Editor et al. (Hrsg.): INFORMATIK 2022 Workshops,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2022 1

Sequential networks for cosmic ray simulations

Pranav Sampathkumar [1], Augusto Antonio Alves Junior [1], Tanguy Pierog [1], Ralf Ulrich for the CORSIKA8 Collaboration

Abstract: A hybrid model of generating cosmic ray showers based on neural networks is presented. We show that the neural network learns the solution to the governing cascade equation in one dimension. We then use the neural network to generate the energy spectra at every height slice. Pitfalls of training to generate a single height slice is discussed, and we present a sequential model which can generate the entire shower from an initial table. Errors associated with the model and the potential to generate the full three dimensional distribution of the shower is discussed.

Keywords: Sequential Neural Networks; Astroparticle Physics; Monte Carlo Simulations

1 Introduction

Cosmic rays are important messengers from the high energy universe. They are also one of the most exciting platforms to study High Energy Physics. They provide us with a window into high energy collisions, which is not possible by current accelerators (of the order $10^{17}eV$). While they provide with extremely high energies, analysing physics using cosmic rays is challenging because of their low intensity and can only be studied using Extensive Air Shower (EAS) which they trigger in the atmosphere. The extensive air showers are a series of cascaded interactions when cosmic rays interact with the earth's atmosphere. Each high energy interaction, triggers a chain reaction, and they produce secondary particle showers. Analysis of such showers is a challenging task because of all the interactions possible in the earth's atmosphere such as, a variety of e/m interactions such as bremsstrahlung, pair production, Compton, Moller, Bhabha processes as well as annihilation and ionization losses. We would also have to simulate similar interactions for high energy muons. All these interactions as well as the demand to be as optimal as possible owing to the large number of times such simulations need to be repeated, for example, in MCMC analysis calls for a demand in more novel approaches to minimize memory overhead and optimize simulations.

The two main approaches currently used are, *thinning* and *hybrid simulations*. In thinning, only a small portion of the shower is explicitly considered for Monte-Carlo simulations while the rest are ignored. The considered showers are re-weighted accordingly. Such an approach, gives rise to artificial fluctuations because of the lower number of particles considered. One way to alleviate this problem is to have a maximum allowable weight to a particle, but this creates limitations on using less detailed sampling and doesn't help speedup the calculation

[1] Institute for Astroparticle Physics, Karlsruhe Institute for Technology, Karlsruhe, Germany
pranav.sampathkumar@kit.edu

doi:12.34567/provided-by-the-editors

process[Ko01]. The second hybrid simulations approach, models the shower using a cascade equation and then after initially running the shower using explicit Monte-Carlo simulations for very high energies. The data is then fed to cascade equations at a threshold. Combining cascade equations and MCMC this way, allows us to get accurate results for the average behaviour of the shower.

The idea of using cascade equations for cosmic ray showers is not new. There have been various attempts at trying to model the longitudinal profile of cosmic ray showers using cascade equations for more than 30 years. The initial shower is usually done explicitly since it involves small number of particles, and it also provides a way to incorporate fluctuations into the simulations since cascade equations are deterministic. We can determine the mean behaviour and the moments of the fluctuations using cascade equations accurately. One of the biggest problems with this approach is, it is limited by our modelling of the cascade equations. Currently, we only have the analytical form of the cascade equations in one dimension. Thus, cascade equations only provide us with the longitudinal profile of the shower. The numerical solution of the cascade, equation can also be avoided by a pre-tabulation of of the secondary showers, using an iterative MC procedure [Al02]. Such pre-tabulation makes the simulations less flexible to changes in the environment and initial conditions.

The current work builds upon the work of CONEX [Be07], in trying to build a neural network model which mimics the behaviour of the analytic form of the cascade equation. Once the network is trained, the network would no longer require any sort of pre-tabulation and is designed to be flexible across various initial conditions. This paper is organised as follows, Sect. 2 serves as an introduction to cascade equations for a general particle shower. Sect. 3 talks about the neural network model. Sect. 4 provides some preliminary results to this approach, and Sect. 5 talks about conclusions and future potential directions for this work.

2 Cascade equations for hybrid simulations

A simple Monte-Carlo process can be described by a linear cascade equation. The current approaches to cascade equations for hybrid simulations are done in a theory driven manner, where our theoretical understanding and ability to derive the analytical form of the cascade equation. Cascade equations in one-dimensional form are derived based on various physics processes, such as hadronic and electromagnetic interactions and they are numerically solved using the tables generated during the explicit simulation as initial conditions.

