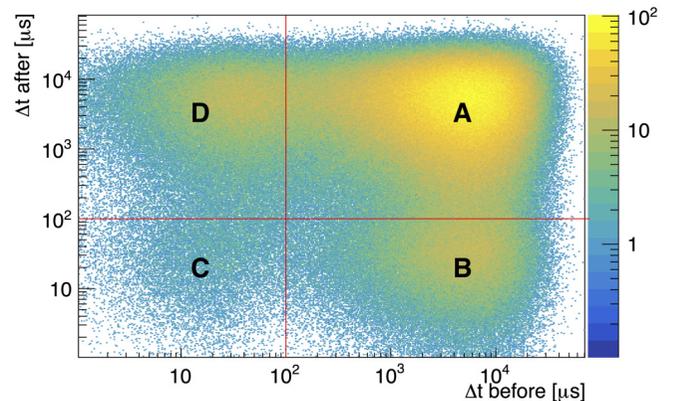


FIG. 4. The distribution of time intervals between detected events. The  $100 \mu\text{s}$  temporal cut is shown in red. Regions B, C, and D likely correspond to the correlated neutron events and region A corresponds to single events.

*before*) and a long time interval to the next event ( $\Delta t$  *after*).

- (ii) The first event of a correlated group appears in region B; these events have a long time interval from the prior event and a short time interval to the next event. Many of these are likely to be caused by either fission gamma rays or neutron captures resulting from a fission event.
- (iii) The last detected event from a correlated group appears in region D; with a short time interval from a prior event and a long time interval to the next event. This subset are preferentially populated by neutron captures following thermalization.
- (iv) Correlated events with a short time interval to both the previous and next events are in region C; this subset is also preferentially populated by neutron captures following thermalization.

Notice that there must be at least two events in a correlated group to occupy regions B and D. There must be at least three events in a correlated group to occupy region C. Region B corresponds to the first event, which could be either a prompt gamma ray event or a delayed neutron event. Region C and region D always correspond to delayed events that are likely due to thermalized neutrons capturing in the water volume. For a  $^{252}\text{Cf}$  calibration run, the rate of events observed in regions A, B, C, and D were 252, 100, 30, and 100 Hz, respectively. The accidental probability of two correlated events within 100  $\mu\text{s}$  time window due to the single events can be approximated as  $(252 \text{ Hz})^2 \times 100 \mu\text{s} = 6.4 \text{ Hz}$ . This implies that accidentals contribute to  $\approx 5\%$  of the total events in regions C and D. Hence, the events that appear in regions C and D are heavily dominated by neutron captures.

#### IV. GEANT4 SIMULATION

The GEANT4-based simulation framework RAT-PAC (Reactor Analysis Tool - Plus Additional Codes) was used to simulate neutron and gamma-ray interactions in the WATCHBOY detector and to determine the neutron capture efficiency. RAT-PAC was developed partly for KamLAND,<sup>1</sup> and partly for the Braidwood detector proposal [19], and has since been used and validated by the SNO+ Collaboration. For this work, GEANT4 version 4.10.3 with the shielding physics list and photon evaporation model were used.

In recent years, large discrepancies have been observed between GEANT4 and measured gamma-ray spectra and multiplicity distributions for neutron captures on gadolinium [20–22]. DICEBOX is an alternative gamma-ray cascade model that appears to model gamma cascades of excited nuclei more accurately [23–25]. Only the DICEBOX gamma cascade for neutron capture on  $^{157}\text{Gd}$  was incorporated into the RAT-PAC-based simulation for this work, since the high thermal neutron capture cross section (254 000 b) [22] dominates the other isotopes. The associated parameters used to define the resonance

structures of the  $^{158}\text{Gd}$  have been measured by the DANCE (Detector for Advanced Neutron Capture Experiments) Collaboration and are available in the literature [24].

