

Deep learning reconstruction of the neutrino energy with a shallow Askaryan detector

Christian Glaser,^{a,*} Stephen McAleer,^b Pierre Baldi^b and Steven Barwick^c

^a*Uppsala University Department of Physics and Astronomy, Uppsala, SE-752 37, Sweden*

^b*Department of Computer Science, University of California, Irvine, CA 92697, USA*

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

E-mail: christian.glaser@physics.uu.se, smcaleer@uci.edu

Cost effective in-ice radio detection of neutrinos above a few 10^{16} eV has been explored successfully in pilot-arrays. A large radio detector is currently being constructed in Greenland with the potential to measure the first cosmogenic neutrino, and an order-of-magnitude more sensitive detector is being planned with IceCube-Gen2. We present the first end-to-end reconstruction of the neutrino energy from radio detector data. NuRadioMC was used to create a large data set of 40 million events of expected radio signals that are generated via the Askaryan effect following a neutrino interaction in the ice for a broad range of neutrino energies between 100 PeV and 10 EeV. We simulated the voltage traces that would be measured by the five antennas of a shallow detector station in the presence of noise. We designed and trained a deep neural network to determine the shower energy directly from the simulated experimental data and achieve a resolution better than a factor of two ($\text{STD} < 0.3$ in $\log_{10}(E)$) which is below the irreducible uncertainty from inelasticity fluctuations. We present the model architecture and study the dependence of the resolution on event parameters. This method will enable Askaryan detectors to measure the neutrino energy.

37th International Cosmic Ray Conference (ICRC 2021)
July 12th – 23rd, 2021
Online – Berlin, Germany

*Presenter

1. Introduction

The detection of ultra-high-energy (UHE) neutrinos is a key to solve the 100-year-old mystery of the origin of cosmic rays and one of the crucial milestones for astroparticle physics [1]. The only cost-efficient way to measure these UHE neutrinos above 30 PeV of energy is via a sparse array of radio antenna stations installed, for instance, in the Arctic or Antarctic ice [2–5]: A neutrino interaction in the ice generates a few-nanoseconds long radio flash that can be detected from kilometer-long distances. Because of the small flux, no UHE neutrino has been observed yet, but the technology has already been shown to work reliably with small test-bed arrays. Much larger detectors are being constructed at the moment and even larger arrays are proposed for the next decade.

This work focuses on a shallow station design as shown in Fig. 1 that has been explored by the ARIANNA test-bed detector on the Ross Ice Shelf and at the South Pole [5]. Each station consists of 4 LPDA antennas installed at a depth of just a few meters below the snow surface, and 1 dipole antenna installed at a depth of 10 m in a narrow borehole. These antennas observe the ice below for neutrino interactions. The dipole antennas was added to help with the reconstruction of the neutrino properties. At this depth, the antenna will observe two signals, one from a direct path to the antenna, and another delayed signal from a reflection off the surface. The time difference between these two signals provides information about the distance to the neutrino interaction which is important to estimate the neutrino energy [6].

Shallow detector stations are part of the RNO-G array that is currently being constructed in Greenland [7], they make up the proposed ARIANNA-200 detector [8, 9] which will consist of 200 autonomous radio detector stations installed on the Ross Ice Shelf in Antarctica, and are an integral part of the plans for IceCube-Gen2 [10].

A missing piece for the success of radio-based detection of neutrinos is the reconstruction of the neutrino properties from the detectable radio flashes. Despite several years of research, no end-to-end reconstruction has been developed so far. This is largely due to the complexity of the problem: All information of the neutrino is compressed in a few nanosecond long radio flash that is observed in just a few antennas. The traditional approach is to extract single parameters, such as the amplitude or the slope of the frequency spectrum, and correlate them with the neutrino properties of interest (see [11] for an overview of the reconstruction strategies). This is a very time consuming process and it often does not take into account all available information.

The development of the NuRadioMC code which precisely simulates the expected signals from neutrinos [12] as well as the corresponding detector response [13], allows us to take a different approach [14]: training deep-neural networks to extract the neutrino properties of interest directly from the raw data. In this work we focus on the most challenging property, the neutrino energy.