Various attempts have been made towards cascade equations, with physics parameters, modelled specifically for hadronic interactions and electromagnetic interactions. The general structure of the cascade equations are as follows.

$$\frac{\partial n_i(E,X)}{\partial X} = \overbrace{\sum_a \int_E^{E_0} \sigma_a n_a(E',X) P_{a\to i}(E,E')dE' - \int_{E_{min}}^{E} \sigma_i n_i(E',X) P_{i\to a}(E,E')dE'}^{\text{Production and Decay Terms}} \\ \underbrace{+S_i(E,X) - \sigma_i n_i(E,X) - \alpha \frac{\partial n_i(E,X)}{\partial E}}_{\text{Source, Interaction and Loss terms}} \tag{1}$$

where, $n_i(E, X)$ is the energy spectra at given depth X. Solution to the cascade equation provides us with the energy spectra $\forall X$. This cascade equation is one dimensional, and helps us calculate the longitudinal profile of the shower.

Production terms in Equation (1), account for the formation of particles by particles at higher energy decaying, and increasing the number of particles at a particular energy bin. *Decay* terms account for the loss of particles at a particular bin by decaying into particles of lower energy. *Source* terms account for sudden appearance and disappearance of particles, not captured by the rest of the cascade equation. It can account for new particles joining the shower during the process of solving it, or particles leaving the cascade equation to be handled differently. *Interaction* and *Loss* terms account for the loss in energy of particles by not performing a explicit cascade beyond a particular energy threshold. This involves ionization losses in the atmosphere, etc.

CONEX solves this system of coupled differential equations, and gets the projection of the shower along the longitudinal axis. One of the problems with extending this approach to higher dimensions is the lack of the analytical form of the cascade equation. In 1D the system is completely described by the energy spectrum at a particular height, but at 3D, we need additional information such as the direction of the subshowers etc in order to fully describe the system. We postulate the existence of such a differential equation which can fully describe the system in 3D though it is not tractable in analytic form. We use a neural network approach where the learning process learns the differential equation which describes the system and can be used to generate the state of the shower at every height.

3 CONEX inspired sequential neural approach

Let us assume that f is the *functional* which takes the energy spectrum at a particular height and gives the energy spectrum at the next height step.

$$n(E, X + \Delta X) = f(n(E, X)) \tag{2}$$

solving the differential equation is equivalent to finding this functional, which helps us iterate through the height and generate the entire shower from an initial condition. The ideology behind the neural network model is that, such a functional can be approximated by the neural network during the training process. This allows us to model the interaction using explicit simulations directly without relying on our understanding of the underlying physics. This also allows us to model higher dimensional distributions since we don't rely on the explicit form of the cascade equations.

As a initial check of the idea, we take a purely electromagnetic cascade in vacuum, and see if it learns the information to perform a single step. The input dataset is generated using CONEX, and we check if given a table at a particular height slice, the network can learn to generate the table at the next height slice. We use a fully connected neural network for this implemented using `PyTorch` [Pa19]. The network has 5 hidden layers, with 512 , 256 , 128 , 256 , 512 nodes each, using MSE as loss function and optimized using ADAM.

4 Preliminary Results

These are the preliminary results for the ability of the network to learn the functional. Fig. 1. shows that the neural network is capable of performing a single step with good accuracy. We find that the neural network learns with around 5% error. When we try to use the learned f, to generate the entire shower, we find that the error soon accumulates and the shower reaches unphysical situations which the network hasn't seen in its training dataset.

A recurrent neural network is ideally suited for such a problem where there are corrective mechanisms for such "time series"data as a hidden layer is passed to the next step. Before using recurrent neural networks, we attempt using sequential neural networks. The neural network is applied iteratively for multiple steps and then the result is back propagated. Such *sequential chains* are useful in helping the network learn the correct physics, so that it doesnt lead to unphysical situations when applied iteratively. The length of the sequential chains is a hyperparameter which needs to be tweaked. The results in this work are done using a sequential chain of ten steps.

Fig. 2 shows the results of using the sequential network to generate the entire shower. We see while the network has higher error earlier in the shower. It has learnt the right physics and has low energy at the later parts of the shower. This is particularly useful given, the later parts of the shower are the ones which are computationally intensive and the part which we hope to replace with our model.

5 Conclusions and Further Work

We see that the neural network can learn the stepping function f, for the one dimensional case. Further, work is needed on analysing the learnt physics by injecting new particles into the shower in between, a situation the network didn't see in the training dataset. This checks if the network learns that showers can be super imposed. One way to solve this is to hard code the linearity of Eq. (1) into the network to improve the network performance in early parts of the shower. This work is ongoing currently.

After this, the aim is to make the network generate the three dimensional distribution. We also hope that the network can predict the higher moments of the fluctuations apart from the mean behaviour, when the higher moments are provided along with the input distribution. The methods outlined in this work, can also be extended to any simple monte-carlo system since they would be described by a linear cascade equation. We hope this will be a valuable contribution to the cosmic-ray physics community where the showers can be quickly generated for high energies for event-by-event analysis in air-shower arrays or fluorescence light detectors.