The RAT-PAC-DICEBOX simulation was used to determine the WATCHBOY detector response for neutron capture throughout the target volume. For tuning purposes, the response was compared against the  $^{252}\text{Cf}$  calibration data taken in October 2013 and March 2014 (around the time period of physics data taking) and the neutrons resulting from muon spallation in the water target (during the time period of physics data taking). The latter was identified as events that pass the photoelectron and charge balance cuts and were within [20,200]  $\mu\text{s}$  since the last muon event [13]. The water absorption length and the target wall reflectivity in the simulation were tuned to maximize agreement with the summed photoelectron and charge balance distributions obtained from the  $^{252}\text{Cf}$  calibration data and the neutron data resulting from muon spallation in the water target. The values of the absorption length and reflectivity are negatively correlated, so constraints on these two values must be analyzed together. To select optimal values and uncertainties for each, the data and RAT-PAC-DICEBOX simulations were compared over a range of input reflectivity and water absorption length values using a  $\chi^2$  analysis (see Figs. 5 and 6), revealing bands of good fits extending from the top left down to the right in both the photoelectron and charge balance plots. An example comparison between the simulation and the calibration data is shown in Fig. 7, and an example comparison between the simulation and the neutron data from muon spallation in water target is shown in Fig. 8. In Fig. 7, wall reflectivity of 98% and water attenuation length of 28 m that yielded  $\chi^2/\text{ndf}$  value of 1.6 for the photoelectron response and 2.1 for the charge balance response were chosen. The poor  $\chi^2/\text{ndf}$  value for the charge balance response could potentially be explained by the target reflective coating that slightly bent, which introduced a disturbance in the light reflection and detection. There could also be variations in the PMT efficiency near the target wall that contributed to the variations in the PMT hit multiplicity. The multiplicity and energy distributions of the gamma rays produced by the deexcitation of  $^{156}\text{Gd}$  nuclei that were not corrected for with DICEBOX could also affect the charge balance response, especially in the case of neutron captures close to the target wall where gamma rays are more likely to escape. Additional subtle detector effects may also have contributed, such as PMT flashers events that somehow passed the charge balance cut, or other electronic noise. The effect of PMT dark rate on the data distributions was checked assuming a 10 kHz dark rate per PMT, which is relatively high, and found to be negligible. In Fig. 8, a wall reflectivity of 90% and a water attenuation length of 45 m that yielded  $\chi^2/\text{ndf}$  values of 1.4 and 1.2 for the photoelectron and charge balance responses were chosen, respectively. Figures 5 and 6 contains more than 700 comparisons between independent simulated data sets and the real data set containing  $6 \times 10^5$  and  $2 \times 10^3$  events, respectively.

The uncertainty in the neutron efficiency that results from the tuning process was determined by propagating the uncertainty in the wall reflectivity and absorption length through the analysis. To determine the uncertainties, we performed a series of one-dimensional (1D) scans of  $\chi^2$  values as a function

<sup>1</sup>RAT incorporates parts of the scintillation and PMT simulation from generic liquid scintillator GEANT4 simulation, written and maintained by Glenn Horton-Smith from the KamLAND Collaboration.

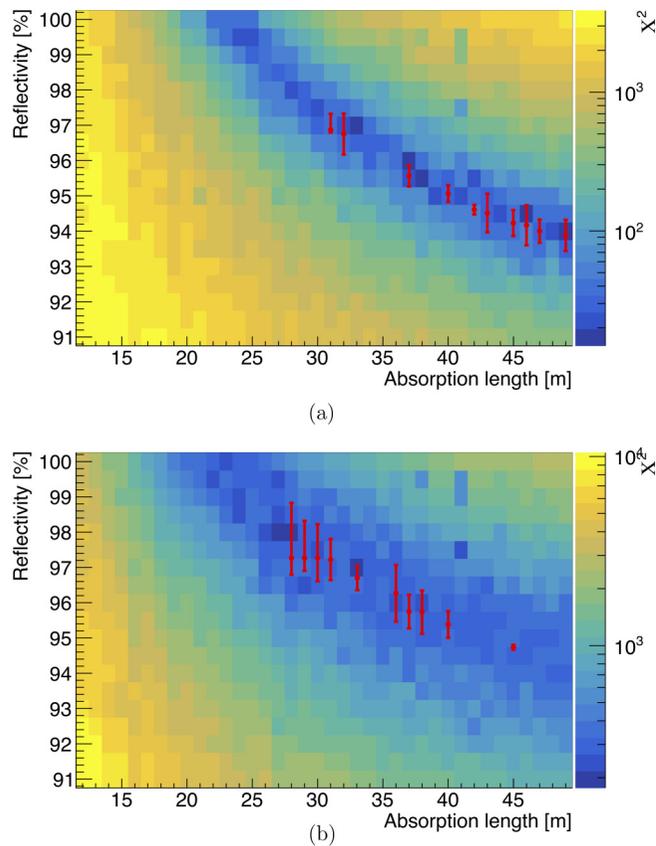


FIG. 5. The distribution of the goodness of fit ( $\chi^2/\text{ndf}$ ) is obtained by comparing (a) the simulated and  $^{252}\text{Cf}$  neutron capture candidate photoelectron distribution and (b) the similar comparison for the charge balance distributions. The lowest  $\chi^2$  values are 15 (26 DOF) and 177 (59 DOF) for the (a) photoelectron and (b) charge balance comparisons respectively.

of reflectivity for a series of absorption values spanning Figs. 5 and 6. Each was fitted with a second degree parabolic function to determine the  $\chi^2$  minimum. An example of a distribution of these  $\chi^2$  minima for Fig. 5 a is shown in Fig. 9. The degrees of freedom (DOF) of the PE response is 26 DOF and the expected distribution of  $\chi^2$  values for 26 DOF is shown in red (Fig. 9). While  $\chi^2$  minima within 68% confidence level (CL) for 26 DOF were observed (data points on the left side of the dashed red line on Fig. 9), we noticed that the distribution of  $\chi^2$  values for 33 DOF fitted our data best, indicating an imperfect match between simulation and data. The range of acceptable reflectivity and absorption length values were the ones that yielded a  $\chi^2$  minimum value within 68% CL for 33 DOF (data points on the left side of the dashed blue line on Fig. 9). An identical method was performed for Figs. 5(b), 6(a), and 6(b) to determine the acceptable range of wall reflectivities and absorption lengths. The accepted ranges are shown in red in Figs. 5 and 6. The detection efficiencies of neutrons uniformly distributed in the target volume were evaluated for the mean, upper bound, and lower bound of each data point (red bars in Figs. 5 and 6) to determine the range of uncertainties. The detection efficiency of neutrons uniformly

