PROCEEDINGS OF SCIENCE

Neural networks for TeV cosmic electrons identification on the DAMPE experiment

D. Droz,[a,*] **A. Tykhonov,**[a] **X. Wu**[a] **and M. Deliyergiyev**[a]

[a] *University of Geneva, CH-1205 Geneva, Switzerland*

E-mail: david.droz@unige.ch, andrii.tykhonov@unige.ch

The past decades have witnessed the deployment of a new generation of cosmic ray (CR) observatories with unprecedented sensitivity and complexity, pushing towards ever-higher energies. To face the challenges of the multi-TeV domain, such instruments must be accompanied by equally powerful analysis techniques, able to exploit as much information as available. For example, the machine learning tool set may provide the needed techniques. We present a neural network optimised for the identification of multi-TeV electrons on DAMPE, a calorimetric spaceborne CR observatory with among other objectives the measurement of cosmic electrons up to 10 TeV. This constitutes a particularly challenging endeavour due to both the soft electron spectrum and the large proton background. The developed neural network significantly outperforms the more traditional cut-based approach, achieving a much lower proton contamination in the multi-TeV domain with a high signal efficiency, and retains its accuracy when transposed from Monte Carlo to real data.

**** The European Physical Society Conference on High Energy Physics (EPS-HEP2021), ****
**** 26-30 July 2021 ****
**** Online conference, jointly organized by Universität Hamburg and the research center DESY ****

*Speaker

PoS(EPS-HEP2021)045

1. Introduction

The Dark Matter Particle Explorer (DAMPE) is a spaceborne cosmic ray detector and γ-ray observatory that has been in stable operations since December 2015 [1]. It carries four subdetectors to measure the energy, absolute charge, and direction of an incoming particle. Its deep Bismuth Germanium Oxide (BGO) calorimeter serves as the primary trigger and for particle identification. Among its main scientific objectives is the precise measurement of cosmic electron+positron (CRE) spectrum up to 10 TeV, a region where several models propose observable features [2]. In this paper, we present a deep learning method for CRE identification on DAMPE. The new method was the object of a recent publication [3] and has been updated since then.

2. Electron identification

The DAMPE detector is able to measure the absolute electric charge $|Z|$ of incoming cosmic rays, allowing to identify and reject ions from Helium up, and the very few γ-rays in orbit at TeV energies. In the absence of a magnet however, the sign of the electric charge is unknown. Therefore, identifying electrons requires to reject the other $|Z| = 1$ particle in space: protons.

Electrons and protons react differently with matter. When interacting inside the BGO calorimeter, the former trigger electromagnetic showers, which are typically shorter and narrower than the hadronic showers produced by the latter. Clasically, one can try and quantify this topological difference. The DAMPE experiment uses an observable called ζ [4] (Figure 1) which is the product between shower width and shower depth. This variable was successfully used for the measurement of the CRE spectrum up to 4.6 TeV [5]. However, ζ loses discrimination power towards highest energies and a more powerful identification method is required.

Deep neural networks [6] have the potential of better exploiting the correlations between topological variables and to extract more information from DAMPE data. We therefore propose a neural network classifier for electron identification.

A neural network is a stack of densely connected mathematical units called neurons, that apply a non-linear activation function on the linear combination of their input. The weights of each neuron connection is determined during the training procedure, where the network is exposed to carefully chosen data. Once the network is fitted to the data, it can take decisions such as in our case the identity of a cosmic ray.

We train a neural network on Monte Carlo (MC) data simulated with the Geant4 package to represent typical cosmic ray interactions inside DAMPE. The training set covers the full energy range of an expected CRE measurement, with 230'000 events equally distributed between signal and background. The input variables are selected first to be representative of the interaction topology, and second to yield reliable, well-simulated results: the energy deposited in the bottom 12 of the 14 layers of the calorimeter, the RMS of the deposition in the same layers, the angle of trajectory, the total energy, the energy along the track in the tracker, and the classical ζ observable.

The network itself is a simple 4-layers dense neural network, with 300-150-75-1 neurons respectively and the ReLU activation function in the hidden layers. It is trained for 50 epochs following the Adam optimizer, using the Keras and Theano libraries on Nvidia Titan X video cards. For more details, see reference [3].

