PROCEEDINGS OF SCIENCE

Deep-learning based event reconstruction for shallow in-ice UHE neutrino detectors

Sigfrid Stjärnholm,[a,*] **Christian Glaser,**[a] **Pierre Baldi,**[b] **Steven Barwick,**[c] **Oscar Ericsson,**[a] **Anton Holmberg,**[a] **Stephen McAleer**[b] **and Ting Wing Choi**[a]

[a] *Department of Physics and Astronomy, Uppsala University, SE-75237, Uppsala, Sweden*

[b] *Department of Information and Computer Science, University of California, CA 92697, Irvine, USA*

[c] *Department of Physics and Astronomy, University of California, CA 92697, Irvine, USA*

E-mail: sigfrid.stjarnholm@physics.uu.se, christian.glaser@physics.uu.se

We present an end-to-end reconstruction of the neutrino energy, direction and flavor from shallow in-ice radio detector data using deep neural networks (DNNs). For the first time, we were able to determine the neutrino direction with a few degrees resolution also for the complicated event class of electron neutrino charged-current interactions where the shower development is impacted by the LPM effect. This result highlights the advantages of DNNs to model the complex correlations in radio detector data. We will present an outlook of extending the model to predict the complex probability distribution of the neutrino direction using Normalizing Flows. Furthermore, we discuss how this work can be used for real-time alerts and an end-to-end detector optimization of, e.g., IceCube-Gen2 radio.

9th International Workshop on Acoustic and Radio EeV Neutrino Detection Activities - ARENA2022
7-10 June 2022
Santiago de Compostela, Spain

*Speaker

© Copyright owned by the author(s) under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

PoS(ARENA2022)019

1. Introduction

The shallow detector station layout that has previously been deployed in the ARIANNA experiment has shown great promise in the technology for detecting ultra-high-energy (UHE) neutrinos [1]. Future UHE neutrino experiments, such as IceCube-Gen2 and its radio component, which is in the technical design phase, still have the possibility of optimizing the detector station layout before it gets deployed. This highlights the need for reconstruction methods to assess station layouts and their performances, which is needed for a station layout optimization process. An end-to-end detector optimization procedure could increase the reconstruction performance and will allow us to build more capable detectors in the future. Reconstruction algorithms will also be needed when the first event has been recorded from the upcoming detectors, but the immediate need lie in the optimization of the detector station layout.

Previous methods have shown that reconstruction of the neutrino's direction and energy is possible, but only for a subset of events that do not originate from ν_e-CC interactions [2, 3]. Furthermore, the reconstruction techniques have taken a lot of work to develop and are often hard to generalize between different tasks. Here, we use advances in the field of machine learning and neural networks to extract the neutrino's energy and direction from simulated UHE neutrino detector data. More details can be found in [4]. Here, we summarize the results of the reconstruction capabilities of deep-learning models for the neutrino direction and energy, and give an outlook of the prospects of also measuring the neutrino flavor. We report on future improvements of our work to use *Normalizing Flows* to predict event-by-event uncertainties. In this work, we focus on shallow detector station layouts, but the methods can also be applied more broadly for in-ice radio detection.

2. Dataset Simulation

The work in this report is based on a detector station layout that used a set of shallow in-ice antennas located at the South Pole. The detector station consists of 4 LPDA antennas buried roughly 2 meters below the surface, along with a dipole antenna buried 15 meters below the surface. This results in 5 measurements of the short radio pulses that emerge from a neutrino interaction in the ice due to the Askaryan effect.

The data used in this work was simulated using the *NuRadioMC* [5] library, which simulates the Askaryan radiation generated by the neutrino interaction, the propagation of the emission through the ice, and the simulation of the detector response and trigger condition. The Askaryan signal that is generated depends on the emission model used. Two emission models will be used in this work, one of which uses a frequency domain parameterization [6], and the dataset that emerges from this model will be denoted as *Alvarez2009 (had.)*. This dataset only contains hadronic showers. The second emission model uses a semi-analytical calculation along with a library of time-domain charge-excess profiles of the cascades [7]. From this model, two different datasets are generated: one dataset with only hadronic showers, denoted as *ARZ2020 (had.)*, and one dataset with hadronic showers and electromagnetic showers emerging from electron neutrino charged-current interactions (ν_e-CC), denoted as *ARZ2020 (had. + EM)*. Events with electromagnetic showers have previously not shown great reconstruction performance with classical methods due to the large stochastic

PoS(ARENA2022)019

variability that comes from the LPM effect. The neutrinos used are generated with energies within the interval 10^{17} to 10^{19} eV, and roughly 20 million events are generated in total.