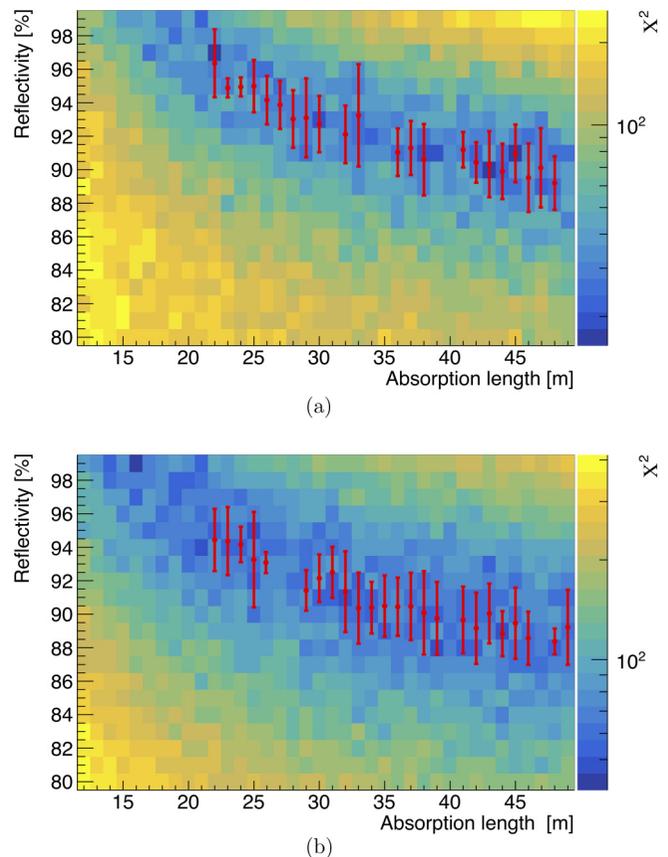


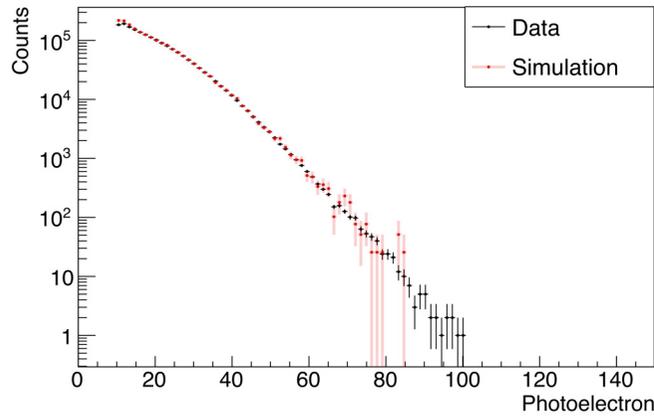
FIG. 6. The distribution of the goodness of fit ( $\chi^2/\text{ndf}$ ) is obtained by comparing (a) the simulated and muon-induced neutron capture candidate photoelectron distribution and (b) the similar comparison for the charge balance distributions. The lowest  $\chi^2$  values are 16 (26 DOF) and 40 (51 DOF) for the (a) photoelectron and (b) charge balance comparisons respectively.

distributed in the target volume was found to be  $30.3\% \pm 5.3\%$  (sys.).

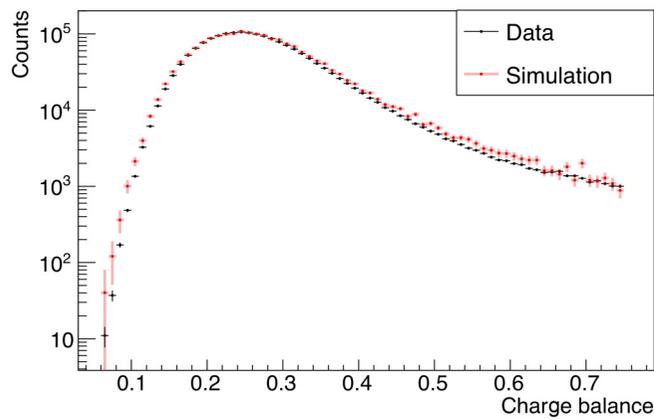
## V. MEASUREMENT OF THE DETECTION RATES OF CORRELATED NEUTRON EVENTS

In Sec. IV we described the RAT-PAC-DICEBOX-GEANT4 simulation used to simulate the response of WATCHBOY to neutron capture events in the target. This simulation was used to determine the neutron detection efficiency. Independent and separate simulations were also done to determine the rate of neutron captures inside the WATCHBOY target. The simulations were performed with the input neutron energy spectrum measurement of MARS, which was incident upon the outer WATCHBOY shield, which also acts as a muon detector. Two such independent simulations were performed. One was based on RAT-PAC-GEANT4, the other was based on FLUKA.

