

# Reconstruction of the muon production longitudinal profiles in extensive air showers

Antonín Kravka<sup>1,\*</sup>, Eva Santos<sup>2,\*\*</sup>, Maximilian Stadelmaier<sup>2</sup>, and Alexey Yushkov<sup>2</sup>

<sup>1</sup>Czech Technical University - Faculty of Nuclear Sciences and Physical Engineering

<sup>2</sup>FZU - Institute of Physics of the Czech Academy of Sciences

## Abstract.

Muons produced in extensive air showers have large decay lengths and small radiative energy losses, enabling a large fraction of them to reach surface and underground detector arrays while keeping relevant information about the hadronic interactions that occurred high in the atmosphere. We can relate a muon's arrival time and position at the detector to its production depth in the atmosphere. The total delay of muons with respect to the shower plane is primarily due to their geometric path and energy, we call these contributions the geometric and kinematic delays, respectively. We are working on the improvement of the current kinematic delay parameterizations using Deep Neural Networks for muons arriving at surface and underground detector arrays. We aim to reconstruct the longitudinal profile of muons for future arrays of buried scintillator detectors at energies from around the second knee to the ankle of the cosmic ray spectrum, where there is an overlap with the nominal energies at the LHC. Given the low acceptance of scintillator detectors to inclined air showers and the richness of the forward physics near the shower core, we aim at applying a radial cut of 200 m instead of the usual 1000 m used in previous works.

## 1 Introduction

Cosmic rays with energies above  $10^{15}$  eV can only be detected indirectly through the secondary particles produced from their interaction in the Earth's atmosphere. Their energy and, particularly, nuclear mass composition are inferred by comparing the air-shower particles' time, lateral, and longitudinal distributions with those of Monte Carlo simulations using hadronic interaction models. However, the most relevant hadronic interactions of the most energetic cosmic rays occur at  $\sqrt{s} \sim 100$  TeV, one order of magnitude above the nominal energy at the LHC, at phase-space regions not well-covered by most detectors. Lacking a theory to calculate multi-particle production from first principles, hadronic interaction models are our most significant source of systematic uncertainty [1].

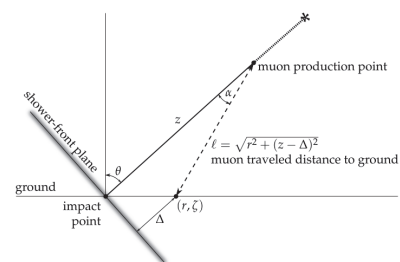
Muons, mostly produced from the decay of charged pions, are the best messengers of the hadronic cascade. After production, they propagate in nearly straight lines to the ground without suffering significant energy losses, keeping relevant information about their mother particle [2–4]. In [5], from a compilation of the eight largest cosmic-ray experiments, a deficit of muons in simulations with respect to data is reported for  $E > 10^{17}$  eV with a significance of more than  $8\sigma$ , right where the center-of-mass energy equals the nominal energy at the LHC. Also, in [6], the Pierre Auger Collaboration observed a tension between

the EPOS-LHC model predictions and data for the maximum depth of the muon production longitudinal profile.

In this work, we extend the method of the reconstruction of the muon production depth (MPD) described in [2, 3], used in [6], for an array of underground muon detectors, as is the case of AMIGA [7], currently being deployed at the Pierre Auger Observatory.

## 2 Muon Production Depth

In [2, 3], a method relating the arrival time of muons with their production distance is proposed. It is assumed that muons are produced in the shower axis and propagate to the ground in straight lines, as shown in Figure 1. Ac-



**Figure 1.** Schematic sketch of the geometry used to calculate the point where a muon was produced in the atmosphere, [6].

cording to Figure 1, muons deviate from the shower axis

\*e-mail: [kravkant@fjfi.cvut.cz](mailto:kravkant@fjfi.cvut.cz)