PoS(EPS-HEP2021)045

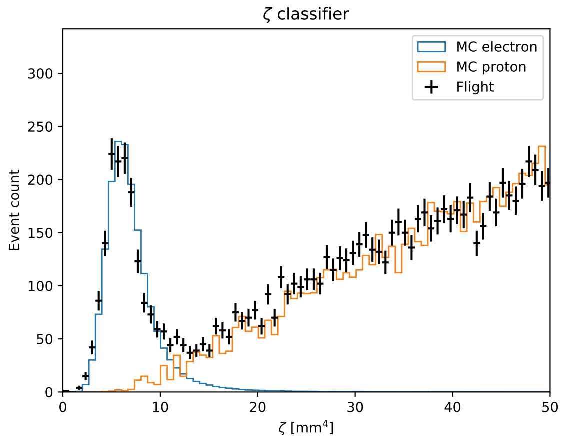

Figure 1: Example distribution of the classical electron identifier ζ, on MC and real data in the 1 to 3 TeV reconstructed energy range.

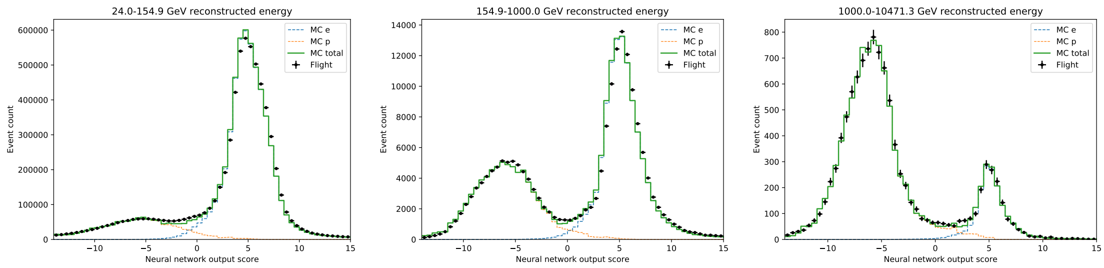

Figure 2: Neural network response distribution comparing MC and real data, on 3 energy bins covering the expected range of DAMPE measurement.

3. Performances

The neural network response distribution is displayed in Figure 2, comparing MC electron, MC proton and real data. MC was scaled to the data in the following way: a control region is set in the range $[-12; -4]$, where the integral of MC proton and data are computed. Their ratio serves as the scaling factor to create a background template. Then, said template is subtracted to the flight data to have a "background-less" template. A signal region $[2; 100]$ is set, and the ratio of integrals yields the electron scaling factor.

One sees that the neural network separates nicely electrons and protons, the former having positive values and the latter negative values. Monte Carlo and real data matching is excellent at all energies, without the need for any smearing although with a small shift of $+0.2$ for MC protons. Therefore the classifier is reliable and generalises well the physics (instead of being biased towards simulation artifacts). Such matching enables the use of baseline methods for background evaluation.

It should be noted that Figure 2 differs from typical machine learning classification distributions. The network was indeed modified by removing its output activation function to have a more convenient shape with background interpolation methods in mind. The performances are not affected as the transformation is bijective. Details in Reference [3].

To assess the performances of the neural network, we define three key metrics:

$$\text{Signal efficiency} = \frac{N_{e^-,pass}}{N_{e^-}} \tag{1}$$

$$\text{False positive rate} = \frac{N_{p,pass}}{N_p} \tag{2}$$

$$\text{Background contamination} = \frac{N_{p,pass}}{N_{p,pass} + N_{e^-,pass}} \tag{3}$$

with N the number of events, and *pass* indicating whether they survive the set discrimination threshold. The first two metrics have the advantage of being blind to the relative fraction of electrons to protons in orbit and can therefore be estimated purely on MC. The last metric requires some fitting to real data (Figure 2) but better assesses the final performances of the classifier in a physical analysis.

Figure 3 shows the evolution of performances with respect to energy, comparing the new method with the classical one ζ. The left panel is only MC data. For a fair assessment, at every energy we set a discrimination threshold such that both classifiers achieve the same signal efficiency (95%), and we record the false positive rate. It can be seen that the neural network achieves a significantly better proton rejection at all energies, especially at the extremes. Above 10 TeV, it reaches 8 times lower false positive rate[1].