The MARS fast neutron spectrum measured at the same depth as the WATCHBOY detector extends from 90 to 400 MeV (Fig. 10, black). No information was available for how the neutron energies scale above 400 MeV at this depth. However we explored the effect of two different scaling models above 400 MeV: A double exponential function (red), chosen



(a)

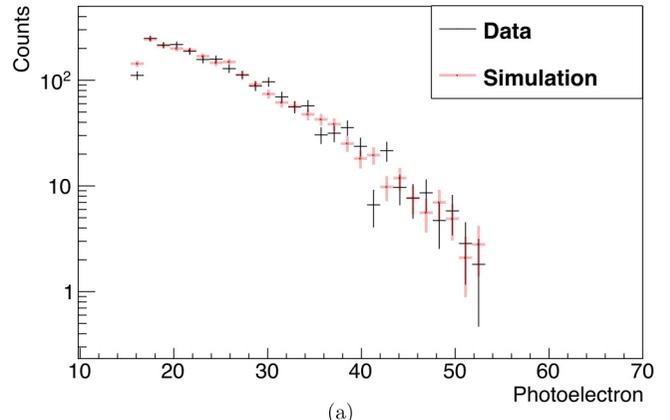


(b)

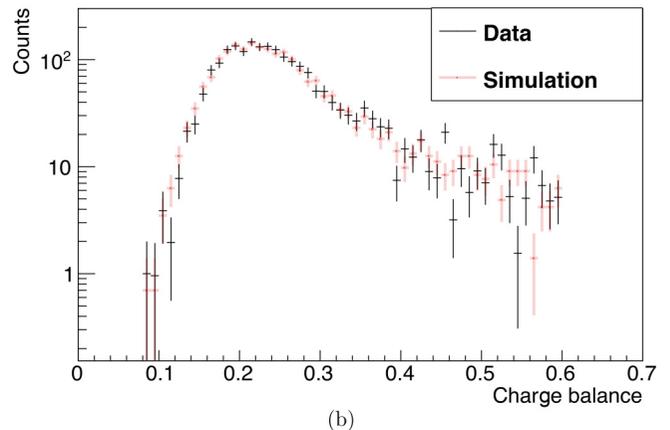
FIG. 7. Example of the comparison between (a) the simulated light collection and (b) the simulated charge balance responses to the WATCHBOY  $^{252}\text{Cf}$  calibration data.

to approximate the same shape at higher energies, and the Mei and Hime model [26] (blue). The two are shown in Fig. 10. The differences in the resulting neutron detection rates were evaluated and used to determine the systematic uncertainty due to the unmeasured neutron spectrum above 400 MeV. The two different scaling assumptions begin to diverge strongly above  $\approx 700$  MeV. We note that according to the simulation using the harder Mei and Hime energy spectrum, only 2% of the neutron captures in the WATCHBOY target result from those above 700 MeV, so the uncertainty in the shape above this energy contributes a relatively small amount to the total uncertainty.

The angular distribution of the neutron flux at  $390 (\pm 12)$  m.w.e. incident upon the WATCHBOY detector was unknown. However, it has been measured above ground and found to be consistent with  $\approx \cos^3 \theta$  [27]. Below ground, we expect the angular distribution to be less peaked along the muon direction than above ground: An energy-dependent combination of isotropic and peaked along the initiating muon direction [26]. The neutron initial directions were therefore simulated at  $\cos \theta^3$ ,  $\cos \theta^2$ , and  $\cos \theta$ , and the differences between the results of these simulations were used to estimate



(a)



(b)

FIG. 8. Example of the comparison between (a) the simulated light collection and (b) the simulated charge balance responses to the WATCHBOY neutron data from muon spallation in the water target.

our systematic uncertainty due to uncertainties in the angular distribution.

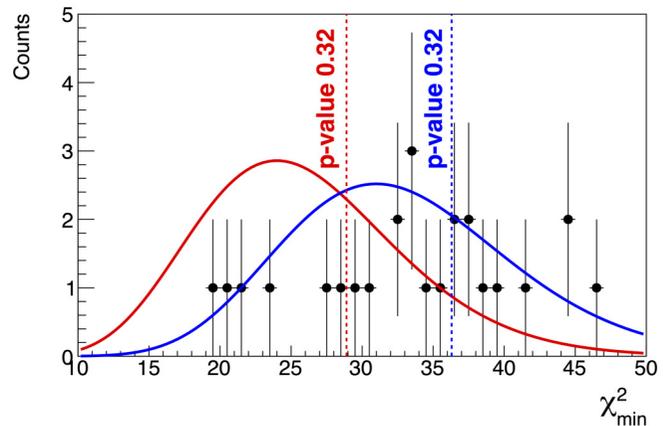


FIG. 9. Distribution of the minimum  $\chi^2$  values. The  $\chi^2$  distributions with 26 DOF (red) and 33 DOF (blue) are shown. In our analysis, we accepted data points that fall within one standard deviation of the mean of  $\chi^2$  distribution with 33 DOF (dashed blue).