2. Data set generation

We use the NuRadioMC code [12] to create a large data set of 39.3 million recorded neutrino event which corresponds to a data volume of 772 GB. We simulate a shallow radio detector station at the ARIANNA site on the Ross ice shelf, i.e., an ice shelf with a thickness of 576 m with an average attenuation length of approx. 500 m and where the ice-water interface at the bottom of the

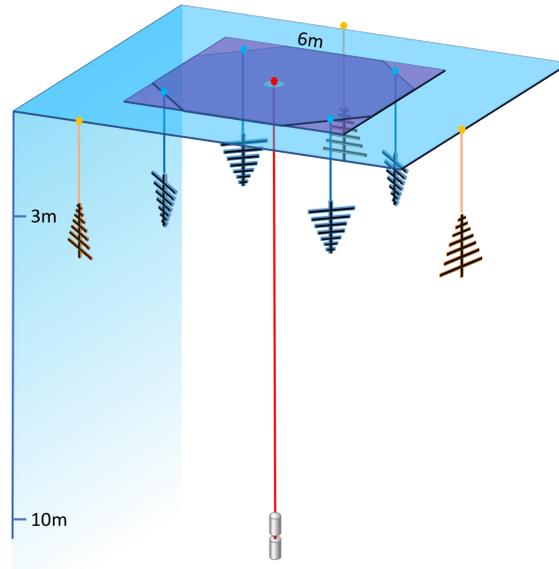


Figure 1: Schematic illustration of the shallow radio detector station. The three upward facing LPDA antennas are for cosmic-ray detection and are ignored in this work. Fig. from [8].

ice shelf reflects radio signals back up with high efficiency. We generate neutrino interactions with fixed energies from 10^{17} eV to 10^{20} eV with a step size of 0.1 in the logarithm of the energy. The neutrino interactions are distributed uniformly in the ice around the detector with random incoming direction.

For this work we limit ourselves to the simulation of hadronic showers, i.e., neutral current interactions of all neutrino flavors, and charged-current interactions of muon and tau neutrinos. The electron created in electron neutrino charged-current interactions will, in addition, initiate an electromagnetic (EM) shower. At high energies, the development of EM showers is effected by the LPM effect which complicates the reconstruction of the shower properties. Reconstruction of EM showers will be studied in a forthcoming publication.

The observed radio signals only depend on the shower properties. Therefore, we will train the deep neural network to determine the energy of the hadronic shower which has a deterministic relation to the observed signal. The amount of energy transferred from the neutrino into the hadronic shower is a stochastic process and can be estimated from theory [6]. The stochastic process also introduces a irreducible uncertainty on the conversion from hadronic shower energy back to the neutrino energy which amounts to a factor of two. This provides a target resolution for the shower energy.

We simulate the resulting radio signals in the five antennas of the shallow detector station by convolving the electric-field pulses with the antenna response and we approximate the rest of the signal chain with an 80–800 MHz bandpass filter. An event is recorded if the signal pulse crossed a high and a low threshold of three times the RMS noise within 5 ns. At such a low trigger threshold, noise fluctuations can fulfil the trigger condition at a non-negligible rate. Therefore, we make sure that the signal before adding noise has at least an amplitude of two times the RMS noise.

The number of events per energy bin is adjusted such that the number of recorded events is approximately the same in each energy bins. Especially at low energies, this requires the simulation