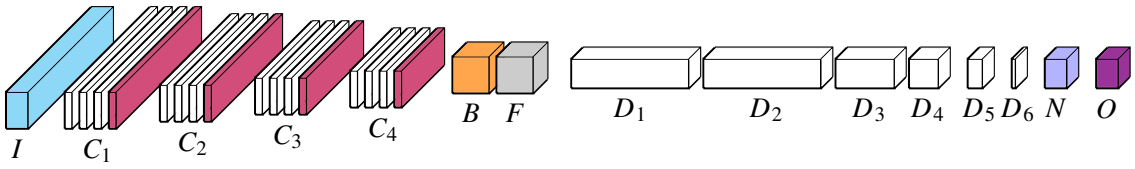

Figure 1: Structure of the neural network that was developed in this work. The labels are specified in section 3.

3. Neural Network Architecture

The network that was developed in this work is presented in Fig. 1, the full details of which can be found in [4]. The network consists of an input layer, I, at which the antenna data is fed to the network. The network then consists of 4 convolutional blocks C_i, each of which have 3 convolutional layers followed by a pooling layer that uses Average Pooling. A batch-normalization layer B and a flattering layer F are present after the convolutional blocks, which helps to stabilize the optimisation procedure [8] and make the network dimensions match what the following layers, the dense layers D_i, are expecting. The dense layers, initially with a size of 1024 nodes, are descending in node count of each layer by a factor of 2, until the last dense layer, which has either 3 nodes for direction reconstruction (x, y, z coordinates), or 1 node for energy reconstruction (E_{shower}), or 2 nodes for flavor reconstruction (one-hot encoding of ν_e-CC versus the rest). For the direction reconstruction, the dense layers are followed by a normalization layer which ensures that the output has a $L2$-norm of 1, i.e., that $x^2 + y^2 + z^2 = 1$. Finally, the output layer yields the prediction of the neural network, which is either the neutrino direction in Cartesian coordinates, the logarithm of the shower energy, or the one-hot encoding of the flavor.

The network was built and trained using the frameworks TensorFlow and Keras [9, 10]. The data, being too large to load into memory in full, was loaded using a custom *tensorflow.data* pipeline, and the training was performed on a NVIDIA Quadro RTX 6000 GPU. A total of 300,000 events per dataset were reserved as a test dataset, while the training and validation split for the rest of the data for each dataset was 87 % versus 13 %. The optimization was done using the *Adam* optimizer [11], using a learning rate of $5 \cdot 10^{-5}$. The loss used is mean absolute error (MAE). The network parameters such as layer sizes, amount of layers, the loss function, activation functions and other hyperparameters were optimized using a trial and error approach with a grid-like search. The same network, apart from small adjustments to fit the task (such as using a L2-normalization layer for direction), was used for the direction, energy, and flavor reconstruction. This points to the fact of great generalizability of deep neural networks, as 3 different tasks can find a solution in very similar architectures.

4. Direction Reconstruction

Using the neural network model presented in section 3, the incoming neutrino direction was predicted using the raw waveforms of the antenna signals. The results for the direction reconstruction