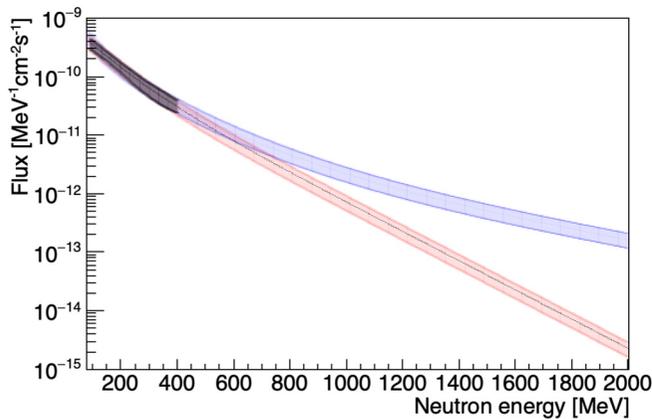


FIG. 10. The MARS neutron energy spectrum (black shaded region) propagated over a broad range of energies (from 20 MeV to 2 GeV), assuming two different scaling models: a double exponential (red), and Mei and Hime (blue). The flux uncertainty of the MARS data was propagated over the whole energy range for both models.

The RAT-PAC-GEANT4 and FLUKA-based distributions of neutron multiplicity in target volume (Fig. 11) were obtained from the above simulations. They were then converted to detected neutron multiplicity distributions (Fig. 12) using the following binomial equation:

$$M_i = \sum_{k=i}^{\infty} \binom{k}{i} \epsilon^i (1 - \epsilon)^{k-i} P_k, \quad (2)$$

where  $M_i$  is the probability that the detected neutron multiplicity equals  $i$ ,  $\epsilon$  is the thermal neutron detection efficiency, and  $P_k$  is the probability that multiplicity of secondary neutrons created per primary equals  $k$ .

The resulting RAT-PAC-GEANT4-based distribution of neutron multiplicity yielded a flatter distribution and higher average (2.3 neutrons) than the FLUKA-based distribution (1.5 neutrons). Likewise, the RAT-PAC-GEANT4-based detected neutron multiplicity distribution was also flatter and had a higher

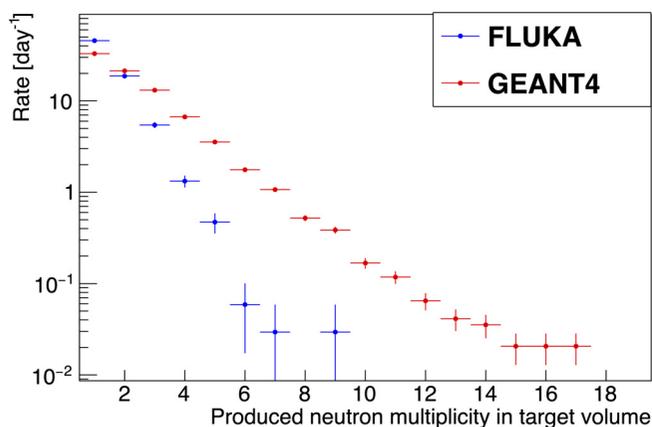


FIG. 11. The distributions of the rate of produced event clusters in the target volume obtained from the fast neutron simulations in FLUKA and GEANT4.

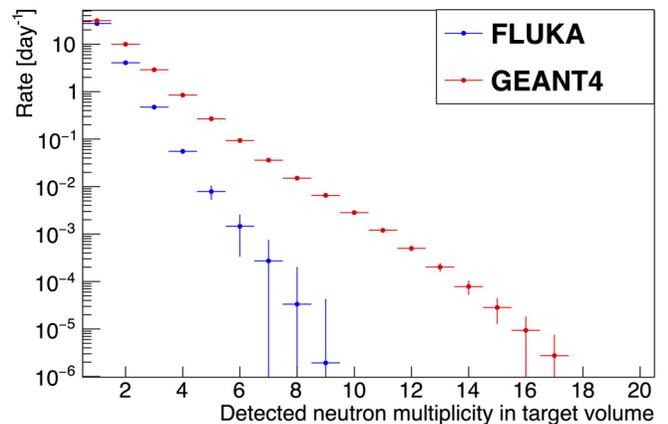


FIG. 12. The distributions of detected event clusters that pass the analysis cuts obtained from the fast neutron simulations in FLUKA and GEANT4. The GEANT4-based distribution has a higher average multiplicity in comparison with the FLUKA-based result.

average (1.45 neutrons) than the FLUKA-based one (1.17 neutrons).