\*\*e-mail: [esantos@fzu.cz](mailto:esantos@fzu.cz)

by an angle  $\alpha$ , arriving at the ground with coordinates  $(r, \zeta)$  expressed on the shower-front plane coordinate system. Here  $r$  is the muon distance from the shower axis, and  $\zeta$  is its azimuthal angle. The total distance traveled by the muon is defined as  $l = \sqrt{r^2 + (z - \Delta)^2}$ , where  $z$  is the muon production distance, measured along the shower axis,  $\Delta = r \cos \zeta \tan \theta$  is the  $z$ -coordinate of the muon at the detector, expressed on the shower-front plane coordinate system, and  $\theta$  is the shower zenith angle. Given the longer path traveled by muons, they will arrive delayed with respect to the arrival of the shower front. Assuming that muons are massless particles that travel at  $c$ , using a purely geometric relation, we can calculate the muon production distance,  $z$ , as

$$z = \frac{1}{2} \left( \frac{r^2}{ct_g} - ct_g \right) + \Delta, \quad (1)$$

where  $t_g$  is the geometric delay of the muon. However, there are additional sources of delay, being the most relevant one the so-called kinematic delay, which is intrinsically associated with the energy of the muon [2–4]. Therefore, Equation (1) is rewritten as  $t_g \simeq t - \langle t_\epsilon \rangle$ , where  $t$  is the total delay, i.e., the one we can measure, and  $\langle t_\epsilon \rangle$  is the kinematic delay. The latter is usually calculated from a parametrization assuming the average energy of muons arriving at a given distance from the shower axis, see Appendix B in [3]. Finally, the Muon Production Depth (MPD) is calculated from the muon production distance  $z$  as

$$X = \int_z^\infty \rho(z') dz', \quad (2)$$

where  $\rho$  is the atmospheric density.

The method proposed in [2, 3] is intended for a detector array, as the one of the Pierre Auger Observatory [8], for  $\theta \sim 60^\circ$  showers, applying a radial cut of  $r > 1000$  m. The reason for these cuts is two-fold; namely, they ensure negligible electromagnetic contamination in the total signal, while the contribution of the geometric delay, from which the MPD is calculated, is the dominant source of delay. In [6], this method was applied to the Pierre Auger data for events with  $E > 10^{19.3}$  eV and  $55^\circ \leq \theta \leq 65^\circ$ , using detectors at  $1700 < r[\text{m}] < 4000$ . Given the 1500 m spacing of the Auger detectors, the array proved too sparse, limiting the method's applicability to fewer than 500 events.

Our goal is to extend the present method's applicability to an array of buried scintillator detectors, as in the one of AMIGA [7]. Another advantage of AMIGA is its 750 m and 433 m detector spacing, allowing us to encompass the  $10^{17}$  eV energy region, where overlap with the LHC energies exists. That enables us to reconstruct the MPD in an energy region where we expect reduced systematic uncertainties from hadronic interaction models. However, since typical scintillator detectors have  $\sim 5$  cm width, their acceptance is limited to  $\theta < 55^\circ$ , [9]. Also, the shower footprint is smaller at low zenith angles and energies, requiring a significantly lower radial cut of  $r \gtrsim 200$  m.

## 2.1 Towards the reconstruction of the MPD

In this subsection, we report on our assessments of the feasibility of the extension of the present method to  $\theta < 60^\circ$ , applying a radial cut of 200 m, as required for an array of buried scintillator detectors as it the case of AMIGA [7]. The AMIGA detectors are being deployed at 2.3 m depth, resulting on an additional mass overburden of  $540 \text{ g cm}^{-2}$  for a local soil density is  $\rho = 2.4 \text{ g cm}^{-3}$  [7]. For our calculations we assume that all muons behave as minimum ionizing particles, i.e., losing a constant fraction of their energy per unit of crossed superficial density of  $a(E) = 1.8 \text{ MeV/g cm}^{-2}$  [10], from which we calculate the minimum kinetic energy for the muons to reach the AMIGA detectors as

$$E_{kin}^\mu > a(E) \rho h / \cos \theta, \quad (3)$$

where  $h = 2.3$  m is the depth at which the detectors are buried, and  $\theta$  is the zenith angle of the shower. From Equation (3), we expect that the minimum kinetic energy for a muon to arrive to the underground detectors to increase as a function of the zenith angle, resulting it to be  $E_{kin,min}^\mu = 1 \text{ GeV}$  ( $2 \text{ GeV}$ ) at  $\theta = 0^\circ$  ( $\theta = 60^\circ$ ). Throughout our work, we will use the results of Equation (3) to select the muons from which we reconstruct the MPD at 2.3 m depth.

### CORSIKA shower libraries

For our studies, we produced two CORSIKA [11] shower libraries, one with a standard thinning of  $10^{-6}$ , combined with a radial thinning of  $r < 50$  m, and a second one in which we did not activate the thinning algorithm. We set the ground level at 1452 m, the average height of the Pierre Auger Observatory. The second library was aimed to study the impact of thinning on the muon distributions close to the shower axis. In both cases, we use a fixed energy of  $10^{17}$  eV, and several zenith angles, assuming proton and iron-initiated showers. We have simulated 200 showers per nuclear species, zenith angle, and high-energy hadronic interaction model.

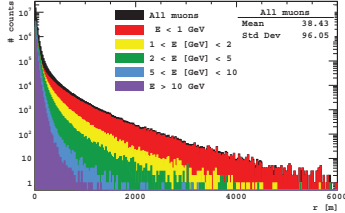
The "thinned" shower library was produced with CORSIKA 7.7402 for the following zenith angles:  $\theta = \{0^\circ, 25^\circ, 35^\circ, 45^\circ, 55^\circ, 60^\circ\}$ . We consider the high-energy hadronic interaction models EPOS-LHC [12], QGSJet-II.04 [13] and Sibyll 2.3d [14], and FLUKA 2020.0.6 [15] as the low-energy hadronic model. For the "unthinned" shower library, we used CORSIKA 7.7410 with FLUKA 2020.0.10. Only Sibyll 2.3d as the high-energy hadronic model and  $\theta = \{0^\circ, 35^\circ, 60^\circ\}$  were considered.

### Testing the original model assumptions

In [2, 3], it was assumed that muons are produced in the shower axis. In an update in [4], the average distance of the production of muons from the shower axis was estimated to be about a few tens of meters for  $10^{19}$  eV showers with  $\theta = 60^\circ$ , using pre-LHC model predictions. We check if these findings hold for the current models and at lower-energy and zenith angles. In this section, for illustration, we show our results for 20 unthinned Sibyll 2.3d

iron-initiated showers at  $10^{17}$  eV and  $\theta = 0^\circ$ . In the text, we compare the results from the figures with the outcome of proton-initiated showers and also at  $\theta = 60^\circ$ , when pertinent.

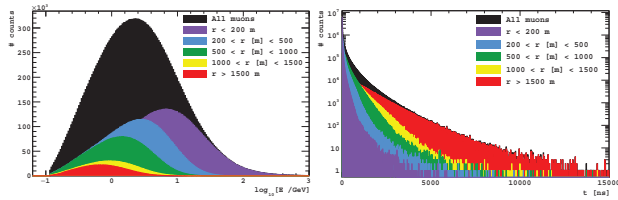
In Figure 2, we plot the distance from the shower axis at which the muons are produced. We observe that, for



**Figure 2.** Distance from the shower axis at which muons are produced. The colors represent several ranges for the initial energy muons, namely: black – all energies; red –  $E_i < 1$  GeV; yellow –  $1 < E_i$  [GeV]  $< 2$ ; green –  $2 < E_i$  [GeV]  $< 5$ ; blue –  $5 < E_i$  [GeV]  $< 10$ ; violet –  $E_i > 10$  GeV.

both primary species, at  $\theta = 0^\circ$ , muons are produced at  $\langle r_{prod} \rangle \sim 40$  m, a distance that reduces to  $\sim 30$  m at  $\theta = 60^\circ$ , confirming the findings of [4]. We also see that at the largest distances we are dominated by a population of low-energy muons. Quantitatively, at  $0^\circ$ , muons with  $1 < E_i$  [GeV]  $< 2$  ( $2 < E_i$  [GeV]  $< 5$ ) are produced at an average distance of 45 m (30 m) from the shower axis.

In a second step, we studied the muon spectra and arrival time distributions of muons at the ground level as a function of the distance to the shower axis, as shown in the left and right panels of Figure 3, respectively. The arrival time of muons was calculated as  $t' = t - t_0$ , where  $t$  is the elapsed time since the first interaction, and  $t_0 = (h_1 - h_g) / (c \cos \theta)$  is the time needed for a hypothetical massless particle to travel, parallel to the shower axis, from the first interaction point to the ground. In the equation,  $h_1$  is the height of the first interaction, as given by CORSIKA,  $h_g$  is the height of the ground level, and  $c$  is the speed of light. From the top panel of Figure 3, we observe that, near

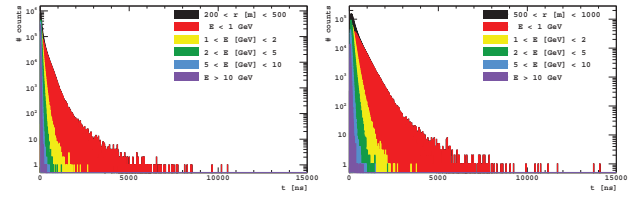


**Figure 3.** Energy spectra (left) and arrival time distribution (right) of muons arriving at the ground plane. We use the same notation for the two plots: black – all muons; violet –  $r < 200$  m; blue –  $200 < r$  [m]  $< 500$ ; green –  $500 < r$  [m]  $< 1000$ , yellow –  $1000 < r$  [m]  $< 1500$ , and red –  $r > 1500$  m.

the core, we have a very rich muon spectrum, ranging from hundreds of GeV to less than 1 GeV. However, as we move further away from the shower axis, we become dominated by a less varied and less energetic muon component. We also find that the muon spectrum hardens with the zenith angle, most likely due to the decay of low-energy muons

in the increased air overburden, which grows by a factor of two at  $\theta = 60^\circ$ . We further observe that for  $\theta = 0^\circ$ , there is a much richer energy spectrum at  $r < 200$  m than the one for  $60^\circ$ . On the bottom panel of Figure 3, we see that the arrival time distribution of muons is narrower closer to the shower axis, broadening as we move further away from it. There is a change in behavior in the arrival time distributions as a function of the zenith angle for the near and far regions from the shower axis. Namely, at  $0^\circ$ , we have a broader arrival time distribution of muons arriving at  $r < 200$  m, than at  $60^\circ$ .

Finally, in Figure 4, we study the impact of the energy on the muon arrival time distributions at  $200 < r$  [m]  $< 500$  (left) and  $500 < r$  [m]  $< 1000$  (right). From Fig-



**Figure 4.** Arrival time distribution of muons arriving at  $200 < r$  [m]  $< 500$  (left) and  $500 < r$  [m]  $< 1000$  (right). The colors represent several ranges for the initial energy muons, namely: black – all energies; red –  $E_i < 1$  GeV; yellow –  $1 < E_i$  [GeV]  $< 2$ ; green –  $2 < E_i$  [GeV]  $< 5$ ; blue –  $5 < E_i$  [GeV]  $< 10$ ; violet –  $E_i > 10$  GeV.

ure 4, we observe that the most (least) energetic muons dominate the early (late) part of the distribution. However, at the beginning of the arrival time distribution, the muon energy spectrum is very diverse, ranging from the lowest to the highest energies, making the proper estimation of the kinematic delay challenging.