The right panel of Figure 3 is a fixed-cut analysis, that uses real DAMPE data to assess the background contamination. The neural network efficiency exhibits very little dependency with energy, by opposition to ζ dropping from 98% down to 50% at 10 TeV. At highest energies, the neural network achieves almost twice higher efficiency for the same background contamination. It should be noted that, while the retained cut allows a higher efficiency for ζ at low energies, it is at the cost of higher background: on this panel, direct comparison is thus only fair when their respective background equalises, above 4 TeV.

In conclusion, the developed neural network classifier exhibits excellent performances, up to 8 times better than the classical method. It is also highly reliable thanks to a near identical response on Monte Carlo and on real data. Such features allow us to better exploit the wealth of information recorded by the DAMPE experiment, and pave the way towards a measurement of the CRE spectrum even beyond 10 TeV.

Acknowledgments

The DAMPE mission was funded by the strategic priority science and technology projects in space science of the Chinese Academy of Sciences. In China, the data analysis was supported in part by the National Key Research and Development Program of China (no. 2016YFA0400200), the National Natural Science Foundation of China (nos. 11525313, 11622327, 11722328, U1738205, U1738207, U1738208), the strategic priority science and technology projects of the Chinese Academy of Sciences (no. XDA15051100), the 100 Talents Program of Chinese Academy of Sciences, and the Young Elite Scientists Sponsorship Program. In Europe, the activities and the

[1]The original model [3] had a factor 4 improvement. The only change that was made since then is additional MC available for training, which illustrates the importance of data size.

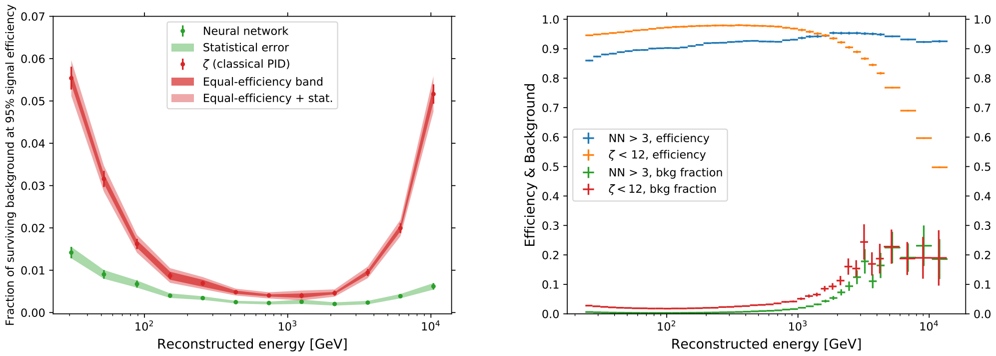

Figure 3: Performances of the neural network (NN) classifier against the classical method ζ. *Left:* Moving-cut analysis. The discrimination thresold is dynamically set to have 95% signal efficiency on each energy bin, and the false positive rate is measured. *Right:* Fixed-cut analysis, showing the evolution of signal efficiency and background contamination with energy.

data analysis were supported by the Swiss National Science Foundation (SNSF), Switzerland, National Institute for Nuclear Physics (INFN), Italy and European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 851103).

The computations presented were performed at University of Geneva on the Baobab cluster. Simulations were performed on INFN CNAF and ReCaS clusters, Italy, and on Swiss National Supercomputing Centre (CSCS) Piz Daint (project s979).

References

[1] Chang, J. et al. (2017) The DArk Matter Particle Explorer mission. *Astropart. Phys.,* **95**, 6–24.

[2] Fan, Y.-Z., Zhang, B., and Chang, J. (2010) Electron/positron excesses in the cosmic ray spectrum and possible interpretations. *International Journal of Modern Physics D,* **19**(13), 2011–2058.

[3] Droz, D., Tykhonov, A., Wu, X., Alemanno, F., Ambrosi, G., Catanzani, E., Santo, M., Kyratzis, D., and Zimmer, S. (jul, 2021) A neural network classifier for electron identification on the DAMPE experiment. *Journal of Instrumentation,* **16**(07), P07036.

[4] Chang, J. et al. (2008) Resolving electrons from protons in ATIC. *Adv. Space Res.,* **42**, 431–436.

[5] Ambrosi, G. et al. (2017) Direct detection of a break in the teraelectronvolt cosmic-ray spectrum of electrons and positrons. *Nature,* **552**, 63–66.

[6] Goodfellow, I., Bengio, Y., and Courville, A. (2016) Deep Learning, MIT Press, `http://www.deeplearningbook.org`.