To calculate the degree to which accidental coincidences of single events contribute to the rate of detected multiple coincidence event clusters, a simple toy Monte Carlo (toy MC) simulation was performed ( $10^8$  events). Note, in the following we will use the term “event cluster” to refer to a collection of correlated events. The toy MC provides a series of timestamps, each corresponding to either a detected single event cluster (which always has a multiplicity equal to 1) or a detected neutron event cluster. To perform this simulation, the detected single rate and the detected neutron multiplicity distribution, presented in Fig. 12, must be known. The detected single rate was approximated based on the observed events that dominated region A in Fig. 13. The toy MC was performed as follows:

- (i) The time to the next event cluster is sampled as  $dt = -\log(\xi)/R_{\text{total}}$ , where  $\xi$  is a random number from 0 to 1 and  $R_{\text{total}}$  is the sum of the detected single rate

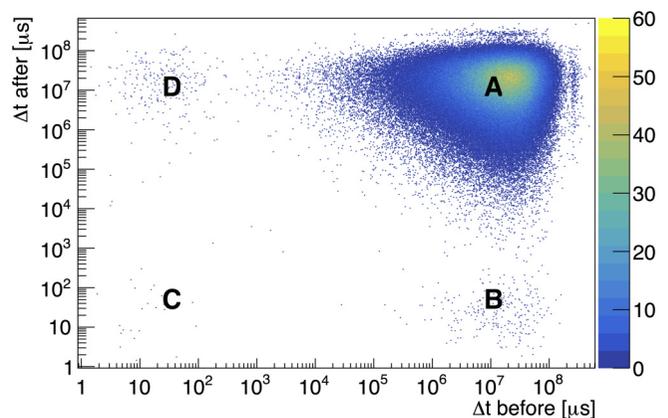


FIG. 13. The measured distribution of time intervals between detected physics events. Regions B, C, and D likely correspond to the correlated neutron events and region A likely corresponds to single events.

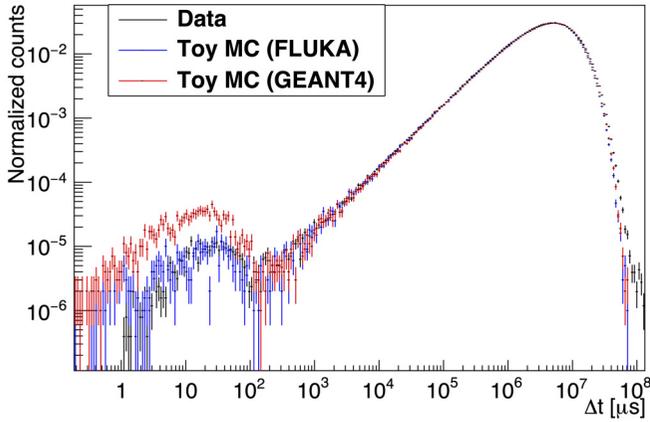


FIG. 14. The  $X$  projection of the distribution of time intervals between the detected physics events is compared against the simulated distribution of the interevent time obtained from a toy Monte Carlo simulation.

( $R_s$ ) and the detected neutron cluster rate. Note, the detected neutron cluster rate is the detected neutron rate divided by the average multiplicity of the detected neutron events ( $R_{mul}$ ).

- (ii) Based on  $R_s$  and  $R_{mul}$ , the type of the next event cluster is determined.
- (iii) If the event cluster type is a detected single event, we proceed to the next event cluster.
- (iv) If the event cluster type is a detected neutron event cluster, the multiplicity is sampled from the distribution shown in Fig. 12. The time of each detected neutron capture within each cluster is sampled as  $dt = -\log(\xi)/26\mu s$  relative to the initial time, where  $26\mu s$  is the experimentally determined average interevent time between neutron captures in regions B, C, and D from Figs. 4 and 13.

The toy MC-based distributions of interevent times could then be realized and compared with the data as shown in Fig. 14. The data interevent time distribution was generated from the  $X$  projection of Fig. 13. The  $\chi^2/\text{NDF}$  comparison between data and the two toy MC distributions between  $10^{-1}$  and  $10^7\mu s$  yielded values of 1.84 and 3.95 for FLUKA and GEANT4 respectively. Interevent times greater than  $10^7\mu s$  were excluded in the  $\chi^2$  analysis because of the appearance of artifacts in the data that were noted at long interevent times due to DAQ resets that occurred when the digitizer buffers were read out.

A  $100\mu s$  time cut (red line on Fig. 4) was applied to both the toy MC and data to obtain the neutron multiplicity distributions shown in Fig. 15. While good agreement was observed between the measured data and FLUKA-based distribution, poor agreement was observed for the GEANT4-based result. This suggests that the Bertini cascade model used in the standard shielding physics list in the RAT-PAC-GEANT4 simulation overestimated the neutron multiplicities in water-based media.

Table I shows the systematic uncertainties that result from the uncertainties in the simulation input parameters. The positive and negative systematic uncertainties for each simulation

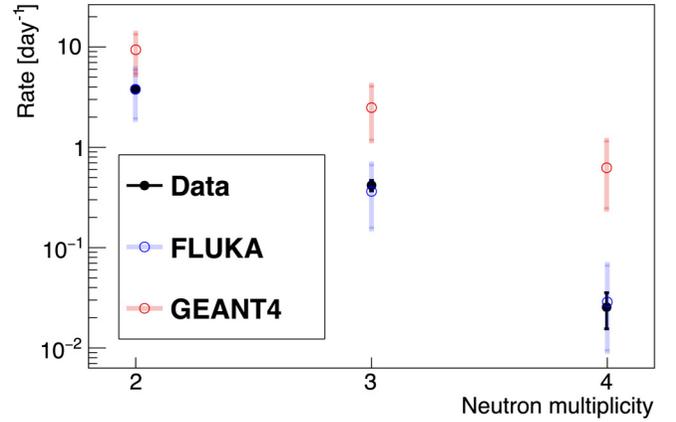


FIG. 15. The distribution of the simulated and measured neutron-like multiplicity after contamination from single events is incorporated. The GEANT4-based detection rate of correlated pairs was approximately 2 times higher than the FLUKA-based one. A good agreement between the measured and the FLUKA-based detected neutron-like event multiplicity rate was observed.

