

# Improving neutrino energy estimation of charged-current interaction events with recurrent neural networks in MicroBooNE

P. Abratenko,<sup>38</sup> O. Alterkait,<sup>38</sup> D. Andrade Aldana,<sup>15</sup> L. Arellano,<sup>21</sup> J. Asaadi,<sup>37</sup> A. Ashkenazi,<sup>35</sup> S. Balasubramanian,<sup>12</sup> B. Baller,<sup>12</sup> A. Barnard,<sup>28</sup> G. Barr,<sup>28</sup> D. Barrow,<sup>28</sup> J. Barrow,<sup>25</sup> V. Basque,<sup>12</sup> J. Bateman,<sup>21</sup> O. Benevides Rodrigues,<sup>15</sup> S. Berkman,<sup>24</sup> A. Bhanderi,<sup>21</sup> A. Bhat,<sup>7</sup> M. Bhattacharya,<sup>12</sup> M. Bishai,<sup>3</sup> A. Blake,<sup>18</sup> B. Bogart,<sup>23</sup> T. Bolton,<sup>17</sup> J. Y. Book,<sup>14</sup> M. B. Brunetti,<sup>41</sup> L. Camilleri,<sup>10</sup> Y. Cao,<sup>21</sup> D. Caratelli,<sup>4</sup> F. Cavanna,<sup>12</sup> G. Cerati,<sup>12</sup> A. Chappell,<sup>41</sup> Y. Chen,<sup>31</sup> J. M. Conrad,<sup>22</sup> M. Convery,<sup>31</sup> L. Cooper-Troeindle,<sup>29</sup> J. I. Crespo-Anadón,<sup>6</sup> R. Cross,<sup>41</sup> M. Del Tutto,<sup>12</sup> S. R. Dennis,<sup>5</sup> P. Detje,<sup>5</sup> R. Diurba,<sup>2</sup> Z. Djurcic,<sup>1</sup> R. Dorrill,<sup>15</sup> K. Duffy,<sup>28</sup> S. Dytman,<sup>29</sup> B. Eberly,<sup>33</sup> P. Englezos,<sup>30</sup> A. Ereditato,<sup>7,12</sup> J. J. Evans,<sup>21</sup> R. Fine,<sup>19</sup> W. Foreman,<sup>15</sup> B. T. Fleming,<sup>7</sup> D. Franco,<sup>7</sup> A. P. Furmanski,<sup>25</sup> F. Gao,<sup>4</sup> D. Garcia-Gamez,<sup>13</sup> S. Gardiner,<sup>12</sup> G. Ge,<sup>10</sup> S. Gollapinni,<sup>19</sup> E. Gramellini,<sup>21</sup> P. Green,<sup>28</sup> H. Greenlee,<sup>12</sup> L. Gu,<sup>18</sup> W. Gu,<sup>3</sup> R. Guenette,<sup>21</sup> P. Guzowski,<sup>21</sup> L. Hagaman,<sup>7</sup> O. Hen,<sup>22</sup> C. Hilgenberg,<sup>25</sup> G. A. Horton-Smith,<sup>17</sup> Z. Imani,<sup>38</sup> B. Irwin,<sup>25</sup> M. S. Ismail,<sup>29</sup> C. James,<sup>12</sup> X. Ji,<sup>26</sup> J. H. Jo,<sup>3</sup> R. A. Johnson,<sup>8</sup> Y.-J. Jwa,<sup>10</sup> D. Kalra,<sup>10</sup> N. Kamp,<sup>22</sup> G. Karagiorgi,<sup>10</sup> W. Ketchum,<sup>12</sup> M. Kirby,<sup>3,12</sup> T. Kobilarcik,<sup>12</sup> I. Kreslo,<sup>2</sup> N. Lane,<sup>21</sup> I. Lepetic,<sup>30</sup> J.-Y. Li,<sup>11</sup> Y. Li,<sup>3</sup> K. Lin,<sup>30</sup> B. R. Littlejohn,<sup>15</sup> H. Liu,<sup>3</sup> W. C. Louis,<sup>19</sup> X. Luo,<sup>4</sup> C. Mariani,<sup>40</sup> D. Marsden,<sup>21</sup> J. Marshall,<sup>41</sup> N. Martinez,<sup>17</sup> D. A. Martinez Caicedo,<sup>32</sup> S. Martynenko,<sup>3</sup> A. Mastbaum,<sup>30</sup> I. Mawby,<sup>18</sup> N. McConkey,<sup>39</sup> V. Meddage,<sup>17</sup> J. Mendez,<sup>20</sup> J. Micallef,<sup>22,38</sup> K. Miller,<sup>7</sup> A. Mogan,<sup>9</sup> T. Mohayai,<sup>16</sup> M. Mooney,<sup>9</sup> A. F. Moor,<sup>5</sup> C. D. Moore,<sup>12</sup> L. Mora Lepin,<sup>21</sup> M. M. Moudgalya,<sup>21</sup> S. Mulleriababu,<sup>2</sup> D. Naples,<sup>29</sup> A. Navrer-Agasson,<sup>21</sup> N. Nayak,<sup>3</sup> M. Nebot-Guinot,<sup>11</sup> J. Nowak,<sup>18</sup> N. Oza,<sup>10</sup> O. Palamara,<sup>12</sup> N. Pallat,<sup>25</sup> V. Paolone,<sup>29</sup> A. Papadopoulou,<sup>1</sup> V. Papavassiliou,<sup>27</sup> H. B. Parkinson,<sup>11</sup> S. F. Pate,<sup>27</sup> N. Patel,<sup>18</sup> Z. Pavlovic,<sup>12</sup> E. Piasetzky,<sup>35</sup> K. Pletcher,<sup>24</sup> I. Pophale,<sup>18</sup> X. Qian,<sup>3</sup> J. L. Raaf,<sup>12</sup> V. Radeka,<sup>3</sup> A. Rafique,<sup>1</sup> M. Reggiani-Guzzo,<sup>11,21</sup> L. Ren,<sup>27</sup> L. Rochester,<sup>31</sup> J. Rodriguez Rondon,<sup>32</sup> M. Rosenberg,<sup>38</sup> M. Ross-Lonergan,<sup>19</sup> I. Safa,<sup>10</sup> G. Scanavini,<sup>42</sup> D. W. Schmitz,<sup>7</sup> A. Schukraft,<sup>12</sup> W. Seligman,<sup>10</sup> M. H. Shaevitz,<sup>10</sup> R. Sharankova,<sup>12</sup> J. Shi,<sup>5</sup> E. L. Snider,<sup>12</sup> M. Soderberg,<sup>34</sup> S. Söldner-Rembold,<sup>21</sup> J. Spitz,<sup>23</sup> M. Stancari,<sup>12</sup> J. St. John,<sup>12</sup> T. Strauss,<sup>12</sup> A. M. Szelc,<sup>11</sup> W. Tang,<sup>36</sup> N. Taniuchi,<sup>5</sup> K. Terao,<sup>31</sup> C. Thorpe,<sup>21</sup> D. Torbunov,<sup>3</sup> D. Totani,<sup>4</sup> M. Toups,<sup>12</sup> A. Trettin,<sup>21</sup> Y.-T. Tsai,<sup>31</sup> J. Tyler,<sup>17</sup> M. A. Uchida,<sup>5</sup> T. Usher,<sup>31</sup> B. Viren,<sup>3</sup> M. Weber,<sup>2</sup> H. Wei,<sup>20</sup> A. J. White,<sup>7</sup> S. Wolbers,<sup>12</sup> T. Wongjirad,<sup>38</sup> M. Wospakrik,<sup>12</sup> K. Wresilo,<sup>5</sup> W. Wu,<sup>29</sup> E. Yandel,<sup>4</sup> T. Yang,<sup>12</sup> L. E. Yates,<sup>12</sup> H. W. Yu,<sup>3</sup> G. P. Zeller,<sup>12</sup> J. Zennamo,<sup>12</sup> and C. Zhang<sup>3</sup>

(MicroBooNE Collaboration)\*

<sup>1</sup>Argonne National Laboratory (ANL), Lemont, Illinois, 60439, USA

<sup>2</sup>Universität Bern, Bern CH-3012, Switzerland

<sup>3</sup>Brookhaven National Laboratory (BNL), Upton, New York, 11973, USA

<sup>4</sup>University of California, Santa Barbara, California, 93106, USA

<sup>5</sup>University of Cambridge, Cambridge CB3 0HE, United Kingdom

<sup>6</sup>Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid E-28040, Spain

<sup>7</sup>University of Chicago, Chicago, Illinois, 60637, USA

<sup>8</sup>University of Cincinnati, Cincinnati, Ohio, 45221, USA

<sup>9</sup>Colorado State University, Fort Collins, Colorado, 80523, USA

<sup>10</sup>Columbia University, New York, New York, 10027, USA

<sup>11</sup>University of Edinburgh, Edinburgh EH9 3FD, United Kingdom

<sup>12</sup>Fermi National Accelerator Laboratory (FNAL), Batavia, Illinois 60510, USA

<sup>13</sup>Universidad de Granada, Granada E-18071, Spain

<sup>14</sup>Harvard University, Cambridge, Massachusetts 02138, USA

<sup>15</sup>Illinois Institute of Technology (IIT), Chicago, Illinois 60616, USA

<sup>16</sup>Indiana University, Bloomington, Indiana 47405, USA

<sup>17</sup>Kansas State University (KSU), Manhattan, Kansas, 66506, USA

<sup>18</sup>Lancaster University, Lancaster LA1 4YW, United Kingdom

<sup>19</sup>Los Alamos National Laboratory (LANL), Los Alamos, New Mexico, 87545, USA

<sup>20</sup>Louisiana State University, Baton Rouge, Louisiana, 70803, USA

<sup>21</sup>The University of Manchester, Manchester M13 9PL, United Kingdom

<sup>22</sup>Massachusetts Institute of Technology (MIT), Cambridge, Massachusetts, 02139, USA

<sup>23</sup>University of Michigan, Ann Arbor, Michigan, 48109, USA

<sup>24</sup>Michigan State University, East Lansing, Michigan 48824, USA

<sup>25</sup>University of Minnesota, Minneapolis, Minnesota, 55455, USA

<sup>26</sup>*Nankai University, Nankai District, Tianjin 300071, China*  
<sup>27</sup>*New Mexico State University (NMSU), Las Cruces, New Mexico, 88003, USA*  
<sup>28</sup>*University of Oxford, Oxford OX1 3RH, United Kingdom*  
<sup>29</sup>*University of Pittsburgh, Pittsburgh, Pennsylvania, 15260, USA*  
<sup>30</sup>*Rutgers University, Piscataway, New Jersey, 08854, USA*  
<sup>31</sup>*SLAC National Accelerator Laboratory, Menlo Park, California, 94025, USA*  
<sup>32</sup>*South Dakota School of Mines and Technology (SDSMT), Rapid City, South Dakota, 57701, USA*  
<sup>33</sup>*University of Southern Maine, Portland, Maine, 04104, USA*  
<sup>34</sup>*Syracuse University, Syracuse, New York, 13244, USA*  
<sup>35</sup>*Tel Aviv University, Tel Aviv, Israel, 69978*  
<sup>36</sup>*University of Tennessee, Knoxville, Tennessee, 37996, USA*  
<sup>37</sup>*University of Texas, Arlington, Texas, 76019, USA*  
<sup>38</sup>*Tufts University, Medford, Massachusetts, 02155, USA*  
<sup>39</sup>*University College London, London WC1E 6BT, United Kingdom*  
<sup>40</sup>*Center for Neutrino Physics, Virginia Tech, Blacksburg, Virginia, 24061, USA*  
<sup>41</sup>*University of Warwick, Coventry CV4 7AL, United Kingdom*  
<sup>42</sup>*Wright Laboratory, Department of Physics, Yale University, New Haven, Connecticut, 06520, USA*



(Received 25 June 2024; accepted 22 September 2024; published 14 November 2024)

We present a deep learning-based method for estimating the neutrino energy of charged-current neutrino-argon interactions. We employ a recurrent neural network (RNN) architecture for neutrino energy estimation in the MicroBooNE experiment, utilizing liquid argon time projection chamber (LArTPC) detector technology. Traditional energy estimation approaches in LArTPCs, which largely rely on reconstructing and summing visible energies, often experience sizable biases and resolution smearing because of the complex nature of neutrino interactions and the detector response. The estimation of neutrino energy can be improved after considering the kinematics information of reconstructed final-state particles. Utilizing kinematic information of reconstructed particles, the deep learning-based approach shows improved resolution and reduced bias for the muon neutrino Monte Carlo simulation sample compared to the traditional approach. In order to address the common concern about the effectiveness of this method on experimental data, the RNN-based energy estimator is further examined and validated with dedicated data-simulation consistency tests using MicroBooNE data. We also assess its potential impact on a neutrino oscillation study after accounting for all statistical and systematic uncertainties and show that it enhances physics sensitivity. This method has good potential to improve the performance of other physics analyses.

DOI: [10.1103/PhysRevD.110.092010](https://doi.org/10.1103/PhysRevD.110.092010)

## I. INTRODUCTION

Neutrino oscillations, which refer to transitions between different neutrino flavor states along their propagation length, are of strong scientific interest as they unequivocally prove the existence of neutrino mass, which was not predicted by the original Standard Model (SM). The precise nature of these oscillations is not yet known and is a focus of many experiments [1]. Some key unanswered questions include the order of neutrino masses, the octant of the mixing angle  $\theta_{23}$ , and the presence of leptonic  $CP$  violations. The latter question is especially relevant to early Universe phenomena such as leptogenesis which has been proposed to

explain the observed baryon asymmetry of our Universe [2]. In addition, neutrino oscillations serve as a probe for physics beyond the SM such as the hypothetical sterile neutrino states within well-motivated extensions of the SM [3].

