Neutrino type identification for atmospheric neutrinos in a large homogeneous liquid scintillation detector

Jiaxi Liu,[1,2,*] Fanrui Zeng,[3,*] Hongyue Duyang,[3,†] Wanlei Guo,[1] Xinhai He,[1] Teng Li,[3,‡] Zhen Liu,[1,§] Wuming Luo,[1,‖] Wing Yan Ma,[3,¶] Xiaohan Tan,[3] Liangjian Wen,[1] Zekun Yang,[3] and Yongpeng Zhang[1,**]

[1]*Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, China*
[2]*School of Physical Sciences, University of Chinese Academy of Science, Beijing 100049, China*
[3]*Shandong University, Jinan, China and Key Laboratory of Particle Physics and Particle Irradiation of Ministry of Education, Shandong University, Qingdao 266237, China*

(Received 29 March 2025; accepted 20 June 2025; published 22 July 2025)

Atmospheric neutrino oscillations are important to the study of neutrino properties, including the neutrino mass ordering problem. A good capability to identify neutrinos' flavor and neutrinos against antineutrinos is crucial in such measurements. In this paper, we present a machine-learning-based approach for identifying atmospheric neutrino events in a large homogeneous liquid scintillator detector. This method identifies features of photomultiplier tube waveforms that reflect event topologies and uses them as input to machine learning models. In addition, neutron-capture information is utilized to achieve neutrino vs antineutrino discrimination. Preliminary performances based on Monte Carlo simulations are presented, which demonstrate such a detector's potential in future measurements of atmospheric neutrinos such as the one planned for the JUNO experiment.

DOI: 10.1103/fznt-z257

I. INTRODUCTION

Atmospheric neutrinos constitute an important natural neutrino source for investigating fundamental neutrino properties. Produced by cosmic rays interacting with Earth's atmosphere, atmospheric neutrinos span wide energy spectra and oscillation baseline lengths, making them well suited for studying neutrino oscillations. The most notable achievement made by atmospheric neutrino measurements is the discovery of neutrino oscillations by the Kamiokande [1] and Super-Kamiokande experiments [2]. One of the biggest questions in neutrino physics that future atmospheric neutrino measurements could solve is the sign of Δm^2_{31} and Δm^2_{32}, commonly known as the neutrino mass ordering (NMO) (or neutrino mass hierarchy) problem. Matter effects, due to charged-current interactions between neutrinos and electrons when atmospheric neutrinos pass through Earth, could help experiments solve the problem by amplifying the differences in oscillation probabilities between normal ordering and inverted ordering [3,4]. The interpretations of such measurements, however, are also affected by uncertainties such as the unknown value of the *CP*-violating phase (δ_{CP}) and the θ_{23} octant ambiguity. The next-generation Cherenkov neutrino experiments, such as Hyper-Kamiokande [5,6], KM3Net/ORCA [7,8], and IceCube Upgrade [9], capable of measuring atmospheric neutrino oscillations and solving the NMO problem with gigantic water detectors, are currently under construction.

A different method that can solve the NMO problem is the measurement of medium baseline (∼50 km) reactor antineutrino oscillation in vacuum or near-vacuum conditions. This method has the advantage that it does not depend on δ_{CP} and the θ_{23} octant. The only experiment on the horizon with this method is the Jiangmen Underground Neutrino Observatory (JUNO), currently under construction in China. It is designed to determine NMO with a 20 kton homogeneous liquid scintillator (LS) detector by measuring reactor electron antineutrino ($\bar{\nu}_e$) oscillations [10,11]. It is also worth noting that the two methods are complementary, and the synergy between them can greatly enhance the total sensitivity in a global fit [12–14]. Therefore, if atmospheric (anti) neutrino oscillations can also be measured in a large homogeneous LS detector like JUNO itself, its NMO sensitivity can be maximized by a joint analysis of

*J. L. and F. Z. contributed equally to this work.
†Contact author: duyang@sdu.edu.cn
‡Contact author: tengli@sdu.edu.cn
§Contact author: liuzhen@ihep.ac.cn
‖Contact author: luowm@ihep.ac.cn
¶Contact author: wingyanma@sdu.edu.cn
**Contact author: ypzhang1991@ihep.ac.cn

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

both reactor and atmospheric neutrino oscillations in the same detector.

Atmospheric neutrino oscillation measurements in an LS detector, however, are very challenging. While such detectors has been utilized by experiments such as KamLAND [15], Borexino [16,17], Daya Bay [18,19], RENO [20,21], and Double Chooz [22,23] for reactor and solar neutrino measurements, they have never been used for atmospheric neutrinos. This is mostly due to two difficulties. First, LS detectors are traditionally believed to lack directional measurement capability, which is mandatory for atmospheric neutrino oscillation measurements to determine the neutrino oscillation baseline length. Second, the atmospheric neutrino flux is a mixture of muon (anti)neutrino ($\nu_\mu/\bar{\nu}_\mu$) and electron (anti)neutrinos ($\nu_e/\bar{\nu}_e$) and hence the neutrino flavor needs to be identified in the detector to determine the oscillation probability. Moreover, the ability to identify neutrinos against antineutrinos is important since matter effects distort their oscillation probability in opposite ways. While experiments lacking such capability still get some sensitivity to NMO from small differences in fluxes and cross sections for neutrinos and antineutrinos, the discriminating power between the neutrinos and antineutrinos certainly helps to disentangle NMO and *CP* violation phase and increase the sensitivity. However, such flavor identification and neutrino/antineutrino discrimination have never been achieved in LS detectors before.

Neutrino flavors are reflected in the flavors of outgoing charged leptons in neutrino-nucleus charged-current (CC) interactions. Therefore, a detector's capability of neutrino flavor identification largely relies on the identification of charged leptons, i.e., $e^{\mp}$ in $\nu_e/\bar{\nu}_e$-CC, and $\mu^{\mp}$ in $\nu_\mu/\bar{\nu}_\mu$-CC interactions. In a water Cherenkov detector, this task is achieved by identifying the sharpness of the Cherenkov ring. However, this approach is not feasible in an LS detector, where Cherenkov light is typically about 2 orders of magnitude weaker than scintillation light. Neutrino vs antineutrino identification is even more challenging. The most prominent difference between neutrino and antineutrino interactions is the charge of outgoing leptons in CC interactions. However, it is impractical to magnetize a modern neutrino detector with a kilotons-scale mass to identify the particle charge. Other handles may help in this task, including the differences in event kinematics and neutron multiplicity in the (anti)neutrino interactions. One recent example in a water Cherenkov detector is the atmospheric neutrino oscillation measurement by Super-Kamiokande with neutron tagging [24]. Super-Kamiokande with gadolinium loading is expected to improve further the neutrino/antineutrino discrimination with better neutron-tagging efficiency [25,26]. Additionally, event inelasticity, defined as the fraction of energy transferred from the incoming neutrino to the hadrons and exhibiting different shapes between neutrinos and antineutrinos, is also proposed to be used in IceCube Upgrade and KM3NET/ORCA to separate neutrinos/antineutrinos and improve the oscillation sensitivity to NMO [27,28]. LS detectors, on the other hand, are known to have excellent neutron-tagging efficiency, but event kinematics such as inelasticity measurements are very difficult with traditional reconstruction methods.

Recently, we reported a novel method of directionality measurement in an LS detector that solves the difficulty of directionality measurement in utilizing an LS detector for atmospheric neutrino oscillation measurements [29]. This method extracts features relevant to the event topology in the detector from photomultiplier tube (PMT) waveforms and uses them as input to machine learning (ML) models trained to find the event directionality. In this paper, we further explore the possibility of flavor identification in an LS detector such as JUNO for atmospheric neutrinos with a methodology based on a similar ML principle. A statistical neutrino vs antineutrinos identification is also demonstrated with additional help from neutron-capture information.

This paper is organized as follows. Section II describes the detector and simulation used in the work. Section III A presents the methodology of 3-label identification, i.e., $\nu_\mu/\bar{\nu}_\mu$-CC vs $\nu_e/\bar{\nu}_e$-CC vs neutral-current (NC) identification, and Sec. III B presents neutrino vs antineutrino identification. The identification performance is presented in Sec. IV, and more details, including the performance's dependence on event generators, score cut optimization, and an alternative classification strategy, are further discussed in Sec. V. Finally, Sec. VI summarizes the study.

II. DETECTOR AND SIMULATION

The JUNO central detector (CD) design and simulation [30,31] are used to demonstrate the method. A schematic view of the detector is shown in Fig. 1. It contains 20 kton LS in an acrylic sphere with a radius of 17.7 m. Installed facing inward are 17,612 (20 in.) PMTs [32] to collect light produced by charged particles with a total PMT coverage of 75%. In addition, 25,600 PMTs (3 in.) [33] provide additional 3% coverage. For simplicity, only 20 in.

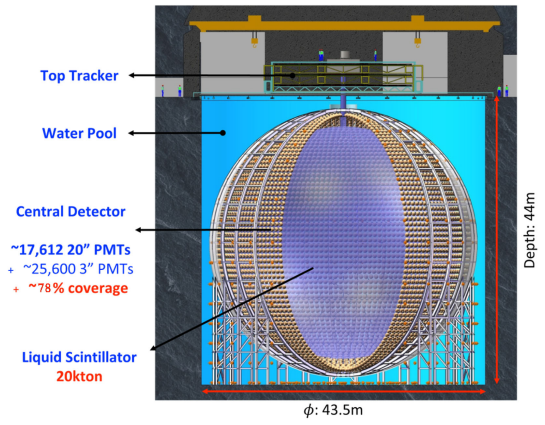


FIG. 1. Drawing of the JUNO detector design.

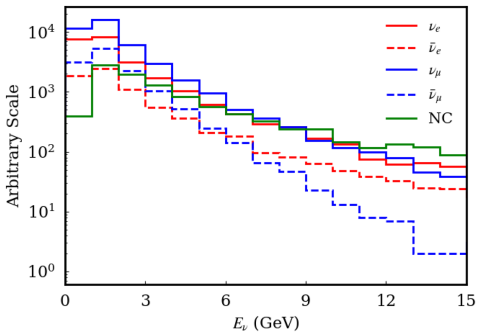


FIG. 2. The neutrino energy distribution of the Honda-flux [40] sample used for performance evaluation in this work. The sample is divided into five classes: $\nu_\mu/\bar{\nu}_\mu/\nu_e/\bar{\nu}_e$-CC and NC, represented by different colors.

PMTs are used in this work. The LS consists of linear alkylbenzene as the detection medium, 2.5 g/L 2,5-diphenyloxazole as the fluor and 3 mg/L p-bis- (o-methylstyryl) benzene as the wavelength shifter [34]. More details about the JUNO detector can be found in Ref. [10].

Neutrino interactions in the detector are simulated mainly by the GENIE event generator (v3.2.0) [35,36], and the NuWro event generator [37] is used as an independent check. The detector response is simulated based on Geant4 [38,39]. Various electronic effects, including PMT dark noise, time transition spread, PMT charge smearing, electronics baseline fluctuation, and PMT response for single photoelectrons, are also implemented. The detailed PMT simulation parameters are the same as in Ref. [29].

Three independent neutrino interaction samples are produced to train and validate the ML models, each including five categories of events: ν_μ-CC, $\bar{\nu}_\mu$-CC, ν_e-CC, $\bar{\nu}_e$-CC, and NC. The first sample is simulated with GENIE and has roughly flat visible energy distributions and similar statistics (about 70 k) for each category to avoid introducing any energy dependence in the model training (referred to as the flat sample). The second is simulated with GENIE and a more realistic flux is taken from the calculation of Honda *et al.* [40] for the JUNO site without the overburden mountain (referred to as the Honda-flux sample). This sample is used for performance evaluation with total statistics of about 95 k. The neutrino energy spectrum of the Honda-flux sample is shown in Fig. 2. The last one is simulated with NuWro with similar statistics to the Honda-flux sample to test the robustness of the performance with a different generator (referred to as the NuWro sample). For each sample, events with visible energy between 0.5 and 15 GeV are selected for this study, which covers the most sensitive energy region of atmospheric neutrinos to NMO [10].

III. METHODOLOGY

Neutrino-nucleus interactions in an LS detector are complicated in the few-GeV energy region. In a typical $\nu_\mu/\bar{\nu}_\mu$-CC event, the $\mu^\mp$ from the primary interaction deposits its energy quickly and leaves a long track in LS. Various hadrons such as n, p, π^0, and $\pi^\pm$ can also be produced through primary and secondary interactions. Most charged hadrons and π^0's produce showers while rapidly depositing their energies in LS, with the exception that $\pi^\pm$'s behave similarly to $\mu^\pm$'s. For a $\nu_e/\bar{\nu}_e$-CC event, the event topology in the LS is similar to that of $\nu_\mu/\bar{\nu}_\mu$-CC except that $e^\mp$ from the primary interaction also produces a shower. The primary charged lepton and hadrons together trigger the detector promptly (the prompt trigger). For an NC event, the hadron part is similar to a CC event, yet the outgoing neutrino is invisible.

Tracks and showers propagate through LS and produce scintillation light. Scintillation light seen by a PMT is the superposition of light from many points along the track or shower. The time evolution of scintillation light seen by PMTs is therefore determined by the event topology in the detector. Figure 3 shows the PE time distributions $[n_{PE}(t)]$

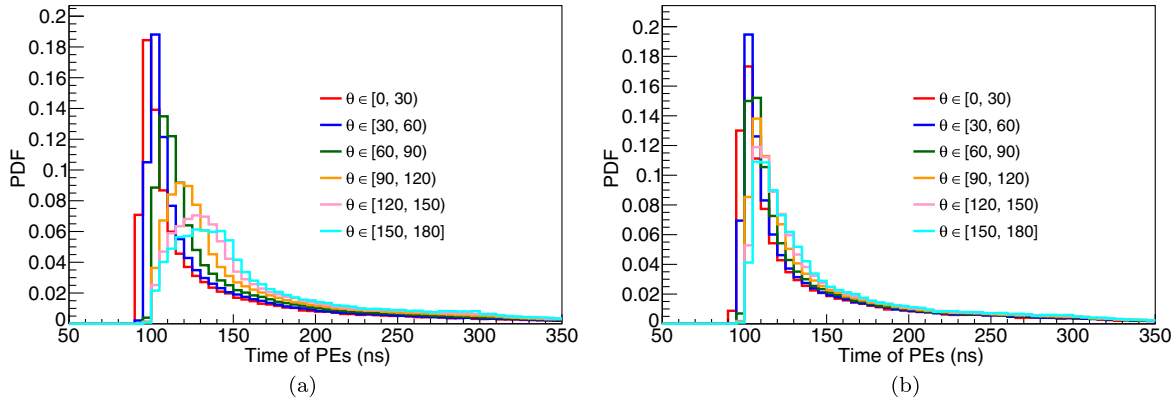


FIG. 3. The probability density function (PDF) of PE time for PMTs at difference angles (θ, in degrees) to (a) a muon and (b) an electron with the same kinetic energy (1 GeV), vertex position (at the center of the detector), and initial direction (along the positive x axis). The PMT angle θ is the intermediate angle between the particle direction and the line connecting the middle point of the particle track and the PMT. Distinct shapes can be observed for the two types of particles.