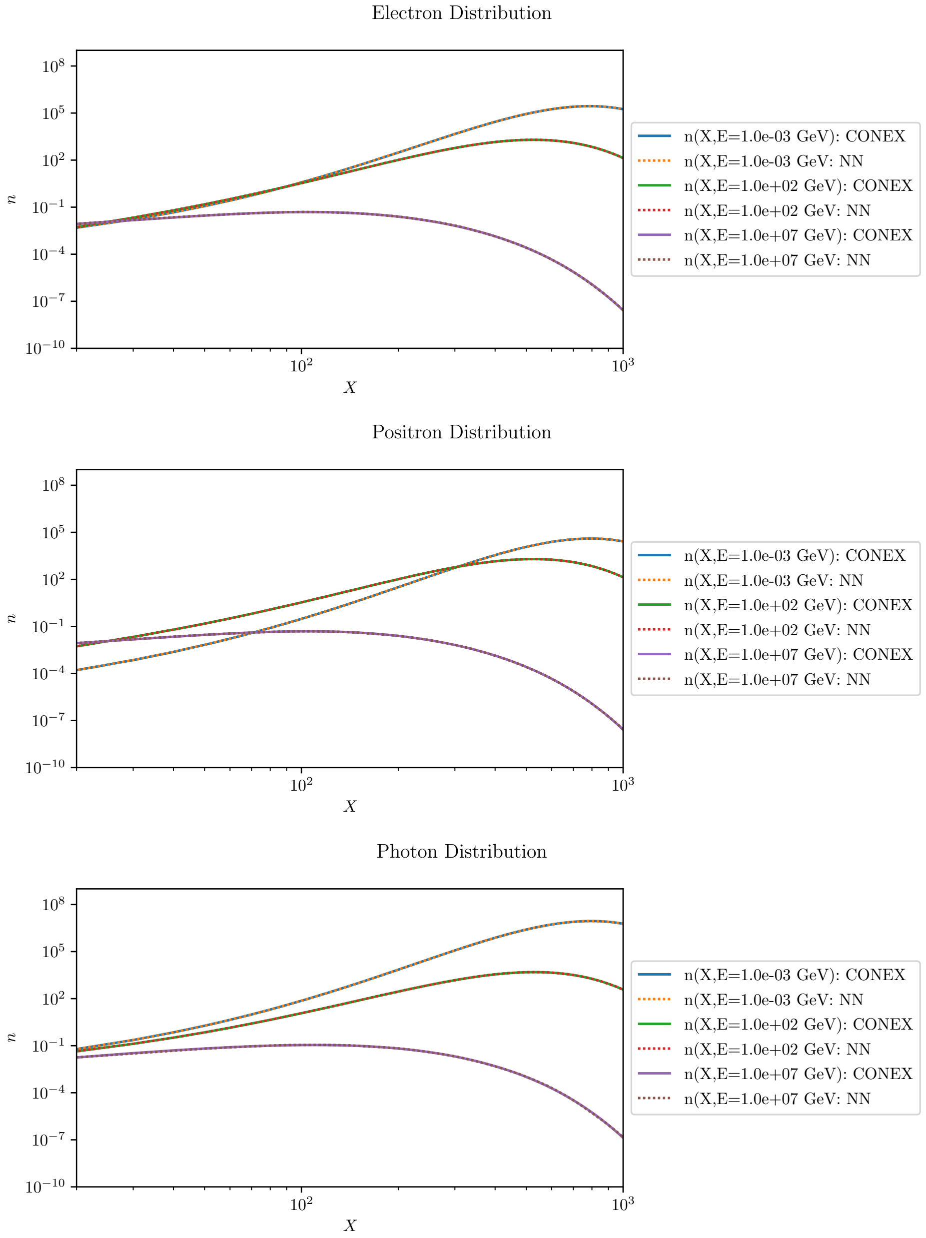

Fig. 1: Results where the neural network performs a single step, the generated showers match with CONEX showers and are accurate to around 5% error.