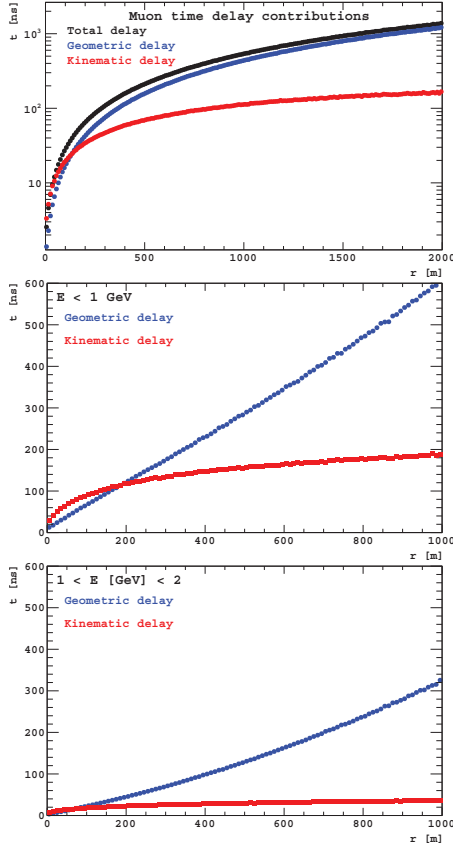
#### Estimating the kinematic delay

In [2, 3], the estimated kinematic delay of muons traveling a distance  $l$  is estimated calculated as:

$$t_\epsilon \approx \frac{1}{2} \frac{(m_0 c^2)^2}{c p a} \left[ \frac{1}{E_f} - \frac{1}{E_i} \right], \quad (4)$$

where  $m_0$  is the muon rest mass,  $a = 2 \text{ MeV/g cm}^{-2}$ , assuming that muons behave as minimum ionizing particles [10].  $E_i$ , and  $E_f$  are the muon energy at production, and at the detector, respectively.

We will use Equation (4) to explicitly calculate the kinematic delay of muons  $t_\epsilon$ . In Figure 5, we plot the total (black), geometric (blue), and kinematic (red) delays of muons as a function from the shower axis. The estimate the total delay as the  $t' = t - t_0$ , as explained above, and the geometric delay as  $t_g = t' - t_\epsilon$ . From the middle and bottom panels of Figure 5, we confirm that the kinematic delay is intrinsically related with the energy of muons, and it is larger for low-energy muons. We also observe that the kinematic delay is dominant close to the shower axis, prevailing over the geometric delay until  $r = 200$  m for  $E < 1$  GeV.



**Figure 5.** Contributions to the total muon delay for an average of 20 iron-initiated showers with  $E = 10^{17}$  eV, and  $\theta = 0^\circ$ . Top panel: Total distributions for the total (black), geometric (blue) and kinematic (red) delay. Middle panel: geometric (blue) and kinematic (red) delays for muons with  $E < 1$  GeV. Bottom panel: geometric (blue) and kinematic (red) delays for muons with  $1 < E$  [GeV]  $< 2$ . Please, mind the logarithmic scale in the  $y$ -axis used in the top panel.

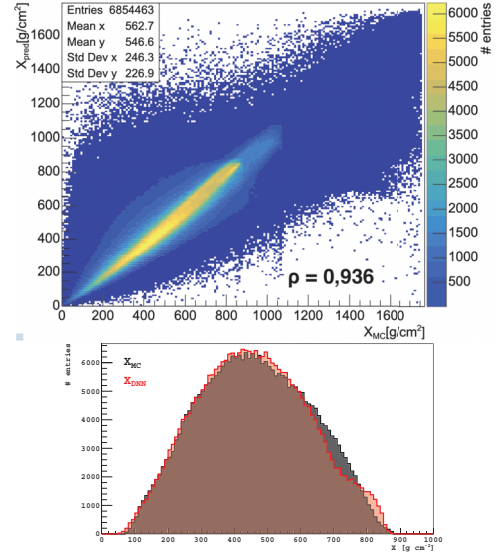
### 3 Calculating the MPD using DNNs

For the MPD reconstruction, we used a Fully Connected Neural Network with 6 hidden layers in a pyramid scheme, where the number of nodes decreases from 1024 to 32 from the first to the last hidden layer. We use ReLU as the activation function and the ADAM optimizer. Our input features to the network are the  $\sec \theta$  of the shower, the ground coordinates of the muon ( $r, \zeta$ ), and its arrival time  $t'$ , as explained above. We also add a boolean flag that tells us if the muons had enough energy to reach the underground detectors, according to Equation (3). Our loss function was derived empirically as

$$L = \left( \frac{z_{DNN} - z_{MC}}{z_{MC} + 1} \right)^2 + [(1 - \rho) + \gamma_3] c, \quad (5)$$

where  $z_{DNN}$  and  $z_{MC}$  are the predicted and Monte Carlo values of muon production distance, respectively.  $\rho$  and  $\gamma_3$  are the Pearson correlation coefficient and the distribution skewness, respectively, and  $c$  is a normalization constant. Finally, we convert the estimated muon production distance  $z$  into the MPD by applying Equation (2). We used 10% of the showers for the training of the DNN.

Our main results are shown in Figure 6. In the top panel, we show the correlation matrix between the reconstructed and Monte Carlo values of the MPD for all muons reaching the ground independently if they had enough energy to reach 2.3 m depth, from which observe a Pearson correlation of 0.936 between the reconstructed and true MPD. On the bottom panel, we show one MPD profile reconstructed at 2.3 m depth for one  $10^{17}$  eV proton-initiated shower at  $\theta = 0^\circ$ , assuming a radial cut of 200 m. We see



**Figure 6.** Top panel: Correlation matrix between the predicted and Monte Carlo MPD. Bottom panel: Example of one reconstructed *apparent* MPD profile using DNNs (red histogram). For comparison, the Monte Carlo profile is also plotted in the figure as the black histogram.

a good agreement between the reconstructed and Monte Carlo profiles up to a production depth of  $500 \text{ g cm}^{-2}$ . However, the DNN seems to perform worse for deeper MPD, probably due to the estimation of the kinematic delay since we expect that, at the deeper regions, the profile is dominated by low-energy muons.

### 4 Conclusions

We extended the MPD reconstruction method presented in [2, 3] to an array of buried scintillator detectors. The applicability region of our work is intended for  $\theta < 60^\circ$ , and  $r > 200$  m, a fully complementary to the previous method. We were able to successfully reconstruct the MPD both for surface and underground muons using a Fully Connected DNN. As we expected, the estimation of the kinematic delay is the largest source of systematic uncertainty in the model. Our preliminary results show that, for the current DNN architecture, we achieve a muon-by-muon reconstruction of the MPD profiles at 2.3 m depth of  $\langle \Delta X \rangle < 10 \text{ g cm}^{-2}$ , and  $\sigma(\Delta X) \sim 80 \text{ g cm}^{-2}$  for  $E = 10^{17}$  eV proton and iron-initiated showers, for  $0^\circ < \theta < 60^\circ$ , when a radial cut of  $r > 200$  m is applied. These results can be improved if a different type of Neural Network or loss function is used. We plan a journal publication with updated results.

## Acknowledgments

This work is fully funded by the Czech Republic grant of MEYS CR GA ĀR 21-02226M.

## References

- [1] R. Engel, et al., *Ann. Rev. Nucl. Part. Sci.* **61**, 467-489 (2011)
- [2] L. Cazon, et al., *Astropart. Phys.* **21**, 71-86 (2004)
- [3] L. Cazon, et al., *Astropart. Phys.* **23**, 393-409 (2005)
- [4] L. Cazon, et al., *Astropart. Phys.* **36**, 211-223 (2012)
- [5] J. Albrecht, et al., *Astrophys. Space Sci.* **367**, 27 (2022)
- [6] The Pierre Auger Coll., *Phys. Rev. D* **90**, 012012 (2014) [erratum: *Phys. Rev. D* **92**, 019903 (2015)]
- [7] The Pierre Auger Coll., *JINST* **11**, P02012, (2016)
- [8] The Pierre Auger Coll., *Nucl. Instrum. Meth. A* **798**, 172-213 (2015)
- [9] The Pierre Auger Coll., *Eur. Phys. J. C* **80**, 751 (2020)
- [10] D. E. Groom, et al., *Atom. Data Nucl. Data Tabl.* **78**, 183-356 (2001)
- [11] D. Heck, et al., FZKA-6019.
- [12] T. Pierog, et al. *Phys. Rev. C* **92**, 034906 (2015)
- [13] S. Ostapchenko, *Phys. Rev. D* **83**, 014018 (2011)
- [14] F. Riehn, et al., *PoS ICRC2015*, **558** (2016)
- [15] G. Battistoni, et al., *Annals Nucl. Energy* **82**, 10-18 (2015)