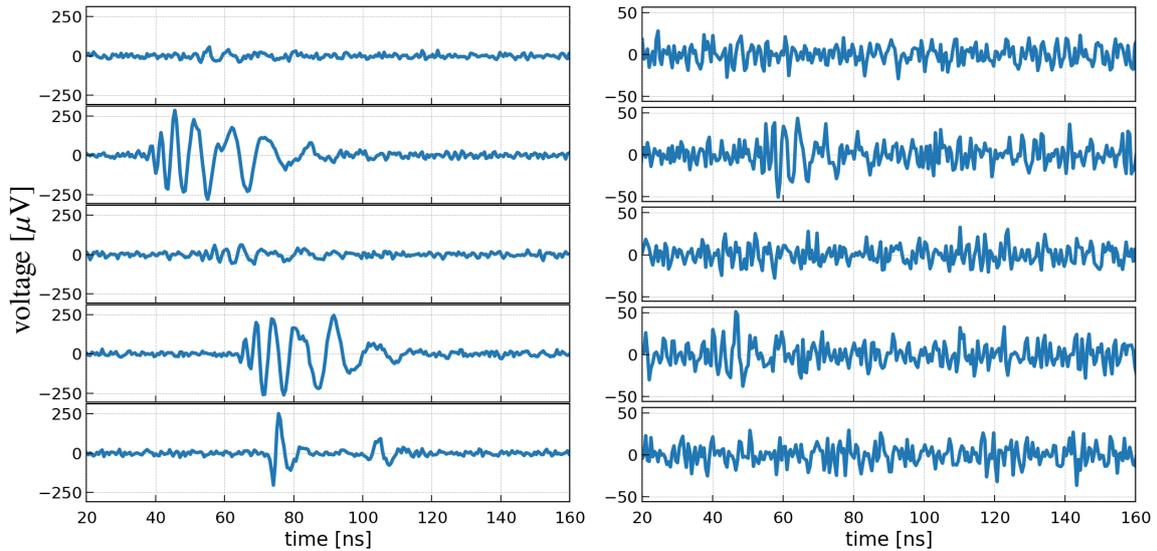


Figure 2: Two examples of the training data set. The traces show the recorded voltages in the 5 antennas as a function of time. The upper 4 panels are the signal in the 4 LPDAs, and the bottom panel shows the signal in the dipole antenna. Left: example at particularly high signal-to-noise ratio, where the direct and reflected pulse is clearly visible in the dipole antenna. Right: a more typical example at low signal-to-noise ratio.

of orders of magnitudes more events which puts a heavy burden on computing resources.

The training data consists of the recorded voltages as a function of time of the 5 antennas. We simulated a length of the time trace of 256 ns with a time binning of 0.5 ns which results in 512 samples per antenna. The shape of one event is (5, 512). Two examples are shown in Fig. 2.

3. Deep learning architecture

After some experimentation, we used a neural network that consists of five convolutional layers followed by six fully connected layers (see Fig. 3 for an overview of the network structure). We decided to use a convolutional neural network to identify and characterize the signal pulses that are invariant in time (cf. Fig. 2). Because the relative time differences of the signals carry information, we apply the convolution filters independently to each antenna, i.e., the input shape is (5, 512, 1) corresponding to 5 antennas and 512 samples. The last component is 1 as the only available information from each antenna is the voltage amplitude (unlike three color channels in the case of images). Each convolutional layer has 100 filters, a kernel of size (1, 10) and stride of size (1, 2). The fully-connected layers have sizes of 2560, 2048, 1024, 256, 128. Each layer (both fully-connected and convolutional) has a relu activation. We used Keras to implement this neural network. The optimizer was Adam with a learning rate of 0.0001 and using the mean squared error (MSE) of the logarithm of the shower energy as objective function. We used Sherpa [15] for further hyperparameter optimization.

We reserved 300,000 events for an independent test data set. We used 80% of the remaining data (31.1 million events) for the training data set and 20% (7.8 million events) for the validation data set. The training data set is so large that it does not fit into memory. To not slow down the training process due to io and data preprocessing times, we developed a data pipeline based on the