PoS(ARENA2022)019

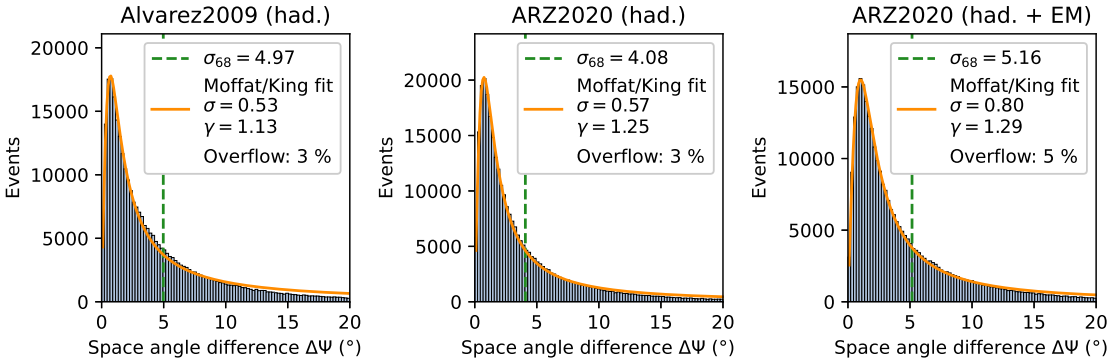

Figure 2: Space angle difference for the *Alvarez2009 (had.)*, *ARZ2020 (had.)*, and *ARZ2020 (had. + EM)* datasets. The point spread function is modeled by the Moffat/King function. The 68 % quantile is marked by a vertical dashed line. The legend specifies the overflow, which is the fraction of events with a space angle difference greater than 20°.

for the three different datasets are presented in Fig. 2 and Fig. 3. The space angle difference ($\delta\Psi$) is the space angle between the true and the predicted direction for a single event. The 68 % quantile (σ_{68}) is defined as being the space angle difference at which 68 % of the test dataset has a space angle difference at or below σ_{68}.

The results shown in Fig. 2 point to the advantages of using neural networks for reconstruction of event properties. The *ARZ2020 (had. + EM)* dataset, which previously has been difficult to reconstruct, shows only a small penalty to the performance of the direction reconstruction when compared to the *ARZ2020 (had.)* dataset. In Fig. 3, the distribution of the reconstruction performance versus neutrino energy is shown, and it is from this figure possible to see that high energy events are reconstructed with better performance when compared to low energy events. This can partly be explained by the fact that the training dataset contains more events at higher energies, as those events are more likely to trigger the detector. Also, the dependence on signal-to-noise radio (SNR) shows that events with higher SNR are reconstructed with better precision, something which is reasonable given that the background noise will have less impact on the signal.

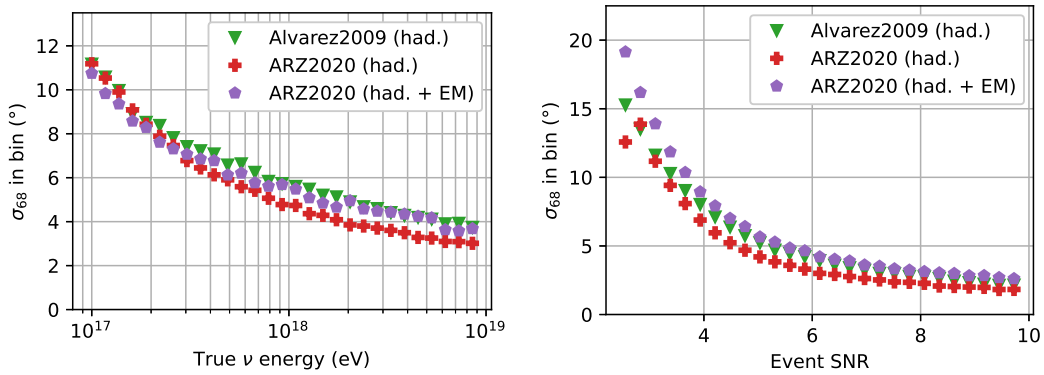

Figure 3: Angular resolution σ_{68} and its dependence on the neutrino energy and signal-to-noise ratio for the *Alvarez2009 (had.)*, *ARZ2020 (had.)*, and *ARZ2020 (had. + EM)* datasets.