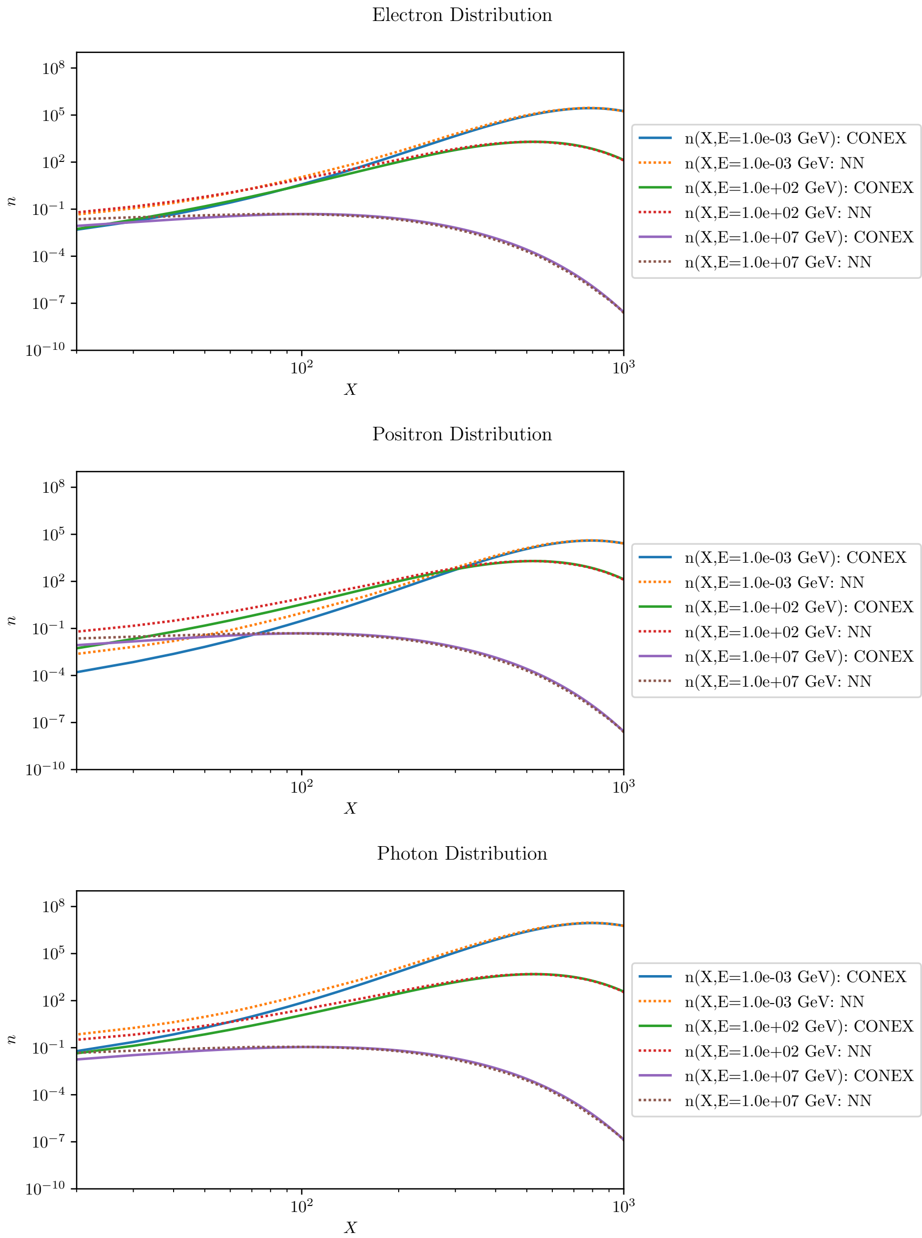

Fig. 2: Results where the neural network generates the entire shower. The generated shower deviates from the CONEX showers earlier in the shower

References

[Al02] Alvarez-Muñiz, J.; Engel, R.; Gaisser, T. K.; Ortiz, J. A.; Stanev, T.: Hybrid simulations of extensive air showers. Phys. Rev. D 66/, p. 033011, Aug. 2002, URL: https://link.aps.org/doi/10.1103/PhysRevD.66.033011.

[Be07] Bergmann, T.; Engel, R.; Heck, D.; Kalmykov, N.; Ostapchenko, S.; Pierog, T.; Thouw, T.; Werner, K.: One-dimensional hybrid approach to extensive air shower simulation. Astroparticle Physics 26/6, pp. 420–432, 2007, ISSN: 0927-6505, URL: https://www.sciencedirect.com/science/article/pii/S0927650506001228.

[Ko01] Kobal, M.: A thinning method using weight limitation for air-shower simulations. Astroparticle Physics 15/3, pp. 259–273, 2001, ISSN: 0927-6505, URL: https://www.sciencedirect.com/science/article/pii/S0927650500001584.

[Pa19] Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; Desmaison, A.; Kopf, A.; Yang, E.; DeVito, Z.; Raison, M.; Tejani, A.; Chilamkurthy, S.; Steiner, B.; Fang, L.; Bai, J.; Chintala, S.: PyTorch: An Imperative Style, High-Performance Deep Learning Library. In (Wallach, H.; Larochelle, H.; Beygelzimer, A.; d'Alché-Buc, F.; Fox, E.; Garnett, R., eds.): Advances in Neural Information Processing Systems 32. Curran Associates, Inc., pp. 8024–8035, 2019, URL: http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf.