PROCEEDINGS OF SCIENCE

Attenuation Models for Extensive Air Showers Derived from Simulations

Fiona Ellwanger and Darko Veberič*

Institute for Astroparticle Physics, Karlsruhe Institute of Technology KIT, Karlsruhe, Germany

E-mail: fiona.ellwanger@kit.edu

At ultra-high energies, the flux of cosmic rays is too low for direct measurements to be meaningful. When a cosmic ray enters the atmosphere, it initiates an extensive air shower, producing a cascade of secondary particles that propagate toward the ground. Large arrays of surface detectors are used to measure these secondary particles upon arrival.

The signal detected at a specific reference distance from the shower core serves as a proxy for the shower size and, consequently, as a reliable estimator of the energy of primary cosmic ray. However, shower development is influenced by attenuation effects: measured signals at the ground depend on the amount of traversed atmospheric density (column density) through which the shower evolves. Since the column density varies with the inclination of the shower, it is important to account for these attenuation effects to ensure accurate energy estimation.

In this study, we derive physics-and-geometry-based functional forms to describe attenuation and propose appropriate expansion terms using simple one-dimensional shower-development models, incorporating one or two main particle-cascade components. We then evaluate the applicability and effectiveness of these functional forms using a Monte-Carlo dataset that includes various primary cosmic-ray particles. By directly calibrating the the shower size derived from ground signals to the Monte-Carlo energy, we characterize attenuation behavior across different primary particles, assess the energy dependence of attenuation, and quantify systematic uncertainties introduced by different functional forms.

39th International Cosmic Ray Conference (ICRC2025)
15 – 24 July 2025
Geneva, Switzerland

*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(ICRC2025)425

1. Motivation

Even though we use in our Monte-Carlo study an example of the surface detector of the Pierre Auger Observatory with a spacing of 1500 m (SD-1500), in general methods presented below apply equally well to all other measurements of extensive air showers using surface detector arrays.

A reconstruction of an event typically combines individual signals in the detector stations with an intention to derive a suitable proxy for the *shower size at the ground*, e.g. by fitting a lateral distribution function to the observed signals. The shower size at the ground, S_{gr}, can then be approximated, for example by the expected signal $S_{\mathrm{gr}}(r_{\mathrm{opt}})$ at an optimal distance r_{opt} [1] from the shower axis. For the SD-1500 of the Pierre Auger Observatory the optimal distance is $r_{\mathrm{opt}} = 1000\,\mathrm{m}$.

Depending on the shower geometry, the "true" or "intrinsic" shower size or simply just *shower size* S_{sh} gets attenuated so that the measured signal at the ground S_{gr} depends on the zenith angle θ of the incoming direction. For example, to describe the attenuation, the Pierre Auger Collaboration is using a cubic polynomial function, see Eq. (14). Using a very generic and flexible ansatz, like the polynomial, might be enough for quick elimination of attenuation effects. However, using a polynomial expansion around some chosen reference point implicitly assumes (a) that the attenuation function is only slowly changing in its vicinity and (b) that its shape can be well described by the lowest orders of power for arguments far from the reference. This might not be the case, especially considering how fast the atmospheric column density can increase with the inclination of the shower. For these reasons, in the following sections we will try to derive more practical functional forms of attenuation based on some simple physical assumptions about the shower development, which, hopefully, will have better extrapolation properties when used outside of the specific fit range.

2. Simulations

We will compare the observed attenuation of simulated proton, iron, and photon-induced extensive air showers. The showers were simulated using Corsika [2] version 7.7420 and the EPOS-LHC hadronic interaction model [3] and were simulated in an energy range from 1 to 160 EeV with a spectral index of -1 and with zenith angles θ between 0 and 60° following a uniform distribution in $\sin^2\theta$. In total, approximately 20 000 showers have been used for each primary particle. For the detector simulation we use the setup of water-Cherenkov detectors for the Pierre Auger Observatory using the $\overline{\mathrm{Off}}\underline{\mathrm{line}}$ software package [4]. In analogy to the energy reconstruction used in the Pierre Auger Collaboration, we exploit the signal measured at 1000 m from the shower axis as a suitable proxy for the shower size at the ground. In the simulation, 12 virtual water-Cherenkov stations are placed in a ring around the shower axis at a distance of exactly 1000 m. The shower size S_{gr} at the ground is then estimated by taking signals S^i_{1000} in these stations i and fitting them to the lowest possible order in azimuthal asymmetry as

$$S^i_{1000} = S_{\mathrm{gr}}\,(1 + \alpha\,\cos\phi^{\mathrm{sp}}_i), \tag{1}$$

where ϕ^{sp}_i is the azimuth angle of the station i in the plane perpendicular to the shower (*shower plane*) and where α is the amplitude of the asymmetry. In this way we take into account the asymmetry of the shower at the ground and still obtain a good estimate of S_{gr} even in the case of several non-triggered stations. Nevertheless, when all 12 stations are present, S_{gr} resulting from

PoS(ICRC2025)425

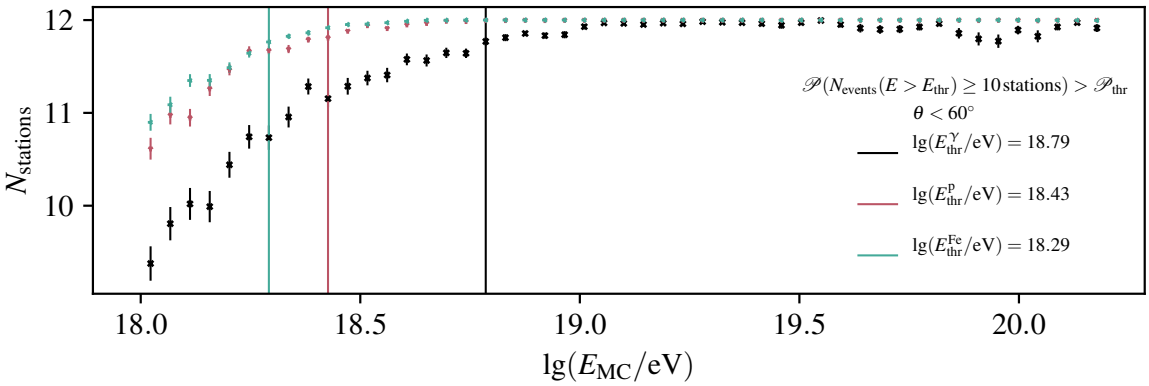

Figure 1: Number of triggered stations in the ring at 1000 m around the shower axis for simulated iron (green), proton (red), and photon-induced showers (black). The vertical lines indicate the corresponding threshold energies defined as the energy at which a fraction of $\mathcal{P}_{\text{thr}}$ of events have at least 10 triggered stations in the ring of 12 simulated stations. For proton and iron-induced showers we chose $\mathcal{P}_{\text{thr}} = 99.9\%$. Since the efficiency for photon-induced showers is not monotonically increasing with the energy, we lower the threshold in the case of photons to reach only $\mathcal{P}_{\text{thr}} = 99\%$.

Eq. (1) is equivalent to a mean over all S^{i}_{1000} values. For the fit we require presence of at least 10 triggered stations and restrict the fit only to the energies above which $\mathcal{P}_{\text{thr}}$ fraction of events satisfy this requirement, see Fig. 1. For proton and iron-induced showers we chose $\mathcal{P}_{\text{thr}} = 99.9\%$. Since for photon-induced showers the efficiency is not monotonically increasing with the energy, for photons we lower this requirement to $\mathcal{P}_{\text{thr}} = 99\%$.

3. Physically motivated forms of the attenuation function

Atmosphere. For attenuation of showers – as observed at the ground – the decisive quantity is the amount of traversed matter i.e. slant column density. In the static and isothermal approximation to the conditions of the atmosphere, the density of air ρ depends only on height, $\rho(h) = \rho_{\text{gr}} \exp(-h/h_{\text{s}})$, where ρ_{gr} is the air density at the ground level ($h = 0$) and h_{s} is the (assumed-constant) scale height of the atmosphere. Assuming a flat Earth and parallel air density layers, we get $X_{\text{gr}}(\theta) = X^{\text{v}}_{\text{gr}} \sec\theta$, where X^{v}_{gr} is the vertical column density. The exact column density in a curved atmosphere is described by the Chapman function [5] but the deviation from the flat approximation only becomes important for very inclined showers, which is beyond the required accuracy for our models in this work.

3.1 Model with two components.

Since the response of either water-Cherenkov or scintillator detectors to secondary shower particles is dominated by the electromagnetic and muonic component, we will from now on assume that showers consist only of these two components. The observed attenuation behavior of the whole shower is then a superposition of the individual attenuation of the two components.

Electromagnetic component. Let us assume that the measured electromagnetic (em) signal at a certain point is proportional to the number of particles $N_{\text{em}} \propto \rho_{\text{em}}$ detected. This number in turn

PoS(ICRC2025)425

depends on a specific stage of the longitudinal development of the shower, which again depends on the traversed column density X until that point. The dependence of the longitudinal particle density $\rho_{\rm em}(X)$ in the electromagnetic cascade on the amount of traversed matter X can be modeled with a Gaisser-Hillas function

$$\rho_{\rm em}(\theta) \propto f_{\rm GH} = \left(\frac{X - X_0}{X_{\rm max} - X_0}\right)^{\frac{X_{\rm max}-X_0}{\lambda}} \exp\left(-\frac{X - X_{\rm max}}{\lambda}\right), \tag{2}$$

where $X_{\rm max}$, X_0, and λ are the essential Gaisser-Hillas shower parameters [6]. We will use this form as a simplified approximation proportional to the signal at the depth X. In doing so, we neglect, in this first stage of derivation, the muonic component of the showers. We can thus write for the signal S at a certain depth X,

$$S(X) = S_{\rm max}\, f_{\rm GH}(X; X_{\rm max}, X_0, \lambda). \tag{3}$$

In this simple model, the particles in the shower continue producing secondary particles until they reach the critical energy $E_{\rm crit}$. This defines the point of maximum shower development, i.e. when $X = X_{\rm max}$, and the number of particles in the shower reaches its peak at $N_{\rm max}$. A detector, approximated as a particle counter, would measure the highest signal, $S_{\rm max}$, when positioned at $X = X_{\rm max}$. Since the energy of the primary particle is distributed among $N_{\rm max}$ particles of energy $E_{\rm crit}$ we can claim that the primary energy E is

$$E \propto N_{\rm max} \propto S_{\rm max}, \qquad \text{and} \qquad E \propto \frac{S(X)}{f_{\rm GH}(X; X_{\rm max}, X_0, \lambda)}. \tag{4}$$

We can now introduce a suitable reference depth $X_{\rm ref}$ where $f_{\rm GH}^{\rm ref} = f_{\rm GH}(X_{\rm ref}; X_{\rm max}, X_0, \lambda)$. The primary energy is thus

$$E \propto \frac{1}{f_{\rm GH}(X_{\rm ref})} \frac{S_{\rm gr}}{f_{\rm GH}(X)/f_{\rm GH}(X_{\rm ref})}, \tag{5}$$

where the first term is just a constant and the zenith-angle dependence of the signal $S_{\rm gr}$ at the ground or the "attenuation" $f_{\rm att}^{\rm em}$ is thus described by

$$f_{\rm att}^{\rm em}(X) = \frac{f_{\rm GH}(X)}{f_{\rm GH}(X_{\rm ref})} = \left(\frac{X - X_0}{X_{\rm ref} - X_0}\right)^{\frac{X_{\rm max}-X_0}{\lambda}} \exp\left(-\frac{X - X_{\rm ref}}{\lambda}\right). \tag{6}$$

Inspired by this result let us first assume that X_0 is close to zero and can thus be safely neglected,

$$f_{\rm att}^{\rm em}(X) \approx \left(\frac{X}{X_{\rm ref}}\right)^{X_{\rm max}/\lambda} \exp\left(-\frac{X - X_{\rm ref}}{\lambda}\right) = \left(\frac{\sec\theta}{\sec\theta_{\rm ref}}\right)^{\mu} \exp(-\gamma(\sec\theta - \sec\theta_{\rm ref})) \tag{7}$$

with $X = X_{\rm gr}\sec\theta$, $X_{\rm ref} = X_{\rm gr}\sec\theta_{\rm ref}$, and $\gamma = X_{\rm gr}/\lambda$, where $\mu = X_{\rm max}/\lambda$. We can thus propose a new form for the attenuation function,

$$f_{\rm att}^{\rm em}(\theta; \gamma, \mu) = \left(\frac{\sec\theta}{\sec\theta_{\rm ref}}\right)^{\mu} \exp[-\gamma(\sec\theta - \sec\theta_{\rm ref})] = \exp[-\gamma\, x + \mu\, y], \tag{8}$$

with $x = [\![\sec\theta]\!]$ and $y = [\![\ln\sec\theta]\!]$ using the notation $[\![z]\!] = z - z_{\rm ref}$. Since $\gamma = X_{\rm gr}/\lambda$ and $\mu = X_{\rm max}/\lambda$ it is advantageous to use $\mu = \gamma - \delta$ as the fitting parameters in order to reduce the regression correlations. This form thus changes into

$$f_{\rm att}^{\rm em}(\theta; \gamma, \delta) = \exp[-\gamma\, x + (\gamma - \delta)\, y]. \tag{9}$$

PoS(ICRC2025)425

Muonic component. From the point of production muons are practically unimpaired by the atmosphere and decay with a typical decay length λ_μ. From the muon profile in Fig. 1 of Ref. [7] where the total number of muons in a shower is given, we can convert the X dependence into distance, obtaining $\lambda_\mu \approx 36\,\mathrm{km}$. We therefore have $\lambda_\mu \gg h_s$ and we may alternatively treat muons as non-attenuating component, i.e. $f_{\mathrm{att}}^{\mu} = 1$.

Total signal. To correctly model the attenuation behavior for the superposition of the two components we have to take into account the differences in the response of the detector to these components. However, for simplification, we describe the total attenuation as a superposition of the attenuation of the two components

$$f_{\mathrm{att}}(\theta) = q_{\mathrm{em}}\, f_{\mathrm{att}}^{\mathrm{em}}(\theta) + (1 - q_{\mathrm{em}})\, f_{\mathrm{att}}^{\mu}(\theta) \tag{10}$$

with the effective fraction of the electromagnetic component q_{em}.

4. Direct Energy Calibration

We can rewrite Eq. (5) as $E = A'\, S_{\mathrm{gr}}/f_{\mathrm{att}}$, where A' takes care of the proportionality, calibration, and units. We also assume that the signal sizes are still dominated by the electromagnetic component, so that the proportionality in Eq. (4) is not broken, which is also why the analysis is limited to zenith angles below 60°. The reference angle θ_{ref} is usually chosen close to the median of the distribution of zenith angles, resulting in a reference angle of around 38°.

As the uncertainties of the shower size are usually log-normally distributed we actually fit the relation between the logarithm of the shower size and primary particle energy. For the signal-to-energy relation we use the model

$$\lg E = A + B\, \lg(S_{\mathrm{gr}}/f_{\mathrm{att}}), \tag{11}$$

in which we allow also for an exponent $B \neq 1$. We take the form of the likelihood $\mathcal{L}$ for the fit from Ref. [8],

$$\mathcal{L} = \prod_i^{E_i > E_{\mathrm{cut}}} \sum_k^{\text{all events}} \frac{1}{\sigma_{\lg E,k}\, \sigma_{\lg S,k}\, \epsilon_k} \exp\left(-\frac{\lg^2(E_i/E(S_k))}{2\sigma_{\lg E,k}^2} - \frac{\lg^2(S_i/S_k)}{2\sigma_{\lg S,k}^2} \right). \tag{12}$$

As we are going to use the "true" Monte Carlo energy for the calibration we have to take the limit $\sigma_{\lg E,k} \to 0$. In this limit the first Gaussian term becomes a δ-function with which we can integrate out the inner sum and derive essentially a least-squares fit for the inverse of Eq. (11),

$$\mathcal{L} = \prod_i^{E_i > E_{\mathrm{cut}}} \frac{1}{\sigma_{\lg S,i}\, \epsilon_i} \exp\left(-\frac{\lg^2(S_i/S(E_i))}{2\sigma_{\lg S,i}^2} \right). \tag{13}$$

If we restrict the energy cut E_{cut} to energies where the respective detector is fully efficient ($\epsilon_i = 1$) the only reason this is not simply a least-squares fit are the intrinsic shower fluctuations σ_{in} which appear as an *additional fit parameter*. This intrinsic fluctuation accounts for the natural shower-to-shower variability that limits the precision of energy reconstruction beyond measurement uncertainties,

PoS(ICRC2025)425

$\sigma^2_{\lg S,i} = \sigma^2_{\lg S_{\mathrm{gr}},i} + \sigma^2_{\mathrm{in}}$, where $\sigma_{\lg S_{\mathrm{gr}},i}$ is derived from the fit to the ring of simulated stations. We observe that ignoring the intrinsic fluctuations, i.e. fixing $\sigma_{\mathrm{in}} = 0$, will lead to biased results.

The minimization is done without bounds on the fit parameters. To ensure the conditions $\sigma_{\mathrm{in}} > 0$ and $q_{\mathrm{em}} \in [0, 1]$ they are wrapped in a softplus and expit[1] functions, respectively.

5. Comparison of fits

Now we want to compare how well the following models describe the attenuation in simulations. We start with a cubic function,

$$f_{abc} = 1 + x(a + x(b + xc)), \quad \text{where} \quad x = [\![\sin^2\theta]\!], \tag{14}$$

and compare the results with our simplified two component model,

$$f_{\delta,\gamma,q_{\mathrm{em}}} = q_{\mathrm{em}}\, \exp(-\gamma\, x + (\gamma - \delta)\, y) + (1 - q_{\mathrm{em}})\, 1 \quad \text{where} \quad x = [\![\sec\theta]\!],\ y = [\![\ln\sec\theta]\!], \tag{15}$$

motivated by which we will suggest an exponential cubic function,

$$f^{\exp}_{abc} = \exp(x(a + x(b + xc)), \quad \text{where} \quad x = [\![\sin^2\theta]\!]. \tag{16}$$

To compare the different forms for the attenuation function, we draw 100 samples with 5000 events above the respective energy threshold and perform the fits. In Fig. 2 *(left)* the difference between the respective attenuation form and the simple cubic function is shown. For each of the samples the difference in the log-likelihoods is also computed and shown in Fig. 2 *(right)*. We observe that the simplified two component model and the exponential cubic function describe the data consistently better than the cubic function. It seems that the exponential cubic function however also outperforms our two component model. This can be caused by to many simplifications in our model, or correlations in the fit parameters. Please note that these conclusions may depend on the chosen hadronic interaction model as well.

6. Comparison of different primary particles

In Fig. 3 we compare the different attenuation behavior of showers induced by protons, iron or photons for the simplified two component model. As expected we observe that for lighter particles same energies are mapped to lower signals. Moreover, as photons and high energy vertical protons will develop their maximum below the detector plane, the attenuation functions increase with zenith angles at low angles before they reach the region of exponential attenuation. The teal curves correspond to a mixed composition linearly going from pure proton at $10^{18.5}$ eV to pure iron at $10^{20.2}$ eV. As expected, we observe the fit for this mix to lay between the curves for pure proton and iron.

An advantage of the simplified two component model is the relation of the fit parameters to physical quantities, see Fig. 4. Looking at the distributions of q_{em}, we observe the expected ordering from heavy to light and an almost pure electromagnetic composition for the photon

[1] $\mathrm{softplus}(x) = \ln(1 + e^x)$, $\mathrm{expit}(x) = (1 + e^{-x})^{-1}$

PoS(ICRC2025)425

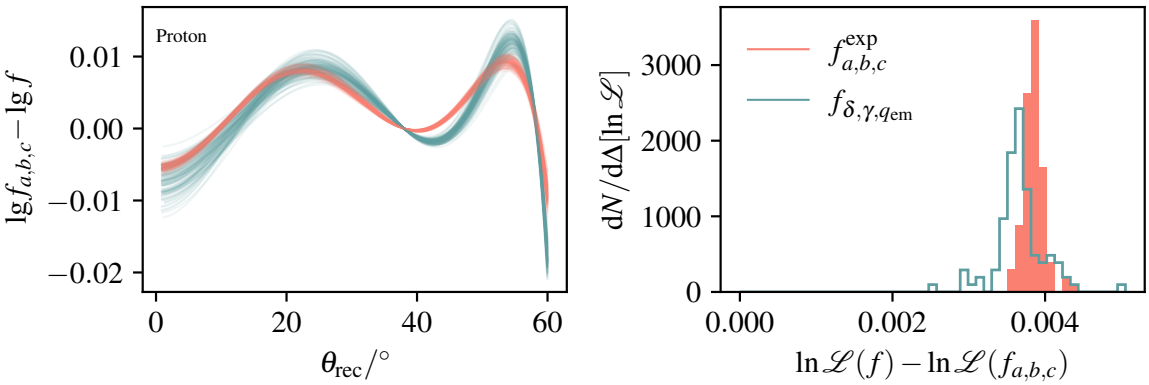

Figure 2: *Left:* Difference in the fitted attenuation functions f^{exp}_{abc} and $f_{\delta,\gamma,q_{\mathrm{em}}}$ compared to the cubic function f_{abc} for each of the 100 samples as a function of the zenith angle. *Right:* Difference in the log-likelihood for the attenuation functions compared to the cubic function. Higher values suggest the respective function to result in a higher likelihood and therefore suggest a better description of the data. A distribution around zero would suggest similar performance.

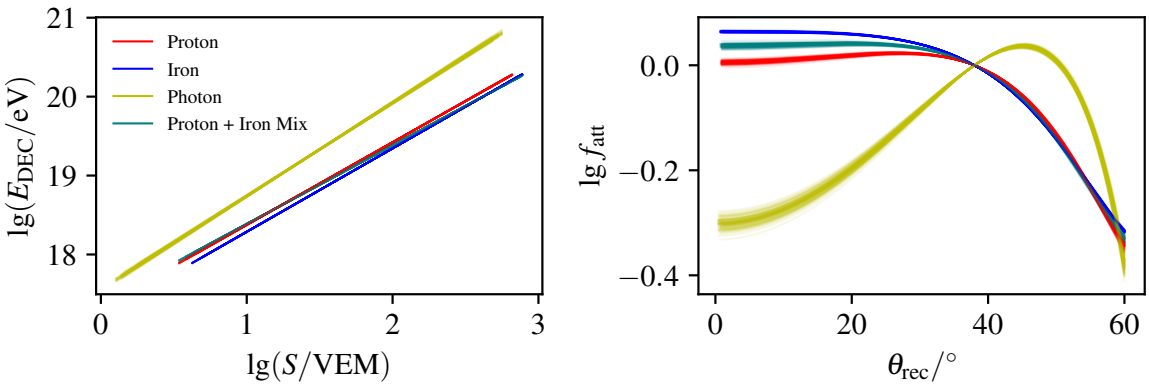

Figure 3: Energy calibration (left) and attenuation function (right) for 100 fits of the simplified two component model to showers induced by the three different primary particles and the mix of proton and iron.

induced shower. As expected the intrinsic shower-to-shower fluctuation are orders of magnitudes higher for photon induced showers and smallest for iron induced showers due to superposition in the shower development.

Looking at the distribution of the parameter B we see a small deviation from $B = 1$ for proton and iron and a distribution peaking around a larger value for photons. As we expect the deviation of B from 1 to be caused by the elongation rate, i.e. X_{max} increasing with energy, this corresponds to photon-induced showers having a larger elongation rate than showers induced by hadrons. For the mixed composition evolving from pure proton to pure iron the elongation rate is effectively reduced, as iron-induced showers have on average lower X_{max} values. Therefore also the fitted value for B moves closer to 1.

PoS(ICRC2025)425

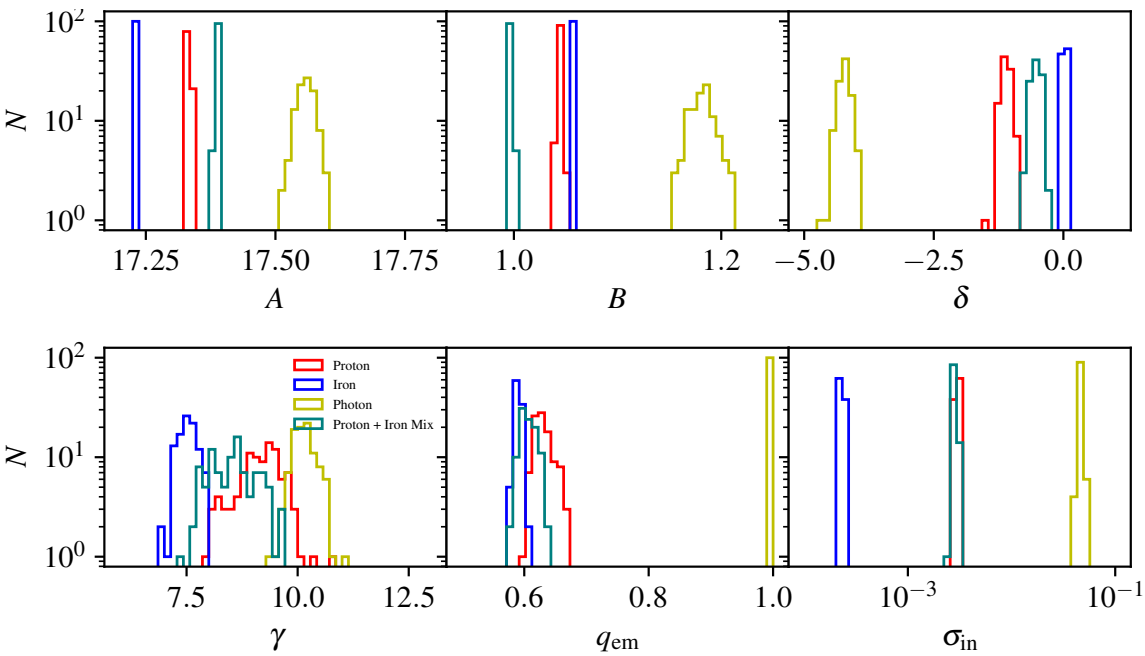

Figure 4: Distribution of the fitted parameters of the simplified two component model for 100 fits to showers induced by different primary particles.

7. Conclusions

We introduced a physically motivated model for shower attenuation, describing the muonic and electromagnetic components with a Gaisser-Hillas function. While this form captures key features of the data, its parameters provide insight into shower dynamics but often act as effective, rather than strictly physical, quantities. Introducing σ_{in}, a fit parameter representing the intrinsic shower-to-shower fluctuations, significantly reduces fit bias. Nonphysical results when extrapolating outside of the specific fit range can be avoided by using either the derived form or an exponential ansatz.

Acknowledgments The authors are very grateful to the Pierre Auger Collaboration for providing the tools necessary to perform the simulations for this contribution and for all the fruitful discussions.

References

[1] D. Newton, J. Knapp and A. Watson, *Astropart. Phys.* **26** (2007) 414.
[2] D. Heck, J. Knapp, J. Capdevielle et al., *Report FZKA* **6019** (1998) .
[3] T. Pierog, I. Karpenko, J.M. Katzy et al., *Phys. Rev. C* **92** (2015) 034906 [1306.0121].
[4] S. Argiro, S.L.C. Barroso, J. Gonzalez et al., *Nucl. Instrum. Meth. A* **580** (2007) 1485 [0707.1652].
[5] D. Vasylyev, *Earth Planets Space* **73** (2021) 112.
[6] T.K. Gaisser and A.M. Hillas, *Proc. Int. Cosmic Ray Conf.* **8** (1977) 353.
[7] Pierre Auger Collaboration, *Astropart. Phys.* **32** (2009) 89.
[8] H. Dembinski, B. Kégl, I. Mariş et al., *Astropart. Phys.* **73** (2016) 44.

PoS(ICRC2025)425