Multi-station Neutrino Vertex Reconstruction with Convolutional Neural Networks for a Simulated RNO-G Sub-array

Santiago Sued and Brian Clark
University of Maryland, College Park

Ultra-high-energy (UHE) neutrinos serve as unique messengers for probing the most energetic environments in the Universe. The search for these elusive particles has driven the development of a new generation of in-ice detectors that utilize coherent radio Cherenkov emission (the Askaryan effect) from neutrino-induced cascades to detect neutrinos. The kilometer-scale attenuation length of radio waves in cold ice allows for the instrumentation of large detector volumes with sparse arrays. Consequently, this extended propagation range significantly increases the likelihood of coincident multi-station detections. Despite this, current reconstruction algorithms predominantly focus on reconstruction with just a single detector's worth of information, ignoring the rich geometric and temporal interplay of multi-station events. In this work, we present a novel machine learning approach utilizing convolutional neural networks to reconstruct the neutrino interaction vertex from four-station coincident events using simulated Radio Neutrino Observatory in Greenland (RNO-G) stations. The architecture is modular, facilitating the rapid adaptation of other single-station reconstruction techniques. We demonstrate that while the model exhibits systematic under-prediction of the vertex radius, the radial precision achieves a relative 1σ (68% containment) error of 0.32, an improvement of a factor of two over the 0.60 achieved by current state-of-the-art single-station methods. These results highlight the transformative potential of deep learning techniques for exploiting multi-station coincidence in next-generation radio neutrino observatories.

I. INTRODUCTION

Historically, the electromagnetic spectrum has been the primary tool for studying the cosmos. To probe the universe's most energetic accelerators, standard model multi-messenger astrophysics utilizes the particles produced by these environments—namely cosmic rays and neutrinos—to provide critical, complementary information on the location and inner workings of these sources. However, in the ultra-high-energy (UHE) regime, interactions during propagation limit the utility of charged particles and photons, making the UHE neutrino a uniquely crucial messenger.

Cosmic rays are massive, charged subatomic particles, typically protons, accelerated by extreme astrophysical sources. However, their utility as directional messengers at ultra-high energies is hindered by two physical factors. First, because they are charged, they are continuously deflected by galactic and extragalactic magnetic fields, effectively destroying their ability to point back to their origins. Second, UHE cosmic ray propagation is constrained by the Greisen–Zatsepin–Kuzmin (GZK) limit [1][2]. At energies exceeding $\sim 5 \times 10^{19}$ eV, protons interact with the ubiquitous photons of the Cosmic Microwave Background (CMB) to produce pions. This interaction downscatters the energy of the proton, limiting its maximum travel distance to roughly 42 million light-years, a distance that encompasses less than 0.1% of the observable universe.

On the other hand, photons have the benefit of being electrically neutral and massless. This allows for straight-line propagation, directly pointing back to the sources that produced them. Unfortunately, at ultra-high energies, photons interact with the Cosmic Microwave Background (CMB) and Cosmic Radio Background (CRB) through the Breit–Wheeler process [3] to produce electron-positron pairs. As detailed by Gould and Schréder [4], for photons between 10^{15} and 10^{16} eV, the mean free path is approximately 10^{22} cm. This distance is actually smaller than the 3×10^{22} cm radius of the Milky Way. Even at higher energies approaching 10^{20} eV, where the CRB becomes the dominant target, the propagation distance remains restricted to scales significantly smaller than the GZK horizon. This effectively renders the extragalactic universe opaque to UHE light, eliminating photons as viable messengers for the distant universe.

This leaves neutrinos, whose neutral charge and negligible mass mimic the attractive characteristics of photons—such as straight-line propagation—yet do not possess the same energy limitations, allowing for the exploration of the UHE universe. We expect these UHE neutrinos to originate from two primary mechanisms. First, cosmogenic neutrinos are generated during propagation as a direct byproduct of the GZK process, when UHE cosmic rays interact with the CMB. Second, astrophysical neutrinos are produced directly within the environments of extreme cosmic accelerators, such as Active Galactic Nuclei (AGN), where accelerated protons interact with ambient gas or radiation fields at the source.

The detection of UHE neutrinos represents one of the most active frontiers in modern astrophysics. Recent milestones, such as the February 2025 detection of a ~ 220 PeV neutrino by the KM3NeT observatory, provide a glimpse into the potential of UHE messengers. However, as illustrated by the flux predictions in Fig. 1, the anticipated neutrino rate drops as a function of energy. This restricts the probability of detection to two main factors: the time spent observing and the physical size of the "net" cast by the detector. At the EeV

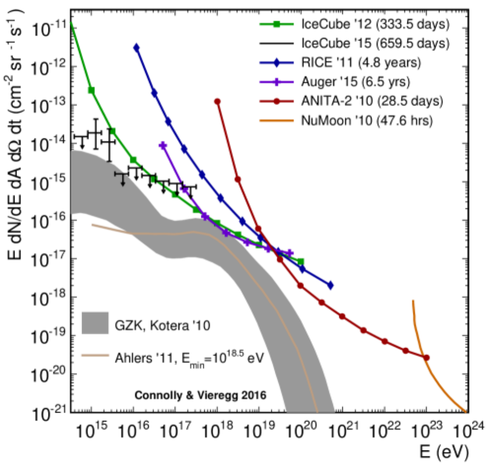

FIG. 1. Most competitive experimental constraints on the all-flavor diffuse flux of the highest energy neutrinos compared to representative model predictions [5].

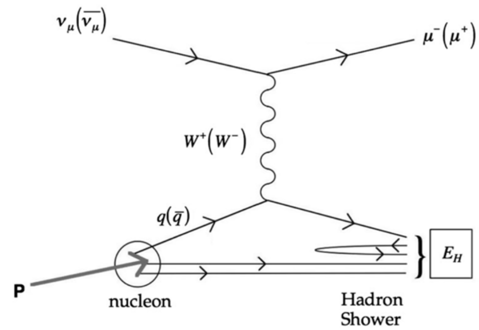

FIG. 2. Charged Current Deep Inelastic Scattering Feynman Diagram. Adapted from Ref. [7]. For the Neutral Current case, the exchange is mediated through the Z^0 boson.

scale, the flux is so low that waiting for a detection with current optical detector volumes is statistically impractical [6]. To bridge this gap, next-generation experiments like the Radio Neutrino Observatory Greenland (RNO-G) and IceCube-Gen2 are designed to scale the detection volume by orders of magnitude.

A. Neutrino Detection

Neutrinos have a small interaction cross section, on the order of 10^{-34} cm^2 for non-UHE neutrinos [7], which is over ten orders of magnitude weaker than electromagnetic interactions. This makes neutrino detection heavily dependent on target volume; because the probability of a collision is so minuscule, instrumenting gigatons of matter is an absolute necessity to capture a statistically significant number of interactions from the already low UHE flux.

Neutrinos interact via the weak force, specifically through neutral-current (NC) and charged-current (CC) interactions. In the UHE regime, these interactions occur primarily through a process known as deep inelastic scattering (DIS) [7]. In a DIS CC interaction (shown in Fig. 2), mediated by the $W^\pm$ boson, the neutrino scatters off a nucleon quark, transferring a large fraction of its momentum and converting into its corresponding charged lepton. The scattered quark then generates a highly energetic hadronic shower because of color confinement. If the produced lepton is a muon or tau, it continues its path across the detector; however, if it is an electron, it immediately initiates a secondary electromagnetic shower. In an NC interaction, mediated by the uncharged Z^0 boson, the neutrino transfers a portion of its energy to knock a quark out of the nucleon, again producing a hadronic shower.

Regardless of the interaction type, the resulting hadronic shower is rich in neutral pions (π^0), which rapidly decay into high-energy photons. These photons then undergo pair production, kick-starting an electromagnetic cascade. As electromagnetic showers propagate through the ice, they strip the medium of electrons via Compton scattering, while their positrons annihilate with ambient electrons. Together, these processes generate a net negative charge within the advancing cascade [8]. Because these relativistic charge carriers travel faster than the phase velocity of light in the ice, they individually emit visible light via the standard Cherenkov effect. Traditional neutrino detection relies on instrumenting massive volumes of naturally transparent material to capture these optical flashes—a technique famously employed by the IceCube Neutrino Observatory, which utilizes a cubic kilometer of deep glacial ice at the South Pole [9].

A consequence of this charge-excess cascade is a process known as Askaryan Radiation, first proposed by Soviet-Armenian physicist Gurgen Askaryan in 1962 [10]. Askaryan predicted that for emission wavelengths larger than the physical dimensions of the shower, the Cherenkov radiation from the individual particles will constructively interfere. This produces a coherent radio-frequency pulse whose power scales quadratically with the shower energy [11]. The emission forms a conical shockwave aligned with the shower trajectory, with a Cherenkov angle of approximately 56° in deep ice (See Fig. 3). In the UHE regime, the pulses produced have immense, easily detectable energy, but because the particle showers reach lengths of up to 10 meters in length the constructive wavelengths correspond to the radio spectrum.

This phenomenon was confirmed in 2001 at the SLAC laboratory using bremsstrahlung photons from 28.5 GeV electrons on a silica sand target [11] and recently confirmed in nature at the ARA detector in the South Pole [12].

To capture the low flux of UHE neutrinos, detectors must instrument hundreds of cubic kilometers of matter. Because constructing artificial detectors of this scale is fi-

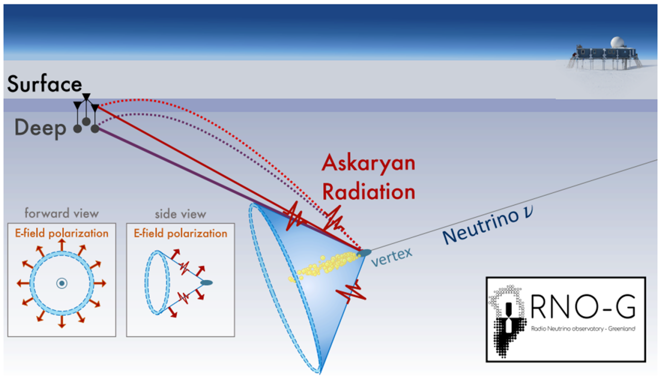

FIG. 3. Side-view diagram of the Cherenkov cone. The signal propagates through the ice and gets refracted before reaching the stations.

nancially and practically impossible, physicists must rely on naturally occurring, dense, and transparent media, such as the polar ice caps. While optical observatories like IceCube have proven the viability of in-ice detection, the standard Cherenkov light they rely on has an attenuation length of approximately 100 meters in glacial ice [13]. But in the UHE regime, Askaryan radiation provides a broadband radio signal that travels an order of magnitude further than optical light—achieving attenuation lengths of approximately 1 kilometer [14]. This long radiation length allows a sparse array of radio antennas, spaced widely apart, to effectively monitor gigatons of ice. Consequently, the radio method is preferred at these energy levels, as it allows researchers to physically and economically instrument the massive volumes necessary to potentially detect UHE neutrinos.

B. Evolution of Radio Neutrino Detectors

Askaryan radiation was the inspiration for the Radio Ice Cherenkov Experiment (RICE) which ran from 1999 to 2012 at the South Pole. Consisting of multiple half-wave dipole antennas distributed within a 200 m x 200 m x 250 m volume [15], RICE established the first long-term upper limits on the ultra-high-energy neutrino flux [16]. It also laid the groundwork for future experiments by characterizing essential properties of the antarctic ice, such as its index of refraction [15].

The Askaryan Radio Array (ARA) emerged as the natural successor to RICE. Beginning construction in 2010, ARA capitalized on RICE's foundation and became the first experiment to implement a distributed, station-based design. The ARA detector is composed of five stations 2 km apart, with each station featuring 16 embedded antennas. ARA's fifth and final station introduced a phased array with an interferometric trigger, which lowered the 50% trigger threshold from a signal-to-noise ratio (SNR) of 3.4 to 2 [17]. This phased array technique has since become a staple of radio-based UHE neutrino detection.

While most experiments thus far have buried their detectors deep under the ice, the Antarctic Ross Ice-Shelf Antenna Neutrino Array (ARIANNA) took a different approach. ARIANNA exploits the reflected radio waves produced from the interface between water and ice in the Ross Ice shelf, allowing detectors to be located at the surface, as seen in Fig. 4[18]. Surface stations enable the use of Log Periodic Dipole Antennas (LPDAs) which offer higher gain and a wider resonant frequency band than their underground counterparts. ARIANNA paved the way for surface-level UHE neutrino technology.

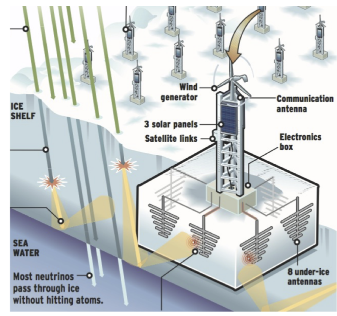

FIG. 4. Single ARIANNA station. Because they are in the surface, large LPDA antennas can be positioned without need for a massive bore hole.

The technological advancements from these precursor experiments have culminated in the Radio Neutrino Observatory in Greenland (RNO-G), which is currently under construction [19]. RNO-G serves as a direct synthesis of its predecessors, explicitly combining the deep interferometric phased array pioneered by ARA with the high-gain surface LPDAs championed by ARIANNA. The observatory will consist of 35 autonomous radio stations laid out in a rectangular grid with a 1.25 km separation. This station spacing is an order of magnitude larger than the 125 m separation of the optical IceCube stations [20], allowing RNO-G to monitor a massive effective volume with a fraction of the infrastructure. Even with this large station separation, UHE neutrinos are expected to be detected at more than one station. Figure 5 shows the expected coincidence fraction (percentage of neutrino signals detected by more than one station) for different grid spacings and different energies. We expect more than 10% of UHE events to be detected by more than one station, highlighting the need for multistation methods in this regime.

As shown in Fig. 6, each station comprises 24 antenna channels distributed across the surface LPDAs, two deep helper stations for vertex triangulation, and the deep phased array trigger first established in ARA.

Crucially, RNO-G serves as the foundational testing ground for the radio component of IceCube-Gen2, the planned next-generation expansion of the IceCube Neutrino Observatory. Because of its pivotal role in the future of UHE neutrino astronomy, RNO-G is the main focus of this paper, and the station simulations utilized in

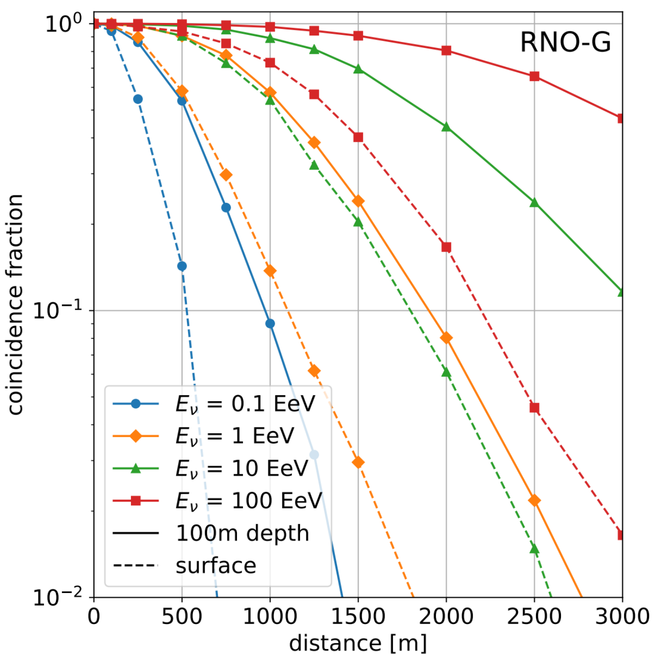

FIG. 5. Coincidence fraction (percentage of events detected by more than one station) with respect to station grid spacing. For in-ice events higher than 1 EeV, the coincidence fraction exceeds 50% for current RNO-G spacing.

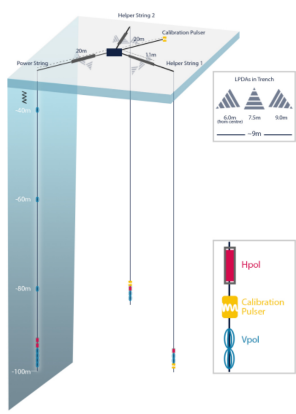

FIG. 6. Schematic of an RNO-G. Each of the three strings has a set of Vertically Polarized (VPOL) and Horizontally Polarized (HPOL) antennas, as well as a calibration pulser.

our subsequent analysis are derived directly from RNO-G open-source models.

C. Neutrino Energy Reconstruction

A main objective of next-generation radio neutrino observatories is not just the detection of UHE neutrinos, but the accurate reconstruction of their primary energy. Because characterizing the high-energy limits of the astrophysical neutrino flux is a major motivation for these experiments, extracting the energy from the recorded radio pulses is critical.

Work by Hanson and Connolly summarizes the state of the art in Askaryan emission modeling from first principles—using complex analysis and cascade equations—vastly improving upon the computationally intractable Monte Carlo simulation methods used previously to simulate the entire particle shower [21].

This analysis provides two key equations detailing the electric field amplitude in frequency space for a neutrino cascade. The first equation defines the field precisely on the Cherenkov cone:

$$|\vec{E}(f,\theta_c)| \propto \frac{E_0}{R}\left[\frac{f/f_0}{1+(f/f_0)^2}\right] \quad (1)$$

The second equation introduces a Gaussian attenuation factor for observing the shower off-cone:

$$|\vec{E}(f,\theta)| = |\vec{E}(f,\theta_c)| \exp\left[-\frac{1}{2}\left(\frac{\theta-\theta_c}{\Delta\theta}\right)^2\right] \quad (2)$$

Here, E_0 is the field normalization constant (a value linearly proportional to the shower energy), f is the frequency of the signal, f_0 is the characteristic decoherence frequency (above which the pulse begins to destructively interfere with itself), R is the distance from the interaction vertex to the detector, θ is the viewing angle, θ_c is the Askaryan angle of maximum coherence, and $\Delta\theta$ is the angular width of the Cherenkov cone.

These two equations demonstrate that the energy of the field normalization constant, E_0, is directly calculable from the measured electric field amplitude, $|\vec{E}|$, and the frequency spectrum, f. However, this calculation is linearly dependent on the distance to the interaction vertex (R) requiring the 3D location of the interaction vertex as a pre-requisite for successful energy reconstruction. Once the field normalization constant is obtained, we can use it to find the shower energy which—depending on the flavor of neutrino—will inform us on the primary neutrino energy.

D. Objective

The goal of this study is to leverage machine learning techniques to reconstruct the exact neutrino interaction vertex from their Askaryan Radiation. This allows us to reconstruct the radius to the detector (R), directly from multi-station voltage traces.

While traditional statistical approaches use likelihood minimization methods to successfully reconstruct the interaction vertex with good accuracy [22], to our knowledge, deep learning has not yet been applied to multi-station vertex reconstruction for in-ice radio detectors.

These multi-station events intrinsically posses more data which ideally allows us to improve neutrino vertex reconstruction.

This work builds heavily upon the established methodologies presented by Aguilar et al. [22]. Their analytical reconstruction framework serves as both the foundational physical model for this study and the main benchmark against which the performance and computational efficiency of our Convolutional Neural Network (CNN) are evaluated.

II. METHODOLOGY

Machine learning models are inherently "data-hungry", requiring large amounts of preprocessed data to train and validate. The most effective way of obtaining clean, structured data for training is through simulation. Thus, we designed an end-to-end data pipeline to convert simulated voltage detections into trainable samples. Using the procedure detailed below, we generated a dataset consisting of approximately 1 million multi-station neutrino events. This section will cover the dataset generation, model architecture, training, and evaluation. The complete codebase used for this pipeline is open-source and available on GitHub at `https://github.com/umd-pa/RNO_vertex_reconstruction_ml`.

A. Dataset Generation and Preprocessing

The dataset generation pipeline is divided into three distinct phases: neutrino injection, detector simulation, and dataset preparation.

1. Neutrino Injection

To simulate the physical particle cascades and subsequent radio emissions we utilized NuRadioMC, a dedicated Monte Carlo simulation framework for radio neutrino detectors [23]. Neutrino interaction vertices were generated uniformly within a cylindrical volume spanning a depth from $z = -2.7$ km to 0 km and a radius of $R = 3.9$ km. Primary neutrino energies were uniformly sampled between 5×10^{16} eV and 1×10^{19} eV from a combination of cosmogenic (GZK-2 [24]) and astrophysical [25] flux model distributions.

2. Detector Simulation

The NuRadioMC pipeline simulates the incident neutrinos, their Askaryan radio emission, the propagation of those radio waves through the Greenlandic ice, and the final receipt and recording of those signals by the detector hardware. For this hardware simulation, we use the publicly-available model of the RNO-G station. We simulate an array of four stations spaced 1.25 km apart, which is the inter-station spacing of the RNO-G detector. We chose this layout as the simplest way to begin studying multi-station reconstruction, as we do not expect more than four stations to see the pulse from a single UHE neutrino.

As outlined by Aguilar et al. [22], we implemented a simple high-low threshold trigger as a proxy for the phased array trigger of RNO-G, with a minimum voltage threshold of -20 mV and a maximum of 20 mV. For each triggered event, we injected Rayleigh-distributed noise with an RMS of 10 mV. This corresponds to an SNR of 2, which is the expected trigger threshold for the phased array [17]. A schematic of the simulated stations and neutrino generation cylinder can be seen in Fig. 7.

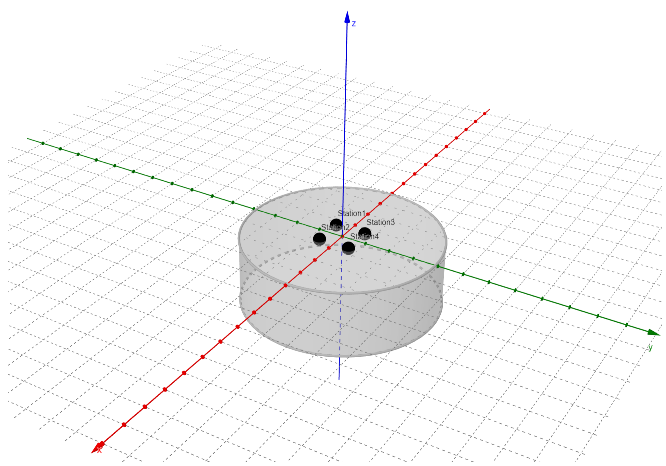

FIG. 7. To-scale schematic of simulated stations and the simulation cylinder that surrounds it, rendered in GeoGebra.

3. Dataset Preparation

Each NuRadioMC simulation outputs a .nur file (a proprietary file format exclusive to NuRadioMC [23]) containing a collection of neutrino events. By default, NuRadioMC logs events strictly on a per-station basis, implying that each station will output a different event even if it is from the same primary. To train a multi-station neural network, it is critical that traces from the same parent neutrino are aggregated across all triggered stations into a single, cohesive sample. To fix this, we grouped all coincident station triggers originating from the same primary particle. Consequently, in this work, an **event** is defined as the collective triggering of *all* stations to a single neutrino interaction.

Because deep learning involves training on hundreds of thousands of events, it would be computationally intractable and highly wasteful to store the entire 2.4 GHz raw voltage trace for every channel of every station as a single input tensor. Therefore, we preprocess the data to extract summary features for vertex reconstruction while reducing the memory requirements. The data is preprocessed into a multidimensional representation of a neutrino event which we call an **image**. This term was chosen because the resulting tensor layout mimics the spatial and depth dimensions of a standard computer vision image sample. An image is produced through the sequential procedure outlined below:

1. **Time Windowing:** Because spatially separated stations observe the radio emission at different times depending on their distance to the interaction vertex, the recorded sample must possess a sufficiently large time window to encompass multiple, time-delayed traces. An analysis of the maximum relative arrival time across all events in 425 .nur files was conducted (see Fig. 8). We determined that 95% of all observed multi-station coincident events occurred within an 8,296 ns window. Therefore, we aligned our traces to a fixed temporal window of $2^{13} = 8,192$ ns.

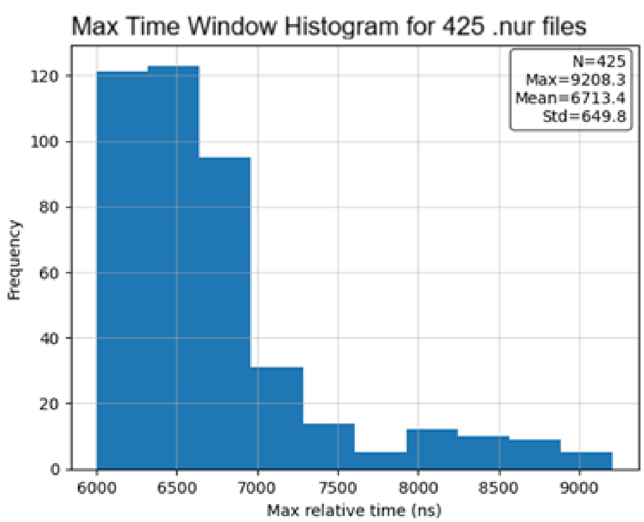

FIG. 8. Distribution of maximum relative time delays across all coincident events. The 95th percentile of time windows was found to be approximately 8,296 ns.

2. **Voltage Clamping:** To accurately mimic the physical constraints of detector hardware, we implemented a hard clamping voltage of 1 V. Any radio signal exceeding this absolute value is clamped, replicating the behavior of a saturated amplifier and digitizer.

3. **Signal Envelope and Binning:** To reduce data dimensionality and complexity, we calculate the Hilbert envelope of the voltage trace—which isolates the voltage amplitude of the signal—and then down sample by taking the average voltage across user-defined time bins.

Following this processing, the traces are spatially structured into images. The binned traces of all channels from a single station are arranged along the first dimension, the time bins along the second, and the distinct stations are stacked along the third. This results in a final tensor of shape $(channels, timebins, stations)$. Each image serves as a complete multi-station summary of the voltage traces from the radio signals of a single particle primary. Figure 9 visually summarizes this generation procedure on a simplified setup.

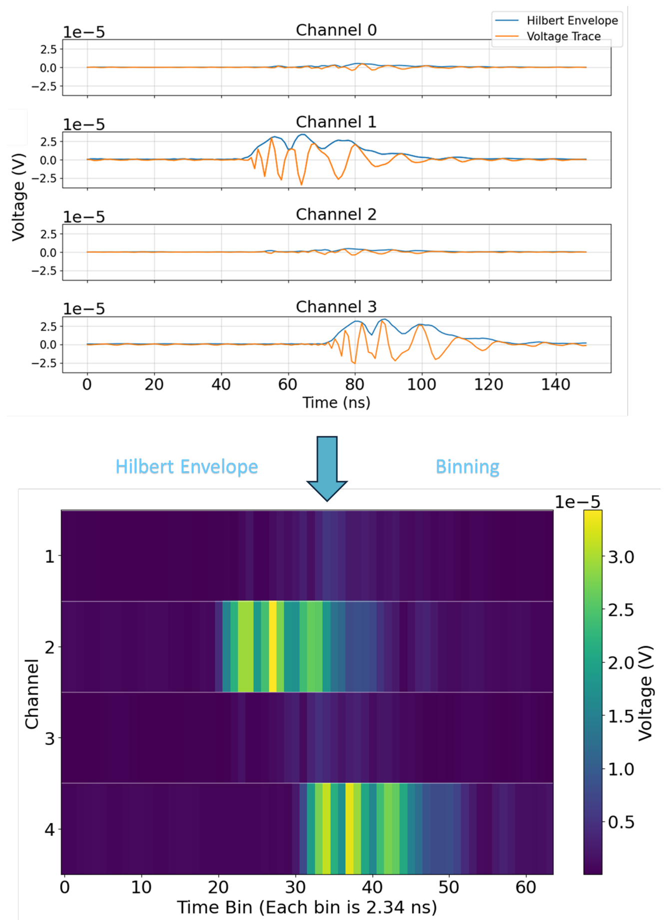

FIG. 9. Simplified image generation diagram for a single station with 4 channels and no injected noise. For the full detector architecture multiple images are stacked along the depth dimension.

For the model to learn, we pair each image with the spatial coordinates of their simulated interaction vertex as the regression target and the station hit counts: the number of stations that saw the event. We define a complete training **sample** as the alignment of three components: the multidimensional image tensor (the network input), the simulated 3D interaction vertex (the target label), and the station hit count (the incidence metadata).

To store these samples efficiently, all events originating from a single .nur file are compiled into a corresponding .hdf5 file, which we define as a **shard**. As illustrated in Fig. 10, each shard maintains three parallel datasets: Images, Vertices, and Hits.

Note that because no trace or noise is stored for a time bin when a station does not trigger, a significant portion of the multi-station image tensors consists of zero values, making the resulting arrays highly sparse. By applying GZIP lvl 1 compression to the .hdf5 image dataset, we reduce the memory footprint to 2% of its original size.

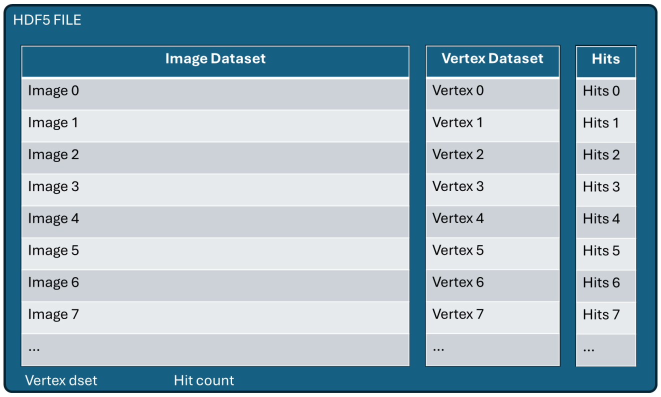

FIG. 10. Internal schematic of a generated .hdf5 data shard. Each file contains three parallel datasets: the multidimensional input images, the true 3D interaction vertices (labels), and the station hit counts. This strict index alignment guarantees consistent input-target pairing during network training.

B. Training Regime and Evaluation Metrics

To effectively train and evaluate the Convolutional Neural Network, the generated dataset was randomly partitioned into three independent subsets: training, validation, and testing. While the data pipeline allows for user-defined splitting ratios, we adopted a standard machine learning distribution of 80% training, 10% validation and 10% testing for this study.

Because our data was distributed across many similar shards, we performed the partitioning at the shard level rather than the event level. We randomly assigned each complete shard to a specific subset, ensuring that the overall 80/10/10 event ratio was conserved. To prevent data leakage, this assignment is recorded in a .json manifest file prior to training.

1. *Dataset Demographics*

We generated exactly 1,017,714 events. To ensure that the model generalizes correctly, it is critical that the training, validation, and testing splits represent identical physical distributions. We verified this by examining the spatial coordinates of the primary neutrino interaction vertices across all three splits. As demonstrated in Fig. 11, the normalized spatial densities for the X, Y, Z coordinates exhibit no appreciable variation between the subsets.

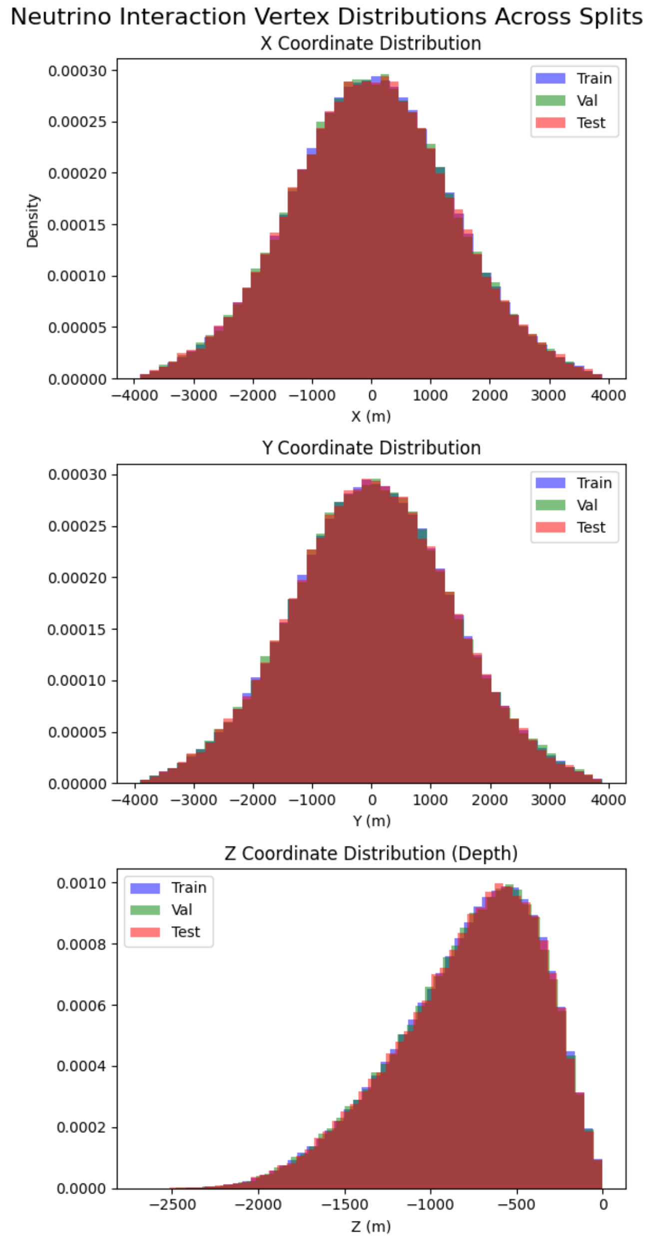

FIG. 11. X,Y,Z distribution of simulated neutrino vertices across testing, training and validation. They all resemble the same distribution.

Furthermore, because multi-station reconstruction is a focus of this study, it is important to quantify the distribution of station hit counts across the dataset, detailed in Table I.

Roughly 88% of the events consist of single station events. This further highlights the necessity of multi-station reconstruction considerations as we expect the radiation of one out of every ten events to reach more than a single station.

TABLE I. Distribution of station incidence counts across the training, testing, and validation sets. The relative proportions remain highly consistent across all three splits, ensuring that the model is evaluated on the same physical distribution it was trained on.

Stations	**Training (%)**	**Test (%)**	**Validation (%)**
1	719,281 (88.39%)	90,162 (88.43%)	90,282 (88.49%)
2	89,522 (11.00%)	11,166 (11.95%)	11,148 (10.93%)
3	4,749 (0.58%)	610 (0.60%)	580 (0.57%)
4	179 (0.02%)	16 (0.02%)	19 (0.02%)
Total	**813,731**	**101,954**	**102,029**

C. Model Training

We implemented our models using the PyTorch framework. To mitigate potential data-generation biases, we dynamically shuffled the simulated data on a per-shard basis using a custom `Dataset` class. While this does not guarantee true global randomness, we consider it sufficient as the data originates from event-independent simulations, making the sequence of samples largely independent.

The network was trained to predict the 3D interaction vertex coordinates by minimizing the `MSE Loss` between the normalized predicted and true simulated vertices. Network parameters were optimized using the AdamW optimizer with a batch size of 256, an initial learning rate of 0.0001, a weight decay of 0.0049 and a dropout rate of 0.1. These values were obtained after a hyperparameter grid-search sweep through the AI developer platform Weights & Biases (W&B).

Given our large dataset of 10^6 events, each training epoch encompasses approximately 3100 gradient updates, resulting in rapid convergence compared to smaller datasets. To exploit early rapid learning while fine tuning optimization in later epochs, we employ PyTorch's built in scheduler, `CosineAnnealingWarmRestarts` with an initial cycle length of $T_0 = 5$ epochs and a cycle multiplier of $T_{\mathrm{mult}} = 2$. The cosine schedule rapidly reduces the learning rate within each cycle before restarting the learning rate decay. This proved more effective than `ExponentialLR` with $\gamma = 0.8$, which decays the learning rate to near-zero over the training run, prematurely suppressing gradient updates before convergence.

During an epoch we perform one complete pass over the training dataset, evaluating the model on a separate validation set at the end of each. To prevent overfitting, we employed an early stopping criterion, terminating training if the validation loss failed to improve for 10 consecutive epochs. All models were trained using an NVIDIA RTX-4090 GPU on the UMD Particle Astrophysics computing cluster.

D. Model Architecture

Convolutional layers have become ubiquitous in modern computer vision due to their ability to extract important features from multidimensional data. These layers utilize a learnable kernel (or filter) that sweeps across the input tensor. At each step, it performs a local dot product between the filter weights and the data elements it is sweeping through. By employing multiple output channels, the layer simultaneously applies numerous independent kernels, each optimizing to detect different patterns from the data.

Because convolutions are fundamentally linear operations, they must be paired with non-linear activation functions to model complex, real-world phenomena. The most standard activation is the Rectified Linear Unit (ReLU), defined mathematically as $f(x) = \max(0, x)$. By mapping all negative values to zero, ReLU enforces sparsity, effectively muting background noise or regions without a physical signal.

During training, the network optimizes its filter weights through backpropagation, utilizing the chain rule to calculate how much each weight contributed to the final error [26]. However, this gradient-based learning introduces the risk of network collapse via "dying neurons" when paired with standard ReLU. If a large weight update pushes a neuron's pre-activation value to a large negative number, the neuron will continuously output zero. Because the local derivative of ReLU is exactly zero for negative inputs, the back-propagated gradient is entirely zeroed out for that node. The weights are therefore never updated, preventing the neuron from ever recovering or learning again.

To mitigate this risk, our model utilizes Batch Normalization (BatchNorm) layers immediately preceding the activation functions. BatchNorm dynamically standardizes the outputs of a convolutional layer, forcing the feature maps to maintain a mean of zero and a variance of one. By keeping the pre-activation values strictly bounded and centered, BatchNorm prevents the inputs from swinging deeply negative. This effectively shields the network from dead neurons while simultaneously smoothing the optimization landscape and allowing for faster convergence.

The sequential combination of a convolutional layer, a batch normalization layer, and a ReLU activation function are so prevalent in CNNs that they receive a special name, the **convolutional block** as seen in Fig. 12.

Even with stable activations, deep networks suffer from the broader vanishing gradient problem. During backpropagation, gradients are multiplied sequentially via the chain rule; as the network scales in depth, these repeated multiplications cause the gradient to decay exponentially, halting the learning process in the earliest layers. To overcome this limitation, He et al. developed the Residual Network (ResNet) [27]. **ResNet blocks** introduce a skip connection that bypasses the convolutional transformations and adds the original input directly to the block's output. Because the derivative of this identity mapping with respect to its input is exactly one, it provides an unimpeded gradient pathway during backpropagation, preserving the learning signal regardless of network depth.

Importantly, unlike usual machine vision architecture, we explicitly omit max pooling layers. Max pooling is ubiquitous in image-based CNNs because discarding intermediate pixel data rarely prevents the network from reconstructing macroscopic visual features. However, in voltage vs. time signal traces, each data point is critical to reconstruction. We stray away from max pooling and rely entirely on strided convolutions (convolutions which skip tensor values as they sweep) to ensure that strict timing intervals are maintained and no fine-grained physical data is arbitrarily discarded.

Our architecture is heavily inspired by 3D ResNet paradigms and separates the data processing into three distinct stages. The full architecture is outlined in Fig. 12.

- **Stage 1: Per-Station Temporal Feature Extraction.** We repeatedly apply one-dimensional kernels strictly to the time-bin dimension. We utilize aggressive strides and kernel sizes (Fig. 12), as the vast majority of voltage time bins consist of zeros. This allows the network to quickly compress the 8,192 time bins into a dense, manageable feature space.

- **Stage 2: Per-Station Channel & Temporal Feature Extraction.** We apply two-dimensional kernels across both the channel and temporal dimensions of each individual station trace. The purpose of this stage is to extract cross-channel correlations and localized features from the distinct channel traces of each station.

- **Stage 3: Cross-Station Time of Arrival Analysis.** Only in the final convolutional phase do we mix the signals from all four distinct stations. This promotes architectural modularity, mitigating the need to develop entirely new network architectures when adapting to different station numbers, and closely mimics independent station-level triggers in real-world data reconstruction.

Following this third stage, the multi-dimensional feature maps are flattened into a single 1D feature vector and passed through a regression head consisting of fully connected (linear) layers, which outputs the final predicted (x, y, z) coordinates of the neutrino interaction vertex.

III. RESULTS

Figure 13 illustrates the training and validation loss curves over the training duration. To demonstrate the impact of the `CosineAnnealingWarmRestarts` learning rate scheduler on the loss landscape, the learning rate is visualized alongside the loss. Training was terminated via early stopping at epoch 71, and the model weights with the minimum validation loss (0.568) at epoch 61 were retained for subsequent evaluation.

To evaluate reconstruction performance, we applied the optimized model to an independent test dataset of $N = 102,029$ events, as outlined in Table I. The model infers the 3D vertex positions directly from the multi-station input image tensors. Figure 14 presents the distribution of the relative error for the reconstructed radial distance.

In Table II, we compare the median relative error and 68% containment interval (1σ) of our proposed neural network against the single-station analytical baseline established in Section I D.

TABLE II. Comparison of relative error for the reconstructed radial distance $(R_{\text{reco}} - R_{\text{true}})/R_{\text{true}}$. The 68% containment interval represents the 16th and 84th percentiles of the distribution.

Work	**Median Error**	$\sigma_{68\%}$
Aguilar et al. [22]	0.01	$[-0.25, 0.35]$
This Work	-0.32	$[-0.48, -0.16]$

Furthermore, to facilitate future machine learning benchmarking in this domain, we report the absolute Euclidean distance between the predicted and true vertex positions for all test events in Fig. 15. This serves as a standard evaluation metric for 3D vertex regression tasks.

IV. DISCUSSION

As observed in Fig. 14, the model exhibits a systematic tendency to under-predict the radial distance. We hypothesize that this behavior arises from the interplay between the `MSELoss` function and signal attenuation in the ice at large distances. As the true radial distance increases, physical attenuation severely degrades the signal, making it blend with the noisy background. Because `MSELoss` penalizes the squared absolute error across all dimensions, it disproportionately punishes incorrect predictions made at extreme distances. To avoid large squared penalties from over-predicting the distance of an uncertain event, the network "hedges" its estimates by pulling them radially inward. Figure 16 provides strong visual evidence for this behavior. While the high-density core of predictions starts off near the ideal $reco = true$ baseline at shorter distances, it distinctly breaks away and under-predicts at higher radii, illustrating this inward pull.

Despite this systematic under-prediction at further vertices, the model still achieves a substantial overall improvement in radial resolution. The network successfully