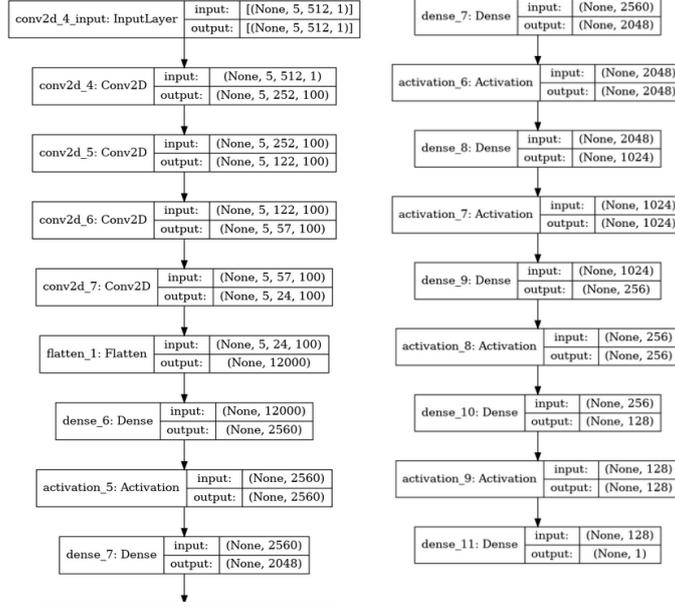


Figure 3: Overview of the network architecture. The tuples indicate the input and output dimensions of the data.

tensorflow.data class that preloads and preprocesses the data on several cores. We used a Nvidia Quadro RTX 6000 GPU to perform the training. We trained the network for 25 epochs after which no further improvement in the validation loss was observed.

4. Results

We evaluated the performance of the DNN using an independent test data set of 100,000 events and present the results in Fig. 4. The scatter plot shows a good correlation between the true and predicted energy over almost 4 orders of magnitude. We quantify the resolution by making a histogram of the logarithm of the difference of the true and predicted shower energy. Using the logarithm leads to a symmetric distribution compared to if a linear scale was used. The standard deviation (which corresponds to the square root of the MSE used during training) is 0.26 which corresponds to a factor of 1.8 on a linear scale. The distribution shows non-Gaussian tails which increases the standard deviation beyond the 68% interval of +0.23 and -0.22 which corresponds to a factor of 1.7 on a linear scale. Thus, we achieved the target resolution of being better than the inelasticity limit of a factor of two.

We inspected how the resolution depends on the event parameters. Unsurprisingly, the resolution improves with increasing signal-to-noise ratio. For signals at the trigger threshold, the resolution is 0.34 and reduces to below 0.2 for high signal-to-noise ratios. More interestingly is the shower energy resolution on the shower energy itself which is shown in Fig. 5. The resolution initially improves up to an energy of 2×10^{17} eV, then worsens towards 10^{18} eV and then improves again towards higher energies. Naively we would expect an improvement of the resolution with energy, because a higher shower energy leads on average also to larger signal-to-noise ratios. However, the Ross Ice Shelf has a special property that radio signals get reflected at the bottom of the ice