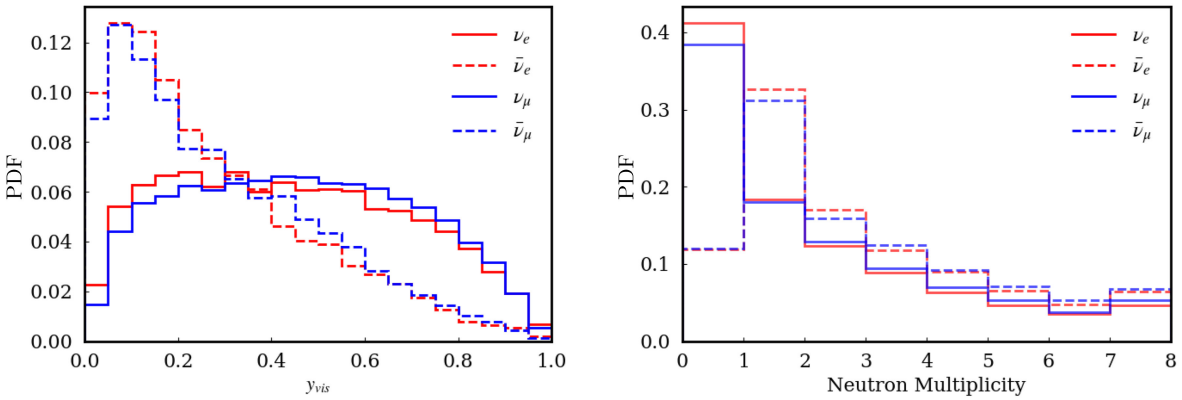


FIG. 4. y_{vis} (left) and neutron multiplicity (right) distributions for $\nu_\mu/\bar{\nu}_\mu/\nu_e/\bar{\nu}_e$-CC interactions in the Honda-flux sample. Full detector simulation is included. The distributions are normalized for a shape-only comparison.

of PMTs at different angles with respect to an electron and a muon simulated at the detector center with the same kinetic energy (1 GeV) and direction. It is observed that the two particles exhibit distinct $n_{PE}(t)$ shapes. This is mostly because electron showers are generally much shorter than muon tracks at the same energy in the LS. The differences in the $n_{PE}(t)$ distribution are then reflected in the PMT waveforms. Therefore, it is possible to use the information from PMT waveforms of the prompt trigger to identify different event topologies and distinguish the $\nu_\mu/\bar{\nu}_\mu$-CC, $\nu_e/\bar{\nu}_e$-CC, and NC atmospheric neutrino events.

While it is impossible to discriminate neutrino and antineutrino interactions event by event without identifying the charge of the outgoing lepton, the task can be done statistically with help from several handles. The first one is their differences in event kinematics. In CC interactions, neutrinos tend to transfer more energy to hadrons than antineutrinos with the same energy due to the V-A structure of $\nu/\bar{\nu}$ interactions. This leads to a larger visible hadronic energy fraction for neutrino interactions (Fig. 4 left), defined as $y_{\text{vis}} = E_{\text{had,vis}}/E_{\text{vis}}$, where $E_{\text{had,vis}}$ is the visible energy caused by hadrons and E_{vis} is the total visible energy of the event. One advantage of LS detectors over water Cherenkov detectors is that the hadrons produced from atmospheric neutrinos are also mostly visible. To help illustrate this point, Fig. 5 shows an example of a simulated neutrino interaction with explicit separation between the PMT charges induced by the final-state lepton and hadrons. Since both the charged leptons and most hadrons in the final state of CC events deposit their energy within the prompt trigger, the y_{vis} information is also reflected in the event topology and finally also in the trigger's PMT waveforms.

The second handle is the neutron information. Neutrons can be produced both in the primary neutrino-nucleus interactions and also in secondary interactions. Antineutrino interactions generally tend to produce more primary neutrons since charge conservation in an antineutrino CC interaction requires a negative charge being transferred from the lepton to the hadrons and is more likely to convert a proton into a neutron. In contrast, a positive charge is transferred in a neutrino CC interaction and is more likely to convert a neutron into a proton. The number and spacial distribution of secondary neutrons are also related to the types and kinematics of the primary particles. Neutron multiplicities (including both primary and secondary neutrons) for different neutrino CC interactions are shown in Fig. 4 (right). The neutrons produced dissipate their kinetic energy and are eventually captured mostly by hydrogen nuclei with a time constant of about 200 μ's. The average distance from the primary neutron-capture position to the neutrino interaction vertex is about 0.5 m in LS according to simulation, and the capture positions for secondary neutrons

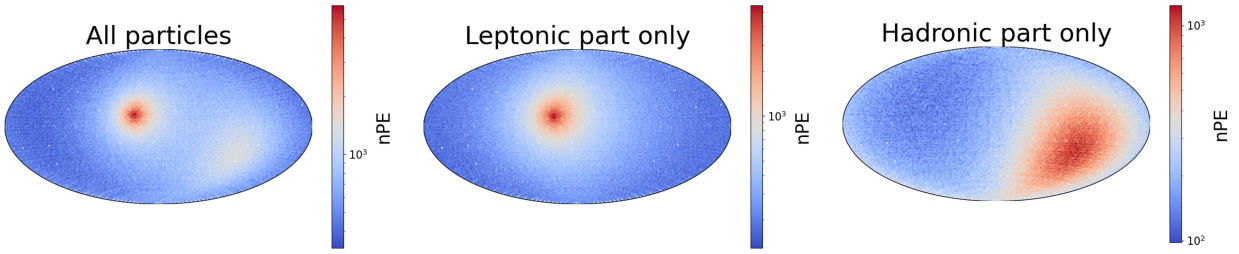


FIG. 5. Event display pictures of a simulated ν_e-CC interaction showing the PMT charges (nPE) from all particles (left), leptonic part only (middle), and hadronic part only (right). The PMT charge values are represented by colors and are projected onto a planer surface following the Mollweide projection method.

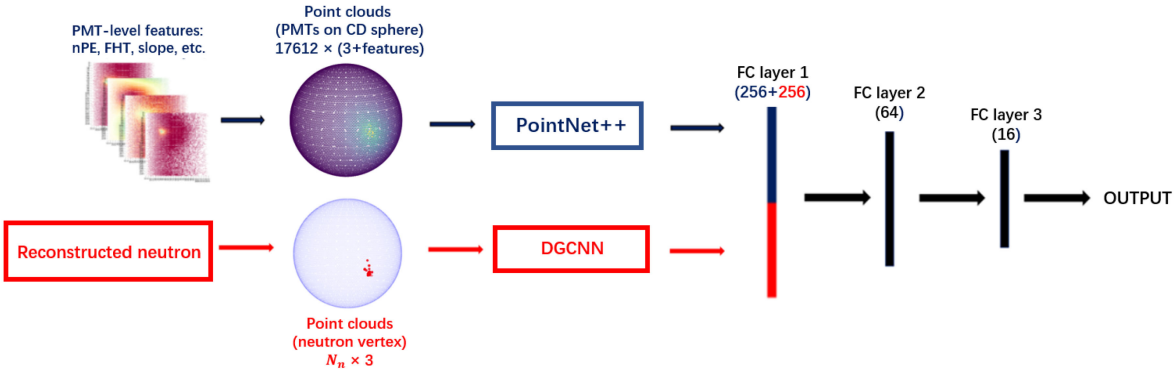


FIG. 6. Schematic diagram of the model architecture used by $\nu/\bar{\nu}$ identification strategy 1. The model consumes two types of inputs: PMT features from the prompt trigger processed by PointNet++ and neutron-capture vertex cloud processed by DGCNN. The outputs of PointNet++ and DGCNN are input to fully connected (FC) layers to predict the final scores representing the probability of the event being caused by neutrinos or antineutrinos.

are much more dispersed. 2.2 MeV γ's are then released and create delayed triggers. Another advantage of LS detectors such as JUNO is the excellent efficiency ($> 90\%$) in tagging such neutron captures compared to Water Cherenkov detectors. The information from delayed triggers caused by neutron captures therefore also helps to distinguish neutrinos from antineutrinos. More details on how to utilize the delayed trigger information will be discussed in Sec. III B.

Based on the discussion above, the identification task of atmospheric neutrino interactions is broken down into two steps: first, to identify the μ-like ($\nu_\mu/\bar{\nu}_\mu$-CC), e-like ($\nu_e/\bar{\nu}_e$-CC), and NC-like events against each other (referred to as the 3-label identification), and second to discriminate ν-like from $\bar{\nu}$-like events (referred to as the 2-label identification). The details of each step are described in the following subsections. Section V also briefly discusses other possible strategies.

A. 3-label flavor identification

In this step, atmospheric neutrino events are classified as μ-, e-, or NC-like using information from the 20 in. PMTs' waveforms in the prompt triggers. Features are extracted from PMT waveforms with a sampling rate of 1 GHz in a 1000 ns trigger window as inputs to ML models. These features include the hit time of the earliest photon (first hit time, FHT), the total charge, the slope of the waveform in the first 4 ns, the charge ratio in the first 4 ns to the total, the time and amplitude of the waveform's maximum bin, and other features such as median time and four moments (mean, standard deviation, skewness, kurtosis) of the waveform which describe the reconstructed waveforms in a more subtle way. A deconvolution and noise-reduction algorithm [41,42] is applied to the waveform before the feature extraction to improve the feature quality.

The PMT features from all 20 in. PMTs of the JUNO CD form spherical imagelike data. Two different ML models are used to deal with these kinds of data: the DeepSphere [43] model, which processes spherical data directly and the PointNet++ [44] model, which inputs data as a three-dimensional (3D) point cloud. The model architectures remain mostly the same except that the activation functions of the last layer are replaced with the softmax function to output three scores representing the probability of the event being μ-, e-, or NC-like, respectively. The sum of these three scores is 1 and an event is classified to the category with the largest score by default.

B. 2-label $\nu/\bar{\nu}$ identification

The identification of $\nu/\bar{\nu}$ interactions relies on their differences in event kinematics, neutron multiplicity, and the spatial distribution of neutron-capture vertices. The event kinematics information contained in the PMT waveforms of the prompt trigger is also reflected in the waveform features described in the previous subsection, including the y_{vis} information discussed in Sec. III. The neutron information, on the other hand, is reflected in the delayed triggers that are selected with energy between 2 and 2.7 MeV and the delay time between 10 μs and 1 ms. Therefore, the information from the prompt and delayed triggers needs to be combined for the task. Two different strategies are developed and described here. For both strategies, the final output is a 2-label classifier, with two scores representing the probability of an event being ν- or $\bar{\nu}$-like, respectively. The models are then trained separately for the ν_μ vs $\bar{\nu}_\mu$ and ν_e vs $\bar{\nu}_e$ identification.

1. Strategy 1

In this strategy, neutron-capture vertices are first reconstructed using the method from Ref. [45]. The reconstructed vertices from multiple selected delayed triggers form a new 3D point cloud that is input to the ML model together with the one formed by PMT waveform features from the prompt trigger. The overall structure of the ML model for strategy 1 is illustrated in Fig. 6. Similar to the

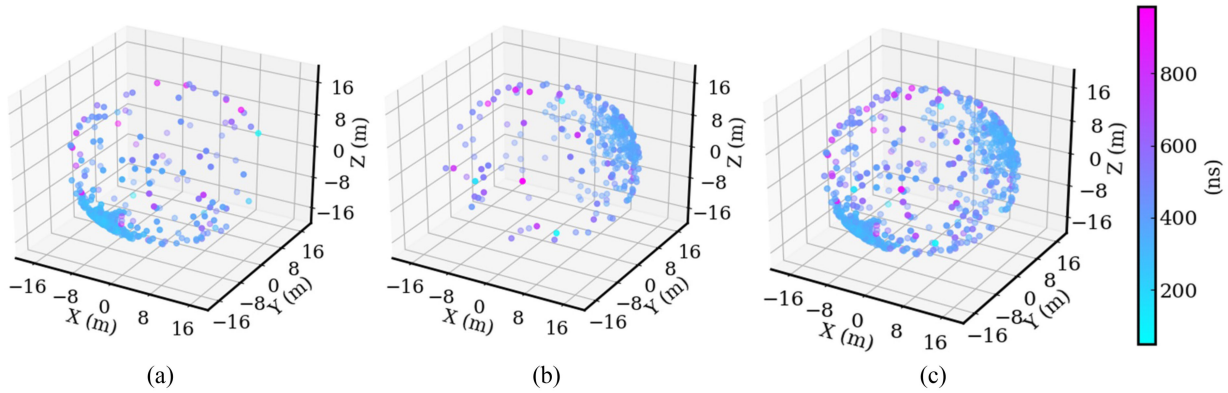


FIG. 7. Illustration of the trigger-merging technique used by strategy 2. (a),(b) Two delayed triggers selected as neutron-capture candidate triggers, and (c) shows the new merged one. Fired PMTs are shown as dots on the figures, with color representing FHT values of the hits.

3-label identification in step 1, the PMT waveform features from the prompt trigger are fed into a model based on PointNet + +. For the neutron-capture vertices, given that the number of neutrons is usually less than 100 for an atmospheric neutrino event, the size of the point-cloud data is fixed to 128 points, with zero padding if the number of selected neutron-capture candidates is less than 128. This point-cloud data of neutrons is fed into an individual dynamic graph convolutional neural-network-(DGCNN) based model [46]. Compared to traditional GCNN models, DGCNN is able to recover the neighborhood topology of point clouds with edge information, making it more suitable for describing spatial relationships of sparse point clouds. The DGCNN-based model output is then concatenated with the PointNet + +-based model output with a set of fully connected layers.

2. Strategy 2

Instead of reconstructing individual neutron-capture vertices, this strategy merges multiple delayed triggers into one by summing up the charges from multiple selected delayed triggers for each PMT and taking the FHT value as the earliest among them (Fig. 7). The two features (FHT and charge) are then used as extra features to the ML model together with features from the prompt trigger. This strategy does not rely on neutron-capture reconstruction algorithms and provides a fast and coherent way of utilizing features from both the prompt trigger and the delayed trigger. A DeepSphere-based model with the same architecture as used in the 3-label identification is used for the demonstration of this strategy.

IV. PERFORMANCE

The 3- and 2-label ML models are trained with the flat sample labeled by event categories. The trained models are then applied to the Honda-flux sample for performance evaluation in a more realistic situation. The efficiency, purity, and accuracy of event selection depend on the cuts on the output scores of ML models, which should be optimized by future analyses. To evaluate the model performances independent from the choices of selection cuts, the receiver operating characteristic (ROC) curve and the area under the ROC curve (AUC) are used [47]. They have the advantage of independence from the score cut choice and do not suffer from a class imbalance in the dataset if there is one. While the typical ROC curve and AUC are designed for binary classifications, the concepts are extended to an N-class classification in this work. For a given class i, the ROC curve and corresponding AUC score are obtained by converting the N-class classification to a "class i vs others" binary classification. The efficiency of class i is defined as the fraction of events in class i with the score associated with the class greater than a certain cut value. In contrast, the selection efficiency of "others" is calculated as the fraction of events with the score lower than the cut value. By varying the cut value from 0 to 1, one can calculate the corresponding two efficiencies above and subsequently draw the ROC curve of class i vs others, from which an AUC value is obtained. Finally, the total AUC score is defined as the arithmetic mean of all class i vs others binary classification scores. The performance of the 3- and 2-label models is presented in the following subsections separately.

A. 3-label flavor identification

Figure 8 shows the distributions of the μ- and e-like score output by the two 3-label ML models. (The NC-like scores are not shown since NC is a background to the oscillation analysis and the NC-like scores can be deduced from the μ- and e-like scores.) Clear separations are observed for the three event classes. ROC curves can be obtained from the score distributions, using the class i vs others method. Figure 9 shows the ROC curves and AUC

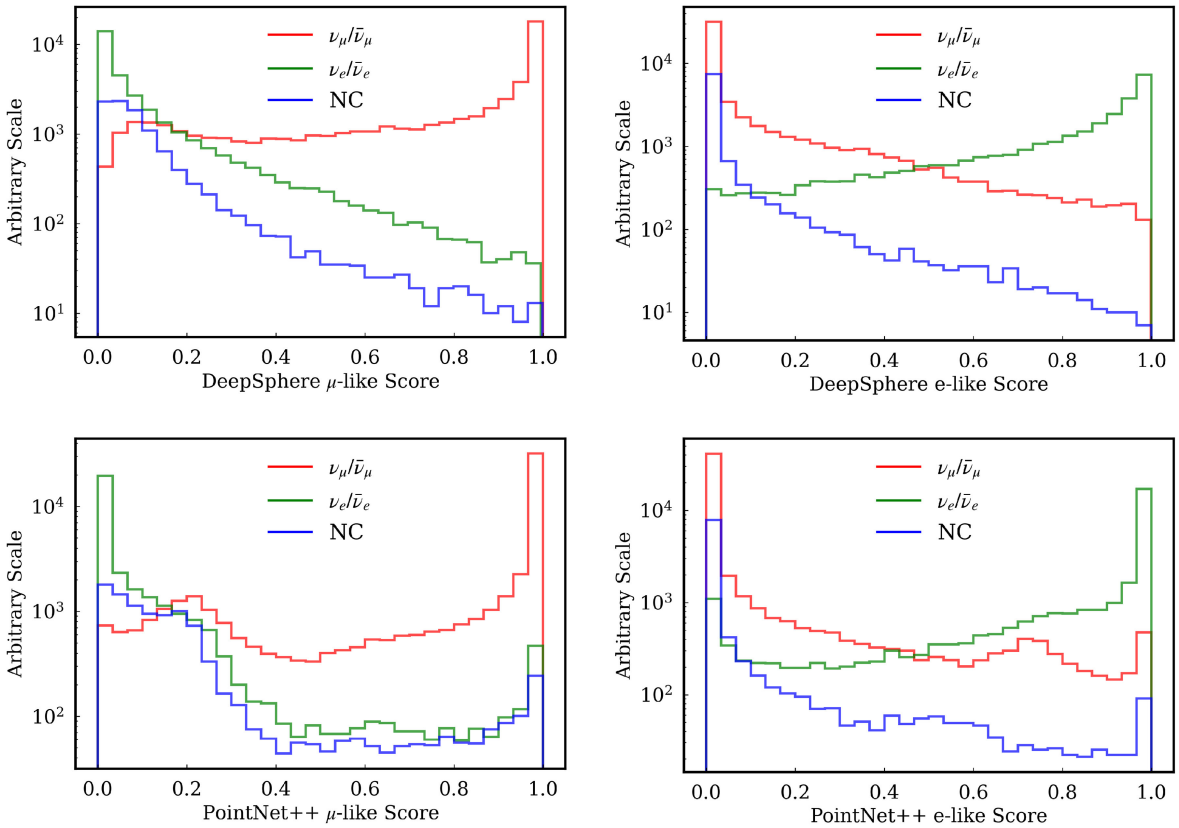


FIG. 8. Distributions of the μ-like score (left) and e-like score (right) by the DeepSphere (top) and PointNet + + (bottom) models. True $\nu_\mu/\bar{\nu}_\mu$-CC, $\nu_e/\bar{\nu}_e$-CC, and NC events are shown in different colors.

values for the two ML models using events across all energies. Such ROC curves are also calculated for events in different visible energy bins and the corresponding total AUC values are obtained by averaging over the AUC value of each class. Figure 10 compares the total AUC values as a function of visible energy for the two ML models. The AUC values are rather consistent, indicating that the performances of the two models are very close.

B. 2-label $\nu/\bar{\nu}$ identification

Similar to the 3-label identification, the performances of the $\nu/\bar{\nu}$ identification are evaluated in terms of ROC curve and AUC values using the Honda-flux sample. The ML model output score distributions are shown in Fig. 11, where distinct shapes are obtained for neutrinos and antineutrinos. The selection efficiencies of (anti)neutrinos for the ROC curve are obtained by selecting events with

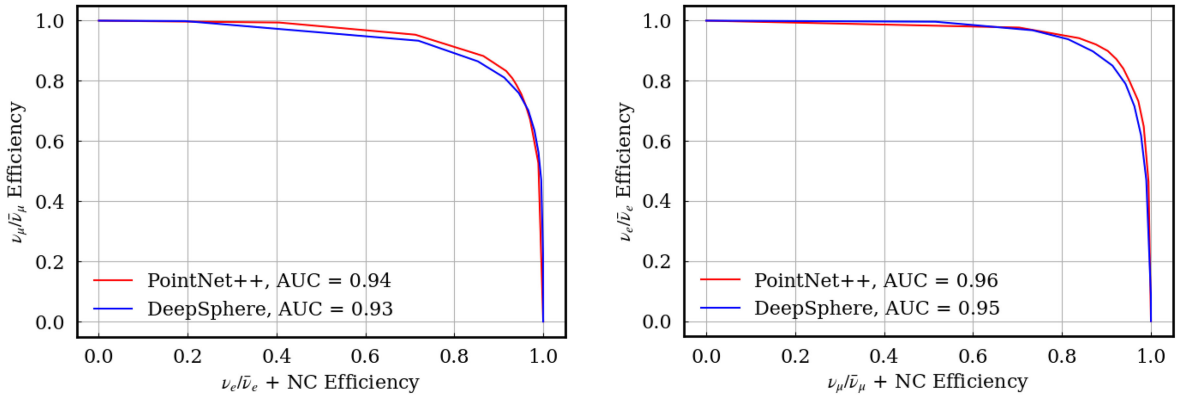


FIG. 9. ROC curves of the 3-label identification by the two ML models using events across all energies. For each ROC plot, the vertical axis shows the selection efficiency of one of the three classes, and the horizontal axis shows the efficiency of the other two combined. The corresponding AUC values are also listed.

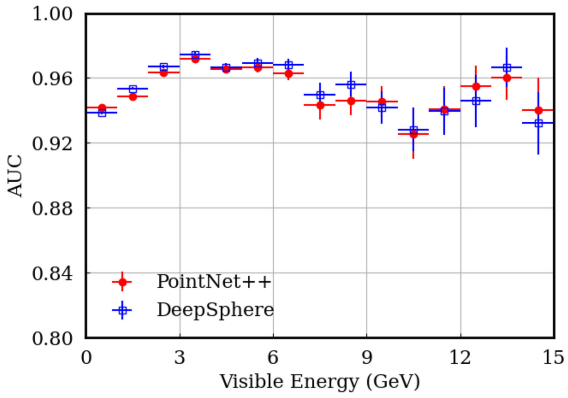


FIG. 10. The total AUC values of the 3-label identification as functions of visible energy for the two models. Error bars show the statistical uncertainties in each energy bin obtained with the bootstrap method [48].

scores above (below) various cut values. The ROC and AUC comparison of the two strategies is shown in Figs. 12 and 13. The two strategies give similar $\nu/\bar{\nu}$ identification performances. The remaining differences could be caused by the differences in the models and how the information of the captured neutron candidates is handled. In addition, the performance for $\nu_\mu/\bar{\nu}_\mu$ identification is better than that of $\nu_e/\bar{\nu}_e$. This could be mainly attributed to more (less) distinctive detector response of muons (electrons) with respect to hadrons in LS.