parameter were estimated by varying the simulation parameter of interest while keeping the other simulation parameters constant. For the fast neutron spectrum and the neutron efficiency parameters, the positive (negative) systematic uncertainty was the difference between the highest (lowest) detection rate and the median detection rate. For the cosine parameter, we cannot assume that the distribution is asymmetric and hence the positive (negative) systematic uncertainty was the difference between the highest (lowest) detection rate and the mean detection rate.

Other potential backgrounds considered were decays due to muon-induced radionuclides (e.g.,  ${}^9\text{Li}$  /  ${}^8\text{He}$ ) and antineutrinos from world reactors. Both of these can produce correlated neutron-like events. A measurement of the rate of muon-induced  ${}^9\text{Li}$  in WATCHBOY was attempted previously and an upper limit was published [13]. The published uncertainty in the dineutron-like event rate in WATCHBOY was

TABLE I. The relative systematic uncertainty associated with variations of the simulation parameters.

Simulation input	FLUKA	GEANT4
Dineutron rate		
Cosine	1.3%	1.2%
Fast neutron spectrum	+22.7% –30.4%	+21.7% –27.7%
Neutron efficiency	+31.7% –28.6%	+20.4% –21.6%
Trineutron rate		
Cosine	1.9%	1.1%
Fast neutron spectrum	+27.4% –36.3%	+22.4% –29.9%
Neutron efficiency	+53.0% –40.3%	+36.6% –32.9%
Quadneutron rate		
Cosine	3.4%	0.95%
Fast neutron spectrum	+38.0% –41.8%	+22.2% –32.9%
Neutron efficiency	+71.3% –47.7%	+54.7% –45.8%

negligible (2.7%) relative to the other uncertainties investigated here and was not included in our analysis. The nearest reactor (McGuire Nuclear Station,  $2 \times 3411$  MW, 216 km away) is too distant to produce a significant rate of antineutrino interactions in the 2 ton WATCHBOY target. According to [28], the rate of IBD events due to the integrated world reactor flux plus geo-antineutrinos in WATCHBOY is only 0.0014 events per day, which produces a negligible contribution to the correlated neutron signal.

## VI. CONCLUSIONS

The rate of correlated neutron-like event clusters in the WATCHBOY target was measured and found to be consistent with the expected rate from a FLUKA simulation of the MARS fast neutron spectrum incident upon the WATCHBOY outer water shield. The observed agreement provides an independent confirmation of the MARS fast neutron spectrum measured at the same depth as WATCHBOY (390 m.w.e.). Which in turn increases confidence in the MARS fast neutron spectra measured at other overburdens. A similar agreement was not found with a GEANT4-based simulation. Since MARS was designed to be transportable, and was deployed at multiple depths in the same configuration, the systematic bias between its measurements was minimized. Hence, a predictive model of the high-energy neutron energy-dependent flux can be obtained at any depth [12]. Confirmation of the MARS fast neutron spectrum places an important constraint on the limited number of depth scaling models that are used to extrapolate fast neutron flux between different overburdens and to validate the simulation framework that would be used to predict the rate of correlated neutrons in large underground water-based detectors planned for future construction, such as the WATCHMAN detector.

## ACKNOWLEDGMENTS

The authors would like to thank S. Valenta for providing the DICEBOX model, J. Xu for the helpful discussion on  $\chi^2$  analysis, and V. Kudryavtsev for helpful discussion on the fast neutron simulation. The authors also acknowledge the AIT-WATCHMAN Collaboration, M. Sweany, P. Marleau, B. Cabrera-Palmer, M. Gerling, K. Vetter, and T. M. Shokair for contributing to the WATCHBOY and MARS detector design, implementation, simulation, and characterization. Part of this

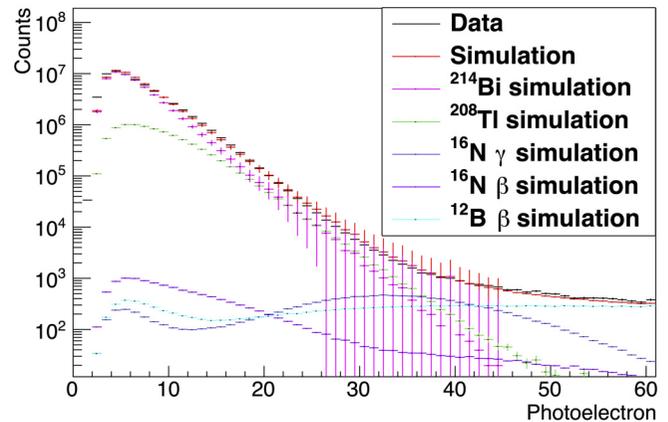


FIG. 16. The simulated detector response to  $^{214}\text{Bi}$  is scaled and compared against the experimental PE response of single events.

work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract No. DE-AC52-07NA27344, release number LLNL-JRNL-796592. The research of F.S. was supported by the Lawrence Livermore Graduate Scholar Program Fellowship and Rackham Merit Fellowship. The work of I.J. and F.S. was partially supported by the Consortium for Verification Technology under U.S. Department of Energy National Nuclear Security Administration Grant No. DE-NA0002534.

## APPENDIX: ESTIMATION OF RATE OF DETECTABLE EVENTS DUE TO PMT RADIOACTIVITY

We also validated our simulation toolkit by using it to predict the rate of detectable events due to PMT radioactivity and then compare the results with the observation.

The detection rate of events due to PMT radioactivity can be approximated by simulating interactions of uranium progeny such as  $^{214}\text{Bi}$  in the WATCHBOY detector, obtaining its simulated detection efficiency, and scaling its simulated PE response to match the experimental PE response of single events as shown in Fig. 16. The estimated rate is  $\approx 7.2$  Bq/PMT which is in decent agreement with the average rate for Hamamatsu R7081 PMT ( $\approx 5.9$  Bq/PMT for the uranium chain).

- [1] R. Abusaidi *et al.*, *Nucl. Instrum. Methods Phys. Res. A* **444**, 345 (2000).
- [2] R. Bernabei *et al.*, *Phys. Lett. B* **480**, 23 (2000).
- [3] J. Klinger and V. Kudryavtsev, *J. Phys.: Conf. Ser.* **718**, 042033 (2016).
- [4] P. S. Barbeau, J. I. Collar, Y. Efremenko, and K. Scholberg, *Phys. Rev. Lett.* **113**, 229001 (2014).
- [5] Y. Abe *et al.*, *J. High Energy Phys.* **10** (2011) 086.
- [6] Y. Abe *et al.*, *Phys. Lett. B* **735**, 51 (2014).
- [7] F. P. An *et al.*, *Phys. Rev. D* **97**, 052009 (2018).
- [8] H. Seo and for the RENO Collaboration, *J. Phys.: Conf. Ser.* **718**, 062053 (2016).
- [9] A. Bernstein *et al.*, *Sci. Global Secur.* **18**, 127 (2010).
- [10] M. Askin *et al.*, [arXiv:1502.01132](https://arxiv.org/abs/1502.01132).
- [11] C. Giunti and C. W. Kim, *Fundamentals of Neutrino Physics and Astrophysics* (Oxford University Press, Oxford, 2007).
- [12] C. Roecker *et al.*, *Nucl. Instrum. Methods Phys. Res. A* **826**, 21 (2016).
- [13] S. Dazeley *et al.*, *Nucl. Instrum. Methods Phys. Res. A* **821**, 151 (2016).
- [14] C. Roecker, Measurement of the high-energy neutron flux above and below ground, Ph.D. thesis, University of California, Berkeley, 2016, <https://escholarship.org/uc/item/8331v2xh>.
- [15] C. Galbiati and J. F. Beacom, *Phys. Rev. C* **72**, 025807 (2005).

- [16] H. Labranche, Time series analysis for the Californium source in sudbury neutrino observatory, Master's thesis, University of Guelph, 2004, <https://inspirehep.net/literature/899722>.
- [17] J. P. Lestone, Los Alamos National Laboratory Report No. LA-UR-05-0288, 2005 (unpublished), <https://mcnp.lanl.gov/pdf/les/la-ur-05-0288.pdf>.
- [18] P. Santi and M. Miller, *Nucl. Sci. Eng.* **160**, 190 (2008).
- [19] T. Bolton, *Nucl. Phys. B, Proc. Suppl.* **149**, 166 (2005).
- [20] T. Yano, *Nucl. Instrum. Methods Phys. Res. A* **845**, 425 (2017).
- [21] Y. Chen, in Conference on Science at the Sanford Underground Research Facility, AARM Workshop 2015 (unpublished), [https://zzz.physics.umn.edu/lowrad/\\_media/meeting8/ychen\\_gdgammas\\_aarm2015.pdf](https://zzz.physics.umn.edu/lowrad/_media/meeting8/ychen_gdgammas_aarm2015.pdf).
- [22] K. Hagiwara *et al.*, *Prog. Theor. Exp. Phys.* **2019**, 023D01 (2019).
- [23] F. Becvar, *Nucl. Instrum. Methods Phys. Res. A* **417**, 434 (1998).
- [24] A. Chyzh *et al.*, *Phys. Rev. C* **84**, 014306 (2011).
- [25] J. Kroll *et al.*, *Phys. Scr.* **2013**, 014009 (2013).
- [26] D. M. Mei and A. Hime, *Phys. Rev. D* **73**, 053004 (2006).
- [27] J. Ziegler, *IBM J. Res. Dev.* **40**(1), 19 (1996).
- [28] A. Barna and S. Dye, [arXiv:1510.05633](https://arxiv.org/abs/1510.05633).