The estimation of neutrino energy  $E_\nu$  is of crucial importance to experiments studying the phenomenon of neutrino oscillation, since the transition probabilities depend on neutrino energy, in particular on  $L/E_\nu$ , where  $L$  is the distance the neutrino has traveled. These experiments typically involve the scattering of neutrinos with a broad distribution of energy (i.e., they are not monochromatic) on a fixed nuclear target, which also serves as a detector. Examples in accelerator experiments can be found in Refs. [4–7] among others. As a result, the neutrino energy is not known *a priori* on a per-interaction basis and must be reconstructed from the interaction itself.

This work focuses on the energy estimation of charged-current (CC) neutrino events since these events are of most interest for oscillation analyses. The  $\nu$ -Ar CC neutrino interactions are associated with the weak CC neutrino

\*Contact author: [microboone\\_info@fnal.gov](mailto:microboone_info@fnal.gov)

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP<sup>3</sup>.

interaction vertex  $\nu_l \rightarrow l^- + W^+$  (and  $\bar{\nu}_l \rightarrow l^+ + W^-$ ) and are characterized by the presence of the primary outgoing lepton  $l^\pm$ . The energy of the incoming neutrino is transferred to the primary lepton  $l^\pm$  and the argon nucleus (through  $W^\pm$  bosons). The latter part of the energy can, in turn, create a number of secondary hadronic particles. By measuring the energy of the primary lepton and the energies of the resulting hadronic particles one can make inferences about the energy of the incoming neutrino.

For simple CC interactions where the scattering is elastic and on quasifree nucleons within the nucleus (“charged-current quasi-elastic” or CCQE), the neutrino energy can be reconstructed from the kinematics of the outgoing charged lepton that is associated with the neutrino, i.e.,  $\mu$  ( $e$ ) for  $\nu_\mu$  ( $\nu_e$ ) [8]. For the CCQE interaction  $\nu_\mu + n \rightarrow \mu^- + p$ , assuming two-body scattering under the energy and momentum conservation laws, we have

$$E_\nu^{\text{QE}} = \frac{m_p^2 - (m_n - E_b)^2 - m_\mu^2 + 2(m_p - E_b)E_\mu}{2(m_n - E_b - E_\mu + p_\mu \cos \theta_\mu)}, \quad (1)$$

where  $m_p$ ,  $m_n$ , and  $m_\mu$  refer to the mass of the proton, neutron, and outgoing muon, respectively.  $E_\mu$  is the muon energy.  $p_\mu$  and  $\theta_\mu$  are the muon momentum and angle with respect to the incoming neutrino momentum direction, and  $E_b$  is the binding energy of the proton within the nucleus, typically  $\sim \mathcal{O}(10)$  MeV for various nuclear targets [9].

However, a significant fraction of neutrino interactions are not quasielastic. The nucleons within the nucleus are generally not quasifree, and the interaction can, therefore, exhibit a strong dependence on the initial nuclear state [10]. In addition, the incoming neutrino can excite baryon resonances leading to a variety of hadronic final states, typically pions and nucleons [11]. Moreover, the mechanism of intranuclear transport of these final-state hadrons is a research topic itself [12] and introduces an additional layer of uncertainty to the presence as well as the kinematics of final-state hadronic particles. These complications subsequently introduce limitations to the use of Eq. (1).

A separate strategy to reconstruct  $E_\nu$  based on calorimetry is commonly adopted for tracking calorimeters, such as that used in the MicroBooNE experiment [13]. The MicroBooNE detector is a liquid argon time projection chamber (LArTPC), which provides mm-scale position resolution, ns-scale timing resolution [14], and sub-MeV energy threshold to resolve neutrino interactions in fine detail. With the reconstructed information of particles (e.g., type and four momentum) in the final state including both the primary and secondary interactions, the energy of the neutrino is estimated by summing up the estimated leptonic, hadronic, and nucleon binding energies [15] based on energy conservation only. This strategy does not limit itself to the CCQE interactions and is more general. However, simple summation methods cannot account for energy lost to undetected particles (e.g., particles below threshold, outgoing neutrinos, or neutrons),

which leads to missing energy. Additionally, binding energy corrections can only be performed on average. Finally, the kinematics of individual particles may be misestimated for a variety of reasons, including biases in the reconstruction algorithms, exiting particles, and reinteractions of charged hadrons. To better estimate neutrino energies, an algorithm needs to be capable of inferring missing energy from the kinematics and topology of the reconstructed particles of individual events.

In this paper, we present a deep learning (DL) approach to estimate neutrino energy by utilizing both energy and momentum information for general neutrino-nucleus interactions at the GeV energy scale. In addition to estimating neutrino energy, we also estimate the lepton energy, a crucial parameter for accurately describing neutrino-nucleus interactions and conducting model validation tests. Deep learning refers to a class of modern machine learning (ML) techniques that perform deep inference by automatically deriving important representations of the input feature set in a high-dimensional space for various tasks such as classification and regression [16]. This is in contrast to more traditional ML-based approaches that rely on significant human inputs for the feature set. Deep learning techniques have been shown to improve performance on a wide array of targeted metrics within high-energy physics [17], including neutrino physics which is often an early adopter on this front [18]. Within neutrino experiments, they have contributed to detector simulation and signal processing [19], particle identification [20,21], estimation of the interaction vertex [22], and reconstruction of energies [23] and directions [24], among other tasks.

The DL approach is especially well suited to energy estimation because of the many multidimensional inputs, which include the various outgoing particles in the interaction and their particle-flow information, as well as the ability to model nonlinear relationships among those inputs in the high-dimensional space. Through training on the simulated event samples, the DL-based approach learns to estimate the neutrino energy considering correlations of the kinematics of the final-state particles embedded in event generators, which are constrained by both the energy and momentum conservation laws. In this work, we employ recurrent neural networks (RNNs), which have found use in various contexts such as natural language processing [25], and are especially suitable for a varying sequence of inputs like the particle flow of a neutrino interaction.

The DL-based neutrino energy estimator is aiming at a better neutrino energy estimation which can benefit future physics analyses. One example is neutrino oscillation measurement [6]. The sensitivity of the neutrino oscillation analysis depends on the quality of reconstructed neutrino energy spectra and on the ability to resolve various features of these spectra. To evaluate the performance of the DL-based neutrino energy estimator, we consider two key metrics related to the quality of the reconstructed

energy spectra. These are the bias and resolution, which refer to the mean and root-mean-square (RMS) of the ratio  $\frac{(E^{\text{reco}} - E^{\text{true}})}{E^{\text{true}}}$  respectively. A large energy resolution (RMS) leads to a smearing of features in the oscillated neutrino energy spectrum, reducing sensitivity in measuring neutrino oscillations. Similarly, a large bias in the energy estimation can skew the oscillation measurement, resulting in a faulty estimate of the oscillation parameters if the bias is not modeled correctly by the simulation. This can be checked by the model validation procedure shown in Sec. V B. Even when the bias in the energy estimation is properly modeled, it can again lead to a reduction of features in the oscillated neutrino energy spectrum, thus decreasing sensitivity.

While the DL-based energy estimator can be shown to outperform traditional energy estimators evaluated with simulation samples, there can be a significant model dependence in mapping from true to reconstructed neutrino energies, given the simulation's task of modeling the complex nuclear physics (i.e., kinematics correlations among final-state particles). While such model dependence may not be apparent in evaluations using simulations, it becomes crucial when applying the DL-based energy estimator to experimental data. In order to mitigate this concern, we perform dedicated model validations with the experimental data from the MicroBooNE experiment to demonstrate that the bias and resolution in the DL-based energy estimator are compatible within the quoted uncertainties of the overall model. These validations rely on the goodness-of-fit metric [15], which is further enhanced by the conditional constraint methodology [26]. These dedicated model validations build confidence not only in our overall simulation but also in the DL model, which often suffers from a lack of interpretability because of its black-box nature and sometimes nonintuitive associations among its inputs.

This paper is organized as follows. In Sec. II, we review various techniques used for neutrino energy estimation and outline the advantages of using the DL-based approach. In Sec. III, we introduce the RNN-based DL energy estimator (DL-EE), the preparation of input particle flow information, and the DL model architecture. In Sec. IV, we outline the training of the DL-EE on MicroBooNE simulation, including methods to control the resulting output bias of the energy estimator. In Sec. V, we evaluate the performance of the DL-EE using simulation and show that this technique is able to improve both the resolution and bias of the neutrino energy estimation. In addition, we perform validation of the DL-based energy estimator using experimental data. Furthermore, we study the impact of incorporating the DL estimator on the sensitivity of searching for a sterile neutrino in the  $\nu_\mu$  disappearance mode using the MicroBooNE detector [6] before concluding in Sec. VI.

## II. REVIEW OF VARIOUS NEUTRINO ENERGY ESTIMATORS

### A. The MicroBooNE experiment

The MicroBooNE detector [13] is a LArTPC consisting of 85 tonnes of liquid argon in the active volume, which is a rectangular volume measuring 10.36 m in length, 2.32 m in height, and 2.56 m in width. The time projection chamber (TPC) is placed inside a larger cylindrical cryostat which has a total capacity of 170 tonnes of liquid argon. The MicroBooNE detector sees the on-axis neutrinos produced from the Booster Neutrino Beam (BNB) at Fermilab National Accelerator Laboratory (FNAL) [27]. The distance between the MicroBooNE detector and the BNB target is about 468 m. The neutrinos from the BNB are predominantly ( $\sim 93.6\%$ )  $\nu_\mu$  with a mean energy of 0.8 GeV.

Ionization electrons are produced by the charged particles from the neutrino interaction traveling through LAr. Under an external electric field of 273 V/cm, these ionization electrons drift horizontally at a constant speed of around 1.1 mm/ $\mu$ s toward the anode plane. The anode plane consists of three planes of wires. The passage of ionization electrons induces bipolar readout signals on two of these planes (“induction planes”) oriented at  $\pm 60^\circ$  with respect to the vertical. The ionized electrons are then collected and induce unipolar signals at the third plane (“collection plane”) oriented vertically. The wire pitch is 3 mm. In addition, a set of 32 photomultiplier tubes (PMTs) is placed behind the wire planes to detect scintillation light from the interaction. The light signal provides a prompt timing signal for when the interaction occurred.

Three 2D views of the detector activity can be obtained from the projective LArTPC wire channel readouts and drift time. Using tomographic reconstruction algorithms [28], the three 2D views can be combined to create a complete 3D image of the detector activity.

### B. Energy estimators in MicroBooNE

In this section, we describe the traditional energy estimation algorithm in MicroBooNE, which serves as a baseline for the performance comparisons. In searching for a  $\nu_e$  low-energy excess [29], MicroBooNE performed three different analyses, each of which relied on completely different reconstruction paradigms, namely, Wire-Cell [15], Pandora [30], and deep learning [31]. Nevertheless, these analyses are based on the same calorimetry neutrino energy estimation strategy [21,28,32]. In the following, we briefly review the neutrino energy reconstruction within the Wire-Cell reconstruction paradigm, upon which the work of this paper is based.

Wire-Cell is a tomography-inspired algorithm that provides a 3D representation of the interaction based on the three 2D projection measurements from the anode planes [28,33]. The particles in the final state and their

4-momenta, including those from secondary interactions, are then reconstructed by various downstream algorithms that perform TPC-charge/PMT-light matching [33], trajectory fitting and particle identification, and neutrino vertex identification, so that the best description of the interaction (“particle-flow”) is produced [22].

The traditional Wire-Cell energy estimator used in MicroBooNE is based on the reconstructed particle flow tree information and employs a straightforward logic [22]. One of three methods is used to reconstruct the deposited energy for each particle: (i) range based method for tracklike particles stopping inside the detector with sufficient length ( $> 4$  cm), (ii) summation of  $dE/dx$  for other tracklike particles, and (iii) scaling of the summed charge for showerlike particles. The primary final-state charged lepton in CC neutrino interactions is selected based on reconstructed energy and particle identification information. The visible neutrino energy is reconstructed by adding up the single-particle kinematic energies, particle masses, and an averaged nucleon binding energy, 8.6 MeV, for each identified proton. More details can be found in Ref. [22]. As discussed in Sec. I, this particle flow summation method is general and robust against neutrino-argon interaction models, yet it suffers from bias induced by missing energy. This motivates our search for alternative algorithms.

### C. Energy estimation with convolutional neural networks

The NOvA experiment introduced a DL energy estimator based on a convolutional neural network (CNN) architecture [23]. This energy estimation method utilizes images of neutrino events and tries to predict the energy of the neutrino from the features present in the image. The CNN-based approach was shown to give superior energy reconstruction compared to the traditional energy estimation methods. However, the use of CNN-based networks is computationally expensive because of large computational costs associated with the convolutional operations. In addition, the CNN-based neutrino energy estimator is expected to demand an accurate detector simulation, as even tiny pixel-size systematic inaccuracies in the simulation could potentially be captured by the convolutional network, leading to the domain shift problem [34]. This problem arises when discrepancies between the training (simulation) data and the deployment data cause a model to perform poorly on new, unseen data when it is deployed. Recently, a CNN-based approach was applied to LArTPC detectors, for the task of shower energy estimation [35]. As a CNN algorithm, it shares similar trade-offs as the NOvA method.