V. DISCUSSION

A. Dependence on event generators

In this work, event identification utilizes the event topology information reflected by the PMT features in the prompt trigger and the neutron-capture information from the delayed triggers. Both pieces of information rely on the simulation of the neutrino-nucleus interaction process, which could be different for various event generators. To check the dependence of the identification performance on the neutrino-nucleus interaction simulation, an independent sample was produced using the NuWro event generator, as mentioned in Sec. II. The models trained with the flat sample produced by GENIE are applied to this NuWro sample. Figure 14 compares the identification performance obtained from the GENIE and NuWro Honda-flux testing samples, using the PointNet++ model and strategy 1 of neutron-capture information handling as an example. Despite some small discrepancies, the performance is mostly consistent between the two samples. Similar results were obtained using the DeepSphere model and strategy 2. This additional check suggests that the ML models in this work do not suffer from apparent dependence on the event generator.

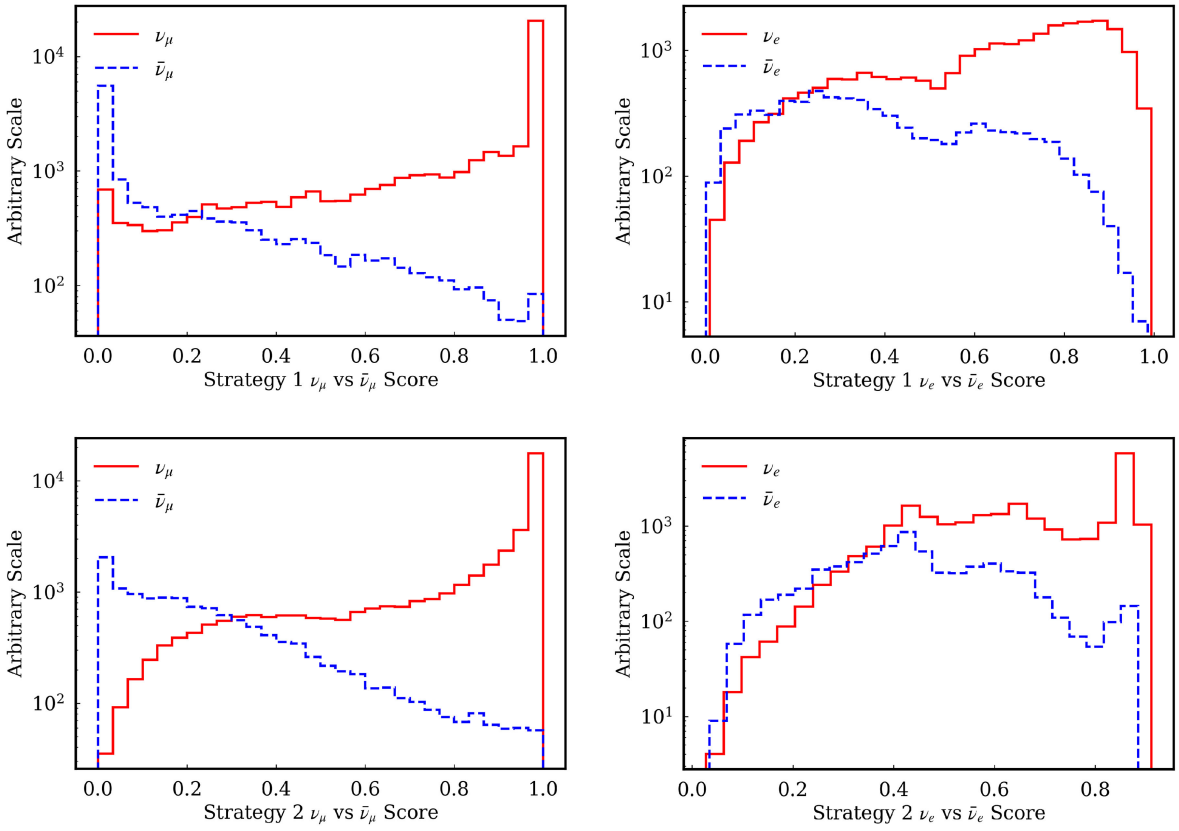


FIG. 11. The distributions of ν_μ vs $\bar{\nu}_\mu$ (left) score and ν_e vs $\bar{\nu}_e$ (right) score by strategy 1 (top) and 2 (bottom).

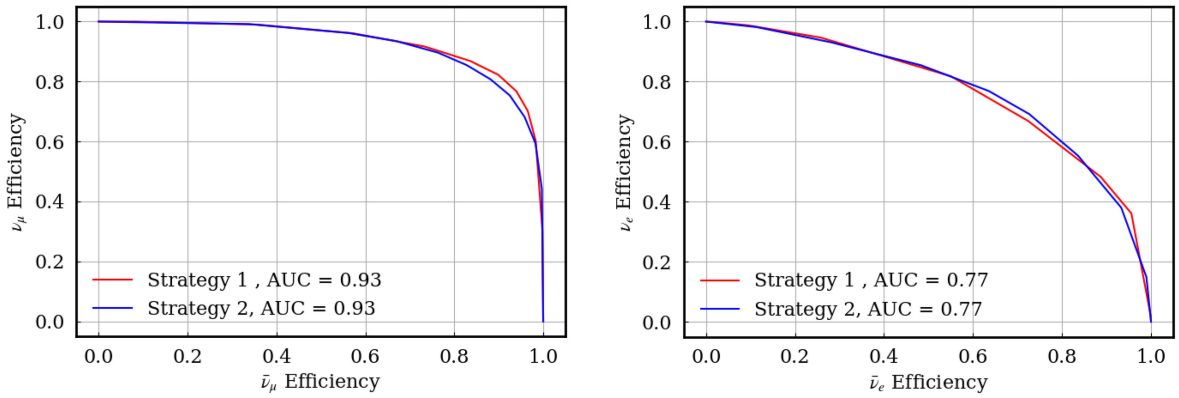


FIG. 12. ROC curves of the ν_μ vs $\bar{\nu}_\mu$ (left) and ν_e vs $\bar{\nu}_e$ (right) identification. The vertical axis represents the selection efficiency of neutrinos by varying cut values on the score shown in Fig. 11, and the horizontal axis represents the efficiency of the antineutrinos. The corresponding AUC values are also listed.

B. Selection criteria optimization

After the flavor and $\nu/\bar{\nu}$ identification, the candidate atmospheric neutrino events shall be categorized into ν_μ-CC, $\bar{\nu}_\mu$-CC, ν_e-CC, $\bar{\nu}_e$-CC, and NC categories, while only the former four CC categories are used for the oscillation analysis. ML models categorize the event into the class with the largest probability score value by default. However, as mentioned previously, this default classification may not be optimal for oscillation analysis. Different selections based on the probability scores can be applied, causing events to migrate among the categories, leading to different signal efficiency and background contamination in each category. For future atmospheric neutrino oscillation analysis, the selection criteria should be optimized to achieve the best sensitivity. It is for the same reason that the ROC curve and AUC values are chosen to represent the identification performances in this work instead of more conventional benchmarks such as signal efficiency and purity.

C. Alternative classification strategy

As described in Sec. III, the baseline strategy of this work is to do the μ-like vs e-like vs NC-like 3-label classification first, followed by the $\nu/\bar{\nu}$ 2-label classification for the identified μ- and e-like events. Thus, this strategy is referred to as the $3+2$-label strategy. Alternatively, the task can also be achieved by doing the ν_e-CC/$\bar{\nu}_e$-CC/ν_μ-CC/$\bar{\nu}_\mu$-CC/NC classification all at once using both the PMT features from the prompt trigger and the delayed trigger information, referred to as the 5-label strategy. To test this 5-label strategy, the models used for the 2-label classifications are retrained with the same primary trigger and neutron-capture information inputs, and the model architectures remain the same except the outputs are changed to five scores representing the probability of the event being $\nu_\mu/\bar{\nu}_\mu/\nu_e/\bar{\nu}_e$/NC-like. The performance is then tested using the same Honda-flux sample and compared with the $3+2$-label results using the total AUC scores

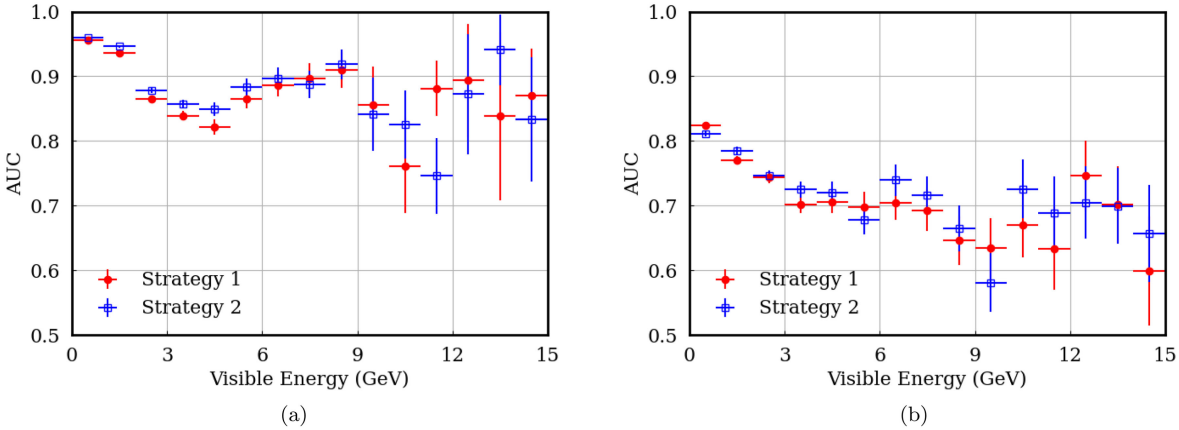


FIG. 13. Comparison of the AUC values for the 2-label (a) $\nu_\mu/\bar{\nu}_\mu$ and (b) $\nu_e/\bar{\nu}_e$ identifications as functions of visible energy between the two strategies.

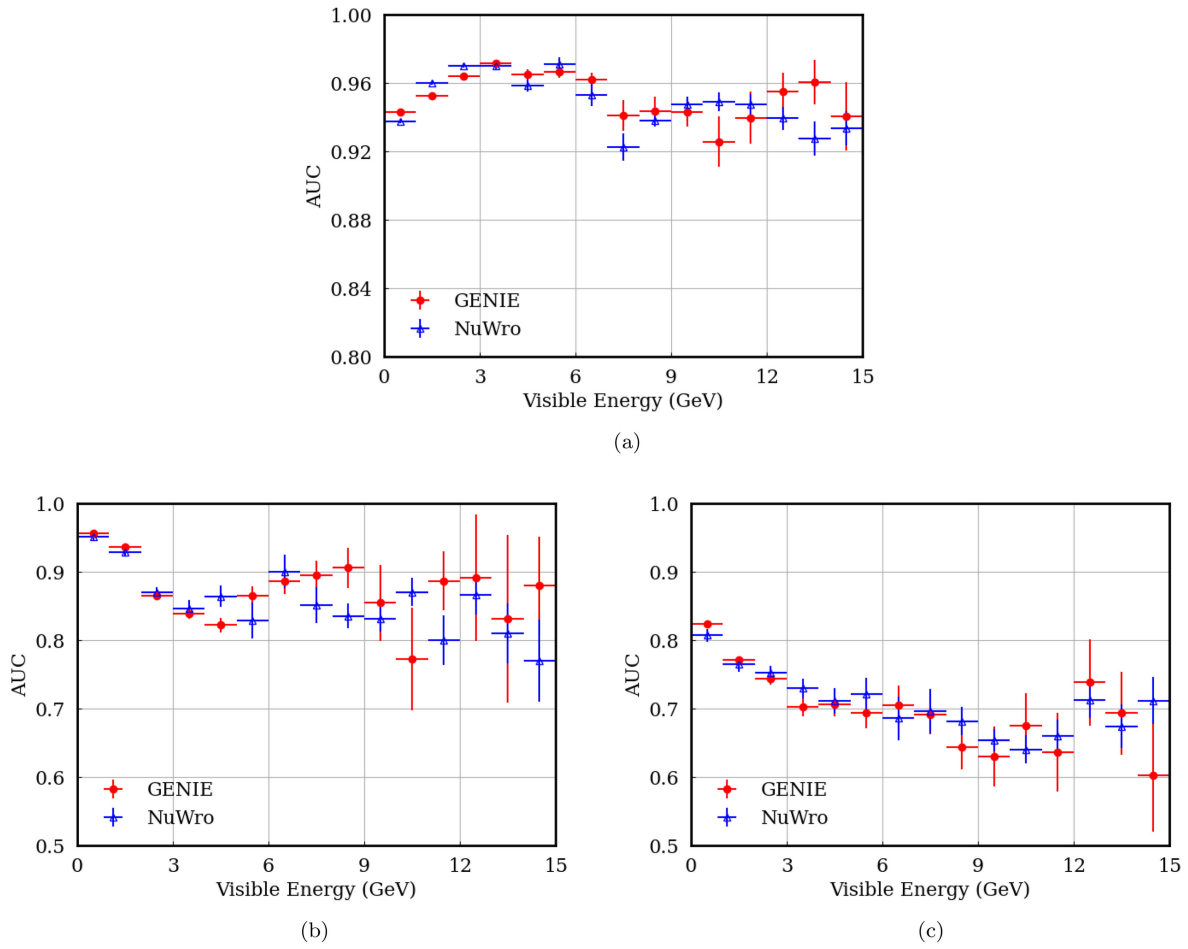


FIG. 14. Comparison of the identification performance using the GENIE-generated Honda-flux sample and the NuWro sample. Both (a) the total AUC values for the 3-label flavor identification and the AUC values for the 2-label (b) $\nu_\mu/\bar{\nu}_\mu$, (c) $\nu_e/\bar{\nu}_e$ identifications are shown. The PointNet++ model and strategy 1 of neutron-capture information handling are used.

described in Sec. IV for the two different neutron-capture information handling strategies individually. To have a fair comparison, both the PMT features from the primary trigger and the neutron-capture information are also used for the 3-label classification in the 3 + 2-label strategy. The total average AUC comparisons of the two strategies are shown in Fig. 15, and consistent performances are observed.

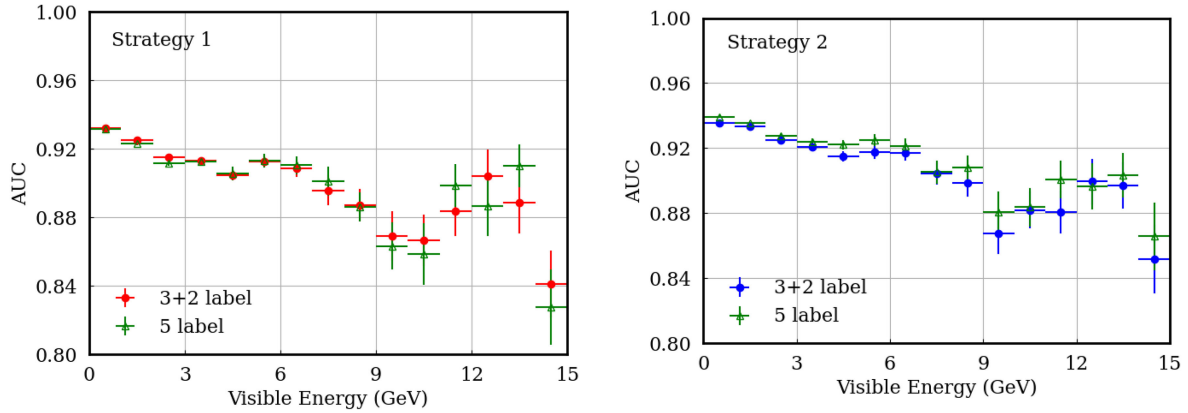


FIG. 15. Comparison of the identification performance (the total AUC values as functions of visible energy) between 5-label and 3 + 2-label strategies for the two different neutron information handling strategies.

VI. SUMMARY

A ML-based method for the flavor identification and $\nu/\bar{\nu}$ discrimination of atmospheric neutrinos in a large homogeneous LS detector is presented in this paper. It utilizes both the features extracted from PMT waveforms in the prompt trigger and the neutron-capture information in the delayed triggers, taking advantage of LS detectors' low threshold measurements of both the charged lepton and hadrons in a neutrino-nucleus interaction and the capability of high-efficiency neutron tagging. PMT features from the prompt trigger can capture different topologies of e-/μ-/NC-like events, leading to excellent neutrino flavor identification. Moreover, they also contain the y_{vis} information, which, combined with the neutron-capture information from the delayed triggers, exhibits good separation power between ν and $\bar{\nu}$. Different ML models, neutron-capture information handling strategies, and event generators are tested and show consistent performances. Using JUNO as an example, we demonstrate for the first time that large homogeneous LS detectors can perform identification of the atmospheric neutrino flavor and neutrino/antineutrino discrimination with their unique advantages. Combined with excellent energy resolution and directional measurement capability for neutrinos reported in [29], LS detectors such as JUNO can provide competitive results for future atmospheric neutrino oscillation measurements.

ACKNOWLEDGMENTS

This work was partially supported by CAS Project for Young Scientists in Basic Research (Grant No. YSBR-099), by the National Natural Science Foundation of China (Grants No. 12275160, No. 12105158, No. 12025502, and No. 12405230), by National Key R&D Program of China (Grant No. 2024YFE0110500), and by Shandong Provincial Natural Science Foundation (Grants No. ZR2022MA062 and No. ZR2024QA127). We would also like to thank the Computing Center of the Institute of High Energy Physics, Chinese Academy of Science, and the Key Laboratory of Particle Physics and Particle Irradiation of Ministry of Education, Shandong University for providing the GPU resources.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available upon publication because it is not technically feasible and/or the cost of preparing, depositing, and hosting the data would be prohibitive within the terms of this research project. The data are available from the authors upon reasonable request.

[1] K. S. Hirata *et al.* (Kamiokande-II Collaboration), Phys. Lett. B **205**, 416 (1988).
[2] Y. Fukuda *et al.* (Super-Kamiokande Collaboration), Phys. Rev. Lett. **81**, 1562 (1998).
[3] L. Wolfenstein, Phys. Rev. D **17**, 2369 (1978).
[4] S. P. Mikheev and A. Y. Smirnov, Nuovo Cimento C **9**, 17 (1986).
[5] K. Abe *et al.*, arXiv:1109.3262.
[6] K. Abe *et al.* (Hyper-Kamiokande Collaboration), arXiv:1805.04163.
[7] S. Adrian-Martinez *et al.* (KM3Net Collaboration), J. Phys. G **43**, 084001 (2016).
[8] S. Aiello *et al.* (KM3NeT Collaboration), Eur. Phys. J. C **82**, 26 (2022).
[9] A. Ishihara (IceCube Collaboration), Proc. Sci. ICRC2019 (**2021**) 1031.
[10] F. An *et al.* (JUNO Collaboration), J. Phys. G **43**, 030401 (2016).
[11] A. Abusleme *et al.* (JUNO Collaboration), Prog. Part. Nucl. Phys. **123**, 103927 (2022).
[12] M. G. Aartsen *et al.* (IceCube-Gen2 Collaboration), Phys. Rev. D **101**, 032006 (2020).
[13] A. Cabrera *et al.*, Sci. Rep. **12**, 5393 (2022).
[14] S. Baum, M. Carena, N. R. Shah, C. E. M. Wagner, and Y. Wang*et al.* (KM3NeT and JUNO Collaborations), J. High Energy Phys. 03 (2021) 055.
[15] K. Eguchi *et al.* (KamLAND Collaboration), Phys. Rev. Lett. **90**, 021802 (2003).
[16] C. Arpesella *et al.* (Borexino Collaboration), Phys. Rev. Lett. **101**, 091302 (2008).
[17] M. Agostini *et al.* (Borexino Collaboration), Nature (London) **562**, 505 (2018).
[18] F. P. An *et al.* (Daya Bay Collaboration), Phys. Rev. Lett. **108**, 171803 (2012).
[19] F. P. An *et al.* (Daya Bay Collaboration), Phys. Rev. Lett. **130**, 161802 (2023).
[20] J. K. Ahn *et al.* (RENO Collaboration), Phys. Rev. Lett. **108**, 191802 (2012).
[21] G. Bak *et al.* (RENO Collaboration), Phys. Rev. Lett. **121**, 201801 (2018).
[22] Y. Abe *et al.* (Double Chooz Collaboration), Phys. Rev. Lett. **108**, 131801 (2012).
[23] H. de Kerret *et al.* (Double Chooz Collaboration), Nat. Phys. **16**, 558 (2020).
[24] T. Wester *et al.* (Super-Kamiokande Collaboration), Phys. Rev. D **109**, 072014 (2024).
[25] K. Abe *et al.* (Super-Kamiokande Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **1027**, 166248 (2022).
[26] K. Abe *et al.* (Super-Kamiokande Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **1065**, 169480 (2024).
[27] M. Ribordy and A. Y. Smirnov, Phys. Rev. D **87**, 113007 (2013).

[28] S. G. Olavarrieta, M. Jin, C. A. Argüelles, P. Fernández, and I. Martínez-Soler, Phys. Rev. D **110**, L051101 (2024).
[29] Z. Yang *et al.*, Phys. Rev. D **109**, 052005 (2024).
[30] A. Abusleme *et al.* (JUNO Collaboration), Eur. Phys. J. Plus **139**, 1128 (2024).
[31] T. Lin *et al.*, Eur. Phys. J. C **83**, 382 (2023); **83**, 660(E) (2023).
[32] A. Coppi *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **1052**, 168255 (2023).
[33] C. Cao *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **1005**, 165347 (2021).
[34] A. Abusleme *et al.* (JUNO Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **988**, 164823 (2021).
[35] C. Andreopoulos, C. Barry, S. Dytman, H. Gallagher, T. Golan, R. Hatcher, G. Perdue, and J. Yarba, arXiv:1510.05494.
[36] Tena-Vidal *et al.* (GENIE Collaboration), Phys. Rev. D **104**, 072009 (2021).
[37] T. Golan, J. T. Sobczyk, and J. Zmuda, Nucl. Phys. B, Proc. Suppl. **229–232**, 499 (2012).
[38] S. Agostinelli *et al.* (GEANT4 Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **506**, 250 (2003).
[39] J. Allison *et al.*, IEEE Trans. Nucl. Sci. **53**, 270 (2006).
[40] M. Honda, M. S. Athar, T. Kajita, K. Kasahara, and S. Midorikawa, Phys. Rev. D **92**, 023004 (2015).
[41] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series* (The MIT Press, Cambridge, MA, 1964).
[42] Y. Huang, J. Chang, Y. Cheng, Z. Chen, J. Hu, X. Ji, F. Li, J. Li, Q. Li, X. Qian, S. Jetter, W. Wang, Z. Wang, Y. Xu, and Z. Yu, Nucl. Instrum. Methods Phys. Res., Sect. A **895**, 48 (2018).
[43] N. Perraudin, M. Defferrard, T. Kacprzak, and R. Sgier, Astron. Comput. **27**, 130 (2019).
[44] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, in *Advances in Neural Information Processing Systems* (Curran Associates, Inc., Red Hook, NY, 2017), Vol. 30.
[45] G.-h. Huang, W. Jiang, L.-j. Wen, Y.-f. Wang, and W.-M. Luo, Nucl. Sci. Tech. **34**, 83 (2023).
[46] Y. Wang, Y. Sun, Z. Liu, S. E. Sarma, M. M. Bronstein, and J. M. Solomon, ACM Trans. Graph. **38**, 1 (2019).
[47] J. Hanley and B. McNeil, Radiology **143**, 29 (1982).
[48] B. Efron, Ann. Stat. **7**, 1 (1979).