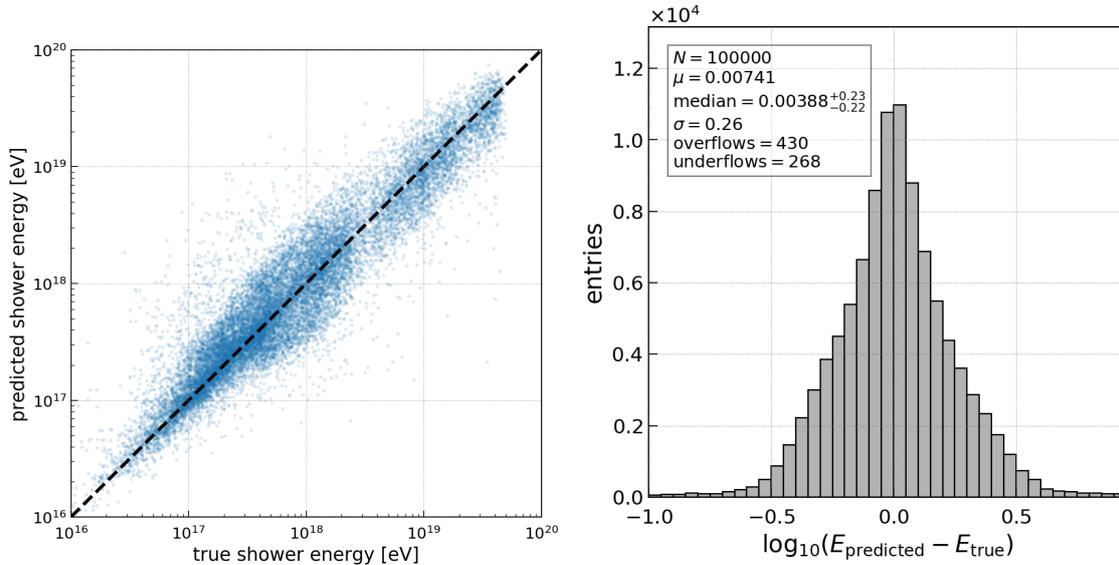


Figure 4: (left) Scatter plot of predicted vs. true shower energy using the first 20,000 events of the test data set. (right) Histogram of the difference of the logarithm of the predicted and true shower energy using 100,000 events of the test data set.

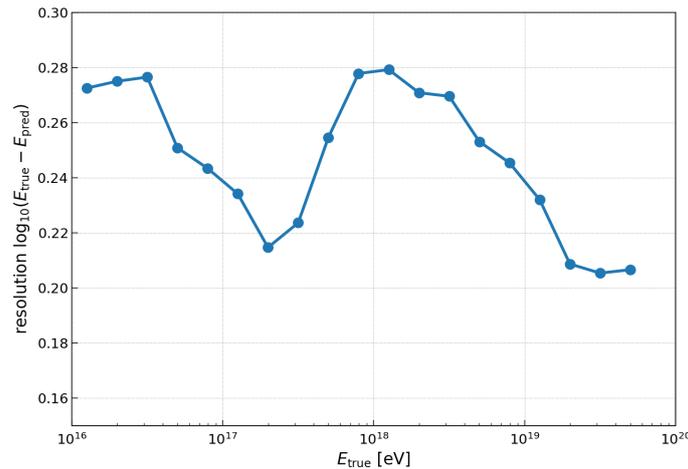


Figure 5: Energy dependence of the energy resolution. The resolution is calculated as the standard deviation of $\log_{10}(E_{\text{true}} - E_{\text{predicted}})$ per bin of true energy. A value of 0.3 corresponds to a factor of two on a linear scale, and a value of 0.2 corresponds to a 60% resolution non a linear scale.

shelf which is the likely reason for this behaviour. If the shower energy is large enough, the emitted radio signal is large enough to propagate downwards to the bottom of the ice shelf, and back up to the receivers. Around 10^{18} eV is exactly the transition region where more and more events are observed via a reflected trajectory.

The large amount of available training data also allowed us to test the performance of the DNN for different training data sizes. As clearly visible in Fig. 6, the mean squared error decreases with increasing training data size as expected. The plot also suggests that we could reach an even better performance by further increasing the training data set.

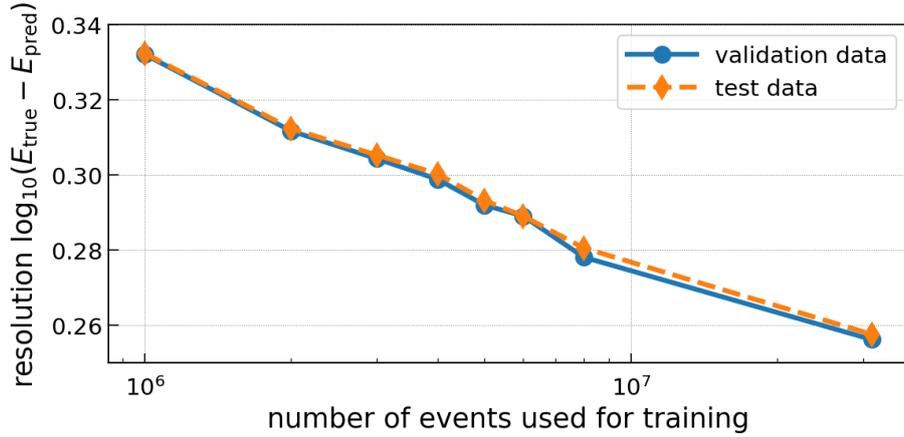


Figure 6: The performance of the DNN reconstruction as a function of the size of the training data set.

5. Conclusions

We presented the first end-to-end reconstruction of the neutrino energy from radio detector data using a deep neural network (DNN). The only information available to the DNN was the raw voltage output of the antenna receivers. The DNN achieved an average resolution of a factor of 1.8 in shower energy which is below the intrinsic uncertainty from inelasticity fluctuations. The dependence on the size of the training data set suggests that an even better resolution can be achieved with an increased training data set.

Although these are promising results, this work is just the beginning of exploiting deep learning for event reconstruction of radio detector data. A critical aspect is the accuracy of the training data. With the development of NuRadioMC and its ongoing improvements (see e.g. [16, 17]) a lot of work has already been done but more work is needed to ensure the validity of the data set, i.e., that it correctly resembles real data. We believe that it is also crucial to compare deep-learning methods with traditional methods on an event-by-event basis to validate the performance such as the work of [18]. We plan to extend the data set to also contain electron neutrinos, which will create both electromagnetic and hadronic showers, further complicating the reconstruction. Furthermore, the DNN needs to be robust against experimental uncertainties which we can test by creating test data sets that are varied according to experimental uncertainties. And last, an event-by-event uncertainty estimation is needed.

The work presented here also opens up new ways to optimize future detectors. Currently, an optimization of reconstruction performance is often not possible as it takes too much time to adapt traditional reconstruction techniques to changing detector layouts. With deep-learning-based reconstruction, different detector designs can be quickly evaluated in terms of reconstruction performance using only simulated data.

Acknowledgements

We thank the ARIANNA collaboration for useful feedback on this work. We thank the University of California, Irvine for providing HPC computing resources to produce the large

training data set. The work of SM and PB in part supported by NSF grant NRT 1633631 to PB.

References

- [1] M. Ackermann *et al.* *Bull. Am. Astron. Soc.* **51**, 185 (2019) .
- [2] **ARA** Collaboration, P. Allison *et al.* *Phys. Rev. D* **102** (2020) 043021.
- [3] **ARIANNA** Collaboration, A. Anker *et al.* *Advances in Space Research* **64** (2019) 2595 – 2609.
- [4] M. G. Aartsen *et al.* *Physical Review Letters* **117** (2016) 241101.
- [5] **ARIANNA** Collaboration, A. Anker *et al.* *JCAP* **03** (2020) 053.
- [6] **ARIANNA** Collaboration, A. Anker *et al.* *JCAP* **11** (2019) 030.
- [7] **RNO-G** Collaboration, J. A. Aguilar *et al.* *JINST* **16** P03025 2021 .
- [8] **ARIANNA** Collaboration, A. Anker *et al.* *arXiv:2004.09841* (2020) .
- [9] S. Barwick for the ARIANNA collaboration *PoS ICRC2021* (these proceedings) 1067.
- [10] S. Hallmann for the IceCube-Gen2 collaboration *PoS ICRC2021* (these proceedings) 1183.
- [11] C. Glaser for the ARIANNA collaboration *PoS(ICRC2019)899* .
- [12] C. Glaser *et al.* *European Physical Journal* **C80** (2020) 77.
- [13] C. Glaser *et al.* *European Physical Journal* **C79** no. 6, (2019) 464.
- [14] P. Baldi, *Deep Learning in Science*. Cambridge University Press, Cambridge, UK, 2021.
- [15] L. Hertel, J. Collado, P. Sadowski, J. Ott, and P. Baldi *SoftwareX* (2020) . Also arXiv:2005.04048. Software available at: <https://github.com/sherpa-ai/sherpa>.
- [16] C. Glaser, D. García-Fernández, and A. Nelles *PoS ICRC2021* (these proceedings) 1231.
- [17] B. Oeyen, I. Plaisier, C. Glaser, A. Nelles, and T. Winchen *PoS ICRC2021* (these proceedings) 1027.
- [18] G. Gaswint, *University of California, Irvine*. PhD thesis, 2021.