PoS(ARENA2022)019

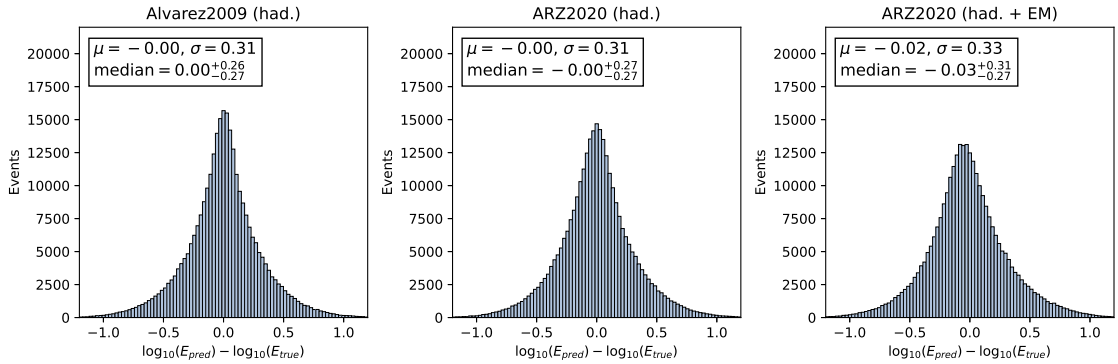

Figure 4: Energy difference histograms for the *Alvarez2009 (had.)*, *ARZ2020 (had.)*, and *ARZ2020 (had. + EM)* datasets.

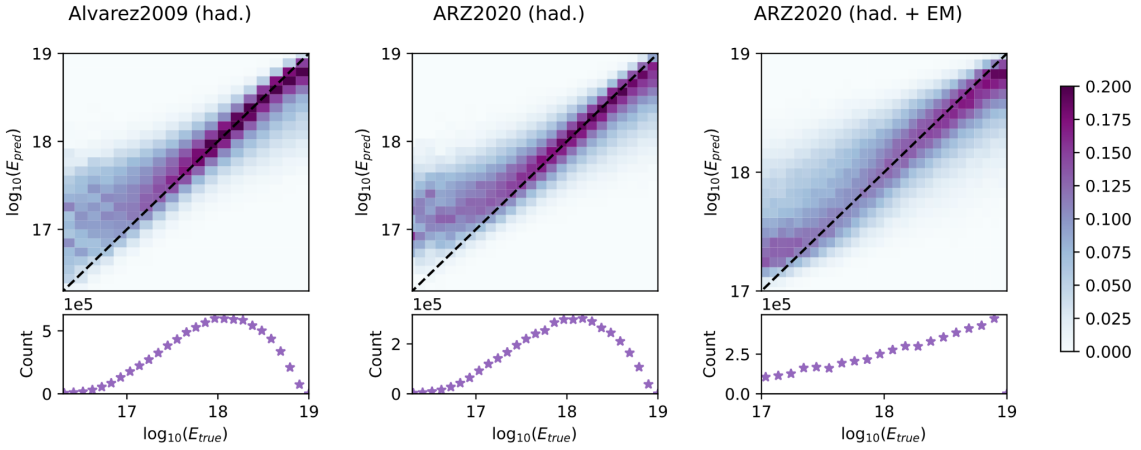

Figure 5: 2D heatmaps of the predicted versus true neutrino shower energies for the *Alvarez2009 (had.)*, *ARZ2020 (had.)*, and *ARZ2020 (had. + EM)* datasets. The columns are normalized to have a sum of 1. The count of events as a function of shower energy are shown in the subplots below the heatmaps.

5. Energy Reconstruction

The energy reconstruction was done using a similar network to that of the direction reconstruction, and the results are presented in Fig. 4 and Fig. 5. Because the shower energies range over several orders of magnitude, the network is trained to reconstruct the logarithm of the shower energy.

The results in Fig. 4 show that the shower energy can be reconstructed within a factor of 2 (0.3 in $log_{10}(E)$). This uncertainty is comparable to the irreducible uncertainty from inelastic fluctuations [12]. Also notable is that this performance metric holds for all 3 datasets, which means that even the complicated *ARZ2020 (had. + EM)* dataset can be reconstructed with similar accuracy. In Fig. 5, the heatmaps of predicted versus true shower energy are presented, along with the count of events versus the shower energy. These heatmaps show that for all 3 datasets, there is a bias at low energies where the model predicts too high values for the energy, and a slight bias at high energies where the model predicts too low values for the energy. This is partly due to the underrepresentation of these

PoS(ARENA2022)019

events in the training dataset which we plan to address in future work. See also [4] for an extended discussion.

6. Flavor Reconstruction

The differences between ν_e-CC interactions and all other neutrino interactions provide a signature to obtain flavor sensitivity. ν_e-CC interactions produce both a hadronic shower (from the breakup of the nucleon) and an electromagnetic shower (from the created electron/positron). Furthermore, the electromagnetic shower is affected by the LPM effect, which delays the shower development leading to interference of both showers, which alters the radio pulses. All other neutrino interactions only initiate a hadronic shower. In Fig. 6, the results from the flavor reconstruction are shown. The figures show confusion matrices in 4 different neutrino energy intervals, with the true flavor label on the vertical axis and the predicted flavor label on the horizontal axis. The performance on reconstructing the neutrino flavor increases greatly with increasing energy. This can be explained by the fact that the LPM effect affects high-energy events to a greater extent than low-energy events, which allows for greater differentiation between ν_e-CC events and non-ν_e-CC events.

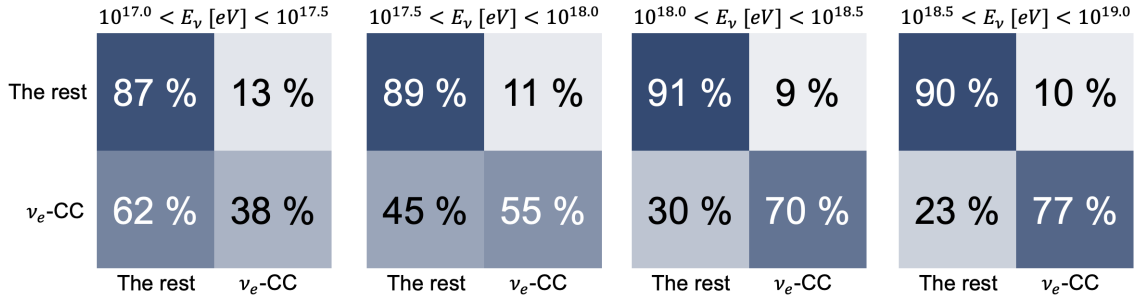

Figure 6: Confusion matrices for the predicted versus true flavor label for 4 different neutrino energy intervals. The vertical axis specifies the true flavor, and the horizontal axis specifies the predicted flavor.

7. Outlook

One important aspect of a successful reconstruction is uncertainty estimation. This is often as important as the nominal value itself, for example, in real-time alerts where it is not only important to know the best fitting neutrino arrival direction but also the sky region from which the neutrino could have originated. A simple approach is to let the network predict also the sigma parameter of a Gaussian distribution. However, often the uncertainties are non-Gaussian and show strong correlations, which is the case for direction reconstruction. One potential solution to this is using normalizing flows [13] to predict event-by-event uncertainties. An example of this is seen in Fig. 7, where the normalizing flows network predicts the probability distribution function of the energy, yielding an estimate of the uncertainty of a particular event [14].

As mentioned previously, there is still time to improve and optimize the station layout for the radio component of IceCube-Gen2. End-to-end detector optimization is currently impossible due

PoS(ARENA2022)019

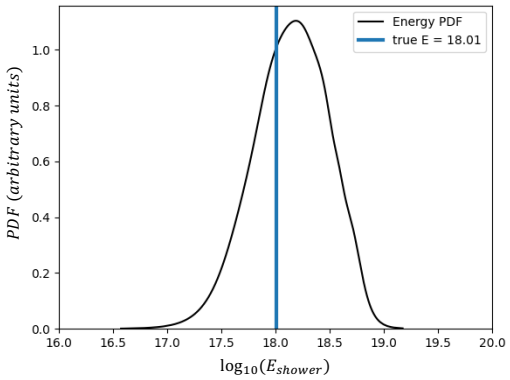

Figure 7: An example PDF of the energy distribution generated using normalizing flows. The x-axis is the energy. From the PDF, an uncertainty estimation can be extracted as the width of the distributing. The true energy of the event is marked with the vertical line.

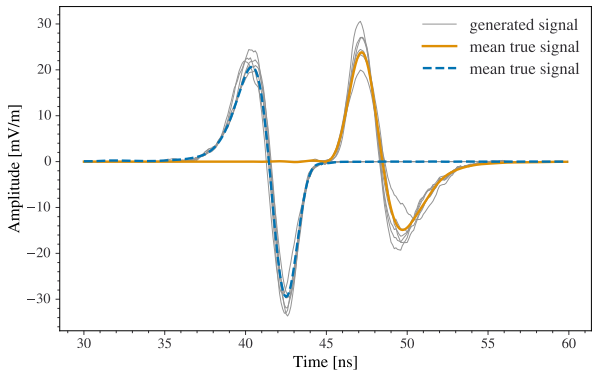

Figure 8: An example of the emitted Askaryan signals generated by a differential surrogate model consisting of a Generative Adversarial Network (GAN). The true signals are marked with the colorized dashed and solid lines, with 10 realizations of the generated signals shown in gray.

to the very time-consuming Monte Carlo (MC) simulation of data. Using the advances in neural networks, this might be overcome by replacing the complete pipeline with fast and differential surrogate models. We achieved promising results for two of the most time-critical parts, namely, the generation of the Askaryan signal as well as the ray tracing. As a function of energy and viewing angle from the Cerenkov cone, a Generative Adversarial Network (GAN) generates the Askaryan signal a lot quicker than it currently takes. An example is shown in Fig. 8, where 10 realizations of the Askaryan signal are shown for two different viewing angles, along with the true signal. Also, see [15] for more on this matter.

Acknowledgements

We thank all developers of the NuRadioMC code for enabling the creation of the training datasets. The computations and data handling were enabled by resources provided by the Swedish

PoS(ARENA2022)019

National Infrastructure for Computing (SNIC) at UPPMAX partially funded by the Swedish Research Council through grant agreement no. 2018-05973. The work of SM and PB in part supported by NSF grant NRT 1633631 to PB.

References

[1] A. Anker, S. Barwick, H. Bernhoff, D. Besson, N. Bingefors, G. Gaswint et al., *Targeting ultra-high energy neutrinos with the arianna experiment*, *Advances in Space Research* **64** (2019) 2595.

[2] G. Gaswint, *University of California, Irvine*, Ph.D. thesis, University of California, Irvine, 2021.

[3] RNO-G collaboration, *Direction Reconstruction for the Radio Neutrino Observatory Greenland (RNO-G)*, *PoS* **ICRC2021** (2021) 1026.

[4] C. Glaser, S. McAleer, S. Stjärnholm, P. Baldi and S. Barwick, *Deep-learning-based reconstruction of the neutrino direction and energy for in-ice radio detectors*, *Astroparticle Physics* (2022) 102781.

[5] C. Glaser, D. García-Fernández, A. Nelles, J. Alvarez-Muñiz, S.W. Barwick, D.Z. Besson et al., *NuRadioMC: simulating the radio emission of neutrinos from interaction to detector*, *The European Physical Journal C* **80** (2020) .

[6] J. Alvarez-Muñiz, C. James, R. Protheroe and E. Zas, *Thinned simulations of extremely energetic showers in dense media for radio applications*, *Astroparticle Physics* **32** (2009) 100.

[7] J. Alvarez-Muñ iz, A. Romero-Wolf and E. Zas, *Čerenkov radio pulses from electromagnetic showers in the time domain*, *Physical Review D* **81** (2010) .

[8] S. Santurkar, D. Tsipras, A. Ilyas and A. Madry, *How does batch normalization help optimization?*, 1805.11604.

[9] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro et al., *TensorFlow: Large-scale machine learning on heterogeneous systems*, 1603.04467.

[10] F. Chollet et al., "Keras." https://keras.io, 2015.

[11] D.P. Kingma and J. Ba, *Adam: A method for stochastic optimization*, 1412.6980.

[12] A. Anker et al., *Neutrino vertex reconstruction with in-ice radio detectors using surface reflections and implications for the neutrino energy resolution*, *JCAP* **11** (2019) 030 [1909.02677].

[13] T. Glüsenkamp, *Unifying supervised learning and vaes – automating statistical inference in (astro-)particle physics with amortized conditional normalizing flows*, 2008.05825.

[14] T.W. Choi, *Reconstruction of the energy of neutrinos with neural networks: Event-by-event uncertainty estimation*, *Master Thesis, Uppsala University* (2022) .

[15] T. Dorigo, A. Giammanco, P. Vischia, M. Aehle, M. Bawaj, A. Boldyrev et al., *Toward the end-to-end optimization of particle physics instruments with differentiable programming: a white paper*, 2203.13818.

PoS(ARENA2022)019