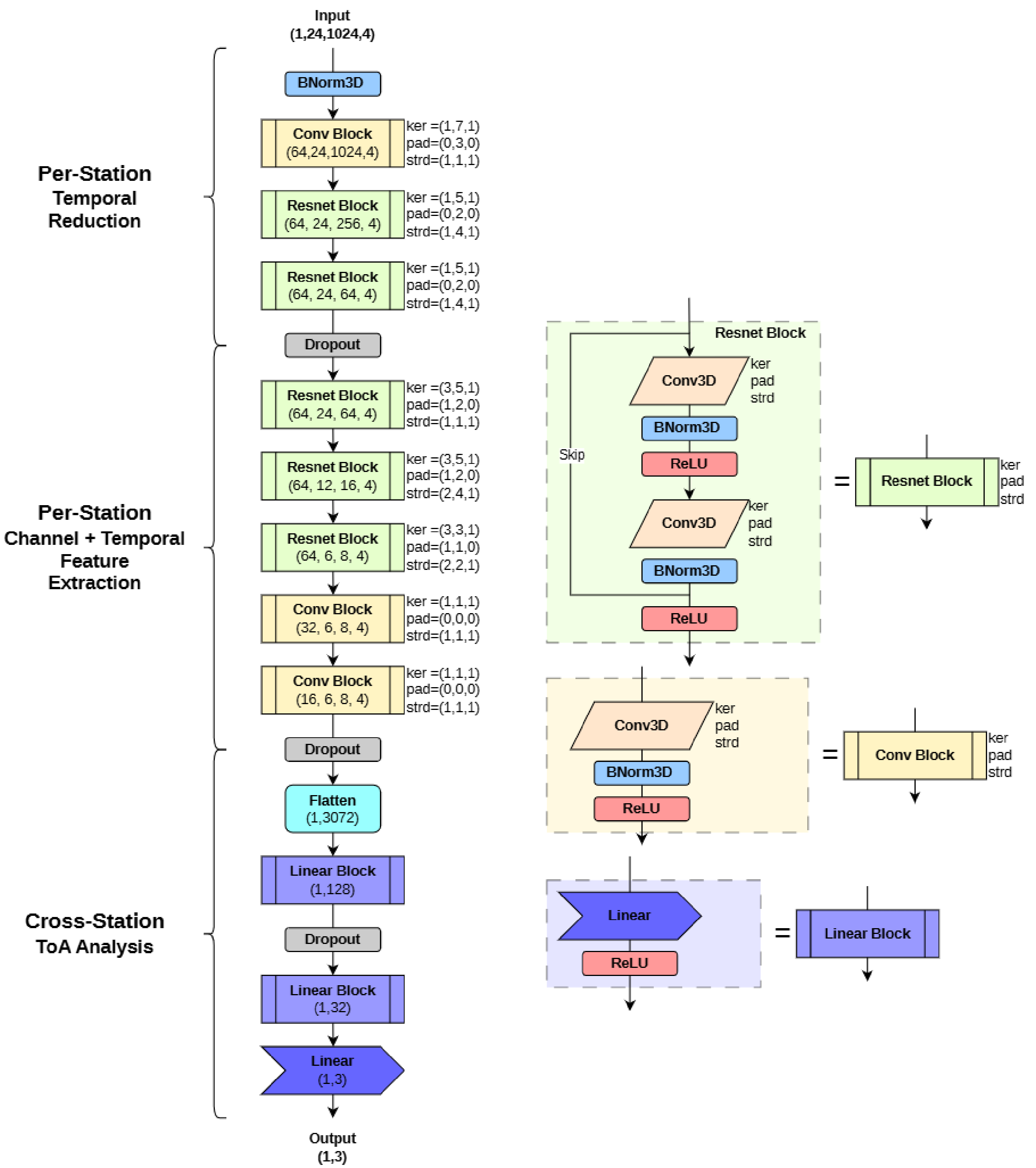

FIG. 12. Architecture of the multi-station neural network model. The pipeline is partitioned into three sequential stages, each designed to extract features across a distinct dimension of the input tensor. Notably, the network employs a late-fusion approach: signals from individual stations are processed independently and are not merged until the final Cross-Station Analysis stage. This preserves a modular structure, enabling single-station feature extraction before multi-station analysis.

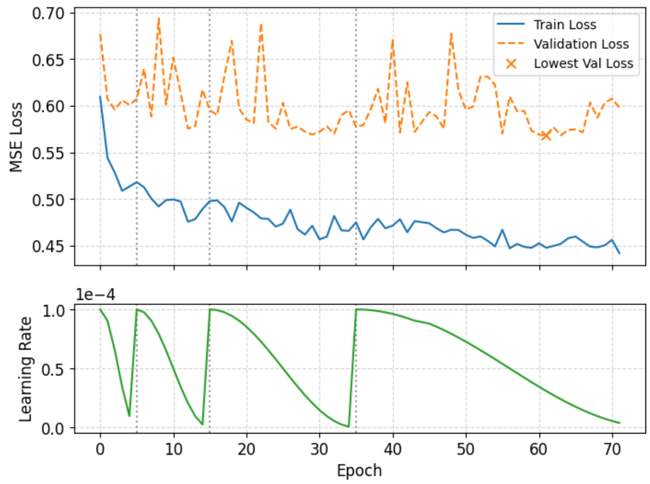

FIG. 13. Training and validation loss (top) and the corresponding cyclical learning rate schedule (bottom) across epochs. The minimum validation loss was achieved at epoch 61.

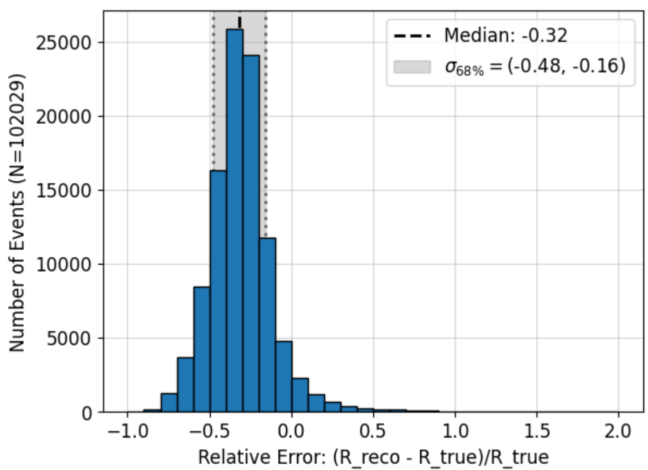

FIG. 14. Distribution of the relative error for the reconstructed radial distance, $(R_{\mathrm{reco}} - R_{\mathrm{true}})/R_{\mathrm{true}}$, evaluated on the independent test set ($N = 102,029$).

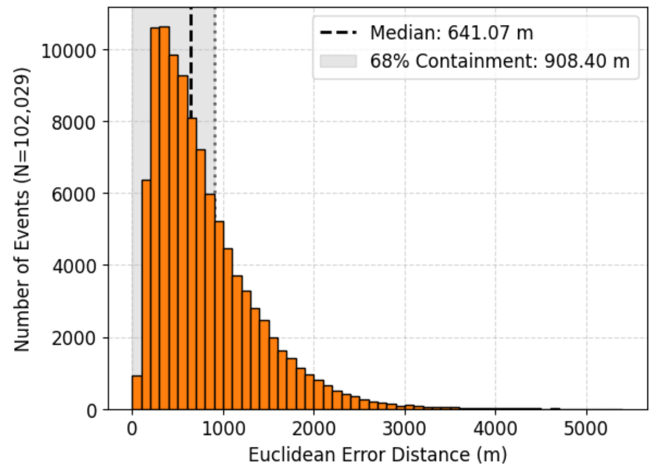

FIG. 15. Distribution of the absolute Euclidean error distance between the true and reconstructed 3D vertex positions. The right-skewed distribution highlights the core resolution of the model alongside extreme outlier events.

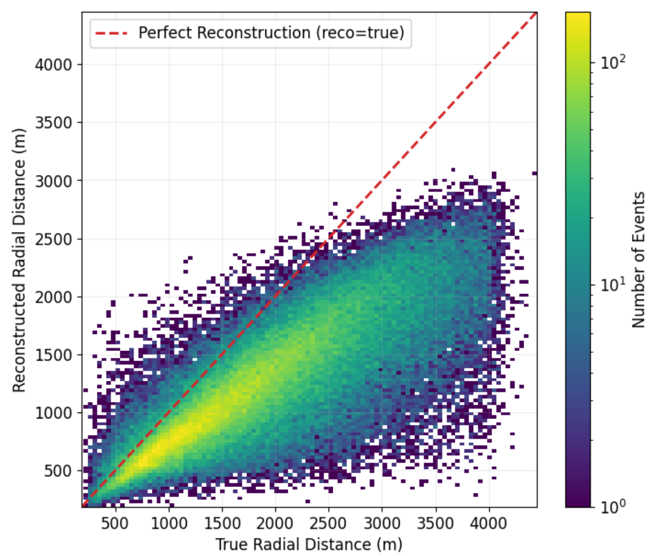

FIG. 16. 2D histogram comparing the reconstructed radial distance against the true radial distance for the test dataset. The red dashed line denotes the ideal baseline of perfect reconstruction ($R_{\mathrm{reco}} = R_{\mathrm{true}}$). The color axis represents the event density on a logarithmic scale. Reconstruction distance appears to decay at higher radii.

shrinks the 68% containment band of the relative radius error by nearly a factor of two, reducing it from 0.60 to 0.32 compared to the baseline.

V. CONCLUSION

Accurate reconstruction of the primary neutrino energy relies critically on determining the physical distance to the interaction vertex. The kilometer-scale attenuation length of radio waves in ice allows for the construction of massive, sparse detector arrays, like RNO-G, which present the opportunity to reconstruct events captured across multiple spatially separated stations. Current analytical algorithms rely heavily on single-station frameworks, largely discarding the rich geometric interplay of multi-station coincidence.

In this work, we presented a novel deep learning approach for multi-station vertex reconstruction in in-ice radio detectors. We trained a convolutional neural network on a simulated four-station RNO-G sub-array to predict the 3D interaction vertex directly from the time-binned Hilbert envelopes of the raw voltage traces.

Our results demonstrate the viability of leveraging multi-station coincidence events through machine learning. The CNN achieved a smaller relative radial error 1σ containment as compared to the baseline achieved by current state-of-the-art single-station statistical frameworks. While the model exhibits a distance-dependent underprediction bias for highly attenuated, low-SNR events, the improvement in precision demonstrates possible advantages of the network's multi-station feature extraction.

This work serves as a starting point for further algorithmic refinements. Specifically, future iterations will focus on mitigating the aforementioned spatial bias by exploring geometry-aware or relative-error loss functions. Because standard `MSELoss` evaluates absolute Cartesian distances, it fails to penalize near-field and far-field fractional errors equally. Implementing tailored loss functions will be critical for forcing the network to maintain its predictive power in low-signal, distant regimes. Ultimately, unlocking the full potential of machine learning for multi-station event reconstruction has the potential to shift the paradigm of neutrino energy reconstruction entirely.

ACKNOWLEDGMENTS

I would like to thank Prof. Brian Clark for the opportunity to work on this project and for his invaluable guidance throughout its development. I also extend my gratitude to my colleagues within the UMD IceCube group, especially Aishwarya Vijai and Taylor St. Jean, for their ongoing support. Finally, I am grateful to Dr. Christoph Welling for his assistance in understanding the analytical framework presented in [22], which ensured a faithful reproduction of the baseline model.

[1] G. T. Zatsepin and V. A. Kuz'min, Upper Limit of the Spectrum of Cosmic Rays, Soviet Journal of Experimental and Theoretical Physics Letters **4**, 78 (1966).

[2] K. Greisen, End to the cosmic-ray spectrum?, Phys. Rev. Lett. **16**, 748 (1966).

[3] G. Breit and J. A. Wheeler, Collision of two light quanta, Phys. Rev. **46**, 1087 (1934).

[4] R. J. Gould and G. P. Schréder, Opacity of the universe to high-energy photons, Phys. Rev. **155**, 1408 (1967).

[5] P. W. Gorham, P. Allison, O. Banerjee, *et al.*, Constraints on the diffuse high-energy neutrino flux from the third flight of anita, Physical Review D **98**, 10.1103/physrevd.98.022001 (2018).

[6] R. Abbasi, M. Ackermann, J. Adams, *et al.* (IceCube Collaboration), Search for extremely-high-energy neutrinos and first constraints on the ultrahigh-energy cosmic-ray proton fraction with icecube, Phys. Rev. Lett. **135**, 031001 (2025).

[7] J. A. Formaggio and G. P. Zeller, From ev to eev: Neutrino cross sections across energy scales, Reviews of Modern Physics **84**, 1307–1341 (2012).

[8] E. Zas, F. Halzen, and T. Stanev, Electromagnetic pulses from high-energy showers: Implications for neutrino detection, Phys. Rev. D **45**, 362 (1992).

[9] S. R. Klein, Icecube: A cubic kilometer radiation detector, IEEE Transactions on Nuclear Science **56**, 1141–1147 (2009).

[10] G. A. Askar'yan, Excess negative charge of an electron-photon shower and its coherent radio emission, Soviet Physics JETP **14**, 441 (1962), original Russian version: Zh. Eksp. Teor. Fiz. 41, 616-618 (1961).

[11] D. Saltzberg, P. W. Gorham, D. Walz, C. Field, R. Iverson, A. Odian, G. Resch, P. Schoessow, and D. Williams, Observation of the Askaryan effect, AIP Conf. Proc. **579**, 225 (2001).

[12] ARA Collaboration, N. Alden, S. Ali, *et al.*, Observation of in-ice askaryan radiation from high-energy cosmic rays (2025), arXiv:2510.21104 [astro-ph.HE].

[13] A. L. Connolly and A. G. Vieregg, Radio Detection of High Energy Neutrinos, in *Neutrino Astronomy: Current Status, Future Prospects* (World Scientific, 2017) pp. 217–244, review article, arXiv:1607.08232 [astro-ph.HE].

[14] J. Avva, J. M. Kovac, C. Miki, D. Saltzberg, and A. G. Vieregg, An in situ measurement of the radio-frequency attenuation in ice at summit station, greenland, Journal of Glaciology **61**, 1005–1011 (2015).

[15] I. Kravchenko *et al.*, In situ index-of-refraction measurements of the south polar firn with the rice detector, Journal of Glaciology **50**, 522–532 (2004).

[16] I. Kravchenko *et al.* (RICE Collaboration), Rice limits on the diffuse ultrahigh energy neutrino flux, Physical Review D **73**, 082002 (2006).

[17] P. Allison, S. Archambault, J. J. Beatty, *et al.*, Low-threshold ultrahigh-energy neutrino search with the askaryan radio array, Physical Review D **105**, 10.1103/physrevd.105.122006 (2022).

[18] S. W. Barwick, Arianna: A new concept for uhe neutrino detection, Journal of Physics: Conference Series **60**, 276–283 (2007).

[19] J. A. Aguilar *et al.* (RNO-G Collaboration), Design and Sensitivity of the Radio Neutrino Observatory in Greenland (RNO-G), JINST **16** (03), P03025, arXiv:2010.12279 [astro-ph.IM].

[20] M. G. Aartsen *et al.* (IceCube), The IceCube Neutrino Observatory: Instrumentation and Online Systems, JINST **12** (03), P03012, arXiv:1612.05093 [astro-ph.IM].

[21] J. C. Hanson and A. L. Connolly, Complex analysis of askaryan radiation: A fully analytic treatment including the lpm effect and cascade form factor, Astroparticle Physics **91**, 75–89 (2017).

[22] J. A. Aguilar, P. Allison, J. J. Beatty, *et al.*, Reconstructing the neutrino energy for in-ice radio detectors: A study for the radio neutrino observatory greenland (rno-g), The European Physical Journal C **82**, 10.1140/epjc/s10052-022-10034-4 (2022).

[23] C. Glaser, D. García-Fernández, A. Nelles, J. Alvarez-Muñiz, S. W. Barwick, D. Z. Besson, B. A. Clark, A. Connolly, J. C. Hanson, and Welling, Nuradiomc: simulating the radio emission of neutrinos from interaction to detector, The European Physical Journal C **80**, 10.1140/epjc/s10052-020-7612-8 (2020).

[24] D. Bergman, Telescope Array Combined Fit to Cosmic

Ray Spectrum and Composition, PoS **ICRC2021**, 338 (2021).

[25] R. Abbasi, M. Ackermann, J. Adams, *et al.* (IceCube Collaboration), Improved characterization of the astrophysical muon–neutrino flux with 9.5 years of icecube data, The Astrophysical Journal **928**, 50 (2022).

[26] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, Learning representations by back-propagating errors, Nature **323**, 533 (1986).

[27] K. He, X. Zhang, S. Ren, and J. Sun, Deep residual learning for image recognition (2015), arXiv:1512.03385 [cs.CV].