### D. Previous works on energy estimation with recurrent neural networks

Large computational costs of DL algorithms are associated in part with a large dimensionality (e.g., a large

number of pixels) of inputs. To reduce such computational costs, one approach is to use alternative inputs that have fewer dimensions. For instance, high-energy physics (HEP) experiments usually run multiple simple and traditional reconstruction algorithms over the events (e.g., event type prediction, particle localization, etc.). The outputs of these algorithms are relatively low dimensional and less affected by the potential systematic differences between the simulation and the real experimental data. Therefore, one can build a robust and less computationally expensive DL energy estimator on top of these outputs.

The application of RNN to the task of neutrino energy estimation was pioneered by the NOvA experiment [36]. The RNN-based energy estimator consumes information from the reconstructed particles in each neutrino event and predicts the neutrino energy. Similar to the CNN-based energy estimation method, the RNN-based method is able to outperform the traditional energy reconstruction methods [37], but without incurring large computational costs.

In this work, we utilize the particle flow output from the state-of-the-art Wire-Cell event reconstruction paradigm [22] to estimate neutrino energy with the RNN. The use of the reconstructed particle flow information allows us to suppress the potential difference between the detector response simulation and experimental data [38,39].

## III. METHOD DESCRIPTION

### A. Energy estimation with recurrent neural networks

In this section, we describe how the recurrent neural network energy estimator operates. The existing Wire-Cell algorithms reconstruct individual particles in the LArTPC volume [22]. These reconstructed particles are hierarchically grouped into a structure called a particle flow (PF). For each particle in the PF, we know its starting and ending coordinates, the best estimate of the particle type, and rough estimates of the particle’s energy and momentum.

For each neutrino interaction event, the RNN energy estimator aggregates information from all particles present in the PF structure in order to make inferences about the neutrino energy and the energy of the primary outgoing lepton. The use of the PF information creates unique challenges and opportunities. First, the number of particles in each event varies, depending on the type of the neutrino interaction. Some neutrino interactions can produce just a few particles, while others create many of them. Therefore, in order to operate on the PF information, the DL model needs to be able to handle inputs of varying lengths. Second, each particle in the PF information has the same semantic meaning and the same format of input variables. This structure of inputs can be exploited to create a sample-efficient deep learning algorithm. In particular, the structure of inputs can be encoded as an inductive bias of the DL algorithm [40]. Adding such an inductive bias can lead to a

model with a smaller number of parameters, faster training times, and better performance with a limited amount of data.

An RNN is a natural deep learning model candidate to handle the PF information. An RNN model operates on a sequence of tokens of arbitrary lengths. It treats each token in the same way, thus ensuring the proper inductive bias. The RNN reads tokens sequentially and maintains a fixed-size memory state of the past tokens. Once all of the tokens have been consumed, one can use the resulting memory state to make inferences about the data. For the energy estimation problem, we consider each particle of the PF to be a separate token, use an RNN to aggregate all particles, and then use another DL model to predict the energy of the neutrino from the RNN's memory. The detailed architecture of this energy estimator is described in Sec. III B.

There is another, more recent, candidate DL model that can be used to work with the PF data—the transformer model [41]. In many domains, transformer-based models achieve state-of-the-art performance [42]. However, in order to achieve outstanding performance they require large amounts of training data, which are not available at MicroBooNE. Therefore, we have chosen to use the RNN-based model in this work, with plans to explore other network architectures in future studies.

## B. Model architecture

The RNN energy estimator is designed to predict both the energy of the neutrino and the energy of the primary final-state lepton in a neutrino interaction event. To make such predictions, it relies on reconstructed particle information available for each neutrino event (PF information). As inputs, the RNN estimator extracts the following

quantities from each particle: (i) particle track starting and ending coordinates, (ii) estimated particle momentum and energy, and (iii) estimated particle type. These quantities are reconstructed by the upstream Wire-Cell 3D pattern recognition algorithms [22]. Besides the particle-level information, the RNN energy estimator incorporates the information about the entire event such as (i) a flag indicating whether the event is fully (FC) or partially contained (PC) in the detector [15] volume and (ii) a prediction on whether the event is a  $\nu_\mu$  CC or a  $\nu_e$  CC event. The FC events are defined as those that have the reconstructed TPC activity fully contained within the fiducial volume (3 cm inside the effective TPC boundary [43]).

The architecture of the RNN is shown in Fig. 1. We chose a long short-term memory (LSTM) [44] neural network cell as a recurrent neural network model since it is rather stable to train. Before feeding the information from each particle into the LSTM cell, we perform a feature extraction step with the help of a fully connected network [45] (depicted as the Preproc branch in Fig. 1). The Preproc is constructed as a stack of three blocks with a similar structure followed by a batch normalization layer [46]. Each block of the Preproc is made out of a batch normalization layer, followed by a linear layer, followed by a ReLU activation function. Empirically, we found that the use of the Preproc improves the performance over the use of a plain LSTM cell. This is likely because the Preproc module can transform the raw particle information into a representation more useful for the LSTM gating mechanism. Additionally, we found that the three-block structure gives the best performance.

After the LSTM cell has finished processing particles in an event, the memory state of the LSTM (vector of 32 features) is concatenated with the event-level information.

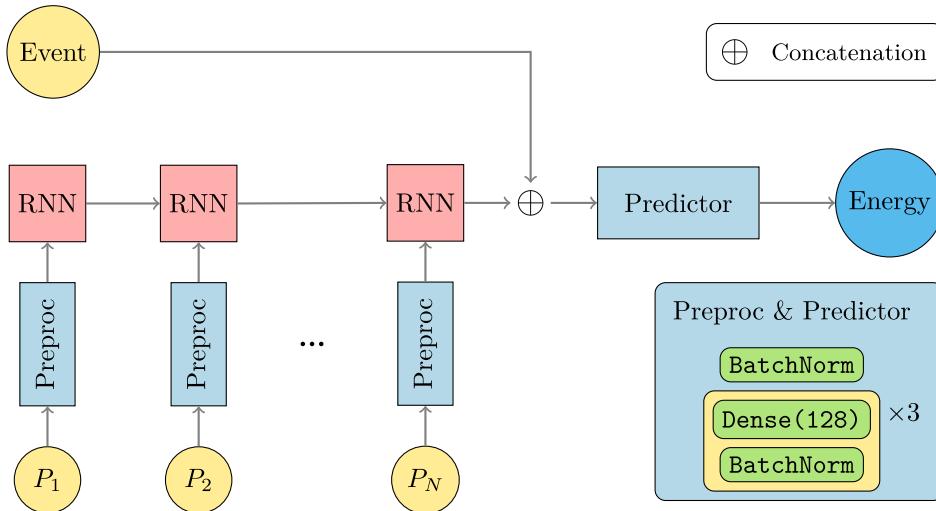


FIG. 1. The schematic representation of the energy estimator architecture. The yellow circle labeled *Event* denotes the event-level inputs. The circles labeled  $P_1 \dots P_N$  represent input variables coming from the reconstructed particles ( $1 \dots N$ ) in the event. *RNN* is a recurrent neural network cell. *Preproc* is a particlewise feed-forward neural network used to perform the feature extraction from the particle-level variables. *Predictor* is another feed-forward neural network that predicts neutrino and lepton energies.

The combined information is fed through another fully connected network (depicted as Predictor). The Predictor is responsible for extracting features useful for energy reconstruction from the combined event-level and particle-level information. The Predictor module has the same structure as the Preproc, which was found to give the best performance. The final energy reconstruction is done using linear layers, which take outputs of the Predictor and estimate the energy of the neutrino and the energy of the primary lepton in the event. In principle, more quantities can be estimated in this way. However, the neutrino energy and the primary lepton energy are of the greatest interest in this paper.

#### IV. DL MODEL TRAINING

The RNN energy estimator is trained on a simulated dataset made of true- $\nu_e$  and true- $\nu_\mu$  CC events. Both FC and PC events are included in the training sample. In this section, we show the performance on  $\nu_\mu$ CC events, where a large number of data events are available to perform dedicated model validations, as described in Sec. V.

##### A. Training dataset

In the MicroBooNE experiment, the simulated BNB neutrino flux [27] is provided to the event generator Genie [47,48] to generate neutrino-argon interactions. Genie v3.0.6, G18\_10a\_02\_11a, was used, which includes improvements on the use of the Valencia model [49–51] for the local Fermi gas nucleon momentum distributions, improvements in the CCQE and CC two-particles-two-holes (CC2p2h) interactions, and improvements in the treatments of final state interaction and pion production with respect to earlier versions. In addition to the default configuration, the parameters governing the CCQE and CC2p2h models are adjusted according to the T2K CC0 $\pi$  cross section results [52] to form the “MicroBooNE Tune” model [53]. The resulting final-state particles of each Monte Carlo (MC) simulated event are processed using the LArSoft [54] software framework, which is a toolkit to perform simulation, reconstruction, and analysis of LArTPC data. The final state particles are propagated through the detector using the Geant4 toolkit [55] v4\_10\_3\_03c. The resulting energy depositions are further processed by dedicated detector simulation programs taking into account detector effects to simulate the ionization charge and scintillation light signals after considering the space charge effect [56,57].

The position and number of ionization electrons modified by space charge and recombination effects are ported to the TPC detector simulation [38,39], which takes into account the charge transportation and diffusion [58]. The induced currents on the wires are simulated by convolving the ionization charge distribution at the wire plane with the position-dependent (at 1/10th of the wire pitch resolution)

TABLE I. Number of neutrino CC events in the training and test datasets obtained from the MicroBooNE simulation.

	Train	Test
$\nu_\mu$ FC	80,256	134,707
$\nu_\mu$ PC	156,628	303,422
$\nu_e$ FC	93,289	112,822
$\nu_e$ PC	66,792	68,061

field response function as well as the electronics response function. The optical detector simulation models the light emitted by charged particles interacting with the detector and produces signals in photomultiplier tubes.

The simulated neutrino interactions are further merged with a dedicated data stream which is collected in a period when there is no neutrino beam, ensuring faithful modeling of cosmic-ray backgrounds and detector noise. At the same time, this choice limits the number of available simulation events due to a finite number of cosmic-only data events. We split the simulated neutrino dataset into training/test partitions following a previously used MicroBooNE procedure [59]. Table I summarizes the final sample sizes used in the training and testing.

##### B. Initial training

Initially, we trained the RNN energy estimator (Fig. 1) on a dataset described in Table I including all four samples ( $\{\nu_\mu, \nu_e\} \times \{\text{FC, PC}\}$ ). The model was trained to predict both the energy of the neutrino and the energy of the primary lepton in the event. As a loss function for each target (neutrino and primary lepton), we used a mean absolute percentage error:

$$\mathcal{L}_{\nu, \text{lep}} = 100 \cdot \left| \frac{E^{\text{reco}} - E^{\text{true}}}{E^{\text{true}}} \right|, \quad (2)$$

where  $E^{\text{true}}$  is the true energy of the particle, and  $E^{\text{reco}}$  is its predicted energy. The total loss function is a sum of losses for each target, i.e.,  $\mathcal{L} = \mathcal{L}_\nu + \mathcal{L}_{\text{lep}}$ . We have found this loss function to perform much better than the traditional regression losses (mean squared error and mean absolute error [45]). The training was performed for 200 epochs (complete passes through the training dataset) with a ReduceLROnPlateau learning rate scheduler [60] allowing the energy estimator to converge. The complete training details can be found in Appendix A.

The sequential nature of the RNN network may make the energy estimator dependent on the particular ordering of the particles used in the training. To ensure that the energy estimator performance does not depend on the particle ordering, we have implemented particle order randomization as a data augmentation strategy.

To assess the performance of the energy estimator, one can consider its energy resolution defined as a ratio of

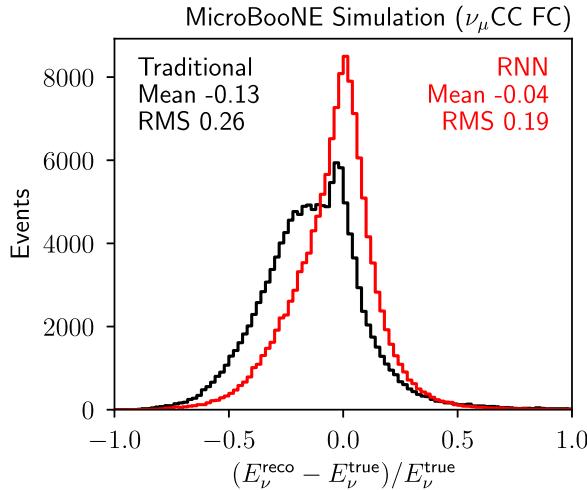


FIG. 2.  $\nu_\mu$  energy resolution histograms for the traditional MicroBooNE energy estimator (black) and the initial training of the RNN energy estimator (red) in the FC CC sample.

$(E^{\text{reco}} - E^{\text{true}})/E^{\text{true}}$ . Figure 2 shows distributions of the  $\nu_\mu$  CC energy resolutions of the FC events. The distribution of the RNN energy resolution (shown in red) has a smaller width compared to the traditional energy estimator (shown in black). This indicates that the RNN energy estimator is able to better predict the true neutrino energy  $E^{\text{true}}$ .

For a more quantitative assessment of the performance, one can consider two characteristics of the energy resolution—its mean and RMS values:

$$\text{Mean} := \mathbb{E} \left[ \frac{E^{\text{reco}} - E^{\text{true}}}{E^{\text{true}}} \right] \quad (3)$$

$$\text{RMS} := \sqrt{\mathbb{E} \left[ \left( \frac{E^{\text{reco}} - E^{\text{true}}}{E^{\text{true}}} \right)^2 \right]}, \quad (4)$$

where  $\mathbb{E}$  is an expectation over the test sample. Generally, smaller RMS values correspond to better energy estimators. The mean value indicates the overall bias of the energy estimator. According to Fig. 2, the RNN energy estimator (RMS  $\sim 0.19$ ) achieves about 26% improvement compared to the traditional MicroBooNE energy reconstruction method (RMS  $\sim 0.26$ ) in the test sample.

Another metric that is commonly considered in HEP experiments is the bias of the energy estimator. To determine a bias of the energy estimator, a binned statistics plot is made, where the x axis represents the true energy and the y axis shows a mean of the energy resolution histogram constructed for each bin. Figure 3 shows the bias plot for the RNN energy estimator (red) and the traditional MicroBooNE energy estimator (black). Any deviation from zero represents the bias of the energy estimator in a particular true energy bin.

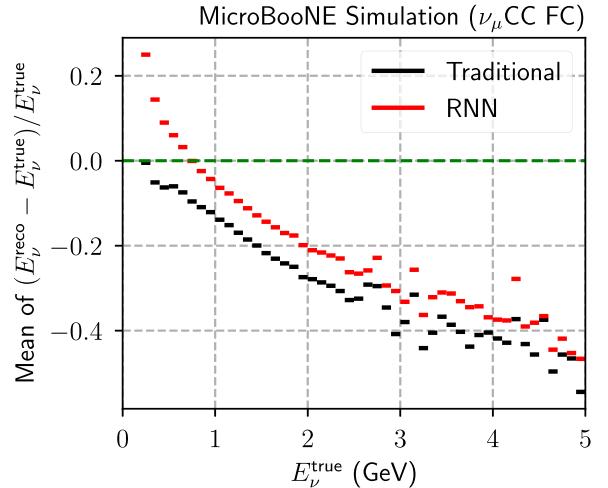


FIG. 3.  $\nu_\mu$  energy bias for the traditional MicroBooNE energy estimator (black) and the initial training of the RNN energy estimator (red) in the FC CC sample.

According to Fig. 3, the RNN energy estimator has a smaller bias compared to the traditional MicroBooNE energy estimator for energies above  $\sim 0.6$  GeV. However, at lower energies the RNN energy estimator quickly acquires a rather significant bias. Such a large bias is not ideal, since it may lead to a reduction in physics sensitivities.

### C. Reducing bias

In this section, we develop a mitigation strategy for the large bias of the RNN energy estimator. Before developing such a strategy, it is instructive to consider the reasons for the appearance of the bias. We believe there are two main sources of bias; one is related to physics, and the other is related to ML.

From the physics point of view, as the energy of the neutrino gets higher, a larger fraction of this energy becomes invisible in neutrino-argon interactions. This happens due to the increased production of various mesons at higher energies. Meson decays have a common byproduct—neutrinos, which easily escape the detector carrying some fraction of the energy of the original interaction away. Therefore, one may expect to see an increasingly negative (missing energy) bias at high energies for all estimators. Figure 3 shows that both energy estimators acquire large negative biases at high energies.

From the ML side, the peaked nature of the energy distribution in the training sample (Fig. 4) can also contribute to the bias. Since the target energy distribution has high population around the 1 GeV peak, energy estimators will prioritize correctly reconstructing the energy of neutrinos around the peak. Moreover, on average, an energy estimator can increase the overall accuracy of predicted energies by slightly pushing all the energies toward the peak. That is, the energy of neutrinos to the left of the peak will be pushed up (positive bias), and the energy

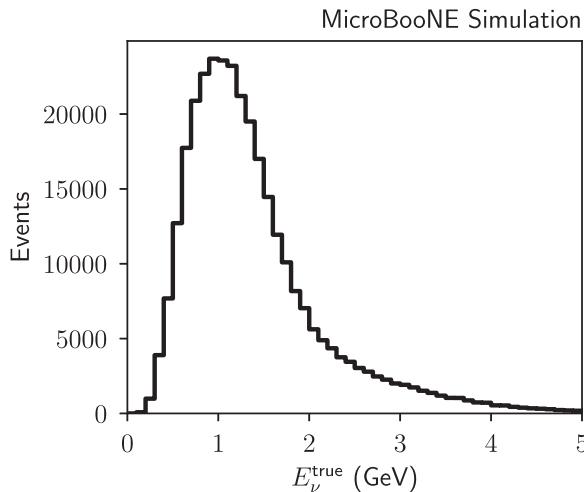


FIG. 4. True neutrino energy distribution in the training sample.

of neutrinos to the right of the peak will be pushed down (negative bias). This prediction is consistent with the observed behavior of the bias in Fig. 3.

The proper way to improve the bias, stemming from the peaked nature of the distribution, is to resimulate the training sample with a flat true neutrino energy distribution. Since we are limited by the total number of simulation events available, we decided instead to use the standard ML approaches to deal with imbalanced data. In particular, we applied event reweighting to flatten the true neutrino energy distribution. The new event weights were constructed using the following procedure.

A histogram of the true neutrino energy  $N_i$  was made in a range of 5 GeV with 50 bins. A weight histogram  $W_i$  was constructed, so that each bin's height is proportional to the inverse of the true neutrino energy bin height, i.e.,  $W_i \propto 1/N_i$ . In training the RNN energy estimator, the weight for each event was determined from the corresponding bin of the weight histogram  $W_i$ . To preserve the scale of the loss function, the weights were normalized to add up to unity.

The weight construction procedure described above guarantees that the true neutrino energy distribution is approximately flat. However, the tails of the distribution in Fig. 4 will have to acquire quite large weights, since there are too few events in those tails. Giving large weights to a few events will result in severe overfitting in the model training. In order to reduce the magnitude of overfitting, we clipped the maximum weight value, ensuring the ratio of the maximum weight to the minimum weight does not exceed 50. Figure 5 shows the true neutrino energy spectrum of the training sample after the reweighting.

Figure 6 shows the  $\nu_\mu$  neutrino energy bias after retraining with the flat weights. It demonstrates that the use of the reweighted training sample improves the bias of the energy estimator.

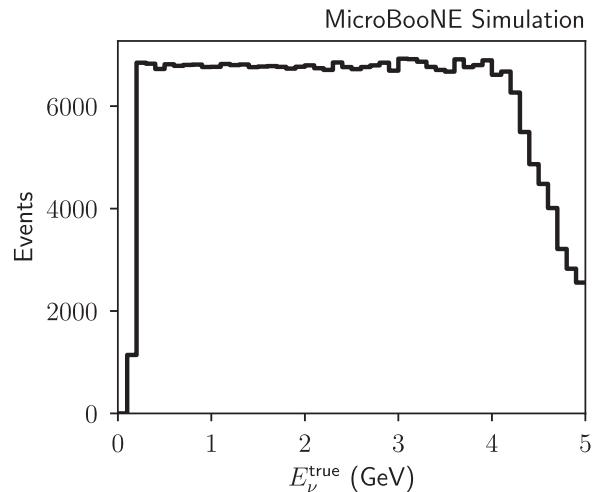
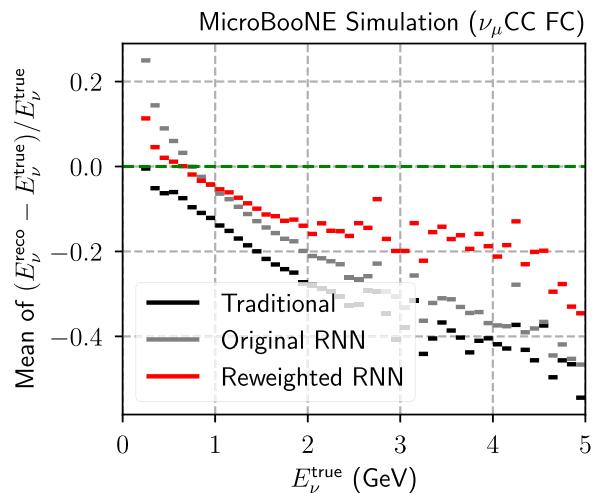


FIG. 5. Reweighted true neutrino energy distribution in the training sample.

FIG. 6.  $\nu_\mu$  energy resolution bias for the traditional MicroBooNE energy estimator (black) and the reweighted training of the RNN energy estimator (red) in the FC CC sample.

## V. DL MODEL EVALUATION

### A. Performance metrics in testing simulation sample

In the previous section, we showed the basic  $\nu_\mu$  energy reconstruction performance plots for the RNN energy estimator in the test sample. We now perform a more detailed evaluation of the RNN energy estimator. Figure 7 demonstrates the  $\nu_\mu$  neutrino energy resolution after retraining with the flat weights. The use of the reweighted training sample produces a small degradation of the RNN energy resolution (RMS increases from  $\sim 19\%$  to  $\sim 20\%$ ).

We believe that the reduction of the bias of the energy estimator, brought by the reweighting, outweighs the associated small degradation of the RMS value. This is partly because the reduction of the bias makes the energy

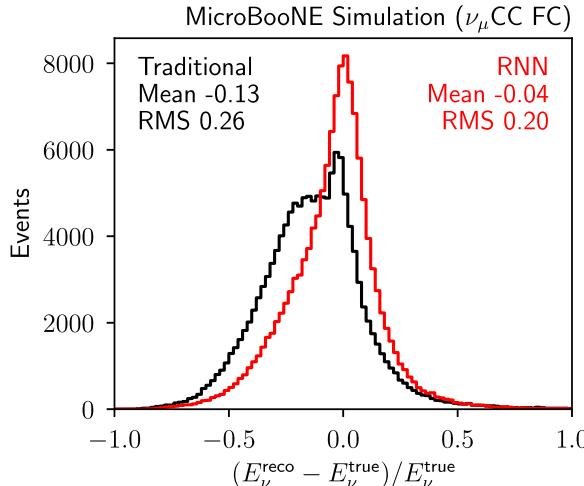


FIG. 7.  $\nu_\mu$  energy resolution histograms for the traditional MicroBooNE energy estimator (black) and the reweighted training of the RNN energy estimator (red) on the FC CC sample. The “shoulder” at the lower reconstructed energy comes from a combined effect of missing hadronic energy and the biased recombination model.

estimator more agnostic to the choice of the neutrino energy spectrum. For instance, the original behavior of the bias stems from the BNB energy spectrum having a peak around 0.8 GeV. However, MicroBooNE can also study the NuMI neutrino beam [61] with higher average energies. Having an unbiased energy estimator would allow us to expect better transferability of its performance between different neutrino beams.

Since the reweighted training results in a better energy estimator overall, we will use the reweighted version in the subsequent analysis. Likewise, each time we refer to the RNN energy estimator below, we will imply its reweighted version.

Apart from predicting the  $\nu_\mu$  energy, the RNN energy estimator is also capable of predicting the energy of the primary lepton. Figure 8 shows energy resolution histograms for the energy reconstruction of the primary outgoing  $\mu$  in the  $\nu_\mu$  CC events. The traditional energy estimator exhibits two peaks in the shape of the muon energy resolution. We discuss the nature of the peaks later in this section. The RNN energy estimator has a well-behaved energy resolution histogram with a single peak. Overall, the RNN energy estimator displays better energy resolution in terms of RMS (15% vs 22%).

The unusual shape of the energy resolution of the traditional estimator is traced back to reconstruction errors. The central (main) peak has a small bias and corresponds to correctly identified muons. The energies of these muons are mostly reconstructed from range information. The resolution of these events is, therefore, good. The second peak with a negative bias is traced back to muons for which the energy was reconstructed by integrating the energy loss per unit length  $dE/dx$ . Biases in the detector modeling,

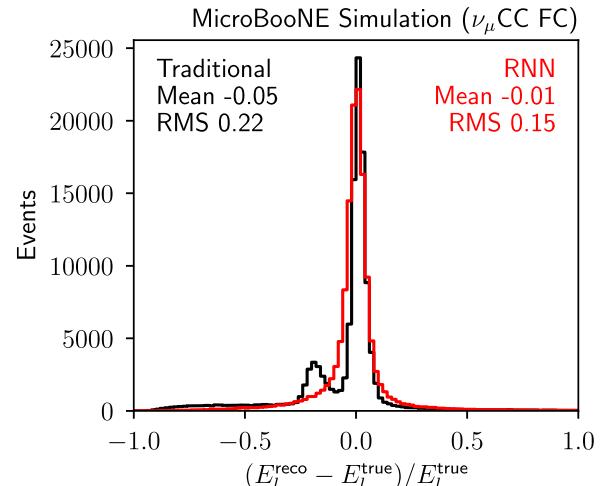


FIG. 8. Primary  $\mu$  energy resolution histograms for the traditional MicroBooNE energy estimator (black) and the reweighted training of the RNN energy estimator (red) on the FC CC sample.

including the inaccuracies in charge recombination, contribute to this offset. This is evident as a second peak in Fig. 8 and an enhanced left “shoulder” in Fig. 7. The same recombination model was used for reconstructing the simulation, and dedicated validation tests have demonstrated its consistency with the data [15,26].

As shown in Fig. 8, the peak of the RNN’s muon energy resolution is shorter and wider than the first (main) peak of the traditional energy estimator. The degradation of the energy resolution in the main peak is likely a result of the inability of the RNN energy estimator to differentiate between various types of the imperfect event reconstruction. Because the RNN energy estimator uses reconstructed particle flow information, it inherits all the shortcomings of the particle flow reconstruction. In addition, unlike the traditional energy estimator, we do not provide any explicit indicators of the reconstruction quality to the RNN model. Therefore, even for the properly reconstructed muons, we may anticipate a degradation of the performance of the RNN method compared to the traditional method.

In principle, in the absence of explicit indicators of the reconstruction failures, a DL model could try to infer them from some other features (e.g., incorrect event topology). However, it is unclear what fraction of information about the reconstruction quality can be inferred from the existing features. Moreover, very large datasets are required for DL models to make such complex inferences. Therefore, the relatively small size of our training dataset may contribute to the degradation of the quality of the main muon peak.

While the RNN energy estimator does not directly predict the hadronic part of the neutrino energy, it can be trivially inferred by subtracting the energy of the primary muon from the total neutrino energy. Figure 9 shows the hadronic energy resolution histograms for the reweighted RNN and the traditional MicroBooNE energy

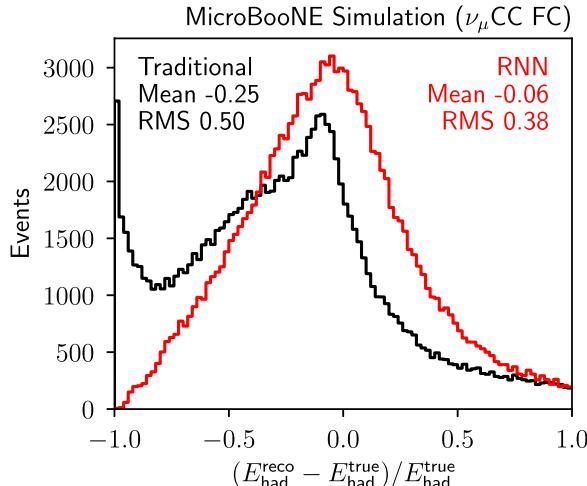


FIG. 9. Hadronic energy resolution histograms for the traditional MicroBooNE energy estimator (black) and the reweighted training of the RNN energy estimator (red) on the FC CC sample.

estimators. The RNN hadronic energy resolution exhibits more Gaussian-like behavior and has a much better RMS ( $\sim 38\%$  vs  $\sim 50\%$ ).

In this section, we have examined the performance of the RNN energy estimator on a sample of fully contained  $\nu_\mu$  CC events. The RNN estimator is also capable of predicting the energy of partially contained  $\nu_\mu$  CC events and  $\nu_e$  CC events. The corresponding performance evaluations are provided in the Appendixes, where the results are consistent with those presented in the main text. Specifically, the performance of the RNN energy estimator for  $\nu_e$  events is discussed in Appendix B. Similarly, Appendix C demonstrates the performance of the RNN energy estimator on partially contained events. Finally, Appendix D explores the behavior of the RNN energy resolution across various ranges of true neutrino energy.

### B. Model validation with experimental data

When developing new ML algorithms for HEP experiments, it should always be kept in mind that the algorithms are trained on a simulated sample, but eventually are applied to experimental data. The simulation data in HEP experiments commonly possess systematic differences from the real experimental data. For example, as briefly reviewed in Sec. I, the response of argon nuclei to a neutrino probe depends on the complex nuclear structure and quantum chromodynamics in the nonperturbative region, which is at the frontiers of nuclear physics research. In addition, understanding LArTPCs' response and calibration has considerable room for improvement. Because of these differences between data and simulation, it is possible that ML algorithms have systematically different performance when applied to real experimental data than those evaluated through simulations.

HEP communities have developed a strict scheme of estimating systematic uncertainties to evaluate and quantify

the differences between data and simulation. Different sources of systematic uncertainties serve as effective degrees of freedom in describing the differences between prediction (e.g., simulation) and data. Instead of requiring that the simulation faithfully reproduce every feature in the data, the HEP community requires that the differences between the data and simulation are within the quoted systematic uncertainties. In other words, the simulation is required to be compatible with data within its quoted uncertainties.

Since the RNN energy estimator combines the calorimetry information, which is sensitive to detector response modeling, and the kinematics information, which is sensitive to the complex neutrino-argon interaction, we expect differences in performance between data and simulation. We utilize a MicroBooNE dataset to demonstrate the compatibility between simulation that is enhanced with the RNN energy estimator and data. This dataset was collected from February 2016 to July 2018 corresponding to an exposure of  $6.369 \times 10^{20}$  protons on target from the BNB at FNAL that was used to search for a low-energy  $\nu_e$  excess [15]. As elaborated in Ref. [15], the sources of systematic uncertainties associated with the simulation include (i) neutrino flux, (ii) neutrino-argon interaction cross sections, (iii) detector effects, and (iv) statistical uncertainties due to a finite number of simulation events with unequal weights.

Following previous work [15,26], the primary tools to test the compatibility between data and simulation (i.e., model validation) are based on goodness-of-fit (GoF) tests, which allow one to quantify the comparison of data and predictions into a single number for evaluation. As detailed in Sec. V of [15], the term “simulation” within the context of MicroBooNE refers to the overall simulation model. This model consists of various components, including the neutrino beam flux model, the neutrino-argon interaction model, the detector model, and the reconstruction algorithms, among others. Therefore, changes in the energy estimator impact the consistency between data and simulation.

While the total GoF test is essential, it may hide some problems of the model when some model uncertainties are overestimated. Additionally, since the missing hadronic energy  $E^{\text{miss}}$  cannot be directly measured, an event generator (or a neutrino-nucleus interaction model) is often required to describe  $E^{\text{miss}}$  accurately in order to ensure a correct mapping from  $E_\nu^{\text{reco}}$  to  $E_\nu^{\text{true}}$ . This mapping is crucial as neutrino oscillation measurements rely on estimations of  $E_\nu^{\text{true}}$ . However, modeling  $E^{\text{miss}}$  remains a challenging theoretical problem, particularly for heavy nuclei such as argon where final state interactions can produce a variety of complex final states that contain significantly different amounts of missing energy, even for a simple quasielastic interaction. In order to mitigate these shortcomings, we implement a conditional constraining procedure. In Refs. [15,26], this procedure was

used to validate the modeling of missing hadronic energy and its associated uncertainties. The validation was performed by comparing the reconstructed hadronic energy distribution between data and an MC prediction after constraining the reconstructed muon kinematic distributions (i.e., energy and polar angle) to those of data. The validation of the mapping between the true and the reconstructed neutrino energy enables the measurement of neutrino-energy dependent total and differential cross sections [26,62] as well as searches for sterile-neutrino-induced oscillations [6].

In the following, we briefly review the model validation procedure, which is based on a covariance matrix formalism in constructing the  $\chi^2$  test statistic:

$$\chi^2 = (M - P)^T \times \text{Cov}_{\text{full}}^{-1}(M, P) \times (M - P), \quad (5)$$

where  $M$  and  $P$  are vectors of measurement and prediction, respectively. The  $\text{Cov}(M, P)$  is the full covariance matrix:

$$\begin{aligned} \text{Cov}_{\text{full}} = & \text{Cov}_{\text{stat}}^{\text{stat}} + \text{Cov}_{\text{MC stat}}^{\text{sys}} + \text{Cov}_{\text{xs}}^{\text{sys}} + \text{Cov}_{\text{flux}}^{\text{sys}} \\ & + \text{Cov}_{\text{det}}^{\text{sys}} + \text{Cov}_{\text{add}}^{\text{sys}}. \end{aligned} \quad (6)$$

The  $\text{Cov}_{\text{stat}}^{\text{stat}}$ ,  $\text{Cov}_{\text{MC stat}}^{\text{sys}}$ ,  $\text{Cov}_{\text{xs}}^{\text{sys}}$ ,  $\text{Cov}_{\text{flux}}^{\text{sys}}$ ,  $\text{Cov}_{\text{det}}^{\text{sys}}$ , and  $\text{Cov}_{\text{add}}^{\text{sys}}$  terms represent the statistical uncertainties of the data sample, the statistical uncertainties corresponding to finite statistics in simulation, systematic uncertainties in cross section modeling, systematic uncertainties from the modeling of the neutrino flux, systematic uncertainties from detector response modeling, and additional systematic uncertainties associated with estimating background events from outside the cryostat, respectively.

The GoF evaluation is performed to test the compatibility between the data and the overall simulation model using Eq. (6). The  $\chi^2$  value can be used to perform a GoF test and to deduce a p value by comparing to the  $\chi^2$  distribution with the associated number of degrees of freedom ( $ndf$ ), which is the total number of bins used in the measurement. For example, Fig. 10 demonstrates a data-MC comparison of the selected FC  $\nu_\mu$  CC events as a function of reconstructed neutrino energy. The p value is above 0.05, which is a predefined threshold for each GoF test. We should note that the MicroBooNE overall simulation model contains many conservative systematic uncertainties (e.g., cross section uncertainties). Therefore, the reduced  $\chi^2$  values, which are the ratios between  $\chi^2$  and the number of degrees of freedom, are generally low suggesting that the overall simulation model describes the data well within its uncertainties.

The conservative estimation of systematic uncertainties aims to determine uncertainties that are large enough to cover all reasonable systematic differences between simulated and real data. However, the resulting uncertainties may be overestimated, potentially obscuring deficiencies in the overall simulation model. To address this shortcoming, the global goodness-of-fit test can be enhanced to study

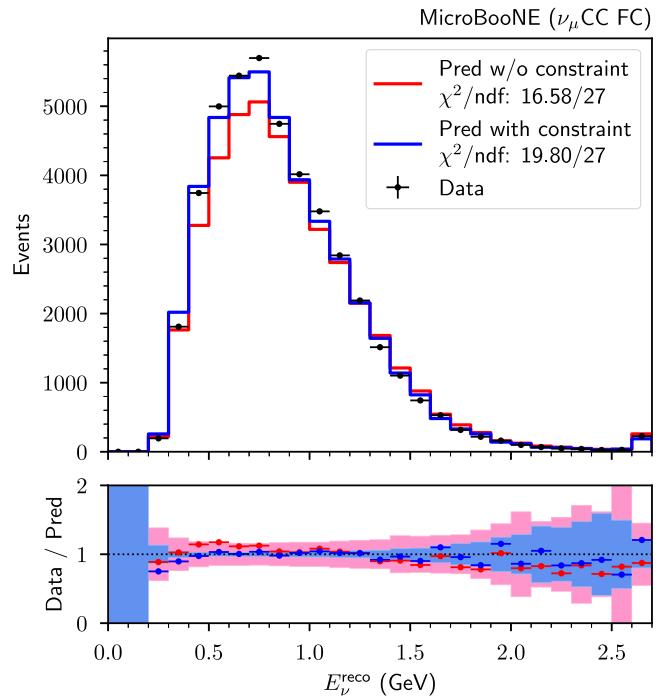


FIG. 10. Top: distribution of the selected FC  $\nu_\mu$  CC events as a function of the reconstructed neutrino energy. The MC prediction after applying constraints on muon kinematics ( $E_\mu^{\text{reco}}$  and  $\cos\theta^{\text{reco}}$ ) and hadronic energy ( $E_{\text{had}}^{\text{reco}}$ ) is shown in blue, and before applying in red. The last bin represents all events with  $E_\nu^{\text{reco}} > 2.6$  GeV. Bottom: the blue (red) points represent the ratio between data and the MC prediction with (without) constraint, and the similarly colored bands (light blue for blue and pink for red) depict the  $\pm 1\sigma$  of the total uncertainty (statistical and systematic) of the MC central prediction.

different parts of the overall simulation model using the conditional covariance matrix formalism [63].

For example, consider two quantities (channels)  $X$  and  $Y$  with the goal of assessing data-simulation differences between their simulated predictions ( $X_{\text{MC}}, Y_{\text{MC}}$ ) and the actual measurements ( $X_{\text{Data}}, Y_{\text{Data}}$ ). These quantities could correspond to any measurements, such as the reconstructed neutrino energy, muon energy, muon angle, etc. One can perform a direct GoF test by comparing  $X_{\text{MC}}$  to  $X_{\text{Data}}$  (and similarly for  $Y$ ), but such a test may suffer from the overestimation of systematic uncertainties. Additionally, if one can simulate conditional probabilities of the form  $P(X|Y)$ , then the difference between  $Y_{\text{MC}}$  and  $Y_{\text{Data}}$  can be used to refine  $X_{\text{MC}}$  and constrain the magnitudes of its systematic uncertainties. This allows us to create a more stringent GoF test. The procedure of refining  $X_{\text{MC}}$  from  $Y$  is referred to as constraining  $X$  on  $Y$  using the conditional constraining method.

Figure 10 illustrates data-simulation comparisons of the RNN neutrino energy for  $\nu_\mu$  CC FC events. The red curve shows the direct prediction of the neutrino energy with the RNN energy estimator. The blue curve is a prediction of the neutrino energy after constraining on muon kinematics

(reconstructed muon energy and angle) and hadronic energy simultaneously. The constraining procedure significantly reduces the magnitudes of the systematic uncertainties (shaded areas in the ratio plot) and provides a more precise GoF test. The  $\chi^2/ndf$  results are favorable both before and after applying the constraints, indicating that the data-simulation differences are well within the model uncertainties.

Following this methodology, we have run a comprehensive and systematic set of MicroBooNE validation tests on the overall simulation model that has been enhanced by the RNN energy estimator using CC  $\nu_\mu$  interactions. We should note that the CC  $\nu_e$  interactions are limited by the data statistics at  $O(100)$ , which is not sufficient to perform precision tests on the models. Table II summarizes the results of these model validation tests, with the  $p$  values above 0.05 indicating a successful test. The RNN energy estimator passes all the tests successfully. The large  $p$  values suggest that the differences between data and simulation are well within the quoted uncertainties.

### C. Sensitivity studies of searching for a sterile neutrino in $\nu_\mu$ disappearance

In this section, we show the impact of the RNN energy estimator on the sensitivity of determining neutrino oscillation parameters using the example of a  $\nu_\mu$  disappearance search. Assuming  $\nu_\mu$  disappearance only, the oscillation probability formula is

$$P_{\nu_\mu \rightarrow \nu_\mu} = 1 - \sin^2 2\theta_{\mu\mu} \sin^2(\Delta m_{41}^2 L/E_\nu), \quad (7)$$

TABLE II. Data vs MC validation results of the reweighted RNN energy estimator. The first column of the table shows the label of the statistical test that was performed. The second and third columns indicate  $p$  values returned by the test for the traditional and RNN energy estimators respectively. A  $p$  value above 0.05 indicates that the respective test was passed.

Test	$p$ value	
	Traditional	RNN
$P(E_\mu)$	0.89	0.92
$P(E_{\text{had}})$	1.00	0.99
$P(E_\nu)$	1.00	0.97
$P(E_\mu^{\text{PC}} E_\mu^{\text{FC}})$	0.95	0.70
$P(E_{\text{had}}^{\text{PC}} E_{\text{had}}^{\text{FC}})$	1.00	0.96
$P(E_\nu^{\text{PC}} E_\nu^{\text{FC}})$	1.00	0.84
$P(E_\mu \cos\theta_\mu)$	0.45	0.50
$P(E_{\text{had}} \cos\theta_\mu)$	1.00	0.99
$P(E_\nu \cos\theta_\mu)$	1.00	0.97
$P(E_{\text{had}} E_\mu)$	1.00	0.97
$P(E_\nu E_\mu)$	1.00	0.99
$P(E_{\text{had}} E_\mu, \cos\theta_\mu)$	1.00	0.99
$P(E_\nu E_\mu, \cos\theta_\mu)$	1.00	1.00
$P(E_\nu E_\mu, \cos\theta_\mu, E_{\text{had}})$	1.00	0.97

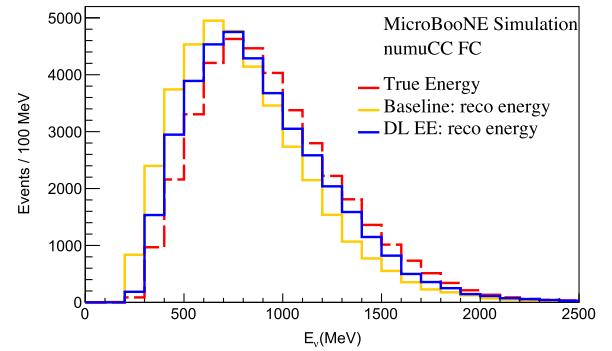


FIG. 11. Reconstructed energy spectra of the selected  $\nu_\mu$  CC FC events assuming no neutrino oscillation in simulation, normalized to  $6.369 \times 10^{20}$  POT.

where  $L/E_\nu$  is the ratio of the neutrino traveling distance and its energy,  $\theta_{\mu\mu}$  is the mixing angle that determines the oscillation magnitude, and  $\Delta m_{41}^2$  is a mass-squared splitting that modulates the oscillation frequency. The better energy resolution of the RNN energy estimator induces less smearing of the oscillating features, which leads to a better sensitivity in determining neutrino oscillation parameters.

The CC  $\nu_\mu$ -argon interaction event selection and corresponding statistical and systematic uncertainties are taken from Ref. [15]. Figure 11 shows the reconstructed energy spectra of the selected  $\nu_\mu$  CC FC events with no oscillations. Compared to the traditional energy estimator, the RNN energy estimator reconstructs a neutrino energy spectrum that is much closer to the truth value. Figure 12 shows the sensitivity contours at the 95% CL in the plane of  $\Delta m_{41}^2$  and  $\sin^2 2\theta_{\mu\mu}$  for the traditional energy reconstruction and the RNN energy estimator, respectively. As expected, the sensitivity is improved especially in the intermediate  $\Delta m_{41}^2$  region where the oscillation pattern

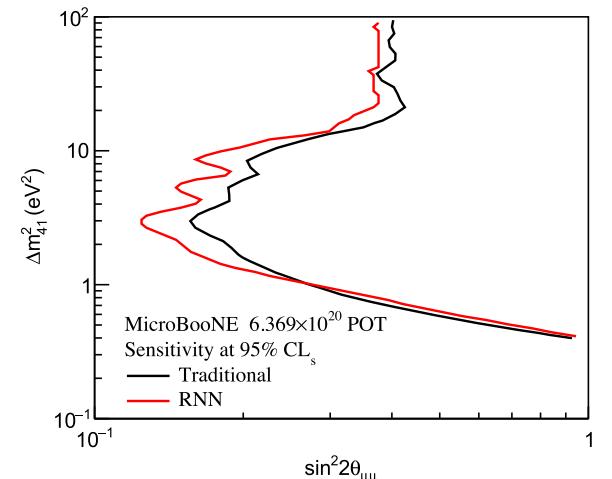


FIG. 12. MicroBooNE Gaussian  $\text{CL}_s$  [6] sensitivity contours at the 95% CL in the plane of  $\Delta m_{41}^2$  and  $\sin^2 2\theta_{\mu\mu}$  from the traditional (black) and the RNN (red) methods.

changes considerably over different neutrino energies. To avoid biases, the study depicted in Fig. 12 does not utilize any training samples from the RNN model.

## VI. CONCLUSIONS

We developed a deep-learning energy estimation method for charged current neutrino interactions in LArTPC detectors. This method is based on an RNN architecture. As inputs it uses reconstructed calorimetry and final-state particle kinematics obtained from the particle flow information. As outputs, it provides inferences about the energy of the neutrino and the energy of the primary outgoing lepton. Evaluating this method with simulations, we have shown that the RNN energy estimator's performance is superior to the Wire-Cell traditional energy estimator in terms of bias and resolution.

In order to make this method more resilient to inconsistencies between simulated and real data, we apply the RNN to the reconstructed particle flow information. This is at the cost of inheriting some inefficiencies from the event reconstruction algorithm. In principle, the RNN is capable of correcting for both miscalculated particle energies and energy from particles that were not reconstructed (either due to inefficiencies or because they do not produce observable signatures). By training an RNN on only the Wire-Cell traditional neutrino energy estimate, we can determine to what extent the RNN leverages individual particle information to determine the missing energy contribution. In this case, the performance is better than the traditional calculation, but falls short compared to the full RNN, indicating that the RNN can indeed learn correlations between the reconstructed particles and the missing energy.

Besides performance evaluations with simulations, a set of dedicated model validation tests was performed to demonstrate that the overall simulation model enhanced by the RNN energy estimator is compatible with MicroBooNE experimental data within the model uncertainties. Using a simple example of searching for a sterile neutrino with  $\nu_\mu$  disappearance oscillations, we show the impact of this RNN energy estimator on the physics sensitivity. Besides the impact on neutrino oscillations, we expect that this technique can be extended to cross-section measurements in both inclusive and exclusive interaction channels. Adaptation of this method for other LArTPC experiments, such as DUNE [64] and SBND [65], is underway. Additionally, we plan to explore other network architectures in future work to further enhance the robustness and applicability of our approach.

## ACKNOWLEDGMENTS

This document was prepared by the MicroBooNE Collaboration using the resources of the Fermi National Accelerator Laboratory (Fermilab), a U.S. Department of

Energy, Office of Science, HEP User Facility. Fermilab is managed by Fermi Research Alliance, LLC (FRA), acting under Contract No. DE-AC02-07CH11359. MicroBooNE is supported by the following: the U.S. Department of Energy, Office of Science, Offices of High Energy Physics and Nuclear Physics; the U.S. National Science Foundation; the Swiss National Science Foundation; the Science and Technology Facilities Council (STFC), part of the United Kingdom Research and Innovation; the Royal Society (United Kingdom); the UK Research and Innovation (UKRI) Future Leaders Fellowship; and the NSF AI Institute for Artificial Intelligence and Fundamental Interactions. Additional support for the laser calibration system and cosmic ray tagger was provided by the Albert Einstein Center for Fundamental Physics, Bern, Switzerland. We also acknowledge the contributions of technical and scientific staff to the design, construction, and operation of the MicroBooNE detector as well as the contributions of past collaborators to the development of MicroBooNE analyses, without whom this work would not have been possible. The initial development of the RNN energy estimator occurred at the NOvA experiment. We are grateful to the NOvA Collaboration for multiple discussions and suggestions to improve the RNN energy estimator.

## APPENDIX A: TRAINING DETAILS

The RNN energy estimator is trained with the help of the VLNE [66] package, which is built on top of the TensorFlow/Keras (v2.9) frameworks. The training is performed for up to 200 epochs using the RMSprop optimizer, and an initial learning rate of 0.001. The learning rate is progressively annealed by the ReduceLROnPlateau scheduler (patience = 5, factor = 0.5).

The RNN training is terminated if the validation loss has not improved over the last 40 epochs, or if the total number of training epochs has reached 200. To reduce overfitting, an  $L_2$  regularization is applied with a strength of 0.008. As an additional regularization technique we have randomly shuffled the order of particles in each event.

For the training we have used a batch size of 1024. When using the flat weights, we applied them only to the neutrino part  $\mathcal{L}_\nu$  of the loss function. We did not reweight the primary lepton part  $\mathcal{L}_{\text{lep}}$  of the loss. The training is run on a single NVIDIA GeForce RTX 3090 GPU. Because of the reliance on high-level input variables, it takes less than 20 minutes to train a single RNN energy estimator.

To determine the final network configuration we have performed several hyperparameter sweeps. We have explored a grid of the following parameters: learning rate, number of training epochs, number of features in the LSTM cell and fully connected layers, depth of the fully connected layers, regularization type ( $L_1$  vs  $L_2$ ), and the strength of the regularization.

## APPENDIX B: RNN FOR $\nu_e$ CC ENERGY ESTIMATION

In this appendix we explore the basic performance of the RNN energy estimator on the  $\nu_e$  events. The training sample of the RNN energy estimator contains charged-current electron neutrino interaction ( $\nu_e$  CC) events. Therefore, the RNN is capable of predicting energies of such events as well. Because of the small statistics of the  $\nu_e$  CC sample we are unable to perform a similar data/MC validation as that with the  $\nu_\mu$  CC sample. This situation is expected to be improved with future experiments.

Figure 13 shows the  $\nu_e$  CC energy resolution histogram of the RNN compared to the traditional energy estimator. In terms of the RMS, the RNN slightly outperforms the traditional energy estimator (18% vs 21%). Figure 14

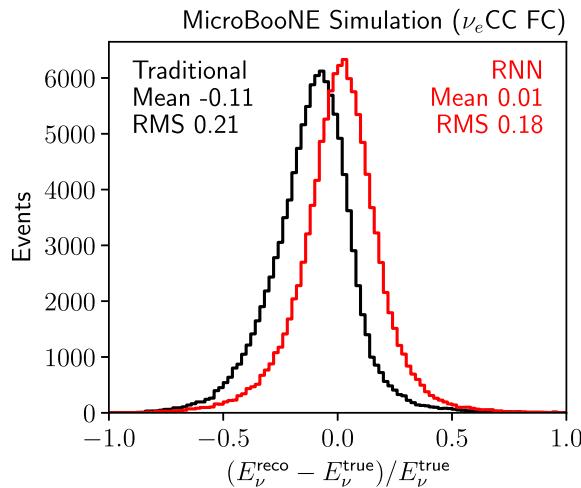


FIG. 13.  $\nu_e$  energy resolution histograms for the traditional MicroBooNE energy estimator (black) and the reweighted training of the RNN energy estimator (red) in the FC CC sample.

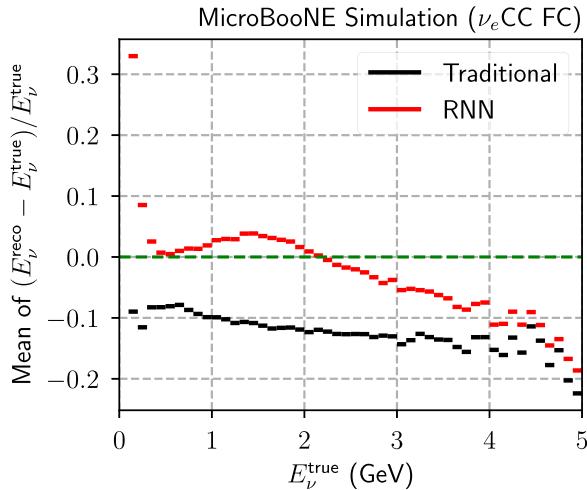


FIG. 14.  $\nu_e$  energy resolution bias for the traditional MicroBooNE energy estimator (black) and the reweighted training of the RNN energy estimator (red) in the FC CC sample.

illustrates biases of the RNN and traditional energy estimators. For the majority of true energy bins, the RNN displays a much smaller bias compared to the traditional energy estimator. Only at the lowest energy bin does the RNN acquire a large bias. While the  $\nu_e$  bias curve exhibits a smaller bias compared to the traditional energy estimator, it presents an unusual feature: a distinct hump around 1.5 GeV that is absent in the  $\nu_\mu$  counterpart. This hump in the  $\nu_e$  energy bias stems from the reweighting procedure, which aims to flatten the total energy distribution of  $\nu_e$  and  $\nu_\mu$  events by increasing the weights of the low-energy and high-energy tails. However, since the  $\nu_e$

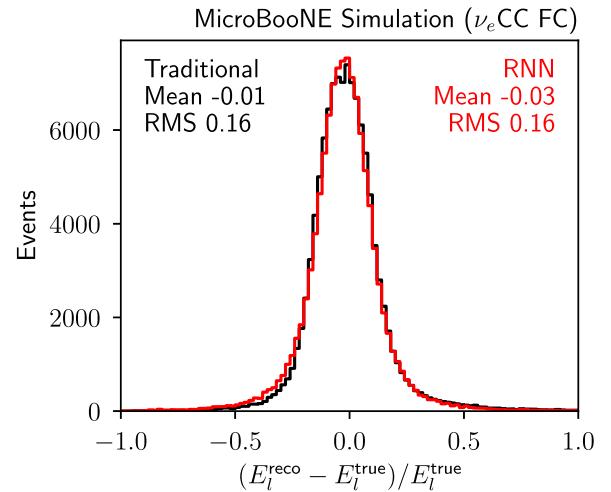


FIG. 15. Primary electron energy resolution histograms for the traditional MicroBooNE energy estimator (black) and the reweighted training of the RNN energy estimator (red) in the FC CC sample.

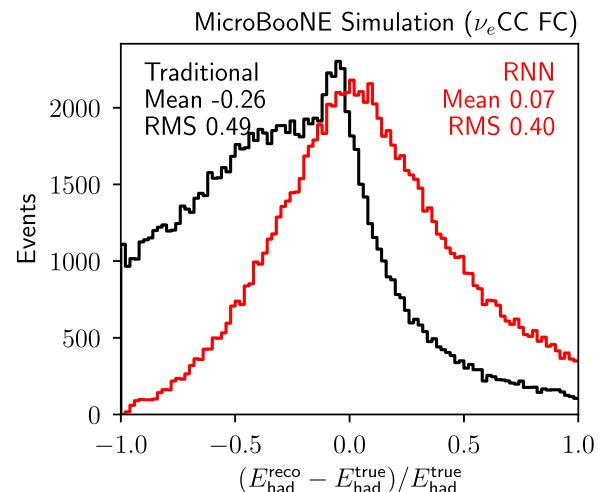


FIG. 16.  $\nu_e$ CC hadronic energy resolution histograms for the traditional MicroBooNE energy estimator (black) and the reweighted training of the RNN energy estimator (red) in the FC CC sample.

energy distribution is nearly flat to begin with, this reweighting disproportionately amplifies the importance of its tails, resulting in a U-shaped bias curve. Using separate weights for  $\nu_e$  and  $\nu_\mu$  datasets could improve performance by reducing interference.

Figures 15 and 16 show the energy resolution histograms of the primary lepton and hadronic energies, respectively. There is little difference in the resolution of the energy of the primary lepton. However, the RNN achieves superior hadronic energy reconstruction, with improvement in RMS from 49% to 40%.

## APPENDIX C: ENERGY ESTIMATION OF THE PARTIALLY CONTAINED EVENTS

In this appendix, we review the performance of the RNN energy estimator on the PC events and compare it to the performance of the traditional MicroBooNE energy estimator. Figure 17 compares neutrino energy resolution for the  $\nu_\mu$  CC and  $\nu_e$  CC events. For the PC events, the RNN energy estimator demonstrates improvement in the energy resolution from 43% to 28% for  $\nu_\mu$  CC, and improvement from 31% to 24% for  $\nu_e$  CC, compared to the traditional MicroBooNE energy estimator.

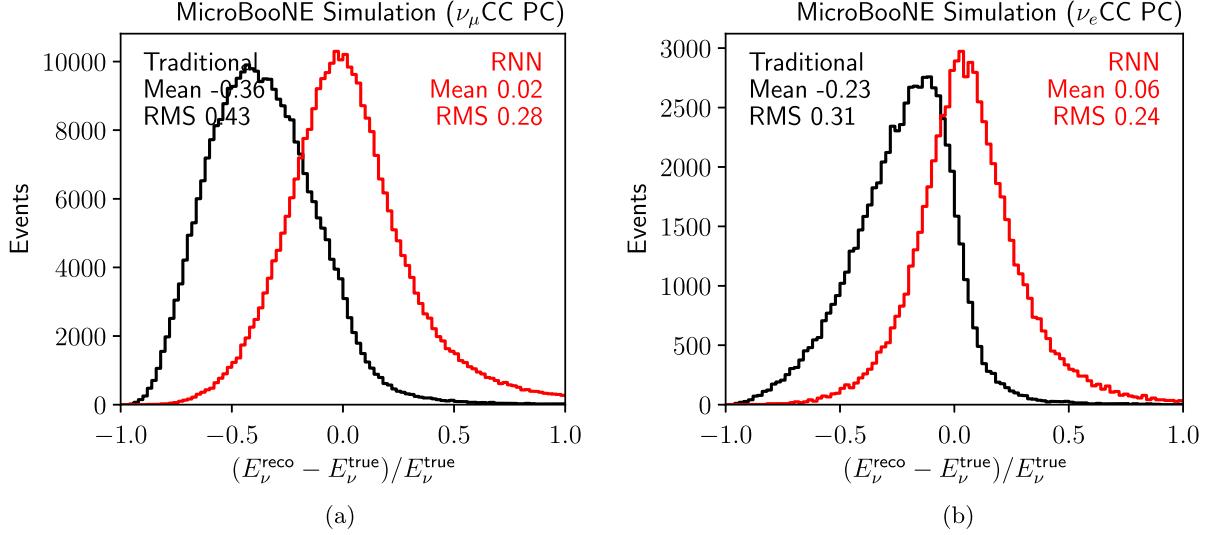


FIG. 17.  $\nu_\mu$  and  $\nu_e$  energy resolution histograms for the traditional MicroBooNE energy estimator (black) and the reweighted training of the RNN energy estimator (red), evaluated on the samples of PC CC events. (a)  $\nu_\mu$  CC energy resolution. (b)  $\nu_e$  CC energy resolution.

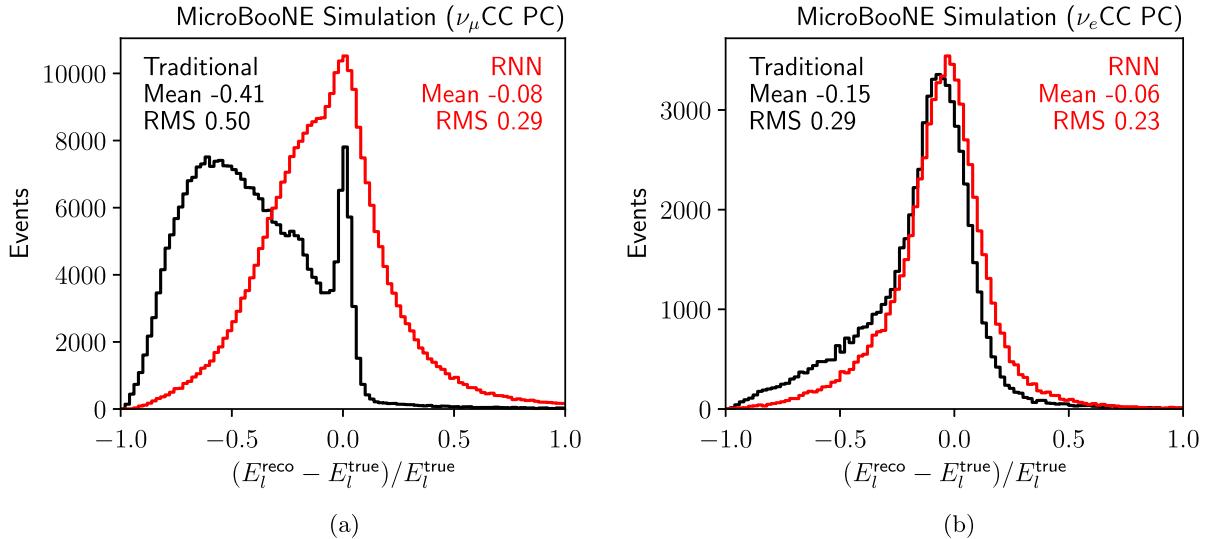


FIG. 18. Primary muon and electron energy resolution histograms for the traditional MicroBooNE energy estimator (black) and the reweighted training of the RNN energy estimator (red), evaluated on the samples of PC CC events. (a) Primary  $\mu$  energy resolution. (b) Primary  $e$  energy resolution.

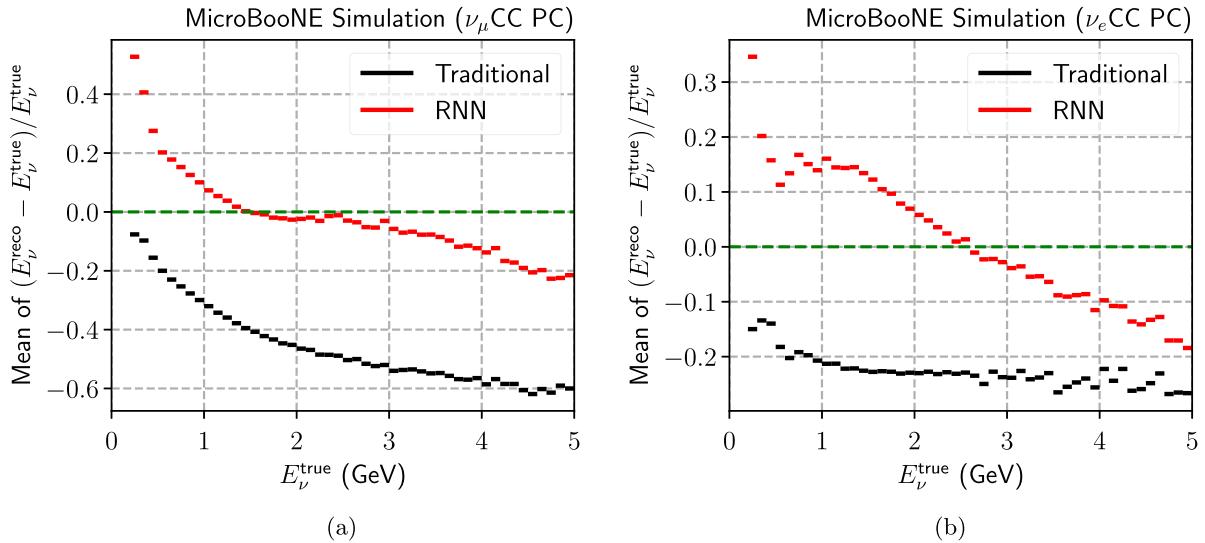


FIG. 19.  $\nu_\mu$  and  $\nu_e$  energy resolution bias plot for the traditional MicroBooNE energy estimator (black) and the reweighted training of the RNN energy estimator (red), evaluated on the samples of PC CC events. (a)  $\nu_\mu$  energy resolution bias. (b)  $\nu_e$  energy resolution bias.

Likewise, Fig. 18 compares energy resolutions of the primary leptons in the  $\nu_\mu$  and  $\nu_e$  CC events. It shows a large improvement in reconstructing the energy of the primary muons with the RMS improving from 50% for the traditional MicroBooNE energy estimator down to 29% for the RNN energy estimator.

The improvement for the energy reconstruction of the primary electron energy is much smaller, with the RMS improving from 29% down to 23%. Moreover, the majority of this improvement comes from the tail of the distribution, indicating that the traditional electron energy reconstruction approach is close to the optimal.

Finally, Fig. 19 shows the bias of the neutrino energy estimators as a function of the true neutrino energy. The  $\nu_\mu$  CC energy bias of the RNN energy estimator is smaller than the bias of the traditional MicroBooNE energy estimator for the majority of neutrino energies. It is, however, larger at small true  $E_\nu$  ( $E_\nu < 0.5$  GeV). This behavior is similar to the energy bias behavior for the FC events, observed in Sec. IV B. The energy reweighting reduces the low energy

bias for the FC events, but evidently, it is not sufficient to remove the bias for the PC events. For the  $\nu_e$  energy estimation, Fig. 19 shows a different behavior. The RNN energy estimator has a smaller bias for virtually the entire energy range, compared to the traditional MicroBooNE energy estimator. However, it also exhibits large oscillations in the low energy region.

#### APPENDIX D: RESOLUTION DEPENDENCE ON ENERGY

In this appendix, we review the binned statistics plots, depicting energy resolution versus true neutrino energy bins. Figure 20 illustrates the behavior of the RMS of the energy resolution as a function of the true neutrino energy. For the FC events, the RNN energy estimator has better neutrino energy resolution compared to the traditional MicroBooNE energy estimator across the entire energy range. For the PC events, however, the RNN loses its performance in the low-energy region.

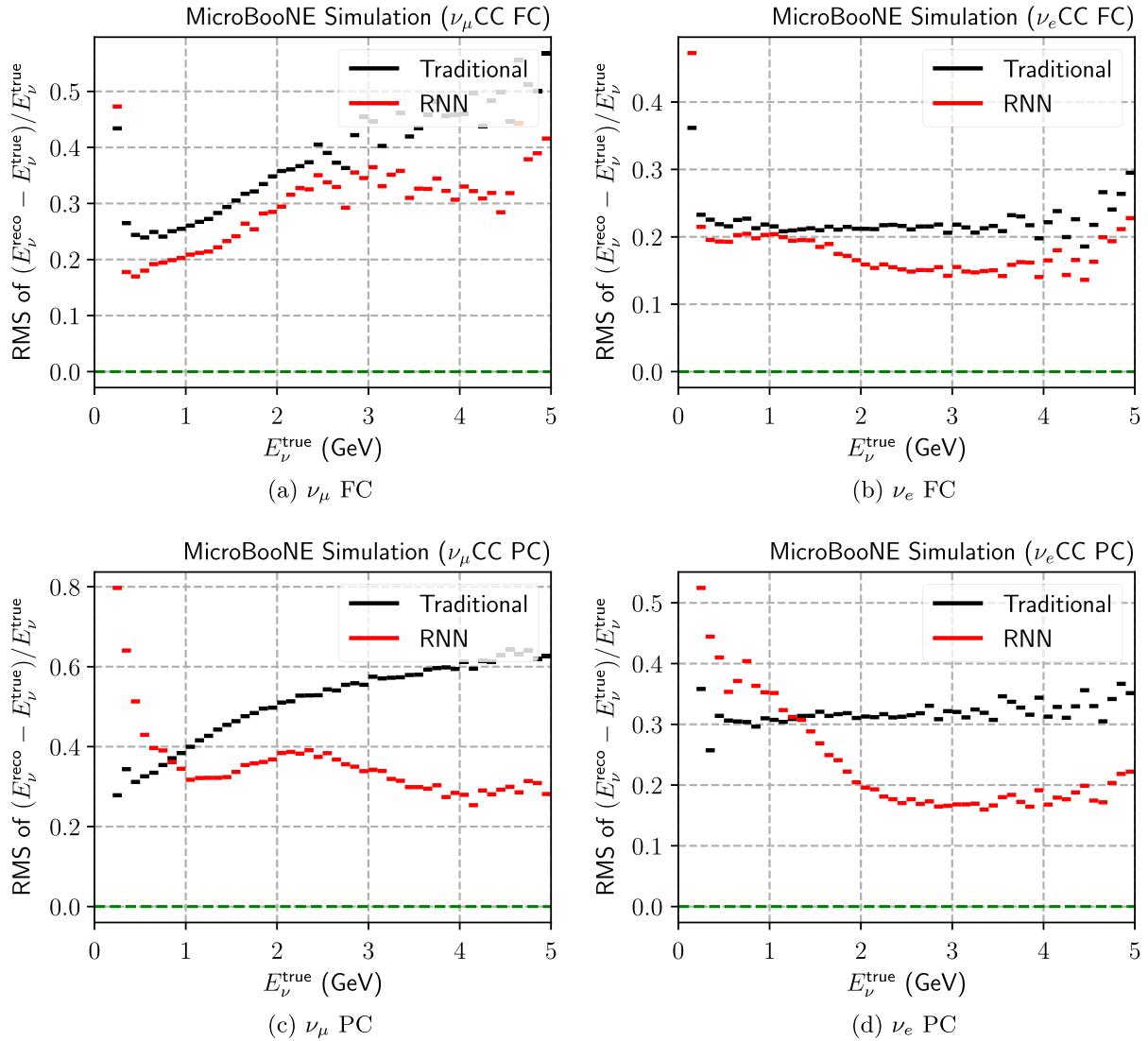


FIG. 20.  $\nu_\mu$  and  $\nu_e$  energy resolution vs true neutrino energy plots for the traditional MicroBooNE energy estimator (black) and the reweighted training of the RNN energy estimator (red), evaluated on the samples of FC and PC events. (a)  $\nu_\mu$  FC. (b)  $\nu_e$  FC. (c)  $\nu_\mu$  PC. (d)  $\nu_e$  PC.

- [1] P. F. de Salas, D. V. Forero, S. Gariazzo, P. Martínez-Miravé, O. Mena, C. A. Ternes, M. Tórtola, and J. W. F. Valle, *J. High Energy Phys.* **02** (2021) 071.
- [2] S. Pascoli, S. T. Petcov, and A. Riotto, *Phys. Rev. D* **75**, 083511 (2007).
- [3] J. Kopp, P. A. N. Machado, M. Maltoni, and T. Schwetz, *J. High Energy Phys.* **05** (2013) 050.
- [4] M. A. Acero *et al.* (NOvA Collaboration), *Phys. Rev. D* **106**, 032004 (2022).
- [5] K. Abe *et al.* (T2K Collaboration), *Phys. Rev. D* **103**, L011101 (2021).
- [6] P. Abratenko *et al.* (MicroBooNE Collaboration), *Phys. Rev. Lett.* **130**, 011801 (2023).
- [7] B. Abi *et al.* (DUNE Collaboration), *Eur. Phys. J. C* **80**, 978 (2020).
- [8] K. S. McFarland (MINERvA Collaboration), *AIP Conf. Proc.* **1405**, 95 (2011).
- [9] <https://www-nds.iaea.org/relnsd/vcharthtml/VChartHTML.html>.
- [10] H. Gallagher, G. Garvey, and G. P. Zeller, *Annu. Rev. Nucl. Part. Sci.* **61**, 355 (2011).
- [11] U. Mosel, *Annu. Rev. Nucl. Part. Sci.* **66**, 171 (2016).
- [12] T. Golan, C. Juszczak, and J. T. Sobczyk, *Phys. Rev. C* **86**, 015505 (2012).
- [13] R. Acciari *et al.* (MicroBooNE Collaboration), *J. Instrum.* **12**, P02017.

[14] P. Abratenko *et al.* (MicroBooNE Collaboration), *Phys. Rev. D* **108**, 052010 (2023).

[15] P. Abratenko *et al.* (MicroBooNE Collaboration), *Phys. Rev. D* **105**, 112005 (2022).

[16] Y. LeCun, Y. Bengio, and G. Hinton, *Nature (London)* **521**, 436 (2015).

[17] D. Guest, K. Cranmer, and D. Whiteson, *Annu. Rev. Nucl. Part. Sci.* **68**, 161 (2018).

[18] A. Radovic, M. Williams, D. Rousseau, M. Kagan, D. Bonacorsi, A. Himmel, A. Aurisano, K. Terao, and T. Wongjirad, *Nature (London)* **560**, 41 (2018).

[19] H. W. Yu *et al.*, *J. Instrum.* **16**, P01036 (2020).

[20] A. Aurisano, A. Radovic, D. Rocco, A. Himmel, M. D. Messier, E. Niner, G. Pawloski, F. Psihas, A. Sousa, and P. Vahle, *J. Instrum.* **11**, P09001 (2016).

[21] C. Adams *et al.* (MicroBooNE Collaboration), *Phys. Rev. D* **99**, 092001 (2019).

[22] P. Abratenko *et al.* (MicroBooNE Collaboration), *J. Instrum.* **17**, P01037 (2021).

[23] P. Baldi, J. Bian, L. Hertel, and L. Li, *Phys. Rev. D* **99**, 012011 (2019).

[24] J. Liu *et al.* (DUNE Collaboration), arXiv:2012.06181.

[25] Y. Yu, X. Si, C. Hu, and J. Zhang, *Neural Comput.* **31**, 1235 (2019).

[26] P. Abratenko *et al.* (MicroBooNE Collaboration), *Phys. Rev. Lett.* **128**, 151801 (2022).

[27] A. A. Aguilar-Arevalo *et al.* (MiniBooNE Collaboration), *Phys. Rev. D* **79**, 072002 (2009).

[28] X. Qian, C. Zhang, B. Viren, and M. Diwan, *J. Instrum.* **13**, P05032 (2018).

[29] P. Abratenko *et al.* (MicroBooNE Collaboration), *Phys. Rev. Lett.* **128**, 241801 (2022).

[30] P. Abratenko *et al.* (MicroBooNE Collaboration), *Phys. Rev. D* **105**, 112004 (2022).

[31] P. Abratenko *et al.* (MicroBooNE Collaboration), *Phys. Rev. D* **105**, 112003 (2022).

[32] R. Acciari *et al.* (MicroBooNE Collaboration), *Eur. Phys. J. C* **78**, 82 (2018).

[33] P. Abratenko *et al.* (MicroBooNE Collaboration), *J. Instrum.* **16**, P06043 (2020).

[34] A. Farahani *et al.*, in *Advances in Data Science and Information Engineering: Proceedings from ICData 2020 and IKE 2020* (Springer, Cham, 2021), p. 877.

[35] K. Carloni, N. W. Kamp, A. Schneider, and J. M. Conrad, *J. Instrum.* **17**, P02022 (2022).

[36] Alexander Radovic (NO $\nu$ A Collaboration), <https://github.com/AlexanderRadovic/rnnNeutrinoEnergyEstimator>.

[37] D. Torbunov, Improving energy estimation at NO $\nu$ A with recurrent neural networks, Ph.D. thesis, University of Minnesota, 2021.

[38] C. Adams *et al.* (MicroBooNE Collaboration), *J. Instrum.* **13**, P07006 (2018).

[39] C. Adams *et al.* (MicroBooNE Collaboration), *J. Instrum.* **13**, P07007 (2018).

[40] S. d'Ascoli, H. Touvron, M. L. Leavitt, A. S. Morcos, G. Birol, and L. Sagun, in *International Conference on Machine Learning, Lille, France* (PMLR, 2021), pp. 2286–2296, <https://proceedings.mlr.press/v139/>.

[41] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, *Adv. Neural Inf. Process. Syst.* **30**, 5998 (2017), [https://papers.nips.cc/paper\\_files/paper/2017/hash/3f5ee243547dee91fb0d053c1c4a845aa-Abstract.html](https://papers.nips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fb0d053c1c4a845aa-Abstract.html).

[42] A. Dosovitskiy *et al.*, arXiv:2010.11929.

[43] P. Abratenko *et al.* (MicroBooNE Collaboration), *Phys. Rev. Appl.* **15**, 064071 (2021).

[44] G. Van Houdt, C. Mosquera, and G. Nápoles, *Artif. Intell. Rev.* **53**, 5929 (2020).

[45] A. Zhang, Z. C. Lipton, M. Li, and A. J. Smola, *Dive into Deep Learning* (Cambridge University Press, Cambridge, England, 2023).

[46] S. Ioffe and C. Szegedy, in *International Conference on Machine Learning, Lille, France* (PMLR, 2015), pp. 448–456, <https://proceedings.mlr.press/v37/>.

[47] C. Andreopoulos *et al.*, *Nucl. Instrum. Methods Phys. Res., Sect. A* **614**, 87 (2010).

[48] L. Alvarez-Ruso *et al.* (GENIE Collaboration), *Eur. Phys. J. Special Topics* **230**, 4449 (2021).

[49] J. Nieves, I. Ruiz Simo, and M. J. Vicente Vacas, *Phys. Lett. B* **707**, 72 (2012).

[50] J. Nieves, J. E. Amaro, and M. Valverde, *Phys. Rev. C* **70**, 055503 (2004); **72**, 019902(E) (2005).

[51] R. Gran, J. Nieves, F. Sanchez, and M. J. V. Vacas, *Phys. Rev. D* **88**, 113007 (2013).

[52] K. Abe *et al.* (T2K Collaboration), *Phys. Rev. D* **93**, 112012 (2016).

[53] P. Abratenko *et al.* (MicroBooNE Collaboration), *Phys. Rev. D* **105**, 072001 (2022).

[54] E. L. Snider and G. Petrillo, *J. Phys. Conf. Ser.* **898**, 042057 (2017).

[55] S. Agostinelli *et al.* (GEANT4 Collaboration), *Nucl. Instrum. Methods Phys. Res., Sect. A* **506**, 250 (2003).

[56] P. Abratenko *et al.* (MicroBooNE Collaboration), *J. Instrum.* **15**, P12037 (2020).

[57] C. Adams *et al.* (MicroBooNE Collaboration), *J. Instrum.* **15**, P07010 (2019).

[58] Y. Li *et al.* (MicroBooNE Collaboration), *Nucl. Instrum. Methods Phys. Res., Sect. A* **816**, 160 (2016).

[59] P. Abratenko, R. An, J. Anthony, L. Arellano, J. Asaadi, A. Ashkenazi, S. Balasubramanian, B. Baller, C. Barnes *et al.* (MicroBooNE Collaboration), *Phys. Rev. D* **105**, 112005 (2022).

[60] [https://pytorch.org/docs/stable/generated/torch.optim.lr\\_scheduler.ReduceLROnPlateau.html](https://pytorch.org/docs/stable/generated/torch.optim.lr_scheduler.ReduceLROnPlateau.html).

[61] L. Aliaga *et al.* (MINERvA Collaboration), *Phys. Rev. D* **94**, 092005 (2016); **95**, 039903(A) (2017).

[62] P. Abratenko *et al.* (MicroBooNE Collaboration), arXiv: 2307.06413.

[63] M. L. Eaton, *Multivariate Statistics: A Vector Space Approach* (John Wiley and Sons, New York, 1983), p. 116.

[64] V. Hewes *et al.* (DUNE Collaboration), *Instruments* **5**, 31 (2021).

[65] P. A. Machado, O. Palamara, and D. W. Schmitz, *Annu. Rev. Nucl. Part. Sci.* **69**, 363 (2019).

[66] D. Torbunov, VLNE—Framework to develop energy estimators for HEP experiments <https://github.com/usert5432/vlne> (2023).